



UNIVERSITE D'ORLEANS

Observatoire des Sciences de l'Univers en région Centre

Licence Terre et Environnement

3<sup>ème</sup> année

2013-2014

***VECTORISATION ET INDEXATION DES CARTES  
PEDOLOGIQUES DE CONDOM ET D'ARLES AU  
1/100 000***

**Par Charles VIYER,**

Sous la direction de M. Laroche B, Mme Richer de Forges A.  
et M. Lehmann S.

Institut National de la Recherche Agronomique

2163, avenue de la pomme de pin

45075 ARDON



# TABLE DES MATIÈRES

<b>I. INTRODUCTION .....</b>	<b>3</b>
1. Présentation de l'INRA.....	3
2. Thème de ce stage.....	3
3. Maître de stage .....	5
<b>II. LE CONTEXTE .....</b>	<b>5</b>
<b>III. LA NUMÉRISATION DES CARTES PÉDOLOGIQUES AU 1/100 000.....</b>	<b>6</b>
1. Avant-propos .....	6
a. Présentation de la carte de Condom.....	6
b. Présentation de la carte d'Arles.....	7
2. Le géoréférencement.....	7
a. Exemple de mise en application .....	7
b. Le cas de Condom .....	10
c. Le cas d'Arles.....	11
3. La digitalisation.....	11
a. Digitalisation manuelle.....	11
b. Digitalisation à partir d'un calque préexistant .....	14
4. La vectorisation .....	14
a. Création des entités.....	14
b. Les erreurs générées .....	17
5. L'indexation.....	22
<b>IV. CONCLUSION .....</b>	<b>26</b>
Références bibliographiques .....	27
Annexe 1 .....	28
Annexe 2 .....	29
Annexe 3 .....	30
Table des illustrations .....	31
Remerciements.....	32

# I. INTRODUCTION

## 1. PRÉSENTATION DE L'INRA

Le centre INRA (Institut National de la Recherche Agronomique) Val de Loire, fort de ses recherches en biologie et gestion durable, de ses capacités expérimentales et de ses 961 agents est le sixième centre de l'institut par son effectif. Il mène des recherches autour de quatre grands pôles :

- Dynamique des sols et gestion de l'environnement
- Biologie intégrative des arbres et organismes associés
- Biologie intégrative animale et gestion durable des productions animales
- Santé animale et santé publique

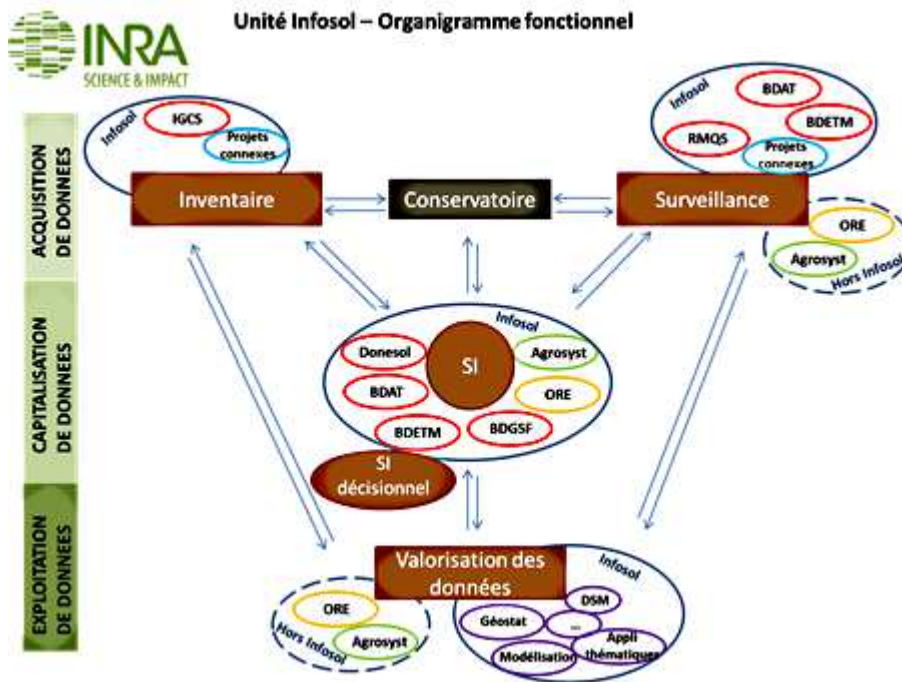
Les recherches menées par l'INRA font suite à l'évolution des questions que le monde scientifique se pose au sujet de l'alimentation, de l'environnement, de l'agriculture et de l'agronomie. Beaucoup d'enjeux, tels que le changement climatique ou encore la nutrition humaine, font de cet institut un pilier du développement sur le plan économique, scientifique et environnemental.

## 2. THÈME DE CE STAGE

Mon stage s'intègre dans le cadre des recherches sur la dynamique des sols et la gestion de l'environnement, et plus particulièrement au sein de l'unité de service InfoSol. Cette unité est chargée de cartographier et de constituer les bases de données sur les sols de France, et l'évolution de leur qualité. On y trouve notamment le conservatoire national d'échantillons de sols. InfoSol est structuré en quatre grandes équipes (figure 1) :

- "Surveillance" : en charge de la mise en place des programmes de surveillance de l'évolution de la qualité des sols ;
- "SI" ou "Système d'Information" : en charge de la gestion des bases de données sur les sols ;
- "Valorisation des données" : en charge des traitements de données pour la réalisation de diverses cartes thématiques ;
- "Inventaire" : en charge des programmes de cartographie des sols (programme "Inventaire, Gestion et Conservation des Sols (IGCS)").

Mon travail s'est effectué au sein de l'équipe Inventaire.



**Figure 1:** Organigramme fonctionnel de l'unité InfoSol.

Ainsi, j'ai utilisé ces bases de données et ces cartes pour de la vectorisation et de l'indexation au 1/100 000. Il m'a été demandé de travailler sur les cartes de Condom et d'Arles (figure 2). La carte de Condom se situe principalement dans le département du Gers en région Midi-Pyrénées. Elle recouvre également une petite zone au sud du département du Lot-et-Garonne et à l'est du département des Landes. La carte d'Arles, quant à elle, se situe principalement sur le département des Bouches-du-Rhône en région Provence-Alpes-Côte-d'Azur avec une petite zone sur le département du Gard.



**Figure 2:** Localisation des coupures de Condom (en vert) et d'Arles (en bleu).

Pour ce faire, le plan de ce rapport suit le protocole qui a été décidé au début du stage. Ainsi, les étapes nécessaires sont donc le géoréférencement, la digitalisation, la vectorisation et l'indexation d'une carte. Les outils nécessaires sont surtout des extensions d'ArcGis, c'est-à-dire ArcMap, ArcCatalog et ArcScan.

### 3. MAÎTRE DE STAGE

Mon maître de stage a été Monsieur Bertrand Laroche, pédologue-cartographe et responsable du programme IGCS. C'est d'ailleurs par son intermédiaire qu'il m'a été donné d'intégrer l'INRA Centre Val de Loire. Monsieur Laroche travaille pour l'Unité InfoSol, sous la direction de Madame Marion Bardy. J'ai par ailleurs été accueilli et épaulé par Madame Anne Richer de Forges, pédologue-cartographe et responsable du programme Connaissance Pédologique de la France (CPF).

Dans le cadre de mon travail de vectorisation, j'ai eu l'appui de Monsieur Sébastien Lehmann, géomaticien à l'INRA.

## **II. LE CONTEXTE**

Le travail fait partie du programme CPF (Connaissance Pédologique de la France) qui est en fait un sous-programme du programme IGCS (Inventaire, Gestion et Conservation des Sols).

Le programme IGCS a pour but de constituer des bases de données sur les sols et leur répartition géographique. Il est axé autour de quatre points :

- Identification, définition, et localisation des types de sols
- Elaboration de documents cartographiques
- Production de données de qualité
- Valorisation par des thématiques

Les données sémantiques des cartographies de sols à différentes échelles sont stockées dans une base de données nationale à structure unique appelée DoneSol. Celle-ci contient toutes les données ponctuelles (description de profils de sols) et surfaciques (description des unités cartographiques de sols) de toutes les cartographies de sols effectuées sur le territoire français. Les données issues du programme CPF ne font pas exception. Ce programme concerne toutes les cartographies de sols réalisées à moyennes échelles (du 1/50 000 au 1/100 000).

D'une manière générale, le volet CPF a trois objectifs principaux :

1. Etablir des lois de répartitions des sols sur la base de leurs facteurs de formation
2. Réaliser des cartes et des bases de données à moyennes échelles (1/50 000 au 1/100 000)
3. Mettre à disposition des connaissances fondamentales sur des travaux de généralisation, à l'échelle régionale ou nationale.

Il est important pour l'utilisation des données que celles-ci soient informatisées tant au niveau sémantique qu'au niveau graphique. Or, parmi les anciennes cartes publiées, peu ont été informatisées. C'est une des missions d'InfoSol avec l'objectif de mettre à disposition ces données en vue de la réalisation de diverses cartes thématiques.

Parmi les cartes au 1/100 000 non informatisées, se trouvent les cartes de Condom et d'Arles. J'ai donc eu à charge de vectoriser et d'indexer ces deux cartes pédologiques au 1/100 000, afin qu'elles rentrent dans les bases de données du programme CPF.

### **III. LA NUMÉRISATION DES CARTES PÉDOLOGIQUES AU 1/100 000**

#### **1. AVANT-PROPOS**

##### *a. Présentation de la carte de Condom*

La carte pédologique au 1/100 000 de Condom (Séguy<sup>1</sup>, 1973) occupe une situation relativement centrale dans le Sud-Ouest de la France, à la limite des régions Midi-Pyrénées -avec le département du Gers- et Aquitaine –avec les départements du Lot-et-Garonne et des Landes-. Cette situation géographique est représentée sur l'annexe 1.

Elle a été établie par la Compagnie d'Aménagement des Coteaux de Gascogne (CACG), pour le compte du Service d'Etudes des Sols et de la Carte Pédologique de France de l'INRA, sur fond topographique au 1/100 000 dressé par l'Institut Géographique National. Les échantillons prélevés ont fait l'objet d'analyses courantes effectuées au laboratoire agronomique de la CACG, sous la responsabilité de P. Divoux, chef du Service des Etudes du Milieu à la CACG. Les travaux d'interprétation et de simplification des études au 1/50 000 ont été effectués par les ingénieurs pédologues de la CACG.

Sur les 235 000 hectares cartographiés, 200 000 correspondent à la réduction et la simplification de la carte des sols au 1/50 000 du périmètre de la CACG, et 35 000 hectares situés dans l'angle Nord-Ouest et Sud-Ouest ont été levés en 1969 à l'échelle du 1/100 000.

---

<sup>1</sup> J. Séguy est un ingénieur pédologue.

### *b. Présentation de la carte d'Arles*

Administrativement, la majeure partie de la coupure est rattachée au département des Bouches-du-Rhône. Seul l'angle Nord-Ouest de la carte, sur la rive droite du Petit Rhône, fait partie du département du Gard.

Géographiquement, la feuille au 1/100 000 d'Arles (Bouteyre<sup>2</sup>, 1994) concerne tout ou partie de quatre unités naturelles distinctes, visibles sur l'annexe 2 :

- Le Delta du Rhône, avec la totalité de la Camargue entre le Grand et le Petit Rhône
- Les Costières du Gard dans leur partie sud-est, entre Générac et Saint-Gilles
- La Crau, ancien delta de la Durance, dans sa quasi-totalité
- Les reliefs calcaires, ou tout du moins les parties occidentales des collines autour de l'Étang de Berre.

La décision de réaliser cette partie de la carte pédologique de France à l'échelle du 1/100 000 a été prise au cours des années 1970. Elle fait suite à des études entreprises par la Direction Départementale de l'Agriculture et de la Forêt des Bouches-du-Rhône concernant la Camargue et la Crau.

