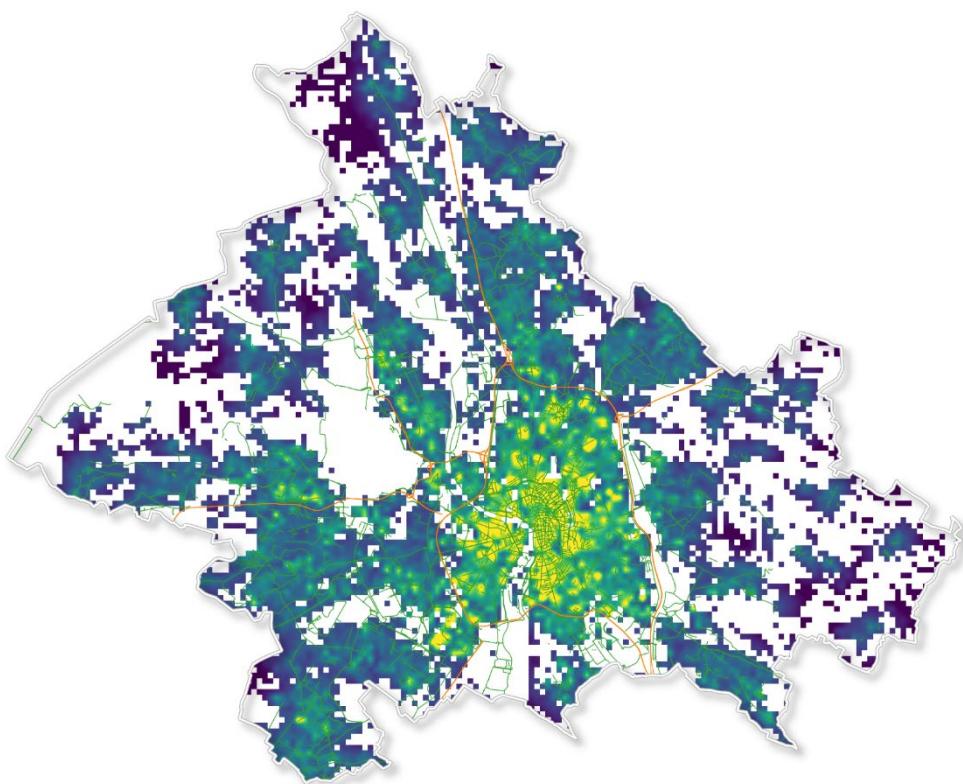


Master 2 SIGMA
“Sciences géomatiques en environnement et aménagement”

UE 901 32 : Modélisation spatio-temporelle

Carte de potentialité par Évaluation MultiCritère (EMC) : L’implantation d’un magasin de produits bio dans la communauté d’agglomération de Toulouse Métropole.



BALLOT Doris, BARBIERO Audrey, HECKENDORN Robin, LIMA Lucas

Année universitaire 2024 - 2025

Table de matières

1.	Introduction	5
1.1.	Choix des contraintes et des facteurs	5
2.	Préparation de données.....	6
2.1.	Les facteurs socio-économiques	6
2.2.	Les données relatives au terrain et à l'architecture urbaine	12
3.	Traitemet des données	13
3.1.	Rastérisation des contraintes.....	13
3.1.1.	Limitation de la zone d'étude.....	13
3.1.2.	Exclusion des carreaux situés dans des zones en eau	15
3.2.	Estimation du nombre d'individus consommateurs de produits biologiques ..	16
3.2.1.	Méthode	16
3.2.2.	Processus avec Python	16
3.3.	Résultats et critique.....	20
3.4.	Pondération des facteurs socio-économiques	20
3.4.1.	Méthode théorique	20
3.4.2.	Processus avec Python	21
3.5.	Différenciation des scénarios avec la prise en compte de l'accessibilité	23
3.5.1.	Premier scénario : magasin de petite surface en ville	25
3.5.2.	Second scénario : magasin de grande surface en périphérie.....	26
4.	Analyse des potentialités d'implantation.....	28
4.1.	Des zones qui se démarquent dans notre analyse	28
4.2.	Prise en compte de la présence actuelle de magasins bio	29
5.	Conclusion	30
6.	Annexes.....	31
6.1.	Bibliothèques Python utilisées	31
6.2.	Cartes des consommateurs potentiels de produits bio à Toulouse Métropole par catégorie - grand format.....	32

Liste de figures

Figure 1 : Répartition des habitudes de consommation de bio chez les français selon quatre facteurs socio-économiques. Source : Agence Bio.....	7
Figure 2 : Tables CSV reprenant les résultats du sondage de l'Agence Bio sur les facteur socio-économiques influençant la consommation de produits bio	7
Figure 3 : Illustration des différences d'échelles géographiques entre les limites d'iris et les carroyages Filosofi.....	11
Figure 4 : Diagramme de flux des traitements pour ramener les données IRIS à l'échelle des carreaux Filosofi	11
Figure 5 : Requête Overpass Turbo pour obtenir la localisation des magasins bio dans Toulouse Métropole depuis la base OpenStreetMap.	12
Figure 6 : Obtention du masque de Toulouse Métropole grâce aux codes INSEE des communes	13
Figure 7 : Diagramme de flux avec tous les géotraitements réalisés	14
Figure 8 : Masque combiné des carreaux de l'INSEE et des surfaces d'eau	15
Figure 9 : Extrait du code Python permettant d'estimer le nombre de consommateurs potentiels de bio grâce aux données INSEE et de l'Agence Bio	17
Figure 10 : Extrait du code Python permettant la normalisation des nombres de consommateurs de bio estimés par chacun des facteurs socio-économiques.	18
Figure 11 : Consommateurs potentiels de produits bio à Toulouse Métropole	23
Figure 12 : Fonction Python permettant de créer un raster de distance depuis une couche de lignes au format vecteur.....	24
Figure 13 : Proximité normalisée aux pistes cyclables.....	25
Figure 14 : Carte des zones prioritaires pour l'implantation d'un magasin de produits bio de petite surface	26
Figure 15 : Proximité normalisée aux routes principales.....	27
Figure 16 : Cartes des zones prioritaires pour l'implantation d'un magasin de produits bio de grande surface	28
Figure 17 : Comparaison entre les zones prioritaires identifiées et les magasins bio actuels .	29

Liste de tableaux

Tableau 1 : Données socio-économiques traitées et sources.	8
Tableau 2 : Relations établies entre les CSP de l'Agence Bio et de l'INSEE.....	9
Tableau 3: Relations établies entre les classes d'âge de l'Agence Bio et de l'INSEE.	9
Tableau 4: Relations établies entre les niveaux de diplôme de l'Agence Bio et de l'INSEE. 10	10
Tableau 5 : Les données géographiques traitées et leur source.	12
Tableau 6 : Écarts-types calculés et poids de chacun des facteurs socio-économiques	21

1. Introduction

Dans le cadre de l'UE “Modélisation spatio-temporelle” de notre seconde année de master SIGMA, nous avons choisi de mener un projet de géomarketing, visant à créer une carte de potentialité basée sur une méthode d’Évaluation MultiCritère (EMC). Cette approche doit nous permettre de modéliser les interactions entre différents critères spatiaux et d'identifier les zones les plus favorables à l'implantation d'un nouveau magasin bio dans la métropole toulousaine.

Le géomarketing consiste à exploiter des données géographiques pour répondre à des enjeux commerciaux et stratégiques. Dans le cas qui nous intéresse, il s'agit d'identifier les emplacements optimaux pour un magasin bio. Notre objectif principal est donc de combiner des critères géographiques, économiques, et sociaux pour fournir une analyse spatiale pertinente.

On envisage deux scénarios : dans un premier temps l'implantation d'un commerce de proximité situé en zone urbaine et dans un second temps l'implantation d'un plus grand espace commercial de type “supermarché” que l'on souhaite planter dans une zone plus rurale ou industrielle.

En fonction du scénario, on fera varier un des facteurs pris en compte pour l'évaluation multicritères.

Pour procéder à cette analyse, nous avons choisi d'utiliser les outils Python et QGIS. D'abord Python pour sa flexibilité d'utilisation dans le cas de manipulation de données (formatage, filtrage, calcul) et ensuite QGIS pour la mise en forme de nos résultats obtenus via les processus Python.

1.1. Choix des contraintes et des facteurs

Nous avons choisi d'appliquer deux contraintes à notre étude. La première concerne l'emprise géographique qui doit donc se limiter aux 37 communes de l'EPCI¹ de Toulouse Métropole. Nous avons aussi envisagé une contrainte de terrain : nous n'allons par exemple pas planter notre magasin sur une zone d'eau (lac, fleuve...).

Nous avons également identifié plusieurs facteurs qui doivent nous permettre de mener notre évaluation.

¹ Établissement Public de Coopération Intercommunale

Pour implanter un commerce bio, il est d'abord important de s'intéresser aux clients potentiels, nous avons donc étudié plusieurs critères socio-économiques afin de cibler au mieux les consommateurs de produits bio. L'*Agence française pour le développement et la promotion de l'agriculture biologique* (aussi appelée “Agence Bio”) publie annuellement un rapport sur la consommation et la perception des produits bio en France. Leur dernier rapport date de 2024, nous avons donc utilisé ces données récoltées entre novembre et décembre 2023² pour notre étude. Plus précisément, nous nous sommes intéressés aux critères de revenus, de niveau de diplôme, de catégorie d'âge et de catégorie socio-professionnelle.

Dans un second temps, nous avons également dégagé certains critères relatifs à l'architecture urbaine et notamment à l'accessibilité du commerce à implanter. La présence de grands axes routiers facilitera l'approvisionnement des magasins et l'accessibilité pour les clients véhiculés. La proximité de pistes cyclables, pour le commerce de proximité notamment, favorisera l'accessibilité pour les clients adeptes de mobilités douces qui peuvent représenter une clientèle importante pour un commerce proposant des produits issus d'une démarche de production durable.

Enfin, on pourrait également s'intéresser à la présence de commerces concurrents et à l'impact, positif ou négatif, que cela pourrait représenter.

2. Préparation de données

Pour démarrer notre analyse multicritère, il s'agit dans un premier temps de se procurer des données et des informations dont nous allons nous servir. En cohérence avec nos choix d'analyse, nous avons identifié des sources de l'INSEE pour les données socio-économiques qui vont être notre base principale de l'analyse pour l'implantation d'un magasin bio sur la métropole toulousaine.

2.1. Les facteurs socio-économiques

Nous avons donc d'abord souhaité étudier quatre facteurs socio-économiques mentionnés par l'Agence Bio dans son rapport : les critères de revenus, le niveau de diplôme, la catégorie d'âge et la catégorie socio-professionnelle.

² Agence Bio, & L'ObSoCo. (2024). *Baromètre des produits biologiques en France : Consommation et perception - 2024*. Agence Bio.

Ces données, résultats d'un sondage réalisé en 2023, sont communiquées par l'Agence Bio sous la forme d'une infographie (Figure 1) :

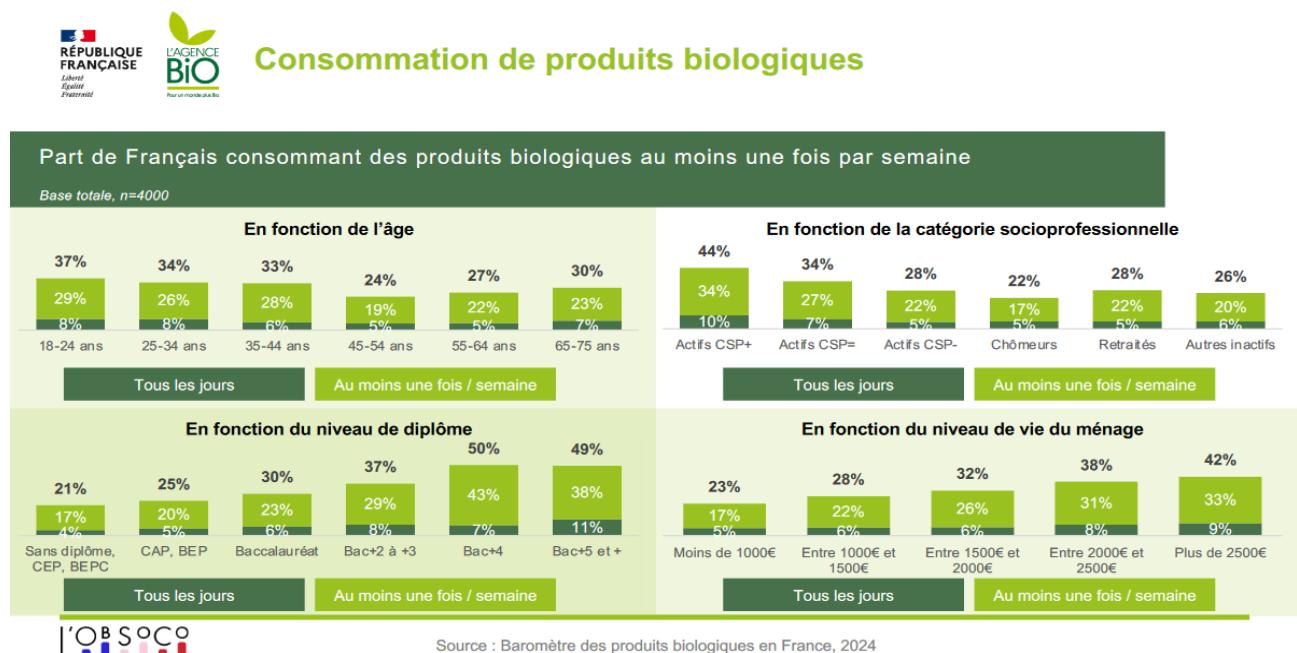


Figure 1 : Répartition des habitudes de consommation de bio chez les français selon quatre facteurs socio-économiques. Source : Agence Bio

Nous avons dans un premier temps transcrit manuellement ces résultats dans quatre fichiers au format .csv afin de pouvoir les traiter en Python (Figure 2) :

1 age,une_fois_par_semaine,tous_les_jours	1 csp,une_fois_par_semaine,tous_les_jours
2 18_24,27,9	2 Actifs_CSP_plus,37,13
3 25_34,32,9	3 Actifs_CSP_egal,35,10
4 35_44,31,9	4 Actifs_CSP_moins,25,7
5 45_54,20,8	5 Chomeurs,17,9
6 55_64,24,8	6 Retraites,19,7
7 65_75,22,7	7 Autres_inactifs,23,6
1 nv_diplome,une_fois_par_semaine,tous_les_jours	1 revenu,une_fois_par_semaine,tous_les_jours
2 Sans_diplome_CEP_BEPC,15,7	2 Moins_de_1000,19,9
3 CAP_BEP,20,6	3 Entre_1000_et_1500,26,5
4 Baccaureat,29,8	4 Entre_1500_et_2000,27,9
5 Bac_plus2_a_plus3,35,10	5 Entre_2000_et_2500,31,12
6 Bac_plus4,40,16	6 Plus_de_2500,37,9
7 Bac_plus5_et_plus,43,12	

Figure 2 : Tables CSV reprenant les résultats du sondage de l'Agence Bio sur les facteur socio-économiques influençant la consommation de produits bio

Les résultats de cette étude, qui donnent la part de consommation de produits bio selon différentes classes de la population et selon différents critères, vont nous permettre d'estimer le poids de chacun de ces critères dans la consommation du bio en France. Nous détaillerons ultérieurement la méthode employée pour établir les poids de ces quatre facteurs.

