

REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE SAAD DAHLAB – BLIDA 1

Faculté de Technologie

Département des Sciences de l'Eau et Environnement



MEMOIRE DE MASTER

Filière : Hydraulique

Spécialité : Ressources Hydrauliques

**La cartographie de la susceptibilité à l'érosion hydrique par
l'intégration de données géospatiales et d'apprentissage
automatique : cas du bassin versant Isser, Algérie du nord**

Présenté par :

ABDERRAHMANE Youcef Abdelaziz

Devant le jury composé de :

Mme. BENKACI Souhila	MCB	USD Blida 1	Présidente
M. BOUZERIA Housseyn	MCA	USD Blida 1	Promoteur
M. KHELFI Mohamed El Amine	MCB	USD Blida 1	Examineur

Promotion 2023/2024

ملخص. في شمال الجزائر، تزداد مشكلة التعرية المائية بشكل خطير، مما يهدد الأراضي الزراعية والمراعي. لفهم هذه المشكلة وإدارتها بفعالية، من الضروري رسم خرائط للتعرية وتحديد المناطق المعرضة للخطر. تتناول هذه الدراسة لتوقع حدوث التعرية المائية في حوض وادي يسر. تم تحليل عوامل متنوعة مناخية مقارنة أربع تقنيات للتعلم الآلي ومورفولوجية وهيدرولوجية وجيولوجية باستخدام نظم المعلومات الجغرافية يدعم صور الأقمار الصناعية تشمل البيانات مؤشرات مثل قابلية التآكل للأمطار، الانحدار، الاتجاه، الارتفاع، استخدام الأراضي، مؤشر الرطوبة الطبوغرافي، المسافة أو البعد عن المجاري المائية والطرق، عامل قابلية التآكل وكذلك التصنيف الجيولوجي. ساهمت الخرجات الميدانية وصور جوية في تحديد مواقع التآكل. تم تقسيم مجموعة البيانات إلى 60% للتدريب والمعايرة، و40% للاختبار. تم تقييم أداء النماذج باستخدام مقاييس إحصائية مثل الحساسية، والتحديد، والدقة، ومنحنى، مع دقة تفوق 94%، وحساسية تبلغ 0.96، وتحديد قدر 0.93. يمكن أن تساهم الخرائط التنبؤية لقابلية التعرية في تعزيز إدارة مستدامة للمياه وحفظ التربة، وبالتالي حماية الأنشطة الزراعية.

الكلمات المفتاحية. التعرية المائية، التعلم الآلي، رسم خرائط نظم المعلومات الجغرافية، حوض وادي يسر، حفظ التربة.

Abstract. In northern Algeria, water erosion is worsening, threatening agricultural lands and pastures. To better understand and manage this problem, it is crucial to map erosion and identify high-risk areas. This research compares four machine learning techniques (XGBoost Classifier, Random Forest, CatBoost Classifier, and LightGBM) to predict erosion in the Isser watershed. Various climatic, morphological, hydrological, and geological factors were analyzed using a GIS, integrating Landsat 8 OLI/TIRS images. The data include rainfall erosivity (MFI), slope, aspect, altitude, land use/Land cover, topographic wetness index (TWI), distance to watercourses and roads, erodibility factor (K), lithology, and geology. Field surveys and Google Earth satellite images helped locate erosion sites. The dataset was divided into 60% for training and calibration and 40% for testing. Model performance was evaluated using statistical measures such as sensitivity, specificity, accuracy, and the AUC-ROC curve, showing an accuracy above 94%, a sensitivity of 0.96, and a specificity of 0.93. Predictive susceptibility maps can help policymakers promote sustainable water management and soil conservation, thus protecting agricultural activities.

Keywords. Water erosion, machine learning, GIS mapping, Isser watershed, soil conservation.

Résumé. Dans le nord de l'Algérie, l'érosion hydrique s'aggrave, menaçant les terres agricoles et les pâturages. Pour mieux comprendre et gérer ce problème, il est crucial de cartographier l'érosion et d'identifier les zones à risque. Cette recherche compare quatre techniques d'apprentissage automatique (XGBoost Classifier, Random Forest, CatBoost Classifier et LightGBM) pour prédire l'érosion dans le bassin versant de l'Isser. Divers facteurs climatiques, morphologiques, hydrologiques et géologiques ont été analysés à l'aide d'un SIG, intégrant des images Landsat 8 OLI/TIRS. Les données incluent l'érosivité des précipitations (MFI), la pente,

l'orientation, l'altitude, l'occupation des terres, l'indice topographique d'humidité (TWI), la distance aux cours d'eau et aux routes, le facteur d'érodibilité (K), la lithologie et la géologie. Des enquêtes de terrain et des images satellite de Google Earth ont aidé à localiser les sites d'érosion. Le jeu de données a été divisé en 60 % pour l'entraînement et la calibration, et 40 % pour les tests. Les performances des modèles ont été évaluées par des mesures statistiques comme la sensibilité, la spécificité, la précision et la courbe AUC-ROC, montrant une précision supérieure à 94 %, une sensibilité de 0,96 et une spécificité de 0,93. Les cartes prédictives de susceptibilité à l'érosion peuvent aider les décideurs à promouvoir une gestion durable de l'eau et la conservation des sols, protégeant ainsi les activités agricoles.

Mots clés. Érosion hydrique, apprentissage automatique, cartographie SIG, bassin versant Isser, conservation des sols.

DEDICACE...

*Je dédie cet humble et modeste travail avec grand amour,
sincérité et fierté :*

*A ma famille, elle qui m'a doté d'une éducation digne, son amour
a fait de moi ce que je suis aujourd'hui,
Particulièrement à ma mère et mon père, aucune dédicace ne
saurait exprimer mon respect, mon amour éternel et ma
considération pour les sacrifices que vous avez consenti pour mon
instruction et mon bien être.*

A mes sœurs et mon frère, qui ont toujours été là pour moi.

A mes proches.

A mes chères amis et collègues.

A tous ceux que j'aime et ceux qui m'aiment.

ABDERRAHMANE Y.A

REMERCIEMENT...

Je tiens tout d'abord à remercier Dieu le Tout-Puissant et Miséricordieux, qui m'a donné la force et la patience d'accomplir ce modeste travail.

Je tiens à remercier mon encadrant, **M. BOUZERIA Housseyn**, pour son aide, son encouragement, ses précieux conseils et sa disponibilité tout au long de ce travail.

Je souhaite exprimer ma gratitude et mes sincères remerciements à tous les membres du jury pour avoir bien voulu donner de leur temps pour lire ce travail et faire partie des examinateurs. Leurs critiques et commentaires seront bénéfiques pour enrichir mes connaissances dans ce domaine.

Je remercie l'ensemble du personnel de **l'Agence Nationale des Ressources Hydrauliques (ANRH)** de m'avoir fourni les données nécessaires qui m'ont aidé à rédiger mon mémoire.

Mes vifs remerciements vont aussi à tous nos professeurs pour leur orientation, leur confiance et leur patience, qui ont constitué un apport considérable pour ma formation et mon enseignement durant toutes ces années d'études.

Enfin, je tiens aussi à remercier toutes les personnes qui ont participé de près ou de loin à la réalisation de ce travail.

Table des matières

Résumé	I
Dédicace	III
Remerciement	IV
Table des matières	V
Liste des figures	VIII
Liste des tableaux	IX
Liste des abréviations	X
Introduction générale.....	1

Chapitre I. Revue De La Littérature

I.1. Généralités sur l'érosion hydrique	2
I.2 Facteurs de l'érosion hydrique	2
I.2.1. Le climat	2
I.2.2. Le sol	2
I.2.3. Couvert végétal	2
I.2.4. Facteur topographique	3
I.2.5. Facteur humain	3
I.3. Différentes formes d'érosion	3
I.3.1 L'érosion en nappe	3
I.3.2 L'érosion linéaire	4
I.3.3 L'érosion en masse	5
I.3.4 Erosion de berger	5
I.3.5. L'érosion par ravinement	6
I.4. Processus de l'érosion hydrique	7
I.5. Conséquence de l'érosion hydrique	8
I.6. La cartographie de la susceptibilité à l'érosion	8
I.6.1. Définition de la cartographie	8
I.6.2. Importance de la cartographie de la susceptibilité à l'érosion.....	9
I.7. Méthodes traditionnelles de cartographie de la susceptibilité à l'érosion	9
I.8. Intégration de données géospatiales dans l'étude environnementale	10
I.8.1. Analyse géospatiale	10
I.8.2. Les données géospatiales	11
I.8.3. L'étude environnementale	11
I.8.4. Intégration de données géospatiales dans l'étude environnementale	11
I.9. Objectifs de la recherche.....	12
I.9.1. Objectif principal	12
I.9.2. Objectifs spécifiques	13
I.10. Justification du choix de la méthodologie	13
I.10.1. Avantages de l'intégration de données géospatiales	13
I.10.1.1. Utilisation de l'analyse géospatiale	13
I.10.1.2. Types de données géospatiales	15
I.10.1.3. Enjeux dans la gestion des informations géospatiales	15
I.10.1.4. Avantage de l'intégration de données géospatiales	16
I.10.2. Rational derrière l'utilisation de l'apprentissage automatique	17
I.10.2.1. L'apprentissage automatique	17
I.10.2.2. Types d'apprentissage automatique	17

I.10.2.3. Avantages de l'apprentissage automatique	18
I.10.2.4. Domaines d'application de l'apprentissage automatique	19
I.10.2.5. La raison fondamentale derrière l'utilisation de l'apprentissage automatique	20
I.11. Conclusion.....	20

Chapitre II. Cadre Théorique Et Méthodologie

II.1. Introduction.....	21
II.2. Description du bassin versant Isser	21
II.2.1. Situation géographique.....	21
II.2.2. Lithologie du bassin versant Isser	23
II.2.3. Paramètre de relief	24
II.2.4. Topographie du bassin versant Isser	25
II.2.5. Les caractéristiques hydrologiques du bassin versant Isser.....	26
II.2.5.1. Réseau hydrographique et principaux cours d'eau	26
II.2.5.2. Régime hydrologique et variations saisonnières.....	27
II.2.5.3. Apports en eau et débits moyens	27
II.2.5.4. Qualité des eaux de surface	28
II.2.5.5. Eaux souterraines et aquifères	28
II.2.6. Impacts des activités humaines sur l'hydrologie	28
II.3. Collecte des données	28
II.3.1. Sources de données géospatiales et méthodes d'acquisition	28
II.3.1.1. Indice de Fournier Modifié (Modified Fournier Index : MFI)	29
II.3.1.2. Pente, Elévation, Aspects et TWI	31
II.3.1.3. BSI, NDBAI et NDBI.....	31
II.3.1.4. NDVI, NDWI et NBR	32
II.3.1.5. Facteur d'érodibilité (K)	32
II.3.1.6. Occupation des sols.....	33
II.3.1.7. Distance par rapport aux routes/oueds	34
II.4. Méthodes d'apprentissage automatique	35
II.4.1. Définition et contexte théorique	35
II.4.2. Types de méthodes d'apprentissage automatique	35
II.4.2.1. Apprentissage supervisé.....	35
II.4.2.2. Apprentissage non supervisé.....	35
II.4.2.3. Apprentissage semi-supervisé.....	36
II.4.2.4. Apprentissage par renforcement.....	36
II.4.2.5. Apprentissage profond.....	36
II.4.2.6. Méthodes de transfert d'apprentissage.....	36
II.4.3. Algorithmes spécifiques	37
II.4.4. Prétraitement des données	38
II.4.5. <i>Evaluation</i> des performances	39
II.5. Conclusion	40

Chapitre III. Résultats Et Discussions

III.1. Introduction.....	41
III.2. Présentation des cartes générées	41
III.2.1. Carte des pentes	41
III.2.2. Carte des élévations.....	42
III.2.3. Carte des aspects	43
III.2.4. Indice d'humidité topographique (TWI)	44
III.2.5. Carte d'occupation du sol (LU/LC)	45
III.2.6. Facteur d'érodibilité du sol (K)	46
III.2.7. La distance par rapport à la route	47

III.2.8. La distance par rapport aux oueds	48
III.2.9. Classes des sols.....	49
III.2.10. Carte de l'Indice de Fournier Modifié (MFI)	50
III.3. Cartographie de la susceptibilité à l'érosion	51
III.3.1. XGBoost Classifier.....	53
III.3.2. Random Forest.....	53
III.3.3. CatBoost Classifier	54
III.3.4. LightGBM.....	54
III.4. Analyse des effets des changements des classes de sol sur la perte de sols.....	55
III.5. Conclusion	56
Conclusion générale.....	57
Références bibliographiques	58

Liste des Figures

Chapitre I. Revue De La Littérature

Figure.I.1. Erosion en nappe au niveau de la route Tazouta-Skoura	4
Figure.I.2. Erosion linéaire	4
Figure.I.3. Erosion en masse.....	5
Figure.I.4. Erosion du berger oued Saida 2013	6
Figure.I.5. Erosion par ravinement dans la région de Mascara, Algérie	7
Figure.I.6. Processus de l'érosion hydrique.....	8

Chapitre II. Cadre Théorique Et Méthodologie

Figure II.1. Situation géographique du bassin versant Isser.....	21
Figure II.2. Sous bassins du bassin versant Isser	23
Figure II.3. Lithologie du bassin versant Isser	24
Figure II.4. Courbe hypsométrique du bassin versant Isser	25
Figure II.5. Distribution des altitudes dans le bassin Isser.....	26
Figure II.6. Réseau hydrographique du bassin versant Isser	27
Figure II.7. Organigramme de la méthodologie	29
Figure II.8. Distribution des pluies moyennes annuelles et MFI.....	31

Chapitre III. Résultats Et Discussions

Figure.III.1. Carte des pentes de bassin versant Isser	42
Figure.III.2. Carte des élévations de bassin versant Isser	43
Figure.III.3. Carte des aspects de bassin versant Isser	44
Figure.III.4. Indice d'humidité topographique (TWI)	45
Figure.III.5. Carte d'occupation du sol (LULC)	46
Figure.III.6. Carte de facteur K.....	47
Figure.III.7. Carte de distance par rapport à la route.....	48
Figure.III.8. Carte de distance par rapport aux oueds	49
Figure.III.9. Classes des sols dans le bassin versant Isser	50
Figure.III.10. Carte de MFI de bassin versant Isser.....	51
Figure III.11. Courbes ROC-AUC pour les quatre modèles.....	52
Figure III.12. Carte de susceptibilité à l'érosion par utilisation des algorithmes : (a) XGBClassifier, (b) Random Forest, (c) CatBoostClassifier et (d) LightGBM.....	53

Liste des Tableaux

Chapitre II. Cadre Théorique Et Méthodologie

Tableau II.1. Caractéristiques morphométriques des cinq sous bassins	22
Tableau II.2. La répartition altimétrique dans le bassin Isser.....	25
Tableau II.3. Description de la plage d'érosivité de MFI.....	29
Tableau II.4. Stations pluviométriques utilisées.....	30
Tableau II.5. Formules de calcul des indices BSI, NDBAI et NDBI	32
Tableau II.6. Sources de données géospatiales	34
Tableau II.7. Caractéristiques des algorithmes choisis	38

Chapitre III. Résultats Et Discussions

Tableau III.1. Classe des pentes	42
Tableau III.2. Classe d'occupation du sol.....	46
Tableau III.3. Classe des sols	50
Tableau III.4. Critères d'évaluation des modèles d'apprentissage automatique	52
Tableau III.5. Importance des caractéristiques des modèles XGBoost Classifier, Random Forest, CatBoost Classifier et LightGBM.....	54
Tableau III.6. Statistiques des zones sujettes à l'érosion dans différentes classes des sols	56

Liste d'abréviation

- Ct** : Le coefficient de drainage qui évalue l'efficacité du réseau hydrographique.
- Dd (km.km²)** : La densité de drainage.
- Hmax (m)** : Les altitudes maximale du bassin
- Hmin (m)** : Les altitudes minimale du bassin, influençant le gradient de pente.
- Hmoy (m)** : L'altitude moyenne du bassin.
- Ig (‰)** : L'indice de gradient.
- Ip (‰)** : L'indice de pente.
- Im (‰)** : L'indice de modulation de la pente.
- Kf** : Un autre facteur de forme qui aide à comprendre la forme du sous bassin.
- Kc** : le coefficient de compacité.
- Lcp (km)** : La longueur totale des cours d'eau dans le bassin.
- L (km)** : La longueur maximale du bassin.
- MFI** : Modified Fournier index
- P (km)** : Le périmètre.
- Pi** : pluie mensuelle.
- Ptotal** : pluie annuelle.
- Rf** : Le facteur de forme qui mesure l'efficacité du sous bassin à transmettre l'eau.
- Re** : l'indice d'élongation.
- Rc** : le coefficient de circularité.
- S (km²)** : la superficie.
- SIG** : système d'information géographique.
- TWI** : topographique waitress index.

L'érosion hydrique, responsable de la détérioration du sol et de la qualité de l'eau, affecte diverses régions mondiales, notamment le Maghreb, l'Amérique du Sud, l'Amérique centrale et l'Europe. Cette dégradation, attribuée aux activités humaines et aux changements climatiques (Aissaoui & Benghazi, 2023), se traduit par une perte de terres fertiles, laissant place à des terrains appauvris. En Afrique du Nord, elle entraîne des dégradations spécifiques dépassant 2000 tonnes par kilomètre carré par an, contribuant à un engorgement annuel moyen de 125 millions de mètres cubes dans les barrages (Remini, 2003).

En Algérie, l'érosion hydrique dans les régions montagneuses, aggravée par les périodes de sécheresse et l'activité humaine, perturbe gravement les écosystèmes. Ce phénomène, répandu dans les pays méditerranéens, prend des proportions considérables en raison des fortes pluies torrentielles, de la vulnérabilité des terrains et des pratiques humaines néfastes telles que la déforestation et l'urbanisation chaotique. Selon la FAO (Food and Agriculture Organization of the United Nations), 45 % des terres algériennes sont menacées par cette érosion (FAO, 1990). Les conditions géomorphologiques, climatiques et socio-économiques en Algérie favorisent également une dégradation élevée des sols, touchant 85 % de la surface cultivable, soit 6 millions d'hectares (Yles et al, 2012).

Le bassin versant Isser est hautement vulnérable à l'érosion, accentuée par les sécheresses et la fragilité des marnes et argiles présentes. Cette érosion entraîne la détérioration de la couche arable, la perte de fertilité du sol, le relâchement des plantes et la pollution par les pesticides dans les eaux de surface. Les facteurs permanents, évolutifs et aléatoires, notamment liés au sol, à la topographie et aux précipitations, contribuent à ce phénomène. Socio-économiquement, les conséquences sont significatives, avec des pertes de sols pouvant atteindre jusqu'à 5000 tonnes par kilomètre carré, une réduction de 0,5 % de la capacité de stockage des barrages, et des impacts négatifs sur l'agriculture et l'environnement local (Bouchelaghem et al, 2012).

Le constat général de la dégradation du bassin versant souligne l'importance cruciale d'une cartographie précise des zones à risque d'érosion pour orienter efficacement les actions de préservation des sols (lahaye & ladet, 2015). Face à cette menace croissante, notre projet propose une méthode utilisant des techniques d'apprentissage automatique pour cartographier les zones sujettes à l'érosion hydrique dans le bassin Isser. En combinant des données de terrain, de télédétection, topographiques et géospatiales, nous utilisons le logiciel ArcGIS pour générer des cartes détaillées. Ces cartes identifient les zones les plus touchées par l'érosion et établissent des priorités pour des mesures de lutte appropriées. En intégrant les parties prenantes locales et en adaptant les actions aux besoins spécifiques du terrain, nous visons une gestion durable des ressources naturelles dans le bassin Isser.

Chapitre I.
Revue de la littérature...

I.1. Généralités sur l'érosion hydrique

L'érosion hydrique, résultant de processus chimiques, physiques et organiques, altère les roches et érode progressivement le sol, principalement par l'action de l'eau. Elle se décline en deux formes : diffuse, due à une faible lame d'eau et vitesse d'écoulement, et concentrée, causée par une lame d'eau à forte vitesse, entraînant des incisions et arrachant des particules de terre. Les zones arides et semi-arides, où la végétation est rare, et les sols sableux profonds, comme en Mauritanie, sont particulièrement vulnérables (Medjber, 2011). Cette érosion, lente ou rapide, diminue le potentiel de production et la qualité de l'eau de surface, influencée par divers facteurs tels que les propriétés des sols, le relief, la couverture végétale, le climat et les pratiques agricoles. Des actions sont nécessaires pour la contenir, notamment en intégrant des interventions paysagères et des pratiques innovantes dans l'agriculture (Roose et Sarrailh, 1990). Les causes de la dégradation environnementale sont multiples et varient selon les contextes (Aouadj et Gasem 2009).

I.2 Facteurs de l'érosion hydrique

I.2.1. Le climat

Le climat exerce une influence déterminante sur le processus d'hydrolyse, avec les précipitations comme principal facteur. L'intensité des précipitations influe sur l'hydrolyse en pouvant désagréger les agrégats et disperser les particules du sol. Les précipitations les plus fortes surviennent généralement de manière brève mais intense. Cependant, même les précipitations moins intenses mais prolongées peuvent entraîner une perte significative de sol (Peigne, 2013 ; Fertisols, 2020).

I.2.2. Le sol

La composition et la configuration du sol jouent un rôle essentiel dans le processus d'érosion hydrique. Les sols présentant une capacité d'infiltration limitée et un potentiel élevé de ruissellement sont particulièrement vulnérables à ce phénomène. En effet, lorsque les sols ne parviennent pas à absorber rapidement l'eau de pluie, celle-ci s'accumule à la surface, augmentant ainsi le risque d'érosion. Par conséquent, ces sols sont plus sujets à être emportés par les précipitations, ce qui peut entraîner une érosion significative du sol (Peigne, 2013 ; Fertisols, 2020).

I.2.3. Couvert végétal

La présence de végétation est plus efficace pour prévenir l'érosion que pour atténuer l'impact des gouttes de pluie ou ralentir le ruissellement pendant certaines saisons. Cependant, évaluer précisément la protection offerte par la végétation nécessite de spécifier les pratiques agricoles utilisées. Les techniques telles que les bandes végétalisées et les paillis jouent un rôle crucial dans la réduction du ruissellement et l'accumulation des particules solides, bien que leur effet sur le volume du ruissellement soit limité. Une couverture végétale complète contribue significativement à la conservation des sols et de l'eau (Peigne, 2013 ; Fertisols, 2020).

I.2.4. Facteur topographique

La pente exerce une influence majeure sur le processus d'érosion, mais il n'est pas nécessaire que la pente soit abrupte pour que l'érosion et le ruissellement intensif se produisent. La configuration, l'inclinaison et la longueur de la pente interviennent dans le développement de l'érosion. Au fil du temps, les parcelles érodées présentent une tendance à devenir de plus en plus concaves sous l'effet d'une érosion intense. La relation entre l'inclinaison de la pente et l'ampleur de l'érosion est exponentielle (Fournier, 1967 ; Arabi et Roose, 1989 ; Aouadj et al, 2009).

I.2.5. Facteur humain

Les activités humaines, notamment l'agriculture intensive, la déforestation et l'urbanisation, peuvent aggraver l'érosion hydrique en perturbant l'équilibre des facteurs naturels évoqués précédemment. Par exemple, les pratiques agricoles telles que le labour intensif ou le surpâturage peuvent dénuder le sol de sa couverture végétale protectrice, augmentant ainsi sa vulnérabilité à l'érosion hydrique. De même, la déforestation, qui supprime les racines qui maintiennent le sol en place, peut entraîner une augmentation du ruissellement et de l'érosion. En outre, l'urbanisation, en remplaçant les sols naturels par des surfaces imperméables telles que le béton ou l'asphalte, peut accélérer le ruissellement des eaux de pluie, exacerbant ainsi les problèmes d'érosion hydrique. Ces activités humaines modifient donc les conditions environnementales naturelles, créant des déséquilibres qui favorisent l'érosion hydrique (Peigne, 2013 ; Fertisols, 2020).

I.3. Différentes formes d'érosion

I.3.1 L'érosion en nappe

Cela se manifeste par une diminution de la visibilité de l'horizon humifère (blanchissement du sol). Les particules se déplacent et se transportent sur toute la surface. Cependant, son action demeure limitée et ne fait que décaper le sol de manière sensiblement homogène. Les horizons supérieurs, généralement de couleur foncée, sont les premiers à être érodés, tandis que les horizons inférieurs sont plus clairs (Roose et al, 2012).



Figure.I.1. Erosion en nappe au niveau de la route Tazouta-Skoura (Faleh et Maktite, 2020)

I.3.2 L'érosion linéaire

L'inclinaison du terrain influe sur le degré d'érosion, mais des processus d'érosion et de ruissellement peuvent se produire même sur des pentes douces. L'influence de l'inclinaison se manifeste par la forme, l'inclinaison et la longueur de la pente. Avec le temps, les zones érodées se creusent davantage, subissant une érosion accrue et devenant de plus en plus concaves. L'augmentation de l'inclinaison du terrain entraîne une croissance exponentielle de l'érosion. (Roose et al, 2012).



