

RAPPORT TECHNIQUE

Version finale

CONSERVATION DES ÎLOTS DE FRAÎCHEUR URBAINS

DESCRIPTION DE LA MÉTHODE SUIVIE POUR IDENTIFIER ET LOCALISER LES ÎLOTS DE FRAÎCHEUR ET DE CHALEUR (MÉTHODE EN 9 NIVEAUX)

Présentée à :

**Institut national de santé publique
du Québec**
Pierre Gosselin

Préparée par :



Centre d'enseignement et de recherche
en foresterie de Sainte-Foy inc.

Emmanuelle Boulfroy, M.Sc.
Jalal Khaldoune, géographe, Ph.D.
Frank Grenon, biol., Ph.D.



Richard Fournier, Ph.D., professeur titulaire
Benoît Talbot, étudiant à la Maîtrise

Mai 2013

Mots-clés : conservation, îlots, fraîcheur, chaleur, SPOT, Landsat, températures, végétation

Référence à citer :

Boulfroy, E., J. Khaldoune, F. Grenon, R. Fournier et B. Talbot. 2013. Conservation des îlots de fraîcheur urbains - Description de la méthode suivie pour identifier et localiser les îlots de fraîcheur et de chaleur (méthode en 9 niveaux). CERFO et Université de Sherbrooke. Rapport 2012-11c. 40 pages.

Ce projet est financé en partie par le Fonds vert dans le cadre du Plan d'action 21 du Plan d'action 2006-2012 sur les changements climatiques (PACC) du gouvernement du Québec.

TABLE DES MATIÈRES

LISTE DES FIGURES.....	III
LISTE DES TABLEAUX	III
RÉSUMÉ LONG	IV
1. MISE EN CONTEXTE DU PROJET	1
2. OBJECTIFS DU PROJET	2
3. LOCALISATION DU TERRITOIRE D'ÉTUDE	3
4. PRINCIPALES ÉTAPES DU PROJET.....	4
5. DONNÉES UTILISÉES POUR LA CONSTRUCTION DU MODÈLE	6
5.1. <i>Sous-échantillonnage du territoire d'étude.....</i>	6
5.2. <i>Choix des données satellitaires utilisées.....</i>	6
5.2.1. <i>Images acquises par le capteur ETM+ de Landsat.....</i>	7
5.2.2. <i>Images SPOT-5</i>	8
5.3. <i>Description de la variable dépendante du modèle.....</i>	9
5.3.1. <i>Méthode de calcul de la température de surface</i>	9
5.3.2. <i>Absence de normalisation des données de température.....</i>	11
5.3.3. <i>Réduction de la dimension des pixels des images Landsat.....</i>	12
5.4. <i>Description des variables explicatives.....</i>	13
5.4.1. <i>Données climatiques</i>	15
5.4.2. <i>Indice de végétation normalisé (NDVI).....</i>	15
5.4.3. <i>Indice d'imperméabilité des surfaces.....</i>	15
5.4.4. <i>Classification de l'occupation du sol</i>	17
5.5. <i>Échantillonnage</i>	20
6. ÉLABORATION DU MODÈLE	22
6.1. <i>Portrait général des variables utilisées</i>	22
6.2. <i>Vérification de la colinéarité entre les variables</i>	22
6.3. <i>Construction du modèle</i>	25
7. VALIDATION DU MODÈLE	27
8. APPLICATION DU MODÈLE À L'ENSEMBLE DU TERRITOIRE D'ÉTUDE	28
8.1. <i>Données SPOT-5 utilisées.....</i>	28
8.2. <i>Réalisation de la classification de l'occupation du sol.....</i>	31
8.2.1. <i>Construction préalable de mosaïques d'images SPOT-5</i>	31
8.3. <i>Production d'une carte de températures relatives de surface.....</i>	31
8.4. <i>Description des plages de températures</i>	32
8.4.1. <i>Méthode utilisée.....</i>	32
8.4.2. <i>Mise en garde concernant les terres agricoles nues</i>	33
8.4.3. <i>Essai de normalisation des valeurs de températures prédites.....</i>	34
8.5. <i>Exemple de la cartographie de température relative.....</i>	36
9. BILAN DES LIMITES MÉTHODOLOGIQUES DE L'ÉTUDE.....	38
10. PISTES D'AMÉLIORATIONS POSSIBLES.....	39
RÉFÉRENCES BIBLIOGRAPHIQUES	40

LISTE DES FIGURES

Figure 1. Principales étapes suivies lors de l'élaboration de la carte de températures de surface.....	5
Figure 2. Exemple de carte des classes de température de surface calculées à partir de la bande thermique du capteur ETM+ de Landsat pour la région de Québec.....	11
Figure 3. Schéma général du jeu de données utilisées comme variables explicatives du modèle de prédiction de la température de surface à partir des images SPOT-5 (20 m).....	14
Figure 4. Température Landsat en fonction de la température prédite pour les 3 509 points d'échantillonnage non utilisés lors de la construction du modèle.....	27
Figure 5. Localisation des images SPOT-5 utilisées pour couvrir tout le territoire d'étude et année de prise de chaque image.....	30
Figure 6. Exemple d'erreur due à la classification de la température de surface avec des données normalisées des températures prédites.....	35
Figure 7. Carte de températures de surface pour les villes de Québec et de Lévis (carte générale et zooms).....	37

LISTE DES TABLEAUX

Tableau 1. Liste des régions socio-sanitaires pour lesquelles au moins une image a été retenue pour l'élaboration du modèle.....	6
Tableau 2. Dates d'acquisition des images SPOT-5 et des images Landsat correspondantes (couvrant le même secteur).....	8
Tableau 3. Classes d'occupation du sol et catégories considérées.....	17
Tableau 4. Matrices de confusion pour chaque classe d'occupation du sol.....	19
Tableau 5. Portrait général des variables.....	22
Tableau 6. Présentation des valeurs de rho pour évaluer la colinéarité.....	24
Tableau 7. Date d'acquisition des images SPOT-5 utilisées pour couvrir tout le territoire d'étude et images Landsat correspondantes.....	29

RÉSUMÉ LONG

1. Cadre méthodologique général

Une première cartographie des îlots de chaleur a été élaborée en 2010. Suite aux demandes des utilisateurs de SUPREME formulées depuis, il a été décidé de bonifier cette première cartographie et d'en développer une nouvelle, élargie aux zones moins peuplées qui sont contigües aux zones urbaines et dont la résolution serait meilleure. La cartographie des îlots de chaleur et de fraîcheur repose sur un modèle spatial de prédiction de la **température relative de surface**, appliqué à l'ensemble des zones habitées de plus de 400 habitants par km² au Québec.

Le projet dans lequel s'inscrit l'élaboration du modèle prédictif présente donc le défi de développer une **méthode qui devra être appliquée à un très grand territoire d'étude** (tout le Québec urbain), **et ce, avec une enveloppe budgétaire limitée**. Dans ce contexte, certains choix méthodologiques ont été faits, de manière à permettre l'élaboration d'un modèle applicable pour tout le Québec urbain, selon un coût raisonnable.

La méthode utilisée est une adaptation d'un précédent modèle de prédiction de la température de surface, utilisant des images Landsat (Kestens *et al.*, 2011¹). Il a été décidé, en premier lieu, d'utiliser le même cadre méthodologique et sensiblement les mêmes variables explicatives que pour le modèle développé par Kestens *et al.* (2011), avec quelques ajustements (par exemple, l'ajout de quelques classes d'occupation du sol). Les variables explicatives à tester dans le nouveau modèle prédictif ont donc été choisies *a priori*, et l'amélioration du pouvoir prédictif du modèle de température devait provenir principalement de l'augmentation de la résolution du type d'image satellitaire utilisé. En effet, la nouvelle approche propose d'utiliser des images satellitaires du capteur SPOT-5 dont la résolution est de 20 m, comparativement à 60 à 100 m dans le cas des bandes thermiques des images Landsat. Notons que parmi le choix de ces variables, deux modifications au modèle initial ont tout de même été apportées, ayant été identifiées par les auteurs de ce dernier comme des pistes intéressantes d'amélioration (ces changements sont présentés dans la section qui décrit les variables explicatives).

2. Données utilisées pour l'élaboration du modèle

Un sous-échantillonnage du territoire d'étude a été réalisé pour l'élaboration et la validation du modèle. Seize images acquises avec le capteur SPOT-5 ont été sélectionnées, de manière à couvrir l'ensemble de l'étendue géographique du territoire d'étude et toute la diversité du tissu urbain retrouvé au Québec. Pour chaque secteur couvert par une scène SPOT-5, la sélection d'une image acquise par le capteur Landsat ETM+ a été réalisée dans le but d'obtenir la mesure réelle de température, étant donné que les images SPOT-5 ne disposent pas de bande thermique. La sélection des 15 images Landsat dans les archives disponibles a été réalisée de manière à ce que leur date d'acquisition soit la plus proche possible (en termes de date journalière) de celle des images SPOT-5 correspondantes. Dans certains cas, il a été nécessaire d'avoir recours à des images prises à des années différentes mais rapprochées le plus possible de l'année de référence.

¹ Kestens, Y., A. Brand, M. Fournier, S. Goudreau, T. Kosatsky, M. Maloley et Audrey Smargiassi. 2011. *Modelling the variation of land surface temperature as determinant of risk of heat-related health events*. International Journal of Health Geographics, p. 2-9.

Le tableau suivant présente les dates d'acquisition des images SPOT-5 et Landsat ETM+ correspondantes. À trois reprises, deux images Landsat ont été nécessaires pour couvrir toute l'étendue de l'image SPOT correspondante.

Date de la scène de l'image SPOT-5				Date de l'image Landsat correspondante		
Année	Mois	Jour	id*	Année	Mois	Jour
2005	8	26		2005	8	6
2006	7	19		2006	7	19
				2007	7	29
2006	8	30	a	2008	8	21
2006	8	30	b	2007	7	25
2007	6	18		2011	6	15
2007	8	2		2007	8	21
				2007	8	5
2007	8	28		2007	8	21
2007	9	4	a	2007	8	21
2007	9	4	b	2008	8	21
2008	7	7		2010	7	23
				2005	7	17
2008	8	21		2008	8	26
2009	8	15		2010	8	20
2009	9	6		2010	6	17
2010	6	21		2010	6	17
2010	9	1		2010	8	20
2011	7	2		2011	7	10

* Deux images ont été prises la même journée mais ne sont pas localisées au même endroit.

La variable dépendante du modèle, soit la température de surface, est calculée à partir de la bande thermique (bande 6) des images Landsat. Le choix d'utiliser des images Landsat pour calculer la variable dépendante du modèle repose sur la décision d'utiliser le même cadre méthodologique que pour le modèle développé par Kestens *et al.* (2011).

Étant en présence de plusieurs images Landsat prises dans des conditions météorologiques pouvant être très différentes, la plage de températures de surface couverte sur chaque image peut être, elle aussi, très variable d'une image à l'autre. Malgré ce constat, il a néanmoins été décidé de ne pas normaliser la donnée de température réelle à l'échelle des 15 images Landsat, mais plutôt d'utiliser les valeurs brutes de température calculées à partir du canal thermique pour chaque image Landsat. Cette décision a été dictée par l'incertitude concernant la plus-value de l'étape de normalisation (dans certains cas, la normalisation nivelle certaines informations qui pourraient influencer la sélection des variables explicatives significatives du modèle) dans un contexte où les budgets ne permettaient pas de faire plusieurs essais. Il est également essentiel de ne pas s'écarter de l'objectif final du modèle, qui consistait à produire une cartographie de température relative permettant de distinguer les zones les plus fraîches des zones les plus chaudes. Dans cette situation, on pose donc l'hypothèse que l'approximation éventuelle du

modèle liée à la présence de plages de températures de surface variables d'une image Landsat à l'autre ne devrait pas affecter de manière importante le résultat final recherché.

Les variables explicatives du modèle qui se sont avérées significatives dans la procédure de régression statistique et qui ont été retenues dans le nouveau modèle sont :

- Indice de végétation NDVI, calculé à partir des images SPOT-5;
- Indice d'imperméabilité, calculé à partir des images SPOT-5 (variable ajoutée par rapport au modèle initial);
- Classes d'occupation du sol provenant des images SPOT-5 : (1) zone urbaine pure correspondant aux chaussées, toits de maisons, édifices de grandes surfaces, pistes d'atterrissage, (2) zone urbaine mixte correspondant aux maisons et édifices avec présence de végétation, (3) forêt correspondant aux zones végétalisées dominées par les arbres, (4) zone de basse végétation correspondant aux champs agricoles, parcs récréatifs, jardins communautaires, terrains de sport, terrains ouverts et (5) eau (le choix d'avoir 5 classes constitue une amélioration comparativement au modèle initial, qui ne comportait que 3 classes : forêt, eau et bâti-sol nu);
- Localisation (coordonnées X et Y);
- Distance par rapport à l'eau;
- Altitude;
- Température moyenne de l'air lors de la journée d'acquisition de l'image Landsat provenant de relevés météorologiques;
- Moyenne des températures 72 heures avant la date d'acquisition des images Landsat provenant de relevés météorologiques.

