

# MÉTHODES INNOVANTES POUR DÉTECTER ET DÉCRIRE LES RÉSEAUX D'EAUX URBAINS

CAROLE DELENNE, NANÉE CHAHINIAN



Cafe'T Observil

26/01/2022



Lemon



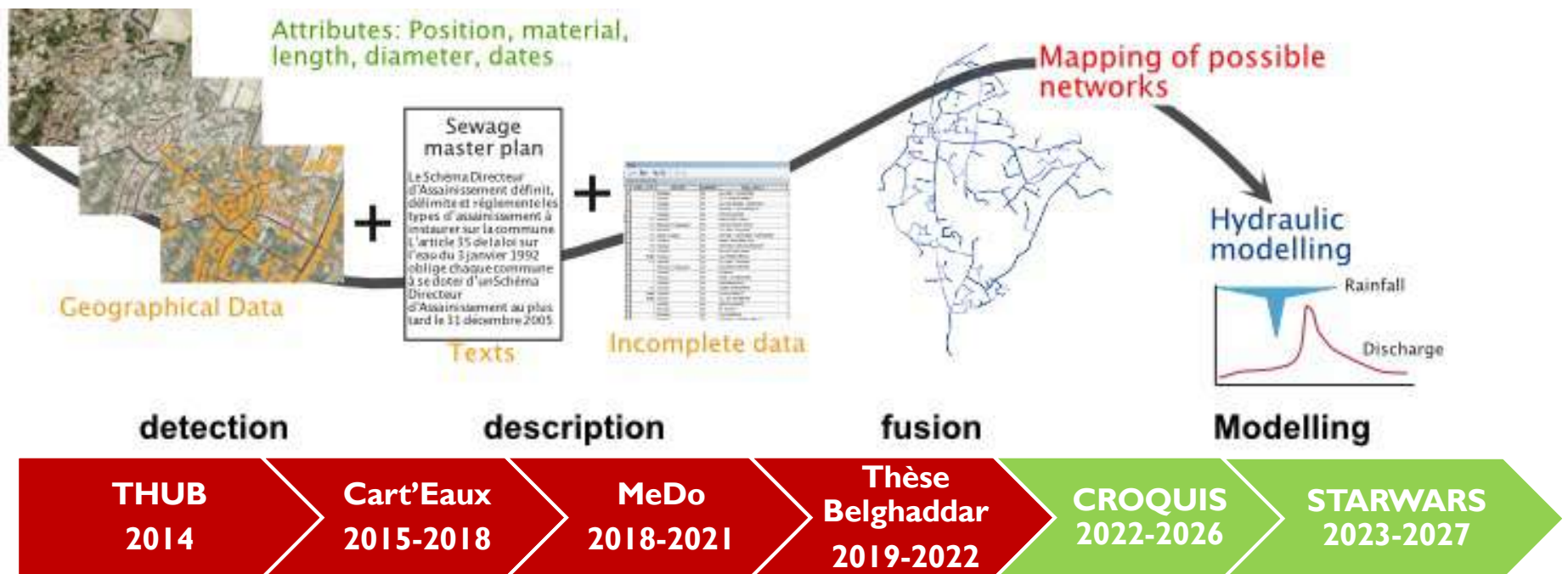
Inria

# MÉTHODES INNOVANTES POUR DÉTECTER ET DÉCRIRE LES RÉSEAUX D'EAUX URBAINS

- Manque d'information sur les réseaux enterrés dans les pays développés comme dans les pays en développement => retards, augmentation des coûts, gestion sous-optimale des ressources.
- Des algorithmes existent pour déterminer la configuration optimale d'un réseau (Walter, 1995; Afshar, 2007 ; Moeini et Afshar, 2012 ; Swamee & Sharma, 2013).
  - Configuration optimale  $\neq$  réseau réel
- La disposition d'un réseau existant peut être reproduite si la topologie est renseignée (Allard et al., 2013 ; Blumensaat et al., 2012).
  - Information souvent manquante et difficile à inférer depuis la surface.

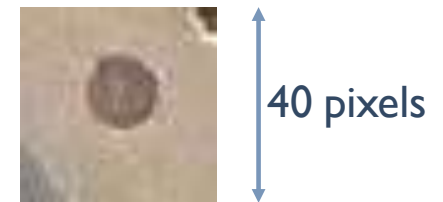
# MÉTHODES INNOVANTES POUR DÉTECTER ET DÉCRIRE LES RÉSEAUX D'EAUX URBAINS

- Comment reconstituer le tracé d'un réseau souterrain à partir de ses composants de surface ?
- Comment renseigner la géométrie du réseau en ayant recours à la fouille de données textuelles ?



# CONTRAINTES DE LA DETECTION DES PLAQUES D'ÉGOUTS SUR DES IMAGES THRS

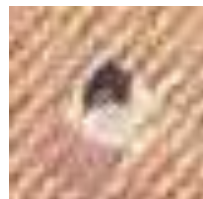
- 80cm seulement => 16 pixels maximum à une résolution de 5cm.
- Généralement très peu de contraste



- De nombreux autres objets circulaires sur l'image.



Lampadaire



Cheminée



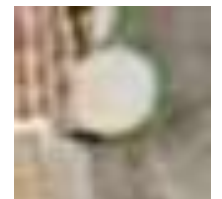
???



Table de jardin



Bouée ??



Parabole?

Commandré et al., 2018

# DETECTION DES PLAQUES D'ÉGOUTS SUR DES IMAGES THRS : METHODE DU FILTRE GEOMETRIQUE (1/2)

- Orthophoto à 5cm de résolution (acquisition) & images Pléiades à 25cm de résolution.