L'étude de la Camargue et des Costières du Gard a été confiée à la Compagnie Nationale d'Aménagement de la Région du Bas-Rhône Languedoc (B.R.L.). Celle de la Crau et des reliefs calcaires autour de l'étang de Berre a relevé de la Société du Canal de Provence et d'Aménagement de la Région Provençale (S.C.P.).

## 2. LE GÉORÉFÉRENCEMENT

### *a. Exemple de mise en application*

InfoSol utilise les outils SIG de la gamme ESRI. Le logiciel utilisé est ArcGis, et plus particulièrement ArcMap –avec l'appui de ArcCatalog. Ainsi, en amont de mon travail sur la carte de Condom, j'ai pu me formaliser à cet outil en réalisant un tutoriel qui est proposé à la formation sur l'utilisation de DoneSol dispensée par InfoSol, sur une autre carte, en Bourgogne, au 1/250 000. Le but de cet exercice a surtout été de me familiariser avec le géoréférencement et les divers outils que pouvaient présenter le logiciel ArcMap. Le but de

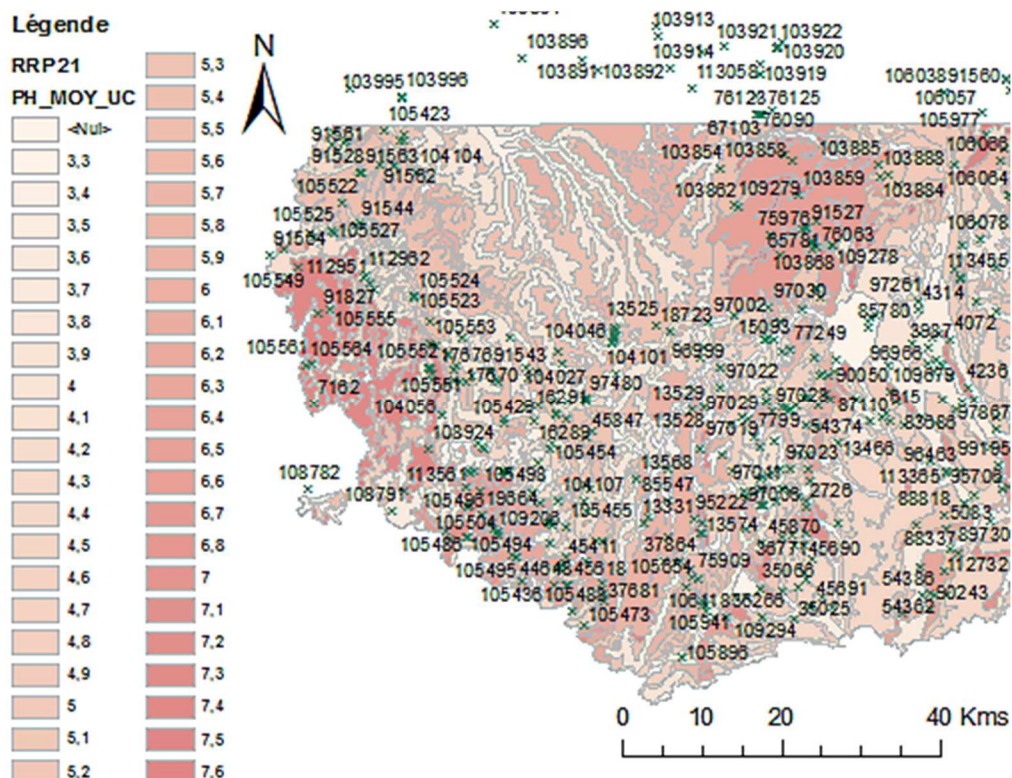
---

<sup>2</sup> G. Bouteyre était un ingénieur pédologue à la Compagnie Nationale d'Aménagement de la Région du Bas-Rhône. Pour la coupure d'Arles au 1/100 000, il a été assisté de G. Duclos, ingénieur pédologue à la Société du Canal de Provence et d'Aménagement de la Région Provençale.

l'exercice est de cartographier la valeur du pHeau, une des thématiques possibles mobilisant les données que je vais numériser dans le cadre de mon stage.

Il faut d'abord ouvrir dans ArcCatalog les données nécessaires à l'analyse de la carte. On peut nommer par exemple les données relatives aux pHeau. Ces-dernières ont été classées par pH croissant selon un code de couleur rouge, lui aussi croissant. Ceci est une étape de l'exercice, mais elle n'est pas obligatoire dans le cadre du géoréférencement puisque c'est une simple jointure attributaire.

Il faut ensuite lier une table de données à une couche de sol et cartographier une donnée issue de cette table. Cette table de données, appelée ici géodatabase, est nommée « RRP21 » dans le travail qui suit. On effectue ce qu'on appelle des jointures par rapport à une table attributaire. C'est au moment de la création de la classe entité de profils que l'on va intégrer les données de géoréférencement de notre carte. La table des profils de DoneSol, base de données comportant toutes les données relatives aux sols, ne contient pas les coordonnées. Elles sont dans une autre table qui permet leur projection sur une carte. Celle-ci est appelée « POI », pour « points of interest », en Lambert II étendu. La jointure entre les données de pH et les données « POI » est donc possible. (figure 3).



**Figure 3: Jointure entre les pHeau moyens et les données « POI ».**



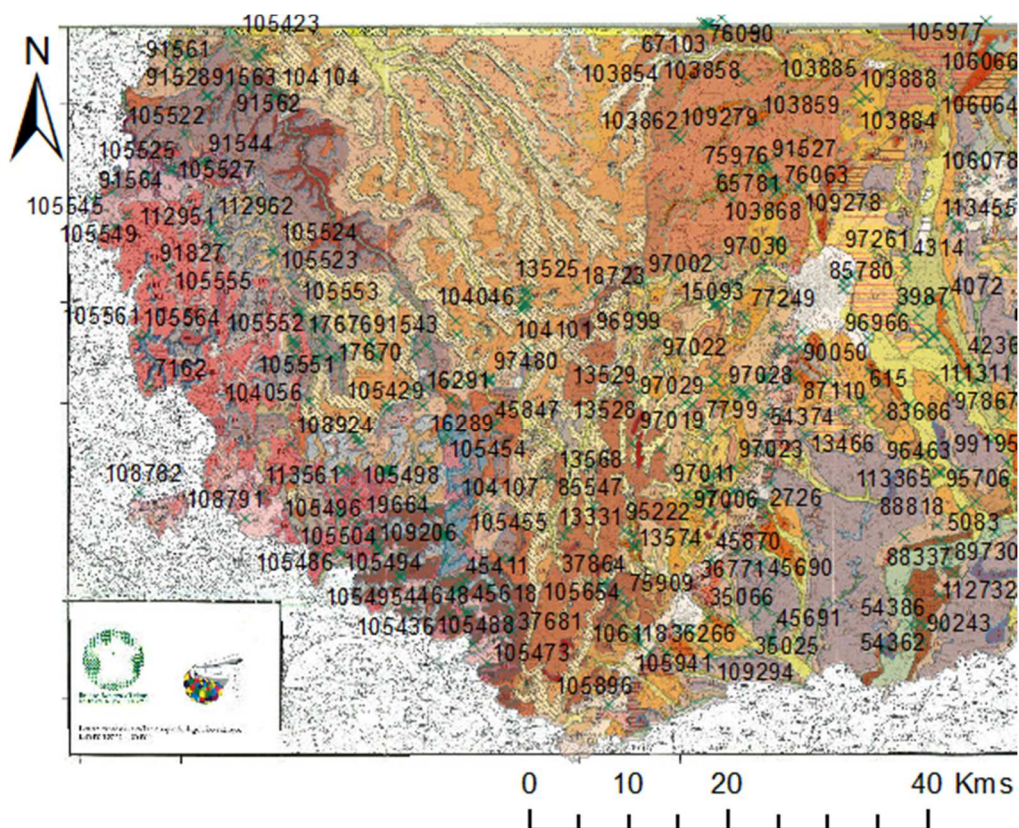
Ce sont des données pré-géolocalisées, que l'on retrouve par exemple dans certains GPS. Les coordonnées seront par la suite générées en Lambert II étendu.

Les données de géoréférencement seront donc en Lambert II étendu. Il s'agit dorénavant de géoréférencer à proprement parler notre carte. Une carte de répartition des pH a été ouverte et synchronisée avec notre carte par le géoréférencement. Cette manipulation s'intègre dans l'exercice de manière à aborder d'autres outils. Pour ce faire, le travail sera de choisir 4 points géographiques, aux quatre extrémités de la carte, et d'aller chercher leurs coordonnées sur le géoportail gouvernemental (<http://www.geoportail.gouv.fr/accueil>). Les données extraites sont dans le tableau Excel suivant, et se présentent en mètres. (tableau 1).

Ville	Moux-en-Morvan	Guillon	Bourberain	Longepierre
Coordonnées en X	737 575	732 170	822 733	818 637
Coordonnées en Y	2 242 927	2 280 729	2 282 111	2 219 150

**Tableau 1:** Coordonnées en mètres des quatre villes pour le géoréférencement de la carte de Bourgogne

J'ai effectué un calage en prenant 4 points géographiques, auxquels on applique leurs coordonnées en Lambert II étendu. En l'occurrence, deux cartes sont ouvertes : celle des pH et celle de la Bourgogne. En intégrant les coordonnées à la carte régionale, on a pu caler la carte des pH avec celle-ci. (figure 4).



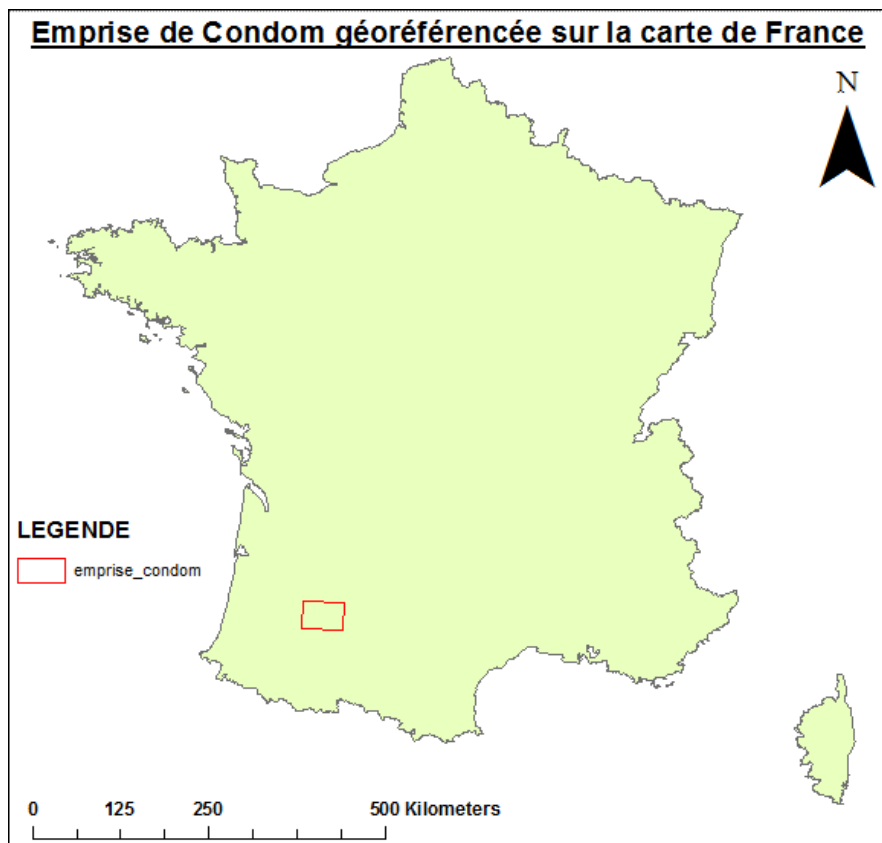
**Figure 4:** jointure des deux cartes après géoréférencement par calage

L'erreur d'approximation RMS, ou erreur quadratique moyenne, du fait qu'elle est la résultante des erreurs conjuguées de chaque ville géoréférencée, est de 112 mètres. Cette erreur qui, de prime abord, semble élevée, ne l'est en réalité pas tellement : au 1/250 000ème, elle représente une marge de 0,45 millimètres ; ce qui est acceptable. Bien entendu, si l'étude est pointue, comme nous le verrons avec les cas de Condom et d'Arles, l'erreur devra être minimisée au maximum.

