

---

## RÉNOVATION DE L'ENQUÊTE TERUTI

Bertrand BALLET

SSP – Ministère de l'Agriculture

[bertrand.ballet@agriculture.gouv.fr](mailto:bertrand.ballet@agriculture.gouv.fr)

**Mots-clés** : Teruti, échantillonnage, allocation sous contrainte locale, données géolocalisées, photo-interprétation, télédétection

---

### Résumé

La couverture et l'utilisation des sols font l'objet d'une forte demande d'informations chiffrées en raison de la montée des préoccupations environnementales. En France, le dispositif législatif mis en place depuis 2010 témoigne de la volonté des pouvoirs publics de limiter la perte des terres agricoles, de réduire le rythme d'artificialisation des sols, de préserver la biodiversité. L'observatoire national des espaces agricoles naturels et forestiers est créé en 2014 et, en 2015, l'artificialisation des sols devient un des 10 indicateurs de richesse nationale.

L'enquête TERUTI est un dispositif statistique d'observation du sol parmi les plus anciens en Europe. Amorcée à la fin des années 1960, elle couvre la totalité du territoire métropolitain depuis 1982 et inclut 3 départements d'outre-mer depuis 2005. C'est une enquête aréolaire par sondage dont l'unité enquêtée est une portion de territoire, généralement une placette circulaire de 3 mètres de diamètre. La couverture et l'utilisation du sol y sont relevées chaque année par un enquêteur. Les surfaces des placettes ainsi qualifiées sont ensuite extrapolées de façon à représenter les 550 000 km<sup>2</sup> du territoire métropolitain.

En 2017, la méthodologie de l'enquête a été fortement renouvelée pour d'une part, prendre en compte l'existence de bases de données géolocalisées plus nombreuses et accessibles et, d'autre part, améliorer la précision des surfaces estimées, à moindre coût. La base de sondage de l'enquête TERUTI est désormais constituée d'une nouvelle grille de points compatible avec la directive Inspire. Chaque point de cette grille dense (1 point tous les 250 mètres) est qualifié par intersection géomatique avec les bases de données géographiques et administratives existantes. Les fichiers de l'IGN (BD TOPO, BD FORET) fournissent de l'information sur les plans et les cours d'eau, les infrastructures routières et ferroviaires, les bâtiments et les massifs forestiers. Le registre parcellaire graphique du Ministère de l'Agriculture géolocalise les parcelles agricoles éligibles aux aides de la PAC. Enfin, un plan de sondage stratifié est élaboré selon la probabilité d'observer à terme un changement d'occupation du sol. Les strates où la couverture est pérenne (eaux, infrastructures, bâti, cœur de forêt) ou connue de façon régulière (parcelles agricoles) ne sont pas enquêtées sur le terrain. A l'inverse, l'échantillon de points à collecter in-situ couvre davantage les anciennes terres agricoles, le littoral ou les zones périurbaines peuplées que les campagnes isolées. Certains points difficilement accessibles sont photo-interprétés en bureau à partir des photos aériennes ou satellites disponibles.

La présentation visera à présenter les principales innovations mises en œuvre dans la conception de la nouvelle enquête et les bénéfices qui en sont escomptés. Elle évoquera aussi les pistes d'évolution envisageables pour l'avenir, en lien avec la mise à disposition d'un large public de données satellitaires et les possibilités offertes par les techniques de télédétection et d'interprétation automatique des données satellites. En particulier, depuis 2017, une carte nationale d'occupation des sols (OSO) est produite automatiquement par le laboratoire Cesbio à partir des photographies du

satellite Sentinel-2 et une confrontation entre les observations de terrain de TERUTI et la carte OSO est prévue en 2018.

## **Abstract**

**TERUTI is a statistical area-frame survey on land use and cover among the oldest in Europe. Conducted since the end of 1960s, it covers all of the French metropolitan territory since 1982 and includes 3 French overseas departments since 2005. The statistical sampling unit is a portion of territory, generally a circular place of 3 meters diameter. Informations on land cover and land use are collected in-situ by a surveyor every year. The surfaces of the qualified places are then extrapolated so as to represent 550 000 km<sup>2</sup> of the metropolitan territory. In 2017, the statistical method of Teruti survey was strongly renovated taking into account the existence of numerous and accessible geo-located databases and trying to improve the accuracy of the estimated surfaces at a lower cost. A stratified sampling scheme has been implemented instead of the old two-stage sample. Some inaccessible places are also photo-interpreted at the office. In 2018, remote sensing data from Sentinel-2 will be compared to in-situ collected data from Teruti survey.**

L'enquête TERUTI est un dispositif statistique d'observation du sol parmi les plus anciens en Europe. Amorcée à la fin des années 1960, elle couvre la totalité du territoire métropolitain depuis 1982 et inclut trois départements d'outre-mer depuis 2005. C'est une enquête aréolaire par sondage dont l'unité enquêtée est une portion de territoire, généralement une placette circulaire de 3 mètres de diamètre. La couverture et l'utilisation du sol y sont relevées chaque année par un enquêteur. Les surfaces des placettes ainsi qualifiées sont ensuite extrapolées de façon à représenter les 550 000 km<sup>2</sup> du territoire métropolitain.

### **1. Bref historique des enquêtes Teruti**

La connaissance et le suivi de l'occupation du territoire sont des préoccupations anciennes de la statistique agricole. Les prémices de l'enquête annuelle «Utilisation du territoire» ont vu le jour en 1946 avec le contrôle des surfaces agricoles basé sur les plans cadastraux. Par la suite, les différents progrès techniques (l'introduction de la photographie aérienne dès 1962, puis à partir de 1969, l'association de relevés de terrain aux photographies aériennes) ont conduit à étendre l'enquête au suivi de l'ensemble du territoire mais avec une densité d'échantillonnage variable selon la diversité et l'intérêt «agricole» des régions étudiées (Fournier, 1972). A partir de 1982, la mise en place d'un échantillon national homogène et obligatoire a permis de stabiliser le système.

L'échantillon permanent a été totalement renouvelé entre 1990 et 1991, afin de corriger certains biais de représentativité introduits lors du tirage de l'échantillon de 1982. Enfin, en 2005, la numérisation des documents cartographiques, le géo-référencement des points et l'adaptation au projet européen Lucas ont donné naissance à un nouvel échantillon de points, avec une nouvelle nomenclature.

De 2005 à 2015, la dernière série des enquêtes Teruti, appelée Teruti-Lucas, a été menée tous les ans à l'exception de l'année 2011, en raison de la priorité accordée à la finalisation du Recensement Agricole 2010. De 2012 à 2015, les données administratives du Registre Parcellaire Graphique (RPG), mises à jour chaque année par les exploitants éligibles aux aides de la PAC, ont été mobilisées pour renseigner l'occupation des points Teruti qui étaient en prairies ou en cultures en 2010 et qui restaient situés dans des parcelles agricoles du RPG. L'occupation des points considérés comme stables à court terme (dunes, rochers, bâtiments, etc.) est aussi reconduite depuis lors. Ainsi, depuis 2012, le nombre de points à visiter se situe entre 50 % et 60 % de l'échantillon total de l'enquête.

L'enquête fournit une observation détaillée de l'occupation et de l'usage du sol à l'échelle nationale y compris DOM (sauf Guyane forestière et Mayotte), régionale et départementale. Depuis 2005, elle comporte une nomenclature détaillée d'occupation (couverture physique de la surface du point observé) de 4 niveaux et 122 postes, et une nomenclature d'usage ou utilisation (fonction socio-économique du point) de 2 niveaux et 38 postes. Bien que les nomenclatures se soient affinées dans la dernière série Teruti-Lucas, les trois séries de Teruti ont été raccordées sur des nomenclatures plus agrégées et permettent d'établir des séries longues disponibles depuis 1982. Enfin, au sein d'une même série, Teruti permet de réaliser des matrices de passage pour comprendre l'origine des changements d'occupation.

