

ÉPURATION DES DONNÉES – PREUVE COMPLÉMENTAIRE

1 À la suite de la lettre de la Régie de l'énergie (« la Régie ») datée du 9 avril 2015 et des échanges
2 précédents avec la Fédération canadienne de l'entreprise indépendante (section Québec) (FCEI),
3 Société en commandite Gaz Métro (Gaz Métro) constate que des questionnements demeurent à
4 propos de la base de données comptables utilisée pour allouer les coûts des conduites
5 principales, plus précisément au niveau du nombre de conduites écartées lors du processus. Le
6 document qui suit tente d'apporter un éclairage supplémentaire sur cette question.

7 Afin de déterminer des coûts moyens représentatifs des différentes conduites, Gaz Métro procède
8 à une épuration de la base de données comptables. L'épuration permet de conserver les données
9 répondant à une loi normale. Les résultats obtenus à partir des approches statistiques et des
10 extrapolations sont alors meilleurs.

11 Lors du processus d'épuration, un certain nombre de données doit être écarté. Malgré un nombre
12 plus faible de conduites par matériel et diamètre, les résultats sont plus précis. L'important n'est
13 pas d'avoir un nombre maximal de conduites, mais bien des conduites qui sont représentatives
14 et qui permettent l'utilisation d'outils statistiques pour extrapoler ou régresser des données.

COMPARAISON DES DONNÉES ÉPURÉES ET NON ÉPURÉES

15 Afin de démontrer que l'épuration des données ne fait que préciser le résultat recherché, sans
16 toutefois le modifier outre mesure, le tableau 1 présente les résultats obtenus à partir des bases
17 de données épurée et non épurée.

Tableau 1

COÛTS COMPARATIFS
(coût moyen pondéré)

N° de ligne	Type	Diamètre	Épuré	Non épuré	Écart	Épuré sans 1979	Non épuré sans 1979	Écart
			(\$/m)	(\$/m)	(\$/m)	(\$/m)	(\$/m)	(\$/m)
			(a)	(b)	(c)	(d)	(e)	(f)
1	Plastique	42,2	157	169	-12	157	169	-12
2	Plastique	60,3	171	209	-38	172	209	-37
3	Plastique	88,9	181	213	-32	181	213	-32
4	Plastique	114,3	206	236	-30	205	233	-29
5	Plastique	168,3	229	263	-34	229	263	-34
6	Plastique	219,1	235	313	-78	235	313	-78
7	Acier	88,9	321	175	146	321	458	-137
8	Acier	114,3	370	448	-78	451	580	-129
9	Acier	168,3	412	459	-47	453	514	-61
10	Acier	219,1	547	643	-97	635	759	-124
11	Acier	323,9	508	560	-52	712	894	-182
12	Acier	406,4	603	664	-61	748	1 005	-257
13	Coût minimal par mètre linéaire		1	-84 422	84 422	1	-84 422	84 422
14	Coût maximal par mètre linéaire		1 942	209 663	-207 720	1 942	209 663	-207 720
15	Réseau de distribution seulement							
16	<u>Intercepte zéro</u>							
17	Valeur		143	150	-7	144	150	-6
18	Accès (%)		62	58	4	60	52	8
19	Capacité (%)		38	42	-4	40	48	-8
20	<u>Réseau minimal</u>							
21	Valeur		171	209	-38	172	209	-37
22	Accès (%)		74	81	-7	72	73	-1
23	Capacité (%)		26	19	7	28	27	1
24	Valeur du réseau		1 769 175 903	1 991 315 882	-222 139 979	1 835 732 555	2 190 336 910	-354 604 355
25	Réseau de distribution et d'alimentation							
26	<u>Intercepte zéro</u>							
27	Valeur		143	150	-7	144	150	-6
28	Accès (%)		53	49	4	49	42	7
29	Capacité (%)		47	51	-4	51	58	-7
30	<u>Réseau minimal</u>							
31	Valeur		171	209	-38	172	209	-37
32	Accès (%)		63	68	-5	58	58	0
33	Capacité (%)		37	32	5	42	42	0
34	Valeur du réseau		2 524 247 049	2 846 361 325	-322 114 276	2 742 503 558	3 312 483 574	-569 980 016
35	Valeurs relatives du réseau de distribution							
36	Distribution (%)		70	70	0	67	66	1
37	Alimentation (%)		30	30	0	33	34	(1)

