

**DÉVELOPPEMENT D'UN OUTIL POUR L'INTÉGRATION DE  
L'ANALYSE DES ALÉAS ASSOCIÉS À LA DYNAMIQUE  
GÉOMORPHOLOGIQUE DES COURS D'EAU (VOLET B)**  
Projet R800.1

Marco Gava<sup>1</sup>, Pascale Biron<sup>1</sup>, William Massey<sup>1</sup> et Thomas Buffin-  
Bélanger<sup>2</sup>

1. Département de géographie, urbanisme et environnement,  
Université Concordia
2. Département de biologie, chimie et géographie, Université du  
Québec à Rimouski

Réalisé pour le compte du  
Ministère des Transports et de la Mobilité durable du Québec  
Direction générale de la planification et de la gestion des infrastructures  
et  
Direction générale de la gestion des actifs routiers et de l'innovation

Mars 2023

DÉVELOPPEMENT D'UN OUTIL POUR L'INTÉGRATION DE L'ANALYSE DES ALÉAS  
ASSOCIÉS À LA DYNAMIQUE GÉOMORPHOLOGIQUE DES COURS D'EAU – PROJET DE  
RECHERCHE (R800.1)

La présente étude a été réalisée à la demande du ministère des Transports et de la Mobilité durable et a été financée par le Ministère.

Les opinions exprimées dans le présent rapport n'engagent que la responsabilité de leurs auteurs et ne reflètent pas nécessairement les positions du ministère des Transports et de la Mobilité durable.

## **REMERCIEMENTS ET NOTE SUR L'AUTEUR(S)**

Nous tenons à remercier la chargée de projet pour ce projet, Ariane Bouffard, biologiste au MTMD, pour son soutien, sa très grande disponibilité et son leadership. Nous remercions également Laurent Saint-Jacques, géographe au MTMD, pour son aide dans la gestion du projet ainsi que Tristan Caron, géomorphologue au MTMD, qui nous a fourni des renseignements précieux sur les dommages causés aux infrastructures routières dans le bassin de la rivière Cascapédia par des processus fluviaux. Merci aux stagiaires (Nicolas Dos Santos et Marc-Antoine Cyr) qui ont aidé à compléter le travail de terrain nécessaire à ce projet ainsi qu'à Alex Delisle Thibeault et Samuel Laroche, auxiliaires de recherche à l'UQAR, pour leur contribution à la collecte de données de terrain pour le bassin versant de la Cascapédia. Merci également aux membres du comité de suivi qui ont pris le temps de lire et de commenter le rapport, ce qui a permis de le rendre aussi clair et informatif que possible.

## SOMMAIRE

Dans le cadre du projet de recherche R800.1 « Développement d'un outil pour l'intégration de l'analyse des aléas associés à la dynamique géomorphologique des cours d'eau » un outil ArcGIS Pro utilisant le modèle d'apprentissage automatique de forêt aléatoire a été développé pour détecter à grande échelle la probabilité que des aléas fluviaux se produisent le long des cours d'eau. Cet outil entraîne trois modèles pour détecter : la présence d'inondations, la présence de mobilité et le type d'érosion (migration latérale ou incision) qui se produira si un potentiel de mobilité est détecté. Il s'agit d'un outil d'évaluation rapide qui peut être exécuté pour l'ensemble du territoire méridional du Québec à faible coût tout en étant facilement modifiable pour être entraîné sur certains bassins versants ou comme modèle global à utiliser sur des bassins versants sans données. Cet outil permet de détecter les sections à forte probabilité d'aléas qui devraient être étudiées plus en profondeur avec des modélisations hydrauliques et des analyses géomorphologiques.

Ce rapport présente la méthodologie utilisée pour entraîner les modèles ainsi que les résultats, plus précisément :

- I. Les variables clés influençant la présence des aléas fluviaux
- II. La méthodologie choisie pour prédire les aléas
- III. Des cartes des endroits où ces aléas sont susceptibles de se produire dans 3 bassins versants pilotes (Cascapédia, Du Gouffre, Nicolet)
- IV. Un test du modèle sur un bassin versant sur lequel il n'a pas été entraîné (Coaticook)
- V. Les limites du modèle et les points importants à considérer lors de l'interprétation des résultats

La source pour les variables clés est principalement le CRHQ (Cadre de référence hydrologique du Québec) puisqu'il s'agit d'un ensemble de données disponibles partout au Québec pour toutes les tailles de rivières. Toutefois, en raison d'un manque de données sur le débit des plus petits cours d'eau qui proviennent de l'Atlas Hydroclimatique du Québec, seules les rivières dont l'aire de drainage est supérieure à 50 km<sup>2</sup> ont été incluses dans l'étude. Des recherches futures pourraient se pencher sur la possibilité d'adapter le modèle à des cours d'eau plus petits en excluant la puissance spécifique du cours d'eau (qui dépend du débit) mais en incluant la pente et la largeur comme approximation puisque ces informations sont disponibles dans le CRHQ pour les plus petits cours d'eau.

Les modèles ont des durées d'exécution courtes (20 minutes au total par bassin versant) et peuvent aussi être facilement modifiés pour incorporer des variables provenant d'autres sources de données et être mis à jour pour se conformer aux nouvelles versions du CRHQ. Au fur et à mesure des améliorations du CRHQ, certains des scripts développés dans ce projet (extraction du débit, estimation du confinement des rivières) ne seront plus requis, ce qui rendra le fonctionnement de l'outil d'autant plus simple et rapide.



Dans les trois bassins versants pilotes, des observations sur le terrain sur la mobilité et l'inondabilité ont été combinées à d'autres sources d'information comme des rapports d'incidents concernant les infrastructures routières afin de produire une banque de données pour entraîner et valider le modèle. Au total, 1807 points ont été classés pour les inondations, 1542 pour la mobilité et 847 pour le type d'érosion.

Pour le modèle de **présence d'inondation**, les valeurs de précision totales sont très bonnes (84,3 %) avec de faibles valeurs de surentraînement et aucun biais majeur pour l'une ou l'autre des classes (présence d'inondation par rapport à l'absence de présence). L'incertitude ne varie pas non plus beaucoup entre les 3 bassins versants pilotes, indiquant que l'outil fonctionne aussi bien pour les trois bassins versants. Il est important de considérer que ces résultats sont relatifs à la taille de la rivière et qu'ils indiquent seulement que le débit a une probabilité de sortir du chenal, et non qu'il inondera une grande surface ou atteindra les infrastructures routières.

Le modèle de **présence de mobilité** a également une bonne valeur de précision totale (84,7 %) et un faible surentraînement, bien que l'incertitude varie davantage entre les bassins versants. Ce modèle possède un biais vers la présence de mobilité, probablement en raison de l'entraînement du modèle sur des bassins versants très mobiles (Du Gouffre et Cascapédia) et cela devrait être pris en compte lors de l'interprétation des résultats.

Le modèle pour le **type d'érosion** a des valeurs de précision totale très élevées (94,7 %) avec de faibles valeurs de surentraînement et une incertitude assez constante d'un bassin versant à l'autre. Ce modèle a un biais vers la classe de migration latérale par rapport à la classe d'incision. Cela est probablement dû au fait que les modèles sont entraînés sur les plus grandes rivières en aval où la migration latérale a tendance à être le type dominant d'érosion par rapport aux cours d'eau de tête de bassin pour lesquels il y a un manque de données de débit dans l'Atlas Hydroclimatique.

Les tests effectués sur un bassin-versant non-entraîné (rivière Coaticook) sont encourageants, mais il est important de poursuivre cette validation pour un grand nombre de secteurs de contextes géomorphologiques variables.

## **SYNTHÈSE DES CONCLUSIONS ET RECOMMANDATIONS**

Les conclusions générales du projet indiquent que les modèles d'intelligence artificielle peuvent être adaptés aux données du Québec pour prédire les lieux d'occurrence potentielle d'aléas fluviaux pour une grande partie du territoire tout en présentant des résultats à une échelle suffisamment fine pour prendre des décisions éclairées sur le potentiel de risque pour les infrastructures routières. L'application des outils développés dans trois bassins-versants contrastés d'un point de vue géomorphologique a permis de bien tester la robustesse des modèles de prévision de l'inondation, de la mobilité et du type de mobilité, le cas échéant. L'étude s'est concentrée sur les plus grands cours d'eau où la migration latérale est le type dominant d'érosion par rapport à l'incision, ce qui résulte en un modèle qui prédit mieux la migration latérale que l'incision dans sa version actuelle. Le choix des bassins versants a également eu un impact, avec deux des trois bassins versants principalement composés de rivières mobiles. Les modèles présentaient également des biais envers certaines classes qui doivent être pris en compte lors de l'interprétation des résultats, mais ces biais se sont avérés relativement mineurs sauf pour la mobilité où il y avait un fort biais vers la classe mobile. Le modèle a prédit correctement les zones stables avec une précision de 77 %, tandis que pour la classe mobile, il était correct 92 % du temps.

Il est recommandé que le MTMD utilise cet outil à des fins d'évaluation préliminaire pour déterminer les infrastructures routières qui pourraient potentiellement être endommagées, tout en procédant à des évaluations plus rigoureuses à l'échelle locale pour déterminer la meilleure façon d'atténuer les aléas fluviaux. Il est aussi recommandé d'améliorer le modèle en continuant à ajouter des données d'entraînement qui peuvent être collectées lors de visites sur le terrain, après des inondations, par des analyses hydrogéomorphologiques ou après avoir réalisé des évaluations de la mobilité. Il est important de recueillir ces données d'entraînement dans une variété de bassins versants avec des caractéristiques géomorphologiques différentes afin d'adapter le modèle au plus grand nombre d'environnements possibles au Québec. Lors de la collecte des données d'entraînement, il est important de s'assurer qu'un nombre égal de points soit collecté dans les aires exposées aux aléas et dans les aires dépourvues d'aléas, afin que le modèle soit en mesure d'identifier ces deux types de zones et de bien les différencier.

## TABLE DES MATIÈRES

LISTE DES TABLEAUX .....	9
LISTE DES FIGURES .....	10
GLOSSAIRE.....	15
1. INTRODUCTION .....	18
1.1 Contexte.....	18
1.2 Évaluation de l'inondabilité et la mobilité : tour d'horizon des méthodes et variables clés .....	18
2 PROBLÉMATIQUE .....	21
3 MÉTHODOLOGIE .....	22
3.1 Données d'entrée pour le modèle.....	24
3.2 Bassins versants pilotes .....	30
3.3 Préparation des données .....	31
3.4 Entraînement du modèle supervisé .....	33
3.4.1 Collecte des données de validation.....	33
3.4.2 Classification des données .....	35
3.4.3 Sélection des variables finales des modèles.....	38
3.4.4 Procédure d'entraînement .....	41
3.5 Prédiction des aléas pour le réseau fluvial d'un bassin versant .....	42
4 ANALYSE DES RÉSULTATS .....	44
4.1 Sélection des paramètres des modèles.....	44
4.1.1 Nombre d'arbres .....	44
4.1.2 Minimum de points de données par nœud .....	44
4.2 Scores de précision des modèles.....	47
4.3 Validation croisée .....	48
4.4 Matrice de confusion .....	48
5. INTERPRÉTATION DES RÉSULTATS.....	52
5.1 Importance de chaque variable .....	52
5.1.1 Inondation.....	53

DÉVELOPPEMENT D'UN OUTIL POUR L'INTÉGRATION DE L'ANALYSE DES ALÉAS  
ASSOCIÉS À LA DYNAMIQUE GÉOMORPHOLOGIQUE DES COURS D'EAU – PROJET DE  
RECHERCHE (R800.1)

5.1.2 Mobilité .....	55
5.1.4 Type d'érosion .....	57
5.2 Résultats de probabilités d'aléas fluviaux pour les 3 bassins pilotes ...	59
5.3 Validation à l'aide des bases de données du MTMD.....	69
5.4 Test sur la Coaticook (bassin versant non entraîné) .....	71
5.4.1 Description de la zone d'étude .....	72
5.4.2 Analyse de sensibilité.....	73
5.4.3 Résultats de probabilités d'aléas fluviaux pour la Coaticook .....	75
6. MISE EN ŒUVRE ET RETOMBÉES, BÉNÉFICES POUR LE MTMD .....	79
7. CONCLUSION.....	82
ANNEXE A.....	86
Annexe A.1 Type d'érosion et mobilité.....	87
Annexe A.2 Inondation .....	97
ANNEXE B .....	101
Annexe B.1 Étapes d'exécution.....	102
Annexe B.2 Étapes d'entraînement.....	104
Annexe B.3 Modifications pour être compatibles avec les futures versions du CRHQ .....	105
ANNEXE C .....	106
ANNEXE D .....	111
RÉFÉRENCES .....	113

## LISTE DES TABLEAUX

<b>Tableau 1:</b> Intrants, banques de données et leur source pour le déploiement du modèle.....	25
<b>Tableau 2:</b> Liste des variables utilisées par le modèle.....	32
<b>Tableau 3:</b> Ensembles de données de validation utilisées dans le projet.....	34
<b>Tableau 4:</b> Schéma de classification utilisé lors de la classification des points de référence pour l'inondation dans le champ "inondation". .....	35
<b>Tableau 5:</b> Schéma de classification utilisé lors de la classification des points de référence pour la mobilité dans le champ "mobilité". .....	36
<b>Tableau 6:</b> Schéma de classification utilisé lors de la classification des points de référence pour le type d'érosion dans le champ "erosion". .....	36
<b>Tableau 7:</b> Tableau des variables conservées dans le modèle pour l'étape d'entraînement/prédiction .....	41
<b>Tableau 8:</b> Tableau montrant les nouveaux champs ajoutés aux points de référence par l'étape de prédiction du modèle pour chaque classe prédite, X étant remplacé par « Inondation », « Mobilité » et « Erosion » sans accents pour ne pas provoquer d'erreurs dans ArcGIS pour un total de 9 nouveaux champs. ....	43
<b>Tableau 9:</b> Résultats de l'entraînement, de la validation et de la précision totale pour le modèle entraîné sur les trois bassins versants pilotes combinés. ....	48
<b>Tableau 10:</b> Résultats de la précision totale du modèle pour chaque bassin versant pilote. ....	48
<b>Tableau 11:</b> Tableau explicatif de la matrice de confusion .....	50
<b>Tableau 12:</b> Matrice de confusion pour les inondations .....	51
<b>Tableau 13:</b> Matrice de confusion pour la mobilité.....	51
<b>Tableau 14:</b> Matrice de confusion pour l'érosion .....	51
<b>Tableau 15:</b> Résultats finaux et conclusions pour chaque modèle.....	84

## LISTE DES FIGURES

<b>Figure 1:</b> Les différentes zones d'un bassin versant, adapté de River Systems and Fluvial Landforms - Geology (U.S. National Park Service), 2022 .....	21
<b>Figure 2:</b> Schéma expliquant le fonctionnement du modèle forêt aléatoire (modifié de Parmar et al., 2019).....	24
<b>Figure 3:</b> Comparaison du polygone de fond de vallée extrait à des simulations d'inondation à grande échelle à l'aide de LISFlood et de la méthodologie de Choné et al. (2021) en utilisant une valeur correspondant à environ une inondation de 20 ans .....	27
<b>Figure 4:</b> Exemple de script de calcul de confinement. A) transects de taille maximale générés tous les 25 m le long du réseau hydrographique. B) transects coupés à l'étendue de la vallée. C) intersection des transects découpés et de la rivière superposée aux points .....	28
<b>Figure 5:</b> La figure montre le fonctionnement de l'outil d'intersection des tronçons modélisés de l'Atlas Hydroclimatique (lignes rouges) et les tronçons UEA de la CRHQ (lignes bleues). Les points sont les points d'intersection et sont colorés en fonction de l'UEA .....	29
<b>Figure 6:</b> Emplacement des trois bassins versants pilotes; A) rivière Cascapédia; B) rivière Nicolet; C) rivière Du Gouffre. ....	31
<b>Figure 7:</b> Distribution des points classés dans chaque catégorie par aléa fluvial et par bassin versant. ....	38
<b>Figure 8:</b> Les variables regroupées en fonction de la distance euclidienne entre leurs corrélations. Plus la hauteur (distance euclidienne entre leurs corrélations) des lignes reliant les variables est faible, plus elles sont corrélées. ....	40
<b>Figure 9:</b> Sensibilité du modèle d'inondation au nombre minimum de points de données par nœud. Le seuil de 2% est identifié par la ligne rouge verticale. .	45
<b>Figure 10:</b> Sensibilité du modèle de mobilité au nombre minimum de points de données par nœud .....	46
<b>Figure 11:</b> Sensibilité du modèle de type d'érosion au nombre minimum de points de données par nœud .....	47
<b>Figure 12:</b> Graphique ROC AUC pour les trois modèles. ....	52
<b>Figure 13:</b> Diminution moyenne de l'impureté pour chaque variable du modèle d'inondation, les lignes noires représentant l'écart type pour chaque variable. ....	54
<b>Figure 14:</b> Diminution moyenne de la précision pour chaque variable du modèle d'inondation, les lignes noires représentant l'écart type pour chaque variable.....	55
<b>Figure 15:</b> Diminution moyenne de l'impureté pour chaque variable du modèle de mobilité, les lignes noires représentant l'écart type pour chaque variable. ....	56

<b>Figure 16:</b> Diminution moyenne de la précision pour chaque variable du modèle de mobilité, les lignes noires représentant l'écart type pour chaque variable.....	57
<b>Figure 17:</b> Diminution moyenne de l'impureté pour chaque variable du modèle de type d'érosion, les lignes noires représentant l'écart type pour chaque variable.....	58
<b>Figure 18:</b> Diminution moyenne de la précision pour chaque variable du modèle de type d'érosion, les lignes noires représentant l'écart type pour chaque variable. ....	59
<b>Figure 19:</b> Carte des prévisions des inondations avec le modèle pour le bassin versant de la Cascapédia .....	60
<b>Figure 20:</b> Carte des prévisions des inondations avec le modèle pour le bassin versant de la Nicolet .....	61
<b>Figure 21:</b> Carte des prévision des inondations avec le modèle pour le bassin versant de la Du Gouffre.....	62
<b>Figure 22:</b> Prédiction de classe d'inondation par bassin versant .....	62
<b>Figure 23:</b> Carte des prévisions de mobilité du modèle pour le bassin versant de la Cascapédia.....	63
<b>Figure 24:</b> Carte des prévisions de mobilité du modèle pour le bassin versant de la Nicolet.....	64
<b>Figure 25:</b> Carte des prévisions de mobilité du modèle pour le bassin versant de la Du Gouffre .....	65
<b>Figure 26:</b> Prédiction de classe de mobilité par bassin versant .....	65
<b>Figure 27:</b> Carte des prévisions du type d'érosion du modèle pour le bassin versant de la Cascapédia .....	66
<b>Figure 28 :</b> Carte des prévisions du type d'érosion du modèle pour le bassin versant de la Nicolet .....	67
<b>Figure 29:</b> Carte des prévisions du type d'érosion du modèle pour le bassin versant de la Du Gouffre .....	68
<b>Figure 30:</b> Prédiction de classe de type d'érosion par bassin versant .....	68
<b>Figure 31:</b> Carte des références aux dommages aux infrastructures liés aux processus fluviaux dans la base de données du MTMD comparée aux résultats de la probabilité de mobilité du modèle (Nicolet). ....	70
<b>Figure 32:</b> Carte des références aux dommages aux infrastructures liés aux processus fluviaux dans la base de données du MTMD comparée aux résultats de la probabilité de mobilité du modèle (Cascapédia).....	71
<b>Figure 33:</b> Zones d'étude de la rivière Coaticook avec les points de référence du CRHQ et le MNT.....	73

<b>Figure 34:</b> Sensibilité des résultats sur le bassin de la Coaticook au nombre minimum de points de données par nœud pour la présence d'inondation montrant à la fois la précision et le pourcentage de points au-dessus du niveau de confiance de 55% .....	74
<b>Figure 35:</b> Sensibilité des résultats sur le bassin de la Coaticook au nombre minimum de points de données par nœud pour la présence de mobilité montrant à la fois la précision et le pourcentage de points au-dessus du niveau de confiance de 55% .....	75
<b>Figure 36:</b> Carte des prévisions de présence d'inondations du modèle pour le bassin versant de la Coaticook .....	76
<b>Figure 37:</b> Prédictions de classe d'inondation pour la Coaticook.....	77
<b>Figure 38:</b> Carte de prévisions de présence de mobilité du modèle pour le bassin versant de la Coaticook .....	77
<b>Figure 39:</b> Prédictions de classe de mobilité pour le Coaticook.....	78
<b>Figure 40:</b> Carte de prévisions du type d'érosion du modèle pour le bassin versant de la Coaticook .....	78
<b>Figure 41:</b> Prédictions de classe de type d'érosion pour le Coaticook.....	79
<b>Figure 42:</b> Carte thermique des points prédits comme étant mobiles, pondérés par la confiance de la prédiction, avec les conflits potentiels entre les rivières mobiles et les routes encerclés en rouge. ....	81
<b>Figure A1:</b> Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Nicolet : 72.0290349°W 45.7900955°N) .....	88
<b>Figure A2:</b> Deux photos prises à l'apex du méandre sur un site où le point a 80% de chances d'être mobile dans la figure A1 .....	88
<b>Figure A3:</b> Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Nicolet : 72.5375146°W 46.1540019°N) .....	89
<b>Figure A4:</b> Deux photos du site prises à partir de Google Earth et sur le site à partir du pont dans la figure A3 .....	89
<b>Figure A5:</b> Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Nicolet : 72.0286840°W 45.7667794°N) .....	90
<b>Figure A6:</b> Deux photos du site prises sur le site depuis la route dans la figure A5 .....	90
<b>Figure A7:</b> Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Cascapédia : 65.9309543°W 48.2969596°N) .....	91
<b>Figure A8:</b> Deux photos du site de la figure A7. Première photo par le MTQ, 2018 la seconde par l'UQAR 2022 .....	91



<b>Figure A9:</b> Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Cascapédia : 65.9576840°W 48.3110689°N).....	92
<b>Figure A10:</b> Deux photos du site de la figure A9. Première photo par le MTQ, 2018 la seconde par l'UQAR 2022.....	92
<b>Figure A11:</b> Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Cascapédia : 66.1466730°W 48.5648444°N) .....	93
<b>Figure A12:</b> Deux photos du site de la figure A11 par l'UQAR 2022 .....	93
<b>Figure A13:</b> Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Du Gouffre : 70.5057976°W 47.4476068°N) .....	94
<b>Figure A14:</b> Une photo du site de la figure A13.....	94
<b>Figure A15:</b> Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Du Gouffre : 70.5164161°W 47.4411945°N).....	95
<b>Figure A16:</b> Deux photos du site de la figure A15 .....	95
<b>Figure A17:</b> Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Du Gouffre : 70.5522495°W 47.5793112°N).....	96
<b>Figure A18:</b> Deux photos du site de la figure A17 .....	96
<b>Figure A19:</b> Carte montrant les prévisions d'inondation du modèle. Les chiffres représentent la probabilité de la classe (Absence d'inondation ou Inondation). (Nicolet : 71.8664060°W 46.1452991°N) .....	97
<b>Figure A20:</b> Carte montrant les prévisions d'inondation du modèle. Les chiffres représentent la probabilité de la classe (Absence d'inondation ou Inondation). (Cascapédia : 65.9307575°W 48.2969121°N) .....	98
<b>Figure A21:</b> Carte montrant les prévisions d'inondation du modèle. Les chiffres représentent la probabilité de la classe (Absence d'inondation ou Inondation). (Du Gouffre : 70.5166415°W 47.4410165°N).....	99
<b>Figure A22:</b> Carte montrant les prévisions d'inondation du modèle. Les chiffres représentent la probabilité de la classe (Absence d'inondation ou Inondation). (Du Gouffre : 70.5123875°W 47.5211689°N) .....	100
<b>Figure B1:</b> Flux de travail de l'outil .....	104
<b>Figure B2:</b> Exemple de ligne à modifier pour changer les champs inclus dans le modèle, la partie qui doit être modifiée est la liste des champs inclus dans les crochets. ....	105

<b>Figure C1:</b> répartition des aires de drainage pour les points dont l'aire de drainage est inférieure à 50km <sup>2</sup> .....	107
<b>Figure C2:</b> Prédiction de classe d'inondation par bassin versant pour les points avec aire de drainage inférieures à 50 km <sup>2</sup> .....	108
<b>Figure C3:</b> Prédiction de classe de mobilité par bassin versant pour les points avec aire de drainage inférieures à 50 km <sup>2</sup> .....	108
<b>Figure C4:</b> Prédiction de classe de type d'érosion par bassin versant pour les points avec aire de drainage inférieures à 50 km <sup>2</sup> .....	109
<b>Figure C5:</b> Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Nicolet : 71.6207918°W 45.8890004°N) .....	109
<b>Figure C6:</b> Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Casapédia : 66.2446110°W 48.7449664°N).....	110
<b>Figure D1:</b> Carte de disponibilité des données du CRHQ .....	112

## GLOSSAIRE

**Aléa fluvial** : Dans ce rapport, le terme « aléa fluvial » désigne les inondations, l'incision, la migration latérale et d'autres événements qui ont un impact sur le niveau de mobilité d'une rivière.

**Avulsion** : Processus naturel de relocalisation du tracé du cours d'eau, souvent en lien avec la dynamique des méandres.

**Classification d'ensemble** : Une méthode d'apprentissage automatique qui utilise plusieurs instances du même modèle, chacune utilisant un sous-ensemble de variables d'entrée et de données d'entraînement.

**Coupure de méandre par dissection**: Processus dans lequel le chenal principal change en faveur d'un chenal secondaire près d'un méandre en raison d'inondation (*Dépret et al., 2017*).

**Entraînement** : Le processus consistant à donner à un modèle de forêt aléatoire des points classifiés et les variables associées à ce point afin qu'il puisse déterminer des modèles dans ces variables qui lui permettent de déterminer la classe à laquelle ce point devrait être assigné sur la base des variables.

**Forêt aléatoire** : Un modèle d'apprentissage automatique qui génère plusieurs arbres de décision utilisant chacun un sous-ensemble aléatoire des variables d'entrée et de données d'entraînement pour prédire la classe à attribuer à chaque point, la décision finale étant la classe déterminée par la majorité des modèles.

**Impureté** : Mesure de précision qui quantifie le nombre de points de données d'entraînement de la bonne classe présents dans un nœud par rapport aux points de données de la mauvaise classe.

**Incision**: Érosion verticale d'une rivière engendrant l'abaissement de l'élévation du lit.

**Migration latérale** : Érosion horizontale de la berge d'une rivière entraînant le déplacement horizontal d'un chenal, en général en lien avec la migration d'une boucle de méandre.

**MNT / LiDAR** : Modèle numérique de terrain produit par des données aéroportées LiDAR (« Light Detection And Ranging » - élévations de haute résolution).

**Nœud** : Sous-ensemble des données d'entraînement déterminé à l'aide d'un seuil de l'une des variables d'entraînement. Un arbre de décision sera constitué de plusieurs de ces sous-ensembles, chacun étant déterminé par des seuils différents. Le premier contiendra toutes les données d'entraînement, tandis que les suivants en contiendront une quantité décroissante afin de tenter d'obtenir

un sous-ensemble ne comprenant que des données d'entraînement d'une classe homogène.

**Précision** : Métrique utilisée pour décrire le pourcentage de prédictions du modèle qui étaient correctes selon les données de validation.

**Processus fluvial** : Processus variés d'érosion, sédimentation et d'inondation qui se produisent dans la plaine alluviale.

**Puissance spécifique** : Mesure d'énergie fournie par l'écoulement d'un cours d'eau par unité de surface ( $W/m^2$ ). La puissance spécifique est un paramètre représentatif des conditions hydrauliques et permet de prédire le taux de transport de sédiments.

**Surentraînement** : Métrique utilisée pour déterminer la différence de précision entre les données d'entraînement par rapport aux données de validation.

**Tronçon** : Unité spatiale qui réfère à une partie du cours d'eau ou du réseau de drainage. Traduction du terme « reach » communément employé en géomorphologie fluviale anglo-saxonne (*Buffin-Bélanger et al., 2015*).

**Tronçon homogène** : Segment de rivière qui a des caractéristiques similaires (pente, largeur, niveaux de confinement...) (*Biron et al, 2014*).

## 1. INTRODUCTION

### 1.1 Contexte

L'objectif global du projet " Développement d'un outil pour l'intégration de l'analyse des aléas associés à la dynamique géomorphologique des cours d'eau (R800.1) " est l'identification des contraintes géomorphologiques le long du tracé proposé de la route 138 et le développement d'outils SIG pour caractériser les aléas géomorphologiques autour des routes à grande échelle. Le projet comporte ainsi deux objectifs distincts permettant de présenter pour des raisons pratiques et pragmatiques deux rapports autoportants :

1. Un rapport se concentrant sur la Basse-Côte-Nord consistant en une étude géomorphologique des aléas potentiels le long de la route 138 et des recommandations pour les éviter ou les atténuer (Volet A).
2. Un rapport se concentrant sur les développements d'un outil SIG consistant à l'élaboration et à la validation d'une méthodologie et d'un outil SIG pour caractériser à grande échelle les aléas fluviaux potentiels pour les routes du Québec (Volet B).

Le présent rapport détaille la méthodologie et expose les résultats du deuxième volet du projet.

### 1.2 Évaluation de l'inondabilité et la mobilité : tour d'horizon des méthodes et variables clés

La mobilité et l'inondabilité sont deux **processus** fondamentaux et naturels de la dynamique des cours d'eau contribuant aux services écosystémiques qu'ils rendent (*Florsheim et al., 2008*). Ces processus contribuent aussi à générer du risque lorsque des enjeux anthropiques se trouvent dans les espaces où ils surviennent. Alors que l'inondabilité est généralement considérée dans les cadres réglementaires et que les outils pour délimiter les espaces d'inondabilité sont reconnus, elle repose en général sur des modèles hydrauliques nécessitant une grande puissance de traitement ainsi que des variables telles que la bathymétrie de la rivière qui ne sont pas disponibles à l'échelle du territoire (*Biron et al., 2014 ; Choné et al., 2021*). Ces modèles produisent des cartes très précises des espaces inondables, mais leur utilisation à l'échelle de la province prendrait beaucoup de temps et serait coûteuse. Des cartes d'inondation à grande échelle utilisant une modélisation hydraulique simplifiée sont déjà en cours d'élaboration au Québec dans le cadre du projet Info-Crue, avec de bons résultats pour l'estimation de la bathymétrie à partir des Modèles Numériques de Terrain (**MNT**) à haute résolution **LiDAR** (*Choné et al., 2021*). Ces cartes seront utilisées pour améliorer notre outil. Cependant, notre outil se veut le moyen le plus simple et le moins coûteux d'identifier les sites potentiels d'aléas pour lesquels des outils

ou des analyses plus précis devraient être utilisés, et n'utilisera donc pas la modélisation hydraulique.

Contrairement à l'inondabilité, il y a moins de consensus scientifique sur les méthodes de détermination de la mobilité des rivières, considérant le nombre élevé de variables et les différentes échelles d'analyses (bassin versant, **tronçon**, et local) nécessaires pour bien l'évaluer. De nombreux outils ont été développés dans les dernières années pour tenter de quantifier le niveau de mobilité en utilisant différentes méthodologies et variables (Beechie & Imaki, 2014; Kasprak et al., 2015; Wheaton et al., 2015; Belletti et al., 2017; Cavalli et al., 2017; Horacio et al., 2017; Crema & Cavalli, 2018; MELCCFP, 2023; Wheeler et al., 2022). Parmi ces outils, on distingue d'un côté ceux utilisant les tracés historiques pour évaluer de manière prospective les espaces de mobilité (Piégay et al. 2005; Biron et al. 2014) et de l'autre les outils qui utilisent les variables morphométriques, géomorphologiques et hydrauliques pour exposer la mobilité potentielle d'un cours d'eau. Dans ce projet, la deuxième approche est préconisée parce qu'elle se concentre sur l'utilisation de données plus facilement accessibles à l'échelle d'un territoire (LiDAR, CRHQ – Cadre de Référence Hydrologique du Québec) et cette partie expose les principales variables et méthodes considérées dans la littérature selon cette approche.

La **puissance spécifique** du cours d'eau ( $\omega$ ), initialement dérivé par Bagnold (1966), est l'un des indicateurs de mobilité les plus utilisés, car elle intègre des variables géomorphologiques et hydrauliques clés, notamment la pente ( $S$ ), le débits de récurrence 2 ans (considérés équivalents au débit plein bord) ( $Q$ ), et la largeur de la rivière ( $W$ ), pour déterminer le niveau d'énergie que l'eau exerce sur les berges et le lit de la rivière (Garther et al., 2015; Bizzi & Lerner, 2015; Garther, 2016; Anderson et al., 2017).

$$\omega = \frac{\rho g Q S}{W}$$

Où  $\rho$  est la densité de l'eau, et  $g$  est l'accélération due à la gravité. La simple utilisation de l'énergie exercée sur les berges et le lit ne tient pas compte des nombreuses autres variables qui influencent la mobilité car elle se concentre uniquement sur les variables locales (pente locale, largeur, et débit) (Parker & Davey, 2023). Il est également important de considérer si les berges sont érodables ou non, car les berges rocheuses peuvent résister à beaucoup plus d'énergie avant de s'éroder par rapport à une berge sableuse. Le matériau des berges et le lit de la rivière sont donc des variables importantes pour quantifier la mobilité des rivières (Fryirs & Brierley, 2018).

Des variables à l'échelle du bassin versant et du tronçon jouent aussi un rôle important dans la détermination de la mobilité en aval (Leopold, 1962; Fryirs & Brierley, 2018; Buffington & Montgomery, 2021; Sear et al., 2021). La largeur de la vallée traversée par la rivière par rapport à la largeur de la rivière est une variable clé à l'échelle du tronçon pour l'évaluation de la mobilité potentielle ainsi que le potentiel d'inondation. Les rivières très confinées ont moins de place pour migrer

latéralement, ce qui peut avoir pour conséquence que l'énergie s'exerce davantage sur le lit. Dans ce cas, l'**incision** (mobilité verticale) sera le principal **aléa fluvial** au lieu de la migration (mobilité latérale) (Demers et al., 2014; Fryirs et al., 2016; Fryirs & Brierley, 2018; Wheeler et al., 2022). Ces types de rivières limitent également le potentiel d'inondation, car le confinement de la rivière retient le débit à l'intérieur des limites du chenal. La différence entre la largeur de la rivière et celle de la vallée peut être quantifiée à l'aide d'un ratio de confinement, qui peut être appliqué au niveau du tronçon ou au niveau local. Avec les données disponibles au Québec, trois ratios de confinement ont pu être calculés pour ce projet en utilisant la méthodologie de Beechie et Imaki (2014) qui ne nécessite que la largeur de la rivière au niveau plein bord et la largeur de la vallée pour être calculée.

$$\text{Ratio de confinement local} = \frac{\text{largeur de la vallée}}{\text{largeur plein bord}}$$

$$\text{Ratio de confinement tronçon} = \frac{\text{moyenne de la largeur de la vallée du tronçon}}{\text{moyenne de la largeur plein bord du tronçon}}$$

$$\text{Ratio de confinement relatif} = \frac{\text{ratio de confinement local}}{\text{ratio de confinement tronçon}}$$

Le ratio de confinement du tronçon est utilisé pour déterminer le type de paysage dans laquelle la rivière se trouve (confiné, partiellement confiné, non confiné), tandis que le ratio local indique le niveau de confinement dans un rayon de 50 m autour de chaque point d'intérêt. Le ratio relatif permet de déterminer si cet endroit est plus ou moins confiné que la moyenne de son tronçon. Cela permet d'identifier les zones où la vallée s'élargit ou se contracte dans un tronçon, ce qui peut modifier les aléas fluviaux locaux.

Une autre variable importante à l'échelle du tronçon est sa sinuosité, définie dans le CRHQ comme suit :

$$\text{Sinuosité} = \frac{\text{Longueur du tronçon}}{\text{Longueur si c'était une ligne droite}}$$

Cette variable est importante car les rivières très sinueuses ont tendance à exercer plus d'énergie sur leurs berges que sur leur lit le long des méandres et les rivières plus linéaires ont tendance à exercer plus d'énergie sur le lit (Beechie et al., 2006). Les rivières très sinueuses sont également plus susceptibles de provoquer des inondations car les zones où se trouvait le chenal précédemment ont tendance à former des plaines alluviales de faible altitude qui sont facilement inondées (Demers et al., 2014). En combinant la sinuosité avec le ratio de confinement du tronçon, c'est possible de déterminer si la sinuosité est associée à une **migration latérale** de plaine alluviale ou si elle est liée à des contraintes géologiques, par exemple la présence de terrasses non érodables.

