

RÉPUBLIQUE DU BÉNIN

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA
RECHERCHE

SCIENTIFIQUE

UNIVERSITÉ D'ABOMEY CALAVI

INSTITUT NATIONAL DE L'EAU

**DÉPARTEMENT D'HYDROLOGIE ET GESTION DES RESSOURCES
EN EAU**

Mémoire de fin de formation présenté par

HOUNKPE Valery Marc-Aurelle Dossou

Pour l'obtention du diplôme de

Master Professionnel en Hydrologie

Option : HYDRO - INFORMATIQUE

**Impact du changement climatique et de l'occupation des terres sur la
probabilité d'inondation en utilisant les méthodes Random-Forest et
Support Vecteur Machine**

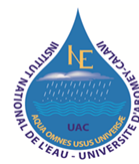
Sous la supervision de :

Prof, Adéchina Éric ALAMOU, Enseignant-Chercheur à l'ENSTP/UNSTIM
d'Abomey, Professeur Titulaire des Universités du CAMES

Sous l'encadrement de :

Dr. Ézéchiél OBADA, Enseignant-Chercheur à l'ENSTP/UNSTIM d'Abomey
Maitre-Assistant des Universités du CAMES

Année académique 2022 – 2023



RÉPUBLIQUE DU BÉNIN

MINISTÈRE DE L'ENSEIGNEMENT SUPÉRIEUR ET DE LA
RECHERCHE

SCIENTIFIQUE

UNIVERSITÉ D'ABOMEY CALAVI

INSTITUT NATIONAL DE L'EAU

DÉPARTEMENT D'HYDROLOGIE ET GESTION DES RESSOURCES
EN EAU

Mémoire de fin de formation présenté par

HOUNKPE Valery Marc-Aurelle Dossou

Pour l'obtention du diplôme de

Master Professionnel en Hydrologie

Option : HYDROINFORMATIQUE

**Impact du changement climatique et de l'occupation des terres sur la
probabilité d'inondation en utilisant les méthodes Random-Forest et
Support Vecteur Machine**

Sous la supervision de :

Prof, Adéchina Éric ALAMOU, Enseignant-Chercheur à l'ENSTP/UNSTIM
d'Abomey, Professeur Titulaire des Universités du CAMES

Sous l'encadrement de :

Dr. Ézéchiél OBADA, Enseignant-Chercheur à l'ENSTP/UNSTIM d'Abomey
Maitre-Assistant des Universités du CAMES

Année académique 2022 – 2023

DEDICACE

Je dédie ce modeste travail en signe de respect et de reconnaissance :

- ⇒ A Dieu, le Tout-Puissant pour sa main sur ma vie et mes études.
- ⇒ A mon père Jacques S. HOUNKPE et ma mère Rachel GANDONOU qui m'ont toujours encouragé durant ces années d'étude.

REMERCIEMENTS

La réalisation de ce travail a été possible grâce à la contribution active et aux efforts louables de plusieurs personnes et institutions. Nous tenons ici à leur témoigner notre profonde gratitude. Nos remerciements s'adressent de façon particulière :

- A notre superviseur Prof. Adéchina Éric ALAMOU et notre encadreur Dr. Ézéchiél OBADA qui nous ont beaucoup aidé par leurs orientations et précieux conseils pour l'élaboration de ce mémoire ;
- A Monsieur Jérôme ZOHOU, monsieur Oscar HOUESSOU et Dr. Halissou YAROU qui nous ont apporté leurs contributions scientifiques tout au long de nos travaux de recherches ;
- Au Directeur et le chef département du Laboratoire d'Hydrologie Appliquée (LHA) pour nous avoir accueillis et conseillés ;
- A tous les enseignants de L'Institut National de L'Eau et au personnel de l'administration ;
- A tous mes camarades de promotion en particulier Freddy HOUNNONDAHO et Gérôme AHOUANDJINO ;
- A mes frères et à tous mes amis. Que ce travail soit l'accomplissement de vos vœux tant allègues, et le fruit de votre soutien infaillible.
- Aux membres du jury qui malgré leurs nombreuses occupations ont accepté d'apporter leurs contributions pour l'amélioration du présent travail.

SOMMAIRES

DEDICACE	3
REMERCIEMENTS.....	4
LISTE DES ABBREVIATIONS.....	6
LISTE DES TABLEAUX.....	6
LISTE DES FIGURES	6
RESUME	8
ABSTRACT.....	9
Introduction Générale.....	10
Chapitre 1 : Revue de littérature.....	13
Introduction	13
1.1. Clarification des concepts	13
1.2. Points des connaissances.....	21
Conclusion	25
Chapitre 2 : Zone d'étude et données	27
Introduction	27
Conclusion	32
Chapitre 3 : Méthodologie.....	33
Introduction	33
Conclusion	43
Chapitre 4 : Résultats et discussion	44
Introduction	44
Discussion	60
Conclusion	61
Conclusion General.....	62
REFERENCES BIBLIOGRAPHIQUES.....	63
ANNEXES	70

LISTE DES ABBREVIATIONS

ANN:	Artificial Neural Network
AUC-ROC:	Area Under the ROC Curve
DEM :	Digital Élévation Model
ML :	Machine Learning
MLP :	Multi-Layer Perceptron
MCGS :	Modèle de Circulation Générale
MOLA:	Mars Orbiter Laser Altimeter
LARS-WG:	Long Ashton Research Station Weather Generator
LCM:	Land Cover Modeler
RCP:	Representative Concentration Pathway
RF:	Random Forest
SVM:	Support Vector Machine

LISTE DES TABLEAUX

Tableau 1 : Description des méthodes de ML dans la prédiction des inondations	22
Tableau 2 : le but de l'utilisation des facteurs	44
Tableau 3: Superficies et taux d'évolution des classes d'occupations entre 1990, 2000 et 2020	51
Tableau 4: Évaluation de la performance des deux modèles	59

LISTE DES FIGURES

Figure 1: Processus général de l'algorithme d'apprentissage supervisé.....	18
Figure 2 : Les forêts des arbres de décision	20
Figure 3 : Carte de la zone d'étude (bassin de la basse vallée de l'Ouémé)	28
Figure 4 : Station de la zone d'etude.....	30

Figure 5 : Structure du modèle LCM.....	35
Figure 6 : La carte des altitudes.....	47
Figure 7 : Carte des facteurs qui influent sur l'inondation.....	48
Figure 8 : Cartes d'occupation des terres/ sols (a. 1990 ; b. 2000 ; c. 2020).....	50
Figure 9 : Intensité de changement par catégorie d'occupation des terres entre 1990 et 2000	52
Figure 10 : Intensité de changement par catégorie d'occupation des terres entre 2000 et 2020	53
Figure 11 : Cartes d'occupation des terres/sols prédites pour l'année 2040	54
Figure 12 : Régimes pluviométriques observés et générés des stations étudiées (RCP4.5)	55
Figure 13: Régimes pluviométriques observés et générés des stations étudiées (RCP8.5)	56
Figure 14 : Cartes de probabilités des inondations	58

RESUME

Cette étude a été menée dans un sous-bassin de la basse vallée de l'Ouémé, et vise à élaborer un modèle intégré d'analyse des risques d'inondation. Cette approche prend en compte le changement climatique et l'évolution de l'occupation des terres en utilisant des techniques d'apprentissage automatique avancées. Dans cette étude les prévisions des changements des terres futures reposent sur la méthode des modeleurs des changements des terres, en parallèle, les projections du changement climatique pour la plage temporelle, basées sur les données climatiques de 1980 à 2016, ont été générées à l'aide du logiciel Lars-WG et en considérant deux scénarios, RCP4.5 et RCP8.5. Deux modèles à savoir le random-forest et les supports vector machine ont été employés pour anticiper les zones potentiellement exposées aux inondations. Les prévisions de risques, obtenues grâce à l'utilisation des modèles tels que le Random-Forest et de Support Vector Machine, renforcées par des modèles de prévision climatique, démontrent une précision remarquable. L'utilisation des chaînes de Markov pour analyser les changements d'occupation des terres jusqu'en 2040 offre des informations cruciales pour la planification urbaine. Les modèles sous les scénarios climatiques RCP4.5 et RCP8.5 permettent de cartographier les tendances des prélèvements, enrichissant la capacité à anticiper les épisodes pluvieux extrêmes. L'évaluation rigoureuse des modèles, avec des critères de performances tels que l'aire sous la courbe ROC (0,88) et les F1-scores (0,80), confirme leur robustesse, mettant en avant l'efficacité du modèle Random-Forest et SVM. Les résultats de cette recherche permettent une approche holistique pour la gestion des risques d'inondation, fournissant des outils précieux pour guider les décisions visant à atténuer l'impact des inondations à l'avenir dans le sous-bassin de la vallée de l'Ouémé.

Mots – clés : Apprentissage automatique ; changement climatique ; random-forest ; support vector machine ; inondation.

ABSTRACT

This study was conducted in a sub-basin of the lower Ouémé Valley and aims to develop an integrated flood risk analysis model. This approach takes into account climate change and land use evolution using advanced machine learning techniques. In this study, predictions of future land changes are based on the method of land change modelers. Simultaneously, climate change projections for the time range, based on climate data from 1980 to 2016, were generated using the Lars-WG software considering two scenarios, RCP4.5 and RCP8.5. Two models, namely random forest and support vector machine, were employed to anticipate areas potentially exposed to floods. Risk predictions, obtained through the use of models such as Random Forest and Support Vector Machine, reinforced by climate prediction models, demonstrate remarkable accuracy. The use of Markov chains to analyze land use changes until 2040 provides crucial information for urban planning. Models under climate scenarios RCP4.5 and RCP8.5 enable mapping of withdrawal trends, enhancing the ability to anticipate extreme rainfall events. Rigorous model evaluation, with performance criteria such as the area under the ROC curve (0.88) and F1-scores (0.80), confirms their robustness, highlighting the effectiveness of the Random Forest and SVM models. The results of this research enable a holistic approach to flood risk management, providing valuable tools to guide decisions aimed at mitigating the impact of floods in the future in the Ouémé Valley sub-basin.

Keywords: Machine learning; climate change; random forest; support vector machine; flood.

Introduction Générale

A. Contexte et Justification

Les catastrophes naturelles figurent parmi les principales menaces pesant sur les pays au cours des dernières décennies (Guha-Sapir et al 2017). Les inondations sont l'un des risques naturels ou catastrophes les plus importants avec un impact élevé et des préoccupations au sein des communautés mondiales. Les inondations sont généralement causées par l'incapacité des cours d'eau naturels et des systèmes de drainage à transporter l'eau en excès après de fortes pluies. Dilley et al. (2005) estime que plus d'un tiers de la superficie terrestre mondiale est sujette aux inondations, ce qui affecte par conséquent environ 82% de la population mondiale. Au cours des dernières décennies, il y a eu une augmentation significative des dommages économiques et des pertes en vies humaines dans de nombreux pays en raison d'inondations. Les données statistiques de 1990 à 2013 indiquent que les inondations ont causé environ 7 millions de morts et plus de 600 milliards de dollars de pertes dans le monde (Alison Freebairn et al 2020). Selon les données statistiques du Bureau des Nations Unies pour la réduction des risques de catastrophe (UNISDR), environ 150 061 inondations se sont produites dans le monde et environ 157 000 personnes sont mortes par inondations de 1995 à 2015, responsables de 11,1 % des victimes des catastrophes mondiales. En 2010, l'ensemble des dommages occasionnés par les inondations sur l'économie béninoise s'élève à près de 78,3 milliards de FCFA (près de 160 millions USD) (Pda, 2015). Il faut noter que le Bénin est, et reste un pays fortement affecté par des inondations cycliques qui touchent généralement vingt-sept (27) des soixante-dix-sept (77) communes que compte le pays et qui sont causées par divers facteurs naturels, anthropiques et organisationnels. Pour l'année 2021, les crues des principaux fleuves du pays (le Niger, l'Ouémé, le Mono) et de leurs affluents, qui ont débuté début septembre, ont entraîné des inondations systématiques dans trente-huit (38) communes (du nord au sud), occasionnant plusieurs dégâts (« IFRC GO - Emergency », s. d.)

Malgré tous les efforts fournis par les différentes autorités et états pour réduire les impacts sur tous les secteurs de ce phénomène, fort est de constater que l'inondation continue de faire des dégâts dans le monde en particulier au Bénin. En effet les régions dont le climat évolue vers des événements pluvieux plus extrêmes et plus fréquents sont plus susceptibles de connaître des conditions d'inondation, ce qui fait accroître l'ampleur de l'inondation. Certains phénomènes comme les changements climatiques, la croissance démographique et les aménagements impactent positivement la vulnérabilité des populations aux

inondations. Par conséquent les études sur la modélisation spatiale des inondations dans un bassin versant ou sur le plan national s'avèrent nécessaires pour d'une part observer les zones à fort risques d'inondations et d'autre prendre des mesures d'adaptation ou d'atténuation contre ce phénomène.

Face à ces défis complexes, l'apprentissage automatique émerge comme un outil prometteur pour analyser les interactions complexes entre le changement climatique, l'occupation du sol et la probabilité d'inondation. Selon Lam et al, (2019) « le machine Learning permet d'extraire des modèles à partir de grandes quantités de données environnementales, climatiques et géospatiales, fournissant ainsi des informations précieuses pour évaluer et prédire les risques d'inondation ». Dans cette optique il est important de construire un modèle de prédiction performant basé sur les modèles d'intelligence artificielle pour déterminer la probabilité des événements d'inondation et de faire la cartographie des régions inondables. Une étude réalisée par Niazkar et al concluent que les méthodes d'IA peuvent simuler et prévoir avec succès des séries chronologiques dans différentes aquifères, de même une étude réalisée sur les modèles hydrologique et l'apprentissage automatique a évalué quatre modèles hydrologiques conceptuels (GR2M, XAJ, SAC, Alpine) et quatre modèles d'apprentissage automatique (RF, GBDT, DNN, CNN) pour la simulation du débit des cours d'eau et la reconstruction des données, les auteurs ont constaté que les modèles d'apprentissage automatique étaient généralement plus performants, mais que les modèles hydrologiques atteignaient un pourcentage plus élevé de stations avec une bonne précision de simulation (« A review of the artificial intelligence methods in groundwater level modeling » 2019). Plusieurs études ont montré que le machine Learning peut être utilisé en hydrologie et dans plusieurs autres domaines. Cela peut permettre de mieux prévoir les inondations, ce qui peut réduire les dommages matériels et les pertes en vie humaine. Aussi, la cartographie des zones à risques peut aider les décideurs dans les prises de décision.

C'est dans cette optique que dans cette étude, nous proposons d'utiliser le machine Learning pour examiner les effets du changement climatique et de l'occupation du sol sur la probabilité d'inondation dans la basse vallée de l'Ouémé. En nous appuyant sur les travaux de recherche précédents et en exploitant les vastes ensembles de données disponibles, nous chercherons à développer des modèles prédictifs capables d'estimer de manière précise la probabilité d'inondation dans la région. Ces modèles contribueront à renforcer la planification et la gestion des risques, en offrant aux décideurs des informations

essentielles pour élaborer des politiques d'adaptation et de prévention plus efficaces.

B. Objectifs de la recherche

Le but principal de notre étude est de développer un modèle intégré d'analyse des risques d'inondation prenant en compte le changement climatique, l'évolution de l'occupation des terres en utilisant des techniques d'apprentissage automatique avancées.

Objectifs spécifiques

De façon spécifique, nous allons :

- Déterminer les paramètres qui influent sur l'inondation dans la basse vallée de l'Ouémé ;
- Analyser l'évolution des précipitations et de l'occupation de sols dans la basse vallée de l'Ouémé ;
- Développer un modèle basé sur l'apprentissage automatique qui prédire la probabilité future des inondations.

C. Organisation du mémoire

Ce document est structuré en 04 chapitres. Le premier chapitre porte sur la revue de littérature. On y présente la définition des différents concepts constituant l'étude et un accent particulier est mis sur les méthodes d'apprentissages d'automatiques. Le chapitre 2 présente notre zone d'étude et les données utilisées pour l'atteinte de nos objectifs. Dans le chapitre 3, nous détaillons la méthodologie pour les différents objectifs spécifiques. Le chapitre 4 nous présente les différents résultats obtenus et discussions proposées.

Chapitre 1 : Revue de littérature

Introduction

L'intégration de la télédétection, des SIG, de la méthode des modeleurs des changements des terres, et de l'apprentissage automatique offre une approche novatrice pour anticiper et gérer les impacts du changement climatique, notamment les risques d'inondation. Dans ce chapitre il est question de définir clairement les termes et concepts utilisés le long de l'étude.

1.1. Clarification des concepts

Nous avons identifié quelques termes à prendre en compte pour la recherche. La combinaison des variantes de ces quatre termes guide donc notre recherche. Dans cette partie il est question de faire une clarification de ces termes et d'ajouter aussi une revue de littérature.

1.1.1. Changement climatique

Le changement climatique : C'est la variation de l'état du climat que l'on peut caractériser (par exemple au moyen des tests statistiques), par des modifications de la moyenne et ou de la variabilité de ses propriétés et qui persiste pendant une longue période, généralement pendant des décennies ou plus. Les principales causes sont les émissions de gaz à effet de serre dues aux activités humaines, telles que la combustion de combustibles fossiles, la déforestation et l'agriculture intensive (Mohamed EL et al 2009).

Les conséquences du changement climatique incluent l'élévation du niveau de mer, des phénomènes météorologiques extrêmes plus fréquents, la fonte des glaces, etc. Tous ces et bien d'autres phénomènes favorisent l'inondation.

1.1.2. SIG et Télédétection

La télédétection

La télédétection, comprend l'ensemble des procédés et techniques qui permettent d'acquérir à distance des informations sur les objets terrestres, en utilisant les propriétés des ondes électromagnétiques émises ou réfléchies par des objets. Elle se rapporte à l'ensemble des techniques liées à l'analyse et à l'exploitation des données obtenues par des satellites d'observation des ressources terrestres et environnementales (tels que Meteosat, NOAA-AVHRR, Landsat Thematic Mapper [TM], SPOT et ERS-SAR) et par photographie aérienne

L'objectif principal de la télédétection est de cartographier les ressources terrestres et d'en effectuer le suivi. Par rapport aux techniques de levée classiques, la télédétection par satellite est précise et rentable, et permet d'obtenir des données en temps utile.

Le Système d'Informations Géographiques (SIG)

On entend par SIG un système informatisé de collecte, de stockage, de traitement, d'analyse et d'extraction de données, composé de matériel et de logiciels spécifiquement conçus pour traiter des données géoréférencées et les informations relatives aux attributs y afférents. En général, les données géoréférencées se trouvent sous forme de cartes reflétant diverses caractéristiques : topographie, disponibilité de l'eau, types de sol, forêts et herbages, climat, géologie, population, propriété foncière, limites administratives, infrastructures (routes, voies ferrées, réseaux électriques ou systèmes de communication), etc. Les principales fonctions des SIG sont la « superposition », procédé consistant à regrouper plusieurs cartes en une, ainsi que la modélisation et la sélection de sites. Cette définition succincte du SIG, introduit la notion de cartographie (Martin P et AL 2004).

La cartographie

Elle se définit comme étant une collecte des données stockées et analysée (Ibrahima et Al 2023). Les différents éléments intervenant dans un document cartographique sont :

- L'objet spatial, délimité par une aire géographique (exemple de la zone d'été) ;
- La ligne, qui est un objet spatial composé de plusieurs points (exemple : les courbes de niveau d'une carte topographique) ;
- Le nœud, qui est un type de point matérialisant la jonction de plusieurs lignes (Croisement des routes) ;
- Le polygone, qui est une surface fermée (exemple du contour d'un bas-fond).

