



Ressources naturelles
Canada

Natural Resources
Canada

**COMMISSION GÉOLOGIQUE DU CANADA
DOSSIER PUBLIC 9270f**

**Données géospatiales à composantes multiples pour la
cartographie prédictive des minéraux critiques
canadiens**

M. Parsa, C.J.M. Lawley, E. Schetselaar, T. Martins et A. Thompson

2026

COMMISSION GÉOLOGIQUE DU CANADA
DOSSIER PUBLIC 9270f

Données géospatiales à composantes multiples pour la cartographie prédictive des minéraux critiques canadiens

M. Parsa¹, C.J.M. Lawley¹, E. Schetselaar¹, T. Martins² et A. Thompson¹

¹Ressources naturelles Canada, Commission géologique du Canada, 601, rue Booth, Ottawa, (Ontario) K1A 0E8

¹Levés géologiques du Manitoba, 360-1395 Ellice Avenue, Winnipeg, (Manitoba) R3G 3P2

2026

© Sa Majesté le Roi du chef du Canada, représenté par le ministre des Ressources naturelles, 2026

Le contenu de cette publication ou de ce produit peut être reproduit en tout ou en partie, et par quelque moyen que ce soit, sous réserve que la reproduction soit effectuée uniquement à des fins personnelles ou publiques mais non commerciales, sans frais ni autre permission, à moins d'avis contraire.

On demande seulement :

- de faire preuve de diligence raisonnable en assurant l'exactitude du matériel reproduit;
- d'indiquer le titre complet du matériel reproduit et le nom de l'organisation qui en est l'auteur;
- d'indiquer que la reproduction est une copie d'un document officiel publié par Ressources naturelles Canada (RNCa) et que la reproduction n'a pas été faite en association avec RNCa ni avec l'appui de celui-ci.

La reproduction et la distribution à des fins commerciales sont interdites, sauf avec la permission écrite de RNCa. Pour de plus amples renseignements, veuillez communiquer avec RNCa à copyright-droitdauteur@rncan-rncan.gc.ca.

On peut télécharger cette publication gratuitement à partir du Dépôt ouvert des sciences et technologie de RNCa (<https://ostrncan-dostrncan.canada.ca/>).

Notation bibliographique conseillée

Parsa, M., Lawley, C.J.M., Schetselaar, E., Martins, T. et Thompson, A., 2026. Données géospatiales à composantes multiples pour la cartographie prédictive des minéraux critiques canadiens; Commission géologique du Canada, Dossier public 9270, 1 fichier .zip. <https://doi.org/10.4095/pnducwy548>

Les publications de cette série n'ont pas été révisées; elles sont publiées telles que soumises par l'auteur.

ISSN 2816-7163
ISBN 978-0-660-97940-3
N° de catalogue M183-2/9270F-PDF
<https://doi.org/10.4095/pnducwy548>

INTRODUCTION

Les cartes de la prospectivité minérale à l'échelle nationale facilitent la gestion des terres et la prise de décision. La disponibilité d'outils d'intelligence artificielle en libre accès avec leurs algorithmes de classification avancés a transformé la cartographie de la prospectivité minérale (CPM) basée sur les données. Cependant, la CPM nécessite des données d'entrée pertinentes pour exploiter au mieux ces algorithmes. Le présent rapport vise à combler la lacune existante en publiant des données géospatiales à composantes multiples qui facilitent la cartographie de la prospectivité minérale des systèmes plurimétallifères à l'échelle pancanadienne.

Les données d'entrée pour la CPM sont doubles : (i) des caractéristiques géospatiales, appelées couches de données probantes, qui servent de covariables au modèle, et (ii) des étiquettes d'apprentissage binaires. Les couches de données probantes sont constituées de multiples ensembles de données géoscientifiques et de leurs dérivés, qui décrivent collectivement les composantes du système minéral soit directement, soit par l'intermédiaire de valeurs représentatives ([McCuaig et al., 2010](#)). Les données géophysiques, géochimiques et géologiques sont des sources courantes pour la génération des couches de données probantes pour la CPM ([Parsa et al., 2024; 2025](#)). En plus de ces entrées de données conventionnelles pour la CPM, ce rapport fournit également les résultats de l'application de techniques de traitement automatique du langage naturel ([Lawley et al., 2023](#)) aux données textuelles des polygones du substratum géologique, ce qui permet de les utiliser en tant que couches de données probantes.

