

**Méthodes de détermination de la ligne des hautes eaux des cours d'eau en milieu agricole et urbain dans un contexte de changements climatiques**

**Rapport d'étape**

**Par**

**Freddy Houndekindo<sup>1</sup>, Mathieu Vaillancourt<sup>2</sup>, Charles Gignac<sup>1</sup>, Sophie Duchesne<sup>1</sup>,  
Monique Poulin<sup>2</sup>, Normand Bergeron<sup>1</sup>, André St-Hilaire<sup>1</sup>**

**<sup>1</sup> INRS-ETE**

**<sup>2</sup> Université Laval**

**Novembre 2020**

**Référence à citer :**

Houndekindo, F., M. Vaillancourt, C. Gignac, S. Duchesne, M. Poulin, N.E. Bergeron, A. St-Hilaire. 2020. Méthodes de détermination de la ligne des hautes eaux des cours d'eau en milieu agricole et urbain dans un contexte de changements climatiques. Rapport d'étape. INRS-ETE #R1989, 89 pages

© INRS, Centre - Eau Terre Environnement, 2020  
Tous droits réservés

ISBN : 978-2-89146-965-4 (version électronique)

Dépôt légal - Bibliothèque et Archives nationales du Québec, 2022  
Dépôt légal - Bibliothèque et Archives Canada, 2022

## Table des matières

1. Contexte et objectif .....	1
2.0 Méthode botanique.....	1
2.1 Sélection des sites	1
2.2 Protocole terrain	3
2.3 Analyses préliminaires	5
2.4 Résultats préliminaires	8
3. Méthode hydrologique/hydraulique.....	11
3.1 Introduction	11
3.2 Étude hydrologique	12
Vérification des données	12
Formation des régions homogènes	13
Test d'adéquation et distribution régionale	15
Estimation des quantiles	17
Régression par modèle additif généralisé (GAM)	17
Pénalisation et sélection des variables significatives	18
3.3. Modélisation hydraulique	19
3.3.1 Méthodologie	19
3.3.3 Résultats	21
3.3.6 Calage du modèle hydraulique	27
3.4 Carte d'inondation	37
4. Approche géomorphologique .....	39
4.1 Utilisation du Lidar pour la méthode géomorphologique	39
4.2 Méthode par débordement	40
4.3 Analyse préliminaire sur les sites tests	43
4.3 Méthode par seuillage de pente	46
4.4 Résultats obtenus sur les sites tests	47
4.5 Création automatisée de modèles bathymétriques pour la méthode hydraulique	49
5. Analyses à venir.....	50
6. Difficultés rencontrées .....	50
Annexes .....	53
Annexe 1 : Résultats de l'analyse fréquentielle locale	54
Annexe 2 : résultats de l'analyse fréquentielle régionale	62

Annexe 3 : Résultats de la modélisation hydraulique 73

Annexe 4 : Pénalisation de la fonction de lissage dans mgcv et sélection des variables

85

## 1. Contexte et objectif

L'équipe de chercheurs de l'INRS-ETE et de l'Université Laval a été mandatée par le Ministère de l'Environnement de la Lutte aux Changements Climatiques (MELCC) afin de procéder à une analyse comparative de différentes méthodes de détermination de la ligne des hautes eaux (LHE) et ce, dans les contextes anthropiques urbanisés et agricoles.

Trois types d'approches sont comparés dans le cadre de ce projet : Les méthodes botanique, hydraulique et géomorphologique. Ce rapport d'étape présente les protocoles et méthodologies de chacune des approches et les résultats préliminaires, qui sont focalisés sur trois cours d'eau parmi la liste des rivières sélectionnées pour le projet. Les sections suivantes présentent ces méthodes et résultats.

## 2. Méthode botanique

### 2.1 Sélection des sites

L'objectif pour le volet botanique est de tester la méthode botanique experte pour délimiter la ligne des hautes eaux, en la comparant notamment à des méthodes alternatives qui font appel à la modélisation hydraulique ou géomorphologique. Nous avons comme cible de sélectionner un minimum de 8 sites correspondant à chacune des catégories de paysage, soit agricole, urbain ou forestier. Un site représente un tronçon de rivière d'environ 100 m de longueur.

Une première sélection de sites a été réalisée à partir d'un fichier de forme (*shapefile*) des rivières dont la ligne des hautes eaux (LHE) avait déjà été modélisée. Chaque tronçon modélisé était ensuite converti en une série de points espacés de 20 m permettant de délimiter des zones tampons de 100 m de rayon. Il a ainsi été possible de faire une analyse de paysage indépendante pour chaque rive des segments modélisés, en tenant compte des superficies agricoles (fichier provenant de la Base de données des parcelles et productions agricoles déclarées), des superficies boisées (fichier provenant de la Base des données topographiques du Québec) et des superficies urbanisées (fichier provenant de DMTI Spatial). Cela nous a permis de décrire les tronçons et de pouvoir par la suite

sélectionner un segment uniforme au sein de ceux-ci. Ainsi, les rives caractérisées par une proportion de plus de 85% de superficies agricoles, urbaines ou forestières étaient sélectionnées. Cette sélection de sites par analyse spatiale a par la suite été validée par une analyse visuelle des cartes satellites pour prendre en compte les changements plus récents dans l'occupation du territoire. Puis, la sélection a été raffinée selon différents critères. Les deux rives opposées devaient correspondre au même type de paysage afin que le milieu soit le plus uniforme possible. Les segments en milieu urbain et en milieu agricole devaient avoir une bande riveraine assez large, qui allait au-delà de la LHE. Les tronçons urbains et agricoles ne devaient pas présenter d'ouvrage de stabilisation des berges. L'amont des tronçons choisis devait être caractérisé le plus possible (en longueur et en proportion de la zone tampon) par le même type de paysage, considérant que les patrons de biodiversité sont structurés non seulement par les dimensions latérales et verticales du cours d'eau, mais également par la dimension longitudinale (Naiman et al. 2005). Pour cette même raison, un maximum d'un site par catégorie a été sélectionné par rivière, pour éviter l'autocorrélation spatiale.

Cette sélection de sites a été complétée par une sélection de cours d'eau non modélisés. Pour ce faire, une sélection a été effectuée à l'aide de Google Maps parmi les rivières situées dans le territoire de la Communauté métropolitaine de Québec (CMQ). Les sites ainsi sélectionnés ont ensuite été filtrés selon la possibilité d'en faire une modélisation hydraulique.

À la suite de ces étapes, les 51 tronçons obtenus ont été visités afin de confirmer leur faisabilité sur le terrain. À la fin de la saison, huit tronçons par catégorie ont pu être inventoriés pour un total de 24 sites (fig. 1). Cette sélection finale comprend des rivières de diverses largeurs (3-30 m) réparties dans le territoire de la CMQ. Le site 17 a été inclus, bien qu'il se trouve plus éloigné des autres, afin d'inclure une rivière de plus de 30 m de largeur dans la catégorie agricole.

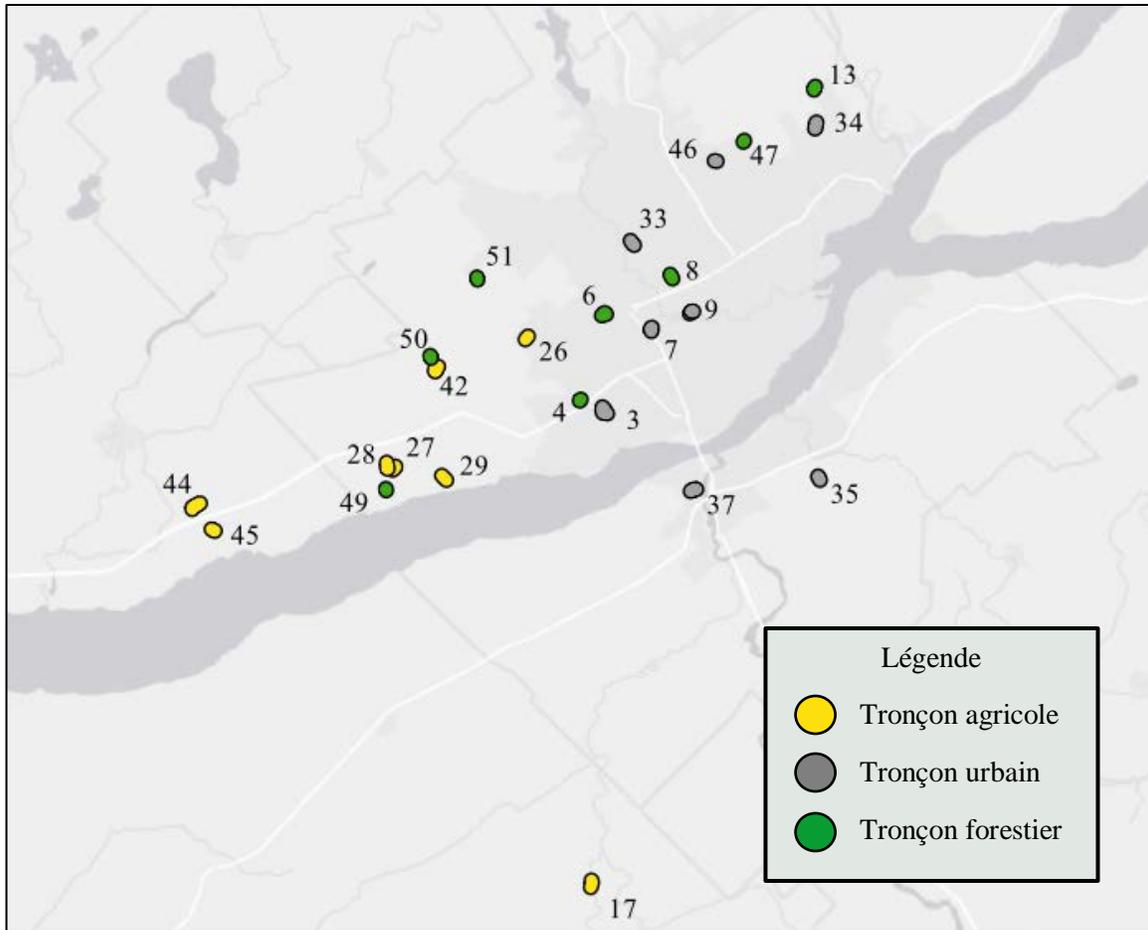


Figure 1. Représentation géographique des 24 sites inventoriés à l'été 2020.

## 2.2 Protocole terrain

Le protocole pour les inventaires a été défini selon la note explicative sur la délimitation de la LHE selon la méthode botanique experte (MDELCC 2015).

À l'arrivée au site, le tronçon était parcouru afin de trouver un segment qui soit positionné le plus loin possible d'un ponceau et qui ne soit pas trop méandreux. Une fois le segment délimité au site, trois transects étaient placés dans trois zones bien végétalisées présentant peu ou pas d'érosion. Au sein du segment, les transects étaient espacés le plus possible les uns des autres (jusqu'à un maximum de 50 m), étaient placés idéalement là où les berges étaient végétalisées, tout en s'assurant qu'il n'y ait entre eux aucune arrivée d'eau (drain, tributaire) et en s'assurant que la sinuosité du segment n'était pas trop grande. Chaque transect était positionné perpendiculairement au cours d'eau, la

limite inférieure correspondant à la limite inférieure de la végétation aquatique et la limite supérieure étant placée au-delà de la LHE. Ainsi, la limite supérieure pouvait correspondre au replat (en zone agricole), à une élévation dépassant le niveau plein-bord lorsque celui-ci était visible, ou encore à une élévation caractérisée par une communauté végétale de milieu terrestre, qui soit distincte du reste de la berge. La largeur du transect était toujours d'un mètre (50 cm de part et d'autre de la ligne tracée). Deux types d'inventaires étaient réalisés, soit par points de contact et par quadrats.

L'inventaire par points de contact était réalisé à tous les 50 cm le long du transect, à raison de six points espacés de 20 cm sur la largeur (parallèle au cours d'eau). Toutes les espèces végétales qui entraient en contact avec ces points le long d'une tige infinie étaient inventoriées. La strate de chaque espèce était également notée (herbacée, arbustive, arborée, lianes).

L'inventaire par quadrats était réalisé à l'aide de quadrats de 0,5 m<sup>2</sup> (100 cm sur la largeur du transect et 50 cm sur la longueur), espacés régulièrement les uns des autres de bas en haut du transect, pour un maximum de huit quadrats par transect. Lorsque ceux-ci faisaient 4 m ou moins de longueur, les quadrats étaient donc placés côte à côte, en continu. Pour les cas de transects faisant plus de 20 m dont les communautés végétales étaient distinctes et homogènes, ou encore lorsque l'élévation n'augmentait pas de façon régulière de bas en haut du transect (par la présence d'un replat au centre, par exemple), les quadrats pouvaient être espacés de façon irrégulière afin de rendre compte de la variabilité du milieu. Dans ce cas, les quadrats étaient placés afin d'être le plus contrastés possible en termes de communautés végétales et d'élévation. Pour chaque quadrat, les données suivantes étaient notées : la position sur le transect, l'ouverture de la canopée, le recouvrement des strates arborée, arbustive, herbacée et muscinale, le recouvrement en sol nu et en litière, le substrat et la présence d'érosion. Comme pour les points de contact, toutes les espèces de toutes les strates présentes dans le quadrat, en bas comme en haut, étaient inventoriées, leur classe de recouvrement absolu notée (<1%, 1-5% 5-10%, 10-25%, 25-50%, 50-75%, 75-100%), ainsi que leur strate. Les espèces qui n'étaient pas identifiées sur place étaient récoltées pour identification ultérieure au laboratoire. Les spécimens qui étaient au stade de plantule et qui étaient par conséquent impossibles à

identifier ont été notées comme telle afin de documenter cette limite de la méthode botanique. Les jeunes arbres et arbustes qui étaient enracinés dans les quadrats étaient considérés comme faisant partie de la strate herbacée.

Les transects étaient géoréférencés sur toute leur longueur à raison d'un point à tous les 50 cm à l'aide d'un système GNSS (Leica Geosystems), ce qui a permis de géoréférencer chaque point d'inventaire. Lorsque l'appareil était incapable de trouver sa position en raison d'un couvert de canopée trop important, le positionnement était fait à partir de points dégagés situés à proximité et l'élévation prise à l'aide d'un altimètre de haute précision (ZIPLEVEL ®, Technidea corporation), ou encore à l'aide d'une station totale (Cansel).

### 2.3 Analyses préliminaires

Les analyses préliminaires ont été réalisées à partir des données de quadrats venant de trois sites. Les sites en question correspondent aux rivières Cap-Rouge (site forestier), des Roches (site agricole) et du Berger (site urbain). Pour chaque quadrat, différents calculs ont été réalisés afin de classer la communauté inventoriée dans la catégorie des plantes de milieux humides ou de milieux terrestres (Tableau 1). Les quadrats ont d'abord été analysés selon (I) la méthode botanique experte du MDDELCC (2015) : le nombre d'espèces de milieux humides présentes dans le quadrat, telles que définies à l'annexe I de la méthode, était divisé par le nombre total d'espèces présentes, toutes strates de végétation confondues, sans tenir compte de leur recouvrement respectif; les quadrats obtenant un score de 50% ou moins étaient classés comme étant sous la LHE.

**Tableau 1.** Liste des calculs réalisés pour établir une LHE à partir de différentes méthodes botaniques.

	<b>Méthode</b>	<b>Calcul</b>
I	Experte	$(Nb_{OBL}^1 + Nb_{FACH}^2) / (Nb_{TOT}^3)$
II	Experte sans les arbres	$(Nb_{OBL} + Nb_{FACH}) / (Nb_{TOT})$
III	Experte, herbacées seulement	$(Nb_{OBL} + Nb_{FACH}) / (Nb_{TOT})$
IV	Experte, pondérée par recouvrements	$(Rec_{OBL}^4 + Rec_{FACH}) / (Rec_{TOT})$
V	Experte, OBL x2	$(2Nb_{OBL} + Nb_{FACH}) / (2Nb_{OBL} + Nb_{FACH} + Nb_{NI})$
VI	Experte, OBL x2, pondérée par recouv.	$(2Rec_{OBL} + Rec_{FACH}) / (2Rec_{OBL} + Rec_{FACH} + Rec_{NI})$
VII	Index de prévalence, trois niveaux	$(Rec_{OBL} + 2Rec_{FACH} + 4Rec_{NI}) / Rec_{TOT}$
VIII	Index de prévalence, cinq niveaux	$(Rec_{OBL} + 2Rec_{FACH} + 3Rec_{FAC} + 4Rec_{FACT} + 5Rec_{UPL}) / Rec_{TOT}$

Légende - Nb Nombre d'espèces, Rec : Somme des recouvrements, OBL : espèces obligées des milieux humides, FACH : espèces facultatives des milieux humides, NI : espèces non indicatrices de milieux humides, FAC : espèces facultatives, FACT : espèces facultatives de milieux terrestres, UPL : espèces de milieux terrestres.

Afin d'éliminer l'influence d'espèces présentes dans le quadrat, mais n'y étant pas enracinées (comme les arbres, dont la canopée peut s'étendre sur une superficie de plusieurs mètres), le calcul a également été réalisé (II) sans tenir compte de la strate arborée et (III) en ne tenant compte que de la strate herbacée, donc sans tenir compte des strates arborées ni arbustives (aucun des quadrats ne contenait de lianes). Puisqu'il est suggéré dans la méthode experte de tenir compte du recouvrement des espèces, le calcul a été réalisé (IV) en utilisant le recouvrement des espèces dans le calcul, plutôt que les données de présence/absence (recouvrement cumulé en espèces indicatrices de milieu humide/recouvrement cumulé en espèces totales), et ce, toujours en ne tenant compte que de la strate herbacée. Considérant que les espèces obligées des milieux humides (OBL) sont, selon la définition donnée par le MDELCC (2015), « presque exclusivement restreintes aux milieux humides », elles sont plus fortement indicatrices que les espèces facultatives des milieux humides (FACH), qui ne sont que « généralement restreintes aux milieux humides » (MDELCC 2015). Le calcul a donc également été fait cette fois (V) en doublant le poids des espèces OBL pour la méthode experte (basée sur le nombre d'espèces, ou présence/absence), et (VI) en doublant le poids des espèces OBL pour la méthode experte basée sur le recouvrement des espèces plutôt que les données de

présence/absence. Ces deux derniers calculs ont également été réalisés en ne tenant compte que de la strate herbacée.

Enfin, la méthode de l'index de prévalence (*prevalence index*; PI) tel que définie dans Lichvar, Gillrich et Ochs (2011) consiste à multiplier le recouvrement des espèces par un facteur correspondant à son statut hydrologique (de 1 à 5 pour les espèces OBL, FACH, facultatives (FAC), facultatives des milieux terrestres (FACT) et des milieux terrestres (UPL) respectivement), puis en divisant le tout par la somme des recouvrements en espèces. Un score de trois ou moins désigne une communauté végétale de milieux humides, les scores de plus de trois, les communautés de milieux terrestres. Selon cette méthode, la ligne des hautes eaux est placée de façon similaire à la méthode experte, à la limite des communautés végétales ayant des scores de  $\leq 3$  et de  $> 3$ . La méthode PI a ainsi été réalisée de deux façons. D'abord, (VII) le statut hydrique des espèces tel que donné à l'annexe I du guide de la méthode experte a été utilisé. Un facteur de 1 était appliqué pour les plantes OBL, de 2 pour les plantes FACH et de 4 pour les plantes non indicatrices (NI), la liste n'incluant pas les indices 3 à 5 et aucune autre classification officielle n'étant connue pour les espèces au Québec. Cependant, la méthode a été reprise (VIII), cette fois en faisant le calcul avec les cinq indices, à partir de classifications existant pour le nord-est des États-Unis; pour ce faire, les espèces pour lesquelles la méthode botanique experte (MDDELCC 2015) ne fournit aucun statut se voyaient attribuer le statut hydrique tel que défini par le US Army Corps of Engineers (Lichvar et al. 2014), le US Department of Agriculture (PLANTS Database; USDA, NCRS 2020) et la flore du Michigan (Michigan Flora Online; Reznicek, Voss et Walters, 2011), en ordre de priorité. Toutes ces analyses permettent de déterminer si un quadrat est caractérisé par une végétation de milieu humide ou de milieu terrestre. Bien que théoriquement, sur un transect uniforme, les communautés végétales passent graduellement de caractéristiques des milieux humides à caractéristiques des milieux terrestres, les communautés rencontrées sur le terrain étaient parfois plus variables, et il arrivait d'inventorier des transects où plusieurs transitions de milieu humide à milieu terrestre étaient repérées. Dans ces cas, la LHE était placée à la transition la plus élevée (Fig. 2).

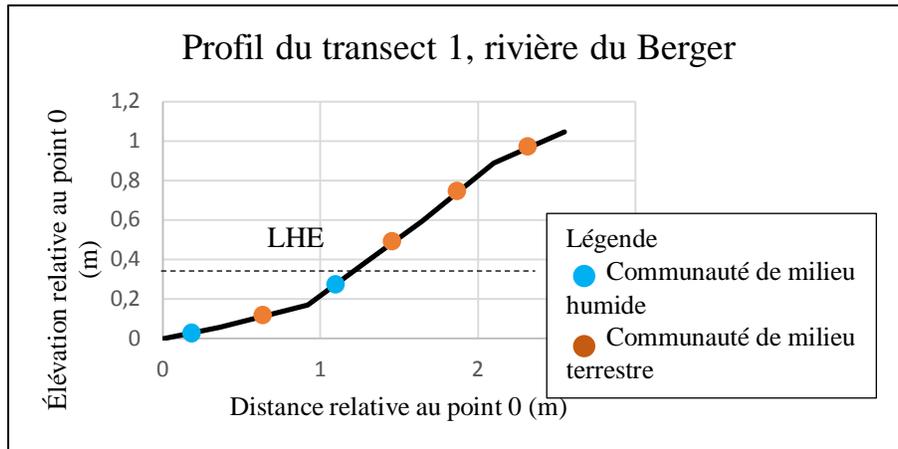


Figure 2. Identité des quadrats le long du profil d'élévation du transect 1, rivière du Berger.

## 2.4 Résultats préliminaires

Les premières analyses permettent d'obtenir des LHE différentes selon la méthode de calcul, incluant également la hauteur de la LHE établie à l'aide de la méthode hydraulique (Tableau 2). Pour un même transect, les hauteurs de la LHE pouvaient aller jusqu'à une différence d'élévation de 2 m selon la méthode utilisée.

La méthode hydraulique donne généralement la ligne des hautes eaux la plus élevée de toutes les méthodes : en moyenne, sur l'ensemble des neuf transects utilisés pour l'analyse préliminaire aux trois sites choisis, les LHE obtenues avec la méthode experte sont 69 cm plus basses. Les méthodes botaniques ayant la plus grande différence sont les méthodes II et III (méthode experte sans les arbres et sans les arbres et arbustes), avec des LHE situées en moyenne 75 cm sous la LHE établie à l'aide de la méthode hydraulique, alors que la méthode ayant la différence la moins grande est la méthode VIII (méthode de l'index de prévalence avec les cinq niveaux), qui place la LHE en moyenne 10 cm en-dessous.

**Tableau 2.** Élévation absolue (en m) de la LHE botanique selon différentes méthodes de calcul

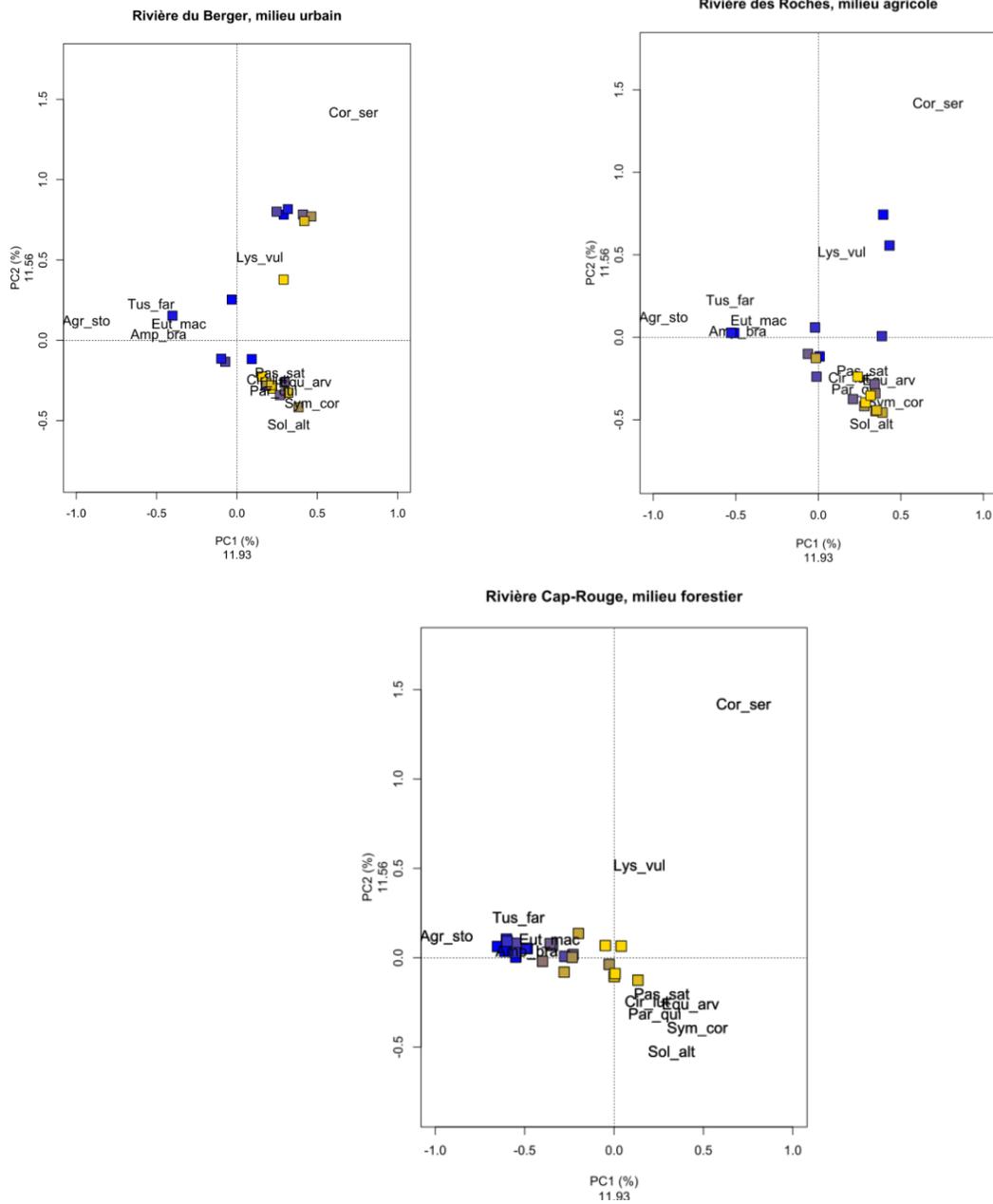
Rivière Site Transect Méthode	Cap-Rouge 4			Du Berger 9			Des Roches 28		
	1	2	3	1	2	3	1	2	3
	<b>Élévation de la LHE (m)</b>								
Experte (E)	21,65	<20,63 <sup>2</sup>	20,57	5,86 <sup>1</sup>	<5,23	<5,21	58,73 <sup>1</sup>	59,03 <sup>1</sup>	59,71 <sup>1</sup>
E sans arbres	21,65	<20,63	20,57	5,86 <sup>1</sup>	<5,23	5,37	<58,38	58,68	59,71 <sup>1</sup>
E sans arbres/arbust.	21,65	<20,63	20,57	5,86 <sup>1</sup>	<5,23	5,37	<58,38	58,68	59,71 <sup>1</sup>
E pondérée	21,50	20,85	20,91 <sup>1</sup>	7,38	<5,23	5,37	59,21	59,43	60,01
E OBL x2	21,65	20,85	20,57	5,86 <sup>1</sup>	<5,23	5,37	59,21	59,43 <sup>1</sup>	<58,96
E pondérée et OBL x2	<21,42	20,85	20,91 <sup>1</sup>	7,38	<5,23	5,37	59,21	59,43	60,01
Index de préval. (IP)	21,50	20,85	20,57	7,38	<5,23	5,37	59,21	59,43	60,01
PI avec 5 niveaux	21,77	21,29 <sup>1</sup>	>21,62 <sup>1</sup>	>7,51	5,65	5,37	59,21	59,43 <sup>1</sup>	>60,08
Méthode hydraulique	<21,42	21,07	20,57	7,38	>7,23	6,91	59,09	59,56	59,63

1 Transect ayant plusieurs transitions milieu humide-milieu terrestre.

2 Les symboles > et < signifient que la LHE obtenue est située au-delà des limites supérieures ou inférieures des transects; la valeur donnée correspond donc respectivement au quadrat le plus élevé ou le plus bas.