### *b. Le cas de Condom*

Le travail de géoréférencement pour cette carte sera sensiblement différent. Comme dans l'exercice d'application, cela se fera par l'intermédiaire de deux cartes. Cependant, la différence est que nous allons utiliser une seule carte pour géoréférencer notre feuille de Condom.

Sous ArcCatalog, il s'agit dans un premier temps d'exporter les données de base. Ici, on a intégré le dossier nécessaire contenant toutes les données pour la manipulation. Nous avons besoin de notre premier cadre, l'emprise de Condom. Il se présente sous la forme d'un rectangle (figure 5). Sa particularité est qu'il est pré-géoréférencé, c'est-à-dire qu'il appartient déjà au système conique Lambert II étendu. Ceci est dû au fait que les cartes au 1/100 000 correspondent aux coupures IGN au 1/100 000.



**Figure 5:** Emprise de Condom géoréférencée sur la carte de France

La suite du géoréférencement ressemble trait pour trait à celle de l'exercice précédent. La seule différence est que cette fois-ci, les quatre points seront les quatre coins de la carte. Il n'y a aucunement besoin d'aller chercher des coordonnées sur le géoportail. On applique ensuite chacun de ces coins aux coins correspondants du cadre importé. Ce-dernier étant déjà géoréférencé, la carte de Condom le sera donc elle aussi.

Par ailleurs, ladite carte est d'autant plus utilisable pour la suite du travail que l'erreur RMS qui en ressort est de seulement 16,6 mètres, soit à peine 0,2 millimètre d'erreur, comme nous pouvons le voir sur l'annexe 3. Ceci est ainsi quasiment négligeable.

### *c. Le cas d'Arles*

De même que pour le géoréférencement de Condom, nous avons pu importer le cadre déjà géoréférencé, c'est-à-dire l'emprise de la carte d'Arles. Le système de projection n'a pas changé, c'est le Lambert II étendu.

Une fois l'emprise et notre carte reliées, cette-dernière se retrouve donc elle aussi géoréférencée. Comme nous l'avons vu pour Condom, la carte se retrouve légèrement inclinée sur ArcMap : c'est normal, la projection étant conique, elle déforme la carte.

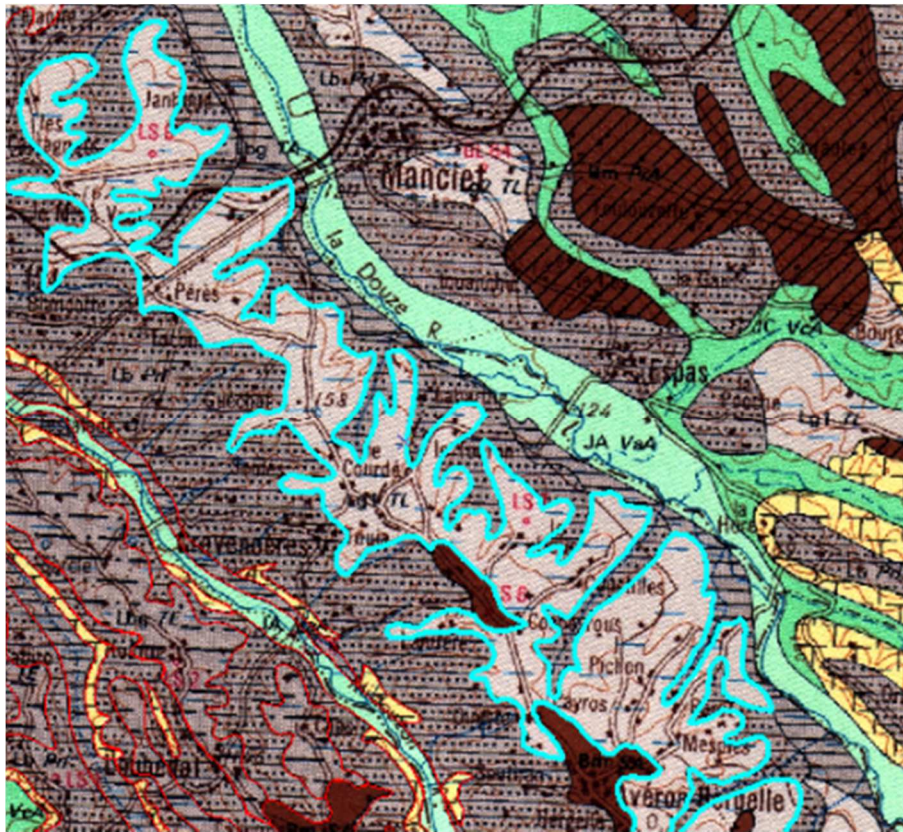
L'erreur RMS est de 34 mètres environ, ce qui demeure aussi raisonnable du fait qu'elle équivaut à une marge de 0,34 millimètre au 1/100 000.

## 3. LA DIGITALISATION

### *a. Digitalisation manuelle*

Notre géoréférencement maintenant effectué, nous pouvons commencer la saisie des polygones. Dorénavant, un fichier suivi est à remplir soigneusement : c'est le fichier que le pédologue-cartographe lira pour suivre l'avancement du travail effectué et connaître la qualité du rendu final qui lui est proposé.

Le principe est de dessiner le contour de chaque unité cartographique de sols (UCS), quelle qu'en soit la difficulté de digitalisation. Par exemple, sur la figure 6, on a à faire à un FLUVIOSOL, c'est-à-dire un sol qui occupe les lits mineurs et majeurs des rivières. Ainsi, l'entité est très ramifiée.



**Figure 6:** exemple de digitalisation d'un polygone pour la carte de Condom

On le voit, le tracé épouse d'une manière générale plutôt bien les courbes. Cependant, la sensibilité du tracé est réglable. Il y a différents modes de tracé possibles. Parmi eux, on compte par exemple le mode continu ou encore le mode par polygone automatique. Il s'exprime en pixels. Plus le nombre de pixels sera élevé, plus le curseur sera « attiré » vers les limites de couches. C'est ce qu'on appelle la tolérance de capture. Ce peut être un avantage, mais en réalité, plus le nombre de polygones voisins est élevé, plus le tracé deviendra fastidieux. Il convient dès lors de jouer avec cette sensibilité en fonction des conditions de digitalisation.

Pour appuyer notre travail, il est préférable d'avoir la carte papier à portée de main. En effet, la qualité de la carte scannée peut varier en fonction du grossissement qu'on lui affecte pour tracer les polygones. Dans certains cas, la lisibilité est médiocre, et il est bon de se référer à la carte papier pour s'assurer des unités en présence. A titre d'exemple, deux couches peuvent avoir la même couleur, mais la trame peut différer légèrement –elle peut passer de pois à des traits fins-. Cela peut être rendu très difficilement détectable à l'œil nu sur ordinateur, et la carte papier nous aide, parfois avec l'utilisation d'une loupe.

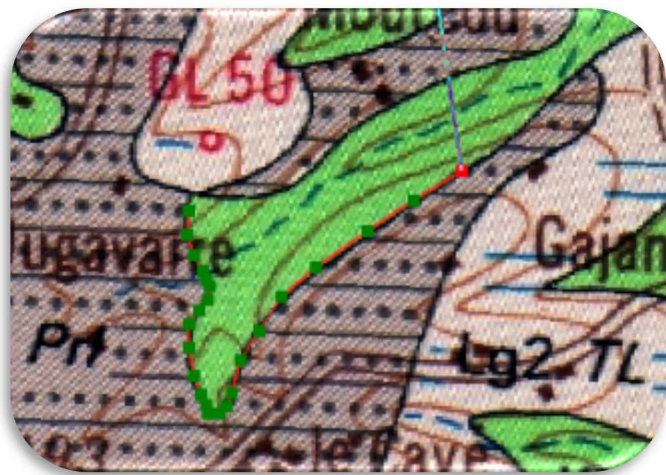
En ce qui concerne les difficultés de travail, certaines unités sont plus difficiles que d'autres à capturer. Un FLUVIOSOL aura, comme nous l'avons vu, un tracé ramifié. La digitalisation est ici rendue fastidieuse puisqu'il faut détourner tous les bords.

Lors de mon travail, j'ai pu constater que certains polygones de la carte étaient difficiles à fermer. En d'autres termes, pour l'indexation finale, tous les polygones doivent être convenablement fermés de manière à former des entités. Or ce n'est pas le cas : certains polygones se juxtaposent, et il est impossible de les fermer dans les règles d'ArcMap.

De ce fait, j'ai choisi de ne pas travailler avec des polygones mais des polylignes. La différence réside dans le fait que les polylignes ne forment pas encore des polygones à proprement parler dans le sens où ils ne se ferment pas forcément pour former des entités à part entière. A posteriori, il s'agira d'une manipulation qui permettra de transformer les polylignes formées en polygones. De cette façon, si toutes les lignes se rejoignent parfaitement, il n'y aura ainsi aucun souci de fermeture des polygones.

A cela, nous avons ajouté un cadre qui épouse parfaitement la carte. Le rôle de ce cadre sera de fermer les polygones ouverts. Les nœuds pendants, c'est-à-dire des morceaux de polylignes non fermés (cf. figure 11), ou tout du moins non rattachés au cadre, disparaissent automatiquement lors de la conversion des polylignes en polygones.

Le fait de saisir des polylignes et non des polygones nous obligent à un travail de précision, qui peut n'être que bénéfique pour le rendu final. En effet, il faut accentuer les zones de courbes avec de nombreux clics. (figure 7).



**Figure 7:** Précision du tracé des polylignes.

Nous traçons des lignes et, au sein d'une courbe, force est de constater qu'il en faut beaucoup pour que la courbe soit respectée –disons qu'il en faut en tout cas plus que dans une zone peu courbée-. On peut aussi zoomer assez fortement pour renforcer la précision dans certaines zones.

### *b. Digitalisation à partir d'un calque préexistant*

Un autre type de digitalisation sera utilisé pour la carte d'Arles : il n'y a aucune digitalisation à effectuer au préalable. En effet, le calque stable déjà digitalisé de notre feuille au 1/100 000 a été importé dans le dossier de travail.

Cependant, nous verrons par la suite qu'il ne s'agit pas que d'une simple vectorisation du raster : après vectorisation, certains polygones disparaîtront car ils sont mal fermés. Il conviendra donc de les reconstituer manuellement.

## 4. LA VECTORISATION

### *a. Création des entités*

Une fois ce travail de fabrication de polygones terminé, il s'agit de vectoriser l'ensemble sous ArcScan, c'est-à-dire de former les « vrais » polygones.

Pour prendre en main cette extension d'ArcMap, un travail a été effectué en amont. La première étape sous ArcScan est de seuiller la carte, c'est-à-dire d'enlever les 255 niveaux de gris pour n'en garder que deux, un blanc et un noir. C'est la méthode la plus simple de segmentation d'image. A partir d'une image en niveaux de gris, le seuillage permet de créer une image comportant uniquement deux valeurs, noir ou blanc. Ce système remplace un à un les pixels de l'image à l'aide d'une valeur seuil fixée. Si un pixel a une valeur supérieure au seuil, il sera blanc, et vice-versa.

Nous partons donc du TIFF de la carte d'Arles, et nous aboutissons à un fichier .grid, plus léger qu'un TIFF, et propre à ArcMap. Le fait d'avoir mis les valeurs dans seulement deux classes permet alors de restreindre de façon considérable le bruit pour la suite du travail. (figure 8).

