

Université du Québec
INRS – Centre Eau, Terre, Environnement

Estimation des concentrations de sédiments en suspension dans le fleuve Saint Jean
(Nouveau-Brunswick) et établissement de liens avec les données météorologiques locales

par
Hélène Higgins

Mémoire présenté pour l'obtention du grade de
Maître ès sciences (M.Sc.)
en Sciences de l'eau

Jury d'évaluation

Président du jury et examinateur interne	Normand Bergeron, centre Eau, Terre, Environnement INRS
Examineur externe	Serge Jolicoeur, Département d'histoire et de géographie, Université de Moncton
Directeur de recherche	André St-Hilaire, centre Eau, Terre, Environnement INRS
Codirecteurs de recherche	Simon C. Courtenay, Pêches et Océans Canada & Département de biologie, Université du Nouveau- Brunswick, Fredericton Katy A. Haralampides, Département de Génie Civil, Université du Nouveau-Brunswick, Fredericton

Résumé

Le Port de Saint John, au Nouveau Brunswick, est situé dans l'estuaire du fleuve Saint Jean et reçoit des apports de sédiments qui s'accumulent et entravent le passage des navires. De coûteuses opérations de dragage sont nécessaires chaque année pour maintenir les profondeurs à un niveau acceptable pour les opérations du port, mais les autorités portuaires sont incapables de prédire l'ampleur des travaux qui varie énormément d'une année à l'autre. Il est donc primordial d'étudier plus en profondeur la proportion de sédiments qui est acheminée au port de Saint John par la rivière, et de déterminer s'il est possible d'en estimer le volume à chaque année à l'aide de variables simples et facilement accessibles. Les variables étudiées comprennent un éventail de données hydrométéorologiques incluant le débit (Q), le niveau d'eau (H) et la précipitation (P). Puisque peu de données de sédiments ont été prises historiquement sur la branche principale du fleuve, une première étape de ce projet consistait à étudier une série chronologique de données journalières de concentration de sédiments en suspension (CSS) allant de 1967 à 1987 sur un tributaire important situé dans la portion en aval du fleuve, la rivière Kennebecasis. Les principales variables explicatives pour la variation des CSS au pas de temps journalier, identifiées à l'aide de corrélations, sont le débit du jour suivant, la pluie du jour précédent et la précipitation cumulée de 6 jours, bien que les coefficients de corrélation sont relativement faibles. Une analyse fréquentielle des maximums annuels de CSS et des variables hydrométéorologiques n'a pas permis d'identifier les quantiles extrêmes coïncidant qui auraient pu indiquer quelles sont les variables explicatives potentielles qui génèrent les valeurs extrêmes de CSS, qui sont typiquement une composante importante des charges annuelles. Afin de déterminer si de meilleures relations existent dans un contexte non linéaire, un réseau de neurone artificiel (RNA) a été développé permettant d'estimer les charges journalières de façon satisfaisante ($r^2 = 0,88$). Un écart absolu moyen de 16% a été ensuite obtenu entre les charges annuelles calculées et celles estimées par le modèle. Le RNA avait pour intrants la température moyenne, Q, Q(t-1), Q(t-2) et le jour julien (JJ). Toutefois, il semble que certains impacts anthropiques et le synchronisme de certains événements (ex. pluie et activités agricoles) devraient être étudiés afin d'améliorer l'estimation.

En parallèle à l'analyse de la rivière Kennebecasis, un suivi de la turbidité a été réalisé à deux sites dans un tronçon du fleuve situé en aval du barrage Mactaquac. Le suivi de la turbidité a été fait à haute fréquence comme mesure indirecte de la CSS. Le site le plus en aval a subi d'importantes pertes de données et a dû être abandonné, mais une courbe de calibration turbidité-CSS a été développée pour le site en amont, situé près de la ville de Fredericton. Une étude de corrélations a permis d'identifier H, P(-1) et la précipitation cumulée de 12 jours comme variables explicatives potentielles, mais encore une fois, les coefficients de corrélation sont relativement faibles. Deux modèles ont été développés et comparés : une régression linéaire multiple pas à pas et une interpolation par krigeage dans un espace multivarié. La seconde méthode permet de mieux estimer le log CSS que la première, mais avec une performance moyenne, comme en témoigne le pourcentage de la variance des observations expliquée par le modèle ($r^2=0,35$). Ces résultats semblent confirmer que le système fluvial étudié est complexe, présente plusieurs non-linéarités, et que d'autres variables que celles étudiées, potentiellement anthropiques, devront peut-être être prises en compte. De plus, une analyse événementielle pourrait être requise pour mieux comprendre la dynamique sédimentaire du fleuve.

L'utilisation de la turbidité en continu pour le suivi du transport de sédiments sur le fleuve Saint Jean est jugée appropriée, bien que comportant plusieurs difficultés. Des améliorations au protocole utilisé dans la présente étude soient nécessaires, incluant le suivi d'une plus longue période, en particulier durant l'hiver et le printemps.

Finalement, une analyse préliminaire où une régression multiple a été utilisée pour estimer les volumes de sédiments dragués dans le port de Saint John, à partir des variables hydrologiques locales (niveau d'eau annuel moyen et volume d'eau total annuel) suggère que cette analyse devrait être approfondie en allongeant la série de données disponibles pour les volumes de sédiments dragués.

Avant-propos

Le présent document est un mémoire par articles, ce qui signifie que les résultats des travaux sont présentés sous forme d'articles rédigés en anglais. Un résumé en français de chacun des manuscrits est présenté avant chacun d'eux, dans la deuxième partie du mémoire. Le premier article, « Suspended sediment dynamics in a tributary of the Saint John River, New Brunswick » a été accepté pour publication dans la revue « Canadian Journal of Civil Engineering ». Le second document, intitulé « Suspended sediment concentrations derived from continuous turbidity measurements in relation to hydrometeorological variables in the Saint John River, New Brunswick » a été soumis à la revue «Hydrological Processes.» Enfin, un rapport scientifique officiel de l'INRS intitulé « Linking sediments dredged from the Saint John Harbour, New Brunswick, and hydrometeorological variables », a été présenté aux autorités portuaires de Saint John.

La première partie du mémoire consiste en une synthèse, en français, mettant en perspective les résultats des trois articles mentionnés ci-haut. Elle comporte une introduction, l'énoncé des objectifs, la description du site d'étude et des méthodes, une discussion et l'intégration des résultats, une conclusion ainsi que des recommandations pour la suite des travaux.

Contribution de l'étudiant versus les autres co-auteurs

Étant première auteure de chacun des articles, mes contributions sont les suivantes :

- Participation majeure au choix des sites et à la méthode de collecte des données de turbidité.
- Élaboration du protocole d'échantillonnage pour les courbes de tarage.
- Mise en place de l'instrumentation sur le terrain et entretien des instruments durant la période de déploiement.
- Principale analyste pour l'ensemble des analyses statistiques.
- Rédaction des articles.
- Soumission des articles.

Les co-auteurs des articles ont contribué à l'élaboration du protocole de recherche, ont eu une participation mineure aux travaux sur le terrain, ont suggéré des méthodes d'analyse selon leur expertise et ont révisé les manuscrits.

Remerciements

Merci à mes directeurs André St-Hilaire, Simon Courtenay et Katy Haralampides pour leur présence et leurs commentaires pertinents. Je suis en voie d'apprendre à poser plus de questions...

À Québec :

Différentes personnes à l'INRS m'ont soit aidé avec l'obtention et l'analyse de données, avec des logiciels, sur le terrain, ou ont simplement supporté mes moments de découragement...

Merci à : Yves Tremblay, Valérie Ouellet, Dan Beveridge, Dae Il Jeong, Anik Daigle, Sandra Proulx-McInnis, Clément Clerc, Dominique Roussel, Philippe Boulanger, Charles Gignac, Jimmy Poulin, Jessy Barrette, ... Je suis certaine que j'en oublie, mais vous savez qui vous êtes.

Pour m'avoir aidée à garder le moral et la motivation nécessaires : Christiane Dupont et Thomas Hodgson.

Un merci spécial à Charles Brouard, Thibault Labarre, Julie-Anaïs Debreil et Didier.

À Fredericton :

Pour la logistique, le travail de terrain et les amitiés gagnées: le Canadian Rivers Institute, Jordan Musetta, Allan Debertin, Jason Hallett et Danika Matheron.

À Ottawa : merci à Raymond Bourdages pour m'avoir dérangée dans mon travail avec tous les documents relatifs à ma maîtrise qu'il pouvait trouver.

Merci à ceux qui m'ont donnée des idées, des commentaires, des données et ont alimenté la discussion; entre autres Ben Whalen du Kennebecasis Watershed Restoration Committee et Environnement Canada.

Table des matières

Résumé	i
Avant-propos.....	iii
Remerciements.....	v
Table des matières.....	vii
Liste des tableaux	<b.xi< b=""></b.xi<>
Liste des figures	xiii
Abréviations	xv
Partie 1- <u>SYNTHÈSE</u>	1
1.1 Introduction.....	1
1.1.1 Le transport sédimentaire en rivière.....	1
1.1.2 La turbidité comme substitut à la CSS.....	2
1.1.3 Le cas des grandes rivières	4
1.1.4 Estimer et modéliser les sédiments en suspension.....	5
1.2 Problématique et site d'étude.....	8
1.2.1 Problématique	8
1.2.2 Site d'étude	9
1.3 Objectifs et hypothèses.....	11
1.4 Méthodes	11
1.4.1 Rivière Kennebecasis	12
1.4.2 Fleuve Saint Jean	13
1.4.3 Volumes dragués dans le port.....	15
1.5 Résultats	15
1.5.1 Rivière Kennebecasis	15
1.5.1.1 Corrélations	15
1.5.1.2 Analyse fréquentielle.....	16
1.5.1.3 Réseaux de neurones artificiels et régressions multiples	16

1.5.2 Fleuve Saint Jean	17
1.5.2.1 Corrélations	18
1.5.2.2 Estimation et comparaison des charges sédimentaires	18
1.5.2.3 Régression pas à pas et interpolation par krigeage dans un espace multivarié	19
1.5.3 Volumes dragués dans le port	20
1.6 Discussion et intégration des résultats	20
1.6.1 Identification de variables explicatives potentielles	20
1.6.2 Estimation des CSS et charges de sédiments en suspension	23
1.6.2.1 Rivière Kennebecasis	23
1.6.2.2 Fleuve Saint Jean	24
1.6.3 Liens entre le transport sédimentaire dans le bassin du fleuve STJ et les volumes dragués dans son estuaire	25
1.7 Conclusions et recommandations	28
1.8 Liste des références	31
Partie 2- ARTICLES.....	35
2.1 Suspended sediment dynamics in a tributary of the Saint John River, New Brunswick.....	35
2.1.1 Résumé.....	37
2.1.2 Abstract.....	39
2.1.3 Introduction	41
2.1.3.1 The Kennebecasis River, tributary of the Saint John River.....	42
2.1.4 Study area and description of data.....	43
2.1.4.1 Study area.....	43
2.1.4.2 Data description	45
2.1.5 Methods	49
2.1.5.1 Correlation analysis	49
2.1.5.2 Frequency analysis	49
2.1.5.3 Artificial Neural Network (ANN)	50

2.1.6 Results	52
2.1.6.1 Correlation analysis	52
2.1.6.2 Frequency Analysis	55
2.1.6.3 Artificial Neural Networks	57
2.1.7 Discussion	61
2.1.7.1 ANNs	62
2.1.8 Conclusion	64
2.1.9 Acknowledgments	65
2.1.10 References	66
 2.2 Suspended sediment concentrations derived from continuous turbidity measurements in relation to hydrometeorological variables in the Saint John River, New Brunswick.....	 69
2.2.1 Résumé.....	71
2.2.2 Abstract.....	73
2.2.3 Introduction	75
2.2.4 Materials and method	76
2.2.4.1 Site Description	76
2.2.4.2 Instrumentation and data acquisition	78
2.2.4.3 Statistical analysis of SSC and potential hydrometeorological predictors	80
2.2.5 Results	82
2.2.5.1 Data Treatment and Analysis	82
2.2.5.2 Statistical analyses.....	86
2.2.6 Discussion	95
2.2.6.1 Data acquisition and treatment	95
2.2.6.2 Statistical analysis.....	97
2.2.7 Conclusions	99
2.2.8 Acknowledgements	100
2.2.9 References	101
 2.3 Linking Sediments Dredged from the Saint John Harbour, New Brunswick, and Hydrometeorological Variables	 105
2.3.1 Résumé.....	105
2.3.2 Introduction	107

2.3.3 Data & Methods	108
2.3.3.1 Estimation of sediment volumes and loads	108
2.3.3.2 Hydro-meteorological data	108
2.3.3.3 Analysis	108
2.3.4 Data	109
2.3.5 Results	110
2.3.6 Discussion	112
2.3.7 References	114
Appendice A – Lois de probabilité	115

Liste des tableaux

Tableau 1.1	Comparaison des critères de performance entre les réseaux de neurones artificiels (RNA) et les régressions multiples utilisant les mêmes intrants (T_{moy} , Q , $Q(t-1)$, $Q(t-2)$, JJ), et ayant soit les CSS soit les charges comme variable dépendante.....	17
Tableau 1.2	Résumé des modèles développés pour l'estimation de données sédimentaires pour la rivière Kennebecasis, le fleuve Saint Jean et l'estuaire où se trouve le port de Saint John.....	20
Table 2.1.1	Land-use comparison of the lower Saint John River watershed, below Mactaquac Dam, and the sub-basin upstream hydrometric station 01AP004 in the Kennebecasis River basin.....	45
Table 2.1.2	Daily statistical parameters (mean $\bar{x}$, median x_{med} , maximum x_{max} , minimum x_{min} , standard variation s_x , skewness coefficient c_{sx} and coefficient of variation c_v) of the hydrometeorological data.....	46
Table 2.1.3	Coefficients of determination (r^2) between hydro-climatic variables using daily data	53
Table 2.1.4	Coefficients of determination (r^2) between hydro-climatic variables using monthly and annually aggregated data	54
Table 2.1.5	Inputs, r^2 and RMSE for different ANN tried, using either SSC (ANN-SSC) or loads (ANN-L) as a target for the model	57
Table 2.2.1	Names, abbreviations and units for the variables used in the statistical analyses	81
Table 2.2.2	Basic statistics of the SSC dataset derived from turbidity measures.....	87
Table 2.2.3	Pearson correlation coefficients (r) among variables available at hourly intervals, with associated lag time, in hours and in parentheses, when different than 0.....	87
Table 2.2.4	Pearson correlation coefficients (r) among variables available at daily intervals, with associated lag time, in days and in parentheses, when different than 0	89
Table 2.2.5	Pearson correlation coefficients between total monthly loads and hydrometeorological variables.....	90
Table 2.2.6	Parameters of the stepwise regression and of the canonical correlation analysis (CCA) developed with log SSC as a dependant variable	91

Table 2.3.1	Sediment and hydro-meteorological data used in the analysis	110
Table 2.3.2	Coefficients of determination (r^2) among all dredging volumes and hydro-meteorological variables	111
Table 2.3.3	Coefficients and statistics of the regression	111

Liste des figures

Figure 1.1	Séries chronologiques journalières des charges estimées sur le fleuve STJ à partir des données de turbidité et sur la rivière Kennebecasis à l'aide du RNA, du 30 septembre au 26 novembre 2008	19
Figure 2.1.1	Kennebecasis River basin, with the location of the WSC hydrological station (01AP004) and weather station in Sussex. Upper inset is the whole Saint John River watershed. (Source: data available from Geobase)	44
Figure 2.1.2	Fraction of the annual load discharged during spring and fall for each year of the time series	48
Figure 2.1.3	Correlation results of SSC with flow (Q) and precipitation (P) for different lags	53
Figure 2.1.4	Return periods for maximum SSC, maximum Q, total precipitation and maximum 12-day precipitation event, for each year	56
Figure 2.1.5	Daily estimates of SSC (A) and loads (B) using the optimal ANNs plotted against observed or calculated values. Both the training and test sets are shown.	59
Figure 2.1.6	Comparison of annual suspended sediment loads obtained from Water Survey Canada and estimated from the ANN model	60
Figure 2.2.1	Saint John River watershed (Data used to create this map available from Geobase)	78
Figure 2.2.2	Flow vs water level at Site 1, 1965-2009	84
Figure 2.2.3	Time series of water levels at station 01AO002 and estimated flows.	84
Figure 2.2.4	Linear regression between suspended sediment concentration (SSC) and turbidity	86
Figure 2.2.5	SSC, water level and precipitation at daily intervals	88
Figure 2.2.6	Correlation coefficients when applying lags between SSC and water level (H), Q and precipitation (P)	90
Figure 2.2.7	Interpolated log SSC using ordinary kriging in canonical Variate (V1, V2) space.	92
Figure 2.2.8	Estimated vs observed log SSC obtained from A) stepwise regression and B) ordinary kriging in a physiographical space built using CCA on 7 variables	93
Figure 2.3.1	Comparison of dredging volumes values provided by SJPA and obtained by the regression. Error bars are $\pm$ RMSE	112

Abréviations

CRI	Canadian Rivers Institute
CSS	Concentration de sédiments en suspension
H	Niveau d'eau
ns	Non significatif
NTU	Nephelometric Turbidity Units
P	Précipitation
Q	Débit
RHC	Relevés hydrologiques du Canada
REMQ	Racine de l'écart moyen quadratique
RK	Rivière Kennebecasis
RMSE	<i>Root mean squared error</i>
RNA	Réseau de neurones artificiel
SJPA	Saint John Port Authorities - Autorités portuaires de Saint John
STJ	Saint John – Saint Jean

Partie 1- SYNTHÈSE

1.1 Introduction

1.1.1 Le transport sédimentaire en rivière

Les cours d'eau transportent des matières solides par la force du courant qui sont généralement connues sous le terme de « sédiments » et qui proviennent de l'érosion du lit, des berges et de toute la surface du bassin versant. Elles sont composées de particules inorganiques comme du sable et de l'argile, ainsi que de divers types de matières organiques (Whiles et Dodds 2002). Les sédiments ont une influence sur la chimie et la biologie des cours d'eau car ils transportent des nutriments, métaux, pathogènes et contaminants organiques (Mailhot et al. 2008; Whiles et Dodds 2002) qui ont un impact direct sur le milieu où ils sont déposés par le courant. De plus, ils jouent un rôle important au niveau de la morphologie des cours d'eau à cause des processus d'érosion et de sédimentation (Gray et Glysson 2002). Ils ont donc des impacts non négligeables sur diverses activités humaines et les infrastructures comme les barrages et les ponts, ainsi que sur les activités portuaires (Meade, Yuzyk et al. 1990; Gray et Glysson 2002; Whiles et Dodds 2002; Grauso, Fatturoso et al. 2008). Une distinction est souvent faite entre « sédiments en suspension » et « matières en suspension », où le terme sédiment réfère uniquement à la partie minérale des particules. Dans le présent mémoire, l'expression « sédiments en suspension » est utilisée de façon générale et réfère à tout type de particules, minérales ou organiques.

La quantité de sédiments en suspension en rivière est très variable, tant au niveau spatial que temporel (Horowitz et al. 2001b). Cette quantité dépend entre autres de l'hydraulicité du cours d'eau et des sources de sédiments, qui sont elles-mêmes liées à la topographie, au type de sol et au couvert végétal du bassin versant, ainsi qu'à l'utilisation du territoire (Gao 2008). Cette variabilité rend difficile l'estimation et la prédiction de la quantité de sédiments se trouvant dans le cours d'eau, d'où l'importance de choisir un endroit et une fréquence adéquats pour les mesurer (Collins et Walling 2004).

Bien que d'autres processus existent, le transport de sédiments passe principalement par la suspension dans plusieurs rivières (Knighton 1998). Au Canada par exemple, il

est estimé que le charriage ne compte que pour 5% à 20% du transport sédimentaire en rivières (Trenhaile 1998). Au États-Unis, le réseau national d'étude des grandes rivières prend uniquement en compte le transport en suspension (Horowitz et al. 2001b). Il est donc raisonnable d'utiliser une mesure de la concentration des sédiments en suspension (CSS) pour quantifier la charge sédimentaire d'un cours d'eau. De plus, cette mesure est plus facile à obtenir que le charriage de fond (bedload) (Meade et al. 1990).

Pour beaucoup de rivières, la majorité des sédiments est transportée lors de quelques événements extrêmes durant l'année (Lenzi et al. 2003; Markus et Demissie 2006; Walling et Webb 1988). Un événement extrême est défini par des conditions hydrologiques (débit) ou une concentration de sédiments beaucoup plus élevés que la moyenne, avec une fréquence d'occurrence relativement faible. Il est donc primordial de prendre des mesures sur un large éventail de débits (Simon et al. 2004) et de s'assurer de ne pas manquer ces événements majeurs, sans quoi la charge annuelle sera grandement sous-estimée (Gao 2008). Bien que les méthodes d'échantillonnage manuelles soient fiables lorsqu'effectuées selon les normes (Gray et Glysson 2002), elles requièrent trop d'investissement pour permettre des mesures journalières ou même hebdomadaires. Pour cette raison, plusieurs études sur les CSS sont basées sur des mesures hebdomadaires à mensuelles (Mailhot et al. 2008; Horowitz et al. 2001; Mano et al. 2006; Picouet et al. 2009), voire même semestrielles (Filizola et Guyot 2009) avec parfois des données un peu plus fréquentes en période de crue (Hooper et al. 2001). Il est probable qu'à cette fréquence, ces données ne peuvent pas être utilisées pour calculer des charges sur de courtes périodes, et ne permettent qu'une estimation beaucoup plus grossière de la charge annuelle. C'est pourquoi les méthodes automatisées qui permettent une mesure à haute fréquence présentent un avantage non négligeable sur les méthodes manuelles.

1.1.2 La turbidité comme substitut à la CSS

La turbidité est une propriété optique de l'eau qui mesure indirectement la quantité de matières qui la troublent. La clarté d'un liquide est effectivement diminuée lorsque la lumière est dispersée et absorbée par les particules en suspension dans l'eau plutôt que transmise en ligne droite. La turbidité en continu permet l'étude du transport sédimentaire sur de courts pas de temps ainsi que par événement, ce que les courbes de tarage classiques sédiment-débit ne permettent pas. Plusieurs études ont montré

que les mesures de turbidité peuvent être efficacement utilisées pour estimer les CSS (Pfannkuche et Schmidt 2003; Brigand et al. 2004; Dodds et Whiles 2004; Pavanelli et Bigi 2005; Gao 2008; Lopez-Tarazon et al. 2010; Oeurng et al. 2010). Parmi ceux-ci, Oeurng *et al.* (2010) ont étudié les dynamiques sédimentaires dans un bassin agricole de taille moyenne (1110 km²) et ont lié les événements de transport avec des variables hydrométéorologiques.

Des mesures de turbidité sont facilement obtenues à haute fréquence à l'aide de néphélomètres ou turbidimètres. Techniquement, il existe une différence entre les deux instruments. Un turbidimètre mesure la lumière transmise dans un échantillon d'eau alors qu'un néphélomètre mesure la lumière dispersée. Cependant, puisque l'utilisation des instruments demeure la même, les deux termes sont typiquement utilisés de façon interchangeable. Il existe plusieurs types d'instruments et chacun a des caractéristiques différentes quant à sa conception; source et puissance de la lumière incidente, sensibilité du détecteur, angle de mesure, etc. Il en résulte des mesures de turbidité spécifiques à chaque instrument sur le terrain malgré leur calibration initiale en laboratoire (Pfannkuche et Schmidt 2003).

De plus, la mesure de turbidité, généralement mesurée en NTU¹ ou en mV, doit par la suite être convertie à une valeur de solides en suspensions, en mg/L. Pour ce faire, une courbe de calibration doit être établie pour chaque instrument. Cette courbe peut être développée manuellement à l'aide d'échantillons d'eau ponctuels (subséquentement filtrés) ou en associant un système de pompe à la sonde (e.g. Eads et Lewis 2003). La courbe est spécifique à chaque site de mesure. En effet, cette méthode repose sur l'hypothèse qu'il existe une seule et unique valeur de CSS pour une valeur de turbidité donnée, ce qui n'est pas réellement le cas puisque la turbidité est très sensible à la composition et à la taille des sédiments (Gao 2008). De plus, d'autres facteurs comme la couleur de l'eau, les obstacles physiques et l'encrassement des capteurs (« biofouling ») influencent la luminosité perçue par les récepteurs et peuvent causer du bruit et/ou une dérive dans les mesures (Gao 2008; Pfannkuche et Schmidt 2003). Les effets de l'accumulation d'un film biologique (biofilm) peuvent être minimisés en nettoyant les lentilles de façon périodique et en calibrant l'appareil périodiquement à l'aide de formazine (Pfannkuche et Schmidt 2003). Il est donc primordial d'effectuer une courbe

¹ À noter que techniquement, les instruments calibrés à l'aide de formazine ont des unités FTU; mais il est typique d'utiliser plutôt la notation NTU pour plus de simplicité.

de calibration de l'instrument *in situ* pour tenir compte de ces facteurs locaux. Idéalement, il faudrait également vérifier si la courbe demeure la même dans le temps, puisque le type de sédiments transportés par le cours d'eau peut varier selon leur source, qui elle peut varier selon les saisons et les années. Par contre, établir différentes relations pour différents types d'évènement exigerait un échantillonnage plus important et n'est pas toujours réalisable. Pfannkuche et Schmidt (2003) ont aussi noté que les courbes de calibration peuvent avoir un faible coefficient de corrélation pour des environnements très turbulents.

1.1.3 Le cas des grandes rivières

Les grandes rivières ont une réponse hydrologique différente des petites, amortie par le temps de transport, l'alternance de zones d'érosion et de sédimentation et la complexité de la morphologie du chenal (Guyot et Walling 2009). De plus, l'efficacité du transport des sédiments jusqu'à l'embouchure est souvent amoindrie pour les larges bassins en terrain plat, à cause du vaste stockage alluvial le long du cours d'eau (Slattery et Phillips 2009). Plusieurs autres défis sont typiquement associés à l'étude et à la modélisation des grandes rivières : l'influence des marées sur de longues distances; la largeur et la profondeur de leur section et la possibilité de variabilité latérale de la turbidité; le phénomène d'hystérèse important; la variabilité spatiale des caractéristiques de leur bassin versant; ainsi que le manque de données (Guyot et Walling 2009). Bien que diverses études utilisent des techniques de régionalisation pour estimer les charges sédimentaires moyennes à long-terme pour de larges bassins non jaugés (Guyot et Walling 2009, Tramblay et al. 2010), il est plus difficile d'estimer la variabilité interannuelle de ces charges.

Malgré les importantes ressources humaines et matérielles que cela demande, plusieurs études utilisent encore des échantillons d'eau ponctuels en au moins un (1) point de la section (e.g. Ashmore et Day 1988; Hooper et al. 2001; Wang et al. 2011), souvent sur une base mensuelle ou semestrielle (Filizola et Guyot 2009; Horowitz et al. 2001b; Mano et al. 2006; Picouet et al. 2009). Il est dans ce cas nécessaire de prendre plusieurs mesures dans le temps sur toute la section de la rivière pour ajuster la mesure prise en un seul point si elle diffère systématiquement de la moyenne sur toute la section (Horowitz et al. 2001b).

Peu d'études ont utilisé la turbidité en continu pour estimer les charges sédimentaires sur des grandes rivières à cause des difficultés associées à l'importante variabilité spatiale des caractéristiques de bassin versant et de la logistique complexe pour l'acquisition de données. En France, par exemple, la turbidité a beaucoup été utilisée dans des bassins de moins de 1000 km² (Mano et al. 2006). Parmi les études de grandes rivières, Grenfell et Ellery (2009) ont examiné la variation intra- et inter-annuelle du transport de SS sur la rivière Mfolozi en Afrique du Sud (11 068 km²) à l'aide, entre autres, d'une série de turbidité de six (6) ans, mais prise dans le sous-bassin de tête. Ils ont conclu que l'utilisation de caractéristiques du bassin comme la superficie et la topographie pour l'estimation des charges moyennes à long-terme résulte en une surestimation des charges annuelles. Wall et al. (2008) ont utilisé un ADCP placé sur la rive pour estimer les CSS, et pris plusieurs mesures sur toute la section pour ajuster les valeurs obtenues à ce point. Ils ont ainsi obtenu une série de quatre (4) ans leur permettant d'estimer le transport sédimentaire juste en amont de l'estuaire de la rivière Hudson, à New York.

1.1.4 Estimer et modéliser les sédiments en suspension

Puisque mesurer les sédiments de façon continue demande beaucoup de ressources humaines et matérielles et de temps, il est pratique de pouvoir les estimer à l'aide de données plus facilement disponibles. Pour ce faire, des modèles sont créés qui établissent une relation entre les données de CSS et d'autres variables explicatives qui peuvent ensuite être utilisées pour estimer les CSS lorsque seules les variables indépendantes sont disponibles. Diverses méthodes existent, comme les régressions simples et multiples, les réseaux de neurones et les modèles déterministes ou empiriques (Gao 2008). Chacune de ces méthodes permet de tirer diverses conclusions sur les sédiments et leurs sources. Il est souhaitable de choisir celle qui est la plus appropriée à une situation et un objectif de gestion donné, pour s'assurer d'obtenir l'information voulue (Gao 2008). Un modèle plus complexe, qui permet de prédire la provenance des sédiments par exemple, peut être intéressant et utile dans certains cas, mais trop lourd dans d'autres cas.

