

# Rapport méthodologique

Cartographie des îlots de chaleur et de fraîcheur dans le  
Québec urbain à l'aide d'imagerie satellitaire Landsat-8  
(2013-2014)

Présenté à

*Institut national  
de santé publique*

Québec 

Ray Bustinza, M. Sc.

Céline Campagna, Ph. D.

Par :



**CERFO**

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

Jean Marchal, Ph. D.

Batistin Bour, M. Sc.

Brindusa Cristina Budei, Ph. D.

Fiston Nininahazwe, M. Sc.

Camille Armellin, B.A.

Mathieu Varin, M. Sc.

Septembre 2022

Ce projet est financé dans le cadre de Plan d'action 2013-2020 sur les changements climatiques.

Mots-clés : îlots de chaleur et de fraîcheur urbains, écoumène du Québec, centres de population, télédétection, traitement d'images, intelligence artificielle, apprentissage automatique, *Google Earth Engine*.

Référence à citer :

Marchal, J., Bour, B., Budei, B.C., Nininahazwe, F., Armellin, C. et Varin, M. 2022. Cartographie des îlots de chaleur et de fraîcheur dans le Québec urbain à l'aide d'imagerie satellitaire Landsat-8 (2013-2014). Centre d'enseignement et de recherche en foresterie de Sainte-Foy inc. (CERFO). Rapport 2022-29. 55 pages.

Réf.: 2022-29 / MV / 03/2022

## Résumé

Au Québec, les dix années les plus chaudes du siècle ont été enregistrées à partir des années 1990 et les prévisions indiquent que les températures moyennes continueront à augmenter durant les prochaines décennies. Cette hausse constante de la température 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. Ce phénomène se définit par des températures plus élevées dans les milieux urbains que dans les zones rurales avoisinantes, entraînant des problèmes de santé importants pouvant aller jusqu'à la mort dans les groupes de populations vulnérables. Il devient dès lors important de localiser à la fois les îlots de chaleurs et de fraîcheur urbains (ICFU) pour établir des plans de gestion et d'intervention du territoire et de la population à risque. Le but du projet était de produire une cartographie des ICFU dans les centres de population (recensement 2021, régions ayant au moins 1 000 habitants et une densité d'au moins 400 habitants au km<sup>2</sup>) à partir de données satellitaires (données 2013-2014) afin de mettre à jour la cartographie de 2012 (données 2005-2011) pour que l'Institut National de Santé Publique du Québec ait en main un outil de suivi. Ainsi, un modèle prédictif de l'écart de température relatif robuste ( $R^2$  : 0,81, erreur moyenne absolue : 1,99 °C, soit un gain de 14 % par rapport à l'étude de 2012) a été produit afin de pouvoir par la suite identifier les ICFU. La cartographie des ICFU, 9 classes, couvre l'ensemble du Québec urbain, soit 14 508 km<sup>2</sup>. La principale limite est que l'amplitude entre l'écart de température relatif le moins élevé et le plus élevé varie selon les régions. Un îlot de chaleur peut donc être identifié avec une différence d'écart relatif plus ou moins élevée. Il est donc important d'interpréter l'identification des îlots de chaleur en relation avec la donnée d'écart de température relatif produite à l'échelle de l'écoumène du Québec (recensement de 2016, 167 764 km<sup>2</sup>). Ainsi, un écart de température relatif de 2°C peut être présent dans la classe 1 dans les régions situées dans le Sud du Québec, mais sera présent dans la classe 9 dans les régions du Nord du Québec. Les îlots de chaleur pourront être représentés dans les classes 8 et/ou 9 selon le territoire ciblé. La principale piste d'amélioration identifiée est l'ajout d'algorithmes d'apprentissage profond afin de rendre plus simple et uniforme la création des mosaïques et de variables. Cette cartographie servira de référence pour une analyse diachronique (2013-2014 c. 2020-2022) prévue dans le sous-projet 3 du présent projet.

## Composition du comité de suivi

Le comité de suivi du projet était constitué de :

- Centre d'enseignement et de recherche en foresterie :  
Mathieu Varin – Chef du laboratoire de télédétection  
Batistin Bour – Chercheur en télédétection forestière  
Jean Marchal – Chercheur en intelligence artificielle  
Emmanuelle Boulfroy – Chargée de projet en foresterie urbaine
- Université de Sherbrooke :  
Yacine Bouroubi – Professeur au département de géomatique appliquée
- Institut National de Santé publique du Québec :  
Ray Bustinza – Conseiller scientifique  
Nathalie Gravel – Analyste en géomatique  
Pierre Gosselin – Conseiller scientifique
- Ministère de l'Environnement et de la Lutte contre les changements climatiques :  
Frédéric Morneau-Vaillancourt – Responsable de projets de recherche  
Loïc Fournier-Simon – étudiant – stagiaire  
Véronique Parent-Lacharité – Conseillère en adaptation aux changements climatiques
- Ministère des Affaires municipales et de l'Habitation :  
Marion Vincens – Conseillère aux programmes en changements climatiques
- Union des Municipalités du Québec :  
Charlotte Legault-Bélanger – Conseillère aux politiques (Changements climatiques et environnement)

Résumé .....	3
1. Introduction.....	9
2. Objectifs du projet .....	10
3. Territoire d'étude.....	10
4. Matériels et méthodes .....	12
4.1. Données et prétraitements .....	13
4.1.1. Données d'imagerie .....	13
4.1.2. Collecte des données.....	14
4.1.3. Prétraitements des images.....	15
4.1.4. Données de température.....	17
4.2. Données auxiliaires .....	20
4.3. Calcul des variables.....	20
4.3.1. Occupation du territoire .....	20
4.3.2. Démixage spectral (indice d'imperméabilité).....	24
4.3.3. Indices spectraux.....	26
4.3.4. Dates d'acquisition .....	27
4.3.5. Position géographique et topographique.....	28
4.3.6. Topologique .....	28
4.4. Échantillonnage des températures .....	29
4.5. Modélisation .....	31
4.6. Prédiction et classement des valeurs .....	34
5. Résultats.....	35
5.1. Variables utilisées.....	35
5.2. Prédiction de l'écart de température relatif prédit .....	42
5.3. Catégorisation des îlots de chaleur et de fraîcheur urbains .....	44
6. Discussion.....	45
6.1. Performances générales .....	45
6.2. Limites.....	47
6.3. Pistes d'amélioration .....	50
7. Conclusion .....	51
8. Références.....	52

## Liste des figures

Figure 1 : Territoire d'étude – écoumène du Québec en 2016 .....	11
Figure 2. Schéma méthodologique des cinq grandes étapes réalisées pour produire la cartographie des îlots de chaleur et de fraîcheur dans l'écoumène du Québec de 2016. SRTM : <i>Shuttle Radar Topography Mission</i> . GRHQ : Géobase du réseau hydrographique du Québec. RHN : Réseau hydro national.....	13
Figure 3 : Exemple de la fusion et correction d'image. ....	16
Figure 4. Exemples d'écart de températures dans des images acquises pendant des mois estivaux et en octobre. ....	18
Figure 5. Comparaison entre les températures au sol calculées avec les relevés des stations du réseau SURFRAD et les températures présentes dans la bande produite par le <i>United States Geological Survey</i> avec les données Landsat. ....	18
Figure 6 : Exemple de la bande 10 du capteur TIRS de Landsat-8 et la température produite par le <i>United States Geological Survey</i> . ....	19
Figure 7 : Exemple d'artéfacts (cercles jaunes) inclus dans la donnée de température produite par le <i>United States Geological Survey</i> . ....	19
Figure 8. Exemple de classes d'occupation du territoire utilisées: a. Eau, b. Sol, c. Urbain Mixte, d. Urbain Pur, e. Végétation arborescente, f. Végétation Basse. ....	21
Figure 9. Nombre de pixels d'entraînement et de validation par classe pour le développement du modèle d'occupation du territoire. ....	22
Figure 10. Caractéristiques et distribution de chaque variable utilisée pour développer le modèle d'occupation du territoire. SI : <i>Shadow Index</i> , NDWI : <i>Normalized Difference Water Index</i> , NDVI : <i>Normalized Difference Vegetation Index</i> .....	22
Figure 11. Matrice d'erreurs du modèle d'occupation du territoire.....	23
Figure 12 : Exemples de polygones homogènes (contour rouge) numérisés pour le démixage spectral sur l'image de Google (gauche) et Landsat-8 (droite) pour les quatre catégories utilisées. ....	25
Figure 13 : Variabilité spatiale d'un indice spectral, le <i>Normalized Difference Vegetation Index</i> , indicateur de la couverture végétale. ....	26
Figure 14 : Extrait de la mosaïque pour la zone de Montréal avec des images d'automne (gauche) et d'été (droite). ....	28
Figure 15 : Exemple de points échantillonnés (points noirs) et de référence (point jaune) pour normaliser la température. ....	29
Figure 16 : Corrélogramme des variables utilisées pour entraîner l'algorithme <i>Random Forest</i> ainsi que l'écart de température observé ( <i>Temperature</i> ).....	33
Figure 17 : Variation des îlots de chaleur (beige-jaune) prédits selon différentes méthodes de classement des valeurs.....	35
Figure 18 : Relation entre l'écart de température relatif en fonction de chacune des variables utilisées pour entraîner l'algorithme <i>Random Forest</i> . ....	41
Figure 19 : Écart de température relatif Landsat-8 observé en fonction de l'écart de température prédit pour le jeu de données de validation. La ligne de régression (bleue et solide) et l'hypothèse nulle (ordonnée à l'origine 0, pente 1 ; ligne pointillée rouge) sont également affichées. ....	43

Figure 20 : Exemple d'écart de température relatif prédit par le modèle (droite) et comparaison avec la température issue de Landsat-8 calculée par le *United States Geological Survey* (centre). Fond de carte Google à gauche. ....44

Figure 21 : Exemples de classement des îlots de chaleur et de fraîcheur urbains en milieu fortement urbanisé (haut) et en milieu rural (bas). Les îlots de chaleur semblent plus diffus en milieu rural.....45

Figure 22 : Exemple de contraste présent sur une mosaïque et écart de température relatif prédit.....46

Figure 23 : Comparaison entre les îlots de chaleur tels que produits par Boulfroy *et al.* (gauche) et ceux produits dans ce projet (droite).....47

Figure 24 : Limites de prédiction de la température normalisée pour les champs agricoles (cercle jaune) et le roc (ellipse bleue), qui représenteront des îlots de chaleur. ....49

Figure 25 : Limites de prédiction de la température normalisée pour les milieux humides (ellipse blanche) et les coupes forestières (rectangle bleu).....50

## Liste des tableaux

Tableau 1 : Description des bandes spectrales de Landsat-8.....	14
Tableau 2 : Liste des indices spectraux calculés. B2 à B7 : numéros des bandes Landsat-8 tels que décrits dans le Tableau 1.....	26
Tableau 3 : Description des images Landsat-8 utilisées pour extraire la température de surface.....	30
Tableau 4 : Liste des variables utilisées pour entraîner l’algorithme <i>Random Forest</i> pour le modèle prédictif de l’écart de température relatif.....	32
Tableau 5 : Liste des hyperparamètres optimaux utilisés pour entraîner l’algorithme <i>Random Forest</i> .....	33
Tableau 6 : Méthodes de classement des valeurs pour définir les îlots de chaleur et de fraîcheur urbains de différentes études réalisées au Québec. ....	34
Tableau 7 : Importance des variables. ....	35
Tableau 8 : Exemples de variables calculées.....	37



## 1. Introduction

Depuis plusieurs années, voire quelques décennies, les effets des changements climatiques à l'échelle mondiale sont indéniables (IPCC, 2022). Au Québec, les dix années les plus chaudes du siècle ont été enregistrées à partir des années 1990 et les prévisions indiquent que les températures moyennes continueront à augmenter durant les prochaines décennies (MELCC, 2020). 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, et même causer des décès chez certains groupes de populations (Giguère, 2009; Kaplan *et al.*, 2018). 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 (accès à des climatiseurs, piscines, cours ombragées, etc.) (Shiab et Bouchard, 2022; Smargiassi *et al.*, 2009).

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 (Lessard et Boulfroy, 2008).

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. La dernière cartographie des îlots de chaleur et de fraîcheur urbains (ICFU) réalisée par le CERFO remonte de plusieurs années (< 2010,  $R^2$  : 0,67) (Boulfroy *et al.*, 2012, 2013) et se basait sur l'étude de Kestens *et al.* (2011) en utilisant l'imagerie Landsat-7 ETM+ et SPOT. Le CERFO a plus récemment (2016) réalisé une cartographie des ICFU de la communauté métropolitaine de Québec avec une méthodologie simplifiée et plus performante ( $R^2$  : 0,84), utilisant uniquement des images satellitaires Landsat-8 OLI/TIRS (Alhawiti et Mitsova, 2016; Varin *et al.*, 2016). D'autres études plus récentes ont également utilisé une approche similaire ((Goldblatt *et al.*, 2021) : F1-score de 0,65 ; (Nadizadeh Shorabeh *et al.*, 2020) :  $R^2$  de 0,87 ; (Guo *et al.*, 2015) :  $R^2$  de 0,61).

L'Institut national de santé publique du Québec (INSPQ) a souhaité, avec la réalisation de ce projet, mettre à jour la cartographie des ICFU pour 2013-2014 sur l'ensemble du Québec

urbain afin d'effectuer ensuite une analyse diachronique avec une future cartographie prévue (2021–2022).

## 2. Objectifs du projet

Le but du projet est de produire une cartographie des ICFU dans l'écoumène (centres de population avec un minimum de 400 habitants par km<sup>2</sup>, recensement de 2021) du Québec de 2016 pour un horizon 2013–2014, afin de mettre à jour la cartographie actuelle datant de 2005–2011 pour que l'INSPQ ait en main un outil de suivi. L'objectif principal du projet est de concevoir un modèle robuste prédictif des écarts de température afin de classer les valeurs en différences de température. Plus précisément, il s'agit de :

- Produire des mosaïques de 2013–2014 couvrant tout le territoire d'étude ;
- Trouver les variables les plus pertinentes pour prédire la température de surface à l'aide d'un modèle d'apprentissage automatique (*Random Forest*) ;
- Cartographier les écarts de température relatifs dans l'écoumène du Québec avec une zone tampon de 5 km ;
- Cartographier les ICFU pour chaque centre de population (densité de population minimale de 400 habitants/km<sup>2</sup>) avec une zone tampon de 2 km.

## 3. Territoire d'étude

Le territoire d'étude couvre l'écoumène de population du Québec, qui désigne la surface habitée et comprend les centres de population ayant une densité de population minimale selon la définition de Statistiques Canada (<https://www150.statcan.gc.ca/n1/pub/92-195-x/2011001/geo/pop/pop-fra.htm>). Cet écoumène est basé sur le recensement de 2016 et après l'ajout d'une zone tampon de 5 km, représente 167 764 km<sup>2</sup>. Afin de simplifier les analyses de préparation des données, ce territoire a été divisé en 37 zones (Figure 1). La cartographie des ICFU a été produite uniquement dans les centres de population (données de 2021) ayant une densité de population de 400 habitants au km<sup>2</sup> avec une zone tampon de 2 km, ce qui représente 701 zones pour une superficie totale de 14 508 km<sup>2</sup>.

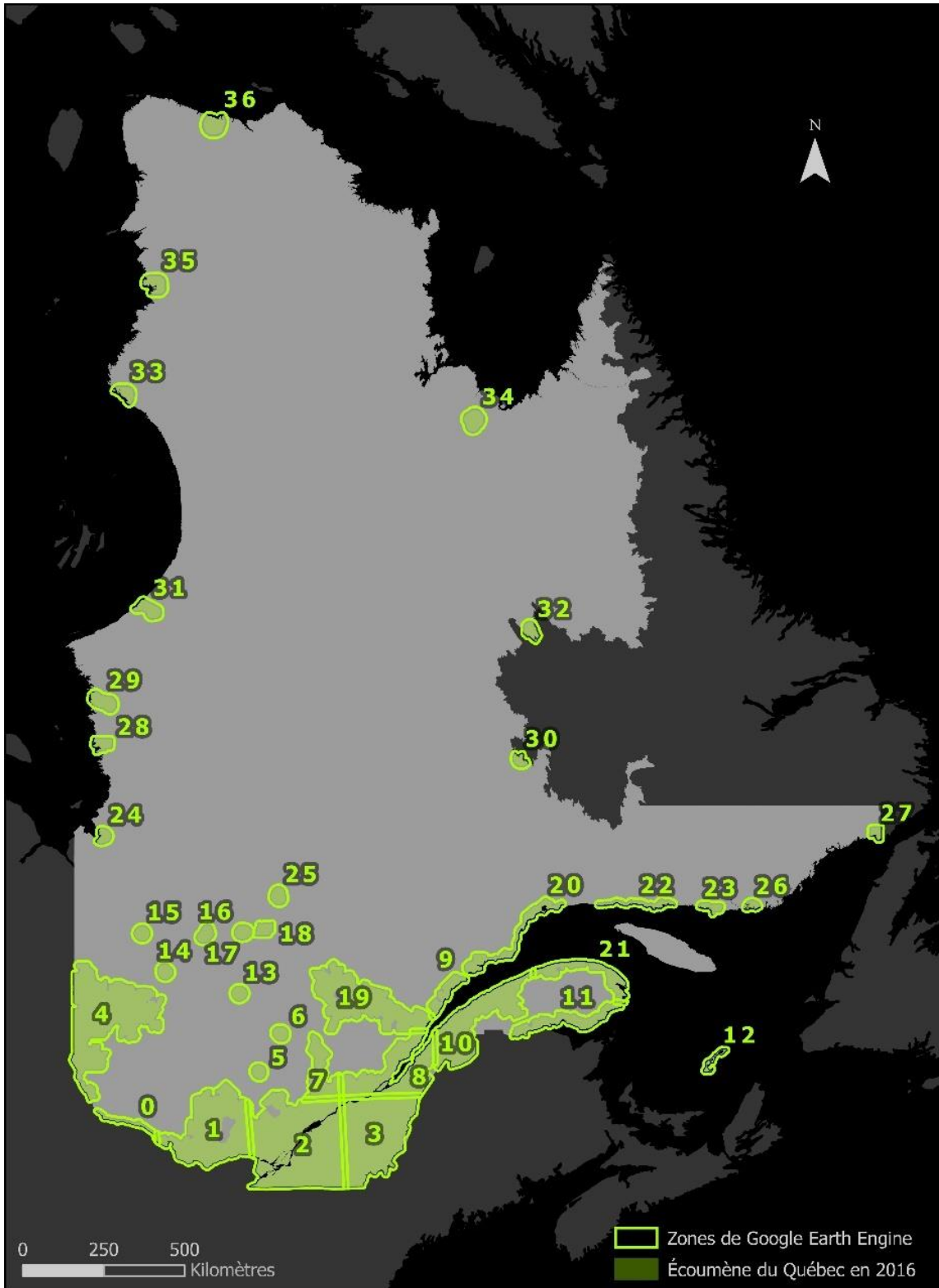


Figure 1 : Territoire d'étude – écoumène du Québec en 2016 et division du territoire pour les analyses.