3. Élaboration et validation du modèle

Un échantillonnage de 35 099 pixels stratifié selon les classes d'occupation du territoire a été réalisé. Le modèle de prédiction de la température a été établi selon une approche par modèle mixte. L'analyse statistique a montré que le NDVI et l'occupation du sol sont les deux variables complémentaires qui présentent les plus fortes valeurs explicatives dans le modèle. Le coefficient de corrélation (R^2) du modèle est de 0,67 avec une erreur quadratique moyenne (RMSE) de 2,35°C, soit 9,6 %.

Le modèle théorique obtenu a été validé à partir de points d'échantillonnage non utilisés pour la construction du modèle. À cette fin, des valeurs prédites ont été obtenues par itération pour l'ensemble des données de l'échantillon. À chaque itération, 90 % des données ont servi à prédire le 10 % restant utilisé pour la validation. La validation du modèle montre également un coefficient de corrélation (R^2) de 0,67 entre les valeurs observées et prédites. Quant à l'erreur quadratique moyenne (RMSE) de validation, elle est de 2,36°C, soit 10,1 %. La relation linéaire entre les valeurs observées et prédites présente une pente de 1 et une ordonnée à l'origine qui n'est pas significativement différente de zéro.

4. Application du modèle à l'ensemble du territoire d'étude pour la production de la carte de température relative

Le modèle développé a été appliqué à chacune des 67 images SPOT-5 acquises entre 2005 et 2011 et couvrant toutes les zones urbaines du Québec. Parmi ces 67 images, la grande majorité (85 %) a été acquise en 2007 ou lors d'années ultérieures. La carte résultante de la température de surface permet de localiser avec une précision aux 20 m près, selon un gradient relatif, les zones couvertes par le spectre de températures les plus chaudes (îlots de chaleur) et les zones couvertes par le spectre de températures les plus froides (îlots de fraîcheur) en milieu habité.

Neuf niveaux de températures relatives ont été retenus pour distinguer les îlots de fraîcheur des îlots de chaleur. La représentation spatiale de température en neuf niveaux utilise la méthode de Jenks, disponible dans ArcGIS, qui est une méthode inductive consistant à choisir les seuils qui maximisent la variance interclasse et minimisent la variance intraclasse. Si la distribution comporte des discontinuités (zones de faible densité des valeurs), celles-ci seront nécessairement détectées par la méthode de Jenks, qui constitue une variante rigoureuse de la méthode bien connue des « seuils naturels ». Notons que plusieurs tests de nombres de classes ont été réalisés et ont permis de confirmer que l'utilisation de 9 niveaux était optimale pour identifier les zones les plus chaudes (niveaux 8 et 9), correspondant à une concentration élevée de surfaces imperméables et recoupant les zones identifiées comme des îlots de chaleur dans l'analyse de Kestens *et al.* (2011).

Un filtre de 3 pixels X 3 pixels (60 m X 60 m) a également été appliqué aux pixels associés aux niveaux 8 et 9 (correspondant aux îlots de chaleur), de manière à ne retenir parmi les îlots de chaleur, que les zones de concentration de plus de 3 600 m².

Avec un découpage en 9 niveaux, il est également apparu qu'il était possible de distinguer deux principales classes d'îlots de fraîcheur, les zones boisées qui sont plus fraîches (correspondant aux niveaux 1 et 2) des zones végétalisées plus ouvertes un peu moins fraîches (correspondant aux niveaux 3 et 4).

La représentation spatiale de la distribution des neuf niveaux de température a été réalisée scène par scène, étant donné que l'amplitude des températures observées au sein d'une même scène varie d'une scène à l'autre. Ces différences entre scènes s'expliquent par le fait que les images n'ont pas toutes été prises dans des conditions météorologiques semblables (dates d'acquisition des images Spot-5 différentes).

Rappelons enfin que cet exercice de cartographie des températures de surface relatives comporte diverses sources d'imprécision souvent non contrôlables (les diverses dates d'acquisition des images satellitaires, qui entraînent des conditions météorologiques qui peuvent être variables d'une image à l'autre) qui doivent être considérées lors de l'interprétation. Par exemple, après vérification des cartes de niveaux de température obtenus, il s'est avéré qu'il existait une confusion entre d'une part, les champs agricoles nus et les îlots de chaleur, et d'autres parts, les étendues d'eau et les îlots de chaleur. La température prédite pour les terres agricoles nues et les étendues d'eau s'apparente en effet à celle des îlots de chaleur, de par leur signature spectrale qui présente de fortes similitudes avec les zones hautement imperméables. Il est donc possible que certains jardins communautaires en ville et certains bassins, lacs ou étangs soient identifiés

comme des îlots de chaleur. Il convient alors d'exclure les étendues d'eau de la cartographie des îlots de chaleur.

Certains choix méthodologiques ont également été faits en lien avec les contraintes budgétaires et temporelles du projet (dont l'adaptation d'un modèle déjà existant). Les classes de températures de surface relatives doivent donc être considérées comme des indicateurs robustes, permettant d'identifier et de localiser les îlots de fraîcheur et de chaleur dans un milieu urbain. Nous sommes confiants que ces résultats pourront être utiles aux fins de planification et d'intervention en santé publique et protection civile, tant par les autorités sanitaires que municipales.

1. MISE EN CONTEXTE DU PROJET

Depuis plusieurs années, voire quelques décennies, on ne peut plus nier la réalité des changements climatiques à l'échelle mondiale. Selon le GIEC, onze des douze années les plus chaudes jamais observées ont été enregistrées depuis 1995. Au Québec, les dix années les plus chaudes du siècle ont été enregistrées à partir des années 1980 et les prévisions indiquent que les températures moyennes continueront à augmenter durant les prochaines décennies. Cette hausse constante de la température (observée et projetée) ainsi que la présence de périodes de températures extrêmes, en particulier en été, accentueront vraisemblablement un problème déjà connu : **l'effet des îlots de chaleur urbains** (Dousset *et al.*, 2011). Ce phénomène se définit par des températures estivales plus élevées dans les milieux urbains que dans les zones rurales avoisinantes, pouvant entraîner des problèmes de santé importants, voire la mort chez certains groupes de populations. Les populations les plus vulnérables sont principalement les individus plus fragiles (personnes âgées, malades, jeunes enfants) et les personnes défavorisées (ayant moins facilement accès à des mesures permettant d'atténuer les effets des îlots de chaleur telles que l'accès à des climatiseurs, des piscines, des cours ombragées, etc.).

Il devient urgent de mettre en place des mesures pour lutter contre les effets nocifs des îlots de chaleur en milieu urbain et minimiser leurs impacts. Une gestion efficace de la végétation et de l'eau en milieu urbain, créant des **zones de fraîcheur urbaines (ou îlots de fraîcheur)**, en est l'une des principales solutions. Des baisses de température de l'ordre de plusieurs degrés peuvent en effet être observées à proximité d'espaces verts, atténuant alors les effets des îlots de chaleur. En plus d'agir comme un climatiseur naturel en diminuant la température ambiante des villes et en favorisant une meilleure ventilation, les espaces boisés urbains sont également responsables de plusieurs autres bienfaits sur la santé : ils réduisent les gaz polluants présents dans l'air, ils filtrent les poussières fines et les aérosols en suspension dans l'air, et ils contribuent à réduire les risques de cancer de la peau et de cataractes et constituent un milieu propice à la tenue d'activités physiques et de plein air. Bien entendu, les espaces verts en milieu urbain remplissent aussi de nombreuses autres fonctions écologiques, sociales et économiques.

La représentativité des îlots de fraîcheur dans les villes ainsi que leur répartition spatiale peuvent être très variables. L'éventualité qu'ils soient détruits dans le cadre de projets de développement immobilier l'est tout autant. L'élaboration d'un outil permettant de localiser les îlots de chaleur et les îlots de fraîcheur à une échelle relativement fine devient très pertinente pour la lutte aux effets des îlots de chaleur, en plus d'être un appui à l'aménagement urbain axé sur la qualité de vie de ses habitants.

2. OBJECTIFS DU PROJET

Quatre principaux objectifs sont visés par la présente étude :

- 1) Améliorer la cartographie des îlots de chaleur générée par l'INSPQ selon la méthode développée par Kestens *et al.* (2011), en utilisant une image SPOT-5 (20 m), ce qui permettra d'augmenter la résolution de la nouvelle cartographie.
- 2) Adapter le modèle développé par l'INSPQ qui permet d'identifier les îlots de chaleur, pour localiser les îlots de fraîcheur.
- 3) Localiser et caractériser les principaux îlots de fraîcheur présents dans les centres urbains de plus de 400 habitants/km² du Québec, et ce, à 20 m près.
- 4) Identifier les secteurs à prioriser pour la protection des îlots de fraîcheur existants et l'implantation de nouveaux îlots.

3. LOCALISATION DU TERRITOIRE D'ÉTUDE

Le territoire couvert par le présent projet comprend :

- l'ensemble des zones urbaines du Québec dont la densité de population est supérieure ou égale à 400 habitants par km²;
- une zone tampon de 2 km de large qui a été ajoutée autour de ces secteurs, afin d'éviter d'exclure certains secteurs situés en milieu urbain qui n'auraient pas la densité de population requise (ex : certains quartiers industriels).

4. PRINCIPALES ÉTAPES DU PROJET

Un premier modèle permettant de prédire la température au sol pour l'ensemble des zones habitées de plus de 400 habitants par km² avait été développé en collaboration avec l'INSPQ en 2010 (Kestens et al., 2011). Dans le cadre du projet actuel, il a été convenu initialement que l'on utiliserait ce modèle, construit à partir d'images Landsat (résolution de 100 m), pour prédire la température de surface à partir d'images SPOT-5 (résolution de 20 m). L'utilisation d'images SPOT et Landsat pour générer une carte de températures de surface a déjà fait l'objet de projets de recherche (Bhang et Seok-Soon, 2009).

Une rencontre avec les auteurs du premier modèle au printemps 2011 (A. Smargiassy et A. Bran) a ouvert la possibilité de tester certaines améliorations possibles du modèle initial, notamment en ce qui concerne certaines variables explicatives telles que les classes d'occupation du sol et certains indices utilisés. Dans ce contexte, il a été décidé de tester *a priori* les mêmes variables explicatives que celles qui ont été retenues dans le modèle de Kestens *et al.* (2011), et de vérifier si l'ajout d'autres variables ciblées permettrait d'améliorer le pouvoir prédictif du modèle. La figure 1 présente les principales étapes suivies lors de l'élaboration du modèle. Chaque étape sera par la suite décrite dans le détail :

- Préparation du jeu de données utilisé pour construire le modèle (variables explicatives et température mesurée au sol);
- Élaboration du nouveau modèle;
- Validation statistique du nouveau modèle;
- Application du modèle sur l'ensemble du territoire d'étude.

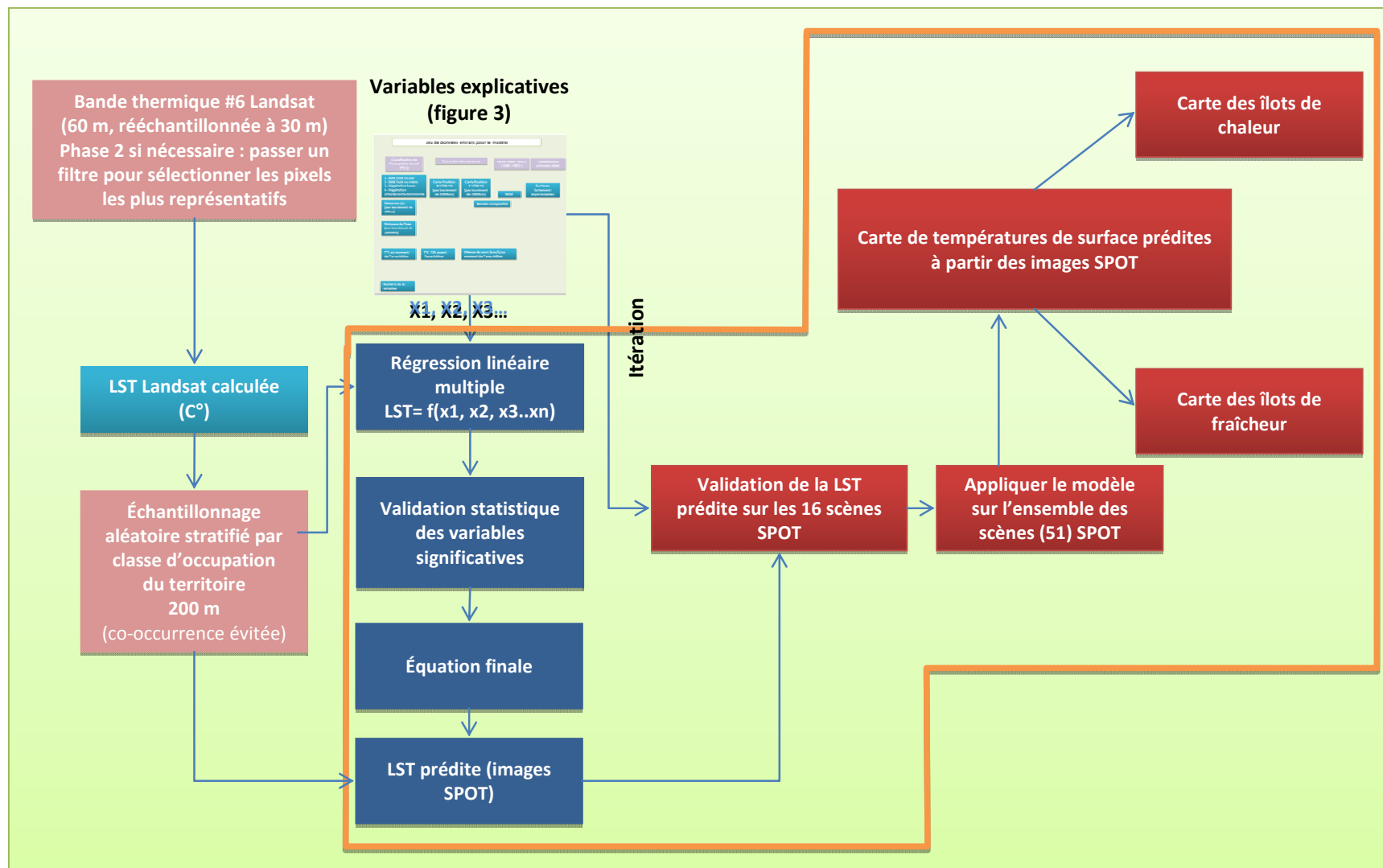


Figure 1. Principales étapes suivies lors de l'élaboration de la carte de températures de surface

