

La Base de données ouverte sur les serres

Base de données ouverte des serres document de métadonnées : concepts, méthodologie et qualité des données, 2025

Version 2.0



Date de diffusion : le 18 août 2025



Statistique
Canada

Statistics
Canada

Canada

Comment obtenir d'autres renseignements

Pour toute demande de renseignements au sujet de ce produit ou sur l'ensemble des données et des services de Statistique Canada, visiter notre site Web à www.statcan.gc.ca.

Vous pouvez également communiquer avec nous par :

Courriel à infostats@statcan.gc.ca

Téléphone entre 8 h 30 et 16 h 30 du lundi au vendredi aux numéros suivants :

- | | |
|---|----------------|
| • Service de renseignements statistiques | 1-800-263-1136 |
| • Service national d'appareils de télécommunications pour les malentendants | 1-800-363-7629 |
| • Télécopieur | 1-514-283-9350 |

Normes de service à la clientèle

Statistique Canada s'engage à fournir à ses clients des services rapides, fiables et courtois. À cet égard, notre organisme s'est doté de normes de service à la clientèle que les employés observent. Pour obtenir une copie de ces normes de service, veuillez communiquer avec Statistique Canada au numéro sans frais 1-800-263-1136. Les normes de service sont aussi publiées sur le site www.statcan.gc.ca sous « Contactez-nous » > « [Normes de service à la clientèle](#) ».

Note de reconnaissance

Le succès du système statistique du Canada repose sur un partenariat bien établi entre Statistique Canada et la population du Canada, les entreprises, les administrations et les autres organismes. Sans cette collaboration et cette bonne volonté, il serait impossible de produire des statistiques exactes et actuelles.

Publication autorisée par la ministre responsable de Statistique Canada

© Sa Majesté le Roi du chef du Canada, représenté par la ministre de l'Industrie, 2025

L'utilisation de la présente publication est assujettie aux modalités de l'[entente de licence ouverte](#) de Statistique Canada.

Une [version HTML](#) est aussi disponible.

This publication is also available in English.

Table des matières

1 Aperçu	4
2 Sources de données.....	5
3 Période de référence	7
4 Géographie.....	7
5 Population cible	7
6 Méthodologie de compilation.....	8
7 Dictionnaire de données	10
8 Exactitude des données.....	13
9 Contactez-nous	14
Références	15

Base de données ouverte des serres document de métadonnées : concepts, méthodologie et qualité des données, 2025

1 Aperçu

En vue d'explorer l'utilisation des données ouvertes pour établir des statistiques officielles et d'appuyer la recherche géospatiale dans divers domaines, la Division de l'agriculture et la Division de la science des données et de l'innovation ont lancé un projet qui repose sur les données d'observation de la Terre pour moderniser les enquêtes traditionnelles menées à Statistique Canada. Cette initiative exploite un ensemble de sources d'observation de la Terre à haute résolution rendu disponible sous forme de données ouvertes par plusieurs ordres de gouvernement au Canada¹. Élaboré en réponse aux efforts de modernisation de Statistique Canada et au projet [AgZéro](#), ce projet rassemble des méthodes avancées, notamment des systèmes d'information géographique, l'intégration de données et des technologies de pointe. Dans le cadre de ces efforts, la Base de données ouvertes sur les serres (BDOS) est utilisée à des fins de soutien et d'entraînement de modèles d'apprentissage automatique visant à automatiser la collecte de renseignements sur les serres à l'échelle du pays. Le présent document décrit le processus de collecte, de traitement et de normalisation des données tirées de l'observation de la Terre et des produits dérivés des imageries des serres numérisées, comme cela a été établi dans la première version de la BDOS. La BDOS est disponible en vertu de la Licence du gouvernement ouvert — Canada².

Dans la première version, la BDOS comprenait 2 476 enregistrements individuels dans 10 municipalités et 4 provinces. La version actuelle (version 2) contient 3 905 serres numérisées manuellement et enregistrements individuels dans 38 municipalités et 4 provinces, ce qui correspond à une augmentation de 58 %. La base de données sera mise à jour périodiquement au fur et à mesure que de nouveaux ensembles de données ouvertes seront rendus disponibles. La BDOS est fournie en fichier géographique de format shapefile.

