Observatoire national sur les incidences des émissions de contaminants sur la santé et l'environnement

Réanalyse de la relation entre la concentration en arsenic des concentrés alimentés au réacteur de la Fonderie Horne et la concentration en arsenic à la station ALTSP1

onicse

Août 2026

Auteur

Sébastien Rivest, agent de recherche en science des données
Observatoire national sur les incidences des émissions de contaminants sur la santé et l'environnement

Sous la direction de

Daniel Proulx, directeur
Observatoire national sur les incidences des émissions de contaminants sur la santé et l'environnement

Réviseur(e)s

Jean-François Boulanger, professeur agrégé à la Faculté de génie
Département de génie chimique et de génie biotechnologique, Université de Sherbrooke

Yohann Moanahere Chiu, professeur adjoint à la Faculté de médecine et des sciences de la santé
Département de médecine de famille et de médecine d'urgence, Université de Sherbrooke

Nolwenn Noisel, professeure agrégée à l'École de santé publique
Département de santé environnementale et santé au travail, Université de Montréal

Observatoire national sur les incidences des émissions de contaminants sur la santé et l'environnement
Université du Québec en Abitibi-Témiscamingue

445 boulevard de l'Université
Rouyn-Noranda (Québec) J9X 5E4
onicse@uqat.ca

Table des matières

Résumé 1
Introduction 2
Méthodologie 3
Résultats et discussion 4
Références 7
Annexe 1 : accessibilité du script est des données 9

Résumé

Glencore a réalisé, en 2022, une analyse visant à évaluer la relation entre la teneur en arsenic des concentrés alimentant le réacteur de la Fonderie Horne et la concentration d'arsenic mesurée dans l'air ambiant à la station légale de la fonderie (ALTSP1). Cette analyse concluait à l'absence de relation statistiquement significative entre ces deux variables. Toutefois, des enjeux méthodologiques liés à l'analyse des données et à l'interprétation des résultats statistiques ont été constatés, justifiant une réanalyse des données utilisées.

Cette réanalyse met en évidence une association positive et statistiquement significative entre la quantité d'arsenic contenue dans les intrants traités par la Fonderie Horne et les concentrations d'arsenic mesurées dans l'air à la station légale ALTSP1. Autrement dit, une augmentation de la quantité d'arsenic dans les concentrés alimentant le réacteur est associée à des concentrations d'arsenic plus élevées mesurées à ALTSP1. Les résultats montrent également que la direction des vents influence les concentrations d'arsenic observées à la station ALTSP1, les concentrations étant plus élevées lorsque les vents soufflent selon l'axe reliant le secteur du réacteur à la station. Ces résultats suggèrent que la quantité d'arsenic traitée par la Fonderie et la direction des vents contribuent toutes deux à expliquer les variations des concentrations d'arsenic mesurées dans l'air à la station légale ALTSP1.

Introduction

Dans le cadre de l'autorisation ministérielle (2023-2028) de Glencore pour la Fonderie Horne, l'entreprise a produit en 2022, à la demande du ministère de l'Environnement, de la Lutte contre les changements climatiques, de la Faune et des Parcs (MELCCFP), une analyse de la relation entre la concentration en arsenic des concentrés alimentés au réacteur de la Fonderie et la concentration en arsenic dans l'air ambiant à la station légale ALTSP1 (Prévost et Viger, 2022). Cette analyse conclut qu'il n'y a pas de relation statistique significative entre ces deux concentrations. Cette conclusion a toutefois été contestée par le MELCCFP qui rapporte que l'analyse de Glencore n'offre pas de preuve de l'absence de relation significative. Le MELCCFP affirme également que selon ses propres analyses, il existe une relation significative[1] (MELCCFP, 2023). Cependant, ces analyses ne sont pas accessibles pour évaluation.