Traitement

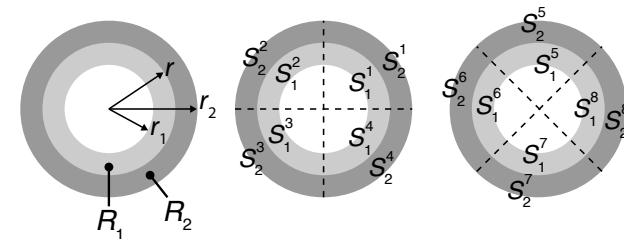
- Détecter les routes
- Filtrer la végétation (NDVI et ExG)
- Eliminer les ombres (Nagao et al., 2015 et ICM)

Détection

- Filtre géométrique (Nigaki et al., 2012)
- Méthode par apprentissage SVM (Chaumont et al., 2014)\*
- Méthode combinée (Pasquet et al., 2015)\*

Evaluation

- Precision =  $TP/(TP+FP)$ ; Rappel =  $TP/(TP+FN)$



# MÉTHODE FILTRE CIRCULAIRE: RÉSULTATS

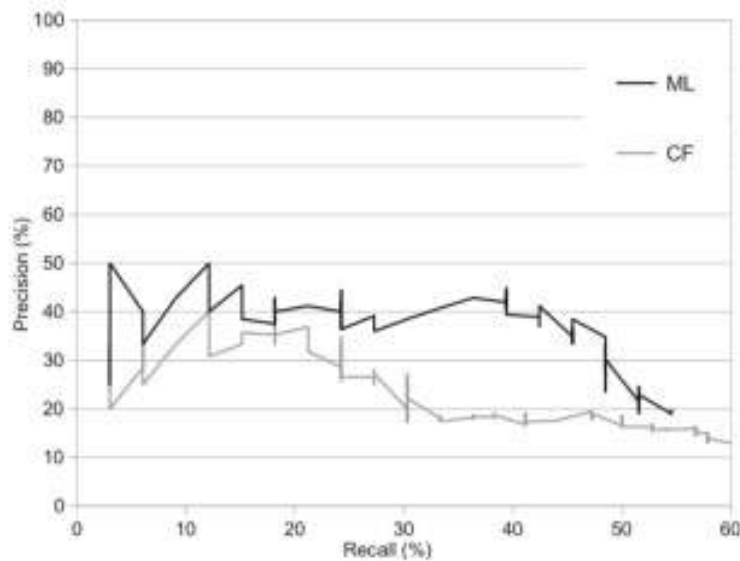


Fig. 9: Precision versus recall, with circular pattern detection (CF) and machine-learning method (ML).

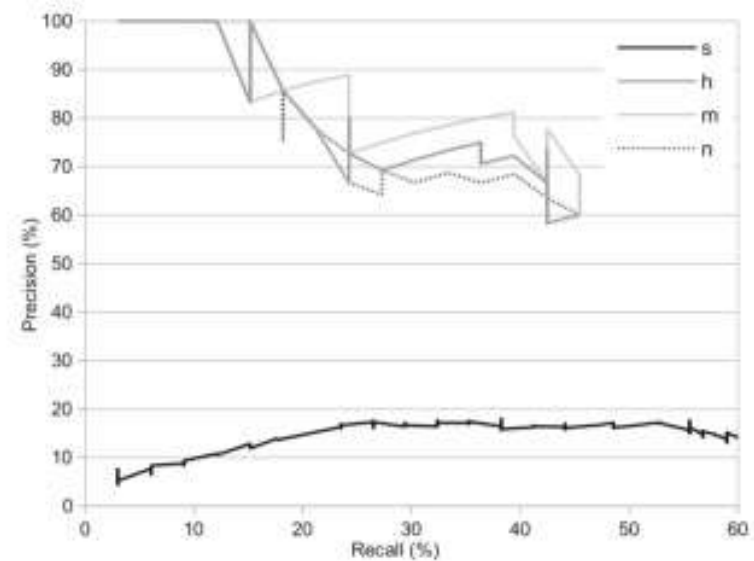


Fig. 10: Results of the combined method: sum (s), product (h), maximum (m), and minimum (n) functions. Comparison with the two approaches (ML: machine learning) and (CF: circular filter).

