Factors contributing to metal endowment in the western Wabigoon and southern Abitibi

subprovinces: A machine learning approach to Precambrian greenstone belts

by

Rebecca M Montsion GIT, MSc, BScH

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

(PhD) in Precambrian and Economic Geology

To

The Office of Graduate Studies

Laurentian University

Sudbury, Ontario, Canada

and

The School of Earth Sciences

The University of Western Australia

Perth, Western Australia, Australia

© Rebecca Montsion, 2023

Abstract

Mineral exploration workflows are including more quantitative techniques and probabilistic targeting to capture subtle or convoluted relationships to gain insight about geological processes. However, related methodological improvements often focus on efficiency and sensitivity, leaving geologically representative feature engineering underdeveloped. Improving modeling capabilities alone is insufficient, and both geological plausibility and representation of complex processes or features are critical to generate robust predictive models. Geological features involved in Magmatic Ni-Cu-PGE, Volcanogenic Massive Sulfide (VMS) Cu-Zn-Pb-Ag(-Au), and Orogenic Au mineral system prospectivity from two Archean greenstone belts from the southern Superior Province near Timmins and Dryden, Ontario are examined and compared using a variety of statistical techniques. Specifically, this PhD research explores 1) current knowledge and characteristic geological features for both greenstone belts, 2) methods to enhance geological knowledge using whole rock geochemistry, 3) methods to reduce bias and improve repeatability when mapping structural complexity, 4) how data science and geological understanding of mineral systems can be integrated for enhanced feature engineering 5) which factors control, or are most strongly associated with mineralization, and 6) why the greenstone belts near Timmins and Dryden, Ontario have contrasting orogenic Au endowment. Data-related outcomes of this research include multi-disciplinary geoscientific databases (e.g., structural, field observations, geochemistry), new bedrock geology maps for each area, and reprocessed aeromagnetic grids. Methodological outcomes of this research include new geochemical classification diagrams for ultramafic to felsic (including tonalite-trondhjemite -granodiorite and lamprophyre) Archean igneous rocks, Igneous Rock Favorability indices, automatic mapping of structural complexity from bedding measurements and aeromagnetic lineaments, mapping pre-deformation fluid path distances, mapping rheological and chemical contrast, semi-discrete interpolation of characteristic element ratios, as well as mapping mobile element gain/loss. Geological knowledge outcomes include the importance ranking of factors controlling magmatic, volcanogenic, and orogenic prospectivity from random forests as well as geological insight about contrasting orogenic Au endowment in the Timmins and Dryden areas. Overall, this research demonstrates the importance of integrative studies that leverage multi-disciplinary data, methods, and knowledge to improve existing geological understanding, maximize data utility, and generate robust exploration targets. These improvements may enhance exploration under difficult conditions, such as in data sparse environments, regions affected by clustered/partial data representation, or targets under cover.

Keywords

Mineral exploration; Machine learning; Geochemical classification; Structural complexity; Magmatic Ni-Cu-PGE; Volcanogenic Massive Sulfide Cu-Zn-Pb-Ag(-Au); Orogenic Au; Southern Abitibi; Western Wabigoon

External summary

Development of techniques that spatially represent and rank geological features controlling Magmatic Ni-Cu-PGE, Volcanogenic Massive Sulfide Cu-Zn-Pb-Ag(-Au), and Orogenic Au mineral system components in Archean greenstone belts near Timmins and Dryden, Ontario. The study summarizes current geological knowledge, applies statistical methods to enhance knowledge using geochemistry, and explores various techniques to map structural complexity, pre-deformation fluid path distances, rheological and chemical contrast, magmatic geochemistry, and hydrothermal alteration. Random forests importance ranking identifies factors controlling mineralization and explain contrasting orogenic Au endowment in the studied greenstone belts. Multi-disciplinary data and methods are needed to improve existing geological understanding and exploration strategy.

Thesis Defence Committee/Comité de soutenance de thèse

Laurentian University/Université Laurentienne

Faculty of Graduate Studies/Faculté des études supérieures

University of Western Australia

Graduate Research School

Title of Thesis **Titre de la thèse**	Factors contributing to metal endowment in the western Wabigoon and southern Abitibi subprovinces: A machine learning approach to Precambrian greenstone belts		
Name of Candidate **Nom du candidat**	Rebecca Montsion		
Degree **Diplôme**	Doctor of Philosophy (PhD) in Precambrian and Economic Geology		
Department/Program **Département/Programme**	Harquail School of Earth Sciences, Laurentian University	**Date of Defence**	June 22, 2023
	School of Earth Sciences, University of Western Australia	**Date de la soutenance**	

APPROVED/APPROUVÉ

Thesis Examiners/Examinateurs de thèse:
Dr. Stéphane Perrouty
(Supervisor/Directeur de thèse)
Dr. Mark Lindsay
(Supervisor/Directeur de thèse)
Dr. Mark Jessell
(Co-supervisor/Directeu de thèse)
Dr. Ross Sherlock
(Co-supervisor/Directeur de thèse)

Dr. Eric Grunsky
(Examiner/Examinateur externe)
Dr. Matthew Cracknell
(Examiner/Examinateur externe)

Approved for The Faculty of Graduate Studies
Approuvé pour la Faculté des études supérieures
Dr. Tammy Eger
Acting Dean, Faculty of Graduate Studies
Doyen intérimaire, Faculté des études supérieures

ACCESSIBILITY CLAUSE AND PERMISSION TO USE

Thesis declaration

I, Rebecca Montsion, certify that:

This thesis has been substantially accomplished during enrolment in this degree.

This thesis is my own work and does not contain any material previously published or written by another person, except where due reference has been made in the text or Authorship Declaration.

This thesis does not contain material which has been submitted for the award of any other degree or diploma in my name, in any university or other tertiary institution.

In the future, no part of this thesis will be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of The University of Western Australia and where applicable, any partner institution responsible for the joint-award of this degree.

This thesis does not violate or infringe any copyright, trademark, patent, or other rights whatsoever of any person.

Signature:

Date: 2023/03/15

Acknowledgements

I would like to begin by acknowledging and paying respect to the Ojibwe/Chippewa, Mushkegowuk (Cree), Algonquin, and Métis Peoples near Timmins, Ontario and the Anishinabewaki and Michif Piyii (Metis) Peoples near Dryden, Ontario who are the cultural stewards of land explored during this research project. A special thank you goes to Adam and Tess Peterson for their assistance with field work logistics (and rescue) as well as for educating me about local cultural traditions and significance.

I would like to express my heartfelt gratitude to my PhD supervisors Drs Stéphane Perrouty, Mark Lindsay, Mark Jessell, and Ross Sherlock, who have gone above and beyond the call of duty to ensure that I complete this thesis and retain some sanity. Thank you for using patience, wit, and encouragement to talk me out of picking potatoes for a living and pursuing a seemingly endless number of research rabbit holes. You have taught me that research is not just about hard work, but also about having a good sense of humor. I promise to pay it forward by using my PhD powers for good, not evil.

Fieldwork was a major component of this graduate experience and is the subject of my most precious PhD memories. Wabigoonies (Metal Earth Dryden transect team members), thanks for keeping me alive and pleasantly insane while we battled the Wapageisies. Honorary members: Mark Dirt and Chubbs the camp dog, field assistants: Austin Goncalves, Alec Graham, Katharina Holt, Luc Leleannec, Lauren Norenberg, Jordan Peterzon, Luke Smith, mapping partners: Amokelani Mavundza, David Downie, Kendra Zammit, and our fearless field leaders: Ben Frieman and Danger Wolf (Stéphane Perrouty), thanks for being an amazing research family. Fieldwork as a Becktatoe would have been impossible and very boring without you. I would also like to thank Craig Ravnaas of the Kenora District Geologist's office, Ontario Geological Survey, for sharing his extensive regional knowledge and assistance with compilation work. A big thank you also goes out to Wesley Webb for being our Gold Rock mud ferry and Aaron DesRoches from the National Waste Management Organization for collaboration during fieldwork.

A very heartfelt thank you goes out to Brandon Smith who survived a BSc thesis under my co-supervision and was a much appreciated assistant digitizing a zillion outcrops. I am also grateful to the Geoscience Laboratories' staff for their assistance in compiling and reconciling legacy geochemical datasets. A special thanks to the Geological Survey of Canada for the use of GIS mapping tools, which assisted in the map interpretation, as well as Seequent for use of Oasis Montaj, PCI Geomatics for use of Focus, REFLEX for use of IoGas, and ESRI for use of ArcGIS/ArcPro.

I also deeply appreciate the feedback given from my thesis examiners Drs Matthew Cracknell and Eric Grunsky as well as the anonymous reviews for my five peer-reviewed publications.

On a personal note, I would like to thank my bearded half, Paul Montsion, for putting up with my constant absences, never-ending deadlines, and extensive hours spent in front of the computer. He deserves a gold medal for listening to me endlessly talk about 'science' while being my sounding board, personal IT department, emotional support, and caffeine supplier. To my life-long role models, Victoria and Anita Hallas, my sulfide-oblivious fluff ball (Auri), and everyone in my family, thank you for cheering me on through the madness. To my friends both near and dear, thank you for indulging my outrageously bad science humour when the big words don't make sense. To my academic compatriots at LU and UWA, thanks for sharing the mental burden of grad-school and talking about rocks instead of statistics when the numbers were scarry. To my mentors Eric de Kemp, Ken Whitherly, Melissa Anderson, Sharron Carr, Sandi Occhipinti, and countless others, thank you for believing in me when I struggled to believe in myself. Everyone, your countless sacrifices made this dream a reality. Without you, I would have been lost in a world of data and confusion. Your support has been everything to me, and I dedicate this life achievement to you.

Authorship declaration and Co-authorship statements

In accordance with Laurentian University and the University of Western Australia's regulations regarding Research for Higher Degrees, this thesis is presented as a series of journal articles and consists of two published articles and three submitted manuscripts for peer-reviewed journals. Spelling and grammatical formats follow American English standards in accordance with the respective journals.

Thesis

The thesis was designed by the candidate, and the supervisory committee Drs Stéphane Perrouty, Mark Lindsay, Mark Jessell, and Ross Sherlock. All fieldwork, data collection, and interpretation of multi-disciplinary geoscientific data was completed by the candidate with guidance from the supervisory committee listed above. Additional collaboration by co-authors listed herein include local geological knowledge, domain specific expertise, constructive criticism, and manuscript editing.

Bibliographic details of publication/manuscript 1

Montsion, R. M., Perrouty, S., & Frieman, B. M. (2021). Geological and geophysical data compilation for the western Wabigoon and southern Abitibi subprovinces of the Superior Province, Ontario, Canada. Data in Brief, 37, 107159.
https://doi.org/10.1016/J.DIB.2021.107159

Description

Methods and results of chapter three have been published in Data in Brief (2021, v. 37, p. 107159). The candidate is first-author of the publication and is co-authored by Drs Stéphane Perrouty, and Ben Frieman.

Fieldwork, data collection, and geological map interpretation was performed by the candidate, Dr Perrouty and Dr Frieman. Additional field observations and open-source data compilation were the result of collaboration with MSc Candidates Kendra Zammit and David Downie. Method integration of all geological, geospatial, and geophysical datasets was a collaboration between the candidate and all co-authors. Financial and logistical support was coordinated by Drs Perrouty and Frieman.

Location in thesis: Chapter 3 (pages 46-57)
Contributions: Candidate 70%, Dr Perrouty 20%, Dr Frieman 10%
Co-author signatures and dates:

Stéphane Perrouty   Ben Frieman

2023/03/16   3/16/23

Copyright statement

Bibliographic details of publication/manuscript 2

Montsion, R. M., Perrouty, S., & Lesher, C. M., (*submitted*). Geochemical data compilation for the western Wabigoon and southern Abitibi subprovinces of the Superior Province, Ontario, Canada. Data in Brief

Description

Methods and results of geochemical compilation in chapter four have been prepared for publication in Data in Brief. The candidate is first-author for both publications and is co-authored by Drs Stéphane Perrouty and Michael Lesher. Fieldwork and data collection were performed by the candidate and Dr Perrouty. Quality control of geochemical data was performed by the candidate, Dr Perrouty, and Dr Lesher. Method development and interpretation of results was performed by the candidate and all co-authors. Financial and logistical support was coordinated by Dr Perrouty.

Location in thesis: Chapter 4 (pages 58-65)

Contributions: Candidate 70%, Dr Perrouty 20%, Dr Lesher 10%

Co-author signatures and dates:

Stéphane Perrouty   Michael Lesher

2023/03/16   2023-03-16

Copyright statement

N/A

Bibliographic details of publication/manuscript 3

Montsion, R M, Perrouty, S, Lesher, C M, Lindsay, M D (*submitted*). Statistical optimization for 5D classification of Archean igneous rock and application for exploration targeting in the Superior Province, Canada

Description

Methods and results of chapter five have been prepared for publication in Precambrian Research. The candidate is first-author for both publications and is co-authored by Drs Stéphane Perrouty, Michael Lesher, and Mark Lindsay. Fieldwork and data collection were performed by the candidate and Dr Perrouty. Quality control of geochemical data was performed by the candidate, Dr Perrouty, and Dr Lesher. Method development and interpretation of results was performed by the candidate and all co-authors. Financial and logistical support was coordinated by Dr Perrouty.

Location in thesis: Chapter 5 (pages 66-102) and supplementary material (pages 198-205)
Contributions: Candidate 70%, Dr Perrouty 20%, Dr Lesher 5%, Dr Lindsay 5%
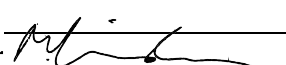**Co-author signatures and dates:**

Stéphane Perrouty   Michael Lesher   Mark Lindsay

2023/03/16   2023-03-16   **16 March 2023**

Bibliographic details of publication/manuscript 4

Montsion, R. M., Perrouty, S., Lindsay, M. D., Jessell, M. W., & Frieman, B. M. (2021). Mapping structural complexity using geophysics: A new geostatistical approach applied to greenstone belts of the southern Superior Province, Canada. Tectonophysics, 228889. https://doi.org/https://doi.org/10.1016/j.tecto.2021.228889

Description

Methods and results of chapter six have been published in Tectonophysics (2021, v. 812, p. 228889). The candidate is first-author of the publication and is co-authored by Drs Stéphane Perrouty, Mark Lindsay, Mark Jessell, and Ben Frieman. Method development and interpretation of results was a collaboration between the candidate and all co-authors. Financial and logistical support was coordinated by Dr Perrouty.

Location in thesis: Chapter 6 (pages 103-138) and supplementary material (pages 206-211)
Contributions: Candidate 70%, Dr Perrouty 15%, Dr Lindsay 5%, Dr Jessell 5%, Dr Frieman 5%
Co-author signatures and dates:

Stéphane Perrouty	Mark Lindsay	Mark Jessell	Ben Frieman
2023/03/16	16 March 2023	16/03/2023	3/10/23

Bibliographic details of publication/manuscript 5

Montsion, R. M., Perrouty, S., Lindsay, M. D., Jessell, M. W., & Sherlock, R. (*submitted*). Development and application of feature engineered geological layers for ranking magmatic, volcanogenic, and orogenic system components in Archean greenstone belts.

Description

Chapter seven is intended for submission to a peer reviewed journal in June 2023. The candidate is first author for this manuscript and is co-authored by Drs Stéphane Perrouty, Mark Lindsay, Mark Jessell, and Ross Sherlock. Financial and logistical support was coordinated by Dr Perrouty.

Location in thesis: Chapter 7 (pages 139-178) and supplementary material (pages 212-256)
Contributions: Candidate 70%, Dr Perrouty 8%, Dr Lindsay 8%, Dr Jessell 8%, and Dr Sherlock 6%
Co-author signatures and dates:
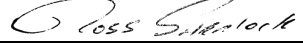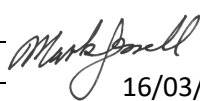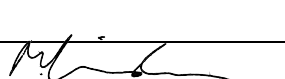
Stéphane Perrouty	Mark Lindsay	Mark Jessell	Ross Sherlock
2023/03/16	16 March 2023	16/03/2023	11-04-2023

Copyright statement			
N/A			

Table of Contents

List of Tables

List of Figures

geological knowledge to attribute value to each dataset. Combined, this information is used to design relevant feature engineering techniques that isolate signals or generate representations of mineral system components. Eight new and existing feature engineering methods are indicated as an inset to the diagram. Target and background locations are synchronously identified and used to constrain training random forests classification and generate a ranked list of feature importance

constrain interpolation and grid results are compared in the legend to assess reprehensively of geochemical trends in data captured by the interpolation. 235

Preface

"The real voyage of discovery consists not in seeking new landscapes, but in having new eyes"

– Marcel Proust

Stylized oil painting modified from photo of a boulder beach on the shores of Wabigoon Lake near Dryden, Ontario. Photo credit: Rebecca Montsion 2021

Chapter 1. Introduction

The mineral exploration industry is undergoing a transformation with the incorporation of quantitative techniques and probabilistic targeting using machine learning. These methods provide the geoscientific community with tools to better identify hidden patterns in convoluted datasets, predict the location of mineralized occurrences, and gain a better understanding of geological processes (Cracknell and Reading, 2014; Rodriguez-Galiano, 2015; Bérubé et al., 2018; Acosta et al., 2019; Karpatne et a., 2019; Zhang et al., 2022). While there is a focus on improving the efficiency/sensitivity of these tools (e.g., Brown et al., 2000; Granek, 2016, Chen and Wu, 2017) or selecting the best method (e.g., Cracknell and Reading, 2014; Rodriguez-Galiano et al., 2015; Schaeben et al., 2019; Sun et al., 2020), the development of geologically representative feature engineering is underdeveloped. In many machine learning applications, feature engineering embeds compounding ambiguity, oversimplification, and subjective bias in the early stages of exploration workflows, which can lead to spurious or misleading results (Feltrin et al., 2016; Karpatne et al., 2019; Zekri et al., 2019;). Although it is crucial to enhance modeling capabilities, it is equally important to ensure that inputs are geologically plausible and reasonably represent the complex geology being investigated. New methods that jointly leverage existing geological knowledge and quantitative approaches are essential to produce more robust predictive models and extract meaningful insights.

1.1.   Research objectives

This PhD research aims to provide the mineral exploration industry with tools that capture representative geological proxies and features (i.e., maps) controlling mineral system prospectivity. These tools were designed to improve on common feature engineering methods that introduce compounding ambiguity, oversimplification, and/or subjective bias in early stages of exploration workflows. New and adapted feature engineering methods are demonstrated by integrating disparate geoscientific datasets, new field-based observations, statistical analyses, and machine learning in two Archean greenstone belts from the southern Superior Province near Timmins and Dryden, Ontario. These greenstone belts are used as comparative case studies and provide a framework to rank factors controlling Magmatic Ni-Cu-PGE, Volcanogenic Massive Sulfide (VMS) Cu-Zn-Pb-Ag(-Au), and Orogenic Au mineralization. The following research questions will be addressed herein:

Chapter 2 - 3

I. What is the current state of knowledge and key geological and structural features for the Archean greenstone belts near Timmins and Dryden, Ontario?

Chapter 4 - 5

II. How can we enhance geological knowledge of the two greenstone belts and their geodynamic setting using statistical analyses of whole rock geochemistry?

Chapter 6

III. How can we reduce bias and improve repeatability when mapping structural complexity?

Chapter 7

IV. How can data science and geological understanding of mineral systems be used to capture critical components governing magmatic, volcanogenic, and orogenic fertility?

V. Which factors control, or are most strongly associated with, magmatic Ni-Cu-PGE and volcanogenic massive sulfide Cu-Zn-Pb-Ag(-Au) mineralization in Precambrian greenstone belts?

VI. Why do the seemingly similar greenstone belts near Timmins and Dryden, Ontario have contrasting orogenic Au endowment?

1.2. Statement of scientific contributions

A summary of the original contributions by the candidate to this study is as follows:

1.2.1. Datasets

- Regional-scale geological maps for the Timmins and Dryden field areas
- Geospatial database of outcrop lithology observations (N ≈ 17,000) for Dryden
- Geospatial database of structural measurements for Dryden (N ≈ 28,000) and Timmins (N ≈ 8,000)
- Re-processed aeromagnetic grids for the Timmins and Dryden areas
- Compilation of whole rock geochemistry of metamorphic, igneous, and sedimentary rocks in the southern Abitibi and western Wabigoon subprovinces of the Superior Province in Canada (N ≈ 23,000)
- Table of idealized/averages rock properties for Archean rocks including: 1) magnetic susceptibly, total porosity, resistivity, grain density, bulk modulus, shear modulus, compression modulus, cohesion, friction coefficient, and uniaxial compression strength

1.2.2. Methodological development

- Classification diagrams for ultramafic to felsic (including tonalite- trondhjemite -granodiorite and lamprophyre) Archean igneous rocks using statistically and geologically defined element ratios that minimize effect of metasomatism as well as closure and are valid for intrusive and extrusive textures. Diagram capture igneous melt composition, initial source characteristics, evolution during ascent (assimilation, fractional crystallization).

- Igneous Rock Favorability indices (IRF) for samples with similar geochemical characteristics as samples nearby magmatic, volcanogenic, and orogenic deposits

- Calculation of structural complexity from bedding measurements and auto-detected aeromagnetic lineaments

- Calculation of non-Euclidean distance to geological map units and structures

- Calculation of rheological and chemical contrast maps

- Semi-discrete interpolation of characteristic element ratios characterizing Archean igneous rocks

- Calculation and interpolation of mobile element gain/loss

1.2.3. Knowledge

- Ranking of factors controlling magmatic, volcanogenic, and orogenic prospectivity

- Identification of factors that may have cause contrasting orogenic Au endowment in the Timmins and Dryden areas

Chapter 2. Geological context

Archean greenstone belts host many world-class base and precious metal deposits, making them an economically significant target for exploration. Generally, these ancient orogenic domains are composed of metamorphosed (greenschist to amphibolite facies) compositionally bi-modal ultramafic to felsic subvolcanic/extrusive igneous sequences and metasedimentary basins. These basins often lay between syn-volcanic to syn-tectonic ultramafic to felsic intrusions and may overly older gneissic crust (Ayer et al., 2005; Bleeker, 2015; Hutchison, 1983; Monecke et al., 2017b; Thurston et al., 2008a). Typically, the deformation history for Archean greenstone belts follows a commonly observed sequence with early-stage opening (rift or back-arc) and progressive to pulsed stages of collision expressed as isoclinal folding, development of brittle-ductile deformation zones, complex fold interference patterns, late transpression and post-orogenic gravitational collapse.

Within these geological environments, optimal conditions for the formation of base and precious metal deposits may be present. The following sections describe the geologic context for each of the study areas as well as three mineralizing systems (i.e., magmatic Ni-Cu-PGE, volcanogenic Cu-Zn-Pb-Ag(-Au), orogenic Au).

The Superior Province is the largest preserved Archean craton and consists of Eo- to Neoarchean granite-greenstone, gneissic, and metasedimentary domains that amalgamated as a series of generally east-trending belts between ~2720 and ~2660 Ma (Percival et al., 2012; Frieman et al., 2017). Both selected regions (Figure 2-1) have a similar geologic history in terms of geodynamic setting, volcanism, deformation, and metamorphism (Beakhouse et al., 2011; Zammit, 2020). Since all rock units in both areas, with the exception of Proterozoic diabase dikes, are metamorphosed to greenschist-amphibolite facies, the term 'meta' will not be used in lithologic descriptions.

Initial phases of greenstone belt formation typically include the accumulation of ultramafic to mafic volcanic rocks and intermittent felsic flows in the upper phases of volcanism. These bimodal volcanic packages were intensely folded and imbricated during an early stage of collision (Ayer et al., 2005, 2002; Bleeker, 2015; Thurston et al., 2008b). Following shortening, a period of uplift and localized extension caused crustal-scale faults and clastic sedimentary basins to form (Bleeker, 2015). The basins consist of turbiditic sequences of wacke, siltstone, and mudstone that unconformably overlay and are folded with older volcanic units (Porcupine-type; Ayer et al, 2002). Later, a period of renewed shortening inverted faults and caused rapid deposition of fluvial-alluvial dominated conglomerate sequences on the footwall of inverted faults. These sequences unconformably overlay the turbiditic and volcanic units (Timiskaming-type; Thurston and Chivers 1990).

Major components of these settings that relate to fertility/prospectivity of mineral systems and relevant exploration criteria are (1) crustal-scale faults that experienced extension followed by transpression (basin inversion), (2) calc-alkaline or alkaline magmatism that is temporally related to extension (magmatic Ni-Cu-PGE, volcanogenic Cu-Zn-Pb-Ag(-Au) or transpression (orogenic Au), and (3) the preservation of syn-orogenic sedimentary basins as potential sources of fluids (orogenic Au).

Broad similarities and differences between the two greenstone belts are summarized below in Table 2-1

2.1. Timmins, southern Abitibi subprovince

The Timmins study area is comprised of ultramafic to felsic volcanic assemblages that are unconformably overlain by two sedimentary successions. Thick glacial cover and overburden limits outcrop exposure; however extensive drilling has sufficiently constrained geological interpretations within the Timmins region. Volcanic successions range in age from ~2750 to ~2695 Ma and include disconformities that are locally marked by iron formation and cherts (Thurston and Chivers, 1990; Thurston et al., 2008; Frieman et al., 2017). The Porcupine sedimentary assemblage (ca. 2690 – 2685 Ma, Frieman et al. 2017) is comprised of fine- to medium-grained beds with rare conglomeratic lenses, iron formations, and mafic to felsic volcanic flows. The Timiskaming coarse- to medium-grained sedimentary assemblage (ca. 2679-2669 Ma; Frieman et al., 2017) is primarily preserved as narrow (<2 km) panels to the north of the Porcupine-Destor deformation zone (PDDZ; Ayer et al., 2005; Bateman et al., 2005; Bleeker, 2015; Montsion et al., 2021a/Chapter 3). A comprehensive geological overview is provided in Ayer et al. (2002, 2005), Bateman et al. (2005), Bleeker (2015), Monecke et al. (2017) and references therein.

Table 2-1 (next page). Summary of broad differences and similarities between the greenstone terranes in the Abitibi near Timmins, ON and western Wabigoon near Dryden, ON

Variable	Feature	Southern Abitibi (Timmins)	Western Wabigoon (Dryden)
Basement	Gneiss	Unknown	Paleo to Mesoarchean crust (ortho and para gneiss)
Volcanic stratigraphy (% area)	Ultramafic to mafic	Present	Absent
	Mafic to intermediate	Dominantly tholeiitic (58.29)	Dominantly tholeiitic (43.27)
	Intermediate to felsic	FIII, rare FI and FII	FI, rare FII and FIII
Sedimentary basins (% area)	Rift basin turbidites/ Fine grained	27.8	9.6
	Molasse basin/ Coarse grained	0.9	2.2
Syn-volcanic intrusions (% area)	Ultramafic to mafic	1.5	Unreported
	Mafic to intermediate	2	4
	TTG to felsic		23.7
Structural evolution	Early isoclinal folding	Present	Present
	Early reverse movement	Present	Unknown
	E-W trending crustal-scale deformation zone(s)	Trans-crustal structure with connected & anastomosing splays	Network of sub-parallel and unconnected (?) anastomosing structures
	Re-folding	Progressive and complex	Progressive and complex
	NE-SW trending crustal-scale deformation zone	Absent	Present
	late-brittle faults	Present	Present
Syn-deformation intrusions (% area)	Ultramafic to mafic	1	0.1
	Mafic to intermediate	0.3	0.2
	Intermediate to felsic	8.1	6.1
Metamorphic grade		Greenschist to amphibolite	Greenschist to amphibolite
Age	Gneisses	Absent	Paleo-Mesoarchean protolith (3550-2800 Ma)
	Volcanic	Neoarchean (2750-2695)	Neoarchean (2750-2715 Ma)
	Syn-volcanic intrusions (ultramafic to felsic)	Neoarchean (2750-2695)	Neoarchean (2750-2715 Ma)
	Turbidites/fine-grained sedimentary basin (± local volcanism)	Neoarchean (2690-2682)	Neoarchean (ca. 2710-2700)
	Molasse basin and locally alkalic magmatism	Neoarchean (2679-2669)	Neoarchean (ca. 2700-2695)
	S-type granitoids & pegmatites	Neoarchean (2668-2632)	Neoarchean (ca. 2685-2640)
Magma source	Ultramafic to intermediate melt source composition	Mainly homogenous, depleted, and metasomatized mantle	Mainly homogenous, depleted, and metasomatized mantle
	Intermediate to felsic melt source composition	Neoarchean supracrustal rocks and TTGs	Paleo to Mesoarchean crust (ortho and para gneiss) and minor Neoarchean supracrustal rocks
	Source depth	Shallow to deep	Mainly deep
Resources *produced + measured + indicated*	Magmatic Ni-Cu-PGE	4.64 (Mt)	Unknown
	VMS Cu-Zn-Pb-Ag(-Au) (± orogenic overprint)	174.3 (Mt)	18.2 (Mt)
	Orogenic Au	93.8 (Moz) at average grade 8 (g/t)	3.6 (Moz) at average grade 6.7 (g/t)
Alteration	CO_2	Pervasive sea floor alteration with anomalies at deformation zone/splay junctions from hydration	Focused at deformation zone/splay junctions
	SiO_2	Gains along unconformities and losses along major deformation zones	Gains near intrusions and contact with basement (gneiss). Losses along deformation zones
	K_2O	Gains along contacts with syn-volcanic ultramafic intrusions and losses along contacts of ultramafic volcanic units	Moderate gains anomies in basement (gneiss) and losses along deformation zones or at contacts with small syn-tectonic porphyries
	Na_2O	Gains along deformation zones and losses spatial associated with ultramafic rocks	Gains within sedimentary basins or small syn-tectonic porphyries and losses along deformation zones

The volcanic and sedimentary packages were intruded by ultramafic to felsic bodies that have been classified based on composition and absolute/relative timing (i.e., syn-volcanic and syn-deformation). Syn-volcanic intrusions are large, batholitic complexes that display gabbroic to granodioritic compositions, similar to their volcanic counterparts, and are generally found west of

the Mattagami River Fault (Corfu, 1993; Mortensen, 1993a,b; Ayer et al., 2005, Montsion et al., 2021/Chapter 3). Syn-deformational intrusions, which dominate the majority of the map area, cross-cut sedimentary and volcanic assemblages, are variably deformed depending on proximity to deformation zones, and are tonalitic to "potassic" (i.e., granite, monzonite, and syenite) in composition (Ayer et al., 2002). Late intrusive intermediate to felsic porphyritic clusters as well as several diabase dike swarms intrude supracrustal units. The latter include the north-northwest trending Matachewan dikes (ca. 2445-2490 Ma; Ernst and Buchan, 2010), northwest-trending Sudbury dikes (~1235 Ma; Ernst and Buchan, 2010), and northeast-trending Abitibi dikes (~1140 Ma; Ernst and Buchan, 2010).

