

Protocole de constitution d'une base de données 1960

Rapport numéro	L2.0-3
Titre	Constitution d'une base de données topographique vectorielle à partir des cartes de 1960 et de la BDTOPO2011
Rédigé par	Lucie Nahassia (COGIT/IGN), Christine Plumejeaud (COGIT/IGN)
Etat (en cours / final)	Final
Relu par	Sébastien Mustière (COGIT/IGN)
Date	Septembre 2012

Nous décrivons notre méthodologie de constitution d'une base de données topographique vectorielle à partir des cartes de 1960. La méthode choisie consiste à dupliquer la BDTOPO 2011, et à modifier les entités (supprimer, créer, modifier) pour faire correspondre cette copie à la carte topographique scannée et géoréférencée datant des années 1960.

Sommaire

1. Description des objectifs	3
2. Méthodologie	4
2.1. Sources de données	4
2.1.1. BDTOPO et classes retenues	4
2.1.2. Données cartographiques	6
2.2. Outils logiciels	8
2.3. Méthode de clarification vectorielle	9
2.3.1. Préparation de la base de données	9
2.3.2. Clarification	11
3. Expérimentation	14
3.1. Test et résultat de la méthode	14
3.2. Problèmes et questions ouvertes	14
4. Expérimentation de méthodologie allégée	16
4.1. Sources de données	16
4.2. Méthodologie allégée	16
4.2.1. Préparation de la base de données	16
4.2.2. Clarification	17
4.3. Test et résultats	18
5. Conclusion	21

1. Description des objectifs

L'objectif est de créer une base de données vectorielles correspondant à l'état du territoire dans les années 1960 pour les quatre zones du projet GéoPeuple et pour les classes qui nous intéressent. Les données produites devront porter une information sur leur évolution entre les années 1960 et 2011 (disparition, transformation, etc).

Cette base de données serait utilisée dans le processus de développement d'outils d'analyses de l'évolution des territoires communaux.

Ces données seraient à mettre en regard avec les objets issus de la vectorisation des cartes de Cassini, des cartes d'Etat-Major et ceux de la BDTOPO 2011.

Dans cette optique, la BDTOPO 2011 possède une résolution et une précision beaucoup plus fine que les données historiques. Néanmoins dans un souci de précision et d'éventuelle réutilisation des données ainsi créées, nous avons choisi de s'appuyer, pour cette mise à jour, sur les cartes topographiques au 1 : 25 000 (TOP25) et sur la BDTOPO 2011.

2. Méthodologie

2.1. Sources de données

2.1.1. BDTOPO et classes retenues

Données vectorielles : BDTOPO 2011

Pour les caractéristiques de la BDTOPO, voir le fichier de description fournie avec la base de données.

Les classes à retenir ont été sélectionnées en fonction de leur intérêt potentiel et/ou remarqué par rapport aux données vectorielles historiques collectées par le projet. Par exemple, dans le cadre de notre étude, au vu de l'importance passé des bâtiments religieux, nous avons sélectionné les CLASSES susceptibles de contenir des croix et des calvaires.

Un premier ensemble de classe a été choisi :

- | | |
|---------------------------|--------------------------------|
| - ROUTE | - CONSTRUCTION_SURFACIQUE |
| - Gare | - ZONE_VEGETATION |
| - TRONCON_VOIE_FERREE | - COMMUNE |
| - POINT_EAU | - CHEF_LIEU |
| - SURFACE_EAU | - PAI_ADMINISTRATION_MILITAIRE |
| - TRONCON_COURS_EAU | - PAI_ESPACE_NATUREL |
| - Hydronyme | - PAI_INDUSTRIEL |
| - BATI_INDIFFERENCIE | - PAI_RELIGIEUX |
| - BATI_REMARQUABLE | - PAI_SANTE |
| - BATI_INDUSTRIEL | - PAI_ZONE_HABITATION |
| - CIMETIERE | - PAI_HYDROGRAPHIE |
| - CONSTRUCTION_LINEAIRE | - Lieu_dit_habite |
| - CONSTRUCTION_PONCTUELLE | - Lieu_dit_non_habite |

Les classes en minuscules ont ensuite été mises de côtés pour plusieurs raisons :

- Les objets de la classe **LIEU_DIT_HABITE** sont tous dupliqués dans la classe **PAI_ZONE_HABITATION** (les deux classes sont des classes d'objets ponctuels), avec le même ID. Pour les mêmes raisons on écarte la classe **GARE** (objets contenus dans **BATI_REMARQUABLE**), **HYDRONYME** (objets contenus dans **PAI_HYDROGRAPHIE**) et **LIEU_DIT_HABITE** (objets contenus dans **PAI_ZONE_HABITATION**).
- Les objets de la classe **LIEU_DIT_NON_HABITE** sont dupliqués dans les classes **PAI_CULTURE_LOISIRS**, **PAI_ESPACE_NATUREL**, **PAI_INDUSTRIEL_COMMERCIAL**, **PAI_ADMINISTRATIF_MILITAIRE** et **PAI_RELIGIEUX** (toutes ces classes sont des classes d'objets ponctuels), avec le même ID. En dehors de la classe **PAI_CULTURE_LOISIRS**, toutes ces classes ont été sélectionnées pour l'étude. Quatre objets de la classe **LIEU_DIT_NON_HABITE** sont contenus dans cette première: trois digues et un jardin des plantes. On décide donc d'écarter la classe **LIEU_DIT_NON_HABITE**, et de faire migrer les quatre objets

contenus dans la classe PAI_CULTURE_LOISIRS dans la classe PAI_ESPACE_NATUREL (c'est cette classe-ci qui est la plus proche de la nature des objets).

D'autres classes se doublent, mais les objets sont différents : points et polygones, ID différents. C'est par exemple le cas d'une partie des objets de la classe PAI_RELIGIEUX : des points de nature église localisent une partie des polygones de nature églises de la classe BATI_REMARQUABLE.

Dans ce cas on garde les deux classes car, d'une part, elles ne se recoupent pas exactement, et, d'autre part, les vecteurs qu'elles contiennent ne sont pas de même nature.

