

Université du Québec
Institut National de la Recherche Scientifique
Centre Eau Terre Environnement

DÉVELOPPEMENT D'UNE MÉTHODOLOGIE D'ESTIMATION DU DÉBIT EN RIVIÈRE POUR LES SITES NON-JAUGÉS À L'AIDE DE L'IMAGERIE RADARSAT DANS L'EST DU CANADA

**Rapport technique
R1683**

Par

Karem Chokmani, Ph.D.

Simon Perreault, M.Sc.

Jimmy Poulin, M.Sc.

Andrès Jacome, Ph.D.

Yves Gauthier, M.Sc.

Monique Bernier Ph.D.

Projet financé par le Programme de Développement d'Applications en Observation de la Terre, Agence Spatiale Canadienne



29 mai 2015

© INRS, Centre - Eau Terre Environnement, 2016
Dépôt légal, Bibliothèque nationale du Québec
Dépôt légal, Bibliothèque et archives Canada
ISBN : 978-2-89146-867-1

Résumé

Les mesures de débit en rivière sont nécessaires, entre autres, pour la planification des ressources en eau et les études sur le climat et l'écologie. Toutefois, les stations du réseau hydrométrique canadien sont généralement localisées dans les zones habitées et/ou d'intérêt socio-économique évident. Ce projet visait à développer une méthodologie d'estimation du débit en rivière dans des sites non-jaugés à l'aide de l'imagerie RADARSAT dans l'Est du Canada. Les travaux ont d'abord permis de développer des courbes de tarage pour les rivières jaugées à partir d'information extraite de l'imagerie RADARSAT. La courbe de tarage est la fonction qui relie la largeur effective de la surface d'eau au débit. Les coefficients de détermination des courbes de tarage varient de 0,50 à 0,90. La méthode d'extraction automatique de la surface d'eau se démarque par le recours à une fonction de probabilité par logique floue au lieu d'utiliser une approche « classique » par seuil. De plus, la faisabilité d'une approche de régionalisation basée sur la notion de similarité hydrologique a été démontrée pour le transfert des courbes de tarage satellitaires des sites jaugés vers les sites non-jaugés en vue d'estimer le débit à ces derniers à partir de l'imagerie RADARSAT. Malgré la preuve du concept qui a été démontrée, le faible nombre de stations de débits utilisables pour la calibration de ce modèle n'a pas permis d'obtenir une évaluation concluante des débits estimés sur les sites non-jaugés.

Table des matières

Résumé	I
Liste des tableaux	IV
Liste des figures	VI
1 Introduction	1
2 Développement de la base de données	3
2.1 Sélection des stations hydrométriques et des images.....	3
2.2 Base de données physiographiques, climatiques et cartographiques.....	8
2.2.1 Données brutes	8
2.2.2 Données dérivées	11
2.3 Traitements des images RS-1&2.....	13
3 Élaboration de la courbe de tarage satellitaire.....	16
3.1 Obtention du masque pour l'application du calcul de probabilité	17
3.2 Intercalibration des orthoimages	18
3.3 Paramétrage de la fonction d'appartenance de l'eau.....	20
3.4 Calcul de la largeur effective	22
3.5 Ajustement d'une fonction pour la courbe de tarage.....	23
3.6 Analyse de sensibilité des courbes de tarage.....	27
4 Développement du modèle d'estimation du débit aux sites non-jaugés	30
4.1 Sélection des stations pour la calibration du modèle	30
4.2 Régionalisation des courbes de tarage.....	35
4.2.1 Exploration de la base de données physiographique.....	35
4.2.2 Analyse de la normalité	36
4.2.3 Analyse de la corrélation entre les variables physiographiques transformées et les paramètres des courbes de tarage.....	39
4.2.4 Analyse de la corrélation croisée entre les variables physiographiques transformées	41
4.2.5 Analyse canonique de corrélation.....	43
4.2.6 Régressions avec les variables canoniques pour prédire les paramètres des courbes de tarage.....	48
4.3 Validation.....	56
4.4 Automatisation de la procédure pour les sites non-jaugés	60

5	Conclusions	62
6	Recommandations.....	64
7	Divulgence de la propriété intellectuelle par l'entrepreneur.....	68
8	Évaluation du succès global du projet.....	69
8.1	Évaluation des résultats du projet par rapport aux indicateurs de rendement.....	69
8.2	Impacts et avantages réels et potentiels.....	73
8.3	Succès global	76
9	Références	78
	Annexes.....	79
	Annexe I : Description détaillée des calculs de variables physiographiques.....	80
	Annexe II : Contenu de la base de données	87
	Annexe III : Analyse de la corrélation	96
	Annexe IV : Procédure pour déterminer le débit sur des images RADARSAT pour un site non jaugé.	99
	Annexe V : Spécification des exigences associées au système et aux interfaces.....	108

Liste des tableaux

Tableau 1 : Décompte des stations hydrométriques initialement identifiées	3
Tableau 2 : Décompte des stations hydrométriques retenues par province	6
Tableau 3 : Liste de stations hydrométriques retenues, et nombre total d'images RS-1&2 par station.....	7
Tableau 4 : Décompte des images d'archives RS-1&2 prétraitées.....	15
Tableau 5 : Exemple de sorties du script pour le calcul de la largeur effective	23
Tableau 6 : Paramètres conservés dans la table courbe de tarage exponentielle.....	27
Tableau 7 : Variation du débit (m ³ /s) engendrée par une variation de 1% des paramètres selon le type de fonction pour une largeur effective de 40 m	28
Tableau 8 : Variation du débit (m ³ /s) engendrée par une variation de 1% des paramètres selon le type de fonction pour une largeur effective de 65 m	28
Tableau 9 : Stations hydrométriques avec couples image-débit non-représentatifs	30
Tableau 10 : Stations hydrométriques sélectionnées pour la calibration du modèle	32
Tableau 11 : Stations hydrométriques non-retenues pour la calibration du modèle ($r^2 < 0.5$).....	33
Tableau 12 : Transformations appliquées aux paramètres des courbes de tarage pour respecter la normalité	39
Tableau 13 : Corrélations des variables physiographiques présentant au minimum une corrélation significative (p-value < 0,05) avec un paramètre de courbe de tarage	40
Tableau 14 : Sélection des variables physiographiques non redondantes présentant une corrélation avec au moins un paramètre des courbes de tarage	42
Tableau 15 : Variables physiographiques retenues pour la régionalisation du modèle exponentiel	43
Tableau 16 : Variables physiographiques retenues pour la régionalisation du modèle polynomial d'ordre 2	43
Tableau 17 : Moyennes des variables physiographiques et coefficients canoniques pour la transformation en variables canoniques (modèle exponentiel).....	44
Tableau 18: Moyennes des variables physiographiques et coefficients canoniques pour la transformation en variables canoniques (modèle polynomial)	44

Tableau 19 : Variables canoniques U, générées pour l'ensemble des stations dans chaque modèle de courbe de tarage	45
Tableau 20 : Coefficients de régressions et ordonnées à l'origine pour l'estimation des paramètres Alpha et Beta du modèle exponentiel de courbe de tarage.....	49
Tableau 21 : Coefficients de régressions et ordonnées à l'origine pour l'estimation des paramètres Alpha et Beta du modèle exponentiel de courbe de tarage.....	52
Tableau 22 : Qualité de l'ajustement des régressions multiples réalisées pour l'estimation de paramètre de courbes de tarage.....	55
Tableau 23 : Variables canoniques générées pour les deux types de modèles de courbes de tarage à partir des données de validation.....	57
Tableau 24 : Paramètres des courbes de tarage générés à partir des modèles régionalisés	58
Tableau 25 : Divulguation de la propriété intellectuelle sur les renseignements originaux (BIP) requis pour utiliser/améliorer la FIP dans le cadre du projet	68
Tableau 26 : Liste de classes d'occupation du sol.	85
Tableau 27 : Définition des classes d'occupation du sol des données d'Observation de la Terre pour le Développement Durable des Forêts (EOSD)	86
Tableau 28 : Structure du fichier Excel nécessaire à l'utilisation de l'outil de création des courbes de tarage.....	94
Tableau 29 : Liste de variables climatologiques, physiographiques et cartographiques produite par station hydrométrique.	95
Tableau 30 : Matrice de corrélation croisée pour les 33 variables physiographiques présentant des corrélations significatives avec au minimum un paramètre de courbe de tarage	96
Tableau 31 : Légende pour la matrice de corrélation croisée.....	98
Tableau 32 : Description des bibliothèques Python utilisées.....	109
Tableau 33 : Bibliothèques Python développées pour le projet.....	109

Liste des figures

Figure 1 : Étapes du projet et livrables correspondants.....	2
Figure 2 : Stations hydrométriques potentielles.....	4
Figure 3 : Probabilité de la présence de neige/glace : (a) 5% printemps et (b) 50% automne (jour julien).....	5
Figure 4 : Distribution spatiale des stations hydrométriques retenues.....	6
Figure 5 : Distribution des 77 bassins versants des sites jaugés retenus.....	12
Figure 6 : Exemple de validation d'une orthoimage RS-2 (station Chateauguay, échelle approximative 1/35 000) : a) à corriger (DK_371116), b) bien orthorectifiée (DK_371266).....	14
Figure 7 : Exemple d'application des filtres de Lee et Enhanced Frost, par rapport à une orthoimage sur la rivière Richelieu non filtrée.....	15
Figure 8 : Diagramme de la méthode de calcul de la courbe de tarage locale.....	16
Figure 9 : Diagramme de la méthode automatisée de calcul de la probabilité d'appartenance à l'eau par logique floue.....	17
Figure 10 : Obtention du masque pour l'application du calcul de la probabilité d'appartenance.....	18
Figure 11 : Ajout d'une zone tampon au polygone du chenal principal.....	18
Figure 12 : Histogrammes de la rétrodiffusion avant (a) et après (b) l'intercalibration en utilisant le 1er percentile et la valeur -23 dB comme référence.....	19
Figure 13 : Exemple d'une Fonction d'appartenance logique floue « Z » (fuzzification).....	20
Figure 14 : Exemple d'un histogramme de densité de probabilité bimodal (en bleu l'histogramme correspondant aux pixels « eau », et en gris ce correspondant aux pixels « terre »).....	22
Figure 15 : Base de données de valeurs de We en fonction de la distance Lmax (longueur par section).....	23
Figure 16 : Interface graphique du script pour le calcul de la courbe de tarage satellitaire et la distance d'intégration.....	25
Figure 17 : Exemple de courbe de tarage satellitaire, avec le coefficient de détermination.....	26
Figure 18 : Exemple d'une courbe de la distance d'intégration, avec les coefficients de détermination par section.....	26

Figure 19 : Calcul de la variable Wn_DI (largeur nominale incluant le buffer de 10 %, à la distance d'intégration).....	34
Figure 20 : Arbre de décision pour la sélection des courbes croissantes ou décroissantes	35
Figure 21 : Exemple d'utilisation de la fonction « Normplot » de Matlab avec le coefficient de massivité.....	37
Figure 22 : Exemple d'utilisation de la fonction « Normplot » de Matlab après la transformation du coefficient de massivité.....	38
Figure 23 : Exemple de matrice de distribution a) avant transformation et b) après transformation normale.....	38
Figure 24 : Espace canonique pour le modèle exponentiel	46
Figure 25 : Espace canonique pour le modèle polynomial avec a) U1 et U2, b) U1 et U3 et c) U2 et U3	48
Figure 26 : Relation entre Alpha et a) U1 et b) U2	50
Figure 27 : Relation entre Beta et a) U1 et b) U2	51
Figure 28 : Relation entre le paramètre A et U2 pour le modèle polynomial.....	53
Figure 29 : Relation entre le paramètre B et U2 pour le modèle polynomial.....	53
Figure 30 : Relations entre le paramètre C et a) U1 et b) U2.....	54
Figure 31 : Validation du modèle de régionalisation des courbes de tarage exponentielles	59
Figure 32 : Validation du modèle de régionalisation des courbes de tarage polynomiales d'ordre 2.....	59
Figure 33 : Étapes pour le calcul de débits pour un site non-jaugé	61
Figure 34 : Exemple de la classification de Strahler (source : http://help.arcgis.com).....	81
Figure 35 : Exemple de la classification de Shreve (source : http://help.arcgis.com)	82
Figure 36 : Diagramme de la structure de données sur le support physique.	88
Figure 37 : Structure relationnelle de la base de données spatiales <i>Finale</i>	90
Figure 38 : Interface d'interrogation dans ArcMap.....	91
Figure 39 : Boîte de dialogue pour la création d'un nouveau projet	99

Figure 40 : Boîte de dialogue pour l'ajout d'un site au projet	100
Figure 41 : Boîte de dialogue pour le téléchargement des données du réseau hydrographique	101
Figure 42 : Boîte de dialogue pour le traitement du réseau hydrographique	102
Figure 43 : Boîte de dialogue pour l'ajout du chenal principal.....	102
Figure 44 : Boîte de dialogue pour le calcul de la largeur du cours d'eau à la station	103
Figure 45 : Boîte de dialogue pour la création des sections de rivière	104
Figure 46 : Boîte de dialogue pour le téléchargement des données d'élévation.....	105
Figure 47 : Boîte de dialogue pour la génération du bassin versant.....	105
Figure 48 : Boîte de dialogue pour le calcul des paramètres physiographiques.....	106
Figure 49 : Boîte de dialogue pour le calcul de la largeur effective	107

1 Introduction

Les mesures de débit en rivière sont nécessaires pour la planification et la gestion durable des ressources en eau et pour les études sur le climat et l'écologie, entre autres. Toutefois, les stations du réseau hydrométrique canadien sont généralement localisées dans les zones habitées et/ou d'intérêt socio-économique évident. Par conséquent, la densité et la répartition spatiale du réseau hydrométrique ne sont pas optimales pour assurer un suivi opérationnel de l'ensemble des ressources hydriques canadiennes. Ainsi, des débits de rivières couvrant de larges territoires (souvent éloignées et/ou isolées) se retrouvent peu ou pas observés. Faut de moyen, de nombreuses sections de rivières localisées dans les zones habitées ne peuvent être couvertes par le programme d'observation de débit.

La disponibilité croissante de données de télédétection offertes dans différentes régions du spectre électromagnétique, ayant diverses résolutions et étendues spatiales augmente la possibilité de suivre le débit des rivières à l'aide de capteurs spatiaux et aéroportés. L'information hydrométrique produite à partir des données de télédétection permet une couverture régionale/globale et répétitive du territoire. L'estimation du débit de rivières à l'aide de telles approches fournit un moyen de maintenir, sinon d'étendre le réseau d'observation hydrométrique et ce, à des coûts raisonnables.

Une des approches préconisées consiste en l'établissement d'une corrélation entre des paramètres hydrauliques et morphologiques des rivières, produits à partir des données de télédétection et d'observations in situ du débit. Cette méthode nécessite l'emploi d'un grand nombre d'images pour la calibration et la validation du modèle. En ce sens, les données radar offrent un grand potentiel en raison de la sensibilité du signal radar à la présence de l'eau dans la scène en plus de sa capacité à opérer indépendamment des conditions atmosphériques.

L'objectif principal du projet est de développer une méthodologie d'estimation du débit en rivière pour les sites non-jaugés à l'aide de l'imagerie d'archives RADARSAT-1&2 dans l'Est du Canada. La zone d'étude du projet couvre les provinces de Québec, du Nouveau Brunswick, de la Nouvelle-Écosse, de l'Île-du-Prince-Édouard et de Terre-Neuve-et-Labrador. Les images RADARSAT (mode Fin) acquises en eau libre, sans glace ni neige, au-dessus de ces provinces, constituent la première base de données du projet, à laquelle se sont ajoutés des données hydrométriques, physiographiques, climatologiques, et cartographiques des sites et de la zone d'étude. Ces données ont été acquises auprès d'Environnement Canada et ses partenaires provinciaux et du Conseil canadien de la géomatique (GéoBase). Les résultats escomptés du projet sont un modèle d'estimation du débit en sites non-jaugés à l'aide de RADARSAT et un modèle conceptuel de traitements pour la mise en place d'une procédure automatisée de la méthode d'estimation de débit.

Le projet a été structuré en trois étapes principales, telles que listées ci-dessous (Figure 1) :

Étape 1 : Développement d'une base de données d'images RADARSAT et de données hydrométriques, physiographiques, climatologiques, et cartographiques.

Étape 2 : Développement de courbe de tarage ($We-Q$) pour chaque site jaugé.

Étape 3 : Développement du modèle d'estimation de débit pour les sites non-jaugés (régionalisation) et le concept de l'automatisation de la procédure.

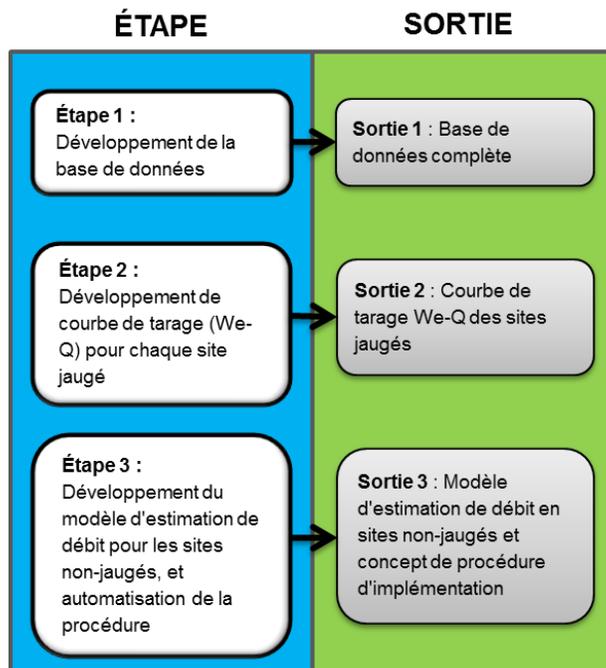


Figure 1 : Étapes du projet et livrables correspondants.

2 Développement de la base de données

La première étape de ce projet était de sélectionner et prétraiter les données nécessaires à la réalisation du projet. Deux types principaux de données ont été utilisées; soit les relevés hydrométriques des rivières (CEHQ, 2003, EC, 2013) et l'imagerie Radarsat.

2.1 Sélection des stations hydrométriques et des images

En ce qui concerne la sélection des stations hydrométriques, trois éléments ont été évalués. D'abord, des critères liés aux stations hydrométriques elles-mêmes, comme :

- la couverture géographique des provinces de Québec, Terre-Neuve et Labrador, Nouveau Brunswick, Nouvelle-Écosse ainsi que l'Île-du-Prince-Édouard;
- la mesure du débit sans interruption lors des deux périodes de disponibilité des images RADARSAT : RS-1 (1997-2013) et RS-2 (2008-2013).

L'application des critères géographiques ainsi que la disponibilité de données de débit a permis d'identifier 332 stations hydrométriques potentielles (

Tableau 1 et Figure 2).

Tableau 1 : Décompte des stations hydrométriques initialement identifiées

Province	Nombre de stations
Québec	172
Terre-Neuve et Labrador	82 (65 TN et 17 L)
Nouveau Brunswick	44
Nouvelle-Écosse	27
Île-du-Prince-Édouard	7
Total	332

Ensuite, des critères associés aux caractéristiques du chenal sont appliqués (avec l'aide de Google Earth), comme :

- la largeur nominale du chenal (minimum 30m, trois pixels radar approximativement);
- l'influence de la marée et/ou de barrages en amont sur le débit du cours d'eau;
- l'influence de l'urbanisme et/ou de la végétation sur les berges du cours d'eau.

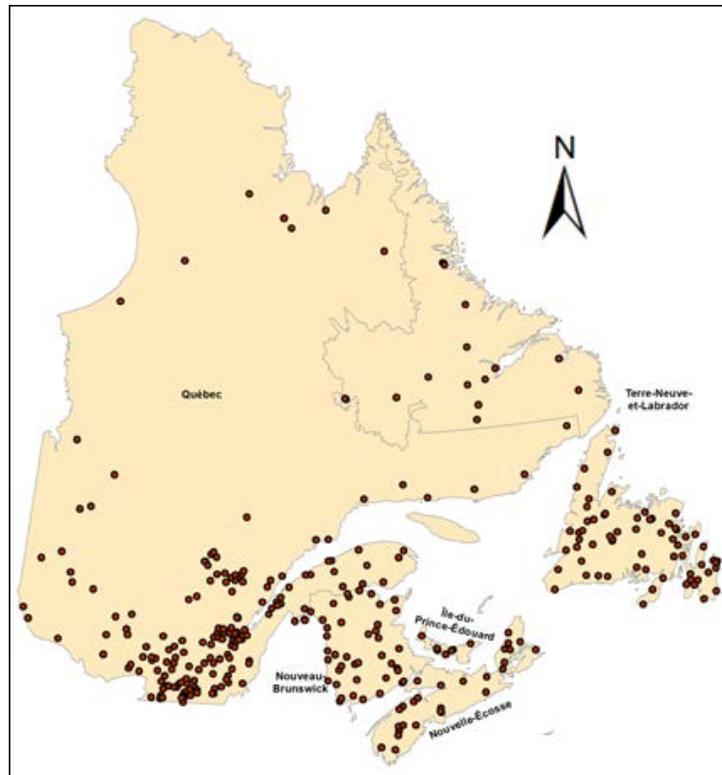


Figure 2 : Stations hydrométriques potentielles

Puis, vu l'énorme quantité d'images RS-1 disponibles en archives (autour de 7000), des critères de sélection d'images ont été appliqués :

- le mode d'acquisition Fin (résolution spatiale 8 m) et la polarisation HH;
- le nombre minimum d'images (RS-1&2) par station :
 - o pour des fins statistiques un minimum de 10 images par station a été fixé pour retenir à la fois une station et ses images. Cependant, afin de garder une bonne répartition de sites expérimentaux sur le territoire, le critère du nombre d'images par station a été abaissé à 4 pour les stations au nord du 50^e parallèle;
 - o les images RS-2 ont été privilégiées étant donné qu'elles sont déjà pré-géoréférencées. Pour une même station, lorsque le nombre d'images RS-2 était supérieur ou égal à 10, la commande RS-1 de cette station a été éliminée;
- l'absence probable de neige et de glace sur les images

Pour ce dernier critère, un algorithme de cartographie d'estimation de la probabilité d'enneigement (Roberge *et al.*, 2010) selon le jour de l'année a été utilisé (Figure 3). Pour chaque station, les dates pour lesquelles la probabilité de la présence neige/glace, au pixel de la station,

tombant en dessous de 5% au printemps et atteignant 50% à l'automne ont été extraites à partir des images de la Figure 3. Seules les images dont la date d'acquisition est comprise entre ces deux dates ont été retenues.

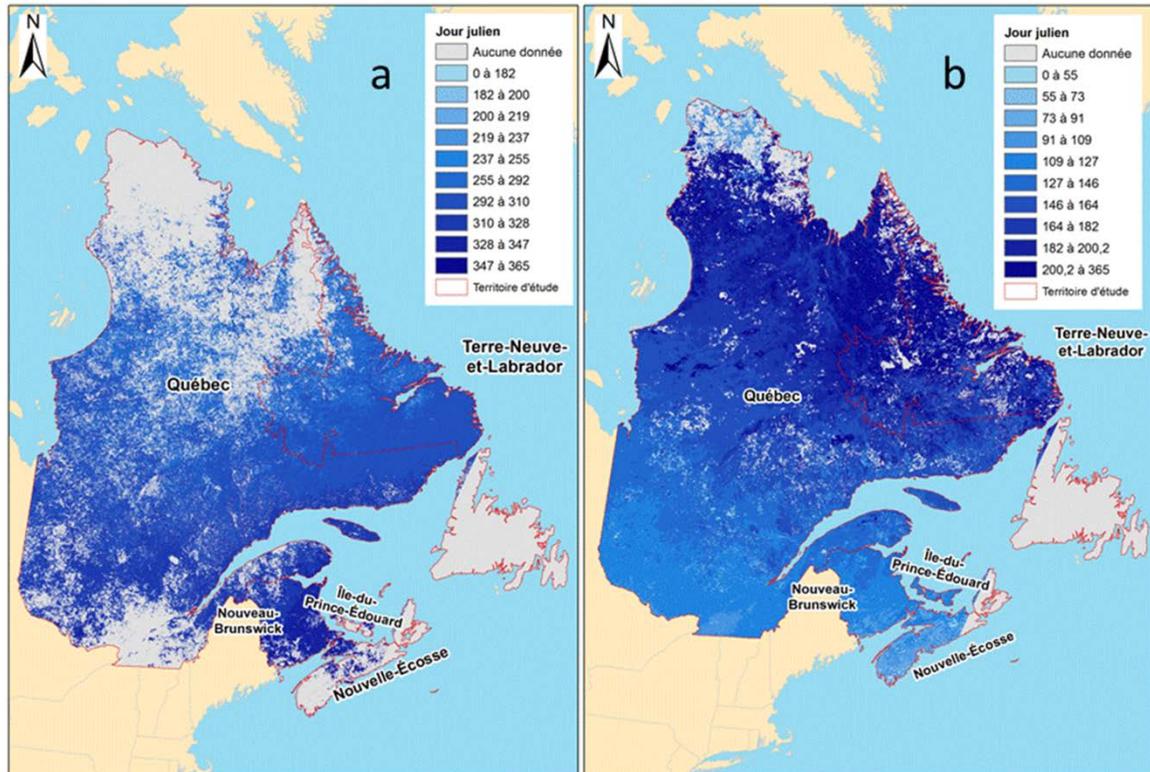


Figure 3 : Probabilité de la présence de neige/glace : (a) 5% printemps et (b) 50% automne (jour julien)

Suite à l'application des critères associés aux caractéristiques du chenal et de sélection des images, il est arrivé pour certaines stations hydrométriques que le nombre d'images RS-1 demeurait très élevé. Par conséquent, un seuil maximal de 30 images par station a été fixé (pour des raisons statistiques et aussi afin de réduire le volume de traitement d'images). Le but était de garder un ensemble d'images de configurations d'acquisition semblables et aussi adapté à l'application prévue (détection de la surface mouillée dans les berges d'un cours d'eau). Un autre critère de sélection visait à ne retenir que les images qui couvraient la portion du cours d'eau en amont de la station (sélection manuelle sur l'écran à l'aide de Google Earth).

Un total de 1 149 images d'archives se sont avérées disponibles (978 RS-1 + 171 RS-2) sur 77 stations hydrométriques. Les images ont été sélectionnées en utilisant l'inventaire complet des données hydrométriques (débits enregistrés) et physiographiques pour les stations et les périodes d'acquisition d'images. Le nombre final de stations hydrométriques retenues par province est présenté au

Tableau 2, de même que le nombre d'images par station au Tableau 3. La Figure 4 montre leur distribution spatiale en fonction du nombre et du type d'images correspondants.

Tableau 2 : Décompte des stations hydrométriques retenues par province

Province	Nombre de stations
Québec	54
Terre-Neuve et Labrador	14 (6 TN et 8 L)
Nouveau Brunswick	7
Nouvelle-Écosse	2
Île-du-Prince-Édouard	0
Totale	77

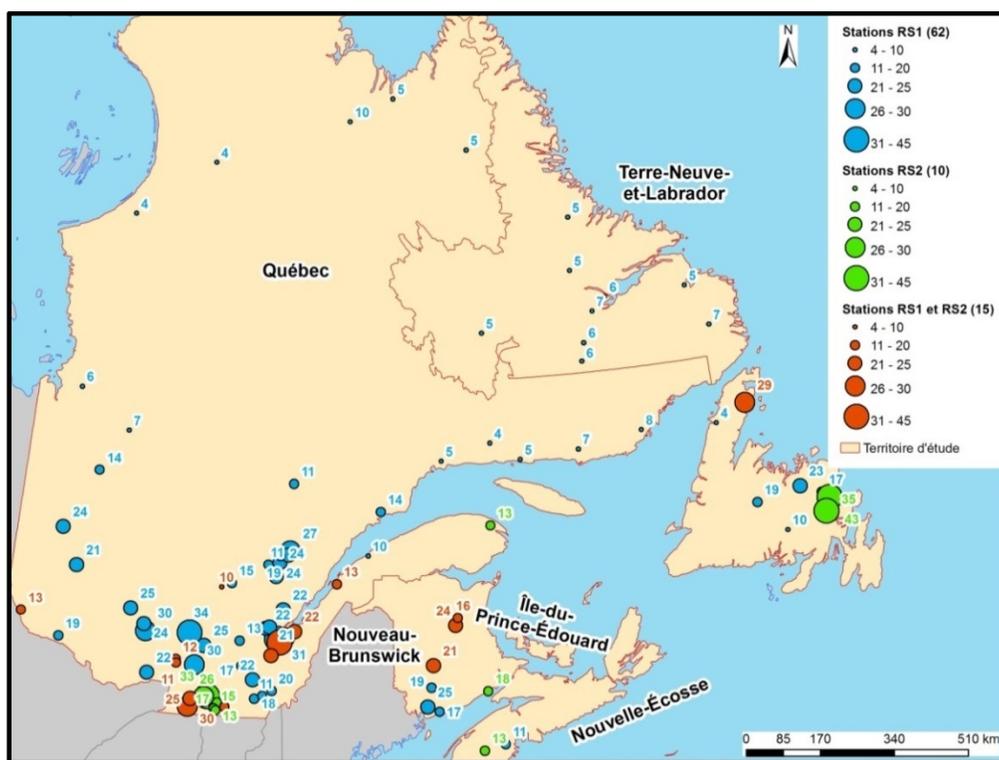


Figure 4 : Distribution spatiale des stations hydrométriques retenues

Tableau 3 : Liste de stations hydrométriques retenues, et nombre total d'images RS-1&2 par station

	ID_ENV_CAN	Province	RS-1	RS-2
1	01AL002	NB	20	1
2	01AM001	NB	19	0
3	01AQ001	NB	17	0
4	01AQ002	NB	25	0
5	01BP001	NB	22	2
6	01BQ001	NB	15	1
7	01BV006	NB	0	17
1	01ED005	NE	0	16
2	01EF001	NE	11	0
1	01BH005	Qc	0	14
2	02AQ002	Qc	10	0
3	02JB009	Qc	21	0
4	02KJ004	Qc	19	0
5	02LC021	Qc	9	2
6	02LD005	Qc	22	0
7	02LE024	Qc	23	0
8	02LE025	Qc	30	0
9	02LG005	Qc	25	0
10	02ND003	Qc	9	1
11	02NE011	Qc	11	0
12	02NF003	Qc	34	0
13	02OA054	Qc	19	11
14	02OA057	Qc	25	5
15	02OB017	Qc	30	1
16	02OC021	Qc	25	0
17	02OD001	Qc	22	0
18	02OD003	Qc	17	0
19	02OE005	Qc	13	0
20	02OE027	Qc	11	0
21	02OE032	Qc	20	0
22	02OG002	Qc	0	17
23	02OG014	Qc	17	6
24	02OH008	Qc	0	14
25	02OH013	Qc	0	13
26	02OJ007	Qc	0	26
27	02OJ026	Qc	0	18
28	02PA007	Qc	13	0
29	02PC002	Qc	21	0
30	02PC010	Qc	22	0
31	02PD004	Qc	18	0
32	02PG001	Qc	8	5
33	02PH010	Qc	16	2
34	02PH011	Qc	26	0
35	02PJ005	Qc	30	1
36	02PJ007	Qc	25	0
37	02PL005	Qc	19	1
38	02RB004	Qc	11	0
39	02RH027	Qc	24	0
40	02RH035	Qc	11	0
41	02RH045	Qc	27	0
42	02RH066	Qc	23	0
43	02UC002	Qc	5	0
44	02VB004	Qc	4	0
45	02VC001	Qc	5	0
46	02WB003	Qc	7	0
47	02XA008	Qc	8	0
48	03AB002	Qc	14	0
49	03BD002	Qc	4	0
50	03BF001	Qc	6	0
51	03ED001	Qc	4	0
52	03FA003	Qc	4	0
53	03LF002	Qc	2	0
54	04NA001	Qc	24	0
1	02XA003	TN	6	0
2	02YC001	TN	4	0
3	02YD002	TN	22	0
4	02YO011	TN	17	0
5	02YQ001	TN	21	0
6	02YS005	TN	0	40
7	02ZE004	TN	10	0
8	03NF001	TN	5	0
9	03OC003	TN	5	0
10	03OE001	TN	7	0
11	03OE003	TN	6	0
12	03PB002	TN	5	0
13	03QC001	TN	5	0
14	03QC002	TN	5	0
77				

Les mesures de débits instantanés associés aux dates et heures d'acquisitions d'images RS-1&2 sont disponibles pour les 77 stations hydrométriques retenues dans un intervalle de +/- 1.5h. Toutes les stations ont des enregistrements du débit quotidien associés aux dates d'acquisitions d'images.

2.2 Base de données physiographiques, climatiques et cartographiques

Des données de diverses sources ont été utilisées dans le cadre de l'étape I du présent projet. Ces données ont parfois été utilisées telles quelles et parfois elles ont permis la création de jeux de données dérivées. La section suivante présente les familles de données utilisées. La majorité de ces données ont été acquises pour être utilisées comme variables explicatives pour la régionalisation des courbes de tarage (voir l'étape III). Certaines d'entre-elles ont également été utilisées pour définir le masque de la rivière lors de la procédure de développement de courbes de tarage pour les sites jaugés (étape II) et la validation du modèle d'estimation du débit pour les sites non-jaugés (étape III). Le détail du calcul des variables physiographique est disponible à l'annexe I. Ces données seront copiées sur un support physique (disque dur externe) et accompagneront ce rapport. L'organisation des données est détaillée à l'annexe II.

2.2.1 Données brutes

Données climatologiques

Le jeu de données climatiques mensuelles à maillage régulier de 10 km pour les années 1901 à 2010 en Amérique du nord a été mis au point par le service national d'information sur les terres et les eaux (SNITE, Agriculture et Agroalimentaire Canada). Ce jeu de données contient les emplacements maillés ponctuels de température maximale (°C), de température minimale (°C) et des précipitations (mm) pour l'Amérique du nord continental. Les grilles sont disponibles en format ASCII géoréférencée (.asc). Ces modèles spatiaux ont été élaborés au moyen des algorithmes de splines de lissage «plaque mince» du progiciel ANUSPLIN, qui est en fait une approche mathématique très élaborée permettant de produire des cartes climatiques à différentes échelles spatiales et temporelles (RNCAN, 2013).

Données hydrométriques

Les données hydrométriques pour la période 1997 à 2013 (débit instantané) correspondantes à la province du Québec ont été acquises auprès du Centre d'expertise hydrique du Québec (CEHQ). Quant aux stations situées dans les provinces de l'Atlantique, elles ont été obtenues via Environnement Canada. Les données ont été regroupées tout d'abord par province dans un répertoire. Ensuite, chaque répertoire de station a été identifié par le code d'identification fédéral de la station. Les fichiers textes correspondant à chaque année d'enregistrement de débits instantanés ont été stockés dans chacun de ces répertoires de stations. Les débits mesurés au moment de l'acquisition des images RS-1&2 ont été compilés et se retrouvent dans la table Stations_Images_RS1_RS2 de la base de données.

Cette table de correspondance a été créée à partir du croisement des dates et heures d'acquisition des images RS commandées (RS-1&2), avec les débits instantanés mesurés sur les rivières figurant sur ces images. Les heures GMT d'acquisition des images ont été converties en

heures locales selon le fuseau horaire de chaque station. Ces données ont été utilisées lors de la calibration des courbes de tarage (instantanées et quotidiennes) et ont servi à la validation du modèle d'estimation du débit pour les sites non-jaugés (instantanées).

Réseau hydrographique

Le réseau hydrographique de la partie canadienne du projet provient du Réseau Hydro National (RHN). Les jeux de données RHN (RNCAN, 2015b) sont produits et distribués par unité de travail RHN correspondant à une aire de drainage. Il en est de même pour les données provenant des États-Unis, où une partie du *National Hydrography Dataset* fut téléchargée. Au total, le Québec et les provinces de l'Atlantique touchent à environ 340 unités de travail. Ces unités se déversant parfois les unes dans les autres, un travail a dû être entrepris afin de déterminer les endroits où cela se produisait, quelles étaient les unités impliquées et quel élément du réseau d'écoulement permettait de faire le lien entre ces dernières. Ce travail était nécessaire afin de pouvoir remonter le réseau hydrographique en amont des stations hydrométriques. Ces données ont permis de définir le masque pour la calibration des courbes de tarage et pour en dériver des variables physiographiques pour la régionalisation des courbes de tarage.

Données d'élévation

Les données d'élévation téléchargées sont les données numériques d'élévation du Canada (DNEC) à l'échelle de 1:50000 (RNCAN, 2015a). Ces données sont distribuées en feuillets du Système national de référence cartographique (SNRC). Le territoire canadien couvert par le projet correspondait à environ 2600 feuillets que nous avons regroupés. À cela s'ajoute environ 50 tuiles du côté des États-Unis, lesquelles proviennent du *National Elevation Dataset*, le principal produit d'élévation de l'USGS. Le regroupement des cellules et tuiles s'est fait dans une base de données spatiales d'ESRI, laquelle contient les informations sur chacune des cellules et non les données d'élévation proprement dites. Ceci a permis de produire une mosaïque à la volée et ainsi éviter de rassembler toutes les données dans un seul fichier de taille énorme et difficile à utiliser. Des fonctions, comme le calcul de la pente ou de l'orientation, peuvent être appliquées dans le processus de production de la mosaïque à la volée de sorte que ces informations n'ont pas à être produites dans des fichiers séparés. Les données d'élévation ont été utilisées pour la correction géométrique des images RS-1 et RS-2. Elles ont été également utilisées pour le calcul de différents paramètres physiographiques associés à chaque station hydrométrique.