#### 4. Matériels et méthodes

La réalisation de ce projet a nécessité la bonification d'une méthodologie employée précédemment par le CERFO (Bouffroy *et al.*, 2013; Varin *et al.*, 2016), basée initialement sur l'étude de Kestens *et al.* (2011). La puissance de cette approche, malgré sa complexité (Barrette *et al.*, 2018), est sa robustesse face aux conditions d'acquisition des images. En effet, malgré que des bandes thermiques ou de températures soient déjà fournies ou relativement simples à calculer (Shiab et Bouchard, 2022; USGS, 2020), il est peu probable de couvrir l'ensemble du territoire d'étude avec des images acquises durant des périodes de canicule sans aucune couverture nuageuse. Il devient alors très complexe d'assembler des images de températures différentes pour catégoriser leurs valeurs. Il est donc proposé d'utiliser un modèle construit à partir de la bande de température (30 m) de Landsat-8 fournie par le *United States Geological Survey* (USGS), pour prédire la température de surface à partir des autres bandes multispectrales fusionnées (résolution de 15 m) en y dérivant plusieurs variables indépendantes (Bouffroy *et al.*, 2013; Varin *et al.*, 2016) (Figure 2). L'objectif de prendre plusieurs variables est de contrebalancer les limites propres à chacune d'entre elles. Ces variables ont été utilisées dans la modélisation pour prédire un écart de température relatif. Toutes les autres données géospatiales ont été réunies (p. ex. hydrographie, modèle numérique de terrain) et projetées en Québec Lambert conique conforme. Les grilles correspondent parfaitement pour s'assurer qu'il n'y ait pas de décalage géométrique entre les données provenant de sources différentes.

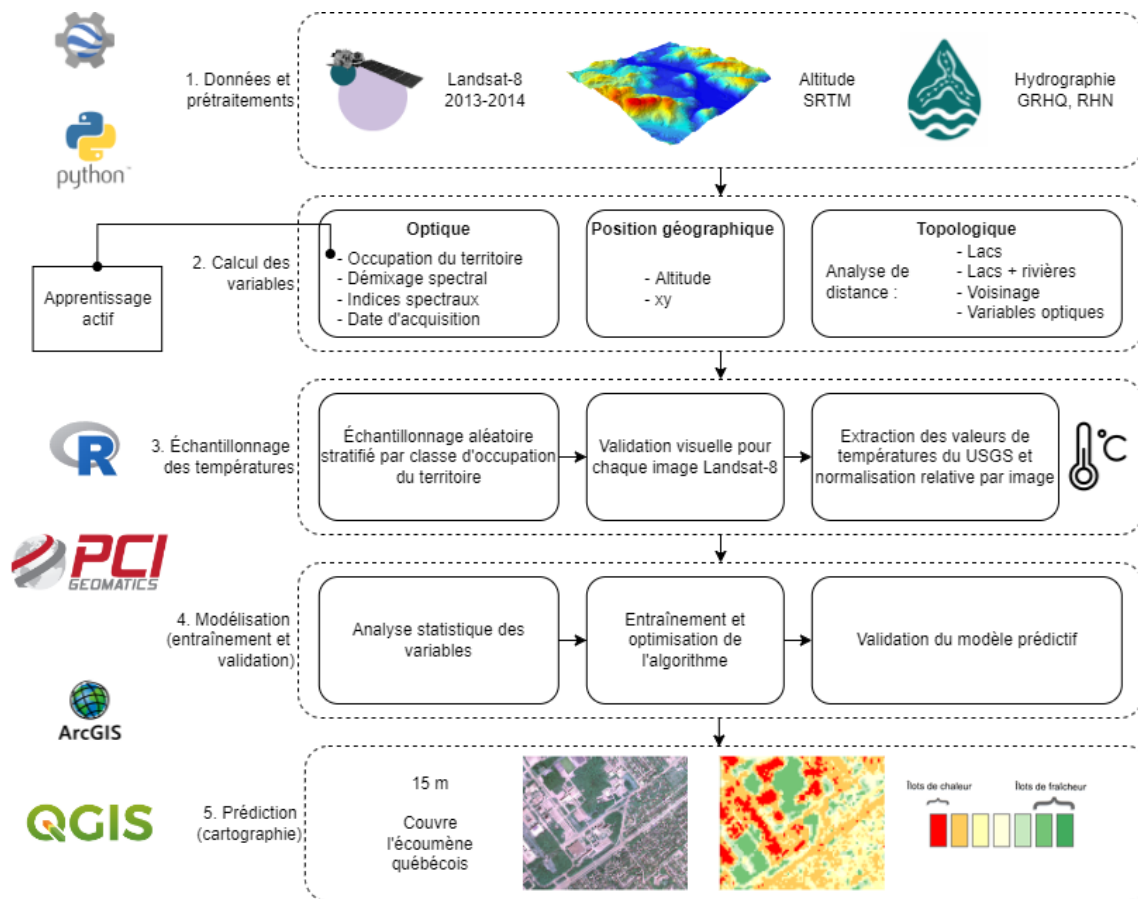


Figure 2. Schéma méthodologique des cinq grandes étapes réalisées pour produire la cartographie des îlots de chaleur et de fraîcheur dans l'écoumène du Québec de 2016. SRTM : Shuttle Radar Topography Mission. GRHQ : Géobase du réseau hydrographique du Québec. RHN : Réseau hydro national.

## 4.1. Données et prétraitements

### 4.1.1. Données d'imagerie

Les données des capteurs OLI (*Operational Land Imager*, bandes visibles et multispectrales) et TIRS (*Thermal InfraRed Sensor*, bandes thermiques) de Landsat-8 ont été utilisées pour dériver d'abord des variables prédictives et ensuite la donnée de température. Les caractéristiques du format des données Landsat-8 sont précisées dans le document du USGS (2020) et l'étendue des bandes spectrales est mentionnée dans le Tableau 1.

Tableau 1 : Description des bandes spectrales de Landsat-8.

Capteur	N° de bande	Description	Plage de longueurs d'onde (nm)	Résolution spatiale (m)
OLI	1	Côtière	435 – 451	30
	2	Bleue	452 – 512	30
	3	Verte	533 – 590	30
	4	Rouge	636 – 673	30
	5	Proche infrarouge	851 – 579	30
	6	Infrarouge à ondes courtes #1	1 566 – 1 651	30
	7	Infrarouge à ondes courtes #2	2 107 – 2 294	30
	8	Panchromatique	503 – 676	15
	9	Cirrus	1 363 – 1 384	30
TIRS	10	Infrarouge thermique #1	10 600 – 11 190	100
	11	Infrarouge thermique #2	11 500 – 12 510	100

OLI : *Operational Land Imager*, TIRS : *Thermal InfraRed Sensor*.

#### 4.1.2. Collecte des données

Plusieurs collections de données sont disponibles à partir des images Landsat et sont rendues disponibles sur le site internet du USGS (*EarthExplorer*). Le choix de la collection doit tenir compte du niveau de traitement et de la qualité de l'image. Les images utilisées, acquises avec le capteur OLI, font partie de la collection 2 *Tier 1 Level 1 TOA Reflectance*. Les valeurs des bandes représentent la réflectance enregistrée au niveau du capteur (*Top of the Atmosphere* - TOA). Cette collection a été choisie parce qu'elle apporte des améliorations concernant l'algorithme de détection des nuages, de la neige et de l'eau, ainsi qu'une meilleure intégration géométrique (précision des points de contrôle, des modèles numériques de terrain) par rapport à la collection 1 (USGS, 2020).

Les images Landsat sont classées par le USGS en deux catégories en fonction de leur qualité concernant la précision géométrique et les corrections radiométriques, les meilleures étant placées dans le *Tier 1*. Seulement les images de *Tier 1* ont été retenues. Le choix de la période d'acquisition a également été fait en fonction de la latitude, par exemple, en choisissant des images du mois d'août pour La Tuque et de juillet pour Montréal, car la phénologie varie selon la latitude (Lechowicz, 1984). L'atlas agroclimatique a aussi été consulté en ce sens et le site web suivant est également un bon indicateur de la perte de feuilles pour les arbres décidus : <https://www.bonjourquebec.com/fr-ca/decouvrir/saisons/automne>. Un effort a été appliqué afin de s'assurer que la végétation présente dans la période ciblée était bien visible. Les images Landsat-8 des capteurs OLI et TIRS sont reconnues comme étant de très bonne qualité pour le traitement d'images, par exemple pour la classification de l'occupation du territoire (Morfitt *et al.*, 2015), mais la disponibilité de ces images ne commence qu'en 2013.

#### 4.1.3. Prétraitements des images

Les images Landsat-8 ont une résolution spatiale est de 30 mètres pour les bandes multispectrales et 15 mètres pour la bande panchromatique. Un algorithme de *pansharpening* permet de fusionner les bandes multispectrales avec la bande panchromatique afin d'augmenter la résolution. Pour des soucis de cohérence, les bandes multispectrales et la bande panchromatique utilisées dans l'algorithme de *pansharpening* devaient avoir subi le même traitement de correction radiométrique. L'USGS met à disposition des utilisateurs les données de réflectance au sol (Collection 2 *Level 2*) (USGS, 2020). Même si très intéressante, cette série de données n'a pas été retenue, car la bande panchromatique n'est pas fournie dans ce jeu de données.

Les images ont été sélectionnées, traitées (masquer les nuages, *pansharpening*) et téléchargées avec l'application *Google Earth Engine* (GEE, <https://earthengine.google.com>) utilisant le *Code Editor* qui permet d'envoyer des requêtes en JavaScript. Plusieurs algorithmes pour réaliser le *pansharpening* (fusion d'images) sont disponibles. Johnson (2014) a fait une comparaison de l'effet de quatre méthodes différentes sur des indices de végétation et a démontré que la méthode HSV (*Hue Saturation Value*) était la plus performante pour conserver les informations spectrales et spatiales. Cette méthode, implémentée dans les fonctions *rgbToHsv* et *hsvToRgb* en GEE, a été utilisée pour augmenter la résolution des bandes rouge, verte et bleue. Les autres bandes spectrales ont été rééchantillonnées à 15 m par l'algorithme du plus proches voisins (*nearest neighbor*). En revanche, cet algorithme modifie la plage des valeurs des bandes spectrales par rapport aux valeurs initiales. En effet, la valeur de l'indice ERGAS (Erreur Relative Globale Adimensionnelle de Synthèse) qui permet d'évaluer la distorsion de l'information spectrale (Belfiore *et al.*, 2016; Li *et al.*, 2017; Park *et al.*, 2020) était très élevée (24,6) alors qu'elle devrait être proche de zéro. Ainsi, les coefficients d'une régression linéaire entre les valeurs des bandes rouge, verte et bleue issues du *pansharpening* et les valeurs des bandes initiales ont été utilisés pour corriger les valeurs après fusion. Cette correction a été appliquée itérativement par image avec R (R Core Team, 2021), avec un  $R^2$  supérieur à 0,9. Nous avons comparé les résultats de cette méthode de correction à ceux de l'algorithme RCS (Ratio Component Substitution) disponible dans OrfeoToolbox (OTB, <https://www.orfeo-toolbox.org>) et qui a produit des images fusionnées avec moins de distorsions spectrales dans d'autres travaux antérieurs (Deur *et al.*, 2021; Mhangara *et al.*, 2020; Nininahazwe *et al.*, 2021). L'indice ERGAS a montré que cette approche de régression linéaire avec HSV était pertinente avec une valeur très faible (2,9) et même inférieure à celle obtenue pour RCS (5,4). L'autre avantage de notre approche réside dans le temps de calcul plus rapide pour le *pansharpening* HSV de GEE et l'application de la correction par régression linéaire, par



rapport au temps de traitement de l'algorithme RCS. La Figure 3 montre un exemple d'image brute (30 m), fusionnée (15 m) et fusionnée-corrigée linéairement (15 m).



Figure 3 : Exemple de la fusion et correction d'image.

La bande de qualité de l'image Landsat (QA\_PIXEL) fournie dans la collection a été utilisée pour éliminer les zones couvertes par les nuages. Une zone plus large a été éliminée autour des nuages pour s'assurer qu'il n'y ait pas de limites de nuages non détectés. Une vérification manuelle a été réalisée pour s'assurer que les images utilisées étaient dépourvues de nuages, des voiles nuageux ou de traînées d'avions.

Plusieurs algorithmes pour réaliser la mosaïque des images peuvent être utilisés lorsqu'il y a superposition d'images (sélectionner le premier pixel, le pixel avec la valeur médiane ou maximale, une sélection manuelle de la provenance des pixels, etc.). De nombreuses techniques de mosaïquage ont été essayées pour obtenir le meilleur résultat, tenant compte de la présence des nuages, de la qualité des images disponibles et de leurs conditions d'acquisition. Par exemple, pour les régions nordiques, la période d'acquisition a été limitée afin d'éviter au maximum la présence de neige ou de glace dans les images.

Une attention particulière a été accordée pour que toutes les bandes spectrales d'un même pixel proviennent d'une même image. Cette condition est requise par le calcul d'indices de végétation, comme le NDVI (*Normalized Difference Vegetation Index*), qui suppose que les deux bandes spectrales utilisées dans la formule mathématique sont acquises en même temps ou en conditions similaires (mêmes conditions d'ensoleillement, d'acquisition, géométriques ou même phénologiques). Calculer les NDVI avec des bandes acquises en des conditions phénologiques et météorologiques différentes pourrait fausser la valeur de l'indice. Dans cette optique, nous avons évité d'utiliser la fonction *Reduce* proposée par GEE qui permet de produire des mosaïques très rapidement, mais qui applique l'algorithme de sélection des pixels pour chaque bande séparément (Google Earth Engine, 2022). De cette manière, l'indice spectral aurait été calculé à partir des images acquises dans des conditions très différentes.



Un autre algorithme proposé dans GEE pour sélectionner le meilleur pixel tenant compte des acquisitions de tous les capteurs Landsat a été proposé par White *et al.* (2014). L'application *Best available Pixel* – GEE (Francini, 2022) sélectionne les bandes appartenant à la même image. Un algorithme semblable au BAP a été utilisé pour cette étude, adapté au choix de capteur et de collection Landsat, et développé dans R.

En effet, le mosaïquage a été réalisé par un algorithme qui priorise les images de 2013 avec faible pourcentage de nuages, en dessous de 25 %. Les zones non couvertes par ces images ont été complétées par des images de 2013 avec un pourcentage de nuages entre 25 et 65 %. Pour les zones encore non couvertes par ces images, celles de 2014 ont été téléchargées et utilisées dans la même approche, pour compléter les pixels manquants. Pour le choix du pixel retenu dans la mosaïque, la valeur maximale de la bande infrarouge a été retenue afin de limiter la présence d'ombre des nuages. Une inspection visuelle des mosaïques, un choix des meilleures images à inclure dans les mosaïques et une vérification des erreurs de commission en ce qui concerne l'algorithme de détection de nuages a été réalisé manuellement dans *PCI Geomatica* (PCI Geomatics, 2018). Dans les zones où une seule image de bonne qualité pouvait être utilisée, elle a été retenue. Dans certains cas exceptionnels, des images de 2015 ont dû être ajoutées manuellement. Ainsi, 37 mosaïques ont été créées pour chacune des zones divisant le territoire d'étude. Ces mosaïques ont ensuite été utilisées pour dériver les variables optiques servant au modèle prédictif d'une résolution spatiale de 15 m.

#### 4.1.4. Données de température

Le choix des images Landsat-8 pour l'obtention de la température de surface (variable dépendante) a été fait en fonction de plusieurs critères, dont une couverture nuageuse minimale et une période de l'année optimale pour les ICFU basée sur les données météorologiques historiques de stations à proximité des régions concernées. Généralement, la température journalière était supérieure à 30 °C pour le sud du Québec. Une vérification de la pluviométrie dans les jours précédents la prise d'images a été faite. La période estivale couverte s'étale donc entre juin et septembre. Dans cette saison, les écarts de température observés sont similaires, mais cet écart se réduit lors d'une journée froide (Figure 4) Une recherche exhaustive des images a été faite pour utiliser les images de la meilleure qualité possible. La bande de température produite par le USGS a été utilisée pour échantillonner des pixels de température à une résolution au sol de 30 m (section 4.4). À des fins de validation, les valeurs d'émissivité radiométrique du réseau SURFRAD (Augustine *et al.*, 2008) ont été relevées sur 50 observations (cinq stations), correspondants aux heures d'acquisition d'images Landsat-8. Les valeurs d'émissivités radiométriques ont été traitées tel que décrit par Yu *et al.* (2014) afin d'obtenir une température au sol en degrés Celsius. Ces mesures ont été comparées avec la valeur du pixel de la bande de température produite par le USGS, là où les stations sont

situées. Cette comparaison montre une très bonne correspondance entre les températures mesurées au sol et calculées dans la bande de température Landsat-8 ( $R^2 = 0,911$  ; Figure 5)

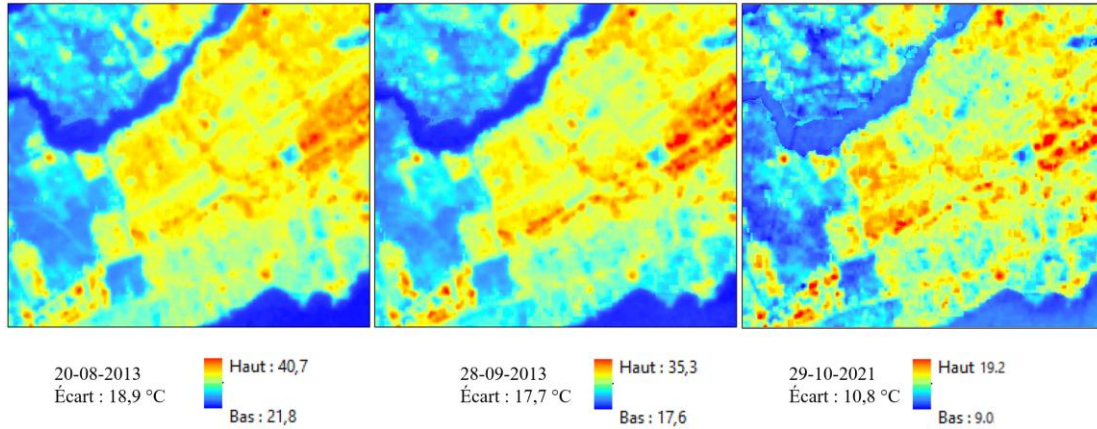


Figure 4. Exemples d'écart de températures dans des images acquises pendant des mois estivaux et en octobre.