Des enjeux apparaissent à la lecture de l'étude de Glencore concernant les analyses et l'interprétation statistique. D'abord, la significativité statistique a été évaluée en utilisant le coefficient de détermination (R^2), alors que cette valeur n'offre aucune information sur la significativité de la relation entre les concentrations. Le R^2 quantifie la proportion de la variance de la variable dépendante (ici la concentration à la station ALTSP1) expliquée par le modèle. Ainsi, un modèle peut présenter un faible R^2 tout en contenant des relations statistiquement significatives, mais peut aussi présenter un R^2 élevé sans que les effets inclus ne soient significatifs (Kutner 2005). De plus, la valeur de R^2 est comparée à des seuils arbitraires qui ne sont pas justifiés et qui varient selon les analyses. Les informations nécessaires à l'interprétation de la significativité statistique de la relation entre la concentration en arsenic à ALTSP1 et la concentration introduite au procédé ne sont pas fournies par Glencore. Ensuite, les comparaisons entre modèles réalisées dans l'étude de Glencore rapportent une diminution du R^2 avec l'ajout de certaines variables dans les modèles. Toutefois, avec les mêmes données, l'ajout de variables explicatives ne peut diminuer le R^2 non ajusté (Kutner 2005). Par ailleurs, malgré la conclusion de l'étude selon laquelle cet ajout n'aurait pas d'impact significatif, le R^2 ne permet pas d'évaluer la significativité de l'inclusion de variables au modèle. En raison de ces enjeux, l'étude de Glencore ne permet pas de déterminer la significativité de la relation entre la concentration d'arsenic mesurée à ALTSP1 et la concentration introduite dans le procédé.

[1] « Les analyses statistiques réalisées par le MELCCFP montrent que la quantité d'arsenic alimentée a une influence statistiquement significative sur la concentration mesurée à la station ALTSP1. Statistiquement significative, dans ce contexte, cela signifie qu'il est improbable d'observer un lien aussi fort entre les deux variables simplement par le fruit du hasard. Par ailleurs, il suffit de montrer qu'une augmentation de l'alimentation en arsenic augmente la concentration mesurée; il n'est pas nécessaire de démontrer que cette relation est linéaire. Également, comme le mentionne Glencore Fonderie Horne (GFH) dans son commentaire, il va sans dire que la direction du vent est une autre variable qui influence de façon importante la concentration mesurée à la station ALTSP1. Cela ne signifie pas pour autant que d''autres facteurs ne peuvent avoir un impact statistiquement significatif sur les concentrations mesurées à la station ALTSP1. Enfin, de l'avis du MELCCFP, l'analyse présentée par GFH ne donne pas la preuve, au sens statistique, que la relation entre la quantité d'arsenic alimentée et la concentration mesurée à la station ALTSP1 est non significative. » (MELCCFP, 2023 ; Volet Émissions atmosphériques, S. 6, condition 1, note 1).

Le présent document présente une réanalyse des données fournies par Glencore afin d'évaluer la relation entre la concentration en arsenic à la station légale (ALTSP1) et la quantité d'arsenic dans les concentrés traités par la Fonderie Horne.

Méthodologie

Afin d'évaluer la relation entre la concentration en arsenic à la station légale ALTSP1 et la concentration en arsenic des concentrés alimentés au réacteur de la Fonderie (voir Figure 1), les données fournies par Glencore dans le document « Étude sur la concentration en arsenic des concentrés alimentés au réacteur et la concentration en arsenic mesurée dans l'air ambiant à la station légale ALTSP1 » ont été utilisées (Prévost et Viger, 2022). Ces données couvrent la période de janvier 2018 à juin 2021 et se présentent sous la forme de deux jeux de données : l'un mensuel et l'autre journalier[2].