Données d'utilisation du sol (NALCMS)

La carte d'occupation des sols de l'Amérique du nord (année 2005, résolution de 250m) a été produite dans le cadre du North American Land Change Monitoring System (NALCMS), un effort

trilatéral entre le Centre canadien de télédétection (CCT), le USGS et trois organisations mexicaines (USGS, 2005). Le produit a été généré à partir des observations acquises par le satellite MODIS/Terra (Moderate Resolution Imaging Spectroradiometer). Chacune des sept bandes spectrales particulières à l'observation de la Terre a été transformée en réflectivité en haut de l'atmosphère à une résolution temporelle de 10 jours et une résolution spatiale de 250m. Ces données ont servi pour le calcul du pourcentage des classes d'utilisation du sol contenues à l'intérieur des bassins versants générés à partir de la station hydrométrique. Elles ont été utilisées comme variables explicatives pour la régionalisation des courbes de tarage.

Données de couverture forestière (EOSD)

La carte de couverture forestière pour le Canada (Observation de la Terre pour le Développement Durable des Forêts) a été créée en 2002 et publiée en 2008 par le Service Canadien des Forêts et l'Agence Spatiale Canadienne (RNCAN, 2008). Le produit a été conçu à partir de l'imagerie Landsat avec une résolution spatiale de 25m. Le produit est distribué par la plate-forme SAFORAH. Ces données ont servi pour le calcul du pourcentage de différentes classes d'occupation du sol contenues à l'intérieur de la zone de calcul pour la calibration des courbes de tarage. Elles ont été utilisées comme variables explicatives pour la régionalisation des courbes de tarage.

Réseau routier

Les données relatives au réseau routier proviennent du réseau Routier National (RNCAN, 2015c) disponible sur le portail Geobase.ca. Ce jeu de données est distribué par provinces. Il a été utilisé pour la validation de la correction géométrique des images RS-1 et RS-2.

Images RADARSAT

Les images RS1 & 2 commandées se retrouvent dans le dossier Brute. Les images brutes ont ensuite été orthorectifiées en lots en utilisant les données d'élévation décrites précédemment et le modèle mathématique accompagnant chaque image. Suite à ces étapes, certaines images jugées insatisfaisantes (<10%) après une évaluation visuelle ont été orthorectifiées de nouveau en utilisant des points de contrôle définis manuellement. L'ensemble des images se retrouve dans le dossier Ortho. Les orthoimages ont été traitées par le passage d'un filtre anti-chatoiement et se retrouvent dans le dossier Filtrees. Les résultats de la procédure d'orthorectification des images sont détaillés à la section suivante.

2.2.2 Données dérivées

Bassins versants

Le bassin versant a été délimité de façon automatique en s'appuyant sur la direction de l'écoulement déterminé par le modèle numérique de terrain et en le contraignant avec le réseau hydrographique filamenteux en amont de la station hydrométrique et les limites de l'unité de travail. Par la suite, chaque bassin versant a été vérifié et corrigé de façon manuelle si nécessaire. Les erreurs corrigées étaient causées principalement par deux facteurs : l'imprécision du modèle numérique de terrain et l'inexactitude du réseau hydrographique. Dans le premier cas, le modèle numérique de terrain ayant une résolution verticale de 1m, et une précision variable pouvant être aussi peu que 10m, les zones de faible pente peuvent être mal représentées dans la matrice d'écoulement. Cette situation cause généralement une sous-estimation de la taille du bassin versant. Dans le deuxième cas, l'inexactitude du réseau hydrographique a principalement causée une surestimation de la taille des bassins versants en considérant dans le bassin des zones s'écoulant vers des tronçons de rivière n'appartenant pas réellement au bassin.

Au final, une proportion plus ou moins grande des limites de chaque bassin provient de la délimitation automatique. Comme le processus automatique repose sur le modèle numérique de terrain qui est au format matriciel, la limite originale du bassin suit le contour des pixels et a une apparence en escalier. Cette structure peut augmenter considérablement la longueur du périmètre si elle n'est pas prise en considération (l'augmentation peut atteindre plus de 40%). Pour éviter ce problème et uniformiser la situation des bassins versants ayant été délimités entièrement automatiquement et celles des bassins délimités de façon plus manuelle, un lissage du contour a été effectué, éliminant ainsi l'effet d'escalier et donnant une apparence plus réelle à la limite.

Les bassins versants associés à chaque station hydrométrique (Figure 5) se retrouvent dans des classes d'entités différentes nommées selon le code de la station mais sont regroupées dans un même jeu de classe d'entité nommé *BV*. Ces bassins versants ont servi comme zone de calcul pour certaines variables physiographiques (annexe I)

Cours d'eau principal

Le cours d'eau principal en amont de chaque station a été déterminé en suivant les tronçons d'ordres les plus élevés. Ce chenal principal est représenté par une polyligne partant de la station et s'étendant jusqu'à l'amont du bassin versant. Le cours d'eau principal de chaque station est enregistré dans une classe d'entité différente nommée en utilisant le code de la station hydrométrique. Tous les cours d'eau principaux se retrouvent dans le jeu de classe d'entité *Main*. Le calcul du cours d'eau principal est détaillé dans la section intitulée courbe de tarage satellitaire.

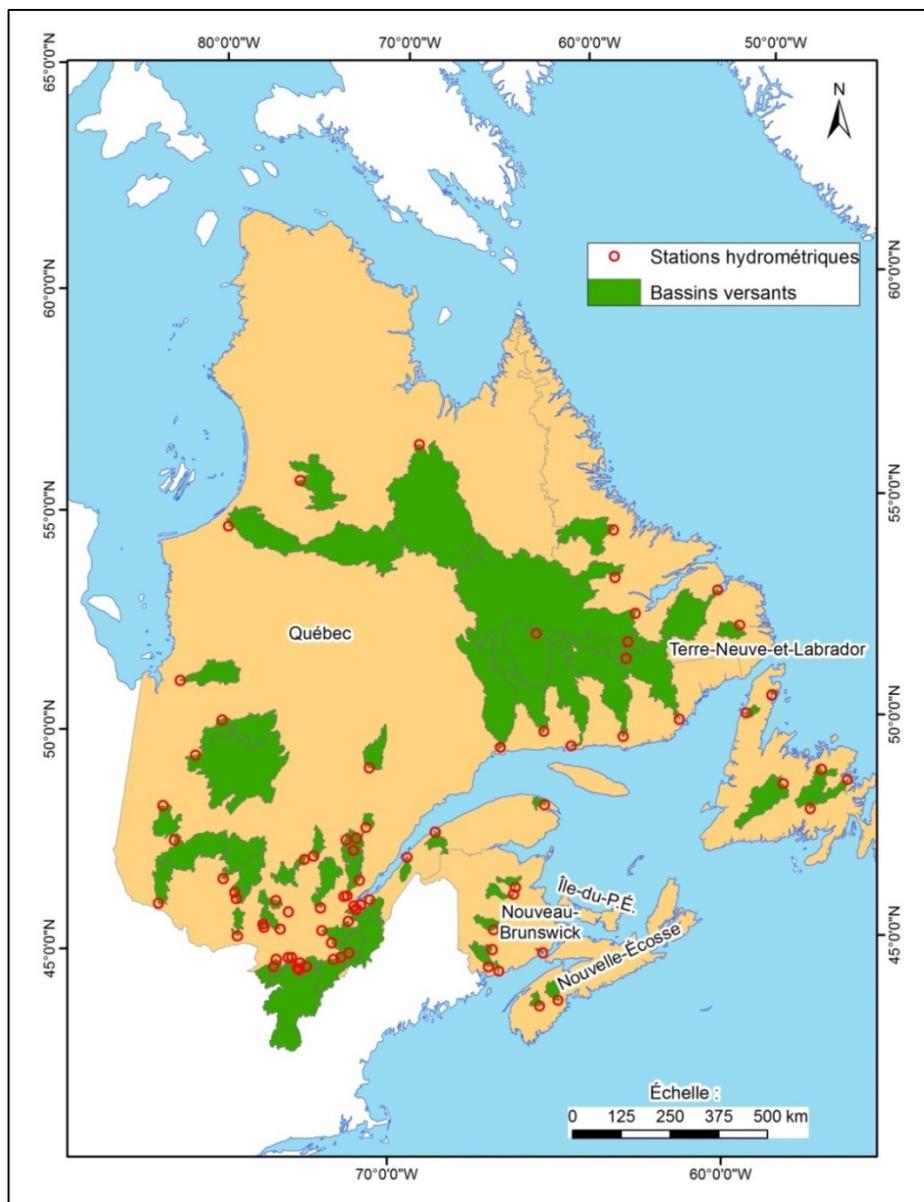


Figure 5 : Distribution des 77 bassins versants des sites jaugés retenus.

Sections

En utilisant conjointement le cours d'eau principal sous forme polyligne et la représentation polygonale du cours d'eau en amont de chaque station, il a été possible de séparer le polygone en tronçons d'une longueur constante de 1km. Le nombre de sections est variable d'une station hydrométrique à l'autre et dépend de la taille de la rivière. Les sections relatives à deux stations différentes se retrouvent dans deux classes d'entités différentes mais sont regroupées dans le même jeu de classes d'entités *Sections*.

2.3 Traitements des images RS-1&2

Les images RS-1 et RS-2 ont été orthorectifiées dans le logiciel Geomatica en utilisant leurs paramètres orbitaux ainsi qu'un modèle numérique de terrain. La procédure a été préalablement automatisée au moyen de scripts en langage Python. La qualité de l'orthorectification a été évaluée manuellement.

Pour l'évaluation, chaque image fut affichée dans le logiciel ArcGIS avec les couches suivantes correspondantes à sa région: routes, réseau hydrographique, station hydrométrique, masque du corridor d'analyse (sections de 1 km). La précision en X et Y du positionnement de l'image, en nombre de pixels, est évaluée près de la station hydrométriques et au niveau de la 5^{ème} section.

Lorsque l'erreur de positionnement observée est de 5 pixels ou moins en X et en Y, l'image est acceptée car le décalage se situe à l'intérieur du corridor de recherche des pixels d'eau (min. 70 m). Lorsque l'erreur de positionnement observée est de plus de 5 pixels en X ou en Y, l'image est recorrectée manuellement en ajoutant des points de contrôle, jusqu'à ce que l'erreur de positionnement soit de 5 pixels et moins en X et Y.

Les résultats de la vérification ainsi que les commentaires associés aux images ont été comptabilisés par stations hydrométriques. Le pourcentage d'efficacité de la correction automatique pour les images RS2 était de 93% alors que pour les RS1, il se situait autour de 80%.

Par la suite, des tests de filtrage ont été effectués sur les images correctement orthorectifiées, afin d'améliorer la performance de la procédure automatisée d'extraction de pixels mouillés. La taille de la fenêtre a été fixée à 3 par 3 pixels pour tous les filtres testés. Les filtres évalués, disponibles dans Geomatica, étaient : Enhanced Frost SAR Speckle Filter (FEFROST), Enhanced Lee SAR Speckle Filter (FELEE), Frost SAR Speckle Filter (FFROST), Gamma SAR Speckle Filter (FGAMMA), Kuan SAR Speckle Filter (FKUAN), Laplacian SAR Speckle Filter (FLAP), Lee SAR Speckle Filter (FLE), SAR Speckle Filter (FSPEC), et Standard Deviation SAR Speckle Filter (FSTDDEV).

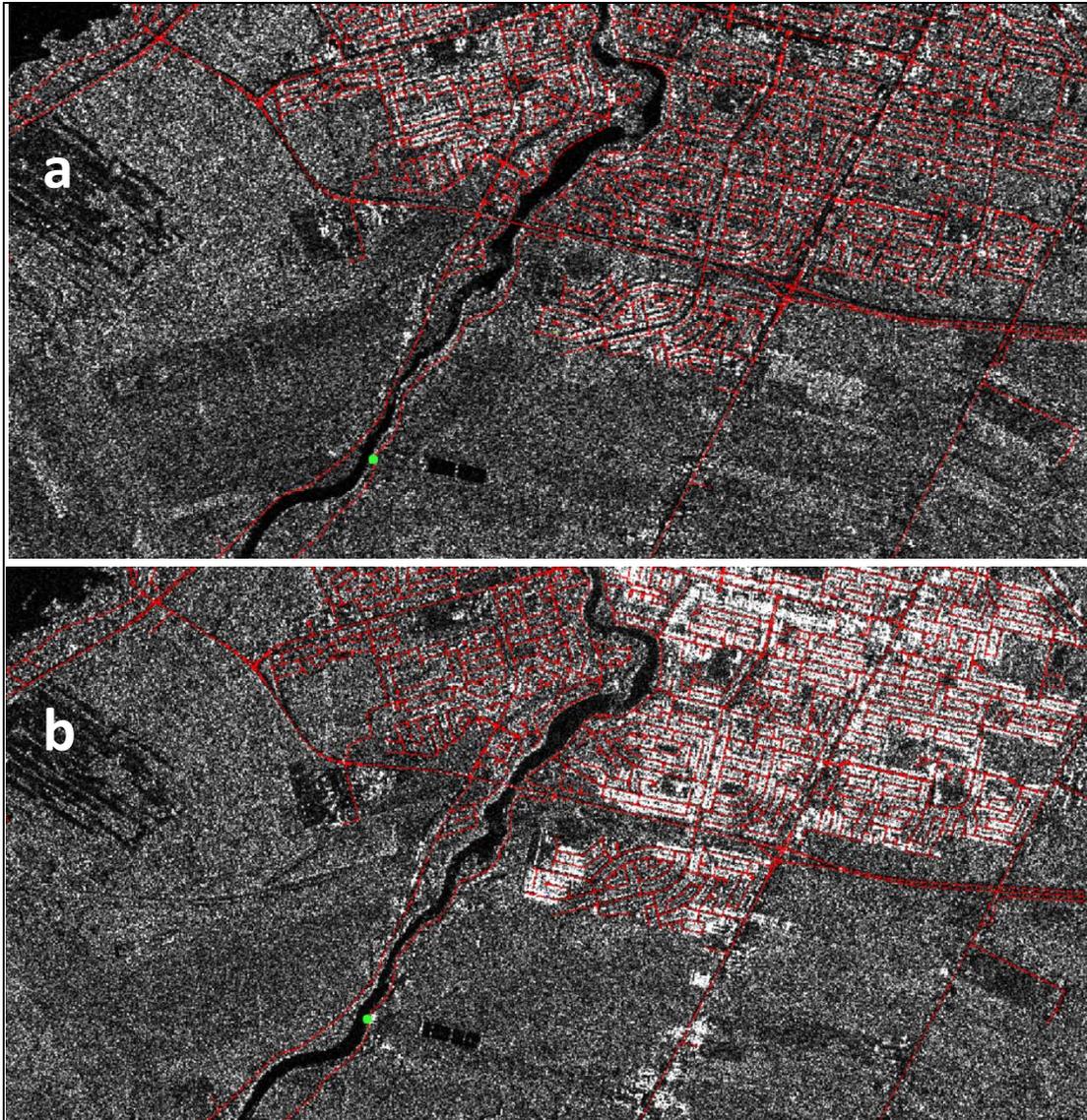


Figure 6 : Exemple de validation d'une orthoimage RS-2 (station Chateauguay, échelle approximative 1/35 000) : a) à corriger (DK_371116), b) bien orthorectifiée (DK_371266)

Le filtre de chatoiement de Lee a été retenu. Il donne en effet un meilleur rapport de réduction du chatoiement et d'augmentation du contraste. Un exemple de l'application des filtres « Lee SAR Speckle » (FLE) et « Enhanced Frost SAR Speckle » (FEFROST) sur une orthoimage RS-2 de la rivière Richelieu est présenté à la Figure 7 .

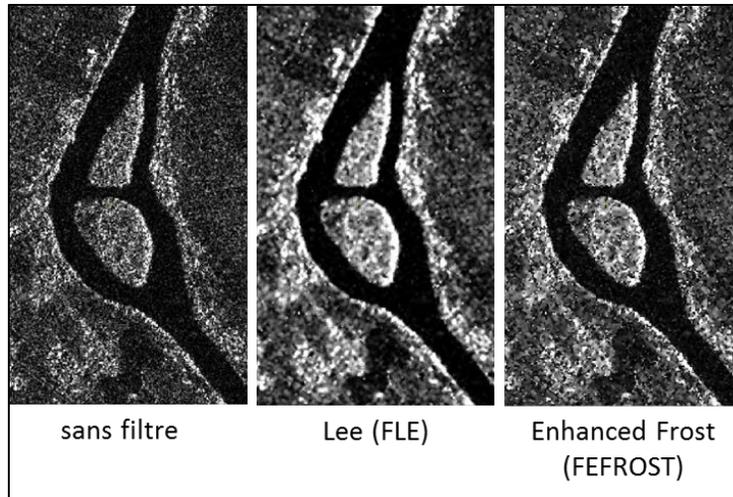


Figure 7 : Exemple d'application des filtres de Lee et Enhanced Frost, par rapport à une orthoimage sur la rivière Richelieu non filtrée

Le décompte du travail effectué sur les images RS-1&2 est détaillé au Tableau 4, de même que le nombre de paires « image-débit » établis pour la totalité de stations hydrométriques sélectionnées.

Tableau 4 : Décompte des images d'archives RS-1&2 prétraitées.

Images d'archives	RS-1	RS-2	Paires RS-1	Paires RS-2	Paires totales
Images commandées	1015	323	-	-	-
Images corrompues (incapable d'être lues)	-13	-	-	-	-
Images corrompues (incapable d'être orthorectifiées)	-3	-	-	-	-
Images reçues (nombre définitif)	999	323	-	-	-
Après le choix de 77 stations	962	171	2007	285	2292
Après l'élimination des images identiques recadrées	928	164	1112	244	1356
Après l'élimination des paires d'images sans débit	872	153	1018	227	1245
Après l'élimination des images ne couvrant pas complètement minimalement la première section	866	152	1008	225	1233

3 Élaboration de la courbe de tarage satellitaire

L'élaboration de la courbe de tarage locale s'est effectuée en plusieurs étapes (Figure 8). D'abord, elle a débutée avec l'analyse de la probabilité d'appartenance à l'eau (à partir d'orthoimages RS 1&2 sélectionnées préalablement), suivi par le calcul de la largeur effective (We) à chacune des sections du cours d'eau analysées. Ensuite, la largeur effective calculée a été associées au débit instantané de la station correspondant à la date et l'heure d'acquisition d'une image. Ces couples largeurs effectives-débits, associés à une distance d'intégration donnée ont été stockés dans une base de données qui a servi pour la construction des courbes de tarage locales. Les procédures d'extraction des pixels d'eau et le calcul de la largeur effective ont été automatisés (Figure 9 : Diagramme de la méthode automatisée de calcul de la probabilité d'appartenance à l'eau par logique floue Figure 9).

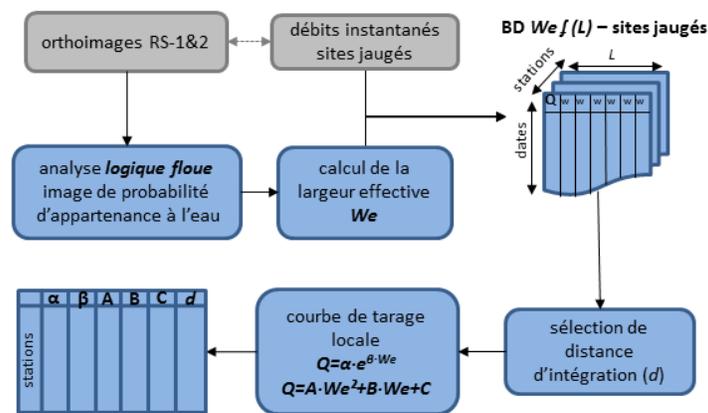


Figure 8 : Diagramme de la méthode de calcul de la courbe de tarage locale

L'obtention d'une courbe de tarage locale s'est déroulée en 5 étapes : soit l'obtention du masque pour l'application du calcul de la probabilité d'eau, l'intercalibration des orthoimages, paramétrage de fonction d'appartenance de l'eau, le calcul de la largeur effective et enfin, l'ajustement d'une fonction pour la courbe de tarage.

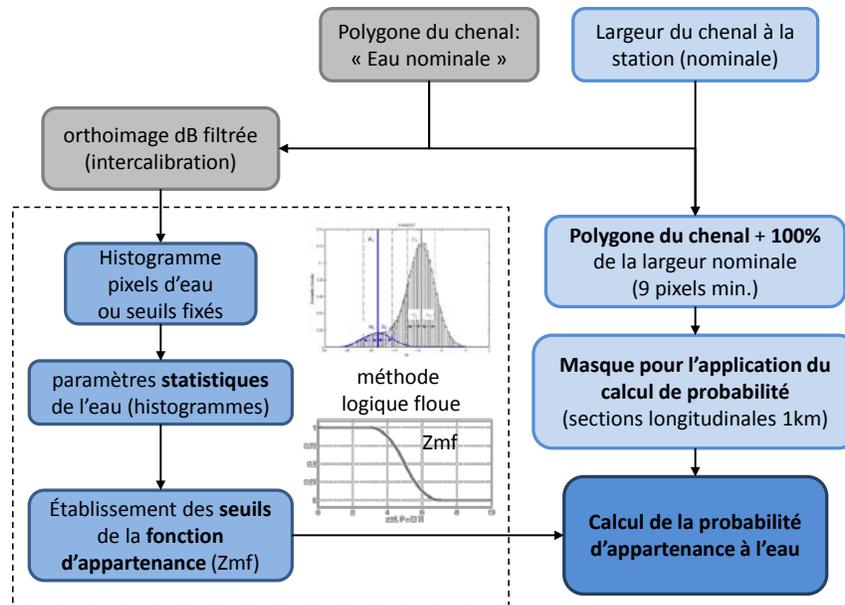


Figure 9 : Diagramme de la méthode automatisée de calcul de la probabilité d'appartenance à l'eau par logique floue

3.1 Obtention du masque pour l'application du calcul de probabilité

La première étape consistait à définir un polygone servant de masque pour la délimitation de la zone de calculs sur les images. La méthodologie utilisée pour délimiter le chenal principal et le masque pour le calcul de la probabilité est schématisée à la Figure 10. Pour chaque station hydrométrique (Figure 10.1), l'élément du réseau hydrographique situé le plus près a été détecté (Figure 10.2). La partie du réseau se situant en amont de la station a été déterminée en utilisant la direction d'écoulement spécifiée dans les données hydrographiques, (Figure 10.3). L'ordre de Shreve a été ensuite calculée sur cette partie du réseau (Figure 10.4).

Le chenal a été déterminé en remontant le réseau hydrographique à partir du tronçon le plus près de la station et en poursuivant, à chaque intersection rencontrée, avec le tronçon d'ordre le plus élevé (Figure 10.5). Ensuite, un polygone a été créé en sélectionnant le ou les polygones du jeu de données hydrographiques représentant en totalité ou en partie le chenal principal (Figure 10.6). Les parties du polygone correspondantes à des tributaires ont été coupées puis supprimées.

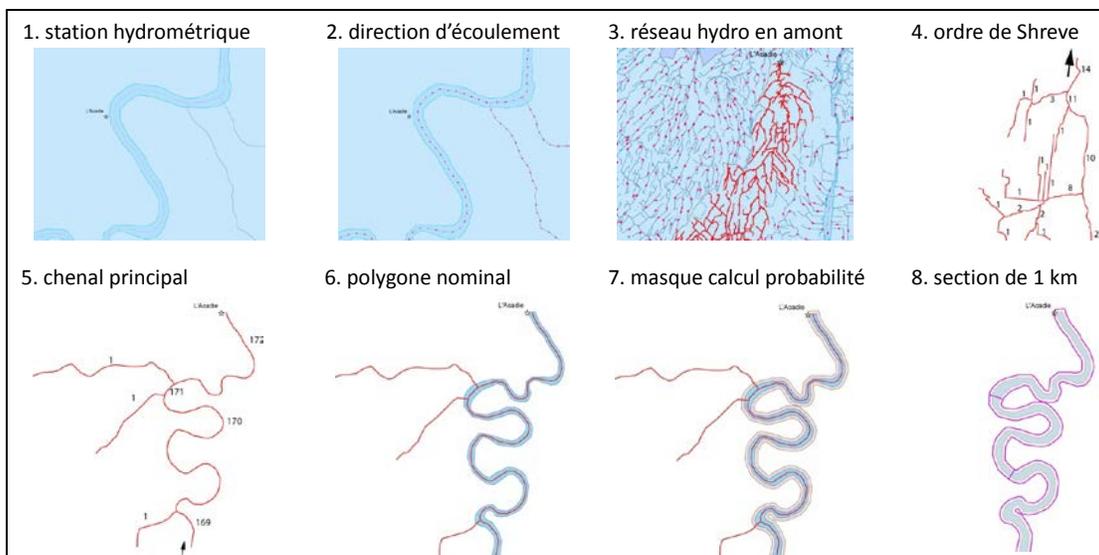


Figure 10 : Obtention du masque pour l'application du calcul de la probabilité d'appartenance

Par la suite, le polygone résultant a été agrandi en créant une zone tampon (masque pour le calcul de la probabilité, Figure 10.7). Cette zone tampon correspondait à 100% de sa largeur nominale à la station (Figure 11). Le nouveau polygone a été forcé à avoir une largeur de neuf pixels minimalement (environ 70m). La zone tampon ajoutée au polygone nominal du chenal a permis de compenser le décalage possible entre les polygones extraits du RHN et les cours d'eau présents dans l'orthoimage.

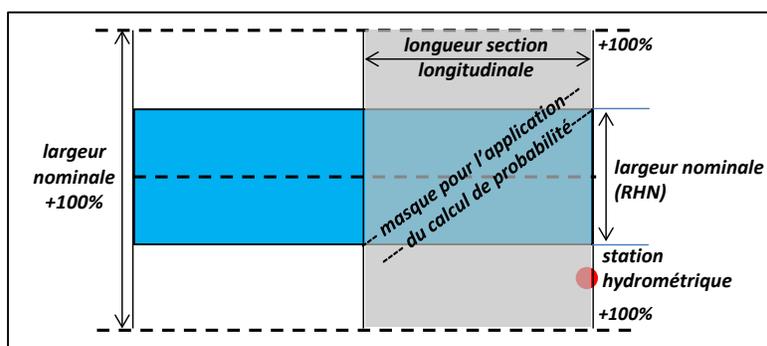


Figure 11 : Ajout d'une zone tampon au polygone du chenal principal

Le polygone a été finalement coupé en sections de 1km (Figure 10.8), au long de la ligne du chenal principal. Ce polygone a formé le masque du calcul de la probabilité.

3.2 Intercalibration des orthoimages

Étant donné que ce projet a eu recours à des images d'archives acquises dans différentes configurations telles que les angles d'incidence, l'orbite, le type satellite Radarsat et les dates depuis le lancement (calibration du gain de l'antenne), il s'est avéré difficile de comparer pour différentes images les signatures Radar de l'eau. Pour pallier à cette lacune, une procédure

intitulée intercalibration des orthoimage a été introduite. L'intercalibration visait en soit à faire concorder les histogrammes de signatures radar de l'eau, extraits à l'intérieur du masque définit à l'étape précédente, à une valeur de rétrodiffusion de référence. La signature de l'eau extraite a été convertie en échelle de décibel (dB) afin d'obtenir une distribution normale. La procédure a utilisé le premier percentile comme valeur de référence des histogrammes afin d'éviter l'intercalibration sur des valeurs étant éloignées ou aberrantes (effet de bordures). Le premier percentile des histogrammes a été ramené à -23 dB, une valeur moyenne de percentiles d'histogramme (rivière Yamaska). La différence entre -23 dB et le premier percentile est additionnée aux pixels extraits, ce qui a pour effet de ramener les histogrammes à une base commune soit le premier percentile à -23 dB (Figure 12). La valeur de -23 dB aurait pu être totalement arbitraire et la procédure aurait tout de même fonctionnée, toutefois, il était intéressant de conserver une valeur nominale représentative de ce qui peut être observé sur les histogrammes. Cette procédure d'intercalibration s'est avérée indispensable dans le calcul de la largeur effective pour l'obtention de rapports entre les valeurs de largeurs effectives et les débits instantanés.

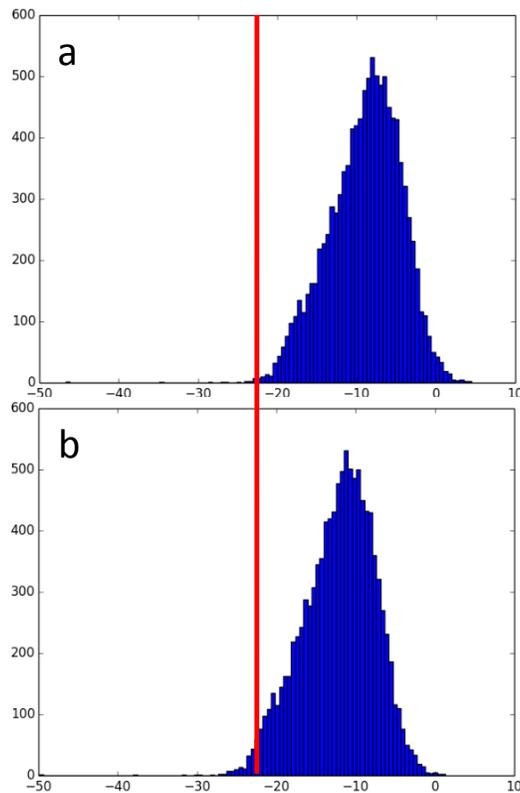


Figure 12 : Histogrammes de la rétrodiffusion avant (a) et après (b) l'intercalibration en utilisant le 1er percentile et la valeur -23 dB comme référence

3.3 Paramétrage de la fonction d'appartenance de l'eau

Pour calculer la probabilité d'un pixel d'eau, une fonction d'appartenance de type logique floue nommée « Z-shaped » (*Zmf built-in membership function*, MATLAB) a été appliquée sur les images intercalibrées à l'intérieur de la zone de calcul. Cette fonction est définie aux équations (1) et (2) :

$$y = zmf(x, [a \ b]) \quad (1)$$

$$f(x; a, b) = \begin{cases} 1, & x \leq a \\ 1 - 2 \left(\frac{x - a}{b - a} \right)^2, & a \leq x \leq \frac{a + b}{2} \\ 2 \left(\frac{x - b}{b - a} \right)^2, & \frac{a + b}{2} \leq x \leq b \\ 0, & x \geq b \end{cases} \quad (2)$$

où x correspond à la valeur de la rétrodiffusion HH (σ^0) du pixel exprimée en décibels, a et b sont les paramètres qui fixent les extrémités de la courbe (Figure 13).

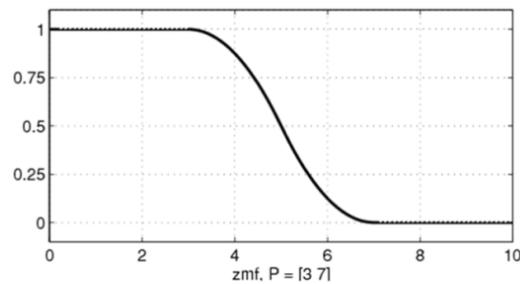


Figure 13 : Exemple d'une Fonction d'appartenance logique floue « Z » (fuzzification).

Puisque le masque pour le calcul de la probabilité correspond au polygone de la rivière plus 100% de sa largeur de part et d'autre, ceci laisse supposer la présence de pixels terrestres et d'eau dans les histogrammes faits à partir des valeurs des pixels sous le masque. Ainsi, deux pics devraient être visibles sur les histogrammes (Figure 14) : soit celui des pixels d'eau (à la gauche) et celui des pixels terrestres (à la droite). La distribution serait dite bimodale. Le contraste entre l'eau et le terrestre (non-eau) explique l'existence de ces deux pics. La zone de chevauchement centrale des deux pics représente la proportion des pixels mixtes (gris). Il a été remarqué que lorsque la rivière avait une largeur de moins de 40m, cette proportion de pixels mixtes augmentait et cachait les deux pics visibles pour ainsi former une distribution unimodale. En effet, en utilisant une zone tampon d'une largeur minimale de 70m pour un cours d'eau de faible largeur, la proportion des pixels correspondante à la surface « terre » est beaucoup plus importante que celle des pixels correspondante à la surface « eau », ce qui cache l'histogramme « eau ». Les images des rivières à faibles débits favorisent les histogrammes unimodaux. Pour l'ajustement des seuils A et B, une approche adaptative à laquelle A égalait le 99^e percentile de l'histogramme de l'eau et B le 1^{er} percentile de l'histogramme de terre avait été envisagée. Pour

réaliser cette approche, deux courbes de distributions distinctes étaient ajustées sur l'histogramme bimodal en utilisant des fonctions de maximum de vraisemblance. Toutefois, le fait d'avoir des histogrammes unimodaux, ce qui touchait près du tiers des images, rendait cette approche inapplicable. La solution choisie a été de fixer les valeurs de A et B. Pour fixer A, l'ensemble des histogrammes bimodaux a été utilisé pour ainsi en calculer la moyenne du 99^e percentile de l'eau à l'intérieur des 5 premières sections de rivières. Les 5 premières sections ont été choisies pour s'assurer que la proportion des pixels d'eau soit semblable à celle à la station. Ainsi le seuil A a été fixé à la valeur moyenne de -15 dB. La valeur du seuil B a été forcée à -14,5 dB. La raison de ce choix est que les histogrammes de pixels terrestres présentaient une plus grande variabilité de rétrodiffusion, causée par l'hétérogénéité des cibles terrestres. L'utilisation d'une valeur statistique comme la moyenne du 1^{er} percentile devenait incohérente avec la valeur fixée pour A. Les images des rivières présentant des histogrammes unimodaux n'ont pas servi à la définition du seuil A, cependant elles ont pu être utilisées tout de même en leur appliquant la fonction ZMF avec les valeurs de A et B déterminées. Des essais ont démontré que les meilleures corrélations entre le débit et les largeurs effectives (calculées à l'étape suivante) ont été obtenues avec un jeu 0,5 dB entre A et B soit des valeurs de -15 et -14,5 dB, respectivement.

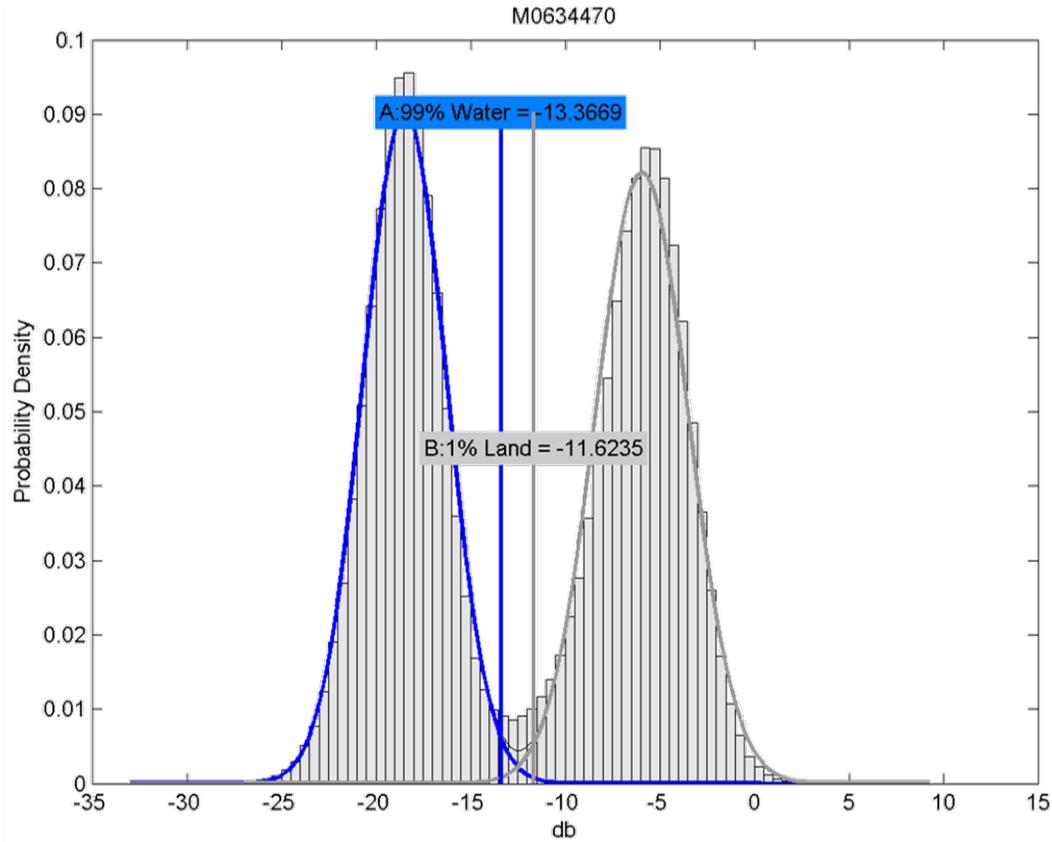


Figure 14 : Exemple d'un histogramme de densité de probabilité bimodal (en bleu l'histogramme correspondant aux pixels « eau », et en gris ce correspondant aux pixels « terre »)

3.4 Calcul de la largeur effective

Tout d'abord, la surface totale (S_{totale}) par section de pixels mouillés a été calculée. La surface totale est la somme des probabilités d'appartenance des pixels à l'eau multipliée par l'aire (A) d'un pixel d'image Radar en mode fin (64 m^2) pour une section comprenant n pixels (équation (3)). Ce calcul de surface s'est effectué de façon cumulative au fur et à mesure qu'on avançait dans les sections du cours d'eau à analyser. Finalement, la largeur effective (We) a été calculée en fonction de la surface totale (S_{totale}) de pixels mouillés et de la distance d'intégration (L_{max}) cumulée en amont de la station (équation (4)).

$$S_{totale} = \sum_1^n p(\text{pixel} = \text{eau}) \cdot A \quad (3)$$

$$We = S_{totale} / L_{max} \quad (4)$$

Les débits instantanés et les largeurs effectives ainsi calculées (We) ont été compilées dans une base de données en fonction de la distance d'intégration utilisée (Figure 15). Ces données ont servi, aux étapes subséquentes, à l'évaluation de la qualité des courbes de tarages en fonction de la distance d'intégration en amont de la station hydrométrique et à en identifier la valeur optimale. Un script intégrant l'intercalibration des images, le calcul de la probabilité d'eau et le calcul de la largeur effective a été développé pour construire cette base de données. Le **Erreur ! Source du renvoi introuvable.** montre un exemple des données produites par le script de calcul de la largeur effective.