Les étiquettes d'apprentissage binaires représentent des caractéristiques géospatiales à l'emplacement des zones minéralisées du type de gisement minéral ciblé, appelées étiquettes positives, et à des emplacements non associés à des zones minéralisées, appelées étiquettes négatives. La sélection des étiquettes positives est simple, mais il existe plusieurs approches pour

sélectionner les étiquettes négatives (par exemple, Carranza, 2008; Nykänen et al., 2015; Parsa et al., 2024). Ce rapport n'introduit toutefois pas d'étiquettes d'apprentissage négatives, laissant ce choix aux praticiens de la CPM. Au lieu de cela, il introduit des étiquettes de classification positives possédant différents niveaux de certitude : les sites minéralisés dont l'importance géométallurgique a été prouvée sont désignés comme des gisements minéraux, y compris les producteurs actuels et passés, tandis que les anomalies géochimiques et les zones minéralisées mineures sont présentées comme des occurrences minérales. Cette classification se justifie par le fait que l'utilisation de la première catégorie dans la CPM permet d'obtenir des modèles prédictifs fiables sur le plan géométallurgique, ce qui facilite la prise de décision. Toutefois, la représentativité de la deuxième catégorie devrait être évaluée avant d'être utilisée en CPM, ce qui n'entre pas dans le champ d'application du présent rapport.

Plusieurs systèmes minéraux, dont la minéralisation Zn-Pb de type Mississippi Valley (MVT) (Leach et al, 2010), la minéralisation magmatique en sulfures Ni-Cu-PGE (Barnes et al., 2016), la minéralisation en éléments de terres rares (ETR) +/- Nb encaissée dans des carbonatites (Verplanck et al, 2016), et les pegmatites Li-Cs-Ta (LCT) (Bradley et al., 2017), ont été ciblés dans cette publication de données, dans le but de fournir aux praticiens la flexibilité nécessaire pour évaluer la prospectivité de différents systèmes minéraux. Ils sont présentés dans le [tableau 1](#), aux côtés du produit critique auquel ils sont associés. Ces systèmes minéraux font partie des principales sources géologiques des minéraux critiques canadiens les plus récemment jugés prioritaires (à la date de rédaction de ce manuscrit).

Tableau 1 - Produits critiques et leur source géologique primaire (système minéral) qui figurent dans le cube de données de ce rapport.

Produit critique	Source géologique primaire
zinc	gisements de zinc-plomb de type Mississippi-Valley (MVT)
nickel	gisements de sulfures magmatiques de Ni-Cu-EGP

éléments de terres rares (ETR) et niobium	gisements d'ETR +/- de niobium (Nb) encaissés dans des carbonatites
niobium	gisements d'ETR +/- Nb encaissés dans des carbonatites
lithium	Pegmatites à Li-Cs-Ta

DONNÉES

Le premier cube de données est fourni dans un format de valeurs séparées par des virgules. Dans ce cube de données, les couches de données probantes (c'est-à-dire les covariables) sont présentées sous forme de colonnes, chacune des 1 820 346 lignes représentant un hexagone H3 de niveau 7, couvrant une zone d'environ 5 km². Il existe une colonne « H3id », contenant un code par entrée représentant l'index géospatial des cellules H3 dans le système d'indexation hiérarchique global H3 (<https://github.com/uber/h3>). Il y a également des colonnes de longitude et de latitude décrivant l'emplacement du point médian de chaque cellule H3 en utilisant le système de projection géodésique mondial de 1984, qui porte le code 4326 de l'European Petroleum Survey Group (EPSG). Les couches de données probantes présentées dans le premier cube de données de ce rapport sont décrites dans le [tableau 2](#). Les caractéristiques 1 à 20 sont des dérivés des données aéromagnétiques et magnétiques et ont été décrites dans des publications antérieures ([Lawley et al., 2021, 2022; Parsa et al., 2024, 2025](#)). Toutefois, les caractéristiques 21-31 sont le résultat de l'application de techniques de traitement automatique du langage naturel aux polygones géologiques, dérivées en utilisant la procédure suivante. Toutes les caractéristiques ont été ramenées à une échelle de [0, 1], ce qui facilite leur analyse à l'aide de méthodes d'apprentissage automatique.

