



**HAL**  
open science

## Évaluation du lien performance – patrimoine des réseaux d'eau potable Rapport intermédiaire

Alain Husson, Yves Le Gat, Caroline Merlet, Eddy Renaud, Christophe Wittner

► **To cite this version:**

Alain Husson, Yves Le Gat, Caroline Merlet, Eddy Renaud, Christophe Wittner. Évaluation du lien performance – patrimoine des réseaux d'eau potable Rapport intermédiaire. [Rapport Technique] Programme 2019-2020 - Action n°01, INRAE. 2019, 48 p. hal-03168242

**HAL Id: hal-03168242**

**<https://hal.inrae.fr/hal-03168242>**

Submitted on 12 Mar 2021

**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.

Public Domain

Programme 2019-2020 – Action n° 01

# **Evaluation du lien performance – patrimoine des réseaux d'eau potable**

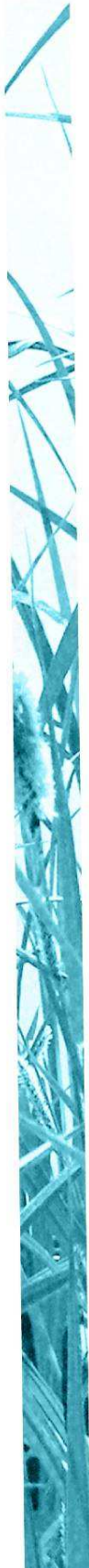
## **Rapport Intermédiaire**

**Alain HUSSON (Irstea)  
Yves LEGAT (Irstea)  
Caroline MERLET (Irstea)  
Eddy RENAUD (Irstea)  
Christophe WITTNER (UMR GESTE)**

**Décembre 2019**

Document élaboré dans le cadre de la convention AFB – Irstea 2019-2020

En partenariat avec le Ministère de la Transition Ecologique et Solidaire



- **AUTEURS**

**Alain HUSSON**, Ingénieur (1), [alain.husson@irstea.fr](mailto:alain.husson@irstea.fr)

**Yves LEGAT**, Ingénieur (1), [yves.legat@irstea.fr](mailto:yves.legat@irstea.fr)

**Caroline MERLET**, Ingénieure (1), [caroline.merlet@irstea.fr](mailto:caroline.merlet@irstea.fr)

**Eddy RENAUD**, Ingénieur (1), [eddy.renaud@irstea.fr](mailto:eddy.renaud@irstea.fr)

**Christophe WITTNER**, Ingénieur (2), [christophe.wittner@irstea.fr](mailto:christophe.wittner@irstea.fr)

(1) INRAE, ETBX  
50 avenue de Verdun Gazinet, F-33612 CESTAS cedex

(2) UMR GESTE ENGEES INRAE  
1quai Koch, 67070 STRASBOURG

- **CORRESPONDANTS**

**AFB** :

**Eric BREJOUX**, Direction de la Connaissance et de l'Information sur l'Eau,  
[eric.brejoux@afbiodiversite.fr](mailto:eric.brejoux@afbiodiversite.fr)

**Irstea** :

**Eddy RENAUD**, responsable de l'équipe GPIE,  
[eddy.renaud@irstea.fr](mailto:eddy.renaud@irstea.fr)

Droits d'usage : accès libre  
Niveau géographique : national  
Couverture géographique : France  
**Niveau de lecture** : Elus, professionnels, experts

- **SOMMAIRE**

<b>1. Introduction</b> .....	<b>5</b>
1.1. Contexte .....	5
1.2. Objectifs et contenu de l'étude .....	5
<b>2. Création et caractéristiques de la base de données patrimoine-performance</b> .....	<b>5</b>
2.1. Les données « performance » .....	5
2.2. Les données « patrimoine » .....	6
2.3. Les données de « contexte » .....	7
2.4. La base de données performance – patrimoine .....	7
<b>3. Corrélations entre les paramètres de patrimoine et de performance</b> .....	<b>11</b>
3.1. Corrélations entre l'IVNC et le patrimoine.....	11
3.2. Corrélations entre les paramètres de patrimoine .....	12
3.3. Analyse en composante principale .....	13
<b>4. Liens entre le patrimoine, la performance et les paramètres de contexte</b> .....	<b>14</b>
4.1. Le mode de gestion .....	14
4.2. Les hydroterritoires .....	14
4.3. Le taux d'urbanisation.....	17
4.4. La taille du service.....	18
<b>5. Etude des courbes cumulées</b> .....	<b>19</b>
5.1. Cumuls après classement décroissant des proportions .....	19
5.2. Cumuls après classement croissant des proportions .....	20
<b>6. Modélisation de l'IVNC par classification</b> .....	<b>22</b>
6.1. Méthode de classification ascendante hiérarchique.....	22
6.2. Résultats sur les matériaux.....	22
6.3. Résultats sur les diamètres.....	23
<b>7. Modélisation du VNC</b> .....	<b>24</b>
7.1. Les modèles linéaires.....	24
7.2. Les modèles « exponentiels ».....	28
7.3. Le modèle linéaire avec modulation sur les abonnés .....	29
7.4. Modèle de prédiction de $\ln(\text{VNC}/N)$ .....	31
<b>8. Etude de l'Indice de Connaissance et de Gestion du Patrimoine (ICGP)</b> .....	<b>33</b>
8.1. La base de données d'étude de l'ICGP.....	33
8.2. Comparaison des ICGP calculés à partir des SIG et de la base SISPEA .....	33
<b>9. Etude à l'échelle du secteur</b> .....	<b>34</b>
9.1. Intérêt.....	34
9.2. Terrains d'étude .....	34
9.3. Données.....	34
<b>10. Evaluation entre la valeur à neuf des réseaux publics d'eau potable et le prix de l'eau potable</b>	<b>34</b>
10.1. Introduction.....	34
10.2. Données mobilisées et caractérisation de l'échantillon.....	34
10.3. Premiers résultats.....	35
10.4. Perspectives.....	37

<b>11. Synthèse .....</b>	<b>38</b>
<b>12. Bibliographie .....</b>	<b>40</b>
<b>13. Table des illustrations .....</b>	<b>41</b>
<b>14. Annexes .....</b>	<b>42</b>
<b>14.1. Annexe 1 : Carte des communes en zone de montagne .....</b>	<b>42</b>
<b>14.2. Annexe 2 : Cartes du paramètre « dénivelé » .....</b>	<b>43</b>
<b>14.3. Annexe 3 : Carte des hydroterritoires (Aquadep, 2012) .....</b>	<b>45</b>

## • EVALUATION DU LIEN PERFORMANCE – PATRIMOINE DES RESEAUX D’EAU POTABLE

### 1. Introduction

#### 1.1. Contexte

La gestion patrimoniale des réseaux d’eau potable est un enjeu actuel important à l’échelle nationale, notamment dans l’objectif de limiter les volumes de pertes et préserver la ressource en eau. Les indicateurs de performance définis dans l’arrêté du 2 décembre 2013 modifiant l’arrêté du 2 mai 2007 relatif aux Rapports annuels sur le Prix et la Qualité des Services (RPQS), ainsi que le descriptif détaillé des réseaux imposé par la loi Grenelle II et son décret d’application n°2012-97 du 27 janvier 2012, sont des éléments destinés à améliorer la gestion patrimoniale à l’échelle des services.

La bancarisation des indicateurs de performance (Système d’Information sur les Services Publics d’Eau et d’Assainissement, SISPEA) permet d’évaluer la performance des réseaux à l’échelle des services, mais sans lien avec les caractéristiques physiques des réseaux (diamètres, matériaux et dates de pose).

Les travaux menés par IRSTEA et l’AFB dans le cadre de l’action 70 de 2016 à 2018 ont permis d’estimer les longueurs de conduites par matériau, par période de pose et par classe de diamètre aux échelles nationale et de bassin par la construction d’un dispositif statistique d’évaluation du patrimoine des réseaux d’eau potable. L’axe 2 de l’action 70 contient une évaluation des besoins des acteurs de l’eau, soulevant la problématique d’un lien entre les caractéristiques des réseaux et les performances, dans l’objectif d’orienter plus intelligemment les actions d’amélioration des réseaux.

#### 1.2. Objectifs et contenu de l’étude

Cette action a pour objectif de relier les indicateurs de performance aux caractéristiques techniques et financières des réseaux, par typologie de services et à des échelles nationales ou régionales (bassins, régions administratives ou géographiques), pour mieux cibler les politiques publiques liées à la gestion patrimoniale des réseaux d’eau potable.

L’objectif de l’action est donc de répondre à la question suivante : peut-on lier la performance des réseaux au patrimoine (type de canalisations) ou à des éléments de contexte (urbanisation, territoires...) ?

L’étude conjointe des indicateurs de performance de la base SISPEA et des données de patrimoine exploitées dans le cadre de l’action 70 a permis d’aborder cette problématique. Les travaux effectués durant l’année 2019 et explicités dans ce rapport ont mené à :

- La création d’une base de données « patrimoine-performance » contenant des données sur le territoire national à l’échelle des services ;
- L’étude des liens entre le patrimoine (diamètre, matériau et période de pose) et la performance, conjointement avec des éléments dits de contexte, utilisant notamment les outils statistiques suivants : coefficients de corrélation linéaires, analyse en composante principale, cumul des données, classification ascendante hiérarchique, et enfin modélisation linéaire et non linéaire.

### 2. Création et caractéristiques de la base de données patrimoine-performance

#### 2.1. Les données « performance »

La base SISPEA fournit des données de performance annuelle à l’échelle des services. Les données de l’année 2016 sont utilisées et complétées lorsqu’elles ne sont pas renseignées pour un service par les bases SISPEA des années 2015 et 2017. Seuls les services de distribution sont conservés (12 415 services de distribution sur 12 799 au total en 2016).

Les indicateurs SISPEA étudiés dans ce rapport sont :

- L’Indice linéaire des Volumes Non Comptés (IVNC), en  $m^3/km/j$  :

$$IVNC = \frac{V_{\text{produit}} + V_{\text{importé}} - V_{\text{exporté}} - V_{\text{comptabilisé domestique}} - V_{\text{comptabilisé non domestique}}}{\text{Linéaire hors branchement}}$$

Le paramètre IVNC a été privilégié car il est bien renseigné et moins biaisé que l'ILP ou le rendement.

- L'Indice de Connaissance et de Gestion Patrimoniale des réseaux d'eau potable (ICGP) et notamment les 45 premiers points de l'indice.

Les données suivantes renseignées dans la base SISPEA sont également utilisées :

- Le nombre d'abonnés (N) ;
- La longueur de réseau hors branchement (L) ;
- Le mode de gestion (régie ou délégation).

Un autre indicateur est créé à partir de l'IVNC pour les besoins de l'étude, le « VNC/N ».

$$\frac{VNC}{N} = \frac{IVNC \times L}{N}$$

Le taux de remplissage des indicateurs de performance et des données utilisées dans la base SISPEA est d'environ 50 %.

## 2.2. Les données « patrimoine »

### 2.2.1. La base de données collectées

Les données patrimoine utilisées sont les mêmes que la base SIG de l'Action 70 menée par l'AFB et IRSTEA de 2016 à 2018. A l'échelle des communes, elle contient les longueurs de tronçons par période de pose (7 classes de périodes de pose), par matériau (10 classes de matériau) et par diamètre (7 classes de diamètres) de 13 346 communes, soit 398 000 km de réseau d'eau potable. L'information « matériau par période de pose » et « matériau par diamètre » est également connue.

Le taux d'inconnu sur le patrimoine n'est pas négligeable puisqu'il est de 4 % en moyenne sur les diamètres, 6 % sur les matériaux et 30 % sur les périodes de pose.

### 2.2.2. La base de données modélisées

La base de données modélisées correspond aux longueurs de réseaux par diamètres, matériaux et périodes de pose estimées par l'action 70. Cette base contient les données de 36 527 communes, soit un linéaire de 874 963 km correspondant à l'intégralité du réseau national estimé.

Ces données ont l'avantage de ne pas avoir d'inconnues sur le patrimoine, et les biais (périodes de pose récentes et gros diamètres mieux renseignés) sont corrigés par le modèle. En raison du plus grand nombre de données contenues dans cette base, elle sera utilisée de manière privilégiée dans toute la suite du rapport (une précision est apportée lorsque ce sont les données collectées qui sont étudiées).

Les données de patrimoine sont réparties selon les classes ci-dessous (regroupement par rapport aux classes définies dans l'action 70) :

**Tableau 1 : Classes sur les données patrimoine et abréviations utilisées**

	Abréviation	Occurrence par rapport à la longueur totale
<b><u>Matériaux</u></b>		
<b>Fonte</b>	<b>F</b>	<b>42 %</b>
Fonte grise	FG	17 %
Fonte ductile	FD	25 %
<b>PVC</b>	<b>PVC</b>	<b>47 %</b>
PVC vieux (< 1980)	PVCV	31 %
PVC récent (> 1980)	PVCR	16 %
<b>Amiante-ciment, acier, PE...</b>	<b>Autres</b>	<b>11 %</b>

<b><u>Périodes de pose</u></b>		
Antérieure à 1965	PP ≤ 1965	28 %
Entre 1965 et 1980	PP 1965 - 1980	37 %
Postérieure à 1980	PP > 1980	35 %
<b><u>Diamètres</u></b>		
Inférieur à 50 mm	≤ 50	27 %
Entre 60 et 100 mm	60 – 100	42 %
Supérieur à 125 mm	≥ 125	31 %

Ces classes ont été choisies de manière à avoir des classes suffisamment grandes et exploitables statistiquement.

### **2.3. Les données de « contexte »**

Les données de contexte sont connues à l'échelle des communes. Les paramètres suivants sont les mêmes que ceux utilisés dans la base communale de l'Action 70 :

- **La population** ;
- **Les longueurs de routes urbaines et rurales**, à partir desquelles sont calculées les proportions de routes urbaines et de routes rurales ;
- **Les zones de montagne** : indicateur des communes 2015 défini par l'Observatoire des territoires (Ministère de l'Agriculture, de l'Agroalimentaire et de la Forêt), la carte est présentée en Annexe 1 ;
- **Les hydroterritoires** : Indicateur Aquadep se déclinant en 5 classes définies selon des caractéristiques des géographies physique et humaine, des ressources en eau, de la pression agricole et de l'organisation de l'AEP des départements (moyenne et haute montagne, France de l'Est et du Nord-Ouest, France de l'Ouest et du Sud-Ouest, Sud urbanisé, Ile-de-France), la carte est disponible en Annexe 3.A. ;
- **L'agence de l'eau** ;
- **Les classes Corinne Land Cover (CLC)**, divisées en quatre classes (zone habitée, zone artificialisée hors zone habitée, zone agricole, zone naturelle).

Un autre paramètre de contexte est calculé :

- **Le dénivelé** : connaissant les altitudes sur le territoire national et les classes CLC, 8 données d'altitude sont enregistrées par commune (altitude maximale et altitude minimale dans chaque classe CLC). Les altitudes de la classe « zone naturelle » sont écartées, permettant de se rapprocher de la différence d'altitude supportée par le réseau d'eau potable. Les valeurs sont divisées par le dénivelé maximal observé pour toutes être comprises entre 0 et 1, la carte des dénivelés est présentée en Annexe 2.A. ;

Ces données sont toutes connues à l'échelle des communes et sont intégrées à la base de données patrimoine. **Le mode de gestion** (délégation ou régie) est aussi un paramètre de contexte et il est disponible dans la base SISPEA à l'échelle du service.

## **2.4. La base de données performance – patrimoine**

### **2.4.1. Description de la méthode de création de la base**

La base de données patrimoine-performance associe les données performance disponibles à l'échelle des services aux données patrimoine disponibles à l'échelle des communes. Le traitement des communes et des services se fait suivant les étapes décrites ci-dessous.

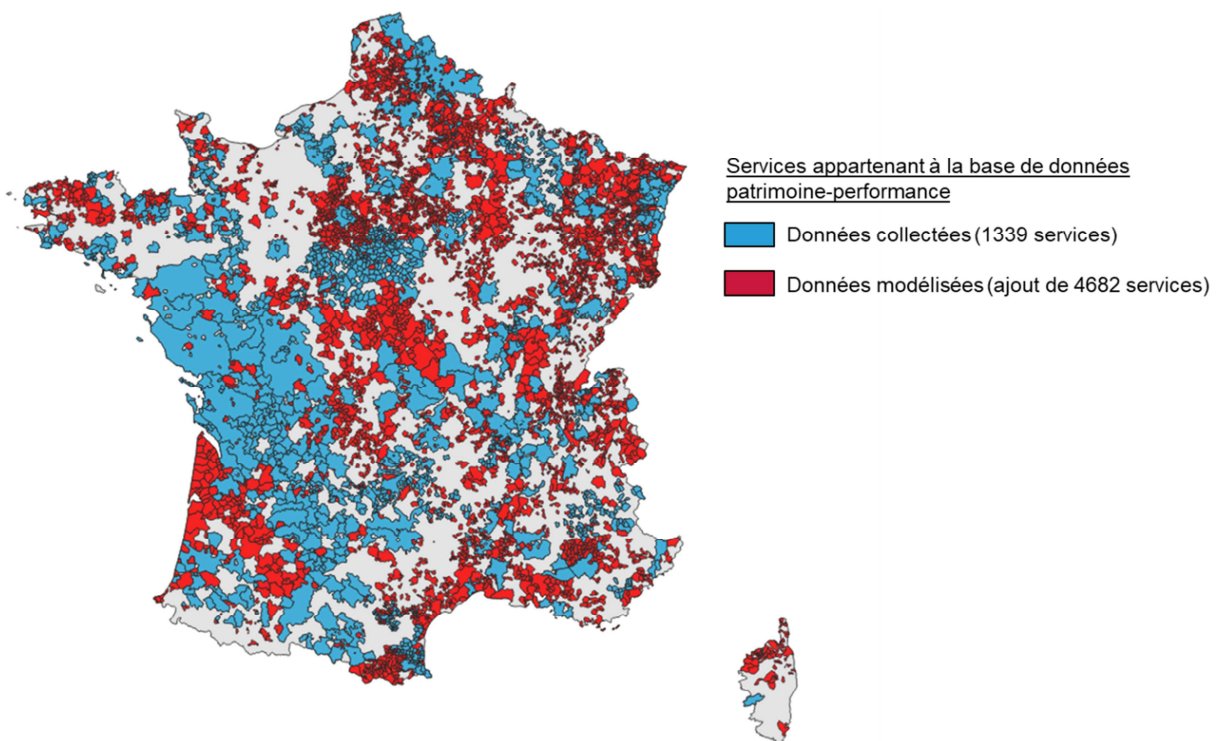
- **Etape 1** : Les services pour lesquels les paramètres « linéaire hors branchement », « nombre d'abonnés » et « IVNC » sont renseignés dans la base SISPEA 2016 sont sélectionnés ;
- **Etape 2** : Les services composés d'une ou plusieurs communes dont le réseau appartient entièrement à un seul service sont d'abord retenus ;
- **Etape 3** : Les services dont au moins une commune est partagée en plusieurs services sont regroupés en mégaservices par concaténation ;



- Etape 4 : Les services dont au moins une commune est partagée en plusieurs services mais pour lesquels on ne construit pas de mégaservice car les 3 paramètres d'intérêt sont manquants pour au moins un service sont traités individuellement : la commune partagée est attribuée à l'un des services dont elle fait partie ;
- Etape 5 : Les étapes 1, 2 et 3 sont répétées sur la base de données SISPEA 2015 pour compléter les données 2016 ;
- Etape 6 : Les étapes 1, 2 et 3 sont répétées sur la base de données SISPEA 2017 pour compléter les données 2015 et 2016.
- Etape 7 :
  - Les données de patrimoine et de contexte sont calculées à l'échelle des services. La carte de la variable « dénivelé » à l'échelle des services est disponible en Annexe 2.B.
  - Dans le cas où un service est sur plusieurs agences, hydroterritoires ou en partie en zone montagneuse, le critère contenant la plus grande longueur de réseau est attribué au service. La carte des hydroterritoires par services est disponible en Annexe 3.B.
  - Dans le cas des mégaservices, les données SISPEA sont recalculées en respectant les consignes de calcul SISPEA de chaque indicateur.
  - Dans le cas des communes appartenant à plusieurs services et attribuées à un seul service, les variables de patrimoine sont corrigées par extrapolation pour correspondre à la longueur de réseau indiquée dans SISPEA.

Cette méthode est appliquée indépendamment aux données SIG collectées et aux données SIG modélisées. Pour les données SIG collectées, la base créée contient 1339 services représentant 11 258 communes pour une longueur totale de 347 348 km. Considérant les inconnues sur le patrimoine, cette base contient peu de données pour le traitement statistique. C'est pourquoi l'utilisation des données SIG modélisées est privilégiée, la base de données patrimoine-performance contenant **6021 services représentant 22 005 communes et 582 108 km**.

Les services faisant partie de la base de données patrimoine-performance sont représentés sur la carte de la Figure 1.



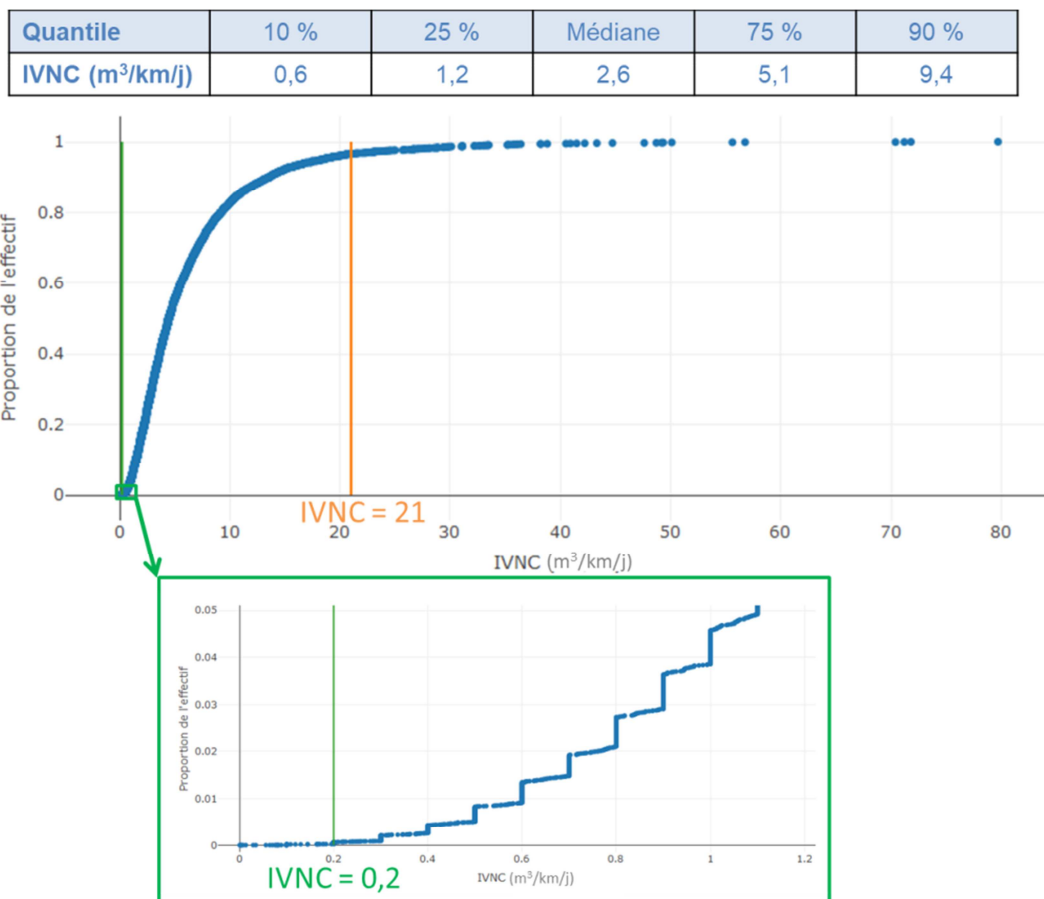
**Figure 1 : Cartographie des services de la base de données patrimoine-performance**

#### **2.4.2. Traitement des valeurs extrêmes**

Cette phase de traitement a pour objectif de supprimer les valeurs extrêmes de la base de données patrimoine-performance. Ces données qui sont très différentes de la majorité des valeurs sont soit suspectes, soit reflètent

des particularités locales. Elles ne sont pas conservées car elles ont un trop gros impact sur les modèles statistiques.