5. DONNÉES UTILISÉES POUR LA CONSTRUCTION DU MODÈLE

5.1. SOUS-ÉCHANTILLONNAGE DU TERRITOIRE D'ÉTUDE

Le modèle prédictif a été construit à partir d'un sous-échantillonnage du territoire d'étude. Parmi l'ensemble des images satellitaires couvrant le territoire d'étude, une sélection de secteurs représentant l'ensemble de l'hétérogénéité du tissu urbain retrouvé au Québec a été réalisée (grandes agglomérations urbaines telles que Montréal et Québec, grandes villes régionales telles que Sherbrooke, Rouyn-Noranda et villes de taille plus réduite). Le choix des scènes a également été fait en s'assurant qu'elles couvrent l'ensemble de l'étendue géographique du Québec habité : du nord (Abitibi) au sud (Estrie, Beauce) ainsi que d'ouest (Témiscamingue) en est (Gaspésie) (tableau 1).

Tableau 1. Liste des régions socio-sanitaires pour lesquelles au moins une image a été retenue pour l'élaboration du modèle

Régions socio-sanitaires considérées dans le cadre du projet		Nombre d'images
01	Bas-St-Laurent	1
02	Saguenay-Lac-St-Jean	2
03	Capitale-Nationale (agglomération de Québec)	1
04	Mauricie-Centre du Québec	2
05	Estrie	1
06-13	Agglomérations de Montréal et Laval	1
07	Outaouais	1
08	Abitibi-Témiscamingue	1
09	Côte-Nord	1
11	Gaspésie	2
12	Chaudière-Appalaches	1
14-15	Lanaudière et Laurentides	1
16	Montréal	1

5.2. CHOIX DES DONNÉES SATELLITAIRES UTILISÉES

Deux jeux de données satellitaires ont été utilisés pour construire le modèle de prédiction de la température de surface prédite :

- La bande 6 (thermique) du capteur ETM+ de Landsat;

- Les 4 bandes du capteur optique de SPOT-5 (Bande 1 : Vert (0,50 - 0,59 μm); Bande 2 : Rouge (0,61 - 0,68 μm); Bande 3 : Proche infrarouge (0,78 - 0,89 μm); Bande 4 : Moyen infrarouge (MIR) (1,58 - 1,75 μm).

5.2.1. Images acquises par le capteur ETM+ de Landsat

Quinze images acquises par le capteur ETM+ de Landsat ont été utilisées pour obtenir une estimation de la température réelle de surface. Ces images ont été acquises entre 2005 et 2011 pendant la saison d'été (du 15 juin au 26 août). Toutes les images Landsat ont été prises entre 14h56 et 15h51, soit la période journalière d'acquisition des images Landsat pour le Québec. . La sélection des images Landsat dans les archives disponibles a été réalisée de manière à ce que leur date d'acquisition soit la plus proche possible (en termes de date journalière) de celle des images SPOT-5 correspondantes. Le choix des images Landsat prend aussi en considération la présence éventuelle d'une couverture nuageuse (les images comportant plus de 20 % de nuages ont été exclues). Dans certains cas, il a été nécessaire d'avoir recours à des images prises à des années différentes mais rapprochées le plus possible de l'année de référence. Le tableau 2 présente les dates d'acquisition des images SPOT-5 et Landsat ETM+ correspondantes. À trois reprises, deux images Landsat ont été nécessaires pour couvrir toute l'étendue de l'image SPOT correspondante.

Tableau 2. Dates d'acquisition des images SPOT-5 et des images Landsat correspondantes (couvrant le même secteur)

Date de l'image SPOT-5				Date de l'image Landsat correspondante			Intervalle de température (°C) mesuré sur les images Landsat	
Année	Mois	Jour	id	Année	Mois	Jour	Min	Max
2005	8	26		2005	8	6	17	32
2006	7	19		2006	7	19	11	23
				2007	7	29	10	25
2006	8	30	a	2008	8	21	15	26
2006	8	30	b	2007	7	25	18	26
2007	6	18		2011	6	15	10	27
2007	8	2		2007	8	21	10	23
				2007	8	5	6	24
2007	8	28		2007	8	21	10	23
2007	9	4	a	2007	8	21	10	23
2007	9	4	b	2008	8	21	10	23
2008	7	7		2010	7	23	12	21
				2005	7	17	8	25
2008	8	21		2008	8	26	7	19
2009	8	15		2010	8	20	13	26
2009	9	6		2010	6	17	14	28
2010	6	21		2010	6	17	14	28
2010	9	1		2010	8	20	13	26
2011	7	2		2011	7	10	12	24

5.2.2. Images SPOT-5

L'utilisation d'images SPOT-5 représente l'une des principales améliorations au modèle initial développé par Kestens *et al.* (2011). En effet, le modèle initial a été construit à partir d'images Landsat, générant alors une cartographie des îlots de chaleur selon une résolution de 100 m. En utilisant des images SPOT-5, la résolution de la cartographie issue du modèle passe de 100 à 20 m. Le choix des images SPOT-5 a également été conditionné par la disponibilité de ces dernières, à un moindre coût, contrairement à des images acquises par d'autres capteurs, dont la résolution aurait pu être encore plus fine, mais dont l'acquisition aurait été plus dispendieuse. Au moment de la soumission de la proposition et lors de sa mise en place, les images SPOT-5 s'avéraient le choix le plus judicieux.

Seize images acquises par le capteur optique de SPOT-5 ont été utilisées pour caractériser certaines variables explicatives du modèle, notamment la présence de végétation. Les images SPOT-5 retenues ont donc été acquises pendant la période où la végétation est en feuilles, entre le 18 juin et le 6 septembre (tableau 2). Le choix des images prend aussi en considération la présence éventuelle d'une couverture nuageuse (les images comportant plus de 20 % de nuages ont été exclues). L'ensemble des plus récentes données disponibles a permis de couvrir les années 2005 à 2011. Les images SPOT-5 n'ont pas fait l'objet d'une correction atmosphérique. Toutefois, une attention particulière était portée au choix d'images acquises dans des conditions atmosphériques favorables. Le manque d'information sur les nombreux paramètres atmosphériques lors de l'acquisition des images limitait l'application d'une correction atmosphérique uniforme pour toutes les images.

5.3. DESCRIPTION DE LA VARIABLE DÉPENDANTE DU MODÈLE

5.3.1. Méthode de calcul de la température de surface

La variable dépendante du modèle est la température de la surface. Cette dernière est calculée à partir de données issues d'une bande thermique, plus spécifiquement la bande 6 du capteur Landsat ETM+. Cette bande thermique est sensible à la radiation émise dans l'intervalle des longueurs d'onde de 10,4-12,5 μm , où l'effet de l'atmosphère est minimal (Bartolucci *et al.*, 1988). L'amélioration de la résolution spatiale à 60 m et de la sensibilité radiométrique sont parmi les avantages de ce capteur. Les paramètres de calibrage pour calculer la radiance sont aussi périodiquement mis à jour et corrigés.

Le choix d'utiliser les images prises par le capteur ETM+ de Landsat repose sur la décision d'utiliser le même cadre méthodologique que le modèle développé par Kestens *et al.* (2011). De plus, les images SPOT-5 ne disposent pas de bandes thermiques.

Les pixels des images ont été convertis en des unités de radiance absolue selon l'équation suivante :

$$\text{Radiance} = \text{Gain} * \text{CN (comptes numériques)} + \text{compensation}$$

Pour les calculs, nous avons utilisé l'expression équivalente :

$$\text{Radiance} = (\text{LMax} - \text{LMin}/\text{QCALMax} - \text{QCALMin}) \times (\text{QCAL} - \text{QCALMin}) + \text{LMin}$$

où :

$$\text{QCALMin} = 1;$$

$$\text{QCALMax} = 255;$$

$$\text{QCAL} = \text{CN};$$

$$\text{Lmin} = 0 \text{ (BRR)}, 3,2 \text{ (HRR)};$$

$$\text{Lmax} = 17,04 \text{ (BRR)}, 12,65 \text{ (HRR)};$$

LMin et LMax sont les paramètres d'étalonnage radiométrique (NASA, 2000) aux CN 1 et 255 (i.e. QCALMin, QCALMax).

La radiance spectrale a ensuite été convertie en la température effective du système terre-atmosphère, avec l'hypothèse d'une émissivité unitaire, et en utilisant les constantes du calibrage pré-lancement (NASA, 2000). La formule de conversion est la suivante :

$$T = (k_2 / \ln(k_1/L_\lambda)) + 1$$

où :

T = température effective au capteur en Kelvins;

$$k_2 = 666,09;$$

$$k_1 = 1\,282,71;$$

L_λ = radiance spectrale en $\text{W m}^{-2} \text{ sr}^{-1} \mu\text{m}^{-1}$.

La figure 2 présente un exemple de températures calculées à partir de l'image Landsat ETM+ pour la région de Québec.

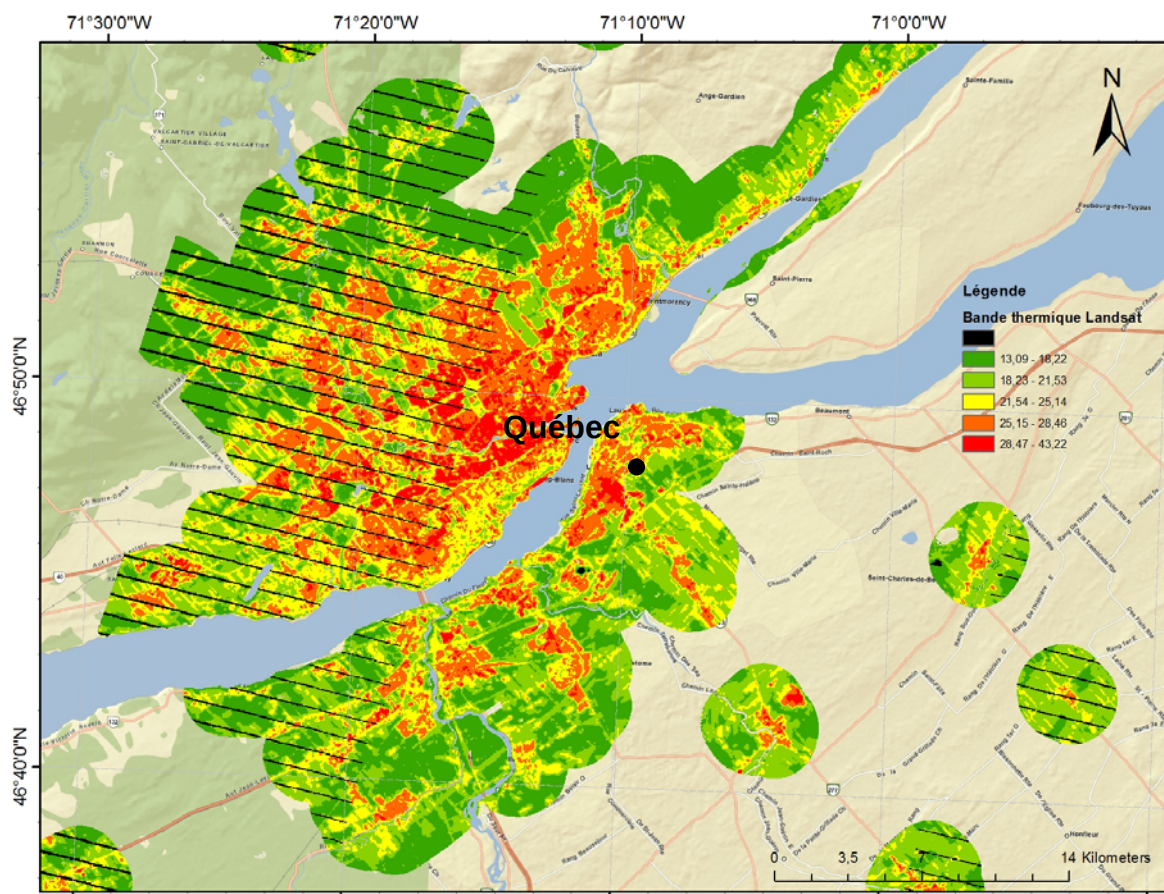


Figure 2. Exemple de carte des classes de température de surface calculées à partir de la bande thermique du capteur ETM+ de Landsat pour la région de Québec

5.3.2. Absence de normalisation des données de température

Il a été décidé de ne pas normaliser la température issue des images provenant du capteur ETM+ de Landsat. Cet exercice aurait permis d'uniformiser la plage des valeurs de température entre chaque image retenue et ces valeurs normalisées auraient ensuite pu être utilisées comme variable dépendante pour construire le modèle. Une telle approche change le paradigme suivi par Kestens *et al.* (2011) où le modèle prédit la température de surface. En effet, la normalisation des valeurs de température implique que chaque valeur de température au sol issue du calcul de la bande 6 du capteur ETM+ est normalisée uniquement selon les températures maximales et minimales. Trois difficultés principales sont à prévoir si une telle approche est suivie.