En plus des serres numérisées manuellement qui fournissent des renseignements à jour sur les serres à partir de la première version, Statistique Canada a également inclus 1 006 serres détectées automatiquement au moyen de méthodes novatrices de science des données dans l'imagerie satellite de source ouverte.

La composante d'apprentissage automatique comprend deux mécanismes : 1) l'imagerie Sentinel à super-résolution; 2) la détection des serres. Le processus de super-résolution est un modèle développé pour utiliser de nombreuses images Sentinel-1 et Sentinel-2, avec une résolution de 10 mètres, d'un même emplacement prises à différents moments, afin de générer une image unique « à super-résolution » d'une résolution de 1,25 mètre. Cette image à super-résolution est ensuite utilisée par le modèle de détection de serres, d'abord pour l'entraînement, puis pour détecter les serres sur de nouvelles images à super-résolution. L'imagerie à super-résolution a permis de supprimer l'obstacle lié au coût de l'imagerie à haute résolution. Le modèle de détection des serres est un modèle U-net conçu pour segmenter les pixels d'une image en deux catégories : « serre » et « non-serre ». Il est entraîné à partir d'une imagerie et d'étiquettes détectant les serres, et apprend ainsi à reconnaître les caractéristiques qui permettent de distinguer les serres des autres objets ou de l'arrière-plan.

Les résultats de l'apprentissage automatique sont expérimentaux et peuvent donc ne pas représenter des données complètes et précises des serres sur le terrain. Comme toutes les techniques de modélisation, il y a certaines limites : (i) le modèle d'apprentissage automatique a été entraîné sur les données du sud de l'Ontario et sa capacité à être appliqué à d'autres régions n'a pas été testée, (ii) la performance du modèle d'apprentissage automatique est affectée par la résolution de l'imagerie et la taille des serres, en particulier les petites serres peuvent être trop petites pour être détectées (comme avec la numérisation humaine).

Cet ensemble de données est diffusé dans le cadre de l'Environnement de couplage de données ouvertes (ECDO). L'ECDO est une initiative qui vise à accroître l'utilisation et l'harmonisation des données ouvertes provenant de sources fournissant une série d'ensembles de données diffusés en vertu d'une licence unique, ainsi que du code source libre pour coupler ces ensembles de données. Il est possible d'accéder aux ensembles de données et au code de l'ECDO sur le site Web de Statistique Canada à l'adresse suivante :

[L'Environnement de couplage de données ouvertes](#)

1. Cela comprend les gouvernements municipaux, régionaux et provinciaux.

2. [Licence du gouvernement ouvert – Canada](#)

2 Sources de données

Plusieurs sources de données ont été recueillies pour créer la version 2 de la BDOS. Les fournisseurs de données de la version 1 comprenaient des sources gouvernementales de divers ordres qui ont offert gratuitement leurs images à haute résolution, ou ces images ont été fournies à Statistique Canada par des entreprises qui détiennent une offre à commande nationale³ avec le gouvernement fédéral, y compris l'attribution à chaque source conformément aux exigences de la licence.

Cependant, Statistique Canada a modifié ses sources d'imagerie satellite pour la deuxième version de la BDOS, qui s'appuie sur des sources d'imagerie satellite gratuites et à fréquence élevée. Pour cette deuxième version, des serres ont été numérisées manuellement à l'aide de Sentinel-2 pour les années 2023 et 2024. Une imagerie à plus haute résolution issue de plateformes à code source ouvert a souvent été utilisée pour affiner le repérage et la numérisation plus précis des structures des serres, dans la mesure du possible. Sentinel-2 a servi à numériser de nouvelles serres dans les cas où une imagerie plus récente correspondant à la date cible n'était pas disponible sur les plateformes à code source ouvert. Sentinel-2 est doté d'un imageur multispectre offrant une résolution spatiale de 10 mètres, qui a été utilisé pour valider la présence ou l'absence de grandes serres, mais dont la précision reste limitée pour la numérisation de structures plus petites.

