

## RT 2.6

# Module 4 : Création des bases de données géographiques historiques

1. Présentation générale du processus .....	2
2. Pré-traitements de données.....	3
2.1 Création des fichiers shapes à différentes dates à partir des cartes topographiques papiers ou numériques – par analyse visuelle .....	3
2.2 Création des fichiers shapes à différentes dates à partir des photos – par analyse visuelle .....	3
3. Etape 1 : Chargement des données shapes .....	3
4. Etape 2 : Génération des identifiants ID_GEO pour la base de données de référence (la plus récente) .....	4
5. Etape 3 : Appariement des données géographiques .....	4
5.1 Choix type d'appariement .....	6
5.2 Définition des liens d'appariement et génération des identifiant ID_GEO.....	6
5.3 Visualisation des résultats d'appariement .....	8

## **1. Présentation générale du processus**

Les règles de simulation définies dans le module GeOpenSim sont construites à partir des données existantes. Afin de suivre les objets géographiques dans le temps il est nécessaire de constituer une base de données géographiques historiques où les objets qui représentent la même entité du monde réel à deux époques différentes sont mis en relations.

D'une manière générale, la création de la base de données historique, il est nécessaire de disposer de données (vectorsielles, cartes topographiques ou photos) à différentes dates ; les objets reconnus comme identiques dans les différentes bases sont ensuite appariés. Ainsi, l'objectif de ce module est de proposer une approche de constitution des bases de données historiques en fonction des données (bases de données géographiques vectorsielles, des cartes topographiques (papier ou numérique)).

Afin de faciliter la tâche des utilisateurs proposons un protocole de création des bases de données géographiques historiques via une interface IHM qui guide les utilisateurs pas à pas. Nous partons de l'idée que pour la constitution d'une base de données historiques il est nécessaire d'avoir au minimum une base de données géographiques vectorsielles qui servira de référence.

Nous présentons dans ce rapport technique les principales étapes nécessaires à la création d'une base de données historique.

La création d'une base de données géographiques historiques dépend des données qu'un utilisateur dispose. Premièrement, l'utilisateur doit choisir, via une interface, le type de données qu'il possède (voir la Figure 1). Notons que les données doivent représenter le même territoire. Plusieurs cas de figures sont possibles :

1. L'utilisateur possède des données vecteur, cartes topographiques papier ou scannées et des photos aériennes ou satellites. Dans ce cas, un pré-traitement de données (2.1 et 2.2) ainsi que les étapes 1, 2, et 3 (3,4,5) sont nécessaires.
2. L'utilisateur possède des données vecteur et des cartes topographiques. Dans ce cas, un pré-traitement de données (2.1) ainsi que les étapes 1, 2, et 3 (3,4,5) sont nécessaires.
3. L'utilisateur possède des données vecteur et des photos aériennes ou satellites. Dans ce cas, un pré-traitement de données (2.2) ainsi que les étapes 1, 2, et 3 (3,4,5) sont nécessaires.
4. L'utilisateur possède uniquement des données vecteur. Dans ce cas, uniquement les étapes 1, 2, et 3 (3,4,5) sont nécessaires.