Pour exploiter ces quatre critères dans notre évaluation multicritère, nous avons dû trouver des données spatialisées correspondant aux facteurs identifiés. Nous avons pour cela utilisé les données de l'INSEE (Tableau 1). Les données relatives à l'âge et aux revenus étaient disponibles dans les données carroyées Filosofi. Pour les diplômes et CSP, nous avons dû utiliser un maillage plus large, celui des IRIS³.

Donnée à représenter	Source de la donnée	Millésime
Âge	INSEE - Filosofi	2019
Revenus	INSEE - Filosofi	2019
Niveau de diplôme	INSEE - IRIS	2020
Catégorie socio-professionnelle	INSEE - IRIS	2020

Tableau 1 : Données socio-économiques traitées et sources.

Les données de l'INSEE et les résultats du rapport de l'Agence Bio, sont répartis par classes. Cependant, les classes de ces deux sources diffèrent, c'est pourquoi il nous a fallu recouper certaines données pour trouver des correspondances entre les classes utilisées.

L'objectif était de s'aligner autant que possible sur les classes présentes dans le rapport sur le bio, de manière à pouvoir appliquer aux données de l'INSEE les pourcentages correspondant pour chaque classe de consommateur.

Pour les CSP, les classes d'âge et les niveaux de diplôme, nous avons établi ces relations :

³ Îlots Regroupés pour l'Information Statistique

Classes du rapport de l'Agence Bio		INSEE - Recensement de la population
Actifs CSP +		Cadres, profession intellectuelle supérieure
Actifs CSP =		Artisans, commerçants, chefs d'entreprise, professions intermédiaires
Actifs CSP -		Agriculteurs exploitants, employés, ouvriers
Chômeurs		Chômeurs
Retraités		Retraités
Autres		Différence entre la somme les personnes de 18 ans ou plus et toutes les catégories précédentes

Tableau 2 : Relations établies entre les CSP de l'Agence Bio et de l'INSEE.

Classes du rapport de l'Agence Bio	Regroupement	Données carroyées
18-24 ans		18-24 ans
25-34 ans	18-39 ans	25-39 ans
35-44 ans		40-54 ans
45-54 ans	40-64 ans	/
55-64 ans		55-64 ans
65-75 ans	Plus de 65 ans	65-79 ans
/		Plus de 80 ans

Tableau 3: Relations établies entre les classes d'âge de l'Agence Bio et de l'INSEE.

Classes du rapport de l'Agence Bio		Données IRIS
	Bio	
Sans diplôme, CEP, BEPC		Diplôme minimal, BEPC
CAP, BEP		CAP, BEP
Baccalauréat		Baccalauréat
Bac+2 à +3		Bac +2, Bac +3 ou 4
Bac +4		Bac +3 ou 4
Bac +5 et plus		Bac + 5

Tableau 4: Relations établies entre les niveaux de diplôme de l'Agence Bio et de l'INSEE.

En ce qui concerne les revenus, le rapport de l'Agence Bio présente des pourcentages de consommateurs de produits bio par tranche de revenus mensuels moyens. La donnée Filosofi sur les revenus se présente en revanche dans un champ "*somme des niveaux de vie winsorisés*" qui correspond aux revenus annuels dont on a traité les valeurs pour atténuer les plus extrêmes.

Pour exploiter ces données, nous avons donc divisé la "somme des niveaux de vie winsorisés" par 12 pour obtenir un revenu mensuel moyen pour chaque et nous les avons associé à la classe correspondante parmi celles du rapport de l'Agence Bio.

Après avoir géré la problématique des classes qui diffèrent, il y a eu un autre travail à faire pour mettre les quatre données au même niveau. En effet, les données socio-économiques à notre disposition se présentent avec deux maillages différents (carreaux Filosofi et IRIS) or, pour pouvoir produire une carte d'évaluation multicritère, il est plus commode de manipuler les données à la même granulométrie. Nous avons choisi d'étudier les données à l'échelle du carreau Filosofi pour plus de précision (Figure 3).

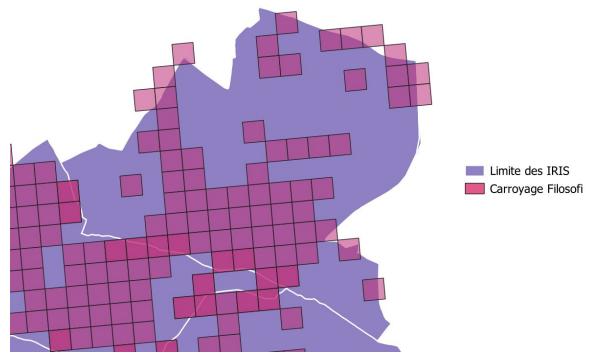


Figure 3 : Illustration des différences d'échelles géographiques entre les limites d'iris et les carroyages Filosofi.

La problématique était donc de ramener au carreau les données de l'IRIS en produisant une estimation. Pour les indicateurs du diplôme et de la catégorie socioprofessionnelle, nous avons commencé par établir la part de chacune des catégories indiquées dans les tableaux précédents au sein de chaque IRIS. Ensuite, pour chaque carreau INSEE, nous avons estimé la population active en sommant les tranches d'âge de 18 à 64 ans.

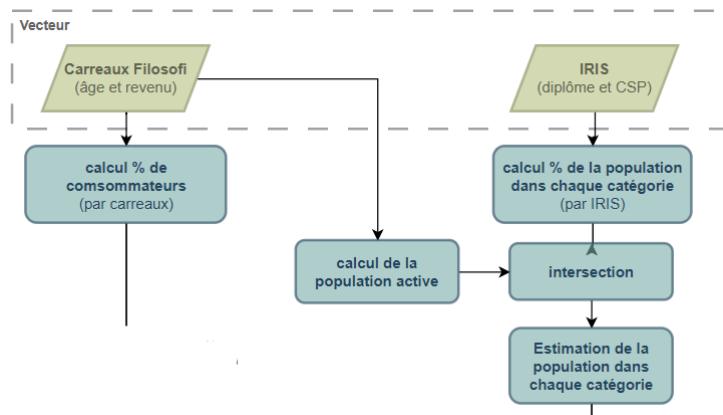


Figure 4 : Diagramme de flux des traitements pour ramener les données IRIS à l'échelle des carreaux Filosofi.

Nous avons intersecté la couche IRIS avec la couche de carreaux pour identifier à quel IRIS se rattache chaque carreau puis avons multiplié la population active du carreau par les parts de chaque catégorie de personnes à l'échelle de l'IRIS pour obtenir ainsi une estimation du nombre de personnes dans chaque catégorie au niveau du carreau. De cette manière, toutes nos données sont à la maille carreau, certaines estimées et d'autres officielles, et nous pouvons les utiliser pour notre analyse multicritère (Figure 4).

2.2. Les données relatives au terrain et à l'architecture urbaine

Les données ont été téléchargées et prétraitées pour être découpées en fonction de l'emprise géographique de Toulouse Métropole. Les limites communales, les réseaux routiers ainsi que les zones en eau sont issues des données de la BDTOPO de l'IGN. Les pistes cyclables ont été extraites de la base des Aménagements cyclables de France Métropolitaine, produite par Géovélo. Les emplacements des magasins bio existants ont été extraits de la base de données OSM avec Overpass Turbo via la requête suivante :

```
1 [out:json];
2 area["name"="Toulouse"]["boundary"]="administrative"]->.a;
3 (
4   node["shop"]="supermarket"]["organic"~"yes|only"](.a);
5   way["shop"]="supermarket"]["organic"~"yes|only"](.a);
6   relation["shop"]="supermarket"]["organic"~"yes|only"](.a);
7 );
8 out body;
9 >;
10 out skel qt;
```

Figure 5 : Requête Overpass Turbo pour obtenir la localisation des magasins bio dans Toulouse Métropole depuis la base OpenStreetMap.

Donnée à représenter	Source de la donnée	Millésime
Limites des communes	IGN - BDTOPO	2024
Routes	IGN - BDTOPO	2024
Zones en eau	IGN - BDTOPO	2024
Pistes cyclables	Géovélo	2024
Magasins bio	OSM	2024

Tableau 5 : Les données géographiques traitées et leur source.

3. Traitement des données

3.1. Rastérisation des contraintes

Deux contraintes étaient à prendre en compte : l'emprise géographique devait se limiter aux 37 communes de Toulouse Métropole et nous avons choisi une contrainte de terrain pour exclure les zones en eau de notre étude.

3.1.1. Limitation de la zone d'étude

Nous avons d'abord traité l'emprise géographique avec QGIS en réalisant un masque raster à partir des limites des communes de la Haute-Garonne en filtrant celles appartenant à Toulouse Métropole (Figure 6) :



Figure 6 : Obtention du masque de Toulouse Métropole grâce aux codes INSEE des communes

Requête utilisé :

```
"inseecom" IN ('31003', '31022', '31032', '31044', '31053',
'31056', '31069', '31088', '31091', '31116', '31149', '31150',
'31157', '31163', '31182', '31184', '31186', '31205', '31230',
'31282', '31293', '31351', '31352', '31355', '31389', '31417',
'31418', '31445', '31467', '31488', '31490', '31506', '31541',
'31555', '31557', '31561', '31588')
```

À partir de ce masque nous avons pu découper toutes nos autres données pour les restreindre à notre zone d'étude. Notamment les données de la BDTOPO relatives aux zones en eau que nous allons être amenés à utiliser pour la contrainte correspondante mais aussi celles relatives aux routes et voies cyclables. Ces données ont été nettoyées, découpées selon le raster d'emprise de Toulouse Métropole, et organisées dans un même GeoPackage (*.gpkg) pour un traitement ultérieur en utilisant le langage Python.

Le flux complet du travail peut être vu dans la figure 7. Dans les sections suivantes, nous détaillerons chacun des processus réalisés.

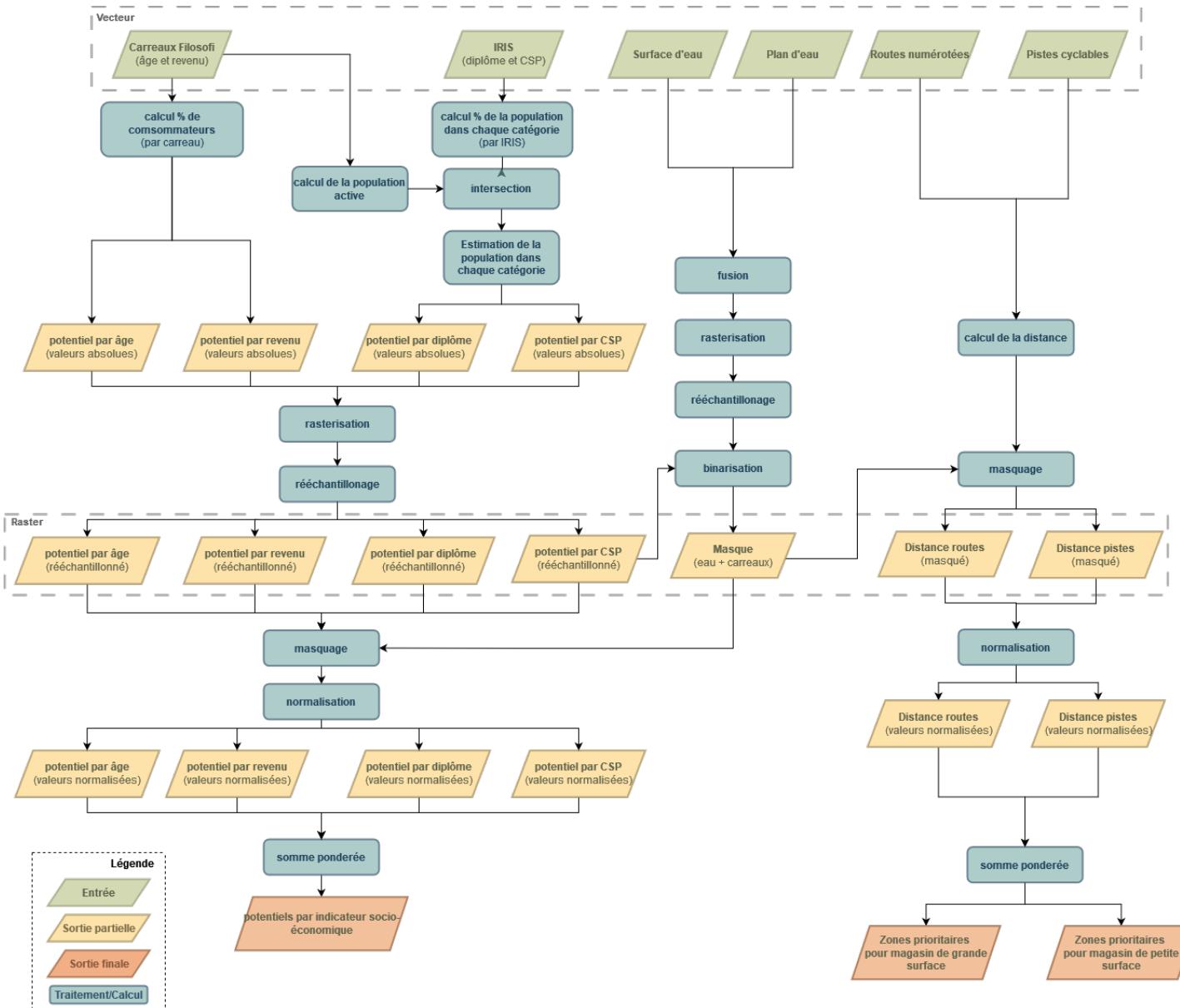


Figure 7 : Diagramme de flux avec tous les géotraitements réalisés

3.1.2. Exclusion des carreaux situés dans des zones en eau

Notre étude doit se concentrer sur les zones “peuplées” de Toulouse Métropole, qui correspondent aux carreaux Filosofi sur lesquels il existe des données INSEE. Nous avons donc affiné notre masque de Toulouse Métropole en y soustrayant ces zones qui ne sont pas concernées par des carreaux Filosofi.