Outre l'utilisation d'une imagerie à résolution moyenne (10 m) issue de sources ouvertes pour la numérisation, Statistique Canada a également adapté une approche innovante pour générer des images à haute résolution (1,25 m) à partir de sources d'imagerie à résolution moyenne, grâce à une méthode appelée super-résolution (Wang, 2022).

Les données d'imagerie Sentinel-1 et Sentinel-2 ont été téléchargées depuis la plateforme Copernicus de l'Agence spatiale européenne (ESA), puis traitées par la méthode de super-résolution. Elles ont ensuite été intégrées à un modèle servant à l'entraînement et à l'évaluation du système de détection des serres. L'imagerie a été recueillie sur deux sites situés dans le sud de l'Ontario, au Canada : Niagara et Essex.

Les renseignements sur les sources de données sont fournis dans le tableau 1 ci-dessous. Au total, 38 municipalités sont couvertes dans 4 provinces.

Pour en savoir plus sur les licences individuelles, les utilisateurs peuvent consulter directement les portails de données ouvertes des fournisseurs de données en question.

3. Une offre à commandes est une offre qu'un fournisseur fait au Canada et qui permet à ce dernier d'acheter régulièrement des biens, des services ou une combinaison des deux, à des prix convenus à l'avance, selon des modalités déterminées et selon le besoin. Une offre à commandes n'est pas un contrat et le Canada n'est nullement obligé d'acheter. Une offre à commandes devient un contrat seulement une fois que le Canada émet une « commande subséquente » à l'offre à commandes. (Voir : [Programme de contrats fédéraux – Offres à commandes et arrangements en matière d'approvisionnement – IPG-085](#)).

Tableau 1
Sources de données

ODG Version	Data Sources	Credits	Image(s) Details	License Agreement
1	MDA Geospatial Services	York, Ontario	WorldView3 - 8 Band, TIFF, DRA off, 30cm Pan+1.2m Date: 2020/06/13	L'Offre à commandes principale et nationale pour MDA a pour numéro de contrat W1786-180002/001/Offre à commandes E60SQ-120001/003/SS
		Laval, Québec	WorldView2 - 8 Band Ortho Tif FTP NN NAD83 UTM 8-bit, DRA off, 40cm Pan+1.6m Date: 2021/05/24	
		St-Eustache, Québec	WorldView2 - 8 Band Ortho Tif, DRA off, 50cm Pan+2m Date: 2020/05/20	
		Medicine Hat, Alberta	WorldView2 - 8 Band Ortho Tif, DRA off, 50cm Pan+2m Date: 2021/05/01	
1	Township of Langley	Canton de Langley — Système d'information géospatiale (SIG)	Orthophotographies de 2017	Contient de l'information visée par la licence de gouvernement ouvert — https://www.tol.ca/en/connect/open-data-licence.aspx (en anglais seulement)
1	Ville de Burnaby	Ville de Burnaby — SIG	Orthophotographies de 2020	Licence de gouvernement ouvert — Colombie-Britannique https://data.burnaby.ca/pages/open-government-licence (en anglais seulement)
1	Ville de Surrey	Ville de Surrey;	Orthophotographies de 2018	Licence de gouvernement ouvert — Ville de Surrey https://data.surrey.ca/pages/open-government-licence-surrey (en anglais seulement)
1	Ministère des Richesses naturelles et des Forêts de l'Ontario	Ministère des Richesses naturelles et des Forêts de l'Ontario	Orthophotographie	Orthophotographie sous licence avec le ministère des Richesses naturelles et des Forêts de l'Ontario et Forestry® King's Printer for Ontario, 2018 et 2020
1	Ville de Chilliwack — SIG	Ville de Chilliwack — SIG;	Orthophotographies de 2021	Les données sont fournies en tant que service public par la Ville de Chilliwack. https://www.chilliwack.com/main/page.cfm?id=2391 (en anglais seulement)
2	L'Agence spatiale européenne	Copernicus Sentinel-2 Multispectre	Multispectre - résolution de 10 mètres. Couverture : Au moins une image est disponible, sans nuages, couvrant les zones agricoles partout au Canada. Date : Images disponibles pour les mois allant d'avril à septembre des années 2020, 2023 et 2024.	Licence d'accès libre de l'ASE https://dataspace.copernicus.eu/terms-and-conditions
2	L'Agence spatiale européenne	Copernicus Sentinel-1 GRD (déecté au sol).	Couverture : Au moins une image est disponible, sans nuages, couvrant les zones agricoles partout au Canada. Date : Images disponibles pour les mois allant d'avril à septembre des années 2020 et 2021.	Licence d'accès libre de l'ASE https://dataspace.copernicus.eu/terms-and-conditions