Finalement, puisque les morphologies et dynamiques des rivières sont le reflet d'un équilibre entre la force exercée par le débit de la rivière et la quantité de

sédiments qu'elles transportent, la quantité et la taille des sédiments transportées de l'amont peuvent influencer les aléas fluviaux qui se produiront en aval (Crema & Cavalli, 2018; Buffington & Montgomery, 2021). Pour s'assurer que l'outil soit simple à utiliser, des proxys ont été utilisés pour déterminer les conditions en amont à échelle du bassin versant en se servant de données disponibles dans le CRHQ. Les quatre proxys fournis sont : la pente moyenne en amont, le ratio de drainage (terme employé dans le CRHQ pour indiquer le ratio d'aire terrestre/aire en eau), l'aire de drainage, et l'altitude moyenne en amont. Ces quatre proxys aideront l'outil à déterminer s'il se trouve dans une zone en tête de bassin, une zone de transition ou une zone de dépôt à l'aval d'un bassin versant (figure 1) (Schumm, 1977; Kondolf, 1997). Ces zones peuvent aider à déterminer les aléas fluviaux qui devraient se produire dans cette zone lorsqu'ils sont combinés avec les variables locales.

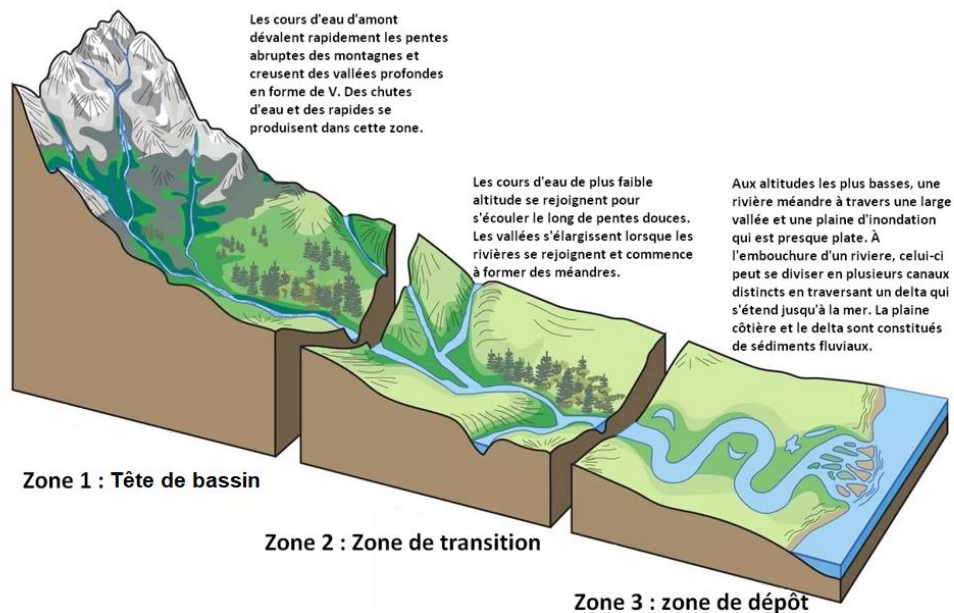


Figure 1: Les différentes zones d'un bassin versant, adapté de River Systems and Fluvial Landforms - Geology (U.S. National Park Service), 2022

## 2 PROBLÉMATIQUE

L'objectif de ce volet du projet est de développer des outils SIG pour caractériser les aléas géomorphologiques autour des routes de façon plus générale, et non seulement pour le contexte spécifique de la Basse-Côte-Nord. Ces outils visent à résoudre le problème suivant : les infrastructures routières du Québec risquent souvent d'être endommagées par le processus naturel des rivières, mais



l'étendue du réseau routier rend difficile l'identification des aires où les tronçons de route risquent le plus d'être endommagés. Cet outil permettra au MTMD d'identifier rapidement et facilement les endroits où des études plus poussées sur la possibilité d'endommagement du réseau routier sont nécessaires.

Ce volet comportait les 3 livrables suivants :

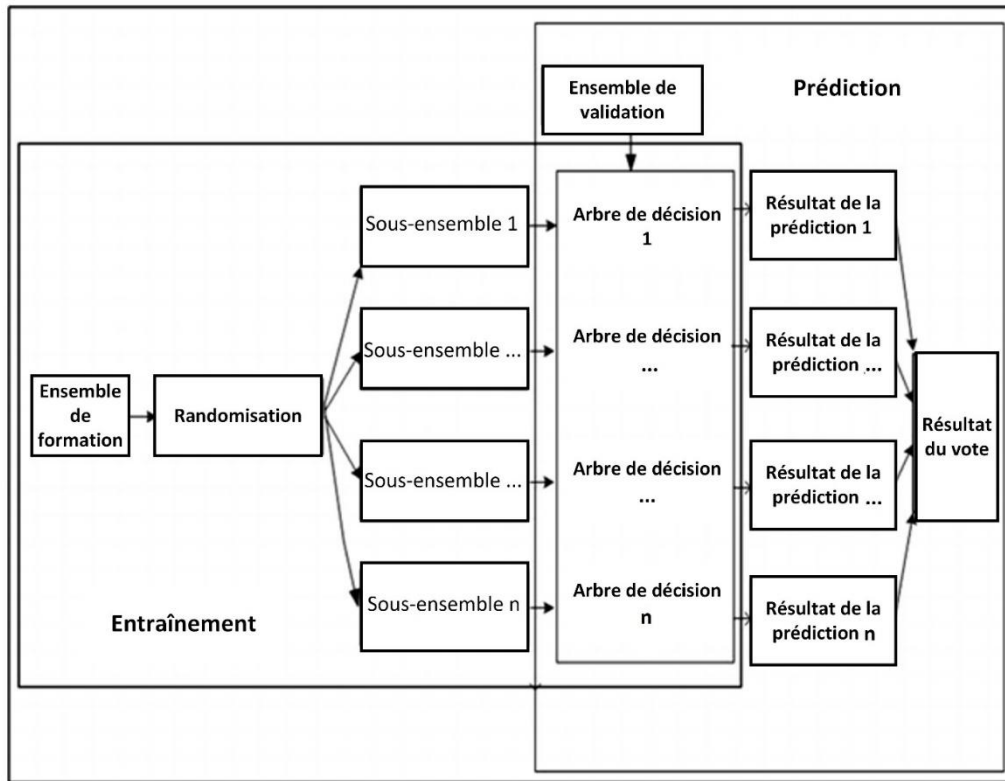
1. Un rapport final comprenant :
  - a. La méthodologie menant à l'élaboration des outils SIG pour identifier les aléas géomorphologiques à grande échelle (échelle du territoire).
  - b. Des cartes où les aléas peuvent se produire selon les outils SIG développés.
2. Un article de vulgarisation scientifique de 2 à 4 pages.
3. Une activité de transfert de connaissances. Cette activité prendra la forme d'un atelier avec une présentation du projet, de ses objectifs, des résultats et des recommandations qui en découlent. L'auditoire sera composé de gestionnaires, de professionnels et d'ingénieurs du ministère ou d'intervenants d'autres ministères et organismes.

### 3 MÉTHODOLOGIE

La méthodologie de ce projet utilise un modèle d'apprentissage automatique appelé **forêt aléatoire** (« random forest ») (Breiman, 2001). Un modèle d'apprentissage automatique est un outil utilisé pour détecter des schémas dans un ensemble de variables permettant de distinguer des classes associées à ces variables. De la même manière que la classification de Rosgen (1994) utilisera un sous-ensemble de variables pour classer les rivières en catégories, ces modèles les utiliseront pour détecter les tronçons de rivière où des inondations ou de la mobilité peuvent se produire. Cet algorithme de classification, composé de nombreux arbres de décision, est un type de modèle de plus en plus populaire dans le domaine de l'hydrogéomorphologie. De nombreux articles ont en effet montré qu'il s'agissait de l'un des modèles d'apprentissage automatique les plus utiles pour prédire les inondations et l'érosion (Wang et al., 2015; Ghosh & Dey, 2021; Fang et al., 2022; Hasan et al., 2023; Razavi-Termeh et al., 2023; Youssef et al., 2023). Ce modèle a été choisi afin de considérer autant de variables influentes que possible et d'aider à déterminer lesquelles sont les plus importantes dans les contextes géomorphologiques présents au Québec. Ce modèle supervisé utilise une **classification d'ensemble** qui présente de nombreux avantages par rapport aux méthodes traditionnelles, comme la correction de déséquilibre dans les données d'**entraînement** et du bruit dans les données (Parmar et al., 2019). Il importe de corriger cela lors de l'utilisation de l'apprentissage automatique pour les applications hydrogéomorphologiques, car le contexte géomorphologique de chaque bassin versant peut créer des conditions qui favorisent une classe plutôt

qu'une autre, ce qui entraîne une abondance de points d'entraînement pour cette classe et un manque de points d'entraînement pour les autres classes. Puisque l'outil est employé pour prédire l'inondabilité et la mobilité aux points de référence du CRHQ qui se situent à tous les 100 m de rivière pour une grande partie de la province, il est également important de considérer le problème du bruit dans les données qui se produit lorsque les points de référence contiennent de mauvaises données d'entrée résultant d'erreurs au cours du processus d'extraction des variables ou d'une mauvaise classification des aléas. Ces types d'erreurs sont assez fréquents à grande échelle lorsque les outils d'extraction de variables ne parviennent pas à extraire la bonne valeur ou que les experts ont des avis contradictoires sur les **processus fluviaux** locaux.

Le mode de fonctionnement de ce modèle est illustré à la figure 2. Le modèle crée un nombre défini d'arbres de décision qui sont indépendants les uns des autres en randomisant la procédure de sous-ensemble de données d'entraînement/validation et des variables de prédiction. L'utilisation de plusieurs arbres au lieu d'un seul améliore la capacité du modèle à prédire avec **précision** les points qui ne faisaient pas partie du processus d'entraînement (Parmer et al., 2019). Ceci est important pour ce projet car la classification de la totalité des points d'un bassin versant n'est pas pratique et le modèle sera éventuellement utilisé sur des bassins versants sur lesquels il n'a pas été entraîné. Les résultats de ce modèle sont bénéfiques à la fois à court terme pour déterminer les points chauds potentiels des aléas fluviaux mais aussi à long terme car les informations fournies sur l'importance de chaque variable dans la détection des zones sujettes aux aléas fluviaux pourront être utilisées dans des projets futurs.



**Figure 2: Schéma expliquant le fonctionnement du modèle forêt aléatoire (modifié de Parmar et al., 2019).**

La méthodologie de l'outil SIG est divisée en 4 sections principales :

- 1) Extraction des variables
- 2) Préparation des données
- 3) Entraînement du modèle supervisé
- 4) Prédiction des aléas pour le réseau hydrographique du bassin versant

### 3.1 Données d'entrée pour le modèle

Le modèle requiert les intrants suivants : un MNT, la délimitation des milieux humides, débits et réseau fluvial de l'Atlas Hydroclimatique, la limite du bassin versant, et les points de référence et réseau fluvial du CRHQ (tableau 1).

**Tableau 1: Intrants, banques de données et leur source pour le déploiement du modèle**

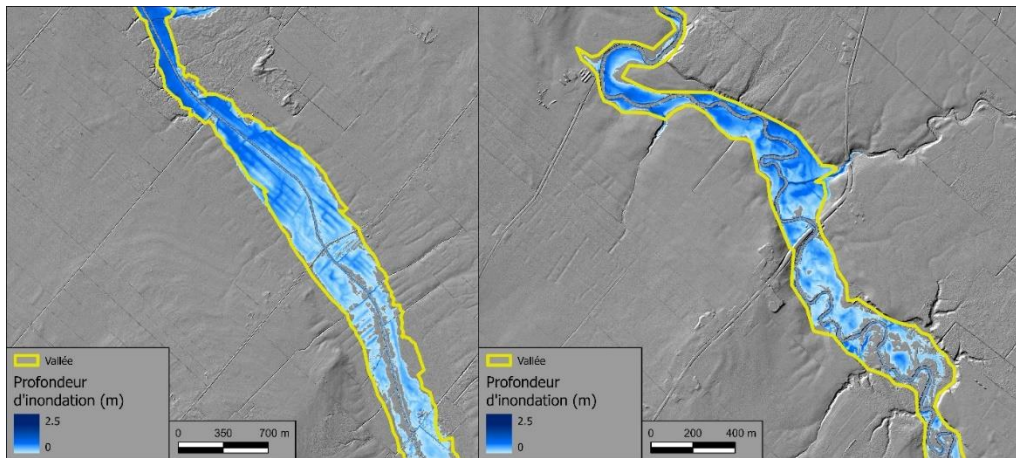
Intrants	Banques de données	Source
Points de référence (aux 100 m)	CRHQ	<a href="https://www.donneesquebec.ca/recherche/dataset/crhq">https://www.donneesquebec.ca/recherche/dataset/crhq</a>
Unités écologiques aquatiques (UEA)		
Débit plein-bord (Q2)	Atlas Hydroclimatique	<a href="https://cehq.gouv.qc.ca/atlas-hydroclimatique/carte-indicateurs/index.htm">https://cehq.gouv.qc.ca/atlas-hydroclimatique/carte-indicateurs/index.htm</a>
Réseau fluvial		
Modèle numérique terrain (MNT) LiDAR	Forêt ouverte	<a href="https://www.foretouverte.gouv.qc.ca">https://www.foretouverte.gouv.qc.ca</a>
Milieux humides potentiels		
Bassin hydrographique		

Les données du CRHQ sont disponibles à deux échelles, soit celle de **tronçons homogènes**, appelés unités écologiques aquatiques (UEA) et pour des points de référence situés à tous les 100 mètres le long des rivières. Ces données sont disponibles pour la majorité du Québec méridional (Annexe D). L'équipe du MELCCFP qui gère le CRHQ prévoit continuer à améliorer la qualité et la couverture de ces données. Les impacts de la qualité actuelle des données et les ajustements potentiels pour améliorer ce modèle sont abordés dans la conclusion. Le modèle utilisera le CRHQ pour le réseau fluvial, les points de

référence à tous les 100 mètres, ainsi que pour la plupart des variables géomorphologiques. Cela réduira la charge de travail du MTMD qui n'aura pas à maintenir des outils d'extraction de variables qui peuvent être difficiles et coûteuses à maintenir car elles doivent être mises à jour lorsque les outils de support passent à une nouvelle version. De cette manière, les variables seront automatiquement mises à jour lorsque de nouvelles versions du CRHQ seront publiées et le MTMD n'aura pas à mettre à jour les scripts utilisés pour extraire les variables.

La variable importante qui n'est pas disponible dans le CRHQ est la largeur de la vallée. Afin de la calculer pour les tronçons homogènes (UEA) ainsi que pour les points de référence, une combinaison d'outils est utilisée. Pour éviter l'utilisation de calculs hydrauliques complexes, nous avons employé l'outil de délimitation du fond de la vallée de la rivière développé par *Sechu et al. (2021)* car il est conçu pour déterminer la vallée inondable à grande échelle. Cet outil prend comme données d'entrée un MNT avec une résolution de 10 m, les milieux humides, le réseau hydrographique et la limite du bassin versant. L'outil calcule d'abord les pentes à partir du MNT, en s'assurant de remplacer les valeurs de pente de zéro par une très faible valeur afin d'éviter des erreurs pour la prochaine étape, soit le calcul par une méthode coût-distance dans laquelle les intrants sont les pentes (sans valeurs nulles) et le réseau hydrographique pour créer un fichier matriciel accumulation coûts-distance. Ce fichier coûts-distance représente les distances à partir de n'importe quel point (pixel) par rapport à la rivière la plus proche, pondérées par la pente locale. Cela donne une mesure utile pour déterminer l'étendue potentielle du fond de la vallée en déterminant les aires de moindre résistance à l'écoulement de l'eau. L'outil utilise ensuite ce nouveau fichier et les milieux humides potentielles pour déterminer statistiquement un seuil sous lequel les valeurs se trouvent dans la vallée, alors que les valeurs supérieures à ce seuil sont en dehors de la vallée. En appliquant ce seuil sur le fichier coût-distance, en reclassant les points au-dessus du seuil à une valeur nulle et en convertissant le fichier matriciel en polygone, il reste un polygone du fond de la vallée pour le réseau hydrographique (*Sechu et al., 2021*). Lors de l'exécution de l'outil sur les petites (aire de drainage < 50 km<sup>2</sup>) et les grandes rivières (aire de drainage > 50 km<sup>2</sup>), il est recommandé de les exécuter séparément pour cette étape, car les exécuter simultanément peut fausser la méthode statistique utilisée pour déterminer le seuil et produire de mauvais résultats sur les plus grandes rivières. Cette considération n'est pas un enjeu pour la présente étude puisque seules les rivières dont l'aire de drainage > 50 km<sup>2</sup> comportent des valeurs de débit dans l'atlas hydro-climatique du Québec. Certains points supplémentaires ont toutefois été inclus en utilisant le débit disponible le plus proche pour tester si cette méthode peut s'appliquer à des rivières plus petites en utilisant des données d'entraînement sur des rivières plus grandes. Ces points représentent 10 % des données de prédiction mais n'ont pas été utilisés pour entraîner le modèle. L'annexe C présente les implications de l'inclusion de ces plus petits cours d'eau et les résultats du modèle dans ces environnements.

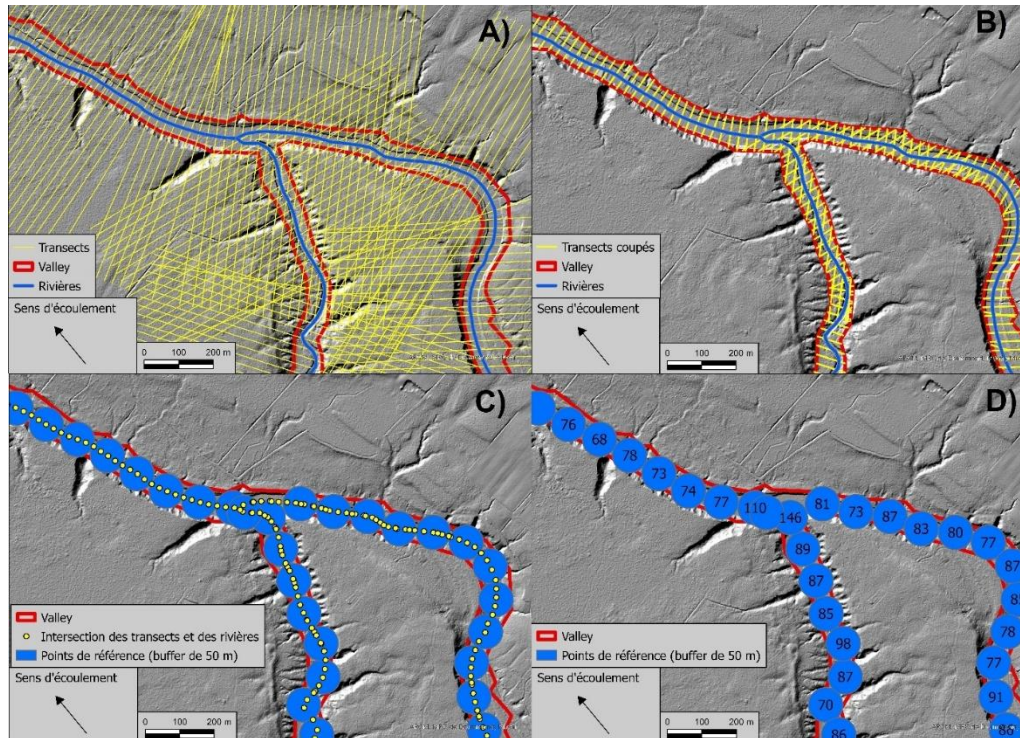
En comparant les résultats de la vallée aux modèles hydrauliques disponibles, notamment la modélisation LISFLOOD large-échelle (Choné et al., 2021) complétée par l'équipe de recherche sur les bassins de la rivière Nicolet et Du Gouffre dans le cadre du projet INFO-CRUE (e.g. figure 3), on constate que le fond de vallée extrait représente approximativement l'espace inondé par une crue de récurrence 20 ans, ce qui est suffisant pour distinguer les rivières confinées des rivières non confinées (figure 3). D'autres outils et méthodes existent pour extraire cette variable et peuvent être trouvés dans Gilbert et al. (2016) tels que les « outils de corridor fluvial » ou VBET (« Valley Bottom Extraction Tool »). Toutefois, lors des essais menés durant cette étude, l'outil de Sechu et al. (2021) s'est avéré le plus simple à utiliser tout en produisant des résultats de qualité suffisante pour distinguer les types de rivières.



**Figure 3: Comparaison du polygone de fond de vallée extrait à des simulations d'inondation à grande échelle à l'aide de LISFlood et de la méthodologie de Choné et al. (2021) en utilisant une valeur correspondant à environ une inondation de 20 ans**

Une fois établi, le polygone de fond de vallée est ensuite entré dans un nouvel outil développé en Python pour ce projet afin de déterminer la largeur de la vallée. Ce nouvel outil prend comme données d'entrée : le polygone du fond de la vallée, le réseau hydrographique, les points de référence aux 100 m, et deux paramètres optionnels pour ajuster le lissage du polygone de la vallée ainsi que la longueur maximale possible des transects. Cet outil va ensuite générer des transects le long de ce réseau hydrographique à des intervalles fixes de 25 m. Ce seuil a été déterminé afin de permettre des mesures multiples par point de référence; c'est donc la moyenne qui est calculée pour réduire les erreurs de surestimation. Des points sont générés au milieu de ces nouveaux transects et les transects sont ensuite découpés sur le polygone original de la vallée. Ces transects sont alors répartis en sous-ensembles en sélectionnant uniquement ceux qui se croisent avec les points centraux générés. Une fois cette opération effectuée, des points sont générés à l'intersection des transects et du réseau fluvial, ces points contenant la longueur du transect. Ensuite, la moyenne de

tous ces points d'intersection situés à moins de 50 m des points de référence est prise et considérée comme la largeur de la vallée locale (figure 4). Cette méthode tend à surestimer d'environ 10 % la largeur de la vallée dans les rivières très sinueuses et aux confluent et ne doit donc être utilisée que pour calculer les ratios de confinement et non pour d'autres études. Cette méthode est une solution temporaire jusqu'à ce que le CRHQ inclue les mesures de la largeur de la vallée.

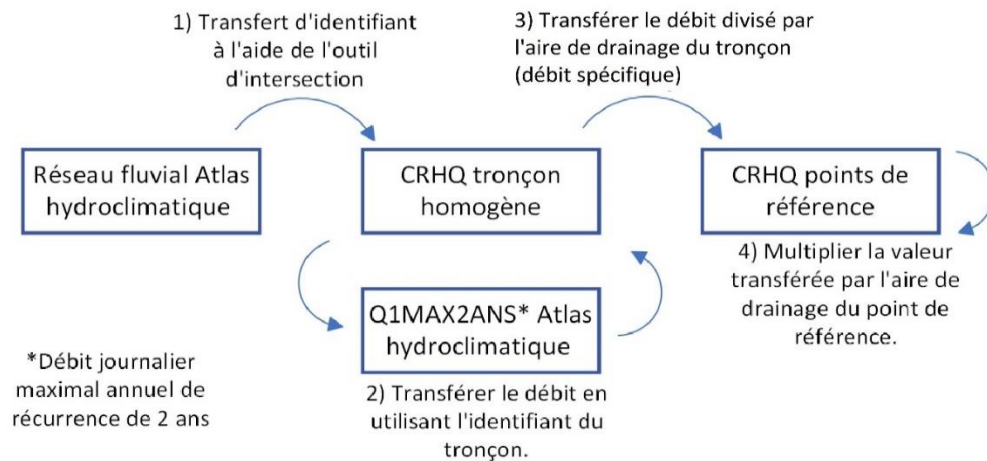
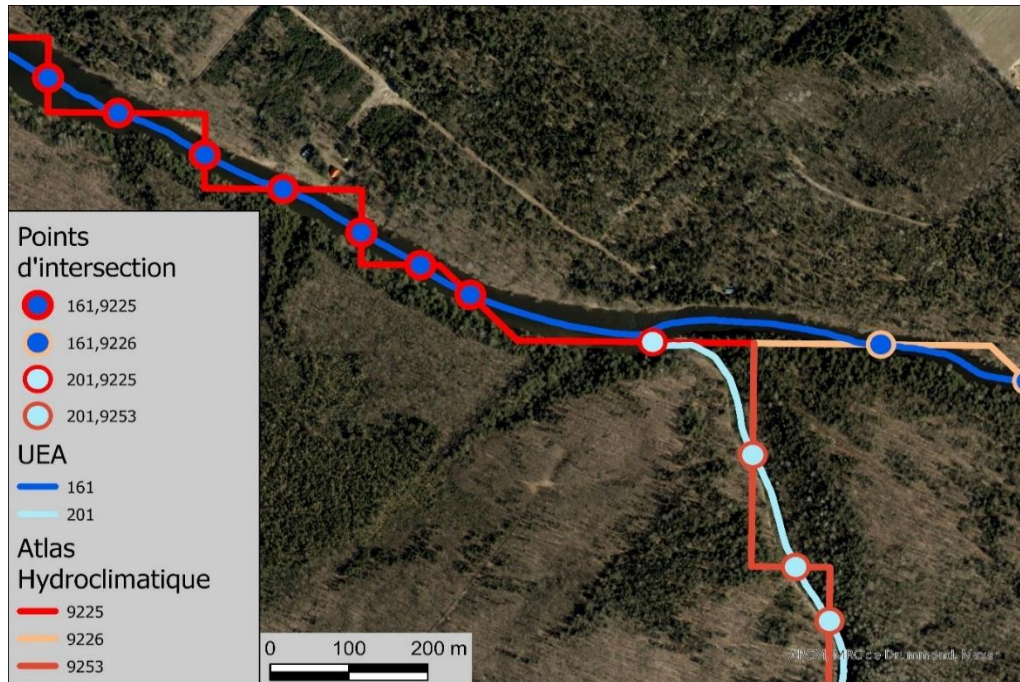


**Figure 4: Exemple de script de calcul de confinement. A) transects de taille maximale générés tous les 25 m le long du réseau hydrographique. B) transects coupés à l'étendue de la vallée. C) intersection des transects découpés et de la rivière superposée aux points**

Les points de référence ne disposent actuellement que des valeurs de débit de l'Atlas Hydroclimatique 2018 qui contient environ 5000 segments, et non celles des débits de la version la plus à jour (2022) qui contient près de 28 000 segments, incluant plusieurs tributaires. Par conséquent, un script a été créé pour extraire et joindre les valeurs de débit de l'atlas aux UEA. Ce script fonctionne en lui fournissant le réseau hydrographique de l'atlas, le réseau hydrographique CRHQ et les débits de récurrence 2 ans (considérés équivalents au débit plein bord) de l'Atlas Hydroclimatique. Il détecte ensuite la fréquence à laquelle chacun des tronçons individuels du réseau UEA (lignes dans des tons de bleu dans la figure 5) croise un tronçon de l'Atlas Hydroclimatique (lignes dans des tons de rouge dans la figure 5) et le tronçon



qu'il croise le plus est considéré comme le tronçon correspondant. Une fois cela fait, il joint la valeur du débit divisée par l'aire de drainage au tronçon correspondant, i.e. la valeur du débit spécifique (figure 5).



**Figure 5:** La figure montre le fonctionnement de l'outil d'intersection des tronçons modélisés de l'Atlas Hydroclimatique (lignes rouges) et les tronçons UEA de la CRHQ (lignes bleues). Les points sont les points d'intersection et sont colorés en fonction de l'UEA



Ce qui explique pourquoi cet outil est nécessaire pour éviter de transférer la valeur d'un mauvais tronçon de l'atlas. Les points d'intersection entre les tronçons des deux réseaux fluviaux (UEA du CRHQ ou Atlas Hydroclimatique) sont identifiés à l'aide de leur numéro d'identification respectif (161 et 162 pour les UEA, 9225, 9226 et 9253 pour l'Atlas Hydroclimatique). La partie du bas explique comment le transfert de débit est mis à l'échelle de l'aire de drainage locale pour éviter que le tronçon entier ait la même valeur de débit.

### **3.2 Bassins versants pilotes**

L'outil SIG développé a été appliqué dans trois bassins versants pilotes (figure 6) sélectionnés pour représenter une diversité de paysages dans les régions peuplées du Québec. Le bassin versant de la rivière Cascapédia (3141 km<sup>2</sup>), dans la région des Appalaches de la Gaspésie, est caractérisé par un fort contrôle géologique où des vallées fluviales confinées s'élargissent périodiquement, ce qui augmente la probabilité d'incision ou d'inondation de la vallée lorsqu'elle s'élargit (figure 6A). Le bassin versant de la rivière Nicolet (3410 km<sup>2</sup>), dont les tronçons amont se trouvent dans la région des Appalaches et les tronçons aval dans les basses terres du Saint-Laurent, se caractérise par de longs tronçons de rivières confinées en amont se déversant dans de vastes plaines d'inondation agricoles non confinées en aval (figure 6B). Finalement, le bassin versant de la rivière Du Gouffre (1000 km<sup>2</sup>), dans le centre du Québec, comportent des cours d'eau confinés et à forte pente en amont se déversant dans de grandes vallées marginalement confinées avec des dépôts épais, ce qui permet une migration latérale importante (figure 6C).

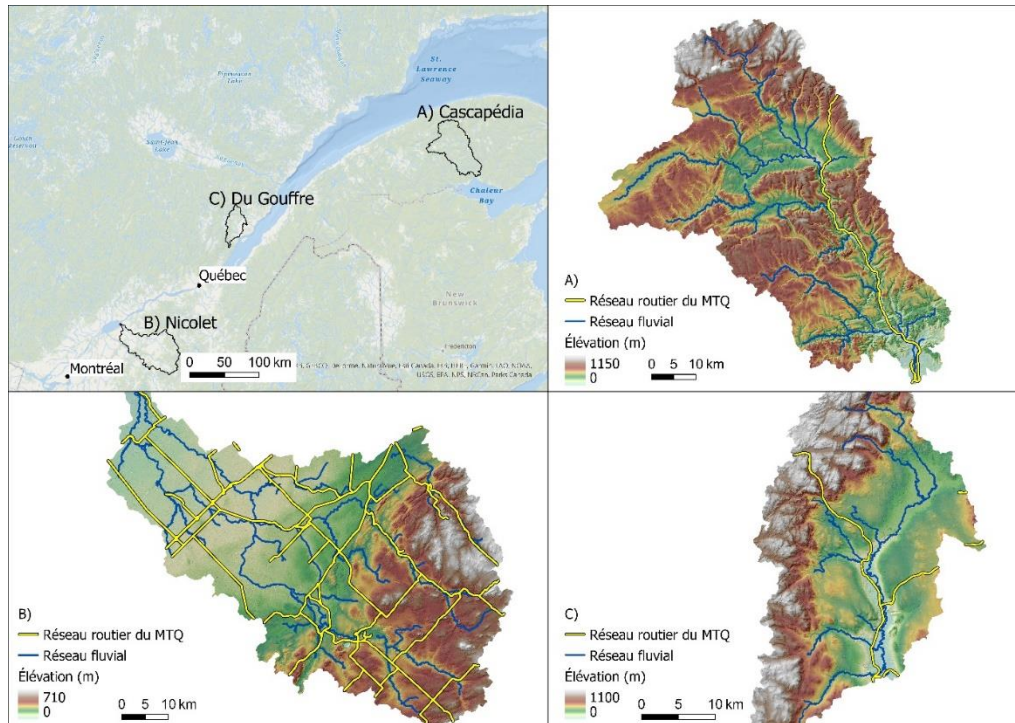


Figure 6: Emplacement des trois bassins versants pilotes; A) rivière Cascapédia; B) rivière Nicolet; C) rivière Du Gouffre.

### 3.3 Préparation des données

L'étape de préparation des données à entrer dans le modèle est effectuée par un script Python séparé. Ce script ne nécessite que les UEAs modifiées qui contiennent maintenant la largeur de la vallée à chaque point et les valeurs de débit spécifique mises à jour. Ce script déterminera ensuite le débit à chaque point de référence en prenant le débit spécifique attaché aux tronçons homogènes et en le multipliant par l'aire de drainage à chacun des points de référence du tronçon. Le modèle obtient ainsi des valeurs de débit à chaque point. Le modèle utilise ensuite cette valeur de débit, la pente et la largeur de la rivière fournies par l'UEA pour calculer la puissance spécifique du cours d'eau à chaque point.

Le script calcule également un ratio de confinement à chaque point en divisant la largeur de la vallée locale par la largeur de la rivière. Ensuite, il agrège les largeurs de la rivière et de la vallée par tronçon, et joint ces informations aux tronçons homogènes. Il divise alors la nouvelle largeur moyenne de la vallée par la largeur moyenne de la rivière pour obtenir un ratio de confinement au niveau du **tronçon homogène** (tableau 2). Il divise également le ratio de confinement

local par le ratio de confinement du tronçon à chaque point pour obtenir un ratio de confinement relatif. Il reste donc trois ratios : d'abord le ratio de confinement local, puis le ratio de confinement du tronçon, et enfin le ratio de confinement relatif qui mesure si le point actuel se trouve dans une zone plus ou moins confinée que la moyenne de son tronçon.

$$\text{Ratio de confinement local} = \frac{\text{largeur de la vallée}}{\text{largeur plein bord}}$$

$$\text{Ratio de confinement tronçon} = \frac{\text{moyenne de la largeur de la vallée du tronçon}}{\text{moyenne de la largeur plein bord du tronçon}}$$

$$\text{Ratio de confinement relatif} = \frac{\text{ratio de confinement local}}{\text{ratio de confinement tronçon}}$$

Pour l'occupation des berges, le CRHQ contient des informations sur l'utilisation des terres sur les deux berges. Pour intégrer ces informations dans l'étude, quatre champs de présence/absence ont été ajoutés pour des catégories différentes d'occupation du sol : les structures anthropiques (routes, ponts, villages), les exploitations agricoles, la végétation et les milieux humides. Ces quatre champs ont reçu une valeur de 1 si cette utilisation du sol était présente sur l'une des rives et de 0 si elle était absente.

**Tableau 2: Liste des variables utilisées par le modèle**

Échelle	Variable	Champ	Unité
Locale	Pente locale au point de référence	Slope_loc	m/m
	Angle moyen entre le point actuel et le point précédent et suivant (sinuosité locale)	Ang_dev_q50	-
	Largeur du chenal au point de référence	Largeur_mod	m
	Aire de drainage du point actuel	Abv_km2	km <sup>2</sup>
	Puissance spécifique du cours d'eau au point actuel	USP	W/m <sup>2</sup>
	Présence d'une exploitation agricole sur au moins une des deux rives du point local	HumanAct	Présence (1)/ absence (0)
	Présence d'un milieu humide sur au moins une des deux rives du point local	Wetland	Présence (1)/ absence (0)

	Présence de végétation sur au moins une des deux rives du point local	Vegetation	Présence (1)/ absence (0)
	Présence d'une structure anthropique (route, stabilisation, ville, ...) sur au moins une des deux rives du point local	Anthro	Présence (1)/ absence (0)
	Ratio local de la largeur de la rivière divisée par la largeur de la vallée locale	LocConfRatio	m/m
	Ratio de confinement local divisé par le ratio de confinement du tronçon pour déterminer le ratio de confinement relatif	RelConfRatio	-
Tronçon	Ratio de la moyenne de la largeur de la rivière pour le tronçon divisée par la largeur moyenne de la vallée pour le tronçon	ValConfRatio	m/m
	La longueur du tronçon divisée par la longueur si le tronçon était une ligne droite	Sinuosite	m/m
Bassin	Pente moyenne en amont du point local	S_moy	%
	Altitude moyenne en amont du point local	Alt_moy	m
	Zone terrestre divisée par zone aquatique en amont du point	Ratio_d	m <sup>2</sup> /m <sup>2</sup>

### 3.4 Entraînement du modèle supervisé

#### 3.4.1 Collecte des données de validation

Pour entraîner et valider le modèle, une combinaison d'informations sur les processus fluviaux et leurs répercussions doit être disponible. Cela peut prendre la forme de rapports d'incidents concernant les infrastructures routières, de l'analyse d'un MNT et/ou de photos historiques, de cartes d'étendue des inondations de récurrence de près de 20 ans (modélisation LISFLOOD large-échelle du projet Info-crue), de sorties d'autres modèles, ainsi que des observations de terrain pour confirmer les informations obtenues par télédétection. Le MNT peut être utilisé pour identifier les inondations et l'érosion en suivant les informations fournies dans *Biron et al. (2014)* et *Dupuis et al. (2017)*. Dans cette étude, les données de terrain ont été recueillies au cours de l'été

2022 avec deux visites dans le bassin versant de la rivière Nicolet et une visite dans les bassins versants de la Cascapédia et de la Du Gouffre. Les signes d'inondation et de mobilité (incluant le type de mobilité) ont alors été documentés ainsi que des photos recueillies par drone et caméra. Les données de validation collectées et utilisées pour ce projet se trouvent dans le tableau 3.

