

Titre: Estimation des courbes de tarage et de leurs incertitudes pour les
Title: cours d'eau du Québec méridional

Auteur: Rahma Jebali
Author:

Date: 2021

Type: Mémoire ou thèse / Dissertation or Thesis

Référence: Jebali, R. (2021). Estimation des courbes de tarage et de leurs incertitudes pour
Citation: les cours d'eau du Québec méridional [Mémoire de maîtrise, Polytechnique
Montréal]. PolyPublie. <https://publications.polymtl.ca/9931/>

 **Document en libre accès dans PolyPublie**
Open Access document in PolyPublie

URL de PolyPublie: <https://publications.polymtl.ca/9931/>
PolyPublie URL:

**Directeurs de
recherche:** Jonathan Jalbert
Advisors:

Programme: Maîtrise recherche en mathématiques appliquées
Program:

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

**Estimation des courbes de tarage et de leurs incertitudes pour les cours d'eau
du Québec méridional**

RAHMA JEBALI

Département de mathématiques et de génie industriel

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
Mathématiques appliquées

Décembre 2021

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

**Estimation des courbes de tarage et de leurs incertitudes pour les cours d'eau
du Québec méridional**

présenté par **Rahma JEBALI**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*
a été dûment accepté par le jury d'examen constitué de :

Antoine SAUCIER, président

Jonathan JALBERT, membre et directeur de recherche

Françoise BICHAI, membre

DÉDICACE

*À l'hommage de l'âme de la femme qui a
guidé mes premiers pas, Fatma Sayedi,
"Rahimaha'ALLAH" . . .*

REMERCIEMENTS

Cette maîtrise a présenté pour moi une chance pour découvrir le domaine des mathématiques appliquées aux sciences de l'eau. Il s'agissait de deux ans de découverte et d'apprentissage qui m'ont demandé beaucoup de motivation et de support. Il m'apparaît important d'exprimer mes profondes reconnaissances et remerciements pour toutes personnes qui m'ont aidé à surmonter tous les défis rencontrés.

Je commence par remercier ALLAH mon créateur qui m'a guidé à poursuivre cette maîtrise et qui m'a doté d'intelligence pour étudier une petite partie de son infini sagesse et de son tout puissance dans la création de ce monde. Merci ALLAH de m'avoir guidé à lever les mains pour demander votre aide et merci de votre réponse immédiate à mon Doua.

Ensuite, je souhaite exprimer mes vifs remerciements à mon directeur de recherche Jonathan Jalbert pour sa direction, sa disponibilité tout au long de mon projet, son orientation et ses conseils. Jonathan a su me redonner confiance à chaque défi rencontré et il m'a consacré beaucoup d'effort durant cette maîtrise.

Un merci spécial à Claudie-Ratté-Fortin pour son aide précieux dans mon travail, sa disponibilité et ses conseils.

Un merci spécial aussi à Gabriel Gobeil pour son assistance, ses conseils de programmation et son grand effort pour m'aider à accomplir mon mémoire dans les meilleurs délais.

Également, je souhaite remercier mes amis Maha Badri et Ala Bejaoui qui ont été toujours disponibles pour répondre à mes questions et pour m'aider dans la programmation.

Je souhaite aussi remercier Wissem Maazoun de m'avoir fourni les meilleures informations sur cette maîtrise.

De façon plus formelle, je remercie le Ministère de l'Environnement et de la Lutte contre les changements climatiques, Ouranos et Mitacs d'avoir financé ce projet.

Finalement, je tiens à rendre grâce à mon cher père Abdelkhalek pour son encouragement à poursuivre mes études, pour son soutien moral et financier et pour son Doua et à ma très chère mère, cette femme exigeante et ambitieuse qui est à l'origine de toutes mes réussites. Merci à mes sœurs Rihab et Mariem Rahimaha'ALLAH, à mon frère Hamza, à ma famille et à mes amis qui m'ont supporté tout au long cette maîtrise, souvent sans même le savoir.

RÉSUMÉ

Les mesures de débits des cours d'eau sont très importantes dans plusieurs domaines en hydrologie, en génie civil et en aménagement du territoire, par exemple pour la cartographie des plaines inondables. Cependant, il est impossible de mesurer le débit facilement et en temps réel. Alors, la hauteur du cours d'eau est mesurée et l'estimation du débit est obtenue avec une relation hydraulique hauteur-débit appelée courbe de tarage. Cependant, ces courbes sont sujettes à des incertitudes importantes : les imprécisions des jaugeages, l'instabilité géomorphologique de la rivière et l'estimation des paramètres de la courbe. Par conséquent, l'estimation de ces incertitudes est une étape cruciale pour l'utilisation de ces courbes. Notamment, après les inondations exceptionnelles vécues au Québec au printemps 2017, un grand projet a été lancé en collaboration avec plusieurs partenaires qui s'appelle "INFO-Crue". Ce projet vise à mettre à jour les cartes des zones inondables. Ce mémoire vient à contribuer dans ce projet et il a pour but de développer une méthode d'estimation des courbes de tarage spécialement adaptée aux stations hydrométriques du Québec méridional et permettant une prise en compte rigoureuse des incertitudes correspondantes.

Ce mémoire répond à cet objectif en développant une méthodologie adaptée pour 15 stations hydrométriques représentatives des différentes caractéristiques du réseau hydrométrique du Québec méridional. Plus spécifiquement, cette méthodologie consiste à estimer les courbes de tarage en se basant sur la relation hydraulique hauteur-débit dans l'espace logarithmique. Cette estimation est effectuée par la méthode des moindres carrés adaptée de la méthode ISO/WMO. La méthode développée constitue un outil efficace pour estimer les courbes de tarage et les incertitudes correspondantes aux prédictions du débit. La méthode développée permet d'ajouter les incertitudes sur les débits mesurés dans l'estimation des débits prédits, ce que ne permet pas la méthode ISO/WMO. Finalement, la méthodologie permet l'estimation de l'incertitude de tous les paramètres de la courbe de tarage en utilisant une méthode de *bootstrap non paramétrique*, encore une fois ce que ne permet pas la méthode ISO/WMO. La méthode a été appliquée à 15 stations hydrométriques du Québec et elle a permis de modéliser adéquatement la relation hauteur-débit ainsi que les incertitudes correspondantes.

ABSTRACT

Streamflow measurements are very important in several areas of hydrology, civil engineering, and urban development, for example for mapping flood plains. However, it is impossible to measure the discharge easily and in real time. Hence, the stage of the streamflow is measured, and the estimate of the discharge is obtained with a hydraulic stage-discharge relationship called the rating curve. However, these curves are subject to important uncertainties: the inaccuracies of the gauges, the geomorphological instability of the river and the estimation of the parameters of the curve. Therefore, estimating these uncertainties is a crucial step in using these curves. In particular, after the exceptional floods experienced in Quebec in the spring of 2017, a big project was launched in collaboration with several partnerships called "INFO-Crue". This project aims to update the maps of flood zones. This work contributes to this project and its aim is to develop a method for estimating rating curves specially adapted to hydrometric stations in southern Quebec and allowing rigorous consideration of the corresponding uncertainties.

This work responds to this objective by developing a methodology adapted for 15 hydrometric stations representative of the different characteristics of the hydrometric network of southern Quebec. More specifically, this methodology consists in estimating the rating curves based on the hydraulic stage-discharge relationship in logarithmic space. This estimate is made by the least squares method adapted from the ISO / WMO method. The developed method constitutes an efficient tool for estimating the rating curves and the uncertainties corresponding to the discharge predictions. It allows to add the uncertainties of measurement in the estimation of the predicted discharge, which the ISO / WMO method does not allow. Finally, the methodology allows the estimation of the uncertainty of all the parameters of the rating curve using a *nonparametric bootstrap* method, what the ISO / WMO method does not allow as well. The method was applied to 15 hydrometric stations in Quebec, and it allowed us to adequately model the stage - discharge relationship as well as the corresponding uncertainties.

TABLE DES MATIÈRES

DÉDICACE	iii
REMERCIEMENTS	iv
RÉSUMÉ	v
ABSTRACT	vi
TABLE DES MATIÈRES	vii
LISTE DES TABLEAUX	ix
LISTE DES FIGURES	x
LISTE DES SIGLES ET ABRÉVIATIONS	xiii
LISTE DES ANNEXES	xiv
CHAPITRE 1 INTRODUCTION	1
CHAPITRE 2 CADRE THÉORIQUE	6
2.1 Fondement hydraulique de la relation hauteur-débit	6
2.2 Incertitude sur la relation hauteur-débit	9
2.3 Régression linéaire	10
2.4 La méthode de rééchantillonnage bootstrap	12
CHAPITRE 3 REVUE DE LITTÉRATURE	14
3.1 Estimation des courbes de tarage basées sur la relation hydraulique	14
3.1.1 Estimation par les moindres carrés	14
3.1.2 Estimation bayésienne	14
3.2 Courbes de tarage non basées sur l'hydraulique	16
3.3 Traitement dynamique de la courbe de tarage	17
3.4 Analyse critique	17
3.4.1 Comparaison des méthodes	18
3.4.2 Choix de la méthode pour les stations du Québec	19
CHAPITRE 4 ANALYSE EXPLORATOIRE DES DONNÉES	20

CHAPITRE 5	MÉTHODOLOGIE	27
5.1	Modélisation de l'erreur des débits mesurés	27
5.2	Ajustement de la fonction de puissance à un segment	27
5.2.1	Estimation des paramètres	27
5.2.2	Estimation de l'incertitude des paramètres des courbes de tarage	28
5.2.3	Estimation de l'incertitude des débits prédits	29
5.3	Ajustement de la fonction de puissance à deux segments	31
5.3.1	Estimation des paramètres	31
5.3.2	Estimation de l'incertitude des paramètres	32
5.3.3	Estimation de l'incertitude des débits prédits	33
5.4	Sélection du nombre de segments	34
CHAPITRE 6	RÉSULTATS ET DISCUSSION	35
6.1	Sélection du nombre de composantes de la courbe de tarage	35
6.2	Incertitudes des paramètres de la courbe de tarage	37
6.3	Résultats des incertitudes sur les débits estimés par les courbes de tarage	40
6.4	Discussion	43
6.4.1	Évaluation du modèle de calcul des incertitudes correspondantes aux débits prédits par les courbes de tarage	43
6.4.2	Évaluation du modèle de calcul des incertitudes correspondantes aux paramètres des courbes de tarage	46
CHAPITRE 7	CONCLUSION ET RECOMMANDATIONS	49
7.1	Synthèse des travaux	49
7.2	Limitations de la solution proposée et améliorations futures	50
RÉFÉRENCES	52
ANNEXES	56

LISTE DES TABLEAUX

Tableau 2.1	Valeurs moyennes du paramètre c , donnés par les formules hydrauliques classiques pour les types géométries les plus courantes [1]	7
Tableau 2.2	Évidence du modèle ayant le plus petit BIC en fonction de la différence.	12
Tableau 6.1	BIC obtenus pour les ajustement de la courbe de tarage à un segment et à deux segment pour les 15 stations hydrométriques sélectionnées. Le plus petit BIC pour chacune des stations est en caractère gras. . .	35
Tableau 6.2	Paramètres des courbes de tarage des stations de la première catégorie. Les intervalles de confiance pour les paramètres à 95% sont obtenus par la méthode bootstrap	41
Tableau A.1	Tableau récapitulatif des caractéristiques des stations de la première catégorie et les composantes d'incertitudes associées à cette catégorie	56
Tableau A.2	Tableau récapitulatif des caractéristiques des stations de la deuxième catégorie et les composantes d'incertitudes associées à cette catégorie	56
Tableau A.3	Tableau récapitulatif des caractéristiques des stations de la troisième catégorie et les composantes d'incertitudes associées à cette catégorie	56
Tableau A.4	Tableau récapitulatif des caractéristiques des stations de la quatrième catégorie et les composantes d'incertitudes associées à cette catégorie	56
Tableau A.5	Tableau récapitulatif des caractéristiques des stations de la cinquième catégorie et les composantes d'incertitudes associées à cette catégorie	57
Tableau B.1	Paramètres des courbes de tarage des stations de la deuxième catégorie. Les intervalles de confiance pour les paramètres à 95% sont obtenus par la méthode bootstrap. La courbe de tarage de la station Yamaska Nord (030340) est ajustée à un seul segment.	58
Tableau B.2	Paramètres des courbes de tarage des stations de la troisième catégorie. Les intervalles de confiance pour les paramètres à 95% sont obtenus par la méthode bootstrap.	58
Tableau B.3	Paramètres des courbes de tarage des stations de la quatrième catégorie. Les intervalles de confiance pour les paramètres à 95% sont obtenus par la méthode bootstrap. Les courbes de tarage des stations Nicolet (030120) et Ha! Ha! (060601) sont ajustées à un seul segment. . . .	58
Tableau B.4	Paramètres des courbes de tarage des stations de la cinquième catégorie. Les intervalles de confiance pour les paramètres à 95% sont obtenus par la méthode bootstrap.	59

LISTE DES FIGURES

Figure 1.1	Schéma illustrant la méthode vitesse surface utilisée par les courantomètres : des verticales sont fixées le long de la section de contrôle, la vitesse et la profondeur sont échantillonnées discrètement à des positions bien définies sur les verticales et illustrée par les points (figure tirée de [2])	2
Figure 1.2	Courbe de tarage de la station Sainte-Anne (050408) : les points en bleu constituent les couples hauteur-débit obtenus par jaugeage, la courbe en rouge présente le traçage de la courbe de tarage	3
Figure 2.1	Géométrie de différentes sections de contrôle.	7
Figure 2.2	Exemple de l'influence de changement d'un contrôle hydraulique sur l'ordre de grandeur du débit. Le schéma de droite montre des longueurs horizontales qui sont les profils en long des lignes de courant pour différentes hauteurs, les valeurs des débits pour ces profils sont illustrées dans le schéma de gauche de la section de contrôle (Schéma tiré de [1], modifié)	8
Figure 2.3	L'influence de l'amplitude de la profondeur de l'eau et de la géométrie de la section de contrôle sur la relation $h - q$ (figure inspirée de l'étude de [1])	9
Figure 2.4	Principe de la méthode bootstrap non paramétrique, les données présentées sont tirées des données de jaugeages de la station Sainte-Anne (figure inspirée de [3])	13
Figure 4.1	Données brutes de la station Sainte-Anne (021407)	20
Figure 4.2	Jaugeages des trois stations considérées pour représenter la catégorie 1.	22
Figure 4.3	Jaugeages des trois stations considérées pour représenter la catégorie 2.	23
Figure 4.4	Jaugeages des trois stations considérées pour représenter la catégorie 3.	24
Figure 4.5	Jaugeages des trois stations considérées pour représenter la catégorie 4.	25
Figure 4.6	Jaugeages des trois stations considérées pour représenter la catégorie 5.	26
Figure 5.1	La relation hauteur-débit correspondante à la station Sainte-Anne (050408) (ajustée à un seul segment)	29
Figure 5.2	L'incertitude du débit prédit par la courbe de tarage pour la station Sainte-Anne (050408)	31
Figure 5.3	La relation hauteur-débit correspondante à la station Sainte-Anne (050408) (ajustée à deux segments)	32

Figure 5.4	L'incertitude du débit prédit par la courbe de tarage pour la station Sainte-Anne (050408) (ajustée à deux segments)	33
Figure 6.1	Courbe de tarage ajustée pour la station Jacques-Cartier (050801).	36
Figure 6.2	Allure des incertitudes sur la courbe de tarage de la station Sainte-Anne (050408), la courbe rouge illustre la courbe de tarage ajustée sur les jaugeages et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage	42
Figure 6.3	Comparaison d'un échantillon bootstrap à son échantillon original de la station Yamaska Nord (030309)	47
Figure C.1	Allure des incertitudes sur la courbe de tarage de la station Jacques-Cartier (050801), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.	60
Figure C.2	Allure des incertitudes sur la courbe de tarage de la station Harricana (080101), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.	61
Figure C.3	Allure des incertitudes sur la courbe de tarage de la station Yamaska Nord (030340), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.	62
Figure C.4	Allure des incertitudes sur la courbe de tarage de la station Saint-Louis (040212), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.	63
Figure C.5	Allure des incertitudes sur la courbe de tarage de la station Des Escoumins (070204), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.	64
Figure C.6	Allure des incertitudes sur la courbe de tarage de la station Sainte-Anne (021407), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.	65

Figure C.7	Allure des incertitudes sur la courbe de tarage de la station Dauphine (120201), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.	66
Figure C.8	Allure des incertitudes sur la courbe de tarage de la station Boyer Sud (023002), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.	67
Figure C.9	Allure des incertitudes sur la courbe de tarage de la station Nicolet (030120), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.	68
Figure C.10	Allure des incertitudes sur la courbe de tarage de la station Capitachouane (043030), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.	69
Figure C.11	Allure des incertitudes sur la courbe de tarage de la station Ha! Ha! (060601), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.	70
Figure C.12	Allure des incertitudes sur la courbe de tarage de la station Yamaska Nord (030309), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.	71
Figure C.13	Allure des incertitudes sur la courbe de tarage de la station Du Chêne (046709), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.	72
Figure C.14	Allure des incertitudes sur la courbe de tarage de la station Suffolk (040409), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.	73

LISTE DES SIGLES ET ABRÉVIATIONS

ISO	<i>International Organization for Standardization</i>
WMO	<i>World Meteorological Organization</i>
MELCC	Ministère de l'Environnement et de la Lutte contre les changements climatiques
SSE	<i>Sum of squares error</i>
BIC	<i>Bayesian information criterion</i>
MCMC	<i>Markov chain Monte Carlo</i>
BaRatin	<i>Bayesian Rating curve</i>
LOWess	<i>LOcally Estimated Scatterplot Smoothing</i>

LISTE DES ANNEXES

Annexe A	Caractéristiques des stations hydrométriques étudiées	56
Annexe B	Résultats des paramètres des courbes de tarage	58
Annexe C	Résultats des incertitudes correspondantes aux courbes de tarage . .	60

CHAPITRE 1 INTRODUCTION

La connaissance du débit des cours d'eau est très importante. Par exemple, en hydrologie, la modélisation hydrologique et l'analyse des phénomènes extrêmes tels que l'inondation et la sécheresse sont basées sur les données du débit. De ce fait, les données du débit constituent l'élément clé dans la cartographie des plaines inondables [4–6]. Elles constituent alors implicitement l'information essentielle pour la gestion des risques naturels et pour la gestion des ouvrages de protection des milieux habités. En outre, les données du débit représentent un élément important dans le domaine de la construction hydraulique tel que le dimensionnement des ouvrages hydrauliques. Elles présentent également un besoin nécessaire dans le domaine de gestion de l'eau. En effet, elles constituent la base du suivie des ressources en eau notamment la base de l'évaluation de sa qualité et sa quantité. Par conséquent, ces données constituent une nécessité pour la planification des ressources en eau [4, 5].

Toutefois, la mesure des débits en temps réel est impossible. La procédure pour mesurer le débit est loin d'être simple et elle présente un exercice coûteux nécessitant l'envoi d'équipe sur le terrain. Cependant, la mesure des hauteurs d'eau des cours d'eau en temps réel est beaucoup plus simple, directe et peut être effectuée d'une façon continue à grande échelle et à coût raisonnable. Cet avantage est exploité dans les appareils de mesure qui mesurent généralement la hauteur puis estiment le débit selon des méthodes de conversion hydraulique. Par exemple, les appareils de mesure des courantomètres tels que le moulinet à godet et le vélocimètre acoustique (ADV) estiment le débit par extrapolation du champ de vitesse (cette méthode est décrite par [7]). En effet, à une section transversale de la rivière appelée section de contrôle, le champ de vitesses de l'écoulement est estimé par une intégration discrète. Cette intégration se base sur des verticales fixées le long de la section de contrôle tel que la hauteur de chaque verticale présente la profondeur de l'eau mesurée. Cela est illustré à la figure 1.1. Ensuite, les vitesses partielles sont estimées sur chaque verticale. Ces vitesses permettent de calculer les débits partiels sur les sous sections associées aux verticales. Finalement, le débit total de la section de contrôle est déduit à partir de la somme des débits partiels selon la relation :

$$Q = \sum_{i=1}^N a_i v_i \quad (1.1)$$

où a_i est la surface de la sous-section en m^2 et v_i est la vitesse moyenne de la sous-section en m/s . Il existe également d'autres techniques de mesures telles que celle du profileur acoustique Doppler (ADCP) qui se base aussi sur la méthode d'extrapolation du champ de vitesse (la méthode est décrite par [8], voir la figure 1.1).

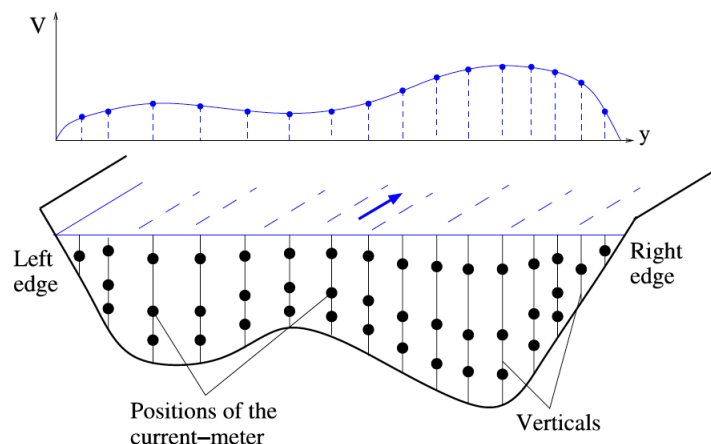


Figure 1.1 Schéma illustrant la méthode vitesse surface utilisée par les courantomètres : des verticales sont fixées le long de la section de contrôle, la vitesse et la profondeur sont échantillonnées discrètement à des positions bien définies sur les verticales et illustrée par les points (figure tirée de [2])

Par ailleurs, il est nécessaire d'avoir un moyen simple pour estimer le débit. Il est alors possible d'estimer le débit en le reliant à la hauteur d'eau. Cette relation entre la hauteur d'eau et le débit est appelée courbe de tarage. Cette relation a été abordée dans de nombreuses études telles que celles de [9–13]. Ces études ont proposé principalement des modèles empiriques tels que la fonction de puissance qui exprime cette relation en se basant sur des connaissances physiques. En effet, la majorité de ces études se sont inspirées des équations de Manning-Strickler [14] de l'écoulement en régime permanent. En hydraulique, la courbe de tarage est la relation entre la hauteur d'eau et le débit à une section de contrôle. Cette courbe modélise à la fois la gamme des débits d'étiage et des débits de crue ainsi que le régime permanent décrit par la gamme des débits moyens. Par conséquent, elle produit une estimation continue du débit. Cela attribue un rôle important à ces courbes dans la résolution du problème de l'absence des mesures du débit en temps réel [9, 12, 15].

Une étape importante qui doit être effectuée avant la construction des courbes de tarage est la mesure des couples hauteur-débit (par les différentes techniques de mesures mentionnées précédemment). Ces couples hauteur-débit sont appelés jaugeages et la procédure de mesure est appelée jauge [1]. En ce qui a trait de l'importance de cette étape, les données de jaugeages sont les seules à être utilisées pour ajuster la courbe de tarage. La figure 1.2 illustre un exemple de courbe de tarage de la station Sainte-Anne (050408) où les points en bleu constituent les couples hauteur-débit obtenus par jaugeage et la courbe en rouge présente le traçage de la courbe de tarage. Par conséquent, des mesures régulières des données de jaugeages par les

stations hydrométriques sont nécessaires non seulement pour la construction de la courbe mais aussi pour suivre son évolution [12, 13].

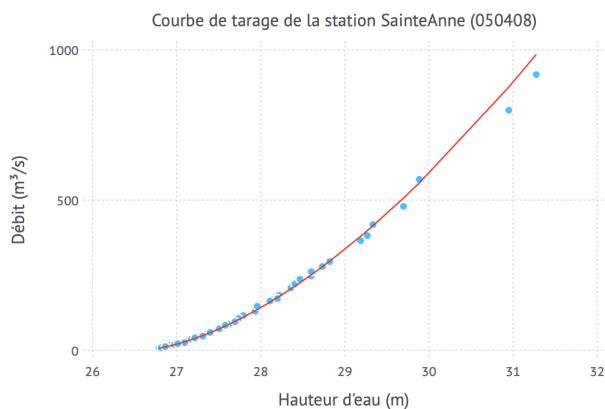


Figure 1.2 Courbe de tarage de la station Sainte-Anne (050408) : les points en bleu constituent les couples hauteur-débit obtenus par jaugeage, la courbe en rouge présente le traçage de la courbe de tarage

Toutefois, ces courbes sont sujettes à de nombreuses sources d'incertitudes notamment dues à des contraintes imposées sur la qualité des services des stations hydrométriques [4]. Ici, il est important de distinguer le terme incertitude et le terme erreur. En effet, l'erreur présente la différence entre la valeur mesurée et la vraie valeur cherchée. Le vrai débit n'est jamais parfaitement connu. Par conséquent, l'erreur est modélisée avec une variable aléatoire. Cette modélisation est effectuée par une estimation de la distribution des erreurs en reproduisant plusieurs fois une mesure donnée tel que la vraie valeur aura une certaine probabilité d'être trouvée [16].

De façon générale, quatre types d'incertitude associées aux courbes de tarage sont introduites. Premièrement, les incertitudes associées à la mesure des couples hauteur-débit servant à construire la courbe de tarage (noté ε_1). En effet, la mesure de la hauteur d'eau est sujette à plusieurs contraintes telles que la précision du niveau de référence, l'erreur de lecture de la hauteur de l'eau, l'erreur systématique de l'instrument de mesure ou l'erreur d'enregistrement des données [7]. La précision des débits jaugés dépend de la technique de mesure employée. Par exemple, lorsque le courantomètre est utilisé, la précision dépend du nombre de verticales et du nombre de mesures de vitesse (comme illustrée à la figure 1.1). De plus, il existe une zone non mesurée dans la section de contrôle en raison de la limitation des techniques de mesure de la profondeur de l'eau aux bords de la section de contrôle. Par conséquent, le débit est estimé par une extrapolation à laquelle correspond une certaine erreur [2]. De ces conditions et de ces limitations de mesure peuvent résulter une incertitude sur les débits

jaugés de 7 – 15% [16]. Deuxièmement, les incertitudes peuvent être aussi causées par des erreurs d’ajustement de la courbe de tarage. Cette source d’erreur est notée ε_2 . En effet, les courbes de tarage sont construites généralement par une minimisation de l’écart entre la courbe et les jaugeages en se basant sur la relation hydraulique reliant la hauteur et le débit. Cette optimisation constitue une approximation incertaine en elle-même. Par conséquent, la forme de la courbe résultante est aussi incertaine [1]; la fonction n’est qu’une modélisation empirique de la relation hauteur-débit. Troisièmement, les changements géomorphologiques de la section de contrôle du lit de la rivière et l’instabilité des conditions hydrauliques peuvent causer des incertitudes importantes au niveau de la courbe de tarage. Elles peuvent produire une dispersion importante au niveau des données et un vieillissement de la courbe même. Cette source d’incertitude est notée ε_3 [1]. La dernière source d’incertitude, notée ε_4 , provient de la forme paramétrique de la relation hauteur-débit utilisée. Le modèle utilisé n’est qu’une approximation de la vraie relation relation qui est inconnue [17].

La présence de toutes ces sources d’incertitude sur la courbe de tarage affecte la qualité des débits estimés avec ces courbes. Par ailleurs, comme mentionné au début de ce chapitre, les données du débit présentent un ingrédient très important dans plusieurs études. Ces incertitudes peuvent alors engendrer des doutes envers les conclusions rapportées aux résultats obtenus. Par conséquent, il est particulièrement crucial de quantifier les incertitudes relatives aux courbes de tarage afin d’assurer la qualité et la fiabilité des résultats et des conclusions obtenues [18]. Par exemple, le processus décisionnel dans le domaine de la cartographie des zones inondables pourrait être plus clair grâce à une meilleur évaluation des risques en comptabilisant les incertitudes correspondantes aux courbes de tarage. L’estimation du risque pourrait avoir une amplification que ce soit au niveau des zones qui peuvent être inondées ou bien au niveau de l’ampleur de l’aléa et sa période de manifestation. Un autre avantage très important de la comptabilisation des incertitudes correspondantes aux courbes de tarage consiste à une diminution des coûts supplémentaires dans plusieurs domaines tel que les pénalités gouvernementales, les dépenses supplémentaires et la mauvaise gestion environnementale [19].

