

Approche géomatique pour la reconstitution des dates de pose des réseaux d'eau potable

Master financé par l'IMU et réalisé au sein d'Eau du Grand Lyon sous l'encadrement de David Poinard (Eau du Grand Lyon) et Frédéric Cherqui (DEEP INSA Lyon / UCBL)

Contexte du projet HIREAU : <https://hireau.org/>

Le projet de recherche HIREAU est un projet de recherche qui a pour objectif d'explorer différentes méthodes pour reconstituer les dates de pose des réseaux d'eau potable et d'assainissement. Il est financé par le Labex IMU. Ce projet regroupe plusieurs acteurs, ayant chacun un rôle différent ou une méthode différente à explorer. L'enjeu pour la Métropole de Lyon est de reconstituer les dates de pose des réseaux d'assainissement, et de fiabiliser les dates pour l'eau potable. Plus largement, l'enjeu pour le consortium est de démontrer que l'on peut reconstituer ces dates pour ainsi permettre à d'autres collectivités de répondre aux contraintes réglementaires et mettre en œuvre une réelle gestion patrimoniale. HIREAU se propose ainsi d'explorer plusieurs approches de reconstitution : la fouille d'archive, la connaissance de l'urbanisation du territoire, la mémoire collective, l'analyse *in situ*, et l'utilisation de règles métier.

Objectif du travail de Master

L'objectif de ce travail de Master est d'explorer une approche géomatique pour reconstituer les dates de pose en exploitant le critère spatial et les relations topologiques entre tronçons. Ce travail consiste à identifier les pratiques de pose et de renouvellement des conduites, ainsi que les caractéristiques des tronçons permettant d'estimer leur date, en exploitant la base de données SIG d'Eau du Grand Lyon. Le principe est de mettre au point un algorithme de déduction de proche en proche, exploitant les relations spatiales entre les tronçons de date inconnue et les tronçons de date connue, en se basant sur des caractéristiques communes entre ces tronçons voisins pour affecter une date de pose à un tronçon inconnu à partir de son voisin connu. Il s'agira donc d'identifier les caractéristiques pertinentes à prendre en compte dans ce processus, en lien avec les experts métiers, afin d'aboutir au final à la réalisation d'algorithmes permettant la déduction automatisée de dates en exploitant ces règles.

Pratiques métiers et règles retenues

Pour réaliser cet algorithme de déduction de proche en proche, il est nécessaire de connaître les pratiques métier locales de pose et de renouvellement, ainsi que le fonctionnement général du réseau. Ces connaissances servent de base à l'élaboration de règles de traitement automatisé.

Les notions les plus importantes à prendre en compte sont les suivantes : tout d'abord, la notion de *bief*, qui constitue un ensemble de tronçons situés entre deux ou plusieurs vannes. D'une manière générale, un bief est renouvelé en entier. Le *matériau* du tronçon est aussi un critère important, les différents matériaux constituant le réseau n'ayant pas été utilisés durant la même période temporelle. Ainsi, un ensemble de tronçons de même matériau et faisant partie du même bief ont probablement la même date de pose. Le principe est de réaliser un parcours de graphe (Figure 1) et, pour chaque tronçon inconnu connecté à un voisin connu, d'utiliser les caractéristiques communes entre ces tronçons pour propager la date ou non. Le *caractère structurant* ou non du tronçon peut aussi être pris en compte. D'autres critères de moindre

importance (mais augmentant la fiabilité de la date déduite) peuvent être utilisés comme le *diamètre* par exemple.