**Tableau 3: Ensembles de données de validation utilisées dans le projet.**

Type de données	Description
Données de terrain	<ol style="list-style-type: none"> <li>1. Observations de la stabilité des berges du bassin versant du Gouffre réalisées par Jean-Philippe Marchand en 2009</li> <li>2. Photos et observations recueillies par l'équipe de Concordia et de l'UQAR durant l'été et l'automne 2022 (tous les bassins versants)</li> <li>3. Données de terrain recueillies par la Direction de l'expertise hydrique (DEH) du MELCCFP dans le bassin versant du Gouffre en 2022</li> </ol>
Résultats du modèle "Planform statistic" de la DEH	Modèle prédisant les taux d'érosion des berges pour le bassin versant du Gouffre en utilisant les tracés historiques des rivières pour une période de 56 ans (2020 par rapport à 1964)
LiDAR	Utilisation du MNT haute résolution pour identifier les preuves d'inondation/érosion anciennes et actuelles, telles que les cicatrices d'inondation et les anciens décrochements de berge et glissements de terrain
Photos historiques	Comparaison des photos historiques datant de plusieurs années pour identifier les endroits où la rivière a pu migrer et déterminer s'il s'agissait d'un changement abrupt ou progressif
Photos satellites	Recherche dans l'imagerie moderne des signes d'érosion des berges tels que l'absence de végétation ou les décrochements dus aux glissements de terrain.
Système de Gestion des Ponceaux et Autres Éléments d'Inventaires (M012)	Données d'inspection des ponceaux effectuées par le MTMD permettant de classer l'état de la structure et du cours d'eau

Réparations routières précédentes	Diverses sources de données historiques sur les réparations routières : Système de Gestion des Structures du Québec (GSQ), données recueillies par Tristan Caron, géomorphologue à la DG de la Gaspésie – Îles-de-la-Madeleine (2017 – 2020), et autres informations transmises par divers employés du MTMD
-----------------------------------	---

### 3.4.2 Classification des données

Une fois que toutes les données identifiées au tableau 2 ont été préparées par l'outil précédent et les données de validation recueillies, le modèle doit être entraîné. Pour ce faire, un sous-ensemble de points de référence doit être classé en fonction de trois catégories : présence d'inondation, présence de mobilité, et type d'érosion si la mobilité est présente (sinon une valeur de - 999 est entrée). La classification se fait en entrant un nombre représentant une classe dans le champ approprié entre les 3 classes prédites :

- présence d'inondation (inondations probables avec une période de retour d'environ 20 ans, basé sur des cartes de modélisation large-échelle d'Info-crue);
- présence de mobilité (possibilité de mobilité justifiant une étude hydrogéomorphologique plus approfondie) et;
- type d'érosion dans les cas des tronçons avec un risque de mobilité (migration latérale ou incision, voir le schéma de classification aux tableaux 4 à 6).

Ces valeurs doivent être saisies dans les champs intitulés respectivement "inondation", "mobilité" et "érosion" (sans accents pour assurer le bon fonctionnement des outils) sur la base d'une évaluation des données de validation et d'une évaluation hydromorphologique de la zone à l'aide du MNT et d'images satellitaires.

**Tableau 4: Schéma de classification utilisé lors de la classification des points de référence pour l'inondation dans le champ "inondation".**

Numéro saisi	Classe de présence d'inondation
0	Inondation peu probable
1	Inondation probable

**Tableau 5: Schéma de classification utilisé lors de la classification des points de référence pour la mobilité dans le champ "mobilité".**

Numéro saisi	Classe de présence de mobilité
0	Stable
1	Mobile

**Tableau 6: Schéma de classification utilisé lors de la classification des points de référence pour le type d'érosion dans le champ "érosion".**

Numéro saisi	Classe type d'érosion
0	Incision
1	Migration latérale
-999*	Stable
*Déterminé dans l'étape de mobilité et donc pas pris en compte dans le modèle de type d'érosion, par conséquent, une valeur manquante de -999 est entrée.	

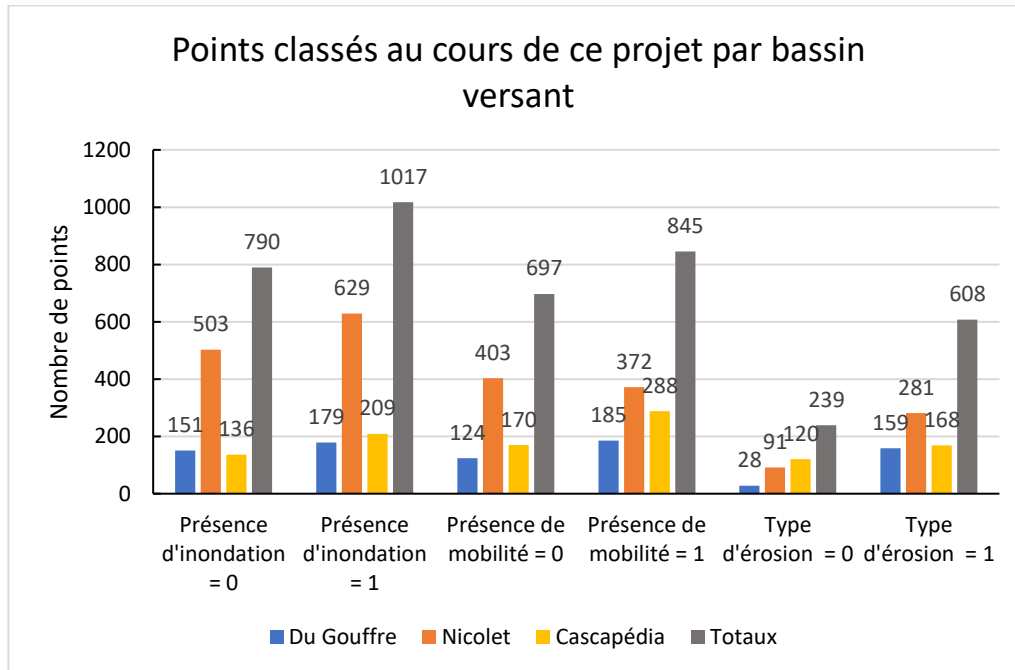
Il importe de maintenir un nombre égal ou similaire de points classés dans chaque classe, car si une classe est représentée de manière disproportionnée, le modèle peut être **surentraîné** pour cette classe et pas assez bien entraîné pour les autres. Au cours du projet, il a également été constaté qu'il est important d'équilibrer le nombre de points dans chaque classe par bassin versant et pas seulement pour les données dans leur ensemble. Ceci est dû aux différents contextes géomorphologiques qui existent dans les différents bassins versants et qui peuvent modifier l'impact de chaque variable sur les aléas fluviaux qui s'y produiront. Par exemple, les principales variables contrôlant l'érosion dans le bassin de la rivière Nicolet peuvent différer de celles qui la contrôlent dans le bassin de la rivière Du Gouffre. Le modèle forêt aléatoire tente de compenser ces différences mais plus les données sont équilibrées, plus le modèle sera en mesure de les différencier (Parmar et al., 2019).

Au total, 1807 points ont été classés pour la présence/absence d'inondations, 1542 pour la présence/absence de mobilité et 847 pour le type d'érosion

(figure 7). Les inondations sont plus faciles à identifier en utilisant le LiDAR et les résultats d'autres modèles, ce qui nous a permis de classer plus de points pour cet aléa par rapport aux autres. Par ailleurs, le nombre plus élevé de points pour la mobilité par rapport au type d'érosion est dû au fait que les zones stables ne sont pas prises en compte dans le modèle de type d'érosion qui vise à distinguer les zones mobiles liées à l'incision ou à la migration latérale. Le nombre de points classés pour chaque catégorie dans les modèles pour la présence d'inondation et la présence de mobilité étaient bien équilibrés mais l'incision est beaucoup plus difficile à identifier que la migration latérale. Il y a donc plus de points classés pour la migration latérale que pour l'incision, en particulier dans le bassin versant de la Nicolet.

Cette classification des points se fait facilement lorsque des données existent et que l'expertise appropriée pour compléter les évaluations hydrogéomorphologiques est disponible. La classification des cours d'eau dans cette étude a consisté en 2 jours de travail sur le terrain et 2 jours d'évaluations hydrogéomorphologiques (à distance) basées sur les données de validation et les données SIG par bassin versant. Des classifications supplémentaires peuvent facilement être ajoutées en modifiant les points de référence et en réexécutant les étapes d'apprentissage et de prédiction du modèle. Ceux-ci peuvent être ajoutés à de multiples occasions, par exemple avant l'ajout d'un nouveau bassin versant au projet, lorsqu'un expert n'est pas d'accord avec une classification, après une grande inondation qui a révélé des emplacements vulnérables (comme sur la rivière du Gouffre au printemps 2023). L'ajout de nouvelles classifications est facultatif mais permet d'améliorer la capacité des modèles à détecter les aléas fluviaux dans de nouveaux environnements, et ce avec une plus grande confiance. Il est également possible de combiner les avis de plusieurs experts en leur demandant d'établir chacun leur propre classification, puis de les agréger par mode, ce qui signifie que la classe choisie par le plus grand nombre d'experts sera utilisée lors de l'entraînement des modèles.





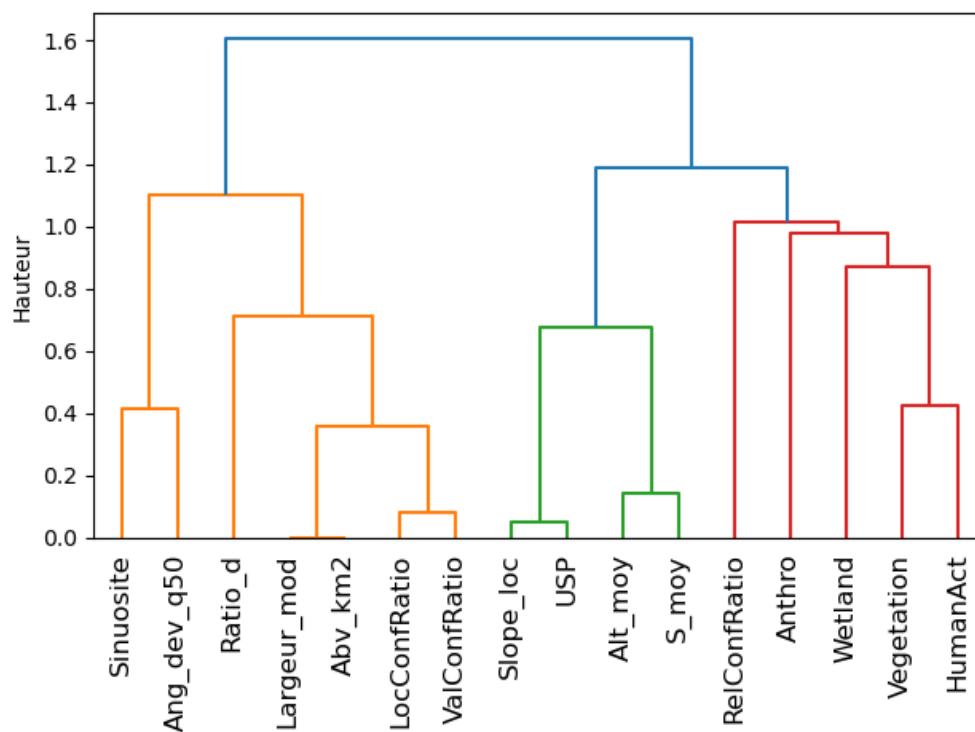
**Figure 7: Distribution des points classés dans chaque catégorie par aléa fluvial et par bassin versant.**

### 3.4.3 Sélection des variables finales des modèles

Lorsque l'on utilise des méthodes d'apprentissage automatique, il est important de tester la multicolinéarité des variables (Kavzoglu & Teke, 2022), car si les variables ne sont pas indépendantes les unes des autres ou si elles sont trop corrélées, cela peut poser des problèmes lors de l'interprétation des résultats du modèle, en particulier lors des tests de permutation des variables décrits ci-dessous (section 4.1). Cela est dû à l'augmentation artificielle de l'importance de certaines variables en donnant au modèle essentiellement la même variable deux fois, ce qui réduit également l'importance d'autres variables qui peuvent en fait être plus importantes qu'elles ne le paraissent si des variables colinéaires sont ajoutées au modèle. Pour tester cet effet, la corrélation de rang de Spearman a été utilisée et ensuite convertie en une matrice de distance pour calculer la méthode de regroupement de Ward afin de créer un dendrogramme des variables les plus corrélées. Un seuil de 0,25 a ensuite été choisi pour créer des groupes de variables et l'une des variables dont la valeur de distance était inférieure à ce seuil a été supprimée.

La figure 8 montre qu'il existe quatre cas de ce type dans le modèle. D'abord, la largeur de la rivière est très fortement corrélée à l'aire de drainage, car dans le CRHQ la largeur est estimée en fonction de l'aire de drainage (relation de géométrie hydraulique), ce qui fait que ces deux variables ne sont pas indépendantes l'une de l'autre. Cela changera dans une prochaine version du

CRHQ, car une nouvelle méthode de détermination de la largeur à partir du LiDAR est en cours de développement. Comme la largeur est utilisée dans le calcul de la puissance spécifique (USP, « Unit Stream Power »), elle a été éliminée du modèle et l'aire de drainage a été conservée. Le deuxième cas est celui des ratios de confinement local et de confinement de la vallée, ce qui est normal puisque le ratio de confinement de la vallée est le ratio de confinement moyen de toutes les mesures locales à l'intérieur de son tronçon homogène; ces deux valeurs ne sont donc pas non plus indépendantes l'une de l'autre. La valeur locale a été conservée car le modèle dispose également du confinement relatif qui permet de mesurer si le confinement du tronçon homogène est plus grand ou plus petit que le confinement local. La pente locale et l'USP sont également fortement corrélées puisque la pente est utilisée pour calculer l'USP ; elles ne sont donc pas totalement indépendantes l'une de l'autre. L'USP a été conservée afin d'incorporer la largeur et la pente locales dans le modèle car elles sont utilisées pour calculer l'USP. Enfin, la pente moyenne et l'altitude moyenne en amont du point local sont corrélées l'une à l'autre, ce qui est probablement dû au fait que les cours d'eau d'amont ont tendance à avoir des pentes plus fortes et que, par conséquent, plus l'altitude du point local est élevée, plus il est probable que la pente en amont le soit également. La pente en amont est aussi un meilleur indicateur pour estimer la quantité et le type de sédiments qui pourraient atteindre la zone locale, ce qui est important pour déterminer la mobilité, et c'est pourquoi elle a été conservée. Le tableau 7 est une version tronquée du tableau 2 montrant uniquement les variables qui ont été conservées pour les étapes de l'entraînement/prédiction du modèle.



**Figure 8: Les variables regroupées en fonction de la distance euclidienne entre leurs corrélations. Plus la hauteur (distance euclidienne entre leurs corrélations) des lignes reliant les variables est faible, plus elles sont corrélées.**

**Tableau 7: Tableau des variables conservées dans le modèle pour l'étape d'entraînement/prédiction**

Échelle	Variable	Champ	Unité
Locale	Angle moyen entre le point actuel et le point précédent et suivant	Ang_dev_q50	-
	Aire de drainage du point actuel	Abv_km2	km <sup>2</sup>
	Puissance spécifique du cours d'eau au point actuel	USP	W/m <sup>2</sup>
	Présence d'une exploitation agricole sur les deux rives du point local	HumanAct	Présence (1)/absence (0)
	Présence d'un milieu humide sur les deux rives du point local	Wetland	Présence (1)/absence (0)
	Présence de végétation sur les deux rives du point local	Vegetation	Présence (1)/absence (0)
	Présence d'une structure anthropique sur les deux rives du point local	Anthro	Présence (1)/absence (0)
	Ratio local de la largeur de la rivière divisée par la largeur de la vallée locale	LocConfRatio	m/m
	Ratio de confinement local divisé par le ratio de confinement du tronçon pour déterminer le ratio de confinement relatif	RelConfRatio	-
Tronçon	La longueur du tronçon divisée pas la longueur si le tronçon était une ligne droite	Sinuosite	m/m
Bassin	Pente moyenne en amont du point local	S_moy	%
	superficie terrestre divisée par superficie des lacs et cours d'eau en amont du point	Ratio_d	m <sup>2</sup> /m <sup>2</sup>

#### **3.4.4 Procédure d'entraînement**

Une fois classé, ce nouvel ensemble de données peut être entré dans le script d'entraînement du modèle fourni dans la boîte à outils. Ce script divise les

données en données d'entraînement et de validation en utilisant une procédure qui assure une distribution uniforme de chaque classe dans chaque sous-ensemble. Ensuite, le script entraîne et exporte un modèle forêt aléatoire pour chaque aléa à utiliser dans l'étape finale (un pour la présence d'inondation, un pour la présence de mobilité, et un pour le type d'érosion). Les statistiques de validation sont également exportées, de même que des graphiques qui montrent quelles variables sont importantes dans le processus de prédiction de ce modèle. Ces graphiques affichent deux mesures différentes : "diminution moyenne de l'impureté (DMI)", qui quantifie l'importance de chaque variable pour limiter chaque **nœud** des arbres de décision à une classe, ce qui augmente la certitude du modèle, et "diminution moyenne de la précision (DMP)", qui quantifie de combien la précision du modèle est réduite si les valeurs de cette variable sont mélangées parmi les points. Pour comprendre la diminution de l'impureté, il est tout d'abord important de comprendre la métrique d'impureté utilisée pour déterminer la confiance du modèle dans la prédiction de cette classe. L'impureté est mesurée à chaque nœud à l'aide de la formule suivante :

$$\text{Impureté (I)} = 1 - \frac{\text{nombre de points avec la bonne classe dans le nœud}}{\text{nombre de points dans le nœud}}$$

Le DMI de chaque variable est calculé à l'aide de la formule :

$$DMI = \frac{1}{n} \sum I_{x-1} - I_x$$

Où n est le nombre de fois où la variable est utilisée pour diviser les nœuds, tandis que  $I_x$  est l'impureté au nœud actuel et  $I_{x-1}$  est l'impureté au nœud précédent. La DMP est calculée pour chaque variable en randomisant cette variable et en déterminant de combien cela réduit la précision du modèle. La précision est le nombre de points correctement prédits par rapport au nombre total de points.

$$\text{Précision} = \frac{\text{nombre de points correctement prédits}}{\text{nombre total de points}}$$

$$DMP = \text{Précision}_{\text{modèle}} - \text{Précision}_{\text{modèle avec randomisation de la variable}}$$

Pour les deux mesures, plus la valeur est élevée, plus la variable est importante pour le modèle.

### 3.5 Prédiction des aléas pour le réseau fluvial d'un bassin versant

Une fois les modèles entraînés, ils peuvent être entrés dans un script final de prédiction. Ce script prend comme données d'entrée les points de référence aux 100 m du CRHQ avec les nouveaux champs ajoutés et les modèles entraînés. Il va ensuite exécuter les trois modèles sur tous les points de référence dans le

bassin versant. Le résultat de cet outil est une version mise à jour des points de référence CRHQ avec trois nouveaux champs pour chaque aléa fluvial prédit pour un total de 9 nouveaux champs. Les trois types de nouveaux champs sont ; "XPred" qui contient la valeur prédite (0 ou 1), "XProb" qui a le niveau de confiance du modèle en pourcentage, et "XClass" qui contient un texte descriptif indiquant quelle classe a été prédite ainsi que le niveau de confiance en utilisant les seuils affichés dans le tableau 8. Les seuils du champ de "XClass" ont été fixés de manière relativement arbitraire à environ 1 écart-type de la moyenne des valeurs de probabilité. Ceci est principalement utilisé pour cartographier les sites potentiels d'aléas, mais le champ de probabilité devrait être employé pour mieux comprendre le potentiel d'aléas à ces endroits. La valeur minimale du niveau de confiance est de 50 %. Cette valeur indique que la moitié des arbres de décision ont voté pour chaque classe et, par conséquent, le modèle est incapable de prédire avec confiance la classe appropriée. Il importe de souligner que les résultats du modèle ne sont pas des espaces d'inondabilité ou de mobilité à proprement parler, mais des probabilités que de la mobilité ou de l'inondabilité se produise aux points de référence CRHQ. Cette probabilité n'informe ni sur la hauteur, ni sur l'étendue latérale de l'inondation et encore moins sur la direction et l'étendue de la mobilité. Les résultats de ce modèle devraient être utilisés pour identifier les sites potentiels où des analyses hydrogéomorphologiques et hydrauliques plus approfondies devraient être menées. En utilisant le champ "classe", nous pouvons déterminer un niveau total d'incertitude pour les modèles en divisant le nombre de points avec un niveau de confiance de 55% ou moins par le nombre total de points.

**Tableau 8: Tableau montrant les nouveaux champs ajoutés aux points de référence par l'étape de prédiction du modèle pour chaque classe prédite, X étant remplacé par « Inondation », « Mobilité » et « Erosion » sans accents pour ne pas provoquer d'erreurs dans ArcGIS pour un total de 9 nouveaux champs.**

XPred - Valeur prédite	XProb - certitude du modèle	XClass - Résultats classés par degré de certitude
0	50% – 55% = Incertain	Classe 0 - probabilité élevée
	55.01% – 75% = Probabilité moyenne	Classe 0 - Probabilité moyenne
		Incertain
1	75.01% - 100% = Probabilité élevée	Classe 1 - Probabilité moyenne
		Classe 1 - Probabilité élevée

## 4 ANALYSE DES RÉSULTATS

### 4.1 Sélection des paramètres des modèles

Le modèle forêt aléatoire a deux paramètres principaux à calibrer : le nombre d'arbres générés et le moment où les arbres doivent cesser de créer de nouveaux nœuds (*Parmar et al., 2019; scikit-learn developers, 2023*). Le nombre d'arbres est simplement contrôlé en entrant une valeur sans décimales représentant le nombre d'arbres à utiliser. Le contrôle du moment où les arbres doivent cesser de générer de nouveaux nœuds peut être effectué à l'aide de différentes techniques telles que la saisie d'une valeur fixe pour le nombre minimum de points de données par nœud, la saisie d'un pourcentage de la base de données d'entraînement totale pour déterminer la quantité minimum de points de données par nœud, ou la définition d'un seuil pour la réduction de l'impureté minimale des nœuds. Pour ce projet, le modèle a été calibré en utilisant un pourcentage de la base de données d'entraînement de 0,25% à 10% pour déterminer les points de données minimums par nœud.

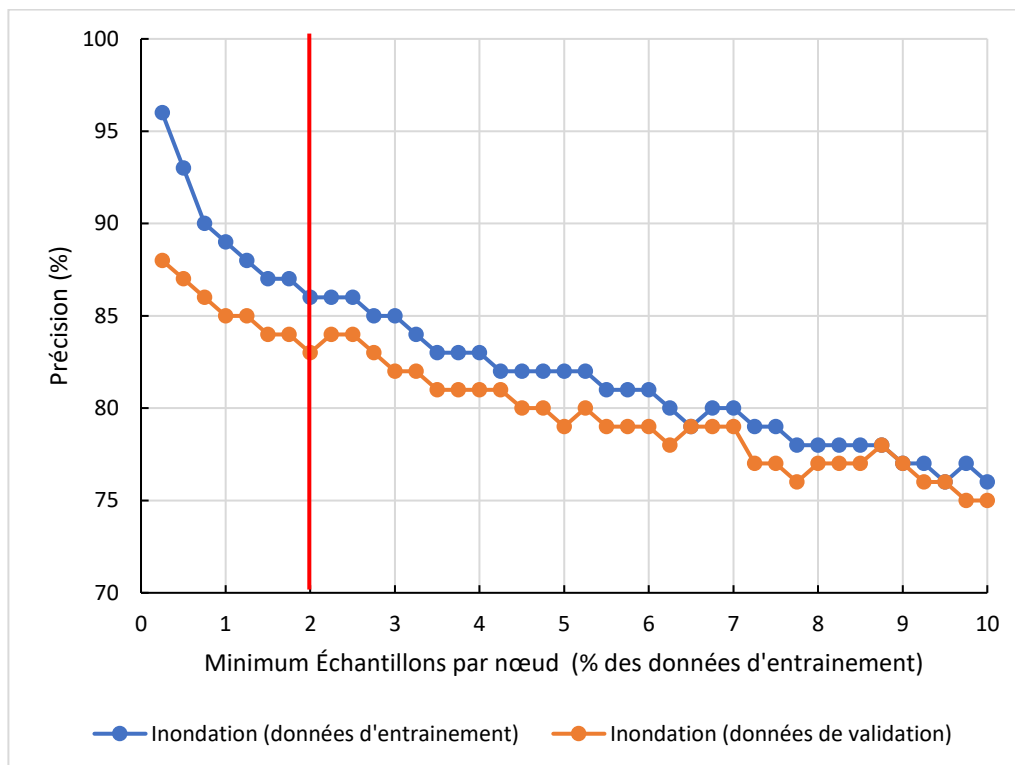
#### 4.1.1 Nombre d'arbres

Pour tester la sensibilité du modèle au nombre d'arbres qu'il contient, il a été exécuté avec de 100 à 3000 arbres, en augmentant le nombre d'arbres de 100 pour chaque exécution. L'augmentation du nombre d'arbres a eu un impact important sur le temps de traitement mais un impact minime sur la précision totale des modèles (1% de différence). Par conséquent, le nombre d'arbres par défaut de 100 a été choisi pour maintenir des temps de traitement courts.

#### 4.1.2 Minimum de points de données par nœud

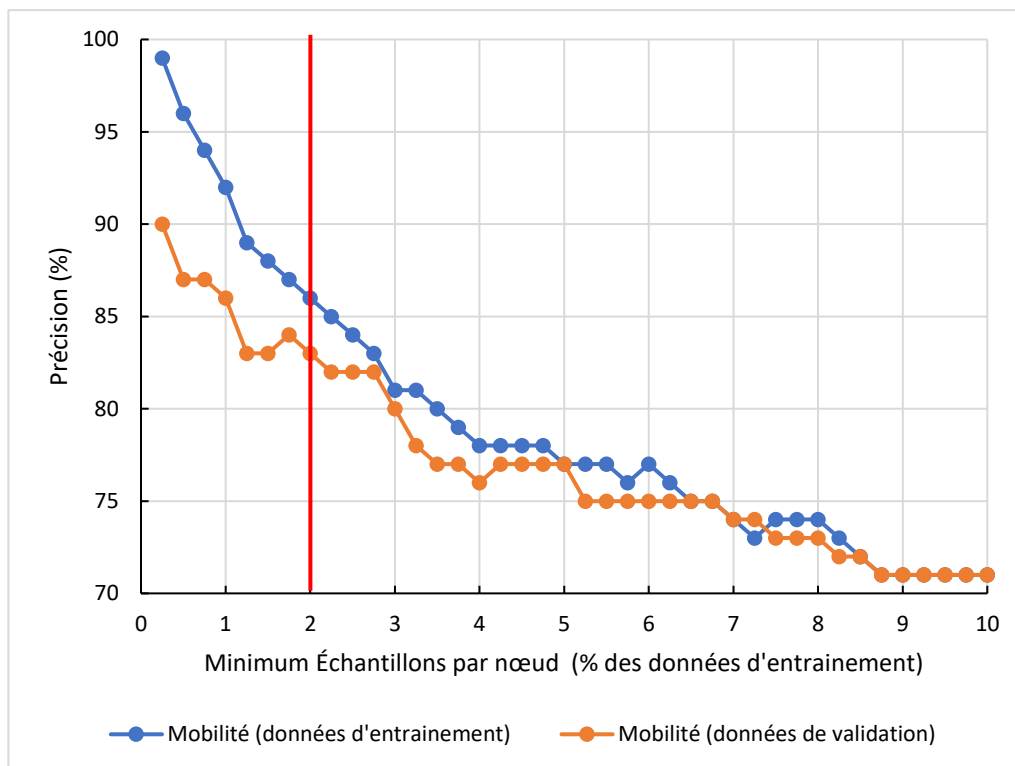
Le paramètre le plus important à calibrer pour le modèle est la quantité minimale de points d'entraînement (points de référence du CRHQ dans notre modèle) par nœud final d'un arbre (*Parmar et al., 2019*), car c'est ce paramètre qui empêche le modèle de se surentraîner. Le modèle continuera donc à diviser les nœuds à l'intérieur de l'arbre de décision jusqu'à ce que les nœuds finaux contiennent le nombre défini par ce paramètre de points de données. Le surentraînement est un problème lorsque le modèle apprend trop bien l'ensemble de données d'entraînement et devient donc seulement bon à détecter ces situations exactes, ce qui peut causer des problèmes lorsqu'on essaie de prédire des points sur lesquels il n'a pas été entraîné. Il est calculé en prenant la différence entre la précision du modèle sur les données d'entraînement et la précision du modèle sur les données de validation. Cet indice permet de s'assurer que notre modèle est capable de se généraliser à des points sur lesquels il n'a pas été entraîné. Pour éviter ce problème, une analyse de sensibilité a été effectuée sur ce paramètre. Le modèle a été exécuté plusieurs fois avec un paramètre d'entrée de minimum de points de données par nœud allant de 0,25% à 10% de la base de données d'entraînement. Les résultats ont ensuite été analysés visuellement pour déterminer le seuil qui permet d'obtenir l'écart le plus faible entre la

précision de l'ensemble de données d'entraînement et la précision de l'ensemble des données de validation, tout en maintenant le niveau de précision le plus élevé. Les figures 9 à 11 montrent qu'à mesure que le nombre minimum de points de données par nœud augmente, la précision du modèle diminue car le modèle n'est plus aussi efficace pour différencier les classes, ce qui réduit également les niveaux de confiance du modèle à chaque point. L'augmentation du nombre minimum de points de données par nœud augmente également la capacité du modèle à généraliser les données à des points sur lesquels il n'a pas été entraîné, ce qui peut améliorer la capacité du modèle à détecter les points chauds dans les bassins versants sur lesquels il n'a pas été entraîné. Il est donc important de trouver un bon équilibre entre les deux. Un seuil de 2 % a ainsi été choisi.

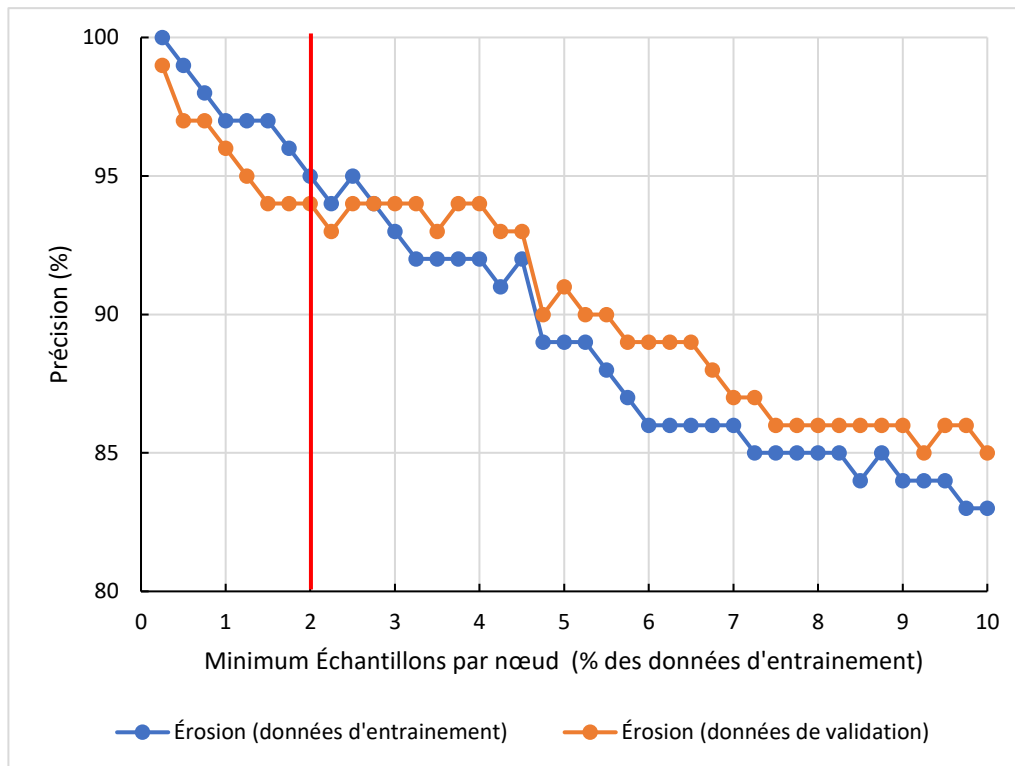


**Figure 9: Sensibilité du modèle d'inondation au nombre minimum de points de données par nœud. Le seuil de 2% est identifié par la ligne rouge verticale.**





**Figure 10: Sensibilité du modèle de mobilité au nombre minimum de points de données par nœud**



**Figure 11: Sensibilité du modèle de type d'érosion au nombre minimum de points de données par nœud**

## 4.2 Scores de précision des modèles

En utilisant le seuil de 2%, le modèle produit de très bons résultats, avec une précision totale de 84,3% pour l'inondation, de 84,7% pour la mobilité et de 94,7% pour le type d'érosion (tableau 9). Les niveaux de surentraînement des modèles varient, la plus grande différence entre les données d'entraînement et de validation étant de 3% pour les modèles d'inondation et de mobilité et la plus faible de 1% pour le type d'érosion (tableau 9). Les modèles de présence d'inondations et de mobilité ont donné de meilleurs résultats dans les bassins versants de la Nicolet et de la Cascapédia que dans celui de la Du Gouffre (tableau 10), bien que les principaux secteurs où les données de prédiction et de validation ne concordaient pas aient été des points où la probabilité de prédiction était faible. Le modèle pour le type d'érosion a donné les meilleurs résultats dans le bassin Du Gouffre, probablement en raison de son contexte géomorphologique unique qui crée des caractéristiques spécifiques (pentes abruptes, berges élevées, taux de confinement principalement élevés ou faibles) permettant à ce type d'outil de les distinguer facilement. Les deux autres bassins versants présentent un ensemble plus diversifié de conditions qui ont provoqué

l'identification des deux types d'érosion, ce qui les rend plus difficiles à distinguer.

**Tableau 9: Résultats de l'entraînement, de la validation et de la précision totale pour le modèle entraîné sur les trois bassins versants pilotes combinés.**

<b>Modèle</b>	<b>Entraînement (%)</b>	<b>Validation (%)</b>	<b>Totale (%)</b>
Inondation	85,8	82,7	84,3
Mobilité	86,4	82,9	84,7
Type d'érosion	95,1	94,3	94,7

**Tableau 10: Résultats de la précision totale du modèle pour chaque bassin versant pilote.**

<b>Bassin</b>	<b>Présence d'inondation (%)</b>	<b>Présence de mobilité (%)</b>	<b>Type d'érosion (%)</b>
Du Gouffre	80,3	80,9	96,8
Nicolet	84,9	87,0	86,6
Cascapédia	89,9	86,2	83,3

### 4.3 Validation croisée

En testant l'importance de la répartition aléatoire entre les données d'entraînement et de validation pour la précision du modèle, il a été constaté que cette répartition n'avait que très peu d'impact sur la précision. La précision moyenne lorsque les données sont réparties dix fois avec des distributions différentes est de 86, 86 et 95% pour le PFM, le PMM et le TEM respectivement, avec des écarts-types très faibles, le maximum étant de 1 %.

### 4.4 Matrice de confusion

Les modèles étant de type binaire, il est possible de distinguer 4 catégories, soit les vrais positifs (VP), les faux positifs (FP), les faux négatifs (FN) et les vrais négatifs (VN) (tableau 11). Si l'on prend par exemple le modèle de mobilité, les VP sont les points où le modèle prédit correctement qu'il y aura de la mobilité dans un secteur mobile. Les FP sont ceux estimés mobiles alors qu'en réalité ils sont stables. Les FN sont les éléments estimés non mobiles alors qu'ils le sont,

tandis que les VN sont les éléments prédits comme étant stables et qui le sont bel et bien (le même raisonnement s'applique au modèle d'inondation et pour le type d'érosion).