Premièrement, il est reconnu que la distribution des températures de surface est fort dépendante du contexte météorologique lors de la prise d'image (valeur de température maximale, humidité relative, conditions de vent, moment durant la saison de croissance, etc.).

Deuxièmement, la normalisation des températures de surface est aussi affectée par les conditions physiques dans les milieux urbains. Une zone très urbanisée comme le centre-ville de Montréal diffusera la température d'une manière différente de celle d'une zone de banlieue à Gatineau par exemple. Lors du développement d'un modèle prédictif sur des champs de température connus, cela implique de mettre en relation les températures réelles avec un ensemble de variables sélectionnées. Toutefois, lorsqu'on met en relation des champs normalisés de températures maximales et minimales, on doit procéder à la remise en question des variables à inclure dans le modèle. Cette considération nous éloignait de l'approche choisie qui consistait à utiliser *a priori* les mêmes variables explicatives que celles provenant de Kestens *et al.* (2011).

Une troisième difficulté est aussi à prévoir en lien avec la normalisation pour l'ensemble des images Landsat acquises. Il est en effet reconnu, dans la mise en place des procédures de normalisation pour la production d'une mosaïque, que plus le territoire à couvrir est grand, plus les chances sont grandes de perdre de la dynamique des valeurs de l'image. C'est pourquoi la mise en place d'un ensemble normalisé requiert l'usage d'une image initiale de référence procurant une bonne dynamique des valeurs, afin de procurer une base pour la normalisation des seize autres images Landsat.

Ce travail de normalisation de l'ensemble des dix-sept images Landsat, quoiqu'attractif dans le cadre d'un développement scientifique, aurait certainement impliqué de nombreux essais et aussi des ajustements de procédures, nous amenant bien au-delà des limites budgétaires du projet. De plus, sachant que la normalisation peut dégrader l'information concernant la température calculée à partir des images Landsat et influencer alors la sélection des variables explicatives qui sont significatives, **il a été décidé de ne pas normaliser la donnée de température du sol pour toutes les images Landsat mais plutôt d'utiliser les valeurs de température calculées à partir du canal thermique pour chaque image Landsat.** La discussion à ce sujet sera reprise dans la section 10 de ce rapport sur les améliorations suggérées.

5.3.3. Réduction de la dimension des pixels des images Landsat

Comme les pixels de la bande thermique de l'image prise par le capteur ETM+ de Landsat ont une résolution de 60 m et ceux de l'image SPOT-5, 20 m, les pixels de l'image Landsat ont été scindés en quatre pixels de 30 m, de manière à traiter des pixels de taille à peu près similaire pour les 2 types d'image utilisés.

5.4. DESCRIPTION DES VARIABLES EXPLICATIVES

Les variables explicatives retenues pour l'élaboration du nouveau modèle prédictif sont, dans un premier temps, celles qui figuraient déjà dans le modèle développé par Kestens *et al.* (2011) :

- des données climatiques;
- l'indice de végétation normalisé (NDVI);
- la distance par rapport à l'eau;
- le positionnement en X et Y;
- le numéro de semaine d'acquisition de l'image Landsat.

De nouvelles variables ont été ajoutées :

- une classification de l'occupation du sol, passant de 2 à 4 classes;
- un indice d'imperméabilité des surfaces (IS);
- la hauteur de la surface par rapport au niveau de la mer (MNA).

La figure 3 présente l'ensemble des données qui ont été testées dans le nouveau modèle.

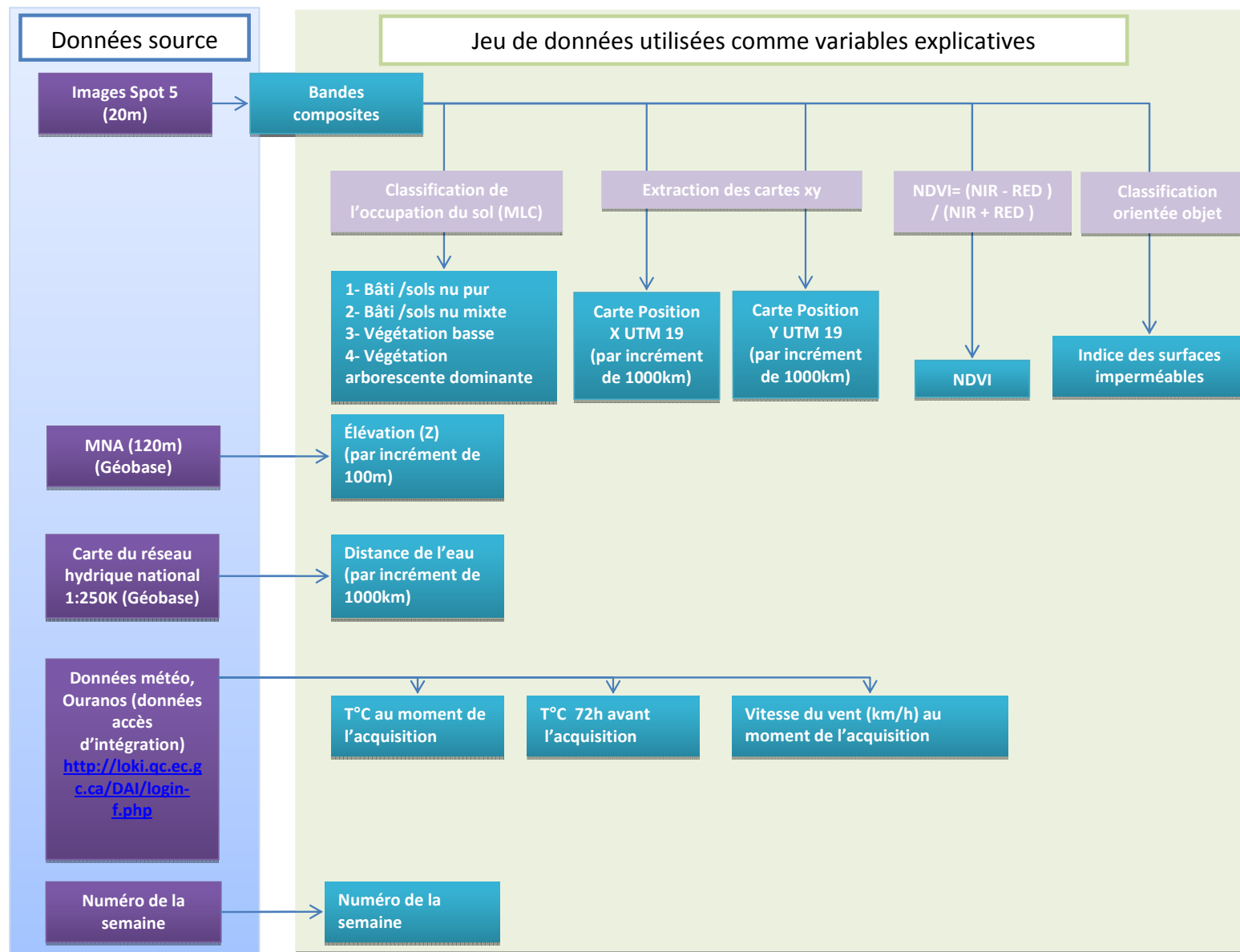


Figure 3. Schéma général du jeu de données utilisées comme variables explicatives du modèle de prédiction de la température de surface à partir des images SPOT-5 (20 m)

5.4.1. Données climatiques

Les données climatiques utilisées sont des valeurs moyennes calculées à partir des données fournies par l'ensemble des stations météorologiques présentes sur chaque scène d'image Landsat. Tous les pixels d'une même image Landsat ont donc les mêmes valeurs de données climatiques.

Les données climatiques ont été fournies par Environnement Canada via le portail Données Accès Intégration (DAI) (<http://loki.qc.ec.gc.ca/DAI/login-f.php>).

Les données intégrées au modèle sont :

- la température moyenne mesurée le jour de l'acquisition de l'image Landsat;
- la température moyenne calculée trois jours avant l'acquisition de l'image;
- la vitesse moyenne journalière du vent calculée le jour de l'acquisition de l'image Landsat.

5.4.2. Indice de végétation normalisé (NDVI)

Le NDVI utilisé pour estimer la quantité de végétation présente est le rapport normalisé de la différence entre la réflectance proche infrarouge et la réflectance acquise dans la bande spectrale rouge du capteur optique de SPOT-5.

$$\text{NDVI} = \frac{\text{proche infrarouge} - \text{rouge}}{\text{proche infrarouge} + \text{rouge}}$$

Les valeurs de NDVI peuvent varier de -1 à +1. En présence de végétation, elles sont généralement comprises entre 0,1 et 0,7. Les valeurs supérieures de l'indice correspondent à la présence d'une couverture végétale dense et en bonne santé, alors que les nuages et la neige entraînent des valeurs de NDVI proches de 0.

5.4.3. Indice d'imperméabilité des surfaces

Le calcul de l'imperméabilité des surfaces informe sur le degré d'anthropisation d'une région, mais aussi sur la quantité de matériaux ayant une forte masse thermique. En effet, les surfaces imperméables, principalement en raison de leur masse élevée, ont une forte capacité à capter et à

stocker l'énergie provenant du rayonnement solaire. Les surfaces imperméables sont principalement des surfaces anthropiques comme les routes, les trottoirs, les cours et allées pavées, ou les stationnements. Elles sont toutes couvertes par de l'asphalte, de la brique, de la pierre ou du ciment. Les toitures, qui sont principalement fabriquées de bardeaux d'asphalte, de tôle ou de cèdre, et le sol nu compacté par le développement urbain sont aussi fortement imperméables.

Le calcul des surfaces imperméables est estimé à partir des données des bandes 2, 3 et 4 de l'image SPOT-5, c'est-à-dire la bande rouge (0,61 - 0,68 μm), la bande du proche infrarouge (0,78 - 0,89 μm) et celle du moyen infrarouge (MIR) (1,58 - 1,75 μm).

Sur chaque scène SPOT-5 à l'étude, une ou plusieurs régions d'intérêt correspondant aux surfaces imperméables dites pures ont été identifiées manuellement en utilisant les données des bandes 2, 3 et 4 ainsi que l'imagerie à haute résolution spatiale disponible dans Google Maps. Ces régions d'intérêt correspondent généralement à des stationnements avec peu ou pas de peinture, à de grandes intersections routières ou à des pistes d'aéroport composées d'asphalte ou de bitume. Les régions d'intérêt ont été reportées dans le logiciel ENVI 4.8 de la compagnie EXEILS et un démixage spectral partiel a ensuite été appliqué sur les scènes, dans le but de déterminer l'abondance des composantes spectrales correspondant aux surfaces imperméables dans chaque pixel. Le démixage est dit partiel car il ne concerne que la fraction de surface imperméable du pixel. Il optimise la réponse des composantes spectrales connues et supprime les valeurs inconnues, selon la signature spectrale de la région d'intérêt. Il fournit un moyen rapide d'identifier les matériaux spécifiques, en se basant sur la similitude des régions d'intérêt préalablement identifiées concernant l'imperméabilité des surfaces.

Les résultats développés par cette méthode ont finalement été appliqués à chaque pixel de la scène. Les valeurs de pixel s'étendent de 0 à 255 et correspondent à la fraction de surface imperméable. Il est possible de transposer ces valeurs pour obtenir un pourcentage d'imperméabilité, et ce, sur chaque pixel de 20 mètres de côté.

La méthode utilisée pour évaluer le pourcentage d'imperméabilité présente certaines limites qu'il est important de connaître. Comme il a été mentionné, cette méthode repose sur l'identification de régions d'intérêt sur chaque scène, qui servent de référence pour le démixage spectral partiel. Une attention particulière a été portée pour obtenir une référence d'imperméabilité maximale pour chacune des scènes et des valeurs spectrales de ces régions d'intérêt idéalement équivalentes d'une scène à l'autre. Mais dans les scènes les moins urbaines où les régions d'intérêt sont plus

petites que 20 mètres de côté, le référent n'est pas aussi pur en surface imperméable que dans les secteurs très urbanisés où l'on retrouve de grandes surfaces imperméables. La résolution spatiale est la cause de cette confusion. Une résolution spatiale plus fine limiterait certainement les différences inter-scènes.