La méthode la plus commune, malgré plusieurs critiques (Gao 2008; Kisi et al. 2006; Picouet et al. 2001), est l'utilisation d'une courbe de tarage sédiment-débit ou « sediment rating curve » (SRC), où les CSS sont estimées à partir d'une corrélation basée sur le débit. Cette méthode statistique simple est souvent utilisée car elle ne

requiert que peu de temps et de données, donc elle est peu onéreuse. La forme la plus fréquente de relation est $C = aQ^b$, où C est la concentration de sédiments (parfois remplacée par la charge de sédiments, Q_s), Q le débit et a et b sont les coefficients de l'équation qu'il faut estimer (Asselman 2000; Dodds et Whiles 2004; Horowitz et al. 2001b). Cette relation est rarement parfaite et montre souvent une large dispersion puisque, comme mentionné précédemment, la variabilité des sédiments ne dépend pas uniquement du débit, mais également de la disponibilité de leur source. Il est aussi possible qu'une SRC donnant un bon coefficient de détermination (r^2) soit fortuite et le fruit de coïncidences reliées à différents processus sédimentaires dans le bassin au moment de l'échantillonnage (Gao 2008). De plus, l'utilisation du logarithme des données crée un biais; plusieurs auteurs notent qu'il en résulte un lissage des variations et une augmentation de l'erreur sur l'estimation (Gao et al. 2008; Mailhot et al. 2008; Asselman 2000). De fait, les SRC tendent à sous-estimer les pics de CSS et à surestimer les faibles concentrations (Horowitz 2003; Horowitz et al. 2001b), surtout sur des périodes de temps inférieures à une saison ou un an (Horowitz 2003). Différentes corrections existent pour améliorer la relation; par exemple, Simon et al. (2004) ont utilisé deux courbes différentes pour représenter les plus faibles et les forts débits. Picouet et al. (2009) ont créé deux courbes, cette fois pour les phases montante et descendante de l'hydrogramme. D'autres corrections sont présentées par Asselman (2000). L'estimation des concentrations avec une SRC est donc de plus en plus utilisée comme premier outil d'exploration des données avant d'élaborer un modèle plus approprié (Dodds et Whiles 2004). Certains systèmes hydriques semblent bien être représentés par cette relation, alors que la corrélation n'est pas significative pour d'autres rivières. Par exemple, Dodds et Whiles (2004) ont trouvé une corrélation significative pour 489 sur 619 rivières, soit 79% d'entre elles, la plus grande possédant un r^2 de 0,91. Ils en ont conclu que cette méthode donne des résultats variables. Une revue de différentes études montre d'ailleurs que les résultats ne sont pas toujours concluants, arborant des erreurs d'estimation variant de -80% à 900% par rapport aux valeurs réelles (Gao 2008).

Une autre méthode statistique consiste à utiliser des régressions multiples, en se basant toujours sur le débit, mais en ajoutant graduellement d'autres variables explicatives représentant d'autres processus physiques affectant les CSS telle que l'utilisation du territoire, la topographie, la superficie du bassin versant, la précipitation, la température

(Gao 2008), ou tout autre facteur expliquant la production et le transport de sédiments (Mailhot et al. 2008). Les facteurs d'intérêt varient selon que l'on tente de comparer plusieurs rivières ou tronçons de rivières entre elles ou que l'on vise à modéliser le transport sur un même cours d'eau, pour différentes périodes de temps.

Lorsque la relation entre les CSS et les variables explicatives est non-linéaire, on peut se tourner vers d'autres types de modèles comme les réseaux de neurones artificiels (RNA). Ceux-ci opèrent un peu comme les neurones biologiques, c'est-à-dire que les différents intrants (variables explicatives) sont transformés par différents opérateurs mathématiques et fonctions de transfert appelés neurones artificiels. Ces neurones peuvent nourrir une autre couche de neurones qui vont continuer à transformer les intrants, de telle sorte qu'à la sortie du réseau, une estimation de la variable dépendante est obtenue. L'ensemble de ces transformations sont ajustées durant l'entraînement du réseau. Cette technique permet souvent plus de flexibilité dans la modélisation que les régressions linéaires car les fonctions de transfert peuvent prendre différentes formes non-linéaires (Partal et Cigizoglu 2008). Les RNA sont de plus en plus utilisés pour différentes applications hydrologiques telle la prédiction de charges sédimentaires (Sarkar et al. 2008).

Plusieurs autres méthodes existent pour estimer l'ordre de grandeur de la charge sédimentaire en suspension ou le potentiel d'érosion, comme l'équation universelle de perte des sols révisée (Revised Universal Soil Loss Equation, ou RUSLE), mais puisque ces modèles sont souvent empiriques, ils demandent souvent une importante quantité de données spatiales et temporelles sur la rivière et son bassin. Ces données ne sont pas toujours facilement disponibles.

Il est important de noter que peu importe la technique choisie pour estimer les CSS, le calcul des charges repose toujours sur une mesure fiable du débit (Gao 2008).

Une autre difficulté associée à la modélisation des CSS provient du concept d'hystérèse associé aux deux variables utilisées, soit les CSS et le débit (Picouet et al. 2001). En effet, les pics de CSS et de débit ne coïncident pas nécessairement. On dit qu'il y a hystérèse positive (clockwise) lorsque le pic de sédiments survient avant que le débit atteigne son niveau maximum. Cela signifie également que pour un même débit, la concentration de sédiments sera différente pendant une crue ou pendant une baisse du niveau d'eau (Collins et Walling 2004). Ce phénomène apparaît parce que la quantité de

matière en suspension ne dépend pas uniquement de la capacité de transport du cours d'eau, mais également des sources de sédiments disponibles (Lenzi et al. 2003). En conséquence, la concentration maximale théorique de CSS dans un cours d'eau n'est typiquement jamais atteinte et les concentrations ne répondent souvent pas instantanément à un changement de débit. Aussi, les plus fortes concentrations de sédiments observées au début d'un hydrogramme sont généralement davantage reliées à la plus grande disponibilité des sédiments suite à l'action érosive des événements de précipitations qu'à l'augmentation de débit (Trenhaile 1998).

1.2 Problématique et site d'étude

1.2.1 Problématique

Les relevés hydrologiques du Canada ont, au fil du temps, fait le suivi de seulement 10 stations hydrométriques incluant des mesures de sédiments en suspension dans le bassin versant du fleuve Saint Jean, dont aucune n'est encore active aujourd'hui. De plus, une seule se trouve sur la branche principale du cours d'eau, en aval du barrage Mactaquac, et ces mesures ont été prises uniquement à l'époque de la construction du barrage, en 1966 et 1967.

L'existence d'un port majeur à Saint John et la nécessité de maintenir une profondeur d'eau minimum pour les navires de gros tonnage à fort tirant d'eau a entraîné diverses études sur le transport de sédiments et leur modélisation, ainsi que sur la sédimentation dans le port (Leys 2007; Neu 1960). Leys (2007) a développé un modèle hydrodynamique 3D incluant le transport sédimentaire dans une section du fleuve en amont du port, au niveau de Reversing Falls. Il en ressort que plus de 50% de la sédimentation dans le port serait dû à un courant de fond provenant de la Baie de Fundy; mais l'étude est basée sur une seule journée d'échantillonnage de concentration de sédiments en suspension (CSS) et sur les données de Neu (1960). Selon ces auteurs, une campagne d'échantillonnage prolongée serait nécessaire pour véritablement tirer des conclusions fiables et afin de tenter de capturer la variabilité spatio-temporelle des CSS. De plus, l'erreur substantielle associée à l'estimation des volumes de sédiments dragués ajoute de l'incertitude au modèle. Enfin, le modèle développé ne permet pas d'estimer la variation interannuelle de la sédimentation; seulement une moyenne à long terme.

En revanche, les autorités portuaires de Saint John (SJPA) sont intéressées à estimer cette variation de la sédimentation dans l'estuaire, puisque les coûteuses opérations de dragage nécessaires à chaque année y sont intimement liées. Plusieurs difficultés dans l'étude des sédiments et de l'hydrologie en général du fleuve Saint Jean sont associées aux très fortes marées dans la Baie de Fundy et à la morphologie compliquée du cours d'eau dans sa partie inférieure. Il est jusqu'à ce jour incertain si ce sont les sédiments provenant de la Baie ou du fleuve qui ont le plus d'impacts sur la sédimentation dans le port. SJPA a donc mandaté l'Institut Canadien des Rivières (CRI) pour étudier le transport de sédiments dans le fleuve Saint Jean ainsi que dans le port.

Le CRI a pour mandat d'évaluer si les sédiments s'accumulant dans le port proviennent directement du fleuve ou de la Baie de Fundy, et d'en évaluer la variabilité. Le présent mémoire s'attarde à la partie de cette problématique concernant le fleuve. Il s'agit, à long terme, de déterminer s'il est possible d'estimer le transport de sédiments dans le cours d'eau à l'aide de variables explicatives facilement disponibles (par exemple, les débits et les variables météorologiques) et de vérifier si les quantités mesurées peuvent être liées aux opérations de dragage dans le port de Saint John.

1.2.2 Site d'étude

Le fleuve Saint Jean (STJ) coule sur environ 673 km entre le Maine (É-U) et la Baie de Fundy au Nouveau-Brunswick (Canada), drainant un territoire de 55 500 km². Environ la moitié du bassin versant (28 860 km²) se trouve au Nouveau Brunswick, le reste passant par le Québec et le Maine. À Saint John, qui se trouve à l'embouchure, les débits minimum de la saison estivale et maximum du printemps sont estimés à 500 m³/s et 3 400 m³/s, respectivement (Leys 2007) et le débit moyen à 991 m³/s (Neu 1960). Le fleuve possède de nombreux tributaires, une série d'îles et îlots, des variations importantes dans la largeur de sa section ainsi que plusieurs lacs sur son bassin versant, en faisant un système hydrique complexe. La rivière Kennebecasis est le tributaire possédant la plus longue période de mesures de sédiments par le service des Relevés hydrologiques du Canada (RHC). La rivière est localisée dans la portion inférieure du bassin, formant la Baie Kennebecasis avant de rejoindre le fleuve STJ et de se déverser dans la Baie de Fundy (Figure 2.1.1).

Dans la portion néo-brunswickoise, le régime hydrique du fleuve Saint Jean est influencé par plusieurs barrages, le plus en aval étant Mactaquac. Bien que ce barrage,

situé environ à 15 km en amont de Fredericton, soit au fil de l'eau, il crée un réservoir (~ 87 km²) où les sédiments peuvent s'accumuler. Cet état de fait est observable sur la série de mesures de sédiments prises pendant la construction du barrage. Les CSS en aval de la structure passent d'une moyenne de 31 mg/l avant la construction à 7 mg/l une fois le barrage complété en 1967. En conséquence, la présente étude considère que la majorité des sédiments transportés vers l'estuaire proviennent de la section du bassin versant en aval du barrage Mactaquac, ce qui représente environ 15 000 km². De plus, bien que ce type de barrage ait généralement peu d'influence sur les niveaux d'eau, Mactaquac est géré conformément aux demandes énergétiques des municipalités locales. En ce sens, il arrive, surtout en période d'étiage, que l'écoulement soit influencé par les opérations du barrage et ne soit ainsi pas lié à d'autres facteurs contrôlant normalement le débit comme la précipitation. L'effet des marées est perceptible jusqu'aux environs de Fredericton (Trites 1960). Dans cette partie du bassin, qui inclut la rivière Kennebecasis, la température de l'air moyenne est de 16°C à 18°C l'été et de -6°C à -9°C l'hiver. Les précipitations annuelles sont de l'ordre de 1200 mm, dont 960 mm sous forme liquide. La géologie est composée principalement de roches sédimentaires et ignées.

Selon Bray et Xie (1993), la portion du bassin en aval du barrage est relativement homogène en termes de climat, de géologie et de conditions hydro-climatiques. L'utilisation du territoire est par contre légèrement différente; l'agriculture, en particulier, est plus concentrée dans le sous-bassin de la rivière Kennebecasis en amont de la station hydrométrique 01AP004 (21,8%, tableau 2.1.1) que dans la portion inférieure du bassin du fleuve Saint Jean (6%). Il est à noter que l'agriculture couvre environ 16% du bassin complet de la rivière Kennebecasis comparativement à 21,8% dans la partie en amont de la station hydrométrique.

Bien que l'agriculture risque de jouer un plus grand rôle dans les processus de transport sédimentaire du tributaire que dans la branche principale du fleuve, la similitude entre les autres caractéristiques, principalement la portion forestière, permet de former l'hypothèse, pour ce projet, que les variables explicatives sous-jacentes au transport sédimentaire décelées pour la rivière Kennebecasis sont de bons indicateurs des variables explicatives dans le bassin inférieur du fleuve Saint Jean. Il ne s'agit pas ici d'appliquer directement au fleuve les résultats obtenus sur son tributaire; par contre, l'utilisation de la longue série de données disponible sur ce tributaire peut permettre une

analyse préliminaire et donner une direction générale à l'étude de la branche principale du fleuve.

1.3 Objectifs et hypothèses

L'objectif principal de ce projet est l'étude de la variation du transport de sédiments dans le fleuve Saint Jean, au Nouveau-Brunswick, et l'établissement de liens avec des données hydrométéorologiques locales ainsi qu'avec le volume de sédiments dragués périodiquement dans l'estuaire où se trouve le port. Plus spécifiquement, les buts de l'étude sont :

- 1) établir des liens entre les sédiments dans le fleuve, son tributaire principal dans la zone en aval et différentes variables hydrométéorologiques ;
- 2) vérifier si les concentrations ou les charges de sédiments peuvent être estimées en utilisant ces mêmes variables hydrométéorologiques;
- 3) établir si les charges sédimentaires annuelles estimées correspondent, ou peuvent être reliées, aux volumes dragués dans l'estuaire.

Puisque les mesures de sédiments en suspension prises sur une courte période, soit la durée normale d'un mémoire, ne permettent pas d'étudier la variabilité interannuelle (qui requiert de longues séries de données), l'étude d'un tributaire important du fleuve, la rivière Kennebecasis, a également été conduite afin d'identifier des variables s'appliquant potentiellement à tout le bassin. Les données de dragage dans le port des dernières années ont également été étudiées. Il est primordial de noter qu'ultimement, l'objectif consiste à obtenir une estimation approximative de l'impact du transport de sédiments dans le fleuve sur les volumes dragués dans le port.

1.4 Méthodes

Afin d'atteindre les objectifs décrits ci-haut, trois séries d'analyses ont été effectuées : 1) une analyse statistique portant sur la rivière Kennebecasis (le tributaire du fleuve), 2) une analyse des liens entre la turbidité et les variables hydrométéorologiques sur le fleuve Saint Jean et 3) une première estimation des liens entre les volumes de sédiments dragués dans le port et les variables hydrométéorologiques.

1.4.1 Rivière Kennebecasis

D'abord, puisque peu de recherches hydrologiques et sédimentaires ont été effectuées dans le bassin du fleuve Saint Jean, une première étape exploratoire consistait à identifier des variables hydrométéorologiques simples et disponibles pouvant expliquer la variabilité du transport de sédiments. Puisqu'aucune longue série chronologique de mesures de sédiments n'existe pour la branche principale du fleuve, la rivière Kennebecasis, qui possède une telle série entre 1967 et 1988 provenant des Relevés hydrologiques du Canada (RHC), a été utilisée à cette fin. Cette série journalière de CSS a servi à : 1) détecter des corrélations (Pearson) avec diverses variables hydrométéorologiques (débit, température, précipitation, précipitation cumulée sur 3, 6 et 12 jours); 2) effectuer une analyse fréquentielle sur les données annuelles extrêmes; et 3) développer un réseau de neurones artificiels (RNA) pour l'estimation de concentrations et de charges sédimentaires journalières.

L'analyse fréquentielle, décrite en détails dans diverses références (e.g. Rao et Hamed 2001), est une méthode permettant de relier l'ampleur d'un événement donné à sa fréquence d'apparition à l'aide de lois de probabilité ajustées à la série de données. Le logiciel HYFRAN (http://www.ete.inrs.ca/activites/groupe/chaire_hydrol/hyfran.html) a été utilisé à cette fin, en suivant la méthode décrite par Tremblay et al. (2008). Dans la présente étude, il s'agissait de vérifier si les maximums annuels de SSC, Q et de précipitation cumulée sur 3, 6 ou 12 jours pouvaient expliquer les années de forte charge sédimentaire totale. La précipitation annuelle totale a également été étudiée.

Un réseau de neurones artificiels (RNA) est un modèle mathématique simulant l'arrangement des réseaux de neurones biologiques, c'est-à-dire une série de neurones hautement interconnectés et arrangés par couches, chacune étant liée à la suivante. Cet arrangement permet la création de puissants modèles non linéaires, qui sont d'abord entraînés à l'aide d'exemples de données afin d'en identifier les patrons, pour être ensuite en mesure de les reconnaître dans de nouvelles séries de données (Jain 2001). La technique des RNA permet une meilleure modélisation de l'hystérèse que les régressions multiples (Alp et Cigizoglu 2007, Jain et Chalisgonkar 2000) et permet l'intégration, dans un modèle non-linéaire, de variables explicatives additionnelles au seul débit. Dans la présente étude, des RNA sont utilisés afin de vérifier si les variables prises en compte sont adéquates et suffisantes pour modéliser les concentrations et les charges de sédiments en suspension dans la rivière Kennebecasis. L'utilisation d'un pas

de temps journalier dans le modèle permet de cibler les moments où le modèle est faux, et donc de détecter de potentielles variables manquantes. Dans un premier temps, un RNA de type rétro-propagation a été construit pour modéliser les CSS, à partir desquelles les charges journalières ont été calculées. Différents essais ont été conduits pour trouver le modèle optimal, c'est-à-dire celui expliquant le plus de variance et étant le moins biaisé, tout en utilisant le moins d'intrants possible. Un autre RNA a ensuite été développé, utilisant directement les charges journalières comme variable dépendante plutôt que les CSS, afin de permettre la comparaison des résultats avec la littérature, où cette pratique est commune (Alp et Cigizoglu 2007; Cigizoglu 2004; Cigizoglu et Kisi 2006; Sarkar et al. 2008). Afin de vérifier si l'utilisation d'un modèle non-linéaire est effectivement digne d'intérêt, des régressions multiples ont été construites en utilisant les mêmes intrants que dans les RNA. Les critères de performance entre les différents modèles sont la racine de l'écart quadratique moyen ($REQM^2$) et le coefficient de corrélation entre les valeurs prédites et observées.

Les détails de ces méthodes, incluant une introduction plus complète à l'analyse fréquentielle et aux RNA, sont disponibles dans la section 2.1.

1.4.2 Fleuve Saint Jean

Afin de pallier au manque de données récentes de sédiments sur le fleuve, un suivi de la turbidité, éventuellement convertie en concentration de sédiments en suspension (CSS) a été effectué. Pour ce faire, des néphélomètres ont été installés, à partir des berges, en deux points du fleuve afin d'estimer les CSS de façon indirecte et en continu. Des données ont été obtenues à toutes les 15 minutes. La présente recherche est la première à utiliser cette méthode pour obtenir des données de CSS sur le bassin versant étudié. La section de la rivière où les instruments ont été installés étant relativement large (> 200 m), des mesures de turbidité ou de CSS ont été prises en différents points de la section, et ce en trois occasions différentes, afin de vérifier la représentativité du point de mesure, tel que suggéré par Wagner et al. (2006). L'utilisation de la turbidité nécessite également le développement d'une courbe de calibration turbidité-concentration dont les détails sont donnés dans la section 2.2. En somme, il s'agit de prendre un échantillon d'eau concomitant avec une mesure de turbidité et ce, pour toute la gamme de turbidité observée au site de mesure.

² La racine de l'écart quadratique moyen est le « root-mean-squared-error » (RMSE) en anglais

L'échantillon d'eau est ensuite filtré en laboratoire pour en déterminer la CSS. Malheureusement, une suite de difficultés techniques ont conduit à la perte de données et à l'impossibilité de développer une courbe de calibration fiable au deuxième site, forçant l'abandon de ces données dans la présente analyse. Aucun problème n'a été décelé au site 1, permettant une transformation satisfaisante des valeurs de turbidité en CSS. Seules les mesures de CSS du site 1 ont donc été étudiées, celles-ci s'étalant sur deux périodes : du 30 septembre au 27 novembre 2008 et du 22 mai au 25 novembre 2009. D'abord, elles ont été agrégées au pas de temps horaire et journalier, puis analysées conjointement avec des données hydrométéorologiques locales (tableau 2.2.1) à l'aide d'une analyse des corrélations de Pearson. À noter que seul le niveau d'eau est mesuré par les relevés hydrologiques du Canada (RHC) près du site (station 01AO002) et que le débit a dû être estimé à l'aide du rapport des superficies drainées. Ensuite, les charges sédimentaires ont été calculées pour les données disponibles à l'aide du débit estimé au site de mesure. Finalement, deux modèles utilisant les CSS comme variable dépendante ont été comparés : une régression pas à pas vers l'avant et une méthode combinée d'analyse multivariée et de géostatistique développée initialement par Chokmani et Ouarda (2004) pour estimer les quantiles de crues dans un cadre régional. Guillemette et al. (2009), qui ont appliqué cette méthode géostatistique pour l'estimation de températures maximum mensuelles dans deux rivières du Québec, en présentent une description plus détaillée. Une introduction est également présentée à la section 2.2.4.3. Brièvement, dans cette approche, l'analyse des corrélations canoniques (ACC) est appliquée sur un ensemble de variables explicatives. De ce fait, toutes les variables explicatives potentielles peuvent être incluses initialement; l'ACC construit des combinaisons linéaires des variables initiales, appelées variables canoniques, en maximisant la corrélation des variables canoniques dépendantes (e.g. les CSS ou leur logarithme) et explicatives. Dans le cas présent, deux groupes de variables sont utilisés; l'un relié aux sédiments et l'autre comprenant diverses variables hydrométéorologiques. Les variables canoniques (*variates*) sont orthogonales et permettent de créer un plan cartésien dans l'espace des variables explicatives. Par la suite, les variables dépendantes sont projetées dans cet espace créé et une interpolation par krigeage ordinaire est effectuée entre elles. Cette étape permet ainsi d'estimer des valeurs de CSS inconnues, connaissant les variables explicatives initialement utilisées dans l'ACC.

1.4.3 Volumes dragués dans le port

Les volumes de sédiments dragués annuellement dans le port entre 2003 et 2009 ont été analysés conjointement avec diverses variables hydrométéorologiques locales ainsi qu'avec les charges sédimentaires annuelles provenant de la rivière Kennebecasis. Ces dernières ont été estimées en appliquant le modèle RNA développé précédemment. Les variables hydrométéorologiques retenues (tableau 2.3.1) sont le niveau d'eau maximum et moyen, le débit estimé maximum et moyen, le volume d'eau total annuel, la pluie, la neige et la précipitation totales annuelles ainsi que la période de retour du volume d'eau total obtenue par analyse fréquentielle. Les données ayant trait aux volumes dragués proviennent de la pesée des résidus de dragage lorsqu'ils sont transportés par barge pour être entreposés à l'extérieur du port. Une régression multiple incluant les meilleures variables explicatives a été développée et l'erreur du modèle a été évaluée à l'aide d'une technique de ré-échantillonnage jackknife pour pallier en partie au petit nombre de valeurs de l'échantillon. Les détails de cette analyse sont disponibles à la section 2.3.

1.5 Résultats

La présente section rapporte les principaux résultats de chaque partie du projet, afin de faciliter la lecture et la compréhension de la discussion qui suit. Les résultats détaillés des différentes analyses se trouvent dans leur article respectif, auxquels une référence est faite au besoin.

1.5.1 Rivière Kennebecasis

1.5.1.1 Corrélations

Au pas de temps journalier, les principales corrélations significatives décelées (tableau 2.1.3) sont celles entre les CSS et le débit du jour suivant ($r^2 = 0,13$, figure 2.1.3) ainsi que la précipitation du jour précédent ($r^2 = 0,09$, figure 2.1.3). La relation entre les diverses variables s'améliore lorsque les données sont regroupées au pas de temps mensuel et annuel (tableau 2.1.4). Parmi les variables sédimentaires, les charges annuelles et le débit moyen (maximum) annuel sont le plus fortement corrélés avec un r^2 de 0,68 (0,69), ce qui était à prévoir puisque les charges sont calculées à partir des débits. Un r^2 de 0,37 (0,39) existe entre les CSS moyens annuels et le débit moyen (maximum) annuel, et un r^2 de 0,48 est présent entre les CSS moyens annuels et l'événement maximum annuel de 12 jours de pluie. De façon générale, les liens sont

plus forts au pas de temps annuel, sauf pour quelques extrêmes comme le débit maximum qui est plus fortement corrélé aux données hydrométéorologiques mensuelles.

1.5.1.2 Analyse fréquentielle

Les distributions sélectionnées (i.e. meilleur ajustement sur les extrêmes) pour les maximums annuels de CSS, débit (Q), précipitation et précipitation cumulée de 12 jours sont les lois lognormale, exponentielle, gamma inverse et gamma, respectivement (Appendice A). Les périodes de retour calculées à partir de ces distributions ne coïncident pas systématiquement pour chacune des variables étudiées (figure 2.1.4). De plus, en comparant les périodes de retour aux charges annuelles représentées par un « + » sur la figure 2.1.4, il n'existe aucun patron systématique reliant les extrêmes hydrométéorologiques aux plus grandes charges annuelles. Les extrêmes de CSS et de débit ne sont pas non plus systématiquement synchrones. Même lorsqu'ils se produisent pendant la même saison, ils concordent à quelques jours près dans seulement quatre cas; le reste du temps, ils ne se produisent même pas durant le même mois.

1.5.1.3 Réseaux de neurones artificiels et régressions multiples

Les deux réseaux de neurones artificiels (RNA) créés à l'aide des séries de données du sous-bassin de la rivière Kennebecasis utilisent les mêmes intrants journaliers, soit la température moyenne (T_{moy}), différents débits Q, Q(t-1), Q(t-2) et le jour julien (JJ). Des essais ont été effectués en incluant d'autres intrants (tableau 2.1.5) mais n'ont pas produit de meilleurs résultats. Le premier RNA tente de modéliser directement les CSS, mais les valeurs estimées et réelles sont faiblement corrélées ($r^2 = 0,23$). De plus les valeurs de CSS élevées sont fortement sous-estimées par le RNA (figure 2.1.5 A).

Le second RNA, qui modélise les charges journalières, permet d'obtenir de meilleurs résultats ($r^2 = 0,88$, figure 2.1.5 B). Les charges annuelles estimées à partir de ce RNA ont un écart absolu moyen de 14% avec les valeurs calculées par les relevés hydrologiques du Canada (RHC, figure 2.1.6).

Les régressions multiples bâties à partir des intrants utilisés dans les RNA produisent des résultats systématiquement moins bons que les RNA, i.e. de plus faibles corrélations et des REQM plus élevées (tableau 1.1). La régression ayant les CSS comme variable dépendante ($r^2 = 0,16$) a des valeurs de REMQ de 44,8 mg/L

(entraînement) et 23,4 mg/L (test) alors que la régression ayant les charges comme variable dépendante ($r^2 = 0,59$) a des valeurs de REMQ de 255 T (entraînement) et 155 T (test). Les charges annuelles calculées à partir des régressions multiples ne sont pas présentées car leur écart absolu moyen avec les valeurs mesurées excédait systématiquement 100%.

Tableau 1.1 Comparaison des critères de performance entre les réseaux de neurones artificiels (RNA) et les régressions multiples utilisant les mêmes intrants (T_{moy} , Q , $Q(t-1)$, $Q(t-2)$, JJ), et ayant soit les CSS soit les charges comme variable dépendante.

Variable dépendante :	CSS			Charges		
	r^2	RMSE		r^2	RMSE	
		Entraînement	Test		Entraînement	Test
RNA	0,23	43,2	20,9	0,88	136	147
Régression multiple	0,16	44,8	23,4	0,59	255	155

Note: r^2 représente la corrélation entre les valeurs estimées par le modèle et les valeurs réelles

1.5.2 Fleuve Saint Jean

La courbe de calibration turbidité-CSS utilisée pour transformer les données au site 1 ($r^2 = 0,92$) est présentée à la figure 2.2.4. Les séries chronologiques journalières de CSS, précipitation et niveau d'eau utilisées dans la présente analyse sur un point du fleuve Saint Jean sont présentées à la figure 2.2.5. Les CSS ont varié entre 0,06 mg/l et 501,72 mg/l, tandis que les niveaux ont varié entre 1,06 m et 3,72 m. Un total de 313,3 mm de précipitations a été reçu durant la période d'échantillonnage en 2008 et 663,9 mm de pluie sont tombés durant la période de suivi en 2009.

Les mesures prises sur la section du fleuve pour valider la représentativité du point de mesure de CSS étaient uniformes lors de deux des trois campagnes d'échantillonnage. Lors de la troisième, un des points de mesure situé près de la berge et du fond de l'eau avait une concentration supérieure (15 mg/l) au reste de la section (moyenne = 1,92 mg/l).

Les débits estimés à l'aide du rapport des surfaces drainées ont d'abord été comparés aux niveaux d'eau mesurés à la station des relevés hydriques du Canada (RHC) en

appliquant avec succès une relation de puissance entre les deux séries ($r^2 = 0,88$, figure 2.2.2). Les pics entre les deux séries semblent généralement concomitants (figure 2.2.3), bien que de meilleurs résultats seraient obtenus en prenant en compte le temps de déplacement de l'eau entre la station en amont et celle au site de mesure. Ainsi, les débits estimés sont jugés satisfaisants pour l'estimation des charges tant que l'incertitude est adéquatement prise en compte, mais les niveaux (effectivement mesurés et non estimés) ont été préférés pour les analyses de corrélations.

1.5.2.1 Corrélations

Les principales variables explicatives décelées à l'aide de corrélations de Pearson pour les CSS du fleuve STJ au pas de temps journalier sont généralement les mêmes que sur son tributaire (rivière Kennebecasis) : le niveau d'eau ($r = 0,43$), la précipitation du jour précédent ($r = 0,27$) et la précipitation cumulée de 12 jours ($r = 0,27$, tableau 2.2.4). Par contre, aucun délai entre les CSS et H n'a été détecté au pas de temps journalier (i.e., la corrélation maximale a été trouvée pour des valeurs journalières concomitantes); seule l'analyse horaire a identifié un délai de +9 heures pour une corrélation maximale ($r = 0,28$, tableau 2.2.3) entre ces deux variables. Il est à noter que des corrélations sur les rangs (de type Spearman) ont également été calculées, sans amélioration importante des coefficients de corrélation. Une analyse annuelle des CSS n'a pas été possible dans le fleuve STJ puisque moins d'un an de données ont été récoltées, mais celles-ci ont tout de même été regroupées par mois. Les principales variables explicatives pour les charges totales mensuelles sont (tableau 2.2.5) : Q(H) maximum avec $r = 0,85(0,79)$, Q(H) moyen avec $r = 0,82(0,73)$ et le maximum de pluie cumulative 12 jours ($r = 0,64$, $p < 0,1$). Les CSS moyennes mensuelles ne sont significativement corrélées ($p < 0,05$) qu'avec le niveau d'eau maximum annuel ($r = 0,74$).