3 Période de référence

Les données portent sur les années 2023 et 2024 pour les serres numérisées manuellement d'un bout à l'autre du Canada. Les sources de données ont été accessibles, consultées et traitées au printemps de 2023 et 2024. L'imagerie satellite utilisée pour la super-résolution dans le cadre de l'apprentissage automatique appliqué à la numérisation des serres couvre les années 2020 et 2021 pour les régions d'Essex et de Niagara, en Ontario.

4 Géographie

Le cadre géographique de la BDOS est cité en référence par les cinq régions clés suivantes : 1) les régions de la vallée du Fraser et du Grand Vancouver, en Colombie-Britannique; 2) Essex et 3) Niagara, dans le sud de l'Ontario; 4) le nord de Montréal, au Québec; 5) les régions de Medicine Hat et d'Edmonton, en Alberta. Ces régions et la BDOS citent en référence les divisions de recensement de Statistique Canada (découpage géographique de Statistique Canada : [Recensement de 2021 — Fichiers des limites](#)).

Les sites des serres générées au moyen de l'apprentissage automatique étaient situés à Essex et à Niagara, en Ontario. L'imagerie Sentinel-1 et Sentinel-2 de ces sites a été téléchargée et utilisée pour générer une image en super-résolution pour chaque site.

Ces sources ne se limitent pas uniquement aux frontières en fonction des régions géographiques de Statistique Canada pour une région particulière et peuvent s'étendre aux municipalités voisines. Il se pourrait que la BDOS ne soit pas complète pour une région indiquée au tableau 1 ou soit limitée aux frontières en fonction des régions géographiques de Statistique Canada.

Représentation géographique

La BDOS est disponible sur le site Web de Statistique Canada, dans la représentation géographique suivante :

- Projection : conique conforme de Lambert
- Abscisse fictive : 6200000,000000
- Ordonnée fictive : 3000000,000000
- Méridien central : -91,866667
- Parallèle de référence 1 : 49,000000
- Parallèle de référence 2 : 77,000000
- Latitude d'origine : 63,390675
- Unité de mesure linéaire : mètre (1,000000)
- Système de référence géodésique : nord-américain 1983 (NAD83)
- Méridien d'origine : Greenwich
- Unité de mesure angulaire : degré
- Sphéroïde : GRS 1980

Le Système de référence géodésique nord-américain de 1983 (NAD83) est une version corrigée du Système de référence géodésique nord-américain de 1927 (NAD27), qui offre un arpentage géodésique plus précis.

5 Population cible

Statistique Canada définit une serre et les produits de serre⁴ comme un espace pour faire pousser des semis, des plantes en pots, des plantes à massif, des boutures et autre matériel de multiplication, ainsi que des fruits et légumes cultivés pour la vente dans une structure permanente, fermée chauffée artificiellement et faite de plastique, de plexiglas, de polyfilm ou de verre.

4. [Enquête annuelle sur les cultures de serre, les pépinières et les gazonnières \(SPG\)](#)

Par conséquent, il se peut que d'autres bâtiments qui ne correspondent pas à la définition de serre mentionnée ci-dessus soient inclus dans l'ensemble de données en fonction de leurs caractéristiques visuelles communes. La base de données ne comprend pas de couplages avec les renseignements des entreprises, qui différencieraient les installations agricoles de celles qui n'ont aucun usage agricole.