Les zones de travail du projet correspondent aux zones suivantes de la BDTOP0 :

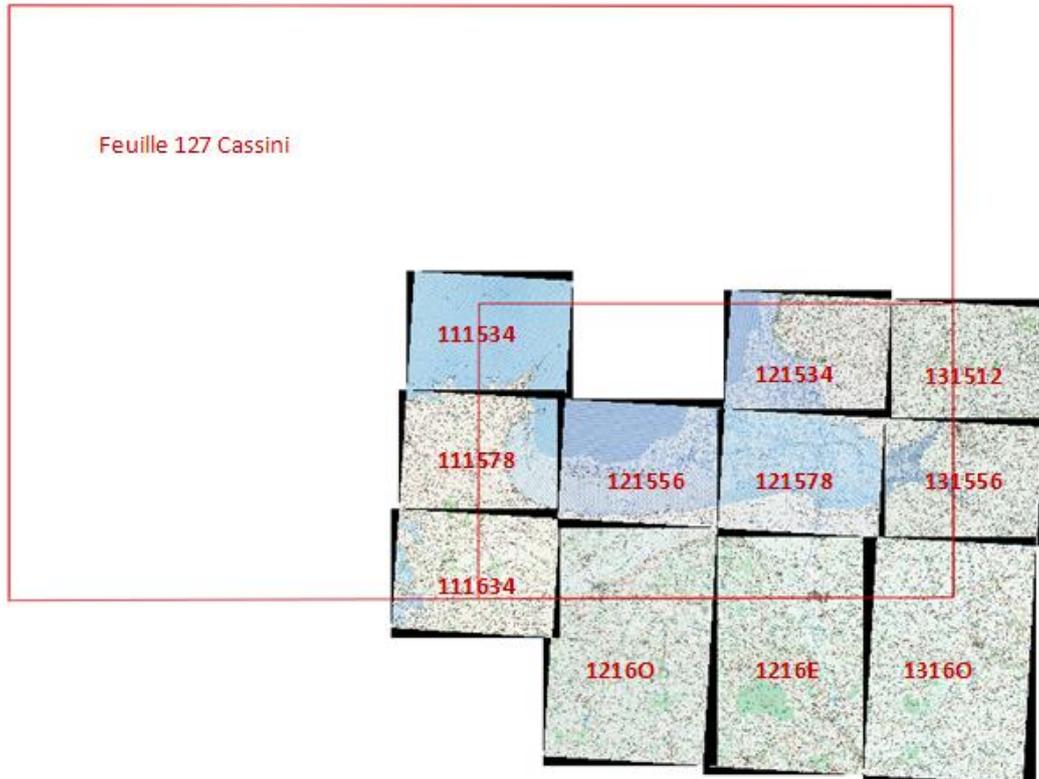
	Zone BDTopo	Référence IGN	Nom du dossier	Emprise temporelle
Reims	Reims Sud	IGNF_BDTOPOr_2-0_SHP_LAMB93_R021	BDT_2-0_SHP_LAMB93_R021-ED111-REIMS-SUD	2002-2011
Saint-Malo	Saint-Malo NE	IGNF_BDTOPOr_2-0_SHP_LAMB93_R025	BDT_2-0_SHP_LAMB93_R025-ED111-StMaloNordEst	2000-2011
	Dinan	IGNF_BDTOPOr_2-0_SHP_LAMB93_R053	BDT_2-0_SHP_LAMB93_R053-ED111-DINAN	1999-2011
Agen	Agen Sud	IGNF_BDTOPOr_2-0_SHP_LAMB93_R073	BDT_2-0_SHP_LAMB93_R073-ED111-AgenSud	1998-2011
Grenoble	Grenoble	IGNF_BDTOPOr_2-0_SHP_LAMB93_R082	BDT_2-0_SHP_LAMB93_R082-ED111-GRENOBLE	1998-2011

Tableau 1. Références et emprise temporelle des shapefiles de la BDTOP0 couvrant les zones d'étude.

La première date correspond à la datation la plus ancienne d'un thème de la base, la deuxième à la révision la plus récente d'un thème de la base.

2.1.2. Données cartographiques

La zone de travail traitée jusqu'à présent ne concerne pas toute la feuille de Cassini mais seulement le quart de feuille (sud-est pour les zones de Saint-Malo, Grenoble et Agen, sud-ouest pour la zone de Reims).



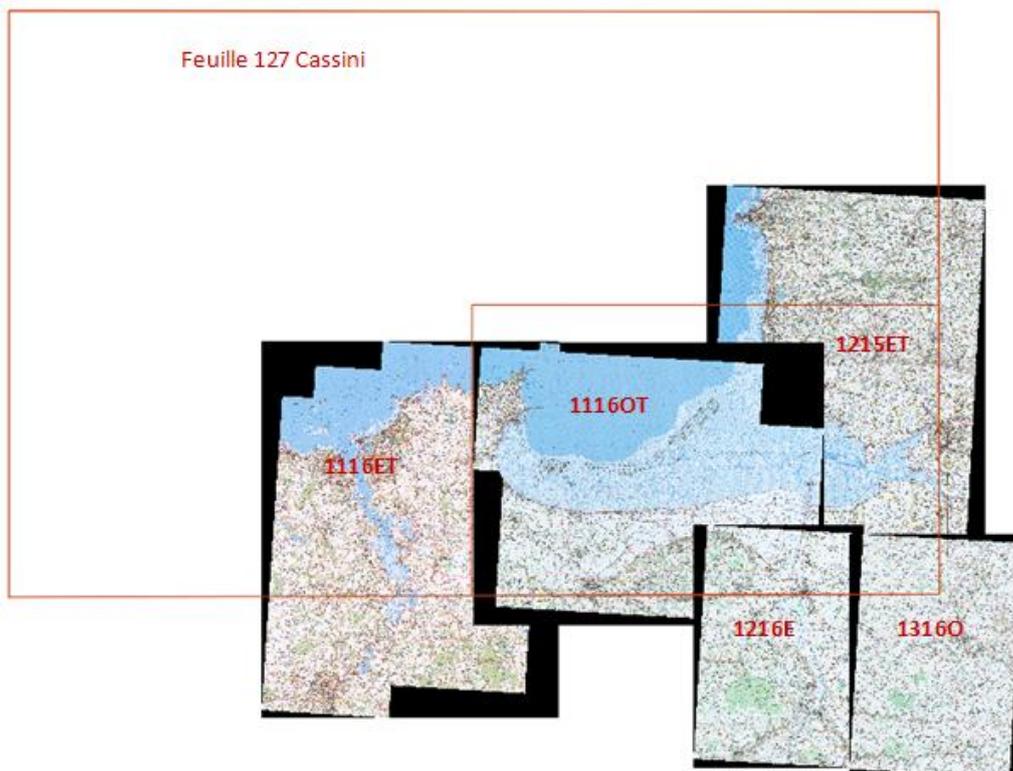


Figure 1. Exemple pour la zone de Saint-Malo, cartes des années 1960 (en haut) et 2010(en bas) Pour effectuer la clarification, on s’appuie sur les cartes au 1 :25000 couvrant la zone d’intérêt.