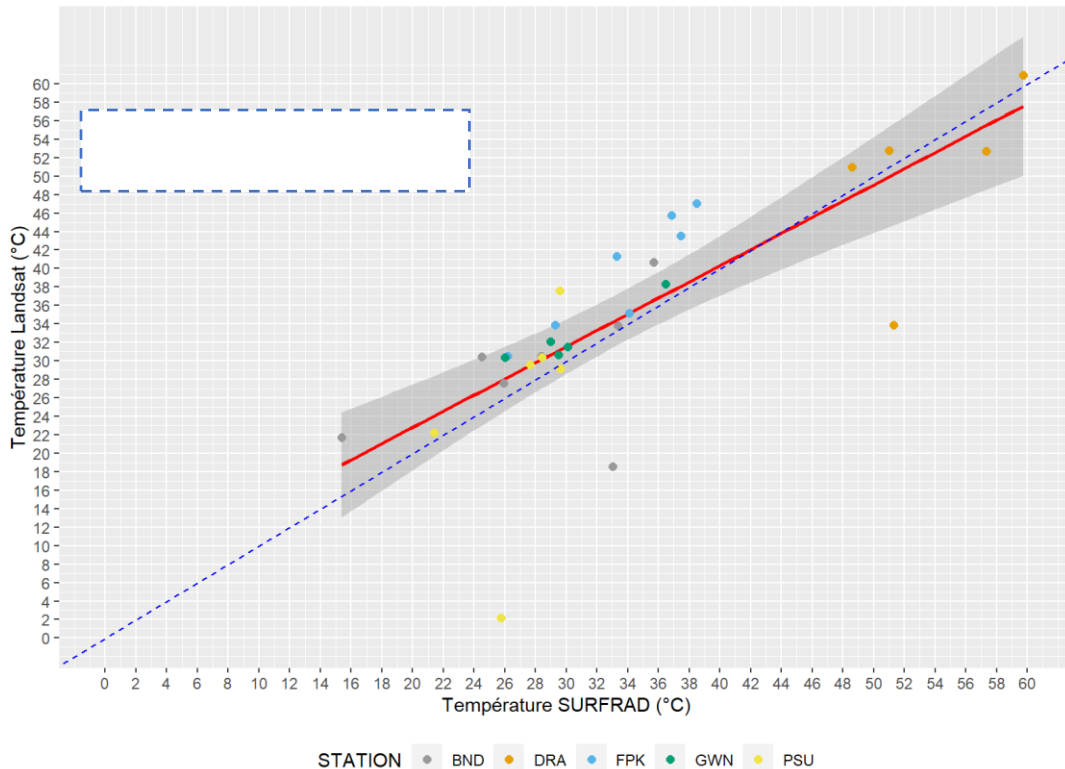


Figure 5. Comparaison entre les températures au sol calculées avec les relevés des stations du réseau SURFRAD et les températures présentes dans la bande produite par le *United States Geological Survey* avec les données Landsat.

La conversion de l'émissivité thermique contenue dans les bandes TIRS de Landsat-8 aurait nécessité des traitements supplémentaires induisant des erreurs potentielles, c'est pourquoi il a été choisi d'utiliser le produit du USGS. L'utilisation d'images Landsat pour générer une carte de températures de surface a effectivement déjà fait l'objet de projets de recherche (Bhang et Park, 2009; Yu *et al.*, 2014). Cette température est calculée à partir de la bande thermique 10 du capteur TIRS (100 m) qui est sensible à la radiation émise dans l'intervalle des longueurs d'onde de 10,6-11,19  $\mu\text{m}$ , où l'effet de l'atmosphère est minimal, ce qui permet d'obtenir des valeurs non bruitées par les aérosols et ne nécessitent pas de correction atmosphérique (Bartolucci *et al.*, 1988; Ghulam, 2010; Storey *et al.*, 2014). La Figure 6 montre un exemple des données Landsat-8 avec la bande thermique 10 et le produit de température généré par le USGS. Cette donnée contient cependant des artéfacts et n'est pas continue, elle ne peut donc être utilisée directement pour obtenir un portrait exhaustif d'une région (Figure 7).

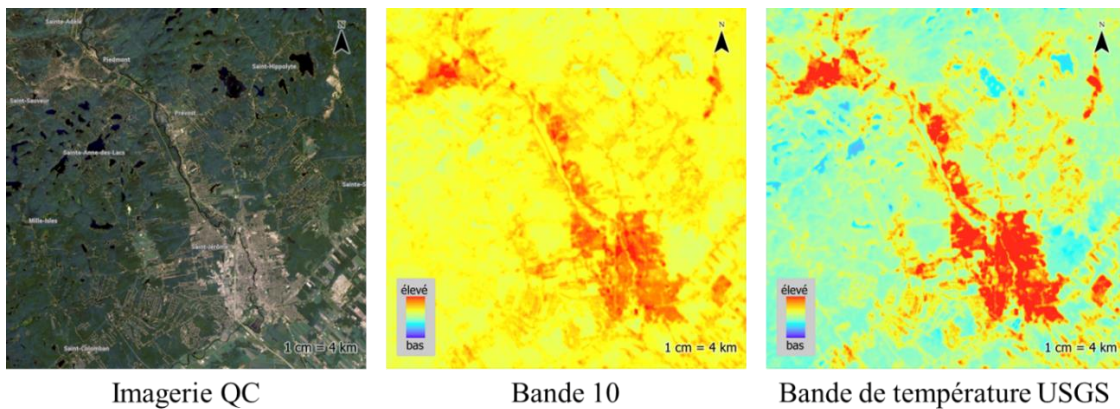


Figure 6 : Exemple de la bande 10 du capteur TIRS de Landsat-8 et la température produite par le *United States Geological Survey*.

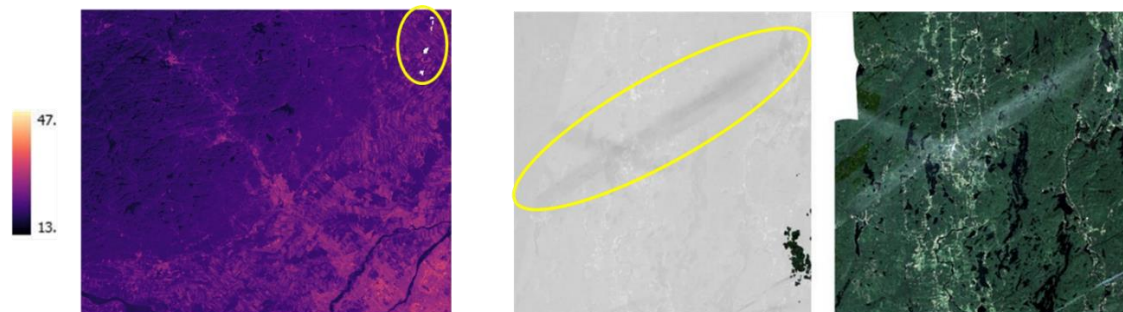


Figure 7 : Exemple d'artéfacts (cercles jaunes) inclus dans la donnée de température produite par le *United States Geological Survey*.

Ces images proviennent de la Collection 2 *Level 2*, ce sont 46 images qui ont ainsi été sélectionnées. Leur date d'acquisition est présentée dans le Tableau 3.

## 4.2. Données auxiliaires

Des données géospatiales autres que Landsat-8 ont été utilisées dans le modèle prédictif. Il s'agit de l'altitude fourni par le *Shuttle Radar Topography Mission* (SRTM) ayant une résolution spatiale de 30 m, complétée par le modèle numérique d'altitude (20 m) de Ressources Naturelles Canada (RNCAN) pour le Nord, l'hydrographie de la Géobase du réseau hydrographique du Québec (GRHQ, produit par le gouvernement du Québec) et complétée par le Réseau hydro national de la série Géobase (RHN, produit par le gouvernement du Canada).

## 4.3. Calcul des variables

D'abord, une classification de l'occupation du territoire en grandes classes a été effectuée et validée. Ensuite, un indice d'imperméabilité à partir d'une technique de démixage spectral a été utilisé (Van Der Meer, 1995). Des indices spectraux, des analyses de distance à l'hydrographie et aux variables d'occupation du territoire et du démixage, la position xyz et la saison et la semaine d'acquisition ont été extraites.

### 4.3.1. Occupation du territoire

Un modèle d'occupation du territoire basé sur l'apprentissage automatique avec l'algorithme *Random Forest* (Breiman, 2001), branche de l'intelligence artificielle, a été développé avec six classes d'occupation considérées, soit : Eau, Sol, Urbain Mixte, Urbain Pur, Végétation Arborescente et Végétation basse (Figure 8). Guo *et al.* (2015) a d'ailleurs démontré un gain de 6 % en précision en utilisant des arbres de régression par rapport à une régression linéaire multiple. Pour chaque zone de l'écoumène (Figure 1), des polygones ont été photo-interprétés pour toutes les classes. Un total de 1 611 polygones a ainsi été obtenu. Par zone, une séparation des polygones a été réalisée pour identifier des zones d'entraînement (80 %) et de validation (20 %). Les valeurs spectrales des pixels de Landsat-8 à 15 m de résolution entièrement inclus dans les polygones identifiés ont été extraites. Les indices NDVI, NDWI (*Normalized Difference Water Index*) et SI (*Shadow Index*) ont été calculés :

$$NDVI = \frac{\text{Bande 5} - \text{Bande 4}}{\text{Bande 5} + \text{Bande 4}}$$

$$NDWI = \frac{\text{Bande 5} - \text{Bande 6}}{\text{Bande 5} + \text{Bande 6}}$$

$$SI = ((1 - \text{Bande 2}) \times (1 - \text{Bande 3}) \times (1 - \text{Bande 4}))^{\frac{1}{3}}$$



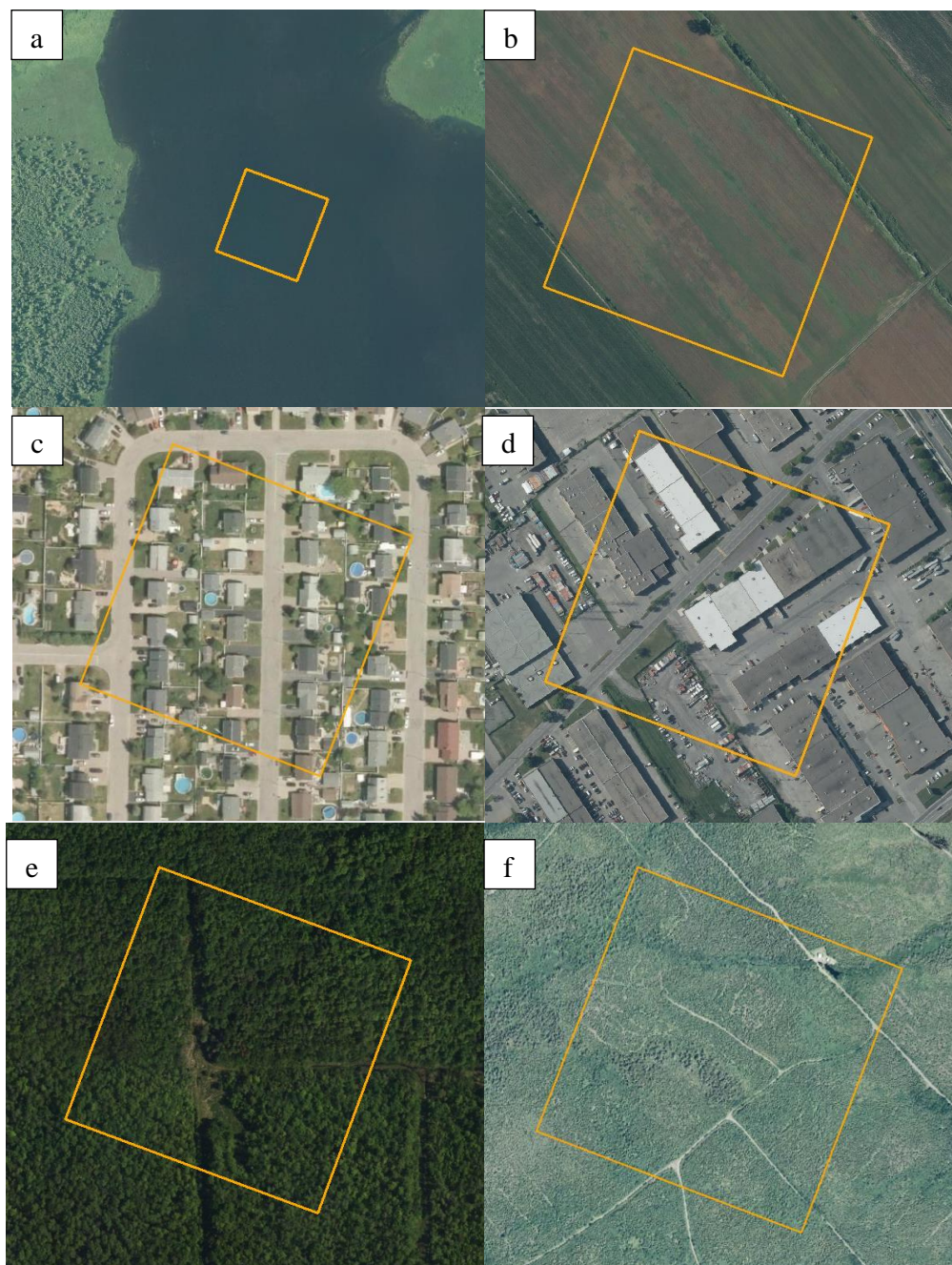


Figure 8. Exemple de classes d'occupation du territoire utilisées: a. Eau, b. Sol, c. Urbain Mixte, d. Urbain Pur, e. Végétation arborescente, f. Végétation Basse.

Un premier modèle global a été développé puis appliqué à l'ensemble de l'écoumène. Les résultats de ce modèle ont été manuellement inspectés et ont servi à créer de nouveaux polygones d'entraînement par une approche d'apprentissage actif, amenant le jeu de données final à 1 803 polygones d'entraînement et de validation (Figure 9 et Figure 10).

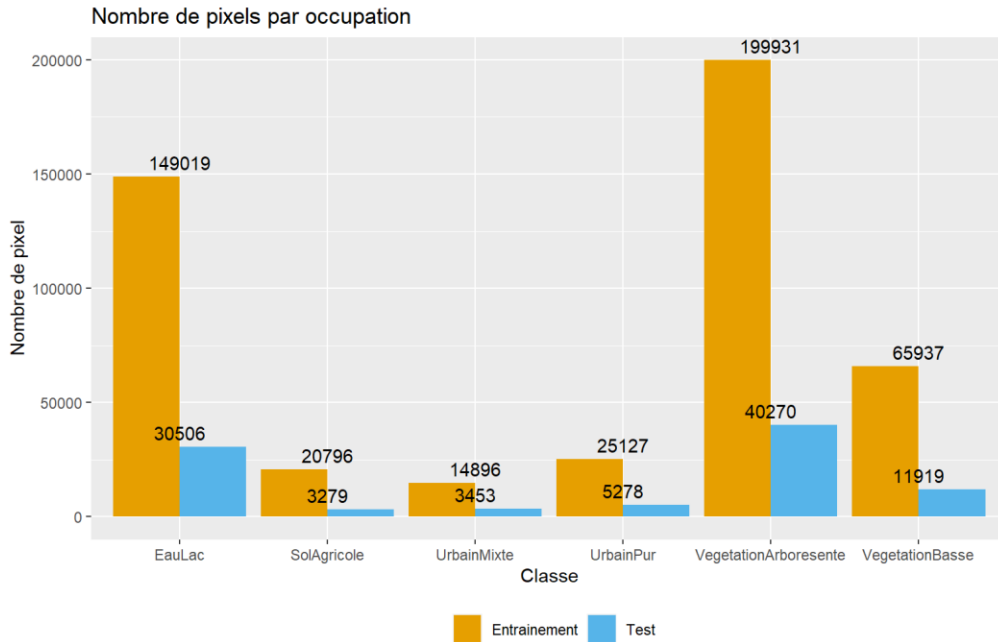


Figure 9. Nombre de pixels d'entrainement et de validation par classe pour le développement du modèle d'occupation du territoire.

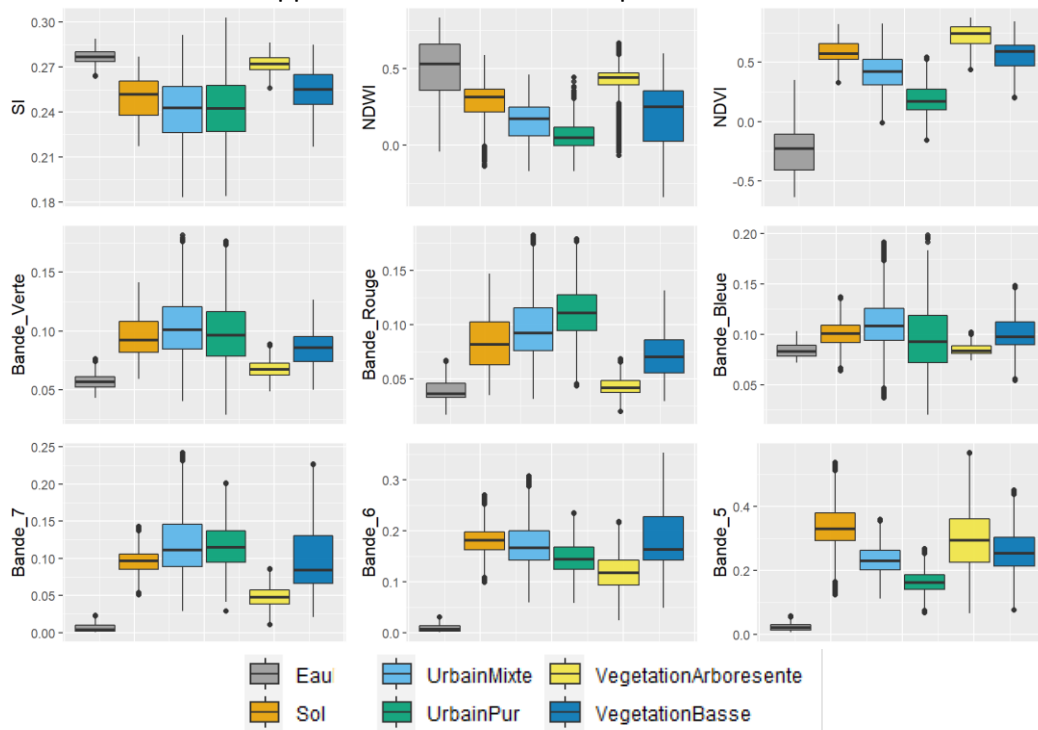


Figure 10. Caractéristiques et distribution de chaque variable utilisée pour développer le modèle d'occupation du territoire. SI : *Shadow Index*, NDWI : *Normalized Difference Water Index*, NDVI : *Normalized Difference Vegetation Index*.

