



**HAL**  
open science

## Gestion Patrimoniale Multi-Echelles des réseaux d'eau potable (GePaME) - Rapport intermédiaire 2021

Nicolas Rodriguez, Alain Husson, Eddy Renaud, Kevin Caillaud, Yves Le Gat, Bénédicte Rulleau

### ► To cite this version:

Nicolas Rodriguez, Alain Husson, Eddy Renaud, Kevin Caillaud, Yves Le Gat, et al.. Gestion Patrimoniale Multi-Echelles des réseaux d'eau potable (GePaME) - Rapport intermédiaire 2021. INRAE UR ETBX; SWDE. 2021. hal-03836820

**HAL Id: hal-03836820**

**<https://hal.inrae.fr/hal-03836820v1>**

Submitted on 2 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Projet GePaME

# Gestion Patrimoniale Multi-Echelles des réseaux d'eau potable

Rapport intermédiaire 2021



Nicolas Rodriguez, Alain Husson, Eddy Renaud, Kevin Caillaud, Yves Le Gat,  
Bénédicte Rulleau  
02/09/2021

## Table des matières

1	Introduction.....	6
1.1	Contexte .....	6
1.2	Contenu du projet GePaME.....	6
1.3	Gouvernance du projet .....	7
1.4	Contenu du rapport.....	8
2	Tâche 01 – Prise en main et traitement des données.....	9
2.1	Objectifs de la tâche 01 .....	9
2.1.1	Préambule .....	9
2.1.2	Première action de la tâche 01 : préparation des fichiers Casses.....	11
2.2	Détail des traitements mis en œuvre pour préparer les fichiers Casses.....	13
2.2.1	Traitement A : Concaténation des fichiers fuites .....	13
2.2.2	Traitement B : Analyse des fichiers de défaillances et de fuites détectées-réparées... 14	14
2.2.3	Traitement C : Extraction des diamètres des pièces de réparation liées aux OT.....	16
2.2.4	Traitement D : Préparation et complétion des conduites en service et abandonnées 16	16
2.2.5	Traitement E : Agrégation des conduites .....	18
2.2.6	Traitement F : Attribution d’ICAR de rue aux conduites .....	20
2.2.7	Traitement G : Affectation des défaillances aux conduites.....	24
2.2.8	Traitement H : Caractéristiques additionnelles des conduites .....	25
3	Tâche 05 - Exploitation des données de la sectorisation .....	27
3.1	Objectifs de la tâche 05 .....	27
3.2	Données caractéristiques des cellules.....	27
3.2.1	Caractéristiques patrimoniales.....	27
3.2.2	Caractéristiques géographiques .....	28
3.2.3	Caractéristiques d’exploitation .....	33
3.3	Débits des cellules .....	33
4	Avancées méthodologiques .....	35
4.1	Tâche 02 - Analyse de la gouvernance de la SWDE en matière de GPI.....	35
4.2	Tâche 07 - Modélisation de la probabilité d’un tronçon d’être fuyard.....	36
5	Conclusion .....	37
6	Annexes .....	38

## Figures

Figure 1 : Vue d'ensemble de l'articulation des tâches et des volets du projet. ....	7
Figure 2 : Planning prévisionnel (version du 24/02/2021). ....	7
Figure 3 : Traitements permettant de produire les fichiers d'entrée du logiciel Casses. ....	13
Figure 4 : Exemples de tronçons abandonnés (lignes rouges) au centre de Mons. ....	17
Figure 5 : Absence de connexité sur certaines conduites HS, aperçue en zoomant. ....	17
Figure 6 : Exemple de tronçons de petite taille localisés dans un carrefour. ....	19
Figure 7 : Exemple d'OT (au centre) dont la conduite concernée peut être ambiguë. ....	21
Figure 8 : Exemple d'affectation des ICAR de rue selon les méthodes par sommets et par tampons. ....	22
Figure 9 : Synoptique des traitements appliqués sur les conduites du SIG. ....	23
Figure 10 : Synoptique des traitements utilisés pour lier chaque défaillance à une conduite du SIG. ....	25
Figure 11 : Synoptique des calculs d'indicateurs pour les conduites et les cellules. ....	26
Figure 12 : Proportions de la longueur de réseau par année de pose. ....	28
Figure 13 : Répartition des 5 classes CLC et du bâti autour de Mons. ....	29
Figure 14 : Exemple d'intersection des routes et trottoirs avec les canalisations. ....	32
Figure 15 : Reconstitution des débits de la cellule 52300016. ....	33
Figure 16 : Articulation de tâches du volet 1. ....	38
Figure 17 : Articulation de tâches du volet 2. ....	39
Figure 18 : Articulation de tâches du volet 3. ....	39
Figure 19 : Exemple de conduite erronée du SIG (superposition). ....	43
Figure 20 : Juxtaposition des graphes des conduites HS (rouge) et ES (noir). ....	46
Figure 21: Illustration des nœuds doubles. ....	49
Figure 22 : Illustration des nœuds triples. ....	49
Figure 23 : Tronçon dont les angles globaux avec ses voisins diffèrent nettement des angles locaux. ....	51
Figure 24 : Tronçons « nœud simple-nœud triple » exclus à l'étape 4. ....	51
Figure 25 : Conduites interceptées ou non par un tampon autour des axes de voirie. ....	52
Figure 26 : Exemple de conduite agrégée qui change d'ICAR de rue du fait de son agrégation. ....	54
Figure 27 : Exemple de projection des couches géographiques relatives au bâti (Cellule 52300041). ....	64

## Tableaux

Tableau 1 : Tâches du projet GePaME.....	6
Tableau 2 : Recrutement des personnels temporaires .....	8
Tableau 3 : Synthèse des données SWDE identifiées pour conduire le projet. ....	9
Tableau 4 : Fichiers concernant les défaillances enregistrées par la SWDE. ....	11
Tableau 5 : Croisement des OT « fuites » et types d'activités (extrait SAP du 27/11/2020). ....	15
Tableau 6 : Résultats de l'agrégation pour les conduites en service. ....	20
Tableau 7 : Statistiques d'attribution d'ICAR de rue aux conduites en fin de traitement. ....	22
Tableau 8 : Construction des classes CLC.....	29
Tableau 9 : Classes de répartition des sols à risque pour les canalisations. ....	32
Tableau 10 : Classes d'usage des sols liées à la couverture des conduites.....	32
Tableau 11 : Tableau de correspondance des colonnes des fichiers fuites. ....	40
Tableau 12 : Extraction des diamètres associés aux pièces de réparation liées aux OT.....	41
Tableau 13 : Correspondance entre Diamètre Nominal (mm) et Nominal Pipe Size (pouces).....	42
Tableau 14 : Conduites erronées du secteur de Mons et corrections apportées.....	44
Tableau 15 : Détail des quatre étapes de l'agrégation des conduites. ....	48
Tableau 16 : Décompte du nombre de tronçons à chaque étape de l'agrégation. ....	51
Tableau 17 : Désignations de matériaux considérées comme équivalentes. ....	62
Tableau 18 : Méthodes d'affectation des défaillances sur conduite pour le secteur de Mons.....	62
Tableau 19 : Nombre de branchements obtenu par points d'adresse et données SWDE.....	65

## Liste des sigles et abréviations

Abréviation	Objet	Description
AEP	Alimentation en Eau Potable	
BDD (BD)	Base De Données	
BI	Borne Incendie	
CR	Compte Rendu	
CSO	Comité de Suivi Opérationnel	Réunion périodique de suivi du projet.
CLC	Corine Land Cover	Occupation du sol
CT	Court Terme	
DDC	Date De Casse	Date à laquelle l'élément du réseau AEP (comme la conduite) a cassé.
DDP	Date De Pose	Date à laquelle la conduite a été posée.
DE	Diamètre Extérieur	diamètre extérieur (mm) d'un tuyau ou d'une conduite.
DHS	Date de mise Hors Service	Date à laquelle la conduite a été mise hors service, ou abandonnée. La DHS de la canalisation est une donnée nécessaire à l'implémentation du modèle LEYP2s.
DIA	Diamètre	Selon le contexte, peut désigner DN ou DE.
DN	Diamètre Nominal	diamètre nominal (mm) d'un tuyau ou d'une conduite.
EPSG	European Petroleum Survey Group	
ES	En Service	Utilisé pour décrire des conduites qui sont encore utilisées à l'heure actuelle pour l'AEP.
FG	Fonte Grise	
GPI	Gestion Patrimoniale des Infrastructures	
HS	Hors Service	Synonyme de « abandonné », utilisé pour les conduites hors service.
ICAR	Inventaire Centralisé des Adresses et des Rues en Wallonie	Système utilisé par le SPW pour identifier de façon unique chaque rue ou point d'adresse par un numéro.
ID	Identifiant	Code utilisé dans un fichier pour identifier de façon unique chaque objet. L'OT SAP est un type d'ID.
INS (code)	Institut National de la Statistique	Code à 5 caractères unique pour chaque commune en Belgique.
LEYP	Linear Extended Yule Process	Modèle probabiliste à base du processus de Yule utilisé pour la prédiction des défaillances (fuites et casses) du réseau AEP (Le Gat, 2009).
LT	Long Terme	
MAT	Matériau	Matériau de l'élément du réseau AEP.

<b>MNT (DTM)</b>	<b>Modèle Numérique de Terrain</b>	Représentation numérisée de la topographie du terrain de surface.
<b>NPS</b>	<b>Nominal Pipe Size</b>	Diamètre de tuyau ou de conduite exprimé en pouces.
<b>OT</b>	<b>Ordre de Travail</b>	Numéro utilisé dans SAP pour identifier de façon unique chaque intervention.
<b>PE</b>	<b>Polyéthylène</b>	
<b>PICC</b>	<b>Projet Informatique de Cartographie Continue</b>	Cartographie des bâtiments de la Wallonie.
<b>PoWalCo</b>	<b>Portail Wallon de Coordination</b>	Portail internet permettant l'échange d'informations sur les intentions de travaux, la coordination des chantiers et la gestion des autorisations en Wallonie.
<b>PVC</b>	<b>Polychlorure de Vinyle</b>	
<b>QGIS</b>	<b>Quantum Geographic Information System</b>	Logiciel de SIG.
<b>RF</b>	<b>Recherche de Fuite</b>	
<b>SAP</b>	<b>Systems Applications and Products in Data Processing</b>	Progiciel de gestion intégré utilisé par la SWDE notamment pour archiver les interventions.
<b>SCR</b>	<b>Système de Coordonnées de Référence</b>	Informations mathématiques sur la méthode de projection de géométrie sphérique sur un plan.
<b>SI</b>	<b>Système d'Information</b>	Système (logiciel) utilisé pour stocker et traiter l'information.
<b>SIG</b>	<b>Système d'Information Géographique</b>	Système (logiciel) utilisé pour stocker et traiter l'information géographique.
<b>SPW</b>	<b>Service Public de Wallonie</b>	Administration publique Wallonne.

# 1 Introduction

## 1.1 Contexte

Partant du constat de leur intérêt commun à rechercher des réponses scientifiques et opérationnelles aux questions soulevées par la gestion patrimoniale des réseaux d'eau potable, SWDE et INRAE ont décidé de s'associer pour conduire des travaux de recherche dans ce domaine. Ils ont ainsi conclu, le 9 juin 2020, un contrat de recherche qui établit une collaboration pour mener à bien le projet « Gestion Patrimoniale Multi-Echelles des réseaux d'eau potable » (GePaME). Le projet GePaME, d'une durée prévisionnelle de trois ans, a démarré le 1<sup>er</sup> septembre 2020. Le contrat prévoit que les résultats du projet sont présentés dans des rapports annuels établis par l'INRAE.

**Le présent rapport est le premier rapport intermédiaire du projet GePaME, il retrace les résultats obtenus au cours de la première année de réalisation du projet.**

## 1.2 Contenu du projet GePaME

Le projet GePaME est organisé en trois volets :

- Volet 1 : compréhension du processus de fuite et efficacité de la recherche de fuite (RF) ;
- Volet 2 : aide à la sélection multi-objectifs et contrainte des chantiers de renouvellements ;
- Volet 3 : simulation long terme de stratégies de gestion patrimoniale.

Il se décompose en treize tâches opérationnelles (Tableau 1) qui doivent aboutir à des avancées scientifiques valorisées par l'INRAE et à des outils d'aide à la décision qui ont vocation à être intégrés dans les pratiques et dans le système d'information de la SWDE.

Tableau 1 : Tâches du projet GePaME.

Tâche	Dénomination	Référent SWDE	Référent INRAE
01	Prise en main et traitement des données	F. Haan	A. Husson
02	Analyse de la gouvernance de la SWDE en matière de GPI	D. Brunet	K. Caillaud
03	Implémentation de Casses	J-G. Baudoin	Y. Le Gat
04	Intégration des outils dans le système d'information	J-G. Baudoin	Y. Le Gat
05	Exploitation des données de la sectorisation	D. Devos	A. Husson
06	Analyse des pratiques de recherche de fuites	O. Pinsar	K. Caillaud
07	Modélisation de la probabilité d'un tronçon d'être fuyard	O. Pinsar	Y. Le Gat
08	Analyse de la vulnérabilité des usagers	D. Brunet	B. Rulleau
09	Modélisation statistique couplée des fuites et des crashes	O. Pinsar	Y. Le Gat
10	Outil multicritère de programmation court terme	D. Brunet	Y. Le Gat
11	Efficacité des actions de recherche de fuite	O. Pinsar	E. Renaud
12	Etude de scénarios long terme et simulation	D. Brunet	Y. Le Gat
13	Coordination du Projet et valorisation	D. Brunet	E. Renaud

La Figure 1 présente une vue d'ensemble de l'articulation des tâches en cohérence avec les trois volets du projet. Une vision détaillée par volet est proposée en Annexe 1.

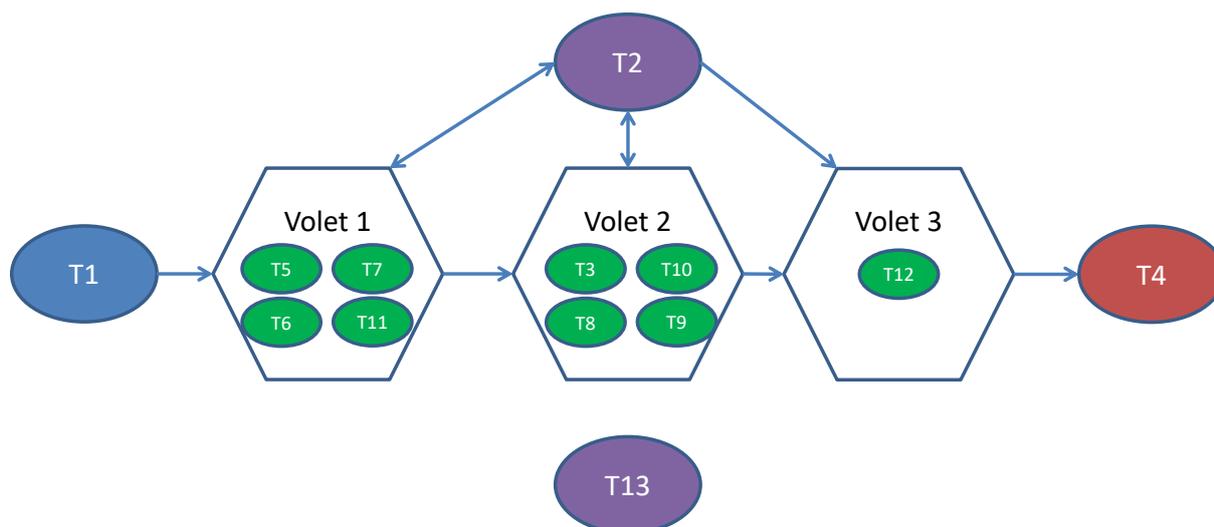


Figure 1 : Vue d'ensemble de l'articulation des tâches et des volets du projet.

### 1.3 Gouvernance du projet

La gouvernance du projet est assurée par deux instances :

- le comité de pilotage stratégique (CPS) : composé de représentants des directions de SWDE et de l'unité « Environnement territoires et infrastructures » d'INRAE, il veille au bon déroulement du projet et peut proposer des réorientations des recherches et/ou des aménagements du contrat et se réunit une fois par an ;
- le comité de suivi opérationnel (CSO) : composé des responsables de l'exécution du projet (dont les référents mentionnés dans le Tableau 1), il assure le suivi et l'animation du programme scientifique et se réunit une fois par mois.

Au cours de la première année d'exécution du projet, le CSO s'est réuni six fois en visioconférence. La première réunion du CPS est prévue d'ici fin 2021.

Le planning du projet est adapté par le CSO en fonction de l'avancement réel des travaux et des contraintes connues ou prévisibles qui peuvent l'influencer. La Figure 2 présente le planning prévisionnel dans sa version actualisée du 24 février 2021.

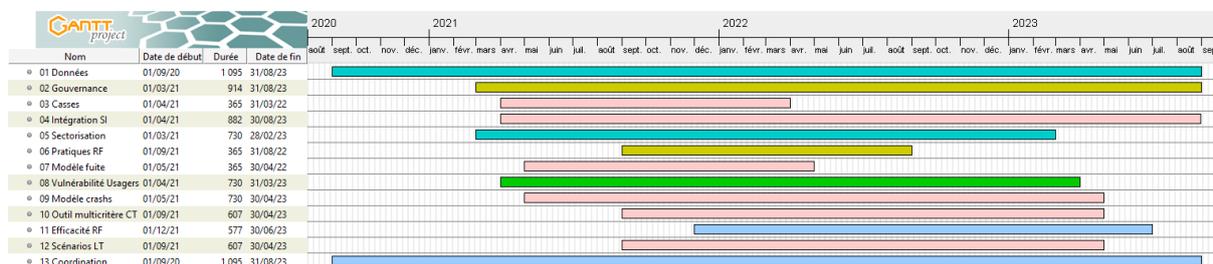


Figure 2 : Planning prévisionnel (version du 24/02/2021).

Le contrat de recherche prévoit que, pour conduire le projet, l'INRAE recrute quatre agents temporaires. Le Tableau 2 fait le point des recrutements réalisés et à venir.

Tableau 2 : Recrutement des personnels temporaires

Profil	Mois	Personne recrutée	Début	Fin
Ingénieur de recherche en eau potable	24	Nicolas Rodriguez	01/12/2020	30/11/2022
Post doctorant en sociologie	12	Selma Baati	01/09/2021	31/08/2022
Post doctorant en statistiques	24	Marie Collet	01/09/2021	31/08/2023
Ingénieur d'étude en économie	6	<i>Prévu en 2022</i>	<i>2022</i>	

## 1.4 Contenu du rapport

Au cours de l'année écoulée, les travaux se sont principalement focalisés sur l'interconnaissance entre les parties prenantes, la préparation des données et le cadrage méthodologique des tâches amont.

En conséquence, ce rapport intermédiaire présente les résultats obtenus dans la réalisation des tâches 01 « Prise en main et traitement des données » et 05 « Exploitation des données de la sectorisation » puis les choix méthodologiques proposés pour les tâches 02 « Analyse de la gouvernance de la SWDE en matière de GPI » et 07 « Modélisation de la probabilité d'un tronçon d'être fuyard ».

## 2 Tâche 01 – Prise en main et traitement des données

### 2.1 Objectifs de la tâche 01

#### 2.1.1 Préambule

Le contrat de collaboration passé entre la SWDE et l'INRAE définit la tâche 01 comme suit : « Identification, analyse, sélection, collecte, prétraitement et traitement des données des systèmes d'information ».

Pour des raisons pratiques, la SWDE et l'INRAE ont convenu, dans un premier temps, de se concentrer sur **le secteur de MONS** qui est une zone assez représentative des différentes configurations que l'on peut rencontrer sur le territoire couvert par la SWDE. Ainsi **les premières données mobilisées pour produire les résultats présentés dans ce rapport, concernent le seul secteur de Mons**. Il est attendu que les méthodes développées pour traiter ces données sur Mons soient adaptables à l'ensemble de la SWDE.

Compte tenu des données disponibles, les périodes prises en compte sont :

- 2011-2021 pour l'analyse des chroniques de défaillances du réseau (fuites et crashes). Cet intervalle résulte du croisement des données de défaillances qui couvrent la période 2007-2021 avec l'archivage des conduites mises hors service dans le SIG de la SWDE qui se met en place de manière stable à partir de 2011 environ.
- 2017-2021 pour l'analyse des débits. En effet, à partir de 2017 les secteurs hydrauliques (cellules) ont été déployés au sein de la SWDE et les données de débits résultantes rapatriées et archivées. De plus, il a été convenu entre l'INRAE et la SWDE, à ce stade d'avancement du traitement des données, de stopper la chronique de données utilisées en 2021 afin d'avoir suffisamment de temps pour exploiter ces données les plus récentes.

Le Tableau 3 récapitule de façon synthétique les données sources identifiées pour conduire le projet, les principales tâches concernées et l'avancement de leur transmission par la SWDE.

**Tableau 3 : Synthèse des données SWDE identifiées pour conduire le projet.**

Tâche	n°	Données requises	Détails	Données fournies	Commentaires
Analyse de la gouvernance de la SWDE en matière de GPI	02	Informations sur la gouvernance	Législation, organisation, rapports d'activité, description des outils...	Partiellement	A compléter lors de l'enquête de terrain.
Implémentation de Cassettes	03	Réseau AEP	Conduites en service. Conduites mises hors service.	Oui	Absence de relation entre nœuds et tronçons. Complétion de données nécessaire pour les conduites HS.
Implémentation de Cassettes	03	Chronique des défaillances	Date, ID de tronçon, méthode d'identification, cause du problème.	Oui	Absence de relation directe entre défaillances et tronçons.
Exploitation des données de sectorisation	05	Structure des cellules	Identifiants de cellules. Compteurs et vannes de sectorisation. Évolution des cellules dans le temps.	Oui	Chronique d'évolution du linéaire de canalisations par cellule à l'échelle annuelle et à partir de 2017 seulement.

Tâche	n°	Données requises	Détails	Données fournies	Commentaires
Exploitation des données de sectorisation	05	Débits de cellules	Données brutes de volumes ou débits des compteurs de sectorisation. Débits ou volumes (pré)traités. Formules de calcul des débits de cellule.	Partiellement	Extraction des données en cours par Olivier Hautier et Guy Deschuyteneer.
Exploitation des données de sectorisation	05	Abonnés et consommations	Nombre, type, localisation d'abonnés/habitants. Consommations des abonnés.	Partiellement	Nombre de raccordements compté par rapport au nombre de bâtis. Pas de données de consommations encore transmises.
Exploitation des données de sectorisation	05	Pose et renouvellement de canalisations	Date et localisation des grandes opérations de renouvellement. Lien entre conduites remplacées et remplaçantes.	Partiellement	Type d'activité 100 dans SAP. Pas de lien direct avec les conduites (lien avec la rue seulement).
Exploitation des données de sectorisation	05	Interventions	Opérations sur réseau pouvant impacter les débits (réparations, purges, essais...)	Partiellement	Types d'activités spécifiques dans SAP. Lien avec la conduite concernée à établir (comme pour la tâche 03).
Exploitation des données de sectorisation	05	Actions de lutte contre les pertes	Facteur déclenchant, technique, date, résultat.	Non	Hormis la RF, aucune action spécifique n'est précisée.
Exploitation des données de sectorisation	05	Environnement des conduites	Type de sol, réseau routier, hauteur de bâti, géologie, aléa argile... Pression de service	Oui	Données publiques éventuellement à récupérer sur sites officiels de Belgique. Données de pression sur quelques hydrants et réducteurs de pression seulement.
Analyse des pratiques de RF	06	Informations sur les moyens humains et matériels de la RF	Organisation des activités de RF (service(s)/unité(s), équipe(s), personnel mobilisé, matériel disponible et utilisé.	Partiellement	Pas d'informations spécifiques sur le matériel disponible et utilisé. L'enquête de terrain devrait apporter des compléments.
Modélisation de la probabilité d'un tronçon d'être fuyard	07	Recherche de fuite	Date, tronçon(s) inspecté(s), méthode, résultat.	Oui	Le lien avec le résultat de la RF et les réparations éventuelles reste à établir.
Analyse de la vulnérabilité des usagers	08	Informations sur les plaintes des usagers	Localisation, cause, conséquences. Informations de contact des abonnés.	Non	
Efficacité des actions de recherche de fuite	11	Gestion des pressions	Appareils, consignes, pressions mesurées	Oui	Données de pression sur quelques hydrants et réducteurs de pression seulement.

Les objectifs de la tâche 01 peuvent être reformulés comme suit :

- Objectif 1 : Élaborer une méthode (automatisée) de préparation des données sur le secteur de Mons qui pourra être rapidement étendue à l'ensemble du territoire couvert par la SWDE.
- Objectif 2 : Constituer une base de données du réseau, des cellules, des pertes en eau, des défaillances et interventions, et des actions de réduction des pertes.
- Objectif 3 : Faciliter l'établissement de relations entre les différentes sources de données (défaillances-caractéristiques du réseau, pertes en eau-défaillances, pertes en eau-environnement du réseau, pertes en eau-réduction des pertes...).

### 2.1.2 Première action de la tâche 01 : préparation des fichiers Cassettes

Pour la première année du projet, les travaux de la tâche 01 ont eu pour but principal de produire les données nécessaires à la réalisation de la tâche 03 : implémentation de Cassettes. En effet, pour fonctionner, le logiciel Cassettes a besoin d'au moins deux fichiers : un fichier de tronçons qui comporte les identifiants, ID, et des caractéristiques obligatoires des tronçons (matériau, longueur, diamètre, date de pose, et si possible date de mise hors service), et un fichier des cassettes subies par ces mêmes tronçons (i.e., reliées à l'ID) et qui mentionne obligatoirement la date de casse (DDC). Bien que très informatives sur ces aspects, les données brutes de la SWDE ne permettent pas de créer directement ces fichiers, car au sein des données de défaillances il n'y a que rarement un lien direct entre la défaillance constatée (fuite réparée après RF, crash) et l'ID de la conduite concernée dans le SIG. Ainsi, les sources de données permettant d'identifier et de caractériser les défaillances sont multiples (Tableau 4). **C'est pourquoi, la première action de la tâche 01 a été de lier explicitement toutes les défaillances constatées aux conduites, dans le but de créer des fichiers des Tronçons et des Cassettes cohérents utilisant les archives les plus complètes possibles.**

Tableau 4 : Fichiers concernant les défaillances enregistrées par la SWDE.

Type de donnée	Date de début d'enregistrement	Nombre d'enregistrements uniques (date du fichier)	Type de localisation du problème	Commentaire
SAP (tableau)	Août 2007	> 27 000 (18/03/2021)	(Rue, numéro d'immeuble) (Rue)	Source la plus complète de données sur les défaillances.
Anomalies (SIG)	Mai 2016 (Liaison SAP le 30/10/2018)	> 600 (22/02/2021)	ID de conduite dans le SIG Point SIG (X, Y)	Une partie des enregistrements a été faite avant la liaison SAP. Depuis la liaison, toutes les anomalies SIG sont aussi dans SAP.
Fichiers Fuites (tableau)	Janvier 2016	> 1 600 (13/01/2021)	(Rue, numéro d'immeuble) (Rue)	Presque toutes les fuites enregistrées dans ces fichiers le sont aussi dans SAP.
Pièces de réparation (tableau)	Août 2007	> 14 000 (25/03/2021)	Non applicable	Chaque pièce est associée à une défaillance de SAP.

Les données de SAP ne contenant aucune information sur la conduite concernée par une défaillance, il est nécessaire de les croiser avec des sources de données qui incluent des caractéristiques du réseau :

- Les fichiers fuites (date de réparation, matériau, et diamètre des conduites ou branchements).
- Les pièces de réparation (diamètre, et éventuellement matériau des conduites ou branchements).
- Le fichier de conduites du SIG (matériau, diamètre, date de pose, date d'abandon, classe, zone de distribution d'eau...). C'est sans nul doute l'information la plus fiable et complète concernant les caractéristiques des conduites.

La démarche mise en œuvre pour lier les ID de conduites du SIG aux défaillances, basée sur l'analyse globale de l'ensemble des données décrites ci-dessus, est la suivante :

1. Utiliser l'extrait SAP comme source de chronique des défaillances. Au sein de ce fichier, chaque Ordre de Travail (OT) est un numéro unique associé à chaque défaillance.
2. Tenter d'identifier la conduite du SIG concernée par l'OT SAP grâce aux anomalies SIG (ce qui est presque immédiat), ou sinon grâce aux informations de localisation dans SAP et/ou dans

les fichiers fuites (adresses postales). Associer les caractéristiques techniques des conduites liées aux défaillances en priorité à partir des informations du SIG.

3. Utiliser les informations techniques des fichiers fuites et/ou des pièces de réparation pour aider l'identification de la conduite concernée par l'OT si elle ne ressort pas de façon évidente.

**Cette démarche nécessite de nombreuses étapes de traitements intermédiaires qui sont détaillés dans les sections suivantes.** De manière résumée, les verrous suivants ont pu être identifiés par l'analyse initiale des différents fichiers :

- Comment trouver de façon fiable la conduite associée à la défaillance lorsque seule une adresse postale est renseignée ? Comment gérer par exemple les adresses de défaillances se situant aux carrefours avec de nombreuses petites conduites proches ?
- Comment trouver la conduite de façon fiable lorsque seule une rue est renseignée pour l'intervention (sachant qu'une rue peut comporter de nombreuses conduites) ?
- Dans quelle rue se situe une conduite donnée ? Comment le déterminer de façon fiable dans les carrefours ?
- Comment doit-on traiter les conduites abandonnées afin de les rendre équivalentes aux conduites en service pour l'attribution des défaillances passées, qui peuvent avoir affecté une conduite mise hors service depuis ?
- De quelle manière doit-on considérer les nombreuses petites conduites du SIG (par exemple, de moins de 6 mètres) dans les traitements intermédiaires et pour l'attribution des défaillances ?
- De quelle façon exactement doit-on utiliser les matériaux et diamètres renseignés dans les fichiers fuites et les pièces de réparation pour aider l'identification de la conduite concernée par une défaillance ?
- Comment combiner les multiples fichiers fuites en un seul fichier directement utilisable dans tous les traitements suivants ?

Les réponses à ces questions font l'objet de traitements synthétisés dans la Figure 3 ci-après et présentés en détail dans la suite.

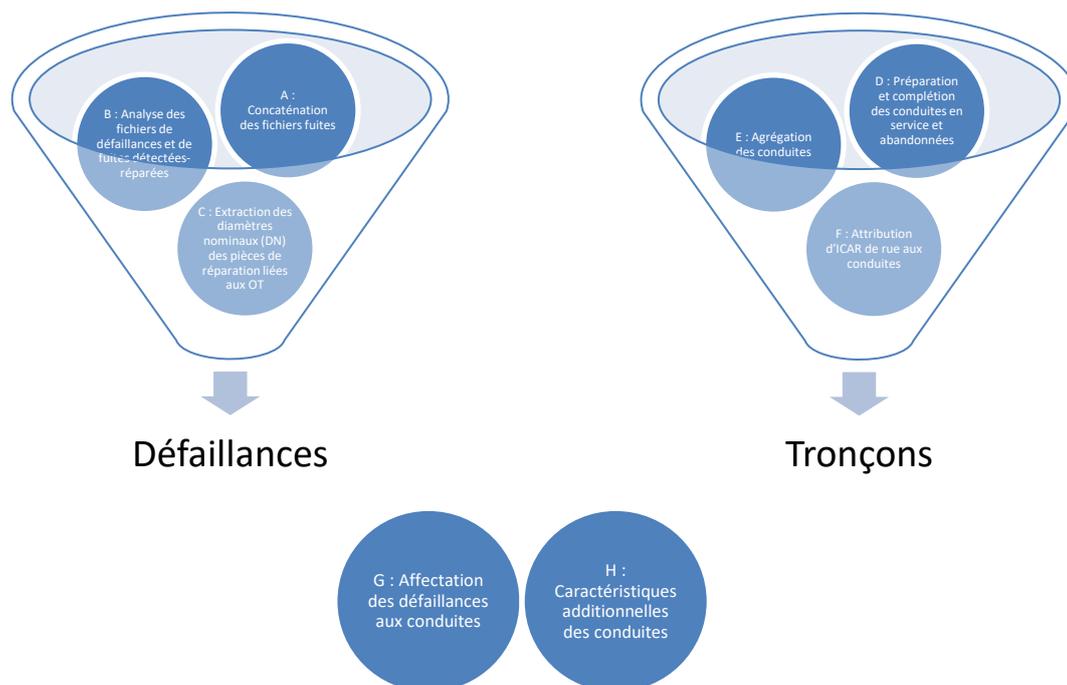


Figure 3 : Traitements permettant de produire les fichiers d'entrée du logiciel Cassees.