Cartes scannées des années 1960, 1 :25000

Références par zones (avec date de publication) :

ST-MALO		AGEN		REIMS		GRENOBLE	
Référence	Date	Référence	Date	Référence	Date	Référence	Date
111534	1969	184156	1960	271234	1964	313412	1959
111578	1969	184178	1960	271278	1964	313434	1959
111634	1968	184212	1962	271334	1962	313456	1959
121534	1970	184234	1962	271378	1962	313478	1959
121556	1970	184256	1962	281212	1964	313512	1959
121578	1970	184278	1962	281234	1964	313534	1959
1216E	1976	194156	1953	281256	1964	313556	1959
1216O	1976	194212	1962	281278	1964	313578	1959
1316O	1976	194256	1962	281312	1962	323412	1959
131512	1971			281334	1962	323434	1965
131556	1971			281356	1962	323456	1965
				281378	1962	323478	1965
				291212	1959	323512	1956
				291234	1949	323534	1956
				291256	1949	323556	1956
				291278	1949	323578	1956
				291312	1949	333412	1953
				291334	1949	333434	1953
				291356	1955	333456	1953
				291378	1955	333478	1953

en rouge : cartes non disponibles
(non scannées ni géoréférencées)

333512	1953
333534	1953
333556	1953
333578	1953

Tableau 2. Références des cartes 1960 couvrant les zones d'étude.

Les cartes en rouges pourraient être par la suite scannées par le service cartographie de l'IGN et géoréférencées par nos soins. Dans la mesure où aucune carte de Grenoble n'est disponible, l'emprise précise de celles-ci par rapport à la zone n'a pas été vérifiée.

Nous disposons aussi de cartes TOP25 au 1 :25000 scannées et géoréférencées sur les mêmes zones de travail, pour les années 2000-2010. Celles-ci nous permettent d'effectuer des vérifications au cours de la clarification, facilitent certaines saisies et aident à la décision. En effet la présence ou non de certains objets, la forme de certains polygones, peut prêter à confusion. L'état des choses sur les cartes représentant le territoire à peu près dans les années correspondant à la BDTOPO 2011 permettent parfois de lever ces ambiguïtés.

L'idéal aurait été de disposer des cartes datées strictement des années 2010, mais les scans à dispositions s'étendent de 1979 (une seule carte) à 2011. La plupart sont toutefois compris entre 2000 et 2011. Cette imprécision n'est pas trop préjudiciable dans la mesure où ces données constituent seulement une aide à la saisie.

Cartes scannées des années 2000-2011, 1 :25000

Références par zones (avec date de publication) :

ST-MALO		AGEN		REIMS		GRENOBLE	
Référence	Date	Référence	Date	Référence	Date	Référence	Date
1116ET	2009	1841E	2004	2712E	2006	3134E	2010
1215ET	2008	1841O	2006	2712O	2008	3134O	2010
1215OT	2008	1842E	2006	2713E	2006	3135E	1979
1216E	2008	1842O	2006	2713O	2006	3135O	2007
1216O	2008	1941E	2000	2812E	2009	3234E	1986
1316E	2009	1941O	2001	2812O	2009	3234O	2010
1316O	2008	1942E	2001	2813O	2009	3235OT	2011
		1942O	2009	2813E	2009	3334OT	2005
				2912E	2007	3335ET	2010
				2912O	2007	3335OT	2005
				2913E	2007		
				2913O	2009		

en rouge : cartes non disponibles
(non scannées ni géoréférencées)

Tableau 3. Références des cartes 2000-2011 couvrant les zones d'étude.

De la même manière que pour les scans de 1960, les cartes en rouges pourraient être scannées et géoréférencées par nos soins, et l'emprise exacte des cartes de Grenoble n'a pas été vérifiée.

2.2. Outils logiciels

Pour la gestion de la base de données et sa mise à jour (ajout de champs, etc) : PostGreSQL (version 9.1) + cartouche spatiale PostGIS (version 1.5), interface pgAdmin III (version 1.14.2)

Pour quelques modifications simples de la base de données (les attributs non spatiaux), édition du

.dbf avec un tableur (Microsoft Excel ou OpenOffice Calc).

Pour l'édition des fichiers vectoriels (suppression/saisie d'objets vectoriels) : Quantum GIS (version 1.7.4 Wrocław)

2.3. Méthode de clarification vectorielle

Le mode opératoire suivant est valable par zone, et permet de constituer 4 bases de données 1960. Il a été testé sur St Malo sur le quart de feuille sud-est, et sur une classe d'objets ponctuels (PAI_RELIGIEUX).

2.3.1. Préparation de la base de données

1. Découper tous les shapefiles des classes sélectionnées selon l'emprise de la zone d'étude (un ou deux fichiers par classe d'objet) avec l'outil `recursive_bat` développé par Eric Grosso basé sur l'utilitaire `ogr2ogr` de la librairie OGR.
2. Pour Saint-Malo : raccorder pour chaque classe les vecteurs présents sur les différents shapefiles couvrant chaque zone :
 - avec l'outil « Vecteur/Outils de gestion de données/Fusionner les shapefiles en un seul » (ne marche pas sous QGIS 1.7.4 Wrocław),
 - en copiant/collant les vecteurs d'une zone vers l'autre.
3. Sauvegarder les shapefiles originellement en Latin9 (ISO8859-15) en UTF8 sous QGIS. Attention aux champs en chaîne de caractères qui peuvent être trop courtes, produisant des erreurs dans les attributs.
 - dans ce cas modifier la longueur maximale de caractères attribuée au champ en question en éditant son titre dans le fichier .dbf correspondant au vecteur (troisième élément). Il pourra aussi ultérieurement être nécessaire de modifier la longueur des champs ID (via un tableur ou pgAdmin III) si, au début du traitement, nous nous apercevons qu'ils sont trop courts pour contenir les nouveaux identifiants.