L'algorithme *Random Forest* entraîné avec ces données a été appliqué sur les données de validation pour évaluer les performances. Une matrice d'erreurs a ainsi été obtenue, permettant d'identifier, par classe, les erreurs d'omission et de commission, mais également des statistiques de performance générale telles que la précision globale et le coefficient de Kappa de Cohen (Cohen, 1960; Congalton, 1991). Avec l'analyse de ces statistiques, il est possible d'affirmer que le modèle d'occupation du territoire développé montre une excellente (Monserud et Leemans, 1992) capacité de discernement des classes d'occupation du territoire avec une précision globale de 0,9133 et une valeur de coefficient de Kappa de 0,8739 (Figure 11). La classe « Eau » présente un succès de classification presque parfait. À l'opposé, la classe « Sol » présente le plus d'erreurs, avec de la commission dans les classes végétation arborescente et surtout végétation basse. Ces erreurs peuvent être attribuées à la présence de cultures dans des sols agricoles, alternant avec des zones de sol nu.

		Référence					
		Eau	Sol	Urbain Mixte	Urbain Pur	Végétation Arborescente	Végétation Basse
Prédiction	Eau	20891	0	0	1	0	0
	Sol	0	914	27	4	33	553
	Urbain Mixte	0	154	2515	301	14	325
	Urbain Pur	0	8	448	4394	0	40
	Végétation Arborescente	0	478	37	0	35286	1439
	Végétation Basse	0	979	106	4	1982	9057

Détails				
Sensitivité	Spécificité	Précision	Prob. détection	Score-F1
1.0000	0.3608	0.9456	0.7935	1.0000
<b>Précision globale</b>		<b>Coefficient de Kappa</b>		
0.9133		0.8739		

Figure 11. Matrice d'erreurs du modèle d'occupation du territoire.

Un post-traitement a été appliqué afin de corriger la classe « Eau » en appliquant un masque du GRHQ et du RHN. Par la suite, le modèle a été appliqué à l'ensemble de l'écoumène pour ainsi constituer la variable d'occupation du territoire.

#### 4.3.2. Démixage spectral (indice d'imperméabilité)

Les images à résolution moyenne telles que celles de Landsat-8 OLI utilisées (15 m après la fusion) sont souvent caractérisées par une réflectance qui est un mélange de plusieurs catégories d'occupation du territoire, lorsque celles-ci occupent une superficie inférieure à celle du pixel (Small, 2001). La réflectance du pixel peut être dans ce cas un mélange de la végétation, du sol nu et de la construction, selon les composantes présentes à l'intérieur du pixel mixte. Pour servir d'indicateur d'imperméabilité des surfaces, les catégories identifiées par interprétation visuelle dans les images sont principalement représentées par les constructions (surfaces imperméables), végétation, sol nu et l'eau. Ainsi, étant donné que la superficie de certains objets de ces catégories est souvent inférieure à celle du pixel, nous avons estimé les proportions relatives d'occupation d'un pixel pour chacune de ces classes (communément appelé « *endmembers* ») en utilisant un modèle de démixage linéaire NNLS (*Non-Negative Linear Least squares*). Ce modèle permet de produire des cartes quantitatives d'abondance (les proportions d'abondances respectives dont la somme est égale ou proche à l'unité) (Ravel, 2017). La librairie RStoolbox du langage R a été utilisée pour cette analyse. Les polygones homogènes de référence ont été numérisés sur les images Landsat-8 avec comme fond de carte de l'imagerie à très haute résolution spatiale de Google, répartis dans chaque zone. Ainsi, un modèle par zone a été développé et appliqué. La valeur spectrale du pixel correspondant au centroïde du polygone a été utilisée dans l'entraînement des modèles NNLS. Les bandes qui ont été utilisées sont la bande verte, rouge, proche infrarouge, infrarouge à onde courte 1 et 2 (Tableau 1). Les bandes bleue et côtière n'ont pas été utilisées, car nous avons constaté qu'il y avait beaucoup de bruit dans les images d'abondance lorsque ces bandes sont utilisées. Quelques exemples de polygones homogènes et centroïdes correspondants utilisés sont présentés à la Figure 12.



Catégorie 1 : Végétation



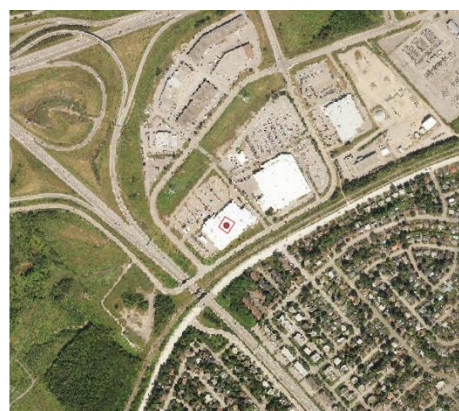




Catégorie 2 : Eau



Catégorie 3 : Sol nu



Catégorie 4 : Construction



Figure 12 : Exemples de polygones homogènes (contour rouge) numérisés pour le démixage spectral sur l'image de Google (gauche) et Landsat-8 (droite) pour les quatre catégories utilisées.

Un post-traitement a été appliqué afin de corriger la classe « Eau » en appliquant un masque avec les données du GRHQ et du RHN.

### 4.3.3. Indices spectraux

Un total de 33 indices spectraux a été calculé sur l'ensemble des mosaïques. Ces indices mesurent par exemple la vigueur de la végétation (Figure 13), le taux d'humidité de celle-ci ou du sol. Ces 33 indices sont décrits dans le Tableau 2.

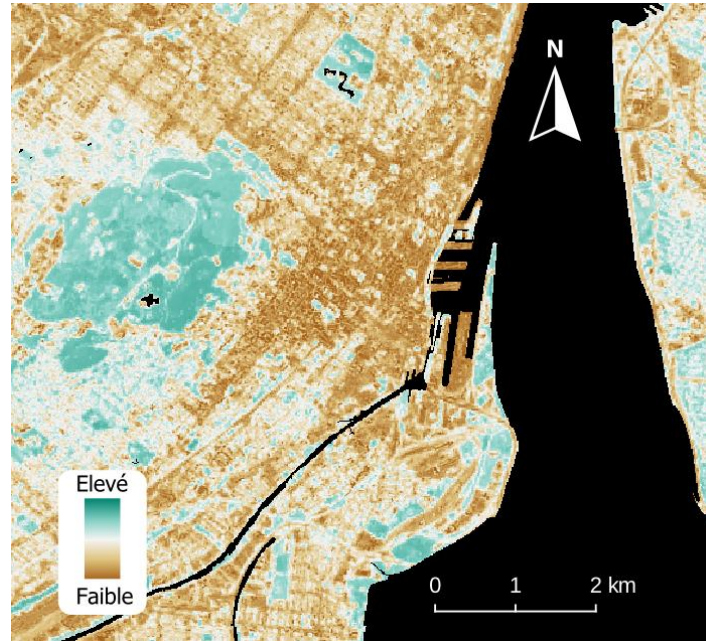


Figure 13 : Variabilité spatiale d'un indice spectral, le *Normalized Difference Vegetation Index*, indicateur de la couverture végétale.

Tableau 2 : Liste des indices spectraux calculés. B2 à B7 : numéros des bandes Landsat-8 tels que décrits dans le Tableau 1.

Indice spectral	Équation	Référence
ADB11	$B5 * \text{atan}((B5 - B6) / (865 - 1609))$	(Liu <i>et al.</i> , 2010)
ADB12	$B5 * \text{atan}((B5 - B7) / (865 - 2200.5))$	(Liu <i>et al.</i> , 2010)
AEWI	$(4 * (B3 - B6) - (0,25 * B5 + 2.75 * B7))$	(Feyisa <i>et al.</i> , 2014)
BAI	$1 / ((0.1 - B4)^2 + (0.06 - B5)^2)$	(Chuvieco <i>et al.</i> , 2002)
GBNDVI	$(B5 - (B3 + B2)) / (B5 + (B3 + B2))$	(Wang <i>et al.</i> , 2007)
GDVI	$B5 - B3$	-
GI	$B3 / B4$	-
GNDVI	$(B5 - B3) / (B5 + B3)$	(Xue et Su, 2017)
MBWI	$3 * B3 - B4 - B5 - B6 - B7$	(Cordeiro <i>et al.</i> , 2021)
MCARI1	$1.2 * (2.5 * (B5 - B4) - 1.3 * (B5 - B3))$	(Haboudane, 2004)
MCARI2	$1.5 * (2.5 * (B5 - B4) - 1.3 * (B5 - B3)) / \text{sqrt}((2 * B5 + 1)^2 - (6 * B5 - 5 * \text{sqrt}(B4)) - 0.5)$	(Haboudane, 2004)
MNDWI	$(B3 - B6) / (B3 + B6)$	(Xu, 2006)

MSAVI	$0.5*((2*B5+1)-\sqrt{(2*B5+1)^2-8*(B5-B4)})$	(Xue et Su, 2017)
MSI	$B6/B5$	(Huntjr et Rock, 1989)
MTVI1	$1.2*(1.2*(B5-B3)-2.5*(B4-B3))$	(Haboudane, 2004)
MTVI2	$1.5*(1.2*(B5-B3)-2.5*(B4-B3))/\sqrt{(2*B5+1)^2-(6*B5-5*\sqrt{B4})}-0.5$	(Haboudane, 2004)
NBR	$(B5-B7)/(B5+B7)$	(García et Caselles, 1991)
NDMI	$(B5-B6)/(B5+B6)$	(Taloor <i>et al.</i> , 2021)
NDTI	$(B6-B7)/(B6+B7)$	(Chaves <i>et al.</i> , 2020)
NDVI	$(B5-B4)/(B5+B4)$	(Gamon <i>et al.</i> , 1995)
NDWI	$(B3-B5)/(B5+B3)$	(Mustafa <i>et al.</i> , 2017)
NHI	$(B6-B3)/(B6+B3)$	(Schindler <i>et al.</i> , 2021)
NISI	$(B2+B3+B4-B5)/(B2+B3+B4+B5)$	(Su <i>et al.</i> , 2022)
NMDI3	$(B4-B6-B7)/(B4+B6+B7)$	(Han <i>et al.</i> , 2019)
NRVI	$(B4/B5-1)/(B4/B5+1)$	(Xue et Su, 2017)
OSAVI	$1.16*(B5-B4)/(B5+B4+0.16)$	(Rondeaux <i>et al.</i> , 1996)
RDVI	$(B5-B4)/\sqrt{B5+B4}$	(Roujean et Breon, 1995)
RVI	$B5/B4$	(Bannari <i>et al.</i> , 1995)
SAVI	$(B5-B4)/(B5+B4+0.5)*1,5$	(Huete, 1988)
STI	$B6/B7$	(Chaves <i>et al.</i> , 2020)
WBI	$(B2-B4)/(B2+B4)$	(Chaves <i>et al.</i> , 2020)

#### 4.3.4. Dates d'acquisition

Les dates d'acquisition des images Landsat-8 ont été regroupées en deux saisons, soit l'été ou l'automne. Ce traitement a été nécessaire pour ajouter une variable captant l'effet de la saison sur la phénologie. En effet, certaines images acquises à la fin septembre démontraient un changement de couleur dans les feuilles des arbres, affectant ainsi les valeurs d'indices spectraux. La saison de l'été a été définie entre le 1er juin et le 15 septembre et la saison d'automne avec toutes les dates suivant le 15 septembre. La variable de la semaine d'acquisition a également été calculée afin d'orienter la prédiction vers la semaine la plus chaude, donc la pire pour les îlots de chaleur. La Figure 14 montre un exemple dans la région de Montréal où deux images d'automne et d'été ont été utilisées. Cela démontre le contraste important que l'on peut retrouver entre les images et les zones.





Figure 14 : Extrait de la mosaïque pour la zone de Montréal avec des images d'automne (gauche) et d'été (droite).

#### 4.3.5. Position géographique et topographique

Les valeurs xy (longitude-latitude) ont été extraites par pixel en supposant que selon la position géographique, l'écart de température relatif de surface est plus ou moins élevé. L'altitude a été extraite à partir des modèles numériques de terrain SRTM et RNCan.

#### 4.3.6. Topologique

Les variables de type topologique sont essentiellement des variables de distance euclidienne. On y retrouve quatre sous-groupes, soit la distance 1) aux lacs (incluant les mers et le fleuve), 2) aux lacs mais en y ajoutant les rivières de type surfacique, 3) de voisinage et 4) des variables optiques. Le 3<sup>e</sup> sous-groupe représente une analyse des superficies d'eau surfacique dans le voisinage (distances de 15, 30 et 45 m autour du pixel

focal). Le 4<sup>e</sup> sous-groupe est une distance euclidienne par rapport aux variables d'occupation du territoire et du démixage spectral. L'ajout des sous-groupes 3 et 4, par rapport à l'étude de Boulfroy *et al.* (2013), a été réalisé afin de tenter de mesurer l'effet de la transmission de la chaleur en fonction de la proximité aux infrastructures anthropiques (Guo *et al.*, 2015).

#### 4.4. Échantillonnage des températures

Une seule image Landsat-8 ne peut couvrir à elle seule l'écoumène entier du Québec. Un total de 46 images de température a été nécessaire. Le Tableau 3 liste les images utilisées ainsi que la température du point de contrôle associé à chaque image. Ces images étant prises à différentes dates et latitudes, les températures observées peuvent varier pour un même territoire. Afin de minimiser ces différences, un point de contrôle a été géoréférencé manuellement pour chacune des images. Ces points de contrôle ont été pris dans des contextes forestiers représentatifs en milieux ruraux, où l'influence des infrastructures humaines sur la température de surface est négligeable ou nulle, et ont été réutilisés entre les images superposées lorsque possible afin d'éviter l'introduction de biais éventuels. Pour chaque image, un écart de température relatif a ensuite été calculé (Nadizadeh Shorabeh *et al.*, 2020) en soustrayant la température de surface à celle prise au point de contrôle, ce qui permet de normaliser les valeurs (Figure 15). Cette méthode permet de corriger le « biais d'indépendance des images » mentionné par Barette *et al.* (2018). Cette approche permet également de mieux représenter la définition d'un îlot de chaleur urbain : « [cette expression] signifie la différence de température observée entre les milieux urbains et les zones rurales environnantes » (Giguère, 2009 ; Guo *et al.*, 2015).

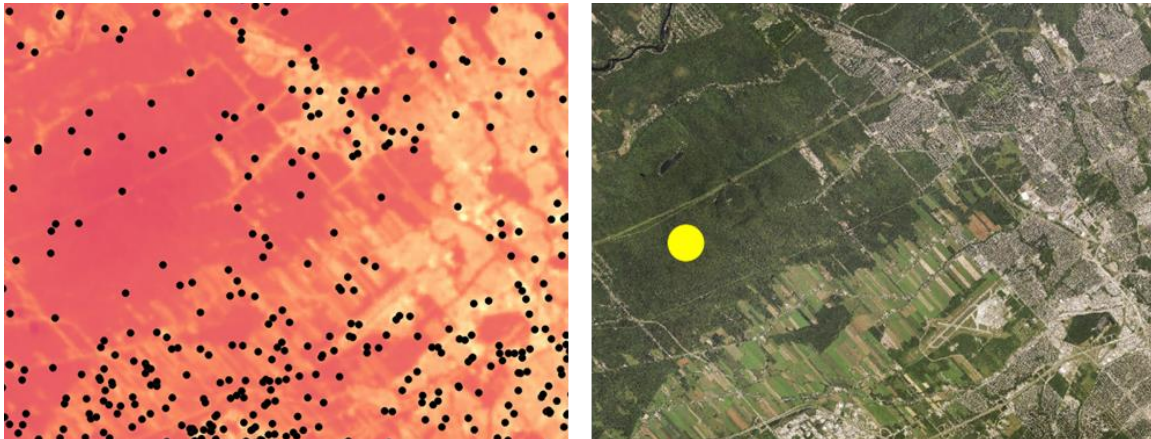


Figure 15 : Exemple de points échantillonnés (points noirs) et de référence (point jaune) pour normaliser la température.

Tableau 3 : Description des images Landsat-8 utilisées pour extraire la température de surface.