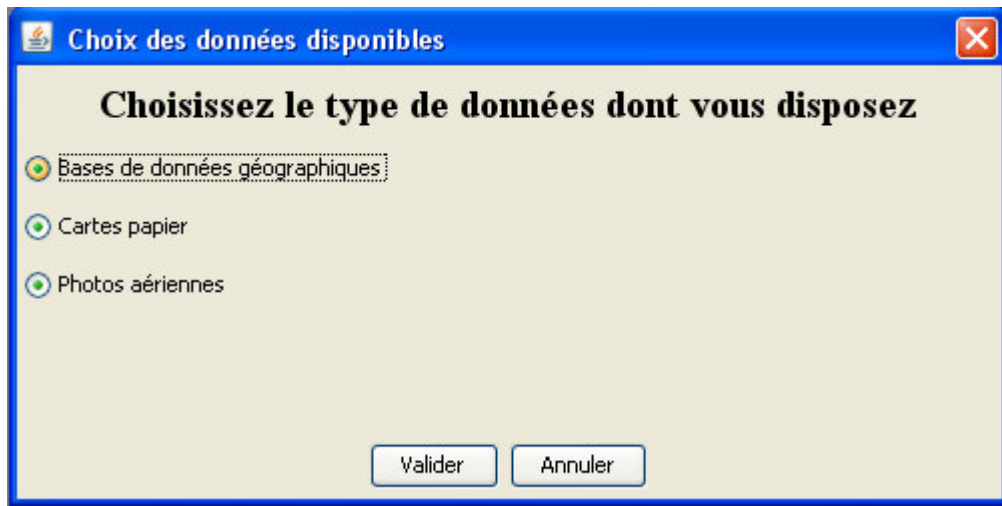


Figure 1. Choix de type de données

## 2. Pré-traitements de données

### 2.1 Création des fichiers shapes à différentes dates à partir des cartes topographiques papiers ou numériques – par analyse visuelle

L'objectif est de créer les données géographiques vectorielles à partir des cartes topographiques. Plusieurs étapes sont nécessaires :

1. **Scannage** de la carte si l'utilisateur possède une carte topographique papier
2. **Géoréférencement** de la carte en utilisant la base de données géographiques de référence (la plus récente).
3. **Création des données vecteur** : Nous proposons une création des bases de données géographiques historiques dite descendante, en superposant une base de données représentant le monde réel au moment (t) sur une carte géoréférencée représentant le monde réel au moment (t-1) (voir la partie 5).

### 2.2 Création des fichiers shapes à différentes dates à partir des photos – par analyse visuelle

L'objectif est de créer les données géographiques vectorielles à partir des photos. Plusieurs étapes sont nécessaires :

1. **Scannage** de la carte si l'utilisateur possède une photo papier
2. **Géoréférencement** de la photo en utilisant la base de données géographiques de référence (la plus récente).
3. **Création des données vecteur** : Nous proposons une création des bases de données géographiques historiques dite descendante, en superposant une base de données représentant le monde réel au moment (t) sur une carte géoréférencée représentant le monde réel au moment (t-1). Pour plus de détail, veuillez voir la partie 5.

## 3. Etape 1 : Chargement des données shapes

L'objectif de cette étape (Figure 2) est de créer les tables de données, c'est-à-dire de les stocker dans un SGBD, par exemple PostGIS.



Figure 2. Etape 1 du processus : Chargement de données shapes

#### **4. Etape 2 : Génération des identifiants ID\_GEO pour la base de données de référence (la plus récente)**

Cette étape (voir Figure 3) vise à générer les identifiants ID\_Geo pour **uniquement** la base de données de référence (la base de données la plus récente). Ces identifiants sont nécessaires pour la création des bases de données historiques (voir **RT\_**).



Figure 3. Etape 2 du processus : Génération des attributs

#### **5. Etape 3 : Appariement des données géographiques**

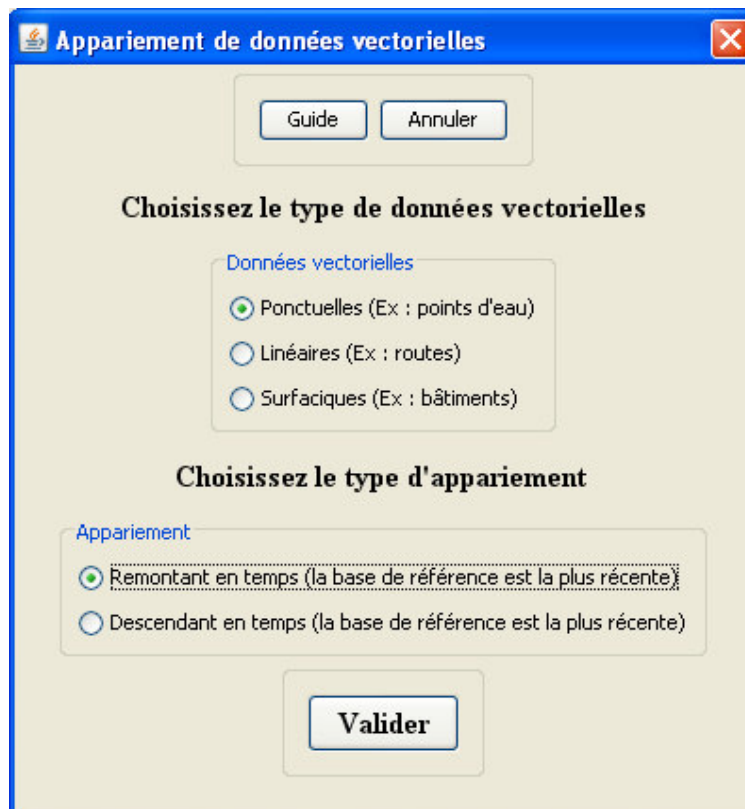
La dernière étape du processus de création des bases de données géographiques historiques est l'appariement de données (voir Figure 4).