The structural history of the Timmins camp is well documented (Ayer et al., 2002, 2005; Bateman, 2008; Bleeker, 2015; Monecke et al., 2017), with two main shortening events. A first north-south shortening event resulted in kilometer-scale isoclinal folds in the volcanic and Porcupine sedimentary assemblages, associated with an axial planar foliation mainly marked by chlorite in volcanic units and micas in sedimentary units. Generally east-trending brittle-ductile deformation zones such as the crustal-scale Porcupine-Destor Deformation Zone (PDDZ; Ayer et al., 2005; Bateman et al., 2008, Montsion et al., 2021/Chapter 3) locally controlled the unconformable deposition of the Timiskaming sediments (Bateman et al., 2008; Bleeker, 2015). A second northwest-southeast shortening event resulted in the refolding of the isoclinal folds, formation of asymmetric z-folds, and the development of a crenulation cleavage axial planar to z-folds. This event is coeval with the formation of world-class orogenic Au deposits along the PDDZ (Dubé and Gosselin, 2007). Late Neoarchean or Proterozoic north-northwest trending faults, such as the Mattagami River, Burrows Benedict, and Buskegau River faults, offset older structural features.

The Timmins mining camp hosts several world-class mineral systems, including the Kidd-Creek Volcanogenic Massive Sulfide (VMS; 9.6 Mt Zn, 3.7 Mt Cu), Dome (27.1 Moz Au), Hollinger (20.5 Moz Au), McIntyre (10.8 Moz Au), Hart (1.86 Moz Ni), Langmuir 1 and 2 (1.21 Moz Ni), as well as 10+ smaller orogenic deposits (>1 Moz Au). Total endowment represents over 2700 t of Au, 10 Mt of zinc, 4 Mt of copper, and 60000 t of nickel (measured and indicated, Ontario Geological Survey, 2019; Montsion et al., 2021a/Chapter 3 and references therein).

2.2.   Dryden, western Wabigoon subprovince

The western Wabigoon subprovince is primarily composed of juvenile Neoarchean volcanic successions that are bounded by the Paleo- to Neoarchean Winnipeg River and Marmion gneissic-plutonic terranes to the north and south, respectively (Percival et al., 2004; Tomlinson et al., 2004). The Dryden study area hosts felsic, intermediate, and mafic volcanic formations (ca. 2750–2715 Ma,

Davis et al., 1989; David et al., 2005), marine turbiditic sequences (ca. 2715–2710 Ma; Davis et al., 2005), and fluvial-alluvial clastic sedimentary units (ca. 2703–2696 Ma; Beakhouse et al. 1995; Dostal et al., 2004; Corcoran and Mueller, 2007).

The volcanic and sedimentary packages were intruded by tonalitic to granodioritic syn-volcanic batholiths (ca. 2750–2680 Ma; Blackburn et al., 1981; Davis et al., 2005; Duguet and Beakhouse, 2012), such as the Revell batholith, Atikwa-Lawrence batholith, and Basket Lake batholith. Smaller dioritic, granitic and syenitic porphyritic intrusions such as the Lost Lake porphyry, Thundercloud porphyry, and Taylor Lake stock (geological map in Montsion et al., 2021a/Chapter 3) intruded during syn-volcanic to post-deformation phases (ca. 2705–2695 Ma). Finally, two generations of Proterozoic diabase dikes crosscut the western Wabigoon subprovince. These are northwest-trending dikes (ca. 1887-1889 Ma; Davis and Hamilton, 2010) and north-trending dikes (~1137 Ma; Heaman and Easton, 2006; Stone et al., 2010)

Deformation began in the Neoarchean (~2700 Ma) with north-south shortening that resulted in kilometer-scale isoclinal to outcrop-scale parasitic folds with an associated axial planar foliation that is mainly marked by chlorite in volcanic units and micas in sedimentary units (Kresz et al., 1982a, 1982b; Montsion et al., 2019; Zammit 2020). Progressive shortening resulted in the localization of strain into several east-trending deformation zones, including the Wabigoon and Mosher Bay-Washeibemaga deformation zones (Montsion et al., 2021s/Chapter 3; Blackburn, 1978; Zammit, 2020), which underwent north-over-south movement (Zammit, 2020) associated with the unconformable deposition of coarse-grained sediments in the Stormy and Manitou basins (Kresz et al., 1982a, 1982b; Corcoran and Mueller, 2007; Montsion et al., 2021a/Chapter 3). A later stage of northwest-southeast shortening, coeval with most orogenic Au mineralization in the area (Melling et al., 1988; Zammit, 2020), refolded the earlier isoclinal folds (i.e., locally increasing structural complexity), and has been interpreted as recording sinistral transpressive strain along the northeast-trending Manitou-Dinorwic deformation zone (geological map in Montsion et al., 2021a/Chapter 3; Parker, 1989; Zammit, 2020).

Although the western Wabigoon subprovince is apparently much less endowed with base and precious metal resources compared to the Abitibi subprovince, it nevertheless hosts several deposits and developed prospects with a variety of metal resources (gold, silver, copper, zinc, iron, lithium, molybdenum, uranium, thorium, and rare earth elements; Parker, 1989). Gold systems include Goliath (1.2 Moz Au), Goldlund (0.8 Moz Au), Gold Rock (0.15 Moz) and Van Horne (0.002 Moz, past production plus measured and indicated resources, Ontario Geological Survey, 2019; Montsion et al., 2021a/Chapter 3 and references therein).

Chapter 3. Geological and geophysical data sources with new geological map interpretations

Copyright statement:

Abstract

The geoscientific data presented in this paper are a foundation for experimental and exploration geological research in the western Wabigoon and southern Abitibi subprovinces of the Superior Province in Ontario, Canada. New geological interpretations, in map and GIS formats, along with compiled mineral deposit information, structural databases, magnetic susceptibility measurements, and reprocessed aeromagnetic grids have been integrated to provide a basis for comparative studies between the two geologically similar yet economically disparate greenstone belts near Dryden and Timmins, Ontario, Canada. Data were acquired from a wide range of publicly sourced data releases and enhanced through the addition of new observations. New geological maps presented for both regions represent the culmination of integrating the multi-disciplinary geoscientific database and recent geological interpretation. Data contained within this publication are co-submitted with Montsion et al. (2021b)/Chapter 6.

Keywords

Value of the Data

- This dataset represents a comprehensive and up-to-date regional, open-access, geoscientific compilation for the western Wabigoon (centred on Dryden, Ontario, Canada) and southern Abitibi (centred on Timmins, Ontario, Canada) subprovinces
- All of the included datasets are intended for use by geologists conducting geoscientific studies in both or either of the areas of interest

- Included data can be used as a foundation for mineral exploration, regional-scale investigations, and comparative studies

Data access

All data-related supplementary material referenced in this chapter is available in its publication in *Data in Brief* and Mendeley Data (https://data.mendeley.com/datasets/rvnk26krjs/2)

Acknowledgments

3.1. Data Description

This comprehensive, multi-disciplinary geoscientific database provides access to observational, measured, and interpreted data for two Archean greenstone belts centred on Dryden, Ontario Canada in the western Wabigoon subprovince and on Timmins, Ontario, Canada in the southern Abitibi subprovince. Layers (e.g. point, line, polygon) with a spatial component are listed within the 'Geology' folder as ESRI shapefiles. Grid files representing reprocessed geophysical data are provided as .tiff and .grd files. These have been spatially projected into the NAD 83 UTM coordinate system (zone 15N for Dryden and 17N for Timmins). For each layer, metadata descriptions relating to symbology, source (if applicable), and other relevant information are recorded in the associated attribute table. Detailed descriptions of compiled layers can be found in their respective source publications.

3.1.1. Geological data

Geological layers consist of georeferenced symbols and shapes representing geological observations (e.g., structural measurements, mineral occurrences, main lithology at outcrop) and interpretations (e.g., faults traces, map units). Digitized and compiled structural measurements ('Structural measurements') and mineral occurrence locations ('Mineral deposit index') are presented as Excel tables in separate folders in a Mendeley Data repository (Montsion et al., 2021a). Interpreted layers such as lithology polygons ('MapUnits') and structural traces ('Geolines') are provided as ESRI shapefiles in the 'Geology' folder and are organized according to their spatial location. References for data that constrained interpretations or contributed to compilations are reported in a separate folder of the Mendeley Data repository ('Data sources for constraint and compiled layers'; Montsion et al., 2021a).

3.1.2. New geological maps

Two new geological maps (Figures 3-1 and 3-2), one for each map area, present a scale appropriate interpretation of geological data and relationships. Maps are available as PDF sheets (in 'PDF Maps') or as ESRI map project file (in 'Geology'). The ESRI .mxd files are provided for comprehensive interrogation or to interact with individual layer files.

The new regional geological map near Timmins, Ontario in the southern Abitibi subprovince of the Superior Province (Figure 3-1) displays poly-deformed Archean bimodal volcanic stratigraphy with overlying sedimentary packages variably intruded by tonalitic to granitic plutons and smaller porphyritic bodies. Proterozoic diabase dikes crosscut stratigraphy as well as intrusive bodies. Geological interpretations were made by integrating compiled observations and interpretations from 33 geological maps and several compilation databases (listed in 'Data sources for constraint

and compiled layers'), geophysical map patterns, and new field observations. Labelled features within the map include the Burrows Benedict fault (BBF), Dome deformation zone (Ddz), Mattagami River Fault (MRF), North Pipestone deformation zone (NPdz), Porcupine Basin (PB), Porcupine-Destor deformation zone (PDdz), Pipestone deformation zone (Pdz), Timiskaming Basin (TB).

The new geological map of the Dryden area, Ontario in the western Wabigoon subprovince of the Superior Province (Figure 3-2) displays poly-deformed Archean bimodal volcanic stratigraphy with overlying sedimentary packages variably intruded by tonalitic to granitic plutons and smaller porphyritic bodies. Proterozoic diabase dikes crosscut stratigraphy as well as intrusive bodies. Geological interpretation integrated observations and interpretations from 64 geological maps (listed in 'Data sources for constraint and compiled layers'), geophysical map patterns, and new field observations. Map coordinates are in NAD 83 UTM zone 15N./ Labelled features within the map include Aiabewatik deformation zone (Adz), Atikwa-Lawrence batholith (Alb), Basket Lake batholith (BLb), Eltrut gneisses (Eg), Goldlund deposit (Gld), Goliath deposit (Gd), Kawashegamuk deformation zone (Kdz), Kenwest deposit (Kd), Larson Bay deformation zone (LBdz), Long Lake River antiform (LLRa), Manitou-Dinorwic deformation zone (MDdz), Melgund north deformation zone (Mndz), Melgund south deformation zone (Msdz), Mosher Bay-Washeibemaga deformation zone (MBWdz), Noonan Lake deformation zone (MLdz), Revell batholith (Rb), Stormy Basin SB), Suzanne Lake antiform (Sla), Thunder Lake antiform (tLa), Thunder Lake synform (TLs), Upper Manitou antiform (uMa), Van Horne deposit (VHd), Vermillion deformation zone (Vdz), Wabigoon deformation zone (Wdz).

3.1.3. Structural observations

In Dryden, 13612 previously published structural measurements (e.g., bedding, younging, lineations, and foliations) were digitized from 64 geological maps and combined with 1047 new observations from 2018/2019 fieldwork.

In Timmins, 8047 structural observations were compiled and integrated from several open access data sources and 33 scanned maps listed in 'Data sources for constraint and compiled layers'. The source of each point is recorded in 'Data sources for constraint and compiled layers'. All structural data are released using a dip direction/dip format.

3.1.4. Economic geology

The location of mineral occurrences and deposits were extracted from the Ontario Geological Survey's Mineral Deposit Index (2020). The compiled table includes 254 locations in Dryden and 537 locations in Timmins ('Mineral deposit index'). Deposit name, grade, and tonnage were extracted where available; however, there were several gaps in quantitative records. Gaps in tonnage records

were filled by extracting information from each deposit's NI43-101 disclosure for mineral projects reports. Gaps in grade (g/t and oz/t) and totals were calculated in the included Excel spreadsheet. The MDI identification number can be used to retrieve additional information from the Ontario Geological Survey website (https://www.geologyontario.mndm.gov.on.ca/MDI_Description.html)

3.1.5. Geophysical data

Two geophysical datasets are included in this publication: 1) petrophysical measurements from recent field work in the Dryden area and 2) an ensemble of reprocessed aeromagnetic grids. These can be assessed from the 'Geophysics' folder in the Mendeley Data repository (Montsion et al., 2021a)

3.1.6. Petrophysical data

A new dataset of magnetic susceptibility measurements from 2018 and 2019 field work in the Dryden area ('Magnetic susceptibility') reports results of 4022 readings collected at 809 outcrops, using a Terraplus KT-10 magnetic susceptibility meter. At each outcrop, several readings of the same rock type were recorded to minimize bias in the collected data and provides a more complete analysis of the rock's magnetic susceptibility. Magnetic susceptibility measurements in the Timmins area can be accessed from the Ontario Geological Survey website (https://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=mrd273-rev2)

3.1.7. Reprocessed aeromagnetic grids

An ensemble of reprocessed aeromagnetic grids (40 m by 40 m cell size) from the Ontario Geological Survey's (2011) Stormy Lake and (2003) Timmins geophysical survey datasets are included as projected geo .tiff and .grd files. The ensemble includes total magnetic intensity reduced to pole with cosine roll off filtering, first and second vertical derivatives, tilt derivative, and dynamic range compression. Hill-shaded relief grids are also included, and their azimuth and angle of incidence are indicated in their filename. The original total magnetic intensity and reduced to pole aeromagnetic grids can be accessed from the Ontario Geological Survey website (OGS, 2003; 2011)

3.2. Experimental Design, Materials and Methods

3.2.1. Geological mapping

Observational geological layers (e.g., structural measurements and mineral occurrence information) were compiled from several open-source geoscientific compilations, digitized from archived scanned maps, extracted from geological reports, and supplemented with new observations. A summary

table of methods, purpose, and sources for each data set is provided ('Data sources for constraint and compiled layers').

3.2.2.   Structural interpretations

For both the Dryden and Timmins databases, geological interpretations (e.g., fault traces, deformation zones, contacts, map units) were generated through an iterative process where interpretations from previous works at various scales (see source map citations from structural observations) were compared to structural data, dominant outcrop lithologies, geophysical grids, geochronological analyses, and mineral occurrence observations. Interpretations were continuously refined by targeted fieldwork and repeated comparisons to observational data. This iterative process is common for regional-scale geological investigations (Gunn et al., 1997; Aitken and Betts, 2009; Metelka et al., 2011; Perrouty et al., 2012; Isles and Rankin, 2013; Blundell et al., 2019).

During comparisons to geophysical grids, spatially continuous, roughly linear anomalies in aeromagnetic grids were assumed to represent variably magnetic volcanic (or sedimentary) layering. This assumption is supported by similarly oriented bedding measurements.

3.2.3.   Map unit polygons

Map units were generated as polygons that outline the extent of a lithology type for a given unit within a litho-tectonic assemblage. Similar to structural interpretations, polygons were interpreted by iteratively cross-referencing all available observational data, geophysical interpretations, and previous works in the (i.e., scanned maps). Map units were classified based on lithology type and stratigraphy group using the schema in Table 3-1.

3.2.4.   New geological maps

Precambrian bedrock and assemblage maps for both areas were generated by overlying observations and interpretations (filtered to be representative of geology at a 1:75 000 map scale) and are presented as geological maps in Figures 3-1 and 3-2.

Table 3-1. Codes and schema for map units in Dryden and Timmins ArcMap project. Units and assemblages matched by approximate age and relative relationships

Wabigoon				Timmins		
Label	**Description**	**Assemblage**		**Assemblage**	**Description**	**Label**
dd	Proterozoic age		**Diabase dikes**		Proterozoic age	dd
ldf	Felsic to intermediate		**Syn-deformation intrusions**		S-Type granites/ porphry suite	ldf
ldm	Mafic to intermediate				Mafic to intermediate	ldm
					Ultramafic to mafic	ldu
lvd	Intermediate to felsic		**Syn-volcanic intrusions**		Intermediate to felsic	lvf
lvm	Mafic to intermediate				Mafic to intermediate	lvm
lvu	Ultramafic to mafic				Ultramafic to mafic	lvu
Stormy_TS	Clastic sedimentary	Stormy	**Volcano-sedimentary units**	Timiskiming	Clastic sedimentary	TS_ms
Stormy_bif	Banded iron formation				Banded iron formation	TS_bif
Stormy_vf	Felsic volcanic				Felsic volcanic	TS_vf
Stormy_vm	Mafic volcanic				Mafic volcanic	TS_vm
TZB_ms	Porcupine-type	Thunder Lake / Zealand/ Brownridge		Porcupine	Porcupine-type	PC_ms
TZB_bif	Banded iron formation				Banded iron formation	PC_bif
TZB_vf	Felsic volcanic				Felsic volcanic	PC_vf
TZB_vm	Mafic volcanic				Mafic volcanic	PC_vm
TZB_vmhi	(low / high mag)					
Boyer_vf	Felsic volcanic	Boyer Lake		Blake River	Clastic metasedimentary	Blake_cls
					Chemical metasedimentary	Blake_chs
Boyer_vm	Mafic volcanic				Felsic volcanic	Blake_vf
Boyer_vmHi	(low / high mag)				Mafic volcanic	Blake_vm
					Clastic metasedimentary	Tisdale_cls
					Chemical metasedimentary	Tisdale_chs
KW_vf	Felsic volcanic	Kawashegamuk / Wapageisi		Tisdale	Felsic volcanic	Tisdale_vf
KW_vm	Mafic volcanic				Mafic volcanic	Tisdale_vm
KW_vmHi	(low / high mag)				Ultramafic volcanic	Tisdale_vu
					Clastic metasedimentary	Kidd_cls
Wab_vf	Felsic volcanic	Wabigoon		Kidd-Munro / Stoughton-Roquemaure / Deloro / Pacaud	Chemical metasedimentary	Kidd_chs
Wab_vm	Mafic volcanic				Felsic volcanic	Kidd_vf
Wab_vmHi	(low / high mag)				Mafic volcanic	Kidd_vm
					Ultramafic to mafic volcanic	Kidd_vu
Og	Orthogneiss		**Basement**			
Pg	Paragneiss					
Bmg	Mafic gneiss					

3.2.5. Geophysical data processing

Raw data was acquired from the Ontario Geological survey and reprocessed as new geophysical grids for inclusion in this comprehensive database. For Dryden, aeromagnetic data were acquired between November 2000 and February 2001, with 200 m line spacing, 1500 m tie line spacing, and a flight elevation of 70 m (Ontario Geological Survey, 2011). Timmins aeromagnetic data was acquired between 1975 and 1992, with 200 m line spacing, 1000 m tie lines spacing, and a flight elevation of 70 m (Ontario Geological Survey, 2003). Total magnetic intensity (TMI) data were gridded at 40 m resolution in both areas, using the minimum curvature algorithm in Seequent's Oasis Montaj™. Reduction to the pole (RTP) was based on the IGRF (International Geomagnetic Reference Field), calculated at the date of the survey (Dryden: declination −0.9°, inclination 75.2°; Timmins: declination −11.2°, inclination 74.8°). First and second vertical tilt derivatives (Miller and Singh 1994; Thurston and Smith 1997; Verduzco et al. 2004), and Phase Preserving Dynamic Range Compression (Kovesi 2012; Holden et al. 2013) were calculated from the RTP. Each technique increases contrast between anomalies and/or act as bandpass filters.
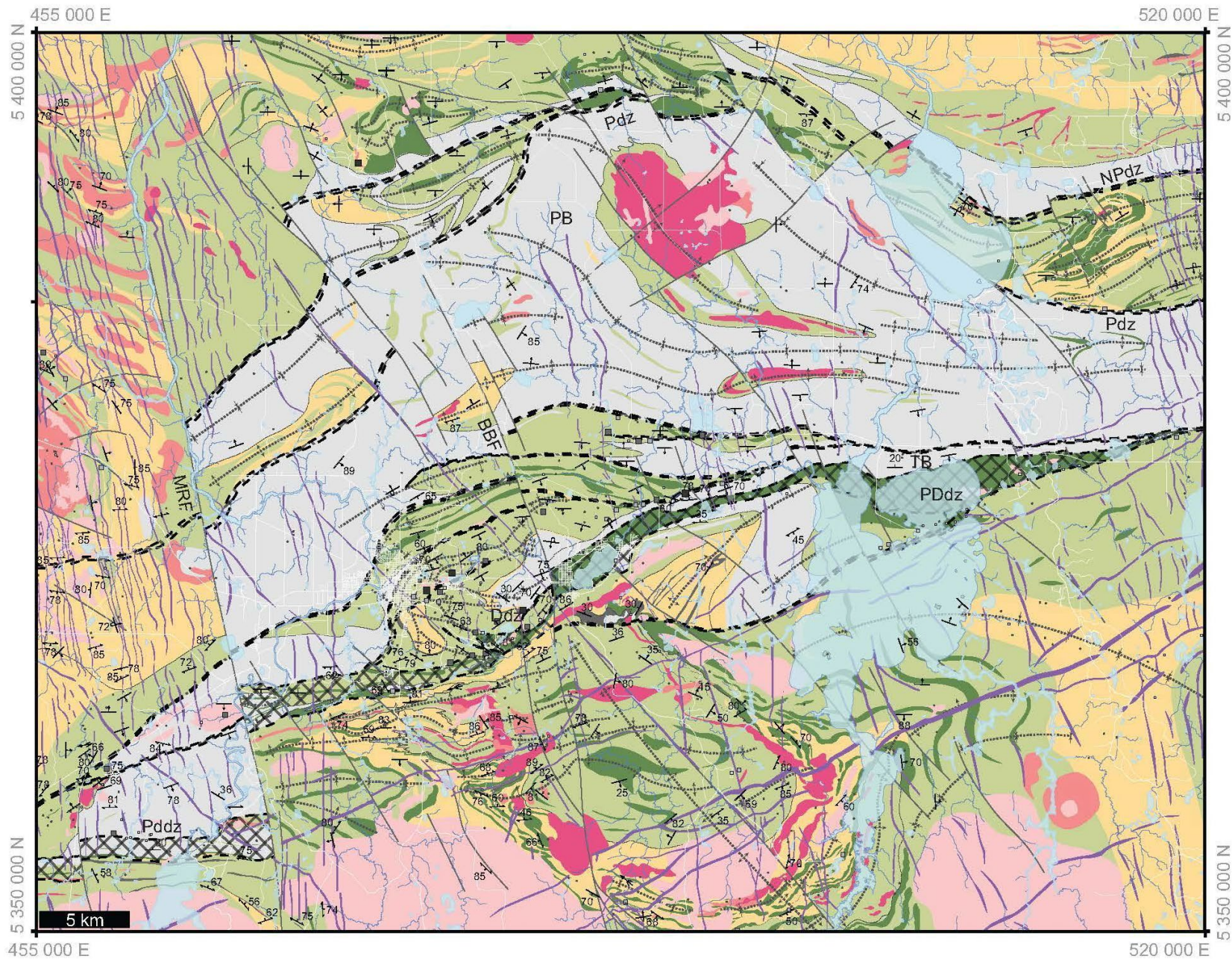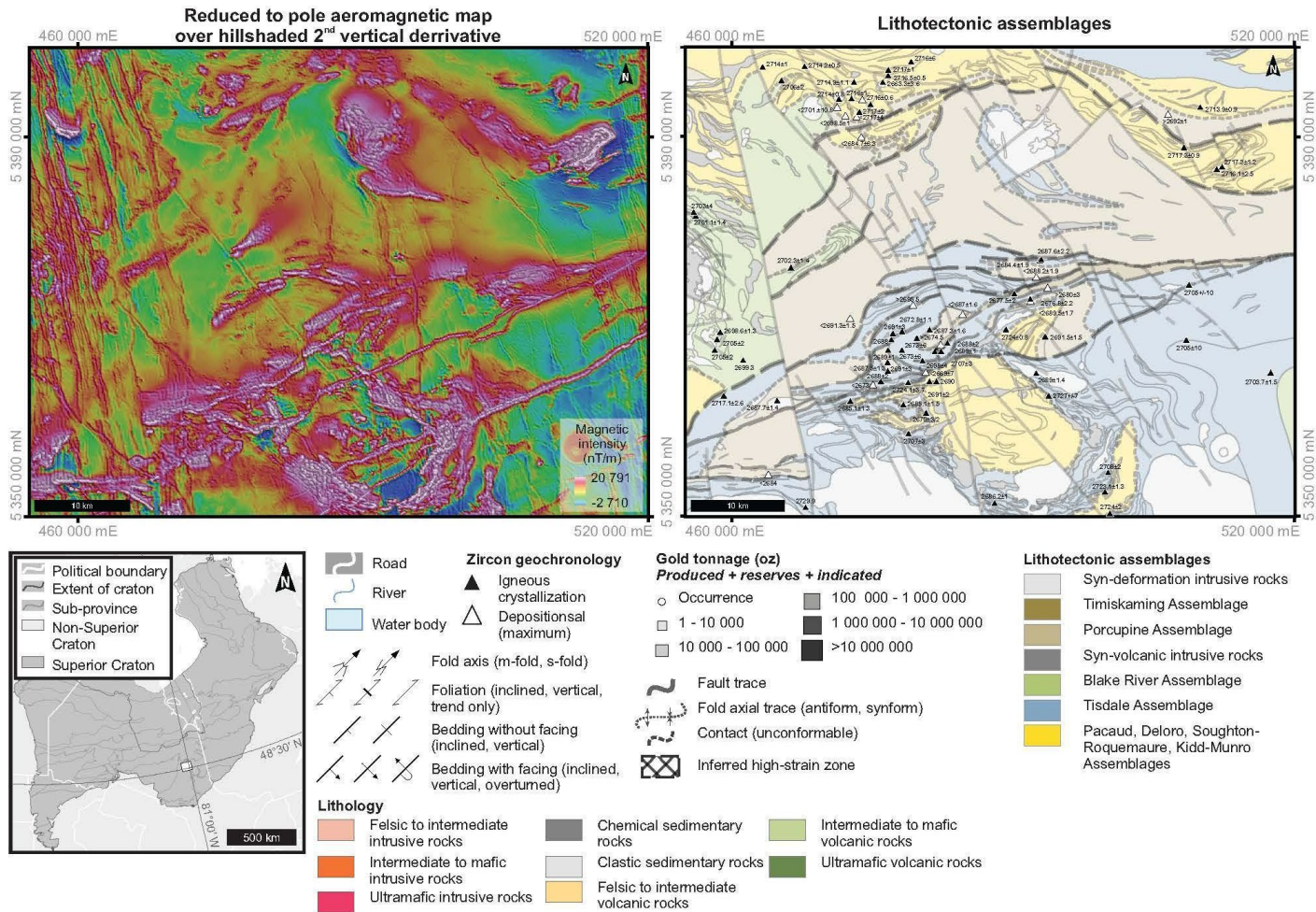
Figure 3-1. Regional geological map near Timmins, Ontario in the southern Abitibi subprovince of the Superior Province; Map displays poly-deformed Archean bimodal volcanic stratigraphy with overlying sedimentary packages variably intruded by tonalitic to granitic plutons and smaller porphyritic bodies. Proterozoic diabase dikes crosscut stratigraphy as well as intrusive bodies; Geological interpretation compiled and integrated from 33 geological maps and several compilation databases (listed in 'Data sources for constraint and compiled layers'); BBF: Burrows Benedict fault; Ddz: Dome deformation zone; MRF: Mattagami River Fault; NPdz: North Pipestone deformation zone; PB: Porcupine Basin; PDdz: Porcupine-Destor deformation zone; Pdz: Pipestone deformation zone; TB: Timiskaming Basin.
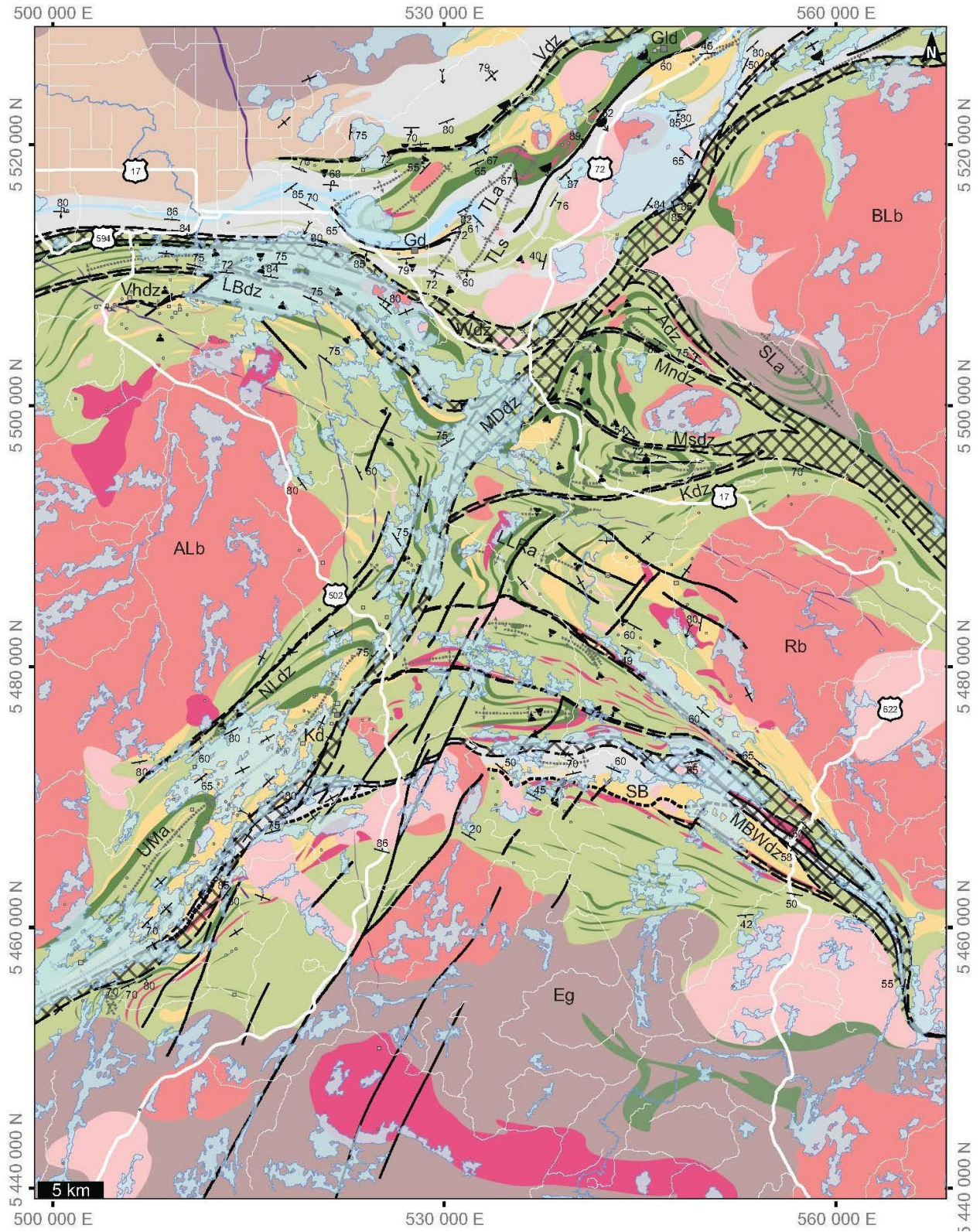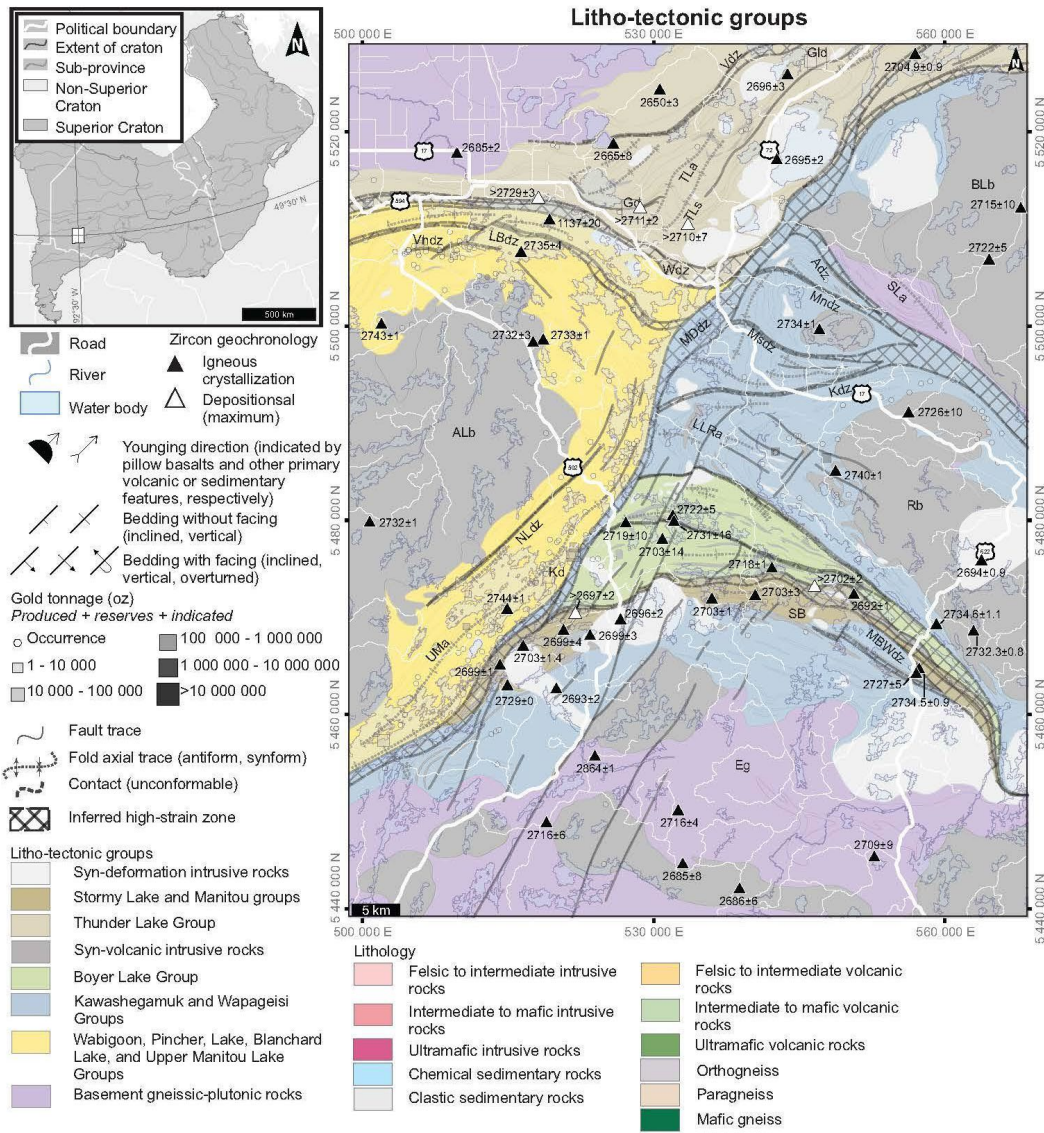
Figure 3-2: Geological map of the Dryden area, Ontario in the western Wabigoon subprovince of the Superior Province; Map displays polydeformed Archean bimodal volcanic stratigraphy with overlying sedimentary packages variably intruded by tonalitic to granitic plutons and smaller porphyritic bodies. Proterozoic diabase dikes crosscut stratigraphy as well as intrusive bodies; Geological interpretation compiled and integrated from 64 geological maps (listed in 'Data sources for constraint and compiled layers'); Coordinates in NAD 83 UTM zone 15N; Adz: Aiabewatik deformation zone; aLb: Atikwa-Lawrence batholith; BLb: Basket Lake batholith; Eg: Eltrut gneisses; Gld: Goldlund deposit; Gd: Goliath deposit; Kdz: Kawashegamuk deformation zone; Kd: Kenwest deposit; LBdz: Larson Bay deformation zone; LLRa: Long Lake River antiform; MDdz: Manitou-Dinorwic deformation zone; Mndz: Melgund north deformation zone; Msdz: Melgund south deformation zone; MBWdz: Mosher Bay-Washeibemaga deformation zone; NLdz: Noonan Lake deformation zone; Rb: Revell batholith; SB: Stormy Basin; sLa: Suzanne Lake antiform; tLa: Thunder Lake antiform; TLs: Thunder Lake synform; uMa: Upper Manitou antiform; VHd: Van Horne deposit; Vdz: Vermillion deformation zone; Wdz: Wabigoon deformation zone.