Notamment, le Québec a vécu des inondations exceptionnelles au printemps 2017. Ces événements ont attiré une attention particulière sur la nécessité d’une mise à jour des cartes des zones inondables du Québec méridional. Par conséquent, un grand projet appelé “INFO-Crue” a été lancé en collaboration avec plusieurs partenariats. Il prévoit créer un système de prévision et d’alerte des zones à risque d’inondation sur un horizon de quelques jours en s’appuyant sur des cartes de zones inondables qui tiennent en compte de l’effet des changements climatiques. Étant donné que la cartographie des zones inondables se base essentiellement sur les données du débit et de la hauteur d’eau, les courbes de tarage des stations hydrométriques

du Québec présentent alors un besoin nécessaire pour ce projet. Ces courbes sont estimées par la direction de l'expertise hydrique et atmosphérique du Ministère de l'Environnement et de la Lutte contre les changements climatiques (MELCC) en se basant sur des méthodes statistiques [20]. Le problème dans ces méthodes d'estimation et auquel ce projet fait face est que dans le cas d'une inondation, la précision de la délimitation des surfaces inondées est affectée non seulement par les incertitudes de mesure de la hauteur d'eau mais aussi par la fonction choisie pour relier le débit et la hauteur d'eau ainsi que l'extrapolation de la courbe de tarage [18]. Le but de ce projet consiste à estimer des courbes de tarage et leurs incertitudes. Ces estimations pourront être intégrées dans la cartographie des zones inondables. Le travail de ce mémoire s'inscrit dans le cadre du projet Info-Crue. Il a pour objectif général de développer une méthode d'estimation des courbes de tarage aux stations hydrométriques du Québec méridional permettant une prise en compte rigoureuse des incertitudes. Afin d'atteindre cet objectif, 15 stations représentatives des différentes caractéristiques de la grande base de données ont été choisies et trois objectifs spécifiques ont été soulignés. Le premier objectif consiste à développer une méthode d'ajustement des courbes de tarage pour les tronçons du Québec méridional, qui constitue en elle-même, une base solide permettant de calculer les incertitudes correspondantes à ces courbes. Le deuxième objectif consiste à développer une méthode qui permet d'estimer les incertitudes correspondantes aux prédictions du débit par la courbe de tarage. Le dernier objectif consiste à développer une méthode qui permet d'estimer les incertitudes de tous les paramètres de la courbe de tarage.

Ce mémoire a été structuré de la façon suivante. Le chapitre 2 concerne le cadre théorique des méthodes utilisées. Il présente notamment la base hydraulique des courbes de tarage et les méthodes d'optimisation utilisées pour estimer les courbes de tarage. Le chapitre 3 présente une revue de littérature qui décrit les différentes méthodes de calcul des incertitudes correspondantes aux courbes de tarage adoptées dans la littérature et qui illustre une analyse critique des critères de sélection entre ces méthodes selon la situation étudiée. Le chapitre 4 consiste à une description des données des 15 stations hydrométriques choisies afin d'identifier la méthode qui assure une modélisation de courbe de tarage. Le chapitre 5 présente la méthode utilisée pour estimer les courbes de tarages ainsi que les incertitudes correspondantes en se basant sur l'analyse des données et sur les éléments théoriques présentée dans le chapitre 1. Le chapitre 6 présente les résultats obtenus pour chaque station hydrométrique étudiée ainsi que la discussion des résultats obtenus. Le dernier chapitre présente une synthèse du travail effectué et des buts atteints dans ce projet et il présente les limites et les points à améliorer.

CHAPITRE 2 CADRE THÉORIQUE

Ce chapitre présente les fondements théoriques de ce mémoire. Il présente dans un premier temps les bases hydrauliques de la relation hauteur-débit d'un cours d'eau. Dans un deuxième temps, les méthodes statistiques nécessaires pour l'estimation des courbes de tarage et de leur incertitude sont présentées.

2.1 Fondement hydraulique de la relation hauteur-débit

Tel que mentionné précédemment dans le chapitre 1, la mesure continue du débit d'un cours d'eau est impossible. La mesure de la hauteur d'un cours d'eau est cependant facile à obtenir de façon continue et automatique. Grâce à une relation hydraulique, il est possible d'estimer le débit (q en m^3/s) correspondant à une hauteur du niveau d'eau donné (h en mètres).

La relation hauteur-débit à une section de contrôle donnée d'un canal régulier dépend de l'hydraulicité du canal et de la géométrie de la section de contrôle. L'hydraulicité dépend, entre autres, de la pente et de la rugosité du canal. Dans le cas d'un régime uniforme (lorsque le tirant du cours d'eau est parallèle au lit de la rivière), les équations de Manning-Strickler [14], tirées de l'hydraulique, indiquent que la relation hauteur-débit possède la forme suivante :

$$q(h) = K_s S(h) R_h^{\frac{2}{3}}(h) I^{\frac{1}{2}} \quad (2.1)$$

où K_s représente le coefficient de Strickler, $S(h)$ est la section mouillée (en m^2), $R_h(h)$ est le rayon hydraulique (en m) et I est la pente du canal (en m/m). Cette relation illustre la dépendance de la relation hauteur-débit aux conditions hydrauliques du canal et à la géométrie de celui-ci.

Dans la nature, les cours d'eau ne correspondent pas à un canal régulier où l'eau s'écoule en régime permanent. Plusieurs approches empiriques basées sur la relation hydraulique (2.1) ont alors été alors proposées pour estimer la forme de la relation hauteur-débit naturelle [1,4]. La fonction de puissance suivante est la relation empirique la plus utilisée [9, 12, 15] :

$$q(h) = \begin{cases} a(h - b)^c & \text{pour } h > b, \\ 0 & \text{sinon ;} \end{cases} \quad (2.2)$$

où a présente un coefficient d'échelle dépendant des caractéristiques de la section de contrôle tel que $a > 0$. b présente le niveau de référence pour lequel le débit est nul tel que $b \in \mathbb{R}$.

c est fonction de la géométrie de la section [1] tel que $c \in \mathbb{R}$. Les valeurs les plus courantes de ce paramètre pour les différentes géométries de sections régulières illustrées à la figure 2.1 sont compilées dans le tableau 2.1.

Tableau 2.1 Valeurs moyennes du paramètre c , donnés par les formules hydrauliques classiques pour les types géométries les plus courantes [1]

Type de section de contrôle	Hypothèse sur la section	c
Canal (généralement au laboratoire)	Rectangulaire large régime stationnaire uniforme	5/3
Rectangulaire	Perpendiculaire au débit moyen, pas de remous	3/2
Triangulaire	Perpendiculaire au débit moyen, pas de remous	5/2
Orifice	Perpendiculaire au débit moyen, pas de remous	1/2

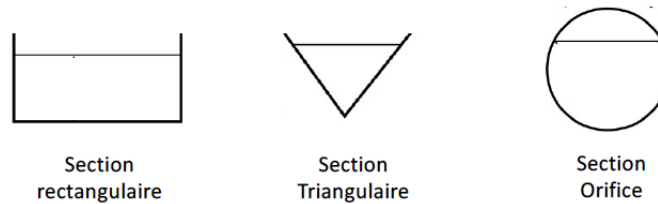


Figure 2.1 Géométrie de différentes sections de contrôle.

En rivière naturelle, la géométrie n'est évidemment pas régulière. La figure 2.2 schématise une section de contrôle d'un cours d'eau naturel. Cette section possède plusieurs contrôles hydrauliques qui peuvent influencer la relation hauteur-débit. Pour des hauteurs très basses, les cailloux et sédiments au fond peuvent influencer le débit. Pour des hauteurs moyennes, l'imperfection du fond ne joue plus un rôle aussi dominant dans la relation hauteur-débit. À partir d'un certain seuil, généralement appelé le lit majeur de la rivière, la largeur de la section de contrôle s'élargit et la relation hauteur-débit est modifiée [1]. Le lit majeur est défini par la plaine inondable de la rivière où se produisent les crues importantes où l'écoulement a une emprise maximale de débordement [21, 22]. Le lit mineur d'une rivière est défini par la partie du lit ordinaire de la rivière délimitée par ses berges. L'écoulement s'effectue la plupart du temps dans le lit mineur et sans déborder de la rivière. La figure 2.3 illustre le changement de la relation hauteur-débit au passage du lit mineur au lit majeur. Dans ce cas, la relation

hauteur-débit peut être décomposée de la façon suivante [23] :

$$q(h) = \begin{cases} a_1(h - b_1)^{c_1} & \text{si } b_1 < h \leq k \\ a_2(h - b_2)^{c_2} & \text{si } h > k \text{ et } h > b_2 \\ 0 & \text{sinon,} \end{cases} \quad (2.3)$$

où k dénote la hauteur de transition entre le lit mineur et le lit majeur, (a_1, b_1, c_1) correspondent aux paramètres de la courbe de tarage pour le lit mineur et (a_2, b_2, c_2) correspondent aux paramètres de la courbe de tarage pour le lit majeur. La plupart du temps, deux segments sont suffisants pour modéliser la relation hauteur-débit d'un cours d'eau [23].

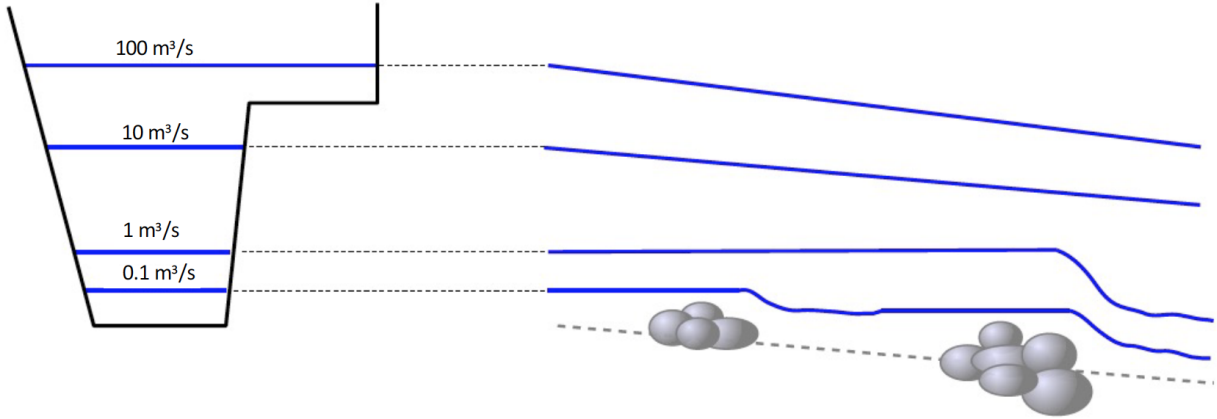


Figure 2.2 Exemple de l'influence de changement d'un contrôle hydraulique sur l'ordre de grandeur du débit. Le schéma de droite montre des longueurs horizontales qui sont les profils en long des lignes de courant pour différentes hauteurs, les valeurs des débits pour ces profils sont illustrées dans le schéma de gauche de la section de contrôle (Schéma tiré de [1], modifié)

De façon générale, les paramètres de la relation hauteur-débit sont estimés à partir de jauges du cours d'eau à une section donnée. Un jaugeage correspond à un couple mesuré de hauteur d'eau et de débit à une section spécialement choisie. La section doit être de la forme la plus régulière possible pour s'assurer d'une bonne adéquation avec la relation hauteur-débit dérivée de l'hydraulique [1].

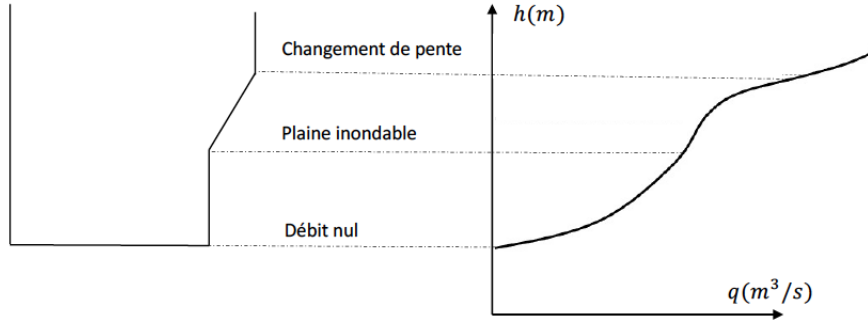


Figure 2.3 L'influence de l'amplitude de la profondeur de l'eau et de la géométrie de la section de contrôle sur la relation $h - q$ (figure inspirée de l'étude de [1])

2.2 Incertitude sur la relation hauteur-débit

L'incertitude des prédictions de débits obtenues par la courbe de tarage peut être modélisée de la façon suivante :

$$Q(h) = q(h) + \varepsilon,$$

où $Q(h)$ est la variable aléatoire modélisant le débit à la hauteur h , $q(h)$ est l'expression de la courbe de tarage donnée par l'équation 2.2 ou 2.3 dans le cas de l'utilisation de la relation hydraulique et ε est l'erreur aléatoire. Le prédiction est donc la somme du signal donné par la courbe de tarage $q(h)$ et du bruit introduit par l'erreur aléatoire ε .

L'erreur aléatoire ε peut être décomposée en plusieurs composantes. En s'inspirant de [1,18], l'erreur sur le débit peut être modélisée par la somme des erreurs suivantes :

$$\varepsilon = \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4 \quad (2.4)$$

où

ε_1 : erreur associée à la mesure des couples niveaux d'eau et débits servant à construire la courbe de tarage ;

ε_2 : erreur associée à l'estimation des paramètres de la courbe de tarage ;

ε_3 : erreur associée aux changements de la géométrie de la section de contrôle ;

ε_4 : erreur associée à l'utilisation du modèle de la relation hauteur-débit.

La première source d'erreur correspond aux erreurs sur les jaugeages. Généralement, on suppose que l'erreur sur la mesure de la hauteur d'eau est négligeable par rapport à l'erreur sur

la mesure du débit [24]. La procédure pour mesurer le débit est assez fastidieuse et est sujette à une imprécision qui peut dépendre de la gamme du débit jaugé [25]. La deuxième source d'erreur est reliée à l'incertitude d'estimation des paramètres de la courbe de tarage basée sur un échantillon fini. C'est une source d'erreur statistique reliée à l'échantillonnage. La troisième source d'erreur est reliée aux changements géomorphologiques du lit de la rivière. Ces changements peuvent se produire, par exemple, à la suite d'une crue qui déplace des débris et des rochers à la section de contrôle [26]. La quatrième et dernière source d'incertitude est reliée à la modélisation de la relation hauteur-débit. Dans le cas de l'utilisation de la relation de puissance dérivée de l'hydraulique, il est possible que cette relation ne représente pas tout à fait la vraie relation entre la hauteur et le débit et c'est ce que modélise cette source d'incertitude [1, 18].

2.3 Régression linéaire

La régression linéaire est utilisée dans ce mémoire pour estimer les paramètres de la courbe de tarage. Il revient ici de présenter les rudiments de la régression.

Soit Y une variable d'intérêt et X , une variable explicative pour Y . La régression linéaire simple suppose qu'il existe un lien linéaire entre l'espérance de la variable Y et la variable explicative X [27] :

$$\mathbb{E}(Y|X = x) = \beta_0 + \beta_1 x,$$

où β_0 et β_1 sont respectivement l'ordonnée à l'origine et la pente de la droite de régression. Cette hypothèse implique le modèle statistique suivant :

$$(Y|X = x) = \beta_0 + \beta_1 x + \varepsilon,$$

où ε est l'erreur aléatoire d'espérance nulle.

De façon classique, les paramètres (β_0, β_1) sont estimés à partir d'un échantillon aléatoire du couple (X, Y) de taille n , $\{(x_i, y_i), 1 \leq i \leq n\}$, en minimisant la somme du carré des erreurs entre la droite et les observations. Les estimations correspondantes sont les suivantes :

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (2.5)$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x} \quad (2.6)$$

où $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ et $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$.

Si on suppose que la variance de l'erreur est constante tout le long de la droite de régression,

i.e. $\text{Var}(Y|X = x) = \sigma^2$, une estimation de celle-ci est donnée par les expressions suivantes :

$$\begin{aligned}\hat{\sigma}^2 &= \frac{1}{n-2} \sum_{i=1}^n (y_i - \beta_0 - \beta_1 x_i)^2, \\ &= \frac{1}{n-2} \sum_{i=1}^n e_i^2, \\ &= \frac{1}{n-2} \text{SSE};\end{aligned}\tag{2.7}$$

où SSE correspond à la somme du carré des erreurs ($e_i : 1 \leq i \leq n$) entre la droite de régression et les observations.

Si on suppose que les observations sont indépendantes, il est possible de calculer la covariance du vecteur des estimateurs $\boldsymbol{\beta} = (\hat{\beta}_0, \hat{\beta}_1)$ [27] :

$$\text{Var}(\boldsymbol{\beta}) = \hat{\sigma}^2 (X^\top X)^{-1},$$

où X est la matrice de structure composée d'une première colonne de "1" et la deuxième colonne correspond aux valeurs de la variable explicative pour chacune des observations. Il est également possible de calculer la covariance de la prédiction $\hat{Y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_0$ correspondante à la valeur x_0 de la variable explicative [27] :

$$\text{Var}(\hat{Y}_0) = \begin{bmatrix} 1 & x_0 \end{bmatrix} \hat{\sigma}^2 (X^\top X)^{-1} \begin{bmatrix} 1 \\ x_0 \end{bmatrix}.\tag{2.8}$$

Si on suppose en plus que les erreurs sont distribuées selon la loi normale, alors les intervalles de confiance de niveau $(1 - \alpha)$ pour les paramètres β_0 et β_1 sont donnés par [27] :

$$\hat{\beta}_j \pm t_{n-2}(1 - \alpha/2) \sqrt{v_{j+1} \hat{\sigma}^2}$$

où $t_{n-2}(1 - \alpha/2)$ correspond au quantile d'ordre $(1 - \alpha/2)$ de la loi de Student à $(n - 2)$ degrés de liberté et où v_j est le j^e élément de la diagonale de la matrice $(X^\top X)^{-1}$. Un intervalle de confiance de niveau $(1 - \alpha)$ peut également être obtenu sur la prédiction \hat{Y}_0 associée à $\mathbf{x}_0^\top = \begin{bmatrix} 1 & x_0 \end{bmatrix}$ [27] :

$$\hat{Y}_0 \pm t_{n-2}(1 - \alpha/2) \sqrt{\hat{\sigma}^2 \{1 + \mathbf{x}_0^\top (X^\top X)^{-1} \mathbf{x}_0\}}.\tag{2.9}$$

Lorsque plusieurs modèles de régression sont possibles, le SSE peut donner une idée sur le modèle optimal, soit celui qui possède le plus petit SSE , mais cela n'est possible que dans

le cas où les modèles ont le même nombre de paramètres. Plusieurs techniques sont alors proposées dans la littérature pour sélectionner le meilleur modèle entre des modèles avec différents nombres de paramètres. Le critère utilisé dans ce mémoire est le critère d'information bayésien (BIC) [28]. Le BIC constitue une mesure de la probabilité *a posteriori* que le modèle soit vrai. Il est construit sous une certaine configuration bayésienne et il est relié à plusieurs arguments asymptotiques. Il tend à avoir une petite valeur lorsque le modèle ajuste bien les données, par conséquent, le modèle optimal est associé au plus faible BIC.

Dans le cas d'un modèle des moindres carrés avec p paramètres (soit les degrés de liberté du modèle), le BIC peut s'écrire sous la forme suivante :

$$BIC = n \log(\hat{\sigma}^2) + p \log(n) \quad (2.10)$$

Le modèle de régression qui minimise le BIC est généralement favorisé.

[29] ont proposé une échelle pour interpréter la différence du BIC entre deux modèles. L'interprétation proposée en faveur du modèle avec le plus petit BIC en fonction de l'écart est résumée au tableau 2.2.

Tableau 2.2 Évidence du modèle ayant le plus petit BIC en fonction de la différence.

Différence du BIC	Évidence du modèle
$[0, 2)$	Faible
$[2, 6)$	Positive
$[6, 10)$	Forte
$[10, \infty)$	Décisive

2.4 La méthode de rééchantillonnage bootstrap

La méthode bootstrap [30, 31] est une méthode de rééchantillonnage permettant d'estimer l'incertitude reliée à l'échantillonnage sur les quantités d'intérêt. Cette méthode est largement utilisée en pratique car elle est simple et particulièrement efficace. Le principe de la méthode est succinctement décrit dans la présente section.

Soit un échantillon aléatoire observé $\mathbf{x} = (x_1, x_2, \dots, x_n)$ de taille n dont les observations sont indépendantes et identiquement distribuées selon la fonction de répartition F . Supposons que l'on souhaite calculer la variance échantillonnale de la statistique $\hat{\theta}$. Si la fonction de répartition F est connue, il est alors possible de générer d'autres échantillons de taille n selon cette

loi et d'analyser la distribution de la statistique $\hat{\theta}$ sur tous les échantillons générés. Cette méthode est appelée *bootstrap paramétrique* car le rééchantillonnage se fait à partir de la fonction de répartition connue F . Or, la plupart du temps, la fonction de répartition F est inconnue. Dans ce cas, la fonction de répartition empirique est utilisée pour le rééchantillonnage. La méthode est alors appelée *bootstrap non paramétrique*. Dans ce cas, les échantillons aléatoires de taille n sont générés en effectuant des tirages avec remise des observations de l'échantillon original. La taille des échantillons n doit être significativement plus grande que le nombre des paramètres pour pouvoir obtenir une approximation valide de la variance échantillonnale.

Une illustration de la méthode pour le rééchantillonnage non paramétrique de trois jaugeages de la rivière Sainte-Anne est présentée à la figure 2.4. Avec l'échantillon original, 1000 échantillons de taille 3 sont générés par tirage avec remise. Le premier échantillon généré contient la première observation et la deuxième observation deux fois. Cette procédure permettra d'étudier l'effet de l'échantillonnage sur l'estimation des paramètres de la courbe de tarage.

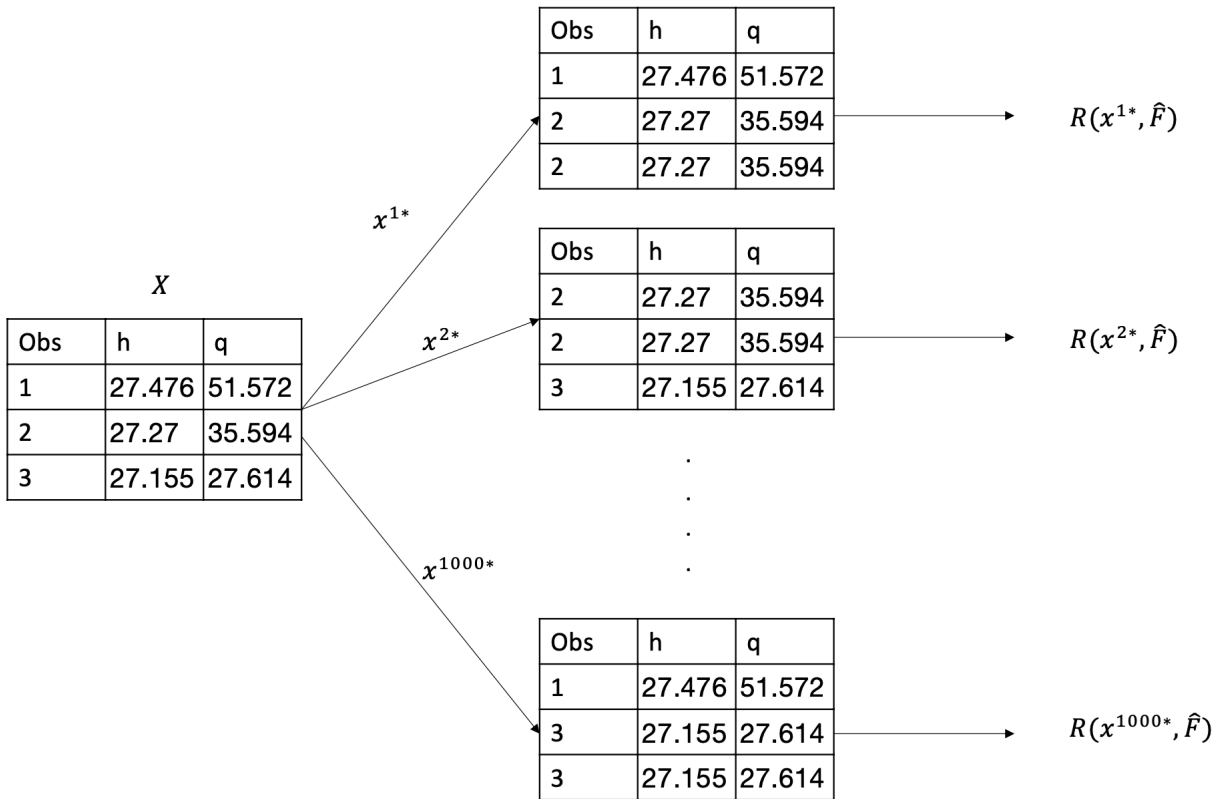


Figure 2.4 Principe de la méthode bootstrap non paramétrique, les données présentées sont tirées des données de jaugeages de la station Sainte-Anne (figure inspirée de [3])

CHAPITRE 3 REVUE DE LITTÉRATURE

Plusieurs techniques ont été proposées dans la littérature pour estimer les courbes de tarage ainsi que les incertitudes associées. Plusieurs approches se basent sur la fonction de puissance pour modéliser la relation hauteur-débit tandis que d'autres utilisent une formulation non paramétrique. Très peu d'études se sont concentrées sur l'incertitude temporelle reliée aux changements géomorphologiques. Le but de ce chapitre est de décrire brièvement les méthodes existantes et d'en faire une analyse critique pour l'estimation des courbes de tarage du Québec.

3.1 Estimation des courbes de tarage basées sur la relation hydraulique

3.1.1 Estimation par les moindres carrés

La méthode ISO/WMO a été développée par R. W. Herschy [32] et elle est décrite dans les documents [13, 15, 33]. Cette méthode se base sur l'approche de régression linéaire classique pour estimer la courbe de tarage et sur la variance résiduelle de la régression pour estimer l'incertitude sur la courbe. La méthode estime la courbe de tarage par la méthode des moindres carrés en se basant sur la fonction de puissance (éq. 2.2) dans l'espace logarithmique. Le modèle de ISO/WMO peut s'exprimer dans l'espace logarithmique comme suit :

$$\log q(h) = \log a + c \log(h - b) + \varepsilon^* \quad (3.1)$$

où ε^* est l'erreur aléatoire et est donné seulement par la composante ε_2 (éq.2.4) dans l'espace logarithmique. L'estimation de la courbe de tarage est effectuée sur plusieurs segments. À chaque segment, une équation de la forme 3.1 est introduite pour un échantillon de jaugeages (h_i, q_i) de taille n . Selon cette norme, la taille des échantillons par segment doit être au minimum 20 jaugeages pour avoir une estimation adéquate de l'incertitude. Ensuite, l'estimation de l'écart type résiduel de la régression linéaire sur chaque segment est obtenue par l'équation 2.7. Les estimations des $\log(\hat{q}_i)$ sont effectuées par l'équation 2.9. Cette méthode suppose que b est connu. Par conséquent, le degré de liberté par segment est donné par $p = 2$.

3.1.2 Estimation bayésienne

A. Petersen-Øverleir *et al.* [4] ont proposé une approche bayésienne pour estimer les courbes de tarage à un segment (éq. 2.2) pour plusieurs stations hydrométriques en Norvège. Le

modèle peut s'exprimer de la même façon que l'équation 3.1 dans l'espace logarithmique. Dans ce cas, l'erreur aléatoire ε est supposée distribuée selon la loi normale d'espérance nulle et de variance inconnue qui doit être estimée. L'erreur ε constitue la somme des trois composantes $\varepsilon_1, \varepsilon_3, \varepsilon_4$ de l'équation 2.4. L'erreur reliée aux changements géomorphologiques n'est pas considérée. Les lois *a priori* utilisées pour l'estimation bayésienne sont informatives et sont spécifiques aux stations hydrométriques à l'étude en Norvège. Un échantillon de la loi *a posteriori* est obtenu par une méthode Monte-Carlo par chaîne de Markov (*MCMC*). Une fois cet échantillon obtenu, des intervalles de confiance sur les paramètres de la courbe peuvent être obtenus et aussi sur les prédictions de débit.