De plus, il s'avère que le choix des bandes spectrales sélectionnées pour l'analyse du démixage spectral fait ressortir certains champs agricoles comme étant fortement imperméables. En effet, les champs labourés, comportant peu de végétation et de matière organique et avec une composition élevée en minéraux, apportent une confusion dans l'analyse. En raison de la confusion concernant les champs agricoles, il est possible que l'analyse statistique des relations entre l'indice d'imperméabilité et les températures de surface ne reflète pas l'efficacité maximale de l'indice pour simuler les températures de surface en milieu urbain.

5.4.4. Classification de l'occupation du sol

La classification de l'occupation du sol a été améliorée par rapport à celle utilisée par Kestens *et al.* (2011). Ainsi, alors que le modèle initial ne distinguait que l'eau, les zones de végétation denses et le reste (sol nu et bâti), le nouveau modèle considère quatre classes d'occupation du sol en plus de l'eau (tableau 3).

Tableau 3. Classes d'occupation du sol et catégories considérées

N°	Classe	Signification
1	Zone urbaine pure	Chaussée, toits de maisons, d'édifices, de grandes surfaces, pistes d'atterrissage ² ...
2	Zone urbaine mixte	Maisons et édifices avec présence de végétation
3	Forêt	Zones végétalisées dominées par les arbres
4	Zone basse végétation	Champs agricoles, parcs (jardins communautaires, aires récréatives...), terrains de sport (golf, soccer, baseball...), terrains ouverts...

La classification de l'occupation du sol a été réalisée sur les images SPOT-5. L'approche utilisée pour la classification du territoire est « le maximum de vraisemblance ». Cette méthode tient compte des fichiers de signature spectrale des bandes composites B2, B3 et B4. Le résultat est

² Il s'est avéré que cette classe inclut, dans certains cas, des champs labourés sans végétation en raison de leur signature spectrale qui s'apparente à celle des zones urbaines pures.

une image-thème sur laquelle se retrouvent les classes définies par les signatures (chaque classe est codée sous une valeur numérique). La classification par maximum de vraisemblance utilise le « seuil gaussien » défini dans chaque fichier signature pour déterminer si un pixel donné fait partie ou non d'une classe. Le seuil correspond au rayon (en unité d'écart-type) d'une hyperellipse et est assigné à la classe. Il est possible d'attribuer une importance différente à une classe, ce qui permet de résoudre le problème de recoupement entre celles-ci en favorisant une classe au dépend d'une autre. Si le pixel n'est associé à aucune classe, il est codé « non classé » et la valeur zéro lui est attribuée. On suppose que les classes ont une distribution gaussienne et que les signatures sont bien sélectionnées.

Afin d'évaluer la précision de la classification réalisée, une matrice de confusion peut être utilisée (Story and Congalton, 1986). Pour chaque classe d'occupation du sol, elle indique le pourcentage de pixels de l'ensemble de référence qui se classe correctement ainsi que la répartition des pixels dont la classe résultante qui diffère de la classe de référence. On peut ainsi analyser les sources de confusion entre les classes.

Pour ce faire, un échantillonnage aléatoire de 30 points de contrôle a été réalisé pour chaque classe d'occupation du sol localisée dans 4 secteurs différents dans des zones de concentration d'au moins 9 pixels identiques et contigus. La validation visuelle a été réalisée avec des photographies aériennes pour chacun des 4 secteurs choisis. L'analyse consistait à examiner l'affichage des erreurs de commission et d'omission et la précision de la classification globale. De ces pourcentages une mesure de performance connue sous le nom de Kappa a été calculée (Bishop et al. 1975).

L'indice de Kappa a été utilisé pour apprécier la précision de la classification adoptée sus-décrite. D'après Pontius (2000), l'indice de Kappa (k) est utilisé pour valider et estimer la précision dans le schéma de classification de l'occupation du sol. Sa formule est la suivante :

$$P_0 - P_c / P_p - P_c \quad (1)$$

où P_0 est égal au pourcentage réel obtenu de classification des éléments de l'occupation du sol; il est égal au quotient de la somme des chiffres de la diagonale de la matrice avec le total du nombre d'observations;

où P_c est l'estimation de la probabilité d'obtenir une classification correcte. Pour calculer P_c , on procède de la façon suivante: on réalise les produits marginaux des valeurs des colonnes et des

rangées au niveau de chaque cellule de la matrice, puis la somme des valeurs de la diagonale est divisée par le total des produits de chaque cellule de la matrice.

Ainsi, la formule (1) peut s'écrire :

$$k = P_0 - P_c / 1 - P_c \quad (2)$$

Les valeurs de Kappa vont de -1 à 1; une valeur de zéro indique que l'effet de classification est à accord égal de chance, et une valeur de 1 indique une classification parfaitement efficace sans une contribution d'un accord de chance. Toute valeur négative indique une mauvaise classification dans laquelle un accord de chance est plus important que l'effet de classification. Par conséquent, une valeur Kappa de 0,75 ou plus indique un très bon classement (Montserud et Leamans, 1992). Le tableau 4 contient la table de contingence complète des 4 classes d'occupation du sol pour l'ensemble des 4 secteurs choisis.

Tableau 4. Matrices de confusion pour chaque classe d'occupation du sol

	Points de contrôle pour la classe : zone urbaine pure			Précision (%)
	zone urbaine pure	Non-zone urbaine pure	Total	
zone urbaine pure	98	22	120	82
Non- zone urbaine pure	22	98	120	82
Total	120	120	240	
Erreur d'omission 18 %				
Erreur de confusion 18 %				
Précision totale 82 %				
Kappa 0,63				

	Points de contrôle pour la classe : zone urbaine mixte			Précision (%)
	zone urbaine mixte	Non-zone urbaine mixte	Total	
zone urbaine mixte	117	3	120	98
Non-zone urbaine mixte	3	117	120	98
Total	120	120	240	
Erreur d'omission 3 %				
Erreur de confusion 3 %				
Précision totale 98 %				
Kappa 0,95				

Points de contrôle pour la classe forêt				
	Forêt	Non-forêt	Total	Précision (%)
Forêt	102	18	120	85
Non-forêt	18	102	120	85
Total	120	120	240	

Erreur d'omission 15 %
 Erreur de confusion 15 %
 Précision totale 85 %
 Kappa 0,70

Points de contrôle pour la classe zone basse végétation				
	zone basse végétation	Non- zone basse végétation	Total	Précision (%)
Zone basse végétation	118	2	120	98
Non-zone basse végétation	2	118	120	98
Total	120	120	240	

Erreur d'omission 2 %
 Erreur de confusion 2 %
 Précision totale 98 %
 Kappa 0,97

La précision moyenne de la classification globale était de 91% pour l'ensemble des 4 secteurs choisis dans la province de Québec. Selon les résultats de classification, la classe : zone basse végétation a une précision de classification moyenne de 98%, avec seulement deux points mal classés, tandis que la classe : zone urbaine pure présente la faible précision 82%, où 22 points des 120 échantillons ont été mal classés. Ceci peut être expliqué par la résolution spatiale des données SPOT-5 (20 m), où dans une zone urbaine, un pixel classé pure peut contenir d'autres entités hétérogènes (pelouse, arbre etc.) qui n'occupe qu'une partie du pixel classé.

En utilisant l'indice de Kappa, dont la valeur moyenne est de 0,81, on peut conclure que les résultats de cette analyse sont statistiquement très satisfaisants.

5.5. ÉCHANTILLONNAGE

Un échantillonnage aléatoire stratifié dans chacune des 16 scènes SPOT-5 a été réalisé en considérant une distance minimale de 200 m entre 2 pixels échantillonnés, de manière à éviter la cooccurrence des points d'échantillonnage. Le but était de s'assurer que ne soient pas retenus 2

pixels qui seraient associés au même pixel initial provenant de l'image Landsat, dont la taille est de 60 m. L'échantillonnage a été stratifié par classe d'occupation du sol. Finalement, 35 099 pixels ont été retenus pour l'élaboration du modèle de température de surface.

6. ÉLABORATION DU MODÈLE

L'ensemble des analyses statistiques a été réalisé à l'aide du logiciel SAS 9.3.

6.1. PORTRAIT GÉNÉRAL DES VARIABLES UTILISÉES

Le portrait général des variables (moyenne, écart-type, minimum, maximum) pour les 35 099 pixels échantillonnés est présenté au tableau 5.

Tableau 5. Portrait général des variables

Variables ³	ID	Moyenne	Écart-type	Minimum	Maximum
Température de surface	LST	22.9	4.131	13.1	41.3
Distance à l'eau (x1000)	D_EAU1000	23.4	29.103	0	114
Altitude	MNA_250K	80.7	77.857	0	566
NDVI	NDVI	0.3	0.241	-0.7228	0.77361
Indice d'imperméabilité	PCT_IMPER	48.6	29.067	0	100
Température moyenne de l'air	T	16.6	2.193	9	25
Moyenne des températures 72h	T72h	17.6	3.905	9	22
Vitesse du vent	VENT	8.9	2.784	4	22
X (x1000)	X_1000	224.2	137.571	-181	840.771
Y (x10000)	Y_10000	512.3	11.333	502.1321	557.4321
Classes d'occupation du sol	cl_occ_sol			1	4

6.2. VÉRIFICATION DE LA COLINÉARITÉ ENTRE LES VARIABLES

La colinéarité entre les variables a été évaluée via une matrice de corrélation de Pearson pour les variables continues (tableau 6).

La valeur seuil couramment utilisée pour établir la présence d'une colinéarité est un $\rho > 0.70$ (soit un $R^2 > 0.50$). Seuls le NDVI et l'indice d'imperméabilité des surfaces (IS) dépassent ce seuil. Cette corrélation était attendue puisqu'ils reflètent tous les deux un certain degré d'urbanisation. En effet, on s'attend à de fortes valeurs de NDVI et à de faibles valeurs d'IS en région peu urbanisée, et au contraire en milieu urbain. Par contre, plusieurs auteurs rapportent que la relation entre la température de surface et le NDVI est non linéaire (Yuan and Bauer, 2007). En effet, lorsque les valeurs de NDVI sont inférieures à 0,6 (moyen à faible), la relation

³ L'échelle de certaines variables a été corrigée afin que les valeurs se situent entre 1 et 1000. Le facteur de correction est affiché entre parenthèses.

entre le NDVI et la température de surface serait non significative. Sachant que l'on retrouve les valeurs de moins de 0,6 dans les secteurs plus urbanisés, il est proposé d'inclure dans le modèle la variable de l'indice d'imperméabilité, qui apporterait une contribution au modèle dans les cas où le NDVI ne serait plus significatif.

Enfin, la vérification d'interactions potentielles entre l'indice d'imperméabilité et les autres variables oblige à considérer cet indice dans l'équation comme variable simple du modèle.

Dans le cas des classes d'occupation du sol, comme il s'agit de variables catégoriques, la colinéarité a été testée par analyse de variance avec chacune des variables continues. Après évaluation, aucune variable ne présentait une relation ayant un $R^2 > 0,50$. Les deux variables les plus corrélées à la classification de l'occupation du sol sont le NDVI (0,49) et le pourcentage d'imperméabilité (0,48).

Tableau 6. Présentation des valeurs de rho pour évaluer la colinéarité

Variable	LST	NDVI	MNA_250K	D_EAU1000	X_1000	Y_10000	PCT_IMPER	T	T72H	VENT
LST	1.00									
NDVI	-0.56	1.00								
MNA_250K	-0.27	0.14	1.00							
D_EAU1000	-0.10	0.18	0.55	1.00						
X_1000	-0.12	0.12	0.23	-0.19	1.00					
Y_10000	-0.01	0.08	0.18	-0.24	0.74	1.00				
PCT_IMPER	0.61	-0.79	-0.01	-0.01	0.05	0.05	1.00			
T	0.27	0.07	-0.38	-0.34	-0.21	0.13	-0.07	1.00		
T72H	0.25	0.02	-0.59	-0.17	-0.61	-0.44	-0.11	0.66	1.00	
VENT	0.24	0.16	-0.42	-0.31	0.11	0.21	-0.09	0.47	0.49	1.00

6.3. CONSTRUCTION DU MODÈLE

La modélisation a été faite selon une approche par modèle mixte où la région de provenance des images et le numéro de la semaine de prise des images LANDSAT ont été utilisés comme effets aléatoires.