	A	B	C	D	E	F	G	H
1	ID,C,24	PREC_PLANI,N,6,1	PREC_ALTI,N,7,1	ORIGINE_BAT,C,8	NATURE,C,25	HAUTEUR,N,4,0	Z_MIN,N,8,2	Z_MAX,N,8,2
2	BATIMENT00	1,5	1	Autre	Gare	10	14,3	16,6
3	BATIMENT00	1,5	1	Autre	Gare	8	23,5	26,8
4	BATIMENT00	1,5	1	Autre	Mairie	6	32,2	33,9
5
6								

Figure 2. Ouverture du fichier attributaire .dbf de la classe BATI_REMARQUABLE en UTF8 sous Open Office Calc : les champs NATURE et ORIGINE_BAT sont de nature chaîne de caractères (deuxième élément du titre : « C »), et il est possible de modifier la longueur maximale du champ si nécessaire (longueur actuelle en rouge).

4. Si nécessaire, sauvegarder les shapefiles en projection Lambert93 (EPSG 2154) sous QGIS.

→ la BDTOPO est bien en Lambert93 mais QGIS, lors de l'importation des shapefiles, leur attribut un SCR « définit par l'utilisateur ». Cette manipulation ne change donc pas le SCR mais normalise seulement son nom.

5. Importer les shapefiles ainsi créés en base de donnée sous PostGreSQL avec le plugin PostGIS « Shapefile and DBF loader »; le schéma des tables est le suivant :
(GID, OID, attribut1, attribut2, ..., attribut_n, géométrie)
GID correspond à l'identifiant attribué par PostGreSQL
OID correspond à l'identifiant natif de la BDTOPO 2011
6. Changer ce schéma pour rajouter quatre colonnes :
(GID, OID, attribut1, attribut2, ..., attribut_n, géométrie, **statut1960**, **idSource**, **commentaire1960**, **incertain**)

Code SQL pour l'insertion de ces quatre champs dans une table de nom "nomX" :

```
ALTER TABLE "NOMX " ADD COLUMN statut1960 integer NOT NULL DEFAULT -1;
ALTER TABLE TABLE "NOMX " ADD COLUMN idSource character varying(255)DEFAULT 'NR';
ALTER TABLE TABLE "NOMX " ADD COLUMN commentaire1960 character varying(255) DEFAULT 'NR';
ALTER TABLE TABLE "NOMX " ADD COLUMN incertain boolean NOT NULL DEFAULT FALSE;
```

7. Modifier le schéma de la base de données pour supprimer la contrainte de géométrie imposée aux objets (contrainte 4 = objets 3D) ce qui nous permettra de tracer nos polygones et lignes sans problèmes sous QGIS (objets 2D correspondant à la contrainte 2).

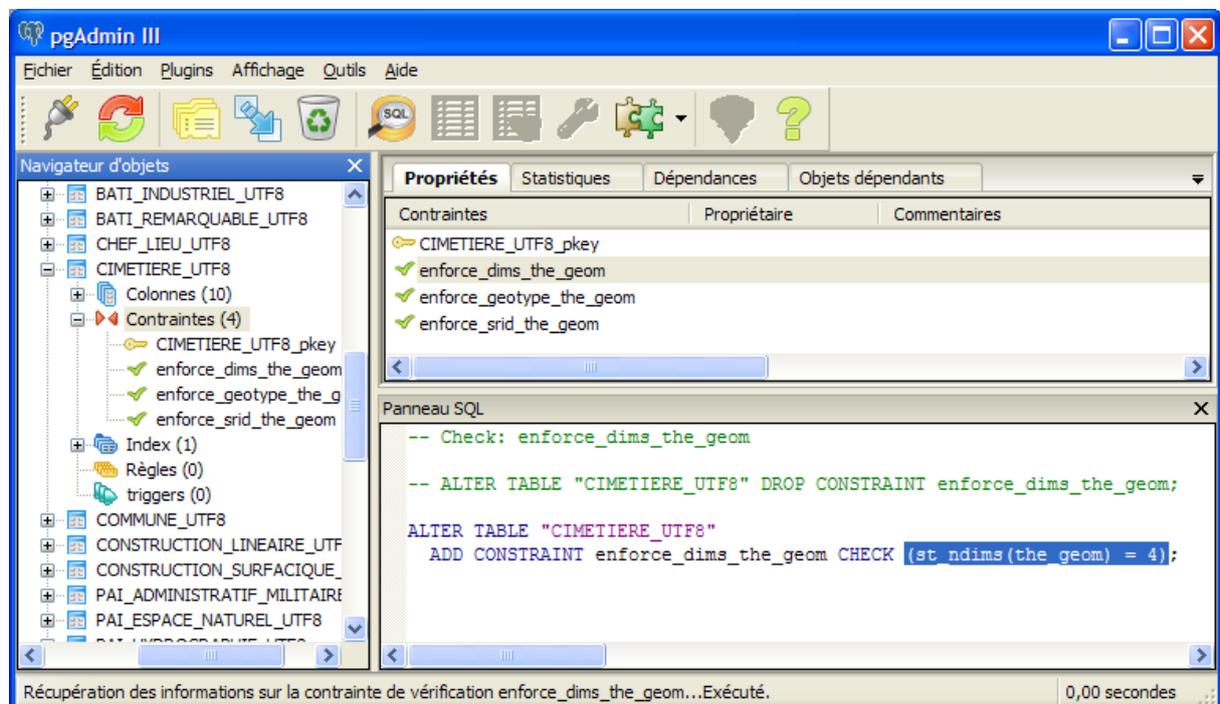


Figure 3. Visualisation sous pgAdmin III des contraintes liées à la table CIMETIERE. Dans le panneau SQL, surligné en bleu, les propriétés de la contrainte de géométrie (« enforce_dims_the_geom ») : les dimensions de l'objet doivent être de type 4 (3D).

