



HAL
open science

Pertes en eau : introduction de la pression dans les indicateurs de performance : rapport final

Eddy Renaud, Julie Pillot

► **To cite this version:**

Eddy Renaud, Julie Pillot. Pertes en eau : introduction de la pression dans les indicateurs de performance : rapport final. irstea. 2013, pp.74. hal-02599323

HAL Id: hal-02599323

<https://hal.inrae.fr/hal-02599323>

Submitted on 16 May 2020

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.

Partenariat Irstea-Veolia Eau

Pertes en eau : Introduction de la pression dans les indicateurs de performance

Rapport Final

Septembre 2013

Eddy RENAUD ⁽¹⁾

Julie PILLOT ⁽¹⁾

⁽¹⁾ Groupement de Bordeaux
Unité de recherches REBX
50, avenue de Verdun, Gazinet
33612 Cestas cedex

Liste des sigles, acronymes et abréviations

AEP	Alimentation en Eau Potable
ANOVA	Analysis of variance
DT	Direction Technique
Irstea	Institut de Recherche en Sciences et Technologies pour l'Environnement et l'Agriculture
IWA	International Water Association
ONSEA	Observatoire National des Services d'Eau et d'Assainissement
RAD	Rapport Annuel du Délégué
RPQS	Rapport sur le Prix et de la Qualité du Service
SEDIF	Syndicat des Eaux d'Ile de France
SISPEA	Système d'Information sur les Services Publics d'Eau et d'Assainissement
VEDIF	Veolia Eau D'Ile de France
VP	Volume de pertes

Table des matières

Table des illustrations - Figures	4
Table des illustrations - Tableaux.....	5
Table des illustrations - Tables.....	6
I Introduction.....	7
II Rôle de la pression de service dans le volume de pertes	9
II.1 Services AEP étudiés.....	9
II.1.1 Calcul des indicateurs de pression de service	9
II.1.2 Constitution de l'échantillon.....	9
II.2 Etude de l'impact de la pression sur les pertes	13
II.2.1 Etude de la distribution des pertes	13
II.2.2 Etude statistique du lien entre les indicateurs de pression et les pertes	14
II.2.3 Etude statistique par classe de pression	21
III Recherche d'un indice de pertes.....	29
III.1 Constitution de l'échantillon des Services AEP.....	29
III.2 Etude de la distribution des pertes.....	31
III.3 Construction d'un modèle des pertes	35
III.3.1 Identification des variables significatives.....	35
III.3.2 Sélection du meilleur modèle	36
III.3.3 Validation sur d'autres données	44
III.3.4 Modèles linéaires simples.....	46
III.4 L'indice de pertes	48
III.4.1 Définition.....	48
III.4.2 Intérêt vis-à-vis du rendement.....	48
IV Conclusion.....	53
Bibliographie	54
Annexe : Article présenté à la conférence Efficient 2013 (Paris 22-24 octobre 2013)	55

Table des illustrations - Figures

Figure 1: Cohérence entre les modèles et les RAD de 122 services AEP Veolia Eau	11
Figure 2: Cohérence des modèles et des RAD après réduction de l'échantillon à 107 services.....	12
Figure 3: Etude des distributions de VP et log(VP)	14
Figure 4: Pertes (m3) en fonction de la pression moyenne pondérée par la demande; 100 services AEP	18
Figure 5: Pertes (m3) en fonction de l'amplitude des distributions des pressions au sein des 100 services AEP.....	18
Figure 6: Pertes (m3) en fonction de la pression de référence des 100 services AEP.....	19
Figure 7: Distributions des pressions moyennes pondérées par la demande des 5 classes de pression	21
Figure 8: Distributions des pertes au sein des 5 classes de pression	22
Figure 9: Distribution des pertes par branchements au sein des 5 classes de pression.....	22
Figure 10: Résidus du modèle linéaire de log(pertes) selon les classes de pression : homoscedasticité et normalité.....	25
Figure 11: ILI en fonction de la pression moyenne de service.....	27
Figure 12: Diagrammes en boîte de ILI pour les 5 classes de pression.....	28
Figure 13: Distribution de VP sur l'échantillon de 926 services AEP	32
Figure 14: Distribution de log(VP) sur l'échantillon de 926 services AEP	33
Figure 15: Répartition de Log(VP) sur l'échantillon de 908 services AEP	34
Figure 16: Distribution du log du volume sortant	37
Figure 17: Distribution du log du nombre de branchements	38
Figure 18: Distribution du log du linéaire	39
Figure 19 : Distribution du log du nombre d'abonnés	40
Figure 20: Prédications et résidus du modèle n°9 de log(VP).....	43
Figure 21: Comparaison des pertes prédites par le modèle sélectionné avec les données pertes.....	44
Figure 22: Application du modèle logarithmique des pertes aux données de l'ONSEA 2009, 2010, 2011	45
Figure 23: Application du modèle des pertes aux données de l'ONSEA 2009, 2010 et 2011.....	46
Figure 24: Modèle linéaire des pertes en fonction du nombre de branchements	47
Figure 25: Modèle linéaire des pertes en fonction du nombre d'abonnés	48
Figure 26: Indice de pertes, rendement et ILC, données de Veolia Eau 2011	49

Figure 27: Indice de pertes, rendement et ILC, données de l'ONSEA 2009.....	50
Figure 28: Indice de pertes, rendement et ILC, données de l'ONSEA 2010.....	51
Figure 29: Indice de pertes, rendement et ILC, données de l'ONSEA 2011.....	52
Figure 30 : Utilisation combinée de ILI et PMI	59
Figure 31 : Exemple de résultats du programme de calcul automatique des indicateurs de pression ...	61
Figure 32 : Etude par secteurs – Volume annuel perdu en fonction du nombre de branchements.....	64
Figure 33 : Etude par secteurs – Volume de pertes par branchement en fonction de la pression moyenne	64
Figure 34 : Evaluation de P_{min} pour le secteur 2 : $P_{min} = 25.98$ m.....	66
Figure 35 : Lien entre P_{min} et $P10\%$, premier décile de la distribution des pressions	67
Figure 36 : Etude par services : Analyse de la distribution des volumes de pertes	67
Figure 37 : Etude par services : Relation entre les logarithmes du volume de pertes et du nombre de branchements.....	68
Figure 38 : Etude par services : Constitution des classes de pression.....	69
Figure 39 : Etude par services. Volume des pertes par branchement selon la classe de pression.....	69
Figure 40 : Etude à l'échelle du service. Lien entre ILI et les classes de pression.....	71

Table des illustrations - Tableaux

Tableau 1: Constitution de l'échantillon de services pour l'étude "Pression-Pertes"	10
Tableau 2: Résultats du test de Shapiro de normalité du $\log(\text{pertes})$ pour les 5 classes de pression.....	25
Tableau 3: Résultat de la méthode ANOVA - Comparaison du $\log(VP)$ entre les classes de pression	26
Tableau 4: Données issues des RAD utilisées.....	29
Tableau 5: Indicateurs calculés avec les données des RAD	30
Tableau 6: Etapes de nettoyage des données des RAD	31
Tableau 7: Résultats des tests de corrélation des variables descriptives des services avec $\log(VP)$	36
Tableau 8: Résultats de 10 essais de modèles linéaire de $\log(VP)$	41
Tableau 9: Coefficients des modèles de $\log(VP)$	41
Tableau 10: Résultats des tests de corrélation de ILC avec l'indice de pertes et le rendement, données Veolia Eau 2011	50
Tableau 11 : Hypothèses de calcul d' U_{ARL} pour une pression moyenne de 50 mce	58
Tableau 12 : Forme des indicateurs recherchés.....	60
Tableau 13 : Caractérisation des secteurs étudiés	63

Table des illustrations - Tables

Table 1: Liste des indicateurs de pression calculés par (ou à partir de) l'applicatif EPAnet.....	9
Table 2: Liste des coefficients de corrélation des grandeurs avec $\log(VP)$	17
Table 3: Liste des coefficients de corrélation des grandeurs avec $\log(VP \text{ par branchement})$	20
Table 4: Résultats des tests du caractère aléatoire des pertes, réalisé sous R.....	24
Table 5: Sortie R du test de Bartlett d'homogénéité des variances de $\log(VP)$ des classes de pression	26
Table 6: Résultat du test de corrélation de Spearman de ILI et de la pression moyenne	28
Table 7: Résultats du modèle de $\log(VP)$ n°9 donnés par R	42

I Introduction

Le présent rapport constitue le rapport final de l'étude menée par Veolia Eau et Irstea de 2011 à 2013 pour développer un nouvel indicateur de pertes des réseaux de distribution d'eau potable faisant intervenir la pression et le nombre de branchements.

L'étude a tout d'abord porté sur une approche à l'échelle du secteur du lien entre volume perdu et pressions. Deux stages ainsi qu'une étude spécifique des secteurs du SEDIF ont été réalisés par Veolia (DT puis VEDIF) sur la recherche de modèles multi-variables des pertes prenant compte, entre autres, des indicateurs de pression et le nombre de branchements. Les résultats de ces études sont détaillés dans les rapports suivants:

- Le Glatin M. (2011), Impact de la pression et des infrastructures sur les réseaux d'eau potable. Veolia Eau. 50p.
- Tangeaoui N. (2012), Introduction de la pression et du nombre de branchements dans les indicateurs de performances caractérisant les pertes en eau d'un réseau. Veolia Eau. 54p.
- Sandraz A.C. (2012), Pertes en eau : Introduction de la pression dans les indicateurs de performance. *Rapport final de VEDIF à l'intention du SEDIF*. Document interne. Veolia Eau D'Ile de France.

En 2012, F. Ben Hassen a réalisé un stage au sein d'Irstea qui a porté sur la caractérisation du potentiel de réduction de pression à l'échelle d'un secteur :

- Ben Hassen F. (2012), Caractérisation et évaluation de la pression moyenne minimale d'une zone de desserte d'un réseau d'alimentation en eau potable. Irstea, 52p.

L'ensemble des résultats obtenus à l'échelle des secteurs font l'objet du rapport de synthèse de la première année d'étude :

- Renaud E., Pillot J. (2012), Pertes en eau : introduction de la pression dans les indicateurs de performance. Rapport de synthèse de la première année d'étude. Irstea, 40p.

Cette étude a nécessité le développement de deux applicatifs EPANet permettant

- d'affecter les nœuds et les tronçons du modèle à un secteur hydraulique afin de travailler à l'échelle des secteurs, applicatif **Sectoriseur.exe**,
- d'extraire les pressions et les demandes à chaque nœud d'un modèle hydraulique et à chaque pas de temps de la simulation, puis de calculer des indicateurs de la pression d'une zone (service ou secteurs si définis), applicatif **CalculPression.exe**.

Ces applicatifs ont été développés à Irstea par D. Gilbert, sur la base d'un cahier des charges défini avec Veolia Eau et sont accompagnés par un mode d'emploi :

- Tangeaoui N., Sandraz A.C. (2012), L'automatisation du calcul de la pression moyenne par secteur sur EPANet. Document interne Veolia Eau, 26p.
- Pillot J. (2013), Utilisation de l'applicatif Epanet « CalculPression_3-4 » créé par Denis Gilbert. Irstea, 4p.
- Pillot J. (2013), Utilisation de l'applicatif Epanet « Sectoriseur_1-1 » créé par Denis Gilbert. Irstea, 6p.

Ces travaux sont synthétisés au sein du rapport intermédiaire de la seconde année d'étude

- Renaud E., Pillot J. (2012), Pertes en eau : introduction de la pression dans les indicateurs de performance. Rapport intermédiaire de la seconde année d'étude. Irstea, 27p

Le travail à l'échelle des secteurs hydrauliques n'a pas permis de définir un modèle satisfaisant du volume perdu et par là, un indicateur de performance prenant en compte la pression du secteur. Une des limites identifiées étant la fiabilité de l'évaluation des pertes annuelles des secteurs, nous avons poursuivi l'étude à l'échelle des services AEP, dont nous présentons les résultats dans le présent rapport final.

L'ensemble des travaux réalisé est synthétisé dans un article annexé au présent rapport.

- Renaud E, Clauzier M., Sandraz A.C., Pillot J., Gilbert D. (2013), Introduction de la pression et du nombre de branchements dans les indicateurs français de pertes des réseaux d'eau potable. Irstea, 15p.

II Rôle de la pression de service dans le volume de pertes

II.1 Services AEP étudiés

II.1.1 Calcul des indicateurs de pression de service

Une des façons d'estimer la pression hydraulique à l'échelle d'un service d'eau potable est de s'appuyer sur le modèle hydraulique (Sissoko, 2010).

Nous avons développé en interne un applicatif (Renaud et Pillot, 2012) permettant:

- d'extraire d'une simulation hydraulique de 24h sous EPAnet les demandes (m³/h) et les pressions (m) à chaque nœud du modèle et à chaque pas de temps,
- de calculer à partir de ces données des indicateurs de la distribution des pressions modélisées, (Table 1).

Indicateur de pression	Notation	Formule
Pression moyenne (m)	p_{moy}	$\frac{\sum P_i \times (t_i - t_{i-1})}{\sum t_i}$
Pression minimale atteinte à un nœud à un pas de temps (m)	p_{crit}	
Pression moyenne pondérée par les 1/2 longueurs (m)	p_{moyL}	$\frac{\sum P_i \times 0.5 \times length_i}{\sum length_i}$
Pression moyenne pondérée par la demande (m)	p_{moyD}	$\frac{\sum P_i \times demand_i}{\sum demand_i}$
1 ^{er} décile de la distribution des pressions (m)	P_{10}	
Médiane de la distribution des pressions (m)	P_{50}	
9 ^{em} décile de la distribution des pressions (m)	P_{90}	
Amplitude de la distribution des pressions entre les 1er et 9e déciles (m)	ΔP	$P_{90} - P_{10}$
Pression de référence techniquement atteignable (m) (Ben Hassen, 2012)	P_{ref}	$p_{moyD} - 0.9 \times (P_{10} - 20)$
Indice de réduction de pression	IRP	$\frac{p_{moyD}}{P_{ref}}$

Table 1: Liste des indicateurs de pression calculés par (ou à partir de) l'applicatif EPAnet

II.1.2 Constitution de l'échantillon

L'échantillon que nous étudions dans ce chapitre est l'ensemble des services AEP exploités par Veolia Eau en France pour lesquels nous avons pu obtenir un modèle hydraulique EPAnet fonctionnant correctement sous la dernière version du logiciel et donnant des résultats cohérents avec les données issues des RAD de 2011.

II.1.2.1 Test des modèles hydrauliques

Les simulations sont réalisées avec les options suivantes pour l'ensemble des modèles :

- Durée de simulation : 24h, début à 12 AM
- Précision des calculs : 0.001
- Arrêt de la simulation si le calcul ne converge pas

Le Tableau 1 ci-dessous récapitule la constitution de l'échantillon de services pour lesquels le modèle hydraulique fonctionne correctement et dont des données fiables des RAD de 2011 nous ont été fournies.

<i>mis à jour le 14/03/2013</i>		Modèles Hydrauliques					<i>Total</i>
		0- Bon (pas de problèmes)	1- Ne converge pas	2- Pressions négatives	3- Autres erreurs	Abandon du modèle	
RAD	Contrat Perdu/Prestation					9	9
	Pas de code d'id du RAD	2	0	0	0	0	2
	Données RAD non fournies	7	2	0	1	0	10
	Données RAD douteuses	0	2	2	0	0	4
	Données RAD cohérentes	123	11	8	10	0	152
	<i>Total</i>	<i>132</i>	<i>15</i>	<i>10</i>	<i>11</i>	<i>9</i>	<i>177</i>

Tableau 1: Constitution de l'échantillon de services pour l'étude "Pression-Pertes"

De plus, un des 123 services réunissant les conditions est écarté car il s'agit manifestement d'un syndicat de d'achat/vente d'eau avec 5 abonnés mais 0 habitants et 0 branchements renseignés.

A ce stade, notre échantillon est de 122 services AEP.

II.1.2.2 Cohérence entre les modèles et les RAD

Nous testons les services sur la cohérence des grandeurs suivantes :

- **Linéaire de canalisation** : Le linéaire de réseau, noté L, en m, donné dans le RAD doit s'approcher de la somme des longueurs des tronçons du modèle (lenght), en m également.
- **Nombre de points de livraison** : Les branchements, noté brt dans les données RAD, sont représentés par des nœuds (notés Nodes) dans le modèle.
- **Volume sortant** : Il est la somme du volume consommé autorisé et du volume exporté enregistrés dans les RAD. Il est modélisé par la somme des demandes en eau aux nœuds du modèle du réseau, ramenée à 365 jours de fonctionnement.

Sur la Figure 1, on constate que l'hypothèse d'égalité des linéaires de réseau est, en général, satisfaite mais que cela n'est pas le cas pour les branchements.

Le nombre de nœuds des modèles est de l'ordre de 20% du nombre de branchements. Cela s'explique par la modélisation des consommations de plusieurs branchements proches par un unique nœud. Il faut de plus noter un usage des nœuds de construction (avec une demande de base nulle) plus ou moins fréquent selon les auteurs des modèles. C'est pourquoi nous n'utiliserons pas le nombre de points de livraison d'eau comme critère de cohérence des modèles et des RAD.

Il semble qu'il y ait effectivement équivalence entre le volume sortant et la demande totale annualisée du modèle, sauf chez un petit nombre de services (15) dont l'écart entre les deux grandeurs est atypique du point de vue de la distribution des écarts sur l'ensemble de l'échantillon. Il s'agit des services représentés par des ronds sur le diagramme en boîte de la Figure 1 (lire ci-dessous le paragraphe *Lecture d'un diagramme en boîte*). La demande étant le facteur de pondération utilisé dans le calcul de la pression moyenne, nous retirons ces services de notre étude.

Lecture des diagrammes en boîte : Le diagramme en boîte permet de représenter les caractéristiques d'une série statistique. La boîte est bornée par le 1^{er} quartile et le 3^e quartile de la distribution, la médiane est représentée par un trait fort. Les extensions en pointillés s'étendent jusqu'à la dernière valeur distante de moins de 1.5 fois la longueur de la boîte à partir des bords de la boîte. Les

occurrences dont les valeurs sont au-delà des pointillés sont repérées individuellement par un point et méritent un examen pour déterminer s'il y a lieu de les considérer atypiques.

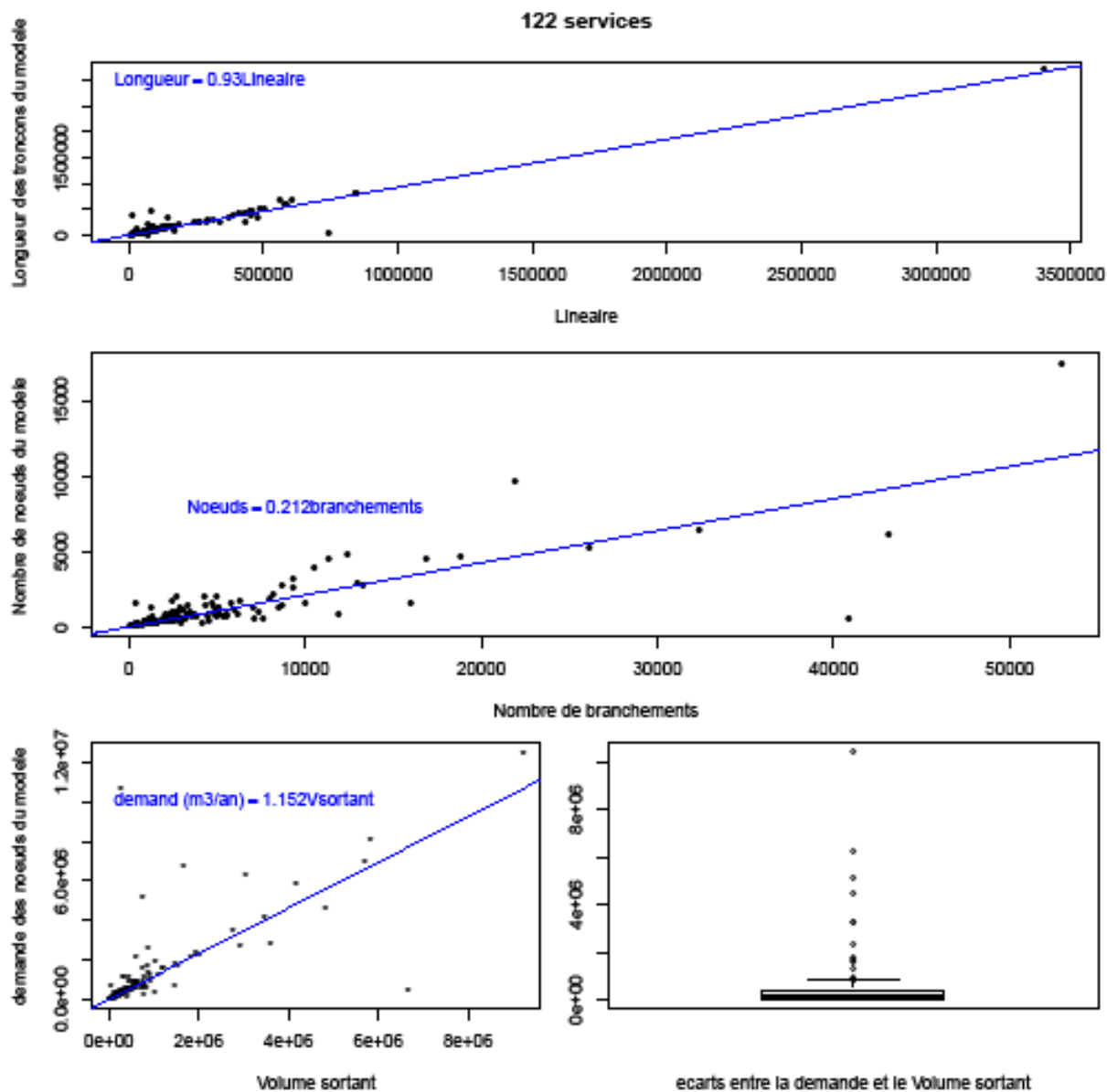


Figure 1: Cohérence entre les modèles et les RAD de 122 services AEP Veolia Eau

On constate sur la Figure 2 que le linéaire et le volume sortant suivent maintenant des droites de tendance de pente 1 et que la répartition des écarts entre la demande totale et le volume sortant est plus homogène.

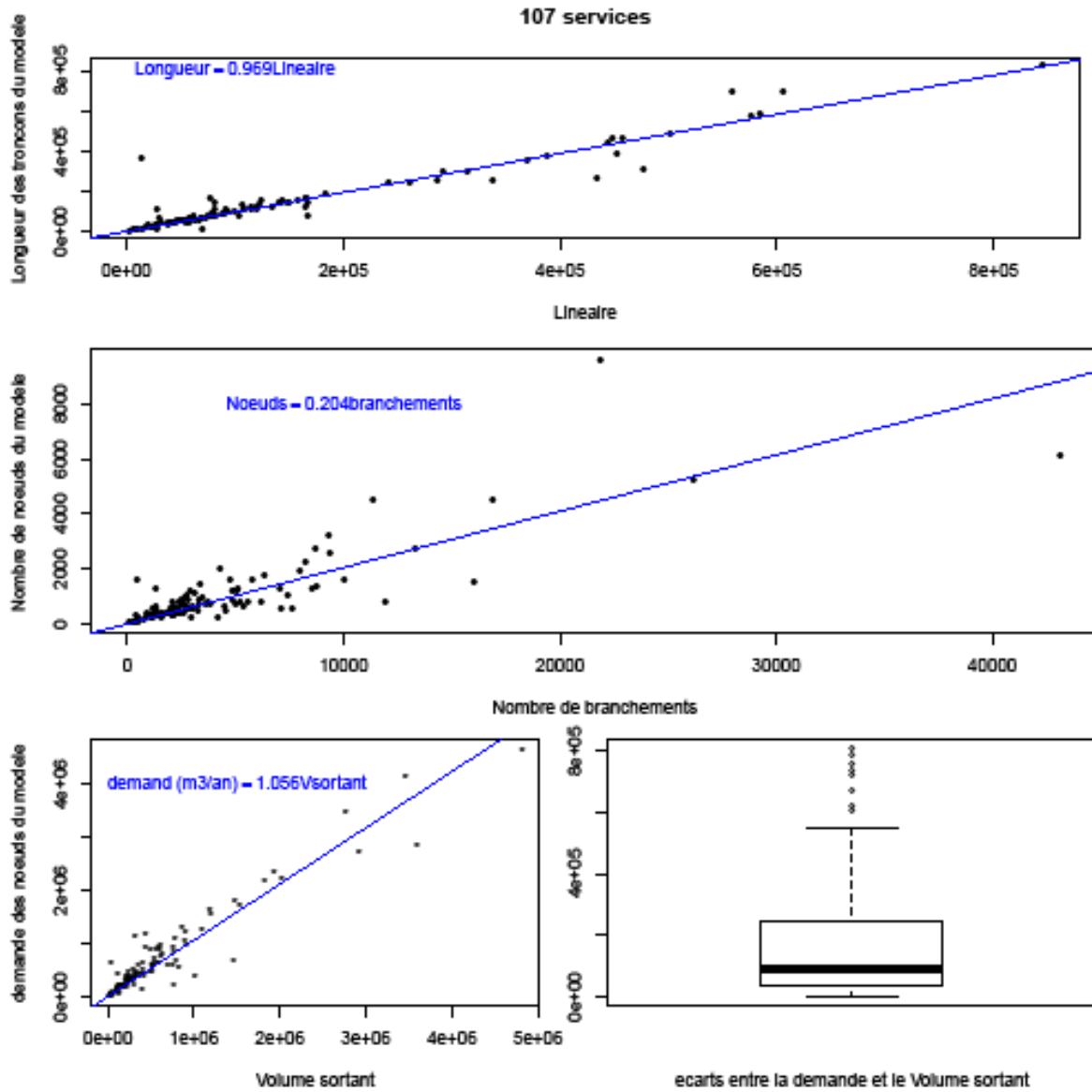


Figure 2: Cohérence des modèles et des RAD après réduction de l'échantillon à 107 services

II.2 Etude de l'impact de la pression sur les pertes

II.2.1 Etude de la distribution des pertes

Il s'agit de s'assurer que la distribution des pertes de notre échantillon de services satisfait les conditions nécessaires à son étude statistique, à savoir :

- la relative symétrie de la distribution (voire la normalité de la distribution),
- l'absence de valeurs extrêmes atypiques qui peuvent influencer fortement les résultats de certaines méthodes statistiques basées sur le carré d'une distance.