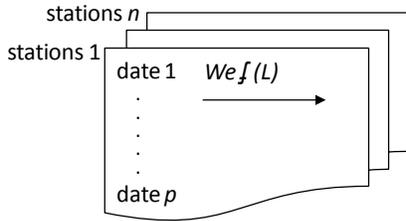


Figure 15 : Base de données de valeurs de We en fonction de la distance L_{max} (longueur par section)

Tableau 5 : Exemple de sorties du script pour le calcul de la largeur effective

Date	Code	Mode	Orbite	Q inst	KM \ Q quot	1	2	3	4
2009-06-14	DK371142	F3	D	24.30	23.90	47.63	44.74	44.67	45.77
2009-06-24	DK371144	F22F	D	24.27	25.40	46.23	44.31	43.43	44.89
2009-07-15	DK371149	F5F	D	24.56	23.90	46.88	44.94	43.97	44.39
2009-08-08	DK371155	F5F	D	14.83	15.10	40.24	37.92	36.51	36.49
2009-09-01	DK371158	F5F	D	63.86	65.00	49.86	48.43	48.44	50.33
2009-09-18	DK371162	F3	D	30.13	31.60	46.39	43.84	43.49	45.06
2009-09-25	DK371182	F5F	D	43.14	40.60	49.54	47.01	46.77	48.46
2009-09-28	DK371184	F22F	D	40.52	40.70	42.62	39.94	40.58	42.40
2009-10-19	DK371188	F5F	D	92.10	91.40	52.14	49.78	49.75	51.59

3.5 Ajustement d'une fonction pour la courbe de tarage

La dernière étape fut l'obtention de la courbe de tarage locale, à partir des valeurs de largeur effective calculées à chaque distance d'intégration et de données de débit instantané correspondant à la date et l'heure d'acquisition des images Radar. Pour ce faire, trois types de fonctions ont été testés; soit la fonction de puissance (équation (5)), l'exponentielle (équation (6)) ainsi qu'une polynomiale d'ordre 2 (équation (7)). Dans le cas de la polynomiale, l'ordre 2 a été choisie puisque les ordres supérieurs engendraient un trop grand nombre de paramètres à régionaliser.

$$Q = \alpha We^\beta \quad (5)$$

$$Q = \alpha e^{\beta We} \quad (6)$$

$$Q = AWe^2 + BWe + C \quad (7)$$

où Q est le débit instantané, α et β représentent le coefficient et l'exposant respectivement, We est la largeur effective, A et B sont les coefficients pour la polynomiale et C est l'ordonnée à l'origine de la polynomiale.

Initialement, il était prévu utiliser le modèle de puissance pour le calcul des courbes de tarages. Cependant, il a été remarqué que le modèle de puissance ne permettait pas de bien représenter les débits au-delà d'une certaine valeur de largeur effective. Pour cette raison, la courbe exponentielle a été introduite et après coup, le coefficient de détermination caractérisant l'ajustement de la fonction s'est amélioré. Pour cette raison, la courbe de puissance a été abandonnée. Toutefois, il a été décidé d'y introduire une forme polynomiale d'ordre 2 afin de vérifier si la courbe s'ajustait mieux aux valeurs plus élevées de largeurs effectives. Puisqu'un polynôme d'ordre 2 est en fait une forme parabolique, il fallait faire en sorte que seule la partie croissante de la parabole soit conservée. Dans ce cas-ci, le modèle exponentiel était utilisé pour estimer les valeurs de débits dans la partie inutilisable de la parabole. La largeur effective correspondante à l'intersection des deux modèles a été définie comme point de changement de fonction.

Un script a été créé afin d'automatiser le calcul des courbes de tarage satellitaire et du coefficient de détermination, ainsi que la distance d'intégration. Ce dernier comprend la mise en place d'une interface graphique pour l'analyse des résultats préliminaires et finaux. L'interface graphique du script pour le calcul des courbes de tarage satellitaire (Figure 16) montre les résultats du calcul de la largeur effective par section et par orthoimage, pour une station donnée (avec des informations pertinentes telles que la date d'acquisition des images, l'angle d'incidence, le type d'orbite et le débit instantané). Il est possible de faire des choix pour construire les courbes de tarage satellitaire, soit en fonction du type d'orbite (ASC ou DESC) et/ou l'angle d'incidence, soit par sélection directe sur la liste d'images disponibles.

76 Utilitaire de regression

Liste des images disponibles:

Date	Code	Mode	Orbite	Q inst	KM \ Q quot	1.0	2.0	3.0	4.0	5.0	6.0	7.0	8.0	9.0	10.0	11.0	12.0	13.0	14.0	15.0	16.0	17.0	18.0	19.0	20.0	21.0	
<input checked="" type="checkbox"/>	2008-09-03	DK371120	F22	D	401.808	401.0	264.21	285.45	227.68																		
<input checked="" type="checkbox"/>	2009-05-01	DK371135	F22	D	754.909	682.0	282.41	312.03	259.70																		
<input checked="" type="checkbox"/>	2009-09-22	DK371163	F22	D	231.104	238.0	254.51	281.98	230.95																		
<input checked="" type="checkbox"/>	2010-05-20	DK371216	F22	D	575.081	570.0	243.40	289.24	261.82																		
<input checked="" type="checkbox"/>	2011-05-05	DK371293	F3N	D	1493.81	1500.0	320.26	351.07	319.72	339.01	340.64	340.46	361.97	366.75	371.86	351.09	341.04	333.62	330.04	325.54	321.48	310.95					
<input checked="" type="checkbox"/>	2011-05-11	DK371298	F22N	A	1399.58	1400.0	244.11	288.70	264.84	268.91	268.58	271.80	287.67	285.92	288.92	270.72	266.98	264.37	264.66	262.58	258.93	255.38	252.16	251.02	252.03	255.46	267.26
<input checked="" type="checkbox"/>	2011-05-12	DK371301	F5N	D	1386.24	1380.0	323.65	354.56	322.70	340.27	341.06	338.23	359.43	364.14	371.86	351.02	344.36	339.81	338.03	333.17	328.04	319.44					
<input checked="" type="checkbox"/>	2011-05-15	DK371303	F23F	D	1302.42	1290.0	285.56	331.97	303.81	319.61	319.03	315.89	330.11	339.21	353.21	337.68	335.92	333.23	308.75								
<input checked="" type="checkbox"/>	2011-05-19	DK371306	F6F	D	1418.33	1420.0	337.26	365.03	330.01	347.23	347.85	346.26	367.41	372.22	382.00	363.87	359.05	354.58	352.96	347.60	342.23	335.79	329.95	322.76			
<input checked="" type="checkbox"/>	2011-06-05	DK371313	F5N	D	1316.8	1310.0	283.89	318.75	293.12	311.46	312.19	307.60	323.70	324.49	325.16	302.83	292.38	287.31	285.24	282.39	279.50	274.86	270.89	269.16	269.20	271.56	282.79
<input checked="" type="checkbox"/>	2011-06-29	DK371328	F5N	D	774.196	771.0	302.38	321.34	294.83	313.42	316.43	312.44	331.41	336.16	340.32	319.70	311.95	306.12	304.58	301.08	297.55	292.61	287.92	285.67	285.32	286.21	296.71
<input checked="" type="checkbox"/>	2011-09-02	DK371341	F3	D	518.487	532.0	309.18	330.98	305.85	327.23	331.15	331.52	356.16	362.05	368.71	346.47	334.26	322.96	317.33	312.70	308.66	303.68	298.88	296.38	295.63	296.76	307.58
<input checked="" type="checkbox"/>	2011-10-20	DK371344	F3	D	537.518	568.0	327.60	339.23	309.97	328.64	331.56	327.87	350.05	356.89	367.07	347.81	339.04	329.72	326.25	322.67	318.61	313.22	308.17	305.15	304.12	304.89	315.59
<input checked="" type="checkbox"/>	2012-09-20	DK371379	F3	D	135.094	143.0	314.11	330.73	306.38	327.81	331.44	331.79	354.51	360.09	366.71	343.10	329.10	316.84	311.94	308.20	304.46	299.50	294.61	292.12	291.27	292.16	302.62
<input checked="" type="checkbox"/>	2012-09-27	DK371380	F5N	D	123.858	127.0	314.57	333.69	308.83	330.14	334.23	334.08	354.52	361.05	371.09	349.85	342.53	336.30	334.09	329.48	324.65	318.79	312.97	295.64			

Orbite: All Desc Asc

Mode: F22 F22N F23F F3 F3N F5N
 F6F

Section:

Répertoire d'enregistrement: E:\Stations\020\007\Output\ZMF\140425_1031\Graphiques\{Date_Hi_Selectionner...} Sauvegarder les résultats

Commentaire:

Figure 16 : Interface graphique du script pour le calcul de la courbe de tarage satellitaire et la distance d'intégration

Ces courbes peuvent être visualisées (Figure 17) et si les résultats sont satisfaisants elles peuvent être aussi enregistrées dans un sous-dossier sous forme de figure. Dans cet exemple, la courbe de tarage exponentielle a été mise en rouge alors que la polynomiale en vert. L'idée était de générer les courbes de tarage dans un premier temps sous une forme exponentielle ce qui constituait la première courbe de tarage. Dans un deuxième temps, la seconde courbe de tarage était la forme exponentielle jumelée avec la polynomiale. Finalement, la courbe de la distance d'intégration (Figure 18) a été calculée selon les coefficients de détermination des courbes de tarage par section.

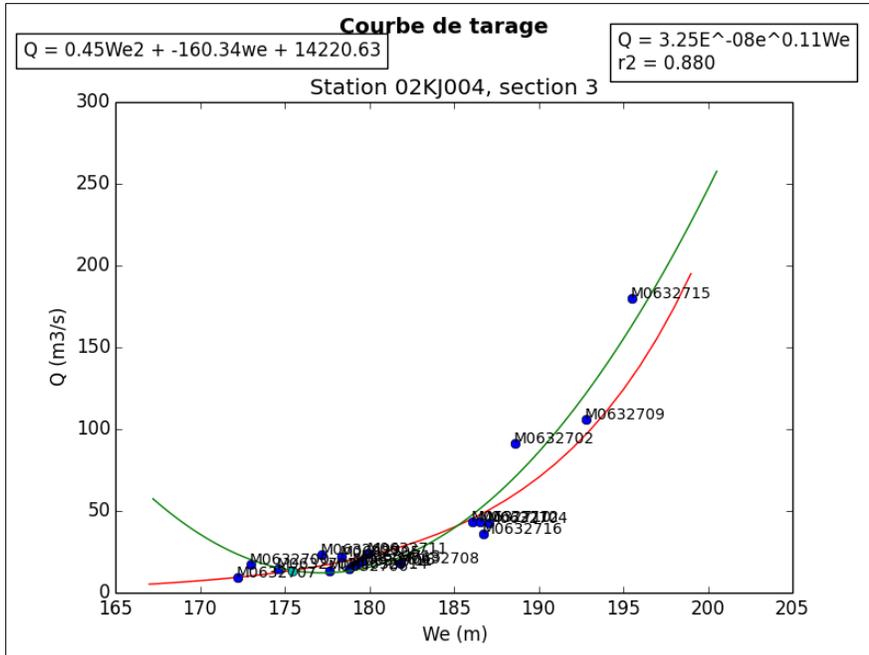


Figure 17 : Exemple de courbe de tarage satellitaire, avec le coefficient de détermination

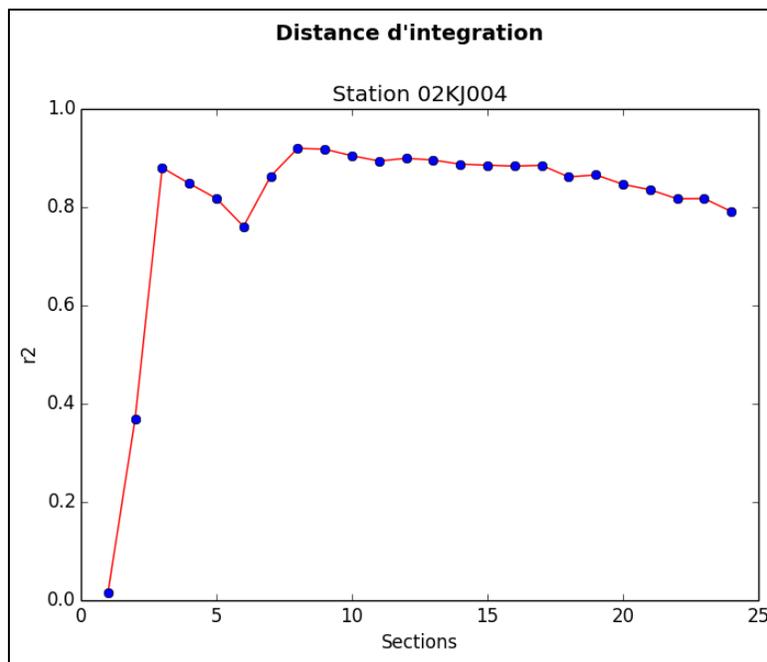


Figure 18 : Exemple d'une courbe de la distance d'intégration, avec les coefficients de détermination par section

Une fois que la distance d'intégration la plus performante a été choisie en fonction du coefficient de détermination le plus élevé, le script permet de récupérer les paramètres de la courbe de tarage pour la distance d'intégration retenue. Ces paramètres ont été conservés dans

un fichier « log » et postérieurement exportés dans une table de la base de données « Finale », accompagnés d'autres informations complémentaires (**Erreur ! Source du renvoi introuvable.**).

Tableau 6 : Paramètres conservés dans la table courbe de tarage exponentielle

Identifiant de la station (Station)
Distance d'intégration (Sections)
Paramètre α (Alpha)
Paramètre β (Beta)
Coefficient de détermination (R^2)
Valeur P (P_value)
Erreur standard (Std_err)
Nombre d'images disponibles (nbImageDisp)
Proportion ascendant (Ascendant)
Proportion descendant (Descendant)
Commentaire sur les images / orbite utilisés (Commentaire)

3.6 Analyse de sensibilité des courbes de tarage

Avant d'entrer dans la procédure de régionalisation des courbes de tarage, des questionnements avaient été soulevés au sujet de l'impact de l'incertitude pour l'estimation des paramètres des courbes de targe. Au début du projet, les courbes de tarage utilisaient une fonction de puissance, puis en cours de route la fonction a été remplacée par une courbe exponentielle. Cette dernière apparaissait mieux adaptée au nuage de points. Or, en remplaçant la fonction ceci a occasionné un changement des paramètres définissant la fonction. Il a donc été décidé d'effectuer une analyse de sensibilité de chacun des paramètres des fonctions afin de savoir comment le débit estimé en était affecté. Pour ce faire, les trois types de fonctions de courbes de tarage, définies aux équations 4 à 6, ont été testés.

L'analyse de sensibilité d'une variable expliquée (fonction ou modèle) à une variable explicative ou un paramètre consiste à calculer la dérivée partielle et de l'évaluer à une valeur nominale. Le calcul des dérivées partielles peut être approximé par des méthodes utilisant de petits intervalles de variation. Dans cette analyse, la variation a été fixée à 1%. La rivière Terra-Nova a été sélectionnée pour la réalisation de cette analyse. Trois courbes de tarage ont été générées; soit une par type de fonction avec les largeurs effectives de la rivière Terra-Nova à la 3^e section. Les valeurs nominales pour évaluer la sensibilité étaient les paramètres de chacune des courbes de tarage de cette rivière. La largeur effective de cette station varie entre 40 et 65 mètres pour la section choisie. La sensibilité a été analysée pour ces deux largeurs afin de savoir si l'impact sur l'estimation du débit est le même à différentes largeurs.

Le Tableau 7 présente les résultats de l'analyse de la sensibilité du débit aux paramètres des trois types de fonctions de courbes de tarage, pour une largeur effective de 40 mètres. Il est

possible de constater que le modèle le moins sensible à la variation de ses paramètres était le modèle exponentiel avec 1,65 m³/s de variation de débit. Le paramètre beta (la pente) ainsi que la largeur effective ont contribué plus fortement à la variation du débit que le paramètre alpha. Les modèles de puissance et polynomiale d'ordre 2 se sont avérés plus sensibles à la variation de 1% que le modèle exponentiel. Ceci indiquait que ce dernier devrait être favorisé pour l'estimation du débit à des largeurs plus faibles.

La même analyse est présentée au

Tableau 8 pour une largeur effective de 65 mètres. Les résultats de l'analyse de sensibilité ont été toutefois différents de ceux obtenus à la largeur de 40 mètres. En effet, le modèle polynomial est celui qui a engendré les plus faibles variations de débits à une largeur effective plus grande (65m). De plus, tous les modèles se sont vu augmenter leur sensibilité de manière importante (Puissance > Exponentielle > Polynomiale).

Tableau 7 : Variation du débit (m³/s) engendrée par une variation de 1% des paramètres selon le type de fonction pour une largeur effective de 40 m

Paramètres	Puissance	Exponentielle	Paramètres	Polynomiale
Alpha	0,29	0,32	A	2,57
Beta	4,90	1,14	B	2,53
We	1,33	1,13	We	2,65
Modèle entier	5,12	1,65	Modèle entier	4,49

Tableau 8 : Variation du débit (m³/s) engendrée par une variation de 1% des paramètres selon le type de fonction pour une largeur effective de 65 m

Paramètres	Puissance	Exponentielle	Paramètres	Polynomiale
Alpha	2,58	3,09	A	6,81
Beta	48,82	18,28	B	3,98
We	11,70	18,09	We	9,45
Modèle entier	50,37	26,14	Modèle entier	12,11

Bien que cette analyse de sensibilité fut limitée à une seule station donnée et a été réalisée uniquement à une valeur de 1% de variation de chaque paramètre, il n'en demeure pas moins qu'elle a permis de révéler que la sensibilité différait selon la fonction utilisée à une largeur effective donnée et augmente avec la largeur effective différemment selon la fonction utilisée. Bien que l'incertitude véritable des paramètres n'ait pas été utilisée dans cette analyse, cette dernière a illustrée d'une certaine façon la propagation de l'erreur à travers la chaîne de traitements. Ainsi, le débit estimé à un site non-jaugé intègrera l'incertitude liée à l'estimation

de la largeur effective ainsi que l'incertitude du modèle utilisé à travers l'erreur d'estimation de ses paramètres.

L'analyse de la sensibilité tend à montrer que, pour l'estimation du débit, l'utilisation conjointe de deux fonctions, soit le modèle exponentiel pour les faibles débits puis une polynomiale d'ordre 2 pour les débits plus élevés, permet d'obtenir de meilleurs résultats. Tel que mentionné précédemment, la fonction polynomiale d'ordre 2 est une fonction parabolique. La valeur de largeur effective correspondante à l'intersection des deux fonctions a été sélectionnée comme point de passage d'une fonction à l'autre.

4 Développement du modèle d'estimation du débit aux sites non-jaugés

4.1 Sélection des stations pour la calibration du modèle

Avant d'établir la sélection des stations pour la calibration du modèle, il fallait s'assurer que l'ensemble des débits observés au moment de l'acquisition des images avaient une dynamique représentative de la dynamique des débits mesurés à la station au fil des ans. Cette analyse de représentativité du débit a été définie comme le rapport entre la différence entre les quartiles à 75% et 25% (Q75-Q25 ou ΔQ , équation (8)) pour les débits associés aux images (d_i) et les débits mesurés en continu à la station (d_s). Le calcul du pourcentage de représentativité (R) est obtenu avec l'équation (9).

$$\Delta Q = Q_{75} - Q_{25} \quad (8)$$

$$R = \frac{\Delta Q_{d_i}}{\Delta Q_{d_s}} * 100 \quad (9)$$

Ce rapport devait être égal ou supérieur à 50% pour que les couples images-débits disponibles d'une station hydrométrique soient considérés comme représentatifs de la variabilité observée à la station en continu depuis son ouverture. Les stations dont les couples images-débits ont été jugés non-représentatifs ont été retirées du groupe de stations de calibration du modèle d'estimation du débit aux sites non-jaugés. Le Tableau 9 présente les stations hydrométriques où les couples images-débits ont été jugés comme non-représentatifs.

Tableau 9 : Stations hydrométriques avec couples image-débit non-représentatifs

	ID	Prov.	RS1	RS2		ID	Prov.	RS1	RS2
1	02JB009	Qc	21	0	14	02XA003	TN	6	0
2	02OC021*	Qc	25	0	15	02XA008	Qc	8	0
3	02OE005**	Qc	13	0	16	02YC001	TN	4	0
4	02OD001	Qc	22	0	17	03AB002	Qc	14	0
5	02OD003	Qc	17	0	18	03BD002	Qc	4	0
6	02OE005	Qc	13	0	19	03FA003	Qc	4	0
7	02PD004	Qc	18	0	20	03LF002	Qc	2	0
8	02PG001	Qc	8	5	21	03NF001	TN	5	0
9	02RH035	Qc	11	0	22	03OC003	TN	5	0
10	02UC002	Qc	5	0	23	03OE001	TN	7	0
11	02VB004	Qc	4	0	24	03OE003	TN	6	0
12	02VC001	Qc	5	0	25	03PB002	TN	5	0
13	02WB003	Qc	7	0	26	03QC001	TN	5	0
					27	03QC002	TN	5	0

* Q instantané

** Q quotidien

Suite à cette analyse de la représentativité, la sélection des meilleures courbes de tarage a pu être réalisée à partir des stations retenues. À chaque station, la courbe de tarage offrant le meilleur coefficient de détermination pour le modèle exponentiel dans les premières 15 sections du cours d'eau, a été sélectionnée. Pour qu'une station soit retenue, celle-ci devait avoir un coefficient de détermination supérieur à 0,5, une *P-value* inférieure à 0,05 (seuil de signification) ainsi qu'une erreur d'estimation minimale. Au total, 37 stations hydrométriques ont été sélectionnées, dont 25 ayant une fonction croissante (pente positive) et 11 ayant une fonction décroissante (pente négative). Le Tableau 10 contient la liste des stations hydrométriques ayant des couples images-débits représentatifs et une courbe de tarage satisfaisant les critères de sélection précédemment énoncés. Le Tableau 11 contient la liste des 14 stations n'ayant pas satisfait les critères de sélection.

Tableau 10 : Stations hydrométriques sélectionnées pour la calibration du modèle

	ID	Prov.	RS1	RS2	Orbite	Fonction	Q instantané		Q journalier	
							Section	r^2	Section	r^2
1	01AL002	NB	20	1	Desc	croissante	1	0.53	1	0.53
2	01AQ002	NB	25	0	Desc	croissante	14	0.65	14	0.64
3	01BP001	NB	22	2	Desc	croissante	4	0.69	4	0.70
4	01BQ001	NB	15	1	Desc	croissante	1	0.69	1	0.69
5	01ED005	NE	0	16	Desc	croissante	9	0.50	9	0.50
6	02LD005	Qc	22	0	Desc	croissante	1	0.65	1	0.65
7	02LE024	Qc	23	0	Desc	croissante	10	0.70	10	0.70
8	02LE025	Qc	30	0	Desc	croissante	2	0.68	2	0.68
9	02OG002	Qc	0	17	Desc	croissante	1	0.79	1	0.07
10	02OH008	Qc	0	14	Desc	croissante	3	0.80	3	0.80
11	02OJ007	Qc	0	26	Desc	croissante	2	0.73	2	0.74
12	02PC002	Qc	21	0	Desc	croissante	1	0.79	1	0.74
13	02PJ005	Qc	30	1	Desc	croissante	3	0.70	3	0.69
14	02PL005	Qc	19	1	Desc	croissante	14	0.69	4	0.63
15	02YD002	TN	22	0	Desc	croissante	2	0.85	2	0.83
16	02YS005	TN	0	40	Desc	croissante	3	0.73	3	0.73
1	01BV006	NB	0	17	Asc	croissante	3	0.85	3	0.87
2	01EF001	NE	11	0	Asc	croissante	1	0.81	1	0.81
3	02KJ004	Qc	19	0	Asc	croissante	3	0.88	3	0.88
4	02LG005	Qc	25	0	Asc	croissante	1	0.82	1	0.83
5	02ND003	Qc	9	1	Asc	croissante	10	0.85	10	0.83
6	02OB017	Qc	30	1	Asc	croissante	3	0.51	4	0.56
7	02OG014	Qc	17	6	Asc	croissante	6	0.61	6	0.60
8	02PH011	Qc	26	0	Asc	croissante	5	0.80	4	0.87
9	04NA001	Qc	24	0	Asc	croissante	12	0.52	12	0.58
1	01AM001	NB	19	0	Desc	décroissante	1	0.75	1	0.74
2	01AQ001	NB	17	0	Desc	décroissante	8	0.50	8	0.50
3	02AQ002	Qc	10	0	Desc	décroissante	3	0.77	3	0.78
4	02OE027	Qc	11	0	Desc	décroissante	1	0.68	2	0.71
5	02OE032	Qc	20	0	Desc	décroissante	1	0.80	1	0.75
1	01BH005	Qc	0	14	Asc	décroissante	2	0.62	2	0.60
2	02NE011	Qc	11	0	Asc	décroissante	2	0.82	2	0.72
3	02NF003	Qc	34	0	Asc	décroissante	3	0.65	3	0.64
4	02OA057	Qc	25	5	Asc	décroissante	1	0.57	1	0.57
5	02OH013	Qc	0	13	Asc	décroissante	1	0.82	1	0.86
6	02RH066	Qc	23	0	Asc	décroissante	9	0.91	9	0.69
*	03BF001	Qc	5	0	Asc	croissante	1	0.59	1	0.63

* station non retenue (seulement 3 images avec orbite ascendante)

Tableau 11 : Stations hydrométriques non-retenues pour la calibration du modèle ($r^2 < 0.5$)

	ID	Prov.	RS 1	RS 2	Q instantané		Q journalier	
					section	r^2	section	r^2
1	02LC021	Qc	9	2	6	0.49	6	0.33
2	02OA054	Qc	19	11	1	0.03	1	0.03
3	02OC021	Qc	25	0	non représentatif		3	0.34
4	02OE005	Qc	13	0	11	0.23	non représentatif	
5	02OJ026	Qc	0	18	5	0.08	5	0.14
6	02PA007	Qc	13	0	5	0.17	5	0.48
7	02PC010	Qc	22	0	1	0.42	3	0.12
8	02PH010	Qc	16	2	13	0.37	12	0.27
9	02PJ007	Qc	25	0	1	0.34	1	0.27
10	02RB004	Qc	11	0	5	0.29	5	0.29
11	02RH027	Qc	24	0	2	0.25	3	0.22
12	02RH045	Qc	27	0	1	0.05	1	0.06
13	02YO011	TN	17	0	4	0.22	4	0.17
14	02YQ001	TN	21	0	3	0.12	3	0.11
15	02ZE004	TN	10	0	2	0.22	2	0.22
16	03ED001	Qc	4	0	1	0.42	1	0.37

Avant de procéder à la régionalisation des courbes de tarage, une nouvelle variable physiographique a été ajoutée; soit la largeur nominale à la distance d'intégration, ou W_n_{DI} (Figure 19). Cette variable met en rapport la largeur nominale du cours d'eau avec une zone tampon de 10% de part et d'autre et la longueur totale en amont de la station, à la section à laquelle la courbe de tarage a été obtenue (la distance d'intégration). Puisque la distance d'intégration n'avait pas été fixée, la largeur nominale à cette distance constitue une information intéressante pour contrer l'effet du changement de largeur d'une section à une autre d'une rivière.

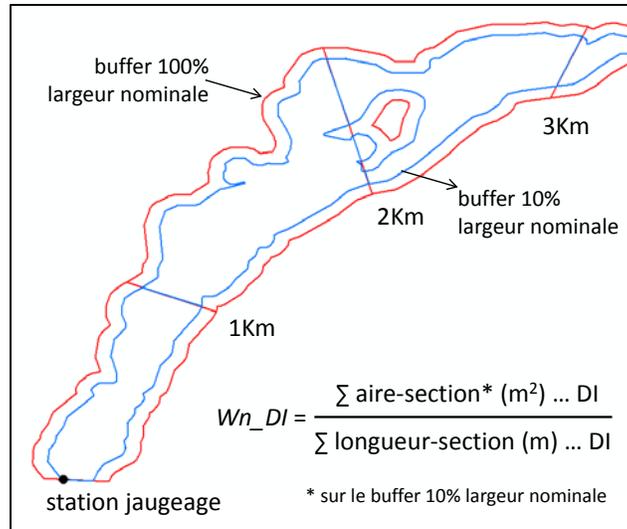


Figure 19 : Calcul de la variable Wn_DI (largeur nominale incluant le buffer de 10 %, à la distance d'intégration)

Comme il a été remarqué au Tableau 10, il existe deux types de courbes de tarage; soit celles à pentes croissantes et celles à pentes décroissantes. Il était donc impératif de créer une méthode permettant de connaître quelle était le type de courbe d'une rivière d'après ses conditions physiographiques. Pour ce faire, la méthode de l'arbre de classification et régression (Classification and Regression Tree ou CART; Breiman *et al.* (1984)) a été utilisée. Cette méthode utilise un échantillon ayant un ensemble de variables explicatives et une classe d'appartenance (variable expliquée). Un arbre de régression est alors construit.

Les variables d'entrées pour la méthode CART étaient les classes de couvertures forestières EOSD extraites à la distance d'intégration, dans la zone tampon de 100% de la largeur nominale. Ces classes ont été résumées en utilisant l'analyse en composantes principales (ACP). Des 36 classes de couvertures existantes d'EOSD, seules 20 classes étaient présentes dans les buffers des rivières du jeu de calibration. Ainsi l'ACP a produit 20 composantes principales.

Bien que les 20 composantes aient été introduites, la méthode CART a sélectionné uniquement les deux premières composantes X1 et X2 qui couvraient 62,9% de la variance cumulée (Figure 20). Il avait été préalablement observé que l'effet de double-rebond dans les berges d'un cours d'eau était associé aux courbes décroissantes. D'un point de vue théorique, le double rebond est associé au type de couverture de la terre présent dans la plaine inondable et est spécifique à chaque section de chaque rivière. La technique proposée a permis de faire cette analyse à partir d'une information canadienne disponible pour toutes les plaines inondables considérées.

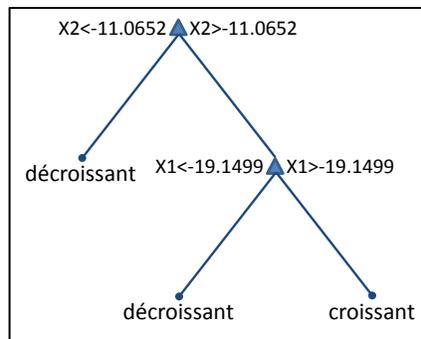


Figure 20 : Arbre de décision pour la sélection des courbes croissantes ou décroissantes

4.2 Régionalisation des courbes de tarage

La régionalisation des courbes de tarage pour le débit instantané ne pouvait pas combiner les modèles de courbes exponentielles croissants avec ceux décroissants. En effet, en plus des signes de pentes différents, les ordonnées à l'origine des modèles étaient contrastants. Les modèles croissants avaient des valeurs d'ordonnées à l'origine qui tendaient vers 0 alors que celles des modèles décroissants tendaient vers l'infini. D'un point de vue statistique, il s'agissait de deux distributions différentes, d'où l'importance de les traiter séparément. La régionalisation est présentée en trois étapes : soit l'exploration de la base de données physiographiques, l'analyse canonique de corrélation ainsi que les régressions multiples entre les variables canoniques et les paramètres des courbes de tarage. Pour cette analyse, seul le débit instantané a été régionalisé. Dans un premier temps, les sections 4.2.1 à 4.2.6 présenteront les résultats pour la régionalisation des courbes de tarage croissantes.

4.2.1 Exploration de la base de données physiographique

La base de données physiographique contient un grand nombre de variables (88). Ces variables appartiennent à des familles selon leur nature telles que la morphologie et morphométrie de la rivière et du bassin versant, la topométrie (pentes), le climat, la couverture du sol (par rapport au bassin versant et au corridor de la rivière). L'extraction de ces variables demande un effort considérable pour un usager alors qu'elles ne sont assurément pas toutes liées aux paramètres des courbes de tarage (Alpha et Beta pour l'exponentielle ainsi que A, B et C pour la polynomiale d'ordre 2). Il était donc essentiel d'étudier d'abord la relation de ces variables avec les paramètres des courbes de tarage par une analyse de corrélation simple et de retirer, dans un premier temps, les variables n'ayant pas présenté de lien significatif avec ces paramètres. Une fois cette première sélection réalisée, il a été important d'évaluer la redondance de l'information de des variables physiographiques conservées pour éviter l'introduction de variables colinéaires. Pour ce faire, la matrice de corrélation croisée de ces variables a été analysée pour identifier les groupes de variables corrélées. Un seul représentant de ces groupes de variables a été sélectionné; soit le plus en lien avec un paramètre de courbe de tarage et le

plus facilement calculable. Une fois ces étapes achevées, cette base de données physiographique épurée a été utilisée dans l'analyse canonique de corrélation.

4.2.2 Analyse de la normalité

Avant de procéder aux analyses de corrélation simple, il est important de s'assurer que les variables physiographiques ainsi que les paramètres des courbes de tarage sont normalement distribués. En cas de non-normalité, les variables devront être transformées pour respecter la distribution normale. Trois types de transformation ont été utilisés :

Transformation en puissance de type « Boxcox » : méthode itérative qui détermine le meilleur exposant (λ) favorisant la normalité (équation (10)).

$$x_{trans} = \frac{x^\lambda - 1}{\lambda} \quad (10)$$

Transformation logarithmique, le logarithme népérien est utilisé directement lorsque les valeurs sont supérieures à 0, si ces dernières contiennent des 0, une valeur de 1 est ajoutée et enfin si les valeurs sont négatives, ces dernières sont ramenées positives. À moins de contre-indications, la forme $x+1$ a été utilisée dans nos analyses (équation (11)).

$$x_{trans} = Ln(x), Ln(x + 1) \text{ ou } Ln(-x) \quad (11)$$

Transformation angulaire, idéale pour les données exprimées en proportions ou pourcentages, dans notre cas, la valeur est divisée par 100 dans la fonction pour ramener le pourcentage en valeur décimale (ex. les classes de couverture du sol) (équation (12)).

$$ArcSin\left(\sqrt{\frac{x}{100}}\right) \quad (12)$$

L'outil « Normplot » du logiciel Matlab a été utilisée pour étudier la normalité de chacune des variables physiographiques et des paramètres des courbes de tarage. Lorsque les données sont normalement distribuées, ces dernières se situent le long de la droite de normalité. La Figure 21 illustre un exemple avec le coefficient de massivité avant l'application d'une transformation pour la normalité. En observant le graphique, il est possible de remarquer que les valeurs du coefficient de massivité inférieures à environ 0,1 ou supérieures à environ 0,4 s'éloignent de la droite de normalité.

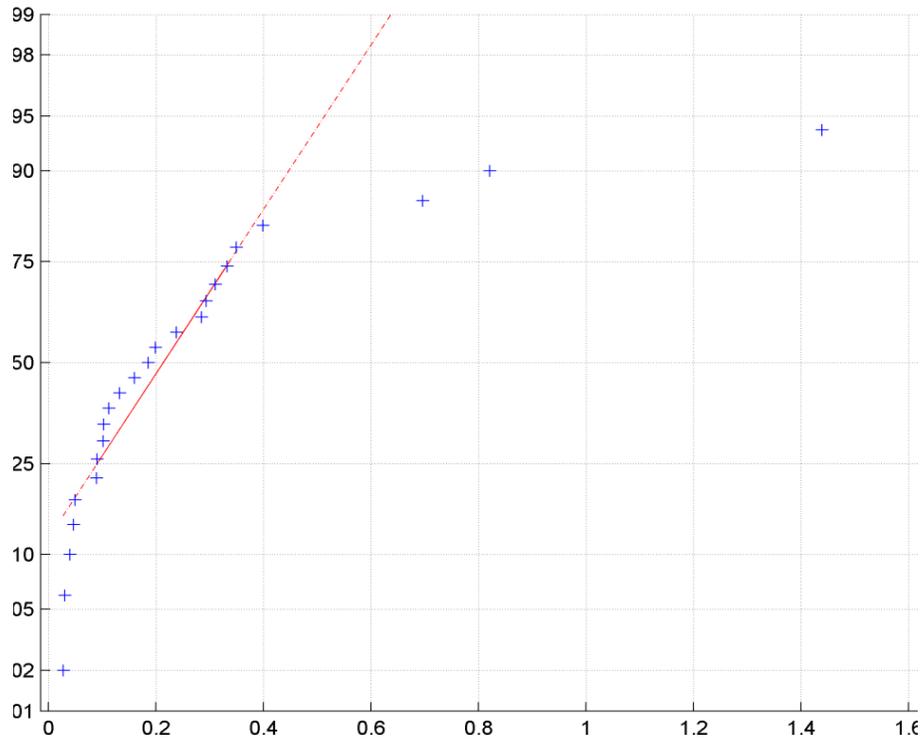


Figure 21 : Exemple d'utilisation de la fonction « Normplot » de Matlab avec le coefficient de massivité

La Figure 22 illustre le même exercice mais cette fois-ci après l'application de la transformation Boxcox ($\lambda = -0,0771875$) sur le coefficient de massivité. La transformation n'a pas été parfaite, cependant elle a permis de rapprocher la majorité des variables près de la droite de normalité en réduisant de manière considérable l'écart observé à la Figure 21. L'exercice a été réalisé auprès des 88 variables physiographiques et des 5 paramètres de courbes de tarage. Pour les besoins du présent rapport, seules les variables retenues à la fin de ce processus de sélection sont détaillées, toutefois les résultats des transformations sont disponibles dans un fichier Excel ci-annexé.