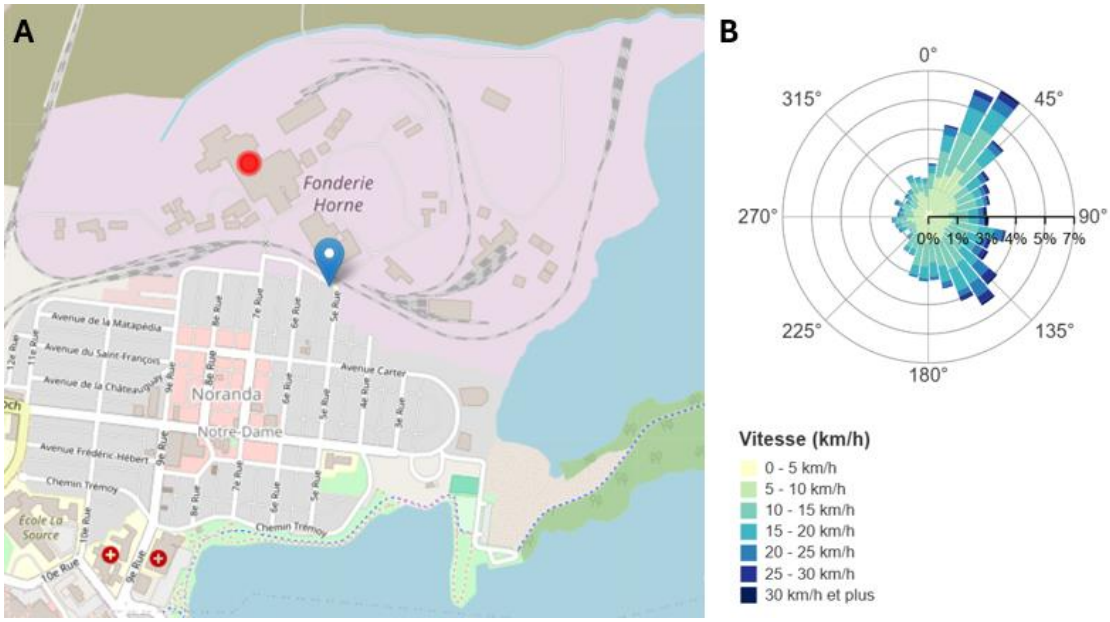

Figure 1. (A) Position de la station ALTSP1 (épingle bleue) par rapport au secteur du réacteur de la Fonderie Horne (point rouge). (B) Rose des vents illustrant la direction vers laquelle soufflent les vents ainsi que leur vitesse pour la période de janvier 2018 à juin 2021.

L'ensemble des analyses a été effectué avec le logiciel R, version 4.4.2 (R Core Team 2025). Les librairies et fonctions utilisées sont présentées à l'annexe 1. Pour les données mensuelles et journalières, la corrélation de Pearson a été évaluée entre la concentration en arsenic à la station ALTSP1 et la quantité d'arsenic dans les concentrés.

Pour l'analyse des données mensuelles, deux modèles linéaires ont été développés avec la concentration en arsenic à la station ALTSP1 comme variable dépendante. Dans le premier modèle,

[2] Les jeux de données sont accessibles à l'adresse suivante : https://donnees.onicse.uqat.ca/dataset/relation-intrants-extrants-fonderie-horne.

la masse d'arsenic alimentée via les concentrés a été utilisée comme variable explicative, alors que pour le second modèle, le pourcentage d'arsenic dans ces mêmes concentrés a été utilisé. Les mois d'octobre et novembre 2018 ainsi que celui de juin 2020 ont été exclus des analyses de Glencore (Prévost et Viger, 2022) puisque des arrêts généraux avaient lieu à la Fonderie Horne. Ils ont également été exclus de la présente analyse par souci de comparabilité. Le nombre moyen d'heures de vent soufflant dans l'axe Fonderie Horne – station ALTSP1 (337-346°) et la vitesse moyenne des vents dans cette direction ont été inclus comme covariables dans chaque modèle pour contrôler l'influence des conditions de vent sur la concentration en arsenic à la station ; l'intervalle angulaire retenu (337-346°) a été déterminé par Glencore (Prévost et Viger, 2022). La vitesse du vent dans l'axe Fonderie Horne – Station a été fixée à zéro pour les jours où le nombre d'heures de vent dans cet axe était nul (à noter que l'inclusion de cette variable n'influence pas les conclusions des modèles). Les valeurs journalières de vitesse des vents ont été agrégées sous forme de moyennes mensuelles. Une transformation par racine carrée a été appliquée aux concentrations en arsenic à ALTSP1 afin d'améliorer la normalité des résidus du modèle. Le respect des hypothèses des modèles a d'abord été vérifié graphiquement. Les hypothèses ont ensuite été validées par le test d'hétéroscédasticité de Breusch-Pagan, le test d'indépendance des erreurs de Durbin-Watson et le test de normalité de Shapiro-Wilk. Finalement, la colinéarité entre variables a été vérifiée en mesurant le facteur d'inflation de la variance (VIF).