En plus de ça, nous avons aussi retiré les carreaux où l'implantation d'un magasin bio sera dans tous les cas impossible à cause d'une contrainte de terrain, les zones en eau par exemple. Pour procéder à ce traitement, nous avons donc produit un raster des zones d'eau sur la zone de Toulouse Métropole à partir des couches correspondantes (`plan_d_eau` et `zone_hydrographique`) de la BDTOPO (Figure 8).

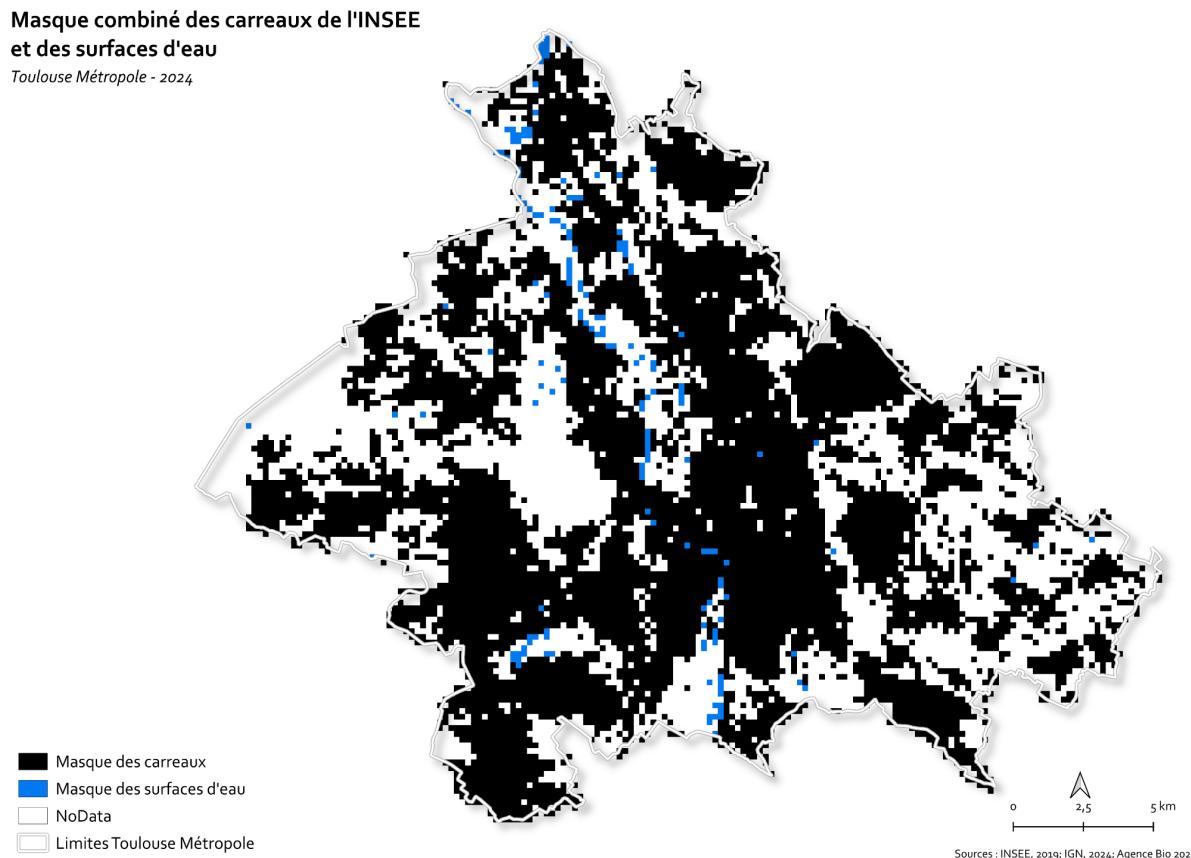


Figure 8 : Masque combiné des carreaux de l'INSEE et des surfaces d'eau

Il aurait été possible de définir d'autres contraintes en excluant certains types de surfaces de la BDTOPO, comme les zones de végétation, les aérodromes ou les réservoirs de biodiversité. Toutefois, nous avons opté pour des critères plus pragmatiques, en restreignant l'analyse à l'emprise de Toulouse Métropole et aux zones hors eau, afin de simplifier le

processus. L'objectif n'est pas de fournir une localisation exacte pour l'implantation d'un magasin bio, mais plutôt d'identifier un secteur favorable à leur développement.

3.2. Estimation du nombre d'individus consommateurs de produits biologiques

À partir des facteurs socio-économiques identifiés dans le rapport de l'Agence Bio, nous avons voulu estimer, sur notre maillage, le nombre de clients potentiels par carreau.

Comme nous avons au préalable ramené les données des IRIS à la maille carreau, toutes les données ont le même format.

Nous avons donc procédé de manière similaire dans les quatre cas, nous n'allons donner qu'un exemple pour le facteur "âge".

3.2.1. Méthode

Pour mesurer le nombre de clients potentiels en fonction d'un facteur, nous avons sommé les produits, pour chaque classe, du taux de consommateurs (en sommant les consommateurs quotidiens et hebdomadaires) par le nombre d'individus correspondant à cette classe dans chaque carreau INSEE.

Ce qui donne, par exemple, pour le facteur âge en trois classes :

```
Nombre de clients potentiels du carreau =  
(Nombre d'individus 18-39 ans * Taux de  
consommateurs bio 18-39 ans)  
+ (Nombre d'individus 40-64 ans * Taux de  
consommateurs bio 40-64 ans)  
+ (Nombre d'individus 65 ans et plus *  
Taux de consommateurs bio 65 ans et plus)
```

- *D'après nos données prétraitées issues des données INSEE*
- *D'après les résultats du rapport de l'Agence Bio*

3.2.2. Processus avec Python

Nous avons effectué cette étape de calcul avec Python (Figure 9), car la manipulation de données y est grandement facilitée. Après avoir établi les classes d'âge indiquées avec les données de l'INSEE, nous avons appliqué les taux de consommation de bio correspondants issus du rapport de l'Agence Bio. Le bloc de code ci-dessous décrit ce procédé :

```

1 # Calculer les tranches d'âge définies
2 # On regroupe les individus en trois tranches d'âge :
3 # 1. 18-39 ans (somme des individus de 18-24 ans et de 25-39 ans)
4 # 2. 40-64 ans (somme des individus de 40-54 ans et de 55-64 ans)
5 # 3. 65 ans et plus (somme des individus de 65-79 ans et de plus de 80 ans)
6 carreaux_insee.loc[:, 'age_18_39'] = carreaux_insee['ind_18_24'] + carreaux_insee['ind_25_39']
7 carreaux_insee.loc[:, 'age_40_64'] = carreaux_insee['ind_40_54'] + carreaux_insee['ind_55_64']
8 carreaux_insee.loc[:, 'age_65p'] = carreaux_insee['ind_65_79'] + carreaux_insee['ind_80p']
9
10 # Appliquer les pourcentages moyens pour estimer les consommateurs bio potentiels par tranche d'âge
11 # Chaque tranche d'âge a un pourcentage associé à la consommation bio.
12 # Par exemple : 35.5% des 18-39 ans sont des consommateurs bio potentiels.
13 carreaux_insee.loc[:, 'potentiel_age'] = (
14     carreaux_insee['age_18_39'] * 0.355 + # Pourcentage pour la tranche 18-39 ans
15     carreaux_insee['age_40_64'] * 0.285 + # Pourcentage pour la tranche 40-64 ans
16     carreaux_insee['age_65p'] * 0.285      # Pourcentage pour la tranche 65 ans et plus
17 )
18
19 # Sélectionner les colonnes pertinentes et créer un nouveau GeoDataFrame avec
20 # les informations sur l'âge, l'individu et le potentiel bio selon l'âge.
21 colonnes_age_calc = ['idcar_200m', 'ind', 'potentiel_age', 'geometry']
22 carreaux_age = carreaux_insee[colonnes_age_calc]

```

Figure 9 : Extrait du code Python permettant d'estimer le nombre de consommateurs potentiels de bio grâce aux données INSEE et de l'Agence Bio

Une fois ce processus exécuté pour les quatre facteurs, on obtient des tableaux de données qui nous donnent le nombre de consommateurs potentiels de bio par carreau, et par critère.

Comme les critères peuvent se recouper entre eux, on a normalisé ce nombre de consommateurs potentiels sur une plage de valeur allant de 0 à 1. On utilise la normalisation min-max avec Python (Figure 10). On obtient donc pour chaque carreau un score de 0 à 1 sur le potentiel de consommation bio de ce carreau pour le critère correspondant.

```

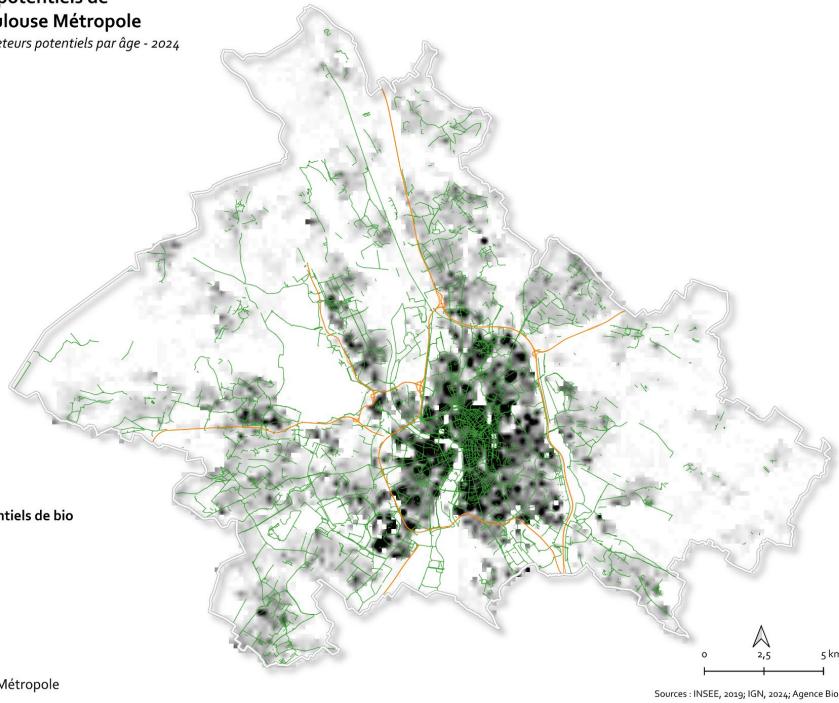
1 def normaliser_raster(input_file, output_file, inverser=False):
2     """
3         Normalise un raster en définissant NaN pour NoData, applique une compression LZW,
4         et enregistre le résultat. Inverse les valeurs si nécessaire.
5     """
6     input_file : Chemin du raster d'entrée
7     output_file : Chemin du raster normalisé
8     inverser : Si True, inverse les valeurs normalisées (1 - valeur)
9     """
10    with rasterio.open(input_file) as src:
11        data = src.read(1)
12
13        # Remplace NoData par NaN pour ignorer ces valeurs dans les calculs
14        if src.nodata is not None:
15            data[data == src.nodata] = np.nan
16
17        # Normalise les données entre 0 et 1, si elles ne sont pas toutes NaN
18        if not np.isnan(data).all():
19            min_val, max_val = np.nanmin(data), np.nanmax(data)
20            print(f"Valeur minimale: {min_val}, Valeur maximale: {max_val}") # Diagnostique
21
22        # Normalisation
23        if min_val != max_val:
24            normalized_data = (data - min_val) / (max_val - min_val)
25        else:
26            normalized_data = np.zeros_like(data) # Cas où tous les valeurs sont identiques
27
28        # Inverser les valeurs si nécessaire
29        if inverser:
30            normalized_data = 1 - normalized_data
31            print("Inversion appliquée aux valeurs normalisées.")
32
33        # Vérifier les valeurs normalisées
34        print(f"Intervalle des données normalisées: {np.nanmin(normalized_data)} à {np.nanmax(normalized_data)}")
35
36        # Met à jour le profil pour NaN comme NoData et ajoute la compression LZW
37        profile = src.profile
38        profile.update(dtype='float32', nodata=np.nan, compress='LZW')
39
40        # Sauvegarde le raster normalisé avec compression
41        with rasterio.open(output_file, 'w', **profile) as dst:
42            dst.write(normalized_data, 1)
43
44        print(f"Raster normalisé sauvé: {output_file}")

```

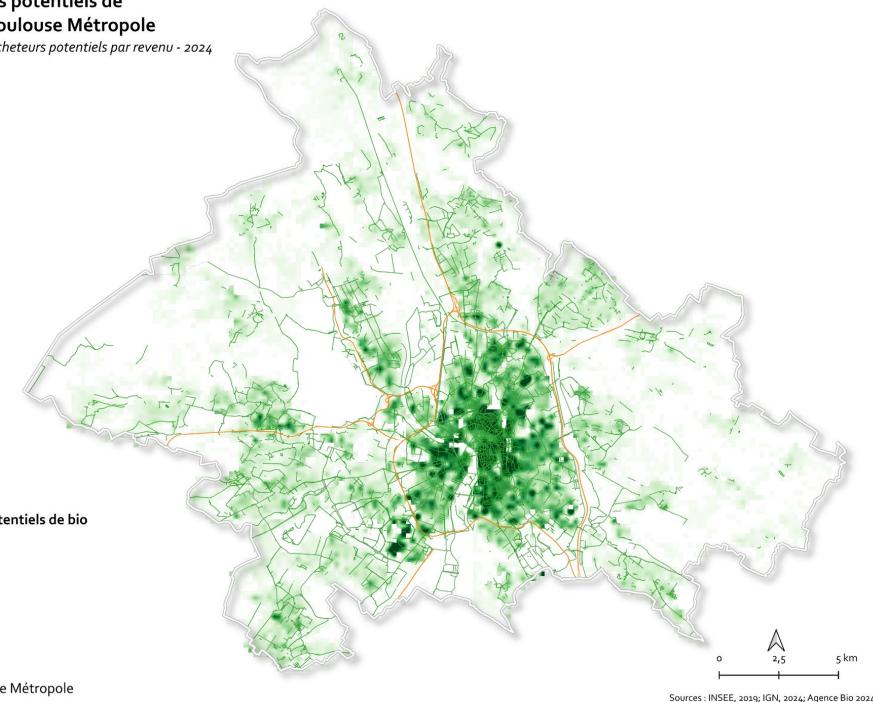
Figure 10 : Extrait du code Python permettant la normalisation des nombres de consommateurs de bio estimés par chacun des facteurs socio-économiques.

