

Etude de faisabilité pour l'estimation du débit en rivière au Québec méridional à l'aide du capteur satellitaire MODIS au sein de l'INRS-ETE

Société d'accueil : **INRS – ETE**

PFE présenté par : **Jennifer NOËL**

Directeur du PFE : **M. Kareem CHOKMANI**

Correcteurs : **Mme Tania LANDES & M. Gilbert FERHAT**



1. Introduction

Ce projet a été mené au sein du centre de recherche ETE (Eau Terre et Environnement) qui fait parti du réseau INRS (Institut National de la Recherche Scientifique).

Ce projet consiste en une étude de faisabilité pour l'estimation du débit en rivière. L'originalité de la méthode tient en l'étude des modifications de réflectance du cours d'eau afin d'établir une corrélation avec un changement de débit. L'étude bibliographique a démontré que cette méthode a été peu développée et que les techniques d'imagerie radar sont à l'heure actuelle les plus fréquentes. Il est donc question d'étudier le potentiel de l'imagerie satellite et plus particulièrement l'imagerie du capteur optique MODIS, afin de développer une méthode d'estimation du débit en rivière et établir, le cas échéant, une approche opérationnelle sur un grand nombre de rivières. Une telle approche serait d'une grande utilité pour un territoire aussi étendu que le Québec, où de nombreux cours d'eau d'intérêts économiques et écologiques se trouvent hors de la portée du réseau d'observation hydro-métrique.

2. Méthodologie

2.1. Zone d'étude

L'étude porte sur des images du capteur MODIS du 15 Mai au 15 Octobre, de 2000 à 2008, prises de jour et dont la région d'étude se situe au Sud du Québec.

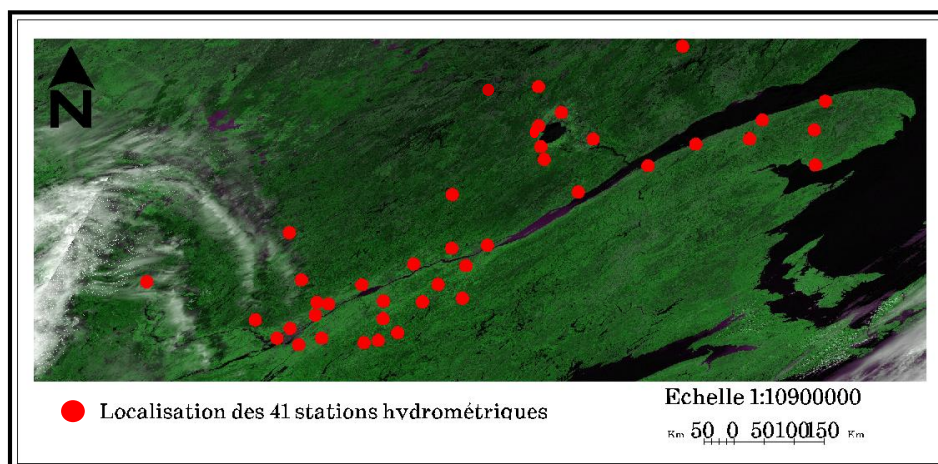


Figure 1 : Image MODIS montrant les 41 stations hydrométriques de l'étude (points rouges)

Trois bases de données ont été créées : une base Carto, Image et Hydro ; c'est à partir de celles-ci que tous les traitements ont été effectués.

2.2. Création des bases de données

➤ Base Carto

La base de données intitulée Carto a été réalisée avec le module ArcMap¹. Cette base doit contenir toutes les informations relatives aux cours d'eau du Sud du Québec qui serviront pour le projet. Ces informations sont de différentes natures : vecteurs des cours d'eau, toponymie des lieux et des rivières, positions des stations hydrométriques et modèles numériques de terrain.

➤ Base Image

Cette base regroupe toutes les images du capteur MODIS qui ont été commandées sur le site de la Nasa. Ces images ont une latitude allant de 39 à 50 degrés Nord et une longitude de -50 à -93 degrés Ouest. Le choix des bandes des images s'est limité à la commande des bandes 1 et 2 qui représentent respectivement le rouge (620-670 nm) et le proche infrarouge (841-876 nm).

Les traitements de ces images consistent en la création de la mosaïque d'images et du masque de nuages. Je suis alors partie sur l'idée de réaliser un masque de nuages par simple seuillage, afin d'obtenir une couche bitmap représentant les pixels d'eau et de terre non contaminés par des nuages. Le seuil choisi pour le masque est de 1845 (valeur en 16-bit signed integer, avec un facteur d'échelle de 0,0001).

