

Réunion du 16 septembre 2014

Présents

Benoit GOURGAND	(Cerema/DTTV)
Thierry SAFFROY	(IGN)
Jacques BOUFFIER	(Cerema/PCI AST)
Emmanuel MAUGEAIS	(IGN)
Sylvie GRAS	(IGN)
Hélène DURAND	(Alisé géomatique)
Laure WATEAU	(Laure Wateau)
Sophie FOULARD	(IAU IDF)
Thierry TOUZET	(IGN)
Julien RAVENEL	(AUCAME)
Pascal BARILLÉ	(Pays d'Auray)
Sylvan COSTE	(SIG L-R)
Jacques AUTRAN	(ENSA Marseille/CRIGE PACA)
Claire AJOUC	(CRIGE PACA)
Marianne GÉRARD	(Auditrice du mastère spécialisé SILAT)
Florence DÉCAUDIN	(CR Nord Pas-de-Calais)
Dominique MESTRESSAT -CASSOU	(ADU Lille Métropole)
Nadège CHAMBON	(CCI France)

Ordre du jour

- Points d'actualité
- Contrôles qualités
- modèle de données

Points d'actualité

L'appel à commentaire concernant les prescriptions nationales est toujours en cours à l'adresse suivante :

http://cnig.gouv.fr/?page_id=8069

il sera clôturé le 1^{er} octobre. Ne pas hésiter à relayer l'information

L'objectif reste de finaliser une V1 des prescriptions pour la fin de l'année. En 2015 le groupe passera sous « étiquette » CNIG afin de faire vivre les prescriptions et de réaliser un guide pour les utilisateurs.

Enfin le Cerema va produire un test sur une zone de la région Nord Pas-de-Calais sur la nomenclature 4 dimensions, sur 2 millésimes et deux

méthodes (un mixte traitement automatique et compléments PIAO et un tout PIAO).

Contrôles qualités

Présentation La présentation de Laure Wateau porte sur les contrôles qualités externes (CQE) fait par un prestataire différent de celui qui produit la base de données.
Laure WATEAU

Le CQE n'a rien à voir avec la qualité externe de la base de données. Le CQE s'intéresse à la qualité interne¹.

On distingue deux types de CQE :

- le CQE structurel qui s'intéresse à la topologie, la sémantique et la géométrie ;
- Le CQE thématique.

Dans tous les cas la prestation autour de la base de données (production et contrôle qualité) commencera par une période de réflexion commune entre maîtrise d'ouvrage, maîtrise d'œuvre de production et maîtrise d'œuvre de contrôle, pour adopter une position commune.

La présentation a ensuite été axée sur les CQE thématiques. Laure Wateau procède par points de contrôle répartis ainsi :

- contrôle orienté (40%)
- contrôle aléatoire (60%)

À la fin la prestation de CQE finira toujours avec l'établissement d'un rapport d'expertise et un PV de recette.

Le CQE réalisé par le prestataire extérieur sur la base d'observation à partir d'image pourra être complété par un CQE terrain réalisé par le maître d'ouvrage lui-même. Attention toutefois à ce que ce contrôle terrain soit en cohérence avec la date et la période à laquelle a été réalisé la prise de vues ou les images satellites.

En conclusion on peut noter l'importance :

- sur les aspects techniques
 - de caler la nomenclature avant le débuter la production ;
 - de choisir des images en multi date si possible ;
 - de l'importance des données exogènes.
- Sur les aspects organisationnels :
 - de l'importance de la collaboration entre les parties ;
 - de la visite terrain et la formation des photos interprètes.

1 Se référer à <http://www.certu-catalogue.fr/la-qualite-des-donnees-geographiques.html>

Groupe National occupation du sol à grande échelle

Pour Laure Wateau, un taux de fiabilité de 90 % d'une BD OCS est un très bon taux.

Présentation IGN Le questionnement posé par l'IGN est la suivant :
(Thierry TOUZET) comment procéder à un contrôle qualité interne d'une base de données départementale en ayant un « taux de fiabilité » final qui tient compte :

- des différents postes ;
- des différentes tailles de polygones
- de la répartition géographique des polygones par postes.