Avec cette sortie de données, une première mise en page peut être réalisée dans QGIS. Les cartes obtenues sont présentées ci-dessous et illustrent chacune des quatre variables socio-économiques étudiées.

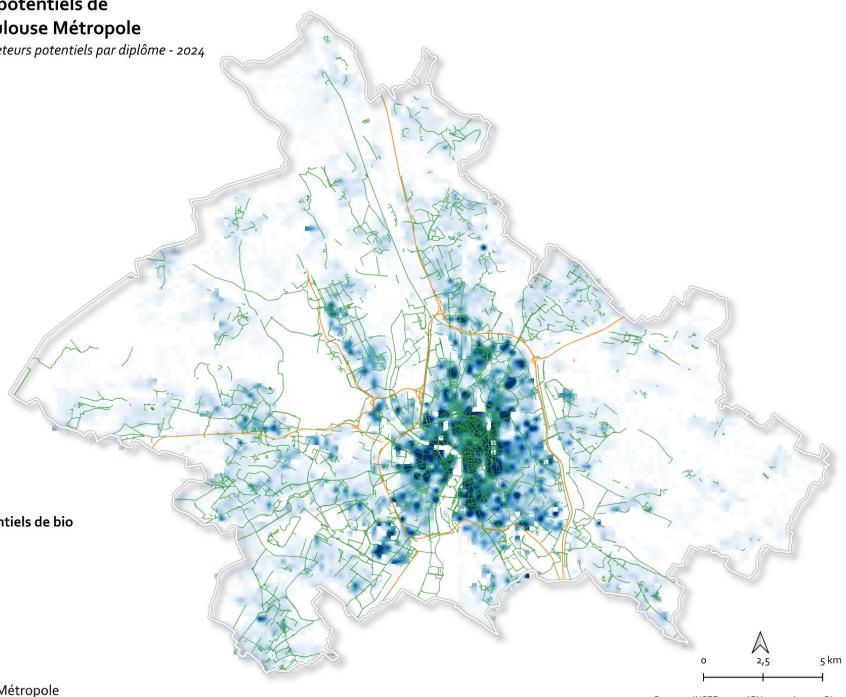
Consommateurs potentiels de produits bio à Toulouse Métropole
Indice normalisé des acheteurs potentiels par âge - 2024



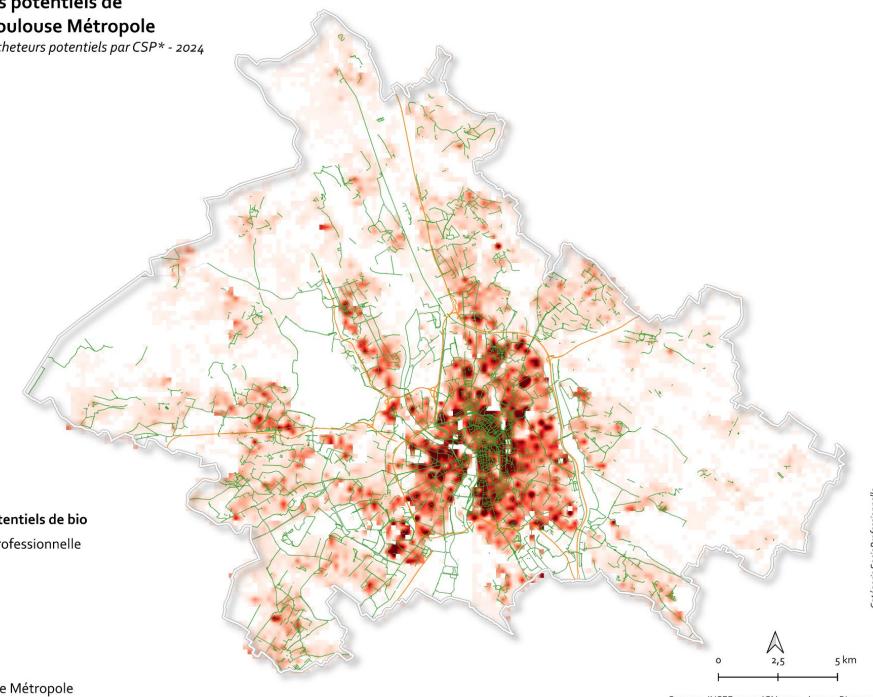
Consommateurs potentiels de produits bio à Toulouse Métropole
Indice normalisé des acheteurs potentiels par revenu - 2024



Consommateurs potentiels de produits bio à Toulouse Métropole
Indice normalisé des acheteurs potentiels par diplôme - 2024



Consommateurs potentiels de produits bio à Toulouse Métropole
Indice normalisé des acheteurs potentiels par CSP* - 2024



3.3. Résultats et critique

Il apparaît clairement que les cartes produites pour chaque facteur sont quasiment identiques. En effet, les quatre facteurs que nous avons étudiés (CSP, âge, niveau de revenu et de diplôme) sont extrêmement corrélés.

Si on ne prend en compte que la clientèle ciblée, les zones les plus propices à l'implantation d'un magasin bio sont les plus peuplées et les plus riches, à savoir la ville de Toulouse dans ses zones urbaines et résidentielles ainsi que certaines zones résidentielles des communes de sa proche couronne.

Pour mener à bien cet exercice théorique, nous avons choisi de passer outre la forte corrélation des facteurs socio-économiques étudiés et de les traiter comme s'ils étaient indépendants. L'étape suivante de notre travail a donc consisté à déterminer le poids à donner à chacun de ces facteurs, autrement dit l'impact de chaque facteur sur les habitudes de consommation.

3.4. Pondération des facteurs socio-économiques

Après avoir préparé comme il se doit toutes les données nécessaires, nous pouvons maintenant prendre en compte le poids des quatre facteurs socio-économiques dans notre étude. Les données relatives à ces facteurs, issues du rapport de l'Agence Bio, sont exprimées en pourcentage d'individus consommant des produits biologiques au moins une fois par semaine. Les variations en fonction des classes sont plus ou moins importantes selon si l'on regarde le facteur âge, la CSP, le niveau de diplôme ou les revenus.

L'objectif est de déterminer quelles variables socio-économiques sont les plus déterminantes dans la consommation hebdomadaire ou non de produits biologiques afin de pouvoir pondérer ces facteurs dans notre étude.

3.4.1. Méthode théorique

Une façon de déterminer les poids de chaque variable aurait pu être de passer par une Analyse en Composantes Principales (ACP) qui aurait permis de mettre en évidence, dans une matrice de corrélation, les variables les plus déterminantes. Les données brutes de l'étude de l'Agence Bio n'étaient cependant pas disponibles ce qui rendait l'idée de faire une ACP irréalisable. Nous avons donc défini une autre méthode.

La méthode que nous avons appliquée consiste à calculer l'écart-type des différents taux de consommation du bio pour chaque critère d'étude, afin de mesurer la dispersion des comportements de consommation au sein de chacun. En effet, on considère que plus l'écart-

type est élevé, plus les variations de consommations sont marquées, donc que la variable est plus pertinente à analyser parce qu'elle a une influence plus significative sur les tendances de consommation de produits bio.

En procédant encore une fois avec Python, nous avons calculé les écarts-types pour chaque série de valeur grâce aux données que l'on a précédemment réunies dans des fichiers CSV. Une fois les écarts-types calculés, nous les avons normalisés en appliquant la formule :

$$Poids\ variable\ i = \frac{\sigma_i}{\sum \sigma}$$

On obtient ainsi, pour chaque variable, une valeur de poids entre 0 et 1 relative à l'importance qu'on lui octroie par rapport aux autres.

Variables	Ecart-type (σ_i)	$Poids\ variable\ i = \frac{\sigma_i}{\sum \sigma}$
Âge	5.05	0.15
Catégorie socio-professionnelle (CSP)	9.41	0.27
Niveau de diplôme	13.13	0.38
Niveau de vie	6.85	0.20

Tableau 6 : Écarts-types calculés et poids de chacun des facteurs socio-économiques

3.4.2. Processus avec Python

Pour procéder à ces calculs nous avons utilisé un script Python. C'est ici que nous utilisons les résultats de l'étude que nous avions retranscrit dans des fichiers CSV.

Ces quatres fichiers ont été chargés dans des DataFrame nommés respectivement selon leur contenu, par exemple “age_data” pour les statistiques de consommation du bio selon l'âge (Figure 12).

```

1 # Calcul des écarts-types des parts de consommation pour chaque facteur
2 # Avec la fonction "std" de numpy
3 var = {"age": np.std(age_data["total_consumption"]),
4         "csp": np.std(csp_data["total_consumption"]),
5         "diplome": np.std(diplome_data["total_consumption"]),
6         "revenu": np.std(revenu_data["total_consumption"])}
7
8 # Normalisation des écarts-types pour en faire des poids (valeurs de 0 à 1
9 # dont la somme fait 1)
10 total_std = sum(var.values())
11 poids_normalises = {key: value / total_std for key, value in var.items()}


```

Figure 12 : Extrait du code Python permettant de calculer les écarts-types et les poids des facteurs socio-économiques

À ce stade, nous disposons d'une couche de carreaux pour chaque facteur, avec des indices normalisés de 0 à 1. Pour obtenir une vue globale et intégrée, nous appliquons, à chaque facteur, la pondération que nous venons de calculer. La méthode consiste à multiplier l'indice de chaque facteur par le poids de la variable correspondante, puis à sommer ces produits. Cette somme pondérée des indices permet de combiner les quatre variables en une seule couche composite, reflétant leur influence relative.

En appliquant cette méthode, nous obtenons une carte synthétique qui montre la distribution potentielle des consommateurs de produits bio en prenant en compte l'ensemble des variables socio-économiques de manière pondérée (Figure 11). Cette première partie révèle surtout qu'une implémentation serait préférable dans les quartiers du centre de Toulouse.

Consommateurs potentiels de produits bio à Toulouse Métropole
Indice normalisé des acheteurs potentiels pour les indicateurs socio-économique confondus - 2024

