Natural Resources Canada | Ressources naturelles Canada

GEOLOGICAL SURVEY OF CANADA
OPEN FILE 9270e

Multicomponent geospatial data for predictive mapping of Canadian critical minerals

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

2026

Canada

GEOLOGICAL SURVEY OF CANADA
OPEN FILE 9270e

Multicomponent geospatial data for predictive mapping of Canadian critical minerals

M. Parsa[1], C.J.M. Lawley[1], E. Schetselaar[1], T. Martins[2], and A. Thompson[1]

[1]Natural Resources Canada, Geological Survey of Canada, 601 Booth Street, Ottawa, Ontario K1A 0E8
[1]Manitoba Geological Survey, 360-1395 Ellice Avenue, Winnipeg, Manitoba R3G 3P2

2026

Recommended citation
Parsa, M., Lawley, C.J.M., Schetselaar, E., Martins, T., and Thompson, A., 2026. Multicomponent geospatial data for predictive mapping of Canadian critical minerals; Geological Survey of Canada, Open File 9270e, 1 .zip file. https://doi.org/10.4095/p80q4bhma6

Publications in this series have not been edited; they are released as submitted by the author.

ISSN 2816-7155
ISBN 978-0-660-77032-1
Catalogue No. M183-2/9270E-PDF
https://doi.org/10.4095/p80q4bhma6

INTRODUCTION

National-scale mineral prospectivity maps help support land management and decision-making. The availability of open-source artificial intelligence tools with their advanced classification algorithms has transformed data-driven Mineral Prospectivity Mapping (MPM). However, MPM requires suitable input data to optimally leverage such classification algorithms. This report aims to bridge this data gap by releasing multicomponent geospatial data that facilitates pan-Canadian, multi-mineral systems MPM.

The input data for MPM is twofold: (i) geospatial features, known as evidence layers, that serve as model covariates, and (ii) binary training labels. Evidence layers comprise multiple geoscientific datasets and their derivatives, which collectively describe mineral system components either directly or through proxies (McCuaig et al., 2010). Common sources for generating MPM evidence layers include geophysical, geochemical, and geological data (Parsa et al., 2024; 2025). In addition to these conventional data inputs for MPM, this report also provides the results of applying natural language processing techniques (Lawley et al., 2023) to geological bedrock polygon text data, allowing them to be used as evidence layers.

Binary training labels represent geospatial features at locations of mineralized zones of the targeted mineral deposit type, known as positive labels, and at locations not associated with mineralized zones, known as negative labels. While selecting positive labels is straightforward, there are various approaches to selecting negative labels (e.g., Carranza, 2008; Nykänen et al., 2015; Parsa et al., 2024). This report, however, does not introduce negative training labels, leaving that choice to MPM practitioners. Instead, this report introduces positive class labels with varying levels of certainty: mineralized locations with demonstrated geometallurgical significance are designated as mineral deposits, including current and past producers, whereas geochemical anomalies and minor mineralized zones are introduced as mineral occurrences. The rationale for this classification is that using the former category in MPM leads to predictive models with geometallurgical reliability, which supports further decision-making. However, the representativeness of the latter category should be assessed before its use in MPM, which is out of the scope of this report.

Multiple mineral systems, including Mississippi Valley-Type (MVT) Zn-Pb mineralization (Leach et al., 2010), magmatic Ni-Cu-PGE sulphide mineralization (Barnes et al., 2016), carbonatite-hosted rare-earth elements (REE) +/- Nb mineralization (Verplanck et al., 2016), and Li-Cs-Ta (LCT) pegmatites (Bradley et al., 2017), have been targeted in this data release, with the aim of providing practitioners with the flexibility to assess the prospectivity of different mineral systems. These mineral systems are listed in Table 1, along with the critical commodity to which they are associated. These mineral systems are among the primary geological sources of the latest (as per the date of writing this manuscript) Canadian priority critical minerals.

Table 1 – Critical commodities and their primary geological source (mineral system) that appear in this report's datacube.

critical commodity	Primary geological source
zinc	MVT zinc-lead deposits
nickel	magmatic Ni-Cu-PGE sulphide deposits
rare earth elements (REEs) and niobium	carbonatite-hosted REE +/- niobium (Nb) deposits
niobium	carbonatite-hosted REE +/- Nb deposits
lithium	LCT pegmatites

DATA

The first datacube is provided in a comma-separated value (CSV) format. In this datacube, evidence layers (i.e. covariates) are presented as columns, wherein each of the 1,820,346 rows represents an H3 level 7 hexagon, covering an area of approximately 5 km^2. There is an "H3id" column, containing a code per entry representing the geospatial index of H3 cells within the H3 global hierarchical indexing system (https://github.com/uber/h3). There are longitude and latitude columns as well that describe the location of each H3 cell's midpoint using World's Geodetic projecting System 1984, which has the European Petroleum Survey Group (EPSG) code of 4326. Evidence layers presented in this report's first datacube are described in Table 2. Features 1 to 20 are derivatives of aeromagnetic and magnetic data and have been described in previous publications (Lawley et al., 2021, 2022; Parsa et al., 2024, 2025). Features 21-31, are, however the results of applying natural language processing techniques to geology polygons, derived using the following procedure. All features have been rescaled to a [0, 1] range, which further facilitates their analysis using machine learning methods.