## **2. Evolution de la demande et des sources disponibles à partir de 2010**

A partir de 2010, le contexte de l'enquête Teruti-Lucas connaît deux évolutions principales. Il y a d'une part, une montée en charge des préoccupations environnementales, et en particulier, celles liées à l'occupation du sol. D'autre part, et parallèlement, se développent la production et la mise à disposition de bases de données géolocalisées d'origine statistique, administrative ou satellitaire.

### **2.1. La demande liée à la protection de l'environnement**

La couverture et l'utilisation des sols font l'objet d'une forte demande d'informations chiffrées en raison de la montée en charge des préoccupations environnementales. Depuis 2010, en France, la loi nationale pour l'environnement, celle de modernisation de l'agriculture ou les mesures de verdissement de la politique agricole commune témoignent de la volonté des pouvoirs publics de limiter la perte des terres agricoles, de réduire le rythme d'artificialisation des sols, de préserver la biodiversité. Ainsi, l'observatoire national des espaces agricoles naturels et forestiers est créé en 2014 et, en 2015, l'artificialisation des sols devient un des 10 indicateurs de la richesse nationale retenus par le gouvernement pour l'évaluation et la définition des politiques publiques.

Dans ce cadre, les résultats de l'enquête Teruti sont utilisés tant pour satisfaire aux obligations européennes et engagements internationaux de la France que pour répondre aux besoins d'utilisateurs nationaux. On peut notamment citer les utilisations suivantes :

- le rapportage annuel relatif à l'utilisation des terres, aux changements d'utilisation et à la forêt (UTCF) au titre de la Convention cadre des Nations-Unies sur le changement climatique (CCNUCC) et du protocole de Kyoto (inventaire des émissions de GES établi par le Citepa) ;
- l'inventaire mondial des ressources forestières de la FAO (Global Forest Resources Assessments – GFRA) et aux indicateurs paneuropéens de gestion durable des forêts (State of Europe's Forests - SoEF) ;
- l'indicateur de l'artificialisation des sols de la stratégie nationale de transition écologique vers un développement durable 2015 – 2020 (SNTEDD) et pour la loi sur les nouveaux indicateurs de richesse ;
- le « Panorama de la quantification de l'évolution nationale des sols agricoles » publié en 2014 pour l'Observatoire national de la consommation des espaces agricoles, l'avis du Conseil économique social et environnemental (Cese) de mai 2015 sur «La bonne gestion des sols agricoles : un enjeu de société» ou plus récemment la synthèse de l'expertise scientifique collective sur les « sols artificialisés et processus d'artificialisation des sols » réalisée par l'Institut national de la recherche agronomique (Inra) et l'Institut français des sciences et technologies des transports, de l'aménagement et des réseaux (Ifsttar) et publiée en décembre 2017.

## 2.2. Les autres sources de mesure des surfaces occupées

L'enquête Teruti est une source statistique annuelle permettant un suivi longitudinal de l'occupation et de l'usage du sol à petite échelle (région, département). Elle trouve sa place à côté d'autres sources statistiques qui fournissent également une mesure des surfaces occupées mais dont les objectifs sont soit dédiés à l'observation de territoires spécifiques (inventaire forestier national), soit à l'observation de périmètres géographiques plus vastes (enquête européenne Lucas) :

L'**inventaire forestier national**, conduit par l'Institut national de l'information géographique et forestière (IGN), vise prioritairement à quantifier les productions de bois à partir d'un échantillon de points représentant le territoire national. Compte tenu de ses objectifs, les nomenclatures d'occupation et d'usage du sol sont peu détaillées en dehors des surfaces boisées.

L'enquête aréolaire sur l'occupation et l'usage des sols (**Land Use and Cover Area frame statistical Survey – Lucas**), conduite au niveau communautaire par Eurostat, produit des données sur l'occupation et l'usage des sols au niveau européen mais sa précision est insuffisante pour des catégories d'occupation détaillées ou une utilisation au niveau infra-national (30 000 points enquêtés pour la France contre 300 000 dans Teruti-Lucas).

A côté de ces sources statistiques, des bases de données géolocalisées sont de plus en plus nombreuses. Ces bases sont d'origine administrative (fichiers fonciers fiscaux – **Majic**, registre parcellaire graphique – **RPG** – lié aux aides PAC, casier viticole – **CVI**), ou fondées sur l'interprétation humaine de photographies aériennes (référentiel IGN, **BD Forêt®**) ou d'images satellitaires (**Corine Land Cover** - CLC). Ces sources sont toutefois mieux adaptées à une connaissance et une analyse du territoire à plus grande échelle (commune) qu'à un suivi annuel des flux d'occupation et d'utilisation des sols à petite échelle (département, région).

En effet, les bases d'origine administrative ne décrivent avec précision qu'une partie des surfaces (parcelles agricoles des exploitations éligibles aux aides pour le RPG, surfaces viticoles pour le CVI, espaces bâtis pour les fichiers fonciers). Bien qu'elles soient mises à jour régulièrement, elles ne permettent pas de disposer d'une couverture exhaustive du territoire.

Les bases résultant de l'interprétation millésimée de photos aériennes ou d'images satellites comportent des informations très précises mais ne peuvent être aisément mobilisées pour disposer d'un suivi longitudinal de l'occupation et de l'usage du sol à l'échelle nationale et régionale.

Ainsi, la base de données géographique Corine Land Cover, disponible tous les 6 ans, retient comme unité d'observation des zones homogènes d'une taille minimum de 25 hectares. Elle décrit donc des dominantes d'occupations, alors que Teruti-Lucas cherche à quantifier plus précisément les types d'occupation.

Le référentiel à grand échelle (**RGE®**) de l'IGN propose, quant à lui, une couverture référente et cohérente du territoire national mais la base reste à vocation topographique. Ainsi, les couches numériques qui le compose ne couvrent pas tous les types d'occupations, notamment les différents types d'occupations agricoles, et ne sont pas toutes mises à jour annuellement.

La difficulté à mettre en cohérence les diverses sources disponibles a conduit à la mise en place en 2012 d'un groupe de travail national chargé d'élaborer une couche référentielle d'occupation du sol à grande échelle sur l'ensemble du territoire (**OCS-GE**). Ce projet vise une cohérence cartographique des bases locales pour une année donnée en s'appuyant sur des nomenclatures d'occupations et d'usages principaux conformes à la directive européenne Inspire. A ce jour, deux régions (Occitanie, Pays-de-la-Loire) ont réalisé une cartographie à grande échelle de ce type.

### 2.3. L'imagerie satellitaire

A partir de 2015, les produits issus du traitement automatique de l'imagerie satellitaire (**téledétection**) sont de plus en plus accessibles. Ainsi, le projet européen **Copernicus** monte en charge et prévoit la mise à disposition libre et gratuite des données issues de l'observation spatiale, notamment sur l'utilisation des sols à partir des satellites Sentinel-2, accompagnée d'actions de formation aux outils permettant l'exploitation de ces données. En France, le projet **Dinamis** (Dispositif Institutionnel National d'Approvisionnement Mutualisé en Imagerie Satellitaire) porté notamment par l'IGN, le Cnes et l'Irstea vise à développer pour 2018 un métaportail des dispositifs d'accès à l'imagerie satellitaire (en particulier à très haute résolution avec les satellites Pléiades et Spot) à destination de la communauté scientifique, des acteurs institutionnels et du secteur de la recherche -développement. Depuis 2016, le Centre d'études spatiales de la biosphère (CESBIO) produit une **carte annuelle d'occupation des sols (OSO)** couvrant le territoire de la France métropolitaine à partir de toutes les images prises par les satellites Sentinel-2A et 2B (entre 10 et 20 m de résolution selon les 13 bandes spectrales avec revisite tous les 5 jours).

### 3. Les objectifs de Teruti et de sa rénovation

Dans ce contexte évolutif, le Service de la Statistique et de la Prospective a réaffirmé l'intérêt de l'enquête Teruti comme outil majeur de suivi de l'évolution des territoires pour les pouvoirs publics ou les scientifiques. En 2016, il a souhaité en rénover le cadre méthodologique afin de mobiliser plus largement les données administratives et géographiques disponibles et permettre de réduire la charge budgétaire de l'enquête tout en améliorant le niveau des précisions sur les taux d'artificialisation et de déprise agricole, au niveau départemental.