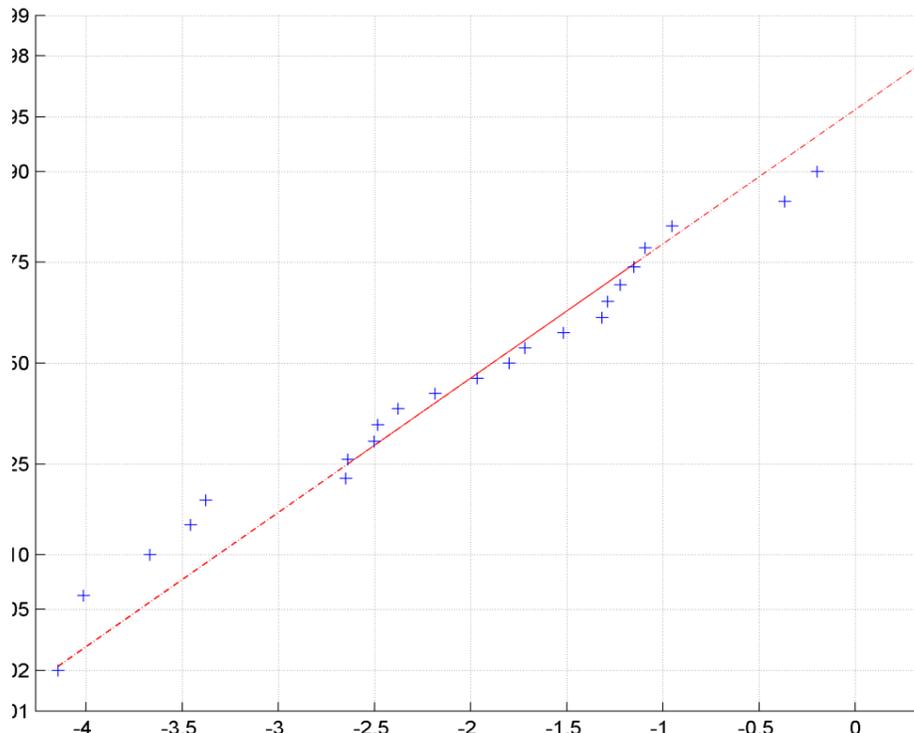


Figure 22 : Exemple d'utilisation de la fonction « Normplot » de Matlab après la transformation du coefficient de massivité

La Figure 23 illustre l'effet de la transformation de 6 variables (à titre d'exemple) sur leur distribution et sur leurs relations entre-elles. Les transformations appliquées aux paramètres des courbes de tarages pour respecter la normalité sont présentés au Tableau 12.

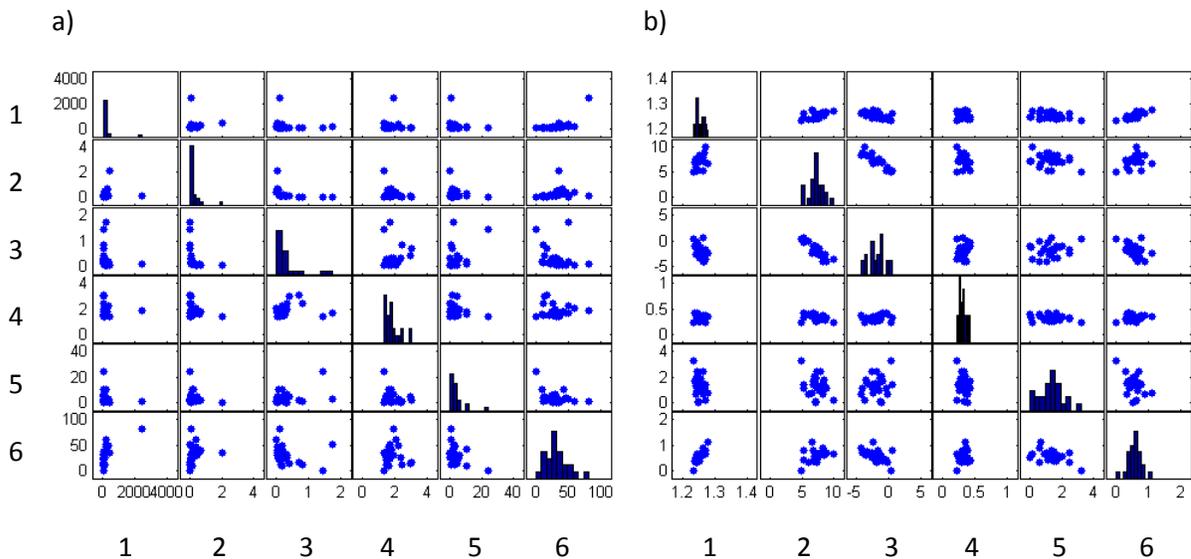


Figure 23 : Exemple de matrice de distribution a) avant transformation et b) après transformation normale

Où : 1 est la largeur effective à la distance d'intégration, 2 est la superficie, 3 est le coefficient de massivité, 4 est la sinuosité du cours d'eau principal, 5 est la pente transversale à la section 3 et 6 est la classe « Eau » extraite à la section 3 de la couverture EOSD.

Tableau 12 : Transformations appliquées aux paramètres des courbes de tarage pour respecter la normalité

Paramètre (type de courbe)	Type de transformation
Alpha (exponentielle)	$\ln(x)$
Beta (exponentielle)	Boxcox : $\lambda = 0,2570625$
A (polynomiale ordre 2)	Boxcox : $\lambda = 0,189375$
B (polynomiale ordre 2)	$\ln(-x)$
C (polynomiale ordre 2)	$\ln(x+1)$

4.2.3 Analyse de la corrélation entre les variables physiographiques transformées et les paramètres des courbes de tarage

Après avoir procédé à la détection des variables physiographiques distribuées non-normalement et à leur transformation, l'analyse de corrélation entre ces-dernières et les paramètres des courbes de tarage a pu débuter. À cette étape le critère de sélection d'une variable était qu'elle soit significativement corrélée, au minimum, avec un des cinq paramètres des courbes de tarage. Pour ce faire, les coefficients de corrélation simple ainsi que leur p-value respective pour chacune des variables physiographiques avec chacun des 5 paramètres ont été générés. Le seuil de signification a été fixé à 5%, un standard en statistique, ainsi seules les corrélations ayant une p-value inférieure à 0,05 ont été conservées. Ceci indique qu'il y a plus de 95% de chances que la corrélation observée soit significative. Les variables retenues ainsi que leurs coefficients de corrélations sont présentées au Tableau 13. Les corrélations significatives pour lesquels une variable physiographique a été retenue sont surlignées en vert. Il y avait 33 variables présentant une corrélation significative avec au moins un paramètre de courbes de tarage. En analysant en détail ce tableau, il est possible de remarquer que les paramètres de la courbe de tarage de type exponentielle ont un plus grand nombre de variables corrélées significativement que les paramètres du modèle polynomial d'ordre 2.

Tableau 13 : Corrélations des variables physiographiques présentant au minimum une corrélation significative (p-value < 0,05) avec un paramètre de courbe de tarage

	Paramètres des courbes de tarage				
	Alpha	Beta	A	B	C
Wn_DI	-0,31	-0,63	-0,42	0,15	0,60
Wn_Station	0,18	-0,37	-0,02	0,34	0,48
L_Axis	0,42	-0,57	0,03	0,26	0,32
Perim	0,42	-0,62	-0,03	0,21	0,31
Superficie	0,44	-0,64	-0,03	0,24	0,34
Coef_compa	0,26	-0,40	-0,05	0,08	0,15
Latitude_m	0,30	0,02	-0,07	-0,28	-0,40
Coeff_mass	-0,40	0,71	0,20	-0,12	-0,32
Coeff_orog	-0,34	0,66	0,33	0,03	-0,22
Shreve	0,45	-0,61	-0,05	0,18	0,26
Strahler	0,41	-0,61	-0,01	0,28	0,38
Pmoy_a	0,47	-0,67	-0,06	0,21	0,32
Pmax_m	0,47	-0,68	-0,07	0,20	0,32
IF	0,46	-0,68	-0,07	0,20	0,33
Qe	0,48	-0,69	-0,08	0,19	0,31
L_drainage	0,44	-0,63	-0,04	0,22	0,31
L_princ	0,43	-0,58	0,00	0,21	0,28
P_princ	-0,28	0,61	0,21	-0,03	-0,23
S_princ	0,00	0,00	-0,27	-0,41	-0,36
P_princ_10	-0,01	0,43	0,00	-0,16	-0,29
v_18	0,27	-0,57	-0,26	-0,07	0,12
s1	-0,04	0,46	0,21	-0,01	-0,21
s2	-0,10	0,45	0,22	0,04	-0,13
s3	-0,13	0,48	0,21	0,04	-0,11
s4	-0,15	0,45	0,17	0,03	-0,09
s5	-0,13	0,49	0,13	-0,05	-0,19
s6	-0,14	0,50	0,16	-0,02	-0,18
s7	-0,13	0,52	0,17	-0,01	-0,17
s8	-0,13	0,50	0,12	-0,06	-0,20
s9	-0,16	0,53	0,11	-0,05	-0,19
s10	-0,17	0,53	0,08	-0,09	-0,22
v_20 (EOSD)	-0,30	-0,55	-0,26	0,19	0,56
v_211 (EOSD)	0,04	0,49	0,09	-0,28	-0,49

4.2.4 Analyse de la corrélation croisée entre les variables physiographiques transformées

Les variables physiographiques identifiées à l'étape précédente ont été utilisées pour en étudier leur corrélation croisée (entre-elles). Cette fois-ci, une corrélation élevée d'une variable à une autre ou encore à un ensemble de variables signifiait qu'il y avait une redondance de l'information et par conséquent, qu'il n'y avait pas de gain à faire en incluant plus d'une variable fortement corrélée. Lors de la sélection d'une variable par rapport à d'autres corrélées, il était important de considérer la facilité à calculer cette variable de même que le sens qu'elle avait d'un point de vue d'interprétation physique. Le Tableau 30 et sa légende (Tableau 31), présentés à l'annexe III, décrivent le résultat de l'analyse de corrélation croisée entre les variables physiographiques sélectionnées. La matrice a été construite de manière symétrique, séparée en deux par la diagonale surlignée en jaune. Les corrélations supérieures à 0,9 en valeur absolue (en caractère de couleur rouge) ont été considérées comme redondantes. Les variables ayant des corrélations supérieures à cette valeur absolue ont été regroupées ensemble par couleurs (bleu, orange et vert). Le groupe orangé a été résumé par la variable superficie du bassin versant, une variable présentant de fortes corrélation avec alpha et beta (Tableau 13) et qui est facilement calculable. Le groupe bleu a été résumé avec le coefficient de massivité puisque ce dernier a présenté une corrélation significative de plus avec Alpha (Tableau 13). Enfin, le regroupement vert constituait les pentes des sections transversales à différentes distances (en km) de la station. La pente à la section 3 a été retenue puisque la 3^e section constitue la distance d'intégration moyenne des courbes de tarage, la distance à laquelle la régionalisation a été effectuée. Le Tableau 14 est un récapitulatif du Tableau 13 qui exclue les variables physiographiques redondantes. Au final, il y a douze variables physiographiques non redondantes qui ont présenté une corrélation significative avec minimalement un paramètre des courbes de tarage.

Tableau 14 : Sélection des variables physiographiques non redondantes présentant une corrélation avec au moins un paramètre des courbes de tarage

	Paramètres des courbes de tarage				
	Alpha	Beta	A	B	C
Wn_DI	-0,31	-0,63	-0,42	0,15	0,60
Wn_Station	0,18	-0,37	-0,02	0,34	0,48
Superficie	0,44	-0,64	-0,03	0,24	0,34
Latitude_m	0,30	0,02	-0,07	-0,28	-0,40
Coeff_mass	-0,40	0,71	0,20	-0,12	-0,32
P_princ	-0,28	0,61	0,21	-0,03	-0,23
S_princ	0,00	0,00	-0,27	-0,41	-0,36
P_princ_10	-0,01	0,43	0,00	-0,16	-0,29
v_18	0,27	-0,57	-0,26	-0,07	0,12
s3	-0,13	0,48	0,21	0,04	-0,11
v_20 (EOSD)	-0,30	-0,55	-0,26	0,19	0,56
v_211 (EOSD)	0,04	0,49	0,09	-0,28	-0,49

Malgré cette réduction importante du nombre de variables physiographiques, un effort supplémentaire de sélection de variables a dû s'ajouter. En effet, le fait d'avoir seulement 25 échantillons dans l'analyse s'avérait limitant dans la résolution d'un système d'équations où le nombre d'échantillons doit être supérieur au nombre d'inconnus à résoudre. Ainsi, si l'analyse canonique de corrélation avait été réalisée sur les 5 paramètres des courbes de tarage, 10 variables canoniques auraient été générées (U1 à U5 et V1 à V5), multipliées par les douze variables sélectionnées, ce qui fait un total de 120 inconnus à résoudre, ce qui dépasse largement notre limite maximale de 25. Avec 10 variables canoniques il fallait sélectionner tout au plus deux variables physiographiques pour ne pas dépasser le cap des 25 inconnues, ce qui s'avérait limitant. La solution a été de mener deux analyses canoniques de corrélations soit une avec les paramètres de la courbe exponentielle et l'autre avec les paramètres de la courbe polynomiale.

Ainsi dans le cas du modèle exponentiel, il n'y avait que 4 variables canoniques à générer ce qui permettait d'avoir au maximum 6 variables physiographiques faisant ainsi 24 inconnus à résoudre avec 25 échantillons. Dans le cas du modèle polynomial il y avait 6 variables canoniques de générées, soit une paire de U et V par paramètre. Il fallait donc limiter le nombre de variables physiographiques à tout au plus 4.

Les variables retenues pour l'analyse canonique de corrélation pour la régionalisation des paramètres du modèle exponentiel sont détaillées au Tableau 15.

Tableau 15 : Variables physiographiques retenues pour la régionalisation du modèle exponentiel

Variable	Transformation
Wn_DI	Boxcox $\lambda = -0,7806875$
Superficie	$\ln(x+1)$
Coeff_mass	Boxcox $\lambda = -0,0771875$
P_princ	Boxcox $\lambda = 0,3216$
P_princ_10	Boxcox $\lambda = -3.5989$
v18	Angulaire
S3	$\ln(x+1)$
v20 (EOSD)	Angulaire
V211 (EOSD)	Boxcox $\lambda = 0,394688$

Le même exercice a été réalisé avec les variables retenues pour l'analyse canonique de corrélation pour la régionalisation des paramètres du modèle polynomial d'ordre 2 (Tableau 16).

Tableau 16 : Variables physiographiques retenues pour la régionalisation du modèle polynomial d'ordre 2

Variable	Transformation
Wn_DI	Boxcox $\lambda = -0,7806875$
Wn_Station	Boxcox $\lambda = -0,206219$
Latitude	Aucune
S_Princ	Boxcox $\lambda = -2,1361875$
V20	Angulaire
V211 (EOSD)	Boxcox $\lambda = 0,394688$

4.2.5 Analyse canonique de corrélation

Suite à la sélection des variables physiographiques, l'analyse canonique de corrélation (ACC) a pu débuter. La fonction « Canocorr » du logiciel Matlab a été utilisée pour effectuer cette analyse.

L'analyse a été réalisée dans un premier temps avec les paramètres des courbes de tarage de type exponentielles ainsi que leur ensemble de variables corrélées.

L'ACC a produit un ensemble de variables canoniques U (U1 et U2) à partir des variables physiographiques utilisées ainsi qu'un ensemble de variables canoniques V (V1 et V2) avec les paramètres Alpha et Beta. Puisqu'il n'y avait que deux variables physiographiques, deux variables ont été générées pour les U et les V. L'avantage d'utiliser une telle méthode provient du fait que les variables canoniques U1 et U2 ont une corrélation minimale, il en est de même avec les variables canoniques V1 et V2. Toutefois les corrélations entre U1 et V1 ainsi que U2 et V2 sont maximales. Dans le cas présent, seules les variables U ont été conservées afin de réaliser une régression directement avec les variables physiographiques. En plus de générer les variables

canoniques, l'algorithme a produit les coefficients canoniques nécessaires pour convertir les variables physiographiques en variables canoniques. L'équation (13) montre le calcul de U à partir des coefficients canoniques a_{ij} ainsi que des variables physiographiques x_j centrés (la moyenne de chaque variable est retranchée).

$$U_i = \sum_{j=1}^N a_{ij} * (x_j - \bar{x}_j) \quad (13)$$

Les coefficients a_{ij} , de même que les moyennes de chaque variables physiographiques (après normalisation) sont présentées au Tableau 17.

Tableau 17 : Moyennes des variables physiographiques et coefficients canoniques pour la transformation en variables canoniques (modèle exponentiel)

Variable physio.	Moyennes	a1 (U1)	a2 (U2)
Wn_DI	1,25	78,59	-13,94
Superficie	7,25	-0,15	0,03
Coeff_mass	-1,89	0,11	-0,79
S3	1,34	-0,10	-0,60
v20 (EOSD)	0,58	0,90	-1,93

Le même exercice a été réalisé pour l'analyse canonique du modèle polynomial d'ordre 2. Puisque ce modèle contenait 3 paramètres au lieu de 2 comme le modèle exponentiel, 3 variables canoniques U ont été générées. Le Tableau 18 présente les moyennes ainsi que les coefficients canoniques pour les variables physiographiques (après normalisation) sélectionnées pour le modèle polynomial d'ordre 2.

Tableau 18: Moyennes des variables physiographiques et coefficients canoniques pour la transformation en variables canoniques (modèle polynomial)

Variable physio.	Moyennes	a1 (U1)	a2 (U2)	a3 (U3)
Wn_DI	1,25	-73,17	-130,01	-66,10
S_Princ	0,32	2,04	-17,01	10,00
V20	0,58	-0,66	7,56	4,79

Les variables canoniques générées pour chaque modèle de courbes de tarage sont présentées au Tableau 19 pour l'ensemble des stations utilisées pour la régionalisation (débit instantané et modèle à pentes croissantes). Les variables canoniques U_i ont toujours une valeur de moyenne égale à 0 ainsi qu'un écart type de 1. Ainsi, en examinant la valeur de chaque variable canonique, en absolue, il est possible d'interpréter son degré d'éloignement à la moyenne. Puisque l'intervalle de confiance d'une distribution normale est défini à $1,96 * \text{l'écart type}$ avec

un seuil de signification à 5%, il est possible de détecter les valeurs éloignées lorsque qu'elles sont supérieures à 1,96 en absolue. Les stations présentant ces variables canoniques élevées ont été surlignées en jaune. Pour éviter d'introduire des valeurs extrêmes dans nos modèles de régressions, ces dernières ont été éliminées du jeu de variables U. Ainsi, pour le modèle exponentiel, deux stations ont été retirées alors que pour le polynomial, le nombre retirées est de trois. Il est à noter que cette analyse des variables canoniques avait été testée avant de déterminer les versions finales des ACC. Pour le modèle exponentiel, les variables suivantes ont été conservées : Wn_DI, Superficie, Coeff_mass, S3 et V20 (EOSD). Pour le modèle polynomial, ce sont plutôt celles-ci : WN_DI, S_princ ET V20 (EOSD).

Tableau 19 : Variables canoniques U, générées pour l'ensemble des stations dans chaque modèle de courbe de tarage

Identifiant de station	Modèle exponentiel		Modèle polynomial		
	U1	U2	U1	U2	U3
01AL002	-1,08	-0,09	0,82	-0,47	-0,85
01AQ002	-0,53	0,61	0,35	0,73	-0,73
01BP001	0,27	-0,76	-0,33	0,06	-0,86
01BQ001	-0,65	-0,87	0,75	0,61	1,00
01BV006	-1,66	-1,59	1,63	-0,12	-2,43
01ED005	2,51	-0,64	-2,08	0,45	1,29
01EF001	0,14	0,37	-0,33	1,04	-0,76
02KJ004	0,71	0,62	-0,96	0,13	-0,01
02LD005	-0,24	0,40	0,06	-0,71	-1,61
02LE024	-0,36	1,01	-0,01	-0,09	0,15
02LE025	1,51	-0,79	-1,16	-1,53	0,54
02LG005	0,57	1,52	-0,90	-0,84	-0,16
02ND003	0,55	-0,54	-0,71	-0,45	0,18
02OB017	-0,83	-0,56	0,79	0,83	0,41
02OG002	-0,15	0,21	0,36	0,68	0,56
02OG014	1,50	-2,53	-0,85	0,43	0,36
02OH008	-0,83	-0,32	1,30	-1,25	0,57
02OJ007	0,91	1,84	-1,48	-0,19	-1,77
02PC002	-0,62	-0,04	0,52	2,66	0,48
02PH011	-0,32	0,36	0,67	-0,78	0,96
02PJ005	0,78	0,39	-1,05	-0,36	-0,79
02PL005	-1,41	-0,34	1,61	-0,93	1,22
02YD002	-0,26	-0,75	0,88	-1,85	0,85
02YS005	-1,15	0,95	0,86	1,63	0,05
04NA001	0,66	1,53	-0,76	0,34	1,36

Le nuage de points illustrant l'espace canonique, sans les stations 01ED005 et 02OG014, pour la régionalisation du modèle exponentiel est illustré à la Figure 24. Tel que prévu, les variables U1 et U2 n'ont été aucunement corrélées. La même tendance a été remarquée avec les variables canoniques pour le modèle polynomial (Figure 25).

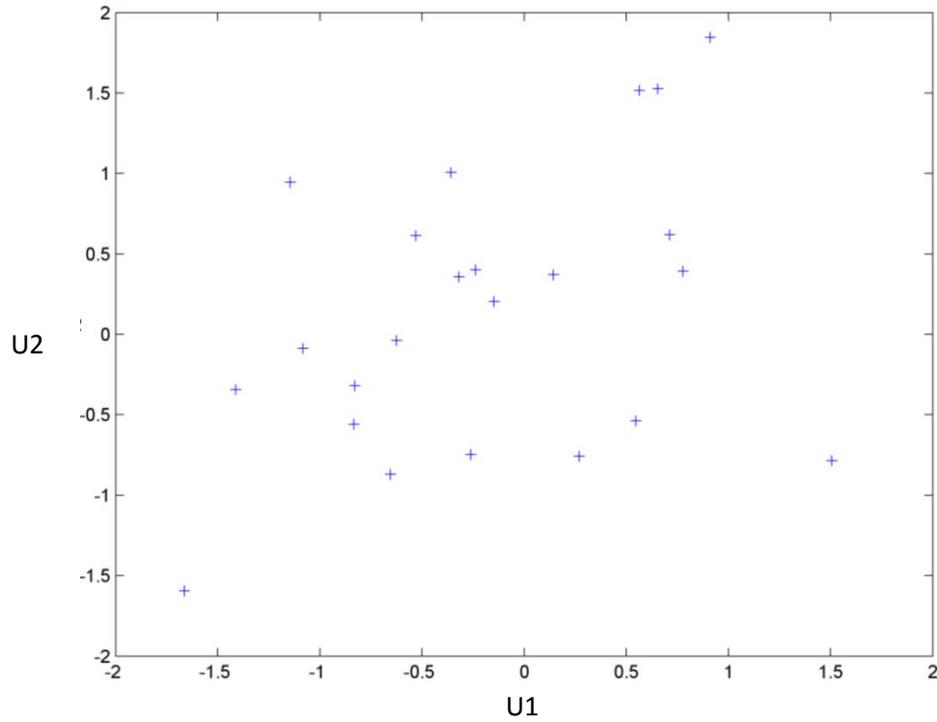
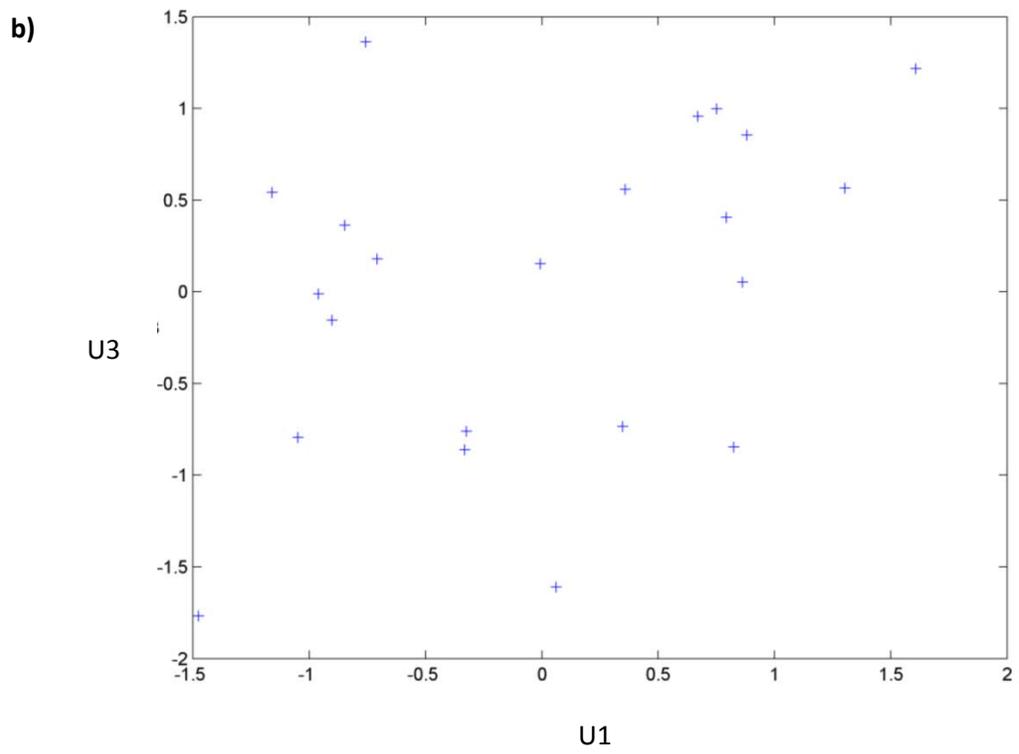
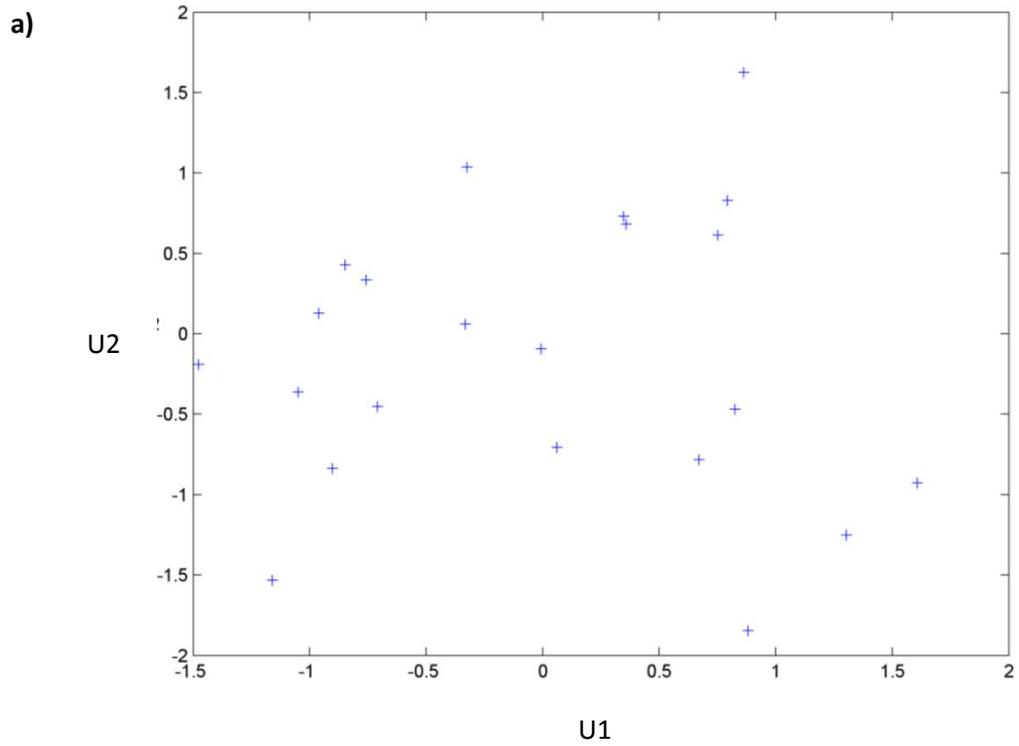


Figure 24 : Espace canonique pour le modèle exponentiel



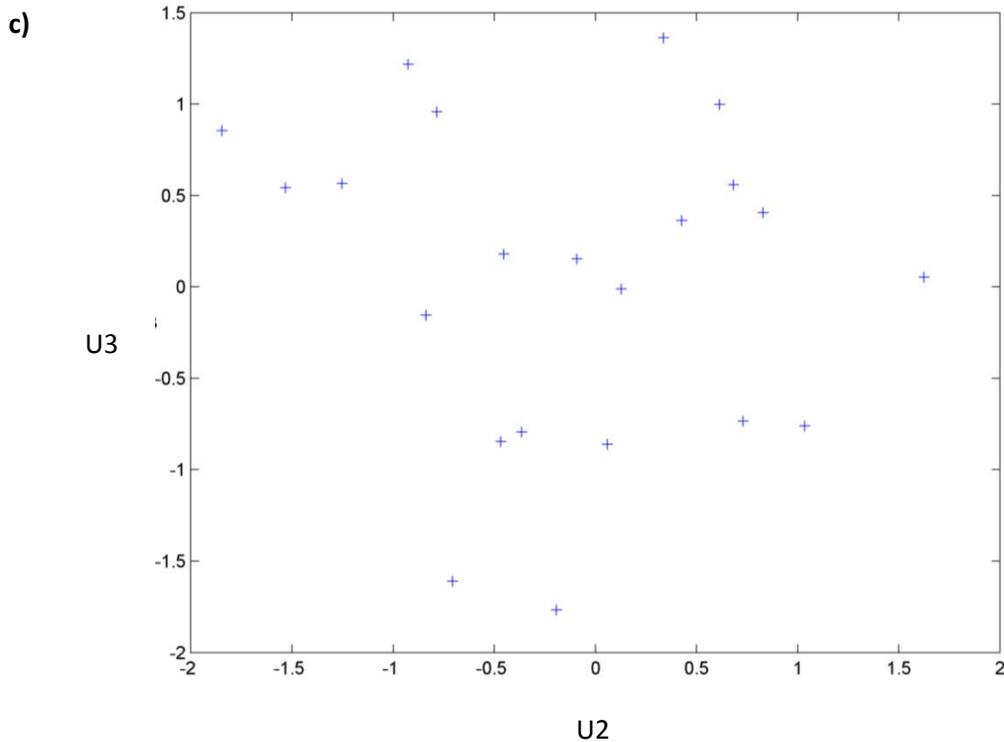


Figure 25 : Espace canonique pour le modèle polynomial avec a) U1 et U2, b) U1 et U3 et c) U2 et U3

4.2.6 Régressions avec les variables canoniques pour prédire les paramètres des courbes de tarage

Cette dernière étape de la procédure de régionalisation visait à développer les équations résultantes de la régionalisation des courbes de tarage pour les modèles croissants, en utilisant le débit instantané. Des régressions multiples ont été réalisées pour chaque paramètre des courbes de tarage à partir des variables canoniques générées. Pour qu'une variable canonique soit retenue dans un modèle de régression, il fallait que l'intervalle de confiance (avec une probabilité de 95%) de son coefficient de régression n'inclue pas une valeur de 0. Si une valeur de 0 était comprise dans cet intervalle de confiance, cela signifiait qu'il y avait plus de 5% de chances que le coefficient soit nul et donc, qu'aucun poids ne soit attribué à la variable canonique en question.

Les coefficients de régressions ainsi que leurs limites d'intervalles de confiance pour l'estimation des paramètres obtenus du modèle exponentiel de courbe de tarage sont présentés au Tableau 20. Aucune valeur de 0 n'est présente dans les intervalles de confiance de ces deux variables. Les deux variables canoniques U1 et U2 ont été retenues dans la régression pour l'estimation d'Alpha et de Beta.

Tableau 20 : Coefficients de régressions et ordonnées à l'origine pour l'estimation des paramètres Alpha et Beta du modèle exponentiel de courbe de tarage

	Alpha ¹			Beta*		
	Coefficient	Borne inf.	Borne sup.	Coefficient	Borne inf.	Borne sup.
Ordonnée à l'origine	-10,47	-13,33	-7,60	-1,57	-1,69	-1,46
U1	-4,80	-8,47	-1,14	-0,24	-0,39	-0,08
U2	5,80	2,37	9,24	-0,27	-0,41	-0,13

La Figure 26 illustre la relation entre le paramètre alpha et a) U1 et b) U2. Pour cette variable, il est possible de remarquer une faible relation négative avec U1 et une relation positive modérée avec U2.

¹ Valeurs transformées d'Alpha : Ln (Alpha), de Beta : Boxcox, $\lambda = 0,2570625$, de A : Boxcox, $\lambda = 0,189375$, de B : Ln(-B) et de C : Ln(C+1)

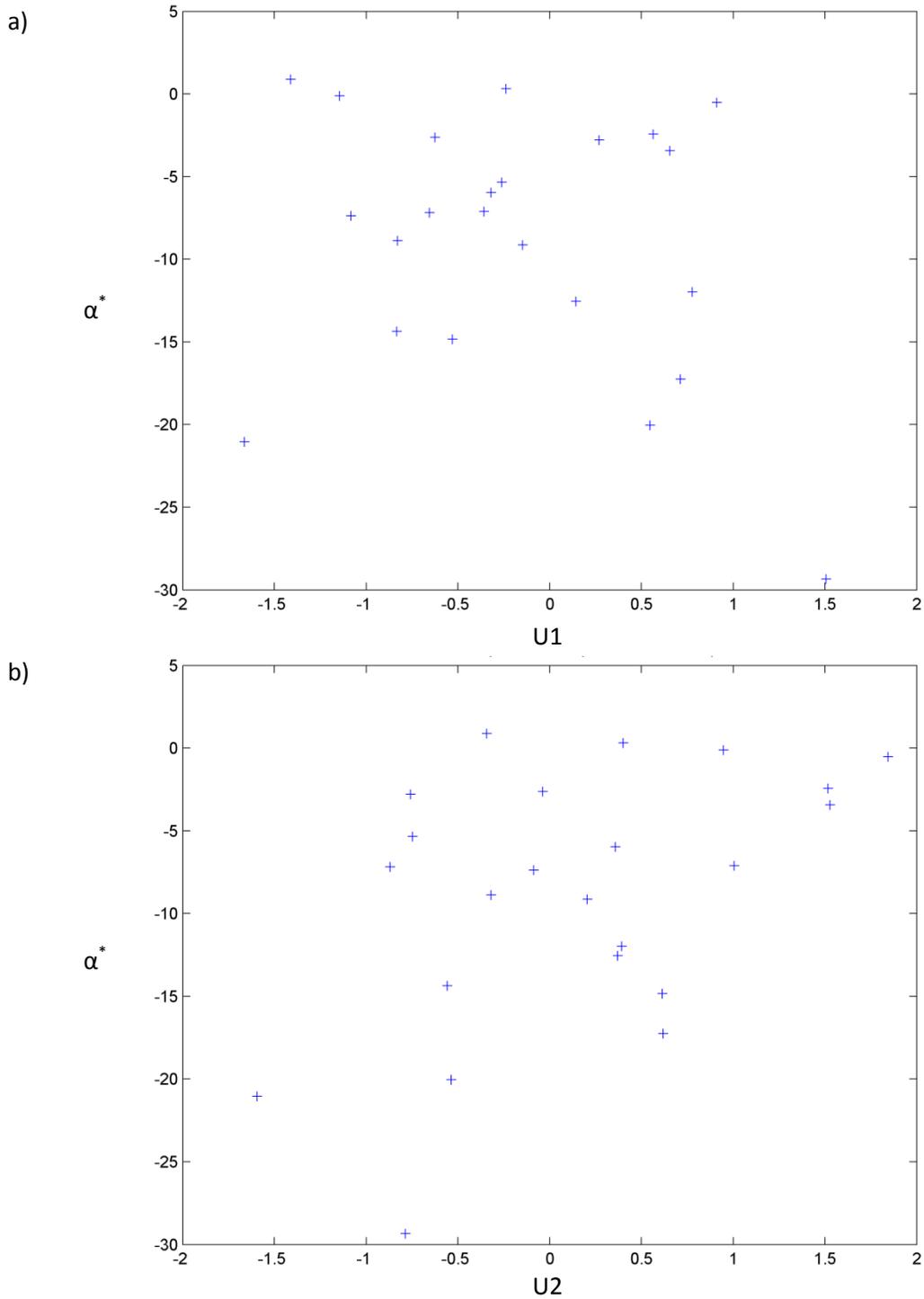


Figure 26 : Relation entre Alpha et a) U1 et b) U2

La Figure 27 illustre la relation entre le paramètre beta et a) U1 et b) U2. Pour cette variable, il est possible de constater des relations négatives avec U1 et U2, mieux définies que celles du paramètre alpha. Le nuage de points d'U2 apparaît moins éparpillé que celui de la variable U1.