**Figure 4. Etape 3 du processus : appariement de données**

L'appariement de données géographiques est un processus qui consiste à mettre en correspondance des objets homologues représentant la même réalité. L'appariement s'appuie sur la notion de ressemblance, c'est-à-dire que deux objets géographiques A et B appartenant à des bases de données différentes sont appariés s'ils se ressemblent. Il consiste donc à mettre en valeur des ressemblances de lieu, de nature, de relation spatiale ou de forme.

Dans le contexte de la création des bases de données historique le processus d'appariement que nous utilisons est composé des étapes suivantes (voir la Figure 5) :



**Figure 5. Choix du type de données à appariées**

## **5.1 Choix de type d'appariement et de données à apparier**

Afin de pouvoir apparier les différents jeux de données que les utilisateurs possède, l'utilisateur doit d'abord choisir le type de données (ponctuel, linéaire ou surfacique) et le type d'appariement (remontant ou descendant dans le temps).

### **5.1.1 Le type d'appariement**

Le processus d'appariement que nous proposons est un processus qui apparie deux bases de données à deux dates différentes. Afin de définir le suivi des objets dans le temps, nous proposons un type d'appariement dit « descendant dans le temps » (notée dans l'interface « Descendant » ) et un autre dit « remontant dans le temps » (notée dans l'interface « Remontant »).

**L'approche descendante dans le temps** est la suivante : nous considérons que la base de données au moment (t) est la base de données de référence. Celle-ci est appariée avec la base de données au moment (t-1), ensuite cette dernière devient elle-même base de référence et elle est appariée avec la base de données au moment (t-2) et ainsi de suite.

Cette approche peut être utilisée lorsqu'un utilisateur possède une base de données vecteur au moment (t) et des photos aériennes et/ou cartes topographiques au moment (t-n).

**L'approche remontante dans le temps** est la suivante : nous considérons la base de données au moment (t) la base de donnée de référence. Celle-ci est appariée avec la base de données au moment (t+1), ensuite cette dernière devient elle-même base de référence et elle est appariée avec la base de données au moment (t+2) et ainsi de suite.

Cette approche peut être utilisée lorsque vous disposez d'une base de données vectorielle au moment (t) et d'une base de données vectorielles au moment (t+1) créée à partir des photos aériennes et/ou cartes topographiques.

### **5.1.2 Le type de données à apparier**

Nous proposons également un appariement en fonctions de types de données, à savoir des données ponctuelles, linéaires et surfaciques.

## **5.2 Définition des liens d'appariement et génération des identifiant ID\_GEO**

Le processus d'appariement utilisé est basé sur l'approche proposée par [Olteanu, 2009], qui consiste à fusionner plusieurs critères d'appariement afin de prendre en compte les imperfections dans les données.

La première étape du processus d'appariement consiste à chercher pour chaque objet géographique appartenant à une base de données à un moment t, des objets homologues potentiels dans la base de données au moment (t-1), appelé candidats à l'appariement. Ensuite, chaque candidat à l'appariement est analysé afin de déterminer le lien d'appariement final. La troisième étape consiste à fusionner les candidats ensembles afin d'avoir une vue d'ensemble. Enfin, la dernière étape est la prise de décision, c'est-à-dire que le meilleur candidat est choisi. Mentionnons que chaque objet de la base de données au moment (t) est traité tout à tour comme ci-dessus, l'ordre n'ayant pas d'importance.

### **5.2.1 Analyse des liens d'appariement pour la génération des identifiants ID\_GEO.**

Après chaque étape d'appariement, par exemple l'appariement de BDG(t) et BDG(t-1), nous générons d'une manière automatique les identifiants ID\_GEO. Cette étape est réalisée en analysant la cardinalité des liens d'appariement, voir les tableaux ci-après.