Une analyse en composantes principales a été faite sur R (fonction RDA de Vegan) en utilisant les données de végétation des quadrats des trois sites analysés (Fig. 3). Les projections montrent que pour ces trois sites, trois tendances caractérisent les communautés. Les communautés dominées par *Agrostis stolonifera* et *Tussilago farfara* sont fortement associées aux bas des transects, alors que les communautés dominées par *Solidago altissima* et *Symphyotrichum cordifolium* sont plutôt associées à des quadrats placés plus haut sur la berge. Les quadrats caractérisés par la présence de *Cornus stolonifera*, une espèce souvent utilisée dans la stabilisation de berge, et par *Lysimachia vulgaris*, une espèce potentiellement envahissante selon le PLANTS Database (USDA, NCRS 2020), ne sont pour leur part pas associés à une position précise le long du transect.

Ces observations n'étant basées que sur trois sites, l'analyse ne permet pas d'aller plus loin dans l'interprétation. Les résultats préliminaires permettent toutefois de penser que ce genre d'analyse permettrait d'identifier quelles espèces sont les plus fortement influencées par l'inondation et de déterminer des communautés représentant le bas et haut de berge.



*Figure 3 PCA des trois sites utilisés pour l'analyse préliminaire. Les points bleus représentent les quadrats situés au bas du transect et les points jaunes, ceux situés en haut. Les valeurs intermédiaires sont colorées en dégradé.*

### 3. Méthode hydrologique/hydraulique

#### 3.1 Introduction

Afin de réaliser une comparaison entre la méthode botanique et la méthode hydraulique pour la détermination de la LHE, une étude du comportement hydraulique de certains tronçons de rivières a été réalisée. Ce rapport porte sur les résultats initiaux obtenus pour les sites 4, 9, et 28 situés respectivement sur les rivières Cap Rouge, du Berger et des Roches.

Les trois sites présentés dans ce rapport sont localisés dans la région de Québec (voir Figure 1 et Tableau 2 ci-haut).

##### Site 4

Le site 4 est localisé dans le parc les Sentiers de la rivière du Cap Rouge. Ce tronçon de la rivière du Cap Rouge traverse notamment un milieu de type forestier. Le bassin versant de la rivière du Cap Rouge occupe une superficie d'environ 78 km<sup>2</sup>. Au site 4, la superficie du bassin versant est d'environ 62 km<sup>2</sup>.

##### Site 9

Le site 9 est un tronçon de la rivière du Berger situé entre le boulevard Central et la rue Labrecque. Cette section de la rivière du Berger traverse une zone urbanisée et draine une superficie d'environ 60,9 km<sup>2</sup>. La superficie totale du bassin versant de la rivière du Berger est de 61 km<sup>2</sup>.

##### Site 28

Situé dans la municipalité de Saint-Augustin-de-Desmaures, le site 28 est une section de la rivière des Roches. Cette rivière draine principalement des milieux agricoles. La superficie du bassin versant de la rivière des Roches au site 28 est de 14,9 km<sup>2</sup>, ce qui représente environ 86 % de l'aire totale du bassin versant de la rivière des Roches.

### 3.2 Étude hydrologique

Les sites présentés dans ce rapport ne sont pas jaugés. De ce fait, l'approche retenue pour la détermination des débits associés aux crues de récurrence 2 ans est l'analyse fréquentielle régionale (AFR). Dans cette étude, l'AFR s'est déroulée en plusieurs étapes, à savoir :

- La vérification des données de débit
- L'analyse fréquentielle locale
- La formation des régions homogènes
- L'ajustement des courbes de croissance régionales par la méthode des L-moments
- L'utilisation des modèles additifs généralisés (GAM) pour la régression entre les débits maximaux annuels moyens et les variables physiographiques.

#### Vérification des données

Les données utilisées pour la régionalisation ont auparavant été utilisées pour une analyse fréquentielle locale. Au cours de cette analyse, les hypothèses de base ont été vérifiées et les sites les rejetant ont été écartés.

Pour chaque série de débits maximums annuels observés aux stations d'intérêt les étapes de l'analyse fréquentielle locale incluent :

- 1) La vérification des hypothèses de variables indépendantes et identiquement distribuées (IID) et de la stationnarité
- 2) L'ajustement des lois statistiques à l'échantillon par la méthode du maximum de vraisemblance
- 3) Le choix de la meilleure distribution basée sur les critères d'adéquation AIC et BIC
- 4) L'estimation des quantiles d'intérêt.
- 5) La détermination des intervalles de confiance par «bootstrap»

En outre, Hosking et Wallis (1997) recommandent une évaluation de la discordance  $D_i$  au sein de l'ensemble des sites sélectionnés. Cette analyse permet de répertorier les sites présentant des erreurs grossières dans leurs données ou ayant des caractéristiques

grandement différentes de celles des autres sites (ex : station hydrométrique en aval d'un barrage). Le calcul de la discordance  $D_i$  est effectué par la formule suivante :

*Équation 1*

$$D_i = \frac{1}{3} N(u_i - \bar{u})^T A^{-1} (u_i - \bar{u})$$

où  $D_i$  est la mesure de discordance pour le site  $i$ ,  $N$  est le nombre de sites,  $u_i$  est un vecteur formé par les ratios de L-moments  $t$ ,  $t_3$ , et  $t_4$  du site  $i$  (voir Hosking et Wallis, 1997, pour une description détaillée de ces ratios), et  $\bar{u}$  est la moyenne non pondérée des vecteurs  $u_i$ .

Formation des régions homogènes

La détermination des régions homogènes est une étape cruciale de la régionalisation. Elle permet de former des régions hydrologiquement homogènes afin de rendre possible le transfert d'informations sur les débits extrêmes de sites jaugés vers des sites d'intérêts non jaugés pour l'estimation des quantiles (Burn 1997). Dans le cadre de la formation de régions homogènes, deux techniques statistiques d'analyse multivariée ont été appliquées aux données, à savoir l'analyse de corrélation canonique (ACC) et la classification ascendante hiérarchique (*hierarchical cluster analysis*). L'ACC permet de comparer deux groupes de variables afin de déterminer si elles décrivent le même phénomène. Dans cette étude, l'ACC a été utilisée pour déterminer les coefficients appropriés à appliquer aux variables physiographiques et climatiques avant la classification ascendante hiérarchique. Les variables (variables canoniques) qui résultent de l'ACC sont des combinaisons linéaires des variables physiographiques et climatiques qui maximisent la corrélation entre ces dernières et les variables hydrologiques et statistiques directement associées au phénomène à étudier (L-Coefficient de variation ( $\tau$ ), L-Asymétrie ( $\tau_3$ ), L-Aplatissement ( $\tau_4$ ) et logarithme du débit maximal annuel moyen). Ultérieurement à l'ACC, les deux premières variables canoniques produites ont été utilisés comme données d'entrée pour la classification ascendante hiérarchique. Les régions formées par classification ascendante hiérarchique des variables canoniques ont été évaluées pour déterminer leur hétérogénéité.

Hosking et Wallis (1997) ont proposé la mesure  $H$  de l'hétérogénéité d'une région. Cette mesure permet de vérifier si la variation  $V$  (Équation 2) des ratios de L-moments observée entre tous les sites de la région analysée est suffisamment similaire à ce qui est attendu d'une région homogène ayant les mêmes ratios de L-Moments. Pour déterminer la variation des rapports de L-moments attendue d'une région homogène similaire, nous avons procédé par simulation. Hosking et Wallis (1997) suggèrent, pour la simulation des régions homogènes, l'utilisation de la distribution Kappa à 4 paramètres, qui est moins restrictive que les distributions à 2 ou 3 paramètres (ex : Gumbel, Generalized Extreme Value Distribution (GEV)). Cela permet d'éviter d'émettre des hypothèses sur la distribution sous-jacente des données lors du test d'hétérogénéité.

Les quatre paramètres de la distribution Kappa (Équation 5) ont été estimés par la méthode des L-Moments en utilisant les rapports de L-Moments régionaux  $\tau^R$ ,  $\tau_3^R$  et  $\tau_4^R$  et le L-Moment d'ordre 1 ( $\lambda_1$ ). Puisque les équations reliant les paramètres de la distribution Kappa et les rapports de L-Moments n'ont pas de solution analytique, une solution numérique a été obtenue par la méthode d'itération Newthou-Raphson. Une fois la distribution ajustée, nous avons tiré aléatoirement de la Kappa  $n$  vecteurs de dimension  $k$ , où  $n$  est le nombre de stations dans la région et  $k$  est le nombre d'années d'enregistrements observés à chaque station de la région. En conséquence, les régions simulées ont le même nombre de sites et le même nombre de données par site que les régions qui ont été préalablement formées. Ces simulations ont été répétées 500 fois par région et nous avons calculé pour chaque région la variation  $V$  des L-Coefficients de variation ( $\tau$ ) suivant l'Équation 2. Nous avons ensuite procédé au calcul de la moyenne et de l'écart-type de la variation  $V$  par région. Finalement, la moyenne et l'écart-type obtenus ont été utilisés pour calculer la mesure  $H$  de l'hétérogénéité suivant l'Équation 4.

*Équation 2*

$$V = \left\{ \frac{\sum_i^N n_i (t^{(i)} - t^R)^2}{\sum_{i=1}^N n_i} \right\}$$

où  $t^R$  est la moyenne régionale des ratios de L-moments pondérés proportionnellement à la longueur d'enregistrement par site, tel qu'exprimé à l'Équation 3 :

Équation 3

$$t^R = \frac{\sum_i^N n_i t^{(i)}}{\sum_{i=1}^N n_i}$$

Équation 4

$$H = \frac{V - \mu_v}{\sigma_v}$$

où  $\mu_v$  et  $\sigma_v$  sont respectivement la moyenne et l'écart-type de la variation des ratios de L-moments des régions simulées.

Équation 5

$$F(x) = [1 - h\{1 - k(x - \xi)/\alpha\}^{1/k}]^{1/h}$$

où  $F(x)$  est la distribution cumulative de distribution Kappa et  $\xi$  et  $\alpha$  sont respectivement les paramètres de position et d'échelle de la distribution Kappa.

Le Tableau 3 présente les valeurs seuils pour la mesure d'hétérogénéité et les conclusions associées selon Hosking et Wallis (1997).

*Tableau 3 : Interprétation de la mesure de l'hétérogénéité d'une région (Hosking et Wallis 1997)*

H	Conclusion
$H < 1$	La région est acceptablement homogène
$1 \leq H < 2$	La région pourrait être hétérogène
$H \geq 2$	La région est définitivement hétérogène

Test d'adéquation et distribution régionale

L'adéquation de l'ajustement des lois de probabilité est évaluée suivant la mesure de l'adéquation  $Z^{DIST}$  proposée par Hosking et Wallis (1997). Cette mesure de l'adéquation se base sur les ratios de L-Moments d'ordre 4 ( $\tau_4$ ). L'adéquation de la distribution est évaluée par l'écart entre le ratio de L-Moments de la distribution ajustée  $t_4^{dist}$  et celui de la région formée  $t_4^R$ . Donc, plus l'écart est proche de zéro, meilleure est l'adéquation. Afin

d'évaluer la signification de cet écart et de tenir compte de la variabilité d'échantillonnage, nous avons simulé un grand nombre (500) de régions homogènes à partir d'une distribution Kappa ajustée aux ratios de L-moments régionaux de chaque région formée. Les régions simulées ont le même nombre de sites et le même nombre de données par site que les régions étudiées. Le test d'adéquation proposé par Hosking et Wallis (1997) a été réalisé afin de déterminer la distribution la plus adéquate entre la GEV et la Pearson 3. Ces deux distributions sont les distributions à trois paramètres qui ont été retenues selon les résultats de l'analyse fréquentielle locale. En effet, Hosking et Wallis (1997) recommandent l'utilisation de distributions ayant 3 à 5 paramètres dans l'analyse fréquentielle régionale. Les paramètres de la distribution retenue ont été estimés pour chaque région homogène par la méthode des L-Moments.

Pour chaque simulation de région avec la distribution Kappa, nous avons calculé les ratios régionaux de L-Moments d'ordre 4 avec l'Équation 3. Le biais ( $B_4$ ) et la déviation standard ( $\sigma_4$ ) ont été calculés avec les formules suivantes :

*Équation 6*

$$B_4 = N_{sim}^{-1} \sum_{m=1}^{N_{sim}} (t_4^{(m)} - t_4^R)$$

*Équation 7*

$$\sigma_4 = \left[ (N_{sim} - 1)^{-1} \left\{ \sum_{m=1}^{N_{sim}} (t_4^{(m)} - t_4^R)^2 - N_{sim} B_4^2 \right\} \right]^{1/2}$$

La mesure de l'adéquation se calcule de la manière suivante :

*Équation 8*

$$Z^{DIST} = (t_4^{dist} - t_4^R + B_4) / \sigma_4$$

### Estimation des quantiles

L'estimation des quantiles à chaque site a été réalisée par la méthode de l'*index flood* (Dalrymple 1960) selon la formule suivante :

#### Équation 9

$$\hat{Q}_i(F) = \hat{\mu}_i \hat{q}(F)$$

où  $\hat{\mu}_i$  correspond au débit maximal annuel moyen au site  $i$ . Pour les sites non jaugés, cette moyenne est estimée par un modèle additif généralisé (GAM) qui prend en entrée les données climatiques et physiographiques du site. Le modèle est initialement calé par région à partir des débits maximaux annuels moyens observés aux sites jaugés de la région (Noto et la Loggia 2009).  $\hat{q}(F)$  correspond au facteur de croissance obtenu à partir de la distribution régionale ajustée par la méthode des L-moments pour une période de retour donnée.

#### Régression par modèle additif généralisé (GAM)

La détermination des débits maximaux annuels moyens aux sites non jaugés a été réalisée par régression avec des modèles additifs généralisés. Les modèles additifs généralisés (Hastie et Tibshirani 1986) offrent une plus grande flexibilité que les modèles linéaires en permettant de modéliser des relations non linéaires. Pour une variable dépendante  $Y$ , les modèles additifs généralisés peuvent être exprimés comme suit :

#### Équation 10

$$g(E(Y|\mathbf{X})) = \alpha + \sum_{j=1}^p f_j(X_j) + \varepsilon$$

où  $f_j$  est la fonction de lissage de la variable indépendante  $X_j$  et peut être définie comme suit :

#### Équation 11

$$f_j(x) = \sum_{i=1}^q \beta_{ji} b_{ji}(x)$$

où les  $b_{ji}(x)$  sont des fonctions de base qui sont combinées linéairement pour l'estimation de la fonction de lissage et  $q$  est la dimension des fonctions de base (le nombre de nœuds).

Pénalisation et sélection des variables significatives

Le package *Mixed GAM Computation Vehicle* (*mgcv*) (Wood 2001) dans R offre la possibilité d'ajuster des modèles GAM pénalisés pour les fonctions de lissage trop oscillantes. Le degré de pénalisation est contrôlé par l'hyperparamètre  $\lambda$ . Pour  $\lambda = 0$ , la fonction de lissage n'est pas pénalisée et est donc sur-ajustée aux données observées. Lorsque  $\lambda \rightarrow \infty$ , la fonction de lissage est sous-ajustée aux données et tend à une ligne droite. Une valeur adéquate de  $\lambda$  permet d'obtenir un modèle qui est un bon compromis entre le modèle sous-ajusté et le modèle sur-ajusté. Wood (2006) propose plusieurs techniques de sélection de la valeur de  $\lambda$  qui sont basées sur des critères de performance du modèle dont, entre autres, l'erreur de prédiction par validation croisée (GCV) et l'espérance de l'erreur quadratique (UBRE).

D'autre part, grâce à l'option « Select », *mgcv* permet de réaliser une sélection automatique (Marra et Wood 2011) des variables significatives telle que décrite à l'Annexe 4. De ce fait, le choix des variables importantes pour la régression a été fait directement dans *mgcv* en activant l'option « Select ». Nous avons testé toutes les variables d'intérêt et retenu celles qui n'ont pas été éliminées par le modèle, et qui ensuite améliorent l'erreur quadratique relative.

Une fois les modèles GAM ajustés avec les variables significatives, l'erreur quadratique relative (RRMSE) et le biais relatif (RBIAS) ont été utilisés pour évaluer les performances des modèles par la validation croisée.

*Équation 12*

$$\text{Erreur quadratique relative : } RRMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N \left( \frac{y_i - \hat{y}_i}{y_i} \right)^2}$$

### Équation 13

$$\text{Erreur relative : } RBIAS = \frac{1}{N} \sum_{i=1}^N \left( \frac{y_i - \hat{y}_i}{y_i} \right)$$

où  $y_i$  est la valeur observée et  $\hat{y}_i$  est la valeur simulée.

### 3.3. Modélisation hydraulique

L'objectif principal de la modélisation hydraulique est la détermination des profondeurs d'eau correspondant au débit de crue de projet. Cette analyse a été réalisée avec le modèle hydrodynamique HEC-RAS.

#### 3.3.1 Méthodologie

La mise en place du modèle hydraulique appliqué dans cette étude s'est déroulée en plusieurs étapes, à savoir :

- La collecte des données sur le terrain
- La construction de la géométrie de la rivière dans le logiciel de SIG Quantum-GIS (QGIS)
- L'importation et la saisie des données dans le logiciel HEC-RAS
- Le calage du modèle
- La simulation des crues de récurrence de 2 ans.

#### *Collecte des données et importation des données dans HEC-RAS*

Les données nécessaires au calage d'un modèle hydraulique 1D ont été recueillies sur les sites à l'étude. Ces données comprennent :

- La bathymétrie de la rivière à des sections transversales du tronçon modélisé. Trois à cinq sections ont été considérées en fonction des conditions du site (ex : ponceau ou apport d'eau en aval)
- Le débit s'écoulant dans la rivière
- Les hauteurs d'eau observées à des endroits précis des sections transversales
- La pente de la surface de l'eau

- Les informations permettant une caractérisation de la rugosité du fond de la rivière et des berges (type de substrat, végétation, présence de blocs de roche, etc.).

Les sections transversales ont été géoréférencées et représentées dans QGIS sous forme de lignes. Ensuite, ces données ont été importées dans HEC-RAS en tant que fichier Shapefile à l'aide du module RAS-MAPPER. Les informations bathymétriques collectées sur le terrain ont été saisies dans l'éditeur de section transversale du logiciel.

#### *Calage du modèle et simulation de la crue de projet*

Le calage du modèle a consisté à ajuster la valeur du coefficient « n » de Manning à chaque section transversale (fond du cours d'eau) afin que les profondeurs d'eau simulées par le modèle s'approchent le plus possible des profondeurs d'eau observées lors des mesures sur le terrain. Le « n » de Manning est un coefficient représentatif de la rugosité de la surface sur laquelle l'eau s'écoule.

L'évaluation du substrat et de la végétation lors des travaux de terrain nous permet une première estimation du coefficient « n » de Manning. Les matériaux qui composent le lit de la rivière sont analysés et classés en fonction de leur taille. Les matériaux fins, tels que le sable, offrent une plus faible résistance à l'écoulement comparé à des matériaux plus grossiers, tels que les blocs de roche. De plus, Chow (1959) propose des valeurs de « n » de Manning correspondant au type de substrat rencontré dans le fond de la rivière et à d'autres caractéristiques du tronçon de rivière telles que :

- Le degré de méandres
- La végétation
- La variation de la largeur? De la section du tronçon
- Le degré d'irrégularité
- Etc. ...

Les valeurs initiales des « n » de Manning définies selon l'analyse du substrat pour le fond de la rivière ont ensuite été ajustées afin que les profondeurs d'eau simulées s'approchent au mieux des profondeurs d'eau observées sur le terrain. Toutefois, les valeurs de « n » de Manning des berges ont été uniquement estimées par les valeurs

théoriques proposées par Chow (1959). En effet, lors des travaux de terrain, les profondeurs d'eau dans les rivières n'étaient pas assez élevées pour un calage des coefficients « n » de Manning des berges.

La simulation a été réalisée en écoulement uniforme. Les paramètres utilisés pour la modélisation sont les suivants :

- Type d'écoulement : fluvial
- Condition limite amont : débit mesuré
- Condition limite aval : profondeur normale (pente du lit de la rivière).

Une fois le modèle calé, les valeurs de débit de récurrence 2 ans ont été entrées et la simulation a été réalisée sans modifier les paramètres du modèle après calage. Les résultats de la simulation ont été exportés en Shapefile dans QGIS pour réaliser les cartes d'inondation.

### 3.3.3 Résultats

#### *Étude hydrologique*

L'analyse fréquentielle régionale a permis de déterminer les débits correspondant à la crue de récurrence 2 ans. Les résultats de cette analyse sont présentés dans les sections suivantes.

#### *Formation des régions homogènes*

Le Tableau 4 présente les variables physico-climatiques utilisées pour l'analyse canonique des corrélations et la classification hiérarchique ascendante. Le logarithme de la superficie du bassin versant (LGA) a le coefficient le plus élevé pour la 1<sup>ère</sup> variable canonique. Ceci est attendu étant donné que la variable LGA est la plus corrélée aux variables hydrologiques.

Le Tableau 5 montre le nombre de sites par région formée, les L-Coefficients de variation ( $\tau$ ) régionaux et la valeur de la mesure d'hétérogénéité obtenue après la formation des régions et l'élimination des stations discordantes. La Figure 1 montre la distribution spatiale des régions formées.

Tableau 4: Variables physio-climatiques (analyse canonique des corrélations et classification ascendante hiérarchique)

Variables	Logarithme de la superficie du bassin versant (LGA)	% de milieu aquatique dans le bassin versant (AQT)	Pente moyenne du bassin versant (PMN)
Coefficient de la 1 <sup>ère</sup> variable canonique	0,93	-0,31	0,14
Coefficient de la 2 <sup>e</sup> variable canonique	0,30	0,94	0,12

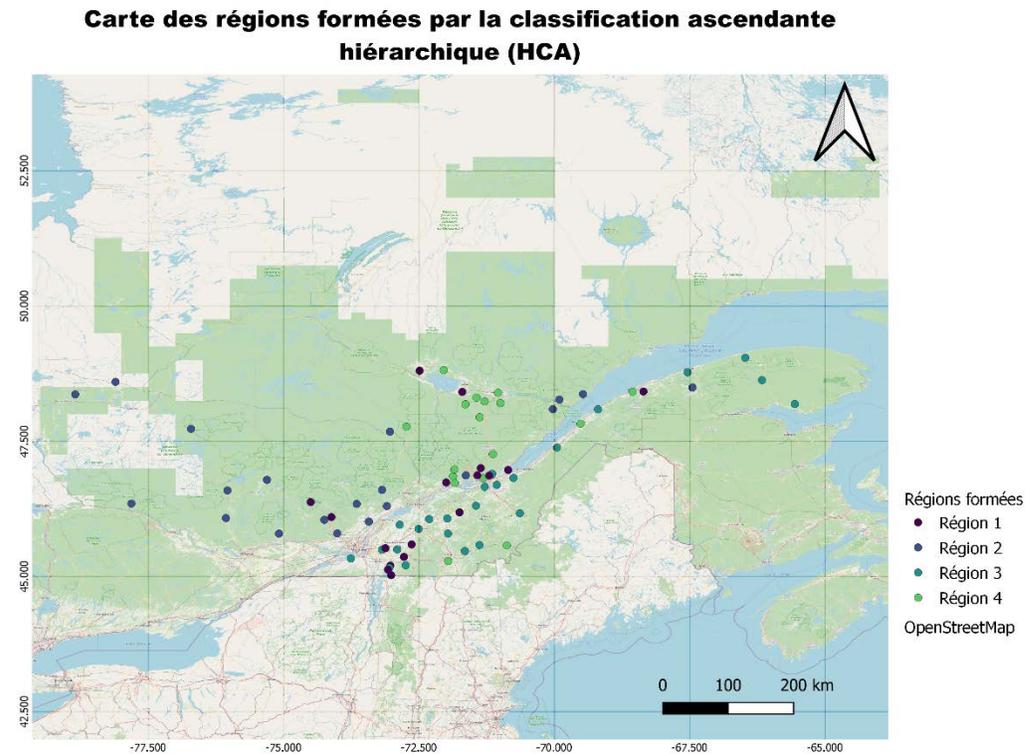


Figure 4 : Carte des régions formées

*Tableau 5 : Mesure d'hétérogénéité (H) des régions formées*

Région	H	Nombre de sites	L-Coefficient de variation
1	0,24	18	0,18
2	1,95	20	0,15
3	1,07	26	0,18
4	0.60	17	0,17

*Test d'adéquation et distribution régionale*

Le Tableau 6 présente les résultats du test d'adéquation proposé par Hosking et Wallis (1997). Il en ressort que la distribution GEV présente le meilleur ajustement pour chaque région; elle est donc retenue pour l'estimation des quantiles.

*Tableau 6 : Résultat du test d'adéquation*

Région	Z <sup>dist</sup> GEV	Z <sup>dist</sup> Pearson 3
1	-0,71	-1,22
2	-0,62	-0,73
3	-0,63	-1,74
4	0,19	-0,97

Le Tableau 7 montre les paramètres de la distribution GEV estimés par la méthode des L-Moments pour chaque région hydrologique formée.

*Tableau 7 : Paramètres estimés des distributions régionales*

Région	A	K
1	0,27	-0,02
2	0,23	0,11
3	0,27	0,04
4	0,28	0,04

La Figure 5 présente les courbes de croissance des régions. La région 1 a les facteurs de croissance les plus élevés.

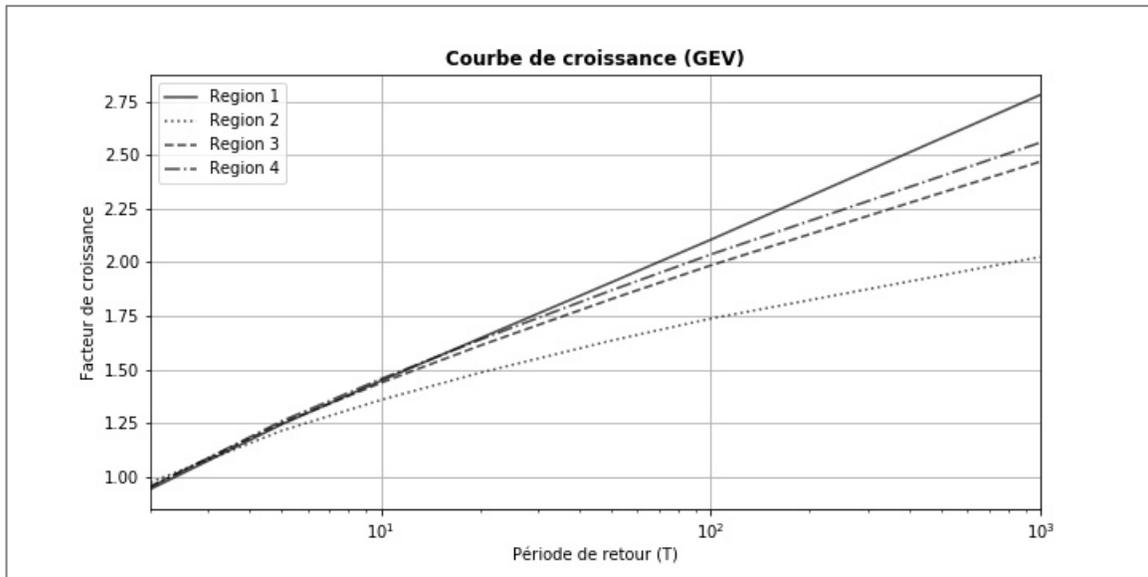


Figure 5 : Courbe de croissance

*Régression par modèle généralisé additif (GAM)*

Les variables réponses retenues dans chaque région et les valeurs des critères de performance (RRMSE et RBIAS) sont présentées au Tableau 8. Les valeurs des critères de performance sont des moyennes obtenues à partir de la validation croisée « *leave-one-out cross-validation* ».

*Tableau 8 : Variables explicatives finales*

	Région 1	Région 2	Région 3	Région 4
Variables explicatives	LGA, AQT, TMN, PMN	LGA, PMN, AQT, AGR	LGA, TMN, AGR	LGA, PPM, PMN
RRMSE	22,8 %	14,6 %	18,1 %	23,4 %
RBIAS	-2,7 %	-0,3 %	-1,6 %	-2,0 %

LGA : Logarithme de la superficie du bassin versant, AQT : % de milieu aquatique dans le bassin versant, PMN : Pente moyenne du bassin versant, AGR : % de milieu agricole dans le bassin versant, TMN : Moyenne des degrés-jours (base de 0°C) cumulés entre 1979 et 2008. PPM : Cumul moyen des précipitations entre le 1<sup>er</sup> avril et le 31 octobre (période de 1974 à 2003)

Les sites 9, 4 et 28 appartiennent tous à la région 1. En effet, la taille du bassin versant ayant été le facteur déterminant dans la formation des régions, la région 1 regroupe les sites de bassins versants de moins de 310 km<sup>2</sup>, ce qui est le cas pour les trois sites à l'étude.