La mise en place du contrôle se fait par la constitution manuelle de strates en fonctions de critères spécifiques (surface relative de types, représentativité des types, proximité des postes, nombres de polygones par types, priorité à certains postes)

Afin de sortir un taux de fiabilité réaliste, les plus grands polygones sont exclus et un calcul de taux d'erreur est obtenu par double saisie des zones retenues.

Reste toutefois a faire attention aux cas particuliers que sont :

- les gros polygones
- les strates peu représentées
- les strates entièrement contrôlées

qui peuvent altérer le résultat. Il faut donc prévoir une prise en compte spécifique.

Attention : le choix de la zone de contrôle à forcement une orientation. De même un contrôle peut être fait en cherchant l'exhaustivité ou les omissions.

Echanges FD : Pour exemple, certains PNR avaient mis en place lors de leur première production de donnée OCS, une méthode interne de CQ en réalisant par placette une interprétation du terrain à la même période que la prise de vue, constituant ainsi une information comparable sans décalage dans le temps,

Est il possible de récupérer les observations de levés terrains annuels dans le cadre des observations statistiques faites pour l'inventaire forestier ou Teruti-lucas ? (chercher la réponse auprès de l'IGN ou du ministère de l'agriculture, de l'agroalimentaire et de la foret).

SF : l'IAU utilise des matrices de confusion pour comparer deux interprétations afin de lever les problèmes strictement dus à l'interprétation des photos (problème de la proximité thématique.)

Groupe National occupation du sol à grande échelle

CA : le CRIGE PACA a également choisi de croiser deux millésimes pour corriger des incohérences et modifier le millésimes d'origine.

HD : La méthodes basée sur des points statistiques n'est pas forcément la méthode la plus adaptée (récupérer autre méthode de contrôle). Concernant les sources exogènes à utiliser il est important de considérer l'impact de l'usage de données très précises (exemple NATURA2000) dans une base de données OCS à moyenne échelle. Enfin, un paramètre de conformité aux spécifications attendues ne doit pas être oublié : le respect de l'UMI. Bien que représentant un coût pour le CQ Externe, ce point est fondamental pour la pertinence des indicateurs qui en découle, tout autant que pour le suivi des dynamiques spatiales des entités fragmentées (mitage urbain en milieu agricole par exemple, petits boisements support de continuité pour la TVB, etc).

FD : Lors d'un marché de production de donnée OCS, la méthode de contrôle interne du prestataire/producteur proposée dans sa réponse est un élément clé que le Maître d'Ouvrage examine avec attention pour analyser les offres et effectuer son choix, Un CQE ne peut se substituer a un contrôle interne du producteur.

Le bon sens du pilotage de la production veut qu'il y ai une rencontre entre maître d'ouvrage, producteur et contrôleur.

Dans le cadre de la réalisation du marché le maître d'ouvrage peut exiger un taux de qualité variable suivant les niveaux d'informations (niveau 1 2 ou 3).

BG : Dans le cadre des prescriptions nationales est ce que l'on doit prescrire une méthode de contrôle ?

Non mais plutôt un processus et des grands principes, incontournables pour un maître d'ouvrage.

Pistes pour préconiser des indicateurs ou autres éléments permettant de qualifier la donnée ?

- regarder les indicateurs du côté INSPIRE ;
- prévoir un document d'accompagnement avec notamment des points de vigilance concernant l'utilisation des données.