Code SQL pour la suppression de la contrainte géométrique dans une table de nom « nomX » :

```
ALTER TABLE "NOMX" DROP CONSTRAINT enforce_dims_the_geom;
```

Les étapes 6 et 7 peuvent être traitées dans un même script .sql.

statut1960 vaut :

- 1 par défaut = la donnée n'a pas été comparée,
- 0 si l'entité est présente en 1960 et en 2011, sans changement,
- 1 si l'entité n'est pas présente en 1960 et est apparue en 2011 (création),
- 2 si l'entité est présente en 1960 et a disparue en 2011 (disparition),
- 3 si l'entité a changé de forme entre 1960 et 2011 (extension de bâtiment par exemple),
- 4 si l'entité a changé de fonction ou de classe entre 1960 et 2011 (exemple : une route a changé de niveau d'importance, un bâti_indifférencié est devenu un bâti_industriel).

idSource vaut :

- NR par défaut = le champ n'est pas utilisé,
- L'ID d'un autre objet d'une autre classe si on enregistre une modification de classe.

commentaire1960 vaut :

- NR par défaut = le champ n'est pas utilisé,
- « Existe en 2011 » si l'objet existe sur le scan 2011 mais pas dans la BDTOPO2011,
- Tout autre commentaire pour lever une ambiguïté.

incertain vaut :

- FALSE par défaut = la donnée est sûre ,
- TRUE si la donnée est incertaine.

2.3.2. Clarification

Comparer le contenu de cette BD avec la carte raster 1960 (placée en dessous sous QGIS), et procéder systématiquement par couche jusqu'au traitement de tous les objets vectoriels. S'aider des scans 2011.

1. Commencer par traiter les objets existant sur le scan 1960 et **n'apparaissant pas dans la BDTOPO 2011**.

Une partie de ces objets correspond aux objets **ayant disparus en 2011** (code 2), mais d'autres sont des objets existant bien sur les scans en 2011 (code 0 ou 1) mais **non renseignés dans la BDTOPO 2011** (cf. commentaire en 2.3.3).

Pour ces deux types d'objet :

- a. créer les objets (point, ligne ou polygone),
- b. leur attribuer un identifiant de type <CLASSE>_n, n variant de 0 à l'infini, défini par l'ordre des opérations dans le champ OID ; l'identifiant GID est géré automatiquement ,

- c. leur attribuer la valeur 2 (objets absent sur le scan 2011 et présent en 1960), 1 (objet présent en 2011 et absent en 1960) ou 0 (objet présent sur les scans 2011 et 1960) dans le champ statut1960,
- d. si incertitude quant à la nature ou certains attributs, et pour tous ceux de code 0, changer le champ incertain en TRUE et remplir le champ commentaire1960.

Pour faciliter le traitement, on modifie la représentation des objets pour qu'ils apparaissent transparents lorsque le champ « statut1960 » est différent de -1.

Tous les objets qui viennent d'être créés disparaissent donc et, au fur et à mesure du traitement, les objets déjà comparés disparaissent à la vue de l'opérateur.

- 2. Traiter les objets qui **existent et sont stables entre 1960 et 2011** : leur attribuer la valeur 0 dans le champ statut1960.
- 3. Traiter les objets **qui n'existent pas en 1960** et existent en 2011 : leur attribuer la valeur 1 dans le champ statut1960.
- 4. Traiter les objets qui **changent de forme entre 1960 et 2011** (ne concerne que les objets linéaires et surfaciques) (code 3).

Cette saisie est limitée car coûteuse. Pour GéoPeuple en effet, ils n'ont d'intérêt que pour une variation supérieure à 20m, qui est le niveau de résolution des cartes d'Etat-Major (on ne peut s'appuyer sur la résolution des cartes de Cassini pour ce traitement car leur résolution, très peu précise, varie en plus de manière trop importante selon les objets : entre 300m et 4km).

- a. modifier la forme de l'objet ,
- b. lui attribuer la valeur 3 dans le champ statut1960,
- c. si incertitude, changer le champ incertain en TRUE et remplir si nécessaire le champ commentaire1960.

- 5. Traiter les objets qui changent de classe (c'est-à-dire de fonction) ou d'importance entre 1960 et 2011 (code 4).

Pour un **changement d'importance** :

- a. lui attribuer la valeur 4 dans le champ statut1960,
- b. attribuer une nouvelle valeur dans le champ correspondant (en général « importance »),
- c. si incertitude, changer le champ incertain en TRUE et remplir si nécessaire le champ commentaire1960.

Pour un **changement de classe/fonction** (par exemple : un bâtiment industriel en 2011 est indifférencié en 1960) :

→ Dans la classe de l'objet en 2011 (BATI_INDUSTRIEL dans notre exemple) :

- a. attribuer la valeur 4 au champ statut1960,
- b. remplir le champ idSource1960 avec l'ID créé à la deuxième étape (voir plus bas),
- c. si incertitude changer le champ incertain en TRUE et remplir si nécessaire le champ commentaire1960.

→ Dans la classe de l'objet tel qu'il est en 1960 (BATI_INDIFFERENCIE dans notre exemple) :

- d. créer un nouvel objet (une simple copie de l'objet de la classe de 2011 suffit en général),
- e. lui attribuer un identifiant de type <CLASSE>_n, n variant de 0 à l'infini, défini par l'ordre des opérations dans le champ OID ; l'identifiant GID est géré automatiquement,

- f. attribuer la valeur 4 au champ statut1960,
- g. remplir le champ idSource avec l'ID de l'objet en 2011
→ ce champ peut contenir plusieurs ID si l'objet existe dans plusieurs classes (sous forme de point et de polygone par exemple),
- h. si incertitude changer le champ incertain en TRUE et remplir le champ commentaire1960.

Si l'objet a changé de forme et de fonction, c'est le changement de fonction qu'on enregistre (il est plus intéressant pour nous).

3. Expérimentation

3.1. Test et résultat de la méthode

Pour établir et tester notre méthodologie nous avons d'abord travaillé sur une classe ponctuelle comprenant peu d'objets : PAI_RELIGIEUX.

Le traitement complet d'1/3 de la zone de travail (180 km² traités pour 580km² au total) a duré environ trois heures, pour un total de 63 objets observés en 1960 (alors que seulement 25 objets sont présents dans la BDTopo 2011). Cette durée correspond surtout à la recherche minutieuse des symboles, qui nécessite d'observer les scans à une échelle très fine. Peut-être cette étape de recherche serait-elle raccourcie et simplifiée si on avait à disposition les cartes papiers (vue plus globale, pas d'attente de chargement du raster de telle ou telle zone...).