1.5.2.2 Estimation et comparaison des charges sédimentaires

À titre d'information supplémentaire à l'article de la section 2.2, les charges journalières estimées sur le fleuve à l'aide de la turbidité et sur son tributaire à l'aide du RNA développé précédemment ont été comparées pour les mois d'octobre et de novembre 2008, seules données concomitantes disponibles sur les deux cours d'eau. Aucune corrélation significative n'existe entre les deux séries de données ($r^2 = 0,05$, $p = 0,08$). Par contre, quelques événements des deux séries sont assez synchrones, comme le montre la figure 1.1. Les quatre (4) pics de charge présents sur la rivière Kennebecasis

coïncident avec une hausse importante de la charge sur la branche principale du fleuve, quoique leur amplitude ne corresponde pas (i.e. le plus gros pic sur la rivière ne survient pas en même temps que celui sur le fleuve). Il y a aussi trois (3) pics observés sur le fleuve qui se sont produits alors qu'aucune hausse de CSS importante a été enregistrée sur la rivière Kennebecasis. Il est à noter que l'évènement sédimentaire majeur du 23 au 27 novembre 2008 sur le fleuve STJ, bien que montré dans les figures, n'a pas été inclus dans les analyses statistiques. Comme discuté dans la section 2.2, sa nature extrême (période de récurrence 9 ans pour cette période de l'année) biaisait les corrélations et masquait les dynamiques présentes lors d'autres évènements au cours de la période de suivi.

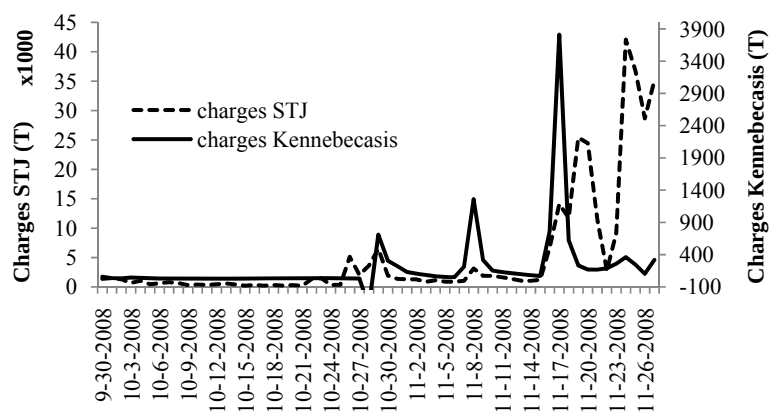


Figure 1.1 Séries chronologiques journalières des charges estimées sur le fleuve STJ à partir des données de turbidité et sur la rivière Kennebecasis à l'aide du RNA, du 30 septembre au 26 novembre 2008

1.5.2.3 Régression pas à pas et interpolation par krigeage dans un espace multivarié

Une régression pas à pas vers l'avant a été construite avec le log CSS journalier comme variable dépendante. Les variables explicatives identifiées sont H, °T et P(-1). Le modèle n'explique que 29% de la variance du log CSS, avec une REMQ de 0,51 (tableau 2.2.6 et figure 2.2.8 A) et un biais³ de 0,05. La méthode de Chokmani et

³ moyenne des écarts entre les valeurs observées et estimées

Ouarda (2004) alliant analyse de corrélation canonique (ACC) et krigeage ordinaire explique légèrement plus de variance ($r^2=0,35$), avec une REMQ de 0,48 (tableau 2.2.6 et figure 2.2.8 B) et un biais de 0,01. Un résumé des modèles est présenté au tableau 1.2.

1.5.3 Volumes dragués dans le port

Malgré la faible taille de l'échantillon disponible, une première régression multiple a été développée en utilisant les meilleures variables explicatives pour la variation des volumes dragués dans l'estuaire, soit le volume d'eau total annuel ($r = 0,85$) et le niveau d'eau moyen ($r = 0,79$). Les autres résultats de corrélations sont présentés dans la section 2.3 (tableau 2.2.3). Le modèle reproduit relativement bien les six (6) données annuelles disponibles, avec un r^2 de 0,87 et un écart absolu moyen de 15,4%.

Tableau 1.2 Résumé des modèles développés pour l'estimation de données sédimentaires pour la rivière Kennebecasis, le fleuve Saint Jean et l'estuaire où se trouve le port de Saint John

Milieu visé	R. Kennebecasis		Fleuve STJ		Estuaire
Type de modèle	RNA	RNA	Régression pas à pas	ACC/krigeage simple	Régression multiple
Variable de sortie	CSS	Charges sédimentaires	Log CSS	Log CSS	Volumes dragués (T)
Intrants	$T_{\text{moy}}, Q, Q(t-1), Q(t-2), JJ$	$T_{\text{moy}}, Q, Q(t-1), Q(t-2), JJ$	$H, T^{\circ}, P(-1)$	$H, Q, T^{\circ}, P(-1), 3d, 6d, 12d$	Volume d'eau total, niveau d'eau moyen annuel
r^2	0,23	0,88	0,29	0,35	0,83
REMQ	-	-	0.51	0.48	77309T
écart absolu moyen.	19%	14%	48%	44%	15%
Détails :	Section 2.1		Section 2.2		Section 2.3

Note : le r^2 représente la corrélation entre les valeurs mesurées et estimées par chaque modèle

1.6 Discussion et intégration des résultats

1.6.1 Identification de variables explicatives potentielles

Les meilleures variables explicatives de la concentration de sédiments en suspension (CSS) identifiées pour le fleuve STJ (deux saisons) et pour son tributaire, la rivière

Kennebecasis (> 20 ans), sont généralement les mêmes au pas de temps journalier. D'abord, il n'est pas surprenant que le débit (ou le niveau d'eau) soit la principale variable explicative puisqu'il est directement lié à la capacité de transport du cours d'eau. D'ailleurs, dans beaucoup d'études, le débit est utilisé comme unique variable pour créer une courbe débit-sédiments (« sediment rating curve », SRC) permettant d'estimer ces derniers malgré plusieurs critiques sur la méthode (Gao 2008; Kisi et al. 2006; Picouet et al. 2001) et le fait que la relation sédiment-débit n'est pas toujours très forte (Dodds et Whiles 2004).

Le délai entre le pic de CSS et celui du débit, le premier précédant le second de 1 jour (ou 9 heures lorsque l'analyse porte sur les données horaires du fleuve), démontre une hystérèse positive, phénomène qui a été observé dans d'autres systèmes hydriques (Asselman 1999; Oeurng et al. 2010). Ce type d'hystérèse peut indiquer que les sédiments proviennent d'une source proche et facilement disponible, comme le chenal et les berges (Oeurng et al. 2010). Cela indique également que les pointes de CSS sont probablement davantage reliées aux précipitations, qui rendent les sédiments facilement disponibles, qu'à l'augmentation de débit subséquente.

La deuxième variable hydrométéorologique identifiée, encore une fois sur le fleuve et son tributaire, est la précipitation du jour précédent. En fait, la portion liquide des précipitations à elle seule semble jouer le plus grand rôle, puisqu'aucune corrélation entre la CSS et la neige n'a été détectée et que la corrélation avec la pluie est à peine plus faible que pour la précipitation totale. Il est à noter que la neige a pu être étudiée seulement sur la rivière Kennebecasis puisque le suivi de la turbidité du fleuve STJ n'a pas été fait durant l'hiver. La corrélation entre la CSS et la précipitation du jour précédent indique que la réponse du système aux événements érosifs n'est pas instantanée. Similairement, l'importance des conditions de sol précédent un événement pluvieux (qui influence le pouvoir érosif de cet événement), l'importance de la lame d'eau ruisselée ainsi que les mécanismes de transport de sédiments en résultant sont des facteurs qui sont partiellement pris en compte par la relation (et la corrélation) entre la CSS et la précipitation cumulée.

Lorsque les données sont regroupées en périodes de temps plus longues (mois, année), les coefficients de corrélation entre la CSS moyenne ou les charges sédimentaires et les

variables explicatives s'améliorent. Avec une série de donnée plus longue, il est possible qu'un meilleur modèle d'estimation des charges puisse être développé sur le fleuve Saint Jean en utilisant les données regroupées par mois plutôt que par jour.

Le fait que les corrélations soient relativement faibles, et ce sur les deux cours d'eau, peut être dû à une multitude de causes.

D'abord, la grande taille des bassins versants, surtout celui du fleuve STJ, augmente la variabilité spatiale et temporelle des processus sédimentaires tel que l'érosion, le transport et le stockage, ce qui rend problématique l'extrapolation de données locales sur de grandes distances ou périodes de temps (Campbell 1992). Conséquemment, il est possible que l'utilisation d'une station météorologique unique, bien que située à proximité des sites de mesure de sédiments, limite le pouvoir explicatif de cette variable. L'utilisation d'une moyenne spatiale provenant de plusieurs stations sur tout le territoire du bassin versant, comme le font certains auteurs (Lopez-Tarazon et al. 2010; Oeurng et al. 2010), pourrait peut-être améliorer la relation, mais au prix d'une analyse plus compliquée et requérant davantage de ressources (temps, données, logiciels). Le débit sur la rivière Kennebecasis n'est que faiblement corrélé aux sédiments, situation qui a été notée sur d'autres cours d'eau (Dodds et Whiles 2004). Dans le cas du débit sur le fleuve, malgré le fait que les valeurs de débit soient jugées satisfaisantes pour la présente étude, la nécessité de recourir à une technique d'estimation très générale diminue la qualité et la fiabilité des données, ce qui joue probablement un rôle dans l'identification de fortes relations entre les variables. Finalement, il n'est pas exclu que des variables explicatives importantes dans le bassin du fleuve STJ autres que celles étudiées ici existent. En effet, différents processus peuvent influencer la mobilisation et le transport de sédiments dans les cours d'eau, tels la disponibilité des sédiments ainsi que les activités anthropiques comme l'agriculture, la foresterie, les mines et les sites de construction (Meade et al. 1990). De plus, la synchronisation de certains événements, comme la fonte des neiges jumelée à une pluie abondante (Lenzi et al. 2003) ou une faible pluie sur un sol agricole à nu (Lecce et al. 2006), influence la susceptibilité du sol à l'érosion ou le transport de sédiments résultant. Enfin, d'autres événements naturels comme la dynamique des glaces et les embâcles peuvent jouer un rôle majeur dans le transport de sédiments; cela a été démontré dans le fleuve STJ par Beltaos et Burrell (2000) et dans une autre rivière canadienne, la rivière Liard dans les Territoires du Nord-

Ouest, par Prowse (1993). Toutefois, ces processus peuvent être difficiles à quantifier, ou nécessitent une grande quantité de données et un traitement spatial beaucoup plus complexe qu'il n'est pas possible d'intégrer dans un modèle statistique relativement simple.

1.6.2 Estimation des CSS et charges de sédiments en suspension

1.6.2.1 Rivière Kennebecasis

Bien que, dans plusieurs systèmes, la majorité du transport de sédiments se produise lors de quelques événements extrêmes (Ashmore et Day 1988; Littlewood et Marsh 2005; Markus et Demissie 2006; Meade et al. 1990), l'analyse fréquentielle effectuée sur la rivière Kennebecasis démontre que l'étude des événements extrêmes (débit, précipitation et CSS) n'est pas suffisante pour prédire les années de fortes charges sédimentaires.

Le réseau de neurones artificiel (RNA) développé avec la série de données de CSS sur la rivière Kennebecasis a permis d'en estimer les charges annuelles de façon satisfaisante. À la base, ce modèle a été utilisé dans le but de vérifier si les variables potentielles identifiés précédemment étaient suffisantes pour expliquer la variation des CSS dans un contexte non-linéaire. Le modèle de CSS ne recrée pas les pics de concentrations de sédiments adéquatement. Comme discuté précédemment, cela tend à montrer que d'autres variables explicatives non prises en compte ici influencent la CSS. Toutefois, les résultats sont grandement améliorés lorsque la variable de sortie du modèle est la charge de sédiments en suspension. Ainsi, la performance du modèle est comparable à d'autres RNA trouvés dans la littérature (Alp et Cigizoglu 2007; Cigizoglu 2004; Cigizoglu et Kisi 2006; Jain 2001; Jain 2008; Sarkar et al. 2008). L'erreur sur l'estimation des charges ($\approx 14\%$) est aussi comparable à celle obtenue par Hansen et Bray (1993) à l'aide de courbes sédiments-débits, bien que notre résultat puisse être considéré meilleur puisque l'hiver et le printemps étaient exclus dans leur étude. Par contre, une meilleure estimation des charges annuelles, avec une erreur d'environ 11%, a déjà été obtenue dans d'autres grands bassins à l'aide de courbes sédiments-débits (Horowitz et al. 2001a).

1.6.2.2 Fleuve Saint Jean

L'utilisation de la turbidité pour l'estimation de la CSS sur le fleuve STJ est une technique prometteuse, bien que le protocole utilisé dans la présente étude puisse bénéficier de quelques améliorations. D'abord, les mesures sont prises en un seul point et doivent être comparées à davantage de valeurs provenant de toute la section du cours d'eau afin d'en vérifier la représentativité (Horowitz et al. 2001b). Dans la présente étude, les mesures prises sur la section du fleuve, généralement homogènes, semblent indiquer que les sédiments ne sont pas transportés de façon préférentielle dans une sous-section de la section transversale. Par contre, l'échantillonnage ponctuel sur plusieurs verticales, où un des échantillons pris à faible profondeur, près de la berge avait une CSS plus élevée, indique que le fait de placer le capteur à cet endroit le rend plus susceptible aux variations de sédiments locales, entre autres dues à la remise en circulation des sédiments du lit. Selon Beltaos et Burrell (2000), une différence systématique de CSS dans la section du fleuve peut exister juste en aval de tributaires, mais cette variation est rapidement dissipée par les processus de mélange. Les mesures de turbidité peuvent donc être considérées représentatives avec un certain degré de confiance, puisque les séries de données ont été examinées avec soin pour identifier et éliminer de la série les événements de remise en circulation locale. Il serait profitable de faire davantage de mesures sur toute la section du fleuve, à différents débits et au site de mesure avant le déploiement des instruments, afin de valider la représentativité du site de mesure. Deuxièmement, la courbe de calibration développée *in situ* afin de transformer les valeurs de turbidité en CSS, adéquate pour le premier site, a été jugée peu fiable pour le site 2, qui a dû être abandonné. Une meilleure méthode de calibration, semblable à celle suggérée par Eads et Lewis (2003), pourrait améliorer ces résultats. Ensuite, différents problèmes liés 1) au développement de biofilm sur le senseur par les algues, bactéries et plantes aquatiques et 2) à l'exondation des instruments lors de baisses de niveau d'eau importantes ont conduit à la perte considérable de données; contribuant entre autres à la décision d'abandonner le second site. Des mesures adéquates liées à un bon entretien des appareils déployés devraient être prises pour éviter cette situation dans le futur. Malgré tout, l'obtention d'une série de CSS avec un si petit intervalle de temps (15 minutes) aurait été impensable avec une méthode d'échantillonnage manuelle.

Bien que comportant plusieurs difficultés discutées ci-haut, l'utilisation de la turbidité en continu a un bon potentiel pour l'étude du transport de sédiments dans le fleuve Saint Jean. L'avantage majeur est la possibilité d'obtenir des données en continu en minimisant la probabilité de rater des événements majeurs. D'un autre côté, le manque de données hydro-météorologiques concomitantes à haute-fréquence force l'agrégation des données en plus longues périodes (au pas de temps journalier ou plus long); mais ces moyennes produisent probablement des données plus fiables qu'un échantillon ponctuel pris à un moment donné dans la journée. L'utilisation de mesures en continu permet également d'obtenir des données tant lors de conditions « normales » que lors d'événements de transport sédimentaire important. Dans son analyse d'un grand éventail de bassins à travers le monde, Meybeck et al. (2003) a montré que le pourcentage du transport total survenant 2% du temps (événements extrêmes) est plus faible pour les grands bassins, ce qui signifie que les conditions normales sont également importantes. Il était donc logique dans la présente étude de ne pas se concentrer uniquement sur ces événements, même s'ils jouent probablement un rôle majeur qui ne pourra être quantifié qu'avec une plus longue série de données. Par exemple, la crue extrême observée sur le fleuve du 23 au 27 novembre 2008, discutée plus en détail dans la section 2.2, indique que le régime de transport de sédiment est peut-être très différent pour les débits faibles et les débits forts, mais une conclusion claire ne peut pas être tirée d'un seul événement.

Finalement, les similitudes entre les variables indépendantes sélectionnées pour les modèles du fleuve STJ et la rivière Kennebecasis permettent de former l'hypothèse qu'un RNA pourrait être développé, pour l'estimation des charges annuelles sur le fleuve, si une série de données suffisamment longue existait. Pour l'instant, les données disponibles couvrent une trop courte période ou sont trop incomplètes pour être utilisées dans un RNA.

1.6.3 Liens entre le transport sédimentaire dans le bassin du fleuve STJ et les volumes dragués dans son estuaire

L'utilisation de données hydrométéorologiques simples et accessibles a permis la création d'une régression qui reproduit les quelques mesures de volumes de sédiments dragués dans l'estuaire de façon encourageante. Le volume d'eau total annuel (somme des débits) et le niveau d'eau moyen annuel à la station de Saint John, où se trouve le port, sont les deux variables incluses dans le modèle. Malgré le fait que les données de

débit proviennent d'une estimation et que seulement six (6) mesures annuelles de volumes dragués aient été utilisées, la régression ($r^2 = 0,87$) produit des valeurs déviant en moyenne de 15,4% de la valeur réelle. Bien qu'il y ait un risque qu'avec davantage de données, la structure de corrélation se dégrade avec une augmentation de la variance des mesures de volumes et que les variables explicatives identifiées ne soient pas les mêmes, il est également possible que la régression soit améliorée et l'erreur du modèle réduite avec une plus longue série de données. De plus, la qualité des données peut également être améliorée; les volumes dragués utilisés ici sont estimés par pesée de sédiments mouillés sur une barge lors de leur transport vers un lieu d'entreposage. Fait important à noter, il arrive que les volumes dragués ne représentent pas exactement l'accumulation dans le port puisque les différents entrepreneurs, qui peuvent changer d'une année à l'autre, ont des techniques différentes et rencontrent parfois des problèmes techniques les empêchant d'enlever tout ce qui devrait l'être. Il est donc possible qu'une année donnée, le volume de sédiments dragués soit découplé de la sédimentation dans le port, réduisant du fait même le lien avec les variables hydrologiques explicatives. Il est donc primordial de connaître tout problème technique lié au dragage lors du contrôle de la qualité des données. Finalement, puisque la régression repose sur des données hydrométéorologiques provenant de la même année que les volumes dragués, une prédiction des volumes de sédiments à draguer l'année suivante n'est pas encore possible, à moins d'obtenir une prévision hydrologique pour l'année à venir.

L'absence de relation entre les charges sédimentaires estimées à l'aide du RNA dans la rivière Kennebecasis et les volumes dragués dans le port peut provenir de diverses causes. D'abord, il n'est toujours pas certain que les sédiments s'accumulant dans l'estuaire proviennent du fleuve Saint Jean. Les marées et courants dans la Baie de Fundy jouent peut-être un plus grand rôle. Il est aussi possible que le sous-bassin versant de la rivière Kennebecasis soit en effet trop différent de celui du fleuve, en termes de taille et d'utilisation du sol, pour être représentatif du transport sédimentaire sur le fleuve. De plus, on note la présence d'un seuil à la confluence de la rivière Kennebecasis avec le fleuve STJ. En amont du seuil, les profondeurs sont relativement élevées, ce qui laisse penser que le tronçon le plus en aval de la rivière Kennebecasis est propice à la sédimentation. Les sédiments ont ainsi peu de chances d'être transportés directement vers l'estuaire lors de conditions hydrologiques moyennes.

Enfin, l'analyse de corrélations met en évidence que les principales variables explicatives pour les charges dans la rivière Kennebecasis et pour les volumes dragués ne sont pas les mêmes, la précipitation n'étant pas identifiée comme variable explicative pour les volumes dragués dans le port.

1.7 Conclusions et recommandations

Cette étude est la première à utiliser la turbidité en continu sur la branche principale du fleuve Saint Jean comme mesure indirecte de la concentration de sédiments en suspension (CSS). Une analyse de corrélation, une régression multiple linéaire pas à pas et une technique d'interpolation par krigeage dans un espace multivarié ont été utilisées afin d'identifier des variables hydrométéorologiques potentielles pour la variation de CSS.

Afin de réaliser une première étude exploratoire pour l'identification de variables potentielles, une série historique de CSS provenant d'un tributaire majeur, la rivière Kennebecasis, a d'abord été étudiée à l'aide de l'analyse fréquentielle, de l'analyse des corrélations et de réseaux de neurones artificiels.

Les données de turbidité et la série historique de CSS sur l'affluent ont permis d'identifier plusieurs variables influençant la concentration de sédiments en suspension (CSS), dont le débit, la précipitation du jour précédent et la précipitation accumulée jusqu'à 12 jours. Par contre, il est aussi évident dans toutes les analyses effectuées que 1) d'autres variables explicatives, potentiellement d'origine anthropique, existent; et que 2) les relations entre diverses variables et la CSS comportent souvent des non-linéarités qui pourraient expliquer en partie la difficulté à estimer adéquatement la CSS à l'aide de techniques de régression linéaire. Bien qu'il n'ait pas été possible, en partie à cause de la courte période de prise de données, de lier les observations dans le fleuve aux sédiments dragués dans l'estuaire, une analyse de la relation entre ces volumes dragués et diverses variables hydrométéorologiques s'est montrée prometteuse. Les principales variables qui semblent pouvoir expliquer une large proportion de la variance des six (6) mesures de volumes annuels disponibles sont le volume d'eau total annuel et le niveau d'eau moyen, mais une plus longue série de données sera nécessaire pour confirmer ces résultats. Il existe d'ailleurs des archives de relevés bathymétriques qui pourraient être utilisés pour estimer les volumes dragués par le passé, allongeant ainsi la série. Ces données n'ont pas pu être incorporées au présent mémoire à cause de difficultés techniques liées à leur format mais il est recommandé que davantage de ressources techniques soient fournies pour permettre l'utilisation de ces relevés.

Dans le contexte du fleuve Saint Jean (STJ), le principal avantage de l'utilisation de la turbidité comme mesure indirecte de la concentration de sédiments en suspension

(CSS) est la possibilité d'avoir des mesures en continu sans être physiquement présent sur le terrain. Même si ces valeurs doivent être agrégées pour des pas de temps plus longs (i.e. journalier, au moins), la prise de mesure à haute fréquence assure une meilleure qualité de données que les mesures ponctuelles. La qualité des données obtenues pourra aussi être améliorée en modifiant la méthode de création de la courbe de calibration, en utilisant par exemple une pompe intégrée au système permettant de prendre des échantillons d'eau au site de mesure même. Une analyse du type et de la taille des sédiments devrait aussi être conduite à quelques reprises lors de différents événements et à différentes concentrations, afin de s'assurer que la courbe créée est adéquate pour toute la période de mesure. Cette analyse granulométrique serait également utile pour fin d'identification des sources de sédiments.

Une plus longue période de suivi sera nécessaire sur le fleuve pour étudier la variabilité interannuelle de la CSS, en plus d'inclure idéalement l'hiver et le printemps. De plus, une plus longue série de mesures de turbidité pourrait permettre de reprendre l'analyse sur une base événementielle, ce qui est intéressant puisqu'il est possible que les sources de sédiments, la dynamique de resuspension et les phénomènes d'hystérèse soient différents pour des événements d'ampleur différente. De plus, il est primordial de s'assurer que les instruments soient placés en un endroit de la rivière représentatif de toute la section. Davantage de mesures sur la section devraient ainsi être prises lors d'un pré-échantillonnage précédant l'installation des instruments afin d'évaluer quantitativement la variabilité latérale de la turbidité et de valider la représentativité du point de mesure. L'utilisation de piliers de ponts et/ou de mouillages en eau plus profonde afin de rapprocher les instruments du chenal pourrait être envisagée, mais un accès facile aux instruments doit être maintenu afin de permettre un entretien plus fréquent.

Finalement, il serait intéressant d'inclure, dans la conception de la suite de l'étude, l'examen du transport de sédiments plus en aval, dans la portion du fleuve entre le site d'étude discuté ici et le port où se déroulent les activités de dragage. Cette section du cours d'eau est caractérisée par une hydrodynamique plus complexe puisque les marées et la présence d'eau saumâtre y jouent un rôle important; cependant, un suivi dans la zone estuarienne s'avère primordial afin de vérifier si les sédiments observés au site plus en amont sont effectivement transportés jusqu'à l'embouchure.

En somme, bien que la présente étude ne permette pas de prédire l'ampleur des opérations de dragage nécessaires une année donnée dans le port de Saint John, le suivi de la turbidité sur le fleuve Saint Jean a permis de mieux comprendre les liens entre les conditions hydroclimatiques et les concentrations de sédiments en suspension sur ce cours d'eau. Ces liens pourraient être établis de manière plus précise avec la poursuite du projet.

1.8 Liste des références

- Alp, M. et Cigizoglu, H. K. 2007. Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. *Environmental Modelling & Software*, 22(1):2-13.
- Ashmore, P. E. et Day, T. J. 1988. Spatial and temporal patterns of suspended-sediment yield in the Saskatchewan River basin. *Canadian Journal of Earth Sciences*, 25:1450-1463.
- Asselman, N. E. M. 1999. Suspended sediment dynamics in a large drainage basin: the River Rhine. *Hydrological Processes*, 13(10):1437-1450.
- Asselman, N. E. M. 2000. Fitting and interpretation of sediment rating curves. *Journal of Hydrology*, 234:228-248.
- Beltaos, S. et Burrell, B. C. 2000. Suspended sediment concentrations in the Saint John River during ice breakup. *In* 2000 Annual Conference Abstracts - Canadian Society for Civil Engineering, p 75.
- Bray, D. I. et Xie, H. 1993. A regression method for estimating suspended sediment yields for ungauged watersheds in Atlantic Canada. *Canadian Journal of Civil Engineering*, 20:82-87.
- Campbell, I. A. 1992. Spatial and temporal variations in erosion and sediment yield. *In* Erosion and Sediment Transport Monitoring Programmes in River Basins, IAHS Publ. no 210:455-465.
- Chokmani, K. et Ouarda, T. B. M. J. 2004. Physiographical space-based kriging for regional flood frequency estimation at ungauged sites. *Water Resources Research*, 40(12):1-13
- Cigizoglu, H. K. 2004. Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. *Advances in Water Resources*, 27(2):185-195.
- Cigizoglu, H. K. et Kisi, Ö. 2006. Methods to improve the neural network performance in suspended sediment estimation. *Journal of Hydrology*, 317(3-4):221-238.
- Collins, A. L. et Walling, D. E. 2004. Documenting catchment sediment sources: problems, approaches and prospects. *Progress in Physical Geography*, 28(2):159-196.
- Dodds, W. K. et Whiles, M. R. 2004. Quality and quantity of suspended particles in rivers: Continent-scale patterns in the United States. *Environmental Management*, 33(3):355-367.
- Eads, R. et Lewis, J. 2003. *Turbidity Threshold Sampling in Watershed Research* (Arcata, CA), (USFS).
- Filizola, N. et Guyot, J. L. 2009. Suspended sediment yields in the Amazon basin: An assessment using the Brazilian national data set. *Hydrological Processes*, 23(22):3207-3215.
- Gao, P. 2008. Understanding watershed suspended sediment transport. *Progress in Physical Geography*, 32(3):243-263.

- Gray, J. R. et Glysson, G. D. 2002. Proceedings of the federal interagency workshop on turbidity and other sediment surrogates. *In* Turbidity and other sediment surrogates, eds J. R. Gray & G. D. Glysson U.S. Geological Survey, Reston, Virginia. 56p.
- Grenfell, S. E. et Ellery, W. N. 2009. Hydrology, sediment transport dynamics and geomorphology of a variable flow river: The Mfolozi River, South Africa. *Water SA*. 35(3) : 271-282.
- Guillemette, N., St-Hilaire, A., Ouarda, T. B. M. J., Bergeron, N., Robichaud, E. et Bilodeau, L. 2009. Feasibility study of a geostatistical modelling of monthly maximum stream temperatures in a multivariate space. *Journal of Hydrology*, 364(1-2):1-12.
- Guyot, J. L. et Walling, D. E. 2009. Preface - Flow and sediment dynamics of large rivers. *Hydrological Processes*, 23:3127-3130.
- Hansen, D. et Bray, D. I. 1993. Single-station estimates of suspended sediment loads using sediment rating curves. *Canadian Journal of Civil Engineering*, 20:133-143.
- Hooper, R. P., Aulenbach, B. T. et Kelly, V. J. 2001. The National Stream Quality Accounting Network: a flux-based approach to monitoring the water quality of large rivers. *Hydrological Processes*, 15(7):1089-1106.
- Horowitz, A. J. 2003. An evaluation of sediment rating curves for estimating suspended sediment concentrations for subsequent flux calculations. *Hydrological Processes*, 17(17):3387-3409.
- Horowitz, A. J., Elrick, K. A. et Smith, J. J. 2001a. Annual suspended sediment and trace element fluxes in the Mississippi, Columbia, Colorado, and Rio Grande drainage basins. *Hydrological Processes*, 15(7):1169-1207.
- Horowitz, A. J., Elrick, K. A. et Smith, J. J. 2001b. Estimating suspended sediment and trace element fluxes in large river basins: methodological considerations as applied to the NASQAN programme. *Hydrological Processes*, 15(7):1107-1132.
- Jain, S. K. 2001. Development of integrated sediment rating curves using ANNs. *Journal of Hydraulic Engineering*, 127(1):30-37.
- Jain, S. K. 2008. Development of Integrated Discharge and Sediment Rating Relation Using a Compound Neural Network. *Journal of Hydrologic Engineering*, 13(3):124-131.
- Kisi, O., Karahan, M. E. et Sen, Z. 2006. River suspended sediment modelling using a fuzzy logic approach *Hydrological Processes*, 20(20):4351-4352.
- Knighton, D. 1998. *Fluvial forms and processes: A new perspective*. Arnold, London p 383.
- Lecce, S. A., Pease, P. P., Gares, P. A. et Wang, J. Y. 2006. Seasonal controls on sediment delivery in a small coastal plain watershed, North Carolina, USA. *Geomorphology*, 73(3-4):246-260.
- Lenzi, M. A., Mao, L. et Comiti, F. 2003. interannual variation of suspended sediment load and sediment yield in an alpine catchment *Hydrological Sciences*, 48(6):899-915.
- Leys, V. 2007. 3D Flow and Sediment Transport Modelling at the Reversing Falls - Saint John Harbour, New Brunswick. in *OCEANS 2007*, Vancouver, BC. 16p.