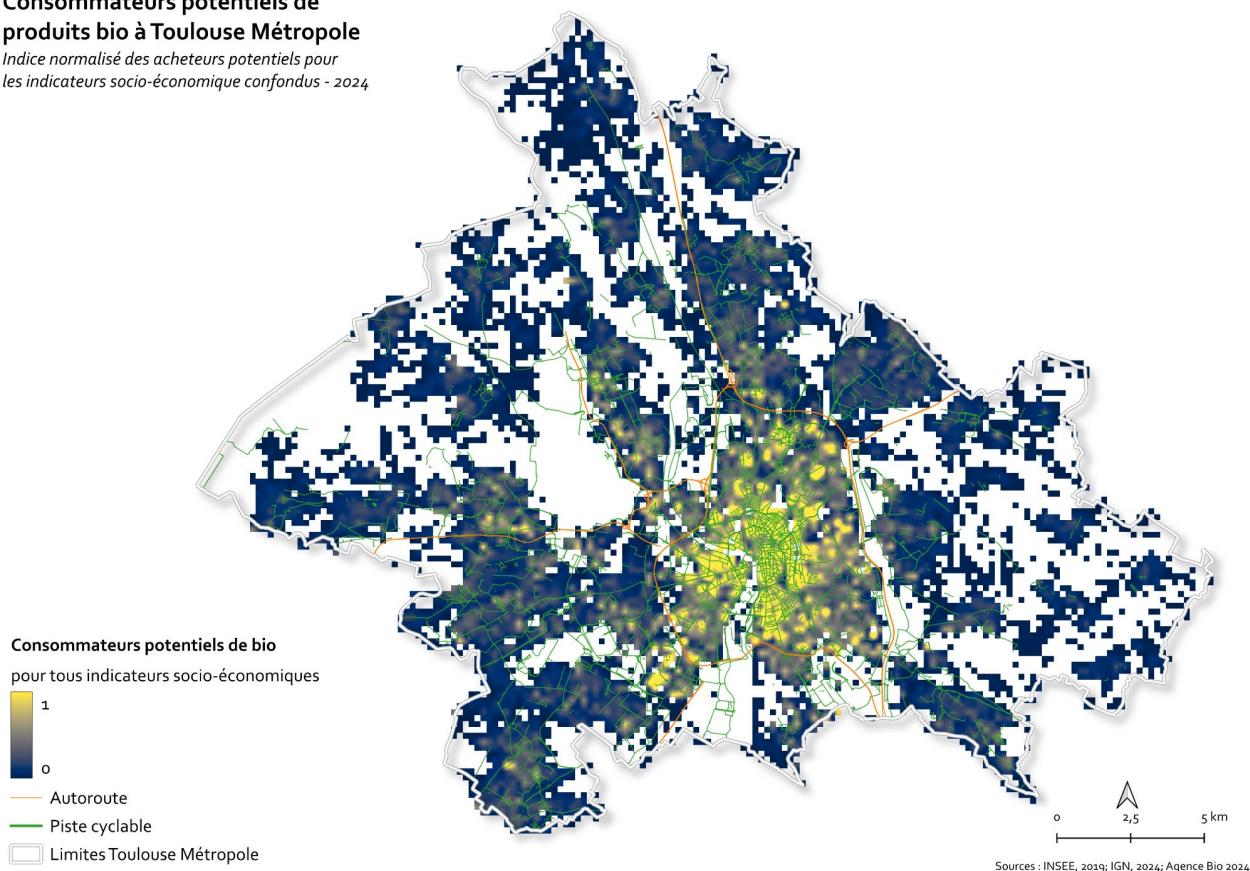


Figure 11 : Consommateurs potentiels de produits bio à Toulouse Métropole

3.5. Différenciation des scénarios avec la prise en compte de l'accessibilité

Les scénarios pour l'implantation d'un magasin bio sont modulés en fonction du type de mobilité préféré pour le site sélectionné. Pour un magasin en centre-ville, l'accent est mis sur l'accessibilité via les pistes cyclables, tandis que pour un site en périphérie, l'accès aux principales voies de circulation de la métropole toulousaine est privilégié. Pour évaluer ces deux scénarios, nous avons donc créé deux rasters de distance avec Python (Figure 12) couvrant la zone de Toulouse Métropole : l'un basé sur les pistes cyclables, l'autre sur les routes principales.

```

1 def calculate_distance(layer, mask_layer, output_name, resolution):
2     # Calculer la taille du raster en fonction de la résolution
3     n_cols = int(np.ceil((maxx - minx) / resolution))
4     n_rows = int(np.ceil((maxy - miny) / resolution))
5     transform = from_origin(minx, maxy, resolution, resolution)
6     raster_shape = (n_rows, n_cols)
7
8     # Rasteriser la couche de Lignes
9     shapes = ((geom, 1) for geom in layer.geometry)
10    rasterized = features.rasterize(
11        shapes=shapes,
12        out_shape=raster_shape,
13        transform=transform,
14        fill=0,
15        all_touched=True,
16        dtype=np.uint8
17    )
18
19    # Calcul de la distance
20    distance = ndimage.distance_transform_edt(1 - rasterized) * resolution
21    distance = distance.astype('float32')
22
23    # Appliquer le masque du département
24    shapes_mask = ((geom, 1) for geom in mask_layer.geometry)
25    raster_mask = features.rasterize(
26        shapes=shapes_mask,
27        out_shape=raster_shape,
28        transform=transform,
29        fill=0,
30        all_touched=True,
31        dtype=np.uint8
32    )
33    distance_masked = np.where(raster_mask == 1, distance, np.nan)
34
35    # Enregistrer le raster avec compression LZW
36    output_filepath = os.path.join(partial_output_path, output_name)
37    compress_options = {
38        'compress': 'LZW',
39        'predictor': 3,
40        'tiled': True,
41        'blockysize': 256,
42        'blockysize': 256
43    }
44
45    with rasterio.open(
46        output_filepath,
47        'w',
48        driver='GTiff',
49        height=n_rows,
50        width=n_cols,
51        count=1,
52        dtype='float32',
53        crs=layer.crs,
54        transform=transform,
55        nodata=np.nan,
56        **compress_options
57    ) as dst:
58        dst.write(distance_masked, 1)
59    print(f"Le raster est enregistré à l'emplacement : {output_filepath}")

```

Figure 12 : Fonction Python permettant de créer un raster de distance depuis une couche de lignes au format vecteur.

La différence entre les deux scénarios réside dans le raster de distance utilisé comme entrée dans l'analyse. Indépendamment du scénario choisi, nous avons décidé de pondérer les

facteurs en appliquant une répartition de $\frac{2}{3}$ pour les facteurs socio-économiques et $\frac{1}{3}$ pour le facteur lié à l'infrastructure de réseau (les pistes cyclables pour un scénario urbain/petite surface, ou les routes pour un scénario périphérique/grande surface). Cette pondération indique que nous avons choisi de donner plus d'importance aux facteurs socio-économiques pour identifier une clientèle potentielle. Toutefois, l'accessibilité reste un critère significatif, d'où sa part de $\frac{1}{3}$ dans le calcul final. Mais avant d'appliquer les pondérations sur ces facteurs, on nous avons normalisé les rasters de distances pour avoir une plage de valeur de 0 à 1.

3.5.1. Premier scénario : magasin de petite surface en ville

Après avoir établi le raster des distances aux pistes cyclables, nous l'avons intégré dans notre analyse en le pondérant avec un poids de $\frac{1}{3}$ avec notre raster de potentiel de consommateurs bio tous critères socio-économiques confondus. La carte de la Figure 13 représente les proximités normalisées aux pistes cyclables dans Toulouse Métropole.

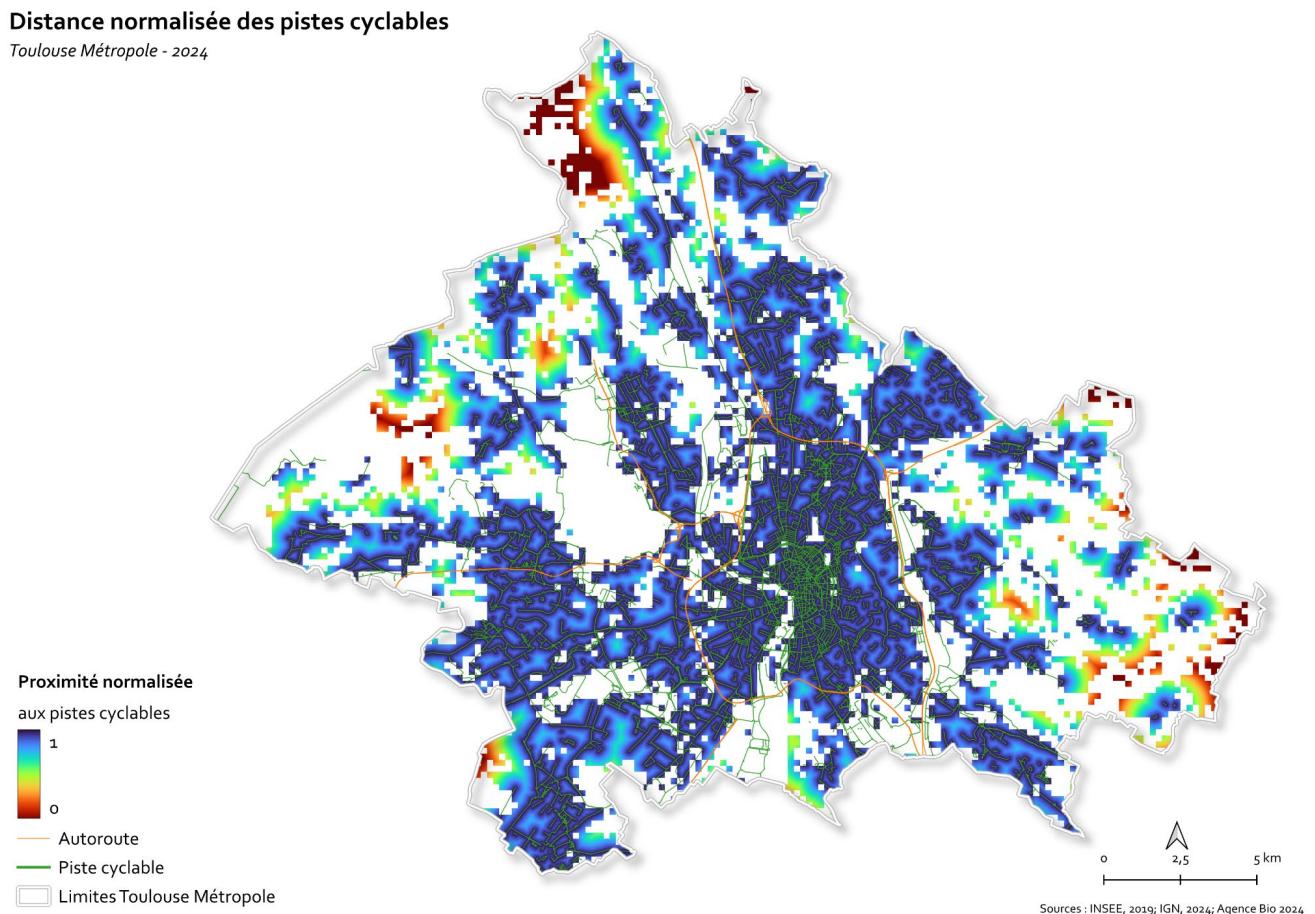


Figure 13 : Proximité normalisée aux pistes cyclables

L'implémentation de ce nouveau facteur nous a permis de générer une carte (Figure 14) des zones propices à l'installation d'un magasin bio. Comme nous pouvons le voir dans cette carte, dans le cadre du premier scénario "petite surface en ville", l'étude multicritère met surtout en évidence les quartiers de la ville de Toulouse, et assez peu les autres villes de la métropole.

Zones prioritaires pour l'installation d'un magasin de produits bio de petite surface

Toulouse Métropole - 2024

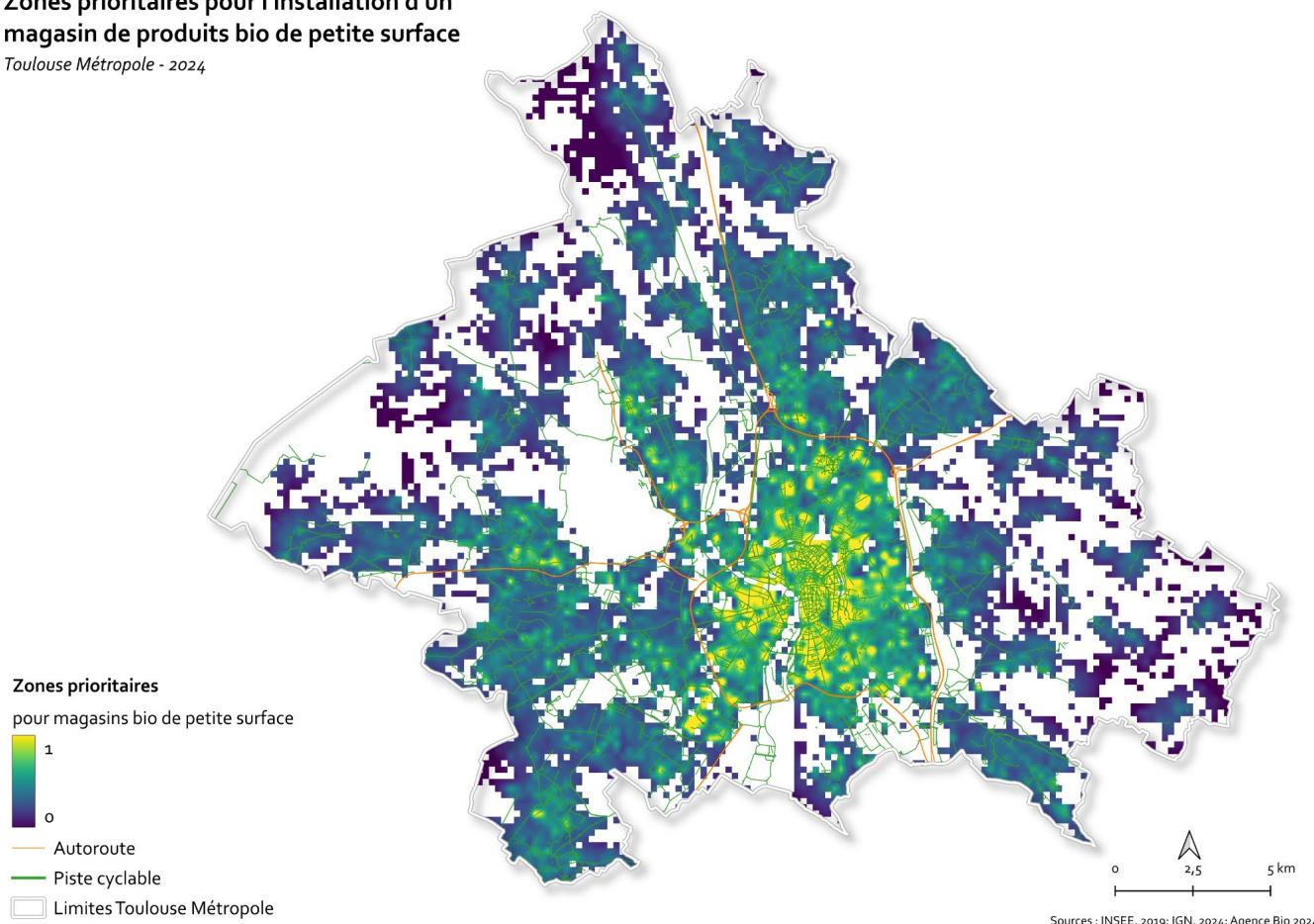


Figure 14 : Carte des zones prioritaires pour l'implantation d'un magasin de produits bio de petite surface

3.5.2. Second scénario : magasin de grande surface en périphérie

Dans ce scénario c'est l'accessibilité par les grandes voies routière qui est privilégiée. On favorise alors les clients qui viendraient en voiture (et correspondrait alors à une population moins urbaine) mais aussi l'accessibilité du magasin pour l'approvisionnement dans la perspective d'implanter une grande surface. La Figure 14 montre cette fois que dans ce second scénario, les quartiers du centre de Toulouse auront tendance à être plus défavorisés par leur distance aux routes principales.