➤ Base Hydro

Cette base de données est composée de fichiers textes contenant toutes les mesures des débits des périodes de l'étude, pour les stations hydrométriques qui ont été choisies. Ces mesures de débit serviront à la construction des courbes de tarage qui mettent en relation la réflectance du cours d'eau à son débit. Pour la sélection des stations, trois critères de sélection ont été retenus :

- Premier critère : position géographique des stations hydrométriques limitées au Sud du Québec.
- Deuxième critère : essentiel pour l'étude, il concerne le régime hydrologique du cours d'eau sur lequel est située la station hydrométrique. Les rivières visibles seulement sur les images satellites sont retenues car cela signifie qu'un changement de régime dans leur débit va modifier la réflectance du pixel.
- Troisième et dernier critère : le nombre d'années d'observation. La taille des séries de données des stations doit respecter un nombre d'années d'observation minimal pour des raisons de fiabilité de l'étude statistique du projet (8 à 9 années d'observation).

2.3. Construction des zones d'analyse et extraction de l'information spectrale

Pour extraire la réflectance concernant les 41 stations hydrométriques, la création de vecteurs sous ArcGis, matérialisant des zones d'analyse, m'a parue évidente. Ces zones sont de deux types : une partie appelée « zone active (ZA) » car elle est placée sur la rivière et une autre partie appelée « zone non active (ZNA) », placée de part et d'autre de la zone active et qui ne contient normalement aucun pixel d'eau (cf. Figure 2). Pour déterminer la largeur de ces zones, j'ai procédé par dichotomie et les largeurs suivantes ont été trouvées : 1250 m / 625 m (ZA et ZNA) pour les rivières avec une largeur inférieure à 150 m et 1500 m / 750 m (ZA et ZNA) pour les rivières supérieures à 150 mètres. Pour la longueur, je me suis basée sur mes recherches bibliographiques et j'ai pris l'option de prendre trois méandres en longueur totale.

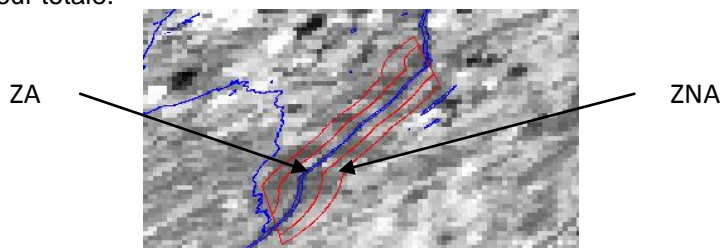


Figure 2 : Image MODIS avec le vecteur du cours d'eau et de la zone

Pour le traitement des zones, j'ai créé un algorithme de traitement d'images avec le module *Easi* de Geomatica. Pour cela, j'ai transformé dans un premier temps les vecteurs des zones en raster puis en bitmap, tout en gardant leur identifiant. Puis, j'ai copié une image MODIS que j'ai nommée « ZAZNA », et avec laquelle tous les traitements ont été effectués (cf. Figure 3).

¹ ArcMap version 9.2 est un module de ArcInfo version de ArcGis : ensemble de logiciels SIG réalisé par la société ESRI.