Chapter 4. Geochemical data compilation and QA/QC

Abstract

This lithogeochemical database presents compiled and new analytical results for the western Wabigoon and southern Abitibi subprovinces of the Superior Province in Ontario, Canada. These data are intended to be applied to a range of geoscientific studies that investigate broad geodynamic processes, mineral system modelling, and to aid in geologic mapping. Additionally, they provide a basis for comparative studies between the two regions, which are geologically similar but economically disparate. Several open access datasets released by the Ontario Geological Survey and the Geological Survey of Canada, as well as new analysis results from recent fieldwork were combined, integrated, and underwent quality assurance to produce a comprehensive lithogeochemical database. Data contained within this publication are co-submitted with Montsion et al. (submitted)/Chapter 5.

Keywords

Geochemical classification, Archean igneous geochemistry, Whole rock analysis, Abitibi, western Wabigoon

Value of the Data

- Lithogeochemical data representing geochemical signatures of rocks from the southern Superior Province. These are relevant to broad tectonic investigations or smaller scale studies or can be leveraged to compare the two geologically similar yet economically contrasting domains.
- Datasets are intended for use by geoscientists conducting exploration or geoscientific studies in both or either of the areas of interest.

These data may also provide a training set for machine learning techniques and mineral prospectivity.

Data access

All data-related supplementary material referenced in this chapter is available in Mendeley Data (https://data.mendeley.com/datasets/f493ctkr38) as an accompaniment with the submitted in *Data in Brief* manuscript.

Acknowledgments

4.1. Data description

The lithogeochemical data (supplementary material 5-3) are available as Excel spreadsheets separated by location (western Wabigoon and southern Abitibi subprovinces) and in a complete format where data from both areas are combined (sample sites with geological context indicated in Figure 4-1). Laboratory brochures that describe packages and methods used to generate each dataset are also included.
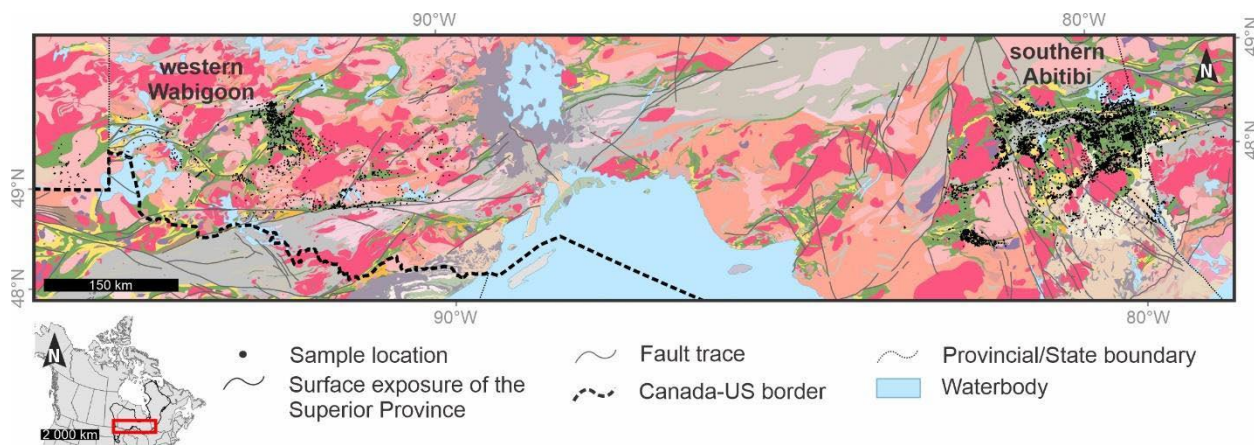
Figure 4-1: Location of samples within the Superior Province. Geology after Montsion et al. (2018) Description of map units in Montsion et al., (2018)

4.1.1. Combined columns (compiling unmodified source data and describing metadata)

Unmodified copies of geochemical data from all sources are appended in one spreadsheet 'Compiled column' along with metadata (e.g., sample ID, coordinates, geological description, element analysis results). Each element column is named according to element, lab package, and unit. For simplicity, values below or above detection limits are reported as "BD" and "AD", respectively, but detailed records of upper and lower detection limits are provided in a separate spreadsheet in the excel document. Additionally, a detailed description of other metadata and laboratory methods (i.e., laboratory, year, analysis type, lab package, preparation details, digestion methods, analytical methods, and limitations related to each component of the analytical workflow) are recorded in a separate spreadsheet in the Excel document. Cells are color-coded so that white cells contain reported information, yellow cells include inferred information from laboratory brochures, and red cells indicate missing information. Indicating inferred or problematic data is intended to provide a means of highlighting records that may be problematic for advanced interrogation or manipulation. Detection limits were recorded where available, usually indicated by "<" or as a negative value in original datasets. In some cases, where sources have already compiled several datasets, more than one detection limit is reported for a single

column. Multiple detection limits for element analysis column are especially common for old data compiled before 1990. Finally, the 'Legend and Abbreviations' tab communicates abbreviations and color codes.

4.1.2. Integrating data

Analytical results reported by all sources were integrated into a single column for each element in preparation for upload in ioGas (a widely used geochemical software) is presented in the 'IntegratingData' excel table for each area. Data from all sources are organized by common elements and the logic (selection formulae) is preserved.

4.2. Experimental Design, Materials and Methods

4.2.1. Compilations

Metadata (e.g. sample ID, coordinates, geological description, element analysis results) were copied from source files into a single Excel table ('CompiledColumns'). Sample metadata was integrated into condensed descriptive columns. Each element column was copied and renamed by element, lab package, and unit. Sample rock descriptions were converted into a uniform rock classification scheme so samples could be compared and queried (see Table 4-1 for rock classification schema).

The laboratory method (e.g., upper and lower detection limits, crushing, preparation, digestions, analytical technique) was recorded or inferred wherever possible in the Excel table ('CompiledColumns'). Where no explicit information was reported by data sources, analytical metadata was inferred from archived laboratory brochures (included in supplementary material 5-3 courtesy of the respective laboratories). Analytical metadata was used to classify results into groups and give priority during integrations.

One exception to inferring laboratory methods from brochures is the PETROCH database (Haus and Pauk, 1993), where element results were grouped by a numerical value rather than a descriptive label (Table 4-2). The following criteria were used to infer laboratory methods for each group: 1) Ensemble of elements reported in each group compared to typical Geolabs packages (information available from scan of 1989 laboratory brochure); 2) Detection limits and units compared to laboratory brochure; 3) Comparing value ranges and distribution of results in probability plots; 4) Element concentrations from different method groups compared using scatter plots (e.g., Figure 4-2). Relationships between point distribution and values in both probability plots and scatter plots suggest completeness of digestion,

precision, detection limits, and accuracy. Observations of these properties permitted the inference of the lab package used to generate data in each PETROCH group.

Table 4-1. Rock description classification scheme

Composition	Rock Type
Clastic	
Chemical	Sedimentary
Undefined	
Ultramafic	
Mafic	Volcanic/
Intermediate	Intrusive/
Felsic	Metamorphic
Undefined	
Quartz	
Carbonate	Vein
Undefined	
Sulfide	Ore
Undefined	Undefined

Table 4-2. Inference of PETROCH laboratory method groups

Group	Included elements	Analysis package*
1	Ag, As, Au, Ba, Bi, Co, Cr, Cu, Li, Ni, Pb, Sb, Zn	4-acid digest with ICP (method unknown)
2	B, Be, Co, Cu, Mo, Ni, Sc, Sr, V, W, Y	4-acid digest with ICP (method unknown)
3	Ba, Cr, Cs, Ga, Nb, Rb, Sn, Sr, Ta, Th, Y, Zr	Pressed pellet with XRF
4	Dy, Er, Eu, Gd, Ho, Lu, Pr, Sm, Tb, Tm, Yb	Fusion (type unknown) with ICP (method unknown)
5	F, Hg, Mo, Sn, U	Unknown
6	Au, Pd, Pt	Pb-Fire assay
8	Ce, La, Nd, Pb, Sc, Th, Y	Fusion (type unknown) with ICP (method unknown)

4.2.2. Integration

Results were converted into elemental concentration (ppm) from their original oxide value or from another concentration unit (e.g., wt% or ppb).

Integration of analyte columns was performed using priority logic using groups of classified analytical methods (Table 4-3). For major elements, the maximum of Group 1 (x-ray fluorescence or atomic absorption spectrometry) results for a given sample was considered to be the most representative value. If no Group 1 result was available, the maximum of any Group 2 (inductively coupled-mass spectrometry, inductively coupled-atomic emission spectrometry, fire assay, or infrared absorption/Leco) result was used. For trace and rare earth elements (REEs), the maximum of Group 2 was considered to be the most appropriate. If no Group
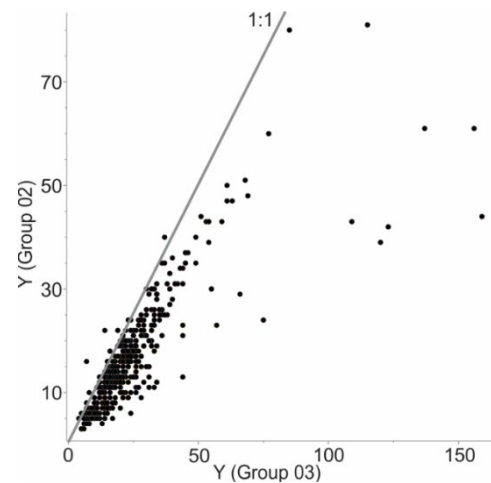
Figure 4-2. Scatter plot of Y results from PETROCH Group 3 vs Y results from PETROCH Group 2 with 1:1 line for reference. Consistently higher Group 3 results, indicated by most points sitting below the 1:1 line, suggests a more complete analytical method for Group 3. Lower values for Group 2 suggest a nearly complete digestion method.

2 value is available, the maximum Group 1 value was recorded. Finally, for metals, the maximum value of any reliable method was used.

Selection criteria for column integration was designed to select the most appropriate or accurate analysis method for each element. Appropriateness for elements was primarily determined based on detection limits vs average crustal abundance of the element as well as precision of analysis. Some digestion and analysis types were determined to be more complete or accurate than others (e.g., a pressed pellet XRF method provides higher values than a slightly incomplete digestion with an ICP analysis method; Figure 4-3). The maximum value of each group, within the logical scheme described above, is interpreted to be the most complete analysis and, therefore, the most representative of the whole rock geochemistry.

For major elements, analytical results from Group 1 methods were consistently higher than results from any other method (Figure 4-3). The higher values indicate a more complete analysis of samples and are a more accurate representation of major element concentrations when compared to other methods.

Trace and REEs have low crustal abundances that require low detection limits and high precision, especially at low concentrations. Additionally, some elements are undetectable using Group 1 methods; therefore, Group 2 methods are considered to be the most appropriate for trace and REEs.
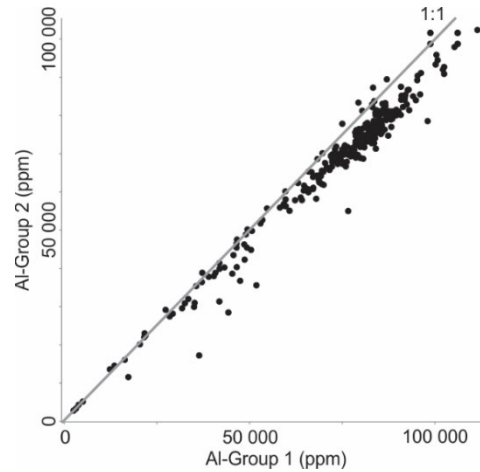
Figure 4-3. Cross plot comparing Al concentrations measured by a pressed pellet XRF method (Group 1) and lithium borate or sodium peroxide fusion ICP analysis method (Group 2) laboratory methods. Consistently higher values (nearly all sample points plot below the 1:1 line) indicate a more complete analysis of samples for Group 1. Similar plots were used to determine the best method for each element and element grouping (e.g., major, trace, REE, ore metal).

Table 4-3. Geochemical analytical method groups. AAS: Atomic Absorption Spectrometry; arICP: Aqua Regia Digestion with Inductively Coupled Plasma-Mass Spectrometry; FA: Fire Assay; IA: Infrared Absorption; ICP-AES/OES: Inductively Coupled Plasma-Atomic Emission Spectrometry; ICP-MS: Inductively Coupled Plasma-Mass Spectrometry; ISE: Ion Selective Electrode; XRF: X-Ray Fluorescence Spectroscopy

Group 1	XRF, AAS
Group 2	ICP-MS, ICP-AES/OES, FA, IA
Group 3	arICP

Most ore metals are frequently found in non-silicate minerals, meaning they are digested in most dissolution solutions and detectable by either Group 1, Group 2 and Group 3 methods. Therefore, the maximum value from any method was used in the final metal element column.

4.3. Quality assurance/Quality control (QA/QC)

Analyses were interrogated to identify quality control issues (e.g., misclassified units, normalized values, comparability of analytical batches). As this database comprises results from numerous studies and with unique characteristics (e.g., age of analysis, quality of instrument, digestion or analysis technique, preservation of metadata) and reporting styles, careful inspection and corrections were required to properly prepare the database. Some batches within a compiled source contained mislabelled units, frequently ppb mislabelled as ppm. These were identified by points plotting at least three orders of magnitude away from other data points in cross plots for an equivalent analytical method. Similarly, ~300 samples were eliminated from the database after identifying previous unknown normalization of their reported REE values (Figure 4-4).

The comparability and need for levelling of data batches was tested by using a series of box and whisker plots (e.g., Figure 4-5), cross plots, and spider diagrams (e.g., Figure 4-4). If the distribution (e.g., range, mean, median) were similar for all batches, they were deemed to be comparable and no levelling was applied. Due to the absence or prevalent ambiguity of metadata for this compilation, quantitative or rigorous determinations of quality based on standards or duplicates (e.g., Thompson and Howarth, 1973; Thompson and Howarth, 1976, Hyslop and White, 2009) were not possible. Therefore, the general strategy for this compilation work was aggressive filtering of suspect datasets/samples or samples with results outside of detection limits to preserve data quality and minimize addition of bias. Using this approach, ~367 samples were removed from the database of (22,965 samples; 1.6%). To minimize sampling bias from batches that sampled different rock compositions, each compositional group (e.g., ultramafic, mafic, intermediate, felsic) were inspected separately. Some natural variability between batches is nevertheless expected..

4.4. Summary presentation

A summary presentation is provided in supplementary material 5-2 and summarizes the workflow and results of this compilation work for an early form of the geochemical database. Since this presentation was created, new data was added and has been used to generate results for this and the co-submitted publication compilation without updating the demonstrated workflow presentation. . This presentation

is intended to add clarity and visually present various aspects of acquisition, compilation, and integration of the lithogeochemical database. These include a description of laboratory methods used to produce the lithogeochemical data, examples of quality control issues observed in the database, and the logic that guided integration.
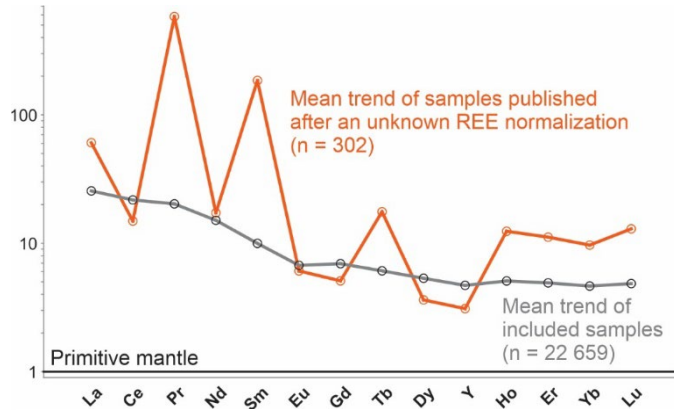
Figure 4-4. Mean trend of eliminated (orange) and included (grey) samples plotted on a REE spider diagram normalized to the primitive mantle (after Sun and McDonough, 1989). The eliminated samples display a REE pattern that suggest these samples were reported as normalized. The normalization factors being unknown, these samples were eliminated from the database.
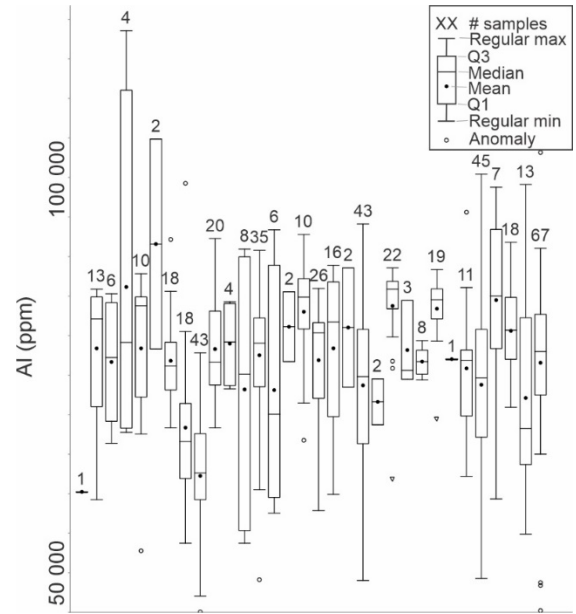
Figure 4-5. Box and whisker plots for felsic volcanic and intrusive rocks from both areas of interest. Each box and whisker along a common Al concentration y-axis represents a single batch where Al was analyzed for a felsic igneous rock. The similar range, median, and mean values from all batches indicate that these data do not require leveling. Some variability is expected as a result of natural changes in different sample populations.

Chapter 5. Knowledge gathering and exploratory analysis

Key words

Geochemical classification, Archean igneous geochemistry, Principal Component Analysis, K-Means clustering, Abitibi, western Wabigoon

Abstract

Archean greenstone belts are highly prospective for a wide range of commodities including magmatic Ni-Cu-PGE, volcanogenic massive sulfide Cu-Zn-Pb-Ag(-Au), and orogenic Au. Geochemical exploration in these settings is hampered by ancient seafloor weathering, regional metamorphism, and/or variable degrees of magmatic-hydrothermal alteration. Current classification diagrams and exploration indices for such environments are not statistically optimized and lack efficiency when applied to regionally representative geochemical databases of Archean igneous rocks. This study applies principal component analysis, k-means cluster analysis, and geological knowledge to propose a new classification scheme that utilizes least-mobile major and trace elements to discriminate between the full compositional range of igneous rocks in Archean greenstone belts. This classification scheme is applicable at low to moderate degrees of alteration/metamorphism regardless of texture or location. The scheme highlights variations in mantle/crustal sources as well as relative degrees of crustal interaction/assimilation, fractional crystallization, and accumulation. Samples proximal to magmatic, volcanogenic, and orogenic deposits in the southern Superior Province were used to define three new Igneous Rock Favorability (IRF) indices near Timmins and Dryden, Ontario, which should be applicable to other Archean cratons may ultimately lead to more efficient exploration targeting.

Acknowledgements

5.1. Introduction

Archean greenstone belts are amongst the most prospective regions for magmatic Ni-Cu-PGE, volcanogenic massive sulfide Cu-Zn-Pb-Ag(-Au), and orogenic Au (e.g., Goldfarb et al., 2001; Galley et al., 2007; Houlé and Lesher, 2011; Mercier-Langevin et al., 2012). Geochemical exploration in these settings is hampered by a lack of classification and targeting tools suitable for use with metasomatized rocks. Some existing classification diagrams (e.g., Winchester and Floyd, 1977; Cox et al., 1979; Le Maitre et al., 1989; Middlemost, 1994; Pearce 1996) assist with determining broad compositional groups (i.e., ultramafic, mafic, intermediate, and felsic) and provide a framework for comparing multiple sample populations. However, they are not ideally suited for Archean rocks for several reasons: 1) they are constrained by pristine samples from modern geological environments, which may exhibit different igneous, metamorphic, and alteration trends than Archean rocks (Condie, 1989; Mathieu, 2018); 2) many utilize too few elements to effectively distinguish between the full range of tectonic settings or prospective/non-prospective rock types (e.g., Li et al., 2005; Condie 2015); and 3) used elements that were selected using geological knowledge alone and are not statistically optimized to capture variance in data distributions.