Images satellites à haute résolution

Utilisées à des fins opérationnelles, les images satellites à haute résolution, telles que celles fournies notamment par Landsat TM, SPOT, ERS-SAR et RADARSAT, offrent de nouveaux débouchés dans les domaines de l'étude et du suivi des ressources naturelles. Par rapport aux informations obtenues par des méthodes classiques, ces données présentent plusieurs avantages (« Série Télédétection pour décideurs N° 21 » 2010) :

- Elles permettent d'obtenir une couverture synoptique d'une zone donnée, et donc de présenter simultanément de grandes superficies en détail.
- Elles peuvent être acquises pour la même zone à une fréquence élevée (deux à trois fois par mois), ce qui permet de choisir les données saisonnières les plus pertinentes.
- Les images satellites sont enregistrées sur plusieurs longueurs d'onde (domaine du visible et du non visible), ce qui permet d'obtenir des informations précises sur les caractéristiques du terrain.
- Il est possible de les obtenir pour toutes les régions du monde, sans restriction administratives.

L'utilisation du système d'information géographique (SIG) et des données de télédétection a été explorée dans plusieurs articles pour l'estimation et la gestion des inondations.

1.1.3. Modeleur des changements des terres

La méthode du Modeleur des Changements des Terres (LCM)

C'est une approche utilisée pour analyser et prédire les changements dans l'utilisation des terres sur une certaine période. Le LCM est une approche couramment utilisée dans le domaine de la télédétection, de la cartographie des terres et de la modélisation des changements environnementaux. La méthode du Modeleur des Changements des Terres (LCM) peut également être appliquée pour étudier le changement des terres lié aux inondations. Cette approche peut être utilisée pour analyser les changements dans l'utilisation des terres causés par les inondations, prédire les zones susceptibles d'être touchées par les inondations. Elle permet également d'évaluer les impacts environnementaux et socio-économiques associés.

C'est une approche utilisée en géographie et en cartographie pour étudier et analyser les modifications qui se produisent dans l'utilisation des terres au fil du temps. Voici un résumé des principales étapes de cette méthode :

- a. Collecte de données : La première étape consiste à rassembler des données historiques sur l'utilisation des terres, généralement sous la forme de cartes, d'images satellites ou d'images aériennes prises à différents moments.
- b. Classification des données : Les données sont classées en différentes catégories d'utilisation des terres, telles que l'agriculture, la forêt, l'urbanisation, etc.

- c. Création d'une chronologie : Les données classifiées sont organisées chronologiquement pour suivre l'évolution des changements au fil du temps.
- d. Analyse des tendances : En utilisant des techniques statistiques et des outils de traitement d'image, on identifie les tendances et les changements significatifs dans l'utilisation des terres.
- e. Modélisation des changements futurs : Sur la base des tendances identifiées, des modèles sont développés pour prédire les changements futurs dans l'utilisation des terres.
- f. Évaluation des impacts : Les impacts environnementaux, économiques et sociaux des changements prévus sont évalués afin de prendre des décisions éclairées en matière de planification urbaine, de conservation des ressources naturelles, etc.

Cette méthode implique souvent l'utilisation de logiciels spécialisés pour la collecte, l'analyse et la modélisation des données sur l'utilisation des terres. Voici quelques-uns des logiciels couramment utilisés pour cette méthode : ArcGIS, ENVI, QGIS, LCM, TerrSet/IDRISI, le langage R avec des packages géospatiaux.

1.1.4. Occupations des sols / utilisations des sols

L'occupation et l'utilisation du sol

L'occupation et l'utilisation du sol sont des éléments déterminants du fonctionnement de la biosphère et contribuent ainsi à la dynamique du système Terre. Les changements d'occupation et d'utilisation du sol sont à la fois une cause et une conséquence du changement global. Au centre d'un système complexe mettant en jeu à la fois des processus humains et des processus « naturels », leur suivi représente un enjeu considérable sur le plan environnemental (Lecerf et Al 2008).

1.1.4.1. L'occupation du sol

Elle désigne un type de couverture physique comme les forêts, les surfaces herbacées ou les surfaces artificialisées. Si nous considérons les terres émergées, l'occupation du sol peut être décomposée en deux parties : le sol et la végétation. Ce sont donc les deux éléments qui permettent de décrire la couverture physique des surfaces terrestres (Lecerf et Al 2008).

1.1.4.2. L'utilisation du sol

Il est plus complexe à déterminer, car elle dépend de l'usage socio-économique qui est fait des terres. Par exemple, une surface herbacée peut avoir

plusieurs fonctions et être utilisée comme terre agricole ou espace vert. Ainsi, la fonction socio- économique de la surface herbacée dépend de l'exploitation que l'Homme en fait et permet de désigner le type d'utilisation du sol. L'utilisation que l'Homme fait du sol peut, dans certaines situations, servir à déterminer l'occupation du sol qui en retour, en fonction du service rendu, peut constituer un facteur explicatif des changements d'utilisation du sol (Lecerf et Al 2008).

Les types d'occupations du sol :

- Territoires urbanisés ;
- Territoires agricoles ;
- Forêts et milieux semi-naturels ;
- Zones humides ;
- Espaces aquatiques.

Les changements d'occupation et d'utilisation sont de plus en plus reconnus comme un élément explicatif du changement (Lambin et al. 2001) (Myers et Patil 2006), Parallèlement, le climat est considéré comme un élément central dans l'explication des changements d'occupation et d'utilisation du sol, comme le soulignent les derniers rapports du GIEC (Groupe d'experts Intergouvernemental sur l'Évolution du Climat) (Lecerf et Al 2008).

1.1.5. Apprentissage Automatique

Selon Tom M. Mitchell (1997) **l'apprentissage automatique** est: « *A computer program is said to learn from experience E concerning some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E.* ».

« On dit qu'un programme informatique apprend de l'expérience E par rapport à une classe de tâches T et mesure de performance P, si sa performance aux tâches de T, mesurée par P, s'améliore avec l'expérience E. », selon Mitchell (1997). Les algorithmes d'apprentissage automatique peuvent être classés en six catégories (Mitchell et Al 1997): l'apprentissage supervisé, l'apprentissage non-supervisé, l'apprentissage semi-supervisé, la méthodologie d'apprentissage par renforcement, l'apprentissage transductif, l'inférence inductive. Les algorithmes d'apprentissage automatique fréquemment appliqués dans les recherches de télédétection appartiennent à deux catégories d'algorithmes : l'apprentissage

supervisé et l'apprentissage non-supervisé. L'efficacité de l'application de ces algorithmes a été largement prouvée dans les recherches de (Blaschke 2010)

1.1.5.1. L'algorithme d'apprentissage supervisé

C'est une méthode d'apprentissage qui déduit la relation sous-jacente entre les données observées (également appelées données d'entrée) et une variable cible (une variable dépendante ou étiquette) qui est soumise à la prédiction. La tâche d'apprentissage utilise les données de formation étiquetées (exemples de formation) pour synthétiser la fonction du modèle qui tente de généraliser la relation sous-jacente entre les vecteurs de caractéristiques (entrée) et les signaux de supervision (sortie) (Kettig et Landgrebe 1976). Les algorithmes d'apprentissage supervisé sont souvent utilisés pour faire des classifications des composants géographiques dans les données d'images de télédétection (par pixel ou par objet). Les méthodes supervisées consistent à prédire une variable Y en fonction de variables explicatives X (MacQueen 1967). Évidemment, les exemples d'applications des modèles prédictifs sont nombreux :

- Estimer la valeur d'une variable
- Déterminer la probabilité
- Évaluer le risque
- Prédire un volume de précipitation
- Prédire

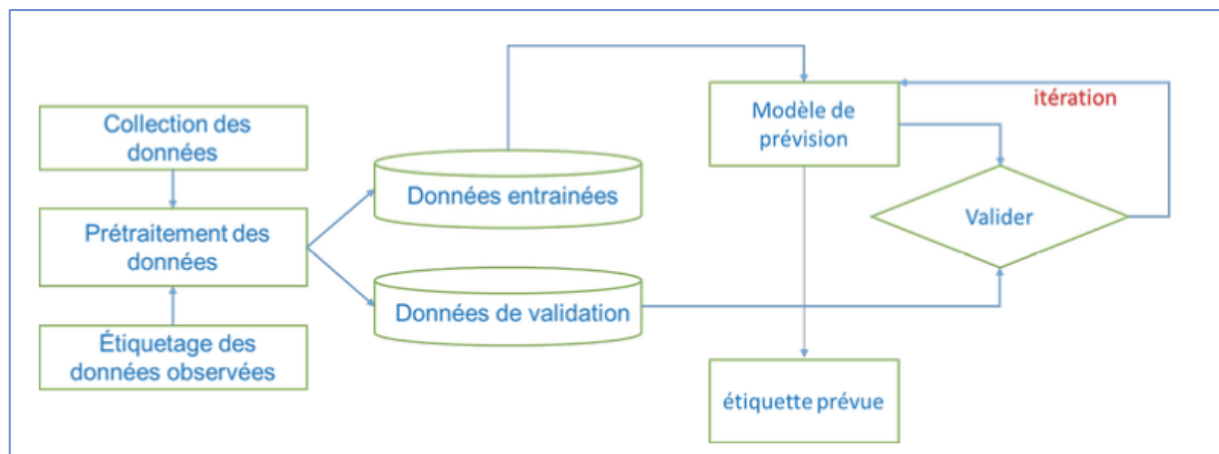


Figure 1: Processus général de l'algorithme d'apprentissage supervisé

1.1.5.1.1. Méthode de classification

La classification est une technique d'apprentissage automatique qui implique la formation d'un modèle pour attribuer une étiquette de classe à une entrée donnée. Il s'agit d'une tâche d'apprentissage supervisé, ce qui signifie que

le modèle est formé sur un ensemble de données étiquetées qui comprend des exemples de données d'entrée et les étiquettes de classe correspondantes. Le modèle vise à apprendre la relation entre les données d'entrée et les étiquettes de classe afin de prédire l'étiquette de classe pour de nouvelles entrées inédites (Awad et Khanna 2015). L'une des principales approches pour obtenir des informations de l'occupation du sol à partir des images satellitaires est la classification. De nombreux algorithmes de classification ont été développés depuis l'acquisition de la première image Landsat au début des années 1970. Parmi eux, on trouve les algorithmes d'apprentissage automatique peu profond (Machine Learning) comme la méthode de maximum de vraisemblance (Maximum Likelihood ML), les algorithmes de réseaux neuronaux (ex. *Multilayer Perceptron* MLP), des arbres de décision (ex. *Random forest*-RF) et l'algorithme de *Support Vector Machine* (SVM).

Support vector machine

Le SVM est considéré comme un des meilleurs algorithmes dans le domaine de l'analyse d'images ; en effet, c'est un outil d'apprentissage automatique très populaire. Les classificateurs SVM sont initialement linéaires à deux classes, reposant sur la maximisation de la marge de séparation entre deux classes. Grâce aux fonctions de noyaux, des transformations sont désormais possibles pour une classification multi-classe (Foody et Mathur 2004). C'est un algorithme d'apprentissage automatique supervisé qui peut être utilisé pour les problèmes de classification ou de régression. Toutefois, il est surtout utilisé dans les problèmes de classification (HUANG et Al 2002). Le principe des SVM consiste à ramener un problème de classification ou de discrimination à un hyperplan (*feature space*) dans lequel les données sont séparées en plusieurs classes dont la frontière est la plus éloignée possible des points de données (ou *marge maximale*). Il existe les SVM linéaires et non-linéaires.

Random Forest

Depuis les années 1980, les algorithmes d'arbre de décision ID3, CART, C45 ont prouvé leur efficacité en utilisant des données numériques et symboliques (« Induction d'arbres de décision | SpringerLink », s. d.). Ces algorithmes donnent des résultats de bonne précision pour les données combinées multi-sources et hétérogènes. Afin de performer ces algorithmes d'arbre de décision et d'éliminer certains de leurs inconvénients, Breiman a développé un algorithme appelé *Random forest*

L'arbre de décision est une structure hiérarchique générée par des fonctions statistiques et probabilistes. Le but du processus de construction d'arbres de

décision est d'utiliser des règles de division et d'arrêt de calcul qui peuvent être partitionnées en ensembles d'éléments homogènes. L'algorithme de référence est celui de (Breiman 2001) décrivant la méthode CART (*Classification et Regression Tree*). Néanmoins, la paternité de cette méthode est attribuée à (Morgan et Sonquist 1963)

Le principe des arbres de décision repose sur la prédiction d'une classe en sortie après un processus de séparation successive des données. L'objectif est de prédire des groupes les plus homogènes possibles par rapport à la variable à prédire. La mise en place de cette méthode est considérée comme étant facile et rapide, mais elle peut être sensible à l'apprentissage exagérée (Thanh N et al 2017).

La forêt des arbres de décision

La forêt des arbres de décision est un ensemble des classificateurs $H = \{H_1, H_2, \dots, H_k\}$ qui sont construits à partir des sous-échantillons aléatoires $E = \{E_1, E_2, \dots, E_k\}$

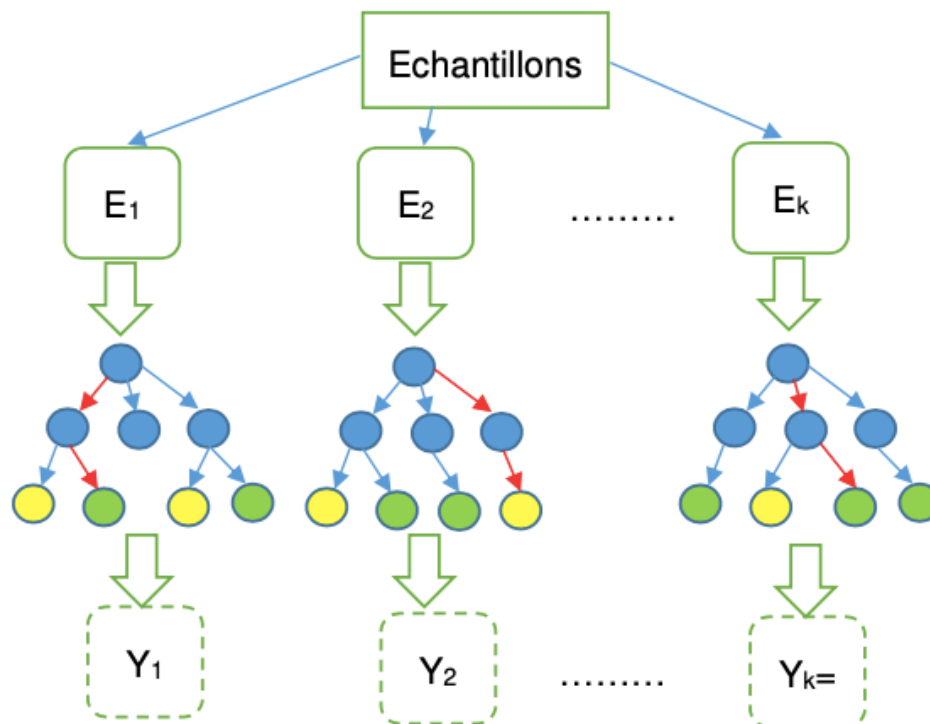


Figure 2 : Les forêts des arbres de décision

Le **forêt** des arbres de décision se fait généralement en deux parties : étape d'entraînement et l'étape d'application.

Voici les différentes étapes d'entraînement :

- Un sous - échantillon est créé avec le nombre m d'éléments aléatoire, le coefficient m est précisé.
- Pour chaque sous - échantillon, un arbre de décision est géré.
- Après k itérations, une forêt de k arbres de décisions est créée. Le coefficient m est précisé.

Voici les différentes étapes d'application :

- Un élément x_i est appliqué sur tous les arbres dans la forêt, et la prévision de chaque arbre est sauvegardée.
- La dernière prévision est calculée par la fonction de vote majorité des résultats provenant des arbres de décision dans la forêt.

1.1.5.2. Méthode de régression

La régression est un type d'apprentissage supervisé dont l'objectif est de prédire une variable dépendante c sur la base d'une ou plusieurs caractéristiques d'entrée (également appelées prédicteurs ou variables indépendantes). Les algorithmes de régression sont utilisés pour modéliser la relation entre les entrées et les sorties et faire des prédictions basées sur cette relation. La régression peut être utilisée pour des variables dépendantes continues ou catégorielles. En général, l'objectif de la régression est de construire un modèle qui peut prédire avec précision la sortie sur la base des caractéristiques d'entrée et de comprendre la relation sous-jacente entre les caractéristiques d'entrée et la sortie (Yariyan et al. 2020). La régression consiste alors à formuler un indicateur sur les valeurs de la variable expliquée dépendant uniquement des valeurs des variables explicatives. On distingue essentiellement deux cas selon la nature de la variable expliquée, représentée ici par une variable aléatoire Y . Les variables explicatives seront notées $X_1, \dots, X_n$. Ainsi, la régression inclut la régression linéaire (Y continu), la régression logistique (Y binaire), la régression multinomiale (Y catégoriel), la régression de Poisson (Jenifa 2022).

1.2. Points des connaissances

1.2.1. Techniques d'intelligence artificielle pour estimer la probabilité des inondations

Des algorithmes d'apprentissage automatique ont été utilisées pour estimer la probabilité d'inondations. Ces techniques ont contribué à la réduction des risques, aux recommandations politiques et à la minimisation des pertes en vies humaines ainsi que des dommages matériels liés aux inondations. Divers algorithmes

d'apprentissage automatique, notamment Multilayer Perceptron Classifier (MLP) et Support Vector Machine (SVM) ont été testés et se sont révélés efficaces dans la prédiction et la classification des inondations. De plus, des méthodologies d'intelligence artificielle, telles que l'apprentissage en profondeur et le traitement d'images, ont été appliquées à la prévision des inondations pluviales en temps réel, fournissant des prévisions opportunes des crues éclair. Des techniques informatiques douces, telles que les réseaux de neurones artificiels (ANN), les machines d'apprentissage extrêmes (ELM) et les arbres modèles M5, ont également été utilisées pour des prévisions d'inondation précises et fiables, même avec une disponibilité limitée des données (Chan et Chan 2021). Dans le domaine de l'atténuation des inondations, l'apprentissage automatique a été utilisé pour la prévision des inondations, la cartographie des risques d'inondation, l'intervention d'urgence en cas d'inondation et la cartographie des dommages causés par les inondations (Ahmed et al. 2022). Outre l'utilisation de techniques mécaniques, l'apprentissage automatique peut être utilisé pour augmenter la résistance et la préparation des communautés au catastrophe d'inondation, en appliquant et en comparant les performances ainsi que la précision de plusieurs techniques telles que : Random Forest (RF), Lazy, J48 arbres, réseaux de neurones artificiels (ANN), Bayes (NB) et régression logistique (LR), la méthode ANN (Réseau de neurones artificiels) utilisé pour développer le modèle inondations régionales (Chang, Lai, et Hwang 2018).

Utilisation de modèles d'apprentissage automatique pour prévoir spatialement les inondations est un domaine de recherche recommandé. Les avancées récentes ont démontré que le modèle d'apprentissage automatique pour l'analyse spatiale des inondations fera révolutionner le domaine de la prévision des crues.

Tableau 1 : Description des méthodes de ML dans la prédiction des inondations

Méthode de Machine Learning	Description et Utilisation dans la Prédiction des Inondations	Avantages

Régression Linéaire	Pour modéliser les relations linéaires utilisées entre les variables telles que les précipitations, le débit des rivières, etc., afin de prédire les niveaux d'eau et les inondations.	Facile à interpréter et à mettre en œuvre, appropriée pour les relations linéaires simples.
Réseaux de Neurones	Les réseaux de neurones peuvent capturer des relations complexes entre de multiples variables climatiques, topographiques et hydrologiques pour prédire les inondations.	Capacité à apprendre des motifs complexes dans les données, adaptabilité à différentes caractéristiques du bassin versant.
Réseaux de Neurones Récurrents (RNN)	Utiles pour modéliser les séquences temporelles, comme les niveaux d'eau en fonction des précipitations passées, pour améliorer la prédiction des inondations.	Prise en compte des dépendances temporelles, aptitude à capturer les variations à long terme.
LSTM (mémoire longue à court terme)	Une variante de RNN qui peut gérer les problèmes de disparition du gradient et peut mieux saisir les motifs temporels complexes dans les séries chronologiques d'inondations.	Adapté à la modélisation des séquences temporelles avec des dépendances à long terme.