Tableau 2 - Description des couches de données probantes apparaissant dans le cube de données de ce rapport

Numéro de la caractéristique	Description	Composante du système minéral
1	1ère dérivée verticale des données gravimétriques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
2	Gravité de Bouguer	Cartographie les environnements géologiques propices au piégeage des fluides hydrothermaux
3	Amplitude du gradient horizontal des données gravimétriques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
4	A une proximité de 5 km des courbes dérivées des données gravimétriques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.

5	A une proximité de 10 km des courbes dérivées des données gravimétriques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
6	A une proximité de 15 km des courbes dérivées des données gravimétriques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
7	A une proximité de 20 km des courbes dérivées des données gravimétriques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
8	A une proximité de 25 km des courbes dérivées des données gravimétriques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
9	A une proximité de 30 km des courbes dérivées des données gravimétriques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
10	1ère dérivée verticale des données magnétiques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
11	Amplitude du gradient horizontal des données aéromagnétiques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
12	Données aéromagnétiques réduites au pôle	Cartographie les environnements géologiques propices au piégeage des fluides hydrothermaux
13	Dérivée inclinée des données magnétiques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
14	âge minimum (géochronologie)	Indique les moteurs du système minéral en fonction de la chronologie des événements géologiques.
15	âge maximum (géochronologie)	Indique les moteurs du système minéral en fonction de la chronologie des événements géologiques.
16	À une proximité de 5 km des courbes dérivées des données magnétiques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
17	À une proximité de 10 km des courbes dérivées des données magnétiques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
18	À une proximité de 15 km des courbes dérivées des données magnétiques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
19	À une proximité de 20 km des courbes dérivées des données magnétiques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
20	À une proximité de 25 km des courbes dérivées des données magnétiques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
21	À une proximité de 30 km des courbes dérivées des données magnétiques	Cartographie les limites géologiques et fournit de l'information sur les voies de passage potentielles des fluides hydrothermaux.
22	NLP_lithology_PC1	Fournit de l'information sur les déclencheurs, la source et les mécanismes de piégeage de la minéralisation
23	NLP_lithology_PC2	Fournit de l'information sur les déclencheurs, la source et les mécanismes de piégeage de la minéralisation
24	NLP_lithology_PC3	Fournit de l'information sur les déclencheurs, la source et les mécanismes de piégeage de la minéralisation
25	NLP_lithology_PC4	Fournit de l'information sur les déclencheurs, la source et les mécanismes de piégeage de la minéralisation
26	NLP_lithology_PC5	Fournit de l'information sur les déclencheurs, la source et les mécanismes de piégeage de la minéralisation
27	NLP_lithology_PC6	Fournit de l'information sur les déclencheurs, la source et les mécanismes de piégeage de la minéralisation
28	NLP_lithology_PC7	Fournit de l'information sur les déclencheurs, la source et les mécanismes de piégeage de la minéralisation
29	NLP_lithology_PC8	Fournit de l'information sur les déclencheurs, la source et les mécanismes de piégeage de la minéralisation
30	NLP_lithology_PC9	Fournit de l'information sur les déclencheurs, la source et les mécanismes de piégeage de la minéralisation
31	NLP_lithology_PC10	Fournit de l'information sur les déclencheurs, la source et les mécanismes de piégeage de la minéralisation

Les données textuelles des polygones géologiques du substratum, qui varient entre des phrases courtes et longues, ont été segmentées en unités lexicales. On a ensuite dérivé un vecteur numérique à 300 dimensions par unité en appliquant le vecteur de référence global de géosciences au modèle de langage Word (Lawley et al., 2023). On a ensuite, pour chaque polygone géologique, calculé la moyenne des plongements lexicaux, pour lesquels on a ensuite réalisé une analyse en composantes principales (Shlens, 2014). Les caractéristiques 21 à 31 de l'ensemble de données 1 sont les dix

premières composantes principales qui expliquent 69 % de la variabilité totale des plongements lexicaux moyens des polygones géologiques.