La structure du modèle complet testé est :

LST= NDVI + pct_imper + MNA_250K + D_EAU1000 + X_1000 + Y_10000 + CL_OCC_SOL + T + T72h + VENT + NDVI*CL_OCC_SOL + pct_imper*CL_OCC_SOL + NDVI*pct_imper + NDVI*pct_imper*CL_OCC_SOL

RANDOM= REGION NS

Dans ce modèle, seule la variable VENT n'est pas significative et a été retirée du modèle final. Les coefficients associés à chaque variable dans le modèle final (Estimate) sont les suivants :

Effets fixes	CL_OCC_SOL	Estimate	StdErr	DF	tValue	Probt
Intercept		116.7354	6.8289	112	17.09	<0.0001
NDVI		-4.9739	0.3458	31494	-14.38	<0.0001
PCT IMPER		0.0493	0.0032	31535	15.41	<0.0001
MNA_250K		-0.0022	0.0004	31538	-5.97	<0.0001
D_EAU1000		-0.0110	0.0015	31281	-7.43	<0.0001
X_1000		-0.0036	0.0009	19723	-3.86	0.0001
Y_10000		-0.1886	0.0104	27998	-18.22	<0.0001
T		-0.9242	0.0579	10753	-15.98	<0.0001
T72h		1.3560	0.0491	27133	27.60	<0.0001
NDVI*PCT IMPER		-0.0564	0.0065	31534	-8.72	<0.0001
CL_OCC_SOL	1	2.6286	0.2295	31532	11.45	<0.0001
CL_OCC_SOL	2	2.0218	0.2948	31531	6.86	<0.0001
CL_OCC_SOL	3	-0.8636	0.3563	31537	-2.42	0.0154
CL_OCC_SOL	4	0				
NDVI*CL_OCC_SOL	1	-7.6396	0.4606	31528	-16.59	<0.0001
NDVI*CL_OCC_SOL	2	-2.3440	0.6589	31529	-3.56	0.0004
NDVI*CL_OCC_SOL	3	-0.8202	0.6267	31535	-1.31	0.1906
NDVI*CL_OCC_SOL	4	0				
PCT IMPER*CL_OCC_SOL	1	-0.0286	0.0036	31532	-7.96	<0.0001
PCT IMPER*CL_OCC_SOL	2	-0.0254	0.0040	31530	-6.32	<0.0001
PCT IMPER*CL_OCC_SOL	3	-0.0443	0.0077	31531	-5.75	<0.0001
PCT IMPER*CL_OCC_SOL	4	0				
NDVI*PCT IMP*CL_OCC	1	0.0989	0.0081	31532	12.25	<0.0001
NDVI*PCT IMP*CL_OCC	2	0.0279	0.0092	31530	3.04	0.0024
NDVI*PCT IMP*CL_OCC	3	0.0543	0.0153	31533	3.56	0.0004
NDVI*PCT IMP*CL_OCC	4	0				

Effets aléatoires	Région	NS	Estimate	StdErrPred	DF	tValue	Probt
region	1		-1.7505	4	9.45	-0.45	0.6617
region	2		4.0749	1	8.07	2.73	0.0257
region	3		-4.4579	2	8.65	-2.93	0.0173
region	4		1.5125	1	7.76	1.02	0.3375
region	5		-2.3207	1	7.86	-1.56	0.1572
region	6		2.8561	1	7.80	1.93	0.0910
region	7		-2.2767	3	8.46	-0.82	0.4350
region	8		-0.8061	4	9.44	-0.21	0.8396
region	9		-4.6026	2	9.36	-2.97	0.0151
region	10		3.9453	3	8.47	1.42	0.1921
region	11		3.8257	1	7.81	2.58	0.0332
NS		24	13.5111	3	6.47	4.03	0.0059
NS		27	1.0796	3	6.43	0.32	0.7573
NS		29	7.5830	4	10.75	1.83	0.0957
NS		31	-7.6867	3	6.33	-2.31	0.0583
NS		33	-7.9548	5	11.56	-1.69	0.1175
NS		34	-2.8692	3	6.41	-0.86	0.4217
NS		35	-3.6631	5	11.72	-0.78	0.4534

Parmi l'ensemble des variables explicatives testées, le NDVI et l'occupation du sol sont les paramètres qui semblent jouer un rôle majeur dans la prédiction de la température de surface, comme le montre la valeur de leur coefficient (champ Estimate).

L'homoscédasticité et la normalité des résidus ont été évaluées de manière graphique et aucune variable n'aurait bénéficié d'une transformation.

Le modèle obtenu a un pseudo R^2 (coefficient de détermination ou degré d'ajustement du modèle entre les valeurs prédites et observées) de 0,67 et un RMSE (erreur quadratique moyenne) de 2,35°C soit 9,6 %.

7. VALIDATION DU MODÈLE

Le modèle théorique obtenu a été validé à partir de points d'échantillonnage non utilisés pour la construction du modèle (figure 4). Pour cela, des valeurs prédites ont été obtenues par itération pour l'ensemble des données de l'échantillon. À chaque itération, 90 % des données ont servi à prédire le 10 % restant utilisé pour la validation. La validation du modèle montre également un coefficient de corrélation (R^2) de 0,67 entre les valeurs observées et prédites. Quant à l'erreur quadratique moyenne (RMSE) de validation, elle est de 2,36°C soit 10,1 %. La relation linéaire entre les valeurs observées et prédites présente une pente de 1 et une ordonnée à l'origine qui n'est pas significativement différente de zéro.

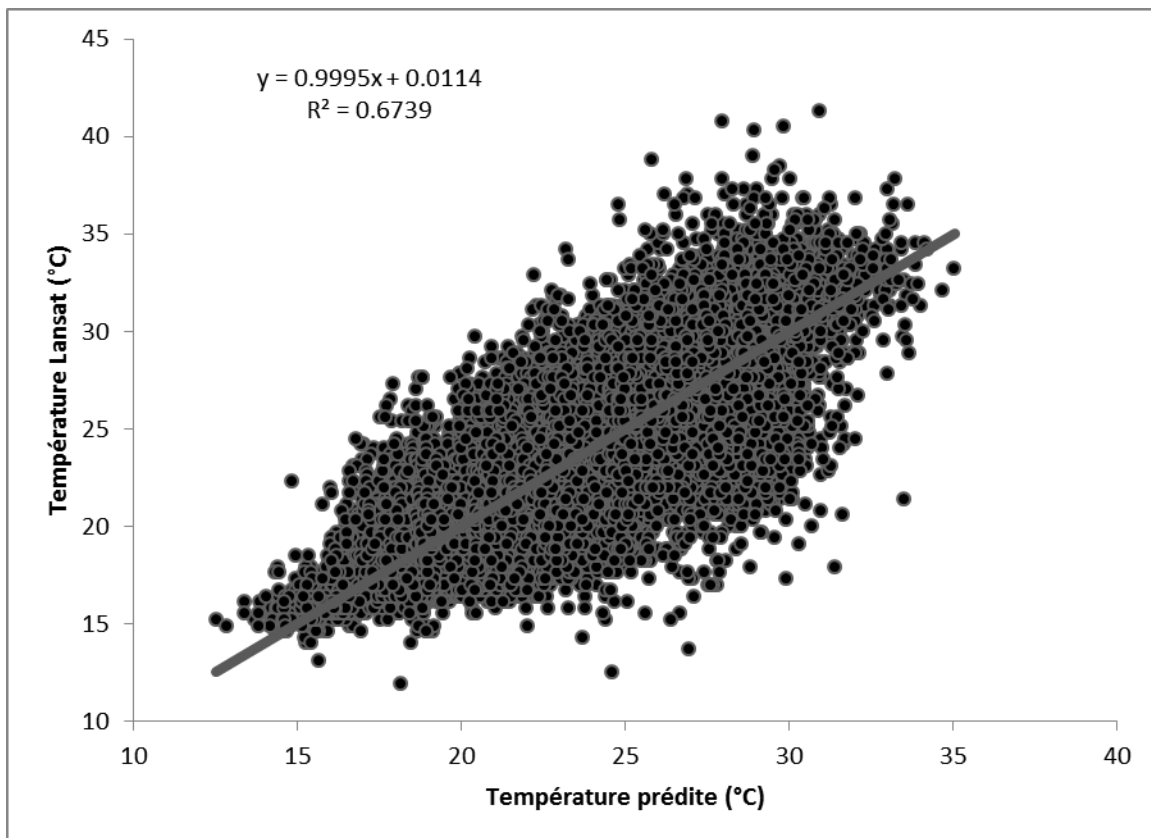


Figure 4. Température Landsat en fonction de la température prédite pour les 3 509 points d'échantillonnage non utilisés lors de la construction du modèle

8. APPLICATION DU MODÈLE À L'ENSEMBLE DU TERRITOIRE D'ÉTUDE

8.1. DONNÉES SPOT-5 UTILISÉES

Soixante-sept images SPOT-5 ont été utilisées pour couvrir l'ensemble du territoire d'étude. Vingt-six proviennent de la banque d'archive gratuite et 41 ont été achetées dans le cadre du présent projet. Tout comme les images utilisées pour la construction du modèle, les images SPOT-5 retenues ont été acquises pendant la période où la végétation est en feuilles, entre le 15 juin et le 9 septembre. Toutes les images SPOT-5 ont été prises à peu près à la même heure de la journée, puisque la période journalière d'acquisition des images SPOT-5 pour le projet oscille entre 14h50 et 16h06. Le choix des images prend aussi en considération la présence éventuelle d'une couverture nuageuse (les images comportant plus de 20 % de nuages ont été exclues). L'ensemble des plus récentes données disponibles a permis de couvrir les années 2005 à 2011. La grande majorité (85 %) a été acquise en 2007 ou pendant les années suivantes :

- 3 datent de 2005;
- 7 de 2006;
- 18 de 2007;
- 14 de 2008;
- 12 de 2009;
- 7 de 2010;
- 6 de 2011.

Ces images SPOT-5 présentent donc l'avantage d'être beaucoup plus récentes que les images Landsat utilisées lors du modèle initial développé par Kestens *et al.* (2011), dont l'année de prise d'image allait de 1987 à 2002. Considérant que le paysage urbain change très vite, en particulier dans un contexte de développement domiciliaire important, l'utilisation d'images plus récentes constitue une grande amélioration de la cartographie des îlots de chaleur produite.

Le tableau 7 présente la date précise d'acquisition de toutes les images SPOT-5 utilisées. La figure 5 présente, quant à elle, la localisation des images SPOT-5 utilisées pour couvrir tout le territoire d'étude ainsi que l'année de prise de chaque image.

Tableau 7. Date d'acquisition des images SPOT-5 utilisées pour couvrir tout le territoire d'étude et images Landsat correspondantes

Date de la Scène de l'image SPOT-5				Date de l'image Landsat correspondante			Date de la Scène de l'image SPOT-5				Date de l'image Landsat correspondante		
Année	Mois	Jour	id	Année	Mois	Jour	Année	Mois	Jour	id	Année	Mois	Jour
2005	8	4		2007	8	21	2008	8	21	_b	2010	6	17
2005	8	26		2005	8	6	2008	8	21		2008	8	26
2005	8	26	_b	2010	8	20	2008	9	16		2007	8	21
2006	7	19		2006	7	19	2008	9	22		2008	8	21
				2007	7	29	2008	6	9	_a	2007	7	29
2006	7	29		2005	8	6	2008	6	9	_b	2007	8	21
2006	7	30		2010	7	23	2008	7	7	_b	2070	8	28
2006	8	21		2007	7	29	2008	9	1	_a	2007	7	25
2006	8	30		2008	8	21	2008	9	1	_b	2006	7	19
2006	8	30	_b	2007	7	25	2008	9	1	_c	2006	7	19
2006	9	14		2007	7	29	2009	8	11		2008	8	21
2006	8	20		2006	7	19	2009	8	15		2010	8	20
2007	6	14		2005	7	18	2009	9	1		2008	8	21
2007	6	17		2008	8	26	2009	9	12		2007	7	29
2007	6	18		2011	6	15	2009	9	26		2007	7	25
2007	6	23		2008	8	21	2009	6	27	_a	2008	8	26
2007	7	23		2007	7	25	2009	6	27	_b	2008	8	26
2007	8	2		2007	8	21	2009	6	27	_c	2008	8	26
				2007	8	5	2009	8	15		2005	8	6
2007	8	10		2011	7	10	2009	9	6		2010	6	17
2007	8	11		2008	8	26	2009	9	6	_a	2008	8	21
2007	8	28		2007	8	21	2009	9	6	_b	2009	9	9
2007	8	29		2007	8	21	2010	6	21		2010	6	17
2007	8	30		2010	6	17	2010	8	19		2007	8	21
2007	9	4		2007	8	21	2010	8	30		2008	8	21
2007	9	18		2007	8	28	2010	9	1		2010	8	20
2007	9	24		2007	8	5	2010	8	29	_a	2007	8	21
2007	7	24	_a	2008	8	21	2010	8	29	_b	2007	8	21
2007	7	24	_b	2006	7	19	2010	8	29	_c	2010	8	20
2007	8	1		2008	8	26	2011	6	15		2010	8	20
2007	9	4	_b	2008	8	21	2011	7	2		2011	7	10
2008	5	14		2007	8	21	2011	8	19		2011	7	10
2008	7	7		2005	7	18	2011	8	26		2010	6	17
				2010	7	23	2011	7	2	_b	2007	8	21
2008	7	17		2007	8	21	2011	7	2	_c	2007	7	29
2008	8	8		2011	6	15							

Notes : 1- À trois reprises, deux images Landsat ont été nécessaires pour couvrir une même scène SPOT-5 (lignes ombragées).

2- Dans certains cas, plusieurs images SPOT-5, localisées à des endroits différents, ont été acquises à la même date (mention d'un id a, b ou c).