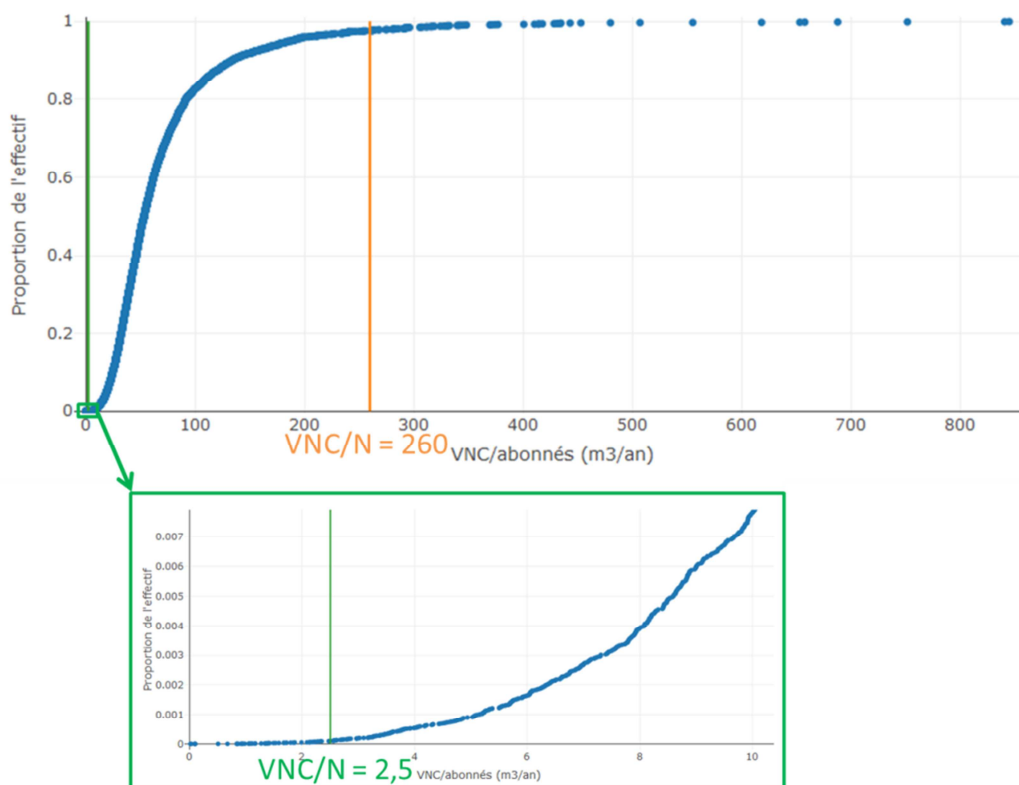
Prenons l'exemple de l'IVNC. La méthode est la suivante : les données sont classées par IVNC croissants, le rang est cumulé et représenté en abscisses sur un graphique et la valeur de l'IVNC est reportée en ordonnées. Cette méthode permet de distinguer visuellement les valeurs très élevées et très faibles. Les limites supérieures et inférieures sont estimées graphiquement sur la Figure 2.



**Figure 2 : Tri des services d'après le paramètre IVNC de la base de données patrimoine-performance**

103 services de la base de données patrimoine-performance ont un IVNC supérieur à 21 m<sup>3</sup>/km/j et 91 services ont un IVNC inférieur à 0,2 m<sup>3</sup>/km/j, ces services seront retirés de la base de données. La même méthode est appliquée au VNC/N :

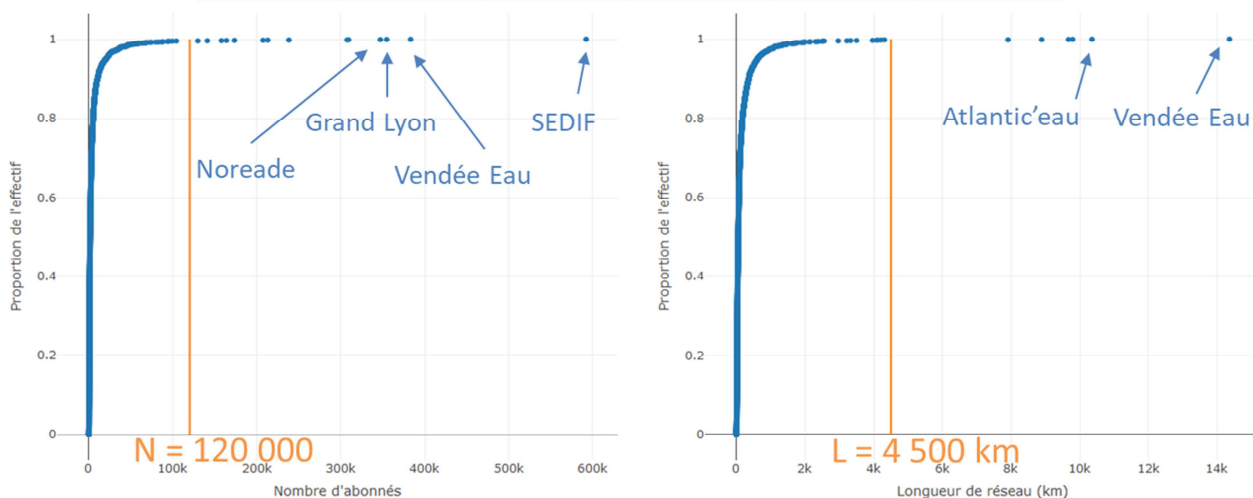
Quantile	10 %	25 %	Médiane	75 %	90 %
IVNC (m <sup>3</sup> /an/N)	10,9	21,0	35,8	58,8	94,9



**Figure 3 : Tri des services d'après le paramètre VNC/N de la base de données patrimoine-performance**

74 services ayant un VNC/N supérieur à 260 m<sup>3</sup>/an/N et 59 services ayant un VNC/N inférieur à 2,5 m<sup>3</sup>/an/N seront ainsi écartés. La même méthode est appliquée au nombre d'abonnés et à la longueur de réseau afin de supprimer les plus gros services qui ont un grand impact sur la modélisation :

Quantile	10 %	25 %	Médiane	75 %	90 %
N	102	215	571	1857	5319
L (km)	4,2	9,3	23,3	64,5	185,9



**Figure 4 : Paliers de sélection des services sur les paramètres N et L**

14 services ayant un nombre d'abonnés supérieur à 120 000 ou une longueur de réseau supérieure à 4500 km seront supprimés. Après l'application des 6 critères, 313 services sont écartés (certains services ne respectent pas plusieurs critères), la base de données patrimoine-performance contient donc finalement **5708 services, soit 484 579 km** de longueur de réseaux d'eau potable.

### 2.4.3. Représentativité de l'échantillon étudié et création de sous-échantillons

Sélectionner les services pour lesquels les données de performance et de patrimoine sont renseignées entraîne une sélection des plus « gros services ». A titre d'exemple, au sein de la base SISPEA 2016, 50 % des services renseignent la longueur de réseau, or la somme des longueurs renseignées est égale à 690 000 km soit près de 80 % de la longueur estimée au niveau national (875 000 km).

Néanmoins, la base de données contient une grande majorité de « petits » services et une faible proportion de « gros » services (comme le montre la Figure 4). Il apparaît dans de nombreux résultats que les services de grande taille se distinguent des services de petite taille, de par leur patrimoine, leur performance, leur contexte et les relations entre ces paramètres.

C'est pourquoi l'échantillon est divisé en quatre sous-échantillons contenant le même nombre d'individus, créés selon la longueur de réseau des services. Le quartile contenant les plus gros services (Q4) représente à lui-seul 80,5 % du VNC total alors que le quartile contenant les plus petits services (Q1) représente seulement 0,2 % du VNC total. La Figure 5 représente de manière schématique les quatre quartiles et les échantillons associés :

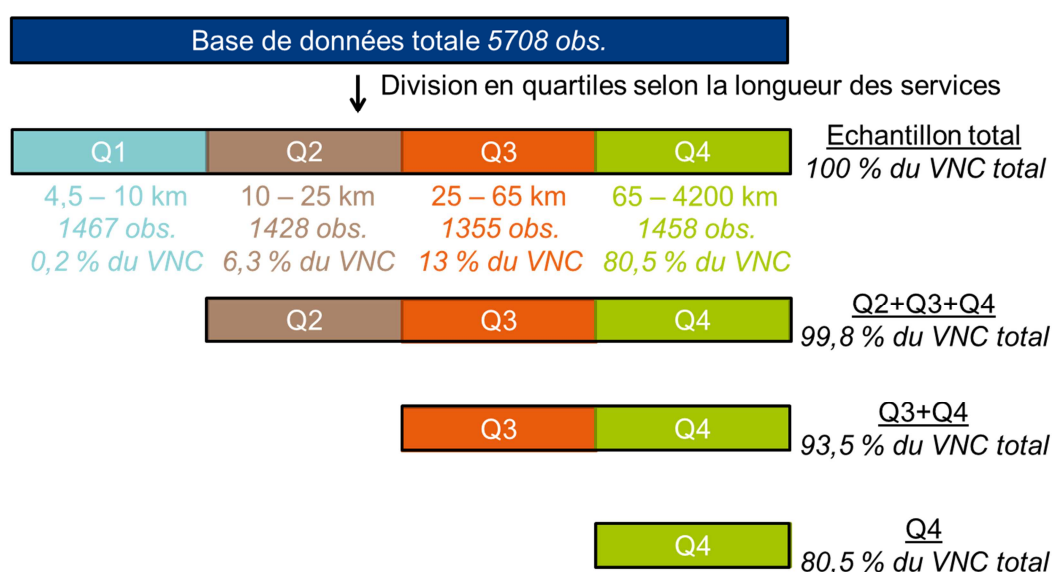


Figure 5 : Représentation schématique de la création des 4 sous-échantillons

Dans la suite de ce rapport, les 4 échantillons présentés seront parfois étudiés séparément.

## 3. Corrélations entre les paramètres de patrimoine et de performance

### 3.1. Corrélations entre l'IVNC et le patrimoine

Les corrélations linéaires entre le patrimoine et la performance sont étudiées grâce à la fonction  $\text{cor}()$  de R. Les corrélations présentées dans les tableaux suivants sont calculées entre les proportions de longueurs par matériau, par diamètre et par période de pose et l'IVNC. Par exemple, la colonne « PVC » correspond à la corrélation linéaire entre la variable  $\frac{L_{\text{PVC}}}{L}$  et l'IVNC, pour chaque échantillon. Les coefficients de corrélation entre les matériaux et l'IVNC sont présentés dans le Tableau 2.

Tableau 2 : Coefficients de corrélation entre les proportions de matériaux et l'IVNC

	Matériaux ( $L/L_{\text{tot}}$ (%))					
	PVC	PVCV	PVCR	F	FG	FD
Echantillon total	- 0,35	- 0,35	- 0,14	0,32	0,11	0,36
Quartiles 2,3 et 4	- 0,47	- 0,46	- 0,24	0,45	0,24	0,47
Quartiles 3 et 4	- 0,54	- 0,53	- 0,31	0,53	0,33	0,53
Quartile 4	- 0,63	- 0,61	- 0,38	0,61	0,41	0,60

Les coefficients calculés sur l'échantillon entier ne permettent pas de conclure. Mais ceux calculés sur les services les plus gros donnent un premier résultat : la proportion de PVC est corrélée négativement à l'IVNC et la proportion de fonte est corrélée positivement à l'IVNC, dans les services de plus de 65 km (Q4). Les coefficients de corrélation indiquent également que le PVCV est d'avantage corrélé à l'IVNC que le PVCR, mais on note aussi que le PVCV représente 66 % du PVC total, et on fait la même observation pour la FD qui représente 60 % de la fonte totale. Les mêmes calculs sont effectués sur les classes de diamètres (Tableau 3).

**Tableau 3 : Coefficients de corrélation entre les proportions de diamètres et l'IVNC**

	Diamètres (L/L <sub>tot</sub> (%))		
	≤ 50	60 - 100	≥ 125
Echantillon total	- 0,35	0,11	0,25
Quartiles 2,3 et 4	- 0,44	0,16	0,38
Quartiles 3 et 4	- 0,50	0,21	0,45
Quartile 4	- 0,58	0,29	0,54

De même que pour les matériaux, les résultats sont intéressants seulement sur les services de plus de 65 km. La proportion de diamètres inférieurs à 50 mm est corrélée négativement à l'IVNC alors que la proportion de diamètres supérieurs à 113 mm est corrélée positivement à l'IVNC. Il ne faut pas négliger cependant le fait que les coefficients de corrélation atteignent à peine 0,6. Les mêmes calculs sont effectués sur les périodes de pose (Tableau 4).

**Tableau 4 : Coefficients de corrélation entre les proportions de périodes de pose et l'IVNC**

	Périodes de pose (L/L <sub>tot</sub> (%))		
	≤ 1965	1965 - 1980	> 1980
Echantillon total	- 0,07	- 0,16	0,26
Quartiles 2,3 et 4	- 0,04	- 0,27	0,33
Quartiles 3 et 4	0,02	- 0,32	0,32
Quartile 4	0,09	- 0,37	0,31

Les coefficients de corrélation n'excédant pas 0,37, ils ne montrent pas de corrélation entre l'IVNC et les proportions de longueurs par périodes de pose.

Précisons que l'étude des corrélations entre le patrimoine et le paramètre de performance VNC/N montre qu'aucun des coefficients de corrélation n'excède 0,2, même sur le quatrième quartile. Ce paramètre semble donc indépendant du patrimoine des réseaux d'eau potable.

### 3.2. Corrélations entre les paramètres de patrimoine

**Les classes de matériaux et de diamètres sont corrélées à l'IVNC. Cependant, les matériaux sont aussi corrélés aux diamètres et aux périodes de pose, comme le montrent les Tableau 5 et**

Tableau 6.

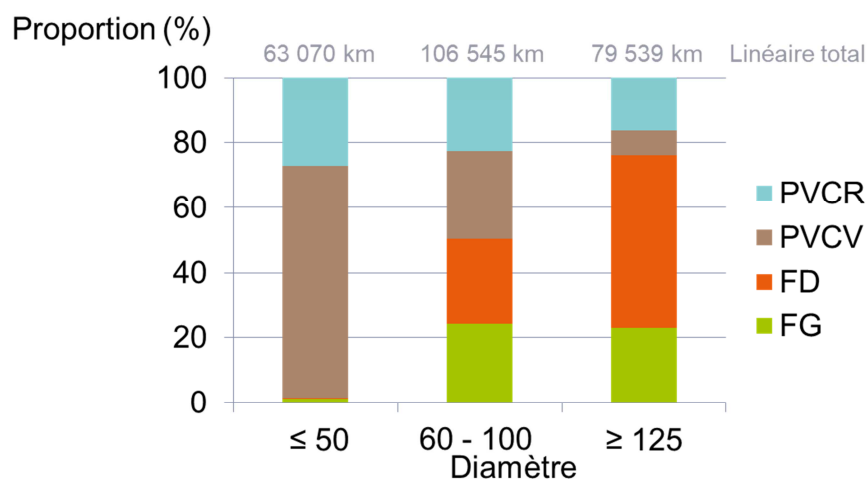
**Tableau 5 : Coefficients de corrélation entre les proportions de matériaux et de périodes de pose (échantillon total)**

		Matériaux (L/L <sub>tot</sub> (%))					
		PVC	PVCV	PVCR	F	FG	FD
Périodes de pose (L/L <sub>tot</sub> (%))	≤ 1965	- 0,29	- 0,12	- 0,39	0,24	<b>0,55</b>	- 0,21
	1965 - 1980	<b>0,58</b>	<b>0,65</b>	0,11	- 0,55	- 0,53	- 0,25
	> 1980	- 0,30	- <b>0,57</b>	0,33	0,32	- 0,07	0,52

**Tableau 6 : Coefficients de corrélation entre les proportions de matériaux et de diamètres (échantillon total)**

		Matériaux (L/L <sub>tot</sub> (%))					
		PVC	PVCV	PVCR	F	FG	FD
Diamètres (L/L <sub>tot</sub> (%))	≤ 50	<b>0,78</b>	<b>0,79</b>	0,29	<b>- 0,74</b>	- 0,51	- 0,54
	60 – 100	- 0,24	- 0,25	- 0,07	0,20	<b>0,38</b>	- 0,10
	≥ 125	- 0,56	-0,56	-0,22	<b>0,55</b>	0,18	<b>0,60</b>

Les périodes de pose antérieures à l'année 1965 correspondent aux fontes grises, les périodes de pose comprises entre 1965 et 1980 correspondent aux PVCV et les périodes de pose postérieures à 1980 correspondent en grande partie aux FD et en moindre partie aux PVCR. Les diamètres inférieurs à 50 mm correspondent aux PVC et les diamètres supérieurs à 125 mm correspondent majoritairement aux fontes. Ces résultats sont repris dans le graphique de la Figure 6. Ce graphique représente les proportions de chaque matériau dans chaque classe de diamètre, calculées sur les données collectées pour lesquelles on possède cette information (1615 observations).

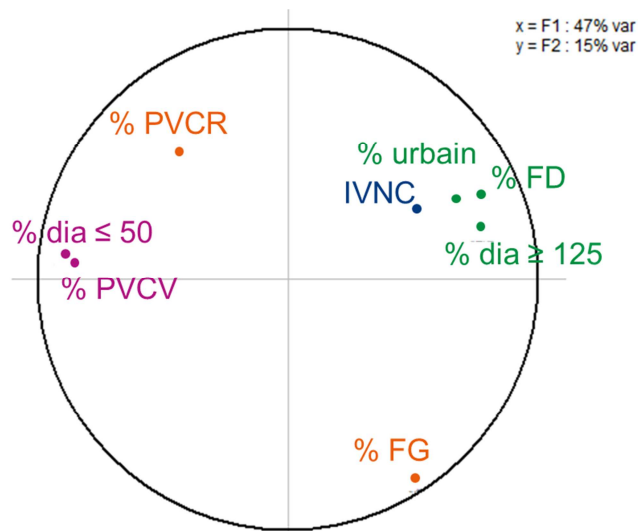


**Figure 6 : Part de matériaux dans chaque classe de diamètres calculées sur les données SIG collectées**

Ce graphique est intéressant car il montre que si l'on étudie les diamètres inférieurs à 50 mm, alors seul le matériau PVC est représenté et les diamètres supérieurs à 125 mm sont majoritairement représentés par la fonte. Si seuls les diamètres compris entre 60 et 125 mm sont considérés alors les proportions de chaque matériaux sont les mêmes.

### 3.3. Analyse en composante principale

Cette représentation permet de visualiser rapidement les corrélations linéaires entre les paramètres d'intérêt.



**Figure 7 : Représentation ACP des corrélations entre le patrimoine, la performance et le contexte**

La Figure 7 montre que les proportions de diamètres supérieurs à 100, les proportions de fonte ductile, les proportions de routes urbaines et l'IVNC sont très corrélées linéairement. Les proportions de diamètres inférieurs à 50 mm et les proportions de PVCV sont très corrélées entre elles et sont corrélées négativement à l'IVNC. Selon cette représentation, les proportions de matériau FG et PVCV ne sont pas corrélées linéairement à l'IVNC. Les matériaux et les diamètres étant liés aux paramètres de contexte, ces éléments doivent être pris en compte dans l'étude.

→ **Résultats de la partie 3 :**

- Les diamètres supérieurs à 100 et les FD sont corrélées positivement à l'IVNC ;
- Les diamètres inférieurs à 50 et les PVCV sont corrélés négativement à l'IVNC ;
- Le patrimoine est corrélé à l'IVNC mais pas au VNC/N ;
- Pas de résultats clairs concernant les périodes de pose.

#### 4. Liens entre le patrimoine, la performance et les paramètres de contexte

##### 4.1. Le mode de gestion

Le Tableau 7 reporte les moyennes et médianes des critères de performance IVNC et VNC/N pour les deux modes de gestion « délégation » et « régie ».

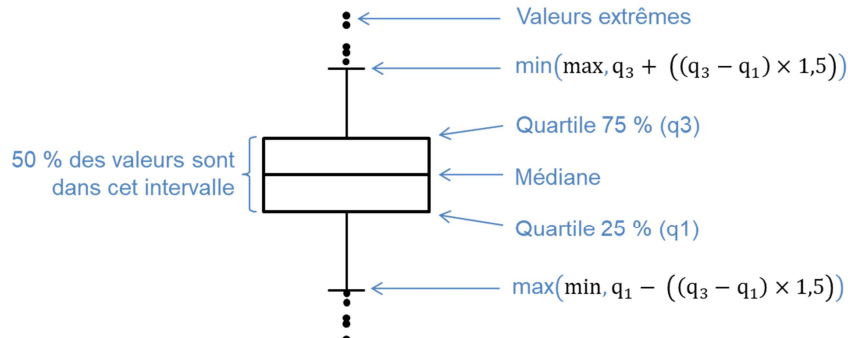
**Tableau 7 : Paramètres de performance en fonction du mode de gestion**

		Délégation (2023 obs.)	Régie (3636 obs.)
IVNC (m <sup>3</sup> /km/l)	Médiane	2.5	2.6
	Moyenne	3.6	3.9
VNC/N (m <sup>3</sup> /an/N)	Médiane	33.3	38.2
	Moyenne	38.8	49.1

Bien que la moyenne des VNC/N soit plus élevée en régie, la différence reste modérée et le mode de gestion ne sera pas un critère déterminant pour la suite de l'étude.