Le Tableau 9 présente les débits obtenus de l'analyse fréquentielle régionale pour la période de retour 2 ans.

*Tableau 9 : Débit théorique de la crue de récurrence 2 ans*

	Site 4	Site 9	Site 28
Débit 2 ans (m <sup>3</sup> /s)	14,60	15,44	4,52

Dans le cadre des projets de détermination des cotes de crues à la rivière du Berger (CEHQ 4132-0509-05-9972) et du Cap Rouge (CEHQ 4132-0539-05-9959), réalisés par le Centre d'expertise hydrique du Québec ((Boucher et Francoeur 2013) (Boucher,

Francoeur et al. 2013)), la détermination des débits de crue a été basée sur un modèle de transformation pluie-débit, en utilisant en entrée des pluies de projet de type SCS (Soil Conservation Service) d'une durée de 24 h. Les résultats obtenus de cette analyse sont présentés dans le tableau 10. Par rapport aux débits résultant de l'analyse fréquentielle régionale, les débits estimés par modélisation pluie-débit sont plus élevés par un facteur d'environ 2,6 pour la récurrence 2 ans.

*Tableau 10 : Quantiles estimés par l'AFR et par la modélisation pluie-débit*

	2 ans		20 ans		100 ans	
	<b>Site 4</b>	<b>Site 9</b>	<b>Site 4</b>	<b>Site 9</b>	<b>Site 4</b>	<b>Site 9</b>
Modèle pluie-débit	38,8	41,0	62,4	70,5	64,0	85,4
AFR	14,6	15,4	25,6	27,0	32,7	34,6

Note : Les débits estimés sont en m<sup>3</sup>/s

### 3.3.6 Calage du modèle hydraulique

Les résultats du calage des modèles sont présentés dans les sections ci-dessous pour quelques sections transversales.

#### Résultats du calage : Site 4

Le Tableau 11 et les Figures 6 et 7 présentent les résultats du calage du modèle hydrodynamique au site 4.

Tableau 11 : Résultat du calage site 4

Section transversale	Point de mesure à partir de la berge gauche (m)	Profondeur d'eau observée (cm)	Profondeur d'eau simulée (cm)	Débit simulé (m <sup>3</sup> /s)	« n » de Manning						
1	15,5	48	43	0,99	<table border="1"> <tr> <td>Rive droite</td> <td>Fond du cours d'eau</td> <td>Rive gauche</td> </tr> <tr> <td>0,15</td> <td>0,0685</td> <td>0,15</td> </tr> </table>	Rive droite	Fond du cours d'eau	Rive gauche	0,15	0,0685	0,15
Rive droite	Fond du cours d'eau	Rive gauche									
0,15	0,0685	0,15									
2	15,5	41	38	0,99	<table border="1"> <tr> <td>Rive droite</td> <td>Fond du cours d'eau</td> <td>Rive gauche</td> </tr> <tr> <td>0,15</td> <td>0,0685</td> <td>0,15</td> </tr> </table>	Rive droite	Fond du cours d'eau	Rive gauche	0,15	0,0685	0,15
Rive droite	Fond du cours d'eau	Rive gauche									
0,15	0,0685	0,15									

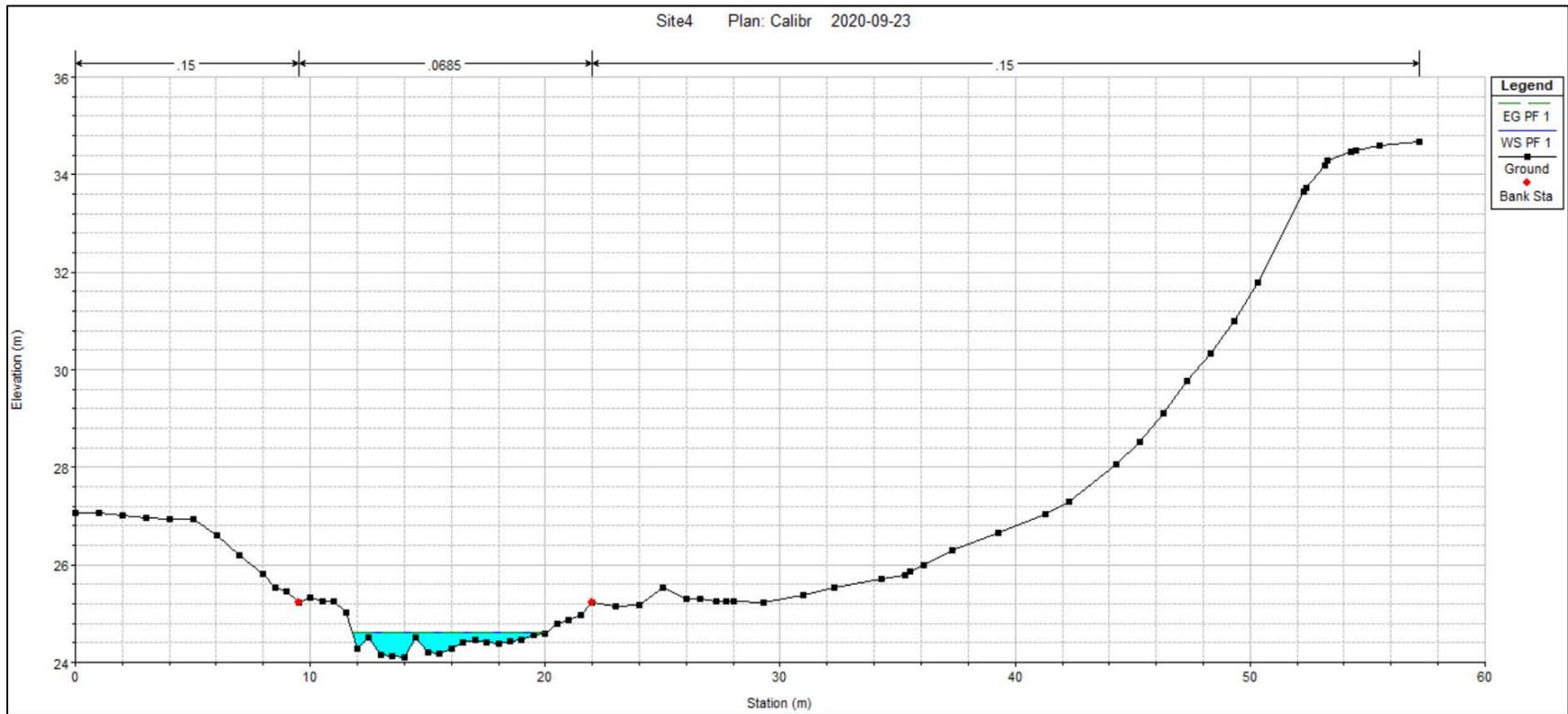


Figure 6 : Section transversale 1 (site 4)

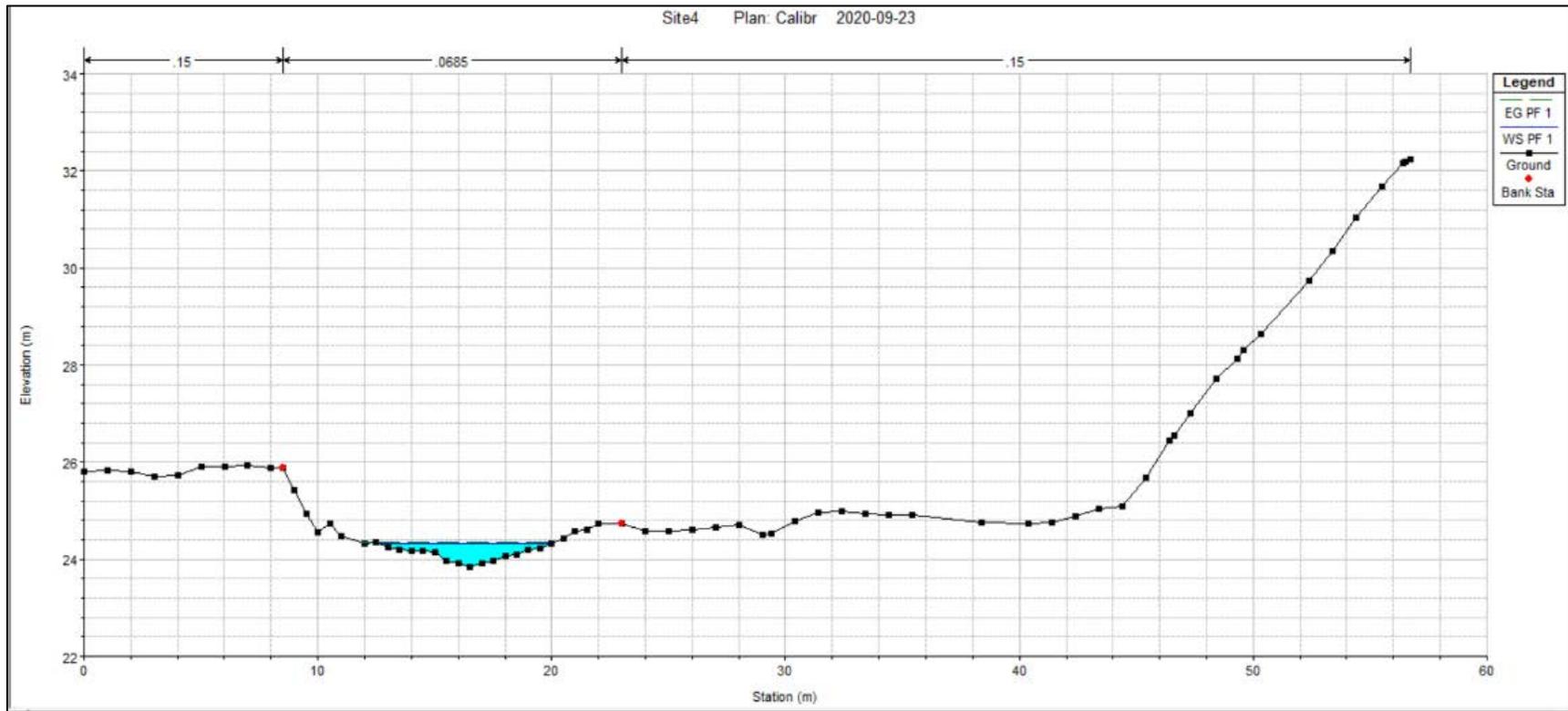


Figure 7 : Section transversale 2 (site 4)

Résultats du calage : Site 9

Les résultats du calage du modèle hydraulique au Site 9 sont présentés dans le Tableau 12 et les Figures 8 et 9.

Tableau 1 : Résultat du calage site 9

Section transversale	Point de mesure à partir de la berge gauche (m)	Profondeur d'eau observée (cm)	Profondeur d'eau simulée (cm)	Débit simulé (m <sup>3</sup> /s)	«n» de Manning						
1	13	33,5	34,0	0,34	<table border="1"> <tr> <td>Rive droite</td> <td>Fond du cours d'eau</td> <td>Rive gauche</td> </tr> <tr> <td>0,096</td> <td>0,05</td> <td>0,096</td> </tr> </table>	Rive droite	Fond du cours d'eau	Rive gauche	0,096	0,05	0,096
Rive droite	Fond du cours d'eau	Rive gauche									
0,096	0,05	0,096									
2	11	48,5	41,0	0,34	<table border="1"> <tr> <td>Rive droite</td> <td>Fond du cours d'eau</td> <td>Rive gauche</td> </tr> <tr> <td>0,096</td> <td>0,05</td> <td>0,096</td> </tr> </table>	Rive droite	Fond du cours d'eau	Rive gauche	0,096	0,05	0,096
Rive droite	Fond du cours d'eau	Rive gauche									
0,096	0,05	0,096									

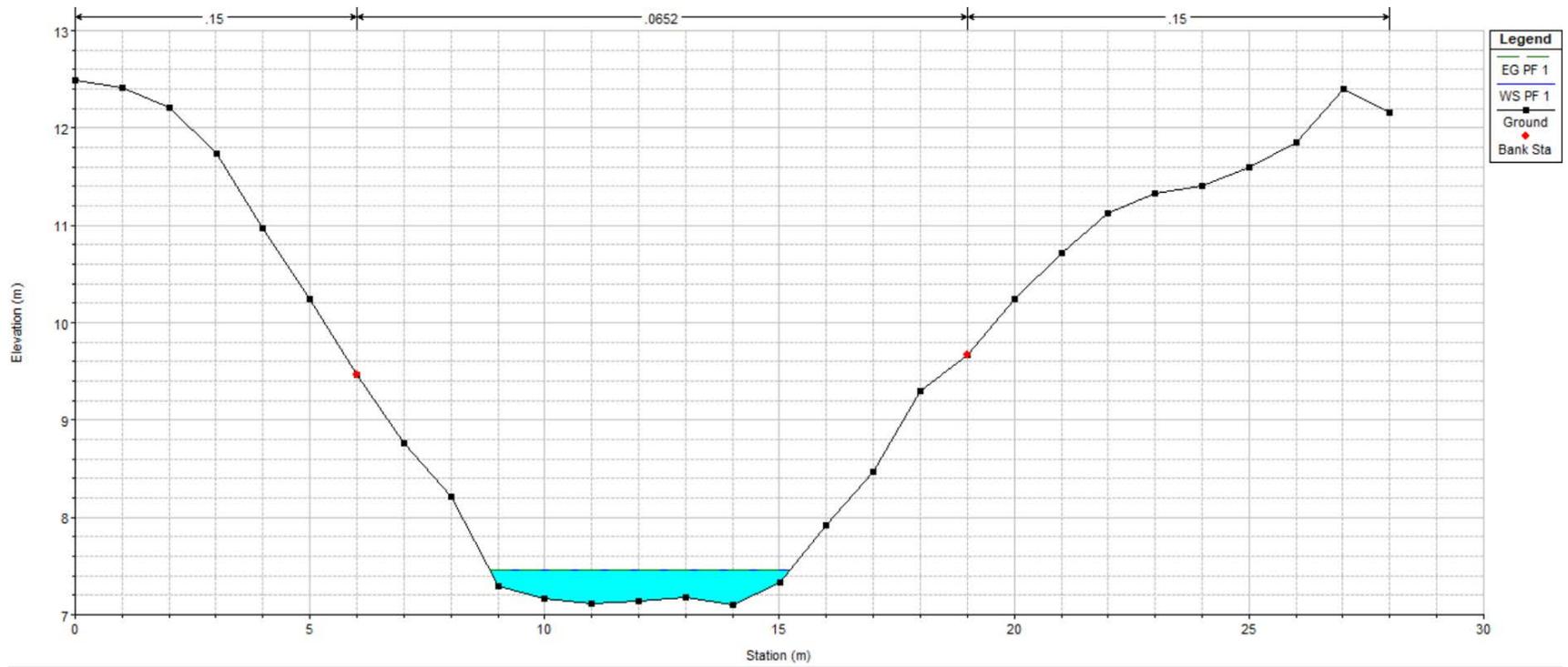


Figure 8 : Section transversale 1 (site 9)

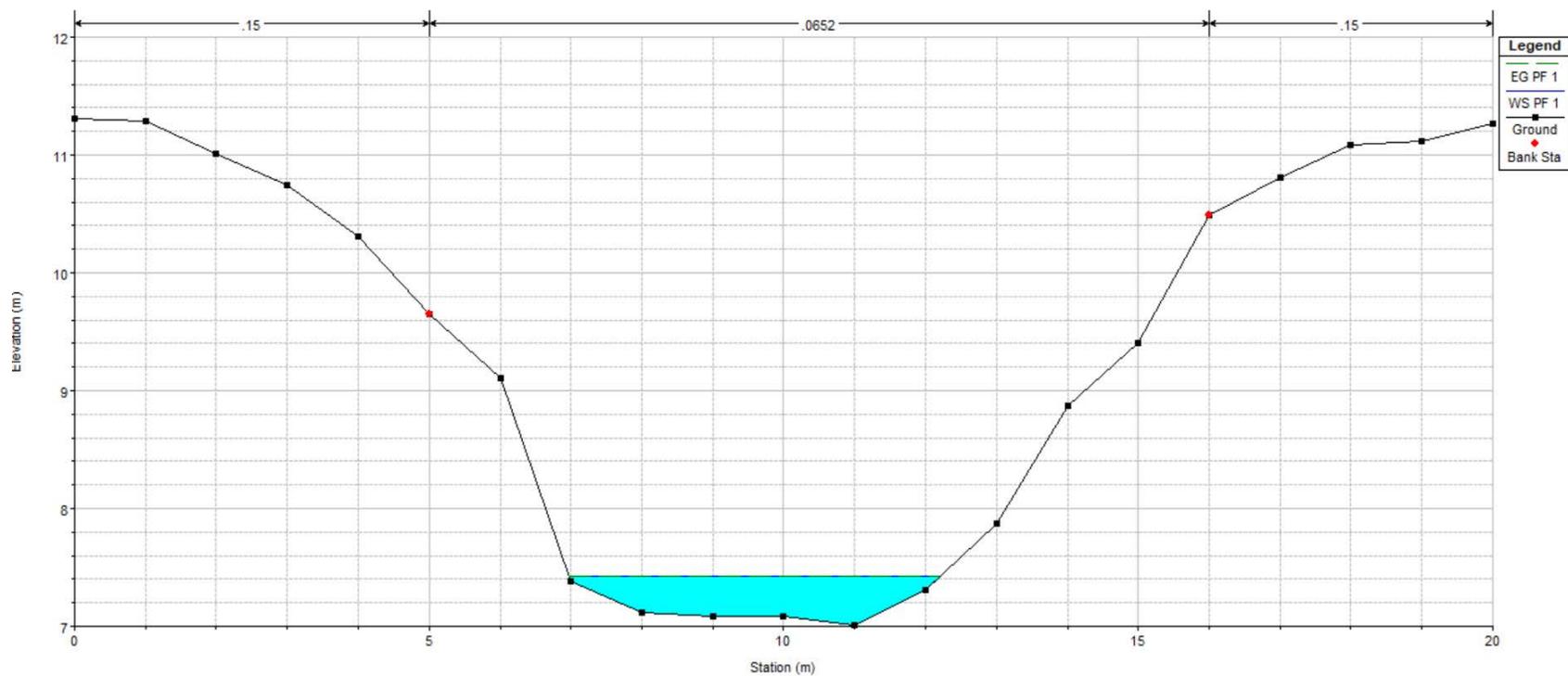


Figure 9 : Section transversale 2 (site 9)

Résultats du calage : Site 28

Le Tableau 13 présente les résultats du calage au Site 28. Les Figures 10, 11 et 12 montrent respectivement les sections transversales 1, 3 et 4 du Site 28 après la simulation du débit de calage.

Tableau 2 : Résultat du calage site 28

Section transversale	Point de mesure à partir de la berge gauche (m)	Profondeur d'eau observée (cm)	Profondeur d'eau simulée (cm)	Débit simulé (m <sup>3</sup> /s)	«n» de Manning						
1	7	44	44	0,03	<table border="1"> <tr> <td>Rive droite</td> <td>Fond du cours d'eau</td> <td>Rive gauche</td> </tr> <tr> <td>0,1</td> <td>0,03</td> <td>0,1</td> </tr> </table>	Rive droite	Fond du cours d'eau	Rive gauche	0,1	0,03	0,1
Rive droite	Fond du cours d'eau	Rive gauche									
0,1	0,03	0,1									
3	5	9	5	0,03	<table border="1"> <tr> <td>Rive droite</td> <td>Fond du cours d'eau</td> <td>Rive gauche</td> </tr> <tr> <td>0,1</td> <td>0,0616</td> <td>0,1</td> </tr> </table>	Rive droite	Fond du cours d'eau	Rive gauche	0,1	0,0616	0,1
Rive droite	Fond du cours d'eau	Rive gauche									
0,1	0,0616	0,1									
4	6	8	8	0,03	<table border="1"> <tr> <td>Rive droite</td> <td>Fond du cours d'eau</td> <td>Rive gauche</td> </tr> <tr> <td>0,1</td> <td>0,0616</td> <td>0,1</td> </tr> </table>	Rive droite	Fond du cours d'eau	Rive gauche	0,1	0,0616	0,1
Rive droite	Fond du cours d'eau	Rive gauche									
0,1	0,0616	0,1									

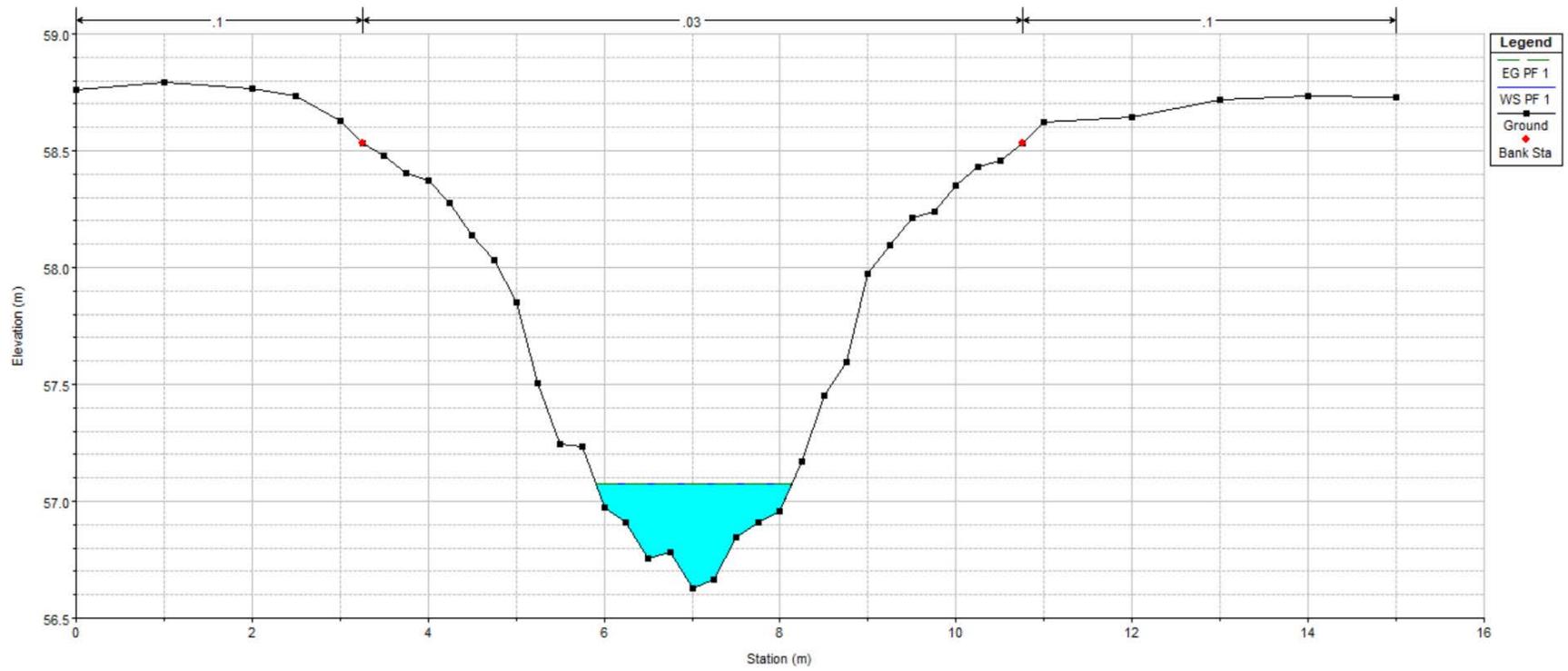


Figure 10 : Section transversale 1 (site 28)

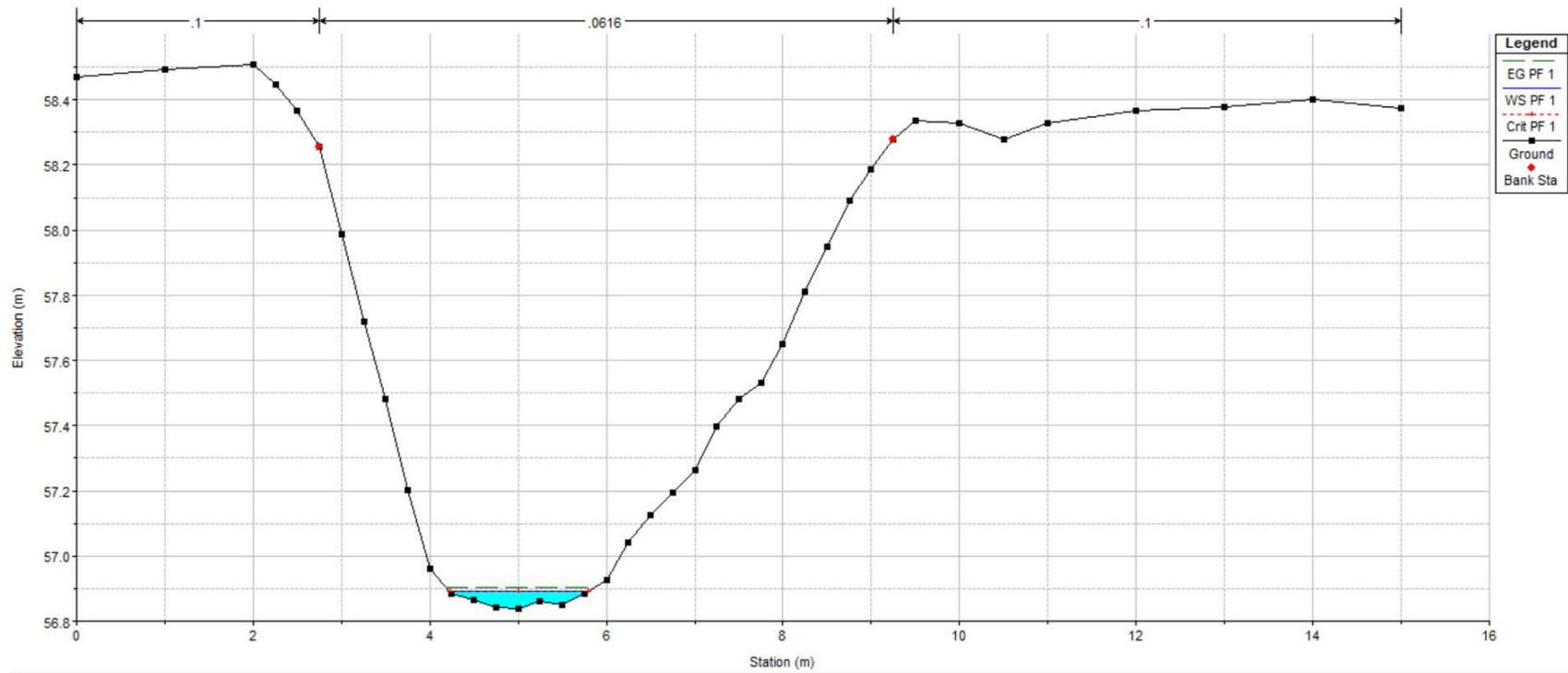


Figure 11 : Section transversale 3 (site 28)