## 2.2 Détail des traitements mis en œuvre pour préparer les fichiers Cassees

### 2.2.1 Traitement A : Concaténation des fichiers fuites

En parallèle de la saisie des OT dans SAP, les fuites détectées et réparées sont en partie enregistrées au sein de la SWDE dans un ensemble de fichiers de type Excel dénommés ici « fichiers fuites ». Pour le secteur de Mons, il existe un fichier par année (années 2016 à 2020). Pour 2016 et 2017, les mois sont séparés par onglet alors que les années 2018, 2019, et 2020 un seul onglet couvre toute l'année en distinguant le mois dans une colonne supplémentaire par rapport aux fichiers précédents. Essentiellement, les fichiers fuites permettent de savoir où et quand certaines fuites ont été détectées puis réparées (ce qui est déjà renseigné dans SAP), mais également sur quel type de conduite (matériau et diamètre, sans pour autant savoir laquelle exactement). Ainsi, près de 70 % des OT de type réparation de conduite dans SAP sont aussi dans les fichiers fuites, qui contiennent souvent les diamètres et matériaux des conduites concernées (matériaux renseignés à plus de 85 %, diamètres renseignés à plus de 28 % pour les données de Mons). Les fichiers fuites sont donc une source utile d'informations techniques sur les conduites pour compléter ce qui n'est pas renseigné dans SAP.

Certaines colonnes ont été ajoutées, supprimées, ou modifiées au cours des mois et des années dans ces fichiers. Cela est particulièrement visible dans le fait que les données de janvier 2016 utilisent 18 colonnes, alors que les données de l'année 2020 en utilisent 49. **Afin de pouvoir utiliser l'ensemble des données des fuites enregistrées sur toutes les années et tous les mois de façon simplifiée, il a été nécessaire de concaténer (c'est-à-dire, de mettre bout à bout) toutes les fuites enregistrées dans ces fichiers, et ce avec les mêmes intitulés de colonnes.** Pour cela, il faut utiliser une table de correspondance (voir Annexe 2) entre les indices (le numéro de position) de colonnes de chacun des fichiers (chacun des onglets pour les années 2016 et 2017) et les indices de colonne du fichier concaténé recevant les données brutes. Cette table est basée sur la comparaison visuelle des

intitulés de colonne, qui a permis de détecter lorsqu'une colonne est ajoutée, supprimée, ou qu'elle change de position (d'indice) au cours du temps. Le fichier de fuites concaténé contient ainsi tous les intitulés de colonnes qui apparaissent dans l'ensemble des fichiers fuites. Enfin, les colonnes qui n'existaient pas pour certains mois ou certaines années sont laissées vides pour les fuites concernées dans le fichier final. Cette méthode sera à adapter à chaque secteur d'exploitation suivant la relation entre colonnes dans les données brutes et colonnes dans le fichier concaténé.

De cette manière, toutes les informations des fichiers fuites sont regroupées dans un seul fichier où il est possible de rapidement recueillir la date de réparation de la fuite ainsi que le matériau et diamètre de la conduite concernée lorsque l'on a besoin de ces informations pour la distinguer parmi plusieurs conduites possibles du SIG. Pour le secteur de Mons, ce fichier représente plus de 1 600 fuites pour les années 2016 à 2020, soit une vingtaine de fuites par mois.

Il est à noter que la colonne OT a été assimilée à l'OT de Recherche de Fuite (OT RF) pour les fichiers de 2016 et 2017 de Mons, qui sont les seuls à distinguer l'OT RF de l'OT de réparation. Les fichiers de 2018, 2019 et de 2020 n'ont que l'OT de réparation (intitulé OT), donc leur colonne OT RF est laissée vide. Il est également à noter que l'identifiant des OT est enregistré dans les fichiers fuites en tant que texte libre (contenant potentiellement des erreurs), alors que dans SAP c'est la clé primaire de type numérique.

Enfin, cette première analyse a été l'occasion de relever des erreurs dans les fichiers fuites, plus précisément :

- Des doublons d'OT (une cinquantaine).
- Des fautes de frappes dans les OT (une cinquantaine).
- Des OT texte ou vides (près de 140).

Les OT texte ou vides sont supprimés. Les OT doublons ne sont pas supprimés à l'avance car parmi les doublons, certains possèdent des informations différentes (comme des adresses différentes, ou des matériaux de conduite différents) qu'il sera utile d'exploiter plus tard. Il sera alors important de se souvenir de ne pas compter les doublons d'OT comme deux défaillances distinctes.

Remarque 1 : Il a été nécessaire d'extraire de ces fichiers les informations numériques sur les diamètres en supprimant le texte éventuellement associé (particulièrement pour 2016 et 2017, où le matériau est renseigné dans le même champ que le diamètre). Pour cela, le DN extrait est toute expression de la forme « Chiffre(s)PonctuationChiffre(s) », ce qui permet de récupérer à la fois les expressions du type « 80 » (DN80), comme du type « 3/4 » (diamètres exprimés en Nominal Pipe Size).

Remarque 2 : la saisie des fuites dans les fichiers sources étant manuelle, un certain nombre d'expressions similaires peuvent désigner le même matériau (par exemple FG et Fonte Grise). Un formatage automatique a donc été appliqué (utilisation du même nom en majuscules évitant les abréviations ambiguës).

### 2.2.2 Traitement B : Analyse des fichiers de défaillances et de fuites détectées-réparées

Le travail précédent nous a rapidement incités à comparer les informations de SAP et des fichiers fuites concernant les défaillances du réseau de distribution. **En particulier, il s'agissait de comprendre quels sont précisément les OT de SAP qui concernent des fuites ou des casses de**

**canalisations, en vue de la préparation des fichiers Casses (tâche 03).** En effet, les OT de SAP sont associés à un type d'activité spécifique (comme l'entretien des bornes/poteaux incendie, la réparation ou le renouvellement de conduites, etc.). Cette comparaison entre fichiers a été effectuée dans Excel en comptant combien d'OT de chaque type d'activité sont présents dans les fichiers fuites.

Il est rapidement ressorti de l'analyse que l'OT est la clé commune des données de défaillances, qui permet de lier une fuite détectée et reportée dans les fichiers fuites à l'opération de réparation enregistrée dans SAP (ainsi qu'à d'autres actions de lutte contre les pertes comme la Recherche de Fuites, mais aussi aux pièces de réparation). La comparaison quantitative a permis de mettre en évidence les types d'activités dans SAP (comme 105 : réparation de conduite) qui sont pertinents pour l'étude des chroniques de défaillances du réseau de distribution. Ces types d'activité ont été sélectionnés sur la base de la présence de l'OT associé au type d'activité en grande proportion dans les fichiers fuites. Les types d'activités et leur proportion dans les fichiers fuites sont présentés dans le Tableau 5.

Tableau 5 : Croisement des OT « fuites » et types d'activités (extrait SAP du 27/11/2020).

Type d'activité	Description	Taux de représentation dans les fichiers fuites
10	Placement BI/PI	< 1 %
11	Remplacement BI/PI	9 %
13	Entretien/réparation BI/PI	< 1 %
30	Problème install. intérieure	< 1 %
31	Autres problèmes/demandes clt	< 1 %
32	Remplac. non systém. de compt.	< 1 %
50	Remplacement raccordement Pb	3.8 %
72	Fin de service	< 1 %
82	Remplacement de raccordement	7 %
<b>92</b>	<b>Réparation raccordement</b>	<b>40 %</b>
100	Renouvellement de conduite	< 1 %
<b>105</b>	<b>Réparation conduite</b>	<b>31 %</b>
112	Placement d'appareil	< 1 %
113	Remplacement d'appareil	1 %
114	Réparation d'appareil	4 %
180	Recherche de fuite	< 1 %
185	RF systématique	< 1 %

De façon intuitive et logique, les types d'activités 92 (réparation de raccordement) et 105 (réparation de conduite) sont les plus pertinents pour l'étude et la prédiction des défaillances du réseau de distribution. En particulier, le type d'activité 105 fournit directement les informations de fuites et casses sur conduites, qui est la donnée principale pour l'utilisation du logiciel Casses (tâche 03). **Ainsi, seuls les OT associés au type d'activité 105 sont utilisés pour la création des fichiers Casses dans un premier temps.**

Remarque : les OT de type 92 (réparation de raccordement) seront par la suite également rattachés aux conduites pour envisager le calage de modèles de prédiction des défaillances sur raccordements subies par les conduites. Enfin, de même que les autres types d'activité présents de façon significative dans les fichiers fuites (11, 50, 82, 113, 114) ces OT seront utilisés pour analyser les variations de débits à l'échelle des cellules.

### 2.2.3 Traitement C : Extraction des diamètres des pièces de réparation liées aux OT

Pour les OT de type 105 qui ne sont pas dans les fichiers fuites (66 % de ces OT), une autre source d'informations techniques sur les conduites reste néanmoins disponible (autre que le SIG). Il s'agit des pièces de réparations liées à l'OT. Parmi ces dernières, on distingue entre autres des pièces assez informatives sur la conduite concernée (grâce à un dictionnaire fourni par la SWDE) :

- Des tuyaux de remplacement
- Des pièces de réparation (comme les manchons de réparation)
- Des pièces de raccord (raccords autobloquants, joints, manchons de liaison, tés, vannes...)
- Des pièces de pose de branchement

Le dictionnaire contient plus de 1 200 pièces, et fournit des informations sur les diamètres et les matériaux des pièces et/ou du réseau recevant les pièces. **L'enjeu était donc de comprendre quelles informations sur les conduites on peut extraire des pièces de réparation, dans le but d'utiliser ces dernières lorsqu'il est difficile d'identifier la conduite du SIG concernée par une défaillance.** La possibilité de retrouver le matériau des conduites recevant les pièces a pu être rapidement écartée car de nombreuses pièces sont compatibles avec plusieurs matériaux (comme les manchons, les joints, les colliers de prise en charge...), et car les tuyaux de remplacement ne garantissent pas que le tuyau remplacé soit du même matériau que le remplaçant (typiquement ce n'est pas le cas, puisque la fonte grise sera souvent remplacée par de la fonte ductile). Il apparaît ainsi que seuls les diamètres des conduites affectées par une défaillance peuvent être retrouvés à partir des pièces de réparation. La méthode utilisée pour extraire les informations de diamètre des conduites à partir des pièces de réparation est décrite en Annexe 3.

Grâce à cette méthode, il est désormais possible d'utiliser les diamètres extraits des pièces de réparation en supplément ou en complément des diamètres et matériaux dans les fichiers fuites pour aider l'identification de la conduite concernée par une intervention enregistrée via un OT dans SAP.

### 2.2.4 Traitement D : Préparation et complétion des conduites en service et abandonnées

Les conduites en service (ES) sont la principale source d'informations du SIG de la SWDE. Lors de l'inspection visuelle de la couche réseau du SIG, et lors de traitements effectués sur les conduites pour la tâche 01 du projet, quelques problèmes de topologie ont pu être constatés. Nous avons donc développé un script R permettant d'effectuer une validation topologique afin de localiser automatiquement toutes les conduites problématiques (en service ou abandonnées). Ces dernières étant en très petit nombre (moins de 10 conduites en service pour les données de Mons), des corrections ont été apportées à la main. Ces corrections ne concernent que les conduites en service car les conduites abandonnées du secteur de Mons ne contiennent aucune erreur topologique. L'Annexe 4 détaille la méthode et les corrections apportées.

D'autre part, les conduites abandonnées (ou conduites mises hors service, i.e., HS), au nombre de 975 pour le secteur de Mons, sont également une source de données indispensables pour l'analyse

de la gestion du patrimoine SWDE (Un exemple de répartition géographique des conduites abandonnées est représenté dans la Figure 4). Les conduites HS sont indispensables pour l'affectation des défaillances aux conduites, car elles peuvent être concernées par des fuites. **En pratique, les conduites abandonnées nécessitent des traitements pour pouvoir être considérées de façon équivalente aux conduites en service pour l'affectation des défaillances.**



Figure 4 : Exemples de tronçons abandonnés (lignes rouges) au centre de Mons.

Il s'agit essentiellement de :

- Corriger des problèmes de connexité (Figure 5). En effet, ces derniers ne peuvent pas être détectés automatiquement par le script de vérification topologique (voir Annexe 4).
- Compléter des données techniques manquantes, comme la Date De Pose (DDP) qui n'est pas renseignée pour 71 % des conduites HS sur Mons. Or, la DDP est essentielle pour calculer l'âge d'une conduite, qui est une variable principale pour le modèle de prédiction de défaillances (Le Gat, 2009).

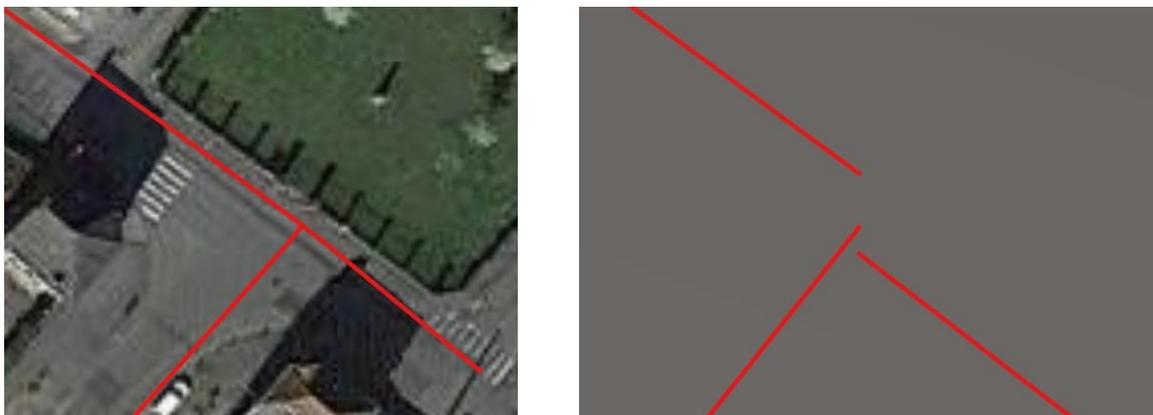


Figure 5 : Absence de connexité sur certaines conduites HS, aperçue en zoomant.

La correction de connexité des conduites HS est réalisée en accrochant les extrémités de conduites proches de moins de 50 cm.

La complétion des DHS est effectuée par la mise en relation des conduites HS et de la conduite ES qui leur est la plus proche. Ainsi, la DHS récupérée est simplement la DDP de la conduite ES la plus proche (sous conditions de compatibilité). Cette étape permet également d'attribuer un numéro de cellule à chaque conduite HS (cet attribut existe uniquement pour les conduites ES). La complétion de la DDP des conduites HS est effectuée par la mise en relation des conduites abandonnées et des conduites en service proches qui ne remplacent pas ces mêmes conduites abandonnées (contrairement à la complétion de DHS). Ainsi, la DDP récupérée pour ces conduites HS est la DDP des conduites ES proches qui ont été posées vraisemblablement au même moment et « autour » des conduites HS dont la DDP est manquante. Pour la complétion de DDP et de DHS des conduites HS, les informations récupérées depuis le réseau ES sont propagées de proche en proche au sein du réseau HS dans le but de réduire au minimum le nombre de conduites HS avec ces attributs manquants. De la même manière, une propagation des autres attributs (diamètre et matériau) de proche en proche est effectuée.

Le détail des traitements effectués est décrit en Annexe 5.

À l'issue de la complétion de données, il reste seulement :

- 4 % de conduites HS sans matériau (MAT)
- 2 % de conduites HS sans date d'abandon (DHS)
- 11 % de conduites HS sans date de pose (DDP)

### 2.2.5 Traitement E : Agrégation des conduites

Une proportion importante du réseau sur le secteur de Mons est constituée de conduites de faible longueur ; parmi les 31 822 conduites du réseau en service de ce secteur (extrait de février 2021), plus de 50 % ont une longueur inférieure à 6.5 m.

**Ces faibles longueurs risquent de poser problème lorsque l'on cherche à attribuer une caractéristique (comme un nom de rue) ou un évènement (fuite, casse, intervention) géo-localisé à une conduite**, puisque la présence d'un grand nombre de petits tronçons au même endroit peut rendre difficile la sélection d'un tronçon particulier (dans le cas d'un point géo-localisé comme une défaillance) ou d'une caractéristique (le nom de rue dans le cas d'un minuscule tronçon situé au milieu d'un carrefour) (Figure 6).



Figure 6 : Exemple de tronçons de petite taille localisés dans un carrefour.

Une inspection du réseau sous SIG révèle que cette taille relativement faible des tronçons ne provient pas d'un éventuel changement fréquent de caractéristiques le long des conduites, mais de la méthode de dessin du réseau. **Ainsi, agréger les tronçons de caractéristiques identiques semble naturel et devrait permettre de faciliter l'utilisation des conduites dans les traitements futurs.**

La méthode détaillée d'agrégation est décrite en Annexe 6. De façon synthétique, le principe de la méthode est le suivant :

- On considère chaque tronçon, en allant du plus petit au plus grand. Tout tronçon de longueur inférieure à un certain seuil (variable) est agrégé à l'un de ses voisins appartenant à la même cellule sous conditions (variables) de similarité en termes de caractéristiques et d'orientation dans le plan afin d'éviter d'agréger des conduites trop différentes (le tronçon voisin sélectionné est appelé tronçon « absorbant »).
- Les tronçons sans cellule peuvent être regroupés entre eux, comme s'ils étaient dans la même cellule.
- Un tronçon pour lequel il manque des caractéristiques (DDP, matériau, diamètre) ne pourra jamais être agrégé.
- Si plusieurs tronçons absorbants sont possibles sur un même nœud, le tronçon agrégable se raccrochera au tronçon voisin minimisant l'angle entre les deux canalisations.
- Un tronçon agrégable s'accrochera toujours de préférence du côté où il a un seul tronçon voisin, et au tronçon compatible le plus long dans le cas où ses deux côtés ont le même nombre de voisins.
- Quatre passes (étapes) sont effectuées séquentiellement afin d'optimiser la réduction du nombre de petits tronçons. Dans chacune des passes, les conditions d'agrégation sont adaptées pour débloquer localement l'agglomération et permettre la propagation du processus d'agrégation sur le réseau entier.

Le Tableau 6 décrit les résultats de l'agrégation pour l'extrait de conduites en service au 22/02/2021 (31 822 tronçons).

Tableau 6 : Résultats de l'agrégation pour les conduites en service.

	Nombre de tronçons	Part des tronçons L ≤ 6 m	Part des tronçons L ≤ 20 m
Fichier initial	31 822	49 %	63 %
Fichier agrégé	16 875	12 %	28 %

La méthode d'agrégation est également appliquée aux conduites HS qui ont été corrigées et complétées, encore une fois dans le but de pouvoir traiter les conduites en service et les conduites HS de la même façon. La seule différence avec l'agrégation des conduites en service vient du critère additionnel de même date d'abandon (attribut propre aux conduites HS uniquement) nécessaire pour autoriser l'agrégation. Dans ce processus, la DHS est considérée exactement de la même façon que la date de pose dans les 4 passes. Ainsi, le nombre de tronçons HS passe de 975 à l'origine à 661 après agrégation.

### 2.2.6 Traitement F : Attribution d'ICAR de rue aux conduites

Comme indiqué en section 2.1.2, il n'y a généralement pas d'identification directe de la conduite concernée par un OT donné dans l'extrait SAP (par exemple via l'ID de conduite). La source la plus directe d'informations reste donc l'adresse associée à l'OT, qui est constituée la plupart du temps d'une rue et d'un numéro d'immeuble. Il faut noter cependant que de nombreuses adresses renseignées n'ont pas de numéro d'immeuble (17 % des OT de type réparation de conduite par exemple), ce qui rend difficile la sélection de la conduite adéquate pour ces OT à partir seulement du nom de rue.

Il paraît a priori possible de retrouver quelle est la conduite concernée par l'OT, par simple relation de proximité avec le point d'adresse (point avec coordonnées X et Y), ou bien, dans certain cas, grâce à la rue lorsque seule cette dernière est renseignée. **Dans les deux cas, il est important de savoir comparer la rue renseignée pour l'intervention et la rue dans laquelle se situe la conduite, ce qui suppose que les conduites possèdent un nom (ou identifiant) de rue comme attribut** (Figure 7). Dans ce sens, le fichier de base des conduites du SIG contient les identifiants ICAR de rue des conduites, attribués par la SWDE grâce à une méthode reposant sur la correspondance entre les rues et les sommets des conduites (méthode « par sommets »). ICAR désigne l'Inventaire Centralisé des Adresses et des Rues en Wallonie. C'est une base de données maintenue par le Service Public de Wallonie (SPW) et qui permet de donner un identifiant unique à chaque rue en Wallonie et à chaque point d'adresse. C'est cette base de données qui a été utilisée pour les rues attribuées aux conduites, en particulier l'ICAR de rue qui est plus simple et plus fiable à manipuler que le nom de rue.

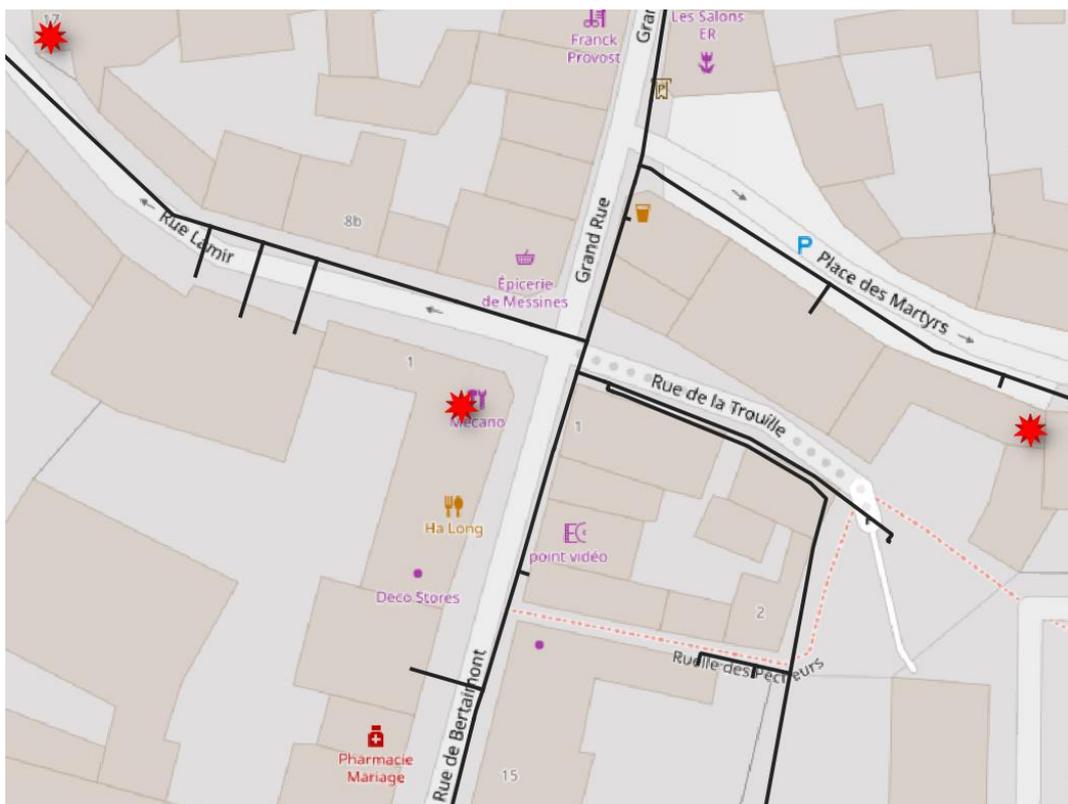


Figure 7 : Exemple d'OT (au centre) dont la conduite concernée peut être ambiguë.

Néanmoins, les ICAR de rue déjà attribués grâce à la méthode basée sur les sommets des conduites ne sont pas toujours suffisamment précis pour permettre la localisation fiable de la conduite grâce au point d'adresse, en particulier dans ou proche des carrefours, où la notion d'appartenance à une rue pour une conduite est parfois ambiguë (Figure 7).

Ainsi, nous avons décidé de réattribuer des ICAR de rue aux conduites en développant une méthode sur mesure dont l'objectif est de lever les ambiguïtés dans les carrefours. La méthode utilisée, dite « par tampons », car utilisant des tampons pour attribuer un ICAR de rue à chaque conduite, est détaillée en Annexe 7. Cette méthode a été validée conjointement par l'INRAE et la SWDE. La Figure 8 compare les ICAR de rue attribués aux conduites par deux méthodes, l'une à partir des sommets des conduites et l'autre (finalement sélectionnée) par tampons autour des axes de voirie.

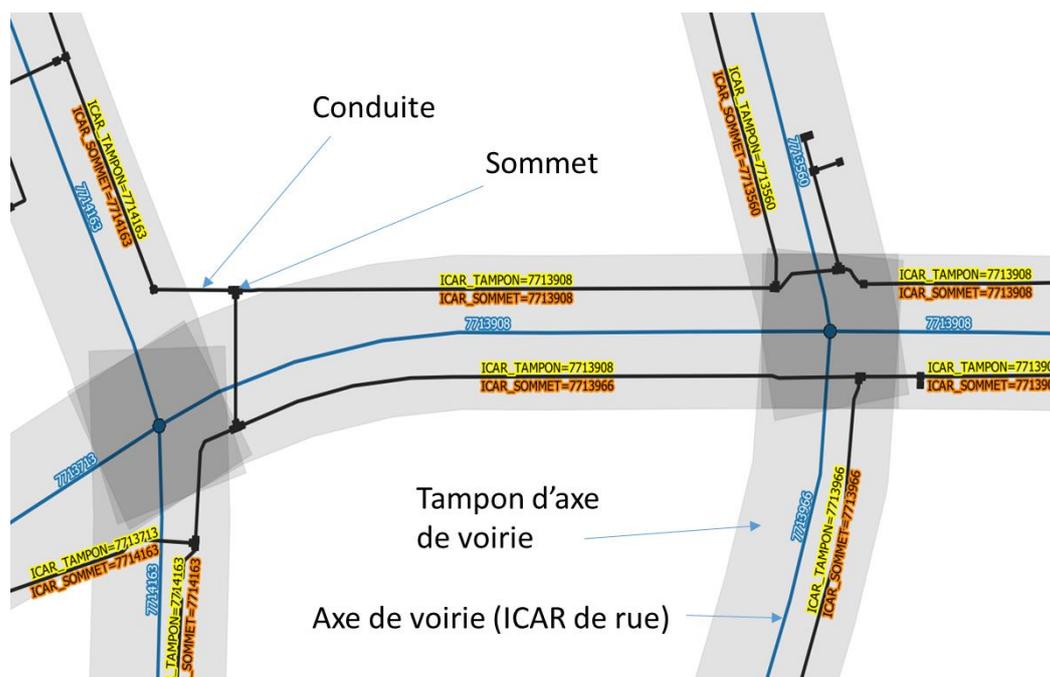


Figure 8 : Exemple d'affectation des ICAR de rue selon les méthodes par sommets et par tampons.

Les ICAR sont attribués avec la même méthode aux conduites agrégées en service, ou abandonnées. L'utilisation de conduites agrégées rend l'attribution des ICAR de rue plus facile en réduisant le nombre des conduites de faible longueur, pour lesquelles il est plus souvent difficile de décider à quelle rue elles appartiennent. En résumé, le principe est d'identifier quelle rue « ressemble » le plus à la conduite. Cette ressemblance est évaluée par la proximité de la conduite avec l'axe de rue/de voirie (par l'intermédiaire de tampons autour de l'axe de rue) et par l'orientation similaire entre la conduite et l'axe de rue (le calcul des angles conduites-voirie est identique au calcul des angles conduites-conduites pour l'agrégation). Ainsi, une conduite qui possède une grande partie de sa longueur dans le tampon d'un axe de voirie et qui forme un angle faible avec l'axe de voirie reçoit l'ICAR de rue de cet axe de voirie (i.e., de cette rue).

Les statistiques sur l'attribution des ICAR aux conduites sont détaillées dans le Tableau 7. Pour les conduites abandonnées, il y a moins de conduites par ICAR car les conduites HS ne sont pas très nombreuses contrairement aux conduites en service qui couvrent la quasi-totalité des rues du secteur de Mons.

Tableau 7 : Statistiques d'attribution d'ICAR de rue aux conduites en fin de traitement.

	Conduites en service (22/02/2021)	Conduites en service agrégées	Conduites abandonnées agrégées
<b>Nombre de conduites</b>	31 822	16 875	661
<b>Taux de conduites sans ICAR de rue</b>	3 %	3 %	2 %
<b>Taux d'ICAR de rue avec une seule conduite</b>	8 %	16 %	48 %
<b>Nombre moyen de conduites par ICAR de rue</b>	14	8	3

L'intérêt de l'agrégation des conduites avant l'attribution des ICAR de rue est évident (Tableau 7), puisqu'il y a presque deux fois moins de conduites par ICAR de rue, et deux fois plus d'ICAR de rue avec une seule conduite. Cela provient simplement du fait d'avoir divisé le nombre de conduites environ par deux, en particulier en éliminant les petits tronçons dans les carrefours qui peuvent potentiellement recevoir de nombreux ICAR différents.

La Figure 9 résume de manière synthétique l'ensemble des traitements intermédiaires D, E, et F appliqués aux conduites en service et aux conduites hors service. Les flèches bleues indiquent l'export et/ou l'import de données tandis que les flèches en pointillés indiquent quels attributs d'un fichier sont associés à quels attributs d'un autre fichier. La signification des attributs est détaillée dans l'Annexe 10.