Image 2013	Année	Mois	Jour	Température du point de contrôle
LC08_L2SP_022016_20130727_20200912_02_T1	2013	7	27	24,6
LC08_L2SP_021019_20130618_20200912_02_T1	2013	6	18	20,9
LC08_L2SP_021019_20130922_20200912_02_T1	2013	9	22	8,3
LC08_L2SP_020023_20130729_20200912_02_T1	2013	7	29	21,7
LC08_L2SP_020023_20130627_20200912_02_T1	2013	6	27	28,1
LC08_L2SP_020022_20130627_20200912_02_T1	2013	6	27	28,9
LC08_L2SP_020021_20130611_20200912_02_T1	2013	6	11	31,3
LC08_L2SP_019026_20130924_20200913_02_T1	2013	9	24	13,5
LC08_L2SP_019024_20130823_20200913_02_T1	2013	8	23	17,3
LC08_L2SP_018028_20130917_20200912_02_T1	2013	9	17	13,8
LC08_L2SP_018027_20130917_20200912_02_T1	2013	9	17	12,8
LC08_L2SP_018026_20130917_20200913_02_T1	2013	9	17	12,9
LC08_L2SP_018025_20130917_20200912_02_T1	2013	9	17	13,9
LC08_L2SP_017028_20130809_20200912_02_T1	2013	8	9	22,3
LC08_L2SP_017025_20130708_20200912_02_T1	2013	7	8	19,4
LC08_L2SP_016028_20130818_20200913_02_T1	2013	8	18	23,5
LC08_L2SP_016027_20130818_20200912_02_T1	2013	8	18	21,6
LC08_L2SP_016026_20130701_20200912_02_T1	2013	7	1	12,2
LC08_L2SP_016025_20130701_20200912_02_T1	2013	7	1	12,8
LC08_L2SP_015029_20130726_20200912_02_T1	2013	7	26	23,8
LC08_L2SP_015028_20130928_20200912_02_T1	2013	9	28	18,9
LC08_L2SP_015028_20130726_20200912_02_T1	2013	7	26	23,8
LC08_L2SP_015019_20130608_20200912_02_T1	2013	6	8	22,5
LC08_L2SP_014029_20130820_20200913_02_T1	2013	8	20	26,4
LC08_L2SP_014028_20130820_20200913_02_T1	2013	8	20	24,7
LC08_L2SP_014027_20130820_20200912_02_T1	2013	8	20	26,6
LC08_L2SP_014026_20130703_20200912_02_T1	2013	7	3	26,7
LC08_L2SP_013028_20130712_20200912_02_T1	2013	7	12	22,4
LC08_L2SP_013027_20130712_20200912_02_T1	2013	7	12	22,3
LC08_L2SP_013026_20130712_20200912_02_T1	2013	7	12	23,6
LC08_L2SP_013023_20130626_20200912_02_T1	2013	6	26	21,9
LC08_L2SP_013022_20130626_20200912_02_T1	2013	6	26	22,3
LC08_L2SP_012027_20130721_20200912_02_T1	2013	7	21	16,7
LC08_L2SP_012026_20130806_20200912_02_T1	2013	8	6	20,4
LC08_L2SP_012023_20130705_20200912_02_T1	2013	7	5	27,8
LC08_L2SP_011027_20130714_20200912_02_T1	2013	7	14	31,6
LC08_L2SP_011026_20130815_20200913_02_T1	2013	8	15	18,6
LC08_L2SP_011025_20130628_20200912_02_T1	2013	6	28	17,4
LC08_L2SP_010027_20130824_20200913_02_T1	2013	8	24	21,1



LC08_L2SP_010026_20130824_20200913_02_T1	2013	8	24	20,0
LC08_L2SP_010025_20130824_20200912_02_T1	2013	8	24	21,0
LC08_L2SP_009026_20130716_20200912_02_T1	2013	7	16	25,3
LC08_L2SP_009025_20130918_20200912_02_T1	2013	9	18	15,7
LC08_L2SP_008025_20130709_20200912_02_T1	2013	7	9	22,8
LC08_L2SP_007027_20130718_20200912_02_T1	2013	7	18	22,6
LC08_L2SP_007025_20130702_20200912_02_T1	2013	7	2	24,0

Pour chaque image Landsat-8 de température, la bande de qualité (QA\_PIXEL) de l'image Landsat-8 fournie dans la collection a été utilisée pour identifier puis éliminer les zones couvertes par les nuages et leurs ombres. Une dilatation du masque de 33 pixels (environ 1 km) a été appliquée sur le masque afin de s'assurer qu'on ne retrouve pas de nuages ou ombres résiduelles aux limites du masque. Une sélection aléatoire d'un maximum de 1 000 pixels par classe d'occupation du territoire et par zone d'analyse a été effectuée afin d'extraire l'écart de température relatif. Un total de 38 196 observations a été considéré pour élaborer le modèle.

#### 4.5. Modélisation

Étant donné que les observations sélectionnées aléatoirement n'ont pas une répartition spatiale homogène, plusieurs tests ont été effectués pour séparer ces points entre les jeux de données d'entraînement et de validation de manière adéquate. D'abord, le calcul de l'indice de Moran a été réalisé pour identifier la présence d'autocorrélation spatiale (Moran, 1948). Ensuite, pour pallier la présence d'autocorrélation spatiale, une analyse de la répartition spatiale des points a été réalisée pour identifier des *clusters* (Chalghaf *et al.*, 2019). Enfin, une sélection aléatoire a été réalisée par domaine bioclimatique, classe d'occupation du territoire et par *cluster*, permettant ainsi d'attribuer au centroïde de chaque polygone d'occupation du territoire un jeu de référence (entraînement – 80 % ou validation – 20 %). Une sélection des 66 variables a ensuite été effectuée en enlevant les variables corrélées à plus de 90 % selon l'indice de corrélation de *spearman*, laissant ainsi 32 variables (Tableau 4 et Figure 16). Par la suite, un modèle de régression basé sur l'apprentissage automatique avec l'algorithme *Random Forest* (Breiman, 2001) a été développé afin de prédire l'écart de température relatif. Une recherche des paramètres optimaux a été effectuée en minimisant l'erreur quadratique moyenne (RMSE). Les hyperparamètres optimaux sont présentés dans le Tableau 5.

Tableau 4 : Liste des variables utilisées pour entrainer l'algorithme *Random Forest* pour le modèle prédictif de l'écart de température relatif.

Variable	Description
ABDI1_B5_B6	Indice spectral
ALT	Altitude
Construction	Probabilité d'appartenance à la classe « Construction » (démixage)
D_EAU_LIN	Distance à l'eau (linéaire)
D_EAU_SURF	Distance à l'eau (surfactive)
DistanceDemix_Construction	Distance aux zones bâties (démixage)
DistanceDemix_Eau	Distance aux pixels de la classe « Eau » (démixage)
DistanceDemix_Sol_nu	Distance aux pixels de la classe « Sol nu » (démixage)
DistanceDemix_vegetation	Distance aux pixels de la classe « végétation » (démixage)
DistanceOccTerr_EauLac	Distance aux pixels de la classe « EauLac » (modèle occupation du territoire)
DistanceOccTerr_SolAgricole	Distance aux pixels de la classe « SolAgricole » (modèle occupation du territoire)
DistanceOccTerr_UrbainMixte	Distance aux pixels de la classe « UrbainMixte » (modèle occupation du territoire)
DistanceOccTerr_UrbainPur	Distance aux pixels de la classe « UrbainPur » (modèle occupation du territoire)
DistanceOccTerr_VegetationArborescente	Distance aux pixels de la classe « VégétationArborescente » (modèle occupation du territoire)
DistanceOccTerr_VegetationBasse	Distance aux pixels de la classe « VégétationBasse » (modèle occupation du territoire)
Eau	Probabilité d'appartenance à la classe de démixage « Eau »
EauVoisinage	Superficies d'eau surfactive dans le voisinage (15 m)
EauVoisinage5x5	Superficies d'eau surfactive dans le voisinage (30 m)
EauVoisinage7x7	Superficies d'eau surfactive dans le voisinage (45 m)
GI	Indice spectral
img_index	Saison (été/automne)
MBWI	Indice spectral
MNDWI	Indice spectral
NBR	Indice spectral
NDVI	Indice spectral
NMDI3	Indice spectral
OccTerr	Classe d'occupation du territoire tel que prédit par le modèle d'occupation du territoire
Semaine	Semaine d'acquisition de l'image de température
Sol_nu	Probabilité d'appartenance à la classe « Sol nu » (démixage)



Vegetation	Probabilité d'appartenance à la classe « Végétation » (démixage)
x	Longitude
y	Latitude

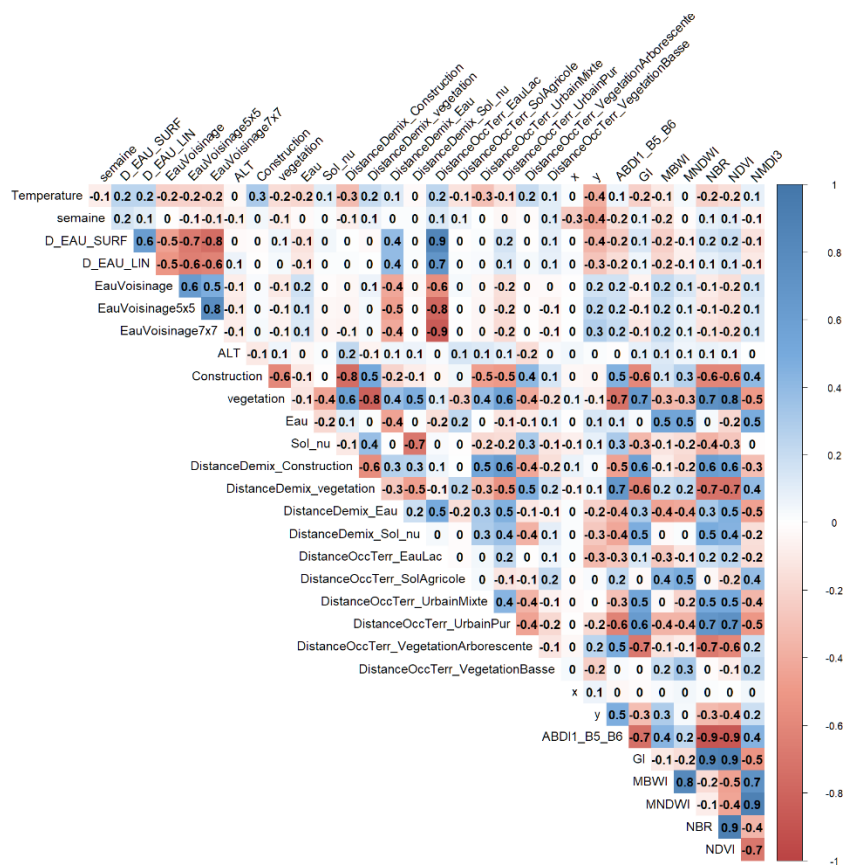


Figure 16 : Corrélogramme des variables utilisées pour entraîner l'algorithme *Random Forest* ainsi que l'écart de température observé (Temperature).

Tableau 5 : Liste des hyperparamètres optimaux utilisés pour entraîner l'algorithme *Random Forest*.

Description	Paramètre
Nombre d'arbres de régression	2000
Nombre de variables échantillonnées au hasard comme candidates à chaque nœud	31
Taille de nœud minimale	5

#### 4.6. Prédiction et classement des valeurs

Une fois le modèle entraîné et validé, il a été utilisé pour prédire l'écart de température relatif sur tous les pixels de l'écoumène du Québec. Cette donnée a ensuite été restreinte aux centres de population ayant au minimum 400 habitants par km<sup>2</sup> avec une zone tampon de 2 km. Les valeurs de cette donnée restreinte ont ensuite été classées en 9 classes par l'algorithme *k-means* (Montolio *et al.*, 1992), après avoir testé différentes méthodes (quantiles, bris naturels, *k-means*, *b-clust*, intervalles égaux) utilisées dans plusieurs études sur les ICFU (Tableau 6). Ce choix a été décidé par suite d'un atelier (24 août 2022) avec des experts de diverses spécialités (statistique, géomatique, santé publique, environnement, foresterie). Cet algorithme est également similaire à la technique des bris naturels utilisés dans des travaux antérieurs (Bouffroy *et al.*, 2013; Varin *et al.*, 2016), mais a l'avantage d'être plus rapide à calculer. Diverses études portant sur le confort humain proposent également d'utiliser 9 classes (Epstein et Moran, 2006 ; Goldblatt *et al.*, 2021). Les deux premières et trois dernières classes sont généralement choisies pour représenter les îlots de fraîcheur et de chaleur, respectivement. L'utilisation de ces classes est abordée dans la section 6. La Figure 17 montre la variation dans l'identification des îlots de chaleur suite à la catégorisation en 9 classes en utilisant ces différentes méthodes. Ce classement a été réalisé indépendamment pour chaque centre de population. Les valeurs d'intervalles pour chaque classe et chaque centre de population de diffusion sont présentées dans un index.

Tableau 6 : Méthodes de classement des valeurs pour définir les îlots de chaleur et de fraîcheur urbains de différentes études réalisées au Québec.

Source et date de l'étude	Période d'acquisition des données	Méthode	Échelle spatiale	Nb de classes
CERFO 2012 <sup>1</sup>	2006–2011	Bris naturels	20 m	7 et 9
CERFO 2016 <sup>2</sup>	2013	Bris naturels	15 m	7
Ville de Montréal 2022 <sup>3</sup>	2016	Moyennes emboîtées et écarts-types	25 m	5
Radio-Canada 2022 <sup>4</sup>	2019–2020	Moyenne pondérée des quantiles	Aire de diffusion	10
Uaval 2018 <sup>5</sup>	2013	Bris naturels	30 m	10

<sup>1</sup>Bouffroy *et al.* (2013), <sup>2</sup>Varin *et al.* (2016), <sup>3</sup>Ville de Montréal (2022), <sup>4</sup>Schiab et Bouchard (2022), <sup>5</sup>Barrette *et al.* (2018)

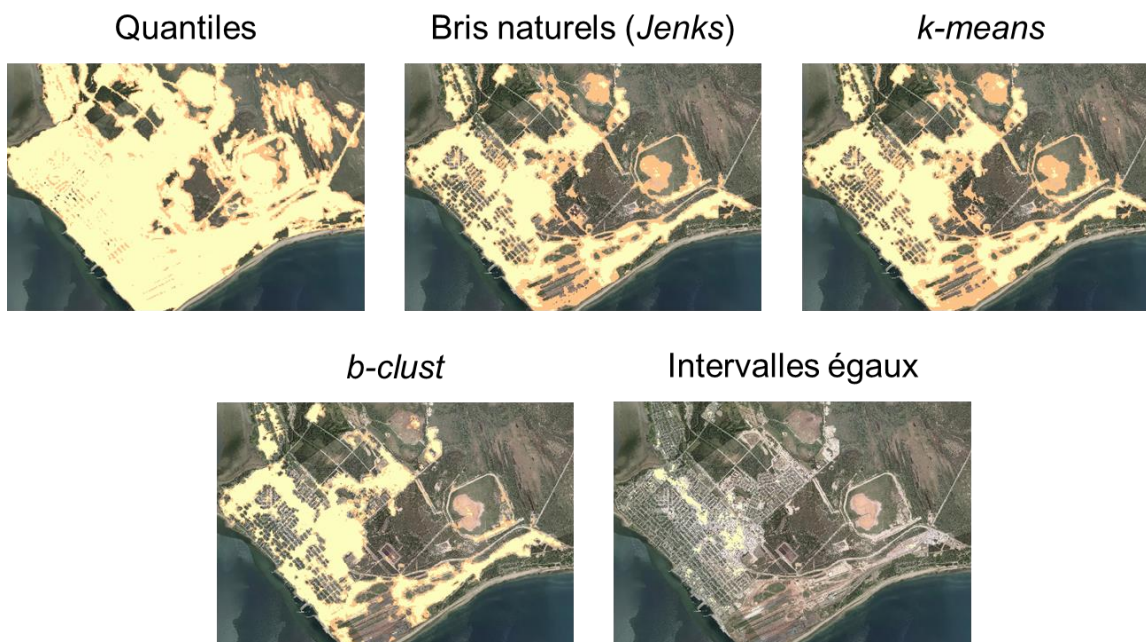


Figure 17 : Variation des îlots de chaleur (beige-jaune) prédits selon différentes méthodes de classement des valeurs.

## 5. Résultats

### 5.1. Variables utilisées

Une analyse statistique a permis d'évaluer la valeur explicative des différentes variables à l'aide de permutations. La latitude est ainsi la variable qui explique le plus les variations des écarts de température relatifs observés, suivi de la semaine d'acquisition de l'image et de la distance aux pixels de la classe « végétation arborescente » (Tableau 7). Ce sont 36 variables qui ont été utilisées dans le modèle.

Tableau 7 : Importance des variables.

Variable	Importance (%)
y	100.00
semaine	42.04
DistanceOccTerr_VégétationArborescente	24.48
X	14.93
ALT	13.64
DistanceDemix_Construction	10.53
GI	7.21
NBR	6.92
MBWI	5.13

DistanceOccTerr_UrbainPur	4.27
Construction	4.08
DistanceDemix_Sol_nu	3.25
ABDI1_B5_B6	3.16
DistanceOccTerr_SolAgricole	2.87
DistanceOccTerr_EauLac	2.32
DistanceOccTerr_VegetationBasse	2.22
DistanceOccTerr_UrbainMixte	2.19
MNDWI	2.03
NDVI	1.89
D_EAU_LIN	1.79
vegetation	1.73
NMDI3	1.53
Sol_nu	1.28
D_EAU_SURF	1.16
EauVoisinage5x5	1.09
DistanceDemix_Eau	0.99
DistanceDemix_vegetation	0.84
EauVoisinage7x7	0.67
Eau	0.63
EauVoisinage	0.56
OccTerrVegetationArborescente	0.40
img_indexAUTOMNE	0.30
OccTerrUrbainPur	0.05
OccTerrUrbainMixte	0.02
OccTerrSolAgricole	0.01
OccTerrVegetationBasse	0.00

Le Tableau 8 illustre quelques exemples des variables calculées. La relation entre les variations des écarts de température relatifs observés et les variables utilisées pour entraîner l'algorithme *Random Forest* est illustrée par la Figure 18. Cela permet de constater que les relations ne sont pas toujours linéaires, d'où l'intérêt d'utiliser un algorithme basé sur des arbres décisionnels.