# METHODOLOGIE : DETECTION DES PLAQUES D'ÉGOUTS SUR DES IMAGES THRS (2/2)

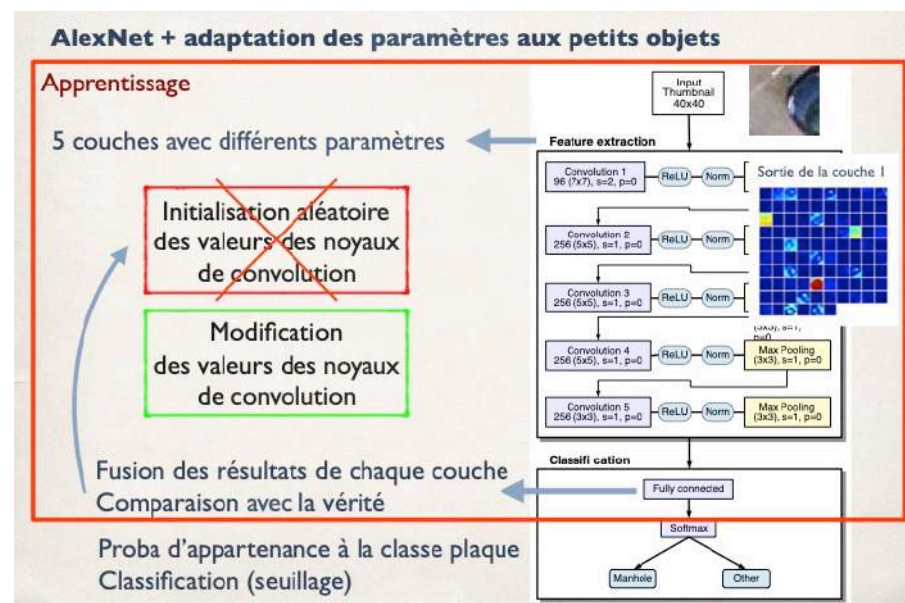
## ■ Méthode Deep Learning

### ■ Base d'apprentissage:

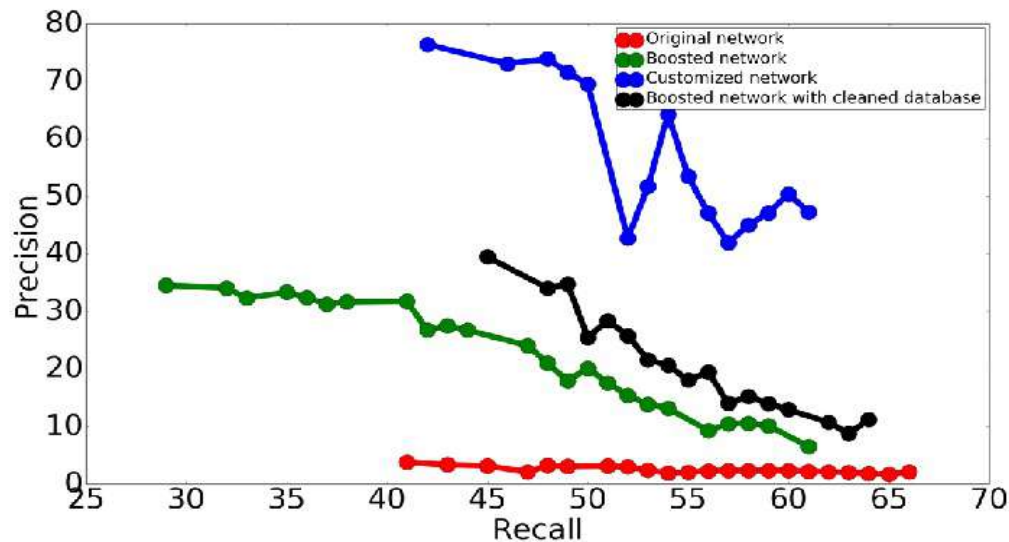
- Prades le Lez : image en couleurs naturelles à 5cm
- 605 plaques => augmentation de la base (env 20 000 plaques 50 0000 autres)

### ■ Réseau AlexNet :

- Convolutional Neural Network CNN
- Paramètres adaptés à la détection des petits objets
- 5 couches
- Fenêtre glissante 40x40 sur toute l'image (pas de 4 pixels)



# MÉTHODE DEEP LEARNING : RÉSULTATS



**Meilleurs taux de détection possible pour un seuil de 0.26**

▶ Rappel = 60% — Précision = 47%

**Beaucoup de faux positifs:**

▶ Plaques non-référencées dans la base de données (=> autre réseau? Anciennes ou nouvelles plaques?)

▶ Tâches sur le sol

=> améliorer le contraste?

=> regrouper les détections trop proches?

▶ Beaucoup de Faux en dehors des routes

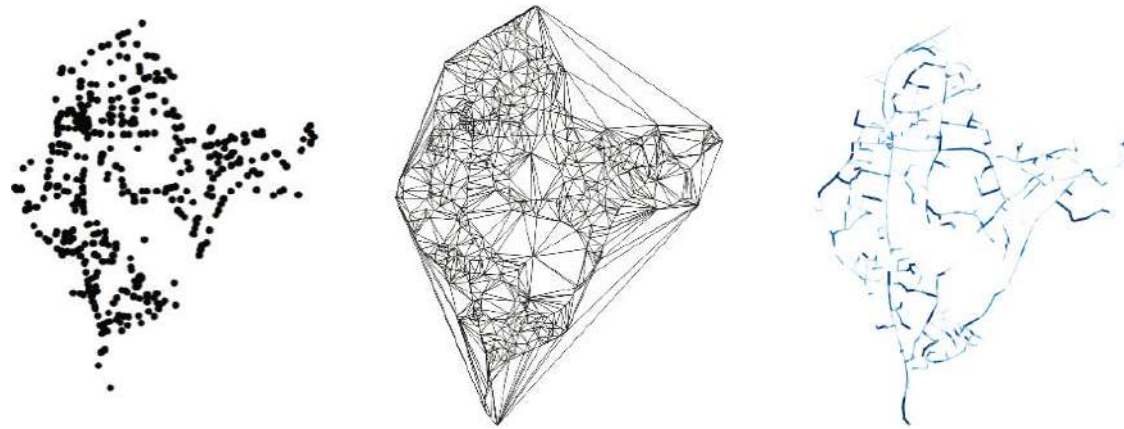
• Filtre sur les routes => Rappel = 54% — Précision = 70%

- Gigan: 100 plaques : - image RVB acquise avec le même capteur, même résolution

Commandré et al., 2017



# MÉTHODOLOGIE : CARTOGRAPHIE DE RÉSEAUX À PARTIR DE LA POSITION DES PLAQUES D'ÉGOUTS



- Nœud du réseau : plaques d'égouts détectées ou positionnées.
- Graphe de Delaunay + sélection des tronçons aux plus faibles coûts  $f$ (distance, pente, croisement de routes ou bâti) (Commandré et al. 2017).
- Algorithme stochastique aval  $\Rightarrow$  amont, avec tirage aléatoire prenant en compte le coût parmi les tronçons possibles (Chahinian et al., 2019).
- Développement d'un GNN spécifique (Belghaddar, 2022).