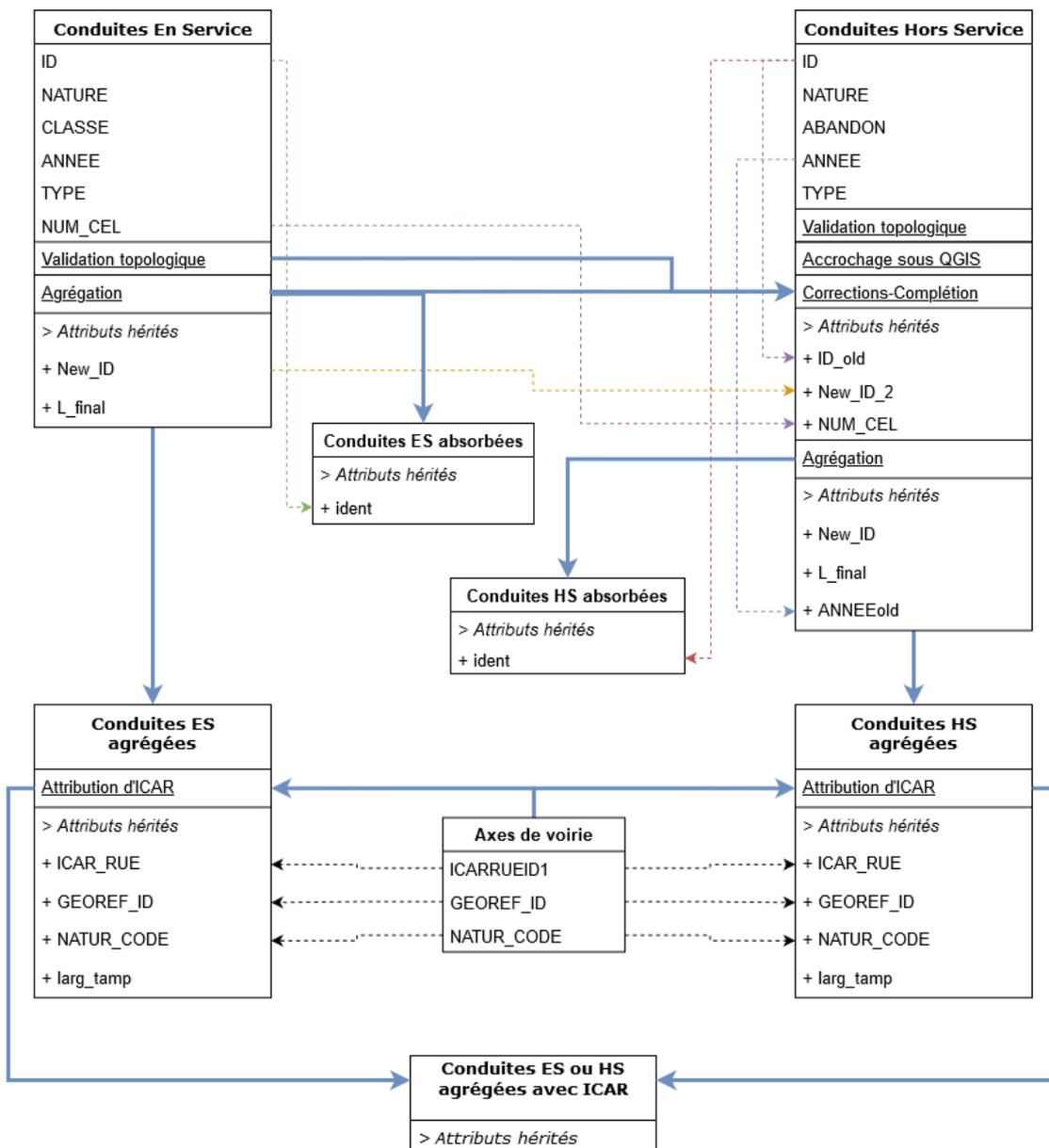


Figure 9 : Synoptique des traitements appliqués sur les conduites du SIG.

### 2.2.7 Traitement G : Affectation des défaillances aux conduites

Tous les traitements précédents ont permis de préparer les données pour faciliter l'affectation d'une fuite ou d'une casse, enregistrée comme un OT dans SAP de type 105, à une conduite du SIG :

- Les liens entre les OT de type 105 (réparation de conduite) dans SAP, les enregistrements des fichiers « fuites » et les « anomalies » du SIG ont été établis.
- On peut identifier dans le SIG l'ensemble des conduites agrégées qui sont dans la même rue que la défaillance, et proches de l'adresse renseignée si le numéro d'immeuble est connu, les conduites HS peuvent être considérées au même titre que les conduites en service.
- Pour une bonne partie de ces OT, on peut directement connaître les matériau et/ou diamètre des conduites via les fichiers fuites et/ou les pièces de réparation pour aider à distinguer la conduite concernée du SIG parmi celles dans cette rue (et proches du point d'adresse si le numéro d'immeuble est connu).

Ainsi, la dernière étape du traitement en vue de la préparation du fichier des défaillances consiste à attribuer un identifiant de conduite (ID) à chaque OT. L'Annexe 8 décrit en détail la procédure utilisée. **La méthode est complexe, car il y a de nombreuses façons de trouver la conduite concernée en croisant les informations de localisation et les caractéristiques techniques.** Ce croisement est notamment nécessaire pour pallier les manques de certaines données pour certains OT (comme des adresses sans numéro d'immeuble). Pour les réparations de conduites, il y a au final plus de 50 manières distinctes d'identifier la conduite réparée. Environ 7 % des conduites identifiées ainsi l'ont été via 39 de ces manières (Tableau 18), qui totalisent chacune moins de 10 interventions et qui représentent donc presque du traitement au cas par cas.

Au final, sur le secteur de Mons, près de 99 % des OT de type réparation de conduite (type 105) ont été liés à une conduite bien identifiée du SIG. Cela représente un peu plus de 900 conduites, soit environ 11 % du réseau de Mons (ou encore 21 % de sa longueur), qui est affecté par une défaillance sur la plage 2007-2020. Sur la base de cette mise en relation, les fichiers Casses ont pu être préparés.

La Figure 10 résume les traitements utilisés (traitements intermédiaires A, B, C et G) pour affecter une conduite à chaque défaillance. Les flèches bleues en trait plein indiquent l'import et/ou l'export de données, tandis que les flèches en pointillés indiquent les attributs qui permettent la jointure entre plusieurs sources de données. L'Annexe 10 explicite les différents attributs qui apparaissent dans la Figure 10.

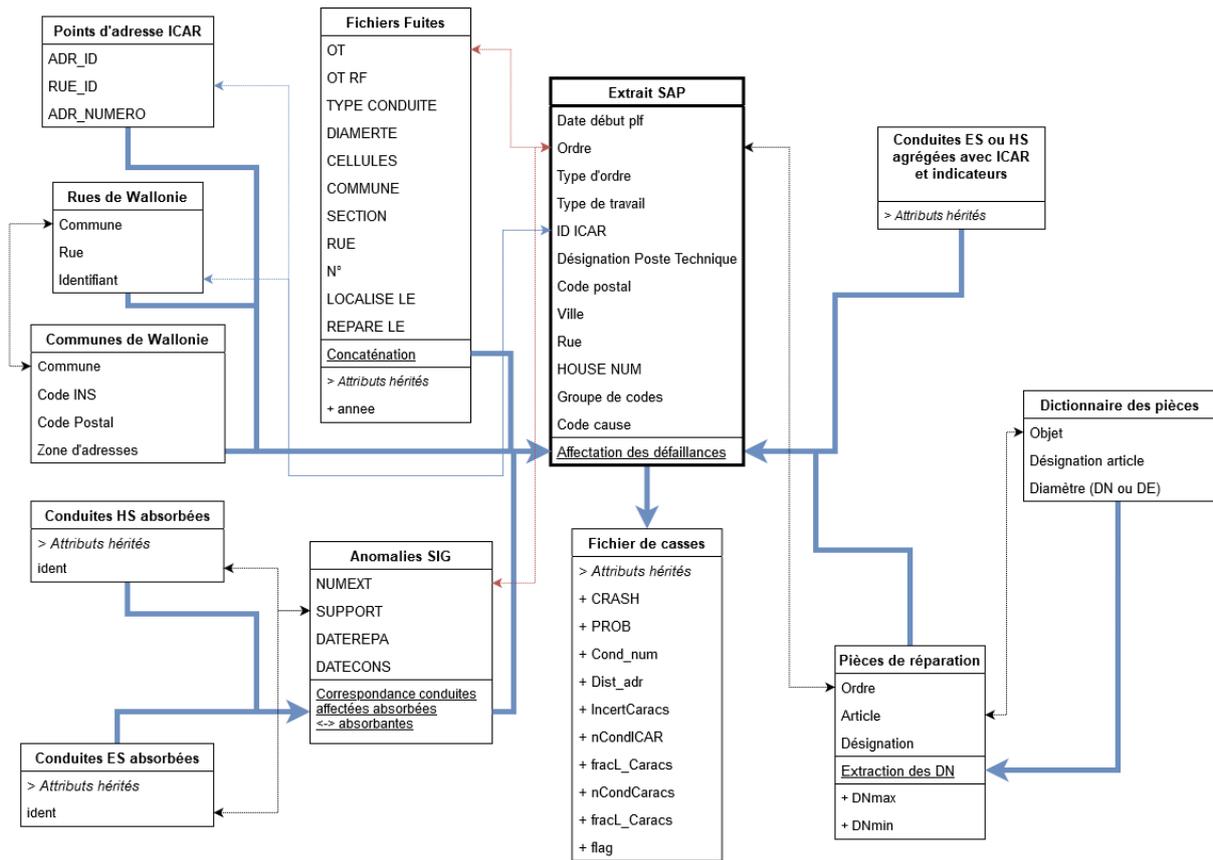


Figure 10 : Synoptique des traitements utilisés pour lier chaque défaillance à une conduite du SIG.

### 2.2.8 Traitement H : Caractéristiques additionnelles des conduites

Les fichiers de conduites en service ou hors service agrégées ont été complétés par des caractéristiques additionnelles (d’environnement, de topographie, de sol, etc.) calculées pour chaque tronçon. Ces informations complémentaires permettent d’un part d’obtenir de nouvelles covariables qui pourront constituer des facteurs de risque dans les modèles de prévision des casses (tâche 03) et d’autre part, en regroupant les conduites par cellule, de produire des indicateurs qui contribueront à l’établissement d’une typologie des cellules (tâche 05). La Figure 11 résume les traitements effectués qui sont présentés au chapitre 3. La signification des attributs est détaillée dans l’Annexe 10.

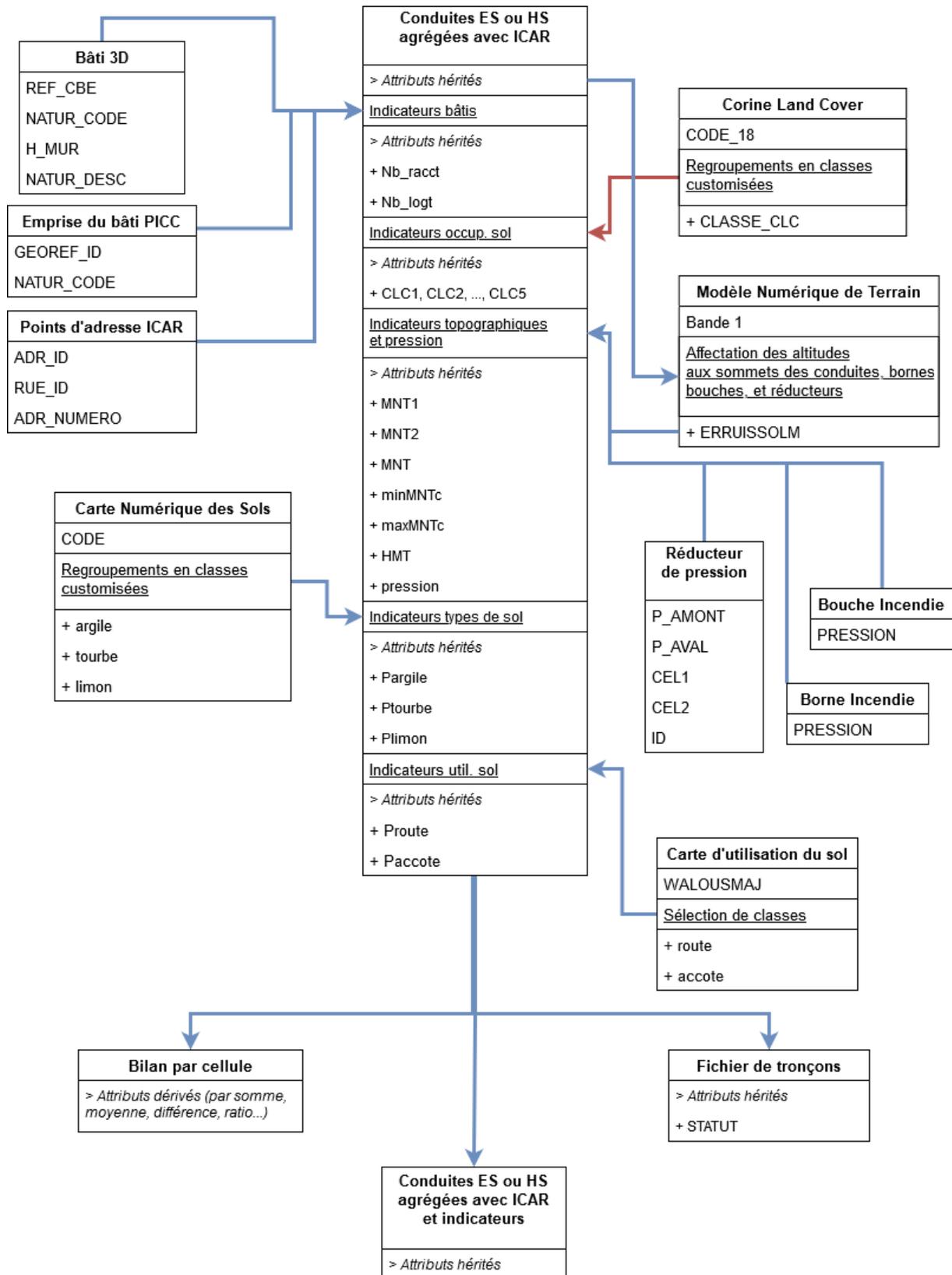


Figure 11 : Synoptique des calculs d'indicateurs pour les conduites et les cellules.

## 3 Tâche 05 - Exploitation des données de la sectorisation

### 3.1 Objectifs de la tâche 05

La tâche 05 contribue essentiellement au volet 1 du projet « compréhension du processus de fuite et efficacité de la recherche de fuite » (Figure 1). Elle se propose de dresser une typologie des cellules pour permettre de moduler les objectifs de performance de chaque cellule. Cette typologie sera construite à partir de données de contexte, de patrimoine et d'exploitation du réseau agrégées à l'échelle de la cellule.

La performance d'une cellule sera appréciée à partir de valeurs de référence prise par des indicateurs à définir et à calculer à partir des données de débit et de contexte associées aux cellules.

### 3.2 Données caractéristiques des cellules

Une première caractérisation des cellules est en cours. Elle s'appuie sur des données géographiques (bâtis, CLC, MNT, nature des sols, utilisation des sols), de patrimoine (données SIG du réseau) et d'exploitation (uniquement les débits pour l'instant, puis les interventions par la suite).

Il est à noter que l'identifiant de la cellule est un attribut du tronçon et que cette information est à la base du calcul des caractéristiques des cellules.

#### 3.2.1 Caractéristiques patrimoniales

Les caractéristiques patrimoniales des cellules seront obtenues à partir des données disponibles dans les couches SIG du réseau AEP. Certaines variables ne posent pas de question sur leur construction :

- Diamètre moyen : moyenne pondérée par la longueur des tronçons ;
- Age moyen : moyenne pondérée par la longueur des tronçons ;

En complément, il reste à définir le détail de chacune des variables suivantes :

- Part de chaque matériau : Définition des regroupements de matériau à réaliser, puis cumul des longueurs de canalisation par type de matériau ;
- Part de chaque classe de diamètre : Définition des différentes classes de diamètre puis cumul des longueurs par classes de diamètre ;
- Part de chaque classe d'âge des canalisations : Définition des différentes classes d'âge puis cumul des longueurs par classes d'âge.

Il est à noter que la définition des classes nécessite une réflexion pour prendre en compte les particularités des données, comme par exemple la surreprésentation de certaines années de pose (Figure 12).

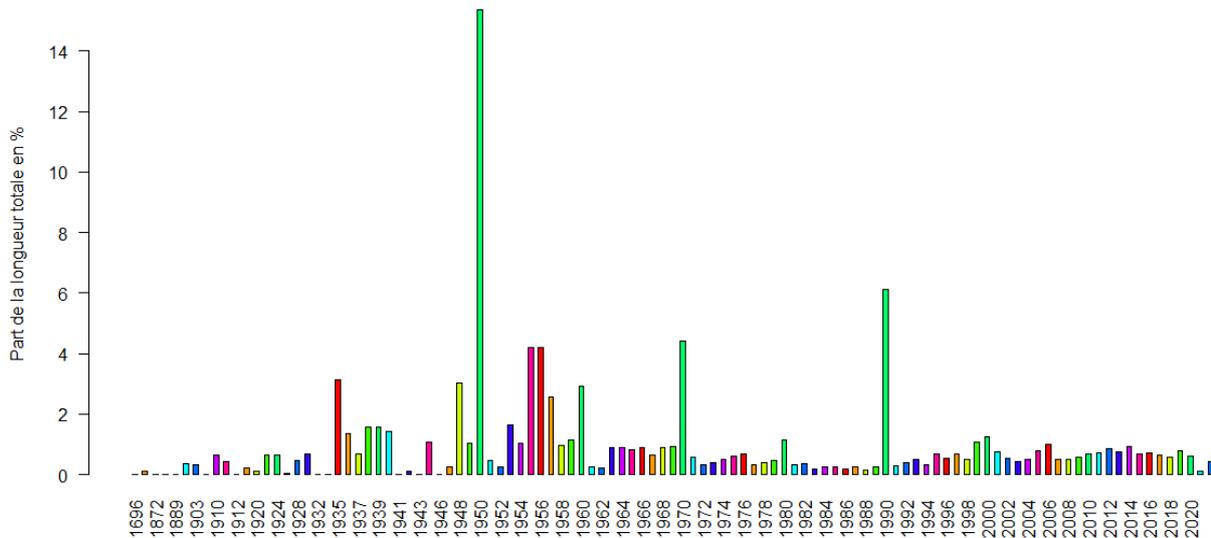


Figure 12 : Proportions de la longueur de réseau par année de pose.

### 3.2.2 Caractéristiques géographiques

#### 3.2.2.1 Bâti

Le nombre, la surface et la hauteur des constructions d'une zone sont des marqueurs de la densité urbaine. Trois couches SIG disponibles auprès du SPW sont utilisées pour extraire ces informations :

- Emprise des bâtiments 2021 (PICC) : cette couche maintenue à jour permet de recueillir la surface au sol du bâti. Il apparaît que la destination des bâtiments mentionnée dans cette couche n'est pas toujours juste ;
- Bâtiments 3D (SPW) créées entre 2013 et 2014 : en complément de la couche précédente, elle permet d'obtenir la hauteur d'une partie du bâti (en l'occurrence la hauteur des murs).
- Les points d'adresse (ICAR) permettent de localiser les locaux à destination d'habitation de la couche « Emprise des bâtiments (PICC) » et d'écartier ainsi les locaux annexes. S'il y a plusieurs logements dans le bâtiment, plusieurs points d'adresse sont superposés (1 pour l'adresse et 1 par logement) ouvrant la possibilité d'estimer le nombre d'abonnés à un point de raccordement.

A l'issue du croisement de ces trois sources d'information, un point d'adresse (qui peut être assimilé à un raccordement au réseau) est lié à une surface de bâti, une destination (habitation, hôpital, industriel, commerce, ...), une hauteur, et un nombre de logements. Ces points d'adresse sont ensuite liés à la conduite la plus proche qui porte la référence de la cellule (voir détail en Annexe 9).

*In fine*, chaque tronçon de conduite porte le nombre de point d'adresse (branchements) et le nombre de logements desservis. Ces informations sont agrégées à l'échelle de la cellule, avec en complément la surface totale des bâtis (avec un point d'adresse ICAR) et la hauteur moyenne pondérée par la surface des bâtis (voir détails en Annexe 9).

#### 3.2.2.2 Occupation des sols via Corine Land Cover

La couche Corine Land Cover (CLC) utilisée pour caractériser l'occupation des sols dans chaque cellule a été importée dans sa version mise à jour en 2018. Les 32 classes initiales d'occupation du sol rencontrées en Wallonie sont regroupées en 5 classes thématiques (Tableau 8).

Tableau 8 : Construction des classes CLC.

Classes		Code Corine Land Cover concernés
1	Zones urbanisées continues	111
2	Zones urbanisées discontinues	112
3	Zones artificialisées	121 à 124, 131 à 133, 141, 142
4	Zones agricoles	211 à 213, 221 à 223, 231, 241 à 244
5	Zones naturelles (forêts, zones humides et surface en eau)	311 à 313, 321 à 328, 331 à 335, 411, 412, 421 à 423, 511, 512, 521 à 523

Pour chaque tronçon de canalisation et par conséquent pour chaque cellule, il est alors possible de déterminer la part de longueur dans chacune des cinq classes tirées de CLC. Cela est réalisé par l'intersection des tracés des conduites avec les polygones délimitant les 5 classes tirées de CLC.

On note que plus on s'éloigne du centre urbain, plus la part de conduite en zone agricole s'accroît. Logiquement, il y a globalement très peu de conduites en zone naturelle (cf Figure 13).

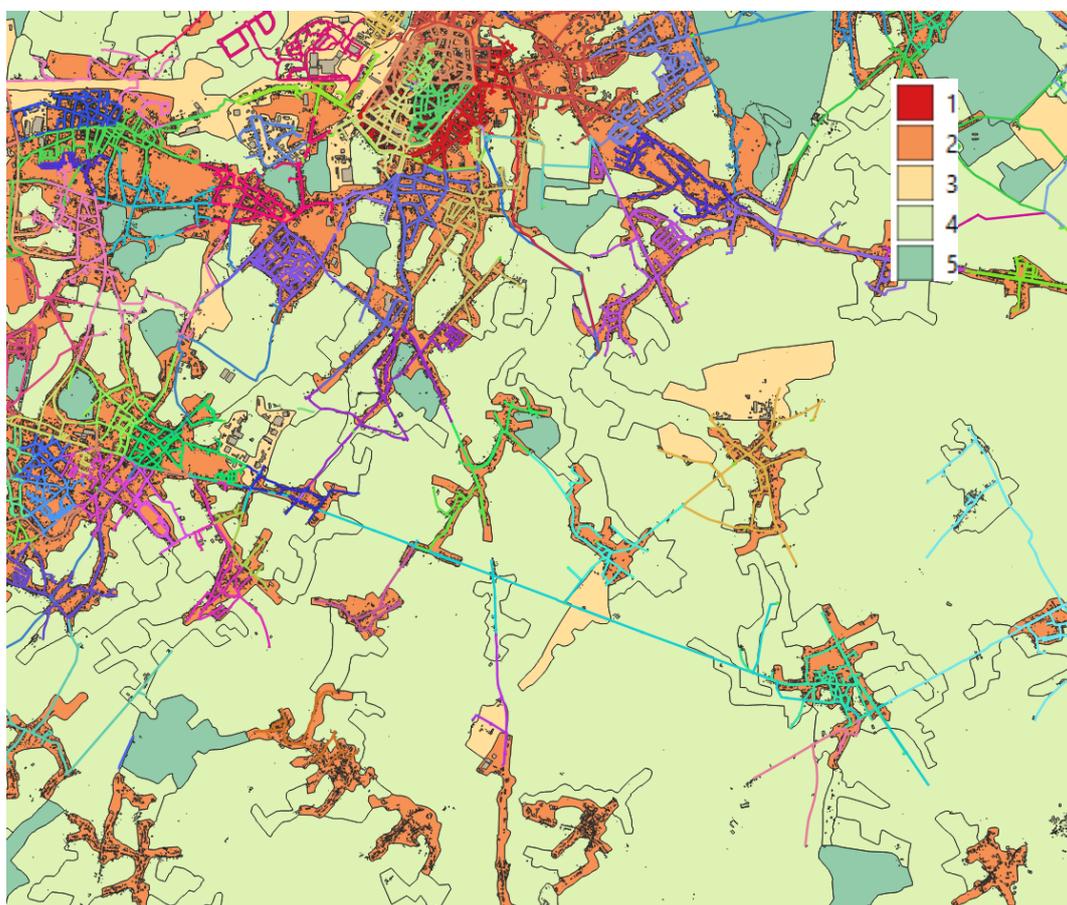


Figure 13 : Répartition des 5 classes CLC et du bâti autour de Mons.

### 3.2.2.3 Dénivelée et pression

Une pression de service élevée constitue un facteur pouvant altérer la durée de vie des canalisations. En première approche, les côtes MNT de chaque conduite de la cellule sont calculées à partir du Modèle Numérique de Terrain ERRUISSOL (MNT à la maille de 10 mètres du Service public de Wallonie) pour évaluer la dénivelée entre la conduite et le point haut de chaque cellule.

Ces données sont ensuite combinées avec les mesures *in situ* sur les hydrants et les vannes de régulation pour calculer des indicateurs de pression lorsque c'est possible.

La couche de référence est fournie par le Service Public de Wallonie à la maille 10 mètres (ERRUISSOL). Les altitudes sont affectées à chaque extrémité des tronçons en croisant le MNT et la couche SIG des tronçons de canalisations agrégés. Pour approcher la pression exercée sur chaque tronçon, les cotes des nœuds d'extrémité, leur moyenne ainsi que les cotes minimum et maximum de la cellule d'appartenance du tronçon sont ajoutées aux attributs des tronçons.

$$MNT_T = (MNT_1 + MNT_2)/2$$

Avec :

- $MNT_T$  cote au sol attribuée au tronçon
- $MNT_1$  cote au sol du premier nœud d'extrémité du tronçon déduite du MNTSPW
- $MNT_2$  cote au sol du second nœud d'extrémité du tronçon déduite du MNTSPW

Pour le calcul de la pression, en l'absence de vannes de régulation, la moyenne des mesures disponibles réalisées sur les hydrants d'une cellule permet d'estimer la HMT (Hauteur Manométrique Totale) de celle-ci puis d'en déduire un indicateur de la pression exercée sur chaque tronçon par différence avec la cote au sol du tronçon telle que calculée précédemment.

$$HMT_C = \frac{\sum_{i=1}^n MNT_{HYi} + P_{HYi}}{n}$$

$$P_T = HMT_C - MNT_T$$

Avec :

- $HMT_C$  hauteur manométrique totale moyenne de la cellule
- $MNT_{HYi}$  cote au sol de l'hydrant i de la cellule
- $P_{HYi}$  pression mesurée pour l'hydrant i de la cellule
- $n$  nombre d'hydrants de la cellule
- $P_T$  pression estimée du tronçon T

Lorsqu'une vanne de régulation de cote au sol  $MNT_{VR}$  se trouve dans la cellule, toutes les mesures de pression des hydrants dont la cote moyenne est supérieure à celle de la vanne ( $MNT_{HY} > MNT_{VR}$ ) ainsi que la pression amont de la vanne sont utilisés pour calculer la valeur moyenne de la HMT à l'amont de la vanne ( $HMT_{Cam}$ ). Alors, pour tous les tronçons de canalisation dont la cote moyenne est supérieure à celle de la vanne de régulation ( $MNT_T > MNT_{VR}$ ) la pression estimée est :

$$P_T = HMT_{Cam} - MNT_T$$

Selon le même principe la HMT à l'aval de la vanne de régulation ( $HMT_{Cav}$ ) est calculée avec les hydrants dont la cote moyenne est inférieure ou égale à celle de la vanne ( $MNT_{HY} \leq MNT_{VR}$ ). Pour tous les tronçons de canalisation dont la cote moyenne est inférieure ou égale à celle de la vanne de régulation ( $MNT_T \leq MNT_{VR}$ ) la pression estimée est :

$$P_T = HMT_{Cav} - MNT_T$$

Lorsqu'il y a plusieurs vannes de régulation en cascade, les étages de HMT sont déterminés par les côtes de chaque vanne de régulation, sa pression amont et sa pression aval. Pour chaque étage, toutes les mesures de pression des hydrants dont la côte moyenne est comprise entre la côte de deux vannes de régulation ainsi que la pression amont de la vanne la plus basse sont utilisées pour calculer la valeur moyenne de la HMT de l'étage. La pression peut alors être calculée pour tous les tronçons de canalisation dont la côte moyenne se situe entre les deux vannes de régulation.

A l'échelle de la cellule, une pression moyenne est calculée à partir de la moyenne des pressions des tronçons pondérée par la longueur des tronçons.

$$P_{moyC} = \frac{\sum_{i=1}^n P_{Ti} \times L_{Ti}}{\sum_{i=1}^n L_{Ti}}$$

Avec :

- $P_{moyC}$  pression moyenne de la cellule
- $P_{Ti}$  pression estimée du tronçon  $T_i$
- $L_{Ti}$  longueur du tronçon  $T_i$
- $n$  nombre de tronçons de la cellule

Si aucune mesure de pression n'est disponible pour une cellule, sa HMT est calculée en ajoutant 20 mètres colonne d'eau (mce) (valeur forfaitaire usuelle de la pression minimum admise qui peut être adaptée) à la côte au sol du nœud le plus élevé de la cellule.

$$HMT_C = MNT_{Cmax} + P_{Cminref}$$

- $HMT_C$  hauteur manométrique totale de la cellule
- $MNT_{Cmax}$  côte au sol du nœud le plus élevé de la cellule
- $P_{Cminref}$  pression minimum de référence de la cellule (par défaut 20 mce)

Les pressions des tronçons et la pression moyenne de la cellule sont alors calculées comme précédemment.

$$P_T = HMT_C - MNT_T$$

### 3.2.2.4 Type de sol via Sols de la Wallonie

La carte numérique des sols disponible sur le Géoportail de la Wallonie (maintenu par le SPW) permet de distinguer 23 types de sol. Parmi ceux-ci, les sols considérés comme potentiellement à risque pour les canalisations (retrait-gonflement, drainage naturel faible) sont répartis en 3 classes (Tableau 9).

Tableau 9 : Classes de répartition des sols à risque pour les canalisations.

Classes	Code et libellé du sol	
Tourbe	1 000	Sols tourbeux ou tourbières
Limon peu drainant	4 020	Sols limoneux à drainage naturel modéré ou imparfait
	4 030	Sols limoneux à drainage naturel assez pauvre à très pauvre
Argile	5 010	Sols argileux à drainage naturel favorable ou imparfait
	5 020	Sols argileux à drainage naturel assez pauvre à très pauvre

La longueur de canalisations dans chacune de ces trois classes est évaluée à l'échelle de la cellule et à l'échelle du tronçon.

### 3.2.2.5 Couverture des conduites via l'utilisation des sols (WALOUS)

A partir de la cinquantaine d'usage des sols cartographiés dans le référentiel WALOUS\_2018, l'objectif est d'évaluer la part des canalisations sous chaussée et sous accotement ou trottoir. Les deux classes retenues pour cette sélection sont présentées dans le Tableau 10.

Tableau 10 : Classes d'usage des sols liées à la couverture des conduites.

Classes	Code et libellé de l'utilisation du sol	
Route	4_1_1	Transport routier
Accotement ou trottoir	6_6_b	usage inconnu non cadastré (trottoir, stationnement ou sol naturel, ...)

Ces 2 zones recourent à elles seules plus de 80 % des conduites (cf Figure 14). Cette donnée constitue une indication qui bien sûr est conditionnée par la qualité de la localisation des canalisations, mais également du découpage des zones d'utilisation des sols.



Figure 14 : Exemple d'intersection des routes et trottoirs avec les canalisations.

### 3.2.3 Caractéristiques d'exploitation

#### 3.2.3.1 Qualité de l'eau

Les tronçons de conduites sont rattachés par un attribut de la couche SIG à la Zone de Distribution d'Eau (ZDE). Pour chaque tronçon de canalisation, il sera possible d'avoir des informations sur la qualité chimique de l'eau (TAC, TH, ...) et ainsi de son caractère corrosif ou non. *A priori*, une cellule est rattachée à une et une seule ZDE, ces informations seront donc également portées par la cellule.

### 3.3 Débits des cellules

L'objectif est d'établir une chronique des volumes d'eau distribués dans chaque cellule. Analysée conjointement avec la chronique des interventions et la typologie de chaque cellule, elle permettra de construire des indicateurs d'aide à la décision pour définir des objectifs et prioriser les actions de maîtrise des pertes en eau.