Modèle de données

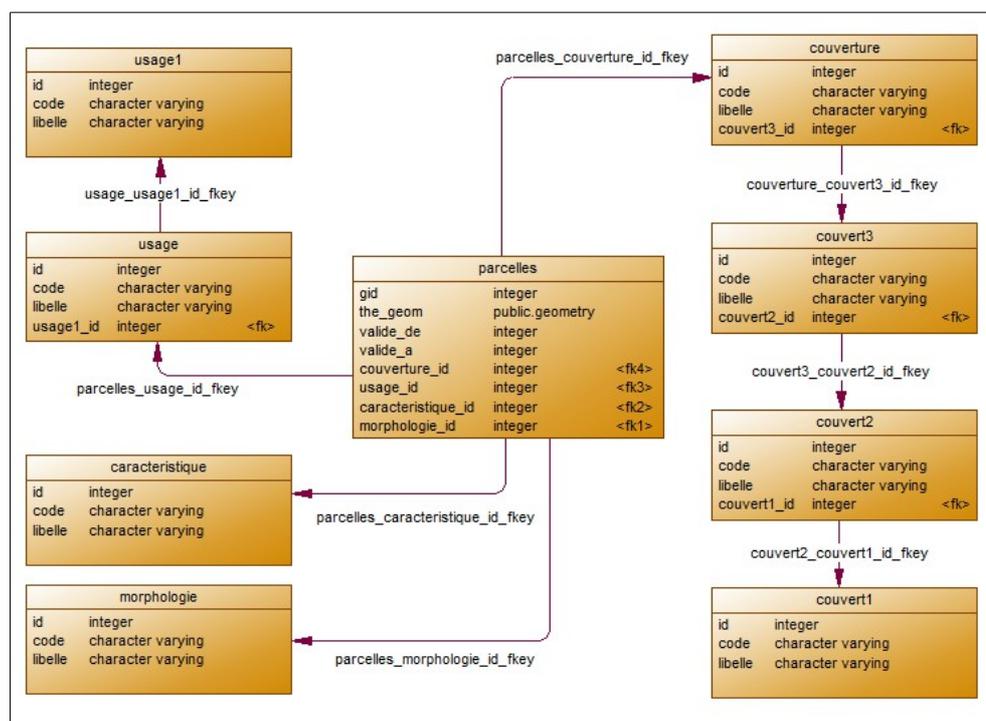
Pour rappel le groupe national a acté lors de la précédente réunion que nous devons réfléchir à une nouvelle modélisation des bases de données occupation du sol dans la mesure où :

- l'introduction de 4 dimensions pour décrire l'espace va compliquer l'exploitation ;

Groupe National occupation du sol à grande échelle

- Le modèle MOS montre ses limites notamment pour les territoires ayant de nombreux millésimes à disposition.

Présentation Présentation des réflexions menées au niveau de la direction technique territoires et ville sur la base d'un besoin simple :
« quelle est l'occupation du sol d'un polygone à telle date »



(suite du document en annexé du CR)

DMC: le modèle est bien pour deux ou trois millésimes mais une modification de géomètre entraîne forcément la modification des géométries postérieures contrairement au MOS.

BG : Sophie Foulard envoi un jeu test sur l'ensemble des millésimes du MOS IAU afin de faire des tests sur la modèle de données.

PB : Comment fait on lorsque l'on a seulement 3 niveaux dans le couvert par exemple ?

À faire pour la prochaine réunion

- Produire le CR ;
- Produire document sur contrôle qualité qui sera ajouté aux « prescriptions nationales » ;
- Production du MCD complet.