Certes, le temps de traitement est également dû au fait qu'il s'agissait de la première clarification, et au caractère incomplet de la couche, ce qui ne sera pas le cas dans le cadre des couches ROUTES ou BATI_INDIFFERENCIE par exemple. Mais ces couches possèdent beaucoup plus d'objets que notre couche de test (respectivement 16 866 et 24 772, contre 25 objets à l'origine pour PAI_RELIGIEUX).

Ce temps de traitement très long nous a donc décidés à mettre de côté cette tâche de clarification de la BDTOPO, quitte à la reprendre, la méthodologie étant déjà mise en place, si son usage nous intéresse par la suite. En revanche, cette expérimentation a soulevé des problèmes non encore résolus que nous exposons ici.

3.2. Problèmes et questions ouvertes

Lors de la création d'objets, il arrive qu'on se retrouve face à un objet manquant dans une classe (par exemple ponctuelle), mais présent sous une autre forme dans une seconde (par exemple surfacique) : c'est parfois le cas pour la classe PAI_RELIGIEUX, où un objet n'existe pas mais est présent dans la classe BATI_REMARQUABLE. Dans ce cas on décide de ne pas recréer l'objet ponctuel correspondant à l'objet surfacique, considérant qu'il n'est pas utile de créer des doublons supplémentaires.

Au sujet des changements d'attributs :

Les changements de valeur du champ « importance » sont souvent difficiles voire impossibles à déterminer. Pour de nombreux objets qui possèdent cet attribut le document de description de la BDTOPO ne donne aucun élément sur lequel s'appuyer pour remplir ce champ : de ce fait les changements éventuels d'importance ne seront quasiment jamais enregistrés (c'est le cas par exemple de toutes les classes PAI_x)

Pour d'autres objets l'importance est largement décrite et documentée : c'est le cas de la classe ROUTE par exemple. Dans ce cas on s'efforcera d'effectuer les modifications d'importance dès lors qu'elles sont identifiables sans ambiguïté à partir de nos sources (changement identifiable de largeur du trait par exemple).

Pour les objets créés, lorsque l'importance est impossible à déterminer pour les raisons exposées ci-dessus, on laisse le champ en « NR ».

Par ailleurs, ce type de traitement très précis, qui demande à ce qu'on s'intéresse de très près à

chaque objet, même lorsque celui-ci n'a pas de changement apparent entre les deux dates, allonge beaucoup le temps de traitement, déjà long (trois heures pour 65 objets, pour environ 1/3 de la zone soit 180 km²).

Cas des objets absents de la BDTOPO :

Le traitement des 20 objets déjà inscrits dans la BDD a été assez rapide. Cependant, on a remarqué que de très nombreux objets existent sur les scans datant de 2008 ou 2009, mais sont absents de la BDTOPO. L'hypothèse que ceci était dû à l'écart entre les scans et la BDTOPO, et que les objets ont bel et bien disparus en 2011 pouvait être émise. Cependant, vu la nature des objets en questions (croix ou calvaires, dans la baie du Mont-Saint-Michel) il semble peu probable qu'ils aient été détruits.

Pour vérifier leur existence ou non, nous avons utilisé plusieurs sources externes :

- dans la base de données d'Open Street Map (<http://www.openstreetmap.org/>), accessible en ligne, aucun calvaire ni croix n'est relevé ;
- dans la base de données photographique d'Open Street View, (<http://openstreetview.org/>) accessible en ligne, aucune photo de la zone n'est disponible ;
- dans la base de données photographique de Google Street View (<http://maps.google.fr/>), accessible en ligne, nous avons pu trouver un ou deux de ces objets ambigus (mais beaucoup de zones, hors agglomération, hameaux et villages ne sont pas couvertes).

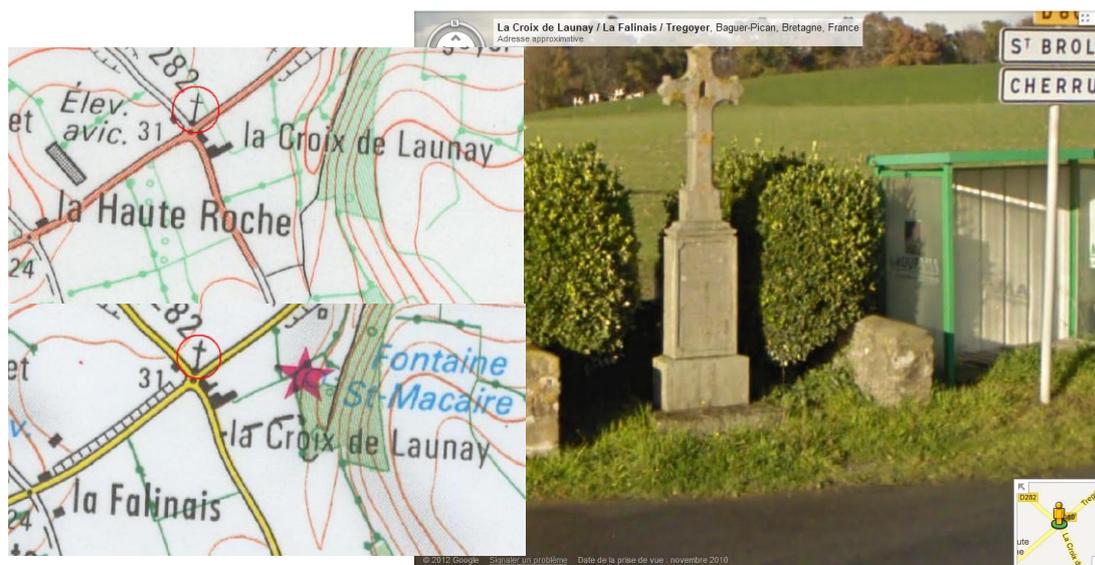


Figure 4. Croix présente en 1976 (carte haut), 2008 (carte bas) et 2010 (photo Google Street View), mais absente de la BDTOPO 2011.

Il semble donc probable que ces objets existent toujours et cette absence reflète un choix de spécification différent de la BDTOPO 2011. En effet, les calvaires ne sont pas sélectionnés pour les mises à jour car ils ne représentent pas une information prioritaire à l'heure actuelle, alors que sous Cassini, ces calvaires étaient une information prioritaire puisqu'ils permettaient aux armées de se repérer sur les chemins¹.

¹ Cenat J.-P., 2008. Stratégie, logistique et propagande : l'usage des cartes militaires par Chamlay. *Le Monde des cartes*, (195), pp.27-37.