Pour l'analyse des données journalières, deux modèles linéaires généralisés ont été développés. La même formulation que pour les modèles mensuels a été reprise, à l'exception de la distribution des erreurs : compte tenu de l'asymétrie marquée vers la droite (peu de valeurs élevées) des concentrations en arsenic mesurées à la station ALTSP1, une distribution log-normale a été utilisée. Cette distribution est appropriée lorsque la variable dépendante est continue, strictement positive et asymétrique (Bolker 2008). La distribution log-normale est fréquemment utilisée pour modéliser les concentrations en contaminants dans l'air (voir Marchant et al. 2013 et références incluses). Les valeurs du nombre d'heures de vent soufflant dans l'axe Fonderie Horne – station ALTSP1 ont subi une transformation par racine carrée afin d'améliorer l'ajustement du modèle. Puisque les données journalières représentent une série temporelle, la présence d'autocorrélation temporelle dans les données a été vérifiée et une composante aléatoire temporelle avec une structure d'autocorrélation de premier ordre (« AR1 ») a été intégrée au modèle. L'intégration de l'autocorrélation rend les tests statistiques plus conservateurs, réduisant le risque de conclure à tort à un effet significatif. La multicolinéarité entre variables explicatives a été examinée à l'aide des facteurs d'inflation de la variance (VIF) ; tous les VIF étaient inférieurs à 1,3 (O'Brien 2007). L'ajustement du modèle et le respect des hypothèses ont été évalués par l'analyse de l'uniformité et de la dispersion des résidus, de la présence de valeurs extrêmes, de l'hétéroscédasticité, de la colinéarité et de l'adéquation prédictive.

Résultats et discussion

La corrélation entre la quantité d'arsenic dans les intrants traités par la Fonderie Horne et la concentration en arsenic à la station légale ALTSP1 est de 0,52 à l'échelle mensuelle et de 0,13 à l'échelle journalière. La corrélation plus élevée pour les moyennes mensuelles peut s'expliquer par

une plus faible influence des variations météorologiques quotidiennes sur les concentrations moyennes mensuelles comparativement aux valeurs journalières (voir plus bas).

L'analyse des données mensuelles et journalières démontre une association positive significative entre la quantité d'arsenic dans les intrants traités par la Fonderie Horne et la concentration en arsenic à la station légale ALTSP1 ($p = 0,0004$ pour les données mensuelles et $p < 0,0001$ pour les données journalières ; Figures 2 et 3). Les mêmes conclusions ont été obtenues en utilisant le pourcentage (concentration), plutôt que la quantité, d'arsenic dans les intrants comme variable explicative ($p = 0,0008$ pour les données mensuelles et $p = 0,0001$ pour les données journalières ; les graphiques ne sont pas montrés pour cette variable explicative puisqu'ils sont très similaires aux Figures 2 et 3). Dans l'ensemble des modèles, un effet significatif du nombre d'heures de vent dans l'axe Fonderie Horne – station ALTSP1 sur la concentration en arsenic à la station ALTSP1 a également été observé ($p < 0,001$ pour l'ensemble des analyses).