- Doodle pour la prochaine réunion
<http://doodle.com/76zix4hg9mti2sff>

Annexe 1

Normalisation de la base

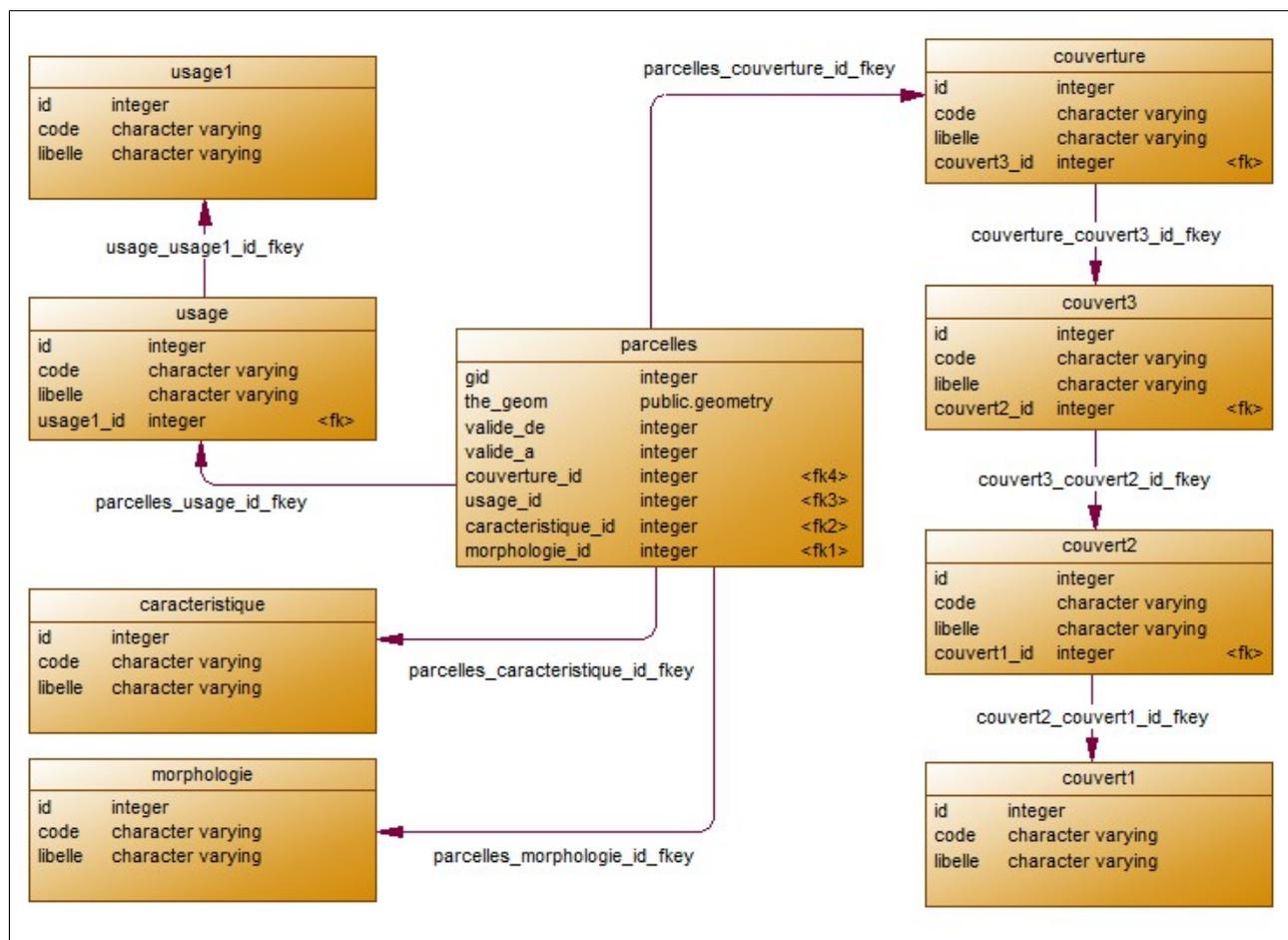


Illustration 1: Modèle de données

Ajout d'information sur les années / dates de validité

Les colonnes « valide_de » et « valide_a » permettent de connaître les millésimes pendant lesquels la donnée est valable.

Une requête ad-hoc permet de ne ramener que les données (géométrie incluse) valides pour l'année n :

```
Select * from parcelles join ..... where (valide_de ≤ n) and (n ≤ valide_a or valide_a is null)
```

Mise à jour des données

Connaissant les géométries à ajouter avec validité à partir de l'année n (G1 à G5 par exemple)

Mise en fin de validité des anciennes géométries remplacées

Les géométries remplacées par les nouvelles géométries à insérer en base doivent recevoir une fin de validité à l'année (n-1)

Groupe National occupation du sol à grande échelle

Elles sont trouvées par intersection avec les nouvelles, moyennant une marge d'erreur que l'on peut ajuster, 50m dans l'exemple ci-dessous :

```
UPDATE parcelles set valide_a = (n-1) where valide_a is null and st_area(st_intersection(the_geom, G1)) > 50;  
UPDATE parcelles set valide_a = (n-1) where valide_a is null and st_area(st_intersection(the_geom, G2)) > 50;  
UPDATE parcelles set valide_a = (n-1) where valide_a is null and st_area(st_intersection(the_geom, G3)) > 50;  
UPDATE parcelles set valide_a = (n-1) where valide_a is null and st_area(st_intersection(the_geom, G4)) > 50;  
UPDATE parcelles set valide_a = (n-1) where valide_a is null and st_area(st_intersection(the_geom, G5)) > 50;
```