Les débits des cellules sont reconstitués à partir des compteurs d'entrée et sortie de la cellule. Les données disponibles issues des comptages se répartissent en deux groupes :

1. Les données sous forme d'index : Ce sont des données brutes au pas de temps du quart d'heure (synchronisées au quart d'heure pile). A chaque compteur sont associées deux séries de données (une pour chaque sens d'écoulement) fournies dans deux fichiers séparés.
2. Les données sous forme de débits : Ce sont des données prétraitées pour être recalées au pas de temps du quart d'heure. Une seule série de données (donc un seul fichier) est fourni pour chaque compteur. Il s'agit de données signées, c'est-à-dire positives pour un sens d'écoulement et négatives pour le sens inverse.

Les débits reconstitués de la cellule le sont au quart d'heure pour calculer dans un premier temps les indicateurs de la SWDE (*Q<sub>nuit</sub>* et Volume jour).

La Figure 15 présente un extrait des débits calculés pour la cellule 52300016 qui est concernée par un compteur « débit » et deux compteurs « index ».

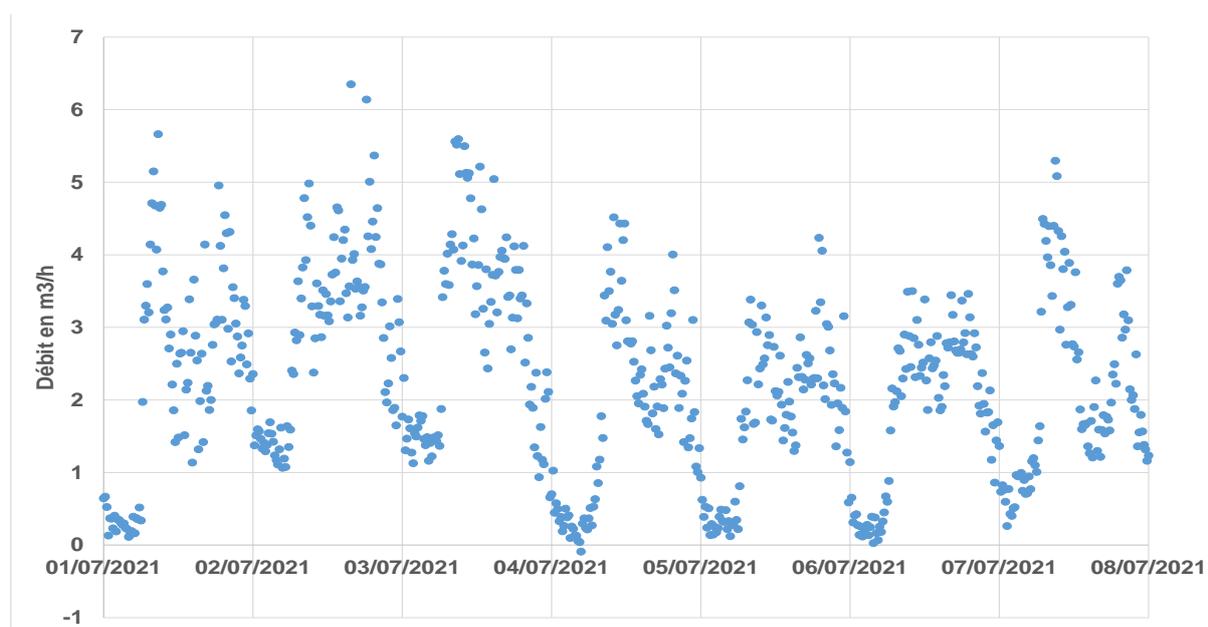


Figure 15 : Reconstitution des débits de la cellule 52300016.

Le volume jour est alors calculé comme la somme des volumes transités par quart d'heure de 05h15 la veille à 05h00.

$$V_j = \sum_{t_i \in PT_j} Q(t_i) \Delta t$$

- $V_j$  Volume du jour j à 5h00 ( $m^3$ )
- $Q(t_i)$  débit au  $\frac{1}{4}$  d'heure  $t_i$  en  $m^3/h$
- $\Delta t = 0.25$  h
- $PT_j$  ensemble des 96 pas de temps au  $\frac{1}{4}$  d'heure de 5h15 le jour j-1 à 5h00 le jour j

Deux calculs sont effectués pour  $Q_{nuit}$ , indicateur caractérisant le débit de nuit de la cellule.

Le premier,  $Q_{nuit1}$ , est calculé selon la formule fournie par SWDE. Il s'agit du minimum de la moyenne glissante des débits horaires mesurés par pas de temps de 15 minutes de 01h00 à 05h00

La moyenne glissante  $Qmg(t_i)$  ( $m^3/h$ ) au pas de temps  $t_i$  est ainsi définie :

$$Qmg(t_i) = \frac{1}{4} (Q(t_{i-3}) + Q(t_{i-2}) + Q(t_{i-1}) + Q(t_i))$$

$$Q_{nuit1_j} = \min_{t_i \in PTN1_j} Qmg(t_i)$$

Avec  $PTN1_j$  ensemble des 14 pas de temps au  $\frac{1}{4}$  d'heure de 1h45 à 5h00 le jour j.

Le second,  $Q_{nuit2}$  est calculé pour être compatible avec les valeurs réellement rencontrées, il s'agit du minimum des débits horaires mesurés chaque quart d'heure entre 01h00 et 05h00

$$Q_{nuit2_j} = \min_{t_i \in PTN2_j} Q(t_i)$$

Avec  $PTN2_j$  ensemble des 17 pas de temps au  $\frac{1}{4}$  d'heure de 1h00 à 5h00 le jour j.

## 4 Avancées méthodologiques

### 4.1 Tâche 02 - Analyse de la gouvernance de la SWDE en matière de GPI

La tâche relative à l'analyse de la gouvernance de la SWDE en matière de Gestion Patrimoniale des Infrastructures (GPI) a démarré le 1<sup>er</sup> avril 2021. Son objectif est d'analyser les processus de décision et de mise en œuvre de la GPI (court et long termes) de la SWDE. Elle vise plus précisément à identifier et à expliciter les déterminants de la gestion patrimoniale de la Société, et leurs logiques (socio-économiques, techniques, professionnelles...). Cela soulève quatre questions centrales :

- Quels sont les processus et les logiques d'élaboration de la GPI à l'échelle de la SWDE ?
- Sur quels dispositifs, outils et critères s'appuient ces processus ?
- Comment ces politiques sont-elles mises en œuvre au cours du temps ?
- Quels effets produisent-elles aux échelles de la SWDE et de la région Wallonne ?

Concrètement, l'engagement de la tâche a donné lieu à un premier travail bibliographique, permettant d'affiner la problématique de la recherche et de préparer le protocole d'enquête. La littérature grise fournie par les partenaires de la SWDE (contrats de gestion, charte de gouvernance, statistiques Aquawal, etc.) et celle disponible sur le site internet de la SWDE (rapports annuels notamment) nous ont permis d'améliorer notre connaissance de l'organisation institutionnelle, professionnelle et territoriale de la Société.

Partant, notre problématique consiste à se demander à quelles échelles et dans quels cadres la GPI s'inscrit-elle, et quelle dialectique (complémentarité, concurrence/opposition, effets rebonds, etc.) s'instaure entre eux. Pour y répondre, nous formulons deux hypothèses :

- La GPI déborde l'unique cadre de la SWDE. Elle s'inscrit dans un système de gouvernance élargi à la région Wallonne, eu égard à la participation d'acteurs politico-administratifs régionaux dans la gestion du service, ainsi qu'aux relations de coopération et d'interdépendance entre acteurs de l'eau. Ce cadrage institutionnel et territorial constitue ainsi un système de contraintes et d'opportunités qui influe sur les stratégies et les orientations de la SWDE.
- La GPI répond également à une organisation (culture, organisation des métiers, règles et procédures...) et à une instrumentation du travail (objectifs, outils...) spécifiques, qui orientent la compréhension des acteurs / agents, et structurent la prise en charge des enjeux infrastructurels. À cela, il convient d'ajouter les effets d'une potentielle « gestionnarisation » du travail, qui se traduit par l'adoption de nouveaux objectifs et une réorganisation de la structure et des conditions de travail, dans le but de maîtriser et d'optimiser les résultats de l'activité des agents. Dès lors, les nouvelles conditions d'exercice nécessitent d'être appropriées par le personnel de la SWDE, ce qui peut se traduire par des difficultés et des résistances, mais également par une nouvelle hiérarchisation des priorités et l'émergence de nouveaux enjeux tant individuels qu'à l'échelle de l'organisme.

Cette première phase préparatoire a donné lieu à une note méthodologique (Annexe 11). Celle-ci sera approfondie durant le deuxième semestre 2021, et enrichie d'une enquête exploratoire. L'année 2022 sera consacrée au principal travail d'enquête et à l'exploitation des résultats.

## 4.2 Tâche 07 - Modélisation de la probabilité d'un tronçon d'être fuyard

L'Annexe 12 détaille l'approche méthodologique proposée pour modéliser la probabilité qu'un tronçon de caractéristiques données soit dans un état fuyard détectable. Cette approche est inscrite dans le cadre conceptuel probabiliste de la régression logistique. Une attention est portée au caractère orienté des campagnes de recherche de fuites, qui ne ciblent pas aléatoirement les cellules, ni les canalisations au sein des cellules, mais plutôt celles que la surveillance des débits de nuit et la pré-localisation désignent comme suspectées a priori de receler des fuites.

## 5 Conclusion

Au cours de la première année du projet GePaME, une part importante des efforts a été consacrée à la prise en main et au traitement des données nécessaires à la conduite du programme de recherche. A l'issue de ce travail, les sources pertinentes de données ont été identifiées et expertisées (SIG, SAP, Fichiers « Fuite », Données de télégestion « Perf'O », ressources géographiques publiques de Wallonie). Pour combler l'écart entre la forme sous laquelle l'information est disponible et les besoins de l'étude, en particulier la nécessité de relier des événements et des caractéristiques aux tronçons et aux cellules, des méthodes et des codes de calcul ont été développés à partir des données du secteur de MONS. **A ce stade, il est considéré que l'on dispose de méthodes opérationnelles et fiables pour produire les fichiers nécessaires pour alimenter le logiciel Casses ainsi que pour calculer les caractéristiques nécessaires à l'établissement d'une typologie des cellules.** Le développement des méthodes de traitement des données de débit de la sectorisation est en cours.

Parallèlement, des approches méthodologiques ont été construites pour servir de fil conducteur à la réalisation des tâches 02 « Analyse de la gouvernance de la SWDE en matière de GPI » et 07 « Modélisation de la probabilité d'un tronçon d'être fuyard ».

Ainsi, à partir des résultats obtenus au terme de cette première année du projet, il est envisagé de poursuivre le programme de recherche dans les directions suivantes :

- appliquer les méthodes de traitement de données aux autres secteurs de la SWDE (ATH et GAURAIN puis l'ensemble du territoire) ;
- analyser et traiter les données de la RF disponibles dans SAP (Tâche 01) ;
- poursuivre le développement des méthodes d'analyse de la sectorisation (tâche 05) ;
- initier la réalisation de la tâche 03 « implémentation de Casses » ;
- poursuivre la conduite de la tâche 02 ;
- mettre en œuvre la tâche 06 « Analyse des pratiques de recherche de fuites » (post-doctorat de Selma Baati) ;
- concevoir le modèle de fuite (tâche 07, post-doctorat de Marie Collet) ;
- lancer l'enquête pour l'analyse de la vulnérabilité des usagers (tâche 08).

Il est en outre prévu de soumettre une communication orale au congrès annuel de l'ASTEE 2022 à Dunkerque pour présenter les méthodes de traitement des données développées au cours de cette première année du projet. Un article dans la revue TSM est également envisagé.

## 6 Annexes

### Annexe 1 : Articulation des tâches – Détail par volet de recherche

L’articulation des tâches par volet est représentée par les figures suivantes :

- Figure 16 : volet 1, compréhension du processus de fuite et efficacité de la recherche de fuite (RF) ;
- Figure 17 : volet 2, aide à la sélection multi-objectifs et contrainte des chantiers de renouvellements ;
- Figure 18 : volet 3, simulation long terme de stratégies de gestion patrimoniale.

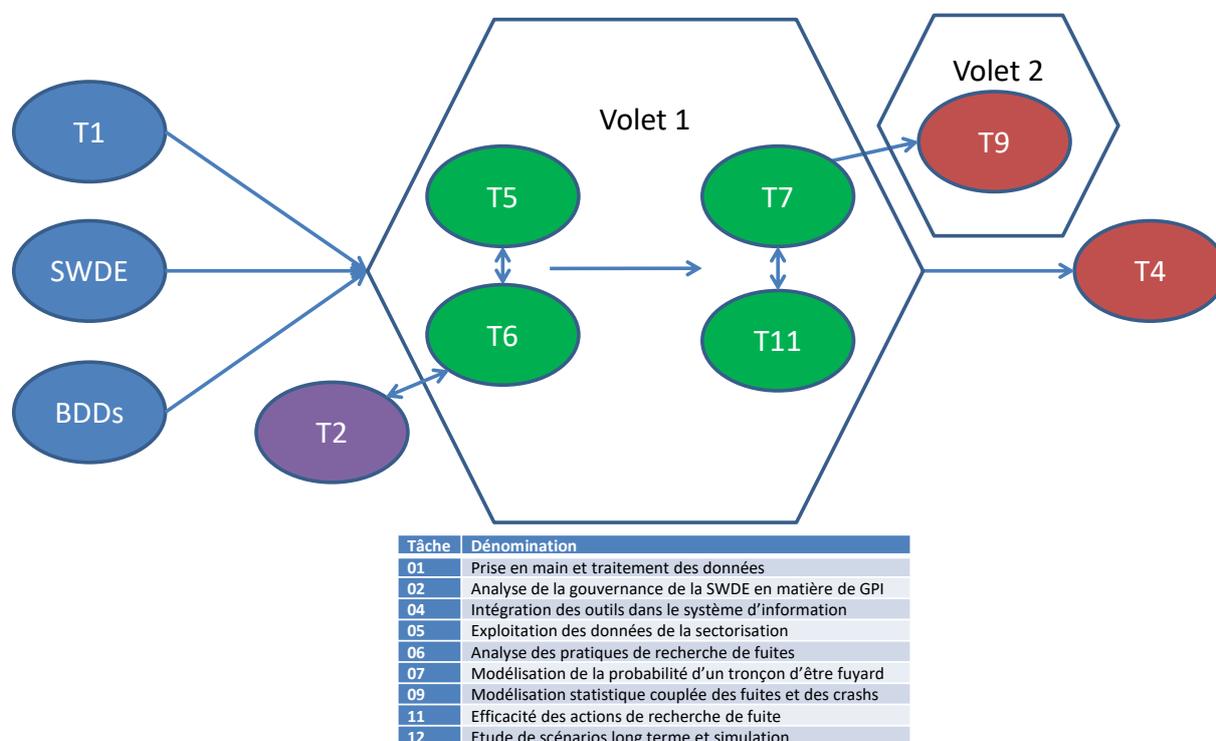


Figure 16 : Articulation de tâches du volet 1.

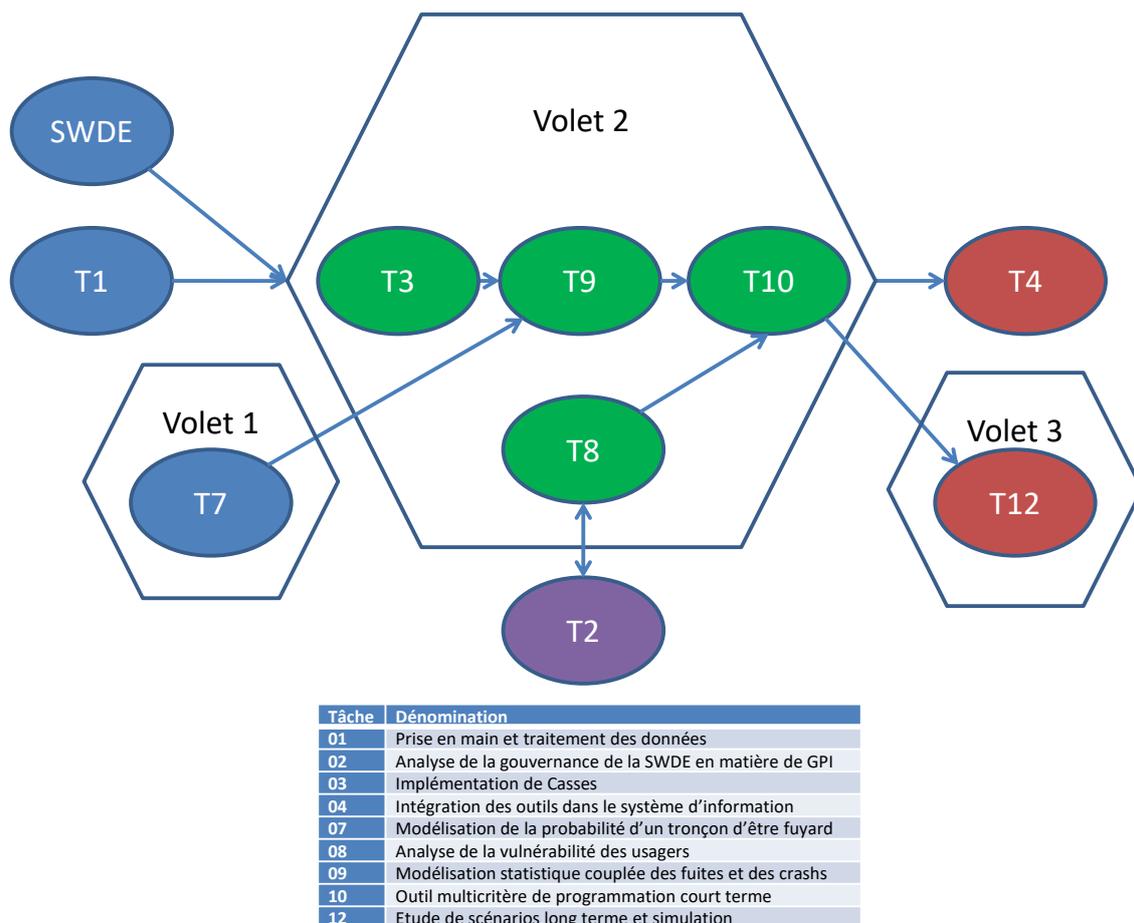


Figure 17 : Articulation de tâches du volet 2.

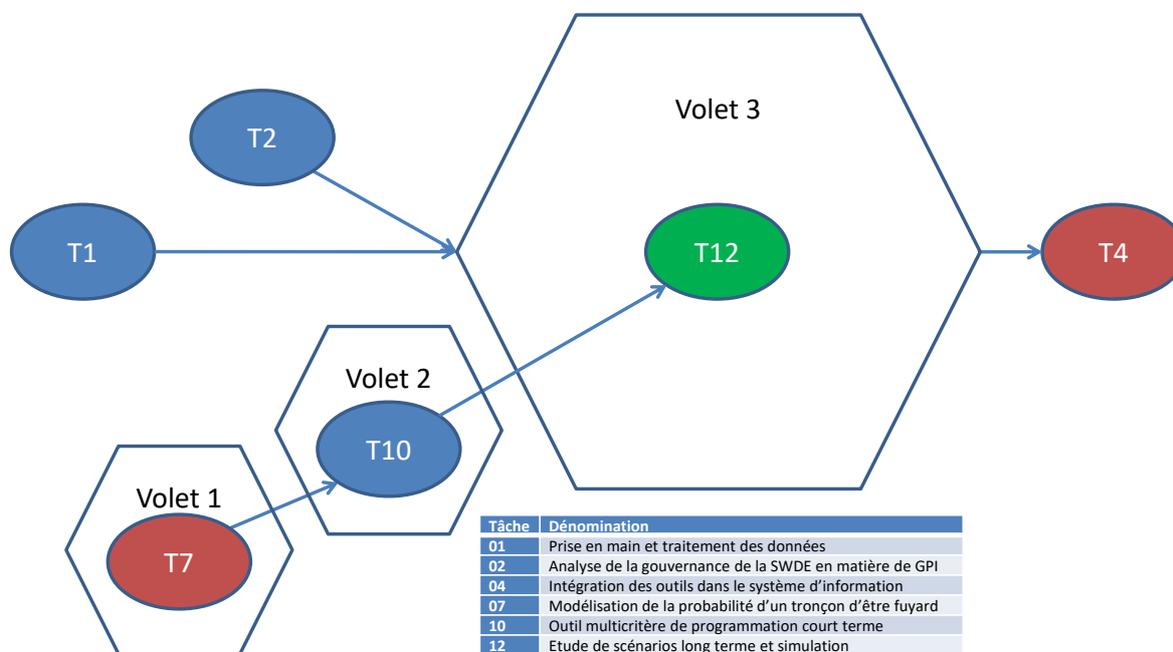


Figure 18 : Articulation de tâches du volet 3.

## Annexe 2 : Table de correspondance de colonnes des fichiers fuites de Mons

Le Tableau 11 montre les correspondances entre les numéros de colonne des fichiers fuites sources et ceux du fichier construit pour les agglomérer.

Tableau 11 : Tableau de correspondance des colonnes des fichiers fuites.

Indice de colonne des données brutes	2016-1	2016-2	2016-3	2016-4	2016-5	2016-6	2016-7	2016-8	2016-9	2016-10	2016-11	2016-12	2017	2018-2020
1	2	2	2	2	2	2	2	2	2	2	2	2	2	2
2	4	4	4	4	4	4	4	4	4	4	4	4	4	3
3	5	5	5	5	5	5	5	5	5	5	5	5	5	4
4	6	6	6	6	6	6	6	6	6	6	6	6	6	5
5	8	8	8	8	8	8	8	8	8	8	8	8	8	6
6	9	9	9	9	9	9	9	9	9	9	9	9	9	7
7	10	10	10	10	10	10	10	10	10	10	10	10	10	8
8	14	14	14	14	14	14	14	14	14	14	14	14	14	9
9	15	15	15	15	15	15	15	15	15	15	15	15	15	10
10	11	11	11	11	11	11	11	11	11	11	11	11	11	12
11	17	17	16	17	17	17	17	17	17	17	17	17	17	13
12	18	18	17	18	18	18	18	18	18	18	18	18	18	14
13	19	19	18	19	19	19	19	19	19	19	19	19	19	15
14	25	25	19	25	25	25	25	25	25	25	25	25	25	16
15	27	27	25	27	27	16	16	16	16	16	16	16	16	17
16	28	28	27	28	28	27	27	27	27	27	27	27	27	18
17	29	29	28	29	29	28	28	28	28	28	28	28	28	19
18			29			29	29	29	29	29	29	29	29	20
19						30	30	30	30	30	30	30	30	21
20						31	31	31	31	31	31	31	31	22
21 à 49														23 à 53

### Annexe 3 : Extraction des DN des pièces de réparation

Pour extraire les informations de diamètre des conduites à partir des pièces de réparation, il a été nécessaire de créer une table de correspondance entre les pièces de réparation et les diamètres des conduites compatibles. En effet, les diamètres ne sont pas toujours indiqués dans un champ séparé dans le dictionnaire de pièces, et quand c'est le cas, l'information peut être incomplète. Par exemple, les pièces de raccord entre DN différents sont indiquées avec un seul des diamètres concernés. Dans un autre exemple, un manchon de réparation donné peut être installé sur des conduites de diamètre extérieur (DE) allant de 88 à 108 mm. Aucun diamètre nominal n'est directement associé à ce manchon dans le dictionnaire, et cette information est à extraire visuellement de la désignation de l'article. Une reconnaissance automatique de texte ne pourrait se substituer à un travail visuel car les chiffres associés aux diamètres dans la désignation ne sont pas identifiables de façon systématique et fiable, ni par leur position, ni par ce qui les précède ou les suit dans la chaîne de caractères. Ces chiffres peuvent être aussi bien des DN que des DE, et seul le jugement humain permet de distinguer les cas.

Ainsi, la table de correspondance (voir Tableau 12) a été créée sur base de l'extraction visuelle des DE ou des DN présents dans la désignation de l'article. Afin de rendre le traitement faisable dans un temps limité, et utile, seules les pièces de réparation utilisées au moins 1 fois sur le secteur de Mons (période 2007-2020) ont été considérées (près de 300 pièces). Le Tableau 12 montre la méthode utilisée pour les 17 premiers articles.

Tableau 12 : Extraction des diamètres associés aux pièces de réparation liées aux OT.

Objet	Désignation article	Diamètre (DN ou DE)	DE min	DE max	DN min	DN max	Nombre d'OT de type 105
102874	PIECE REPA. 3 ELEM. FD 88-108MM		88	108	80	100	183
101782	BANDE GRASSE (ROULEAU DE 10M)						147
102108	TUYAU PE DE32 COURONNE 100 M HAUT DENS	80 mm					98
101175	RACCORD FD AUTOBL. PE DE32 X FE1"	32x1"					96
102100	TUYAU FONTE DUCTILE DN80 - L: 6 M	80 mm			80	80	88
102400	MANCHON LT PL:84-106MM		84	106	80	100	75
103575	Boulon inox TH - M16x70						66
101176	"RACCORD FD AUTOBL. PE DE32 X FI 4/4"	32x1"					60
102873	PIECE REPA. 3 ELEM. FD 70-89MM		70	89	60	80	60
101121	JOINT PLAT ELM DN80	80 mm			80	80	55
103482	"BOUCHON LAITON FE-1""	1"					49
102953	"PRISE EN CHARGE DN80 89-108MM FI-DN1""	80 mm	80	108	80	100	44
102099	TUYAU FONTE DUCTILE DN60- L: 6 M	60 mm			60	60	43
101174	MANCHON FD EE AUTOBL. PE DE32	32 mm					36
101609	MANCHON LT PL 68-85 MM		68	85	60	80	35
101227	"RACCORD VERROU. MECA. PE34 X FE1""	34x1"					31
102872	PIECE REPA. 3 ELEM. FD 65-82MM		65	82	60	80	30

Le remplissage de la table respecte les règles décrites ci-dessous :

- Pour les tuyaux de remplacement en métal ou béton, on note directement le DN : 80-80 par exemple (plage à intervalle nul).
- Pour les tuyaux de remplacement plastiques (PVC ou PE), on autorise un intervalle non nul de DN, comme par exemple DN60-63, car les matériaux plastiques sont parfois indiqués avec des DN similaires mais différents de quelques mm (comme DN60 et DN63 pour les mêmes matériaux).
- Pour les pièces de réparation comme les manchons, on note la plage de DE, comme 89-108 mm, que l'on convertit en plage de DN 80-100 en prenant pour le DE min et DE max les DN inférieurs ou égaux les plus proches qui existent dans le SIG. Si le DN est indiqué tel quel dans la désignation d'article et s'il est légèrement plus faible que le DE min indiqué, on prend le DN indiqué en désignation comme DN min.
- Certaines pièces comme les tés comportent plusieurs diamètres, comme DN80 x DN100. Dans ce cas on note la plage DN80-100 (idem si DE x DE renseigné, mais les DE sont d'abord convertis en DN).
- D'autre part, on ne considère pas les pièces de réparation qui concernent directement les branchements (robinets, compteurs, regards...). En pratique, toute pièce utilisée pour DN inférieur à 32 mm n'est pas considérée car le SIG de la SWDE ne comporte pas les tuyaux de DN inférieur à 32 mm (donc a priori pas les branchements de clients particuliers).

Remarque : de nombreux diamètres sont indiqués en pouces (symbole " , par exemple 3/4") dans le dictionnaire, il s'agit du Nominal Pipe Size (NPS), qui a été converti en DN (mm) grâce à une table de correspondance présentée dans le Tableau 13. Cette conversion est également utilisée pour les fichiers fuites lorsqu'une comparaison est requise avec les diamètres du SIG, qui sont tous en DN.

**Tableau 13 : Correspondance entre Diamètre Nominal (mm) et Nominal Pipe Size (pouces).**

NPS (pouces)	DN (mm)
3/8	10
1/2	15
3/4	20
4/4	25
1	25
5/4	32
3/2	40
6/4	40
2	50

## Annexe 4 : Validation topologique des fichiers de conduites

Certaines conduites du SIG présentent des erreurs qui sont difficiles à détecter visuellement et qui peuvent créer des blocages lors des traitements automatisés. L'exemple ci-dessous (Figure 19) montre deux conduites superposées (STARAPIC-399596 et STARAPIC-399600). Ces dernières peuvent par exemple poser des problèmes dès lors que les traitements/calculs sur les conduites du SIG partent du principe que chaque conduite est uniquement identifiée par une paire de nœuds terminaux (ce qui est supposé dans plusieurs traitements, comme l'agrégation des tronçons ou encore la complétion de données).

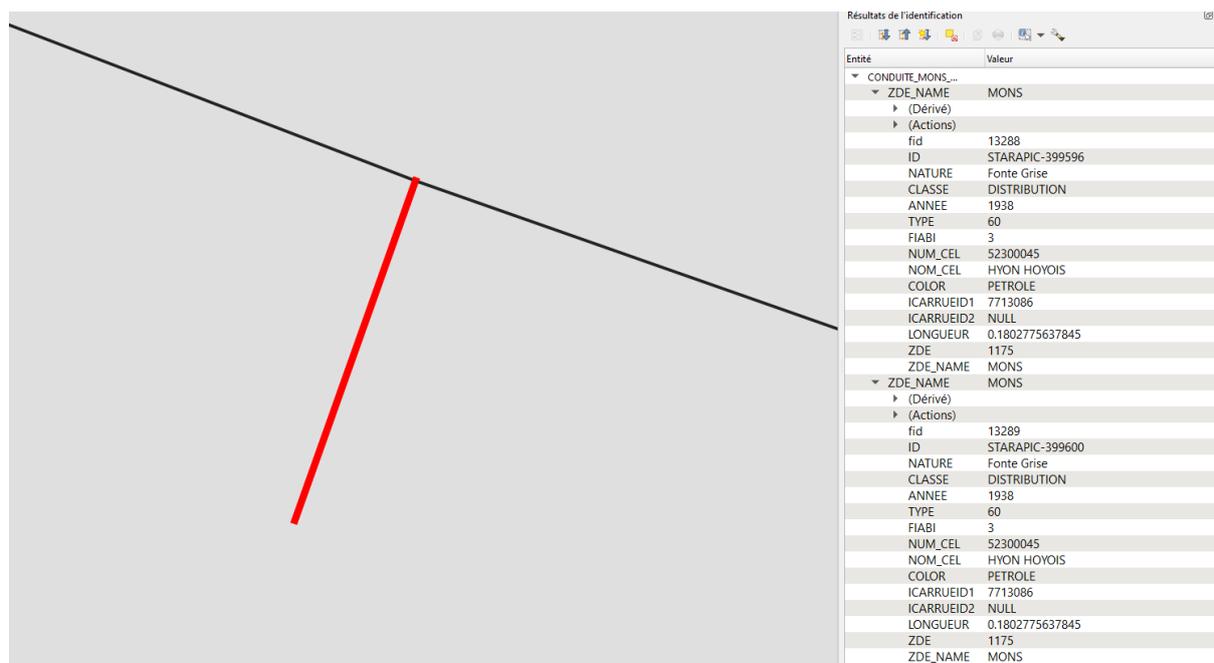


Figure 19 : Exemple de conduite erronée du SIG (superposition).