Ainsi, l'enquête Teruti rénovée vise à répondre à deux enjeux déjà anciens :

- établir des séries annuelles historiques sur les statistiques d'occupation et d'usage des sols à petite et moyenne échelle (échelon national, régional et départemental) pour évaluer sur le moyen et long terme les politiques d'aménagement du territoire et notamment les politiques de préservation du milieu agricole et naturel. Il s'agit, en particulier, de mesurer :
  - le rythme du processus d'artificialisation des sols, le mode de consommation des terres agricoles et des espaces naturels...
  - la part d'imperméabilisation des sols, le type de bâti, sa densité, ses usages principaux et secondaires...
  - le développement des infrastructures agro-écologiques
- quantifier les principaux flux en réalisant des matrices de passage entre grandes catégories d'occupation et d'usage des sols entre 2 années d'une même série d'enquêtes .

A ces deux enjeux, sont assortis de deux objectifs complémentaires spécifiques à la nouvelle génération d'enquête :

- accroître les précisions sur les dynamiques de changements en concentrant davantage de points observés dans les zones à forte probabilité de changement d'occupation : îlots sortants du RPG et pourtour des parcelles déjà artificialisées, par exemple ;
- établir une synthèse statistique de référence en mobilisant davantage les sources administratives (RPG, Fichiers MAJIC, CVI) et les bases de données géographiques (BD TOPO®, BD FORET®, CLC) pour la stratification du territoire et le calage des résultats.

### 4. Description succincte de l'ancien plan de sondage Teruti-Lucas

Le plan de sondage de l'enquête Teruti-Lucas a été présenté lors des JMS 2012 (Amorich et al. ,2012). C'est un sondage aréolaire à deux niveaux de tirage, segments et points, ces derniers étant regroupés en grappes (segments) pour rentabiliser les déplacements.

#### 4.1. Les segments

Le segment est le premier niveau de tirage (Unité Primaire : UP). Il s'agit d'une portion de territoire dont la taille peut varier de 1,5 km sur 600 m (si 10 points ont enquêtés) à 1,5 km sur 1,5 km (si 25 points sont enquêtés). Le point est le second niveau de tirage (Unité Secondaire : US). Il s'agit également d'une portion de territoire, un cercle de 3 m de diamètre. Les points sont espacés de 300 m à l'intérieur d'un segment.

#### 4.2. Les points

Chaque segment contient 5 lignes de 5 points orientés comme les segments N-S, E-W et espacés régulièrement de 300 m. Les deux premières lignes (points 11 à 25) correspondent aux points des enquêtes européennes Lucas de 2001 et 2003, les trois suivantes ont été ajoutées pour répondre à d'éventuelles extensions. Tout segment enquêté a obligatoirement ses deux premières lignes enquêtées.

Figure 1 : les points composant un segment

○	○	○	○	○
11	12	13	14	15
○	○	○	○	○
21	22	23	24	25
●	●	●	●	●
31	32	33	34	35
●	●	●	●	●
41	42	43	44	45
●	●	●	●	●
51	52	53	54	55

#### 4.3. L'échantillon

Dans la plupart des départements, l'échantillon de segments correspond au doublement en quinconces du nombre de segments de la grille 6 km. Pour les départements les moins étendus (75, 90, 92, 93, 94, 971, 972, 974), l'échantillon de segments correspond à ceux de la grille de base (3 ou 2 km). Au total, l'échantillon annuel de Teruti-Lucas comprend 32 498 segments de 10 points (11 à 25).

Figure 2 : Positionnement des segments et des points de l'échantillon Teruti-Lucas pour une zone géographique donnée (département 66)





#### 4.4. Précision des estimations

L'intérêt du plan de sondage Teruti-Lucas à deux degrés (segments-points) est de limiter les coûts de déplacement des enquêteurs chargés de se rendre sur chaque point et de collecter les informations de couverture et d'usage du sol. L'échantillon de Teruti-Lucas est ainsi constitué de «grappes» le plus souvent éloignées d'au plus 6 km et regroupant 10 points espacés de 300 m.

En contrepartie, les surfaces estimées à partir de ce type d'échantillon sont moins précises par rapport à un échantillon de points issus d'un sondage aléatoire simple (Ardilly, 2006). Ainsi, le plan de sondage de Teruti-Lucas subit un effet de grappe défavorable qui limite son efficacité en termes de

précision. En effet, il conduit à recueillir de l'information sur des points géographiquement proches les uns des autres et donc à relever plusieurs fois la même couverture du sol dans un même segment (effet de grappe).

En 2007, une étude menée sur le plan de sondage de l'enquête européenne LUCAS conclut à une efficacité considérablement améliorée d'une méthode d'échantillonnage de points non groupés dans le cadre d'un sondage à deux phases avec post-stratification par rapport à des méthodes alternatives ou un plan de sondage par grappe (Gallego, 2007).

## **5. La nouvelle méthodologie**

Le changement méthodologique de l'enquête Teruti concerne essentiellement le plan de sondage. L'objectif est d'améliorer la précision des indicateurs-cibles de l'enquête en passant d'un plan de sondage à 2 degrés (celui de l'ancien dispositif Teruti-Lucas) à un plan de sondage en deux phases avec post-stratification.

Teruti reste une enquête aréolaire dont l'unité statistique enquêtée est une portion de territoire délimitée par un cercle de 3 m de diamètre, centrée sur un point géolocalisé (X,Y).

Le nouvel échantillon Teruti est issu d'un plan de sondage en deux phases avec post-stratification afin d'améliorer la précision des estimateurs par rapport à l'ancien plan de sondage à deux degrés. Ce plan de sondage est similaire à celui de l'enquête européenne LUCAS conçu et mis en œuvre par Eurostat depuis 2006.

La base de sondage Teruti est constituée d'une grille systématique de points couvrant l'ensemble du territoire. La nouvelle méthodologie fait largement appel aux bases de données géographiques, administratives et statistiques géo-localisables disponibles. Ces bases de données permettent d'imputer directement la couverture sur sol pour une partie des points Teruti et d'éviter ainsi le déplacement d'un enquêteur sur le terrain. Elles sont également mobilisées pour constituer des groupes de points homogènes (stratification) avant d'effectuer un tirage aléatoire parmi ceux nécessitant une collecte sur le terrain.

La taille de l'échantillon à enquêter est fixée par la contrainte budgétaire. L'allocation des points par département et strate est optimisée dans le but de maximiser la précision de l'indicateur-cible retenu (taux de déprise agricole) par département.

### **5.1. La base de sondage**

La base de sondage de l'enquête TERUTI est une grille de points espacés de 250 m dans le système de projection et de coordonnées géographiques de référence recommandé par la directive européenne INSPIRE (INfrastructure for SPatial Information in Europe) : ETRS 1989 LAEA (EPSG:3035). Chaque point de la grille est identifié par un code numérique unique indiquant le pas de la grille et les coordonnées du point (ex: «250mE1025N22000»).

La base de sondage TERUTI contient ainsi 8 780 567 points pour le territoire de France métropolitaine.

L'intérêt de tirer un échantillon dans une grille de points aussi dense permet, a priori, d'augmenter la précision sur les surfaces de chaque strate, et donc sur la précision de chaque type d'occupation, et en particulier, hors agriculture et hors forêt. L'objectif d'une telle densité réside également dans la richesse du système d'information ainsi créé, permettant de croiser facilement et à moindre coût des sources hétérogènes de données (sources géographiques ou administratives) pour décrire statistiquement le territoire à partir de ces sources.



## Qualification de la base de sondage

Chaque point de la base de sondage est qualifié en fonction des bases de données géographiques, administratives et statistiques disponibles.