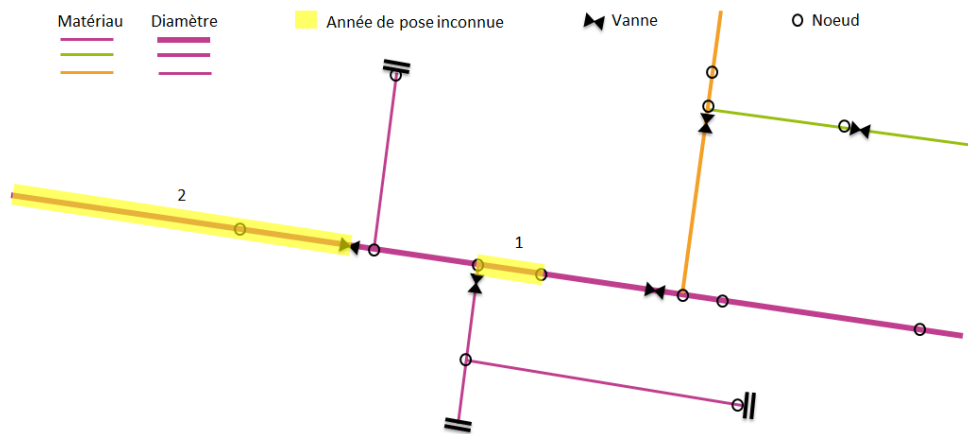


Figure 1 : illustration d'un parcours de graphe sur un réseau d'eau potable. Le tronçon de date inconnue numéroté « 1 » a probablement la même date que les tronçons qui l'encadrent car ces tronçons ont le même matériau, font partie du même bief et sont du même diamètre.

Une autre donnée très importante doit être prise en compte : les plans de récolement. Ils sont réalisés à l'issue d'un chantier afin de rendre compte des travaux réellement effectués. Lorsque ce chantier est un chantier de renouvellement, la date du plan constitue la date de pose du tronçon. Cette donnée permet donc de déduire des dates, mais aussi d'interdire la propagation d'une date : un tronçon renouvelé n'a aucun lien avec un tronçon voisin non renouvelé, et deux tronçons connectés faisant partie d'un plan de récolement différent n'ont aussi aucun lien entre eux, même si ils possèdent tous les autres points communs. Il est donc très important de connaître les tronçons renouvelés et les tronçons non renouvelés.

Méthode utilisée

Une fois les règles établies, elles doivent être testées afin de déterminer leur fiabilité. La méthode de test retenue est la suivante :

- établir une zone de test dont les dates de pose sont entièrement connues, représentative de l'ensemble du réseau, et dont les tronçons sont correctement connectés topologiquement ;
- rendre inconnu un certain pourcentage de dates de la zone (de manière aléatoire et diffuse spatialement, ou de manière concentrée comme représenté sur la figure 2a ci-dessous), puis exécuter les algorithmes de déduction décrits précédemment ;
- analyser les dates retrouvées afin de déterminer si elles sont correctes.

Les tests ont été réalisés avec des taux de connaissance de départ sur la zone test différents : 20%, 40%, 60% et 80% du linéaire connu. Ainsi, au final, il est possible de déterminer :

- la fiabilité de la méthode, via l'analyse de l'exactitude des dates retrouvées ;
- son efficacité dans différentes circonstances : inconnus diffus, concentrés, taux de connaissance de départ faible ou élevé.

Les règles sont testées individuellement tout d'abord pour déterminer leur fiabilité et leur efficacité, puis ensemble lors d'un test combiné. Le principe est de propager d'abord les dates en exigeant un grand nombre de critères communs entre tronçons, puis de « dégrader » la règle pour

trouver d'autres dates manquantes, avec l'affectation d'un indice de fiabilité pour chaque date déduite suivant les critères ayant permis de la déterminer. Enfin, le test combiné est ensuite réalisé sur le réseau réel (Figure 2b) pour lequel 70% du linéaire est connu, et qui est composé d'inconnus diffus et concentrés.