- Littlewood, I. G. et Marsh, T. J. 2005. Annual freshwater river mass loads from Great Britain, 1975-1994: estimation algorithm, database and monitoring network issues. *Journal of Hydrology*, 304(1-4):221-237.
- Lopez-Tarazon, J. A., Batalla, R. J., Vericat, D. et Balasch, J. C. 2010. Rainfall, runoff and sediment transport relations in a mesoscale mountainous catchment: The River Isabena (Ebro basin). *Catena*, *in press*. doi: 10.1016/j.catena.2010.04.005
- Mailhot, A., Rousseau, A. N., Talbot, G., Gagnon, P. et Quilbé, R. 2008. A framework to estimate sediment loads using distributions with covariates: Beaurivage River watershed (Québec, Canada). *Hydrological Processes* 22(26): 4971-4985.
- Mano, V., Moatar, F., Coynel, A., Etcheber, H., Ludwig, W., Meybeck, M., Nemery, J., Poirel, A., Blanc, G. et Schafer, J. 2006. Space and time variability of suspended particulate matter (SPM) transport in 32 French rivers (100 to 100 000 km²; daily to yearly). *Sediment Dynamics and Hydromorphology of Fluvial Systems*, ICCE IAHS International Symposium, 3-7 July 2006, Dundee, Scotland, (Poster Report Booklet):29-37.
- Markus, M. et Demissie, M. 2006. Predictability of annual sediment loads based on flood events. *Journal of Hydrologic Engineering*, 11(4):354-361.
- Meade, R. H., Yuzyk, T. R. et Day, T. J. 1990. Movement and storage of sediment in rivers of the United States and Canada. *In* *Surface Water Hydrology*, Geological Society of America, Boulder, Colorado, USA, Vol O-1, pp 255-280.
- Meybeck, M., Idlafkih, Z., Fauchon, N. et Andreassian, V. 1999. Spatial and temporal variability of total suspended solids in the Seine basin. *Hydrobiologia*, 410:295-306.
- Meybeck, M., Laroche, L., Durr, H. H. et Syvitski, J. P. M. 2003. Global variability of daily total suspended solids and their fluxes in rivers. *Global and Planetary Change*, 39(1-2):65-93.
- Neu, H. A. 1960 Hydrographic of Saint John Harbour, NB. National Research Council of Canada, Mechanical Engineering Report MH-97. 264p
- Oeurng, C., Sauvage, S. et Sanchez-Perez, J.-M. 2010. Dynamics of suspended sediment transport and yield in a large agricultural catchment, southwest France. *Earth Surface Processes and Landforms* 35(11):1289-1301.
- Partal, T. et Cigizoglu, H. K. 2008. Estimation and forecasting of daily suspended sediment data using wavelet-neural networks. *Journal of Hydrology*, 358(3-4):317-331.
- Pfannkuche, J. et Schmidt, A. 2003. Determination of suspended particulate matter concentration from turbidity measurements: particle size effects and calibration procedures. *Hydrological Processes*, 17(10):1951-1963.
- Picouet, C., Hingray, B. et Olivry, J. C. 2001. Empirical and conceptual modelling of the suspended sediment dynamics in a large tropical African river: the Upper Niger river basin. *Journal of Hydrology*, 250:19-39.
- Picouet, C., Hingrey, B. et Olivry, J. C. 2009. Modelling the suspended sediment dynamics of a large tropical river: the Upper Niger River basin at Banankoro. *Hydrological Processes*, 23:3193-3200.

- Prowse, T. D. 1993. Suspended sediment concentration during river ice breakup. *Canadian Journal of Civil Engineering*, 20:872-875.
- Rao, A. R. et Hamed, K. H. 2001. *Flood Frequency Analysis*. New York, USA. 376 p.
- Sarkar, A., Kumar, R., Jain, S. et Singh, R. D. 2008. Artificial neural network models for estimation of sediment load in an alluvial river in India. *Journal of environmental hydrology*, 16:1-12.
- Simon, A., Dickerson, W. et Heins, A. 2004. Suspended-sediment transport rates at the 1.5-year recurrence interval for ecoregions of the United-States: transport conditions at the bankfull and effective discharge. *Geomorphology*, 58:243-262.
- Slattery, M. C. et Phillips, J. D. 2009. Controls on sediment delivery in coastal rivers. *Journal of Environmental Management*, *in press*.
- Tramblay, Y., T. B.M.J. Ouara A. St-Hilaire, J. Poulin. 2010. Regional estimation of extreme suspended sediment concentrations using watershed characteristics. *Journal of Hydrology* 380: 305-317.
- Trenhaile, A. S. 1998. *Geomorphology: a Canadian perspective*. Oxford University Press, Don Mills, Ont pp.190-191
- Trites, R. W. 1960. An oceanographical and biological reconnaissance of Kennebecasis Bay and Saint John River estuary. *Journal of the Fisheries Research Board of Canada*, 17(3):377-408.
- Wagner, R. J., Boulger Jr., R. W., Oblinger, C. J. et Smith, B. A. 2006. Guidelines and Standard Procedures for Continuous Water-quality Monitors: Station Operation, Record Computation, and Data Reporting. (USGS).
- Wall, G. R., Nystrom, E. A. et Litten, S. 2008. Suspended Sediment Transport in the Freshwater Reach of the Hudson River Estuary in Eastern New York. *Estuaries and Coasts*, 31:542-553.
- Walling, D. E. et Webb, B. W. 1988. The reliability of rating curve estimates of suspended sediment yield: some further comments. *In* *Sediment Budgets* (Proceedings of the Porto Alegre Symposium), IAHS publ. no. 174, pp 337-350.
- Wang, J. J., Lu, X. X. et Kummu, M. 2011. Sediment load estimates and variations in the lower Mekong River. *River Research and Applications* 27(1):33-46.
- Whiles, M. R. et Dodds, W. K. 2002. Relationships between Stream Size, Suspended Particles, and Filter-Feeding Macroinvertebrates in a Great Plains Drainage Network. *Journal of Environmental Quality*, 31:1589-1600.

Mis en forme : Police : (Par défaut)
Arial, Anglais (Canada)

Mis en forme : Police : (Par défaut)
Arial, Anglais (Canada)

Mis en forme : Police : (Par défaut)
Arial

Partie 2- ARTICLES

2.1 Suspended sediment dynamics in a tributary of the Saint John River, New Brunswick

Canadian Journal of Civil Engineering, 38(2):221-232.

Hélène Higgins, M.Sc. candidate, Canadian Rivers Institute, INRS-ETE, 490 De la Couronne St. Quebec City, Qc, G1K 9A9, Canada, helene.higgins@ete.inrs.ca,

André St-Hilaire, Professor, Canadian Rivers Institute, INRS-ETE, 490 De la Couronne St. Quebec City, Qc, G1K 9A9, Canada.

Simon C. Courtenay, Professor, Canadian Rivers Institute, University of New Brunswick, Department of Biology, 10 Bailey Drive, P.O. Box 4400 Fredericton, NB, E3B 5A3, Canada

Katy A. Haralampides, Associate Professor, Canadian Rivers Institute, University of New Brunswick, Department of Civil Engineering, 17 Dineen Drive, P.O. Box 4400, Fredericton, NB, E3B 5A3, Canada

2.1.1 Résumé

Des données hydrométéorologiques et de concentration de sédiments en suspension (CSS) historiques de la rivière Kennebecasis, un tributaire du fleuve Saint Jean au Nouveau-Brunswick, Canada, ont été examinées afin de déterminer les contrôles hydrométéorologiques de la variabilité des concentrations et charges de sédiments dans ce système hydrique. L'analyse de corrélations entre les CSS et des variables explicatives potentielles à divers pas de temps suggère que des régressions multiples ne seraient pas optimales pour estimer les CSS ou la charge, et que le décalage entre les pics de débit (Q), les précipitations et les CSS doit être pris en compte. Une analyse fréquentielle incluant les maximums annuels de CSS, Q et les événements de précipitation maximums a révélé qu'il n'existe pas d'explication unique aux événements extrêmes de CSS ou aux charges sédimentaires annuelles. Enfin, deux réseaux de neurones artificiels (RNA) ont été développés pour vérifier si les mêmes variables étudiées précédemment produisent de meilleurs résultats dans un contexte non-linéaire. Les résultats obtenus en utilisant la température moyenne, Q , $Q(t-1)$, $Q(t-2)$ et le jour julien comme intrants, pour estimer la charge journalière sont satisfaisants, avec 88% de la variance expliquée ainsi qu'une déviation moyenne absolue de 14% entre les charges annuelles calculées et estimées par le modèle.

2.1.2 Abstract

Historical hydro-meteorological and suspended sediment concentration (SSC) data from the Kennebecasis River, a tributary of the Saint John River in New Brunswick, Canada, were investigated to help understand what drives high sediment transport in that system. Analysis of correlations coefficients between SSC and potential drivers at various time steps suggested that multiple regressions would not be optimal for this purpose, and that lagged flow (Q) and precipitation should be taken into account in any model. A frequency analysis involving annual maxima of SSC, Q and precipitation events revealed there is no systematic unique driver of extreme annual SSC or high annual loads. Finally, Artificial Neural Network (ANN) models were developed to verify whether the variables examined previously would yield better results in a non-linear context. Network inputs were mean temperature, Q , $Q(t-1)$, $Q(t-2)$ and day-of-year. Using daily loads directly as a target in the network yielded satisfactory results, with 88% of the variance explained by the model and a mean absolute deviation between estimated and real annual loads of 14%. The ANN model systematically outperformed multiple linear regressions.

Subject headings: suspended sediment, correlation, frequency analysis, artificial neural networks.

2.1.3 Introduction

Sediments in rivers have a wide array of impacts that have been extensively studied for decades, for small and large rivers alike. In addition to being recognized as the most important source of pollution in North American rivers by the USEPA, sediments also have direct impacts on human activities and infrastructures such as impoundments (reservoir filling) and ports (Gray and Glysson 2002; Meade et al. 1990).

Suspended sediment concentrations (SSC) are both spatially and temporally highly variable (Horowitz et al. 2001b). They depend on river hydraulics, deposition and re-suspension processes as much as on sediment sources, which are linked to topography, soil type, vegetation cover and land use (Gao 2008). This variability makes it difficult to estimate or predict the total amount of sediments transported in the river for a given time period. In addition, sediment-related work in rivers is resource-intensive. Hence, over time, different methods have been developed to estimate sediment loads and yields in rivers without having to measure them continuously.

While it is generally accepted that no one single method is appropriate for all situations, the sediment-discharge rating curve (SRC) is the most frequently used approach to estimate SSC because it requires less time and fewer data than some deterministic models or sophisticated statistical approaches. Due to much criticism regarding its efficiency (Gao 2008), this method is increasingly used as a first exploratory tool before determining a more appropriate model (Horowitz 2003). In fact, for some hydrological systems a SRC appears to be adequate, while the relationship is weak for many rivers. For example, Dodds and Whiles (2004) found a significant correlation between flow (Q) and suspended sediment concentration (SSC), for 489 out of 619 rivers in the United States. They observed correlation coefficients (r^2) ranging from 0.01 to 0.91 with a median r^2 of 0.35, concluding that this method yields variable results. A review of the literature indeed shows that the results are not always satisfactory, with relative errors ranging between -80% and 900% compared to real values (Gao 2008). Furthermore, the SRC relies on a single explanatory variable, discharge, which cannot account for all the variability in sediment transport.

Extreme SSC are also of interest when investigating annual loads in a river as a few events, such as those triggered by floods or snowmelt, can sometimes represent the bulk of the annual suspended sediment transport (e.g. Ashmore and Day 1988). SSC

extremes can be characterized in a probabilistic context by estimating their return periods (average number of years between two extremes of same amplitude). Therefore, another interesting exploratory tool is frequency analysis, which was used in a study by Trambly et al. (2008) to examine the magnitude and occurrence of extreme annual SSC events in 179 rivers in North America. They were able to calculate return periods and quantiles of extreme annual SSC for these rivers with different watershed characteristics, showing that this method can be suitable for a wide array of river systems. Their results also emphasize the site-specificity of extreme SSC events.

Many other methods have been developed to palliate the inaccuracies in sediment estimation and prediction. Gao (2008) reviewed the techniques and models for sediment monitoring as well as transport and load estimation such as sediment rating curves, multi-linear regressions, empirical models and artificial neural networks (ANNs). The latter can also be particularly useful to verify whether the selected predictors are appropriate, as it is a powerful non-linear tool that is typically able to detect even the slightest links between variables. In addition, when daily or higher frequency data are used, ANNs allow targeting of specific points where the model is wrong, which can help detect missing explanatory variables.

2.1.3.1 The Kennebecasis River, tributary of the Saint John River

A better understanding of inter-annual variability of sediment loads coming from the river system would benefit future sediment-related studies in the Saint John estuary. There is also a need to determine which explanatory variables can explain some of the variance in the magnitude of suspended sediment loads in that river system. The objective of this study was thus to examine the potential causes of suspended sediment concentrations and load inter-annual variability.

One major difficulty lies in the lack of long-term concomitant sediment and flow data for the main Saint John River. Therefore, a first exploratory analysis relies on historical data from the Kennebecasis River, a tributary of the Saint John River, which are analyzed to help understand what drives high sediment transport in that river system.

Specifically, the objectives of this analysis are to:

- 1) Describe the links between different hydro-climatic variables, SSC and loads by means of correlations
- 2) Determine annual extremes of SSC and other variables, and detect whether they coincide through a frequency analysis
- 3) Detect whether using non-linear relationships would yield better results than linear regressions, using the same variables, via the use of artificial neural networks

2.1.4 Study area and description of data

2.1.4.1 Study area

The Saint John River (SJR) flows for about 673 km between Maine (U.S.) and the Bay of Fundy, in New Brunswick (Canada), draining an area of 55 500 km². About half of the watershed (28 860 km²) lies in New Brunswick the rest being in Quebec (Canada) and Maine (US). While multiple tributaries flow into the SJR, the Kennebecasis River is the one having the most extensive sediment data record. It is also located in the lower portion of the basin, forming first the Kennebecasis Bay before joining the SJR and flowing into the Bay of Fundy (Figure 2.1.1).

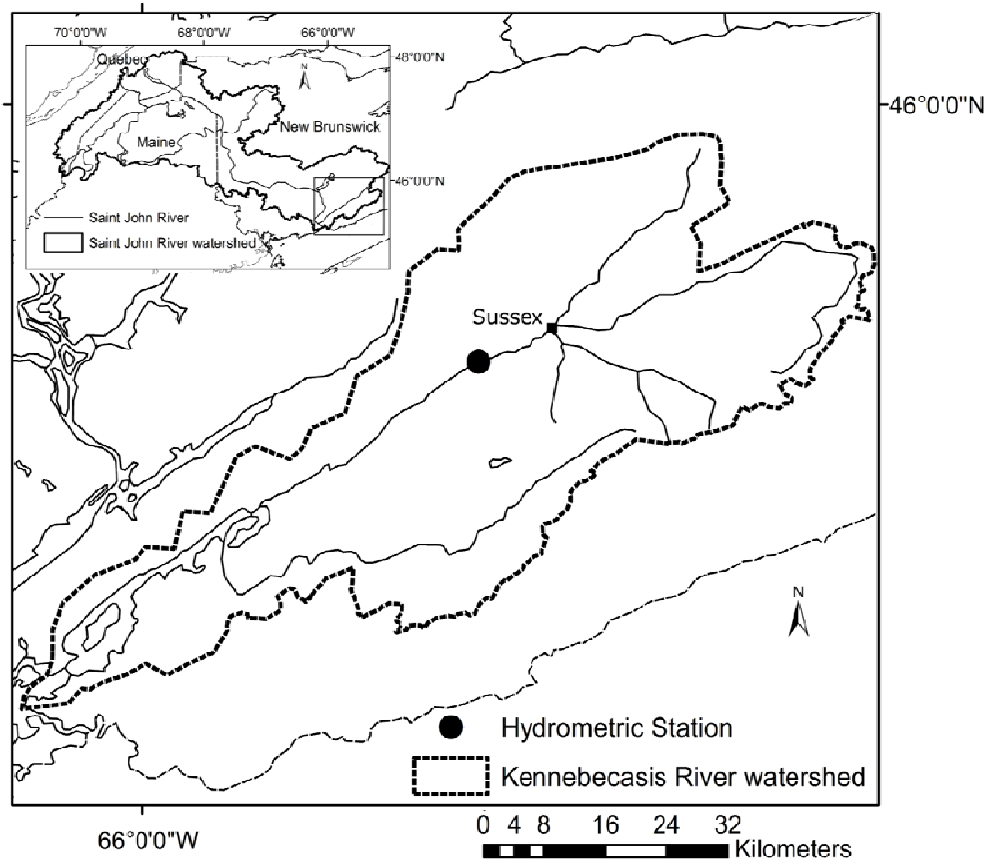


Figure 2.1.1 Kennebecasis River basin, with the location of the WSC hydrological station (01AP004) and weather station in Sussex. Upper inset is the whole Saint John River watershed. (Source: data available from Geobase)

The lower portion of the Saint John River watershed, which includes the Kennebecasis River sub-basin, experiences average air temperatures between 16°C and 18°C during summer and between -6°C and -9°C during winter. It receives 1200 mm of precipitation annually. The geology is composed mainly of sedimentary and some igneous rocks.

According to Bray and Xie (1993), the SJR watershed below Mactaquac Dam is approximately homogeneous in terms of climate, geological and hydro-climatic conditions. Land-use, presented in Table 2.1.1, is also similar between the Kennebecasis and the lower Saint John, except for agriculture which occupies a greater proportion of the territory in the Kennebecasis sub-basin (21.8%,) than in the whole New Brunswick section of the Saint John River basin (6%). While these differences in terms

of land use and basin size are likely to affect sediment processes, the two basins share enough similarities to use the tributary data to conduct a first investigation about potential explanatory variables for sediment dynamics in the Saint John River.

Table 2.1.1 Land-use comparison of the lower Saint John River watershed, below Mactaquac Dam, and the sub-basin upstream hydrometric station 01AP004 in the Kennebecasis River basin

Land use	Lower Saint John River	Upper Kennebecasis River
Forest	80.5%	73.8%
Agriculture	6%	21.8%
Urban	2.3%	1.4%
Wetlands	0.3%	0.2%
Water	4.9%	0.1%
Other	5.9%	2.8%

Source: data Land Cover circa 2000 available from GeoBase

2.1.4.2 Data description

The daily flow (Q) and suspended sediment concentration (SSC) data, available through Water Survey Canada (WSC), came from Environment Canada HYDAT station 01AP004 on the Kennebecasis River at Apohaqui (45°42'5"N 65°36'6"W). The station drains a territory of 1100 km² (Figure 2.1.1). Both datasets were complete between July 1st, 1966 and December 31st, 1987. While a previous study assumed the data contained no unacceptable errors (Bray and Xie 1993), there is one period, from September 19th 1981 to October 17th, 1981, that was considered dubious in the present study due to an unusually long-lasting SSC peak. The data were kept in the analysis, but conclusions pertaining to that period were considered potentially tainted by erroneous SSC values. Table 2.1.2 presents some basic descriptive statistics of the datasets. SSC exhibits a skewed ($c_{sx}=8.4$) and quite variable ($c_v= 2.02$) distribution compared to the other variables. Alp and Cigizoglu (2007) pointed out that these characteristics can make the modeling of SSC quite complex. Observation of annual hydrographs reveals there are usually two floods per year, in spring and fall.

Table 2.1.2 Daily statistical parameters (mean $\bar{x}$, median x_{med} , maximum x_{max} , minimum x_{min} , standard variation s_x , skewness coefficient c_{sx} and coefficient of variation c_v) of the hydrometeorological data. The inter-annual range of values, the values for the whole dataset as well as the specific statistics for the two datasets used in the ANN are presented.

	Precipitation (mm)		Flow (m ³ /s)		SSC (mg/L)		Load (T/day)	
	range	all data	range	all data	range	all data	range	all data
$\bar{x}$	2.4-4.1	3.1	16.4-40.2	25.7	11-56	21.9	39-190	90.6
x_{med}	0	0.0	7.4-28.8	14.0	3-17	10.0	2.6-38	13.2
x_{max}	37.3-89.2	89.2	129-520	520.0	93-814	814.0	1400-11457	11457
x_{min}	0	0.0	1.4-4.8	1.0	0-4	0.0	0-1.8	0
s_x	5.6-9.1	7.2	20.3-47.5	32.1	16.4- 136.4	44.3	124-685	369
c_{sx}	2.6-5.7	3.8	2.0-7.0	3.8	1.4- 8.9	8.4	4.1-13.4	15.8
c_v	2.0-2.8	2.32	0.96-1.8	1.25	1.0-2.4	2.02	1.9-5.9	4.07
	training	Test	training	test	training	test	training	test
$\bar{x}$	3.1	3.1	27.1	21.4	23.7	16.8	96.9	71.7
x_{med}	0.0	0.0	15.1	10.7	11.0	8.0	15.4	6.9
x_{max}	89.2	52.1	520.0	310.0	814.0	419.0	11457	3909
x_{min}	0.0	0.0	1.0	1.6	0.0	1.0	0	0.17
s_x	7.4	6.6	33.5	27.2	48.2	29.0	400.8	250.7
c_{sx}	3.9	3.4	3.8	3.1	8.2	5.5	15.8	7.7
c_v	2.39	2.13	1.24	1.27	2.03	1.73	4.14	3.50

Daily loads were obtained from the daily Q and SSC data using the following equation:

$$[1] \quad L_t = b C_t Q_t$$

where L is the load in tonnes, C and Q are sediment concentration in mg/L and discharge in m^3/s representative of the period t , and b is a constant to convert the units for the appropriate time period. For daily values, $b = 0.0864$. For the Kennebecasis river, most (80%) of the sediments are discharged in approximately 40% of the days in any one year. The computation of seasonal and annual loads allows investigating the seasonality of transport and determining when most of the transport is occurring, i.e. during the relatively wet fall period (Oct-Dec) or during the snowmelt season (approximately springtime, March-May). On average, 49.5% of the annual load occurs during the snowmelt season and 24% during fall (Figure 2.1.2).

The closest meteorological station within the Kennebecasis River watershed, station number 8105200 from Environment Canada, is located at Sussex ($45^\circ 43' 1.1994''\text{N}$ $65^\circ 31' 58.8''\text{W}$). All daily precipitation (snow and rain) and temperature (maximum and mean) data used in this study come from this station, which is located near the center of the Kennebecasis sub-basin (Figure 2.1.1). Any missing days (41 days in the whole series) were gap-filled using multiple regressions with three other surrounding stations, namely at Gagetown, Oromocto and Moncton.

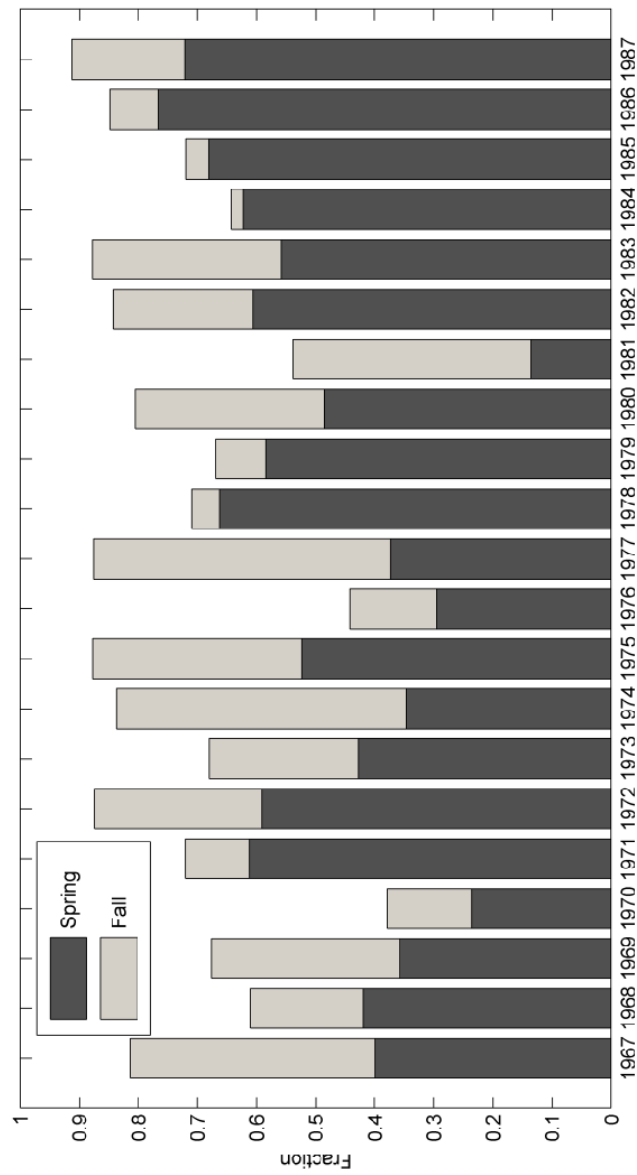


Figure 2.1.2 Fraction of the annual load discharged during spring and fall for each year of the time series

2.1.5 Methods

2.1.5.1 Correlation analysis

As a first step in detecting explanatory variables, Pearson product-moment correlation coefficients were calculated between SSC descriptive statistics (mean, maximum), water flow, precipitation and temperature data at different time steps. Daily data with temporal lags were first investigated between SSC, flow and precipitation. Then, aggregated data (monthly and annual) were explored. For each hydrological variable, the mean and maximum of all daily values for the time period studied (month or year) were extracted. Different trials were made with daily rain, snow and total precipitation. Precipitation events of longer duration (3, 6 and 12 consecutive days) were also computed.

2.1.5.2 Frequency analysis

To characterize probabilistically extreme flow and precipitation events, suspended sediment concentrations and annual sediment loads, a frequency analysis was performed. This is a widely-used method described in many textbooks (e.g. Rao and Hamed 2001). It allows relating the magnitude of events to their frequency of occurrence through the use of probability distributions. For a given event q with a return period T , the amplitude of which is exceeded only once in T years, the probability of exceedance is:

$$[2] \quad P(Q_T > q) = 1 / T$$

and the cumulative probability of occurrence $F(Q_T)$ is:

$$[3] \quad F(Q_T) = P(Q_T \leq q) = 1 - P(Q_T > q) = 1 - (1 / T)$$

$F(Q_T)$ is estimated by fitting a known statistical distribution to annual or seasonal maxima. One can estimate a quantile Q_T for a given return period T (e.g. 10 years, 100 years etc.).

The following annual maxima were extracted from the daily data sets: SSC, Q and 3, 6 and 12 days precipitation event. Total annual precipitation was also included to verify whether wet years caused high annual loads. Each time series were then analyzed using the HYFRAN software (http://www.ete.inrs.ca/activites/groupe/chaire_hydrol/hyfran.html), following the method described by Trambly et al. (2008). This software allows testing the data for the three

hypotheses that must be verified prior to performing a frequency analysis: homogeneity (Wilcoxon test), stationarity (Kendall test) and independence (Wald-Wolfowitz test). Subsequently, sixteen different statistical distributions available in HYFRAN (see Trambly et al. 2008) were fitted to each time series, and the best distribution was chosen first by using the Akaike (Akaike, 1974, cited in Trambly et al. 2008) and Bayesian (Schwarz 1978) information criteria and by visual observation of the goodness of fit. Return periods for the annual events can then be estimated from the fitted distribution. The maximum likelihood method was used to estimate the parameters of the statistical distribution fitted to series of annual extremes.

2.1.5.3 Artificial Neural Network (ANN)

An ANN is a mathematical model that tries to mimic the arrangement of biological neural networks, i.e., highly interconnected neurons arranged in a series of layers, each one linked to the next. It allows the creation of powerful non-linear models that are trained with examples (past data) to learn the patterns and are then able to recognize them in new datasets (Jain 2001). Typically, the model requires input data that are transformed, summed and fed to the subsequent layers containing weighted non-linear function (e.g. sigmoid, exponential). Weights are adjusted in order to yield the desired output and to minimize the error on the model. It is a technique that is increasingly used in hydrology in general, including suspended sediment studies (Alp and Cigizoglu 2007) as it allows to better model the hysteresis effect than multi-linear regressions (Alp and Cigizoglu 2007; Jain and Chalisgaonkar 2000), and allows to integrate variables other than flow.