Or, la distribution des pertes de notre échantillon apparaît nettement asymétrique avec des valeurs extrêmes atypiques, sur le diagramme en boîte de gauche de la Figure 3. La représentation des quantiles de cette distribution selon les quantiles de la loi normale à gauche sur la Figure 3 **Erreur ! Source du renvoi introuvable.** (lire ci-dessous le paragraphe Lecture du graphique Quantile-Quantile) montre que VP ne suit pas une loi normale. Cette distribution est plutôt de type exponentiel, c'est pourquoi nous choisissons d'effectuer une transformation logarithmique sur VP.

La distribution de $\log(VP)$ s'approche d'une distribution symétrique mais les valeurs les plus faibles sont atypiques (lire le paragraphe Lecture d'un diagramme en boîte), au centre de la Figure 3. Nous choisissons de réduire notre échantillon à 100 services afin de travailler sur une distribution symétrique et sans valeurs atypiques (graphique quantile-quantile et diagramme en boîte à droite de la Figure 3).

Lecture du graphique Quantile-Quantile (QQ plot): Ce graphique permet de comparer la distribution de l'échantillon étudié à une distribution théorique. Pour cela, il trace les quantiles de la variable étudiée en fonction des quantiles de la loi normale. Ainsi, le nuage de points est aligné sur une droite dans le cas où la variable étudiée est distribuée selon cette loi normale.

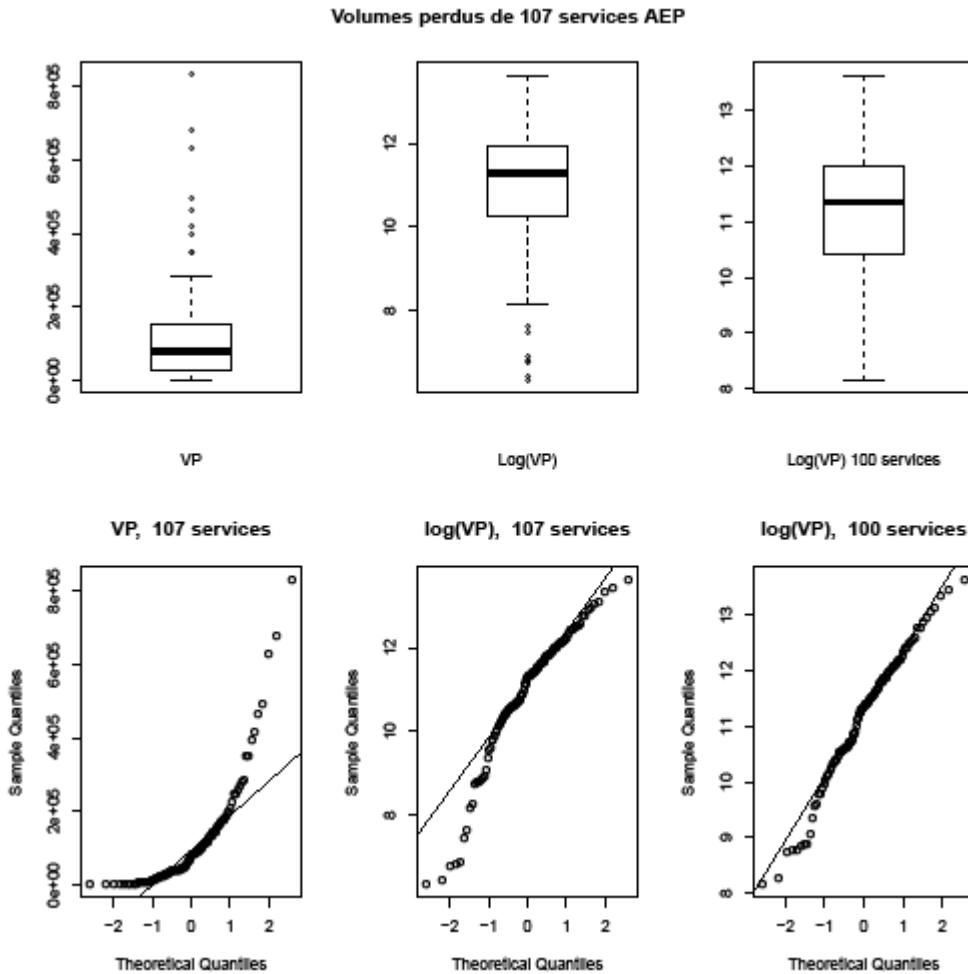


Figure 3: Etude des distributions de VP et log(VP)

II.2.2 Etude statistique du lien entre les indicateurs de pression et les pertes

Parmi les données et les indicateurs du service disponibles (issus des RAD et calculés avec les modèles hydrauliques) qui n’incluent pas les volumes de pertes, nous cherchons à identifier ceux ayant un lien avec le log(VP).

Interprétation des tests de corrélation de Pearson et de Spearman (Rakatomalala, 2012):

Les coefficients de corrélation de Pearson et de Spearman permettent de tester si deux séries sont corrélées entre elles dans les hypothèses respectives où la relation serait linéaire ou monotone.

- Méthode paramétrique : Bravais-Pearson

Soient $X(x_1, \dots, x_n)$ et $Y(y_1, \dots, y_n)$ deux séries, supposées normalement distribuées, dont on souhaite caractériser la liaison au moyen du coefficient de corrélation de Bravais-Pearson.

$r_p = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$ est le coefficient de corrélation, rapport entre :

- la covariance de X et Y :

$$\sigma_{xy} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x}) \cdot (y_i - \bar{y})$$

- et le produit des écart-types des deux séries σ_x et σ_y :

$$\sigma_x = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2} \quad \sigma_y = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2}$$

Avec \bar{x} et \bar{y} les moyennes de X et Y

r_p est compris entre -1 et 1. Le signe indique si la corrélation linéaire est croissante (+) ou décroissante (-). La valeur reflète l'intensité du lien entre les variables. Dans le cas où les séries suivent une loi normale bi-variée, $r_p=0$ signifie que les deux séries sont indépendantes.

r_p^2 est le coefficient de corrélation généralement noté R^2 . Il s'interprète comme la proportion de la variance de l'une des variables linéairement expliquée par l'autre variable.

Le test de corrélation de Pearson a pour hypothèses :

$H_0 : r_p = 0$, c'est-à-dire qu'il n'y a pas de corrélation entre les deux séries.

H_1 , l'hypothèse alternative $r_p \neq 0$ signifie qu'il existe une corrélation entre les deux séries.

Sous H_0 , la statistique de test t suit une loi de Student¹ à $(n - 2)$ degrés de liberté. $t = \frac{r_p}{\sqrt{\frac{1-r_p^2}{n-2}}}$;

La région critique de rejet de l'hypothèse H_0 avec un risque α est

R.C. : $|t| > t_{1-\frac{\alpha}{2}} \times (n-2)$, quantile d'ordre $1-\frac{\alpha}{2}$ de la loi de Student à $n-2$ degrés de liberté.

Les tests statistiques donnent une valeur, notée p -value, que l'on doit comparer au seuil α que l'on s'est fixé. Une p -value inférieure à α entraîne le rejet de l'hypothèse H_0 avec un risque d'erreur α .

- Méthode non paramétrique de Spearman

Le coefficient de corrélation de Spearman est identique au coefficient de Pearson sinon qu'il substitue aux éléments de X et Y leurs rangs dans chacune des séries. Il s'agit d'une méthode non paramétrique, c'est-à-dire qu'elle ne repose pas sur l'hypothèse de normalité des séries.

Cette méthode est plus robuste vis-à-vis des valeurs atypiques. Une valeur proche de 1 du coefficient de corrélation signifie qu'il existe une relation monotone, mais pas nécessairement linéaire.

Le test statistique de Spearman teste l'égalité (H_0) ou la différence (H_1) du coefficient de corrélation à 0. Il s'interprète de la même manière que le test de Pearson, à l'aide de la statistique de test t exprimée avec le coefficient de Spearman et comparée au quantile d'ordre $1-\frac{\alpha}{2}$ de la loi de Student à $n-2$ degrés de liberté. La p -value permet de rejeter ou d'accepter l'hypothèse H_0 avec le risque d'erreur α .

¹ Loi de t-student : La loi de Student est une loi de probabilité, faisant intervenir le quotient entre une variable suivant une loi normale centrée réduite et la racine carrée d'une variable distribuée suivant la loi du χ^2 .

Pour identifier les variables liées à $\log(VP)$ nous commençons par calculer les coefficients de corrélation de Pearson et de Spearman et les tests associés (avec un seuil $\alpha=0.05$) à ces deux méthodes sur les variables et leurs logarithmes.

On constate dans la Table 2 que la transformation logarithmique se justifie sur un grand nombre de variables (augmentation du coefficient de corrélation et diminution de la p-value associée au test de Pearson, approchant des résultats obtenus avec la méthode de Spearman), mais pas pour le diamètre nominal moyen, le nombre d'abonnés par km et le taux de fuites des branchements.

La méthode de Spearman est basée sur le rang des observations dans l'échantillon, les résultats ne dépendent pas d'une transformation de la variable. Cette méthode est donc plus robuste.

On ne peut pas exclure l'absence de corrélation de $\log(VP)$ avec les variables en rouge dans la Table 2. On constate que les trois méthodes utilisées ne permettent pas de conclure à une corrélation des indicateurs de pression, surlignés en gris, avec $\log(VP)$.

Méthode	Pearson		Pearson		Spearman	
	Log(VP)-Variable		log(VP)-log(Variable)		Log(VP)-rang (Variable)	
Variables	Coef. de corrélation	p-value	Coef. de corrélation	p-value	Coef. de corrélation	p-value
Volume consommé autorisé	0.67	3.9E-14	0.86	0.00	0.86	5.2E-31
Volume sortant	0.60	4.0E-11	0.83	0.00	0.86	3.5E-30
Volume consommé compté	0.66	7.7E-14	0.85	0.00	0.86	4.5E-30
Demande en eau	0.62	4.8E-12	0.79	0.00	0.80	9.2E-24
Nombre d'abonnés	0.60	2.7E-11	0.64	0.00	0.80	5.2E-23
Nombre d'habitants	0.60	3.1E-11	0.79	0.00	0.79	8.2E-23
Volume de service	0.54	8.1E-09	0.76	0.00	0.77	1.5E-20
Nombre de branchements	0.56	9.8E-10	0.77	0.00	0.75	3.7E-19
Nombre de fuites sur branchements	0.64	1.1E-12	0.74	0.00	0.71	7.7E-17
Linéaire en fonte	0.60	4.9E-08	0.53	0.00	0.71	6.5E-12
Infrastructure Leakage Index	0.46	1.6E-06	0.63	0.00	0.59	7.6E-11
Indice de Pertes Structurelles	0.36	2.2E-04	0.61	0.00	0.57	7.2E-10
Nombre de fuites sur canalisations	0.50	1.1E-07	0.54	0.00	0.55	2.3E-09
Linéaire de réseau	0.50	1.2E-07	0.60	0.00	0.55	2.4E-09
Longueur des tronçons du modèle	0.44	3.6E-06	0.46	0.00	0.45	2.4E-06
Diamètre Nominal moyen	0.35	3.2E-03	0.10	0.41	0.40	7.1E-04
Indice Linéaire de Consommation [au sens du Grenelle]	0.36	2.8E-04	0.34	0.00	0.38	1.0E-04
Volume consommé non compté	0.41	2.5E-05	0.77	0.00	0.38	1.1E-04
Indice Linéaire de Consommation	0.38	8.1E-05	0.34	0.00	0.37	1.6E-04
Nombre d'abonnés par km de réseau	0.32	1.3E-03	0.18	0.08	0.32	1.0E-03
Volume exporté	0.19	6.4E-02	0.44	0.01	0.26	7.8E-03
Nombre de branchements au km de réseau	0.20	4.8E-02	0.21	0.03	0.25	1.3E-02
Taux de fuites des branchements (pour 1000 branchements)	0.27	6.6E-03	0.19	0.07	0.20	4.7E-02
Taux de fuites des canalisations (par km de canalisations)	0.08	4.5E-01	0.03	0.77	0.13	2.0E-01
Linéaire en acier	-0.08	5.0E-01	-0.41	0.03	0.04	7.6E-01
Indice de Réduction de Pression	-0.07	4.8E-01	-0.05	0.61	0.01	9.1E-01
9 ^e décile des pressions	-0.06	5.8E-01	-0.04	0.69	-0.01	9.2E-01
Ecart entre le 9 ^e et le 10 ^e déciles des pressions	-0.03	7.3E-01	-0.01	0.92	-0.02	8.7E-01
Pression de référence	-0.10	3.1E-01	-0.04	0.69	-0.03	7.7E-01
1 ^{er} décile des pressions	-0.05	6.3E-01	-0.05	0.63	-0.05	6.0E-01
Pression moyenne pondérée par la demande	-0.12	2.2E-01	-0.09	0.38	-0.07	5.2E-01
Pression moyenne pondérée par le linéaire	-0.11	2.6E-01	-0.09	0.37	-0.07	4.7E-01
Pression moyenne	-0.11	2.9E-01	-0.10	0.33	-0.11	2.6E-01
Médiane des pressions moyennes	-0.11	2.6E-01	-0.11	0.29	-0.14	1.7E-01
Pression minimale au point critique	-0.20	4.4E-02	-0.15	0.13	-0.16	1.2E-01
Taux de branchements par abonnés	0.03	7.7E-01	-0.06	0.57	-0.24	1.8E-02

Table 2: Liste des coefficients de corrélation des grandeurs avec $\log(VP)$

Ce constat est renforcé par l'absence de forme des nuages de points (Figure 4, Figure 5, Figure 6) entre les pertes et :

- La pression moyenne pondérée par la demande, notée $pmoyD$:

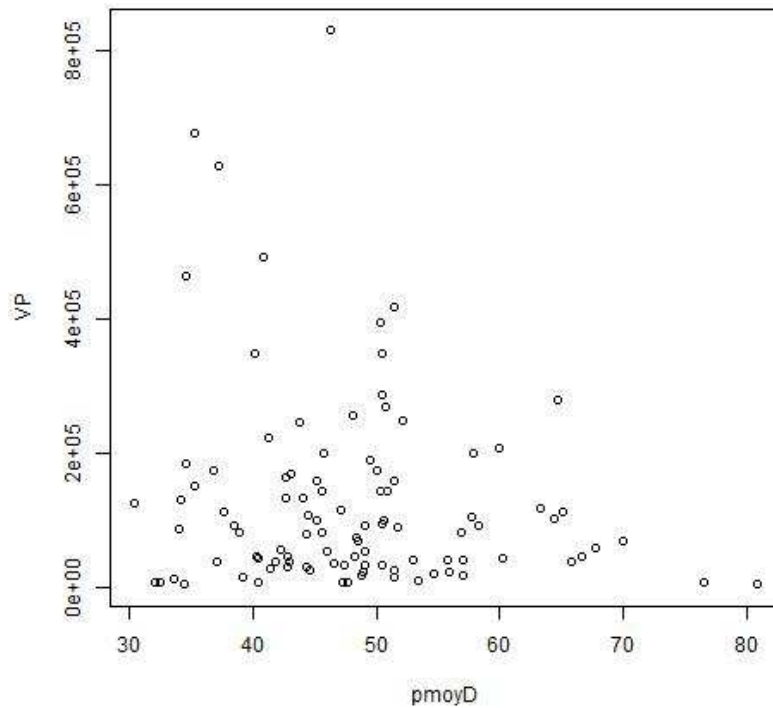


Figure 4: Pertes (m3) en fonction de la pression moyenne pondérée par la demande; 100 services AEP

- L'amplitude de la distribution des pressions, notée ΔP :

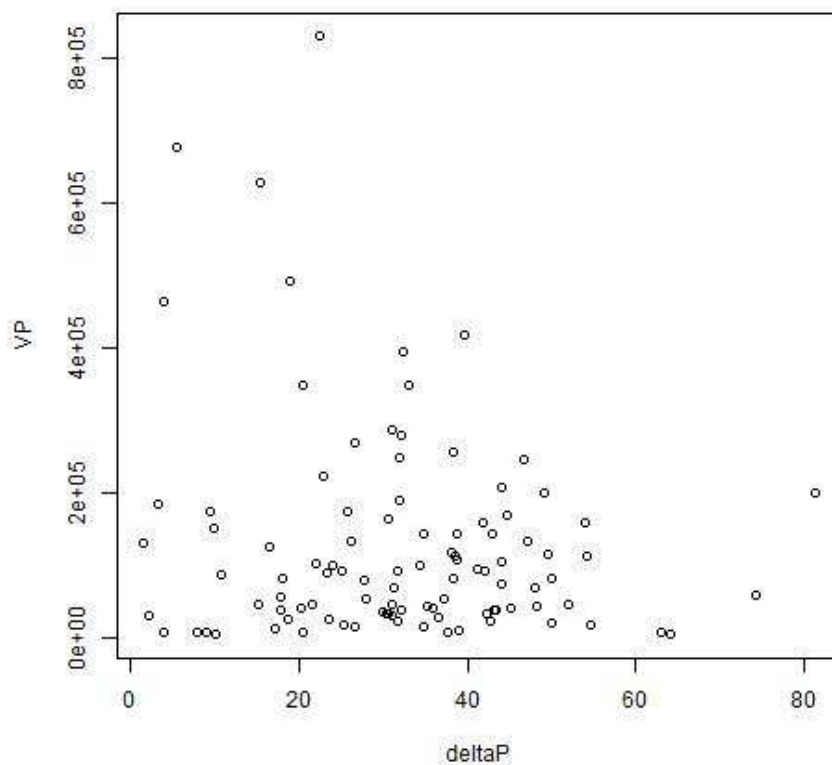


Figure 5: Pertes (m3) en fonction de l'amplitude des distributions des pressions au sein des 100 services AEP

- Le 9^{ème} décile de la distribution des pressions, notée P90 :

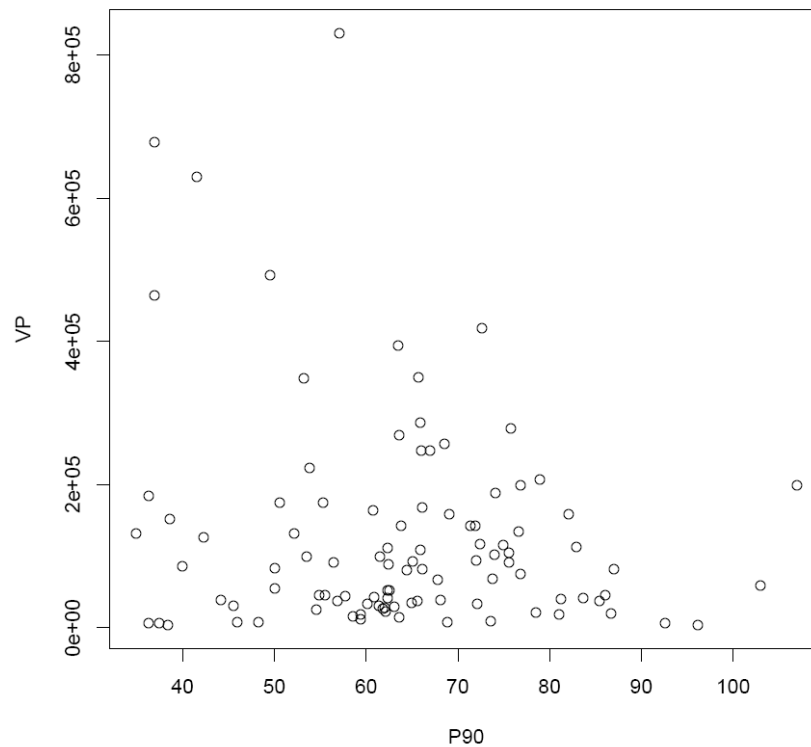


Figure 6: Pertes (m3) en fonction de la pression de référence des 100 services AEP

Afin de s'assurer que le lien entre les pertes et les indicateurs de pression n'est pas occulté par l'effet de taille, nous testons les corrélations des variables avec le logarithme des pertes par branchement, Table 3. A nouveau, aucune corrélation entre les indicateurs de pression ou leurs logarithmes avec les pertes par branchements n'a pu être mise en évidence.

Méthode	Pearson Log(VP/brcht)~Variable		Pearson log(VP/brt)~log(Variable)		Spearman Log(VP/brcht)~rang (Variable)	
	Coef. de corrélation	p-value	Coef. de corrélation	p-value	Coef. de corrélation	p-value
Variables						
Indice de Pertes Structurelles	0.666	0.000	0.971	0.000	0.960	0.000
Infrastructure Leakage Index	0.768	0.000	0.929	0.000	0.906	0.000
Nombre de fuites sur branchements	0.529	0.000	0.498	0.000	0.410	0.000
Volume de service	0.395	0.000	0.421	0.000	0.407	0.000
Diamètre Nominal moyen	0.317	0.008	0.143	0.244	0.376	0.001
Indice Linéaire de Consommation	0.444	0.000	0.344	0.000	0.342	0.001
Indice Linéaire de Consommation [au sens du Grenelle]	0.327	0.001	0.309	0.002	0.321	0.001
Taux de fuites des branchements (pour 1000 branchements)	0.322	0.001	0.360	0.000	0.309	0.002
Taux de fuites des canalisations (par km de canalisations)	0.247	0.013	0.254	0.014	0.255	0.010
Nombre d'abonnés par km de réseau	0.303	0.002	0.215	0.031	0.229	0.022
Demande en eau	0.093	0.357	0.228	0.022	0.227	0.023
Volume consommé autorisé	0.098	0.330	0.217	0.030	0.183	0.069
Linéaire en fonte	0.139	0.255	0.164	0.211	0.181	0.137
Volume sortant	0.048	0.636	0.192	0.055	0.178	0.076
Volume consommé compté	0.089	0.377	0.209	0.037	0.173	0.084
Nombre de branchements par km de réseau	0.032	0.752	0.048	0.633	0.110	0.278
Volume consommé non compté	0.486	0.000	0.423	0.002	0.099	0.326
Nombre d'abonnés	0.000	1.000	0.140	0.165	0.092	0.362
Nombre d'habitants	0.038	0.710	0.114	0.261	0.069	0.496
Nombre de fuites sur canalisations	0.032	0.755	0.115	0.268	0.046	0.647
Pression de référence	-0.048	0.635	-0.026	0.796	0.012	0.908
Linéaire en acier	-0.056	0.648	-0.439	0.015	0.007	0.952
P90-P10	-0.033	0.741	-0.088	0.384	0.005	0.960
Volume exporté	-0.083	0.414	0.050	0.788	-0.013	0.895
Nombre de branchements	-0.093	0.360	-0.018	0.861	-0.028	0.779
P90 : 9 ^e décile des pressions	-0.092	0.360	-0.117	0.246	-0.032	0.749
Linéaire de réseau	-0.112	0.269	-0.046	0.651	-0.090	0.373
Pression moyenne pondérée par la demande	-0.124	0.218	-0.125	0.216	-0.091	0.368
Indice de Réduction de Pression	-0.133	0.187	-0.125	0.215	-0.092	0.361
Longueur des tronçons du modèle	-0.111	0.270	-0.106	0.296	-0.115	0.256
Pression moyenne pondérée par le linéaire	-0.140	0.164	-0.147	0.145	-0.131	0.195
Pression minimale au point critique	-0.046	0.648	-0.056	0.585	-0.136	0.176
Pression moyenne	-0.124	0.218	-0.141	0.162	-0.137	0.175
P10 : 1 ^{er} décile des pressions	-0.145	0.149	-0.157	0.120	-0.141	0.163
P50 : médiane des pressions moyennes	-0.111	0.273	-0.130	0.196	-0.156	0.120
Taux de branchements par abonnés	-0.092	0.364	-0.218	0.029	-0.171	0.089

Table 3: Liste des coefficients de corrélation des grandeurs avec $\log(VP \text{ par branchement})$

II.2.3 Etude statistique par classe de pression

II.2.3.1 Découpage et observation des distributions des classes

Afin de s'assurer qu'à pression équivalente des services ont effectivement des niveaux de pertes variables et que cette variation n'est pas imputable au niveau de la pression moyenne, nous effectuons un découpage de notre échantillon en classes de 20 services de pressions moyennes les plus proches.

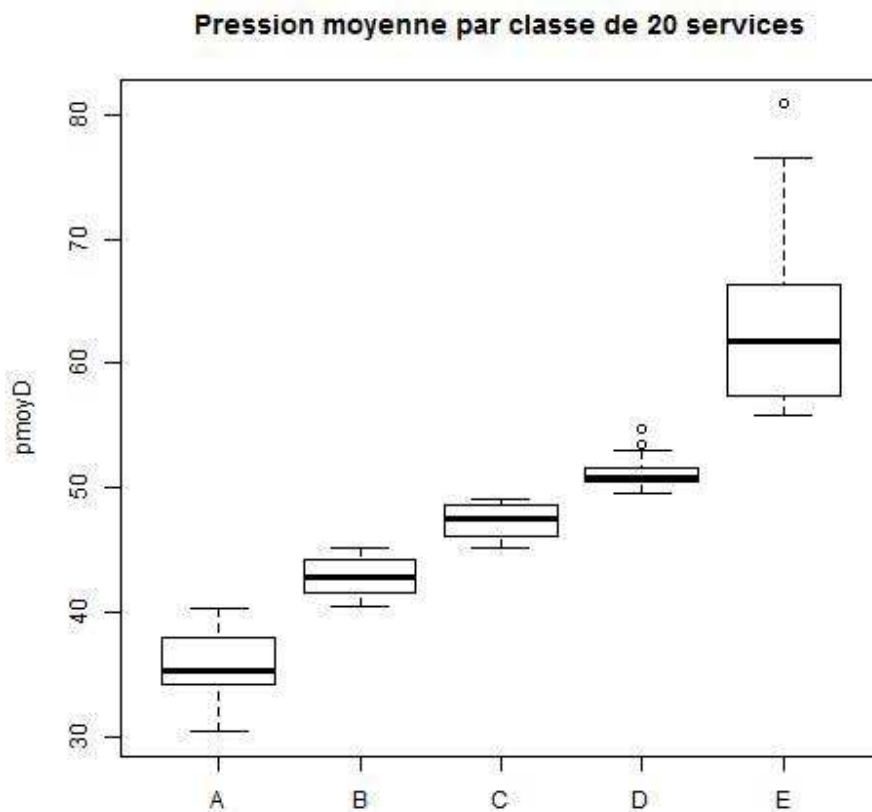


Figure 7: Distributions des pressions moyennes pondérées par la demande des 5 classes de pression

On constate que les distributions des pressions moyennes sont relativement symétriques pour chaque classe, Figure 7, avec des amplitudes plus importantes pour les extrêmes et très resserrées au centre. Les classes représentent bien 5 niveaux croissants de pression moyenne.