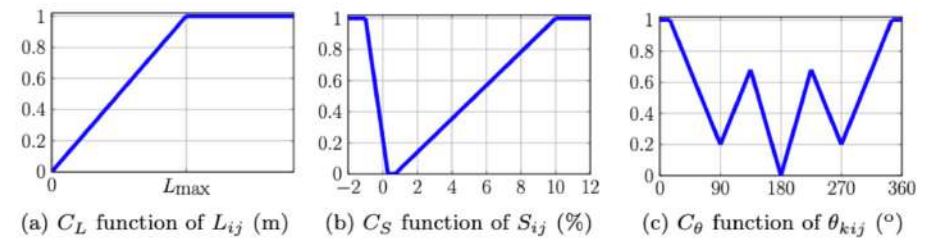
# MÉTHODOLOGIE : RÈGLES DE RECONSTRUCTION

## ■ Évaluation

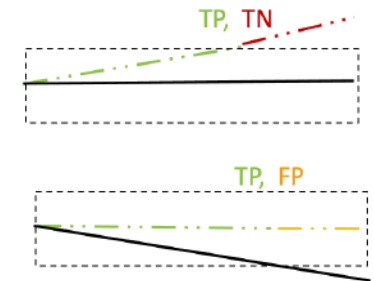
- Exactitude, complétude, qualité (Heipke et al., 1997)
- Arborescence du réseau (Shreve, 1966)
- Distance à l'exutoire (Allard et al., 2013)
- Erreur

## ■ Tests sur les communes de

- Gigean : 792 plaques d'égouts et 23,45 km de conduites (M3M)
- Ramonville St Agne : 1878 plaques d'égouts et 58,9 km km de conduites (Open Data)
- Données cartographiques : BD TOPO et RGE ALTI (IGN).



Chahinian et al., 2019

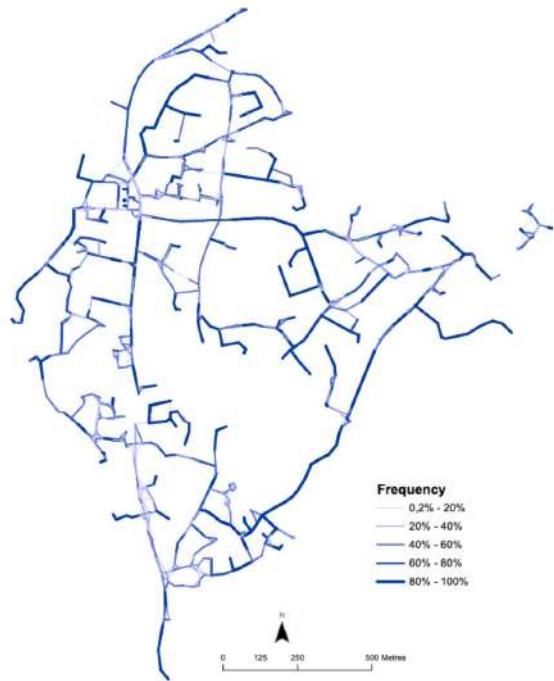


— Verité terrain  
- - - Reconstruction

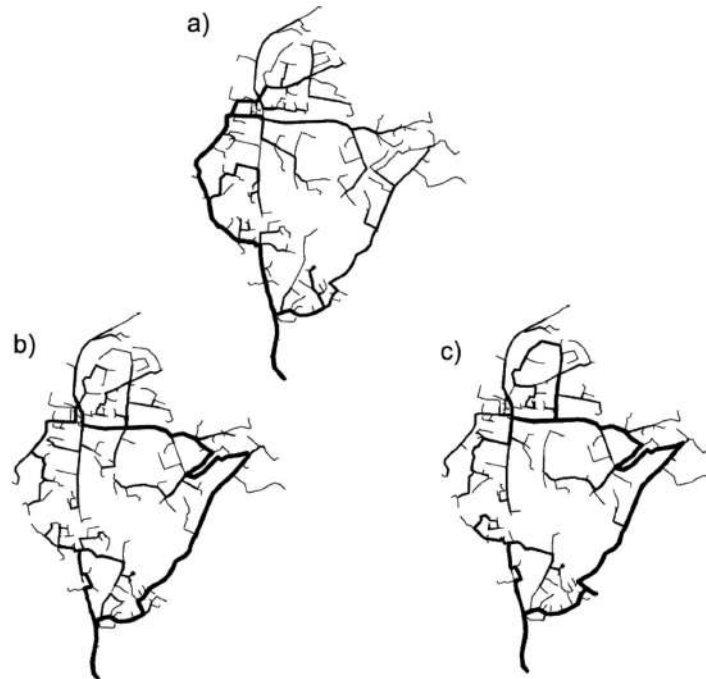


Pente des conduites supposée // surface.