Le script développé en langage R a donc pour but de détecter tous les types d'erreurs topologiques sur le réseau de conduites et d'enregistrer les ID de conduites concernées. Il est important de noter que ces vérifications ne corrigent pas les problèmes de connectivité (rencontrés principalement pour les conduites hors service. Les vérifications effectuées (et les fonctions utilisées dans R) sont :

1. Géométrie non vide (`st_is_empty`). Ceci se produit si l'attribut `geometry` (ou `geom`) de la conduite est vide mais qu'elle possède tout de même d'autres attributs (comme un ID, etc.).
2. Géométrie valide (`st_is_valid`). Typiquement, une conduite réduite à un point est non valide.
3. Pas de multi-lignes (`MULTILINESTRING`) connexes. Ces dernières sont transformées en lignes simples (`LINestring`). Cela signifie que les conduites faites par exemple de deux morceaux qui se touchent sont alors faites d'un seul morceau.
4. Pas de multi-lignes restantes malgré la fusion l'étape précédente. Cela se produit si les conduites multi-lignes restantes sont non connexes.
5. Géométrie simple (`st_is_simple`). Typiquement, une conduite qui s'auto-intersecte n'est pas une géométrie dite « simple ».
6. Longueur de conduite non nulle. Certaines conduites peuvent avoir leurs deux nœuds terminaux confondus.

7. Pas de lacet (terme topologique pour boucle). Une conduite qui se referme sur elle-même n'est pas valide.
8. Pas de conduite(s) superposée(s). On considère qu'une seule conduite peut posséder les mêmes nœuds terminaux et faire un arc de même longueur entre ces nœuds. Dans le cas contraire, on considère qu'il s'agit d'une erreur de dessin SIG, où une conduite a été dessinée deux fois ou plus par erreur.

À chaque étape, les conduites erronées sont écartées temporairement du fichier pour ne pas impacter les vérifications suivantes (par exemple, une ligne réduite à un point est une géométrie non valide, sans l'éliminer à l'étape 2, on la retrouve à l'étape 5 comme étant erronée). Les erreurs rencontrées sont alors enregistrées, et les ID de conduites correspondants sont exportés pour un traitement manuel. En effet, seulement 3 erreurs ont été repérées par cet algorithme (Tableau 14) et corrigées.

**Tableau 14 : Conduites erronées du secteur de Mons et corrections apportées.**

<b>Conduite concernée (ID)</b>	<b>Problème</b>	<b>Action effectuée</b>
STARAPIC-399600	Superposition avec une conduite identique.	Suppression de STARAPIC-399600.
200005170	Superposition avec une conduite non identique (200005178).	Suppression de l'autre conduite superposée (200005178) qui montre un DN80 moins cohérent par rapport au réseau environnant (DN150).
200000933	Superposition avec une conduite non identique (200000932).	Suppression de l'autre conduite superposée (200000932) qui montre un DN100 moins cohérent par rapport au réseau environnant (DN150).

## Annexe 5 : Correction et complétion des conduites HS

### *Correction des problèmes de connexité*

Ce traitement est effectué dans QGIS, car les bibliothèques de fonctions SIG dans R ne possèdent pas d'outil adapté pour cette tâche spécifique. La méthode utilisée est la suivante, dans QGIS :

*Onglet Traitement -> Boîte à outils -> Accrocher les géométries à la couche -> Fichiers : Conduites HS avec Conduites HS -> Tolérance = 50 cm -> Comportement = « Accrocher les points terminaux aux points terminaux uniquement ».*

La tolérance de 50 cm signifie que seuls les sommets de conduites HS distants de 50 cm au maximum (distance euclidienne classique) seront accrochés. Ce seuil est choisi sur base de tests variés de différentes valeurs et inspection du réseau résultant. Cette valeur choisie permet de corriger le réseau sans modifier sa topologie de façon notable.

Le comportement choisi permet d'éviter la création de nouvelles conduites par section d'une ou plusieurs conduites existantes via la création de nouvelles extrémités terminales. C'est encore une fois une façon de modifier au minimum la topologie initiale du réseau tout en apportant les corrections nécessaires.

Cette correction de connexité est indispensable pour le traitement suivant qui identifie les conduites HS qui sont connexes.

### *Complétion des données manquantes*

#### Caractéristiques autres que la DDP

Les conduites HS montrent parfois des manques importants d'informations techniques dans leurs attributs. Les informations suivantes sont manquantes dans les proportions indiquées :

- 5 % des matériaux (MAT).
- 20 % des dates de mise hors service (DHS).
- 71 % des dates de pose (DDP).
- 2 conduites n'ont pas de diamètre (DIA) renseigné.

De plus, contrairement aux conduites en service, les conduites HS n'ont pas reçu de numéro de cellule (les cellules ont commencé à être définies après 2016-2017 par la SWDE), ce qui pose problème pour l'étude à venir de la sectorisation (tâche 05) qui repose sur le fait qu'une conduite appartient à une cellule.

La même méthode est employée pour compléter toutes les informations manquantes des conduites HS, autres que la DDP. Cette méthode part du principe que le réseau remplaçant est généralement situé au même endroit que le réseau remplacé et qu'il contient des informations valides sur le réseau remplacé (Figure 20). On considère aussi le réseau « encadrant », qui se situe autour du réseau remplaçant et possède généralement les mêmes caractéristiques que le réseau remplacé (Figure 20). Il est important de noter que le réseau en service utilisé pour cette étape spécifique est le réseau de

conduites en service agrégées (voir 2.2.5) car ces dernières permettent d'accéder aux informations manquantes de façon plus simple (conduites ES moins nombreuses et plus longues).

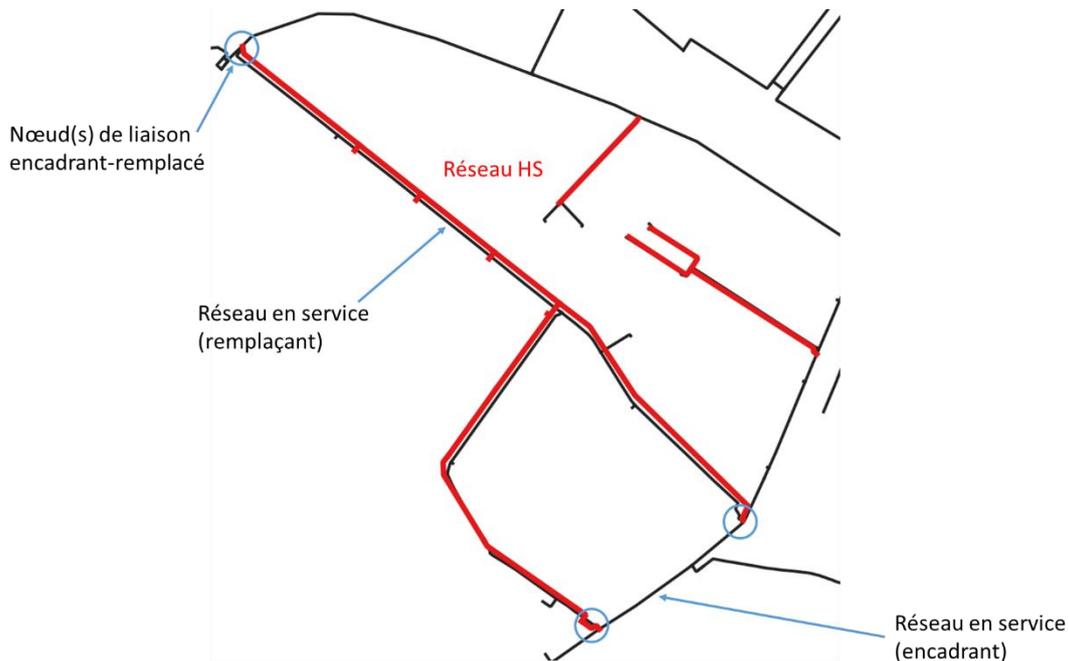


Figure 20 : Juxtaposition des graphes des conduites HS (rouge) et ES (noir).

Les opérations suivantes sont effectuées :

1. Création d'un tampon de largeur totale 8 m autour des conduites HS.
2. Interception des conduites en service (agrégées) dans ce tampon. Il s'agit du réseau remplaçant.
3. Si la conduite HS possède une DHS renseignée, la DDP du tronçon remplaçant est comparée. Si  $DHS \leq DDP + 1$ , le numéro de cellule de la conduite ES est attribué à la conduite HS (une tolérance de 1 an est utilisée afin de prendre en compte l'incertitude des dates). Si plusieurs conduites remplaçantes sont possibles sous ces conditions, celle qui possède la plus grande longueur dans le tampon est choisie pour cette étape.
4. Si la conduite HS n'a pas de DHS renseignée, la DHS est supposée égale à la DDP du tronçon remplaçant et le numéro de cellule de ce dernier lui est attribué. Cependant, seules les conduites remplaçantes dont  $DDP \geq 2011$  sont utilisées à cette étape. En effet, le réseau de conduites HS semble contenir principalement les conduites abandonnées à partir de 2011 environ (DHS antérieure à 2011 pour seulement 11 conduites sur 975 dans les données brutes).
5. Récupération du numéro de cellule pour les conduites HS ayant intercepté des conduites en service mais n'ayant pas encore récupéré de numéro de cellule aux étapes précédentes (attestant de l'incompatibilité entre DDP et DHS). Si plusieurs conduites en service sont possibles sous ces conditions, celle qui possède la plus grande longueur dans le tampon est choisie pour cette étape.
6. Propagation des numéros de cellule et des DHS de proche en proche dans le réseau HS. Les tronçons dont l'attribut recherché (cellule, DHS...) est inconnu reçoivent l'information de leur voisin le plus long qui possède cette information. Le processus est répété itérativement jusqu'à ce que le nombre de tronçons dont l'information est manquante ne change plus d'une itération à l'autre.

Enfin, pour compléter les autres informations parfois manquantes sur le réseau HS, les diamètres et matériaux sont aussi propagés au sein du réseau de conduites HS avec la même méthode que le point 6 ci-dessus.

Remarque : la complétion des diamètres de cette façon pourrait causer des incohérences si l'information n'était pas suffisamment présente à la base (notamment parce que le diamètre décroît généralement avec la distance réservoir-abonné). Cependant ce n'est pas le cas puisque seulement 2 diamètres sont manquants.

### Complétion de la DDP

La méthode précédente ne permet pas de récupérer la DDP du réseau HS. Cependant, il est assez facile de voir que la DDP du réseau en service situé autour des conduites HS (i.e., le réseau « encadrant », voir Figure 20) est un bon indicateur de la DDP de ces dernières. Il est possible d'obtenir cette information par un lien de proximité entre les nœuds du réseau HS et les nœuds du réseau ES encadrant (voir Figure 20). La méthode suivante a été utilisée :

1. Chaque nœud terminal de conduites HS est associé au nœud terminal du réseau ES le plus proche dans un rayon de 10 m maximum autour du nœud HS. Il s'agit du réseau encadrant.
2. Les caractéristiques des conduites HS et des conduites encadrantes ainsi liées sont comparées. Parmi la (les) DDP(s) disponible(s) associée(s) au nœud terminal de la (des) conduite(s) encadrante(s), la DDP de la conduite encadrante la plus longue qui satisfait les critères suivants par ordre de préférence est attribuée à la conduite HS :
  - a.  $DHS - DDP \geq 5$  ans, même matériau et même diamètre entre conduite HS et conduite encadrante.
  - b.  $DHS - DDP \geq 5$  ans et même matériau entre conduite HS et conduite encadrante.
  - c.  $DHS - DDP \geq 5$  ans et même diamètre entre conduite HS et conduite encadrante.
  - d.  $DHS - DDP \geq 5$  ans.
3. Propagation de(s) DDP(s) ainsi obtenue(s) aux extrémités du réseau HS de proche en proche dans le réseau HS. Les tronçons dont la DDP est inconnue reçoivent l'information de leur voisin le plus long qui possède cette information. Le processus est répété itérativement jusqu'à ce que le nombre de tronçons dont la DDP est manquante ne change plus d'une itération à l'autre.

Il est important de noter que pour cette complétion de DDP, contrairement aux traitements de complétion précédents (DHS, cellule, MAT, DIA), les conduites ES non-agrégées (i.e., données brutes) sont employées comme réseau encadrant, car l'agrégation supprime des nœuds terminaux du réseau ES et rend la recherche de correspondance entre nœuds de réseau HS et nœuds de réseau ES plus difficile et moins fiable.

Remarque : Certaines désignations de matériaux sont considérées comme équivalentes lors de la comparaison de matériaux entre conduites HS et conduites en service afin de ne pas limiter les correspondances entre conduites HS et conduite ES pour des désignations vraisemblablement similaires. Voir Tableau 17.

## Annexe 6 : Agrégation des conduites

Le Tableau 15 résume les quatre étapes employées pour agréger les tronçons.

Tableau 15 : Détail des quatre étapes de l'agrégation des conduites.

	<b>Etape 1</b>	<b>Etape2</b>	<b>Etape 3</b>	<b>Etape 4</b>
<b>Longueur</b>	≤ 20 m	≤ 3 m	≤ 6 m	≤ 3 m
	Tous les tronçons	Tronçons sauf déjà agglomérés	Tous les tronçons	Sauf tronçons [Nœud triple – nœud simple]
<b>Cellule identique</b>	oui	oui	oui	oui
<b>MAT identique</b>	oui	oui	non	non
<b>DIA identique</b>	oui	oui	non	non
<b>DDP identique</b>	oui	oui	non	non
<b>Angle nœud double</b>	≤ 45°	Pas de condition	≤ 45°	Pas de condition
<b>Angle nœud triple</b>	≤ 15° (le plus petit si plus d'un)	≤ 15° (le plus petit si plus d'un)	≤ 15° (le plus petit si plus d'un)	≤ 45° (le plus petit si plus d'un)
<b>Caractéristiques du tronçon agrégé</b>			Prend les caractéristiques du plus long	Prend les caractéristiques du plus long
<b>Priorité des nœuds</b>	On privilégie toujours une fusion via un nœud double par rapport à un nœud triple. Si deux nœuds de même types, on fusionne avec le tronçon le plus long			
<b>Règle de continuité des tronçons</b>	Il n'y a jamais deux fusions sur un même nœud (un tronçon est toujours une ligne continue entre deux nœuds)			

La signification des différents intitulés de lignes est détaillée ci-dessous (par convention, on appelle « tronçon absorbant » un tronçon sur lequel on agrège un autre tronçon plus court, « nœud double » un nœud qui réunit deux tronçons et « nœud triple un nœud qui réunit trois tronçons ou plus) :

- Longueur : la longueur maximale (en m) que peut avoir un tronçon pour pouvoir être agrégé à cette étape.
- Cellule identique : le tronçon considéré doit-il avoir le même numéro de cellule que l'absorbant pour pouvoir être agrégé (oui/non) ? Idem pour les autres caractéristiques (MAT = matériau, DIA = diamètre, DDP = date de pose).
- Angle nœud double : angle maximum (en degrés) que peut former le tronçon considéré avec son unique voisin de ce côté pour pouvoir être absorbé par ce dernier (voir Figure 21).
- Angle nœud triple : angle maximum (en degrés) que peut former le tronçon considéré avec un de ses multiples voisins de ce côté pour pouvoir être absorbé par un voisin de ce côté (voir Figure 22).
- Caractéristiques du tronçon agrégé : comment sont mises à jour les caractéristiques (cellule, MAT, DIA, DDP) du tronçon agrégé lorsqu'il est accroché à son absorbant ?
- Priorité des nœuds : comment choisit-on d'agréger un tronçon par un de ses côtés plutôt que par l'autre ?
- Règle de continuité des tronçons : règle imposée pour empêcher la création de tronçons avec plus de deux nœuds terminaux.

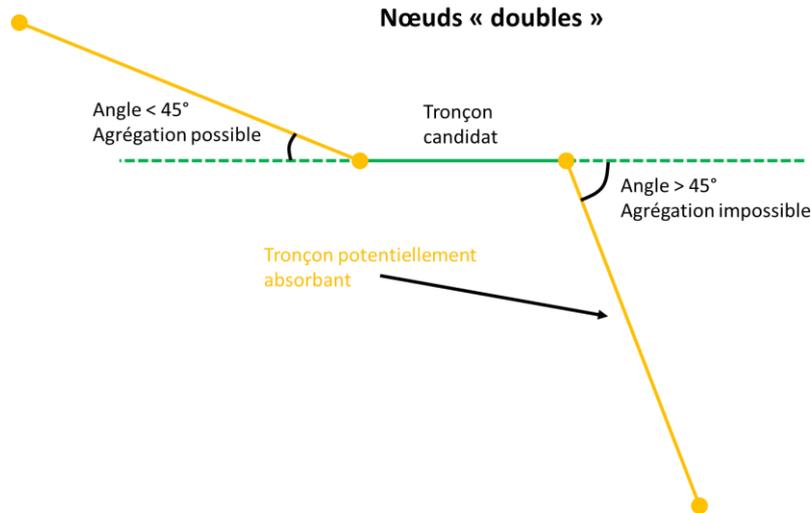


Figure 21: Illustration des nœuds doubles.

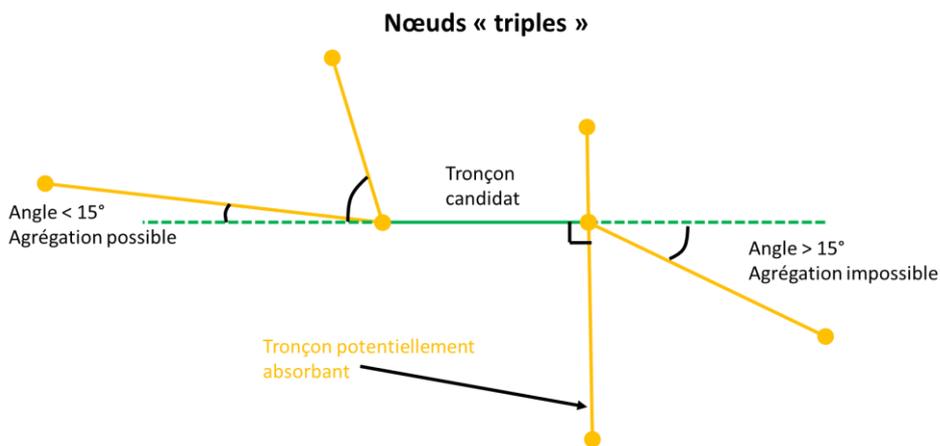


Figure 22 : Illustration des nœuds triples.

Le calcul des angles entre tronçons est effectué par différence absolue entre les deux angles par rapport à la flèche du Nord formés par les tronçons considérés. On soustrait si nécessaire 180 degrés aux angles afin d'éviter de différencier les orientations N->S et S->N, et O->E et E->O entre elles. Nommons  $A_1$  et  $A_2$  les angles entre 0 et 360 degrés formés par chacun des deux tronçons avec la flèche Nord. Alors, on calcule :

$$A_1' = A_1 - 180 \text{ si } A_1 \geq 180^\circ, \text{ sinon } A_1' = A_1.$$

$$A_2' = A_2 - 180 \text{ si } A_2 \geq 180^\circ, \text{ sinon } A_2' = A_2.$$

$$A = A_1' - A_2'$$

$$A' = \left| |A_1' - A_2'| - 180 \right|.$$

Lorsque l'on souhaite vérifier que deux conduites font un angle de moins de  $\theta$  entre elles, on vérifie alors que :

$$A < \theta \text{ ou } A' < \theta.$$

La logique d'adaptation des 4 étapes est la suivante.

1. La première étape est la plus générale, mais aussi la plus restrictive afin de ne regrouper que les tronçons très semblables. Les critères d'angles stricts sont utilisés pour ne pas agréger des tronçons situés dans des rues différentes et pour conserver le caractère rectiligne des tronçons autant que possible. Une majorité des tronçons sont correctement agrégés à leur voisin à cette étape.
2. La deuxième étape s'affranchit seulement du critère d'angle sur nœud double et se concentre sur les tronçons non agrégés de moins de 3 mètres de long afin d'agréger un grand nombre de petits tronçons qui forment un angle global (créé par les nœuds terminaux uniquement) trop élevé avec leur voisin pour être agrégés lors de la première étape bien que l'inspection visuelle indique un angle local (créé par les nœuds intermédiaires) faible suggérant naturellement une agrégation (voir Figure 23). Le critère de longueur maximale de 3 m est choisi pour ne pas permettre à une conduite formée de plusieurs morceaux à l'issue de cette étape de posséder une grande partie de sa longueur dans une autre rue (ce que le critère d'angle sur nœud double empêche normalement). Pour les étapes 1 et 2, toutes les caractéristiques sont identiques entre tronçons éventuellement regroupés donc elles ne changent pas lors de l'accrochage de tronçons entre eux.
3. La troisième étape s'affranchit seulement des critères sur les caractéristiques identiques (en conservant le critère d'angle) et se focalise sur les tronçons de moins de 6 mètres de long afin de se débarrasser des petites conduites aux caractéristiques différentes de leurs voisins empêchant la propagation de l'agglomération de proche en proche. En effet, il arrive fréquemment qu'un paquet de petits tronçons similaires ne soit pas regroupé à cause d'une seule conduite aux caractéristiques différentes dans ce paquet. Puisqu'il n'y a plus nécessairement de concordance sur les caractéristiques à cette étape, ces dernières sont mises à jour à cette étape pour le tronçon absorbé, qui reçoit les caractéristiques du plus long tronçon dans le groupe qui l'absorbe.
4. La dernière étape est la plus souple. Elle se focalise sur les plus petits tronçons (moins de 3 mètres) qui ont encore résisté à l'agrégation à cause de leur angle important formé avec les conduites voisines. Ainsi, il n'y a pas de critère de caractéristiques identiques, ni d'angle sur les nœuds doubles, et l'angle maximal sur les nœuds triples est rehaussé à 45 degrés. Les caractéristiques des tronçons sont mises à jour comme dans l'étape précédente. Cependant, les tronçons candidats à l'agrégation à cette étape qui sont délimités par un nœud simple d'un côté (c'est-à-dire, pas de voisin de ce côté) et un nœud triple, qui en général correspondent à des piquages, ne sont pas agrégés pour empêcher la création de tronçons sans réalité technique qui forment des angles presque droits à leur extrémité (voir Figure 24).

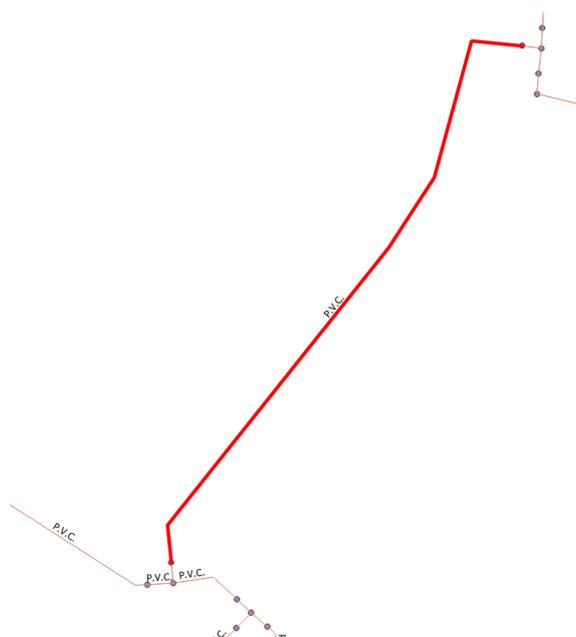


Figure 23 : Tronçon dont les angles globaux avec ses voisins diffèrent nettement des angles locaux.



Figure 24 : Tronçons « nœud simple-nœud triple » exclus à l'étape 4.

Le Tableau 16 suivant récapitule comment le nombre de tronçons évolue à chaque étape. Au final, le nombre total de tronçons est presque divisé par 2, et le nombre de tronçons de longueur inférieure à deux seuils jugés adéquats (6 et 20 m) est grandement réduit (par un facteur allant de 4 à 7).

Tableau 16 : Décompte du nombre de tronçons à chaque étape de l'agrégation.

		Etape 1	Etape 2	Etape 3	Etape 4
Longueur	SIG du 21-02-22	≤ 20 m	≤ 3 m	≤ 6 m	≤ 3 m
Total Tronçons	31 822	19 648	18 629	17 082	16 875
Dont moins de 20 m	19 964	7 544	6 522	4 957	4 745
Dont moins de 6 m	15 706	4 896	3 856	2 276	2 064

## Annexe 7 : Attribution d'ICAR de rue aux conduites

### Description de la méthode

Cette méthode consiste à dessiner des tampons à bouts rectangulaires autour des axes de voirie puis à faire l'intersection entre les conduites et ces tampons pour donner à chaque conduite les attributs d'un axe de voirie, en particulier l'ICAR de rue. Pour une largeur de tampon choisie, certaines conduites ne sont pas interceptées par les tampons comme l'illustre la Figure 25 (les conduites sont représentées en noir et les tampons en rose). Face à cette difficulté, la solution proposée est de répéter l'opération ci-dessus de façon itérative avec des tampons de largeur croissante, afin que la quasi-totalité des conduites reçoivent un ICAR de rue.



Figure 25 : Conduites interceptées ou non par un tampon autour des axes de voirie.

Dès la première itération où une conduite est située pour au moins à 50 % de sa longueur dans le tampon d'un axe de voirie, elle en reçoit l'ICAR de rue. Ce seuil de 50 % permet de sélectionner parmi plusieurs rues possibles, celle dans laquelle la conduite est majoritairement située (ou le plus proche). Une conduite ayant déjà reçu un ICAR le conserve (ses intersections avec d'autres tampons de plus grande taille lors des itérations suivantes ne le modifie pas). Ainsi, l'ICAR de rue du plus petit tampon est prioritaire, car c'est dans cette rue que la conduite se situe le plus vraisemblablement. Lorsque plusieurs ICAR de rue sont possibles pour une taille de tampon donnée (c'est-à-dire, lorsqu'une conduite n'a pas encore reçu d'ICAR de rue aux largeurs de tampon inférieures), on choisit l'ICAR du tampon qui intercepte la plus grande longueur de conduite, puis celui dont l'axe de voirie fait l'angle le plus faible avec la conduite si le premier critère est insuffisant pour faire un choix (angle calculé comme pour l'agrégation des conduites présenté en Annexe 6).

Le pseudo-code (implémenté en langage R) est décrit ci-dessous.

Charger les couches de données :

- Conduites agrégées
- Axes de voirie

[Tous les axes de voirie avec ICAR de rue vide ou identifiant GEOREF\_ID doublon sont éliminés]

### **Pour des tailles de tampon croissant de 2 à 100 m de largeur totale**

*Calculer les intersections des tampons avec les conduites pour la taille de tampon actuelle. On obtient des morceaux de conduites (1 ou plus par conduite intersectée) portant chacun l'ICAR du tampon qui le découpe sur la conduite*

*Parmi tous les morceaux de conduite obtenus suite à l'intersection, garder seulement le(s) morceau(x) le(s) plus long(s), par ID de conduite*

*Parmi tous les morceaux de conduite restants, garder seulement celui (ceux) qui fait (font) l'angle le plus faible avec l'axe de voirie en relation, par ID de conduite*

***S'il reste plus d'un ICAR de rue parmi les morceaux restants pour un ID de conduite donné malgré les critères de longueur et d'angle ci-dessus***

*Prendre un ICAR aléatoirement parmi ceux disponibles pour cette taille de tampon  
[Note : ceci a été utilisé pour une seule conduite pour les données de Mons]*

### **Fin de si**

*Pour chaque ID de conduite, éliminer l'ICAR de rue candidat pour cette taille de tampon si la longueur d'intersection totale de la conduite avec le tampon est  $< 50\%$  de la longueur de cette conduite [cette longueur d'intersection totale compte en multiple les morceaux de conduite qui seraient dans des tampons superposés et de même ICAR de rue, mais elle reste bornée à la longueur de conduite]*

*Pour chaque ID de conduite, lui attribuer l'unique ICAR de rue sélectionné pour la taille de tampon actuelle si celui-ci n'est pas déjà attribué par des tampons plus petits*

### **Fin de boucle**

## **Validation de la méthode**

Une question soulevée lors des échanges avec la SWDE est de savoir dans quelle mesure l'agrégation des conduites ne mène pas à attribuer un ICAR de rue erroné à une conduite agrégée qui a une partie importante de sa longueur dans une autre rue que celle qui donne son ICAR à la conduite absorbante. La conduite représentée en rouge Figure 26 illustre cette situation.



Figure 26 : Exemple de conduite agrégée qui change d'ICAR de rue du fait de son agrégation.

Ces changements de rue pourraient compromettre la bonne attribution de certaines défaillances. A priori, ce phénomène est limité par les critères d'angles au cours de l'agrégation, qui évitent au maximum un changement de rue pour une conduite agrégée.

Afin de vérifier que ce n'est pas souvent le cas, un décompte du nombre de conduites dans cette situation a été effectué. Il y a 3439 conduites (soit 12 km) sur Mons qui ont changé d'ICAR suite à l'agrégation, soit 11 % des conduites en nombre, et seulement 0.8 % du linéaire de Mons, car ce sont souvent de petites conduites dans les carrefours qui changent d'ICAR. Parmi ces 3439 conduites, 107 (totalisant 2.4 km) sont des conduites ayant absorbé d'autres conduites, ce qui représente une fraction négligeable du réseau de Mons (> 1500 km). Ainsi, la majorité des conduites ayant changé d'ICAR sont les conduites ayant été absorbées par une autre conduite (au nombre de 3332, totalisant 9.6 km).

Parmi les conduites absorbées ayant changé d'ICAR (3332), seulement 492 font plus de 6 m de longueur (totalisant 5.3 km de longueur). Les autres 2840 conduites (totalisant 4.3 km) sont de taille réduite (< 6 m) et généralement situées dans les carrefours où la notion d'ICAR de rue est à la base ambiguë, et donc où le changement d'ICAR causé par l'agrégation est un effet souhaité. En effet, un des buts de l'agrégation est d'aider la sélection de l'ICAR de rue pour ces conduites.

Parmi ces 492 conduites absorbées de longueur supérieure à 6 m et ayant changé d'ICAR, il reste à déterminer le nombre de conduites qui n'auraient pas dû changer d'ICAR. La méthode employée est la suivante :

- Attribution des ICAR de rue avec la méthode décrite ci-dessus, pour les conduites en service non agrégées (donnant un ICAR de rue nommé « ICAR\_SIMPLE », Figure 26).
- Attribution des ICAR de rue avec la méthode décrite ci-dessus, pour les conduites en service agrégées (donnant un ICAR de rue nommé « ICAR\_AGR », Figure 26).

- Sélection des conduites absorbées de longueur supérieure à 6 m et dont l'ICAR de rue a changé avant/après absorption (ICAR\_AGR  $\neq$  ICAR\_SIMPLE, Figure 26). Ce seuil de longueur est choisi pour éviter de considérer les petits tronçons situés au milieu de carrefours.
- Création de tampons de largeur totale 20 m autour des axes de voirie dont l'ICAR est dans la liste des ICAR de rue après agrégation (ICAR\_AGR) des conduites absorbées du point ci-dessus.
- Intersection des conduites et des tampons décrits ci-dessus.
- Décompte du nombre de conduites qui ont moins de 50 % de leur longueur dans le tampon qui a donné l'ICAR après absorption (ICAR\_AGR) ou qui ne sont pas du tout dans ce tampon. Ces conduites sont bien celles qui ont indûment changé d'ICAR de rue.

Au final, seulement 120 conduites sont comptées comme décrit ci-dessus, ce qui représente moins de 1 % du total de conduites sur Mons, mais tout de même 7 % des conduites absorbées de plus de 6 m de long. Ainsi, l'agrégation a un impact limité sur l'attribution des ICAR pour les conduites agrégées à l'interface de plusieurs rues.