Réseaux de Neurones Convolutifs (CNN)	Pour traiter des données géospatiales utilisées, tels que des images satellites, pour prédire les zones à risque d'inondation en analysant les caractéristiques visuelles.	Excellente pour la reconnaissance de motifs spatiaux, utile pour la cartographie des zones inondables.
Méthodes ensemblistes (Random Forest, XGBoost)	Utilisées pour combiner plusieurs modèles simples et capturer des relations non linéaires complexes entre les variables hydrologiques, améliorant ainsi les prédictions d'inondation.	Réduisent le risque de surajustement, offrent des mesures d'importance des caractéristiques, et sont adaptées à la parallélisation, permettant ainsi de traiter de vastes ensembles de données géographiques.

Des études ont montré que la qualité de la prévision des inondations peut être améliorée de plusieurs manières à savoir :

- Faire de la nouveauté dans l'hybridation, soit par l'intégration de deux ou plusieurs méthodes apprentissage automatique ou intégration de méthodes l'apprentissage automatique de plusieurs manières conventionnelles et/ou informatiques.
- Utilisation de techniques de décomposition de données dans l'objectif d'améliorer la qualité de l'ensemble de données, c'est-à-dire grandement contribué à l'amélioration précision des prédictions.
- L'utilisation de la méthode d'ensemble, peut considérablement améliorer modéliser les capacités et réduire les incertitudes dans la réalisation des prévisions.
- Faire un choix d'algorithmes pour améliorer la qualité des algorithmes d'apprentissage automatique, par exemple, en proposant des ajustements mieux sur les ANN pour atteindre une architecture neuronale optimale.

1.2.2. Prédiction des inondations avec les données de SIG et d'occupation du sol

Grace à ce thème nous avons effectués nos recherches sur les divers outils de recherche pour pouvoir mieux appréhender les différents facteurs qui influent sur la susceptibilité à l'inondation, bien qu'il n'y ait pas de méthode standard pour sélectionner les facteurs qui pourraient le mieux prédire les inondations futures. Hakim et al., sur la base de la revue effectuée dans leur ouvrage on trouve et exposé les 10 principaux facteurs ou paramètres de conditionnement largement utilisés pour identifier les zones à fortes risques d'inondation. Il s'agit de la pente (98%), l'altitude (92%), la couverture terrestre/ utilisation des terres (79%), la distance à la rivière (76%), la précipitation (73%), la géologie (61%), le TWI (61%), la densité des rivières (43%), l'aspect (41%) et le SPI (35%). Ces mêmes facteurs/ paramètres ont été considérés très souvent dans certains ouvrages comme (Fujiki et Finance 2022); (Roy et al. 2020); (Sahana et Patel 2019); Yousefi et al. (2020) ; and (Zaman et al. 2021).

Chul Kim, Souk-gyu Kim et al ont utilisés les SIG et de la profondeur d'inondation par périodes de retour pour prédire les zones d'inondation vulnérables dans le bassin de la rivière Musan, On peut en déduire que les facteurs tels que :

- La topographie,
- L'utilisation des terres et
- La profondeur des inondations par périodes de retour

Sont des paramètres importants qui peuvent affecter l'occurrence et la gravité des inondations.

Le document de (Ahmed et al. 2022), utilise des données de télédétection et SIG ainsi que des données hydrologiques et météorologiques pour générer une carte des risques d'inondation dans le bassin de Marand en Iran. Ce document mentionne plusieurs paramètres qui affectent l'inondation, notamment :

- Intensité des précipitations
- Facteurs physiographiques tels que la hauteur et la pente du terrain
- Longueur du cours d'eau principal
- Différence de hauteur entre les points les plus hauts et les plus bas du bassin
- Type d'utilisation du sol

Conclusion

Dans ce chapitre nous avons défini clairement les concepts. En combinant la puissance de la télédétection, des SIG, des modèles de changements des

terres, et de l'apprentissage automatique, nous disposons d'outils avancés pour comprendre et gérer les défis environnementaux, tels que les changements climatiques et les inondations. Cette approche holistique offre des solutions sophistiquées pour une gestion durable des ressources et une résilience accrue face aux menaces environnementales.

Chapitre 2 : Zone d'étude et données

Introduction

Avant d'entreprendre toute démarche de modélisation, une étape fondamentale consiste à acquérir les séries de données relatives aux variables hydrométéorologiques essentielles dans le contexte de la zone d'étude. Cette section offre une vue d'ensemble détaillée de la zone d'étude, englobant des éléments tels que la localisation géographique, le climat, l'hydrographie, le relief, et les caractéristiques géologiques et pédologiques. De plus, une présentation approfondie des données utilisées dans cette étude sera fournie, jetant ainsi les bases nécessaires à la modélisation.

2.1. Description de la zone d'étude

2.1.1. Localisation de la zone d'étude

Le Bénin bénéficie d'un réseau assez important de cours d'eau plus ou moins permanents. Ces cours d'eau sont modestes par leur débit et leur longueur et ont un régime irrégulier (Adam et Boko 1993). Il est constitué de quatre grands bassins versants : la Volta, le Niger, le Mono-Couffo et l'Ouémé. Le fleuve a une longueur de 510 km et la superficie de son bassin versant est 51000 km² environ. Le bassin inférieur de l'Ouémé est compris entre les parallèles 8° 15' et 6° 33' N et les méridiens 2° 00' et 1° 50' Est.

En effet, la basse vallée de l'Ouémé est comprise entre les parallèles 7°12' et 6°23' Nord et les méridiens 2°21' et 2°33' Est. Localisée dans la région Sud-Est du Bénin, la basse vallée de l'Ouémé est une plaine d'inondation mesurant environ 90 km du Nord au Sud. Elle est limitée au Sud par le lac Nokoué et la lagune de Porto-Novo ; au Nord, à l'Est et à l'Ouest, ses limites sont imprécises car elles varient énormément avec l'importance des crues. Sa superficie peut donc varier de 1000 à 9000 km² selon le moment où les observations ont été faites (« Laleye, P. (1995). Ecologie comparée de deux espèces de *Chrysischtys*, poissons siluriformes (Claroteidae) du complexe lagunaire lac Nokoué-lagune de Porto-Novo au Bénin (152 p. + Annexes). Thèse de doctorat. Belgique Université de Liège. - Références - Publication de recherche scientifique », s. d.).

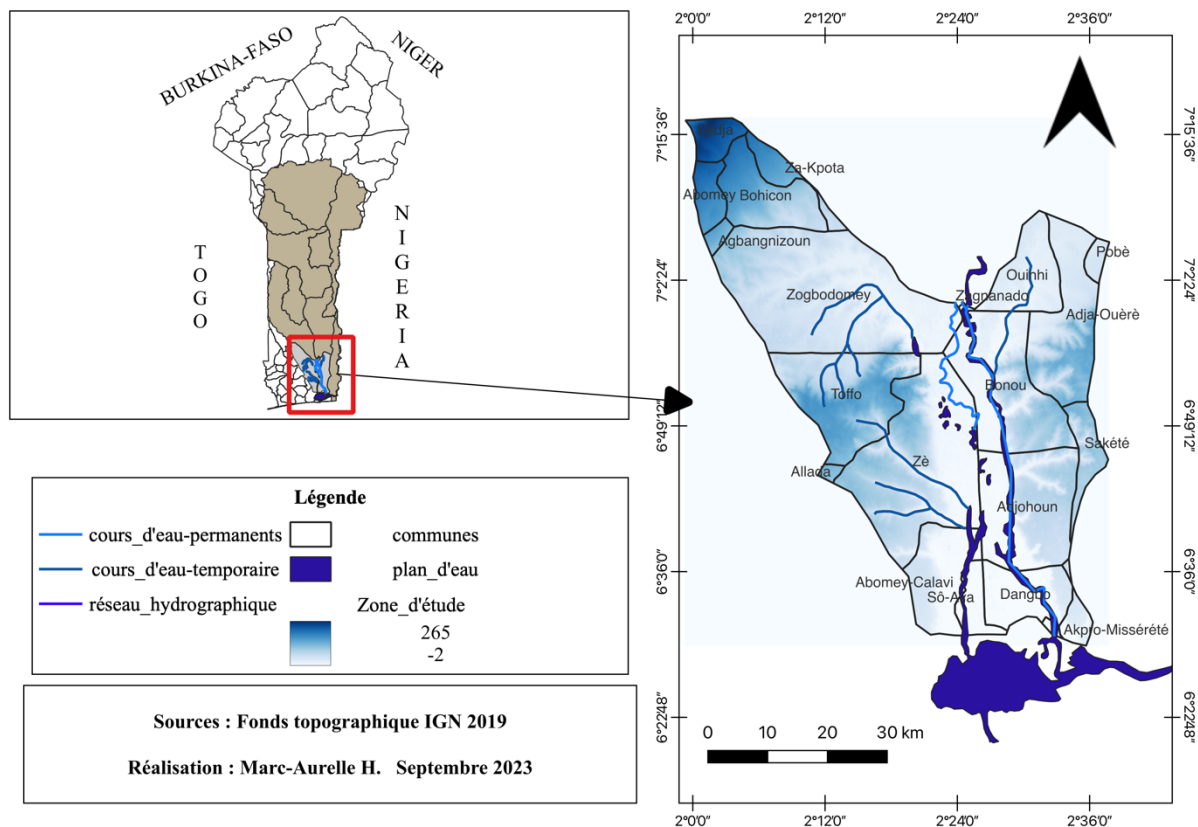


Figure 3 : Carte de la zone d'étude (bassin de la basse vallée de l'Ouémé)

2.1.2. Le climat et l'hydrographie

Le bassin de l'Ouémé a une pluviométrie très variable selon les zones climatiques et les saisons. La pluviométrie décroît du Nord au Sud jusqu'à l'entrée du delta : 1100 à 1300 mm en moyenne annuelle. Dans le delta, elle croît de 1100 mm jusqu'à 1400 mm surtout au voisinage du lac Nokoué et de la lagune de Porto-Novo (Adam et Boko 1993).

Le climat de la zone d'étude est de type subéquatorial avec une succession annuelle de quatre saisons, deux pluvieuses et deux sèches réparties comme suit :

- Une grande saison des pluies de mi-mars à mi-juillet ;
- Une petite saison sèche de mi-juillet à mi-septembre ;
- Une petite saison des pluies de mi-septembre à mi-novembre ;
- Une grande saison sèche de mi-novembre à mi-mars.

A la station synoptique de Bohicon, la température moyenne annuelle pour la période 1980-2014 est de 27,74°C tandis qu'elle est de 27,48°C à la station synoptique de Cotonou. Les températures les plus élevées s'observent en février et mars et les plus basses en juillet et août. En général, les températures moyennes

maximales sont élevées sur toute l'année et varient entre 29°C et 36°C. En revanche, les minima s'échelonnent entre 21°C et 24°.

Le fleuve Ouémé prend sa source dans la chaîne de l'Atacora au Nord-Ouest du Bénin et coule vers le Sud. La pente du lit est de l'ordre de 10 m / km dans la plus haute partie du cours d'eau et atteint 0,5 m/km, parfois moins, en arrivant dans le delta. Là, pente diminue fortement puisqu'elle n'est en moyenne que de 10,6 cm / km entre Sagon et Houédomé. Le fleuve se divise en fin de parcours en deux branches l'une se jetant dans le lac Nokoué faisant partie du delta de l'Ouémé et l'autre dans la lagune de Porto-Novo. Dans sa partie inférieure juste après sa confluence avec le Zou c'est-à-dire dans la zone sédimentaire, l'Ouémé forme une zone inondable dite Delta de l'Ouémé dans laquelle se jette certaines rivières telles que la Sô, la Zounga, l'Agbagbe, le Wovi, la Zouvi. Encore appelé Basse Vallée de l'Ouémé (BVO), le delta de l'Ouémé est un ensemble formé par trois bassins (Ouémé, Sô, Zou) et deux plans d'eau. Le système hydrographique de la basse vallée est un fort complexe en aval du confluent de l'Ouémé et du Zou, car la Sô, rivière deltaïque, parallèle à l'Ouémé, draine une partie des eaux du fleuve qu'elle alimente par ses bras d'eau, qui joignent ces rivières de second ordre. Le triangle dont la pointe est le confluent de l'Ouémé et du Zou, forme une première plaine d'inondation de 80 km², selon les études de (Adam et Boko 1993) et cité par (Amoussou 2014). Cette rivière Sô naît dans les marais du lac Hlan, et est alimentée par le débordement du Zou. Sa crue se produit pendant les fortes crues de l'Ouémé. En basses eaux, le Ouovi draine la plaine vers l'Ouémé. Mais pendant les hautes eaux, lorsque le niveau de l'eau à Bonou se situe entre 4 et 6,3 m, les eaux de la Sô refluent vers l'Ouémé. Le lac Nokoué et la lagune de Porto – Novo ne sont que les réceptacles des eaux du fleuve Ouémé et de la rivière Sô. Elles communiquent entre elles et se réalimentent en période de basses eaux, lors des marées hautes de l'Atlantique. Elles deviennent très saumâtres pendant une longue période de 4 à 5 mois dans l'année (Sossou-Agbo 2013).

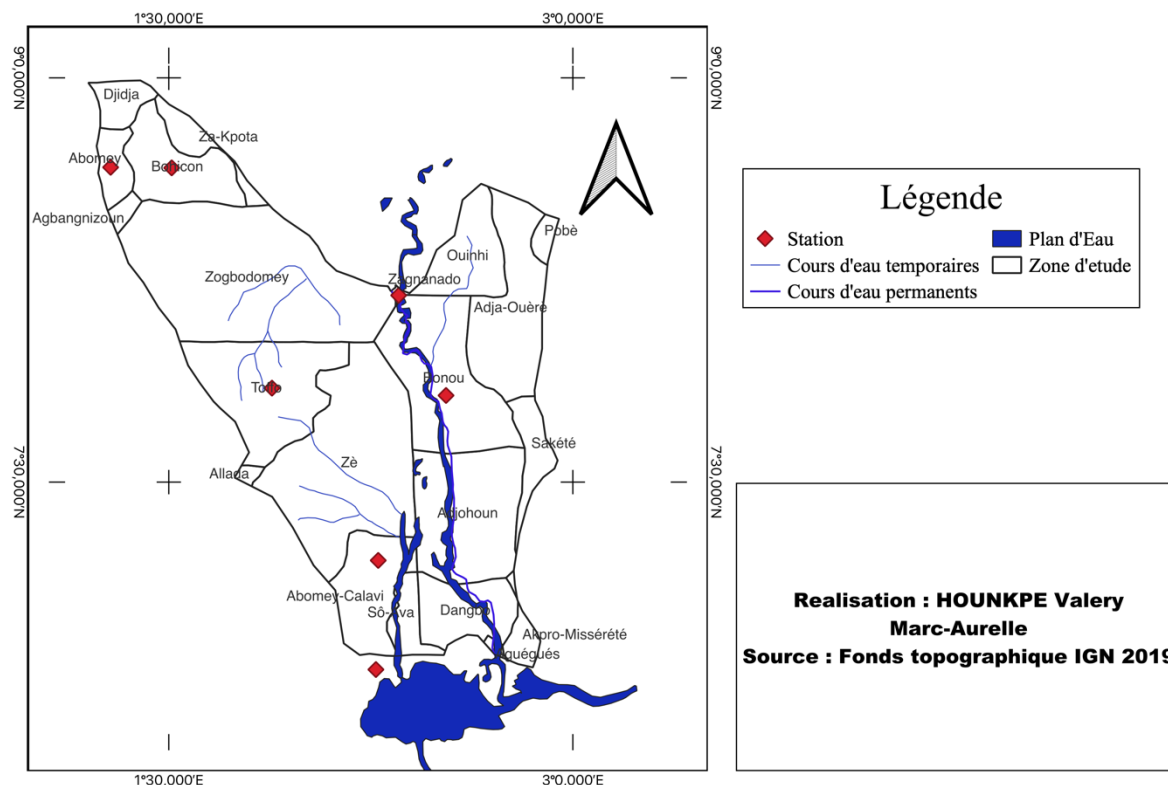


Figure 4 : Station de la zone d'étude

2.1.3. Le relief

Du point de vue du relief, deux grands domaines topographiques caractérisent le Delta de l'Ouémé :

- La plaine côtière, basse, sableuse, souvent marécageuse ne dépasse nulle part 10 m d'altitude. Elle est un complexe de plusieurs cordons littoraux séparés par des bas-fonds marécageux et des lagunes.
- Les plateaux : Les plateaux de terre de barre du Bas-Bénin se trouvant en contact avec le Delta de l'Ouémé font suite à la plaine côtière. D'altitude comprise entre 20 et 126 m, ces plateaux sont légèrement inclinés vers le sud (Adam & Boko, 1993).

2.1.4. Phénomène d'inondation dans la basse vallée de l'Ouémé

L'Ouémé est l'un des principaux fleuves du Bénin, et ses crues peuvent être influencées par divers éléments, notamment les prélèvements, la topographie, l'utilisation des terres et les pratiques de gestion des ressources en eau. La basse vallée de l'Ouémé, située au Bénin, est susceptible de connaître des inondations en raison de plusieurs facteurs (Pélissier 1962). En 2010, une forte pluviométrie a été observée sur l'ensemble du territoire national, entraînant ainsi une élévation significative du niveau d'eau, notamment dans la basse vallée du fleuve Ouémé.

Selon les données relayées, les eaux de ce fleuve ont atteint le seuil critique. Plusieurs maisons et habitations ont été ravagées par cette montée d'eau obligeant le déplacement précoce de certains habitants (Seidou et al. 2021). C'était presque tout le Sud Bénin qui était inondé en cette année-là, au point où la rentrée scolaire 2010-2011 n'a été effective dans ces localités qu'en janvier 2011 (Centre 2016).

2.2. Données d'étude

2.2.1. Données planimétriques

Les données utilisées pour l'étude de la dynamique de l'occupation des terres ici sont essentiellement des images satellitaires Landsat TM (date), ETM + (Thematic Mapper) (date) et OLI- TIRS (date) de 30 m de résolution. Ces images ont été téléchargées gratuitement sur le site earthexplorer.usgs.gov au format GEOTIFF (Path 193 et Row 53). L'image SRTM de 2000 (résolution 90 m) a été utilisée pour la réalisation des Modèles Numériques de Terrain (carte d'altitude et pente). Les cartes topographiques et fond de carte de l'Institut Géographique National (IGN) ont été également utilisés dans cette étude. A ces données planimétriques s'ajoutent quelques données socio-économiques (densité, superficies emblavées, pédologie, etc.) collectées sur le terrain et indispensables pour la simulation de l'occupation des terres. Le logiciel Terret à travers son outil Modeleur des changements des terres (LCM) a été utilisé pour la simulation de l'occupation des terres.

2.2.2. Données météorologiques

Elles concernent les hauteurs de pluie journalière, les températures minimales et maximales journalières et les insolation de la période allant de 1980 à 2016 collecté à l'Agence Nationale de la Météorologie du Bénin (Météo Bénin) qui est le principal gestionnaire sur le territoire béninois. Les différentes données nous ont permis de faire la prédiction du changement climatique. Elles ont été recueillies sur différentes stations de la zone d'étude (Figure 4).

2.2.3. Données de débits journaliers

Les données de débits journaliers utilisées dans cette étude sont celles des stations hydrométriques Hlan et Bonou (Figure 4) issues de la zone d'étude et provenant de la banque de données de la Direction Générale de l'Eau qui est le principal gestionnaire sur le territoire béninois. Aussi la direction a mis à notre disposition un document qui nous a permis de faire la classification des débits, et de déterminer les différentes probabilités dans la zone d'étude.