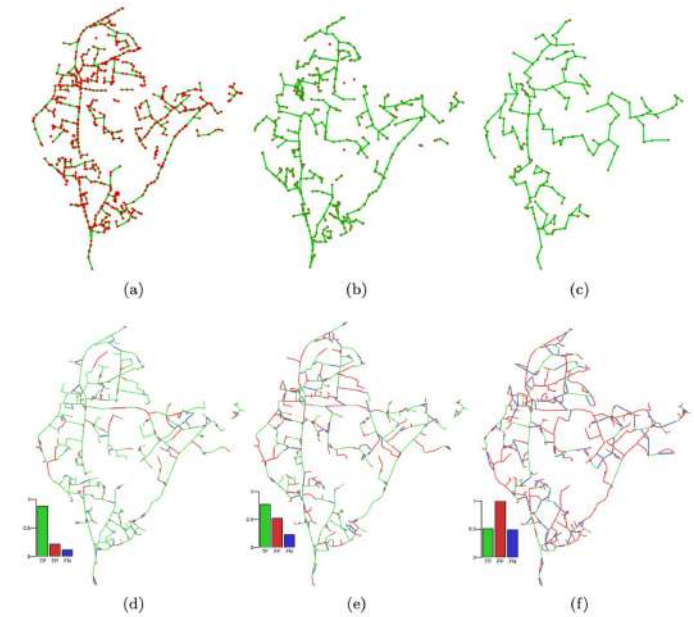
# RÉSULTAT : CARTES CRÉÉES



Fréquence ++ pour tronçons proches de l'exutoire ou en tête de bassin



Topologie du réseau à améliorer  
=> Belghaddar et al., 2022b

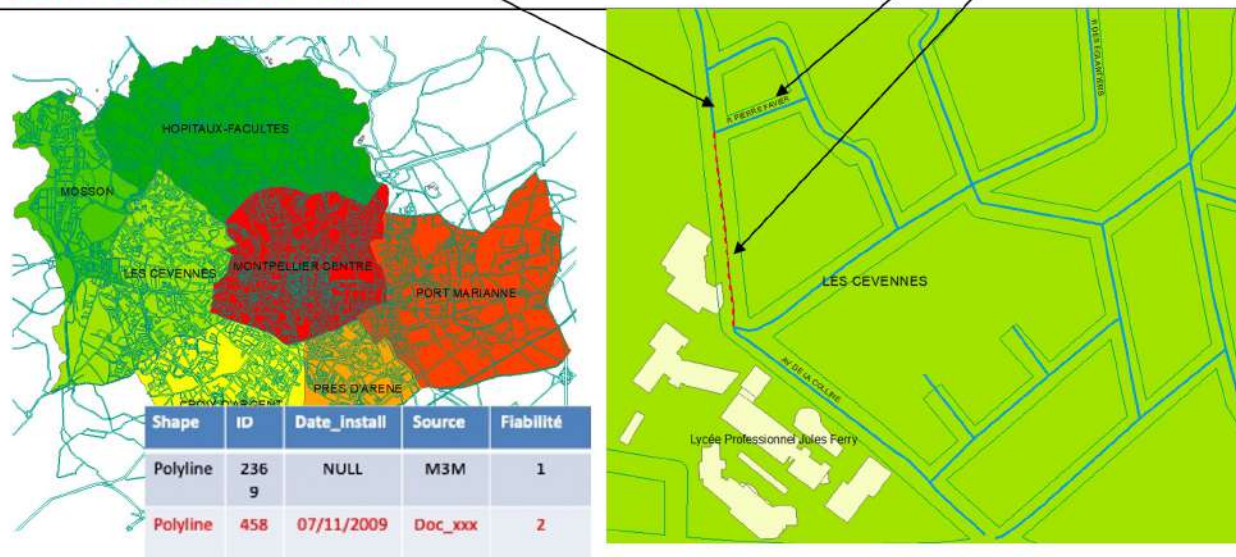


Chahinian et al., 2019

Au moins 50% points pour tracé réaliste

# LA FOUILLE DE DONNÉES POUR COLLECTER DES INFORMATIONS SUR LES RÉSEAUX D'EAUX URBAINS

La **Métropole** va procéder à des travaux d'extension du **réseau d'eaux pluviales** **avenue de la Colline** dans le **quartier Cévennes**, à partir du **jeudi 20 octobre 2009**. Ces **travaux**, qui débutent pour une **durée de 3 semaines**, consistent à réaliser une **extension** de la **conduite d'évacuation des eaux pluviales** de l'**avenue de la Colline**, notamment pour **raccorder le nouveau bâtiment du lycée Jules Ferry**. Cette **conduite** s'arrête **actuellement au-dessus de la rue Favier** et elle sera **prolongée jusqu'à la rue des Eglantiers**.



Documents collectés :  
requête Google  
(2 mots clés+  
localisation)