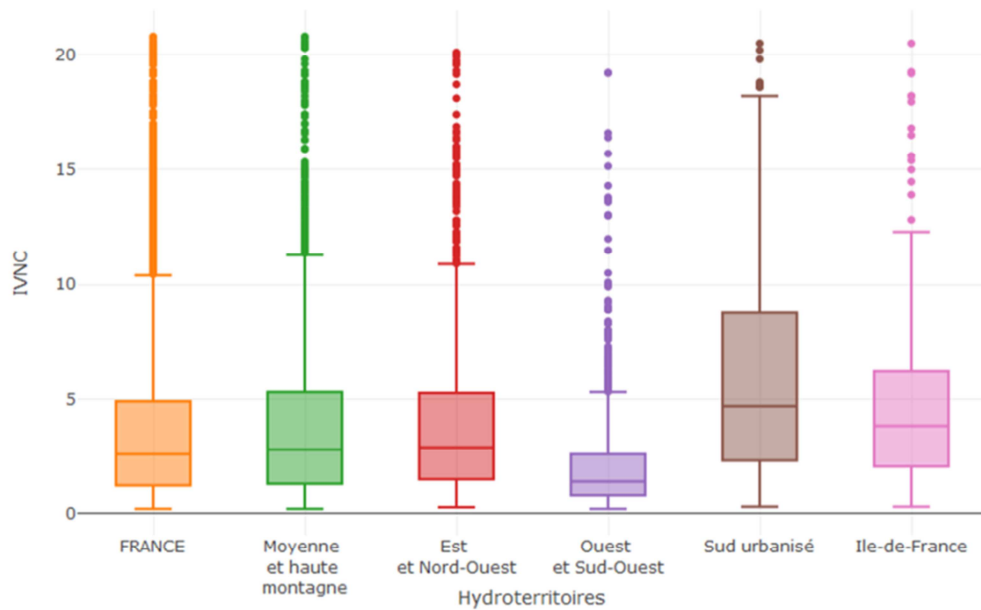
##### 4.2. Les hydroterritoires

Les boîtes à moustache présentées dans la suite du rapport doivent être lues comme indiqué sur la Figure 8 :



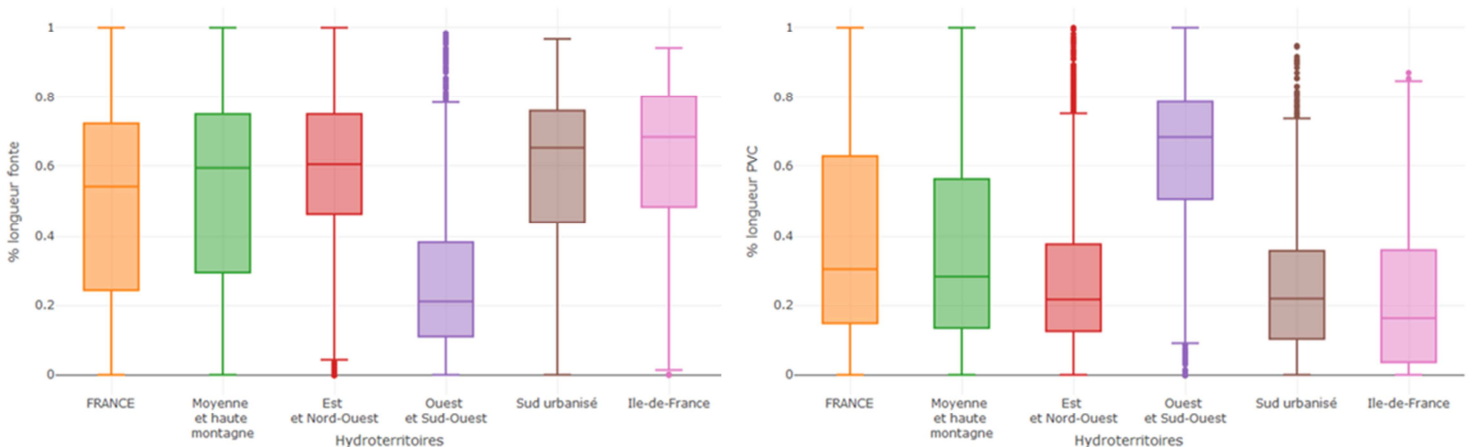
**Figure 8 : Interprétation d'une boîte à moustache**

On s'intéresse à la distribution des IVNC dans les 5 classes d'hydroterritoires. Les boîtes orange de la Figure 9 à la Figure 11 correspondent à la France entière et les autres boîtes aux hydroterritoires.



**Figure 9 : Distribution de l'IVNC (en  $m^3/km/j$ ) dans les 5 classes d'hydroterritoires**

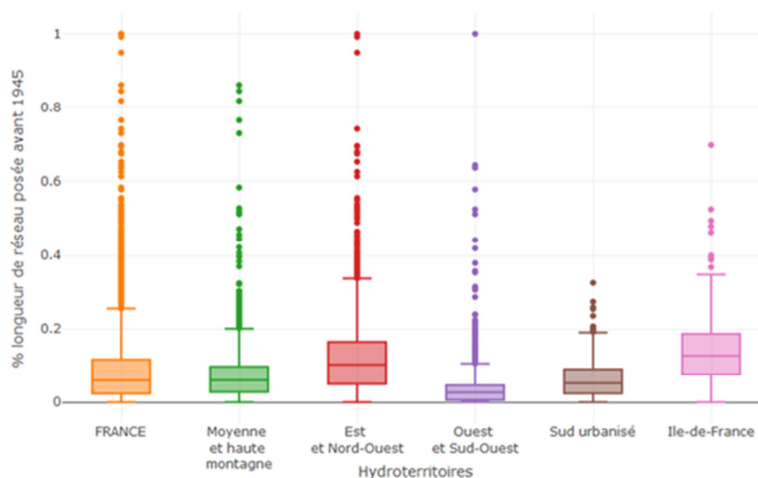
On constate que la médiane des IVNC dans les départements d'Ouest et du Sud-Ouest ( $1,4 m^3/km/j$ ) est plus faible que sur les autres hydroterritoires et la France entière où la médiane est de  $2,6 m^3/km/j$ . Notons que le critère VNC/N ne montre pas de variations selon les hydroterritoires. On étudie ensuite les proportions de fonte ( $\frac{L_F}{L}$ ) et de PVC ( $\frac{L_{PVC}}{L}$ ) selon les hydroterritoires (Figure 10).



**Figure 10 : Distribution des proportions de fonte et de PVC dans les 5 classes d'hydroterritoires**

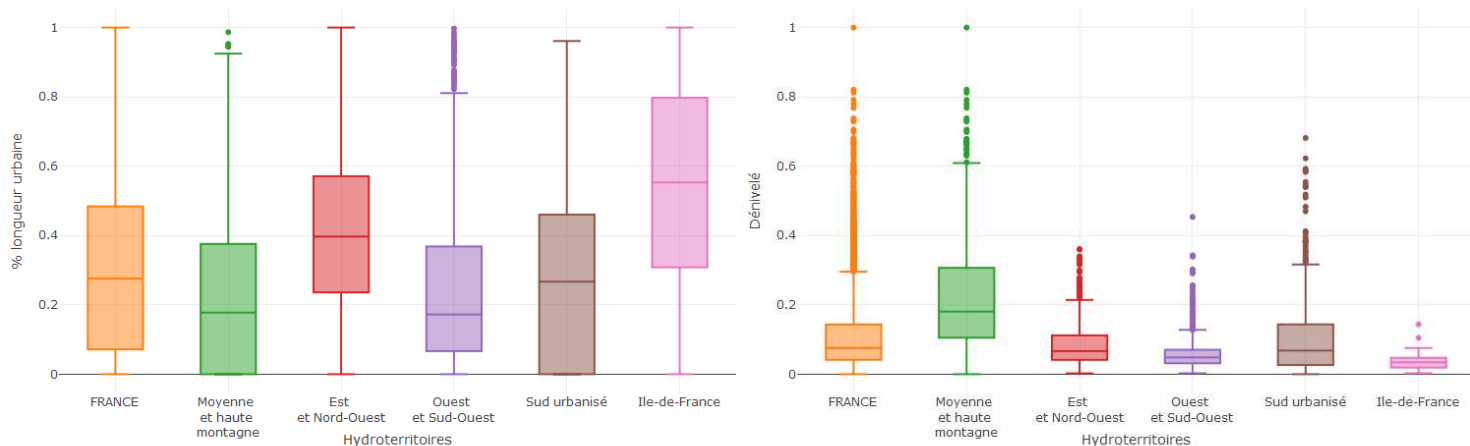


Les proportions de fonte sont bien plus faibles dans les départements d'Ouest et du Sud-Ouest (médiane à 20 % de fonte) et celles de PVC bien plus élevées (médiane à 68 % de PVC) que dans les autres hydroterritoires et sur la France entière (médianes à 54 % de fonte et 31 % de PVC). L'étude des distributions des proportions de période de pose antérieure à 1945 ( $\frac{L_{PP \leq 1945}}{L}$ ) montre aussi des disparités (Figure 11).



**Figure 11 : Distribution des proportions de longueurs posées avant 1945 dans les 5 classes d'hydroterritoires**

La Figure 12 montre les variations de contexte par hydroterritoire. A gauche, la proportion de routes urbaines et à droite, le dénivelé tel que défini dans la partie 2.3.



**Figure 12 : Distribution des proportions d'urbain et du dénivelé dans les 5 classes d'hydroterritoires**

La figure de gauche montre que les départements de l'Ouest et du Sud-Ouest sont moins urbanisés que le territoire national, mais c'est aussi le cas des départements en moyenne et haute montagne pour lesquels l'IVNC n'est pas plus faible que les valeurs nationales, mais pour lesquels le dénivelé est bien plus grand (figure de droite). On peut donc supposer un lien entre le faible taux d'urbanisation et les faibles IVNC, qui est atténué en zone de montagne par le fort dénivelé.

Les départements de l'Ouest et du Sud-Ouest se distinguent donc du reste de la France par des IVNC plus faibles, des plus grandes proportions de PVC et des plus faibles proportions de fonte, une plus faible proportion de périodes de pose antérieures à 1945 et une faible urbanisation.

Les mêmes tendances (IVNC, proportions de fonte et de PVC) sont observées si on s'intéresse aux agences, les médianes des IVNC des agence Loire-Bretagne et Adour-Garonne étant plus faibles que la médiane à l'échelle nationale.

### 4.3. Le taux d'urbanisation

Les données SIG collectées croisées avec les 4 classes CLC permettent d'attribuer à chaque commune une longueur de réseau par CLC. Deux nouvelles variables sont ainsi créées :

- La longueur de réseau urbain (regroupement des CLC « zone habitée » et « zone artificialisée hors zone habitée ») ;
- La longueur de réseau rural (regroupement des CLC « zone agricole » et « zone naturelle »).

Les paramètres « longueur de réseau urbain » et « longueur de réseau rural » ne sont pas disponibles pour les données de patrimoine modélisées, mais ont un coefficient de corrélation linéaire de 0,98 avec les paramètres « routes urbaines » et « routes rurales ». Ce sont donc ces deux paramètres qui sont utilisés dans la suite de l'étude.

La distribution des IVNC selon le caractère urbain ou rural des services est représentée sur la Figure 13. Les services possédant plus de 70 % de routes urbaines sont qualifiés d'urbains (520 services) et les services possédant plus de 70 % de routes rurales sont qualifiés de ruraux (3029 services).

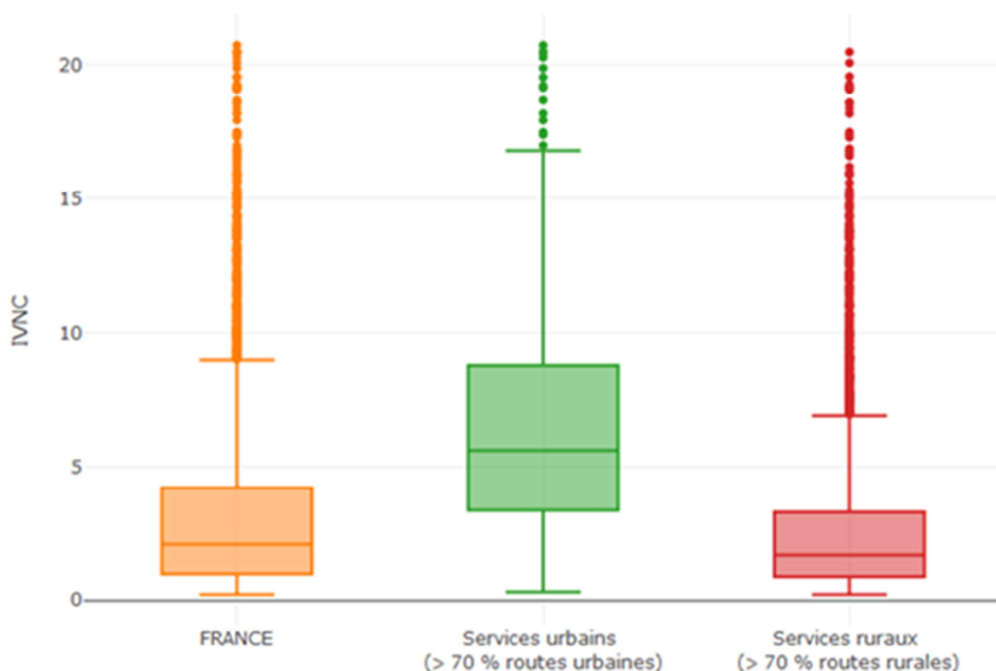
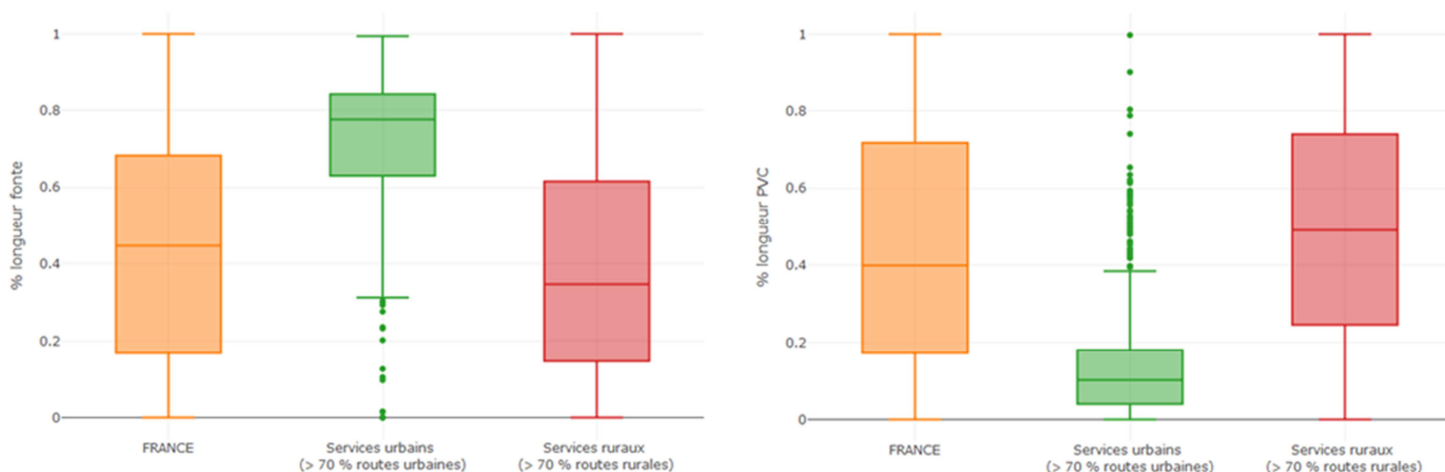


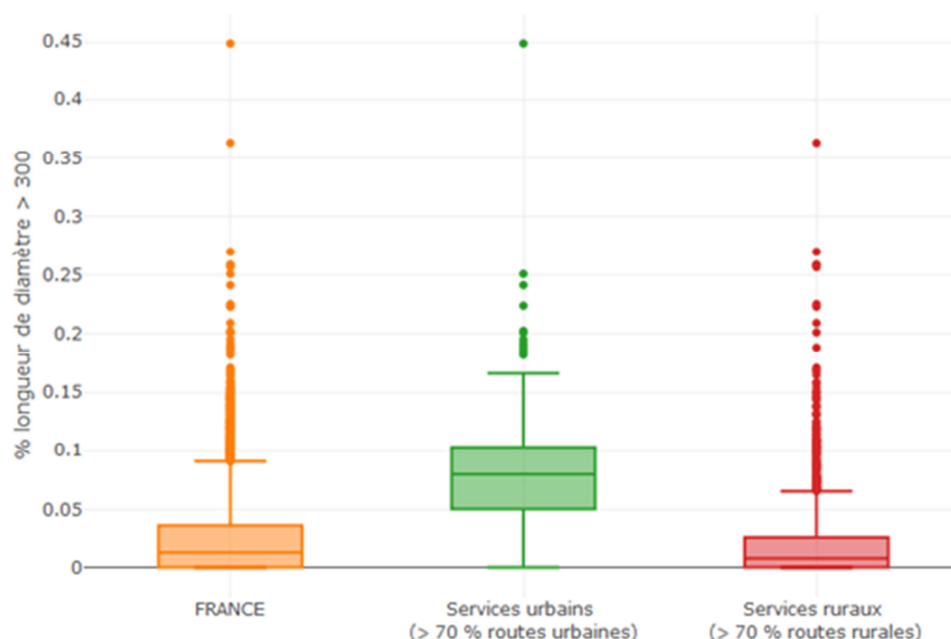
Figure 13 : Distribution des IVNC selon le caractère urbain ou rural

La médiane pour les services urbains ( $5,6 \text{ m}^3/\text{km}/\text{j}$ ) est bien supérieure à celle des services ruraux ( $1,7 \text{ m}^3/\text{km}/\text{j}$ ). Notons que la distribution du VNC/N ne dépend pas du caractère urbain ou rural. Les proportions de matériau fonte et PVC varient également selon le caractère urbain ou rural (Figure 14).



**Figure 14 : Distribution des proportions de fonte et de PVC selon le caractère urbain ou rural**

La proportion de fonte est beaucoup plus grande dans les services urbains et celle en PVC beaucoup plus faible. De plus, les proportions de gros diamètres dépendent aussi du caractère urbain ou rural (Figure 15).

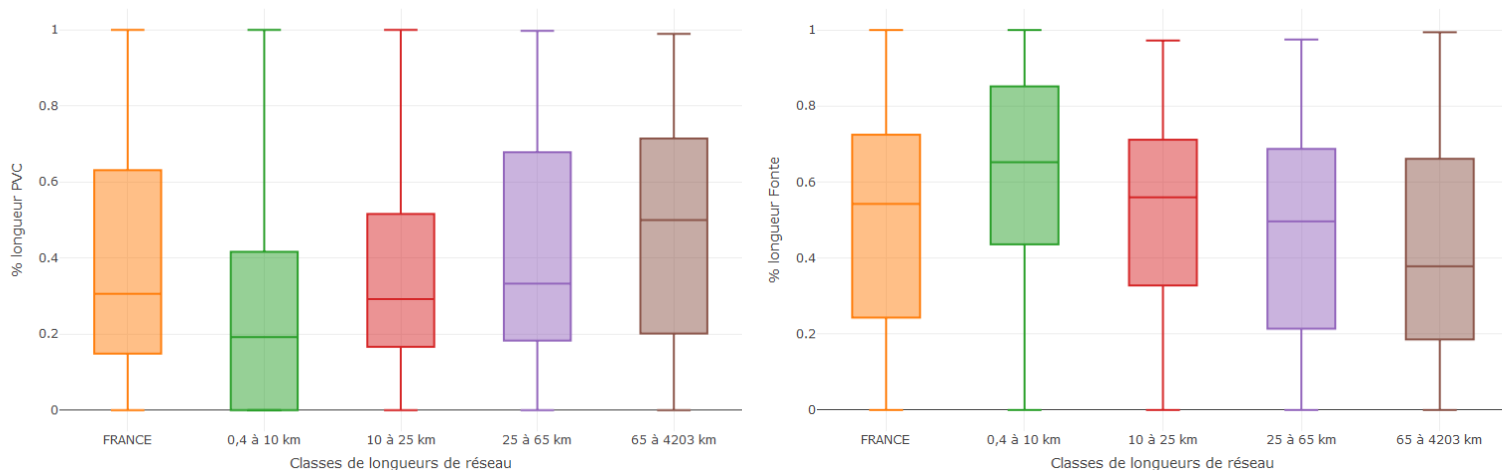


**Figure 15 : Distribution des proportions de diamètres supérieurs à 300 mm selon le caractère urbain ou rural**

Les services très urbanisés ont donc en commun des IVNC plus élevés que les services ruraux, une plus grande proportion de fonte et de diamètres supérieurs à 300 mm et une plus faible proportion de PVC.

#### 4.4. La taille du service

Les services de taille différente ont des patrimoines et des contextes différents. La distribution des proportions de fonte et de PVC est étudiée pour les 4 quartiles définis dans la partie 2.4.3. (Figure 16).



**Figure 16 : Distribution des proportions de fonte et de PVC selon la taille des services**

Plus la longueur de réseau est grande, plus la proportion de PVC est élevée et la proportion de fonte faible. La distribution des IVNC et des VNC/N est la même selon la taille des services, mais le VNC est un paramètre qui dépend de certains paramètres de taille du service (Tableau 8).

**Tableau 8 : Coefficients de corrélation entre la performance et les paramètres de contexte**

	Longueur	Population	Nombre	Surface	Routes	Routes
--	----------	------------	--------	---------	--------	--------

	totale		d'abonnés	urbaine	urbaines	rurales
VNC	0.64	0.82	0.81	0.80	0.82	0.46

Le VNC est d'avantage corrélé au nombre d'abonnés et à la population qu'à la longueur de réseau. Il est corrélé aux routes urbaines et surfaces urbaines mais peu corrélé aux routes rurales et surfaces rurales.

→ **Résultats de la partie 4 :**

- Le taux d'urbanisation est lié à l'IVNC et au patrimoine ;
- Le patrimoine et le contexte diffèrent avec les territoires (hydroterritoires ou agences) ;
- Les services ont un patrimoine différent selon leur taille ;
- Le VNC est plus corrélé au nombre d'abonnés qu'à la longueur de réseau.