Figure.I.2. Erosion linéaire (Binet et Gonnot, 2005)

I.3.3 L'érosion en masse

L'érosion de masse englobe divers mouvements de terrain imprévisibles, comme les glissements de terrain, les coulées de boue et les déplacements de masse. Ces événements altèrent des quantités importantes de sol et peuvent résulter de plusieurs facteurs, tels que la saturation excessive des pentes instables, l'humidification des surfaces lors de fortes précipitations rares ou même l'énergie des gouttes de pluie en amont. Ces mouvements sont catégorisés en glissements lents, glissements rapides et brutaux, et coulées de boue. Leur impact peut être dévastateur, causant des inondations, des dommages aux bâtiments et cultures, des pertes humaines, ainsi que des perturbations des systèmes de drainage et de la qualité de l'eau. Une compréhension approfondie de ces mécanismes est cruciale pour élaborer des stratégies efficaces d'atténuation des risques. (Roose et al, 2012).



Figure.I.3. Erosion en masse (Silvia,et al,2020).

I.3.4 Erosion de berger

L'érosion des berges est un processus naturel influencé par divers éléments tels que la gravité, les précipitations, le ruissellement et les vagues. Ce phénomène se caractérise par le déplacement graduel de matériaux des zones amont vers l'aval, ce qui contribue à modeler les paysages. L'action de l'eau, l'érosion causée par le vent et les mouvements de terrain sont parmi les principales causes de l'érosion. Les berges non consolidées, comme les sédiments d'abrasion, sont particulièrement vulnérables à ce phénomène. La morphologie du cours d'eau, la pente des berges et la composition des matériaux jouent un rôle crucial dans ce processus. De plus, les activités humaines telles que la déforestation et l'aménagement du territoire peuvent exacerber l'érosion, entraînant des conséquences telles que la perte de terres, la détérioration de la qualité de l'eau et l'érosion côtière. (Roose et al, 2012).



Figure.I.4. Erosion du berger oued Saida 2013 (Remmas, 2014)

I.3.5. L'érosion par ravinement

L'érosion par ravinement est un phénomène d'érosion causé par l'action de l'eau, qui affecte les voies de drainage sur les pentes et peut entraîner des dommages significatifs. Ce processus se produit lorsque le ruissellement devient intense et que l'érosion en rigoles devient incontrôlée, générant ainsi des courants plus puissants qui creusent des ravins de plus en plus profonds. Les ravins se forment lorsque le ruissellement en nappe se concentre, créant des fissures qui érodent progressivement le sol pour former des structures linéaires en profondeur. Bien que moins fréquentes que les rigoles, les ravines sont de plus grande taille et peuvent entraîner des problèmes tels que des inondations, des perturbations dans le drainage et des pertes de terres cultivables. Pour contrer l'érosion par ravinement, des mesures telles que la plantation de couvertures végétales, la construction de terrasses ou de canaux de drainage, ainsi que la réduction de l'impact du travail du sol sont recommandées. (Roose et al,2012).



Figure.I.5. Erosion par ravinement dans la région de Mascara, Algérie (Rebai et al, 2012).

I.4. Processus de l'érosion hydrique

L'érosion hydrique est un processus complexe résultant du détachement, du transport et du dépôt des particules de sol. Ce phénomène, influencé par l'action combinée de la pluie et du ruissellement, varie selon la résistance du milieu (sol, couvert végétal, pratiques agricoles) et la topographie. Principalement causée par la pluie, l'érosion hydrique peut être accentuée par divers facteurs tels que le climat, le type de sol, le couvert végétal, le relief et la configuration du paysage. Les précipitations, selon leur intensité et leur durée, déterminent leur capacité à provoquer cette érosion, appelée érosivité. Ce processus commence par la désintégration des agrégats de sol en surface, formant une croûte qui limite l'infiltration de l'eau et favorise le ruissellement, pouvant entraîner une érosion diffuse ou concentrée selon la pente et la taille de la surface. Les activités humaines, notamment dans les zones agricoles, peuvent aggraver ce phénomène, notamment par des pratiques telles que le labourage intensif, la monoculture et la réduction du couvert végétal.

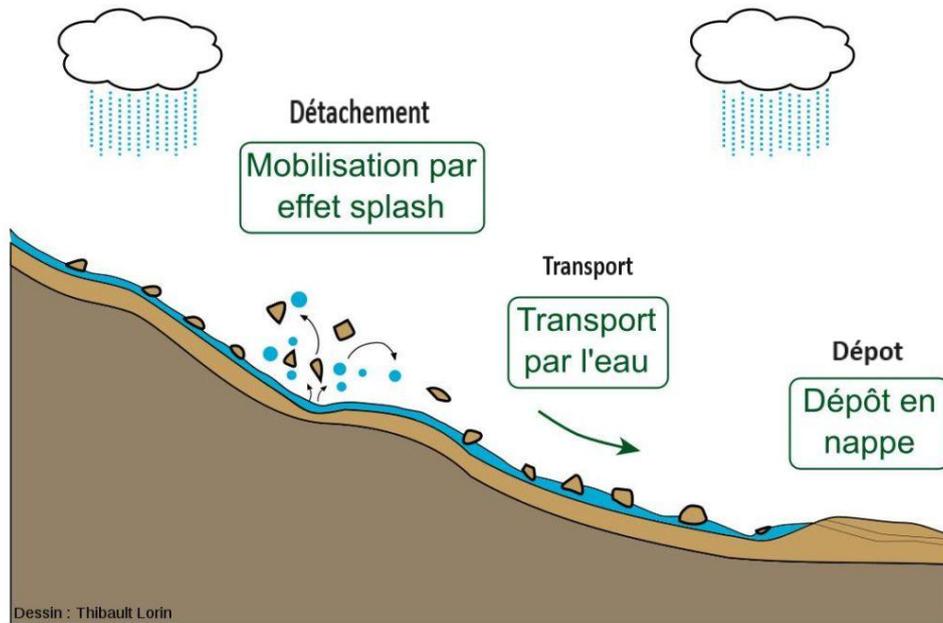


Figure.I.6. Processus de l'érosion hydrique (Rebai et al, 2012).

I.5. Conséquence de l'érosion hydrique

Les effets de l'érosion hydrique comprennent plusieurs conséquences néfastes telles que l'envasement des barrages, la détérioration du bilan hydrique et l'altération de la qualité des sols. D'autres répercussions incluent le déchaussement des cultures, la fragmentation accrue des parcelles, la diminution du drainage profond menant à une moindre recharge des nappes phréatiques, ainsi que des dommages aux infrastructures routières et hydrauliques. En outre, cette érosion entraîne la dégradation des sols, la perte de leur fertilité, la diminution des rendements agricoles, la dégradation de la qualité de l'eau, et le transport de sédiments chargés en éléments nutritifs et en substances phytosanitaires dans les cours d'eau et les nappes souterraines, provoquant une turbidité nuisible. (Medjber, 2011).

I.6. La cartographie de la susceptibilité à l'érosion

I.6.1. Définition de la cartographie

La cartographie englobe l'ensemble des méthodes et des techniques utilisées pour créer, représenter et étudier des cartes géographiques. Elle consiste à simplifier et à conventionnaliser la représentation de la surface terrestre sur des supports plats tels que des cartes ou des graphiques. Cette discipline combine à la fois la pratique, la science et l'art, et elle est utilisée pour diverses applications, allant de la représentation des zones géographiques à l'analyse des relations spatiales.

En plus de son utilisation traditionnelle pour représenter des territoires et des frontières politiques, la cartographie est également employée dans d'autres domaines tels que la biologie pour la disposition des gènes, la génétique pour la cartographie génétique, et la représentation des chromosomes. De nos jours, les systèmes d'information géographique (SIG) sont souvent utilisés pour créer, analyser

et visualiser des données cartographiques de manière informatisée. (Boughalem et al, 2012)

Au-delà de sa fonction de simple représentation, la cartographie nous aide à comprendre notre position dans le monde, à analyser les relations spatiales et à réfléchir à l'impact de la géographie sur notre vie quotidienne. Chaque carte est façonnée par l'intention du cartographe et l'objectif spécifique de la carte, ce qui en fait à la fois un outil pratique et une forme d'expression artistique. (Bkhairi,2012 ; Ouallali et al,2016)

I.6.2. Importance de la cartographie de la susceptibilité à l'érosion

L'importance de cartographier la susceptibilité à l'érosion réside dans sa contribution cruciale à la gestion des ressources naturelles et à la prévention des catastrophes liées à l'érosion. Cette démarche permet de spatialiser et de hiérarchiser l'intensité de l'érosion sur l'ensemble du bassin versant, ce qui est essentiel pour planifier et mettre en œuvre des mesures de conservation et de lutte contre l'érosion. (Boughalem et al, 2012)

Dans le cadre de l'étude du bassin versant de l'oued Isser, la cartographie de la susceptibilité à l'érosion a été effectuée en intégrant divers paramètres, notamment la pente, l'élévation, les précipitations, la géologie, le sol, la direction des pentes, l'occupation des sols, la distance par rapport à la route et la distance par rapport au plan d'eau. Cette approche a permis de quantifier l'érosion annuelle des sols et d'identifier les zones les plus vulnérables à l'érosion. (Boughalem et al, 2012)

En outre, la cartographie de la susceptibilité à l'érosion peut également être réalisée à partir de techniques de télédétection et de l'utilisation de données de terrain dans un environnement SIG.

En parallèle, la modélisation par intelligence artificielle (IA) offre une autre approche pour évaluer les risques d'érosion hydrique. En prenant en compte une gamme variée de facteurs qui influencent le ruissellement et l'érosion dans divers contextes de sol et d'utilisation des terres, cette méthode permet de générer des cartes de risque d'érosion hydrique grâce aux systèmes d'information géographique (SIG). De plus, elle ouvre des perspectives pour une gestion durable des ressources en eau et des sols, ainsi que pour la réhabilitation des zones touchées par l'érosion, en tenant compte des paramètres de contrôle et en permettant une estimation quantitative de l'érosion. (Boughalem et al, 2012)

I.7. Méthodes traditionnelles de cartographie de la susceptibilité à l'érosion

Les méthodes traditionnelles de cartographie de la susceptibilité à l'érosion sont généralement basées sur une combinaison de données topographiques, géologiques, hydrologiques et climatiques. Voici quelques-unes des méthodes traditionnelles couramment utilisées (Boughalem et al, 2012) :

- **Méthode basée sur la pente et la forme du terrain** : Cette méthode utilise des données topographiques pour évaluer la pente du terrain. Des pentes plus raides sont généralement plus susceptibles à l'érosion. De plus, la forme du

terrain, comme la concavité ou la convexité, peut également être prise en compte pour évaluer la susceptibilité à l'érosion.

- **Méthode basée sur l'utilisation du sol et la végétation** : Cette méthode prend en compte le type d'utilisation des sols (agriculture, forêt, zone urbaine, etc.) ainsi que la couverture végétale. Les sols non protégés par une végétation dense sont plus susceptibles à l'érosion.
- **Méthode basée sur l'hydrologie** : Cette méthode évalue les caractéristiques hydrologiques de la région, telles que la densité du réseau de drainage, la taille des bassins versants et la concentration des précipitations. Les zones avec un drainage concentré et des précipitations intenses sont plus susceptibles à l'érosion.
- **Méthode basée sur les sols** : Cette méthode prend en compte les propriétés physiques et chimiques des sols, telles que la texture, la structure, la perméabilité et la capacité d'infiltration. Les sols plus susceptibles à la fragmentation et à la compaction sont plus à risque d'érosion.
- **Méthode basée sur l'analyse historique** : Cette méthode utilise des données sur les événements passés d'érosion, tels que les glissements de terrain, les crues éclairs et les ravins, pour identifier les zones à risque élevé.

Ces méthodes peuvent être combinées et pondérées pour produire des cartes de susceptibilité à l'érosion qui aident à identifier les zones les plus à risque et à orienter les mesures de conservation et de gestion des terres. Toutefois, il convient de noter que ces méthodes traditionnelles peuvent être complétées voire remplacées par des approches plus avancées basées sur la télédétection, les SIG (systèmes d'information géographique) et la modélisation numérique.

I.8. Intégration de données géospatiales dans l'étude environnementale

I.8.1. Analyse géospatiale

L'analyse géospatiale, également connue sous le nom d'analyse de données géographiques, exploite les données liées à des lieux ou des régions spécifiques pour faciliter la prise de décisions en temps réel. Ce domaine englobe la collecte, la présentation, la manipulation et l'interprétation des données géographiques dans le but de déceler des modèles et des tendances pour éclairer les processus décisionnels (DeRunz et al, 2008 ; Miquel et al, 2002).

Les systèmes d'information géographique (SIG) ainsi que les outils d'analyse statistique et spatiale sont largement employés dans cette discipline pour explorer les relations spatiales entre différentes données. L'application de l'analyse géospatiale s'étend à divers domaines tels que la planification urbaine, la gestion environnementale, la gestion des ressources naturelles, la planification des interventions d'urgence et bien d'autres (DeRunz et al, 2008 ; Miquel et al, 2002).

Diverses données, telles que l'utilisation des sols, la densité de population, la topographie, les conditions météorologiques et l'infrastructure, sont examinées dans le cadre de cette analyse. Elle peut être effectuée à différentes échelles, du local au

mondial, et peut varier en complexité en fonction des données et des objectifs spécifiques (DeRunz et al, 2008 ; Miquel et al, 2002).

I.8.2. Les données géospatiales

Les données géospatiales jouent un rôle crucial dans la gestion de l'érosion hydrique. Elles représentent des informations localisées sur une carte, relatives aux lieux ou aux événements liés à ce phénomène spécifique sur ou près de la surface terrestre. Ces données incluent généralement des informations de localisation telles que des coordonnées géographiques, des caractéristiques des zones affectées par l'érosion, ainsi que des indicateurs temporels sur le moment et la durée de l'événement (Arenas et al, 2020).

Dans le domaine de l'hydrologie, les données géospatiales sont essentielles pour comprendre et prévenir l'érosion hydrique. Elles permettent d'identifier les zones à risque, de suivre l'évolution des phénomènes d'érosion dans le temps, et de concevoir des stratégies de gestion adaptées pour atténuer ses effets néfastes sur les sols et les ressources en eau (Arenas et al,2020).

I.8.3. L'étude environnementale

L'analyse environnementale, ou évaluation environnementale, évalue les impacts écologiques des activités humaines en développement, notamment via des études détaillées telles que les études d'impact environnemental (EIE). Le suivi environnemental surveille les changements écologiques avant, pendant et après un projet pour prévenir les impacts négatifs. Les études environnementales contribuent à la biodiversité, la gestion des déchets et sensibilisent aux enjeux écologiques, offrant des opportunités de carrière variées. L'EIE identifie et évalue les risques environnementaux d'un projet, proposant des mesures pour les atténuer, adoptant une approche proactive. L'évaluation environnementale des projets intègre les préoccupations écologiques dans la planification des projets, guidant la prise de décision en amont. (Fertisols,2020)

I.8.4. Intégration de données

L'intégration de données géospatiales dans les études environnementales revêt une importance capitale pour évaluer et gérer les risques environnementaux, incluant notamment l'hydrologie et l'érosion. Ces données englobent une gamme variée d'informations telles que la topographie, l'hydrographie, et les conditions atmosphériques, permettant ainsi une analyse complète des paramètres environnementaux. Pour garantir une utilisation efficace de ces données, l'interopérabilité est essentielle, impliquant l'adoption de normes et de standards pour assurer leur compatibilité syntaxique et sémantique, facilitant ainsi leur combinaison et leur analyse. (Sheeren,2006 ; Lloyd,2008)

Cette intégration de données géospatiales trouve des applications diverses, allant de la gestion des ressources en eau à la prévention de l'érosion des sols. Elle permet d'identifier les zones à risque et de mettre en place des mesures de gestion adaptées, contribuant ainsi à la préservation des écosystèmes et à la réduction des impacts environnementaux. Par exemple, dans le domaine de l'hydrologie, l'intégration

de données géospatiales aide à cartographier les bassins versants, à modéliser les flux d'eau et à prévoir les risques d'inondation, permettant ainsi une gestion plus efficace des ressources en eau et une meilleure préparation aux événements extrêmes. (Sheeren,2006 ; Lloyd,2008).

L'intégration de données géospatiales dans les études environnementales offre une approche holistique pour évaluer et gérer les risques environnementaux, incluant l'hydrologie et l'érosion. En adoptant des normes et des standards pour assurer l'interopérabilité des données, cette pratique favorise une utilisation efficace des informations géographiques, contribuant ainsi à une meilleure compréhension et à une gestion durable de notre environnement. (Sheeren,2006 ; Lloyd,2008).

I.9. Objectifs de la recherche

I.9.1. Objectif principal

Cette recherche se concentre sur la cartographie de la susceptibilité à l'érosion hydrique dans le bassin versant Isser, au nord de l'Algérie. Son objectif principal est d'exploiter les données géospatiales et l'apprentissage automatique afin de mieux comprendre et prédire les modèles d'érosion dans cette région. Elle revêt une importance majeure pour la gestion durable des terres, la préservation des ressources en eau et l'évaluation des risques environnementaux.

Le travail s'appuiera sur des techniques avancées telles que la télédétection, les systèmes d'information géographique (SIG) et les algorithmes d'apprentissage automatique pour analyser une gamme de facteurs contribuant à l'érosion, tels que la topographie, la couverture terrestre, le type de sol et les précipitations.

Le bassin versant de l'Isser a été choisi comme zone d'étude en raison de ses caractéristiques climatiques, topographiques, géologiques et des modes d'utilisation des terres, qui peuvent tous influencer les taux d'érosion. De plus, les activités humaines telles que la déforestation, le surpâturage et les pratiques agricoles ont aggravé l'érosion dans cette région.

Des travaux antérieurs en Algérie ont déjà montré que l'intégration des données géospatiales et de l'apprentissage automatique peut être efficace pour cartographier le risque d'érosion. Par exemple, une étude antérieure a utilisé pente, exposition, altitude, LULC, indice d'humidité topographique (TWI), distance à la rivière, distance aux routes, rapport minéraux argileux, lithologie par machine learning pour identifier les zones à risque d'érosion dans le bassin versant de Macta. L'inventaire de l'érosion des sols a été créé en intégrant des données de positionnement global, en analysant des images satellite de Google Earth et en divisant l'ensemble de données en deux groupes pour la formation et l'étalonnage, et 40 % pour tester les modèles. Des mesures statistiques telles que la sensibilité, la spécificité et le ROC ont été utilisées pour évaluer la validité des modèles. (Hamza Bouguerra et al, 2023).

Ce travail comprendra plusieurs étapes, notamment la collecte, le traitement et l'analyse des données, ainsi que le développement de modèles prédictifs. Les données géospatiales utilisées comprendront des informations sur la topographie, la couverture terrestre, le type de sol et les précipitations. Ces données seront traitées à

l'aide de techniques SIG et d'apprentissage automatique pour identifier les facteurs contribuant à l'érosion.

Ce travail vise à fournir une analyse approfondie de la cartographie de la susceptibilité à l'érosion hydrique dans le bassin versant de l'Isser en intégrant des données géospatiales et l'apprentissage automatique, avec pour objectif d'informer la prise de décision et l'élaboration de politiques pour une gestion durable des terres et des ressources en eau dans la région semi-aride.

1.9.2. Objectifs spécifiques

Les objectifs spécifiques de cette étude visent à développer une approche de cartographie de la susceptibilité à l'érosion hydrique en utilisant des données géospatiales et des techniques d'apprentissage automatique, en se concentrant particulièrement sur le bassin versant de l'Isser. Voici les points clés :

- ❖ **Évaluation de la sensibilité à l'érosion hydrique** : Il s'agit d'évaluer le degré de sensibilité à l'érosion hydrique dans la zone d'étude en utilisant des modèles traditionnels ainsi que des données climatiques et de télédétection. Les résultats soulignent l'urgence d'interventions pour protéger les terres agricoles et l'environnement.
- ❖ **Cartographie des zones sensibles** : Une méthodologie spécifique est appliquée pour cartographier les zones présentant différents degrés de sensibilité à l'érosion. Cette approche qualitative permet de prendre en compte les spécificités locales et de générer des cartes de sensibilité à l'érosion hydrique.
- ❖ **Intégration des données et modélisation** : L'incorporation de données de télédétection spatiale et de systèmes d'information géographique (SIG) permet de fusionner divers éléments afin de classer les zones par ordre de priorité et de créer une carte consolidée illustrant les niveaux de sensibilité à l'érosion.
- ❖ **Identification des facteurs de contrôle** : Il s'agit de déterminer les facteurs influençant l'érosion hydrique, tels que la pente, l'érodibilité des sols et la couverture végétale, afin de mieux gérer les pratiques antiérosives.
- ❖ **Création d'une carte de vulnérabilité** : À partir des données recueillies, une carte de vulnérabilité à l'érosion des sols est élaborée, permettant une meilleure compréhension des zones à risque et des interventions nécessaires.

I.10. Justification du choix de la méthodologie

I.10.1. Avantages de l'intégration de données géospatiales

I.10.1.1. Utilisation de l'analyse géospatiale

Dans différents domaines, l'analyse des données géospatiales permet d'analyser et d'interpréter les données en fonction de leur localisation géographique. Les illustrations d'analyse géospatiale montrent comment l'analyse des données géospatiales peut être utilisée dans la réalité :

✦ **Gestion des ressources naturelles et de l'environnement** :

L'analyse des données géospatiales est utilisée pour surveiller et gérer les ressources naturelles ainsi que pour évaluer les risques environnementaux. Ces informations permettent aux gestionnaires des ressources de prendre des décisions éclairées concernant la préservation et la gestion durable de ces ressources. Ils ont la capacité d'assister les responsables de l'environnement dans la prise de décisions éclairées pour prévenir les problèmes environnementaux (Fortin et Dale, 2005).

- ✦ **Aménagement urbain** : La géospatialisation est utilisée pour analyser et cartographier la répartition spatiale de la population, des infrastructures et de l'utilisation des sols dans les zones urbaines. Ces données sont précieuses pour éclairer les urbanistes dans leurs décisions d'aménagement et de développement des villes (DeRunz et al, 2008).
- ✦ **Logistique** : L'analyse des données géospatiales est employée pour examiner les schémas de circulation et les infrastructures de transport. Les planificateurs des transports tirent parti de ces informations pour prendre rapidement des décisions concernant la conception et la mise en place des systèmes de transport (Fotheringham et Rogerson, 2010).
- ✦ **Agriculture** : Les données géospatiales sont utilisées pour étudier la qualité des sols, les rendements des cultures et les données climatiques. Ces informations peuvent aider les agriculteurs à prendre des décisions éclairées concernant la sélection des cultures, la gestion de la fertilisation et le calendrier des récoltes (Fortin et Dale, 2005).
- ✦ **Planification de l'implantation des entreprises et des interventions d'urgence** : L'analyse des données géographiques est utilisée pour comprendre les données démographiques des clients, la demande du marché et l'emplacement des concurrents. Elle facilite également la cartographie et l'analyse des ressources d'intervention d'urgence telles que les stations de pompage, les hôpitaux et les itinéraires d'évacuation. Ces informations aident les planificateurs d'urgence à prendre des décisions éclairées sur la planification des interventions et assistent les entreprises dans leurs choix stratégiques de déploiement et de commercialisation (Fortin et Dale, 2005 ; DeRunz et al, 2008) .
- ✦ **Géologie et géophysique** : L'exploitation de l'analyse des données géospatiales permet de cartographier et d'analyser les caractéristiques géologiques telles que les failles, les minéraux et les aquifères souterrains. Les géologues et les géophysiciens peuvent tirer parti de ces données pour prendre des décisions concernant l'exploitation des ressources naturelles, la gestion des risques géologiques et d'autres activités liées à la géologie (Fortin et Dale, 2006 ; Fotheringham et Rogerson, 2010) .

I.10.1.2. Types de données géospatiales

Il existe plusieurs types de données géospatiales utilisées dans les SIG, la télédétection et d'autres applications d'analyse géospatiale. Voici quelques-uns des principaux de ces types :

✦ **Données vectorielles :**

On utilise les données vectorielles afin de représenter des éléments géographiques tels que des **points**, des **lignes** et des **polygones**. Les coordonnées et les relations topologiques entre les caractéristiques sont les formes de données de ce genre (Sheeren, 2006 ; Lloyd, 2008).

✦ **Données matricielles :**

Les données matricielles en géomatique sont des représentations de données spatiales organisées selon une grille régulière de cellules, où chaque cellule détient une valeur représentant une caractéristique géographique spécifique. Les images satellites et les cartes topographiques sont des exemples couramment utilisés de données matricielles en géomatique. (Sheeren, 2006)

✦ **Données GPS :**

Les données GPS (Global Positioning System) sont utilisées pour enregistrer et stocker la localisation ainsi que les déplacements d'objets ou de personnes. Ces données sont largement exploitées dans des applications de navigation, de cartographie et de suivi (Lloyd, 2008).

✦ **Données de recensement :**

Les données du recensement sont utilisées pour représenter les caractéristiques démographiques et socio-économiques d'une population. Elles sont couramment exploitées dans les domaines de la planification urbaine, du marketing et de la recherche en sciences sociales.

✦ **Données météorologiques et climatiques :**

La représentation des conditions atmosphériques et des schémas climatiques repose sur les données météorologiques et climatiques. Les prévisions météorologiques, la modélisation climatique et la surveillance environnementale font souvent appel à ce type de données. (Sheeren, 2006)

I.10.1.3. Enjeux dans la gestion des informations géospatiales

Les données géospatiales ont connu une croissance rapide dans les secteurs gouvernementaux, privés, académiques et même parmi le grand public, souvent impliqué dans la collecte involontaire de données. Cependant, avec la prolifération rapide des données d'Observation de la Terre (OT) provenant de satellites et l'émergence de plateformes basées sur le cloud pour l'OT, accessibles via le web pour de nombreuses organisations et entreprises, la gestion de ces données devient un défi majeur. La multiplicité des sources peut entraîner des incohérences et des problèmes de stockage (Innerebner et al, 2017).