Ajout des nouvelles données

On peut ensuite ajouter les nouvelles données avec une validité à partir de l'année n :

```
INSERT INTO parcelles VALUES (DEFAULT, G1, 8,8,4,2,n,null);  
INSERT INTO parcelles VALUES (DEFAULT, G2, 8,8,4,2,n,null);  
INSERT INTO parcelles VALUES (DEFAULT, G3, 8,8,4,2,n,null);  
INSERT INTO parcelles VALUES (DEFAULT, G4, 8,8,4,2,n,null);  
INSERT INTO parcelles VALUES (DEFAULT, G5, 8,8,4,2,n,null);
```

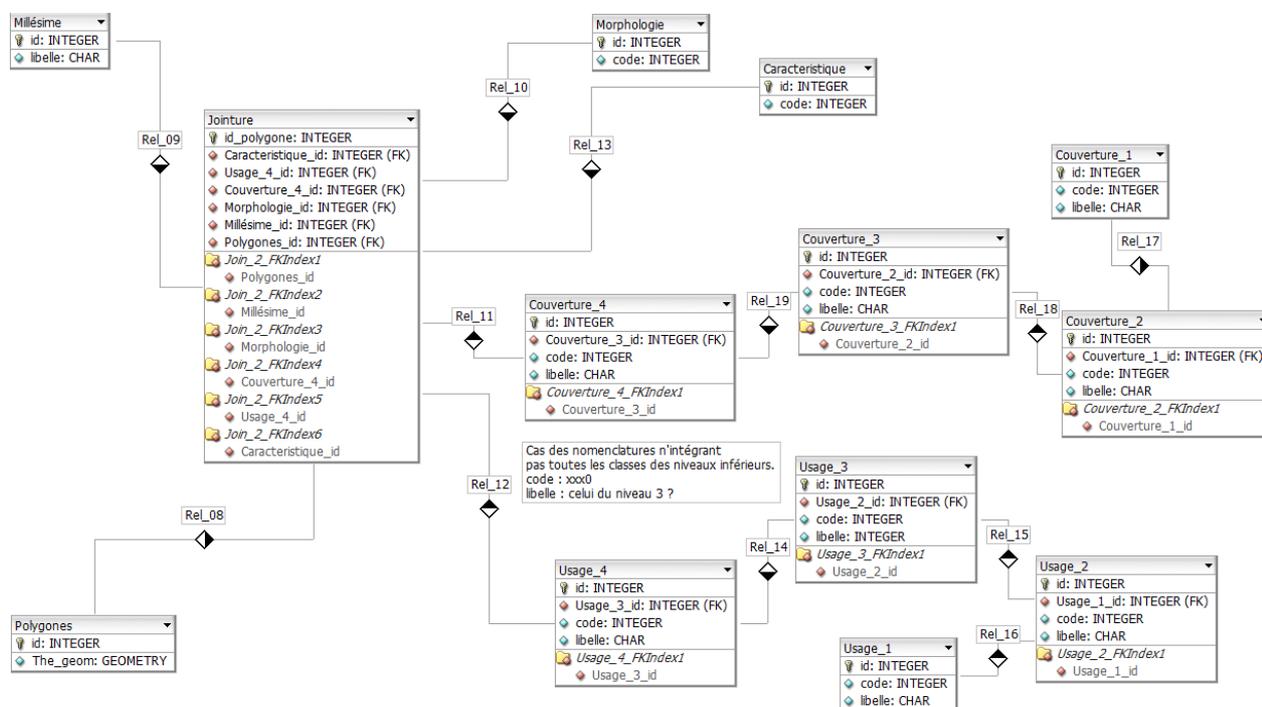
Exploitation

Ces mises à jour, décrites en SQL, peuvent tout à fait être automatisées dans un script produit par « ETL » (Extract-Transform-Load) qui saura gérer les formats SIG usuels dans lesquels seront vraisemblablement fournies les nouvelles données et les insérer dans une base de données spatialisée.

À noter qu'à l'inverse, il est très simple de produire un fichier SHP depuis une table, une vue ou une requête PostgreSQL (outil pgsq2shp).

Annexe 2 :

Proposition de modèle réalisée en séance par Pascal Barillé (Pays d'Auray) pendant nos discussions



- Une table polygone servant à stocker les contours des zones de MOS sans doublons
- Une table millésime stockant les années de production
- les tables couverture_x, usage_x, morphologie_x et caractéristique_x stockant la nomenclature
- La table jointure sert à assembler le tout aux différents millésimes