Un travail minimal de vérification et de validation de la forme des bâtiments est effectué au moyen de méthodes de numérisation et d'apprentissage automatique dans les cas où la validation des bâtiments saisis dans la base de données présente des caractéristiques visuelles comparables. L'ensemble de données détecte les serres sans faire de distinction quant à leur type, aux cultures qui y sont cultivées, ni à toute autre caractéristique pouvant aider à leur classification. La base de données ne comprend pas de couplages avec les renseignements des entreprises et ne renvoie pas aux enquêtes de Statistique Canada, aux registres des entreprises, aux données fiscales ou à d'autres sources, ce qui lui permet de maintenir une composante de base de données ouverte.

6 Méthodologie de compilation

La création de la version 2 de la BDOS comprenait deux étapes principales de traitement de l'information : premièrement, le traitement des données d'observation de la Terre; deuxièmement, la création et le formatage de l'ensemble de données superposant les données d'observation de la Terre et mettant en correspondance les attributs de l'ensemble de données original avec les noms des variables standards (colonne). La version 2 de la BDOS comporte deux méthodes d'élaboration, à savoir la numérisation manuelle et l'apprentissage automatique.

Un dictionnaire de données des variables utilisées est présenté à la section 7. Les deux méthodes permettant de compiler les données dans la base de données géographiques en fichier de format shapefile final sont décrites ci-dessous :

Numérisation

- Chaque donnée d'observation de la Terre a été extraite, décompressée et convertie en format TIF si elle n'était pas déjà au format lors de l'acquisition.
- L'imagerie Sentinel-2 a été visualisée dans le logiciel du SIG à l'aide de services Web, ce qui a permis de supprimer les étapes de téléchargement et de traitement de l'imagerie, et de créer de nouvelles classes d'entités géographiques pour chaque région d'intérêt.
- Les serres visuellement comparables à des serres connues ont été repérées dans l'observation de la Terre, et un nouvel enregistrement a été numérisé au sein de la classe d'entités.
- Dans les cas où un bâtiment abritant une serre figurait dans la première version de la BDOS, mais n'était pas visible dans la **seconde** version, les polygones de serre ont été supprimés.
- Une fois chaque région traitée, les tableaux d'attributs ont été complétés avec les renseignements corrects et à jour.
- Une fois l'ensemble des régions couvertes, les différentes classes d'entités ont été fusionnées pour constituer le fichier final de la BDOS.

Imagerie en super-résolution

- Le modèle de super-résolution est un réseau de neurones présentant l'architecture suivante :
 - ▶ une couche CGR (Convolution+GroupNorm+unité linéaire rectifiée) de taille 1 x 3 x 3, suivie d'un bloc DenseNet (réseau de neurones densément connectés) 1 x 3 x 3;
 - ▶ une couche CGR 1 x 1 x 1, suivie d'une couche d'attention à plusieurs têtes;
 - ▶ une couche CG (Convolution+GroupNorm) 1 x 1, suivie d'un bloc DenseNet 3 x 3;
 - ▶ une couche CG 1 x 1, suivie d'un suréchantillonnage, d'une autre couche CG 1 x 1, puis d'un autre bloc DenseNet 3 x 3;
 - ▶ une couche CGR 1 x 1, suivie d'une couche de convolution 1 x 1 et d'une fonction sigmoïde.
- L'architecture du modèle est fondée sur une analyse documentaire (Wang et coll., 2022; Dong et coll., 2016; Kawulok et coll., 2020; Fuoli et coll., 2021).

- La première étape lors de l'exécution du modèle consiste à télécharger l'imagerie Sentinel-1 et Sentinel-2. Veiller à une projection commune, puis fournir les images au modèle de super-résolution afin d'obtenir des images en super-résolution des sites demandés.
- Les images en super-résolution sont reprojctées avec la même projection que celle du fichier de format shapefile des serres numérisées manuellement.