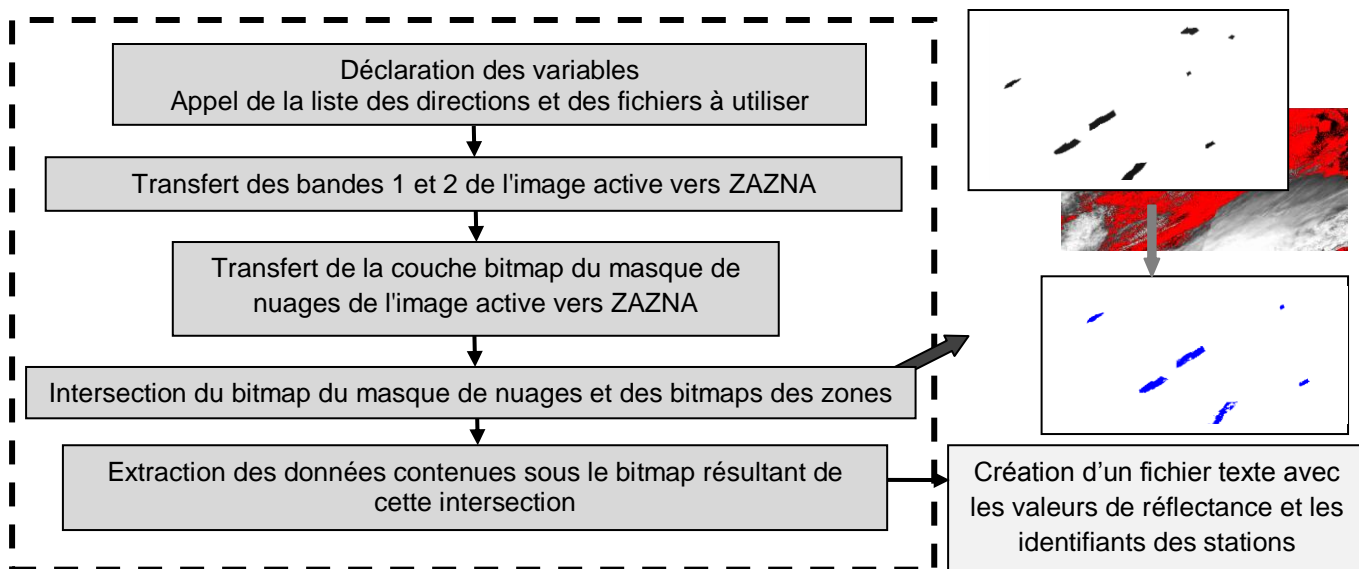


Figure 3 : Schéma explicatif des traitements réalisés pour l'extraction de l'information spectrale sous Geomatica

2.4. Construction des courbes de tarage et validation

Pour la création des courbes de tarage, j'ai utilisé le logiciel Matlab. Les différents algorithmes créés et utilisés sont décrits ci-dessous :

1. Calcul de la moyenne ou de la médiane des réflectances par jour et pour chaque zone
-> Création des matrices Matza et Matzna pour les valeurs respectives des ZA et des ZNA

2. Traitement des valeurs des moyennes et des médianes calculées pour la bande 1, 2 et le NDVI (Normalized Difference Vegetation Index) et regroupement des valeurs dans des matrices, pour cela :

-> Appel des matrices Matza et Matzna

-> Rapport des valeurs obtenues de la ZNA sur celles de la ZA pour la bande en question :

$$\text{Bande 1 : } R_1 = \frac{R_{1ZNA}}{R_{1ZA}} \quad \text{Bande 2 : } R_2 = \frac{R_{2ZNA}}{R_{2ZA}} \quad \text{NDVI: } R_{NDVI} = \frac{(R_{2ZNA} - R_{1ZNA})}{(R_{1ZNA} + R_{2ZNA})}$$

-> Création des matrices Rf_R1, Rf_R2 et Rf_NDVI pour le groupement des calculs

3. Construction des courbes de tarage débit/réflectance

-> Calcul de la courbe de régression, du r^2 , du F statistique, du P-value et de l'erreur de variance

4. Validation croisée des résultats avec :

-> Calcul de l'équation exponentielle de la courbe de régression avec une valeur en moins

-> Calcul du débit de la valeur qui a été écartée

-> Comparaison avec les valeurs des débits téléchargés et calcul des écarts moyens quadratiques

3. Résultats

3.1. Choix du mode de traitement

Le masque de nuages a tout d'abord été modifié. Un seuil à 500 (valeur en 16-bit signed integer, avec un facteur d'échelle de 0,0001) a été choisi car l'ancien seuil à 1845 prenait en compte des pixels contaminés par des voiles de nuages (nuages avec une réflectance plus faible que la moyenne). Ces pixels de nuages ont pour effet de modifier la réflectance totale de la zone. Ce nouveau seuil laisse cependant de côté un certain nombre de pixels du sol, pour des zones de cultures qui présentent une réflectance moyenne voisine de celle des nuages.

Pour le choix du mode de traitement, j'ai décidé d'effectuer des tests sur toutes les zones du projet puis de comparer les résultats. Ces tests consistent à effectuer les traitements pour la bande 1, 2 et pour le NDVI. Pour chacun d'eux, j'ai testé l'influence des nuages, de la médiane et de la moyenne des valeurs. L'influence des nuages a été mesurée en tolérant 5 à 10 % de pixels contaminés dans les zones. Après traitements, il est apparu que les meilleures réponses sont obtenues en prenant la bande 1, la médiane des valeurs et en acceptant 10% de pixels contaminés dans les zones.