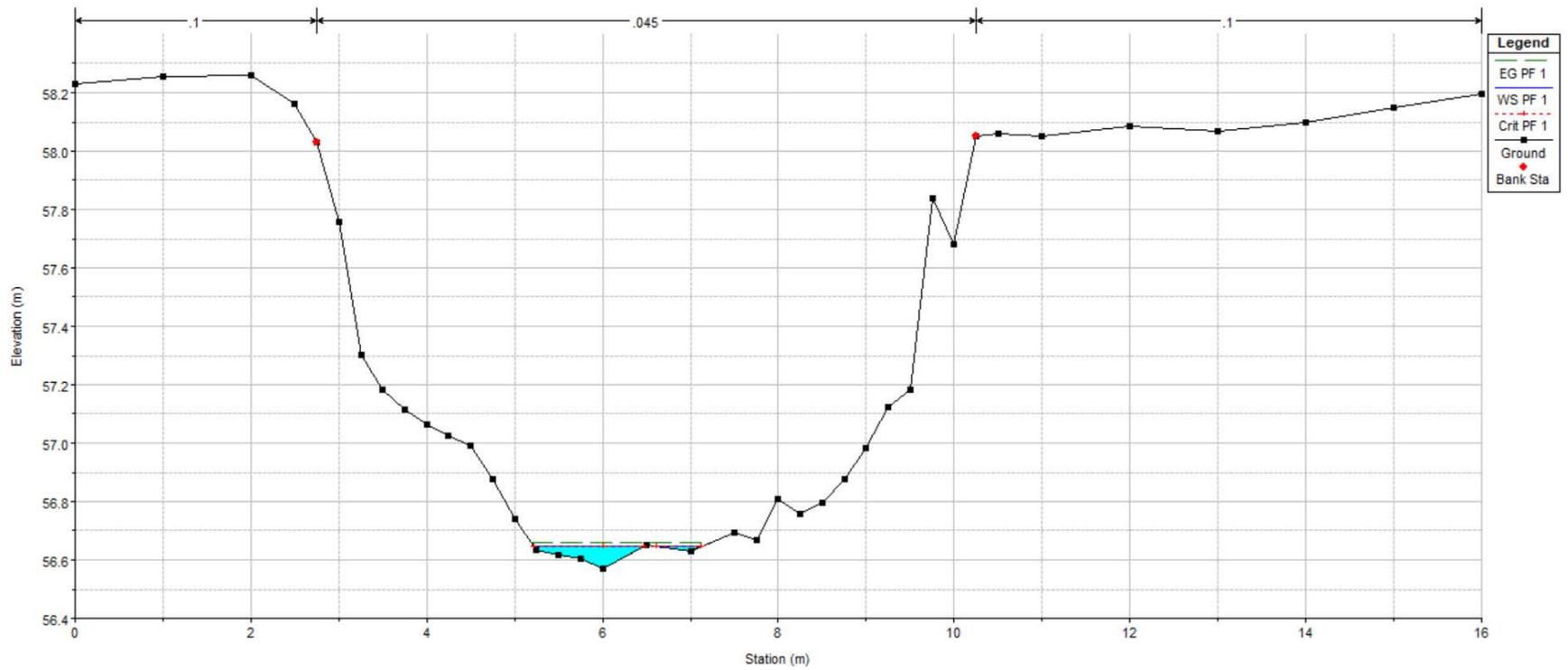
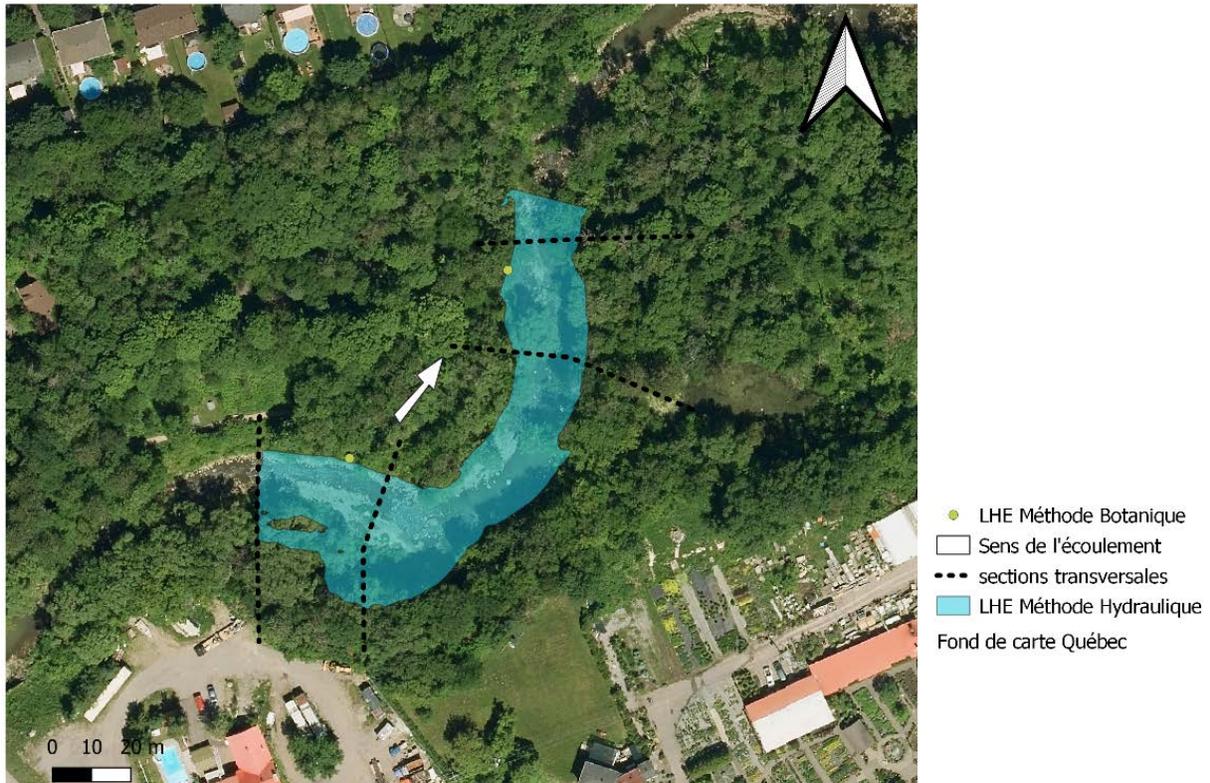


Figure 12 : Section transversale 4 (site 28)

### 3.4 Carte d'inondation

Les Figures 13, 14 et 15 montrent la ligne des hautes eaux délimitée par la modélisation hydraulique et par la méthode botanique pour chaque site étudié.

#### **Carte d'inondation Site 4 Rivière du Cap-Rouge (2 ans)**



*Figure 1 : Carte d'inondation du site 4*

### Carte d'inondation Site 9 Rivière Du Berger (2 ans)

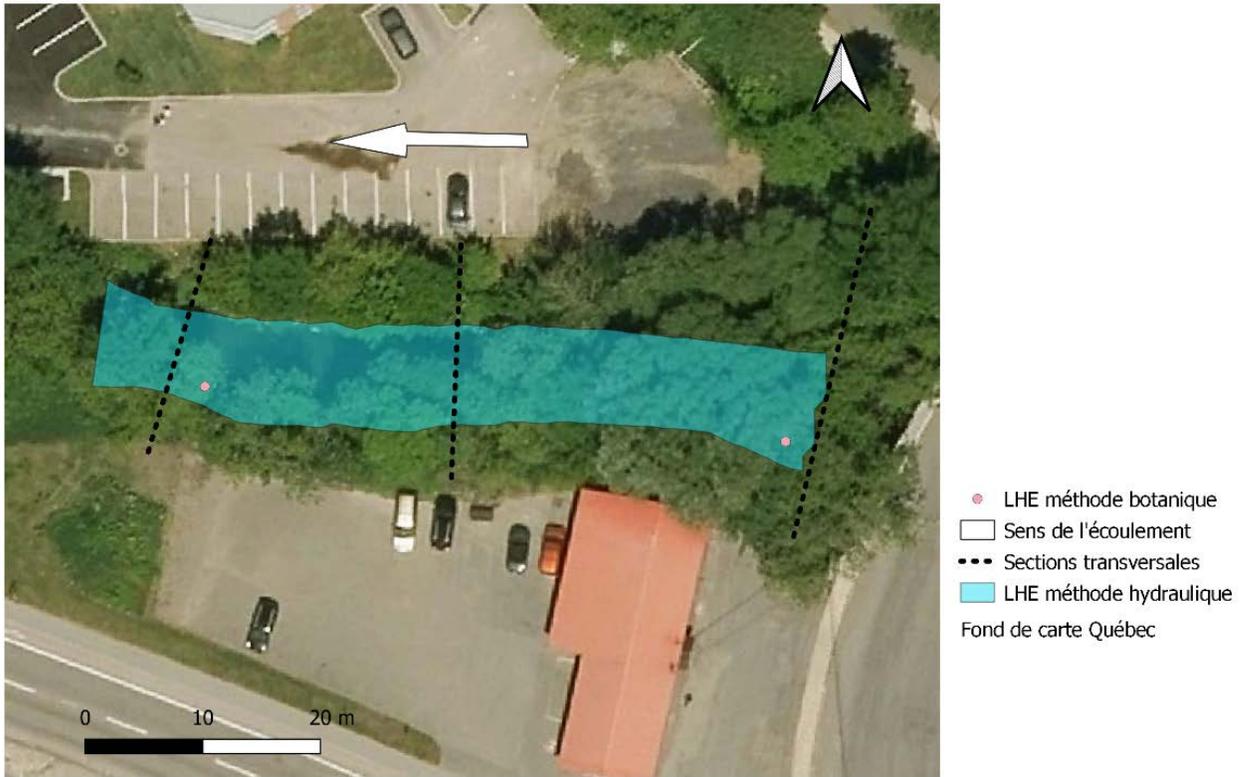


Figure 14 : Carte d'inondation du site 9

## Carte d'inondation Site 28 Rivière Des Roches (2 ans)

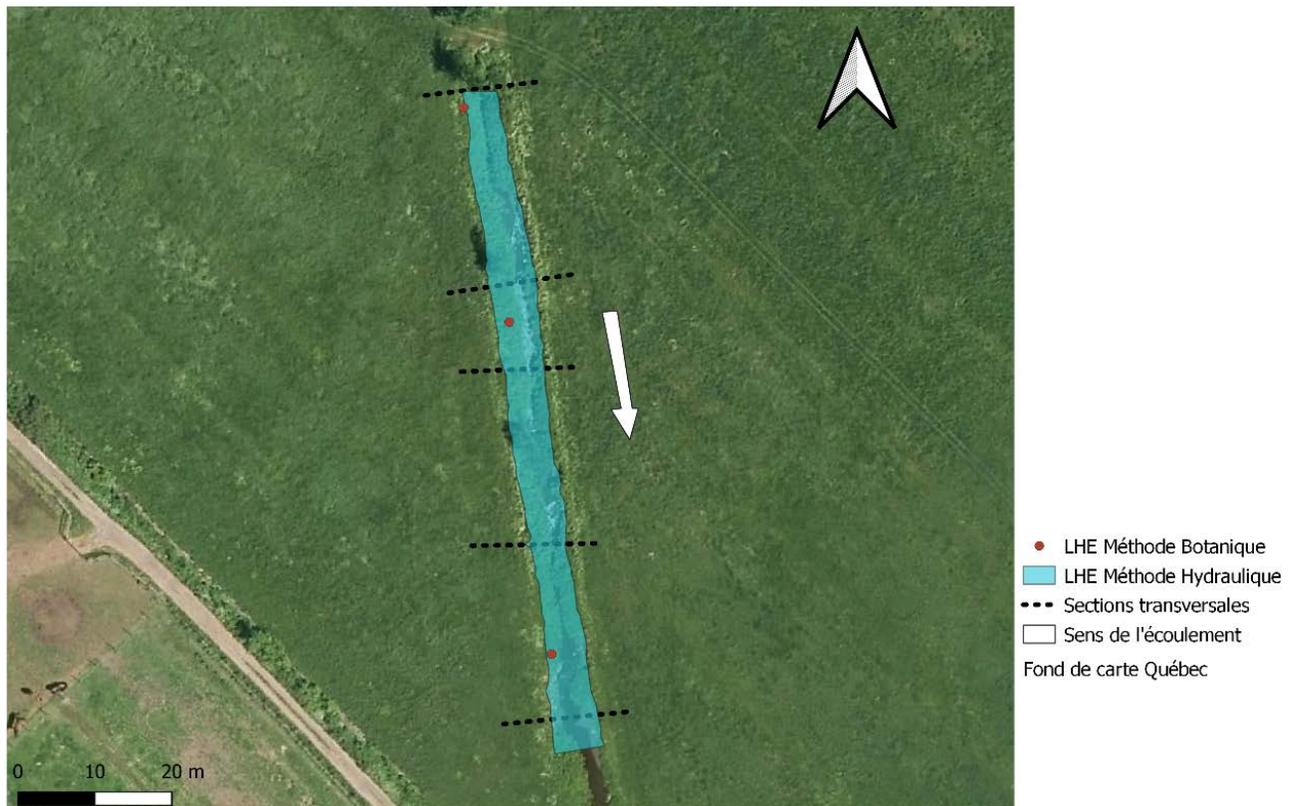


Figure 2 : Carte d'inondation site 28

### 4. Approche géomorphologique

Cette partie du projet a pour objectif d'évaluer les possibilités offertes par les données Lidar pour 1) l'identification automatisée des indicateurs géomorphologiques permettant la délimitation de la ligne des hautes eaux (LHE) et 2) la création automatisée de modèles bathymétriques servant d'intrants à l'identification de la LHE par la méthode hydraulique.

#### 4.1 Utilisation du Lidar pour la méthode géomorphologique

L'approche géomorphologique repose essentiellement sur l'identification, sur le terrain, du niveau plein bord des cours d'eau. Ce niveau correspond au niveau d'eau maximum qu'un cours d'eau peut atteindre avant de déborder sur la plaine d'inondation. Le débit au niveau plein bords est souvent appelé débit morphogène et il possède une récurrence pouvant varier entre 1 et 2 ans (Wolman et Leopold 1957).

Plusieurs études se sont penchées sur le développement d'une méthode objective de détermination du niveau plein bord d'un cours d'eau, mais principalement à partir de données mesurées sur le terrain. Nous nous sommes inspirés de ces travaux afin de développer deux approches automatisées de détermination de la LHE à partir des données d'élévations topographiques obtenues par des relevés Lidar. Les modèles numériques de terrain (MNT) produits à partir des données Lidar ont une résolution spatiale et verticale de 1m et 0,15m respectivement, ouvrant ainsi la porte à leur utilisation possible pour la détection automatisée de la LHE au moyen d'indicateurs géomorphologiques.

Nous avons débuté l'implantation et le test de deux méthodes d'identification automatisées de la LHE.

#### 4.2 Méthode par débordement

La première méthode vise à établir, pour une coupe transversale donnée, le niveau d'eau à partir duquel la rivière débordera du chenal principal. Il s'agit d'une méthode itérative où la largeur du chenal ( $W$ ) est d'abord mesurée au point le plus profond de la coupe transversale, puis à des incréments successifs de hauteurs de niveau d'eau au-dessus du point le plus profond ( $H$ ) (Figure 16). En suivant l'évolution de la largeur en fonction de l'augmentation du niveau d'eau, le niveau de débordement correspondra à celui où une très forte augmentation de largeur sera mesurée par rapport à la précédente. Le niveau plein bord correspondra donc au niveau précédent le débordement. L'exemple de la Figure 1, effectué en utilisant un incrément de 25 cm, montre que la précision de la détermination du niveau plein bord dépend de la valeur de l'incrément utilisé. La méthode étant automatisée, il semblerait souhaitable d'utiliser le plus petit incrément possible, mais la précision sera alors limitée par la résolution verticale du modèle de terrain utilisé. La précision verticale des MNT Lidar étant de 15 cm, nous avons opté pour l'utilisation de cette valeur d'incrément pour l'application de cette méthode.

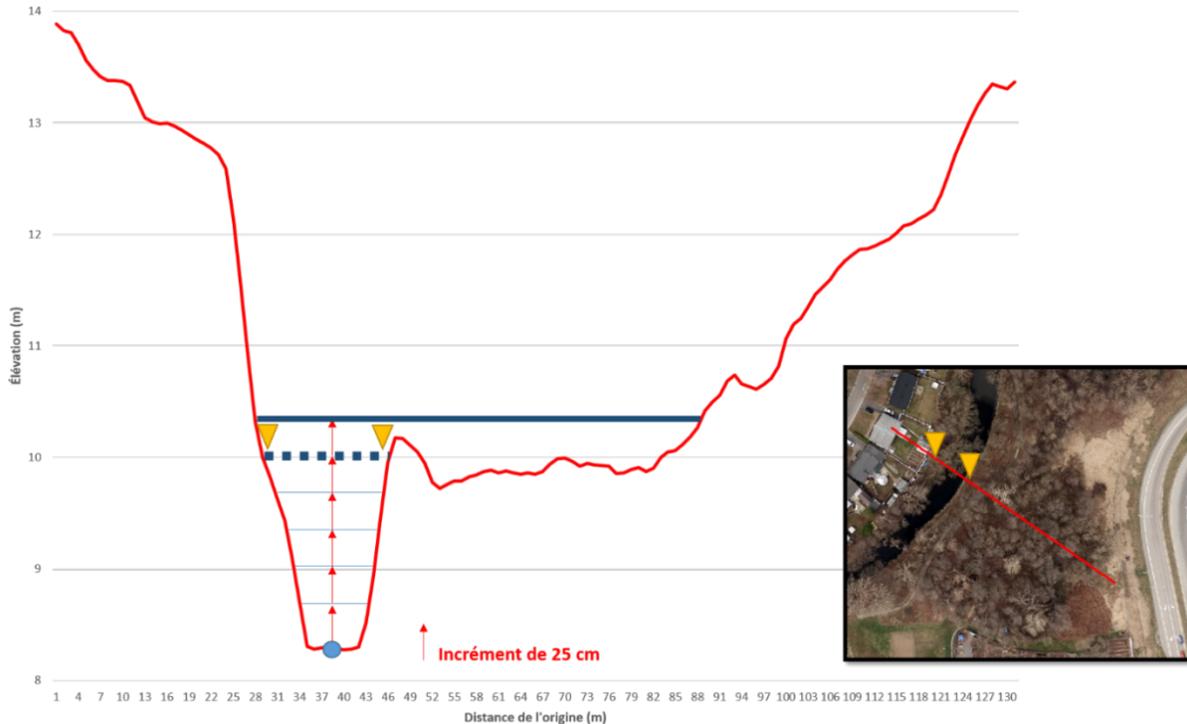


Figure 16 : Exemple d'application de la méthode de débordement à une coupe transversale en utilisant un incrément de 25 cm du niveau d'eau (Transect 30 – Site #4). La ligne bleue pleine correspond au niveau de débordement. La ligne bleue pointillée correspond au niveau plein bord identifié. Les triangles jaunes identifient la localisation de la LHE sur la coupe transversale et la photo aérienne correspondante.

L'identification objective du niveau de débordement peut être facilitée par la création d'une figure qui décrit l'évolution de  $W$  en fonction de  $H$  (Figure 17). Il arrive souvent, tel qu'illustré à la Figure 16, que plus d'un point d'inflexion soient présents. Des tests démontrent que le point d'inflexion le plus important est généralement celui qui correspond au niveau plein bord, en autant que le transect d'analyse ne se prolonge pas trop loin de part et d'autre du chenal. Dans ce cas, un point d'inflexion important pourrait correspondre à d'autres éléments du paysage, tel que la limite de la plaine d'inondation ou le mur de la vallée. Une figure illustrant le gradient de changement de  $W$  en fonction de  $H$  (Figure 18) est encore plus efficace pour identifier objectivement et facilement le niveau plein bord.

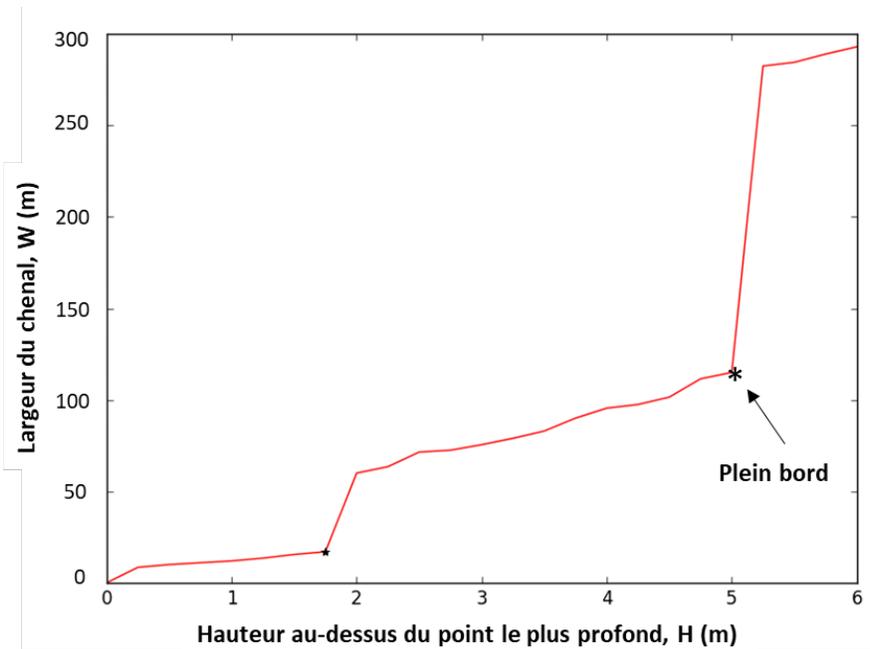


Figure 17 : Exemple de courbe d'augmentation de la largeur du chenal en fonction de la hauteur du niveau d'eau (Transect 30 – Site #4). Des deux points d'inflexions possibles identifiés par des \*, celui qui correspond à la plus grande augmentation de largeur est généralement celui du niveau pleins bords.

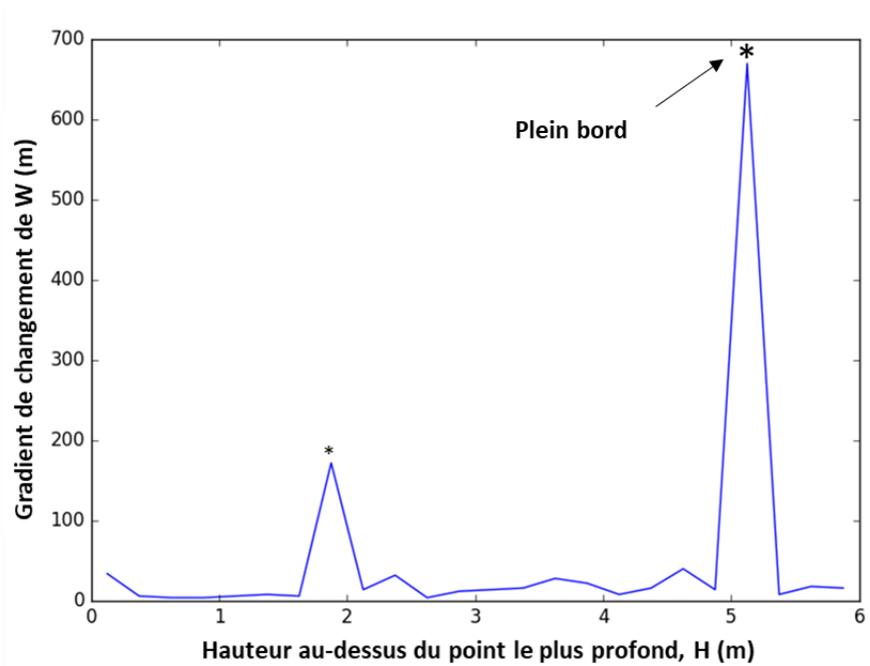


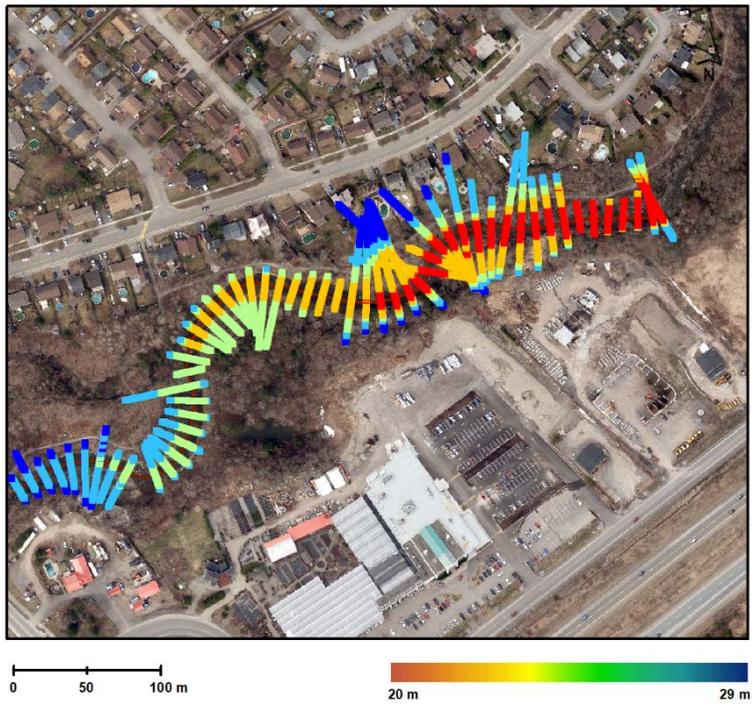
Figure 18 : Illustration de la méthode utilisant le gradient de changement de W en fonction de H pour identifier le niveau plein bord (Transect 30 - Site # 4).

#### 4.3 Analyse préliminaire sur les sites tests

Les figures 19, 20 et 21 montrent les résultats préliminaires obtenus suite à l'application de la méthode par débordement sur les sites 4, 9 et 28 respectivement. À chacun des sites, des transects espacés de 10 m ont été générés automatiquement le long de la ligne centrale du chenal et analysés en utilisant une valeur seuil de gradient de changement pour la détection du niveau plein bord. Sur ces figures, les extrémités des lignes colorées correspondent à la localisation du niveau plein bord et la palette de couleur décrit les valeurs d'élévations du MNT le long des transects.

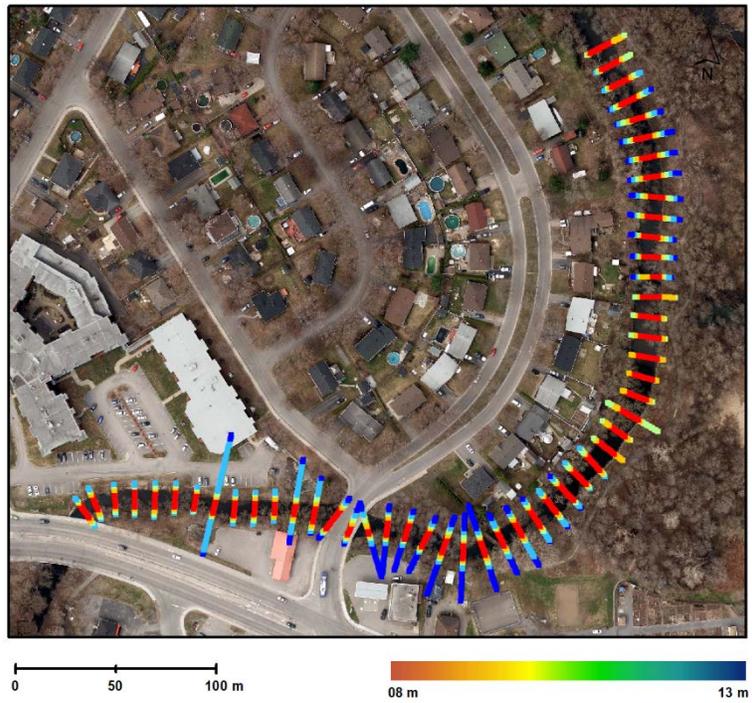
La cohérence spatiale des résultats est généralement bonne, mis à part quelques transects (plusieurs au site 4), où des valeurs singulières de limites de la LHE se démarquent largement des valeurs des transects voisins. L'incohérence de ces transects reflète fort probablement une rive fortement perturbée dont les berges anthropisées n'ont pas été façonnées par les processus géomorphologiques. Une possibilité serait de ne pas tenir compte de ces valeurs singulières dans la délimitation de la LHE. Ceci pourrait par exemple être fait en effectuant une régression entre l'élévation du sol aux extrémités du niveau plein bord et la distance le long du chenal. Ceci équivaldrait à modéliser la pente du cours d'eau au niveau plein bord. Puisque cette pente du niveau plein bord ne peut faire autrement que démontrer une grande cohérence spatiale, les valeurs singulières d'élévations seraient identifiées et retirées de l'analyse.

La comparaison avec les LHE obtenues par les méthodes botaniques et hydrauliques aux mêmes sites permettra également d'évaluer la valeur de la méthode par débordement. La méthode est programmée sous Python 2.7 et sera intégrée dans l'interface ArcGIS dans une boîte d'outils.



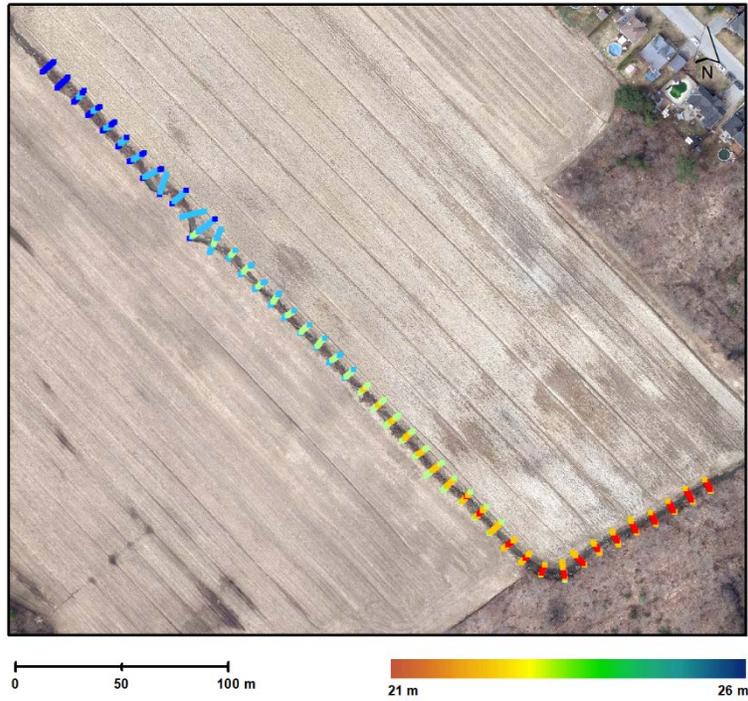
LHE - Méthode de la hausse du niveau d'eau - Site # 04

Figure 19 : Limites de la LHE estimés au site #4 avec la méthode par débordement.