4. Expérimentation de méthodologie allégée

Il s'agit ici de tester une clarification plus simple, sur une classe contenant plus d'objets.

Si cette méthodologie allégée s'avère efficace (temps de traitement plus court), nous pourrions envisager de l'appliquer ultérieurement à tout ou partie des classes initialement sélectionnées.

4.1. Sources de données

Les données restent les mêmes : scans géoréférencés au 1 :25 000 et BDTOPO 2011.

Cependant, nous allons effectuer nos tests sur une seule classe d'objet : la classe BATI_INDIFFERENCIE.

Nous avons choisi cette classe tout d'abord car elle est thématiquement importante dans notre étude. Nous formulons aussi l'hypothèse que cette classe est représentée de manière plus complète que la classe PAI_RELIGIEUX dans la BDTOPO.

D'autre part, il s'agit d'une classe surfacique, ce qui implique un traitement différent de la classe PAI_RELIGIEUX (objets ponctuels) et elle contient beaucoup d'objets (24 772 individus).

4.2. Méthodologie allégée

Ce mode opératoire est testé sur une seule classe (BATI_INDIFFERENCIE), sur le quart de feuille sud-est de la zone de Saint-Malo.

Il s'agit de ne pas enregistrer tous les changements affectant les objets entre les années 1960 et 2011, mais seulement les disparitions /apparitions de bâti.

4.2.1. Préparation de la base de données

Les étapes 1 à 5 de préparation de la base de données restent les mêmes (bien qu'elles ne concernent ici que la classe BATI_INDIFFERENCIE).

Les modifications de la méthode commencent à l'étape 6 :

1.

...

6. Changer le schéma de la base de données sous PostGreSQL pour rajouter trois colonnes : (GID, OID, attribut1, attribut2, ..., attribut_n, géométrie, **statut1960**, **commentaire1960**, **incertain**)

Code SQL pour l'insertion de ces quatre champs dans une table de nom "nomX" :

```
ALTER TABLE "NOMX " ADD COLUMN statut1960 integer NOT NULL DEFAULT -1;
ALTER TABLE TABLE "NOMX " ADD COLUMN commentaire1960 character varying(255) DEFAULT 'NR';
ALTER TABLE TABLE "NOMX " ADD COLUMN incertain boolean NOT NULL DEFAULT FALSE;
```

L'opération n° 7 est inchangée.

statut1960 vaut :

- 1 par défaut = la donnée n'a pas été comparée,
- 0 si l'entité est présente en 1960 et en 2011, sans changement,
- 1 si l'entité n'est pas présente en 1960 et est apparue en 2011 (création),
- 2 si l'entité est présente en 1960 et a disparue en 2011 (disparition).

commentaire1960 vaut :

- NR par défaut = le champ n'est pas utilisé,
- « existe_<date> » si l'objet existe sur le scan <date> (années 2000) mais pas dans la BDTOPO2011,
- Tout autre commentaire pour lever une ambiguïté.

incertain vaut :

- FALSE par défaut = la donnée est sûre,
- TRUE si la donnée est incertaine.

4.2.2. Clarification

Comparer le contenu de cette BD avec la carte raster 1960 (placée en dessous sous QGIS), et procéder systématiquement par couche jusqu'au traitement de tous les objets vectoriels. S'aider des scans 2011.

1. Commencer par traiter les objets existant sur le scan 1960 et **n'apparaissant pas dans la BDTOPO 2011**.

Une partie de ces objets correspond aux objets **ayant disparus en 2011** (code 2), mais d'autres sont des objets existant bien sur les scans en 2011 (code 0 ou 1) mais **non renseignés dans la BDTOPO 2011** (cf. commentaire en 2.3.3).

Pour ces deux types d'objet :

- a. créer les objets (point, ligne ou polygone),
- b. leur attribuer un identifiant de type <CLASSE>_n, n variant de 0 à l'infini, défini par l'ordre des opérations dans le champ OID ; l'identifiant GID est géré automatiquement,
- c. leur attribuer la valeur 2 (objets absent sur le scan 2011 et présent en 1960), 1 (objet présent sur le scan 2011 et absent en 1960) ou 0 (objet présent sur les scans 2011 et 1960) dans le champ statut1960,
- d. si incertitude quant à la nature ou certains attributs, et pour tous ceux de code 0, changer le champ incertain en TRUE et remplir le champ commentaire1960.

Pour faciliter le traitement, on modifie la représentation des objets pour qu'ils apparaissent transparents lorsque le champ « statut1960 » est différent de -1.

Tous les objets qui viennent d'être créés disparaissent donc et, au fur et à mesure du traitement, les objets déjà comparés disparaissent à la vue de l'opérateur.

2. Traiter les objets qui **existent et sont stables entre 1960 et 2011** : leur attribuer la valeur 0 dans le champ statut1960.
3. Traiter les objets **qui n'existent pas en 1960** et existent en 2011 : leur attribuer la valeur 1 dans le champ statut1960.

4.3. Test et résultats

La BDTOPO 2011 a une meilleure précision en termes de coordonnées des objets que les scans des cartes topographiques au 1 :25000 (nombreuses distorsions géométriques pour améliorer la lisibilité de la carte). De ce fait, les objets vectoriels représentant le bâti et leur représentation sur les cartes de 1960 et 2011 sont souvent décalé (la carte de 2009 est néanmoins mieux calée avec la BDTOPO, dans la mesure où elle est issue de cette dernière). Il faut bien prendre en compte ce décalage pour le traitement, même s'il est parfois très fort :



Figure 5. A gauche : décalage entre la BDTOPO et le scan de 1969, à droite : le scan de 2009 et la BDTOPO sont parfaitement calés.

Les scans 2008 présentent plus ou moins le même décalage que les scans 1960 (ils sont le résultat d'actualisation de cartes plus anciennes). Ceci introduit une imperfection dans le traitement dans la mesure où le travail est beaucoup plus précis sur le territoire couvert par le scan 2009 (environ 70km²) que sur les zones couvertes par les scans 2008 (environ 510km²).

En 6h de travail, 535 objets ont été traités sur environ 1/4 de la zone (148km² de traités sur 580km² au total). Seule l'étape de création d'objets n'apparaissant pas dans la BDTOPO mais existant sur les scans (tant de 1960 que de 2011) a été effectuée, ainsi que le traitement de ces objets (attribution du code correspondant à leur statut, incertitude et commentaire si nécessaire).