## Annexe 8 : Attribution des défaillances aux conduites

### *Hiérarchie des sources d'informations*

Les informations dans les fichiers sources (SAP, fichiers fuites, anomalies SIG, pièces de réparation...) sont utilisées dans cet ordre de priorité :

#### *Concernant la localisation de la défaillance*

1. Anomalies reportées au SIG
2. Adresse complète liée à l'OT (ICAR de rue + numéro d'immeuble)
3. ICAR de rue seul lié à l'OT

#### *Concernant les sources d'adresse postale*

1. Issues du « poste technique » (objet géo-localisé défini dans SAP)
2. Issue des fichiers fuites
3. Issues du champ nommé « Rue » de SAP

*Concernant la sélection d'une conduite parmi les conduites du SIG sélectionnées par l'ICAR de rue ou par proximité du point d'adresse de la défaillance, pour un OT donné*

1. Même ICAR de rue que l'OT et caractéristiques compatibles entre SIG et fichiers fuites et pièces de réparation
2. Même ICAR de rue que l'OT et caractéristiques compatibles entre SIG et fichiers fuites
3. Même ICAR de rue que l'OT et caractéristiques compatibles entre SIG et pièces de réparation
4. Même ICAR de rue que l'OT
5. Étapes 1 à 4 sans conditions d'ICAR de rue mais un rayon de 200 mètres maximum autour du point d'adresse de la défaillance pour les OT avec numéro d'immeuble

#### *Concernant la date de casse*

1. Date de réparation du fichier d'anomalies SIG
2. Date de constatation de l'anomalie SIG
3. Date de réparation dans les fichiers fuites
4. Date de localisation dans les fichiers fuites
5. Date de début planifié dans SAP

#### *Concernant le type de défaillance*

Pour le type d'activité 105 (réparation de conduite), on considère que les OT de type « CLT », c'est-à-dire issus d'un appel client, sont des crashes et que les OT de type « RFU » ou « DIST » sont des fuites réparées après RF.

### *Pseudo-code*

Le pseudo-code suivant décrit la méthode employée pour lier chaque OT présent dans SAP à une conduite du SIG, i.e., à son identifiant (ID). Il correspond à une implémentation dans un code R.

Le texte respecte une indentation montrant la hiérarchie des blocs de traitements et les conditions pour accéder à ces blocs de traitement. Des mots clés comme « si », « sinon », « ou », « et », qui ont

une forte signification logique, sont surlignés pour rendre les expressions formulées comme des conditions plus visibles. Des marqueurs (flags), comme [05], ont été utilisés pour marquer les nœuds, c'est-à-dire les situations où plusieurs choix sont possibles, ou bien les actions résultant de conditions préalables. En particulier, certains blocs d'instructions débutent et finissent par un mot clé, comme [SAP], qui identifie une étape clé du traitement avec un libellé approprié (dans ce cas, l'utilisation de données dans l'extrait SAP). Ces flags sont essentiels pour la compréhension du code R, qui se base majoritairement sur la reconnaissance de ces marqueurs de situation afin de ne pas rendre les tests logiques (si... alors...) de plus en plus longs et complexes au fur et à mesure du parcours de l'algorithme.

Remarque 1 : Il est important de bien exécuter dans l'ordre d'écriture les conditions en [SI...FIN DE SI] qui se succèdent au sein d'un même niveau d'indentation. L'ordre y est important, au contraire des instructions de type [SI...SINON...FIN DE SI] qui peuvent être évaluées dans n'importe quel ordre.

Remarque 2 : de nombreuses lignes (donc numéros de nœuds) ont été supprimées lors des versions successives au cours du développement de cette méthode, et les numéros de nœuds n'ont pas été ajustés au fur et à mesure afin de ne pas avoir à modifier ces derniers de façon rétroactive dans le code R, ce qui mènerait très probablement à des erreurs.

De façon simplifiée, l'algorithme se résume ainsi :

Considérer chaque OT de SAP :

1. Tester si l'OT est déjà dans le SIG (ANOMAL). Si oui, c'est terminé.
2. Sinon, chercher son adresse aussi complète que possible. Considérer d'abord le poste technique de SAP, puis le fichier FUTTES, puis le champ « Rue » de SAP.
3. Une fois le point d'adresse ICAR identifié, considérer les conduites proches de longueur  $\geq 6$  m qui possèdent le même ICAR de rue.
4. Si (2) échoue, utiliser seulement l'ICAR de rue comme adresse de la conduite (priorité : poste technique, fichier FUTTES, puis « Rue »).
5. Si une seule conduite est identifiée par (3) ou (4), c'est terminé.
6. Si plusieurs conduites sont possibles pour cet ICAR (point et/ou rue), tenter de les discriminer selon caractéristiques dans le fichier FUTTES et/ou dans les pièces de réparation si nécessaire, en éliminant les conduites de moins de 6 m.
7. Si un conflit persiste, choisir la conduite la plus longue de la rue, et sauvegarder divers indicateurs d'incertitude de l'attribution.

### Début Pseudo code

Partir du fichier SAP (considérer chaque ligne)

**SI** le type d'activité est 92 **OU** 105 **OU** parmi (11 ; 82 ; 114 ; 50 ; 113 ; 30 ; 180 ; 100 ; 13 ; 31 ; 32 ; 112 ; 72 ; 185 ; 10) **[01]**

**SI** l'OT est présent dans le fichier ANOMAL **[02]** **(act)** pour conduites actuelles // **(ab)** pour conduites abandonnées

Étape **[SIG]**

**FIN DE TRAITEMENT [OK]**

Fin de l'étape SIG

**SINON [03]**

Étape **[SAP]**

**SI** le type d'activité est [92 **OU** 105] **[08]** **OU** l'OT est dans le fichier fuites **[09]**

**SI** l'adresse du poste technique est complète (rue + numéro) **[13]** :

Utiliser l'adresse du Poste Technique **(FULL)** si correspondance totale avec BD ICAR, **(DIG)** si le numéro sans complément a dû être utilisé pour correspondance avec BD ICAR, et **(NA)** si pas de correspondance BD ICAR

**SI** pas de correspondance BD ICAR (13(NA)) aller au nœud 14 sans l'évaluer **FIN DE SI**

Étape Adresse **[AD]**

Identifier les N (valeur proposée : 5) conduites plus proches du point d'adresse qui sont de longueur suffisante (valeur proposée : 6 mètres)

Supprimer les conduites qui n'ont pas le même ICAR de rue que le point d'adresse

Supprimer les conduites en service avec une DDP postérieure à la défaillance, avec une tolérance de A années (valeur proposée : 1 an)

Supprimer les conduites abandonnées avec une DHS antérieure à la défaillance, avec une tolérance de A années (valeur proposée : 1 an)

**SI** l'activité n'est pas spécifiquement de la réparation de conduite (type 105)

Prendre la conduite la plus proche du point d'adresse [SIG.1]

Aller à l'étape SIG

**SINON**

Parmi les conduites candidates, sélectionner dans cet ordre de préférence, la conduite la plus proche qui :

- Possède des caractéristiques cohérentes avec le fichier FUITES et avec les PIECES de réparation

- Possède des caractéristiques cohérentes avec le fichier FUITES seulement, si aucune conduite n'est trouvée à l'étape précédente

- Possède des caractéristiques cohérentes avec les PIECES de réparation seulement, si aucune conduite n'est trouvée à l'étape précédente

- (aucun critère, prendre simplement la plus proche si aucune n'est trouvée aux étapes précédentes)

**SI** une conduite a été trouvée malgré les filtres ci-dessus [SIG.1]

Aller à l'étape SIG

**SINON**

Effectuer une seconde passe comme ci-dessus (reprendre à « étape Adresse ») **mais sans le critère de l'ICAR de rue**

**SI** une seule conduite a été trouvée [SIG.2]

Aller à l'étape SIG

**SINON**

**FIN DE TRAITEMENT [END]**

FIN DE SI

FIN DE SI

FIN DE SI

Fin de l'étape Adresse

FIN DE SI

**SI** l'adresse du poste technique est incomplète (rue seule, ou vide) [14]

**SI** l'OT SAP est dans le fichier fuites [22]

**SI** une seule fuite est associée à l'OT **OU** il y a cohérence des adresses entre les doublons de la fuite [32]

**SI** l'adresse dans le fichier fuites est complète [43]

Utiliser l'adresse du fichier fuites (FULL) si correspondance totale avec ICAR, (DIG) si le numéro sans complément a dû être utilisé pour correspondance avec ICAR, (APPR) si l'ICAR est tiré d'un match partiel sur les noms de rue, (ADIG) si le numéro sans complément a dû être utilisé pour correspondance avec ICAR, et (NA) si pas de correspondance ICAR

**SI** pas de correspondance BD ICAR (43(NA)) aller au nœud 44 sans l'évaluer **FIN DE SI**

Aller à l'étape Adresse

**SI** l'adresse dans le fichier fuites est incomplète [44]

**SI** le champ « Rue » de SAP est complet [60]

**SI** la rue de « Rue » et la rue du poste technique sont cohérentes **OU**

l'ICAR du poste technique est vide [66]

Utiliser l'adresse du champ « Rue » de SAP (FULL) si correspondance totale avec ICAR, (DIG) si le numéro sans complément a dû être utilisé pour correspondance avec ICAR, (APPR) si l'ICAR est tiré d'un match partiel sur les noms de rue, (ADIG) si le numéro sans complément a dû être utilisé pour correspondance avec ICAR, et (NA) si pas de correspondance ICAR

**SI** pas de correspondance BD ICAR (66(NA))

Tenter successivement de récupérer ICAR de rue du PT [iP], de FUITES [iF], et de « Rue » [iR], puis aller à l'étape ICAR

SI tous les ICAR sont vides

**FIN DE TRAITEMENT [END]**

FIN DE SI

Aller à l'étape Adresse

SINON [66bis]

Prendre l'ICAR du poste technique

Aller à l'étape ICAR

FIN DE SI

SINON [60bis]

Tenter successivement de récupérer ICAR de rue du PT [iP], de FUITES [iF], et de « Rue » [iCC], puis aller à l'étape ICAR

SI tous les ICAR sont vides **FIN DE**

**TRAITEMENT [END]**

FIN DE SI

SI seule la rue est disponible dans le fichier fuites [61]

SI la rue du poste technique est disponible [67]

Prendre en compte l'ICAR du poste technique

SINON [68]

Prendre la rue du fichier fuites ([APPR] si ICAR

par match approximatif)

FIN DE SI

Aller à l'étape ICAR

FIN DE SI

SI l'adresse du fichier fuites est vide [62]

Aller à l'étape « Rue » de SAP

FIN DE SI

FIN DE SI

SINON [33]

SI le champ « Rue » de SAP est complet [45]

SI la rue de « Rue » et la rue du poste technique sont cohérentes OU l'ICAR poste technique est vide [63]

Utiliser l'adresse du champ « Rue » de SAP (FULL) si correspondance totale avec ICAR, (DIG) si le numéro sans complément a dû être utilisé pour correspondance avec ICAR, (APPR) si l'ICAR est tiré d'un match partiel sur les noms de rue, (ADIG) si le numéro sans complément a dû être utilisé pour correspondance avec ICAR, et (NA) si pas de correspondance ICAR

SI pas de correspondance BD ICAR (63(NA)) aller au nœud 46 et l'évaluer FIN DE SI

Aller à l'étape Adresse

SINON [63bis]

Aller au nœud 46 et l'évaluer

FIN DE SI

FIN DE SI

SI la rue est cohérente entre les doublons de fuites ET au moins un doublon a un numéro d'immeuble [46]

Supprimer les doublons avec ICAR+numéro d'immeuble sans correspondance ICAR, sauf s'il n'y a aucun doublon avec correspondance ICAR

Prendre aléatoirement un numéro dans la rue parmi ceux disponibles

(FULL) si correspondance totale avec ICAR, (DIG) si le numéro sans complément a dû être utilisé pour correspondance avec ICAR, (APPR) si l'ICAR est tiré d'un match partiel sur les noms de rue, (ADIG) si le numéro sans complément a dû être utilisé pour correspondance avec ICAR, et (NA) si pas de correspondance ICAR

SI pas de correspondance BD ICAR (46(NA)) aller au nœud 47 sans l'évaluer FIN DE SI

Aller à l'étape Adresse

SINON [47]

SI l'adresse du poste technique possède au moins la rue [64]  
**Utiliser la rue du poste technique**  
 Aller à l'étape ICAR

SINON [65]  
 Aller à l'étape « Rue » de SAP

FIN DE SI

FIN DE SI

FIN DE SI

SINON [23]

Fin Pseudo code

## Détails des traitements

### Informations liées aux adresses

#### Codification des communes

L'extrait SAP ne comporte pas de codes INS (code unique associé à chacune des 262 communes Wallonnes), mais des codes postaux. Il est préférable de travailler avec les codes INS que les codes postaux car ces premiers sont moins nombreux et plus pérennes, et car les bases de données du SPW utilisent plutôt le code INS comme clé que les codes postaux. Les codes INS sont donc récupérés pour chaque OT grâce au nom de commune présent dans SAP. Ce dernier a dû en premier lieu être corrigé car il mélangeait parfois les noms de communes et les noms de « zones d'adresse ». Cette correction a été possible grâce au code postal présent dans SAP qui a permis de retrouver le nom de commune lorsque celui-ci a été remplacé par un nom de zone d'adresse (ceci est valide car un code postal n'a qu'une seule commune). Une fois le nom de commune corrigé, ce dernier a pu être utilisé pour identifier le code INS par simple correspondance avec la BD de communes du SPW. Un traitement similaire a été fait pour les fichiers fuites, mais la zone d'adresse a dû être utilisée au lieu du code postal (non présent dans les fichiers fuites). Au préalable, il a été impératif d'identifier les zones d'adresse dont le nom est partagé par plusieurs communes, et de les éliminer du traitement (contrairement aux codes postaux, certaines zones d'adresses apparaissent dans plusieurs communes).

#### Numéros d'immeuble

Les numéros d'immeuble ont été extraits des champs contenant ces derniers (comme le champ « Rue », ou le numéro d'immeuble lié au poste technique dans SAP) à partir de la reconnaissance automatique de toute expression régulière du type « Numéro(s) » ou « Numéro(s)Lettre(s) » ou « Numéro(s) Lettre(s) » en fin de chaîne de caractères, à laquelle on a par la suite enlevé les espaces inutiles. Ainsi, le numéro formaté est du type NuméroComplémentÉventuelEnLettres sans espaces, en majuscules, comme « 58TER ».

#### Noms de rue

Les noms de rue ont été extraits des champs contenant ces derniers (comme le champ « Rue » dans SAP, ou la rue renseignée dans les fichiers fuites) en plusieurs étapes. Seul le code R permet de décrire de façon correcte et exhaustive l'ensemble des traitements. De façon simplifiée, un grand nombre de mots et/ou symboles n'appartenant vraisemblablement pas à un nom de rue (comme

« bâtiment », « à gauche », « en face de », « près de », >, <, ...) sont supprimés, puis certaines abréviations sont remplacées, et enfin les espaces inutiles sont supprimés.

#### *ICAR de rue*

Les ICAR de rue associés aux noms de rues extraits comme décrit au paragraphe précédent sont récupérés par correspondance entre les couples (code INS, nom de rue) dans l'extrait SAP et ces mêmes couples dans la base de données (BD) ICAR du SPW (fichier texte qui associe à chaque couple un numéro ICAR unique). Afin de ne pas être impacté par les légères mais relativement fréquentes différences qu'il peut exister entre les noms de rue dans les fichiers SWDE et dans la BD ICAR, une correspondance par ressemblance approximative de texte (*fuzzy string matching*) est également effectuée. Cette dernière consiste à associer les deux noms de rue les plus proches (toujours sous contrainte du même code INS) tout en imposant une distance maximale (exprimée grosso modo en nombre de caractères différents). Pour la recherche de correspondance exacte ou approximative, les noms de rue comparés sont toujours mis en majuscule et ne comportent pas d'accents, afin de limiter les différentes orthographes d'un même nom de rue.

#### *Points d'adresse*

Les points d'adresse des défaillances sont ceux de la BD ICAR. Ils sont identifiés pour chaque OT SAP par comparaison entre une adresse de défaillance du type (ICAR de rue, numéro d'immeuble) et ce même couple dans la base de points d'adresse (chaque tel couple est unique). Il peut arriver que l'adresse associée à la défaillance soit erronée, par exemple lorsqu'elle indique le numéro 4C dans la rue X, alors que seul le numéro 4 de la rue X existe dans la BD ICAR. Ainsi, lors de l'absence de correspondance immédiate comme décrit ci-dessus, on cherche à retrouver la correspondance avec la même adresse sans complément de numéro d'immeuble. Deux cas sont alors à distinguer :

- Lorsqu'il s'agit d'une adresse issue du poste technique, il n'y a pas d'autre tentative de récupérer le point d'adresse si la méthode ci-dessus échoue.
- Lorsqu'il s'agit d'une adresse issue du champ « Rue » de SAP ou des fichiers fuites, on peut encore tenter de faire la correspondance grâce à l'ICAR de rue extrait de la correspondance approximative entre les noms de rues de l'OT et de la BD rues du SPW. Enfin, en dernier recours, on recommence le test précédent mais avec un numéro d'immeuble sans complément.

### Informations liées aux matériaux et diamètres

Lors du chargement du fichier de fuites concaténé, le matériau « FONTE » a été remplacé par « FONTE GRISE » car le matériau « FONTE » seul n'existe pas dans le SIG. C'est une étape nécessaire pour faciliter la comparaison de matériaux entre fichiers fuites et conduites du SIG afin de sélectionner les conduites du SIG compatibles en termes de matériau. Notons que ce choix résulte de la nette prédominance de la fonte grise dans les réparations sur fonte lorsque le type de fonte est connu.

Dans le même ordre d'idées, lors de comparaisons de matériaux entre conduites abandonnées et conduites en service (voir Annexe 5) et pour l'affectation des défaillances aux conduites, certains noms de matériaux du SIG ont été traités comme équivalents afin ne pas traiter des matériaux

similaires comme étant incompatibles et de ne pas limiter les correspondances entre conduites et conduites ou entre conduites et défaillances (Tableau 17).

Tableau 17 : Désignations de matériaux considérées comme équivalentes.

Nom final	Noms initiaux
ACIER	ACIER ANCIENNE NORME
FONTE DUCTILE	FONTE DUCTIL
AMIANTE CIMENT	AMIANTE CIMENT EB15 AMIANTE CIMANT EB15 AMIANTE CIMENT ISO
PVC	P.V.C.
PE	POLYETHYLENE POLYETHYLENE EXPANSE HAUTE DENSITE

Cette liste devra être mise à jour lorsque l'ensemble du SIG de la SWDE sera disponible.

Concernant les pièces de réparation, il est clair qu'un OT donné peut correspondre à plusieurs pièces de réparation avec des DN pré-identifiés (voir Annexe 3). On considère donc qu'une conduite du SIG est compatible avec les pièces de réparation si son DN est compris entre le plus petit DN et le plus grand DN de l'ensemble des pièces de réparation associées à l'OT.

### Résultats de l'attribution des réparations sur conduite

La répartition des attributions des défaillances sur conduite (type d'activité 105) aux tronçons en fonction des méthodes est présentée dans le Tableau 18.

Tableau 18 : Méthodes d'affectation des défaillances sur conduite pour le secteur de Mons.

Méthode	Nombre d'OT	% d'OT
Adresse liée au poste technique	673	51 %
ICAR de rue du poste technique	239	18 %
Point d'anomalie dans le SIG	119	9 %
Adresse dans le champ « Rue »	105	8 %
Adresse du fichier fuites	71	5 %
Autre méthode	93	7 %
Non attribué	21	1.6 %
<b>Total</b>	<b>1321</b>	<b>100 %</b>

## Annexe 9 : Caractérisation des cellules - bâti

L'objectif est de caractériser les cellules par le nombre de bâtiments, leur surface au sol, la hauteur moyenne.

Ces informations sont recueillies dans 3 couches SIG :

1. *Bâtiments 3D* (PICC - année de création 2013-2014). Cette couche permet de récupérer la hauteur du bâti, en l'occurrence la hauteur des murs, qui correspond à la différence obtenue entre la côte MNS (modèle numérique de surface 2013 - 2014) et la côte MNT (modèle numérique de terrain)
2. *Emprise des bâtiments et destination (habitation, annexe, administration, ...)* (PICC). Cette couche est maintenue à jour contrairement à la précédente. Les identifiants étant identiques, l'information hauteur est récupérée pour une part du bâti. Cependant, la destination des bâtis n'est pas toujours correctement attribuée : Nombres d'annexes sont considérées comme des habitations et des commerces sont identifiés comme des habitations.
3. *Points d'adresse* (ICAR). Le fichier adresse permet de localiser les locaux à destination d'habitation et d'écartier ainsi les locaux annexes sélectionnés dans la couche précédente. L'information ICAR de rue est présente dans ce fichier. Plusieurs points d'adresse sont superposés s'il y a plusieurs logements dans le bâtiment, ouvrant la possibilité d'estimer le nombre d'abonnés. Si le nombre N de points d'adresse superposés est supérieur à 1, le nombre de logement à ce point sera égal à N-1. Lorsque sur un même bâtiment il y a plusieurs points d'adresse distincts, la surface bâtie est répartie au prorata du nombre de points.

Après traitement de ces 3 couches SIG, pour chaque point d'adresse les informations suivantes sont connues : la surface du bâti, sa nature, la hauteur des murs (si elle existe), le nombre de logements, l'adresse avec l'ICAR de rue.

Il reste à déterminer la cellule d'appartenance du point d'adresse qui porte maintenant les caractéristiques des bâtiments. Cette information est disponible dans la couche conduite. La conduite la plus proche est sélectionnée en éliminant au préalable les conduites de production, d'eau brute, de liaison, de décharge et d'adduction qui ne sont pas utilisées dans cette étape de traitement :

- Si la conduite est dans la même rue que le point d'adresse, sa distance, son ICAR de rue, sa cellule et son identifiant sont liés au point d'adresse.
- Si elle n'est pas dans la même rue et que la distance au point d'adresse est inférieure ou égale à 50 mètres, alors sa distance, son ICAR rue, sa cellule et son identifiant sont retenus, sinon, sa cellule n'est pas retenue.

Par regroupement, le nombre de bâtiments, nombre d'abonnés, surface de bâti, hauteur pondérée du bâti, et le linéaire de canalisations de distribution sont estimés pour chaque cellule. L'information du nombre de raccordements et du nombre de logements est également affectée à chaque tronçon de canalisation.

### Cas particulier des conduites abandonnées :

Pour les conduites HS, il n'est pas possible de ne sélectionner que les conduites de distribution. Aucune distinction n'est donc faite à priori. La démarche est la même que pour les conduites en service, avec une contrainte supplémentaire : la distance maximum de rattachement d'une adresse à une canalisation abandonnée est de 20 mètres, même s'il y a correspondance des ICAR de rue entre la canalisation et l'adresse. Cette précaution est nécessaire car la conduite peut être abandonnée sur une petite partie de la rue et on ne peut alors lui attribuer l'ensemble des branchements de la rue.

A titre de comparaison, le Tableau 19 juxtapose le nombre de branchements résultant du traitement par cellule sur MONS avec les données de raccordements transmises par la SWDE. Les écarts sont faibles (1,8 % en moyenne) avec vraisemblablement un retard dû au retard de mise à jour de la base adresse dans les secteurs en forte expansion où, comme le montre la Figure 27, environ 20 points d'adresse sont absents sur une zone de 8 ha et 66 logements.



Figure 27 : Exemple de projection des couches géographiques relatives au bâti (Cellule 52300041).

Tableau 19 : Nombre de branchements obtenu par points d'adresse et données SWDE.

NUM_CEL	Croisement bâti - conduites MONS 22/02/2021			Rapport cellule MONS SWDE 15/11/2020		Ecart en %	
	Nombre branchements	Nombre logements	Longueur distribution	Nombre Raccordements	Longueur Canalisations	Branchements	Longueur
52300001	387	1 317	7,61	380	8,36	1,8	9,0
52300002	841	2 164	9,48	826	9,48	1,8	0,0
52300003	1 142	2 741	11,29	1 109	11,28	3,0	0,1
52300004	1 287	2 943	17,27	1 249	17,27	3,0	0,0
52300005	472	584	9,37	478	9,58	1,3	2,2
52300006	1 407	1 771	23,2	1 429	23,19	1,5	0,0
52300007	1 333	1 465	32,33	1 346	31,79	1,0	1,7
52300008	840	1 332	23,38	849	22,61	1,1	3,4
52300009	1 036	1 720	25,04	1 089	25,03	4,9	0,0
52300010	661	738	21,37	668	21,37	1,0	0,0
52300011	840	998	10,5	821	10,5	2,3	0,0
52300012	499	993	7,12	486	7,12	2,7	0,0
52300013	217	229	7,35	212	7,35	2,4	0,0
52300015	524	572	13,26	529	13,26	0,9	0,0
52300016	216	224	7,61	222	7,61	2,7	0,0
52300017	445	457	11,22	423	11,21	5,2	0,1
52300018	153	161	3,72	155	3,72	1,3	0,0
52300019	701	802	12,1	712	12	1,5	0,8
52300020	937	1 034	21,95	931	21,93	0,6	0,1
52300021	852	877	16,45	822	16,45	3,6	0,0
52300022	428	440	11,22	430	11,21	0,5	0,1
52300023	337	382	8,88	337	8,75	0,0	1,5
52300024	210	237	7,63	206	7,62	1,9	0,1
52300025	1 457	1 696	16,81	1 406	16,54	3,6	1,6
52300026	426	486	3,57	429	3,57	0,7	0,0
52300029	2 377	2 760	22,75	2 392	22,74	0,6	0,0
52300030	1 609	1 707	22,58	1 617	21,82	0,5	3,5
52300031	355	376	4,1	355	4,1	0,0	0,0
52300032	596	646	11,14	603	11,14	1,2	0,0
52300033	99	286	1,81	100	1,81	1,0	0,0
52300035	67	75	6,16	70	6,16	4,3	0,0
52300036	300	638	21,58	290	21,58	3,4	0,0
52300037	136	138	3,79	136	3,79	0,0	0,0
52300039	439	453	11,96	455	11,93	3,5	0,3
52300041	596	615	13,16	643	13,16	7,3	0,0
52300042	1 293	1 994	17,47	1 299	17,46	0,5	0,1
52300043	330	353	8,71	345	8,7	4,3	0,1
52300045	1 080	1 159	18,88	1 080	18,88	0,0	0,0
52300046	862	952	21,83	874	21,83	1,4	0,0
52300047	598	788	6,6	603	6,6	0,8	0,0
52300048	1 111	1 185	15,76	1 110	15,76	0,1	0,0
52300049	801	837	19,41	797	19,41	0,5	0,0
52300050	253	275	3,54	246	3,53	2,8	0,3
52300052	282	393	6,13	286	6,13	1,4	0,0
52300053	362	382	12,51	359	12,51	0,8	0,0
52300054	146	148	4,67	150	4,67	2,7	0,0
52300055	850	913	11,87	856	11,86	0,7	0,1
52300056	1 437	1 745	17,68	1 434	17,68	0,2	0,0
52300057	1 106	1 130	11,11	1 125	11,11	1,7	0,0
52300058	1 249	1 500	9,3	1 249	9,3	0,0	0,0
52309022	274	274	6,56	276	6,56	0,7	0,0
Total	36 256	48 085	650,79	36 294	649,02	0,1	0,0

## Annexe 10 Dictionnaire simplifié des attributs des fichiers de données

Fichier	Type donnée	Attribut	Type Attribut	Unité	Description
Conduites En Service	Brute	ID	caractère		Identifiant de conduite
Conduites En Service	Brute	NATURE	caractère		Matériau
Conduites En Service	Brute	CLASSE	caractère		Classe d'utilisation (production, distribution...)
Conduites En Service	Brute	ANNEE	date		Année de pose
Conduites En Service	Brute	TYPE	numérique	mm	Diamètre Nominal
Conduites En Service	Brute	NUM_CEL	numérique		Numéro de cellule
Conduites Hors Service	Brute	ID	caractère		Identifiant de conduite
Conduites Hors Service	Brute	NATURE	caractère		Matériau
Conduites Hors Service	Brute	TYPE	numérique	mm	Diamètre Nominal
Conduites Hors Service	Brute	ABANDON	date		Année d'abandon
Conduites Hors Service	Brute	ANNEE	date		Année de pose
Anomalies SIG	Brute	NUMEXT	numérique		Numéro d'Ordre de Travail
Anomalies SIG	Brute	SUPPORT	caractère		Type et ID d'objet SIG concerné par l'anomalie
Anomalies SIG	Brute	DATEREPA	date		Date de réparation
Anomalies SIG	Brute	DATECONS	date		Date de constatation de l'anomalie
Anomalies SIG	Brute	ETAT	caractère		État de l'anomalie (détectée ou réparée)
Axes de voirie	Brute	GEOREF_ID	caractère		Identifiant unique d'objet géoréférencé
Axes de voirie	Brute	NATUR_CODE	caractère		Code du type de voirie (route, chemin, piste...)
Axes de voirie	Brute	ICARRUEID1	numérique		ICAR de rue de l'axe de voirie
Points d'adresse ICAR	Brute	ADR_ID	numérique		Identifiant ICAR unique du point d'adresse
Points d'adresse ICAR	Brute	RUE_ID	numérique		ICAR de rue du point d'adresse

Fichier	Type donnée	Attribut	Type Attribut	Unité	Description
Points d'adresse ICAR	Brute	ADR_NUMERO	caractère		Numéro d'immeuble du point d'adresse
Emprise de bâti du PICC	Brute	GEOREF_ID	caractère		Identifiant unique d'objet géoréférencé
Emprise de bâti du PICC	Brute	NATUR_CODE	caractère		Code du type de bâtiment (commercial, hab...)
Bâti 3D	Brute	REF_CBE	caractère		Identifiant unique d'objet géoréférencé (identique à GEOREF_ID)
Bâti 3D	Brute	NATUR_CODE	caractère		Code du type de bâtiment (commercial, hab...)
Bâti 3D	Brute	NATUR_DESC	caractère		Description du type de bâtiment
Bâti 3D	Brute	H_MUR	numérique	m	Hauteur de mur du bâtiment
Corine Land Cover	Brute	CODE_18	numérique		Code Corine Land Cover 2018
Modèle Numérique de Terrain	Brute	Bande 1	numérique	m	Cote altimétrique du terrain naturel
Carte Numérique des Sols	Brute	CODE	numérique		Code du type de sol (pédologie)
Carte d'utilisation du sol (WALOUS)	Brute	WALOUSMAJ	caractère		Code du type d'utilisation du sol
Fichiers fuites	Brute	OT	numérique		OT de réparation, N° D'OT
Fichiers fuites	Brute	OT RF	caractère		OT de recherche de fuite (OT RF)
Fichiers fuites	Brute	TYPE CONDUITE	caractère		Matériau (et diamètre pour certains fichiers)
Fichiers fuites	Brute	DIAMERTE	numérique	mm	Diamètre Nominal
Fichiers fuites	Brute	CELLULES	numérique		N° DE CELLULE (secteur hydraulique)
Fichiers fuites	Brute	COMMUNE	caractère		Commune
Fichiers fuites	Brute	SECTION	caractère		Zone d'adresses (souvent contenues dans une commune)
Fichiers fuites	Brute	RUE	caractère		Nom de rue
Fichiers fuites	Brute	N°	caractère		Numéro d'immeuble
Fichiers fuites	Brute	LOCALISE LE	date		Date de localisation
Fichiers fuites	Brute	REPARE LE	date		Date de réparation
Communes de Wallonie	Brute	Code INS	numérique		Code unique associé à chaque commune
Communes de Wallonie	Brute	Code Postal	numérique		Code d'une localité au sein d'une commune