Le tableau ci-dessous indique les principales caractéristiques des bases de données utilisées pour qualifier la base de sondage TERUTI :

*Tableau 1 : Bases de données géographiques, administratives et statistiques utilisées pour qualifier la grille de points (base) TERUTI*

Type	Source	Nom	Caractéristique d'occupation	Millésimes disponibles
Géographique (vectorisée)	IGN	BD TOPO	Surfaces en eaux	2013-2016 selon la date de prise de vue aérienne
Géographique (vectorisée)	IGN	BD TOPO	Infrastructures (réseau routier, voies ferrées,...)	2016 pour le réseau routier
Géographique (vectorisée)	IGN	BD TOPO	Bâtiments	2013-2016 selon la date de prise de vue aérienne
Géographique (vectorisée)	IGN	BD Forêt	Forêts	2009-2016
Géographique (raster)	IGN	BD Alti	Altitude	
Administrative (vectorisée)	ASP (PAC)	RPG (Registre Parcellaire Graphique)	Parcelles cultivées	2016
Administrative (géolocalisée*)	Douanes	CVI (Casier Viticole Informatisé)	Parcelles viticoles	2016
Statistique	Insee	Carreaux de population 1 km	Densité	2011
Statistique	CGET	Typologie des campagnes françaises	Dynamisme démographique et économique des communes	2011

\*CVI géolocalisé par rapprochement avec les parcelles vectorisées du cadastre (fichiers fonciers).

## 5.2. Stratification de la base de sondage

Le croisement de la grille de points avec ces différentes couches géographiques permet d'obtenir une première répartition en 13 classes.

Tableau 2 : Classification (qualification séquentielle) de la grille de points TERUTI à partir du croisement avec les bases de données disponibles

Types d'occupation	Nombre de points	%	Classe	Traitement dans l'enquête
Zones sous les eaux	139 742	1,6%	C1	imputée
Sols bâtis et infra structures (sols imperméabilisés)	273 437	3,1%	C2	imputée
Sols agricoles ou viticoles	4 435 144	50,5%	C3	imputée
Sols ex-agricoles (vraies sorties* du RPG entre 2010 et 2015)	66 254	0,8%	C4	enquêtée
Haies sorties du RPG	8 203	0,1%	C5	imputée
Cœur de forêt fermée	2 076 503	23,6%	C6	imputée
Orée de forêt fermée	131 378	1,5%	C7	enquêtée
Tâche urbaine dense	212 818	2,4%	C8	enquêtée
Ripisylve	31 402	0,4%	C9	enquêtée
Communes à dynamisme démo.-éco. <u>fort</u>	298 856	3,4%	C10	enquêtée
Communes à dynamisme démo.-éco. <u>moyen</u>	298 959	3,4%	C11	enquêtée
Communes à dynamisme démo.-éco. <u>faible</u>	436 675	5,0%	C12	enquêtée
Points en altitude, difficile d'accès	371 196	4,2%	C13	photo-interprétée
Ensemble France métr.	8 780 567	100%		

\*les vraies sorties du RPG ne sont pas liées à un redécoupage administratif des parcelles déclarées en culture.

Source : SSP, enquêtes Teruti

Ces 13 classes sont ensuite regroupées en 6 strates, dans un premier temps selon le mode de collecte (imputation / enquête terrain /photo-interprétation) puis, pour les classes «enquêtées» , selon la probabilité d'y observer un changement d'occupation du sol. Cette probabilité est estimée en fonction des flux observés entre les espaces artificiels, agricoles et naturels dans les enquêtes Teruti-Lucas entre 2010 et 2015, et par grands types d'occupation du sol.

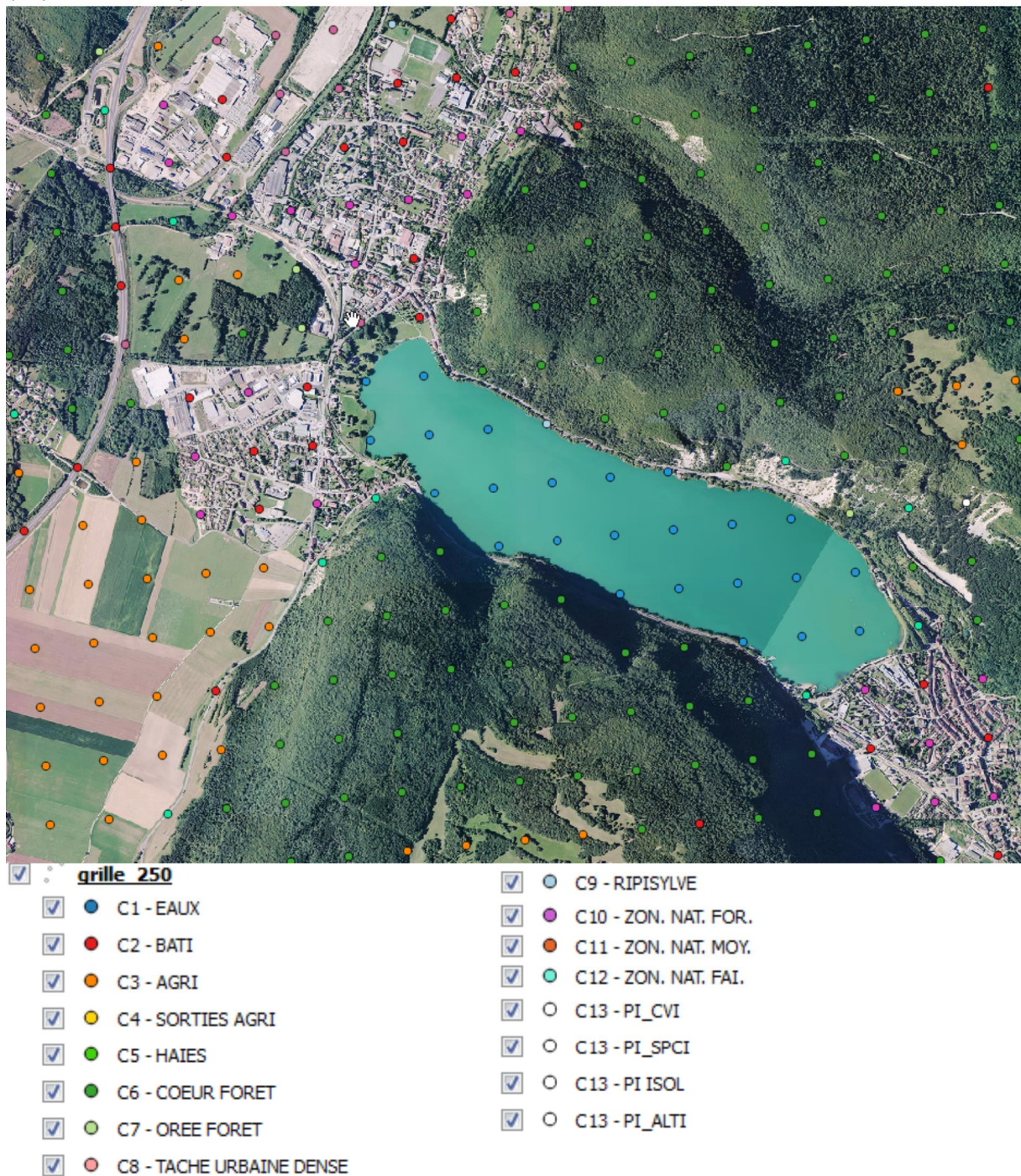
On retient ainsi une stratification en 6 strates (par département).

Tableau 3 : Poids des 6 strates de l'enquête TERUTI au niveau national

Strates	Classes	Nombre de points	%
S1 – Sorties du RPG	C4	66 254	0,8 %
S2 – Tache urbaine dense	C8	212 818	2,4 %
S3 – Communes dynamiques	C10	298 856	3,4 %
S45 – Autres espaces à proba . de changement moyen ou faible	C7, C9, C11, C12	898 414	10,2 %
S6 – Eaux, bâti, agricoles, forêt	C1, C2, C3, C5, C6	6 933 029	79,2 %
S10 – Photo-interprétation	C13	371 196	4,2 %
France métropolitaine		8 780 567	100 %

Source : SSP, enquêtes Teruti

Figure 3 : grille de points (base de sondage) Teruti classifiée pour une zone géographique donnée (département 01).