## 5. Etude des courbes cumulées

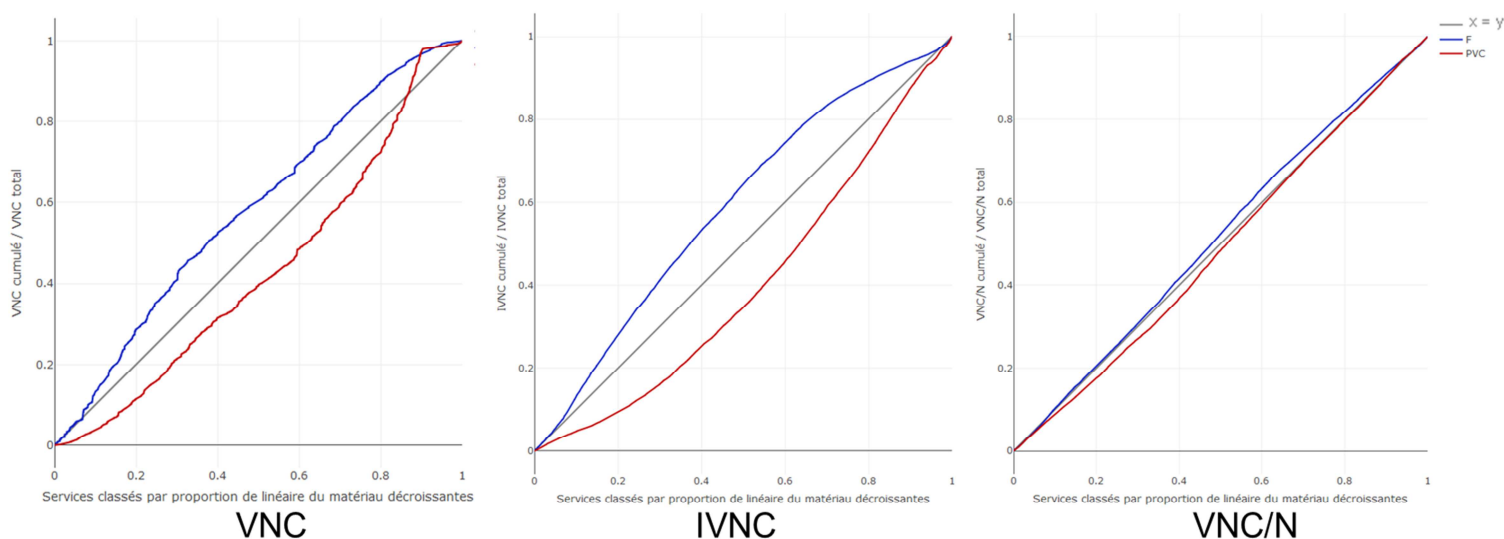
### 5.1. Cumuls après classement décroissant des proportions

#### 5.1.1. Principe de la méthode

Les courbes sont tracées de la manière suivante : les services sont classés par proportions décroissantes du premier matériau et le rang relatif du service est reporté en abscisses. En ordonnées, le cumul des VNC (ou IVNC ou VNC/N) est divisé par la somme totale des VNC (ou IVNC ou VNC/N). Les services sont ensuite reclassés par proportions décroissantes du deuxième matériau afin de tracer la deuxième courbe. On effectue ensuite la même démarche sur les diamètres.

#### 5.1.2. Résultats

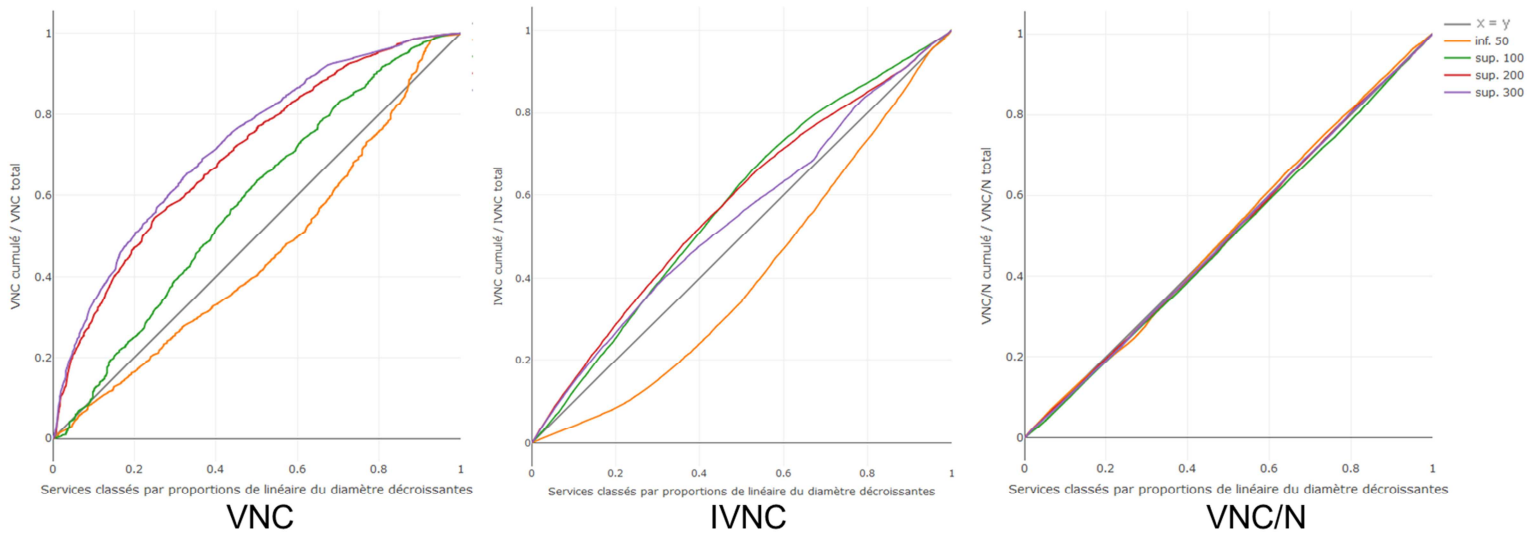
La Figure 17 représente les courbes pour les paramètres VNC, IVNC et VNC/N et pour les deux matériaux fonte (en bleu) et PVC (en rouge).



**Figure 17 : Courbes cumulées (classement décroissant) sur les matériaux fonte et PVC**

D'une part, les VNC et les IVNC s'accroissent de manière différente sur les grandes et les petites proportions de matériau et, d'autre part, ils s'accroissent de manière différente selon le matériau fonte ou PVC. Cependant, le matériau ne semble avoir qu'une très faible influence sur le paramètre VNC/N.

La Figure 18 représente les mêmes courbes obtenues pour les classes de diamètres, qui ont été détaillées par rapport aux classes utilisées dans les autres parties du rapport (orange :  $\leq 50$  mm, vert :  $\geq 100$  mm, rouge :  $\geq 200$  mm et violet :  $\geq 300$  mm).



**Figure 18 : Courbes cumulées (classement décroissant) sur les diamètres**

Les courbes « diamètres inférieurs à 50 mm » et « diamètres supérieurs à 100 mm » sont proches des courbes obtenues avec le PVC et la fonte respectivement pour les trois paramètres de performance. L'étude des diamètres croissants ( $\geq 100$  mm,  $\geq 200$  mm et  $\geq 300$  mm) montre que l'accumulation des VNC est de plus en plus accentuée avec l'augmentation de la taille des diamètres, ce qui n'est pas le cas pour l'accumulation des IVNC.

Ces courbes témoignent donc de l'importance de l'effet de la taille du réseau observée sur le VNC, moins visible sur l'IVNC, et qui n'est pas observé lorsqu'on s'intéresse au VNC/N.

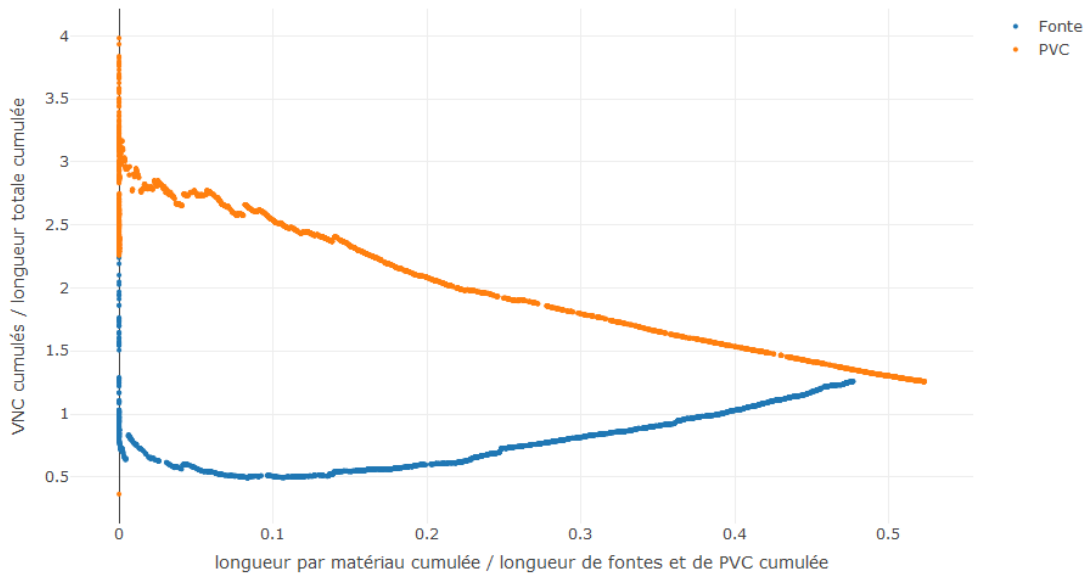
## 5.2. Cumuls après classement croissant des proportions

### 5.2.1. Principe de la méthode

Les courbes de cette partie sont tracées suivant la méthode ci-après : les services sont classés par proportions croissantes du premier matériau, les longueurs sont ensuite cumulées et divisées par la longueur totale et reportées en abscisses, les VNC sont cumulés et divisés par la longueur totale et reportés en ordonnées. Puis les services sont reclassés pour tracer la courbe du second matériau.

### 5.2.2. Résultats

Les courbes obtenues sont présentés sur la Figure 19.



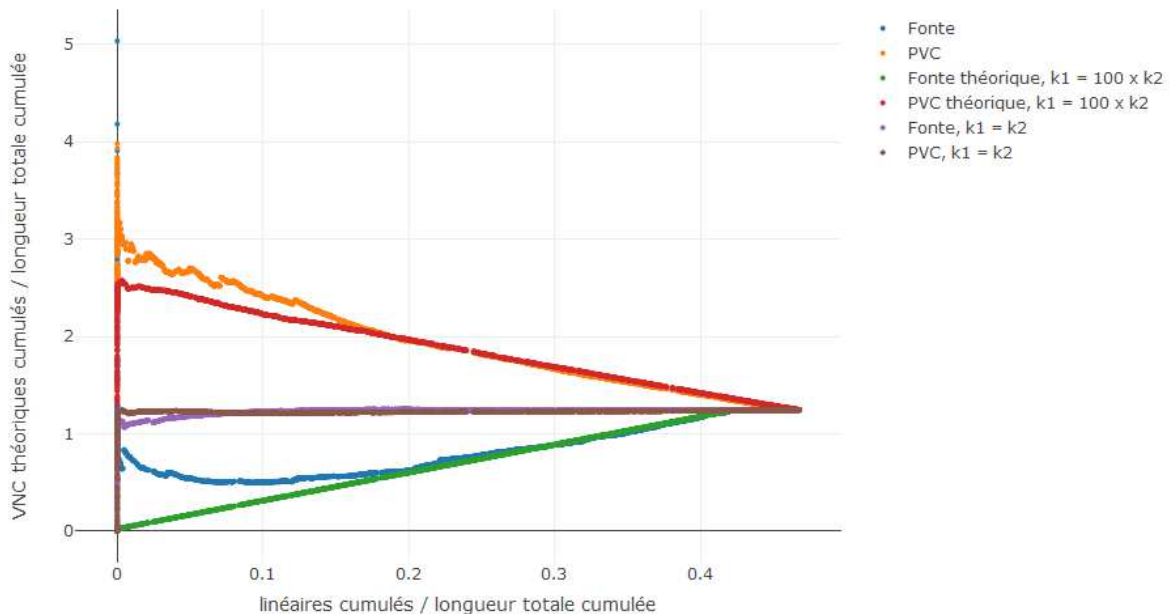
**Figure 19 : Courbes cumulées (classement croissant) des matériaux fonte et PVC**

Ces courbes permettent d'identifier deux tendances différentes : une droite croissante pour la fonte et une droite décroissante pour le PVC. L'objectif est de retrouver ces droites par expérimentation, en approximant un  $VNC_{théorique}$  s'exprimant :

$$VNC_{théorique} = k_F \times L_{Fonte} + k_{PVC} \times L_{PVC}$$

On pose  $k_F = \alpha \times k_{PVC}$ .

- Si  $\alpha = 1$  alors cela signifie que le matériau n'a pas d'influence sur le VNC et alors  $k_F = k_{PVC} = 1,25$  (courbes violette et marron sur la Figure 20) ;
- Si  $\alpha = 100$ , alors  $k_{PVC} = \frac{\sum VNC}{100 \times \sum L_{Fonte} + \sum L_{PVC}} = 0,03$  et  $k_F = 100 \times k_{PVC} = 2,9$  (courbes rouges et vertes sur la Figure 20).



**Figure 20 : Courbes cumulées (classement croissant) et expérimentation**

Les courbes constatées (bleu et orange) permettent d'affirmer que les matériaux PVC ou fonte ont une influence sur le VNC. Le cas  $k_F = 100 \times k_{PVC}$  permet de se rapprocher des données constatées, mais c'est aussi le cas de toutes les valeurs d' $\alpha > 100$ , on sait donc que  $k_F$  est supérieur à  $k_{PVC}$  mais il est impossible de déterminer graphiquement un  $\alpha$  optimal pour lequel la courbe théorique est proche de la courbe constatée.

→ **Résultats de la partie 5 :**

- Les VNC et les IVNC s'accroissent de manière différente en fonction du matériau (PVC ou fonte) pris en compte pour le classement des proportions, et le même résultat est observé avec les diamètres. Cependant, ces courbes ne permettent pas de s'affranchir de l'effet de la taille du service ;
- Les VNC/N s'accroissent indépendamment du matériau et des diamètres.

## 6. Modélisation de l'IVNC par classification

### 6.1. Méthode de classification ascendante hiérarchique

La méthode utilisée pour réaliser les classes est la Classification Ascendante Hiérarchique (CAH). La classification est ascendante car à la situation initiale, il y a autant de classes que d'individu, puis les individus sont regroupés selon leur similarités jusqu'à obtenir le nombre de classes désiré.

Dans le cas des matériaux, les classes sont réalisées selon leurs similarités de proportions de matériaux « fonte » et « PVC » et l'échantillon est séparé en 20 classes. Seuls les services possédant au moins 80 % de longueur en fonte ou en PVC cumulées sont conservés (les services possédant plus de 20 % de matériau « Autres » sont écartés), soit 5131 services conservés. Le dendrogramme résultant de la classification hiérarchique ascendante est représenté sur la Figure 21 :

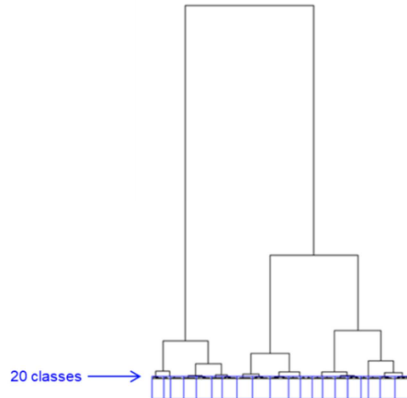


Figure 21 : Dendrogramme obtenu par CAH et représentation des 20 classes

L'IVNC et les proportions de fonte et de PVC sont recalculées pour les 20 classes. Pour l'étude des diamètres, la classification est faite selon les similarités de proportions de diamètres divisées en 3 groupes (inférieur à 50 mm, entre 60 et 100 mm, supérieur à 125 mm). Tous les services sont conservés et l'échantillon est regroupé en 20 classes.

### 6.2. Résultats sur les matériaux

La Figure 22 représente l'IVNC en fonction de la proportion de fonte. Sur le graphique de gauche, les points bleus correspondent à l'ensemble des services et les points orange aux 20 classes créées, les points orange sont reportés sur le graphique de droite.

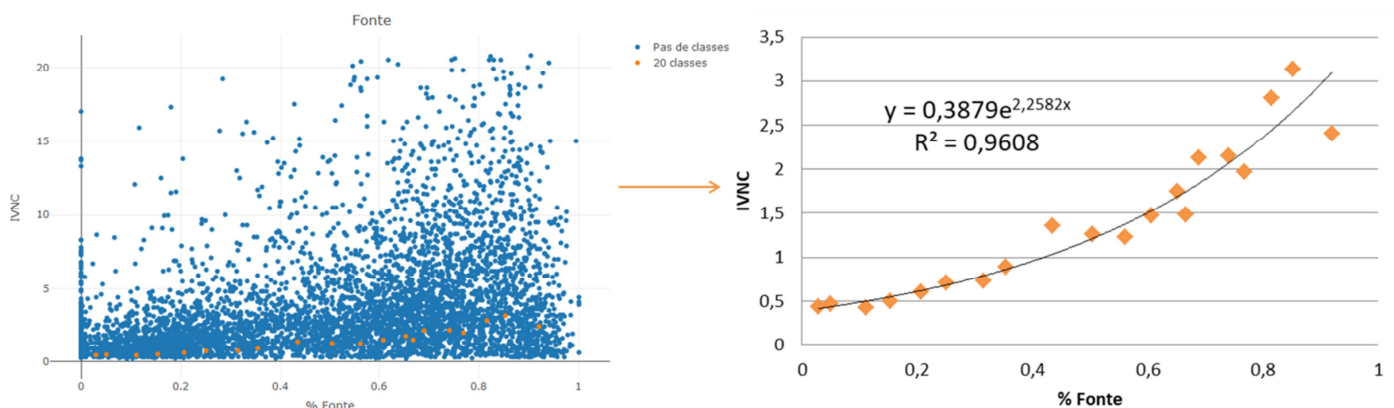
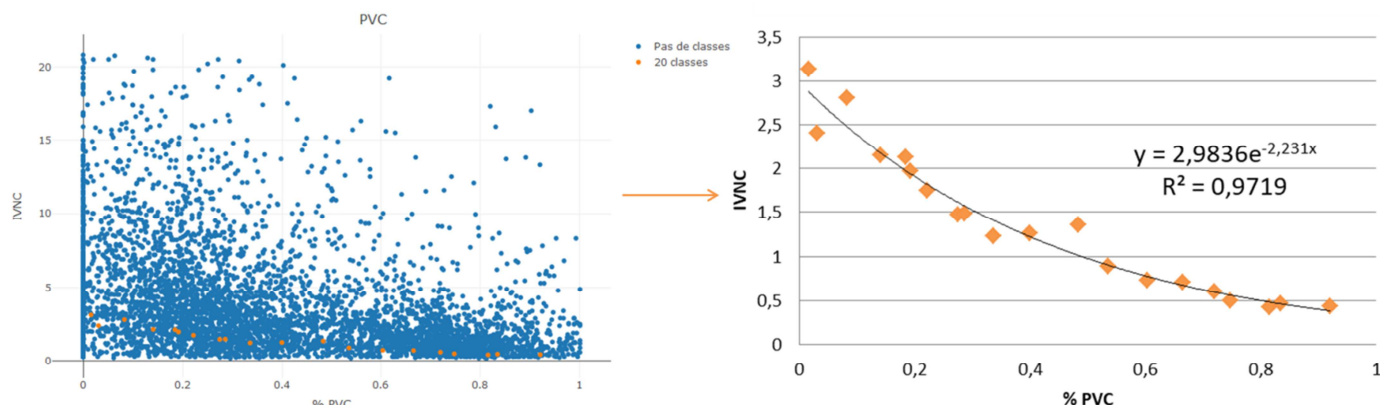


Figure 22 : IVNC en fonction de la proportion de fonte après CAH

Sans classes, aucune relation entre l'IVNC et la proportion de fonte n'est détectable. Avec 20 classes, l'IVNC est lié à la proportion de fonte par une relation exponentielle. Les mêmes graphiques sont représentés pour les proportions de PVC sur la Figure 23.

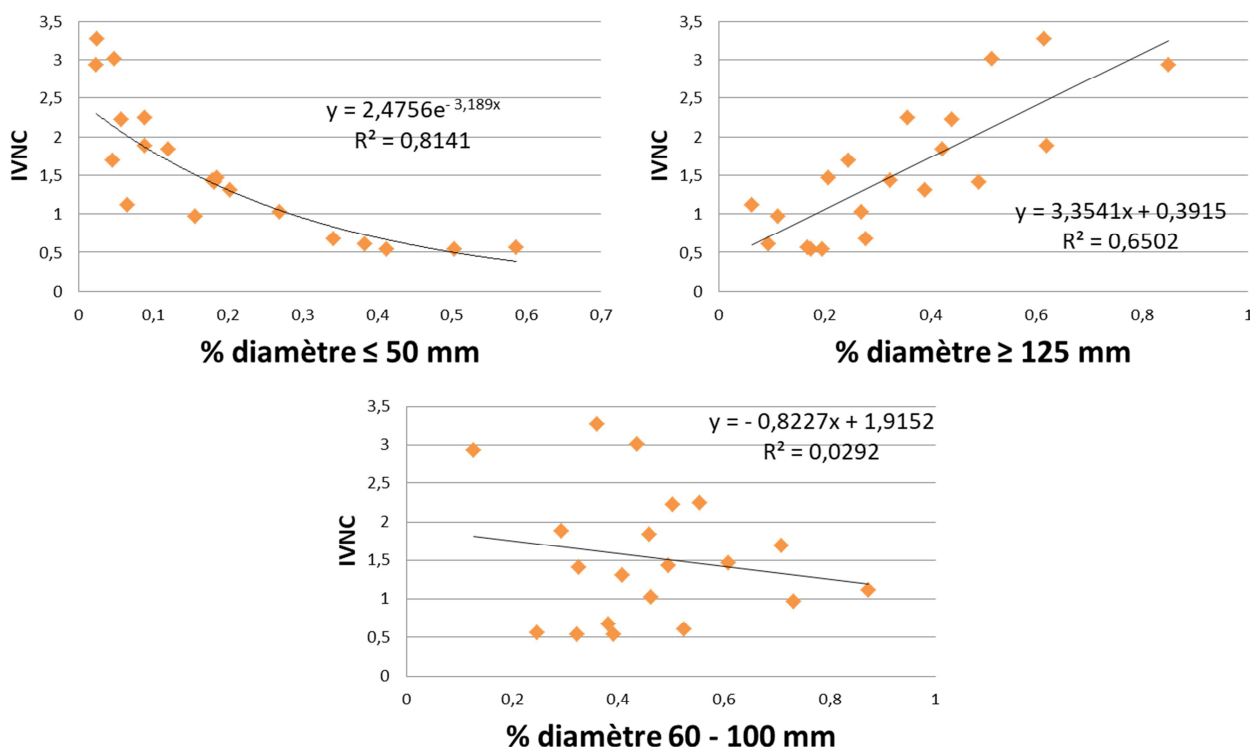


**Figure 23 : IVNC en fonction de la proportion de PVC après CAH**

L'IVNC est aussi lié à la proportion de PVC par une relation exponentielle qui est décroissante.