Geochemical classification is inherently multi-dimensional, as multiple elements are required to properly characterize lithological/fertility/alteration signatures. Dimension reduction methods, such as Principal Component Analysis (PCA), are best suited to identify which input criteria provide the most discriminating power in classification applications (e.g., Vermeesch, 2006; Iwamori et al., 2017; Caté et al., 2018; Ordóñez-Calderón and Gelcich, 2018; Ueki et al., 2018).

Mineral exploration indices are typically calculated using concentrations or ratios of least-mobile chalcophile elements in magmatic systems or most-mobile lithophile/chalcophile elements in hydrothermal systems (e.g., Ishikawa Alteration Index from Ishikawa et al, 1976; Chlorite-Carbonate-Pyrite Index from Large et al., 2007). However, their application to regional-scale greenfield exploration programs is limited since most of these indices are only applicable to individual compositional groups under restricted geologic conditions (e.g., similar/low metamorphic grade; Mathieu, 2018 and references therein).

Herein, PCA and k-means cluster analysis are applied to define a statistically optimized classification scheme, and determine igneous rock favorability for magmatic, volcanogenic, and orogenic deposits hosted in Archean greenstone belts. The proposed classification scheme minimizes metasomatic changes and discriminates rock compositions in all major Archean cratons. The Igneous Rock Favorability

(IRF) indices identify rock types that are spatially associated with mineralization in the southern Superior Province; however, they are not alteration or mineralization indices (i.e., being a favorable host rock for mineralization does not mean being a mineralized rock).

5.1.1. Material

A new database of lithogeochemistry was compiled, quality checked, and interrogated for the southern Abitibi and western Wabigoon subprovinces of the Superior Province, Ontario, Canada (location of samples indicated in supplementary material Figure 5-1.1; n(Abitibi)= 17,637 and n(Wabigoon) = 4,875). These two Neoarchean greenstone belts display abundant outcrops with minimal weathering and benefited from decades of sampling and drilling, resulting in numerous open-source geochemical databases. A summary and detailed description of database compilation and workflow is given in Chapter 4 (Montsion et al., submitted a).

Additionally, samples from the eastern Yilgarn (Smithies and Lowrey 2020) and other Archean cratons (GEOROC database from Sarbas and Nohl, 2009) were used to test the applicability of the proposed classification scheme and compare results from the southern Abitibi and western Wabigoon databases. Within the GEOROC database, biased sampling of more economically significant, more accessible, or better exposed regions caused an imbalance in representation. To mitigate bias in PCA and subsequent results, only regions with >300 igneous samples were included for statistical interrogation. The retained areas include the Amazonia/Sao Francisco, Baltic, Dharwar, Kaapvaal/Zimbabwe, North Atlantic, Pilbara, Superior (excluding the Wabigoon and Abitibi subprovinces), and Yilgarn (see supplementary material 5-1.2 for geographic locations). For each of these cratons, the GEOROC database was randomly sub-sampled to 300 to not overweight any specific region in the Archean craton compilation.

5.1.2. Background

Rock classification diagrams and mineralization indices have been used in a wide range of applications. Classification diagrams such as Le Maitre et al (1989)'s Total Alkali Silica (TAS; Figure 5-4A), Pearce (1996; 2004; Figure 5-4C), Winchester and Floyd (1977 Figure 5-4E), and subsequent studies have provided a means of separating compositional groups given a small subset of discriminatory elements. The most common approach to define axes for classification diagrams uses a ratio or formula of elements according to expert knowledge of relevant chemical and mineralogical processes as well as overcoming the compositional nature of geochemical data (i.e. closure and summing to a constant/whole). Once the axes have been defined, samples are plotted and field boundaries interpreted/superimposed based on point distributions. Field boundaries vary from qualitative natural

breaks in compositional populations (e.g., Winchester and Floyd, 1977) to quantitative such as probability contours (e.g., Pearce, 1996).

Most of the established diagrams were defined using minimally altered samples from modern type localities, limiting their application in other geological settings (e.g., Archean greenstone belts, Pearce, 2008; Condie, 2015) or with metasomatized rocks (e.g., MacLean, 1990). In Figure 5-4, Archean samples from various compositional groups are coloured based on visual hand sample identification and plotted on common classification diagrams. In these diagrams, compositional groups are poorly constrained by labelled fields and overlap with one another. One of the main contributors to this issue is misidentification of hand sample rock type, which is often influenced by alteration/deformation/metamorphism obscuring primary mineralogy, variable experience of individual geologists, and/or diagram limitations including the use of mobile elements as discriminatory variables. Samples from Archean greenstone belts, where rocks have undergone seafloor alteration, polyphase deformation, and regional metamorphism ± hydrothermal alteration are difficult to classify and no diagram has been designed specifically for regional databases (i.e., containing a range of rock types and compositions) of such samples. Some diagrams proposed by Le Bas (2000) for high-Mg rocks and Moyen and Martin (2012) for Tonalite-Trondhjemite-Granodiorite are used to identify rock types within a single compositional group, but none effectively display the full range of lithological compositions at once.

5.2. Methods

Diagram creation could benefit from statistical tools like PCA, which maximize the spread amongst compositional groups by quantifying variance amongst multi-dimensional data (Pearson, 1901; Loughlin, 1991). A 'short list' of the most discriminatory elements can be defined and then geological knowledge can be applied to select the most relevant elements within this list (Step 1 and 2 in Figure 5-1), and therefore maximize efficiency and applicability of the diagrams. On this basis, a new scheme is proposed for classification of Archean igneous rocks by performing PCA using 30 Centered Log Ratio transformed least-mobile elements to select the most appropriate elements to define diagram axes. The proposed scheme is minimally influenced by alteration or texture. Field boundaries were interpreted using natural breaks between k-means clusters of southern Abitibi and western Wabigoon samples and validated by comparing independently generated clusters with the same workflow from the Yilgarn and other Archean cratons samples. Finally, new Igneous Rock Favorability (IRF) indices were developed for magmatic, volcanogenic, and orogenic deposits based on compositional similarity (measured as 5D

Euclidean distance) to mode points of samples located near known mineralization. This workflow is outlined in Figure 5-1 and detailed in subsequent sections.

Classification diagrams

01 STEP	02 STEP	03 STEP	04 STEP	05 STEP
Element selection	**PCA**	**k-Means clustering**	**Diagram design**	**Validation**
Input: Lithogeochemical database	**Input:** Least mobile element analyses (CLR)	**Input:** PCA results (for samples)	**Input A:** PCA loadings (for elements)	**Input A:** PCA loadings (Archean world)
Method: Knowledge-based selection	**Method:** Principal Component Analysis (PCA)	**Method:** k-means clustering	**Input B:** Compositional groups (clusters)	**Input B:** Compositional groups (Archean world clusters)
Output: Least mobile element list	**Output A:** PCA loadings (for elements)	**Output:** Compositional groups (clusters)	**Method:** Knowledge-based interpretation	**Method:** Visual comparison to defined fields
	Output B: PCA results (for samples)		**Output:** Classification diagrams	

Igneous rock favorability indices

06 STEP	07 STEP	08 STEP
Distance to mineralization	**Mode coordinates**	**IRF**
Input A: Igneous samples	**Input A:** Distance between each sample and nearest deposit/occurrence of each type	**Input A:** Mode value(s) for each element ratio for each deposit type
Input B: Mineral deposits	**Input B:** Mg/Zr, Ce/Y, Th/Nb, Eu/Eu*, Al/Ti	**Input B:** Mg/Zr, Ce/Y, Th/Nb, Eu/Eu*, Al/Ti
Method: Euclidean distance (on map)	**Method:** Kernel density estimate	**Method:** Euclidean distance (data space) to nearest mode coordinate
Output: Distance between each sample and nearest deposit/occurrence of each type	**Output:** Mode value(s) for each element ratio for each deposit type	**Output:** Minimum distance to mode

Figure 5-1: Flow chart of techniques and steps involved in creating the proposed classification scheme and Igneous Rock Favorability (IRF) indices for magmatic Ni-Cu-PGE, volcanogenic Cu-Zn-Pb-Ag(-Au), and orogenic Au hosted in Archean igneous terranes.

5.2.1. Element selection

To minimize the effect of metasomatism obscuring primary compositional signatures while designing new classification diagrams for Archean igneous rocks, six major (Ca, Mg, Mn, Fe, Ti, Al) and 24 least-mobile elements were selected for further investigation, including minor transition elements (Ni, Co, Sc, V, Cr), rare-earth elements (La, Ce, Pr, Nd, Sm, Eu, Gd, Tb, Dy, Ho, Y, Er, Tm, Yb, Lu), and high field strength elements (HFSE; Th, Zr, Hf, Nb). Excluded major elements (e.g., Si, Na, K, C, S) are extremely

mobile in hydrothermal systems and commonly relate to silicification, albitization or sericitization, carbonatization, and sulfidation. Trace elements that can easily substitute for the above major elements (e.g., Ba, Rb, Cs that substitute for K) are also excluded.

Because many samples from the Archean craton compilation (GEOROC database) do not contain the full suite of elements listed above, a reduced list was selected to maximize the number of samples for PCA. The reduced element list (Ca, Mg, Mn, Fe, Ti, Al, Ce, Sm, Eu, Gd, Y, Th, Zr, Nb) was used for an independent PCA and subsequent analyses.

5.2.2. Transformation and closure

Multi-variate statistical methods such as the ones employed here in (i.e., PCA) are best used on normally distributed datasets and have limited success when applied to strongly skewed data. The reality of most geochemical datasets is that the majority of samples contain background (i.e., average crustal abundance) concentrations of most elements, with a few enriched in some locally important element. Geochemical analyses are also variable in range where major elements have orders of magnitude larger ranges and/or values compared to trace elements. This disparity presents another challenge for multi-variate statistical methods. To overcome both limiting characteristics of geochemical data, transformation (e.g., centered log ratio, additive log ratio; transformations explained in Greenacre et al. (2021), normalization, and/or scaling (demonstrated in Baxter, 1995) are performed. However, geochemical analyses represent elemental abundance as a proportion of a whole (i.e., compositional data). When dealing with such data, it is important to account for appropriate relevant changes when performing transformations (termed 'closure problem'). Herein, we apply the Centered Log Ratio transform in ioGas for the selected less mobile elements.

5.2.3. Principal Component Analysis

Principal Component Analysis (PCA) is a dimension-reduction technique that captures and characterizes variance within a cloud of multi-dimensional data as principal components (Pearson, 1901; Hotelling, 1933). In geochemical applications, principal components have been used to characterize sample stoichiometries and mineral compositions (e.g., Grunsky 2010; Yu et al., 2012; Iwamori et al., 2017; Grunsky and Caritat, 2019) as they are a logical choice to integrate high dimension data (suite of analysed elements). Other, kernel-based dimension reduction methods that are more sensitive to local or subtle distributions are also applied to similar use cases involving clustering (e.g., tSNE and UMAP); however, they do not preserve the eigenvectors (ie., variable relationships) needed to define elements for the proposed diagram axes. Here, PCA was performed in Reflex ioGAS software (http://

www.reflexnow.com/) to explore relationships between each of the elements listed above (Figure 5-2A) and identify the most appropriate ratio to maximize discrimination potential. Ideally, a ratio of elements representing endmembers of each principal component would satisfy this requirement.

From the list of principal component element loadings (R-mode), elements used for diagram axes were selected using the following logic: 1) the element is routinely analyzed and is commonly recorded in most geochemical databases; 2) the element can be quantitatively analyzed with a portable X-ray fluorescence device (enhancing the utility in exploration programs); 3) the element is rarely contaminated during common preparation or analytical techniques; 4) expert knowledge indicates that the element is relevant to document igneous processes in Archean rocks.

5.2.4. K-means Clustering

K-means clustering is the simplest and most common partitional clustering algorithm (Jain et al., 1999). K-mean clustering defines isotropic clusters based on random initial centroids, where the number of centroids is k. This clustering technique iteratively assesses the distance between each datapoint and cluster center using squared error (MacQueen, 1967). When the distance between all data points and the closest centroid meets a given threshold, cluster memberships are assigned. Other clustering methods are also valid for this use case (e.g., hierarchical, DBSCAN, spectral) and may be more aptly suited to capture locally important distributions within the simplex; however, for simplicity and rapid implementation, they were not applied here.

K-means clustering was performed on igneous samples from the southern Abitibi, western Wabigoon, and other Archean cratons databases using principal component scores for each sample (Q-mode) using in-built 'Analysis' tools in Reflex's ioGAS software (http:// www.reflexnow.com/). A k = 10 was selected using elbow analysis from the delta and sum of squares error (Figure 5-2C), which evaluates proximity of all points to their assigned cluster centroid, and was used to partition the data into lithologic compositional groups. These groups were interpreted using common knowledge and compared to traditional plots (Figure 5-3). To mitigate k-means' sensitivity to seed location, 5000 tries and 15 maximum clusters were used as hyperparameters.
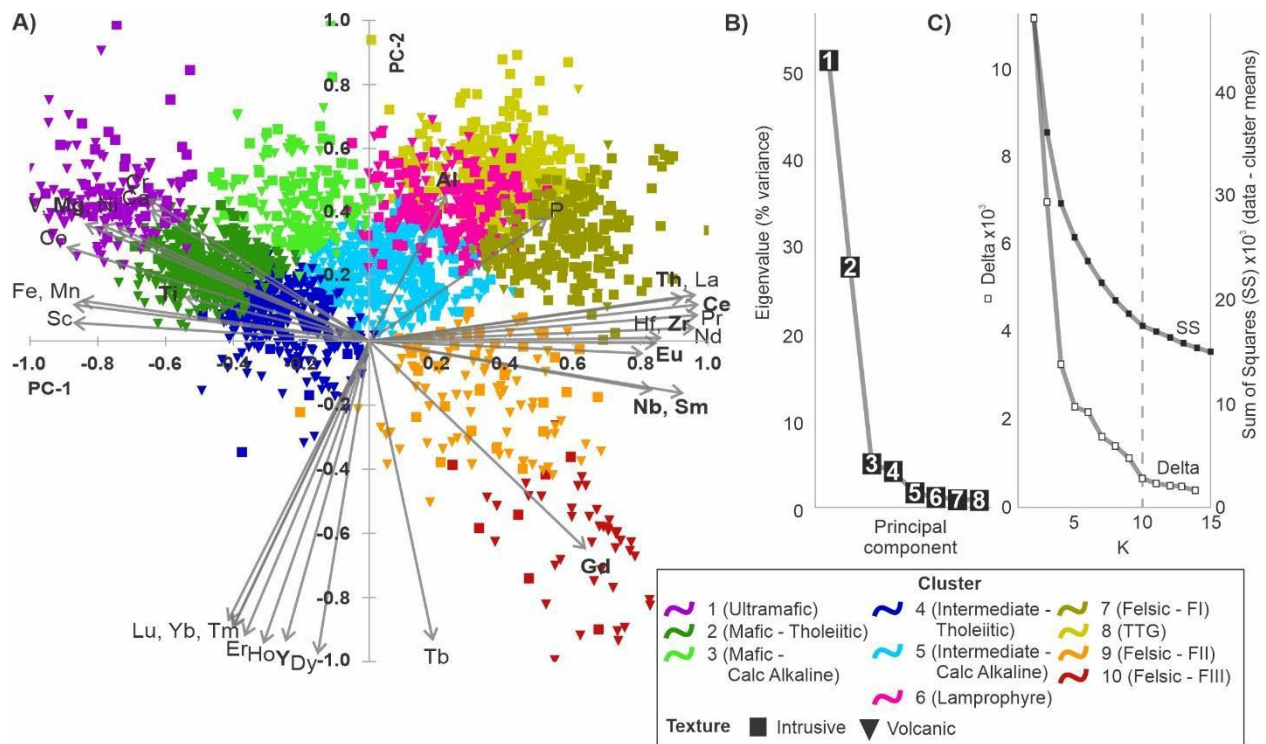
Figure 5-2. A) Scaled principal component (PC-1 vs PC-2) loadings (R-mode) statistical relationship selected element concentrations in samples from the southern Abitibi and western Wabigoon. Elements compatible in mafic minerals (Co < Fe < Sc < Mn < Mg < Ni < V < Cr < Ca) oppose incompatible elements (Gd < Eu < Nb < Hf < Zr < Sm < Th < Nd < La < Pr < Ce), separating samples according to compatibility in mafic minerals/melts along PC-1. Groups of HREEs (Dy < Ho < Y < Tb < Er < Tm < Yb < Lu) are opposed to LREEs (Sm < Nd < Pr < Ce < Th < La), HFSEs (Nb < Hf < Zr < Ti), and major elements (Al < P) along PC-2, separating samples mostly by fractionation of REEs. The PC scores for individual samples (Q-mode) are represented as points plotting on the PC-1 vs PC-2 graph and are colored according to k-means clusters outlined in section 5.3.2. Shape indicated sample texture as volcanic or intrusive. Bold element abbreviations indicate elements used to define diagram axes in the context of PC-1 vs PC-2; B) Scree plot of eigenvectors representing the percentage of variance for each principal component indicates that the first two principal components cumulatively explain >80% of variance. C) Error line plot where the 'elbow' is used to select the appropriate number of k clusters

5.2.5. Igneous Rock Favorability indices

The foundational logic behind the proposed new indices for magmatic, volcanogenic, and orogenic exploration is that samples near existing deposits/prospects/occurrences of each deposit type represent characteristic and favorable lithogeochemical signatures for their respective mineral system. The distance (geochemical similarity) to these characteristic signatures, here reported as mode coordinates, is evaluated using Euclidean distance in multi-dimensional 5D space (Mg/Zr, Ce/Y, Th/Nb, Eu/Eu*, Al/Ti). Based on the chemical similarity (5D Euclidean distance) to samples from mineralized areas, any igneous rock can be evaluated for favorability. Euclidian distance was evaluated using following formula:

$$IRF = \sqrt{\frac{\sum (\log\log x_m - \log\log x_i)^2}{d}} \qquad [1]$$

where x_m is the mode coordinate for each discriminatory dimension, x_i is the corresponding sample element ratio, and d is the number of dimensions used in the calculation. In some cases, only two or more element ratios could be calculated as some elements were not reported in the original geochemical databases.

To calculate mode values, the map distance between all igneous samples within the Dryden area of the western Wabigoon and the Timmins area of the southern Abitibi (extents indicated in supplementary material 5-1.1), and the nearest deposit/prospect/occurrence for each deposit type (compiled in Montsion et al., 2021a/Chapter 3), was calculated. Kernel Density Estimate (Rosenblatt, 1956; Parzen, 1962) 80% contours were calculated for samples within a) 2000 m, b) 1000 m, and c) 500 m of known mineralization for each deposit type (Figure 5-7).

5.3.   Results and interpretations

5.3.1. Principal Component Analysis

Principal Component (PC) element loadings (R-Mode) and sample scores (Q-mode) of igneous samples from the southern Abitibi and western Wabigoon characterize the full range (ultramafic to felsic) of compositional groups and their relationships based on geochemical processes (Figure 7-2A). The first and second principal components represent the majority (80%) of variance within the data while the remaining principal components represent very little; however, there does appear to be some, minor/subtle significance for PC 3 to PC 5 as they corroborate findings from established geochemical proxies.

PC-1 represents 52% of data variability (Figure 7-2B) and is characterized by opposing loadings of elements that are compatible in mafic minerals (Co < Fe < Sc < Mn < Mg < Ni < V < Cr < Ca) and elements that are incompatible in mafic minerals (Gd < Eu < Nb < Hf < Zr < Sm < Th < Nd < La < Pr < Ce) (Figure 7-2A; Table 5-1), which correspond to fractional crystallization processes. Although Co/Ce have the greatest discriminating power to separate ultramafic-mafic rocks from intermediate-felsic rocks according to PC loadings, Mg/Zr was selected as the best discriminant for the first axis of the new classification scheme using the criteria outlined in Section 5.2.1. To a lesser degree, LREEs (Sm < Th < Nd < La < Pr < Ce) are opposed to Heavy Rare-Earth Elements (HREEs; Lu < Tm < Yb < Er < Ho < Y < Dy < Tb)

along PC-1, following a linear relationship between ionic radius and PC-1 loadings, such that larger elements (e.g., LREEs) have higher PC-1 loadings (Table 5-1).

PC-2 represents 28% of data variability (Figure 7-2B) and is characterized by HREEs (Dy < Ho < Y < Tb < Er < Tm < Yb < Lu) opposed to LREEs (Sm < Nd < Pr < Ce < Th < La), HFSEs (Nb < Hf < Zr < Ti), and major elements (Figure 7-2A; Table 5-1), and therefore relates to source composition, partial melting, and/or fractional crystallization processes. While Dy/Al provide the most discriminating power along PC-2, Ce/Y was selected as a geologically significant alternative with comparable discrimination power along PC-2 according to criteria outlined in Section 5.2.1. Th/Nb was selected for the third axis based on subtle opposition on PC-2 and geological knowledge of crustal interaction and fractional crystallization (Pearce, 2008).

PC-3 represents 5% of data variability (Figure 7-2B) and discriminates samples based on feldspar±titanite±zircon-related elements (Ca, Ti, Al, Zr, Hf, Nb) opposing olivine±pyroxene-related elements (Ni, Cr, Mg, Co; Table 5-1) and may be related to accumulation processes. However, the low Eigenvalues for all principal components from 3 onwards may indicate very subtle geological significance or that variance is, at least in part, representative of noise rather than signal.

PC-4 and PC-5 represent 4% and 2% of the variability (Figure 7-2B), respectively, and are characterized primarily by opposed Al, Ti, Ca or Eu, which suggest partial melting or fractional crystallization processes primarily related to Ca-bearing minerals. Eu/Eu* and Al/Ti were subsequently selected from existing literature as the best discriminants for the fourth and fifth axis to identify felsic rocks based on plagioclase/titanite/apatite fractionation and to separate ultramafic rocks derived from different mantle settings (Taylor and McLennan, 1985; Sproule et al., 2002).

All other PCs are considered to be statistically insignificant as their eigenvalues are <1% (Figure 7-2B, Scree plot).

Table 5-1(next page). Scaled coordinates of principal component loadings (Q-mode) for clustered southern Abitibi and western Wabigoon igneous samples. Scaled bars, where color indicates the sign of loading values (red denotes positive and blue denotes negative) and length indicates normalized to principal component loadings, are used to highlight trends in variables. Bar length is normalized maximum and minimum loadings within each principal component. Elements are listed in order of common ionic radius in minerals and the six-fold coordination (Site), effective ionic radii (Ionic radius; Shannon, 1976), and element grouping are reported for reference. REE: Rare Earth Elements; HFSE: High Field Strength Element

Eigen vector	Charge	Site	Ionic radius	Element group	PC1	PC2	PC3	PC4	PC5
Fe	2+	VI	0.61	Major	-0.87	0.11	0.07	-0.11	-0.04
Co	2+	VI	0.65	Trace	-0.89	0.29	-0.11	-0.09	-0.05
Mn	2+	VI	0.67	Major	-0.85	0.13	0.09	-0.10	0.15
Ni	2+	VI	0.69	Trace	-0.79	0.36	-0.34	0.17	-0.12
Mg	2+	VI	0.72	Major	-0.84	0.37	-0.22	0.07	-0.04
Ca	2+	VI	1.00	Major	-0.65	0.41	0.19	-0.03	0.43
Cr	3+	VI	0.62	Trace	-0.65	0.44	-0.32	0.34	0.01
V	3+	VI	0.64	Trace	-0.74	0.36	0.13	-0.29	-0.05
Sc	3+	VI	0.75	Trace	-0.87	0.06	0.04	-0.10	0.08
Lu	3+	VI	0.86	REE	-0.42	-0.87	0.12	0.09	-0.05
Yb	3+	VI	0.87	REE	-0.40	-0.89	0.11	0.09	-0.05
Tm	3+	VI	0.88	REE	-0.40	-0.89	0.08	0.05	-0.04
Er	3+	VI	0.89	REE	-0.37	-0.92	0.06	0.04	-0.02
Y	3+	VI	0.90	REE	-0.25	-0.94	0.08	0.03	0.01
Ho	3+	VI	0.90	REE	-0.31	-0.95	0.02	-0.01	0.01
Dy	3+	VI	0.91	REE	-0.15	-0.98	-0.03	-0.05	0.05
Tb	3+	VI	0.92	REE	0.19	-0.94	-0.13	-0.13	0.12
Gd	3+	VI	0.94	REE	0.64	-0.65	-0.22	-0.20	0.17
Eu	3+	VI	0.95	REE	0.81	-0.04	-0.08	-0.33	0.32
Sm	3+	VI	0.96	REE	0.93	-0.16	-0.23	-0.16	0.12
Nd	3+	VI	0.98	REE	0.96	0.04	-0.20	-0.09	0.05
Pr	3+	VI	0.99	REE	0.97	0.08	-0.16	-0.05	0.02
Ce	3+	VI	1.01	REE	0.97	0.12	-0.12	-0.02	0.00
La	3+	VI	1.03	REE	0.97	0.15	-0.10	0.01	-0.01
Al	4+	VI	0.54	Major	0.23	0.46	0.64	0.32	0.32
Ti	4+	VI	0.67	HFSE	-0.55	0.14	0.50	-0.54	-0.14
Hf	4+	VI	0.71	HFSE	0.85	0.00	0.40	0.22	-0.07
Zr	4+	VI	0.72	HFSE	0.86	0.01	0.38	0.17	-0.10
Th	4+	VI	0.94	REE	0.93	0.14	0.01	0.16	-0.07
P	5+	VI	0.38	Major	0.53	0.39	0.11	-0.57	-0.21
Nb	5+	VI	0.64	HFSE	0.84	-0.15	0.23	0.12	-0.24

5.3.2. Cluster analysis

The results of the k-means cluster analysis of PC sample scores (Figure 7-2A) were interpreted using common cross-plots and diagrams (Figure 7-3) including primitive mantle-normalized spider diagrams (normalization values from Williams et al., 2020), MgO vs Zr, Th/Yb vs Nb/Yb basalt classification diagram (Pearce, 2008), Alkali-Iron-Magnesium (AFM) diagram, and Zr/Y vs Y rhyolite fertility diagram (Lesher et al., 1986). Sample classification from hand sample and interpreted clusters are compared in the context of typical classification diagrams (Figure 5-4). These comparisons indicate that samples identified from hand samples are poorly constrained by field boundaries, which may be the result of misidentification. Interpreted clusters generally agree with labelled fields; however, the earlier stated limitations with such diagrams remain. The advantage of using multi-variate analysis for high dimensional data (i.e., whole rock geochemistry) is highlighted in the Figure 5-4 comparison as there is

more sensitivity to subtle geochemical, and thus mineralogical, differences between a greater number of sub-populations (i.e., compositional groups).

Cluster 1 is characterized by very high abundances of Mg-Ni-Co-Cr (Figures 5-2A, 5-3B), very low abundances of Highly Incompatible Lithophile Elements (HILE; Cs, U, Th, Nb, Ta, LREE), and low Moderately Incompatible Lithophile Elements (MILE; middle-REE, Y, Zr, Hf) (Figure 5-3A), and plots within or toward the field of Mid-Ocean Ridge Basalt (Figure 5-3C). Most contain >18% MgO, a threshold that has long been used to separate ultramafic and mafic rocks (e.g., Sun and Nesbitt, 1976; Arndt et al. 1977; Nisbet et al. 1977; see also Le Bas, 2000). Based on these characteristics, samples attributed to cluster 1 are considered ultramafic rocks.

Cluster 2 is characterized by moderately high abundances of Mg-Ni-Co-Cr (Figures 5-2A, 5-3B) and moderately low abundances of HILE-MILE, a flat REE profile (Figure 5-3A), <18% MgO, and plots within or toward the Mid-Ocean Ridge Basalt field (MORB; Figure 5-3C) with minimal amount of crustal interaction and fractional crystallization. Based on these characteristics, cluster 2 is interpreted as tholeiitic mafic rocks.

Cluster 3 is similar to cluster 2 in terms of Mg-Ni-Co-Cr (Figures 5-2A, 5-3B), moderately low abundances of HILE-MILE and <18% MgO. However, characteristics of this cluster differs from cluster 2 with a negative REE trend, negative Ta-Nb and Ti anomalies (Figure 5-3A), higher Th and Zr concentrations (Figures 5-3A, 5-3C), and samples that plot above the MORB field (Figure 5-3C). Based on these characteristics, cluster 3 is interpreted as calc-alkaline mafic rocks or tholeiitic mafic rocks with significant crustal interactions and fractional crystallization.

Cluster 4 is also similar to cluster 2 with slightly lower abundances of Mg-Ni-Co-Cr (Figures 5-2A, 5-3B), moderate HILE-MILE abundances, a flat REE profile, and subtly negative Ta-Nb and Ti anomalies (Figure 3A). Cluster 4 plots within or toward the MORB field (Figure 5-3C) and sits along the tholeiitic trend towards Fe2Ototal and MgO (Figure 5-3D). Based on these characteristics, cluster 4 is interpreted as tholeiitic intermediate rocks.