EXPLICATION DES ÉCARTS POUR LE CALCUL DU RÉSEAU MINIMAL

- 1 Au niveau du réseau minimal, l'utilisation de la base de données non épurée fait augmenter
- 2 initialement l'accès et diminuer la capacité par rapport aux résultats obtenus à partir de la base

1 de données épurée (lignes 20 à 23 ainsi que lignes 30 à 33, colonnes (a), (b) et (c)). Par contre,
2 cet écart pourrait être le résultat d'un biais important dans la base de données non épurée.

3 Effectivement, comme il a été soulevé par l'Association des consommateurs industriels de gaz
4 (ACIG), les conduites posées avant 1979 peuvent causer un biais puisque leur prix moyen est
5 mis à jour à partir de l'année 1979 plutôt qu'à partir de l'année réelle de mise en terre. Dans sa
6 réponse à la demande de renseignements n° 3 de la Régie, question 2 (pièce B-0083,
7 Gaz Métro-3, Document 11), Gaz Métro a expliqué que le processus d'épuration éliminait toutes
8 les données empêchant l'obtention d'une répartition normale des données; par conséquent, si les
9 résultats de ces conduites étaient hors norme, ceux-ci étaient déjà éliminés.

10 Donc, pour que les résultats soient comparables, Gaz Métro a également calculé dans le
11 tableau 1 les résultats de la base de données épurée et de la base de données non épurée en
12 retirant toutes les conduites de 1979 (colonnes (d), (e) et (f)). Une fois les conduites enlevées,
13 l'écart entre les résultats obtenus est presque inexistant (lignes 20 à 23 ainsi que lignes 30 à 33).

EXPLICATION DES ÉCARTS POUR LE CALCUL DE L'INTERCEPTE ZÉRO

14 Dans le cas de l'intercepte zéro, les résultats obtenus à l'aide des bases de données épurée et
15 non épurée sont différents (lignes 16 à 19 et lignes 26 à 29).

16 L'intercepte zéro est une approche qui nécessite un niveau de qualité des données élevé afin
17 d'assurer la cohérence des résultats. Avec une telle approche, l'épuration de la base de données
18 permet d'obtenir des données représentatives par conduite sous forme de courbe normale. La
19 cohérence des données est importante pour que l'intercepte soit valide.

20 Les coûts minimum et maximum observés dans la base de données non épurée de -84 422 \$ à
21 207 720 \$ par mètre linéaire sont passablement hors-norme. L'origine de ces coûts impossibles
22 ne peut être identifiée précisément, mais Gaz Métro fournit des pistes sur les raisons qui peuvent
23 expliquer ces incohérences dans la réponse à la demande de renseignements n° 1 de la Régie,
24 question 8.2 (pièce B-0048, Gaz Métro-3, Document 1). Par contraste, la dispersion est plus
25 probable dans la base de données épurée avec des conduites ayant des coûts de 1 \$ par mètre
26 linéaire jusqu'à des coûts de 1 942 \$ par mètre linéaire. *A priori*, l'approche épurée devrait donc
27 donner un meilleur résultat.

1 Pour confirmer cette hypothèse, les résultats doivent être comparés avec des coûts connus. Lors
 2 de ses analyses initiales, Gaz Métro a tenté d'établir un intercepte théorique basé sur les coûts
 3 de la construction. Afin d'y arriver, Gaz Métro a demandé à ce que les coûts de matériel soient
 4 séparés des coûts d'installation et des frais généraux.