Dans cette classe nous avons également trouvé de très nombreux objets apparaissant sur les scans de 2009 ou 2008, mais n'existant pas dans la BDTOPO 2011. Nous avons rejeté l'hypothèse de leur disparition entre ces deux dates dans la mesure où nous avons trouvé plusieurs de ces objets dans une source externe. Sur l'exemple suivant, nous avons utilisé les images satellites accessibles depuis Google Maps (<http://maps.google.fr/>) :

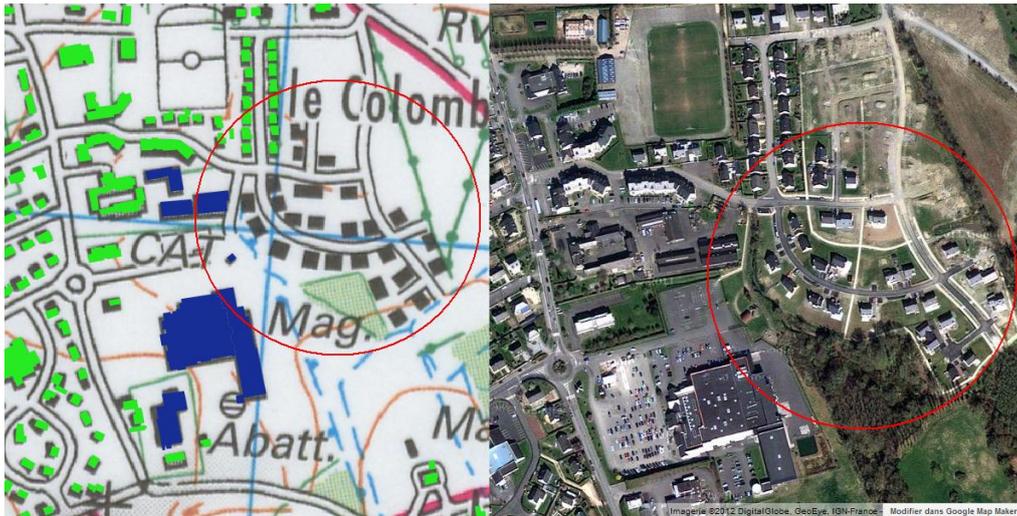


Figure 6. A gauche : ensemble de bâtiments non représentés dans la BDTOPO (BATI_INDIFEREBNCIE = en vert, BATI_INDUSTRIEL = en bleu), mais présent sur le scan 2008. A droite : même ensemble de bâtiments visibles sur une image satellite GeoEye datée de 2012 (source : Google Maps).

A partir de cette constatation, nous avons décidé de saisir tous ces objets, et pour conserver une unité méthodologique, nous avons traité de même les bâtiments isolés non renseignés dans la BDTOPO et présents sur les scans 2011.

Cette méthodologie allégée demande donc également un temps de traitement très long, du en partie au fait que la BDTOPO 2011 n'est pas complète. Cependant ce type de traitement est par définition très chronophage dans la mesure où il requiert une observation minutieuse des scans.

Nous pouvons donc constater également des problèmes de complétion de la BDTOPO pour ce thème, dont nous avons recherché l'origine. Cette constatation a été rapportée à François Lecordix, qui pilote le projet « Nouvelle carte de base » à l'IGN, projet qui vise à harmoniser le contenu des cartes avec celui de la BDTopo. Voici ses explications concernant le processus de production de la carte:

Il apparait que les différences entre la carte 2008 et la BDTopo sont en fait justifiées, parce que nous nous intéressons au bâti, qui a un processus de mise à jour particulier.

Pour les cartes, avant le projet « Nouvelle carte de base », dont les premières productions viennent de sortir cette année, il existait un processus spécifique pour le bâti où les bâtiments étaient insérés directement sur la carte à partir de photos aériennes au moment de la production de la carte, et n'ayant pas les informations sémantiques ni d'altitude, la BDTopo n'était alors pas mise à jour.

De plus, le bâti est mis à jour lentement dans la BDTopo (plus que pour les routes) car il faut attendre un vol aérien sur la zone (c'est ce qu'ils appellent la « restitution », et surtout la fusion avec la BD parcellaire). En revanche, les routes de la BDTopo sont elles toujours plus à jour que celles des cartes.

Concernant les photos satellitaires de GeoPortail, elles font l'objet d'un processus de renouvellement tous les 5 ans. C'est la raison pour laquelle seule la photo de GeoEye montre le nouveau lotissement.

A partir de début 2011, ces différences de mise à jour disparaîtront : sortie des premières données du projet de « la carte de base », qui cherche à rendre strictement conformant la BDTOPO avec les cartes produites, mais la production de cartes n'a été faite que sur 100 cartes de 60 km par 60 km.

Par ailleurs les métadonnées concernant la zone de St Malo et la couche de bâti utilisée indiquent ceci :

Actualité du thème Bâti :

- * 2005-06-08
- * 2009-06-02

Actualité des points d'activités et d'intérêts administratifs

- * 2009-06-10
- * 2010-07-01

Pour résumer les différences entre les cartes et la base de données sont expliquées. Il ne s'agit donc pas ici d'un problème de choix de représentation, comme cela était probablement le cas pour les calvaires présentés ci-dessus, mais de gestion des mises à jour.

Il s'agit donc de s'interroger, au regard des analyses prévues dans le projet, sur l'intérêt ou non de saisir ces objets absents dans la BDTOPPO, qui ne se justifierait que si on voulait s'assurer de disposer de données strictement aux mêmes dates sur toutes les zones, ce qui n'est probablement pas le cas dans le cadre du projet.

5. Conclusion

Après avoir testé deux méthodes de saisie assez fines, permettant d'établir des liens de changement entre l'époque 1960 et 2012 dans la BDTopo, il apparaît que l'examen et la confrontation de scans anciens et données nouvelles est chronophage, en particulier pour marquer les éléments qui existaient en 1960 et ont disparus à l'heure actuelle.

Il est donc nécessaire de choisir un compromis entre précision et rapidité de la saisie, compromis qui doit être fait en regard de l'utilisation prévue des données, et qui dépend donc des analyses des évolutions réalisées dans le projet.

Le seul processus de clarification raisonnable, relativement au temps de saisie, semble être le signalement des objets nouveaux en 2012 par rapport à 1960.