Cette section présente des matrices de confusion pour les trois modèles ainsi que les indices de validation associés, soit la précision de chaque classe ( $P_{0/1}$ ), la précision totale ( $P_t$ ), le rappel (R), la spécificité (S), le score kappa, l'aire sous la courbe (AUC) de la courbe de la fonction d'efficacité du récepteur (ROC) et le score F1 qui sont des indices de validation importants pour les modèles de classification (Tiwari, 2022). Les formules de ces indices figurent dans le tableau 11. La précision quantifie le pourcentage de points correctement prédits par rapport à tous les points prédits dans cette classe, tandis que le rappel et la spécificité mesurent le pourcentage de VP et VN, respectivement, par rapport au nombre de points observés dans cette classe. Le score F1 est une mesure utilisée pour incorporer ces deux mesures en une seule qui donne un portrait plus général de la qualité de la prédiction. Le coefficient kappa est similaire à la précision mais intègre la probabilité de sélectionner au hasard la bonne classe. Les valeurs supérieures à 60 % sont considérées comme une forte concordance, tandis que les valeurs supérieures à 80 % sont considérées comme une concordance presque parfaite (Kavzoglu & Teke, 2022). L'AUC est une mesure couramment utilisée pour comparer les modèles d'apprentissage automatique. Cette méthode est utilisée plus pour les modèles de régression, mais comme notre modèle est une classification binaire, elle peut être utilisée pour évaluer notre modèle. Tout d'abord, un diagramme ROC est généré, dans lequel le seuil déterminant le pourcentage de votes nécessaires pour être classé dans la classe positive est modifié et le rappel est comparé au taux de faux positifs ( $FP/(FP + VP)$ ). L'aire sous cette courbe est calculée et utilisée comme métrique AUC pour comparer la capacité des modèles à prédire les classes de manière fiable (figure 12). Les valeurs comprises entre 80 et 90 % sont considérées comme très bonnes et les valeurs supérieures à 90 % sont considérées comme excellentes (Youssef et al., 2023). Puisque dans cette étude nous voulons nous assurer qu'aucun secteur risquant d'être affecté par les aléas fluviaux ne soit prédit comme n'étant pas à risque (nécessité d'éviter les faux négatifs (FN)), la mesure la plus importante à prendre en compte est le rappel, qui quantifie le pourcentage des secteurs à risque observés qui ont été correctement prédits comme étant à risque. En maximisant la valeur de rappel, nous nous assurons que le modèle ne prédise pas des secteurs à risque comme n'étant pas à risque. Pour le modèle de type d'érosion, la spécificité et le rappel sont tous deux importants car il ne s'agit pas de distinguer risque/pas de risque, mais plutôt le type d'aléa. Il est donc important que les deux soient prédits avec précision, même si les conséquences d'une mauvaise prédiction sont sans doute moins graves que les conséquences de ne pas avoir identifié correctement des secteurs mobiles.

Pour les inondations, Le tableau 12 indique que la précision pour les deux classes (présence/absence) est approximativement la même avec des scores de 81 et 84 % pour l'absence et la présence d'inondations, respectivement. Cependant, le rappel était supérieur à la spécificité avec une bonne valeur de

85% des secteurs observés présentant un risque d'inondation qui ont été correctement prédits. Par conséquent, le modèle est légèrement plus performant pour prédire les zones susceptibles d'être inondées que les zones non susceptibles d'être inondées, comme le montrent les scores F1.

Pour la mobilité, la précision est 85 et 82 % pour l'absence et la présence de mobilité, respectivement (tableau 13). Le rappel présentait une grande différence par rapport à la spécificité, la présence de mobilité étant correctement prédite 90 % du temps par rapport à 73 % pour les zones stables. Avec un rappel de 90%, ce modèle prédit avec précision la majorité des secteurs mobiles, mais il produit encore un grand nombre de FP qui devront être améliorés en ajoutant plus de données d'entraînement pour les secteurs stables. Par conséquent, le modèle est meilleur pour détecter les zones mobiles que pour détecter les zones stables, ce qui se reflète dans les scores F1 (tableau 13).

Les indices étaient meilleurs pour la migration latérale que pour l'incision, avec une précision de 96 % et un rappel de 97 % pour la migration latérale, contre une précision de 89 % et une sensibilité de 85 % pour l'incision (tableau 14). On peut donc conclure que le modèle détecte mieux la migration latérale que l'incision.

Ces résultats indiquent que bien que les modèles puissent être biaisés vers certaines classes, ils le sont vers la classe qui présente le risque le plus élevé d'endommager les infrastructures routières. S'il existe un risque d'aléas fluviaux, il y a donc de fortes chances que le modèle le détecte, mais il créera également des FP. En gestion de risque, les faux positifs sont préférables aux faux négatifs puisqu'ils incitent à un excès de prudence plutôt qu'à conclure qu'un site n'est pas à risque alors qu'il l'est en réalité.

Dans l'ensemble, les indices de validation (tableaux 12 – 14 et figure 12) sont d'une qualité similaire à celle d'autres études récentes portant sur les applications de l'intelligence artificielle pour les aléas naturels (Fang et al., 2022; Kavzoglu & Teke, 2022; Avand et al., 2023; Youssef et al., 2023).

**Tableau 11: Tableau explicatif de la matrice de confusion et des variables permettant de juger de la qualité des modèles**

Aléa				
Prédictions	Observations			
		0	1	Précision ( $P_x$ )
	0	VN	FN	$VN/(VN+FN) (P_0)$
	1	FP	VP	$VP/(FP+VP) (P_1)$
		$VN/(VN+FP)$ Spécificité (S)	$VP/(FN + VP)$ Rappel (R)	$(VN + VP)/$ $(VN+VP+FN+FP) (P_t)$
Score F1		$2 * P_t * R /$ $(P_t + R)$		
Kappa		$2 * (VP * VN - FN * FP) /$ $((VP + FP) * (FP + VN) + (VP + FN) * (FN + VN))$		
AUC		Aire sous la courbe		

**Tableau 12: Matrice de confusion pour les inondations**

Inondations				
	Observations			
Prédictions		0	1	Précision
	0	155	36	81%
	1	42	219	84%
		79%	86%	83%
Score F1		84.5%		
Kappa		64.8%		
AUC		90%		

**Tableau 13: Matrice de confusion pour la mobilité**

Mobilité				
	Observations			
Prédictions		0	1	Précision
	0	120	22	85%
	1	44	200	82%
		73%	90%	83%
Score F1		86.4%		
Kappa		64.4%		
AUC		90%		

**Tableau 14: Matrice de confusion pour le type d'érosion**

Type d'Érosion				
		Observations		
Prédictions		0	1	Précision
	0	39	5	89%
	1	7	161	96%
		85%	97%	94%
Score F1		95.4%		
Kappa		83.1%		
AUC		94%		

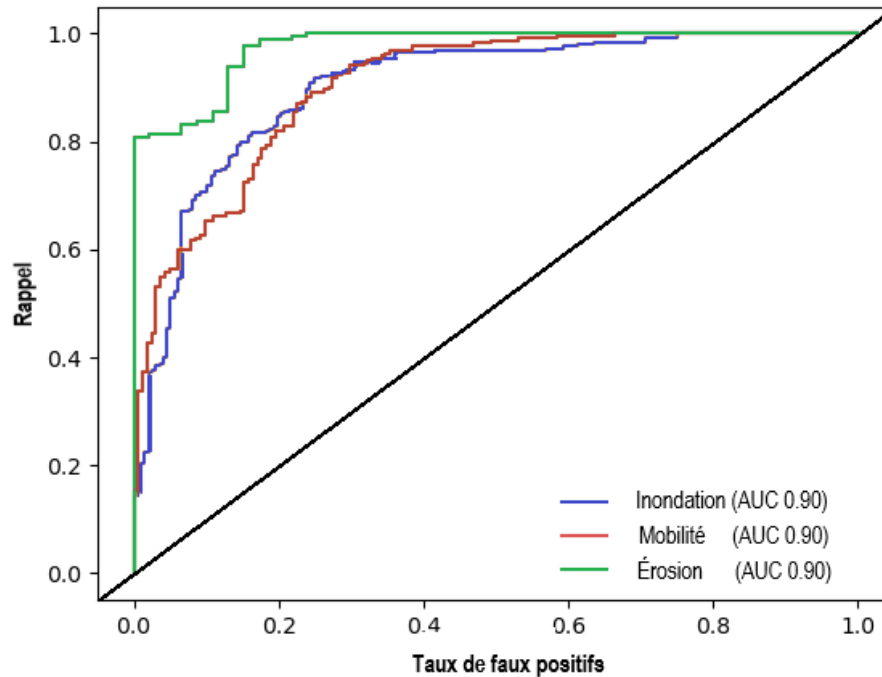


Figure 12: Graphique ROC AUC pour les trois modèles.

## 5. INTERPRÉTATION DES RÉSULTATS

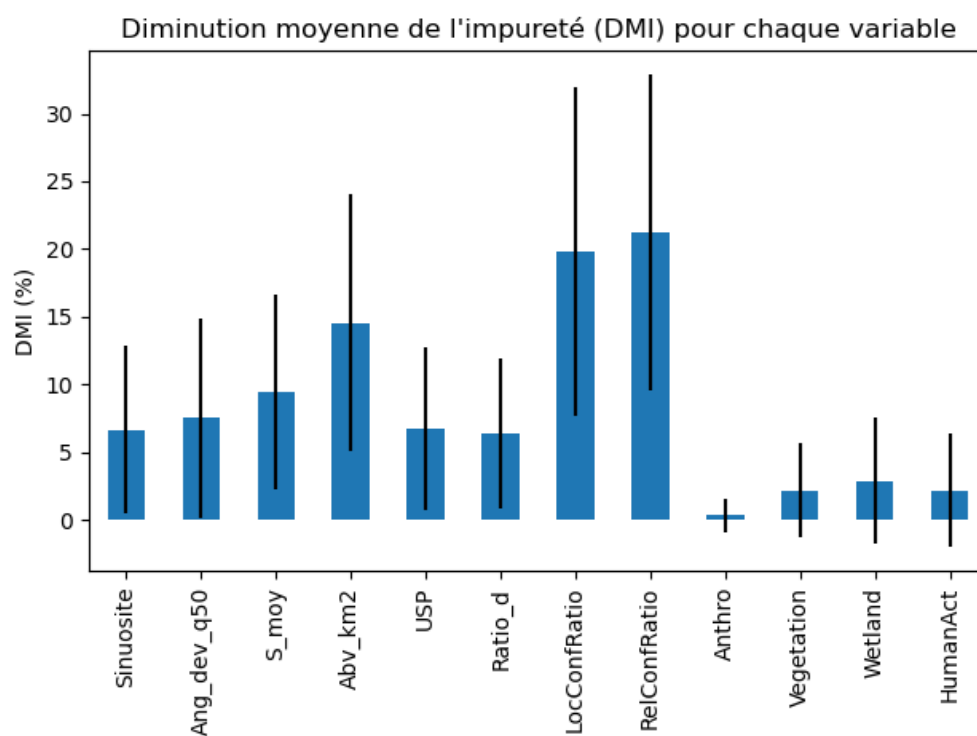
### 5.1 Importance de chaque variable

Comme mentionné dans la section 3.4.4, l'outil produit 2 graphiques pour chaque modèle qui montrent le DMI et le DMP pour chaque variable. Le DMI montre l'importance de chaque variable dans la division des données pendant le processus de création de l'arbre et garantit que chaque nœud contient des points de données d'une seule classe. De son côté, le DMP indique l'importance d'une variable dans l'étape de prédiction en montrant à quel point la précision du modèle sera affectée si de mauvaises données sont entrées pour cette variable. Par conséquent, des valeurs plus élevées de DMP indiquent que le modèle dépend fortement de la saisie de bonnes données dans ces champs, sinon la précision sera considérablement réduite. Ces indices nous permettent d'interpréter quelles variables le modèle utilise afin de prédire la présence d'aléas.

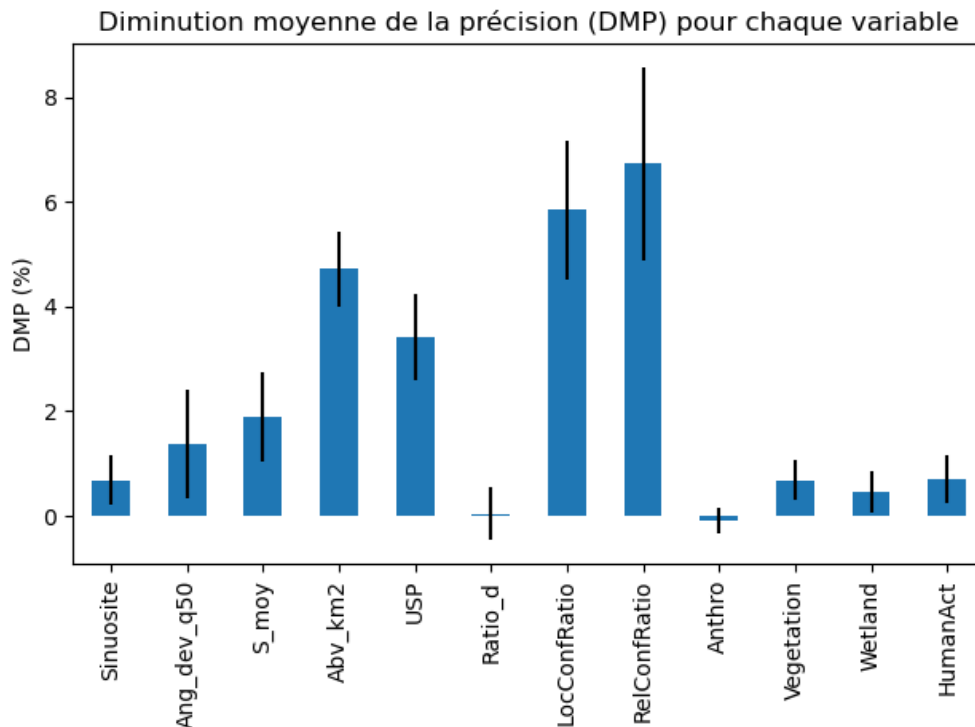
### **5.1.1 Inondation**

Pour le modèle d'inondation, les deux variables de confinement (LocConfRatio et RelConfratio) sont assez importantes pour déterminer si une inondation se produira, ce qui est prévisible car plus la rivière est confinée, moins il y a de chance qu'une inondation survienne (figure 13). Le confinement relatif est important car il aide le modèle à trouver les parties les moins confinées des rivières confinées où le débit accumulé peut alors s'étendre (figure 13). Les autres variables qui sont importantes pour différencier les régions sujettes aux inondations de celles qui ne le sont pas sont la localisation dans le bassin versant (aire de drainage (Abv\_km2) et pente moyenne en amont (S\_moy)). Les variables de présence/absence sont moins puissantes dans ce type de modèle car un seul seuil est possible. Par ailleurs, le niveau de végétation, la présence de terres agricoles et de milieux humides sont également des variables importantes dans ce type de modèle pour déterminer les inondations, ce qui est également un résultat attendu puisque les milieux humides se trouvent fréquemment dans les zones inondables et que la végétation peut limiter les inondations. Les forêts ou les activités humaines qui éliminent ces forêts auront donc un impact sur les inondations. La figure 14 montre que le confinement de la rivière, la puissance spécifique et l'aire de drainage sont les facteurs les plus importants pour s'assurer que le modèle prédit la classe appropriée pour l'inondation pour chaque point de référence car ils réduisent la précision du modèle de 3,5 % à 7,5 % lorsqu'ils sont randomisés, les deux variables de confinement étant les plus importantes. L'importance des variables de présence/absence n'apparaît pas dans ce type de graphique en raison de la forte probabilité qu'elles reçoivent la même valeur que celle qu'elles avaient précédemment lors du processus de randomisation des variables. Cependant, les trois mêmes éléments (le niveau de végétation, la présence de terres agricoles et de milieux humides) qu'identifiés précédemment sont importants pour déterminer l'inondation.





**Figure 13: Diminution moyenne de l'impureté pour chaque variable du modèle d'inondation, les lignes noires représentant l'écart type pour chaque variable.**



**Figure 14: Diminution moyenne de la précision pour chaque variable du modèle d'inondation, les lignes noires représentant l'écart type pour chaque variable.**

### 5.1.2 Mobilité

La figure 15 montre que lors de l'entraînement du modèle de mobilité, les variables les plus importantes étaient l'emplacement du point dans le bassin versant et les caractéristiques en amont (pente moyenne, ratio de drainage et aire de drainage). Ceci est attendu pour le modèle de mobilité car ces variables représentent la quantité et le type de sédiments qui atteignent le point de référence, ce qui est un facteur important pour déterminer la mobilité du cours d'eau. De plus, l'angle entre le point de référence local et les deux points adjacents, ainsi que la sinuosité du tronçon, sont utilisés pour déterminer la mobilité, ce qui aide le modèle à identifier les méandres dans lesquels l'énergie exercée sur les berges augmente et où les chances de mobilité sont accrues. Le confinement est toutefois moins important pour le modèle de mobilité que pour celui d'inondation. Ceci est prévisible puisque la mobilité comprend à la fois l'incision et la migration latérale et peut donc se produire dans les rivières confinées (principalement de l'incision) et non confinées (principalement de la mobilité latérale). Bien que la variable de confinement relatif conserve son importance lors de la vérification de la DMP (figure 16), cela est probablement dû au fait qu'elle est utilisée pour déterminer les aires où le confinement de la rivière diminue, permettant à l'énergie accumulée d'éroder les berges plus

érodables par rapport à la zone géologiquement confinée en amont. Les variables les plus importantes lors de l'étape de prédiction sont l'angle moyen de déviation et les variables proxy pour les caractéristiques en amont, ce qui indique que la présence de méandres et le type d'environnement en amont (qui a un impact sur le volume et la taille des sédiments qui pourraient atteindre ce point) sont fondamentaux dans la capacité du modèle à détecter la mobilité.

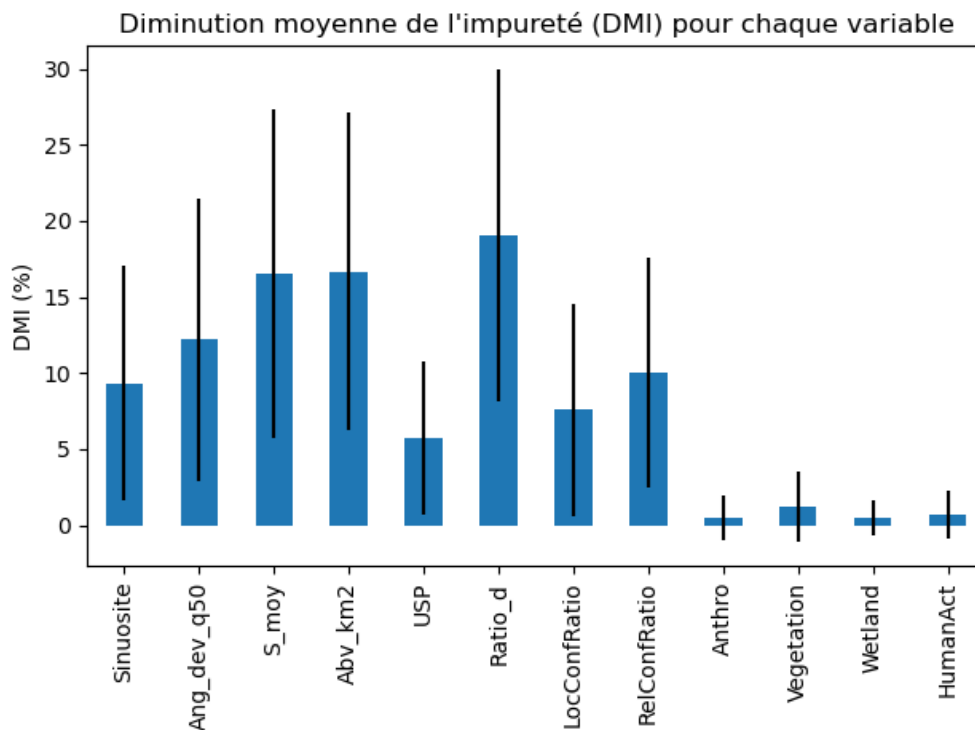


Figure 15: Diminution moyenne de l'impureté pour chaque variable du modèle de mobilité, les lignes noires représentant l'écart type pour chaque variable.

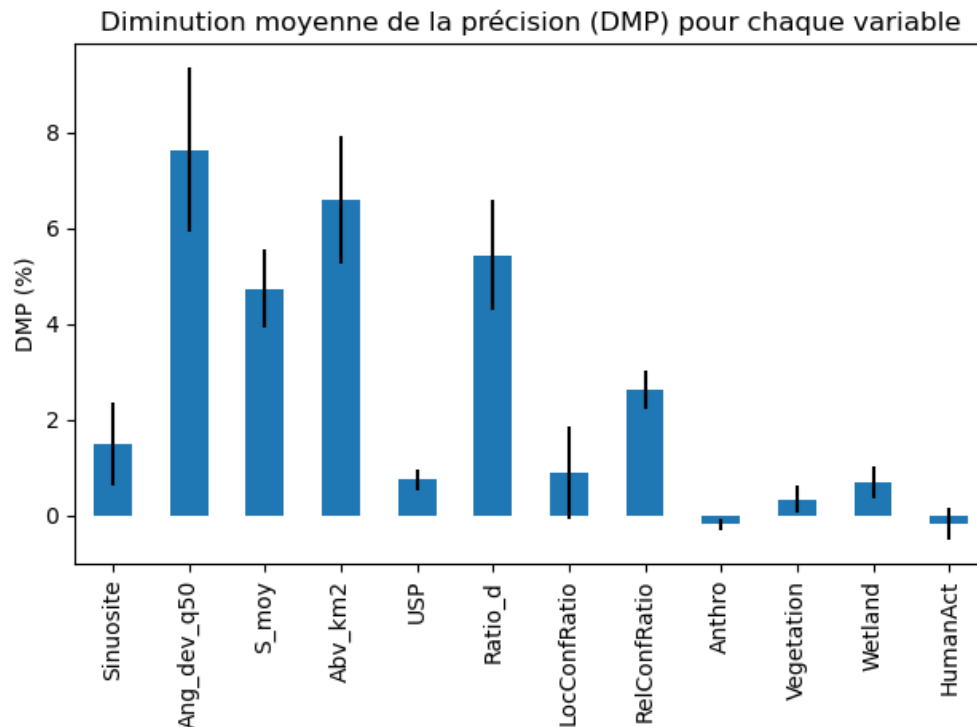
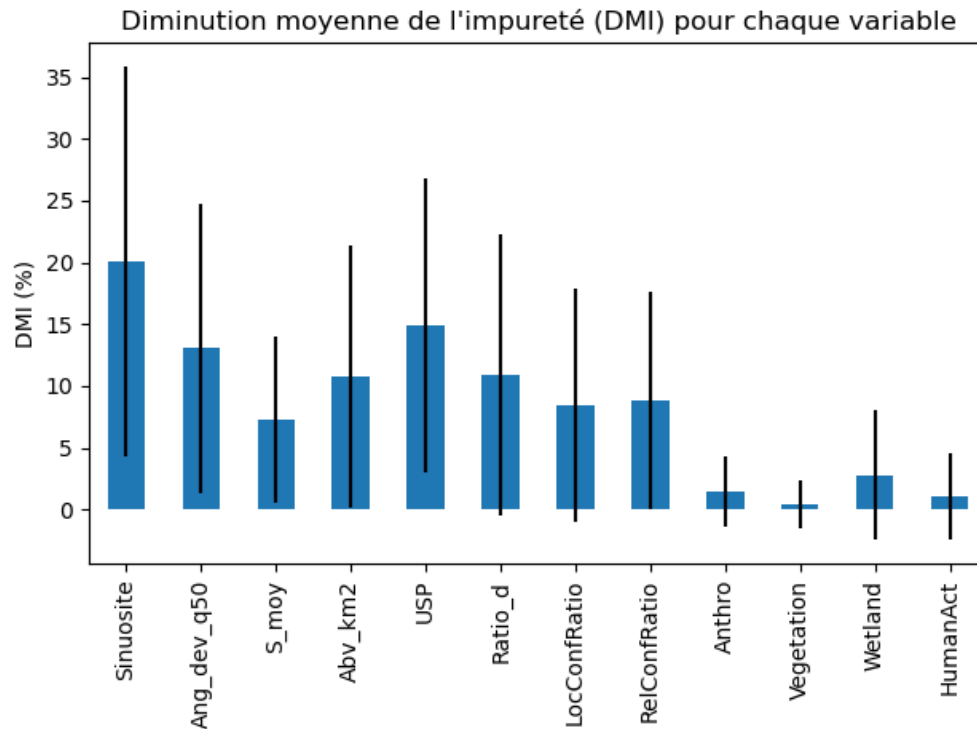


Figure 16: Diminution moyenne de la précision pour chaque variable du modèle de mobilité, les lignes noires représentant l'écart type pour chaque variable.

#### 5.1.4 Type d'érosion

Durant l'entraînement du modèle pour différencier la migration latérale et l'incision, les variables ont toutes des niveaux d'importance à peu près identiques. Seule la sinuosité du tronçon homogène montre des niveaux d'importance significativement plus élevés, suivie par la puissance spécifique (figure 17). Il s'agit probablement des variables les plus importantes pour ce modèle, car l'angle entre les points permet au modèle d'identifier les méandres susceptibles de migrer latéralement, tandis que les zones où se produit de l'incision sont susceptibles d'avoir une puissance spécifique élevée, ce qui aide le modèle à distinguer les zones d'incisions et de migrations latérale. Bien que toutes les variables restent importantes pour réduire l'impureté du modèle. La figure 18 révèle que la sinuosité et l'angle entre les points sont les plus importants pour s'assurer que le modèle prédit la bonne classe, car ils contribuent à bien distinguer les tronçons linéaires susceptibles d'être incisés des tronçons sinueux où la migration latérale est plus probable. La puissance spécifique est la deuxième variable la plus importante, mais elle reste proche des autres, qui ont toutes la même importance pour garantir la prédiction de la bonne classe. Par conséquent, bien que de nombreuses variables soient également importantes pour distinguer la migration latérale de l'incision, la

sinuosité et l'angle moyen de déviation sont les variables prédictives les plus utiles.



**Figure 17: Diminution moyenne de l'impureté pour chaque variable du modèle de type d'érosion, les lignes noires représentant l'écart type pour chaque variable.**

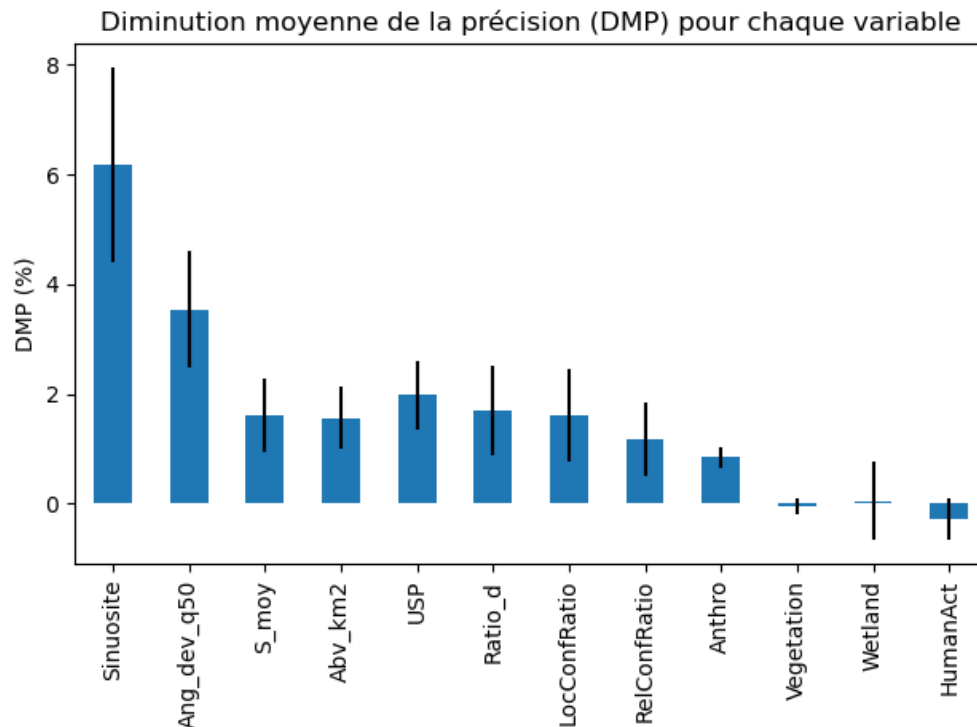
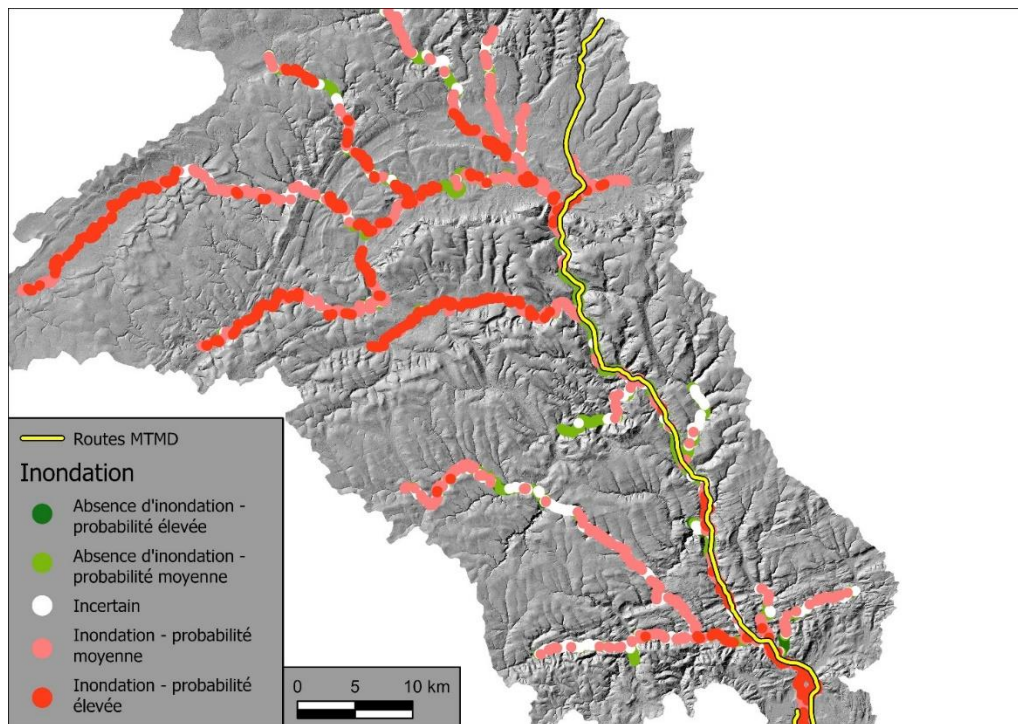


Figure 18: Diminution moyenne de la précision pour chaque variable du modèle de type d'érosion, les lignes noires représentant l'écart type pour chaque variable.

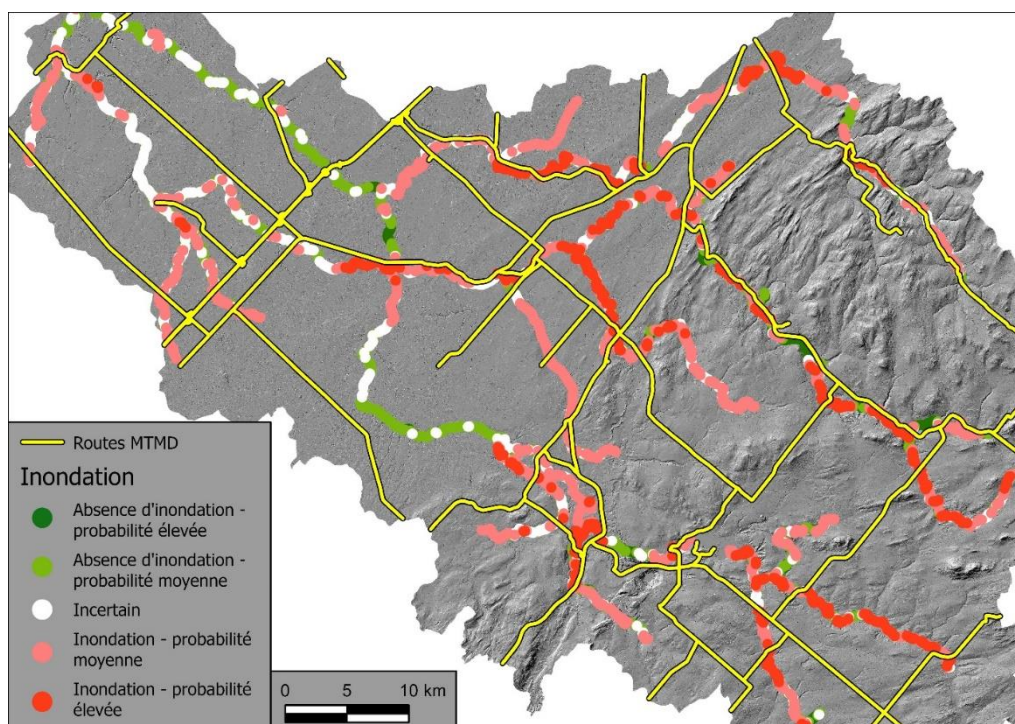
## 5.2 Résultats de probabilités d'aléas fluviaux pour les 3 bassins pilotes

Cette section présente des cartes montrant les prédictions pour la totalité des bassins versants, des cas de validation plus ponctuels avec des exemples des limites du modèle qui sont présentés à l'annexe A. Il est important de tenir compte du fait qu'à cette échelle, de nombreux points se chevauchent et que la classe située au bas de la légende est générée en dernier et donc au-dessus des autres points, ce qui peut rendre l'interprétation des résultats difficile. Il est recommandé d'utiliser des méthodes de regroupement (regroupement dynamique, cartes thermiques, etc.) lors de l'interprétation des résultats à grande échelle. Les figures 19 à 22 montrent le rendu cartographique de la modélisation des probabilités d'inondation pour les bassins de la Cascapédia, Nicolet et Du Gouffre, respectivement. Dans la plupart des cas, il est possible de distinguer cet aléa, mais il reste un certain nombre de points de référence qui demeurent incertains. Les résultats sont dans l'ensemble assez équilibrés, les données étant répartis de manière bimodale autour de la classe incertaine (faible niveau de confiance), avec une légère préférence pour la classe de

présence d'inondation. Cela est probablement dû au fait que notre modèle a été entraîné principalement sur les grandes rivières en aval des bassins versants, qui sont plus sujettes aux inondations que les cours d'eau situés en amont. L'incertitude était relativement stable entre les bassins versants avec 18%, 15% et 19% pour la Cascapédia, Du Gouffre et Nicolet, respectivement.