J. Le Coz *et al.* [1] ont étendu l'approche proposée par A. Petersen-Øverleir *et al.* [4] à des courbes de tarage à plusieurs segments pour l'ajustement bayésien à plusieurs stations hydrométriques françaises. L'erreur liée aux changements de la géométrie de la section de contrôle n'est pas considérée. À l'instar de A. Petersen-Øverleir *et al.* [4], l'erreur combinée est modélisée par une loi normale de moyenne nulle et de variance inconnue dans l'espace logarithmique. J. Le Coz *et al.* [1] ont proposé des lois *a priori* informatives basées sur l'hydraulique en particulier pour les paramètres b et c . Ces lois *a priori* permettent de favoriser des valeurs de paramètres qui possèdent un sens physique au sens de l'hydraulique. Selon eux, les valeurs de c supérieures à 2,33 ne sont pas raisonnables d'un point de vue hydraulique et la loi *a priori* informative utilisée attribue une probabilité 0 à l'événement $c > 2,33$. Cette approche est nommée *BaRatin bayesian Rating curve*.

A. E. Sikorska *et al.* [34] ont développé une approche bayésienne pour estimer les courbes de tarage pour un bassin versant en Pologne (ces courbes se basent également sur l'équation 2.2 à plusieurs segments). Selon A. E. Sikorska *et al.* [34], l'intégration de la composante d'incertitude ε_2 (éq.2.4) dans le calcul de l'incertitude prédictive des hauteurs d'eau, est important pour une meilleure estimation des incertitudes. A. E. Sikorska *et al.* [34] ont attribué des lois *a priori* informatives aux paramètres de la courbe de tarage. Ces lois se basent sur les mesures des vitesses moyennes transversales et sur les hauteurs d'eau correspondantes. Par conséquent, les caractéristiques du site d'étude ne sont pas nécessaires pour la modélisation de ces lois. L'erreur correspondantes aux paramètres de la courbe de tarage est supposée comme une loi normale multivariée et la distribution *a posteriori* des paramètres est estimée par la méthode *MCMC*. Cette méthode a permis de considérer l'interaction entre les paramètres. Par conséquent, la corrélation entre les paramètres qui estime la composante ε_2 a été comptabilisée dans les lois *a priori*. Finalement, cette méthode néglige la composante ε_3 en considérant des échantillons de jaugeages de courte période.

H. McMillan et I. Westerberg [35] ont développé une approche bayésienne pour estimer les

incertitudes sur les débits prédits par les courbes de tarage. Cette étude diffère des études de J. Le Coz *et al.* [1] et de A. Petersen-Øverleir *et al.* [4] principalement par l'ajout de la composante ε_3 (éq.2.4) liée aux changements de la géométrie de la section de contrôle dans l'estimation des incertitudes. Pour ce fait, H. McMillan et I. Westerberg [35] se base sur une méthode nommée *Voting Point*. Cette méthode adapte des courbes de tarage à plusieurs sous-échantillons de jaugeages dans l'estimation de la fonction de vraisemblance. Ces courbes se basent sur la fonction de puissance (éq. 2.2) multi-segment avec des contraintes de continuités entre les points de ruptures et leurs estimations sont basées sur la méthode *MCMC*. La composante ε_3 est modélisée par les différentes formes de courbe de tarage produites d'un sous-échantillon à un autre. Les lois *a priori* des paramètres sont non informatives. Par conséquent cette méthode est indépendante des caractéristiques du lit de la rivière et elle a été appliquée sur quatre bassins versants de différentes caractéristiques. L'erreur de mesure ε_1 (éq.2.4) pour chaque point de jaugeages est modélisée pour les mesures du débit par une loi logistique ou bien par une loi normale et pour les mesures de la hauteur de l'eau par une distribution uniforme.

3.2 Courbes de tarage non basées sur l'hydraulique

Comme dans la section 2.1, la forme des relations (2.2) et (2.3) résultent d'une simplification empirique des formules hydrauliques élémentaires et elles ne sont pas toujours appropriées pour la section de contrôle. Selon J. D. Fenton [36, 37], cette forme est plus représentative dans le cas de déversoir d'une largeur importante, dans le cas d'un écoulement à profondeur convenable ou bien dans le cas d'un écoulement uniforme dans un canal rectangulaire. Dans ces cas, il est possible d'estimer la relation hauteur-débit par des approches non-paramétriques [36, 37].

Parmi les approches non-paramétriques la plus utilisée est la méthode de spline cubique par morceaux. Les splines cubiques présentent des polynômes de degré 3 qui interpolent une série successive de données. En effet, soit $n + 1$ paires de jaugeages, entre deux points de jaugeages successifs (h_{i-1}, q_{i-1}) et (h_i, q_i) , une bande est interpolée par une fonction spline cubique donnée par [36] :

$$s_i(h) = c_0 + c_1 (h - h_{i-1}) + c_2 (h - h_{i-1})^2 + c_3 (h - h_{i-1})^3 \quad \text{tel que } h_{i-1} \leq h \leq h_i$$

Les paramètres c_0 , c_1 , c_2 et c_3 sont estimés par des conditions de continuités entre deux splines aux points de raccordements. Pour ce faire, à chaque raccordement, trois conditions sont imposées : les valeurs des splines sont supposées égales $s_i(h_i) = s_{i+1}(h_i) = q_i$, les dérivées

premières sont supposées égales $s'_{i-1}(h_i) = s'_i(h)$ et les dérivées secondes sont supposées égales $s''_{i-1}(h_i) = s''_i(h_i)$. Ces conditions sont imposées sur les deux extrémités de la bande, par conséquent quatre contraintes sont imposées par spline cubique.

G. Coxon *et al.* [5] ont proposé d'utiliser la méthode non paramétrique de régression locale *LOESS LOcally Estimated Scatterplot Smoothing*, développée par W. S. Cleveland [38]. Cette méthode repose sur la méthode de lissage par les plus proches voisins. Étant donné que la méthode est non paramétrique, au moins 20 jaugeages sur plusieurs gammes de débits sont nécessaires pour l'estimation de la relation hauteur-débit [38]. G. Coxon *et al.* [5] ont appliqué la régression *LOESS* à plusieurs stations hydrométriques situées en Angleterre et au Pays de Galles. Les auteurs mentionnent notamment que la méthode *LOESS* ne permet pas d'extrapoler la relation hauteur-débit pour des débits supérieurs à ceux observés. Par conséquent, les estimations du débit au delà de l'étendue des débits est impossible, autant pour les étiage que les crues. De plus, dans les extrémités de la relation, l'estimation des incertitudes est très imprécise. Les auteurs proposent d'effectuer des simulations hydrauliques pour obtenir des estimations sur les débits en dehors de l'étendue des observations.

3.3 Traitement dynamique de la courbe de tarage

Le traitement de l'incertitude temporelle des courbes de tarage liée à la modification de la géométrie de la section de contrôle est peu étudié dans la littérature. J. Jalbert *et al.* [39] et T. Morlot *et al.* [18] ont proposé de quantifier l'incertitude temporelle en considérant l'historique de changements des courbes dans le temps. Le terme ε_3 (éq. 2.4) correspond à une variance de vieillissement de la courbe. Cette variance en fonction du temps est estimée avec un variogramme.

3.4 Analyse critique

Cette revue aborde des études qui traitent des méthodes d'estimation de courbes de tarage et des incertitudes correspondantes. Chacune de ces études a défini une méthode appliquée sur des sites de différentes sources de jaugeages et de différentes conditions hydrauliques et géomorphologiques. Par conséquent, une comparaison entre ces méthodes serait d'une grande utilité et ce, afin d'avoir une idée des avantages et des limites de chaque approche et de déduire un critère de sélection de méthode.

3.4.1 Comparaison des méthodes

Pour commencer, l'amplitude de l'incertitude des débits estimés est influencée par les principes de base des méthodes (paramétrique ou non paramétrique) et les informations *a priori* qui peuvent être intégrées dans les méthodes bayésiennes. De plus, l'extrapolation de la courbe de tarage au-delà des zones jaugées n'est possible que par les méthodes qui utilisent la fonction de puissance, ce qui n'est pas le cas pour la méthode non paramétrique *LOESS* [5]. Selon J. E. Kiang *et al.* [40], les méthodes non-paramétriques nécessitent un nombre minimal de jaugeages pour assurer la robustesse de calcul des incertitudes.

Ensuite, en ce qui concerne la différence entre la méthode ISO/WMO [13, 15, 33] et les méthodes bayésiennes, premièrement la méthode ISO est plus simple dans son implantation. En effet, l'algorithme des méthodes bayésiennes nécessite de fixer un choix adéquat pour les lois *a priori*. Le choix des lois *a priori* des études abordées dans cette revue sont divisées en deux types, celles qui se basent sur les conditions du site d'étude (proposées par [1, 4]) et celles qui se base seulement sur les jaugeages (proposée par A. E. Sikorska [34] et H. McMillan et I. Westerberg [35]). Par contre, la méthode ISO/WMO se basent seulement sur les jaugeages pour ajuster la courbe et sur l'information des conditions du site d'étude pour estimer le paramètre b . De plus, l'algorithme de la méthode des moindres carrés est beaucoup plus simple que celui de la méthode *MCMC*. Deuxièmement, la méthode ISO n'inclut pas l'incertitude de mesure dans son modèle ni la composante ε_3 . Elle se limite à la variance résiduelle de la régression. Cependant, les méthodes bayésiennes telles que celle proposée par J. Le Coz *et al.* [1] incluent l'incertitude de mesure dans la fonction de vraisemblance des paramètres et la méthode H. McMillan et I. Westerberg [35] inclut la composante ε_3 dans l'estimation des incertitudes. Troisièmement, la méthode ISO exige un nombre de jaugeages par segment ce qui n'est pas le cas dans les méthodes bayésiennes [40].

Par la suite, les méthodes bayésiennes utilisent toutes une distribution normale pour modéliser l'erreur dans l'espace logarithmique. Dans le cas de l'étude de A. Petersen-Øverleir *et al.* [4], la relation hauteur-débit est ajustée indépendamment pour un ou deux segments alors que dans BaRatin, le nombre de segments est déterminé dans l'algorithme bayésien. De plus, BaRatin utilise l'expertise des hydrologues dans la loi *a priori* pour déterminer le nombre de segments.

Puis, la méthode bayésienne de H. McMillan et I. Westerberg [35] est la seule qui inclut la composante ε_3 dans l'estimation totales des incertitudes parmi les études bayésiennes citées dans cette revue. Également, les deux études non bayésiennes J. Jalbert *et al.* [39] et T. Morlot *et al.* [18] incluent la composante ε_3 dans l'estimation des incertitudes. La différence entre ces deux études et l'étude de H. McMillan et I. Westerberg [35] et que H. McMillan et

I. Westerberg [35] se limite à modéliser cette composante par les différentes configurations possibles des courbes de tarage et elle ne considère pas la variabilité temporelle de la courbe. Par contre, les études de J. Jalbert *et al.* [39] et T. Morlot *et al.* [18] intègrent cette variabilité temporelle dans l'estimation de ε_3 par un varigramme.

Finalement, les méthodes qui ont insérée l'erreur de mesure ε_1 dans leurs approches sont les méthodes bayésiennes A. Petersen-Øverleir *et al.* [4], J. Le Coz *et al.* [1], A. E. Sikorska [34] et H. McMillan et I. Westerberg [35] et la méthode proposée par T. Morlot *et al.* [18].

3.4.2 Choix de la méthode pour les stations du Québec

Puisque le but du travail est d'estimer l'incertitude des débits estimés par la courbe de tarage, notamment pour les crues pour la définition des zones inondables, la relation de puissance est utilisée afin de pouvoir extrapoler la relation hauteur-débit au-delà des observations. Dans le cadre de ces travaux, la méthode ISO [13, 15, 33] sera adaptée et améliorée pour les stations hydrométriques du Québec méridional. La fonction de puissance dans sa forme (éq. 2.2) et dans sa forme composite (éq. 2.3) est alors la base de la méthode utilisée pour estimer les courbes de tarage. Bien que les approches bayésiennes soient très prometteuses, elles sont beaucoup plus difficiles à implémenter. Ce serait néanmoins une avenue intéressante pour des travaux futurs.

CHAPITRE 4 ANALYSE EXPLORATOIRE DES DONNÉES

Cette section présente les données analysées dans le cadre de ce mémoire. Le ministère de l’Environnement et de la Lutte aux Changements climatiques (MELCC) exploite 230 stations hydrométriques permettant de convertir la hauteur d’un cours d’eau mesurée en temps réel en débit avec les courbes de tarage de ces stations [41]. Le MELCC a mis à notre disposition les données de tous les jaugeages et pour toutes les stations. Il a également mis à disposition certaines informations sur les conditions de jaugeage, notamment s’ils ont été effectués en eau libre, s’il y avait un effet de glace ou d’obstruction qui pourrait compromettre l’écoulement régulier.

Plusieurs jaugeages avec la présence d’effet de glace ont été relevés par le MELCC et ceux-ci ne peuvent pas être considérés pour estimer la relation hauteur-débit. Ces données non représentatives de l’écoulement non contraint peuvent influencer la qualité des courbes de tarage. Cela est illustrée par la figure 4.1 illustrant la station Sainte-Anne (021407). Trois jaugeages de hauteurs moyennes possèdent un débit pratiquement nul, sont dû à l’effet de glace. Une élimination de ces données a été alors effectuée sur l’ensemble des stations. Cela a été effectué grâce à l’information disponible sur les conditions de jaugeage.

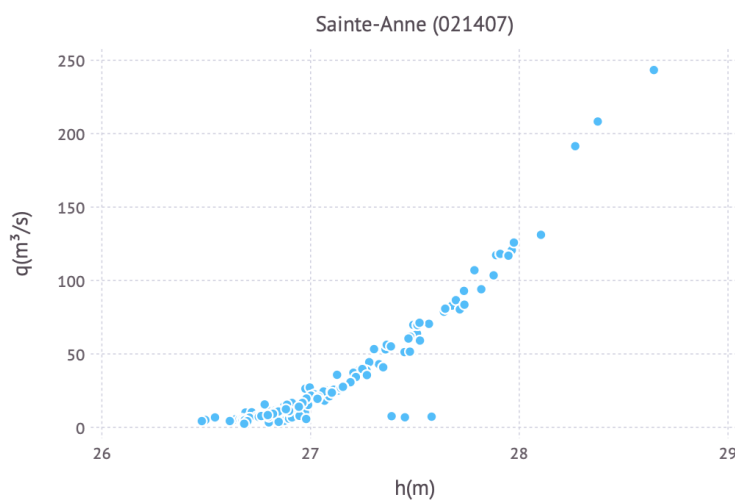


Figure 4.1 Données brutes de la station Sainte-Anne (021407)

Il a été convenu avec le MELCC que l’analyse se concentrerait dans un premier temps sur 15 stations issues de 5 catégories distinctes, soit 3 stations par catégorie. La catégorisation a été effectuée en collaboration avec le MELCC basée sur les caractéristiques des stations. Les

5 catégories peuvent être résumées ainsi :

Catégorie 1 – cette catégorie correspond aux stations idéales. La section de contrôle de la station est stable dans le temps et la géométrie est régulière. Beaucoup de jaugeages sont disponibles et pour toutes les gammes de débits (régime permanent, étiage et crue).

Catégorie 2 – cette catégorie correspond aux stations pour lesquelles l'estimation des débits de crue est difficile.

Catégorie 3 – cette catégorie correspond aux stations où la section de contrôle est géomorphologiquement instable. La relation hauteur-débit n'est alors pas constante dans le temps.

Catégorie 4 – cette catégorie correspond aux stations récentes dont seulement un nombre limité de jaugeages sont disponibles pour l'estimation de la courbe de tarage.

Catégorie 5 – cette catégorie correspond aux stations où la courbe de tarage est atypique ; possiblement dû à la présence d'une plaine inondable qui modifie considérablement la forme de la courbe en fonction de la hauteur de l'eau.

Les stations de la première catégorie correspondent à Sainte-Anne (050408), Jacques-Cartier (050801) et Harricana (080101). La figure 4.2 illustre les jaugeages pour ces trois stations classées comme idéales. On peut dénoter qu'il existe une certaine variabilité des débits mesurés pour une hauteur d'eau semblable, plus visible pour la station Harricana.

Les trois stations sélectionnées de la deuxième catégories sont Yamaska Nord (030340), Saint-Louis (040212) et Des Escoumins (070204). Les jaugeages pour ces trois stations sont illustrés à la figure 4.3. La variabilité des débits semble en effet être plus grande pour les débits de crue pour les deux premières stations tandis que pour la troisième station, la relation semble avoir une cassure pour les grands débits.

Les 3 stations représentant la catégorie 3 où la section de contrôle est géomorphologiquement instable sont Sainte-Anne (021407), Dauphine (120201) et Boyer Sud (023002). Les jaugeages de ces stations sont illustrés à la figure 4.4. La variabilité des débits mesurés pour une même hauteur d'eau est grande pour ces stations, ce qui illustre le fait que la section de contrôle peut être modifiée entre les jaugeages. Ce comportement influence la précision des débits prédits par la courbe de tarage [5]. Cela n'est pas surprenant puisque cette catégorie est associée à une source d'incertitude supplémentaire par rapport aux autres catégories, celle de la modification temporelle de la section de contrôle ε_3 .

Les stations choisies pour représenter la catégorie 4 pour les stations récentes sont Nicolet (030120), Capitachouane (043030) et Ha! Ha! (060601). Les jaugeages pour ces trois stations sont illustrés à la figure 4.5. Lors de la première visualisation des données de la station

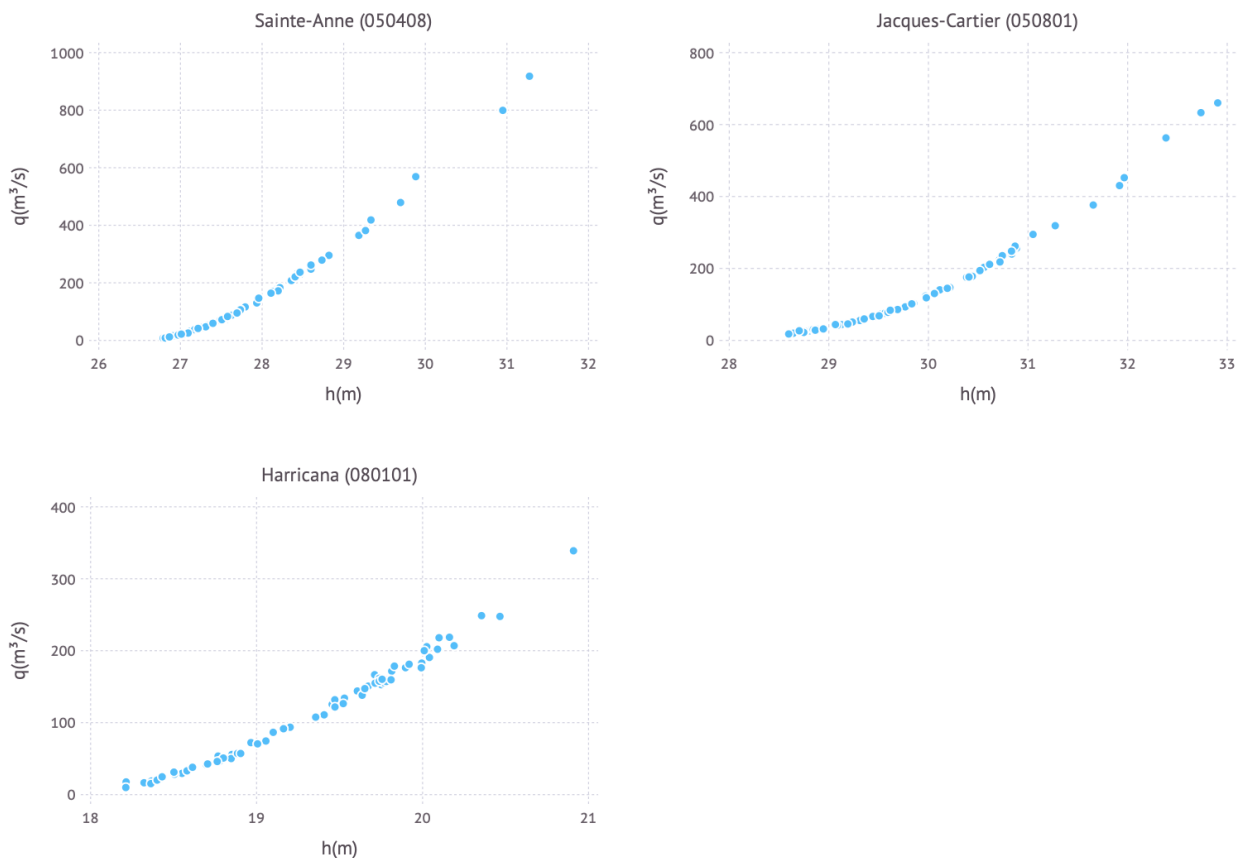


Figure 4.2 Jaugeages des trois stations considérées pour représenter la catégorie 1.

Ha!Ha!, les données étaient hétérogènes. Il y avait en effet un décalage de 50 m pour la hauteur d'eau. En examinant la documentation sur cette station, il est mentionné que cette station a été fermée en 1995. Ce décalage est survenu après sa réouverture en 2019 parce que la hauteur d'eau a été calculée à partir d'une nouvelle élévation géodésique. Par conséquent, les jaugeages avant cette date ne sont pas considérés. Une remarque générale sur les trois stations est le faible nombre de mesures sur chaque gamme de débits. Par exemple, la station Capitachouane (043030) comporte seulement deux couples de mesure dans la gamme $50 \text{ m}^3/\text{s} - 100 \text{ m}^3/\text{s}$. Cela peut influencer la qualité des courbes de tarage associées à ces stations et les incertitudes correspondantes peuvent être élevées.

Les stations sélectionnées pour représenter la cinquième catégorie sont Yamaska Nord (030309), Du Chêne (046709) et Suffolk (040409). Les jaugeages de ces trois stations sont illustrés à la figure 4.6. Ces trois stations comportent un nombre modéré de couples de hauteur-débit, cependant la gamme des débits élevés comportent moins de débit mesuré. La relation hauteur-débit pour ces stations ne semble pas être lisse. Une cassure dans la relation semble visible.

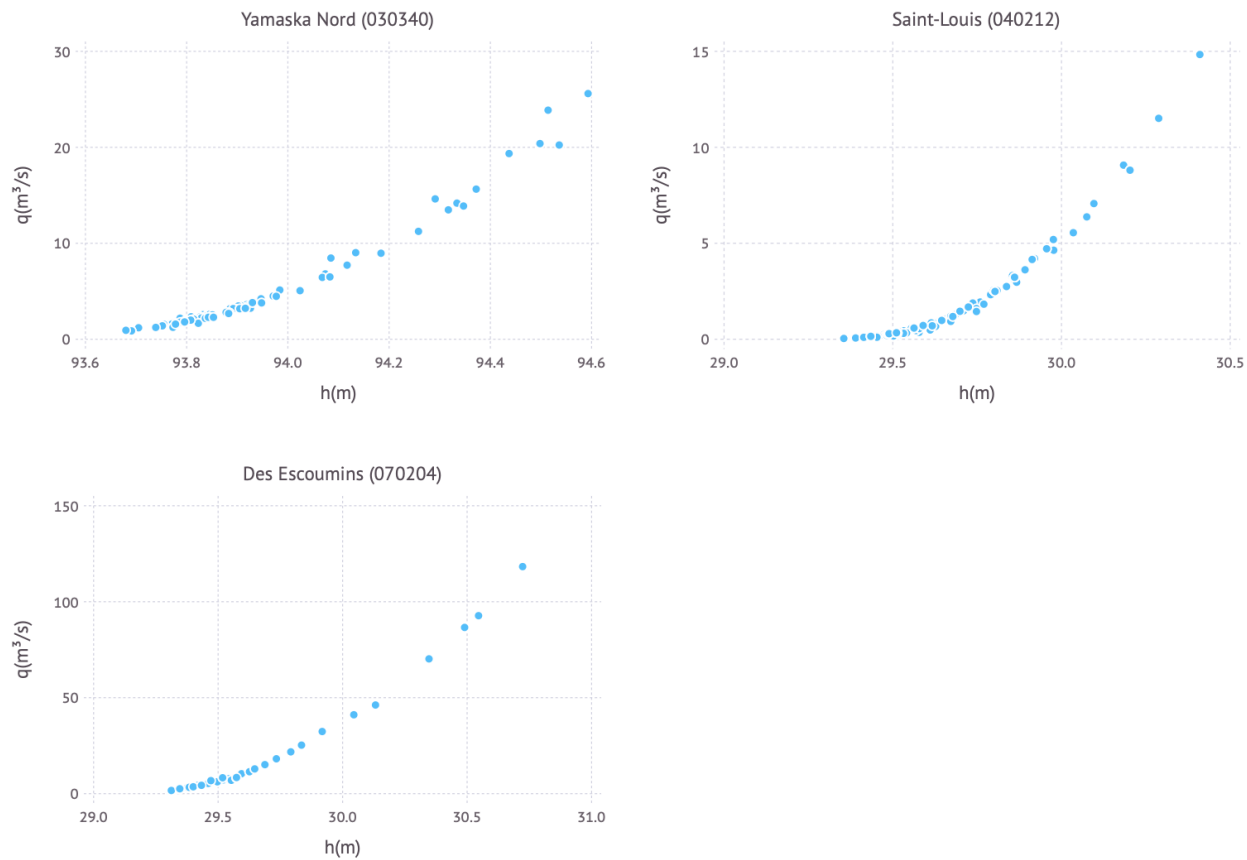


Figure 4.3 Jaugeages des trois stations considérées pour représenter la catégorie 2.

En effet, cette cassure est observée pour la station Yamaska Nord (030309) à partir de la hauteur 30.55 m . Pour la station Du Chêne (046709), cette cassure est observée à partir de la hauteur 30.8 m . Pour la station Suffolk (040409), cette cassure est observée à partir de la hauteur 30.95 m .

Les informations disponibles sur les 15 stations choisies sont compilées dans les tableaux A.1, A.2, A.3, A.4 et A.5 disponible en annexe A. Selon les propriétés des catégories définies, seulement la troisième catégorie comporte les 4 sources d'incertitudes ε_1 , ε_2 , ε_3 et ε_4 définies à l'équation 2.4. Cela est dû au fait qu'elle est la seule catégorie associée à des changements géomorphologiques importants. Les autres catégories comportent les trois sources d'incertitudes ε_1 , ε_2 et ε_4 . La composante ε_3 peut également exister dans les autres catégories mais elle est d'une faible amplitude.

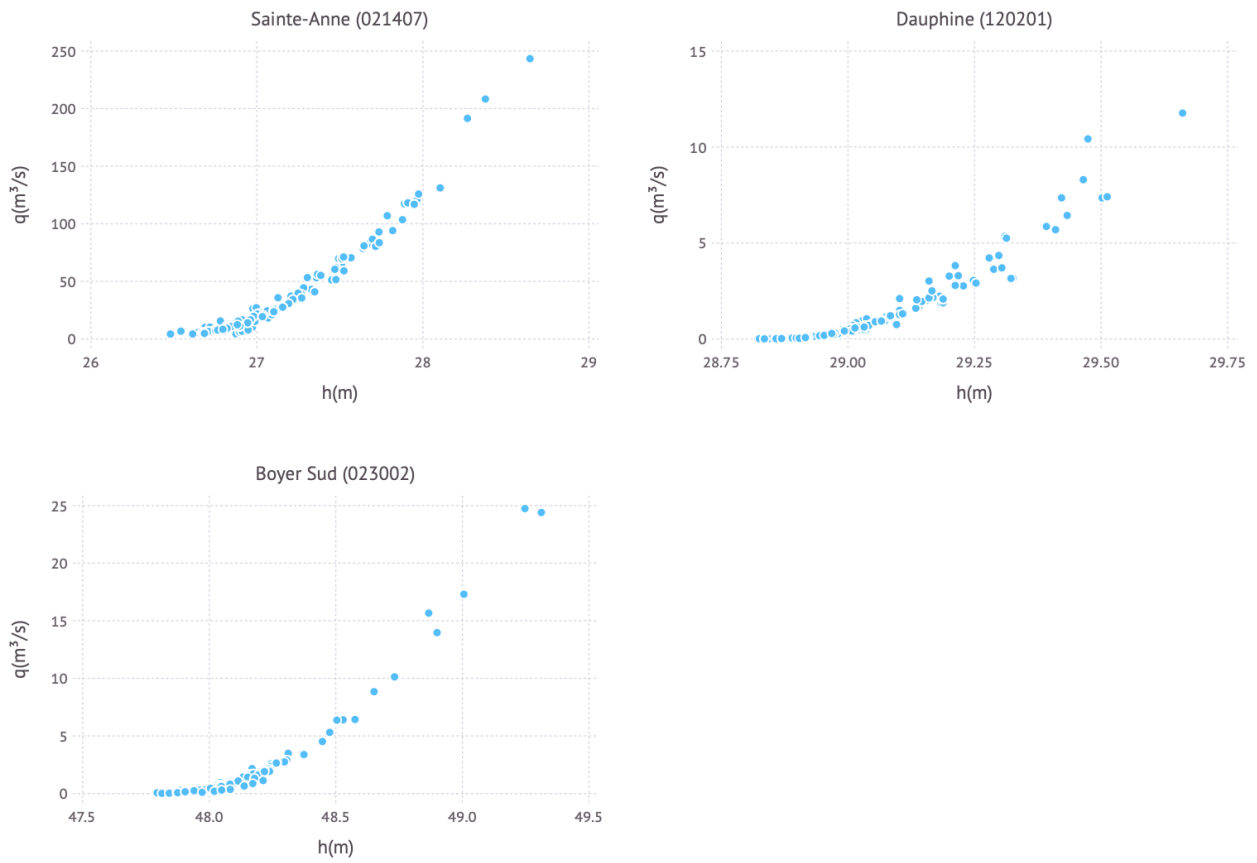


Figure 4.4 Jaugeages des trois stations considérées pour représenter la catégorie 3.