### 6.3. Résultats sur les diamètres

L'interprétation des résultats pour les diamètres est plus délicate. Les résultats sont reportés sur la Figure 24 pour les trois classes de diamètres.



**Figure 24 : IVNC en fonction des proportions de diamètres (<50, 50 à 100, >100) après CAH**

La relation exponentielle entre l'IVNC et les proportions de diamètres inférieurs à 50 mm est proche de celle obtenue pour les proportions de PVC, ce qui est cohérent sachant que les canalisations de diamètres inférieurs à 50 mm sont quasiment exclusivement en PVC. La relation entre les diamètres supérieurs à 125 mm et l'IVNC est croissante et linéaire, le fait que l'on ne retrouve pas la relation exponentielle comme dans le cas de la fonte peut s'expliquer par le fait que les diamètres supérieurs à 125 mm comprennent 75 % de fonte et 25 % de PVC. Il n'y a pas de relation entre les diamètres compris entre 60 et 100 mm et l'IVNC.

Connaissant ces résultats, il paraît judicieux de modéliser le logarithme de l'IVNC en fonction des proportions de matériaux, mais sans la création de classes, ces modèles fournissent une prédiction médiocre de l'IVNC. Le VNC/N n'étant que très faiblement corrélé au patrimoine, c'est le VNC qui est modélisé dans la partie suivante.



→ **Résultats de la partie 6 :**

- Le logarithme de l'IVNC dépend de la proportion de fonte selon une équation de la forme  $\ln(\text{IVNC}) = \alpha \times \%F + \beta$  et de la proportion de PVC selon une équation de la forme  $\ln(\text{IVNC}) = -\alpha \times \%PVC + \beta$  ;
- On retrouve ce résultat pour les diamètres inférieurs à 50 mm (qui sont quasi-exclusivement en PVC) mais c'est plus compliqué pour les autres classes de diamètres.

## 7. Modélisation du VNC

### 7.1. Les modèles linéaires

#### 7.1.1. Présentation des modèles

Le VNC est très corrélé au nombre d'abonnés du service et à sa longueur. Le coefficient de régression du modèle ci-dessous est de 0,77 (les longueurs sont exprimées en mètres).

$$\text{VNC} = 33,5 \times N + 0,1 \times L$$

Si la longueur est distinguée par matériau, le modèle peut être considérablement amélioré, comme présenté ci-dessous (coefficient de régression de 0,83).

$$\text{VNC} = 21,6 \times N + 0,7 \times L_{FG} + 2,0 \times L_{FD} + 0,02 \times L_{PVCV} - 0,7 \times L_{PVCR} + 0,6 \times L_{autres}$$

Le coefficient multiplié au PVCV est proche de zéro, celui de la FD est positif et celui du PVCR est négatif, ce qui est cohérent avec les coefficients de corrélation. Notons que le coefficient négatif multipliant la longueur de PVCR signifie que le VNC calculé peut théoriquement être négatif, ce qui est impossible en pratique. Ce coefficient négatif permet de mieux estimer le VNC d'un point de vue statistique mais n'a aucun sens « métier ».

La longueur peut aussi être distinguée par diamètres, mais le coefficient de régression est plus faible (coefficient de régression de 0,79).

$$\text{VNC} = 25,3 \times N - 0,6 \times L_{\text{dia}[\leq 50]} + 0,6 \times L_{\text{dia}[50-125]} + 1,0 \times L_{\text{dia}[\geq 125]}$$

Le coefficient négatif sur les diamètres inférieurs à 50 mm et positif sur les diamètres supérieurs à 125 mm est cohérent avec le modèle précédent et les coefficients de corrélation.

Ces modèles sont également calés sur les données collectées seules (environ 1000 services, les services possédant plus de 80 % d'inconnus sur les matériaux ou les diamètres sont écartés) et permettent de valider les modèles précédents. Sur ces données, il est possible de distinguer les longueurs selon le matériau par diamètre. Les diamètres inférieurs à 50 ne sont pas distingués par matériaux car ils correspondent majoritairement aux PVC et de même pour les diamètres supérieurs à 125 qui correspondent majoritairement aux fontes.

$$\begin{aligned} \text{VNC} = & 20,1 \times N + 0,1 \times L_{\text{dia}[\leq 50]} \\ & + 0,9 \times L_{FG,\text{dia}[60-100]} + 1,4 \times L_{FD,\text{dia}[60-100]} + 1,8 \times L_{FI,\text{dia}[60-100]} \\ & - 0,7 \times L_{PVCV,\text{dia}[60-100]} - 1,3 \times L_{PVCR,\text{dia}[60-100]} - 0,2 \times L_{PVCI,\text{dia}[60-100]} \\ & + 0,9 \times L_{\text{dia}[\geq 125]} + 1,7 L_{\text{dia inconnu}} \end{aligned}$$

Ce modèle a un coefficient de régression de 0,81. Il est intéressant car il permet d'étudier le matériau sans influence du diamètre mais il s'appuie sur un échantillon restreint (1152 observations).

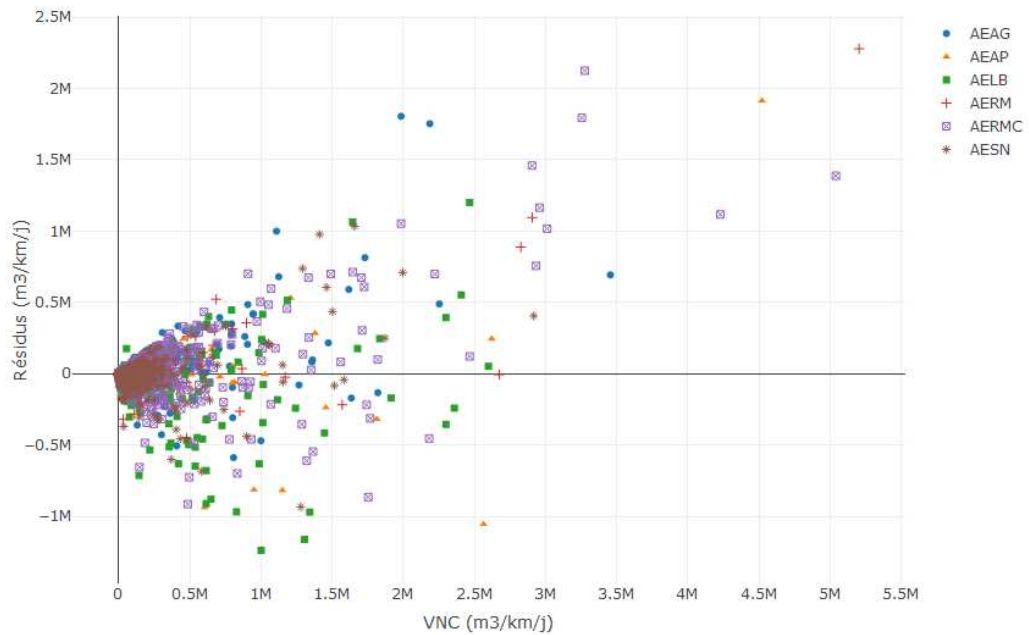
#### 7.1.2. Observations sur les modèles

Dans cette partie les illustrations correspondent au modèle ci-dessous, présenté dans la partie précédente :

$$\text{VNC} = 21,6 \times N + 0,7 \times L_{FG} + 2,0 \times L_{FD} + 0,02 \times L_{PVCV} - 0,7 \times L_{PVCR} + 0,6 \times L_{autres}$$

Pour caler les modèles, on utilise les outils suivants :

- **Le graphique des résidus ( $VNC_{\text{observé}} - VNC_{\text{prédit}}$ ) en fonction des  $VNC_{\text{observé}}$  :**

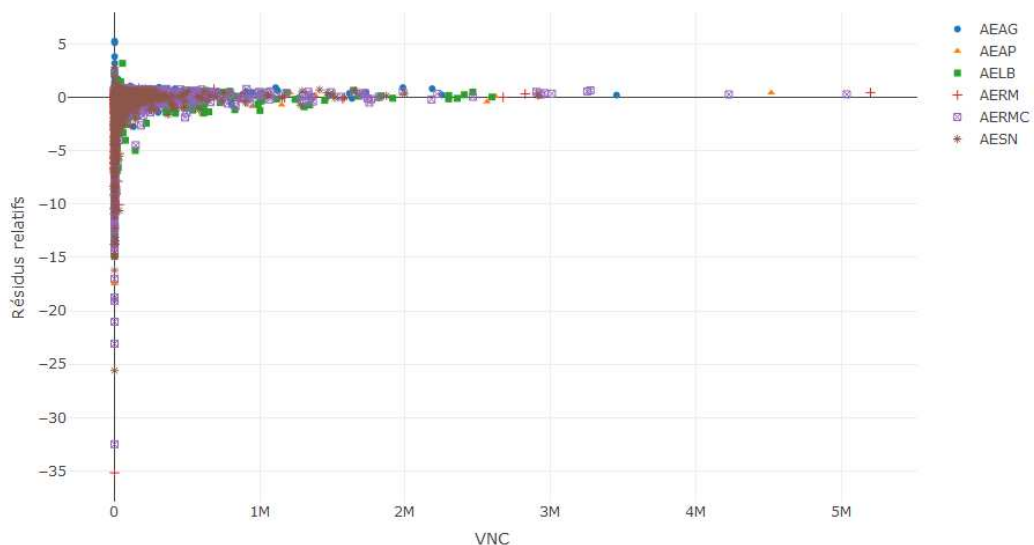


**Figure 25 : Résidus en fonction des VNC observés**

Les points sont distingués selon les agences. La Figure 25 montre que plus les VNC sont grands, plus les résidus sont grands. La variance des résidus dépend donc de la variable modélisée, ce qui n'est pas pris en compte dans le modèle linéaire (on parle d'hypothèse d'homoscédasticité dans le cas d'un modèle linéaire, qui n'est pas respectée ici). On ne peut pas modifier la variance des résidus : plus le VNC modélisé est grand, plus l'erreur est grande, mais on peut introduire un terme dans le modèle qui prend en compte la variance des résidus.

- **Les résidus relatifs  $\left(\frac{VNC_{\text{observé}} - VNC_{\text{prédit}}}{VNC_{\text{observé}}}\right)$  :**

Un résidu relatif égal à 0 signifie que le VNC prédit est égal au VNC observé et un résidu relatif égal à 1 signifie que l'erreur est de 100 %.

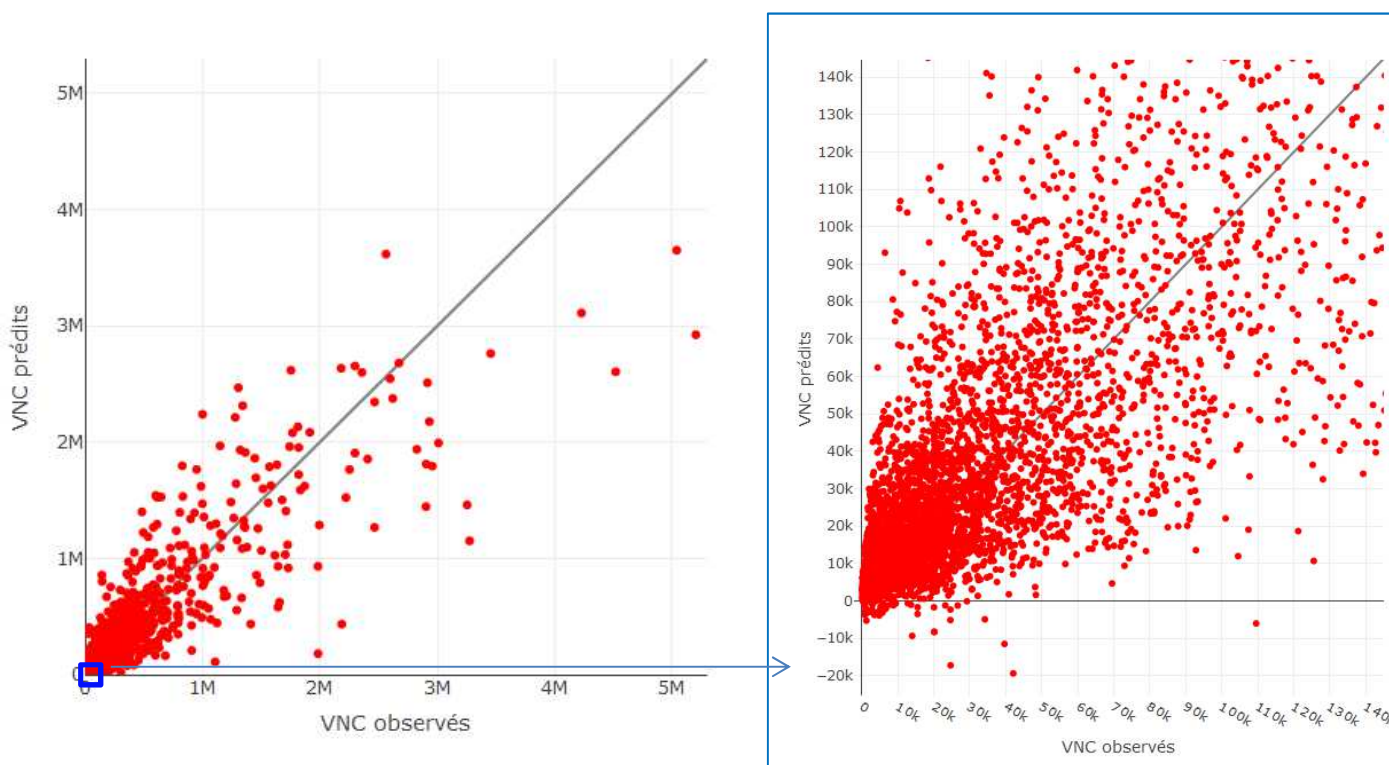


**Figure 26 : Résidus relatifs en fonction du VNC observé**

L'étude des résidus relatifs nous apprend que 50 % des valeurs prédites ont un résidu relatif entre -0,5 et 0,5, soit une erreur inférieure à 50 %, et 25 % des valeurs prédites ont une erreur de plus de 100 % (résidu relatif

inférieur à  $-1$  ou supérieur à  $1$ ). Ces valeurs peuvent être améliorées. De plus, cette représentation permet de détecter des comportements différents selon les agences lorsqu'il y en a.

– **Les valeurs prédites en fonction des valeurs observées :**



**Figure 27 : VNC prédits en fonction des VNC observés**

La Figure 27 permet d'observer que certains VNC prédits sont négatifs, de plus, les prédictions ne sont pas très précises et le modèle peut être amélioré.

### **7.1.3. Validation croisée par Monte-Carlo**

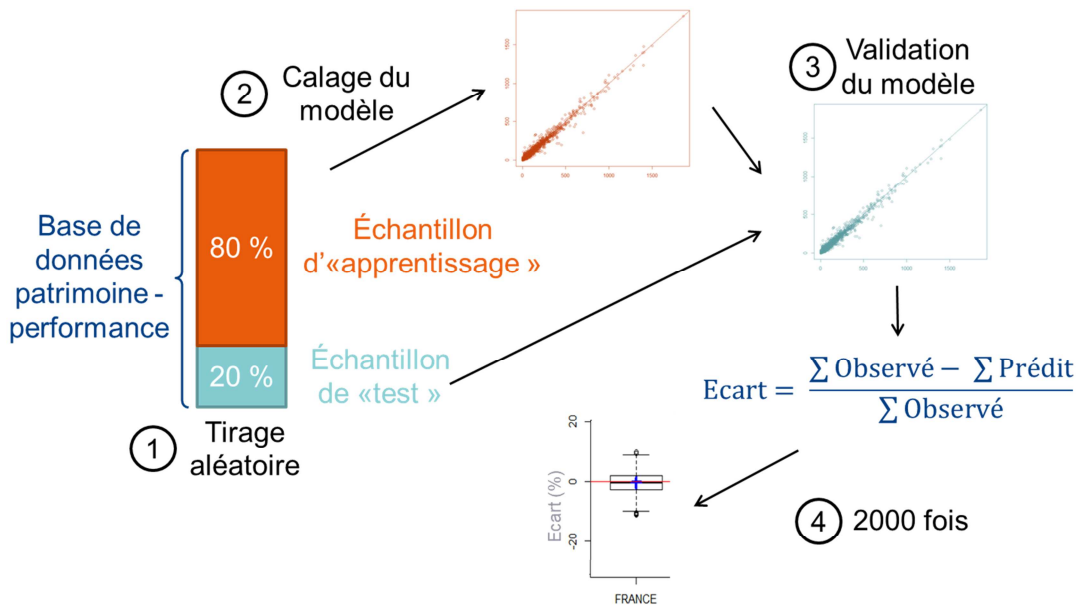
#### **7.1.3.1. Principe de la méthode de validation croisée**

Pour vérifier la robustesse des modèles, la méthode utilisée est celle de Monte-Carlo qui permet d'évaluer la stabilité de l'estimation à partir d'un nombre important de tirages aléatoires.

A partir de la base initiale, une succession de tirages aléatoires est effectuée. A chaque tirage, la base est partagée en deux échantillons, un échantillon d'« apprentissage » et un échantillon « test ». Le modèle est calé sur l'échantillon d'apprentissage puis testé sur l'échantillon test. L'écart relatif entre le volume connu et le volume prédit est ensuite calculé, à l'échelle nationale et à des sous-échelles (agences de l'eau et quartiles selon la longueur de réseau). Les étapes de la procédure sont décrites ci-dessous :

- Tirage aléatoire d'un échantillon d'apprentissage (80 % de la base de données sur laquelle a été calé le modèle) ;
- Calage du modèle sur l'échantillon d'apprentissage puis extrapolation sur les 20 % de la base restant ;
- Calcul de l'écart relatif (plus celui-ci est proche de zéro plus le modèle est pertinent) ;
- Les trois premières étapes sont répétées 2000 fois.

La figure ci-dessous schématise ces étapes :

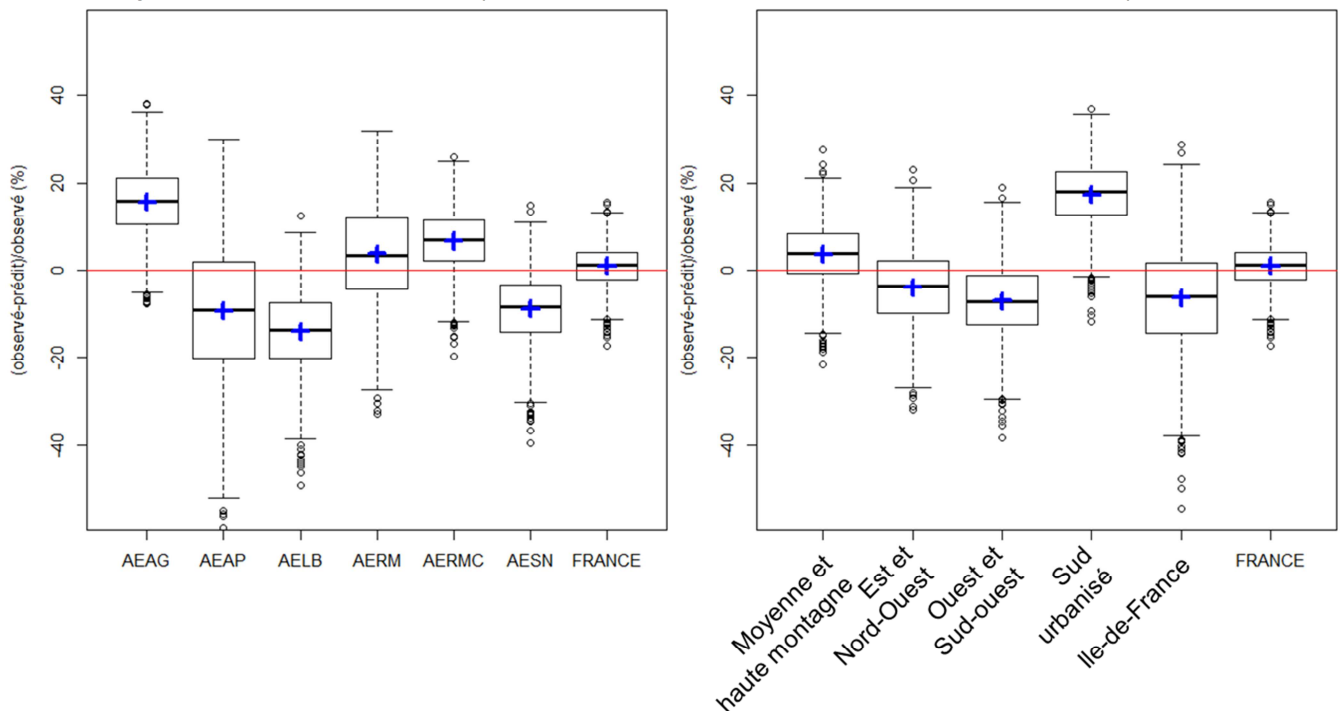


**Figure 28 : Schématisation de la validation croisée par la méthode de Monte-Carlo**

Les résultats sont synthétisés au sein de boîtes à moustache, à l'échelle nationale, par agences et par hydroterritoires.

### 7.1.3.2. Résultats

La Figure 29 présente les résultats de la validation croisée pour le modèle de la partie précédente, à gauche selon les agences et à droite selon les hydroterritoires. Les croix bleues correspondent à la moyenne.



**Figure 29 : Validation croisée par la méthode Monte-Carlo du modèle linéaire**

Le test indique qu'à l'échelle nationale, les VNC sont estimés de manière suffisamment précise mais à l'échelle des agences, certaines sont systématiquement surestimées (Artois-Picardie, Loire-Bretagne, Seine-Normandie) et d'autres sont sous-estimées (Adour-Garonne, Rhône-Méditerranée-Corse). L'apport d'éléments de contexte au modèle devrait permettre de l'affiner.