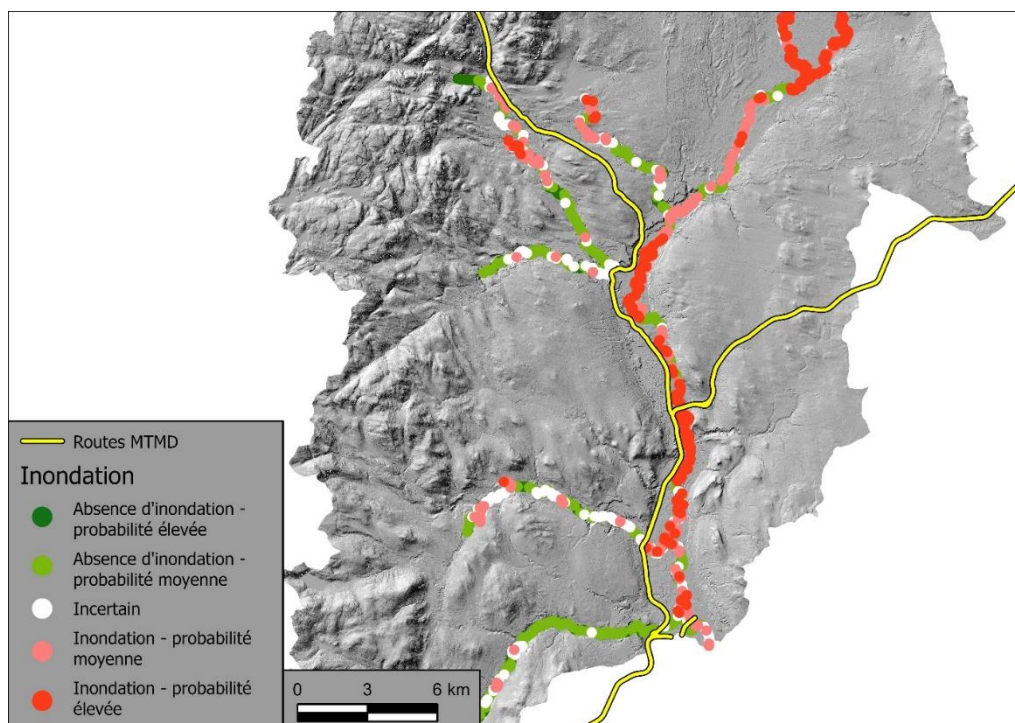


**Figure 19: Carte des prévisions des inondations avec le modèle pour le bassin versant de la Cascapédia**

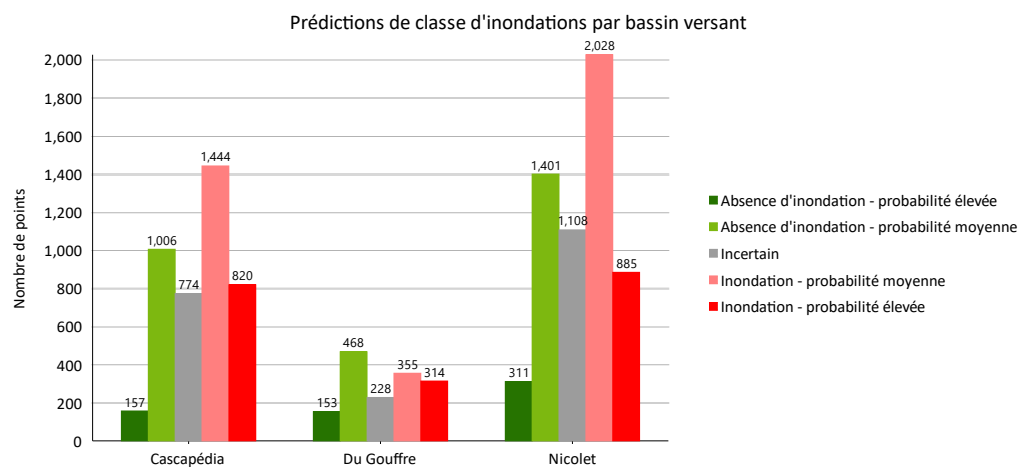


**Figure 20: Carte des prévisions des inondations avec le modèle pour le bassin versant de la Nicolet**



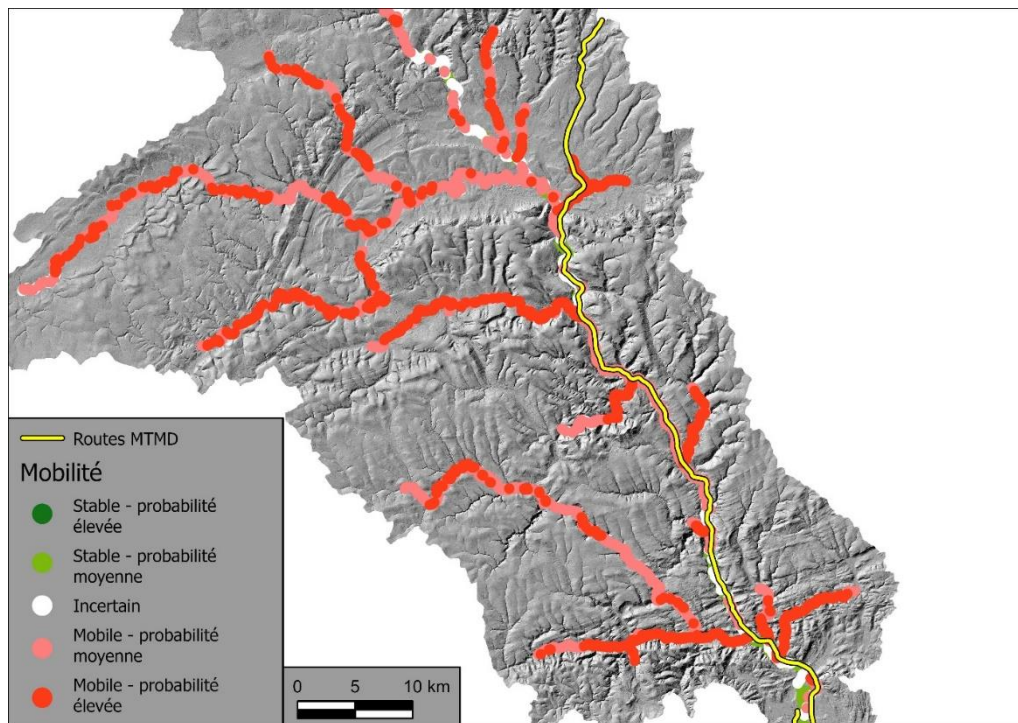


**Figure 21: Carte des prévisions des inondations avec le modèle pour le bassin versant de la Du Gouffre**

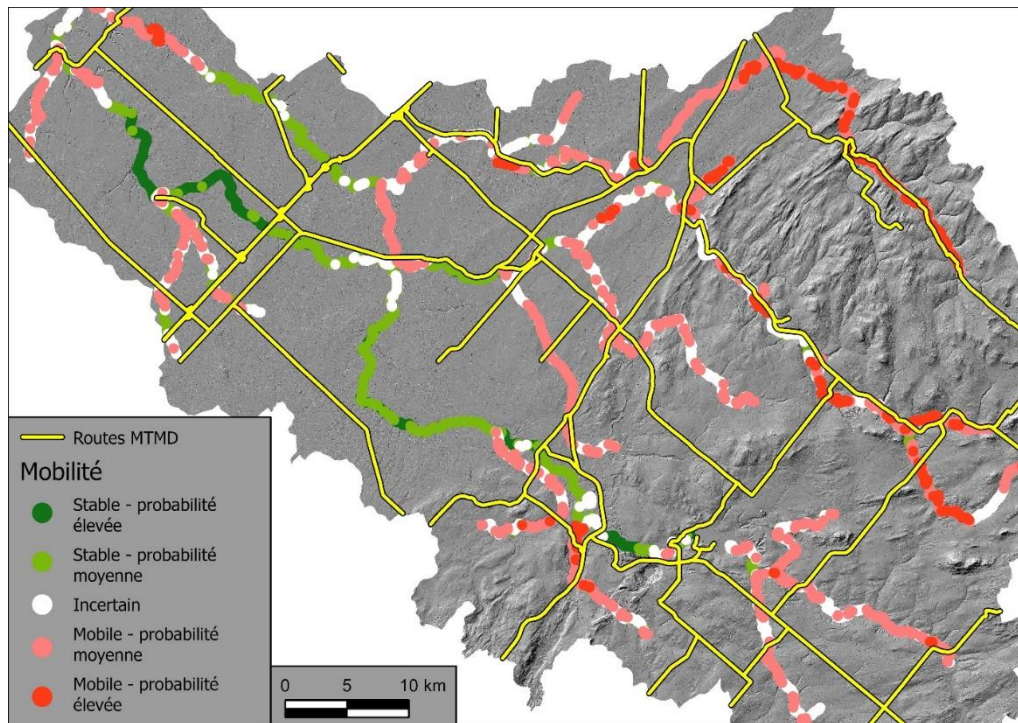


**Figure 22: Prédictions de classe d'inondation par bassin versant**

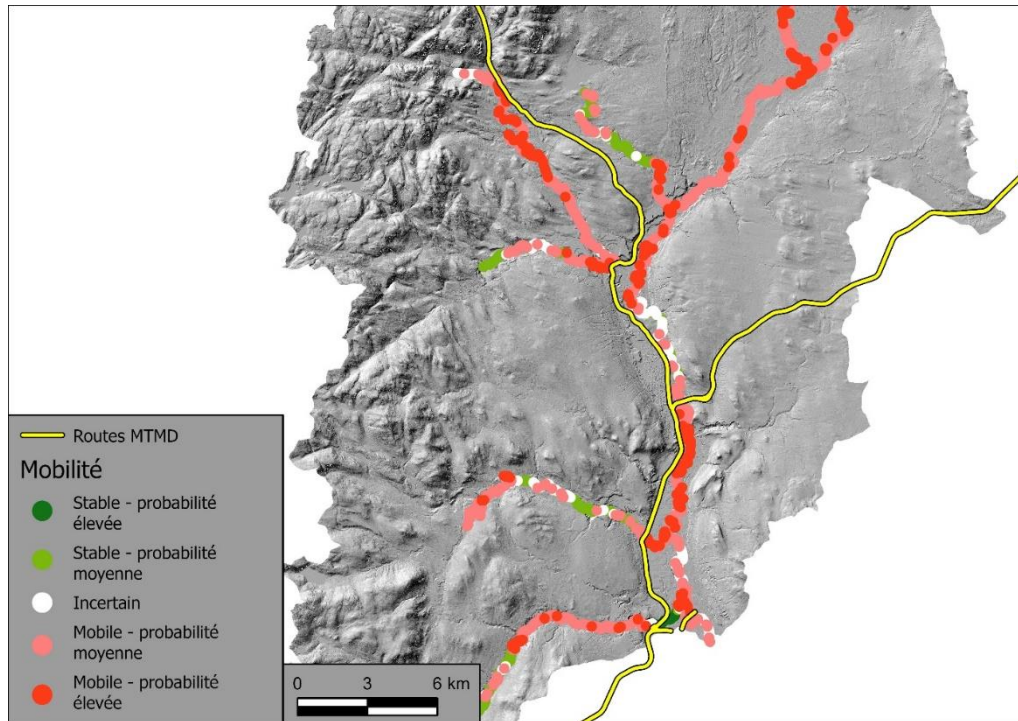
Les résultats du modèle prédisant la mobilité sont présentés aux figures 23 à 26. La majorité des points de référence dans le bassin-versant de la Cascapédia (figures 23 et 26) et Du Gouffre (figures 25 et 26) indiquent une probabilité d'aléa de mobilité. La rivière Nicolet (figures 24 et 26) est moins dynamique, avec de nombreuses zones classées comme incertaines et stable. Le modèle de mobilité était moins distribué d'une façon bimodale autour de la classe incertaine, favorisant la classe mobile, cela pourrait être dû au fait que le modèle a été entraîné et utilisé sur des bassins versants très mobiles (Du Gouffre et Cascapédia). Par conséquent, lors de l'interprétation des résultats de ce modèle, il est important de tenir compte du fait que de nombreux sites stables tomberont dans la catégorie incertaine, ou alors faible probabilité de mobilité. La prédiction brute des sites stables par rapport aux sites mobiles est indiquée dans le champ MobPred et le niveau de confiance indiqué dans le champ MobProb doit être pris en compte lors de l'interprétation des résultats. Ce problème pourrait être résolu en ajoutant davantage d'observations dans les zones stables. L'incertitude du modèle de mobilité varie d'ailleurs fortement selon les bassins versants : 9 %, 17 % et 29 % pour la Cascapédia, Du Gouffre et Nicolet, respectivement.



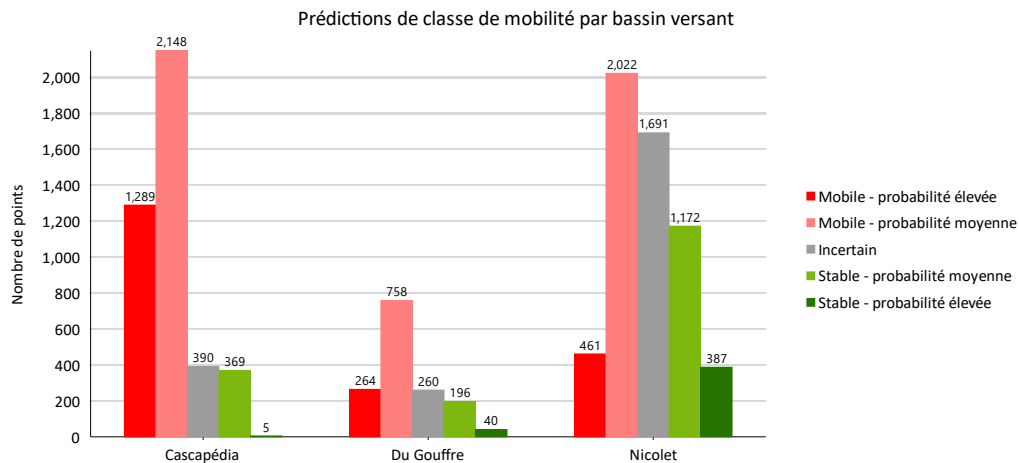
**Figure 23: Carte des prévisions de mobilité du modèle pour le bassin versant de la Cascapédia**



**Figure 24: Carte des prévisions de mobilité du modèle pour le bassin versant de la Nicolet**



**Figure 25: Carte des prévisions de mobilité du modèle pour le bassin versant de la Du Gouffre**

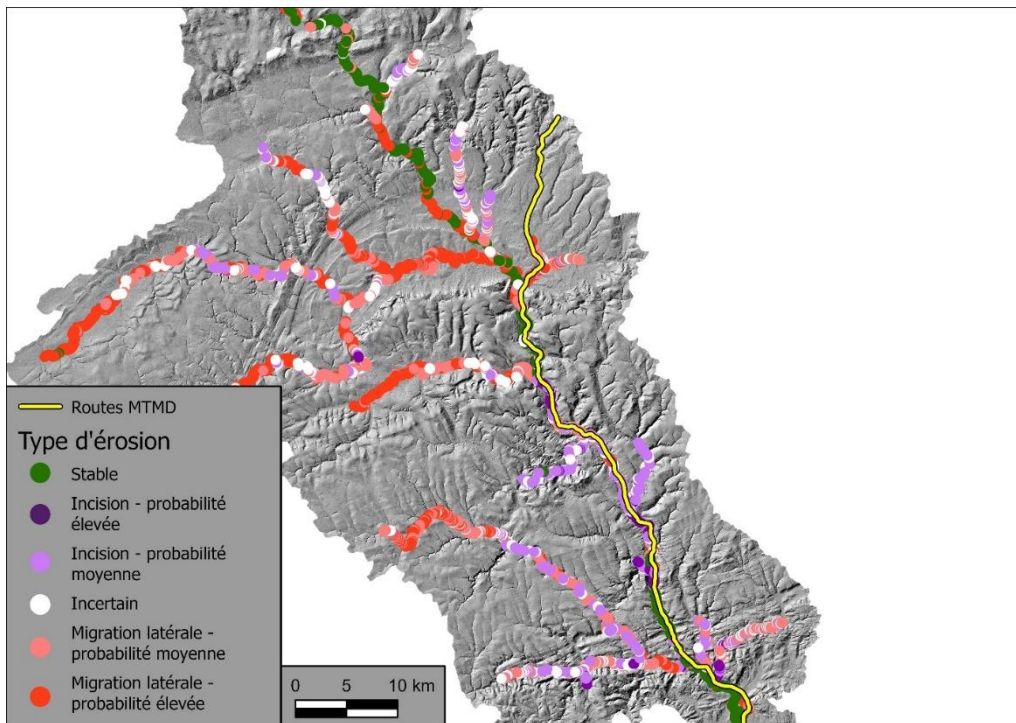


**Figure 26: Prédictions de classe de mobilité par bassin versant**

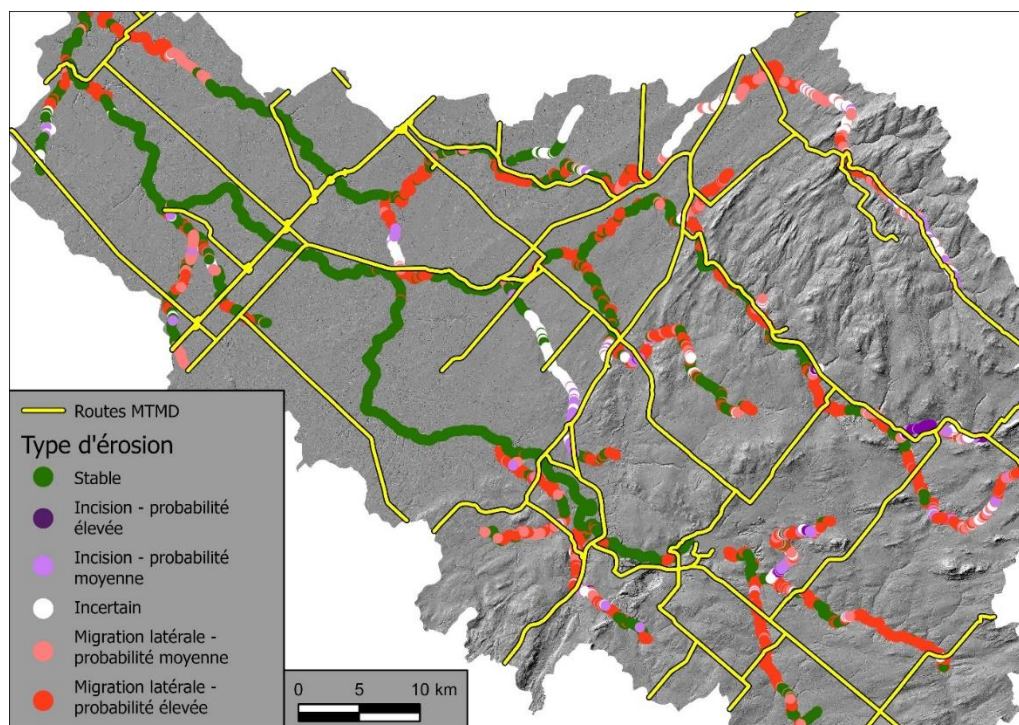
Enfin, le type d'érosion est présenté aux figures 27 à 30. La migration latérale domine largement dans le bassin-versant de la rivière Nicolet (figure 28) alors que la probabilité d'incision est assez élevée dans le chenal principal de la rivière



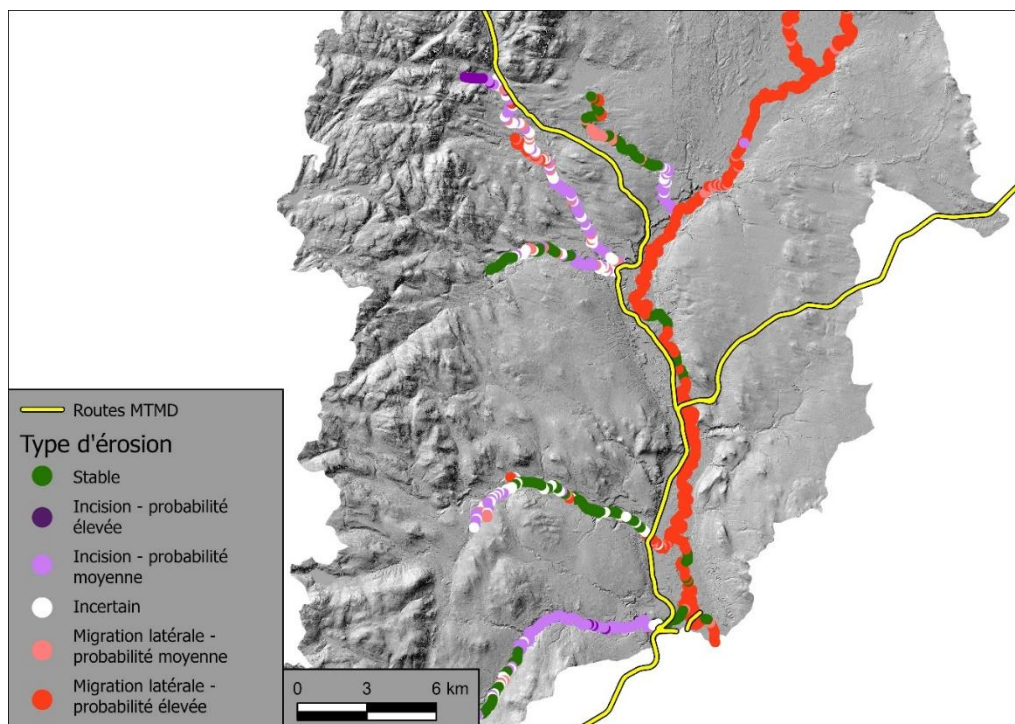
Cascapédia (figure 27) et dans certains tributaires de la rivière du Gouffre (figure 29). Comme indiqué à la figure 30, la migration latérale est prédite par le modèle à des taux beaucoup plus élevés que l'incision. Cela s'explique par le fait qu'il est beaucoup plus facile d'identifier les sites où la migration latérale pourrait se produire et que le modèle est entraîné sur des rivières plus grandes situées plus en aval, où la migration latérale est plus souvent un problème, plutôt que sur des cours d'eau de tête où l'incision est plus fréquente. Il est donc difficile d'obtenir un équilibre entre les deux aléas dans ce modèle. L'incertitude dans le modèle de type d'érosion était assez uniforme parmi les 3 bassins avec 14 %, 10 % et 8 % pour la Cascapédia, Du Gouffre et Nicolet, respectivement.



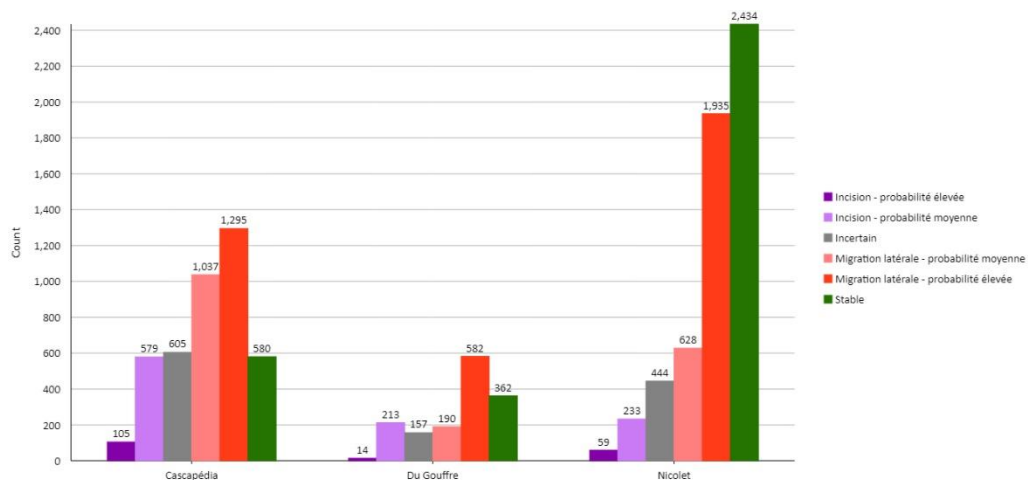
**Figure 27: Carte des prévisions du type d'érosion du modèle pour le bassin versant de la Cascapédia**



**Figure 28 : Carte des prévisions du type d'érosion du modèle pour le bassin versant de la Nicolet**



**Figure 29: Carte des prévisions du type d'érosion du modèle pour le bassin versant de la Du Gouffre**

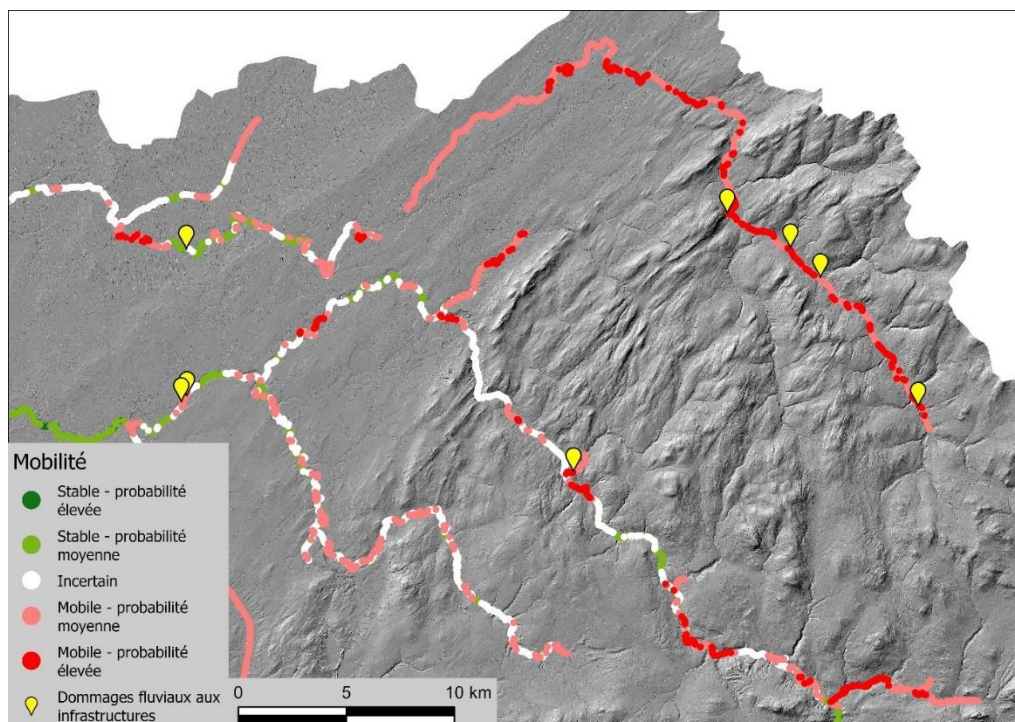


**Figure 30: Prédictions de classe de type d'érosion par bassin versant**

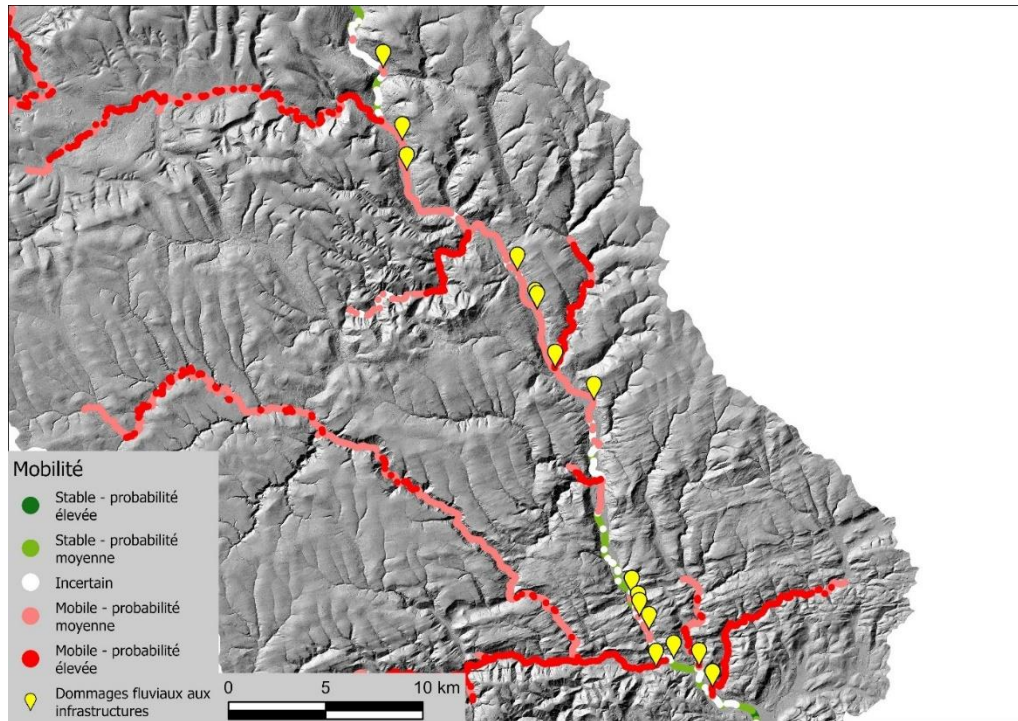
### **5.3 Validation à l'aide des bases de données du MTMD**

Pour valider davantage le modèle, 24 références aux dommages causés aux infrastructures par les processus fluviaux ont été trouvées dans les bases de données du MTMD, soit 8 pour le Nicolet et 16 pour la Cascapédia. Ces références sont tirées de la base de données Planification, Programmation, et Suivi (PPS) pour la Nicolet et des travaux de Tristan Caron (Direction de la coordination et des relations avec le milieu - Gaspésie-Îles-de-la-Madeleine) pour la Cascapédia. Sept des 8 points de la Nicolet ont été trouvés dans des zones prédites comme ayant une probabilité moyenne ou élevée d'être mobiles, tandis que le dernier a été trouvé dans une zone prédite comme étant stable bien qu'avec une certaine incertitude juste en amont du pont endommagé (figure 31). Dans la Cascapédia, tous les points se trouvaient dans des zones mobiles, à l'exception de cinq points situés près du centre, qui présentaient tous des niveaux élevés d'incertitude, avec trois des cinq points pointant vers la mobilité et deux vers la stabilité (figure 32). Cela montre que le modèle fonctionne bien dans ces deux bassins versants, bien qu'avec une certaine incertitude dans la Cascapédia, probablement en raison des problèmes de qualité des données dans le CRHQ pour ce bassin versant interférant avec la formation du modèle. Malheureusement, les bases de données du MTMD ne contiennent aucune référence claire sur les dommages causés aux infrastructures dans le bassin de la Du Gouffre qui auraient été causés par des processus fluviaux.





**Figure 31 : Carte des références aux dommages aux infrastructures liés aux processus fluviaux dans la base de données du MTMD comparée aux résultats de la probabilité de mobilité du modèle (Nicolet).**



**Figure 32 : Carte des références aux dommages aux infrastructures liés aux processus fluviaux dans la base de données du MTMD comparée aux résultats de la probabilité de mobilité du modèle (Cascapédia).**

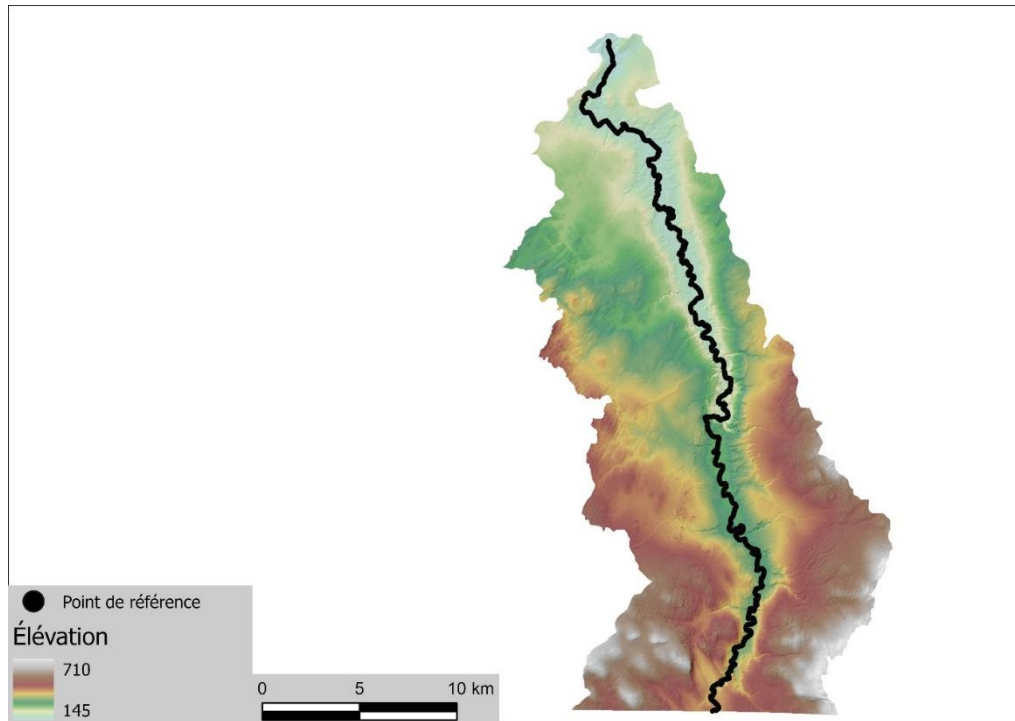
#### 5.4 Test sur la Coaticook (bassin versant non entraîné)

Pour tester le modèle sur un bassin versant sur lequel il n'a pas été entraîné, nous avons utilisé la rivière Coaticook où les équipes de recherche de Concordia et de l'UQAR avaient déjà effectué une analyse de l'espace de liberté de la rivière (ELR), ce qui nous a permis de valider les résultats en lien avec l'espace d'inondabilité et l'espace de mobilité de cette rivière. Les résultats de l'ELR ne permettent toutefois pas de déterminer les zones où l'incision est susceptible de se produire et, par conséquent, le modèle pour déterminer le type d'érosion n'a pu être validé. Il est important de prendre en compte les différences entre les deux approches lors de l'interprétation des résultats de cette analyse. L'analyse ELR cartographie les zones qui sont susceptibles d'être inondées ou érodées dans les 50 prochaines années (M50), tandis que notre modèle prédit où ces aléas fluviaux pourrait survenir le long de la rivière. Par conséquent, lorsque l'on classe les points en utilisant les résultats de l'ELR, les résultats sont des grands groupes homogènes de prédiction d'inondation, tandis que notre modèle crée des groupes plus hétérogènes aux points de référence du CRHQ. Par exemple, dans notre modèle, les zones plus sinueuses ont une plus grande chance d'initier un événement d'inondation et sont donc classées comme ayant une

forte probabilité d'inondation, tandis que les portions plus linéaires qui peuvent être affectées par cet événement d'inondation mais qui sont moins susceptibles de contribuer directement à l'inondation sont prédites comme ayant une faible probabilité d'inondation, ou pas d'inondation. En ce qui concerne les résultats de mobilité, les résultats de l'ELR cartographient la zone susceptible de s'éroder au cours des 50 prochaines années, ce qui signifie que les zones où l'incision se produit n'apparaissent pas comme mobiles car elles n'engendreront pas de migration latérale significative. Des problèmes similaires apparaissent dans les zones fortement peuplées où les structures anthropiques réduisent la capacité de migration latérale, ce qui fait que la zone semble stable. Notre modèle étant entraîné à détecter à la fois l'incision et les zones où l'érosion peut se produire, incluant dans les secteurs avec des structures anthropiques, il produira des résultats différents de ceux de l'ELR, car il pourra prédire que ces zones sont mobiles, alors que d'après le M50, elles semblent stables. Par conséquent, les résultats de précision doivent être considérés avec prudence, en particulier la présence de résultats de mobilité.

#### ***5.4.1 Description de la zone d'étude***

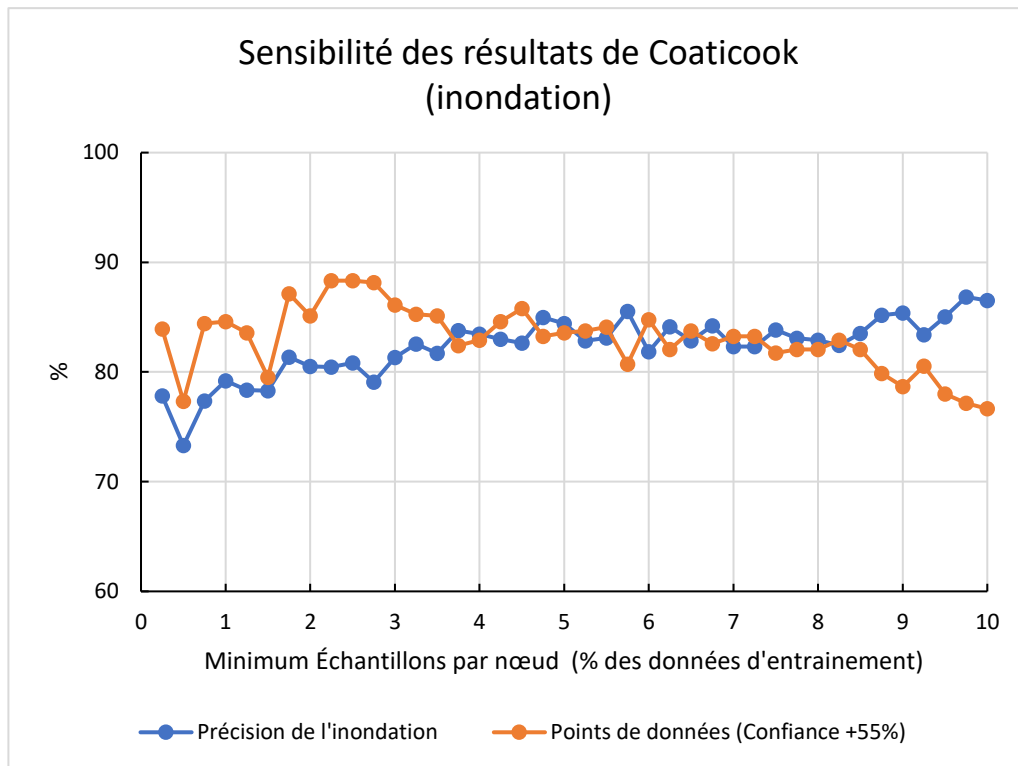
La zone d'étude est située dans les Appalaches et comprend le tronçon canadien de la rivière Coaticook (figure 33). Il s'agit d'une rivière partiellement confinée avec plusieurs grandes plaines inondables non confinées ayant des tronçons très sinueux qui sont sujets à la migration latérale et aux inondations, interrompues par des segments linéaires géologiquement confinés où l'aléa d'inondation est moindre et qui sont plus sujets à l'incision.