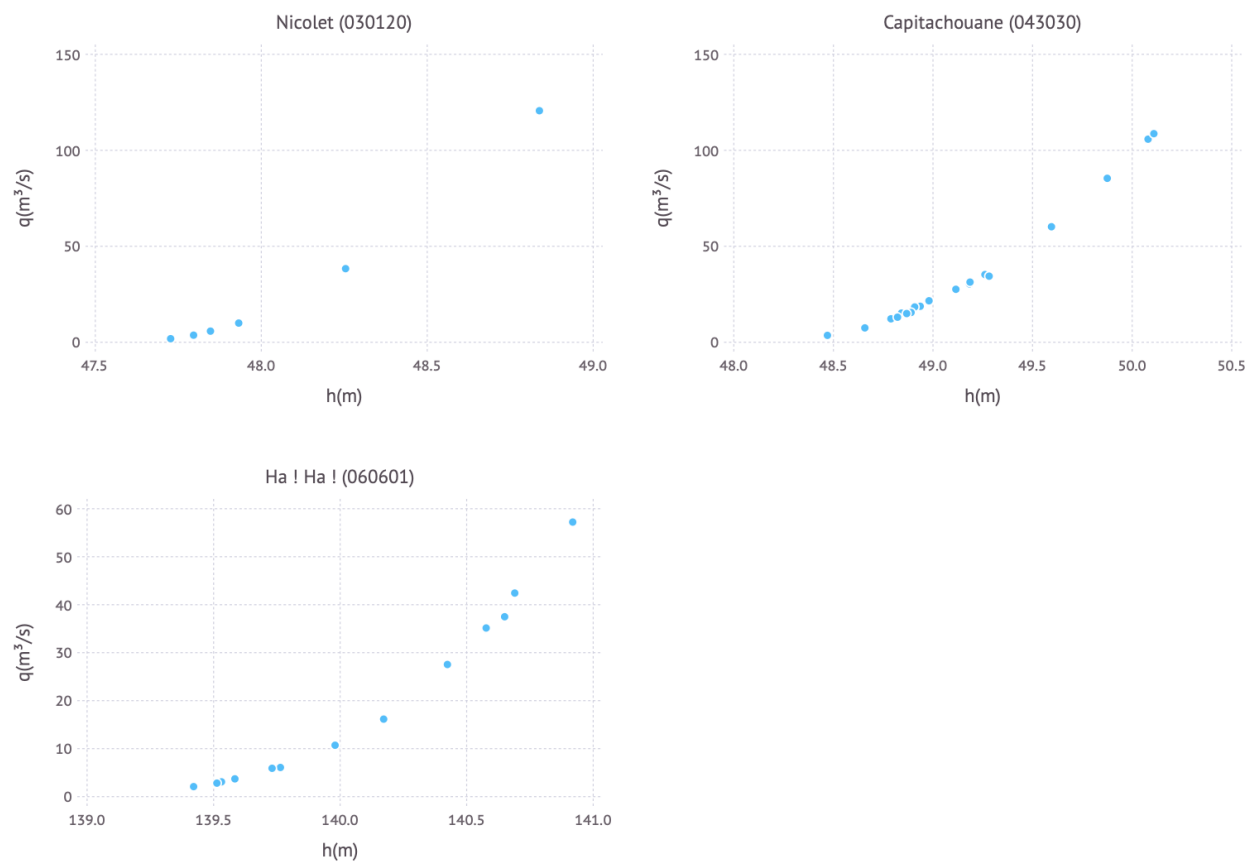


Figure 4.5 Jaugeages des trois stations considérées pour représenter la catégorie 4.

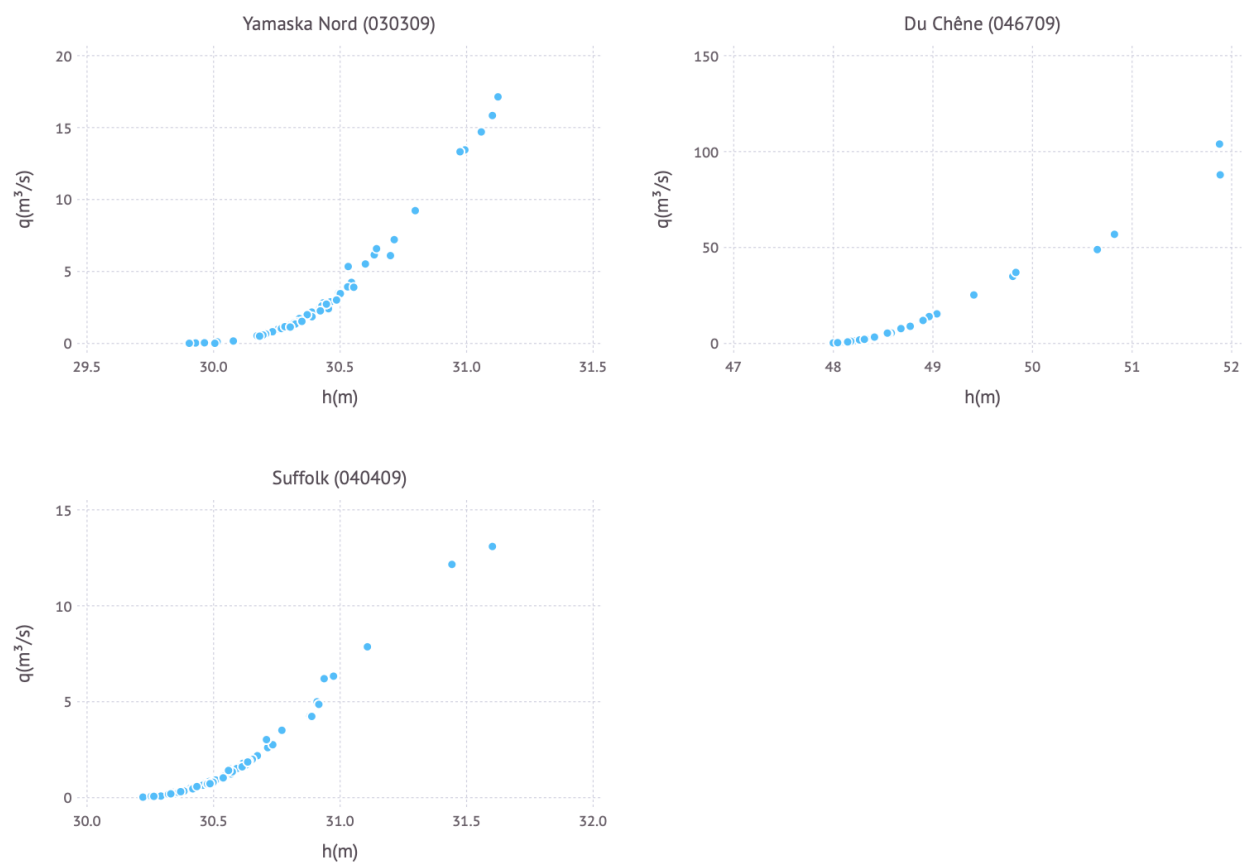


Figure 4.6 Jaugeages des trois stations considérées pour représenter la catégorie 5.

CHAPITRE 5 MÉTHODOLOGIE

Cette section présente la méthodologie développée pour estimer la relation hauteur-débit d'un cours d'eau naturel et son incertitude. La méthode d'estimation des paramètres de la fonction de puissance est d'abord présentée, puis celle pour l'estimation de l'incertitude des débits prédits.

5.1 Modélisation de l'erreur des débits mesurés

Dans la littérature, l'incertitude sur la hauteur mesurée est généralement considérée négligeable par rapport à l'erreur sur le débit mesuré. Il est alors commun de ne considérer que l'erreur sur les débits. De plus, l'erreur de mesure du débit est couramment modélisée avec une loi normale d'espérance nulle (voir par exemple [1, 18]). Les arguments qui justifient cette hypothèse se résument au fait que cette incertitude est une somme de différentes sources d'erreur telles que l'approximation de la surface de la section transversale, à l'erreur d'étalonnage des instruments, à l'estimation discrète du profil de vitesse, etc. Selon le théorème limite central, la somme de ces sources d'erreur tend vers une distribution normale indépendamment de la distribution de chaque erreur [42].

En s'appuyant alors sur le rapport de [43] et sur l'expertise du MELCC, l'incertitude relative de la mesure du débit par jaugeage est d'au plus 5%. Il est supposé que le vrai débit Q_i est distribué selon la loi normale centrée au débit mesuré q_i et avec une variance pour laquelle 95% des valeurs sont plus petites que l'erreur relative de 5% :

$$Q_i \sim \mathcal{N} \left\{ q_i, \left(\frac{0.05 q_i}{1.96} \right)^2 \right\}. \quad (5.1)$$

5.2 Ajustement de la fonction de puissance à un segment

5.2.1 Estimation des paramètres

Tel que mentionné dans le chapitre 2, l'équation (2.2) est la plus utilisée pour décrire la relation hauteur-débit pour un seul segment et c'est celle-ci qui sera utilisée dans le cadre de ce mémoire. Cependant, l'estimation des paramètres de cette relation n'est pas immédiate en raison de la non-linéarité de la formulation. Dans l'espace logarithmique, la relation s'écrit

par un développement et une approximation de série Taylor de la façon suivante :

$$\log q(h) = \log a + c \log(h - b) \quad (5.2)$$

Si on connaissait le paramètre b , alors les paramètres a et c qui minimisent la somme du carré des erreurs pourraient être obtenus explicitement en utilisant la régression linéaire, *i.e.* les équation (2.5) et (2.6). Or, le paramètre b est inconnu et doit également être estimé.

Afin d'exploiter la solution explicite pour les paramètres a et c dans le cas de la minimisation de la somme du carré des erreurs, on pose la fonction objectif suivante à minimiser :

$$f(a, b, c) = \sum_{i=1}^n \{\log q_i - \log a - c \log(h_i - b)\}^2 \quad (5.3)$$

à l'aide de l'ensemble des n jaugeages $\{(h_i, q_i) : 1 \leq i \leq n\}$ de la section de contrôle. La méthode d'optimisation adaptée à cette fonction objectif consiste à parcourir les valeurs possibles de $b \in (-\infty, \min_{1 \leq i \leq n} h_i)$ avec l'algorithme L-BFGS (*limited-memory Broyden—Fletcher—Goldfarb—Shanno algorithm*, [44]). Pour chacun des points candidats pour b , les paramètres a et c sont calculés explicitement et la valeur de la fonction objectif est efficacement évaluée.

À titre d'exemple, la figure 5.1 illustre l'ajustement de l'équation (5.3) aux 192 jaugeages de la rivière Sainte-Anne. Les estimations sont les suivantes :

$$\hat{a} = 76.39, \quad \hat{b} = 26.55 \quad \text{et} \quad \hat{c} = 1.65.$$

5.2.2 Estimation de l'incertitude des paramètres des courbes de tarage

Bien que les méthodes d'optimisation donnent les meilleures estimations des paramètres des courbes de tarage, les différentes sources d'incertitude de l'équation 2.4 constituent un enjeu important afin d'obtenir des estimations de l'incertitude de ces paramètres. L'utilisation du bootstrap non paramétrique constitue un outil de choix pour estimer l'incertitude des paramètres de la courbe. Le rééchantillonnage obtenu par bootstrap permet de tirer un échantillon aléatoire des jaugeages selon la fonction de répartition empirique de l'échantillon empirique. La courbe de tarage est ajustée pour chacun des échantillons bootstrap obtenus et la distribution des paramètres peut être étudiée. Comme mentionné dans la section 2.4, dans le cas non paramétrique la méthode bootstrap permet de calculer des intervalles de prédiction sans considérer des hypothèses sur la fonction de répartition générée.

Avec les n observations de couple hauteur-débit $\{(h_i, q_i) : 1 \leq i \leq n\}$, 1000 échantillons

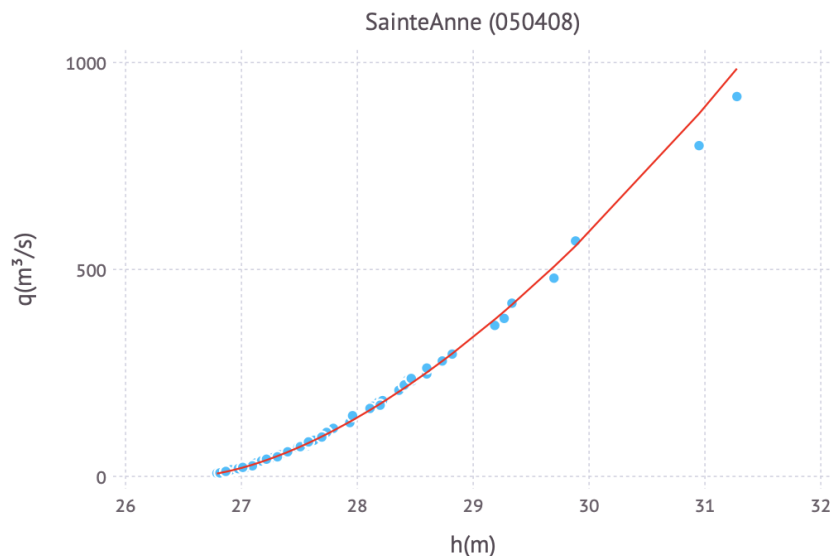


Figure 5.1 La relation hauteur-débit correspondante à la station Sainte-Anne (050408) (ajustée à un seul segment)

bootstrap de taille n sont générés à partir des observations par échantillonnage avec remise. Afin de prendre en compte l'incertitude de mesure des jaugeages, une nouvelle réalisation du débit q_i est générée selon la distribution exprimée à l'équation (5.1) pour chaque jaugeage sélectionné dans l'échantillon bootstrap. Aussi, la plus petite valeur de hauteur d'eau est toujours sélectionnée de façon à s'assurer que la courbe de tarage soit définie sur toute l'étendue des observations originales.

Avec les 1000 ajustements de la courbe de tarage, des intervalles de confiance sur les paramètres peuvent être obtenus empiriquement. Par exemple, pour la station de la rivière Sainte-Anne, 95% des estimations de a se trouvent dans l'intervalle $[72.32, 77.62]$, 95% des estimations de b se trouvent dans l'intervalle $[26.52, 26.55]$ et 95% des estimations de c se trouvent dans l'intervalle $[1.64, 1.71]$. Puisque l'ajustement des courbes ne peut pas être effectué que pour des échantillons de tailles supérieurs aux nombres des paramètres, le succès de l'ajustement prouve que les échantillons bootstrap obtenus sont toujours de tailles supérieurs aux nombres des paramètres. Par conséquent, l'approximation de la variance échantillonnage est significativement valide.

5.2.3 Estimation de l'incertitude des débits prédits

L'estimation de l'incertitude des débits prédits par la courbe de tarage est obtenue en utilisant la variance de prédiction de la régression linéaire dans l'espace logarithmique. La variance de

prédiction est la somme de la variance associée à l'estimation des paramètres de la courbe et à la variance de mesure des jaugeage. La variance d'estimation du logarithme du débit d'estimation est obtenue avec l'équation (2.8) appliquée à la hauteur h_0 pour laquelle une estimation est requise :

$$\mathbb{V}ar(\log \hat{q}(h_0)) = \begin{bmatrix} 1 & \log(h_0 - b) \end{bmatrix} \hat{\sigma}^2 (X^\top X)^{-1} \begin{bmatrix} 1 \\ \log(h_0 - b) \end{bmatrix},$$

où la matrice de structure X est composée d'une colonne de «1» pour l'ordonnée à l'origine de la régression linéaire et de la colonne $\{\log(h_i - b) : 1 \leq i \leq n\}$.

En considérant l'incertitude de mesure des débits jaugés, un intervalle de confiance de niveau $(1 - \alpha)$ du logarithme du débit peut être développé à partir de l'équation (2.9) de la façon suivante :

$$\log \hat{q}(h_0) \pm t_{n-1}(1 - \alpha/2) \sqrt{\sigma_{\log Q}^2 + \begin{bmatrix} 1 & \log(h_0 - b) \end{bmatrix} \hat{\sigma}^2 (X^\top X)^{-1} \begin{bmatrix} 1 \\ \log(h_0 - b) \end{bmatrix}}$$

où $\sigma_{\log Q}^2$ est la variance des débits mesurés dans l'espace logarithmique. Ce terme est calculé à partir de la méthode delta [45] appliquée à l'équation (5.1) :

$$\sigma_{\log Q}^2 = \left(\frac{0.05}{1.96} \right)^2.$$

En effet, pour la fonction logarithmique, la méthode delta stipule que si la variable aléatoire X possède l'espérance μ et la variance σ^2 , si $\sigma \ll \mu$ alors l'espérance et la variance de la variable aléatoire $\log X$ peuvent être approximées par :

$$\mathbb{E}(\log X) \approx \log \mu \quad \text{et} \quad \mathbb{V}ar(\log X) \approx \frac{\sigma^2}{\mu^2}.$$

À titre d'exemple, l'incertitude du débit prédit par la courbe de tarage pour la station Sainte-Anne est illustré à la figure 5.2.

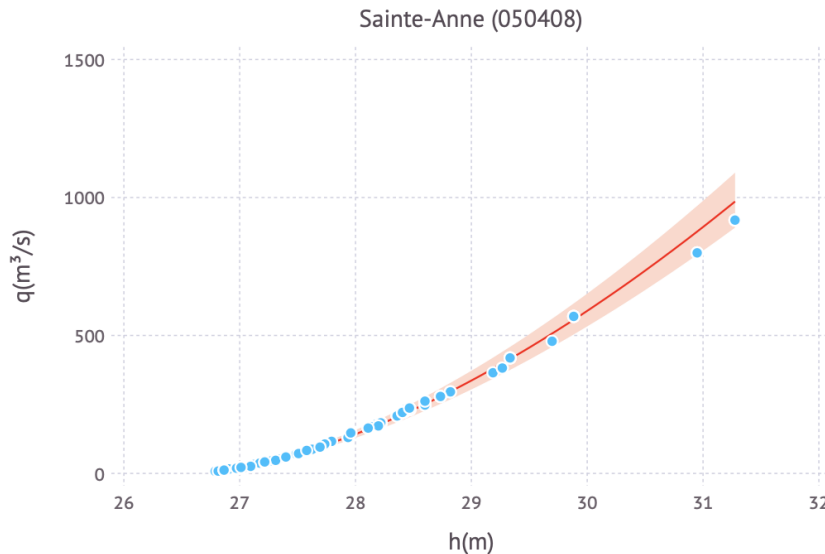


Figure 5.2 L'incertitude du débit prédit par la courbe de tarage pour la station Sainte-Anne (050408)

5.3 Ajustement de la fonction de puissance à deux segments

5.3.1 Estimation des paramètres

Le logarithme du débit $q(h)$ à la hauteur h d'une courbe à deux segments (équation 2.3) est donné par l'équation suivante :

$$\log q(h) = \begin{cases} \log a_1 + c_1 \log(h - b_1) & \text{si } b_1 < h \leq k, \\ \log a_2 + c_2 \log(h - b_2) & \text{si } h > k \text{ et } h > b_2, \\ -\infty & \text{sinon.} \end{cases} \quad (5.4)$$

La courbe composite possède 7 paramètres, les paramètres du premier segment (a_1, b_1, c_1) et du deuxième (a_2, b_2, c_2) ainsi que le point de rupture k . À l'instar de la courbe à un segment, l'estimation des paramètres s'effectue en minimisant la somme du carré des erreurs entre la courbe et les n jaugeages $\{(h_i, q_i) : 1 \leq i \leq n\}$ de la section de contrôle. La méthode d'optimisation adaptée à cette fonction objectif consiste parcourir les valeurs possibles de la rupture k avec l'algorithme L-BFGS. Pour les hauteurs supérieures au candidat pour k , la courbe de tarage est ajustée avec la méthode décrite à la section 5.2.1 avec les jaugeages effectués à une hauteur supérieure à k .

Pour les hauteurs inférieures au k candidat, le premier segments est ajusté avec les jaugeages

dont la hauteur mesurée est inférieure à k . Une contrainte de continuité est cependant ajoutée. La courbe du segment inférieur doit absolument passer par le point $(k, \log a_2 + c_2 \log(k - b_2))$ pour assurer la continuité du débit prédit en fonction de la hauteur. Parmi toutes les courbes passant par le point $(k, \log a_2 + c_2 \log(k - b_2))$, la courbe du premier segment est celle qui minimise la somme du carré des erreurs. À titre d'exemple, la figure 5.3 illustre la courbe de tarage à deux segments ajustée au jaugeages de la rivière Sainte-Anne. On remarque que les jaugeages suggèrent une rupture dans la relation hauteur-débit à $k = 27.77 \text{ m}$.

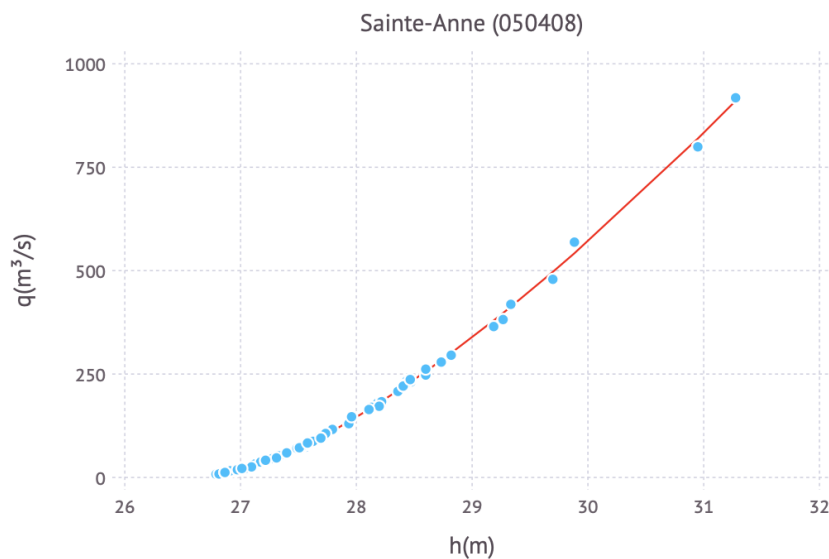


Figure 5.3 La relation hauteur-débit correspondante à la station Sainte-Anne (050408) (ajustée à deux segments)

5.3.2 Estimation de l'incertitude des paramètres

À l'instar de la courbe à un segment, les incertitudes des paramètres sont estimées avec un bootstrap non paramétrique. La courbe à deux segments est ajustée à tous les échantillons obtenus par le rééchantillonnage avec remise. Dans le cas de la rivière Sainte-Anne,

les intervalles de confiance de niveau 95% sont les suivants pour chacun des paramètres :

$$\mathbb{P}(k \in [27.5, 28.0]) = 0.95;$$

$$\mathbb{P}(a_1 \in [55.4, 68.4]) = 0.95;$$

$$\mathbb{P}(b_1 \in [26.4, 26.5]) = 0.95;$$

$$\mathbb{P}(c_1 \in [1.83, 2.13]) = 0.95;$$

$$\mathbb{P}(a_2 \in [104, 118]) = 0.95;$$

$$\mathbb{P}(b_2 \in [26.7, 26.8]) = 0.95;$$

$$\mathbb{P}(c_2 \in [1.36, 1.44]) = 0.95.$$

5.3.3 Estimation de l'incertitude des débits prédits

Pour la courbe de tarage à deux segments, la variance de l'erreur σ^2 est estimée avec la somme du carré des erreurs (équation 2.7) séparément pour chacun des segments. Les intervalles de prédictions pour chacun des segments sont construits de la façon décrite à la section 5.2.3. À titre d'exemple, l'intervalle de prédiction de niveau 95% pour les débits prédits par la courbe de tarage à deux segments est illustré à la figure 5.4.

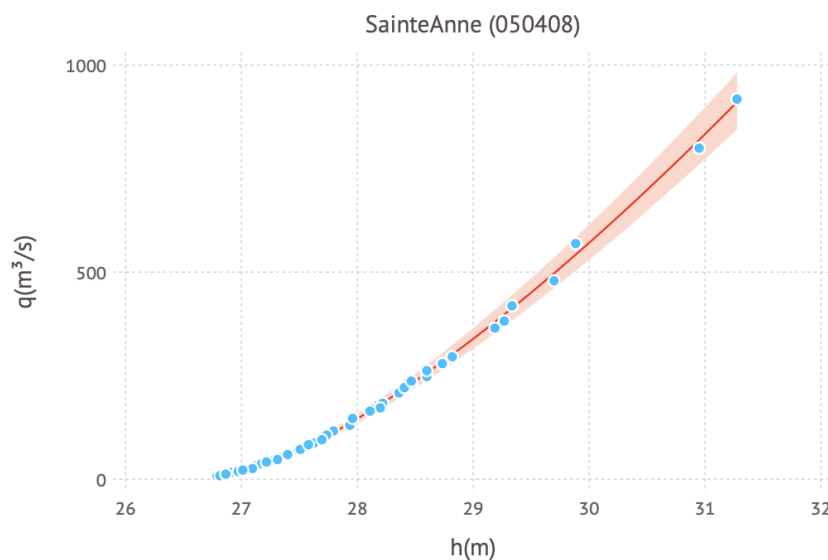


Figure 5.4 L'incertitude du débit prédit par la courbe de tarage pour la station Sainte-Anne (050408) (ajustée à deux segments)

5.4 Sélection du nombre de segments

La figure 5.1 illustre la courbe à un segment ajustée aux jaugeages de la rivières Sainte-Anne et la figure 5.3 illustre la courbe à deux segments ajustée à ces même jaugeages. L'ajustement de la courbe à deux segments est bien sûr meilleur car le modèle possède plus de flexibilité mais il convient de vérifier si le modèle est significativement meilleur. Le critère d'information bayésien (*BIC*) peut être utilisé pour comparer le modèle à un segment et à deux segments à l'instar de [23]. Le BIC est une mesure d'adéquation du modèle aux données mais qui pénalise par le nombre de paramètres. Le nombre de segments requis est donné par le modèle de courbe de tarage qui minimise le BIC. Ce critère est basé seulement sur les données pour comparer les différents modèles possibles. Par conséquent, son utilisation serait avantageuse lorsque l'information sur la section de contrôle est non disponible. Ce critère a été utilisé aussi par A. Petersen-Øverleir et T. Reitan [23] pour le choix du nombre de segments.

Pour la courbe de tarage à un segment, le BIC s'écrit de la façon suivante :

$$\text{BIC} = n \log \hat{\sigma}^2 + 3 \log n,$$

où le nombre 3 correspond au nombre de paramètres du modèle. Dans le cas de la courbe de tarage à deux segments, le BIC s'écrit de la façon suivante :

$$\text{BIC} = n_1 \log \hat{\sigma}_1^2 + n_2 \log \hat{\sigma}_2^2 + 6 \log n,$$

où n_1 et n_2 correspondent respectivement aux nombres de jaugeages attribués au premier segment et au deuxième segment, $\hat{\sigma}_1^2$ et $\hat{\sigma}_2^2$ correspondent à la variance de l'erreur pour chacun des segments et le nombre 6 correspond aux degrés de liberté du modèle. Le modèle possède 7 paramètres mais il perd un degré de liberté en raison de la contrainte de continuité imposée.

Dans le cas des jaugeages de la rivière Sainte-Anne, la courbe à deux segments est choisie car le BIC associé est de -1262, ce qui est plus petit que celui de la courbe à un segment évalué à -1168.

CHAPITRE 6 RÉSULTATS ET DISCUSSION

Cette section présente les résultats relatifs à l'estimation des courbes de tarage ainsi que les incertitudes correspondantes des 15 stations sélectionnées et précédemment décrites à la section 4.