## Tableau 8 : Exemples de variables calculées

Landsat-8

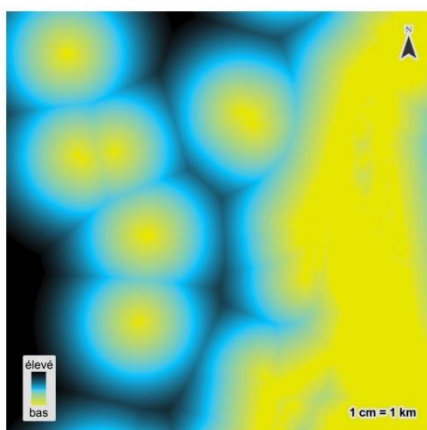
Imagerie QC



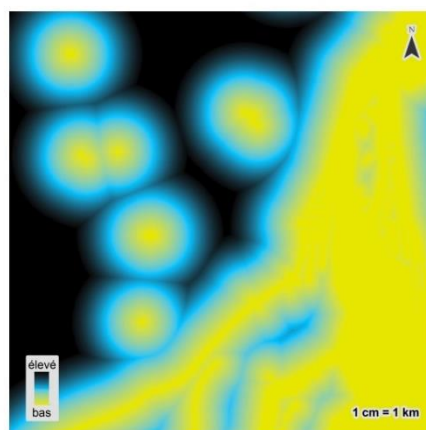
Distance eau, lac, fleuve et mer



Distance eau, lac, fleuve, mer et rivière

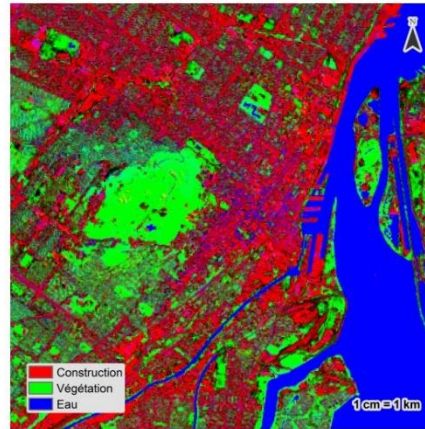
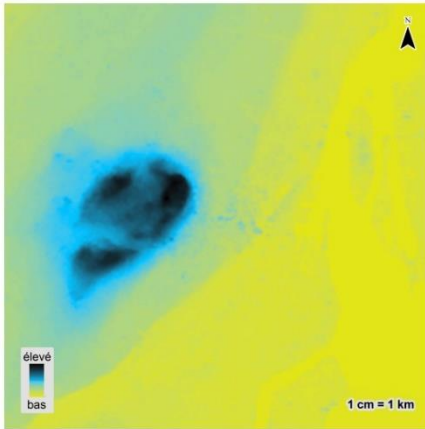


Modèle numérique de terrain



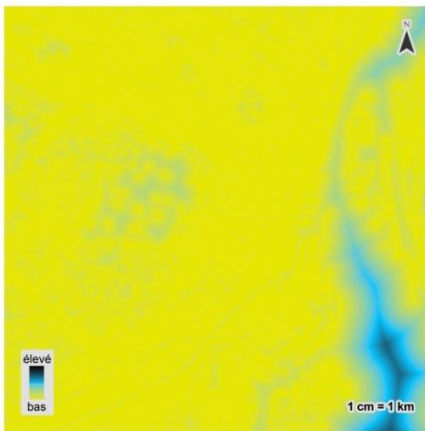
Démixage avec masque d'eau



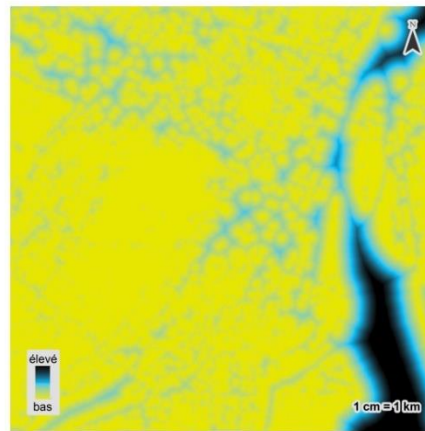


Distance aux classes de démixage

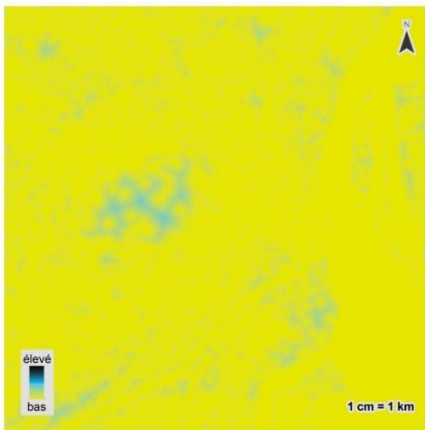
Classe 1 - Construction



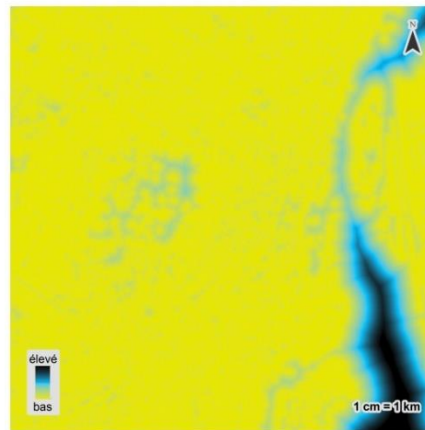
Classe 2 - Végétation



Classe 3 - Eau



Classe 4 - Sol nu

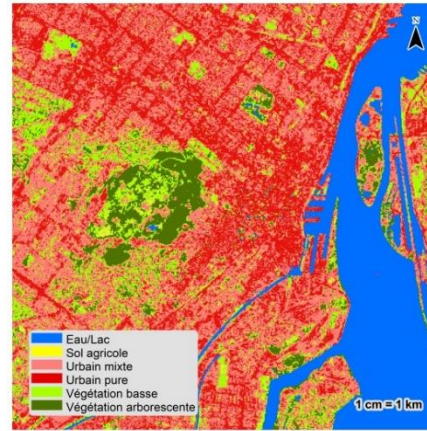


Eau voisinage



Occupation du territoire

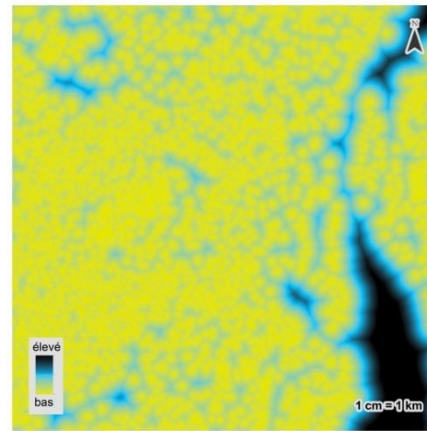
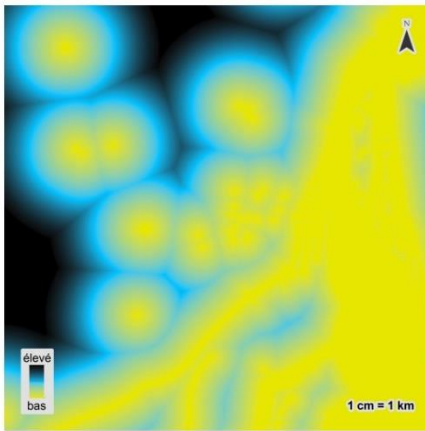




Distance aux classes d'occupation du territoire

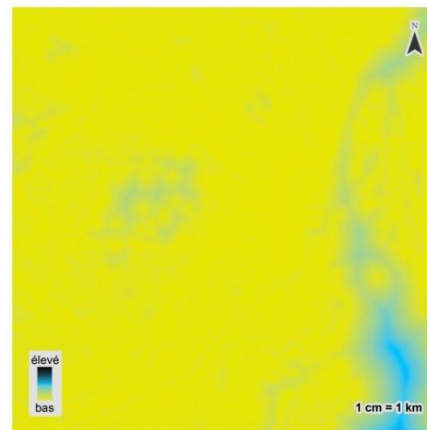
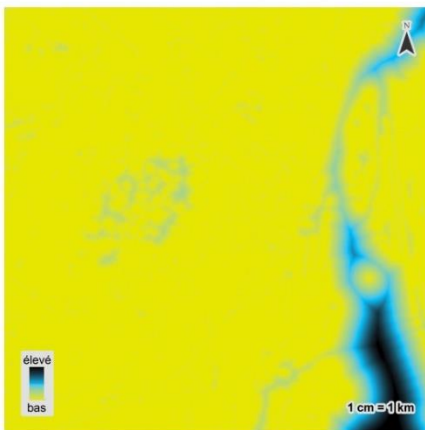
Classe 1 – Eau/Lac

Classe 2 – Sol agricole



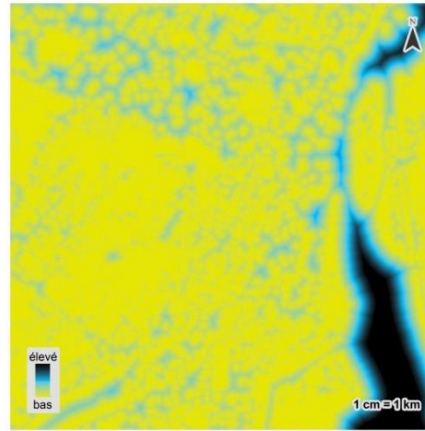
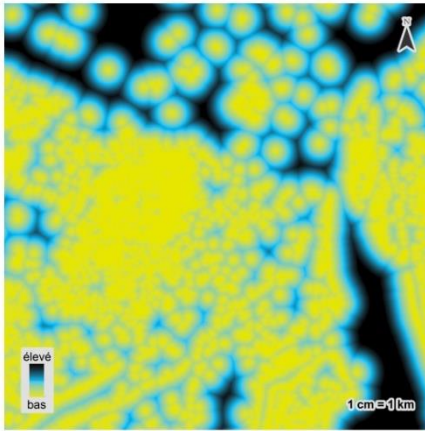
Classe 3 – Urbain mixte

Classe 4 – Urbain pur



Classe 5 – Végétation arborescente

Classe 6 – Végétation basse





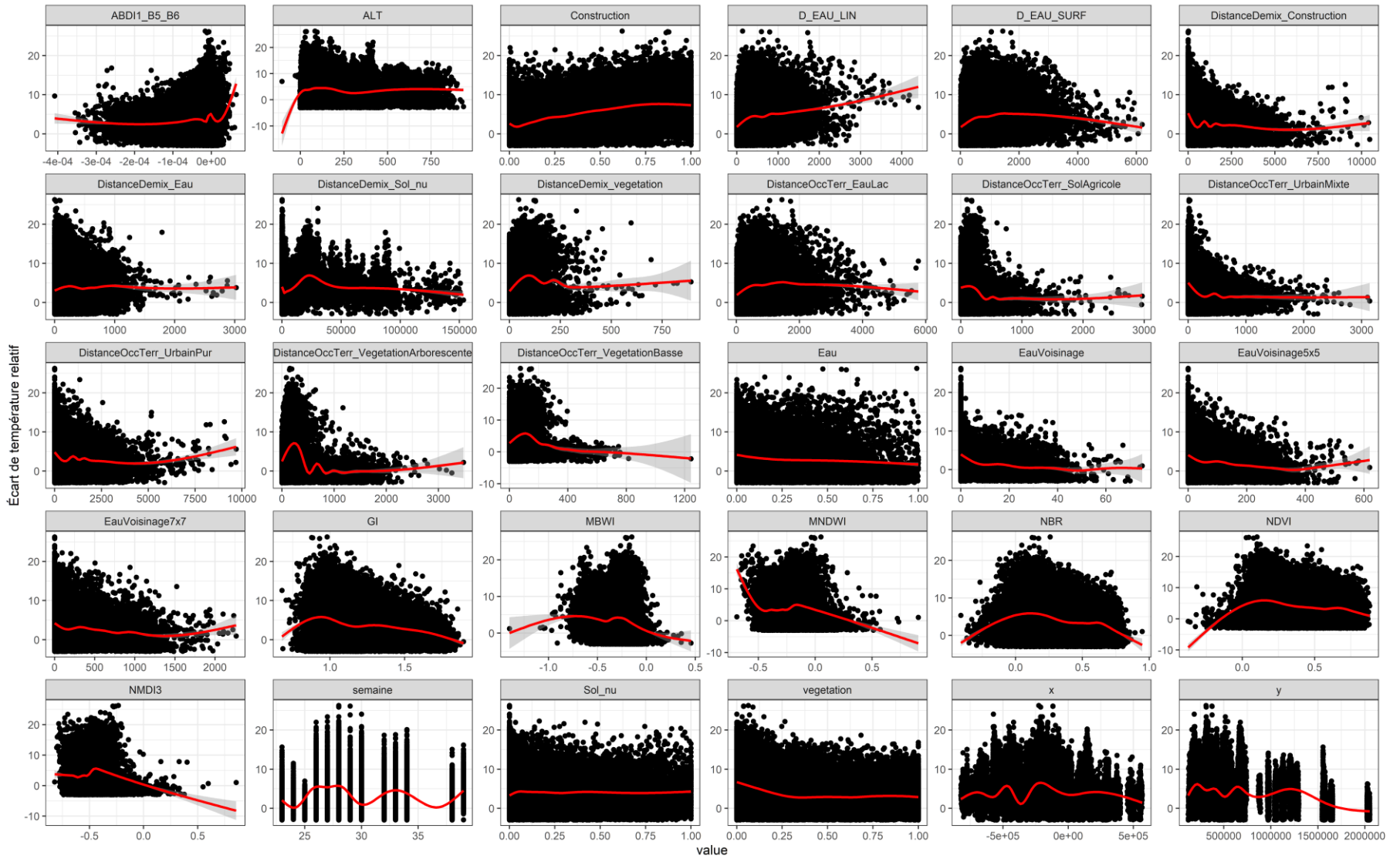


Figure 18 : Relation entre l'écart de température relatif en fonction de chacune des variables utilisées pour entraîner l'algorithme *Random Forest*.

## 5.2. Prédiction de l'écart de température relatif prédit

Le RMSE pour le jeu de données de validation a été estimé à 1,99 °C de différence de température. La distribution des échantillons du jeu de données de validation montre un  $R^2$  de 0,81, ce qui signifie que dans 81 % des cas, l'écart de température relatif prédit concorde avec celui issu des données de température du USGS calculées à partir de Landsat-8. Ce résultat illustre la cohérence globale entre les données prédites et réelles. Il confirme également que des différences existent entre les deux données, ce qui répond à l'objectif ciblé du projet. En effet, on souhaitait augmenter la précision des valeurs de température en générant une donnée prédite à une résolution spatiale plus fine (15 m) que celle de la donnée du USGS (30 m). La Figure 19 représente la distribution des observations du jeu de données de validation avec l'écart de température relatif réel et prédit.

La Figure 20 illustre un exemple visuel de température de Landsat-8 (données USGS) et la prédiction de l'écart de température relatif. On remarque que cette méthode permet de mettre davantage en évidence les différences de température, en lien avec les différentes occupations du territoire visibles sur l'imagerie.

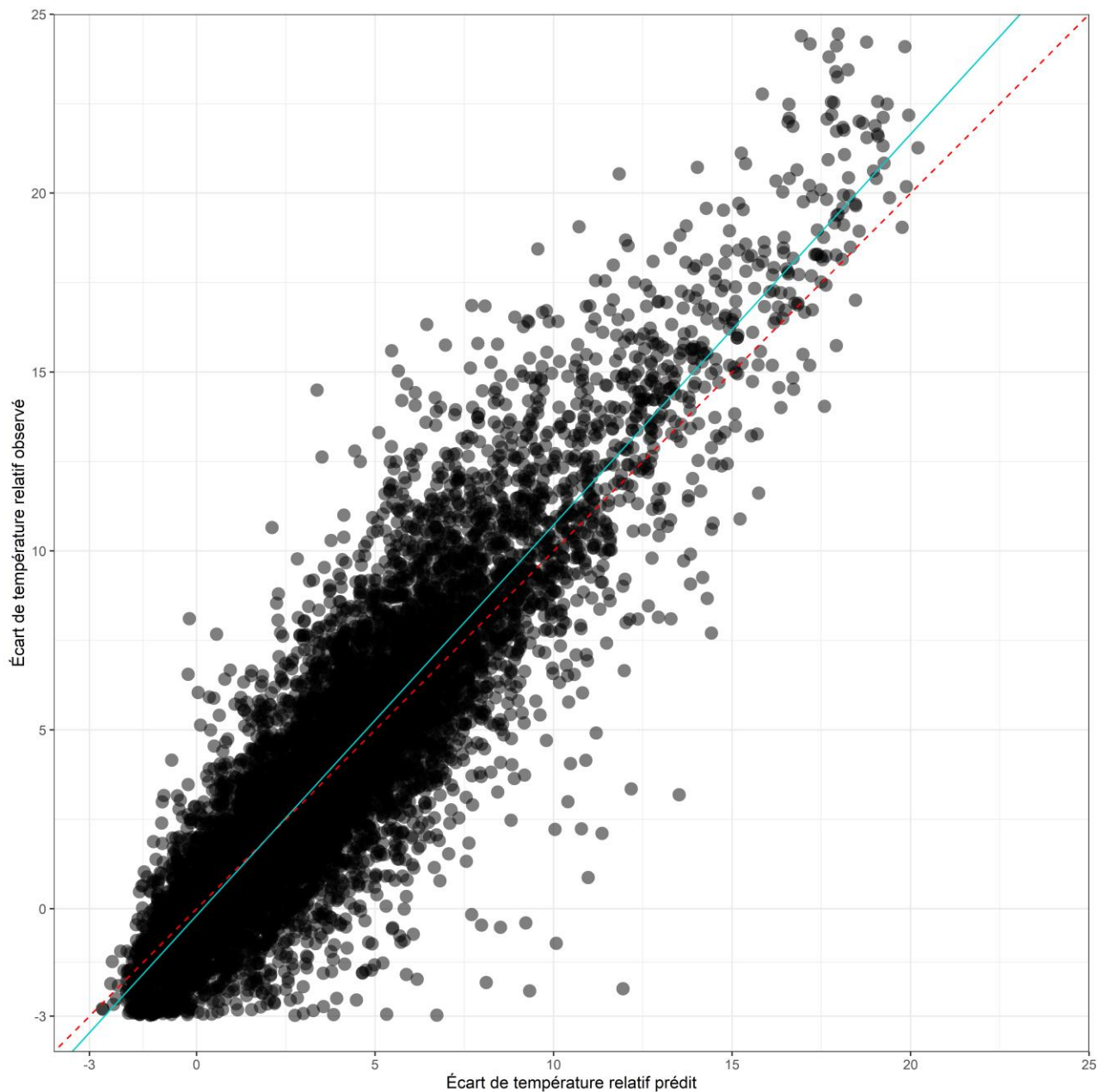


Figure 19 : Écart de température relatif Landsat-8 observé en fonction de l'écart de température prédit pour le jeu de données de validation. La ligne de régression (bleue et solide) et l'hypothèse nulle (ordonnée à l'origine 0, pente 1 ; ligne pointillée rouge) sont également affichées.

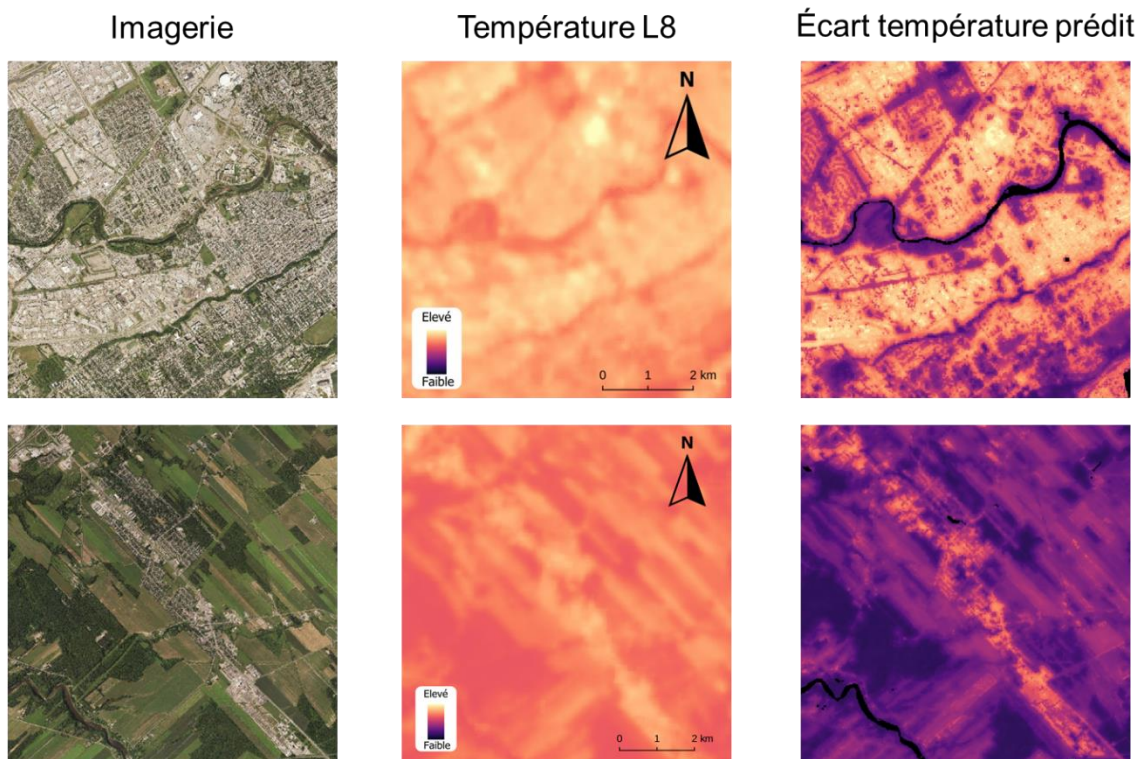


Figure 20 : Exemple d'écart de température relatif prédit par le modèle (droite) et comparaison avec la température issue de Landsat-8 calculée par le *United States Geological Survey* (centre). Fond de carte Google à gauche.