Cluster 5 is similar to cluster 3 with slightly lower abundances of Mg-Ni-Co-Cr (Figures 5-2A, 5-3B), moderate HILE-MILE abundances, a negative REE trend, and negative Ta-Nb and Ti anomalies (Figure 5-3A). Cluster 5 also plots above the MORB field (Figure 5-3C) and sits along the calc-alkaline trend towards $Na_2O + K_2O$ (Figure 5-3D). Based on these characteristics, cluster 5 is interpreted as calc-alkaline

intermediate rocks or tholeiitic intermediate rocks with significant crustal interactions and fractional crystallization.

Cluster 6 is characterized by high abundances of Al-P (Figure 5-2A), high abundances of HILE, a negative REE trend, strongly negative Ta-Nb and Ti anomalies (Figure 5-3A), elevated but <18% MgO, high Zr concentrations, and plot above the MORB field and off set from other trends (Figure 5-3C). Based on these characteristics (Rock, 1977), cluster 6 is interpreted as lamprophyres and other alkaline rocks.

Cluster 7 is characterized by low concentrations of compatible elements (Mg-Ni-Co-Cr; Figure 2A), high abundances of HILEs, a strongly negative REE trend, strongly negative Ta-Nb and Ti anomalies (Figures 5-2A, 5-3A), and a slightly negative Eu anomaly (Figure 5-3A). Cluster 7 plots within the FI infertile field for felsic volcanic rocks (Figure 5-3E, Lesher et al., 1986) and is, therefore, interpreted as FI felsic rocks.

Cluster 8 is similar to cluster 7 with moderate to low abundances of compatible elements (Mg-Ni-Co-Cr; Figure 5-2A), an overall lower abundance of REE, including extremely low abundance of HREE, strongly negative Ta-Nb and Ti anomalies (Figure 5-3A), and a subtly positive Eu anomaly (Figure 5-3A), indicative of plagioclase (and/or titanite and/or apatite) mineral phase. Additionally, most of cluster 8 samples are described as intrusive. Based on these characteristics, cluster 8 is interpreted as Tonalite-Trondhjemite-Granodiorite (TTG) rocks.

Custer 9 is characterized by very low abundances in compatible elements (Mg-Ni-Co-Cr) (Figure 5-2A), high concentrations of HILE-MILE, a negative REE trend including a near flat HREE trend, strongly negative Ta-Nb and Ti anomalies (Figure 5-3A), and a moderately negative Eu anomaly (Figure 5-3A). Cluster 9 plots within the FII moderately fertile field for felsic volcanic rocks (Figure 5-3E, Lesher et al., 1986) and is, therefore, interpreted as FII felsic rocks.

Cluster 10 is similar to cluster 9 with an overall higher abundance of REE, flat HREE trend, moderately negative Ta-Nb and Ti (Figure 5-3A), and pronounced negative Eu anomaly (Figure 5-3A). Cluster 10 plots within the FIII fertile field for felsic volcanic rocks (Figure 5-3E, Lesher et al., 1986) and is interpreted as FIII felsic rocks.

5.3.3. New classification diagrams

The new classification diagrams discriminate samples across the full range of Archean igneous rock compositions at low and moderate degrees of alteration/metamorphism, regardless of texture (Figures 5-5A, 5-5D, 5-5G, 5-5J), and highlight variations in the source composition, degree of partial melting,

crustal interaction, fractional crystallization, accumulation (Figures 5-5B, 5-5E, 5-5H, 5-5K), and plagioclase/titanite/ apatite fractionation (Figure 5-5C, 5-5F, 5-5I, 5-5L).

Table 5-2. Population size of clusters from the Abitibi, Wabigoon, eastern Yilgarn, and the sub-sampled data from the other Archean cratons.

		Population size			
Cluster	Interpretation	Abitibi	Wabigoon	Yilgarn	Other Archean Cratons
1	Ultramafic	114	37	606	228
2	Mafic - Tholeiitic	106	89	797	279
3	Mafic - 'Calc Alkaline'	222	281	312	281
4	Intermediate - Tholeiitic	193	145	437	359
5	Intermediate - 'Calc Alkaline'	192	238	257	171
6	Lamprophyre	66	134	454	0
7	Felsic - FI	56	284	187	165
8	TTG	39	193		
9	Felsic - FII	67	53	129	110
10	Felsic - FIII	42	20	0	73
Total		1097	1474	3179	1666

Figure 5-3 (next page). Commonly applied interpretation diagrams are used to interpret k-means clusters. A) Average extended spider plot for each k-means cluster on primitive mantle-normalized diagram (after Williams et al., 2020); B) Zr vs MgO discriminates cluster 1 (ultramafic) from clusters 2, 3, and 6 (tholeiitic mafic, calc-alkaline mafic, and lamprophyre, respectively) using a 18% MgO threshold; C) Pearce (2008)'s Nb/Yb vs Th/Yb Basalt Classification Diagram is used to observe the relationship and trends present in the clusters used previously (1,2,3,6). This diagram highlights the compositional evolution of samples as a result of magma/crust interaction and fractional crystallization; D) Subclasses of felsic rocks (FI, FII, FIII) are defined using Lesher et al. (1986)'s Zr/Y vs Y Rhyolite Fertility Classification diagram based on plagioclase fractionation of REEs; E) Alkali-Iron-Magnesium (AFM) diagram is used to outline tholeiitic vs calc-alkaline affinity of each cluster. Two boundaries for thoeiites vs calc-alkaline rocks from Kuno (1968) and Irvine and Bragar (1971); F) Jensen Cation Plot for classifying the full compositional range of sub-alkalic volcanic rocks (Jensen, 1976). A* = Andesite, B* = Basalt, D* = Dacite, R* = Rhyolite.

A) Multi-element spider diagram (normalized to Pyrolite) with elements: Th, U, Ta, Nb, La, Ce, Pr, Nd, Sm, Hf, Zr, Ti, Eu, Gd, Tb, Dy, Ho, Y, Er, Tm, Yb, Lu, Al, Ca, Sc

B) MgO (%) vs Zr (ppm); 18% reference line

C) Th/Yb vs Nb/Yb; Crust/Magma interaction, Volcanic arc array, OIB, E-MORB, N-MORB

D) Zr/Y vs Y (ppm); FI, FII, FIII, Calc-alkaline, Transitional, Tholeiitic

E) Fe₂O_total – Na₂O + K₂O – MgO ternary; Tholeiitic, Calc-alkaline, Kuno (1968), Irvine and Baragar (1971)

F) Ti+Fe (total) Cation % – Al Cation % – Mg Cation% ternary; Tholeiitic, Fe-rich, Mg-rich, Calc-alkalic, Komatiitic, Basaltic, Ultramafic, A*, B*, D*, R*

Texture
- ■ Intrusive
- ▼ Volcanic

Cluster
- 1 (Ultramafic)
- 2 (Mafic - Tholeiitic)
- 3 (Mafic - Calc Alkaline)
- 4 (Intermediate - Tholeiitic)
- 5 (Intermediate - Calc Alkaline)
- 6 (Lamprophyre)
- 7 (Felsic - FI)
- 8 (TTG)
- 9 (Felsic - FII)
- 10 (Felsic - FIII)

Figure 5-4. A comparison of southern Abitibi and western Wabigoon samples visually classified from hand samples (A, C, E) and interpreted clusters (B, D, F) using Le Maitre (1989)'s Total Alkali Silica diagram (A, B), Pearce (2004)'s (C, D), Winchester and Floyd (1977)'s (E, F) Nb/Y vs Zr/Ti classification of volcanic rocks as a framework. Samples identified from hand sample are poorly constrained by field boundaries, which may be the result of misidentification. Clustered samples are better defined by indicated fields; however, there is room for improvement and the inclusion of lamprophyre and ultramafic rocks.

5.3.3.1. Rock classification

The Ce/Y vs Mg/Zr diagram (Figure 5-5A, 5-5D) represents a first-order rock classification (PC1 and PC2), from ultramafic to felsic rocks, and is largely influenced by the presence of mafic vs felsic minerals. The negatively sloping boundaries that separate ultramafic from mafic/intermediate are interpreted to reflect increasing Mg and decreasing/diluted Zr during partial melting of the mantle to produce ultramafic-mafic magmas and/or decreasing Mg and increasing Zr during fractional crystallization of ultramafic-mafic magmas. The vertical boundaries reflect thresholds of decreasing Ce/Y with accumulation of Ol-Opx±Cpx and increasing Ce/Y with enrichment of the mantle source. FI felsic rocks can be produced by ~90% fractional crystallization of mafic magmas (rare) or ~10% partial melting of mafic rocks (e.g., Lesher et al., 1986; Hart et al., 2004), but the steeply sloping boundaries between FI/FII/FIII felsic rocks reflect variations in the composition and depth of the source (and therefore mineralogy of the residue; e.g., Lesher et al., 1986; Hart et al. 2004).

Interestingly, samples coloured by classification from hand sample broadly agree with labelled fields in Figure 5-6, which is a distinct improvement from the same samples and symbology from traditionally used classification diagrams (Figure 5-4). In the new classification scheme, compositional groups are distributed more evenly across the diagrams (i.e., less overlapping or superimposed groups) and misclassified samples or those that plot outside of the correct labelled field in traditional diagrams sit within expected fields in the new scheme

Samples were colored based on reported rock types (i.e., visual identification by geologist or interpretation is the source publication) in Figure 5-6 to investigate if there was a systematic or notable issue with PCA and cluster analysis. While it was not possible to differentiate the felsic sub-classes using reported rock types, most compositional classes match well between statistical results and geological interpretations. Exceptions to this are rock interpretations that rely on mobile elements for identification (e.g., Na-rich TTGs versus K-rich F1 rhyolites, silicified basalt vs andesite/dacite), as silicification, albitization and/or potassic alteration may obliterate primary mineralogy and lithogeochemistry (Robert et al., 2001; Perrouty et al., 2019).

S. ABITIBI (N = 1097)

W. WABIGOON (N = 1474)

E YILGARN (N = 3179)

OTHER ARCHEAN CRATONS (N = 1666)

S. ABITIBI (N = 1097)

W. WABIGOON (N = 1474)

E YILGARN (N = 3179)

OTHER ARCHEAN CRATONS (N = 1666)

S. ABITIBI (N = 1097)

W. WABIGOON (N = 1474)

E YILGARN (N = 3179)

OTHER ARCHEAN CRATONS (N = 1666)

S. ABITIBI (N = 1097)

W. WABIGOON (N = 1474)

E YILGARN (N = 3179)

OTHER ARCHEAN CRATONS (N = 1666)

Clusters

1 - Ultramafic
2 - Mafic *Tholeiitic*
3 - Mafic *Calc Alkaline*
4 - Intermediate *Tholeiitic*
5 - Intermediate *Calc Alkaline*
6 - Lamprophyre
7 - Felsic (*FI*)
8 - TTG
9 - Felsic (*FII*)
10 - Felsic (*FIII*)

% Kernel Density Estimate (*all samples*)

10
50
90

Texture
■ Intrusive ▼ Volcanic

C*: Cumulates
D*: Depleted
E*: Enriched
CI: Crustal Interaction
FC: Fractional Crystallization
MMC: Mafic Mineral Cumulates

Figure 5-5 (previous four pages). Interpreted compositional clusters plotted on the classification scheme is used as a framework to compare samples from the southern Abitibi (A, B, C), western Wabigoon (D, E, F), eastern Yilgarn (G, H, I), and other Archean cratons (J, K, L). Shaded Kerned Density Estimate (KDE) contours indicate the density of all sample as means of assessing sampling bias for each cluster. A normalized histogram of each compositional group is indicated for each axis. 4.1) Ultramafic rocks. 4.2) mafic rocks (no lamprophyre and alkaline rock data for the "other Archean cratons"). 4.3) Intermediate rocks. 4.4) Felsic rocks and TTGs. The first order classification (A, D, G, J) discriminates samples based on ratio of compatible (Mg) to incompatible (Zr) elements in a mafic mineral/melt (i.e., PC-1) vs fractionation between LREEs (Ce) and HREEs (Y) (i.e., PC-2). A second order of classification (B, E, H, K) separates samples based on crustal interaction/fractional crystallization/depth of source (Th/Nb) based on those elements' opposition along PC-2 and geological knowledge of melt modification processes (Pearce, 2008). A third order of classification (C, F, I, L) uses elements affected by feldspar/titanite-related partial melting or fractional crystallization processes (Eu, Gd, Sm, Al, Ti) based on opposition along PC-4 and PC-5 to separate felsic subclasses and ultramafic rocks derived from differing sources.

5.3.3.2. Degree of partial melting, crustal interaction/assimilation, fractional crystallization

The Ce/Y vs Th/Nb diagram (Figure 5-5B, 5- 5E) represents a second-order classification (PC2) and highlights variations in source compositions and degrees of crustal interaction and fractional crystallization, which typically result in increasing Ce/Y and Th/Nb. For ultramafic to intermediate rocks, melt compositions are indicated by Ce/Y where 1) Ce/Y decreases with accumulation of Ol-Opx±Cpx and 2) increases with enrichment of the mantle source. For felsic and TTG rocks, Ce/Y likely reflects variations in the composition and depth of the source where increasing Ce/Y indicates deeper crustal sources for magmas (e.g., presence of mineral fractionating the HREE such as garnet).

In addition to subtle opposition on PC-2, Th/Nb leverages conclusions drawn from Pearce (2008) that report Th and Nb as proxies for crustal interaction and fractional crystallization. While Pearce (2008) limited the use of Th and Nb to mafic volcanic rocks, sample patterns on the proposed classification scheme suggest that Th/Nb can be applied to all compositional groups (ultramafic to felsic) of Archean age. The origin of melts differs between mafic (mantle source) and felsic rocks (mantle or crustal source); however, the crustal interaction and fractional crystallization processes/trends are broadly similar: e.g., an igneous suite may evolve from depleted mantle-sourced ultramafic rocks (cluster 1) to mafic to intermediate rocks (cluster 2 and 3). Within each compositional cluster, similar linear trends

with steeper slopes compared to the global trend indicate influence from the same geologic processes. The slope and trajectory of these trends may provide insight into tectonic processes governing melt formation and evolution in the Archean crust, which likely differs from one locality to another (e.g., eastern Yilgarn vs southern Abitibi and western Wabigoon subprovinces in Figure 5-5).

5.3.3.3. Plagioclase/Titanite/Apatite Fractionation

Figure 5-5C and 5-5F (Eu/Eu* vs Al/Ti) represent a third-order classification based on compositional variability relating to the formation of Ca-bearing minerals (Eu/Eu* and Al/Ti). Eu-anomalies (Eu/Eu*) are frequently attributed to Eu2+ tendency to substitute for Ca2+, such as in plagioclase (Drake and Weill, 1975; Wilke and Behrens, 1999), titanite, or apatite (Fleet and Pan, 1995). The remaining melt fraction displays negative Eu-anomalies (Eu/Eu* <1; Taylor and McLennan, 1985). A vertical threshold at 1 Eu/Eu* marks the boundary between positive and negative Eu-anomalies. Similarly, Al/Ti is a proxy used to document partial melting and source-related processes (e.g., in ultramafic rocks, Sproule et al., 2002) or plagioclase/titanite/apatite accumulation. Three ultramafic populations were identified using southern Abitibi, western Wabigoon, and other Archean craton samples (supplementary material 5-1.5). The other mafic and intermediate rocks do not display distinct clusters or trends. FIII and, to a lesser extent FII felsic rocks, show a pronounced negative Eu anomaly, which is commonly interpreted as plagioclase, sometimes titanite/apatite, fractionation (Lesher et al., 1986). The southern Abitibi and western Wabigoon felsic samples show that elevated Al/Ti ratios are associated with stronger Eu-anomalies and contrast with the other Archean cratons' samples where the FIII and FII distribution is more diffuse, but still present.
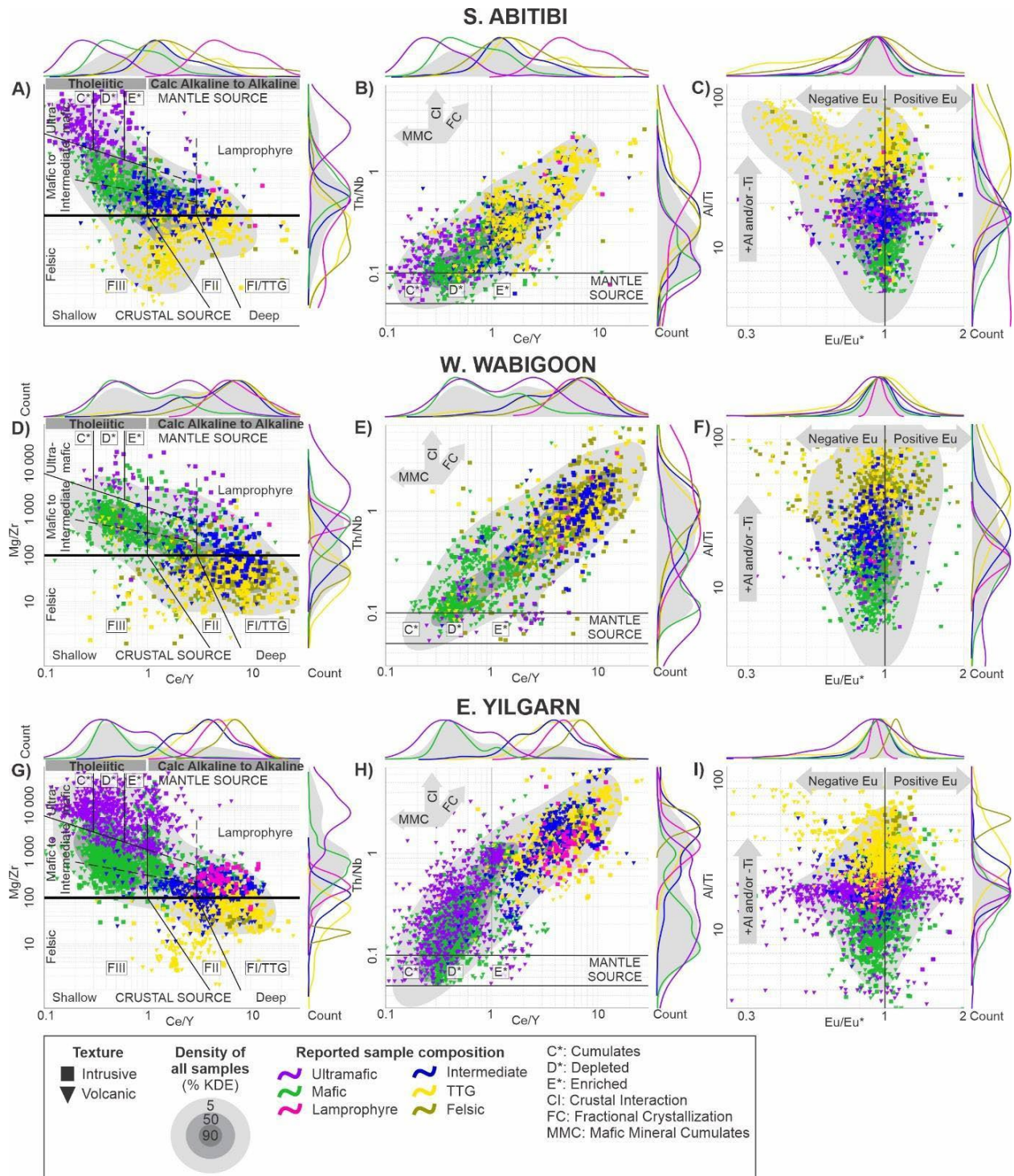
Figure 5-6. A comparison of samples visually identified from hand samples collected in the southern Abitibi (A, B, C), western Wabigoon (D, E, F), and eastern Yilgarn (G, H, I) in the context of the proposed classification scheme indicates that while some samples may be misidentified, the vast majority are well defined by the proposed compositional fields. A normalized histogram of each compositional group is indicated for each axis.

5.3.4. Igneous Rock Favorability indices

5.3.4.1. Kernel Density Estimates of mineralized samples

Kernel Density Estimates (KDE) are a non-parametric estimation of Probability Density Functions (PDF). that represents the shape of underlying data without assuming any specific distribution. KDE curves and contours are similar to histograms but provide a more accurate representation of the underlying distribution while minimizing noise. However, caution should be used for data with outliers as tail density can be overestimated (Silverman, 2986; Scott, 1992, Waskom, 2022).

The distribution and mode values from 80% KDE contours of samples within 2000 m, 1000 m, and 500 m of known deposits/prospects/occurrences characterize favorable compositions for each deposit type based on proximity to mineralization (Figure 5-7). Samples >2000 m from mineralization are assumed to be geologically unrelated and are not considered. Mode values for each element ratio are reported in Table 5-3. Samples proximal to known magmatic Ni-Cu-PGE mineralization display a unimodal distribution, therefore, only one set of mode coordinates is reported. Samples proximal to volcanogenic Cu-Zn-Pb-Ag(-Au) and orogenic Au display a bimodal distribution and, thus, two sets of mode coordinates are reported.

Table 5-3. Mode values and extent of the 80% Kernel Density Estimate (KDE) contour of samples within 2000 m of known magmatic Ni-Cu-PGE, volcanogenic, and orogenic Au mineralization. Values indicate the most favorable rock composition using axes (element ratios) from the proposed classification scheme.

| Deposit type | Mode | Mode values | | | | | KDE density contours | | | | |
		Mg/Zr	Ce/Y	Th/Nb	Eu/Eu*	Al/Ti	Mg/Zr	Ce/Y	Th/Nb	Eu/Eu*	Al/Ti
Magmatic		5900	0.3	0.2	0.95	17	700-30000	0.09-0.50	0.1-0.4	0.85-1.2	18-20
Volcanogenic	i	640	0.5	0.13	0.95	17	200-1000	0.3-0.7	0.1-0.2	0.8-1.1	8-25
	ii	21	1.3	0.29	0.4	54	7-50	0.9-1.0	1-2	0.3-0.5	40-80
Orogenic	i	800	0.45	0.11	0.9	14	400-1500	0.3-0.5	0.08-0.16	0.8-1.1	11-21
	ii	52	6.5	0.84	0.95	45	20-300	1-8	0.28-0.90	0.9-1.1	36-60

5.3.4.1.1. Magmatic Ni-Cu-PGE

KDE contours for volcanic and intrusive rock samples proximal to magmatic Ni-Cu-PGE occurrences/deposits are tightly clustered with a uni-modal distribution. Based on results from Figure 5-7 (A,B,C), samples most proximal (500 m) to Ni-Cu-PGE prospects/deposits in the Timmins region (southern Abitibi) are ultramafic (Q1 = 16 wt%; Q3 = 39 wt% MgO), appear to have been mainly derived from a cumulate-rich/depleted mantle source (mode = 0.3 Ce/Y), and have interacted with some crustal material (mode = 0.2 Th/Nb). According to supplementary material 5-1.5, these samples are mainly derived from an Al-undepleted magma (mode = 17 Al/Ti). No samples with complete geochemical

analyses were reported within 2000 m of the few known occurrences in the Dryden area (western Wabigoon), however, rocks nearby these rare occurrences were documented as ultramafic cumulates within mafic intrusions.

Figure 5-7 (next page). Characteristic favorability for each mineral system (magmatic, volcanogenic, orogenic) is evaluated using mode values and the 80% Kernel Density Estimate (KDE) along each of the proposed classification scheme's 5 element ratio axes for igneous samples within 2000, 1000, and 500m of known deposits/prospects/occurrences. Both volcanic and intrusive samples were used to define favorability in magmatic (A, B, C) and volcanogenic (D, E, F) as their host rocks are often genetically related to mineralization. Orogenic (G, H, I) mineralization can occur in any metamorphic protolith but are favorable in areas with significant rheological/competency/chemical contrast and are often spatially associated with syn-volcanic intrusions. Therefore, only intrusive samples were used to define orogenic favorability. The two dominant rock types related to orogenic Au (mafic and TTG intrusion) are disproportionately represented in the database. To preserve compositional information of both, mafic samples were statistically examined (i.e., mode and KDE contour) separately and indicated by a dashed line. Favorable intrusions were further investigated by separately examining the two rock types for the southern Abitibi and western Wabigoon (J, K, L).
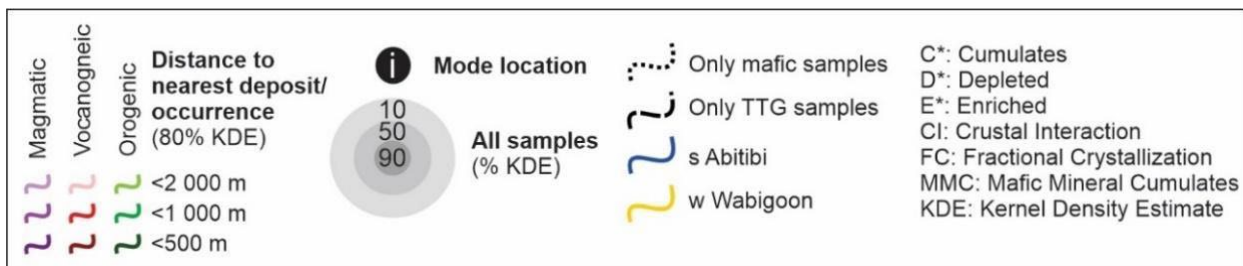
Magmatic Ni-Cu-PGE (volcanic + intrusive rocks)
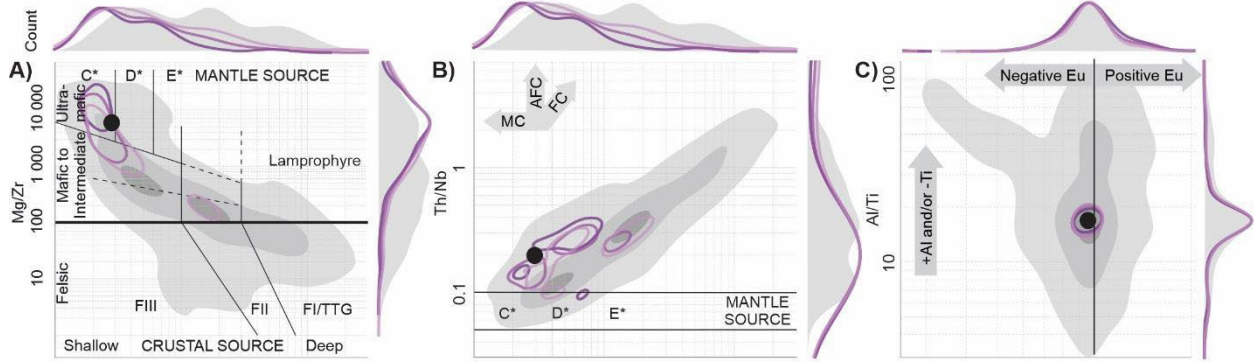
Volcanogneic Cu-Zn-Ag±Au (volcanic + intrusive rocks)
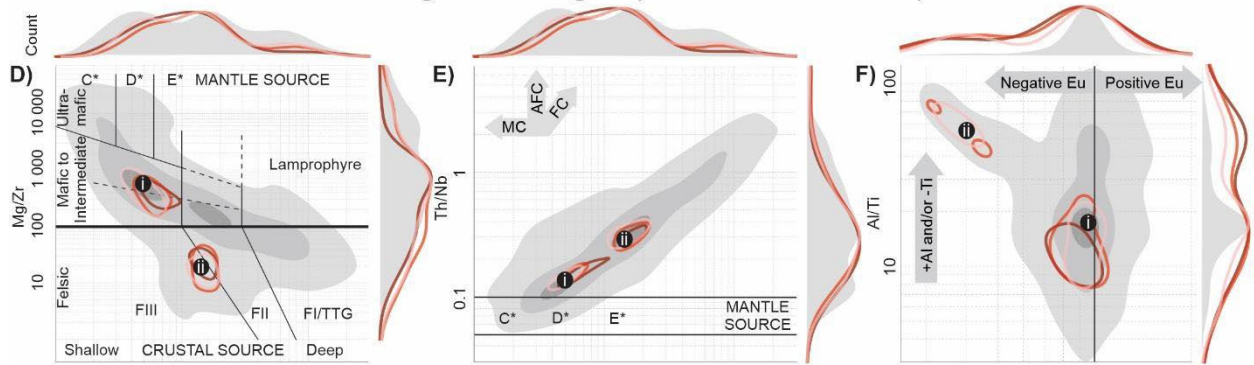
Orogenic Au (intrusive rocks)
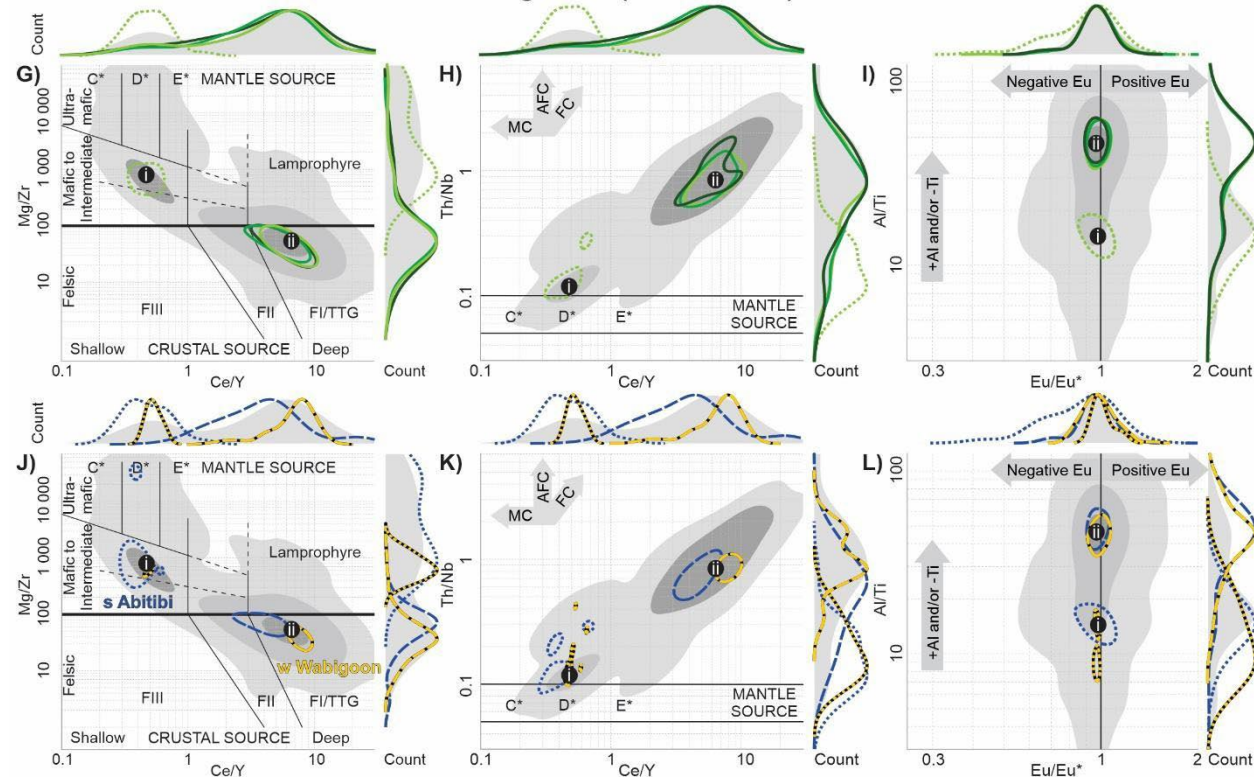
5.3.4.1.2.   Volcanogenic Cu-Zn-Pb-Ag(-Au)

KDE contours of samples within 2000 m of known volcanogenic mineralization in the Timmins and Dryden areas display a bi-modal distribution with modes centered on tholeiitic mafic and FIII felsic clusters. The mafic mode ('i' in Figure 5-7 G,H,I) is characterized by a moderate Mg/Zr value (640; as expected for a mafic rock), a low Ce/Y value (0.5; indicating a flat REE pattern), a very low Th/Nb value (0.13; indicating very low crustal interaction and fractional crystallization and eruption through a thin crust). While Pearce (2012) reports that mafic rocks associated with Archean volcanogenic systems have elevated TiO_2 concentrations, there appears to be no trend or difference amongst Al/Ti and Eu anomaly values. Al/Ti, especially, is highly variable (Q1 = 9.1; Q3 = 19.4). Conversely, the FIII felsic mode ('ii' in Figure 5-7 G,H,I) is characterized by low Mg/Zr (21; as expected for a felsic rock), low Ce/Y (1.3; indicating a flat REE curve), the lowest Th/Nb values for felsic rock (0.29; indicating shallow source) as well as a very low Eu anomaly (0.4) and high Al/Th value (54; indicating strong plagioclase fractionation).