6.1 Sélection du nombre de composantes de la courbe de tarage

Pour chacune des stations sélectionnées, l'ajustement de la courbe de tarage à un et à deux segments a été effectué tel que décrit à la section 4. La qualité de l'ajustement de ces modèles a été évaluée par le BIC et ces valeurs sont compilées dans le tableau 6.1.

Tableau 6.1 BIC obtenus pour les ajustement de la courbe de tarage à un segment et à deux segment pour les 15 stations hydrométriques sélectionnées. Le plus petit BIC pour chacune des stations est en caractère gras.

Station	Un segment	Deux segments
SainteAnne (050408)	-1182.99	-1253.58
Jacques-Cartier(050801)	-563.99	-605.99
Harricana (080101)	-1068.99	-1115.73
Yamaska Nord (030340)	-674.20	-670.25
Saint-Louis (040212)	-1141.62	-1221.03
Des Escoumins (070204)	-264.02	-295.06
Sainte-Anne (021407)	-884.74	-1073.77
Dauphine (120201)	-1040.61	-1192.79
Boyer Sud (023002)	-388.21	-486.38
Nicolet (030120)	-31.68	-
Capitachouane (043030)	-223.52	-274.56
Ha! Ha!(060601)	-86.88	-75.72
Yamaska Nord (030309)	-407.07	-433.61
Du Chêne (046709)	-312.54	-384.27
Suffolk (040409)	-670.56	-711.16

Pour la station Nicolet (030120), le nombre de jaugeages disponibles (6 jaugeages) n'était pas suffisant pour estimer la courbe de tarage à deux segments. De façon générale, le BIC permet de sélectionner de façon décisive le meilleur modèle entre une courbe de tarage à un segment et à deux segments. En effet, à part le cas de la station Yamaska Nord (030340) qui a un écart entre les deux BIC de 3.95, l'écart entre les deux modèles est toujours supérieur à 10 et il peut aller jusqu'à plus de 100 comme dans le cas de la station Sainte-Anne (021407).

Selon le BIC, la courbe de tarage à un seul segment n'est retenue que pour les stations Ha! Ha! (060601) et Yamaska Nord (030340). Pour les autres stations, la courbe de tarage à deux segments est privilégiée. Par ailleurs, les stations de troisième catégorie ont marqué l'écart le plus important. À titre de rappel, il était difficile d'identifier à l'analyse exploratoire le besoin de segmentation pour cette catégorie à cause de la dispersion des données. Les courbes ajustées pour chacune des stations sont disponibles à l'annexe C (ces figures présentent les courbes de tarages et l'incertitude du débit prédit par les courbes de tarage à la fois). À titre d'exemple, la courbe de tarage de la station Jacques-Cartier (050801) est présentée à la figure 6.1. Le point de rupture entre les segments est estimé à 30.98 m.

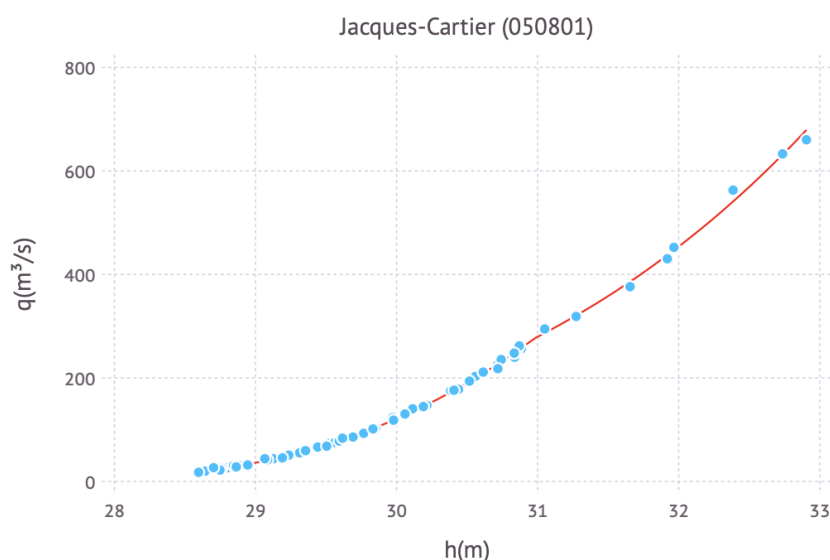


Figure 6.1 Courbe de tarage ajustée pour la station Jacques-Cartier (050801).

L'information sur la nature du lit de la rivière n'étant pas toujours disponible, la conclusion sur le meilleur modèle est basée seulement sur le BIC. C'est l'information apportée par les jaugeages qui permet de conclure qu'il y a une rupture dans la relation hauteur-débit. En effet, étant donné que l'écart important entre les deux modèles, il est possible que ce résultat ait révélé une nécessité importante de la segmentation des courbes de tarage des stations étudiées. Cependant, dans le cas des stations de la cinquième catégorie, ce résultat est confirmé hydrauliquement. En effet, cette catégorie est caractérisée par l'existence d'une plaine inondable ou bien d'une discontinuité dans la rivière. De plus, elle a montré des signes clairs d'un certain déphasage dans les données dans l'analyse exploratoire. Ensuite, pour le seul résultat différent de la station Ha! Ha! (060601), il semble qu'effectivement, un seul segment est adéquat à ce qui a été observé au niveau des données qui n'ont pas montré un

signe de déphasage à une certaine hauteur d'eau.

D'après les courbes de tarage obtenues, il est visiblement clair que le modèle ajuste bien les courbes par rapport aux données de jaugeages. Cela est valide pour les différentes situations hydrauliques traitées. En effet, c'est le cas pour les stations stables de la première et de la deuxième catégorie. D'ailleurs au niveau de la deuxième catégorie, il semble que l'ajustement au débit élevé est satisfaisant et aucune problématique n'est associée à cette gamme de débit. Même la station Yamaska Nord (030340) (figure C.3) de cette catégorie, sujette à une légère dispersion de données a un bon ajustement. C'est aussi le cas des stations qui ont une configuration spécifique des données. En effet, pour les stations de la troisième catégorie, qui sont caractérisées par une dispersion des données, les courbes obtenues semblent satisfaisantes. Pour la quatrième catégorie, il semble aussi que l'ajustement soit très satisfaisant malgré la faible quantité de données. Finalement, pour le cas de la cinquième catégorie, seulement pour la station Du Chêne (046709) (figure C.13), il semble que la courbe sous-estime les débits élevés. Mais cela est non surprenant puisque cette station n'a que 3 jaugeages dans la gamme des débits élevés et c'est bien clair qu'elles sont sujettes à des incertitudes importantes. Comme mentionné précédemment, le manque de jaugeages dans une gamme influence la qualité de la courbe ajustée. Pour les deux autres stations de cette catégorie, la courbe de tarage semble d'une meilleure qualité. Il est possible que cette différence dans la qualité par rapport à la station Du Chêne (046709) est due, dans le cas de la station Suffolk (040409) (figure C.14), à des incertitudes moins importantes au niveau de la gamme des débits élevés. Pour la station Yamaska-Nord (030309) (figure C.12), cette différence peut être due au nombre de jaugeages plus important dans la gamme des débits élevés.

La validité hydrauliques des paramètres obtenus ne peut pas être évaluée avec les jaugeages seulement. L'expertise d'un hydrologue serait nécessaire sur notamment la forme des sections de contrôles des stations hydrométriques étudiées. Néanmoins, le très bon ajustement des courbes sur les jaugeages permet de présumer que la relation estimée est adéquate.

6.2 Incertitudes des paramètres de la courbe de tarage

Les incertitudes des paramètres de la courbe de tarage ont été estimées avec la méthode bootstrap décrite à la section 5.2 et les intervalles de confiance de niveau 95% sont compilés pour trois stations au tableau 6.2. Les résultats pour les autres stations sont disponibles aux tableaux B.1, B.2, B.3 et B.4 de l'annexe B. Avec ces intervalles, il est possible de vérifier si l'ordre de grandeur de ces paramètres peut être hydrauliquement acceptable. À titre de rappel, selon [1], le coefficient a doit être strictement positif, la limite du coefficient c varie de 0.5 jusqu'à 2.5 dans le cas d'une section perpendiculaire au débit moyen. C'est vrai que cet

intervalle est spécifique pour ce cas mais il peut donner une idée sur l'ordre de grandeur de ce paramètre. Par conséquent, plus la valeur obtenue est loin de ces bornes, plus il y aura des doutes sur la validité du résultat. Ensuite, le paramètre b doit être inférieur à $\{\text{minimum}(h_i) : 1 \leq i \leq n\}$. Finalement, une valeur réaliste du point de rupture k ne peut pas être inférieure à $\{\text{minimum}(h_i) : 1 \leq i \leq n\}$ ni supérieure à $\{\text{maximum}(h_i) : 1 \leq i \leq n\}$.

Pour le paramètre a , il semble que le modèle a bien ajusté des valeurs strictement positives pour les 15 courbes. Cela n'est pas surprenant puisque c'était une condition imposée pour l'ajustement du modèle. Par exemple, l'estimation obtenue pour la courbe de tarage de la station Sainte-Anne (050408) est de $a_1 = 68.29$ avec un intervalle de confiance de $[62.09, 73.10]$ et de $a_2 = 112.97$ avec un intervalle de confiance $[100.52, 125.33]$. Cependant, il y a quelques cas où des valeurs très proches de zéro sont obtenues. Les valeurs les plus irréalistes sont notamment aux stations Sainte-Anne (021407) et Dauphine (120201) de la troisième catégorie. En effet, pour la station Sainte-Anne (021407), l'estimation donne une valeur de $a_1 = e^{-322}$ avec un intervalle de prévision de $[e^{-324}, e^{-5}]$ et pour la station Dauphine (120201), une valeur de $a_2 = e^{-323}$ avec un intervalle de prévision de $[e^{-323}, 1137.57]$. Ces valeurs très faibles peuvent être causées par le nombre important de jaugeages de débits très faibles (voir figure 4.4) ou par des incertitudes importantes possibles pour cette catégorie. Cette deuxième hypothèse est remarquable dans le cas de la station Dauphine (120201) par le très large intervalle de prévision du paramètre a_1 . Deux autres stations présentent des valeurs de a proche de 0. D'abord, l'estimation pour la station Jacques-Cartier (050801) donne un $a_2 = 0.001$ avec à un intervalle de prévision de $[e^{-319}, 82.17]$, puis, l'estimation pour la station Saint-Louis (040212) donne un $a_2 = 0.15$ avec à un intervalle de prévision $[e^{-8}, 14.46]$. La raison de ces valeurs faibles reste difficile à expliquer puisque ces deux stations sont assez stables. Il est possible que ces valeurs soient causées par les incertitudes importantes au niveau de ces stations qui sont remarquables au niveau des intervalles de prévisions très larges. De plus, les bornes supérieures de ces intervalles sont très éloignées de la valeur 0. Par exemple, pour la station Jacques-Cartier (050801), la borne supérieure est de 82.17, ce qui donne une large gamme de valeurs possibles réalistes pour du paramètre a . De plus, d'une façon générale, l'estimation de ce paramètre pour l'ensemble des stations correspond à de larges intervalles de prévisions surtout au deuxième segment, ce qui implique possiblement que des incertitudes importantes sont reliées à ce paramètre. Par exemple, à la station Harricana (080101), l'estimation de ce paramètre donne un intervalle de prévision de $[19.20, 63.22]$ pour le premier segment et un intervalle de prévision de $[e^{-8}, 99.42]$ pour le deuxième segment. La dernière constatation à propos des intervalles de prévision du paramètre a est l'existence de bornes inférieures très proches de 0. Par exemple, l'estimation pour la station Du Chêne (046709) au deuxième segment donne une borne inférieure de e^{-323} . Ce résultat suggère l'existence d'un

volume important de faibles débits incertains qui influence la qualité des courbes de tarage. Pour le paramètre b , aucune valeur supérieure au minimum des hauteurs d'eau n'a été obtenue. Cela est valide aussi pour les intervalles de confiance. Ce résultat n'est pas surprenant puisque l'intervalle d'optimisation a été fixé à $(-\infty, \min_{1 \leq i \leq n} h_i)$ pour ce paramètre. Cependant, il y a trois cas où la borne inférieure de l'intervalle de confiance du paramètre b_2 est très faible. En effet, l'estimation pour les stations Jacques-Cartier (050801) et Du Chêne (046709) donne respectivement des intervalles de confiance de $[-245.33, 28.69]$ et de $[-253.29, 48.39]$. Ces bornes sont loin d'être réalistes. À vrai dire, la borne inférieure $-\infty$ de l'intervalle d'optimisation permet théoriquement d'atteindre des valeurs très faibles. D'ailleurs, elle a été fixée de cette façon pour avoir un algorithme standard pour tous les cas étudiés. Mais à l'échelle pratique, il n'est pas attendu d'atteindre de telles valeurs irréalistes, grâce à l'hypothèse que la fonction de puissance est capable de modéliser adéquatement des paramètres acceptables hydrauliquement et qui ajuste le mieux possible les données. Par conséquent, il est possible que l'échantillonnage bootstrap ne constitue pas toujours des échantillons adéquatement modélisables par la fonction de puissance comme ici dans le cas des courbes atypiques ou bien dans le cas où les incertitudes de mesures de hauteur de l'eau sont élevées. Une dernière remarque à propos de ce paramètre est que dans la plupart des cas, l'intervalle de confiance obtenu est étroit tel que la station Sainte-Anne (050408) qui correspond à l'intervalle de confiance du paramètre b_1 de $[26.44, 26.52]$. Cela semble adéquat puisqu'il n'est pas attendu d'avoir une forte incertitude du niveau de référence. Cependant, il y a quelques cas où les intervalles de confiance sont larges, notamment les deux stations mentionnées au début de ce paragraphe, Jacques-Cartier (050801) et Du Chêne (046709), la station Sainte-Anne (021407) au paramètre b_1 et la station Yamaska Nord (030309) aux paramètres b_1 et b_2 . Cela signifie qu'il est possible d'une part, que les données sur les hauteurs d'eau correspondent à des incertitudes importantes et d'autre part, que la méthode bootstrap surestime l'incertitude correspondante à ce paramètre pour ces cas.

En ce qui concerne le paramètre c , pour les stations de la première, deuxième, quatrième et cinquième catégorie, aucune valeur irréaliste n'est remarquable. En effet, la plupart des valeurs obtenues sont dans l'intervalle $[1, 4]$, sauf pour les deux stations Saint-Louis (040212) et Jacques-Cartier (050801), pour lesquelles des estimations respectives de $c_2 = 5.43$ avec un intervalle de confiance $[1.44, 130.85]$ et de $c_2 = 5.25$ avec un intervalle de confiance $[1.43, 16.55]$ ont été obtenues. Ainsi, ces deux stations ont eu des estimations du paramètre c un peu élevées. Ils ont d'ailleurs été discutées précédemment pour le paramètre a qui a eu une estimation très proche de 0, ce qui indique que ces deux stations présentent une configuration incertaine des données qui empêche le modèle d'ajuster des valeurs réalistes de ces deux paramètres. Cela est apparent dans les larges intervalles de confiance correspondants. De

plus, la borne supérieure de l'intervalle de confiance Jacques-Cartier (050801) semble très irréaliste et elle reflète des incertitudes importantes. Pour le cas des stations de la troisième catégorie, l'estimation pour les deux stations Sainte-Anne (021407) et Dauphine (120201) donne des valeurs loin d'être réalistes. Ces valeurs sont respectivement données par $c_1 = 185.30$ avec intervalle de confiance $[9.22, 186.97]$ et $c_1 = 312.07$ avec un intervalle de confiance $[3.38, 317.33]$. Ces valeurs correspondent à des intervalles de confiance irréalistes et trop larges. De même, ces deux stations ont également été discutées précédemment pour le paramètre a qui a eu une estimation de pratiquement 0. L'intervalle de confiance de la station Sainte-Anne (021407) n'arrive même pas à avoir une borne inférieure réaliste.

Les estimations larges et irréalistes pour les intervalles de confiance n'ont pas été données que pour les stations sujettes à une dispersion des données mais aussi pour plusieurs autres cas. Cela a été remarqué surtout au niveau du deuxième segment des courbes de tarage comme le cas de la station Jacques-Cartier (050801) mentionné précédemment. Il y a aussi l'exemple des deux stations Du Chêne (046709) et Yamaska Nord (030309) de la cinquième catégorie qui ont donné au deuxième segment des intervalles de confiance respectivement donnés par $[1.02, 133.07]$ et $[1.06, 158.03]$. Ces larges intervalles de confiance impliquent que la composante d'incertitude ε_3 reliée à une instabilité géomorphologique est élevée. En effet, le paramètre c dépend principalement de la forme de la section de contrôle [1].

Pour le point de rupture k , aucun cas n'a présenté de valeur inférieure à $\{\text{minimum}(h_i) : 1 \leq i \leq n\}$ ni de valeur supérieure à $\{\text{maximum}(h_i) : 1 \leq i \leq n\}$ et les intervalles de confiance sont étroits. Ce résultat n'est pas surprenant puisque l'intervalle d'optimisation vérifie ces deux conditions.

Finalement, une dernière remarque générale à propos des intervalles de confiance est que dans le cas des courbes à un seul segment, les intervalles de confiance sont beaucoup plus étroits et les bornes de ces intervalles sont plus réalistes. Par exemple, la station Yamaska Nord (030340) de la deuxième catégorie a comme estimations de paramètres : $a = 15.58$ avec un intervalle de confiance $[13.23, 16.76]$, $b = 93.37$ avec un intervalle de confiance $[93.31, 93.40]$ et $c = 2.50$ avec un intervalle de confiance $[2.36, 2.76]$.

6.3 Résultats des incertitudes sur les débits estimés par les courbes de tarage

Le but principal de ce mémoire est d'estimer les incertitudes sur les débits prédits par la courbe de tarage. À la section précédente, les courbes de tarage obtenues ainsi que les incertitudes des paramètres ont été présentées afin de valider le modèle. Dans la présente section, les incertitudes correspondantes des prédictions obtenues avec les courbes de tarage sont étu-

Tableau 6.2 Paramètres des courbes de tarage des stations de la première catégorie. Les intervalles de confiance pour les paramètres à 95% sont obtenus par la méthode bootstrap

Station	a_1	b_1	c_1	k	a_2	b_2	c_2
SainteAnne (050408)	68.29 ~ [62.09, 73.10]	26.48 ~ [26.44, 26.52]	1.84 ~ [1.72, 1.98]	27.77 ~ [27.51, 28.11]	112.97 ~ [100.52, 125.33]	26.80 ~ [26.71, 26.88]	1.39 ~ [1.34, 1.46]
Jacques-Cartier (050801)	6.17 ~ [1.46, 22.71]	27.13 ~ [26.55, 27.73]	2.82 ~ [1.93, 3.58]	30.98 ~ [29.57, 31.05]	0.001 ~ [e^{-319} , 82.17]	20.64 ~ [-245.33, 28.69]	5.25 ~ [1.44, 130.85]
Harricana (080101)	51.23 ~ [19.20, 63.22]	17.80 ~ [17.36, 17.92]	1.75 ~ [1.47, 2.64]	19.46 ~ [18.82, 19.87]	35.07 ~ [e^{-8} , 99.42]	17.46 ~ [6.56, 18.26]	1.82 ~ [1.22, 8.71]

diées.

Tel que mentionné à la section 5.2.3, l'incertitude sur le débit prédit dans l'espace logarithmique est donnée par la somme de la variance de prédiction de régression linéaire et la variance de l'incertitude mesure. Les limites de l'intervalle de prévision sur les débits prédits dans l'espace original (en m^3/s) sont données par l'exponentielle de ces limites dans l'espace logarithmique. Ensuite, les limites de l'intervalle de prévision sur les débits prédits présentent une enveloppe d'incertitude sur la courbe de tarage et sur les mesures correspondantes à la courbe. Cela est illustré dans l'exemple de la station Sainte-Anne (050408) à la figure 6.2 qui superpose la courbe de tarage et l'intervalle de prédiction à 95% pour chacune des hauteur d'eau. Les figures des autres stations sont présentées à l'annexe C.

De façon générale, la largeur de l'intervalle de prédiction augmente avec le débit, ce qui est normal car l'incertitude de mesure augmente également avec le débit (voir l'équation 5.1.) Pour les courbes à deux segments, il arrive parfois que l'incertitude du débit prédit diminue au point de rupture entre les deux segments (voir par exemple la figure C.4 de la station Saint-Louis (040212)). L'incertitude reprend son augmentation jusqu'au débit le plus élevé. De plus, la bande d'incertitude est plus étroite au deuxième segment que celle au premier segment. Pour les stations de la première catégorie, cette dernière configuration de bande d'incertitude est observable pour la station Harricana (080101) (figure C.2). Cependant, la station Sainte-Anne (050408) (figure 6.2) possède une bande plus uniforme sur les trois gammes de débit. La station Jacques-Cartier (050801) possède une bande légèrement plus large au premier segment. Cela semble incohérent avec ce qui a été remarqué au niveau des intervalles de prévision très larges des paramètres de la courbe de cette station au deuxième segment.

Ensuite, pour les stations de la deuxième catégorie, les bandes d'incertitude présentent une configuration particulière pour chaque station. La station Yamaska Nord (030340) (figure C.3) possède une bande d'incertitude plus uniforme et large sur les deux segments que les deux

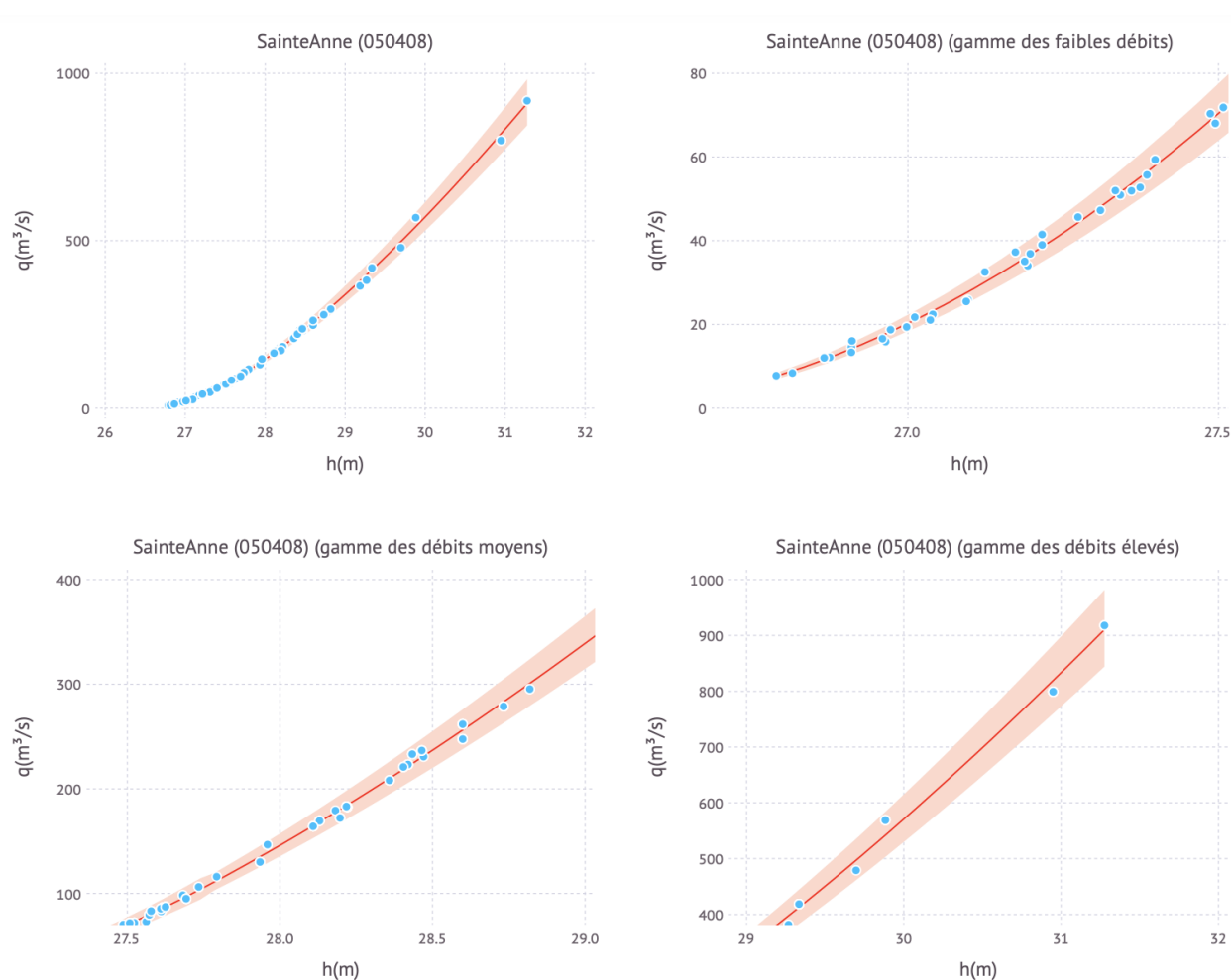


Figure 6.2 Allure des incertitudes sur la courbe de tarage de la station Sainte-Anne (050408), la courbe rouge illustre la courbe de tarage ajustée sur les jaugeages et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage

autres stations. La station Saint-Louis (040212) (figure C.4) possède une bande d'incertitude beaucoup plus étroite au deuxième segment avec quelques mesures dispersées autour de la bande d'incertitude à la gamme des faibles débits. Ce résultat est incohérent avec les intervalles de confiance des paramètres de la courbe qui sont plus larges au deuxième segment. La station Des Escoumins (070204) (figure C.5) possède une bande d'incertitude plus étroite que les deux autres stations avec une partie un peu plus large juste avant le point de rupture. Par la suite, les stations de la troisième catégorie ont eu des bandes d'incertitude les plus larges. Notamment, la station Dauphine (120201) (figure C.7) a la largeur la plus grande. Cela est non surprenant puisque ces stations sont sujettes à une dispersion importante des données. De plus, ce résultat est cohérent aux intervalles de confiance larges des paramètres

de leurs courbes de tarage. Il est visible aussi que la bande d'incertitude est plus large au premier segment qu'au deuxième segment dans le cas des stations Sainte-Anne (021407) (figure C.6) et Boyer Sud (023002) (figure C.8). Cela est cohérent à ce qui a été observé à l'affichage des données que la dispersion plus importante aux gammes des faibles débits et des débits moyens. La station Dauphine (120201) (figure C.7) a une très large bande d'incertitude sur les trois gammes notamment aux débits élevés (sauf la partie près du point de rupture est un plus étroite). Cette configuration confirme aussi ce qui a été observé à l'affichage des données de la dispersion des données plus importantes aux gammes de débits moyens et des débits élevés. Une dernière remarque à propos des stations de cette catégorie est l'existence de quelques jaugeages dispersés autour de la bande d'incertitude.

Les stations de la quatrième catégorie possèdent les bandes d'incertitude les plus étroites. Pour les deux stations Nicolet (030120) (figure C.9) et Ha! Ha! (060601) (figure C.11), la bande d'incertitude est plus uniforme que pour la station Capitachouane (043030) (figure C.10) sur les trois gammes. Cela est non surprenant puisque l'ajustement de leurs courbes a été effectué à un segment. La station Capitachouane (043030) (figure C.10) possède la même configuration de bandes d'incertitudes que celles des courbes à deux segments

Enfin, les bandes d'incertitudes des stations de la cinquième catégorie ont des allures similaires aux autres stations à deux segments, mais avec une largeur plus importantes pour la gamme des débits élevés. La station Du Chêne (046709) (figure C.13), qui possède une courbe qui sous-estime les débits élevés, a une bande d'incertitude qui englobe le jaugeage du débit le plus élevé. Cependant, la largeur de la bande des incertitudes n'est pas de la largeur attendue par rapport aux largeurs des intervalles de confiance des paramètres obtenues de sa courbe au deuxième segment. La station Yamaska Nord (030309) (figure C.12) a eu une large bande d'incertitudes aux faibles débits et aux débits élevés. Cela est cohérent avec les largeurs des intervalles de confiance sur les paramètres.

6.4 Discussion

Cette section évalue les résultats obtenus dans ce mémoire et elle discute la validité de la méthode employée.

6.4.1 Évaluation du modèle de calcul des incertitudes correspondantes aux débits prédits par les courbes de tarage

Le modèle développé dans la section 5 qui combine les méthodes d'optimisation des moindres carrés et L-BFGS afin de construire des intervalles de prévisions sur les débits prédits à partir

des courbes de tarage a réussi à atteindre le but principal de ce mémoire. À titre de rappel, ce but consiste à estimer les incertitudes sur les débits estimés par les courbes de tarage des 15 stations hydrométriques choisies.