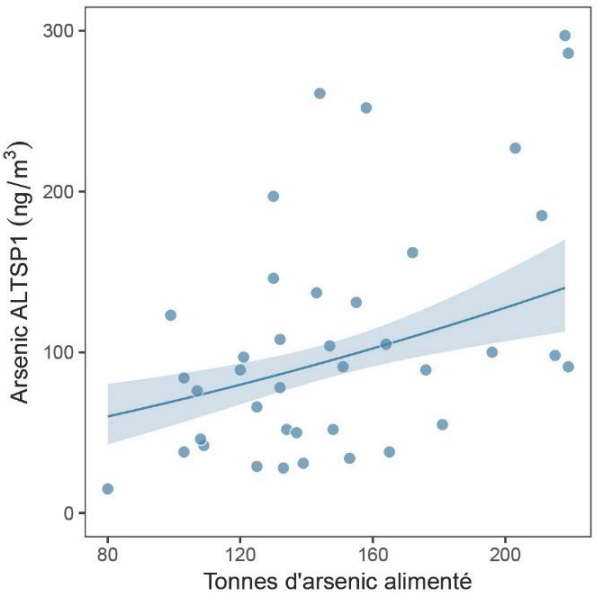

Figure 2. Relation entre la quantité mensuelle d'arsenic alimentée au réacteur de la Fonderie Horne et la concentration moyenne mensuelle en arsenic à la station ALTSP1. La ligne continue représente la prédiction du modèle, et la surface ombragée représente l'intervalle de confiance à 95 % autour de la prédiction, alors que les points représentent les valeurs mesurées.

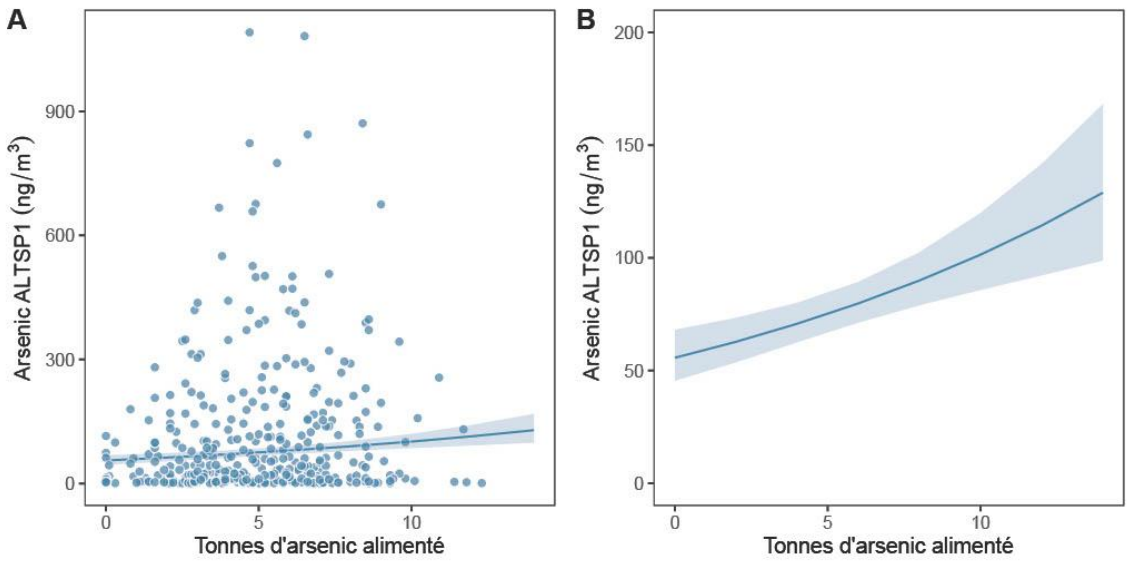

Figure 3. Relation entre la quantité journalière d'arsenic alimentée au réacteur de la Fonderie Horne et la concentration journalière en arsenic à la station ALTSP1 (A). La ligne continue représente la prédiction du modèle, et la surface ombragée représente l'intervalle de confiance à 95 % autour de la prédiction, alors que les points représentent les valeurs mesurées. Les prédictions seules (B) sont présentées afin de faciliter la visualisation des tendances.