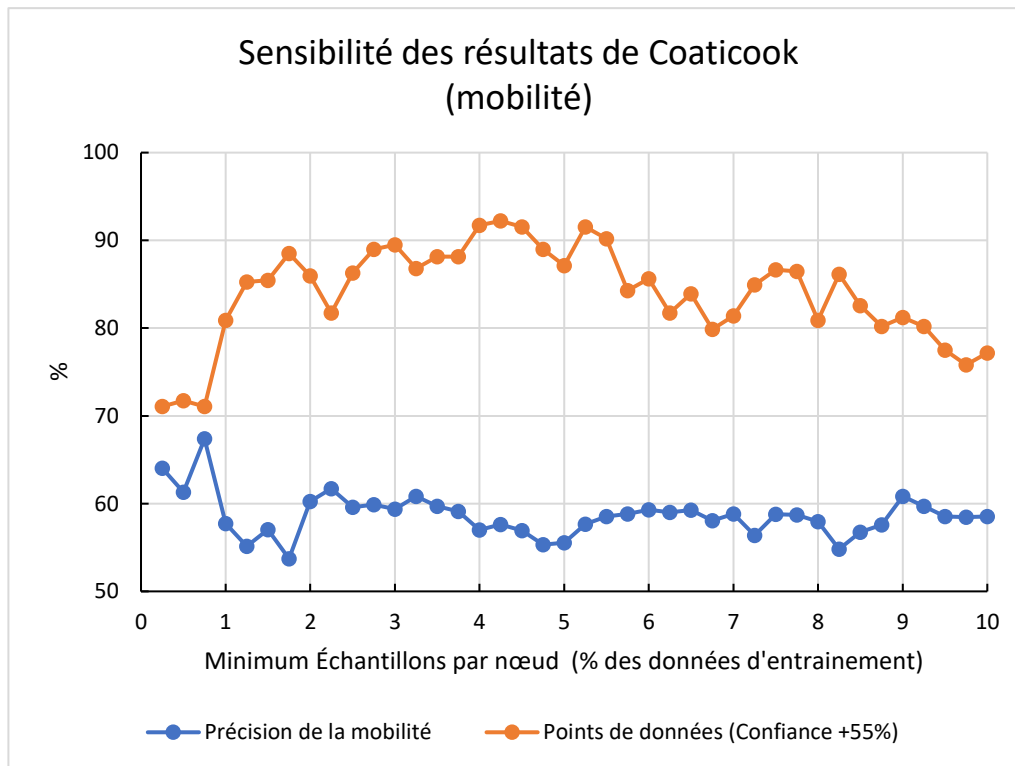
**Figure 33: Zones d'étude de la rivière Coaticook avec les points de référence du CRHQ et le MNT**

#### **5.4.2 Analyse de sensibilité**

La figure 34 indique que la précision des résultats d'inondation était bonne avec des niveaux de précision variant entre 73% et 86%, et ce malgré le fait que le modèle n'a pas été entraîné sur ce bassin versant et les différences attendues entre le modèle ELR utilisé pour la validation et notre modèle. L'incertitude était relativement faible, variant entre 12% et 23%. Les résultats de validation de la mobilité (figure 35) étaient moins bons avec des résultats de précision variant de 53% à 67%. Comme indiqué ci-haut, il est important d'interpréter ces résultats avec prudence en raison des différences dans la façon dont les données de validation (modèle ELR) et les données prédites (notre modèle) sont générées pour la mobilité. Lors de l'examen visuel des résultats, la majorité des points d'inadéquation se trouvaient dans des zones géologiquement ou anthropiquement confinées que le modèle ERL affichait comme stables (migration latérale minimale) alors que notre modèle les a prédites comme potentiellement problématiques pour les infrastructures ou des sites avec un potentiel d'incision. Le nombre de points avec des niveaux de confiance supérieurs à 55% varient davantage pour le modèle de mobilité avec une valeur minimale de 8% et maximale de 29%.



**Figure 34: Sensibilité des résultats sur le bassin de la Coaticook au nombre minimum de points de données par nœud pour la présence d'inondation montrant à la fois la précision et le pourcentage de points au-dessus du niveau de confiance de 55%**

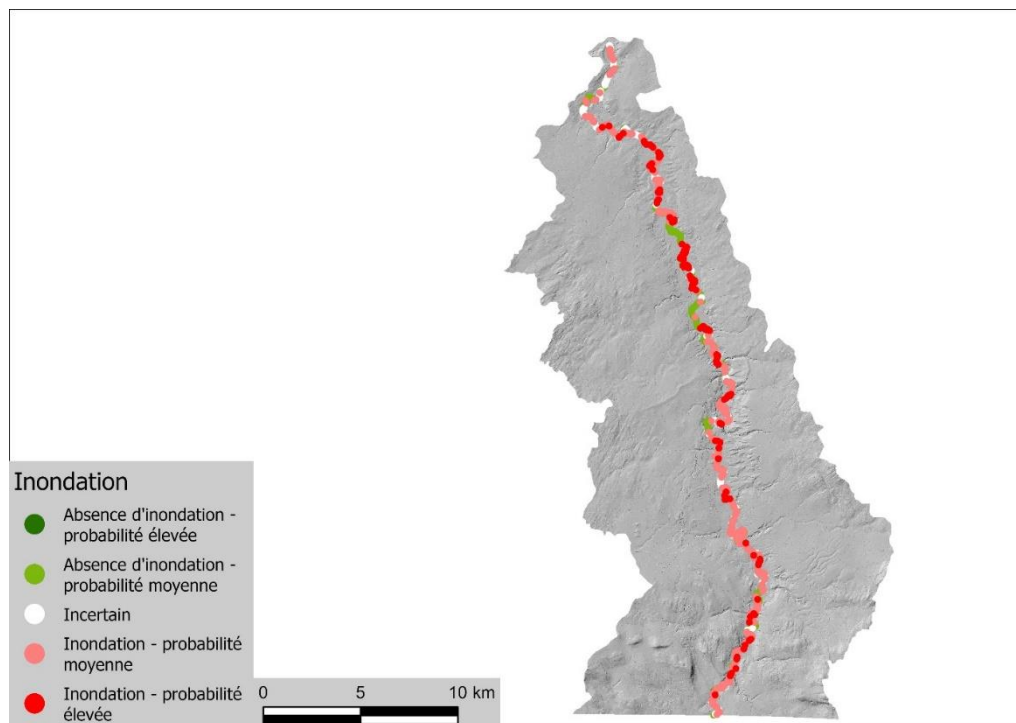


**Figure 35: Sensibilité des résultats sur le bassin de la Coaticook au nombre minimum de points de données par nœud pour la présence de mobilité montrant à la fois la précision et le pourcentage de points au-dessus du niveau de confiance de 55%**

#### **5.4.3 Résultats de probabilités d'aléas fluviaux pour la Coaticook**

Les figures 36 et 37 présentent les résultats pour l'inondation, qui est distribuée d'une manière bimodale autour de la classe incertaine, avec toutefois une plus grande fréquence de points indiquant la présence d'inondations. C'est un résultat attendu pour ce site d'étude en raison de ses vastes plaines inondables non confinées. Le potentiel d'aléa de mobilité est présenté dans les figures 38 et 39. Les résultats favorisent fortement la classe mobile. Compte tenu de la tendance du modèle de mobilité en faveur de la classe mobile, nous pouvons interpréter que les sites les plus susceptibles d'être mobiles seront ceux qui sont classés comme tels, tandis que les sites incertains ne connaîtront probablement pas d'érosion significative. Le champ de probabilité doit être pris en compte lors de l'interprétation des résultats et les sites prédits comme mobiles avec de faibles valeurs de probabilité auront probablement une érosion minimale. Ces résultats montrent que les zones les plus aléatoires pour la mobilité dans ce bassin versant se trouvent dans les grandes plaines inondables avec des

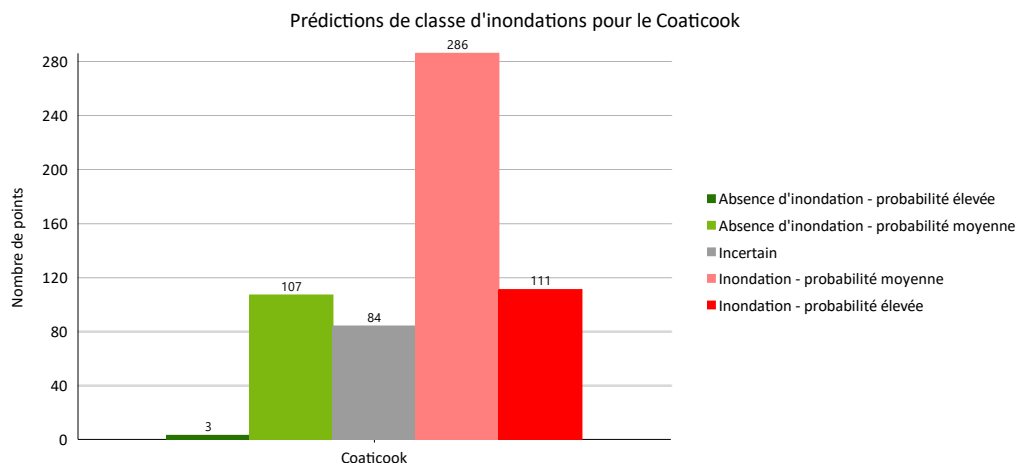
rivières sinueuses. Les résultats pour le type d'érosion sont présentés dans les figures 40 et 41. Ces résultats montrent que ce bassin versant est dominé par la migration latérale avec quelques zones potentiellement stables ou en incision, ces zones étant souvent là où des contrôles géologiques ou anthropogéniques existent. En combinant ces résultats avec ceux de la présence de mobilité, nous sommes en mesure de déterminer que les zones à plus haut potentiel d'aléa de mobilité dans ce bassin versant sont en effet les grandes plaines inondables et que le type d'érosion dominant sera la migration latérale dans ces zones.



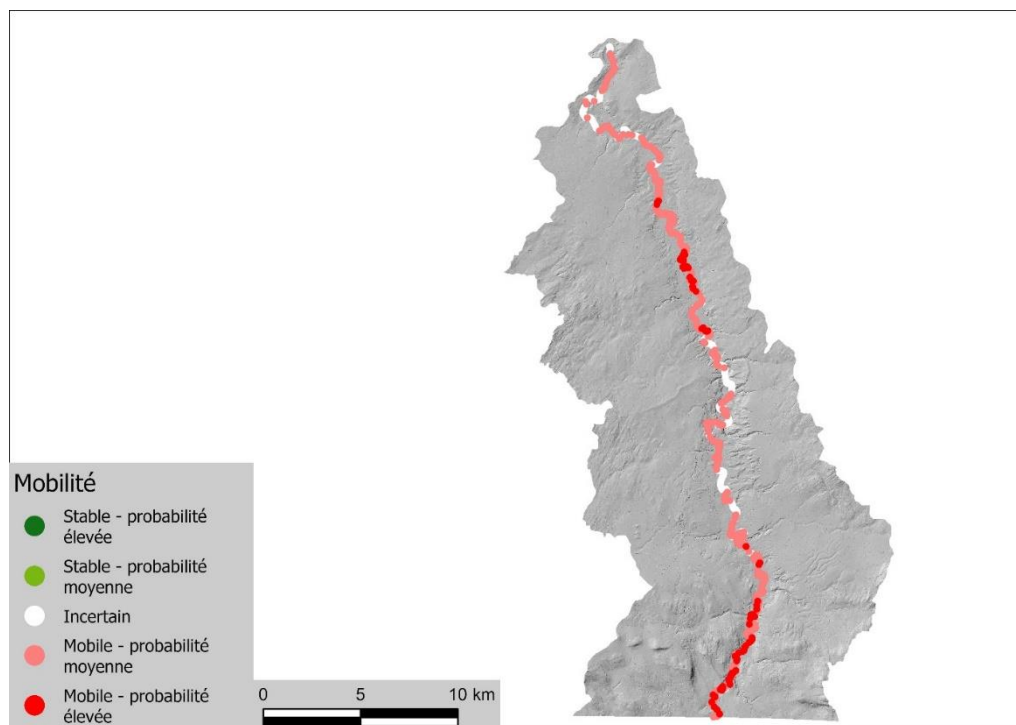
**Figure 36: Carte des prévisions de présence d'inondations du modèle pour le bassin versant de la Coaticook**



**DÉVELOPPEMENT D'UN OUTIL POUR L'INTÉGRATION DE L'ANALYSE DES ALÉAS  
ASSOCIÉS À LA DYNAMIQUE GÉOMORPHOLOGIQUE DES COURS D'EAU – PROJET DE  
RECHERCHE (R800.1)**



**Figure 37: Prédictions de classe d'inondation pour la Coaticook**



**Figure 38: Carte de prévisions de présence de mobilité du modèle pour le bassin versant de la Coaticook**



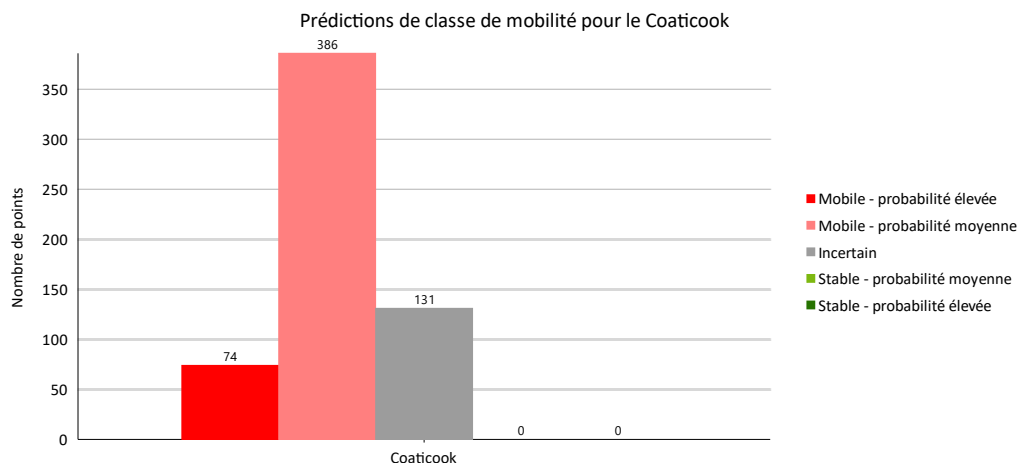


Figure 39: Prédictions de classe de mobilité pour le Coaticook

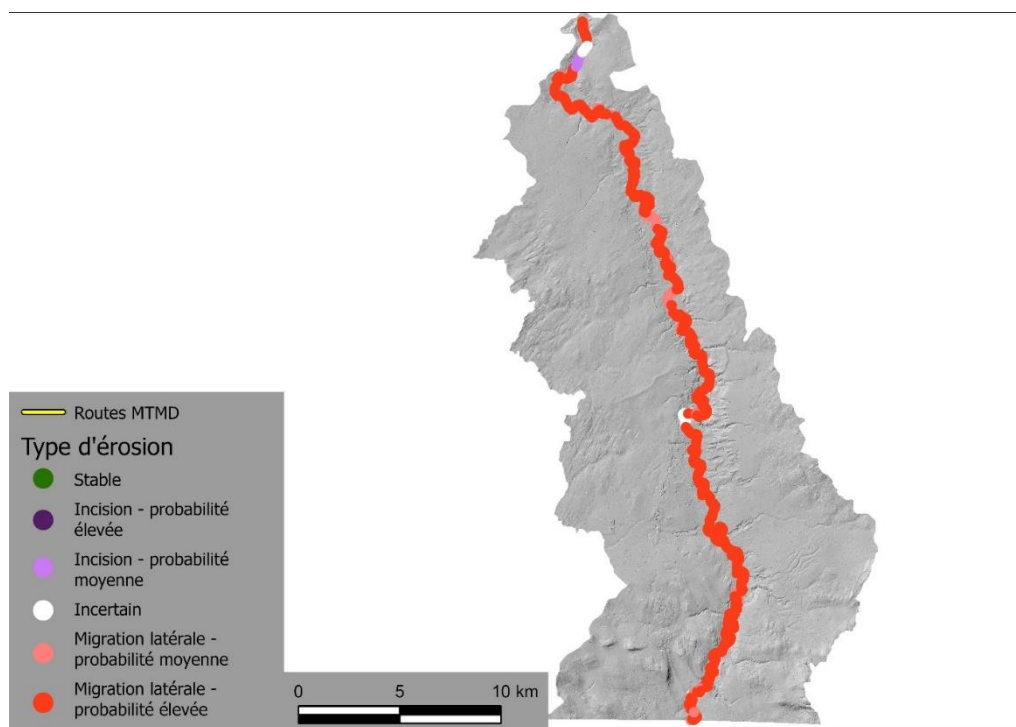


Figure 40: Carte de prévisions du type d'érosion du modèle pour le bassin versant de la Coaticook

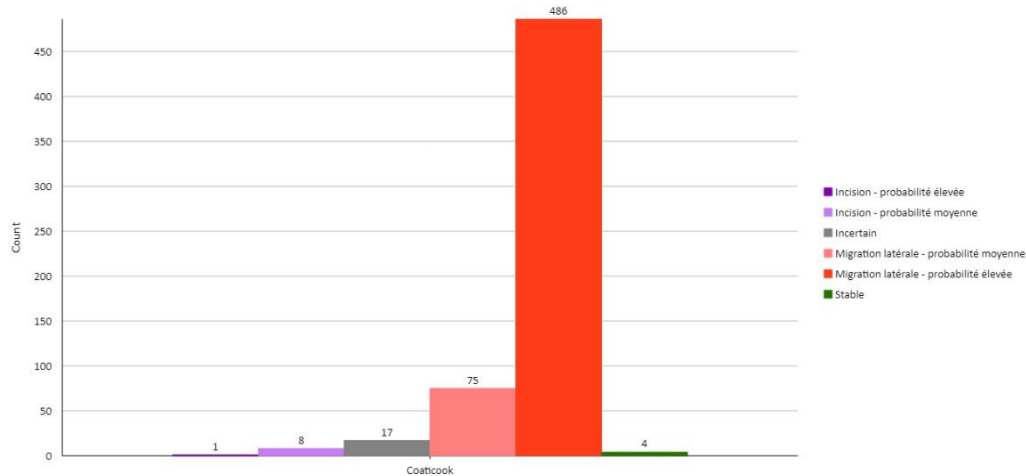


Figure 41: Prédictions de classe de type d'érosion pour le Coaticook

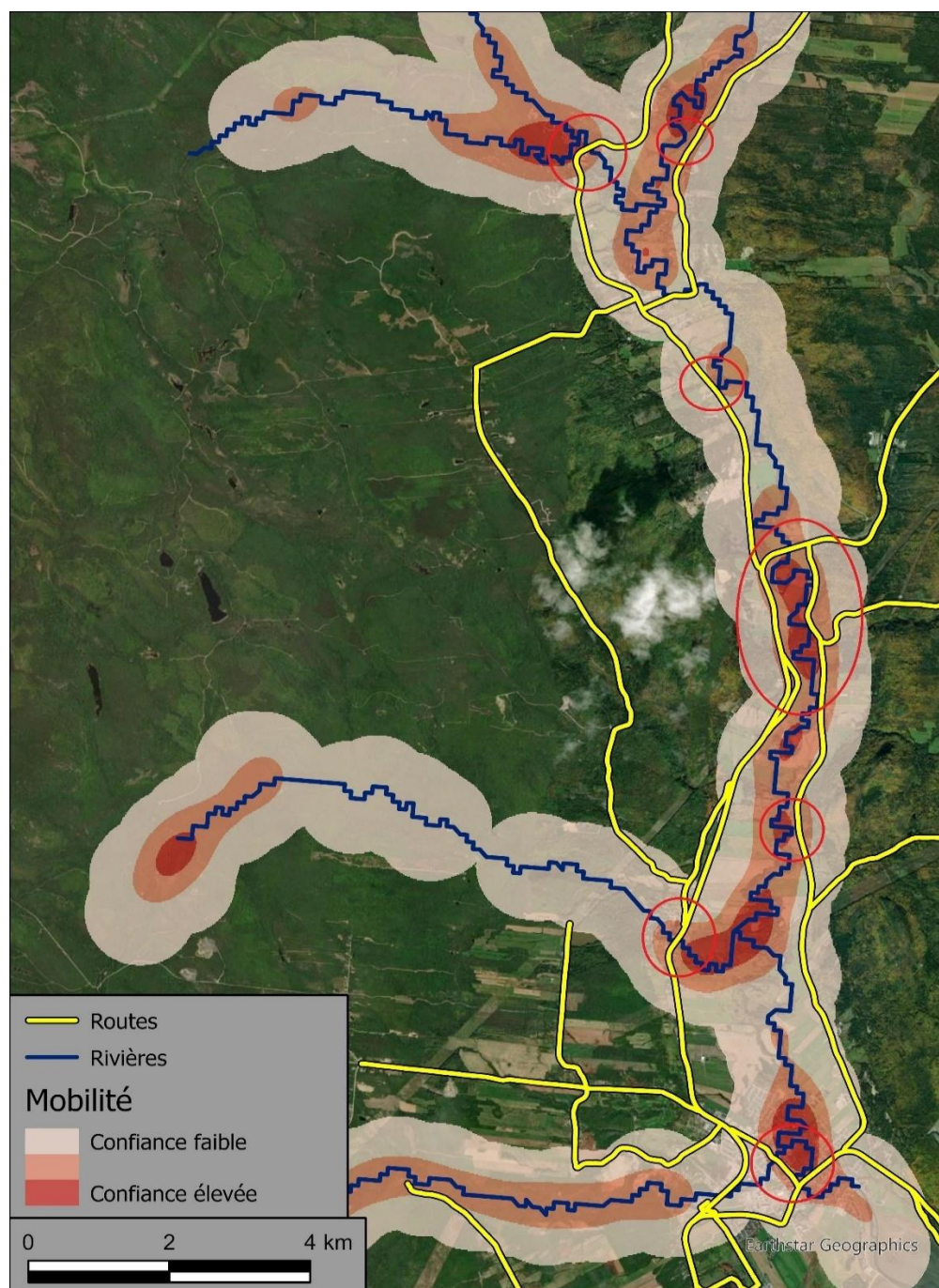
## 6. MISE EN ŒUVRE ET RETOMBÉES, BÉNÉFICES POUR LE MTMD

Les résultats du modèle peuvent être utilisés pour fournir des informations sur les aléas fluviaux potentiels à trois échelles d'analyse. À l'échelle du bassin versant, d'abord, l'outil d'évaluation préliminaire peut identifier les sections potentiellement touchées par les aléas (figure 42). À l'échelle du tronçon, ensuite, il peut être utilisé pour identifier les processus fluviaux dominants de ce tronçon et la manière dont ils peuvent interagir avec l'infrastructure routière. Finalement, l'outil peut aussi servir à l'échelle locale en fournissant des informations sur le potentiel d'un aléa fluvial se produisant dans des secteurs ponctuels (tous les 100 mètres) de la rivière (annexe A).

Cette méthodologie d'identification à grande échelle des rivières présentant des niveaux élevés d'aléa permet au MTMD de déterminer les segments de route qui devraient faire l'objet d'une évaluation rigoureuse de la mobilité des cours d'eau proximaux afin de mieux se préparer à des dommages futurs. Cela permet au MTMD d'être proactif dans la protection de ses infrastructures routières, plutôt que d'attendre que les dommages surviennent lors de crues, afin d'identifier les zones de conflit potentiel entre la rivière et la route. La cible de cette analyse étant les rivières, les résultats peuvent également être utilisés pour les projets d'infrastructure qui sont encore au stade de la planification, ce qui permet au MTMD d'éviter les rivières potentiellement problématiques lors de la conceptualisation de nouvelles routes et de nouveaux ponts ou ponceaux. Dans ces perspectives, les résultats de l'application du modèle pourraient être intégrés dans les avis de mobilité en complément et en support de l'avis fournis.

Dans une autre perspective, les résultats fournissent également des informations cruciales sur la charge sédimentaire potentielle et relative en amont

des infrastructures routières en prédisant le potentiel d'incision ou de migration latérale en amont. Le modèle étant conçu de manière que la majorité du travail puisse être effectué par une petite équipe, cela limite le nombre d'experts nécessaires pour utiliser le modèle afin de générer des résultats. De plus, les résultats du modèle sont conçus de manière à permettre aux personnes sans expertise en géomorphologie ou en informatique de les comprendre et de les utiliser efficacement, sans nécessairement comprendre le modèle lui-même. Cela permet donc à une grande variété d'experts d'intégrer les aléas géomorphologiques dans leurs projets sans avoir besoin d'une formation en géomorphologie.



**Figure 42: Carte thermique des points prédits comme étant mobiles, pondérés par la confiance de la prédiction, avec les conflits potentiels entre les rivières mobiles et les routes encerclés en rouge.**

## 7. CONCLUSION

Dans l'ensemble, le modèle de prévision des aléas fluviaux fonctionne bien, avec des valeurs de précision variant de 84,3% à 94,7% et de faibles niveaux de surentraînement (1 à 3%) (tableau 14). Certaines limites dans l'approche de modélisation ont été notées, comme le fait que le modèle a été entraîné principalement sur les plus grandes rivières en aval des bassins versants puisque ce sont les rivières pour lesquelles l'Atlas Hydroclimatique fournit des valeurs de débit. Cela crée un biais dans le modèle vers des aléas fluviaux plus susceptibles de se produire dans ces types de rivières par rapport aux cours d'eau de tête. Il existe également des problèmes de qualité des données, telles que la largeur de la rivière qui dépend entièrement de l'aire de drainage, bien que l'équipe du CRHQ soit en train de résoudre ce problème. Cela pose le problème de ne pas prendre en compte les fluctuations réelles de la largeur de la rivière qui existent et qui auraient un impact sur des variables telles que le ratio de confinement et la puissance spécifique. Cela pose également des problèmes lors du calcul des ratios de confinement, car la largeur est parfois surestimée, ce qui se traduit par des ratios de confinement artificiellement bas. Un autre problème est que parfois les tronçons homogènes du CRHQ (UEA) ne sont pas générés correctement (très longs ou extrêmement courts), ce qui est prévalent dans le chenal principal de la rivière Cascapédia. Cela pose des problèmes lors du calcul de la sinuosité du chenal ainsi que du ratio de confinement du tronçon et du ratio de confinement relatif.

Le champ "Class" ne doit être utilisé que pour identifier les points chauds, car le seuil de ces classes (incertain, moyen, élevé) a été fixé de manière relativement arbitraire sur la base de la moyenne et de l'écart-type des probabilités des modèles. De nouveaux seuils pourraient aisément être générés en utilisant les champs "pred" et "prob". Il s'agit également d'indicateurs plus précis du potentiel d'occurrence des aléas. Toutefois, les points présentant de faibles valeurs de "prob" (plus proches de 50 %) doivent être considérés avec précaution car le modèle était moins fiable dans ces sections et la probabilité peut avoir été influencée par le biais du modèle. Ces points doivent être considérés par rapport aux autres points d'un tronçon homogène. Par exemple, si les points de référence dans un tronçon sont prédits comme ayant l'absence d'inondation avec une grande confiance mais qu'un point près du milieu du tronçon a une très faible confiance d'absence d'inondation, alors il y a un potentiel qu'une inondation se produise à cet endroit. Essentiellement, le modèle a une faible valeur de confiance parce que les variables à grande échelle (tronçon et bassin) sont en conflit avec les variables locales. Par conséquent, le point peut être situé à un endroit du bassin versant où l'on ne s'attend généralement pas à des inondations, mais les variables locales indiquent qu'il y a un risque d'inondation. Cela peut se produire dans des environnements tels que des tronçons très confinés où la vallée s'élargit en un point. Par conséquent, les variables à l'échelle du tronçon indiqueront un faible risque d'inondation, tandis que la variable locale indiquera un potentiel d'inondation. La figure A21 de l'annexe A en donne un exemple. Le champ "prob" peut également être utilisé comme un

pseudo proxy pour le niveau d'intensité de l'aléa, des valeurs plus élevées indiquant non seulement un potentiel plus élevé d'aléa mais aussi des inondations ou une érosion possiblement plus sévère. Il est important de considérer que les résultats sont relatifs à la taille de la rivière et que, par conséquent, les petites rivières peuvent avoir la même probabilité de mobilité ou d'inondation que les grandes rivières, mais l'impact réel peut être plus faible. Il faut également considérer que le modèle de présence d'inondation est entraîné pour détecter les zones où le débit a une chance d'excéder le niveau plein bord, et non pas qu'il causera l'inondation d'une grande zone. Par ailleurs, les résultats de mobilité ne prédisent que la migration latérale et l'incision, mais il est important de considérer que la migration d'un cours d'eau peut également se produire en raison de **coupure de méandre par dissection** (« chute cutoff ») par le développement de chenaux secondaires dans des zones inondables entraînant des **avulsions** (Dépret et al., 2017). Pour détecter ces événements, il serait utile dans le futur de considérer conjointement les résultats du modèle d'inondation et de mobilité.

Les modèles ont également du mal à détecter les zones de forte hétérogénéité des processus parce qu'il utilise également des variables non locales pour faire la prédiction, ce qui peut influencer les résultats locaux. Par conséquent, s'il existe un point où une inondation est susceptible de se produire dans un tronçon homogène qui a une très faible probabilité d'inondation pour le reste des points de référence, le modèle aura tendance à prédire une absence d'inondation, bien qu'avec un niveau de confiance assez faible. Lors de l'interprétation des résultats pour des zones spécifiques, il est donc important de tenir compte de la probabilité des autres points environnants, les baisses ou augmentations soudaines de la probabilité indiquant probablement un changement majeur de la probabilité d'aléa. Des exemples de ces cas et de ces limites sont présentés à l'annexe A et les résultats sur les petits cours d'eau sont présentés à l'annexe C.

**Tableau 15 : Résultats finaux et conclusions pour chaque modèle**

Modèle	précision totale (%)	Niveaux de sur- entraîne- ment (%)	Gamme de pré- cision pour les bassins versant (%)	Gamme d'incer- titude pour les bassins versant (%)	Conclusion pour le modèle
Présence d'inonda- tion	84,3	3	80 - 90	15 - 19	<ul style="list-style-type: none"> <li>• Fonctionne bien</li> <li>• Pas de biais majeur pour une classe de prédiction</li> <li>• Les résultats sont relatifs à la taille de la rivière/vallée</li> </ul>
Présence de mobilité	84,7	3	81 - 87	9 - 29	<ul style="list-style-type: none"> <li>• Efficace pour déterminer les zones d'aléa élevé</li> <li>• Biais en faveur de la classe mobile</li> <li>• Les probabilités de mobilité faibles doivent être considérées comme plutôt stables en raison du biais du modèle</li> </ul>
Type d'érosion	94,7	1	83 - 97	8 - 14	<ul style="list-style-type: none"> <li>• Les résultats ne reflètent pas la probabilité de mobilité, mais bien la probabilité du type d'aléa de mobilité pour les tronçons considérés comme dynamiques</li> <li>• En raison de la prédominance de l'érosion latérale dans les bassins à l'étude : <ul style="list-style-type: none"> <li>○ Précision artificiellement élevée</li> <li>○ Biais en faveur de l'érosion latérale</li> </ul> </li> <li>• Les points avec des niveaux de confiance plus faibles doivent être considérés avec précaution</li> </ul>

Les tests effectués sur un bassin-versant non-entraîné (rivière Coaticook) sont encourageants, mais il serait évidemment très utile de poursuivre cette validation pour un grand nombre de secteurs, dans des contextes géomorphologiques variables. Idéalement, des données exhaustives de mobilité (migration latérale et incision) et d'inondabilité devraient être disponibles dans ces bassins versants, ce qui demeure une contrainte à l'heure actuelle, surtout pour la mobilité encore à ce jour peu documentée au Québec.

Les outils ont également des temps de traitement très faibles, avec un temps d'exécution total d'environ 20 minutes par bassin versant, ce qui, combiné à la rapidité avec laquelle les données d'entraînement peuvent être modifiées et collectées, permet aux modèles d'être rapidement améliorés et testés. Plus la qualité et la quantité des points d'entraînement fournis au modèle sont élevées, plus le modèle sera capable de différencier les processus fluviaux. Une méthode plus standardisée de catalogage des dommages causés aux infrastructures par les eaux fluviales, telle que celle menée par le MTMD dans le bassin versant de la rivière Cascapédia, améliorerait grandement la capacité d'entraîner et de valider ce modèle, et permettrait de plus de valider d'autres modèles. Il est donc recommandé de maintenir une base de données contenant ces informations car cela permettra au modèle d'être mieux entraîné à détecter non seulement les aires mobiles mais plus spécifiquement les niveaux de mobilité qui peuvent avoir un impact sur les infrastructures routières.



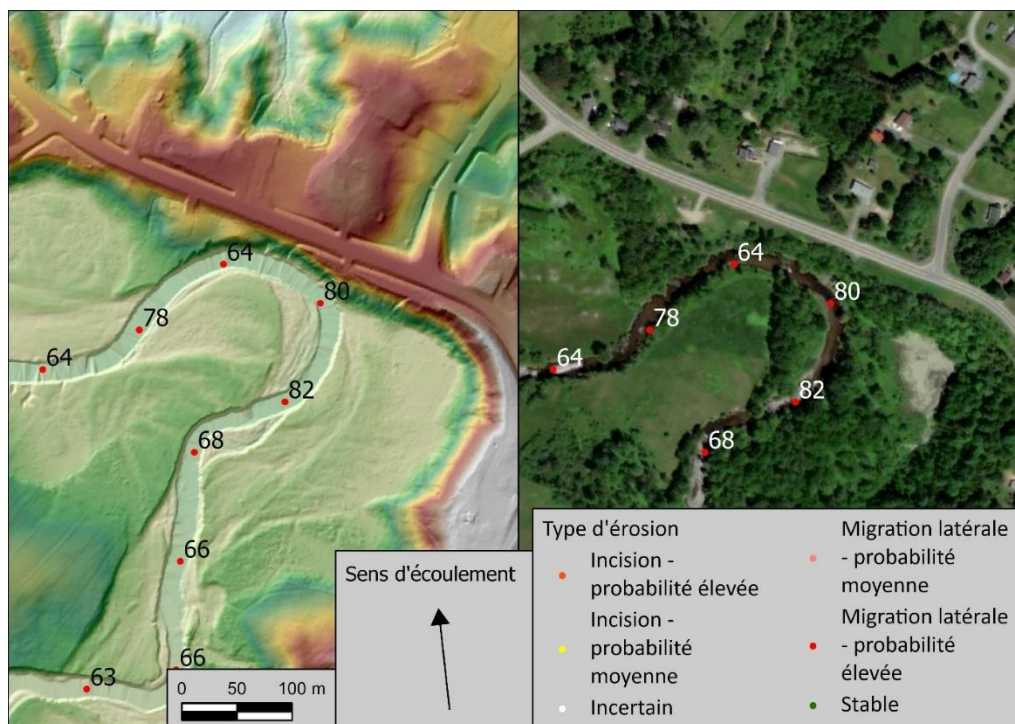
## **ANNEXE A**

### **RÉSULTATS PONCTUELS DE PROBABILITÉS D'ALÉAS FLUVIAUX DANS LES 3 BASSINS PILOTES**

---

## **Annexe A.1 Type d'érosion et mobilité**

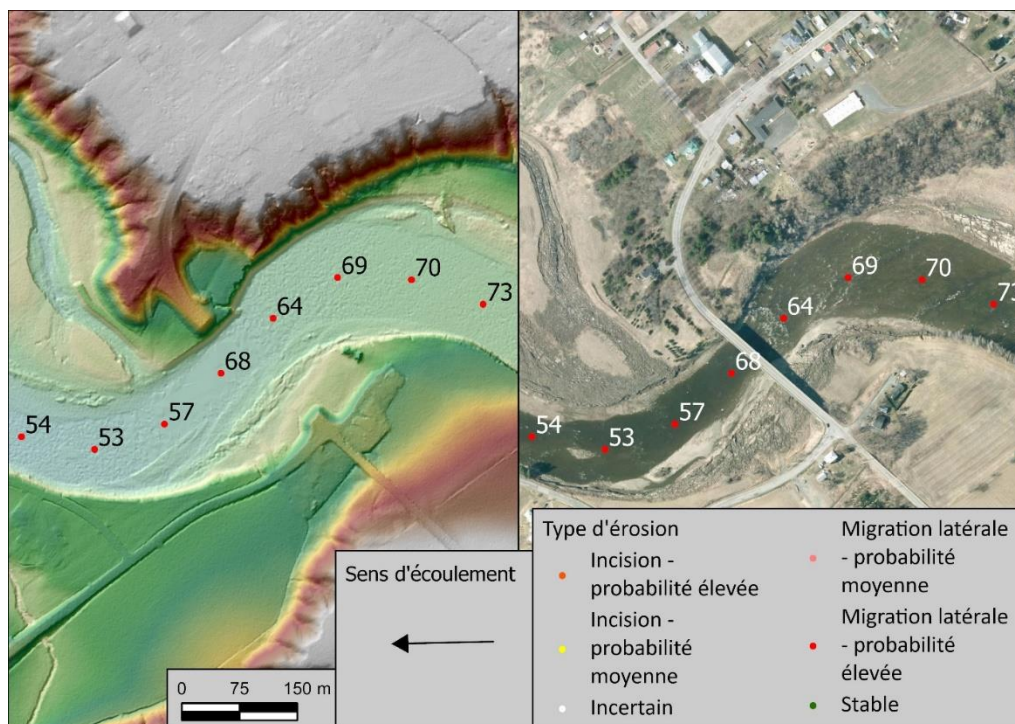
Les figures A1 à A18 montrent les prédictions du modèle pour le type d'érosion avec la probabilité de mobilité/stabilité comme étiquettes, ainsi que des photos de ces sites. Dans l'ensemble, ces sections ont donné de bons résultats. Les figures A9, A15 et A17 présentent toutefois une certaine ambiguïté. La figure A9 montre une section qui est clairement en érosion dans la Cascapédia, bien que le modèle n'ait pas été capable de prédire cet aléa avec certitude, probablement en raison des problèmes de qualité des données dans le chenal principal de la Cascapédia (surestimation de la largeur du chenal et tronçons homogènes trop petits). Cela interfère à la fois avec l'entraînement et la prédiction du modèle, ce qui se traduit par une faible confiance pour le chenal principal de la Cascapédia. La figure A15 montre une section dans le bassin de la Du Gouffre où des niveaux élevés de stabilisation et de confinement anthropogénique font que le modèle prédit la zone comme stable, bien que les photos de la figure A16 montrent qu'il y a encore une érosion mineure sur les berges non protégées et que la stabilisation des berges a dû être réparée plusieurs fois. La figure A17 montre une aire où le modèle prédit la mobilité avec un niveau de confiance élevé, bien que le type d'érosion qui se produit à cet endroit demeure incertain. Les figures A7 et A20 montrent également les problèmes de qualité des données où le point final de l'affluent prédit une incision plutôt qu'une migration latérale et une absence d'inondation (alors que l'inondation est probable). Cela est dû au fait que ce point se voit assigné la largeur du chenal principal bien qu'il soit censé faire partie de l'affluent.