Fichier	Type donnée	Attribut	Type Attribut	Unité	Description
Communes de Wallonie	Brute	Zone d'adresses	caractère		Zone d'adresses (souvent contenues dans une commune)
Communes de Wallonie	Brute	Commune	caractère		Nom de commune
Rues de Wallonie	Brute	Commune	caractère		Nom de commune
Rues de Wallonie	Brute	Rue	caractère		Nom de rue
Rues de Wallonie	Brute	Identifiant	numérique		ICAR de rue
Pièces de réparation	Brute	Ordre	numérique		Numéro d'OT de réparation
Pièces de réparation	Brute	Article	numérique		Numéro de pièce (SAP)
Pièces de réparation	Brute	Désignation	caractère		Description de la pièce
Dictionnaire de pièces de réparation	Brute	Objet	numérique		Numéro de pièce (SAP)
Dictionnaire de pièces de réparation	Brute	Désignation article	caractère		Description de la pièce
Dictionnaire de pièces de réparation	Brute	Diamètre (DN ou DE)	numérique	mm	Diamètre de la pièce ou de la conduite compatible
Extrait SAP	Brute	Date début plf	date		Date de début planifié de l'intervention
Extrait SAP	Brute	Ordre	numérique		Numéro d'OT
Extrait SAP	Brute	Type d'ordre	caractère		Type d'OT
Extrait SAP	Brute	Type de travail	numérique		Type d'activité lié à l'OT
Extrait SAP	Brute	ID ICAR	numérique		ICAR de rue
Extrait SAP	Brute	Désignation Poste technique	caractère		Adresse liée au poste technique
Extrait SAP	Brute	Code postal	numérique		Code d'une localité au sein d'une commune
Extrait SAP	Brute	Ville	caractère		Commune
Extrait SAP	Brute	Rue	caractère		Nom de rue
Extrait SAP	Brute	HOUSE NUM	caractère		Numéro d'immeuble
Extrait SAP	Brute	Groupe de codes	caractère		Code de type panne

Fichier	Type donnée	Attribut	Type Attribut	Unité	Description
Extrait SAP	Brute	Code cause	caractère		Code de cause panne au sein d'un type de panne
Borne Incendie (SIG)	Brute	PRESSION	numérique		Pression mesurée sur l'hydrant
Bouche Incendie (SIG)	Brute	PRESSION	numérique		Pression mesurée sur l'hydrant
Réducteur de pression (SIG)	Brute	ID	caractère		Identifiant d'objet SIG
Réducteur de pression (SIG)	Brute	P_AMONT	numérique		Pression amont de la régulation
Réducteur de pression (SIG)	Brute	P_AVAL	numérique		Pression aval de la régulation
Réducteur de pression (SIG)	Brute	CEL1	caractère		Cellule amont
Réducteur de pression (SIG)	Brute	CEL2	caractère		Cellule aval
Conduites traitées (en service ou HS)	Traitée	New_ID	caractère		Identifiant de conduite après agrégation
Conduites traitées (en service ou HS)	Traitée	L_final	numérique	m	Longueur de conduite après agrégation
Conduites traitées (en service ou HS)	Traitée	GEOREF_ID	caractère		Identifiant GEOREF de l'axe de voirie qui a donné son ICAR
Conduites traitées (en service ou HS)	Traitée	NATUR_CODE	caractère		Code du type de voirie (chemin, route...) qui a donné son ICAR
Conduites traitées (en service ou HS)	Traitée	ICAR_RUE	numérique		ICAR de rue attribué à la conduite
Conduites traitées (en service ou HS)	Traitée	larg_tamp	numérique	m	Largeur du tampon autour de l'axe de voirie utilisée pour attribuer l'ICAR de rue à la conduite
Conduites traitées (en service ou HS)	Traitée	Nb_racct	numérique		Nombre de raccordements sur cette conduite
Conduites traitées (en service ou HS)	Traitée	Nb_logt	numérique		Nombre de logements sur cette conduite
Conduites traitées (en service ou HS)	Traitée	CLC1-CLC5	numérique		Pourcentage de longueur de conduite dans la classe Corine Land Cover n (avec n allant de 1 à 5)
Conduites traitées (en service ou HS)	Traitée	MNT1	numérique	m	Cote topographique du nœud 1 de la conduite
Conduites traitées (en service ou HS)	Traitée	MNT2	numérique	m	Cote topographique du nœud 2 de la conduite

Fichier	Type donnée	Attribut	Type Attribut	Unité	Description
Conduites traitées (en service ou HS)	Traitée	MNT	numérique	m	Cote topographique moyenne de la conduite
Conduites traitées (en service ou HS)	Traitée	HMT	numérique	m	HMT moyenne dans la conduite déterminée par méthode topographique
Conduites traitées (en service ou HS)	Traitée	minMNTc	numérique	m	Cote topographique minimum de la cellule de cette conduite
Conduites traitées (en service ou HS)	Traitée	maxMNTc	numérique	m	Cote topographique maximum de la cellule de cette conduite
Conduites traitées (en service ou HS)	Traitée	pression	numérique	m	Pression moyenne dans la conduite déterminée par méthode topographique
Conduites traitées (en service ou HS)	Traitée	Pargile	numérique		Pourcentage de longueur de conduite dans sol argileux
Conduites traitées (en service ou HS)	Traitée	Ptourbe	numérique		Pourcentage de longueur de conduite dans sol tourbeux
Conduites traitées (en service ou HS)	Traitée	Plimon	numérique		Pourcentage de longueur de conduite dans sol limoneux
Conduites traitées (en service ou HS)	Traitée	Proute	numérique		Pourcentage de longueur de conduite sous chaussée
Conduites traitées (en service ou HS)	Traitée	Paccote	numérique		Pourcentage de longueur de conduite sous accotement
Conduites HS traitées	Traitée	ID_old	caractère		Identifiant de conduite avant renommage moins ambigu (ajout de ABAN_ en préfixe, voir m.CHS.IDC)
Conduites HS traitées	Traitée	New_ID_2	caractère		Identifiant de la conduite en service utilisée pour la complétion de données
Conduites HS traitées	Traitée	NUM_CEL	numérique		Numéro de cellule rattaché grâce à la correspondance avec les conduites en service
Conduites HS traitées	Traitée	ANNEEold	date		Année de pose avant la complétion de données
Conduites En Service absorbées	Traitée	ident	caractère		Identifiant de la conduite en service qui absorbe cette conduite
Conduites HS absorbées	Traitée	ident	caractère		Identifiant de la conduite abandonnée qui absorbe cette conduite
Bilan par cellule	Traitée	Nb_racct	numérique		Nombre de raccordements dans cette cellule
Bilan par cellule	Traitée	Nb_racct_hab	numérique		Nombre de raccordements comptés avec bâti de type habitation ou building dans cette cellule
Bilan par cellule	Traitée	Nb_logt	numérique		Nombre de logements dans cette cellule
Bilan par cellule	Traitée	Nb_logt_hab	numérique		Nombre de logements comptés avec bâti de type habitation ou building dans cette cellule

Fichier	Type donnée	Attribut	Type Attribut	Unité	Description
Bilan par cellule	Traitée	Ssurf	numérique	m2	Somme de surface de bâti dans cette cellule
Bilan par cellule	Traitée	Ssurf_hab	numérique	m2	Somme de surface de bâti de type habitation ou building dans cette cellule
Bilan par cellule	Traitée	SsurfXh	numérique	m3	Somme de surface de bâti multipliée par la hauteur dans cette cellule
Bilan par cellule	Traitée	SsurfXh_hab	numérique	m3	Somme de surface de bâti de type habitation ou building multipliée par la hauteur dans cette cellule
Bilan par cellule	Traitée	H_pondere	numérique	m	Hauteur de bâti pondérée par la surface pour cette cellule
Bilan par cellule	Traitée	H_pondere_hab	numérique	m	Hauteur de bâti de type habitation ou building pondérée par la surface pour cette cellule
Bilan par cellule	Traitée	L_distri	numérique	m	Longueur totale de conduite dans la cellule, hors classe PRODUCTION, EAU BRUTE, LIAISON, DECHARGE, et ADDUCTION
Bilan par cellule	Traitée	LONG_tot	numérique	m	Longueur totale de conduite dans la cellule
Bilan par cellule	Traitée	LONG_CLC1- LONG_CLC5	numérique	m	Longueur totale de conduite dans classe Corine Land Cover n (avec n allant de 1 à 5) pour cette cellule
Bilan par cellule	Traitée	Part_CLC_1- Part_CLC_5	numérique		Pourcentage de longueur totale de conduite dans classe Corine Land Cover n (avec n allant de 1 à 5) pour cette cellule
Bilan par cellule	Traitée	Nb_pt_MNT	numérique		Nombre de sommets terminaux de conduites dans la cellule
Bilan par cellule	Traitée	Nb_pt_unique _MNT	numérique		Nombre de sommets terminaux uniques (non superposés) de conduites dans la cellule
Bilan par cellule	Traitée	min_MNT	numérique	m	Cote topographique minimum la cellule
Bilan par cellule	Traitée	max_MNT	numérique	m	Cote topographique maximum de la cellule
Bilan par cellule	Traitée	moy_MNT	numérique	m	Cote topographique moyenne de la cellule
Bilan par cellule	Traitée	med_MNT	numérique	m	Cote topographique médiane de la cellule
Bilan par cellule	Traitée	deniv_MNT	numérique	m	Dénivelé entre cote min et cote max de cette cellule
Bilan par cellule	Traitée	LONG_argile	numérique	m	Longueur totale de conduite dans sol argileux pour cette cellule
Bilan par cellule	Traitée	LONG_tourbe	numérique	m	Longueur totale de conduite dans sol tourbeux pour cette cellule
Bilan par cellule	Traitée	LONG_limon	numérique	m	Longueur totale de conduite dans sol limoneux pour cette cellule
Bilan par cellule	Traitée	Part_argile	numérique		Pourcentage de longueur totale de conduite dans sol argileux pour cette cellule

Fichier	Type donnée	Attribut	Type Attribut	Unité	Description
Bilan par cellule	Traitée	Part_tourbe	numérique		Pourcentage de longueur totale de conduite dans sol tourbeux pour cette cellule
Bilan par cellule	Traitée	Part_limon	numérique		Pourcentage de longueur totale de conduite dans sol limoneux pour cette cellule
Bilan par cellule	Traitée	LONG_route	numérique	m	Longueur totale de conduite sous chaussée pour cette cellule
Bilan par cellule	Traitée	LONG_accote	numérique	m	Longueur totale de conduite sous accotement pour cette cellule
Bilan par cellule	Traitée	Part_route	numérique		Pourcentage de longueur totale de conduite sous chaussée pour cette cellule
Bilan par cellule	Traitée	Part_accote	numérique		Pourcentage de longueur totale de conduite sous accotement pour cette cellule
Bilan par cellule	Traitée	Pmoy	numérique	m	Pression moyenne sur la cellule
Bilan par cellule	Traitée	L_Psup40	numérique	m	Pourcentage de longueur de conduite pour lesquelles la pression est supérieure à 40 mètres
Bilan par cellule	Traitée	P_q80	numérique	m	80 % de la longueur des tronçons ont une pression inférieure à cette valeur
Bilan par cellule	Traitée	Nb_hyd	numérique		Nombre d'hydrants dans la cellule
Bilan par cellule	Traitée	Nb_hyd_avec_P	numérique		Nombre d'hydrants de la cellule avec information sur la pression
Bilan par cellule	Traitée	Nb_reduc	numérique		Nombre de vannes de régulation dans la cellule
Bilan par cellule	Traitée	Nb_reduc_avec_P	numérique		Nombre de vannes de régulation de la cellule avec information sur la pression
Fichier de casses	Traitée	IDT	caractère		Identifiant de tronçon
Fichier de casses	Traitée	DDC	date		Date de casse
Fichier de casses	Traitée	CRASH	booléen		Crash sur conduite ?
Fichier de casses	Traitée	ACT	numérique		Numéro d'activité SWDE (exemple : 105 = réparation de conduite)
Fichier de casses	Traitée	OT	numérique		Ordre de travail (OT) SWDE
Fichier de casses	Traitée	TYPE_OT	caractère		Type d'OT SAP
Fichier de casses	Traitée	OT_RF	numérique		Ordre de travail recherche de fuite (RF) SWDE
Fichier de casses	Traitée	DateLoc	date		Date de localisation (fichiers fuites)
Fichier de casses	Traitée	DateSIG	date		Date de localisation (Anomalies SIG)
Fichier de casses	Traitée	PROB	caractère		Code de cause de panne SWDE (combinant type et cause)

Fichier	Type donnée	Attribut	Type Attribut	Unité	Description
Fichier de casses	Traitée	ETAT	caractère		État de la fuite, détectée ou réparée (info du SIG)
Fichier de casses	Traitée	Cond_num	numérique		Nombre ordinal de la proximité conduite--adresse de défaillance
Fichier de casses	Traitée	Dist_adr	numérique	m	Distance euclidienne entre la conduite affectée et l'adresse de la défaillance renseignée
Fichier de casses	Traitée	IncertCaracs	numérique		Indice d'incertitude (0 à 4) d'attribution de défaillance vis-à-vis des caractéristiques de conduite renseignées
Fichier de casses	Traitée	nCondICAR	numérique		Nombre de conduites portant le même ICAR que la conduite affectée par la défaillance
Fichier de casses	Traitée	fracL_ICAR	numérique		Ratio entre longueur de conduite affectée et longueur totale de conduites avec le même ICAR de rue
Fichier de casses	Traitée	nCondCaracs	numérique		Nombre de conduites portant le même ICAR et ayant les mêmes caractéristiques que la conduite affectée par la défaillance
Fichier de casses	Traitée	fracL_Caracs	numérique		Ratio entre longueur de conduite affectée et longueur totale de conduites avec le même ICAR de rue et les mêmes caractéristiques
Fichier de casses	Traitée	flag	caractère		Cheminement utilisé pour attribuer la défaillance à une conduite
Fichier de tronçons	Traitée	IDT	caractère		Identifiant de tronçon
Fichier de tronçons	Traitée	DDP	date		Date de pose
Fichier de tronçons	Traitée	LNG	numérique	m	Longueur
Fichier de tronçons	Traitée	MAT	caractère		Matériau
Fichier de tronçons	Traitée	DHS	date		Date de mise hors service
Fichier de tronçons	Traitée	DIA	numérique	mm	Diamètre Nominal
Fichier de tronçons	Traitée	CLASSE	caractère		Type de conduite (production, distribution...)
Fichier de tronçons	Traitée	STATUT	caractère		Statut de la conduite (en service / abandonnée)
Fichier de tronçons	Traitée	NUMCEL	numérique		Numéro de cellule
Fichier de tronçons	Traitée	ICAR	numérique		ICAR de rue
Fichier de tronçons	Traitée	L_TAMP	numérique	m	Largeur du tampon utilisée pour l'attribution d'ICAR
Fichier de tronçons	Traitée	NB_RACCT	numérique		Nombre de raccords
Fichier de tronçons	Traitée	NB_LOGT	numérique		Nombre de logements

Fichier	Type donnée	Attribut	Type Attribut	Unité	Description
Fichier de tronçons	Traitée	CLC1	numérique		Pourcentage de longueur de conduite dans classe CLC1
Fichier de tronçons	Traitée	CLC2	numérique		Pourcentage de longueur de conduite dans classe CLC2
Fichier de tronçons	Traitée	CLC3	numérique		Pourcentage de longueur de conduite dans classe CLC3
Fichier de tronçons	Traitée	CLC4	numérique		Pourcentage de longueur de conduite dans classe CLC4
Fichier de tronçons	Traitée	CLC5	numérique		Pourcentage de longueur de conduite dans classe CLC5
Fichier de tronçons	Traitée	Z1	numérique	m	Cote topographique du nœud 1 de la conduite
Fichier de tronçons	Traitée	Z2	numérique	m	Cote topographique du nœud 2 de la conduite
Fichier de tronçons	Traitée	Zmin	numérique	m	Cote topographique minimum de la cellule de la conduite
Fichier de tronçons	Traitée	Zmax	numérique	m	Cote topographique maximum de la cellule de la conduite
Fichier de tronçons	Traitée	Pargile	numérique		Pourcentage de longueur de conduite dans sol argileux
Fichier de tronçons	Traitée	Ptourbe	numérique		Pourcentage de longueur de conduite dans sol tourbeux
Fichier de tronçons	Traitée	Plimon	numérique		Pourcentage de longueur de conduite dans sol limoneux
Fichier de tronçons	Traitée	Proute	numérique		Pourcentage de longueur de conduite sous chaussée
Fichier de tronçons	Traitée	Paccote	numérique		Pourcentage de longueur de conduite sous accotement
Fichier de tronçons	Traitée	TYPE_VOI	caractère		Type de voirie (route, chemin...)
Fichiers fuites concaténés	Traitée	annee	date		Année concernée par le fichier de fuites
Fichiers fuites concaténés	Traitée	OT_reparatio n	numérique		OT de réparation, N° D'OT
Fichiers fuites concaténés	Traitée	OT_RF	caractère		OT de recherche de fuite
Fichiers fuites concaténés	Traitée	materiau	caractère		Matériau
Fichiers fuites concaténés	Traitée	diametre	numérique	mm	Diamètre Nominal
Fichiers fuites concaténés	Traitée	num_cell	numérique		N° DE CELLULE (secteur hydraulique)

Fichier	Type donnée	Attribut	Type Attribut	Unité	Description
Fichiers fuites concaténés	Traitée	commune	caractère		Commune
Fichiers fuites concaténés	Traitée	section	caractère		Zone d'adresses
Fichiers fuites concaténés	Traitée	rue	caractère		Nom de rue
Fichiers fuites concaténés	Traitée	num_boite	caractère		Numéro d'immeuble
Fichiers fuites concaténés	Traitée	date_localise	date		Date de localisation
Fichiers fuites concaténés	Traitée	date_repar	date		Date de réparation
Pièces de réparation traitées	Traitée	DE min	numérique	mm	Diamètre Extérieur minium de la conduite recevant la pièce de réparation
Pièces de réparation traitées	Traitée	DE max	numérique	mm	Diamètre Extérieur maximum de la conduite recevant la pièce de réparation
Pièces de réparation traitées	Traitée	DNmin	numérique	mm	Diamètre Nominal minimum de la conduite recevant la pièce de réparation
Pièces de réparation traitées	Traitée	DNmax	numérique	mm	Diamètre Nominal maximum de la conduite recevant la pièce de réparation
Corine Land Cover traité	Traitée	CLASSE_CLC	numérique		Code (1 à 5) du type d'occupation du sol créé à partir des codes de base de la classe CLC
Carte Numérique des Sols Wallons traitée	Traitée	argile	booléen		Indicateur de la présence de sol argileux (vaut 1 si le type de sol correspond).
Carte Numérique des Sols Wallons traitée	Traitée	tourbe	booléen		Indicateur de la présence de sol tourbeux (vaut 1 si le type de sol correspond).
Carte Numérique des Sols Wallons traitée	Traitée	limon	booléen		Indicateur de la présence de sol limoneux (vaut 1 si le type de sol correspond).
Carte d'utilisation du sol (WALOUS) traitée	Traitée	route	booléen		Indicateur de la localisation de la conduite sous chaussée (vaut 1 si l'utilisation du sol correspond)
Carte d'utilisation du sol (WALOUS) traitée	Traitée	accote	booléen		Indicateur de la localisation de la conduite sous accotement (vaut 1 si l'utilisation du sol correspond)
Sommets de conduites traités	Traitée	ERRUISSOLM	numérique	m	Cote altimétrique du sommet de conduite interpolée grâce au MNT qui a été rasterisé au préalable



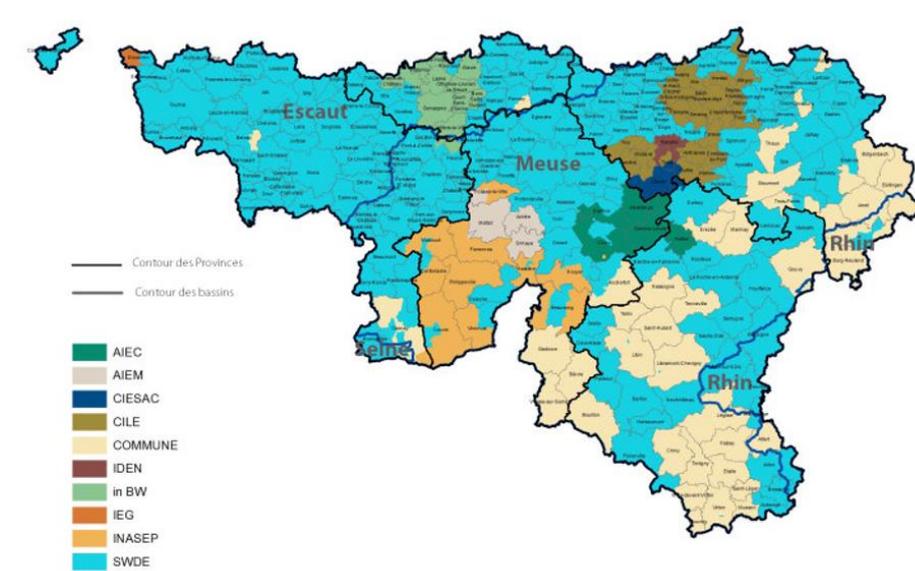
## Annexe 11 Note de cadrage méthodologique de la tâche 02

### 1. Contexte de départ

#### 1.1. La SWDE en quelques chiffres

La SWDE constitue le principal opérateur de service d'eau potable de la Wallonie. Elle assure la production et la distribution d'eau potable pour près de 2,5 millions d'habitants (soit 1,1 millions de clients abonnés), répartis sur 190 communes<sup>1</sup> (cf. Carte 1). Son réseau technique représente plus de 38 000 km de canalisations (dont 25 000 km dédiés à la distribution), et s'étend sur un territoire d'environ 13 000 km<sup>2</sup> (SWDE, 2019).

Carte 1 : Les sociétés de distribution d'eau potable en Wallonie



Source : AquaWal, 2018

En 2019, la SWDE a distribué plus de 104 millions m<sup>3</sup> d'eau potable, pour un coût de 2,62 €/m<sup>3</sup> HTVA. Ce prix est stable depuis 2014, conformément à l'objectif de stabilité dont s'est doté la SWDE. En 2020, la Société a investi 116 millions € dans son réseau. Le taux de réhabilitation des canalisations connaît toutefois une baisse constante depuis 2014, atteignant le taux de 0,44% en 2019. Une légère amélioration est intervenue en 2020, suite à la mise en place d'une nouvelle organisation décentralisée donnant pouvoir aux équipes locales d'engager les études et les travaux de réalisation, et en mutualisant les contrats d'achats. Les objectifs de performance du réseau ne sont pas non plus atteints, puisque l'Indice Linéaire de Volumes Non Enregistrés (ILVNE = VLAR – Volumes facturés) s'élève en 2020 à 5,42 m<sup>3</sup>/km/j, contre le seuil souhaité de 4,4 m<sup>3</sup>/km/j d'ici 2022. L'Indice Linéaire de Perte est par ailleurs le plus élevé depuis 2012, malgré une stabilité apparente des fuites sur le réseau et le maintien des taux de réparation. Depuis 2008, la SWDE participe également au benchmark européen sur la performance.

<sup>1</sup> Le périmètre géographique d'intervention continue régulièrement de s'étendre, grâce à l'adhésion de nouvelles communes.

Enfin, la SWDE s'est engagée dans le déploiement progressif de compteurs communicants (notamment en direction des établissements communaux), en vue de proposer une offre intitulée « *EasyConso* » permettant d'effectuer de la télé-relève et d'alerter les clients quant à d'éventuelles fuites ou consommations anormales.

## 1.2. Éléments historiques, institutionnels et organisationnels

La Société a été créée dans les années 1980, suite à la restructuration de la Société Nationale Des Eaux intervenue lors de la décennie précédente dans le but de doter les régions Nord et Sud du pays d'organes opérationnels. Avec la délégation de compétences aux régions, la Société Nationale a été scindée en deux entités (SWDE pour la Wallonie, VMW pour la Flandre), avant d'être dissoute (SWDE, 2014).

La SWDE est une entreprise publique autonome sous statut de société coopérative. Elle compte dans son capital social : 207 communes, 12 intercommunalités, 5 provinces, la Région Wallonne et 3 sociétés publiques (2 financières, 1 de gestion de l'eau). Elle est administrée de manière bicéphale par :

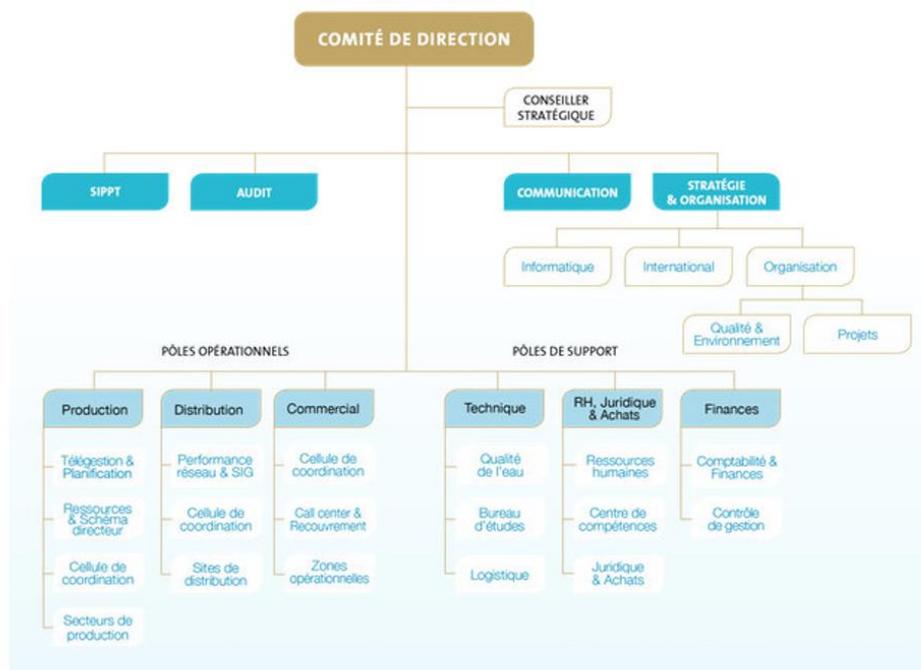
- un conseil d'administration, qui a pour mission de définir la politique et la stratégie générale de l'entreprise (valeurs ; stratégie à long terme, moyens alloués, contrôle de la performance, etc.). Il détient aussi un pouvoir de supervision de la gestion opérationnelle. Il se structure autour d'un comité d'audit, d'un comité de rémunération et d'un comité stratégique. Ce dernier appuie le conseil d'administration dans la définition de la politique générale SWDE (dont la stratégie à long terme) ;
- un comité de direction, qui assure la gestion opérationnelle au quotidien.

La SWDE s'organise autour de sept grandes activités :

- la protection des ressources
- la production d'eau potable
- la distribution d'eau potable
- la relation aux usagers (relation commerciale)
- le contrôle de la qualité (laboratoire interne)
- la formation-conseil
- la conception-développement (R&D)

L'organigramme ci-après présente l'organisation et la structuration des métiers au sein de la Société :

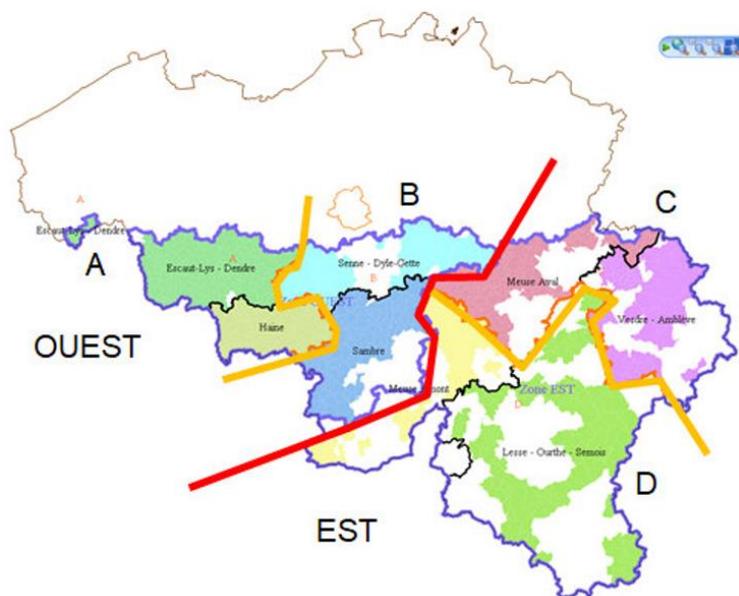
Figure 1 : Organigramme de la SWDE



Source : <https://www.swde.be>

Au plan opérationnel, la SWDE a récemment fait évoluer son organisation par zone géographique Ouest/Est, en la complétant d'une logique « fonctionnelle » fondée sur une approche hydraulique du territoire (cf. Carte 2). Chaque pôle opérationnel (production, distribution, commercial) a deux directeurs (Ouest, Est), et sur chacun des quatre sites fonctionnels interviennent : un responsable de Recherche de Fuite (RF) ; un ingénieur d'exploitation en charge de la production ; plusieurs ingénieurs dédiés à la distribution ; un bureau d'étude décentralisé ; un centre administratif. Au total, la Société compte plus de 1 300 agents ; 90% d'entre eux sont employés (près de 50%) ou ouvriers (environ 40%).

Carte 2 : Organisation géographique de la SWDE en 2019-2020



Source : SWDE, 2020, Organisation géographique

Dans le cadre de son projet quinquennal 2022-2027, la SWDE a défini trois objectifs stratégiques :

- Réduire les coûts d'exploitation
- Assurer la pérennité des infrastructures
- Maîtriser du prix de service (le « coûts-vérité distribution » doit rester inchangé)

Parmi les avantages que la SWDE identifie en son sein, on peut noter : la conception et gestion des infrastructures techniques, la gestion du cycle de l'eau potable (dont les ressources), la relation clientèle et la « stratégie informatique intégrée ». Les opportunités devant aider à atteindre les objectifs ont notamment trait à la dimension gouvernancière de la GPI (schéma régional des ressources en eau, « gouvernance du secteur », coopération avec d'autres opérateurs). Les faiblesses identifiées pour y parvenir relèvent quant à elles en premier lieu de l'état des infrastructures. La SWDE considère également la complexité des normes et la « surréglementation » comme étant sources de menaces pour ses missions et l'atteinte des objectifs (cf. Figure 2).

Figure 2 : Matrice SWOT de la SWDE



Source : SWDE, Notre stratégie 2022-2027

## 2. Problématisation, hypothèses, angles de travail

L'objectif de la tâche 2 consiste à analyser les processus de décision et de mise en œuvre de la GPI (court et long termes) de la SWDE. Elle vise à identifier et à expliciter les déterminants de la gestion patrimoniale de la Société, et leurs logiques (socio-économiques, techniques, professionnelles...). Cela soulève quatre questions centrales :

- Quels sont les processus et les logiques d'élaboration de la GPI à l'échelle de la SWDE ?
- Sur quels dispositifs, outils et critères s'appuient ces processus ?
- Comment ces politiques sont-elles mises en œuvre au cours du temps ?
- Quels effets produisent-elles aux échelles de la SWDE et de la région Wallonne ?

Pour y répondre, nous formulons deux hypothèses :

- La GPI déborde l'unique cadre de la SWDE. Elle s'inscrit dans un système de gouvernance élargi à la région Wallonne, eu égard à la participation d'acteurs politico-administratifs régionaux dans la gestion du service, ainsi qu'aux relations de coopération et d'interdépendance entre acteurs de l'eau.
- La GPI répond également à une organisation (culture, organisation des métiers, règles et procédures...) et à une instrumentation du travail (objectifs, outils...) spécifiques, qui orientent la compréhension des acteurs / agents, et structurent la prise en charge des enjeux infrastructurels.