En se fondant sur les données mensuelles, le modèle prédit qu'un doublement du tonnage d'arsenic introduit est associé à une augmentation moyenne d'environ 80 % de la concentration en arsenic mesurée à la station ALTSP1, toutes choses égales par ailleurs (Figure 2). En comparant les valeurs minimales et maximales du tonnage mensuel en arsenic introduit (80 et 219 tonnes, respectivement), cela correspond à une concentration en arsenic environ 2,3 fois plus élevée à la station ALTSP1. À l'échelle journalière, l'écart de 0 à 12 tonnes par jour, soit les valeurs minimale et maximale rapportées par Glencore, est associé à un doublement des concentrations en arsenic mesurées à la station ALTSP1 (Figure 3). Toutefois, cette estimation journalière pourrait sous-estimer l'effet réel, en raison d'un décalage de périodicité entre les mesures de concentration en arsenic à ALTSP1 (cycle 12 h – 12 h) et celles des procédés pyrométallurgiques (cycle 6 h – 6 h) (Prévost et Viger, 2022). Dans l'ensemble, ces résultats suggèrent que la quantité d'arsenic introduite dans le procédé contribue aux concentrations en arsenic dans l'air ambiant.

Comme attendu, la direction des vents exerce un effet important sur les concentrations en arsenic mesurées à la station ALTSP1. Cette influence se manifeste particulièrement pour les données journalières, en raison de leur sensibilité accrue aux variations météorologiques quotidiennes comparativement aux moyennes mensuelles. Puisque, selon Glencore, l'arsenic émis dans l'air provient majoritairement du secteur de la pyrométallurgie (Prévost et Viger, 2022), les concentrations en arsenic sont élevées lorsque les vents soufflent selon l'axe reliant ce secteur à la station ALTSP1. Cela indique qu'il est important de considérer la direction des vents dans l'interprétation des concentrations en arsenic à cette station. Cependant, un effet considérable de la

direction des vents sur les concentrations en arsenic mesurées dans l'air n'amoindrit pas l'effet ou la significativité des autres variables comme la quantité d'arsenic ajoutée au réacteur.

Références

Bolker, Benjamin M. 2008. « Ecological Models and Data in R ». Princeton University Press. https://www.degruyterbrill.com/document/doi/10.1515/9781400840908/html.

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). « glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling ». The R Journal, 9(2), 378–400. https://doi.org/10.32614/RJ-2017-066.

Fonderie Horne de Glencore. 2022. « Relation entre les intrants et les extrants de la Fonderie Horne de Glencore » [Jeu de données]. Dépôt de données ONICSE. https://donnees.onicse.uqat.ca/dataset/relation-intrants-extrants-fonderie-horne

Fox, John, Sanford Weisberg, et Brad Price. 2024. « Car: Companion to Applied Regression ». septembre 27. https://CRAN.R-project.org/package=car.

Hartig, Florian. 2025. « DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models ». janvier 20, 0.4.7. https://doi.org/10.32614/CRAN.package.DHARMa.

Hothorn, Torsten, Achim Zeileis, Richard W. Farebrother, et Clint Cummins. 2022. « Lmtest: Testing Linear Regression Models ». octobre 13. https://CRAN.R-project.org/package=lmtest.

Kutner, Michael H., éd. 2005. « Applied Linear Statistical Models ». 5th ed. The McGraw-Hill/Irwin Series Operations and Decision Sciences. McGraw-Hill Irwin.

Lüdecke, Daniel, Mattan Ben-Shachar, Indrajeet Patil, Philip Waggoner, et Dominique Makowski. 2021. « Performance: An R Package for Assessment, Comparison and Testing of Statistical Models ». Journal of Open Source Software 6 (60): 3139. https://doi.org/10.21105/joss.03139.