Détection par apprentissage automatique

- Le modèle de détection est un réseau de neurones qui repose sur l'architecture U-Net pour la segmentation d'images, fondé sur l'article scientifique de Ronneberger O. et al., 2015.
- L'architecture a été modifiée pour utiliser ResNet34 comme partie pour l'encodage afin d'améliorer la stabilité de l'entraînement.
- Le modèle a été entraîné à l'aide d'imagerie en super-résolution et d'étiquettes de serres numérisées manuellement.
- En ce qui concerne la détection, la première étape consiste à transmettre l'imagerie en super-résolution au modèle de détection des serres fondé sur l'apprentissage automatique.
- On obtient alors en sortie un fichier de format shapefile contenant les détections de serres.
- Pour cette version, les faux positifs ont été supprimés en croisant les serres détectées avec les fichiers de format shapefile des serres connues (numérisées manuellement).
- À l'avenir, des étapes de traitement des données après la collecte permettront de supprimer les faux positifs sans devoir compter sur les polygones de serres connues. Par exemple, cela pourrait se faire en utilisant une couche de l'Indice de végétation par différence normalisée.
- Enfin, le fichier de format shapefile des serres numérisées manuellement est fusionné avec les détections issues du modèle d'apprentissage automatique.

Les champs de données originaux étaient l'identificateur et la forme uniques indiqués automatiquement à partir du logiciel. De nouveaux champs ont été créés pour fournir des renseignements sur la source des données d'imagerie, la méthode de collecte, l'emplacement du centroïde X et Y et la province. Bien que des efforts aient été déployés pour s'assurer que toutes les serres étaient détectées et que d'autres types de bâtiments n'étaient pas inclus, certains bâtiments pourraient être identifiés par erreur, ou il se pourrait que des serres aient été omises de l'image source. Si de telles erreurs sont relevées, elles seront corrigées dans les versions futures de la BDOS.

Les données incluses dans la BDOS sont attribuables à une inspection visuelle seulement, et elles ne sont pas couplées à des bases de données, des enquêtes ou des sources privées officielles.

Géocodage

Les enregistrements de la version 2 de la BDOS comprennent l'année, la source d'information, la méthode de collecte, la longitude du centroïde x du bâtiment, la latitude du centroïde y du bâtiment, l'identifiant unique de la province et le nom. Les enregistrements ne comprennent aucun autre renseignement sur l'emplacement, comme l'adresse ou le code postal. Ils n'incluent aucun autre renseignement concernant le matériau de couverture de la serre, la culture à l'intérieur de la serre ou la propriété.

Les images utilisées dans le processus d'entraînement du modèle de détection des serres sont reprojctées dans le système de coordonnées de référence (EPSG:3347), correspondant à celui des étiquettes, afin de garantir une bonne harmonisation avec les serres numérisées.

Normalisation des données

En raison des différentes normes adoptées dans les données originales, les mesures prises pour normaliser les données ont pu entraîner des erreurs. Les principes clés de la méthodologie utilisée étaient d'éviter les faux positifs et les modifications importantes des données. La méthodologie et les limites de chaque technique sont décrites ci-dessous. Les techniques de nettoyage simples, comme la suppression des espaces et de la ponctuation, ne sont pas décrites.

Pour ce qui est de la méthodologie d'apprentissage automatique, le processus de normalisation consistait à traiter les étiquettes et les images selon la même projection et à garantir l'harmonisation pour l'entraînement du modèle d'apprentissage automatique.

Comparaisons avec l'Enquête annuelle sur les cultures de serre, les pépinières et les gazonnières

L'Enquête annuelle sur les cultures de serre, les pépinières et les gazonnières (EACSPG) de Statistique Canada est un recueil d'information sur la production des serres, les stocks des pépinières et les gazonnières produits au Canada, et les données de cette enquête sont souvent utilisées pour mener des analyses sur les tendances du marché. Puisque l'EACSPG n'est pas fondée sur les renseignements de la BDOS, et que la BDOS ne repose pas non plus sur les données de l'EACSPG, il est peu probable que les renseignements et la superficie totale pour une province ou une région soient comparables. Les données demeurent séparées les unes des autres afin de permettre à la BDOS d'être publiée et utilisée par le public grâce à la licence d'utilisation de données ouvertes.

Suppression des enregistrements en double

En ce qui concerne la BDOS, seules les entrées qui semblaient être des doublons évidents, chevauchant les formes de serres, ont été choisies pour être supprimées.

7 Dictionnaire de données

Le dictionnaire de données décrit les variables contenues dans la BDOS.