Hysteresis is the lagging effect in a system where SSC and flow do not peak at the same time. It means that for a given flow, SSC will be different depending on whether it occurs during the rising or falling limb of the hydrograph. Sarkar et al. (2008) found the ANNs superior to regular rating curves for the estimation of sediment loads in an alluvial river. Alp and Cigizoglu (2007) estimated sediments loads using ANNs and obtained a similar overall result as with multi-linear regressions, but found the sediment hydrograph obtained with ANN to be closer to real values than the one obtained with the multiple regressions. To estimate SSC at time t , many authors use suspended sediment concentration at time $t-1$ as an input in the ANN (Jain 2001; Jain 2008; Sarkar et al. 2008). This is akin to a non-linear autoregressive model.

In this study, artificial neural networks were used to verify whether the explanatory variables investigated were adequate and sufficient to model SSC and sediment loads in the Kennebecasis River. By looking at daily predictions of SSC, it will be possible to detect where the model is wrong and target specific causes, and thus possibly missing explanatory variables. A feed-forward backpropagation neural network was used to predict daily SSC, which were later used to calculate daily loads. Different combinations of inputs and number of hidden neurons were tried to obtain the optimal network structure, which is defined as the most powerful model using the least data. Inputs considered were mean air temperature, flow and precipitation from same and previous days, rain, snow, 3, 6 and 12-day precipitation events and day-of-year (DOY). Using flow difference (ΔQ) instead of past flow ($Q(t-1)$) was also investigated, as relative change in flow may be as important as the initial flow value for sediment transport. Hidden layers contained a number of tangent sigmoid functions and one final linear function to allow the outputs to exceed the -1 to 1 range (Demuth et al. 2006). Trials were made with two or three layers containing between 1 and 20 log sigmoid or tangent sigmoid neurons. The network contains only one output layer, SSC. Three-quarters of the original data (from July 1st, 1966 to August 7th, 1982) were used to train the network using the Levenberg-Marquardt optimization algorithm (Demuth et al. 2006). The training was pursued until an acceptable level of error was attained or until the number of iterations exceeded 100. The rest of the dataset (from August 8th, 1982 to December 1st, 1987) was used to test the performance of the network after training. All inputs were standardized (0 mean, standard deviation of 1) before being fed to the network. The output (a) was back transformed to unnormalized units. The root mean squared errors (RMSE) of both subsets of data (training and test) were calculated, as well as the overall coefficient of determination (r^2) between the model output and the target value.

There is an ongoing debate (see Gilroy et al. 1990; McBean and Al-Nassri 1988; Shivers and Moglen 2008) as to whether using directly the load instead of SSC is an acceptable practice when using sediment rating curves. It does not seem to be an issue in ANNs, and using loads as a target is common in the literature (Alp and Cigizoglu 2007; Cigizoglu 2004; Cigizoglu and Kisi 2006; Sarkar et al. 2008). This technique was thus applied in this study to allow comparison of the results with the literature. Hence, another ANN was developed, using the same method as described above except for the output which was daily loads (L).

In order to verify whether using a non-linear model is worthwhile, multiple regressions were investigated between SSC and the strongest explanatory variables; i.e. the results of the best ANN models were compared with a multiple regressions including the same variables as the neural network. The aim was not to find the best linear model, but to verify that ANN indeed has an added value compared to linear regressions. RMSE and the correlation coefficient between the model and target value were calculated and compared to the ANN results.

2.1.6 Results

2.1.6.1 Correlation analysis

Daily correlations between SSC and flow as well as loads and flow reveal that, as expected, Q explains more of the load variation ($r^2=0.44$) than SSC variation ($r^2=0.12$, Table 2.1.3). Daily precipitation or temperature data were not significantly correlated with SSC and Q. The relationships improve when lagged data (i.e. from previous or subsequent days) were used; Figure 2.1.3 shows the strength of the correlation, using daily data, as a function of temporal lags between SSC and a) precipitation (P) and b) discharge (Q). On the x axis, the lags in days are expressed as negative when SSC is correlated with hydroclimatic values from previous days and positive when the hydroclimatic variables are measured on the following days. The maximum correlation occurs with a lag of +1 for flow ($r^2=0.13$) and -1 for precipitation ($r^2=0.09$). Hence, SSC peaks lead flow maxima (positive, or clockwise, hysteresis). These two variables, even with the appropriate lag, do not seem sufficient to accurately predict SSC in a multiple regression: daily SSC as a dependant variable and Q(t+1) and P(t-1) as predictors only yielded a r^2 of 0.17.

Table 2.1.3 Coefficients of determination (r^2) between hydro-climatic variables using daily data

Daily data	SSC r^2	Q r^2	Load r^2
<i>a) hydrological variables</i>			
SSC	1		
Q	0.12	1	
Load	0.35	0.44	1
<i>b) climatic variables</i>			
Mean temp.	-	-	-
Rain	0.01	-	-
Snow	-	-	-
Precipitation	0.01	-	-
3-day precipitation	0.01	0.09	0.01
6-day precipitation	0.05	0.17	0.07
12-day precipitation	0.04	0.16	0.05

Note: All values shown are significant with $p < 0.05$

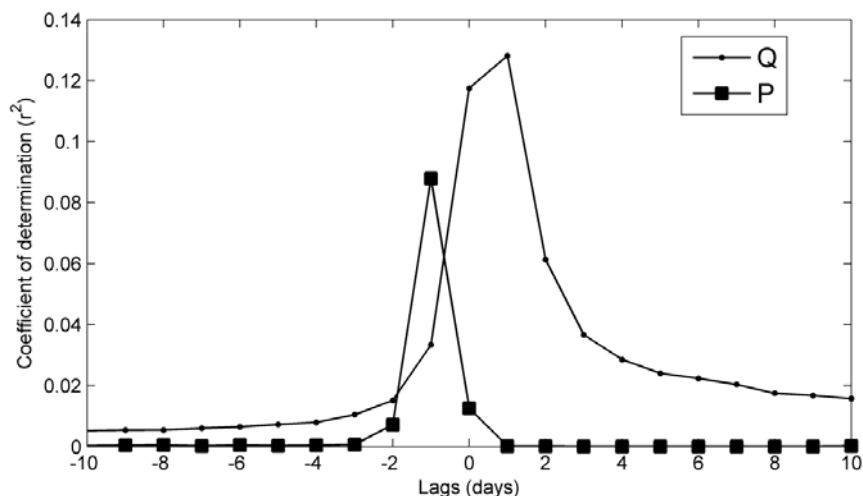


Figure 2.1.3 Correlation results of SSC with flow (Q) and precipitation (P) for different lags

Table 2.1.4 shows the improvement of the relationships when data were aggregated using a monthly or annual time period between a) the hydrological variables together and b) the hydrological and climatic variables. The strongest relationship overall ($r^2=0.71$) was observed between total precipitation and mean discharge on an annual basis. Unsurprisingly, the strongest relationship among sediment and hydrological

variables was observed between load and flow ($r^2=0.69$), but it must be kept in mind that loads were calculated by multiplying Q by SSC. A coefficient of determination of 0.37 was found between mean annual SSC and mean annual discharge. Using log (Q) and log (SSC) did not yield better results than using untransformed variables. Mean air temperature exhibits an r^2 of 0.25 with mean Q when looking at annual averages. In general, the relationships were better with data aggregated by year. This was not true, however, for extremes such as maximum flow which tend to have a stronger relationship with all hydrometeorological variables on a monthly basis. There was also a very weak ($r^2 < 0.1$) but significant relationship between hydrological variables and snow that was not detected on an annual basis.

Table 2.1.4 Coefficients of determination (r^2) between hydro-climatic variables using monthly and annually aggregated data

	Mean SSC		Max SSC		Mean Q		Max Q		Total load	
	monthly	annual	monthly	annual	monthly	annual	monthly	annual	monthly	annual
<i>hydrological variables</i>										
Mean SSC	1	1								
Max SSC	0.55	0.61	1	1						
Mean Q	0.12	0.37	0.16	0.29	1	1				
Max Q	0.12	0.39	0.20	0.21	0.68	0.28	1	1		
Total load	0.39	0.58	0.45	0.59	0.65	0.68	0.69	0.41	1	1
<i>climatic variables</i>										
Mean temp.	0.03	-	0.03	-	0.04	0.25	0.05	-	0.04	-
Total rain	0.03	0.21	0.06	-	0.06	0.64	0.09	-	0.06	0.48
Total snow	0.01	-	-	-	0.03	-	0.03	-	0.02	-
Tot precipitation	0.08	0.25	0.10	-	0.12	0.71	0.19	0.2	0.12	0.28
Max 3-days precip	0.03	0.30	0.06	-	0.05	0.19	0.17	-	0.09	-
Max 6-days precip	0.05	0.46	0.06	-	0.07	0.25	0.16	0.18	0.10	-
Max 12-day precip	0.04	0.26	0.05	-	0.09	0.35	0.13	0.3	0.10	0.31
Max 12-day rain	-	0.48	0.03	0.28	0.03	0.34	0.04	0.21	0.03	0.48

All values shown are significant with $p < 0.05$

The seasonal relationships were also investigated by analyzing the monthly data for specific seasons. Two time periods were tried (Sept-Dec and Feb-May, inclusively). The results were similar to those obtained when using the whole year, except for total monthly loads that tend to be more strongly correlated with mean and maximum SSC for the Sept-Nov period ($r^2 = 0.76$ and 0.62).

2.1.6.2 Frequency Analysis

All series of annual maxima and total precipitation met the hypotheses of stationarity, homogeneity and independence at the 5% level. Three two-parameter distributions were selected: the gamma (precipitation), the exponential (flow) and the lognormal (SSC).

Figure 2.1.4 shows the return periods associated to each variable for annual extremes extracted from the data set. The first observation is that the return periods of extreme SSC, Q or precipitation do not match systematically. In addition, when compared to the load values, illustrated as “+” on Figure 2.1.4, there was no systematic pattern linking high extremes with the highest loads. In general, the high annual loads (>40000T) correspond to either infrequent maximum SSC or Q, or a combination of variables having a somewhat higher return period than average (>5years). There were some exceptions to this pattern. The highest annual load occurred in 1981; that year corresponds to the highest return period for maximum SSC, maximum annual Q and maximum 12-days precipitation. It was also characterized by the second highest return period for total precipitation. However, the second highest load, in 1979, does not correspond to a high quantile for maximum SSC, but had an unusually high amount of precipitation (with a return period > 30 years). The most exceptional 3-day precipitation event occurred in 1970, which also corresponds to a somewhat higher load than average.

The timing of maximum SSC and maximum flow occurrence was also investigated. For 17 out of the 21 years, maximum discharge occurred during springtime. Maximum SSC occurred during spring for only 13 years, two of which (1969 and 1973) the maximum annual flow occurred in fall. The most frequent month for maximum SSC was April (n=8) while it was March (n=5) or April (n=5) for maximum flow. Even when the season in which they each occur was the same, there were only four occasions on which the timing was in the same day or in consecutive days; in most other cases, maximum SSC and Q did not even occur during the same month.

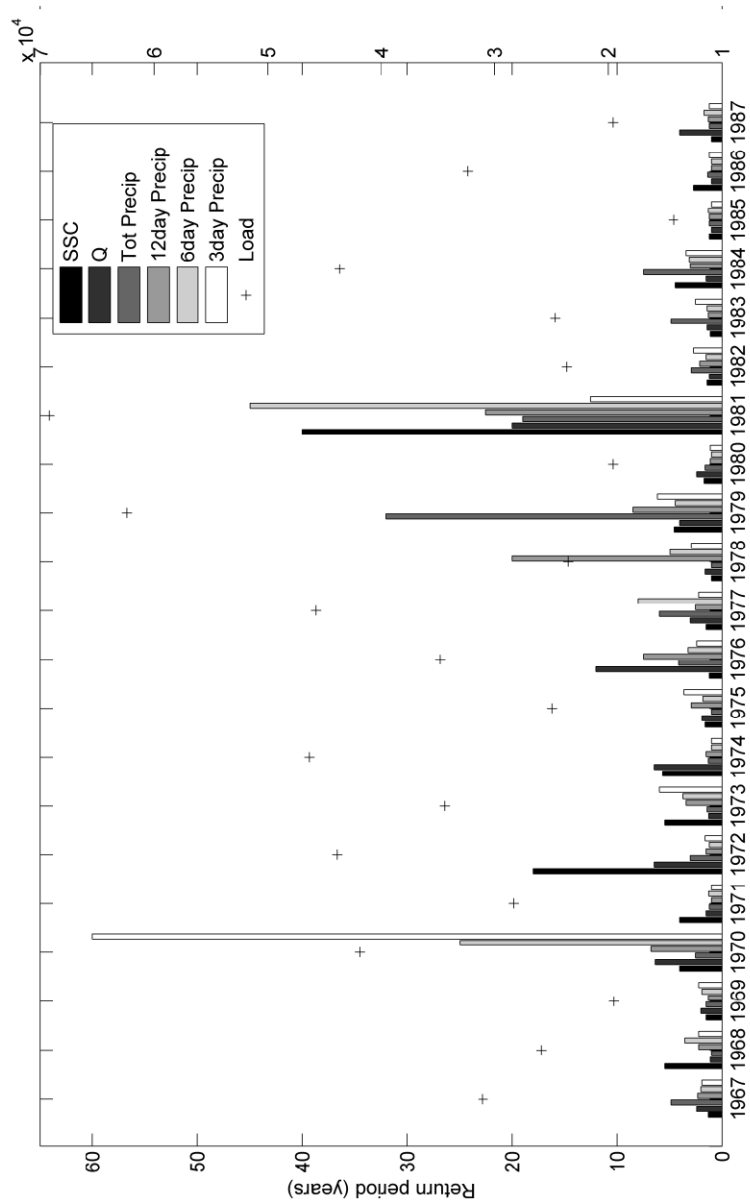


Figure 2.1.4 Return periods for maximum SSC, maximum Q, total precipitation and maximum 12-day precipitation event, for each year

2.1.6.3 Artificial Neural Networks

The following results pertain to one ANN modeling SSC (ANN-SSC), one ANN modeling loads (ANN-L) and the corresponding linear regressions. The statistics of the training and test datasets (Table 2.1.2) were similar for precipitation, flow and SSC, with the training set being slightly more scattered. Different sets of inputs were tried to model SSC or loads. The most relevant are presented in Table 2.1.5. Other combinations involving precipitation and difference in flow are not shown here since they did not yield better results, nor did adding a dummy variable to represent seasons or splitting the data between high and low flows. The best, most parsimonious models, for both ANN-SSC and ANN-L, include mean air temperature, Q, Q(t-1), Q(t-2) and day of year (DOY) as inputs. The models contain two hidden layers, the first having six (6) tangent sigmoid neurons and the second one (1) linear neuron.

Table 2.1.5 Inputs, r^2 and RMSE for different ANN tried, using either SSC (ANN-SSC) or loads (ANN-L) as a target for the model

Model inputs	ANN-SSC			ANN-L		
	r^2	RMSE		r^2	RMSE	
		Train	Test		Train	Test
Q	0.09	46.84	25.02	0.77	246	204
Q, Q(t-1), Q(t-2)	0.13	44.69	21.41	0.82	171	149
T _{mean} , Q, Q(t-1), Q(t-2)	0.20	43.95	21.72	0.82	178	148
T _{mean} , Q, Q(t-1), Q(t-2), P(t-1)	0.22	43.78	21.87	0.78	186	159
T _{mean} , Q, Q(t-1), Q(t-2), 12dayP	0.24	42.89	25.96	0.88	138	170
T _{mean} , Q, Q(t-1), Q(t-2), DOY	0.23	43.21	20.93	0.88	136	147
T _{mean} , Q, ΔQ_1 , ΔQ_2 , DOY	0.19	44.92	23.06	0.84	159	183

Note: r^2 represents the coefficient of determination between the model and the target value

Figure 2.1.5A shows that there was a noteworthy underestimation of high SSC values, especially for observed concentrations above 400 mg/L. Most of those points correspond to the dubious 1981 event mentioned in "Data description". The correlation improved from $r^2=0.23$ to $r^2=0.41$ when this event was eliminated from the database, and RMSE for the training set decreased to 27.04 mg/L, compared to 42.89 mg/L. In addition, when the analysis was limited to SSC < 400 mg/L, the model showed a much improved performance, with r^2 of 0.39 and RMSE of 26.02 mg/L and 20.33 mg/L for the training and test sets. When the model targets were loads, high events were better estimated, as seen in Figure 2.1.5-B. The model resulted in an r^2 of 0.88 and RMSE of 136T and 147T for the training and test sets. In that case, removing the 1981 event did not change the

results significantly. To allow direct comparison between the two models, relative RMSE was calculated and yielded values of 0.79 and 0.71 for ANN-SSC, while ANN-L performed better with 0.33 and 0.56 for the training and test sets, respectively.

The linear multiple regressions corresponding to each ANN, i.e. using same inputs and output, had systematically much lower r^2 and higher RMSE. These values were, respectively, $r^2 = 0.16$ and 0.59 and $RMSE = 44.8$ mg/L (training), 23.4 mg/L (test), $255T$ (training) and $155T$ (test) for SSC and loads respectively.

Since the interest lies in the inter-annual variability, model outputs were used to calculate annual loads that were compared to those calculated from the original Water Survey Canada (WSC) data. For the ANN-SSC, loads were calculated by multiplying the daily SSC output of the model by daily Q from WSC. Loads from the multiple regressions yielded poor results and are thus not presented here.

On an annual basis, the mean absolute deviation between the WSC data and estimated annual loads was 19% and 14% for ANN-SSC and ANN-L, respectively. ANN-SSC tended to underestimate the load more often (19 years) than ANN-L (11 years, Figure 2.1.6). Not taking the 1981 event into account only improved slightly the results for the ANN-SSC model (18% mean absolute deviation). The results of the linear regressions are not shown because the mean deviation was above 100%.

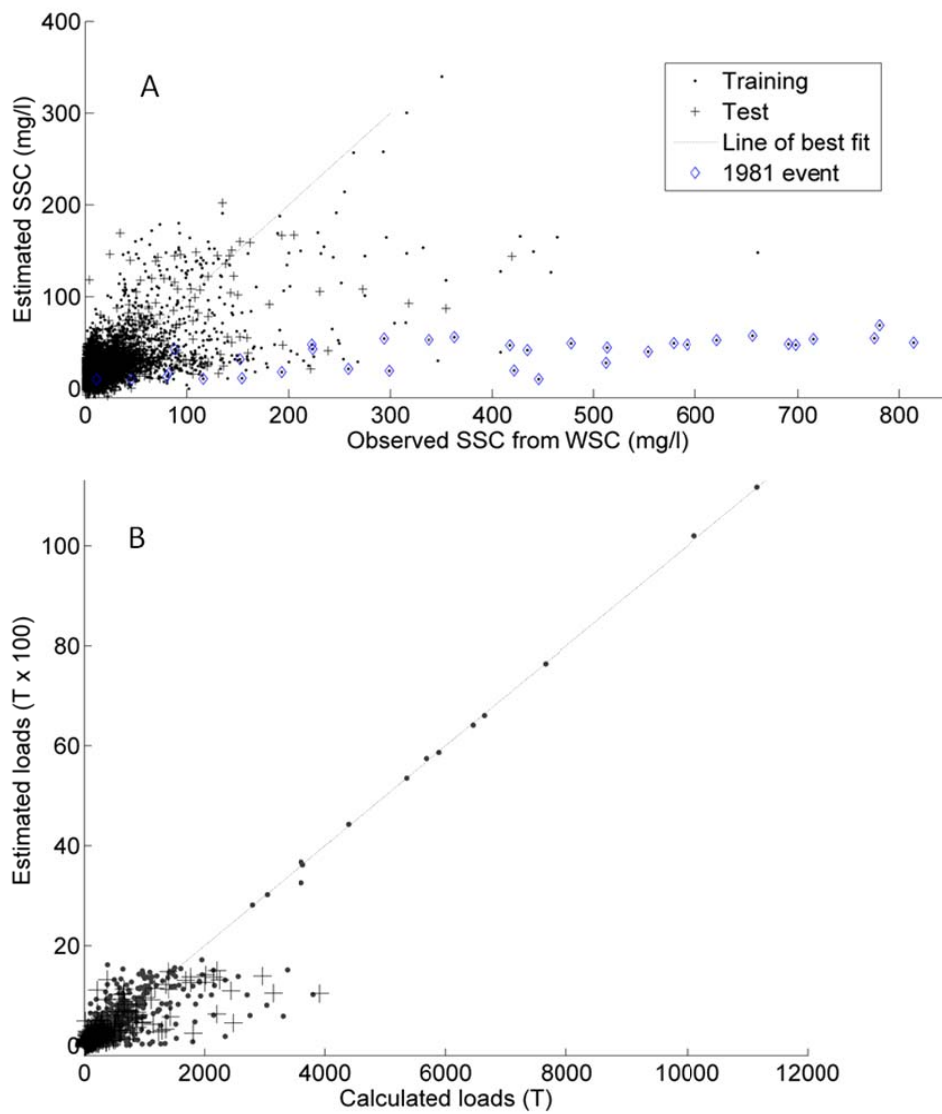


Figure 2.1.5 Daily estimates of SSC (A) and loads (B) using the optimal ANNs plotted against observed or calculated values. Both the training and test sets are shown.

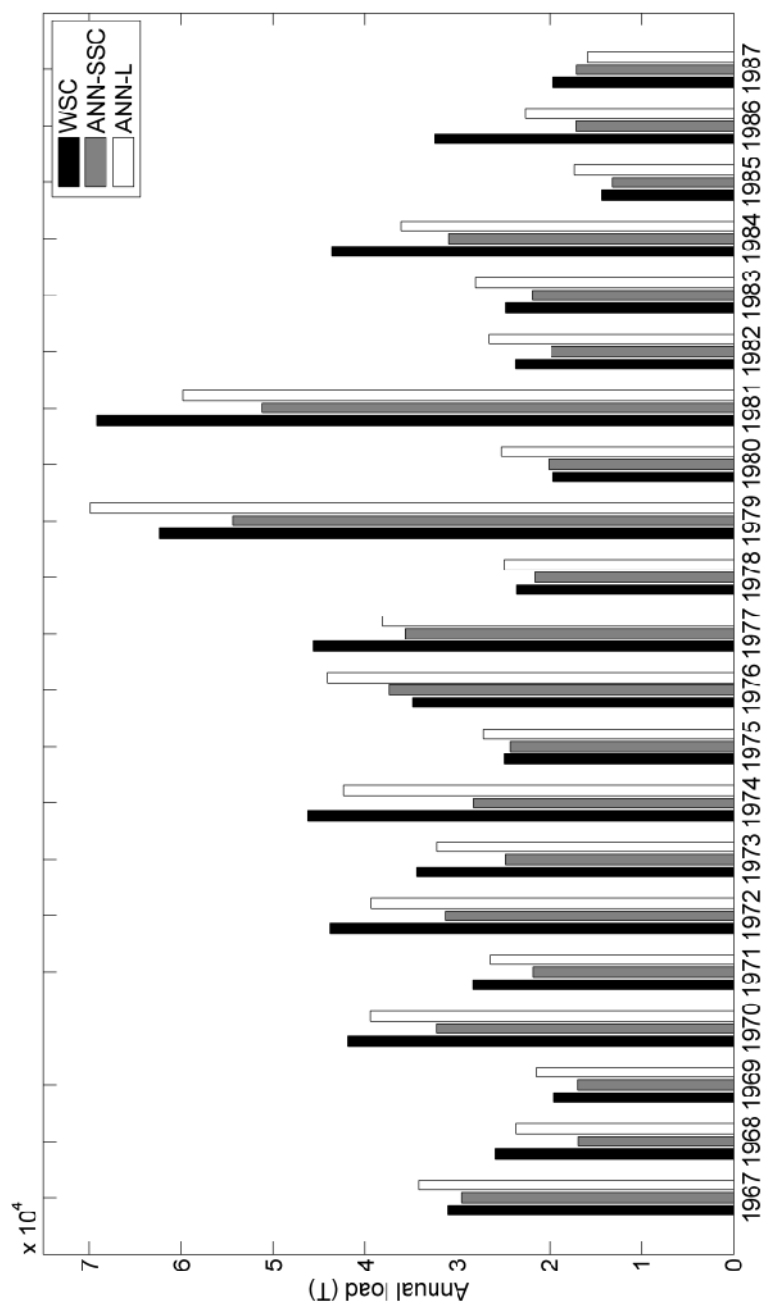


Figure 2.1.6 Comparison of annual suspended sediment loads obtained from Water Survey Canada and estimated from the ANN model

2.1.7 Discussion

Similarly to what was found by Hansen and Bray (1993) on a daily basis, the correlations between annual SSC and Q were significant but not very strong, with r^2 between mean annual SSC and mean (maximum Q) of 0.37 (0.39). Dodds and Whiles (2004) also noted that many rivers do not exhibit a strong relationship between discharge and total suspended solids. Precipitation and temperature were thus examined and were also found to be linked with either mean SSC or mean Q, but to a lesser extent.

The clockwise hysteresis existing in the system makes the prediction of SSC with flow only more difficult because the same flow can result in very different SSC values whether the water levels are going up or down. The lag between SSC and precipitation also highlights the importance of taking previous days into account when trying to model SSC, which was done in both the ANN approach and multiple regressions.

While it is typical in many rivers to observe that most of the sediment transport occurs in a few events during the year (Ashmore and Day 1988; Littlewood and Marsh 2005; Markus and Demissie 2006; Meade et al. 1990) the study of annual extremes in the Kennebecasis River was not sufficient to detect high annual loads using hydro-meteorological predictors. High, but not necessarily extreme SSC events can be quite scattered throughout the year, sometimes coinciding with high flows, which can cause important load events. However, the timing of annual maximum SSC and Q did not coincide systematically. For instance, a closer observation of the 1981 data revealed that extreme events for SSC, Q and 12-day precipitation do not coincide in time. This is consistent with the results of Trambly et al. (2008), who found only 49 rivers out of the 208 they studied with maximum SSC and Q occurring in the same month, and 151 in the same season. This indicates that the highest annual loads are likely caused by a combination of different events happening throughout the year. Hence in this river system, it is not only the snowmelt-driven spring flood that drives high annual loads. Other events must be taken into account.

The frequency analysis confirmed that there is no systematic unique driver of extreme annual SSC or high annual loads. High sediment loads were not systematically observed in conjunction with one extreme SSC, Q or P event. While the highest annual loads seemed to be caused by at least one of those variables being extreme, the recurrence of the variables examined cannot be used to predict the recurrence of the loads because

no clear pattern was detected in the analysis. Thus, while frequency analysis remains a useful tool to predict return periods of given events, it cannot be used in this case to detect potential explanatory variables for high annual loads or predict the magnitude of the load for a specific year.

2.1.7.1 ANNs

As expected, the results of the ANN were better when using loads directly rather than SSC. The results of ANN-L were comparable to other studies where acceptable results were obtained with r^2 ranging between 0.71 and 0.93 (Alp and Cigizoglu 2007; Cigizoglu 2004; Cigizoglu and Kisi 2006; Jain 2001; Jain 2008; Sarkar et al. 2008). It must be noted, however, that in most cases, authors used sediment data either from previous days or an upstream station as inputs in the model (Cigizoglu 2004; Cigizoglu and Kisi 2006; Jain 2001; Jain 2008; Sarkar et al. 2008). This is a technique that was avoided here since the objective was to identify hydro-meteorological variables that can be used to estimate SSC. In that regard, while there is no physical process explaining a relationship between SSC and air temperature, the latter can be a proxy for the time of year (season). A non-linear relationship can thus be found as warmer air temperatures occur in summer when there is also typically less sediment transport in the studied river, while the most important transport events occur during spring and fall when temperatures are lower than in summer.

In addition to looking at the RMSE and r^2 statistics, computing the annual loads is a way to assess the performance of the ANN. When using the modeled daily SSC and daily Q data to do so, annual loads were systematically underestimated, because SSC peaks were not well represented. A method such as rating curves may provide better results in some cases, but for the wrong reasons, e.g. a systematic overestimation of low SSC values and underestimation of high values (Horowitz et al. 2001a). Over a long time period, those two errors tend to cancel each other. As such, calculating the error of the annual load estimation is a first step, but looking at the results of the ANN on a daily basis really allows targeting specific events for which the model is wrong. The ANN models more closely the low SSC values, which overall results in a greater underestimation of the load for a given period. This is what was observed when comparing the results with those of Hansen and Bray (1993), who created rating curves for the months of April to October with data collected from 1970 to 1979 on the Kennebecasis River. The deviation (19%) between the estimated annual load from ANN-

SSC and measured data from WSC indicated that there was room for improvement in the method, which was achieved partly by ANN-L. The annual loads calculated from ANN-L were better than ANN-SSC yet slightly poorer (mean absolute deviation of 14%) than what was achieved in a few large American watersheds using rating curves, where the average deviation with observed data was 11% (Horowitz et al. 2001b). Hansen and Bray (1993) also obtained results within 16.4% of the WSC values for the April-October period using rating curves.

The SSC peaks that were identified when there is no corresponding high Q forces us to hypothesize that the model was likely incomplete and that land use information or other data linked to sediment availability should be included. This is especially true if understanding the variability in SSC rather than in annual loads was the focus of the study. While discharge is heavily affecting sediment transport, the availability of sediment also plays a major role, resulting in large SSC variations and substantial transport in some cases where flows are not extreme (Tramblay et al. 2008) or without major fluctuations in discharge (Horowitz, 1995, cited in Horowitz et al. 2001a). Those high SSC could be linked to anthropogenic impacts such as agriculture, forestry, mining, major construction sites or natural events such as ice-jam and ice-melt effects, which are all recognized as affecting sediment supply (Meade et al. 1990). For instance, Lecce et al. (2006) observed that more sediments are transported when agricultural practices cause the soil to be bare, even when rainfalls are neither intense nor abundant. Lenzi et al. (2003) noted that a rapid snowmelt combined with abundant rainfall causes significantly more sediment transport than a continuous snow cover during winter followed by a gradual snowmelt and scant rainfall. Ice breakup can also cause large increases in SSC with only moderate increases in flow (Prowse 1993). Hence, the real loads might be difficult to determine without taking into account anthropogenic and ice-related factors that are affecting sediment generation.