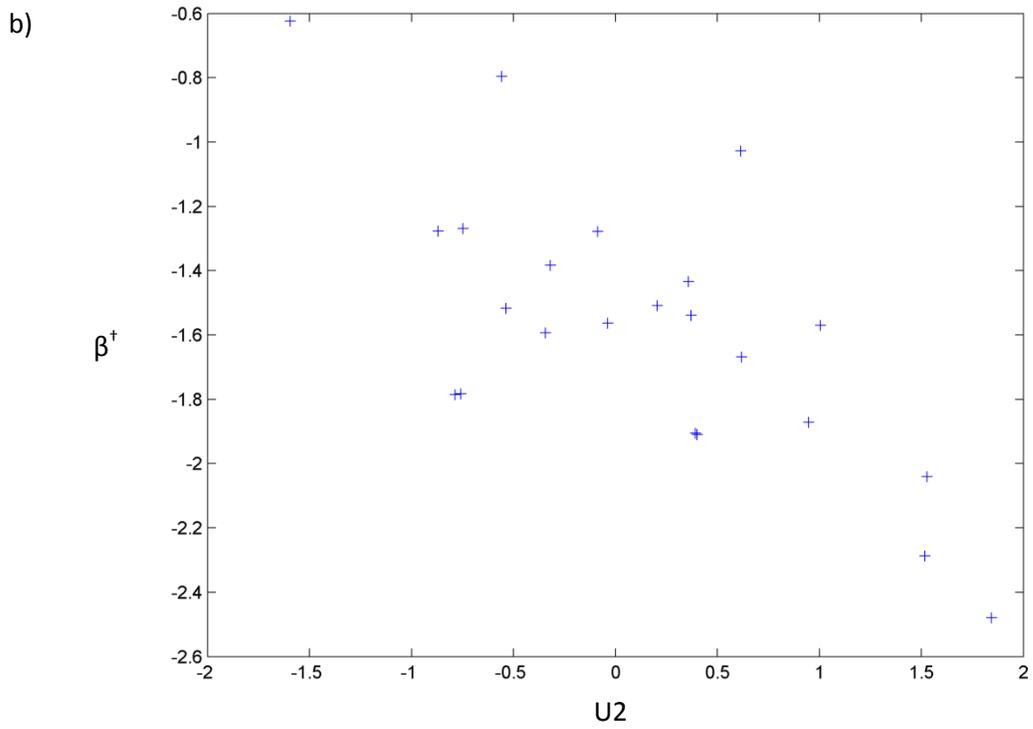
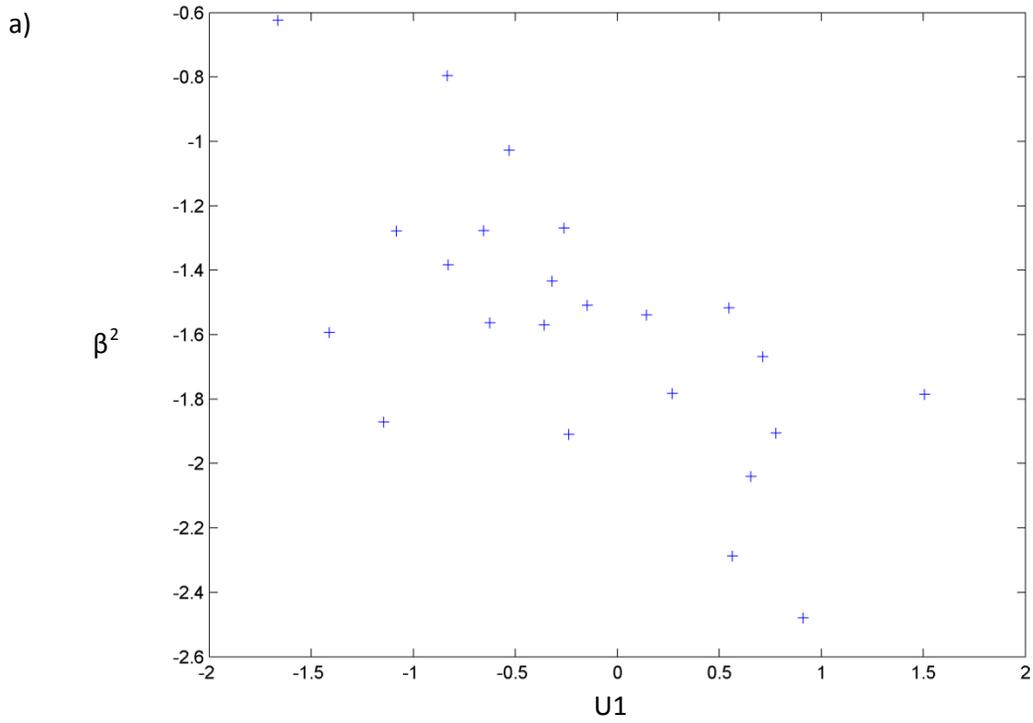


Figure 27 : Relation entre Beta et a) U1 et b) U2

² Beta transformé

Le même exercice a été réalisé pour estimer les paramètres des courbes de tarage du modèle polynomial (Tableau 21). Cependant, contrairement aux régressions pour le modèle exponentiel, ce ne sont pas toutes les variables canoniques qui ont été retenues dans les régressions. Les variables canoniques non retenues (ayant une valeur de 0 comprise dans l'intervalle de confiance) pour l'estimation d'un paramètre ont été mises en rouge. Ainsi, la variable U3 n'a été d'aucune utilité pour les paramètres du modèle polynomial. Il n'a donc pas été nécessaire de calculer cette dernière lors de l'application de la méthode de régionalisation.

Tableau 21 : Coefficients de régressions et ordonnées à l'origine pour l'estimation des paramètres Alpha et Beta du modèle exponentiel de courbe de tarage

	A ¹			B ¹			C ¹		
	Coefficient	Borne inf.	Borne sup.	Coefficient	Borne inf.	Borne sup.	Coefficient	Borne inf.	Borne sup.
Ordonnée à l'origine	-1,01	-1,41	-0,60	3,73	3,17	4,29	7,35	6,70	8,00
U1	0,22	-0,28	0,71	-0,62	-1,20	0,02	-1,47	-2,34	-0,74
U2	0,60	0,14	1,07	0,84	0,21	1,45	0,85	0,12	1,62
U3	0,00	-0,50	0,51	0,07	-0,63	0,77	0,17	-0,65	0,99

La Figure 28 illustre la relation entre le paramètre A du modèle polynomial et la variable canonique U2 sélectionnée. Une tendance positive est observée.

La Figure 29 illustre la relation entre le paramètre B du modèle polynomial et la variable canonique U2. Tout comme pour le paramètre A, le paramètre B a une tendance positive avec U2.

Enfin, la Figure 30 a) et b) illustre les relations entre le paramètre C et les variables canoniques sélectionnées dans la régression. Une relation inverse est notée en a) entre le paramètre C et la variable U1 alors qu'une relation négative est observée avec U2. Le nuage de points entre C et U1 s'est avéré mieux défini que celui avec U2.

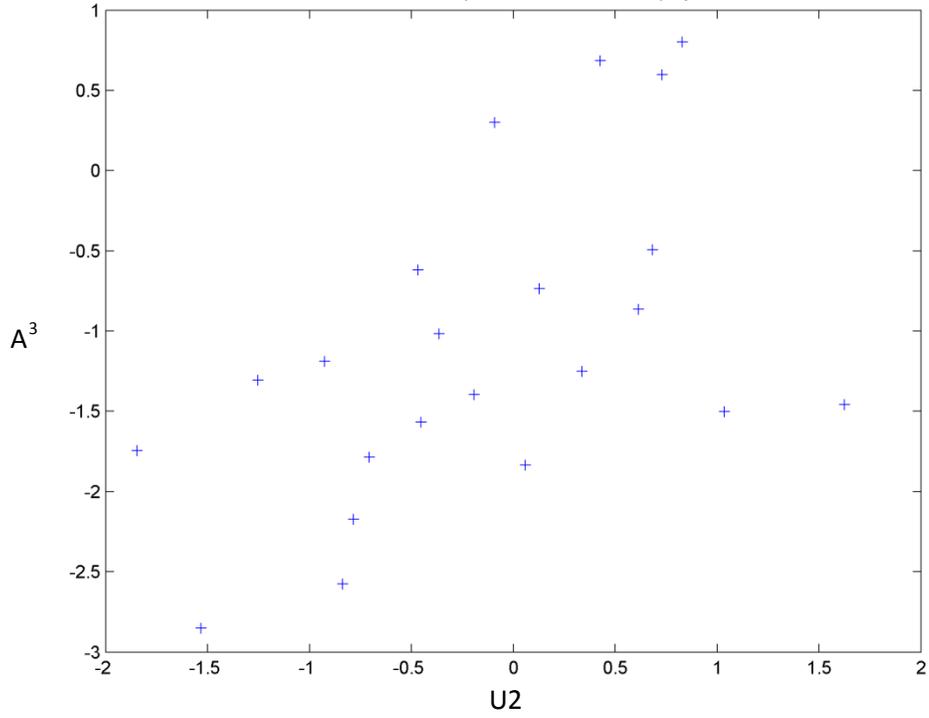


Figure 28 : Relation entre le paramètre A et U2 pour le modèle polynomial

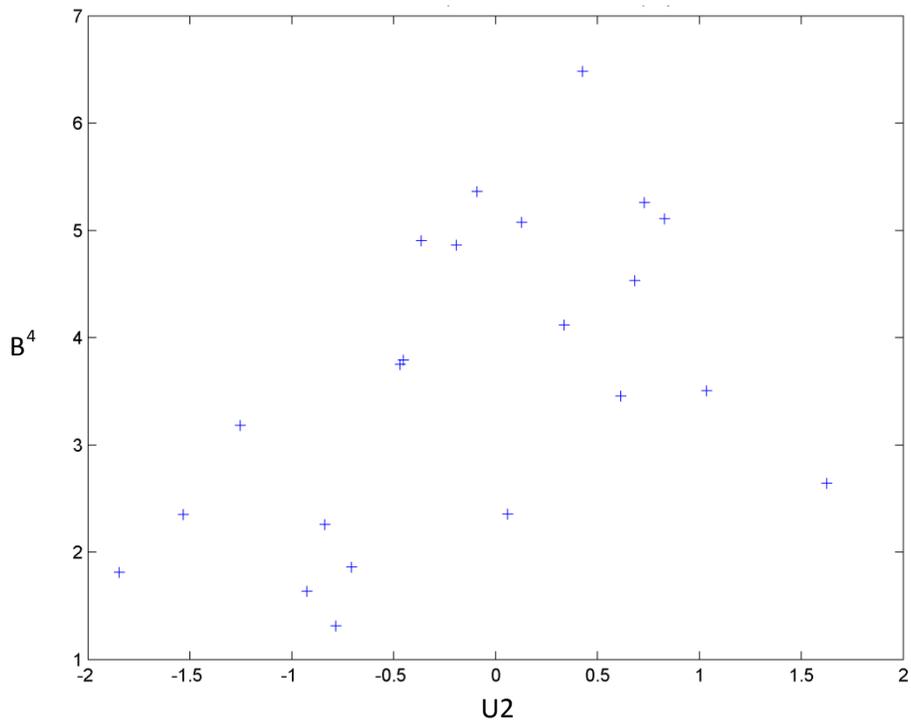
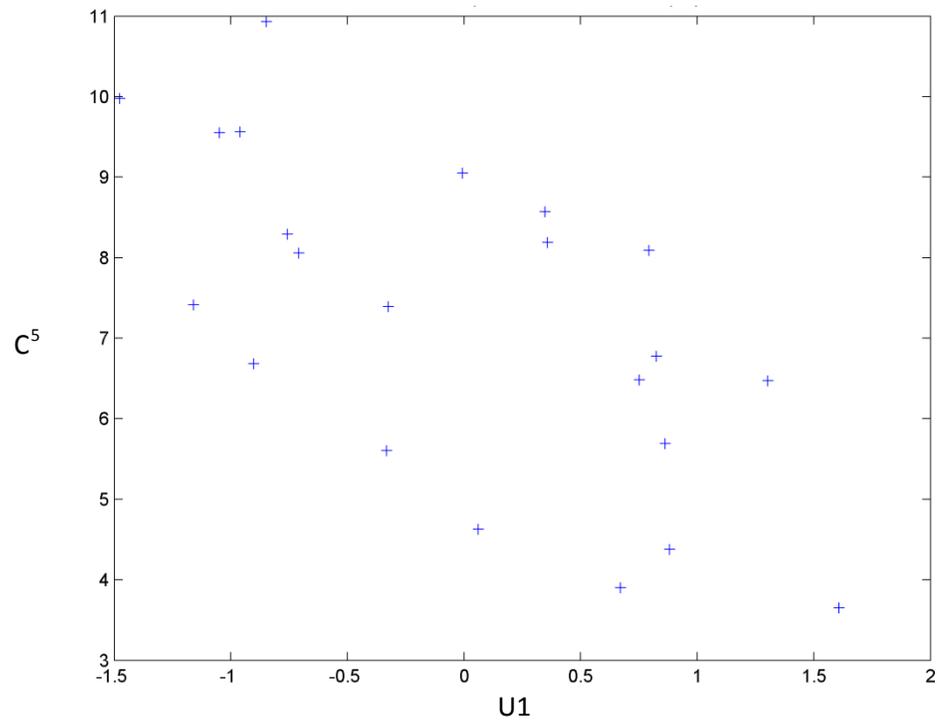


Figure 29 : Relation entre le paramètre B et U2 pour le modèle polynomial

³ Valeurs transformées

⁴ Valeurs transformées

a)



b)

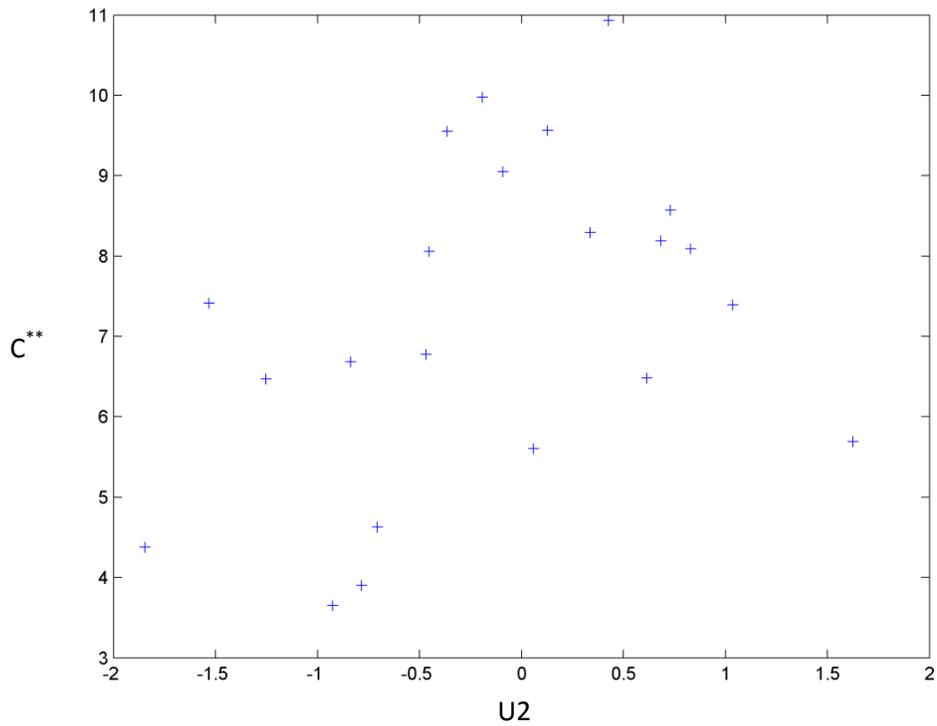


Figure 30 : Relations entre le paramètre C et a) U1 et b) U2

⁵ Valeurs transformées

L'analyse de la qualité de l'ajustement des régressions pour l'estimation des paramètres des courbes de tarage est présentée au Tableau 22. Trois indicateurs ont été identifiés; soit le coefficient de détermination, sa *p-value* ainsi qu'une estimation de l'erreur de la variance. Cette dernière est difficilement comparable d'un paramètre à un autre puisque les unités diffèrent. Toutefois, lors de la calibration de chacune des régressions, le meilleur modèle choisis était celui qui a comporté l'erreur la plus faible. Le coefficient de détermination a permis de comparer la qualité des modèles de régressions développés entre chaque paramètre. Dans le cas présent, les modèles n'ont pas fourni les mêmes performances. Ainsi la qualité d'estimation des paramètres suit l'ordre décroissant suivant $Beta > C > Alpha > A > B$. Tous les modèles se sont avérés significatifs avec une *p-value* n'excédant jamais le seuil de signification de 5%. En termes de choix de modèle, le modèle exponentiel s'est avéré un peu mieux calibré que le polynomial. Il avait été remarqué au Tableau 13, que le nombre de variables corrélées avec les paramètres du modèle polynomial était moins élevé et par conséquent, il était difficile de trouver une bonne relation entre les paramètres du modèle et les variables physiographiques. Bien que le modèle exponentiel fût le meilleur, la qualité de la calibration des équations a été moyenne. Avec 25 stations au départ, dont 2 ayant été retirées au cours de l'analyse canonique de corrélation, le nombre d'échantillons (nombre de courbe de tarage pouvant être utilisées pour la régionalisation) s'est avéré la plus grande limitation. Ceci a également imposé une restriction quant au nombre de variables physiographiques pouvant être retenues même si plusieurs présentaient un bon potentiel (5 au lieu de 9 et 3 au lieu de 6, respectivement pour les modèles exponentiel et polynomial).

Tableau 22 : Qualité de l'ajustement des régressions multiples réalisées pour l'estimation de paramètre de courbes de tarage

Paramètre	R ²	p-value	Estimation de l'erreur de la variance
Alpha	0,42	0,004	38,94
Beta	0,68	0,00001	0,066
A	0,28	0,01	0,75
B	0,26	0,02	1,65
C	0,57	0,0004	1,97

Suite à ces résultats, il apparaissait évident que la régionalisation des courbes de tarage à pentes décroissantes ne pouvait s'effectuer en raison de la faible taille de l'échantillon. En effet, avec 11 courbes de tarage au départ et en présumant qu'aucune données ne sont contrastantes, il y aurait au maximum 2 variables physiographiques utilisables pour le modèle exponentiel et 1 pour le modèle polynomial, ce qui est très limitant.

4.3 Validation

La validation de la méthode s'est effectuée à partir d'un jeu de données indépendant, c'est-à-dire n'ayant pas servi à la calibration. Lors de ce projet, certaines rivières ont été écartées de l'analyse puisque que les images disponibles ne couvraient pas une assez grande plage de débit par rapport à la dynamique enregistrée à la station. Ces stations et leurs images ont servi à la validation puisque des mesures de débits instantanés sont disponibles. Le nombre total de stations disponibles pour la validation était de 38.

La première étape de la validation consistait à fixer la distance d'intégration. Il a été décidé d'utiliser la distance d'intégration moyenne (3 km) obtenue à partir des meilleurs modèles de courbes de tarage sélectionnés dans l'étape de calibration.

Ensuite, un jeu de stations hydrométriques indépendant (non utilisées dans la calibration des modèles) a été sélectionné. Les variables prédictives utilisées au niveau du bassin versant ont été récupérées de la base de données « bassin versant ». Les nouvelles variables prédictives, comme Wn_DI et les classes de couverture EOSD, ont été calculées à la distance d'intégration moyenne préalablement établie.

Afin d'identifier les stations de validation présentant des modèles croissants et décroissants, les 20 classes de couverture EOSD ont été converties en composantes principales à partir des coefficients obtenus à la calibration et les moyennes des classes de EOSD à la calibration. Une fois ces composantes principales calculées, la règle de décision CART a été appliquée. Ainsi, 29 stations se sont avérées classifiées à pentes croissantes et 9 ont été classifiées à pentes décroissantes.

L'étape suivante consistait à appliquer les transformations pour respecter la normalité (Tableau 15 et Tableau 16) à chacune des variables explicatives (s'il y a lieu) du jeu de variables de validation. Ensuite, les variables canoniques pour le jeu de données de validation ont été calculées pour chaque modèle, en utilisant les coefficients canoniques ainsi que les moyennes des variables obtenues à la calibration. La séquence a été réalisée deux fois, soit une pour le modèle exponentiel et la seconde pour le modèle polynomial.

L'ACC est une méthode de régionalisation qui se base sur l'interpolation. Il était donc essentiel que le jeu de données de validation soit inclus dans le domaine d'application du jeu de calibration. Si ce n'avait pas été le cas, le modèle aurait été en situation d'extrapolation et la qualité de la performance en aurait été affectée et peu fiable. Il était impératif d'éliminer les rivières dont les variables prédictives physiographiques causaient la situation d'extrapolation.

Pour ce faire, les stations présentant des variables canoniques ayant une valeur en absolue excédant 2 ont retirées. À la calibration, ce seuil avait été utilisé pour retirer les données aberrantes ou éloignées. Le Tableau 23 présente les variables canoniques calculées à chacune des stations pour chaque type de modèle. Les cas d'extrapolations ont été mis en rouge. Ainsi,

22 stations sur les 29 ont été utilisées pour la validation du modèle exponentiel alors que seulement 14 des 29 l'ont été pour la validation du modèle polynomial.

Tableau 23 : Variables canoniques générées pour les deux types de modèles de courbes de tarage à partir des données de validation

Station	Modèles exponentiels		Modèles polynomiaux	
	U1	U2	U1	U2
02JB009	2.7	1.0	-2.7	-0.9
02OC021	2.7	-2.7	-2.1	0.2
02OD003	0.4	-0.4	-0.5	-0.6
02PA007	1.1	0.6	-1.3	-1.0
02PC010	3.5	-3.4	-2.7	0.8
02PG001	-0.4	0.7	0.4	0.3
02PJ007	-0.5	0.2	0.8	-0.5
02RB004	1.2	-0.1	-1.3	-0.4
02RH045	0.3	-1.4	0.1	-2.5
02UC002	1.3	0.3	-2.0	0.2
02VC001	1.8	1.4	-2.0	-2.7
02WB003	1.9	0.7	-2.5	-1.8
02XA003	1.6	0.2	-1.7	-1.2
02XA008	1.2	0.9	-1.8	-0.8
02YC001	1.4	-1.7	-1.2	-0.2
02YO011	1.1	0.7	-1.8	0.1
02YQ001	0.5	1.2	-0.8	-0.8
02ZE004	2.7	-1.0	-1.9	-0.8
03AB002	2.7	1.8	-3.0	-0.6
03BD002	1.8	1.5	-2.0	-4.1
03ED001	1.3	1.7	-2.0	-3.4
03FA003	2.5	1.8	-2.6	-4.5
03LF002	1.5	1.9	-2.3	-2.2
03NF001	1.5	0.4	-1.9	-1.6
03OC003	1.8	2.2	-2.0	-5.3
03OE001	1.6	1.9	-2.3	-3.5
03OE003	0.7	-0.3	-0.9	-0.2
03QC001	1.9	0.4	-2.0	-5.2
03QC002	-0.6	0.2	0.1	-1.6

Les estimations des paramètres des courbes de tarage exponentielles et polynomiales ont été générées à partir des variables canoniques et en appliquant les régressions. Les courbes de tarage ont été alors construites pour les rivières de validation. Il ne suffisait que de calculer la

largeur effective à partir des images couvrant les rivières, à la distance d'intégration de 3 Km, ensuite de calculer le débit estimé en appliquant ces courbes et de comparer ces valeurs avec le débit mesuré. Le Tableau 24 présente les paramètres de courbes de tarage pour les deux modèles à partir des données de validation.

Tableau 24 : Paramètres des courbes de tarage générés à partir des modèles régionalisés

Station	Beta	Alpha	A	B	C
02OD003	0.1332485	0.0000005	0.2280393	-24.8018742	1833.5
02PA007	0.0603582	0.0000036	0.1170984	-18.1111121	4264.8
02PG001	0.1126135	0.0096623	0.4475052	-54.5996209	1073.6
02PJ007	0.1471603	0.0009956	0.5815669	-28.0702676	309.4
02RB004	0.0865680	0.0000000	0.1098382	-29.2268452	7691.7
02RH045	0.2147277	0.0000000			
02UC002	0.0655359	0.0000003			
02VC001	0.0251116	0.0000172			
02WB003	0.0380072	0.0000001			
02XA003	0.0598976	0.0000000	0.0757747	-15.8467784	7380.6
02XA008	0.0490543	0.0000134	0.0690634	-20.6380040	11083.7
02YC001	0.1667902	0.0000000	0.1273896	-36.4564884	7547.8
02YO011	0.0572650	0.0000086	0.0713013	-45.7165411	23689.3
02YQ001	0.0573577	0.0029448	0.1698075	-20.8114885	2614.3
02ZE004			0.0628300	-21.0279121	12978.9
03BD002	0.0230021	0.0000264			
03ED001	0.0280299	0.0009427			
03LF002	0.0226387	0.0012361			
03NF001	0.0572457	0.0000002	0.0632903	-11.3033015	6840.5
03OE001	0.0200689	0.0005640			
03OE003	0.1175826	0.0000002	0.1543557	-34.2853809	5122.7
03QC001	0.0440858	0.0000000			
03QC002	0.1499318	0.0014561	0.3562426	-11.3255931	347.9

L'étape suivante visait à appliquer les équations régionalisées à chacune des stations et de les alimenter avec les largeurs effectives obtenues avec leurs images respectives. Une fois ceci réalisé, les débits estimés ont été comparés aux débits mesurés aux stations. La Figure 31 et la Figure 32 illustrent les graphiques de validation indépendante pour les modèles de régionalisation du débit instantané avec les fonctions exponentielles et polynomiales, respectivement. En observant les graphiques, il est possible de constater que les modèles d'estimation du débit pour les sites non-jaugés sont non-concluants. La régionalisation du modèle exponentiel sous-estime largement l'estimation du débit alors que le modèle polynomial la surestime largement. Dans le cas où les résultats auraient été plus prometteurs, la

prochaine étape aurait été de trouver la largeur effective à laquelle le modèle exponentiel aurait passé au modèle polynomial.

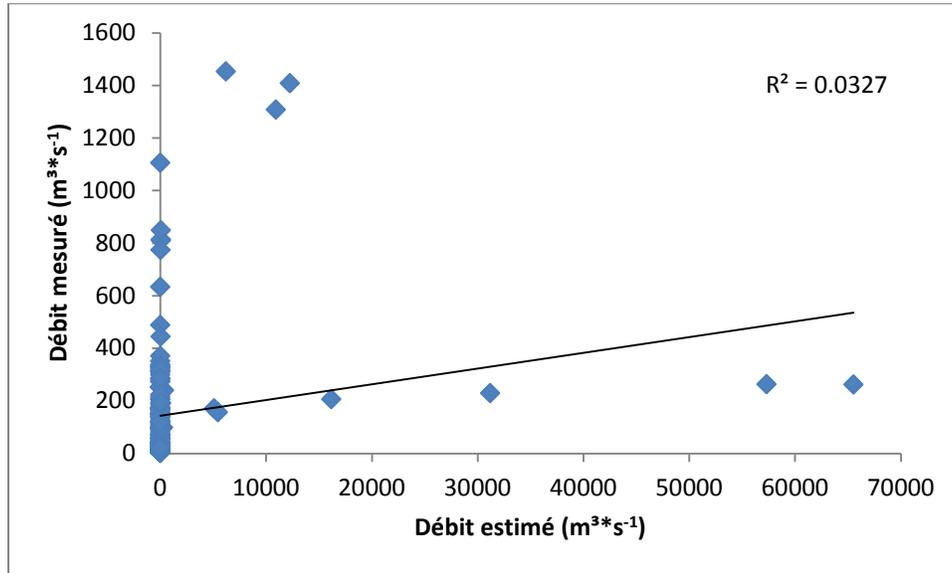


Figure 31 : Validation du modèle de régionalisation des courbes de tarage exponentielles

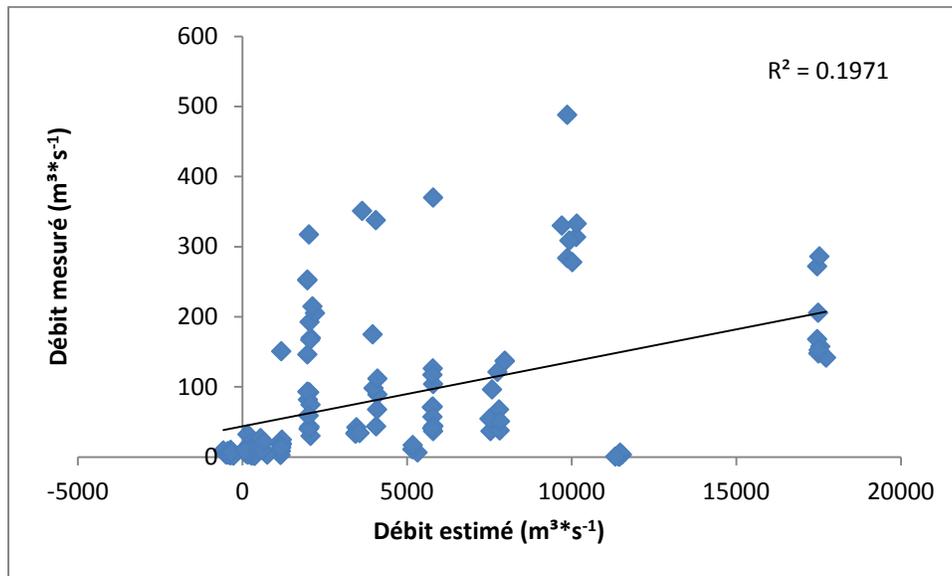


Figure 32 : Validation du modèle de régionalisation des courbes de tarage polynomiales d'ordre 2

La procédure de régionalisation des courbes de tarage intègre la propagation de l'incertitude du modèle du début jusqu'à la fin. Le fait d'avoir un échantillon de 25 stations ayant des courbes de tarage croissantes était limitant pour développer un modèle d'estimation pour les sites non-

jaugés, ce qui a affecté la qualité ainsi que la robustesse de la régionalisation. Ceci explique les résultats non-concluants pour cette partie du projet.

4.4 Automatisation de la procédure pour les sites non-jaugés

Une séquence partiellement automatique a été développée pour la détermination du débit à partir d'images Radarsat pour un site non-jaugé (Figure 33). Cette séquence a nécessité la programmation de plusieurs modules dont la plupart pouvant être intégrés dans ArcGIS. Cette séquence est présentée à l'annexe IV de façon détaillée. Les spécifications des exigences associées au système et aux interfaces sont présentées à l'annexe V. Dans cette annexe, on y retrouve des détails concernant certains logiciels ayant été utilisés uniquement pour la calibration et ne figurant pas dans la séquence automatisée pour estimer le débit d'un site non-jaugés.

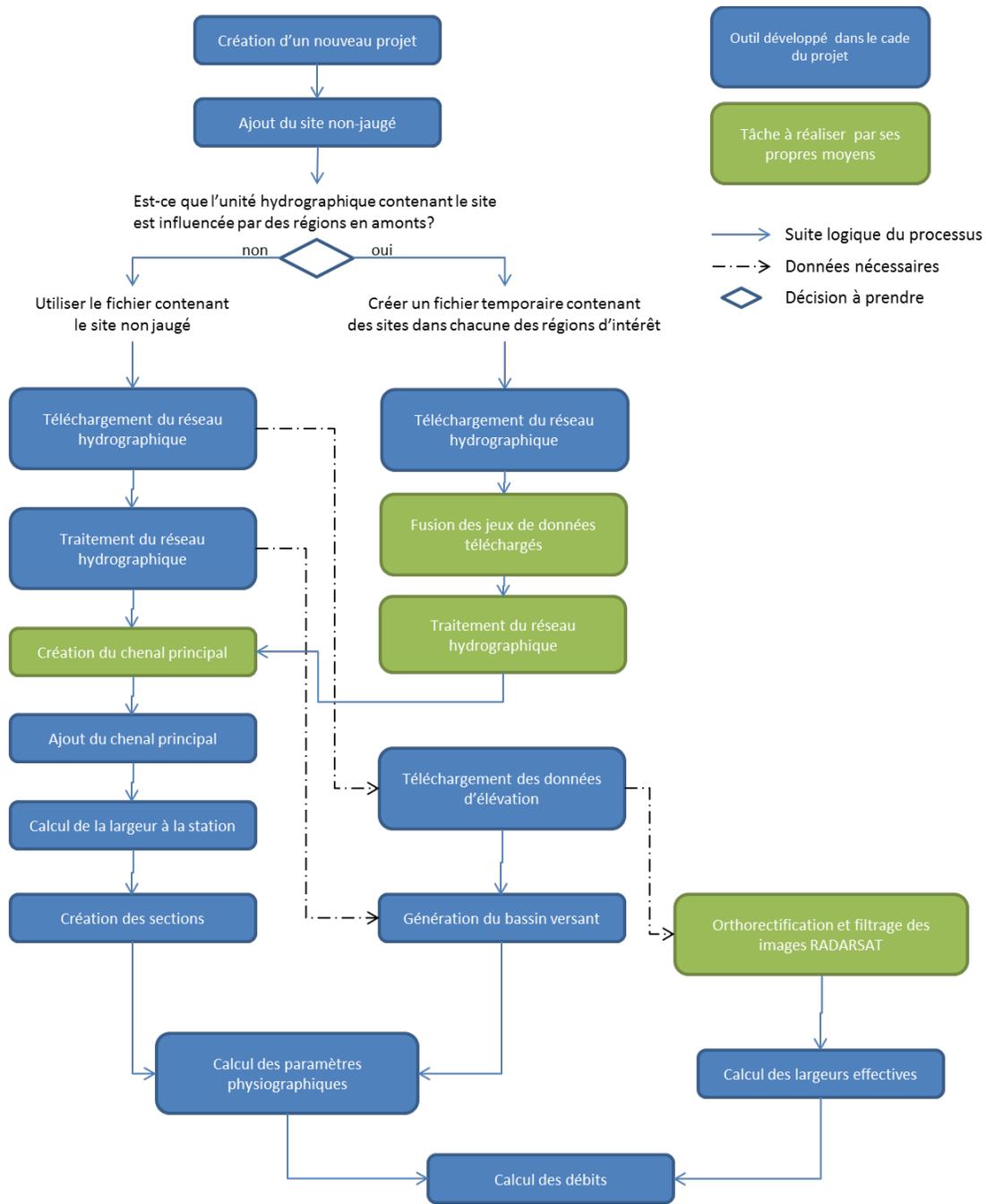


Figure 33 : Étapes pour le calcul de débits pour un site non-jaugé

5 Conclusions

Ce projet avait pour objectif de développer une méthodologie d'estimation du débit en rivière pour des sites non-jaugés à l'aide de l'imagerie RADARSAT dans l'Est du Canada. Pour tenter d'accomplir cet objectif, il a fallu développer une base de données intégrant des informations de différente nature (imagerie Radarsat, hydrométrie et physiographie). Avant d'arriver à l'étape du développement des courbes de tarage pour les sites jaugés, plusieurs sous-étapes non-planifiées ont dû être ajoutées pour cette phase de calibration. Ces sous-étapes concernaient l'orthorectification manuelle de certaines images, la validation manuelle du chenal principal de chaque rivière, l'intercalibration pour contrer l'effet des conditions d'acquisitions différentes, le paramétrage des seuils de la fonction de logique floue ainsi que le développement d'interfaces pour faciliter la production de courbes de tarage. Sans ces sous-étapes, le développement des courbes de tarage n'aurait pas pu être réalisé avec succès.

Lors du développement des courbes de tarage pour des sites jaugés, d'autres situations imprévues sont survenues. En effet, les courbes de tarage ne pouvaient pas mélanger des images provenant des deux orbites. Ceci a occasionné une fragmentation supplémentaire de la base de données puisque le nombre d'images utilisable pour construire une courbe de tarage a été diminué. Suite à ce constat, une analyse de la représentativité de la dynamique du débit du jeu d'images par rapport à la variabilité à la station s'est ajoutée. Ensuite, l'existence de courbes de tarage décroissantes s'est ajoutée aux imprévus. Ceci constituait un obstacle important à franchir avant la régionalisation. Il a donc fallu identifier la cause de ces fonctions (double-rebond avec des cibles dans la plaine inondable) et ainsi déterminer une méthode permettant de savoir dans quel cas une rivière non-jaugée se situerait. Ceci a impliqué une fragmentation supplémentaire de la base de données puisque deux modèles de régionalisations devaient maintenant être construits séparément. Ensuite, pendant le déroulement de ce projet, le type de fonction pour la courbe de tarage a été modifié pour une forme exponentielle, mieux adapté aux données que la forme logarithmique. Ceci a mené à la réalisation d'une analyse de sensibilité du débit aux paramètres des courbes de tarage et par le fait même d'introduire la courbe polynomiale d'ordre 2. La conclusion de cette analyse suggère d'utiliser la forme exponentielle pour les faibles largeurs et à l'intersection avec la forme polynomiale de transférer sur cette dernière.