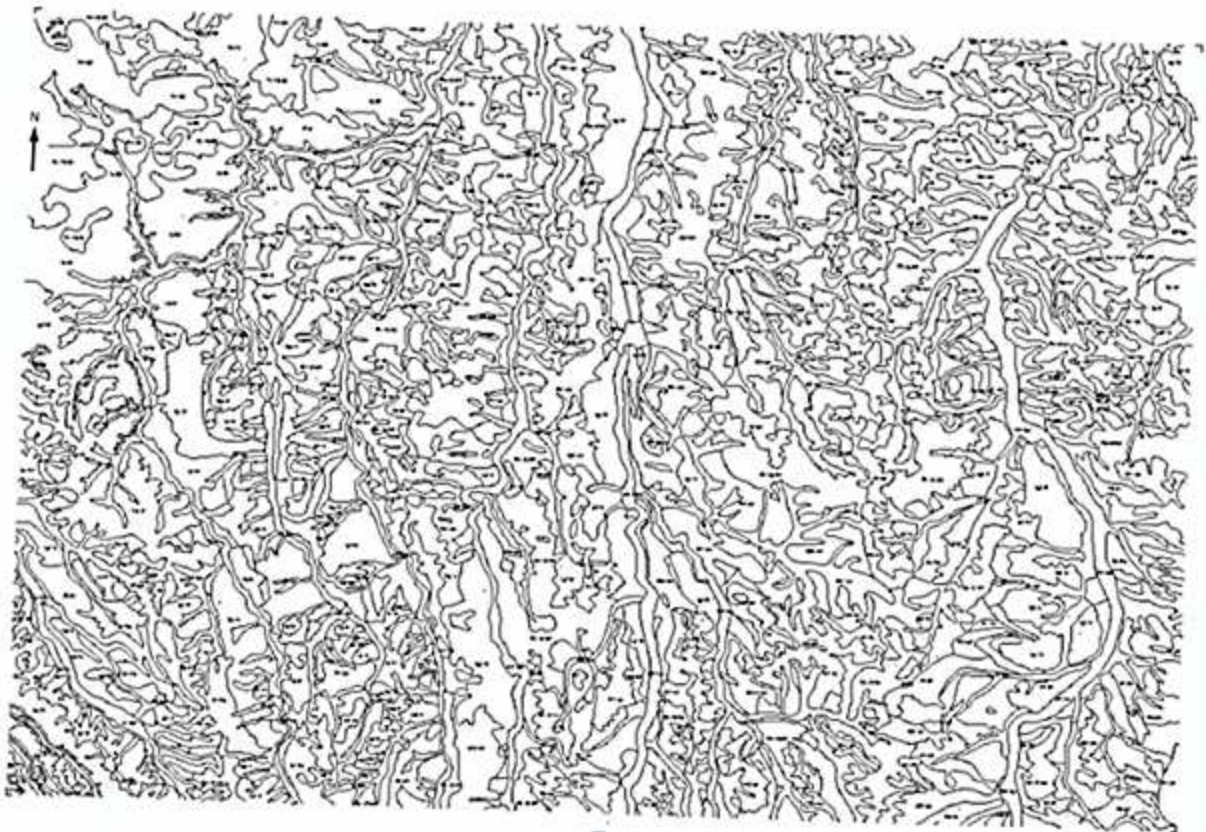
Dans mon cas, sur l'histogramme de luminance, j'ai placé la barre pour une valeur de seuil la plus proche de 255. Ainsi, nombreux seront les pixels qui seront noirs, puisque beaucoup de valeurs seront inférieures à 255. En somme, nous ne voulons que du noir et du blanc, sans aucun niveau de gris. On aura ainsi deux classes, 0 pour le noir et 1 pour le blanc. Ce faisant, il ne reste plus qu'à exporter le fichier au format .grid, très allégé. C'est le format de fichiers qui supportent un jeu de données Raster.



**Figure 8: Illustration d'un grid seuillé.**

Les étapes précédentes sont primordiales pour vectoriser une carte, surtout pour la suite des travaux à venir. Les étapes suivantes peuvent être listées comme suit :

- Paramétrer la vectorisation : on dicte quelles données raster peuvent être vectorisées et on spécifie les données vectorielles en sortie qui doivent être générées. Dans notre cas nous avons opté pour une largeur de vectorisation de 10, un bruit de 65 et une tolérance de compression de 0,1. Ces paramètres permettent d’effacer les « paquets », tout en gardant un maximum d’entités.
- Créer un fichier de formes polygones et générer les entités après avoir créé un aperçu. Ce fichier de type polyligne intègre le résultat de la vectorisation.
- Afin d’affiner notre raster, il convient de le nettoyer. L’utilisation d’une gomme permet de supprimer les artefacts les plus visibles.
- Pour optimiser le nettoyage, nous allons le paramétrer dorénavant de sorte qu’ArcScan sélectionne les cellules connectées selon une aire de 500 pixels. D’une manière générale, le nettoyage sera facilité par le fait que nous sommes en présence de vecteurs. Nous avons alors un premier jet de vecteurs retravaillés. (figure 9).
- Il faut ensuite ajuster le cadre si cela est nécessaire. Pour ce faire, on procède à un ajustement caoutchouté. Cependant, celui-ci est en mode polygones. Il faut faire un passage de polygones à polygones, et le copier-coller sur notre cadre.



**Figure 9: Entités créées sur la carte de Condom.**

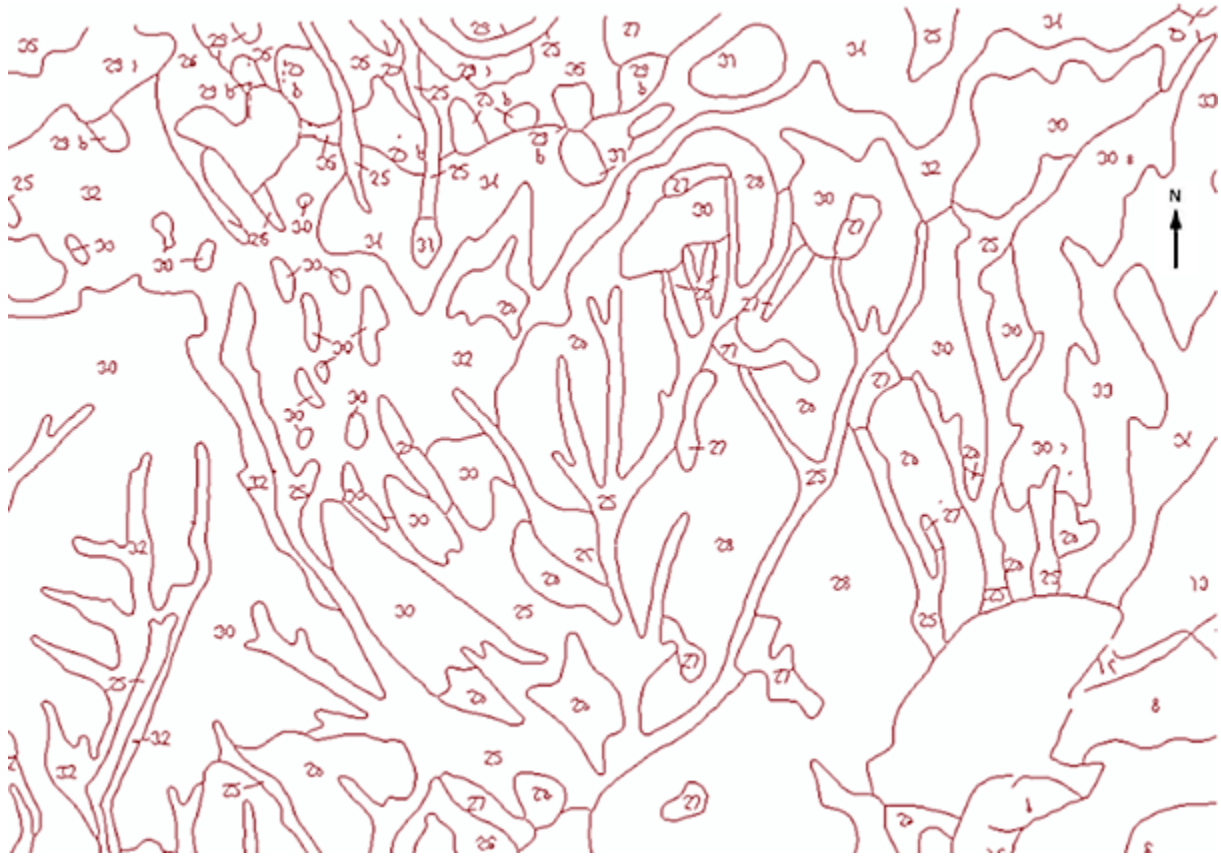
A cette étape, nous avons fait des travaux de géoréférencement, digitalisation, vectorisation et de nettoyage d'un raster ; lequel raster peut encore être affiné grâce à la topologie. C'est ce que nous verrons plus loin. Pour l'heure, voyons la suite des manipulations de vectorisation appliquées à la carte d'Arles au 1/100 000.

Dans un premier temps, on seuille la carte par les mêmes procédés employés que pour celle de Condom. De même, on se retrouve alors avec seulement deux classes, une pour le noir et une pour le blanc. Reste ensuite à créer une nouvelle base de données comportant des polygones dans un jeu de bases de données appelé communément « pedo ». Pour ce faire, on exporte notre carte en raster (.grid) ; en effet, la vectorisation n'est possible que sur un fichier raster. On peut désormais créer un aperçu pour se rendre compte de la vectorisation à venir.

Si l'aperçu ne nous satisfait pas, on peut toujours régler les paramètres de vectorisation. Il faut jouer avec ces paramètres de façon à effacer au maximum les impuretés



de la carte, tout en faisant attention à ne pas supprimer trop de lignes. Une fois ces travaux effectués, on peut générer nos entités. (figure 10).



**Figure 10:** Entités créées sur la carte d'Arles.

Le problème que la carte comporte encore trop d'éléments créant des gênes pour la suite, comme le numéro de légende au sein de la couche, nous amène à nettoyer le fichier vecteur.

#### *b. Les erreurs générées*

Le logiciel a travaillé avec des polygones. De ce fait, certaines lignes ne font pas la jonction avec le cadre, ou le dépassent. On appelle cela des "nœuds pendants" dans le cas où les branches ne sont pas correctement raccordées. (figure 11). Cependant, il peut y avoir d'autres artefacts à l'intérieur de la carte, notamment avec la présence de micro-polygones. Au départ, la légende pédologique est notée sur les couches de la carte. Une fois numérisée et retravaillée, il se peut qu'ArcMap transforme le 0 de la couche « Bl UO » par exemple en un micro-polygone. Une erreur qu'il convient d'éviter faute de quoi la suite du travail serait erronée.

Pour remédier à ce problème, on fait travailler l'extension « topologie », en chargeant les données pédologiques et topologiques que l'on a dans notre géodatabase initiale. Comme on le voit sur la figure 11 pour la carte de Condom, il en résulte des endroits où les jonctions ne sont pas faites : ce sont les nœuds pendants cités précédemment.