Variable — Numéro d'enregistrement

Nom

FID

Format

Chaîne de caractères

Source

Générée à l'interne lors du traitement des données

Description

Numéro d'enregistrement unique généré automatiquement lors du traitement des données

Variable — Forme

Nom

Shape

Format

Géométrie

Source

Générée à l'interne lors du traitement des données

Description

Générée automatiquement à partir des propriétés géométriques lors du traitement des données

Variable — Date de l'image

Nom

ImageDate

Format

Long

Source

Fourni dans la source d'imagerie

Description

Année d'acquisition de l'imagerie

Variable — Province ou territoire

Nom

PROV_TERR

Format

Chaîne de caractères

Source

Province ou territoire d'enregistrement

Description

Nom de la province ou du territoire

Variable — Identificateur unique de la province

Nom

PRUID

Format

Long

Source

Converti à partir du code de province

Description

Identificateur unique de la province conformément à l'identificateur officiel de Statistique Canada

Variable — Longitude

Nom

Longitude

Format

Double

Source

Géométrie calculée du centroïde x du bâtiment pour chaque enregistrement, exprimée en degrés décimaux

Description

Longitude

Variable — Latitude

Nom

Latitude

Format

Double

Source

Géométrie calculée du centroïde y du bâtiment pour chaque enregistrement, exprimée en degrés décimaux

Description

Latitude

Variable — Source de données

Nom

Source

Format

Chaîne de caractères

Source

Créée à partir des origines des données d'observation de la Terre

Description

Nom de l'entité qui a fourni les données d'observation de la Terre

Variable — Méthode

Nom

Method

Format

Chaîne de caractères

Source

Créé en fonction de la méthode de collecte utilisée pour enregistrer les serres

Description

Méthode de collecte utilisée pour enregistrer les serres. Seules deux méthodes de collecte sont possibles, à savoir la numérisation et l'apprentissage automatique. La méthode de numérisation s'applique lorsque la numérisation manuelle a été utilisée pour tracer le contour du bâtiment de la serre. L'apprentissage automatique s'applique lorsque l'algorithme d'apprentissage automatique a été utilisé pour détecter et créer automatiquement le bâtiment de la serre dans l'imagerie source.

8 Exactitude des données

Toutes les serres numérisées de la version 2 de la BDOS étaient en référence à l'imagerie acquise durant une période spécifique, fournie par l'imagerie satellite de l'ESA. En général, outre le traitement et la numérisation des caractéristiques dans l'ensemble de données, l'imagerie a été utilisée telle quelle et elle peut ainsi créer des erreurs dans la base de données finale, où les caractéristiques n'ont pu être déterminées correctement dans certains cas. Compte tenu de la nature de l'acquisition des données et de l'élaboration de la base de données, il est possible que certaines erreurs soient détectées dans le produit géographique final.

Modèle de super-résolution

L'évaluation de la qualité des prédictions en super-résolution a été réalisée de la manière suivante :

- En examinant l'erreur absolue moyenne et la racine de l'erreur quadratique moyenne pour chaque bande.
 - ▶ Un seuil de tolérance coulissant a été appliqué à chaque fenêtre pour l'erreur absolue moyenne et la racine de l'erreur quadratique moyenne. Pour chaque fenêtre de 100 x 100 pixels, l'erreur absolue moyenne et la racine de l'erreur quadratique moyenne minimales ont été prises en compte, en permettant un décalage jusqu'à 8 pixels.
 - ▶ Cette méthode par fenêtre coulissante permet d'accepter des images légèrement décalées (p. ex. un décalage constant de 5 pixels vers la gauche) par rapport aux images au sol.
 - ▶ Le fait d'autoriser de tels décalages favorise considérablement les images plus détaillées. Le modèle ne peut pas déterminer avec certitude le décalage entre les images Sentinel à 10 m et les images au sol. Sans cette tolérance de décalage, une image floue sans réelle amélioration serait favorisée par rapport à une image détaillée légèrement décalée de quelques pixels.
- On compare le ratio de l'erreur absolue moyenne des images à super-résolution à une valeur de référence, constituée d'un suréchantillonnage cubique d'une image Sentinel-2 ressemblant le plus possible à l'image cible (même ratio pour la racine de l'erreur quadratique moyenne). Si ce ratio est inférieur à un, cela signifie que la super-résolution dépasse le rendement de la méthode de référence pour la bande correspondante. Le même seuil de fenêtre coulissante a été utilisé à la fois pour la méthode de référence et la super-résolution.
- Une inspection visuelle des résultats en super-résolution a également été effectuée, en les comparant aux images d'entrée et aux images au sol à haute résolution.