## 7.2. Les modèles « exponentiels »

### 7.2.1. Modèle exponentiel utilisant les longueurs par matériaux

Le premier modèle exponentiel correspond au modèle linéaire présenté précédemment auquel on multiplie un terme exponentiel permettant d'introduire les paramètres de contexte.

$$VNC = (81,3 \times N + 5,8 \times L_{FG} + 8,2 \times L_{FD} + 2,5 \times L_{PVCV} + 1,7 \times L_{PVCR} + 7,3 \times L_{autres}) \times e^{\left( \begin{array}{l} -2,5 \times AELB - 2,2 \times AEAG - 2,3 \times AE(AP-SN) - 2,0 \times AERM - 2,0 \times AERMC \\ + 0,9 \times \frac{Routes_{urbaines}}{Routes} + 0,2 \times Dénivelé_{[0,1]} \end{array} \right)}$$

Les termes de contexte introduits qui rendent le modèle plus proche de la réalité sont :

- Les agences de l'eau avec un regroupement des agences Artois-Picardie et Seine-Normandie, permettant de prendre en compte les disparités régionales ;
- La proportion de routes urbaines qui permet d'apporter une correction d'autant plus grande que le service est urbain ;
- Le dénivelé (variable définie dans la partie 2.3) qui apporte une correction d'autant plus grande que le dénivelé est élevé.

Ce modèle présente plusieurs faiblesses :

- La validation croisée n'est pas satisfaisante : les boîtes à moustache sont centrées autour de zéro pour toutes les agences sauf l'agence Rhin-Meuse qui est très surestimée ;
- Lorsqu'on cale ce modèle sur les données collectées, le coefficient sur la fonte ductile est beaucoup plus faible que celui sur la fonte grise, les coefficients sont donc peu stables.

De plus, ce modèle introduit les paramètres de patrimoine dépendamment de leur taille, on exprime donc un deuxième modèle pour lequel les variables de patrimoine sont indépendantes de la longueur de réseau.

### 7.2.2. Modèle exponentiel utilisant les proportions par matériaux

Le deuxième modèle exponentiel utilise la très forte corrélation entre le volume et le nombre d'abonnés, qui est ensuite modulé par les paramètres de patrimoine (proportions de matériaux FG, FD, PVCV et PVCR) et de contexte insérés dans l'exponentielle.

$$VNC = 38,7 \times N \times e^{\left( \begin{array}{l} -0,2 \times \frac{L_{FG}}{L_{tot}} - 0,05 \times \frac{L_{FD}}{L_{tot}} + 0,5 \times \frac{L_{PVCV}}{L_{tot}} - 1,3 \times \frac{L_{PVCR}}{L_{tot}} \\ - 1,0 \times AELB + 0,3 \times AEAG + 0,2 \times AE(AP-SN) + 0,6 \times AERM + 0,4 \times AERMC \\ - 0,2 \times \frac{Routes_{urbaines}}{Routes} + 0,2 \times Dénivelé_{[0,1]} \end{array} \right)}$$

Les termes de contexte introduits sont les mêmes que dans le modèle précédent. On remarque deux différences notables avec le modèle précédent :

- Le coefficient sur le PVCV est positif alors que les coefficients sur les autres matériaux sont négatifs ;
- Le coefficient sur le terme traduisant la proportion d'urbain est négatif alors qu'il était positif dans le modèle précédent. L'effet de l'urbanisation sur le VNC se traduit différemment s'il est appliqué aux abonnés ou aux longueurs par matériau.

Ce modèle présente plusieurs faiblesses :

- La validation croisée n'est toujours pas satisfaisante : l'agence Rhin-Meuse est surestimée et l'agence Rhône-Méditerranée-Corse est sous-estimée, les plus petits services (premier quartile) sont très surestimés ;
- Les éléments de contexte « proportion d'urbain » et « dénivelé » n'ont qu'un très faible effet sur le modèle. Plusieurs « bons » modèles utilisant des paramètres de contexte différents sortent du lot mais aucun ne satisfait la condition de boîtes centrées autour de zéro pour toutes les agences après la validation croisée.

Les modèles exponentiels étant plus difficiles à caler que les modèles linéaires, le modèle linéaire présenté dans la partie suivante a été d'avantage développé que les modèles exponentiels.

## 7.3. Le modèle linéaire avec modulation sur les abonnés

### 7.3.1. Présentation du modèle

Pour ce modèle, six nouvelles classes de territoires sont créées à partir des hydroterritoires et des agences. Ces six classes sont représentées sur la Figure 30.

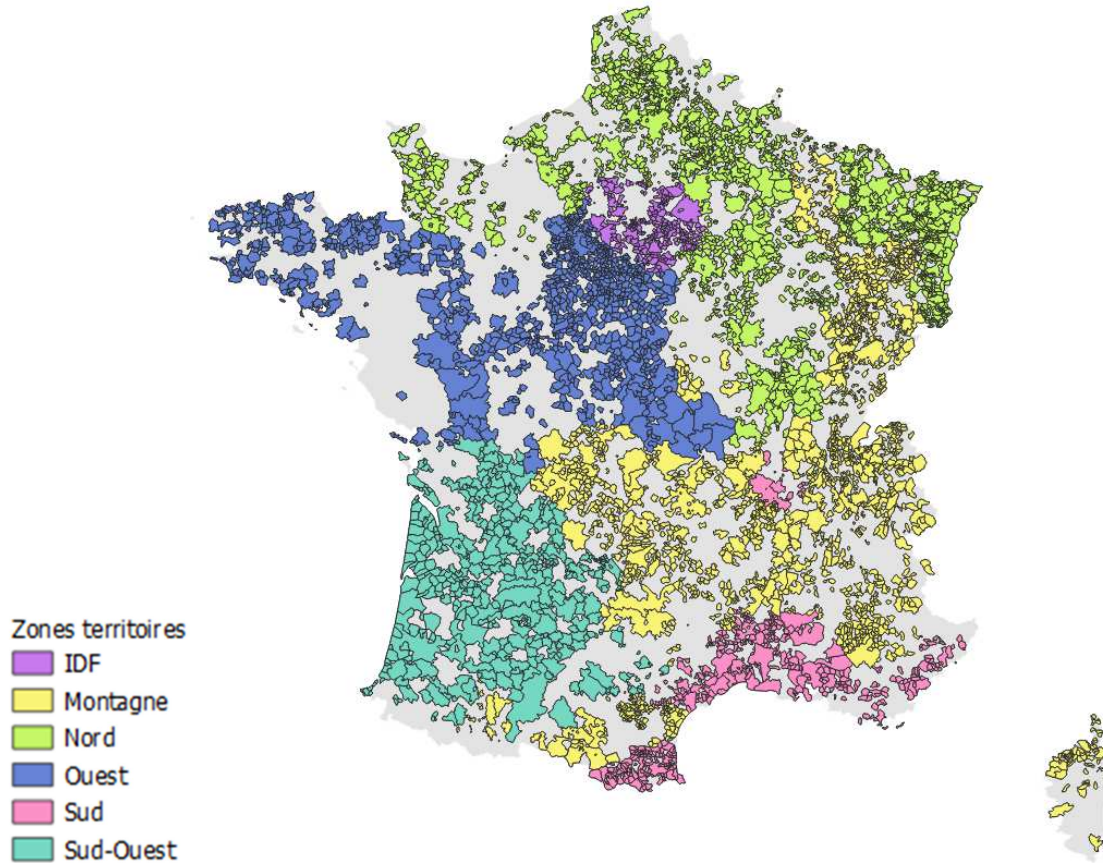


Figure 30 : Cartographie des six classes de territoires créées

Le modèle linéaire le plus robuste, qui prend en compte les éléments de contexte et satisfait au mieux la validation croisée est celui-ci ( $r^2 = 0.88$ ) :

$$\text{VNC} = \beta_{0,t} \cdot \text{taille} + \beta_1 \times N + \left[ N \times \left( \beta_2 \times \frac{L_{FG}}{L} + \beta_3 \times \frac{L_{FD}}{L} + \beta_4 \times \frac{L_{PVCV}}{L} + \beta_5 \times \frac{L_{PVCR}}{L} + \beta_6 \times \frac{L_{\text{dia}[60-100]}}{L} + \beta_7 \times \frac{L_{\text{dia}[\geq 125]}}{L} + \beta_8 \times \frac{\text{Routes}_{\text{urbaines}}}{\text{Routes}} + \beta_9 \times \text{Dénivelé}_{[0;1]} \right) \right] : \text{territoire}$$

Le VNC est égal à une constante  $\beta_{0,t}$  qui dépend des 4 classes de longueur de réseau (de 4,5 à 10 km :  $\beta_{0,1} = 744,7$  ; de 10 à 25 km :  $\beta_{0,2} = 4791,1$  ; de 25 à 65 km :  $\beta_{0,3} = 8773,1$  ; de 65 à 4200 km :  $\beta_{0,4} = 29\,521,2$ ) à laquelle on ajoute le nombre d'abonnés (N) multiplié par la constante  $\beta_1 = 12,9$ , auxquels on ajoute le nombre d'abonnés modulé par les paramètres de patrimoine et de contexte, pour lesquels on attribue un coefficient par territoire.

Les paramètres de patrimoine sont :

- Les proportions de matériaux (FG, FD, PVCV, PVCR) ;
- Les proportions de diamètres (diamètres entre 50 et 125 mm et diamètres supérieurs à 125 mm).

Les paramètres de contexte sont :

- La proportion d'urbain ;
- Le dénivelé.

Chaque coefficient ( $\beta_2$  à  $\beta_9$ ) est multiplié au terme de contexte et au nombre d'abonnés, et estimé pour chaque agence. Les coefficients sont reportés dans les Tableau 9 et Tableau 10.

**Tableau 9 : Coefficients du modèle linéaire sur les variables de patrimoine**

	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\beta_7$
	$\frac{L_{FG}}{L}$	$\frac{L_{FD}}{L}$	$\frac{L_{PVCV}}{L}$	$\frac{L_{PVCR}}{L}$	$\frac{L_{dia[60-100]}}{L}$	$\frac{L_{dia[\geq 125]}}{L}$
IDF	128.5	79.8	46.8	-81.9	-90.2	-31.0
Montagne	5.5	6.2	-22.3	40.4	66.6	-20.0
Nord	-6.1	-38.8	-4.7	-127.0	27.4	53.6
Ouest	14.2	-14.4	-0.8	-31.7	6.9	50.9
Sud	54.3	-32.6	94.2	-143.7	22.0	112.0
Sud-Ouest	45.0	91.6	73.8	10.3	-41.1	-86.5

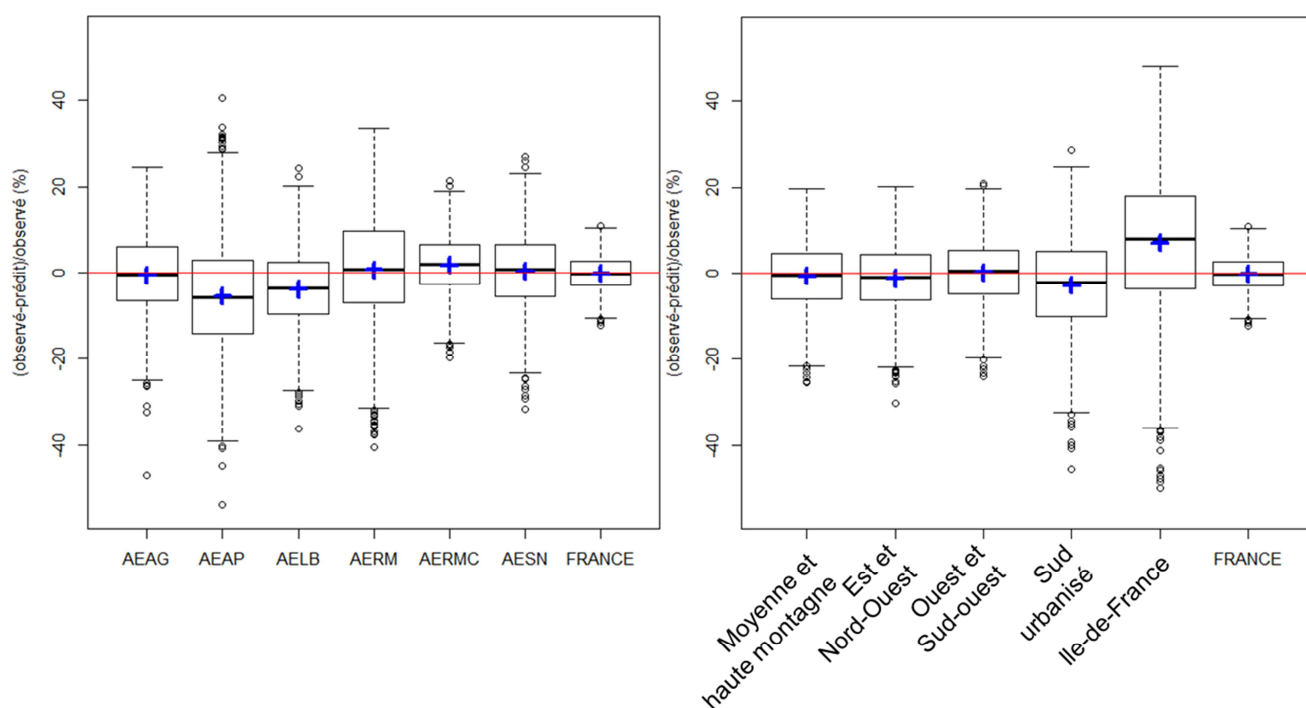
**Tableau 10 : Coefficients du modèle linéaire sur les variables de contexte**

	$\beta_8$	$\beta_9$
	$\frac{\text{Routes}_{urbaines}}{\text{Routes}}$	Dénivelé <sub>[0;1]</sub>
IDF	-6.6	198.0
Montagne	-8.9	33.7
Nord	12.3	132.4
Ouest	-17.3	93.4
Sud	-14.3	-17.7
Sud-Ouest	17.6	73.8

Les coefficients varient beaucoup d'une agence à l'autre, ne permettant pas de conclure sur une tendance globale pour un matériau ou une classe de diamètre. Les coefficients négatifs entraînent des prédictions de VNC négatifs, ce qui ne reflète pas la réalité. Le modèle de la partie 7.4 permet de ne prédire que des VNC positifs.

### 7.3.2. Validation croisée par la méthode Monte-Carlo

Les résultats de la validation croisée par la méthode de Monte-Carlo sont présentés sur la Figure 31.



**Figure 31 : Validation croisée du modèle linéaire**

Le graphique de gauche montre qu'il n'y a pas de surestimation ou sous-estimation pour une agence entière. Le graphique de droite montre que les services d'Ile-de-France sont légèrement sous-estimés.

## 7.4. Modèle de prédiction de $\ln(\text{VNC}/N)$

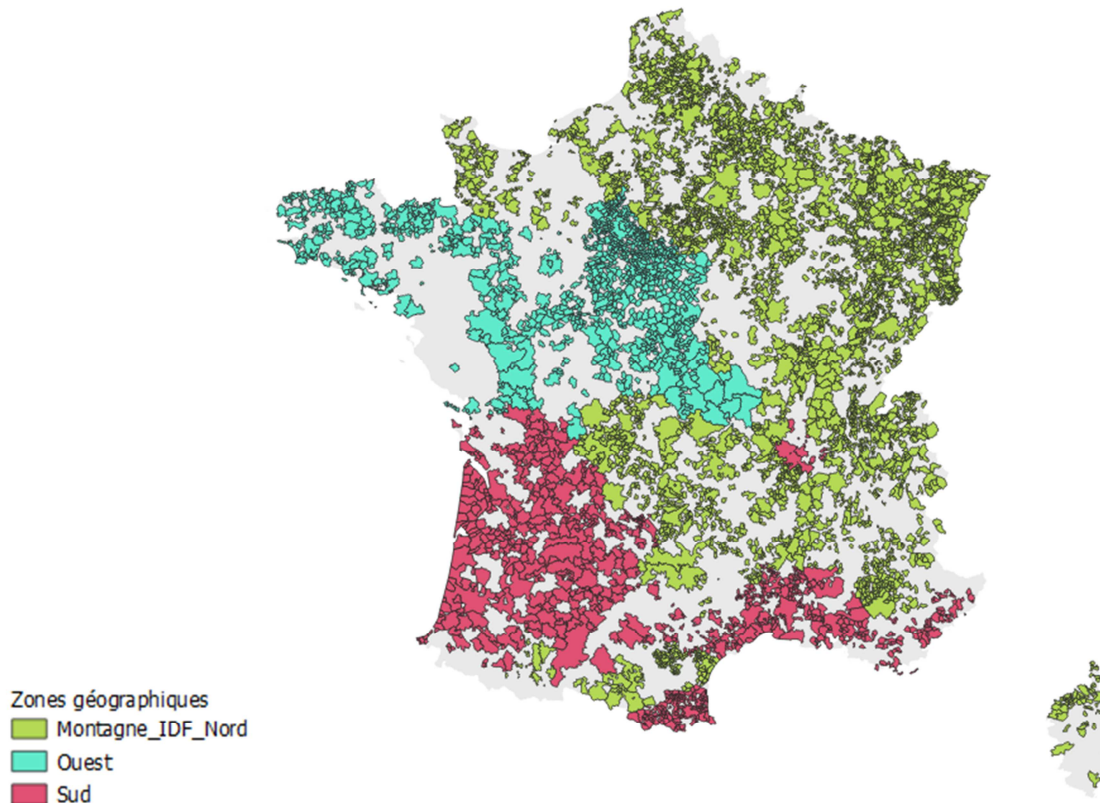
Le modèle exponentiel de la partie 7.2.2 est rappelé ci-dessous :

$$\text{VNC} = 38,7 \times N \times e^{\left( \underbrace{\begin{aligned} & -0,2 \times \frac{L_{FG}}{L_{tot}} - 0,05 \times \frac{L_{FD}}{L_{tot}} + 0,5 \times \frac{L_{PVCV}}{L_{tot}} - 1,3 \times \frac{L_{PVCR}}{L_{tot}} \\ & - 1,0 \times \text{AELB} + 0,3 \times \text{AEAG} + 0,2 \times \text{AE}(\text{AP} - \text{SN}) + 0,6 \times \text{AERM} + 0,4 \times \text{AERMC} \\ & - 0,2 \times \frac{\text{Routes}_{urbaines}}{\text{Routes}} + 0,2 \times \text{Dénivelé}_{[0,1]} \end{aligned}}_{= \varphi} \right)}$$

Ce modèle à l'avantage de pouvoir s'écrire sous la forme linéaire :

$$\ln\left(\frac{\text{VNC}}{N}\right) = \beta_0 + \varphi$$

Le modèle linéaire est calé et pour prédire au mieux le paramètre «  $\ln(\text{VNC}/N)$  », trois nouvelles classes de zones géographiques sont créées à partir des 5 classes d'hydroterritoires et des 6 agences, afin de prendre au mieux en compte les divergences selon les territoires et la tailles des services. Ce sont donc 6 classes qui sont créées correspondant aux 3 zones géographiques dans lesquelles on distingue les services de moins de 10 km de ceux de plus de 10 km. Les 3 classes géographiques sont représentées sur la Figure 32.



**Figure 32 : Cartographie des trois classes de zones géographiques**

Ce modèle prend aussi en compte les proportions des longueurs par diamètres et s'écrit comme ci-dessous :

$$\ln\left(\frac{\text{VNC}}{N}\right) = \beta_{0,i[1;6]} \cdot (\text{territoire} \times \text{taille}) + 0,37 \times \frac{L_{FD}}{L} - 0,34 \times \frac{L_{PVCR}}{L} - 0,17 \times \frac{L_{\text{dia}[\geq 125]}}{L} - 0,17 \times \frac{\text{Routes}_{urbaines}}{\text{Routes}} + 1,25 \times \text{Dénivelé}_{[0;1]}$$

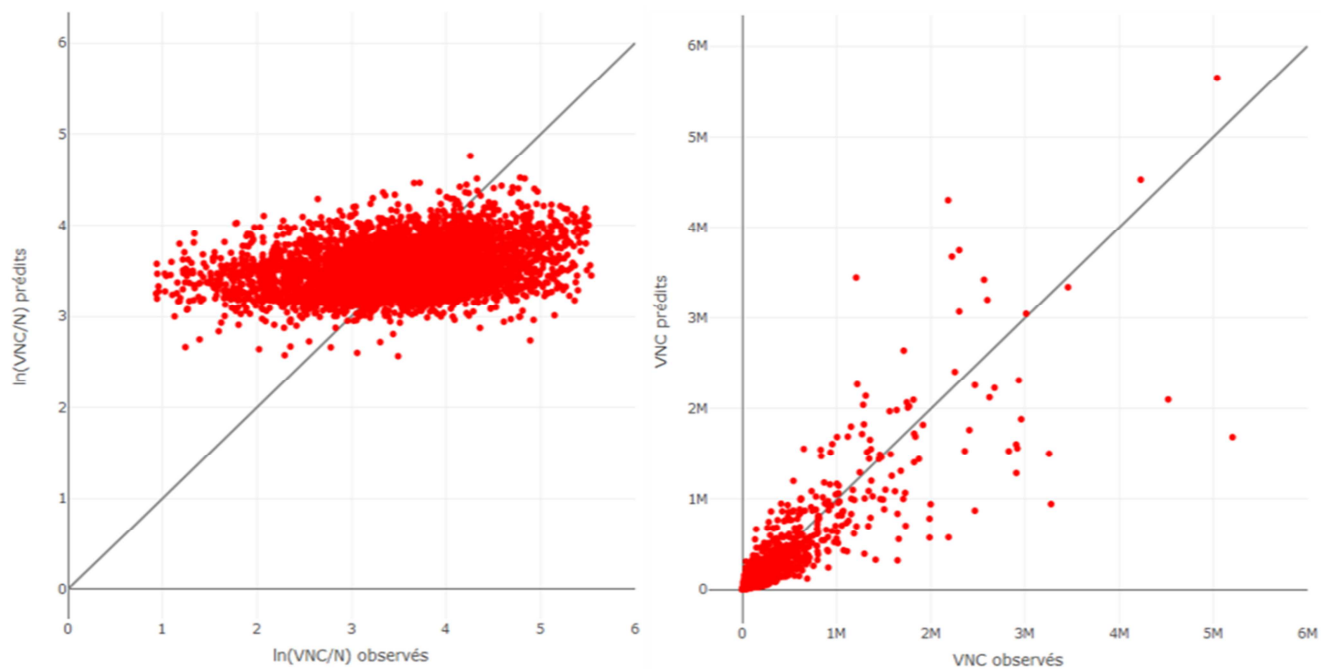
Le coefficient de régression de ce modèle est de 0,96 et les valeurs des coefficients  $\beta_{0,i[1;6]}$  qui dépendent de la classe du service sont reportées dans le Tableau 11.