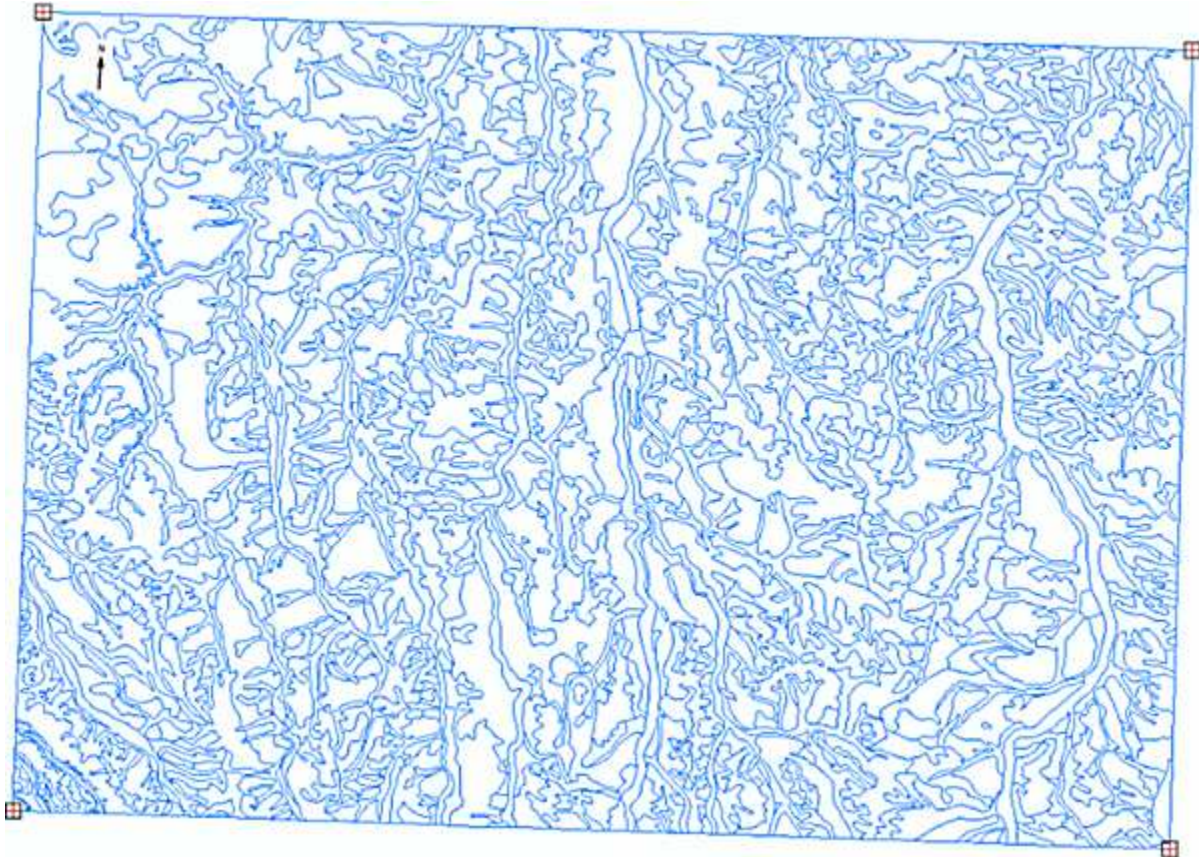


**Figure 11:** Exemple de nœud pendante sur la carte de Condom.

Pour le moment, le texte de légende est noté comme exception, nous verrons comment l'effacer une fois nos artefacts supprimés. On emploie pour ce faire l'outil « validation de la topologie dans l'étendue courante » qui permet de supprimer ces artefacts.

On repasse ensuite de polygones à polygones afin de corriger les derniers artefacts manuellement, y compris les légendes. On redécoupe certains polygones afin de faire en sorte qu'il n'y ait plus du tout de micro-polygones polluants la carte. En général, il ne reste qu'une poignée de micro-polygones car le logiciel a tout de même bien nettoyé la carte. Dans ceux-ci la plupart du temps, nous retrouvons des fragments de texte de légende (des

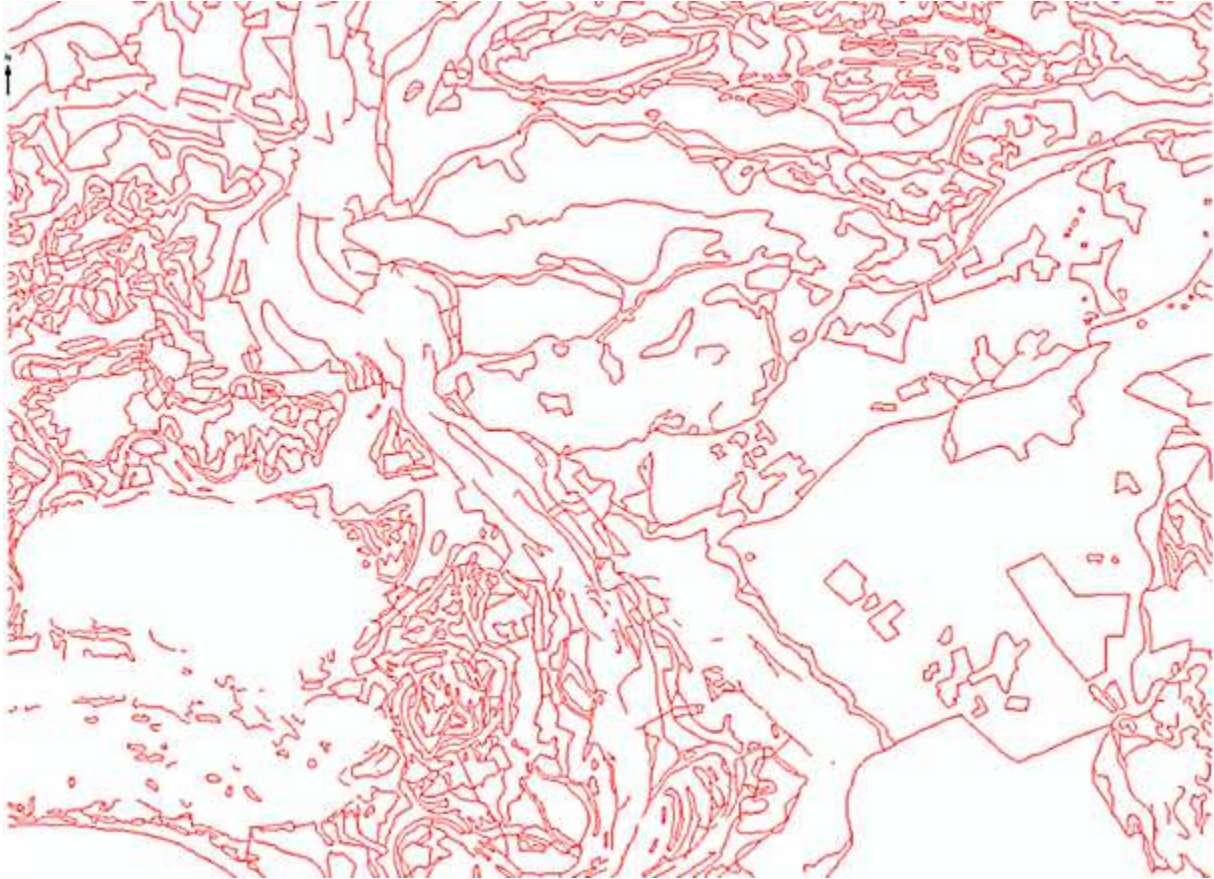
restes de 0 par exemple). Nous finalisons alors avec un nettoyage selon une sélection sous 9000 m<sup>2</sup> afin d'enlever les petits résidus. Le rendu, après vectorisation et nettoyage complet, est visible sur la figure 12. Il ne reste plus qu'à attribuer un numéro de label à chaque polygone.



**Figure 12:** Rendu final après réparation des erreurs sur la carte de Condom.

Voyons maintenant ce qu'il en est du nettoyage du raster d'Arles. Pour commencer, il s'agit de nettoyer la partie entourant la carte. Elle comporte les légendes et les cartes annexes d'études ; ladite légende sera refaite par la suite sous ArcMap. Les cartes ne nous intéressent pas dans le cadre de l'inventaire des cartes pédologiques. Ainsi, il suffit de sélectionner ces éléments et de les supprimer.

L'opération suivante consiste à nettoyer l'intérieur de la carte, et notamment les légendes du type « 3b » par exemple ; ainsi que les petites lignes qui polluent la carte. Le rendu est visible sur la figure 13. Cette manipulation se fait manuellement pour pallier à tout risque de pertes de données.



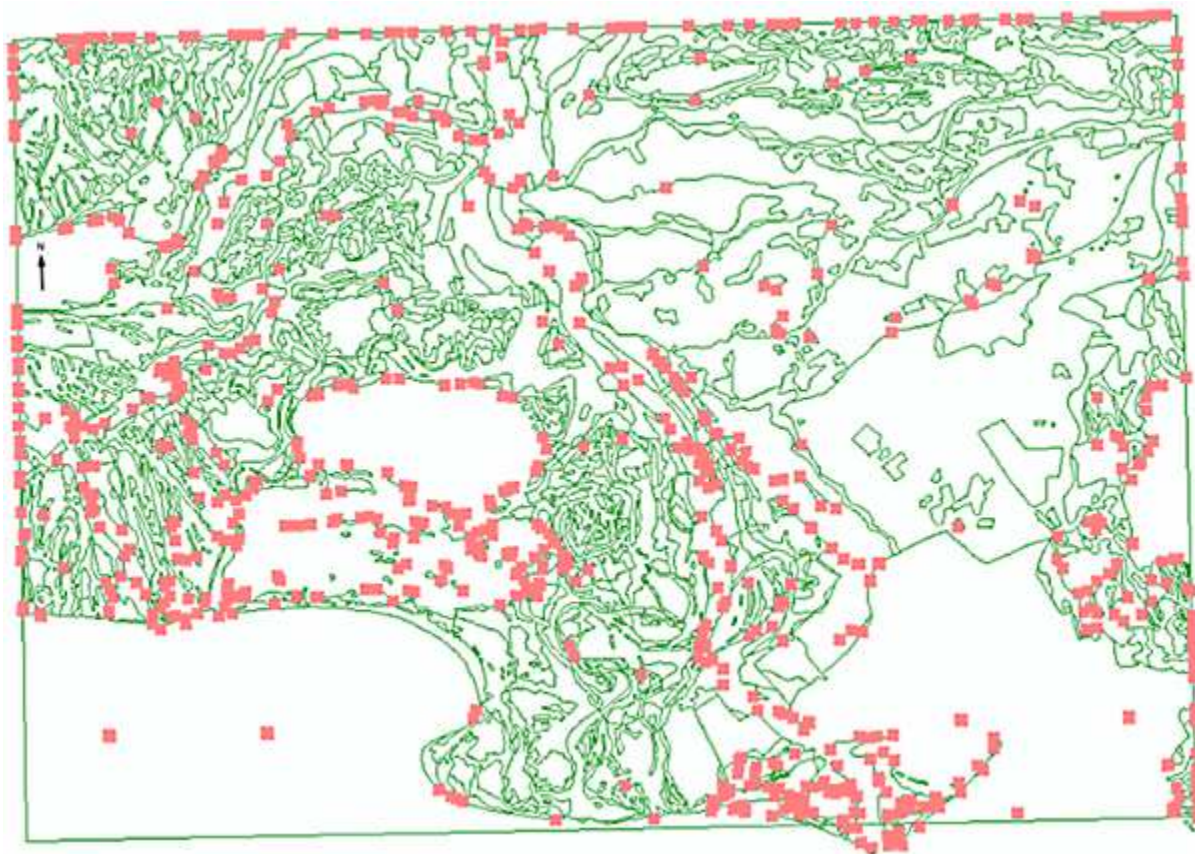
**Figure 13: nettoyage terminé de la carte d'Arles.**

Notre carte étant nettoyée, on peut importer la topologie pour aussitôt vérifier les erreurs.

Dans un premier temps, il faut exporter la couche contenant la carte nettoyée vers une géodatabase unique pour qu'elle se retrouve sélectionnable lors de la création de la topologie. En effet, c'est cette couche qui sera la source pour obtenir la topologie.

En vue d'un réel confort de travail, il est important de faire intervenir un cadre autour de notre carte. Lors de la numérisation, et après la vectorisation, le cadre de la carte papier a disparu. Comme nous allons faire travailler ArcMap avec la topologie de sorte qu'il n'autorise pas les nœuds pendants, il importe qu'un cadre vienne fermer les entités du bord de carte. De même, il faut exporter le rectangle d'emprise de la carte sous une géodatabase unique, la copier, puis la coller sur le cadre vectorisé.