Marchant, Carolina, Víctor Leiva, M. Fernanda Cavieres, et Antonio Sanhueza. 2013. « Air Contaminant Statistical Distributions with Application to PM10 in Santiago, Chile ». In Reviews of Environmental Contamination and Toxicology Volume 223, édité par David M. Whitacre. Springer. https://doi.org/10.1007/978-1-4614-5577-6_1.

Ministère de l'Environnement, de la Lutte contre les changements climatiques, de la Faune et des Parcs (MELCCFP). 2023. « Commentaires de l'établissement Fonderie Horne portant sur le préavis incluant les réponses du MELCCFP ». Gouvernement du Québec. https://www.environnement.gouv.qc.ca/air/fonderie-horne/commentaires-fonderie-horne-preavis-incluant-reponses-melccfp.pdf

O'Brien, Robert M. 2007. « A Caution Regarding Rules of Thumb for Variance Inflation Factors ». Quality & Quantity 41 (5): 673-90. https://doi.org/10.1007/s11135-006-9018-6.

Prévost, Yves, et Marie-Élise Viger. 2022. « Étude sur la concentration en arsenic des concentrés alimentés au réacteur et la concentration en arsenic mesurée dans l'air ambiant à la station légale ALTSP1 ». https://www.bibliotheque.assnat.qc.ca/DepotNumerique_v2/AffichageFichier.aspx?idf=308473

R Core Team. 2025. « R: A language and environment for statistical computing ». R Foundation for Statistical Computing, Vienna, Austria Édition. https://www.R-project.org/.

Rivest, S. 2026. « Script R pour la réanalyse de la relation entre la concentration en arsenic des concentrés alimentés au réacteur de la Fonderie Horne et la concentration en arsenic à la station ALTSP1 » [Script R]. Observatoire national sur les incidences des émissions de contaminants sur la santé et l'environnement, Université du Québec en Abitibi-Témiscamingue. https://donnees.onicse.uqat.ca/dataset/script-r-reanalyse-relation-concentration-arsenic-concentres-station-altsp1

Annexe 1 : accessibilité du script est des données

Les jeux de données et le script utilisés pour produire les analyses statistiques sont accessibles dans le dépôt de données de l'ONICSE (Fonderie Horne de Glencore, 2022 ; Rivest, 2026).

Pour l'analyse des données mensuelles, deux modèles linéaires ont été développés en utilisant la fonction « lm » de la librairie « stats » du logiciel R (R Core Team 2025). Les hypothèses des modèles ont été validées par le test d'hétéroscédasticité de Breusch-Pagan en utilisant la fonction « bptest » de la librairie « lmtest » du logiciel R (Hothorn et al. 2022), le test d'indépendance des erreurs de Durbin-Watson avec la fonction « dwtest » de la librairie « lmtest » et le test de normalité de Shapiro-Wilk avec la fonction « shapiro.test » de la librairie « stats ». Finalement, la colinéarité entre variables a été vérifiée en mesurant le facteur d'inflation de la variance (VIF) à l'aide de la fonction « VIF » de la librairie R « car » (Fox et al. 2024).

Pour l'analyse des données journalières, les modèles ont été développés en utilisant la fonction « glmmTMB » de la librairie « glmmTMB » (Brooks et al., 2017). La multicolinéarité entre variables explicatives a été examinée à l'aide des facteurs d'inflation de la variance (VIF) calculés avec la fonction « check_collinearity » de la librairie « performance » (Lüdecke et al. 2021). L'ajustement du modèle et le respect des hypothèses ont été évalués à l'aide de la fonction « simulateResiduals », « testUniformity », « testDispersion » et « testOutliers » de la librairie « DHARMa » (Hartig, 2025) ainsi que la fonction « check_heteroscedasticity » de la librairie « performance ».