Distance normalisée des tronçons de route

Toulouse Métropole - 2024

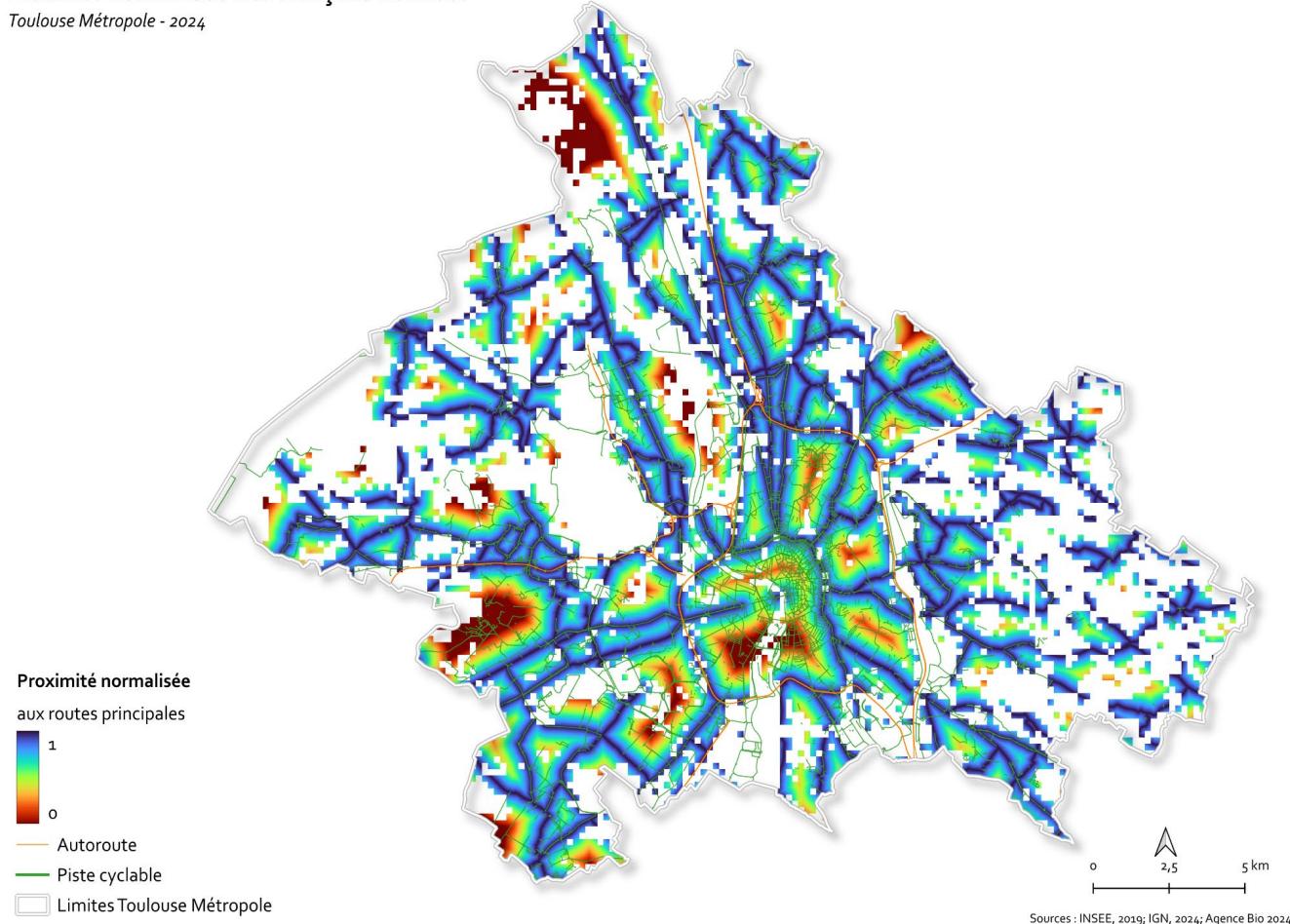


Figure 15 : Proximité normalisée aux routes principales

La carte de la Figure 16 illustre les résultats de l'analyse multicritères dans le cadre du notre second scénario en ajoutant aux critères socio-économiques le facteur de la proximité aux routes principales. Bien que l'accessibilité ait été repensée en mettant davantage l'accent sur le réseau routier, les zones identifiées comme favorables pour l'implantation d'un magasin bio restent globalement similaires. Cela s'explique par la part prépondérante ($\frac{2}{3}$) laissée aux facteurs socio-économiques dans l'analyse mais aussi par l'assez bonne répartition des routes à travers Toulouse Métropole.

Zones prioritaires pour l'installation d'un magasin de produits bio de grande surface
Toulouse Métropole - 2024

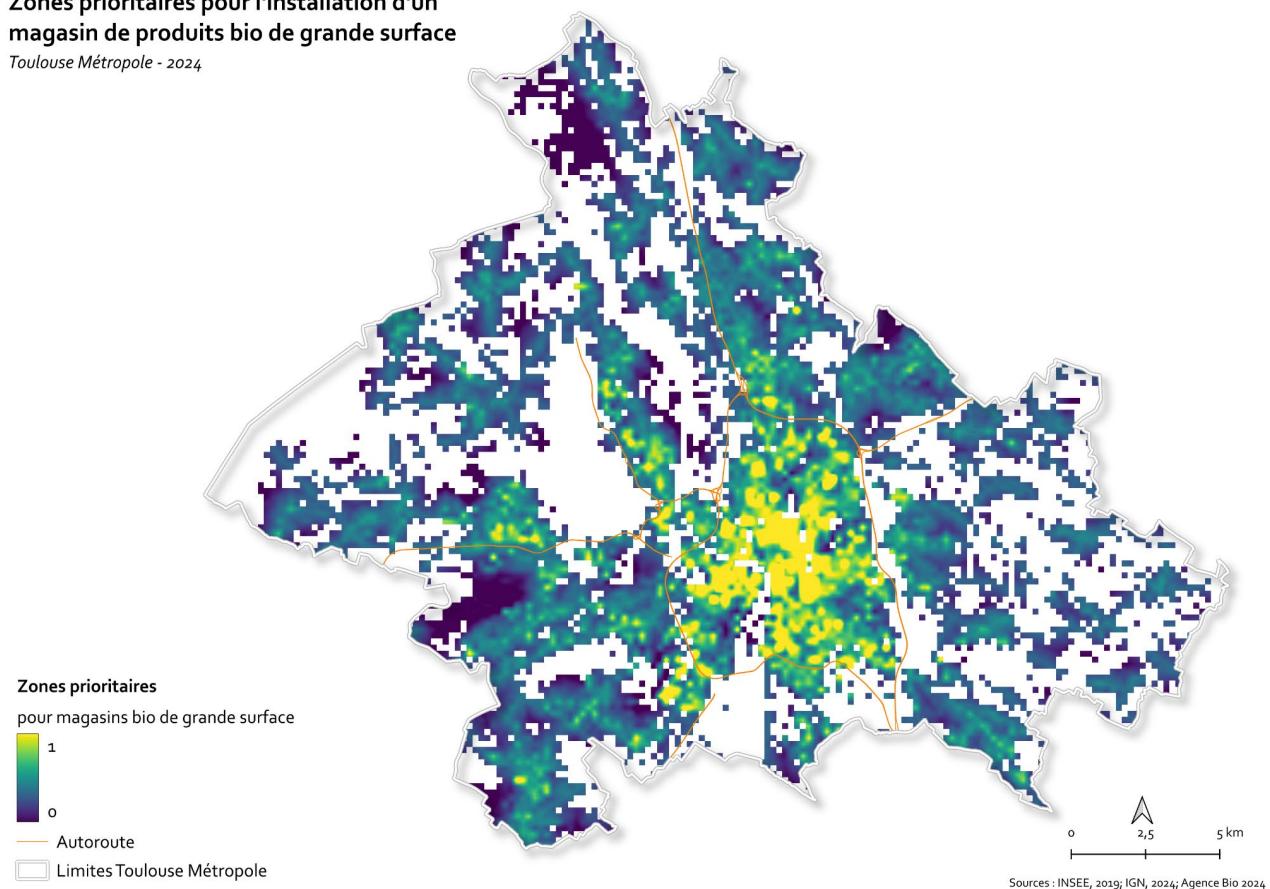


Figure 16 : Cartes des zones prioritaires pour l'implantation d'un magasin de produits bio de grande surface

4. Analyse des potentialités d'implantation

4.1. Des zones qui se démarquent dans notre analyse

Certaines zones se distinguent dans notre évaluation pour l'implantation d'un magasin bio sur la métropole toulousaine. La zone privilégiée qui ressort clairement dans les deux scénarios est le centre-ville de Toulouse, caractérisé par une forte densité de population et un niveau de vie élevé. Cependant, l'analyse met également en lumière des opportunités intéressantes en périphérie, notamment dans des communes telles que Blagnac, Portet-sur-Garonne ou Plaisance-du-Touch. Ces zones bénéficient d'une proximité avec le centre urbain tout en offrant un potentiel pour capter une clientèle locale spécifique.

Bien que les différences entre les scénarios étudiés soient relativement faibles, les zones mentionnées, tant en centre-ville qu'en périphérie proche, semblent les plus prometteuses pour l'implantation d'un magasin bio.

4.2. Prise en compte de la présence actuelle de magasins bio

Pour évaluer la cohérence de notre analyse, il est pertinent de la confronter à la répartition actuelle des magasins bio dans la métropole toulousaine. Comme on peut le voir sur la carte de la Figure 17, cette comparaison montre qu'en centre-ville, la densité de magasins bio est effectivement élevée, ce qui confirme l'attrait de cette zone.

En périphérie, on observe une répartition plus hétérogène. Par exemple, la commune de Blagnac compte déjà deux magasins bio, ce qui reflète un certain dynamisme local. En revanche, à Portet-sur-Garonne, on ne recense qu'un seul magasin bio, malgré un grand nombre de carreaux présentant des scores élevés dans notre analyse. Cette situation suggère un potentiel intéressant pour planter un nouveau magasin bio dans cette commune, afin de répondre à une demande potentielle non satisfaite. Ainsi, en complément de l'attractivité du centre-ville, Portet-sur-Garonne pourrait représenter une opportunité stratégique pour renforcer l'offre en périphérie, particulièrement dans des zones identifiées comme ayant un fort potentiel.

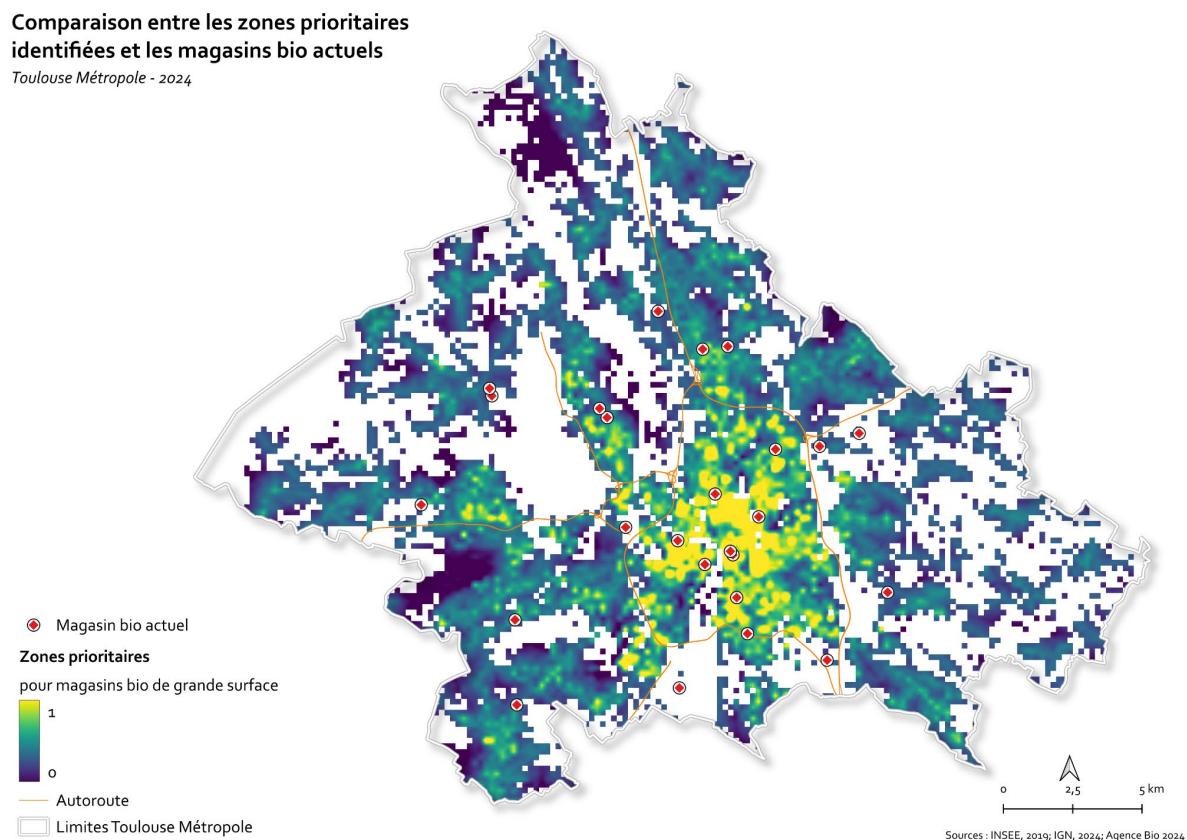


Figure 17 : Comparaison entre les zones prioritaires identifiées et les magasins bio actuels

5. Conclusion

En conclusion, nous avons identifié une zone prometteuse vers Portet-sur-Garonne, mais notre analyse comporte certaines limites. En effet, nous aurions pu enrichir notre démarche en intégrant davantage de critères plutôt que de nous limiter à l'exclusion des zones en eau. Concernant les facteurs socio-économiques, bien qu'ils soient fortement corrélés entre eux, nous les avons traités individuellement. Une analyse en composantes principales (ACP) aurait été pertinente, mais l'absence de données brutes nous en a empêché. De plus, l'estimation du nombre de personnes potentielles dans chaque carreau repose sur une hypothèse simplifiée : une même personne peut cumuler plusieurs critères (par exemple, être titulaire d'un Bac+5, appartenir à une catégorie socio-professionnelle spécifique et correspondre à une tranche d'âge). Cette redondance constitue une limite notable, mais nous avons fait au mieux avec les ressources disponibles.

Cela dit, il est important de rappeler qu'il s'agit d'un exercice théorique. Dans ce cadre, nous nous sommes permis de nombreuses approximations sur les données, ce qui produit des résultats approximatifs. Ces approximations s'expliquent par notre choix de privilégier la méthode plutôt que la précision des résultats. En travaillant avec Python plutôt que TerrSet, nous avons mis l'accent sur une approche adaptable et réutilisable, qui pourrait être appliquée à des données de meilleure qualité si nécessaire.

Ce projet enrichissant nous a permis de travailler notre maîtrise de Python et de différentes bibliothèques en automatisant des traitements géospatiaux avec ce langage. Nous avons également pris conscience de la complexité de ce type d'analyse, tant dans l'identification des facteurs et critères que dans la collecte des données et leur application concrète.