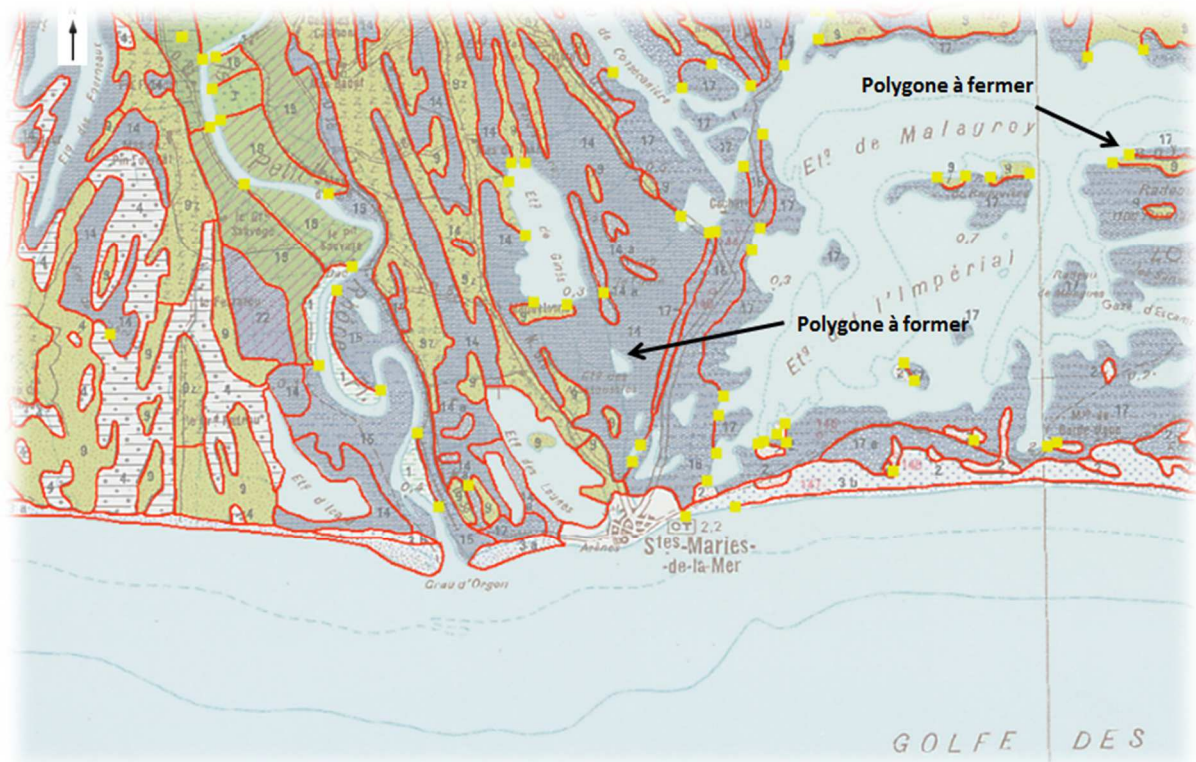
Nous pouvons à présent faire fonctionner l'outil topologie. Il met à jour les erreurs identifiées comme étant principalement des nœuds pendants, c'est-à-dire des tronçons de segments non reliés. Il s'agit principalement d'erreurs ponctuelles, dues à ces tronçons non rattachés. (figure 14)



**Figure 14:** erreurs dues à la topologie sur la carte d'Arles.

La tâche est désormais de réparer ces erreurs. Comme pour Condom, le principe est de fermer les polygones qui ne le seraient pas, et, soit de fermer les nœuds pendants, soit de les effacer. Cependant, cela ne suffira pas : lors du nettoyage, certains polygones pré-vectorisés ont été effacés. Peut-être les paramètres de vectorisation étaient-ils mal ajustés. Ainsi, il va falloir reformer ces entités, avec l'outil « créer des entités », et plus particulièrement les « lignes ».

Tous ces derniers constats sont regroupés dans la figure 15. On y retrouve des polygones à former, d'autres à fermer; ainsi que des nœuds pendants.



**Figure 15:** Polygones à former et à fermer sur la carte d'Arles.

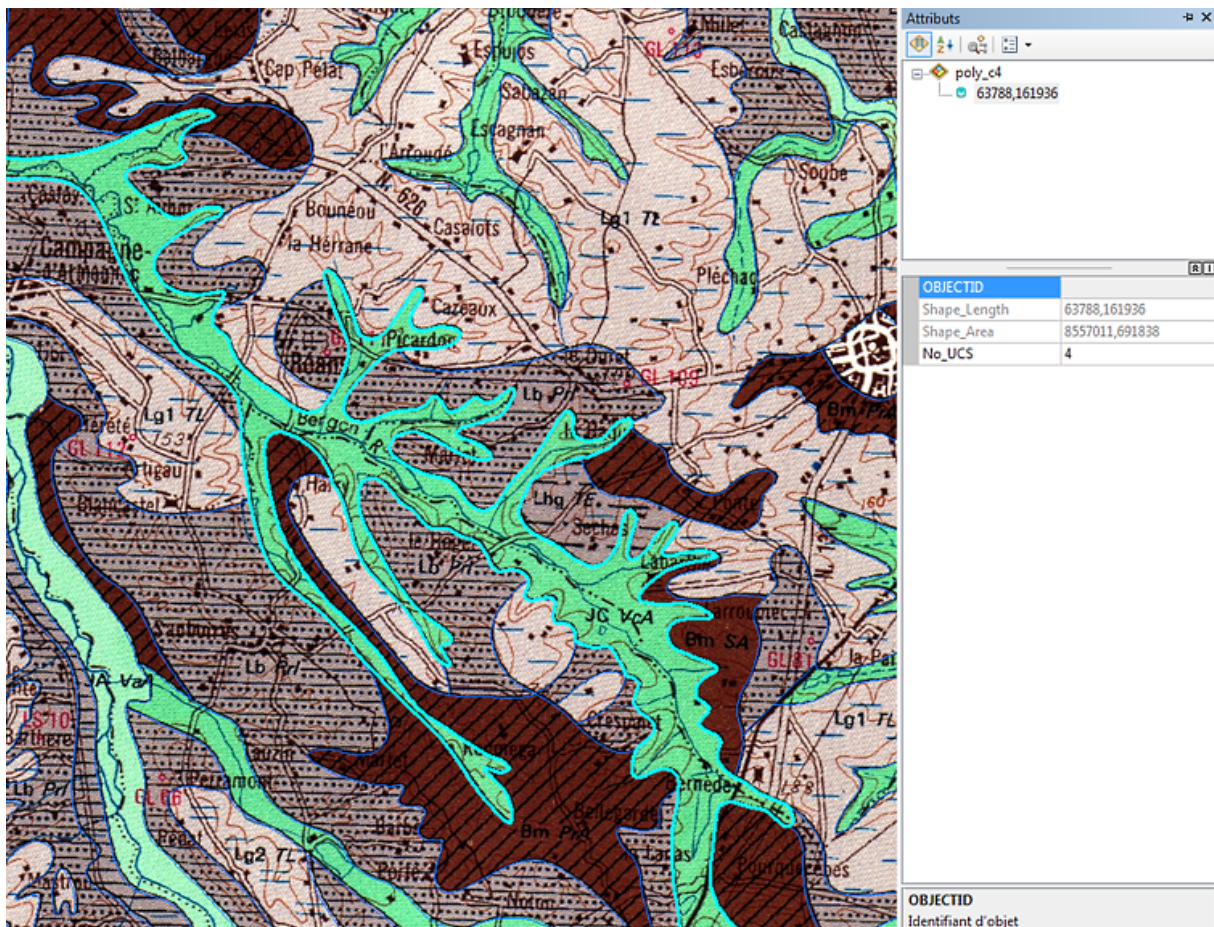
## 5. L'INDEXATION

Cette étape est importante pour mettre à jour et éclaircir la table attributaire de notre géodatabase de polygones. Maintenant que tous les polygones de la carte de Condom sont référencés, il s'agit de leur attribuer un numéro d'UCS (Unité Cartographique de Sols). Ce numéro est extrêmement important car il est le seul élément permettant de faire la liaison entre la couche graphique et la base sémantique de la carte. Par convention, ce doit être un entier. Or, sur les anciennes cartes, comme Condom ou Arles, les UCS ne sont pas toujours des entiers. Aussi, nous allons faire correspondre chaque légende de la carte avec un numéro UCS, de 1 à 22 (donc de type entier). Cette correspondance se trouve dans le tableau 2 ci-contre.

Légende carte	No_UCS
JAC VaA	1
JA VaA	2
JCC VcA	3
JC VcA	4
Crc MA	5
Crm KA	6
Cbm MA	7
Bm SA	8
BM PrA	9
BI SI	10
BI UO	11
Lb PrI	12
Lb TL	13
Lbg TA	14
Lbg TE	15
La SS	16
Lg1 TL	17
Lg2 TL	18
Phm SS	19
JC PrA	20
Cb+Cr MA	21
Pp+Pq SS	22

**Tableau 2:**  
correspondance  
entre la légende  
et le numéro UCS

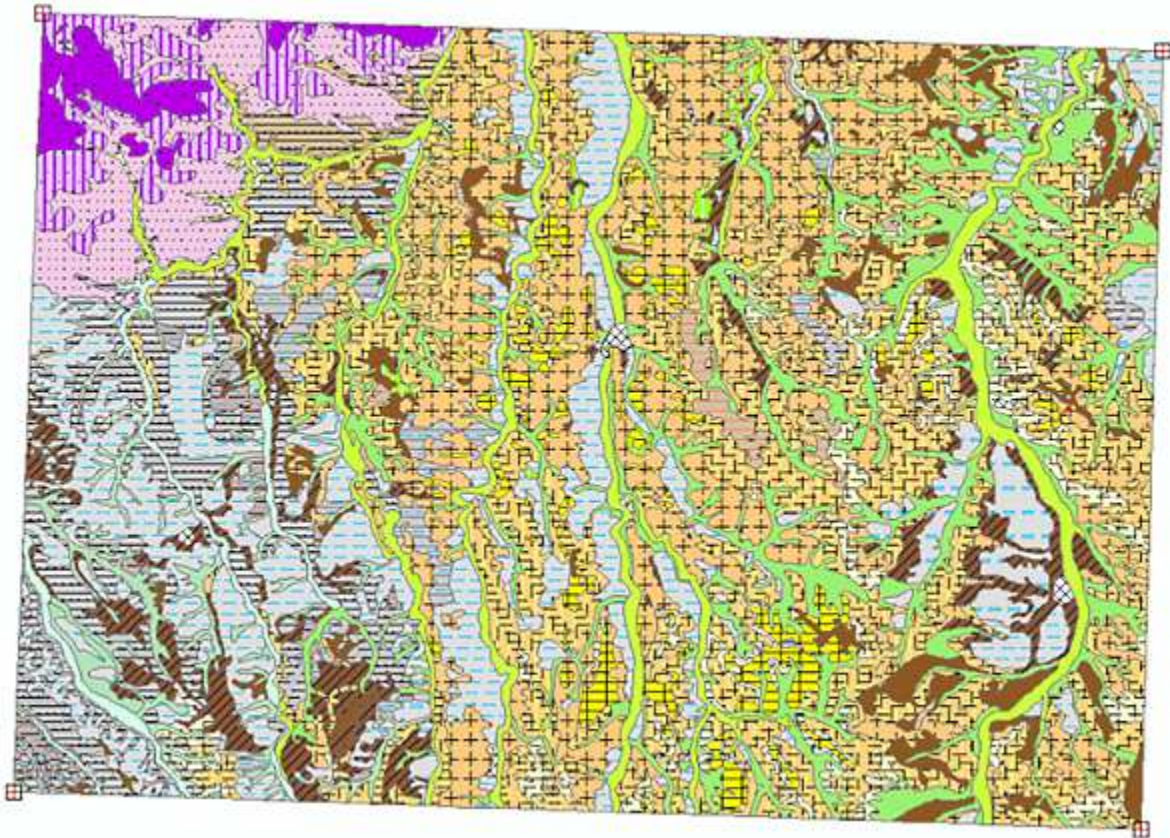
Une fois ce travail effectué, il s'agit de donner un attribut, c'est-à-dire le numéro d'UCS précédemment choisi, à chaque entité de notre carte. La figure 16 représente un polygone sélectionné entouré de bleu. En ouvrant la table des attributs, nous constatons que le champ « NO\_UCS » est vide, il faut donc le remplir par le numéro d'UCS correspondant à la couche. Ici, notre couche est un sol argilo-limoneux sur colluvions non calcaires, légendé "JC VcA". Si on se reporte au tableau de correspondance (tableau 2), "JC VcA" correspond au numéro UCS 4. Donc on rentrera le chiffre 4 dans cet attribut, et ainsi de suite pour les 1052 polygones constituant notre carte.