La capacité des satellites à capturer des données en grande quantité et à haute résolution spatiale et spectrale représente un défi majeur. Cette avancée technologique nécessite la gestion de volumes de téléchargement considérables, nécessitant une infrastructure informatique interne encore plus robuste. Si le traitement des données en local est privilégié par rapport au stockage sur le cloud, cette infrastructure devra être encore plus étendue. (Fortin et Dale, 2005 ; Innerebner et al, 2017)

Un défi supplémentaire réside dans la diversité des façons dont les éléments de données, tels que la localisation, peuvent être définis, incluant des aspects tels que

les noms, les coordonnées, les frontières administratives, les codes NUTS, les plans d'étage, les polygones, les arrêts de bus, etc. Il devient encore plus difficile de rendre ces données et services de localisation partageables, réutilisables et interopérables. L'incertitude, l'imprévisibilité et l'irrégularité quant à l'accès et au type de données auxquelles on peut accéder ajoutent une couche supplémentaire de complexité au problème (DeRunz et al, 2008).

De plus, l'absence d'implication des dirigeants, de financement adéquat, de consensus entre les parties prenantes et de gouvernance précise peut également présenter des défis pour la standardisation. Il est également crucial de comprendre la valeur ajoutée d'une approche basée sur des normes, ainsi que les lacunes en termes de connaissances et d'expérience dans la mise en œuvre de ces normes (Lloyd, 2008 ; Arenas et al, 2020).

Ainsi, la standardisation des données géospatiales revêt une importance capitale : une donnée n'a de réelle valeur que si elle est accessible et utilisable. Étant donné la diversité des sources de données et des utilisateurs, il est essentiel de faciliter la découverte et le partage de ces informations (Miquel et al, 2002).

I.10.1.4. Avantage de l'intégration de données géospatiales

L'intégration de données géospatiales offre de nombreux avantages pour les entreprises et les organisations :

- Précision améliorée : Les technologies avancées telles que le GPS et la télédétection fournissent des informations précises pour la navigation, la gestion des ressources et l'évaluation des risques.
- Réduction des coûts : L'automatisation de la collecte de données et la rationalisation des processus permettent des économies de coûts significatives.
- Intégration des données : La fusion de données provenant de différentes sources permet une meilleure compréhension et analyse, conduisant à des décisions plus éclairées.
- Enrichissement des données : Le géocodage et la validation des données d'adresse ajoutent du contexte et améliorent la compréhension des relations entre les ensembles de données.
- Visualisation des données : La représentation sur des cartes facilite la compréhension, l'identification des modèles et des tendances.
- Évolutivité : La gestion de données à grande échelle permet l'utilisation de modèles d'apprentissage automatique pour l'analyse prédictive.
- Sécurité des données : Les plates-formes d'intégration garantissent la confidentialité et la sécurité des données.
- Accessibilité des données : Les données de haute qualité deviennent plus accessibles pour diverses applications.

En résumé, l'intégration de données géospatiales facilite la recherche, la traçabilité et la représentation des caractéristiques, jouant ainsi un rôle crucial dans la fourniture d'informations vitales pour les entreprises dans un marché de plus en plus numérique.

I.10.2. Rational derrière l'utilisation de l'apprentissage automatique

I.10.2.1. L'apprentissage automatique

L'apprentissage automatique, une branche de l'intelligence artificielle, consiste à élaborer des séquences d'actions pour résoudre des problèmes, en optimisant automatiquement à partir de l'expérience avec une intervention humaine minimale. Il trouve des applications dans divers domaines comme la classification, la détection, la génération de contenu à partir de bases de données, la conception, l'analyse, l'optimisation, le développement et la mise en œuvre de méthodes. Ses variantes incluent l'apprentissage supervisé, non supervisé et par renforcement, selon le degré d'intervention humaine nécessaire (Essamari, 2018).

Dans le secteur financier, l'apprentissage automatique est utilisé pour évaluer la valeur fondamentale des actifs, améliorant ainsi la gestion des risques de portefeuille. Il promet également d'optimiser les processus des services financiers pour les clients et de favoriser les interactions entre les systèmes et le personnel utilisant l'IA, afin d'améliorer la prise de décision.

Par ailleurs, il contribue à améliorer la qualité des données en identifiant automatiquement les anomalies pour les signaler aux statisticiens et/ou aux sources. Cela favorise des rapports de meilleure qualité et moins coûteux, ainsi qu'un traitement des données et des analyses plus efficaces.

Cependant, l'apprentissage automatique présente des limites telles que la précision des prédictions qui ne peut jamais être de 100 %, la complexité des algorithmes et la dépendance des performances du modèle aux données d'entraînement. Le fléau de la dimension est un problème récurrent, surtout dans l'analyse de données à grande échelle, où la dimension de l'espace augmente rapidement et les données représentatives deviennent éparses (Janny et al, 2022).

I.10.2.2. Types d'apprentissage automatique

Les algorithmes employés dans l'apprentissage automatique sont des modèles mathématiques élaborés à partir de données afin d'accomplir des tâches précises. Ils sont classés généralement en trois catégories principales : l'apprentissage supervisé, l'apprentissage non supervisé et l'apprentissage par renforcement. (Perkowitz et Etzioni, 1997 ; Perkowitz et Etzioni, 2000)

A. L'apprentissage supervisé : L'apprentissage supervisé est une approche de l'apprentissage automatique où un modèle est formé sur un jeu de données étiquetées, où chaque exemple est associé à une étiquette de sortie spécifique. Ces données étiquetées fournissent des réponses correctes, ce qui permet au modèle d'apprendre à prédire des résultats pour de nouvelles données non étiquetées. Cette méthode est couramment utilisée pour des tâches telles que la classification, la régression et la détection d'objets. Les algorithmes d'apprentissage supervisé comprennent des techniques telles que la régression linéaire, la régression logistique, les arbres de décision, les machines à vecteurs de support (SVM), les réseaux neuronaux, et bien d'autres (Perkowitz et Etzioni, 1997).

- B. L'apprentissage non supervisé :** L'apprentissage non supervisé est une partie du machine learning qui consiste à examiner et à regrouper des données qui ne sont pas étiquetées. Contrairement à l'apprentissage supervisé où le modèle reçoit des données étiquetées pour s'entraîner, l'apprentissage non supervisé apprend à partir de données non étiquetées pour découvrir des motifs ou des groupes dans les données. Les algorithmes d'apprentissage non supervisé sont utilisés pour des tâches comme le regroupement et l'association, afin de rassembler des données selon des similitudes ou des structures communes. On utilise ce type d'apprentissage lorsque l'étiquetage des données prend trop de temps ou d'argent. Il est utilisé dans divers domaines comme la détection d'anomalies, la segmentation d'image, les recommandations personnalisées et la publicité ciblée, ainsi que dans d'autres applications où découvrir des schémas cachés dans les données est important ([Perkowitz et Etzioni, 1997](#) ; [Perkowitz et Etzioni, 2000](#)).
- C. L'apprentissage par renforcement :** L'apprentissage par renforcement est une technique de machine learning où un agent apprend à prendre les meilleures décisions en interagissant avec son environnement. Dans ce processus, l'agent utilise une stratégie pour choisir une action, et son objectif est d'améliorer cette stratégie en recevant des récompenses positives ou des pénalités. Cette méthode est idéale pour les problèmes où il faut jongler entre les récompenses à court terme et celles à long terme. Les types principaux de problèmes d'apprentissage par renforcement incluent les bandits algorithmiques, les problèmes de décision markoviens et les arbres de jeu. On trouve des applications de cette méthode dans des domaines comme la robotique, les jeux vidéo, les systèmes de recommandation et la gestion des feux de circulation. ([Janny, et al, 2022](#)).

I.10.2.3. Avantages de l'apprentissage automatique

Adaptation Automatique : Les systèmes peuvent s'ajuster automatiquement aux changements sans nécessiter de révisions manuelles.

Amélioration de la Précision : Avec un entraînement continu, les modèles deviennent plus précis et fiables, ce qui renforce la qualité des décisions.

Traitement de Grandes Quantités de Données : L'apprentissage automatique est capable de gérer des volumes massifs de données, ce qui s'avère bénéfique dans des domaines comme la santé, les finances et l'automobile.

Analyse de Données Complexes : Ces algorithmes peuvent repérer des tendances et des motifs dans des données complexes et volumineuses que les humains pourraient ne pas détecter.

Recommandations personnalisées : L'apprentissage automatique peut fournir des recommandations adaptées à chaque utilisateur, que ce soit dans le domaine du commerce électronique ou des services de streaming.

Réduction des Coûts : L'automatisation des tâches répétitives permet de minimiser les coûts de main-d'œuvre et d'améliorer l'efficacité opérationnelle ([Thierry et al, 2022](#) ; [Margarida et al, 2023](#)).

I.10.2.4. Domaines d'application de l'apprentissage automatique

Les applications de l'apprentissage automatique sont vastes et touchent divers secteurs. Voici quelques domaines clés où il est largement utilisé :

Finance :

- Détection de fraude dans les transactions financières.
- Offres personnalisées basées sur les profils clients.
- Prédire la solvabilité des emprunteurs potentiels.
- Anticiper les fluctuations du marché pour de meilleures stratégies d'investissement.

Commercialisation :

- Offres commerciales personnalisées.
- Analyse des publications sur les réseaux sociaux pour une publicité ciblée.
- Optimisation des campagnes publicitaires.

Environnement :

- Analyse des niveaux de pollution.
- Détection des nuisances affectant la faune.
- Évaluation des taux de déforestation et de surpêche.

Sécurité :

- Reconnaissance faciale et vocale.
- Détection de fraude vocale potentielle.
- Sécuriser les transactions bancaires et les données sensibles.

Soins de santé :

- Détection des maladies avec une haute précision.
- Analyse des données médicales pour des diagnostics plus précis.
- Prédiction des tendances et des résultats médicaux.

Ces exemples illustrent la polyvalence et l'importance croissante de l'apprentissage automatique dans des domaines tels que la finance, le marketing, l'environnement, la sécurité et la santé. ([Perkowitz et Etzioni, 1997](#) ; [Perkowitz et Etzioni, 2000](#) ; [Janny et al, 2022](#)).

I.10.2.5. La raison fondamentale derrière l'utilisation de l'apprentissage automatique

L'utilisation principale de l'apprentissage automatique réside dans sa capacité à permettre aux ordinateurs d'apprendre à partir de données sans nécessiter de

programmation explicite, ce qui améliore leurs capacités prédictives. Cette approche marque un changement significatif par rapport à la programmation traditionnelle, où chaque instruction doit être explicitement définie pour l'ordinateur. L'apprentissage automatique, ou intelligence artificielle statistique, est appliqué dans divers domaines tels que la classification, la détection, la génération de contenu à partir de bases de données, la conception, l'analyse, l'optimisation, le développement et la mise en œuvre de méthodes.

Dans le domaine de l'hydrologie et de l'érosion, par exemple, l'apprentissage automatique présente des opportunités prometteuses. Il peut être utilisé pour modéliser et prédire les variations des schémas de précipitations, les débits des rivières et les processus d'érosion du sol. L'intégration de l'IA dans ces domaines permet d'améliorer la compréhension des processus hydrologiques, de détecter les anomalies dans les données collectées sur le terrain et d'optimiser les décisions relatives à la gestion des ressources en eau et à la prévention des risques d'érosion. Cependant, plusieurs défis persistent, notamment la nécessité de garantir la précision des prédictions dans des environnements naturels complexes et la compréhension limitée des mécanismes sous-jacents des modèles d'apprentissage automatique dans de tels contextes. De plus, la gestion de vastes ensembles de données spatiales et temporelles présente des défis en termes de dimensionnalité, nécessitant des approches innovantes pour assurer des analyses efficaces et significatives (Koedinger et Corbett, 2006 ; Kirat et al, 2022 ; Romero et al, 2023).

I.11. Conclusion

Cette recherche, axée sur la modélisation de l'érosion hydrique dans le bassin versant Isser en Algérie en utilisant des techniques de machine learning, revêt une importance cruciale pour la gestion durable des ressources naturelles. En offrant un outil de prise de décision précis et fiable, elle permet d'évaluer et de diminuer de façon proactive les risques d'érosion hydrique. Par conséquent, cette méthode joue un rôle important dans la préservation de l'environnement, la préservation des terres agricoles indispensables et la sécurité des communautés locales qui dépendent de ces ressources pour leur subsistance. Grâce à l'amélioration des compétences d'intervention et de planification, cette étude ouvre la voie à une gestion plus efficace et durable des bassins versants, ce qui favorise un équilibre harmonieux entre les activités humaines et la préservation des natures.

Chapitre II.

Cadre théorique et méthodologique...

II.1. Introduction

Ce chapitre établit les bases théoriques et méthodologiques de notre recherche. Nous décrivons d'abord la zone d'étude avec ses caractéristiques géographiques, hydrologiques et environnementales. Ensuite, nous détaillons la collecte et le prétraitement des données, en abordant les défis rencontrés et les solutions mises en place pour garantir la qualité des données. Enfin, nous présentons les méthodes d'apprentissage automatique choisies, en expliquant nos critères de sélection et les mesures d'évaluation des performances.

II.2. Description du bassin versant Isser

II.2.1. Situation géographique

Le bassin versant de l'Oued Isser, identifié par le code 09 dans le répertoire de l'ANRH, est situé dans le nord de l'Algérie. Ce territoire est délimité au nord-est par la daïra de Drâa El Mizane, au sud-est par la wilaya de Bouira, au sud par la daïra d'Ain Boucif, au sud-ouest par la daïra de Ksar El Boukhari et la wilaya de Médéa, au nord-ouest par les daïras de Tablat et Larbaâ, et au nord par la mer Méditerranée. La figure ci-dessous présente une vue d'ensemble de la zone d'étude ainsi que les particularités du bassin versant Isser. Ce bassin couvre une superficie de 4126 km² et se situe à environ 70 km au sud-est d'Alger. Sa forme est presque quadrilatérale, orientée sud-ouest/nord-est. Administrativement, il s'étend sur plusieurs wilayas : Médéa, Bouira, Tizi Ouzou et Boumerdès. Géographiquement, le bassin se développe dans un environnement comprenant l'Atlas tellien algérois au nord, avec une altitude de 1130 mètres au Djebel Tamesguida, et la chaîne des Bibans au sud, culminant à 1810 mètres au Djebel Dira. La plaine des Aribis, située à 550 mètres d'altitude, sépare ces deux chaînes. Les lignes de crête du bassin versant de l'Oued Isser varient entre 90 et 1810 mètres d'altitude, avec une altitude moyenne de 710 mètres.

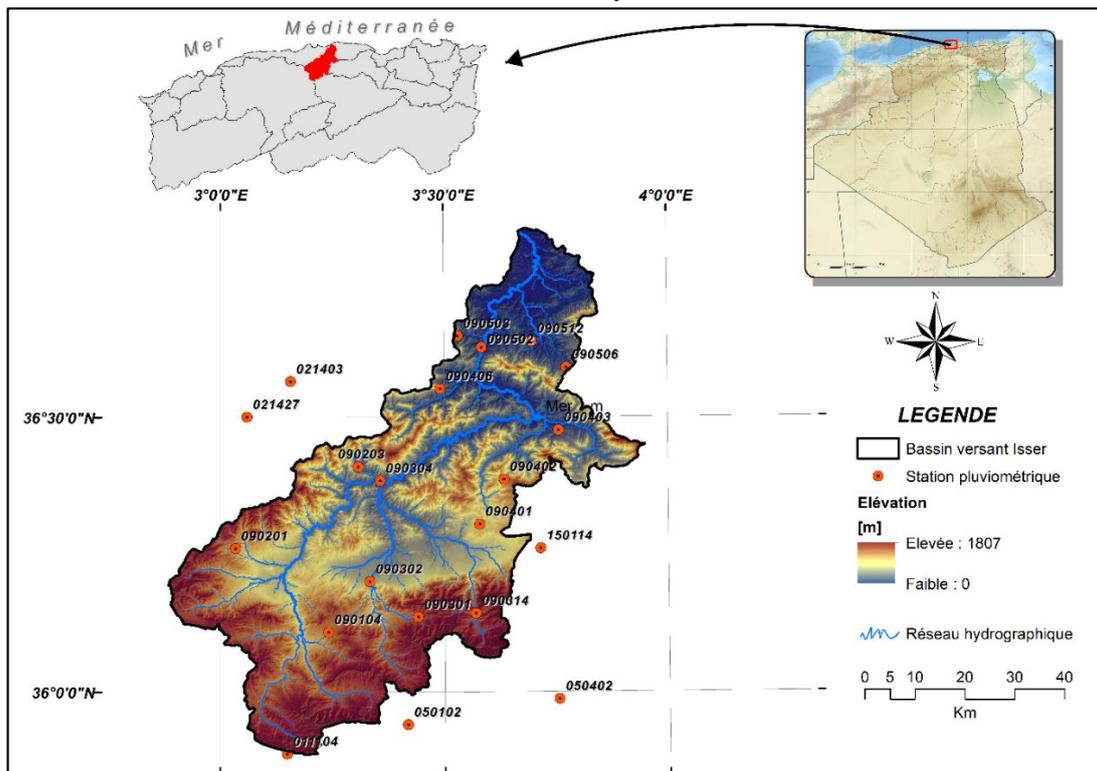


Figure II.1. Situation géographique du bassin versant Isser.

Le bassin versant Isser est divisé en cinq sous-bassins versants, chacun ayant une forme allongée.

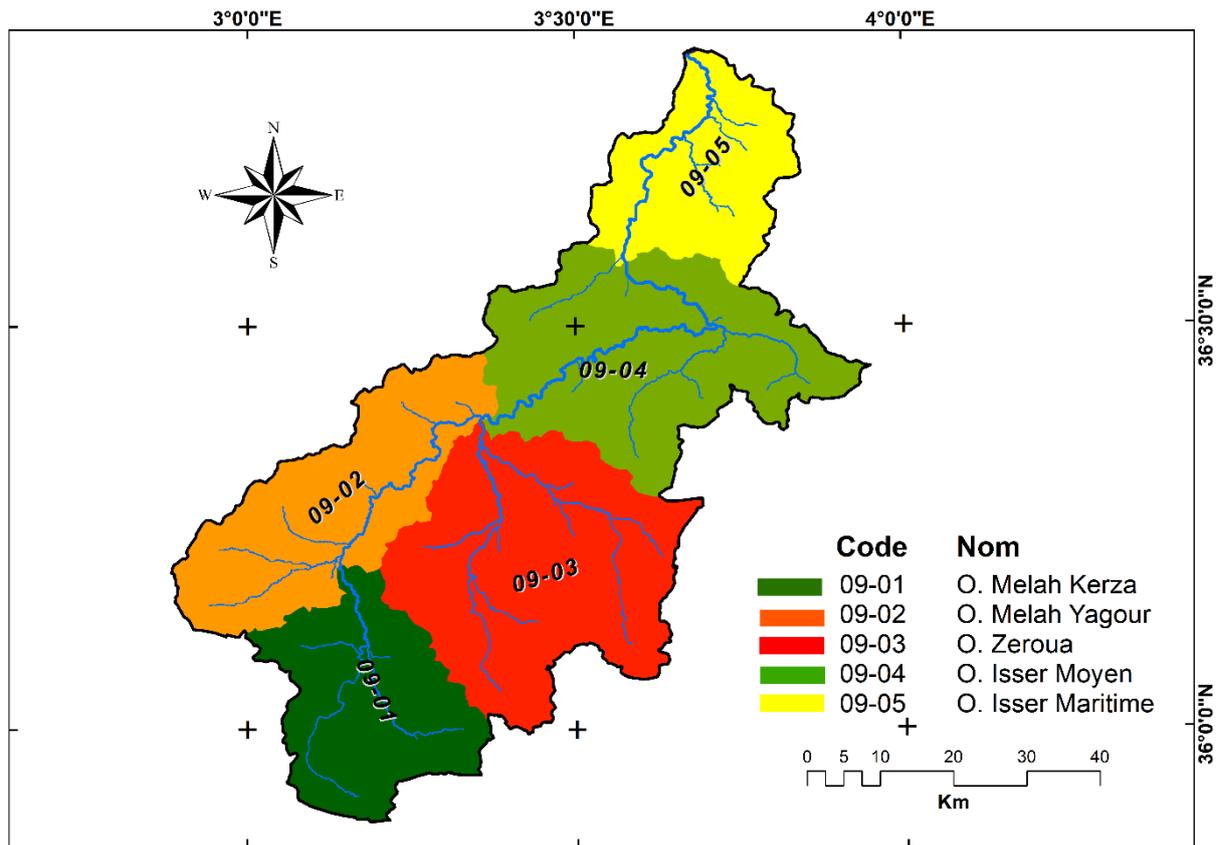


Figure II.2. Sous bassins du bassin versant Isser.

Voici les caractéristiques morphométriques de ces cinq sous-bassins présentées dans le tableau qui suit :

Tableau II.1. Caractéristiques morphométriques des cinq sous bassins.

	Paramètres		Sous bassin				
	Symbole	Unité	09-01	09-03	09-03	09-04	09-05
Caractéristiques de forme	S	km ²	665	73	273	2570	3615
	P	km	110	41	71	210	270
	Kc	-	1,19	1,1	1,1	1,16	1,26
	Rf	-	0,33	0,55	0,4	0,22	0,09
	Kf	-	0,40	0,66	0,43	0,36	0,18
	Re	-	0,65	0,84	0,72	0,52	0,34
	Rc	-	0,69	0,55	0,68	0,73	0,62
	L	km	37	9,55	24,2	66,1	98,2
Caractéristiques topographiques	Hmax	m	1467	1810	1371	1810	1810
	Hmin	m	595	810	600	349	90
	Hmoy	m	750	940	940	843	710
	Ig	‰	19,38	86,96	29,65	13,40	8,60
	Ip	‰	15	22	13,00	12,00	11,00

Hydrographiques	Im	‰	54	60	53	51	48,5
	Dd	km.km ⁻²	3,03	2,4	2,4	3,2	3,6
	Ct		19,39	17,5	17,5	23,98	29,5
	Lcp	km	45	11,5	26	109	200

(zouaoua mehdi,2021)

II.2.2. Lithologie du bassin versant Isser

Sur le plan géologique, le bassin versant Isser, est composé de 39 % de schiste calcaire et de 5 % d'argilite et d'air (Zeggane, 2017). Les schistes calcaires sont des roches datant du Crétacé moyen, du Crétacé supérieur et de l'Éocène. Elles sont très fragiles, tendres et érodables. Les eaux de ruissellement peuvent provoquer des ravinements sur les schistes, surtout lorsque leur pendage est faible (généralement entre 10 et 20°). Les argilites et les argiles sont des formations du Miocène et du Pliocène inférieur. Elles sont tendres, avec des pentes douces et convexes selon des gradients allant de 5° à 20°. L'eau de ruissellement entraîne rapidement le ravinement de ces formations.

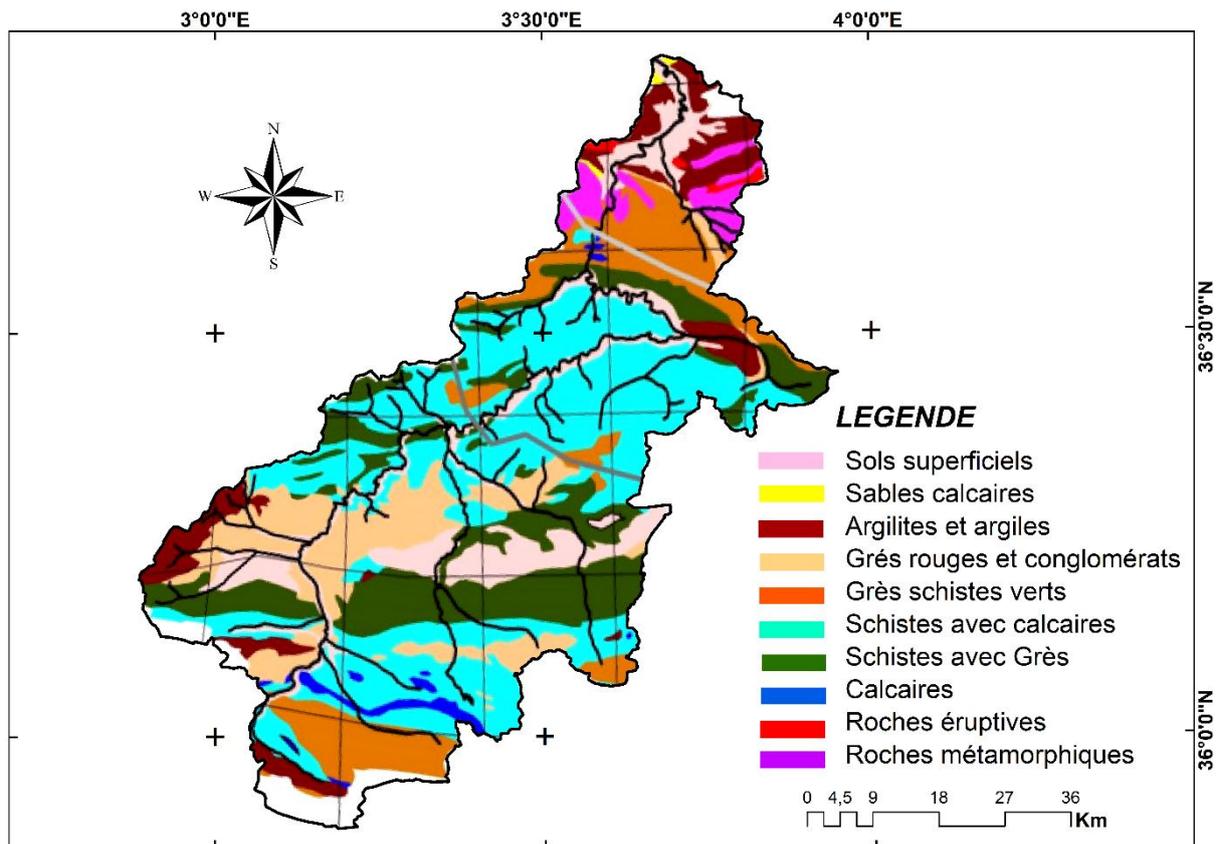


Figure II.3. Lithologie du bassin versant Isser.