The results seem to indicate that air temperature, Q, Q(t-1), Q(t-2) and day of year (DOY), might not be sufficient to correctly estimate SSC variation using an ANN. However, it seems an adequate method to investigate inter-annual variation in loads. Although the findings of the present study focus on a tributary of the Saint-John River, the results offer a first understanding of potential drivers of sediment loads and the associated variability on the main river using the same type of hydrological variables. SSC in large basins tend to be less variable (Meybeck et al. 2003), and transport events

need to be longer to have a significant effect on the annual loads (Horowitz et al. 2001a). Thus, it is also possible that high SSC peaks caused by small, local events could be less frequent in the remainder of the Saint John River watershed or at least have a smaller impact on the overall annual load. This could help obtain yet better results.

Further work could include using another method such as neuro-fuzzy networks to account for the anthropogenic and ice-related effects that are hard to simulate in regular ANNs because they often require some kind of classification. This method has been effectively used by Cobaner et al. (2009) to predict daily SSC either with flow data only or by adding sediment data from previous days. Sediment peaks were also better represented by neuro-fuzzy models than with ANNs. Special care should also be taken in using data (flow, precipitation, land-use, etc.) that are representative of the whole watershed. While precipitation was not found to affect SSC in the Kennebecasis River significantly and high SSC values that were missed by the model did not seem to be caused by coinciding precipitation events, care must be taken in rejecting precipitation as an explanatory variable because all data came from only one weather station. More weather stations could be necessary to make a representative average of the whole watershed. Spatial distribution of precipitation may be another factor to be considered. Local rain events of high intensity may generate high SSC if the rain falls on bare, mobile soil surface. Finally, Hansen and Bray (1993) were unable to improve their sediment rating curve on the Kennebecasis River by using a different equation for high and low flows, but it worked for other studies, such as the one reported by Cigizolgu and Kisi (2006). Hence partitioning the data between different flows or seasons remains a technique worth exploring for future studies.

2.1.8 Conclusion

No strong linear relationships were found between SSC and Q, precipitation or temperature, on a daily, monthly or annual basis on the Kennebecasis River.

The system exhibits a clockwise hysteresis that could explain the poor correlations with SSC. Lags observed between SSC, flow and precipitation emphasize the need to take previous days into account when modeling the system.

A frequency analysis of annual maximums of SSC, Q and precipitation shows extreme events of those variables do not necessarily coincide in time. It also confirms that there is no unique driver of high SSC events or annual loads, and that frequency analysis was

not an appropriate tool in this case to detect potential explanatory variables or predict the magnitude of high annual loads for a specific year. It remains, nonetheless, useful for the prediction of return periods for events of a given magnitude.

A feedforward backpropagation neural network with mean temperature, Q , $Q(t-1)$, $Q(t-2)$ and day of year (DOY) as inputs was not able to model correctly all SSC peaks, especially those corresponding to low flows. This suggests there might be missing explanatory variables related to sediment sources such as land-use. Using categories as variables such as in adaptive neuro-fuzzy approaches (e.g. Cobaner et al. 2009) could be explored to better estimate peaks missed by a regular ANN.

The same ANN, with daily loads as target, was adequate for estimation of annual loads and yielded much better results than a linear regression using the same inputs. Hence explanatory variables that might not be adequate in a multi-linear regression can in fact be sufficient to model a system using non-linear models.

Given the similarities in land cover and climatology between the Kennebecasis sub basin and the Saint-John River basin, an ANN using hydrometeorological variables seems a promising method for the estimation of annual loads on the latter river, keeping in mind explanatory variables must be representative of the whole area under study.

2.1.9 Acknowledgments

The authors wish to thank Yves Trambly and Anik Daigle for their technical support, as well as Ben Whalen from the Kennebecasis Watershed Restoration Committee for relevant comments and discussions. This project was funded in part by NSERC.

2.1.10 References

- Akaike, H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19: 716-723.
- Alp, M. and Cigizoglu, H. K. 2007. Suspended sediment load simulation by two artificial neural network methods using hydrometeorological data. *Environmental Modelling & Software*, 22(1):2-13.
- Ashmore, P. E. and Day, T. J. 1988. Spatial and temporal patterns of suspended-sediment yield in the Saskatchewan River basin. *Canadian Journal of Earth Sciences*, 25:1450-1463.
- Bray, D. I. and Xie, H. 1993. A regression method for estimating suspended sediment yields for ungauged watersheds in Atlantic Canada. *Canadian Journal of Civil Engineering*, 20:82-87.
- Cigizoglu, H. K. 2004. Estimation and forecasting of daily suspended sediment data by multi-layer perceptrons. *Advances in Water Resources*, 27(2):185-195.
- Cigizoglu, H. K., and Kisi, Ö. 2006. Methods to improve the neural network performance in suspended sediment estimation. *Journal of Hydrology*, 317(3-4): 221-238.
- Cobaner, M., Unal, B., and Kisi, O. 2009. Suspended sediment concentration estimation by an adaptive neuro-fuzzy and neural network approaches using hydro-meteorological data. *Journal of Hydrology*, 367(1-2):52-61.
- Demuth, H., Beale, M., and Hagan, M. 2006. *Neural Network Toolbox for use with Matlab: user's guide*, MathWorks, Inc., Natick, MA.
- Dodds, W. K. and Whiles, M. R. 2004. Quality and quantity of suspended particles in rivers: Continent-scale patterns in the United States. *Environmental Management*, 33(3):355-367.
- Gao, P. 2008. Understanding watershed suspended sediment transport. *Progress in Physical Geography*, 32(3):243-263.
- Gilroy, E. J., Kirby, W. H., Cohn, T. A., and Glysson, G. D. 1990. Uncertainty in suspended sediment transport curves. *Journal of Hydrologic Engineering*, 116(1):143-145.
- Gray, J. R. and Glysson, G. D. 2002. Proceedings of the federal interagency workshop on turbidity and other sediment surrogates. *In Turbidity and other sediment surrogates, edited by J. R. Gray & G. D. Glysson* U.S. Geological Survey, Reston, Virginia.
- Hansen, D. and Bray, D. I. 1993. Single-station estimates of suspended sediment loads using sediment rating curves. *Canadian Journal of Civil Engineering*, 20:133-143.
- Horowitz, A.J. 1995. *The Use of Suspended Sediment and Associated Trace Elements in Water Quality Studies*. IAHS Special Publication No. 4. IAHS Press: Wallingford, UK; 58 pp.
- Horowitz, A. J. 2003. An evaluation of sediment rating curves for estimating suspended sediment concentrations for subsequent flux calculations. *Hydrological Processes*, 17(17):3387-3409.

- Horowitz, A. J., Elrick, K. A. and Smith, J. J. 2001a. Estimating suspended sediment and trace element fluxes in large river basins: methodological considerations as applied to the NASQAN programme. *Hydrological Processes*, 15(7):1107-1132.
- Horowitz, A. J., Elrick, K. A. and Smith, J. J. 2001b. Annual suspended sediment and trace element fluxes in the Mississippi, Columbia, Colorado, and Rio Grande drainage basins. *Hydrological Processes*, 15(7):1169-1207.
- Jain, S. K. 2001. Development of integrated sediment rating curves using ANNs. *Journal of Hydraulic Engineering*, 127(1):30-37.
- Jain, S. K. 2008. Development of Integrated Discharge and Sediment Rating Relation Using a Compound Neural Network. *Journal of Hydrologic Engineering*, 13(3):124-131.
- Jain, S. K. and Chalisgaonkar, D. 2000. Setting up stage-discharge relations using ANN. *Journal of Hydrologic Engineering*, 5(4):428-433.
- Lecce, S. A., Pease, P. P., Gares, P. A., and Wang, J. Y. 2006. Seasonal controls on sediment delivery in a small coastal plain watershed, North Carolina, USA. *Geomorphology*, 73(3-4):246-260.
- Lenzi, M. A., Mao, L., and Comiti, F. 2003). Interannual variation of suspended sediment load and sediment yield in an alpine catchment. *Hydrological Sciences*, 48(6): 899-915.
- Littlewood, I. G., and Marsh, T. J. 2005. Annual freshwater river mass loads from Great Britain, 1975-1994: estimation algorithm, database and monitoring network issues. *Journal of Hydrology*, 304(1-4): 221-237.
- Markus, M., and Demissie, M. 2006. Predictability of annual sediment loads based on flood events. *Journal of Hydrologic Engineering*, 11(4):354-361.
- McBean, E. A., and Al-Nassri, S. 1988. Uncertainty in suspended sediment transport curves. *Journal of Hydraulic Engineering*, 2(1): 63-74.
- Meade, R. H., Yuzyk, T. R. and Day, T. J. 1990. Movement and storage of sediment in rivers of the United States and Canada. In *Surface Water Hydrology*, Geological Society of America, Boulder, Colorado, USA, Vol O-1, pp 255-280.
- Meybeck, M., Laroche, L., Durr, H. H., and Syvitski, J. P. M. 2003. Global variability of daily total suspended solids and their fluxes in rivers. *Global and Planetary Change*, 39(1-2):65-93.
- Prowse, T. D. 1993. Suspended sediment concentration during river ice breakup. *Canadian Journal of Civil Engineering*, 20:872-875.
- Rao, A. R. and Hamed, K. H. 2001. *Flood Frequency Analysis*. New York, USA.
- Sarkar, A., Kumar, R., Jain, S. and Singh, R. D. 2008. Artificial neural network models for estimation of sediment load in an alluvial river in India. *Journal of environmental hydrology*, 16:1-12.
- Schwarz, G. 1978. Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461-464.
- Sciences and Reporting Branch. 2007a. *New Brunswick Watersheds- Kennebecasis River*. Available from the Department of Environment of New-Brunswick, Fredericton.

- Sciences and Reporting Branch. 2007b. New Brunswick Watersheds- Saint John. Available from the Department of Environment of New-Brunswick, Fredericton.
- Shivers, D. E., and Moglen, G. E. 2008. Spurious Correlation in the USEPA Rating Curve Method for Estimating Pollutant Loads. *Journal of Environmental Engineering*, 134(8):610-618.
- Tramblay, Y., St-Hilaire, A. and Ouarda, T. 2008. Frequency analysis of maximum annual suspended sediment concentrations in North America. *Hydrological Sciences Journal*, 53(1):236-252.
- Trites, R. W. 1960. An oceanographical and biological reconnaissance of Kennebecasis Bay and Saint John River estuary. *Journal of the Fisheries Research Board of Canada*, 17(3):377-408.

2.2 Suspended sediment concentrations derived from continuous turbidity measurements in relation to hydrometeorological variables in the Saint John River, New Brunswick

Soumis pour publication à la revue Hydrological Processes

Hélène Higgins* M.Sc. candidate, Canadian Rivers Institute, INRS-ETE, 490 De la Couronne St. Quebec City, Qc, G1K 9A9, Canada helene.higgins@ete.inrs.ca, tel : 418-654-2530 ext.4457

*Author for correspondance

André St-Hilaire, Professor, Canadian Rivers Institute, INRS-ETE, 490 De la Couronne St. Quebec City, Qc, G1K 9A9, Canada.

Katy A. Haralampides, Associate Professor, Canadian Rivers Institute, University of New Brunswick, Department of Civil Engineering, 17 Dineen Drive, P.O. Box 4400, Fredericton, NB, E3B 5A3, Canada

Simon C. Courtenay, Fisheries and Oceans Canada & Professor, Canadian Rivers Institute, University of New Brunswick , Department of Biology, 10 Bailey Drive, P.O. Box 4400 Fredericton, NB, E3B 5A3, Canada

2.2.1 Résumé

Dans le cadre d'une étude visant à améliorer la compréhension des phénomènes de transport sédimentaire menant à l'ensablement de son estuaire, des mesures de turbidité ont été prises à haute fréquence comme substitut à la concentration de sédiments en suspension (CSS) dans la portion inférieure du fleuve Saint Jean, un large cours d'eau situé principalement au Nouveau-Brunswick. Une courbe de calibration linéaire a été utilisée ($r^2=0,92$) pour convertir les données de turbidité en CSS, qui ont ensuite été agrégées en données horaires et journalières. La structure de corrélation entre la CSS et différentes variables hydrométéorologiques a été évaluée. Au pas de temps horaire le niveau d'eau (H) est la variable la plus fortement corrélée avec les CSS, avec un délai de +9 heures ($r = 0,28$). Cependant, la meilleure corrélation entre ces deux variables au pas de temps journalier ne montre aucun délai ($r = 0,43$). Bien que faiblement, la CSS journalière était aussi corrélée avec la température ($r = -0,15$), la précipitation du jour précédent ($r = 0,27$) et la précipitation cumulée sur 3, 6 et 12 jours ($r = 0,23, 0,17$ et $0,27$ respectivement). Les charges sédimentaires mensuelles ont également été étudiées. Afin de vérifier si les variables identifiées peuvent adéquatement expliquer la variance du logarithme de la CSS, une régression pas à pas a été comparée à une méthode alliant analyse multivariée et géostatistique. Une analyse de corrélation canonique (ACC) a été utilisée afin de construire un espace physiographique dans lequel la variable sédimentaire d'intérêt peut être projetée. Le krigeage ordinaire est ensuite utilisé pour estimer la valeur voulue par interpolation dans cet espace. Cette seconde méthode a donné des résultats un peu meilleurs (RMSE = 0,48, $r^2= 0,35$) que la régression (RMSE = 0,50, $r^2= 0,29$). Cette étude montre que la turbidité a un bon potentiel comme mesure indirecte des CSS permettant un suivi en continu. De plus longues séries d'observations des CSS et d'autres variables explicatives, de même que des modèles améliorés seront nécessaires pour mieux comprendre la dynamique sédimentaire dans ce système fluvial.

2.2.2 Abstract

In order to improve the understanding of sedimentary transport processes leading to deposition in its estuary, continuous turbidity measurements were used as a surrogate for suspended sediment concentration (SSC) in the lower reaches of the Saint John River, a large watercourse located mostly in New Brunswick, Canada. Turbidity measurements were successfully converted to SSC using a linear calibration curve ($r^2 = 0.92$). High frequency SSC measurements were aggregated into hourly and daily values. The correlation structure between SSC and hydrometeorological variables was assessed. At an hourly time step, SSC and water level (H) were best correlated with a lag of +9 hours ($r = 0.28$), but the best correlation at a daily time step showed no lag ($r = 0.43$). Albeit weakly, daily SSC was also correlated with temperature ($r = -0.15$), precipitation from the previous day ($r = 0.27$) and 3, 6 and 12 day cumulative precipitation ($r = 0.23$, 0.17 and 0.27 respectively). Monthly sediment loads were also investigated. In order to verify whether the identified variables could adequately explain the variance in SSC, a stepwise regression was compared to a multivariate and geostatistical approach. The latter method uses canonical correlation analysis (CCA) to construct a physiographical space in which SSC data can be projected. Ordinary kriging is subsequently used to estimate the desired values through interpolation. The latter method yielded better results than a stepwise linear regression, with a RMSE of 0.48 and r^2 of 0.35 compared to a RMSE of 0.50 and r^2 of 0.29 for the regression. This study shows turbidity has the potential to facilitate sediment monitoring, though other explanatory variables and models will be needed to further improve the understanding of suspended dynamics in this river.

Keywords: turbidity, suspended sediments, suspended load, geostatistical, multivariate analysis, large watershed

2.2.3 Introduction

In many rivers, washload is the main process by which sediments are transported (Knighton, 1998). Hence, sediment transport is often estimated by taking the suspended load only and disregarding bedload and dissolved particles. Suspended sediment loads are typically driven by 1) the sources of available washload that may be abundant or limited and 2) hydrometeorological forcings that route SS to the river and eventually downstream, with potential deposition and re-suspension. In most cases, there is limited information on sediment source availability, with the exception of land use data that can now be extracted from GIS databases.

Another important limitation in sediment load estimation is sampling frequency. In many studies, weekly to monthly (Horowitz *et al.* 2001; Mano *et al.* 2006; Picouet *et al.* 2009) or even quarterly (Filizola and Guyot 2009) time steps are used. Some calculations have been done using daily SSC and flow (Sivakumar *et al.* 2005; Meybeck *et al.* 2004). The use of hourly data has been mostly limited to applications and research on very small drainage basins. Even for such small basins, data are often integrated at the daily time scale (e.g. Pavey *et al.* 2007).

Many authors have successfully used turbidity as a surrogate for SSC (e.g. Birgand *et al.* 2004; Dodds and Whiles 2004; Oeurng *et al.* 2010; Pfannkuche and Schmidt 2003; Pavanelli and Bigi 2005; Lopez-Tarazon *et al.* 2010) because sediment-discharge rating curves are in many cases inadequate for the study of short-term transport (Horowitz 2003) and also because they do not allow studying individual events when flows are measured at a coarse (e.g. weekly, monthly) temporal scale. An added benefit to turbidity measurements is their continuous nature and relative high frequency which facilitates investigating sediment transport patterns at a relatively short time scale to gain a better understanding of the processes involved. SSC can be derived from the turbidity readings with an empirical site-specific calibration curve. Site specific calibration is preferred as water color, as well as the type, shape and size of sediment affect the relationship between turbidity measurements and associated SSC values (Gippel 1995; Marttila 2010). Studies based on turbidity measurements are less frequent in large rivers than in streams due to difficulties associated with large spatial variability of catchment characteristics and data acquisition logistics. In the USA, sediments in large rivers are still mostly studied through periodic grab samples (Horowitz *et al.* 2001). In Canada, Environment Canada conducted daily sampling on many rivers in the past, including

some large ones such as the Mackenzie, but none of these programs exist anymore. In France also, turbidity has been mostly used in basins $<1000 \text{ km}^2$ (Mano *et al.* 2006). Oeurng *et al.* (2010) studied the dynamics of sediment transport in a medium-sized (1110 km^2) agricultural catchment in France using continuous turbidity measurements and successfully linked transport events with precipitation and discharge variables. Still, the use of turbidity has the potential to simplify the investigation of sediment transport in large rivers, especially those that have been widely understudied in the past.

The Port of Saint John, a major transport hub for the province of New Brunswick (Canada), is located downstream of the Saint John River estuary in the southwestern part of this province. As many navigable waterways, frequent dredging is necessary to allow ship passage. In that context, having a sound knowledge of the sediment transport and deposition processes is critical in order to plan these operations. However, it is difficult to determine the proportion of sediments accumulating in the harbour that originate from the river versus from the Bay of Fundy. The sediment regime of the Saint John River (SJR) and its estuary have been insufficiently studied and no long term records of suspended or bedload transport exist on its main stem.

The main objective of this study was therefore to investigate the relationships between suspended sediment and hydrometeorological variables on the Saint John River drainage basin, which is the largest drainage basin in the province of New Brunswick. More specifically, the study investigates potential relationships at various time scales, including the hourly time step. To achieve this, specific objectives included the monitoring of turbidity in the lower reaches of the Saint John River, followed by calibration of these turbidity measurements against suspended sediment concentrations and performance of statistical analyses that assess the relationship between concomitant as well as lagged SSC, rainfall and flow.

2.2.4 Materials and method

2.2.4.1 Site Description

The Saint John River (SJR) flows for about 673 km between Maine (U.S.A.) and the Bay of Fundy, in New Brunswick (Canada), draining an area of $55\,500 \text{ km}^2$ (Figure 2.2.1). Estimated minimum summer and maximum spring flows at Saint John are $500 \text{ m}^3/\text{s}$ and $3400 \text{ m}^3/\text{s}$, respectively (Leys 2007), and mean annual discharge was estimated to be $991 \text{ m}^3/\text{s}$ (Neu 1960). About half of the watershed ($28\,860 \text{ km}^2$) lies in New Brunswick,

with the remainder being in Quebec (Canada) and Maine. The river system is relatively complex, with multiple tributaries flowing into the SJR, and with the presence of many small islands and a few large lakes in the watershed that are open to mixing with the river. In addition, river enlargements, low slopes and low currents in the downstream reaches produce areas in which deposition is likely to occur. One such area is Long Reach, where the main stem meets with the Kennebecasis River tributary before flowing into the Bay of Fundy.

In the NB portion of the Saint John River, flows are affected by several dams, the most downstream being the Mactaquac Dam Generating Station. While this dam, located about 19 km upstream of Fredericton, is a “run of the river” design, it creates a small reservoir (approximately 87 km²) where sediments can accumulate. This is obvious when looking at daily SS data taken downstream of the structure by the Water Survey of Canada (WSC) during the construction of the dam: after completion in 1967, daily measurements of concentrations decreased from about 31mg/l to 7 mg/l. Thus in this study, we consider that most of the sediments flowing towards the estuary originate from the portion of the basin drained downstream of the dam, which covers approximately 15 000 km². The head of tide is 3 km below Mactaquac Dam (Trites 1960). That portion of the watershed experiences average air temperatures between 16°C and 18°C during summer and between -6°C and -9°C during winter. It receives on average 1200 mm of precipitation annually, with about 960 mm as rainfall. The geology is composed mainly of sedimentary and some igneous rocks.

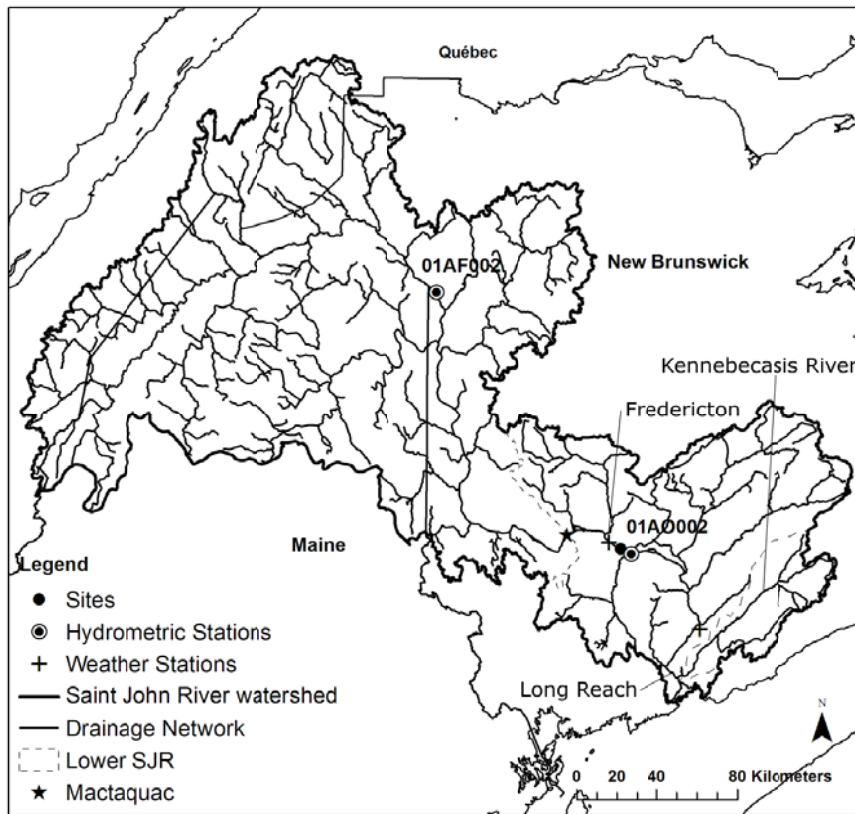


Figure 2.2.1 Saint John River watershed (Data used to create this map available from Geobase)

2.2.4.2 Instrumentation and data acquisition

The instrumentation included a turbidity meter with integrated wiper (Analite 395), a datalogger, a 12V battery and a solar panel installed approximately 12m away from the shore. The initial survey design included two stations, one upstream and one near the estuarine portion of the river. The downstream station was plagued with numerous problems, including excessive probe fouling and data had to be discarded. The remaining monitoring site (Site 1), located in Maugerville (45° 53' 18.8" N 66° 31' 35.4" W), had a cross section width of about 700 m and a maximum depth of about 14 m in the central channel. The area drained above Site 1 is approximately 42 668 km², but only 2850 km² is downstream from Mactaquac Dam. Turbidity readings were taken every

15 minutes. Time series of hourly and daily averages were subsequently used. Monitoring occurred from September 30th 2008 to November 27th 2008 and from May 22nd 2009 to November 26th, 2009.

In order to transform the turbidity values measured in nephelometric turbidity units (NTU) by the turbidity meter into SSC values in mg/L, site-specific NTU-SSC curves were developed by mixing local water and different volumes of local sediments in a plastic box. For each concentration, 3 or 4 turbidity readings were taken along with a 1L water sample. To obtain a sufficient number of points on the curves while keeping the number of filtrations necessary to a manageable number, a replicate water sample was taken for only about a quarter of the concentrations. The procedure was first done in early July 2009, and was repeated in late August 2009 as it was found that the July calibration points did not cover the full range of measured turbidity. Water bottles were kept refrigerated and were filtrated within one week. Water samples were filtered using pre-weighed 0.45µm Millipore mixed cellulose filters. Whenever possible, the whole 1-litre sample was filtered. When it was not possible due to filter clogging, a subsample (>250mL) was used and the volume noted. The filters with the suspended sediments were then dried at 65°C for at least 12 hours, and weighed again. SSC was determined by dividing the weight of sediments on the filter by the volume of water filtered to obtain values in mg/L.

In an attempt to estimate the river cross-section variability as suggested in USGS standard methods (Wagner *et al.* 2006), turbidity measurements were taken along a cross-section encompassing the usual probe location, and using the same probe. Measurements were taken from a small flat-bottomed boat equipped with a 15hp outboard motor and a sampling winch with a 10kg lead weight. The boat was kept at approximately the same position using the motor and vertical profiles were measured as equidistant as possible, using the three-points method (WMO 1980), in which values are obtained at 0.2, 0.6 and 0.8 of total depth. Two surface measurements located on each side of the channel were also obtained in May 2009 using grab samples, with 2 replicates each, taken from a bridge approximately 10km upriver. Finally, 5 verticals along a cross section in Fredericton were sampled (0.2 and 0.8 of total depth) in October 2010 using a Kemmerer water sampler.

The measurement location is within approximately 2km of a Water Survey of Canada (WSC) level recording station (Saint John River at Maugerville, 01AO002) which provided quasi complete hourly data during the monitoring periods. In this dataset, gaps smaller than 3 hours were filled using a linear interpolation and an elliptic filter followed by zero-phase digital filtering was used to remove the effect of tides (Matlab software, the Mathworks Inc.). In addition, one gauging station, located upstream at Grand Falls (01AF002, 21900km²) provided daily flows. The ratio area method, where specific discharge (m³/s/km²) is multiplied by the drainage area at the sampling station (Murdock and Gulliver 1993), was used to estimate flows at the turbidity monitoring station. While this monitoring station is located downstream from a hydro-electric dam, the latter is of the run-of-the-river type and a hydrograph comparison showed that there is generally no significant impact from the Mactaquac dam on levels or flows. However, the Mactaquac Generating Station is managed to meet fluctuating regional energy requirements, and in this regard, there are periods within a day and within a season that flow is locally substantively altered and is not directly correlated with other more dominant flow-controlling factors such as precipitation and snow melt. Although these can explain some water flow increases in the absence of precipitation events, it was not possible to obtain data on daily water releases from the dam.

An autonomous tipping-bucket rain and temperature gauge (HOBO Rain Gauge) was also installed and programmed to store values every 15 minutes, but data were only obtained from July 12th to November 26th 2009. Meteorological data were obtained from the closest Environment Canada weather station, at the Fredericton Airport (approximately 1km away).

2.2.4.3 Statistical analysis of SSC and potential hydrometeorological predictors

Pearson correlations were used to quantify relationships between SSC and different hydrometeorological variables, presented with their abbreviations in Table 2.2.1.

Table 2.2.1 Names, abbreviations and units for the variables used in the statistical analyses

	Abbreviation	Unit
Mean suspended sediment concentration †	SSC	mg/l
Maximum suspended sediment concentration †	SSC _{max}	mg/l
Suspended sediment load, in tonnes	Ls	tonnes
Water level	H	m
Temperature	T	°C
Daily precipitation at Fredericton weather station	1P	mm
Cumulative precipitation from 3,6,12 previous days	3P,6P,12P	mm
Estimated mean daily discharge	Q	m ³ /s
Maximum rain intensity (from rain gauge at Site 2)	Max I	mm/h

†Mean hourly SSC or mean daily SSC, as appropriate

Suspended sediment loads were calculated from the mean SSC and discharge (Q) using the following equation:

$$Ls = 0.0864 \cdot Q \cdot SSC \quad (1)$$

where Ls is the load in tonnes, Q and SSC are discharge in m³/s and sediment concentration in mg/L for the day and 0.0864 is a constant used to convert the units into tonnes/day.

The variables investigated at hourly and daily time steps were different as they were dependant on available data. For the two time steps, both maximum and mean SSC were considered. Hourly precipitation was only available from the rain gauge installed for this study, and not from WSC weather stations. Antecedent conditions were characterized by the accumulated precipitation from the previous 3, 6 and 12 days. Cumulative precipitation from the previous 1 to 15 days was initially tested, but longer periods are not presented as they did not yield better results. Lags between sediment, water level and precipitation were also examined.

In addition to the correlation matrices, two methods were compared to identify the best explanatory variables: forward stepwise linear regression and a geostatistical method initially developed by Chokmani and Ouarda (2004). This method, described in greater detail by Guillemette *et al.* (2009) and Daigle *et al.* (2010) who used it to model monthly temperature maximums in streams, consists of estimating the desired value using an interpolation method, ordinary kriging, in a physiographical space built from the outputs of a canonical correlation analysis (CCA). Canonical correlation analysis builds linear combinations of two groups of variables; in the present case, SSC-derived variables

form one group (log SSC and max SSC) and hydrometeorological predictors form the other. Within each group, original variables are combined so that the first linear combination (called canonical variate) in one group is most correlated with the first linear combination of the second group. A second linear combination is then performed under that constraint of orthogonality with the first one, and so on. Each pair of sample canonical variates is uncorrelated with the previous $k-1$ sample canonical variate. The canonical correlation that must be maximized is defined as:

$$r_{U,V} = \frac{a' S_{12} b}{\sqrt{a' S_{11} a} \sqrt{b' S_{22} b}} \quad (2)$$

where U and V are the matrices of canonical variates. The canonical coefficients vectors a and b in equation (2) are respectively the coefficients of the linear combinations of the first and second group of variables (SSC and hydrometeorological, respectively). S_{12} , S_{11} and S_{22} are the sample variance covariance matrices of the groups. (Johnson and Wichern 2007).