Les Figure 8 et Figure 9 représentent respectivement les distributions des pertes et des pertes par branchement au sein de ces mêmes classes. Visuellement on ne constate pas 5 niveaux de pertes. La classe D, avec une très faible amplitude des pressions moyennes (Figure 7), a la plus forte disparité des pertes et la médiane la plus élevée. Le minimum et le premier quartile des pertes sont en outre équivalents pour toutes les classes.

Il ne semble pas y avoir de lien entre la pression moyenne d'un service et ses pertes.

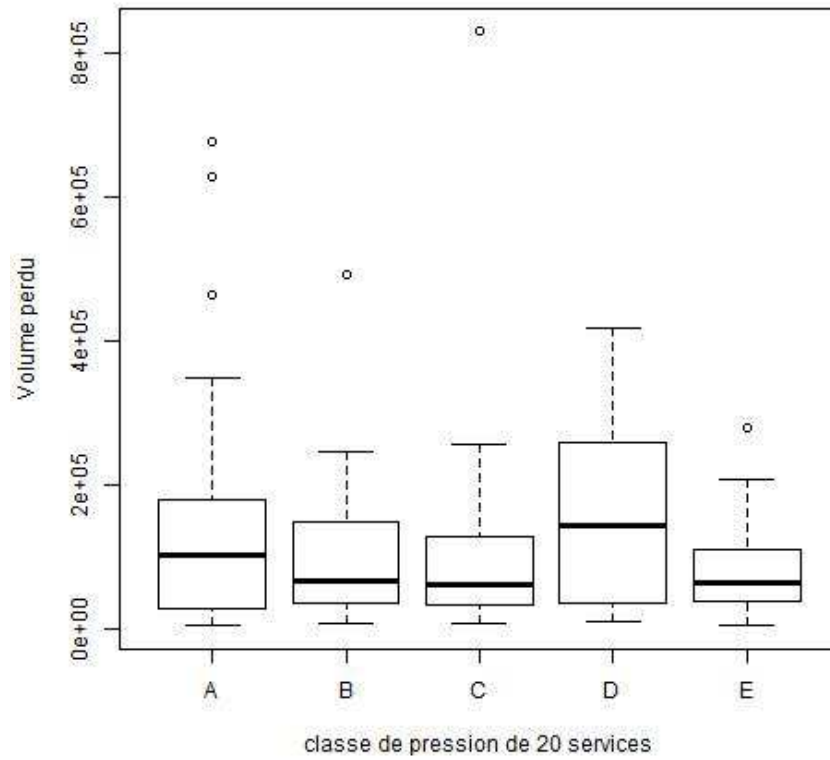


Figure 8: Distributions des pertes au sein des 5 classes de pression

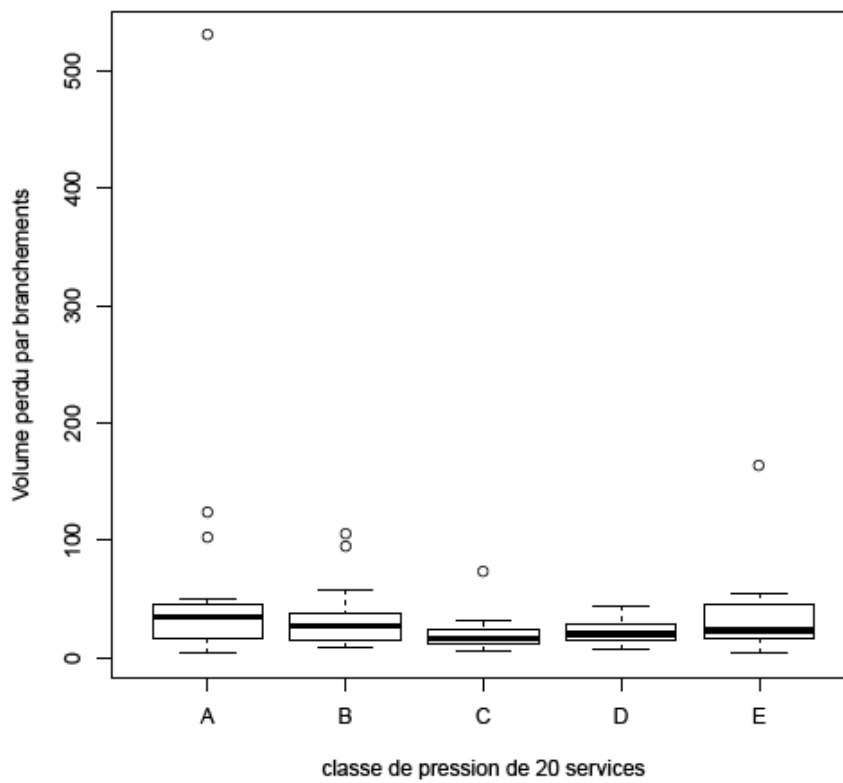


Figure 9: Distribution des pertes par branchements au sein des 5 classes de pression

II.2.3.2 Réalisation d'une ANOVA

Nous allons confirmer notre intuition par la réalisation d'une analyse de variance des classes (méthode « **AN**alysis **O**f **V**ariance »). L'ANOVA teste l'influence d'un facteur sur une variable.

Méthode de l'ANOVA (Bienvenue, 2011):

Soit F un facteur, variable quantitative, à k niveaux. Soient k échantillons de taille n_i avec $i \in [1..k]$, **indépendants** et de **loi normale**. Soit une variable quantitative **aléatoire** Y à étudier, série constituée de $n = \sum_{i=1}^k n_i$ observations. Y suit une loi normale $N(m_i, \sigma^2)$. Pour chaque niveau on calcule la moyenne de Y_i de l'échantillon. L'écart-type (théorique) est le même pour tous les niveaux. La moyenne (théorique) peut varier avec le niveau.

$$Y_{ij} = m_i + \varepsilon_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

où μ est l'effet moyen, α_i est l'effet du niveau i du facteur F , et ε_{ij} le résidu de la réalisation Y_{ij} suivant une loi normale $N(0, \sigma^2)$.

On a de plus la contrainte $\sum_{i=1}^k n_i \alpha_i = 0$.

Le problème se résume au modèle linéaire $Y = \beta X + \varepsilon$ avec

- Y le vecteur des n observations Y_{ij} ,
- β le vecteur de taille i des moyennes de Y des i échantillons,
- X une matrice de taille (n, i) contenant à chaque ligne des 0, sauf un 1 à la $i^{\text{ème}}$ colonne selon la valeur de i de l'observation Y_i ,
- ε le vecteur des résidus ε_{ij} .

On cherche des estimateurs de μ , α , σ^2 qui minimisent l'erreur du modèle et qui permettent de décomposer la variabilité de Y :

variabilité totale (ST) = variabilité résiduelle (SR) + variabilité due au modèle (SM)

Le résultat de l'ANOVA se présente dans un tableau :

Source de variation	ddl	S.C.	Carré moyen
Régression	$k-1$	SM	$SM/(k-1)$
Résiduelle	$n-k$	SR	$SR/(n-k)$
Totale	$n-1$	ST	

L'ANOVA test l'hypothèse $H_0: \forall i, \alpha_i = 0$, qui est équivalent à $\forall i, m_i = 0$

Le test utilise la statistique de test $Z = \frac{SM/(k-1)}{SR/(n-k)}$ qui doit suivre une loi de Fisher $F(k-1, n-k)$ sous H_0 .

a) Hypothèses à vérifier sur la variable testée : les pertes

La première hypothèse que doit satisfaire l'échantillon des volumes perdus de nos 100 services est d'être une série **aléatoire**. Ce caractère est vérifié par les tests de Bartels et de Runs. Les p-value des tests > 5%, Table 4, permettent de conclure positivement.

Bartels Test - Two sided

data: pertes\$VP

Standardized Bartels Statistic = -1.0365, RVN Ratio = 1.793, p-value = 0.2999

Runs Test - Two sided

data: pertes\$VP

Standardized Runs Statistic = -0.8041, p-value = 0.4214

Table 4: Résultats des tests du caractère aléatoire des pertes, réalisé sous R

Comme nous avons déjà vérifié les hypothèses de normalité et de valeurs atypiques de la série de $\log(VP)$ que nous étudions, nous pouvons définir le modèle linéaire de $\log(VP)$ en fonction des classes de pression.

b) Modèle linéaire $\log(VP)$ ~classes de pression : Contrôle des hypothèses

L'ANOVA nécessite que les **résidus** du modèle soient **normaux** et de **variances homogènes** selon les classes. Graphiquement, ces deux conditions semblent être satisfaites, respectivement à droite et à gauche de la Figure 10.

Lecture du graphique des résidus : Les résidus sont les valeurs des écarts entre la valeur de la variable étudiée et la valeur prédite par le modèle. Le graphique des résidus en fonction des valeurs prédites par le modèle permet de mettre en évidence des résidus trop grands (on essaye usuellement d'avoir des résidus entre -2 et 2) et/ou fonction des valeurs prédites (hétéroscédasticité² montrant qu'un facteur caché n'est pas pris en compte dans le modèle).

² La « scédasticité » désigne le caractère constant (homo-) ou non (hétéro-) de la variance de l'erreur de la variable aléatoire

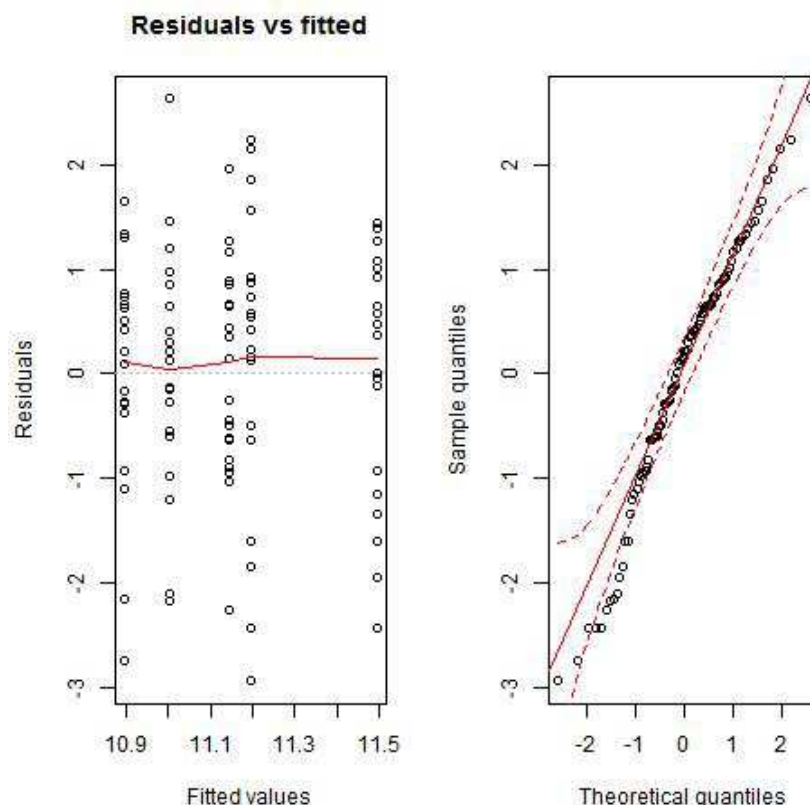


Figure 10: Résidus du modèle linéaire de log(perces) selon les classes de pression : homoscedasticité et normalité

Nous réalisons deux tests statistiques pour confirmer l'impression visuelle.

Le test de normalité de Shapiro-Wilk est réalisé sur chaque classe de pressions, Tableau 2, et ne remet pas en cause leur normalité.

Test de Normalité de Shapiro-Wilk :

Ce test a pour hypothèse H_0 : « L'échantillon testé est issu d'une population normalement distribuée ». Il a pour statistique de test W dont une valeur trop petite entraîne le rejet de l'hypothèse H_0 . Usuellement, on rejette H_0 avec un risque $< 5\%$, le risque étant donné par p -value.

Classe de pression	W	p.value
A	0.93	0.19
B	0.97	0.81
C	0.98	0.94
D	0.92	0.08
E	0.93	0.17

Tableau 2: Résultats du test de Shapiro de normalité du log(perces) pour les 5 classes de pression

Le test de Bartlett s'intéresse à l'homogénéité des variances des classes d'un facteur.

Test d'homogénéité des variances de Bartlett :

Le test de Bartlett teste l'hypothèse H_0 « les variances des échantillons sont égales ». Il a pour statistique de test Bartlett's K -squared.

La p -value du test étant supérieure à 5%, Table 5, nous ne pouvons rejeter H_0 à faible risque : aucune

classe n'a sa variance significativement différente de celle des autres.

Bartlett test of homogeneity of variances

data: log(VP) by classe de Pression

Bartlett's K-squared = 4.1768, df = 4, p-value = 0.3826

Table 5: Sortie R du test de Bartlett d'homogénéité des variances de log(VP) des classes de pression

c) Résultat de l'ANOVA

Nous pouvons finalement réaliser l'ANOVA. La p-value du test est de 0.59, Tableau 3. Cela nous conduit à accepter H_0 « les distributions suivent une même loi normale ». On ne peut donc pas conclure à l'influence des classes de pression sur $\log(VP)$.

Variable réponse : log(pertes)	Degrés de liberté	Sum of Squares	Mean Square	F value	Pr(>F)
Classes de pression	4	4.145	1.0362	0.7029	0.5919
Résidus	95	140.037	1.4741		

Tableau 3: Résultat de la méthode ANOVA - Comparaison du $\log(VP)$ entre les classes de pression

II.2.3.3 Influence de la pression dans ILI

ILI (Infrastructure Leakage Index) est l'indicateur internationalement utilisé pour caractériser la part des pertes due à l'infrastructure théoriquement incompressible (fuites non détectables avec les moyens actuels et délais de réparation des fuites).

$$ILI = \frac{CARL}{UARL}$$

$$CARL = VP - V_{cv} - V_{sc}$$

$$UARL = P_{moy} \times (18 \times L + 0.8 \times brt + 0.025 \times L_{brt} \times brt)$$

Avec :

- CARL Current Annual Real Losses (L/j)
- UARL Unavoidable Annual Real Losses (L/j)
- VP le volume de pertes (L/j)
- V_{cv} le volume volé (L/j)
- V_{sc} le volume sous compté (L/j)
- P_{moy} la pression moyenne du service (m)
- L le linéaire de réseau (km)
- brt le nombre de branchements du service
- L_{brt} le linéaire moyen d'un branchement, pris égal à 6 m

La formule de l'UARL tirée de (Lambert *et al*, 1999) considère la pression comme un facteur de pertes. Cela signifie que l'indicateur ILI représente la part de pertes que l'on ne peut pas attribuer à l'effet de la pression, du linéaire et des branchements.

Or d'après notre étude, les pertes sont indépendantes de la pression de service. Nous vérifions que par conséquent l'ILI est lié à la pression moyenne sur notre échantillon de services. La Figure 11 montre que nous avons des ILI très faibles : les pertes de 63 des 100 services de cet échantillon sont inférieures aux pertes incompressibles UARL, soit $ILI < 1$. Une tendance des ILI plus forts pour des faibles pressions moyennes de service se dessine. Elle se confirme sur les diagrammes en boîte des ILI par classes de pression, Figure 12.

Enfin, l'existence du lien entre la pression et *ILI* est confirmée par un test de corrélation de Spearman, Table 6. Le coefficient est significativement négatif.

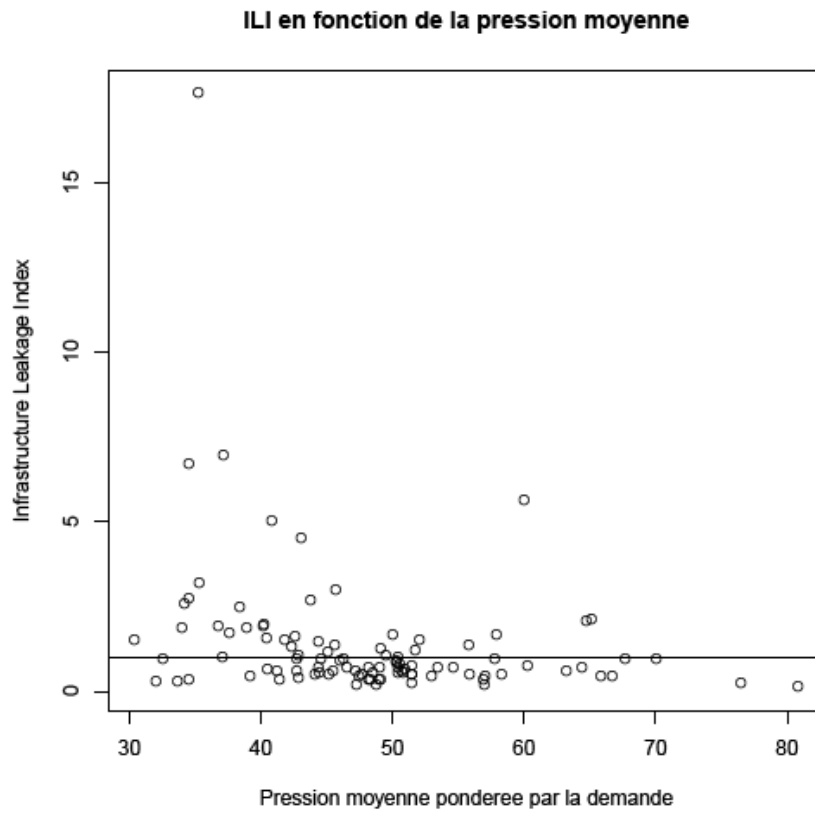


Figure 11: ILI en fonction de la pression moyenne de service

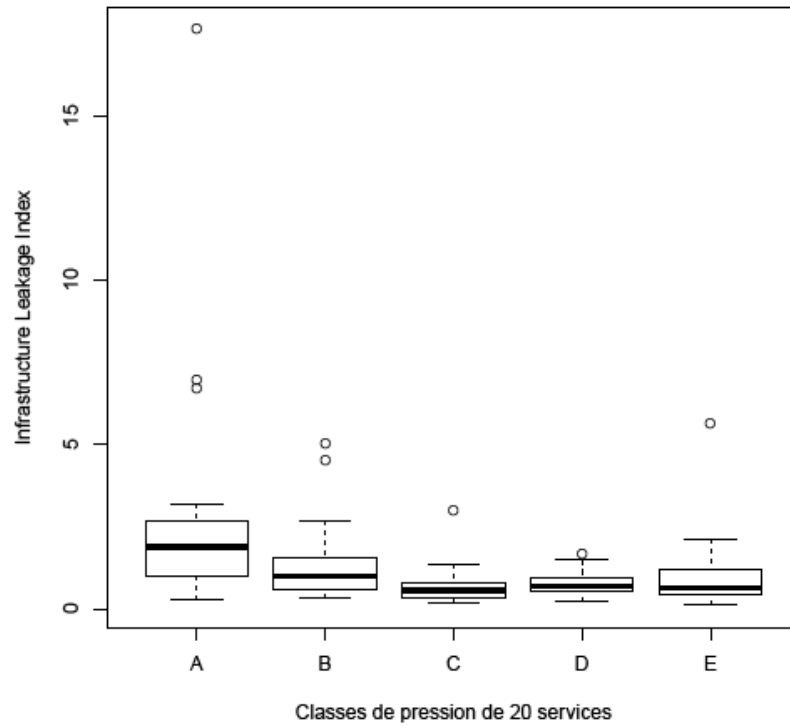


Figure 12: Diagrammes en boîte de ILI pour les 5 classes de pression

<p>Spearman's rank correlation rho</p> <p>data: pertes\$ILI and pertes\$pmoy</p> <p>S = 229572.1, p-value = 0.0001077</p> <p>alternative hypothesis: true rho is not equal to 0</p> <p>sample estimates:</p> <p>rho</p> <p>-0.3775706</p>

Table 6: Résultat du test de corrélation de Spearman de ILI et de la pression moyenne

III Recherche d'un indice de pertes

Sur la base d'une extraction des données et indicateurs servant à l'élaboration des RAD des 1849 services AEP français exploités par Veolia Eau, nous menons une étude statistique (à l'aide du logiciel R) dont l'objectif est l'identification des facteurs de pertes en eau et la construction d'un modèle du volume perdu par ces services.

III.1 Constitution de l'échantillon des Services AEP

Les Tableau 4 et Tableau 5 récapitulent respectivement les données des services et les indicateurs calculés pour alimenter les RAD (Veolia Eau) ou pertinents pour l'étude (Irstea).

Donnée N°	Donnée	Notation
D1	Nombre d'habitants	<i>hab</i>
D2	Nombre d'abonnés	<i>ab</i>
D3	Nombre de branchements	<i>brt</i>
D4	Linéaire de canalisation (m)	<i>L</i>
D5	Volume produit total (m3/an)	<i>V_p</i>
D6	Volume acheté en gros (m3/an)	<i>V_{imp}</i>
D7	Volume vendu en gros (m3/an)	<i>V_{exp}</i>
D8	Volume consommé comptabilisé sur 52 semaines (m3/an)	<i>V_{cc}</i>
D9	Volume consommé non compté sur 52 semaines (m3/an)	<i>V_{nc}</i>
D10	Volume de service (m3/an)	<i>V_s</i>
D11	Nombre de fuites sur branchement	<i>F_b</i>
D12	Nombre de fuites sur canalisations	<i>F_c</i>
D13	Linéaire dont le matériau et le diamètre sont renseignés (m)	<i>L_r</i>
D14	Linéaire dont le matériau est déterminé (m)	<i>L_m</i>
D15	Linéaire dont le diamètre est déterminé (m)	<i>L_d</i>
D16	Linéaire en fonte (m)	<i>L_{fonte}</i>
D17	Linéaire en pvc (m)	<i>L_{pvc}</i>

Tableau 4: Données issues des RAD utilisées

N°	Variable calculée	Notation	Calcul
I1	Volume mis en distribution (m3/an)	V_{md}	$V_p + V_{imp} - V_{exp}$
I2	Volume consommé sur 52 semaines (m3/an)	V_{ca}	$V_{cc} + V_{cnc} + V_s$
I3	Volumes non comptés	V_{NC}	$V_{md} - V_{cc}$
I4	Volume de perte (m3/an)	V_P	$V_{md} - V_{ca}$
I5	Volume sortant (m3/an)	$V_{sortant}$	$V_{exp} + V_{ca}$
I6	Rendement RPQS (%)	R	$(V_{ca} + V_{exp}) / (V_p + V_{imp})$
I7	Rendement primaire (%)	R_{dt}	V_{ca} / V_{md}
I8	Indice Linéaire des Volumes Non Comptés (m3/j.km)	$ILVNC$	$(V_{md} - V_{cc}) / (365 \times L)$
I9	Indice Linéaire de Pertes (m3/j.km)	ILP	$(V_{md} - V_{ca}) / (365 \times L)$
I10	Indice Linéaire de Consommation (m3/km.j)	ILC	$V_{ca} / (365 \times L)$
I11	Indice Linéaire de Consommation au sens du grenelle (m3/km.j)	ILC_g	$V_{sortant} / (365 \times L)$
I12	Taux de fuites sur branchement (Nb/1000 brcht/an)	txF_b	$F_b / (1000 \times brt)$
I13	Taux de fuites sur canalisation (Nb/km canalisations/an)	txF_c	$F_c / (L / 1000)$
I14	Taux de branchements par abonné	$Txbrt$	brt / ab
I15	Part du linéaire dont le matériau et/ou le diamètre sont renseignés	$\%L_r$	L_r / L
I16	Part du linéaire dont le matériau déterminé (%)	$\%L_m$	L_m / L
I17	Part du linéaire dont le diamètre est déterminé (%)	$\%L_d$	L_d / L
I18	Part du linéaire en fonte (%)	Tx_{fonte}	L_{fonte} / L
I19	Part du linéaire en pvc (%)	Tx_{pvc}	L_{pvc} / L
I20	Diamètre moyen pondéré par le linéaire (mm)	DN	$\Sigma(di \times Li) / L$
I21	Densité d'abonnés	Ab_{km}	ab / L
I22	Densité de branchements	Brt_{km}	brt / L

Tableau 5: Indicateurs calculés avec les données des RAD

Nous avons effectué 6 étapes de « tri » afin de travailler sur un échantillon de services pour lesquels nous disposons de l'ensemble des données, jugées fiables.

Tout d'abord, les quelques contrats avec un rendement de RPQS $\geq 100\%$ présentent manifestement une erreur.

Ceux dont le volume consommé autorisé n'est pas donné sont écartés, ce qui revient à éliminer les services dont le rendement primaire est nul (le rendement RPQS ne l'est pas du fait du volume exporté positif).

Nous avons enlevé de l'étude les services dont le nombre de branchements n'est pas renseigné ou est très éloigné du nombre d'abonnés.

Comme nous souhaitons étudier l'influence du nombre de fuites, des matériaux et des diamètres, nous ne conservons que les services pour lesquels des fuites ont été renseignées et un pourcentage élevé du linéaire détaillé en termes de matériaux et de diamètres.

L'ensemble de ces étapes de tri ainsi que le nombre de services écartés par chacune est récapitulé dans le Tableau 6.

Etape	Grandeur de tri	Ecarté si	Seuil	Nombre de services
1	R	\geq	100%	12
2	rdt	=	0%	11
3	brt	=	vide ou 0	13
4	Tx_{brt}	< >	0.06 et 1.2	14 96
5	$Fc+Fb$	=	vide	63
6	$\%Lr$	< >	90% et 110%	554 58
7	$\%Ld$	< >	90% et 110%	101 1

Tableau 6: Etapes de nettoyage des données des RAD

Au final, nous disposons d'un échantillon constitué de **926 services AEP Veolia Eau**.