D'abord, la méthodologie développée permet d'intégrer les différentes sources d'incertitude présentées dans l'équation 2.4 dans l'estimation du débit. En effet, la méthode développée permet d'estimer d'une part les composantes d'incertitudes ε_2 et ε_4 et d'autre part la composante ε_1 a été ajouté au modèle. Pour la composante ε_3 , il est clair que les stations de la troisième catégorie sujette à cette composante ont les bandes d'incertitudes les plus large. Cela confirme d'une part l'hypothèse que c'est la seule catégorie qui comporte cette composante et d'autre part que le modèle est capable d'estimer implicitement cette composante. Cependant, les quelques jaugeages observés hors de la bande d'incertitude montrent la possibilité d'une faiblesse dans la précision de l'estimation. Par ailleurs, comme mentionné à la section 4, cette composante peut exister dans les stations des autres catégories mais avec une faible amplitude. Cela a été illustrée dans les stations qui comportent une légère dispersion de données comme la station Saint-Louis (040212) (figure C.4) à la gamme des faibles débits et comme la station Yamaska Nord (030340) (figure C.3). Ces deux stations ont des bandes d'incertitude plus larges là où il y a la dispersion de données ce qui confirme l'efficacité de la méthode à estimer cette composante d'incertitude.

Ensuite, les résultats obtenus aux sections 6.3 et 6.2 ont indiqué que les données de jaugeages de ces 15 stations sont sujettes à des incertitudes considérables. Cela est à cause des larges bandes d'incertitudes obtenues pour la plupart des cas et à cause des larges intervalles de prévisions qui correspondents aux paramètres des courbes ajustées. Comme mentionné précédemment, la grande base de données des stations hygrométriques du Québec méridional comporte une classification préliminaire des stations selon leurs différentes caractéristiques. Selon cette classification, il est attendu d'avoir une configuration commune de l'amplitude des incertitudes dans les stations sélectionnées. Cependant, d'après les résultats obtenus, il n'y a pas vraiment une relation solide entre la catégorie et l'amplitude des incertitudes pour toutes les catégories. Il est vrai que dans le cas des stations de la troisième catégorie et de la cinquième catégorie, il y a une amplitude d'incertitude proportionnellement similaire, mais dans le cas des stations de la première, de la deuxième catégorie et de la quatrième catégorie, chaque station possède sa propre amplitude d'incertitude. Par conséquent, il n'est pas possible d'avoir une interprétation générale des résultats, mais plutôt il est plus convenable de les traiter au cas par cas.

Par contre, il est clair qu'il y a une tendance similaire des incertitudes entre les stations de courbes de tarage ajustées à deux segments et celles qui sont ajustées à un seul segment.

En effet, dans le cas de traitement à deux segments, le deuxième segment a des incertitudes moins importantes avec une diminution brusque des incertitudes au point de rupture. Toutefois, cette tendance d'incertitude ne peut pas être affirmée avec confiance. En effet, elle pourrait être expliquée par deux raisons. D'une part, il est possible que le modèle sous-estime les incertitudes à cause de la diminution de la variance σ au deuxième segment. Cette diminution est reliée au fait que le volume le plus important des données est au premier segment notamment à la gamme des faibles débits. D'autre part, il est possible que la segmentation améliore la courbe de tarage et donc les incertitudes diminuent au deuxième segment. À vrai dire, d'après ce qui a été observé que la bande d'incertitude inclut bien la plupart des jaugeages, il est possible de conclure que le modèle a réussi à prédire le mieux possible la tendance des incertitudes mais avec une limitation causée par la diminution de la variance au deuxième segment. De plus, l'obtention des incertitudes plus importantes au premier segment notamment à la gamme des faibles débits est adéquat à ce qui a été obtenu dans des études similaires (voir [4,5]). Ces études affirment que la gamme des faibles débits est la plus affectée par les incertitudes à cause des difficultés des instruments de jaugeages à mesurer les faibles débits. Il est possible alors de conclure que la partie la plus affectée par les incertitudes dans le cas des stations hydrométriques étudiées dans ce mémoire est celle de leurs premiers segments de courbes de tarage. Dans des travaux futurs, il pourrait être approprié de traiter d'un seul coup l'incertitude de prédiction plutôt que segment par segment tel que proposé ici. Peut-être que le traitement global assurerait une certaine continuité de l'incertitude au point de rupture.

Finalement, il est possible de tirer des conclusions spécifiques pour quelques catégories. En effet, les résultats des stations de la deuxième catégorie n'ont pas de configuration spécifique à la gamme des débits élevés. Par conséquent, la catégorisation de station à un faible nombre de jaugeage aux débits élevés n'a pas donné une influence particulière. Cela est peut-être à cause que les stations choisies ne vérifient pas vraiment cette idée. Cela a été constaté dans l'analyse exploratoire des données à la section 4. Il est possible alors que ces stations reviennent à la même classification de la première catégorie. Pour traiter cette catégorisation pour des futurs travaux, il est possible de chercher des cas de station qui vérifient mieux le critère de la deuxième catégorie. Par ailleurs, les stations de la quatrième catégorie ont les incertitudes les plus faibles. Cela suggère alors que lorsque le nombre de jaugeages est faible, les incertitudes sont de plus en plus faibles. En particulier, cette hypothèse n'est valide que lorsque les données ne sont pas sujettes à une grande dispersion. Un autre aspect qui a été observé est que les stations de la cinquième catégorie ont eu des incertitudes importantes à la gamme des débits élevés. Cela suggère que la présence d'une plaine inondable où d'une discontinuité du lit de la rivière exerce une grande influence sur la qualité des données dans

cette gamme et que des scénarios des inondations importantes peuvent être rencontrés au niveau de ces stations.

6.4.2 Évaluation du modèle de calcul des incertitudes correspondantes aux paramètres des courbes de tarage

La méthode développée et décrite dans la section 5 est adaptée d'une part pour l'ajustement des courbes de tarage aux jaugeages et d'autre part pour estimer l'incertitude des paramètres. Cela est dû principalement à l'efficacité de la méthode des moindres carrés à ajuster adéquatement la fonction de puissance (éq. 2.2) dans sa forme logarithmique. De ce point de vue, le premier et le troisième sous objectifs de ce mémoire sont atteints.

Par contre, il est intéressant d'évaluer la réussite de ce modèle à donner des résultats réalistes, c'est à dire des valeurs de paramètres qui possèdent un sens du point de vue de l'hydraulique (précisées dans [1]). En effet, cette vérification permet de qualifier la validité de l'ajustement des jaugeages des stations hydrométriques étudiées ainsi que les échantillons bootstrap par la fonction de puissance dans sa forme 5.2 et dans sa forme 5.4 pour le cas de courbe multi-segments. Les résultats des paramètres correspondants aux courbes de tarage ont montré que la méthode est plus efficace dans le cas des stations qui ne sont pas sujettes à une sur-dispersion des données. En effet, l'estimation pour les stations de la troisième catégorie, sujette à cette dernière contrainte, est celle qui a donné des valeurs irréalistes des paramètres a et c . Le modèle a tendance à amplifier le paramètre c et à négliger le paramètre a . Cette tendance est remarquée aussi au deuxième segment des courbes des deux stations supposées stables Jacques-Cartier (050801) et Saint-Louis (040212), mais avec un effet moins important.

De même, dans le cas de courbe de tarage à deux segment, l'estimation des intervalles de confiance des paramètres a , c et même le paramètre b , pour lequel il y a eu deux valeurs très faibles, n'est pas réaliste pour toutes les stations. Cela se produit au niveau de ces deux dernières stations (Jacques-Cartier (050801) et Saint-Louis (040212)), des stations aux relations hauteur-débit atypiques et des stations à dispersion de données. Ces bornes irréalistes des intervalles de confiance peuvent être expliquées par la diminution du nombre total des points de jaugeage engendrée par l'échantillonnage bootstrap. En effet, il existe une absence de couples de jaugeage dans des gammes de débits par rapport aux autres qui se manifeste par des écarts observables entre deux points de jaugeages consécutifs. Cette inégalité de distribution de points entre les gammes du débits a des conséquences plus remarquables dans la qualité de la courbe lors d'un ajustement par segment. En effet, le nombre total de données par segment est beaucoup plus faible que celui d'un échantillon au complet, d'où la diminution de la qualité de l'optimisation. À titre d'exemple, la figure 6.3 illustre l'absence de point

de jaugeages au deuxième segment à la gamme de débit ($7 m^3/s, 12 m^3/s$). Les estimations des paramètres de courbe de tarage de cet échantillon bootstrap sont les suivantes :

$$\hat{a}_1 = 13.15 \quad \hat{b}_1 = 29.81 \quad \hat{c}_1 = 3.41 \quad \hat{k} = 30.70$$

$$\hat{a}_2 = e-319 \quad \hat{b}_2 = -81.17 \quad \text{et} \quad \hat{c}_2 = 155.90$$

Il est remarquable que le deuxième segment sujet à cette dernière contrainte est celui qui a des valeurs irréalistes des trois paramètres a_2 , b_2 et c_2 .

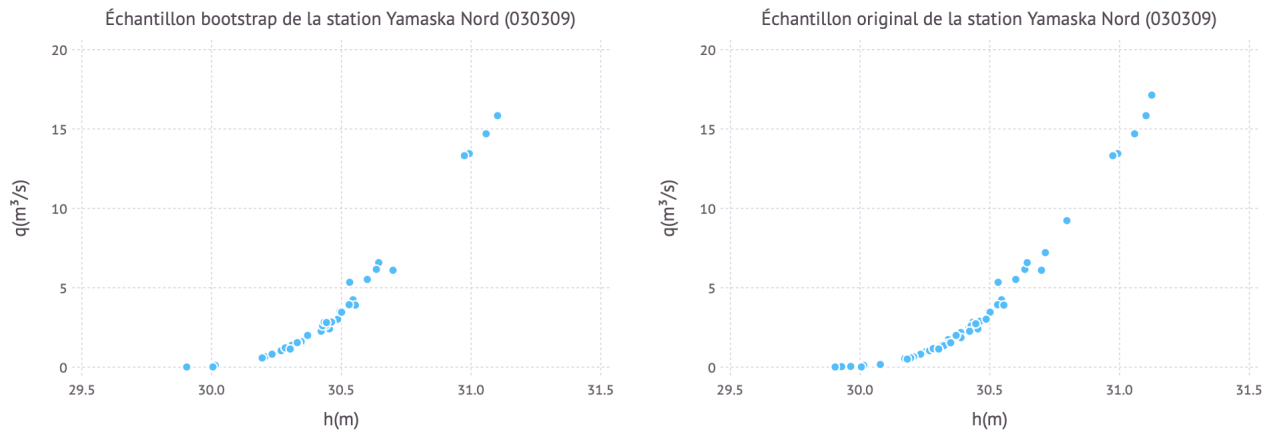


Figure 6.3 Comparaison d'un échantillon bootstrap à son échantillon original de la station Yamaska Nord (030309)

Les valeurs irréalistes obtenues, que ce soit au niveau des paramètres des courbes de tarage ou bien de leurs intervalles de confiance, peuvent être améliorées dans des futurs travaux en utilisant une optimisation sous contraintes. En effet, même si l'intervalle d'optimisation inclut des bornes basées sur la théorie hydraulique des courbes tarage, l'ajout des contraintes sur les paramètres imposera au modèle de trouver des estimations plus raisonnables au sens hydraulique. Une première contrainte peut être imposée sur le paramètre c pour que sa valeur ne soit pas trop loin des bornes 0.5 et 2.5. Par exemple, l'ajout de cette contrainte dans le cas de la station Sainte-Anne (021407), qui a eu une estimation donnée par $c_1 = 185.30$, va diminuer cette valeur. Elle va également augmenter l'estimation très faible obtenue pour le paramètre a_1 sinon les débits prédits seront sous-estimés. En effet, l'ordre de grandeur de ces deux paramètres dépend l'un de l'autre. Toutefois, il est possible qu'une contrainte seulement sur un des deux paramètres ne satisfasse pas encore les conditions hydrauliques. Par conséquent, une deuxième contrainte sur le paramètre a pourrait améliorer les résultats.

Enfin, pour le paramètre b , il est possible d'ajouter une troisième contrainte sur sa valeur en se basant sur la connaissance de la topographie de chaque station traitée.

CHAPITRE 7 CONCLUSION ET RECOMMANDATIONS

7.1 Synthèse des travaux

Ces travaux ont permis de développer une approche d'estimation des paramètres de la courbe de tarage ainsi que d'estimer les incertitudes qui y sont associées. L'approche intègre les différentes sources d'incertitudes, basée sur la relation hydraulique hauteur-débit et sur la méthode des moindres carrées. La méthode a été appliquée à 15 stations aux caractéristiques et aux localisations variées du réseau hydrométrique exploité par le MELCC.

La première étape a été d'introduire les bases hydrauliques de la relation hauteur-débit et les méthodes statistiques essentielles pour l'estimer par *régression hydraulique* ainsi que les intervalles de confiance correspondantes. Puis, une présentation de la méthode bootstrap utilisée pour estimer les incertitudes correspondantes aux paramètres de la courbe de tarage a été effectuée.

Pour la deuxième étape, il a été nécessaire d'effectuer une revue des études antérieures afin de choisir une méthode spécialement adaptée pour les données du Québec. Pour ce faire, les méthodes qui se basent sur l'estimation des courbes de tarage par la relation hydraulique et celles qui les estiment par des méthodes non paramétriques ont été présentées. Cette étape conclut que la méthode ISO/WMO [13, 15, 33] est adaptée pour ce mémoire grâce à la simplicité de son implémentation, mais des améliorations selon les composantes d'incertitudes correspondantes aux données doivent être effectuées.

La troisième étape était constituée d'une analyse exploratoire des données de jaugeages étudiées. De cette analyse, 15 stations ont été sélectionnées selon cinq catégories qui couvrent les différentes caractéristiques du réseau hydrométrique du Québec méridional. Globalement, les stations hydrométriques étudiées comportent un nombre modéré de jaugeages distribués sur toutes les gammes.

L'étape finale a été de développer la méthode de calcul de la relation hauteur-débit et des incertitudes. D'abord, la relation hauteur-débit dans l'espace logarithmique a été ajustée par la méthode des moindres carrées. Ensuite, les incertitudes sont estimées par les intervalles de prédiction sur les débits estimés par cette relation tout en ajoutant la variance de l'incertitude de mesure du débit à ces intervalles. Puis, la méthode bootstrap non paramétrique a présenté une solution pour estimer les incertitudes sur les paramètres qui n'ont pas de fonctions de répartition connues. Les résultats ont montré que la base de données traitée est sujette à des incertitudes considérables notamment les stations instables géomorphologiquement.

En conclusion, la méthode développée dans ce mémoire a permis d'améliorer l'estimation des courbes de tarage des stations traitées, d'estimer les incertitudes d'ajustement de ces courbes, les incertitudes du modèle de la relation hauteur-débit et les incertitudes de mesure. Elle a permis aussi d'estimer au mieux possible des instabilités géomorphologiques. Par conséquent, cette méthode améliore la méthode ISO/WMO [13, 15, 33] par l'ajout de la variance de l'incertitude de mesure du débit. Cette méthode appliquée sur 15 stations hydrométriques peut être appliquée facilement sur les autres stations du réseau hydrométrique du Québec méridional.

7.2 Limitations de la solution proposée et améliorations futures

L'approche développée dans ce mémoire est sujette à des limitations. D'abord, dans le cas de courbes multi-segments, le modèle semble sous-estimer les incertitudes du deuxième segment en raison de la diminution de la deuxième variance de l'erreur. Cette limitation ne peut être résolue facilement à l'heure actuelle car cela nécessiterait le développement d'une approche globale plus complexe. Néanmoins, bien que l'incertitude soit discontinue au point de rupture des segments, l'incertitude estimée semble généralement adéquate car les jaugeages se trouvent à peu près tous dans l'intervalle de prédiction.

Ensuite, le modèle n'inclut pas la composante d'incertitude reliée aux instabilités géomorphologiques. D'ailleurs, la conclusion tirée que le modèle est capable d'approximer cette composante implicitement a été retenue après l'interprétation des résultats des stations instables géomorphologiquement. Il est possible dans des prochaines travaux de considérer cette composante avec une approche bayésienne ou bien avec une approche similaire à celle proposée par [18, 39]. Le modèle néglige aussi les erreurs d'approximation des paramètres k et b dans l'estimation des incertitudes correspondantes aux débits prédits. Cette limitation peut être résolue dans des futurs travaux par des algorithmes plus complexes pour estimer ces erreurs. Par contre, l'incertitude correspondantes à ces paramètres a été bien estimée par la méthode bootstrap.

Par la suite, dans le cas des stations sujettes à une dispersion de données et à des incertitudes importantes, le modèle a estimé des paramètres de courbes de tarage irréalistes. Cela a été aussi observé pour plusieurs bornes des intervalles de prévisions. Cette limitation peut être améliorée dans le futur par une optimisation sous des contraintes imposées sur ces paramètres. Par ailleurs, il est possible que l'approche bayésienne estime des paramètres plus réalistes. En effet, des études [1, 4, 34, 46] se basent dans leurs lois *a priori* des paramètres sur les connaissances hydrauliques de la rivière. Par conséquent, leurs estimations sont adéquates. Cela peut être également une solution pour la sous-estimation des incertitudes au deuxième

segment. En outre, il est possible d'utiliser l'approche non paramétrique *LOESS* [5] dans le cas des stations de la troisième catégorie. En effet, selon [5], la validité de la relation hydraulique paramétrique n'est pas garantie dans le cas de données de jaugeages dispersées et elle propose cette méthode comme solution.

Finalement, des améliorations pourraient être apportées sur la méthode *bootstrap non paramétrique*. Cela en utilisant la méthode *bootstrap non paramétrique stratifiée* [47]. En effet, le principe de la stratification est d'effectuer un échantillonnage sur des sous-échantillons homogènes du jeu de données principal. Ces sous-échantillons peuvent être les trois gammes de débits (faibles, moyens et élevés). Par conséquent, il serait plus garanti d'avoir le plus possible des couples de jaugeages sur toute la courbe.

RÉFÉRENCES

- [1] J. Le Coz, B. Renard, L. Bonnifait, F. Branger et R. Le Boursicaud, “Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves : A bayesian approach,” *Journal of Hydrology*, vol. 509, p. 573–587, 2014. [En ligne]. Disponible : <https://doi.org/10.1016/j.jhydrol.2013.11.016>
- [2] J. Le Coz, B. Camenen, X. Peyrard et G. Dramais, “Uncertainty in open-channel discharges measured with the velocity–area method,” *Flow Measurement and Instrumentation*, vol. 26, p. 18–29, 2012. [En ligne]. Disponible : <https://doi.org/10.1016/j.flowmeasinst.2012.05.001>
- [3] G. James, D. Witten, T. Hastie et R. Tibshirani, *An introduction to statistical learning*. Springer, 2013, vol. 112. [En ligne]. Disponible : <https://link.springer.com/book/10.1007/978-1-4614-7138-7>
- [4] A. Petersen-Øverleir, A. Soot et T. Reitan, “Bayesian rating curve inference as a streamflow data quality assessment tool,” *Water Resources Management*, vol. 23, n^o. 9, p. 1835–1842, 2009. [En ligne]. Disponible : <https://doi.org/10.1007/s11269-008-9354-5>
- [5] G. Coxon, J. Freer, I. K. Westerberg, T. Wagener, R. Woods et P. Smith, “A novel framework for discharge uncertainty quantification applied to 500 uk gauging stations,” *Water resources research*, vol. 51, n^o. 7, p. 5531–5546, 2015. [En ligne]. Disponible : <https://doi.org/10.1002/2014WR016532>
- [6] H. McMillan, T. Krueger et J. Freer, “Benchmarking observational uncertainties for hydrology : rainfall, river discharge and water quality,” *Hydrological Processes*, vol. 26, n^o. 26, p. 4078–4111, 2012. [En ligne]. Disponible : "<https://doi.org/10.1002/hyp.9384>"
- [7] D. P. Turnipseed et V. B. Sauer, “Discharge measurements at gaging stations,” US Geological Survey, Rapport technique, 2010.
- [8] J. Le Coz, F. Larrarte, G. SAYSSET, G. Pierrefeu, J.-F. Brochot et P. Marchand, “Mesures hydrologiques par profileur à effet doppler (adcp) en france : application aux cours d’eau et aux réseaux urbains,” *La Houille Blanche*, n^o. 3, p. 115–122, 2009.
- [9] R. W. Herschy, *Streamflow measurement*. CRC press, 2008.
- [10] S. Rantz, “Measurement and computation of streamflow. vol. 1-measurement of stage and discharge, usgs water supply paper 2175, vol 1, 313,” 1982. [En ligne]. Disponible : <https://pubs.usgs.gov/wsp/wsp2175/wsp2175.pdf4>
- [11] A. R. Schmidt, *Analysis of stage-discharge relations for open-channel flows and their associated uncertainties*. University of Illinois at Urbana-Champaign, 2002.

- [12] WMO, *Manual on Stream Gauging. Computation of discharge*. WMO-No. 1044, vol. II : World Meteorological Organization, 2010, p. 195. [En ligne]. Disponible : https://library.wmo.int/index.php?lvl=notice_display&id=538#.YNPHZC-cY_U
- [13] *Hydrometry — Determination of the stage-discharge relationship (Revision of ISO 1100-2 :2010)*, Norme ISO/PWI 18 320, 2015.
- [14] R. Manning, J. P. Griffith, T. Pigot et L. F. Vernon-Harcourt, *On the flow of water in open channels and pipes*, 1890.
- [15] *Hydrometry – measurement of liquid flow in open channels – Part 2 : Determination of the stage-discharge relationship*, Norme ISO 1100-2, 2010.
- [16] J. Le Coz, B. Camenen, G. Dramais, M. Ferry, J. Rosique et J. Ribot-Bruno, “Contrôle des débits réglementaires, application de l’article l. 214-18 du code de l’environnement,” *Guide pratique Onema/Cemagref*, 2011.
- [17] R. T. Clarke, “Uncertainty in the estimation of mean annual flood due to rating-curve indefintion,” *Journal of Hydrology*, vol. 222, n°. 1-4, p. 185–190, 1999.
- [18] T. Morlot, C. Perret, A.-C. Favre et J. Jalbert, “Dynamic rating curve assessment for hydrometric stations and computation of the associated uncertainties : Quality and station management indicators,” *Journal of Hydrology*, vol. 517, p. 173–186, 2014. [En ligne]. Disponible : <https://doi.org/10.1016/j.jhydrol.2014.05.007>
- [19] H. McMillan, J. Seibert, A. Petersen-Overleir, M. Lang, P. White, T. Snelder, K. Rutherford, T. Krueger, R. Mason et J. Kiang, “How uncertainty analysis of streamflow data can reduce costs and promote robust decisions in water management applications,” *Water Resources Research*, vol. 53, n°. 7, p. 5220–5228, 2017. [En ligne]. Disponible : "<https://doi.org/10.1002/2016WR020328>"
- [20] “Actions d’ouranos dans info-crue cartographie dynamique des zones inondables dans un contexte de changements climatiques,” Rapport technique, 2018. [En ligne]. Disponible : https://www.ouranos.ca/wp-content/uploads/Fiche_projet_INFO_Crue.pdf
- [21] P. Saffache, *Dictionnaire simplifié de la géographie*. Editions Publibook, 2003.
- [22] P.-L. Viollet, J.-P. Chabard et P. Esposito, *Mécanique des fluides appliquée : écoulements incompressibles dans les circuits, canaux et rivières, autour des structures et dans l’environnement*. Presses des Ponts, 2003.
- [23] A. Petersen-Overleir et T. Reitan, “Objective segmentation in compound rating curves,” *Journal of Hydrology*, vol. 311, n°. 1-4, p. 188–201, 2005. [En ligne]. Disponible : <https://doi.org/10.1016/j.jhydrol.2005.01.016>
- [24] J. R. Dymond et R. Christian, “Accuracy of discharge determined from a rating curve,” *Hydrological Sciences Journal*, vol. 27, n°. 4, p. 493–504, 1982.

- [25] H. McMillan, T. Krueger et J. Freer, “Benchmarking observational uncertainties for hydrology : rainfall, river discharge and water quality,” *Hydrological Processes*, vol. 26, n°. 26, p. 4078–4111, 2012.
- [26] S. Rantz, *Hydrometry, Principles and Practices 2nd ed.*, 1999, p. 376.
- [27] T. Hastie, R. Tibshirani et J. Friedman, “The elements of statistical learning,” *Cited on*, p. 33, 2009.
- [28] G. Schwarz, “Estimating the dimension of a model,” *The annals of statistics*, p. 461–464, 1978.
- [29] R. E. Kass et A. E. Raftery, “Bayes factors,” *Journal of the american statistical association*, vol. 90, n°. 430, p. 773–795, 1995.
- [30] B. Efron, “Bootstrap Methods : Another Look at the Jackknife,” *The Annals of Statistics*, vol. 7, n°. 1, p. 1 – 26, 1979. [En ligne]. Disponible : <https://doi.org/10.1214/aos/1176344552>
- [31] B. Efron et R. J. Tibshirani, *An introduction to the bootstrap*. CRC press, 1994.
- [32] R. W. Herschy, “Hydrometry : principles and practices,” Rapport technique, 1999.
- [33] WMO, *Technical regulations, Vol. III Hydrology, Annex VII — Determination of the stage-discharge relation*. World Meteorological Organization, 2005, p. 110. [En ligne]. Disponible : https://library.wmo.int/doc_num.php?explnum_id=4564
- [34] A. E. Sikorska, A. Scheidegger, K. Banasik et J. Rieckermann, “Considering rating curve uncertainty in water level predictions,” *Hydrology and Earth System Sciences*, vol. 17, n°. 11, p. 4415–4427, 2013.
- [35] H. McMillan et I. Westerberg, “Rating curve estimation under epistemic uncertainty,” *Hydrological Processes*, vol. 29, n°. 7, p. 1873–1882, 2015. [En ligne]. Disponible : "<https://doi.org/10.1002/hyp.10419>"
- [36] J. D. Fenton, “Generating stream rating information from data,” *Alternative Hydraulics Paper*, vol. 8, 2015.
- [37] —, “On the generation of stream rating curves,” *Journal of Hydrology*, vol. 564, p. 748–757, 2018.
- [38] W. S. Cleveland, “Robust locally weighted regression and smoothing scatterplots,” *Journal of the American statistical association*, vol. 74, n°. 368, p. 829–836, 1979. [En ligne]. Disponible : <https://www.tandfonline.com/doi/abs/10.1080/01621459.1979.10481038>

- [39] J. Jalbert, T. Mathevet et A.-C. Favre, “Temporal uncertainty estimation of discharges from rating curves using a variographic analysis,” *Journal of Hydrology*, vol. 397, n^o. 1-2, p. 83–92, 2011.
- [40] J. E. Kiang, C. Gazoorian, H. McMillan, G. Coxon, J. Le Coz, I. K. Westerberg, A. Belleville, D. Sevrez, A. E. Sikorska, A. Petersen-Øverleir *et al.*, “A comparison of methods for streamflow uncertainty estimation,” *Water Resources Research*, vol. 54, n^o. 10, p. 7149–7176, 2018. [En ligne]. Disponible : "<https://doi.org/10.1029/2018WR022708>"
- [41] Ministère du Développement durable, Environnement et Lutte contre les changements climatiques. Le réseau hydrométrique québécois. [En ligne]. Disponible : <https://www.cehq.gouv.qc.ca/hydrometrie/reseau/index.htm>
- [42] A. Petersen-Øverleir, “Accounting for heteroscedasticity in rating curve estimates,” *Journal of Hydrology*, vol. 292, n^o. 1-4, p. 173–181, 2004.
- [43] W. W. M. Organization), “Guide to hydrological practices. volume i : Hydrology—from measurement to hydrological information,” *WMO Report No. 168*, p. 296, 2008.
- [44] J. Nocedal et S. Wright, *Numerical optimization*. Springer Science & Business Media, 2006.
- [45] G. W. Oehlert, “A note on the delta method,” *The American Statistician*, vol. 46, n^o. 1, p. 27–29, 1992. [En ligne]. Disponible : <http://www.jstor.org/stable/2684406>
- [46] A. E. Sikorska, D. Del Giudice, K. Banasik et J. Rieckermann, “The value of streamflow data in improving tss predictions—bayesian multi-objective calibration,” *Journal of Hydrology*, vol. 530, p. 241–254, 2015.
- [47] A. C. Davison et D. V. Hinkley, *Bootstrap methods and their application*. Cambridge university press, 1997, n^o. 1.