Avant même de débiter le développement du modèle d'estimation d'un site non-jaugé, la taille de l'échantillon était déjà problématique d'un point de vue statistique. Malgré cet échantillon à taille réduite, des corrélations significatives avec les paramètres des courbes de tarage et plusieurs variables physiographiques ont été obtenues, prouvant ainsi que les variables physiographiques peuvent être utilisées pour la régionalisation. Pour la suite des choses, l'analyse canonique de corrélation et les régressions n'ont que fait ressortir le phénomène de propagation de l'incertitude dans la chaîne de traitement en raison de la fragmentation de la base de données. Enfin, toute la chaîne de traitement pour l'estimation du débit d'une rivière non-jaugée a été automatisée.

Néanmoins, même si les résultats du modèle d'estimation du débit d'une rivière non-jaugée se sont avérés non-concluants, il s'avère que le concept de station hydrométrique virtuelle serait une alternative prometteuse. Ainsi, en suivant les recommandations de ce projet (voir la section recommandations), il serait possible de construire une courbe de tarage satellitaire à partir de mesures de débits ponctuelles au moment de l'acquisition de l'image à des périodes clés. Une fois que la courbe de tarage est bien calibrée, de nouvelles images sont utilisées pour en estimer le débit sans aucune mesure *in situ*.

6 Recommandations

Méthode d'ensemble ou à plusieurs estimateurs

Dans des conditions de modélisation complexes, telles que l'estimation du débit pour les sites non jaugés par rétrodiffusion radar, il est souvent recommandé d'employer une approche de méthode d'ensemble ou à plusieurs estimateurs. Ceci permet de réduire l'impact de la sensibilité ou de la propagation d'erreurs d'une méthode basée sur un seul estimateur. Cette technique, utilisable dans le contexte de classification et/ou estimation, est plus robuste lorsque l'erreur d'estimation dans la modélisation est élevée.

Intercalibration des images

L'intercalibration a permis de ramener les images à une base comparable afin de corriger la dépendance de la radiométrie des images aux conditions d'acquisition. Sans cette procédure, il était impossible d'appliquer aux images Radar une classification par logique floue avec une fonction d'appartenance de type Z en utilisant des seuils fixes. En effet, sans l'intercalibration, les bornes A et B pour l'application de la fonction d'appartenance changeaient d'une image à l'autre, venant ainsi fausser le calcul de la probabilité d'eau. Malgré cette procédure, la qualité des courbes de tarage développées demeure meilleure lorsque l'ensemble des images ont des configurations d'acquisition semblables. La méthode ne corrige pas parfaitement le signal surtout pour les effets d'ombres portées et du double-rebond dues aux angles d'incidences plus élevés.

La stratégie utilisée pour intercaler les images consistait à ramener le premier percentile de l'histogramme de la signature des images contenue dans la rivière et sa zone tampon équivalente à 100% de sa largeur à une valeur de référence commune. Or, après coup, il a été constaté que l'intercalibration rendait certaines images inutilisables pour des rivières généralement de faible largeur (autour de 40m) en conditions d'étiage. Puisque les signatures des pixels d'eau pure devenaient absentes ou très peu abondantes, il arrivait que le premier percentile de l'histogramme de ces images soit un mixel, ce qui provoquait une surestimation des probabilités de pixels d'eau. Les pistes de solutions qui devraient être testées seraient d'effectuer l'intercalibration sur le premier percentile de l'image entière ou encore sur le premier percentile de la signature de toutes les étendues d'eau contenues dans l'image. L'objectif étant d'accroître le nombre de pixels d'eau pure extrait pour en construire un histogramme ayant les mêmes conditions d'acquisition de l'image. L'amélioration de l'intercalibration pourrait permettre d'augmenter le nombre d'images et par le fait même de stations utilisables.

Accroître le nombre de courbes de tarage retenues pour la calibration

Les résultats non-concluants de la conception du modèle d'estimation du débit pour un site non-jaugé sont principalement expliqués par la taille de l'échantillon des courbes de tarage retenues. Ainsi, pour augmenter le nombre de stations (courbes de tarage retenues), un plus grand

nombre de d'images représentatives de la dynamique du débit de la rivière est nécessaire pour obtenir une qualité d'ajustement de la courbe acceptable ($R^2 > 0,5$). Ainsi, l'amélioration de la qualité d'ajustement est dépendante du nombre d'images disponibles. La recommandation d'améliorer l'intercalibration énumérée précédemment permettrait d'améliorer la qualité et la quantité de courbes de tarage. Il serait possible de réviser les archives d'images en éliminant la contrainte quant au nombre maximal d'images pouvant être utilisées. Toutefois, les stations avec peu d'images au départ n'en seront pas bonifiées. Les nouvelles images acquises depuis le printemps 2013 à aujourd'hui pourraient être intégrées si jamais il y avait une suite au projet. De nouvelles acquisitions pourraient également être réalisées sur les stations non retenues ayant un faible nombre d'images. En ce sens, les « background missions » permettraient d'acquérir des images sur plusieurs stations, ce qui serait bénéfique pour l'ensemble d'entre elles. Enfin, il serait envisageable d'intégrer le territoire de l'Ontario dans le projet pour accroître le nombre de stations, si l'intérêt pour l'utilisation d'images d'archives est prédominant. Évidemment, la disponibilité de la donnée de débit doit être considérée pour toute commande d'images (archives et nouvelles acquisitions). Il est à noter que certaines stations utilisées pour ce projet peuvent être actuellement fermée, temporairement hors service ou en service, d'où l'importance de vérifier la disponibilité de la donnée de débit. Bien que le décompte n'ait pas été réalisé, il est arrivé plusieurs fois qu'une image disponible pour une station, n'a pas pu être commandée faute de mesure du débit.

La distance d'intégration

Un autre élément à analyser est la distance d'intégration. Des changements significatifs ont été constatés dans les courbes de tarage d'une section à l'autre pour un même cours d'eau en raison du changement dans la largeur nominale entre les sections.

Le calcul de largeur effective dépend directement de la qualité de l'estimation de la probabilité des pixels d'eau. Le concept de distance d'intégration a été introduit pour que le calcul de la largeur effective s'effectue à partir du cumule des probabilités d'eau d'une section à l'autre. La réalité est que les rivières n'ont pas la même variabilité d'une section à une autre. Certaines sections de rivières en amont de la station sont plus encaissées (pentes fortes et lit étroit) d'autres moins (pentes faible et lit large), d'autres comportent des plages, des îles ou bancs de sable, des éléments de nature anthropique (ponts, aménagements urbains, lignes à haute tension, etc.). Il arrive aussi que des rivières comportent des sections homogènes et que la largeur nominale de la rivière varie très peu. Ainsi il est quasi-impossible de fixer une distance d'intégration pour la calibration/validation des modèles. La problématique est que les sections ayant peu de variabilité de largeur d'eau selon le débit sont quand même intégrées au calcul de la largeur effective.

Il aurait été intéressant d'effectuer le calcul de manière non-cumulée d'une section à la suivante pour voir quel section présente la meilleure dynamique avec le débit à la station. Cette recherche aurait pu s'effectuer de la station jusqu'au premier tributaire en amont étant donné

que le débit peut être considéré équivalent. Par la suite, il aurait été possible d'établir une règle pour identifier les zones où le calcul de la largeur effective est utile considérant les caractéristiques physiographiques les plus prometteurs. Dans ce cas-ci, il sera important de tenir compte de la largeur nominale des sections choisies pour l'intégrer au modèle de régionalisation.

Tester la régionalisation des modèles sans confondre les courbes de tarage d'orbite différente

Les courbes de tarage ont été calibrées avec les images d'une seule orbite puisque la qualité des résultats en était nettement améliorée. Toutefois, puisque le nombre de courbe de tarage retenues était réduit en raison des modèles croissants/décroissants, la régionalisation des modèles a été réalisée en intégrant les meilleures courbes de tarage peu importe l'orbite à laquelle elles ont été développées. L'impact est probablement moins marqué que l'inclusion des images des deux orbites pour calibrer une courbe de tarage. Il aurait été intéressant de savoir s'il y avait un effet sur la régionalisation. Dans une telle analyse, l'influence de l'angle entre la rivière et l'orbite du satellite pourrait être évaluée.

Concept de la station hydrométrique virtuelle

Bien que les résultats de la validation du modèle d'estimation du débit pour un site non-jaugé se sont avérés non-concluants, il en est tout autrement pour les sites jaugés. En effet, les courbes de tarage calibrées sur les rivières jaugées ont confirmées le potentiel de l'utilisation de l'imagerie Radarsat pour estimer le débit d'une rivière. Ainsi, il serait possible de construire des courbes de tarage satellitaire sur des rivières actuellement non jaugées en procédant à des mesures de débits ponctuelles simultanément à l'acquisition des images. Pour ce faire, des images Radarsat-2 seraient acquises à des moments clés durant la période où les eaux sont dégelées de manière à couvrir la gamme de débits entre les périodes de crues et d'étiage. Un jaugeage de la rivière serait réalisé au moment de l'acquisition de chaque image afin d'obtenir le débit correspondant. Ensuite, la largeur effective de la rivière serait calculée sur chacune des images afin de produire des couples image-débit permettant l'établissement de la courbe de tarage.

Les images seraient acquises avec les mêmes paramètres d'acquisition (orbite et angle d'incidence) pour limiter les corrections à appliquer. Il faudrait favoriser les angles d'incidence plus faibles puisque les rivières et leur environnement en sont moins affectés par les ombres portées et le phénomène de double rebond de l'onde. La courbe de tarage pourrait être construite au fur et à mesure que les couples images-débits sont acquis et traités. À la fin de la saison ou lorsque la courbe de tarage sera jugée acceptable, elle constituera le cœur de la station hydrométrique virtuelle. De nouvelles images pourront être acquises dans les mêmes conditions que celles ayant servie à la calibration mais cette fois-ci, aucune mesure de débit ne sera acquise simultanément. En appliquant l'équation de la courbe de tarage sur la largeur effective obtenue sur l'image, une mesure de débit serait alors obtenue.

Par ailleurs, les courbes de tarage développées au cours de ce projet pourraient éventuellement devenir des stations hydrométriques virtuelles si les stations hydrométriques actuelles venaient qu'à être fermées définitivement ou temporairement pour des raisons d'entretien.

7 Divulgence de la propriété intellectuelle par l'entrepreneur

La propriété intellectuelle sur les renseignements de base (BIP) nécessaire à la réalisation ou au soutien du projet, de même que la propriété intellectuelle sur les renseignements originaux (FIP) qui découlent des travaux réalisés sont illustrées dans le Tableau 25.

L'INRS restera propriétaires des droits sur la propriété intellectuelle développée dans le cadre de ce projet.

Tableau 25 : Divulgence de la propriété intellectuelle sur les renseignements originaux (BIP) requis pour utiliser/améliorer la FIP dans le cadre du projet

Élément de la BIP	Titre la BIP	Type de PI (logiciel, concept, algorithmes)	Type d'accès à la BIP requis pour utiliser/améliorer la FIP	Description de la BIP	Document de référence (rapport, document de conception)	Origine de la BIP (projet R-D interne,)	Propriétaire de la BIP (entrepreneur, sous-traitant)
1	Quadrillages des moyennes climatiques à long terme pour le Canada et les États-Unis (accès sur demande)	Base de données	Licence d'utilisation sans restriction	Données climatiques historiques mensuels pour les années 1901-2003 couvrant le Canada et les États-Unis. Elles sont spatialement continues et ont été cartographiées à une résolution d'environ 10 km. Les données climatiques du Canada et des États-Unis ont été produites à l'aide des algorithmes de lissage par plans affinés (ANUSPLIN) mis au point par Michael Hutchinson (de la Australian National University). Ce travail en partenariat a été effectué par M. Hutchinson, ainsi que par les divisions des Stratégies de surveillance et de gestion des données et de la Recherche climatique d'Environnement Canada, de même que par Custom Climate Services Inc.	McKenney, et al. (2011). Bulletin of American Meteorological Society-BAMS December: 1612-1622. Hopkinson, et al. (2011). Journal of Applied Meteorology and Climatology 50: 1654-1665. Hutchinson, M. F. et al. (2009). American Meteorological Society(April): 725-741.	Projet R-D	Ressources naturelles du Canada

8 Évaluation du succès global du projet

8.1 Évaluation des résultats du projet par rapport aux indicateurs de rendement

Corrections géométriques des images

La totalité des images reçues (non-corrompues et de bonne qualité radiométrique) pour ce projet ont été orthorectifiées avec succès, selon le degré de précision requis. Le degré de précision était plus ou moins cinq pixels au maximum, évalué à la station hydrométrique et à la 5^e section de la rivière en amont de la station. La correction automatique des images Radarsat-2 à partir des paramètres orbitaux a été satisfaisante dans 93% des cas. Il en est tout autrement pour les images Radarsat-1 ayant des paramètres orbitaux beaucoup moins précis. En effet, environ 20% des images Radarsat-1 ont été recorrigées manuellement jusqu'à ce qu'elles soient acceptables.

Détection des pixels d'eau par logique floue

Cet objectif a été accompli en sa totalité. La méthode de détection des pixels d'eau utilise une fonction de logique floue intitulée *Z-Membership function*. Pour appliquer une telle fonction, il a fallu introduire une procédure d'intercalibration des images permettant de faire concorder le premier percentile des histogrammes de rétrodiffusion de l'eau à une valeur de référence en décibel. L'intercalibration a permis de corriger l'effet du gain du signal du au changement d'angle d'incidence d'une image à autre. Grâce à cette procédure, les seuils A et B, nécessaires à l'application de la fonction d'appartenance, sont devenus universels c'est-à-dire fixes peu importe l'angle d'incidence des images. La fonction a été appliquée pour les pixels des images situés à l'intérieur du corridor élargi à 100% de la valeur de la largeur à la station. Plusieurs éléments présents dans le corridor de calcul peuvent affecter positivement ou négativement la détection des pixels d'eau pour une rivière. En effet, la présence d'îles, de bancs de sable et de plages ont favorisé la détection de l'eau dans la zone d'intégration. D'autre part, la présence de rapides, d'infrastructures routières, de lignes à hautes-tensions, de canalisations et tout autre milieu de nature anthropique dans la plaine inondable nuisent à la détection de l'eau. La méthode a fonctionnée pour la plupart des rivières disponibles pour le projet.

Calcul de la largeur effective (W_e)

La largeur effective a pu être calculée à chacune des images disponibles pour l'ensemble des rivières à partir des probabilités d'eau calculées. Le calcul a été réalisé à l'intérieur de chaque section cumulée de 1 km du corridor élargi de la rivière.

Développement de courbes de tarage satellitaires

Des courbes de tarage satellitaires ont été générées pour l'ensemble des stations disponibles sur une série de distances d'intégration. La qualité des courbes de tarage développées dépend

de la bonne détection des pixels d'eau pour la distance d'intégration où la courbe est calculée. De plus, pour toutes les rivières à l'étude, la qualité des courbes de tarage a été grandement améliorée en ne sélectionnant que les images acquises dans une même orbite. En considérant une seule orbite, les effets d'ombres portées dans une même direction, généralement confondues avec l'eau, sont systématiques alors qu'en incluant les deux orbites celles-ci deviennent aléatoires. La qualité des courbes de tarage dépend également du nombre de paires image-débit disponible et de leur représentativité en matière de dynamique du débit enregistré à la station depuis son ouverture. Les courbes de tarage ont fonctionné tant pour celles développées pour l'estimation du débit instantané que celle du débit quotidien. Cependant, les modèles de courbes de tarage pour le débit instantané se sont avérés légèrement meilleures que celles du débit quotidien. La qualité de l'intercalibration a eu un impact sur le nombre de paires image-débit disponibles pour la construction des courbes de tarage. En effet, certaines rivières de faibles largeurs se sont retrouvées en conditions d'étiage faisant en sorte que la méthode ramène un histogramme contenant peu ou pas de pixels d'eau purs au premier percentile. Ceci a occasionné une surestimation de la largeur effective et donc un retrait de ces images aux courbes de tarage. Les courbes de tarage ont été développées en utilisant une fonction exponentielle venant ainsi remplacer la forme de puissance qui était moins bien ajustée aux paires image-débit. Une forme polynomiale d'ordre 2 a aussi été générée pour être utilisée conjointement avec la fonction exponentielle. Enfin, deux types de relations ont été identifiées; soit les courbes à pente croissante (de signe positif) et celles à pentes décroissantes (de signe négatif). Intuitivement, les courbes de tarage devraient être croissantes selon le concept de changement de largeur effective en fonction du débit. Toutefois, la présence de milieux humides, d'arbuste et d'arbres sur la plaine inondable engendre le phénomène de double-rebond du signal se traduisant par une réduction la largeur effective. Pour ces modèles, la dynamique de l'eau induit une dynamique du double-rebond perceptible sur les images et la courbe de tarage de la rivière en question.

Développement d'une méthode de détection du signe de la pente d'une courbe de tarage

Bien que les modèles de courbes de tarage à pente décroissantes ne figuraient pas dans la proposition initiale du projet, une méthode de détection du signe de leur pente s'imposait pour arriver à régionaliser les modèles. Pour résoudre ce problème, le produit *EOSD land cover* a été employé pour identifier les pourcentages de couvertures des 20 classes présentes dans le corridor élargi des rivières utilisé pour le calcul de la largeur effective. Une analyse en composantes principales a été réalisée sur les données extraites des rivières. Puis un arbre de décision a été développé à partir des 2 premières composantes principales permettant de déterminer le signe de la pente d'une courbe de tarage d'un site non-jaugé.

Étude des relations entre les paramètres des courbes de tarage et les variables physiographiques

La corrélation simple a été utilisée pour étudier la relation entre les variables physiographiques et les paramètres des courbes de tarage (pentes croissantes et débit instantané). Des 88 variables physiographiques disponibles, 33 se sont avérées significativement corrélées avec les paramètres des courbes de tarage. Sur ces 33, en excluant les variables physiographiques fortement corrélées entre-elles, 12 ont ressorties comme étant pertinentes à inclure dans le processus de régionalisation. Ce résultat démontre l'existence d'un lien véritable entre certaines variables physiographique et les paramètres des courbes de tarage. Le nombre de variables physiographiques corrélées avec les paramètres des courbes de tarage à fonctions exponentielles est plus important que ceux corrélés avec les paramètres des fonctions polynomiales.

Données de débit (Q) et largeurs effectives (We) couvrant une large espace d'extrapolation dans le modèle physiographique

Cet objectif n'a pas été atteint en raison du nombre limité de stations utilisées pour la calibration du modèle de régionalisation des courbes de tarage. La base de données images-stations (débits disponibles) a été plus fragmentée que ce qui avait été anticipé.

Développement du modèle d'estimation du débit pour les sites non-jaugés

Le modèle d'estimation du débit pour les sites non-jaugés a été construit en sa totalité pour les courbes de tarage à pentes croissantes. En effet, l'analyse canonique a permis de construire deux variables canoniques exploitant le maximum d'information contenu dans les variables physiographiques en lien avec les paramètres des courbes de tarage. Cette analyse a été menée deux fois; soit une première analyse pour les paramètres des courbes exponentielles et une seconde fois avec les paramètres des modèles polynomiaux. Ensuite, des régressions multiples avec les variables canoniques ont été développées pour estimer chaque paramètre des courbes de tarage. La chaîne de traitements nécessaires pour l'accomplissement de cette tâche a été réalisée en sa totalité en termes d'étapes à franchir. Cependant, en raison du faible nombre de stations présentant des courbes de tarage à pente décroissante, le modèle d'estimation n'a pas pu être complété. Ainsi, lors de l'application de ce modèle, si la méthode de détection du signe de la pente des courbes de tarage indique une pente à signe positif, le modèle peut être appliqué. Dans le cas d'une détection d'une pente à signe négatif, le modèle d'estimation du débit serait inexistant. Cette tâche pourrait être qualifiée de réussie à 75% étant donné que pour obtenir un modèle d'estimation à pente de signe négatif il ne suffirait que de reprendre la chaîne de traitements développée en y ajoutant plus de courbes de tarages à pentes décroissantes.

Validation du modèle d'estimation du débit pour les sites non-jaugés à partir de données indépendantes

Tout comme pour le développement du modèle d'estimation du débit pour les sites non-jaugés, sa validation a été réalisée en sa totalité. Les résultats de cette validation ont été non-concluants. En effet, les valeurs de débits produites dépassaient les limites plausibles pour une rivière. L'analyse de sensibilité des fonctions avait révélé qu'une simple variation de 1% des paramètres des courbes de tarage engendrait une forte variabilité du débit estimé, particulièrement pour la courbe exponentielle. Bien que le débit estimé par la fonction polynomiale se soit montré moins sensible à la variation des paramètres de la courbe, les résultats n'en ont pas été meilleurs. En effet, les régressions pour la prédiction des paramètres de la courbe polynomiale ont été de moins grande qualité que celles obtenues pour les paramètres de la courbe exponentielle. Le modèle d'estimation du débit pour un site non jaugé intègre la propagation de l'incertitude de l'ensemble des traitements générés en amont. Même à 25 échantillons initialement disponibles pour la régionalisation des paramètres des courbes de tarage à pentes croissantes, ce nombre n'a pas été suffisant pour permettre à l'analyse canonique de corrélation d'exploiter les 12 variables physiographiques présentant un potentiel. Tout comme pour la phase de développement, cet objectif a été atteint à 75%.

Conceptualisation et documentation de l'implémentation de la procédure automatisée

L'objectif de conceptualisation et de documentation de l'implémentation de la procédure automatisée a été atteint et même dépassé. Ainsi, toutes les fonctions de base ont été automatisées et ce, aussi bien en mode de calibration qu'en mode d'application.

8.2 Impacts et avantages réels et potentiels

1- *Développer des méthodes, systèmes, produits ou autres services d'OT novateurs qui abordent les domaines prioritaires des ressources en eau du Canada de manière à :*

a) *prendre des mesures à l'égard du changement climatique et de la production d'énergie*

Les changements climatiques ont d'importantes répercussions sur le bilan hydrique d'un bassin versant. Les rivières et les lacs font partie de la composante stockage de l'eau. Ainsi, la méthode développée permettrait de mesurer les variations temporelles des largeurs effectives des rivières et d'en estimer le débit si des courbes de tarage ont été produites pour les rivières en question et ce, même si la station n'est plus en fonction. Dans le cas du suivi du débit des affluents alimentant des réservoirs hydroélectriques, il serait envisageable de développer des courbes de tarage satellitaires à des moments clés durant la saison, témoignant de la dynamique de la rivière, en utilisant des mesures de débits ponctuelles. Une fois la courbe de tarage satellitaire développée, le débit pourrait être estimé par l'entremise des images Radarsat. Le concept de station hydrométrique virtuelle pourrait être mis en place. Ce concept novateur aurait recours à une phase de calibration pour développer la courbe de tarage, puis une fois bien calibrée, l'imagerie Radar alimenterait le modèle pour en estimer le débit sans avoir recours à une mesure hydrométrique. Dans le cas d'un cours d'eau problématique, une telle station virtuelle pourrait être implémentée à différents endroits en amont pour avoir un meilleur contrôle des apports en eau provenant des différents affluents de cette rivière. Le travail effectué dans ce projet donne également de très bonnes pistes pour atteindre l'objectif initial, soit la régionalisation du débit. Une fois cet objectif atteint, la prise de mesure de débits pour une station virtuelle ne serait plus nécessaire ce qui faciliterait et augmenterait la rapidité de sa mise en place.

b) *développer des compétences scientifiques de calibre international.*

La démarche expérimentale réalisée au cours de ce projet est novatrice et a impliquée des connaissances avancées en ce qui a trait à la physique de l'onde et son interaction avec les différents milieux, la programmation ainsi que des méthodes d'analyses statistiques avancées. Le projet a impliqué principalement un associé de recherche, trois professionnels de recherche et une étudiante au doctorat. L'expertise développée au sein du laboratoire est de calibre international et pourra servir à accélérer d'autres avancées technologiques en OT.

c) *protéger les écosystèmes aquatiques et les droits autochtones sur l'eau et prévenir les conflits entre les diverses autorités et lutter contre les exportations massives de l'eau*

Au cœur de ce projet, le calcul de la largeur effective a eu recours entre autre à une fonction permettant de calculer la surface d'eau comprise dans une zone de calcul, basée sur une fonction de logique floue. Ce module pourrait être utilisé pour faire le suivi spatio-temporel des surfaces d'eau incluant à la fois les lacs et les rivières. Le changement de surface d'eau serait un indicateur d'inondation ou d'assèchement d'un milieu aquatique. Des images

pourraient être acquises spécifiquement pour faire ce suivi dans l'optique de protéger les écosystèmes aquatiques et les droits autochtones sur l'eau et prévenir les conflits entre les diverses autorités et lutter contre les exportations massives de l'eau. De plus, pour la protection des écosystèmes aquatiques, une connaissance du débit est importante car il s'agit d'un paramètre clé au niveau des rivières. Que ce soit par une station virtuelle, ou ultimement par la régionalisation du débit, les outils développés permettent l'acquisition de cette connaissance.

2- inciter les différents intervenants du secteur de l'eau à utiliser sur une base permanente des données d'OT pour leurs applications;

Le concept de station hydrométrique virtuelle, alimentée par l'imagerie Radarsat, peut inciter les différents intervenants du secteur de l'eau à utiliser sur une base permanente les données d'observation de la terre. Les principaux intervenants qui pourraient être intéressés sont les Services Hydriques et climatiques d'Environnement Canada qui gèrent entre autre les données hydrométriques canadiennes via la base de données HYDAT, Hydro-Québec pour connaître le potentiel de certaines rivières avant l'implantation d'un barrage, le Centre d'Expertise Hydrique du Québec (gère les stations hydrométriques québécoises) ainsi que Ressources Naturelles Canada et Sécurité Publique Canada qui tous les deux ont des mandats visant à augmenter la résilience des organisations gouvernementales et municipales face aux inondations. À ce niveau, à partir d'une série d'images, l'algorithme de détection de l'eau pourrait servir à déterminer la probabilité d'un endroit à être immergé ou inondé.

3- développer, dans le secteur de l'eau, des applications dont l'utilisation opérationnelle à long terme pourrait permettre une surveillance soutenue, compte tenu des besoins particuliers des utilisateurs et des changements dans les ressources de l'infrastructure de l'utilisateur

Le concept de la station virtuelle expliquée dans les recommandations rejoint très bien cet objectif. En effet, dans le projet, les courbes de tarage calibrées sur les rivières jaugées ont confirmées le potentiel de l'utilisation de l'imagerie Radarsat pour estimer le débit d'une rivière. Dans le cas d'une fermeture temporaire ou permanente de l'une des stations hydrométriques utilisées, les courbes obtenues permettraient de continuer à mesurer le débit au moyen de nouvelles acquisitions Radarsat. Il serait également possible de construire des courbes de tarage satellitaire sur des rivières actuellement non jaugées en procédant à des mesures de débits ponctuelles simultanément à l'acquisition des images. Une fois calibrée, la station virtuelle serait opérationnelle et permettrait la surveillance à long terme du débit de la rivière.

En absence de débit pour le calcul de nouvelles courbes de tarage, le développement fait sur la détection de la surface d'eau peut néanmoins permettre le suivi temporelle à long terme des étendues d'eau.

4- améliorer l'accès à des produits et à des services du secteur de l'eau issus de données d'OT, ainsi que la livraison de ces produits et de ces services;

L'approche qui a été préconisée dans ce projet consistait à automatiser le plus possible les calculs en développant différents modules correspondants aux étapes de la chaîne de traitements. Une interface graphique a été créée pour qu'un utilisateur d'un module n'ait pas à adapter le code Python à ses données. En ce sens, la solution proposée est beaucoup plus accessible à un usager intéressé que si la séquence était fermée et sans interface.

5- mettre au point des produits et des services normalisés liés au secteur de l'eau et créés à partir de données d'OT.

Tous les intrants (données) utilisés pour ce projet ont été produits selon les standards canadiens. Les données produites en sortie lors de ce projet ont été exprimées dans les mêmes unités que les intrants.

8.3 Succès global

Principales découvertes

L'existence du lien entre la rétrodiffusion radar et le débit ($We-Q$) mesuré en rivière a été confirmé. La variation du coefficient de rétrodiffusion du signal Radar a permis de suivre la dynamique de la variation de We en fonction du débit et la construction des courbes de tarage satellitaires testées dans plusieurs configurations physiographiques. Avec cette méthode il serait possible de continuer le suivi du débit d'un cours d'eau calibré avec une courbe de tarage satellitaire, même si la station est en arrêt temporaire ou permanent.

Le regroupement établi par l'analyse canonique de corrélation a démontré qu'il existe des liens mathématiques entre les variables canoniques et les paramètres des courbes de tarage satellitaire (exponentielles et polynomiales).

L'information contenue dans le produit EOSD distribué par la plate-forme Saforah (classes forestières de couvertures de la terre pour le Canada) a servi à séparer les fonctions croissantes et décroissantes des courbes de tarage satellitaire.

Remarques

Ce qui a véritablement limité le développement du modèle d'estimation du débit pour les rivières non-jaugées concerne les problèmes dus à la taille de l'échantillon pour l'analyse ainsi que le fractionnement de la base de données de calibration. Dans le cas des 25 stations avec fonctions croissantes, il y a 16 stations où les courbes de tarage ont été développées avec des images en orbite descendante et 9 stations avec images en orbite ascendante. Dans le cas des 11 stations avec fonction décroissante, il y a 5 stations où les courbes de tarage ont été développées avec des images en orbite descendante et 6 stations avec images en orbite ascendante. Idéalement, les modèles de prédiction du débit dans le site non jaugés ne devraient pas combiner les courbes de tarage générées avec des images en orbite ascendantes et descendantes.

Afin de construire les meilleurs courbes de tarage, les paramètres d'acquisition d'images radar devraient être les plus homogènes possibles (orbite, angle d'incidence moyen). En effet, même si la procédure d'intercalibration des images réduit l'effet de l'angle d'incidence, il reste qu'il s'agit d'une correction imparfaite.

Le rapport entre la quantité et l'homogénéité de l'ensemble des images disponibles utilisées dans le développement des courbes de tarage satellitaire n'est pas idéal. Les courbes de tarage n'ont pas la même qualité dans le jeu de calibration, ce qui a un impact sur la propagation d'erreurs des modèles établis.

L'incertitude du calcul de la pente de la courbe de tarage (α) a un impact important sur la prédiction du débit. La méthode développée est très sensible aux changements de We .

Enfin, pour tous les développements futurs, il serait important de considérer les éléments abordés dans la section intitulée « Recommandations ».

9 Références

- Breiman L, Friedman J, Stone CJ & Olshen RA (1984) *Classification and regression trees*. CRC press,
- CEHQ (2003) *Données historiques des niveaux et des débits*. Centre d'expertise hydrique, Ministère du Développement Durable, de l'Environnement et de la Lutte contre les Changements Climatiques, Québec, <http://www.cehq.gouv.qc.ca/hydrometrie/> (Consulté le 29 mai 2015)
- EC (2013) *Archives nationales des données hydrologiques, Base de données HYDAT*. Environnement Canada), <https://ec.gc.ca/rhc-wsc/default.asp?lang=Fr&n=9018B5EC-1> (Consulté le 29 mai 2015)
- RNCAN (2008) *SAFORAH - EOSD Data and Products*. Natural Ressources Canada, Canadian Forest Service, <http://www.saforah.org/> (Consulté le 29 mai 2015)
- RNCAN (2013) *Regional, national and international climate modeling*, <http://cfs.nrcan.gc.ca/projects/3/1> (Consulté le 29 mai 2015)
- RNCAN (2015a) *Données numériques d'élévation du Canada (DNEC)*. Secteur des sciences de la Terre, Centre Canadien de la Cartographie et d'Observation de la Terre.
- RNCAN (2015b) *Réseau hydro national*. Secteur des sciences de la Terre, Centre Canadien de la Cartographie et d'Observation de la Terre.
- RNCAN (2015c) *Réseau routier national*. Secteur des sciences de la Terre, Centre Canadien de la Cartographie et d'Observation de la Terre.
- Roberge S, Chokmani K & Danielle DS (2010) Mise à jour de l'algorithme de cartographie de la présence de neige au sol à partir des données optiques du capteur NOAA-AVHRR, Rapport final, Rapport de recherche-R1171. Institut national de la recherche scientifique, Centre Eau Terre Environnement, Québec.
- USGS (2005) *2005 North American Land Cover at 250 m spatial resolution*. Natural Resources Canada/Canadian Center for Remote Sensing (NRCAN/CCRS), United States Geological Survey (USGS); Instituto Nacional de Estadística y Geografía (INEGI), Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO) and Comisión Nacional Forestal (CONAFOR)., <http://landcover.usgs.gov/nalcms.php> (Consulté le 29 mai 2015)

Annexes

Annexe I: Description détaillée des calculs de variables physiographiques

La définition et la méthode utilisée dans le calcul de chaque variable sont présentées ci-dessous. Le nom est accompagné entre parenthèse de sa dénomination abrégée dans la base de données et dans ce rapport :

- **Superficie (Superficie)**

Correspond à la superficie, exprimée en km², couverte par le bassin versant ayant comme exutoire la station de jaugeage.

- **Périmètre (Perim)**

Correspond à la longueur de la ligne de partage des eaux définissant la limite du bassin versant, exprimée en km.

- **Longueur de l'axe longitudinal (L_Axis)**

Distance maximale entre deux points appartenant à la limite du bassin versant, exprimée en km.

- **Facteur de forme (Facteur_forme)**

Correspond à la superficie du bassin versant divisée par la longueur de l'axe longitudinal élevé au carré (équation (14)).

$$Facteur_forme = \frac{S}{L_Axis^2} \quad (14)$$

- **Indice de compacité de Gravelius (Coef_compacite)**

Correspond au rapport du périmètre du bassin à celui d'un cercle de même surface. L'indice a une valeur proche de 1 pour un bassin versant de forme quasiment circulaire et supérieur à 1 lorsque le bassin est de forme allongée (équation (15)).

$$coef_compacite = \frac{P}{2\sqrt{\pi A}} \approx 0.28 \frac{P}{\sqrt{A}} \quad (15)$$

- **Pente moyenne (pente_moy)**

Correspond à la pente moyenne de la région couverte par le bassin versant. La pente est calculée à chaque pixel du modèle numérique de terrain puis la moyenne de tous les pixels à l'intérieur des limites du bassin est calculée. Pour un pixel donné, la pente correspond à la différence d'élévation maximale entre le pixel et ses 8 voisins, divisée par la distance centre à centre entre les deux pixels en question. La pente est exprimée en pourcentage (%). La pente moyenne donne une bonne indication sur le temps de parcours du ruissellement direct - donc sur le temps de concentration - et influence directement le débit de pointe lors d'une averse.

- **Altitude moyenne (Altitude_moy)**

Correspond à l'élévation moyenne, exprimée en mètre, de la région couverte par le bassin versant par rapport au niveau moyen des mers.

- **Latitude moyenne (Latitude_moy)**

Correspond à la latitude moyenne, exprimée en degrés décimaux, de la région couverte par le bassin versant par rapport au plan équatorial.

- **Coefficient de massivité (Coeff_massiv)**

Correspond à la moitié de la différence d'élévation entre le point le plus haut et le point le plus bas de la région couverte par le bassin versant, divisée par la superficie du bassin en m². S'exprime en m⁻¹. Une valeur élevée du coefficient de massivité reflète une plus grande augmentation de la hauteur par unité de surface et donc des risques accrus de ruissellement, d'inondations et d'érosion (équation (16)).

$$coeff_massiv = \frac{h_{max} - h_{min}}{2} \times \frac{1}{superficie} \quad (16)$$

- **Coefficient orographique (Coeff_oro)**

Correspond à la moitié de la différence d'élévation entre le point le plus haut et le point le plus bas de la région couverte par le bassin versant, élevée au carré puis divisée par la superficie du bassin en m². S'exprime sans unité (équation (17)).

$$coeff_oro = \left(\frac{h_{max} - h_{min}}{2} \right)^2 \times \frac{1}{superficie} \quad (17)$$

- **Ordre de Strahler à la station (Strahler)**

Dans la méthode Strahler, tous les cours d'eau sans affluents sont classés avec la valeur 1 et sont connus sous le nom du premier ordre. L'ordre d'écoulement augmente lorsque des cours d'eau du même ordre se croisent. L'intersection de deux cours d'eau de premier ordre crée donc un cours d'eau de deuxième ordre. L'intersection de deux cours d'eau de deuxième ordre crée un cours d'eau de troisième ordre, et ainsi de suite. En revanche, l'intersection de deux cours d'eau d'ordres différents n'incrémente pas l'ordre.

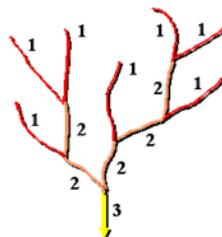


Figure 34 : Exemple de la classification de Strahler (source : <http://help.arcgis.com>)

- **Ordre de Shreve à la station (Shreve)**

La méthode de classification de Shreve tient compte de l'ensemble des cours d'eau du réseau. Comme avec la méthode de Strahler tous les cours d'eau externes sont associés à un ordre 1. Toutefois, les cours d'eau internes dans le cadre de la méthode de Shreve s'additionnent. Par exemple, l'intersection de deux liaisons de premier ordre crée une liaison de deuxième ordre, l'intersection d'une liaison de premier ordre et d'une liaison de deuxième ordre crée une liaison de troisième ordre et l'intersection d'une liaison de deuxième ordre et d'une liaison de troisième ordre crée une liaison de quatrième ordre. L'ordre d'un cours d'eau, dans la méthode de Shreve, est directement proportionnel au nombre de cours d'eau en amont.