5.3.4.1.3.   Orogenic Au

The KDE contours and mode values for orogenic Au in Figure 5-7 (G,H,I) confirm that there is a spatial association with gabbroic and TTG intrusions; however, the index's effectiveness is limited by only 3/5 element ratios being discriminatory. The favorable mafic population (mode 'i' in Figure 5-7 G,H,I) has relatively low Mg/Zr (800), appears to be a near depleted mantle source (Ce/Y = 0.45) that underwent minimal crustal interaction and fractional crystallization (Th/Nb = 0.11), and typically represents syn-volcanic feeder dikes. The intermediate/felsic population (mode 'ii' in Figure 5-7 G,H,I) is classified as FI felsic or TTG rocks with moderate Mg/Zr values (52), was derived from either weak partial melting of a metasomatized mantle, partial melting of the base of an intermediate crust, or for the common TTG suites, from partial melting of a mafic crustal source (as proposed by Moyen and Martin, 2012). In any case, this population was sourced from a deep setting and experienced significant crustal interaction and fractional crystallization (Ce/Y = 6.5; Th/Nb = 0.84). It should be noted that there was a disparity between mafic and FI/TTG samples near mineralization. Therefore, KDE contouring was performed independently for each compositional group in Figure 5-7 G,H,I.

Further associations may be present; however, sampling biases can be noted: in the northern Timmins region (4 in Figure 5-7D), and, in the Dryden's Upper Manitou (6 in Figure 5-8D) and Van Horne (7 in Figure 5-8D) areas, where no intrusive samples with the needed elemental analyses were available. Small intermediate to felsic dikes are also commonly under sampled yet possibly prospective for orogenic Au mineralization. Furthermore, sampling intrusions is not systematic in most exploration

programs. The disparity between abundant favorable intrusions in Timmins compared to Dryden may be the result of this bias, as Timmins has benefited from >100 years of exploration and extensive drilling while Dryden has largely been neglected since Timmins was discovered.

5.3.4.2. Application to the Timmins and Dryden areas

Favorable rocks, identified by thresholds along probability density functions of igneous rock favorability values, are highlighted as large red and yellow points in Figures 5-8 and 5-9 for the Timmins and Dryden map areas, respectively. In general, Timmins contains more favorable rocks than Dryden, regardless of deposit type. Note that the presence of a favorable rock type does not indicate mineralization, rather it represents the chemical similarity between a sample and host rocks commonly found nearby known mineralization. Mineralization in any of the three investigated systems are the result of favorable geochemical, structural, and hydrothermal factors.

5.3.4.2.1.   Magmatic Ni-Cu-PGE

In the Timmins area, favorable host rocks for magmatic Ni-Cu-PGE mineralization are spatially associated with ultramafic volcanic and/or intrusive rocks near deformation zones or unconformities. For example, near the Alexo mine north of the Pipestone fault (1 in Figure 5-8B) as well as along the southern margin of the Shaw dome near the abandoned Langmuir and McWaters mines (2 in Figure 5-8B). These favorable rocks are only observed in higher level/younger volcanic assemblages including the Kidd-Munro and Tisdale (see also Sproule et al., 2002; Houlé and Lesher, 2011).

In contrast, no primary (red) nor secondary (yellow) rocks that are favorable for Ni-Cu-PGE mineralization were identified in the Dryden area (Figure 5-9B), despite a few historical discretionary occurrences, as defined by the Ontario Geological Survey (OGS, 2018), being reported along the northern margin of the Atikwa-Lawrence batholith.

5.3.4.2.2.   Volcanogenic Cu-Zn-Pb-Ag(-Au)

Several favorable host rocks in the Timmins area were identified near the active Kidd Creek mine (1 in Figure 5-8C), near the former Kam Kotia mine (3 in Figure 5-8C), and in the weakly mineralized Dundonald area (4 in Figure 5-8C), but also in non-mineralized areas within the Tisdale assemblage near the city of Timmins (2 in Figure 5-8C) and along the southern margin of the Shaw dome (5 in Figure 5-8C). In mineralized zones, favorable rocks from both cluster 2 (mafic) and cluster 10 (FIII) are present, within the Kidd-Munro/Blake River volcanic assemblages. In non-mineralized zone, favorable rocks from

only one cluster are present, within the Kidd-Munro/Tisdale/Blake River volcanic assemblages. No favorable rocks were identified in the Deloro assemblage.

In the Dryden area, only two favorable rocks were identified in Figure 5-9C. The most favorable one, located near the Goliath prospect, is associated with the Thunder Lake Group volcano-sedimentary stratigraphy north of the Wabigoon deformation zone and west of the Manitou-Dinorwic deformation zone (1 in Figure 5-9C). The second one is observed in the Wapageisi volcanic assemblage south of the Stormy Basin (2 in Figure 5-9C). Moderately favorable rocks (yellow points) are randomly distributed across the map area in all lithostratigraphic volcanic assemblages and/or intrusions. Additionally, a new linear zone of favorability was identified along the northern margin of the Vermillion deformation zone in Thunder Lake mafic volcanic rocks (3 in Figure 5-9C).

When validating results, favorable igneous rock samples were spatially compared to areas with training points (known volcanogenic mineralization) and field observations from 2018-2021. In Timmins, observations were limited by thick glacial overburden and restricted access to deposit sites. In Dryden, the newly identified linear zone of favorability was confirmed in the field and coincides with previously undocumented sulfide-rich horizons.

5.3.4.2.3.   Orogenic Au (spatially associated with intrusions)

Three regions of favorable intrusive rocks for orogenic Au mineralization can be documented in the Timmins area. They are located in the core of the Timmins mining camp (1 in Figure 5-8D), along the margin of the igneous complex near Kamiskotia Lake (2 in Figure 5-8D), and along the southern margin of the Shaw dome (3 in Figure 5-8D). In addition, a handful of isolated favorable igneous rock samples are associated with late granitoids, especially in the NE quadrant of the map. Favorable intrusions are present in all assemblages, indicating that favorability for orogenic Au is not related to specific supracrustal units. Note that this favorability index only applies for orogenic Au systems that are spatially associated with intrusions.

In the Dryden area, three favorable samples were identified, two of which are located in small intrusive bodies affected by nearby ductile deformation zones, the third sample is located within the Eltrut gneisses in Figure 5-9D. Some moderately favorable rocks (yellow points) are observed near Au occurrences at the periphery of Revell batholith (Parker, 1989; 1 in Figure 5-9D), at the Katisha prospect near the Thundercloud porphyry (2 in Figure 5-9D), and in sheared margins of intrusions along the Wabigoon (3 in Figure 5-9D), Manitou-Dinorwic (4 in Figure 5-9D), and Vermillion (5 in Figure 5-9D)

deformation zones. A sampling bias should also be noted near the Upper Manitou (6 in Figure 5-9D) and Van Horne (7 in Figure 5-9D) orogenic gold prospects. Lithogeochemical data of intrusive samples for these historical mines are not present in existing open-access databases and/or do not contain samples with relevant elemental analyses.

5.4. Discussion

5.4.1. Applications to compare Precambrian regions

The new 5D classification scheme provides a comprehensive framework for comparing and discussing samples from several locations in the context of geodynamic setting, melt source, crustal interaction, fractional crystallization, and mineral system fertility. Samples from greenstone belts in the southern Abitibi, western Wabigoon, eastern Yilgarn, and other Archean cratons display equivalent patterns for most of the compositional range. Despite variable abundance of key rock types (e.g., komatiites, FIII rhyolites), which may reflect slightly different geological environments, the overall similarity in geochemical distribution indicates that the proposed classification scheme is broadly applicable to classify and interpret Archean igneous rocks. Only differences across the four datasets and their implications for metallogeny will be discussed hereafter.

5.4.1.1. Southern Abitibi vs Wabigoon

First order differences between the two regions of the southern Superior Province include 1) Significantly less ultramafic rocks reported/sampled in the western Wabigoon; 2) An abundance of FIII felsic rocks in the southern Abitibi and general absence in the western Wabigoon; 3) An abundance of FI felsic in the western Wabigoon and general absence in the southern Abitibi.

Ultramafic rocks provide a means of comparing plume-related melt characteristics (e.g., depth, temperature, crustal interaction, fractional crystallization). Notable differences include higher abundance and higher average MgO ($\bar{x}$ = 24 wt% MgO) as well as less crustal interaction/fractional crystallization ($\bar{x}$ = 0.3 Ce/Y) in the southern Abitibi compared to western Wabigoon ($\bar{x}$ = 21 wt% MgO; $\bar{x}$ = 0.5 Ce/Y; Figure 5-5). Ultramafic rocks from both regions are dominated by Al – undepleted signatures (supplementary material 5-1.5), with minor occurrence of Ti-depleted and Al-depleted/Ti-enriched flavors in the southern Abitibi (Sproule et al., 2002). The rarity, textural, and geochemical differences of ultramafic rocks between the two regions illustrates different geodynamic settings: high temperature and thin crust in the southern Abitibi versus lower temperature (less partial melting) and likely thicker crust (i.e., more crustal interaction and fractional crystallization) in the western Wabigoon.

Tholeiitic mafic rocks are abundant in both regions and outnumber those with "calc-alkaline" signatures (66% in southern Abitibi and 76% in western Wabigoon; Figure 5-5). Mafic rocks from the southern Abitibi show a continuous trend from tholeiitic to "calc-alkaline" compositions, derived from a depleted mantle source ($\bar{x}$ = 0.4 Ce/Y) and experienced variable amounts of crustal interaction/fractional crystallization (up to $\bar{x}$ = 3 Ce/Y; $\bar{x}$ = 0.4 Th/Nb). These observations suggest that "calc-alkaline" signatures in the southern Abitibi does not necessarily reflect the nature of the melt source but instead are the result of a continuous significant interaction between a parental tholeiitic magma with existing (mafic) crust. Conversely, the western Wabigoon hosts two populations of mafic rocks. First, a tholeiitic trend derived from a depleted mantle source ($\bar{x}$ = 0.4 Ce/Y) with a range of crustal interaction and fractional crystallization ($\bar{x}$ = 0.3 Ce/Y; $\bar{x}$ = 0.1 Th/Nb) and second, a trend evolving from a more enriched mantle source ($\bar{x}$ = 1.5 Ce/Y; $\bar{x}$ = 0.3 Th/Nb) and experiencing variable amount of crustal interaction and fractional crystallization (up to $\bar{x}$ = 4 Ce/Y; $\bar{x}$ = 1 Th/Nb). Similar trends can be documented on intermediate rocks (Figure 5-5), which are primarily derived from fractional crystallization of a mafic parental magma.

Lamprophyre dikes and other alkaline rocks are relatively common in the southern Superior Province. They form a very distinct ultramafic/mafic population grouped around high Ce/Y (4.9), high Th/Nb (1.4) and average Eu/Eu* (0.9) values (Figure 5-5). These rocks originate from moderate-low degree partial melting of a metasomatized (i.e., volatile-rich) enriched mantle source, experience significant crustal interaction and fractional crystallization, and are temporally (younger) and genetically (different source composition) unrelated to the ultramafic and mafic rocks documented above.

Figure 5-8 (next page). A comparison of favorable samples in the Timmins area of the southern Abitibi. A geologic map and assemblage map were provided for reference when examining Igneous Rock Favorability (IRF) results for magmatic (B), volcanogenic (C), and orogenic (D) mineralization. IRF values for each system are indicated by colored sample location points. Large red and yellow points indicate samples above high favorability thresholds defined using inflections of a probability density function. Locations of known mineralization (i.e., deposits/prospects/occurrences) are indicated as black squares in each map. Map coordinates given in NAD 83 UTM zone 17N; BBF: Burrows Benedict fault; Ddz: Dome deformation zone; MRF: Mattagami River Fault; NPdz: North Pipestone deformation zone; PB: Porcupine Basin; PDdz: Porcupine-Destor deformation zone; Pdz: Pipestone deformation zone; TB: Timiskaming Basin. Locations of notable results discussed in text are indicated by circled numbers.

Legend:

- Mineralization
- Road
- River
- Water body
- Fault trace
- Contact
- Inferred high-strain zone

Lithology
- Felsic to intermediate intrusive rocks
- Intermediate to mafic intrusive rocks
- Ultramafic intrusive rocks
- Chemical sedimentary rocks
- Clastic sedimentary rocks
- Felsic to intermediate volcanic rocks
- Intermediate to mafic volcanic rocks
- Ultramafic volcanic rocks

Scaled proximity (Favorability)	VMS Cu-Zn±Au	Magmatic Ni-Cu-PGE	Orogenic Au
(red)	< 0.03	< 0.02	< 0.05
(yellow)	0.03 - 0.05	0.02 - 0.04	0.05 - 0.10
(light green)	0.05 - 0.10	0.04 - 0.08	0.10 - 0.20
(light blue)	0.10 - 0.20	0.08 - 0.16	0.20 - 0.40
(dark blue)	> 0.20	> 0.16	> 0.40

Lithotectonic assemblages
- Syn-deformation intrusive rocks
- Timiskaming Assemblage (<2679-2669 Ma)
- Porcupine Assemblage (<2690-2695 Ma)
- Syn-volcanic intrusive rocks
- Blake River Assemblage (2704-2695 Ma)
- Tisdale Assemblage (2710-2704 Ma)
- Kidd-Munro Assemblage (2719-2711 Ma)
- Stoughton-Roquemaure Assemblage (2723-2720 Ma)
- Deloro Assemblage (2734-2724 Ma)

A)

B) Magmatic Ni-Cu-PGE

C) Volcanogenic Cu-Zn-Ag±Au

D) Orogenic Au (intrusions only)

E)

	Mineralization		Fault
	Road		Contact
	River		Inferred high-strain zone
	Water body		

Lithology
- Felsic to intermediate intrusive
- Intermediate to mafic intrusive
- Ultramafic intrusive
- Chemical sedimentary
- Clastic sedimentary
- Felsic to intermediate volcanic
- Intermediate to mafic volcanic
- Magnetic mafic volcanic
- Orthogneiss
- Paragneiss
- Mafic gneiss

Litho-tectonic group
- Syn-deformation intrusion
- Stormy Lake and Manitou
- Thunder Lake
- Syn-volcanic intrusion
- Boyer Lake
- Kawashegamuk and Wapageisi
- Wabigoon, Pincher, Lake, Blanchard Lake, and Upper Manitou Lake
- Basement gneissic-plutonic

Scaled proximity (Favorability)

	VMS Cu-Zn±Au	Magmatic Ni-Cu-PGE	Orogenic Au
(red)	< 0.03	< 0.02	< 0.05
(yellow)	0.03 - 0.05	0.02 - 0.04	0.05 - 0.10
(green)	0.05 - 0.10	0.04 - 0.08	0.10 - 0.20
(light blue)	0.10 - 0.20	0.08 - 0.16	0.20 - 0.40
(dark blue)	> 0.20	> 0.16	> 0.40

Figure 5-9 (previous page). A comparison of favorable samples in the Dryden area of the western Wabigoon (location indicated in supplementary material 5-1.1). A geologic map and assemblage map (A; Montsion et al. 2021a/Chapter 3) were provided for reference when examining Igneous Rock Favorability (IRF) results for magmatic (B), volcanogenic (C), and orogenic (D) mineralization. IRF values for each system are indicated by colored sample location points. Large red and yellow points indicate samples above high favorability thresholds defined using inflections of a probability density function. Locations of known mineralization (i.e., deposits/prospects/occurrences) are indicated as black squares in each map. Map coordinates in NAD 83 UTM zone 15 N; Adz: Aiabewatik deformation zone; ALb: Atikwa-Lawrence batholith; BLb: Basket Lake batholith; Eg: Eltrut gneisses; Gld: Goldlund deposit; Gd: Goliath deposit; Kdz: Kawashegamuk deformation zone; Kd: Kenwest deposit; LBdz: Larson Bay deformation zone; LLRa: Long Lake River antiform; MDdz: Manitou-Dinorwic deformation zone; Mndz: Melgund north deformation zone; Msdz: Melgund south deformation zone; MBWdz: Mosher Bay-Washeibemaga deformation zone; NLdz: Noonan Lake deformation zone; Rb: Revell batholith; SB: Stormy Basin; SLa: Suzanne Lake antiform; TLa: Thunder Lake antiform; TLs: Thunder Lake syn- form; UMa: Upper Manitou antiform; VHd: Van Horne deposit; Vdz: Vermillion deformation zone; Wdz: Wabigoon deformation zone.

Felsic rocks, including FI, FII, and FIII rhyolites (Lesher et al., 1986) are common in both regions but their relative abundance varies. The southern Abitibi is dominated by FIII felsic rocks and contains very few FI felsic samples by comparison to the western Wabigoon (Figure 5-5). Despite a small population (mainly FIII) in the southern Abitibi that may be interpreted as fractional crystallization from mantle-sourced magma, most felsic rocks in both regions originate from partial melting of crustal material at shallow (FIII) to moderate (FII) to deep (FI) levels ($\bar{x}$ = 1.5 Ce/Y; $\bar{x}$ = 0.3 Th/Nb) and evolve through moderate crustal interaction and fractional crystallization (up to $\bar{x}$ = 11 Ce/Y; $\bar{x}$ = 3 Th/Nb). The general lack in FIII felsic rocks, abundance of FI felsic rocks, and more prevalent crustal interaction and fractional crystallization in western Wabigoon samples indicates a much thicker crust during volcanism compared to the southern Abitibi.

5.4.1.2. Applicability to other Archean cratons

The applicability of the proposed classification scheme was tested using a high-resolution dataset from the eastern Yilgarn (Smithies and Lowrey, 2020) and a subset of samples from several other Archean greenstone belts (GEOROC database from Sarbas and Nohl, 2009). While abundances of individual compositional classes varied by region, each displayed similar patterns as samples from the southern Superior Province.

In comparison to the southern Abitibi, the eastern Yilgarn displays: 1) significantly more ultramafic rocks reported and sampled in the eastern Yilgarn, with an overall higher MgO ($\bar{x}$ = 26 wt%) and a larger range of Ce/Y (up to $\bar{x}$ = 1.5) and Th/Nb (up to $\bar{x}$ = 1) values. Both regions are dominated by Al-undepleted komatiites (supplementary material 5-1.5). 2) Similar trends of crustal interaction and fractional

crystallization from tholeiitic magma toward "calc-alkaline" signatures in mafic samples, although Th/Nb trends are steeper (up to $\bar{x}$ = 1) suggesting more significant crustal interaction and fractional crystallization in the eastern Yilgarn (Figure 5-5). 3) Two trends of intermediate rocks, a first from a depleted mantle source ($\bar{x}$ = 0.4 Ce/Y) and a second from a slightly enriched mantle source ($\bar{x}$ = 1 Ce/Y), both with large ranges of crustal interaction and fractional crystallization (up to $\bar{x}$ = 1.5 Th/Nb). These two trends are also distinguishable using Al/Ti ratio. 4) A lack of FIII felsic rocks. While some felsic samples may have sufficiently low Ce/Y and Eu/Eu* values to be classified as FIII, their rarity in the database prevented them to show up as a separate cluster, and suggest an overall thicker crust in the eastern Yilgarn.

Patterns in clusters from the Amazonia, Sao Francisco, Baltic, Dharwar, Kaapvall, Zimbabwe, North Atlantic, and Pilbara cratons are broadly similar to all three of the locations described above, suggesting some commonalities in processes forming Archean greenstone belts; however, cluster distribution is more diffuse, which is expected for such a diverse global database. Notable differences in the global dataset compared to the southern Superior and eastern Yilgarn are a lack of lamprophyre and TTG rocks, which may reflect a sampling bias rather than actual abundances. Regardless of abundance variations in each setting, the similar distribution of samples from all plotted Archean igneous rocks indicates that the classification is applicable to a wide range of samples and may facilitate comprehensive discussion about geodynamic settings.

5.4.2. Limitations in application

The proposed classification is comprehensive in terms of capturing geologic processes relevant to geochemical discrimination of Archean igneous rocks; however, a few limitations affect its application in some use cases.

5.4.2.1. Single rock type analysis

The classification scheme is intended for broad/comprehensive investigations involving multiple compositional groups. For investigations involving processes unique to a given rock type, diagrams specifically designed for that rock type (e.g., Sproule et al., 2002, for komatiite, Moyen and Martin, 2012, for TTG) should be used to enhance the general information proved by this newly proposed classification scheme.

5.4.2.2. Data availability

As with any statistical investigation, representative sampling is recommended to properly capture trends and draw broad geodynamic conclusions. When applying this classification scheme to a dataset, balanced sample sets from all rock types in each supracrustal assemblage and intrusive complex is needed. Despite best intentions, however, sampling biases are prevalent in geology. Field access, density of outcrop exposure, and weathering often limit sampling capabilities. In such cases, a statement of bias and careful consideration during interpretation is required.

Additionally, available element analyses may be incomplete, especially for legacy geochemical data. When applying this classification scheme, each element from the ratios used to define diagram axes (Mg, Zr, Ce, Y, Th, Nb, Eu, Gd, Sm, Al, Ti) are needed to fully classify a sample. Mg/Zr and Ce/Y are commonly analyzed using a portable x-ray fluorescence device or traditional laboratory methods, and can assist in providing a first order rock classification during field work.-

5.4.2.3. Alteration or weathering

While this classification scheme leverages minimally/least mobile elements, strong alteration or weathering may obscure primary geochemical signatures. The application of the classification scheme is not recommended for rocks sampled in an environment that will affect one or more of the key variables. However, large populations of reliable samples can minimize the effect strongly altered outliers have on global trends.

5.4.2.4. Favorability indices

Samples constraining favorability indices were selected by spatial association with known mineralization (i.e., deposits and occurrences). The mode values indicated by these samples should be used with caution and may be unique for each locality. Certainly, further testing with a global database would provide a more robust constraint for exploration.

Additionally, rock favorability results do not capture the bimodal nature of some systems (e.g. volcanogenic). Instead, samples are simply measured against geochemical similarity to either end-member. Knowing this, the favorability should be assessed based on the presence of both modes in a given litho-stratigraphic assemblage.

5.4.3.   Application for exploration targeting in the Superior Province

The new 5D classification scheme provides a comprehensive framework to characterize favorable geologic processes and signatures for rocks spatially associated with known mineralization. In fact, the

element ratios used for the proposed classification scheme capture criteria frequently applied to Archean base and precious metal exploration.

5.4.3.1. Magmatic systems

When evaluating a region for Ni-Cu-PGE fertility, three factors are critical: 1) sufficient magma flux via continuous plumes extending from the melting zone to shallow crustal levels. Magma flux also governs flow rates and indirectly the amount of crustal interaction and fractional crystallization during ascent along 2) trans-lithospheric pathways/deformations zones and through 3) sulfur-rich source rocks such as chemical sediments and shales (Sproule et al., 2002; Begg et al., 2010; Barnes et al., 2012; Le Vaillant et al., 2016). Sulfur-isotopes and enrichment in incompatible elements (i.e., Zr, Th, LREEs) identify such favorable stratigraphy (Lesher et al., 2001; Le Vaillant et al., 2016). However, sulfur-rich rocks are sometimes too small/difficult to capture in regional exploration programs. Frequently, elevated crustal interaction in ultramafic volcanic rocks is used as a secondary proxy to track stratigraphic level such that prospective samples have higher crustal interaction after ascending through more crust (including sulfur-rich sources; Sproule et al., 2002; Begg et al., 2010; Barnes et al., 2012; Barnes and Fiorentini, 2012; Le Vaillant et al., 2016). The ideal tectonic setting for a Ni-Cu-PGE rich magmatic system is a (micro-)plate boundary where the lithosphere is thin and crustal-scale structures are present to deflect ascending mantle plumes and focus magma (Sleep et al., 2002; Barnes et a., 2012).

MgO concentrations are commonly used as a proxy for magma flux and mantle temperature; however, a limitation in only using MgO to assess favorability is that magnesian melts form from a wide range of volcanic rocks (i.e., komatiitic to andesitic magmas; Lesher, 1989, 2017; Naldrett, 1999, 2011), thus additional constraints are needed to successfully indicate favorability (Barnes et al., 2012; Le Vaillant et al., 2016). Fractionation of incompatible elements (e.g., Th, U, Y) identifies mafic mineral cumulate-rich rocks (Sproule et al., 2002; Davidson et al., 2007; Otamendi et al., 2016), suggesting that ratios between incompatible elements can be used as proxies for identifying favorable rocks. Additionally, major elements ratios, such as Al/Ti, describe plume temperature and residue in melts as a means of differentiating komatiite populations (Nesbitt et al., 1979; Sproule et al., 2002). Sproule et al. (2002) showed that there is progression in melt composition from Ti-depleted rocks in the older and baren volcanic packages (Pacaud) of the southern Abitibi to Al-depleted/Ti-enriched and finally to Al-undepleted in younger and magmatically mineralized volcanic assemblages (e.g., Kidd-Munro and Tisdale). Based on this information, Sproule et al. (2002) concludes Al and Ti indicate plume morphology as it ascends and extracts melts from progressively shallower depths.

The proposed classification scheme, and thus the favorability indices, incorporates most of the commonly used geochemical exploration tools for magmatic Ni-Cu-PGE deposits. The first-order classification, Mg/Zr and Ce/Y (section 5.3.3.1), identifies ultramafic rocks that are derived from high flux, high temperature magma (indicated by high MgO), and mafic mineral cumulate-rich rocks (indicated by low Ce/Y values). The second order classification, Th/Nb (section 5.3.3.2) captures crust (that may locally include sulfur sources) interaction, with Th and Nb inversely concentrated during crustal interaction (Pearce, 2008). Finally, the third-order classification, Al/Ti (section 5.3.3.3) utilizes Nesbitt et al. (1979)'s proxy for degree of melting in mantle derived material and tracks the most prospective plume conditions for mineralization in a magmatic system. In the southern Abitibi, moderate MgO and moderate Th/Nb indicates sub-optimal conditions/plume morphology to generate large Ni-Cu-PGE deposits but sufficient to generate a few localized zones of mineralization. In contrast, samples from the eastern Yilgarn display higher MgO and a much larger range of Th/Nb values, suggesting more optimal conditions for Ni-Cu-PGE deposits. The western Wabigoon lacks ultramafic rocks (low MgO, moderate Th/Nb) and is therefore not prospective for magmatic systems.