### 5.3. Taille de l'échantillon et précisions attendues

#### 5.3.1. Contrainte budgétaire et démarche adoptée

Le budget annuel de l'enquête Teruti est limité à 1 million d'euros par an contre 1,2 million pour Teruti-Lucas. Des tests sur le terrain ont permis d'estimer à environ 14 euros le coût de collecte d'un point de la nouvelle enquête Teruti. La taille globale de l'échantillon annuel enquêté France entière est d'environ 71 000 points dont 67 000 pour la métropole.

L'allocation de cet échantillon par département et par strate est recherchée selon la démarche suivante :

- définition d'un indicateur-cible pour lequel une certaine précision est attendue au niveau départemental ;
- calcul de cet indicateur et de sa dispersion au niveau départemental et par strate à partir des enquêtes Teruti-Lucas 2010 et 2015 ;
- allocation de Neyman sous contrainte de précision locale (Koubi, Mathern, 2009)

### 5.3.2. Définition et calcul d'un indicateur-cible

Les gains de précision attendus de la nouvelle enquête Teruti concernent les indicateurs de changement d'occupation du sol entre deux années (enquêtes) différentes. Plusieurs indicateurs de changement entre les grands types de sols (artificiels, agricoles et naturels) ont été calculés à partir des enquêtes Teruti-Lucas 2010 et 2015 :

- taux d'artificialisation globale : nombre de points qui étaient non artificiels en 2010 et qui le sont devenus en 2015 sur le nombre total de points enquêtés ;
- taux de désartificialisation globale : nombre de points qui étaient artificiels en 2010 et qui ne le sont plus en 2015 sur le nombre total de points enquêtés ;
- taux de changement artificiel : nombre de points qui étaient non artificiels en 2010 et qui le sont devenus en 2015 + nombre de points qui étaient artificiels en 2010 et qui ne le sont plus en 2015 sur le nombre total de points enquêtés. C'est donc la somme des deux taux précédents ;
- taux de perte agricole : nombre de points qui étaient agricoles en 2010 et qui ne le sont plus en 2015 sur le nombre total de points enquêtés ;
- taux de gain agricole : nombre de points qui n'étaient pas agricoles en 2010 et qui le sont devenus en 2015 sur le nombre total de points enquêtés ;
- taux de changement agricole : nombre de points qui étaient agricoles en 2010 et qui ne le sont plus en 2015 + nombre de points qui n'étaient pas agricoles en 2010 et qui le sont devenus en 2015 sur le nombre total de points enquêtés. C'est donc la somme des deux taux précédents.

La précision souhaitée pour ces indicateurs est celle qui permet d'affirmer qu'un changement d'occupation du sol est significatif au niveau départemental. Cette précision correspond à un rayon de l'intervalle de confiance autour de l'indicateur suffisamment petit pour exclure significativement une valeur nulle de cet indicateur. Or, ramenée en moyenne annuelle, la valeur de ces indicateurs est très faible. La recherche de précision sur ces indicateurs conduit à des tailles d'échantillon très supérieures à la contrainte budgétaire. Ainsi, il faudrait enquêter près de 250 000 points pour estimer un taux d'artificialisation annuel significativement non nul dans tous les départements de France métropolitaine.

*Tableau 4 : taux annuels de perte agricole et d'artificialisation dans les régions de métropole entre 2010 et 2015*

	Perte agricole	Artificialisation
Ile-de-France	0,25%	0,32%
Champagne-Ardenne	0,18%	0,20%
Picardie	0,15%	0,13%

Haute-Normandie	0,24%	0,20%
Centre	0,26%	0,19%
Basse-Normandie	0,35%	0,21%
Bourgogne	0,16%	0,12%
Nord-Pas-de-Calais	0,38%	0,36%
Lorraine	0,12%	0,14%
Alsace	0,19%	0,19%
Franche-Comté	0,20%	0,18%
Pays de la Loire	0,34%	0,26%
Bretagne	0,31%	0,26%
Poitou-Charentes	0,18%	0,14%
Aquitaine	0,34%	0,24%
Midi-Pyrénées	0,25%	0,14%
Limousin	0,19%	0,10%
Rhône-Alpes	0,33%	0,20%
Auvergne	0,29%	0,15%
Languedoc-Roussillon	0,48%	0,20%
Provence-Alpes-Côte d'Azur	0,48%	0,39%
Corse	0,50%	0,13%
<b>France métropolitaine</b>	<b>0,28%</b>	<b>0,20%</b>

Source : SSP, enquêtes Teruti-Lucas 2010-2015

Compte tenu de la contrainte budgétaire annuelle et des niveaux de précision recherchés, il a été décidé :

1 - de cumuler 2 voire 3 années d'enquête-terrain avant de revenir sur les mêmes points et pouvoir produire des estimateurs de changement d'occupation du sol statistiquement significatifs au niveau départemental ;

2 – de retenir comme indicateur cible le **taux annuel de perte agricole** par département. Il correspond aux préoccupations du Ministère de l'Agriculture et il est également corrélé aux indicateurs d'artificialisation.

### 5.3.3. Recherche d'une allocation optimale de l'échantillon de l'enquête-terrain

La méthode d'allocation optimale sous contrainte de précision locale a été mise en œuvre pour répartir l'échantillon de points Teruti entre les départements. Cette méthode, développée par Malik Koubi et Sandrine Mathern pour l'enquête du Ministère du Travail sur l'activité et les conditions de main d'œuvre (Koubi, Mathern, 2009), améliore l'allocation de Neyman en assurant une précision « locale » minimale dans les regroupements de publication, sans trop détériorer sa précision « globale » au niveau de l'ensemble de la population.

Pour Teruti, l'allocation classique de Neyman a été effectuée sur la dispersion du taux annuel de perte agricole calculée par département et strate, selon les formules suivantes :

$N_D^h$  nombre de points de la base de sondage (grille) dans le département D et la strate h

$P_D^h$  probabilité de changement observée dans la strate h du département D (indicateur-cible) selon les résultats des enquêtes Teruti-Lucas entre 2010 et 2015

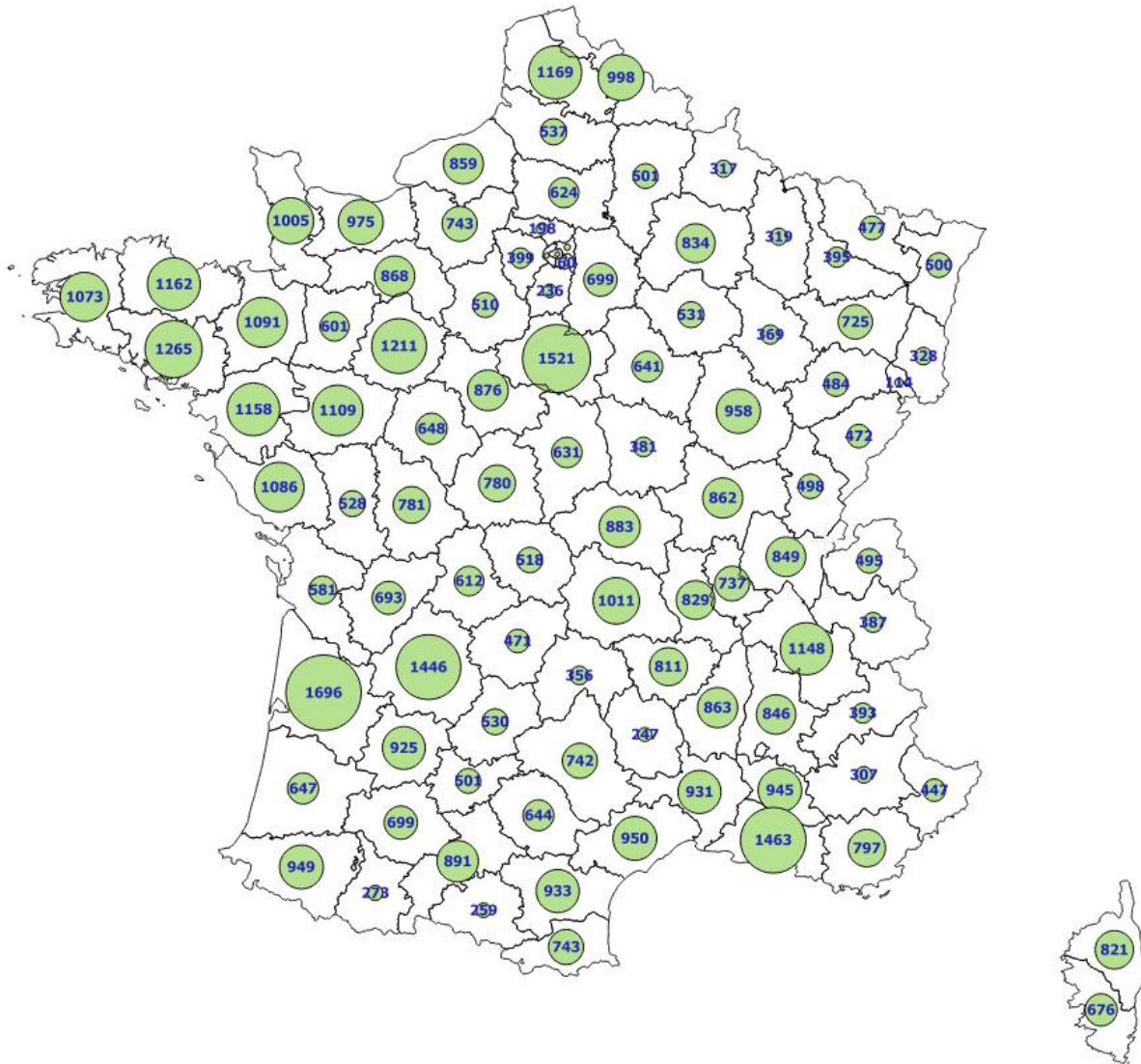
$S_{Dh}^2$  dispersion corrigée de la probabilité de changement observée dans la strate h du département D