Conclusion

Au terme de cette section, les données recueillis, ainsi que les caractéristiques distinctives de la zone d'étude, notamment le climat, le sol, et la végétation, ont été exposées en détail. Malgré certaines lacunes constatées dans les données relatives aux précipitations et aux débits sur quelques années spécifiques, il est à noter que le sous bassin de la vallée de l'Ouémé considéré révèle des traits particulièrement propices à l'atteinte de nos objectifs. Ces conditions offrent une base solide pour les investigations à venir dans le cadre de notre analyse.

Chapitre 3 : Méthodologie

Introduction

Nous présenterons dans ce chapitre une approche novatrice de modélisation intégrée des inondations. Cette étude multidisciplinaire combine une analyse historique des paramètres d'inondation. Elle utilise la méthode Modeleur des changements des terres (LCM) pour quantifier les changements du terrain, la simulation des changements climatiques avec le logiciel LARS-WG et une méthodologie de prédiction des débits basée sur l'apprentissage automatique (Mikhail Semenov et Barrow 2002).

3.1. Détermination des paramètres qui influent sur l'inondation dans la basse vallée de l'Ouémé

Cette étape vise à effectuer une analyse approfondie des méthodes existantes dans la détermination des paramètres qui influent sur l'inondation. Pour ce faire nous allons suivre quelques étapes :

Recherche documentaire : Une recherche approfondie dans des bases de données scientifiques, telles que Google Scholar, IEEE Xplore, SCOPUS, Web of Science, Science Direct, ou des bases de données spécialisées en hydrologie et en géographie tels que **Hydrology and Earth System Sciences (HESS) Discussions, Hydrological Sciences Journal**.

Mots-clés pertinents : Trouver des mots-clés pertinents pour la recherche, tels que "facteurs d'inondation", "causes d'inondation", "paramètres hydrologiques", "modélisation des inondations", ou d'autres thèmes liés à notre recherche etc.

Sélection des articles : Identifier les articles pertinents qui abordent la question des paramètres influant sur les inondations. S'assurer que les articles sont basés sur des études scientifiques et des données fiables.

Analyse des articles : Lorsque les séries d'articles sont réunis, faire une lecture attentivement et examinez des paramètres mentionnés comme ayant une influence sur les inondations. Ces paramètres peuvent inclure la pluviométrie, la topographie, la capacité d'absorption des sols, la gestion des eaux pluviales, le changement climatique, etc.

Synthèse des résultats : Faire une synthèse des résultats et des conclusions des articles après lecture. Après cette synthèse déterminer les paramètres qui reviennent fréquemment et qui sont largement reconnus comme ayant un impact significatif sur les inondations.

Validation des résultats : Si possible, vérifier si les résultats des articles sont cohérents avec d'autres recherches similaires. Cela peut aider à confirmer la validité des paramètres identifiés.

Les paramètres influant sur les inondations peuvent varier en fonction de la région géographique, du climat et d'autres facteurs locaux. Il est donc important de prendre en compte le contexte géographique de la zone d'étude

3.2. Analyse de l'évolution des précipitations et de l'occupation de sols dans la basse vallée de l'Ouémé

3.2.1. Méthode des modeleurs des changements des terres

C'est une technique utilisée dans les systèmes d'information géographique (SIG) pour analyser et quantifier les changements du terrain au fil du temps. Il s'agit de comparer des modèles numériques d'élévation (DEM) de différentes périodes de temps pour identifier et mesurer les changements topographiques. Voici une explication du fonctionnement de la méthode Modeleur des changements des terres :

Préparation des données : Obtenir des DEM pour la zone d'étude à partir de différentes périodes. Les DEM représentent les valeurs d'altitude du terrain et sont généralement dérivés d'images satellite, de données LiDAR (Light Detection and Ranging) ou d'autres sources d'altitude. S'assurer que les DEM ont une résolution et une projection spatiales cohérentes.

Prétraitement : Il est crucial de prétraiter les DEM pour supprimer les erreurs systématiques ou les artefacts. Cela peut impliquer d'effectuer des corrections pour les biais systématiques, tels que les décalages verticaux ou les corrections de pente, pour s'assurer que les DEM sont correctement alignés.

Co-enregistrement : une fois le prétraitement fini, aligner les DEM dans l'espace pour vous assurer qu'ils sont géographiquement alignés et qu'ils ont la même étendue et la même taille de cellule. Cette étape consiste à effectuer des transformations géométriques pour aligner avec précision les pixels ou les cellules de la grille des DEM.

Détection des changements : comparer les DEM alignés pixel par pixel pour détecter les changements d'élévation. Cela se fait généralement en soustrayant les valeurs d'élévation des cellules correspondantes dans les deux DEM. Le résultat obtenu est une nouvelle couche raster représentant les différences d'altitude (ou changements) entre les deux périodes.

Seuil : appliquer un seuil au raster de différence d'altitude pour distinguer les changements de terrain significatifs des fluctuations mineures ou des erreurs de

mesure. Le seuil peut être déterminé en fonction de la résolution des DEM et du niveau de changement jugé significatif pour l'objectifs de l'étude.

Analyse et visualisation : analyse le raster de détection de changement résultant pour identifier les zones de changement de terrain. Visualiser les changements en superposant le raster sur une carte ou en générant des cartes thématiques pour mettre en évidence des types de changements spécifiques, tels que l'érosion, le dépôt ou les modifications du relief.

Quantification : pour quantifier l'étendue et l'ampleur des changements de terrain, calculer diverses mesures statistiques, telles que la surface totale affectée par le changement, la différence d'altitude moyenne ou le volume de matériau déplacé. Ces métriques fournissent des informations précieuses pour comprendre et évaluer la dynamique du terrain.

La méthode modelleur des changements des terres peut être appliquée à diverses applications, notamment la surveillance des changements de paysage, l'analyse des schémas d'érosion ou de dépôt, l'étude de l'évolution du relief ou l'évaluation des impacts des processus naturels ou anthropiques sur le terrain.

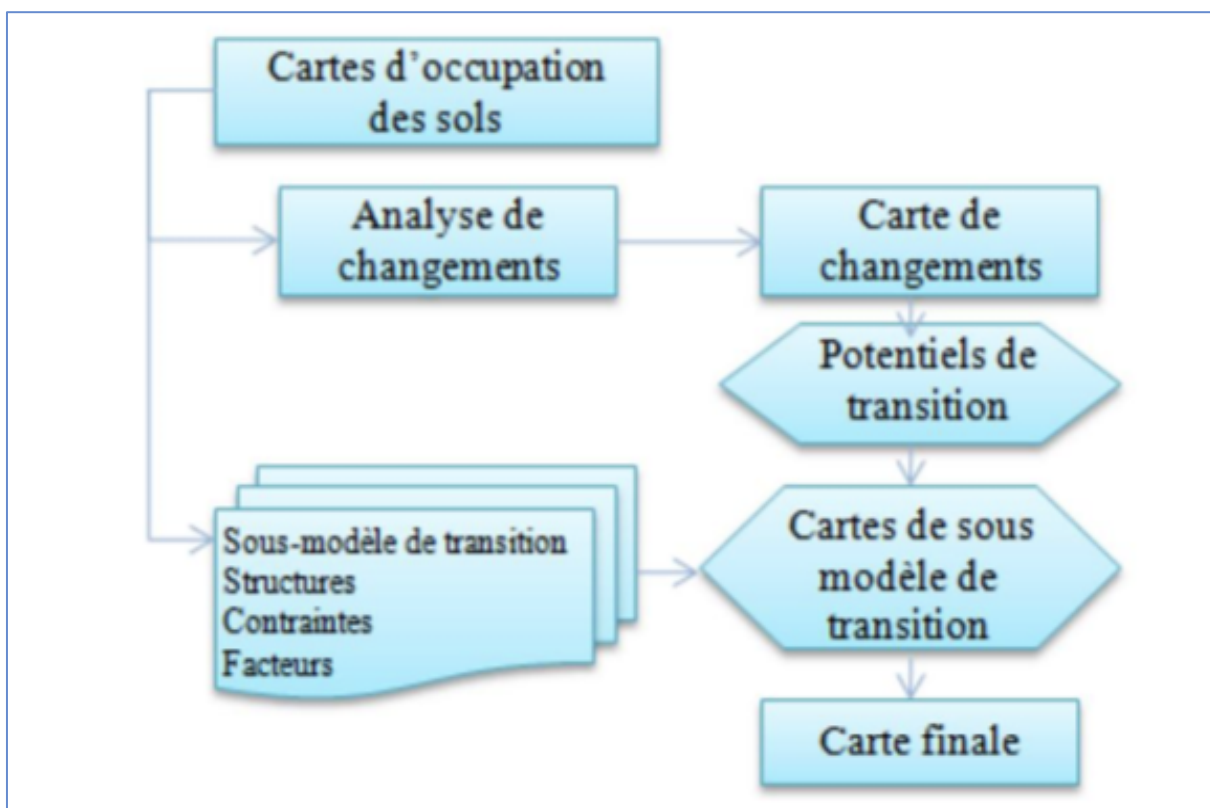


Figure 5 : Structure du modèle LCM

3.2.1.1. Dynamique de l'occupation des terres

Le logiciel TerrSet a permis de réaliser trois cartes d'occupation des terres à partir des images satellitaires de 1990, 2000 et 2020. Ici les bandes radiométriques téléchargées et susceptibles de nous fournir des informations sur le couvert végétal ont été superposées par synthèse additive des couleurs primaires. Ainsi la composition colorée en fausses couleurs avec la combinaison des bandes 234 pour l'images TM/ETM+ et 345 pour l'images OLI TIR ont été utilisées. La méthode de classification supervisée selon l'algorithme du maximum de vraisemblance a été utilisée. L'indice de Kappa a également permis de valider les classifications. Les cartes d'occupation des terres réalisées ont fourni les données de base pour l'analyse et la quantification du paysage. Les différents indices suivants ont été calculé pour mieux apprécier la dynamique observée (Thierry et al. 2018):

Calcul du LUCC budget

La matrice de transition des trois cartes d'occupation des terres (1990, 2000 et 2020) permet de faire le bilan des gains, des pertes et des persistances en termes de superficie des catégories d'occupation des terres entre deux dates différentes. Le gain brut pour chaque catégorie est calculé en soustrayant la persistance du total de la ligne, tandis que la perte brute est calculée en soustrayant la persistance du total de la colonne.

L'équation (1) et (2) donne la méthode de calcul du gain (G) et de la perte (L) observé au niveau de chaque catégorie d'occupation des terres entre deux dates.

$$G_j = \frac{\{(\sum_{i=1}^j S_{jj}) - S_{jj}\} \times 100}{\sum_{i=1}^j S_{ij}} \quad \text{ou} \quad G_j = (\sum_{i=1}^j S_{jj}) - S_{jj} \quad (1)$$

$$L_i = \frac{\{(\sum_{j=1}^j S_{ij}) - S_{ii}\} \times 100}{\sum_{j=1}^j S_{ij}} \quad \text{ou} \quad L_i = (\sum_{j=1}^j S_{ii}) - S_{ii} \quad (2)$$

G_j : Gain de superficie (ou en proportion) d'une catégorie d'occupation des terres j

L_i : Perte de superficie (ou en proportion) d'une catégorie d'occupation des terres i

S_{ii} : Superficie stable d'une catégorie d'occupation i entre deux dates

S_{jj} : Superficie stable d'une catégorie d'occupation i entre deux dates

S_{ij} : Superficie d'une catégorie d'occupation i transformée en j entre deux dates
 S_{ji} : Superficie d'une catégorie d'occupation j transformée en i entre deux dates

L'équation (3) exprime le changement total. La superficie totale changée est calculée en additionnant toutes les entrées de la matrice puis en soustrayant les persistances.

$$\text{Changement Total} = \frac{\sum_{j=1}^J \{(\sum_{i=1}^J S_{ij}) - S_{jj}\}}{\sum_{j=1}^J \sum_{i=1}^J S_{ij}} \quad (3)$$

L'équation (4) indique comment calculer le changement net pour une catégorie j en ajoutant les transitions dans lesquelles la catégorie j gagne, c'est-à-dire C_{ij} , puis en soustrayant la transitions dans lesquelles la catégorie j perd, c'est-à-dire C_{ji} . Autrement dit, pour une catégorie donnée, dans la matrice elle correspond à la différence entre le gain brut et la perte brute.

$$\text{Changement Net} = \frac{\sum_{j=1}^J \{ \text{MAXIMUM}[0, \sum_{i=1}^J (S_{ij} - S_{ji})] \}}{\sum_{j=1}^J \sum_{i=1}^J S_{ij}} \quad (4)$$

L'équation (5) permet de calculer la balance qui est la différence entre le changement total et le changement net. Cette variable exprime un changement de lieu sans changement de quantité. (Alo et Pontius 2008) l'expliquent comme un désajustement qui pourrait être résolu en réorganisant les pixels d'une carte pour s'adapter autant que possible avec l'autre carte.

$$\text{Swap} = \text{Change Total} - \text{Changement Net} \quad (5)$$

Calcul du taux d'évolution

La dynamique de chaque catégorie d'occupation sera appréciée en calculant le taux d'évolution $E(i, k)$ des superficies d'occupation des terres de la façon suivante :

Soit S_i la superficie d'une catégorie d'occupation des terres de l'année i et S_k , celle d'une catégorie d'occupation des terres de l'année k , avec $k > i$. $E(i, k)$ sera égale à :

$$E(i, k) = \frac{S_k - S_i}{S_i} \times 100 \quad (6)$$

Si $E(i,k) = 0$, on conclut qu'il y a stabilité de cette catégorie d'occupation des terres;

Si $E(i,k) < 0$, on conclut qu'il y a régression de cette catégorie;

Si $E(i,k) > 0$, il y a extension ou évolution de cette catégorie.

3.2.1.2. Simulation de l'occupation des terres

Le modèle de prédiction utilisé ici est Modeleur des changements des terres (LCM) implémenté dans le logiciel TerrSet (ex Idrisi). Ce modèle s'appuie sur les connaissances passées et actuelles de l'occupation des terres pour prédire le futur. Plusieurs facteurs influençant les changements d'occupation des terres ont été intégrés dans ce modèle. Il s'agit de la pente, l'altitude, la pédologie, la distance aux agglomérations, la distance aux champs, la distance aux plantations, la distance aux routes et la densité. L'indice V de Cramer a été calculé pour chaque facteur explicatif et a permis de retenir ceux qui contribuent le mieux à la mise en place de chaque catégorie d'occupation des terres. Les variables explicatives utilisés ici sont de deux types : dynamiques et statiques. Les variables statiques : (altitude, pente, pédologie) (Thierry et al. 2018).

Les cartes de prédiction ont été élaborées par le réseau de neurones multicouche à partir des cartes de transitions observées et des variables explicatives. Ceux-ci assignent à chaque pixel une catégorie de détérioration à l'aide du processus de décision multi-objectif (MOLA). Pour attribuer la quantité de changement, l'algorithme s'appuie sur la matrice de probabilité de transition 2020-2040 générée.

Deux types de carte ont été générée à la sortie du modèle : une carte de simulation dite « hard », qui alloue à chaque pixel de la carte un type d'occupation des terres et un seul, et un résultat de simulation dit « soft », qui associe à chaque pixel une valeur d'intensité de changement au cours du temps, une sorte de propension générale au changement d'occupation des terres quelle qu'en soit la nature. Cette propension est calculée par cumul des probabilités de changement à chaque pas de la simulation.

3.2.2. Utilisation du logiciel Lars-WG pour la prédiction des changements climatiques

LARS-WG simule une série d'années climatiques sans logique de succession temporelle. Des séries climatiques synthétiques ont été produites avec le générateur climatique LARS - WG (Stochastic Weather Generator). LARS - WG est un outil numérique conçu par le Dr Mikhail Semenov (Long Ashton

Research Station, IACR, Université de Bristol) (Mikhail Semenov et Barrow 2002). Il produit des séries climatiques journalières artificielles mais admissibles de précipitations, températures minimales et maximales et rayonnement solaire. Ces séries sont valables pour un lieu donné ; elles sont produites à partir des distributions statistiques caractéristiques pour ce lieu des différentes variables du climat. L'intérêt de cet outil est de pouvoir fournir une infinité d'années climatiques probables en ce lieu.

Propriété très intéressante dans le cas des études sur le changement climatique, où les séries, elles-mêmes simulées par les MCGs, sont courtes et chères. Les scénarios des MCGs vont apporter au générateur les paramètres de changement climatique qu'il utilise comme suit : le paramètre de variabilité relatif à la moyenne des températures d'un mois donné est additionné à l'ensemble des valeurs des températures journalières minimales et maximales de ce mois ; LARS - WG multiplie chaque valeur qu'il choisit aléatoirement pour les précipitations et l'écart type des températures, par le paramètre de changement climatique correspondant. L'humidité et le vent ne sont pas générés par LARS - WG. L'estimation de l'humidité relative a été faite selon l'hypothèse que la quantité d'eau dans l'atmosphère est constante. La part d'humidité relative sera alors modifiée par le changement de la température de l'air. Pour le vent, les prévisions scientifiques concernant la modification du nombre de jours de vent et de sa force sont quasi inexistantes. Afin d'éviter toutes affirmations erronées, l'hypothèse d'une variation nulle du vent dans les décennies à venir a été posée

LARS-WG stochastic weather generator est un modèle qui simule les séries chronologiques météorologiques journalières. Il peut être utilisé dans les cas suivants :

- Pour générer de longues séries chronologiques servant dans l'évaluation des risques climatiques ;
- Pour combler les lacunes d'observations manquantes ;
- Pour servir comme un outil qui peut produire des scénarios climatiques au pas de temps quotidien pour l'évaluation de l'impact du changement climatique ;

LARS - WG nous a permis de produire des séries climatiques journalière. Le procédé de génération de données synthétiques peut être divisé en trois étapes distinctes :

- a) **Model Calibration - SITE ANALYSIS** : Les données d'observation météorologiques sont analysées pour déterminer leurs caractéristiques

statistiques. Ces informations sont stockées dans deux fichiers de paramètres (*.stx et *.wgx).

- b) **Model Validation - QTEST** - : Les caractéristiques statistiques des données observées et synthétiques sont analysées pour déterminer s'il y a des différences statistiquement significatives entre elles (*.tst).
- c) **Generation of Synthetic Weather Data – GENERATOR** : Les fichiers paramètres dérivés à partir des données météorologiques observées au cours de la procédure d'étalonnage sont utilisés pour générer des données météorologiques synthétiques ayant les mêmes caractéristiques statistiques que les données observées d'origine, mais différentes de jour en jour.

Étalonnage du modèle

L'étalonnage du modèle est effectué dans LARS-WG en mettant en œuvre l'analyse du site qui analyse les données météorologiques quotidiennes observées de 1986 à 2016 (précipitations, température maximale et minimale et rayonnement solaire) à des stations de la zone d'étude afin de déterminer leurs caractéristiques statistiques. L'information résultante est stockée dans deux fichiers à savoir un fichier **wgx** qui est le fichier de paramètres du site et un fichier **stx** qui est le fichier contenant des tests statistiques supplémentaires. Les deux fichiers ne doivent pas être modifiés. Ils sont utilisés plus tard dans le processus de génération (Benaichata 2016).

Validation ou performance du modèle

Dans cette étape, les caractéristiques statistiques des données observées sont comparées à celles des données synthétiques générée. Le Q-TEST est appliqué en utilisant les paramètres dérivés des données observées d'origine. Cela implique plusieurs tests statistiques dont le Kolmogorov-Smirnov (K- S) Test et le test du Khi-carré pour comparer la probabilité des distributions, le Student-test (ou test t) pour comparer les moyennes et le Fisher-test (ou test F) pour comparer les écart- types. Les deux derniers tests déterminent, à travers la p-value, si les valeurs moyennes et les écarts-types des données synthétiques sont significativement différentes de celles de l'ensemble de données observées d'origine (Benaichata 2016) (Ma Semenov et al. 1998).