LHE - Méthode de la hausse du niveau d'eau - Site # 09

Figure 20 : Limites de la LHE estimés au site #9 avec la méthode par débordement.



LHE - Méthode de la hausse du niveau d'eau - Site # 29

*Figure 21 : Limites de la LHE estimés au site #28 avec la méthode par débordement.*

### 4.3 Méthode par seuillage de pente

La deuxième méthode testée utilise une approche matricielle basée sur le calcul de la pente locale du sol aux abords du cours d'eau. Inspirée de Johansen et al. (2011), la méthode identifie le niveau plein bord du cours d'eau en recherchant les limites pour lesquelles la pente est supérieure à un seuil fixé par l'analyste.

Dans un premier temps, les données d'élévations LiDAR (Figure 22) sont utilisées pour calculer une carte des pentes locales à l'intérieur d'une zone tampon située de part et d'autre du talweg du cours d'eau (Figure 22). La pente locale maximale est calculée dans une fenêtre de convolution 3 x 3 centrée sur le pixel d'intérêt.

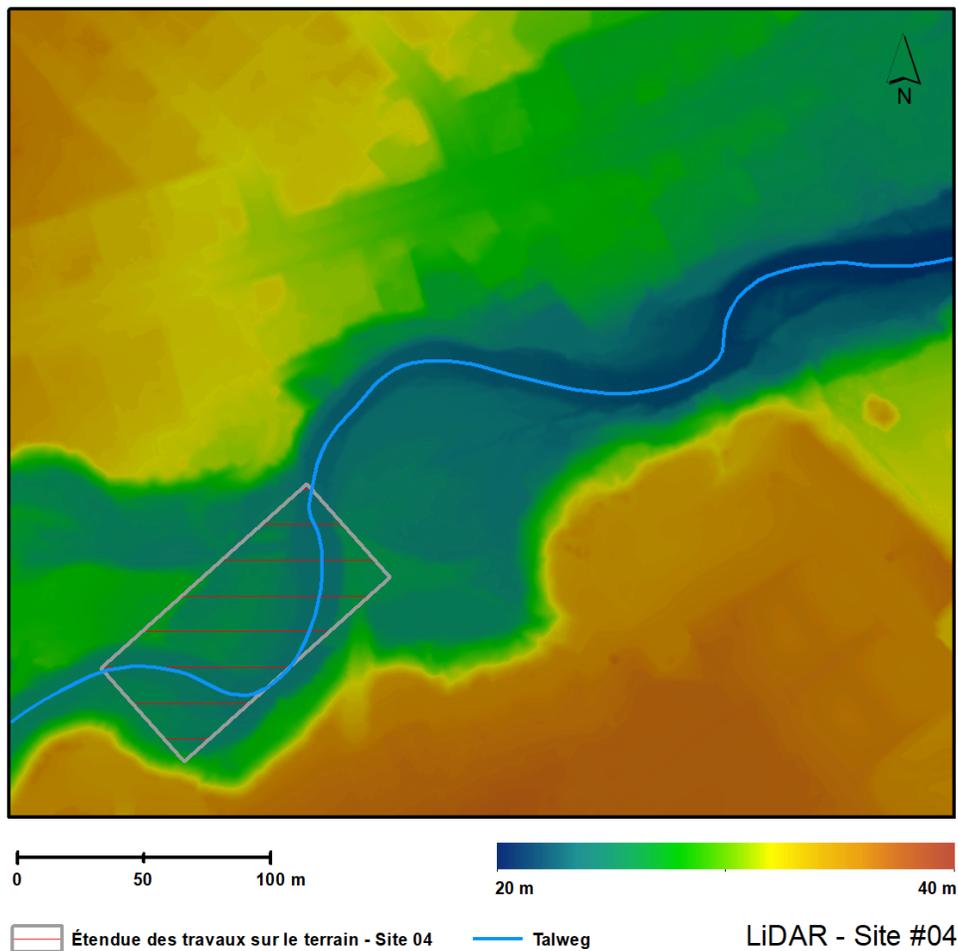


Figure 22 : Exemple d'un MNT provenant de données LiDAR pour le site # 04.

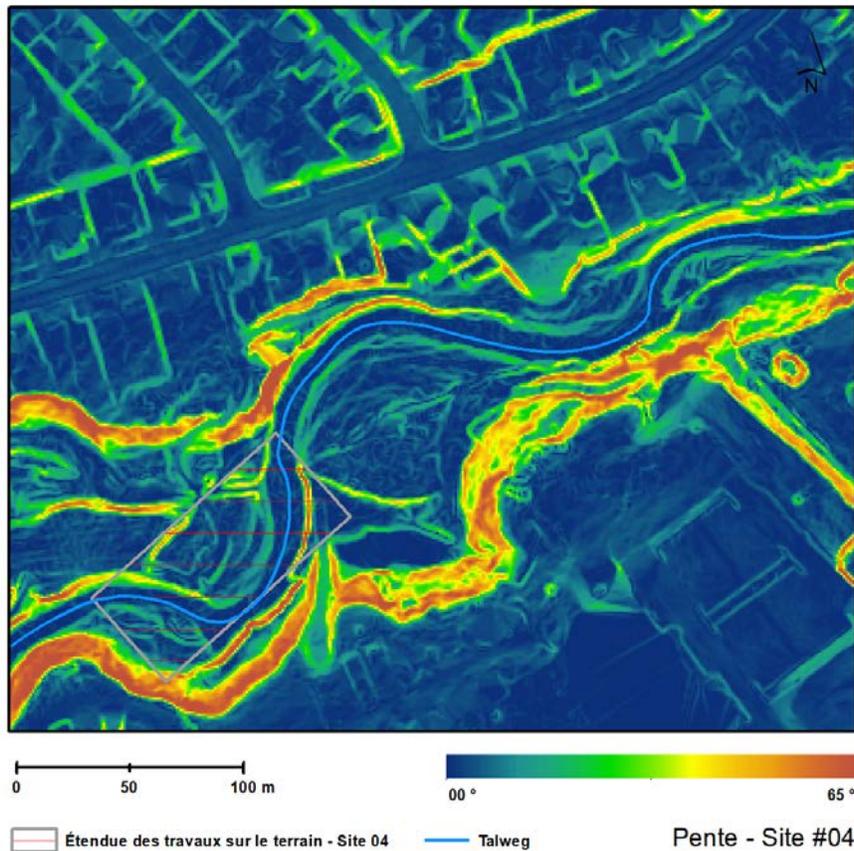


Figure 23 : Exemple d'une carte de pente calculée pour le site #04.

Deuxièmement, les limites de la LHE sont obtenues en localisant les valeurs les plus proches du chenal qui excèdent une certaine valeur seuil de pente déterminée par l'opérateur. La localisation finale de la position de la LHE est ensuite déterminée en trouvant l'endroit, sur le haut du talus, où la pente locale redescend en dessous de la valeur seuil.

#### 4.4 Résultats obtenus sur les sites tests

Les figures 24, 25 et 26 montrent des exemples d'utilisation de la méthode par seuillage de pente pour les sites tests 4, 9 et 28 respectivement. Ces analyses préliminaires ont été effectuées avec des valeurs seuils allant de 5° à 12°. Les résultats indiquent que la valeur optimale de seuil est variable selon le type de milieu analysé et que des analyses plus poussées doivent être effectuées afin d'identifier la meilleure valeur de seuil de pente à utiliser selon les circonstances. La comparaison avec les LHE obtenues par les méthodes botaniques et hydrauliques sera très utile à cette fin.

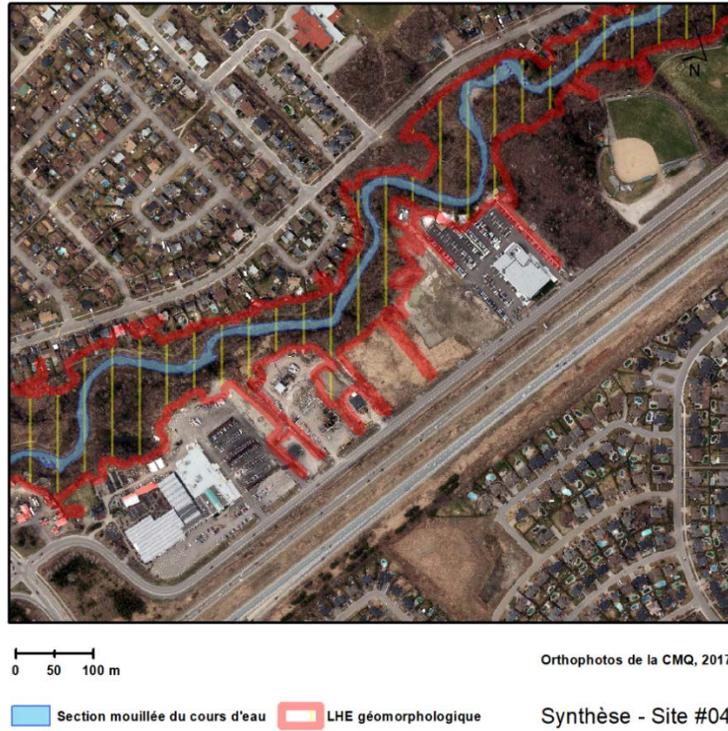


Figure 24 : Exemple de délimitation de la LHE au site #4 à l'aide de la méthode par seuillage de pente.

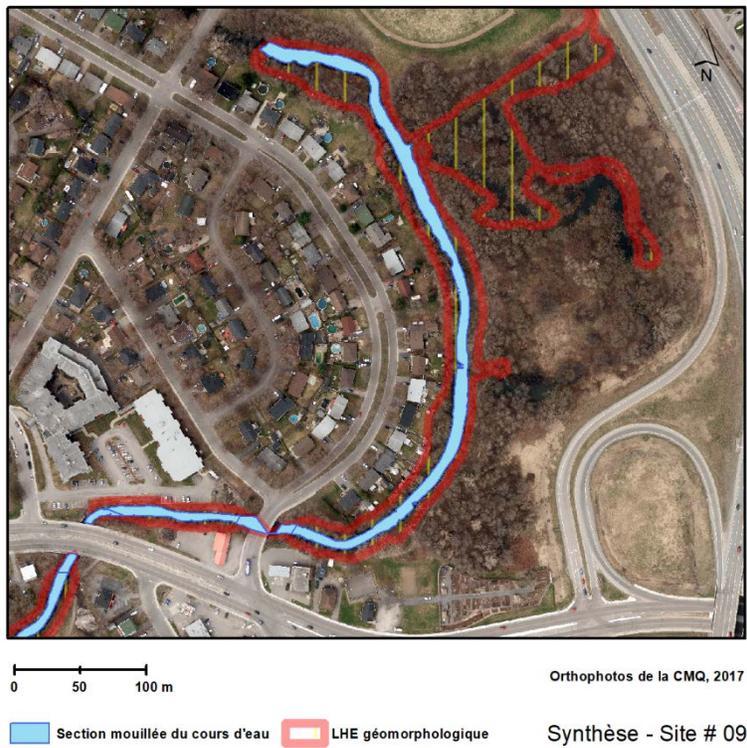


Figure 25 : Exemple de détermination de la LHE au site #9 à l'aide de la méthode par seuillage de pente.

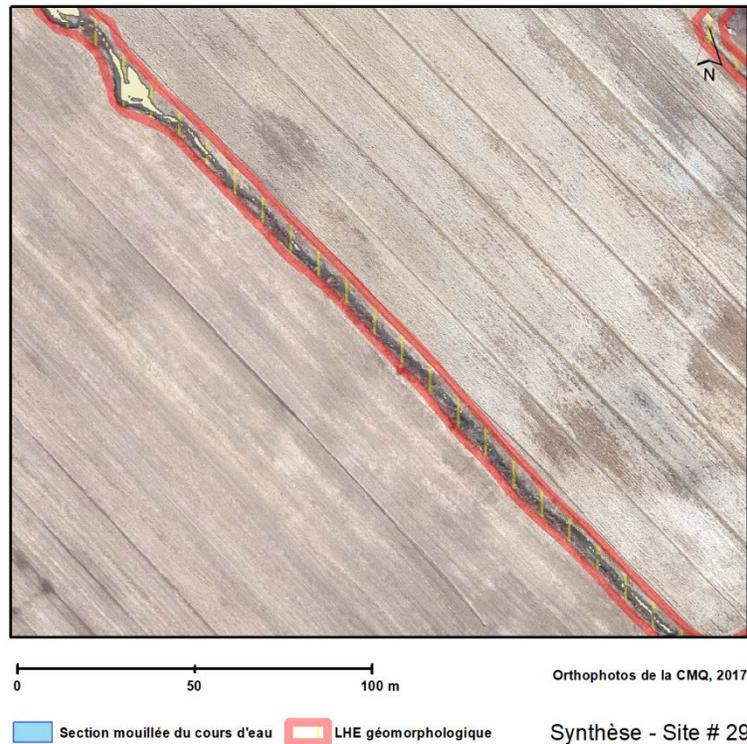


Figure 26 : Exemple de détermination de la LHE au site #28 à l'aide de la méthode par seuillage de pente.

#### 4.5 Création automatisée de modèles bathymétriques pour la méthode hydraulique

Cette partie du projet est actuellement en cours de réalisation. L'approche générale est basée sur l'idée d'utiliser les données LiDAR comme intrant à la création automatisée des modèles bathymétriques nécessaires à la méthode hydraulique d'identification de la LHE. La cueillette des données bathymétriques sur le terrain est une étape longue et coûteuse qui n'est souvent appliquée que sur de courts tronçons de cours d'eau. L'approche proposée pourrait permettre d'obtenir rapidement et simplement les données bathymétriques sur l'ensemble des cours d'eau à analyser.

Lorsque le débit le jour du survol LiDAR est connu (ou si il peut être estimé), il devient alors possible de calculer la profondeur hydraulique du cours d'eau puisque la largeur mouillée et la pente de la surface de l'eau sont mesurables à partir des données LiDAR, et que le coefficient de Manning peut être estimé. Cette estimation de la profondeur hydraulique permettrait alors d'approximer une coupe transversale de forme trapézoïdale standard pour le débit au moment du relevé LiDAR. Cette coupe transversale serait combinée aux données topographiques LiDAR de la berge et de la plaine d'inondation de

part et d'autre du chenal. Cette bathymétrie serait alors utilisée pour réaliser les simulations hydrodynamiques nécessaires à la méthode hydraulique de détermination de la LHE.

## 5. Analyses à venir

Les analyses botaniques présentées dans la section 2 seront reprises avec le jeu de données complet comprenant tous les sites, ce qui représente un total de 71 transects (3 types de berges x 8 tronçons par type x 3 transects; moins un transect pour un site qui était fortement colonisé par la renouée du japon). De plus, elles seront appliquées au jeu de données prises par points de contact. La méthode du ratio de dominance (dominance ratio; Lichvar, Gillrich et Ochs 2011) sera également testée. Cette méthode permet de ne tenir compte que des espèces dominantes dans un quadrat, ce qui pourrait constituer une économie de temps lors des inventaires.

Outre la PCA, une recherche d'espèces indicatrices des différents niveaux de récurrence des inondations sera effectuée, ainsi qu'une RDA avec les facteurs environnementaux, afin de comprendre ce qui structure les communautés au sein des trois types de paysage à l'étude.

La question des spécimens qui ne sont pas identifiables à l'espèce pour des raisons de stade de maturité sera aussi explorée, afin de comprendre leur impact sur la détermination hydrophytique des communautés végétales.

De plus, les comparaisons avec les autres méthodes seront poussées plus loin, et permettront de déceler la présence de biais qui seraient associés à certains types de paysages plus qu'à d'autres.

Il en va de même pour les deux autres approches (hydraulique et géomorphologique) qui seront appliquées sur l'ensemble des tronçons sélectionnés.

## 6. Difficultés rencontrées

Certaines difficultés ont été rencontrées lors de la phase terrain de l'expérience et ont causé des retards dans la réalisation du protocole, ce qui explique que l'équipe ait dû se limiter au nombre de 8 sites par catégorie de paysage.

Le plus grand obstacle rencontré cet été vient du système GNSS, qui présente des difficultés lorsqu'utilisé sous un couvert forestier. Cela a nécessité l'utilisation d'alternatives comme l'approche par station totale et le ZIPLEVEL®, une méthode qui est beaucoup plus longue à utiliser sur le terrain.

La méthode botanique en elle-même présente des contraintes pour sa réalisation. Une fois les tronçons de rivières sélectionnés, il arrivait que nous devions les rejeter ou nous

déplacer afin de trouver des sections bien végétalisées. Cette végétation ne devait pas être aménagée (pas de fauche, pas de plantations). Les zones colonisées par la renouée du japon (*Reynoutria japonica*) ne pouvaient être inventoriées, puisque cette espèce exclut toutes les autres par son couvert très dense. Ces validations ont demandé plusieurs journées de terrain supplémentaires.

Enfin, les contraintes imposées par la méthode hydraulique s'ajoutaient à ces dernières, c'est-à-dire le besoin d'avoir un tronçon raisonnablement rectiligne, dont la longueur ne présentait aucune arrivée d'eau (tributaire ou tuyau de drainage). Ce critère-ci était particulièrement contraignant en milieu agricole, où les fossés sont parfois assez rapprochés et les tuyaux de drainage parfois difficiles à voir.

Un certain retard est également dû à la disponibilité de l'équipement. Afin de convertir un cours donné à l'automne en cours à distance, le GNSS a été réquisitionné pour un tournage et n'a été disponible pour le projet qu'une semaine après le début des inventaires, ce qui a nécessité une deuxième visite pour plusieurs sites.

Pour le volet hydrologique, outre les difficultés associées aux mesures de débit correspondant à différents niveaux en rivière, la méthode d'estimation des crues de récurrences 2 ans par analyse régionale peut parfois poser un défi. En effet, comme mentionné précédemment, les débits estimés par modélisation pluie-débit sont plus élevés par un facteur d'environ 2,6 comparée à l'approche statistique pour la récurrence 2 ans.

Pour le volet géomorphologique, les données LiDAR communes ne pénètrent pas sous la surface de l'eau. Vis-à-vis des cours d'eau, les valeurs d'élévations correspondent donc à la surface de l'eau au moment de la réalisation des relevés. À un transect, la partie avec la plus faible élévation (souvent plus ou moins plane en raison de l'erreur verticale des données) constitue donc la surface de l'eau, et les parties adjacentes correspondent aux berges exondées du chenal. Afin de compléter la bathymétrie, il faudrait donc être en mesure de caractériser la coupe transversale le long du périmètre mouillé. Pour les plus grands cours d'eau, ceci peut être fait à partir d'une analyse quantitative des photo-aériennes. Dans les petits cours d'eau de l'étude actuelle, ceci sera le plus souvent impossible en raison de la végétation riveraine qui masque la majeure partie de la surface des cours d'eau.

## Références

- Boucher, M.-È. et J. Francoeur (2013). Détermination des cotes de crues des rivières du Berger, des Commissaires et des Sept Ponts. Québec, Centre d'expertise hydrique du Québec (CEHQ).
- Boucher, M.-È., et al. (2013). Détermination des cotes de crues de la rivière du Cap Rouge. Québec, Centre d'expertise hydrique du Québec (CEHQ).
- Burn, D. H. (1997). "Catchment similarity for regional flood frequency analysis using seasonality measures." *Journal of Hydrology* **202**(1): 212-230.
- Chow, V. T. (1959). *Open-channel hydraulics*. New York, Mcgraw-Hill.
- Dalrymple, T. (1960). Flood-frequency analyses, Manual of Hydrology: Part 3. *Water Supply Paper*.
- Hastie, T. et R. Tibshirani (1986). "Generalized Additive Models." *Statist. Sci.* **1**(3): 297-310.
- Hosking, J. R. M. et J. R. Wallis (1997). *Regional frequency analysis an approach based on L-moments*. Cambridge etc., Cambridge University Press.
- Lichvar, R.W., J. Gillrich et W. Ochs. 2011. Discrepancies in Hydrophytic Determinations Produced by Three Vegetation Formulas Used for Wetland Delineations. *Wetlands*, 31: 603-611. Doi: 10.1007/s13157-011-0166-7.
- Lichvar, R.W., M. Butterwick, N.C. Melvin et W.N. Kirchner. 2014. The National Wetland Plant List: 2014 Update of Wetland Ratings. *Phytoneuron* 2014-41: 1-42.
- Marra, G. et S. N. Wood (2011). "Practical variable selection for generalized additive models." *Computational Statistics & Data Analysis* **55**(7): 2372-2387.
- Michigan Flora Online. A. A. Reznicek, E. G. Voss, & B. S. Walters. February 2011. University of Michigan. Web. October 10, 2020. <https://michiganflora.net/>
- Ministère du Développement durable, de l'Environnement et de la Lutte contre les changements climatiques, Québec, 2008, version révisée 2015. Politique de protection des rives, du littoral et des plaines inondables. Note explicative sur la ligne des hautes eaux : la méthode botanique experte, 9 pages + annexes.
- Naiman, R. J., H. Décamps, et M. E. McClain. 2005. *Riparia: Ecology, Conservation and Management of Streamside Communities*. Elsevier/Academic Press, San Diego, 130 pp.
- USDA, NRCS. 2020. The PLANTS Database (<http://plants.usda.gov>, 10 October 2020). National Plant Data Team, Greensboro, NC 27401-4901 USA.
- Noto, L. et G. la Loggia (2009). "Use of L-Moments Approach for Regional Flood Frequency Analysis in Sicily, Italy." *Water Resources Management* **23**: 2207-2229.
- Wood, S. (2001). "mgcv: GAMs and generalized ridge regression for R." *R News* **1**.
- Wood, S. (2006). *Generalized Additive Models: An Introduction With R*.

Annexes

## Annexe 1 : Résultats de l'analyse fréquentielle locale

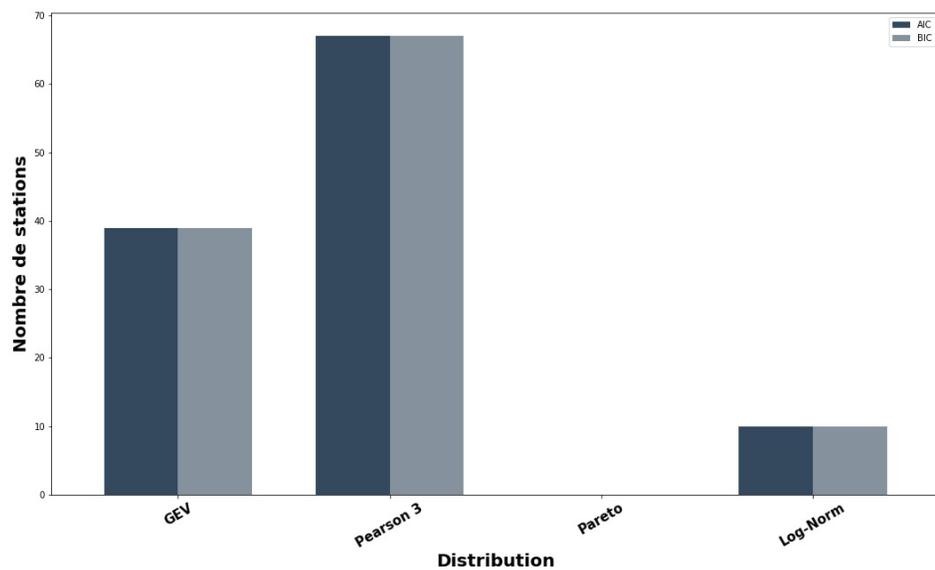


Figure A3: résultat des tests d'adéquation AIC et BIC (Analyse fréquentielle locale)

Tableau A1 : Paramètre estimé de la distribution Pearson 3 par la méthode du maximum de vraisemblance

Statio n	$\lambda$	m	$\alpha$	Statio n	$\lambda$	m	$\alpha$	Statio n	$\lambda$	m	A	Statio n	$\lambda$	m	$\alpha$
10802	2,0 6	438,8 9	237,4 0	<b>24015</b>	1,7 4	16,61	4,89	<b>40122</b>	0,1 7	53,15	11,11	<b>52212</b>	- 0,2 2	219,0 7	50,53
11003	1,6 7	417,5 2	167,7 0	<b>30101</b>	0,6 1	139,1 7	45,68	<b>40129</b>	- 0,4 2	23,55	5,42	<b>52219</b>	0,2 8	203,3 9	52,01
11204	1,3 8	300,2 9	122,9 5	<b>30103</b>	0,4 1	425,2 2	126,2 5	<b>40132</b>	- 0,9 7	7,94	1,83	<b>52228</b>	0,7 2	36,07	7,56
11508	0,6 3	58,97	14,84	<b>30106</b>	1,5 6	130,8 2	48,60	<b>40212</b>	1,4 1	13,19	5,03	<b>52233</b>	0,5 4	129,2 6	39,45
<b>11509</b>	1,2 9	407,8 8	131,5 3	<b>30220</b>	0,6 0	66,53	22,03	<b>40406</b>	0,8 9	96,06	29,51	<b>52601</b>	0,6 3	121,6 6	34,24
20404	0,8 6	148,2 2	55,51	<b>30234</b>	1,0 1	178,6 0	56,98	<b>40409</b>	1,7 4	21,25	10,74	<b>52606</b>	- 0,2 5	42,46	9,10
<b>20602</b>	1,3 8	215,3 9	84,91	<b>30262</b>	1,3 9	54,61	21,43	<b>40619</b>	0,9 8	57,48	17,76	<b>52805</b>	0,6 5	95,41	29,95
21407	1,8 4	231,6 5	84,83	<b>30278</b>	1,5 1	194,4 5	91,86	<b>40624</b>	1,0 0	334,7 5	91,85	<b>54001</b>	2,3 5	5,42	1,87

Statio n	$\lambda$	m	$\alpha$	Statio n	$\lambda$	m	$\alpha$	Statio n	$\lambda$	m	A	Statio n	$\lambda$	m	$\alpha$
21601	0,8 4	383,9 2	121,6 6	<b>30282</b>	0,7 8	225,0 5	66,86	<b>40627</b>	0,9 3	32,79	7,72	<b>60102</b>	0,9 9	92,00	25,23
<b>21702</b>	<b>1,3 7</b>	<b>53,51</b>	<b>16,90</b>	<b>30304</b>	0,3 1	352,7 2	105,3 5	<b>40840</b>	2,1 4	108,3 7	34,35	<b>60704</b>	2,1 3	85,13	33,22
21916	2,0 6	17,48	6,17	<b>30309</b>	<b>1,3 0</b>	<b>19,51</b>	<b>9,65</b>	<b>40841</b>	0,5 0	131,5 7	35,61	<b>60901</b>	- 2,1 7	53,29	14,60
22003	0,9 4	270,6 7	86,15	<b>30314</b>	1,1 8	74,58	28,56	<b>41902</b>	0,5 6	204,8 5	63,45	<b>61004</b>	0,9 5	278,8 8	107,3 2
22301	0,8 9	217,7 4	76,71	<b>30316</b>	0,5 4	94,85	30,04	<b>43012</b>	1,0 2	267,0 0	69,55	<b>61020</b>	1,0 7	185,9 3	65,84
<b>22505</b>	<b>0,9 9</b>	<b>47,16</b>	<b>16,57</b>	<b>30340</b>	1,7 1	49,53	21,85	<b>43030</b>	0,1 7	116,2 1	23,10	<b>61022</b>	0,7 1	93,60	28,19
<b>22507</b>	<b>1,3 4</b>	<b>100,1 9</b>	<b>35,22</b>	<b>30343</b>	<b>1,0 6</b>	<b>5,89</b>	<b>2,16</b>	<b>46709</b>	2,1 8	45,64	10,55	<b>61024</b>	0,7 6	64,50	19,88
22513	0,3 8	169,0 1	55,68	<b>30348</b>	0,9 8	37,07	14,51	<b>50119</b>	<b>0,6 8</b>	<b>144,3 1</b>	<b>42,10</b>	<b>61028</b>	- 0,3 2	408,0 0	109,6 9
22601	2,2 3	39,24	11,54	<b>30415</b>	1,2 4	97,76	34,81	<b>50135</b>	1,2 7	234,6 6	87,79	<b>61307</b>	0,1 0	46,76	11,99
22704	0,9	213,1	76,56	<b>30421</b>	1,7	81,11	43,57	<b>50144</b>	0,8	236,5	77,44	<b>61801</b>	1,1	120,5	35,52