**Tableau 11 : Coefficients  $\beta_{0,i}$  du modèle de prédictions de  $\ln(\text{VNC}/N)$**

Classe (géographie et taille)		$\beta_{0,i}$
1	Nord, Ile-de-France et montagne, L < 10 km	3,30
2	Nord, Ile-de-France et montagne, L > 10 km	3,52
3	Ouest, L < 10 km	3,06
4	Ouest, L > 10 km	3,31
5	Sud, L < 10 km	3,35
6	Sud, L > 10 km	3,71

Bien que le coefficient de régression de ce modèle soit très élevé, les prédictions ne sont pas satisfaisantes, comme le montre la distribution des  $\ln(\text{VNC}/N)$  prédits en fonction des  $\ln(\text{VNC}/N)$  observés (Figure 33, graphique de gauche). Bien que les  $\ln(\text{VNC}/N)$  observés varient entre 1 et 5,5, les prédictions sont pour la grande majorité comprises entre 2,8 et 4,2. Cependant, si l'on calcule les VNC prédits selon la formule  $\text{VNC}_{\text{prédits}} = \exp\left(\ln\left(\frac{\text{VNC}}{N}\right)_{\text{prédits}} \times N\right)$ , alors la prédiction des VNC paraît autant satisfaisante que celle obtenue avec les modèles de la partie 7.1 (Figure 33, graphique de droite).



**Figure 33 : Prédictions en fonction des observations des  $\ln(\text{VNC}/N)$  à gauche et des  $\text{VNC}/N$  à droite**

Cela confirme que le nombre d'abonnés explique en très grande partie le VNC et, bien que les paramètres de patrimoine et de contexte permettent de préciser le modèle, leur contribution est tellement faible qu'elle est difficilement identifiable.

Ce modèle a fait l'objet d'investigations récentes et la suite de l'action permettra de le préciser.

→ **Résultats de la partie 7 :**

- Le VNC est très corrélé au nombre d'abonnés ;
- Le patrimoine et le contexte permettent d'améliorer le modèle de prédiction du VNC en moindre mesure par rapport au nombre d'abonnés ;
- Les modèles montrent de grandes disparités selon les territoires et ne permettent pas de conclure sur l'influence du patrimoine sur le VNC.

## 8. Etude de l'Indice de Connaissance et de Gestion du Patrimoine (ICGP)

### 8.1. La base de données d'étude de l'ICGP

La base SISPEA fournit l'Indice de Connaissance et de gestion du Patrimoine (ICGP) à l'échelle des services. On distingue l'ICGP sur 120 points, de l'ICGP sur 45 points, étudié dans ce rapport. Les 45 points de l'ICGP45 sont attribués de la manière suivante :

- 15 points pour l'existence d'un plan de réseau et sa mise à jour ;
- 10 points si les matériaux et les diamètres sont connus à au moins 50 % et un point supplémentaire par tranche de 10 % connus ;
- 10 points si les dates ou périodes de pose sont connues pour au moins 50 % et un point supplémentaire par tranche de 10 % connus.

Les données SIG collectées donnent accès, pour les communes possédant un SIG, au niveau de connaissance du patrimoine et permet de calculer l'ICGP directement à partir du SIG.

Les services de la base SISPEA 2016 ayant renseigné l'ICGP sur 45 sont sélectionnés et cette base de données est croisée avec la base de données SIG collectées. Seuls les services composés de communes appartenant à un seul service et dont les données SIG collectées recouvrent au moins 90 % de la longueur de réseau indiquée dans la base SISPEA sont retenus. Cela permet de comparer les ICGP renseignés dans SISPEA et calculés à partir des SIG de 761 services, soit 116 844 km de canalisations.

### 8.2. Comparaison des ICGP calculés à partir des SIG et de la base SISPEA

On fait l'hypothèse que les services dont le SIG a été collecté possèdent les 15 points pour l'existence et la mise à jour du plan de réseau. Les résultats obtenus sur les 761 services étudiés sont synthétisés dans le Tableau 12:

**Tableau 12 : Comparaison des ICGP renseignés dans SISPEA et des ICGP calculés avec les SIG**

	Renseignement de l'ICGP120 dans la base SISPEA (%)	Moyenne ICGP45 (nombre de points)		Part de services possédant un ICGP45 supérieur à 40 (%)	
		SISPEA 2016	SIG	SISPEA 2016	SIG
Loire-Bretagne	63,7	42.46	36.1	94.6	51.0
Adour-Garonne	46,5	39.82	31.47	89.2	32.4
Artois-Picardie	75,8	42.44	35.67	100	55.5
Rhin-Meuse	54,6	42.67	42.84	94.7	87.7
Seine-Normandie	47,1	42.41	35.16	96.8	44.2
Rhône-Méditerranée-Corse	42,8	40.08	33.89	86.2	40.8
<b>National</b>	<b>50,1</b>	<b>41.52</b>	<b>35.31</b>	<b>92.5</b>	<b>47.3</b>

Les informations sont contradictoires : alors que les bases SISPEA indiquent que plus de 90 % des services ont un ICGP45 supérieur à 40 à l'échelle nationale, les données SIG indiquent qu'un peu moins de 50 % des services atteignent ce score. Notons que l'échantillon étudié ne regroupe que les services pour lesquels un SIG est disponible. La valeur d'ICGP45 moyen indiquée est donc une surestimation de la réalité car elle ne prend pas en compte les services ne possédant pas de SIG dont la connaissance des réseaux est vraisemblablement plus lacunaire.

## 9. Etude à l'échelle du secteur

### 9.1. Intérêt

L'étude du lien entre la performance et le patrimoine à l'échelle des services a permis d'identifier des liens, surtout pour les matériaux fonte et PVC. Il est difficile d'affiner les relations ou d'étudier les autres paramètres de patrimoine, notamment pour les raisons suivantes :

- Le grand nombre d'inconnues sur les périodes de pose ;
- La fiabilité interrogée de certaines données ;
- L'hétérogénéité des territoires étudiés.

Analyser les données de patrimoine et de performance à l'échelle des secteurs devrait permettre une plus grande fiabilité des données sur le patrimoine et les débits horaires. De plus, les données seront plus homogènes (même exploitant, patrimoines semblables...).

### 9.2. Terrains d'étude

Environ 140 secteurs répartis sur 12 collectivités ont été présélectionnés sur le département de la Gironde. Ces secteurs seront sélectionnés pour l'analyse statistique si suffisamment de données de compteurs sont disponibles, si les SIG sont complets et si le nombre de compteurs dans un même secteur n'est pas trop grand (pour éviter une incertitude trop élevée).

Les secteurs du département de la Charente-Maritime exploités par la régie RESE 17 vont potentiellement pouvoir être exploités, les discussions sont en cours avec Eau 17.

### 9.3. Données

Pour chaque secteur, les débits de compteur de sectorisation et les marnages des réservoirs ainsi que les consommations des abonnés permettront de calculer un IVNC attribué à chaque secteur. Les données de casses et de renouvellement, lorsqu'elles sont disponibles, permettront d'étudier les taux de casses et de renouvellement en lien avec le patrimoine connu grâce aux SIG collectés.

## 10. Evaluation entre la valeur à neuf des réseaux publics d'eau potable et le prix de l'eau potable

### 10.1. Introduction

La fourniture d'un service à l'utilisateur est l'objectif principal des services publics d'eau potable. Cette fourniture de service se fait par l'interface d'une infrastructure publique constituée d'ouvrages et d'équipements permettant le prélèvement d'eau brute, le traitement, le stockage et le transport jusqu'au droit des immeubles à desservir. Le réseau public d'eau potable constitue en termes de valeur à neuf 85 % de la valeur à neuf de l'ensemble des biens publics des services publics d'eau potable, soit environ 134,8 milliards d'euros.

L'investissement conséquent dans des infrastructures caractérise les services publics d'eau potable comme une activité fortement capitalistique : elle nécessite lors de sa création puis ensuite en phase d'amélioration, d'extension et de renouvellement des capitaux financiers conséquents.

Par conséquent, la question du lien entre la valeur financière du réseau et le prix de l'eau acquitté par les usagers se pose d'une manière évidente. La présente étude se propose d'investiguer ce sujet et de le caractériser en fonction des variables descriptives habituelles des services publics.

### 10.2. Données mobilisées et caractérisation de l'échantillon

Dans notre première approche, les données mobilisées sont les suivantes :

- La base de données de l'observatoire national des services d'eau potable et d'assainissement (SISPEA) en y mobilisant les données relatives au prix 120 m<sup>3</sup> et la composition communale des services d'eau potable au titre de l'année 2018.

- La base de données relative à la valeur à neuf du patrimoine établie par commune dans le cadre de l'action n°70 de la convention AFB/IRSTEA 2016-2018.

Les données disponibles le sont pour 3264 services d'eau en raison du taux de remplissage de la base de données SISPEA 2018. Il apparaît qu'en termes de types de collectivités locales porteuses de la compétence eau potable (figure n°1), les communes sont sous-représentées dans notre échantillon de travail au regard de l'effectif national des services d'eau (70,5 % contre 73,7 %) et les syndicats (EPCI) surreprésentés (26,9 % contre 24 %), tout comme les EPCI-FP (2,5 % contre 2,1 %). Ce décalage trouve principalement son explication par le taux de remplissage de SISPEA qui est plus élevé pour les grands services (EPCI-FP et EPCI) que pour les petits services (les communes en général).

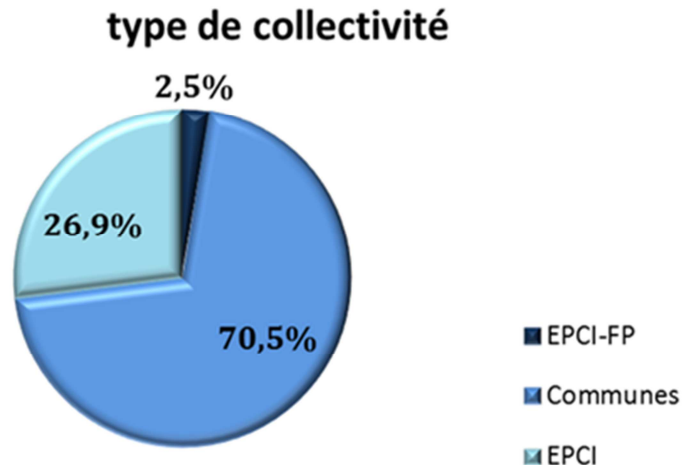


Figure 34 : Caractéristique de l'échantillon de collectivités par types de collectivité locale

### 10.3. Premiers résultats

#### 10.3.1. Echantillon global

En premier lieu, nous avons recherché pour l'ensemble de notre échantillon de service une relation statistique macroscopique entre le prix 120 m<sup>3</sup> et la valeur à neuf du patrimoine du service (Figure 35). Pour gérer l'effet de taille des services, nous avons mobilisé le ratio valeur à neuf par habitant desservi dans notre première analyse.

Les tentatives de recherche de lien statistique entre le prix moyen pour une consommation d'eau de 120 m<sup>3</sup> tel que définie par l'INSEE et la valeur à neuf du réseau de distribution d'eau potable des collectivités se sont révélées infructueuses. Qu'il s'agisse d'une régression linéaire, logarithmique, exponentielle ou de puissance, aucun lien ne peut être mis en évidence. Les coefficients de corrélation sont très proches de zéro et mettent à jour une indépendance statistique entre les deux variables mobilisées.

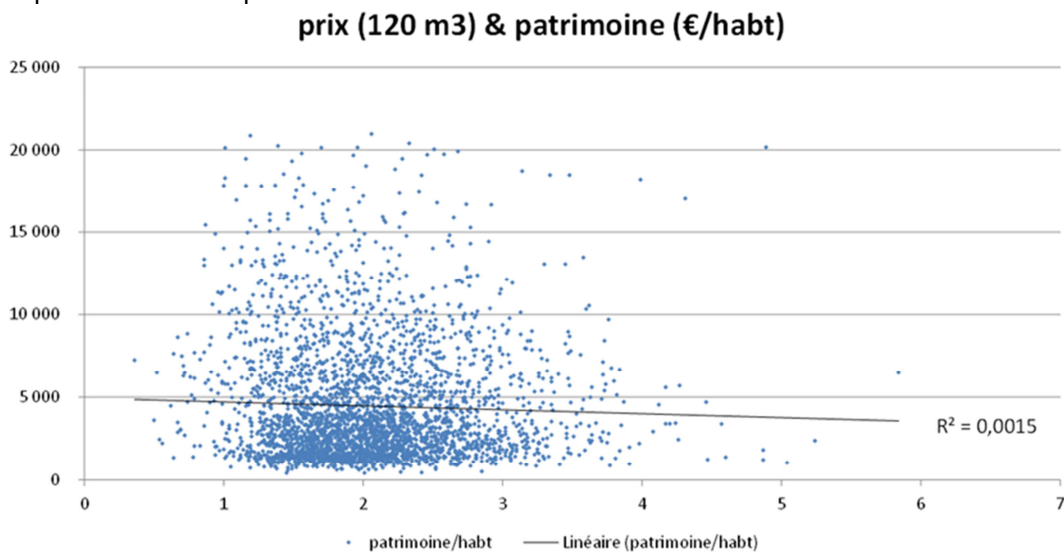
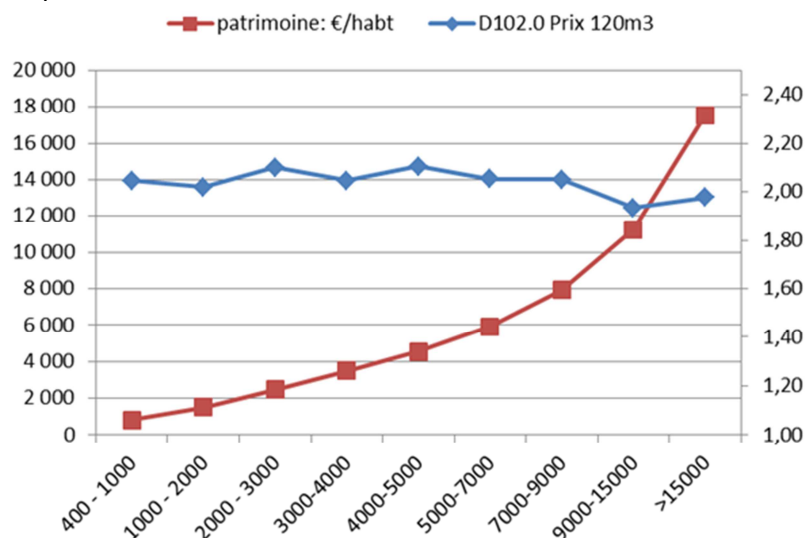


Figure 35 : Régression linéaire entre prix 120 m<sup>3</sup> et valeur à neuf du patrimoine réseau par habitant

### 10.3.2. Par classes de valeur à neuf du réseau

En second lieu, une approche segmentée a été entreprise : il s'agit d'observer par classe de valeur à neuf du réseau (les valeurs moyennes de chaque classe apparaissent en rouge dans la Figure 36) les valeurs moyennes par classe du prix de l'eau.

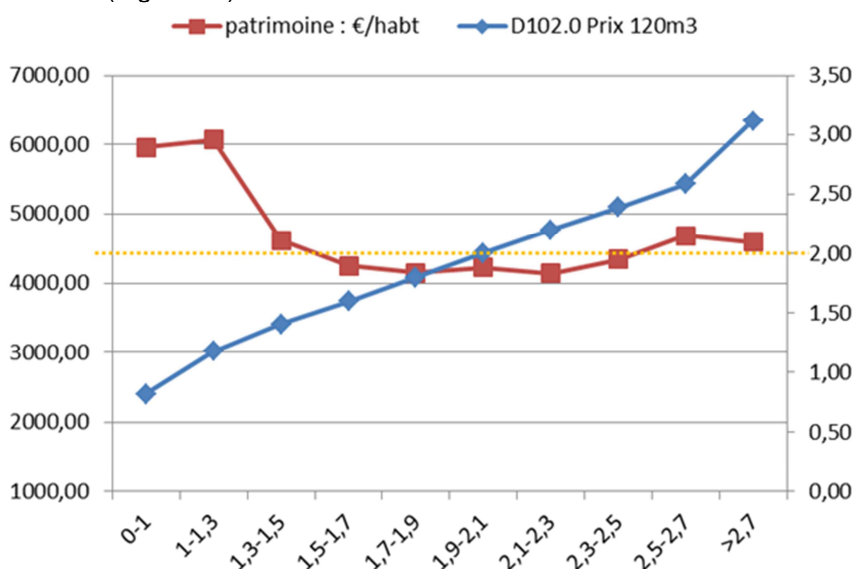


**Figure 36 : Prix 120 m<sup>3</sup> suivant les classes de valeur à neuf du réseau d'eau par habitant desservi**

Il apparaît dans cette nouvelle approche une indépendance entre le prix 120 m<sup>3</sup> et les classes de valeur à neuf par habitant desservi. Qu'il s'agisse des services présentant une valeur de ratio très faible (moins de 1000 €/habitant) ou très élevés (plus de 15 000 €/habitant), le prix 120 m<sup>3</sup> reste proche de sa valeur moyenne sur l'ensemble de l'échantillon, soit 2,04 € TTC. Le minimum observé est de 1,93 € (classe de patrimoine : 9000-15000 €/hab) et le maximum s'élève à 2,10 € (pour la classe 4000 – 5000 €/hab). Cette approche vient conforter l'invariance constatée entre les deux variables lors de l'approche sur l'échantillon global.

### 10.3.3. Par classe de prix de l'eau

Cette fois-ci, nous avons observé les valeurs prises par le ratio valeur du patrimoine par habitant en fonction des classes de prix de l'eau (Figure 37).



**Figure 37 : Valeur à neuf du patrimoine du réseau par habitant desservi en fonction des classes de prix de l'eau**

Dans cette approche, il apparaît deux groupes de services :

- Les services ayant un prix de l'eau très faible (moins de 1,3 € soit les deux premières classes) et qui ont un ratio de valeur à neuf du réseau par habitant de l'ordre de 6 000 €/habitant ;
- Les services ayant un prix de l'eau supérieur à 1,3 € (les huit autres classes) et qui présentent eux une valeur proche de 4300 €/habitant pour la seconde variable mobilisée.

Le premier groupe correspond en moyenne à des petits services communaux ruraux qui ont des prix de l'eau faibles en raison d'infrastructures à gérer avec peu de degré de complexité, une capacité budgétaire à financer

de l'investissement très faible voire inexistante et un prix de l'eau parfois subventionné par le budget général. Ces services présentent une densité linéaire d'habitant faible caractéristique d'un contexte rural. L'action n°70 avait mis en avant que les services ruraux avaient des valeurs élevées du ratio patrimoine par habitant par rapport aux autres types de services (urbains et ultra-urbains). Nous retrouvons ici, pour les deux classes de prix de l'eau les plus faibles, le même constat.

Pour le second groupe, il apparaît que les classes de prix ne discriminent pas la valeur du patrimoine réseau par habitant.

#### **10.4. Perspectives**

Les premiers résultats ne mettent pas en avant de lien clair entre prix de l'eau et valeur à neuf du réseau d'eau. Pour l'année à venir de la présente étude, les actions suivantes sont envisagées :

- Elargir le nombre de données en mobilisant une base de données SISPEA antérieure (2017 ou 2016) ;
- Avec l'augmentation du nombre de services, on pourra explorer plus finement les caractéristiques des services et avoir une approche plus pertinente par classe de prix de l'eau ou de valeur patrimoniale ;
- Etudier les liens entre les indicateurs de performance relatifs au réseau de distribution d'eau potable et le prix du service.

## 11. Synthèse

L'étude à l'échelle des services sur le territoire national a montré un lien entre la performance des réseaux d'eau potable et le patrimoine :

- L'IVNC est corrélé positivement aux proportions de fonte, et négativement aux proportions de PVC, et cela est d'autant plus visible que la longueur des réseaux étudiés est grande ;
- Les proportions de diamètres inférieurs à 50 mm, qui sont quasi-exclusivement en PVC, sont aussi corrélées négativement à l'IVNC. Les proportions de diamètres supérieurs à 100 mm, qui sont majoritairement en fonte, sont corrélées positivement à l'IVNC mais cette corrélation est moins nette que pour la fonte seule.

Ces résultats doivent cependant être interprétés avec prudence car ils ne prennent pas en compte l'effet de la densité d'abonnés, l'IVNC étant également très lié au taux d'urbanisation.

Des distinctions entre la fonte grise et la fonte ductile, et les PVC posés avant 1980 et après 1980 sont perceptibles mais pas assez nettes pour comprendre les différences entre les matériaux. Aucun lien entre le VNC/N et le patrimoine n'a été identifié. Les études menées à ce jour n'ont pas permis de montrer un lien entre la performance des réseaux et leur période de pose.

Certains paramètres de contexte sont très liés au patrimoine et à la performance :

- Le taux d'urbanisation ;
- L'aire de localisation qui est mise en évidence par les agences de l'eau et par les hydroterritoires ;
- Les différences d'altitude ;
- La taille du service.

Le VNC est très corrélé à la taille du service et notamment au nombre d'abonnés, il est également lié au patrimoine dans une moindre mesure. Le VNC et le patrimoine sont aussi liés à des paramètres de contexte, mais la modélisation statistique ne permet pas de caractériser les liens à l'échelle du service.

Le VNC étant davantage corrélé au nombre d'abonnés qu'à la longueur de réseau, et le VNC/N n'étant pas corrélé au patrimoine, le VNC/N peut être un indicateur fort utile pour juger de la performance d'un réseau.

L'étude à l'échelle du secteur permettra de confirmer ou non les hypothèses formulées, affiner les résultats et en apporter de nouveaux.

Aucun lien clair n'a été mis en évidence entre le prix de l'eau et la valeur à neuf du réseau, l'élargissement de la base de données devrait permettre d'explorer plus finement les caractéristiques des services.

## Sigles & Abréviations

**AEAG** : Agence de l'Eau Adour Garonne

**AEAP** : Agence de l'Eau Artois Picardie

**AELB** : Agence de l'Eau Loire Bretagne

**AERM** : Agence de l'Eau Rhin Meuse

**AERMC** : Agence de l'Eau Rhône Méditerranée Corse

**AESN** : Agence de l'Eau Seine Normandie

**AEP** : Alimentation en Eau Potable

**AFB** : Agence Française pour la Biodiversité

**ICGP** : Indice de Connaissance et de Gestion Patrimoniale

**IRSTEA** : Institut National de Recherche en Sciences et Technologies pour l'Environnement et l'Agriculture

**SISPEA** : Système d'Information sur les Services Publics d'Eau et d'Assainissement



## 12. Bibliographie

CHESEL, D., THIOULOUSE, J., & DUFOUR, A. (Septembre 2004). *Introduction à la classification hiérarchique*.