II.2.3. Paramètre de relief

Le relief est un facteur déterminant, car il influence largement la capacité des terrains à permettre le ruissellement, l'infiltration et l'évaporation de l'eau. Il joue un rôle essentiel dans le fonctionnement hydrologique d'un bassin. La majorité des paramètres météorologiques (comme la pluie et la température) et hydrologiques dépendent de l'altitude. Les altitudes dans le bassin varient de 100 mètres au niveau de l'embouchure à plus de 1800 mètres dans les zones montagneuses, créant des dénivelés importants. Ces caractéristiques soulignent la complexité et la richesse hydrologique du bassin versant Isser, nécessitant une gestion intégrée et durable.

La courbe hypsométrique offre une représentation synthétique de la pente du bassin, c'est-à-dire du relief. Elle illustre la répartition de la surface du bassin versant en fonction de l'altitude et permet d'évaluer l'état du bassin en termes de potentiel d'érosion.

Tableau II.2. La répartition altimétrique dans le bassin Isser.

Classe d'altitude (m)	Moyenne (m)	Surface partielle		
		km ²	%	Cumulée (%)
0-200	100	2,14	0,05	0,05
200-400	300	6,63	0,16	0,21
400-600	500	86,40	2,08	2,29
600-800	700	450,10	10,83	13,12
800-1000	900	893,06	21,48	34,60
1000-1200	1100	1243,41	29,91	64,51
1200-1400	1300	734,61	17,67	82,18
1400-1600	1500	395,01	9,50	91,68
1600-1807	1703	345,77	8,32	100,00

D'après le tableau II.2, la courbe hypsométrique de notre bassin versant peut être observée, comme illustré dans la figure suivant :

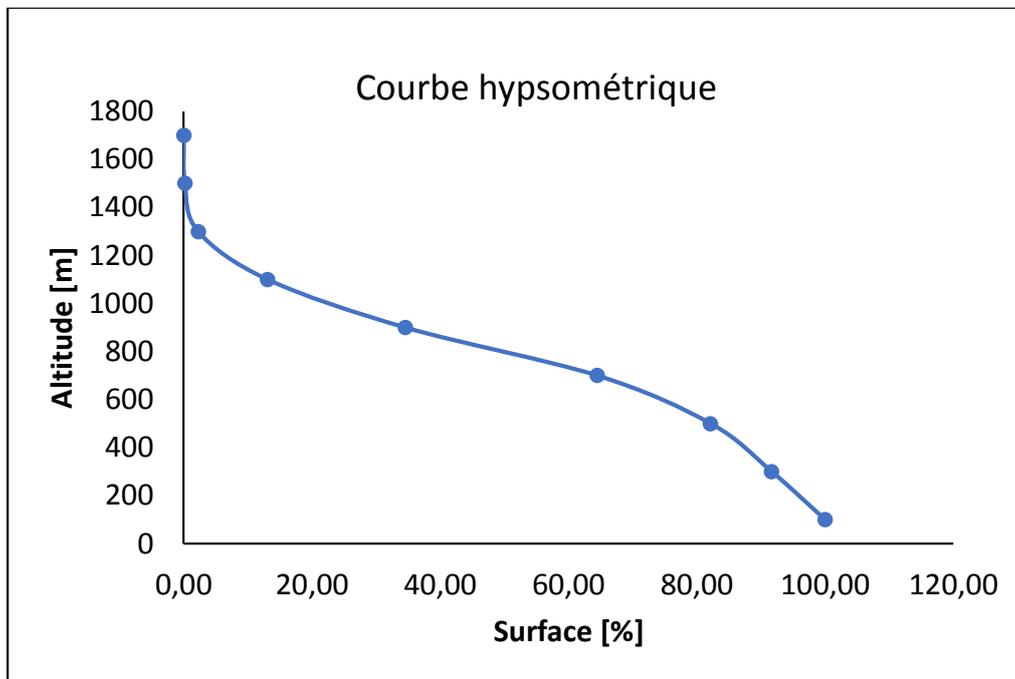


Figure II.4. Courbe hypsométrique du bassin versant Isser.

II.2.4. Topographie du bassin versant Isser

La topographie générale du bassin versant de l'oued Isser se distingue par un relief très escarpé et fortement découpé, présentant des variations importantes de pentes selon les différentes zones.

Au nord du bassin, les pentes sont modérées, variant entre 10 % et 40 %. La zone médiane, quant à elle, présente des pentes relativement douces, ne dépassant pas 10%. En contraste, la partie sud du bassin est caractérisée par des pentes extrêmement abruptes, pouvant atteindre jusqu'à 67°.

L'indice global de pente du bassin versant, évalué à 13,61 m/km, reflète une topographie particulièrement accidentée. Le réseau hydrographique du bassin est dense et présente une forme dendritique générale, ce qui favorise des crues importantes durant les saisons hivernale et printanière. Ces crues provoquent une érosion significative au sein du bassin versant.

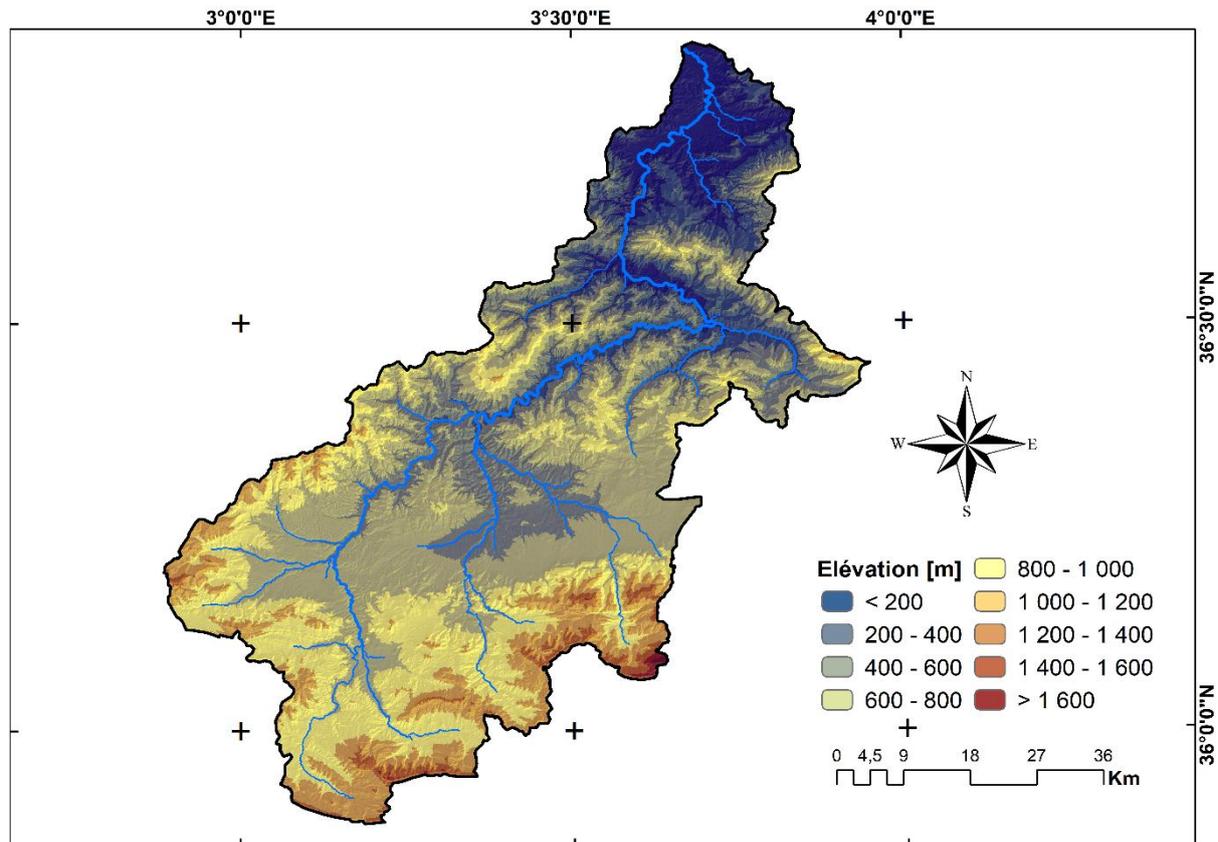


Figure II.5. Distribution des altitudes dans le bassin Isser.

II.2.5. Les caractéristiques hydrologiques du bassin versant Isser

Les caractéristiques hydrologiques du bassin versant Isser englobent divers aspects relatifs à l'écoulement et à la distribution de l'eau, essentiels pour la gestion des ressources hydriques. Les principaux éléments hydrologiques sont détaillés ci-dessous.

II.2.5.1. Réseau hydrographique et principaux cours d'eau

Le bassin versant Isser est traversé par un réseau hydrographique complexe comprenant des rivières, des oueds et de nombreux affluents qui drainent les eaux de surface. Le réseau hydrographique présente une forme générale dendritique avec une densité de drainage élevée. Deux barrages principaux se trouvent dans ce bassin : le barrage de Koudiat Acerdoun et le barrage de Beni Amrane.

Les principaux cours d'eau du bassin versant Isser sont :

- **L'Oued Isser** : Le cours d'eau principal, d'une longueur de 77,2 km, qui se jette dans la Méditerranée et sert d'exutoire pour le bassin ;
- **L'Oued Soufflat** : Un affluent de 26,2 km ;
- **L'Oued Bou Hamoud** : Un autre affluent de 18,3 km.

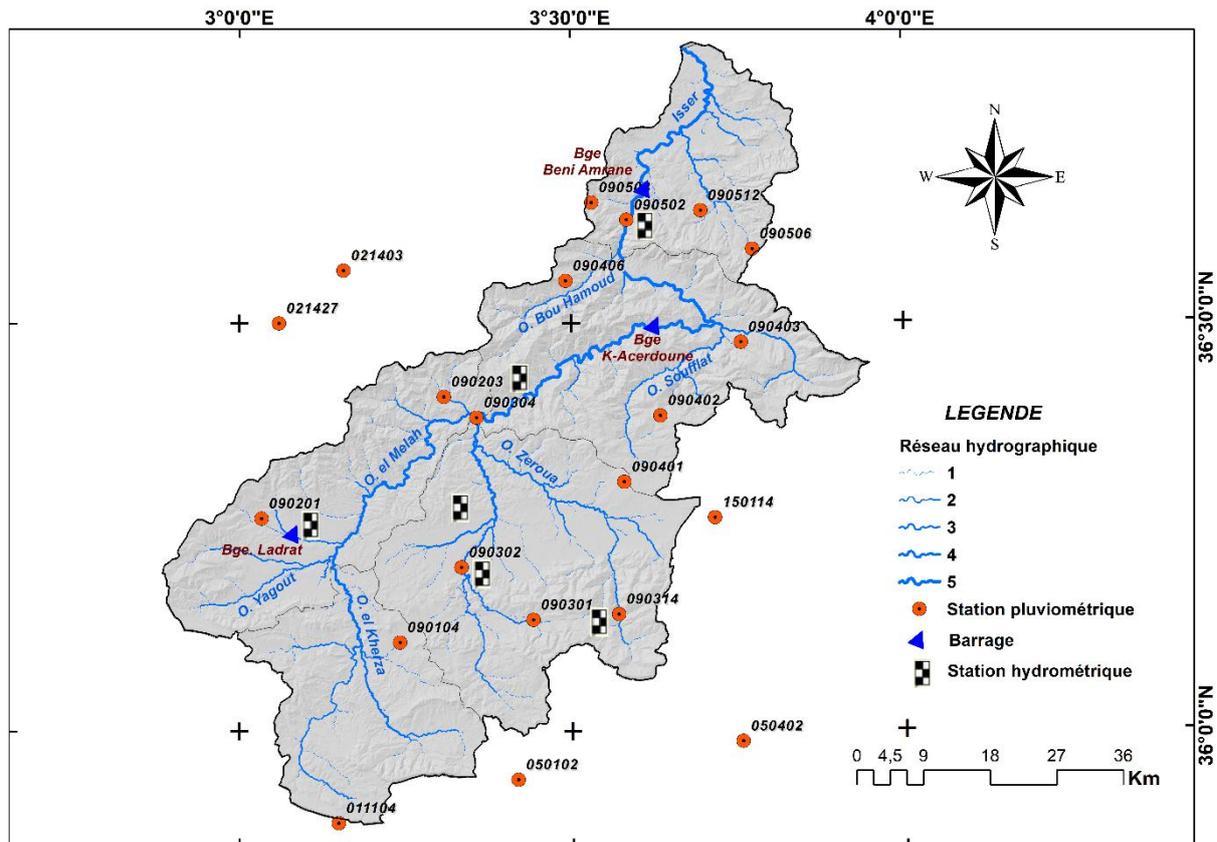


Figure II.6. Réseau hydrographique du bassin versant Isser.

II.2.5.2. Régime hydrologique et variations saisonnières

Le régime hydrologique décrit les variations saisonnières du débit d'eau, marquées par des différences entre les niveaux d'eau bas et hauts. Différents types de régimes hydrologiques existent :

- ✦ **Monomodal** : Un seul maximum et minimum de débit par an.
- ✦ **Polymodal** : Plusieurs maximums et minimums de débit par an.
- ✦ **Nival de transition** : Un régime influencé par la fonte des neiges.
- ✦ **Pluvial tropical** : Un régime influencé par les précipitations tropicales.

Ces variations saisonnières influencent significativement la biologie des espèces aquatiques et la dynamique des populations de poissons en raison de la répartition saisonnière des précipitations et de la variabilité climatique.

II.2.5.3. Apports en eau et débits moyens

Le bassin versant Isser bénéficie d'un apport moyen annuel en eau estimé à 14 millions de mètres cubes. Ce volume se traduit par un débit moyen du fleuve Isser à son exutoire de 3,8 mètres cubes par seconde. Ces données soulignent l'abondance hydrologique du bassin, constituant une ressource précieuse pour les divers usages de l'eau dans la région ([Keddar et al, 2011](#)).

II.2.5.4. Qualité des eaux de surface

Les eaux de surface de l'oued Isser montrent des paramètres physico-chimiques stables, incluant la température, le pH et la conductivité. Cependant, les activités agricoles et domestiques contribuent parfois à des niveaux élevés de nutriments, tels que les nitrates et les phosphates. Les analyses révèlent également la présence de micropolluants, comme des métaux lourds et des résidus de pesticides, provenant des déchets industriels et agricoles (*Djafer Khodja, 2020*).

II.2.5.5. Eaux souterraines et aquifères

Le bassin versant Isser possède des ressources en eaux souterraines importantes, avec plusieurs aquifères majeurs. Les nappes phréatiques sont alimentées par l'infiltration des eaux de pluie et des cours d'eau, jouant un rôle crucial pour l'approvisionnement en eau potable et l'irrigation agricole. Les aquifères principaux se trouvent dans les formations calcaires et alluviales, avec des profondeurs et débits variables selon la géologie locale. Des études hydrogéologiques détaillées ont permis de mieux caractériser ces réserves en eau souterraine et leurs potentiels d'exploitation durable.

II.2.6. Impacts des activités humaines sur l'hydrologie

Les activités humaines exercent une pression croissante sur les ressources en eau du bassin versant Isser. Les prélèvements excessifs, la pollution des eaux et les aménagements hydrauliques perturbent gravement l'équilibre hydrologique naturel, fragilisant les écosystèmes aquatiques et compromettant la disponibilité en eau. Une gestion durable s'impose pour préserver ces ressources vitales.

II.3. Collecte des données

II.3.1. Sources de données géospatiales et méthodes d'acquisition

Dans cette section, nous offrons une analyse détaillée de l'environnement des Systèmes d'Information Géographique (SIG) ainsi que des techniques de prétraitement des données géospatiales que nous avons employées dans notre recherche (Figure II.7). L'utilisation des SIG nous a permis de manipuler et de traiter efficacement les données collectées, qui comprenaient une gamme variée de facteurs climatologiques, morphologiques, hydrologiques et géologiques. L'intégration de ces multiples facteurs dans le cadre des SIG a grandement facilité la production de cartes thématiques, offrant ainsi une représentation visuelle précise de leur répartition spatiale à travers la zone d'étude.

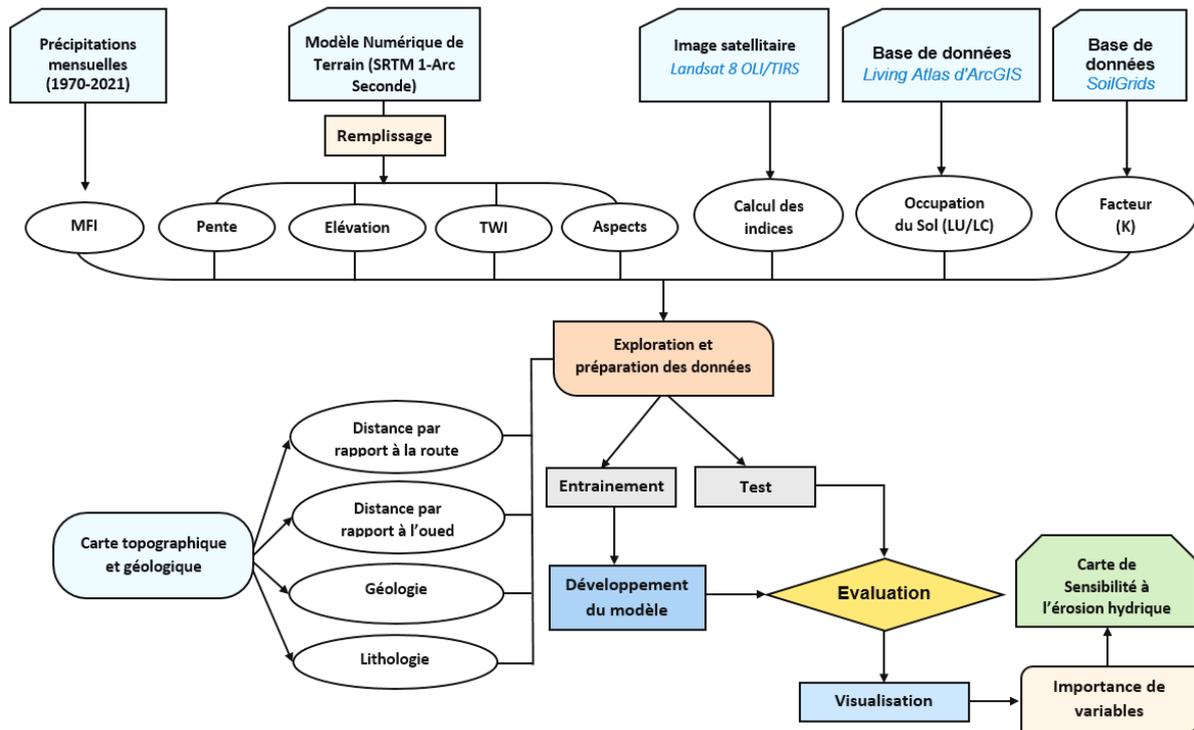


Figure II.7. Organigramme de la méthodologie.

II.3.1.1. Indice de Fournier Modifié (Modified Fournier Index : MFI)

La capacité des gouttes de pluie à provoquer l'érosion est désignée par l'érosivité de pluie. Pour évaluer cette érosivité, il est nécessaire de mesurer l'intensité et la durée des précipitations (Irvem et al, 2007). L'indice de Fournier Modifié (MFI), proposé par Arnoldus en 1977, est couramment utilisé pour estimer l'érosivité de pluie. Cet indice se base sur les précipitations mensuelles (p_i) et les précipitations annuelles totales (P_{total}) :

$$MFI = \sum_{i=1}^{i=12} \frac{p_i^2}{P_{total}} \quad \dots(1)$$

Le résultat de l'équation (1) peut être classifié comme présenté dans le tableau II.3. Sur cette base, le MFI aura des valeurs élevées là où les valeurs des précipitations sont élevées. En ce sens, les régions avec de grandes quantités de précipitations annuelles totales et une concentration des précipitations auront une valeur de MFI élevée (De Luis et al, 2010). Globalement, le calcul de MFI pourrait fournir une estimation réaliste du potentiel du facteur d'érosivité des précipitations (Apaydin et al, 2006 ; Harsányi et al, 2022).

Valeur de MFI	Risque érosif
<60	Très faible
60-90	Faible
90-120	Modéré
120-160	Elevée
>160	Très élevé

Tableau II.3. Description de la plage d'érosivité de MFI (Fenta et al, 2017)

Pour notre région d'étude, l'érosivité des précipitations a été évaluée en utilisant une série de données mensuelles couvrant une période de 51 ans, de 1970 à 2021. Ces données ont été recueillies à partir de 15 stations pluviométriques réparties dans la zone d'étude. Chaque station a fourni des enregistrements détaillés des précipitations mensuelles, permettant ainsi d'analyser les variations de l'érosivité sur une période prolongée et de mieux comprendre l'impact des précipitations sur l'érosion des sols dans la région. Les données de précipitations nécessaires pour le calcul de ce paramètre ont été obtenues auprès de l'Agence Nationale des Ressources Hydrauliques (ANRH). Le tableau suivant récapitule les détails des stations utilisées :

Tableau II.4. Stations pluviométriques utilisées.

Station	Code	Altitude (m)	Latitude (DD)	Longitude (DD)	MFI	Pmoy annuelle (mm)	Période de mesures
AIN BOUCIF	011104	1250	35,89	3,15	36,03	315,50	1970-2021
MEDEA SECTEUR	011509	935	36,27	2,75	74,20	646,20	1977-2020
LARBAA	021403	100	36,57	3,16	76,46	652,01	1970-2021
DOUAR MAKLI	021427	130	36,50	3,06	76,81	638,23	1979-2021
BOGHNI	021903	250	36,54	3,95	81,09	712,80	1970-2012
CHELLALET EL ADAOURA	050102	990	35,94	3,42	26,99	239,58	1970-2021
DIRAH CENTRE	050402	800	35,99	3,76	30,63	307,20	1970-2021
SOUAGUI	090104	810	36,11	3,24	34,48	326,48	1974-2019
EL OMARIA	090201	790	36,26	3,03	53,00	439,01	1970-2020
TABLAT DRS (A.N.R.H)	090203	450	36,41	3,31	68,55	593,36	1970-2021
DJOUAB	090301	825	36,14	3,44	54,59	494,03	1970-2021
BENI SLIMANE	090302	600	36,20	3,33	38,76	384,43	1970-2021
PONT DE LA TRAILLE	090304	370	36,38	3,36	53,21	474,31	1970-2021
DECHMYA	090314	850	36,14	3,57	54,64	502,73	1975-2021
KHABOUZIA	090401	720	36,31	3,58	50,64	452,87	1969-2021
SOUK EL KHEMIS	090402	782	36,39	3,63	63,37	477,37	1970-2021
DJEBAHIA	090403	320	36,48	3,76	71,35	548,52	1970-2021
TAZEROUT	090406	520	36,55	3,49	93,59	795,22	1970-2018
LAKHDARIA GORGES	090502	50	36,63	3,58	85,09	726,83	1970-2018
OULED BOUHADDADA	090503	400	36,65	3,53	90,83	767,63	1970-2018
TIZIGHENIF	090506	390	36,59	3,77	93,44	657,44	1970-2018
CHABET EL AMEUR	090512	235	36,64	3,70	84,80	668,01	1970-2018
AIN TASTA	150114	690	36,26	3,71	45,73	427,45	1970-2014
CRETE ROUGE	150115	410	36,30	4,12	43,96	340,99	1975-2021

DD : Degré Décimaux

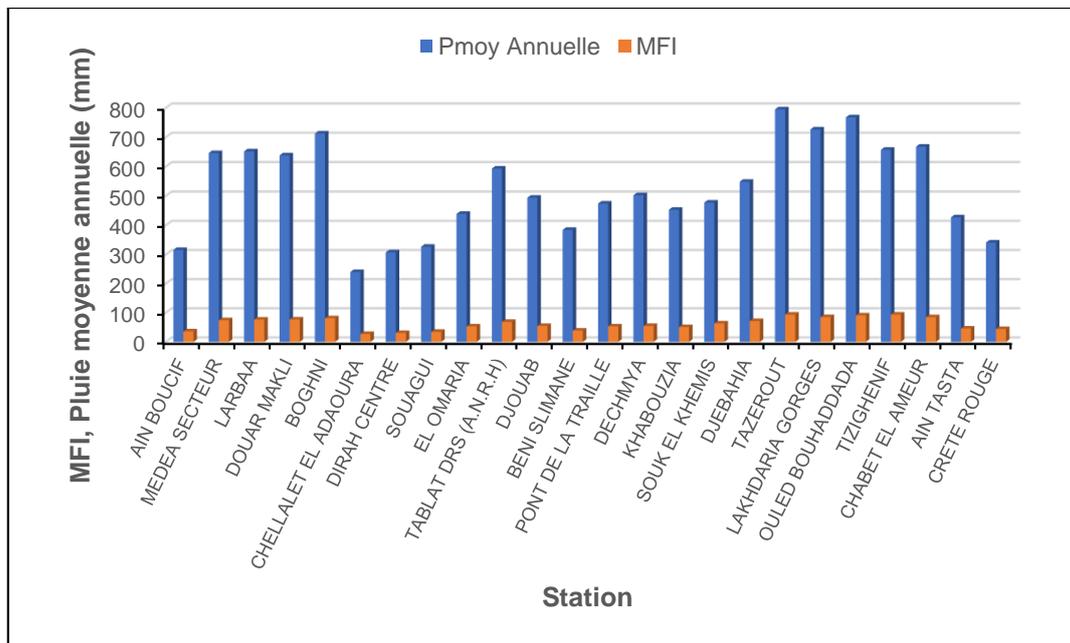


Figure II.8. Distribution des pluies moyennes annuelles et MFI.