Liens BDG(t) : BDG(t-1)	Interprétation pour BDG(t)	Opération à faire sur l'objet du BDG(t-1)
<b>1 : 1</b>	Aucune modification	Copier l'ID_GEO de l'objet du BDG(t)
<b>1 : 0</b>	Création	Pas d'opération à faire
<b>0 : 1</b>	Destruction	Créer ID_GEO
<b>1 : n</b>	Fusion	-Créer <b>n</b> ID_GEO pour les <b>n</b> objets du BDG(t-1) -Définir la relation entre l'objet du BDG(t) et les <b>n</b> objets du BDG(t-1)
<b>m : 1</b>	Désagrégation	-Créer un ID_GEO -Définir la relation entre les <b>m</b> objets du BDG(t) et l'objet du BDG(t-1)
<b>n : m</b>	Restructuration	-Créer <b>m</b> ID_GEO pour les <b>m</b> objets du BDG(t-1) -Définir la relation entre les <b>n</b> objets du BDG(t) et les <b>m</b> objets du BDG(t-1)

Tableau 1. Analyse des liens d'appariement d'après l'approche descendante dans le temps

Liens BDG(t) : BDG(t+1)	Interprétation pour BDG(t)	Opération à faire sur l'objet du BDG(t+1)
<b>1 : 1</b>	Aucune modification	Copier l'ID_GEO de l'objet du BDG(t)
<b>1 : 0</b>	Destruction	Pas d'opération à faire
<b>0 : 1</b>	Création	Créer ID_GEO
<b>1 : n</b>	Désagrégation	-Créer <b>n</b> ID_GEO pour les <b>n</b> objets du BDG(t+1) -Définir la relation entre l'objet du BDG(t) et les <b>n</b> objets du BDG(t+1)
<b>m : 1</b>	Fusion	-Créer un ID_GEO -Définir la relation entre les <b>m</b> objets du BDG(t) et l'objet du BDG(t+1)
<b>n : m</b>	Restructuration	-Créer <b>m</b> ID_GEO pour les <b>m</b> objets du BDG(t+1) -Définir la relation entre les <b>n</b> objets du BDG(t) et les <b>m</b> objets du BDG(t+1)

Tableau 2. Analyse des liens d'appariement d'après l'approche remontant dans le temps