Quatre corrélations ont ainsi été trouvées : trois significatives et une moins forte (cf. *Tableau 1*) et ceci pour des rivières avec une largeur moyenne supérieure à la taille d'un pixel. Pour la précision demandée, je me suis reportée à une étude réalisée par Pavelsky et Smith en 2008, dans laquelle est utilisé le même type d'images, à savoir des images optiques du capteur MODIS. Pour leur projet, l'erreur moyenne entre le débit estimé et le débit observé était de $\pm 325 \text{ m}^3/\text{s}$, soit 10% du débit moyen, précision que l'on retrouve dans les résultats ci-dessous (cf. *Tableau 1*) :

Tableau 1 : Résultats avec la bande 1, la médiane et 10% de pixels contaminés

Id_Zone	r ²	P-value	Nombre_Images	Q moyen été/aut. (m ³ /s)	Emq (m ³ /s)
360	0,609	1,18e-008	37	840	± 89
370	0,380	1,11e-16	173	641	± 91
390	0,417	3,11e-007	51	643	± 67
400	0,234	2,40e-012	194	339	± 58

3.2. Analyse des résultats

Dans un premier temps, j'ai essayé d'améliorer les réponses des zones d'analyse qui ne présentent aucune corrélation et qui sont situées sur des rivières dont la largeur moyenne est inférieure à la taille du pixel. Pour cela, j'ai cherché à déterminer d'autres calculs pour estimer la réflectance de l'eau dans ces zones. Ces calculs ont été effectués dans trois zones représentatives de l'étude mais n'ont finalement donné aucun résultat. Par la suite, j'ai essayé de déterminer s'il y avait une relation entre le modèle d'élévation de la zone et l'absence de corrélation. Mais là aussi, aucune conclusion n'a pu être trouvée. Il semblerait au final, que quand la rivière n'est pas assez large (inférieure à la taille d'un pixel), l'information spectrale des pixels d'eau dans la zone d'analyse est noyée dans celle des pixels de la végétation.

Je me suis ensuite penchée sur six autres zones où la largeur moyenne du cours d'eau est supérieure à la taille du pixel mais où il n'y a pas de corrélation. Outre le fait que pour certaines de ces six zones, je manquais d'information spectrale car le masque de nuages choisi n'était pas adapté, il semblerait qu'une morphologie en trapèze du cours d'eau soit indispensable à la présence d'une corrélation avec le débit. Les zones situées en plaine inondable, même avec une largeur supérieure à la taille du pixel, ne donnent aucun résultat concluant pour une corrélation débit/réflectance (cf. *Figures 4 et 5*).

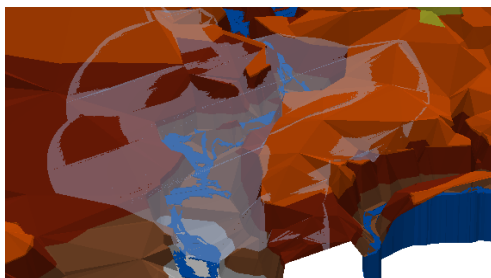


Figure 4 : Topologie en trapèze - corrélation existante

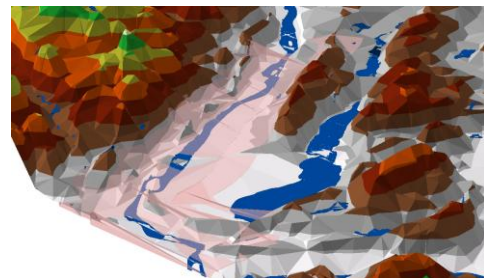


Figure 5 : Plaine inondable - corrélation absente

4. Conclusion

Après analyse des résultats, la conclusion est que le premier objectif du projet est presque atteint. On peut dire qu'une corrélation existe bien entre la réflectance au sol et le débit, mais sous certaines conditions que l'on peut nommer optimales. Il faut en effet être en présence de rivières avec un fort débit (supérieur à $300 \text{ m}^3/\text{s}$ en moyenne), dont la largeur moyenne est supérieure à la résolution spectrale de l'image et présentant une topologie en trapèze du terrain. Le second objectif ne sera pas atteint car je n'ai pas obtenu assez de zones présentant des corrélations débit/réflectance.

Cette étude de faisabilité a donné des résultats encourageants mais n'est pas idéale en l'état actuel. Il y a des points de recherche à approfondir et des pistes à explorer pour une amélioration de la méthode. On retient ainsi :

1. Amélioration du masque de nuages avec l'introduction d'autres bandes
2. Utilisation des données des débits horaires et non des débits journaliers comme cela a été fait
3. Sélection d'un plus grand nombre de rivières avec une largeur moyenne équivalente à celle d'un pixel
4. Possibilité d'utiliser des images radar pour palier les problèmes liés aux précipitations.