II.3.1.2. Pente, Elévation, Aspects et TWI

La pente joue un rôle crucial en accélérant ou en ralentissant le ruissellement, ce qui impacte directement l'énergie cinétique de l'eau et sa capacité à éroder le sol. À des pentes raides, le ruissellement est rapide, favorisant une érosion accrue, tandis qu'à des pentes douces, l'infiltration est privilégiée, réduisant l'érosion de surface. L'élévation modifie le climat local et la végétation, affectant indirectement l'érosion : les zones élevées, souvent soumises à des précipitations abondantes et à des fontes de neige, présentent un risque accru d'érosion. L'aspect des pentes, orienté par rapport aux points cardinaux, influence les microclimats locaux et donc la sécheresse du sol ou l'intensité du ruissellement. Enfin, l'indice d'humidité topographique (Topographic Wetness Index : TWI) différencie les zones où l'eau stagne de celles où elle s'écoule rapidement, les premières étant plus sujettes à l'érosion en affaiblissant la structure du sol lors des ruissellements.

Notre étude utilise des données du Modèle Numérique de Terrain SRTM-1 arc seconde à 30 mètres de résolution, traitées avec ArcGIS pour évaluer la pente, l'élévation, l'aspect et le TWI. Ces analyses fournissent des informations précises sur la vitesse d'écoulement de l'eau, la répartition du ruissellement et les risques d'érosion associés à la saturation du sol et au potentiel de ruissellement, essentielles pour la gestion et la préservation des sols.

II.3.1.3. BSI, NDBAI et NDBI

Pour notre étude, nous avons utilisé plusieurs indices dérivés des images Landsat 8, calculés à l'aide du logiciel ArcGIS. Ces indices comprennent le BSI (Bare Soil Index), utilisé pour détecter les zones de sol nu ou non couvertes par la végétation, ainsi que le NDBAI (Normalized Difference Built-Up Index) et le NDBI (Normalized Difference Built-Up Index). Le NDBAI permet de différencier les zones construites des autres types de terrains, tandis que le NDBI évalue la densité urbaine en détectant les zones bâties. Ces indices sont cruciaux pour notre analyse, fournissant des informations détaillées sur la répartition spatiale et les caractéristiques des zones urbaines étudiées. Le tableau II.5 récapitulatif des formules de calcul pour les indices BSI, NDBAI et NDBI directement à partir des bandes des images Landsat 8.

II.3.1.4. NDVI, NDWI et NBR

Les indices de végétation et de qualité de l'eau obtenus à partir des images Landsat 8 dans ArcMap (Tableau II.5) offrent des perspectives cruciales pour l'analyse environnementale et la gestion des ressources naturelles. Le NDVI (Normalized Difference Vegetation Index) est utilisé pour évaluer la densité et l'activité photosynthétique de la végétation en mesurant la différence normalisée entre les réflectances dans les bandes rouge et infrarouge proche. De son côté, le NDWI (Normalized Difference Water Index) permet de détecter la présence d'eau en comparant les réflectances dans les bandes verte et infrarouge. Enfin, le NBR (Normalized Burn Ratio) évalue les impacts des incendies sur la végétation en mesurant la différence normalisée entre les réflectances dans les bandes infrarouge moyen et infrarouge proche. Ces indices jouent un rôle crucial dans la surveillance des écosystèmes terrestres et aquatiques, fournissant des données précieuses pour prendre des décisions informées en matière de conservation et de gestion des ressources naturelles.

Tableau II.5. Formules de calcul des indices BSI, NDBAI et NDBI.

Indice	Formule de calcul
BSI	$BSI = \frac{(Band5 - Band4)}{(Band5 + Band4)}$
NDBAI	$NDBAI = \frac{(Band5 - Band2)}{(Band5 + Band2)}$
NDBI	$NDBI = \frac{(Band5 - Band6)}{(Band5 + Band6)}$
NDVI	$NDBI = \frac{(Band5 - Band4)}{(Band5 + Band4)}$
NDWI	$NDBI = \frac{(Band3 - Band5)}{(Band3 + Band5)}$
NBR	$NDBI = \frac{(Band5 - Band7)}{(Band5 + Band7)}$

II.3.1.5. Facteur d'érodibilité (K)

Le facteur d'érodibilité des sols K est une mesure clé de la susceptibilité d'un sol à l'érosion sous l'effet des précipitations et du ruissellement. Il dépend de plusieurs propriétés du sol, telles que la texture, la structure, la perméabilité et la teneur en matière organique. Par exemple, les sols limoneux ont un facteur K élevé en raison de leur facilité à être détachés et transportés par l'eau, tandis que les sols sableux, qui favorisent l'infiltration de l'eau, ont un facteur K plus faible. Une bonne structure du sol et une teneur élevée en matière organique réduisent également l'érodibilité. Dans l'Équation Universelle de Perte de Sol (USLE), le facteur K est utilisé pour estimer la perte de sol annuelle moyenne en conjonction avec d'autres facteurs tels que l'érosivité des précipitations, la topographie, la couverture végétale et les pratiques de gestion.

Ainsi, le facteur « K » influence directement le taux d'érosion, soulignant l'importance de la gestion des propriétés du sol pour prévenir l'érosion et conserver la qualité des sols.

Dans cette étude, la valeur de référence du facteur K est obtenue à partir de la base de données SoilGrids (<https://soilgrids.org/>), qui fournit les pourcentages de sable, de limon, d'argile et de carbone organique à différentes profondeurs. Pour notre cas, une profondeur superficielle inférieure à 30 cm a été choisie. Ensuite, le facteur d'érodabilité du sol, K, est évalué pour la couche arable (0-30 cm) en utilisant l'équation suivante (Sharpley, A.N., et J.R. Williams, éd. 1990) :

$$K = A * B * C * D$$

Avec:

$$A = \left[0.2 + 0.3 \exp(-0.0256 SAN) \left(1 - \frac{SIL}{100} \right) \right]$$

$$B = \left(\frac{SIL}{CLA + SIL} \right)^{0.3}$$

$$C = \left(1.0 - \frac{0.25 * C}{C + \exp(3.72 - 2.95 * C)} \right)$$

$$D = \left(1.0 - \frac{0.7 * SN1}{SN1 + \exp(-5.51 + 22.9 * SN1)} \right)$$

$$SN1 = 1 - SAN/100$$

Où *SAN*, *SIL*, *CLA* et *C* sont respectivement les pourcentages de sable, limon, argile et carbone organique.

II.3.1.6. Occupation des sols

L'occupation du sol joue un rôle crucial dans l'évaluation et la gestion de l'érosion hydrique, car la manière dont les terres sont utilisées influence directement le risque d'érosion. Une couverture végétale dense, par exemple, protège le sol en réduisant le ruissellement et en augmentant l'infiltration de l'eau, tandis que des terres agricoles mal gérées ou des zones urbaines imperméabilisées augmentent le risque d'érosion en favorisant le ruissellement. Les pratiques de conservation des sols, telles que les terrasses et les bandes enherbées, ainsi que la planification urbaine et rurale, sont essentielles pour atténuer l'érosion.

La carte d'occupation du sol de notre zone d'étude est extraite de la base de données disponible sur Living Atlas d'ArcGIS (livingatlas.arcgis.com). Cette ressource fournit des informations détaillées et actualisées sur l'occupation des sols, essentielles pour évaluer l'érosion hydrique. L'utilisation de ces données permet de comprendre les types d'occupation du sol, tels que les forêts, les terres agricoles, les zones urbaines, et leur distribution spatiale. Ces informations sont intégrées dans les

modèles prédictifs pour élaborer des cartes de sensibilité à l'érosion et développer des stratégies de gestion et de conservation des sols.

II.3.1.7. Distance par rapport aux routes/oueds

Les zones proches des routes, en particulier celles non pavées, et des oueds sont souvent sujettes à des taux d'érosion plus élevés en raison de la perturbation du sol, de la réduction de la couverture végétale, des modifications des schémas de drainage et des fluctuations des débits d'eau. La végétation, agissant comme une barrière naturelle, est essentielle pour stabiliser le sol et réduire l'érosion, mais son absence près des routes et des oueds augmente le ruissellement et l'érosion, exacerbée par le compactage du sol et les pentes raides. Les routes modifient les schémas de drainage naturels, concentrant l'écoulement de l'eau et augmentant l'érosion, tandis que les oueds, particulièrement en période de crues, déstabilisent le sol avec l'impact direct et intense de l'eau courante. Plus loin des routes et des oueds, les taux d'érosion diminuent grâce à une couverture végétale plus stable, des schémas de drainage naturels et une réduction de l'influence des crues. Des mesures telles que la plantation de bandes tampons végétatives, la conception adéquate du drainage et les techniques de stabilisation du sol peuvent aider à atténuer l'érosion dans ces zones.

Pour obtenir ces cartes, nous avons initié la création de l'hydrographie en utilisant le Modèle Numérique de Terrain (MNT) comme source de données de base. Ensuite, dans ArcMap, nous avons employé l'outil Euclidean Distance pour calculer et produire la carte correspondante. Concernant la distance par rapport aux routes, nous avons récupéré la carte des routes de la zone d'étude à partir du site OpenStreetMap (www.openstreetmap.org) et nous avons appliqué également le même outil dans ArcMap pour générer une carte indiquant cette distance spécifique par rapport aux routes.

Tableau II.6. Sources de données géospatiales.

Paramètre	Source	Lien	Résolution spatiale	Périodes temporelles
MFI	(ANRH)	-	Pas mensuel	(1970-2021)
Classe des sols	Grilles de sol	https://soilgrids.org/	160 m	2006
Occupation des sols	Esri Sentinel-2	https://livingatlas.arcgis.com/landcover/	10	2021
MNT	USGS	https://earthexplorer.usgs.gov/	1 arc-seconde	2014
Carte Géologique	Institut National de Cartographie	-	1/50 000	-
Image satellitaire	Landsat 8-OLI/TIRS	https://earthexplorer.usgs.gov/	30 m	05/2024

II.4. Méthodes d'apprentissage automatique

II.4.1. Définition et contexte théorique

Les algorithmes se définissent comme des séries d'instructions structurées permettant de résoudre des problèmes par l'exécution d'un nombre fini d'opérations. En informatique, ils sont essentiels à la création de programmes capables de réaliser des tâches spécifiques de manière efficace. Fondamentalement, un algorithme peut être comparé à un plan détaillé, composé d'étapes précises ordonnées de manière à atteindre un résultat défini. Ce concept trouve son utilisation non seulement dans le domaine de la programmation, mais aussi en mathématiques, où il désigne des méthodes systématiques pour la résolution de problèmes.

Sur le plan théorique, les algorithmes de Machine Learning sont classés en trois principales catégories : supervisés, non-supervisés et semi-supervisés. Chaque type repose sur une approche distincte d'apprentissage permettant de traiter des ensembles de données complexes et de découvrir des modèles utiles. En utilisant ces algorithmes, il devient possible d'automatiser des processus décisionnels, de prédire des comportements futurs, ou encore de découvrir des structures cachées dans les données, ouvrant ainsi la voie à de nombreuses applications pratiques dans des domaines variés tels que la reconnaissance vocale, la recommandation de produits, la classification d'images, et bien d'autres encore (Nemmich, 2014).

II.4.2. Types de méthodes d'apprentissage automatique

L'apprentissage automatique (ML) est un domaine vaste avec plusieurs méthodes pour entraîner des modèles à partir de données. Ces méthodes peuvent être classées en différentes catégories selon leur approche et les types de problèmes qu'elles résolvent. Voici les principales catégories de méthodes d'apprentissage automatique :

II.4.2.1. Apprentissage supervisé

Dans l'apprentissage supervisé, le modèle est entraîné sur un ensemble de données étiquetées. Chaque exemple d'entraînement est composé d'une entrée et d'une sortie souhaitée. L'objectif est d'apprendre une fonction qui mappe les entrées aux sorties.

- **Régression** : Prédit une valeur continue.
 - Exemples : Régression linéaire, régression polynomiale, régression Ridge.
- **Classification** : Prédit une catégorie ou classe.
 - Exemples : Régression logistique, k-plus proches voisins (k-NN), arbres de décision, forêts aléatoires, machines à vecteurs de support (SVM), réseaux de neurones.

II.4.2.2. Apprentissage non supervisé

Dans l'apprentissage non supervisé, le modèle est entraîné sur des données sans étiquettes. L'objectif est de découvrir des structures ou des motifs cachés dans les données.

- **Clustering (Regroupement)** : Groupe les données en clusters ou groupes.
 - Exemples : k-means, algorithme des attentes-maximisation (EM), clustering hiérarchique, clustering spectral.
- **Réduction de dimensionnalité** : Réduit le nombre de variables tout en préservant autant que possible la variance des données.
 - Exemples : Analyse en composantes principales (PCA), analyse en composantes indépendantes (ICA), t-SNE, UMAP.
- **Apprentissage d'associations** : Trouve des relations intéressantes entre les variables.
 - Exemples : Algorithme apriori, FP-growth.

II.4.2.3. Apprentissage semi-supervisé

L'apprentissage semi-supervisé utilise à la fois des données étiquetées et non étiquetées pour l'entraînement. Cela est utile lorsque l'étiquetage des données est coûteux ou difficile.

- **Méthodes** : Autoencodeurs, méthodes de co-entraînement, générative adversarial networks (GANs) avec semi-supervision.

II.4.2.4. Apprentissage par renforcement

Dans l'apprentissage par renforcement, un agent apprend en interagissant avec un environnement et en recevant des récompenses ou des punitions en fonction de ses actions. L'objectif est de maximiser la récompense cumulée.

- **Méthodes** : Q-learning, Deep Q-Network (DQN), politique de gradient, acteur-critique, Monte Carlo Tree Search (MCTS).

II.4.2.5. Apprentissage profond

L'apprentissage profond est une sous-catégorie de l'apprentissage supervisé et non supervisé utilisant des réseaux de neurones artificiels avec plusieurs couches (profondes).

- **Réseaux de neurones convolutifs (CNN)** : Utilisés principalement pour les données image.
- **Réseaux de neurones récurrents (RNN)** : Utilisés pour les données séquentielles comme le texte et le langage.
- **Transformateurs** : Utilisés pour le traitement du langage naturel (NLP) et les tâches séquentielles.

II.4.2.6. Méthodes de transfert d'apprentissage

Le transfert d'apprentissage utilise un modèle pré-entraîné sur une grande quantité de données et l'adapte à une tâche spécifique avec peu de données.

- **Exemples** : Fine-tuning des modèles comme BERT, GPT, VGG, ResNet.

Chaque catégorie de méthodes a ses propres avantages et inconvénients et est choisie en fonction du problème spécifique à résoudre et des données disponibles.

II.4.3. Algorithmes spécifiques

✓ **Categorical Boosting (CatBoost)**

CatBoost, abréviation de "Categorical Boosting", est un algorithme de gradient boosting développé par Yandex, conçu pour améliorer les performances sur les données catégorielles tout en réduisant le temps d'entraînement et en prévenant l'overfitting. Il est particulièrement efficace pour les ensembles de données contenant des caractéristiques catégorielles, couramment trouvés dans des domaines comme le marketing, la finance, et les sciences sociales.

✓ **Extreme Gradient Boosting (XGBoost)**

XGBoost, est un algorithme de machine learning basé sur la technique de boosting par gradient. Développé par Tianqi Chen, XGBoost est connu pour sa performance exceptionnelle, sa vitesse, et son efficacité dans la compétition de machine learning. Il est largement utilisé pour les tâches de classification et de régression dans divers domaines tels que la finance, le marketing, et la bioinformatique.

✓ **Random Forest (RF)**

Random Forest (RF) est une méthode d'apprentissage automatique qui combine plusieurs arbres de décision pour des tâches de classification ou de régression. Chaque arbre est construit sur un sous-ensemble aléatoire des données d'entraînement, avec une sélection aléatoire des caractéristiques à chaque division de l'arbre. RF utilise ensuite un processus de "bagging" pour agréger les prédictions des différents arbres, offrant ainsi une prédiction finale robuste et moins sensible au surajustement par rapport aux arbres de décision individuels. Grâce à sa capacité à gérer efficacement de grands ensembles de données et à fournir une mesure d'importance des caractéristiques, Random Forest est largement utilisé dans divers domaines pour ses performances fiables et sa capacité à traiter des problèmes complexes en apprentissage automatique.

✓ **LightGBM**

LightGBM (Light Gradient Boosting Machine) est à la fois une bibliothèque et un algorithme d'apprentissage automatique. C'est un algorithme de gradient boosting qui est implémenté dans la bibliothèque LightGBM développée par Microsoft. Il utilise le framework de boosting par gradient pour construire des ensembles de modèles prédictifs, généralement des arbres de décision, de manière séquentielle. Contrairement à d'autres implémentations de gradient boosting comme XGBoost qui utilisent une croissance de type niveau par niveau, LightGBM utilise une stratégie de croissance de type feuille par feuille. Cela permet une optimisation plus efficace et une meilleure performance, notamment sur de grands ensembles de données.

Tableau II.7. Caractéristiques des algorithmes choisis.

Algorithme	Caractéristiques Clés	Avantages	Inconvénients	Applications
XGBoost Classifier	<ul style="list-style-type: none"> - Boosting par Gradient - Optimisation de la vitesse et de l'efficacité - Gestion des données manquantes - Régularisation - Importance des variables 	<ul style="list-style-type: none"> - Performance élevée - Efficacité - Flexibilité pour divers types de tâches - Gestion des valeurs manquantes - Régularisation intégrée 	<ul style="list-style-type: none"> - Complexité de l'hyperparamétrage - Consommation mémoire - Temps d'entraînement pour de très grands ensembles de données 	<ul style="list-style-type: none"> - Finance (détection des fraudes, modélisation du risque) - Marketing (segmentation de la clientèle, analyse du churn) - Bioinformatique - Santé
Random Forest	<ul style="list-style-type: none"> - Ensemble d'arbres de décision pour des prédictions stables. - Bagging pour réduire la variance. - Sélection aléatoire de caractéristiques pour chaque nœud. - Robustesse face aux valeurs manquantes et données hétérogènes. - Importance des caractéristiques. - Résistance au surapprentissage par l'agrégation de nombreux arbres. 	<ul style="list-style-type: none"> - Très précis et stable. - Facilité d'utilisation avec peu de prétraitement des données. - Permet d'estimer l'importance des caractéristiques. - Efficace pour une variété de tâches de classification et de régression. 	<ul style="list-style-type: none"> - Gourmand en mémoire et en temps de calcul. - Moins interprétable qu'un seul arbre de décision. - Peut surapprendre si les hyperparamètres ne sont pas correctement ajustés. - Risque de biais. 	<ul style="list-style-type: none"> - Finance : Prédiction des risques de crédit, détection de fraudes. - Santé : Prédiction des maladies. - Marketing : Segmentation de marché. - Écologie : Classification des espèces.
CatBoost Classifier	<ul style="list-style-type: none"> - Gestion des données catégorielles - Réduction de surajustement - Traitement efficace - Gestion des données manquantes 	<ul style="list-style-type: none"> - Performances élevées sur les données catégorielles - Réduction de surajustement - Facilité d'utilisation - Optimisé pour CPU et GPU 	<ul style="list-style-type: none"> - Ressources intensives - Complexité de configuration 	<ul style="list-style-type: none"> - Marketing et publicité - Finance - Sciences sociales - Médecine
LightGBM	<ul style="list-style-type: none"> Algorithme de boosting par gradient - Utilisation d'une stratégie de croissance feuille par feuille - Optimisation du binning pour réduire la complexité du modèle 	<ul style="list-style-type: none"> - Haute performance en termes de vitesse d'entraînement et de prédiction - Gestion efficace de la mémoire - Bonne précision sur de grands ensembles de données 	<ul style="list-style-type: none"> - Nécessite des hyperparamètres sensibles pour optimiser les performances - Peut nécessiter plus de données pour éviter le surajustement 	<ul style="list-style-type: none"> - Classification - Régression - Recommandation - Ranking

II.4.4. Prétraitement des données

Avant de procéder à la modélisation par apprentissage automatique, une étape cruciale de préparation des données a été nécessaire. Tout d'abord, nous avons extrait les valeurs des points érosifs et non érosifs pour chaque paramètre étudié. Pour cette étape, nous avons sélectionné 250 points érosifs et 240 points non érosifs répartis sur l'ensemble du bassin versant. Ces points ont été intégrés dans deux tableaux distincts.

Ensuite, avant de passer à la modélisation et à l'application des différents algorithmes d'apprentissage, il est essentiel de normaliser les données. La normalisation est un processus qui ajuste les valeurs des données pour qu'elles se situent toutes dans une plage uniforme, ici de 0 à 1. Cela permet d'assurer que chaque paramètre contribue de manière égale au modèle et améliore ainsi la performance des algorithmes d'apprentissage automatique en facilitant leur convergence.

La normalisation des données est particulièrement importante lorsque les paramètres ont des échelles différentes, car les algorithmes d'apprentissage automatique sont sensibles aux variations de l'échelle des données. En normalisant, nous réduisons l'impact des grandes différences d'échelle et améliorons la robustesse et la précision des modèles.

Ainsi, la préparation des données comprend l'extraction et la sélection des points érosifs et non érosifs, leur insertion dans des tableaux distincts, et enfin la normalisation de ces données. Ce processus assure que les algorithmes d'apprentissage automatique que nous allons utiliser pourront fonctionner de manière optimale, offrant des résultats plus précis et fiables pour la modélisation de l'érosion sur le bassin versant.

Normalisation des données

La formule de normalisation standard que l'on utilise souvent dans le contexte de l'apprentissage automatique est la normalisation min-max. Cette méthode permet de transformer les données pour qu'elles se situent dans une plage définie, généralement entre 0 et 1. Voici la formule de normalisation min-max :

$$X_{\text{norm}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}} \quad \text{Où :}$$

- x est la valeur à normaliser,
- X_{norm} est la valeur normalisée,
- X_{min} est la valeur minimale de la plage de données initiale,
- X_{max} est la valeur maximale de la plage de données initiale.

II.4.5. Evaluation des performances

Pour évaluer la performance des modèles, plusieurs métriques ont été utilisées selon le type de tâche. Pour les problèmes de classification, les critères tels que l'exactitude (accuracy), la précision, le rappel (recall), le F1-score et l'AUC-ROC ont été essentiels. L'accuracy mesure la proportion de prédictions correctes parmi l'ensemble des prédictions. La précision quant à elle indique la proportion de prédictions positives correctes parmi toutes les prédictions positives faites par le modèle. Le recall, aussi appelé rappel, mesure la proportion de vrais positifs identifiés parmi tous les exemples réels positifs. Le F1-score est une mesure combinée de précision et de recall, fournissant un équilibre entre les deux. L'AUC-ROC (Area Under the Receiver Operating Characteristic curve) évalue la capacité du modèle à distinguer entre les classes positives et négatives. Ces métriques permettent une évaluation

approfondie et objective des performances des modèles, essentielle pour leur sélection et leur amélioration.

II.5. Conclusion

Ce chapitre a jeté les bases essentielles pour comprendre et aborder la cartographie de la susceptibilité à l'érosion hydrique dans le bassin versant Isser à travers une approche intégrée de données géospatiales et d'apprentissage automatique. En explorant la géographie et la topographie diversifiées de la région, nous avons identifié les zones à risque potentiel d'érosion, en tenant compte des caractéristiques hydrologiques qui influencent la dynamique des processus érosifs.

La collecte et le prétraitement minutieux des données géospatiales, incluant des images satellites et des modèles numériques de terrain, ont constitué une étape cruciale pour préparer le terrain à l'application des algorithmes d'apprentissage automatique. Ces méthodes, telles que Extreme Gradient Boosting (XGBoost), Random Forest (RF), Categorical Boosting (CatBoost) et LightGBM ont été sélectionnées et entraînées avec rigueur pour modéliser la susceptibilité à l'érosion hydrique.

En parvenant à valider nos modèles à l'aide de techniques robustes, nous avons assuré leur capacité à généraliser au-delà des données d'entraînement, offrant ainsi une perspective fiable pour la prédiction et la gestion prospective des risques d'érosion. Cette approche méthodologique ne se contente pas de cartographier les zones vulnérables, mais ouvre également la voie à des stratégies adaptatives de conservation des sols et de gestion des ressources en eau.