### 5.3. Catégorisation des îlots de chaleur et de fraîcheur urbains

La méthode utilisée pour la catégorisation des ICFU est le classement en 9 classes des écarts de température relatifs prédits. Ceci permet de discriminer les zones ayant un fort potentiel d'être des îlots de chaleur (dernière classe) des îlots de fraîcheur (deux premières classes). Sur l'ensemble du territoire d'étude, ce sont ainsi 4 764 km<sup>2</sup> et 1 046 km<sup>2</sup> qui représentent des îlots de chaleur et de fraîcheur, respectivement. Compte tenu des différences observées dans les écarts de températures relatifs prédits entre le Nord (-1,6 à 2 °C) et le Sud du Québec (0,2 à 17,8 °C), il est important d'interpréter ces classes en relation avec les intervalles d'écarts de températures qui y sont associés. Ainsi, dans certains contextes, il peut être pertinent d'utiliser également l'avant-dernière (8) classe dans la représentation des îlots de chaleur pour un territoire donné.

La Figure 21 illustre le résultat de la catégorisation en 9 classes dans deux contextes différents, soit une zone fortement urbanisée et une zone composée d'une mosaïque de milieux urbanisés, de champs et de forêts. Dans le cas présent, les écarts de température relatifs maximaux varient de 14,1 °C dans le cas du contexte rural comparé à 17,8 °C dans la zone fortement urbanisée.

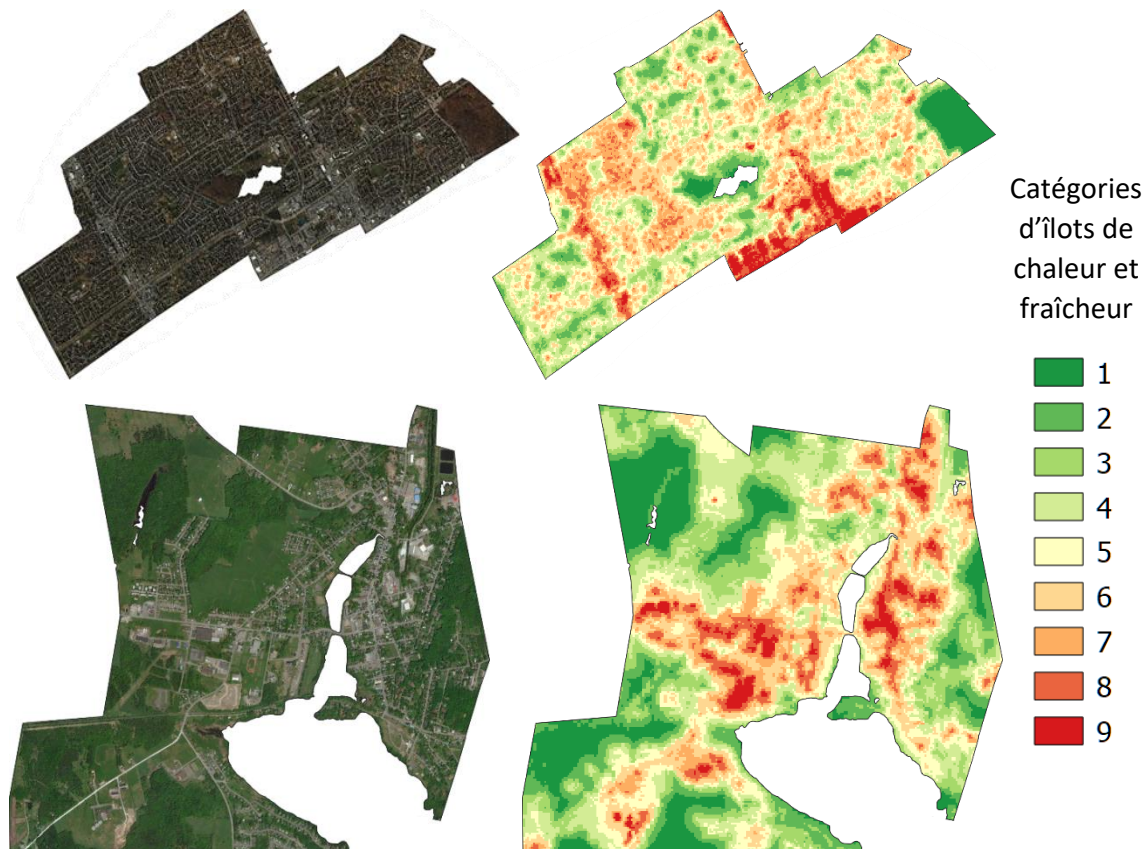


Figure 21 : Exemples de classement des îlots de chaleur et de fraîcheur urbains en milieu fortement urbanisé (haut) et en milieu rural (bas). Les îlots de chaleur semblent plus diffus en milieu rural.

## 6. Discussion

### 6.1. Performances générales

Plusieurs constats peuvent être relevés pour donner suite à l'analyse des résultats :

- L'écart de température relatif prédit est généralement très bon. Les écarts de températures sont réalistes et correspondent bien à la donnée de référence du USGS. On observe également que l'écart observé au niveau de l'écart de



- température relatif est plus grand dans le Sud du Québec que dans le Nord, ce qui transparait au niveau de la latitude qui explique en grande partie cet écart. Par exemple, dans la région de Montréal, on peut observer un écart jusqu'à 17 °C alors qu'à Roberval, on observe plutôt un écart de 12 °C. Ce phénomène apparaît à la fois dans la prédiction et la donnée de référence ;
- Les mosaïques étant composées de plusieurs dates d'acquisition d'images, des contrastes importants y sont visibles. Néanmoins, la donnée prédite semble moins affectée par cette problématique (Figure 22) ;
  - Le modèle est également robuste, en étant appliqué sur l'ensemble du territoire d'étude, obtenant un  $R^2$  de 0,81, soit un gain de 14 % par rapport à Boulfroy *et al.* (2013). Ce type de modèle fonctionnant à l'aide de variables physiques pourrait s'appliquer à une résolution spatiale beaucoup plus fine (p. ex. 5, 3 voire 1 m). Il suffit dans ce cas de calculer les variables nécessaires à la résolution voulue. Cela assure également une reproductibilité et une comparaison diachronique pertinente ;
  - Bien que le modèle ait été créé avec une approche légèrement différente au niveau des variables et du type de modèle, les résultats sont relativement comparables avec ceux produits par Boulfroy *et al.* (2013). Les résultats du présent projet démontrent en général moins de sous-estimation des îlots de chaleur et sont aussi plus diffus dû aux variables topologiques utilisées et à la résolution spatiale plus grossière (15 m au lieu de 10 m) (
  - Figure 23).

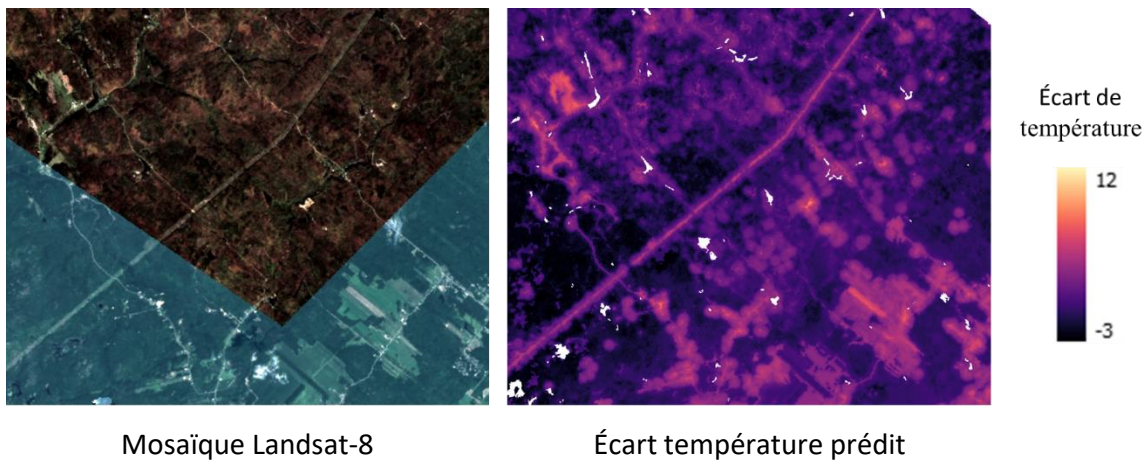


Figure 22 : Exemple de contraste présent sur une mosaïque et écart de température relatif prédit.

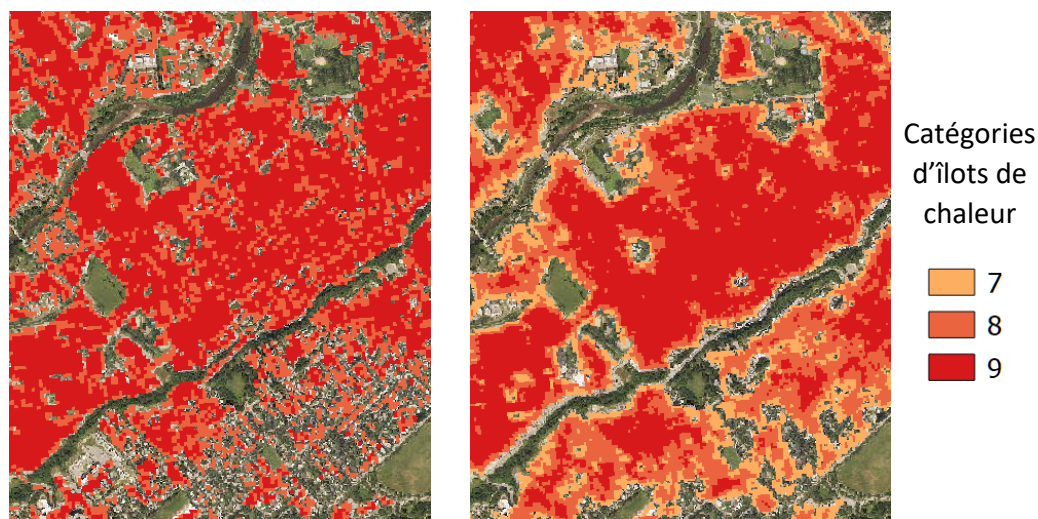


Figure 23 : Comparaison entre les îlots de chaleur tels que produits par Boulfroy *et al.* (gauche) et ceux produits dans ce projet (droite).

## 6.2. Limites

Plusieurs limites de la méthode et des résultats peuvent être notées :

- Il est essentiel de dissocier la température de surface de la température de l'air ambiante pouvant être mesurée par les stations météorologiques. En effet, la température de surface fait plutôt référence à une température estimée au niveau du sol, sur les bâtiments ou la canopée. Celle-ci est surtout influencée par le type de surface (p. ex. les matériaux), mais est souvent supérieure à la température de l'air ambiante, malgré une forte corrélation positive entre les deux types de mesures (Cui et Foy, 2012 ; Good *et al.*, 2017 ; Soltani et Sharifi, 2017). Elle ne correspond donc pas à la température ressentie ou au confort des habitants ;
- Sable identifié comme une surface imperméable : lors de l'étape du démixage spectral, certaines plages détiennent une forte probabilité d'appartenance à la catégorie des constructions. Le sable représente une surface brillante et se confond donc avec certaines surfaces anthropiques. Cela aura comme effet de prédire une température plus chaude à ces endroits. Néanmoins, ces zones sont souvent exclues dans le masque hydrographique appliqué ;
- Présence de neige dans les secteurs du Nord : la neige présente un aspect brillant qui peut se confondre avec des surfaces anthropiques. À l'inverse, le démixage identifiera plutôt cet élément comme un mélange de plusieurs catégories. Ainsi, la température prédite pour cet élément, qui se retrouve surtout dans les creux de vallées, pourrait être erronée. La variable de l'altitude vient limiter cette erreur ;

- Forte présence de roc et d’affleurements rocheux dans les secteurs du Nord : le même constat que la présence de neige peut être établi. Généralement, ces secteurs devraient présenter une température prédite plus élevée que la moyenne (Figure 24) ;
- Présence de milieux humides : les milieux humides sont des zones transitoires entre les milieux terrestres et aquatiques. Ils n’ont pas été identifiés comme une classe ou catégorie en soi. Ainsi, dans la variable d’occupation du territoire, ils sont considérés comme soit de la végétation basse, arborescente ou du sol. Dans la variable du démixage spectral, ils forment un mélange des catégories Eau, Végétation et Sol nu. Ces milieux peuvent être prédits avec un écart de température relatif élevé (Figure 25) ;
- Certains champs agricoles sont identifiés dans la classe d’occupation du territoire « Urbain pur ». Comme ces champs sont très brillants, ils sont confondus avec les surfaces anthropiques (Figure 24). L’écart de température relatif prédit est donc relativement élevé. Le même constat a été identifié dans Boulfroy *et al.* (2013) ;
- Les coupes forestières récentes sont identifiées dans la classe d’occupation du territoire « Urbain pur ». Ces coupes apparaissent très brillantes et peuvent obtenir une température prédite relativement chaude (Figure 25) ;
- Toute méthode de classement des valeurs étant subjective, dépendante de la superficie de la zone d’intérêt et du contexte d’application (p. ex. milieux urbains contre milieux ruraux), il est important de bien comprendre leurs limites pour faciliter l’interprétation des résultats. Ainsi, les classes produites dans ce projet devraient être interprétées en prenant en compte les valeurs des écarts de températures relatifs utilisées pour faire le classement. Il est ainsi suggéré d’utiliser soit la dernière (9) ou les deux dernières classes (8-9) pour représenter les îlots de chaleur selon le territoire ciblé.

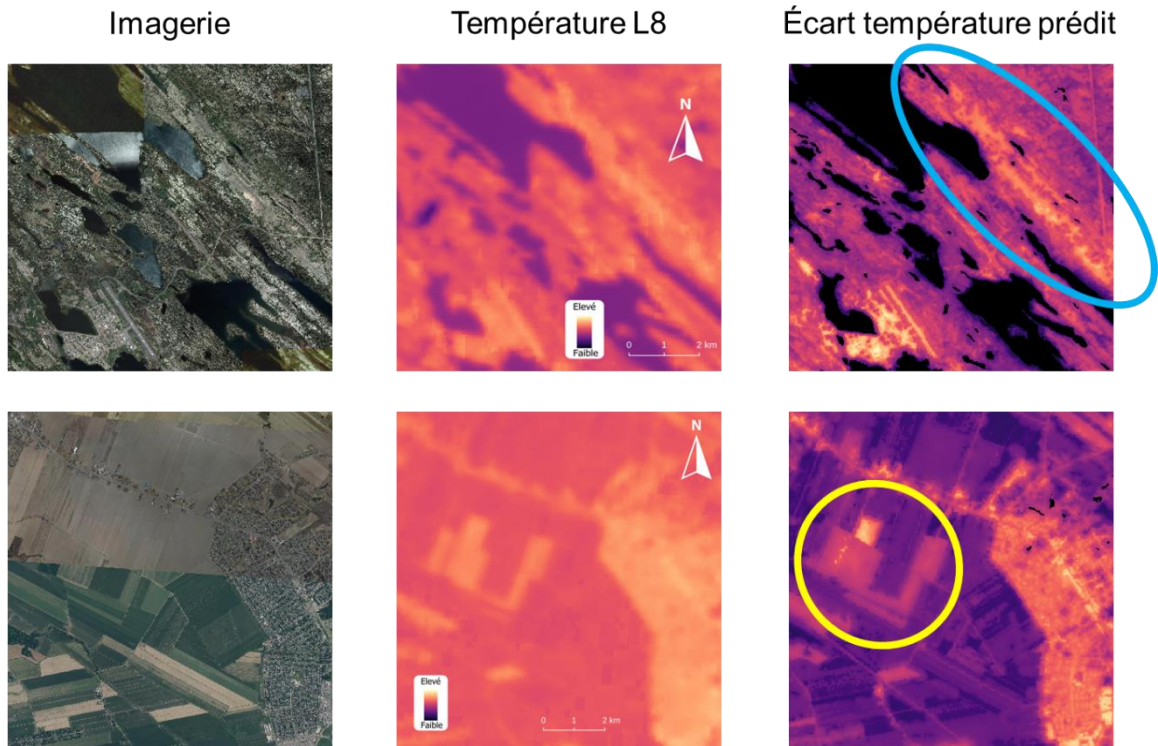


Figure 24 : Limites de prédiction de la température normalisée pour les champs agricoles (cercle jaune) et le roc (ellipse bleue), qui représenteront des îlots de chaleur.