**Figure A1: Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Nicolet : 72.0290349°W 45.7900955°N)**



**Figure A2: Deux photos prises à l'apex du méandre sur un site où le point à 80% de chances d'être mobile dans la figure A1**

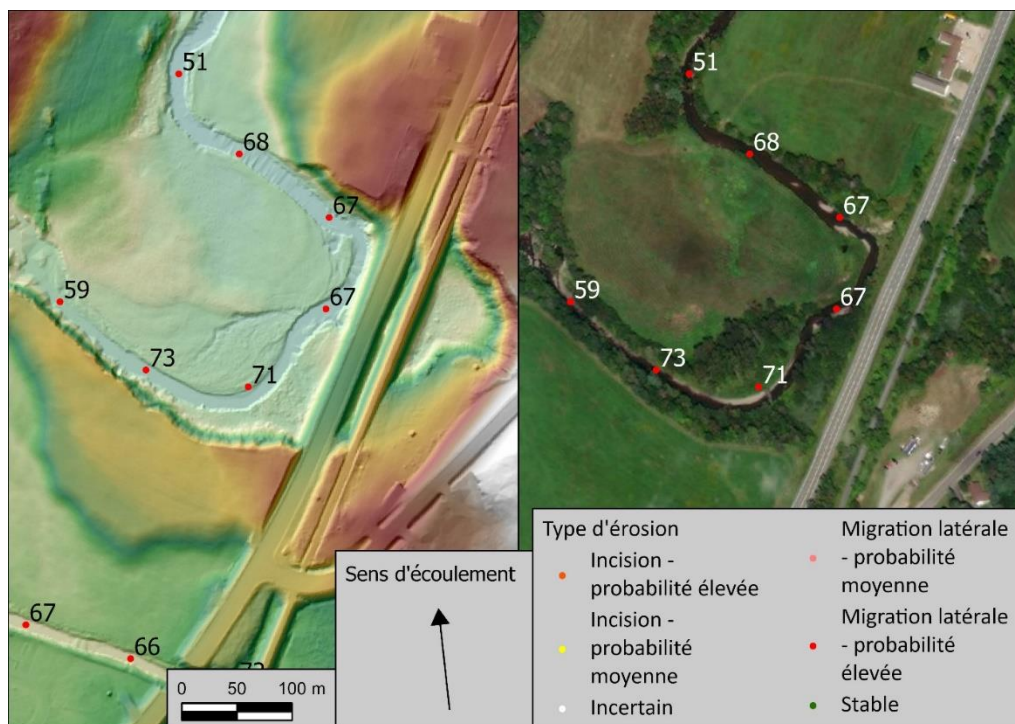


**Figure A3: Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Nicolet : 72.5375146°W 46.1540019°N)**



**Figure A4: Deux photos du site prises à partir de Google Earth et sur le site à partir du pont dans la figure A3**



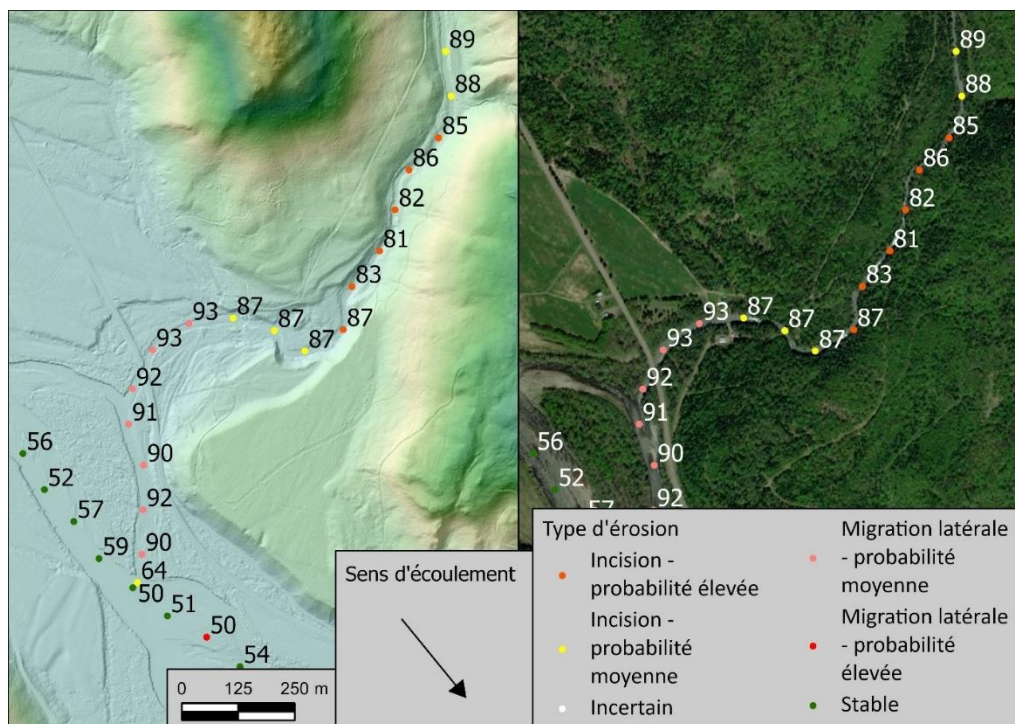


**Figure A5: Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Nicolet : 72.0286840°W 45.7667794°N)**

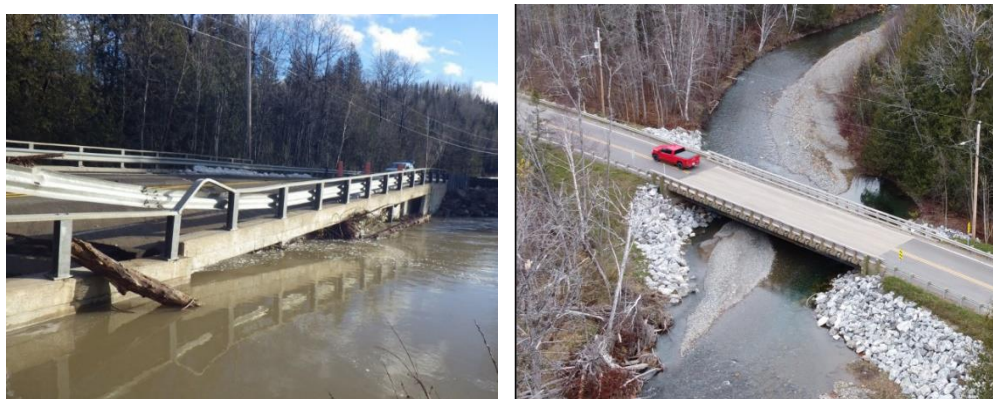


**Figure A6: Deux photos du site prises sur le site depuis la route dans la figure A5**

DÉVELOPPEMENT D'UN OUTIL POUR L'INTÉGRATION DE L'ANALYSE DES ALÉAS ASSOCIÉS À LA DYNAMIQUE GÉOMORPHOLOGIQUE DES COURS D'EAU – PROJET DE RECHERCHE (R800.1)

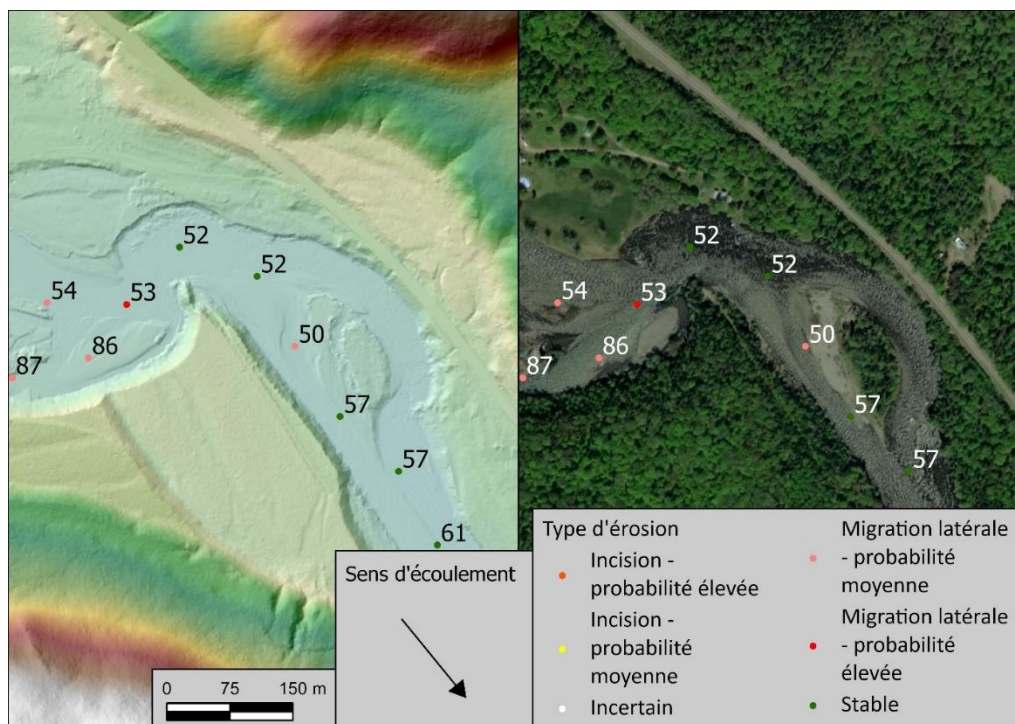


**Figure A7: Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Cascapédia : 65.9309543°W 48.2969596°N)**



**Figure A8: Deux photos du site de la figure A7. Première photo par le MTMD, 2018 la seconde par l'UQAR 2022**

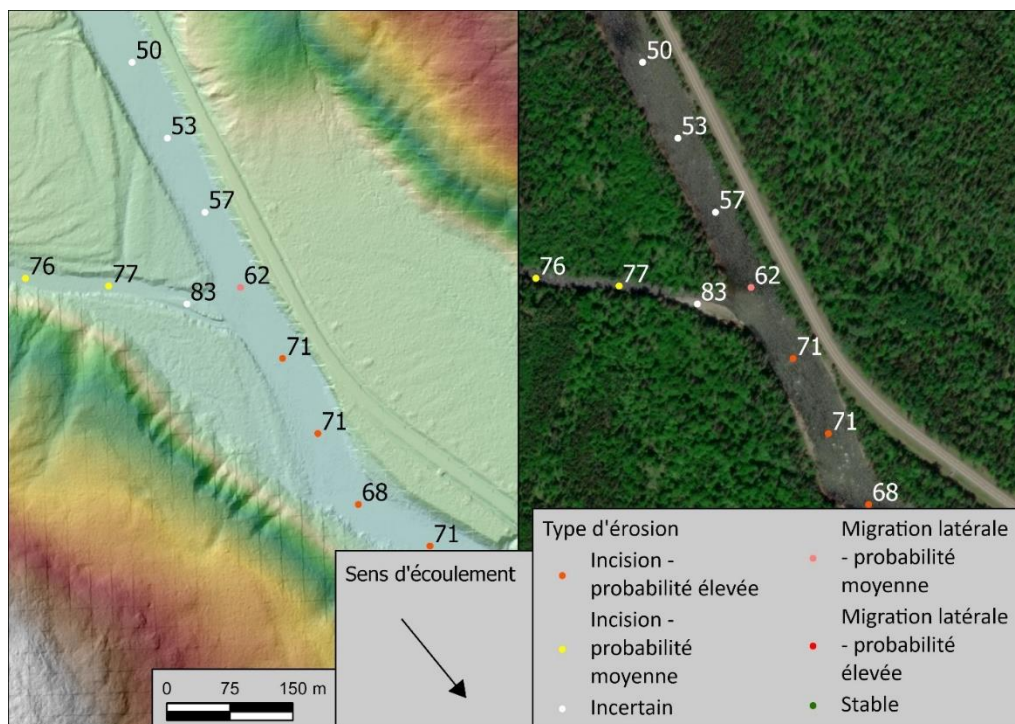




**Figure A9: Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Caspédia : 65.9576840°W 48.3110689°N)**



**Figure A10: Deux photos du site de la figure A9. Première photo par le MTMD, 2018 la seconde par l'UQAR 2022**

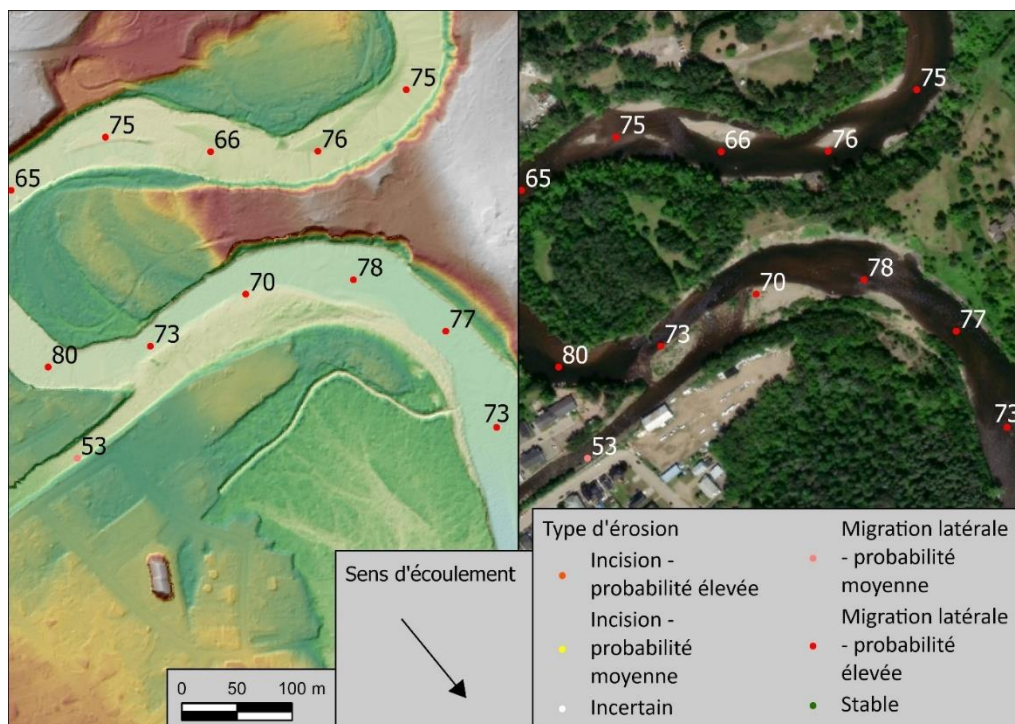


**Figure A11: Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Cascapédia : 66.1466730°W 48.5648444°N)**



**Figure A12: Deux photos du site de la figure A11 par l'UQAR 2022**

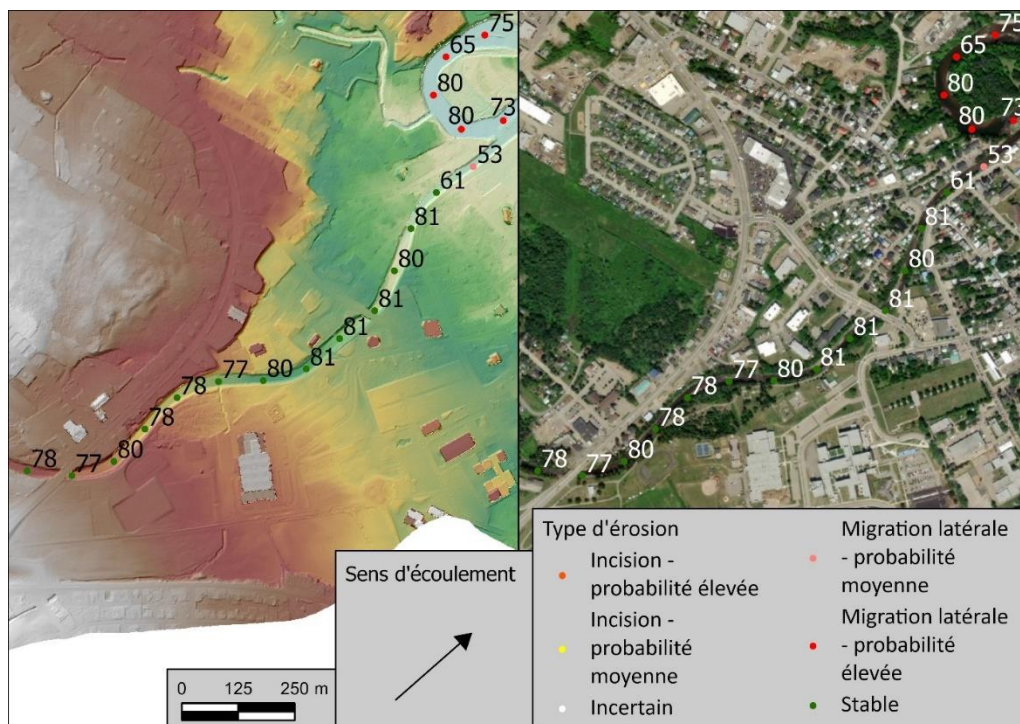




**Figure A13: Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Du Gouffre : 70.5057976°W 47.4476068°N)**



**Figure A14: Une photo du site de la figure A13**

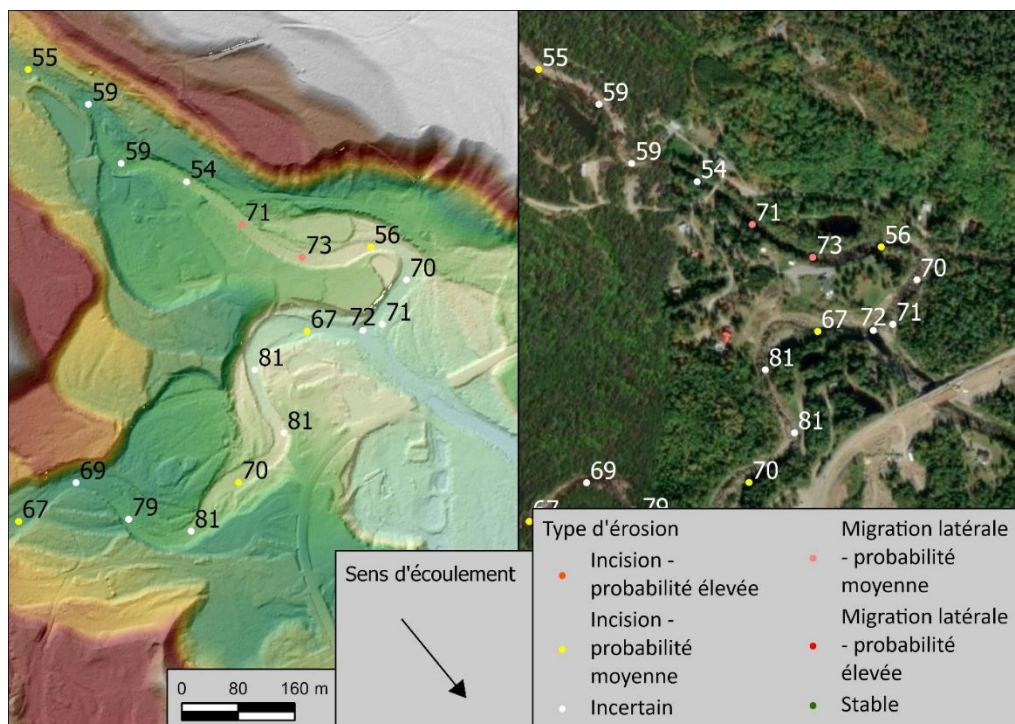


**Figure A15: Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Du Gouffre : 70.5164161°W 47.4411945°N)**



**Figure A16: Deux photos du site de la figure A15**





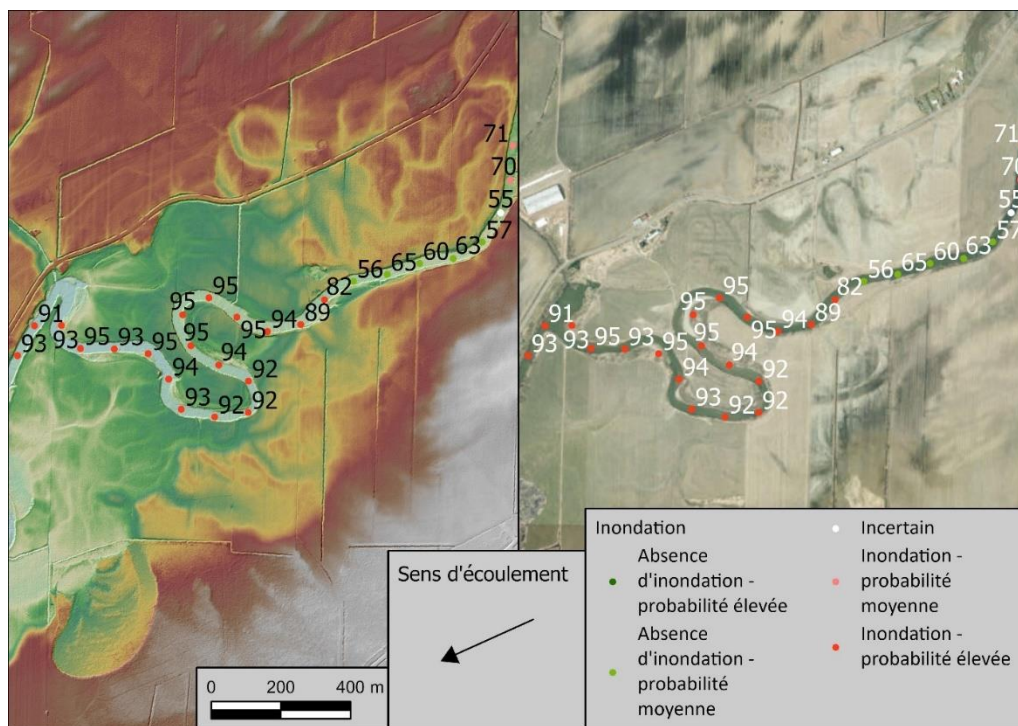
**Figure A17: Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Du Gouffre : 70.5522495°W 47.5793112°N)**



**Figure A18: Deux photos du site de la figure A17**

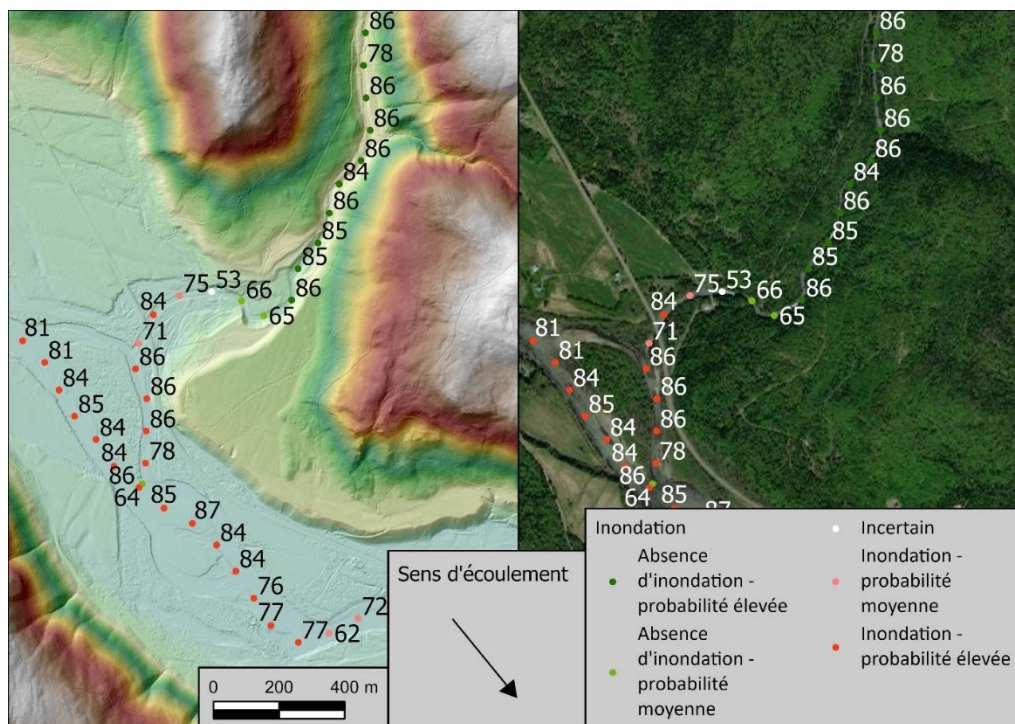
## Annexe A.2 Inondation

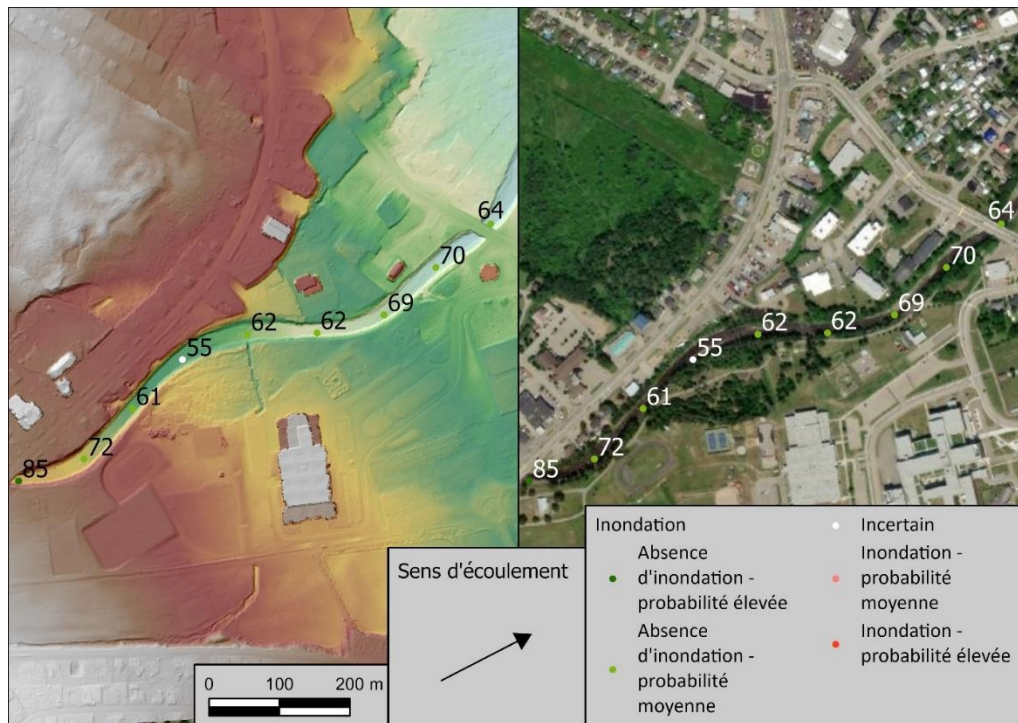
Les figures A19 à A22 montrent quelques exemples du modèle de présence d'inondation. Les figures A19 et A20 montrent des aires de transition dans les rivières Cascapédia et Nicolet pour illustrer la différence entre les secteurs qui sont prédits comme étant à haut risque d'inondation par rapport à ceux à faible risque d'inondation. La figure A19 montre également une zone présentant un aléa élevé d'avulsions liées aux inondations. La figure A21 est située dans le bassin versant de la rivière Du Gouffre, dans une section confinée par des activités humaines. Cette section présente une faible probabilité d'inondation, bien qu'à proximité du point où la valeur de prédiction est de 55, le potentiel est plus élevé et il y a des signes d'inondations antérieures dans cette section. Il s'agit d'un exemple où une baisse de la probabilité peut indiquer un potentiel d'aléa dans une petite section d'un tronçon à faible probabilité d'aléa. La figure A22 est située dans le chenal principal de la rivière Du Gouffre et montre une zone où un méandre est marqué comme présentant un risque d'inondation et où, en raison de l'inondation dans ce méandre, un événement de coupure de méandre par dissection s'est produit.



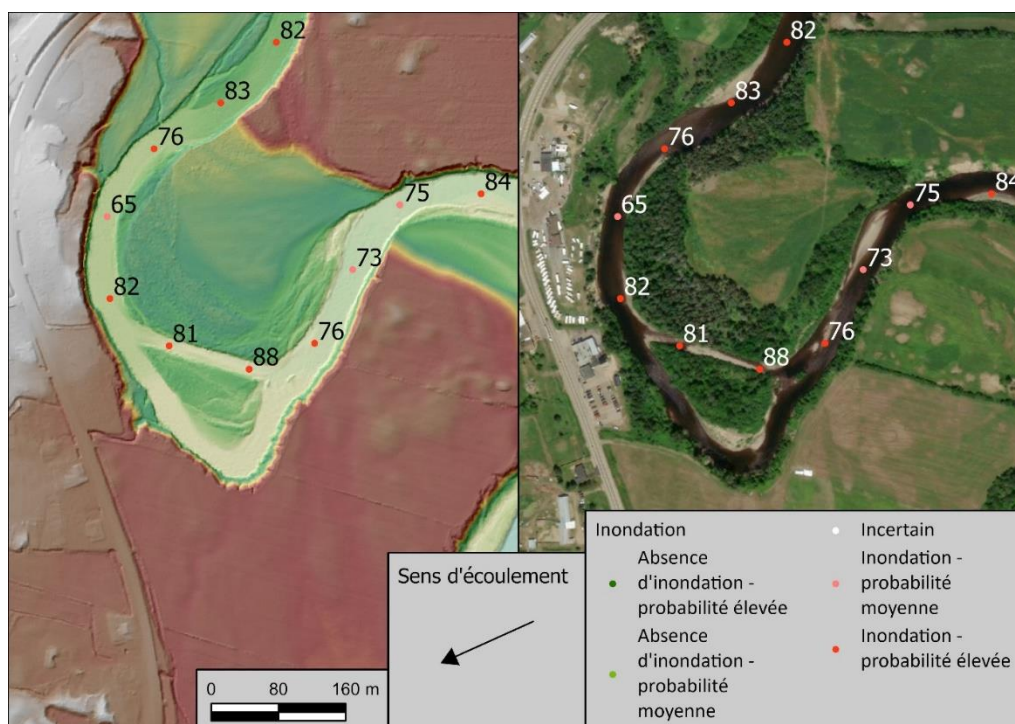
**Figure A19: Carte montrant les prévisions d'inondation du modèle. Les chiffres représentent la probabilité de la classe (Absence d'inondation ou Inondation). (Nicolet : 71.8664060°W 46.1452991°N)**







**Figure A21: Carte montrant les prévisions d'inondation du modèle. Les chiffres représentent la probabilité de la classe (Absence d'inondation ou Inondation). (Du Gouffre : 70.5166415°W 47.4410165°N)**



**Figure A22: Carte montrant les prévisions d'inondation du modèle. Les chiffres représentent la probabilité de la classe (Absence d'inondation ou Inondation). (Du Gouffre : 70.5123875°W 47.5211689°N)**

## **ANNEXE B**

### **EXÉCUTION ET MODIFICATION DES MODÈLES**

---



## Annexe B.1 Étapes d'exécution

Dépendances :

- Python 3.9.13
- pandas 1.35
- geopandas 0.9.0
- matplotlib 3.4.3
- arcpy 3.0
- scikit-learn 1.1.3
- numpy 1.20.1
- Licence d'analyste spatial ESRI

\*Avant d'exécuter ces scripts, vous devrez vous assurer qu'ArcGIS Pro utilise l'environnement Python fourni ou que toutes les dépendances sont installées

Étape 1 : téléchargez les données d'entrée à partir des sources du tableau 1.

Étape 2 : exécutez l'outil 1 : Découpage des données (figure B1).

Étape 3 : Vérifiez manuellement que l'outil de découpage des données a fonctionné correctement en vérifiant que le réseau fluvial est bien limité aux rivières de plus de 50 km<sup>2</sup>.

Étape 4 : exécutez l'outil 2 : Extraction du débit (figure B1).

Étape 5 : l'outil d'extraction de débit supprime les segments de rivière qui n'intersectent pas le réseau de l'Atlas Hydroclimatique. Assurez-vous qu'il l'a fait correctement et qu'il n'a pas supprimé de petits segments de rivière qui peuvent prévaloir dans les zones où le CRHQ est toujours en cours (le bassin versant de Cascapédia en avait plusieurs). Si ceux-ci ont été supprimés, vous devrez les copier et coller spécialement et ajouter une valeur de débit unitaire appropriée (copier-coller une valeur du segment en amont ou en aval). À ce stade, vous pouvez également supprimer les petits affluents qui ont été inclus en raison d'une intersection avec le réseau de l'Atlas Hydroclimatique. Ces erreurs résultent des différentes échelles de MNT à partir desquelles ces deux réseaux ont été générés.

Étape 6 : lissez le MNT LiDAR de résolution 1 m disponible sur « Forêt ouverte » à l'aide d'un filtre passe-bas avant de l'agréger à 10 m en utilisant la valeur minimale. Cela garantit que les terrains irréguliers tels que les zones de dépôt sur les méandres n'interfèrent pas avec le processus de création de polygones de vallée en augmentant artificiellement les valeurs de pente cumulées, ce qui entraînerait des polygones de vallée qui n'incluent pas toute la vallée ou les bancs d'accumulation dans le chenal.

Étape 7 : exécutez l'outil 3 : Extraction de vallée (figure B1) à l'aide du MNT de 10 m.

Étape 8 : exécutez l'outil 4 : Calculateur de confinement (figure B1).

Étape 9 : exécutez l'outil 5 : Préparation des données (figure B1).

Étape 10 : Confirmez que la feuille de calcul Excel que ce script exporte ne contient pas de valeurs nulles ou de valeurs irréalistes (grands nombres négatifs) pouvant résulter de données incorrectes à un point de référence. Cela pourrait se produire dans des bassins versants qui n'ont pas encore été testés.

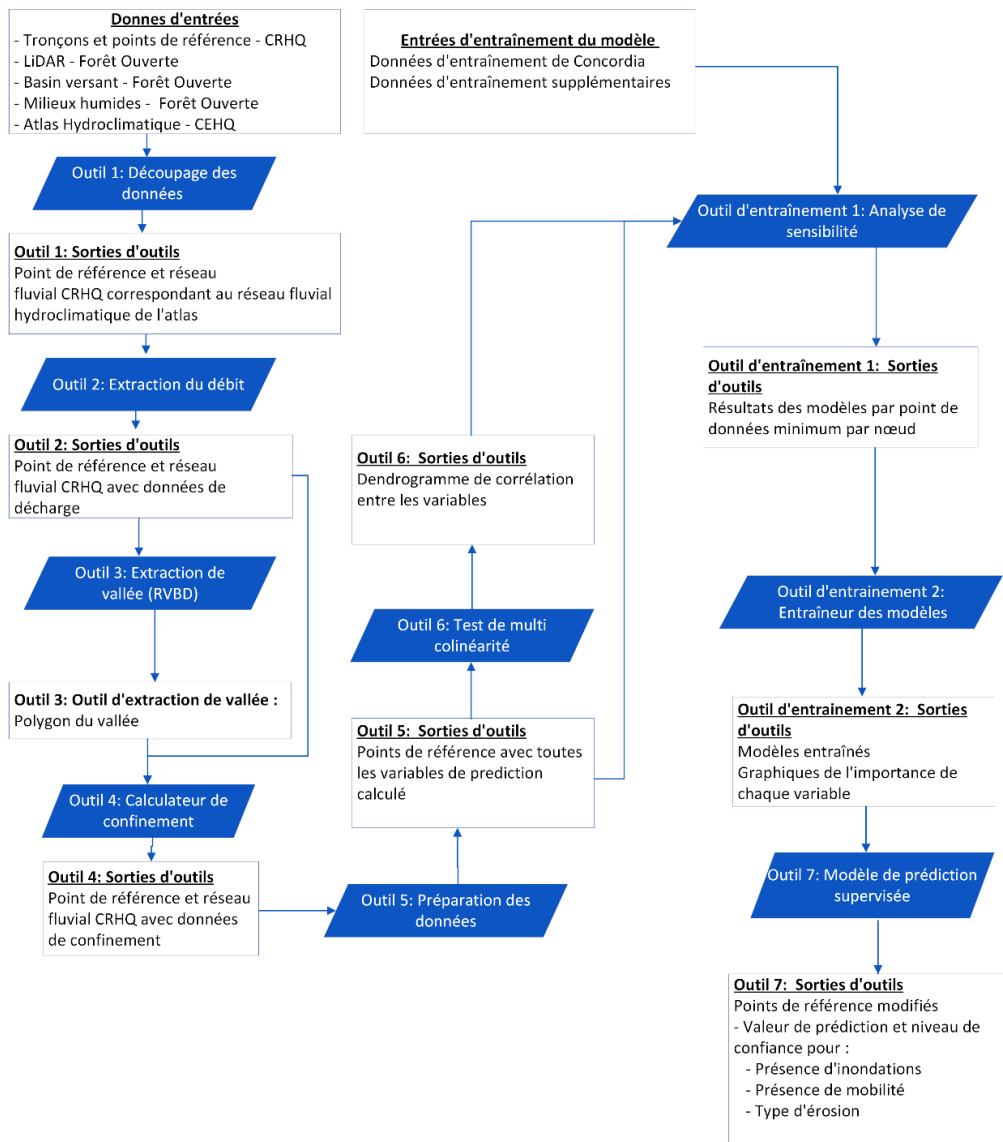
Étape 11 : ajoutez toute nouvelle variable non testée précédemment à la liste des variables de ce script ou modifiez les noms des variables modifiées (expliqué en B.3)

Étape 12 : exécutez l'outil 6 : Test de multi colinéarité (figure B1).