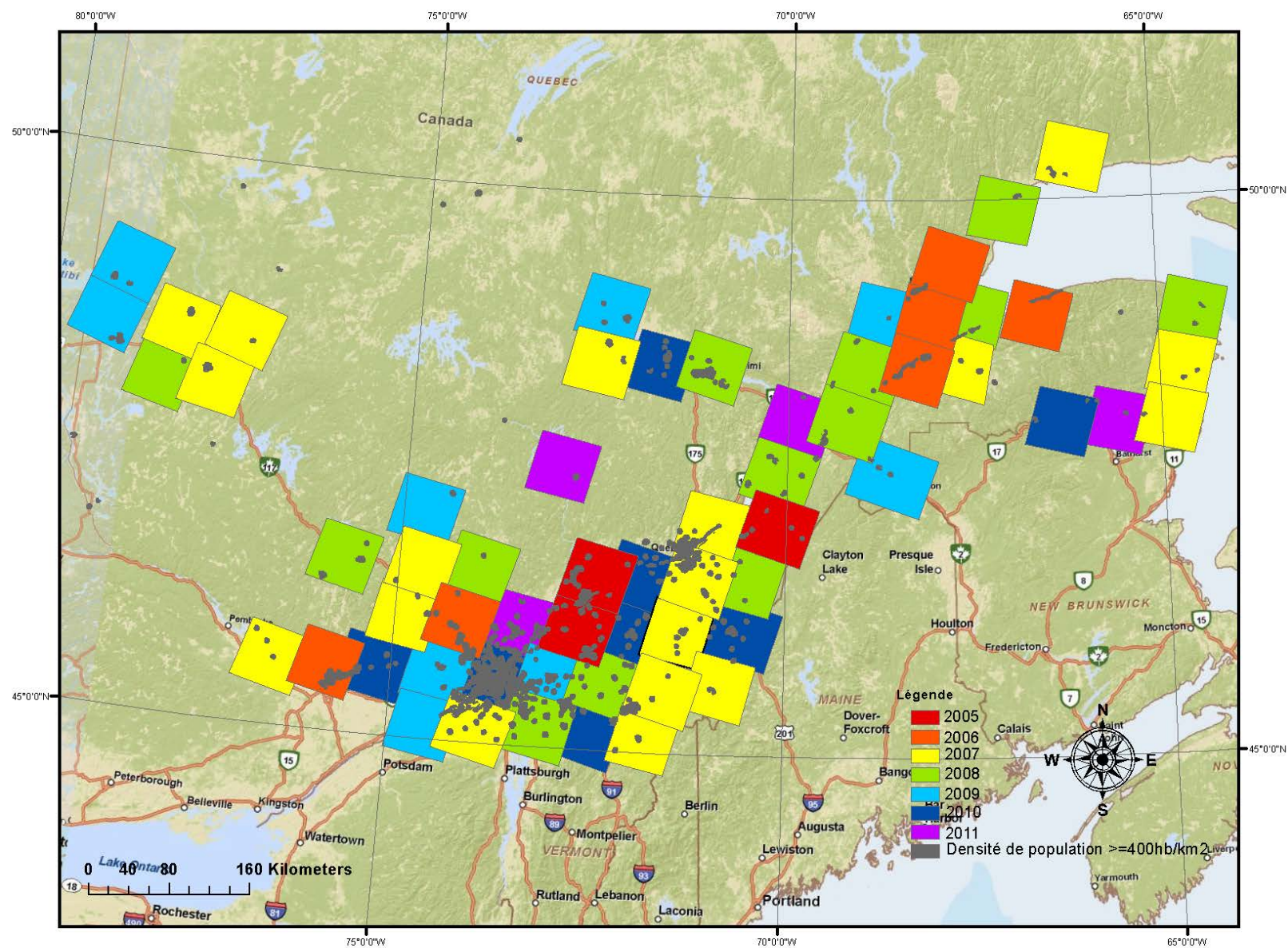


Figure 5. Localisation des images SPOT-5 utilisées pour couvrir tout le territoire d'étude et année de prise de chaque image

8.2. RÉALISATION DE LA CLASSIFICATION DE L'OCCUPATION DU SOL

8.2.1. Construction préalable de mosaïques d'images SPOT-5

Préalablement à l'étape de classification de l'occupation du sol, une correction radiométrique a été appliquée pour la mise en place de mosaïques d'images SPOT-5 dans la portion du territoire d'étude où des images se chevauchent.

La normalisation des images a été effectuée à l'aide du logiciel PCI Geomatica et du module Focus (*raster calculator*). Pour se faire, une image de référence a été choisie au préalable, soit celle qui montrait la plus grande dynamique radiométrique. Dans la zone commune à 2 images qui se chevauchaient, les statistiques (moyenne et écart-type) de la zone de recouvrement ont été calculées pour chacune des bandes, à l'aide de l'algorithme HIS (*Histogramming Database Image*).

La mosaïque des images normalisées a été réalisée avec le module OrthoEngine de PCI Geomatica. Ce module intègre 2 techniques de mosaïquage (*Minimum Difference* et *Minimum Relative Difference*) et permet d'optimiser la ligne de coupure des images. Quand l'opération automatique ne donnait pas un bon résultat, elle était réalisée manuellement, tout en gardant les mêmes options que la méthode automatique. Afin d'ajuster le contraste des images, la méthode d'égalisation d'histogramme, ou *Histogram Matching*, a été appliquée.

L'ensemble des images traitées a donc permis de créer 12 mosaïques pour des régions plus urbanisées (par ex. Montréal, Québec, Sherbrooke, Gatineau) et 10 images indépendantes SPOT-5 pour des villes géographiquement isolées (par ex. Rouyn-Noranda, Gaspé, Sept-Îles). Par la suite, la classification de l'occupation du sol a été réalisée à l'échelle de chaque mosaïque et des scènes isolées.

8.3. PRODUCTION D'UNE CARTE DE TEMPÉRATURES RELATIVES DE SURFACE

Pour chaque pixel des portions de scènes SPOT-5 couvrant le territoire d'étude, l'ensemble des variables explicatives ont été calculées. Le modèle développé a ensuite été appliqué à chaque pixel et une cartographie de la température de surface prédite a ainsi été générée scène par scène. La carte résultante de la température relative de surface permet de localiser, avec une précision

aux 20 m près, les zones les plus chaudes (îlots de chaleur) et les zones les plus froides (îlots de fraîcheur) en milieu habité.

Mise en garde :

Étant donné que l'ensemble des images SPOT-5 et Landsat utilisées pour couvrir le territoire d'étude au complet ont été prises à des dates différentes, les conditions climatiques des journées de prise d'images étaient variables. Par conséquent, l'étendue des températures prédites au sein d'une même image varie d'une image à l'autre. On peut citer l'exemple de 2 scènes mitoyennes, une première prise une journée fraîche générant des températures de surface variant de X à Y et une deuxième prise une journée de forte chaleur générant des températures variant de Y à Z.

Il est donc essentiel de ne pas considérer la température prédite à l'échelle d'un pixel comme une valeur absolue, qui peut être comparée à celle de tous les autres pixels du territoire d'étude. La donnée de température prédite doit impérativement être considérée comme une VALEUR RELATIVE calculée au sein d'une même scène, permettant de distinguer les zones potentiellement les plus chaudes des zones potentiellement les plus fraîches.

8.4. DESCRIPTION DES PLAGES DE TEMPÉRATURES

8.4.1. Méthode utilisée

À l'échelle de chaque scène, neuf plages de températures relatives ont été définies, dans le but de distinguer les zones les plus fraîches (îlots de fraîcheur) des zones les plus chaudes (îlots de chaleur). Le seuillage utilise la méthode de Jenks d'ArcGIS, qui est une méthode inductive consistant à choisir les seuils qui maximisent la variance interclasse et minimisent la variance intraclasse. Si la distribution comporte des discontinuités (zones de faible densité des valeurs), celles-ci seront nécessairement détectées par la méthode de Jenks, qui constitue une variante rigoureuse de la méthode bien connue des « seuils naturels ». Notons que plusieurs tests de nombres de plages ont été réalisés, de manière à comparer les éléments que l'on retrouve dans chaque plage. Le contenu de chaque plage de température provient d'une analyse visuelle de l'occupation du sol utilisant des images de haute résolution (Google Earth). Ces tests ont permis de confirmer que l'utilisation de 9 niveaux était optimale pour définir les zones les plus chaudes correspondant à une concentration élevée de surfaces imperméables et recoupant les zones identifiées comme des îlots de chaleur dans l'analyse de Kestens *et al.* (2011), mais avec une

précision plus fine (20 m). Avec un découpage en neuf plages, il est également apparu qu'il était possible de distinguer, parmi les niveaux les plus frais, les zones boisées des zones végétalisées plus ouvertes.

Ainsi, les neuf niveaux de température obtenus se résument ainsi :

Niveaux de température	Description
1 et 2 regroupés	Îlots de fraîcheur les plus frais (généralement les zones boisées)
3 et 4 regroupés	Îlots de fraîcheur moins frais (généralement des zones végétalisées plus ouvertes)
5, 6 et 7	Continuum de températures sans qualificatif particulier
8	Îlots de chaleur un peu moins chauds (souvent mitoyens aux îlots de niveau 9)
9	Îlots de chaleur les plus chauds

D'autres méthodes de seuillage ont également été essayées (écart-type, intervalle égale, intervalle géométrique mais n'ont pas été aussi concluantes que la méthode de Jenks.

Un filtre de 3 pixels par 3 pixels (ou 60 m X 60 m) a été appliqué aux pixels épars des classes 8 et 9, de manière à ne garder que les concentrations d'îlots de chaleur de plus de 3 600 m².

8.4.2. Mise en garde concernant les terres agricoles nues et l'eau

Après vérification de la classification obtenue, il s'est avéré qu'il existait une confusion entre d'une part les champs agricoles nus et les îlots de chaleur, et d'autre part les étendues d'eau et les îlots de chaleur. La température prédite pour les terres agricoles nues et l'eau s'apparente en effet à celle des îlots de chaleur, de par leur signature spectrale qui présente de fortes similitudes avec les zones hautement imperméables. Ainsi, en zone urbaine, il est fort possible que certains jardins communautaires soient identifiés comme des îlots de chaleur, en particulier si l'image SPOT-5 a été prise lorsque les cultures n'étaient pas encore présentes. Enfin, il convient d'exclure de la cartographie des îlots de chaleur toutes étendues d'eau (bassins, étangs, rivières, etc.), pour éviter leur confusion avec des îlots de chaleur.

8.4.3. Essai de normalisation des valeurs de températures prédites

Un essai de normalisation des températures prédites a été réalisé, en posant l'hypothèse que cette transformation de la donnée brute permettrait de diminuer les écarts de températures observés entre les différentes images. Ces écarts s'expliqueraient par des conditions de prises d'images satellitaires qui peuvent être très variables d'une image à l'autre (ex : une image prise en juillet en période de chaleur et une autre image prise en août en période fraîche)

Trois méthodes de normalisation ont été testées, soit les méthodes STD (Écart type), MIDRANGE (étendue/2) et MAD (Déviation absolue de la médiane) de la procédure STDIZE du logiciel d'analyse statistique SAS. Les tests réalisés sur 2 images de la région de Québec ont donnés des résultats très similaires et ne se sont pas avérés concluants comme le montre l'exemple présenté à la figure 6. En effet, les plages de températures des valeurs normalisées de ces 2 images présentaient des différences, faisant en sorte que les valeurs des pixels situés à la limite commune de ces 2 images n'étaient pas associées aux mêmes plages de températures... Une telle situation n'était par contre pas présente en réalisant le seuillage sur les valeurs brutes de températures image par image. Suite à ce constat, il a été décidé de ne pas normaliser les données de températures prédites.