Table 2 – Description of evidence layers appeared in this report's datacube.

Feature Number	**Description**	**Mineral system component**
1	1st vertical derivative of gravity data	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
2	Bouguer gravity	Maps geological environments suitable for trapping hydrothermal fluids
3	Horizontal gradient magnitude of gravity data	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
4	5km proximity to gravity worms	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
5	10km proximity to gravity worms	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
6	15km proximity to gravity worms	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
7	20km proximity to gravity worms	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
8	25km proximity to gravity worms	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
9	30km proximity to gravity worms	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
10	1st vertical derivative of magnetic data	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
11	Horizontal gradient magnitude of aeromagnetic data	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
12	Reduced to pole aeromagnetic data	Maps geological environments suitable for trapping hydrothermal fluids
13	Tilt derivative of magnetic data	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
14	minimum age (geochronology)	Indicates mineral system drivers based on the timing of geological events
15	maximum age (geochronology)	Indicates mineral system drivers based on the timing of geological events
16	5km proximity to magnetic worms	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
17	10km proximity to magnetic worms	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
18	15km proximity to magnetic worms	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
19	20km proximity to magnetic worms	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
20	25km proximity to magnetic worms	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
21	30km proximity to magnetic worms	Maps geological boundaries and provides insights into potential pathways for hydrothermal fluids
22	NLP_lithology_PC1	Provides insights into mineralization drivers, source, and trapping mechanism
23	NLP_lithology_PC2	Provides insights into mineralization drivers, source, and trapping mechanism
24	NLP_lithology_PC3	Provides insights into mineralization drivers, source, and trapping mechanism
25	NLP_lithology_PC4	Provides insights into mineralization drivers, source, and trapping mechanism
26	NLP_lithology_PC5	Provides insights into mineralization drivers, source, and trapping mechanism
27	NLP_lithology_PC6	Provides insights into mineralization drivers, source, and trapping mechanism
28	NLP_lithology_PC7	Provides insights into mineralization drivers, source, and trapping mechanism
29	NLP_lithology_PC8	Provides insights into mineralization drivers, source, and trapping mechanism
30	NLP_lithology_PC9	Provides insights into mineralization drivers, source, and trapping mechanism
31	NLP_lithology_PC10	Provides insights into mineralization drivers, source, and trapping mechanism

The text data of geology bedrock polygons, which varies between short and long phrases, were tokenized. This was followed by deriving a 300-dimensional numeric vector per token by applying the Geoscience Global Vector to Word language model (Lawley et al., 2023). Word embeddings were then average for each geology polygon, which were then subjected to principal component analysis (Shlens, 2014). Features 21 to 31 in Dataset 1 are the first ten principal components that explain 69% of the total variability of average word embeddings of geology polygons.

Each mineral system has been described in the datacube by two dichotomous variables, describing the presence and absence of their deposits and occurrences in individual cells (Table 3). The distinction between the former and the latter is linked to their economic viability.

Table 3 – Critical commodities and their primary geological source (mineral system) that appear in this report's datacube and separately as Geopackage (GPKG) files.

Mineral system	columns describing deposits	columns describing occurrences
MVT zinc-lead deposits	MVT_a	MVT_b
magmatic Ni-Cu-PGE sulphide deposits	MNCP-a	MNCP-b
carbonatite-hosted REE +/- niobium (Nb) deposits	CHRN_a	CHRN_b
LCT pegmatites	LCT_a	LCT-b

The above datacube can collectively be used for pan-Canadian MPM using spectrum-based classifiers, such as one-dimensional convolutional neural networks (Kiranyaz et al., 2020). These methods analyze the data on a pixel-by-pixel basis, overlooking potential relationships between mineral deposits and neighboring pixels (Zuo and Carranza, 2023). Such assumptions, however, may not appeal to MPM practitioners, as mineral deposits are inherently associated with their surrounding geological environment. Spatial-spectrum-based classifiers, such as conventional convolutional neural networks (Gu et al., 2018), however, analyze patches of pixels, thereby respecting the potential relationship between mineralization and its surrounding geological environment (Zuo and Carranza, 2023).

The authors, therefore, were inclined to further present the evidence layers as a multi-channel raster adhering to a TIFF format, which allows the use of spatial-spectrum-based classifiers. The multi-channel raster image contains 31 bands, each representing one of the features listed in Table 2, with the order of features from Table 2 preserved in the raster. This raster file comes with a resolution of 2000m × 2000m.