III.2 Etude de la distribution des pertes

Afin que l'étude statistique soit pertinente, il faut s'assurer que notre échantillon de volumes perdus est correctement distribué pour représenter au mieux l'ensemble des services AEP français.

Les diagrammes en boîte, sont les plus appropriés pour rendre compte de la répartition d'une variable continue (Lire le paragraphe Lecture d'un diagramme en boîte).

Les graphiques « QQplot » représentent la normalité d'une distribution (Lire le paragraphe Lecture du graphique Quantile-Quantile).

La Figure 13 représente la distribution des pertes de notre échantillon. On constate qu'elle est fortement dissymétrique avec les pertes les plus importantes considérées comme atypiques. La distribution n'est manifestement pas normale. L'allure des quantiles de VP est exponentielle.

Cela nous donne l'intuition qu'une transformation logarithmique de VP peut permettre d'obtenir une variable avec une distribution symétrique, voire normale. De fait, la Figure 14 montre que $\log(VP)$ a une distribution symétrique et régulière (médiane au centre de la boîte) mais qu'un certain nombre de services a des pertes atypiques, éloignées de la distribution théorique.

Afin d'effectuer certains tests statistiques qui requièrent un échantillon normalement distribué, nous choisissons de retirer ces services de notre échantillon. Nous constatons sur la Figure 15 que la condition de normalité est alors satisfaite, graphiquement puis avec le test de Shapiro-Wilk dont la p-value de 0.151 indique que l'on ne peut pas rejeter l'hypothèse de normalité.

Notre échantillon est maintenant de **908 services AEP**.

926 services AEP

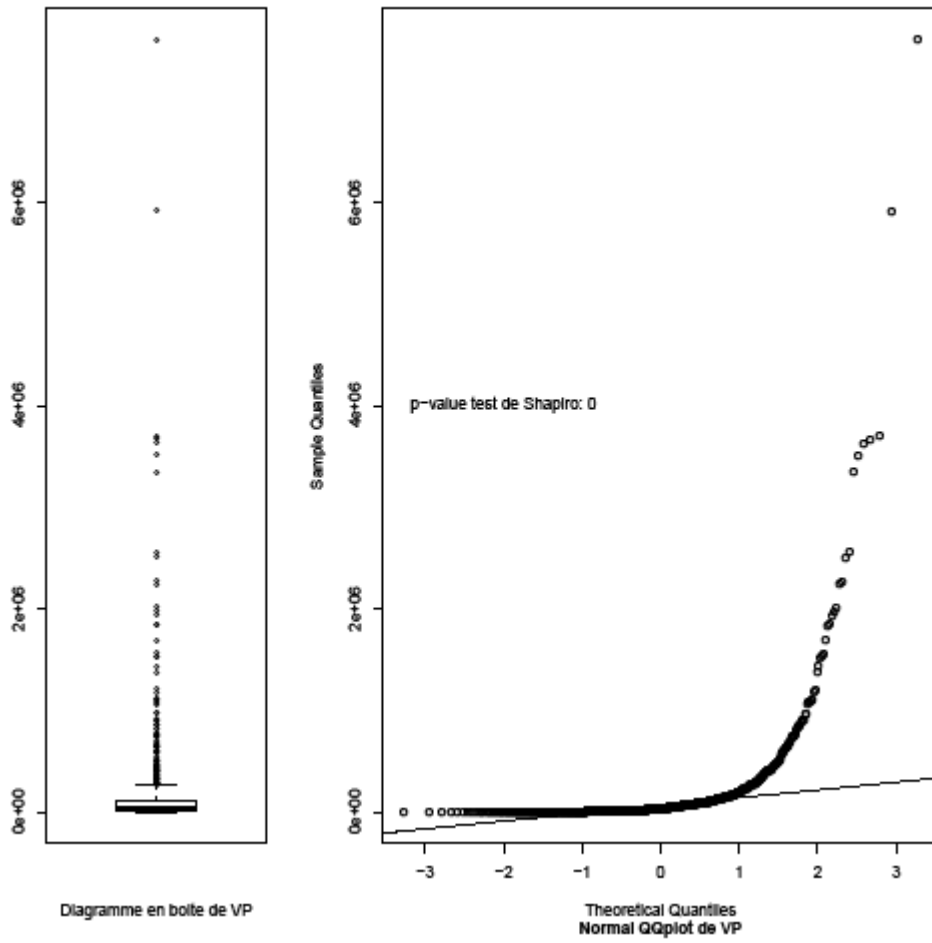


Figure 13: Distribution de VP sur l'échantillon de 926 services AEP

926 services AEP

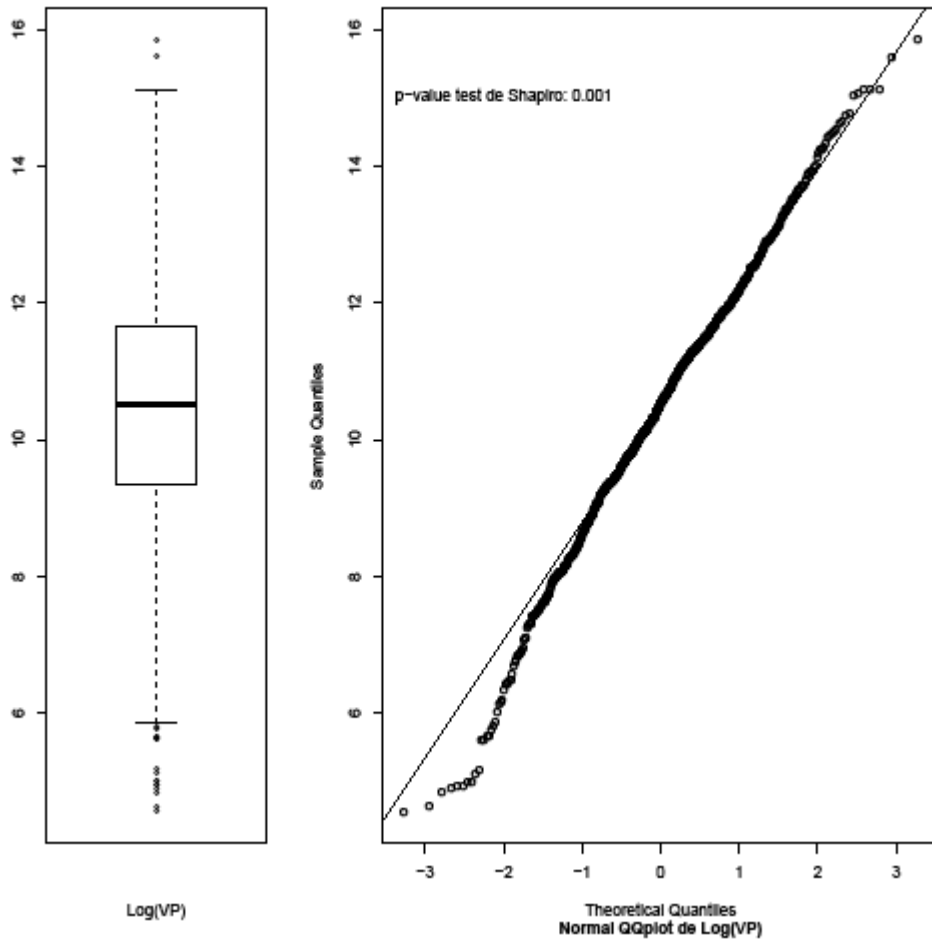


Figure 14: Distribution de log(VP) sur l'échantillon de 926 services AEP

908 services AEP

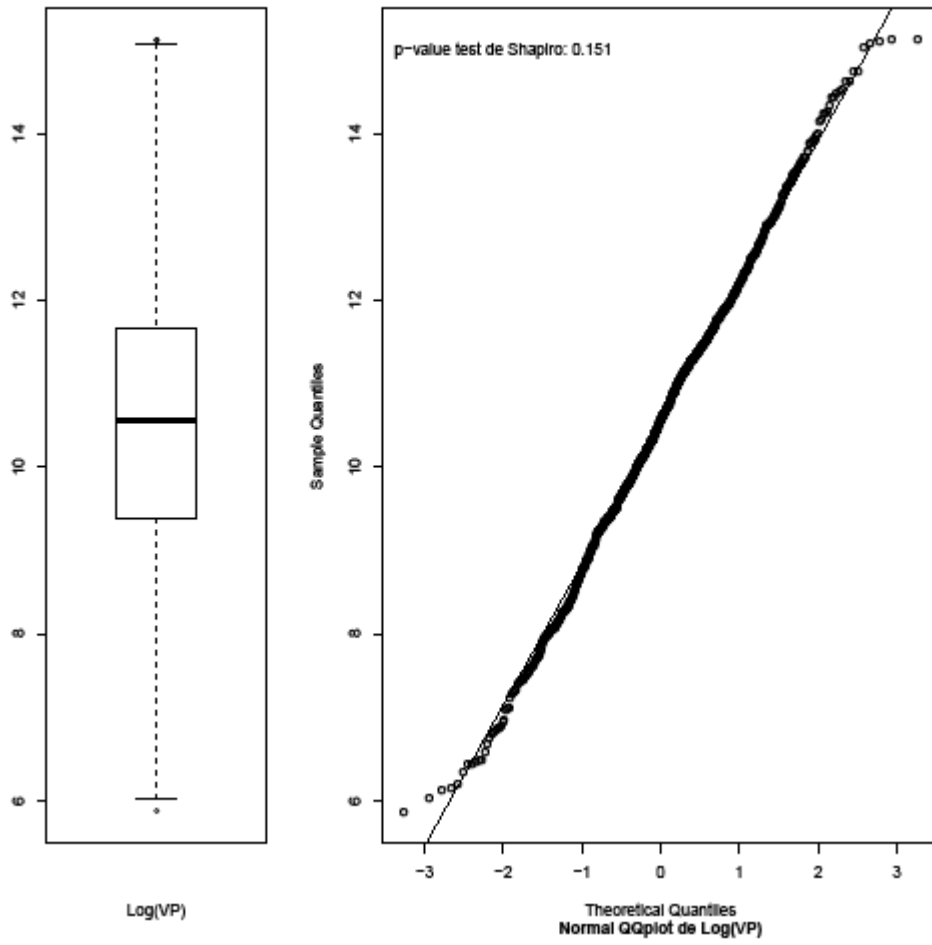


Figure 15: Répartition de Log(VP) sur l'échantillon de 908 services AEP

III.3 Construction d'un modèle des pertes

Parmi les données des services mises à notre disposition, Tableau 4 et Tableau 5, nous cherchons à identifier et quantifier les liens avec le $\log(VP)$.

III.3.1 Identification des variables significatives

Pour commencer nous éliminons les grandeurs qui nous ont servi à contrôler la complétude des données : $Lr, Lm, Ld, \%Lr, \%Lm, \%Ld$.

Ensuite nous écartons les grandeurs obtenues avec, ou incluses dans VP , c'est à dire : $Vp, Vimp, Vmd, Ventrant, VNC, ILVNC, ILP, R, rdt$.

Pour les variables restantes, nous testons successivement leurs corrélations avec VP et $\log(VP)$ par les méthodes de Pearson et de Spearman, dont les résultats sont présentés dans le Tableau 7.

Méthode	Pearson		Pearson		Spearman	
	Log(VP)~Variable		log(VP)~log(Variable)		Log(VP)~rang (Variable)	
Variables	Coef.de Corrélation	p-value	Coef.de Corrélation	p-value	Coef.de Corrélation	p-value
Vsortant	0.563	0	0.91	0.00	0.911	0
Vca	0.583	0	0.89	0.00	0.901	0
Vcc	0.584	0	0.91	0.00	0.899	0
ab	0.623	0	0.88	0.00	0.889	2.79e-310
hab	0.596	0	0.88	0.00	0.881	1.06E-294
brt	0.657	0	0.87	0.00	0.879	3.00E-294
L	0.577	0	0.82	0.00	0.827	1.28E-228
Fc	0.546	0	0.77	0.00	0.767	4.27E-172
Fb	0.362	0	0.74	0.00	0.745	2.03E-159
Lfonte	0.584	0	0.72	0.00	0.740	3.93E-158
Vs	0.295	0	0.75	0.00	0.706	5.35E-138
DN	0.526	0	0.54	0.00	0.533	1.04E-67
ILCg	0.410	0	0.47	0.00	0.469	7.61E-51
ILC	0.400	0	0.39	0.00	0.440	2.37E-44
Lpvc	0.357	0	0.42	0.00	0.406	2.06E-37
Vexp	0.318	0	0.48	0.00	0.392	8.82E-35
Ab_km	0.374	0	0.31	0.00	0.340	5.29E-26
Venc	0.274	0	0.69	0.00	0.305	4.58E-21
Brt_km	0.279	0	0.24	0.00	0.265	4.27E-16
Txfonte	0.161	1.14E-06	0.12	0.00	0.183	2.67E-08
txFc	0.064	0.056	-0.05	0.12	0.122	0.00024
txFb	-0.072	0.031	-0.14	0.00	0.021	0.53
Txpv	-0.174	1.22E-07	-0.17	0.00	-0.153	3.72E-06
Txbrt	-0.316	1.86E-22	-0.25	0.00	-0.358	7.23E-29

Tableau 7: Résultats des tests de corrélation des variables descriptives des services avec $\log(VP)$

On met ainsi en évidence trois groupes de variables :

- En rouge, les variables représentant les volumes consommés : elles sont les plus fortement corrélées avec $\log(VP)$.
- En vert, les variables représentant la taille du service.
- En bleu, les variables d'exploitation du service : nombre de fuites, linéaire en fonte, volume de service.

Le log des pertes a une relation monotone croissante avec toutes ces variables (coefficients de corrélation de Spearman > 0). Cette relation est fortement linéaire avec les transformations logarithmiques des variables (coefficients de corrélation de Pearson plus élevés après transformation logarithmique et voisins des coefficients de corrélation de Spearman).

D'après les tests sur nos données, on ne peut pas s'engager quant à l'existence d'un lien entre les variables $txFb$ et $txFc$ (en gris dans le Tableau 7) avec $\log(VP)$.

NB : Nous ne disposons pas des données sur les fuites pour 100% de l'échantillon étudié et l'on peut s'interroger sur la significativité de 0 fuites (vrai 0 ou non renseignées) de certains services. En conséquence, l'interprétation de ces variables Fb , Fc , $txFb$, $txFc$ est délicate.

Il apparaît déjà d'une manière générale que plus le réseau est important en termes de volume livré, de personnes desservies et de longueur de canalisations, plus les pertes sont importantes. Afin d'expliquer les pertes, il va falloir s'affranchir de cet effet de taille (sur lequel il n'est pas possible de jouer pour réduire les pertes).

III.3.2 Sélection du meilleur modèle

La démarche de construction d'un modèle statistique linéaire, les hypothèses et les tests permettant de les vérifier sont résumés dans (Besse et Laurent, 2013).

D'après les coefficients de corrélation du Tableau 7, nous choisissons de tester des modèles ayant pour variables :

- $\log(V_{sortant})$: le coefficient de $V_{sortant}$ est plus important que ceux de V_{ca} et de V_{cc} et les volumes exportés sont également compris dans l'expression du rendement (au sens RPQS) ;
- $\log(brt)$: le nombre de branchements est représentatif de la taille du service et de sa sollicitation ;
- $\log(L)$: le linéaire de canalisation est également représentatif de la taille du service et varie beaucoup selon le caractère urbain ou rural du service ;
- $\log(ab)$: le nombre d'abonnés représente la sollicitation du service.

La justification de la transformation logarithmique de chacune de ces variables est présentée sur les, Figure 16, Figure 17, Figure 18 et Figure 19. Elle permet d'obtenir des variables distribuées symétriquement, presque normalement.

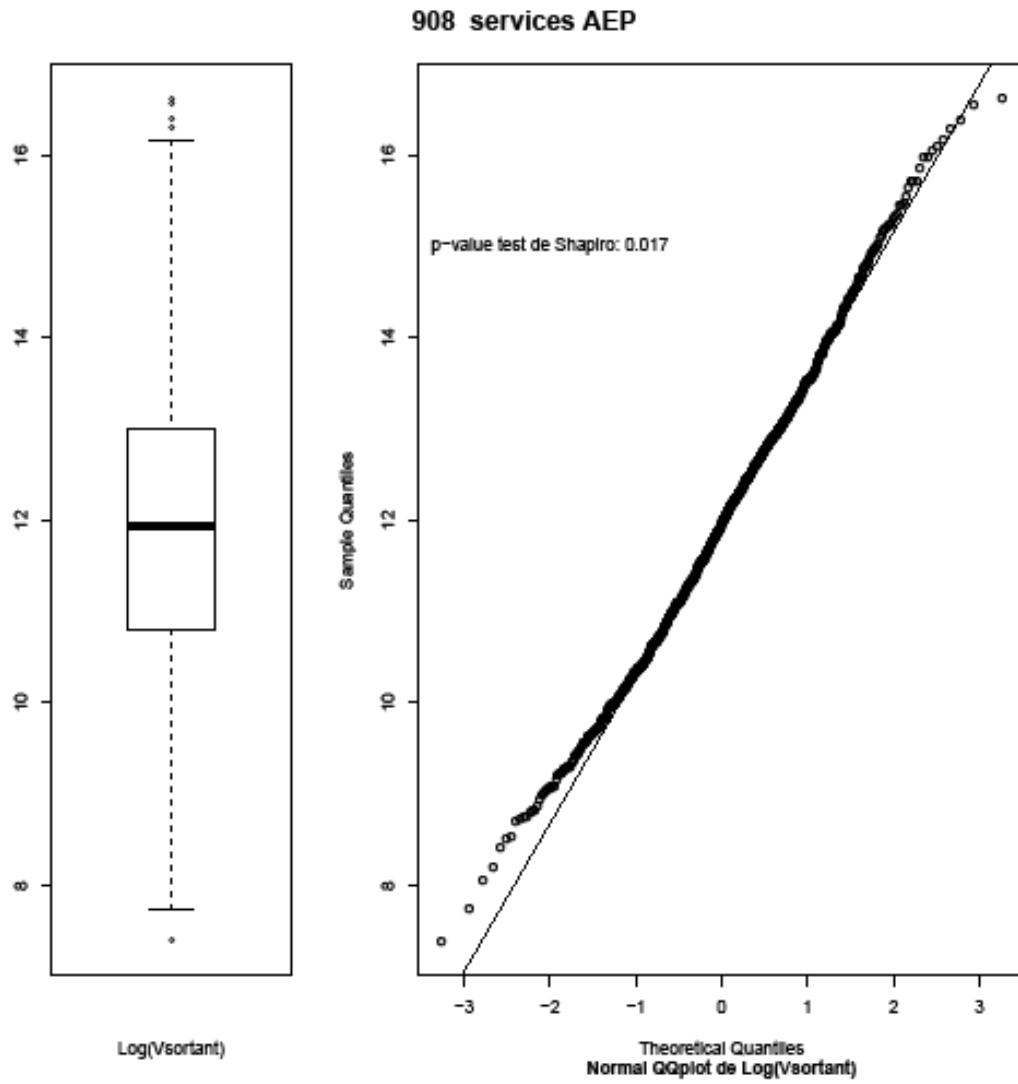


Figure 16: Distribution du log du volume sortant

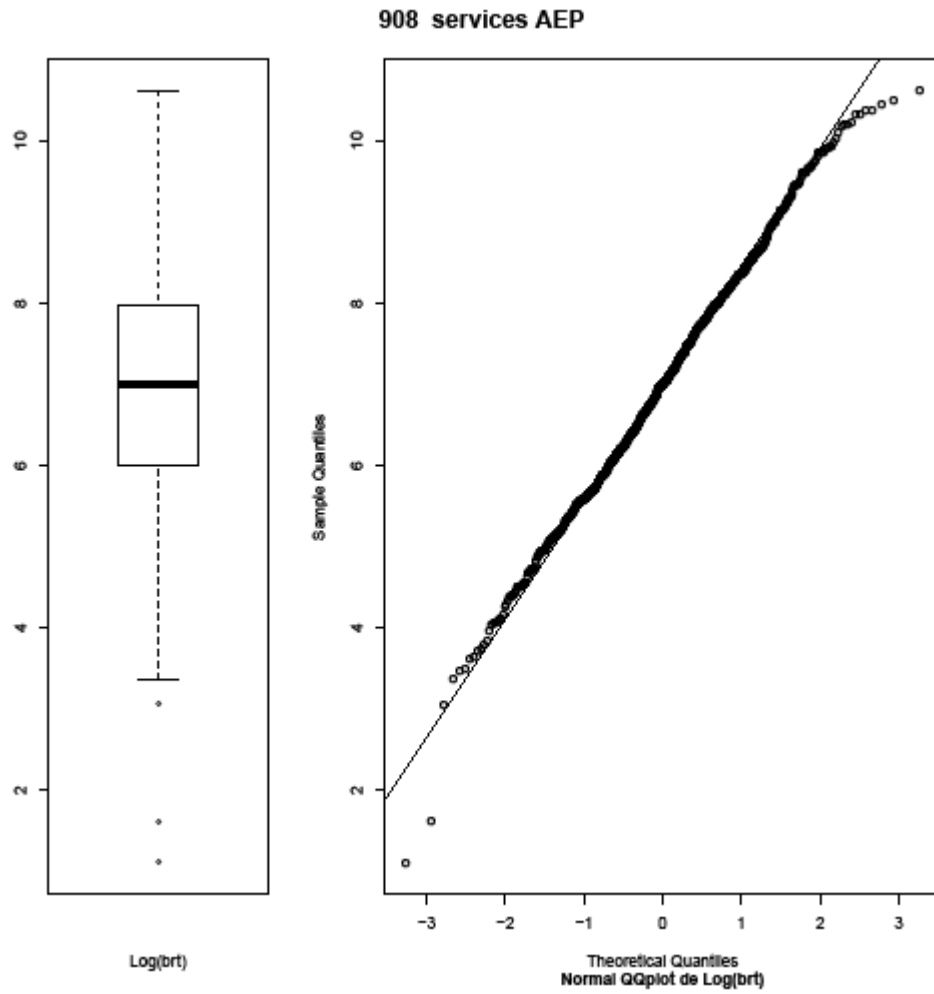


Figure 17: Distribution du log du nombre de branchements

908 services AEP

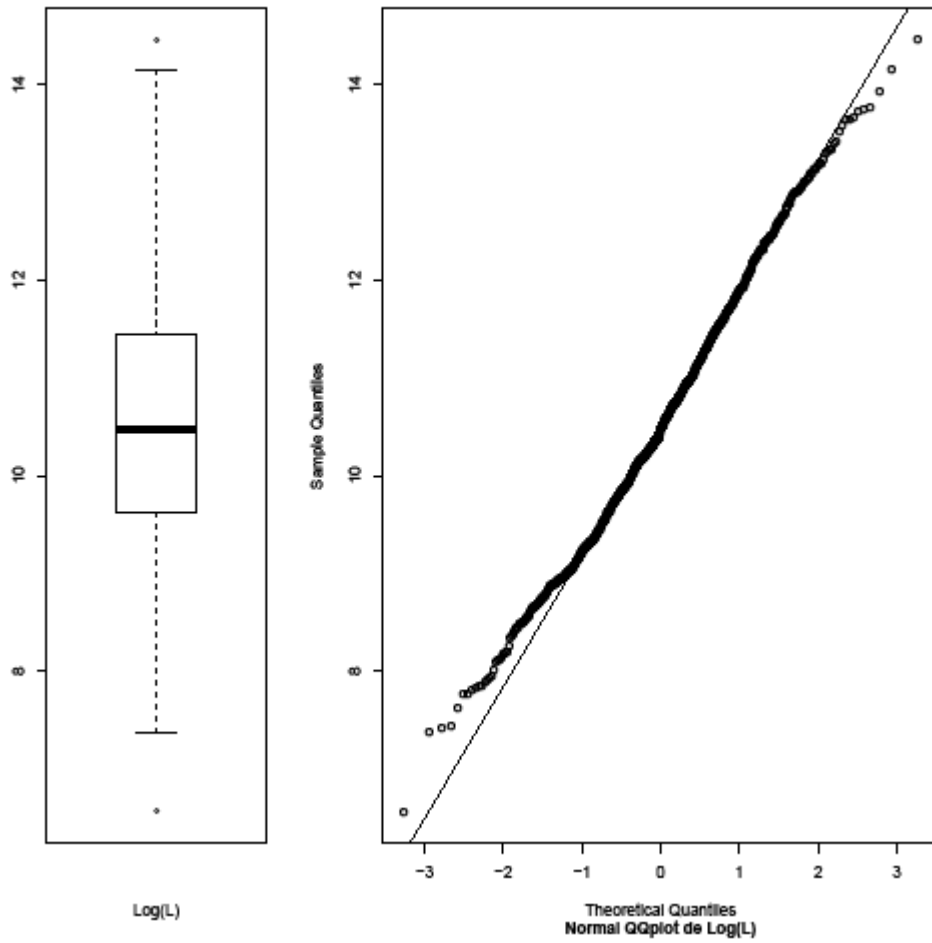


Figure 18: Distribution du log du linéaire

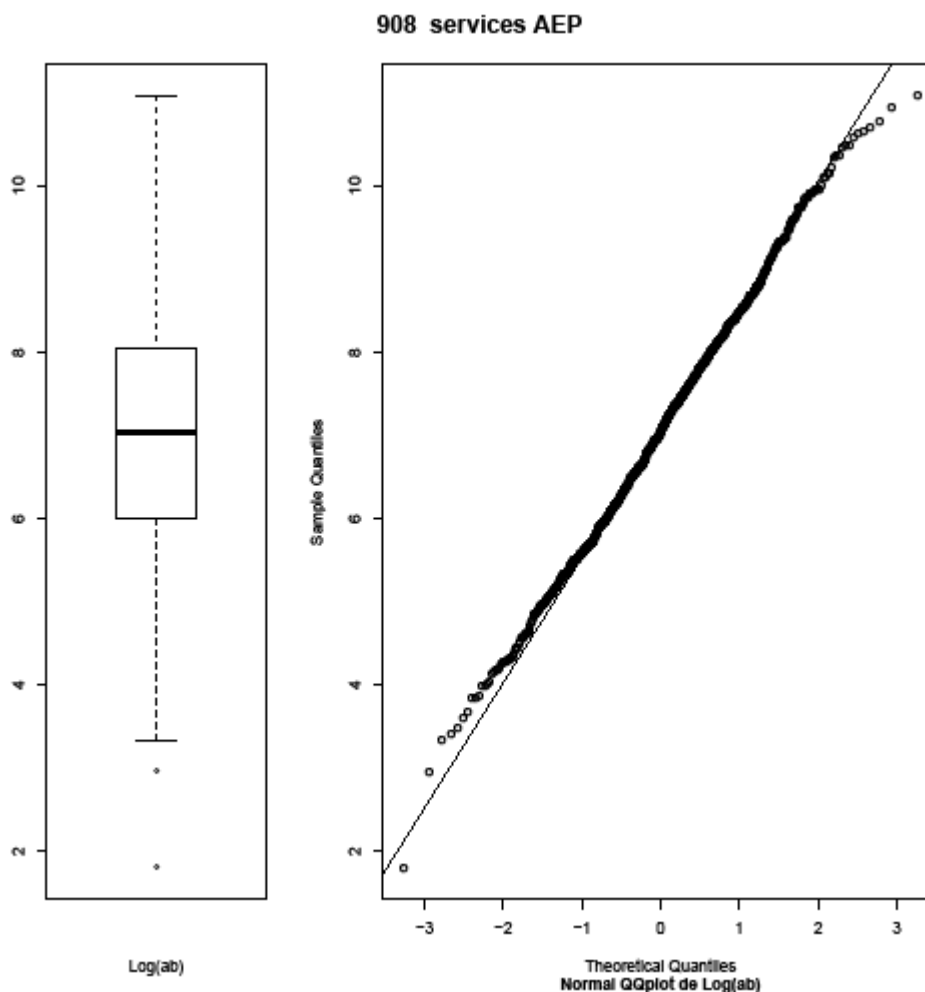


Figure 19 : Distribution du log du nombre d'abonnés

Le Tableau 8 récapitule les modèles linéaires testés à une, deux ou trois variables parmi les quatre ci-dessus.