Statio n	$\lambda$	m	$\alpha$	Statio n	$\lambda$	m	$\alpha$	Statio n	$\lambda$	m	A	Statio n	$\lambda$	m	$\alpha$
	3	6			8				0	8			6	8	
23002	1,8 2	22,35	9,62	30423	1,4 1	22,29	8,98	50408	1,3 4	448,2 2	136,6 9	61909	0,9 7	74,29	28,81
23004	0,5 6	74,05	21,58	30424	0,6 4	115,2 0	29,36	50409	1,1 5	165,5 0	47,84	62002	0,9 3	4,38	1,43
23106	0,4 3	266,4 7	80,91	30425	1,0 4	17,44	5,11	50501	- 0,2 6	4,49	0,93	62209	0,1 5	492,7 1	161,8 7
23303	0,6 7	278,9 9	73,21	30426	1,3 3	8,34	3,35	50702	0,0 7	69,32	16,91	62701	1,5 2	168,9 9	49,92
23401	0,4 6	187,1 5	57,71	30429	0,6 3	17,38	5,79	50807	0,6 0	48,91	9,31	62803	2,0 9	251,4 7	91,47
23422	1,1 2	195,8 5	67,36	30905	1,1 2	434,3 2	154,5 1	50812	1,1 0	21,04	6,35	70204	0,8 2	122,7 8	39,45
23427	0,7 4	117,0 7	30,95	30907	1,8 1	125,2 7	66,07	50904	1,1 1	65,63	14,04	71203	0,8 8	79,70	20,01
23432	0,7 4	60,92	25,29	30919	- 2,1 9	145,2 6	17,01	50915	- 0,4 8	13,33	3,41	71401	0,9 0	313,9 3	114,3 4
23702	2,0 5	121,1 2	38,73	30920	1,6 5	9,45	4,02	50916	2,5 8	29,66	12,24	80101	0,5 3	184,7 4	42,42

Statio n	$\lambda$	<b>m</b>	$\alpha$	Statio n	$\lambda$	<b>m</b>	$\alpha$	Statio n	$\lambda$	<b>m</b>	A	Statio n	$\lambda$	<b>m</b>	$\alpha$
24003	0,7 5	224,5 0	66,58	<b>30921</b>	0,6 9	40,35	13,49	<b>51001</b>	0,8 2	351,0 5	88,70	<b>80106</b>	- 2,5 1	50,35	14,77
<b>24014</b>	<b>0,7 7</b>	<b>489,7 4</b>	<b>141,5 5</b>	<b>40110</b>	0,4 7	192,1 9	52,51	<b>51005</b>	1,0 4	91,87	29,30	<b>120201</b>	1,2 2	10,10	3,38

$\lambda$ : paramètre de forme; **m** : paramètre de position;  $\alpha$  : paramètre d'échelle; Note : Les stations rejetant l'hypothèse de stationnarité sont surlignées en rouge, les stations rejetant l'hypothèse d'indépendance sont surlignées en bleu et les stations rejetant l'hypothèse de stationnarité et d'indépendance sont en vert

Tableau A2 : Quantiles estimés par l'analyse fréquentielle locale (m<sup>3</sup>/s)

Station	IC 2,5% (m <sup>3</sup> /s)	Q2ans (m <sup>3</sup> /s)	IC 97,5% (m <sup>3</sup> /s)	Station	IC 2,5% (m <sup>3</sup> /s)	Q2ans (m <sup>3</sup> /s)	IC 97,5% (m <sup>3</sup> /s)
10802	301,3	364,3	425,1	<b>40122</b>	49,3	52,8	56,2
11003	317,7	373,2	417,4	<b>40129</b>	21,9	23,9	26,0
11204	206,3	273,0	333,8	<b>40132</b>	7,4	8,2	9,1
11508	52,1	57,4	62,6	<b>40212</b>	10,6	12,0	13,3
11509	313,3	380,4	439,7	<b>40406</b>	82,9	91,8	100,3
20404	121,0	140,3	160,3	<b>40409</b>	15,0	18,3	21,2
20602	170,9	196,5	219,0	<b>40619</b>	48,3	54,6	59,8
21407	183,0	207,3	227,5	<b>40624</b>	287,9	319,7	348,2
21601	341,3	367,1	392,0	<b>40627</b>	28,1	31,6	34,5
21702	45,0	49,8	53,8	<b>40840</b>	85,0	97,3	110,0
21916	13,2	15,5	18,4	<b>40841</b>	100,1	128,6	152,4
22003	233,8	257,3	280,3	<b>41902</b>	181,0	199,0	215,8
22301	190,8	206,6	221,5	<b>43012</b>	235,6	255,4	275,1
22505	40,6	44,5	47,9	<b>43030</b>	96,1	115,5	134,4
22507	81,3	92,6	103,2	<b>46709</b>	38,3	42,2	47,4
22513	150,2	165,5	181,7	<b>50119</b>	130,3	139,6	148,2
22601	31,2	35,4	41,7	<b>50135</b>	189,7	216,6	240,8
22704	173,2	201,4	228,1	<b>50144</b>	198,9	226,4	253,4
23002	15,9	19,6	22,8	<b>50408</b>	380,8	418,7	452,9
23004	60,2	72,0	83,0	<b>50409</b>	143,0	156,5	168,6
23106	237,1	260,6	286,8	<b>50501</b>	3,7	4,5	5,3
23303	242,7	270,9	295,9	<b>50702</b>	58,2	69,1	80,6

Station	IC 2,5% (m <sup>3</sup> /s)	Q2ans (m <sup>3</sup> /s)	IC 97,5% (m <sup>3</sup> /s)	Station	IC 2,5% (m <sup>3</sup> /s)	Q2ans (m <sup>3</sup> /s)	IC 97,5% (m <sup>3</sup> /s)
23401	170,1	182,7	195,0	<b>50807</b>	44,9	48,0	51,0
23422	163,7	183,6	201,1	<b>50812</b>	17,2	19,9	22,3
23427	102,5	113,3	122,6	<b>50904</b>	58,8	63,1	67,0
23432	45,5	57,8	69,0	<b>50915</b>	10,7	13,6	16,3
23702	93,2	109,0	125,3	<b>50916</b>	22,0	25,2	31,2
24003	197,0	216,2	236,4	<b>51001</b>	317,4	339,1	361,4
24014	393,9	471,8	547,7	<b>51005</b>	76,0	86,9	96,0
24015	13,0	15,3	17,6	<b>52212</b>	206,9	220,9	237,3
30101	125,3	134,6	144,1	<b>52219</b>	184,1	201,0	217,8
30103	380,6	416,5	453,3	<b>52228</b>	32,6	35,2	38,1
30106	103,7	118,7	130,5	<b>52233</b>	111,0	125,7	138,4
30220	58,4	64,4	70,5	<b>52601</b>	109,7	118,1	125,1
30234	155,8	169,2	183,4	<b>52606</b>	38,2	42,8	47,4
30262	40,0	49,8	57,3	<b>52805</b>	83,8	92,2	100,9
30278	143,5	172,3	198,5	<b>54001</b>	4,2	4,8	5,5
30282	193,2	216,4	236,5	<b>60102</b>	73,5	87,9	99,3
30304	317,1	347,4	379,4	<b>60704</b>	61,7	74,4	91,7
30309	14,5	17,5	19,9	<b>60901</b>	46,4	58,1	64,2
30314	60,2	69,1	76,7	<b>61004</b>	241,9	262,1	283,1
30316	82,9	92,2	101,8	<b>61020</b>	153,8	174,4	194,4
30340	36,6	43,6	49,9	<b>61022</b>	81,2	90,3	99,3
30343	4,6	5,5	6,2	<b>61024</b>	52,1	62,0	70,9
30348	26,3	34,7	41,7	<b>61028</b>	326,8	413,8	498,5
30415	80,4	90,7	99,4	<b>61307</b>	40,2	46,6	53,0
30421	55,5	69,0	80,3	<b>61801</b>	102,7	113,9	123,7
30423	15,6	20,3	24,3	<b>61909</b>	58,1	69,7	80,9

Station	IC 2,5% (m <sup>3</sup> /s)	Q2ans (m <sup>3</sup> /s)	IC 97,5% (m <sup>3</sup> /s)	Station	IC 2,5% (m <sup>3</sup> /s)	Q2ans (m <sup>3</sup> /s)	IC 97,5% (m <sup>3</sup> /s)
30424	92,7	112,1	129,0	<b>62002</b>	3,1	4,2	5,5
30425	13,6	16,6	19,1	<b>62209</b>	427,1	488,7	555,3
30426	5,8	7,6	9,2	<b>62701</b>	140,6	156,8	169,9
30429	12,2	16,8	21,7	<b>62803</b>	192,0	222,5	251,5
30905	360,0	406,1	446,1	<b>70204</b>	99,4	117,5	133,2
30907	87,5	106,5	122,0	<b>71203</b>	60,6	76,8	93,7
30919	143,9	150,9	156,7	<b>71401</b>	260,2	297,0	333,7
30920	6,3	8,4	10,5	<b>80101</b>	170,8	181,0	190,3
30921	29,5	38,8	49,3	<b>80106</b>	45,0	55,7	60,7
40110	176,6	188,1	200,3	<b>120201</b>	8,5	9,4	10,4

Notes : IC 2,5% et IC 97,5% sont respectivement les bornes inférieures et supérieures de l'intervalle de confiance à 95% estimé par « bootstrap ». Q2ans est le débit de récurrence 2 ans

Annexe 2 : résultats de l'analyse fréquentielle régionale

Tableau A3 : Données de l'analyse fréquentielle Régionale

Station	ABV (m <sup>2</sup> )	LGA (log m <sup>2</sup> )	AGR (%)	AQT (%)	PMN (degrés)	PPM (mm)	TMN (degrés jours)	$\tau$	$\tau_3$	$\tau_4$	Q <sub>obs</sub> (m <sup>3</sup> /s)	Q <sub>sim</sub> (m <sup>3</sup> /s)	Région
10802	1911	7,56	0,00	0,41	12,07	663,14	2196,64	0,24	0,21	0,07	448,10	486,22	3
11003	1687	7,43	0,00	0,72	9,69	690,54	1922,96	0,21	0,28	0,14	417,52	388,31	3
11508	555	6,32	14,10	8,31	6,06	644,32	2192,35	0,14	0,11	0,15	58,97	69,01	2
20404	670	6,51	0,01	0,98	9,43	654,12	1941,43	0,21	0,19	0,18	148,22	174,34	3
21407	772	6,65	0,00	1,31	14,01	660,43	1981,00	0,17	0,20	0,09	231,65	201,13	3
21601	1650	7,41	3,09	1,56	8,35	647,38	2146,55	0,18	0,13	0,14	383,92	418,28	3
21916	98,6	4,59	18,14	4,70	6,68	647,69	2244,94	0,16	0,27	0,25	17,59	16,22	1
22003	1595	7,37	3,68	5,09	6,42	651,27	2296,26	0,18	0,21	0,18	270,67	183,04	4
22301	930	6,84	19,07	2,15	5,93	662,94	2301,37	0,20	0,13	0,12	217,74	260,67	3
22513	1042	6,95	8,67	3,41	5,10	649,36	2321,75	0,19	0,05	0,11	169,01	124,07	4
22704	787	6,67	2,08	2,15	4,10	679,43	2372,56	0,20	0,14	0,12	213,16	229,29	3
23106	821	6,71	7,45	1,35	5,63	748,24	2379,33	0,17	0,07	0,21	266,47	236,11	3
23303	1152	7,05	19,54	1,16	5,04	746,73	2403,56	0,15	0,06	0,12	278,99	316,05	3

Station	ABV (m <sup>2</sup> )	LGA (log m <sup>2</sup> )	AGR (%)	AQT (%)	PMN (degrés)	PPM (mm)	TMN (degrés jours)	$\tau$	$\tau_3$	$\tau_4$	Q <sub>obs</sub> (m <sup>3</sup> /s)	Q <sub>sim</sub> (m <sup>3</sup> /s)	Région
23401	708	6,56	37,32	0,34	3,67	719,48	2695,92	0,17	0,06	0,14	187,15	195,11	3
23422	696	6,55	8,35	0,91	3,26	710,12	2502,00	0,19	0,21	0,21	195,85	200,55	3
23427	781	6,66	2,98	5,79	5,84	757,78	2467,83	0,15	0,08	0,04	117,07	100,50	4
24003	918	6,82	22,93	2,11	5,77	753,04	2607,59	0,17	0,10	0,10	224,50	245,28	3
24015	48,1	3,87	45,99	0,04	5,74	734,94	2657,01	0,15	0,29	0,25	16,61	15,11	1
30101	550	6,31	27,34	1,03	4,73	738,78	2670,02	0,18	0,14	0,22	139,17	152,38	3
30103	1550	7,35	41,70	0,98	4,96	710,15	2736,73	0,17	0,08	0,12	425,22	386,73	3
30106	344	5,84	34,22	0,77	6,21	730,16	2676,83	0,19	0,19	0,07	130,82	103,27	3
30220	610	6,41	34,24	4,83	5,37	727,01	2771,83	0,19	0,12	0,15	66,53	105,48	4
30234	646	6,47	18,53	0,75	4,73	738,26	2679,80	0,18	0,18	0,17	178,60	171,66	3
30262	262	5,57	42,70	0,15	3,10	670,99	2938,66	0,21	0,25	0,18	54,61	67,28	3
30282	734	6,60	6,36	0,82	5,13	761,72	2566,88	0,17	0,13	0,10	225,05	205,88	3
30304	1505	7,32	44,49	0,66	4,00	687,08	2934,28	0,17	0,05	0,13	352,72	317,42	3
30314	211	5,35	16,87	1,18	6,96	746,83	2794,57	0,21	0,18	0,06	74,58	58,62	3
30316	323	5,78	70,36	0,16	2,73	646,05	3009,10	0,18	0,08	0,14	94,85	99,31	3
30340	235	5,46	20,93	4,13	5,18	722,86	2899,72	0,21	0,14	0,01	49,53	52,21	1

Station	ABV (m <sup>2</sup> )	LGA (log m <sup>2</sup> )	AGR (%)	AQT (%)	PMN (degrés)	PPM (mm)	TMN (degrés jours)	$\tau$	$\tau_3$	$\tau_4$	Q <sub>obs</sub> (m <sup>3</sup> /s)	Q <sub>sim</sub> (m <sup>3</sup> /s)	Région
30348	102	4,62	32,96	0,22	4,37	687,37	2954,36	0,22	0,26	0,38	37,07	31,40	1
30415	308	5,73	74,60	0,33	3,52	644,32	3097,77	0,19	0,24	0,19	97,76	90,28	3
30423	94,1	4,54	65,00	0,21	2,75	682,03	3087,13	0,21	0,17	0,12	22,29	27,07	1
30424	584	6,37	45,18	0,64	3,47	700,73	3035,53	0,15	0,11	0,26	115,20	123,93	3
30425	72,6	4,28	8,29	0,00	1,02	672,87	3121,35	0,16	0,14	0,23	17,44	18,39	1
30426	26,3	3,27	89,85	0,03	2,44	662,48	3115,44	0,22	0,18	0,14	8,34	8,86	1
30429	35,1	3,56	78,65	0,13	3,93	640,59	3106,12	0,20	0,11	0,17	17,38	13,39	1
30905	2492	7,82	51,88	0,88	2,33	612,78	3135,46	0,19	0,22	0,24	434,32	417,11	3
30921	246	5,51	20,37	0,57	1,92	618,11	3120,19	0,20	0,07	0,07	40,35	46,92	3
40110	1163	7,06	0,79	7,07	9,35	688,25	2589,45	0,15	0,07	0,12	192,19	187,25	2
40122	309	5,73	0,62	9,15	9,75	691,24	2500,88	0,12	0,03	0,15	53,15	52,66	2
40129	106	4,66	0,75	3,64	8,53	693,08	2469,98	0,13	- 0,09	0,13	23,55	27,87	1
40132	40,3	3,70	0,56	8,16	8,41	691,55	2473,16	0,13	- 0,19	0,07	7,94	8,61	1
40212	40	3,69	0,00	5,08	10,41	694,73	2281,89	0,20	0,24	0,17	13,19	11,43	1
40406	1327	7,19	1,23	15,08	9,02	644,94	2619,81	0,17	0,13	0,13	96,06	93,16	2

Station	ABV (m <sup>2</sup> )	LGA (log m <sup>2</sup> )	AGR (%)	AQT (%)	PMN (degrés)	PPM (mm)	TMN (degrés jours)	$\tau$	$\tau_3$	$\tau_4$	Q <sub>obs</sub> (m <sup>3</sup> /s)	Q <sub>sim</sub> (m <sup>3</sup> /s)	Région
40624	4560	8,43	0,45	9,69	6,37	650,42	2413,88	0,15	0,18	0,09	334,75	353,59	2
40840	1277	7,15	1,50	12,34	7,88	579,60	2729,57	0,17	0,20	0,10	116,00	107,52	2
40841	1646	7,41	0,19	12,02	7,45	610,99	2587,12	0,16	0,08	0,20	131,57	132,19	2
41902	3756	8,23	0,00	12,01	5,96	686,98	2518,00	0,18	0,11	0,15	204,85	232,45	2
43012	2593	7,86	3,96	9,05	4,31	614,83	2273,71	0,15	0,12	0,16	264,55	199,82	2
43030	1090	6,99	0,00	10,16	6,75	686,98	2518,00	0,12	- 0,01	0,09	116,21	106,46	2
50135	1563	7,35	0,05	5,74	7,95	634,81	2331,14	0,20	0,24	0,20	234,66	211,56	4
50144	2662	7,89	0,00	8,28	6,73	641,27	2348,70	0,18	0,16	0,19	236,58	267,64	2
50408	1546	7,34	1,50	4,62	11,68	804,95	2255,39	0,16	0,21	0,12	448,22	328,91	4
50409	646	6,47	0,51	5,31	11,87	802,83	2197,10	0,16	0,17	0,13	165,50	142,32	4
50501	19,4	2,97	12,35	0,25	3,52	743,69	2704,98	0,13	- 0,04	0,01	4,49	5,39	1
50702	328	5,79	20,56	2,66	5,41	787,54	2586,46	0,15	0,01	0,06	69,32	61,39	4
50807	223	5,41	0,76	7,83	9,28	815,88	2490,83	0,11	0,05	0,07	48,91	44,04	2
50904	365	5,90	1,70	2,81	9,05	812,52	2343,58	0,12	0,12	0,11	65,63	83,48	4
50915	58,4	4,07	4,87	0,73	6,45	803,34	2566,94	0,16	-	0,03	13,33	17,20	1

Station	ABV (m <sup>2</sup> )	LGA (log m <sup>2</sup> )	AGR (%)	AQT (%)	PMN (degrés)	PPM (mm)	TMN (degrés jours)	$\tau$	$\tau_3$	$\tau_4$	Q <sub>obs</sub> (m <sup>3</sup> /s)	Q <sub>sim</sub> (m <sup>3</sup> /s)	Région
									0,01				
50916	75,5	4,32	0,65	1,09	12,88	818,89	2164,39	0,20	0,19	0,00	32,23	31,91	1
51001	1107	7,01	0,18	2,59	11,08	811,34	2148,46	0,14	0,14	0,16	351,05	296,34	3
51005	267	5,59	0,00	2,15	10,43	821,74	2051,89	0,18	0,19	0,12	91,87	61,28	4
52219	1288	7,16	3,07	6,71	9,65	675,23	2449,02	0,15	0,05	0,10	203,39	216,20	2
52228	188	5,24	0,03	8,92	11,00	674,64	2434,83	0,12	0,08	0,10	36,07	39,46	2
52601	1027	6,93	4,74	8,36	7,83	666,49	2558,72	0,16	0,11	0,12	121,66	131,89	2
52805	769	6,65	0,00	9,62	7,93	660,71	2443,28	0,18	0,07	0,06	95,41	92,48	2
54001	23,9	3,17	5,10	0,26	3,95	777,19	2657,24	0,18	0,37	0,12	5,58	6,54	1
60102	433	6,07	0,02	7,03	12,58	663,49	2176,12	0,15	0,22	0,21	92,00	107,56	2
60704	450	6,11	0,00	2,45	9,49	668,40	2238,41	0,21	0,41	0,29	85,49	99,12	4
60901	358	5,88	5,32	2,81	8,20	652,45	2313,91	0,17	-	-	49,67	72,41	4
									0,08	0,08			
61020	1115	7,02	0,00	4,21	8,54	663,40	2288,74	0,18	0,22	0,33	185,93	181,18	4
61022	490	6,19	0,00	4,90	7,46	686,98	2518,00	0,17	0,09	0,16	93,60	89,60	4
61024	350	5,86	0,00	2,45	8,45	698,23	2168,38	0,17	0,18	0,25	64,50	66,49	4

Station	ABV (m <sup>2</sup> )	LGA (log m <sup>2</sup> )	AGR (%)	AQT (%)	PMN (degrés)	PPM (mm)	TMN (degrés jours)	$\tau$	$\tau_3$	$\tau_4$	Q <sub>obs</sub> (m <sup>3</sup> /s)	Q <sub>sim</sub> (m <sup>3</sup> /s)	Région
61028	2639	7,88	0,00	4,35	8,49	685,03	2214,83	0,16	- 0,04	0,12	408,00	356,66	4
61307	306	5,72	3,45	5,23	6,20	657,34	2298,35	0,15	0,03	0,10	46,76	41,70	1
61801	1020	6,93	5,84	6,51	5,94	633,51	2297,24	0,16	0,09	0,03	120,58	106,04	4
62002	26,8	3,29	47,73	0,58	2,70	583,15	2383,36	0,19	0,21	0,27	4,38	4,87	1
62803	1102	7,00	0,15	7,04	10,42	694,22	2128,60	0,15	0,11	0,07	253,01	195,75	2
70204	768	6,64	0,09	8,84	9,95	686,89	2126,03	0,18	0,15	0,18	122,78	117,22	2
80101	3722	8,22	5,69	10,72	2,81	612,89	2308,34	0,13	0,10	0,22	184,74	203,11	2
120201	26,6	3,28	57,53	0,19	2,65	756,87	2714,64	0,18	0,31	0,28	10,10	6,50	1

ABV : superficie du bassin versant; LGA : logarithme de la superficie du bassin versant; AGR : %Milieu agricole; PMN : pente moyenne; PPM : Cumul moyen des précipitations entre le 1<sup>er</sup> avril et le 31 octobre (période de 1974 à 2003); TMN : Moyenne des degrés-jours (base de 0°C) cumulés entre 1979 et 2008 ;  $\tau$  : L-Coefficient de variation;  $\tau_3$  L-Asymétrie;  $\tau_4$  : L-Aplatissement, Q<sub>obs</sub> : débit maximal annuel moyen observé; Q<sub>sim</sub> : débit maximal annuel moyen simulé; Reg : région.

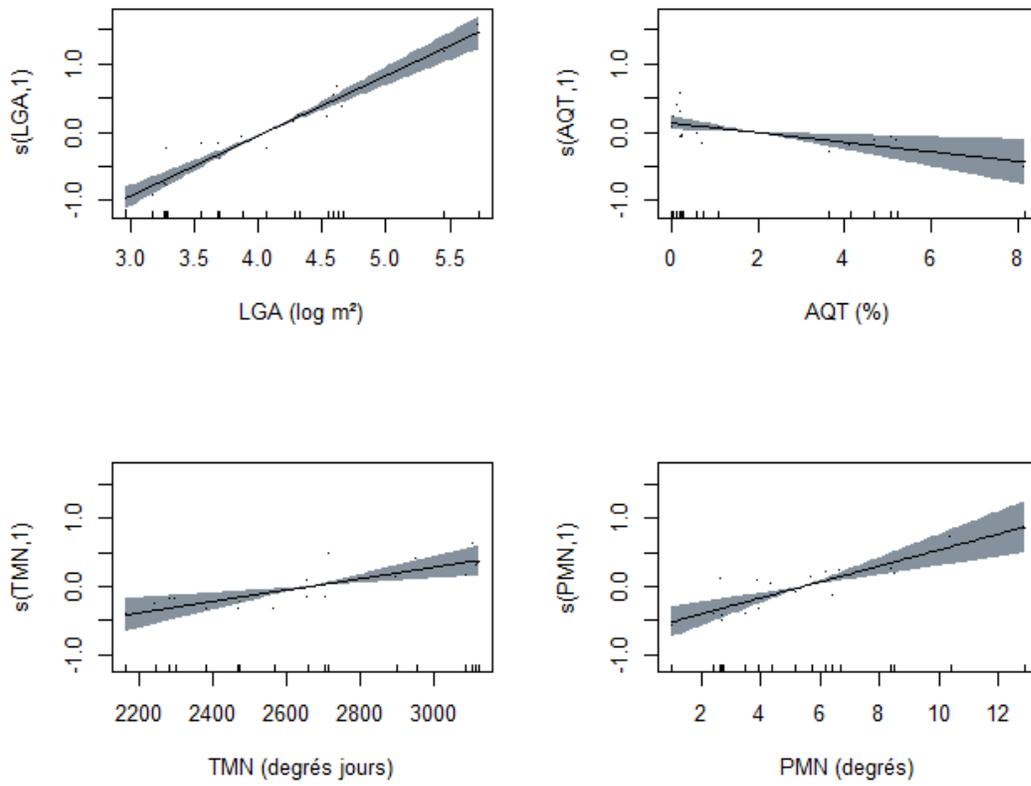


Figure 4 : Relations de variation partielle (Région 1)

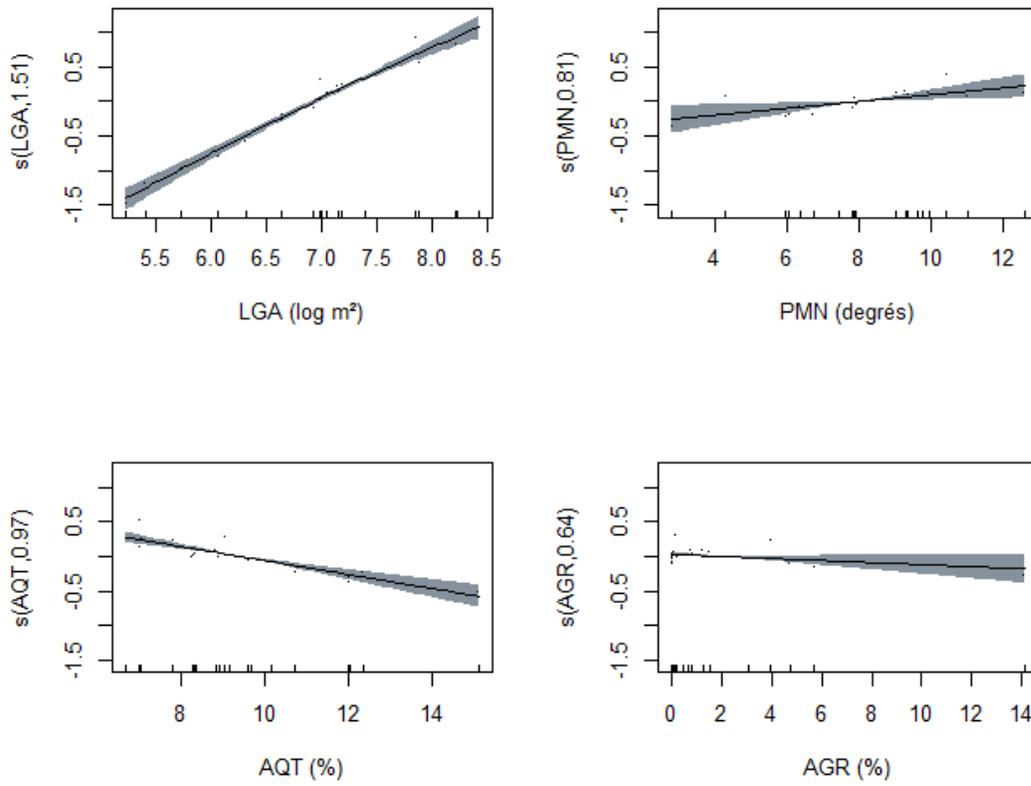


Figure A3 : Relations de variation partielle (Région 2)