5.4.3.2. Volcanogenic systems

Key features for Cu-Zn-Pb-Ag(-Au) fertility include 1) a thin crust with 2) a shallow heat source (2-5 km depth) to drive convection of seawater from surface to 3) a high-temperature reaction zone where fluids leach metals and ascend along 4) trans-lithospheric brittle-ductile structures (Campbell et al., 1981; Franklin et al., 2005; Galley, 2003; Lesher et al., 1986) all present during 5) active volcanism. The geodynamic setting is critical for volcanogenic deposit development: rift or back-arc basins are considered optimal environments with thin crust and active magmatism (Allen et al 2002). Exploration for volcanogenic systems often targets characteristic alteration and metamorphism associated with shallow heat sources, reaction zones along fluid pathways, and silicified caps (Campbell et al., 1981; Franklin et al., 2005; Galley, 2003; Lesher et al., 1986). While alteration mapping/indices is outside the scope of the proposed classification scheme, the combined presence of favorable geochemical signatures in bi-modal volcanic assemblages can indicate a favorable geodynamic setting (e.g., Kidd-Munro and Blake River assemblages in the southern Abitibi). If only one volcanic composition is present, the assemblage is considered to be less favorable (e.g., Tisdale assemblage in in the southern Abitibi).

Favorable mafic volcanic rocks (boninite, island arc tholeiites, mid-ocean ridge basalt, and back-arc basin basalt) have distinct signatures but were generally derived from a depleted mantle source (low Ce/Y) and experienced little or no crustal contamination (low Th/Nb). Favorable FIII felsic (typically rhyolite)

rocks have characteristic flat HREE curves, negative Eu-anomalies, and intermediate to high abundances of HFSE that suggest modification in high-level magma chambers, which likely supplied heat for seawater convection (Lesher et al., 1986).

The proposed classification scheme and favorability indices uses Mg/Zr to differentiate mafic from felsic rocks: Mg is compatible in many mafic minerals/melts while Zr is compatible in felsic minerals/melts. The Ce/Y ratio discriminates between the three types of felsic rocks defined by Lesher et al. (1986) such that favorable FIII felsic rocks have Ce/Y values near 1 since REEs are generally unfractionated while FII and especially FI felsic rocks are enriched in LREEs (i.e., Ce) and depleted in HREEs (i.e., Y). The Th/Nb ratio tracks crustal interaction, and, along with Ce/Y, can be interpreted as a proxy for crustal thickness/depth of source (i.e., presence of or absence of garnet as a residual phase or as a fractional crystallization product during melt evolution). Finally, the Eu/Eu* ratio documents plagioclase fractionation in magmatic chamber and help in discriminating FIII from other felsic rocks.

5.4.3.3. Orogenic systems

Unlike magmatic and volcanogenic, Au fertility is not genetically related to the formation of their host rocks. Instead, it is the result of 1) orogenesis driving deformation, magmatism, and metamorphism that 2) generates metal-rich fluids to be transported along 3) pathways and concentrated in 4) traps (e.g., Robert and Poulsen, 2001; Robert et al., 2005; Dubé and Gosselin, 2007; Groves and Santosh, 2015; Groves et al., 2018). Orogenic deposits can occur in any metamorphic rock protoliths, including volcanic (e.g., Bateman et al., 2008; Dubé and Mercier-Langevin, 2020), intrusive (e.g., Robert et al., 2001; Katz et al. 2017), and sedimentary (e.g., De Souza et al., 2020; Pitcairn et al., 2021).

As with many other mineral systems, alteration is a priority exploration tool to track fluid mobility and interaction with country rock. Since discrimination by mobile elements is not captured by this classification scheme and host rock compositions vary widely and thus not discriminatory, only intrusions are assessed as possible fluid and/or metal sources. Intrusions that cooled before mineralization may have also provided sufficient rheological/chemical contrast to act as traps for precipitating metals (e.g., Burrows and Spooner, 1988; Zhang et al., 2014; Perrouty et al., 2017) or exploited the same structural as mineralized fluids. Lamprophyre (Taylor et al., 1994) and other alkalic intrusions (Robert et al., 2001), as well as late S-type leucogranites and pegmatites, are often temporally, if not genetically, related with gold mineralization and late-stage orogenesis. They may have exsolved fluid and metals, and commonly exploited existing fluid pathways, leading to a spatial association with orogenic gold deposits.

The proposed classification scheme and favorability indices identifies intrusions that are commonly spatially associated with orogenic systems. As stated previously, the Mg/Zr ratio differentiates rocks by capturing opposing compatibility in mafic and felsic minerals, and the Ce/Y ratio identifies compositions using fractionation of LREEs vs HREEs. Mafic bodies are distinguished from other intrusive rocks by their elevated Mg/Zr and relatively low Ce/Y (i.e., flat REE pattern). Lamprophyre and other alkali intrusions display similar Mg/Zr but a much higher Ce/Y. Tonalite-Trondhjemite-Granodiorite (TTG) and late S-type leucogranite or pegmatites also display high Ce/Y, with a significant fractionation of HREE indicative of garnet-bearing source (Moyen and Martin, 2012). Other ratios, such as Th/Nb and Eu/Eu*, address processes (crustal interaction and fractional crystallization) that are less relevant for orogenic gold exploration. Nevertheless, they are useful at documenting magmatic evolution and crustal settings, in particular for TTGs with high-Al2O3, which are derived from fractionating a hydrated basalt or partial melting of basalt where little or no plagioclase remains in the residue, effectively meaning no Eu-anomaly (Barker and Arth, 1976). Alternatively, Al/Ti is another useful tool to distinguish TTGs from other rocks because of their characteristically high Al2O3 abundances and low TiO2 anomalies (Barker and Arth, 1976; Moyen and Martin, 2012).

5.5. Conclusions

Most currently available tools for classification and exploration are ill-suited for use with metasomatized Precambrian rocks or with compositionally diverse sample populations. Additionally, existing tools have mainly relied on expert interpretation of geochemically-related systems and subjective techniques to define diagram axes. Using statistical approaches such as PCA and k-means cluster analysis, a new classification scheme was defined for Archean igneous rocks. These techniques were constrained by whole rock geochemical databases compiled for the southern Abitibi and western Wabigoon subprovinces, Canada and then verified using samples from the eastern Yilgarn and other Archean cratons. The proposed classification scheme, which utilizes less mobile major elements and least mobile trace, HFSE, REEs, successfully discriminates between the full range of compositional groups (i.e., ultramafic to felsic) at low to moderate degrees of alteration/metamorphism. It also facilitates comprehensive discussion on igneous processes such as mantle/crustal sources, crustal interaction, fractional crystallization and accumulation.

In addition to providing a new approach for rock classification, the proposed diagrams suggest first-order differences in crustal formation and evolution between the southern Abitibi, eastern Yilgarn and western Wabigoon regions, with implication for base and precious metal endowment. Geochemical data

for the southern Abitibi suggest plume-related magma (e.g., presence of komatiites) erupting on top of thin (e.g., presence of FIII rhyolites) mafic-intermediate crust with minimal crustal interactions (e.g., moderate range of Th/Nb), leading to an environment moderately favorable for magmatic systems and very favorable for volcanogenic and orogenic deposit formation. In the eastern Yilgarn, similar plume-related magma (e.g., presence of komatiites), slightly higher in temperature (e.g., higher MgO), erupting on top of moderately thick (e.g., lack of FIII rhyolites) mafic-intermediate crust with significant crustal interactions (e.g., large range of Th/Nb), lead to an environment very favorable for magmatic and orogenic deposits but poorly suitable for volcanogenic systems. In the western Wabigoon, lower temperature magma (e.g., lack of komatiites) erupting on top of a thick (e.g., lack of FIII rhyolites) intermediate to felsic crust (i.e., Paleo to Mesoarchean gneisses) with significant crustal interactions (e.g., large range of Th/Nb), lead to poorly favorable setting for either magmatic, volcanogenic or orogenic systems.

Similarity of samples distribution and trends (e.g., continuum between tholeiitic and "calc-alkaline"/crustal interaction signatures) between independent cratons implies some commonalities in the Archean crust-forming processes and outlines the broad applicability of the proposed classification scheme to Archean, and perhaps modern, questions. Its statistically driven design provides Precambrian geoscientists with a comprehensive framework for comparing and discussing igneous processes that would otherwise be much more challenging to interpret using existing classification diagrams.

Chapter 6. Mapping structural complexity – example of new feature engineering method

Copyright statement

Key words

Superior Craton, Mineral exploration, Structural complexity, Orogenic gold, Fisher statistics, Aeromagnetic auto-detection

Abstract

With the increasing size and complexity of geoscientific databases, statistical tools and machine learning algorithms are needed to highlight subtle patterns within dense data clouds and reduce human-related bias in geological mapping. This study presents a new tool that maps structural complexity using circular variance and spherical dispersion in two Archean greenstone belts near Dryden and Timmins in the Superior Province, Ontario. Structural complexity quantitatively highlights areas where structures are folded, juxtaposed, and/or parallel. Bedding measurements and autodetected aeromagnetic lineaments are used as inputs to test if structural complexity can refine or add confidence to existing geological interpretations. Linear, low variance trends frequently delineate the location and extent of deformation zones, which likely reflect parallel transposed fabrics in high strain corridors. Areas of high variance spatially correlate with fold traces. Sharp boundaries between high and low variance anomalies are interpreted as faults and/or shear zones that juxtapose structural blocks. The near continuous and representative nature of autodetected magnetic lineaments was effective in capturing structural complexity at a regional scale; however, spherical dispersion captured the highest resolution of structural complexity at local scales (<1:30000). Key considerations when performing structural complexity analysis are scale, data density, and optimization of input parameters (e.g., neighborhood

radius, lineament length, minimum neighborhood populations). Gold deposits are found <~1700 m from low (< μ - 1σ) and < ~1500 m from high circular variance anomalies (> μ + 1σ). Additionally, the gold grade of large deposits (> 1 Moz) increases with proximity to high variance anomalies. When paired with expert knowledge, these techniques will increase repeatability in future exploration endeavors, making exploration a more rigorous process with increased confidence.

Acknowledgements

6.1. Introduction

Geological modeling commonly integrates a wide range of geoscientific data (e.g., field observations, geophysics, geochemistry, geochronology). In recent years, secondary or derivative products such as fault density and heat flow models enhance characterization of various components in crustal architecture (e.g., Holden et al., 2012; Hronsky, 2019) using increasingly data-rich and complex geoscientific databases. These new products act as effective constraints that provide additional confidence and a means of further refining qualitative interpretations. Herein, a new derivative parameter, 'structural complexity', from variance of bedding measurements and aeromagnetic lineaments, is calculated to support geological interpretation of shear-zones, faults, and folds in the western Wabigoon and southern Abitibi subprovinces of the Canadian Shield. Structural complexity here refers to an increased variance of planar feature orientations, such as bedding, in response to deformation events.

In addition to improving geological interpretations, structural complexity mapping may also be applied to exploration for mineral deposits. Frequently, mineral prospectivity studies use proximity or density calculations of favorable features (e.g., fault traces, fold axes, favorable lithological units) to represent structurally complex zones as a proxy for both fluid pathways and traps, which are likely to be genetically related to ore formation (Knox-Robinson and Wyborn, 1997; Carranza and Hale, 2001; Bierlein et al., 2006; Kreuzer et al., 2015; Zhang and Zhou, 2015; Campos et al., 2017; Ford, 2019). Within orogenic and other mineralizing systems, fluid pathways act as the plumbing system that carries mineralized fluids through the crust to structural and/or chemical traps where fluids are concentrated and metals are precipitated. The preparation of derivative products that highlight structural traps, such as antiform fold hinges and brittle faults, are often used as inputs in such prospectivity analysis; however, they frequently oversimplify and heavily bias geological features. Efforts to diminish human-related biases and increase repeatability are being developed to make prospectivity analysis more rigorous and robust (An et al., 2001; Ford and McCuaig, 2010; Kreuzer et al., 2015; Lindsay et al., 2016; Yousefi and Carranza, 2016; Yousefi and Nykänen, 2016; Almasi et al., 2017; Burkin et al., 2019; Hronsky and Kreuzer, 2019; Schaeben et al., 2019). Automated techniques of mapping structural complexity may provide a valuable addition to these efforts.

Statistical mapping of structural complexity is applied to two Archean greenstone belts in the western Wabigoon subprovince near Dryden, Ontario, Canada and southern Abitibi subprovince near Timmins, Ontario Canada. These study sites include representative areas of juvenile volcanic crust that record

similar deformational histories but contrasting known base and precious metal resources, while having sufficient geological and geophysical data to support statistical analyses.

Compiled structural measurements, geological interpretations, new regional geological maps for each area of interest, and mineral deposit information are provided in a joint data-in-brief publication (Montsion et al., 2021a/Chapter 3).

6.1.1. Regional geology

The Superior Province is the largest preserved Archean craton and consists of Eo- to Neoarchean granite-greenstone, gneissic, and metasedimentary domains that amalgamated as a series of generally east-trending belts between ~2720 and ~2660 Ma (Percival et al., 2012; Frieman et al., 2017). Both selected regions (Figure 6-1) have a similar geologic history in terms of geodynamic setting, volcanism, deformation, and metamorphism (Beakhouse et al., 2011; Zammit, 2020). Since all rock units in both areas are, with the exception of Proterozoic diabase dikes, metamorphosed to greenschist to amphibolite facies, the term 'meta' will not be used in lithologic descriptions.

6.1.1.1. *Dryden study area in the western Wabigoon* subprovince

The western Wabigoon subprovince (geological map in Montsion et al., 2021a/Chapter 3) is primarily composed of juvenile Neoarchean volcanic successions that are bounded by the Paleo- to Neoarchean Winnipeg River and Marmion gneissic-plutonic terranes to the north and south, respectively (Percival et al., 2004; Tomlinson et al., 2004). The Dryden study area (Figure 6-1; geological map in Montsion et al., 2021a/Chapter 3) hosts felsic, intermediate, and mafic volcanic formations (ca. 2750–2715 Ma, Davis et al., 1989; David et al., 2005), marine turbiditic sequences (ca. 2715–2710 Ma; Davis et al., 2005), and fluvial-alluvial clastic sedimentary units (ca. 2703–2696 Ma; Beakhouse et al. 1995; Dostal et al., 2004; Corcoran and Mueller, 2007).

The volcanic and sedimentary packages were intruded by tonalitic to granodioritic syn-volcanic batholiths (ca. 2750–2680 Ma; Blackburn et al., 1981; Davis et al., 2005; Duguet and Beakhouse, 2012), such as the Revell batholith, Atikwa-Lawrence batholith, and Basket Lake batholith (Montsion et al., 2021a/ Chapter 3). Smaller dioritic, granitic and syenitic porphyritic intrusions such as the Lost Lake porphyry, Thundercloud porphyry, and Taylor Lake stock (geological map in Montsion et al., 2021a/Chapter 3) intruded during syn-volcanic to post-deformation phases (ca. 2705–2695 Ma). Finally, two generations of Proterozoic diabase dikes crosscut the western Wabigoon subprovince. These are

northwest-trending dikes (ca. 1887-1889 Ma; Davis and Hamilton, 2010) and north-trending dikes (~1137 Ma; Heaman and Easton, 2006; Stone et al., 2010)

Deformation began in the Neoarchean (~2700 Ma) with north-south shortening that resulted in kilometer-scale isoclinal to outcrop-scale parasitic folds with an associated axial planar foliation that is mainly marked by chlorite in volcanic units and micas in sedimentary units (Kresz et al., 1982a, 1982b; Montsion et al., 2019; Zammit 2020). Progressive shortening resulted in the localization of strain into several east-trending deformation zones, including the Wabigoon and Mosher Bay-Washeibemaga deformation zones (Montsion et al., 2021a/Chapter 3; Blackburn, 1978; Zammit, 2020), which underwent north-over-south movement (Zammit, 2020) associated with the unconformable deposition of coarse-grained sediments in the Stormy and Manitou basins (Kresz et al., 1982a, 1982b; Corcoran and Mueller, 2007; Montsion et al., 2021a/Chapter 3). A later stage of northwest-southeast shortening, coeval with most orogenic gold mineralization in the area (Melling et al., 1988; Zammit, 2020), refolded the earlier isoclinal folds (i.e., locally increasing structural complexity), and has been interpreted as recording sinistral transpressive strain along the northeast-trending Manitou-Dinorwic deformation zone (geological map in Montsion et al., 2021a/Chapter 3; Parker, 1989; Zammit, 2020).

Although the western Wabigoon subprovince is much less endowed in base and precious metal resources compared to the Abitibi subprovince, it nevertheless hosts several deposits and developed prospects with a variety of metal resources (gold, silver, copper, zinc, iron, lithium, molybdenum, uranium, thorium, and rare earth elements; Parker, 1989). Gold systems include Goliath (1.2 Moz Au), Goldlund (0.8 Moz Au), Gold Rock (0.15 Moz) and Van Horne (0.002 Moz, past production plus measured and indicated resources, Ontario Geological Survey, 2019).

6.1.1.2. The Timmins study area in the southern Abitibi subprovince

The Timmins study area (Figure 6-1; geological map in Montsion et al., 2021a/Chapter 3) comprises ultramafic to felsic volcanic assemblages that are unconformably overlain by two sedimentary successions. Thick glacial cover and overburden limits outcrop exposure; however extensive drilling has sufficiently constrained geological interpretations within the Timmins region. Volcanic successions range in age from ~2750 to ~2695 Ma and include disconformities that are locally marked by iron formation and cherts (Thurston and Chivers, 1990; Thurston et al., 2008; Frieman et al., 2017). The Porcupine sedimentary assemblage (ca. 2690 – 2685 Ma, Frieman et al. 2017) comprisesfine to medium grained beds with rare conglomeratic lenses, iron formations, and mafic to felsic volcanic flows. The Timiskaming coarse- to medium-grained sedimentary assemblage (ca. 2679-2669 Ma; Frieman et al., 2017) is

primarily preserved as narrow (<2 km) panels to the north of the Porcupine-Destor deformation zone (PDDZ; Ayer et al., 2005; Bateman et al., 2005; Bleeker, 2015; Montsion et al., 2021a/Chapter 3). A comprehensive geological overview is provided in Ayer et al. (2002, 2005), Bateman et al. (2005), Bleeker (2015), Monecke et al. (2017) and references therein.

The volcanic and sedimentary packages were intruded by ultramafic to felsic bodies that have been classified based on composition and relative timing (i.e., syn-volcanic and syn-deformation). Syn-volcanic intrusions are large, batholitic complexes that display gabbroic to granodioritic compositions, similar to their volcanic counterparts, and are generally found west of the Mattagami River fault (Corfu, 1993; Mortensen, 1993a,b; Ayer et al., 2005, Montsion et al., 2021a/Chapter 3). Syn-deformational intrusions, which dominate the majority of the map area, cross-cut sedimentary and volcanic assemblages, are variably deformed depending on proximity to deformation zones, and are tonalitic to "potassic" (i.e., granite, monzonite, and syenite) in composition (Ayer et al., 2002). Late intrusive intermediate to felsic porphyritic clusters as well as several diabase dike swarms intrude supracrustal units. The latter include the north-northwest trending Matachewan dikes (ca. 2445-2490 Ma; Ernst and Buchan, 2010), northwest-trending Sudbury dikes (~1235 Ma; Ernst and Buchan, 2010), and northeast-trending Abitibi dikes (~1140 Ma; Ernst and Buchan, 2010).

The structural history of the Timmins camp is well documented (Ayer et al., 2002, 2005; Bateman, 2008; Bleeker, 2015; Monecke et al., 2017), with two main shortening events. A first north-south shortening event resulted in kilometer-scale isoclinal folds in the volcanic and Porcupine sedimentary assemblages, associated with an axial planar foliation mainly marked by chlorite in volcanic units and micas in sedimentary units. Generally east-trending brittle-ductile deformation zones such as the crustal-scale PDDZ (Figure 6-2C; Ayer et al., 2005; Bateman et al., 2008, Montsion et al., 2021 a/Chapter 3) locally controlled the unconformable deposition of the Timiskaming sediments (Bateman et al., 2008; Bleeker, 2015). A second northwest-southeast shortening event resulted in the refolding of the isoclinal folds, formation of asymmetric z-folds, and the development of a crenulation cleavage axial planar to z-folds. This event is coeval with the formation of world-class orogenic gold deposits along the PDDZ (Dubé and Gosselin, 2007). Late Neoarchean or Proterozoic north-northwest trending faults, such as the Mattagami River, Burrows Benedict and Buskegau River faults (geological map in Montsion et al., 2021 a/Chapter 3), offset older structural features.

The Timmins mining camp hosts several world-class mineral systems, including the Kidd-Creek Volcanogenic Massive Sulfide (VMS; 9.6 Mt Zn, 3.7 Mt Cu), Dome (27.1 Moz Au), Hollinger (20.5 Moz

Au), McIntyre (10.8 Moz Au) and 10+ smaller orogenic deposits (>1 Moz Au). Total endowment represents over 2700 t of gold, 10 Mt of zinc, 4 Mt of copper, and 60000 t of nickel (measured and indicated, Ontario Geological Survey, 2019).

6.2.   Data

6.2.1.   Structural measurements

In Dryden, structural measurements (e.g., bedding, younging, lineations, and foliations) were digitized from 64 geological maps (complete list in Montsion et al., 2021 a/Chapter 3). These were integrated with one digital compilation database (Beakhouse et al., 2011) and new field observations from 2018 and 2019 fieldwork. Combined, the database is composed of roughly 28000 data points. Figure 6-2A presents a subset of the data overlain on mapped geology. A rose diagram of bedding in this area shows that bedding is generally oriented northeast-southwest and the calculated axial plane of the Upper Manitou anticline has a strike of 050° and dip of 85°. The axial plan was calculated using pi-analysis (e.g., Ragan, 2009).

In Timmins, several digital databases consisting of drill logs, structural measurements, and multi-scale geological interpretations (Ayer and Chartrand, 2011; Bateman, 2005; Berger, 2010; Hathway, 2005; Houlé and Hall, 2007; MIRA, 2005), along with digitized points from 33 scanned maps (complete list in Montsion et al., 2021 a/Chapter 3), were compiled and integrated to generate a comprehensive structural measurement database with roughly 8000 measurements. The number of Timmins measurements is restricted by thick glacial overburden and limited outcrop exposure. Most of the compiled measurements are concentrated around the city of Timmins and other, smaller mineralized properties. Frequently, data points are clustered around outcrops and appear superimposed in regional-scale maps. Figure 6-2C presents representative compiled structures as well as a rose diagram of bedding for the Timmins camp. Bedding in this area trends east-northeast and the axial plane of the most dominant fold is 260/88; however, several progressive deformations events, strain from the nearby PDDZ, and intrusions have resulted in complex fold interference patterns.
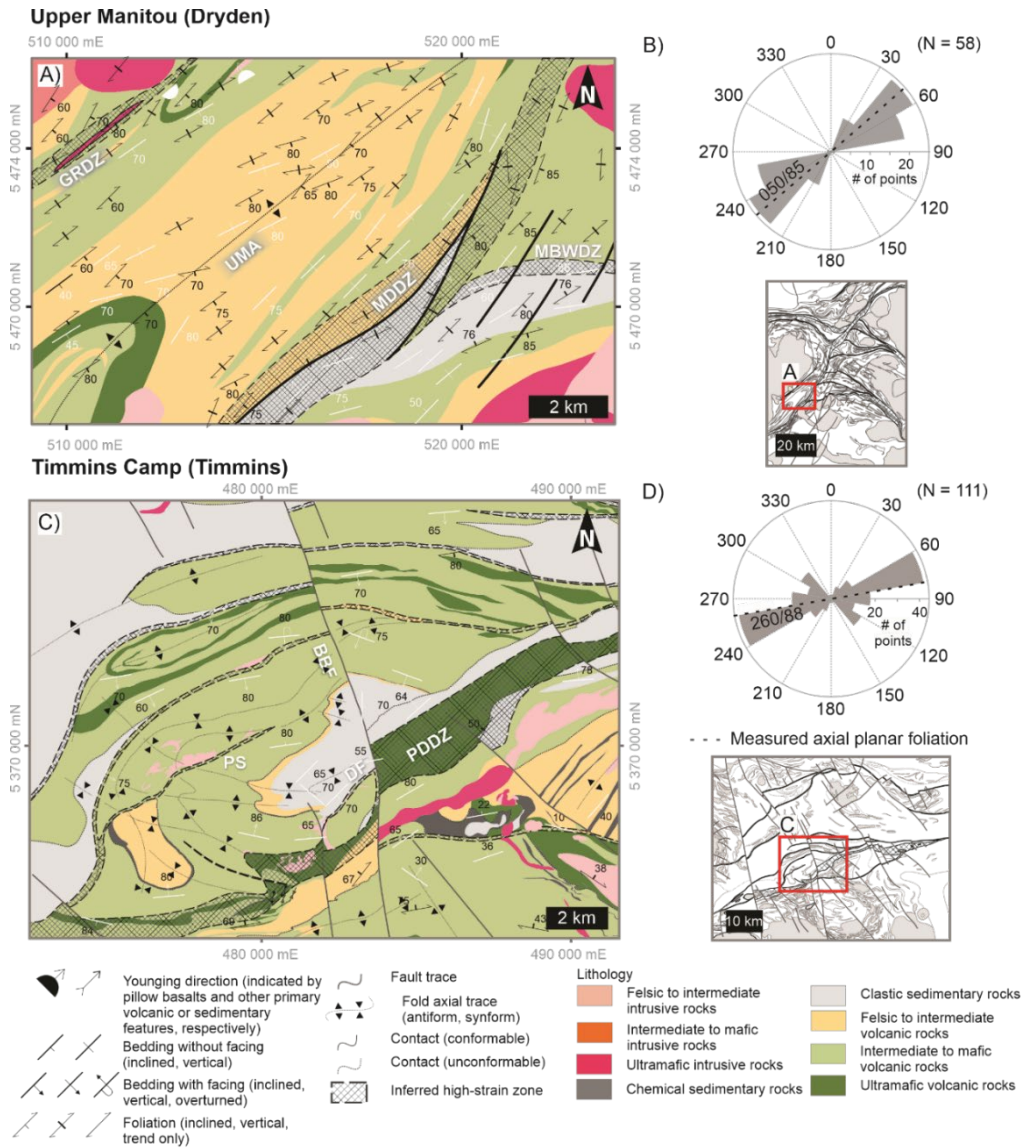
Figure 6-2. Representative bedrock geology map for A) Upper Manitou region near Dryden, Ontario and C) the Timmins gold camp near Timmins Ontario; Bedding and parallel S1-foliations compiled from open sources and new field observations are superimposed on similarly symbolized rock units; Rose diagrams of compiled bedding for both the B) Upper Manitou and D) Timmins Camp demonstrate the main bedding orientation and the dominant axial plane orientation; Population size indicated as N; For the Upper Manitou areas, the axial plane was calculated using pi analysis (050/85; strike/dip) is superimposed on the rose diagrams. For the Timmins camp area, the axial plane of the Porcupine Syncline is interpreted to be the mean axial planar foliation value measured in field investigations (260/68; strike/dip); Simplified maps of both B) Dryden and D) Timmins indicate the location of representative bedrock maps; Coordinates in NAD 83 UTM zone 15N in Dryden and Zone 17N in Timmins; Legend symbols and lithologic units described in Figure 6-1; (BBF) Burrows Benedict faults, (DDZ) Dome deformation zone, (GRDZ) Gold Rock deformation zone, (MBWDZ) Mosher Bay-Washeibemaga deformation zone, (PDDZ) Porcupine-Destor deformation zone, (PS) Porcupine synform, (Uma) Upper Manitou antiform.