Modèle de détection des serres

L'ensemble de données comprenait 1 209 serres, dont 749 ont été utilisées pour l'entraînement et 460 pour les essais. Le modèle de détection des serres a obtenu une cote F1 au niveau des pixels de 80 % lors des essais. Le taux de rappel était de 87 %, ce qui signifie que le modèle détectait systématiquement les vraies serres, et le taux de précision s'établissait à 75 %. La cote F1 au niveau des objets⁵ était plus élevée, atteignant 91 %. La majorité des erreurs de rendement provenaient des faux positifs. Après un examen plus approfondi, les faux positifs se sont avérés être des erreurs dans les images à super-résolution qui apparaissent sous des formes visuellement distinctes du paysage environnant, mais qui présentaient des caractéristiques similaires à celles des serres. Il y avait aussi des cas où certains bâtiments devenaient flous pendant le processus de super-résolution, donnant l'apparence des toits de serres; à l'inverse, certaines serres dans l'imagerie d'origine devenaient floues après le processus de super-résolution. Ces erreurs dans les images à super-résolution ont donc contribué aux erreurs dans les résultats de détection des serres.

5. La cote F1 est une mesure utilisée pour évaluer la précision d'un essai ou d'un modèle dans les tâches de classification. Il s'agit de la moyenne harmonique de la précision et du rappel, fournissant une mesure unique qui équilibre ces deux aspects. La précision désigne la proportion de résultats réellement positifs parmi toutes les prédictions positives, tandis que le rappel indique la proportion de résultats réellement positifs parmi l'ensemble des cas réellement positifs. La cote F1 est particulièrement utile lorsque la répartition des classes est déséquilibrée, car elle accorde une importance égale à la précision et au rappel. Une cote F1 plus élevée indique de meilleurs rendements en détectant correctement les cas pertinents tout en évitant d'inclure des éléments non pertinents.

9 Contactez-nous

Les bases de données ouvertes de l'ECDO sont conçues pour être améliorées de façon continue. Pour fournir des renseignements sur les ajouts, les mises à jour, les corrections ou les omissions, ou pour en savoir davantage, veuillez communiquer avec nous à l'adresse statcan.lode-ecdo.statcan@statcan.gc.ca. Veuillez inclure le titre de la base de données ouvertes dans l'objet du courriel.

Comment citer cet ensemble de données

Statistique Canada. (2025). La Base de données ouvertes sur les serres, version 2. <https://www.statcan.gc.ca/fr/ecdo/bases-donnees/bdos>

Références

- Dong, C., Change Loy, C. et Tang, X. (2016). *Accelerating the Super-Resolution Convolutional Neural Network*.
- Fuoli, D., Van Gool, L. et Timofte, R. (2021). *Fourier Space Losses for Efficient Perceptual Image Super-Resolution*.
- Kawulok, M., Benecki, P., Piechaczek, S., Hrynczenko, K., Kostrzewa, D. et Nalepa, J. (2020). Deep Learning for Multiple-Image Super-Resolution. *IEEE Geoscience and Remote Sensing Letters*.
- Li, Y., Wang, Y., Li, B. et Wu, S. (2022). Super-Resolution of Remote Sensing Images for 4 Resolution without Reference Images. *Electronics 2022*, 11, 3474.
- Wang, P. Bayrem, B. et Sertel, E. (2022). A comprehensive review on deep learning based remote sensing image super-resolution methods. *Earth-Science Reviews*, 232, 104110. 10.1016/j.earscirev.2022.104110.
- Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. MICCAI 2015. Lecture Notes in Computer Science(), vol 9351.