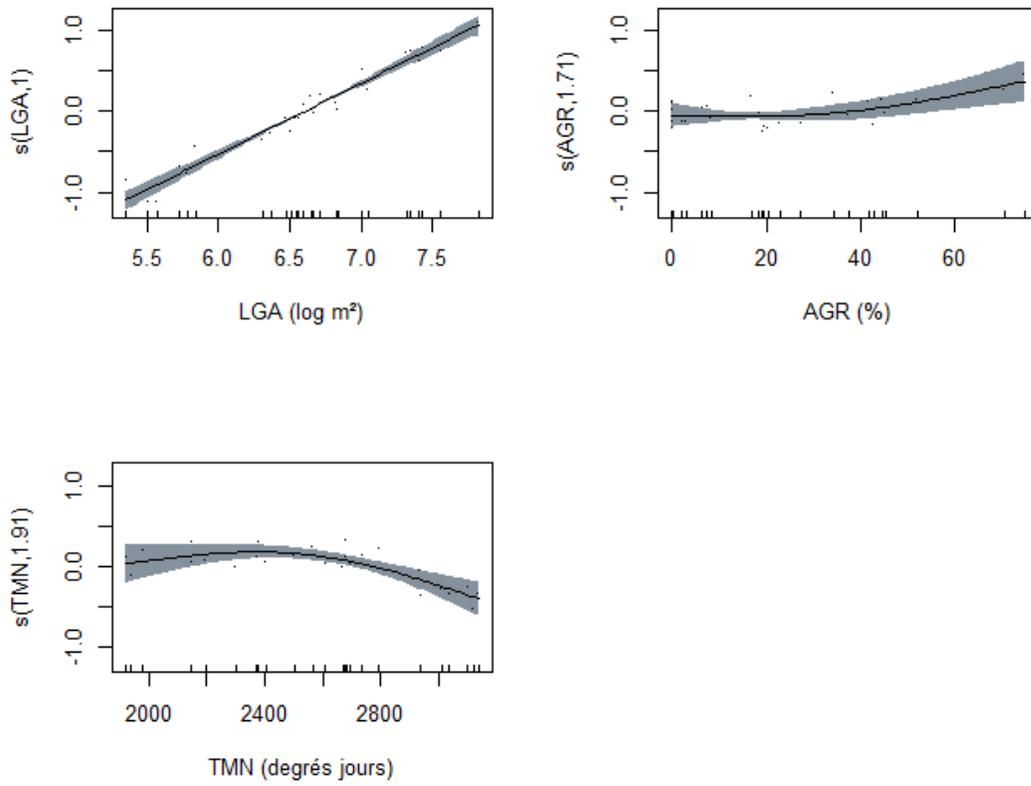


Figure 5 : Relations de variation partielle (Région 3)

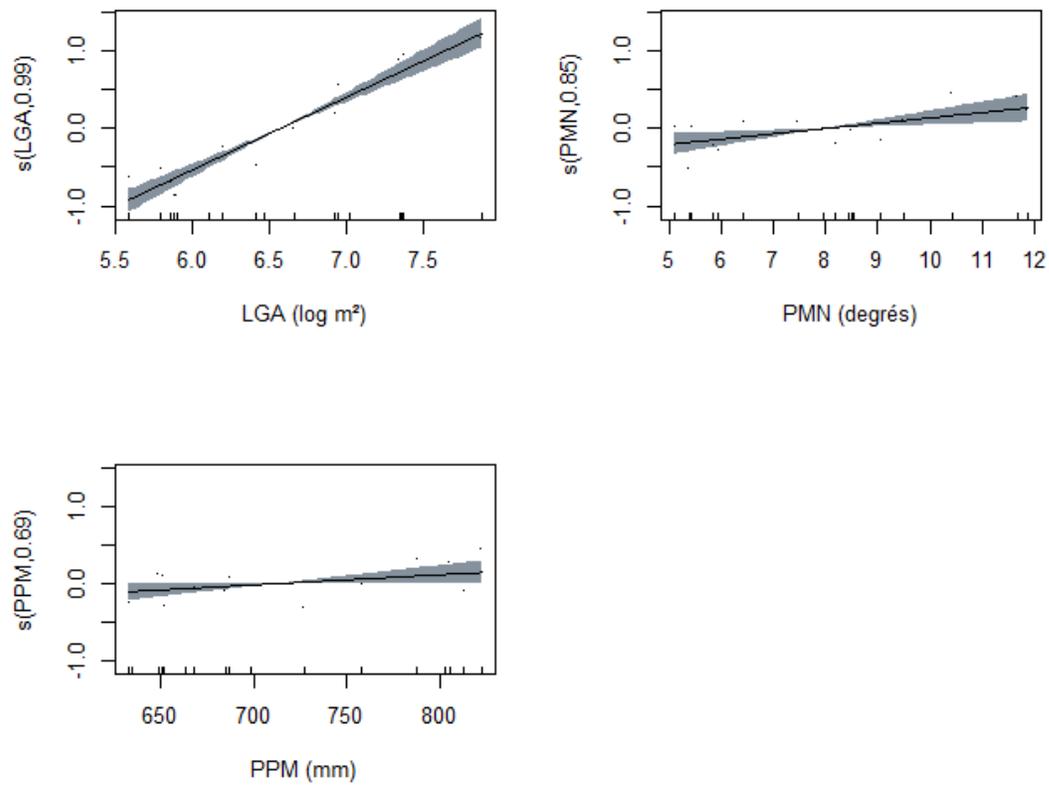


Figure A5 : Relations de variation partielle (Région 4)

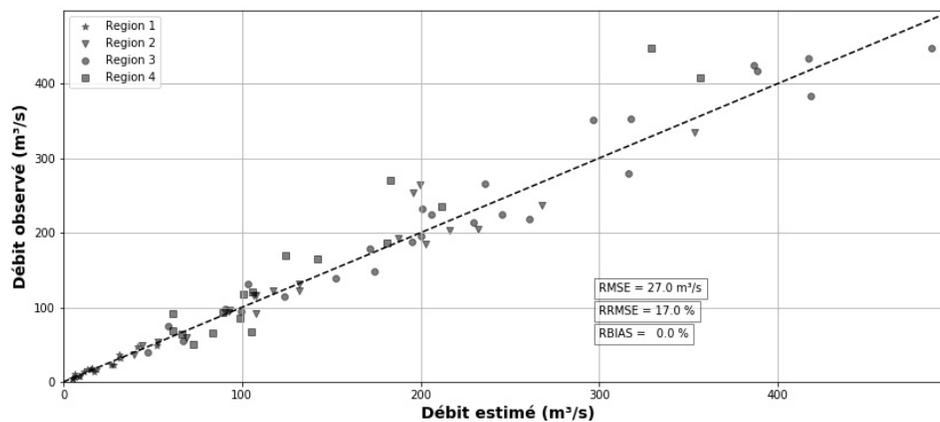


Figure A6 : Nuage de points des débits simulés vs débits observés

*Tableau A4 : Données des sites à l'étude*

Site	ABV (m <sup>2</sup> )	LGA (log (m <sup>2</sup> ))	AGR (%)	AQT (%)	PMN (degrés)	PPM (mm)	TMN (degrés jours)	Q <sub>sim</sub> (m <sup>3</sup> /s)	Région
Site 9	60,92	4,10	0,87	0,92	5,63	790,53	2597,77	16,29	1
Site 28	14,9	2,70	30,58	0,0	4,3	753,53	2703,48	4,83	1
Site 4	62	4,12	38,71	0,23	4,44	773,53	2622,07	15,50	1

# Annexe 3 : Résultats de la modélisation hydraulique

## 1. Site 4

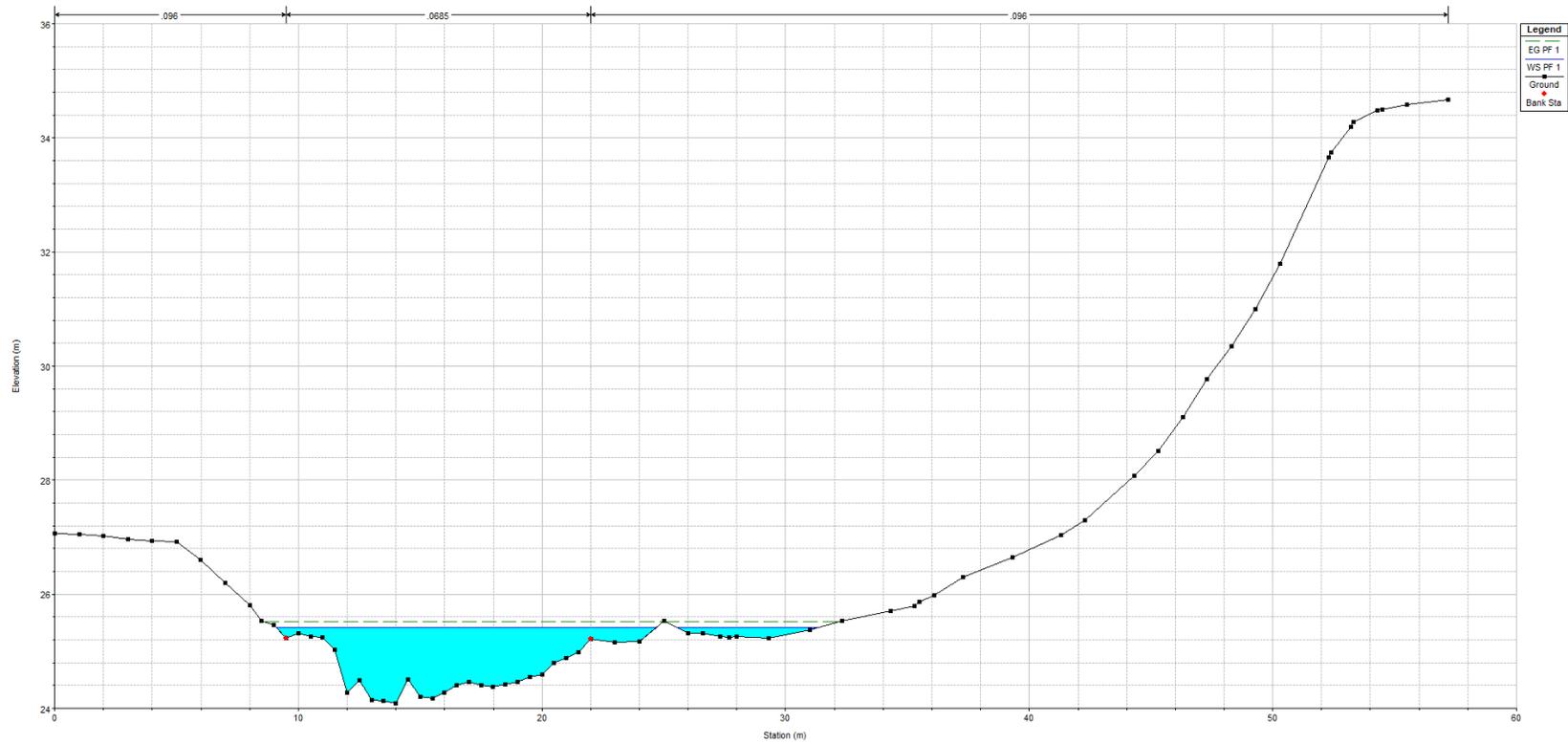


Figure A7 : Section transversale 1 (Site 4)

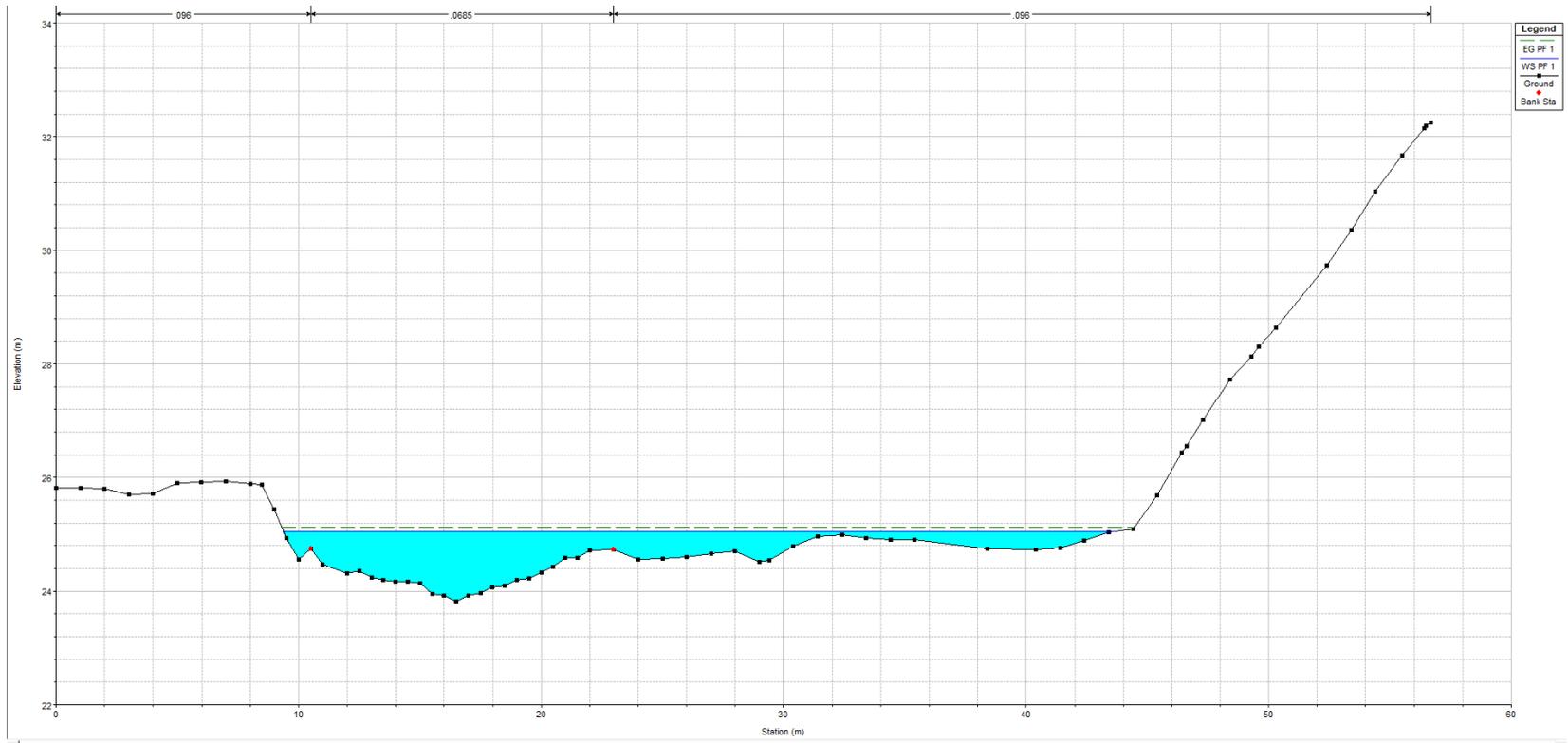


Figure A8 : Section transversale 2 (Site 4)

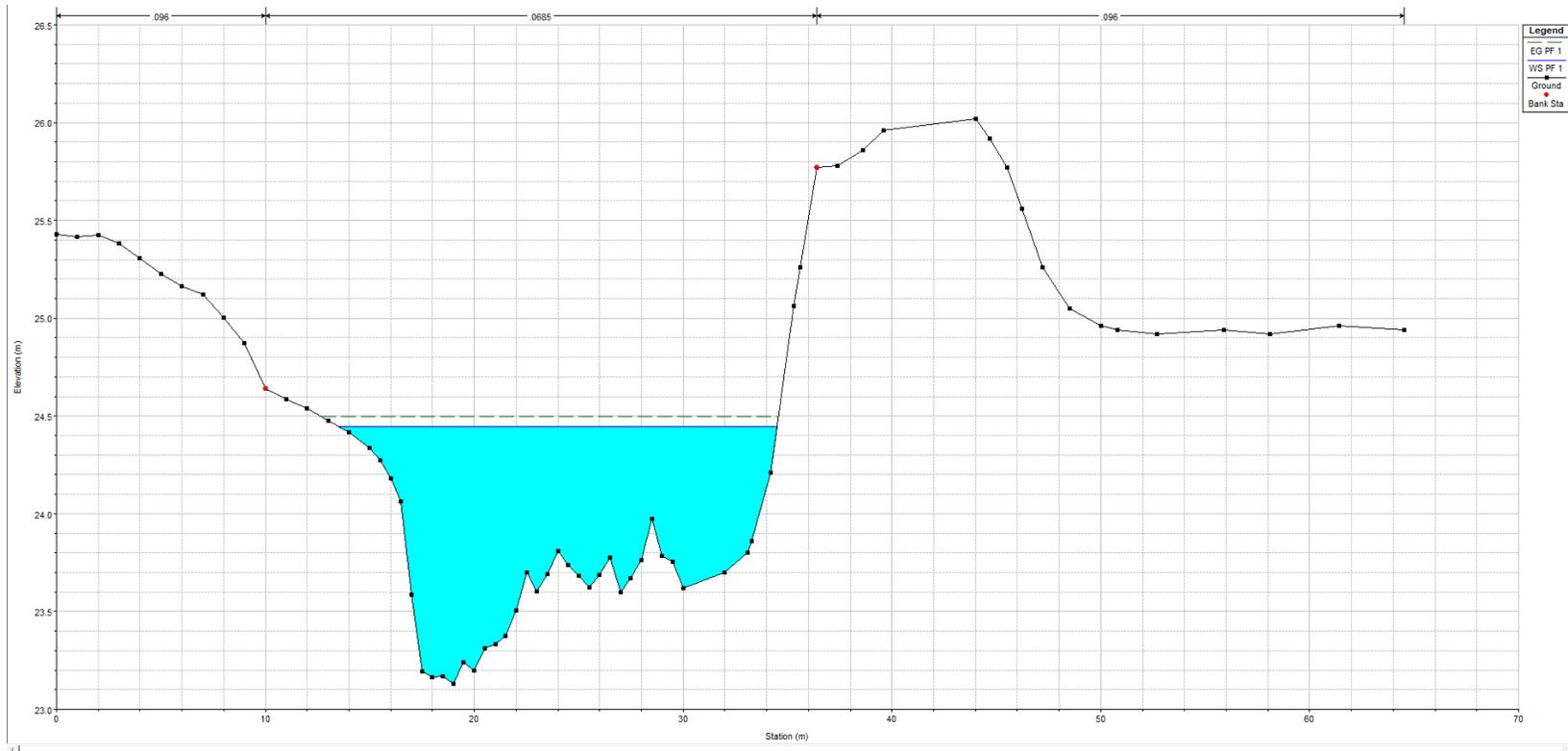


Figure A9 : Section transversale 3 (Site 4)

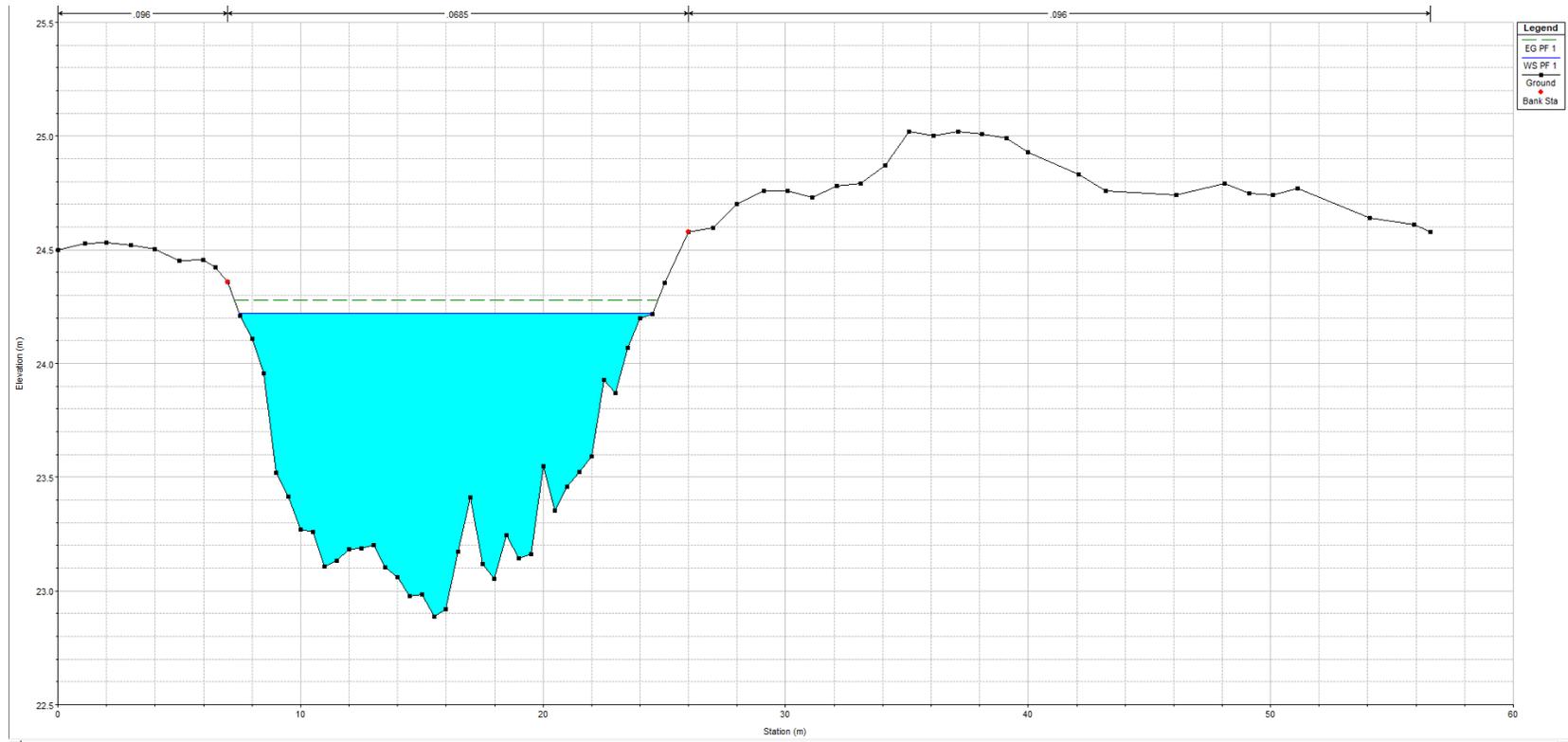


Figure A10 : Section transversale 4 (Site 4)

## 2. Site 9

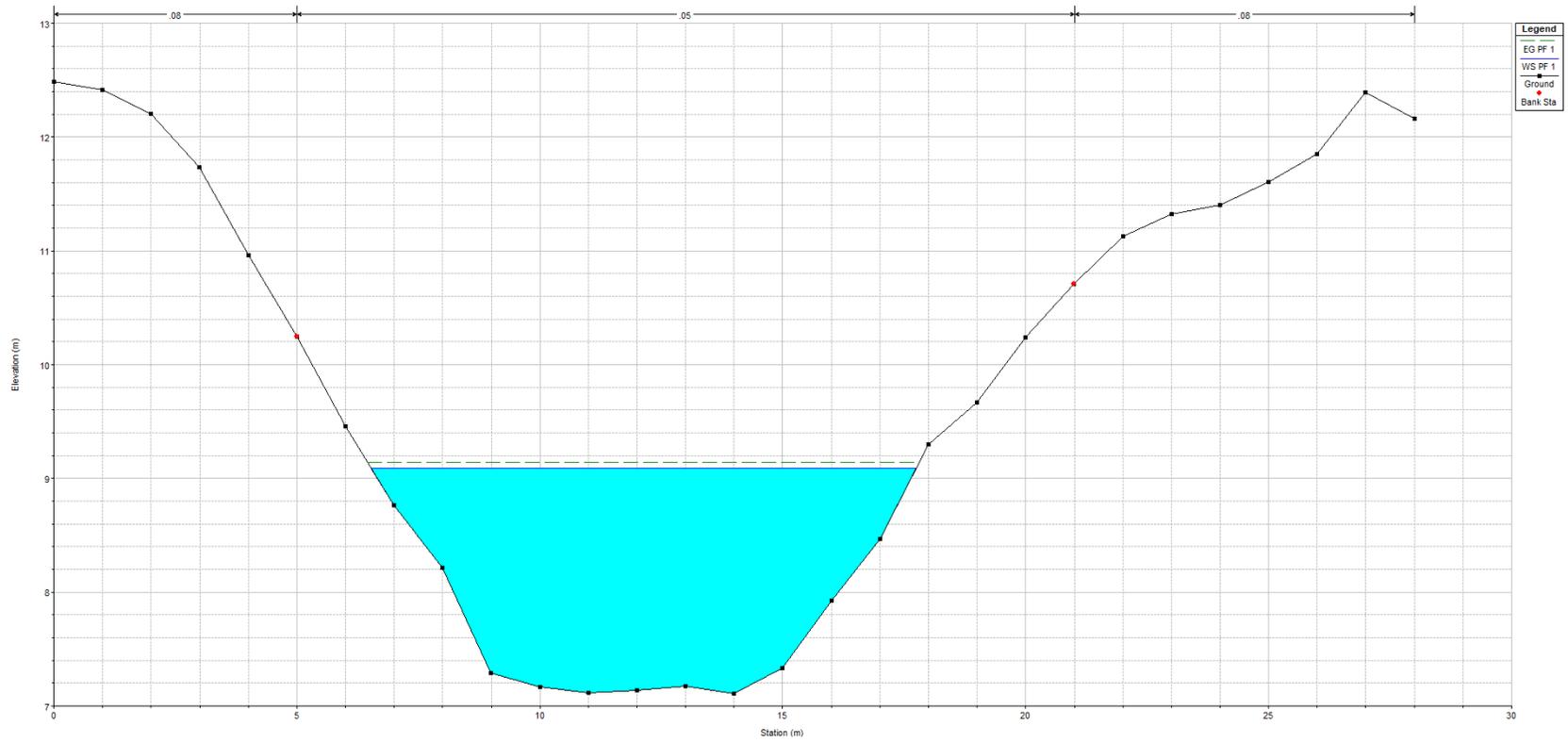


Figure A11 : Section transversale 1 (Site 9)

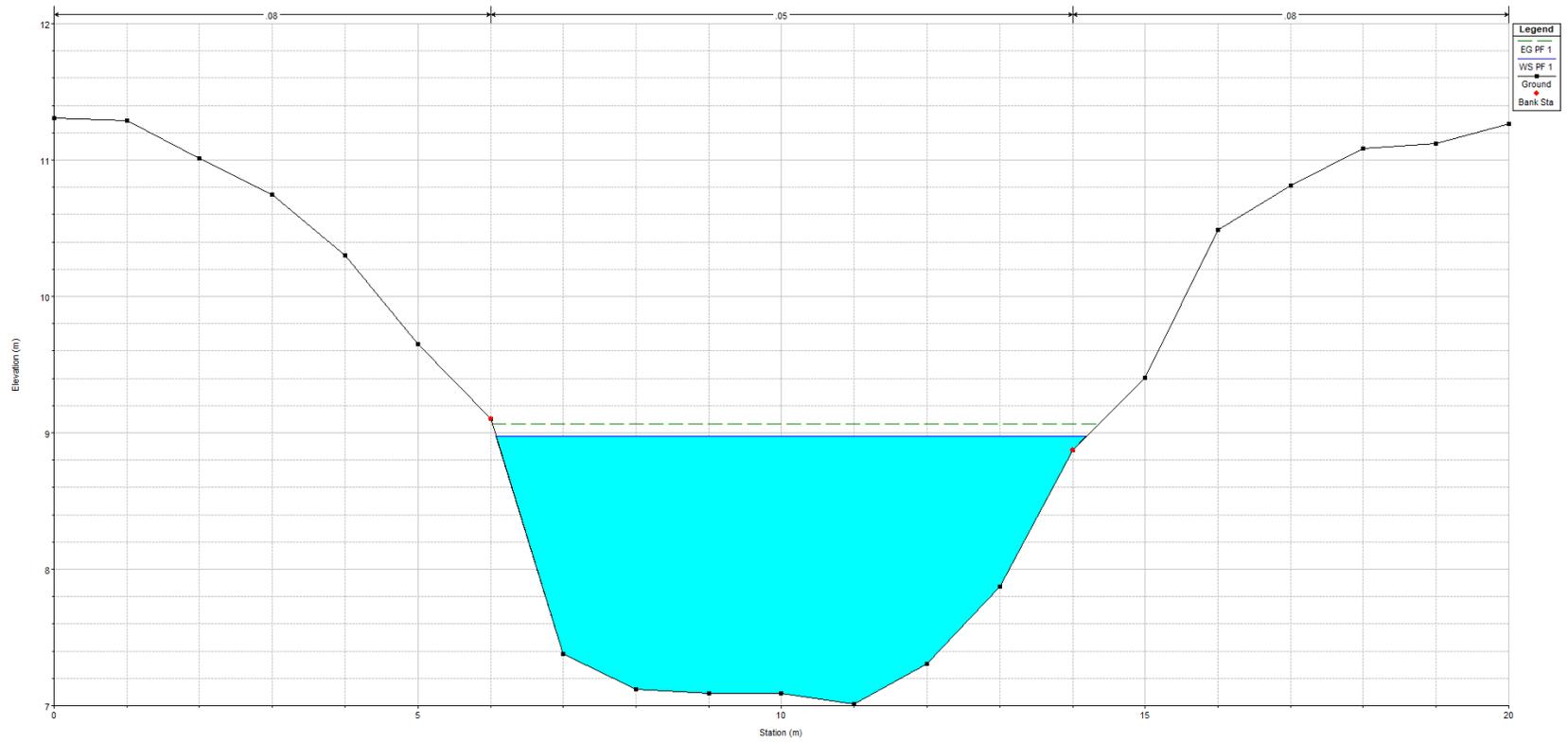


Figure A12 : Section transversale 2 (Site 9)