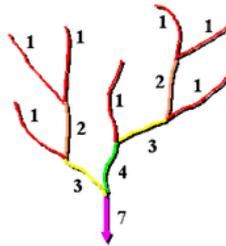


Figure 35 : Exemple de la classification de Shreve (source : <http://help.arcgis.com>)

- **Longueur du cours d'eau principal (L_princ)**

Correspond à la distance curviligne depuis l'exutoire jusqu'au cours d'eau le plus en amont, en suivant toujours le segment d'ordre de Shreve le plus élevé lorsqu'il y a un embranchement. Si les deux segments à l'embranchement sont de même ordre, on suit celui qui draine la plus grande surface. Exprimée en km.

- **Sinuosité du cours d'eau principal (S_princ)**

Correspond à la longueur réelle du cours d'eau divisée par la distance en ligne droite entre le point amont et le point aval. Lorsque le résultat est compris entre 1,25 et 1,5 on dit que le cours d'eau est sinueux. Au-delà de 1,5 il est méandriforme.

- **Sinuosité du cours d'eau principal sur les 10km en amont de la station (S_princ_10)**

Correspond à la longueur réelle du cours d'eau (10km) divisée par la distance en ligne droite entre le point amont et le point aval. Lorsque le résultat est compris entre 1,25 et 1,5 on dit que le cours d'eau est sinueux. Au-delà de 1,5 il est méandriforme.

- **Longueur du réseau hydrographique (L_Hydro)**

Correspond à la somme de la longueur de tous les cours d'eau en amont de la station hydrométrique, exprimée en km.

- **Pente moyenne du cours d'eau principal (P_princ)**

Correspond à la différence d'altitude entre les points extrêmes du cours d'eau divisée par sa longueur totale (équation (18)).

$$Pm_{cp} = \frac{\Delta H_{max}}{L} \quad (18)$$

- **Pente moyenne du cours d'eau principal sur 10km (P_princ_10)**

Correspond à la différence d'altitude entre la station et un point situé à 10 km en amont de celle-ci divisée par la distance les séparant (10km).

- **Densité de drainage (Dd)**

Correspond à la longueur totale du réseau hydrographique par unité de surface du bassin versant. S'exprime en km /km² (équation (19)).

$$Dd = \frac{\sum L_i}{A} \quad (19)$$

- **Densité hydrographique (Dh)**

Correspond au nombre de cours d'eau par unité de surface (équation (20)).

$$Dh = \frac{\sum N_i}{A} \quad (20)$$

- **Précipitations moyennes annuelles (Pmoy_a)**

Correspond à la moyenne des précipitations reçues par année entre 1960 et 2010 sur l'ensemble du bassin versant, exprimé en mm (équation (21)).

$$Pmoy_a = \frac{\sum_{a=1960}^{2010} \sum_{m=1}^{12} P_{am}}{(2010 - 1960) + 1} \quad (21)$$

- **Précipitations maximales moyennes mensuelles (Pmax_m)**

Correspond à la moyenne, pour les années entre 1960 et 2010, des précipitations reçues pendant le mois où les précipitations ont été maximales dans l'année, exprimé en mm (équation (22)).

$$P_{max_m} = \frac{\sum_{a=1960}^{2010} \max(P_{a1}, P_{a2}, \dots, P_{a12})}{(2010 - 1960) + 1} \quad (22)$$

- **Précipitations solides moyennes annuelles (Pmoy_sol_a)**

Correspond à la moyenne annuelle des précipitations lorsque la température moyenne est inférieure à 0.

- **Précipitations liquides moyennes annuelles (Pmoy_liq_a)**

Correspond à la moyenne annuelle des précipitations lorsque la température moyenne est supérieure à 0.

- **Indice de Fournier (IF)**

Correspond à la précipitation maximale moyenne mensuelle élevée au carré et divisée par la précipitation moyenne annuelle (équation (23)).

$$IF = \frac{P_{max_m}^2}{P_{moy_a}} \quad (23)$$

- **Indice de Fournier, version 2 (IF_unite)**

Version modifiée de l'indice de Fournier où les précipitations sont ramenées par unité de surface (pixel) (équation (24)).

$$IF_{unite} = \frac{P_{max_m}^2}{P_{moy_a}} * \frac{100}{S} \quad (24)$$

- **Quotient pluviothermique d'Emberger (Qe)**

L'indice d'Emberger prend en compte les précipitations annuelles P, la moyenne des maxima de température du mois le plus chaud (M en K) et la moyenne des minima de température du mois le plus froid (m en K) (Emberger, 1955) (équation (25))

$$Qe = \frac{2000P}{M^2 - m^2} \quad (25)$$

- **Quotient pluviothermique d'Emberger, version 2 (Qe_unite)**

Version modifiée du quotient pluviothermique d'Emberger où les précipitations totales sont ramenées par unité de surface (pixel) (équation (26)).

$$Qe_{unite} = \frac{2000P}{M^2 - m^2} * \frac{100}{S} \quad (26)$$

- **Occupation du sol (LCLU_pct)**

Le pourcentage de chaque classe d'occupation du sol (Tableau 26) pour le territoire couvert par le bassin versant de chaque station a été calculé.

Tableau 26 : Liste de classes d'occupation du sol.

Code	Classe d'occupation du sol
1	Temperate or sub-polar needle leaf forest
2	Sub-polar taiga needle leaf forest
3	Tropical or sub-tropical broadleaf evergreen forest
4	Tropical or sub-tropical broadleaf deciduous forest
5	Temperate or sub-polar broadleaf deciduous forest
6	Mixed forest
7	Tropical or sub-tropical shrubland
8	Temperate or sub-polar shrubland
9	Tropical or sub-tropical grassland
10	Temperate or sub-polar grassland
11	Sub-polar or polar shrubland-lichen-moss
12	Sub-polar or polar grassland-lichen-moss
13	Sub-polar or polar barren-lichen-moss
14	Wetland
15	Cropland
16	Barren lands
17	Urban
18	Water

- **Occupation du sol (EOSD)**

Le pourcentage de chaque classe d'occupation du sol (Tableau 27) selon les classes d'occupation du sol des données d'Observation de la Terre pour le Développement Durable des Forêts (EOSD) a été calculé à l'intérieur du masque de calcul pour la probabilité d'appartenance de l'eau (défini à la section II).

- **Pente transversale à la rivière (s1 à s10)**

La pente transversale est déterminée en traçant, dans chaque section, 3 profils orthogonaux à la ligne centrale de la rivière et en calculant la différence d'élévation entre un point situé à la périphérie de la section et un point situé à l'intersection du profil et de la ligne centrale. La mesure se prenant de part et d'autre de la ligne centrale, les 3 profils génèrent 6 mesures de pente lesquelles sont ensuite utilisées pour calculer une pente moyenne. Les valeurs sont ensuite cumulées par section, jusqu'à la section 10.

Tableau 27 : Définition des classes d'occupation du sol des données d'Observation de la Terre pour le Développement Durable des Forêts (EOSD)

Code	Classe d'occupation du sol
0	Pas de données
12	Nuages
11	Ombre
20	Eau
31	Neige/Glace
32	Roche/Blocaille
33	Terrain découvert
40	Bryophytes et lichens
51	Grands arbustes
52	Petits arbustes
81	Zone humide boisée
82	Zone humide arbustive
83	Zone humide herbacée
100	Plantes herbacées
211	Coniférien dense
212	Coniférien ouvert
213	Coniférien clairsemé
221	Feuillu dense
222	Feuillu ouvert
223	Feuillu clairsemé
231	Mixte dense
232	Mixte ouvert
233	Mixte clairsemé

Annexe II : Contenu de la base de données

L'organisation des données qui seront remises sur le support physique (disque dur) avec ce rapport est dans la Figure 36. La partie de la structure en verte correspond à des dossiers tandis que la partie en violet correspond à une base de données spatiales (file GeoDataBase).

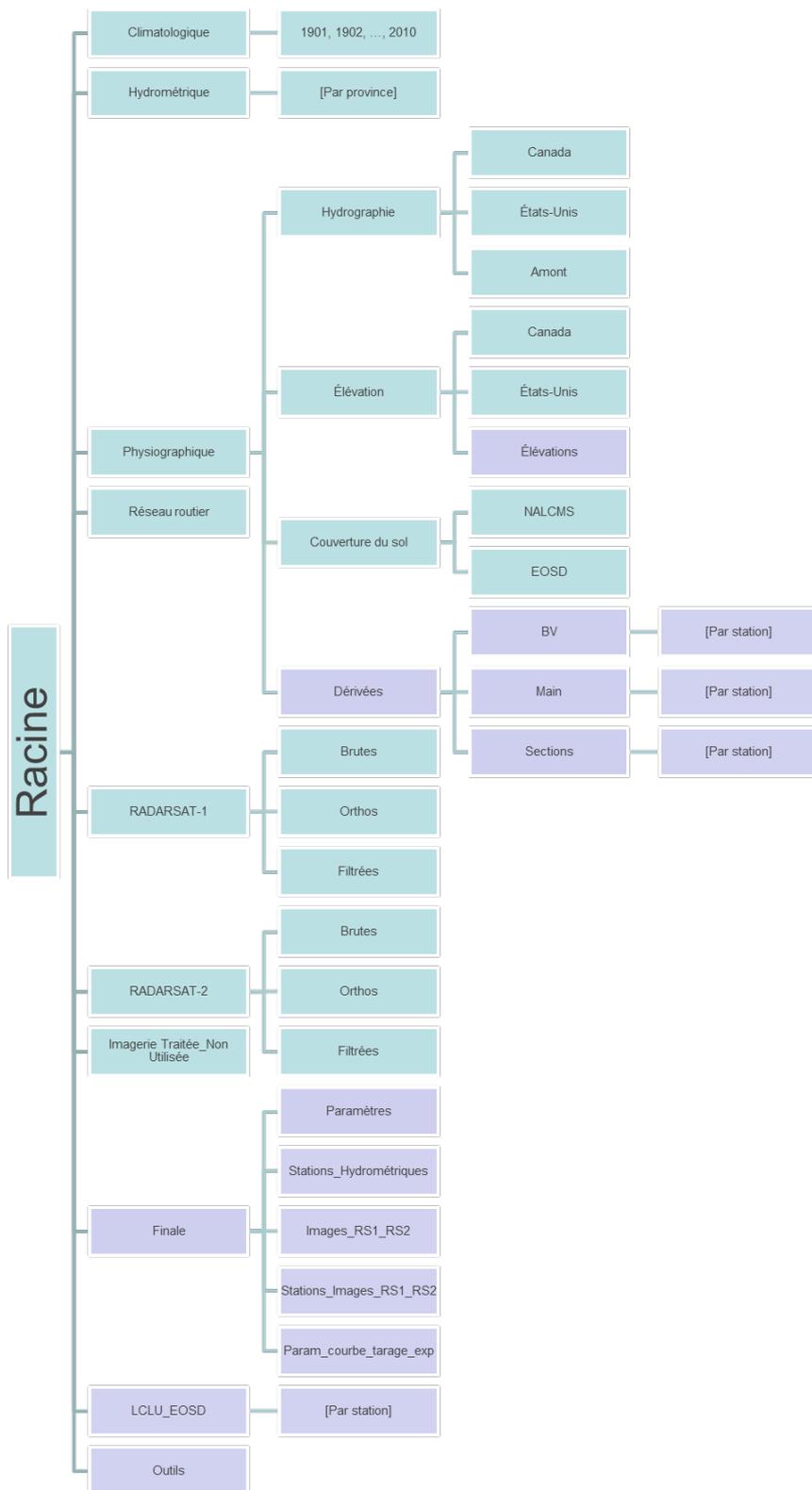


Figure 36 : Diagramme de la structure de données sur le support physique.

BD finale

Paramètres climatologiques, physiographiques et cartographiques (Parametres)

Afin de caractériser l'environnement de chaque station hydrométrique retenue dans le projet, plusieurs paramètres et indices climatologiques, physiographiques et cartographiques ont été calculés (Tableau 29 **Erreur ! Source du renvoi introuvable.**). Ces paramètres ont été obtenus à partir du traitement des différents jeux de données décrit précédemment. Les couches de référence utilisées sont les bassins versants et les chenaux principaux. Une description plus étendue de chaque paramètre se retrouve à l'Annexe I : Description détaillée des calculs de variables physiographiques.

Stations hydrométriques (Stations_hydrometriques)

Ce jeu de données regroupe la position de chaque station hydrométrique utilisée de même que quelques informations pertinentes permettant de les décrire. Est également présent, le nombre d'images (RS-1&2) associées à la station ainsi que la largeur nominale du cours d'eau mesurée vis-à-vis la station.

Images RS-1 RS-2 (Images_RS1_RS2)

Pour chaque image RS-1 et RS-2 obtenue dans le cadre du projet, un polygone correspondant à l'empreinte au sol de l'image a été réalisé. À ce polygone a été ajouté comme attribut le nom du satellite ayant acquis l'image, le code de l'image, le mode et la date d'acquisition. Ces polygones ont été regroupés dans le jeu de données Images_RS1_RS2.

Stations par images RS-1&2 (Stations_Images_RS1_RS2)

Une table de relation a été créée afin de faire le lien entre les images RS-1&2 commandées et les stations hydrométriques. Dans cette table, les images sont liées à chaque station qu'elle contienne. Il s'agit d'une relation « plusieurs à plusieurs », i.e. une image peut contenir plusieurs stations et une station peut se retrouver sur plusieurs images. La relation entre une station et une image est accompagnée des valeurs de débits instantanés et de débit quotidien pour la station en question au moment de l'acquisition de l'image, de même que de la liste des sections couvertes par l'image.

Paramètres des courbes de tarage satellitaire (Param_courbe_tarage_exp)

La base de données des courbes de tarage satellitaire a été établie en sélectionnant le meilleur ajustement dans le rapport largeur effective (We) et débit instantané (Q) en fonction de la distance d'intégration, dans chaque site jaugé. Les paramètres dans cette base de données sont entre autres le coefficient et l'exposant de la largeur effective (α et β , respectivement) pour les courbes exponentielles ainsi que la distance d'intégration (*Sections*), pour chaque site jaugé.

Relations

À l'intérieur de la base de données spatiales *Finale*, plusieurs tables sont reliées les unes aux autres (Figure 37 **Erreur ! Source du renvoi introuvable.**). La table centrale est celle des stations hydrométriques. Tel que mentionné précédemment, elle est reliée à la table

Images_RS1_RS2 dans une relation « plusieurs à plusieurs », i.e. qu'une station hydrométrique peut être incluse dans une ou plusieurs images RS alors qu'une image RS peut inclure une ou plusieurs stations. Cette relation nécessite la création d'une nouvelle table mettant en correspondance l'identifiant des stations et celui des images RS. La relation entre une image et une station étant valide pour un moment précis dans le temps, soit la date et l'heure d'acquisition de l'image, les débits instantanés et quotidiens mesurés à la station à l'heure et au jour en question ont été ajoutés à cette table. De plus, chaque image couvrant un nombre différent de sections de rivière, un champ liste les sections couvertes par l'image pour cette station.

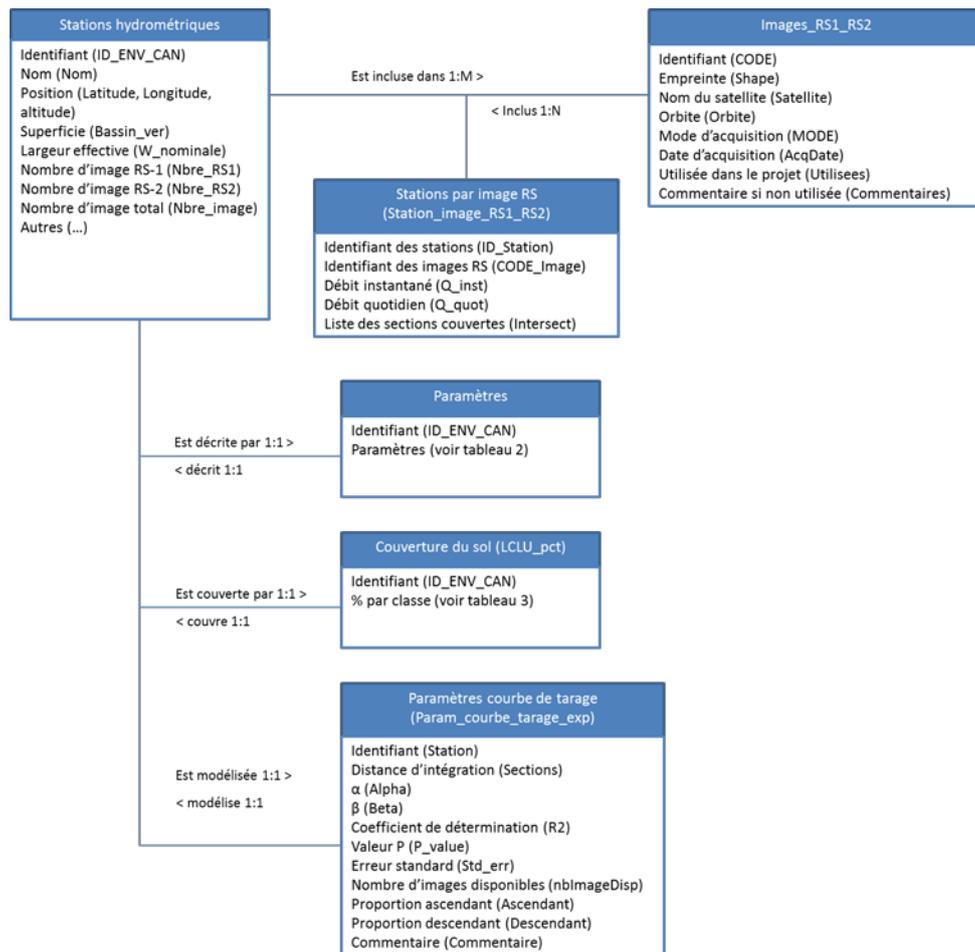


Figure 37 : Structure relationnelle de la base de données spatiales *Finale*.

La table Stations hydrométriques est également reliée aux deux tables de paramètres descriptifs (Paramètres et Couverture du sol) ainsi qu'à la table des paramètres de courbes de tarage dans une relation « un à un », i.e. qu'une station est décrite par un seul ensemble de paramètres alors qu'un ensemble de paramètre ne décrit qu'une seule station.

La création de ces relations permet de facilement interroger les différentes tables une fois dans ArcGIS en cliquant sur une station ou une empreinte d'images RS (Figure 38).

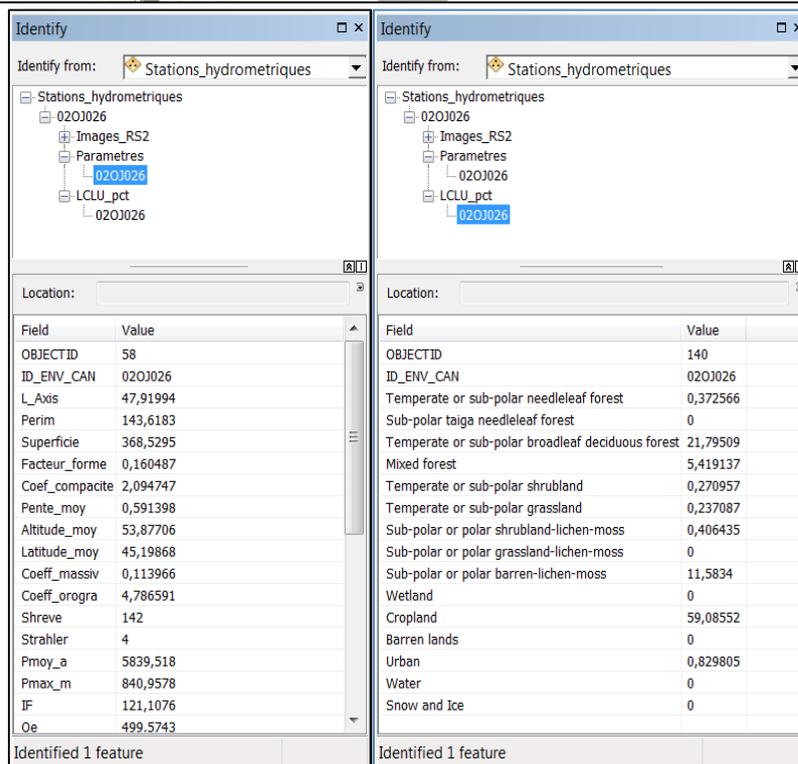
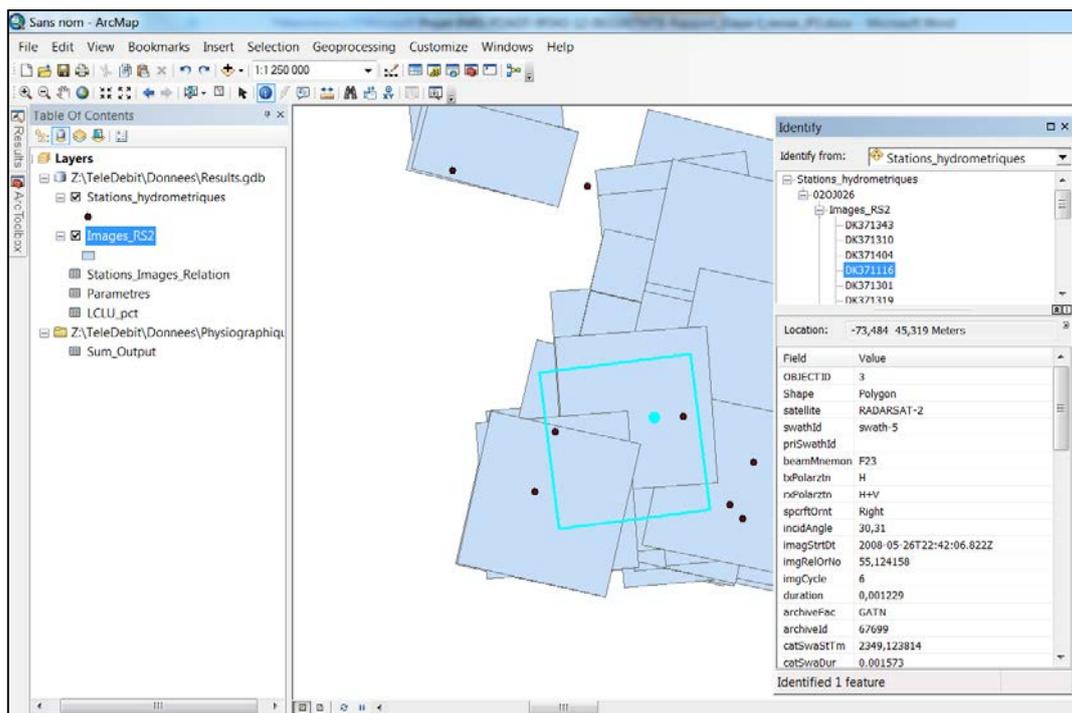


Figure 38 : Interface d'interrogation dans ArcMap.

BD EOSD

La base de données EOSD contient, pour chaque station, la couverture du sol pour chacune des 10 premières sections déterminée en utilisant les données EOSD. Les valeurs présentées

sont exprimées en pourcentage de représentativité pour chaque classe. La signification des codes des champs se trouve dans le Tableau 27. Les pourcentages de chaque type d'utilisation du sol par sections de rivière ont été conservés dans une base de données séparée de celle des autres paramètres car ils renferment beaucoup d'information et forment un ensemble de paramètres à eux seuls.

BD Dérivées

Bassin versant (BV)

Les bassins versants associés à chaque station hydrométrique se retrouvent dans des classes d'entités différentes nommées selon le code de la station mais sont regroupées dans un même jeu d'entité nommé *BV*. Ces bassins versants ont servi comme zone de calcul pour certaines variables physiographiques (Annexe I)

Cours d'eau principal (Main)

Le cours d'eau principal en amont de chaque station a été déterminé en suivant les segments d'ordres les plus élevés. Ce cours d'eau principal est représenté par une polyligne partant de la station et s'étendant jusqu'à l'amont du bassin versant. Le cours d'eau principal de chaque station est enregistré dans une classe d'entité différente nommée en utilisant le code de la station hydrométrique. Tous les cours d'eau principaux se retrouvent dans le jeu d'entité *Main*.

Sections de la rivière (Sections)

En utilisant conjointement le cours d'eau principal sous forme polyligne et la représentation polygonale du cours d'eau en amont de chaque station, il a été possible de séparer le polygone en tronçons d'une longueur constante de 1km. Le nombre de sections est variable d'une station hydrométrique à l'autre et dépend de la taille et de la longueur de la rivière. Les sections relatives à deux stations différentes se retrouvent dans deux classes d'entités différentes mais sont regroupées dans le même jeu d'entités *Sections*. Plusieurs classes d'entités sont présentes pour une même station car différentes valeurs de zone tampon ont été appliquées sur le polygone du cours d'eau. Ces classes d'entités présentent des suffixes différents représentant le pourcentage de la largeur nominale utilisé pour la création de la zone tampon.

BD Élévations

La base de données Élévations est composée de deux mosaïques de données (Mosaic Dataset). Les mosaïques de données permettent de stocker, gérer, visualiser et interroger des collections de données matricielles. Elles sont utilisées pour gérer un ensemble de données matricielles stocké comme un catalogue et visualisé comme une mosaïque. Les données numériques d'élévation du Canada étant distribuées par tuiles, de même que celles du National Elevation Dataset, il est possible de les visualiser comme si elles avaient été

fusionnées par l'entremise de la mosaïque de données. Dans le processus de visualisation, certaines fonctions mathématiques et géométriques peuvent être appliquées aux données. Dans notre base de données Élévations, l'une des mosaïques n'a aucune fonction d'appliquée (Brute) alors que la seconde a une fonction de projection la transformant dans une projection conique conforme de Lambert (LCC).

Outils

Une boîte à outils compatibles avec ArcGIS contenant les outils développés dans le cadre du projet est présente dans ce répertoire. Outre les outils permettant de calculer le débit pour un site non-jaugé qui seront décrit à l'annexe suivante, d'autres outils complémentaires accompagnent les données. Ces outils sont parfois des versions autonomes de ceux présents dans la procédure pour un site non-jaugé.

Ces outils ont été développés dans le but de démontrer la faisabilité du calcul du débit pour un site non-jaugé. Ils ne sont pas exempt de bugs et ne possèdent pas la robustesse d'une version commerciale.

Création des sections

Cet outil permet la division d'un polygone représentant une rivière en sections de longueur définie. Les sections sont créées en coupant, à distance constante, le polygone perpendiculairement à la ligne centrale de la rivière. Cette distance doit être spécifiée en mètre.

Calcul de la largeur effective

Cet outil permet le calcul de la largeur effective de sections de rivière pour une ou plusieurs images Radarsat. Il produit un fichier Excel contenant les résultats de chaque section et chaque image. L'outil présume que la longueur des sections est de 1000m.

Création des courbes de tarage

L'outil permettant la création des courbes de tarage peut être lancé à partir de la boîte à outils d'ArcGIS ou fonctionner de façon autonome en exécutant le script associé (InterfaceCourbeTarage.py) situé dans le sous-dossier *Scripts*. L'outil permet la visualisation et l'analyse des données de débits et de largeurs effectives et la création des courbes de tarage. Pour l'utiliser, un fichier Excel doit d'abord être ouvert en utilisant le bouton *Charger*. L'outil s'attend à retrouver une structure particulière à l'intérieur de ce fichier (Tableau 28). Les deux premières lignes peuvent contenir de l'information d'intérêt pour l'utilisateur ou être vides. La troisième ligne n'est pas considérée par l'outil mais correspond aux titres des colonnes sous-jacentes. La quatrième ligne contient dans l'ordre la Date de l'image, son identifiant, le mode d'acquisition, l'orbite, le débit instantané de la rivière au moment de l'acquisition de l'image, le débit quotidien de la journée d'acquisition, puis les largeurs effectives par section, de la section 1 à la dernière section. Les lignes suivantes doivent avoir la même structure mais correspondre à des images différentes.

Une fois le fichier chargé, l'utilisateur peut en visualiser le contenu et sélectionner les images qu'il juge d'intérêt à la création de la courbe de tarage, de même que le type de débit à utiliser. L'utilisateur peut également déterminer les sections qu'il veut traiter en modifiant la boîte de texte *Section*. Les choix possibles pour cette option sont : *All*, qui correspond à toutes les sections, le # de la section pour l'analyse d'une seule section ou les # de deux sections séparés par un « - » pour l'analyse d'une série de sections. L'utilisateur a le choix de faire afficher les graphiques résultant de la régression et montrant la courbe de tarage ou de les sauvegarder en cochant la case *Sauvegarder les résultats*. Dans la boîte de texte *Répertoire d'enregistrement*, l'utilisation du texte {Date_Heure} comme nom de dossier indique à l'outil de créer un nouveau dossier nommé selon la date et l'heure du calcul à chaque fois qu'un calcul est lancé et que l'enregistrement des données est désiré.

Les résultats des calculs sont conservés dans le fichier spécifié dans la boîte de texte *Fichier log*.

Tableau 28 : Structure du fichier Excel nécessaire à l'utilisation de l'outil de création des courbes de tarage

	A	B	C	D	E	F	G	H	...
1									
2									
3	Date	Code	Mode	Orbite	Q inst	KM \ Q quot	1	2	...
4	2008-09-03	DK371120	F22	D	401,808	401	338,83	330,75	...
5	2009-05-01	DK371135	F3N	D	754,909	682	389,03	375,33	...
6	2009-09-22	DK371163	F6F	A	231,104	238	326,47	322,80	...
...

Tableau 29 : Liste de variables climatologiques, physiographiques et cartographiques produite par station hydrométrique.

Source	#	Variable
MNT	1	Superficie du BV
	2	Périmètre
	3	Longueur de l'axe longitudinal
	4	Facteur de forme
	5	Indice de compacité de Gravelius
	6	Pente moyenne du BV
	7	Altitude moyenne
	8	Latitude moyenne
	9	Coefficient de massivité
	10	Coefficient orographique
	11	Orientation principale du BV
RHN	12	Ordre de Strahler
	13	Ordre de Shreve
	14	Longueur du cours d'eau principal
	15	Sinuosité du cours d'eau principal
	16	Sinuosité du cours d'eau principal 10km en amont de la station
	17	Longueur du réseau hydrographique
	18	Pente moyenne du cours d'eau principal
	19	Pente moyenne du cours d'eau principal sur 10km
	20	Densité de drainage
	21	Pente transversale au cours d'eau principal
Climat	22	Précipitations moyennes annuelles
	23	Précipitations maximales moyennes mensuelles
	24	Indice de Fournier (versions 1 & 2)
	25	Quotient pluviothermique d'Emberger (versions 1 & 2)
	26	Précipitation liquide annuelle moyenne
	27	Précipitation solide annuelle moyenne
LULC	28	Pourcentage de couverture par type d'occupation dans le bassin versant
	29	Pourcentage de couverture par type d'occupation par sections

Annexe III : Analyse de la corrélation

Tableau 30 : Matrice de corrélation croisée pour les 33 variables physiographiques présentant des corrélations significatives avec au minimum un paramètre de courbe de tarage

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	BB	CC	DD	EE	FF	GG
A	1.00	0.43	0.40	0.47	0.47	0.35	0.18	0.56	0.53	0.39	0.39	0.45	0.46	0.47	0.44	0.44	0.40	0.47	0.08	0.39	0.39	0.45	0.34	0.31	0.24	0.27	0.27	0.28	0.27	0.26	0.26	0.82	0.49
B	0.43	1.00	0.68	0.66	0.69	0.38	0.02	0.59	0.40	0.68	0.73	0.68	0.68	0.69	0.69	0.69	0.66	0.30	0.34	0.20	0.13	0.23	0.24	0.22	0.16	0.17	0.19	0.15	0.13	0.11	0.11	0.29	0.54
C	0.40	0.68	1.00	0.98	0.98	0.73	0.01	0.84	0.59	0.94	0.90	0.97	0.97	0.96	0.96	0.97	0.97	0.53	0.37	0.35	0.34	0.22	0.19	0.18	0.13	0.16	0.15	0.16	0.15	0.18	0.19	0.29	0.43
D	0.47	0.66	0.98	1.00	0.98	0.80	0.00	0.86	0.62	0.94	0.90	0.96	0.96	0.96	0.95	0.98	0.97	0.60	0.30	0.39	0.39	0.29	0.25	0.23	0.19	0.21	0.21	0.21	0.24	0.24	0.34	0.44	
E	0.47	0.69	0.98	0.98	1.00	0.66	0.03	0.88	0.64	0.96	0.92	0.99	0.99	0.99	0.98	0.99	0.97	0.61	0.28	0.43	0.38	0.33	0.30	0.29	0.25	0.28	0.28	0.28	0.30	0.31	0.37	0.53	
F	0.35	0.38	0.73	0.80	0.66	1.00	0.08	0.55	0.39	0.65	0.59	0.61	0.61	0.61	0.59	0.67	0.70	0.41	0.28	0.16	0.30	0.10	0.02	0.03	0.06	0.05	0.06	0.04	0.04	0.03	0.03	0.17	0.08
G	0.18	0.02	0.01	0.00	0.03	0.08	1.00	0.10	0.19	0.02	0.02	0.02	0.03	0.03	0.02	0.03	0.08	0.13	0.38	0.02	0.20	0.16	0.20	0.12	0.14	0.19	0.19	0.16	0.18	0.14	0.15	0.04	0.36
H	0.56	0.59	0.84	0.86	0.88	0.55	0.10	1.00	0.92	0.78	0.77	0.88	0.89	0.90	0.88	0.84	0.84	0.83	0.20	0.44	0.59	0.31	0.22	0.23	0.19	0.23	0.23	0.24	0.23	0.27	0.28	0.48	0.44
I	0.53	0.40	0.59	0.62	0.64	0.39	0.19	0.92	1.00	0.50	0.51	0.64	0.66	0.67	0.65	0.58	0.60	0.87	0.10	0.38	0.67	0.25	0.11	0.13	0.09	0.14	0.14	0.16	0.14	0.18	0.20	0.49	0.28
J	0.39	0.68	0.94	0.94	0.96	0.65	0.02	0.78	0.50	1.00	0.92	0.94	0.94	0.93	0.92	0.99	0.94	0.51	0.25	0.35	0.36	0.24	0.25	0.26	0.23	0.26	0.26	0.26	0.25	0.26	0.26	0.23	0.48
K	0.39	0.73	0.90	0.90	0.92	0.59	0.02	0.77	0.51	0.92	1.00	0.92	0.91	0.91	0.91	0.93	0.90	0.50	0.22	0.39	0.31	0.29	0.30	0.31	0.29	0.32	0.33	0.33	0.34	0.36	0.37	0.33	0.53
L	0.45	0.68	0.97	0.96	0.99	0.61	0.02	0.88	0.64	0.94	0.92	1.00	1.00	1.00	1.00	0.98	0.96	0.59	0.29	0.45	0.38	0.33	0.31	0.30	0.26	0.29	0.29	0.30	0.29	0.32	0.33	0.37	0.52
M	0.46	0.68	0.97	0.96	0.99	0.61	0.03	0.89	0.66	0.94	0.91	1.00	1.00	1.00	0.97	0.95	0.61	0.28	0.45	0.38	0.33	0.31	0.30	0.26	0.28	0.29	0.29	0.29	0.32	0.33	0.38	0.52	
N	0.47	0.69	0.96	0.96	0.99	0.61	0.03	0.90	0.67	0.93	0.91	1.00	1.00	1.00	0.99	0.97	0.95	0.61	0.28	0.44	0.39	0.33	0.31	0.30	0.25	0.28	0.28	0.29	0.28	0.31	0.32	0.38	0.52
O	0.44	0.69	0.96	0.95	0.98	0.59	0.02	0.88	0.65	0.92	0.91	1.00	1.00	0.99	1.00	0.96	0.94	0.59	0.30	0.45	0.39	0.34	0.32	0.31	0.27	0.30	0.30	0.31	0.30	0.34	0.35	0.37	0.52
P	0.44	0.69	0.97	0.98	0.99	0.67	0.03	0.84	0.58	0.99	0.93	0.98	0.97	0.97	0.96	1.00	0.97	0.57	0.27	0.40	0.38	0.29	0.28	0.28	0.24	0.27	0.27	0.26	0.26	0.28	0.28	0.31	0.49
Q	0.40	0.66	0.97	0.97	0.97	0.70	0.08	0.84	0.60	0.94	0.90	0.96	0.95	0.95	0.94	0.97	1.00	0.57	0.21	0.37	0.36	0.27	0.26	0.25	0.21	0.22	0.22	0.22	0.21	0.24	0.23	0.35	0.49
R	0.47	0.30	0.53	0.60	0.61	0.41	0.13	0.83	0.87	0.51	0.50	0.59	0.61	0.61	0.59	0.57	0.57	1.00	0.14	0.57	0.70	0.35	0.24	0.29	0.27	0.34	0.33	0.34	0.34	0.37	0.39	0.53	0.37
S	0.08	0.34	0.37	0.30	0.28	0.28	0.38	0.20	0.10	0.25	0.22	0.29	0.28	0.28	0.30	0.27	0.21	0.14	1.00	0.26	0.03	0.17	0.24	0.25	0.25	0.18	0.19	0.19	0.18	0.17	0.14	0.20	0.10
T	0.39	0.20	0.35	0.39	0.43	0.16	0.02	0.44	0.38	0.35	0.39	0.45	0.45	0.44	0.45	0.40	0.37	0.57	0.26	1.00	0.42	0.47	0.38	0.39	0.37	0.47	0.49	0.51	0.52	0.54	0.55	0.59	0.41
U	0.39	0.13	0.34	0.39	0.38	0.30	0.20	0.59	0.67	0.36	0.31	0.38	0.38	0.39	0.39	0.38	0.36	0.70	0.03	0.42	1.00	0.43	0.30	0.37	0.36	0.44	0.41	0.44	0.42	0.45	0.47	0.38	0.30

V	0.45	0.23	0.22	0.29	0.33	0.10	0.16	0.31	0.25	0.24	0.29	0.33	0.33	0.33	0.34	0.29	0.27	0.35	0.17	0.47	0.43	1.00	0.95	0.92	0.88	0.86	0.86	0.86	0.86	0.83	0.82	0.78	0.56	0.65
W	0.34	0.24	0.19	0.25	0.30	0.02	0.20	0.22	0.11	0.25	0.30	0.31	0.31	0.31	0.32	0.28	0.26	0.24	0.24	0.38	0.30	0.95	1.00	0.98	0.95	0.89	0.89	0.87	0.84	0.82	0.78	0.48	0.64	
X	0.31	0.22	0.18	0.23	0.29	0.03	0.12	0.23	0.13	0.26	0.31	0.30	0.30	0.30	0.31	0.28	0.25	0.29	0.25	0.39	0.37	0.92	0.98	1.00	0.98	0.94	0.93	0.91	0.89	0.88	0.84	0.46	0.62	
Y	0.24	0.16	0.13	0.19	0.25	0.06	0.14	0.19	0.09	0.23	0.29	0.26	0.26	0.25	0.27	0.24	0.21	0.27	0.25	0.37	0.36	0.88	0.95	0.98	1.00	0.97	0.96	0.94	0.92	0.91	0.88	0.42	0.62	
Z	0.27	0.17	0.16	0.21	0.28	0.05	0.19	0.23	0.14	0.26	0.32	0.29	0.28	0.28	0.30	0.27	0.22	0.34	0.18	0.47	0.44	0.86	0.89	0.94	0.97	1.00	0.99	0.98	0.98	0.96	0.95	0.46	0.68	
AA	0.27	0.19	0.15	0.21	0.28	0.06	0.19	0.23	0.14	0.26	0.33	0.29	0.29	0.28	0.30	0.27	0.22	0.33	0.19	0.49	0.41	0.86	0.89	0.93	0.96	0.99	1.00	0.99	0.99	0.97	0.95	0.46	0.70	
BB	0.28	0.15	0.16	0.21	0.28	0.04	0.16	0.24	0.16	0.26	0.33	0.30	0.29	0.29	0.31	0.26	0.22	0.34	0.19	0.51	0.44	0.86	0.87	0.91	0.94	0.98	0.99	1.00	0.99	0.99	0.97	0.95	0.48	0.68
CC	0.27	0.13	0.15	0.21	0.28	0.04	0.18	0.23	0.14	0.25	0.34	0.29	0.29	0.28	0.30	0.26	0.21	0.34	0.18	0.52	0.42	0.83	0.84	0.89	0.92	0.98	0.99	0.99	1.00	0.99	0.98	0.46	0.68	
DD	0.26	0.11	0.18	0.24	0.30	0.03	0.14	0.27	0.18	0.26	0.36	0.32	0.32	0.31	0.34	0.28	0.24	0.37	0.17	0.54	0.45	0.82	0.82	0.88	0.91	0.96	0.97	0.99	0.99	1.00	0.99	0.47	0.65	
EE	0.26	0.11	0.19	0.24	0.31	0.03	0.15	0.28	0.20	0.26	0.37	0.33	0.33	0.32	0.35	0.28	0.23	0.39	0.14	0.55	0.47	0.78	0.78	0.84	0.88	0.95	0.95	0.97	0.98	0.99	1.00	0.46	0.64	
FF	0.82	0.29	0.29	0.34	0.37	0.17	0.04	0.48	0.49	0.23	0.33	0.37	0.38	0.38	0.37	0.31	0.35	0.53	0.20	0.59	0.38	0.56	0.48	0.46	0.42	0.46	0.46	0.48	0.46	0.47	0.46	1.00	0.57	
GG	0.49	0.54	0.43	0.44	0.53	0.08	0.36	0.44	0.28	0.48	0.53	0.52	0.52	0.52	0.49	0.49	0.37	0.10	0.41	0.30	0.65	0.64	0.62	0.62	0.68	0.70	0.68	0.68	0.65	0.64	0.57	1.00		

Tableau 31 : Légende pour la matrice de corrélation croisée

Wn_DI	A	P_princ	R
Wn_Station	B	S_princ	S
L_Axis	C	P_princ_10	T
Perim	D	v_18	U
Superficie	E	s1	V
Coef_compa	F	s2	W
Latitude_m	G	s3	X
Coef_mass	H	s4	Y
Coef_orog	I	s5	Z
Shreve	J	s6	AA
Strahler	K	s7	BB
Pmoy_a	L	s8	CC
Pmax_m	M	s9	DD
IF	N	s10	EE
Qe	O	v_20 (EOSD)	FF
L_drainage	P	v_211 (EOSD)	GG
L_princ	Q		

Annexe IV : Procédure pour déterminer le débit sur des images RADARSAT pour un site non jaugé.