Enfin, ce chapitre pose les fondations nécessaires pour le chapitre suivant de cette thèse, qui exploreront en détail les résultats obtenus et leur application pratique dans le contexte spécifique du bassin versant Isser. À travers cette étude, nous visons à apporter une contribution significative à la compréhension scientifique et à la gestion durable des écosystèmes fragiles confrontés aux défis croissants de l'érosion hydrique et du changement environnemental.

Chapitre III.

Résultats et Discussions...

III.1. Introduction

Ce chapitre présente une analyse approfondie des résultats obtenus sur la cartographie de la susceptibilité à l'érosion hydrique du bassin versant Isser, essentielle pour la gestion environnementale et la préservation des sols. Il démontre comment la cartographie de l'érosion permet de visualiser les zones à risque et de comparer l'efficacité des nouvelles méthodes proposées. Les cartes générées offrent une vue détaillée des zones susceptibles à l'érosion, fournissant une base solide pour des interventions ciblées. Enfin, les implications pratiques des résultats sont discutées avec des recommandations pour la gestion environnementale et des suggestions d'améliorations futures, visant à optimiser l'application des résultats dans des contextes réels et à contribuer à une gestion plus efficace et durable des ressources naturelles.

III.2. Présentation des cartes générées

III.2.1. Carte des pentes

La carte des pentes du bassin versant Isser présente les différentes inclinaisons du terrain, classées en cinq catégories distinctes. Les pentes douces, inférieures à 5%, couvrent une superficie de 1006,67 km², représentant 24,11% du bassin. Les pentes modérées, entre 5% et 10%, couvrent 1071,69 km², soit 25,67% de la surface totale. Les pentes plus prononcées, entre 10% et 18%, sont les plus étendues, représentant 1303,11 km², soit 31,21%. Les pentes plus raides, entre 18% et 25%, occupent 553,58 km², correspondant à 13,26%. Enfin, les pentes très raides, supérieures à 25%, couvrent une superficie de 247,56 km², représentant 5,93% du bassin. Cette répartition des pentes est cruciale pour comprendre les dynamiques d'érosion, la gestion des ressources en eau et l'aménagement du territoire dans cette région.

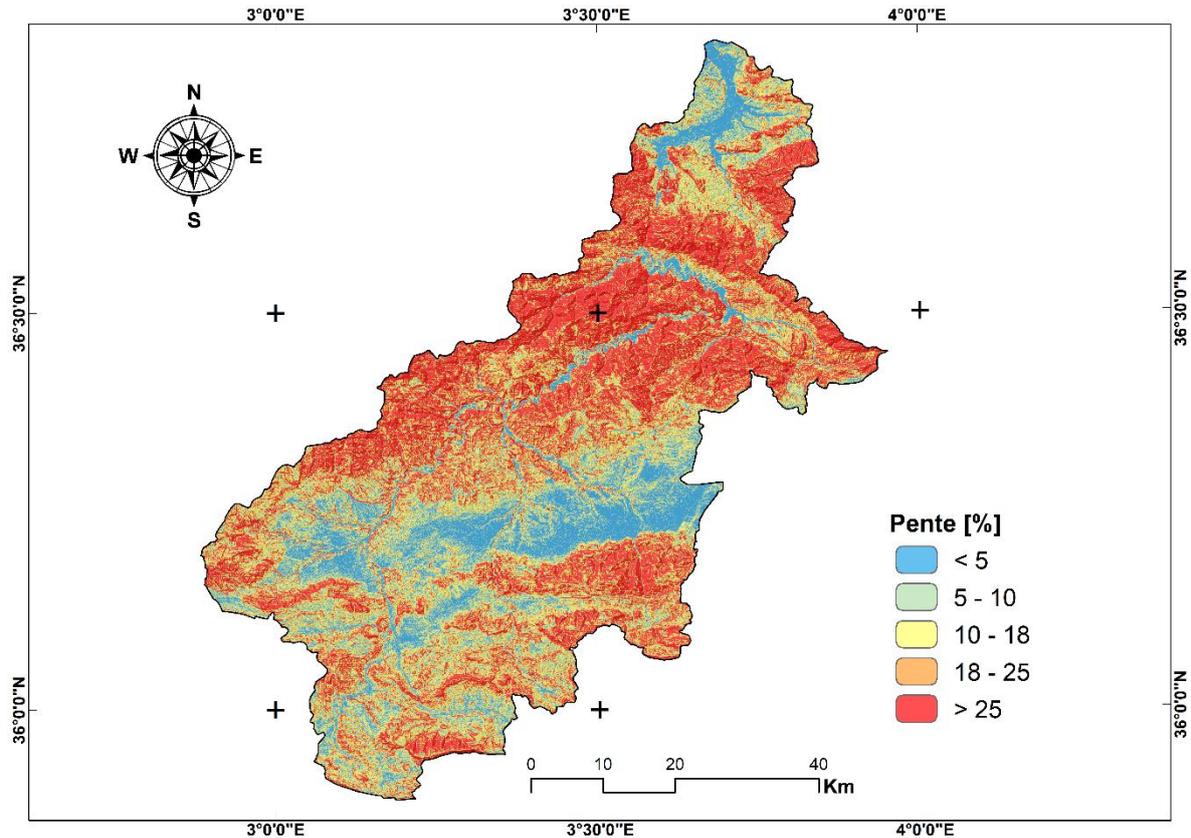


Figure.III.1. Carte des pentes de bassin versant Isser.

Tableau III.1. Classe des pentes.

Classes	Surface (km ²)	Surface (%)
<5	1006.67	24.11
5 -10	1071.69	25.67
10 -18	1303.11	31.21
18 - 25	553.58	13.26
>25	247.56	5.93

III.2.2. Carte des élévations

La carte d'élévation montre la distribution des altitudes dans le bassin versant Isser, divisée en cinq classes distinctes. Les zones situées à moins de 300m d'altitude couvrent 531,71 km², représentant 12,71% de la superficie totale. Les terrains situés entre 300m et 550m d'altitude couvrent 678,87 km², soit 16,23% du bassin. La classe d'altitude la plus dominante, entre 550m et 750m, s'étend sur 1360,39 km², représentant 32,53% de la surface totale. Les altitudes comprises entre 750m et 1000 m couvrent 1046,20 km², correspondant à 25,01% de la région. Enfin, les zones situées à plus de 1000m d'altitude occupent 565,17 km², soit 13,51% du bassin. Cette répartition altitudinale est essentielle pour comprendre les dynamiques hydrologiques, les variations climatiques et les potentialités agricoles du bassin versant.

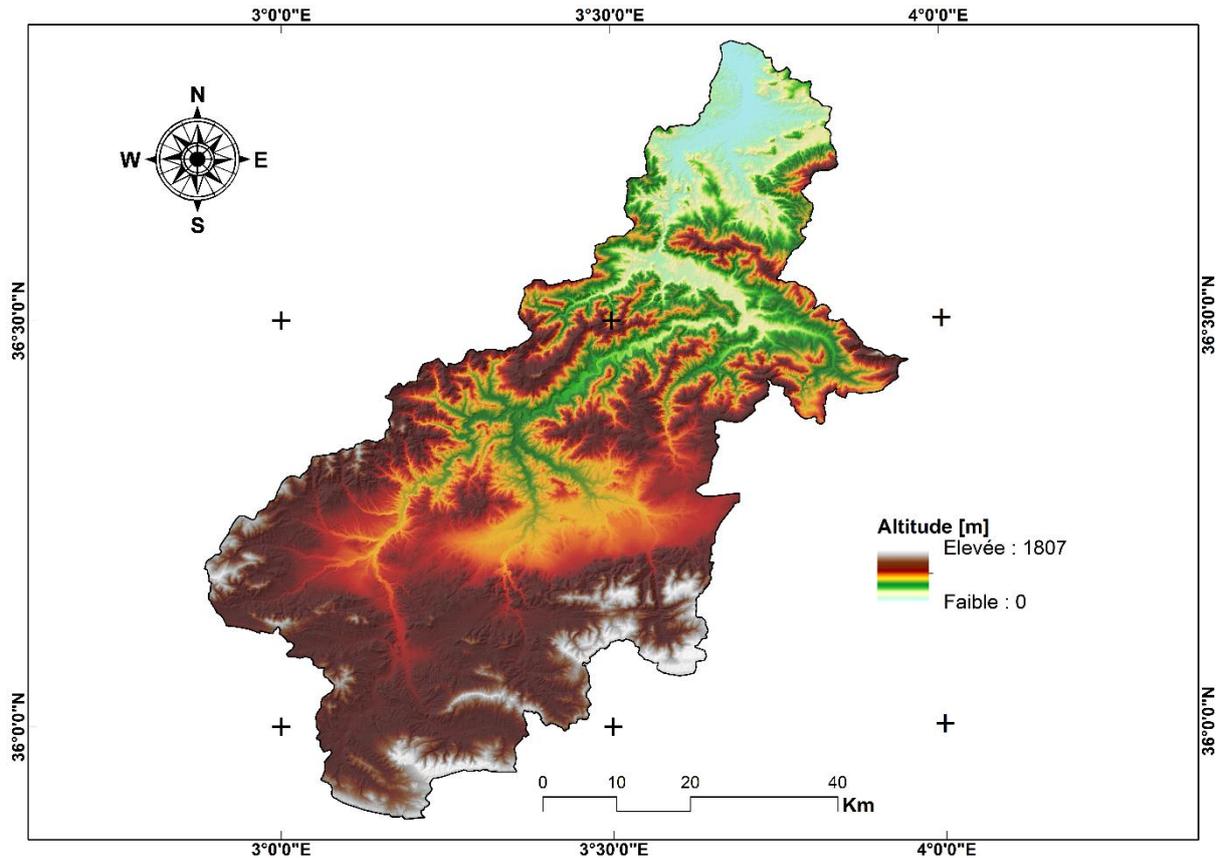


Figure.III.2. Carte des élévations de bassin versant Isser.

III.2.3. Carte des aspects

La carte présente les différentes orientations des pentes (aspects) dans le bassin versant Isser, représentées par différentes couleurs. Chaque aspect, ou orientation de pente, est crucial pour comprendre la répartition de la lumière solaire, l'humidité et les microclimats dans la région. Cette carte fournit des informations précieuses sur la distribution des orientations des pentes, ce qui est crucial pour la gestion des terres, la conservation des sols et l'optimisation de l'utilisation agricole. Chaque orientation présente des caractéristiques uniques qui influencent la lumière solaire, l'humidité et la susceptibilité à l'érosion.

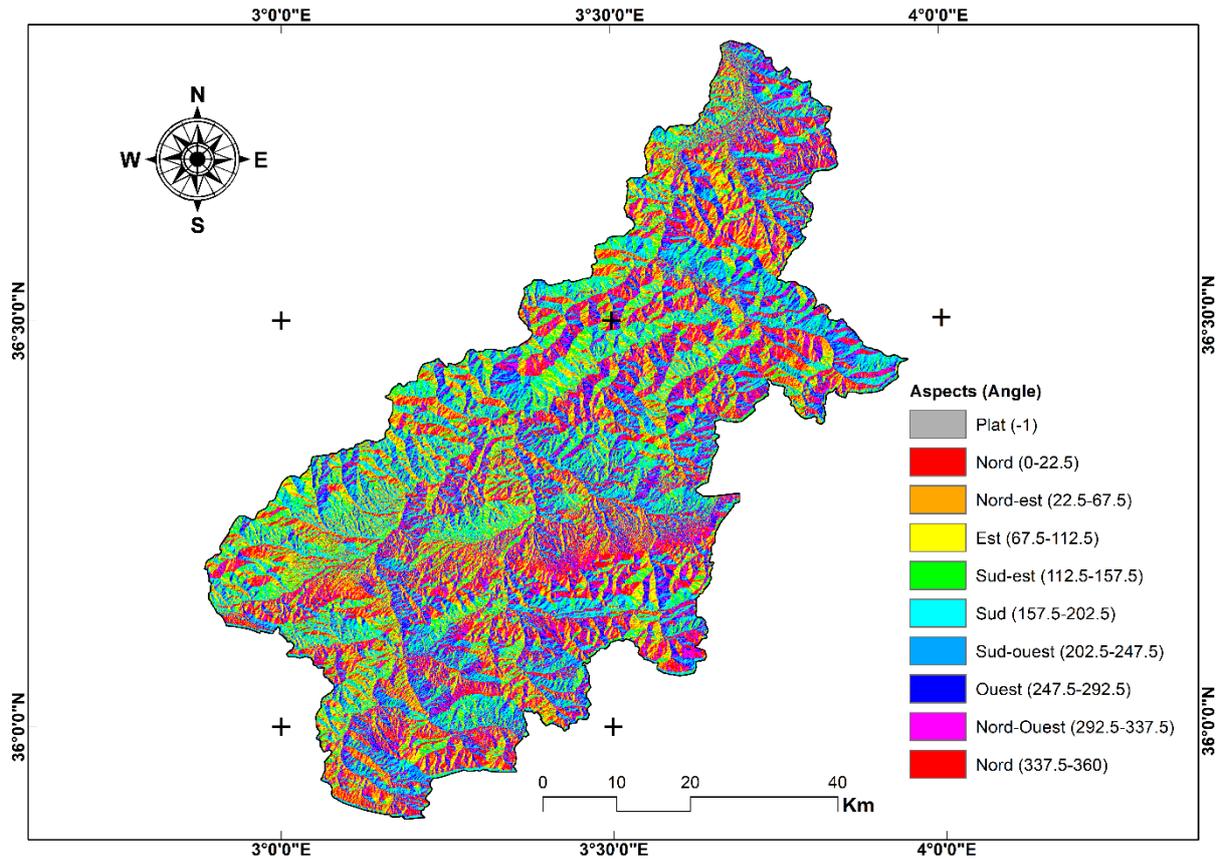


Figure.III.3. Carte des aspects de bassin versant Isser.

III.2.4. Indice d'humidité topographique (TWI)

La carte de l'indice d'humidité topographique (TWI) du bassin versant Isser fournit des informations cruciales pour la gestion des terres et des ressources en eau. Les zones à faible TWI nécessitent une gestion pour prévenir l'érosion et la dégradation des sols, tandis que les zones à TWI élevé sont plus propices à l'agriculture mais doivent être protégées contre les inondations et l'érosion. Une approche intégrée de gestion des bassins versants, comprenant des pratiques de conservation du sol et de l'eau, est essentielle pour maintenir la santé environnementale et la productivité agricole de la région.

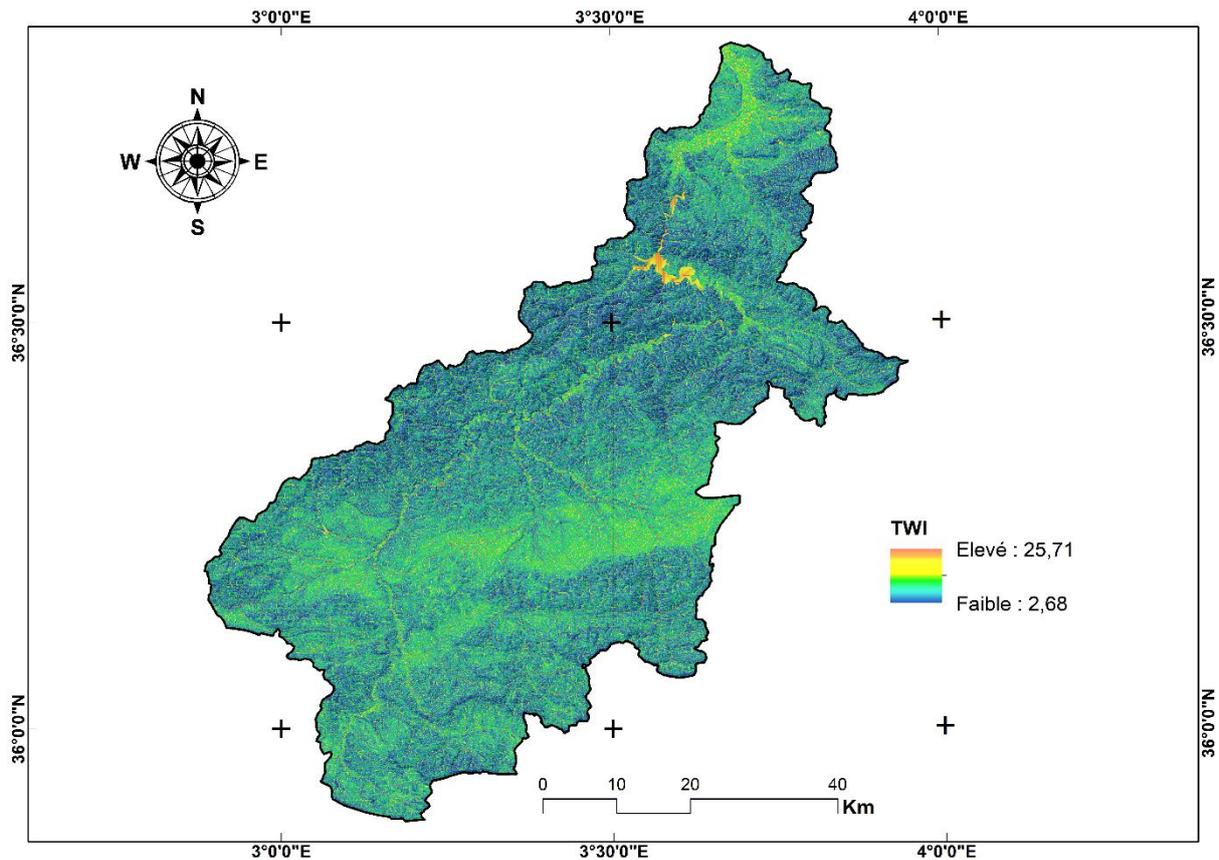


Figure.III.4. Indice d'humidité topographique (TWI).

III.2.5. Carte d'occupation du sol (LU/LC)

Le tableau II.2 illustre la répartition des types d'occupation des sols dans le bassin Isser. Les données révèlent que la majorité de la surface est constituée de sols nus, couvrant 2108,44 km², soit 50,72 % de la superficie totale. Les terres cultivées forment la deuxième catégorie la plus étendue, avec 978,31 km², représentant 23,53 % de la surface. La végétation couvre 639,82 km², soit 15,39 % de la superficie totale. Les zones d'agglomération occupent 425,62 km², soit 10,24 % de la surface. Les plans d'eau représentent une petite fraction, avec seulement 4,84 km², soit 0,12 % de la surface. Cette répartition indique une prédominance inquiétante des sols nus, signalant un potentiel élevé d'érosion et de dégradation des sols. Cela nécessite des interventions urgentes pour améliorer la couverture végétale et promouvoir des pratiques agricoles durables.

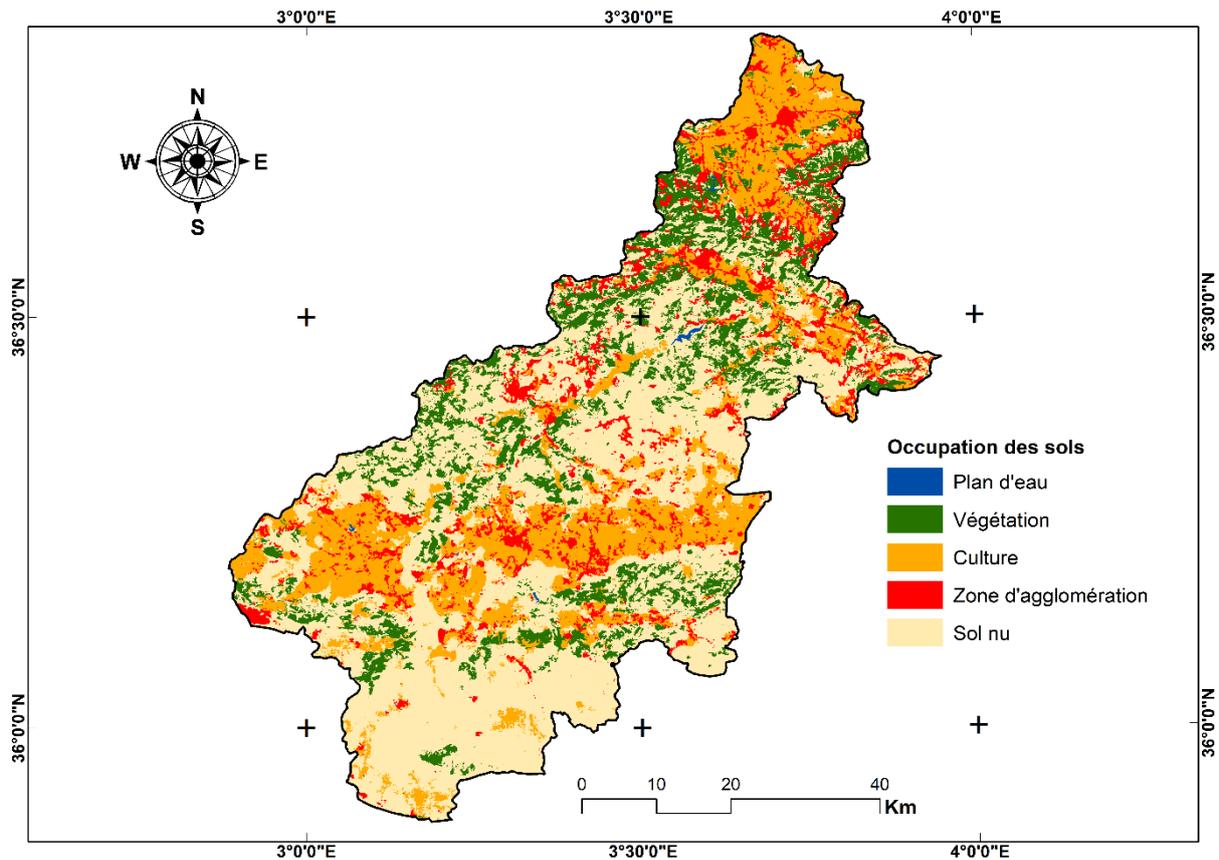


Figure.III.5. Carte d'occupation du sol (LULC).

Tableau III.2. Classe d'occupation du sol.

Classe	Surface (Km ²)	Surface (%)
Plan deau	4,84	0,11
Végétation	639,81	15,39
Culture	978,31	23,53
Zone d'agglomération	425,62	10,23
Sol nu	2108,44	50,72

III.2.6. Facteur d'érodibilité du sol (K)

Le Facteur K, représentant l'érodabilité des sols, est un élément essentiel de cette étude. Il permet de comprendre la susceptibilité des sols à l'érosion hydrique dans le bassin versant Isser. Ce facteur varie en fonction de la texture du sol, de sa teneur en matière organique, de sa structure et de sa perméabilité. L'analyse de la carte révèle que les sols du bassin versant Isser ayant des valeurs élevées de Facteur K sont particulièrement vulnérables à l'érosion. Les valeurs du Facteur K dans cette région varient entre 0,027 et 0,036, avec des valeurs plus élevées prédominant dans le sud et le sud-ouest, tandis que le nord et le nord-est affichent des valeurs relativement faibles.

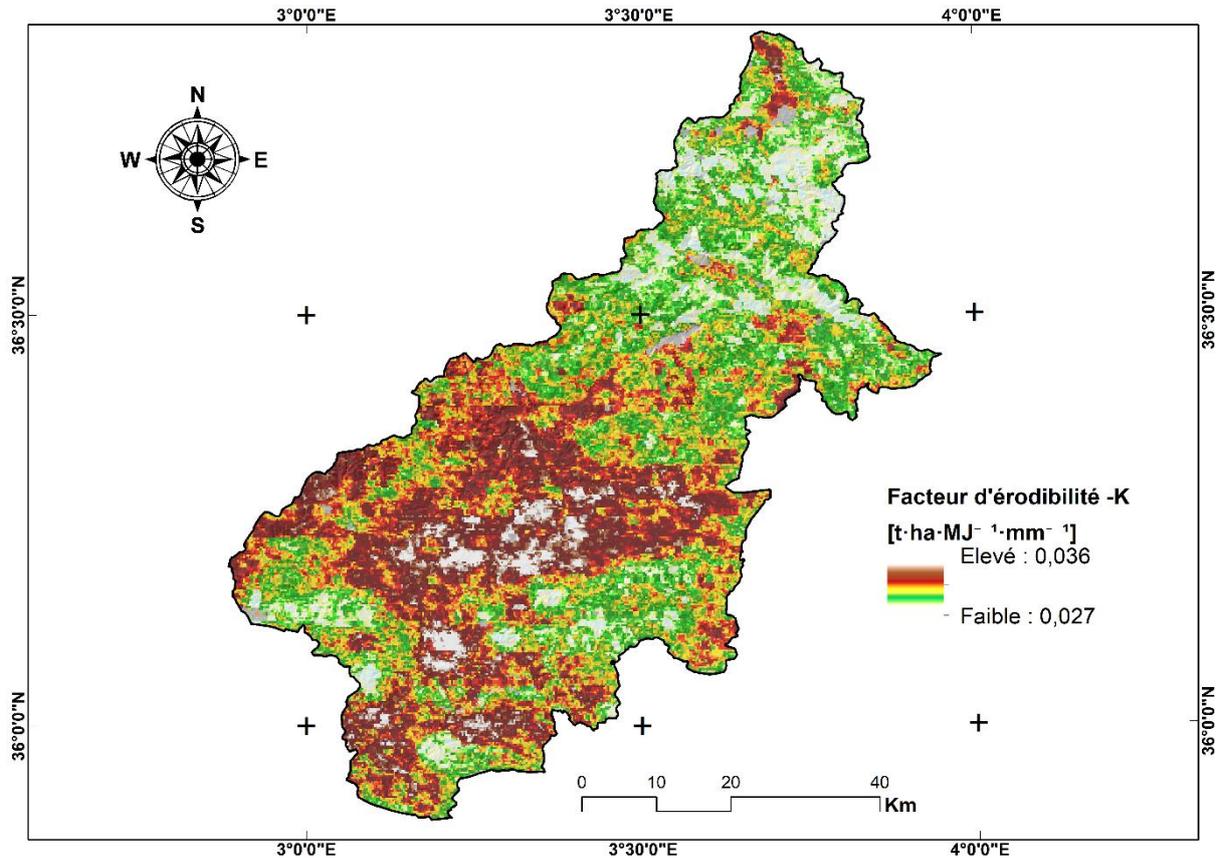


Figure.III.6. Carte de facteur K.