DEQUESNE, J., & BREJOUX, E. (Septembre 2018). *Observatoire des services publics d'eau et d'assainissement - Panorama des services et de leur performance en 2015*.

HUSSON, A., VACELET, A., STRICKER, A.-E., LATOUR, J., LEGAT, Y., & RENAUD, E. (Février 2019). *Construction d'un dispositif permanent d'évaluation du patrimoine des réseaux d'eau potable aux échelles nationale et de bassin. Axe 1 : Approche technique*.

LORENZ, M. O. (Janvier 1905). *Methods of Measuring the Concentration of Wealth. Publications of the American Statistical Association, Vol. 9, No.70*.

*Arrêté du 2 décembre 2013 modifiant l'arrêté du 2 mai 2007 relatif aux rapports annuels sur le prix et la qualité des services publics d'eau potable et d'assainissement*. (2013). Récupéré sur Legifrance: <https://www.legifrance.gouv.fr/eli/arrete/2013/12/2/DEVL1323309A/jo/texte>

*Arrêté du 2 mai 2007 relatif aux rapports annuels sur le prix et la qualité des services publics d'eau potable et d'assainissement*. (2007). Récupéré sur Legifrance: <https://www.legifrance.gouv.fr/affichTexte.do?cidTexte=JORFTEXT000000274838>

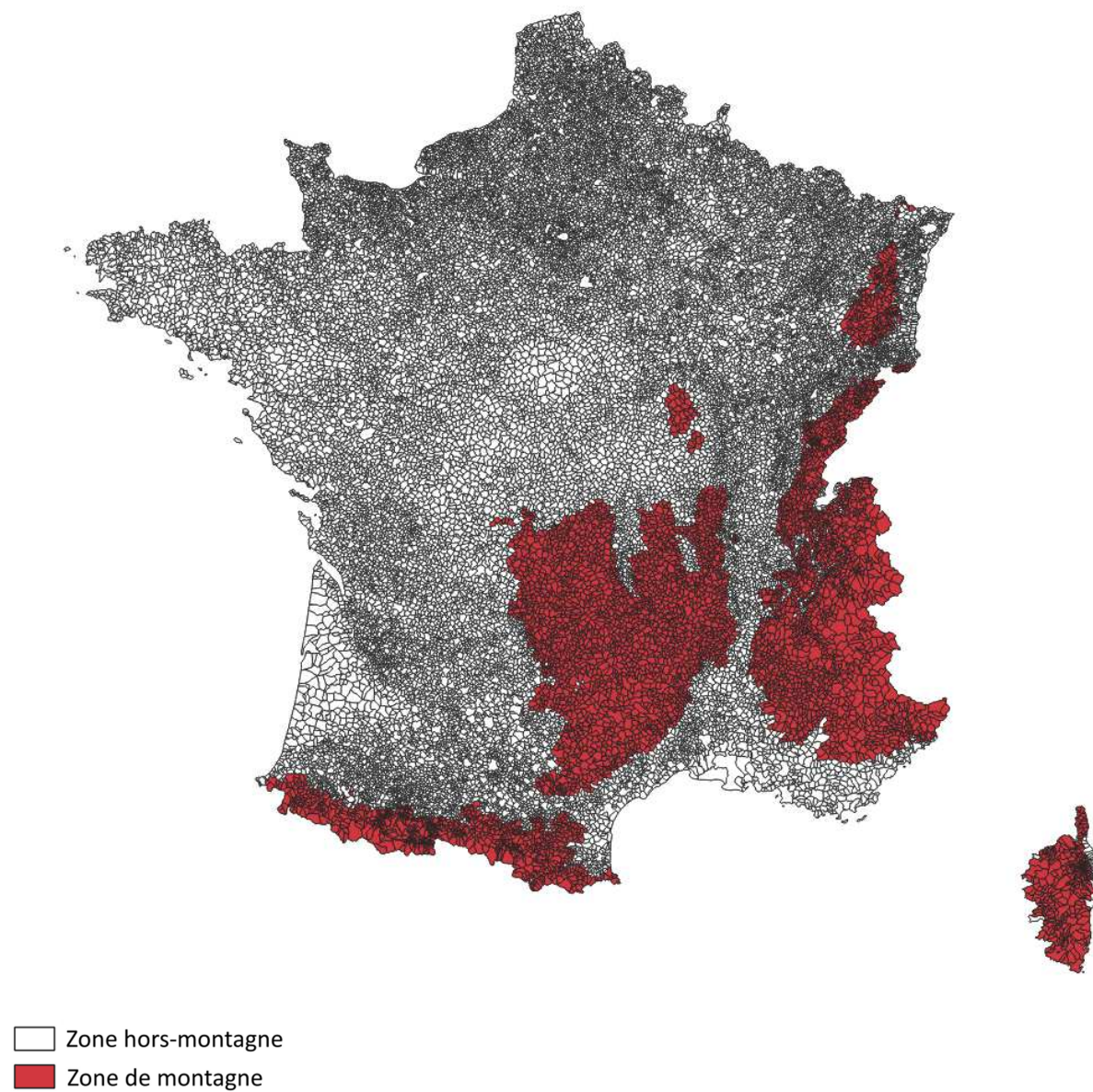
*Décret n° 2012-97 du 27 janvier 2012 relatif à la définition d'un descriptif détaillé des réseaux des services publics de l'eau et de l'assainissement et d'un plan d'actions pour la réduction des pertes d'eau du réseau de distribution d'eau potable*. (2012). Récupéré sur Legifrance: <https://www.legifrance.gouv.fr/eli/decret/2012/1/27/DEVL1132866D/jo/texte>

### 13. Table des illustrations

Figure 1 : Cartographie des services de la base de données patrimoine-performance .....	8
Figure 2 : Tri des services d'après le paramètre IVNC de la base de données patrimoine-performance .....	9
Figure 3 : Tri des services d'après le paramètre VNC/N de la base de données patrimoine-performance .....	10
Figure 4 : Paliers de sélection des services sur les paramètres N et L .....	10
Figure 5 : Représentation schématique de la création des 4 sous-échantillons.....	11
Figure 6 : Part de matériaux dans chaque classe de diamètres calculées sur les données SIG collectées .....	13
Figure 7 : Représentation ACP des corrélations entre le patrimoine, la performance et le contexte .....	14
Figure 8 : Interprétation d'une boîte à moustache.....	15
Figure 9 : Distribution de l'IVNC (en m3/km/j) dans les 5 classes d'hydroterritoires.....	15
Figure 10 : Distribution des proportions de fonte et de PVC dans les 5 classes d'hydroterritoires .....	15
Figure 11 : Distribution des proportions de longueurs posées avant 1945 dans les 5 classes d'hydroterritoires .....	16
Figure 12 : Distribution des proportions d'urbain et du dénivelé dans les 5 classes d'hydroterritoires .....	16
Figure 13 : Distribution des IVNC selon le caractère urbain ou rural.....	17
Figure 14 : Distribution des proportions de fonte et de PVC selon le caractère urbain ou rural.....	18
Figure 15 : Distribution des proportions de diamètres supérieurs à 300 mm selon le caractère urbain ou rural .....	18
Figure 16 : Distribution des proportions de fonte et de PVC selon la taille des services.....	18
Figure 17 : Courbes cumulées (classement décroissant) sur les matériaux fonte et PVC.....	19
Figure 18 : Courbes cumulées (classement décroissant) sur les diamètres .....	20
Figure 19 : Courbes cumulées (classement croissant) des matériaux fonte et PVC.....	21
Figure 20 : Courbes cumulées (classement croissant) et expérimentation .....	21
Figure 21 : Dendrogramme obtenu par CAH et représentation des 20 classes .....	22
Figure 22 : IVNC en fonction de la proportion de fonte après CAH .....	22
Figure 23 : IVNC en fonction de la proportion de PVC après CAH.....	23
Figure 24 : IVNC en fonction des proportions de diamètres (<50, 50 à 100, >100) après CAH.....	23
Figure 25 : Résidus en fonction des VNC observés .....	25
Figure 26 : Résidus relatifs en fonction du VNC observé .....	25
Figure 27 : VNC prédits en fonction des VNC observés .....	26
Figure 28 : Schématisation de la validation croisée par la méthode de Monte-Carlo.....	27
Figure 29 : Validation croisée par la méthode Monte-Carlo du modèle linéaire .....	27
Figure 30 : Cartographie des six classes de territoires créées .....	29
Figure 31 : Validation croisée du modèle linéaire .....	30
Figure 32 : Cartographie des trois classes de zones géographiques .....	31
Figure 33 : Prédications en fonction des observations des ln(VNC/N) à gauche et des VNC/N à droite .....	32
Figure 34 : Caractéristique de l'échantillon de collectivités par types de collectivité locale .....	35
Figure 35 : Régression linéaire entre prix 120 m <sup>3</sup> et valeur à neuf du patrimoine réseau par habitant.....	35
Figure 36 : Prix 120 m <sup>3</sup> suivant les classes de valeur à neuf du réseau d'eau par habitant desservi .....	36
Figure 37 : Valeur à neuf du patrimoine du réseau par habitant desservi en fonction des classes de prix de l'eau .....	36
.....	
Tableau 1 : Classes sur les données patrimoine et abréviations utilisées.....	6
Tableau 2 : Coefficients de corrélation entre les proportions de matériaux et l'IVNC.....	11
Tableau 3 : Coefficients de corrélation entre les proportions de diamètres et l'IVNC.....	12
Tableau 4 : Coefficients de corrélation entre les proportions de périodes de pose et l'IVNC.....	12
Tableau 5 : Coefficients de corrélation entre les proportions de matériaux et de périodes de pose (échantillon total).....	12
Tableau 6 : Coefficients de corrélation entre les proportions de matériaux et de diamètres (échantillon total) ...	13
Tableau 7 : Paramètres de performance en fonction du mode de gestion .....	14
Tableau 8 : Coefficients de corrélation entre la performance et les paramètres de contexte.....	18
Tableau 9 : Coefficients du modèle linéaire sur les variables de patrimoine .....	30
Tableau 10 : Coefficients du modèle linéaire sur les variables de contexte .....	30
Tableau 11 : Coefficients $\beta_{0,i}$ du modèle de prédictions de ln(VNC/N) .....	32
Tableau 12 : Comparaison des ICGP renseignés dans SISPEA et des ICGP calculés avec les SIG .....	33

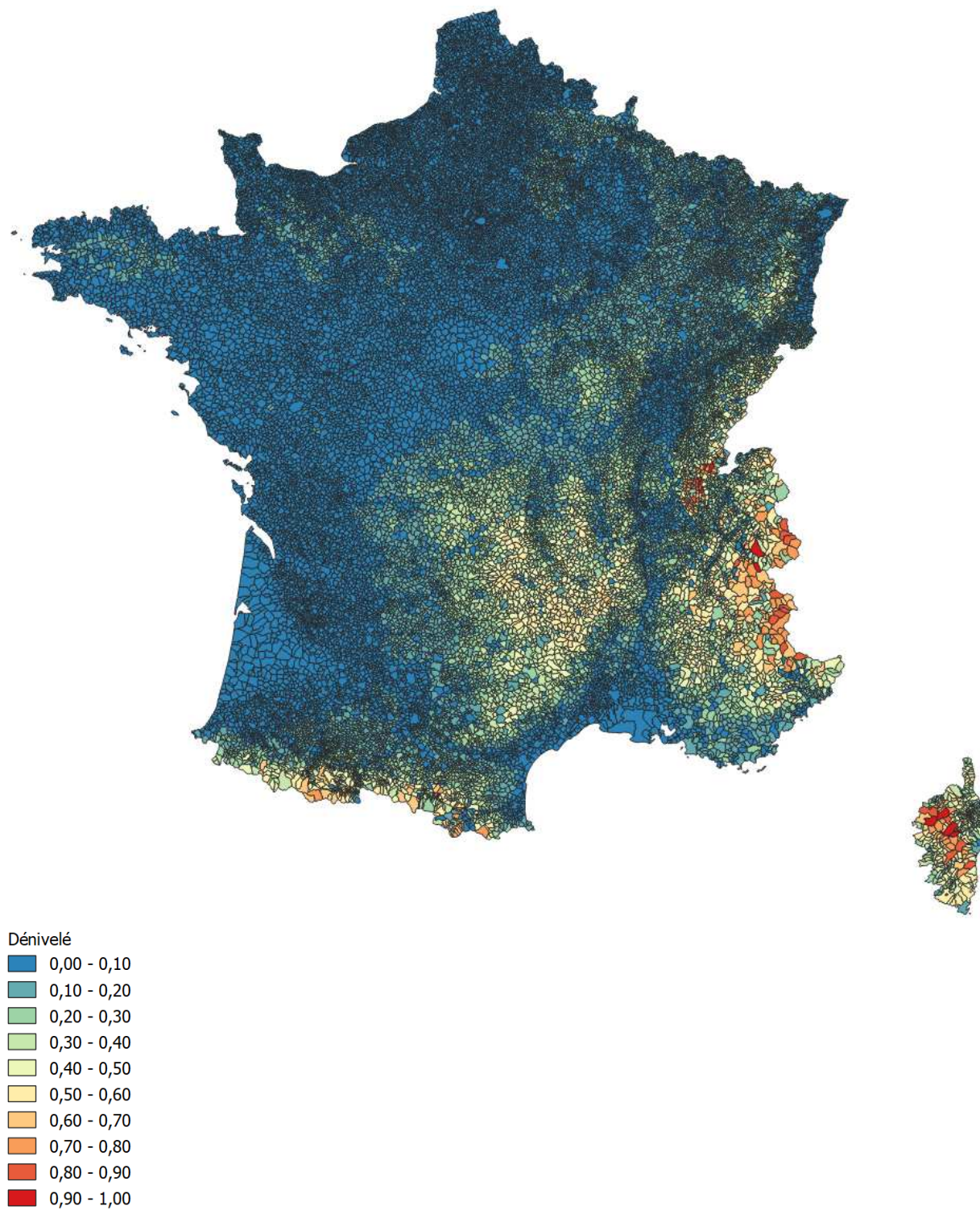
## 14. Annexes

### 14.1. Annexe 1 : Carte des communes en zone de montagne

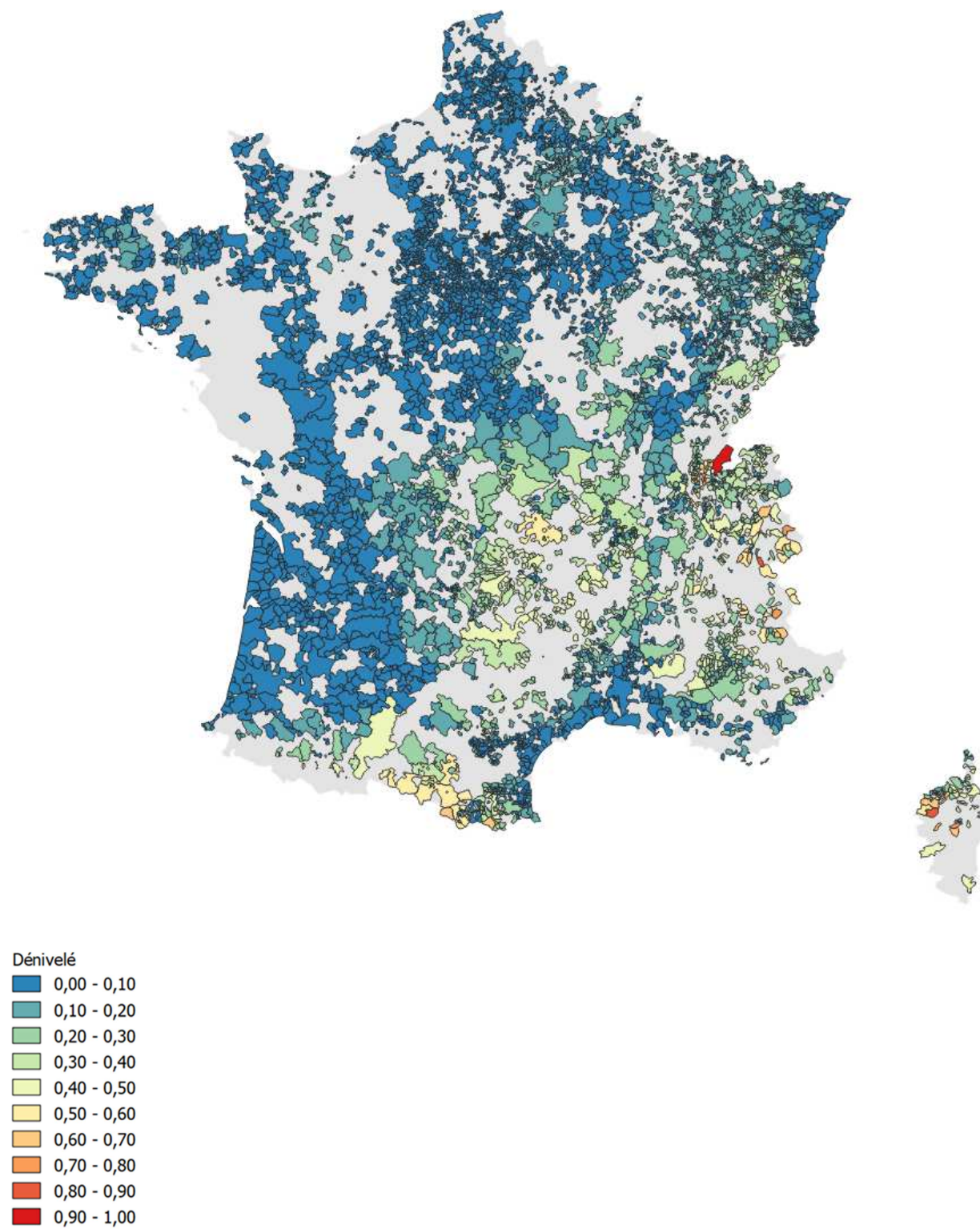


## 14.2. Annexe 2 : Cartes du paramètre « dénivelé »

### A. Par communes sur le territoire national

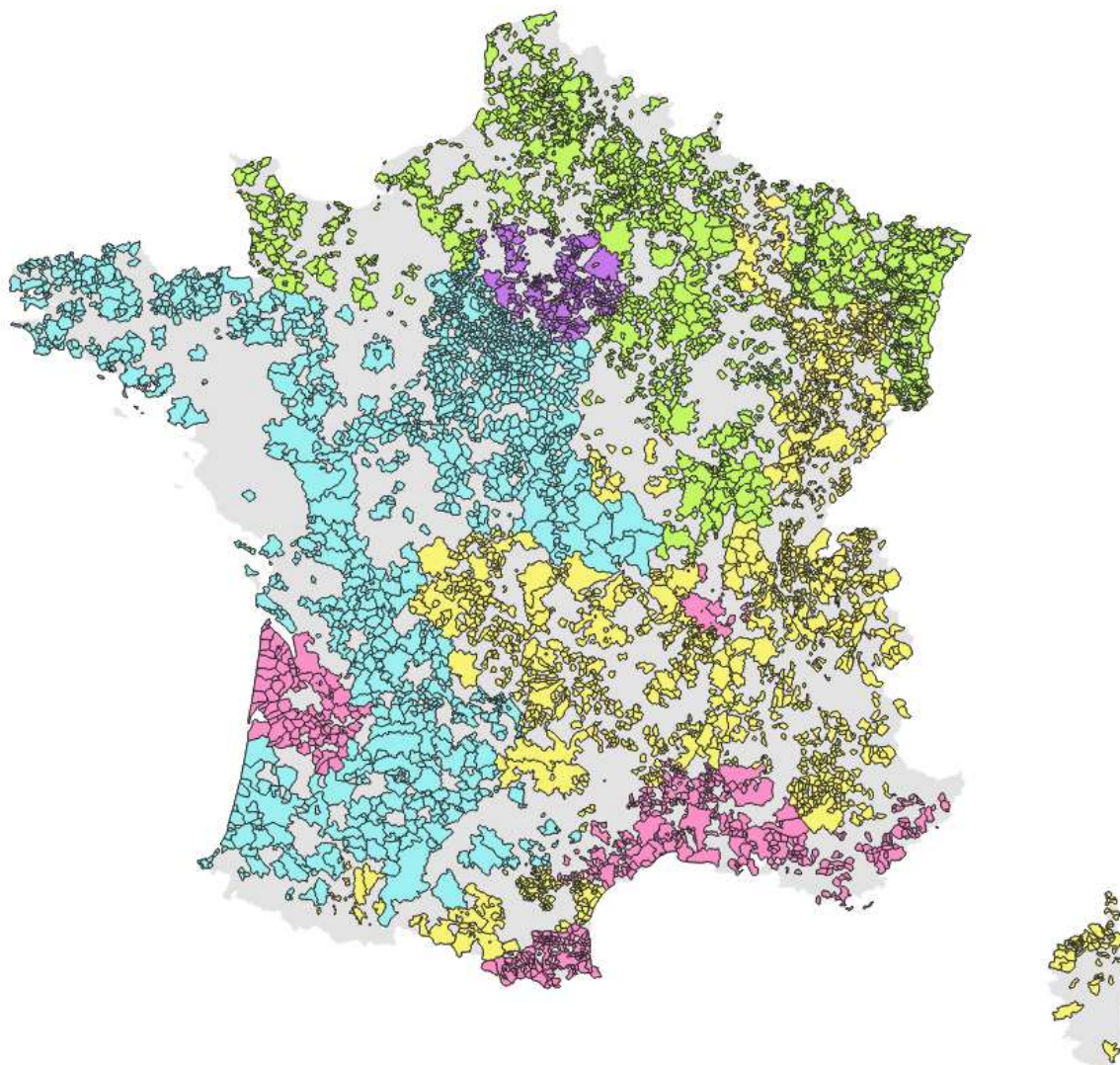


**B. Par services de la base de données patrimoine-performance (5708 services)**





**B. Par services de la base de données patrimoine-performance (5708 services)**



- Moyenne et haute montagne
- France de l'Est et du Nord-ouest
- France de l'Ouest et du Sud-Ouest
- Sud Urbanisé
- Ile-de-France

**Irstea**

1, rue Pierre-Gilles de Gennes  
CS 10030  
92761 Antony Cedex

**01 40 96 61 21**

[www.irstea.fr](http://www.irstea.fr)

**Agence Française pour la Biodiversité**

Hall C – Le Nadar  
5, square Félix Nadar  
94300 Vincennes

**01 45 14 36 00**

[www.afbiodiversite.fr](http://www.afbiodiversite.fr)