Since matrix V is composed of linear combinations of the physiographic variables that are orthogonal to one another, the set $[V1, V2]$ provides a Cartesian coordinate plane in which SSC variables can be projected. Using this projection, interpolation techniques can be used to estimate the SSC variable for any hydrometeorological conditions included in the $[V1, V2]$ domain. As per Guillemette *et al.* (2009), ordinary kriging was subsequently used to interpolate the log SSC variable. A leave-one-out procedure was applied to obtain estimations of log SSC for each sample. Goodness of fit statistics (r^2 and RMSE) were subsequently calculated in order to compare this approach with the stepwise regression.

2.2.5 Results

2.2.5.1 Data Treatment and Analysis

Visual inspection of the raw turbidity time series identified a number of spurious spikes. A correction was applied to every unique 15min datum above 300NTU, if the previous and following data were very low, as long as no precipitation was observed on that day. These values were assumed to be local events caused by a fish or boat passage and were replaced by the average of the previous and next measurement. A local event here is defined as the re-suspension of sediment near the shoreline and most likely not

representative of rest of the cross section. Visual inspection of the data did not reveal a significant drift in minimum turbidity measurements as was noticed in other studies (e.g. Birgand et al, 2004).

The final dataset spanned from September 30th to November 23th 2008 and from May 22nd to November 25th 2009, with five incomplete days: October 23rd 2008, and July 9th, 11th, August 24th - 26th 2009. Data from September 12th to 28th, 2009 (17 days) were removed from the analysis since water levels were excessively low and it was suspected the probe was in and out of the water depending on the tide, winds and waves. In addition, the values from September 24th to 27th 2008 were excluded from the analysis. While a Grub test for outliers failed to reject them, the turbidity data were higher than those used in the calibration curve, forcing extrapolation to transform them into SSC. A frequency analysis performed on the maximum water level for November using 1965-2009 data revealed that flood event (3.716 m) had a return period of 9 years and was well above the mean (2 m) and upper quartile (2.5 m) for those dates.

Flow estimation

The daily flow values estimated from the ratio area method were compared to levels at the corresponding WSC hydrometric station. A power fit was successfully applied between Q and level data ($r^2=0.88$, Figure 2.2.2) using 44 years of daily data. While most peaks seem to coincide (Figure 2.2.3), a correlation analysis using Pearson correlation coefficients and applying lags to the data revealed that Q from one day before H was as strongly related to H as Q from the same day ($r = 0.81$, $p < 0.05$). This indicates that the method introduces a small lag in the Q data of at least a few hours because the transport time between the two monitoring stations is not taken into account. The ratio area method was still used to estimate discharge (Q) values as a potential explanatory variable for sediment variability and subsequently to estimate daily sediment loads, but keeping in mind the short lag explained above.

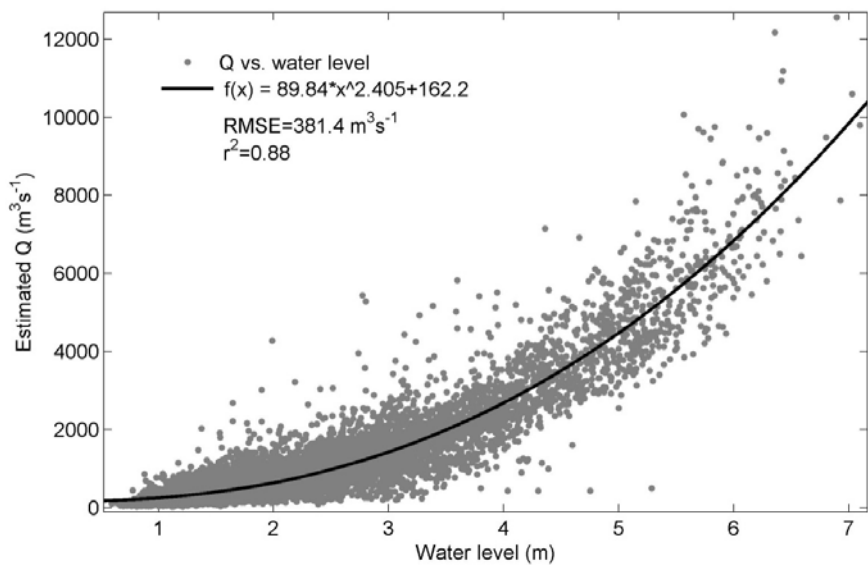


Figure 2.2.2 Flow vs water level at Site 1, 1965-2009

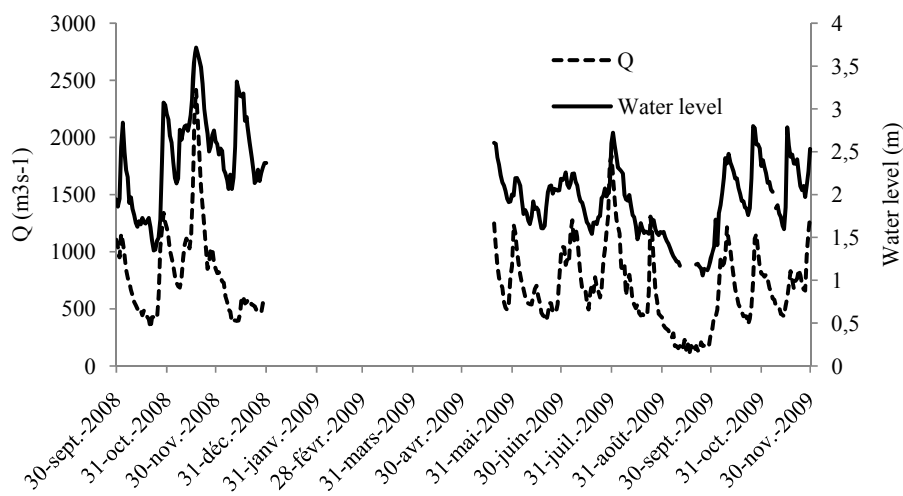


Figure 2.2.3 Time series of water levels at station 01AO002 and estimated flows.

Cross-section sampling

Turbidity measurements taken throughout the cross-section at Site 1 in August 2009 were uniform, but at the constant value of 0 NTU. Despite a large rain event 1.5 days prior to the measurements, the water was clear, even in the channel section and near the banks; hence variations in the SSC across the river could not be detected. Grab samples with replicates taken from two points near each shore approximately 10km upriver from site 1 during the spring freshet of 2009 yielded similar SSC values, all four of them between 25 and 30 mg/L. Finally, the 5 vertical profiles sampled also about 10km upriver from Site 1 in October 2010 were almost uniform, with a mean of 1.92mg/L. Only one point, taken near the shore and close to the bottom, had a value of 15 mg/L.

Transfer from turbidity to suspended sediment concentration

Even though sampling for the NTU-SSC curves was conducted on two different days, a Mann-Whitney test confirmed that both sets of data had similar distributions and could be assumed to be coming from the same population. They were therefore combined to produce a unique calibration curve. In total, 29 points were taken between 0.2 and 440 NTU, nine of them having replicates. Three (3) samples were rejected due to bad manipulations during filtration. On average, replicates differed by 14%, and three quarters of the sample pairs had SSC values that differ by 20% or less. When available, the replicates' average was used for the calibration curve. After screening for outliers using the 95% confidence interval when performing a linear regression one point was removed. The curves were forced through the origin (i.e. no initial bias due to water colour or any other factor was assumed and $0 \text{ NTU} \rightarrow 0 \text{ mg/L}$), as null values were observed in the field. The final regression is shown in Figure 2.2.4. The displayed equation was used to transform the turbidity data into SSC.

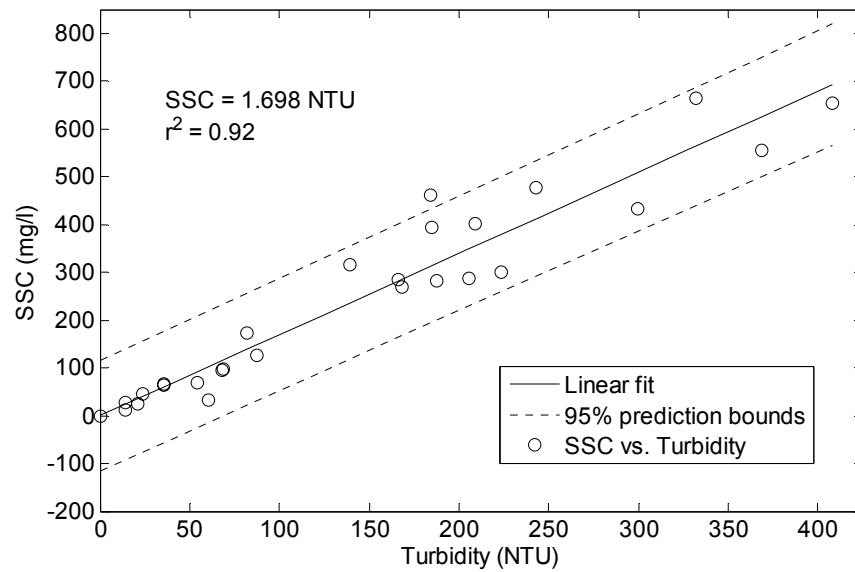


Figure 2.2.4 Linear regression between suspended sediment concentration (SSC) and turbidity

2.2.5.2 Statistical analyses

The basic descriptive statistics of the daily suspended sediment concentration (SSC) derived from turbidity measures are presented in Table 2.2.2.

Table 2.2.2 Basic statistics of the SSC dataset derived from turbidity measures

	Daily SSC* (mg/L)
Mean	17.14
Median	7.85
Maximum	178.20
Minimum	0.06
Standard deviation	29.20
Skewness	3.45

*excluding November 24rd to 27th 2008

Pearson correlation coefficients (r) for all variables at hourly interval are presented in Table 2.2.3. The only available hydrological variable available at this time step was water level (H). All shown values are significant with $p < 0.01$. When better results were obtained by applying a lag to the data, the appropriate lag time is shown in parentheses. The lag is expressed as negative when the variable of interest is correlated with variables from previous days and positive when looking at variables from following days. Mean and maximum SSC were strongly related ($r = 0.91$). Mean SSC is best explained by water level from 9 hours following the SSC value ($r = 0.28$), which is indicative of a hysteresis effect.

Table 2.2.3 Pearson correlation coefficients (r) among variables available at hourly intervals, with associated lag time, in hours and in parentheses, when different than 0

	SSC	SSC _{max}	H
SSC	1		
SSC _{max}	0.91	1	
H	0.28 (+9)	0.24(+5)	1

Note: All values shown are significant with $p < 0.05$.

Figure 2.2.5 shows daily precipitation, water level and SSC time series for the period of record. While some suspended sediment events are obviously related to precipitation and increases in water level (starting July 28th, 2009, for example), there is a lot of variability in the response of sediment concentrations to similar precipitation events or water level changes.

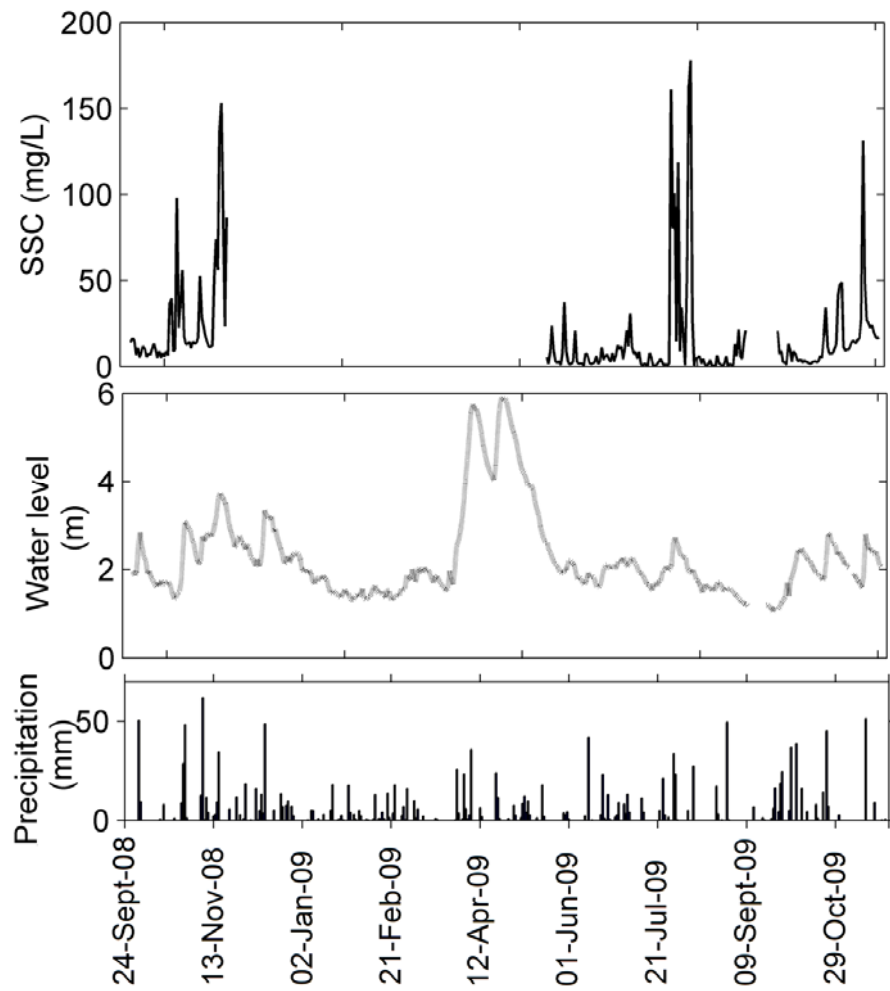


Figure 2.2.5 SSC, water level and precipitation at daily intervals

Table 2.2.4 shows the correlations between SSC and various hydrometeorological variables and Figure 2.2.6 shows the correlation results when applying lags between SSC and water level (H), discharge (Q) and precipitation (P). Correlation coefficients between mean daily SSC and H (Q) were 0.43(0.44) with a lag of 0 for H and -1 for Q. However, the relationship between SSC and Q at lag 0 was roughly the same as that of SSC and H ($r = 0.44$, Figure 2.2.6). Hence, it can be concluded that no hysteresis is detected at a daily time step. SSC was best correlated with P from the day before ($r = 0.27$) and with cumulative precipitation from the 12 previous days ($r = 0.27$). Generally, using the log of SSC did not yield much better results.

It was expected that Ls would be correlated to Q, as discharge is a variable in the equation used to calculate loads. Sediment loads were correlated to water level ($r = 0.55$) and discharge ($r = 0.64$) with 0 and -1 days, respectively. They were also correlated with total precipitation ($r = 0.22$) from the day before, as well as with cumulative precipitation from 3 and 12 previous days ($r = 0.12$ and 0.23).

Table 2.2.4 Pearson correlation coefficients (r) among variables available at daily intervals, with associated lag time, in days and in parentheses, when different than 0

Site 1	SSC	SSC _{max}	Log SSC	Ls	H	Q	1P	Max I	3P	6P	12P
SSC	1										
SSC _{max}	0.70	1									
Log SSC	0.72	0.62	1								
Ls	0.90	0.65	0.59	1							
H	0.43	0.27	0.5(+1)	0.55	1						
Q	0.44(-1)	0.41(-1)	0.36	0.64(-1)	0.81(0 & -1)	1					
1P	0.27(-1)		0.22(-1)	0.22(-1)	0.24(-2)	0.21(-3)	1				
Max I	0.22(-1)		0.18†	0.2(-1)		0.18(-2)	0.59	1			
3P	0.23		0.24	0.12†	0.2	0.31		0.19	1		
6P	0.17		0.20		0.43	0.40			0.68	1	
12P	0.27	0.21	0.32	0.23	0.56	0.48			0.51	0.68	
T	-0.15		-0.36	-0.29	-0.5	-0.14				-0.15	-0.23

Note: all values significant with $p < 0.05$ unless otherwise stated

† Significant with $p < 0.1$

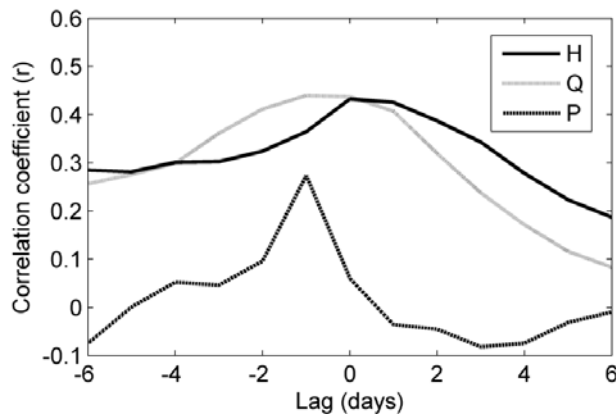


Figure 2.2.6 Correlation coefficients when applying lags between SSC and water level (H), Q and precipitation (P)

There was not enough data to calculate annual loads. However, monthly loads were computed. Correlations with precipitation and hydrological variables are shown in Table 2.2.5. The best explanatory precipitation variable was maximum 12 day cumulative precipitation ($r = 0.64$) and the best hydrological variable was maximum Q ($r = 0.85$).

Total estimated sediment loads for the period of record amounted to 1.94×10^5 tonnes, which can also be expressed as a yield of 68 tonnes/km² when spread over the area between the Mactaquac Dam and Site 1.

Table 2.2.5 Pearson correlation coefficients between total monthly loads and hydrometeorological variables.

<i>Variables</i>	Loads	
	r	p
Precipitation	0.33	0.38
Max daily P	0.69	0.04
Max 3P	0.50	0.17
Max 6P	0.43	0.26
Max 12P	<i>0.64</i>	0.07
Mean H	0.73	0.03
Max H	0.81	0.01
Mean Q	0.79	0.01
Max Q	0.85	0.00

Note: significant correlations with $p < 0.05$ are shown in **bold**
significant correlations with $p < 0.1$ are shown in *italic*

Water level (H), temperature (T) and precipitation from the previous day (P(-1)) were the identified predictors using a forward stepwise regression ($r^2 = 0.29$, RMSE = 0.51), for which the associated parameters are found in Table 2.2.6.

The canonical variates were constructed using log SSC and max SSC in the first matrix and H, Q, T, P(-1), 3P, 6P and 12P in the predictors matrix as they were found to be all significantly correlated with SSC in the first correlation analysis. Observation of the coefficients of the physiographic canonical variates revealed the most influential variables were H, T and 6P. Figure 2.2.7 is the developed kriging map. Cross-validation between observed and modeled log SSC shows slightly better results than the ones from the regression, with RMSE = 0.48 and $r^2 = 0.35$ (Figure 2.2.8).

Table 2.2.6 Parameters of the stepwise regression and of the canonical correlation analysis (CCA) developed with log SSC as a dependant variable

Parameter	Regression	CCA
Water level (H)	0.4660	0.2800
Discharge (Q)	-	0.0023
Temperature (T)	-0.0170	0.0522
Precipitation (P(t ₁))	0.0089	0.0095
3day precip	-	0.0059
6day precip	-	-0.0215
12day precip	-	0.0083
<i>statistics</i>		
intercept	0.05	-
RMSE	0.51	0.48
r^2	0.29	0.35

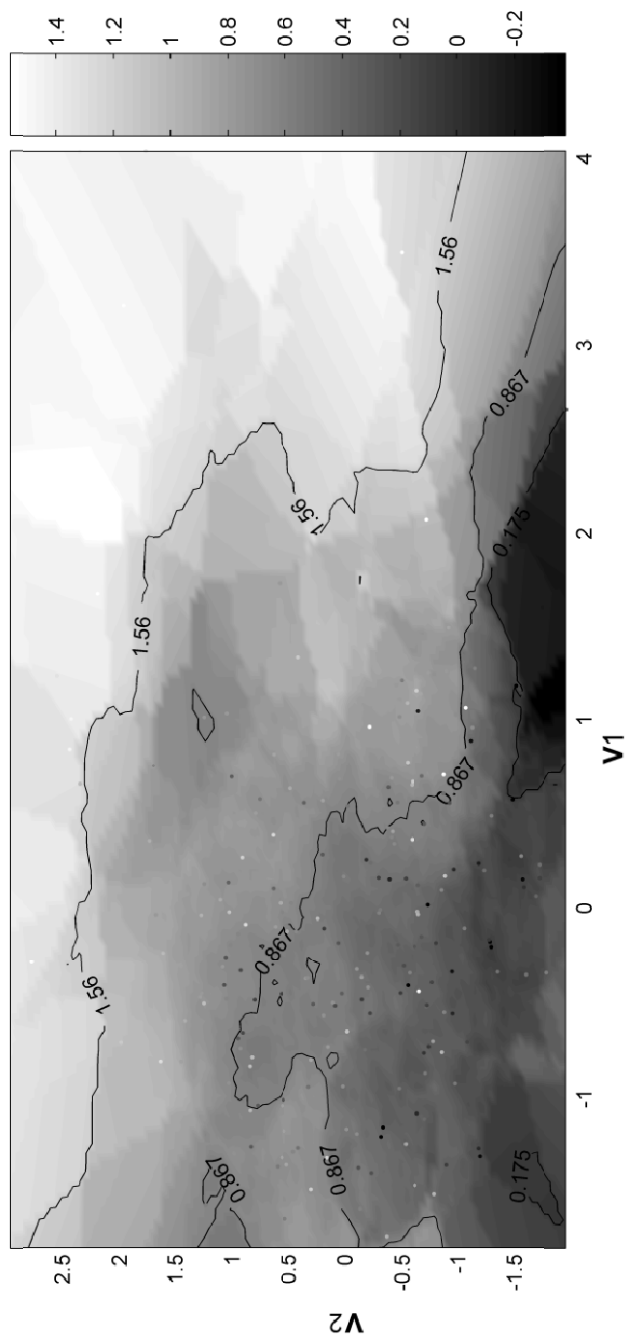


Figure 2.2.7 Interpolated log SSC using ordinary kriging in canonical Variate ($V1$, $V2$) space.

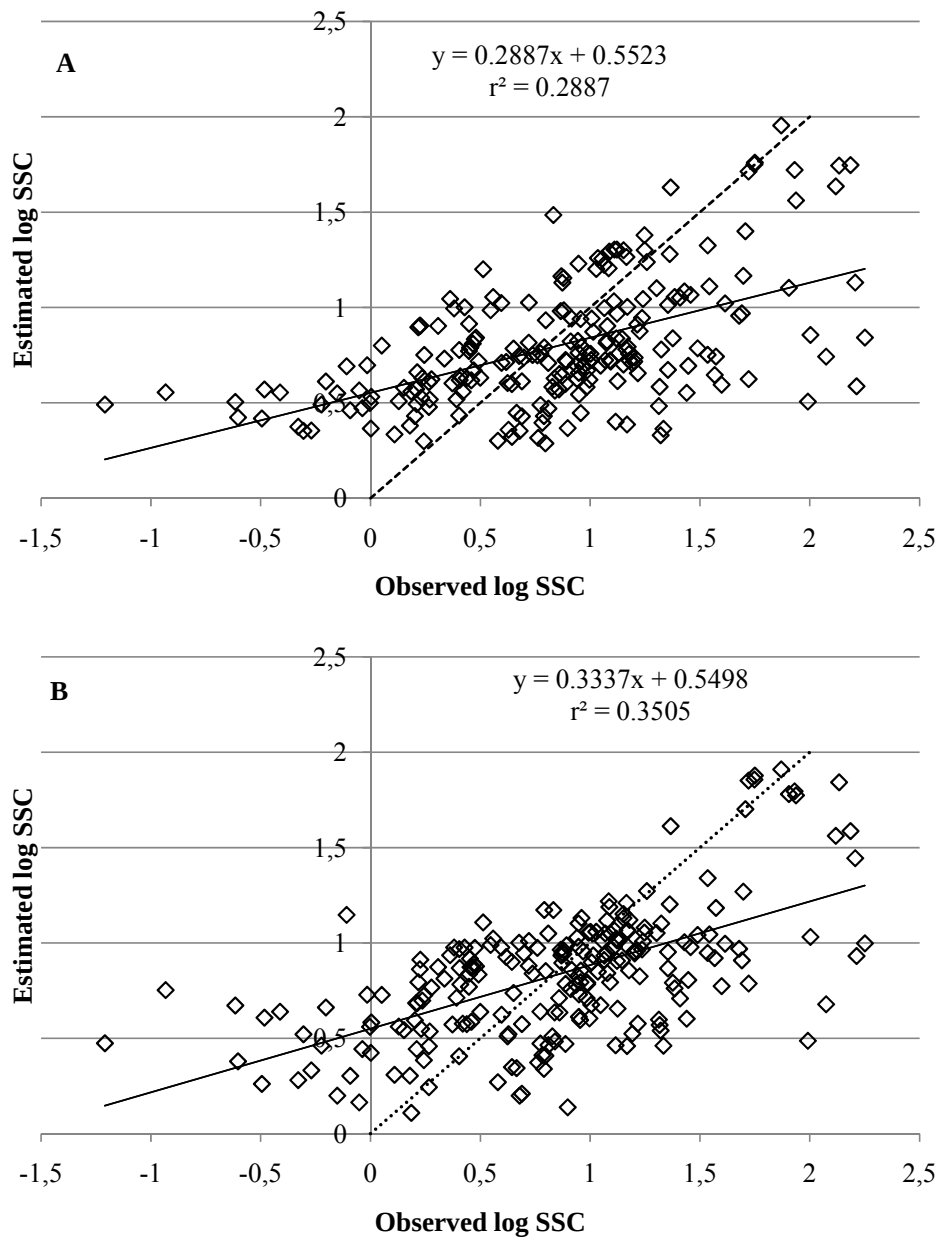


Figure 2.2.8 Estimated vs observed log SSC obtained from A) stepwise regression and B) ordinary kriging in a physiographical space built using CCA on 7 variables

2.2.6 Discussion

2.2.6.1 Data acquisition and treatment

A site-specific curve to transfer turbidity values (NTU) into suspended sediment concentration (mg/L) was successfully created for the period of record at the monitoring site. The uncertainty on the curve could be further reduced by increasing the number of samples filtered and by avoiding the filtration of sub-samples in the lab. Gray *et al.* (2000) found that filtering sub-samples yields values (typically referred to as total suspended solids, TSS) that were less accurate than when the entire sample was filtered due to the difficulty of adequately stirring the sample to ensure the homogeneity of sub-samples. Using an automated sampling system triggered by turbidity peaks (see Eads and Lewis 2003) could also allow deriving event-specific curves throughout the year, as well as obtaining water samples directly from the water column rather than from artificially created mixes. In addition, some error in the higher end of the curve can be caused by the high concentration itself; Birgand *et al.* (2004) noted that the calibration curve is linear up to a point after which some particles within the IR beam can be hidden by others, making it difficult to obtain a linear relationship. Marttila *et al.* (2010) also found, in a laboratory study looking at the influence of type, size and shape of particles on SSC-NTU relationships, that larger particles “were concealed under the influence of the smaller particles in the turbidity value, resulting in error at high concentrations.” This phenomenon could explain why it was technically difficult in the present study to obtain values for the calibration curve above 425 NTU.

Using fixed turbidity measuring devices in large rivers poses obvious challenges related mostly to potential SSC variability across the width and depth of the cross-section (Horowitz *et al.* 2001). In the Saint John River, the logistic challenges associated with probe installation made it impossible to install multiple probes on structures such as bridges to measure the spatial variability within the section. The measuring point was limited by the length of the instrument cable. However, Beltaos and Burrell (2000) found that, in a reach of the SJR between Dickey, Maine and St-Leonard, NB, variations within the cross-section can be found just downstream of tributaries having very different suspended sediment concentrations, but are typically dissipated quickly by mixing processes. Cross-section sampling in the present study also indicates sediments are not transported preferentially in a sub-area of the cross-section. However, the grab samples taken at different verticals indicate that locating the sensor near shore and close to the

bottom makes it more vulnerable to local variations and to resuspension. Hence it is believed the observed turbidity values and derived SSC can be considered representative of the river cross-section when aggregated over daily or longer time steps, and that high-frequency data series must be carefully examined to detect obvious short-duration, local SSC peaks.

Other drawbacks of locating the probes near shore include possible vegetation growth close to the instruments, biofouling by algae, the danger of flooding the datalogger on shore, potential sensor exposure to air during dry spells and the impossibility of obtaining winter and freshet data, as the equipment would almost certainly be washed away by ice and high flows.

Estimated daily discharge (Q) values from the ratio area method were used for estimating suspended sediment loads for the period of record, as the effect of the small lag between estimated Q and real water level is not as important when loads are summed over monthly periods. Yet level data were preferentially used as an explanatory variable since they were available at an hourly time step and more reliable than the estimated Q .

The usefulness of high-frequency turbidity data in rivers seems related to the availability of other concomitant data that are expected to be analyzed along with it. In the absence of other variables taken at high-frequency, such as 15 minutes or hourly intervals, one is forced to aggregate the data into daily time steps, which might hide some processes occurring over a few hours. For instance, in the present analysis, best correlations were found between SSC and water level data lagged by 9 hours for hourly data, but the best correlation values between the same variables had no lag when using daily data. In addition, rain intensity data were only available for a short portion of the period of study. Consequently, it can be questioned whether collection of high-frequency data is merited in study sites where no other reliable data are taken at short intervals. In addition, it is possible that high-frequency data in large basins reflects mostly “noise” or quick, local variations in conditions that are only representative when aggregated on longer time frames. On the other hand, having instruments deployed facilitates obtaining data for every major sediment transport event. In addition, averaging 96 measurements to obtain one daily SSC value makes that daily value more reliable than one grab sample taken at any time of the day, as it could well happen that this grab sample is taken during a local,

short-duration SSC peak. Such events can often be identified in series of data taken at 15 minutes time step and on that aspect, the USGS has documented some of the most commonly observed odd values in turbidity series, caused by nonsubmergence of the sensor, interference, bedload burial, fouling, debris, etc. (Lewis and Eads 2008).