Bonnabaud La Bruyère et al., 2022

Organisation, Type réseau, Spatial, Temporel, Travaux, Élément réseau

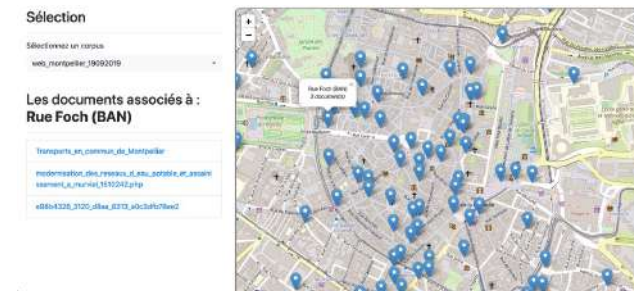
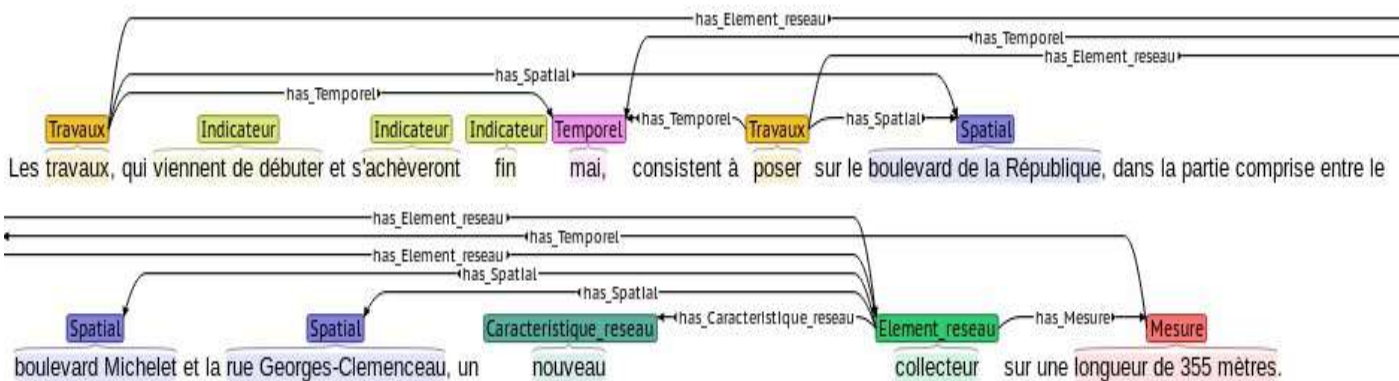
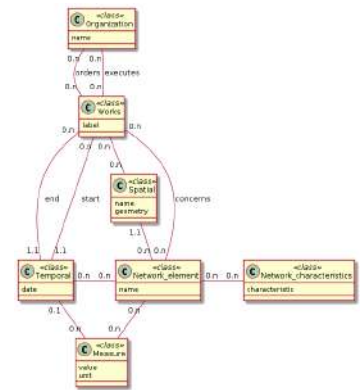
# MÉTHODOLOGIE : EXTRACTION D'ENTITÉS NOMMÉES

- Classification de relevance : Scikit-learn, Multinomial Naive Bayes, IDF-TF
- Découpage des phrases, tokenisation et étiquetage morpho-syntaxique : TreeTagger
- Évaluation : MUC-5 (Chinchor& Sundheim, 1993)
- Ensemble module REN : 13 entités
  - Core-NLP (meilleure precision, F-score= 50.42) ⇒ En cas de conflit, choix du modèle le plus performant
  - SpaCy (Meilleur rappel, F-score=52.29) ⇒ Pour entités temporelles, Heideltime outrepassé les autres.
  - Heideltime (temporels)

Adaptation d'outils existants au domaine de l'assainissement urbain

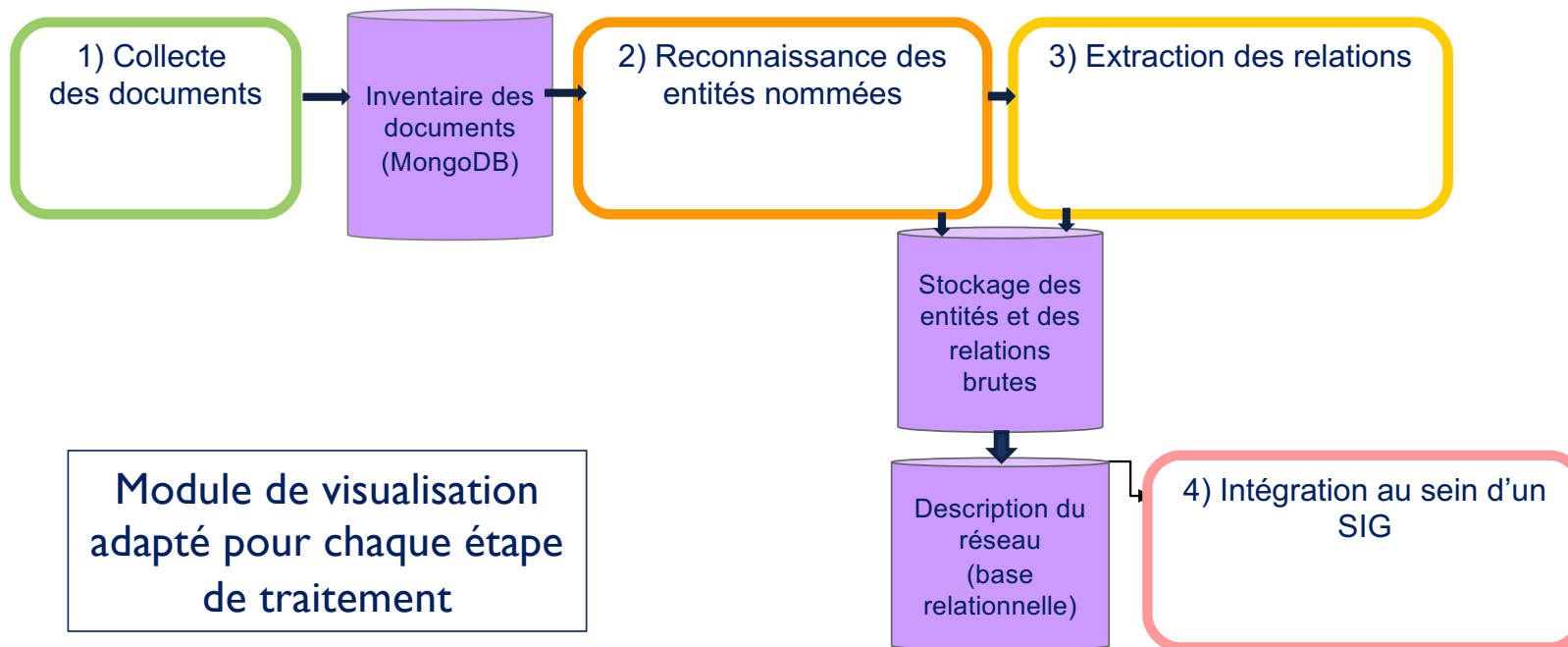
# MÉTHODOLOGIE : EXTRACTION DES RELATIONS