Chaque système minéral a été décrit dans le cube de données par deux variables dichotomiques, décrivant la présence et l'absence des gisements et occurrences dans des cellules individuelles (Tableau 3). La distinction entre la présence et l'absence est liée à leur viabilité économique.

Tableau 3 - Produits critiques et leur source géologique primaire (système minéral) qui figurent dans le cube de données de ce rapport et séparément en tant que fichiers GeoJSON.

Système minéral	colonnes décrivant les gisements	colonnes décrivant les occurrences
gisements de zinc-plomb de type Mississippi-Valley (MVT)	MVT_a	MVT_b
gisements de sulfures magmatiques de Ni-Cu-EGP (MNCP)	MNCP-a	MNCP-b
gisements d'ETR +/- de niobium (Nb) dans des carbonatites (CHRN)	CHRN_a	CHRN_b
Pegmatites à Li-Cs-Ta	LCT_a	LCT-b

Le cube de données ci-dessus peut être utilisé collectivement pour la CPM pancanadienne à l'aide de classifieurs basés sur le spectre, tels que les réseaux de neurones à convolution unidimensionnels (Kiranyaz et al., 2020). Ces méthodes analysent les données pixel par pixel, en négligeant les relations potentielles entre les gisements minéraux et les pixels voisins (Zuo et Carranza, 2023). Ces hypothèses peuvent toutefois ne pas intéresser les praticiens de la CPM, car les gisements minéraux sont intrinsèquement associés au milieu géologique qui les entoure. Les classifieurs basés sur le spectre spatial, tels que les réseaux de neurones à convolution conventionnels (Gu et al., 2018), cependant, analysent des blocs de pixels, respectant ainsi la relation potentielle entre la minéralisation et le milieu géologique environnant (Zuo et Carranza, 2023).

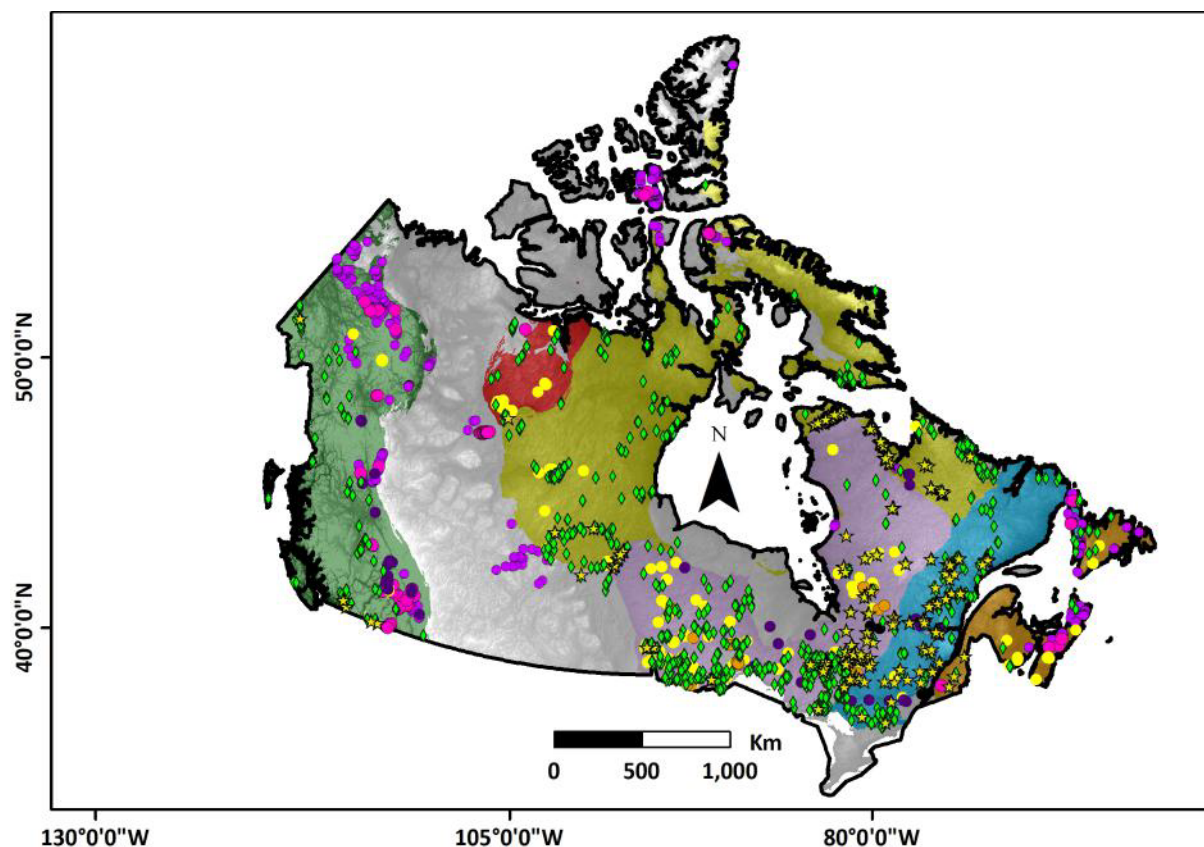
Les auteurs ont donc décidé de présenter les couches de données probantes sous la forme d'une trame multicanal au format TIFF, ce qui permet d'utiliser des classifieurs basés sur le spectre spatial.