2.2.6.2 Statistical analysis

The strongest relationship was detected between water level and mean SSC, with a lag of +9 hours translating into no lag at the daily interval. While this indicates that positive (clockwise) hysteresis is present at this site, where SSC peak precedes an increase in discharge (here substituted by water level), this process is concealed when using daily data. Clockwise hysteresis has been frequently observed in different watersheds (Asselman 1999; Oeurng *et al.* 2010) and can indicate that the sources of sediments were easily available, such as local sediments transported by nearby drainage and resuspension within the river (Gao 2008; Oeurng *et al.* 2010). Other clear relationships between suspended sediment and hydrometeorological variables were hard to detect, due to the highly variable nature of the data, even though coinciding peaks between SSC, water level and precipitation are obvious from a visual inspection of the time series

It is also possible that a different sediment regime exist at high and low flows, even though it is not possible to draw such a conclusion, given the short monitoring period. The high event of November 24th-27th, which was the highest flood and the highest SSC event in the whole period of record, indicates it might be the case. Including it in the correlation analysis biased all results, since the values were much higher than the remaining of the monitoring period and the lag between SSC and water level was different. . Other such events in future monitoring periods might occur, allowing more information to be derived from them.

Relatively weak correlations coefficients between SSC and precipitation could be caused by the use of only one weather station as opposed to a spatial average from many stations over the whole watershed, as used by some authors (e.g. Lopez-Tarazon *et al.* 2010; Oeurng *et al.* 2010). It must also be kept in mind that rainfall intensity data were only available from July 12th to November 27th 2009, and at a location approximately 70km away from the site. The absence of correlation between precipitation and sediment has previously been interpreted as the response of sediments being linked to some

other processes such as sediment availability (Lopez-Tarazon *et al.* 2010). While precipitation typically makes sediments more available, exhaustion can sometimes occur after a series of precipitation events. This can explain the absence of SSC peaks following some major precipitation events, such as in June 2009. However, in the present study, the fact that both levels and sediments (H and SSC) are correlated to precipitation with a lag of -1 day, albeit weakly, indicate precipitation drives at least some of the SSC variation in this system, and that the response to erosion events might not be instantaneous. Cumulative precipitation from previous days was also found to be correlated to water levels, sediment concentrations and loads, which again indicates that the soil moisture condition prior to any rain event has an impact on the erosion, runoff and transport of sediment to the river.

Both the stepwise regression and the geostatistical approach identified water level, temperature and either the precipitation from the day before or 6 days cumulative P as the main explanatory variables for log SSC. Yet, only 35% of the variance can be explained by even the better of the two models. This indicates: 1) non-linearities exist between the variables that should be handled by other means than simply taking the logarithm of data; and 2) other variables than those examined here could play a role in the sediment transport processes of the Saint John River.

It should be remembered that the Mactaquac Dam Generating Station has some impact on flows, especially during the hottest, driest part of the summer when natural flows are especially low. Then, water can be stored and released quickly to meet the energy demand for air conditioning systems, creating a rise in water level that is not caused by precipitation. Unfortunately, no data about water releases during the period of study were available.

The collected data allowed the calculation of sediment loads. Specific yield for the monitoring period during 2009 (68 tonnes/km²) was above the range of observed annual yields on the downstream comparative tributary, the Kennebecasis River, between 1967 and 1988 (between 13 and 63 tonnes/km²/year, Water Survey Canada), even though the whole year was not sampled. This is a surprising result, given the potential trapping effect of Mactaquac Dam. Possible explanations include that: 1) the loads are overestimated due to unreliable Q data; 2) yields are higher on the Saint John River at Maugerville than on the Kennebecasis because of the presence of Fredericton some

10km upriver, a city with large impervious areas that cannot trap sediments; 3) some sediments come from upstream of the dam or resuspension is more frequent in the lower reaches of the Saint John than in the Kennebecasis. At any rate, the underlying cause is likely not the proportion of agricultural land use, which is higher in the Kennebecasis sub-basin (16%) than on the Saint John (4%).

Overall, the present analysis indicates that a more complex model, with additional spatial (i.e. more stations) and temporal (i.e. longer time series) data, would be necessary to investigate the inter-annual variability of sediment transport within the large Saint John River watershed. Many different processes have been found to play an important role in sediment transport in river systems. These can include sediment availability, anthropogenic impacts such as agriculture, forestry, mining, and major construction sites which are all recognized as affecting sediment supply (Meade *et al.* 1990). Specific characteristics of land use such as the timing of agricultural activities (Lecce *et al.* 2006) also have an impact on the susceptibility of soils to erosion. Unfortunately, the data needed to include these effects is typically hard to access, sometimes difficult to express in a quantitative manner (Meade *et al.* 1990), and, for some of it, excessively variable in large watersheds. In fact, much of the difficulties in the present study were due to the lack of data and the need to revert to estimated proxies, such as the estimated discharge (Q). Yet explanatory studies such as this one are needed to help identifying what additional data would be most useful to obtain.

2.2.7 Conclusions

High-frequency turbidity data were obtained, for the fall of 2008 and summer /fall of 2009, at two sites on the Saint John River, New Brunswick. It was only possible to derive suspended sediment concentration (SSC) data, using a site-specific regression, for one of them.

Relationships between SSC and different hydrometeorological variables were investigated at two temporal scales (hourly and daily) using Pearson correlation coefficients. Strongest correlations were detected with precipitation from the previous day and with water level from the following 9 hours or the same day. A geostatistical approach combined with multivariate analysis yielded slightly better results for estimation of log SSC than a forward stepwise regression. Yet much of the variation remains unexplained, which may be attributed, in part, to lack of reliable and representative

hydrometeorological data and in part to other, potentially anthropogenic, processes at play.

Hourly or 15-minutes interval data have the potential to help understanding short-term processes driving sediment transport in rivers. The erratic nature of turbidity may make it difficult to identify proper predictors at a short time scale in large systems such as the Saint John River. Other hydrologic or meteorological variables are seldom available at the same time scale, forcing the averaging of the high-frequency SSC data into hourly or daily intervals, these averages are more reliable than one grab sample assumed to be representative of the whole day. Some recommendations for further work include the use of an automated sampling system for developing the site-specific regressions, as well as ensuring that the instruments are located at a representative point in the river cross-section.

This research shows the potential of turbidity measurements to improve the currently poor knowledge of sediment transport dynamics on the Saint John River, though a longer monitoring period for all variables including the spring freshet and additional land-use related data might be required to address the issue of sediment load inter-annual variability.

2.2.8 Acknowledgements

The authors wish to acknowledge the help of many people regarding the collection of field data, as well as Environment Canada for providing weather data. This research was funded in part by NSERC and the Saint John Port Authority.

2.2.9 References

- Asselman, N. E. M. 1999. Suspended sediment dynamics in a large drainage basin: the River Rhine. *Hydrological Processes*, 13(10):1437-1450.
- Beltaos, S. and Burrell, B. C. 2000. Suspended sediment concentrations in the Saint John River during ice breakup. *In* 2000 Annual Conference Abstracts - Canadian Society for Civil Engineering, p 75.
- Birgand, F., Lefrançois, J., Grimaldi, C., Novince, E., Gilliet, N. and C., G.-O. 2004. Mesure des flux et échantillonnage des matières en suspension sur de petits cours d'eau. *Ingénieries*, 40:21-35.
- Chokmani, K., and T. B. M. J. Ouarda. 2004. Physiographical space-based kriging for regional flood frequency estimation at ungauged sites. *Water Resour. Res.*, 40(12):1-13. doi:10.1029/2003WR002983 .
- Daigle, A. St-Hilaire, A. D. Peters, D. Baird. 2010. Multivariate water temperature modeling in a semi-arid watershed. *Canadian Water Resources Journal* 35(3) : 237-258.
- Dodds, W. K. and Whiles, M. R. 2004. Quality and quantity of suspended particles in rivers: Continent-scale patterns in the United States. *Environmental Management*, 33(3):355-367.
- Eads, R. and Lewis, J. 2003. *Turbidity Threshold Sampling in Watershed Research*. Arcata, CA, US Forest Services.
- Gao, P. 2008. Understanding watershed suspended sediment transport. *Progress in Physical Geography*, 32(3):243-263.
- Gippel, C. J. 1995. Potential of turbidity monitoring for measuring the transport of suspended solids in streams. *Hydrological Processes*, 9(1):83-97.
- Gray, J. R., Glysson, G. D., Turcios, L. M. and Schwarz, G. E. 2000. Comparability of suspended-sediment concentration and total suspended solids data. US Geological Survey, Water-Resources Investigations Report 00-4191, Reston, Virginia.14p.
- Guillemette, N., St-Hilaire, A., Ouarda, T. B. M. J., Bergeron, N., Robichaud, E. and Bilodeau, L. 2009. Feasibility study of a geostatistical modelling of monthly maximum stream temperatures in a multivariate space. *Journal of Hydrology*, 364(1-2):1-12.
- Horowitz, A. J. 2003. An evaluation of sediment rating curves for estimating suspended sediment concentrations for subsequent flux calculations. *Hydrological Processes*, 17(17):3387-3409.
- Horowitz, A. J., Elrick, K. A. and Smith, J. J. 2001. Estimating suspended sediment and trace element fluxes in large river basins: methodological considerations as applied to the NASQAN programme. *Hydrological Processes*, 15(7):1107-1132.
- Johnson RA, Wichern DW. 2007. *Applied multivariate statistical analysis*. Pearson Education, Inc: Upper Saddle River, New-Jersey.
- Knighton, D. 1998. *Fluvial forms and processes: A new perspective*. Arnold, London, 382p.

- Lecce, S. A., Pease, P. P., Gares, P. A. and Wang, J. Y. 2006. Seasonal controls on sediment delivery in a small coastal plain watershed, North Carolina, USA. *Geomorphology*, 73(3-4):246-260.
- Lewis, J. and Eads, R. 2008. Implementation guide for turbidity threshold sampling: principles, procedures, and analysis. Gen. Tech. Rep. PSW-GTR-212. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station. 86 p.
- Leys, V. 2007 3D Flow and Sediment Transport Modelling at the Reversing Falls - Saint John Harbour, New Brunswick. in *OCEANS 2007*, Vancouver, BC.
- López-Tarazón, J. A., Batalla, R. J., Vericat, D. and Balasch, J. C. 2010. Rainfall, runoff and sediment transport relations in a mesoscale mountainous catchment: The River Isabena (Ebro basin). *Catena*, *in press*. doi :10.1016/j.catena.2010.04.005
- Mano, V., Moatar, F., Coynel, A., Etcheber, H., Ludwig, W., Meybeck, M., Nemery, J., Poirel, A., Blanc, G. and Schafer, J. 2006. Space and time variability of suspended particulate matter (SPM) transport in 32 French rivers (100 to 100 000 km²; daily to yearly). *Sediment Dynamics and Hydromorphology of Fluvial Systems*, ICCE IAHS International Symposium, 3-7 July 2006, Dundee, Scotland, (Poster Report Booklet):29-37.
- Marttila, H., Postila, H. and Klove, B. 2010. Calibration of turbidity meter and acoustic doppler velocimetry (Triton-ADV) for sediment types present in drained peatland headwaters: focus on particulate organic peat. *River Research and Applications*, 26:1019-1035.
- Meade, R. H., Yuzyk, T. R. and Day, T. J. 1990. Movement and storage of sediment in rivers of the United States and Canada. *in Surface Water Hydrology*, Geological Society of America, Boulder, Colorado, USA, Vol O-1, pp 255-280.
- Meybeck, M., Laroche, L., Durr, H. H. and Syvitski, J. P. M. 2003. Global variability of daily total suspended solids and their fluxes in rivers. *Global and Planetary Change*, 39(1-2):65-93.
- Murdock, R. U. and Gulliver, J. S. 1993. Prediction of river discharge at ungauged sites with analysis of uncertainty. *Journal of Water Resources Planning & Management - ASCE*, 119(4):473-487.
- Neu, H. A. 1960. Hydrographic of Saint John Harbour, NB. Mechanical Engineering Report MH-97, Ottawa, On. National Research Council of Canada, 264 p.
- Oeurng, C., Sauvage, S. and Sanchez-Perez, J.-M. 2010. Dynamics of suspended sediment transport and yield in a large agricultural catchment, southwest France. *Earth Surface Processes and Landforms*, *in press*.
- Pavanelli, D. and Bigi, A. 2005. Indirect Methods to Estimate Suspended Sediment Concentration: Reliability and Relationship of Turbidity and Settleable Solids. *Biosystems Engineering*, 90(1):75-83.
- Pavey, B., Saint-Hilaire, A., Courtenay, S., Ouarda, T. and Bobee, B. 2007. Exploratory study of suspended sediment concentrations downstream of harvested peat bogs. *Environmental Monitoring and Assessment*, 135(1-3):369-382.

- Pfannkuche, J. and Schmidt, A. 2003. Determination of suspended particulate matter concentration from turbidity measurements: particle size effects and calibration procedures. *Hydrological Processes*, 17(10):1951-1963.
- Trites, R. W. 1960. An oceanographical and biological reconnaissance of Kennebecasis Bay and Saint John River estuary. *Journal of the Fisheries Research Board of Canada*, 17(3):377-408.
- Wagner, R. J., Boulger Jr., R. W., Oblinger, C. J. and Smith, B. A. 2006. *Guidelines and Standard Procedures for Continuous Water-quality Monitors: Station Operation, Record Computation, and Data Reporting*, (USGS).
- WMO 1980. Manual on stream gauging - Field Work. World Meteorological Organization, Geneva, Switzerland p 308.

2.3 Linking Sediments Dredged from the Saint John Harbour, New Brunswick, and Hydrometeorological Variables

2.3.1 Résumé

Les volumes de sédiments dragués dans l'estuaire du fleuve Saint-Jean, où se trouve le port de Saint John, sont très variables d'une année à l'autre. Puisqu'il n'a pas été possible jusqu'à présent d'en prédire la variation à l'aide de modèles de transport de sédiments provenant de la Baie de Fundy ou du fleuve Saint Jean, l'utilisation de variables hydrométéorologiques locales pour expliquer cette variation a été tentée. Des corrélations de Pearson ont d'abord été développées entre les volumes de sédiments dragués et d'autres variables annuelles pour les années hydrologiques 2003-2004 à 2008-2009: le niveau d'eau moyen, le niveau d'eau maximum, le débit moyen, le débit maximum, le volume d'eau total annuel, la pluie, la neige et la précipitation totale. La période de récurrence du volume d'eau total annuel (débits cumulés), calculée à l'aide d'une analyse fréquentielle, a également été utilisée, ainsi qu'une estimation des charges sédimentaires provenant de la rivière Kennebecasis. Ces dernières ont été obtenues à l'aide d'un réseau de neurones artificiels (RNA) développé pour ce tributaire du fleuve. Pour pallier au petit nombre de données (6), la méthode jackknife a été utilisée pour estimer l'erreur du modèle. Bien qu'il soit prématuré d'utiliser cette régression comme telle pour le moment, les résultats sont prometteurs ($r^2 = 0,83$) et suggèrent que cette analyse devrait être approfondie en allongeant la série de données disponibles.

2.3.2 Introduction

Dredging is required annually in the Saint John Harbour, NB. However, the amounts dredged are variable from one year to the next, without the Saint John Port Authorities (SJPA) being able to detect why and predict dredging needs accordingly. Preliminary studies of suspended sediments on the main SJR branch, using turbidity as a surrogate, were undertaken during the fall 2008 and from May 22nd to November 27th 2009. It was not possible to obtain winter or freshet measurements during these years of monitoring. The lack of long-term sediment data on the SJR complicates the study of inter-annual suspended sediments variation. Fortunately, such long-term sediment data exist on the Kennebecasis River (KR), a major tributary located near the mouth of the SJR. Results from investigations on that river could give an insight as to processes also occurring at the whole SJR watershed scale.

In the 1990's, Bray and Xie (1993) defined regressions to estimate long-term and inter-annual yields in ungauged Atlantic watersheds, based on physical characteristics of the regions. However, inter-annual variation was solely dependent on mean annual discharge. Long-term sediment record from the KR was also used to estimate sediment loads and identify hydro-meteorological variables that can explain SSC variability with rating curves (Hansen and Bray 1993) and an artificial neural network (ANN; Higgins et al. 2011). Results of the latter showed that while the developed ANN is not able to adequately represent SSC peaks, annual loads are satisfactorily estimated. In addition, it was suggested that modelling sediment yields might be feasible in the SJR using hydro-meteorological variables, although the task would be complicated by anthropogenic and ice effects.

The above-mentioned results have not been used yet to try understanding the inter-annual variability of the sediment loads, or extrapolating the results to the whole lower SJR. These steps would be necessary in order to investigate the relationship between SJR sediment loads and dredging volumes from the port.

The present study aims at investigating relationships between dredging in the harbour, hydro-meteorological variables, and sediment loads from the Kennebecasis River.

2.3.3 Data & Methods

2.3.3.1 Estimation of sediment volumes and loads

Estimates of sediment volumes dredged from the harbour every year were obtained from the Saint John Port Authorities for the summers 2004 to 2009. Estimates of sediment loads for the main branch of the Saint John River were not available; however, the artificial neural network (ANN) developed for the Kennebecasis River was used to verify whether loads coming from that tributary are related to dredging in the harbour. A value for the year 2008-2009 was not available for lack of associated Q data on the Kennebecasis.

2.3.3.2 Hydro-meteorological data

Water levels were obtained from station 01AP005 in Saint John. Water discharge (Q) is not measured at this station by the Water Survey of Canada (WSC) hence the ratio area method was used with station 01AF002 in Grand Falls with the following equation:

$$Q_{01AP005} = 2.51Q_{01AF002} \quad (1)$$

Where 2.51 is the ration of the SJR watershed drained at Saint John over the area drained at station 01AF002. A visual inspection of both Q and water level data sets confirms peaks and lows coincide, and are correlated with a r^2 of 0.74 ($p < 0.0001$). Due to technical problems with WSC monitoring stations, no flow data was available at the Grand Falls station for the first 4 months of 2009. The Q values were estimated for that period using an exponential fit with 80 years of daily levels from the station at Fredericton (01AK003; $r^2 = 0.83$, $p < 0.01$). Since the month of April (typically part of the snowmelt) was included in this period, and because extreme values tend to have the greatest errors associated to them, the Q values from 2009 will be used with caution.

Precipitation data were obtained from Environment Canada Weather Office for Saint John (station 8104900).

2.3.3.3 Analysis

Correlation analysis was used to detect relationships among variables. A multiple linear regression was subsequently built, using the variables with the strongest correlation coefficients. Due to the small number of years in the data set, a jackknife procedure was used to estimate the error of the model. With this technique, the regression model is

repeatedly built leaving one different data point out every time. The value of the left out point is then estimated each time, and the error and bias are estimated from these estimated values. It allows a better estimation of the error when only very few data points are available in a regression.

To help assessing whether years with extreme annual water yields could be linked with higher dredged volumes, a frequency analysis was performed on that hydrological variable. In summary, a frequency analysis allows characterizing data in a probabilistic context by estimating their return periods (average number of years between two extremes of same amplitude). This is a widely-used method described in many textbooks (e.g. Rao and Hamed 2001). It allows relating the magnitude of events to their frequency of occurrence through the use of probability distributions.

The HYFRAN software⁴ was used along with 78 years of annual water yields (in million cubic meters per year) from the monitoring station at Grand Falls (01AF002) to verify whether wet years caused high dredged volumes. A more detailed application of the HYFRAN software can be found in Trambly et al. (2008).

2.3.4 Data

It can be hypothesized that sediments dredged from the harbour in any one year accumulated due to transport occurring since the last dredging. Because dredging typically occurs from the month of July to November, using hydro-meteorological data from the hydrological year (Oct-Sept) seemed more appropriate than the calendar year. This implies that sediments flowing from the river to the harbour during the dredging season, which coincides approximately with the end of the hydrological year, is taken into account in that same season. This seems realistic, since dredging operations are carried out in different phases over one summer. A first survey of the bathymetry is made approximately in June to estimate the dredging needs; then dredging operations are undertaken, followed by another bathymetry survey, then more dredging as needed, and finally a post-dredging survey. Sediments can thus accumulate between the first and second dredging. Hence the data used in the analysis is presented in that format (for example, from October 2003 to September 2004), and the expression “annual” used in this report refers to a hydrological year.

⁴ http://www.ete.inrs.ca/activites/groupe/chaire_hydrol/hyfran.html

Table 2.3.1 Sediment and hydro-meteorological data used in the analysis

year	Sediment		Hydro-meteo Data								T _Q
	ANN	Dredged volumes	Mean Level	Max Level	Qmean	Qmax	Qtot	Rain	Snow	Precip	
	(T)	(m ³)	(m)	(m)	(m ³ s ⁻¹)	(m ³ s ⁻¹)	(m ³ x10 ⁶)	(mm)	(cm)	(mm)	
2003-04	53298	146065	1.654	2.919	1398	7007	41399	941.5	204.7	1121.7	1.43
2004-05	72393	165660	1.481	4.738	1024	11377	37805	989.6	206.5	1157.5	5
2005-06	59544	375880	1.966	3.280	1667	7183	47133	1263.8	100.7	1359.7	50
2006-07	39999	116225	1.582	3.298	1226	8062	35576	1086.2	165.2	1249.6	1.3
2007-08	54633	271958	1.641	5.076	1350	16199	48500	1136.5	295.6	1432.1	75
2008-09	N/A	207384	1.776	4.080	1001*	9694*	42578*	1447	296	1748	10

*Contains an estimation of the values for Jan-April 2009, as described in section 2.3.3.2

2.3.5 Results

The correlation matrix (Pearson r^2) between all variables is presented in Table 2.3.2. The best relationships occurred when using data belonging to the same hydrological year, with no lag; for example, the total annual water yield and sediment volumes both from 2004-2005. The volumes of sediments dredged from the port are correlated with mean water level ($r^2=0.63$) and total water yield ($r^2=0.73$). No relationship was detected between dredged volumes in the harbour and annual sediment load from the KR estimated with the ANN, either from the same year or the year before. A multiple linear regression was created using mean level and total water yield. The resulting parameters and statistics are found in Table 2.3.3.

Table 2.3.2 Coefficients of determination (r^2) among all dredging volumes and hydro-meteorological variables

	Dredge volumes	ANN	Mean Level	Max Level	Qmean	Qmax	Qtot	Rain	Snow	Precipitation	T _Q
Dredge vol.	1.00										
ANN	0.08	1.00									
Mean Level	0.63	0.00	1.00								
Max Level	0.02	0.24	0.15	1.00							
Qmean	0.42	0.05	0.40	0.21	1.00						
Qmax	0.02	0.06	0.13	0.87	0.07	1.00					
Qtot	0.73	0.03	0.41	0.06	0.33	0.15	1.00				
Rain	0.24	0.01	0.45	0.01	0.01	0.00	0.16	1.00			
Snow	0.05	0.00	0.11	0.41	0.37	0.47	0.03	0.05	1.00		
Precipitation	0.12	0.03	0.23	0.08	0.07	0.06	0.17	0.89	0.29	1.00	
T _Q	0.64	0.01	0.17	0.20	0.29	0.37	0.78	0.06	0.02	0.08	1.00

Note: Bold values are significant with $p < 0.1$

Table 2.3.3 Coefficients and statistics of the regression

Parameter	Value
Intercept	-654674
Mean level	237279
Qtot	11.12
Statistic	
RMSE	77309 T
RRMSE	0.34
Bias	18620 T
R square	0.83
F statistic	7.43
P value	0.07

Due to the small data set, the regression could not be tested with values not used to create the model, but the jackknife technique used allows a better estimate of the error (RMSE) and bias. Annual estimates of dredging volumes obtained from the jackknife procedure are compared to the values provided by the SJPA in Figure 1. The model successfully predict an appropriate range of values encompassing the observed one, although the error (RMSE=77309 T) can be large compared to some of the values estimated (i.e. 2006-2007). Overall, the error is acceptable with a relative root-mean-

squared error (RRMSE) of 34%. The regression, using all years, yield estimates deviating on average 15.4% from the real value.

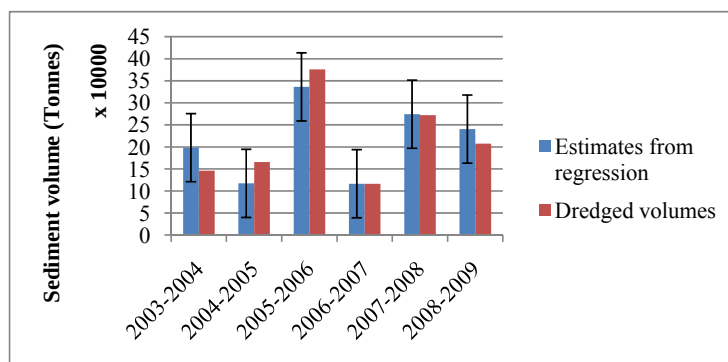


Figure 2.3.1 Comparison of dredging volumes values provided by SJPA and obtained by the regression. Error bars are $\pm$ RMSE

2.3.6 Discussion

The regression presented here was built from very few data points and should hence be applied to other years with caution. Longer-term dredging data in the harbour would be necessary to verify its accuracy. However, the strong relationships detected between the dredging in the harbour and hydro-meteorological variables from the Saint John River (SJR) watershed, especially total water yield and mean water level, are encouraging. In addition, a strong correlation was also detected between the dredged volumes and rain for the years 2004-2008. For some reason, this relationship significantly decreased in 2009, but should definitely be kept in mind for future analysis.

It is likely that a simple model could be built, with additional data, to allow an approximate prediction of dredging requirements. However, long-term prediction will not be possible as long as the model requires hydro-meteorological data from the year of interest. In addition, results will always be dependent on the availability of input hydro-meteorological data, mainly from Environment Canada. Years with missing or unreliable data can make an estimate difficult.

The lack of relationship between the estimated sediment loads in the Kennebecasis River (KR) and dredged sediments from the harbour could be related to the fact that loads were calculated for a hydrological station located in the middle portion of the

watershed. This upstream location, with a drainage area of 1100 km², might be too small to represent accumulated effects at the SJR scale. In addition, it is known that the morphology of the Kennebecasis Bay, which is much deeper than the SJR itself and experiences little freshwater flux (Haigh and Clarke, Online), is favourable to sediment accumulation, which are likely not being discharged to the main river during normal hydrological events.

As hinted by previous analysis of sediment loads on the KR, the annual load is much influenced by transport throughout the year, not only by one important event such as the spring freshet, as is observed in many other river systems. While it might still be the case that most transport occurs during a few events, these might not be particularly high-magnitude events, and can thus go unnoticed without a closer look at the hydrograph. The result can be an unforeseen volume of sediments to remove from the harbour, especially when high amounts are expected only as a result of major flooding events.

As such, the general idea expressed by the SJPA that dredging volumes should be related to the magnitude of the previous spring flood does not seem to hold true, according to the limited data set analysed in this project. Rather, it seems that total water yield is the single most important parameter to take into account, followed by mean water level. For instance, the hydrological year 2005-2006, for which dredging volumes were especially high, did not have an extreme freshet. However, three other unusually high peaks for their respective season occurred in the previous fall and in mid-January. This was reflected in the total annual water yield, with a high value that had not been observed since 1983. It should be noted that precipitation was the second most important parameter, until data from the 2009 dredging campaign was included in the analysis. While a higher water yield was observed in 2007-2008, lower rain occurred that year. In addition, almost twice as much snow fell in Saint John in 2007-2008, likely protecting the ground from erosion for a longer period. While snowfall was not identified as a major variable in the present regression, it should be considered in future models. Using degree-days in the spring could, for instance, give an idea of the timing and speed of snowmelt. As long as the dataset will remain as small as it is, the explanatory variables are likely to change every time a new year is added.

It has been acknowledged before that the timing of different processes affecting sediment transport plays a major role in the magnitude of annual loads. For instance, for similar total water yield between two years, Oeurng et al. (2010) observed a 470% increase in sediment load in the second year, partially due to the timing of large floods and tilling activities in the fall of that year. This indicates that hydro-meteorological variables alone might not be sufficient to investigate inter-annual variation in sediment transport, and that anthropogenic influences should ideally be somehow taken into account. Yet, the results obtained in the present study seem to indicate an appropriate predictive model could be eventually developed for the SJPA using simple, easy-to-access hydro-meteorological variables. Finally, the results presented here should be examined in the context of the studies done in parallel in the estuary and the Bay of Fundy, looking at the impact of tides and currents on sediment deposition and accumulation.

2.3.7 References

- Bray, D. I. and Xie, H. 1993. A regression method for estimating suspended sediment yields for ungauged watersheds in Atlantic Canada. *Canadian Journal of Civil Engineering*, 20:82-87.
- Haigh, S. and Clarke, J. H. [Online] Numerical Modelling of Kennebecasis Bay. Ocean Mapping Group, UNB. Accessed from <http://www.omg.unb.ca/people/haigh/kenneb.html> on August 10th, 2010
- Hansen, D. and Bray, D. I. 1993. Single-station estimates of suspended sediment loads using sediment rating curves. *Canadian Journal of Civil Engineering*, 20:133-143.
- Higgins, H., A. St-Hilaire, S.C. Courtenay & K.A. Haralampides. 2011. Suspended sediment dynamics in a tributary of the Saint John River, New Brunswick. *Canadian Journal of Civil Engineering*, 38(2):221-232.
- Oeurng, C., Sauvage, S. and Sanchez-Perez, J.-M. 2010. Dynamics of suspended sediment transport and yield in a large agricultural catchment, southwest France. *Earth Surface Processes and Landforms* 35(11):1289-1301.
- Rao, A. R. et Hamed, K. H. 2001. *Flood Frequency Analysis*. New York, USA. 376 p.
- Tramblay, Y., T. B.M.J. Ouarda A. St-Hilaire, J. Poulin. 2010. Regional estimation of extreme suspended sediment concentrations using watershed characteristics. *Journal of Hydrology* 380: 305-317.

Appendice A – Loïs de probabilité