III.2.7. La distance par rapport à la route

La distance par rapport aux routes et la proximité à celles-ci sont des facteurs essentiels pour évaluer la vulnérabilité à l'érosion d'un lieu. Les routes peuvent limiter l'absorption d'eau dans le sol, ce qui entraîne une augmentation du ruissellement de surface et favorise ainsi l'érosion des zones avoisinantes. Par conséquent, les zones situées à proximité des routes sont plus exposées à l'érosion en raison d'une infiltration réduite et d'un ruissellement accru.

La carte de distance par rapport aux routes pour notre bassin d'étude a été générée à partir de la carte des routes en utilisant l'outil ArcMap. Vous trouverez ci-dessous la carte illustrant les différentes classes de distance par rapport aux routes

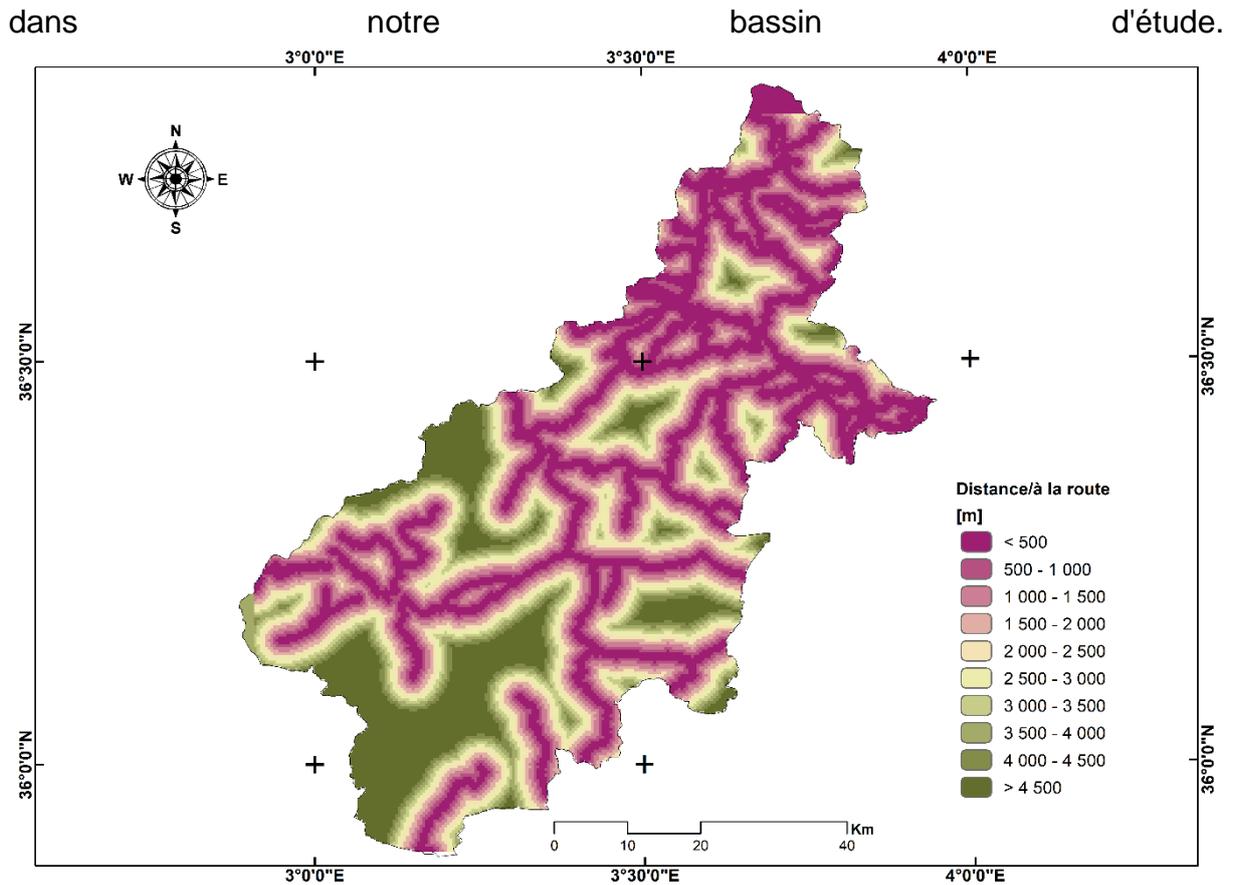


Figure.III.7. Carte de distance par rapport à la route.

III.2.8. La distance par rapport aux oueds

La carte des distances par rapport aux oueds (cours d'eau) indique la proximité des terres aux cours d'eau dans le bassin versant Isser. Les zones proches des oueds sont plus susceptibles à l'érosion en raison du débit des cours d'eau, surtout lors des crues. Les zones proches des rivières (0-250 m) sont plus susceptibles de subir une érosion fluviale en raison de l'action directe de l'eau courante. Et les zones plus éloignées peuvent être moins affectées par l'érosion directe des rivières, mais elles peuvent encore subir une érosion due au ruissellement des eaux pluviales.

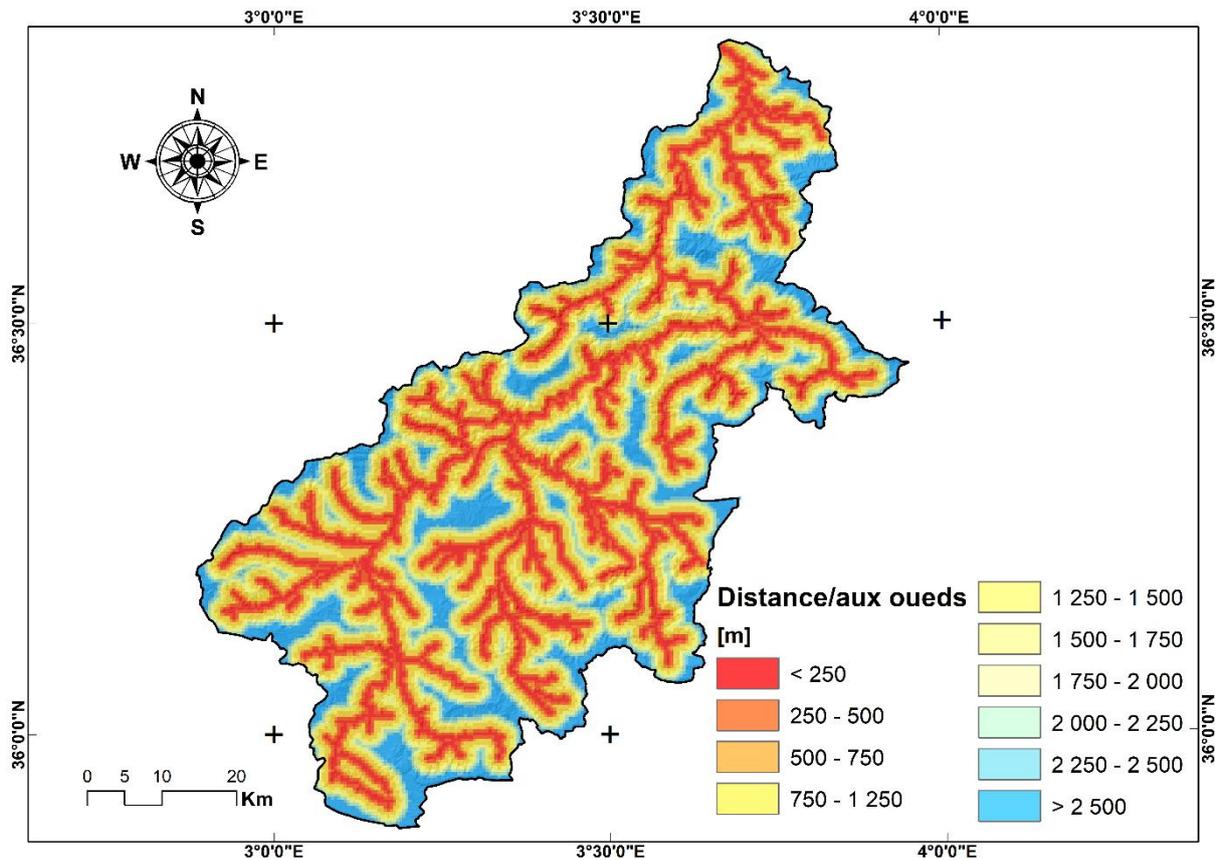


Figure.III.8. Carte de distance par rapport aux oueds.

III.2.9. Classes des sols

L'étude de la susceptibilité à l'érosion hydrique dans le bassin versant Isser révèle que les Luvisols dominent largement, couvrant 64,67% de la surface. Ces sols, riches en argile et souvent fertiles, sont modérément susceptibles à l'érosion et nécessitent une bonne couverture végétale et des pratiques agricoles adéquates. Les Vertisols, représentant 15,82% de la surface, sont des sols argileux lourds qui se fissurent en période sèche et se gonflent lorsqu'ils sont humides, augmentant ainsi leur risque d'érosion en présence de fortes pluies. Les Leptosols, bien que couvrant seulement 5,60%, sont très vulnérables à l'érosion hydrique en raison de leur faible profondeur et de leur localisation sur des terrains en pente. Les Cambisols (6,31%), jeunes et peu développés, sont modérément susceptibles à l'érosion, surtout sur des pentes non protégées. Les Calcisols (6,12%), riches en carbonate de calcium, voient leur susceptibilité à l'érosion varier selon leur structure et la couverture végétale. Les Regosols (1,38%), sols jeunes et non consolidés, sont très susceptibles à l'érosion en raison de leur faible cohésion. Les Phaeozems (0,10%), bien drainés et riches en matière organique, sont généralement moins susceptibles à l'érosion mais leur faible étendue limite leur impact global. Ainsi, une gestion durable des sols et des pratiques de conservation efficaces sont essentielles pour réduire la vulnérabilité à l'érosion dans cette région.

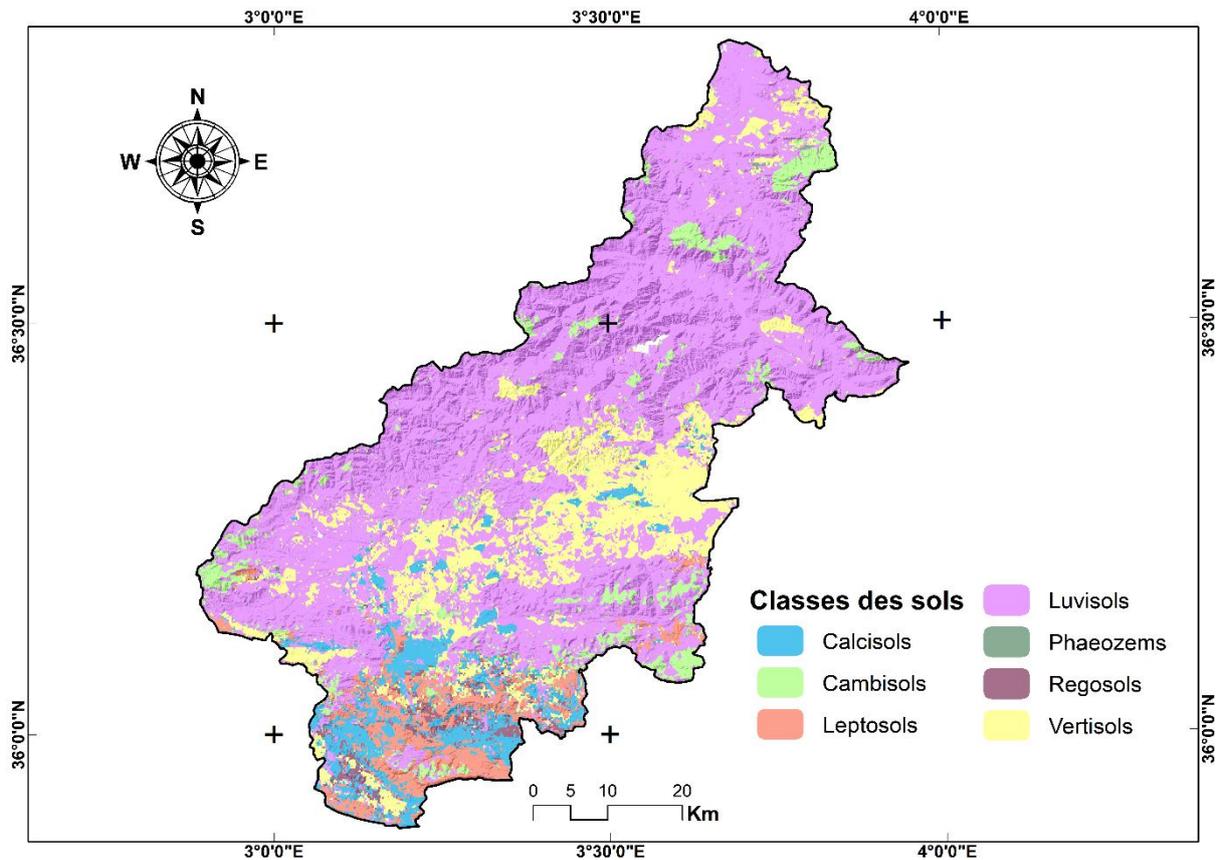


Figure.III.9. Classes des sols dans le bassin versant Isser.

Tableau III.3. Classe des sols.

Classe	Surface (Km ²)	Surface (%)
Calcisols	254,08	6,12
Cambisols	262,12	6,31
Leptosols	232,25	5,60
Luvisols	2684,58	64,67
Phaeozems	4,05	0,10
Regosols	57,24	1,38
Vertisols	656,64	15,82

III.2.10. Carte de l'Indice de Fournier Modifié (MFI)

Cette carte montre la distribution spatiale de l'Indice de Fournier Modifié (MFI) dans le bassin versant Isser. Le MFI est une mesure utilisée pour évaluer l'érosivité des précipitations. Il est basé sur l'intensité et la fréquence des précipitations, ce qui permet de déterminer la capacité de la pluie à provoquer l'érosion des sols. Et est un outil précieux pour identifier les zones à risque d'érosion dans le bassin versant. En combinant cette information avec d'autres facteurs, il est possible de développer des stratégies de gestion intégrée des ressources en eau et des sols pour minimiser l'érosion et protéger les écosystèmes locaux.

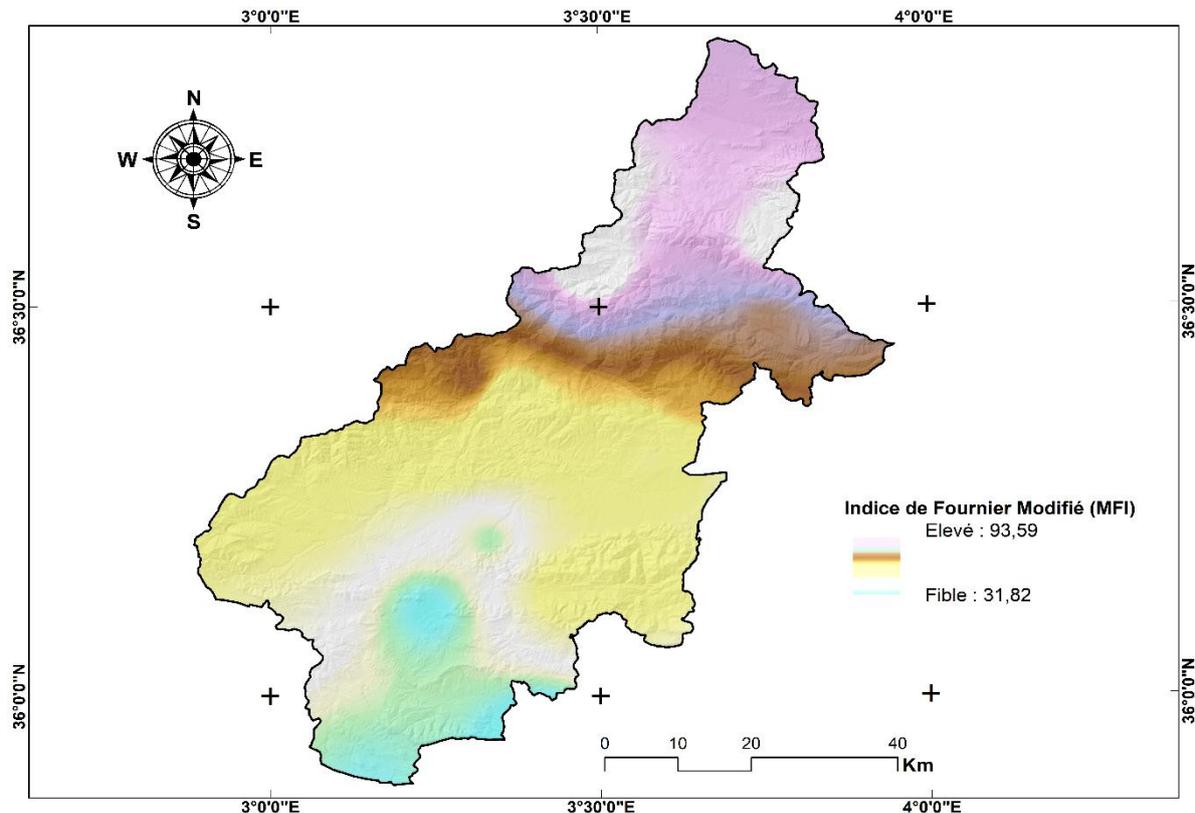


Figure.III.10. Carte de MFI de bassin versant Isser.

III.3. Cartographie de la susceptibilité à l'érosion

Le XGBClassifier, CatBoostClassifier, RandomForest, et DecisionTree Classifier ont été entraînés sur 60% des échantillons de données, tandis que les 40% restants ont été utilisés pour la validation. Les courbes AUC-ROC pour chaque méthode sont présentées dans la (Figure III.11). Suite à cela, les valeurs de susceptibilité ont été classées en cinq niveaux de risque (très faible, faible, modéré, élevé et très élevé) en utilisant la méthode de classification par quantiles. La carte de risque résultante est montrée dans la (Figure III.12).

Les résultats comparatifs des quatre modèles de classification ont été évalués à l'aide de plusieurs métriques. CatBoostClassifier et LightGBM se distinguent avec une AUC de 0,98, tandis que XGBClassifier et RF sont légèrement en retrait. En termes d'exactitude, CatBoostClassifier est en tête avec 0,97, suivi de RF à 0,9653, et les autres à 0,9583. XGBClassifier se démarque par son rappel et sa sensibilité à 0,99, mais présente une précision et une spécificité légèrement inférieures (0,94 et 0,93 respectivement). CatBoostClassifier et RF montrent un excellent équilibre entre précision, rappel et score F1 (tous deux à 0,97), avec une spécificité élevée de 0,97. LightGBM, bien que performant avec des scores F1 et de rappel de 0,96 et 0,97 respectivement, est légèrement en retrait par rapport à CatBoost et RF. En résumé, CatBoostClassifier et RandomForest se démarquent par leurs performances équilibrées et excellentes sur la plupart des métriques importantes. CatBoostClassifier a un léger avantage en termes d'AUC et d'exactitude, ce qui en fait potentiellement le meilleur choix général. Cependant, si la sensibilité (minimisation des faux négatifs) est prioritaire, XGBClassifier serait le meilleur choix. LightGBM reste une bonne option, mais il est légèrement en retrait par rapport à CatBoostClassifier et RF sur certaines métriques.

Tableau III.4. Critères d'évaluation des modèles d'apprentissage automatique.

Statistique	XGBClassifier	CatBoost Classifier	RandomForest	LightGBM
AUC	0.98	0.97	0.97	0.96
Exactitude	0.96	0.97	0.97	0.96
(TP)	74	73	72	73
(TN)	64	67	67	65
(FP)	5	2	2	4
(FN)	1	2	3	2
Précision	0.94	0.97	0.97	0.95
Rappel	0.99	0.97	0.96	0.97
Score F1	0.96	0.97	0.97	0.96
Sensibilité	0.99	0.97	0.96	0.97
Spécificité	0.93	0.97	0.97	0.94

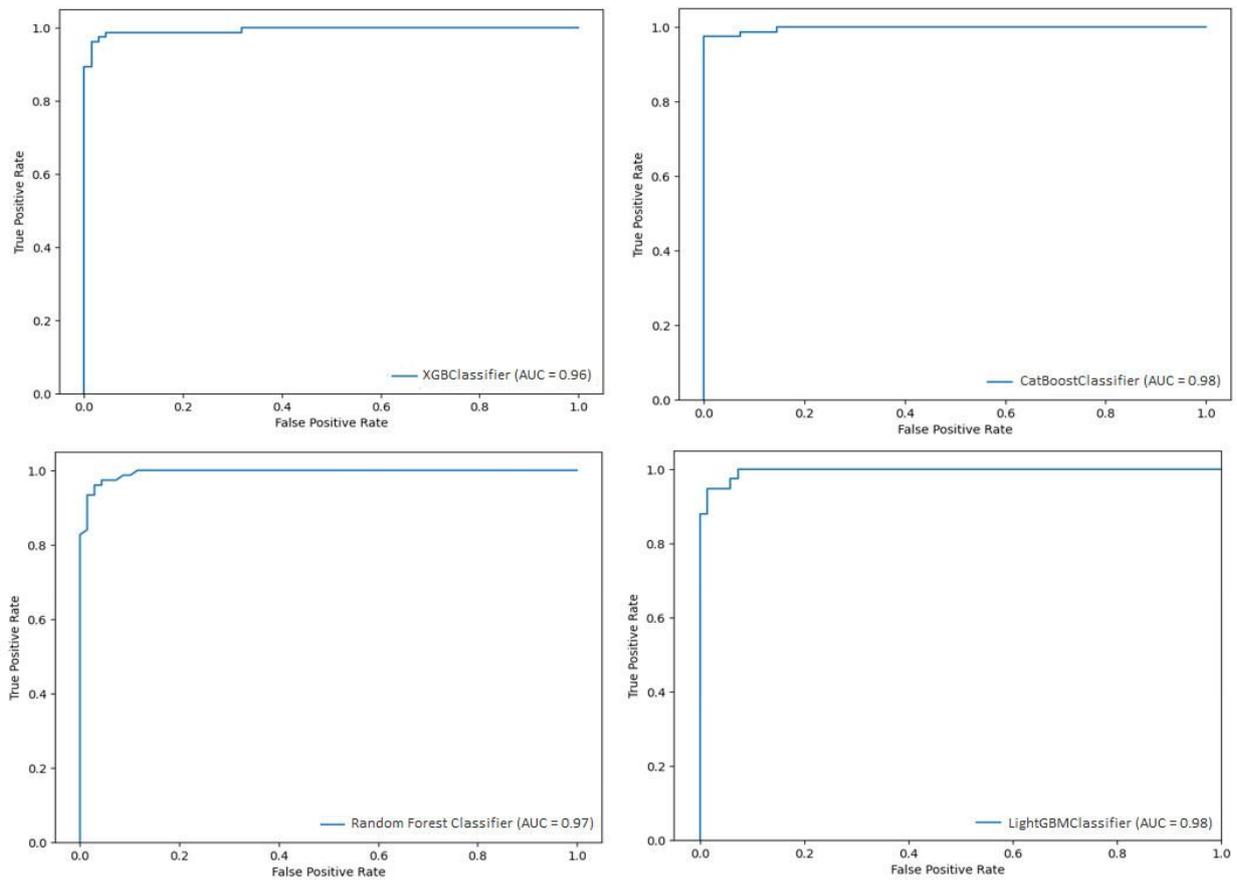


Figure III.11. Courbes ROC-AUC pour les quatre modèles.