## 2.1. La SWDE au sein de la gouvernance territoriale : contraintes et opportunités du système politico-administratif Wallon dans la GPI

Entendue comme un « processus de coordination d'acteurs, de groupes sociaux, d'institutions pour atteindre des buts discutés et définis collectivement » (Le Galès, 2006, p5), la gouvernance constitue un phénomène d'organisation socio-politique établi à différentes échelles. En ce qui concerne la SWDE, il est intéressant d'interroger ce processus et le cadre qu'il confère à la Société directement à l'échelle de la Wallonie, au moins pour 3 raisons :

- Le statut de coopérative de la SWDE place cette dernière en position de redevable à l'endroit de ses sociétaires, lesquels correspondent à des institutions politico-administratives (région, provinces, intercommunalités, communes) et à des entreprises publiques. Leurs représentants siègent au sein du conseil d'administration, ce qui leur offre collectivement un pouvoir d'orientation et de décision sur les stratégies d'action et les opérations de la SWDE. L'analyse du processus de décision et d'organisation en matière de gestion patrimoniale doit donc tenir compte du positionnement des représentants des sociétaires et de leur influence sur les orientations de la SWDE, et notamment des représentants de la Région Wallonne, laquelle considère la SWDE « comme son bras opérationnel » (Gouvernement Wallon, 2018).
- L'action de la SWDE s'inscrit également dans des dispositifs de coordination élargis ayant trait à tout ou partie des missions de la Société. Il s'agit en particulier du schéma régional de gestion des ressources et du décret PowalCo, qui organisent les relations entre acteurs parties prenantes de la gestion de l'eau (potable) et d'autres secteurs (notamment la voirie en ce qui concerne le décret PowalCo). De par leurs principes, leurs règles et leurs lignes directrices, ces dispositifs constituent des cadres structurants pour l'opérateur SWDE. En effet, celui-ci est contraint d'adapter sa stratégie de GPI (à court comme à long terme) aux objectifs, aux préconisations (ou injonctions) et aux procédures prévus par ces dispositifs, ce qui constitue autant de contraintes que de marges de manœuvre et de leviers d'action. Par conséquent, il est intéressant de comprendre comment la SWDE est non seulement intégrée à ces dispositifs, mais également comment elle y contribue directement, et comment elle se les approprie pour pouvoir correctement mener à bien l'ensemble de ses missions, dont celle de GPI. Cette interrogation est d'autant plus importante, que les agents responsables de la SWDE perçoivent le caractère « d'outil régional » de leur structure comme étant source d'opportunités (*cf. supra*, matrice SWOT).
- Enfin, de manière complémentaire, la SWDE développe des relations de coopération avec les autres opérateurs d'eau potable intervenant en Wallonie. Il convient donc de comprendre ce qui contribue à établir ces relations, mais également les opportunités ou les effets d'aubaine ainsi créés, ainsi que les contraintes que cela fait peser sur l'action de la SWDE. Cette perspective peut d'ailleurs servir à interroger plus largement l'évolution du périmètre géographique d'intervention et de la stratégie de GPI de la SWDE. Les transferts de compétences et l'intégration de nouvelles communes sont alors vus comme un prolongement des relations de coopération, désormais fondus au sein d'un même organisme de gestion de l'eau potable et des infrastructures. Par conséquent, cet « héritage » technopolitique doit s'insérer au sein de la Société Wallonne et de sa politique, ce qui est

susceptible d'affecter si ce n'est la stratégie de la SWDE, au moins certaines de ses priorités d'action.

Il s'agira donc d'analyser les rapports de force et de pouvoir qui se jouent autant à l'échelle de la Société (au sein du conseil d'administration et en direction du comité de direction), qu'à celle de la Wallonie. Sachant que les dispositifs de coordination peuvent à la fois servir des objectifs partagés entre acteurs et nourrir des desseins plus politiques et/ou catégoriels (Caillaud, 2013), et que les infrastructures d'eau constituent des « ressources de pouvoir » (Lambelet, Pflieger, 2016), il nous semble important d'appréhender le positionnement des différents acteurs à l'endroit de la politique stratégique de la SWDE. L'analyse des logiques politiques permettra ainsi de révéler les dimensions afférentes aux jeux d'« associés-rivaux » (Bourricaud, 1961) autour desquels s'établissent les relations de coopération-concurrence, et qui sont susceptibles de travailler le système de gouvernance.

## 2.2. La SWDE comme organisation : effets de la gestionnarisation du travail sur la culture professionnelle et l'organisation de l'entreprise

La GPI doit également être interrogée dans une perspective de sociologie des organisations et du travail, en ce qu'elle représente à la fois le contenu et le produit de l'activité organisée de la SWDE (Alter, 2006 ; Bernoux, 2010). Ce faisant, l'élaboration des politiques de GPI et leur mise en œuvre ont partie liée à la structuration des métiers au sein de l'entreprise, à la culture et aux règles professionnelles en vigueur (valeurs, procédures, temporalités du travail, distribution du pouvoir, mode de management – fonctionnement en équipe, transversalité des échanges, autonomie –, etc.), aux objectifs politiques et gestionnaires, aux outils de gestion et de suivi, mais aussi aux aspects plus sociaux du travail et des agents : âge, formation, expérience professionnelle, conflictualité entre personnes, appropriation des conditions de travail, etc.

À cela, il convient de remarquer une modification non négligeable du cadre et des modalités d'exercice du travail au sein des services publics (marchands et non marchands), depuis la fin des Trente Glorieuses. Au cours des quarante dernières années, on observe une mutation des services publics désormais orientés, non plus vers un souci d'uniformité et d'égalité de traitement des usagers, mais plutôt dans une optique d'individualisation et de marchandisation des prestations, engagée au profit – dit-on – des utilisateurs (Strobel, 1993 ; Warrin, 1997 ; Weller, 1998 ; Jeannot, 1998). C'est d'ailleurs un axe de travail clairement affiché par la SWDE dans son projet stratégique 2022-2027, en faisant apparaître l'« Orientation client » comme valeur centrale, puis en la déclinant en « programme » à part entière. Ce mouvement d'individualisation et de marchandisation a pour effet de transformer profondément le fonctionnement de la bureaucratie locale et la figure de l'utilisateur : requalification du travail, « rationalisation » des organisations, recours croissant aux instruments techniques et technologiques, etc.

Cette évolution s'inscrit plus globalement dans l'imposition du nouveau management public, qui promeut une rationalisation (organisationnelle et instrumentale) du secteur public en vue d'améliorer son efficacité et sa performance (Bezès, Musselin, 2015). Parmi l'ensemble des instruments mobilisés pour y parvenir, les indicateurs de performance prennent notamment de l'importance. Ils constituent, de manière ambivalente, une source d'informations et de savoirs au service de l'action, une ressource de pouvoir, ainsi qu'un sous-bassement à la construction et à la

publicisation de problèmes plus spécifiques (Bezes, Chiapello, Desmarez, 2016). Pris dans la dynamique néo-libérale, ils contribuent à la gestionnarisation des services (Craipeau, Metzger, 2011), en mettant au centre des préoccupations le souci de maîtriser et d'améliorer – par une certaine standardisation et le recours aux nouvelles technologies de l'information et de la communication – les pratiques et les activités sociales et professionnelles, ainsi que les subjectivités. Cette vision semble ici aussi congruente avec la culture professionnelle et le projet stratégique de la SWDE, qui placent la « performance » comme valeur cardinale de l'organisation, et qui se déclinent entre autres par la recherche et l'implantation de nouvelles technologies aptes à suivre le rendement des réseaux, et par la participation de l'entreprise au benchmark européen sur la performance des services d'eau.

Ainsi, la GPI est en partie guidée – dans son élaboration et dans sa mise en œuvre – par cette quête de performance et de maîtrise de l'incertitude. Elle doit toutefois composer avec un certain nombre de facteurs qui échappent le plus souvent au fonctionnalisme de l'organisation, et qui peuvent avoir trait : à l'innovation sociale et à la création de nouvelles marges de manœuvre relevant de l'appropriation sociale du travail (Alter, 2006) ; au sens donné à celui-ci (« vrai » Vs « sale » boulot, politisation des agents, etc.) (Bidet, 2010 ; Caillaud, 2018 ; Caillaud, Nougazol, 2021) ; à son ergonomie (outils, procédures, règles, activités, conditions de l'objectivation de l'activité...) (*Travail et emploi*, 2003) ; etc.

Il s'agira donc d'analyser le cadre organisationnel de la GPI, ainsi que ce qui s'y joue de manière plus ou moins informelle, diffuse et subjective, en s'attachant à mettre en relation les dimensions socioprofessionnelle, technique et organisationnelle.

## Bibliographie

### Littérature académique

Alter, N., 2006, *Sociologie du monde du travail*, PUF, Coll. Quadrige.

Bernoux, Ph., 2010 [2004], *Sociologie du changement dans les entreprises et les organisations*, Éditions du Seuil.

Bezes, Ph., Chiapello, E., Desmarez, P., 2016, « Introduction : la tension savoirs-pouvoirs à l'épreuve du gouvernement par les indicateurs de performance », *Sociologie du travail*, Vol. 58, n°4, p. 347-369.

Bezes, Ph., Musselin, C., 2015, Le New Public Management : entre rationalisation et marchandisation ?, in : Boussaguet, L., Ravinet, P. (sous la direction de), *Une « French Touch » dans l'analyse des politiques publiques ?*, Paris, Presses de Science Po, p. 128-151.

Bidet, A., 2010, « Qu'est-ce que le vrai boulot ? Le cas d'un groupe de technicien », *Sociétés contemporaines*, Vol. 2, n°78, p. 115-135.

Bourricaud, F., 1961, *Esquisse d'une théorie de l'autorité*, Plon.

Caillaud, K., 2018, « Les conditions de mise au travail des usagers. Le cas de la gestion des déchets », *Gouvernement et action publique*, Vol. 7, n°3, p. 57-81.

Caillaud, K., 2013, *Vers une gouvernance territoriale de l'environnement ? Analyse comparée des politiques départementales de gestion de l'eau destinée à la consommation humaine et des déchets municipaux*, Thèse de doctorat de sociologie, Université de Strasbourg.

Caillaud, K., Nougarol, R., 2021, « La triple politisation de la tarification incitative. Rapports de force, réagencements et effets d'un instrument politique », *Géocarrefour* [En Ligne], Vol. 95, n°1.

Craipeau, S., Metzger, J.-L., 2011, « La gestionnarisation de la société : de la volonté de maîtrise à la déréalisation », *Mana : revue de sociologie et d'anthropologie*, n°17-18, p. 23-40.

Jeannot, G., 1998, *Les usagers du service public*, PUF, Coll. QSJ.

Lambelet, S., Pflieger, G., 2016, « Les ressources du pouvoir urbain », *Métropoles* [En ligne], Vol. 18.

Le Galès, P., 2006 « Gouvernement et gouvernance des territoires », *Problèmes politiques et sociaux*, n°922.

Strobel, P., 1993, « L'usager, le client et le citoyen : quels rôles dans la modernisation du service public ? », *Recherches et prévisions*, n°32, p. 31-44.

*Travail et emploi*, 2003, Dossier « Le travail aujourd'hui. Regards d'ergonomes et de sociologues », n°94.

Warrin, P. (dir.), 1997, *Quelle modernisation des services publics ? Les usagers au cœur des réformes*, La Découverte.

Weller, J.-M., 1998, « La modernisation du service public : évolution des approches ces dix dernières années », *Recherches et prévisions*, n°54, p. 85-92.

#### Littérature grise

AquaWal, 2018, *Statistiques de l'eau potable et de l'assainissement des eaux usées en Wallonie. Rapport 2018*.

Gouvernement Wallon, 2018, *Contrat de gestion entre le gouvernement Wallon et la Société Wallonne des Eaux, 2018-2022*.

SWDE, 2019, *Rapport annuel d'activité*.

SWDE, 2014, *Eaux plurielles. 100 ans d'une histoire singulière*.

## Annexe 12 Modélisation de la probabilité d'état fuyard. Note méthodologique de la tâche 07

### 1 Introduction

La rationalisation des dépenses de surveillance, maintenance et renouvellement des réseaux exploités par la Société Wallonne Des Eaux (SWDE) implique la mise au point de méthodes et outils de Gestion Patrimoniale des Infrastructures (GPI) qui exploitent au mieux l'information disponible au sein des différents services, et soient adaptés au cadre organisationnel et de gouvernance de la SWDE. L'exploitation optimale de l'information suppose une bonne appréhension de la chaîne de production des données depuis l'opération de terrain jusqu'à la validation au sein du Système d'Information (SI) de la SWDE; l'adaptation des méthodes et outils au contexte de la SWDE dépend de la connaissance de l'organisation des services concernés, ainsi que du fonctionnement de la gouvernance stratégique à l'échelle de la SWDE.

Le programme scientifique « Gestion Patrimoniale Multi-Echelles (GePaME) » proposé par l'INRAE se base sur une conception de la GPI pilotée par la performance du service offert à l'usager, principalement en termes de préservation des ressources en eau potable (maîtrise des pertes en réseaux) et de continuité du service (maîtrise des défaillances, causes d'interventions non programmées).

La maîtrise des pertes suppose une production préalable de connaissances du processus de fuite sur canalisations et branchements de service (questions de la « durée de vie des fuites », et de leurs relations avec le processus de casse), pour permettre in fine la modélisation de la probabilité de détection de fuite comme fonction des caractéristiques (âge, matériau, diamètre, pression, etc.) de l'élément de réseau (conduite ou branchement), et de la méthode de détection mise en œuvre. Il importe de mettre à profit les efforts importants de sectorisation et de recherche de fuites déployés par la SWDE, et les données d'observation résultantes, pour lever ce verrou de connaissance.

Ce document propose, dans le cadre de la tâche 7 du projet GePaME, une approche méthodologique visant à la modélisation de la probabilité qu'un tronçon de caractéristiques (intrinsèques et extrinsèques) données soit dans un état fuyard détectable, combinant le cadre conceptuel probabiliste de la régression logistique et une analyse de l'activité de recherche de fuite (RF) relevant des cadres épistémologiques de la sociologie du travail et de la sociologie des sciences et techniques. Cette approche interdisciplinaire vise à garantir que les hypothèses fondant la modélisation probabiliste sont raisonnablement satisfaites, et le cas échéant à adapter cette dernière aux imperfections possiblement mises en évidence dans les données résultant de la chaîne d'opérations allant de la programmation des campagnes de RF à l'alimentation du SI.

### 2 Cadre conceptuel de la régression logistique

La régression logistique est une méthode statistique, de la classe des « modèles linéaires généralisés », dot la théorie et la mise en œuvre pratique sont présentées par McCullagh and Nelder [1989]. Cette méthode d'analyse vise à modéliser la probabilité de réponse d'un ensemble d'individus soumis à des expériences aléatoires de Bernoulli, comme une fonction, dite fonction de « lien », d'une combinaison linéaire de variables caractéristiques de ces individus. Ces variables, souvent dénommées « variables indépendantes » ou « covariables » dans le cadre des modèles de régression, sont considérées comme « explicatives » du comportement des individus. Les covariables caractérisant un individu quelconque sont dénotées par le vecteur  $Z$  dans

la suite de ce document.

Une expérience aléatoire de Bernoulli offre deux résultats possibles, formalisés par une variable aléatoire binaire de « réponse », dénotée dans la suite  $Y \in \{0, 1\}$ .

Le modèle de régression logistique considère la probabilité conditionnelle\* :

$$p = \Pr\{Y = 1 \mid \mathbf{Z}\} \in [0, 1] \quad (2.1)$$

qui est aussi l'espérance conditionnelle de  $Y$  :

$$p = E(Y \mid \mathbf{Z}) \quad (2.2)$$

et la transformation *logit* :

$$\forall x \in ]0, 1[, \quad g(x) = \ln\left(\frac{x}{1-x}\right) \quad (2.3)$$

L'effet des covariables  $\mathbf{Z}$  sur la probabilité de réponse  $p$  est alors formalisé par le modèle de régression :

$$g(p) = \mathbf{Z}^T \boldsymbol{\beta} \quad (2.4)$$

où  $\boldsymbol{\beta}$  est un vecteur de coefficients de régression dont les valeurs mesurent les effets des covariables sur la réponse  $Y$ .

La prédiction du modèle de régression logistique est donc obtenue par :

$$E(Y \mid \mathbf{Z}) = \frac{e^{\mathbf{Z}^T \boldsymbol{\beta}}}{1 + e^{\mathbf{Z}^T \boldsymbol{\beta}}} \quad (2.5)$$

Le calage statistique du modèle de régression logistique consiste à estimer le vecteur de paramètres  $\boldsymbol{\beta}$  à partir d'un échantillon aléatoire de  $n$  observations  $y_i \in \{0, 1\}$  de la variable de réponse  $Y$ , caractérisées chacune par le vecteur de covariables  $\mathbf{Z}_i, i \in \{1, \dots, n\}$ . La méthode canonique d'estimation de  $\boldsymbol{\beta}$  procède en construisant la fonction de vraisemblance :

$$L(\boldsymbol{\beta}) = \prod_{i=1}^n p_i^{y_i} (1 - p_i)^{1-y_i} \quad (2.6)$$

et en cherchant numériquement la valeur optimale  $\hat{\boldsymbol{\beta}}$  qui maximise le logarithme naturel de  $L(\boldsymbol{\beta})$  :

$$\hat{\boldsymbol{\beta}} = \underset{\boldsymbol{\beta}}{\operatorname{argmax}} \ln L(\boldsymbol{\beta}) \quad (2.7)$$

### 3 Application de la régression logistique aux données de recherche de fuites

Est envisagé dans un premier temps le schéma le plus simple dans lequel est observé, comme exposé en fin de section précédente 2, un échantillon aléatoire de tronçons de réseau d'eau, supposé représentatif de la population des tronçons du système d'alimentation en eau potable (SAEP) étudié.

La prise en compte du mode de programmation des campagnes de recherche de fuites (RF), fonction de la sectorisation et de la surveillance des cellules, est abordée dans un second temps.

\*Le choix de considérer l'événement  $Y = 1$  plutôt que l'événement contraire  $Y = 0$  est purement conventionnel, mais conditionne cependant l'interprétation des résultats de l'analyse.

### 3.1 Campagne aléatoire de recherche de fuite

Dans ce schéma simple,  $n$  tronçons de réseau du SAEP, indexés par  $i \in \{1, \dots, n\}$ , sont inspectés dans le but de détecter la présence d'une fuite sur la canalisation; l'échantillon inspecté peut résulter de plusieurs campagnes de RF effectuées à des dates différentes dans des conditions comparables, ou dont les différences sont documentées. Le tronçon  $i$  est inspecté à l'âge  $t_i$ , et caractérisé par les variables rassemblées dans le vecteur  $\mathbf{Z}_i$ ; ces variables caractérisent la canalisation - matériau, diamètre, longueur, nombre de raccordements, *etc.* -, son environnement - position sous chaussée, trottoir, ou espace naturel, intensité du trafic dans la voie, *etc.* -, ou ses conditions d'exploitation - pression de service, type d'eau, mode de désinfection, *etc.* Il est aussi possible d'introduire dans  $\mathbf{Z}_i$  des variables transformées, telles que *e.g.* le logarithme de la longueur ou de l'âge, ou des combinaisons non linéaires de certaines caractéristiques du tronçon.

La mise en oeuvre du modèle formalisé par l'Eq. (2.5), détaillé en section 2 permet directement d'estimer la probabilité conditionnelle  $\Pr\{Y = 1 \mid \mathbf{Z}\}$ , où  $Y$  désigne l'indicatrice aléatoire de l'état fuyard du tronçon inspecté.

Il convient ici d'insister sur l'importance de la complétude de l'échantillon d'observations de RF, dont dépend la validité de l'analyse statistique mise en oeuvre; l'échantillon d'observations doit impérativement référencer la totalité des conduites incluses dans le périmètre investigué par la campagne de RF, y compris celles dont les équipes techniques en charge des inspections ont considéré qu'elles étaient trivialement non fuyardes. Ce point d'attention est à la croisée de l'ingénierie et de la sociologie, respectivement dans les tâches 5 et 6 du projet GePaME.

### 3.2 Prise en compte du mode de surveillance des fuites

En pratique les campagnes de RF ne sont pas effectuées au hasard, mais ciblent :

- les cellules dont la surveillance du débit nocturne permet de détecter une dérive soudaine à la hausse,
- les tronçons au sein de ces cellules qui sont dans une zone de la cellule qu'une opération de pré-localisation a désigné comme susceptible de recéler des fuites.

La démarche technique de la surveillance des pertes d'eau en réseau sectorisé est présentée en détails par Renaud et al. [2014].

La prise en compte du mode de surveillance des fuites nécessite d'introduire, en sus de la variable aléatoire de réponse  $Y$  indicatrice de la détection d'une fuite sur le tronçon, une nouvelle variable aléatoire notée  $D$  indicatrice de l'exposition\* du tronçon à la campagne de RF. Cette exposition est susceptible de dépendre de covariables caractérisant les cellules et le résultat de leur suivi, telles que *e.g.* un indicateur de la dérive du débit de nuit; ces covariables sont rassemblées dans le vecteur noté  $\mathbf{X}$ .

Le modèle formalisé par l'Eq. (2.5) peut alors être appliqué à la probabilité conditionnelle  $\Pr\{Y = 1 \mid D = 1, \mathbf{Z}, \mathbf{X}\}$ .

\*Le terme « exposition », issu du champ lexical de l'épidémiologie, signifie que le tronçon a été soumis à la campagne de RF.

Cependant, la prise en compte du risque de fuite comme critère de programmation annuelle des renouvellements de canalisations, ainsi que comme composante d'un système de simulation long terme de politique patrimoniale, nécessite de disposer d'une modélisation de la probabilité conditionnelle  $\Pr\{Y = 1 \mid \mathbf{Z}, \mathbf{X}\}$ . En utilisant la formule des probabilités totales, cette quantité peut être calculée comme :

$$\Pr\{Y = 1 \mid \mathbf{Z}, \mathbf{X}\} = \Pr\{Y = 1 \mid D = 1, \mathbf{Z}, \mathbf{X}\} \Pr\{D = 1 \mid \mathbf{X}\} + \Pr\{Y = 1 \mid D = 0, \mathbf{Z}, \mathbf{X}\} \Pr\{D = 0 \mid \mathbf{X}\} \quad (3.1)$$

Il est ainsi nécessaire de compléter la modélisation :

- en procédant à des campagnes de RF expérimentales, à seul but de production de connaissances, sur des cellules dont le suivi du débit de nuit ne laisse pas suspecter qu'elles recèlent des fuites, afin de modéliser  $\Pr\{Y = 1 \mid D = 0, \mathbf{Z}, \mathbf{X}\}$ ;
- en collectant les données adéquates relatives à la sectorisation, afin d'établir le modèle complémentaire  $\Pr\{D = 1 \mid \mathbf{X}\}$ .

### 3.3 Processus de fuite et de casse

Lors de sa mise en oeuvre sur données réelles de réparations de canalisations, la modélisation des défaillances, telle que présentée en annexe 4 en suivant l'approche du modèle LEYP, ne peut généralement pas distinguer, faute de documentation adéquate des événements, les fuites réparées suite à leur détection lors d'une campagne de RF, des casses imprévues de conduites se manifestant par un écoulement abondant en surface et conduisant à une intervention non programmée sur le réseau.

Un point de vue fréquemment adopté, et formalisé par le travail doctoral de Chesneau [2006], considère que le défaut d'étanchéité est un processus en temps continu transitant par trois états discrets :

- état de fuite diffuse, de débit trop faible pour être détectable par les moyens couramment mobilisés en RF,
- état de fuite détectable par la technologie de RF,
- état de casse manifeste, nécessitant une intervention rapide.

Cette conception, éminemment questionnable comme toute hypothèse scientifique, est séduisante en ce qu'elle offre un cadre unifié des phénomènes de fuite et de casse, et fournit une base de réflexion dont le travail de modélisation prévu en tâche 9 du projet GePaME devrait permettre soit d'affirmer la pertinence, soit de remettre en cause le bien-fondé.

L'événement considéré par le modèle LEYP est le passage de l'état de fuite détectable à celui de casse manifeste ; la régression logistique, formalisée par l'Eq. 2.5, modélise la présence dans l'état de fuite détectable, qui dépend à la fois de l'intensité du passage de l'état de fuite diffuse à celui de fuite détectable, et du temps de séjour dans ce dernier. En outre le LEYP est un processus doté de mémoire des événements passés, contrairement à la régression logistique qui en est dépourvue. L'enjeu théorique de la tâche 9 est ainsi d'inventer une synthèse des

deux modèles, une forme possible enchaînant une entrée dans l'état de fuite détectable selon un LEYP, et un temps de séjour dans cet état suivant une distribution à déterminer (e.g. une distribution exponentielle).

La section suivante clôt ce document en rappelant les bases théoriques du modèle LEYP.

## 4 Le modèle LEYP de défaillances répétées

Comme exposé par Le Gat [2014], les travaux internationaux consacrés au phénomène de défaillance des conduites d'eau sous pression s'accordent sur les points suivants :

- le taux instantané de défaillance (nombre de défaillances susceptibles d'affecter une canalisation par unité de temps) tend à augmenter avec l'âge de la canalisation ;
- à un âge donné, le taux de défaillance varie selon des caractéristiques propres à la canalisation (longueur, diamètre, matériau, type de joints, *etc.*), ainsi que selon la nature de son environnement (position sous chaussée ou sous trottoir, trafic dans la voie, nature du sol encaissant, *etc.*) ou de son exploitation (pression de service, variation de cette dernière, agressivité de l'eau transportée, *etc.*) ;
- les défaillances présentent une certaine agrégation spatiale, qui se traduit par leur tendance à s'accumuler sur les mêmes canalisations, le délai entre deux défaillances successives tendant en outre à se raccourcir quand le rang de la défaillance augmente.

Cela a conduit à proposer un modèle probabiliste du processus de défaillance issu du cadre théorique des *processus stochastiques de comptage* (cf. Ross [1996]), basé sur le processus de Yule (cf. Greenwood and Yule [1920]), et conçu comme une synthèse des approches précédemment proposées par Eisenbeis [1994], Pelletier [1999] et Røstum [2000].

Comme illustré par la figure 4.1, le processus de défaillance d'une canalisation, dont l'âge est noté  $t$ , est formalisé par la fonction de comptage  $N(t) \in \mathbb{N}$ , qui prend la valeur 0 à  $t = 0$  (i.e. à l'installation de la canalisation), et s'incrémente d'une unité à chaque défaillance ; la différentielle  $dN(t)$  de  $N(t)$  est nulle partout sauf à chaque instant de défaillance  $t_j$ , où elle vaut 1.

L'Extension Linéaire du Processus de Yule, *Linear Extension of the Yule Process* (LEYP), est définie par son intensité, qui est l'espérance conditionnelle de la différentielle de  $N(t)$  :

$$\forall t \in \mathbb{R}_+, \alpha \in \mathbb{R}_+^* : \\ E(dN(t) | N(t-), \mathbf{Z}) = (1 + \alpha N(t-))\lambda(t, \mathbf{Z})dt$$

où  $N(t-)$  est le nombre de défaillances subies par la canalisation entre sa pose et l'instant précédant juste l'âge  $t$ , et  $\lambda(t, \mathbf{Z})$  est une fonction réelle de  $t$  et des caractéristiques de la canalisation rassemblées dans le vecteur  $\mathbf{Z}$  (covariables).

Le LEYP contient comme cas particuliers le processus de Poisson non homogène (NHPP), pour  $\alpha = 0$ . A noter que le processus de Poisson a la propriété importante d'être sans mémoire de son passé, contrairement au processus de Yule ( $\alpha > 0$ ).

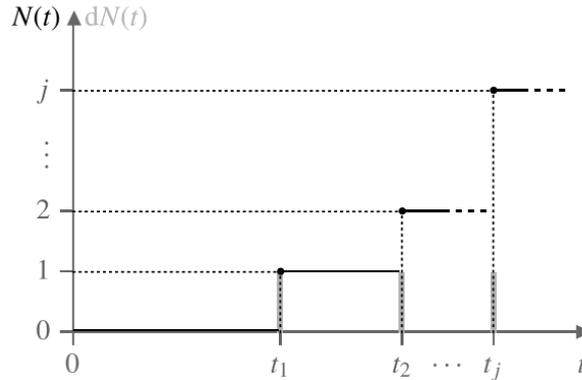


FIGURE 4.1 : Fonction de comptage  $N(t)$  et sa dérivée  $dN(t)$

Alors que la fonction de comptage du Processus de Poisson suit une distribution de Poisson, celle du LEYP suit une loi binomiale négative :

$$\forall t \in \mathbb{R}_+, \quad N(t) \sim \mathcal{NB}(\alpha^{-1}, \mu(t)^{-1})$$

où, pour alléger la notation :

$$\mu(t) = \exp\left(\alpha \int_0^t \lambda(u) du\right)$$

La distribution binomiale négative est conservée pour le nombre de défaillances susceptibles de se produire dans un intervalle de temps  $[c, d]$ , connaissant le nombre de défaillances qui se sont produites dans un intervalle antérieur  $[a, b]$ , comme illustré par la figure 4.2 :

$$[N(d) - N(c) \mid N(b) - N(a) = m] \sim \mathcal{NB}\left(\alpha^{-1} + m, \frac{\mu(b) - \mu(a) + 1}{\mu(d) - \mu(c) + \mu(b) - \mu(a) + 1}\right)$$

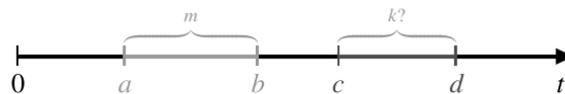


FIGURE 4.2 : Fenêtres d’observation et de prédiction

Cette propriété est importante pour l’utilisation pratique du modèle LEYP, car elle permet de prédire explicitement le nombre de défaillances futures en se basant sur une connaissance incomplète de l’historique de défaillances (*i.e.* quand  $a > 0$ ), ce qui le plus souvent le cas pour des canalisations anciennes, seulement observées sur une fenêtre temporelle récente et étroite.

## Références

- O. Chesneau. *Un outil d'aide à la maîtrise des pertes dans les réseaux d'eau potable : la modélisation dynamique de différentes composantes du débit de fuite (in French). A tool to help to control losses in drinking water networks : dynamic modelling of leakage flow components*. PhD thesis, University Louis Pasteur, Strasbourg, France, 2006.
- P. Eisenbeis. *Modélisation statistique de la prévision des défaillances sur les conduites d'eau potable (in French). Statistical modelling of failure prediction on water mains*. PhD thesis, University Louis Pasteur, Strasbourg, France, 1994.
- M. Greenwood and G. U. Yule. An inquiry into the nature of frequency distributions representative of multiple happenings with particular reference to the occurrence of multiple attacks of disease or of repeated accidents. *Journal of the Royal Statistical Society, Series A*, 83 : 255–279, 1920.
- Y. Le Gat. Extending the Yule process to model recurrent pipe failures in water supply networks. *Urban Water Journal*, 11(8) :617–630, 2014.
- P. McCullagh and J.A. Nelder. *Generalized Linear Models*. Monographs on Applied Probability and Statistics 37. Chapman and Hall, 2nd edition, 1989.
- G. Pelletier. *Impact du remplacement des conduites d'aqueduc sur le nombre annuel de bris*. PhD thesis, Université du Québec, INRS-Eau, 1999.
- E. Renaud, J. Pilot, A. Auckenthaler, and C. Aubrun. Réduction des pertes d'eau des réseaux de distribution d'eau potable - guide pour l'élaboration du plan d'actions (décret 2012-97 du 27 janvier 2012). Technical report, Office National de l'Eau et des Milieux Aquatiques, 2014.
- S. Ross. *Stochastic Processes*. John Wiley and Sons, Inc., New York, second edition, 1996.
- J. Røstum. Statistical modelling of pipe failures in water networks. Doctor engineer dissertation, Norwegian University of Science and Technology, Department of Hydraulic and Environmental Engineering, Trondheim, Norway, 2000.