$$S_{Dh}^2 = \frac{(N_D^h)}{(N_D^h - 1)} x (P_D^h)(1 - P_D^h)$$

Dans le cadre de l'allocation de Neyman, les contraintes de précision locale correspondent à une précision absolue de  $\pm 750$  ha autour de l'estimateur de perte annuelle de surface agricole dans chaque département.

Figure 4 : allocation optimale de Neyman de l'échantillon d'enquête sur le terrain sous contrainte de précision départementale des pertes de surfaces agricoles



Sous l'hypothèse que la consommation des surfaces agricoles observée entre 2010 et 2015 se poursuive au même rythme entre 2015 et 2020, cette allocation par strate et département des 67 000 points observés sur le terrain permet de mesurer des pertes de surfaces agricoles significativement non nulles dans la plupart des départements .

Tableau 5 : nombre de départements (sur 96) pour lesquels l'indicateur estimé est significativement non nulle selon le nombre de points observés sur le terrain

Nombre d'enquêtes annuelles réalisées	Cumul de points enquêtés	Taux annuel de perte de terres agricoles	Taux annuel d'artificialisation
1	67 000	92	80
2	134 000	95	90
3	201 000	96	93*

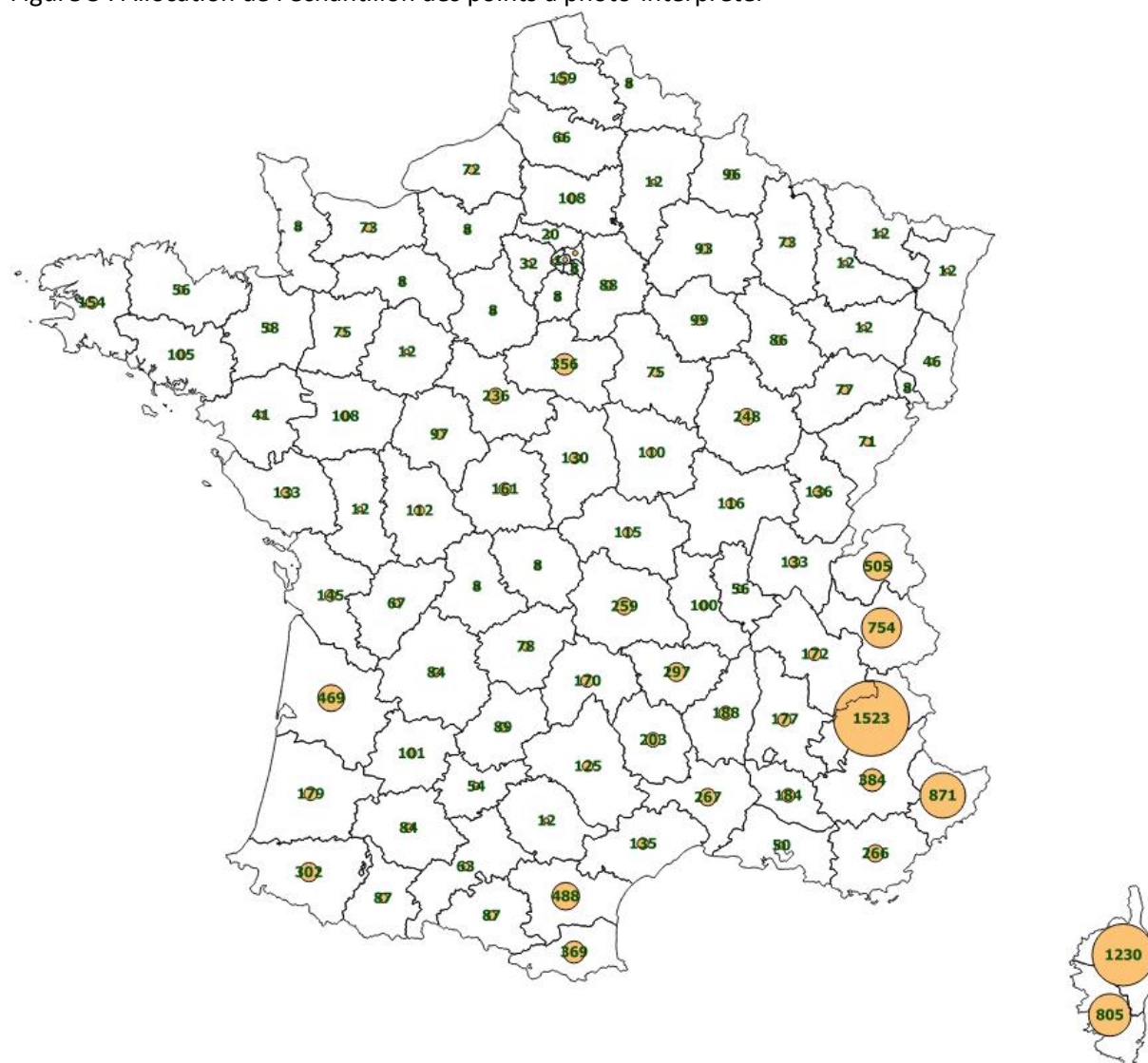
Source : SSP, enquêtes Teruti

\*avec un échantillon de 201 000 points enquêtés sur le terrain, le taux d'artificialisation estimé reste significativement non différent de 0 dans les départements 65, 93 et 94

### 5.3.4. Allocation optimal de l'échantillon de points à photo-interpréter

Les moyens humains disponibles pour réaliser la photo-interprétation en bureau des points en altitude (> 1500 m) ou difficiles d'accès limite la taille de l'échantillon à environ 15 000 points sur les 371 196 de la strate S10-Photo-interprétation.

Figure 5 : Allocation de l'échantillon des points à photo-interpréter



L'opération de photo-interprétation est « mutualisée » sur l'ensemble des 13 services régionaux du Ministère de l'Agriculture (Srise), quelle que soit la localisation du point à photo-interpréter.