ANNEXE A CARACTÉRISTIQUES DES STATIONS HYDROMÉTRIQUES ÉTUDIÉES

Tableau A.1 Tableau récapitulatif des caractéristiques des stations de la première catégorie et les composantes d'incertitudes associées à cette catégorie

Station	Période de jaugeage	Nombre de jaugeages	Composantes d'incertitudes
SainteAnne (050408)	1965-2020	192	ε_1
Jacques-Cartier (050801)	1925-2018	93	ε_2
Harricana (080101)	1944-2019	193	et ε_4

Tableau A.2 Tableau récapitulatif des caractéristiques des stations de la deuxième catégorie et les composantes d'incertitudes associées à cette catégorie

Station	Période de jaugeage	Nombre de jaugeages	Composantes d'incertitudes
Yamaska Nord (030340)	1983-2019	137	ε_1
Saint-Louis (040212)	1968-2019	280	ε_2
Des Escoumins (070204)	1992-2018	53	et ε_4

Tableau A.3 Tableau récapitulatif des caractéristiques des stations de la troisième catégorie et les composantes d'incertitudes associées à cette catégorie

Station	Période de jaugeage	Nombre de jaugeages	Composantes d'incertitudes
Sainte-Anne (021407)	1978-2020	283	ε_1
Dauphine (120201)	1970-2020	441	ε_2
Boyer Sud (023002)	1993-2019	210	ε_3 et ε_4

Tableau A.4 Tableau récapitulatif des caractéristiques des stations de la quatrième catégorie et les composantes d'incertitudes associées à cette catégorie

Station	Période de jaugeage	Nombre de jaugeages	Composantes d'incertitudes
Nicolet (030120)	2020/05-2020/08	6	ε_1
Capitachouane (043030)	2005-2020	37	ε_2
Ha! Ha! (060601)	1977-2020	46	et ε_4

Tableau A.5 Tableau récapitulatif des caractéristiques des stations de la cinquième catégorie et les composantes d'incertitudes associées à cette catégorie

Station	Période de jaugeage	Nombre de jaugeages	Composantes d'incertitudes
Yamaska Nord (030309)	1977-2020	139	ε_1
Du Chêne (046709)	2010-2020	90	ε_2
Suffolk (040409)	1976-2019	172	et ε_4

ANNEXE B RÉSULTATS DES PARAMÈTRES DES COURBES DE TARAGE

Tableau B.1 Paramètres des courbes de tarage des stations de la deuxième catégorie. Les intervalles de confiance pour les paramètres à 95% sont obtenus par la méthode bootstrap. La courbe de tarage de la station Yamaska Nord (030340) est ajustée à un seul segment.

Station	a_1	b_1	c_1	k	a_2	b_2	c_2
Yamaska Nord (030340)	15.58 ~ [13.23 , 16.76]	93.37 ~ [93.31, 93.40]	2.50 ~ [2.36, 2.76]	-	-	-	-
Saint-Louis (040212)	11.95 ~ [9.98, 14.19]	29.23 ~ [29.09, 29.24]	2.84 ~ [2.70, 4.51]	29.91 ~ [29.51,29.92]	0.15 ~ [e^{-8} 14.46]	28.09 ~ [25.54, 29.42]	5.43 ~ [1.86,12.28]
Des Escoumins (070204)	30.90 ~ [25.48,47.37]	29.22 ~ [28.90, 29.25]	1.23 ~ [1.02, 3.14]	29.57 ~ [29.55,30.33]	67.21 ~ [e^{-16} , 70.88]	29.31 ~ [18.60, 29.35]	1.53 ~ [1.43,16.55]

Tableau B.2 Paramètres des courbes de tarage des stations de la troisième catégorie. Les intervalles de confiance pour les paramètres à 95% sont obtenus par la méthode bootstrap.

Station	a_1	b_1	c_1	k	a_2	b_2	c_2
Sainte-Anne (021407)	e^{-322} ~ [e^{-324}, e^5]	-27.89 ~ [-43.39, 22.86]	185.30 ~ [9.22,186.97]	27.25 ~ [26.92, 27.34]	29.54 ~ [0.45,87.75]	26.11 ~ [24.47, 26.69]	2.31 ~ [1.45, 4.52]
Dauphine (120201)	e^{-323} ~ [$e^{-323}, 1137.5$]	18.23 ~ [16.85, 28.76]	312.07 ~ [3.38, 317.33]	28.91 ~ [28.88, 29.07]	20.58 ~ [17.99, 21.73]	28.83 ~ [28.82, 28.89]	2.10 ~ [1.66, 2.23]
Boyer Sud (023002)	11.77 [0.0003, 14.29]	47.62 [46.44, 47.70]	3.42 [2.37,16.71]	48.26 [47.98, 48.31]	12.01 [8.76, 15.35]	47.81 [47.67, 47.94]	1.95 [1.62, 2.33]

Tableau B.3 Paramètres des courbes de tarage des stations de la quatrième catégorie. Les intervalles de confiance pour les paramètres à 95% sont obtenus par la méthode bootstrap. Les courbes de tarage des stations Nicolet (030120) et Ha! Ha! (060601) sont ajustées à un seul segment.

Station	a_1	b_1	c_1	k	a_2	b_2	c_2
Nicolet (030120)	79.12 [71.26,101.93]	47.58 [47.32, 47.59]	1.97 [1.90,4.44]	-	-	-	-
Capitachouane (043030)	26.55 [0.002,29.51]	48.11 [46.21, 48.16]	2.00 [1.81,9.11]	49.40 [48.91,49.65]	53.47 [4.04,105.75]	48.51 [47.30,49.08]	1.52 [0.93, 3.19]
Ha! Ha! (060601)	3.41 [1.96, 4.81]	138.56 [138.40,138.67]	3.31 [3.04,3.71]	-	-	-	-

Tableau B.4 Paramètres des courbes de tarage des stations de la cinquième catégorie. Les intervalles de confiance pour les paramètres à 95% sont obtenus par la méthode bootstrap.

Station	a_1	b_1	c_1	k	a_2	b_2	c_2
Du Chêne (046709)	9.32 [6.26, 10.15]	47.77 [47.61, 47.82]	2.51 [2.23, 3.41]	48.77 [48.32, 49.47]	21.46 [e^{-323} , 24.21]	48.29 [-253.29, 48.39]	1.11 [1.02, 133.07]
Suffolk (040409)	10.09 [9.43, 11.76]	30.13 [30.05, 30.13]	2.47 [2.40, 3.57]	30.77 [30.36, 30.77]	12.89 [9.39, 13.37]	30.53 [30.16, 30.58]	0.93 [0.8, 2.14]
Yamaska Nord (030309)	11.09 [e^{-324} , 12.12]	29.77 [10.46, 29.89]	3.56 [0.70, 261.24]	30.53 [30.04, 30.75]	11.11 [e^{-320} , 20.21]	29.90 [-77.89, 30.30]	2.05 [1.06, 158.03]

ANNEXE C RÉSULTATS DES INCERTITUDES CORRESPONDANTES AUX COURBES DE TARAGE

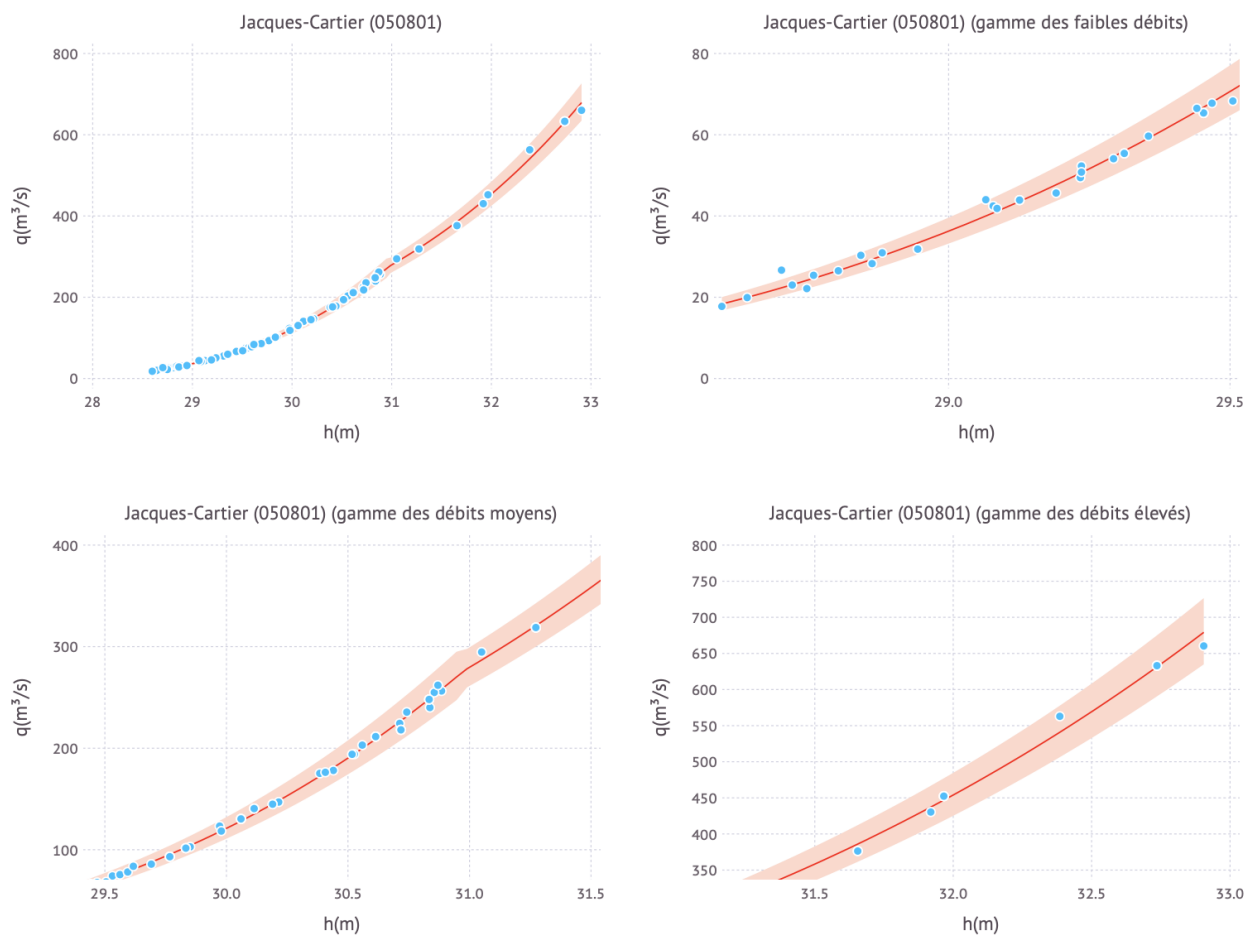


Figure C.1 Allure des incertitudes sur la courbe de tarage de la station Jacques-Cartier (050801), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.

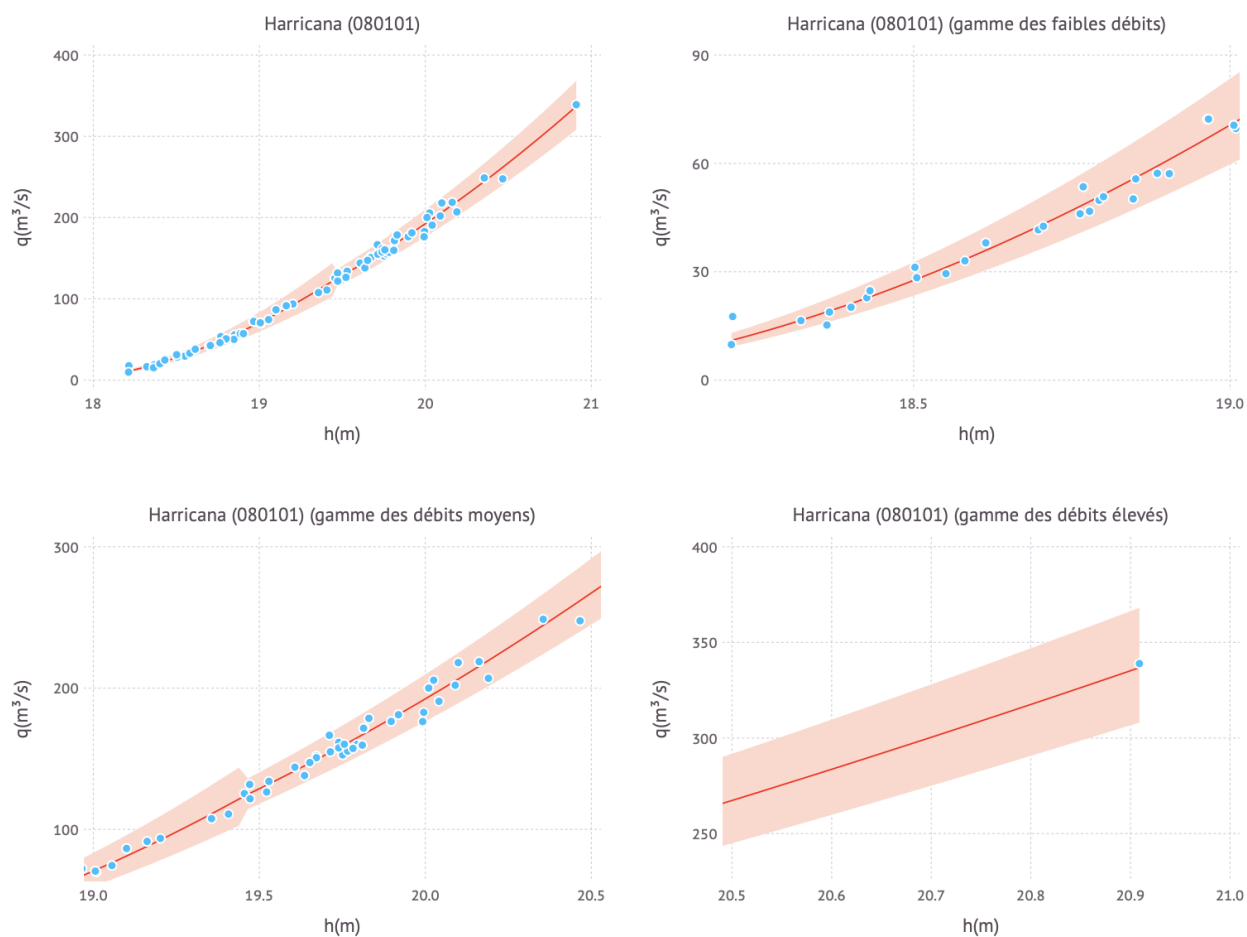


Figure C.2 Allure des incertitudes sur la courbe de tarage de la station Harricana (080101), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.

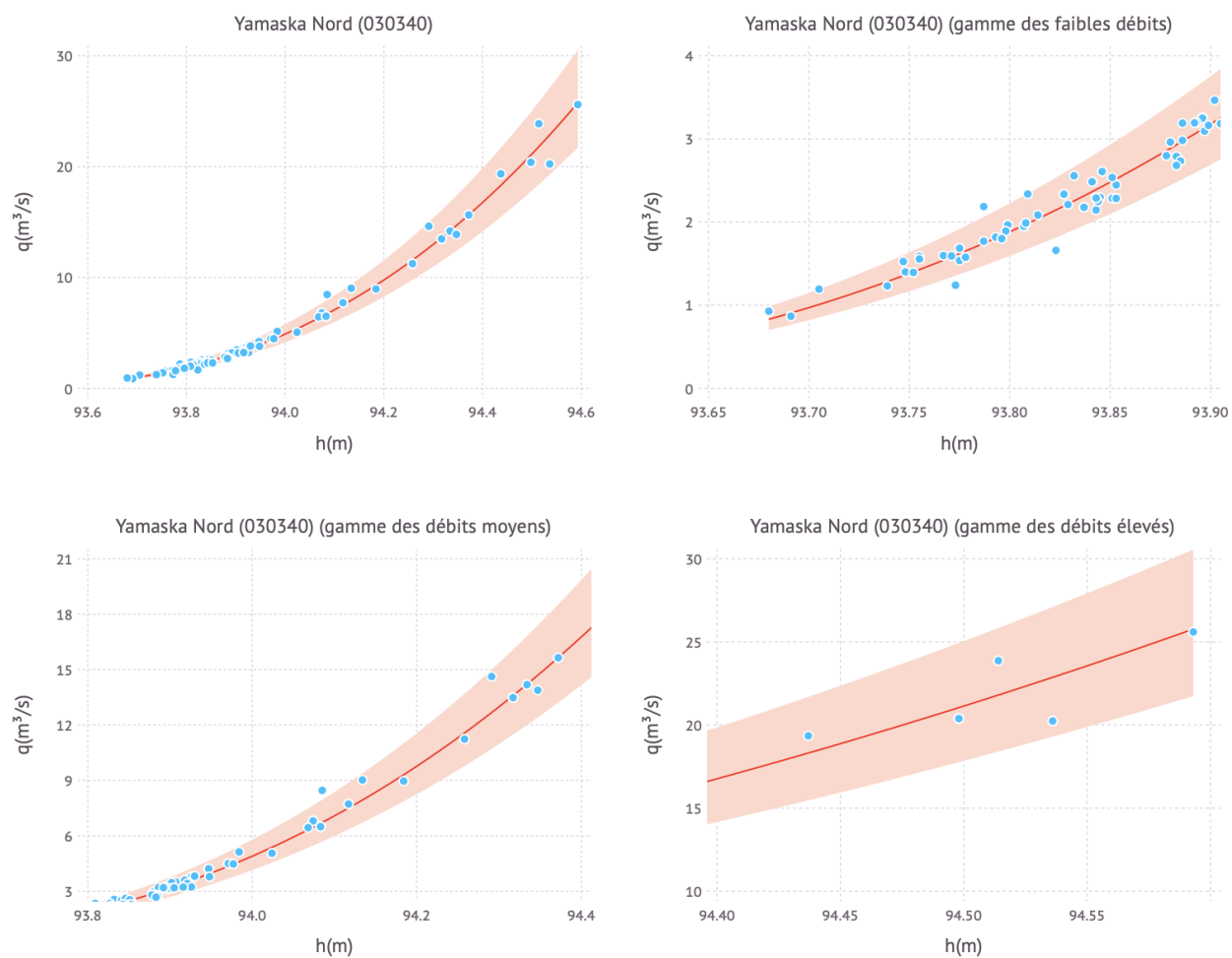


Figure C.3 Allure des incertitudes sur la courbe de tarage de la station Yamaska Nord (030340), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.

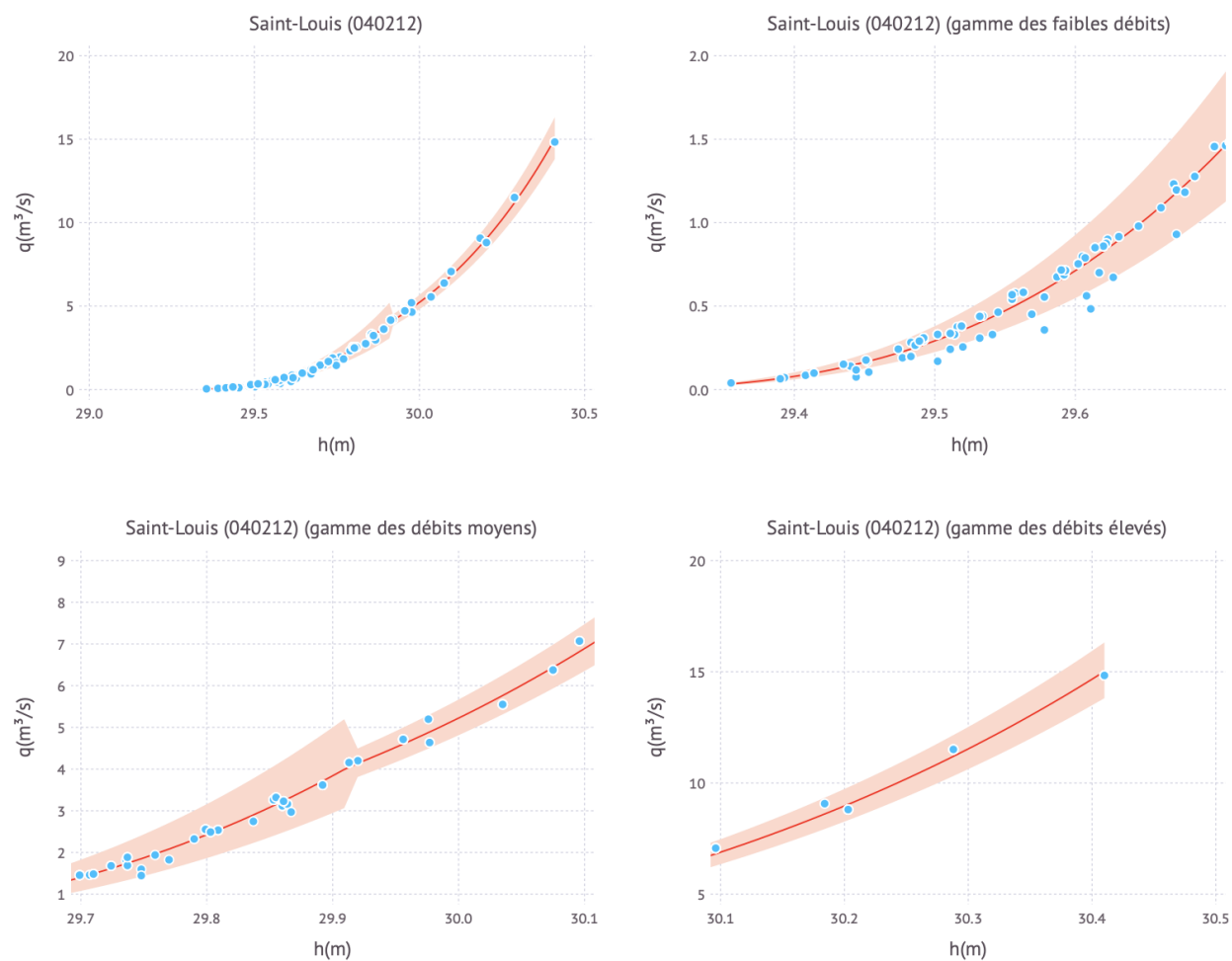


Figure C.4 Allure des incertitudes sur la courbe de tarage de la station Saint-Louis (040212), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.

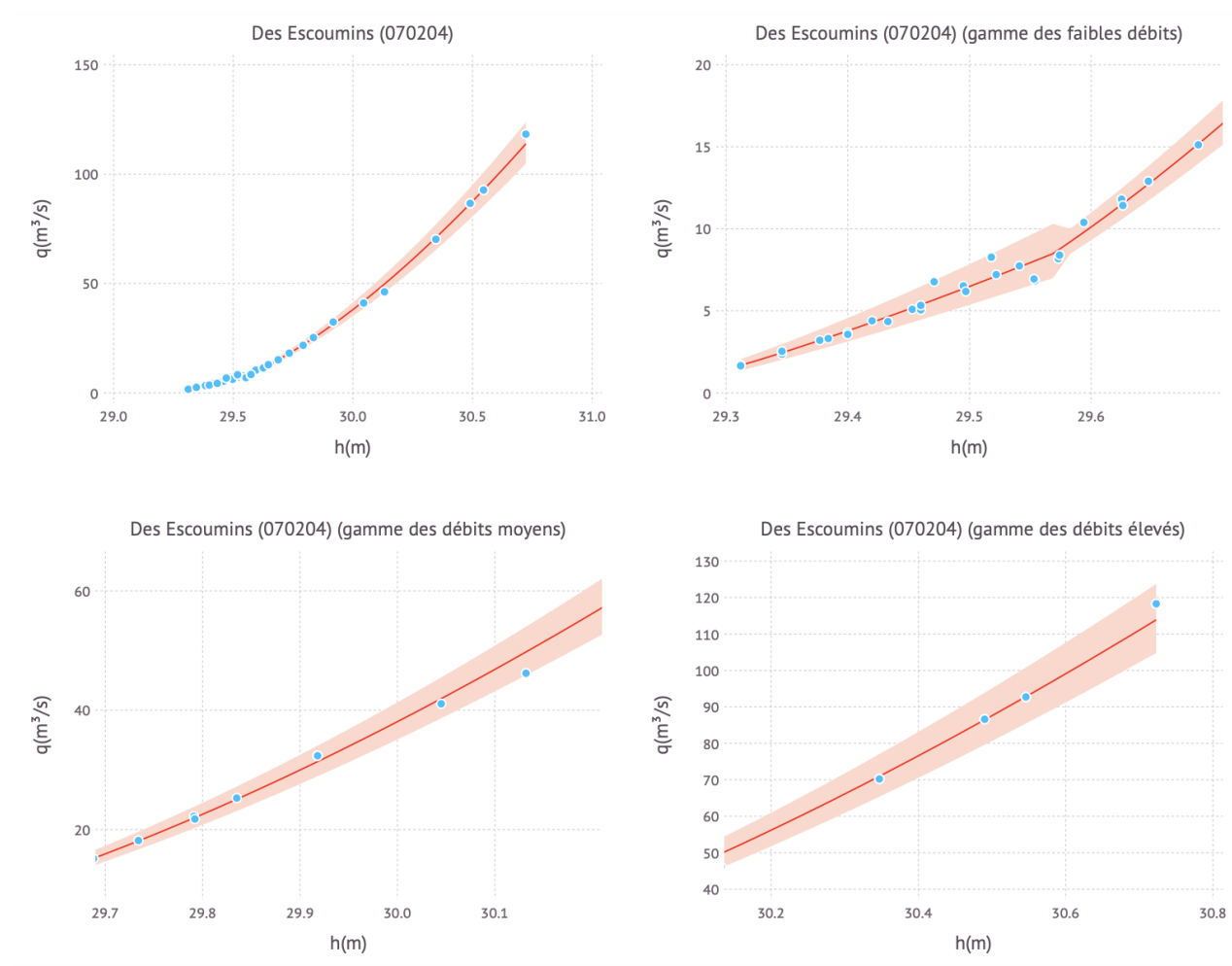


Figure C.5 Allure des incertitudes sur la courbe de tarage de la station Des Escoumins (070204), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.

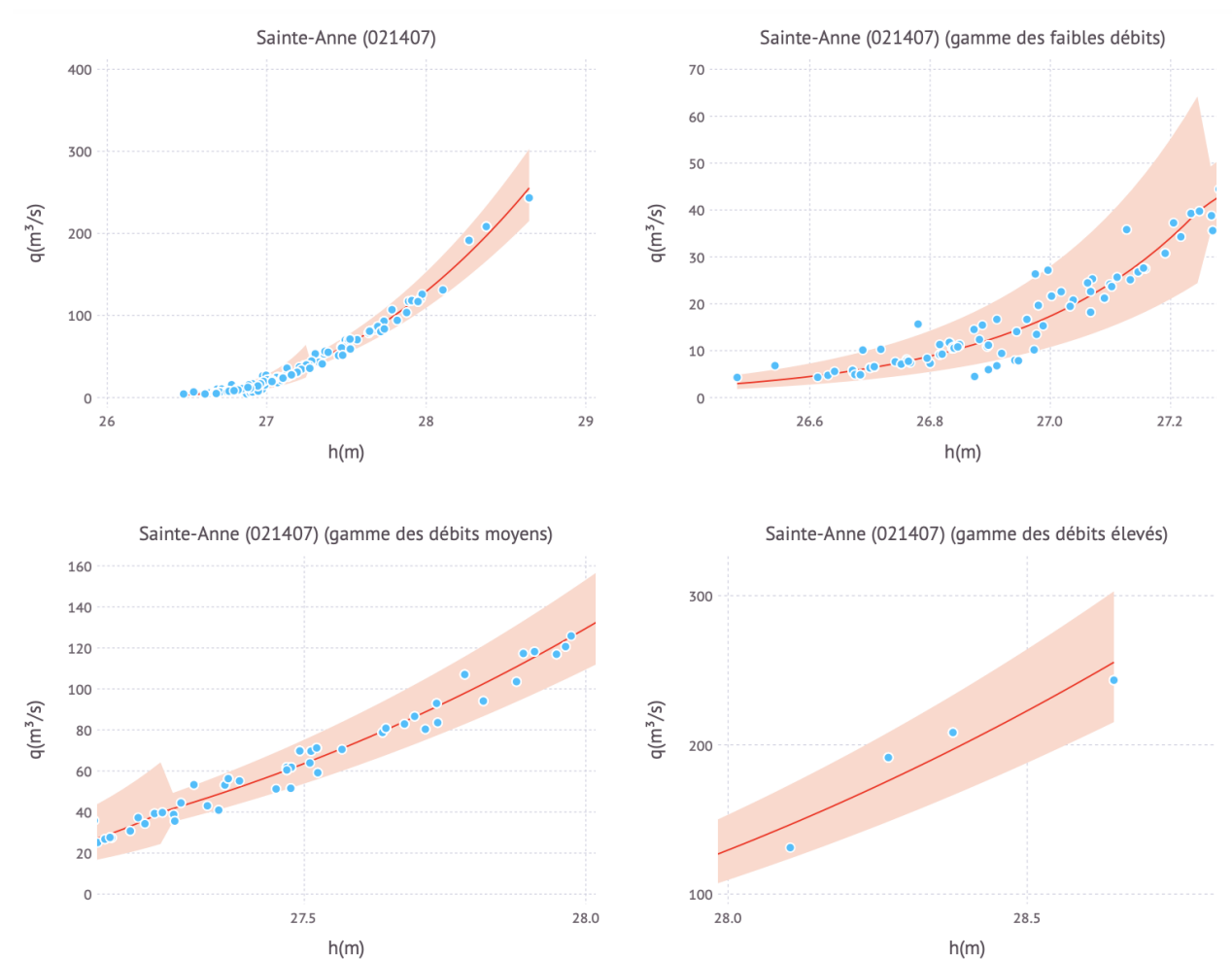


Figure C.6 Allure des incertitudes sur la courbe de tarage de la station Sainte-Anne (021407), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.