Interprétation des résultats d'un modèle linéaire :

L'erreur standard résiduelle représente l'écart entre la variable à expliquer (ici $\log(VP)$) et la variable réponse, résultat du modèle.

Le R^2 du modèle représente la proportion de variance de $\log(VP)$ expliquée par le modèle.

Le R^2 ajusté est calculé de manière à supprimer l'influence des prévisions du modèle réalisées avec l'échantillon ayant servi au calage du modèle.

AIC (Akaike Information Criterion) est un critère destiné à choisir entre plusieurs modèles testés sur le même échantillon en faisant un « compromis entre ajustement des données et parcimonie du modèle ». On aura tendance à rechercher un modèle qui minimise AIC.

Modèle de log(VP)	Variable(s)	Residual Standard error	R ²	R ² ajusté	AIC
1	Vsortant	0.7215	0.8275	0.8273	1988
2	Nombre de branchements	0.8621	0.7538	0.7535	2311
3	Nombre d'abonnés	0.8185	0.778	0.7778	2217
4	Linéaire	0.9886	0.6762	0.6759	2559
5	- Volume sortant - Nombre de branchements	0.7091	0.8336	0.8332	1957
6	- Volume sortant - Nombre d'abonnés	0.7061	0.835	0.8346	1949
7	- Volume sortant - Linéaire	0.7097	0.8333	0.833	1958
8	- Volume sortant - Nombre de branchements - Linéaire	0.705	0.8357	0.8352	1947
9	- Volume sortant - Nombre d'abonnés - Linéaire	0.7021	0.8371	0.8365	1940
10	- Volume sortant - Nombre de branchements - Nombre d'abonnés	cf modele 6 (nombre de branchements non significatif)			

Tableau 8: Résultats de 10 essais de modèles linéaire de log(VP)

Modèle	Ordonnée à l'origine	Log de Volume sortant	Log de Nombre de branchements	Log de Nombre d'abonnés	Log de Linéaire
1	-1.43189	0.99843			
2	3.01571		1.07378		
3	3.03395			1.06126	
4	-1.19591				1.11109
5	-0.74065	0.79641	0.24714		
6	-0.49363	0.74568		0.29593	
7	-1.87123	0.85657			0.20272
8	-1.24481	0.76014	0.17361		0.13773
9	-0.99476	0.71160		0.22791	0.13173
10	<i>Nombre de branchements non significatif</i>				

Tableau 9: Coefficients des modèles de log(VP)

Le modèle 9 est celui qui présente l'erreur résiduelle standard la plus faible, ainsi que le meilleur R². En outre il a le plus faible AIC.

R donne le résumé suivant du modèle logarithmique n°9:

```

Call:
lm(formula = pertes$IVP ~ log(pertes$Vsortant) + log(pertes$ab) + log(pertes$L))

Residuals:
    Min       1Q   Median       3Q      Max
-2.97312 -0.36692  0.03423  0.44917  1.83946

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -0.99476   0.27367  -3.635  0.000294 ***
log(pertes$Vsortant) 0.71160   0.04314  16.493 < 2e-16 ***
log(pertes$ab)      0.22791   0.05020   4.540  6.38e-06 ***
log(pertes$L)       0.13173   0.03899   3.378  0.000760 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7021 on 904 degrees of freedom
Multiple R-squared:  0.8371,    Adjusted R-squared:  0.8365
F-statistic: 1548 on 3 and 904 DF,  p-value: < 2.2e-16

```

Table 7: Résultats du modèle de log(VP) n°9 donnés par R

L'analyse des résidus du modèle, Figure 20, montre qu'ils sont homoscédastiques (absence de variance des résidus studentisés selon les valeurs prédites) et linéaires (absence de forme particulière du nuage de points des résidus).

On observe par contre que les résidus ne suivent pas une loi normale à gauche. Ceci se voit également sur le graphique Log(VP) en fonction des valeurs prédites par le modèle. Les faibles valeurs de log(VP) ne sont pas alignées sur la droite de pente 1.

Aucun service n'influence sensiblement le modèle (Distance de Cook >1).

Lecture d'un graphique de Distance de Cook (StatSoft, 2013):

« Il s'agit d'une autre mesure de l'impact de l'observation respective sur l'équation de régression. Elle représente la différence entre les coefficients calculés et les valeurs qui auraient été obtenues si l'observation correspondante avait été exclue de l'analyse. Toutes les distances doivent être du même ordre de grandeur ; si ce n'est pas le cas, il y a de bonnes raisons de penser que la ou les observations respectives biaisent l'estimation des coefficients de régression. »

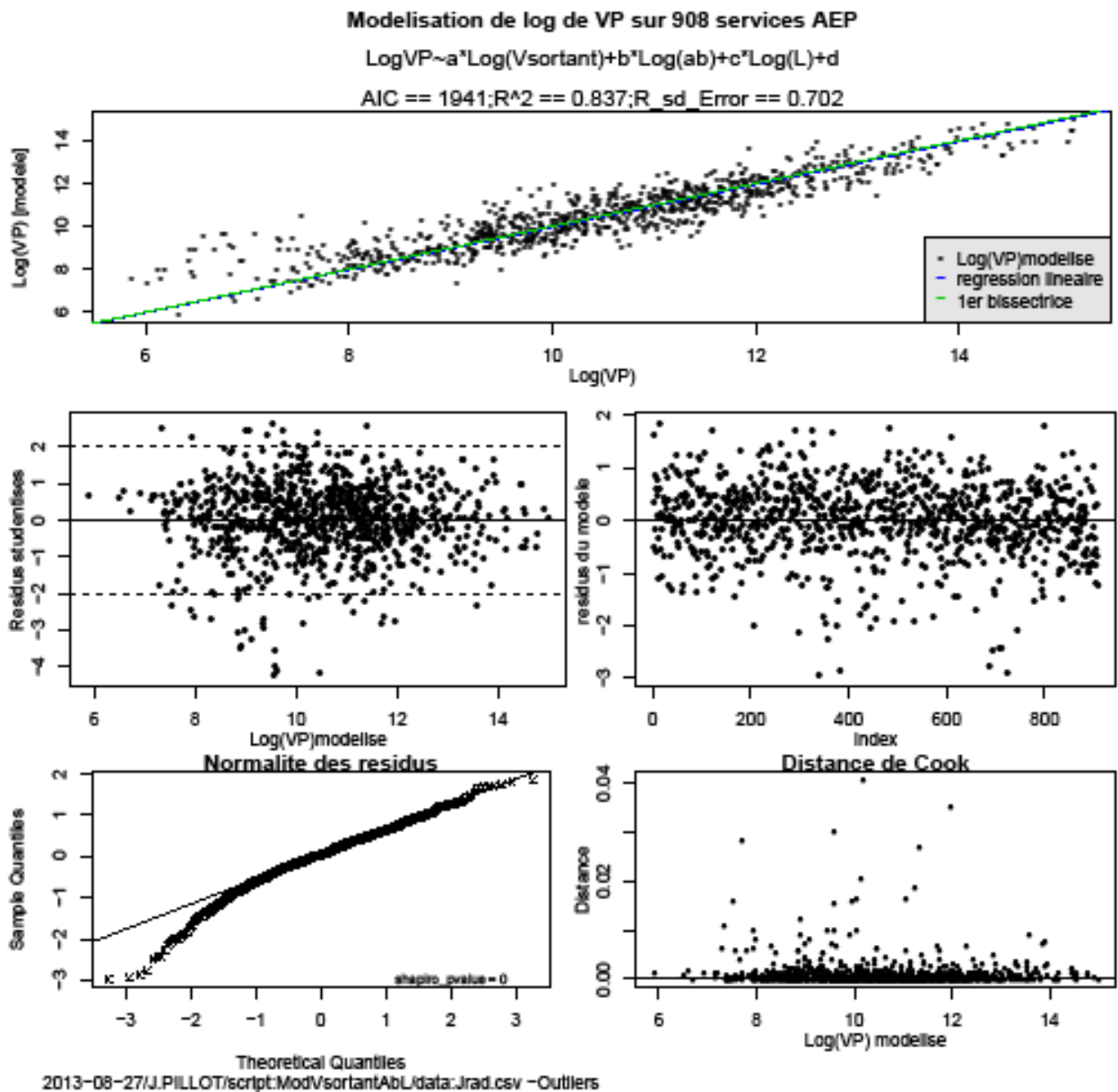


Figure 20: Prédications et résidus du modèle n°9 de log(VP)

Le modèle logarithmique peut être transformé en modèle puissance. On a alors :

$$VP = e^{-0.99476} \times V_{\text{sortant}}^{0.7116} \times ab^{0.22791} \times L^{0.13173},$$

qui peut s'arrondir à : $VP = 0.37 \times V_{\text{sortant}}^{0.71} \times ab^{0.23} \times L^{0.13}$

On constate sur la Figure 21, graphique de VP en fonction des valeurs prédites par le modèle, que le nuage de points est centré sur la droite de régression rouge (de pente 0.79) qui s'écarte de la 1^{er} bissectrice (droite verte). Plus les pertes sont importantes, plus l'erreur du modèle est importante.

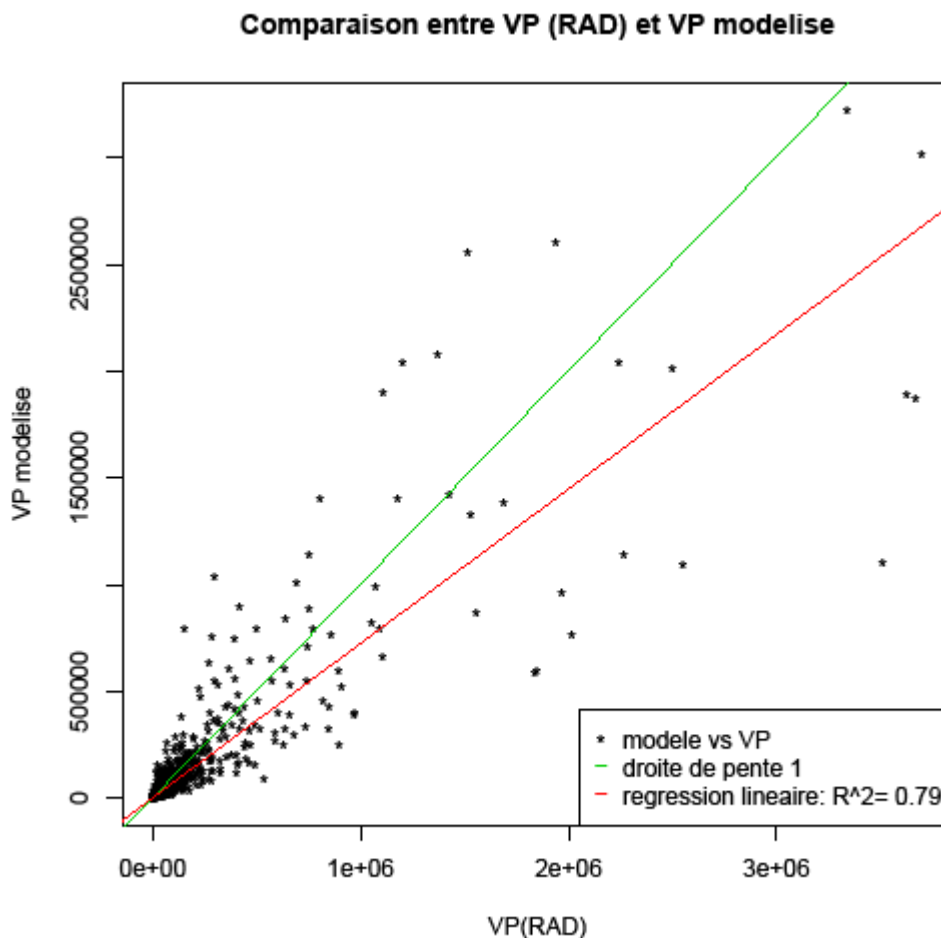


Figure 21: Comparaison des pertes prédites par le modèle sélectionné avec les données pertes

III.3.3 Validation sur d'autres données

Afin de vérifier si notre modèle, calé sur notre échantillon de services AEP exploités par Veolia Eau, donne de bons résultats sur l'ensemble des services AEP français, nous utilisons une extraction de la base de données nationale de l'ONSEA des années 2009, 2010 et 2011.

Tout d'abord, nous éliminons les services pour lesquels nous ne disposons pas de l'ensemble des données ab , L , $V_{entrant}$, $V_{sortant}$.

Nous calculons le rendement ($V_{sortant}/V_{entrant} * 100$) ainsi que VP ($V_{entrant} - V_{sortant}$) et $\text{Log}(VP)$.

Nous éliminons les services ayant des rendements $\geq 100\%$ ou nuls. Nous écartons également les services ayant un volume de pertes négatif.

Enfin nous appliquons notre modèle logarithmique (avec une précision de 10^{-5}) aux données « nombre d'abonnés », « linéaire de réseau » et « volume sortant ».

La Figure 22 illustre la validité de notre modèle logarithmique des pertes. Notre modèle sous-estime légèrement les pertes (pente de 0.93). L'ajustement aux données est très bon (0.99 sur les trois ans).

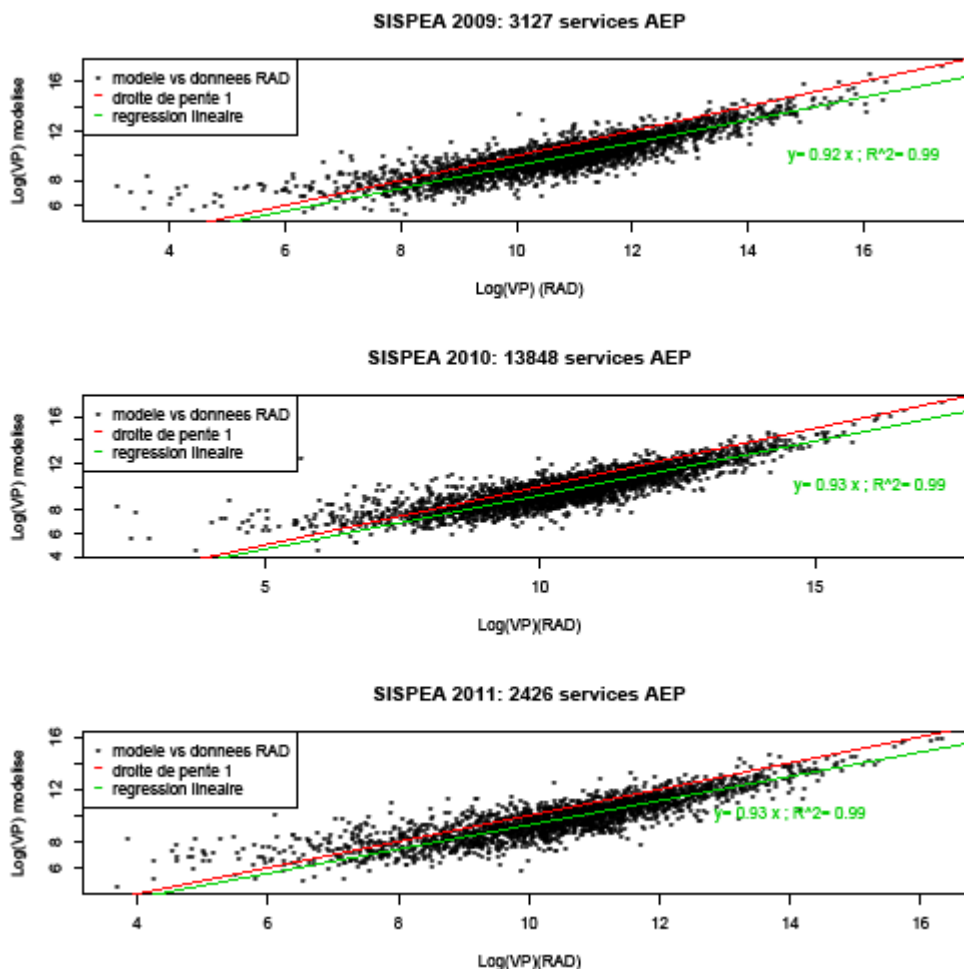


Figure 22: Application du modèle logarithmique des pertes aux données de l'ONSEA 2009, 2010, 2011

La Figure 23 présente le volume de pertes obtenu grâce à notre modèle (transformation exponentielle du log des pertes calculé précédemment) en fonction des pertes déclarées dans la base de l'ONSEA. On constate que l'écart est plus significatif avec une pente allant de 0.59 à 0.77 selon les années. Les ajustements des droites de régression aux nuages de points est bon en 2010 et 2011 ($R^2=0.9$) mais plus faible en 2009 ($R^2=0.77$).

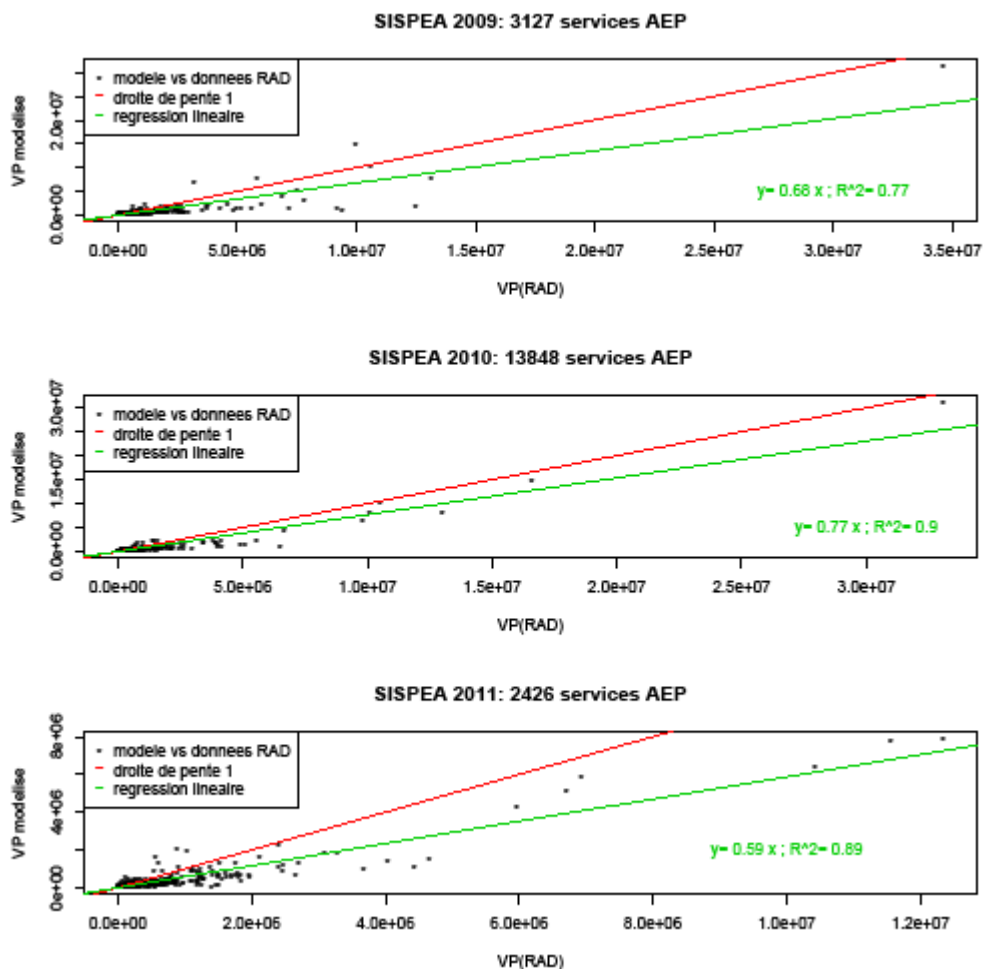


Figure 23: Application du modèle des pertes aux données de l'ONSEA 2009, 2010 et 2011

Le modèle est très sensible à la précision des coefficients :

- avec les coefficients du modèle log arrondis à 10^{-2} , on obtient une pente de 1.06 au lieu de 0.93 avec une précision de 10^{-5} ;
- Une pente d'environ 3 au lieu d'environ 0.7 avec une transformation exponentielle du modèle log arrondis à 10^{-2} .

III.3.4 Modèles linéaires simples

Comme le modèle n° 9 nous paraît difficile à utiliser en pratique, du fait de sa grande variabilité lorsqu'on effectue la transformation exponentielle pour exprimer VP , nous nous intéressons aux modèles n° 2 et n° 3 qui sont respectivement des fonctions affines du nombre de branchements et du nombre d'abonnés.

III.3.4.1 Nombre de branchements : modèle n° 2

Les pertes prédites par le modèle ainsi que les résidus du modèle sont présentés sur la Figure 24.

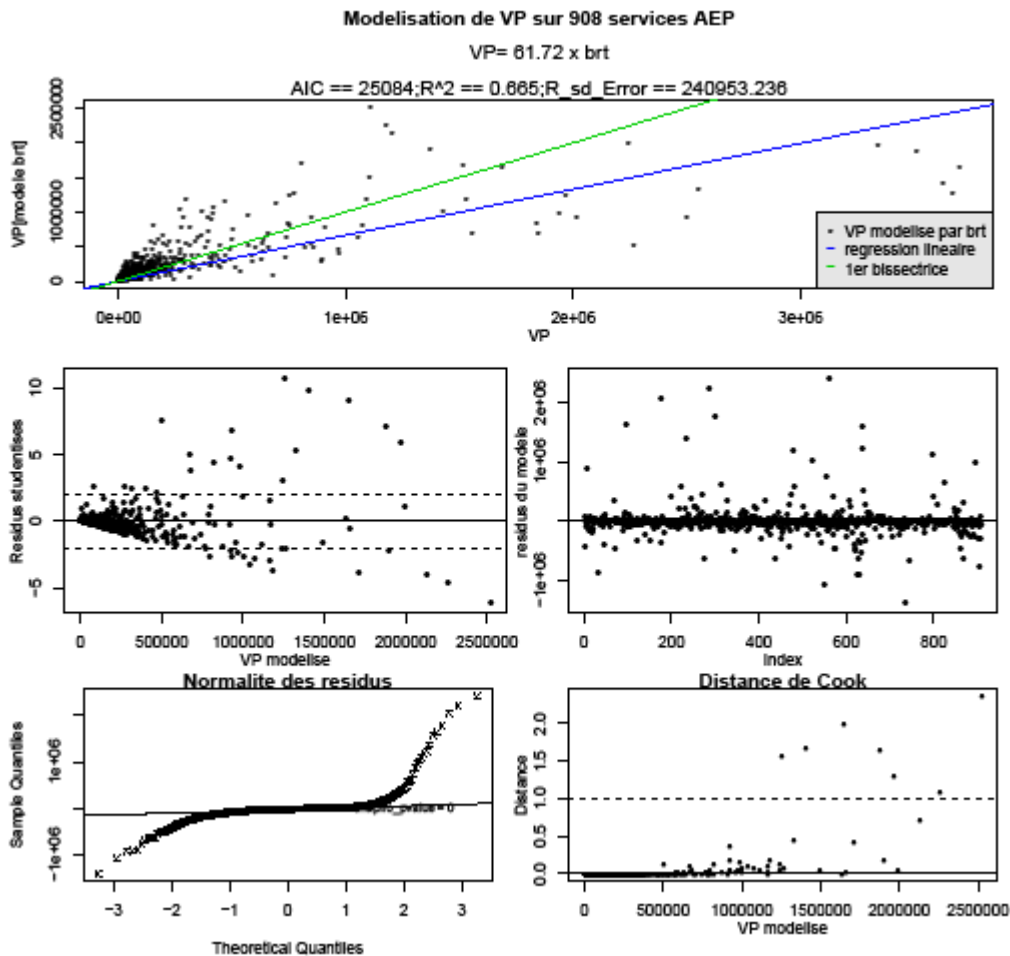


Figure 24: Modèle linéaire des pertes en fonction du nombre de branchements

On constate que les résidus de ce modèle présentent une hétéroscédasticité, linéaire et ne suivent pas une loi normale.

Ce modèle a des pouvoirs prédictif (Residual standard error élevée) et explicatif assez faibles ($R^2=0.665$).

L'analyse des résidus montrent que le modèle est sensible à notre échantillon (sept services ont une distance de Cook supérieure à 1 et influencent fortement le modèle) et qu'un paramètre n'est pas pris en compte (hétéroscédasticité : augmentation de la variance des résidus studentisés avec la valeur de VP modelisé).