3.3. Développement d'un modèle basé sur l'apprentissage automatique qui prédire la probabilité future des inondations

Cette méthodologie est basée sur l'apprentissage supervisé notamment les méthodes de classification (Random-Forest et Support vector machine). Les principales étapes de la méthodologie proposée sont entre autres : i) collecte de données ii) prétraitement des données et analyse statistique, iii) standardisation des données pour normaliser les données collectées, iv) la construction et la validation des modèles, v) amélioration des performances du modèle. Enfin, les deux modèles ont été utilisé pour la prévision des débits futurs (Ouedraogo 2023).

Collecte de données : la première étape de notre méthodologie est la collecte de données. C'est l'étape la plus importante de tout projet de gestion de données de masse.

Prétraitement de données : La deuxième étape consiste à traiter les valeurs incohérentes et /ou manquantes. Dans notre étude, cette étape a été nécessaire pour générer des données de qualité afin d'améliorer les performances des modèles SVM et Random Forest ;

Normalisation des données : la troisième étape est la normalisation des données collectées. L'objectif de la normalisation est de modifier les valeurs des données (celle de précipitations, de débit, de températures et d'insolations) du jeu de données pour utiliser une échelle commune, sans que les différences de plages de valeurs ne soient faussées et sans perte d'informations ;

Division de l'ensemble de données : Nous utilisons dans notre étude les données des stations pluviométriques, des stations hydrométriques et synoptiques (Annexe 2) de la zone d'étude sur la période 1980-2016. Pour élaborer notre modèle de prédiction basée sur les modèles de classifications de l'apprentissages supervisé (Random-Forest et Support vector machine), notre jeu de données a été reparti en trois grandes parties :

- **Ensemble d'entraînement :** un premier jeu de donnée servira à l'entraînement des modèles (training set). Cet ensemble couvre 80% du jeu de données. Il permet d'apprendre les différents poids des neurones constituant notre réseau.
- **Ensemble de test :** un troisième jeu de données servira à tester les performances réelles de notre modèle. Il s'agit de l'échantillon de test (test set). Il n'a été utilisé qu'une fois le modèle complètement entraîné (à l'aide des ensembles d'entraînement et de validation). Il est utilisé pour fournir une évaluation impartiale de l'ajustement du

modèle final sur l'ensemble de données de formation. Ce jeu de données représente aussi 20% de notre base de données.

Validation des modèles : Les modèles sélectionnés sont évalués en utilisant des données de validation. Ces données servent à évaluer et valider le modèle, assurant ainsi des prévisions adéquates.

Voici quelques-unes des métriques couramment utilisées, qui peuvent être appliquées aux modèles de classification tels que Random Forest et SVM dans notre contexte de prédiction des inondations (Chloé et Al 2019b; 2019a) :

Matrice de Confusion

La matrice de confusion montre le nombre de vrais positifs, de vrais négatifs, de faux positifs et de faux négatifs. Cela offre une vue détaillée des performances du modèle.

Précision (Precision)

La précision mesure le nombre de vrais positifs parmi les instances prédites positives. C'est le rapport entre les vrais positifs et la somme des vrais positifs et des faux positifs.

$$\text{Précision} = \frac{\text{Vrais Positifs}}{\text{Vrais Positifs} + \text{Faux Positifs}} \quad (7)$$

Une précision de 80% ou plus peut être considérée comme bonne dans de nombreuses situations

Rappel (Recall)

Le rappel mesure le nombre de vrais positifs parmi toutes les instances réellement positives. C'est le rapport entre les vrais positifs et la somme des vrais positifs et des faux négatifs.

$$\text{Rappel} = \frac{\text{Vrais Positifs}}{\text{Vrais Positifs} + \text{Faux Négatifs}} \quad (8)$$

F1-Score

Le F1-Score est la moyenne harmonique entre la précision et le rappel. Il fournit une mesure équilibrée entre les deux.

$$\text{F1 - Score} = 2 * \frac{\text{Précision} * \text{Rappel}}{\text{Précision} + \text{Rappel}} \quad (9)$$

Exactitude (Accuracy)

L'exactitude mesure le pourcentage total d'instances correctement classées par le modèle.

$$\text{Exactitude} = \frac{\text{Vrais Positifs} + \text{Vrais Négatifs}}{\text{Total d'instances}} \quad (10)$$

Une exactitude supérieure à 80% peut être considérée comme bonne dans de nombreux cas.

Courbe ROC (Receiver Operating Characteristic)

La courbe ROC et l'Aire sous la courbe ROC (AUC-ROC) sont utilisées pour évaluer la capacité d'un modèle à discriminer entre les classes. Cela est particulièrement utile pour les modèles de classification. Une AUC-ROC supérieure à 0.8 est souvent considérée comme acceptable, tandis qu'une valeur supérieure à 0.9 est généralement considérée comme excellente.

Conclusion

Pour aboutir à notre objectif principal qui est de développer un modèle intégré d'analyse des risques d'inondation prenant en compte le changement climatique, l'évolution de l'occupation des terres en utilisant des techniques d'apprentissage automatique avancées, nous avons présenté différentes méthodologie. De plus nous avons présenté les étapes pour la conception de notre modèle.

Chapitre 4 : Résultats et discussion

Introduction

Ce chapitre présente les conclusions d'une analyse approfondie sur les paramètres influençant les inondations, explorant des facteurs tels que l'altitude, la pente, l'occupation des sols et les précipitations. L'utilisation de techniques d'apprentissage automatique et l'intégration de scénarios climatiques projettent une lumière significative sur la susceptibilité future aux inondations dans la région d'étude.

4.1. Détermination des paramètres qui influent sur l'inondation dans la basse vallée de l'Ouémé

La détermination des paramètres qui influent sur l'inondation dans la basse vallée de l'Ouémé implique une analyse approfondie de divers facteurs physiques, météorologiques, géographiques et humains. Ainsi le tableau 3 les facteurs qui influent sur l'inondation et leurs buts d'utilisation. Il ressort de tableau 3 que neuf facteurs influent sur l'inondation, il s'agit de la pente, de la précipitations, de l'indice d'humidité topographiques, de la densité fluviale, de l'indice de puissance de flux, de l'élévation, de l'utilisation des terres, de la distance des rivières et de l'aspect.

Tableau 2 : le but de l'utilisation des facteurs

Facteur de conditionnement	But de l'utilisation
Pente	Il contrôle la vitesse du ruissellement de surface et les opportunités de perméation (Tien Bui et al. 2019). Dans les zones à fortes pentes le temps d'infiltration diminue ce qui entraîne une augmentation du débit d'eau. La pente peut contrôler la quantité d'eau qui s'infiltre dans le sol et elle a un impact sur la profondeur et la capacité de rétention d'eau (Dodangeh et al. 2020). La pente a un impact sur la quantité de ruissellement par ses effets sur la profondeur du sol et la longueur de la pente. Des pentes plus raides génèrent une vitesse plus élevée et favorisent l'augmentation des volumes vers le bas.

Indice d'humidité topographique (TWI)	Le TWI est utilisé pour quantifier l'effet topographique sur les processus hydrologiques. Il montre la tendance de l'eau à s'accumuler à un endroit spécifique ou à se déplacer vers le bas en raison de la force gravitationnelle (Tehrany, Jones, et Shabani 2018). De plus TWI montre des schémas spatiaux d'humidité du sol et d'intensité d'écoulement (Tien Bui et al. 2019)
Densité fluviale	Les systèmes de drainage influencent souvent le débordement des mers et les inondations connues dans une zone, Une densité de drainage plus élevée entraîne une infiltration plus faible et un ruissellement plus élevé (Arabameri et al. 2020).
Aspect	L'aspect a un effet indirect sur les inondations en contrôlant certains facteurs environnementaux tels que la quantité de précipitations, le développement de la végétation, évapotranspiration et la teneur en humidité du sol en définissant où la foudre frappe (Dodangeh et al. 2020).
Indique de puissance de flux (SPI)	Le SPI est un facteur pour exprimer le pouvoir érosif de l'eau qui coule (Tehrany, Jones, et Shabani 2018).
Élévation	L'altitude joue un rôle important dans la direction et la quantité de ruissellement de surface dans une zone (Tiryaki et Karaca,2018). De plus, différentes altitudes ont modifié les caractéristiques climatiques et peuvent entraîner des différences dans les conditions de la végétation et du sol (Natarajan et al. 2021 : Rahmati et al,2016). Elle est inversement proportionnelle à l'inondation : plus l'altitude

	est élevée, plus l'inondation est faible probabilité (Tehrany et al., 2014b ; Youssef et al., 2016)
Utilisation des terres	Il peut affecter indirectement ou directement le ruissellement de surface, l'évapotranspiration, la perméabilité et le taux de pénétration (Yarivan et al. 2020), Dans les zones urbaines, le ruissellement augmente en raison d'un sol imperméable étendu. Dans les terres agricoles en jachère, le manque de couverture végétale pour contrôler et empêcher écoulement rapide de l'eau à la surface du sol augmente le ruissellement (Ullah et Zhang.,2020).
Distance de la rivière	Les débordements de rivière sont critiques pour le déclenchement d'une Inondation. La probabilité d'inondation est plus faible dans les zones éloignées des lits des rivières (Das, 2020 ; Kazakis et al., 2015 : Tehrany et Kumar 2018).
Précipitations	Les précipitations sont le facteur d'inondation le plus fiable qui ne Peut imaginer la survenue d'inondations qu'elle (Chowdhuri et al.,2020). Une inondation s'est produite en raison d'un volume énorme de ruissellement et d'un excès de berge résultant de pluies excessivement fortes ou des pluies prolonges (ahmadou et al. ,2019 ; Shahiri Tabarestani et Afzalimehr,2021).

Ces différents facteurs ne sont pas fixes et aucune formule n'est disponible pour la sélection de ces facteurs. Le choix des différents facteurs ou paramètres résultent juste de la revue de littérature et de paramètre géologie, physique et naturelle de la zone a étudiée. Dans le cas de notre étude nous avons faire le choix de cinq grands paramètres ou facteurs qui influent sur l'inondation (Altitude, Pente, Indice Topographique, l'occupation du sol et la précipitations). Les autres facteurs de conditionnement d'inondations n'ont pas été prise en compte en raison de la non disponibilité des sources de données claires et fiables pour la cartographie de ces dernières.

La figure 6 présente la carte des altitudes. De cette figure il ressort que plus on s'éloigne de la mer, l'altitude augmente or qu'elle régresse lorsque on s'approche

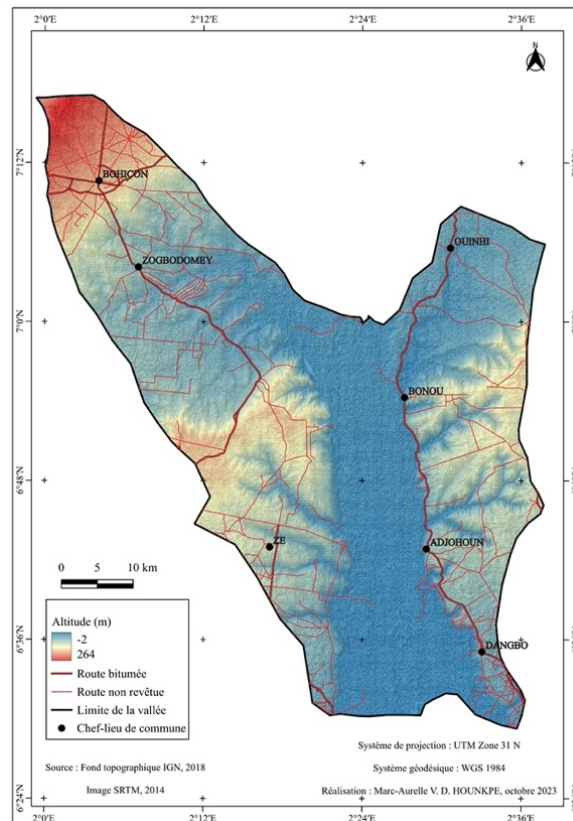


Figure 6 : La carte des altitudes

La figure 7.A présente la carte des pentes. De cette figure il ressort qu'au niveau 0 nous sommes proches de la mer et plus on s'éloigne pour la partie nord de notre zone d'étude la pente augmente considérablement de 264m. La figure 7.B présente la carte de l'indice topographique d'humide. On remarque sur cette figure que le nombre d'indices topographique d'humide élevé est de 19.5 et se retrouve proche des zones marécageuses, cet indice diminue lorsqu'on s'éloigne des zones marécageuses.

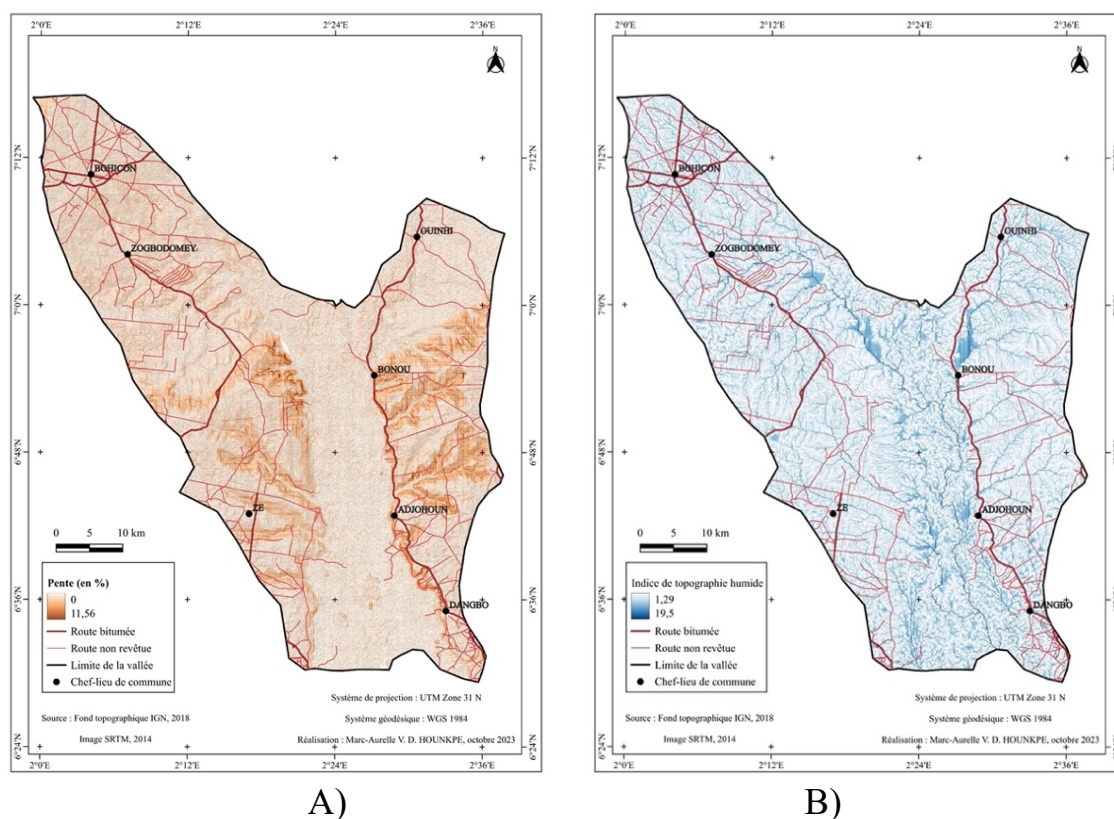


Figure 7 : Carte des facteurs qui influent sur l'inondation

A) Cartes des pentes ; B) Cartes des indices topographiques d'humidité

4.2. Analyse de l'évolution des précipitations et de l'occupation de sols dans la basse vallée de l'Ouémé

4.2.1. Changements d'occupation des terres

La précision globale obtenue est de 85 % avec un coefficient de Kappa de 0,85 pour l'ensemble des images cartographié. Ces valeurs signifient que plus de 85 % des pixels des deux images ont été correctement classés conformément aux données de vérité-terrain. De façon globale, les valeurs des différents indicateurs de précision de la classification supervisée obtenues pour les différentes images, traduisent d'une part, de la bonne qualité des échantillons et d'autre part, de la bonne correspondance entre le résultat de la classification et la réalité spatiale contenue dans les images. Les cartes passées et contemporaines ont été classées en suivant des classes : Agglomération, agricoles, forestières, champs jachères, plantation, Route etc. (Fig. 8).

La figure 8 et le tableau 5 présentent les résultats des cartes d'occupation du sol en 1990 (a), 2000 (b) et 2020 (c). De l'analyse de cette figure, il est remarqué une augmentation des superficies, entre les années 1990 et 2000 des

classes d'occupations comme les champs-jachères (174%), les agglomérations (35%), et l'apparition de la classe des plantations et agroforêts. Par contre, il est noté que durant cette même période, les superficies des unités d'occupation des terres telles que les forêts, les savanes et les zones de plantation ont régressé respectivement de 32% ; 72% et 61%. Pour la période 2000-2020, des diminutions annuelles de 25% pour toutes les unités d'occupations du sol sauf les agglomérations, les plan d'eau et savanes saxicoles sont observées. On note une progression de ces unités soit respectivement de 94%, 5% et 25%. Ce qui a conduit à une diminution des superficies des forêts, des plantations et agroforêts, des champs et jachères ainsi que des forêts claires savanes boisé et savane dense.

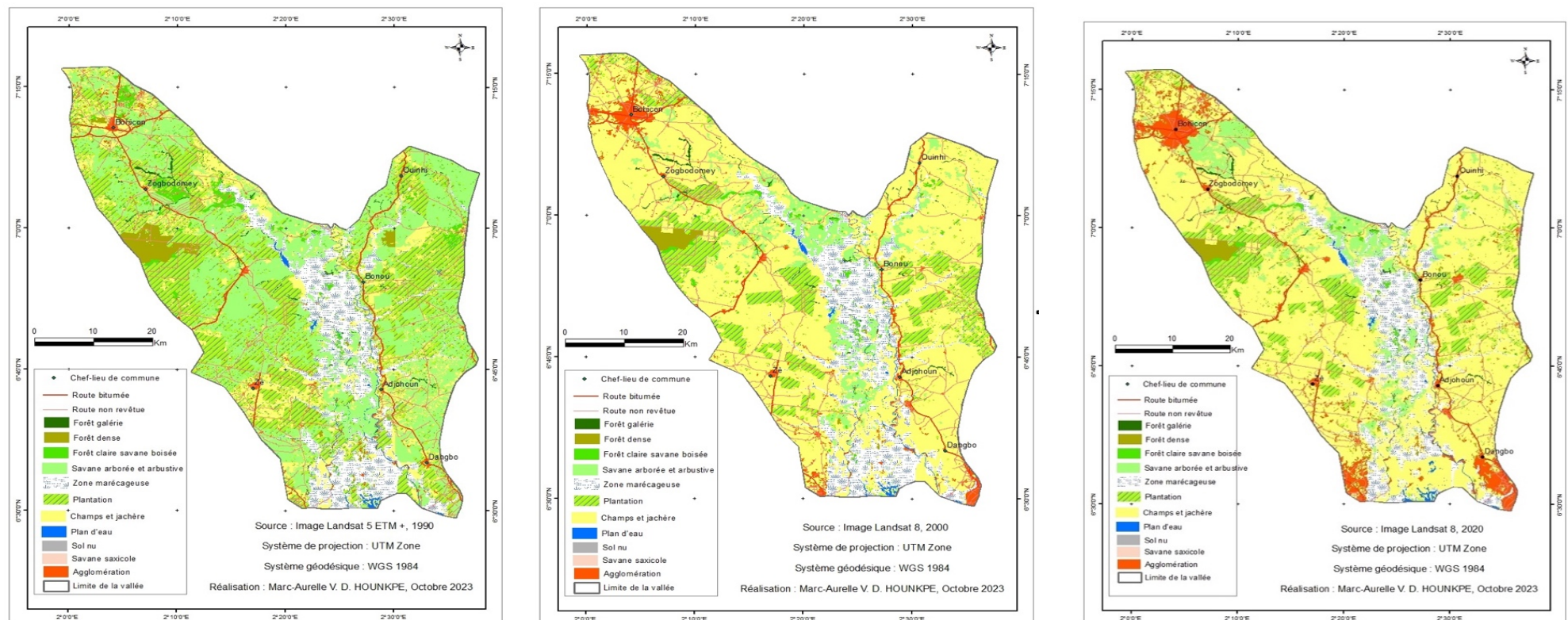


Figure 8 : Cartes d'occupation des terres/ sols (a. 1990 ; b. 2000 ; c. 2020)

Tableau 3: Superficies et taux d'évolution des classes d'occupations entre 1990, 2000 et 2020

	Superficie 2020	Superficie 2000	Superficie 1990	Taux évolution 1990-2000 (%)	Taux évolution 2000-2020 (%)
Champ jachère	196544,22	196752,51	71569,62	174,91065	-0,105864
Forêt claire savane boisée	3038,24	3086,31	4541,62	-32,043852	-1,5575234
Forêt dense	3134,64	3350,94	4384,46	-23,572344	-6,4549052
Forêt galerie	1651,53	1627,5	1691,58	-3,7881744	1,4764977
Zone marécageuse	36148,84	39241,12	39297,2	-0,1427074	-7,8802032
Agglomération	20875,86	10733,84	7957,07	34,8968904	94,486409
Plan d'eau	2125,33	2021,18	2021,18	0	5,1529304
Plantation	33599,29	39912,02	104308,45	-61,736542	-15,816614
Savane arborée et arbustive	22682,88	23091,43	84045,67	-72,52514	-1,7692711
Sol nu	10,86	10,86	10,86	0	0
Savane saxicole	77,99	61,97	61,97	0	25,8512183
Total	319889,7	319889,7	319889,7		
Indice kappa	0,85	0,85	0,85		

4.2.2. Intensité des changements par catégorie d'occupation des terres entre 1990-2000 et 2000-2020

La figure 9 présente l'intensité de changement par catégorie d'occupation des terres entre 1990 et 2000 dans la zone d'étude. L'observation de cette figure 9 montre des changements au niveau des unités d'occupation des terres. Ces changements se caractérisent par les pertes, les gains et de la stabilité. Ainsi, les plus forts gains ont été observés au niveau des mosaïques de champs et jachères

(70%). Les plus fortes pertes ont été observées respectivement au niveau des champs jachères (24 %), les savanes saxicoles (5%). Les forêts denses (6 %) ; les plantations (15%) ; les forêts galeries (4 %) et les forêts claires et savanes boisées (5 %) sont restées plus ou moins stables.