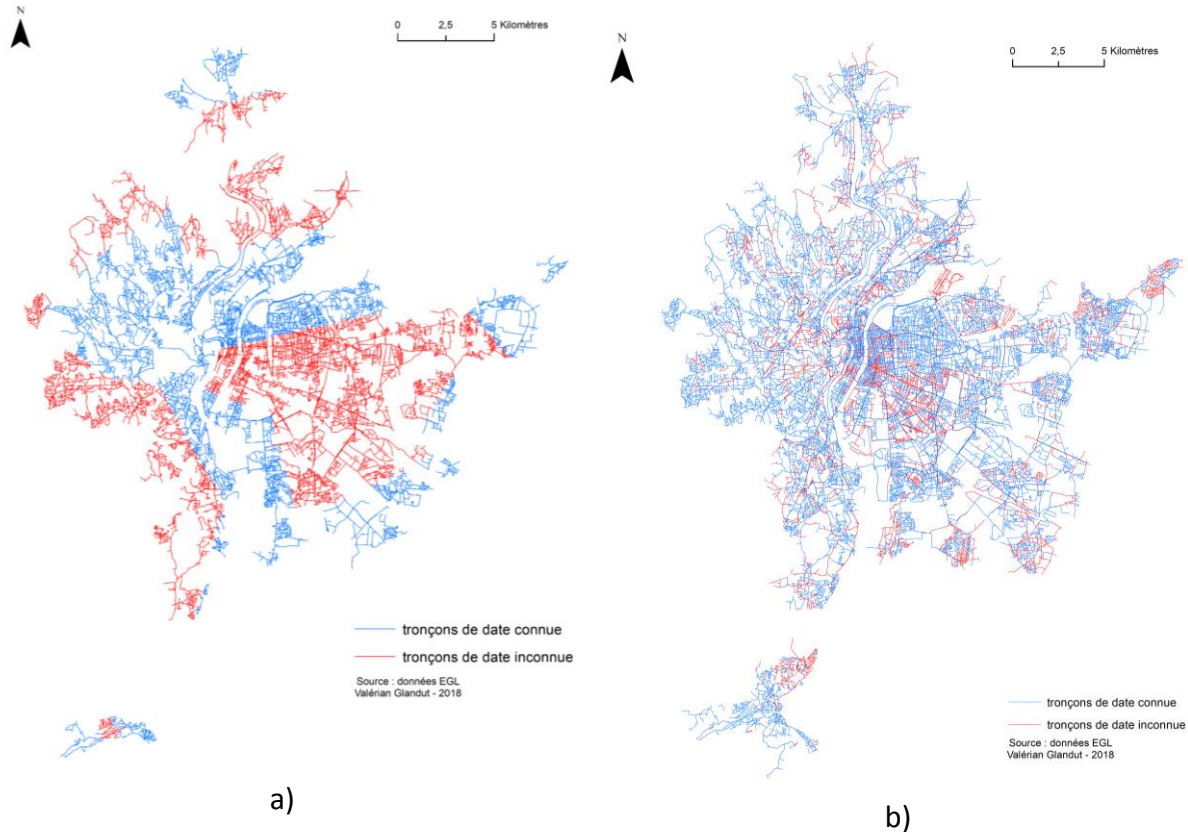


Figure 2 : a) zones de test des règles avec des tronçons de date de pose inconnue concentrés, et b) ensemble réseau réel.

Résultats obtenus

Les résultats des tests ont démontré que la propagation de dates était beaucoup plus efficace sur des inconnus diffus que sur des inconnus concentrés, et que l'algorithme utilisé (assez exigeant en termes de critères de propagation) avait tendance soit à propager des dates de manière très fiable, soit à ne pas les propager du tout. Les meilleurs résultats pour les inconnus diffus s'expliquent par la méthode qui fonctionne de proche en proche. Un point très positif est que la méthode s'arrête avant de propager des dates erronées. La Figure 3 ci-dessous présente la pyramide des âges obtenus pour un réseau avec 60% d'inconnus diffus.

On constate sur cette figure que même avec un faible pourcentage de dates connues au départ (40%), la méthode retrouve près de 40% des dates avec un écart de moins de 5 ans, et les dates retrouvées sont réparties dans le temps. Les dates avec un écart supérieur à 5 ans restent marginales (moins de 2%).