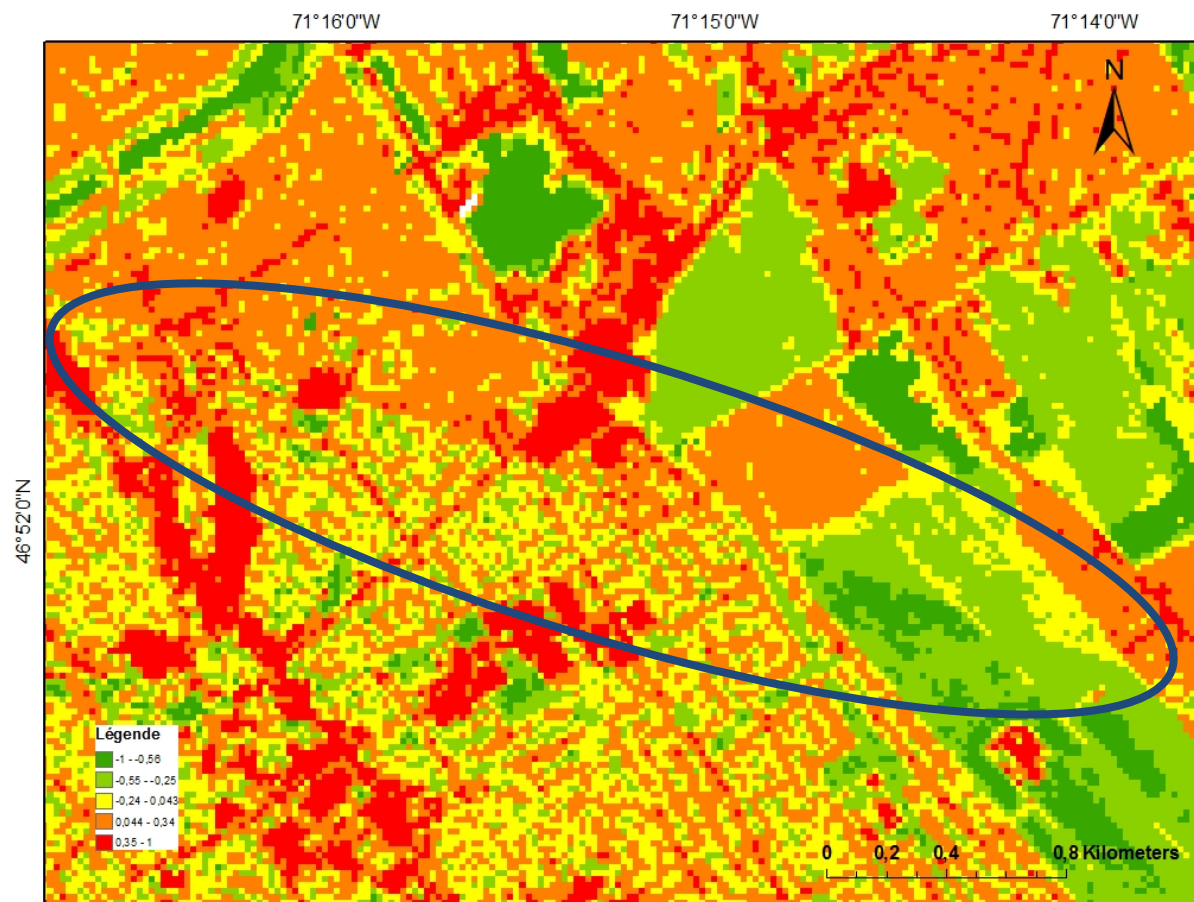


Figure 6. Exemple d'erreur due à la classification de la température de surface avec des données normalisées des températures prédites

8.5. EXEMPLE DE LA CARTOGRAPHIE DE TEMPÉRATURE RELATIVE

La figure 7 présente la cartographie de la température relative de surface au sol pour les secteurs des villes de Québec et Lévis. En zoom, on peut comparer, pour une même petite portion :

- l'ancienne cartographie des îlots de chaleur produite par Kestens *et al.* (2011) sur fond d'images issues de Google Maps;
- la cartographie de température de surface au sol réelle extraite de l'image Landsat (résolution à 100 m près),
- la cartographie de température de surface prédite à partir de l'image SPOT-5 (résolution à 20 m près).

On remarque que la nouvelle cartographie permet une caractérisation plus précise des îlots de chaleur que l'ancienne carte produite à partir de Landsat. On remarque également que la cartographie de la température réelle issue du canal thermique de Landsat présente un portrait moyen de la température au sol, dont les nuances sont clairement visibles sur la carte issue de l'image SPOT-5.

Distribution de la température de surface pour les villes de Québec et de Lévis
Températures prédites à partir des données SPOT-5 (20 m) 28/08/2007

Pour un même secteur

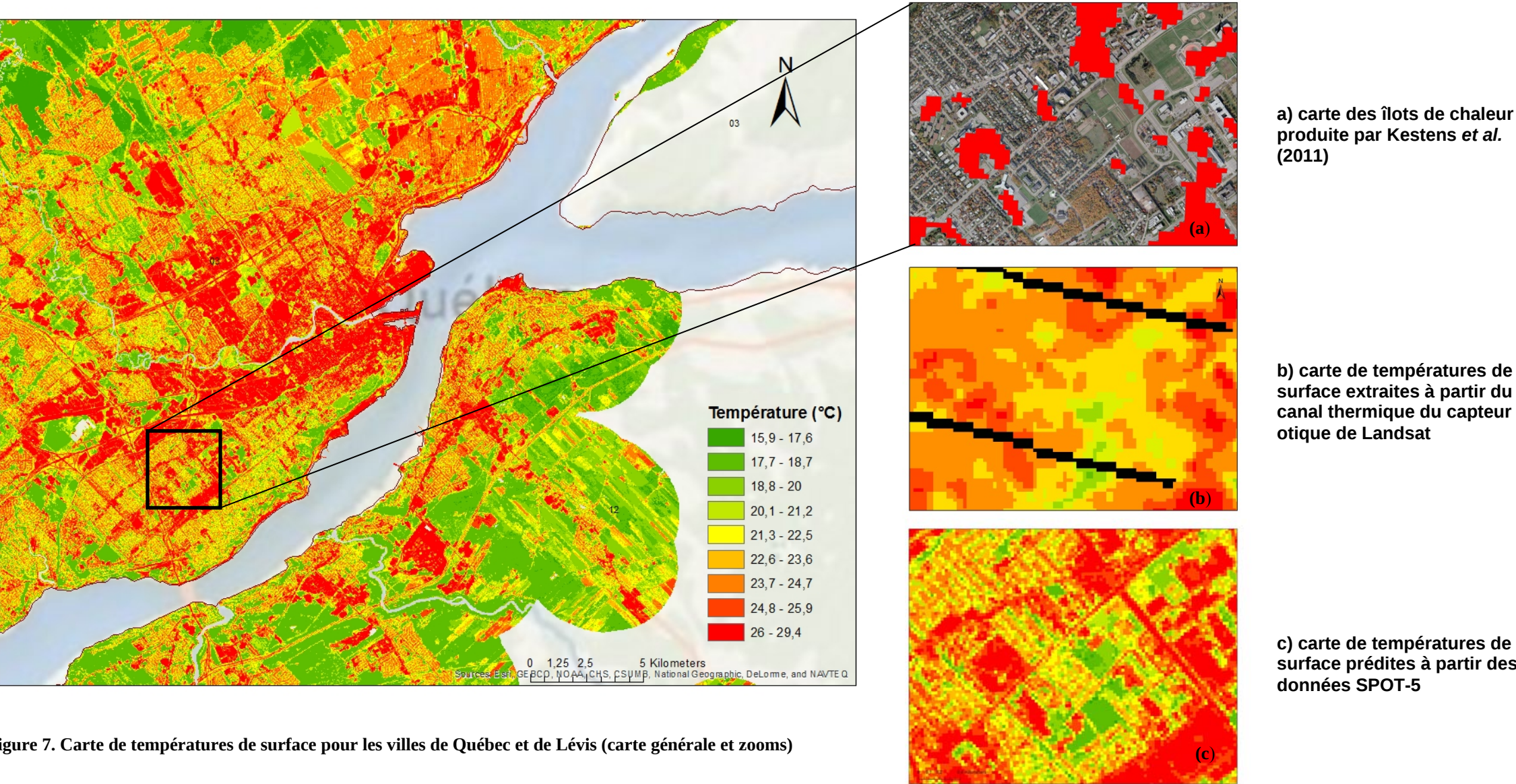


Figure 7. Carte de températures de surface pour les villes de Québec et de Lévis (carte générale et zooms)

9. BILAN DES LIMITES MÉTHODOLOGIQUES DE L'ÉTUDE

Ce projet présente un premier défi, qui consiste à développer une **méthode qui devra être appliquée à un très grand territoire d'étude** (tout le Québec urbain), **et ce, avec une enveloppe budgétaire limitée**. Dans ce contexte, certains choix méthodologiques ont été faits, de manière à permettre l'élaboration d'un modèle prédictif de la température de surface et son application à tout le Québec urbain, selon un coût raisonnable.

Le cadre méthodologique suivi pour l'élaboration du nouveau modèle de prédiction de la température de surface était très clairement défini au démarrage du projet et comportait certaines limites qui ont été acceptées comme un compromis satisfaisant : il s'agissait en effet d'adapter le modèle de Kestens *et al.* (2011) et de tester les mêmes variables que celles utilisées dans ce dernier, en utilisant une imagerie satellitaire à plus haute résolution que les images Landsat. **Les variables explicatives utilisées dans le nouveau modèle prédictif ont donc été choisies a priori, en se référant au modèle de Kestens et al. (2011)** (consulter la section 5 pour davantage de détails). L'amélioration du pouvoir prédictif du modèle de température devait provenir principalement de l'augmentation de la résolution du type d'image satellitaire utilisé.

De plus, il est essentiel de garder en tête que l'objectif poursuivi par ce modèle n'est pas de produire une cartographie précise de températures absolues, mais une **cartographie de températures relatives, permettant de localiser les zones potentiellement les plus chaudes et les plus fraîches**.

Dans ce contexte, il a été décidé de **ne pas normaliser la température issue des images Landsat**, qui a été utilisée comme variable dépendante dans le modèle prédictif. En effet, il n'y avait aucune certitude que le temps investi pour cet exercice aurait permis d'améliorer l'exactitude prédictive du modèle.

10. PISTES D'AMÉLIORATIONS POSSIBLES

Dans le contexte du projet actuel, dont l'objectif était de produire un outil opérationnel et utilisable pour toutes les municipalités du Québec, la carte produite à partir du nouveau modèle constitue une amélioration importante par rapport à la cartographie issue du modèle de Kestens *et al.* (2011).

Par contre, la mise en relation de 17 images Landsat et d'un grand nombre d'image SPOT-5 (67) est très exigeante. Certaines recherches ou certains essais n'ont pu être réalisés, étant donné les limites financières du projet. Dans un contexte qui serait davantage de l'ordre de la recherche, certaines améliorations pourraient être tentées :

- 1- **Réaliser une cartographie des îlots de chaleur et de fraîcheur en utilisant des images issues du capteur RapidEye**, dont la résolution spatiale est plus fine (résolution native de 6,5 m, ré-échantillonnée à 5 m) par rapport à celle des images SPOT-5 (20 m). Ces images contribueraient très certainement à préciser la cartographie des îlots de chaleur et de fraîcheur si un nouveau modèle était développé avec ces nouvelles données. Par contre, le prix à payer pour couvrir l'ensemble des zones urbaines du Québec est élevé car il n'existe aucune banque d'images gratuites contrairement aux images SPOT.
- 2- **Utiliser l'image panchromatique haute résolution (10 m)** pour améliorer la résolution des images SPOT-5. Dans le cadre du présent travail, les images panchromatiques n'ont pas été utilisées, à cause de l'importance du temps de traitement qu'elles impliquent. L'utilisation de ces images a par contre été testée pour le territoire de la ville de Québec. Les résultats obtenus sont très encourageants et la carte de température prédite a une résolution bien plus fine. L'utilisation d'une résolution de 10 m peut donc améliorer significativement la classification des îlots de chaleur/fraîcheur à l'échelle des zones urbaines au Québec, mais implique un travail de traitement d'images supplémentaire non négligeable, lorsqu'il faut le réaliser à l'échelle de tout le Québec urbain.
- 3- **Tester les effets de la normalisation des données de température** provenant de l'image Landsat sur le choix des variables explicatives du modèle et le pouvoir prédictif du nouveau modèle ainsi développé. Cette étape implique par contre un travail exigeant à grande échelle (utilisation d'au moins 17 images Landsat pour couvrir tout le Québec urbain) et de nombreux essais et ajustements de procédures qu'il serait intéressant de tester dans un contexte de développement scientifique moins limité par les contraintes financières de notre projet.

RÉFÉRENCES BIBLIOGRAPHIQUES

- Bartolucci, L.A., M. Chang, P.E. Anuta, and M.R. Graves. 1988. Atmospheric effects on Landsat TM thermal IR data. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, p. 171-175.
- Bhang, K.J. and S.-S. Park. 2009. Park evaluation of the surface temperature variation with surface settings on the urban heat island in Seoul, Korea, using Landsat-7 ETM+ and SPOT. *IEEE Geosciences and Remote Sensing Letters*, vol. 6, n° 4, p. 708-712.
- Bishop Y. M. M., Fienberg S. E., and Holland P. W. *Discrete Multivariate Analysis: Theory and Practice*, MIT Press, Cambridge, UK, 1975.
- Dousset, B, F. Gourmelon, K. Laaidi, A. Zeghnoun, E. Giraudet, P. Bretin, E. Maurid and S. Vandentorren. 2011. Satellite monitoring of summer heat waves in the Paris metropolitan area. *International Journal of Climatology*, vol. 31, n° 2, p. 313-323.
- Kestens, Y., A. Brand, M. Fournier, S. Goudreau, T. Kosatsky, M. Maloley and A. Smargiassi. 2011. Modelling the variation of land surface temperature as determinant of risk of heat-related health events. *International Journal of Health Geographics*, vol. 10, n° 7, p. 2-9.
- Montserud R. A. and Leamans R. Comparing global vegetation maps with the Kappa statistic, *Ecological Modelling*, vol. 62, no. 4, pp. 275–293, 1992.
- NASA. 2000. Landsat 7, Science Data Users Handbook, Data Products. Chapter 11. Site web : http://landsathandbook.gsfc.nasa.gov/pdfs/Landsat7_Handbook.pdf
- Pontius, Jr.R.G. (2000) Quantification error versus location in comparison of categorical maps. *Photogrammetric Engineering and Remote Sensing*, vol. 66, n° 8, p.1011–1016.
- Story, M. and Congalton, R.G. (1986) Accuracy assessment: A user's perspective. *Photogrammetric Engineering and Remote Sensing*, vol. 52, p. 397–399.
- Yuan, F. and M.E. Bauer. 2007. Comparison of impervious surface area and normalized difference vegetation index as indicators of surface heat island effects in Landsat imagery. *Remote Sensing of Environment*, vol. 106, n° 3, p. 375–386.