6. Annexes

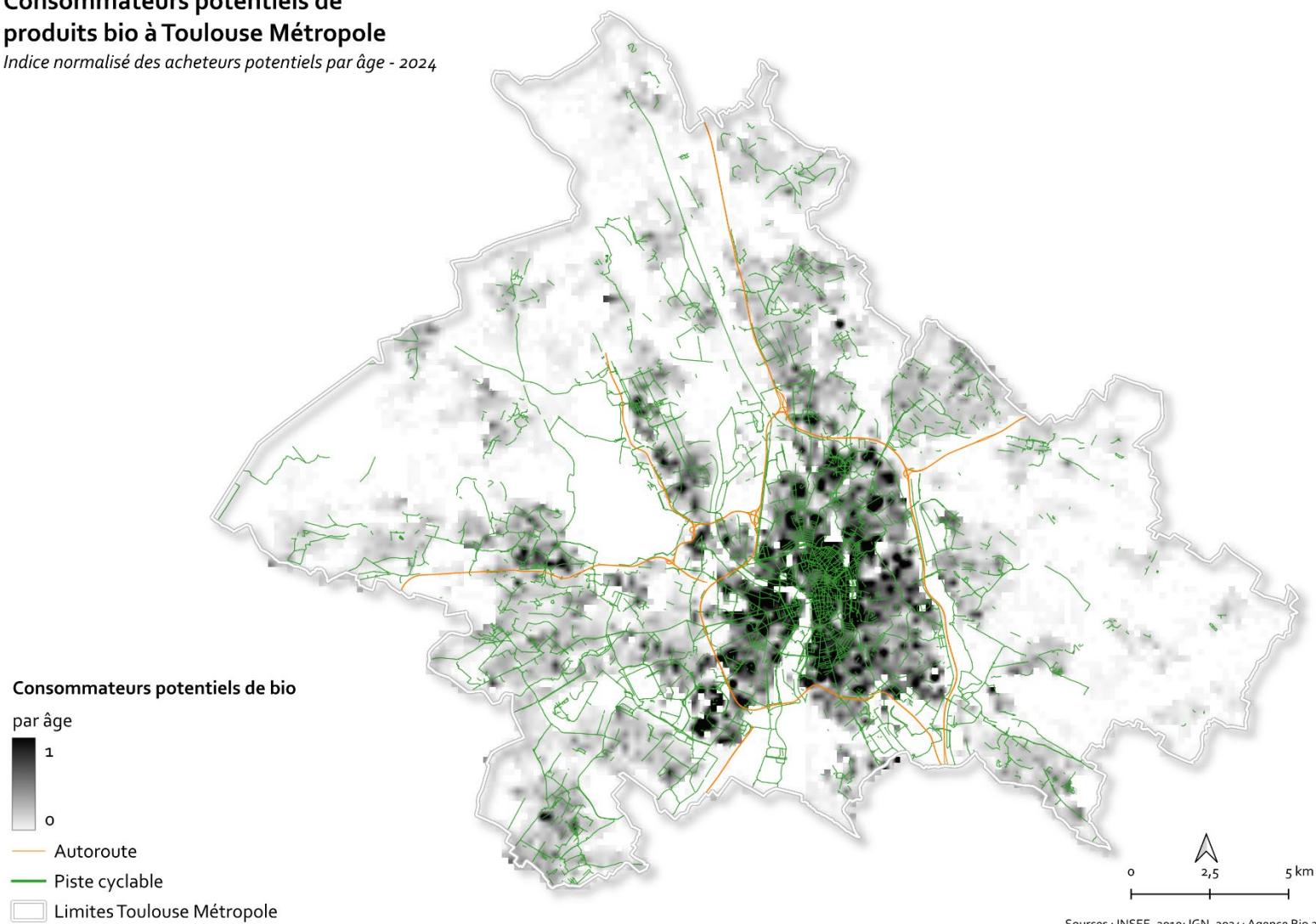
6.1. Bibliothèques Python utilisées

Librairie	Description
Fiona	Utilisée pour lire et écrire des fichiers géospatiaux
Geopandas	Permet de manipuler des données géospatiales en Python avec des opérations sur des DataFrames géographiques.
Matplotlib	Utilisée pour la visualisation des données, en créant des graphiques et des cartes.
Numpy	Fournit des structures de données de tableau multidimensionnel, utilisé ici pour le calcul numérique et le masquage des données raster.
Os	Gère les opérations système, comme la gestion des chemins de fichiers et la création de dossiers.
Rasterio	Manipule les données raster, permettant la lecture, l'écriture et la manipulation des pixels.

6.2. Cartes des consommateurs potentiels de produits bio à Toulouse Métropole par catégorie - grand format

Consommateurs potentiels de produits bio à Toulouse Métropole

Indice normalisé des acheteurs potentiels par âge - 2024



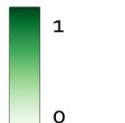
Sources : INSEE, 2019; IGN, 2024; Agence Bio 2024

Consommateurs potentiels de produits bio à Toulouse Métropole

Indice normalisé des acheteurs potentiels par revenu - 2024

Consommateurs potentiels de bio

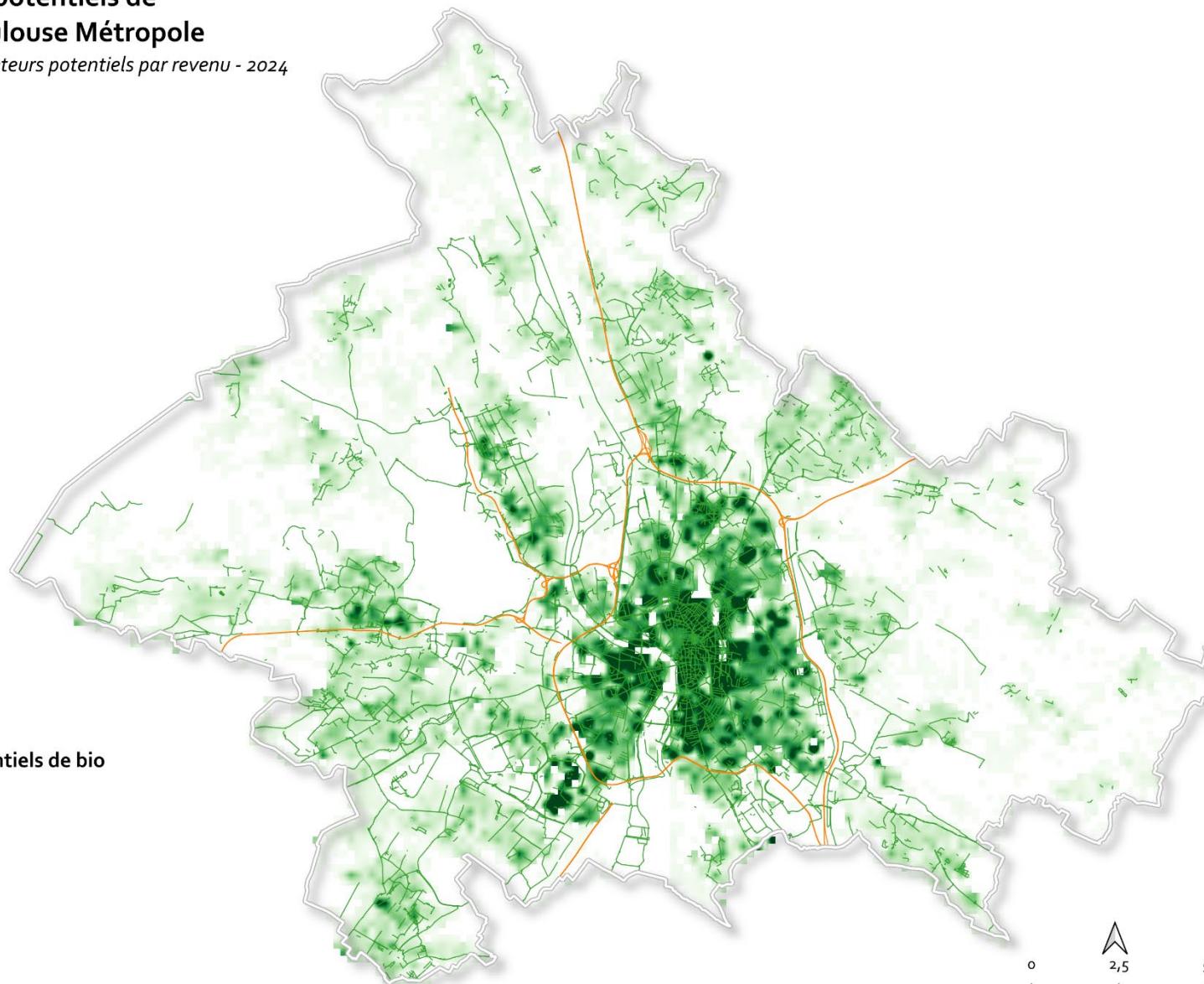
par revenu



Autoroute

Piste cyclable

Limites Toulouse Métropole

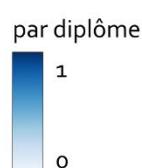


Sources : INSEE, 2019; IGN, 2024; Agence Bio 2024

Consommateurs potentiels de produits bio à Toulouse Métropole

Indice normalisé des acheteurs potentiels par diplôme - 2024

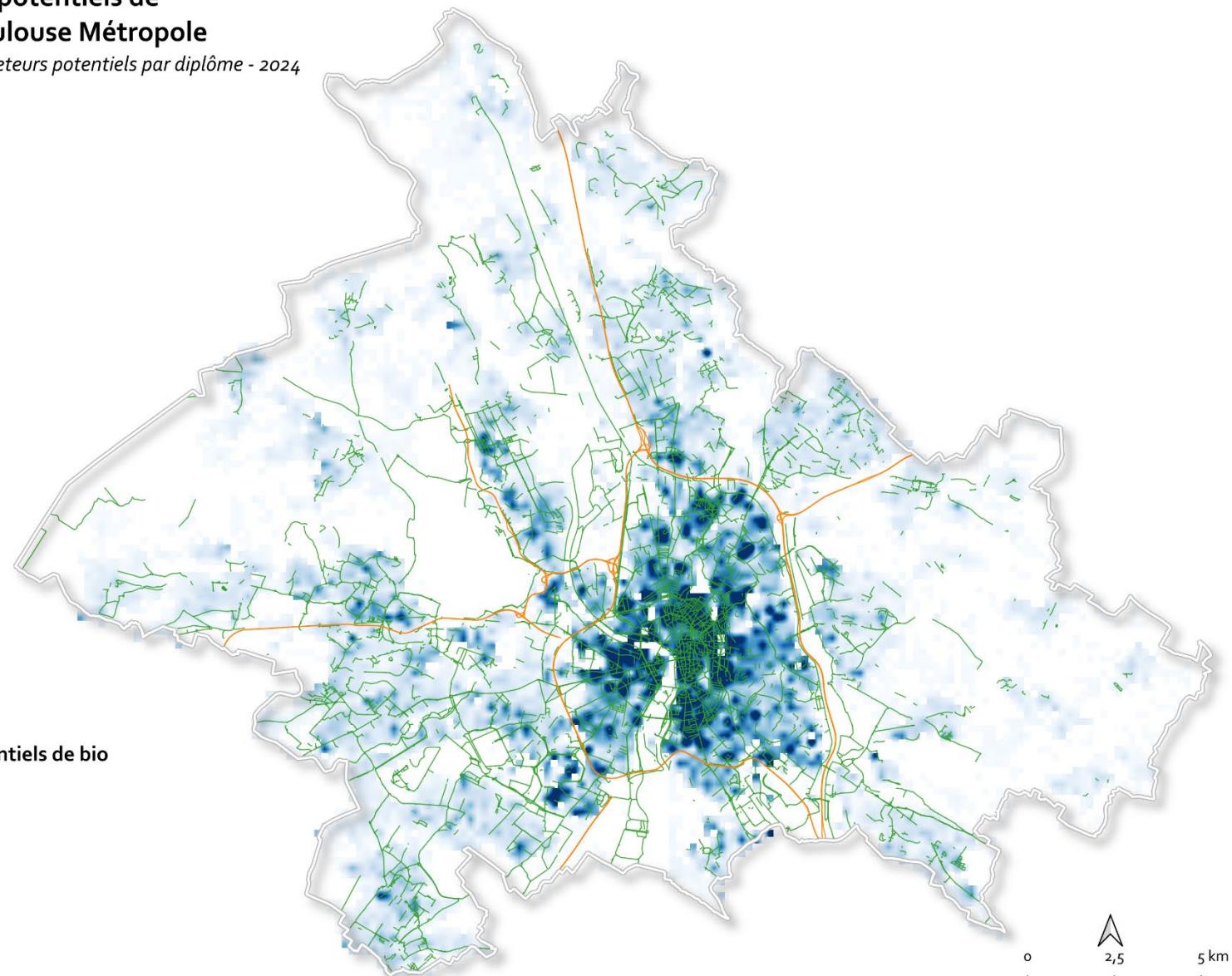
Consommateurs potentiels de bio par diplôme