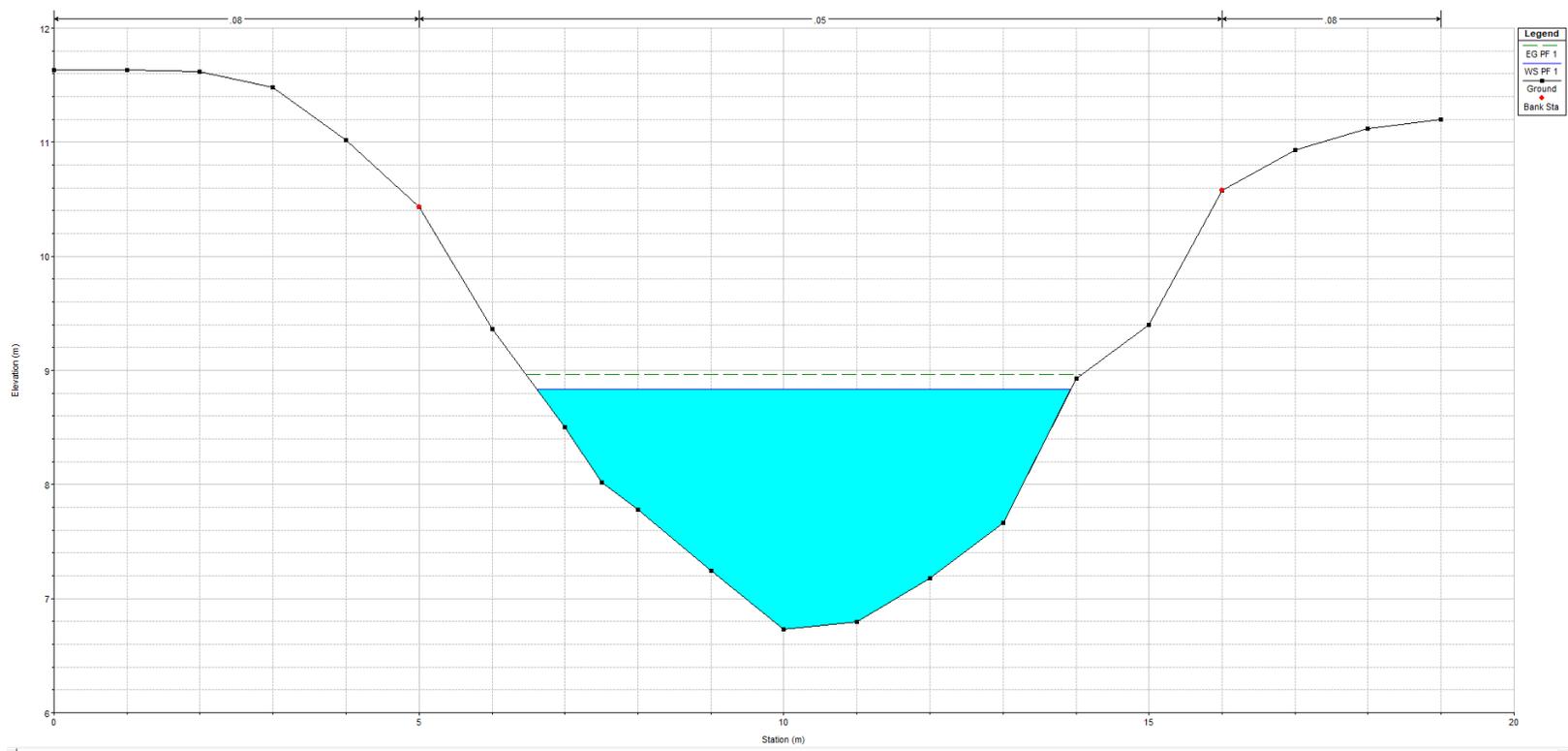


Figure A13 : Section transversale 3 (Site 9)

### 3. Site 28

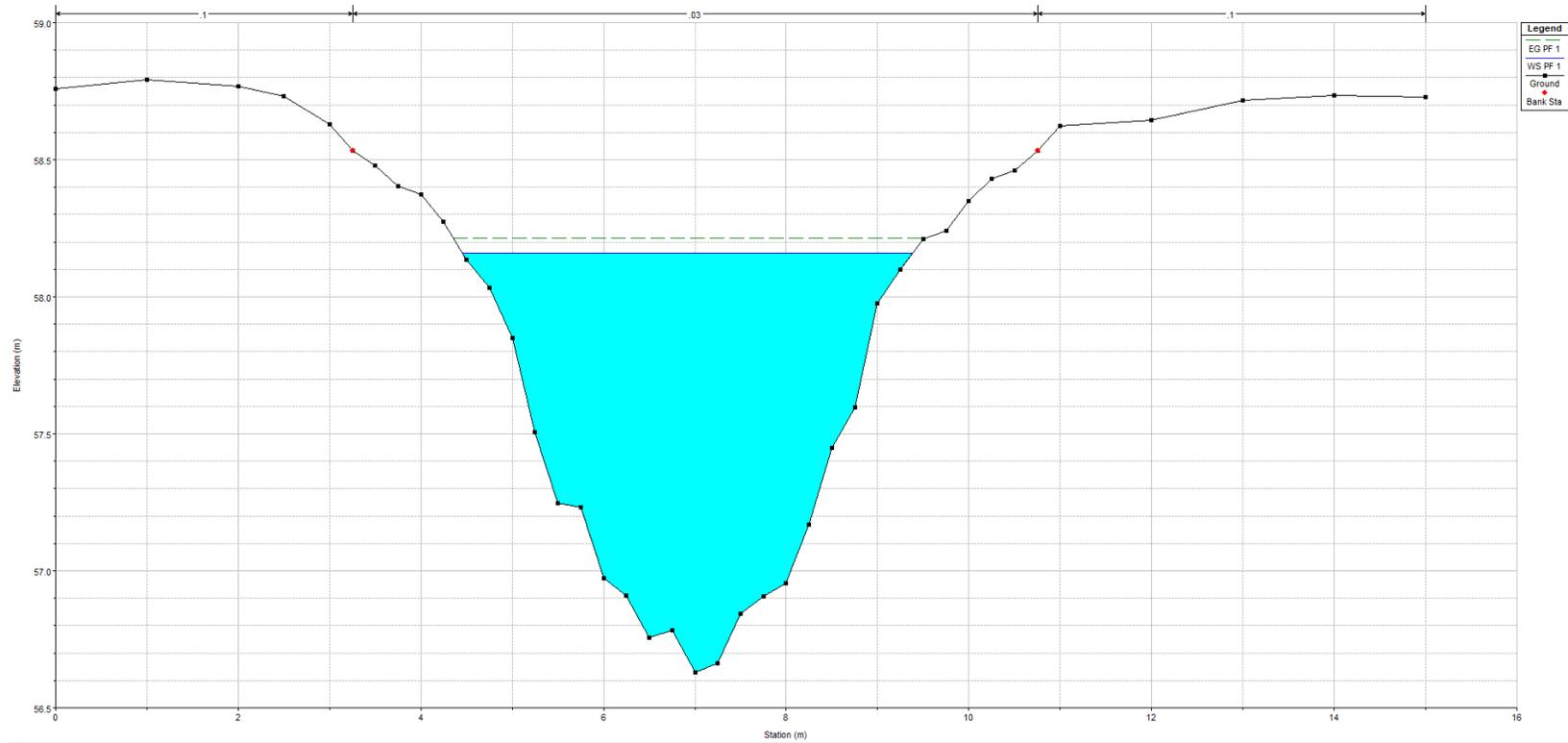


Figure A14 : Section transversale 1 (Site 28)

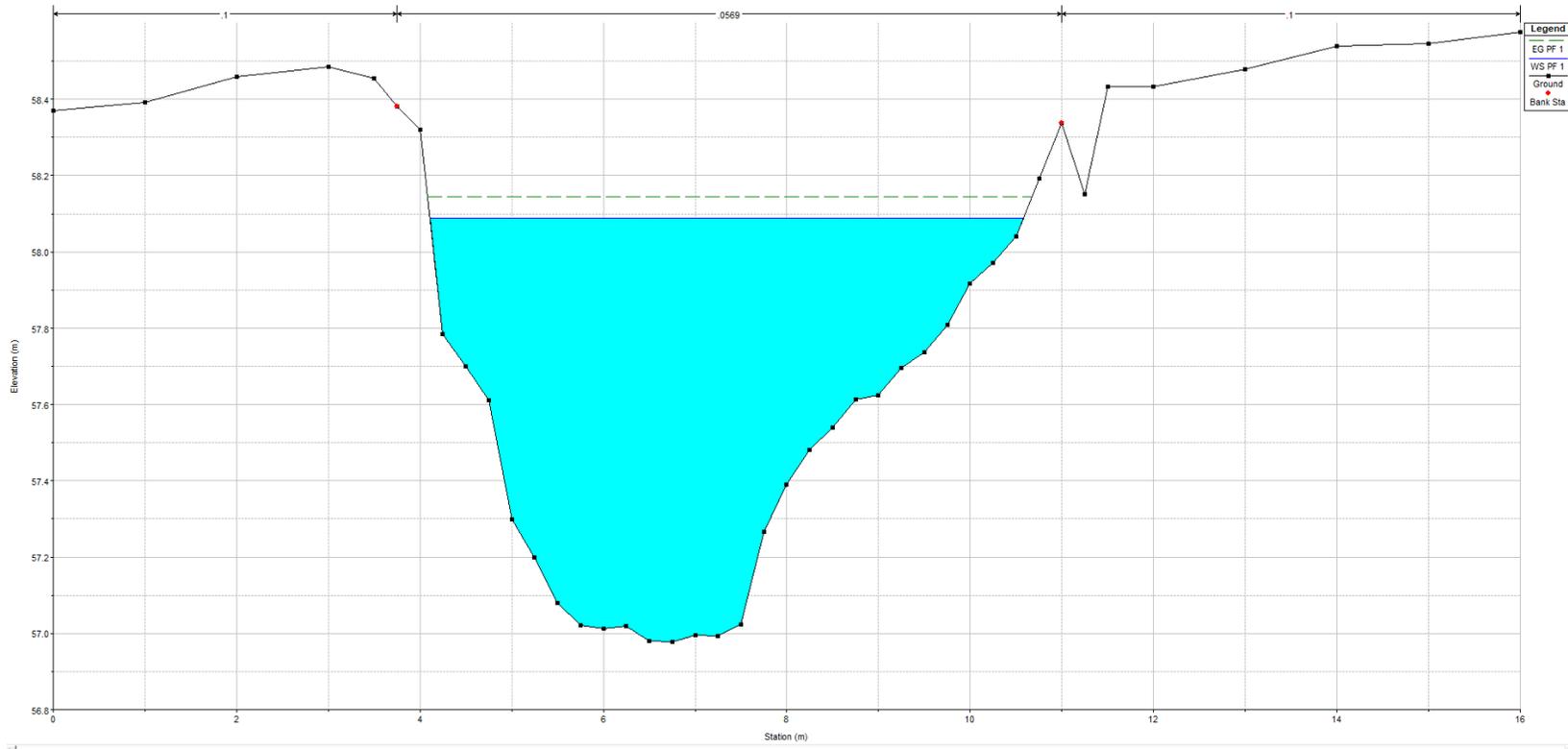


Figure A15 : Section transversale 2 (Site 28)

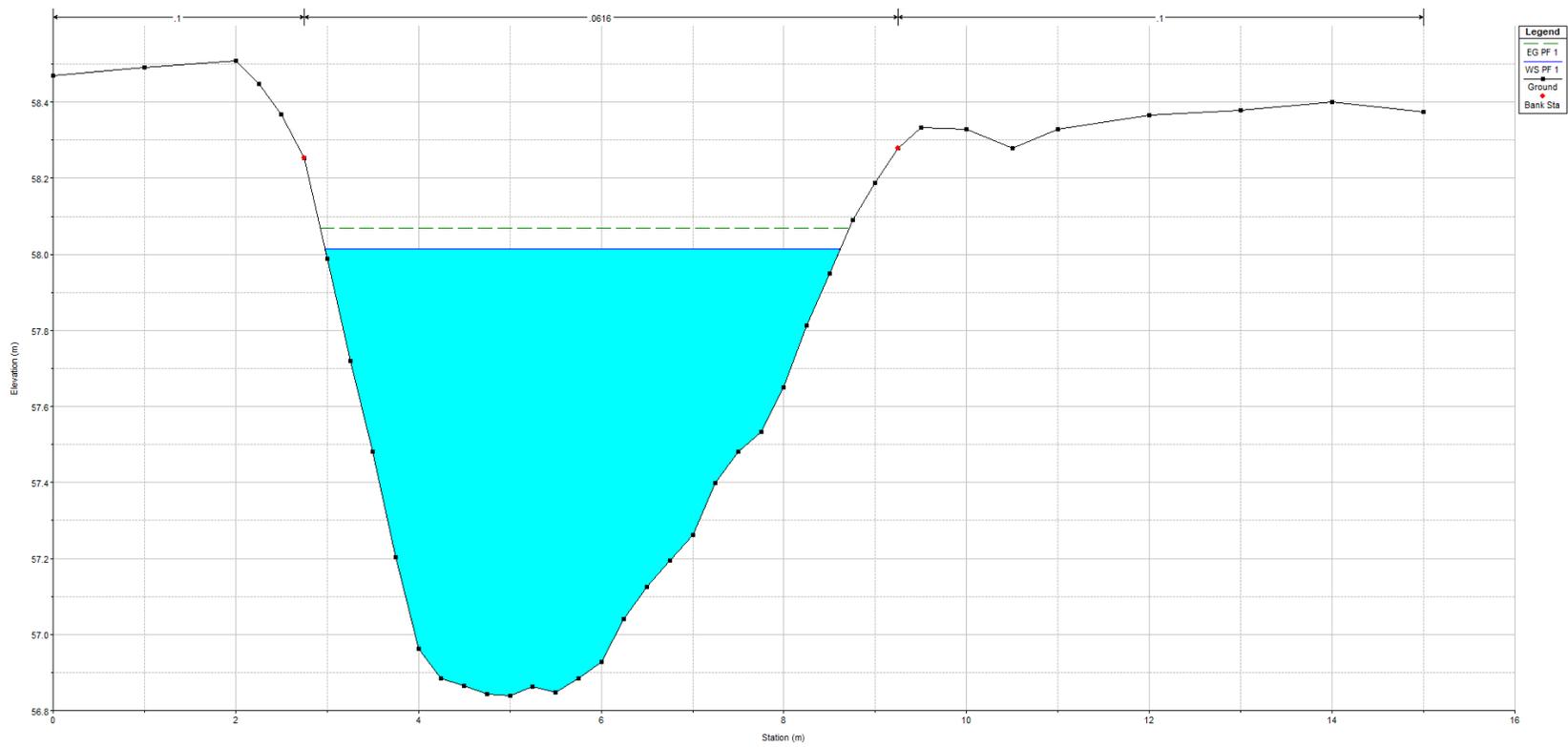


Figure A16 : Section transversale 3 (Site 28)

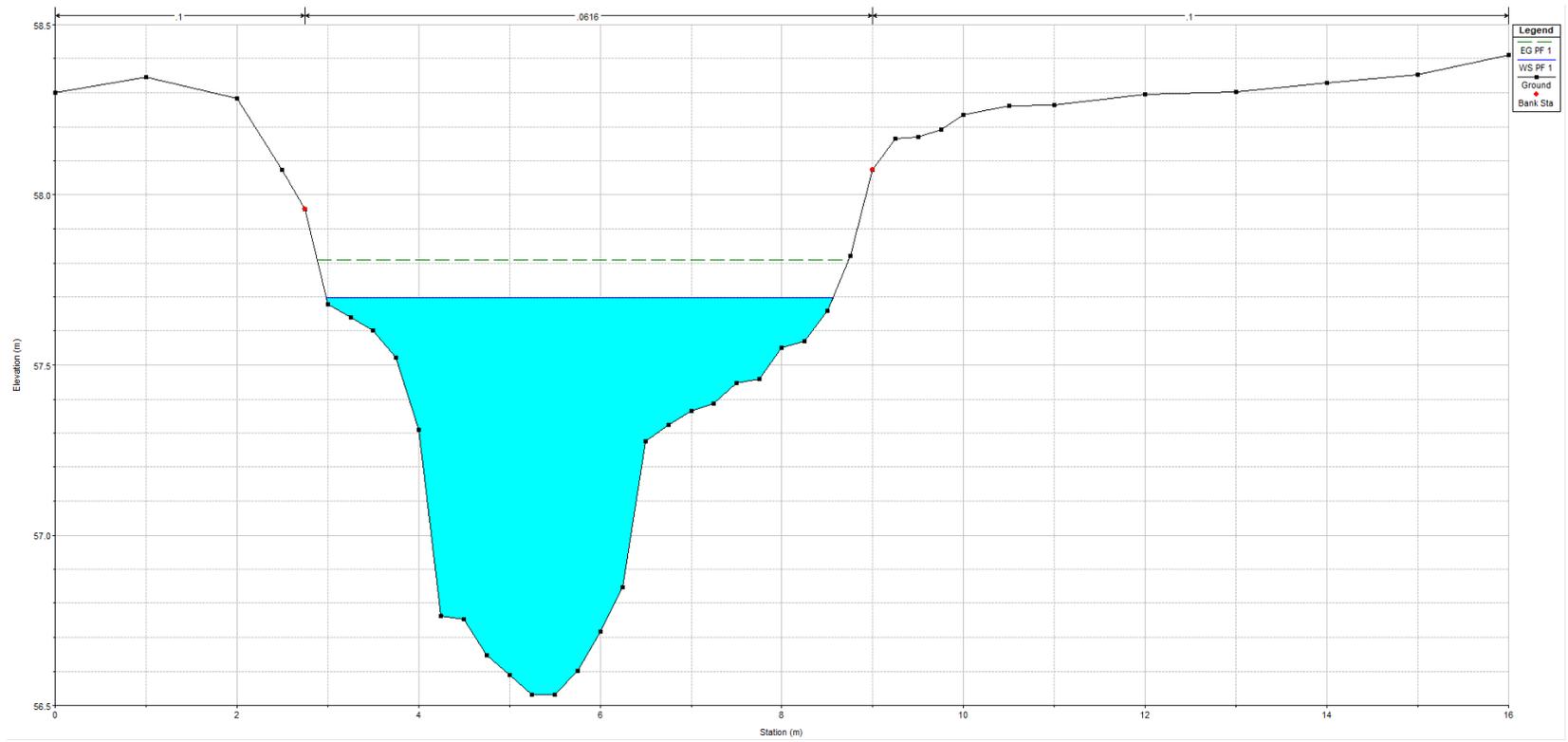


Figure A17 : Section transversale 4 (Site 28)

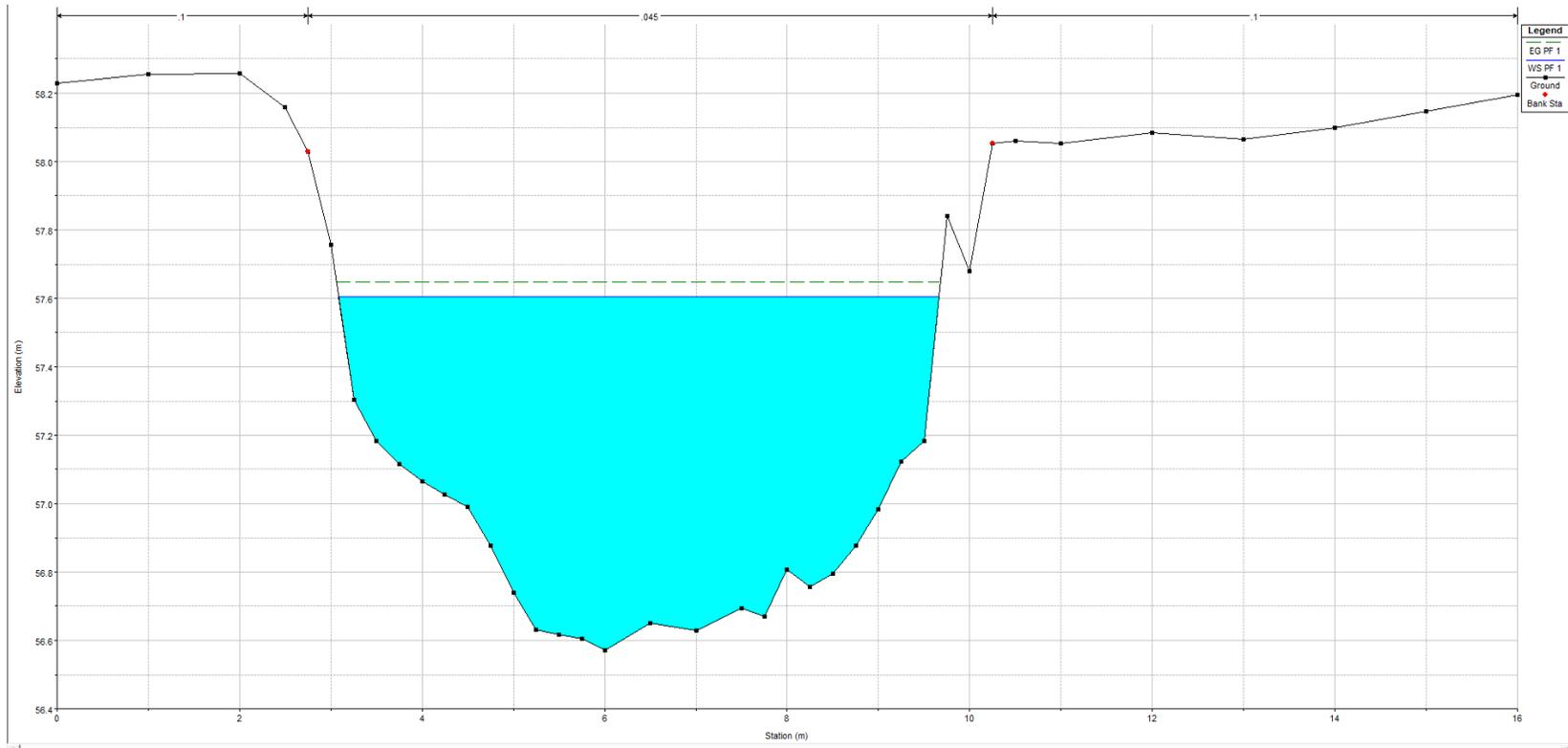


Figure A18 : Section transversale 5 (Site 28)

## Annexe 4 : Pénalisation de la fonction de lissage dans mgcv et sélection des variables

### **Pénalisation de la fonction de lissage**

Une manière proposée par Wood (2006) pour la pénalisation de la fonction de lissage est l'ajout du terme suivant à la fonction objectif initiale :

*Équation 14*

$$\int_{i=1}^n [f''(x_i)]^2 dx$$

Où  $n$  correspond au nombre total d'observations.

Les fonctions de lissage étant doublement dérivables, le terme à l'Équation 14 permet donc de mesurer leur courbure (oscillation) totale sur tout le domaine.

La fonction objectif est alors définie de la manière suivante :

*Équation 15*

$$J(w) = \|y - X\beta\|^2 + \lambda \int_{i=1}^n [f''(x_i)]^2 dx$$

où  $y$  est le vecteur d'observations.

Wood (2006) a démontré qu'il est possible de réécrire le terme de pénalisation comme suit :

*Équation 16*

$$\int_{i=1}^n [f''(x_i)]^2 dx = \beta^T S \beta$$

L'Équation 12 devient :

*Équation 17*

$$J(w) = \|y - X\beta\|^2 + \lambda \beta^T S \beta$$

où  $S$  est la matrice des pénalités et  $\beta$  est le vecteur des paramètres de la régression à estimer. Le terme de pénalité permet de pénaliser les modèles excessivement oscillants

(Wood 2006).  $\lambda_i$  est le paramètre permettant de contrôler le degré de lissage. Pour  $\lambda = 0$ , la fonction de lissage n'est pas pénalisée et est donc sur-ajustée aux données observées. Lorsque  $\lambda \rightarrow \infty$ , la fonction de lissage est sous-ajustée aux données et tend à une ligne droite. Une valeur adéquate de  $\lambda$  permet d'obtenir un modèle qui est un bon compromis entre le modèle sous-ajusté et le modèle sur-ajusté. Wood (2006) propose plusieurs techniques de sélection de la valeur de  $\lambda$  qui sont basées sur des critères de performance du modèle dont, entre autres, l'erreur de prédiction par validation croisée (GCV) et l'erreur quadratique (UBRE). Dans le cas de la méthode GCV, par exemple, la valeur de  $\lambda$  est choisie de manière à maximiser la performance du modèle sur des données qui n'ont pas été utilisées pour le calage du modèle. La performance du modèle est évaluée avec la métrique suivante :

*Équation 18*

$$v_g = \frac{n\|y - \hat{y}\|^2}{[\text{tr}(I - A)]^2}$$

où  $\hat{y}$  est le vecteur de prédiction,  $\text{tr}$  est un opérateur qui retourne la trace d'une matrice et  $A$  est la matrice «chapeau» de la régression. De l'Équation 15 on retrouve que :

*Équation 19*

$$A = X(X^T X + \lambda S)^{-1} X^T$$

L'approche UBRE consiste à choisir un  $\lambda$  qui minimise l'équation suivante :

*Équation 20*

$$v_u(\lambda) = \frac{\|y - Ay\|^2}{n} - \sigma^2 + 2\text{tr}(A)\sigma^2/n$$

où  $v_u(\lambda)$  est l'espérance de l'erreur quadratique moyenne du modèle,  $n$  est le nombre d'échantillons et  $\sigma^2$  est la variance de la fonction lien.

## Méthode de sélection des variables significatives

Le terme de pénalité tel que défini à l'Équation 14 pénalise l'oscillation de la fonction de lissage à travers ses courbures (intégrale de la dérivée seconde). Cependant, certaines fonctions de base  $b_{ji}(x)$  de la fonction de lissage sont linéaires et n'ont donc pas des courbures qui peuvent être pénalisées en utilisant leurs dérivées secondes ( $b_{ij}''(x) = 0$ ). Cela évite donc que des variables non significatives soient complètement éliminées de l'équation de régression même pour des valeurs très élevées de  $\lambda$ .

Les fonctions de base non linéaires appartiennent à l'espace colonne de la matrice de pénalité et celles linéaires appartiennent à son espace nul (Marra et Wood 2011). Afin de pénaliser toutes les composantes de la fonction de lissage, il faut ajouter un autre terme qui permet de pénaliser les éléments de l'espace nul.

La matrice de pénalité peut être décomposée en éléments propres comme suit :

*Équation 21*

$$S = U\Lambda U^T$$

où  $U$  est la matrice des vecteurs propres de la matrice  $S$  et  $\Lambda$  est la matrice diagonale des valeurs propres de  $S$ .  $\Lambda$  contient des valeurs propres nulles étant donné que la matrice  $S$  a un espace nul. Cela ne permet pas de pénaliser à zéro certains termes lorsque  $\lambda$  devient trop élevé. Marra et Wood (2011) ont proposé l'ajout d'une pénalité supplémentaire permettant de pénaliser les composantes de l'espace nul de  $S$  :

*Équation 22*

$$S^* = U^*U^{*T}$$

où  $U^*$  est la matrice des vecteurs propres associés aux valeurs propres nuls.

Nous pouvons donc réécrire La fonction objectif comme suit :

*Équation 23*

$$J(w) = \|y - X\beta\|^2 + \lambda\beta^T S\beta + \lambda^*\beta^T S^*\beta$$

La fonction objectif telle que définie à l'Équation 23 permet aussi bien de pénaliser la fonction de lissage que de sélectionner les variables significatives lors de l'ajustement du modèles.