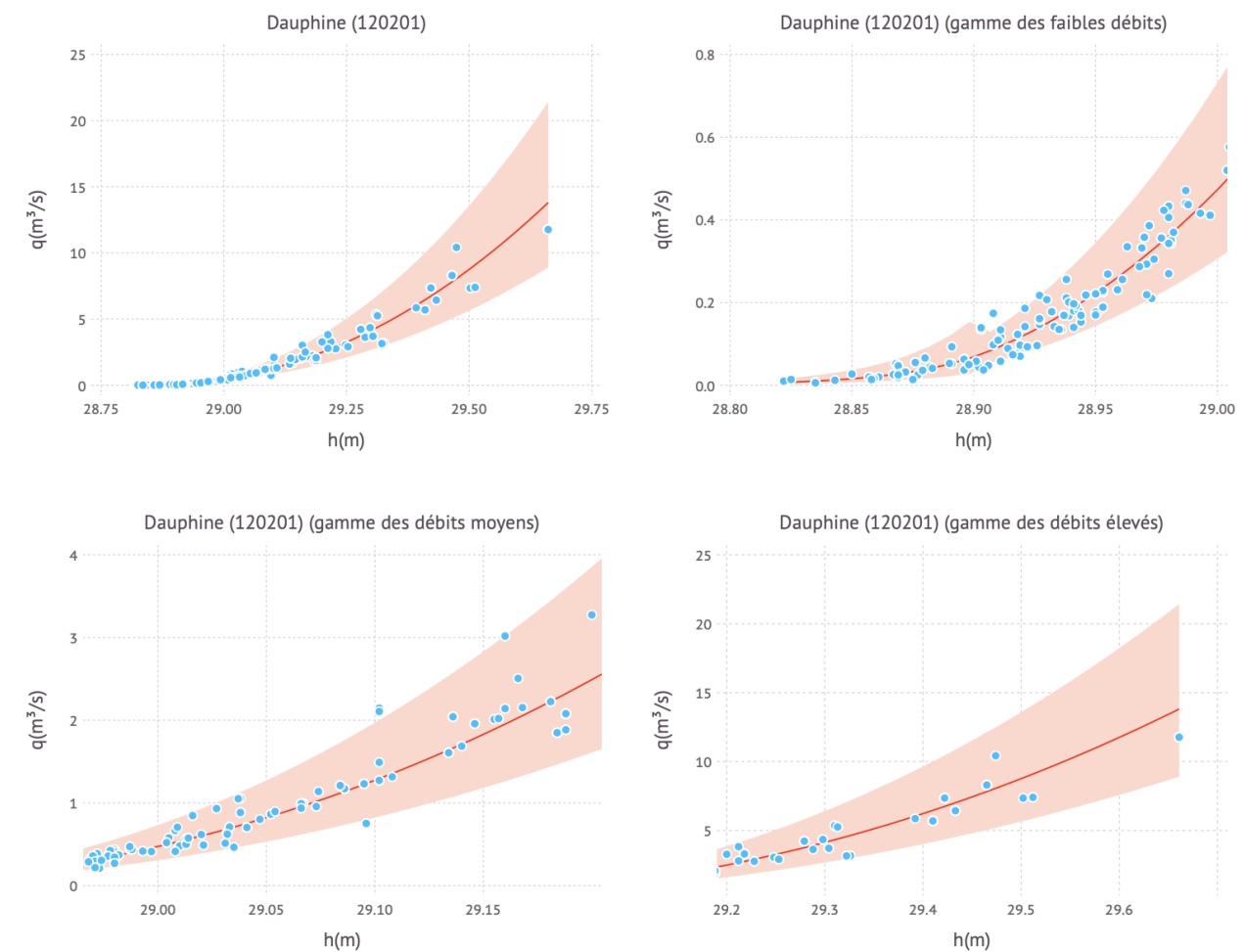


Figure C.7 Allure des incertitudes sur la courbe de tarage de la station Dauphine (120201), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.



Figure C.8 Allure des incertitudes sur la courbe de tarage de la station Boyer Sud (023002), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.

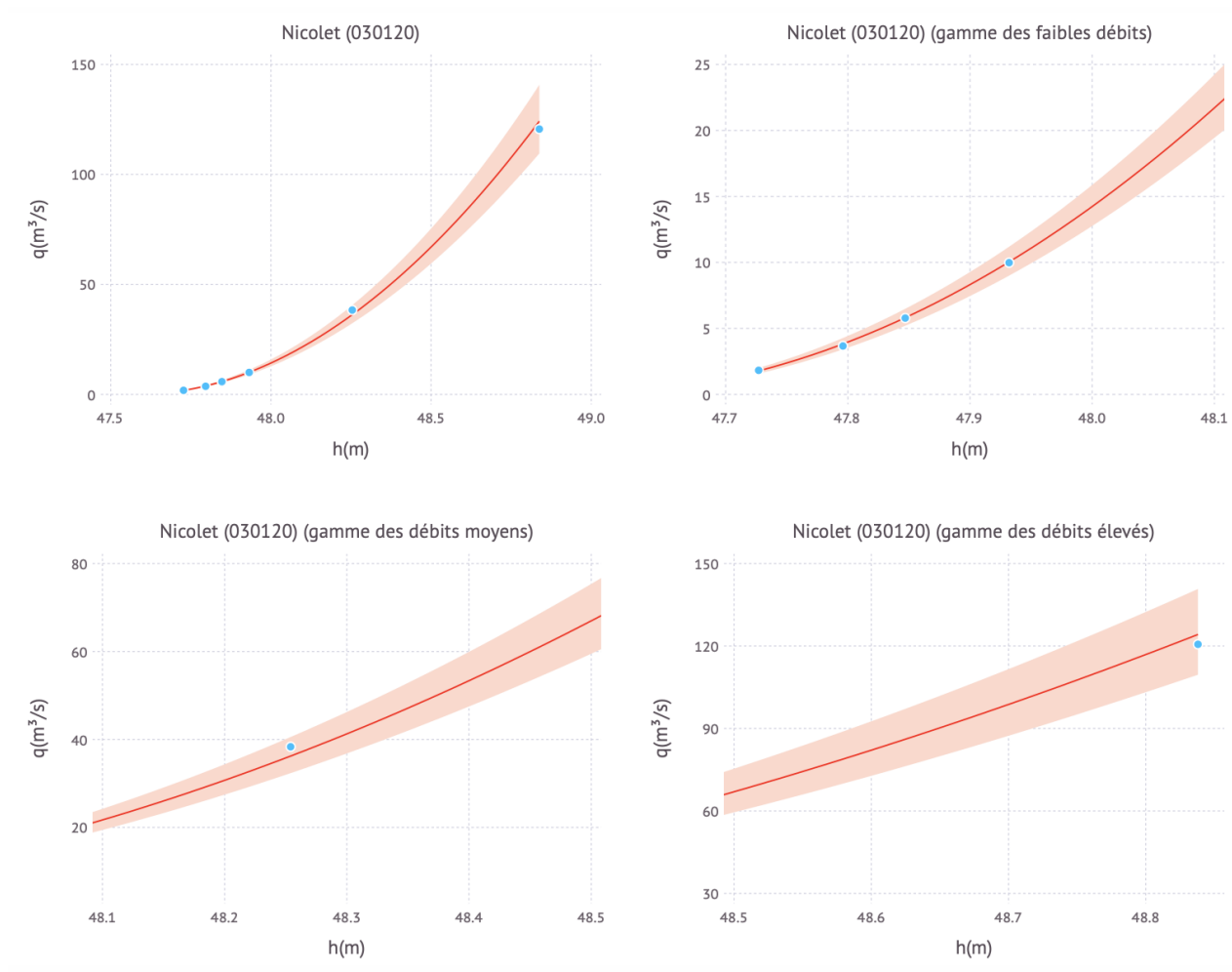


Figure C.9 Allure des incertitudes sur la courbe de tarage de la station Nicolet (030120), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.

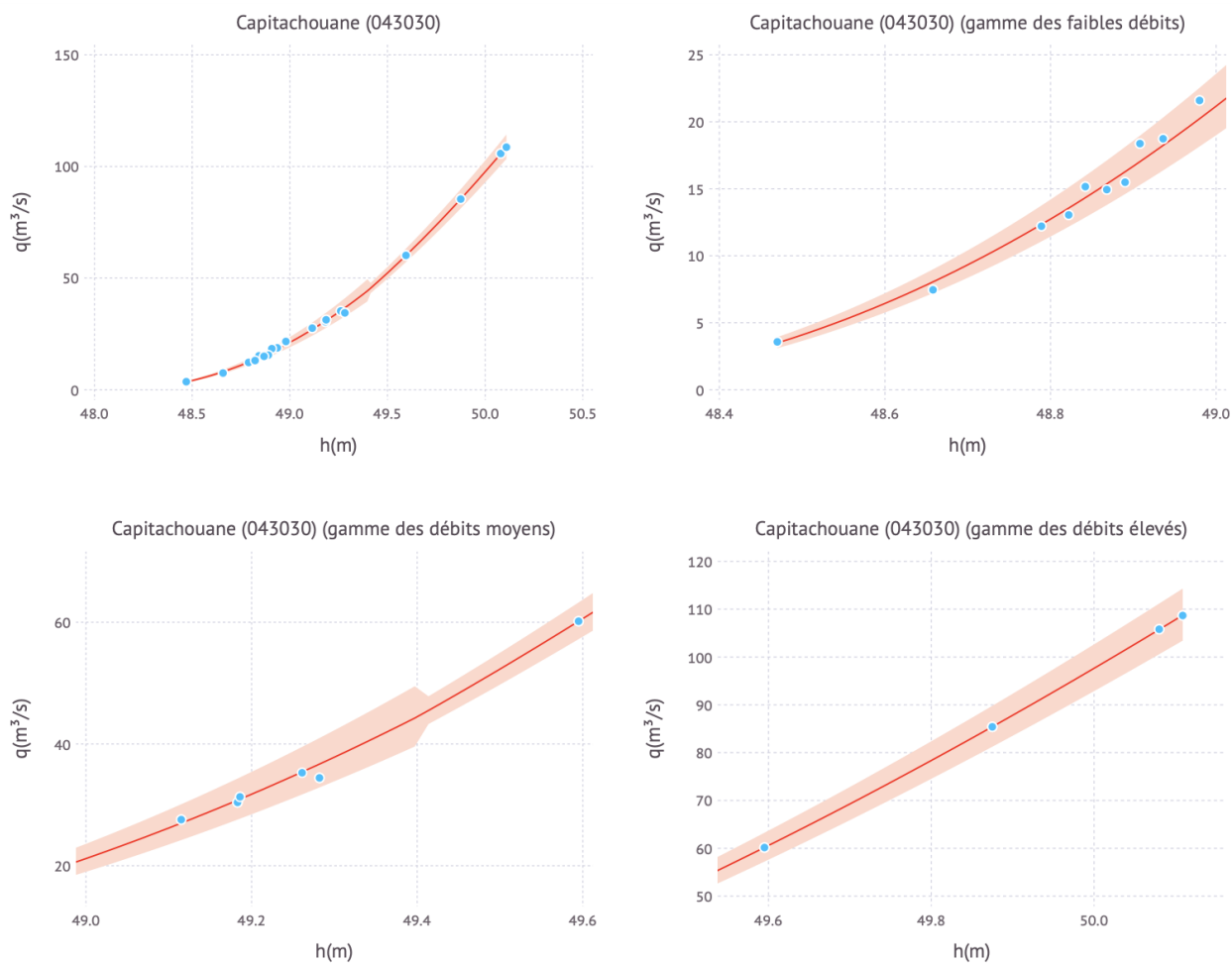


Figure C.10 Allure des incertitudes sur la courbe de tarage de la station Capitachouane (043030), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.

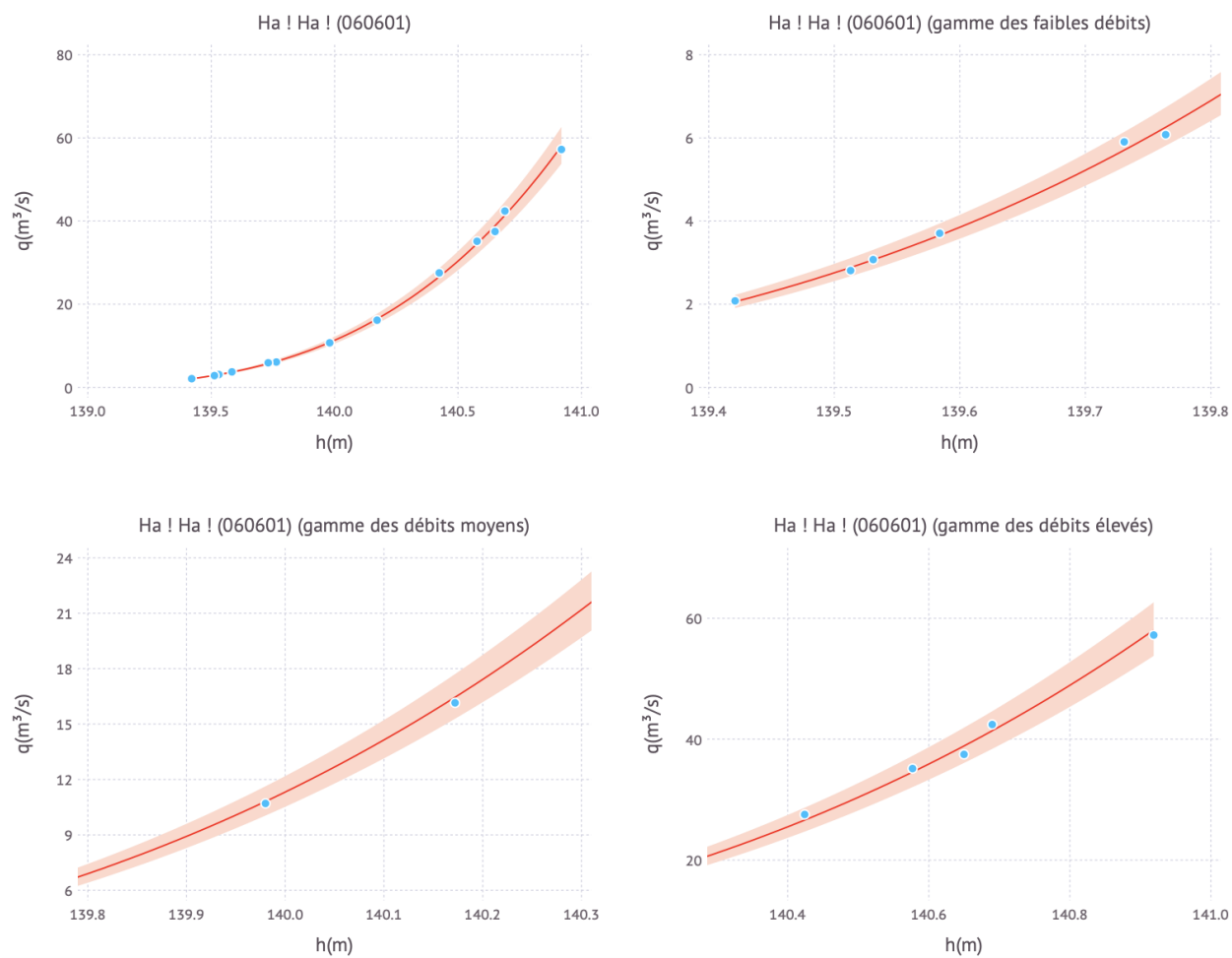


Figure C.11 Allure des incertitudes sur la courbe de tarage de la station Ha! Ha! (060601), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.

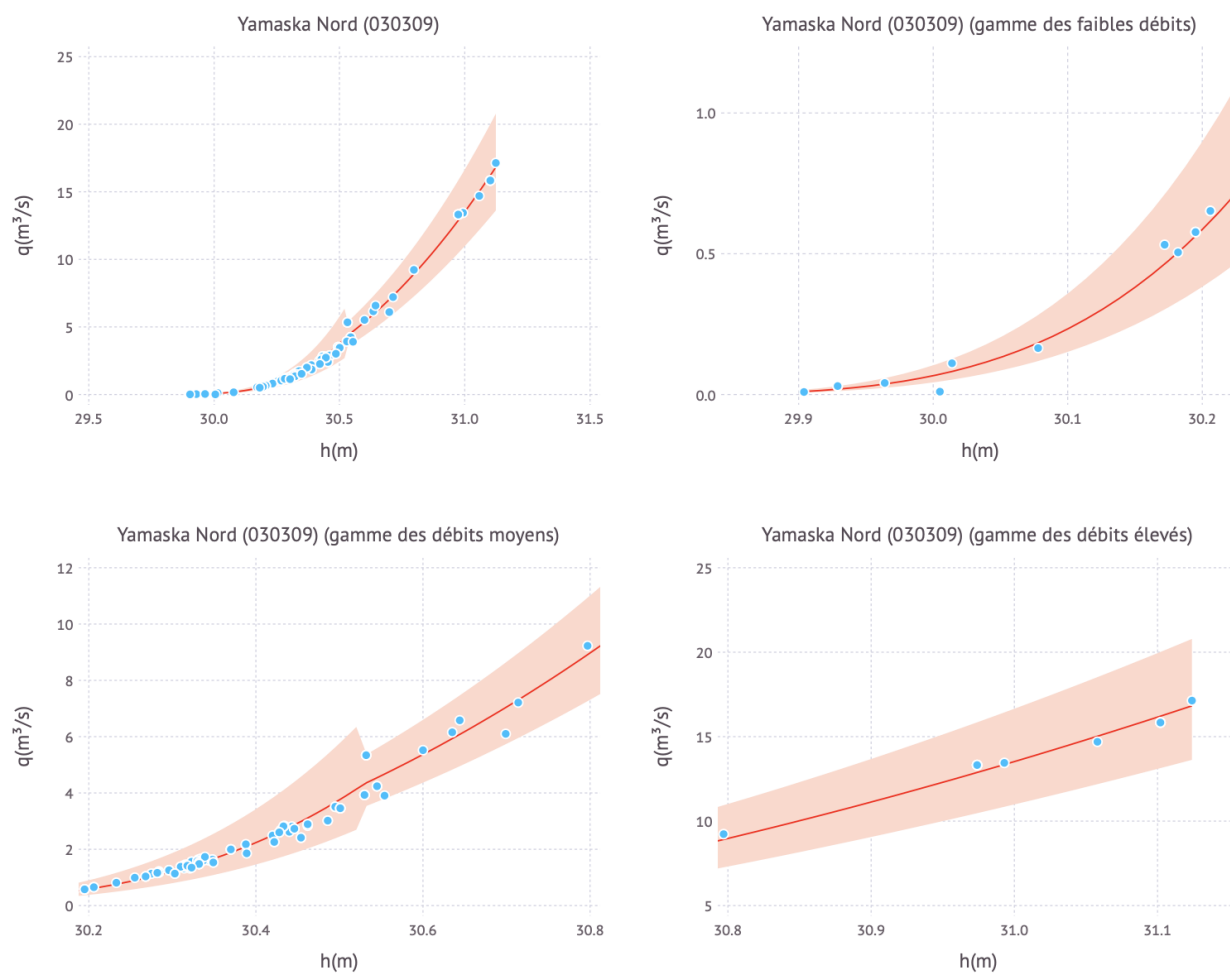


Figure C.12 Allure des incertitudes sur la courbe de tarage de la station Yamaska Nord (030309), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.

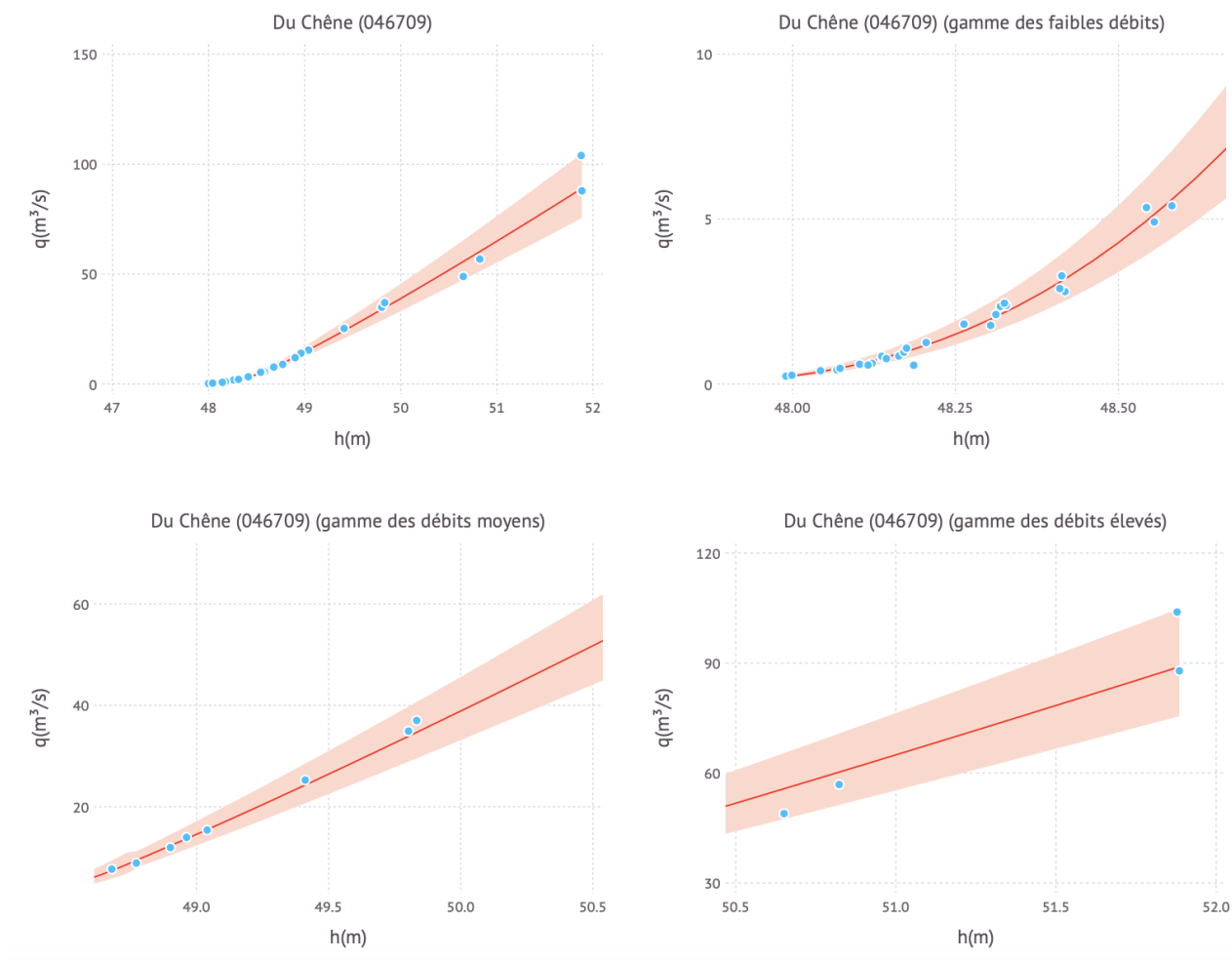


Figure C.13 Allure des incertitudes sur la courbe de tarage de la station Du Chêne (046709), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.

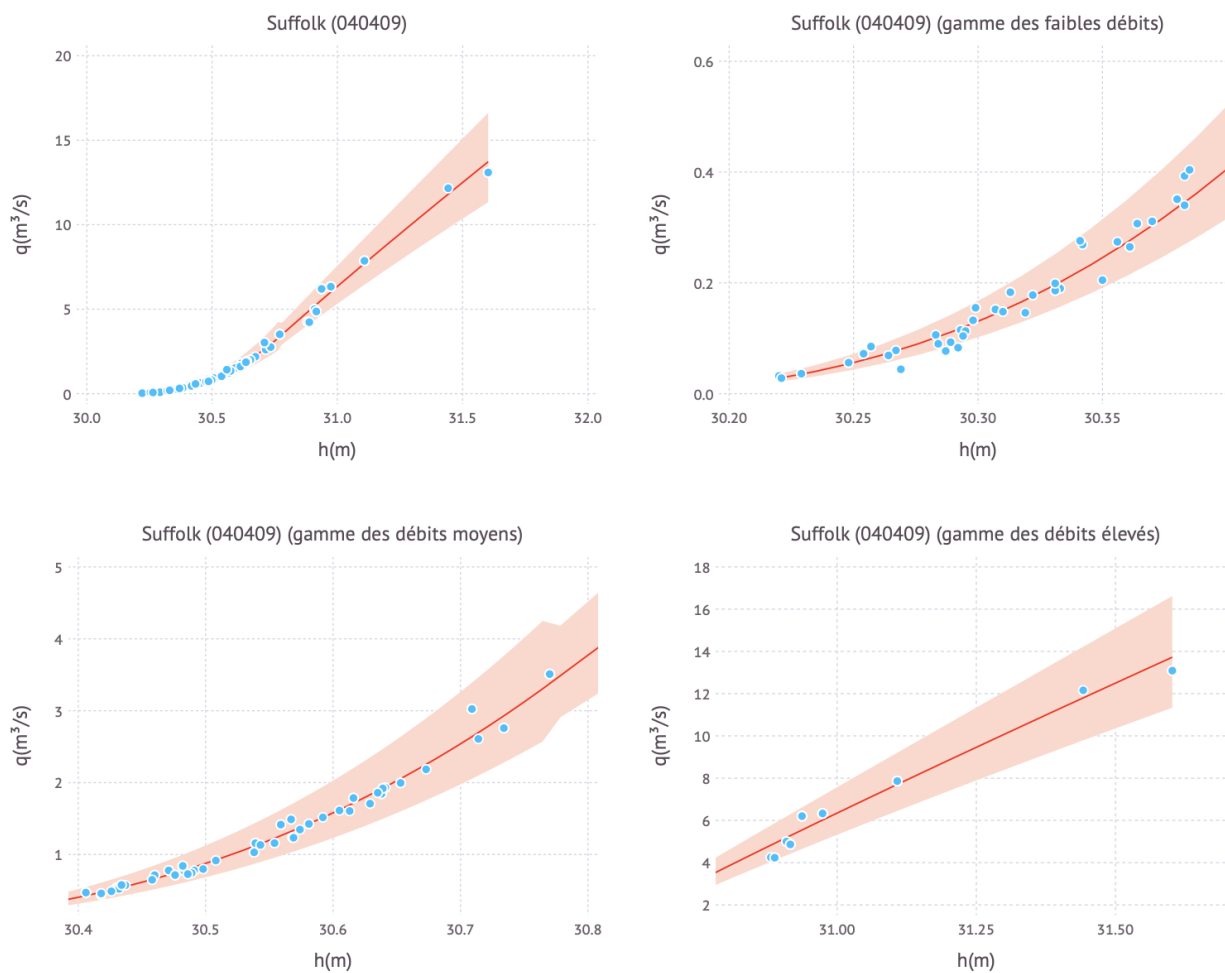


Figure C.14 Allure des incertitudes sur la courbe de tarage de la station Suffolk (040409), la courbe rouge est la courbe de tarage de l'échantillon original des données et la bande rouge présente l'incertitude du débit prédit par la courbe de tarage.