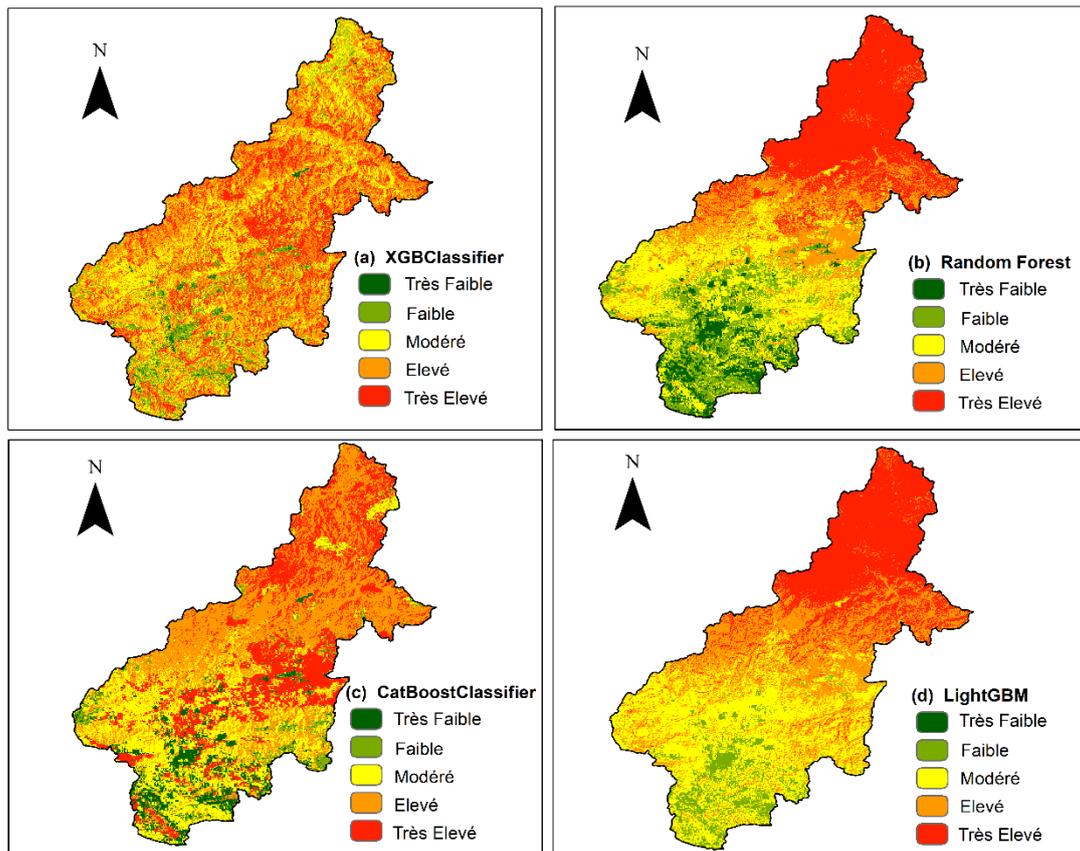


Figure III.12. Carte de susceptibilité à l'érosion par utilisation des algorithmes : (a) XGBClassifier, (b) Random Forest, (c) CatBoostClassifier et (d) LightGBM.

Le tableau III.5 présente l'importance des caractéristiques pour divers modèles de classification, dont XGBoost Classifier, Random Forest, CatBoost Classifier et LightGBM. Voici une synthèse des résultats :

III.3.1. XGBoost Classifier

La caractéristique la plus déterminante est la Géologie, avec un score d'importance de 0,56, représentant 55,56 % de l'importance totale. Viennent ensuite le pourcentage de pente et le NDVI, avec des scores respectifs de 0,17 (17,22 %) et 0,13 (13,12 %). D'autres caractéristiques, comme NDBI, NDWI, Lithologie, LULC et Distance par rapport à la route, ont une importance nulle, indiquant leur faible contribution aux prédictions du modèle.

III.3.2. Random Forest

La Géologie reste prédominante avec une importance de 42,85 %, suggérant une forte influence sur les prédictions. D'autres caractéristiques significatives incluent NDBAI (9,15 %), pourcentage de pentes (7,42 %), et le facteur d'érosivité (5,94 %). Des caractéristiques comme l'ombrage (5,53 %), NDVI (4,34 %), BI (3,45 %), et les aspects (3,37 %) montrent également une contribution notable. En revanche, des caractéristiques telles que NDWI (1,81 %), NDBI (1,68 %), TWI (1,44 %), et la lithologie (0,33 %) ont une importance relativement faible.

III.3.3. CatBoost Classifier

La Géologie domine largement avec une contribution de 77,37 %. Les autres caractéristiques importantes, bien que moins influentes, incluent la pente (3,51 %), le NDVI (2,68 %), le NDBAI (2,40 %), l'ombrage (2,13 %), l'altitude (1,94 %), le facteur R (1,87 %), et le MFI (1,52 %). Les caractéristiques restantes, telles que l'utilisation des terres, la distance par rapport à la route, l'aspect, et la distance par rapport aux oueds, ont toutes une importance inférieure à 1 %, la lithologie étant la moins influente avec 0,36 %.

III.3.4. LightGBM

La Géologie reste la plus influente avec une contribution de 23,35 %. La pente (14,26 %), le NDBAI (9,83 %), le NDVI (9,23 %), et l'altitude (7,13 %) suivent. Les autres caractéristiques importantes incluent la distance par rapport à la route (5,48 %), le MFI (5,33 %), le TWI (5,18 %), l'ombrage (5,11 %), et l'aspect (4,50 %). Les facteurs comme le BI (3,45 %), le facteur R (2,85 %), le NDBI (2,18 %), l'utilisation des terres (1,20 %), et le NDWI (0,68 %) ont une importance moindre, et la distance par rapport aux oueds n'a aucune importance dans ce modèle.

Tableau III.5. Importance des caractéristiques des modèles XGBoost Classifier, Random Forest, CatBoost Classifier et LightGBM.

Paramètre	Importance (%)			
	XGBClassifier	CatBoostClassifier	Random Forest	LightGBM
Géologie	55,56	77,37	42,85	23,35
Pente (%)	17,22	3,51	7,42	14,26
NDVI	13,12	2,68	4,34	9,23
TWI	4,15	0,57	1,44	5,18
Aspect	2,36	0,90	3,37	4,50
Ombrage	1,66	2,13	5,53	5,11
NDBAI	1,34	2,40	9,15	9,83
BI	1,29	0,77	3,45	3,45
Distance/à la route	1,23	0,94	2,77	5,48
MFI	0,80	1,52	2,79	5,33
facteur R	0,79	1,87	5,94	2,85
Élévation	0,48	1,94	2,65	7,13
NDBI	0,00	0,59	1,68	2,18
Occupation du sol	0,00	0,99	3,08	1,20
NDWI	0,00	0,64	1,81	0,68
Lithologie	0,00	0,36	0,33	0,23
Distance/Aux oueds	0,00	0,81	1,39	0,00

Ces résultats mettent en évidence l'importance variable des caractéristiques selon le modèle utilisé. La Géologie apparaît systématiquement comme une caractéristique clé, tandis que d'autres facteurs comme la pente, le NDVI, et divers indices de végétation et d'élévation jouent également des rôles significatifs mais à des degrés différents. Cette analyse peut guider les efforts de sélection de caractéristiques et d'amélioration des modèles pour des applications futures.

III.4. Analyse des effets des changements des classes de sol sur la perte de sols

Le bassin versant Isser, situé au nord-ouest de l'Algérie, fait face à un grave problème d'érosion, entraînant une perte significative de sols et compliquant la mise en œuvre de pratiques de gestion durable des terres. Il est donc essentiel d'identifier les zones de forte vulnérabilité en utilisant les techniques de modélisation les plus efficaces pour mettre en place des mesures appropriées de conservation des sols et de l'eau.

Pour estimer avec précision l'érosion dans cette zone et déterminer le modèle le plus approprié, nous avons appliqué quatre modèles d'intelligence artificielle : XGBoost Classifier, Random Forest, CatBoost Classifier et LightGBM.

Notre recherche visait à identifier les principaux facteurs contribuant à l'érosion, un phénomène influencé par divers paramètres. Il a été établi que la géologie est un facteur crucial, comme prévu, et que la pente, le NDVI, le TWI, le NDBAI et plusieurs autres facteurs sont dominants dans la cartographie des zones sensibles à l'érosion. Bien que des variations dans l'importance des variables aient été observées entre les modèles, la précision des résultats reste exceptionnelle pour chacun des modèles élaborés.

Les résultats obtenus montrent les pourcentages de surfaces classées selon le degré de risque érosif à l'aide de quatre modèles de classification : XGBClassifier, CatBoostClassifier, Random Forest, et LightGBM. Globalement, il y a des variations notables entre les modèles quant à la manière dont ils évaluent le risque érosif. Le modèle XGBClassifier tend à être le plus pessimiste, classant la majorité des surfaces dans les catégories de risque élevé (39,77%) et très élevé (24,58%), avec un faible pourcentage en très faible (1,45%). CatBoostClassifier montre une tendance similaire avec des pourcentages élevés pour les catégories élevées (37,69%) et très élevé (27,16%), mais classe davantage de surfaces en très faible (6,95%).

En revanche, le modèle Random Forest présente une répartition plus équilibrée des surfaces parmi les différentes catégories de risque, avec des pourcentages notables en faible (14,50%), modéré (26,26%), et très élevé (30,54%). Ce modèle semble moins pessimiste comparé à XGBClassifier et CatBoostClassifier. LightGBM, pour sa part, montre une tendance à classer une grande proportion des surfaces dans les catégories modéré (33,60%), élevé (29,58%), et très élevé (28,84%), tout en affichant le pourcentage le plus faible pour le risque très faible (0,19%).

Ces divergences peuvent être attribuées aux caractéristiques spécifiques des algorithmes utilisés par chaque modèle, à la manière dont ils gèrent les données et à leur sensibilité variable aux facteurs contribuant à l'érosion des sols. XGBClassifier et CatBoostClassifier montrent une plus grande prudence, voire un pessimisme, en attribuant des risques plus élevés à une majorité de surfaces. À l'opposé, Random Forest offre une vision plus variée, tandis que LightGBM met l'accent sur les risques modérés et élevés, avec une classification minimale en très faible risque. Ces différences d'approche soulignent l'importance de choisir le modèle le plus adapté en fonction des objectifs spécifiques et des caractéristiques des données disponibles.

Tableau III.6. Statistiques des zones sujettes à l'érosion dans différentes classes des sols.

Modèle	Classe	Calci-sols	Cambi-sols	Lepto-sols	Luvi-sols	Phaeozems	Rego-sols	Verti-sols
XGB Classifier	T. Faible	37,46	3,67	2,18	4,93	1,01	1,05	3,73
	Faible	117,17	55,34	15,38	97,49	2,22	4,73	16,07
	Modéré	62,39	85,77	86,32	810,93	1,65	7,38	48,76
	Elevé	23,97	78,77	92,54	1174,03	1,69	25,31	250,64
	T. Elevé	11,34	35,15	33,38	583,87	0,55	18,51	333,43
Random Forest	T. Faible	37,66	3,39	2,31	4,71	0,00	0,93	4,42
	Faible	118,25	53,99	16,01	97,31	0,26	4,93	15,59
	Modéré	60,90	86,58	85,50	809,55	2,08	6,72	50,66
	Elevé	23,47	78,10	92,19	1169,30	1,63	25,97	248,42
	T. Elevé	10,47	34,26	33,05	581,85	0,67	18,41	331,29
CatBoost Classifier	T. Faible	37,65	3,39	2,31	4,71	0,00	0,93	4,42
	Faible	118,24	53,98	16,01	97,30	0,26	4,93	15,59
	Modéré	60,90	86,58	85,49	809,49	2,08	6,72	50,65
	Elevé	23,47	78,06	92,18	1169,15	1,63	25,97	248,40
	T. Elevé	10,47	34,26	33,04	581,77	0,67	18,41	331,26
LightGBM	T. Faible	37,65	3,39	2,31	4,71	0,00	0,93	4,42
	Faible	118,24	53,98	16,01	97,30	0,26	4,93	15,59
	Modéré	60,90	86,58	85,49	809,49	2,08	6,72	50,65
	Elevé	23,47	78,06	92,18	1169,15	1,63	25,97	248,40
	T. Elevé	10,47	34,26	33,04	581,77	0,67	18,41	331,26

La compréhension des facteurs influents par chaque modèle diffère, ce qui implique que des modèles de grande précision peuvent être dérivés même à partir de facteurs moins précis. Pour mieux saisir ce phénomène, il est recommandé d'utiliser des algorithmes complexes de sélection de caractéristiques, qui permettent de rechercher et d'identifier les associations entre les phénomènes étudiés et leurs facteurs d'influence.

La création d'une carte de susceptibilité à l'érosion est un outil précieux pour atténuer les risques d'érosion des sols induite par l'eau. Les zones identifiées comme ayant une sensibilité élevée ou très élevée, mais n'ayant pas encore subi d'érosion, révèlent des conditions favorables au développement de ce phénomène. Ces zones sont donc particulièrement vulnérables, ce qui souligne l'importance de les inclure dans la carte de susceptibilité à l'érosion. Il est crucial de promouvoir des pratiques durables, tant dans l'agriculture que dans la foresterie, afin de préserver la qualité des sols. En mettant en œuvre des mesures préventives et de gestion appropriées, il est possible de réduire significativement les risques d'érosion et d'assurer la durabilité des terres agricoles et forestières.

III.5. Conclusion

L'étude présentée dans ce chapitre a permis d'élaborer une cartographie détaillée de la susceptibilité à l'érosion hydrique dans le bassin versant Isser, situé dans la partie centrale du nord de l'Algérie. Grâce à l'utilisation de techniques avancées de modélisation, notamment le XGBClassifier, le CatBoostClassifier, la forêt aléatoire (Random Forest) et le LightGBM, nous avons pu identifier avec une grande précision les zones à risque d'érosion.

Ces cartes de susceptibilité à l'érosion se révèlent être des outils extrêmement précieux pour la gestion environnementale. Elles offrent une base solide pour la

planification et la mise en œuvre de mesures de conservation des sols. En identifiant les zones particulièrement vulnérables, ces cartes permettent de cibler les interventions spécifiques nécessaires pour prévenir l'érosion future et préserver la qualité des sols.

Les résultats de cette étude soulignent l'importance de l'intégration de plusieurs algorithmes de modélisation pour accroître la précision des prédictions. Cette approche combinée permet de fournir des informations fiables, essentielles pour une gestion durable des ressources naturelles. En somme, la cartographie élaborée grâce à cette étude constitue une avancée significative dans la protection et la gestion des sols du bassin versant Isser, contribuant ainsi à la conservation des écosystèmes et au développement durable de la région.

Conclusion Générale...

En combinant les prédictions des quatre algorithmes d'apprentissage automatique, XGBClassifier, CatBoostClassifier, RandomForest et LightGBM, cette recherche a généré des cartes de susceptibilité à l'érosion pour le bassin versant Isser, au nord de l'Algérie. Les résultats mettent en évidence que la géologie et la pente sont les deux paramètres ayant une importance élevée selon ces modèles.

CatBoostClassifier a identifié la géologie comme le principal prédicteur avec une importance significative de 77,37%, tandis que XGBClassifier a souligné l'importance de la pente avec 17,22%. LightGBM a également montré que la pente joue un rôle crucial, bien que moins prononcé que dans XGBClassifier, avec 14,26%. RandomForest, quant à lui, a attribué une importance relative à la géologie (42,85%) et à la pente (7,42%).

L'utilisation de ces méthodes a permis de développer des cartes précises de susceptibilité à l'érosion, essentielles pour une gestion durable des ressources en sols et des bassins versants. En intégrant ces résultats dans un modèle d'ensemble, cette étude a produit une estimation robuste et fiable de la susceptibilité à l'érosion, exploitant les forces de chaque algorithme tout en atténuant leurs limitations respectives.

Ces cartes fournissent des informations critiques pour les décideurs et les gestionnaires environnementaux, facilitant une planification proactive et efficace pour la conservation des terres et la réduction des risques d'érosion. Pour l'avenir, il est recommandé de continuer à perfectionner ces techniques d'apprentissage automatique et à intégrer plus efficacement les projections climatiques afin d'améliorer encore la précision des modèles de prédiction d'érosion. Cela renforcera la résilience des écosystèmes locaux et soutiendra des pratiques de gestion durable des terres dans la région du bassin versant Isser.

Références Bibliographiques...

REFERENCES BIBLIOGRAPHIQUES

- [1] **Aouadj, S., Merzoug, R., Bouzid, N. 2009.** Utilisation du SIG pour la cartographie du risque d'érosion hydrique du sol dans la région de Saïda. Thèse d'ingénieur en biologie, Université de Saïda.
- [2] **Apaydin, H., Erpul, G., Bayramin, I., Gabriels, D. 2006.** Evaluation of Indices for Characterizing the Distribution and Concentration of Precipitation: A Case for the Region of Southeastern Anatolia Project, Turkey. *Journal of Hydrology*, 328, 726–732.
- [3] **Arenas, L.F., Ponce de León, C., Walsh, F.C. 2017.** Engineering aspects of the design, construction and performance of modular redox flow batteries for energy storage. [DOI: 10.1016/j.est.2017.02.007](https://doi.org/10.1016/j.est.2017.02.007).
- [4] **Arnoldus, H.M.J. 1977.** Methodology Used to Determine the Maximum Potential Average Annual Soil Loss Due to Sheet and Rill Erosion in Morocco. *FAO Soils Bulletin*, 5, 39–48.
- [5] **Binet, T., Gonnot, C. 2005.** Étude de l'érosion sur des plantations d'ananas à Moorea (Polynésie française). [DOI: 10.13140/RG.2.1.4944.2806](https://doi.org/10.13140/RG.2.1.4944.2806).
- [6] **Bkhairi, A. 2012.** Suffosion et érosion hydrique en milieux semi-arides : le cas des Hautes Steppes tunisiennes (Tunisie centrale). *Physio-Géo*, volume 6. [DOI: 10.4000/physio-geo.2246](https://doi.org/10.4000/physio-geo.2246).
- [7] **Boughalem, M., Bounoua, L., Amrouni, H. 2012.** Évaluation par analyse multicritères de la vulnérabilité des sols à l'érosion : cas du bassin versant de l'Isser – Tlemcen – Algérie.
- [8] **Bouguerra, H., Taleb, M., Hamimed, A. 2023.** Integration of High-Accuracy Geospatial Data and Machine Learning Approaches for Soil Erosion Susceptibility Mapping in the Mediterranean Region: A Case Study of the Macta Basin, Algeria.
- [9] **De Luis, M., González-Hidalgo, J.C., Longares, L.A. 2010.** Is Rainfall Erosivity Increasing in the Mediterranean Iberian Peninsula? *Land Degradation & Development*, 21, 139–144.
- [10] **De Runz, C., Bellot-Gurlet, L., Guillon, E. 2008.** Typologie de l'imperfection des données archéologiques.
- [11] **Essamari, O. 2018.** Introduction au Machine Learning « Apprentissage Automatique » et ses implications pour la stabilité financière. LinkedIn.
- [12] **FAO, 1990.** Conservation des sols et des eaux dans les zones semi-arides. Rome, *Bulletin Pédologique* n°57.
- [13] **Fenta, A.A., Yasuda, H., Shimizu, K., Haregeweyn, N., Kawai, T., Sultan, D., Ebabu, K., Belay, A.S. 2017.** Spatial Distribution and Temporal Trends of Rainfall and Erosivity in the Eastern Africa Region. *Hydrological Processes*, 31, 4555–4567.
- [14] **Fertisols, 2020.** L'érosion hydrique des sols agricoles. Fiche 15 : l'érosion des sols.
- [15] **Fortin, M.J., Dale, M.R.T. 2005.** *Spatial Analysis: A Guide for Ecologists*. [DOI: 10.1017/CBO9780511542039](https://doi.org/10.1017/CBO9780511542039).

- [16] **Fotheringham, A.S., Rogerson, P.A. 2010.** The SAGE Handbook of Spatial Analysis. [DOI: 10.1111/j.1467-9787.2010.00709.6.x](https://doi.org/10.1111/j.1467-9787.2010.00709.6.x).
- [17] **Fournier, F. 1967.** Factors of Soil Erosion by Water.
- [18] Fournier Index by Using Artificial Neural Network in Central Europe. International Journal of Environmental Research and Public Health, 19, 10653. [DOI: 10.3390/ijerph191710653](https://doi.org/10.3390/ijerph191710653).
- [19] **Goetz, J.N., Brenning, A., Petschko, H., Leopold, P. 2015.** Evaluating machine learning and statistical prediction techniques for landslide susceptibility modeling. Computers & Geosciences. [DOI: 10.1016/j.cageo.2015.04.007](https://doi.org/10.1016/j.cageo.2015.04.007).
- [20] **Harsányi, E., Bashir, B., Alsilibe, F., Moazzam, M.F.U., Ratonyi, T., Alsalman, A., Széles, A., Nyeki, A., Takács, I., Mohammed, S. 2022.** Predicting Modified Cazorla, G. 2005. Planification, programmation et périodisation de l'entraînement. AREAPS.
- [21] **H. REBAI , D. RACLOT , H. BEN OUEZDOU,2012,**Érosion et Dégradation des Sols
- [22] **Innerebner, H., Schobesberger, D., Nagler, T. 2017.** Organizing Earth Observation Data inside a Spatial Data Infrastructure.
- [23] **Irvem, A., Topaloğlu, F., Uygur, V. 2007.** Estimating Spatial Distribution of Soil Loss over Seyhan River Basin in Turkey. Journal of Hydrology, 336, 30–37. **Longtin, J. 2002.** PR-GQ-011, Version 05.
- [24] **Janny, S., Petit, M., Ducroquet, M. 2022.** Introduction à l'apprentissage automatique. Culture Science de l'Ingénieur.
- [25] **Keddar, A., Boukharouba, N., Cherchali, S. 2011.** Degré d'impact des aménagements hydrauliques sur le ruissellement et le transport solide dans une région semi-aride : cas des micro-bassins dans le haut Isser, Algérie. Afrique SCIENCE, 07(3).
- [26] **Khodja, D.H. 2020.** Contribution à la gestion des ressources en eau du bassin versant de l'oued Isser à l'aide d'un système informatisé.
- [27] **Kirat, T., Lévy, J., Labarthe, F. 2022.** Équité et explicabilité des algorithmes d'apprentissage automatique : un défi technique et juridique. HALSHS-03667000.
- [28] **Koedinger, K.R., Corbett, A.T. 2006.** Cognitive Tutors: Technology Bringing Learning Science to the Classroom.
- [29] **Lloyd, C.D. 2008.** Local Models for Spatial Analysis. [DOI: 10.1111/j.1467-9671.2008.01123.x](https://doi.org/10.1111/j.1467-9671.2008.01123.x).
- [30] **Medjber, S. 2011.** Influence de la variabilité des précipitations sur le taux de sédimentation dans plusieurs barrages algériens. Revue Scientifique et Technique LJEE, n°19.
- [31] **Miquel, M., Pécher, A., Chauveau, N. 2002.** Conception d'entrepôts de données géospatiales à partir de sources hétérogènes : Exemple d'application en foresterie. [DOI: 10.3166/isi.7.3.89-111](https://doi.org/10.3166/isi.7.3.89-111).
- [32] **Mosavi, A., Ozturk, P., Chau, K.W. 2020.** Susceptibility Mapping of Soil Water Erosion Using Machine Learning Models. Water. [DOI: 10.3390/w12071995](https://doi.org/10.3390/w12071995).

- [33] **Nemmich, M.A. 2014.** Algorithme de clustering spectral pour la classification textuelle.
- [34] **Ouallali, A., Chahid, M., Oumoussad, L., et al. 2016.** Évaluation et cartographie des taux d'érosion hydrique dans le bassin versant de l'Oued Arbaa Ayacha (Rif occidental, Nord Maroc). Bulletin de l'Institut Scientifique, Rabat, Section Sciences de la Terre, n°38.
- [35] **Pal, N.R., Jain, L.C. 2005.** Advanced Techniques in Knowledge Discovery and Data Mining. Springer. [DOI: 10.1007/1-84628-183-0](https://doi.org/10.1007/1-84628-183-0).
- [36] **Peigne, J. 2013.** Services écosystémiques et processus écologiques des sols. ISARA Lyon, UVED.
- [37] **Perkowitz, M., Etzioni, O. 1997.** Towards Adaptive Web Sites: Conceptual Framework and Case Study. Artificial Intelligence, Volume 118, Issues 1-2, April 2000, Pages 245-275.
- [38] **Remini, B., Hallouche, M., Hadj-Brahim, K. 2003.** Les barrages du Maghreb face au phénomène de l'envasement. Revue Vecteur Environnement.
- [39] **Renod, J., Lahaye, J., Ladet, S., 2015.** SIG et GPS / GNSS: Concepts de base. Slideshare.
- [40] **Roose, É., Benkobi, L., Sahli, A. 2012.** Soixante années de recherches en coopération sur l'érosion hydrique et la lutte antiérosive au Maghreb. Physio-Géo, volume 06. [DOI: 10.4000/physio-geo.2319](https://doi.org/10.4000/physio-geo.2319).
- [41] **Romero, M., Boucher, D., Fortier, S. 2023.** Enseigner et apprendre à l'ère de l'intelligence artificielle. CDFT.
- [42] **Sheeren, D., Dubois, A., Joost, S. 2006.** Discovering Rules with Genetic Algorithm to Classify Remotely Sensed Data. [DOI: 10.1109/IGARSS.2006.1005](https://doi.org/10.1109/IGARSS.2006.1005).
- [43] **Sur, F. 2024.** Introduction à l'apprentissage automatique.
- [44] **Silvia liberata Uilo, Alessandro Sebastianelli, Amrita Mohan, Shaik Ejaz Ahamed, 2020,** A New Mask R-CNN Based Method for Improved Landslide Detection
- [45] **Taye, G., Yirgu, T., Zegeye, A.D., Steenhuis, T.S., Hurni, H. 2013.** Effects of Land Use, Slope Gradient, and Soil and Water Conservation Structures on Runoff and Soil Loss in Semi-Arid Northern Ethiopia. Geomorphology. [DOI: 10.1080/02723646.2013.832098](https://doi.org/10.1080/02723646.2013.832098).
- [46] **United Nations, 2002.** In Yousef, 2011.
- [47] **Yles, F., Abidi, A., Bensaoula, N., Kadri, A. 2012.** Quantification et modélisation du transport solide dans le bassin-versant de l'oued Saïda (hauts plateaux algériens). Sécheresse, vol 23, n°4.
- [48] **zouaoua mehdi, 2021.** Etude de la variabilité spatio-temporelle des précipitations annuelles dans le bassin versant d'Isser.