La présente procédure utilise des outils développés en Python en utilisant la librairie d'ArcGIS, ArcPy, et ajoutés à une boîte à outils dans ArcGIS. Seule la partie *Orthorectification des images RADARSAT* nécessite d'autres logiciels comme Geomatica. Ces outils ont été développés dans le but de démontrer la faisabilité du calcul du débit pour un site non-jaugé. Ils ne sont pas exempt de bugs et ne possèdent pas la robustesse d'une version commerciale.

1- Création d'un nouveau projet

Cet outil (Figure 39) permet de créer un nouveau projet (NomDeProjet.pro) et un répertoire de données associé (NomDeProjet.data). Le fichier de projet est un fichier texte contenant un ensemble de paramètres à définir dans les différentes étapes de cette procédure afin de permettre le calcul de débits pour un site non-jaugé. Le répertoire de données est un dossier permettant de regrouper les données nécessaires à la procédure. Plusieurs traitements devant être fait subséquentment nécessitant que les données soient projetées, un système de coordonnées projetées doit être sélectionné pour le projet. Le fichier de projection (.prj) est copié dans le répertoire de données.

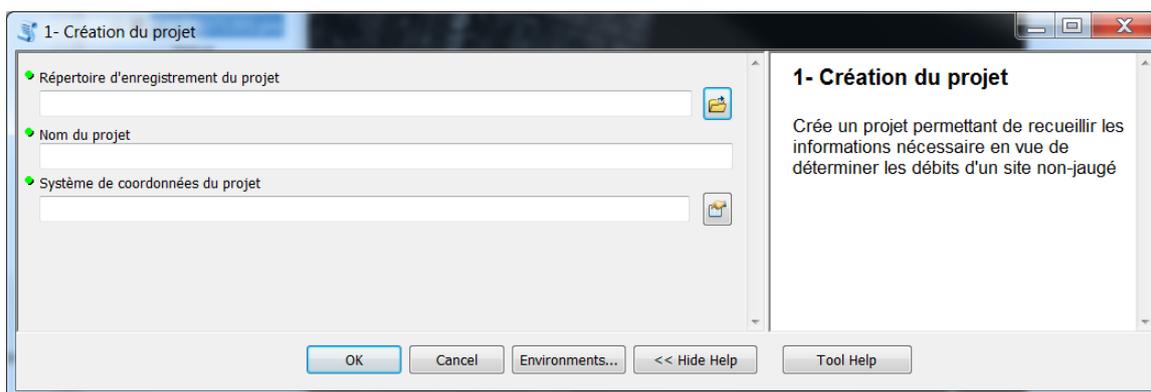


Figure 39 : Boîte de dialogue pour la création d'un nouveau projet

2- Ajout du site non jaugé

Permet de sélectionner un jeu de données de type Point contenant la position du site non-jaugé (Figure 40). Si le jeu de données n'est pas dans le même système de coordonnées que le projet, une transformation est faite avant la copie des coordonnées. Le jeu de données ne doit contenir qu'un seul point (site). À partir des coordonnées de ce point, la région hydrographique correspondante selon l'index du Réseau hydrographique national est déterminée. L'index accompagne la boîte à outils de la présente procédure et se situe dans le sous-dossier /data. Une brève analyse de la valeur du champ NOMWSCSSDA de l'index est faite afin de vérifier si plus d'une région hydrographique pourraient être impliquées dans le

traitement. Si le champ NOMWSCSSDA contient les termes « inférieur » ou « moyen », comme par exemple « Cours inférieur de la rivière X », cela indique que la partie en amont est représentée par au moins une autre unité hydrographique. Dans ce cas, le résultat de l'analyse est porté à l'attention de l'utilisateur et l'outil lui recommande de créer un nouveau jeu de données qui contiendrait des points dans chacune des unités hydrographiques reliées au site non jaugé. Ce nouveau jeu sera nécessaire pour le téléchargement des données hydrographiques.

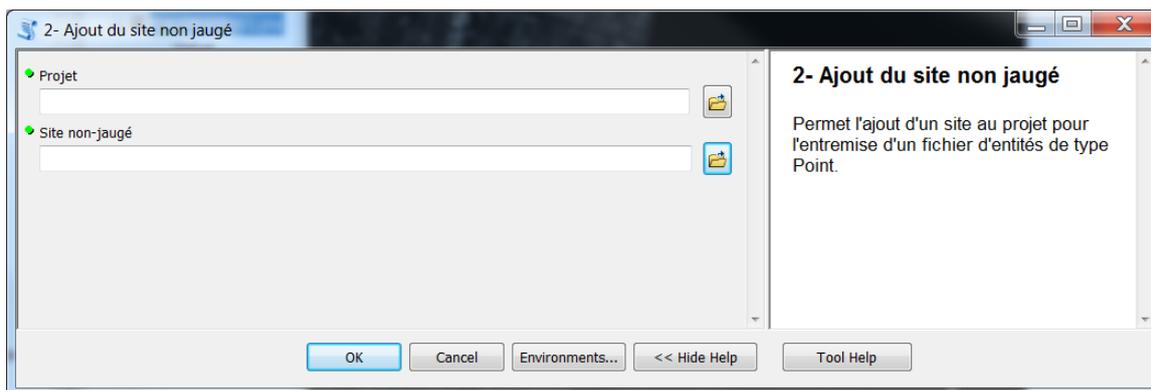


Figure 40 : Boîte de dialogue pour l'ajout d'un site au projet

3- Téléchargement du réseau hydrographique

L'outil procède au téléchargement automatique des données du réseau hydrographique pour la ou les régions hydrographiques chevauchant une zone d'intérêt fourni par l'utilisateur (Figure 41). En absence d'avertissement à l'étape précédente, l'utilisateur devrait reprendre le même fichier de point comme zone d'intérêt. S'il a été déterminé que plusieurs unités hydrographiques étaient impliquées, un nouveau jeu de données contenant des points dans chacune des unités hydrographiques reliées au site non jaugé doit être utilisé. Les données du réseau hydrographique peuvent provenir de deux sources, soit du Réseau hydrographique national (ancienne disponible sur le portail Geobase mais maintenant disponible sur le portail Geogratis) ou du National Hydrography Dataset du USGS. Les deux sources permettent le téléchargement des données par site FTP. Les fichiers d'intérêt sont téléchargés puis les archives sont décompressées. Si l'utilisateur le désire, les archives sont ensuite supprimées.

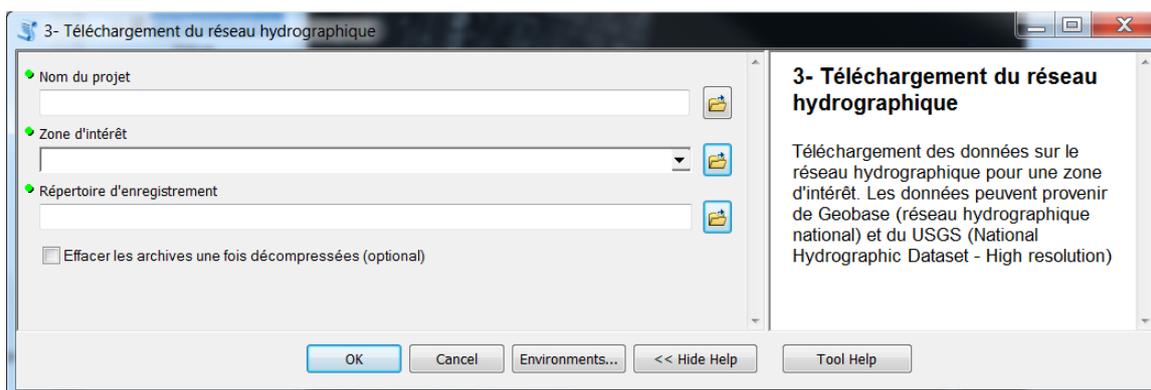


Figure 41 : Boîte de dialogue pour le téléchargement des données du réseau hydrographique

4- Traitement du réseau hydrographique

À cette étape, le jeu de données sur le filamentaire d'écoulement appartenant au réseau hydrographique téléchargé est transformé sous forme de graphe afin d'effectuer une analyse réseau et calculer les ordres de Shreve et Strahler pour chaque segment. Pour les données du Réseau hydrographique national, l'information présente dans la table d'attribut permet de facilement déterminer la connectivité des segments ce qui accélère la création du graphe et diminue considérablement le temps nécessaire à cette étape. Pour les données du National Hydrography Dataset, cette information n'est pas présente de sorte que la connectivité est déterminée de façon itérative par intersection spatiale (i.e les segments touchant le segment de départ sont sélectionnés, puis les segments touchant ces segments, etc.) ce qui peut demander énormément de temps pour des bassins de grande taille. Suite à l'analyse du graphe, l'ordre de Shreve est utilisé afin de déterminer les segments du filamentaire qui appartiennent au cours d'eau principal. Pour ce faire, à partir du site non jaugé vers l'amont, à chaque intersection, le segment sélectionné est celui ayant l'ordre de Shreve le plus élevé, donc celui ayant le plus de tributaire en amont. Tous les segments sélectionnés sont fusionnés pour former le cours d'eau principal lequel est enregistré dans le répertoire de données du projet sous le nom *trib_princ_station.shp*. Ce dernier servira ultérieurement de ligne centrale de la rivière. L'analyse du graphe permet aussi de sélectionner tous les segments en amont du site non jaugé. Une copie de ces segments est faite dans le répertoire de données du projet.

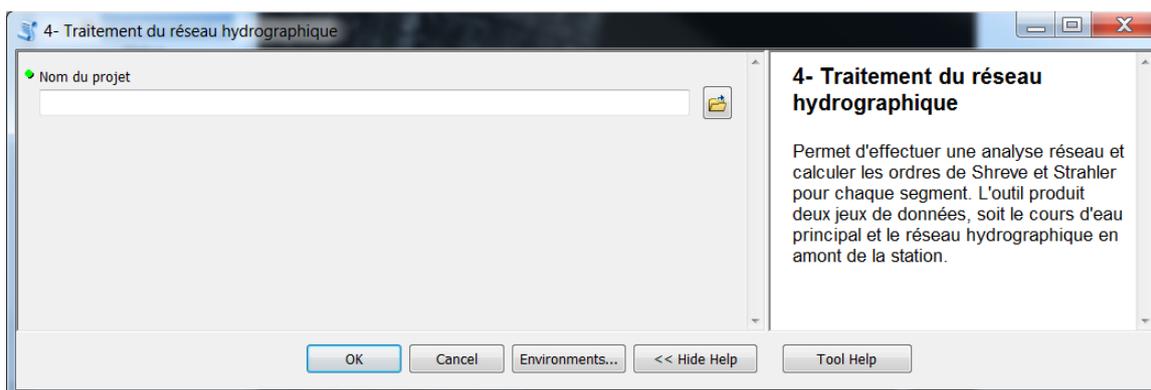


Figure 42 : Boîte de dialogue pour le traitement du réseau hydrographique

5- Création du chenal principal

Le chenal principal correspond au polygone de la rivière auquel a été coupé et éliminé les parties appartenant à des tributaires secondaires. Cette partie de nettoyage doit se faire manuellement dans un logiciel comme ArcGIS. Le cours d'eau principal créé à l'étape précédente peut servir à déterminer les parties du chenal à conserver et celles à éliminer. Dans les données du Réseau hydrographique national, le polygone de rivière se trouve dans le jeu de données *RHN_0XXX00X_X_X.gdb/Dataset/RHN_HD_REGIONHYDRIQUE_2* où les X indiquent des valeurs spécifiques à chacune des unités hydrographiques. Pour les données du National Hydrography Dataset, les polygones se trouvent dans *NHDHXXX.gdb/Hydrography/NHDArea*.

6- Ajout du chenal principal au projet

Une fois que le chenal principal a été créé, il suffit de l'ajouter au registre du projet au moyen de cet outil (Figure 43). Ce dernier permet de lier le fichier existant au projet ou d'en faire une copie dans le répertoire de données du projet.

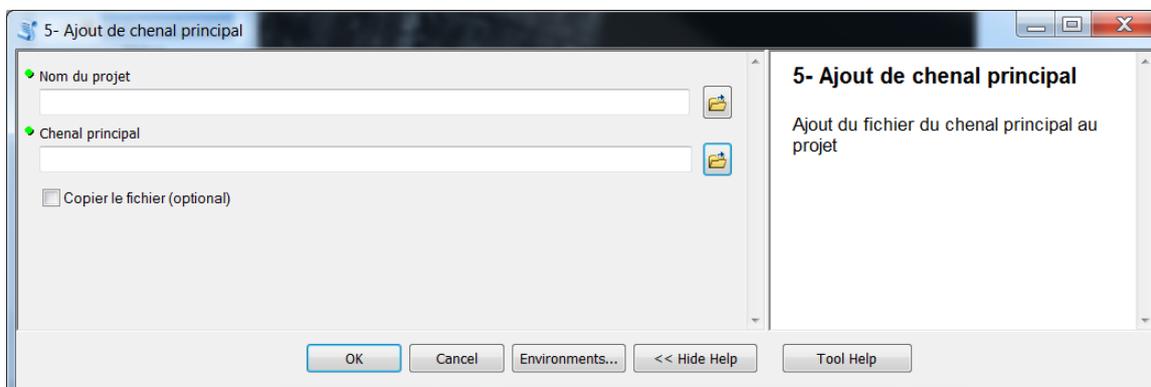


Figure 43 : Boîte de dialogue pour l'ajout du chenal principal

7- Calcul de la largeur à la station

La largeur du chenal principal vis-à-vis du site non jaugé peut être mesurée manuellement par l'utilisateur ou déterminée par l'outil (Figure 44). Si l'utilisateur entre une valeur dans la boîte de texte prévue à cet effet dans l'outil, cette valeur est directement ajoutée au projet. En absence de valeur, l'outil déterminera la largeur en mesurant la distance d'une berge à l'autre le long d'une droite perpendiculaire à la ligne centrale de la rivière (cours d'eau principal) et passant par le site non-jaugé. La mesure de la largeur par l'outil est donc dépendante du niveau de détail et de lissage de la ligne centrale. Une ligne centrale non lissée peut être composée de petits segments ayant des directions différentes de la direction générale du cours d'eau. Ainsi, la droite orthogonale pourrait ne pas être orthogonale à la direction présumée de l'écoulement.

Une fois la largeur obtenue (par l'utilisateur ou calculée), l'outil fera une copie du chenal principal de la rivière à laquelle il appliquera une zone tampon (buffer) équivalente à 100% de la largeur, ou 70m, la plus élevée des deux valeurs.

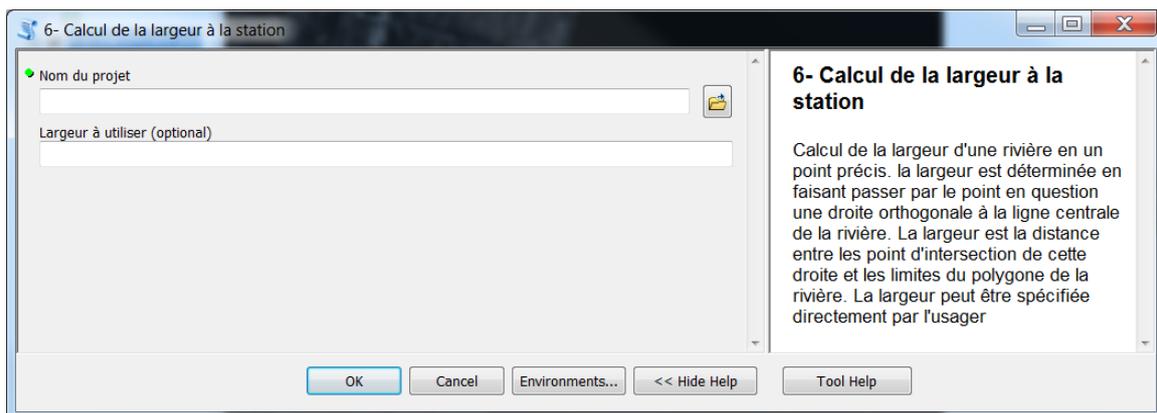


Figure 44 : Boîte de dialogue pour le calcul de la largeur du cours d'eau à la station

8- Création des sections

Lors de la création des sections de rivière, le chenal principal auquel a été appliquée une zone tampon est divisé en tronçon de longueur fixe, soit 1000m dans le cadre de ce projet (Figure 45). La longueur est mesurée le long de la ligne centrale de la rivière. À chaque 1000m, le polygone du chenal principal est coupé le long d'une droite orthogonale à la ligne centrale. Ici encore, la direction de la coupure est dépendante de la ligne centrale. Les sections sont produites pour toute la longueur où il y a un chevauchement de la ligne centrale et du chenal principal. Habituellement, la ligne centrale est plus longue que le chenal principal car à partir d'une certaine largeur, le cours d'eau n'est plus cartographié sous forme de polygone mais par une simple ligne, laquelle n'est plus considérée dans le traitement. Une fois produites, les sections doivent être validées manuellement par l'utilisateur et éditées au besoin.

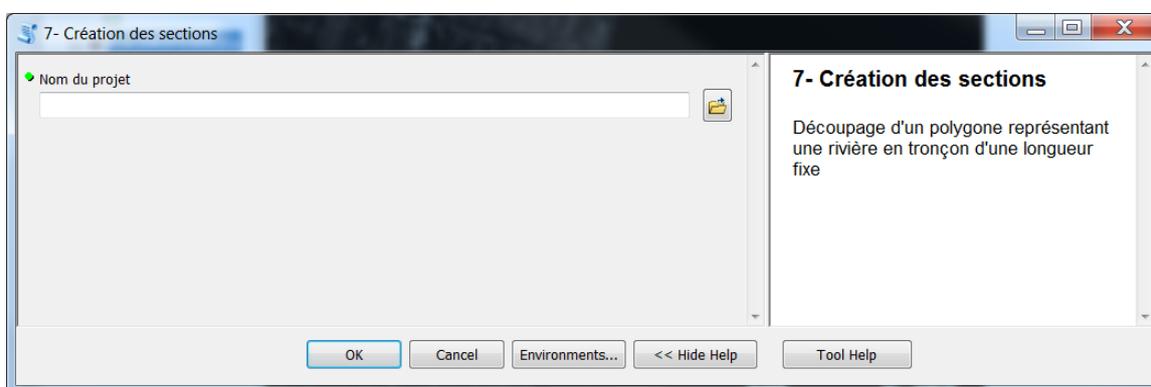


Figure 45 : Boîte de dialogue pour la création des sections de rivière

9- Téléchargement des données d'élévation

Les données d'élévation servent à la détermination des limites du bassin versant et au calcul de certains paramètres physiographiques. Dans le cadre du projet, deux sources de données ont été utilisées, l'une pour le territoire canadien et l'autre pour la partie nord-est des États-Unis. La première est la collection Données numériques d'élévation du Canada à l'échelle 1:50000 distribuée par le portail Geogratis et la seconde est le National Elevation Dataset à une résolution de 1/3 de seconde d'arc. Un index des données des deux collections accompagne la boîte à outils de la présente procédure et se situe dans le sous-dossier /data. Les données canadiennes peuvent être téléchargées automatiquement par l'outil qui en crée ensuite une mosaïque si nécessaire (Figure 46). Les données américaines doivent être téléchargées manuellement par l'utilisateur sur le site <http://viewer.nationalmap.gov/viewer/?p=ned>. Si des données américaines sont nécessaires, l'outil en averti l'utilisateur.

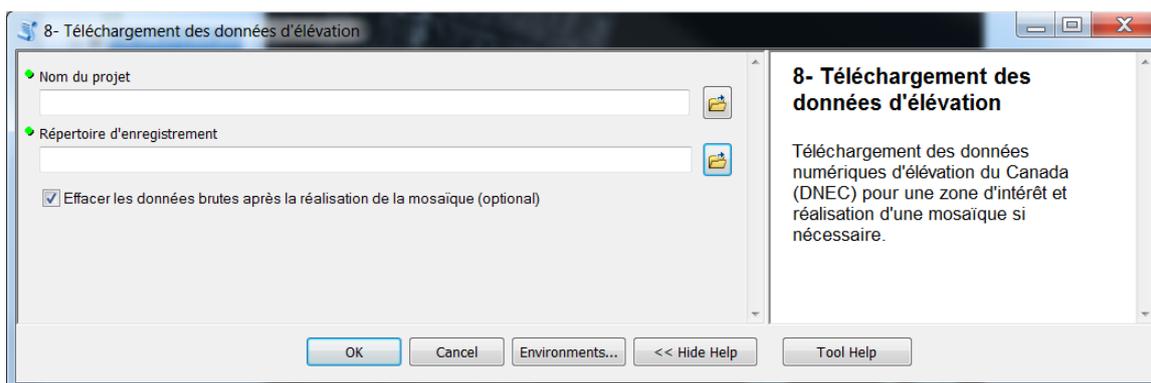


Figure 46 : Boîte de dialogue pour le téléchargement des données d'élévation

10- Génération du bassin versant

Le bassin versant du site non jaugé est déterminé en utilisant les données d'élévation préalablement téléchargées et assemblées (Figure 47). Afin d'améliorer le résultat produit, certains traitements ou choix sont fait en arrière-plan. Tout d'abord, les données d'élévation peuvent être légèrement modifiées afin de supprimer des dépressions locales qui mèneraient à des petits bassins versants à l'intérieur du bassin associé au site non-jaugé. Ensuite, l'algorithme de délimitation du bassin versant nécessite de positionner un exutoire vers lequel l'eau du bassin doit s'écouler. Plutôt que d'utiliser la position du site non jaugé, tout le réseau hydrographique en amont du site est utilisé comme « exutoire ». Ceci évite qu'un léger décalage entre la position du site non jaugé et l'endroit exact où l'écoulement devrait se faire selon les données d'élévation ne vienne fausser le résultat. Finalement, afin d'être cohérent avec le réseau hydrographique, les limites du bassin sont contraintes à ne pas excéder les limites des unités hydrographiques impliquées.

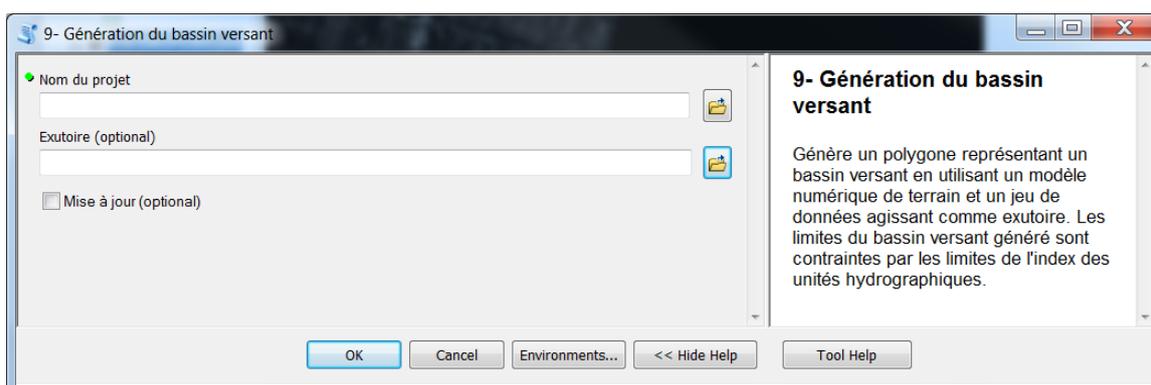


Figure 47 : Boîte de dialogue pour la génération du bassin versant

11- Calcul des paramètres physiographiques

Les données sont maintenant toutes structurées afin de permettre le calcul des paramètres physiographiques d'intérêt pour l'établissement de la courbe de tarage pour le site non jaugé. Cet outil (Figure 48) procède à l'évaluation des paramètres suivants :

- Superficie du bassin versant
- Périmètre du bassin versant
- Coefficient de compacité
- Sinuosité du cours d'eau principal
- Pente transversale à la rivière à l'intérieur des trois premières sections de rivière
- Pourcentage de classe eau à l'intérieur des trois premières sections de rivière
- Largeur moyenne du cours d'eau pour les trois premières sections de rivière

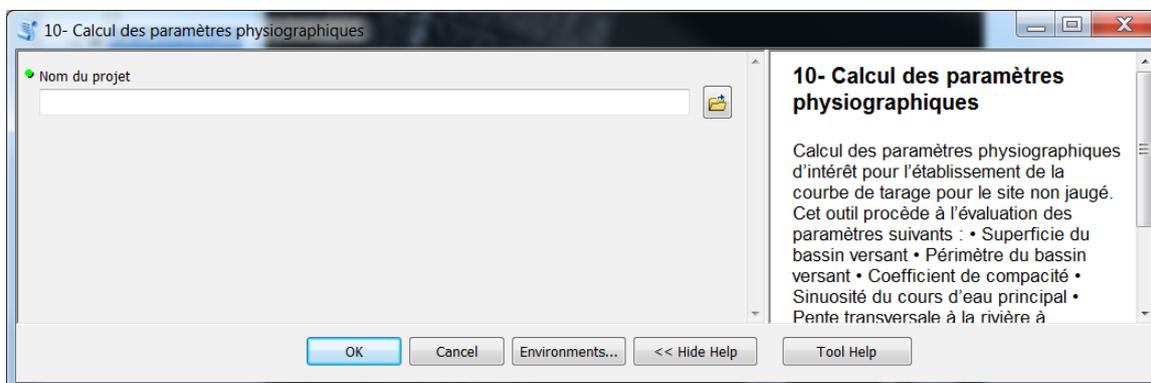


Figure 48 : Boîte de dialogue pour le calcul des paramètres physiographiques

12- Orthorectification et filtrage des images RADARSAT

Les prétraitements des images RADARSAT sont laissés au soin de l'utilisateur et peuvent être réalisés à partir d'un logiciel comme Geomatica. Le format de fichier étant différent pour RADARSAT-1 et RADARSAT-2, les traitements à appliquer pour chacun des formats est légèrement différent. Pour RADARSAT-1, certains prétraitements sont nécessaires pour rendre le fichier brut apte à l'orthorectification dans Geomatica. Pour l'orthorectification, des données d'élévation couvrant toute l'étendue de l'image sont nécessaires. Ces données proviennent des mêmes sources que pour la détermination des limites du bassin versant. Les données canadiennes peuvent donc être téléchargées automatiquement si nécessaire, et assemblées en mosaïque au besoin.

13- Calcul des largeurs effectives

La largeur effective correspond à une largeur de rivière déterminée à partir de la classification des pixels des images RADAR se trouvant à l'intérieur des sections du chenal principal. Suite à la classification, chaque pixel se voit attribué une probabilité entre 0 et 1 d'être de l'eau. La somme des probabilités, divisée par la longueur des sections donne la largeur effective. Dans le processus, les images sont tout d'abord alignées spectralement en faisant correspondre le 1^{er} percentile à une valeur de -23 dB puis une fonction d'appartenance est appliquée pour déterminer la probabilité de chaque pixel (Figure 49).

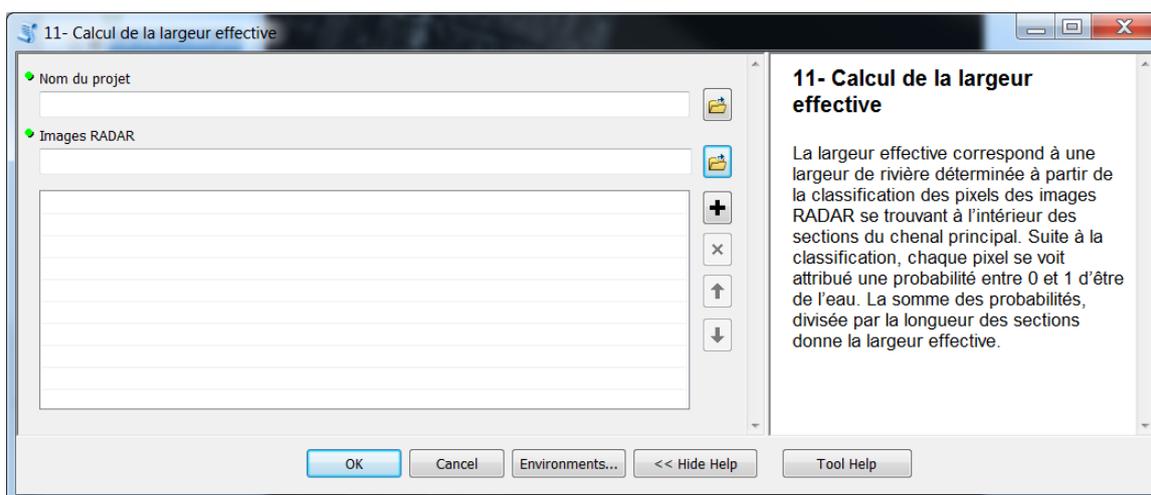


Figure 49 : Boîte de dialogue pour le calcul de la largeur effective

14- Calcul des débits

À partir des largeurs effectives et des paramètres physiographiques, il est possible de déterminer le débit sur chacune des images RADAR en utilisant le fichier de calibration qui accompagne la boîte à outils de la présente procédure et se situe dans le sous-dossier /data. L'outil produit un rapport regroupant l'information sur les données utilisées dans le projet, les traitements effectués et les résultats obtenus.

Annexe V : Spécification des exigences associées au système et aux interfaces

Dans le cadre de ce projet, trois logiciels commerciaux ont été utilisés pour le développement. Il s'agit de Geomatica de PCI Geomatics, Matlab de Mathworks et ArcGIS de ESRI.

Geomatica

Le logiciel Geomatica a été utilisé pour la lecture et les traitements de base des images RADARSAT-1 et RADARSAT-2. Il a permis de lire les images brutes, de les orthorectifier et de les filtrer. Au besoin, les images orthorectifiées ont été recorrigées afin de satisfaire les critères de qualité fixés. Toutes ces opérations sont possibles avec la version de base du logiciel.

Matlab

L'utilisation faite de Matlab se rapporte principalement à la procédure de régionalisation des courbes de tarage. Dans le contexte d'une application de la méthode, Matlab n'est pas requis. Toutefois, pour re-calibrer le modèle il faudra utiliser les mêmes fonctions que celles employées pour la calibration. Dans le cas présent, les toolboxes nécessaires sont : *Statistic Toolbox* et *Financial Toolbox*. *Statistic toolbox* est requis pour l'analyse de la normalité des données, l'analyse en composantes principales, l'analyse canonique de corrélation, la méthode CART pour la développer des arbres de régressions et faire des régressions multiples. Bien qu'il s'agisse de statistiques, *Financial Toolbox* a été requis pour la transformation de puissance BoxCox.

ArcGIS

ArcGIS, et le langage de programmation Python, ont été utilisés à de nombreuses fins tout au long du projet. Ils ont permis le traitement des données spatiales nécessaires au calcul des variables physiographiques décrivant chaque station, la visualisation de ces mêmes données, la production de carte, l'automatisation des processus (même ceux impliquant Geomatica) et la mise en place de la procédure de traitement pour un site non-jaugé. La version *Advanced* du logiciel ArcGIS a été nécessaire à la réalisation du projet et à l'application sur un site non-jaugé. Pour ce qui est de Python, il s'agit d'un langage de programmation orienté-objet sous une licence libre. Différentes bibliothèques ont dû être installées pour réaliser les travaux désirés (Tableau 32).

Tableau 32 : Description des bibliothèques Python utilisées

Librairie	Fonction	Version utilisée
NumPy	Calcul numérique. La librairie NumPy v1.6.1 est installée par défaut avec ArcGIS mais une version plus récente était nécessaire pour être compatible avec la librairie SciPy.	v1.9.0
SciPy	Ajustement des fonctions de régression, réalisation d'histogrammes	v 0.14.0
xlrd, xlwt	Lecture et écriture des fichiers Excel	v0.9.3; v0.7.5
networkx	Traitement des réseaux hydrographiques; calcul des ordres de Shreve et Strahler	v1.7

D'autres bibliothèques ont été produites dans le cadre du projet (Tableau 33)

Tableau 33 : Bibliothèques Python développées pour le projet

Librairie	Fonction
teledebit_fct	Permet de réaliser différentes opérations mathématiques, géométriques et spatiales particulièrement utiles au projet
teledebit_projet	Gestion et manipulation du fichier de projet pour l'application à un site non-jaugé.