III.3.4.2 Nombre d'abonnés : modèle n° 3

Les résultats du modèle n° 3 sont présentés en détail sur la Figure 25.

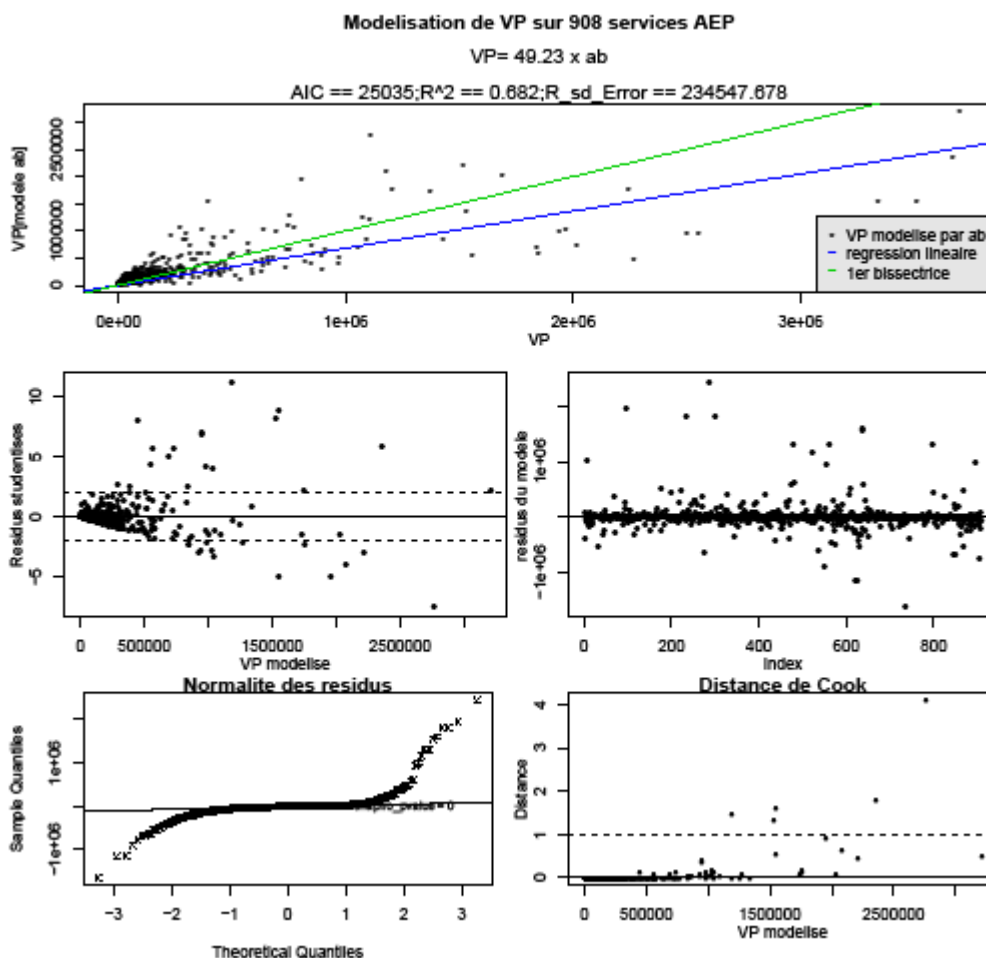


Figure 25: Modèle linéaire des pertes en fonction du nombre d'abonnés

Le modèle n° 3 est sensiblement équivalent au modèle n° 2 avec les mêmes problématiques : hétéroscédasticité, absence de normalité et points influents.

D'après les résultats du modèle faisant intervenir des puissances précédemment étudié, nous pouvons supposer qu'il est nécessaire de prendre en compte le volume sortant pour avoir un modèle homoscédastique.

III.4 L'indice de pertes

III.4.1 Définition

On définit l'indice de pertes comme le rapport du volume perdu et du volume issu de la prédiction du modèle n° 9, fonction du volume sortant, du linéaire et du nombre d'abonnés. Notre modèle définit un volume de pertes moyen par rapport aux pertes actuellement enregistrées par les services AEP. Le ratio entre le volume de pertes et la prédiction du modèle permet donc aux services de situer leur performance. Cet indicateur a le même sens que le rendement, c'est pourquoi nous étudions le lien entre ces deux indicateurs.

III.4.2 Intérêt vis-à-vis du rendement

La comparaison de cet indice de pertes au rendement (au sens RPQS), 1^{er} graphique des Figure 26, Figure 27, Figure 28 et Figure 29, montre un très fort lien entre les deux indicateurs.

Le rendement est un bon indicateur mais il n'est pas indépendant de la densité d'abonnés (Renaud, 2009). C'est ce que montrent les derniers graphiques des Figure 26, Figure 27, Figure 28 et Figure 29 : le rendement croît avec l'ILC. Ce constat visuel est confirmé par les tests de corrélation entre le rendement et l'ILC. Les méthodes de Pearson et de Spearman concluent à l'existence d'une corrélation positive (p-value <5%), bien que relativement faible (coefficients autour de 0.1), pour tous les échantillons.

Visuellement, il est difficile de déterminer s'il y a un lien entre l'ILC et notre indice de pertes du fait du petit nombre de services avec un grand ILC. Les tests de corrélations de Pearson sont non significatifs pour Veolia Eau 2011, l'ONSEA 2009 et l'ONSEA 2010 et concluent à une relation linéaire décroissante très faible (-0.024) pour 2011. Avec la méthode de Spearman, la p-value passe sous le seuil de 5% pour l'ONSEA 2009, 2010 et 2011 et les tests concluent à une très faible corrélation positive (coefficients de 0.063, 0.035 et 0.055). Le test de corrélation avec la méthode de Spearman ne met pas en évidence de corrélation sur les données Veolia Eau de 2011, Tableau 10.

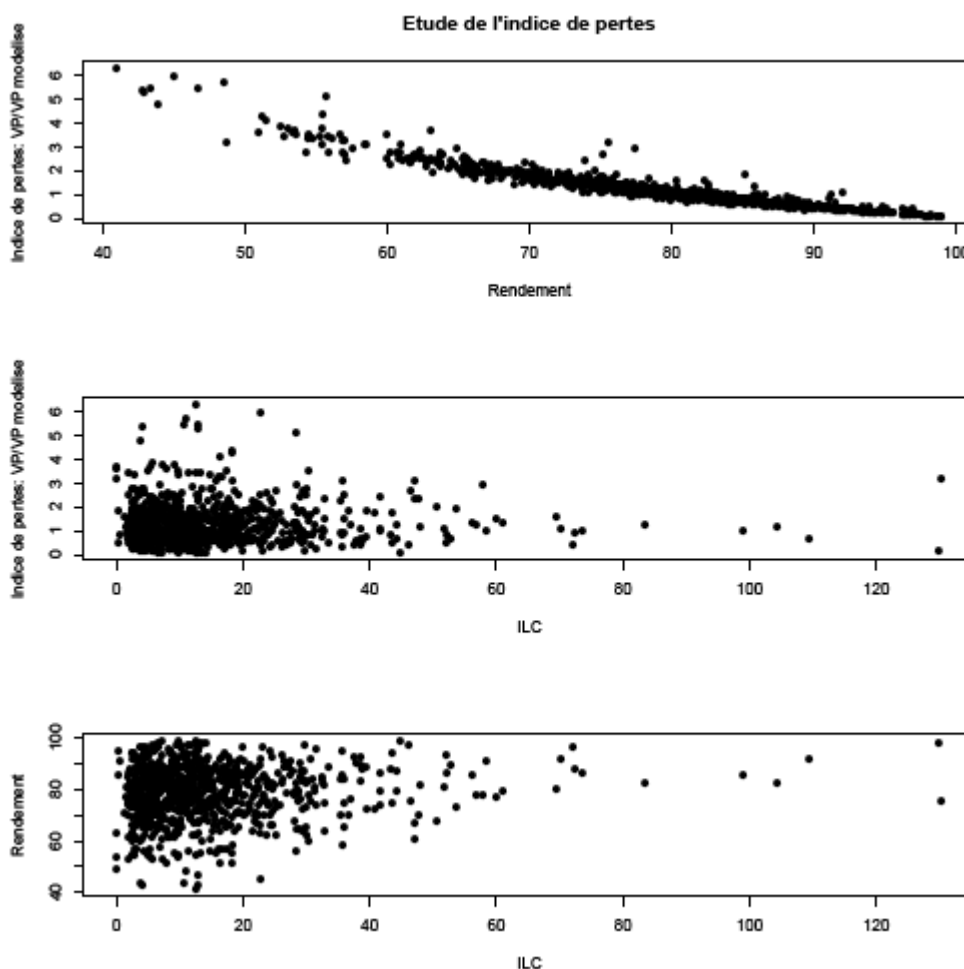


Figure 26: Indice de pertes, rendement et ILC, données de Veolia Eau 2011

Corrélation de <i>ILC</i> avec :	Méthode	P-value	Coefficient de corrélation
Indice de pertes	Pearson	0.3091	0.034
	Spearman	0.1599	0.047
Rendement	Pearson	0.001883	0.103
	Spearman	0.01571	0.080

Tableau 10: Résultats des tests de corrélation de *ILC* avec l'indice de pertes et le rendement, données Veolia Eau 2011

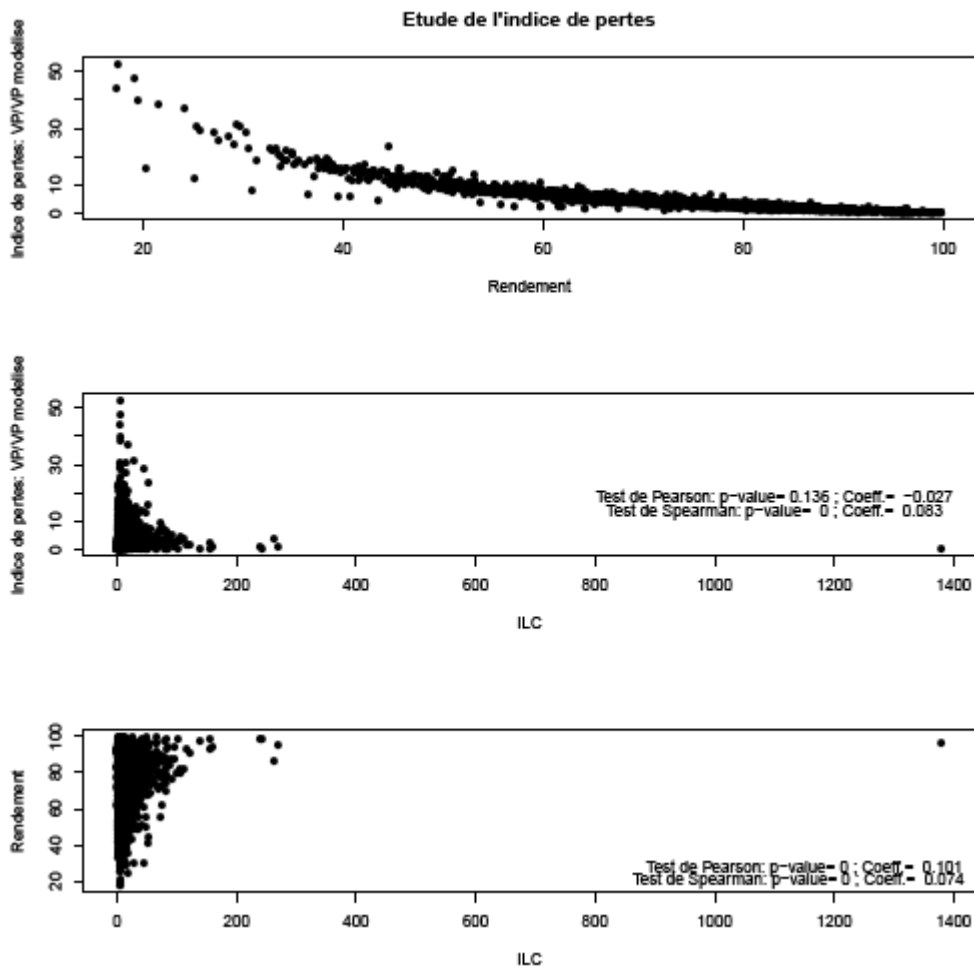


Figure 27: Indice de pertes, rendement et ILC, données de l'ONSEA 2009

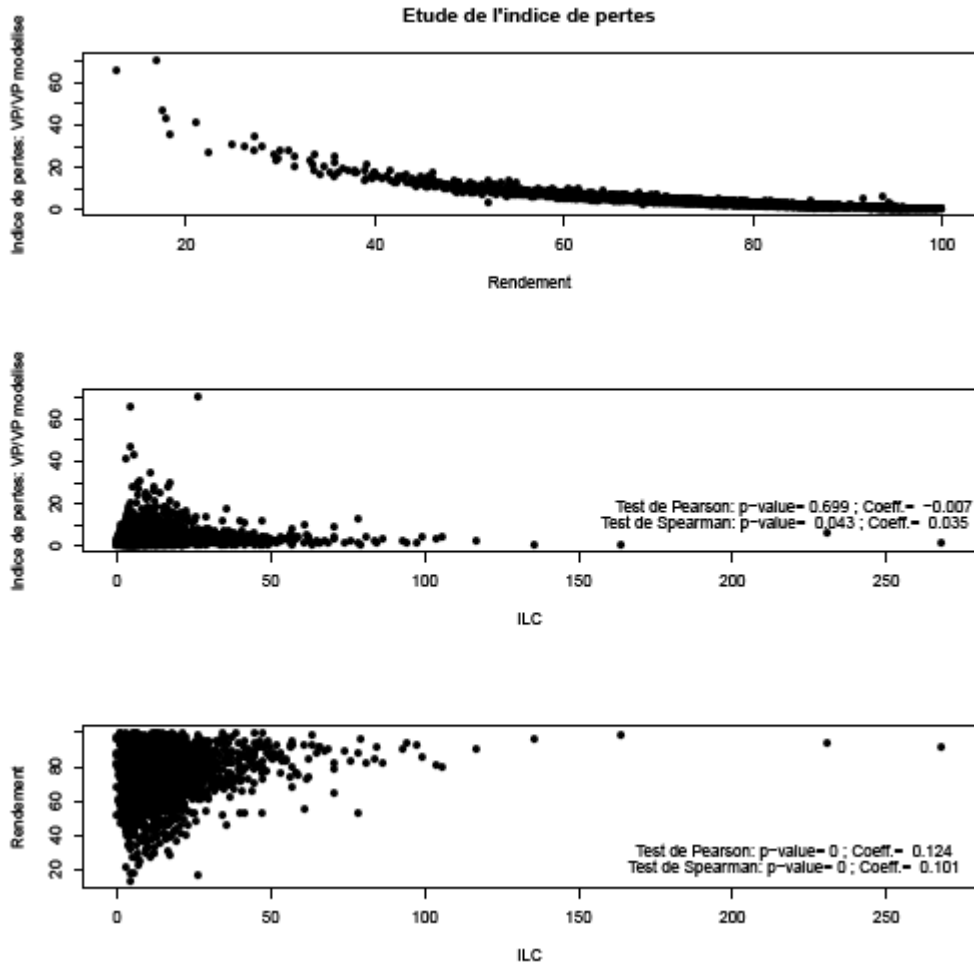


Figure 28: Indice de pertes, rendement et ILC, données de l'ONSEA 2010

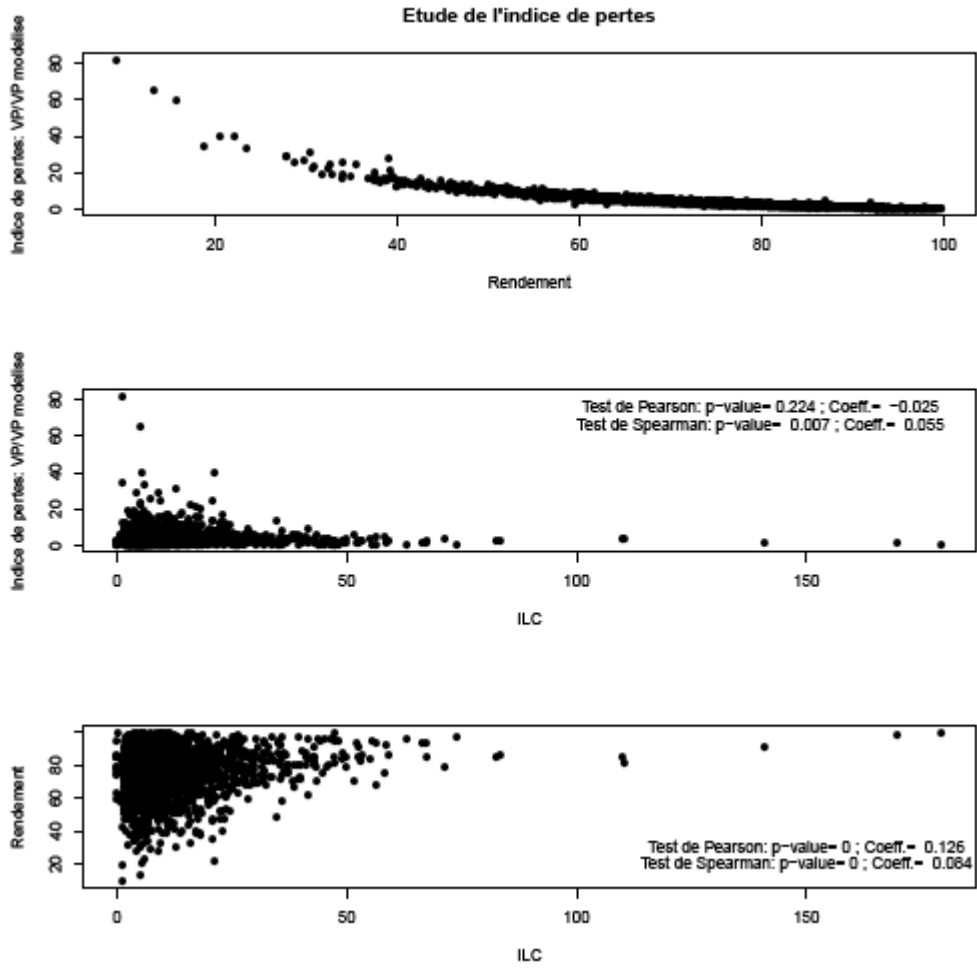


Figure 29: Indice de pertes, rendement et ILC, données de l'ONSEA 2011

IV Conclusion

Partant des expériences réussies de réduction des pertes obtenues grâce à la régulation de la pression, notre étude présupposait que la pression moyenne de service impacte significativement le niveau des pertes et qu'ainsi il était opportun de l'introduire dans les indicateurs de pertes français. Le système envisagé consistait à décomposer un indicateur global en deux composantes, l'une liée au potentiel de réduction des pertes par la régulation de pression et l'autre liée aux caractéristiques structurelles du réseau, dont le nombre de branchements.

Les investigations que nous avons menées, tant à l'échelle du secteur qu'à celle du service ont confirmé la pertinence de rapporter le volume annuel global de pertes au nombre de branchements, les deux grandeurs étant, en moyenne, liées par une relation de proportionnalité. En revanche, tous les résultats convergent pour conclure à une absence de relation significative entre le niveau des pertes et la pression moyenne de service.

Ce constat, bien qu'inattendu pour les raisons précédemment évoquées, nous semble robuste vu la variété et la profondeur de nos explorations (travail par secteur et par service, analyses statistiques individuelles et par classes).

En conséquence, nous pensons que l'idée de définir à l'échelle du service des indicateurs distincts dédiés aux aspects structurels et fonctionnels des pertes doit être abandonnée. Cependant, nous préconisons l'utilisation d'un indicateur de pertes par branchement pour évaluer de façon globale les pertes d'un service. Un tel indicateur existe dans le système proposé par l'IWA (Alegre et al, 2006).

Concernant le potentiel de réduction des pertes par la régulation de pression, nous proposons d'utiliser l' $IRP = \frac{P}{P_{\min}}$, défini dans (Ben Hassen, 2012), à l'échelle du secteur. Cet indicateur nous semble intéressant pour identifier les zones présentant le plus fort potentiel.

Enfin, nos résultats concernant l'absence d'impact significatif de la pression moyenne sur le niveau des pertes nous conduisent à renouveler nos réserves sur la pertinence d'*ILI* puisque cet indicateur considère que les pertes incompressibles sont proportionnelles à la pression. Ceci est confirmé par l'étude du lien entre *ILI* et les classes de pression qui montre une tendance décroissante, significative selon les tests statistiques réalisés (Figure 12).

Bibliographie

- Alegre H., Baptista J.M., Cabrera E., Cubillo F., Duarte P., Hirner W., Merkel W. and Parena R. (2006), Performance Indicators for Water Supply Services. *IWA Publishing*, 289p.
- Ben Hassen F. (2012), Caractérisation et évaluation de la pression moyenne minimale d'une zone de desserte d'un réseau d'alimentation en eau potable. Irstea, 52p.
- Besse P. et Laurent B. (2013), Apprentissage Statistique : modélisation, prévision et data mining, Cours de 5^{ème} année GMM-MMS. 142p. Institut de mathématique de Toulouse, INSA de Toulouse [en ligne] <http://wikistat.fr/pdf/st-m-app-select.pdf>
- Bienvenue F. (2011), UE MASS 42 Statistiques. Université Claude Bernard Lyon 1 Année Universitaire 2011-2012. [en ligne] http://math.univ-lyon1.fr/homes-www/duheille/MASS42_poly.pdf
- Lambert A.O., Brown T.G. , Takizawa M. , Weimer D. (1999) A Review of Performance Indicators for Real Losses from Water Supply Systems. *Aqua 48 (1999) 227-237*. IWA Publishing.
- Le Glatin M. (2011), Impact de la pression et des infrastructures sur les réseaux d'eau potable. Veolia Eau. 50p.
- Rakatomalala R.(2012), Analyse de corrélation. Etude des dépendances – Variables quantitatives, Version 1.0. Université Lumière Lyon 2. 89p. [En ligne] : http://eric.univ-lyon2.fr/~ricco/cours/cours/Analyse_de_Correlation.pdf
- Renaud E. (2009), Valeurs de référence de l'indice linéaire de pertes des réseaux d'alimentation en eau potable. SMEGREG. 63p.
- Renaud E., Pillot J. (2012) Pertes en eau : Introduction de la pression dans les indicateurs de performance. Rapport de synthèse de la première année d'étude. Irstea, 40p.
- Renaud E., Pillot J. (2012) Pertes en eau : Introduction de la pression dans les indicateurs de performance. Rapport intermédiaire de la seconde année d'étude. Irstea, 33p.
- Renaud E, Clauzier M., Sandraz A.C., Pillot J., Gilbert D. (2013), Introduction de la pression et du nombre de branchements dans les indicateurs français de pertes des réseaux d'eau potable. Irstea, 15p.
- Sandraz A.C. (2012), Pertes en eau : Introduction de la pression dans les indicateurs de performance. *Rapport final de VDIF à l'intention du SEDIF*. Document interne. Veolia Eau D'Ile de France, 39p.
- Sissoko M. (2010), Maîtrise des pertes en eau potable : Méthode d'évaluation de la pression caractérisant une zone de desserte. Cemagref, 98p.
- StatSoft (2013), Concepts Fondamentaux en Statistique et Data Mining, Glossaire.[en ligne] <http://www.statsoft.fr/concepts-statistiques/glossaire/d/distance-cook.html>
- Tangeaoui N. (2012), Introduction de la pression et du nombre de branchements dans les indicateurs de performances caractérisant les pertes en eau d'un réseau. Veolia Eau, 54p.
- Tangeaoui N, Sandraz A.C. (2012), L'automatisation du calcul de la pression moyenne par secteur sur EPAnet. Document interne Veolia Eau, 26p.

Annexe : Article présenté à la conférence Efficient 2013 (Paris 22-24 octobre 2013)

Introduction de la pression et du nombre de branchements dans les indicateurs français de pertes des réseaux d'eau potable

Eddy RENAUD³, Marion CLAUZIER⁴, Anne-Claire SANDRAZ⁵, Julie PILLOT⁶, Denis GILBERT⁷

INTRODUCTION

La lutte contre les pertes en eau des réseaux de distribution d'eau potable nécessite un diagnostic de l'origine des pertes afin d'adapter au mieux les actions mises en œuvre pour les réduire. Il est donc important de se doter d'indicateurs de performance qui permettent, en première approche, de quantifier et de qualifier les volumes perdus selon leurs sources.

Les actions de lutte contre les fuites peuvent essentiellement porter sur les infrastructures et sur la pression de service. Il semble donc naturel que des informations attachées à ces deux aspects, telles le nombre de branchements et la pression moyenne de service, soient prises en compte pour caractériser le niveau des pertes.

En France, un système d'indicateurs de performance des réseaux d'eau potable est fixé par la réglementation. En matière d'évaluation des pertes, ce système rapporte le volume annuel perdu au volume annuel consommé ou à la longueur des canalisations mais aucun indicateur ne prend en compte la pression moyenne de service ou le nombre de branchements. A l'inverse, l'International Water Association (IWA) propose des indicateurs qui font intervenir ces paramètres. Des études antérieures ont cependant montré que ces indicateurs ne sont pas toujours adaptés aux services des eaux français dont les tailles sont très variables et dont les densités d'usagers peuvent être très faibles.

Dans ce contexte, Irstea (organisme public de recherche) et Veolia Eau (opérateur des services de l'eau) ont décidé d'étudier ensemble la mise au point d'indicateurs de pertes adaptés au contexte français prenant en compte la pression et le nombre de branchements.

Après avoir passé en revue les indicateurs de pertes usuels puis présenté les bases méthodologiques de nos travaux, nous présenterons les résultats obtenus sur nos terrains à l'échelle du secteur puis à l'échelle du service. Pour conclure nous nous appuieront sur nos résultats pour formuler des propositions sur les indicateurs et leur utilisation.

1 LES INDICATEURS DE PERTES USUELS

Les indicateurs usuels sont présentés en détail dans (Renaud et al., 2012).