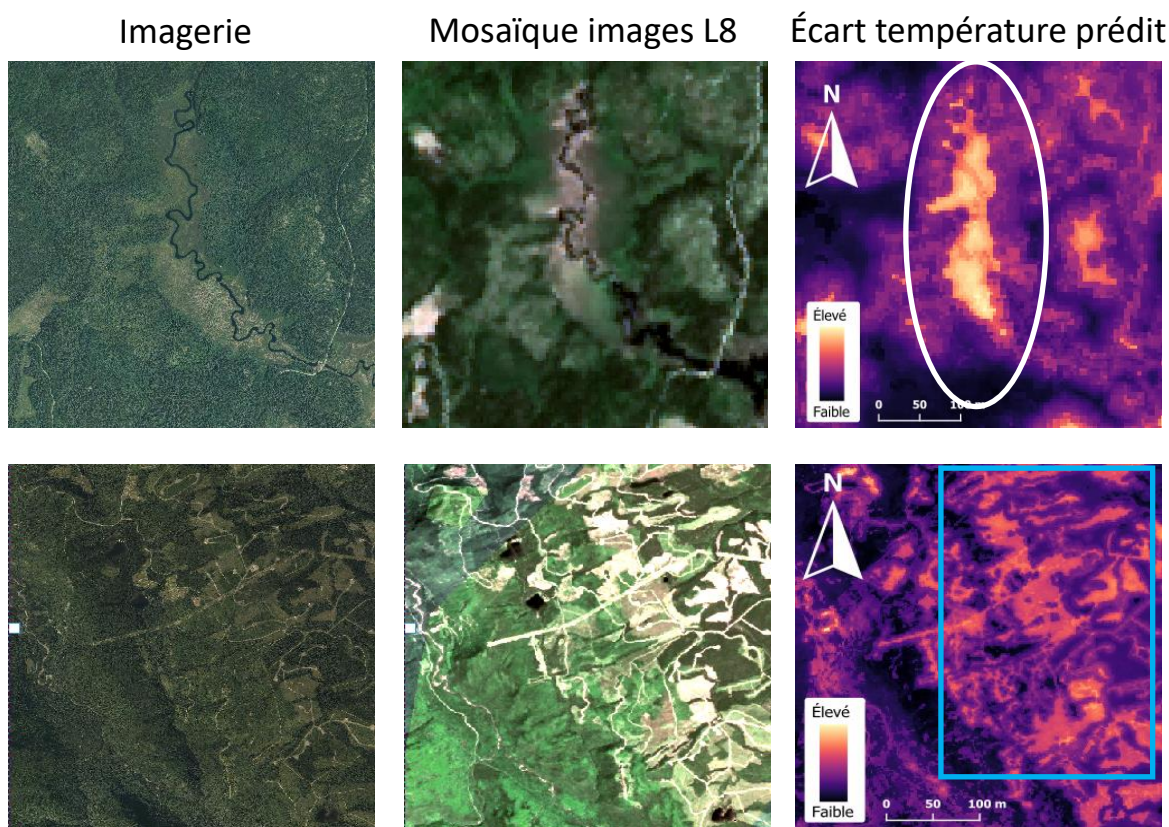


Figure 25 : Limites de prédiction de la température normalisée pour les milieux humides (ellipse blanche) et les coupes forestières (rectangle bleu).

### 6.3. Pistes d'amélioration

Une première piste d'amélioration est de simplifier la création de mosaïques pour couvrir tout le territoire d'étude. L'utilisation de GEE est limitée, notamment pour créer un masque de nuages et d'ombre de ceux-ci sans erreurs, et pour le choix des images lorsqu'il y a des superpositions. L'utilisation d'un modèle d'apprentissage profond pour créer ce masque permettrait d'atteindre une plus grande précision. Des tests réalisés par le CERFO ont démontré une meilleure performance en utilisant cette technique que celle proposée par GEE. Ainsi, utiliser une telle approche permettrait de mieux automatiser la création de mosaïques, nécessitant moins de manipulations manuelles.

La deuxième piste d'amélioration proposée réside dans la création des variables. Par exemple, l'occupation du territoire et l'indice d'imperméabilité pourraient également être



généralisés avec une meilleure précision en utilisant des algorithmes plus performants, tels que ceux disponibles en apprentissage profond (Bui *et al.*, 2021 ; Vali *et al.*, 2020). En effet, des gains de 6 % peuvent être observés entre *Random Forest* (algorithme utilisé dans ce projet) et un réseau de neurones par convolution (Kussul *et al.*, 2017). Cependant, il faut noter que ces techniques plus performantes nécessitent beaucoup de données de référence, il serait alors nécessaire d'adapter des banques de données existantes avec les besoins du projet ou de bonifier au besoin celles déjà créées pour des projets similaires par le passé. Également, d'autres variables pourront être ajoutées à la modélisation, telles qu'un indice d'ombrage en fonction de l'ensoleillement et des structures verticales urbaines (Lindsay, 2014).

Une troisième piste d'amélioration serait de réaliser l'analyse de classement des valeurs à l'échelle d'autres limites administratives d'intérêt telles que le périmètre urbain. Ceci permettrait d'obtenir un contraste plus adapté au niveau de la discrimination des ICFU pour ces territoires d'intérêt. La diffusion de la donnée d'écart de température relatif permet déjà aux différentes organisations de faire leur propre classement à l'échelle d'intérêt de leur choix. Toutefois, cela reste une opération qui peut être complexe.

## 7. Conclusion

Le but du projet était de produire une cartographie des ICFU dans l'écoumène (centres de population de 2021 ayant un minimum de 400 habitants par km<sup>2</sup>) du Québec de 2016 pour un horizon 2013-2014 afin de mettre à jour la cartographie actuelle datant de 2005-2011 pour que l'INSPQ ait en main un outil de suivi. L'objectif principal du projet a été atteint, soit de concevoir un modèle robuste prédictif d'écarts de température relatifs permettant le classement (9 classes) de ses valeurs en différences de température. Cette donnée permet d'identifier les zones potentielles les plus chaudes et les plus fraîches. Les classes d'écarts de température relatifs obtenues peuvent donc être considérées comme des indicateurs robustes qui permettent d'identifier et de localiser les ICFU. Ces classes devraient être interprétées en prenant en compte la donnée d'écarts de température relatifs utilisée pour faire le classement. Il est ainsi suggéré d'utiliser soit la dernière (9) ou les deux dernières classes (8-9) pour représenter les îlots de chaleur selon le territoire ciblé. La cartographie des ICFU couvre tout le Québec urbain, soit 14 508 km<sup>2</sup>. Le modèle prédictif est très performant, avec un R<sup>2</sup> est 0,81 et une erreur moyenne en absolue de 1,99 °C d'écart de température relatif, soit un gain de 14 % par rapport à l'étude de Boulfroy *et al.* (2013). Des pistes d'amélioration ont été identifiées pouvant être appliquées à plusieurs stades du processus de développement, notamment par l'ajout d'algorithmes d'apprentissage profond afin de rendre plus simple et uniforme la création des mosaïques et de variables. Cette cartographie servira de référence pour une analyse diachronique (2013-2014 c. 2020-2022) à l'échelle des centres de population prévue dans le sous-projet 3 du présent projet.

## 8. Références

- Alhawiti, R. et Mitsova, D. (2016) USING LANDSAT-8 DATA TO EXPLORE THE CORRELATION BETWEEN URBAN HEAT ISLAND AND URBAN LAND USES. *International Journal of Research in Engineering and Technology*, vol. 5, p. 457-466.
- Augustine, J., Hodges, G., Dutton, E., Michalsky, J. et Cornwall, C. (2008) An aerosol optical depth climatology for NOAA's national surface radiation budget network (SURFRAD). *Journal of Geophysical Research*, vol. 113.
- Barrette, N., Vandersmissen, M.-H. et Roy, F. (2018) Atlas web de la vulnérabilité de la population québécoise aux aléas climatiques (p. 290). Département de géographie, Université Laval.
- Bartolucci, L. A., Chang, M., Anuta, P. E. et Graves, M. R. (1988) Atmospheric effects on Landsat TM thermal IR data. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, n°2, p. 171-176.
- Belfiore, O. R., Meneghini, C., Parente, C. et Santamaria, R. (2016) APPLICATION OF DIFFERENT PAN-SHARPENING METHODS ON WORLDVIEW-3 IMAGES, vol. 11, n°1, p. 7.
- Bhang, K. J. et Park, S.-S. (2009) Evaluation of the Surface Temperature Variation With Surface Settings on the Urban Heat Island in Seoul, Korea, Using Landsat-7 ETM+ and SPOT. *IEEE Geoscience and Remote Sensing Letters*, vol. 6, n°4, p. 708-712.
- Bouffroy, E., Khaldoune, J., Grenon, F., Fournier, R. et Talbot, B. (2012) Conservation des îlots de fraîcheur urbains - Description de la méthode suivie pour identifier et localiser les îlots de fraîcheur et de chaleur (p. 38). CERFO et Université de Sherbrooke.
- Bouffroy, E., Khaldoune, J., Grenon, F., Fournier, R. et Talbot, B. (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) (p. 40). Centre d'enseignement et de recherche en foresterie de Sainte-Foy inc. (CERFO) et Université de Sherbrooke.
- Breiman, L. (2001) Random forests. *Machine Learning*, vol. 45, n°1, p. 5-32.
- Bui, Q.-T., Chou, T.-Y., Hoang, T.-V., Fang, Y.-M., Mu, C.-Y., Huang, P.-H., Pham, V.-D., Nguyen, Q.-H., Anh, D. T. N., Pham, V.-M. et Meadows, M. E. (2021) Gradient Boosting Machine and Object-Based CNN for Land Cover Classification. *Remote Sensing*, vol. 13, n°14, p. 2709.
- Chalghaf, B., Varin, M. et Joanisse, G. (2019) Cartographie fine des essences individuelles par une approche de modélisation de type « Random Forest », à partir du lidar et de RapidEye (Rapport scientifique No. 2019-04). Centre d'enseignement et de recherche en foresterie de Sainte-Foy inc. (CERFO).
- Cohen, J. (1960) A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, vol. 20, p. 37-46.
- Congalton, R. G. (1991) A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, vol. 37, n°1, p. 35-46.
- Cui, Y. Y. et Foy, B. de (2012) Seasonal Variations of the Urban Heat Island at the Surface and the Near-Surface and Reductions due to Urban Vegetation in Mexico City. *Journal of Applied Meteorology and Climatology*, vol. 51, n°5, p. 855-868.

- Deur, M., Gašparović, M. et Balenović, I. (2021) An Evaluation of Pixel- and Object-Based Tree Species Classification in Mixed Deciduous Forests Using Pansharpened Very High Spatial Resolution Satellite Imagery. *Remote Sensing*, vol. 13, n°10, p. 1868.
- Dousset, B., Gourmelon, F., Laaidi, K., Zeghnoun, A., Giraudet, E., Bretin, P., Mauri, E. et Vandentorren, S. (2011) Satellite monitoring of summer heat waves in the Paris metropolitan area. *International Journal of Climatology*, vol. 31, n°2, p. 313-323.
- Epstein, Y. et Moran, D. S. (2006) Thermal Comfort and the Heat Stress Indices. *Industrial Health*, vol. 44, n°3, p. 388-398.
- Francini, S. (2022) BAP-GEE. JavaScript.
- Ghulam, A. (2010) Calculating Surface Temperature Using Landsat Thermal Imagery. Thèse de doctorat, Department of Earth & Atmospheric Sciences, and Center for Environmental Sciences, Saint Louis University, St. Louis, MO.
- Giguère, M. (2009) Mesures de lutte aux îlots de chaleur urbains revue de littérature (p. 1 ressource en ligne). Québec : Institut national de santé publique du Québec. Direction des risques biologiques environnementaux et occupationnels.
- Goldblatt, R., Addas, A., Crull, D., Maghrabi, A., Levin, G. G. et Rubinyi, S. (2021) Remotely Sensed Derived Land Surface Temperature (LST) as a Proxy for Air Temperature and Thermal Comfort at a Small Geographical Scale. *Land*, vol. 10, n°4, p. 410.
- Good, E. J., Ghent, D. J., Bulgin, C. E. et Remedios, J. J. (2017) A spatiotemporal analysis of the relationship between near-surface air temperature and satellite land surface temperatures using 17 years of data from the ATSR series. *Journal of Geophysical Research: Atmospheres*, vol. 122, n°17, p. 9185-9210.
- Google Earth Engine (2022) Reducer Overview | Google Earth Engine. Google Developers.
- Guo, G., Wu, Z., Xiao, R., Chen, Y., Liu, X. et Zhang, X. (2015) Impacts of urban biophysical composition on land surface temperature in urban heat island clusters. *Landscape and Urban Planning*, vol. 135, p. 1-10.
- IPCC (2022) Climate Change 2022: Impacts, Adaptation, and Vulnerability (Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama (eds.)]) (p. 3676). GIEC, Cambridge University Press.
- Johnson, B. (2014) Effects of Pansharpening on Vegetation Indices. *ISPRS International Journal of Geo-Information*, vol. 3, n°2, p. 507-522.
- Kaplan, G., Avdan, U. et Avdan, Z. Y. (2018) Urban Heat Island Analysis Using the Landsat 8 Satellite Data: A Case Study in Skopje, Macedonia. *Proceedings*, vol. 2, n°7, p. 358.
- Kestens, Y., Brand, A., Fournier, M., Goudreau, S., Kosatsky, T., Maloley, M. et Smargiassi, A. (2011) Modelling the variation of land surface temperature as determinant of risk of heat-related health events. *International journal of health geographics*, vol. 10, p. 7.
- Kussul, N., Lavreniuk, M., Skakun, S. et Shelestov, A. (2017) Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geoscience and Remote Sensing Letters*, vol. 14, n°5, p. 778-782.
- Lechowicz, M. J. (1984) Why do temperate deciduous trees leaf out at different times? Adaptation and ecology of forest communities. *American Naturalist*, vol. 124, n°6, p. 821-842.

- Lessard, G. et Boulfroy, E. (2008) Les rôles de l'arbre en ville (p. 21). Québec : Centre d'enseignement et de recherche en foresterie de Sainte-Foy inc. (CERFO).
- Li, H., Jing, L. et Tang, Y. (2017) Assessment of Pansharpening Methods Applied to WorldView-2 Imagery Fusion. *Sensors*, vol. 17, n°1, p. 89.
- Lindsay, J. (2014) The Whitebox Geospatial Analysis Tools project and open-access GIS. The University of Glasgow.
- MELCC (2020) : septième année la plus chaude en 106 ans au sud de la province et neuvième plus chaude à l'échelle du Québec. Ministère de l'Environnement et de la Lutte contre les changements climatiques (MELCC).
- Mhangara, P., Mapurisa, W. et Mudau, N. (2020) Comparison of Image Fusion Techniques Using Satellite Pour l'Observation de la Terre (SPOT) 6 Satellite Imagery. *Applied Sciences*, vol. 10, n°5, p. 1881.
- Monserud, R. A. et Leemans, R. (1992) Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling*, vol. 62, n°4, p. 275-293.
- Montolio, P., Gasull, A., Monte, E., Torres, L. et Marqués, F. (1992) Analysis and optimization of the k-means algorithm for remote sensing applications. *Pattern Recognition and Image Analysis, Series in Machine Perception and Artificial Intelligence (Vol. Volume 1, p. 155-170)*. WORLD SCIENTIFIC.
- Moran, P. A. P. (1948) The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society*, vol. series B 10, p. 243-251.
- Morfitt, R., Barsi, J. A., Levy, R., Markham, B. L., Micijevic, E., Ong, L., Scaramuzza, P. et Vanderwerff, K. (2015) Landsat-8 Operational Land Imager (OLI) radiometric performance on-orbit. *Remote Sensing*, vol. 7, n°2, p. 2208-2237.
- Nadizadeh Shorabeh, S., Hamzeh, S., Zanganeh Shahraki, S., Firozjaei, M. K. et Jokar Arsanjani, J. (2020) Modelling the intensity of surface urban heat island and predicting the emerging patterns: Landsat multi-temporal images and Tehran as case study. *International Journal of Remote Sensing*, vol. 41, n°19, p. 7400-7426.
- Nininahazwe, F., Varin, M. et Théau, J. (2021) Caractérisation et suivi des écosystèmes riverains de l'agglomération de Québec - Volet 2 développement d'une méthode de télédétection pour l'identification d'espèces exotiques envahissantes ( No. 2021-22) (p. 59 + 10 annexes). Centre d'enseignement et de recherche de Sainte-Foy inc. (CERFO) et Université de Sherbrooke.
- Park, H., Kim, N., Park, S. et Choi, J. (2020) Sharpening of Worldview-3 Satellite Images by Generating Optimal High-Spatial-Resolution Images. *Applied Sciences*, vol. 10, n°20, p. 7313.
- PCI Geomatics (2018) Using PCI Software. Richmond Hill, Ontario.
- R Core Team (2021) R A Language and Environment for Statistical Computing. Vienna, Austria : R Foundation for Statistical Computing.
- Ravel, S. (2017) Démixage d'images hyperspectrales en présence d'objets de petite taille. Thèse de doctorat, Physique et Science de la Matière, Institut Fresnel, l'ÉCOLE CENTRALE de MARSEILLE, Marseille, France, 110 p.
- Shiab, N. et Bouchard, I. (2022) Voici qui vit dans les pires îlots de chaleur de votre ville.
- Small, C. (2001) Estimation of urban vegetation abundance by spectral mixture analysis. *International Journal of Remote Sensing*, vol. 22, n°7, p. 1305-1334.
- Smargiassi, A., Goldberg, M. S., Plante, C., Fournier, M., Baudouin, Y. et Kosatsky, T. (2009) Variation of daily warm season mortality as a function of micro-urban heat islands. *Journal of epidemiology and community health*, vol. 63, p. 659-64.

- Soltani, A. et Sharifi, E. (2017) Daily variation of urban heat island effect and its correlations to urban greenery: A case study of Adelaide. *Frontiers of Architectural Research*, vol. 6, n°4, p. 529-538.
- Storey, J., Choate, M. et Moe, D. (2014) Landsat 8 Thermal Infrared Sensor Geometric Characterization and Calibration. *Remote Sensing*, vol. 6, p. 11153-11181.
- USGS (2020) Landsat 8-9 Operational Land Imager (OLI) - Thermal Infrared Sensor (TIRS) Collection 2 Level 1 (L1) Data Format Control Book (DFCB) ( No. LSDS-1822 Version 6.0) (p. 64).
- Vali, A., Comai, S. et Matteucci, M. (2020) Deep Learning for Land Use and Land Cover Classification Based on Hyperspectral and Multispectral Earth Observation Data: A Review. *Remote Sensing*, vol. 12, n°15, p. 2495.
- Van Der Meer, F. (1995) Spectral unmixing of Landsat Thematic Mapper data. *International Journal of Remote Sensing*, vol. 16, n°16, p. 3189-3194.
- Varin, M., Grenon, F. et Boulfroy, E. (2016) Cartographie des îlots de chaleur et de fraîcheur de la Communauté métropolitaine de Québec.
- Ville de Montréal (2022) Îlots de chaleur - Jeu de données. Données ouvertes.
- White, J. C., Wulder, M. A., Hobart, G. W., Luther, J. E., Hermosilla, T., Griffiths, P., Coops, N. C., Hall, R. J., Hostert, P., Dyk, A. et Guindon, L. (2014) Pixel-Based Image Compositing for Large-Area Dense Time Series Applications and Science. *Canadian Journal of Remote Sensing*, vol. 40, n°3, p. 192-212.
- Yu, X., Guo, X. et Wu, Z. (2014) Land Surface Temperature Retrieval from Landsat 8 TIRS—Comparison between Radiative Transfer Equation-Based Method, Split Window Algorithm and Single Channel Method. *Remote Sensing*, vol. 6, n°10, p. 9829-9852.