La figure 10 présente l'intensité de changement par catégorie d'occupation des terres entre 2000 et 2020 dans la zone d'étude. L'observation de cette figure montre des changements au niveau des unités d'occupation des terres. Ces changements se caractérisent par les pertes, les gains et de stabilité. Ainsi, les plus forts gains ont été observés au niveau des mosaïques de champs et jachères (55%). Les plus fortes pertes ont été observées respectivement au niveau des champs jachères (7 %), et des savanes saxicoles (5%). Les forêts denses (6 %) ; les plantations (15%) ; les forêts galeries (4 %) et les forêts claires ; savanes boisées (5 %) ; les agglomérations (10%) et les forets claires et savanes boisés (3%) sont restées plus ou moins stables.

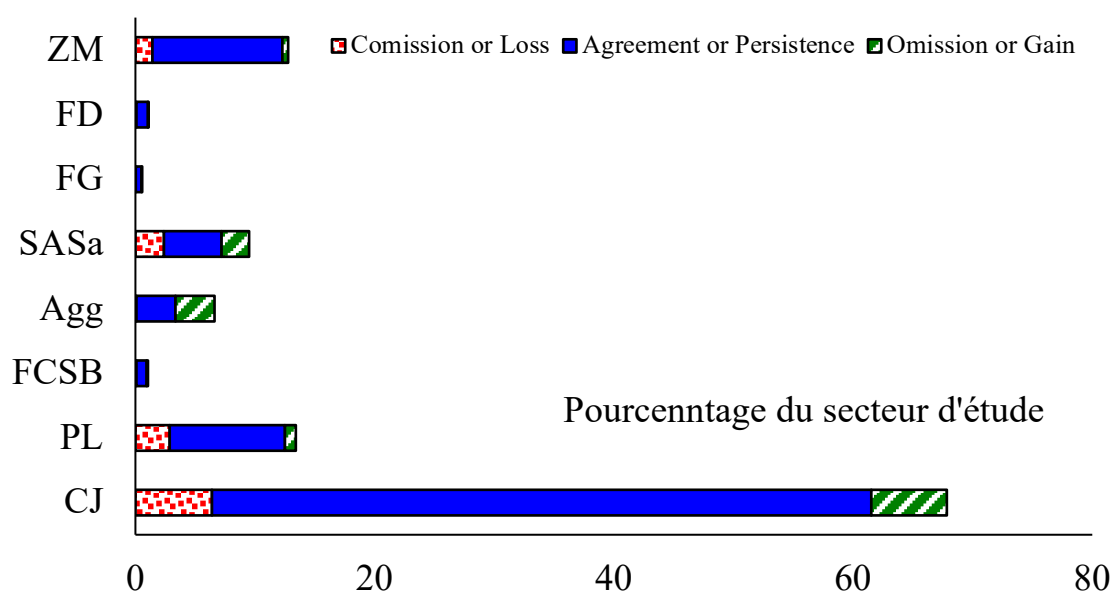


Figure 9 : Intensité de changement par catégorie d'occupation des terres entre 1990 et 2000

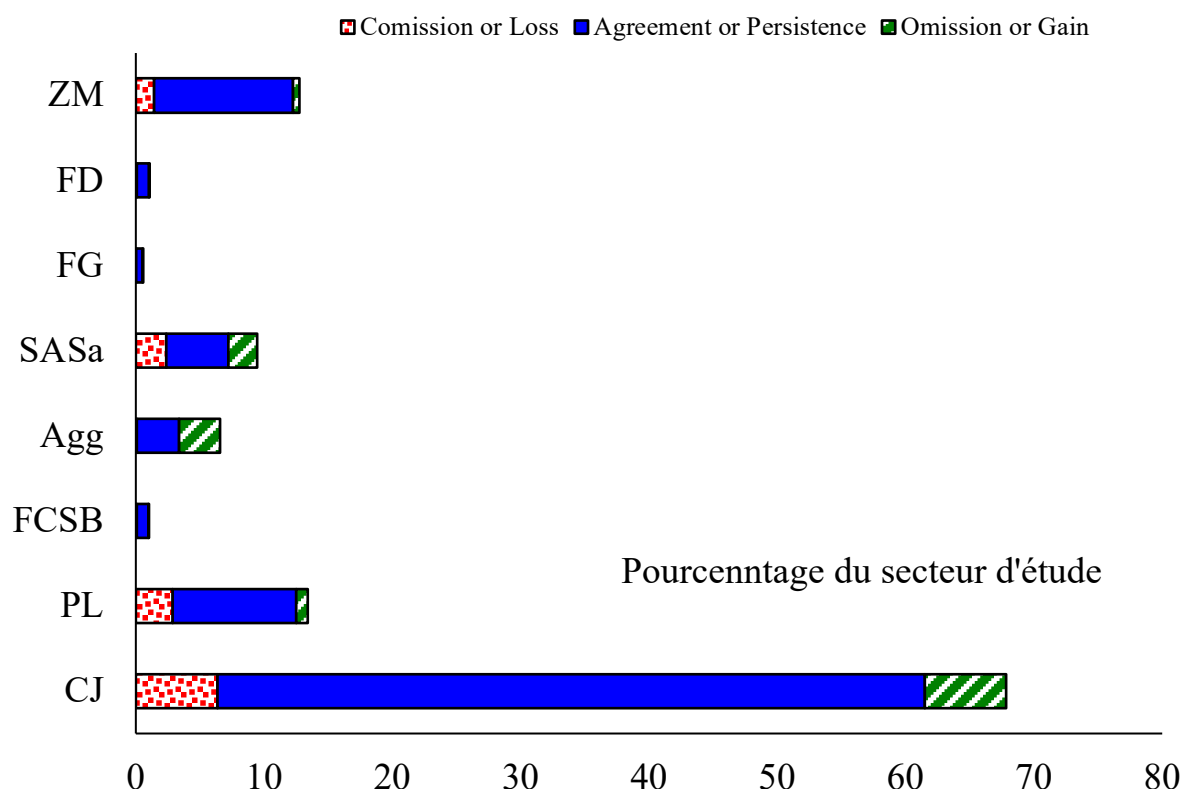


Figure 10 : Intensité de changement par catégorie d'occupation des terres entre 2000 et 2020

En outre, la diminution des zones humides peut réduire la capacité naturelle de rétention d'eau, laissant moins d'espace pour que l'eau soit absorbée et stockée. Les zones humides agissent souvent comme des zones tampons naturelles, aidant à réguler les débits d'eau et à prévenir les inondations (FAO 2021). L'expansion agricole peut également jouer un rôle, car les changements dans les pratiques agricoles, tels que le drainage excessif, peuvent influencer le ruissellement et contribuer aux inondations (Scarwell et Laganier 2017).

4.2.3. Occupation du sol projeté

Après avoir confirmé la précision des modèles pour la carte de 2020, la méthode de la chaîne de Markov a été utilisée pour prédire une carte future d'occupation pour l'année 2040. La prévision du modèle des changements des terres pour 2040 indique une régression des terres pour toutes les unités d'occupations sauf au niveau des champs jachère et agglomérations qui ont augmenté respectivement de 11% et 40% dans la zone d'étude (Annexe 1). Cela peut indiquer des changements dans l'évolution de l'urbanisation. Cette urbanisation entraîne l'abandon des champs ayant pour conséquence l'augmentation de la jachère (figure 11).

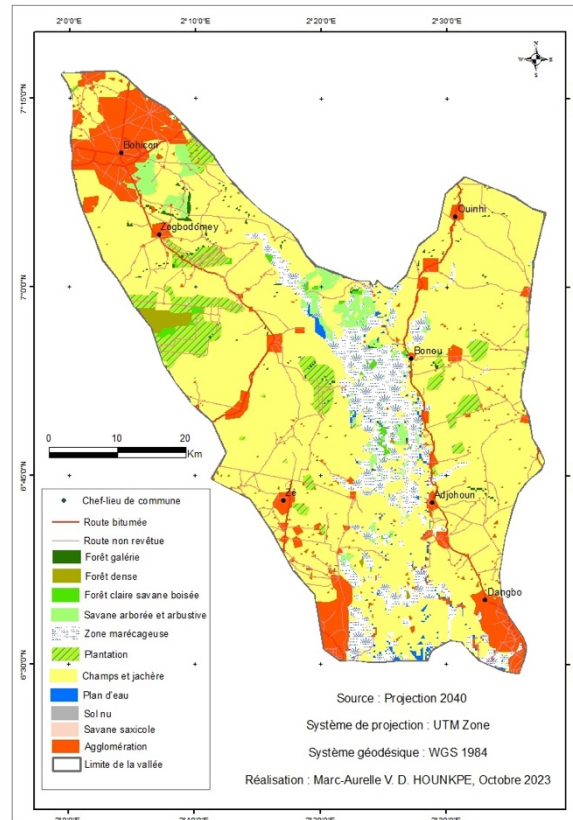


Figure 11 : Cartes d'occupation des terres/sols prédites pour l'année 2040

4.2.4. Changement climatique projeté

Les stations qui ont fourni des mesures météorologiques des 20 et 25 derniers années ont été utilisées pour déterminer l'effet que le changement climatique peut avoir sur les précipitations futures dans la zone d'étude. La figure 12 présente les régimes pluviométriques observés et générés des stations étudiées (RCP4.5). L'état moyen simulé pour les précipitations correspond à celui observé pour les stations de Allada, Toffo, Zagnanado et Bohicon. On observe une diminution des précipitations moyennes mensuels projeter aux stations de Bonou et Porto-Novo. Les modèles utilisés ont réussi à capturer ces variations avec précision. La figure 13 présente les régimes pluviométriques observés et générés des stations étudiées (RCP8.5). L'état moyen simulé pour les précipitations correspond à celui observé pour les stations de Bonou, Toffo et Zagnanado. On observe une diminution des précipitations moyennes mensuels projeter aux stations de Bohicon, Zagnanado et Porto-Novo. Ces résultats rentrent en concordance avec ceux de (Lambin et al. 2001) qui ont montré une tendance positive des précipitations dans la plupart des mois. Cela est aussi confirmé par (Yariyan et al. 2020) qui stipule que les altitudes inférieures et proche de la mer recevrons plus de précipitations contrairement aux altitudes supérieures.

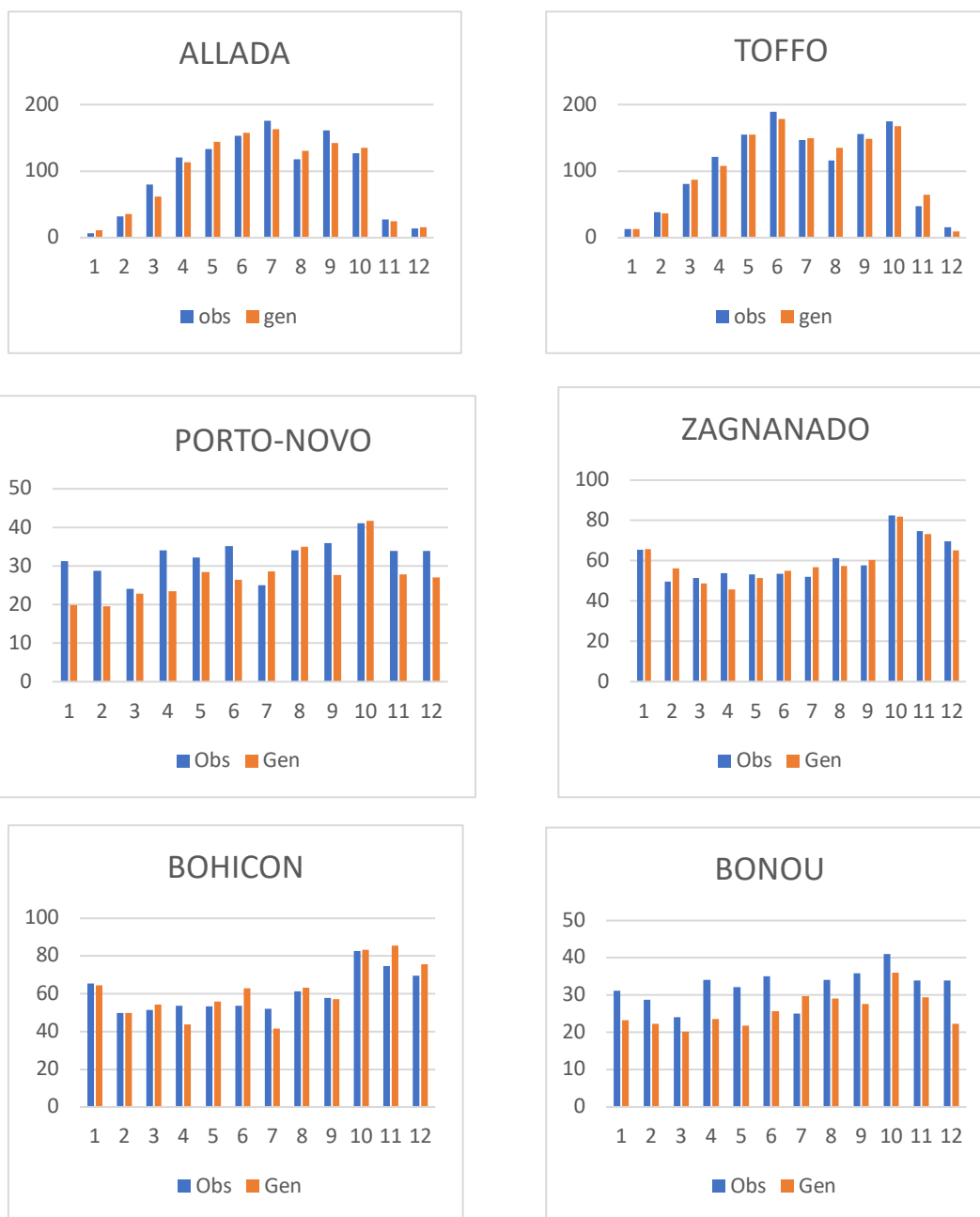


Figure 12 : Régimes pluviométriques observés et générés des stations étudiées (RCP4.5)

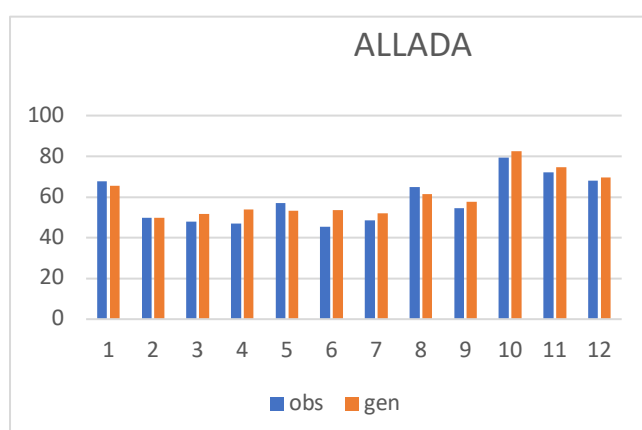
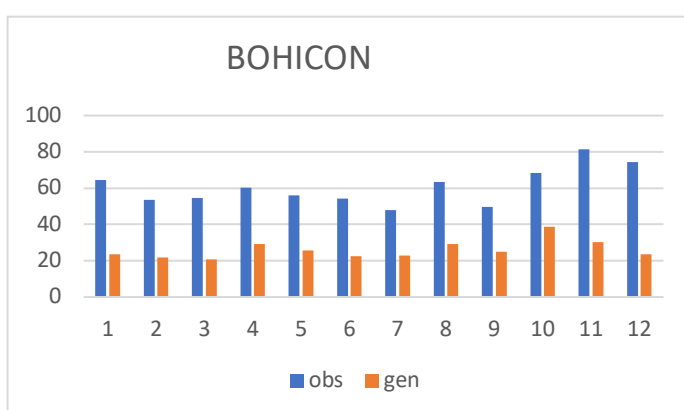
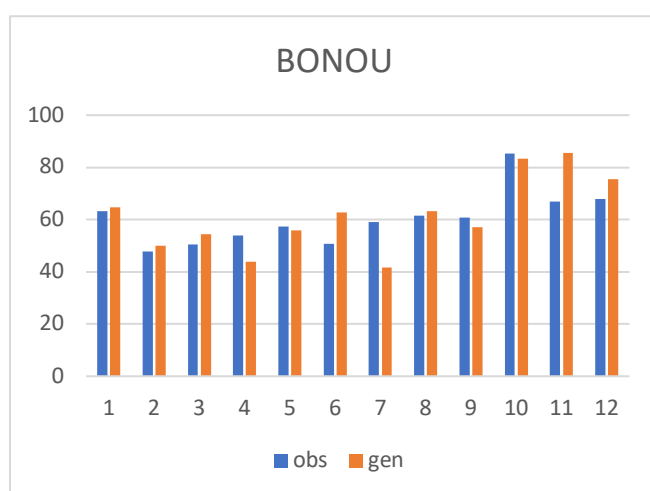
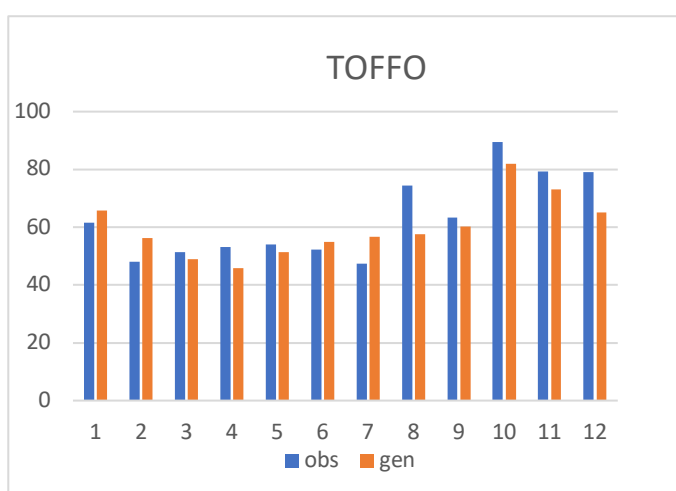
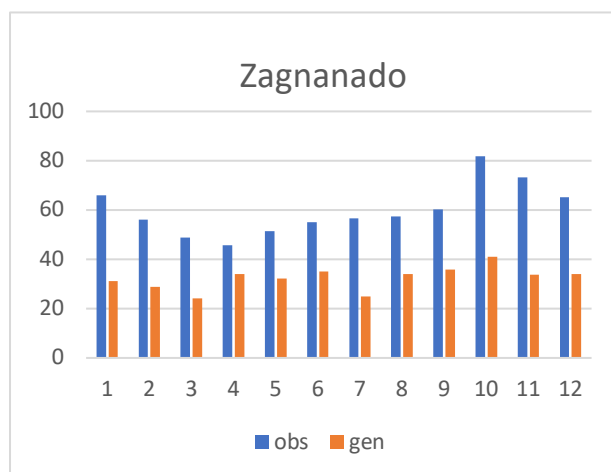
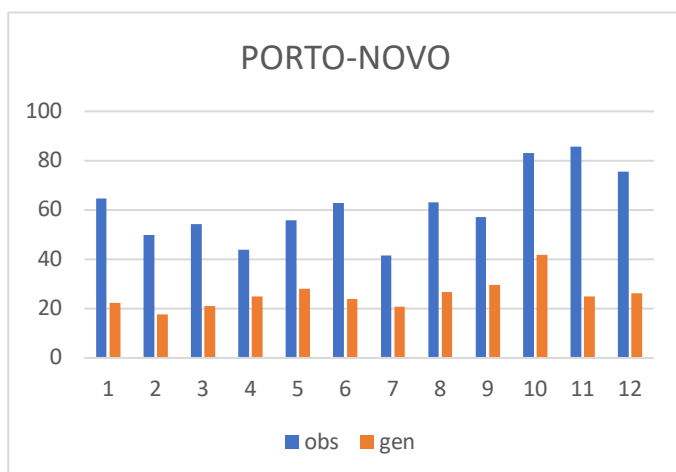