Autoroute

Piste cyclable

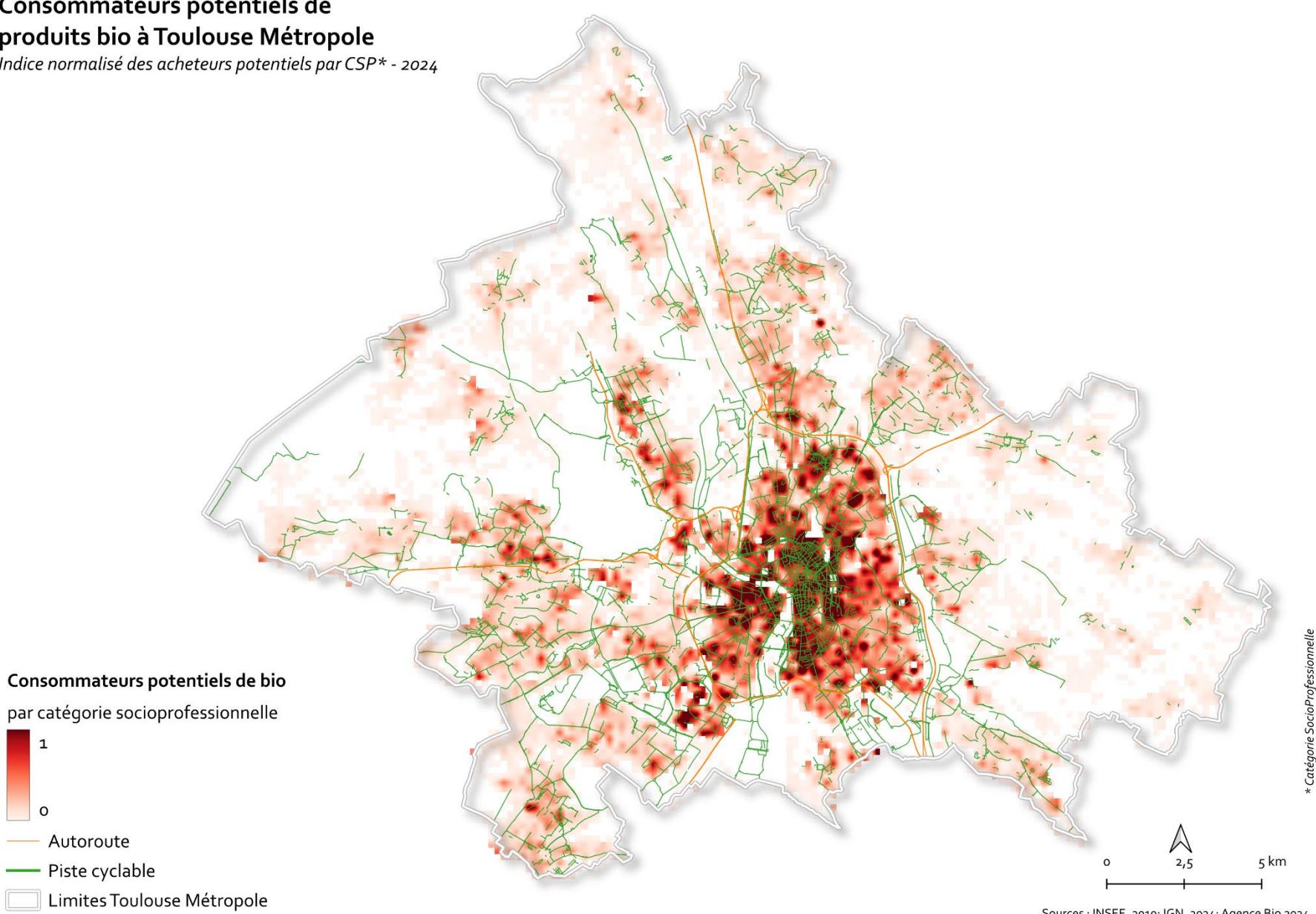
Limites Toulouse Métropole



Sources : INSEE, 2019; IGN, 2024; Agence Bio 2024

Consommateurs potentiels de produits bio à Toulouse Métropole

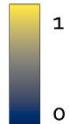
Indice normalisé des acheteurs potentiels par CSP* - 2024



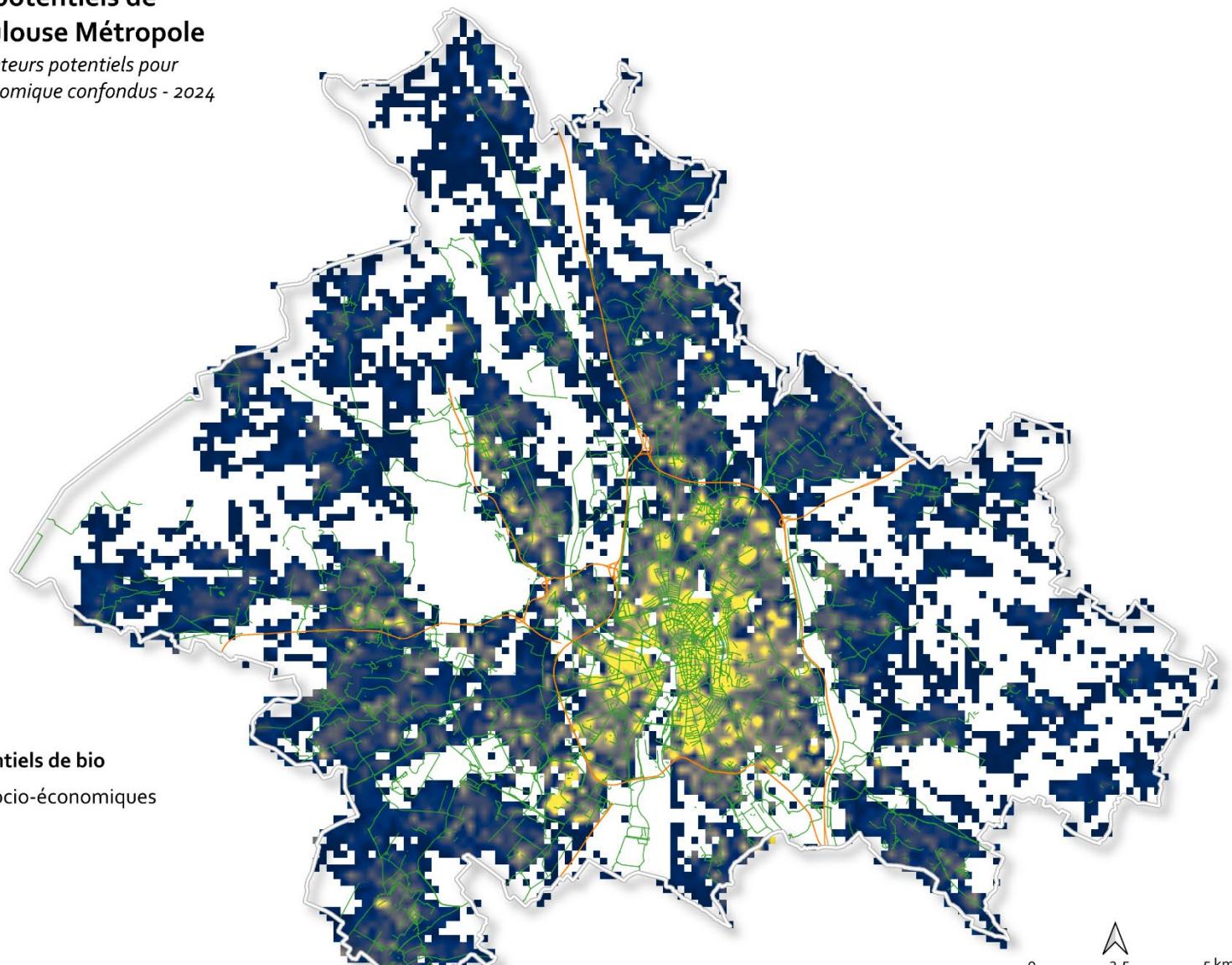
Consommateurs potentiels de produits bio à Toulouse Métropole

Indice normalisé des acheteurs potentiels pour les indicateurs socio-économique confondus - 2024

Consommateurs potentiels de bio
pour tous indicateurs socio-économiques



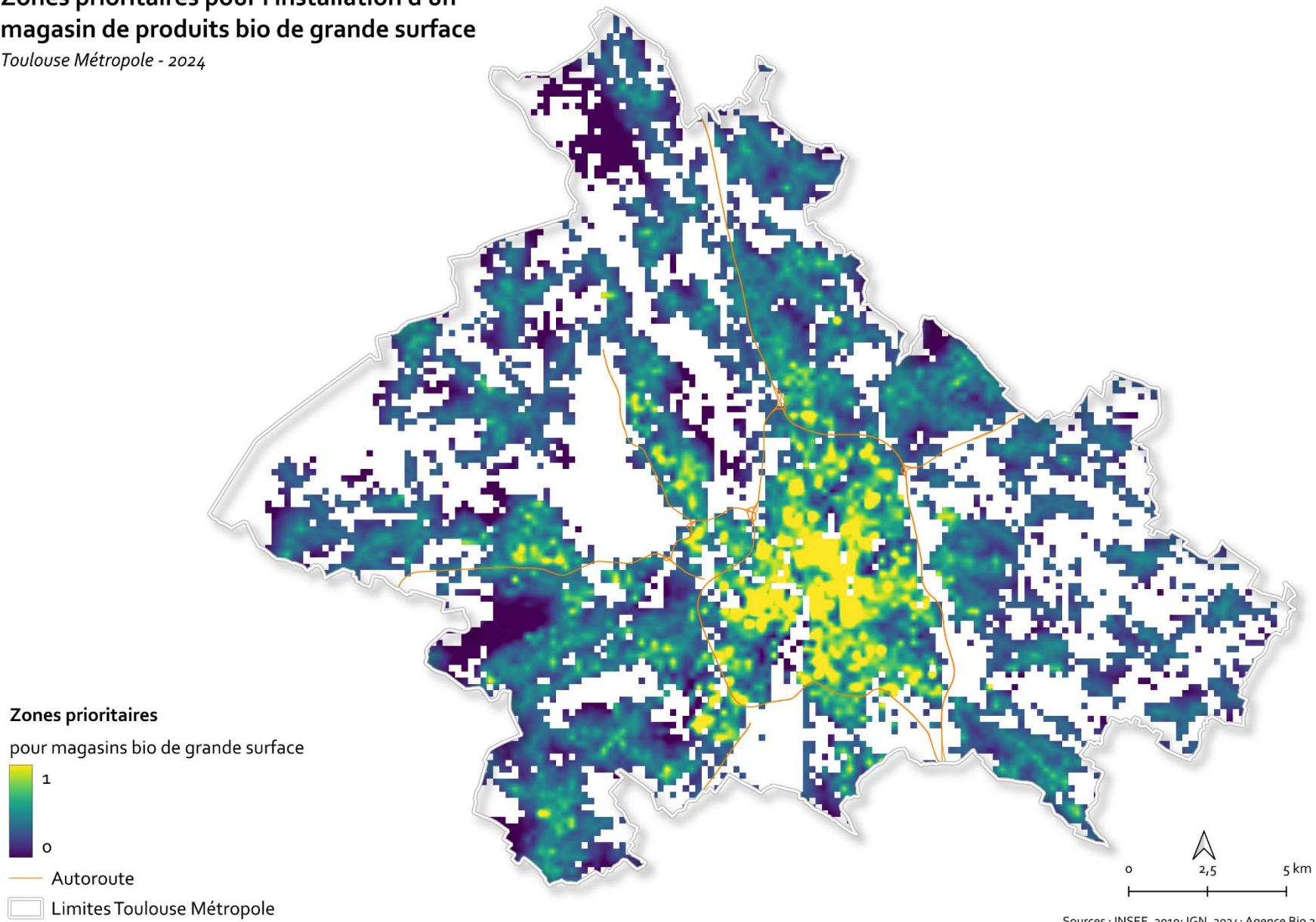
- Autoroute
- Piste cyclable
- Limites Toulouse Métropole



Sources : INSEE, 2019; IGN, 2024; Agence Bio 2024

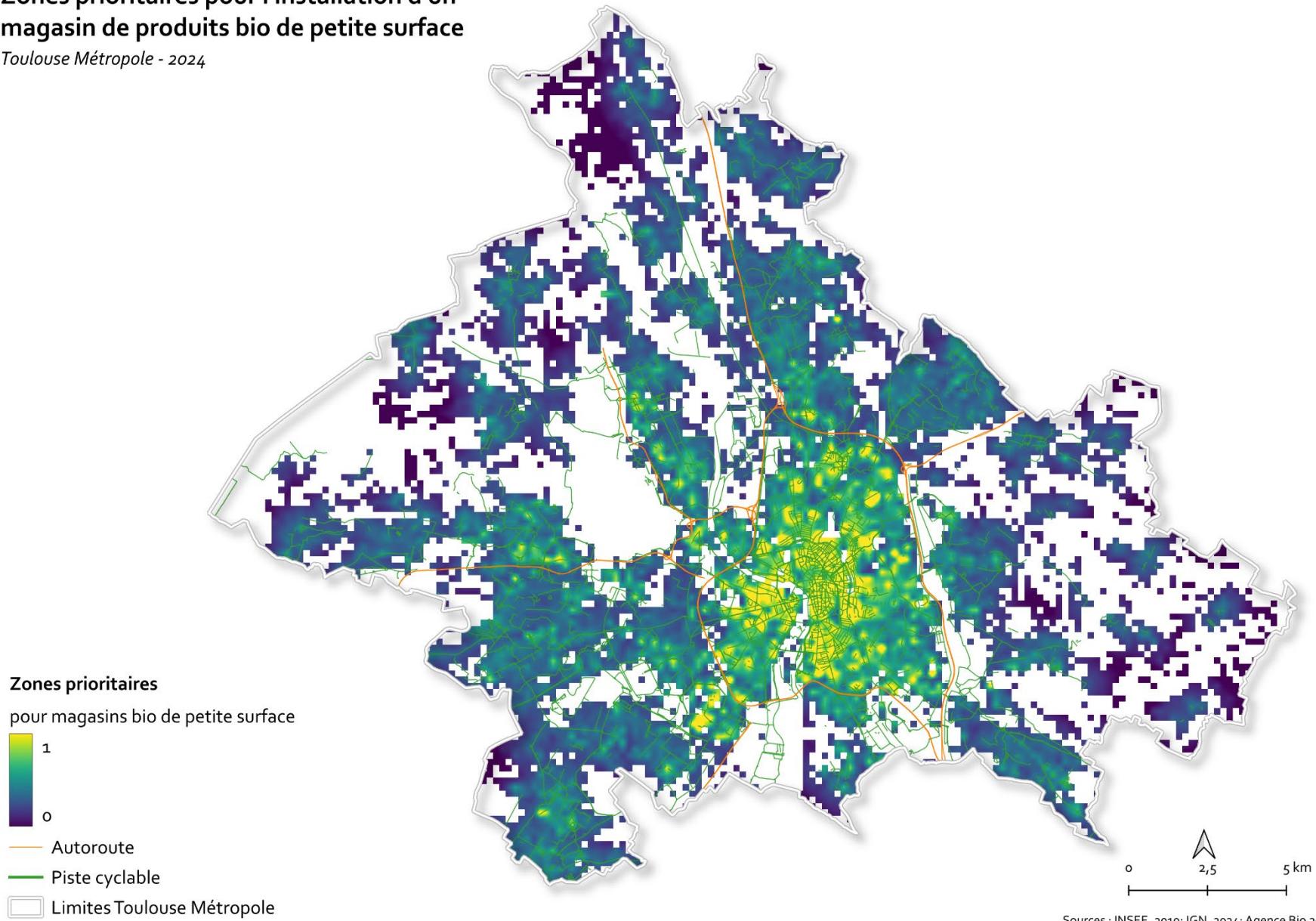
Zones prioritaires pour l'installation d'un magasin de produits bio de grande surface

Toulouse Métropole - 2024



Zones prioritaires pour l'installation d'un magasin de produits bio de petite surface

Toulouse Métropole - 2024



Comparaison entre les zones prioritaires identifiées et les magasins bio actuels

Toulouse Métropole - 2024