### **Exemple d'utilisation :**

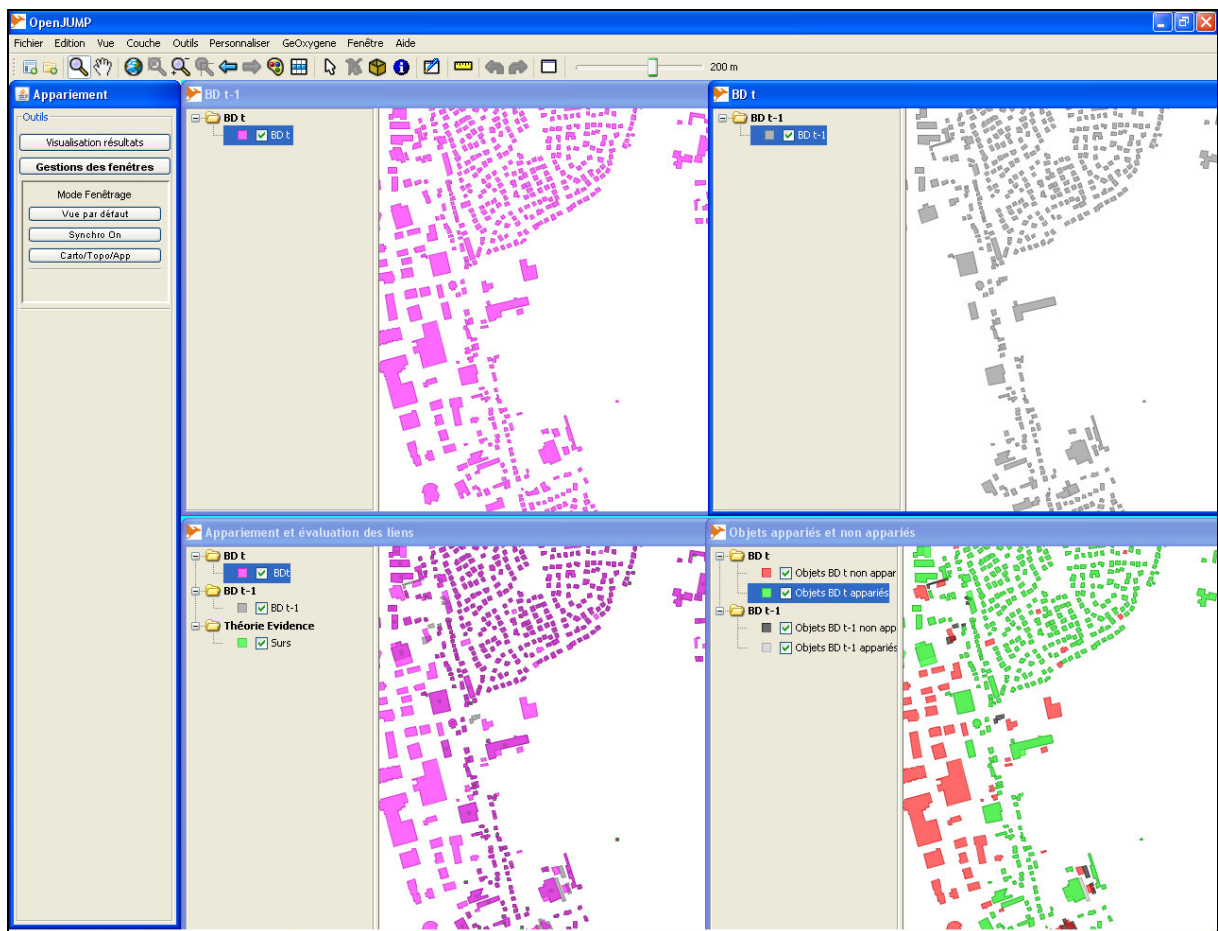
Supposons qu'un utilisateur souhaite appairier quatre bases de données BDG(2009), BDG(2002), BDG(1989) et BDG(1978), où les bases de données BDG(2009) BDG(1989) et BDG(1978) ont été créés à partir de photos aériennes et cartes topographiques. La base de données de référence est la BDG (2002) étant donné que les autres bases ont été créées à partir des images, voir cartes topographiques. Afin de réaliser la base de données historiques un utilisateur devrait réaliser les tâches suivantes.

1. Un appariement de type descendant
  - a. Appariement des BDG(2002) et BDG(1989). Mentionnons qu'à l'étape 2 du processus général nous avons déjà généré les identifiants ID\_GEO pour la base BDG(2002),
  - b. Création automatique des ID\_GEO pour la base BDG(1989),

- c. Appariement de la base BDG(1989) qui **contient les ID\_GEO créés lors de l'étape b**, et la base BDG(1978)
  - d. Création automatique des ID\_GEO pour la base BDG(1978)
2. Un appariement de type ascendant
    - a. Appariement des BDG(2002) et BDG(2009). Mentionnons qu'**à l'étape 2 du processus général nous avons déjà généré les identifiants ID\_GEO pour la base BDG(2002)**
    - b. Création automatique des ID\_GEO pour la base BDG(2009)

### 5.3 Visualisation des résultats d'appariement

Les résultats d'appariement peuvent être visualisés via l'interface OpenJump. A la fin d'un type d'appariement cette interface est composée de 5 fenêtres principales :



**Figure 6. Interface de visualisation des résultats d'appariement**

- Fenêtres d'affichage des données
  - Les données à appairer sont affichées dans les fenêtres « *BD t-1* » et « *BD t* » situées en haut.
  - Les liens d'appariement entre les données appariées sont affichés dans la fenêtre « *Appariement et évaluation des liens* » en bas à gauche
  - Le bilan d'objets appariés et non-appariés des deux bases de données est affiché dans la fenêtre « *Objets appariés et non-appariés* » en bas à droite.
- Fenêtre « *Appariement* » (à gauche) contenant deux boutons : « *Visualisation résultats* » et « *Gestion des fenêtres* ».



- Le premier bouton permet de visualiser, un par un, les résultats (objets appariés, non-appariés, liens d'appariement, etc). Lorsqu'on appuie sur ce bouton une barre permettant de choisir les résultats à visualiser s'affiche en bas de l'interface.
- Le bouton « *Gestion des fenêtres* » permet de réaliser la synchronisation des 4 fenêtres expliquées ci-dessus, d'afficher uniquement trois fenêtres et de revenir sur la vue par défaut (4 fenêtres).