Figure 13: Régimes pluviométriques observés et générés des stations étudiées (RCP8.5)

4.3. Développement d'un modèle basé sur l'apprentissage automatique qui prédire la probabilité future des inondations

4.3.1. Cartes de probabilités d'inondation

La susceptibilité aux inondations de la zone d'étude a été cartographiée par deux modèles d'apprentissage automatique le Random-Forest et le SVM et par deux scénarios RCP4.5 et RCP8.5, puis en utilisant un algorithme de coupure naturelle dans un environnement logiciel SIG pour 4 classes comprenant le vert (tout va bien), le jaune (l'eau vient), l'orange (l'eau est là) et le rouge (danger) (Annexe 4). Il existe différents algorithmes dans un logiciel SIG pour classer les classes de cartes raster. Ils comprennent la rupture naturelle, l'intervalle égal, le quintile, l'intervalle régulier, l'écart type et la technique manuelle. L'algorithme de coupure naturelle est l'algorithme de classification de cartes le plus connu utilisé dans la plupart des études. La carte de susceptibilité aux inondations des modèles utilisés est présentée dans la figure 13. La figure 13 présente les cartes de probabilité des inondations obtenus pour les modèles SVM et Random-forest avec les scénarios RCP 4.5 et RCP 8.5. De cette figure il ressort que les zones d'inondation obtenus à l'aide des deux modèles (Random-Forest et de SVM) sont influencées par de différents scénarios climatiques (RCP4.5 et RCP8.5). Ces résultats montrent que les zones proches des stations marquées en rouge (BONOU, ZAGNANADO, OUIHI) de la zone d'étude ont été identifiées respectivement par les modèles RF et SVM comme zones en danger (rouge) pour le scénario RCP8.5, par contre dans le scénario RCP4.5, les zones proches de la station de ZAGNANADO sont beaucoup plus susceptibles aux inondations sur la carte. Aussi, sur les cartes de probabilité d'inondation, les zones de Bohicon, Zogbodomey, Zè et Dangbo sont à l'abri du phénomène, car tout est normal dans cette zone. De plus, pour chaque modèle (RF et SVM) les scénarios RCP8.5 ont détecté plus de superficie inondable contrairement aux scénarios RCP4.5, soit des superficies comprises entre 62667 et de 87328m² pour les scénarios RCP 8.5 contre des superficies comprises entre 1691 et 3202m² pour les scénarios RCP4.5.

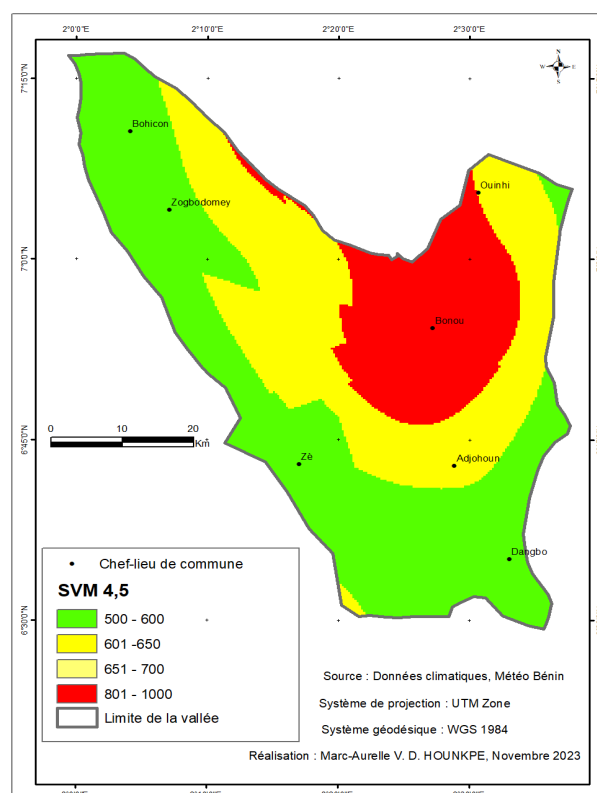
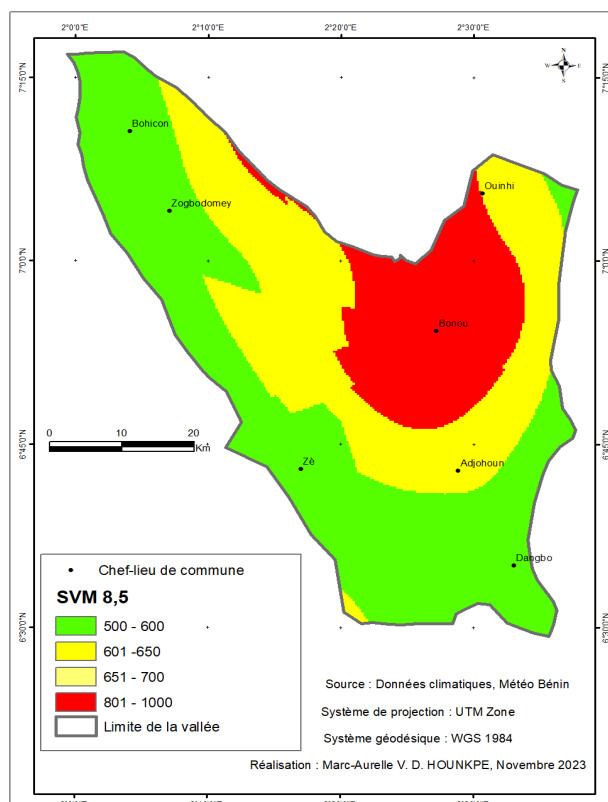
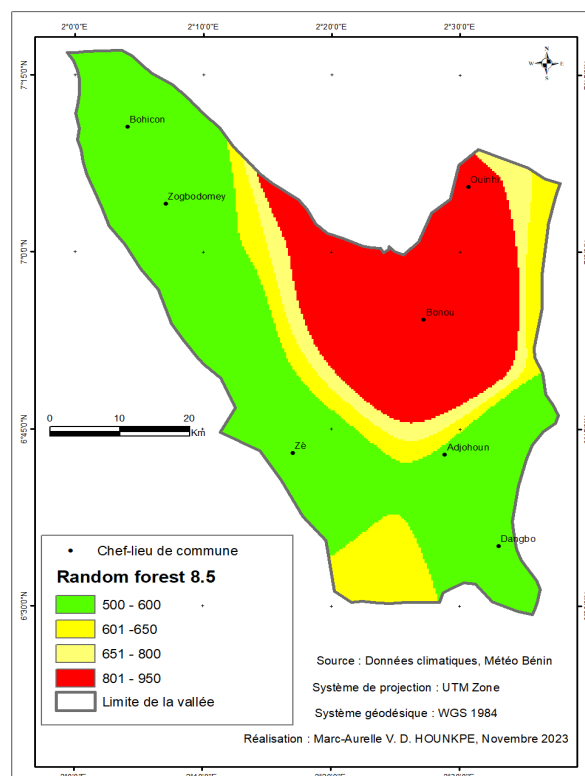
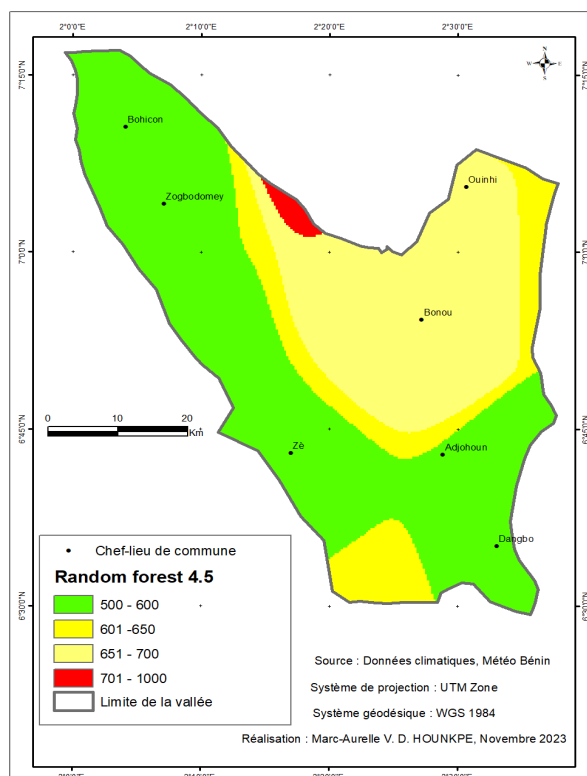


Figure 14 : Cartes de probabilités des inondations

4.3.2. Évaluation et comparaison du modèle

Le tableau 5 présente l'évaluation de la performance des modèles. Ces performances ont été validé par les prédictions des modèles qui ont été validées avec la courbe ROC et d'autres critères. Les résultats montrent que le modèle SVM qui présente Accuracy = 0,86, AUC- ROC= 0,88, F1-scores = 0.80 est un prédicteur plus précis des probabilités aux inondations que le modèle RF qui présente Accuracy = 0,85 ; AUC- ROC= 0,83, F1-scores = 0.79, (Tableau 5). Les modèles ont été exécutés à l'aide de l'ensemble de données de validation basé sur les données climatiques futur et sur les modèles climatiques pessimistes et optimistes.

Tableau 4: Évaluation de la performance des deux modèles

Critères de performances	SVM	RF
AUC-ROC	0.88	0.83
F1-Scores	0.80	0.79
Recall	0.85	0.82
Accuracy	0.86	0.85

Discussion

Nous discuterons dans cette partie de l'interprétation des images satellitaires multi-dates, la cartographie de la dynamique, et la modélisation de l'occupation du sol aux divers horizons ont été discutées.

Les deux (2) tableaux de matrice de confusion montrent pour chacune des classes, le niveau de fiabilité et les principales confusions faites lors de la classification des images. Pour l'ensemble des classes définies, il y a eu de confusion c'est-à-dire les pixels de certaines unités d'occupation des terres ont été confondus à d'autres. Mais avec l'indice de Kappa (un indice qui permet de « retirer » la portion de hasard ou de subjectivité de l'accord entre les techniques), 0.90, 0.93, et 0,98 que nous avons eu respectivement pour les trois images de 1990, 2000 et 2020, nous permettent de conclure que les résultats de ces classifications sont statistiquement acceptables. En effet ,(Landis K et Al 2020) cet indice est **Excellent** quand il est égal à 0.81; il est **Bon** quand il compris entre 0.80 et 0.61; il est **Modéré** quand il est compris entre 0.60 et 0.21; il est **Mauvais** quand il est compris entre 0.20 et 0.0 et il est **Très Mauvais** quand il est inférieur à 0.0.

Les cartes d'occupation des terres réalisées ont permis de mettre en évidence la tendance générale de la dynamique de l'occupation des terres. Cette tendance générale révèle une augmentation de la superficie des formations anthropiques (cultures et jachères, agglomérations, plantations) au détriment des formations végétales naturelles (forêts, savanes boisées). (Roose et al. 2012) et (Houndagba, Tente, et Guedou 2007) étaient déjà parvenus à la même conclusion concernant le bassin versant de la rivière Zou et l'étude de la dynamique des petits bassins lacustres du Sud-Bénin. Plus inquiétant, cette tendance de dégradation de la couverture végétale n'épargne même pas les terres dans le milieu d'étude, ce résultat peut être comparé à celui de (Morsli, Habi, et Meddi 2013) (Khalid et al. 2015) qui ont montré que la dégradation est le résultat des incendies et des défrichements qui se sont succédés tout au long de l'histoire. Suite à cette dégradation, le sol a subi à son tour une modification. Situé sur pente forte et occupé par une végétation clairsemée, le sol est très affecté par l'érosion en nappe. Le décapage du sol est si important que l'érosion a fini par affleurer la roche-mère gênante ainsi fortement l'infiltration des eaux.

Deux modèles d'apprentissage automatique, RF et SVM, ont été utilisés pour prédire la probabilité d'inondation dans la zone d'étude au sud-Benin. L'évaluation des performances des modèles montre que le SVM était le modèle le plus efficace. L'analyse de la modélisation a montré que le modèle

SVM, par rapport à certains modèles, présente une précision de prédiction plus élevée, ce qui confirme les résultats de cette étude. De plus, les résultats de classification des données futures des modèles RF et SVM sous deux scénarios climatiques RCP4.5 et RCP8.5 à l'aide des différentes métriques ont montré que la carte modèle SVM sous le scénario RCP4.5 avec $AUC-ROC = \pm 0,89$ pour presque tous les stations de la zone d'étude, par rapport à d'autres modèles et scénarios climatiques, est la plus précis pour l'élaboration des probabilités pour la détermination des cartes de susceptibilité aux inondations de la zone d'étude. Elle est suivie par la carte modèle RF RCP4.5 avec $AUC = 0,893$.

Conclusion

Dans ce dernier chapitre nous avons faire une analyse approfondie des données. Ensuite nous avons procédé à conception des modèles RF et SVM et l'entraînement et test de ces derniers. Enfin, le meilleur modèle retenu grâce aux métriques ayant donnés les meilleures performances sur les données de validations a permis de renforcer ainsi notre capacité à anticiper et à atténuer les risques d'inondation. Le modèle SVM est la plus précis pour la détermination des probabilités.

Conclusion General

En conclusion, cette étude exhaustive des facteurs influençant les inondations a permis de cerner les éléments clés contribuant à la susceptibilité aux inondations. L'analyse minutieuse des données, intégrant des paramètres variés tels que la topographie, la couverture terrestre, la pluviométrie, et l'occupation des terres, a offert une compréhension approfondie des dynamiques en jeu.

L'application de techniques d'intelligence artificielle, notamment le RF et le SVM, combinées à des modèles de prévision climatique, a considérablement amélioré la capacité de prévision des inondations. Ces approches ont démontré une précision et une fiabilité remarquables, ouvrant la voie à des prévisions plus précises des risques futurs. L'analyse des changements d'occupation des terres au fil des décennies, conjuguée à des projections basées sur la chaîne de Markov, a permis de visualiser les évolutions probables jusqu'en 2040. Ces résultats offrent des perspectives cruciales pour la planification urbaine et la gestion des zones à risque. En parallèle, l'examen des modèles climatiques, notamment sous les scénarios RCP4.5 et RCP8.5, a permis de cartographier les tendances des prélèvements, renforçant la capacité à anticiper les épisodes pluvieux extrêmes. Enfin, l'évaluation des modèles, à travers des critères rigoureux comme la courbe ROC, confirme la solidité des prédictions, démontrant l'efficacité, notamment, du modèle Random-Forest et svm.

Dans son ensemble, cette étude offre une approche holistique pour la gestion des risques d'inondation, intégrant des données historiques, des modèles climatiques, et des méthodes d'intelligence artificielle, fournissant ainsi des outils précieux pour guider les décisions et les actions visant à minimiser l'impact des inondations dans l'avenir. Nos résultats peuvent servir d'outils d'améliorations du système SAP d'alerte précoce de la DGEau.

REFERENCES BIBLIOGRAPHIQUES

- Alison Freebairn et al. 2020. : « : Alison Freebairn et al ». 2020. <https://go.ifrc.org/emergencies/5670/details>.
- « A review of the artificial intelligence methods in groundwater level modeling ». 2019. *Journal of Hydrology* 572 (mai): 336-51. <https://doi.org/10.1016/J.JHYDROL.2018.12.037>.
- Adam et Boko. 1993. « Le Bénin et son réseau hydrographique (Adam & Boko, 1993). | Télécharger le diagramme scientifique ». 1993. https://www.researchgate.net/figure/Le-Benin-et-son-reseau-hydrographique-Adam-Boko-1993_fig3_334828680.
- Ahmed, Fahad, Ho Huu Loc, Edward Park, Muhammad Hassan, et Panuwat Joyklad. 2022. « Comparison of Different Artificial Intelligence Techniques to Predict Floods in Jhelum River, Pakistan ». *Water* 14 (21): 3533. <https://doi.org/10.3390/w14213533>.
- Alo, Clement Aga, et Robert Gilmore Pontius. 2008. « Identifying Systematic Land-Cover Transitions Using Remote Sensing and GIS: The Fate of Forests inside and Outside Protected Areas of Southwestern Ghana ». *Environment and Planning B: Planning and Design* 35 (2): 280-95. <https://doi.org/10.1068/b32091>.
- Amoussou, Ernest. 2014. « Analyse hydrométéorologique des crues dans le bassin-versant du Mono en Afrique de l'Ouest avec un modèle conceptuel pluie-débit ». <https://shs.hal.science/halshs-01143318>.
- Arabameri, Alireza, Sunil Saha, Wei Chen, Jagabandhu Roy, Biswajeet Pradhan, et Dieu Tien Bui. 2020. « Flash flood susceptibility modelling using functional tree and hybrid ensemble techniques ». *Journal of Hydrology* 587 (août): 125007. <https://doi.org/10.1016/j.jhydrol.2020.125007>.
- Awad, Mariette, et Rahul Khanna. 2015. « Support Vector Machines for Classification ». In *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*, édité par Mariette Awad et Rahul Khanna, 39-66. Berkeley, CA: Apress. https://doi.org/10.1007/978-1-4302-5990-9_3.
- Benaichata, Lazreg. 2016. « Génération Stochastique des données climatiques pour les modèles de prévision d'incendies de forêts ».
- Blaschke, Thomas. 2010. « Object based image analysis for remote sensing. ISPRS J Photogramm Remote Sens ». *ISPRS Journal of Photogrammetry and Remote Sensing* 65 (janvier): 2-16.