### 5.3.5. Calcul de précision a posteriori

Le plan de sondage retenu pour Teruti est celui d'un tirage en deux phases avec post-stratification. Dans la 1<sup>ère</sup> phase, un point est tiré de manière systématique tous les 250 mètres : c'est l'application de la grille de points dans le système de projection et de coordonnées géographiques de référence recommandé par la directive européenne INSPIRE. Un point (cercle de 3m de diamètre) ayant une superficie théorique de 7,06 mètres carrés, le nombre total de points constituant l'univers de tirage en 1<sup>ère</sup> phase est obtenu en divisant la superficie du territoire national (550 000 km<sup>2</sup>) par 7,06 mètres carrés. L'échantillon de points issus de la 1<sup>ère</sup> phase (grille de 8,8 millions points) représente ainsi 1 / 8 859 de l'ensemble des points de l'univers. Tous les points échantillonnés en 1<sup>ère</sup> phase sont qualifiés et affectés à une strate (post-stratification). En 2<sup>e</sup> phase, l'échantillon est tiré par département et par strate selon une allocation de Neyman sous contrainte locale.

Les formules de calcul qui s'appliquent pour évaluer la précision du sondage en deux phases sont les suivantes (Cochran, 1977) :

#### Notations utilisées :

$N$  Nombre total de points dans l'univers

$H$  strates indicées par  $h$

$n'$ ,  $n'_h$  le nombre d'unités échantillonnées en première phase (total et par strate)

$n$ ,  $n_h$  le nombre d'unités échantillonnées en deuxième phase (total et par strate)

$Y_{ih}$  la présence de l'occupation considérée dans l'unité  $i$  de la strate  $h$

$$w_h = n'_h / n'$$

L'estimateur de la moyenne est donc dans ces conditions :

$$\bar{Y} = \sum_{h=1}^H \left( w_h \sum_{i=1}^{n_h} \frac{Y_{ih}}{n_h} \right) = \sum_{h=1}^H w_h \bar{y}_h$$

L'estimateur classique de variance pour l'estimation en deux phases est :

$$V\left(\frac{ss}{Y}\right) = \sum_{h=1}^H \frac{w_h^2 s_h^2}{n_h} + \frac{1}{n'} \sum_{h=1}^H w_h (\bar{y}_h - \bar{y})^2$$

avec  $s_h^2$  désignant l'estimateur de la dispersion de la variable  $Y$  et en négligeant le terme en  $\frac{1}{N}$

dans la mesure où  $N$  est très grand.

Dans le cadre d'un tirage systématique il est préférable d'utiliser l'estimateur suivant concernant la dispersion

$$s_h^2 = \left( 1 - \frac{n_h}{n'_h} \right) \frac{\sum_{i \neq j} (Y_i - Y_j)^2}{n_h (n_h - 1)}$$

Dans l'expression de la variance ci-dessus, le second terme en  $1/n'$  représente la perte de précision liée au tirage en deux phases par rapport à un plan de sondage stratifié à une seule phase (Ardilly, 1994). Dans le cas de Teruti, ce terme peut être négligé compte tenu de la taille de l'échantillon de 1<sup>ère</sup> phase ( $n'$ ) constitué de la grille des 8,8 millions points espacés de 250 m. Les calculs réalisés en appliquant les formules de précision d'un sondage en deux phases et celles d'un sondage en une phase ont montré un écart très faible (inférieur à 1 %) au niveau départemental.

## **6. Évolutions liées à la télédétection**

Dans la période actuelle, la disponibilité accrue de l'imagerie satellitaire et des outils d'exploitation de ces images (Copernicus, Dinamis, CESBIO, etc.) laisse envisager plusieurs évolutions possibles du dispositif Teruti. Il convient de rappeler que les techniques de télédétection et d'exploitation d'images satellites s'appuient sur des méthodes d'apprentissage et des stratégies de classification qui nécessitent de disposer de données de références récentes et de qualité. De ce point de vue, les données relevées sur le terrain par l'enquête Teruti sont particulièrement intéressantes pour valider les modèles de télédétection de l'occupation des sols. En retour, Teruti pourra bénéficier de la télédétection pour cibler davantage les déplacements d'enquêteur sur le terrain vers des points du territoire ayant probablement subi un changement d'occupation du sol. La télédétection d'images satellites pourrait aussi permettre d'imputer l'occupation du sol sur des points difficiles d'accès. Pour s'approprier ces données issues de la télédétection, une première expérimentation est menée en 2018 visant à confronter les occupation du sols relevées par les enquêteurs Teruti avec celles de la couche d'occupation du sol OSO produite par le CESBIO à partir des images des satellites Sentinel2.

### **6.1. Le Centre d'études spatiales de la biosphère (CESBIO)**

Le Centre d'études spatiales de la biosphère (CESBIO) est un laboratoire de recherche dédié au développement des connaissances sur le fonctionnement et la dynamique de la Biosphère continentale à différentes échelles spatiales et temporelles. Cette unité conduit des recherches dans le domaine de l'observation et de la modélisation des surfaces continentales. Elle participe à la définition de missions spatiales et au traitement des données de télédétection et développe des méthodes d'analyse et des modélisations. Celles-ci peuvent être industrialisées et mises en application de façon opérationnelle. Le CESBIO est unité de recherche mixte entre l'Université Paul Sabatier (Toulouse III), le Centre National de la Recherche Scientifique (CNRS), le Centre National d'Études Spatiales (CNES) et l'Institut de Recherche pour le Développement (IRD).

Depuis 2016, le CESBIO produit une carte d'occupation des sols de la France métropolitaine à partir des images des satellites Sentinel-2 (2A puis 2B à partir de 2017).

### **6.2. La mission Sentinel-2**

La mission Sentinel-2 fait partie du programme d'observation et de surveillance de la Terre, Copernicus, conduit par l'Union Européenne. L'Agence spatiale européenne (ESA) est responsable du développement des deux satellites, de leurs instruments et du segment sol.

Le satellite d'observation européen Sentinel-2A a été mis en orbite le 23 juin 2015. Son jumeau, Sentinel-2B, l'a rejoint le 6 mars 2017 sur la même orbite à 786 km d'altitude. A eux deux, ils fournissent tous les 5 jours des images d'une largeur au sol de 290 km et d'une résolution de 10 à 60 m selon 13 bandes spectrales allant du visible au moyen infrarouge.

Suite à un accord de coopération, le CNES a développé le système de traitement des données : qualité d'image, détection des nuages et corrections atmosphériques. Le CNES met à disposition de

manière libre et gratuite les données des Sentinel-2 sur sa plate-forme Internet PEPS, et fournit des produits corrigés des effets atmosphériques par l'intermédiaire du pôle thématique THEIA.

Ces images servent un large éventail d'applications : suivi du développement des cultures et forêts, de l'occupation des sols, des pratiques et productions agricoles, des littoraux,...

### 6.3. La couche OSO et Teruti

Le CESBIO produit une couche d'occupation des sols (OSO) millésimée à partir de l'ensemble des images Sentinel-2 prises pendant une année. Le centre a développé en logiciel libre une chaîne de traitement automatique (iota-2) réalisant la classification supervisée des téra-octets de pixels des images Sentinel. La chaîne iota-2 met en œuvre des méthodes d'apprentissage profond et s'appuie notamment sur une stratification spatiale selon les régions éco-climatiques comme stratégie de classification. Les données de référence utilisées pour l'apprentissage et la validation sont principalement la BD TOPO de l'IGN, Corine Land Cover et le Registre Parcellaire Graphique.

La couche d'occupation de sols produite à partir des pixels de 10 m des images Sentinel 2 contient 17 postes de nomenclature : 11-culture été, 12-culture hiver, 31-forêt feuillus, 32-forêt conifères, 34-pelouses, 36-landes ligneuses, 41-urbain dense, 42-urbain diffus, 43-zones industrielles et commerciales, 44-surfaces routes, 45-surfaces minérales, 46-plages et dunes, 51-eau, 53-glaciers ou neige, 211-prairies, 221-vergers, 222-vignes.

La matrice de confusion entre occupation prédite et occupation réelle indique que 88 % des pixels de validation sont bien classés (ou plutôt, classés comme la donnée de référence, qui elle-même peut contenir des erreurs). Le CESBIO indique qu'avec un tel taux d'erreur, il n'est pas conseillé de détecter des évolutions de l'occupation des sols au niveau des pixels, à moins de regrouper les classes qui sont souvent confondues entre elles (les surfaces artificialisées, par exemple).