L'image matricielle multicanal contient 31 bandes, chacune représentant une des caractéristiques présentées dans le [tableau 2](#). L'ordre des caractéristiques du [tableau 2](#) est préservé dans la trame. Ce fichier matriciel a une résolution de 2000 m × 2000 m.

Pour l'utilisation de classifieurs basés sur le spectre spatial, des étiquettes positives sont créées en recadrant la matrice multicanal centrée sur l'emplacement des zones minéralisées ([Tableau 3](#)). Pour faciliter le processus de recadrage, huit fichiers vectoriels (des fichiers Esri Shape), représentant des emplacements potentiels pour différents systèmes minéraux avec différents niveaux de certitude ([Tableau 3](#)), ont également été présentés dans ce rapport. Il s'agit de vecteurs qui correspondent au système de géoréférencement par coordonnées de la matrice multicanal, qui est le NAD 1983/Atlas du Canada Lambert. Le code EPSG pour ce système de projection est le 3978. [La figure 1](#) montre ces huit vecteurs superposés sur la carte des provinces géologiques canadiennes ([Wheeler et al., 1996](#)).

REMERCIEMENTS

Les auteurs souhaitent remercier les Levés géologiques du Manitoba pour avoir fourni les données nécessaires à cette recherche. Cette recherche a été appuyée par le Programme de données géoscientifiques sur les minéraux critiques de la Commission géologique du Canada. Les auteurs souhaitent remercier Jean-Sébastien Bouffard pour ses commentaires sur une version antérieure de ces travaux.



- | | |
|--|--|
| <ul style="list-style-type: none"> ● MVT deposits ● MVT occurrences ★ magmatic Ni deposits ◆ magmatic Ni occurrences ● carbonatite-hosted REE+/-Nb deposits ● carbonatite-hosted REE+/-Nb occurrences ● LCT pegmatite deposits ● LCT pegmatite occurrences | <p>Geological provinces</p> <ul style="list-style-type: none"> ■ Appalachian ■ Cordilleran ■ Grenville ■ Churchill ■ Superior ■ Slave |
|--|--|

Figure 1. Carte des provinces géologiques du Canada (Wheeler et al., 1996) montrant la distribution spatiale des divers groupes d'étiquettes de classification utilisés dans cette étude.