**Figure 16:** indexation des couches de la carte de Condornet.

La table attributaire contient seulement les propriétés physiques de nos polygones, c'est-à-dire leur aire et leur périmètre. Or, il est absolument nécessaire que les numéros d'UCS soient indiqués car ils permettent de faire le lien avec la base sémantique stockée sous DoneSol. Toujours dans un souci d'avoir un rendu final le plus soigné possible, nous pouvons représenter le plus fidèlement possible sous ArcMap la légende de la carte papier. Il s'agit d'importer dans les propriétés de la couche les différentes légendes, de 1 à 22 ; puis de

rentrer manuellement la trame de chaque légende papier. Cette manipulation reste malgré tout fort complexe du fait que certaines trames ne sont pas dans les bases de données de références de styles du logiciel. Une fois ce travail terminé, on obtient le rendu de la figure 17.

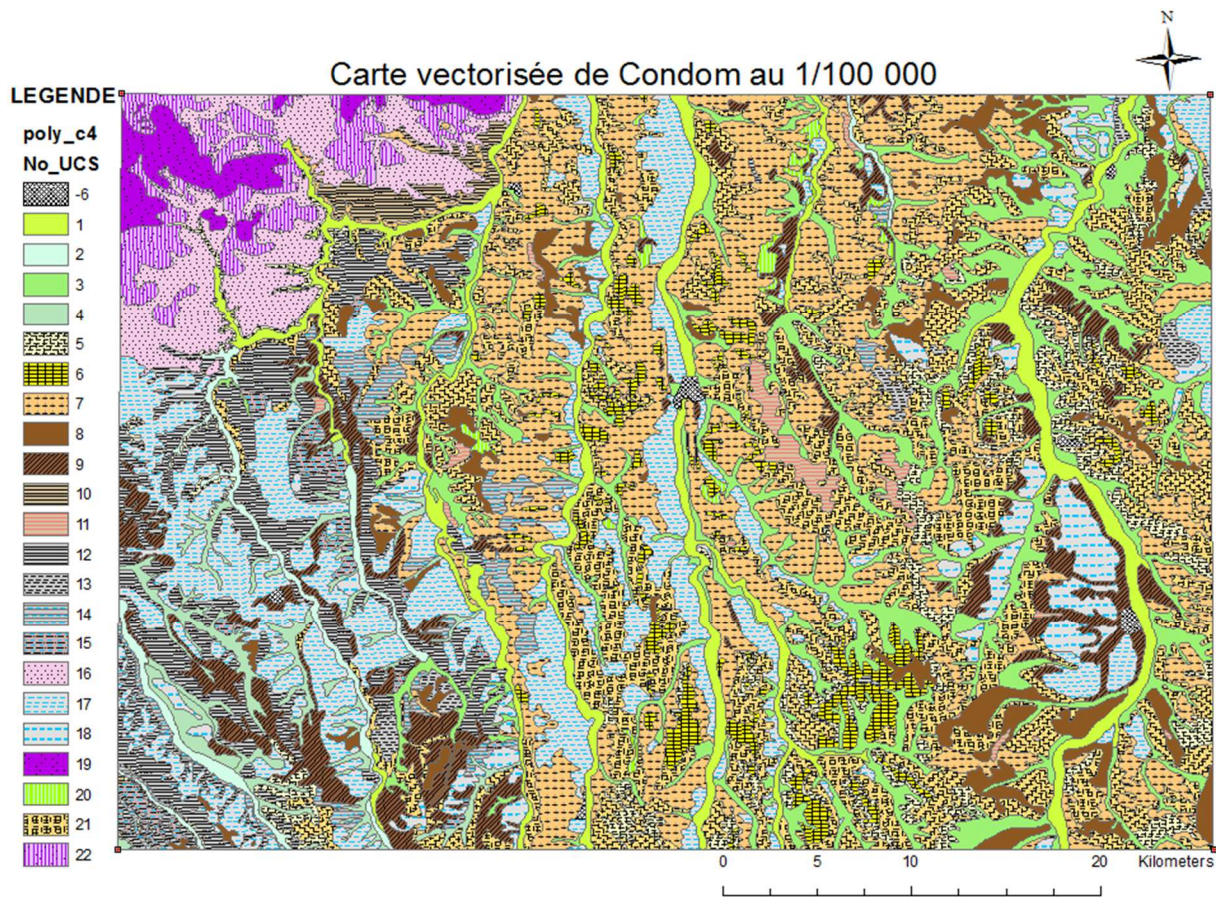


**Figure 17:** rendu après indexation des légendes de la carte de Condom.

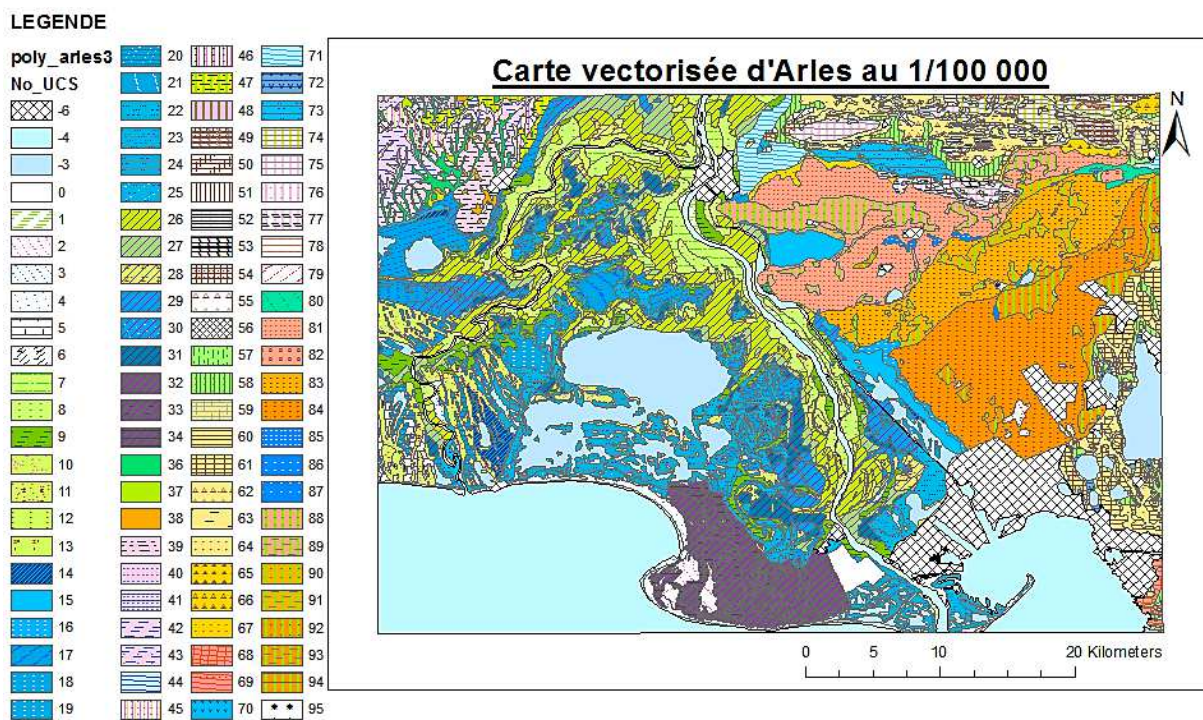
Le rendu final de nos travaux est représenté sur la figure 18. Une mise en page a été faite pour l'occasion, incluant un titre, une légende, une flèche d'orientation vers le Nord et une barre d'échelle. Une rotation de l'image a par ailleurs été effectuée afin d'en rendre confortable sa lecture. Celle-ci est due au géoréférencement initial. Notre système de projection étant conique, il est ainsi normal d'obtenir une légère rotation de la carte située à l'Ouest du méridien central.

Pour l'indexation de la carte d'Arles, nous avons à faire à des procédés similaires. Dans un premier temps, il convient d'attribuer un numéro UCS à chaque légende ; travail fastidieux s'il en est car générant pas moins de 95 légendes différentes. On assimile ensuite chaque polygone avec son numéro UCS correspondant. Une fois les 1720 polygones indexés, il ne reste plus qu'à établir une légende en tout point fidèle à la carte papier. (figure 19).





**Figure 18:** Carte pédologique de Condom au 1/100 000, vectorisée et indexée.



**Figure 19:** Carte pédologique d'Arles au 1/100 000, vectorisée et indexée.

## **IV. CONCLUSION**

Le but de ce stage était de vectoriser et d'indexer deux cartes pédologiques au 1/100 000, celles de Condom et d'Arles. Le cheminement à suivre pour arriver à un tel résultat est souvent le même pour ce genre de cartes. Il faut tout d'abord commencer par le géoréférencement du scan : sans celui-ci, le travail est inutile car on ne pourra replacer géographiquement cette carte dans son contexte. Arrive ensuite l'étape de digitalisation avec la création des polylignes, qui donneront les polygones au moment de la vectorisation. Enfin, il faut attribuer un label à chaque polygone, afin de pouvoir faire le lien entre les unités de sols représentées sur la carte et ceux de la légende ou de la base de données liée.

Ce travail étant achevé, ces deux cartes numérisées seront archivées dans le répertoire CPF de l'unité InfoSol. Maintenant que les données sont acquises et capitalisées, elles peuvent être valorisées par des thématiques.

# Références bibliographiques

---

INRA Orléans unité InfoSol. (2013). *Dictionnaire de données DoneSol version 3.4*.

Séguy, J. (1973). Carte Pédologique de la France à 1/100000 - Condom H21 - *Notice explicative*. (INRA, Éd.) . 152 p

<http://www.quae.com/fr/r608-carte-pedologique-de-france-a-1-100-000.html>

Bouteyre, G. (1994). Carte Pédologique de la France à 1/100 000 -Arles N22- *Notice explicative*. (INRA, Éd) . 302 p

<http://www.quae.com/fr/r534-carte-pedologique-de-france-a-1-100-000.html>

<http://www.geoportail.gouv.fr/accueil>. (s.d.).