*Figure 6 : regroupement des classes préconisé par le CESBIO selon la qualité de la classification OSO*

$$Rappel_i = \frac{\text{nombre de pixels correctement attribués à la classe } i}{\text{nombre de pixels appartenant à la classe } i}$$

$$Precision_i = \frac{\text{nombre de pixels correctement attribués à la classe } i}{\text{nombre de pixels attribués à la classe } i}$$

$$F\text{-score} = 2 \cdot \frac{Precision \times Rappel}{Precision + Rappel}$$



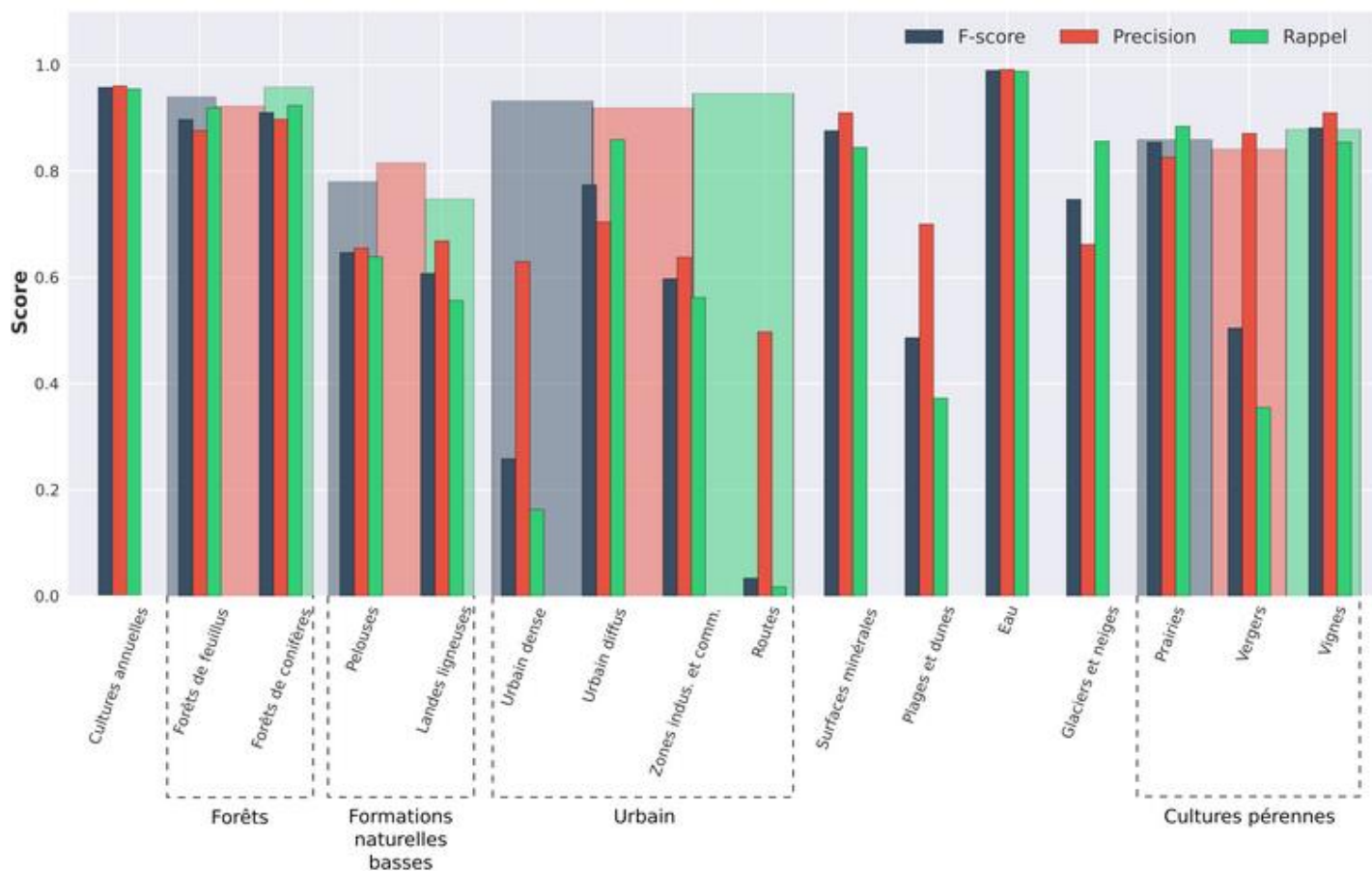


Figure 7 : grille de cohérence entre les nomenclatures OSO et Teruti

	COHERENCE	INCERTAIN	ANOMALIE	Nomenclature agrégée OSO		
Nomenclature agrégée Teruti	Cultures annuelles	Forêts	Formations naturelles basses	Tâche urbaine	Surfaces naturelles nues	Cultures permanentes
Surface naturelle sans végétation	ANOMALIE	ANOMALIE	ANOMALIE	ANOMALIE	COHERENCE	ANOMALIE
Surface artificialisée	ANOMALIE	ANOMALIE	ANOMALIE	COHERENCE	ANOMALIE	ANOMALIE
Surface cultivée	COHERENCE	ANOMALIE	ANOMALIE	ANOMALIE	ANOMALIE	COHERENCE
Formation herbeuse	COHERENCE	ANOMALIE	COHERENCE	INCERTAIN	ANOMALIE	COHERENCE
Formation arbustive	ANOMALIE	ANOMALIE	COHERENCE	INCERTAIN	ANOMALIE	ANOMALIE
Formation arborée	ANOMALIE	COHERENCE	ANOMALIE	ANOMALIE	ANOMALIE	ANOMALIE

En 2018, un rapprochement de la collecte Teruti avec la couche OSO aura lieu, dans le cadre de la validation de l'enquête Teruti 2018. Une vérification de la cohérence entre l'occupation relevée sur le terrain par les enquêteurs Teruti et la couche OSO 2017 sera organisée. Ce contrôle interviendra lorsque les 2 modes d'occupation du sol divergeront au niveau agrégé des nomenclatures de couverture du sol Teruti et OSO, selon le principe général suivant :

L'occupation «urbaine» de la couche OSO correspond à la notion de «tâche urbaine» qui peut recouvrir plusieurs occupations ponctuelles différentes comme le montre l'image ci-dessous. Ce type d'occupation OSO ne peut pas être confronté directement avec l'occupation «artificialisée» des points Teruti.

Figure 8 : la codification OSO en bâti dense recouvre des points qui seraient classés différemment dans Teruti



A l'avenir, la couche OSO pourrait être utilisée comme source d'imputation des points Teruti pour les classes d'occupation du sol jugées robustes, en fonction du résultat des contrôles Teruti-OSO effectués en 2018. La couche OSO devrait aussi intervenir en amont de l'enquête, comme source de stratification de la base de sondage Teruti. Elle permettrait notamment de cibler efficacement le retour des enquêteurs sur des points enquêtés 2 ou 3 ans auparavant et susceptibles d'avoir changé d'occupation depuis la visite précédente.

## Bibliographie

- [1] Fournier P., « Etude sur l'utilisation du territoire – Méthodologie – Résultats 1969 – 1970 – 1971 », supplément « Série Etudes », n° 104, novembre 1972
- [2] Amorich S., Michel P., Mirouze B, Mary A., « Les développements méthodologiques réalisés dans le cadre de l'enquête Teruti-Lucas », Journées de méthodologie statistique, 2012
- [3] Ardilly P., « Les techniques de sondage », éditions Technip, 2006
- [4] Gallego J. , « Sampling efficiency of the EU point survey LUCAS 2006 ». Proceedings of the 56th ISI session. Lisbon, 22-29 september 2007
- [5] Koubi M., Mathern S., « La nouvelle méthode d'échantillonnage de l'enquête trimestrielle Acemo depuis 2006, amélioration de l'allocation de Neyman », Dares, Documents d'études n°146, février 2009
- [6] Cochran W., « Sampling Techniques », 1977