RÉFÉRENCES

- Barnes, S.J., Cruden, A.R., Arndt, N., Saumur, B.M., 2016. The mineral system approach applied to magmatic Ni– Cu–PGE sulphide deposits. *Ore Geol. Rev.* 76, 296–316.
- Bradley, D.C., McCauley, A.D., Stillings, L.L., 2017. Mineral-deposit model for lithium-cesium-tantalum pegmatites. Carranza, E.J.M., 2008. Geochemical anomaly and mineral prospectivity mapping in GIS. Elsevier.
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., 2018. Recent advances in convolutional neural networks. *Pattern Recognit.* 77, 354–377.
- Kiranyaz, S., Zabihi, M., Rad, A.B., Ince, T., Hamila, R., Gabbouj, M., 2020. Real-time phonocardiogram anomaly detection by adaptive 1D convolutional neural networks. *Neurocomputing* 411, 291–301.
- Lawley, C.J.M., Gadd, M.G., Parsa, M., Lederer, G.W., Graham, G.E., Ford, A., 2023. Applications of Natural Language Processing to Geoscience Text Data and Prospectivity Modeling. *Nat. Resour. Res.* 32, 1503–1527.
- Lawley, C.J.M., McCafferty, A.E., Graham, G.E., Huston, D.L., Kelley, K.D., Czarnota, K., Paradis, S., Peter, J.M., Hayward, N., Barlow, M., Emsbo, P., Cohan, J., San Juan, C.A., Gadd, M.G., 2022. Data-driven prospectivity modelling of sediment-hosted Zn–Pb mineral systems and their critical raw materials. *Ore Geol. Rev.* 141, 104635.
- Lawley, C.J.M., Tschirhart, V., Smith, J.W., Pehrsson, S.J., Schetselaar, E.M., Schaeffer, A.J., Houlié, M.G., Eglington, B.M., 2021. Prospectivity modelling of Canadian magmatic Ni ($\pm$ Cu $\pm$ Co $\pm$ PGE) sulphide mineral systems. *Ore Geol. Rev.* 132, 103985.
- Leach, D.L., Taylor, R.D., Fey, D.L., Diehl, S.F., Saltus, R.W., 2010. A deposit model for Mississippi Valley-type lead-zinc ores (No. 2328–0328). US Geological Survey.
- McCuaig, T.C., Beresford, S., Hronsky, J., 2010. Translating the mineral systems approach into an effective exploration targeting system. *Ore Geol. Rev.* 38, 128–138.
- Nykänen, V., Lahti, I., Niiranen, T., Korhonen, K., 2015. Receiver operating characteristics (ROC) as validation tool for prospectivity models—A magmatic Ni–Cu case study from the Central Lapland Greenstone Belt, Northern Finland. *Ore Geol. Rev.* 71, 853–860.
- Parsa, M., Lawley, C.J., Cumani, R., Schetselaar, E., Harris, J., Lentz, D.R., Zhang, S.E., Bourdeau, J.E., 2024. Predictive Modeling of Canadian Carbonatite-Hosted REE+/- Nb Deposits. *Nat. Resour. Res.* 1–25.
- Parsa, M., Lawley, C.J.M., Cawood, T., Martins, T., Cumani, R., Zhang, S.E., Thompson, A., Beyer, S., Lentz, D.R., Harris, J., Fam, H.J.A., Voinot, A., 2025. Pan-Canadian Predictive Modeling of Lithium–Cesium–Tantalum Pegmatites with Deep Learning and Natural Language Processing. *Nat. Resour. Res.* <https://doi.org/10.1007/s11053-024-10438-x>.
- Shlens, J., 2014. A tutorial on principal component analysis. *ArXiv Prepr. ArXiv14041100*.
- Verplanck, P.L., Mariano, A.N., Mariano Jr., A., 2016. Rare earth element ore geology of carbonatites, in: *Rare Earth and Critical Elements in Ore Deposits*. Society of Economic Geologists, Inc., Littleton, CO, pp. 5–32.
- Wheeler, J., Hoffman, P., Card, K., Davidson, A., Sanford, B., Okulitch, A., & Roest, W. (1996). Geological map of Canada/Carte géologique du Canada, Geological Survey of Canada, “A” Series Map 1860A, 1996, 3 sheets; 1 CD-ROM.
- Zuo, R., Carranza, E.J.M., 2023. Machine Learning-Based Mapping for Mineral Exploration. *Math. Geosci.* 55, 891–895.