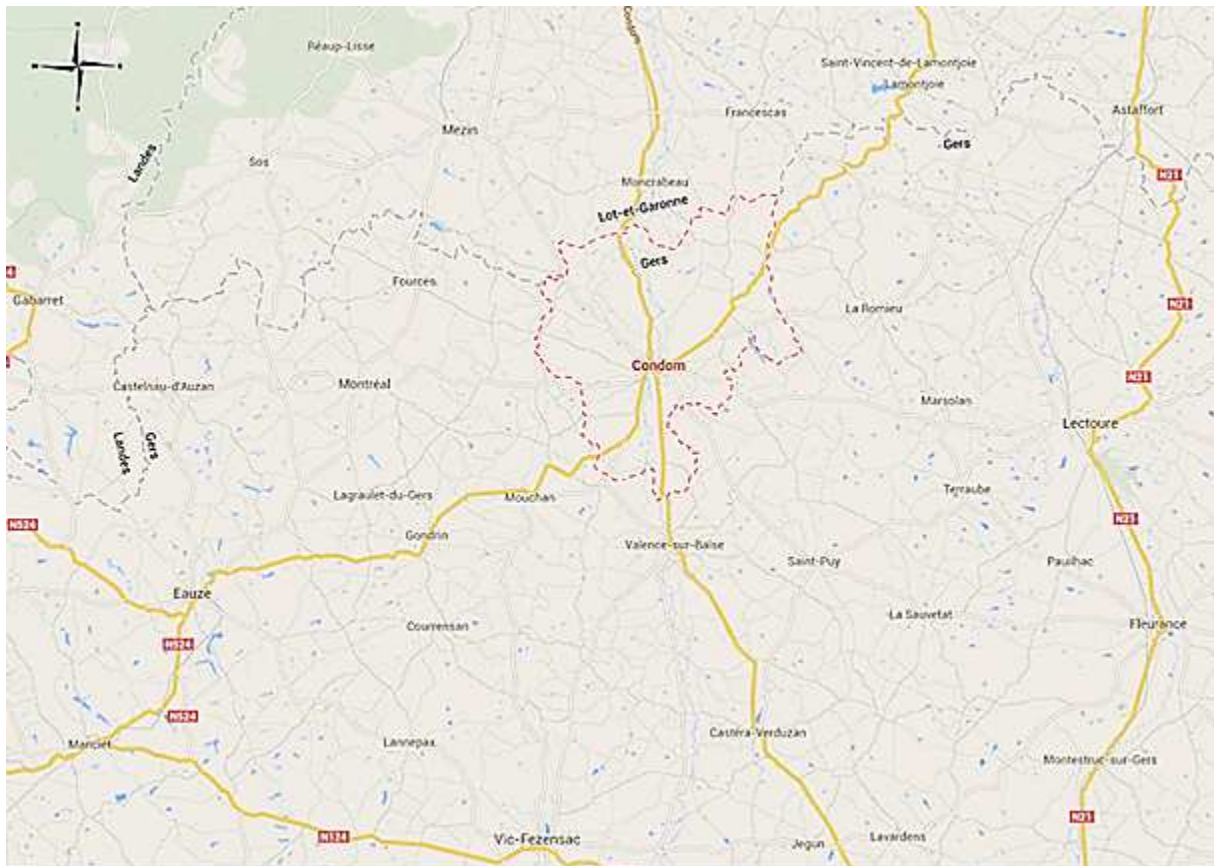
<https://fr.wikipedia.org/wiki/Arles>. (s.d.).

[https://fr.wikipedia.org/wiki/Condom\\_\(Gers\)](https://fr.wikipedia.org/wiki/Condom_(Gers)). (s.d.).

<https://www.google.fr/maps>. (s.d.).

# Annexe 1

---



**Annexe 1:** localisation géographique précise de la carte de Condom au 1/100 000.

# Annexe 2

---



**Annexe 2:** Localisation précise de la carte d'Arles au 1/100 000.

# Annexe 3

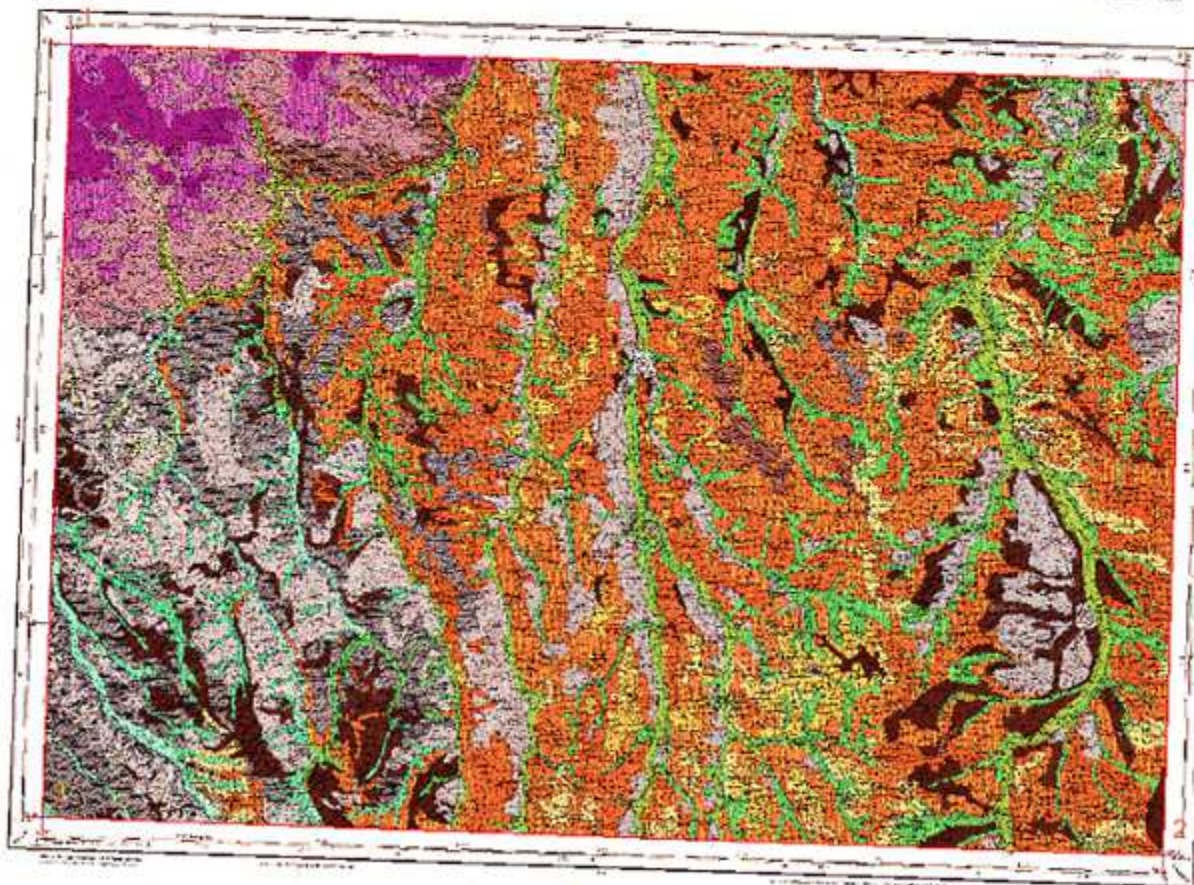
Lier

Erreur QM : Forward:16,5973

	Lier	Source X	Source Y	X Carte	Y Carte	Residual_x	Residual_y	Résiduel
<input checked="" type="checkbox"/>	1	2343,324101	-1376,077096	459578,288326	6338012,956146	15,8444	5,10081	16,6452
<input checked="" type="checkbox"/>	2	8021,607954	-5326,484099	516039,634939	6296098,649479	15,753	5,07137	16,5492
<input checked="" type="checkbox"/>	3	2327,000083	-5326,353263	458057,518758	6298039,924411	-15,7618	-5,0742	16,5584
<input checked="" type="checkbox"/>	4	8008,216902	-1374,010719	517195,553000	6336083,743824	-15,8356	-5,09797	16,636

Ajustement auto      Transformation : Transformation de 1er ordre (affine)

Degrés Minutes Secondes      Forward Residual Unit : Unknown



**Annexe 3: Géoréférencement terminé de la carte de Condom au 1/100 000.**

# Table des illustrations

---

Figure 1: Organigramme fonctionnel de l'unité InfoSol. ....	4
Figure 2: Localisation des coupures de Condom (en vert) et d'Arles (en bleu). ....	4
Figure 3: Jointure entre les pHeau moyens et les données « poi ». ....	8
Figure 4: Jointure des deux cartes après géoréférencement par calage.....	9
Figure 5: Emprise de Condom géoréférencée sur la carte de France.....	10
Figure 6: Exemple de digitalisation d'un polygone pour la carte de Condom.....	12
Figure 7: Précision du tracé des polygones. ....	13
Figure 8: Illustration d'un grid seuillé. ....	15
Figure 9: Entités créées sur la carte de Condom. ....	16
Figure 10: Entités créées sur la carte d'Arles. ....	17
Figure 11: Exemple de nœud pendant sur la carte de Condom. ....	18
Figure 12: Rendu final après réparation des erreurs sur la carte de Condom. ....	19
Figure 13: Nettoyage terminé de la carte d'Arles. ....	20
Figure 14: Erreurs dues à la topologie sur la carte d'Arles. ....	21
Figure 15: Polygones à former et à fermer sur la carte d'Arles. ....	22
Figure 16: Indexation des couches de la carte de Condom.....	23
Figure 17: Rendu après indexation des légendes de la carte de Condom. ....	24
Figure 18: Carte pédologique de Condom au 1/100 000, vectorisée et indexée. ....	25
Figure 19: Carte pédologique d'Arles au 1/100 000, vectorisée et indexée. ....	25
Annexe 1: Localisation géographique précise de la carte de Condom au 1/100 000. ....	28
Annexe 2: Localisation précise de la carte d'Arles au 1/100 000.....	29
Annexe 3: Géoréférencement terminé de la carte de Condom au 1/100 000.....	30
Tableau 1: Coordonnées en mètres des quatre villes pour le géoréférencement de la carte de Bourgogne.....	9
Tableau 2: Correspondance entre la légende de la carte et le numéro UCS attribué.....	22

# Remerciements

---

Je tiens à remercier dans un premier temps Monsieur Bertrand Laroche, pédologue-cartographe et responsable du programme IGCS, qui m'a permis d'intégrer l'institut dans la section Connaissance Pédologique de la France, du programme Inventaire, Gestion, et Conservation des Sols. Il m'a apporté toute son expérience, tant du point de vue pédologique que géomatique. Par ailleurs, je le remercie pour son encadrement professionnel et pour sa confiance dont j'ai bénéficié tout au long du stage.

D'autre part, je remercie tout particulièrement les pédologues qui ont m'ont fait partager leurs expériences et leurs intérêts, à savoir Madame Anne Richer de Forges, pédologue-cartographe et responsable du programme Connaissance Pédologique de France, et Monsieur Sébastien Lehmann, géomaticien au sein du même volet.

Enfin, je remercie Monsieur Max Vidal et Monsieur Yannick Branquet, professeurs à l'Observatoire des Sciences de l'Univers en région Centre, pour leur encadrement avant, et pendant le stage.



# **Vectorisation et indexation des cartes pédologiques de Condom et d'Arles au 1/100 000**

**Charles VIYER**

**Avril 2014**

**Mots-clefs :** INRA, IGCS, CPF, vectorisation, indexation, ArcMap, Condom, Arles.

Dans le cadre de ma 3<sup>ème</sup> année de Licence Terre et Environnement, j'ai effectué un stage d'un mois au sein de l'Institut National de Recherche Agronomique d'Orléans (INRA). C'est un institut menant des recherches sur les problèmes environnementaux : l'alimentation, l'agriculture, et l'agronomie. Ces recherches font de cet institut un pilier du développement scientifique du point de vue économique, scientifique, et environnemental.

L'INRA, dans le cadre du programme Inventaire, Gestion, et Conservation des Sols (IGCS), étudie et répertorie les sols de France. Des archives sont faites depuis des années afin de collecter un maximum d'informations sur tout le territoire français.

Mon stage s'est donc intégré dans la partie IGCS, et plus particulièrement dans le volet Connaissance Pédologique de France (CPF). Le but a été de vectoriser et d'indexer deux cartes pédologiques au 1/100 000, celles de Condom et d'Arles. Le logiciel moteur de ce travail a été ArcMap et ses extensions. Ces deux cartes numérisées ont été archivées avec les autres dans le répertoire CPF.

**Outils utilisés :** ArcMap, Excel, Word, Internet, cartes pédologiques au 1/100 000.

## **Encadrement :**

Bertrand LAROCHE (service InfoSol)

2163, avenue de la pomme de pin

45075 ARDON

[bertrand.laroche@orleans.inra.fr](mailto:bertrand.laroche@orleans.inra.fr)