Étape 13 : supprimez les variables des scripts d'entraînement et de prédiction (expliqués à la section 3.4.4 et B.3) qui sont trop corrélées en laissant seulement 1 d'entre elles pour chaque groupement ou ajouter les nouvelles variables

Étape 14 : exécutez l'outil 7 : Prédiction d'aléa (figure B1).

# **DÉVELOPPEMENT D'UN OUTIL POUR L'INTÉGRATION DE L'ANALYSE DES ALÉAS ASSOCIÉS À LA DYNAMIQUE GÉOMORPHOLOGIQUE DES COURS D'EAU – PROJET DE RECHERCHE (R800.1)**



**Figure B1: Flux de travail de l'outil**

## **Annexe B.2 Étapes d'entraînement**

Étape 1 : collecte des données de validation suivant les instructions de la section 3.4.1.

Étape 2 : classer les points de référence suivants les instructions de la section 3.4.2

Étape 3 : combiner les nouvelles données d'entraînement avec les anciennes.

Étape 4 : exécuter l'outil 6 : Test de multi colinéarité (figure B1).

Étape 5 : Supprimer les variables des scripts d'entraînement et de prédiction (expliqué à la section 3.4.3 et B.3) qui sont trop corrélées en laissant seulement 1 d'entre elles pour chaque groupement

Étape 6 : exécuter l'outil d'entraînement 1 : Analyse de sensibilité (figure B1).

Étape 7 : choisir un seuil pour le nombre minimum de points de données par nœud suivant les recommandations de la section 4.1.2.

Étape 8 : exécuter l'outil d'entraînement 2 : Entraîneur de modèle (figure B1).

### **Annexe B.3 Modifications pour être compatibles avec les futures versions du CRHQ**

Dans les futures versions du CRHQ, les valeurs de débit extraites de l'Atlas Hydroclimatique seront mises à jour et il ne sera donc plus nécessaire d'utiliser l'outil 2 : Extraction du débit. Il est également prévu d'inclure des mesures de la largeur de la vallée ou du confinement directement dans le CRHQ et il ne sera donc plus nécessaire d'utiliser les outils 3 et 4. Par conséquent, seuls les outils 1, 5 et 6 seront nécessaires une fois que ces nouveaux champs auront été ajoutés ou mis à jour. Ces outils devront toutefois être modifiés pour tenir compte des nouvelles conventions de dénomination. Cela peut facilement être fait en recherchant les anciens noms de champs qui ont été modifiés dans ces trois scripts et en les remplaçant par les nouveaux noms. Pour exclure des champs ou ajouter de nouveaux champs au modèle, il suffit d'aller modifier la liste des champs suivant "X =" dans les outils d'entraînement 1 et 2 ainsi que dans l'outil 6 et 7 (figure B2). Il y aura 3 lignes de ce type à modifier par outil.

```
X = data.loc[data["Inondation"] >= 0, ["Slope_loc", "Sinuosite",  
"Ang_dev_q50", "Alt_moy", "S_moy", "Largeur_mod", "Abv_km2",  
"USP", "Ratio_d", 'LocConfRatio', 'RelConfRatio',  
'ValConfRatio', 'Anthro', 'Vegetation', 'Wetland', 'HumanAct']]
```

---

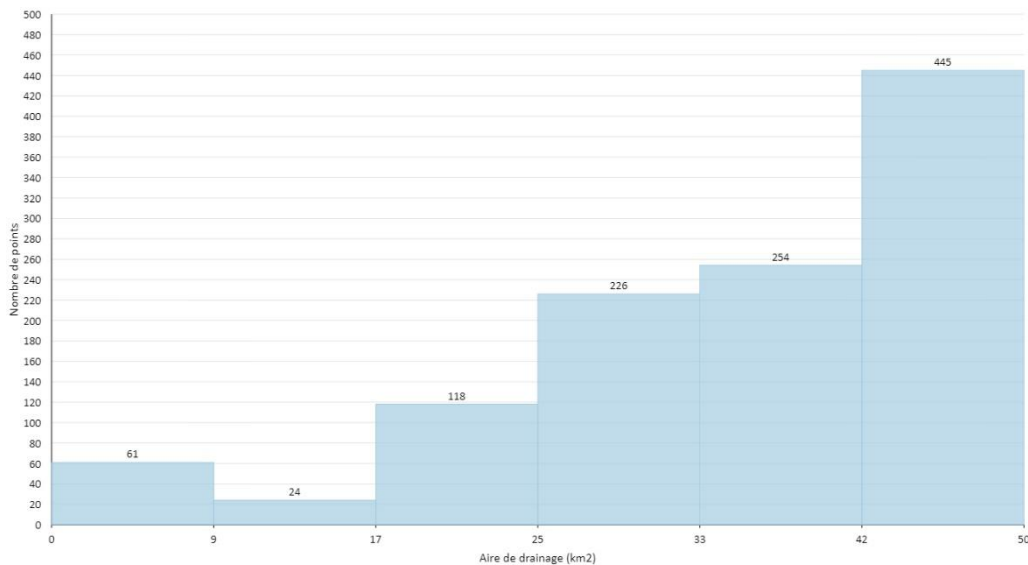
**Figure B2: Exemple de ligne à modifier pour changer les champs inclus dans le modèle, la partie qui doit être modifiée est la liste des champs inclus dans les crochets.**

## **ANNEXE C**

### **ANALYSE DES COURS D'EAU DONT L'AIRE DE DRAINAGE EST INFÉRIEURE À 50 KM<sup>2</sup>**

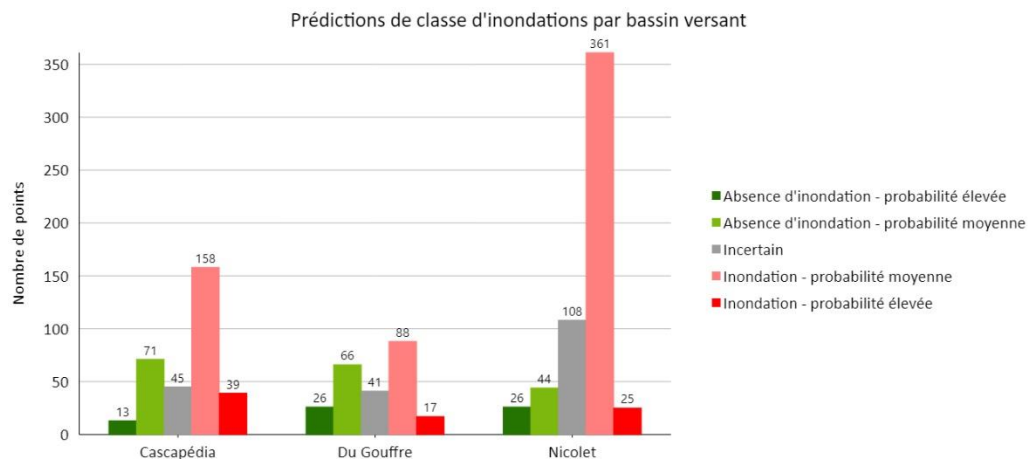
---

Pour pouvoir tester la capacité du modèle à extrapoler sur des cours d'eau plus petits sans avoir été entraîné sur ceux-ci, un sous-ensemble de points avec des aires de drainage inférieures à 50 km<sup>2</sup> qui faisaient partie d'un des trois bassins versants pilotes a été inclus. Ce sous-ensemble a utilisé le débit spécifique de l'Atlas Hydroclimatique le plus proche pour calculer la puissance spécifique (1128 points). La distribution des aires de drainage incluses indique que la majorité est supérieure à 25 km<sup>2</sup>, bien que certains affluents plus petits aient également été inclus (figure C1). Les résultats des trois modèles sur ces cours d'eau sont présentés dans les figures C2, C3 et C4. Ils suivent la même tendance que les données totales présentées dans la section 5.2 bien qu'avec des niveaux de confiance plus faibles. Une validation manuelle effectuée sur ces points à l'aide du LiDAR et de données satellitaires révèle que le modèle fonctionne bien pour différencier les processus fluviaux de mobilité et le type d'érosion sur ces cours d'eau (figures C5 et C6). Cela montre que ce type de modèle est effectivement adapté à la correction du bruit dans les données et à la généralisation à des environnements sur lesquels il n'a pas été formé. Il est à noter que l'inclusion de ces points de données n'a pas eu d'impact sur la valeur de prédiction des rivières de plus de 50 km<sup>2</sup> car ils n'ont pas été utilisés lors de l'étape d'entraînement du modèle.

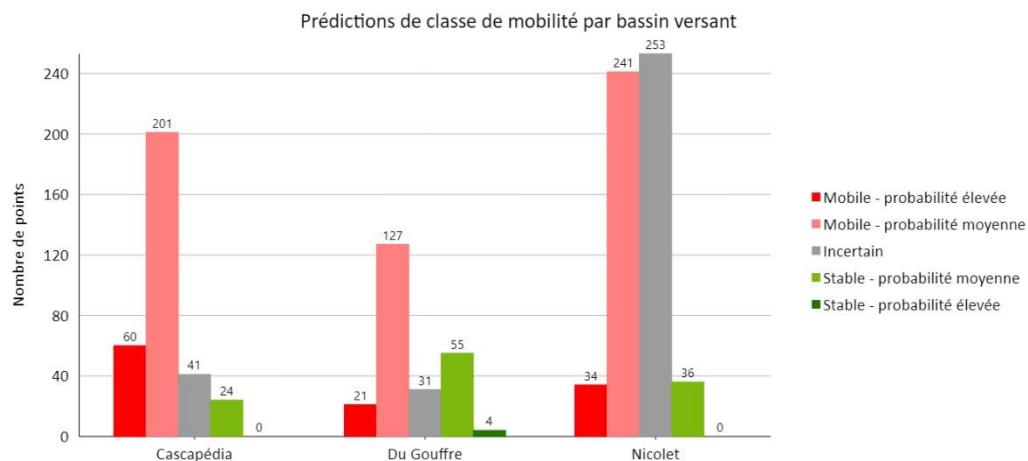


**Figure C1: Répartition des aires de drainage pour les points dont l'aire de drainage est inférieure à 50 km<sup>2</sup>**

DÉVELOPPEMENT D'UN OUTIL POUR L'INTÉGRATION DE L'ANALYSE DES ALÉAS  
ASSOCIÉS À LA DYNAMIQUE GÉOMORPHOLOGIQUE DES COURS D'EAU – PROJET DE  
RECHERCHE (R800.1)

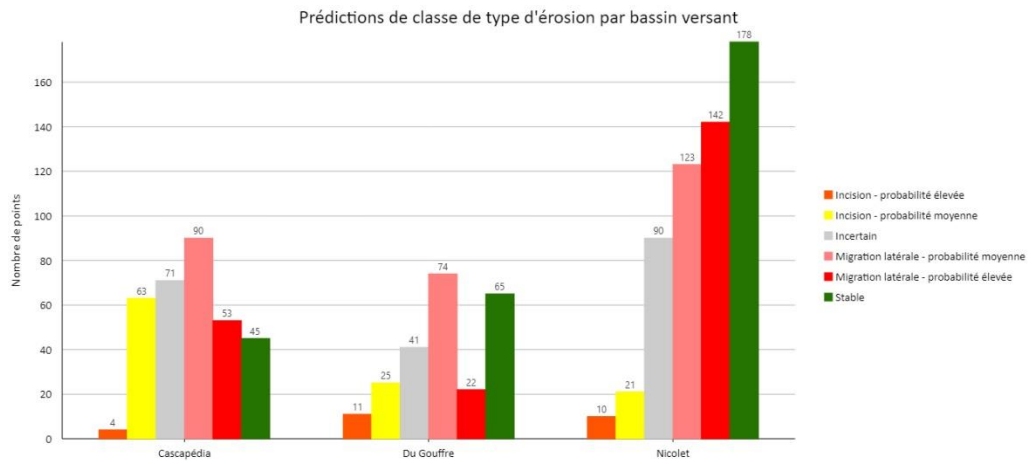


**Figure C2: Prédictions de classe d'inondation par bassin versant pour les points avec aires de drainage inférieures à 50 km<sup>2</sup>**

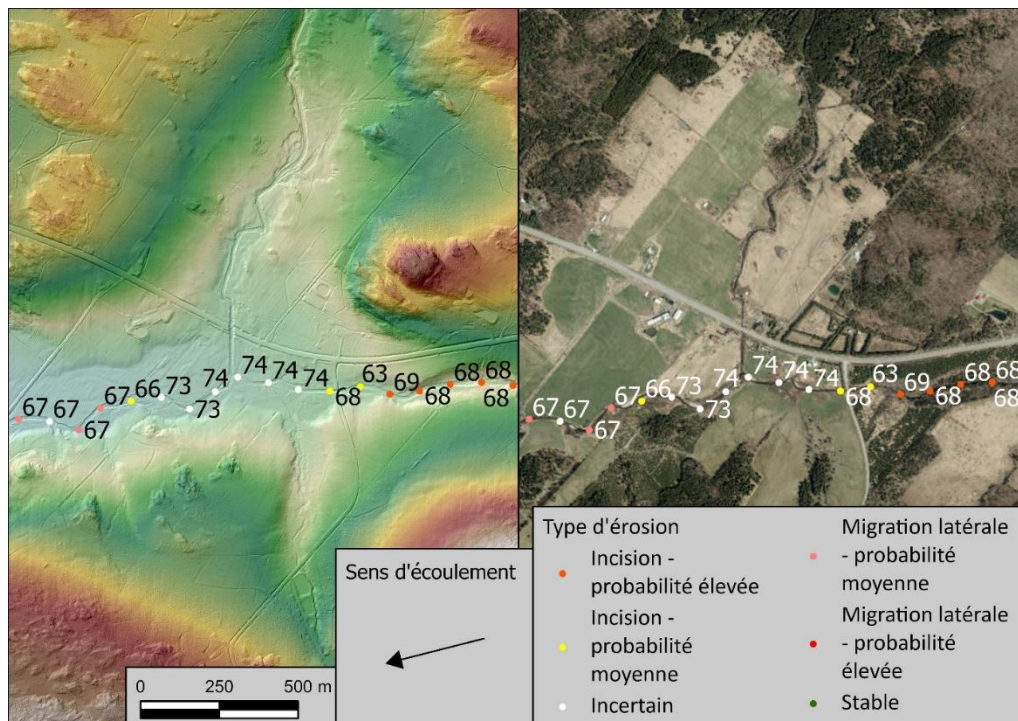


**Figure C3: Prédictions de classe de mobilité par bassin versant pour les points avec aires de drainage inférieures à 50 km<sup>2</sup>**

**DÉVELOPPEMENT D'UN OUTIL POUR L'INTÉGRATION DE L'ANALYSE DES ALÉAS  
ASSOCIÉS À LA DYNAMIQUE GÉOMORPHOLOGIQUE DES COURS D'EAU – PROJET DE  
RECHERCHE (R800.1)**

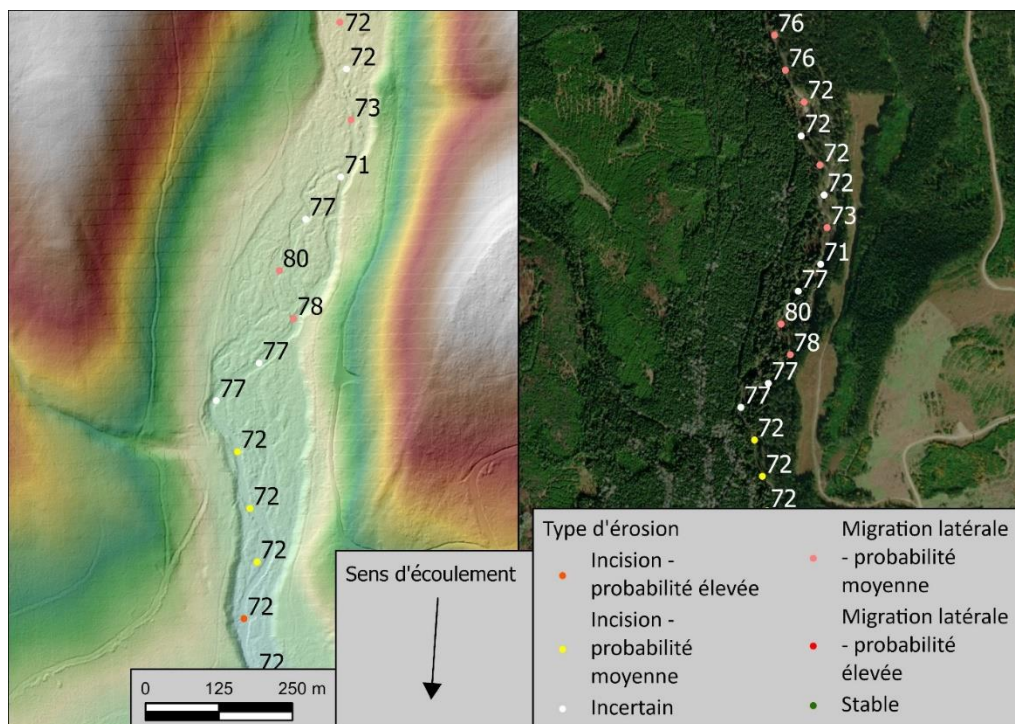


**Figure C4: Prédictions de classe de type d'érosion par bassin versant pour les points avec aires de drainage inférieures à 50 km<sup>2</sup>**



**Figure C5: Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Nicolet : 71.6207918°W 45.8890004°N)**





**Figure C6: Carte montrant les prévisions du modèle de type d'érosion. Les chiffres représentent la probabilité de présence de mobilité du modèle. (Cascapédia : 66.2446110°W 48.7449664°N)**

## **ANNEXE D**

### **DISPONIBILITÉS DE LA BASE DE DONNÉES CRHQ**

---

DÉVELOPPEMENT D'UN OUTIL POUR L'INTÉGRATION DE L'ANALYSE DES ALÉAS ASSOCIÉS À LA DYNAMIQUE GÉOMORPHOLOGIQUE DES COURS D'EAU – PROJET DE RECHERCHE (R800.1)

Cette annexe présente une carte des bassins versants où le CRHQ est disponible (figure D1). Il est prévu d'améliorer la couverture et la qualité de cette base de données.

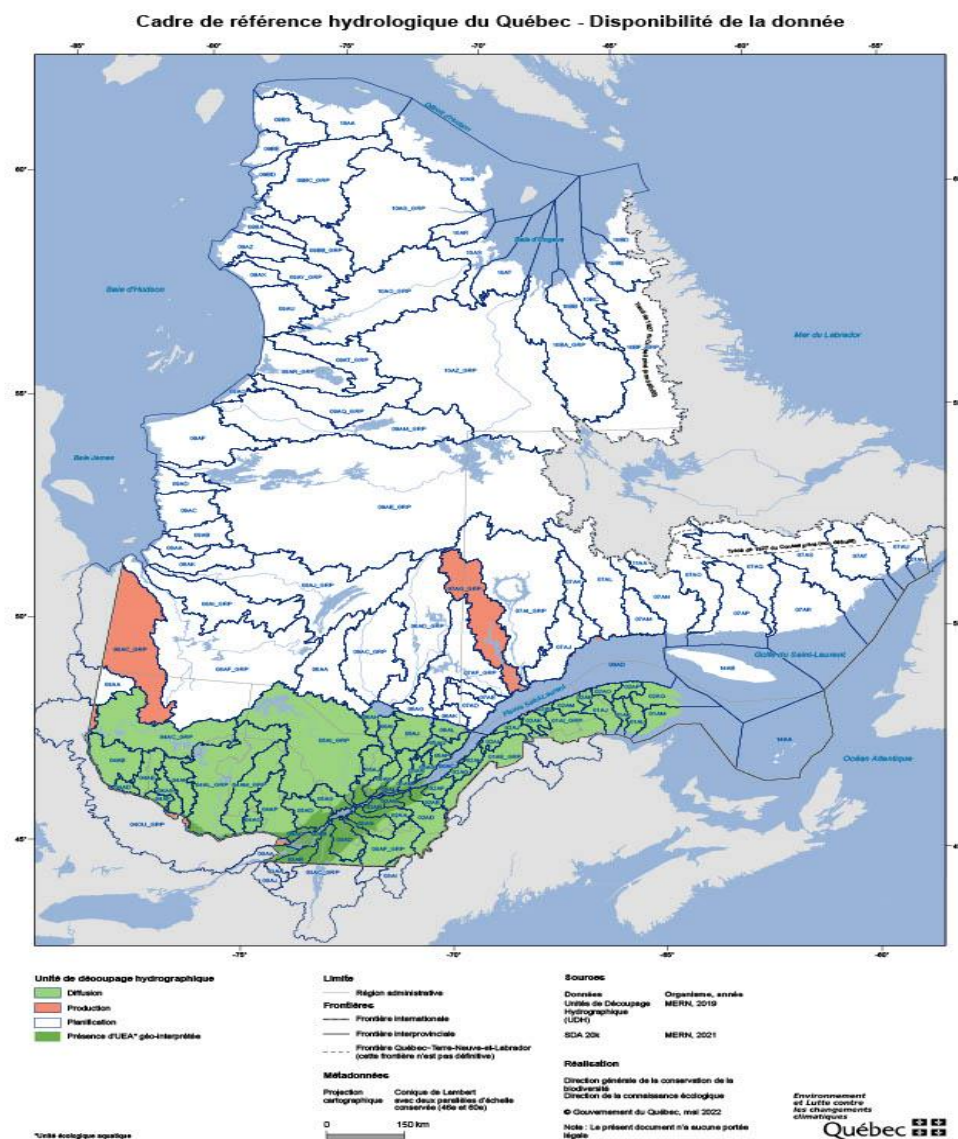


Figure D1: Carte de disponibilité des données du CRHQ

## RÉFÉRENCES

1. ANDERSON, I., RIZZO, D. M., HUSTON, D. R., & DEWOOLKAR, M. M. (2017). Stream Power Application for Bridge-Damage Probability Mapping Based on Empirical Evidence from Tropical Storm Irene. *Journal of Bridge Engineering*, 22(5), 05017001. [https://doi.org/10.1061/\(ASCE\)BE.1943-5592.0001022](https://doi.org/10.1061/(ASCE)BE.1943-5592.0001022)
2. AVAND, M., MOHAMMADI, M., MIRCHOOI, F., KAVIAN, A., & TIEFENBACHER, J. P. (2023). A New Approach for Smart Soil Erosion Modeling: Integration of Empirical and Machine-Learning Models. *Environmental Modeling & Assessment*, 28(1), 145–160. <https://doi.org/10.1007/s10666-022-09858-x>
3. BAGNOLD, R. A. (1966). An approach to the sediment transport problem from general physics (No. 422–I). U. S. Govt. Print. Off., <https://doi.org/10.3133/pp422I>
4. BEECHIE, T. J., LIERMANN, M., POLLOCK, M. M., BAKER, S., & DAVIES, J. (2006). Channel pattern and river-floodplain dynamics in forested mountain river systems. *Geomorphology*, 78(1–2), 124–141. <https://doi.org/10.1016/j.geomorph.2006.01.030>
5. BEECHIE, T., & IMAKI, H. (2014). Predicting natural channel patterns based on landscape and geomorphic controls in the Columbia River basin, USA: Predicting Channel Patterns in the Columbia Basin. *Water Resources Research*, 50(1), 39–57. <https://doi.org/10.1002/2013WR013629>
6. BELLETTI, B., RINALDI, M., BUSSETTINI, M., COMITI, F., GURNELL, A. M., MAO, L., NARDI, L., & VEZZA, P. (2017). Characterising physical habitats and fluvial hydromorphology: A new system for the survey and classification of river geomorphic units. *Geomorphology*, 283, 143–157. <https://doi.org/10.1016/j.geomorph.2017.01.032>
7. BREIMAN, L. (2001). Random Forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
8. BIZZI, S., & LERNER, D. N. (2015). The Use of Stream Power as an Indicator of Channel Sensitivity to Erosion and Deposition Processes: SP AS AN INDICATOR OF EROSION AND DEPOSITION. *River Research and Applications*, 31(1), 16–27. <https://doi.org/10.1002/rra.2717>
9. BIRON, P. M., BUFFIN-BÉLANGER, T., LAROCQUE, M., CHONÉ, G., CLOUTIER, C.-A., OUELLET, M.-A., DEMERS, S., OLSEN, T., DESJARLAIS, C., & EYQUEM, J. (2014). Freedom Space for Rivers: A Sustainable Management Approach to Enhance River Resilience.

- Environmental Management*, 54(5), 1056–1073.  
<https://doi.org/10.1007/s00267-014-0366-z>
10. BUFFIN-BÉLANGER, T., DEMERS S. et OLSEN T. (2015). Diagnostic hydrogéomorphologique pour mieux considérer les dynamiques hydrosédimentaires aux droits des traverses de cours d'eau : guide méthodologique. Laboratoire de géomorphologie et de dynamique fluviale, Université du Québec à Rimouski. Remis au ministère des Transports du Québec, mars 2015, 55 pages
  11. BUFFINGTON, J. M., & MONTGOMERY, D. R. (2021). Geomorphic Classification of Rivers: An Updated Review. In *Treatise on Geomorphology* (pp. 1143–1190). Elsevier. <https://doi.org/10.1016/B978-0-12-818234-5.00077-8>
  12. CAVALLI, M., CREMA, S., TREVISANI, S., & MARCHI, L. (2017). GIS tools for preliminary debris-flow assessment at regional scale. *Journal of Mountain Science*, 14(12), 2498–2510. <https://doi.org/10.1007/s11629-017-4573-y>
  13. CHONÉ, G., BIRON, P. M., BUFFIN-BÉLANGER, T., MAZGAREANU, I., NEAL, J. C., & SAMPSON, C. C. (2021). An assessment of large-scale flood modelling based on LiDAR data. *Hydrological Processes*, 35(8). <https://doi.org/10.1002/hyp.14333>
  14. CREMA, S., & CAVALLI, M. (2018). SedInConnect: a stand-alone, free and open-source tool for the assessment of sediment connectivity. *Computers & Geosciences*, 111, 39–45. <https://doi.org/10.1016/j.cageo.2017.10.009>
  15. DEMERS, S., OLSEN, T., BUFFIN-BÉLANGER, T., MARCHAND, J.-P., BIRON, P. M., & MORNEAU, F. (2014). L'hydrogéomorphologie appliquée à la gestion de l'aléa d'inondation en climat tempéré froid : l'exemple de la rivière Matane (Québec). *Physio-Géo*, Volume 8, 67–88. <https://doi.org/10.4000/physio-geo.3813>
  16. DÉPRET, T., RIQUIER, J., & PIÉGAY, H. (2017). Evolution of abandoned channels: Insights on controlling factors in a multi-pressure river system. *Geomorphology*, 294, 99–118. <https://doi.org/10.1016/j.geomorph.2017.01.036>
  17. DUPUIS, M., LEBOEUF, A., VAILLANCOURT, E., & LACHANCE, D. (2017). *Guide d'utilisation du lidar pour l'identification des dépôts de surface et des bancs d'emprunts*. Ministère des forêts, de la faune et des parcs, Secteur des Forêts, Direction des inventaires forestiers.
  18. FANG, L., HUANG, J., CAI, J., & NITIVATTANANON, V. (2022). Hybrid approach for flood susceptibility assessment in a flood-prone mountainous



- catchment in China. *Journal of Hydrology*, 612, 128091.  
<https://doi.org/10.1016/j.jhydrol.2022.128091>
19. FLORSHEIM, J., MOUNT, J., CHIN, A. (2008). Bank erosion as a desirable attribute of rivers. *Bioscience*, 58, 521-529.
  20. FRYIRS, K. A., WHEATON, J. M., & BRIERLEY, G. J. (2016). An approach for measuring confinement and assessing the influence of valley setting on river forms and processes: Measuring Confinement along Fluvial Corridors. *Earth Surface Processes and Landforms*, 41(5), 701–710.  
<https://doi.org/10.1002/esp.3893>
  21. FRYIRS, K. A., & BRIERLEY, G. J. (2018). What's in a name? A naming convention for geomorphic river types using the River Styles Framework. *PLOS ONE*, 13(9), e0201909.  
<https://doi.org/10.1371/journal.pone.0201909>
  22. GARTNER, J. D., DADE, W. B., RENSHAW, C. E., Magilligan, F. J., & Buraas, E. M. (2015). Gradients in stream power influence lateral and downstream sediment flux in floods. *Geology*, 43(11), 983–986.  
<https://doi.org/10.1130/G36969.1>
  23. GARTNER, J. (2016). *Stream Power: Origins, Geomorphic Applications, and GIS Procedures*. Water Publications.  
[https://scholarworks.umass.edu/water\\_publications/1](https://scholarworks.umass.edu/water_publications/1)
  24. GHOSH, A., & DEY, P. (2021) «Flood Severity assessment of the coastal tract situated between Muriganga and Saptamukhi estuaries of Sundarban delta of India using Frequency Ratio (FR), Fuzzy Logic (FL), Logistic Regression (LR) and Random Forest (RF) models», *Regional Studies in Marine Science*, 42, 101624. <https://doi.org/10.1016/j.rsma.2021.101624>
  25. GILBERT, J. T., MACFARLANE, W. W., & WHEATON, J. M. (2016) «The Valley Bottom Extraction Tool (V-BET): A GIS tool for delineating valley bottoms across entire drainage networks». *Computers & Geosciences*, 97, 1–14. <https://doi.org/10.1016/j.cageo.2016.07.014>
  26. HASAN, Md. H., AHMED, A., NAFEE, K. M., & HOSSEN, Md. A. (2023) « Use of machine learning algorithms to assess flood susceptibility in the coastal area of Bangladesh ». *Ocean & Coastal Management*, 236, 106503. <https://doi.org/10.1016/j.ocecoaman.2023.106503>
  27. HORACIO, J., OLLERO, A., & PÉREZ-ALBERTI, A. (2017) « Geomorphic classification of rivers: a new methodology applied in an Atlantic Region (Galicia, NW Iberian Peninsula) ». *Environmental Earth Sciences*, 76(21), 746, 2017. <https://doi.org/10.1007/s12665-017-7072-0>

28. KASPRAK, A., HOUGH-SNEE, N., BEECHIE, T., BOUWES, N., BRIERLEY, G. J., CAMP, R., FRYIRS, K. A., IMAKI, H., JENSEN, M. L., O'BRIEN, G., ROSGEN, D. L., & WHEATON, J. M. (2015) « Choosing the right tool for the job: comparing stream channel classification frameworks (No. e1093) ». *PeerJ PrePrints*. <https://doi.org/10.7287/peerj.preprints.885v1>
29. KAVZOGLU, T., & TEKE, A. (2022). Predictive Performances of Ensemble Machine Learning Algorithms in Landslide Susceptibility Mapping Using Random Forest, Extreme Gradient Boosting (XGBoost) and Natural Gradient Boosting (NGBoost). *Arabian Journal for Science and Engineering*, 47(6), 7367–7385. <https://doi.org/10.1007/s13369-022-06560-8>
30. KONDOLF, G. M. « PROFILE: (1997) Hungry Water: Effects of Dams and Gravel Mining on River Channels ». *Environmental Management*, 21(4), 533–551,. <https://doi.org/10.1007/s002679900048>
31. LEOPOLD, L.B. (1962) «Rivers». *American Scientist*, 50, 4, 511-537.
32. HACK, J.T. (1972) « Dynamic Equilibrium and Landscape Evolution, In Melhorn, W.N. and Flemal, R.C. (eds.) » *Theories of Landform Development, Proceedings of the Sixth Annual Binghampton Symposium*, 87-102.
33. MELCCFP.(2023) « Projet INFO-Crue - Zones inondables ». *MELCCC [En ligne]*. [<https://www.cehq.gouv.qc.ca/zones-inond/info-crue/index.htm>]
34. PARMAR, A., KATARIYA, R., & PATEL, V. (2019) « A Review on Random Forest: An Ensemble Classifier ». In J. Hemanth, X. Fernando, P. Lafata, & Z. Baig (Eds.), *International Conference on Intelligent Data Communication Technologies and Internet of Things (ICICI) 2018* (pp. 758–763). Springer International Publishing. [https://doi.org/10.1007/978-3-030-03146-6\\_86](https://doi.org/10.1007/978-3-030-03146-6_86)
35. PARKER, C., & DAVEY, J. (2023) « Stream power indices correspond poorly with observations of alluvial river channel adjustment ». *Earth Surface Processes and Landforms*, esp.5550. <https://doi.org/10.1002/esp.5550>
36. PIÉGAY, H., DARBY, S.E., MOSSELMAN, E. and SURIAN, N. (2005) « A review of techniques available for delimiting the erodible river corridor: a sustainable approach to managing bank erosion ». *River Res. Applic.*, 21: 773-789. <https://doi.org/10.1002/rra.881>
37. RAZAVI-TERMEH, S. V., SEO, M., SADEGHI-NIARAKI, A., & CHOI, S.-M. (2023) « Application of genetic algorithm in optimization parallel ensemble-based machine learning algorithms to flood susceptibility

- mapping using radar satellite imagery ». *Science of The Total Environment*, 162285. <https://doi.org/10.1016/j.scitotenv.2023.162285>
38. ROSGEN, D. L. (1994). « A classification of natural rivers ». *CATENA*, 22(3), 169–199, 1994. [https://doi.org/10.1016/0341-8162\(94\)90001-9](https://doi.org/10.1016/0341-8162(94)90001-9)
  39. SCHUMM, S. A. (1977) « The fluvial system ». *John Wiley & Sons, New York*, 1977.
  40. SCIKIT-LEARN DEVELOPERS. SCIKIT-LEARN. 1.11. *Ensemble methods*. [En Ligne] [<https://scikit-learn/stable/modules/ensemble.html>]
  41. SECHU, G. L., NILSSON, B., IVERSEN, B. V., GREVE, M. B., BØRGESSEN, C. D., & GREVE, M. H. (2021) « A Stepwise GIS Approach for the Delineation of River Valley Bottom within Drainage Basins Using a Cost Distance Accumulation Analysis ». *Water*, 13(6), 827. <https://doi.org/10.3390/w13060827>
  42. SEAR, D., HORNBLY, D., WHEATON, J., & HILL, C. (2021) « Understanding river channel sensitivity to geomorphological changes ». *Environment Agency*.
  43. TIWARI, A. (2022) Supervised learning: From theory to applications. In *Artificial Intelligence and Machine Learning for EDGE Computing*. Elsevier. <https://doi.org/10.1016/B978-0-12-824054-0.00026-5>, 23–32. p.
  44. WANG, Z., LAI, C., CHEN, X., YANG, B., ZHAO, S., & BAI, X. (2015) « Flood hazard risk assessment model based on random forest ». *Journal of Hydrology*, 527, 1130–1141. <https://doi.org/10.1016/j.jhydrol.2015.06.008>
  45. WHEATON, J. M., FRYIRS, K. A., BRIERLEY, G., BANGEN, S. G., BOUWES, N., & O'BRIEN, G. (2015) « Geomorphic mapping and taxonomy of fluvial landforms ». *Geomorphology*, 248, 273–295. <https://doi.org/10.1016/j.geomorph.2015.07.010>
  46. WHEELER, N., PINGRAM, M., DAVID, B., MARSON, W., TUNNICLIFFE, J., & BRIERLEY, G. (2022) « River adjustments, geomorphic sensitivity and management implications in the Waipā catchment, Aotearoa New Zealand ». *Geomorphology*, 108263. <https://doi.org/10.1016/j.geomorph.2022.108263>
  47. YOUSSEF, A. M., MAHDI, A. M., AL-KATHERI, M. M., POUYAN, S., & POURGHASEMI, H. R. (2023) « Multi-hazards (landslides, floods, and gully erosion) modeling and mapping using machine learning algorithms ». *Journal of African Earth Sciences*, 197, 104788. <https://doi.org/10.1016/j.jafrearsci.2022.104788>