1.1 Indicateurs de pertes utilisés en France

Les deux indicateurs des pertes les plus utilisés en France sont issus de la réglementation relative au rapport annuel sur le prix et la qualité du service (arrêté et décret du 2 mai 2007). Il s'agit du rendement (R) et de l'indice linéaire de pertes (ILP). Ils sont définis comme suit :

$$R = \frac{V_{ca} + V_{exp}}{V_{prod} + V_{imp}} (\%)$$

³ Irstea, Groupement de Bordeaux, Unité REBX, eddy.renaud@irstea.fr

⁴ Veolia Eau, Direction technique, marion.clauzier@veoliaeau.fr

⁵ Veolia Eau d'Ile-de-France, anne-claire.sandraz@veoliaeau.fr

⁶ Irstea, Groupement de Bordeaux, Unité REBX, julie.pillot@irstea.fr

⁷ Irstea, Groupement de Bordeaux, Unité REBX, denis.gilbert@irstea.fr

$$ILP = \frac{(V_{prod} + V_{imp}) - (V_{ca} + V_{exp})}{365 \times L} \left(\frac{m^3}{km \cdot jour} \right)$$

Avec :

V_{ca} , Volume annuel consommé autorisé (m^3/an)

V_{prod} , Volume annuel produit (m^3/an)

V_{exp} , Volume annuel exporté (ou vendu) (m^3/an)

V_{imp} , Volume annuel importé (ou acheté) (m^3/an)

L , Longueur du réseau de distribution hors branchements (km)

Ces deux indicateurs présentent plusieurs avantages. Ils sont clairement définis et font intervenir un nombre limité de données, ils sont régulièrement utilisés par l'ensemble des praticiens et sont aisément compréhensibles. Ils ne sont toutefois pas sans inconvénients.

Le rendement ne reflète pas uniquement le niveau des pertes et de leurs variations, il est très sensible aux consommations. (Agence de l'eau Adour Garonne, 2005) propose un exemple qui illustre clairement ce point.

« Une entreprise consomme $8\,000\ m^3/mois$ et le réseau urbain utilise $10\,000\ m^3/mois$, le niveau des pertes est de $7\,000\ m^3/mois$. Quand l'entreprise fonctionne, il doit être produit $25\,000\ m^3/mois$ pour couvrir l'ensemble des besoins, le rendement est alors égal à $(10\,000 + 8\,000) / 25\,000 = 72\ %$. Quand l'entreprise est fermée, la production nécessaire est de $17\,000\ m^3/mois$, le rendement vaut alors $10\,000 / 17\,000 = 58\ %$. »

Cet exemple simple démontre que pour un même niveau de pertes, le rendement peut varier fortement en fonction du niveau des consommations. La valeur de cet indicateur est donc délicate à interpréter et ses variations ne sont pas forcément représentatives des variations de performance du réseau du point de vue des pertes.

L'ILP, contrairement au rendement, est indépendant des variations de la consommation. Il permet en outre de rapporter les pertes à une caractéristique structurelle du réseau (le linéaire des canalisations). Il présente cependant un inconvénient majeur : sa valeur ne peut pas être analysée indépendamment du caractère urbain ou rural du réseau.

Il existe de multiples référentiels d'interprétation de l'ILP pour lesquels le caractère urbain ou rural est selon les cas basé sur les valeurs de l'Indice Linéaire de Consommation (ILC, consommation annuelle par kilomètre de canalisation et par jour) ou de la densité d'abonnés (D, nombre d'abonnés par kilomètre de canalisation). (Renaud, 2009) montre d'une part que l'appréciation de la performance du réseau est très variable selon le système de référence utilisé et d'autre part que les référentiels sont tous établis par classe (d'ILC ou de D), ce qui génère des effets de seuils importants : par exemple, l'ILP d'un service jugé « acceptable » une année peut devenir « mauvais » l'année suivante sans réelle dégradation de sa performance, si une faible variation de D ou ILC fait passer le réseau de la classe « urbain » à « semi urbain ».

1.2 Indicateurs de pertes de l'IWA

Parmi les indicateurs de pertes recommandés par l'IWA (Alegre et al., 2006), l'Infrastructure Leakage Index (ILI) est l'un de ceux dont il est fait un large usage à travers le monde depuis plus de dix ans. Sa formulation est la suivante :

$$ILI = \frac{CARL}{UARL} = \frac{CARL}{(18 \times L + 0.8 \times N_b + 25 \times L_b) \times P}$$

Avec

CARL, *Current Annual Real Losses (Litres/jour)*

UARL, *Unavoidable Annual Real Losses (Litres/jour)*

L, *Longueur du réseau hors branchements (km)*

N_b , *Nombre de branchements*

L_b , *Longueur totale des branchements de la voirie au compteur (km)*

P, *Pression moyenne de service, en mètres de colonne d'eau (mce)*

CARL exprime les pertes « réelles », c'est-à-dire les pertes au sens du système français diminuées des volumes liés aux vols d'eau et au sous-comptage des compteurs domestiques.

UARL représente les pertes dites « incompressibles » évaluées selon la formule proposée par (Lambert et al, 1999)

L'UARL repose sur l'idée qu'il existe un niveau minimal de pertes en dessous duquel il n'est pas envisageable de vouloir descendre. Ce seuil minimal, UARL, est basé sur les hypothèses récapitulées dans le Tableau 11, formulées dans le cas de réseaux en bon état et exploités dans les règles de l'art.

Tableau 11 : Hypothèses de calcul d'UARL pour une pression moyenne de 50 mce

	<i>Fuites indétectables</i>	<i>Fuites apparentes</i>	<i>Fuites apparentes non détectées</i>
Conduites	20 l/km/h	0.124 réparation/km/an à 12 m ³ /h pendant 3 jours	0.006 réparation/km/an à 6 m ³ /h pendant 50 jours
Branchement partie publique	1.25 l/brcht/h	2.25 réparation/1000 brcht/an à 1.6 m ³ /h pendant 8 jours	0.75 réparation/1000 brcht/an à 1.6 m ³ /h pendant 100 jours
Branchement partie privée	0.5 l/brcht/h pour 15 m de longueur de branchements	1.5 réparation/1000 brcht/an à 1.6 m ³ /h pendant 9 jours	0.5 réparation/1000 brcht/an à 1.6 m ³ /h pendant 101 jours

L'intérêt majeur d'ILI est qu'il s'agit d'un indicateur adimensionnel qui prend en compte plusieurs caractéristiques essentielles d'un réseau pouvant expliquer son niveau de pertes. Il est conçu pour représenter la part des pertes due aux infrastructures, indépendamment de la pression de service.

Son inconvénient majeur est qu'il ne s'applique pas à tous les réseaux. Ses concepteurs le préconisent pour les réseaux de plus de 3 000 branchements présentant une densité de branchements supérieure à 20 branchements/km. (Renaud, 2009) montre que de nombreux réseaux français ont des valeurs d'ILI inférieurs à 1 (ce qui signifie des pertes plus faibles qu'UARL et ne devrait pas être) et qu'au-delà, ILI est en moyenne proportionnel à la densité d'abonnés, alors que son interprétation se veut indépendante de la typologie du réseau.

En complément d'ILI, conçu pour mesurer les pertes liées aux infrastructures, (Trow, 2009) propose un indicateur pour mesurer l'impact de la pression sur les pertes, Pressure Management Index (PMI) :

$$PMI = \frac{CASP}{MARP}$$

Avec

CASP, Current Annual System Pressure (mce)

MARP, Minimal Annual Reference Pressure (mce)

CASP représente la pression moyenne de service tandis que MARP est la pression minimale au point de livraison (réglementaire ou conventionnelle) majorée de 3 mce.

(Trow, 2009) propose une utilisation combinée d'ILI et PMI pour mesurer les potentiels respectifs de réduction des pertes par la recherche active des fuites et par la gestion des pressions de service (Figure 30).

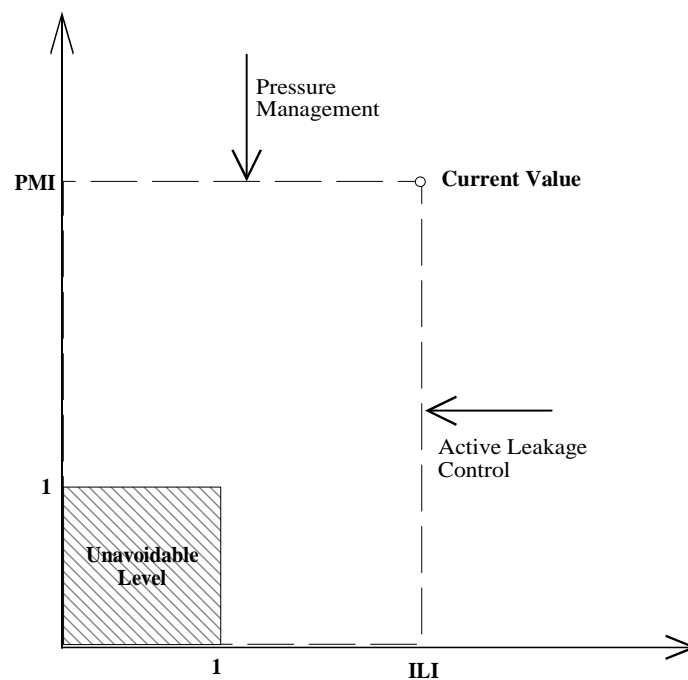


Figure 30 : Utilisation combinée de ILI et PMI

PMI présente l'intérêt incontestable de consacrer la pression comme paramètre à prendre en compte pour évaluer le potentiel de réduction des pertes. Sa faiblesse majeure, est de rapporter la pression de service à une pression minimale au point de desserte (MARP). Pour d'évidentes raisons hydrauliques, MARP ne peut pas être atteinte sur l'ensemble du réseau, une telle référence ne prend pas en compte l'impact déterminant de la configuration et du contexte topographique d'un réseau sur son potentiel de réduction de la pression.

2 BASES POUR LA CONSTRUCTION DE NOUVEAUX INDICATEURS DE PERTES

2.1 Objectifs

Il ressort de ce qui précède que, d'une part, la pression et le nombre de branchements ne sont pas pris en compte dans les indicateurs de performance français et que, d'autre part, les indicateurs internationaux qui les intègrent ne sont pas pleinement satisfaisants. Il semble donc opportun de rechercher de nouveaux indicateurs qui soient adaptés (entre autre) au contexte français et qui, comme ILI et PMI mais sans en comporter les inconvénients, intègrent le nombre de branchements et la pression.

(Renaud, 2010) propose un indicateur global de pertes (GLI, Global Leakage Index), qui est le produit d'ILI et de PMI. C'est en repartant de ce principe que nous avons défini la forme des trois indicateurs recherchés dans le cadre de l'étude.

Tableau 12 : Forme des indicateurs recherchés

Indicateur	Formulation	Données
Indice de Pertes Structurelles	$IPS = \frac{VP}{v_{ref} \times P^\alpha \times N_b^\beta}$	VP, volume annuel de pertes (m ³) P, pression moyenne de service (mce)
Indice de Réduction de la Pression	$IRP = \frac{P}{P_{min}}$	N _b , nombre de branchements v _{ref} , constante à déterminer (m ³)
Indice Global de Pertes	$IGP = IPS \times IRP^\alpha$	α et β exposants à déterminer P _{min} pression minimum qui peut être atteinte sur le service (mce). Méthode d'évaluation à définir.

2.2 Calcul de la pression moyenne d'une zone

Le calcul des indicateurs recherchés nécessite de disposer d'une évaluation de la pression moyenne de service de la zone étudiée. En s'appuyant sur les travaux de la Water Services Association of Australia (WSAA, 2009), (Sissoko, 2010) propose plusieurs méthodes d'évaluation de la pression d'une zone, selon le contexte et les données disponibles. Dans le cadre de l'étude et dans l'objectif de travailler sur un nombre important de services ou de secteurs, nous avons choisi d'utiliser la méthode « hydraulique ». Cette méthode est utilisable dès lors que l'on dispose d'un modèle hydraulique correctement calé avec des données de terrain. Elle utilise les pressions moyennes journalières de chaque nœud, déduites des pressions calculées à chaque pas de temps par le modèle. La pression moyenne de la zone est ensuite obtenue en calculant la moyenne pondérée des pressions moyennes journalières des nœuds. Trois systèmes de pondération sont envisagés : uniforme (poids identique pour tous les nœuds), selon la demande au nœud et selon la longueur de canalisations (chaque nœud est affecté de la demi-longueur des tronçons qui lui sont connectés).

Pour fiabiliser et automatiser le calcul de la pression moyenne d'une zone, nous avons développé des applications informatiques dédiées aux réseaux modélisés avec le logiciel EPANet.

Une première application permet d'affecter les nœuds du réseau aux secteurs. Pour cela un travail préalable de représentation des compteurs de sectorisation dans le modèle est requis. Il s'agit d'une opération simple qui consiste à attribuer la valeur « compteur » au champ « TAG » des nœuds concernés. A partir du modèle ainsi complété et enregistré dans le format « .inp », le programme, qui

inclut un algorithme de parcours de graphe, permet l'identification des secteurs et l'affectation de chaque nœud et de chaque tronçon du réseau à un secteur.

Une seconde application est dédiée au calcul des indicateurs de pression. L'identification des secteurs d'appartenance des nœuds ayant été mentionné dans le champ « TAG » de chacun d'eux, à partir du modèle au format « .inp », le programme lance un calcul hydraulique EPANet, en extrait les informations et résultats utiles et calcule, pour chaque secteur et pour le réseau dans son ensemble, les indicateurs suivants (cf. Figure 31) :

- A. « tag », identification du secteur. « tag » = « Network » correspond à l'ensemble du réseau
- B. « nbnodes », nombre de nœuds du groupe
- C. « demand m3 », demande totale du groupe (m3)
- D. « length », longueur totale de canalisations du groupe (m)
- E. « pmoy », moyenne non pondérée des pressions des nœuds du groupe
- F. « pmoyL », moyenne des pressions pondérées par la longueur de canalisations des nœuds du groupe
- G. « pmoyD », moyenne des pressions pondérées par la demande des nœuds du groupe dont la demande est strictement positive
- H. « pcrit », pression minimale parmi les pressions calculées à tous les pas de temps pour les nœuds dont la demande est strictement positive
- I. « IDcrit », identifiant du nœud dont la pression est égale à pcrit à un instant donné
- J. « P10 », 1er décile de la distribution des pressions des nœuds du groupe
- K. « P50 », 5ème décile de la distribution des pressions des nœuds du groupe
- L. « P90 », 9ème décile de la distribution des pressions des nœuds du groupe

	A	B	C	D	E	F	G	H	I	J	K
1	tag	nbnodes	demand m3	length	pmoy	pmoyL	pmoyD	pcrit	IDcrit	P10	P50
2	Network	791	1706.4636	253186.5	48.6	54	48.34	12.05	582	32.67	
3		713	1456.4886	233151	48.98	53.82	48.57	19.2	2	32.77	
4	che	1	6.477034	180	68.85	68.85	68.85	68.04	405	68.85	
5	PI1	3	3.598352	724	59.2	69.25	68.92	64.6	421	65.44	
6	PI17	2	7.436595	987.5	55.56	65.44	65.49	64.67	597	65.49	

Figure 31 : Exemple de résultats du programme de calcul automatique des indicateurs de pression

2.3 Echelles et terrains d'étude

L'évaluation du volume des pertes se fait par différence entre les volumes introduits dans le réseau et les volumes utilisés. Ces informations peuvent être obtenues à deux échelles, celle du secteur et celle du service.

A l'échelle du secteur, le volume mis en distribution est connu à partir des mesures des compteurs de sectorisation tandis que le volume consommé peut être obtenu avec les données de facturation, moyennant le rattachement des clients aux secteurs. A cette échelle, l'analyse des débits de nuit peut également permettre une estimation du volume des pertes.

A l'échelle du service, les données nécessaires à l'évaluation des pertes font partie de celles produites dans le cadre des rapports annuels. Les volumes mis en distribution résultent des relevés de production, d'import et d'export tandis que les données de consommation proviennent du relevé annuel des compteurs des usagers.

Pour mener nos investigations sur la recherche d'indicateurs de pertes pertinents, il est nécessaire de disposer d'informations complémentaires sur les facteurs susceptibles d'impacter le niveau des pertes, en particulier des données relatives au patrimoine et à la pression.

Au regard de ces contraintes, des services des eaux ont été sélectionnés parmi ceux exploités par Veolia Eau. A l'échelle du secteur, cinq services disposant d'une sectorisation opérationnelle, d'un Système d'Information Géographique (SIG) et d'un modèle hydraulique calé récemment ont été choisis. A l'échelle du service, afin de permettre le calcul automatique des indicateurs de pression, l'étude a porté sur 122 services disposant d'une modélisation hydraulique complète et cohérente réalisée avec le logiciel EPAnet.

Les investigations ont été menées successivement à l'échelle du secteur puis à l'échelle du service.

3 ETUDE A L'ECHELLE DU SECTEUR

3.1 Acquisition et traitement des données

Des données suffisantes ont été obtenues sur 100 secteurs parmi les 196 secteurs que comptent les 5 services sélectionnés. En effet, certains secteurs ont dû être écartés et d'autres regroupés en raison de l'absence ou de l'imprécision de certaines informations. Le cas le plus courant résulte de manques importants dans la chronique de mesure des débits, liés à des dysfonctionnements de la chaîne de mesure.

Si plus de la moitié des secteurs comptent entre 500 et 2000 branchements, on note toutefois que certains secteurs sont très grands (3 secteurs de plus de 90 000 branchements). Comme le montre le Tableau 13, le panel étudié est varié mais n'est pas homogène, il est plutôt urbain et présente des niveaux de pertes et de pression assez élevés.

Tableau 13 : Caractérisation des secteurs étudiés

Répartition des secteurs	Pression moyenne (mce)	Densité de branchements (branchements/km)	ILP m3/km/j	Rendements (%)
1^{ère} décile	38,4	15,3	2,0	53,7
Médiane	52,9	57,2	7,3	71,5
9^{ème} décile	75,3	165,7	18,2	90,5

Les données nécessaires à l'étude proviennent essentiellement des quatre sources que sont le SIG, la base de données clientèle, le modèle hydraulique et le système de gestion de la sectorisation. D'importants traitements de données ont été nécessaires pour mettre en cohérence les informations et calculer les données et indicateurs recherchés (Le Glatin, 2011), (Tangeaoui, 2012).

L'affectation des clients aux secteurs, nécessaire à l'évaluation des consommations par secteur, a été réalisée à partir des adresses. L'imprécision de certaines d'entre elles et l'existence de voies partagées entre plusieurs secteurs ont conduit à réaliser certaines affectations au prorata des longueurs de réseau concernées. Pour encadrer les incertitudes liées à ces traitements, des indices de fiabilité mesurant la proportion des affectations incertaines ont été construits, et les secteurs ayant des valeurs de ces indices supérieures à 30 % ont été écartés de l'étude.

Les volumes annuels distribués par secteurs ont été tirés des chroniques de débit. Ici encore, les secteurs avec une proportion de valeurs estimées ou corrigées supérieure à 30 % ont été écartés. Le calcul de la pression moyenne des secteurs a nécessité une affectation des nœuds aux secteurs, celle-ci a été réalisée de façon graphique. Les calculs effectués à partir des modèles hydrauliques ont servi de base à l'élaboration des spécifications des outils de calculs automatiques mentionnés au 2.2 (Tangaoui et Sandraz, 2012).

3.2 Analyse des facteurs explicatifs des pertes

Comme le montre la Figure 32, une relation nette existe entre le volume annuel perdu et le nombre de branchements. Une relation similaire existe entre le volume perdu et le linéaire de réseau ce qui est logique dans la mesure où le linéaire de réseau et le nombre de branchements sont deux variables indicatrices de la taille du secteur très corrélées entre elles.

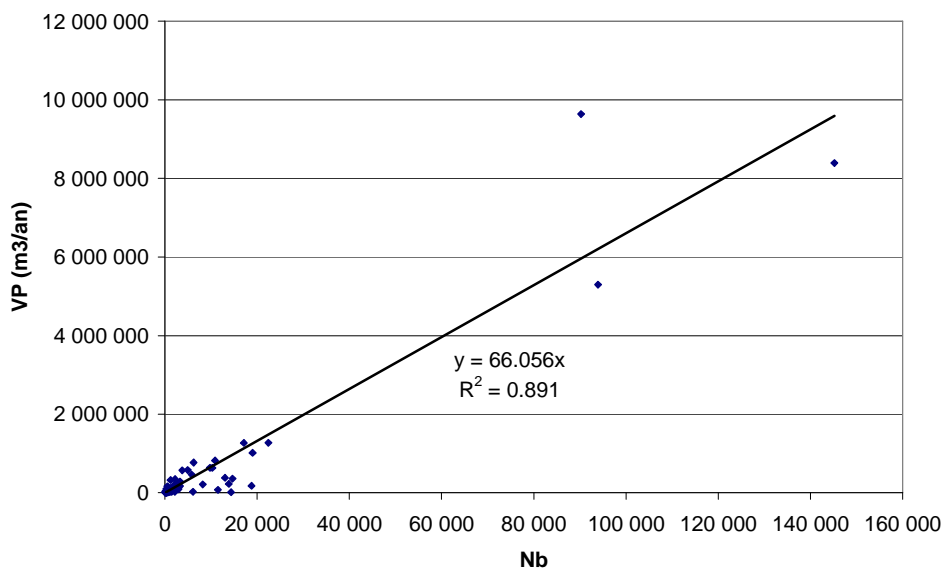


Figure 32 : Etude par secteurs – Volume annuel perdu en fonction du nombre de branchements

Ce résultat nous a conduits à considérer le volume annuel de pertes par branchement pour rechercher l'impact d'autres facteurs sur les pertes, indépendamment de la taille des secteurs.

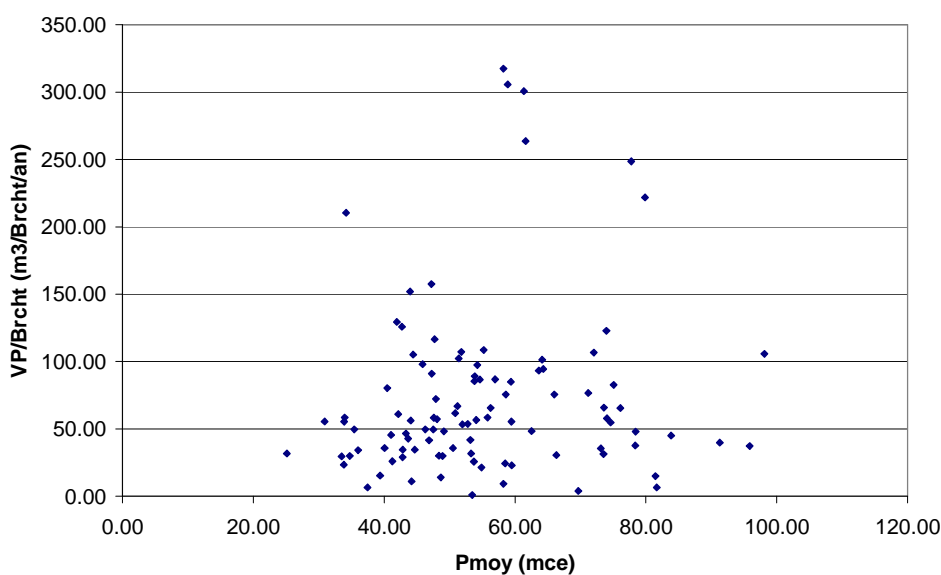


Figure 33 : Etude par secteurs – Volume de pertes par branchement en fonction de la pression moyenne

La Figure 33 montre qu'il n'y pas de lien significatif entre le volume annuel de pertes par branchement et la pression moyenne de service des secteurs. Ce résultat va à l'encontre de nos attentes. De nombreuses expériences ayant montré qu'en réduisant la pression moyenne d'un secteur on en diminue les volumes de pertes, il aurait semblé logique que le niveau de pertes par branchement soit corrélé positivement et significativement avec la pression moyenne. Ce n'est pas le cas sur nos données.

Des relations ont également été recherchées entre le niveau de pertes par branchement et le diamètre moyen, le pourcentage de canalisation en fonte, l'âge moyen des canalisations et les taux de casses

sur les secteurs où ces informations sont disponibles. De nouveau, aucun lien vraiment significatif n'a été trouvé.

Ces résultats nous ont amenés à nous interroger sur la fiabilité de nos données. En effet, comme mentionné au 3.1, l'évaluation du volume annuel de pertes des secteurs fait intervenir plusieurs sources de données et nécessite des traitements successifs qui sont autant de sources potentielles d'imprécisions. Lorsque cela était possible, les volumes de pertes utilisés ont été comparés à des estimations faites à partir des débits de nuits issus de la sectorisation. Ces contrôles ont montré dans certains cas des écarts sensibles entre les deux modes d'évaluation des pertes. Face à ces incertitudes, nous avons donc décidé de travailler à l'échelle du service afin de disposer d'une estimation plus fiable des volumes de pertes. Ces travaux font l'objet du chapitre 4.

3.3 Etude de la pression minimale d'une zone

Comme indiqué au 2.1, nous nous proposons de construire l'indicateur $IRP = \frac{P}{P_{\min}}$ (indice de réduction de pression), P_{\min} étant la pression moyenne minimale susceptible d'être atteinte sur le secteur avec les technologies usuelles et en garantissant à tout moment une pression suffisante à chacun des usagers (Ben Hassen, 2012).

Pour définir et évaluer P_{\min} , 10 secteurs parmi ceux précédemment évoqués ont été étudiés. Pour chacun des secteurs, la gestion de la pression a été simulée à l'aide d'un modèle hydraulique. La réalisation préalable d'un état de l'art des appareils de régulations (caractéristiques, domaines d'application) et la comparaison de leur représentation dans différents logiciels de modèles hydrauliques courants ont permis de s'assurer de la représentativité des simulations. Pour chacun des secteurs, après une phase de diagnostic, une même méthodologie a été mise en œuvre pour simuler des étapes progressives de réduction de la pression. Chaque solution envisagée doit d'une part, être possible technologiquement (conditions satisfaisantes de fonctionnement d'un appareil) et d'autre part, garantir une pression au moins égale à 20 m à l'ensemble des points de consommation.

Les solutions successives sont déterminées en analysant à la fois les pressions minimales en différents points stratégiques du réseau (réduction à consigne fixe), les variations journalières de la pression (modulation de la pression) et la configuration hydraulique du réseau (possibilité de délimiter certaines zones). La démarche est arrêtée lorsque la diminution locale de pression devient inférieure à 5 m ou lorsque la densité d'appareils à mettre en œuvre devient très importante. Pour chacune des solutions, la valeur de la pression moyenne P est calculée selon la méthode avec pondération par la demande au nœud, proposée par (Sissoko, 2010).