- Sur la base d'un modèle conceptuel développé par le projet MeDo.
- Extraction des relations depuis chaque entité avec :
  - Motifs syntaxiques
  - Annotations environnantes (jusqu'à 2 phrases selon type)
  - Titre du document
- Liage spatial gazetiers : Base Adresse Nationale, Geonames & Nominatim



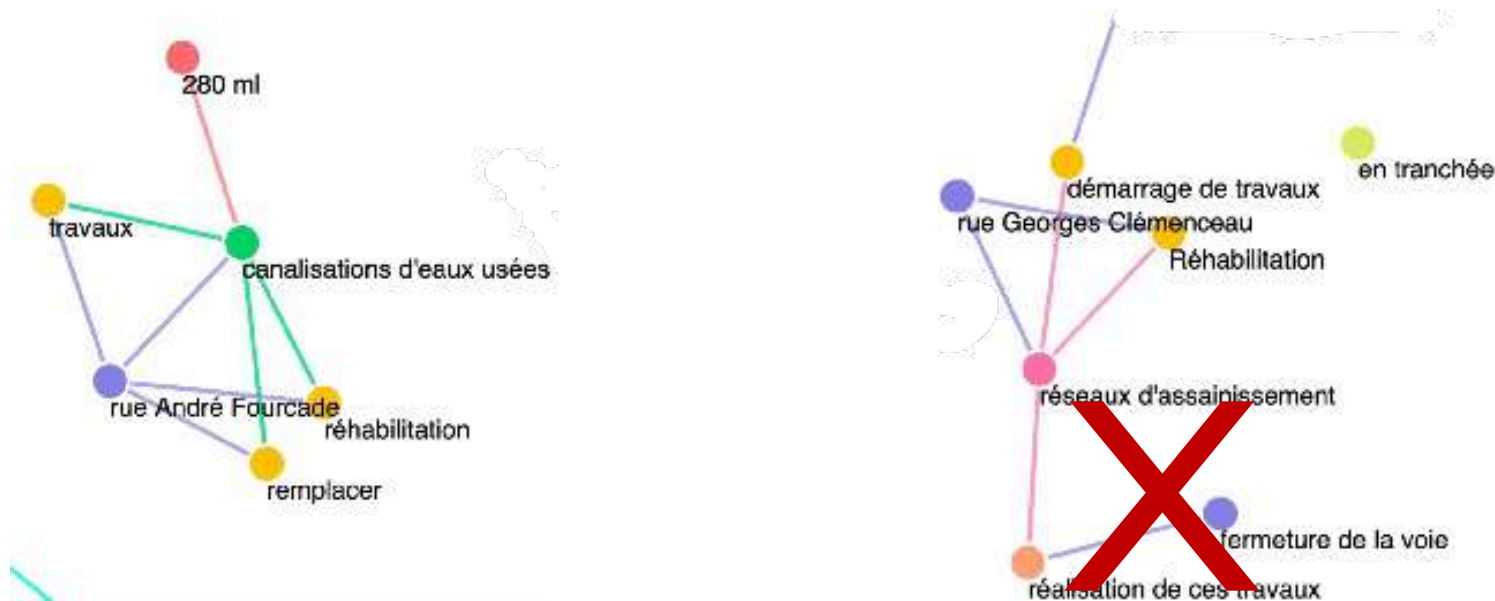
Chahinian et al., 2022

# ARCHITECTURE PLATEFORME WEIR-P\*



\*Wastewater Information extraction Platform

# RÉSULTATS : EXEMPLE DE MISE EN RELATION



Temporel

Spatial

Organisation

Indicateur

Travaux

Traitement

Accident

Type\_reseau

Element\_reseau

Gestion

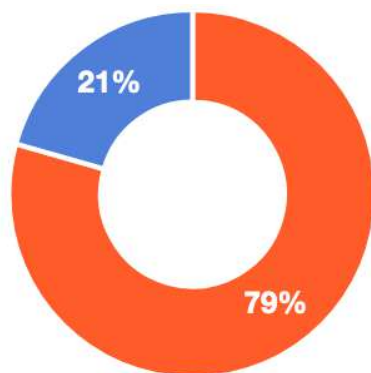
Mesure



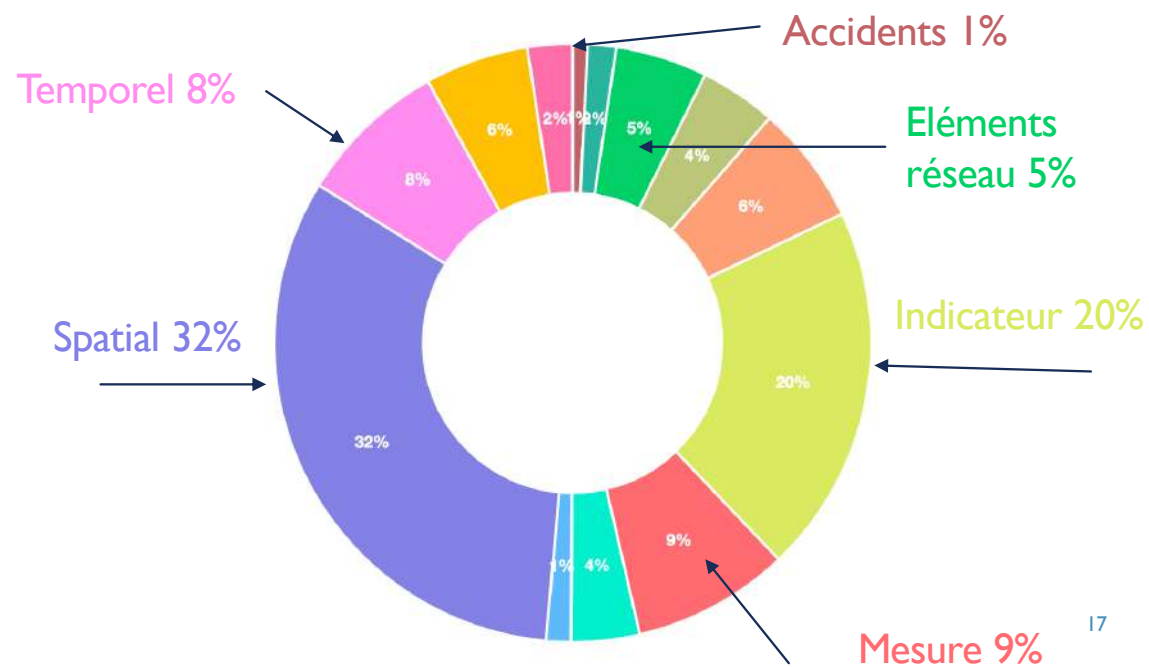
# VALIDATION : JACOU

Jacou (Hérault), collecté en juillet 2020, 728 documents, 59,318 EN, 42359 relations

## Pertinence



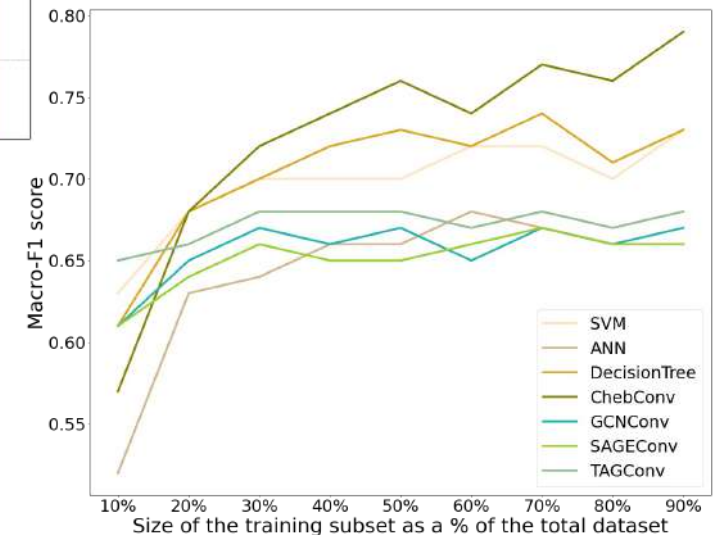
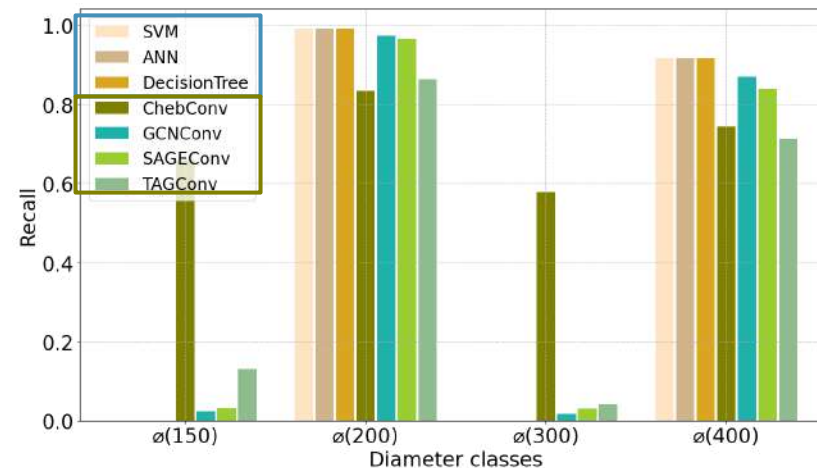
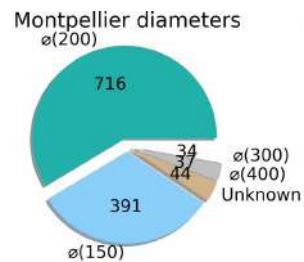
## Entités nommées



# MÉTHODES INNOVANTES POUR DÉTECTER ET DÉCRIRE LES RÉSEAUX D'EAUX URBAINS : BILAN

- Si résolution < 25 cm, détection directe des plaques sur des images très difficile. 1/3 des objets détectés à résolution < 40 cm par filtre géométrique (Bartoli et al., 2015)
- Possible de cartographier un réseau à partir des positions des plaques d'égouts et d'avaloirs
  - Si topologie du réseau non-respectée => procédure de correction (Belghaddar et al., 2022b)
- Résultats encourageants pour la fouille de données
  - Les données quantitatives moins mentionnées que les événements (Chahinian et al., 2019)
  - La plateforme peut être utilisée pour retrouver des dates mais il reste des améliorations à faire = > projets CROQUIS et STARWARS.
- Autres pistes prometteuses de l'IA : les GNNs pour combler les lacunes ou cartographier les réseaux (Belghaddar et al., 2021 et 2022)

# PERSPECTIVES : LES GNNs POUR COMBLER LES DONNÉES MANQUANTES



- 3 modèles non-topologiques

- 4 GNNs

= > Les classes minoritaires ne sont pas détectées par les modèles « classiques »

= > Chebconv a la meilleure performance quelle que soit la taille de la base d'entraînement.

Belghaddar et al., 2022

# REMERCIEMENTS



Merci aussi

- aux collectivités qui partagent leurs données en open data,
- aux mairies qui font des annonces de travaux sur le web,
- aux habitants qui postent, tweetent et râlent sur les réseaux d'eaux urbains...