Tableau 2

Coût moyen des conduites
(2012)

Matériel	Diamètre	Coût moyen matériel	Coût moyen d'installation et frais généraux	Coût moyen
Plastique	60,3	5	174	179
Plastique	114,3	16	186	202
Plastique	168,3	32	195	227

5 En observant ce tableau, il est possible de constater que la hausse des coûts de matériel est
 6 différente de la hausse des coûts moyens d'installation et des frais généraux. Cela s'explique par
 7 le fait que le coût de la machinerie et des salaires est à peu près le même peu importe le diamètre
 8 à installer puisque la tranchée est la même pour tous les diamètres de conduites de plastique.

9 Dans la méthode de l'intercepte zéro, la régression est effectuée sur le coût moyen total.
 10 L'extrapolation pour déterminer le point zéro considère donc, de façon combinée, autant la
 11 hausse des coûts de matériaux que les coûts d'installation. Or, dans le cas des coûts moyens
 12 d'installation et de frais généraux, la relation en fonction du diamètre est relativement linéaire,
 13 alors que ce n'est pas du tout le cas pour les coûts moyens liés aux matériaux.

14 Il est donc possible de constater que le passage de 60,3 à 114,3 et à 168,3 mm engendre une
 15 hausse d'environ 10 \$ des coûts d'installation et des frais généraux seulement pour chaque
 16 augmentation de diamètre de 54 mm. Une extrapolation simple de ces données permet donc
 17 d'estimer que le coût moyen d'installation et de frais généraux d'une conduite à diamètre zéro
 18 devrait être d'environ 10 \$ de moins que le coût d'une conduite de 60,3 mm.

19 Dans le cas du matériel, le coût moyen devrait être nul, dans la mesure où une conduite de
 20 diamètre nul n'a pas de matériel qui la compose. Le coût moyen en matériel serait donc 5 \$ moins
 21 cher que celui d'une conduite de 60,3 mm. Ainsi globalement, la pose d'une conduite à diamètre

1 zéro devrait donc coûter environ 15 \$ de moins par mètre linéaire que la pose d'une conduite de
2 plastique de 60,3 mm.

3 Voici maintenant un tableau qui compare la valeur de l'intercepte en utilisant la base de données
4 épurée et non épurée et qui permet de porter un jugement sur l'écart constaté.

Tableau 3

	Intercepte	60,3	Écart	Écart
			(\$)	(%)
Coût moyen construction - Intercepte calculé	164	179	15	8,4 %
Base de données épurée	143	171	28	16,4 %
Base de données non épurée	150	209	59	28,2 %
Base de données épurée sans 1979	144	172	28	16,3 %
Base de données non épurée sans 1979	150	209	59	28,2 %

5 Comme la variation des coûts d'installation est stable à environ 10 \$ par accroissement de 54 mm
6 de diamètre et comme la réduction maximale du coût du matériel peut être établie à 5 \$, alors le
7 résultat le plus fiable sera celui dont l'écart entre l'intercepte et la conduite de 60,3 mm est le plus
8 près de 15 \$.

9 L'hypothèse de départ peut alors être confirmée, soit que la base de données épurée permet le
10 calcul d'un intercepte plus significatif puisque l'écart relatif entre l'intercepte et la conduite de
11 60,3 mm de 28 \$ est plus près de 15 \$ que pour la base de données non épurée.

12 Donc, la plus grande cohérence de la base de données épurée ainsi que son écart plus petit par
13 rapport au comparatif permet de déterminer que son calcul de l'intercepte est plus précis, et ce
14 avec ou sans les conduites de 1979.

CONCLUSION

15 Gaz Métro estime que ce complément de preuve permet d'écarter l'idée que l'épuration des
16 données amène une variance importante et inexplicable sur le calcul de l'accès et de la capacité.
17 Par ailleurs, bien que la base de données épurée donne des résultats se rapprochant plus des
18 coûts de construction et des résultats permettant de meilleurs calculs statistiques, Gaz Métro ne
19 s'opposerait pas à l'utilisation de la base de données sans épuration autre que les conduites de

- 1 1979 si la méthode du système minimal calculée de façon globale était approuvée. Les autres
- 2 approches ayant des besoins de cohérences statistiques plus élevées nécessitent cependant
- 3 l'épuration des données.