For the use of spatial-spectrum-based classifiers, positive labels are created from cropping the multi-channel raster centered at the location of mineralized zones (Table 3). To facilitate the cropping process, eight vector-based files (Geopackage files, representing locations for different mineral systems with varying levels of certainty (Table 3), have been also presented in this report. These are vectors that match the georeferencing coordinate system of the multi-channel raster, which is NAD 1983 Canada Atlas Lambert. The EPSG code for this projection system is 3978. Figure 1 shows these eight vectors overlaid on the map of Canadian geological provinces (Wheeler et al., 1996).

ACKNOWLEDGMENT

The authors wish to thank Manitoba Geological Survey for providing data to this research. This research was supported by the Critical Minerals Geoscience and Data program of Geological Survey of Canada. The authors wish to thank Jean-Sebastien Bouffard for his comments on an earlier version of this work.

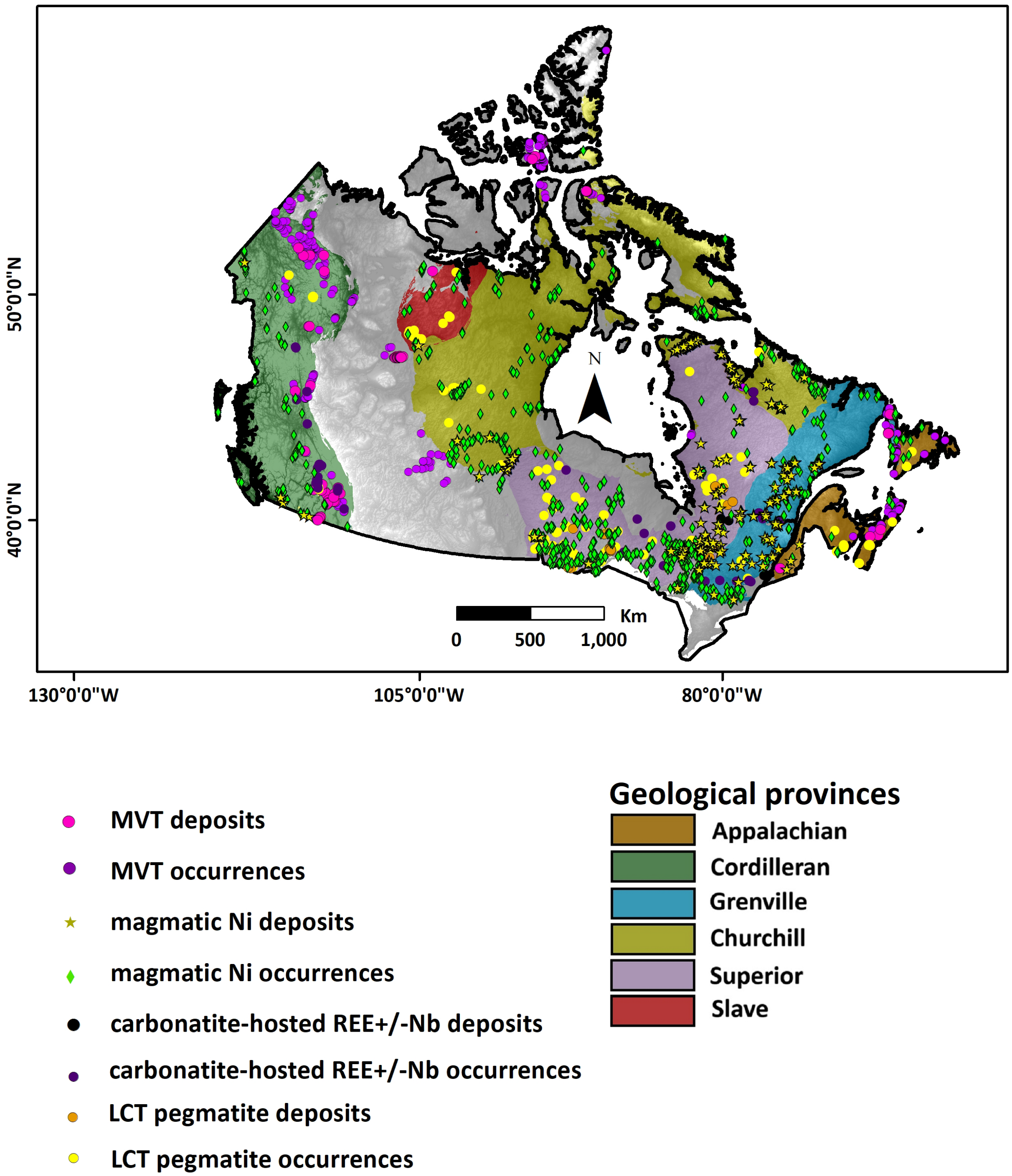

Figure 1. Map of Canada's geological provinces (Wheeler et al., 1996) showing the spatial distribution of various class label groups used in this study.

REFERENCES

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., Coyan, 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., Houlé, M.G., Eglington, B.M., 2021. Prospectivity modelling of Canadian magmatic Ni (±Cu ± Co ± 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 g´eologique 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.