Concernant les résultats du test combiné sur la totalité du réseau de la Métropole, on constate que l'algorithme a permis de déduire des dates sur environ 5% du linéaire total. Le pourcentage faible de dates déduites s'explique par le pourcentage élevé de dates connues au départ, ainsi que par l'exigence forte de correspondance définie par les règles (le maximum de caractéristiques est

pris en compte). Dans 80% des cas où la propagation s'est arrêtée, le tronçon inconnu était fait d'un matériau différent que ses voisins connus (différence de matériau qui peut aussi être associée à une différence de plan de récolement). Il est également intéressant de noter que l'efficacité de l'algorithme est très concentrée spatialement au Sud-Est et au Nord de l'agglomération.

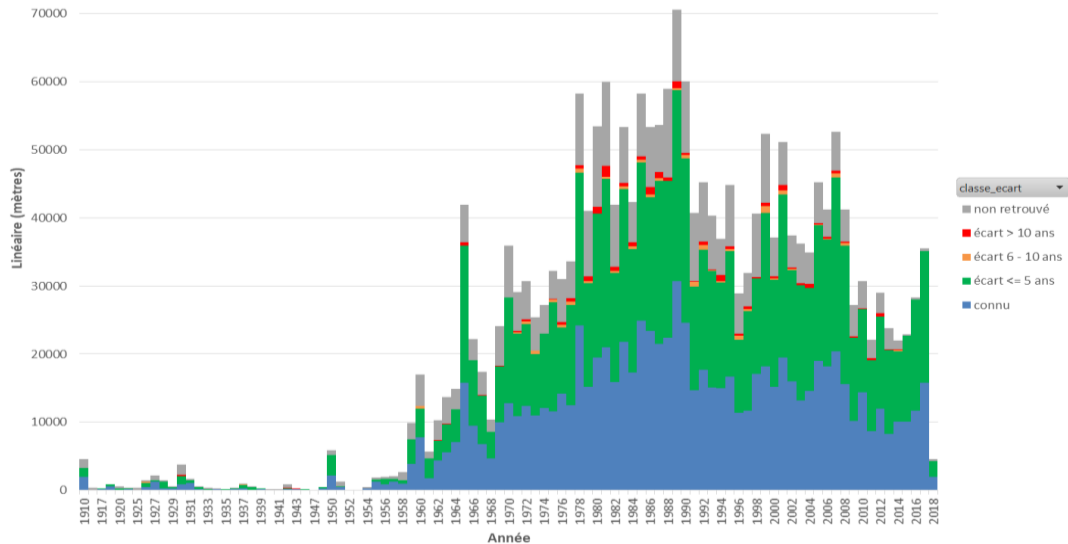


Figure 3 : représentation des résultats par pyramide des âges de la propagation avec 40% de linéaire connu, avec des inconnus diffus.

Conclusion

En conclusion, l'efficacité de la méthode géomatique explorée lors de ce travail est dépendante de plusieurs facteurs :

- L'exhaustivité de la base de données (matériaux, plans de récolement, bief...)
- La qualité des connexions topologiques entre tronçons (permettant de mettre en relation un tronçon avec son voisin), ainsi que les points communs entre tronçons connus et inconnus (si tous les tronçons inconnus ont des caractéristiques différentes des tronçons connus, cette méthode sera peu efficace)
- La qualité du lien entre les plans de récolement et les tronçons (connaissance des tronçons renouvelés ou non, connaissance des tronçons concernés par chaque plan, caractérisé par un lien attributaire entre les tronçons et les plans de récolement)

Sur une base de données correctement mise à jour, les résultats sont très prometteur même lorsque que le pourcentage de départ de dates de pose connues est faible.

L'algorithme de propagation, est également fortement adaptable : avec un paramétrage prenant en compte un nombre important de caractéristiques, le pourcentage de date déduite sera plus faible mais plus fiable. En réduisant le nombre de caractéristiques prises en compte, on peut augmenter le pourcentage de dates reconstituées (en réduisant la fiabilité). Ce paramétrage doit être adapté aux objectifs du gestionnaire (pourcentage du linéaire à reconstituer, précision attendue) et au patrimoine lui-même (fiabilité de la base de données, répartition des dates de pose inconnues).