Nous avons pu constater que la diminution de pression entre deux solutions successives est décroissante et tend vers une limite. Pour déterminer cette limite, à chaque solution est associé un couple $\left(\frac{1}{1+n}, P\right)$, n étant le nombre total d'appareils de régulation de pression nécessaire à la mise en œuvre de la solution. P_{\min} est alors l'ordonnée à l'origine de la régression linéaire des points représentatifs des couples (Figure 34).

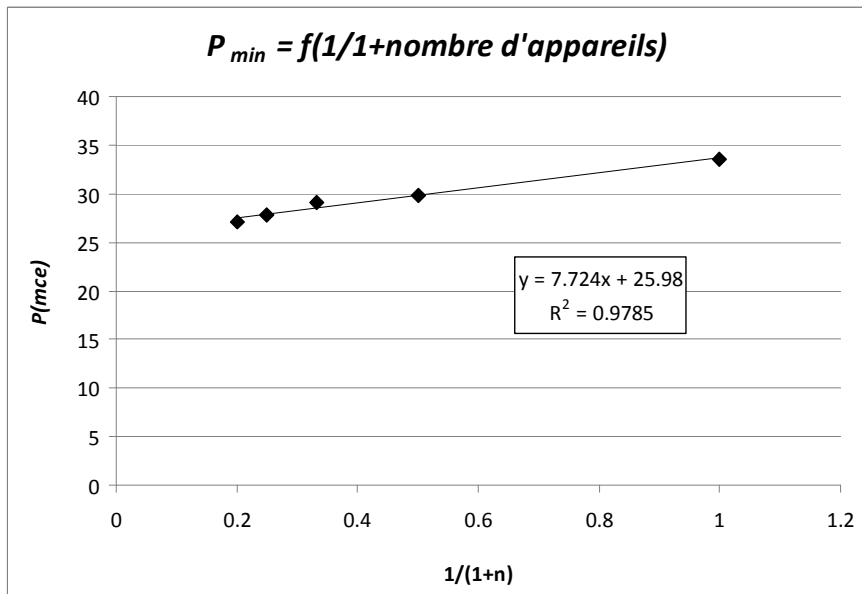


Figure 34 : Evaluation de Pmin pour le secteur 2 : $P_{min} = 25.98$ m

Nous avons ensuite recherché un lien entre les valeurs de P_{min} obtenues avec les méthodes précédemment définies et des indicateurs caractéristiques de la distribution spatiale des pressions sur la zone concernée. L'objectif était de disposer d'un moyen simple d'évaluer P_{min} sans recourir à des simulations hydrauliques.

Sur les 10 secteurs étudiés, un lien significatif est apparu entre $(P - P_{min})$ et $(P10\% - 20)$. $P10\%$ est le premier décile de la distribution des pressions, c'est-à-dire la pression au dessous de laquelle se situent 10 % des nœuds du réseau (Figure 35).

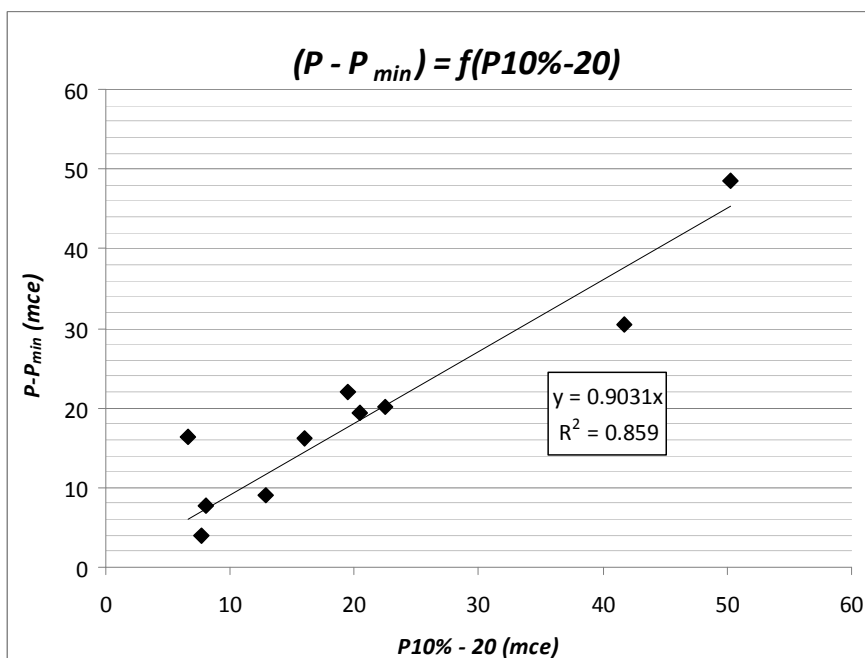


Figure 35 : Lien entre P_{min} et $P10\%$, premier décile de la distribution des pressions

Cette corrélation a permis de proposer la formulation suivante pour une évaluation approchée de P_{min} :

$$P_{min} = P - 0.9 \times (P10\% - 20)$$

4 ETUDE A L'ECHELLE DU SERVICE

4.1 Acquisition et traitement des données

Pour cette partie, afin de pouvoir calculer automatiquement les indicateurs de pression à l'aide de l'outil présenté au 2.2, nous nous sommes intéressés à des services exploités par Veolia Eau dotés d'une modélisation hydraulique du réseau avec le logiciel EPAnet. Nous avons ainsi recueillis les modèles hydrauliques et les informations du Rapport Annuel du Délégué (RAD) 2011 de 177 services. Après vérification de la qualité des modèles et des données issues des RAD, 122 services ont été conservés. Ensuite, la cohérence entre les données issues des RAD et les modèles a été vérifiée (linéaire de canalisations, nombre de branchements, demande), elle a été jugée satisfaisante pour 107 services. La distribution des pertes au sein de l'échantillon est exponentielle ce qui invite à privilégier la transformation logarithmique $\text{Ln}(VP)$. La distribution de $\text{Ln}(VP)$ s'approche d'une distribution symétrique, sous condition d'éliminer 7 services atypiques. Après élimination des services problématiques, 100 services ont finalement été retenus pour conduire l'étude (Figure 36).

Lecture des diagrammes en boîte : la boîte est bornée par le 1er quartile et le 3e quartile de la distribution, la médiane est représentée par un trait fort. Les extensions en pointillés s'étendent jusqu'à la dernière valeur distante de moins de 1.5 fois la longueur de la boîte à partir des bords de la boîte. Les occurrences dont les valeurs sont au-delà des pointillés sont repérées individuellement par un point.

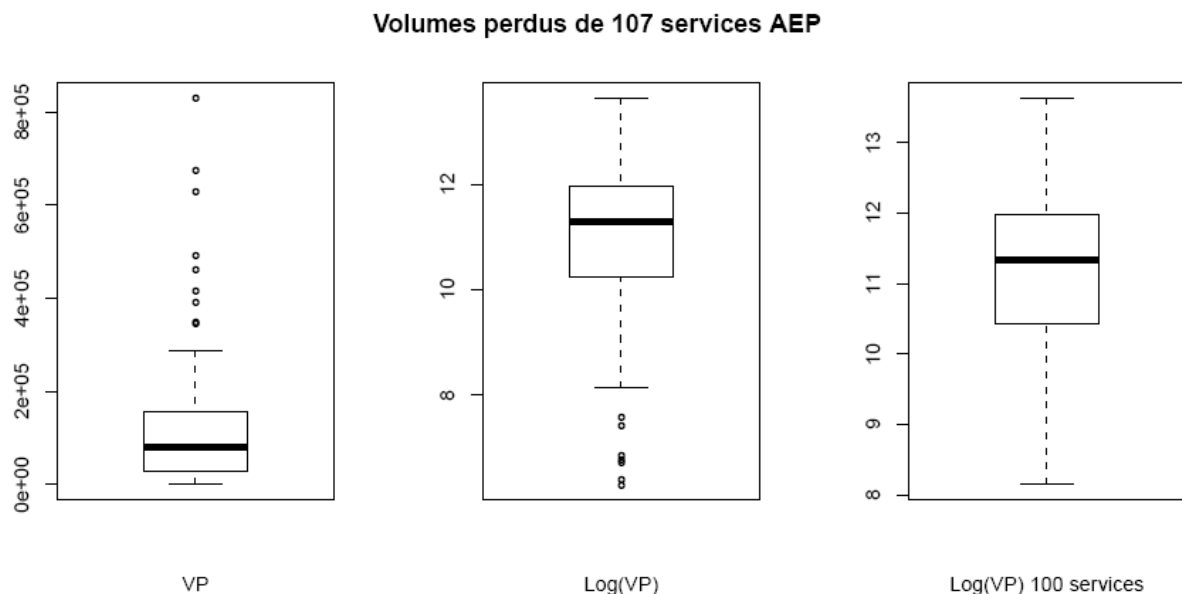


Figure 36 : Etude par services : Analyse de la distribution des volumes de pertes

Les données disponibles sur les services étudiés sont : l'ensemble des indicateurs de pression mentionnés au 2.2, le détail des volumes annuels mis en œuvre, des informations descriptives du service (linéaire, matériaux, diamètres, abonnés, branchements, ...) et les nombres annuels de fuites sur canalisation et sur branchement.

4.2 Analyse des facteurs explicatifs des pertes

L'analyse des corrélations entre le volume de pertes et les autres variables montre un lien significatif avec le linéaire, le nombre de branchements, le nombre d'abonnés et les volumes mis en œuvre. Aucune relation significative n'est mise en évidence avec les indicateurs de pression, les informations qualitatives sur le réseau (matériau, diamètre) ou les taux de fuite.

La corrélation entre le logarithme du volume de pertes et le logarithme du nombre de branchements est parmi les plus significatives, la pente proche de 1 de la courbe de tendance confirme une relation de quasi proportionnalité entre le volume de pertes et le nombre de branchements (Figure 37).

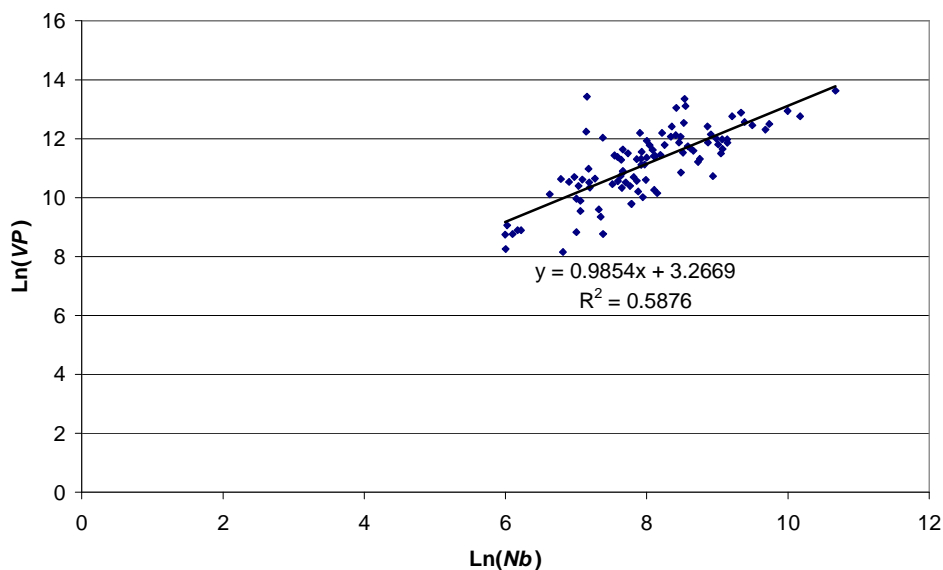


Figure 37 : Etude par services : Relation entre les logarithmes du volume de pertes et du nombre de branchements

De la même façon que dans l'analyse par secteur, nous avons donc recherché les liens potentiels entre le volume de pertes par branchement et les autres facteurs.

Aucune relation intéressante n'a été mise en évidence et en particulier, des tests statistiques ont montré qu'il n'y a pas de corrélation significative entre le volume de pertes par branchement et les indicateurs de pression.

Pour pousser plus loin les investigations, une analyse par classe de pression (pression moyenne pondérée par la demande) a été réalisée en constituant des groupes de 20 services (Figure 38).

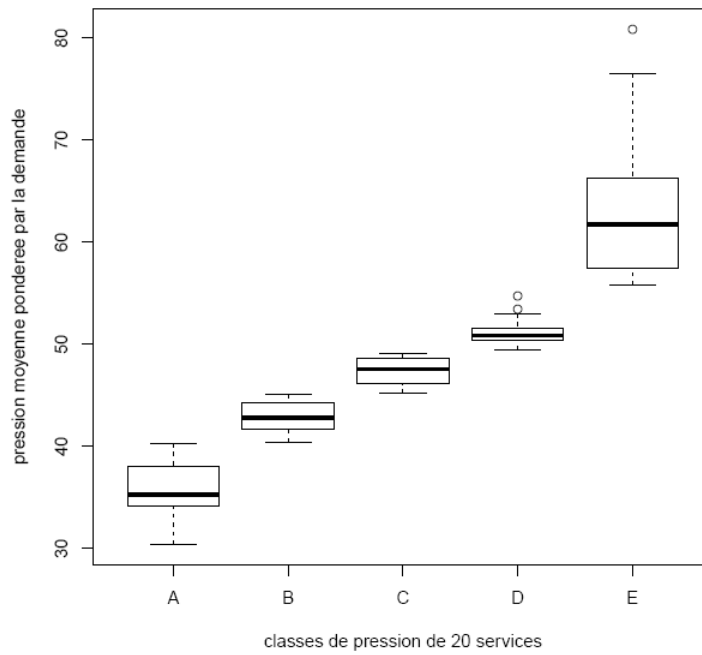


Figure 38 : Etude par services : Constitution des classes de pression

Cette analyse par classe confirme une absence de lien significatif entre le volume de pertes par branchement et la pression (Figure 39).

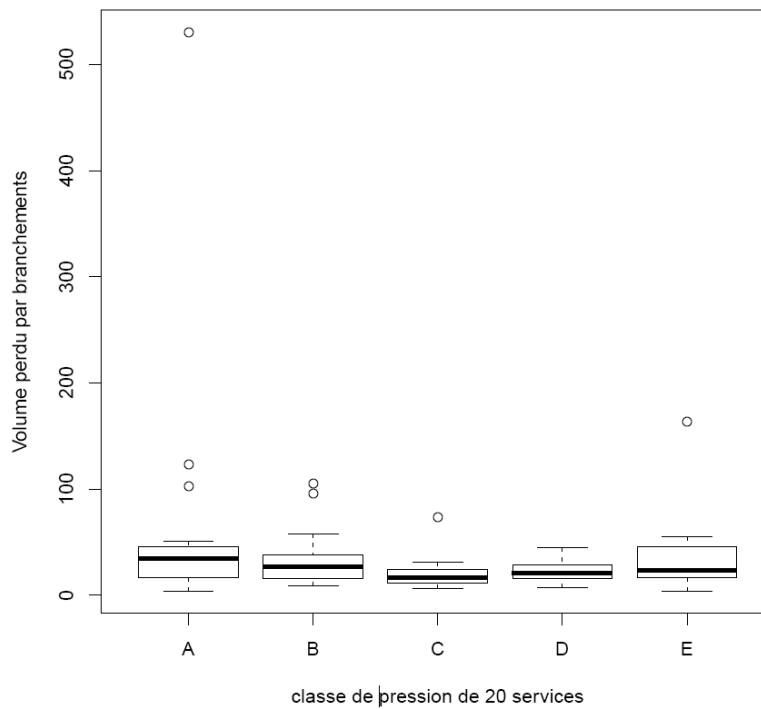


Figure 39 : Etude par services. Volume des pertes par branchement selon la classe de pression

CONCLUSION

Partant des expériences réussies de réduction des pertes obtenues grâce à la régulation de la pression, notre étude présupposait que la pression moyenne de service impacte significativement le niveau des pertes et qu'ainsi il était opportun de l'introduire dans les indicateurs de pertes français. Le système envisagé consistait à décomposer un indicateur global en deux composantes, l'une liée au potentiel de réduction des pertes par la régulation de pression et l'autre liée aux caractéristiques structurelles du réseau, dont le nombre de branchements.

Les investigations que nous avons menées, tant à l'échelle du secteur qu'à celle du service ont confirmé la pertinence de rapporter le volume annuel global de pertes au nombre de branchements, les deux grandeurs étant, en moyenne, liées par une relation de proportionnalité. En revanche, tous les résultats convergent pour conclure à une absence de relation significative entre le niveau des pertes et la pression moyenne de service.

Ce constat, bien qu'inattendu pour les raisons précédemment évoquées, nous semble robuste vu la variété et la profondeur de nos explorations (travail par secteur et par service, analyses statistiques individuelles et par classes).

En conséquence, nous pensons que l'idée de définir à l'échelle du service des indicateurs distincts dédiés aux aspects structurels et fonctionnels des pertes doit être abandonnée. Cependant, nous préconisons l'utilisation d'un indicateur de pertes par branchement pour évaluer de façon globale les pertes d'un service. Un tel indicateur existe dans le système proposé par l'IWA (Alegre et al, 2006).

Concernant le potentiel de réduction des pertes par la régulation de pression, nous proposons d'utiliser $IRP = \frac{P}{P_{\min}}$ tel que nous l'avons défini au 3.3 à l'échelle du secteur. Cet indicateur nous semble intéressant pour identifier les zones présentant le plus fort potentiel.

Enfin, nos résultats concernant l'absence d'impact significatif de la pression moyenne sur le niveau des pertes nous conduisent à renouveler nos réserves sur la pertinence d'ILI puisque cet indicateur considère que les pertes incompressibles sont proportionnelles à la pression. Ceci est confirmé par l'étude du lien entre ILI et les classes de pression qui montre une tendance décroissante qui est significative selon les tests statistiques réalisés (Figure 40).

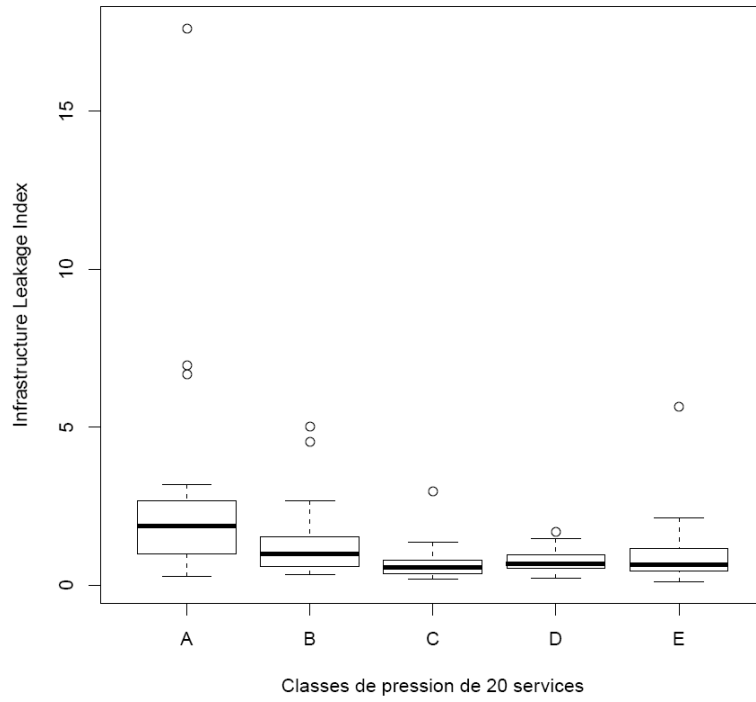


Figure 40 : Etude à l'échelle du service. Lien entre ILI et les classes de pression

BIBLIOGRAPHIE

- AEAG, SMEGREG, OIEau (2005), *Connaissance et maîtrise des pertes dans les réseaux d'eau potable*. Agence de l'eau Adour Garonne. 89p.
- Alegre H., Baptista J.M., Cabrera E., Cubillo F., Duarte P., Hirner W., Merkel W. and Parena R. (2006), *Performance Indicators for Water Supply Services*. IWA Publishing. 289p.
- Ben Hassen F. (2012), *Caractérisation et évaluation de la pression moyenne minimale d'une zone de desserte d'un réseau d'alimentation en eau potable*. Irstea, ENGEES, 52p.
- Lambert A.O., Brown T.G., Takizawa M., Weimer D. (1999) *A Review of Performance Indicators for Real Losses from Water Supply Systems*. Aqua 48 (1999) 227-237. IWA Publishing.
- Le Glatin M. (2011), *Impact de la pression et des infrastructures sur les réseaux d'eau potable*. Veolia Eau, ENSE3. 50p.
- Renaud E. (2009), *Valeurs de référence de l'indice linéaire de pertes des réseaux d'alimentation en eau potable*. SMEGREG. 63p.
- Renaud E. (2010), *Towards a global performance indicator for losses from water supply systems*. IWA Water Loss Conference. Sao Paulo, Brasil.
- Renaud E., Khedhaouiria D., Clauzier M., Nafi A., Wittner C, Wery, C (2012), *Réduction des fuites dans les réseaux d'alimentation en eau potable : Systèmes d'indicateurs et méthodologies pour la définition, la conduite et l'évaluation des politiques de lutte contre les fuites dans les réseaux d'eau potable*. Fiches pratiques. ONEMA
- Sissoko M. (2010), *Maîtrise des pertes en eau potable : Méthode d'évaluation de la pression caractérisant une zone de desserte*. Cemagref. 98p.
- Tangeaoui N. (2012), *Introduction de la pression et du nombre de branchements dans les indicateurs de performances caractérisant les pertes en eau d'un réseau*. Veolia Eau, Politech'Nice. 54p.
- Tangeaoui N, Sandraz A.C. (2012), *L'automatisation du calcul de la pression moyenne par secteur sur Epanet*. Document interne Veolia Eau
- Trow SW. (2009), *Development of a pressure Management Index (PMI)*, Water Loss Conference. IWA Cape Town, South Africa.
- WSAA (2009) *Guidelines relating to the calculation of Average Pressures in Water Distribution Systems and Zones*. Prepared by Wide Bay Water Corporation for Water Services Association of Australia, as part of WSAA Asset Monitoring Project PPs-3, November 2009. jaimie.hicks@wsaa.asn.au

Introduction de la pression et du nombre de branchements dans les indicateurs français de pertes des réseaux d'eau potable

E. RENAUD, M. CLAUZIER, A.C. SANDRAZ, J. PILLOT, D. GILBERT

RÉSUMÉ

En France, un système d'indicateurs de performance des réseaux d'eau potable est fixé par la réglementation. En matière d'évaluation des pertes, ce système rapporte le volume annuel perdu au volume annuel consommé ou à la longueur des canalisations mais aucun indicateur ne prend en compte la pression ou le nombre de branchements. A l'inverse, le système d'indicateurs de performance de l'IWA prend en compte ces paramètres, cependant, des études antérieures ont montré que son application est délicate dans le contexte français, caractérisé par des services de tailles très variables et des densités d'utilisateurs pouvant être très faibles. Face à ces constats, Irstea (organisme public de recherche) et Veolia Eau (opérateur des services de l'eau) ont décidé d'étudier ensemble la mise au point d'indicateurs de pertes pertinents adaptés au contexte français prenant en compte la pression et le nombre de branchements.

Dans un premier temps, l'étude a été menée à l'échelle du secteur. Les données de 100 secteurs répartis dans 5 réseaux exploités par Veolia Eau ont été recueillies, traitées et analysées. Ces travaux à l'échelle du secteur ont confirmé l'intérêt de rapporter les pertes au nombre de branchements, en revanche, ils n'ont pas permis de mettre en évidence un lien significatif entre le niveau de pertes et la pression moyenne. Une étude spécifique a défini la variable « pression moyenne minimum d'une zone » et propose des méthodes pour la calculer.

Dans un second temps, un module de calcul automatique de la pression moyenne d'un réseau modélisé avec le logiciel EPANet a été développé. Cet outil a rendu possible des investigations à l'échelle du réseau dans sa globalité sur 122 services des eaux. L'intérêt de travailler à cette échelle a été de fiabiliser l'évaluation du volume de pertes. Les résultats à l'échelle du service sont similaires à ceux obtenus à l'échelle du secteur.

En conclusion, l'étude a montré l'intérêt d'utiliser un indicateur de pertes par branchement. Elle a, à l'inverse, montré l'absence de lien entre le niveau de pertes et la pression moyenne, ce qui va à l'encontre de ce qui est couramment admis.

Mots clés : Réseau de distribution d'eau potable, lutte contre les fuites, indicateurs de performance, pression, branchement

Introducing pressure and number of connections into water loss indicators for French drinking water supply networks.

E. RENAUD, M. CLAUZIER, A.C. SANDRAZ, J. PILLOT, D. GILBERT

SUMMARY

In France, a system of performance indicators for drinking water supply networks is defined within the relevant regulations. This system assesses water loss based on the ratio of water lost per year to that consumed per year, or to pipe length. However, there are no indicators relating to pressure, or the number of connections. While the IWA performance indicator system does take these parameters into account, previous studies have shown that applying it to French water networks is somewhat difficult, due to the varying sizes of water utilities, along with some very low levels of user density. For these reasons, Irstea (French research institute) and Veolia Water (service operator) decided to create a new system of water loss indicators, specially designed for use in France, and taking into account pressure and the number of connections.

The study was initially carried out at District Metered Area (DMA) level. Data from 100 DMAs across five Veolia networks were collected, processed, and analysed. These analyses at DMA level showed the advantages of considering the ratio of losses to number of connections, but did not highlight a significant link between loss levels and average pressure. A parallel study on this topic defined the variable "minimum average zone pressure", as well as providing methods of calculation.

The second part of our study was to develop a module that automatically calculates average pressure based on an EPANet model of a network. Using this tool, it was possible to carry out network-wide analyses of around 150 different water systems. Working at this scale was particularly useful in better evaluating the volume of water losses. The results at the network scale were similar to those obtained at the DMA level.

In conclusion, the study demonstrates that it is relevant to use a performance indicator of losses per connection. On the other hand, it showed no link between the volume of losses and the average pressure, which is contrary to what is commonly admitted.

Keywords: Leakage management, Performance Indicators, Pressure, Service connections, Water supply systems