- <https://doi.org/10.1016/j.isprsjprs.2009.06.004>.
- Breiman, Leo. 2001. « [No title found] ». *Machine Learning* 45 (1): 5-32. <https://doi.org/10.1023/A:1010933404324>.
 - Centre, UNESCO World Heritage. 2016. « Basse vallée de l'Ouémé - Centre du patrimoine mondial de l'UNESCO ». UNESCO World Heritage Centre. 2016. <https://whc.unesco.org/en/tentativelists/6481/>.
 - Chan, Keung Sum, et Yee Shan Chan. 2021. « Exploring Hong Kong Nursing Students' Experience of Using Smartphones in Clinical Practicum ». *Journal of Biosciences and Medicines* 9 (6): 194-207. <https://doi.org/10.4236/jbm.2021.96018>.
 - Chang, Ching-Yi, Chiu-Lin Lai, et Gwo-Jen Hwang. 2018. « Trends and Research Issues of Mobile Learning Studies in Nursing Education: A Review of Academic Publications from 1971 to 2016 ». *Computers & Education* 116 (janvier): 28-48. <https://doi.org/10.1016/j.compedu.2017.09.001>.
 - Chloé et Al. 2019a. « Évaluez les performances d'un modèle de machine learning ». OpenClassrooms. 2019. <https://openclassrooms.com/fr/courses/4297211-evaluez-les-performances-dun-modele-de-machine-learning>.
 - ———. 2019b. « Explication de la régression et de la classification dans l'apprentissage automatique - Geekflare ». 2019. <https://geekflare.com/fr/regression-vs-classification/>.
 - Dodangeh, Esmaeel, Mahdi Panahi, Fatemeh Rezaie, Saro Lee, Dieu Tien Bui, Chang-Wook Lee, et Biswajeet Pradhan. 2020. « Novel hybrid intelligence models for flood-susceptibility prediction: Meta optimization of the GMDH and SVR models with the genetic algorithm and harmony search ». *Journal of Hydrology* 590 (novembre): 125423. <https://doi.org/10.1016/j.jhydrol.2020.125423>.
 - FAO. 2021. « Les zones humides et l'eau | Zones Humides ». 2021. <https://www.zones-humides.org/les-zones-humides-et-l-eau>.
 - Foody, G. M., et A. Mathur. 2004. « Toward intelligent training of supervised image classifications: directing training data acquisition for SVM classification ». *Remote Sensing of Environment* 93 (1-2): 107-17. <https://doi.org/10.1016/j.rse.2004.06.017>.
 - Fujiki, Kenji, et Olivier Finance. 2022. « Exposition et vulnérabilité sociale des villes françaises au risque inondation : une analyse spatiotemporelle à fine échelle (1999-2017) ». *Cybergeo: European Journal of Geography*,

- juin. <https://doi.org/10.4000/cybergegeo.39179>.
- Guha-Sapir et al, Guha-Sapir et al. 2017. « Annual reports ». 2017. <https://www.emdat.be/categories/adsr/>.
 - Houndagba, Cossi Jean, Agossou Brice Hugues Tente, et Raoul Guedou. 2007. « Dynamique des forêts classées dans le cours moyen de l'Ouémé au Bénin ». In *Quelles aires protégées pour l'Afrique de l'Ouest ?*, édité par Anne Fournier, Brice Sinsin, et Guy Apollinaire Mensah, 369-80. IRD Éditions. <https://doi.org/10.4000/books.irdeditions.8069>.
 - HUANG et Al. 2002. « (PDF) Une évaluation des machines à vecteurs de support pour la classification de la couverture terrestre ». 2002. https://www.researchgate.net/publication/248591610_An_assessment_of_support_vector_machines_for_land_cover_classification.
 - Ibrahima et Al, Dr Sylla. 2023. « Eléments de correction Carto-SIG (1ère Session – 2022/2023) ». *Sylla Academie* (blog). 17 juin 2023. <https://syllaacademie.com/correction-carto-sig-1ere-session-2022-2023/>.
 - « IFRC GO - Emergency ». s. d. Consulté le 6 novembre 2023. <https://go.ifrc.org/emergencies/5670/details>.
 - « Induction d'arbres de décision | SpringerLink ». s. d. Consulté le 6 novembre 2023. <https://link.springer.com/article/10.1007/BF00116251>.
 - Jenifa, Ashlin. 2022. « Explication de la régression par rapport à la classification dans l'apprentissage automatique ». Geekflare. 19 décembre 2022. <https://geekflare.com/fr/regression-vs-classification/>.
 - Kettig, R. L., et D. A. Landgrebe. 1976. « Classification of Multispectral Image Data by Extraction and Classification of Homogeneous Objects ». *IEEE Transactions on Geoscience Electronics* 14 (1): 19-26. <https://doi.org/10.1109/TGE.1976.294460>.
 - Khalid, Fariza, Mazalah Ahmad, Aidah Abdul Karim, Md. Yusoff Daud, et Rosseni Din. 2015. « Reflective Thinking: An Analysis of Students' Reflections in Their Learning about Computers in Education ». *Creative Education* 06 (20): 2160-68. <https://doi.org/10.4236/ce.2015.620220>.
 - « Laleye, P. (1995). Ecologie comparée de deux espèces de Chrysichthys, poissons siluriformes (Claroteidae) du complexe lagunaire lac Nokoué-lagune de Porto-Novo au Bénin (152 p. + Annexes). Thèse de doctorat. Belgique Université de Liège. - Références - Publication de recherche scientifique ». s. d. Consulté le 6 novembre 2023. [https://www.scirp.org/\(S\(lz5mqp453edsnp55rrgjct55.\)\)/reference/referencespapers.aspx?referenceid=2955616](https://www.scirp.org/(S(lz5mqp453edsnp55rrgjct55.))/reference/referencespapers.aspx?referenceid=2955616).

- Lambin, E. F., B. L. Turner, H. J. Geist, S. B. Agbola, A. Angelsen, J. W. Bruce, O. T. Coomes, et al. 2001. « The Causes of Land-Use and Land - Cover Change : Moving beyond the Myths ». *Global Environmental Change* 11 (4): 261-69. [https://doi.org/10.1016/S0959-3780\(01\)00007-3](https://doi.org/10.1016/S0959-3780(01)00007-3).
- Landis K et Al. 2020. « Landis and Koch interpretation of Kappa Cohen scores Download Table ». 2020. https://www.researchgate.net/figure/Landis-and-Koch-interpretation-of-Kappa-Cohen-scores_tbl3_267416171.
- Lecerf et Al, Rémi. 2008. « Suivi des changements d'occupation et d'utilisation des sols d'origine anthropique et climatique à l'échelle régionale par télédétection moyenne résolution (application à la Bretagne) ».
- MacQueen, J. 1967. « Some methods for classification and analysis of multivariate observations ». In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Statistics*, 5.1:281-98. University of California Press. <https://projecteuclid.org/ebooks/berkeley-symposium-on-mathematical-statistics-and-probability/Proceedings-of-the-Fifth-Berkeley-Symposium-on-Mathematical-Statistics-and/chapter/Some-methods-for-classification-and-analysis-of-multivariate-observations/bsmsp/1200512992>.
- Martin P et AL. 2004. « Logiciel SIG : Tout savoir sur les Systèmes d'Information Géographique ». 2004. <https://www.esrifrance.fr/produits/Logiciel-SIG.aspx>.
- Mitchell et Al, Tom M. 1997. « Does Machine Learning Really Work? » *AI Magazine* 18 (3): 11-11. <https://doi.org/10.1609/aimag.v18i3.1303>.
- Mohamed EL et al. 2009. « Memoire Online - Les changements climatiques - Mohamed El Moctar Agweidess ALATY OULD ». Memoire Online. 2009. <https://www.memoireonline.com/10/13/7594/Les-changements-climatiques.html>.
- Morgan, James N., et John A. Sonquist. 1963. « Problems in the Analysis of Survey Data, and a Proposal ». *Journal of the American Statistical Association* 58 (302): 415-34. <https://doi.org/10.1080/01621459.1963.10500855>.
- Morsli, Boutkhil, Mohammed Habi, et Mohamed Meddi. 2013. « Dynamique de l'érosion en zone méditerranéenne algérienne : facteurs explicatifs de variation du ruissellement et de l'érosion sous différentes occupations du sol ». *Revue des sciences de l'eau / Journal of Water*

Science 26 (2): 89-105. <https://doi.org/10.7202/1016061ar>.

- Myers, Wayne L., et Ganapati P. Patil, éd. 2006. « Comparative Change and Pattern Perturbation ». In *Pattern-Based Compression of Multi-Band Image Data for Landscape Analysis*, 79-97. Sur Les Statistiques Environnementales et Écologiques. Boston, MA: Springer US. https://doi.org/10.1007/978-0-387-44439-0_5.
- Ouedraogo, Par Issouf. 2023. « AUTOMATISATION DU PRE-TRAITEMENT DES DONNEES PAR L'OPTIMISATION METAHEURISTIQUE ».
- Péliissier, Paul. 1962. « Les pays du Bas-Ouémé (premier article) ». *Les Cahiers d'Outre-Mer* 15 (59): 204-54. <https://doi.org/10.3406/caoum.1962.2254>.
- Roose, Éric, Mohamed Sabir, Mourad Arabi, Boutkhil Morsli, et Mohamed Mazour. 2012. « Soixante années de recherches en coopération sur l'érosion hydrique et la lutte antiérosive au Maghreb ». *Physio-Géo. Géographie physique et environnement*, n° Volume 6 (février): 43-69. <https://doi.org/10.4000/physio-geo.2319>.
- Roy, Paramita, Subodh Chandra Pal, Rabin Chakraborty, Indrajit Chowdhuri, Sadhan Malik, et Biswajit Das. 2020. « Threats of climate and land use change on future flood susceptibility ». *Journal of Cleaner Production* 272 (novembre): 122757. <https://doi.org/10.1016/j.jclepro.2020.122757>.
- Sahana, Meheub, et Priyank Patel. 2019. « A comparison of frequency ratio and fuzzy logic models for flood susceptibility assessment of the lower Kosi River Basin in India ». *Environmental Earth Sciences* 78 (mai). <https://doi.org/10.1007/s12665-019-8285-1>.
- Scarwell, Helga-Jane, et Richard Laganier. 2017. « Chapitre 2 : L'intégration matérielle du risque par le biais des mesures structurelles ». In *Risque d'inondation et aménagement durable des territoires*, 117-62. Environnement et société. Villeneuve d'Ascq: Presses universitaires du Septentrion. <https://doi.org/10.4000/books.septentrion.15687>.
- Seidou, Sophiatou, Pierre Ouassa, Gervais A. A. Atchade, et Expédit W. Vissin. 2021. « Caractérisation Des Risques Hydroclimatiques Dans La Basse Vallée De L'Oueme Au Benin (Afrique De L'Ouest) ». *International Journal of Progressive Sciences and Technologies* 25 (2): 334-48. <https://doi.org/10.52155/ijpsat.v25.2.2716>.
- Semenov, Ma, Rj Brooks, Em Barrow, et Cw Richardson. 1998.

- « Comparison of the WGEN and LARS-WG Stochastic Weather Generators for Diverse Climates ». *Climate Research* 10: 95-107. <https://doi.org/10.3354/cr010095>.
- Semenov, Mikhail, et Elaine Barrow. 2002. « LARS-WG A Stochastic Weather Generator for Use in Climate Impact Studies », janvier.
 - « Série Télédétection pour décideurs N° 21 ». 2010. 2010. <https://www.fao.org/3/Y3642F/Y3642F00.htm>.
 - Sossou-Agbo, Anani Lazare. 2013. « La mobilité dans le complexe fluvio-lagunaire de la basse vallée de l'Ouémé au Bénin, en Afrique de l'Ouest ». Phdthesis, Université de Grenoble ; Université d'Abomey-Calavi (Bénin). <https://theses.hal.science/tel-00995697>.
 - Tehrany, Mahyat, Simon Jones, et Farzin Shabani. 2018. « Identifying the essential flood conditioning factors for flood prone area mapping using machine learning techniques (Acceptance date 10/12/2018) ». *Catena* 175 (décembre). <https://doi.org/10.1016/j.catena.2018.12.011>.
 - Thanh N et al. 2017. « Sensors | Free Full-Text | Comparison of Random Forest, k-Nearest Neighbor, and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery ». 2017. <https://www.mdpi.com/1424-8220/18/1/18>.
 - Thierry, Agbanou, Paegelow Martin, Toko Imorou Ismaïla, et Tente Brice. 2018. « Modelisation Des Changements D'occupation Des Terres En Region Soudanienne Au Nord-Ouest Du Benin ». *European Scientific Journal, ESJ* 14 (12): 248. <https://doi.org/10.19044/esj.2018.v14n12p248>.
 - Tien Bui, Dieu, Nhat-Duc Hoang, Tien-Dat Pham, Phuong-Thao Thi Ngo, Pham Viet Hoa, Nguyen Quang Minh, Xuan-Truong Tran, et Pijush Samui. 2019. « A new intelligence approach based on GIS-based Multivariate Adaptive Regression Splines and metaheuristic optimization for predicting flash flood susceptible areas at high-frequency tropical typhoon area ». *Journal of Hydrology* 575 (août): 314-26. <https://doi.org/10.1016/j.jhydrol.2019.05.046>.
 - Yariyan, Peyman, Mohammadtaghi Avand, Rahim Ali Abbaspour, Ali Torabi Haghighi, Romulus Costache, Omid Ghorbanzadeh, Saeid Janizadeh, et Thomas Blaschke. 2020. « Flood susceptibility mapping using an improved analytic network process with statistical models ». *Geomatics, Natural Hazards and Risk* 11 (1): 2282-2314. <https://doi.org/10.1080/19475705.2020.1836036>.
 - Zzaman, Rashed, Sara Nowreen, Maruf Billah, et Akm Islam. 2021.

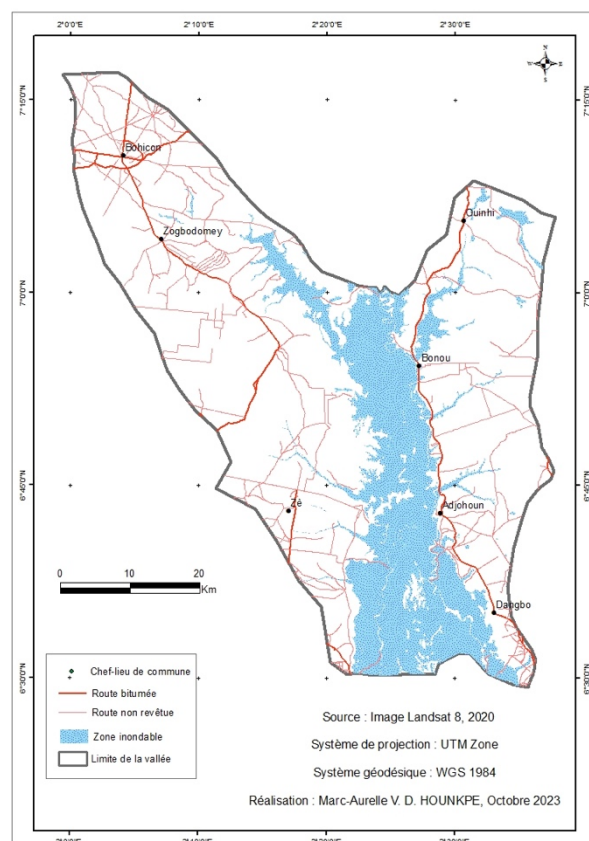
« Flood hazard mapping of Sangu River basin in Bangladesh using multi-criteria analysis of hydro-geomorphological factors ». *Journal of Flood Risk Management* 14 (mars). <https://doi.org/10.1111/jfr3.12715>.

ANNEXES

Annexe 1 : Superficies et taux d'évolution des classes d'occupations entre 1 2020 et 2040

	Superficie 2020	Superficie 2040	Taux evolution 2020-2040 (%)
Champ jachère	196544,22	219899,71	11
Forêt claire savane boisée	3038,24	1544,81	-97
Forêt dense	3134,64	1970,55	-59
Forêt galerie	1651,53	1396,49	-18
Zone marécageuse	36148,84	33012,32	-10
Agglomération	20875,86	34748,39	40
Plan d'eau	2125,33	1865,92	-14
Plantation	33599,29	16626,82	-102
Savane arborée et arbustive	22682,88	8777,92	-158
Sol nu	10,86	5,37	-102
Savane saxicole	77,99	41,4	-88
Total_2040	319889,7	319889,7	0
indice Kappa	0.85	0.85	

Annexe 2: Zone inondable



Annexe 3: Valeurs seuils utilisées pour chaque classe

Valeurs seuils de classifications(débit)	Observation	Couleur
550	Situation normale (tout va bien)	Vert
550- 700	L'eau vient	Jaune
700- 800	L'eau est la	Orange
>800	Danger	Rouge

Table des matières

DEDICACE	3
REMERCIEMENTS.....	4
LISTE DES TABLEAUX.....	6
LISTE DES FIGURES	6
RESUME	8
ABSTRACT.....	9
Introduction Générale.....	10
A. Contexte et Justification	10
B. Objectifs de la recherche	12
C. Organisation du mémoire.....	12
Chapitre 1 : Revue de littérature.....	13
Introduction	13
1.1. Clarification des concepts	13
1.1.1. Changement climatique	13
1.1.2. SIG et Télédétection	13
1.1.3. Modeleur de changements des terres	15
1.1.4. Occupations des sols / utilisations des sols	16
1.1.4.1. L'occupation du sol.....	16
1.1.4.2. L'utilisation du sol	16
1.1.5. Apprentissage Automatique	17
1.1.5.1. L'algorithme d'apprentissage supervisé	18
1.1.5.1.1. Méthode de classification.....	18
1.1.5.2. Méthode de régression	21
1.2. Points des connaissances.....	21
1.2.1. Techniques d'intelligence artificielle pour estimer la probabilité des inondations.....	21

1.2.2. Prédiction des inondations avec les données de SIG et d'occupation du sol.....	25
<i>Conclusion</i>	25
<i>Chapitre 2 : Zone d'étude et données</i>	27
<i>Introduction</i>	27
2.1. Description de la zone d'étude	27
2.1.1. Localisation de la zone d'étude	27
2.1.2. Le climat et l'hydrographie	28
2.1.3. Le relief.....	30
2.1.4. Phénomène d'inondation dans la basse vallée de l'Ouémé.....	30
2.2. Données d'étude	31
2.2.1. Données planimétriques	31
2.2.2. Données météorologiques.....	31
2.2.3. Données de débits journaliers.....	31
<i>Conclusion</i>	32
<i>Chapitre 3 : Méthodologie</i>	33
<i>Introduction</i>	33
3.1. Détermination des paramètres qui influent sur l'inondation dans la basse vallée de l'Ouémé	33
3.2. Analyse de l'évolution des précipitations et de l'occupation de sols dans la basse vallée de l'Ouémé	34
3.2.1. Méthode modeleur des changements des terres	34
3.2.1.1. Dynamique de l'occupation des terres.....	36
3.2.2. Utilisation du logiciel Lars-WG pour la prédiction des changements climatiques.....	38
3.3. Développement d'un modèle basé sur l'apprentissage automatique qui prédire la probabilité future des inondations	41
<i>Conclusion</i>	43
<i>Chapitre 4 : Résultats et discussion</i>	44
<i>Introduction</i>	44
4.1. Détermination des paramètres qui influent sur l'inondation dans la basse vallée de l'Ouémé	44

4.2. Analyse de l'évolution des précipitations et de l'occupation de sols dans la basse vallée de l'Ouémé.....	48
4.2.1. Changements d'occupation des terres.....	48
4.2.2. Intensité des changements par catégorie d'occupation des terres entre 1990-2000 et 2000-2020.....	51
4.2.3. Occupation du sol projeté.....	53
4.2.4. Changement climatique projeté.....	54
4.3. Développement d'un modèle basé sur l'apprentissage automatique qui prédire la probabilité future des inondations	57
4.3.1. Cartes de probabilités d'inondation.....	57
4.3.2. Évaluation et comparaison du modèle.....	59
<i>Discussion</i>	60
<i>Conclusion</i>	61
<i>Conclusion General</i>.....	62
<i>REFERENCES BIBLIOGRAPHIQUES</i>.....	63
<i>ANNEXES</i>	70