6.2.2. Geophysical

The Dryden aeromagnetic data were acquired between November 2000 and February 2001, with 200 m line spacing, 1500 m tie lines spacing, and a flight elevation of 70 m (Ontario Geological Survey, 2011). The Timmins aeromagnetic data were acquired between 1975 and 1992, with 200 m line spacing, 1000 m tie lines spacing, and a flight elevation of 70 m (Ontario Geological Survey, 2003). Both data sets were reprocessed as part of this study. Total magnetic intensity (TMI) data were gridded at 40 m resolution in both areas, using the minimum curvature algorithm in Geosoft Ltd. Oasis Montaj™. Reduction to the pole (RTP) was based on the IGRF (International Geomagnetic Reference Field), calculated at the date of the survey (Dryden: declination −0.9°, inclination 75.2°; Timmins: declination −11.2°, inclination 74.8°). First and second vertical derivatives, tilt (see Figure 6-3A and 6-4A; Miller and Singh 1994; Thurston and Smith 1997; Verduzco et al. 2004), and Phase Preserving Dynamic Range Compression (see Figure 6-3B and 6-4B; Kovesi 2012; Holden et al. 2013) were calculated from the RTP. Each technique increases contrast and/or act as bandpass filters, thereby highlighting trends.
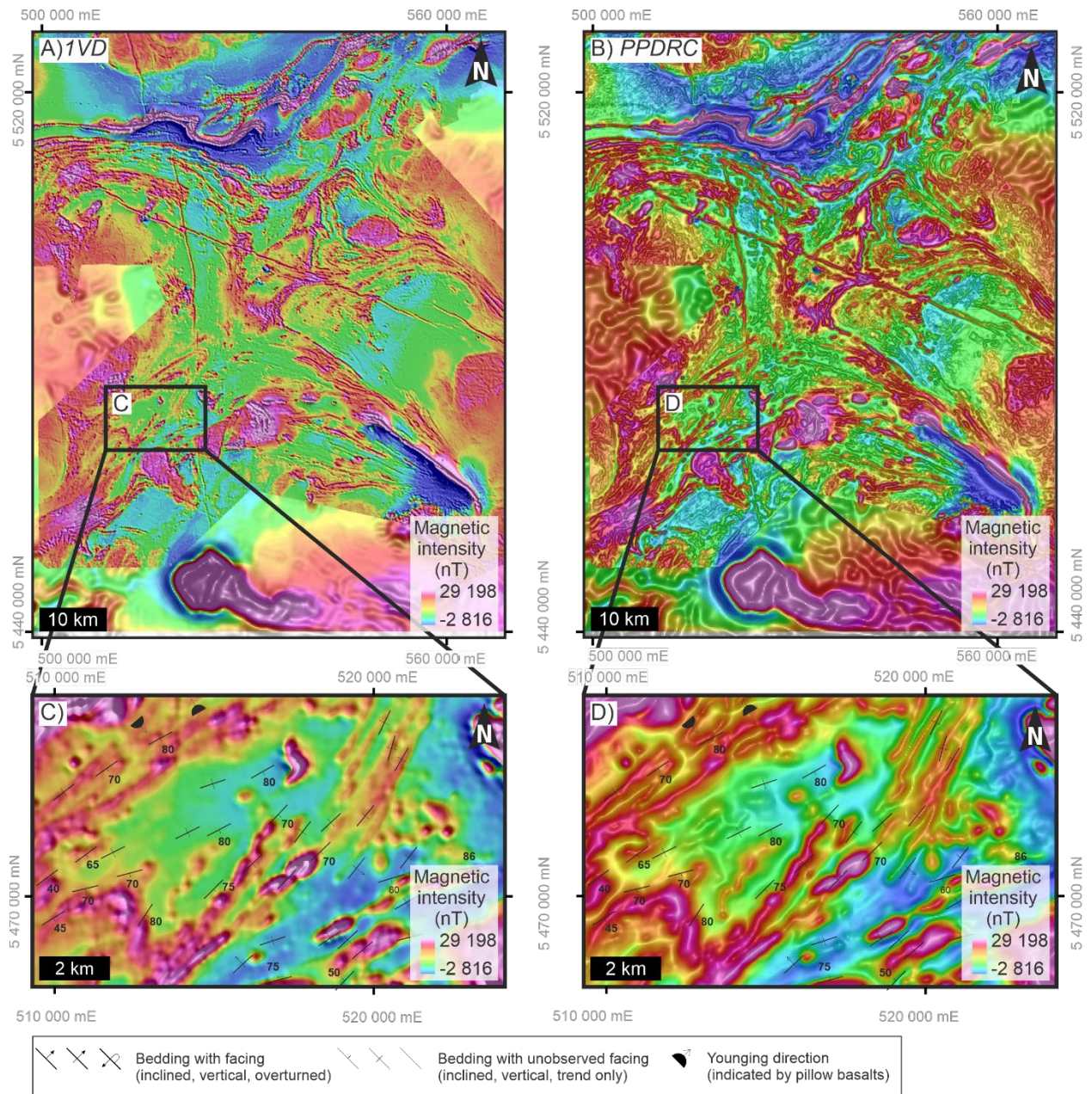
Figure 6-3. Reduced to pole (RTP) of aeromagnetic maps for the Dryden area transparently overlain on A) hill-shaded first vertical derivative (1VD) and B) Dynamic Range Compression (PPDRC) filtered grids; An inset grid of the Upper Manitou area near Dryden visually compares C) RTP over 1VD and D) RTP over PPDRC to representative bedding measurements. For both C and D, linear magnetic trends are sub-parallel to bedding measurements; Coordinates in NAD 83 UTM zones 15N.
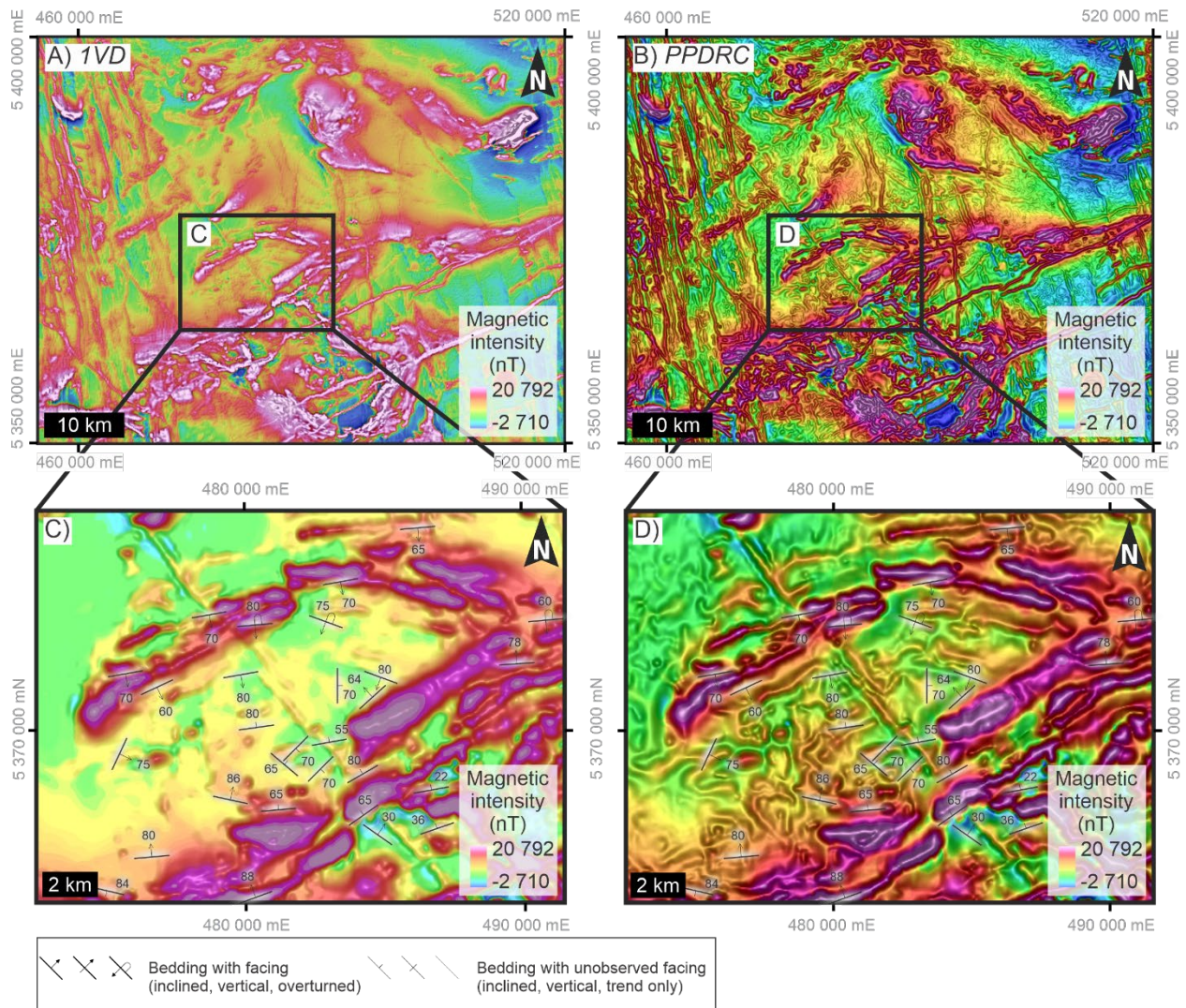
Figure 6-4. Reduced to pole (RTP) aeromagnetic maps for the Timmins area transparently overlain on A) hill-shaded first vertical derivative (1VD) and B) Dynamic range compression (PPDRC) filtered grids; An inset grid of the Timmins gold camp visually compares C) RTP over 1VD and D) RTP over PPDRC to representative bedding measurements. For both C and D, linear magnetic trends are sub-parallel to bedding measurements; Coordinates are NAD 83 UTM zones 17N.

Figure 6-5 (next page). Histograms of magnetic susceptibility (MS) compiled from Biswas (2019) and new field measurements; (A and B) MS for volcanic rocks with ultramafic, mafic, intermediate compositions; (C and D) MS for intrusive rocks with ultramafic, mafic, intermediate compositions; (E) MS for diabase dikes; A and B also display enlarged histograms of volcanic rocks to identify trends with low range MS (0.1 to 1 SI). Histograms and trendlines distinguish distinct magnetic signatures for rock units based on composition and type; (n) population size of analyses in each lithologic group. All plots represent Magnetic Susceptibility along the x-axis. For visual simplicity, only the lower-most plots are labelled.
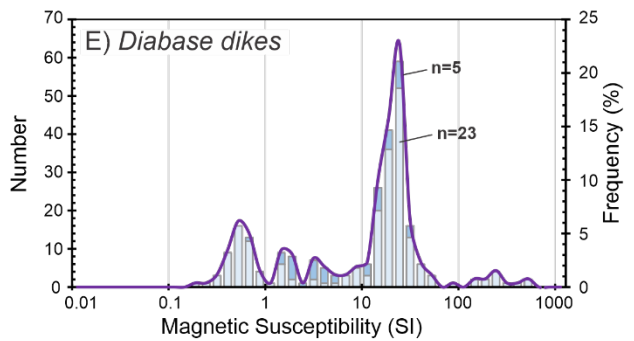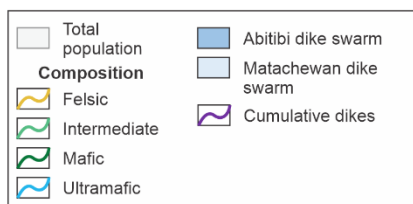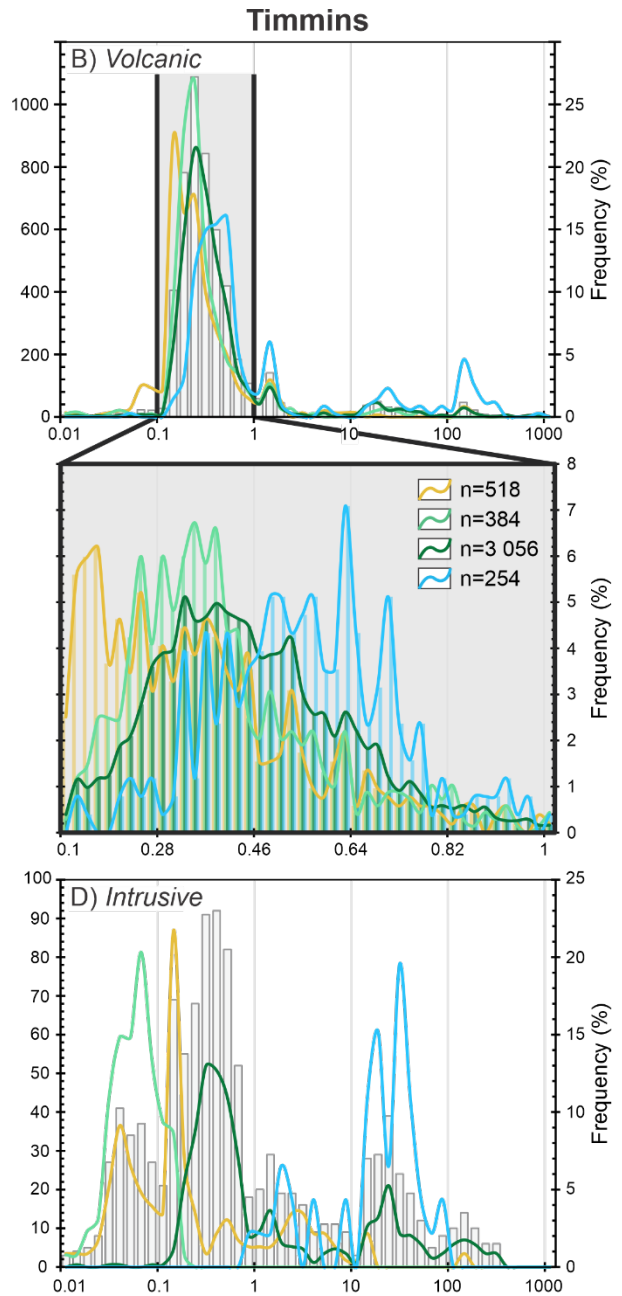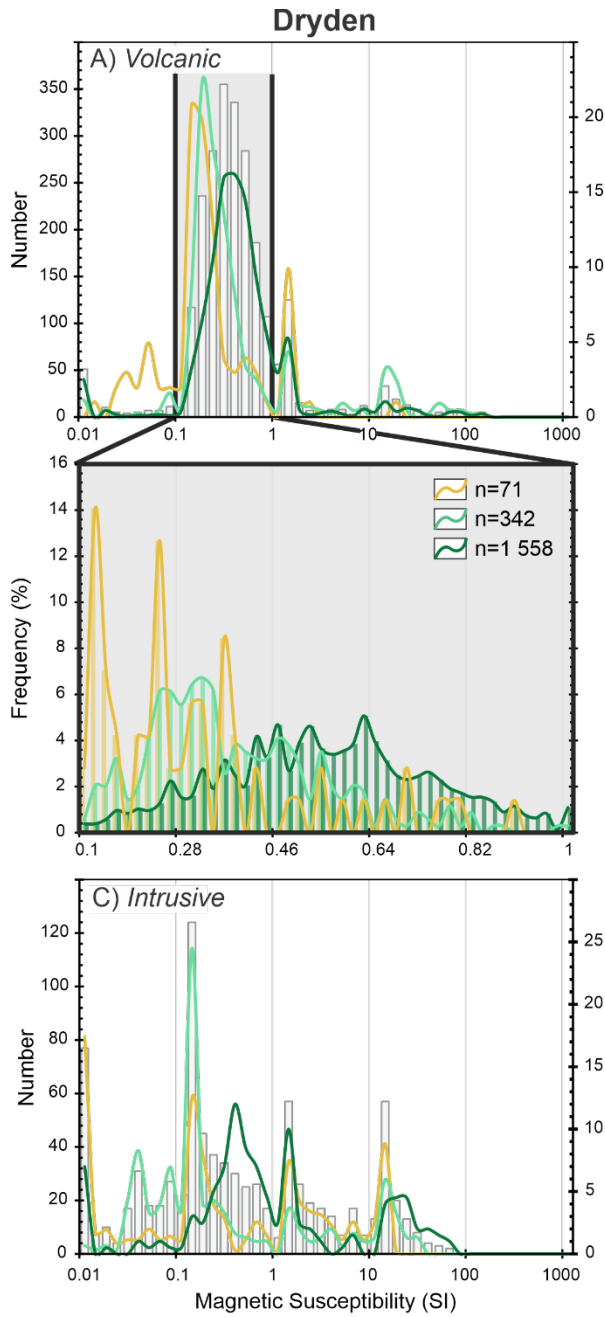
6.3. Methods

The proposed workflow for mapping structural complexity was performed using two inputs, field bedding measurements and auto-detected aeromagnetic lineaments. Techniques relating to pre-processing and lineament extraction from aeromagnetic grids as well as methods of calculating complexity are described below and illustrated in Figure 6-6. Parameter sensitivity testing was also performed as part of the workflow to determine optimal parameters and the effect each has on results.

01 STEP	02 STEP	03 STEP	04 STEP	05 STEP	06 STEP
Gridding (Section 3.1.1)	**Lineament Extraction (Section 3.1.2)**	**Lineament Post-Processing (Section 3.1.3)**	**Calculate (Section 3.2)**	**Interpolate (Section 3.2)**	**Validation (Section 3.3)**
Input: Geophysical points	**Input:** Geophysical Grids	**Input:** Lineaments	**Input:** Structural points (lineaments & bedding)	**Input:** Structural complexity points	**Input:** Structural complexity grids
Method: RTP, 1VD, Wavelength filtering	**Method:** Edge detection & thresholding (iterative testing)	**Method:** Segment & measure trend	**Method:** Circular variance & spherical dispersion (iterative testing)	**Method:** Minimum curvature	**Method:** Kolmogorov-Smirnov test
Output: Geophysical Grids	**Output:** Lineaments	**Output:** Structural Points	**Output:** Structural complexity points	**Output:** Structural complexity grids	

Figure 6-6. Illustrated summary of steps involved for the presented structural complexity analysis method. Panels include methods and outputs to grid geophysical data, detect lineaments, and calculate complexity maps as described in Methods (section 6.2.1 to 6.2.3); Relevant sections where further details are discussed in the text are indicated in each panel; Circular arrows indicate iterative processes where parameters and inputs were tested for sensitivity and visually assed to select an optimum combination.

6.3.1. Geophysical grids

6.3.1.1. Pre-processing for lineament extraction

Filtering methods such as vertical and tilt derivatives are commonly used to outline high amplitude anomalies, which are often interpreted as intrusive contacts or magnetite-rich stratigraphic layers. Despite their prevalence in traditional geophysical studies, these methods provide poor constraints in regions where contrast in physical properties, such as magnetic susceptibility, between adjacent lithological units is low. Signals can be masked by nearby strongly magnetic units (e.g., Fe-rich units like magnetite-bearing dikes or banded iron formations) or are attenuated by overburden (e.g., Vallée et al., 2019). These methods consider the full range in magnetic intensity across the map area causing

sometimes irrelevant and anomalous outliers to dominate outputs while suppressing detection and characterization of relevant anomalies.

To mitigate these effects and to highlight local magnetic trends, a high pass filtering technique termed Phase Preserving Dynamic Range Compression (PPDRC) from the CET Grid Analysis tool in Oasis Montaj™ was applied (PPDRC; Kovesi, 2012; Seequent 2019). This processing diminishes the effect of spatially distant anomalous values while amplifying subtle local trends, thus refining lineament extraction results. The PPDRC method applies monogenic filters to separate local phase and amplitude statistics within the moving windows. The grid is reconstructed while the filter preserves phase values and applies a range of reducing functions to amplitude values (Wong et al., 2006; Kovesi, 2012; Holden et al., 2013). The most significant parameter to note for this tool is the high pass cut-off wavelength. This cut-off is used to control the scale of interest such that spatial frequencies in the grid with wavelengths greater than indicated will be suppressed so residual features can be captured more effectively (Wong et al., 2006).

6.3.1.2. Lineament extraction

Linear continuity in magnetic anomalies was captured as vectorized polylines using the Line Extract tool from the Focus image analysis package provided by PCI Geomatics (2018, 2019). Extracted lineaments are defined by contrasts between spatially continuous magnetic anomalies using a suite of parameters, described below and reported in Table 6-1. Lineament extraction is a two-stage process: 1) edge detection and 2) thresholding.

6.3.1.2.1.   Edge detection

The Canny approach produces an edge strength image of the hill shaded, gray-scale PPDRC grid (Canny, 1986; Liang et al., 2011) and involves filtering with a Gaussian function whose radius is defined by the 'Radius of Filter' parameter. Hill shading was used to increase contrast between minimum and maximum anomalies. Several sun directions and angles for the hill shades were tested, which all yielded similar results. A gradient within the radius of the filtering window is computed and any pixels whose gradient is not the local maximum are suppressed (i.e., edge strength set to 0 or 'OFF'). This lineament extraction technique generates a binary image where each remaining grid cell (i.e., edge strength = 1 or 'ON') represents an edge element.

6.3.1.2.2.   Thresholding

As part of the Focus lineament detection routine, a thinning algorithm is applied to binary grids, producing pixel-wide traces that undergo testing against control parameters such as curve length, fitting error and angular distance. Traces with a minimum defined extent are extracted when the number of pixels satisfies the user-defined "Minimum for curve length" parameter. These traces are vectorized by fitting line segments as an approximation of the original pixel trace, where the maximum fitting error is defined by the "Maximum line fitting error" parameter. If the distance between pixels and the main trace exceeds this error, they will not be used to constrain the vectorized line. The resulting lines are linked and further filtered according to the "Maximum angular difference", which controls the maximum angle between two linked segments, and "Maximum linking distance", which controls the distance allowed between two segments (Salui, 2018).

While testing this extraction method, various combinations (>100) and ranges in each parameter were applied. Based on a qualitative analysis of the different scenarios, three grids from each area were selected for further analysis (Figure 6-7; Table 6-1). Of these, scenario 2 (Figure 6-7C) for Dryden and scenario 3 (Figure 6-7D) for Timmins were used in subsequent stages of the workflow as they balance resolution as well as sensitivity to mapped geological units.

6.3.1.3. Post-processing

Extracted lineaments, in the form of vectorized polylines, were resampled and segmented in 200 m intervals using ESRI's ArcMap™ software. The average trend of each segment was then measured and used as input for later variance calculations.

Table 6-1. Values of parameter settings used for the three best lineament extraction results in each map area (40 m cell size for input aeromagnetic grid). Settings used to produce grids presented in Figure 6-7.

Variable	Unit	Dryden			Timmins		
		1	2	3	1	2	3
PPDRC highpass cutoff wavelength	Cells/m	72/2880	72/2880	72/2880	37/1480	72/2880	139/5560
Radius of filter	Pixels	15	20	25	7	15	5
Minimum edge gradient	Degrees	125	20	150	15	5	5
Minimum curve length	Cells	5	40	20	15	15	15
Maximum line fitting error	Cells	7	3	5	1	2	2
Maximum angular difference	Degrees	75	30	5	30	30	30
Maximum linking distance	Cells	5	30	10	2	3	2

As an additional step in post-processing, segments within 300 m of intrusive rocks were removed from the dataset (e.g., segments encompassed by the 300m buffer zones in Figure 6-7) as the primary focus for this study is to map variance related to deformation in volcanic and sedimentary packages rather

than local structural or metamorphic effects of intrusions. Only the orientation of segments that met such filtering criteria were used in the variance calculation
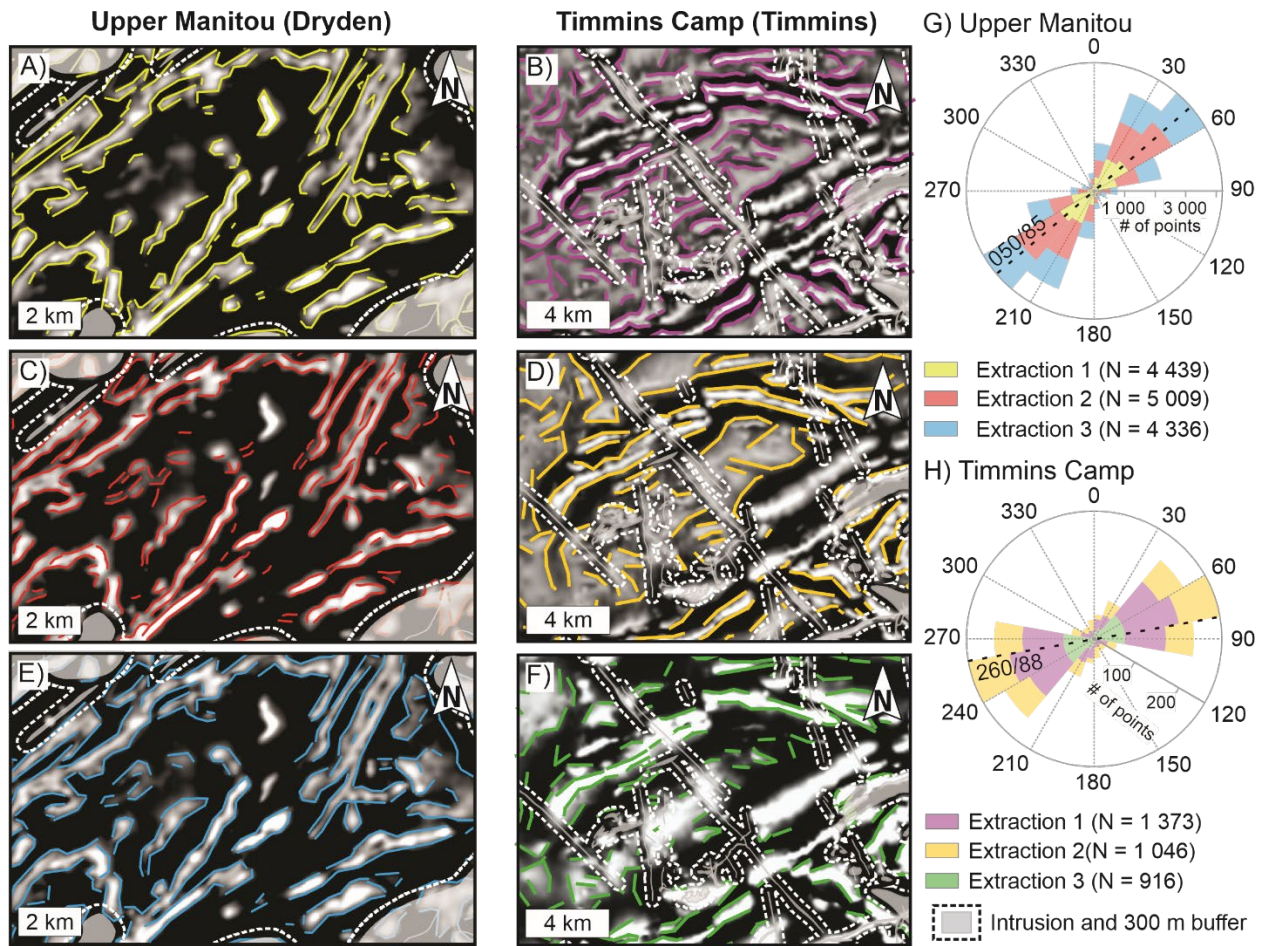
Figure 6-7. Lineament extraction results for (A, C, E) the Upper Manitou region near Dryden, Ontario and (B, D, F) the Timmins gold camp near Timmins, Ontario. Lineaments were extracted using a range of parameter settings (see Table 6-1); Lineaments within 300m buffer zones, indicated by white dashed lines, around mapped intrusions, indicated by transparent gray polygons, were discarded from subsequent calculations; Areas were selected as they host well-defined, kilometer-scale isoclinal folds and can be easily compared to representative bedrock geology maps in Figure 6-2; Rose diagrams of 200 m lineament segments in the G) Upper Manitou and H) Timmins camp areas display similar distributions to bedding in Figure 6-2 and are roughly parallel to representative fold axial planes. Population size for each extraction result is indicated as N; For the Upper Manitou areas, the axial plane was calculated using pi analysis (050/85; strike/dip) is superimposed on the rose diagrams. For the Timmins camp area, the axial plane of the Porcupine Syncline is interpreted to be the mean axial planar foliation value measured in field investigations (260/68; strike/dip); The similarity in orientation indicates that the extracted lineaments are a good approximation of primary layering and can be used in structural complexity analysis; The location of both regions is indicated in Figures 6-2G and 6-2H; Coordinates in NAD 83 UTM zone 15N in Dryden and Zone 17N in Timmins.

6.3.2. Mapping structural complexity

To map structural complexity, circular variance (in 2D) and spherical dispersion (in 3D) of primary layering in volcanic and sedimentary stratigraphy was calculated. The trend of lineament segments and the strike of bedding measurements were inputs for circular variance while strike and dip from bedding measurements were inputs for spherical dispersion. The following sections describe how each calculation method was applied to both input datasets. For outcrop bedding measurements, both variance and dispersion were performed as >90% of the compiled datasets record both the strike and dip, permitting the calculation of the dispersion of the down dip vector. For magnetic lineaments, only strike was measured, and thus, only variance was calculated.

6.3.2.1. Variance

Circular variance (Davis, 2002) was calculated for neighbourhoods around each point in the database using Fisher statistics (Fisher, 1959; Fisher et al., 1987). Neighborhoods are defined as circular regions within a fixed radius centered on each data point. Calculations determined variance for each neighborhood using radii from 250 to 70000 m and a minimum population of n > 3. Variance (V) is the average square deviation of all strike observations in radians ($\theta_1, \theta_2, \theta_3 \ldots. \theta_i$) from the population mean (μ) and is represented by Equation 1, where n equals the number of points in the neighborhood.

$$V = 1 - \frac{\sqrt{\left(\sum_{i=1}^{n} cos(2 * mod(\theta_i, \pi))\right)^2 + \left(\sum_{i=1}^{n} sin(2 * mod(\theta_i, \pi))\right)^2}}{n} \quad \text{(EQ 1)}$$

For this study, low variance ($V \approx 0$) represents regions where bedding orientation is sub-parallel. Variance increases as bedding or lineament strikes display a large relative difference to one another, such as around fold hinges or across fault offsets ($V \approx 1$).

Variance was calculated for each point in the dataset and the resulting values were individually assigned to those points. To create a map of complexity, variance was interpolated among the data points using the minimum curvature method, a 250 m grid cell size, and 1000 m blanking distance in Oasis Montaj™ (Figure 6-8A and 6-8C).
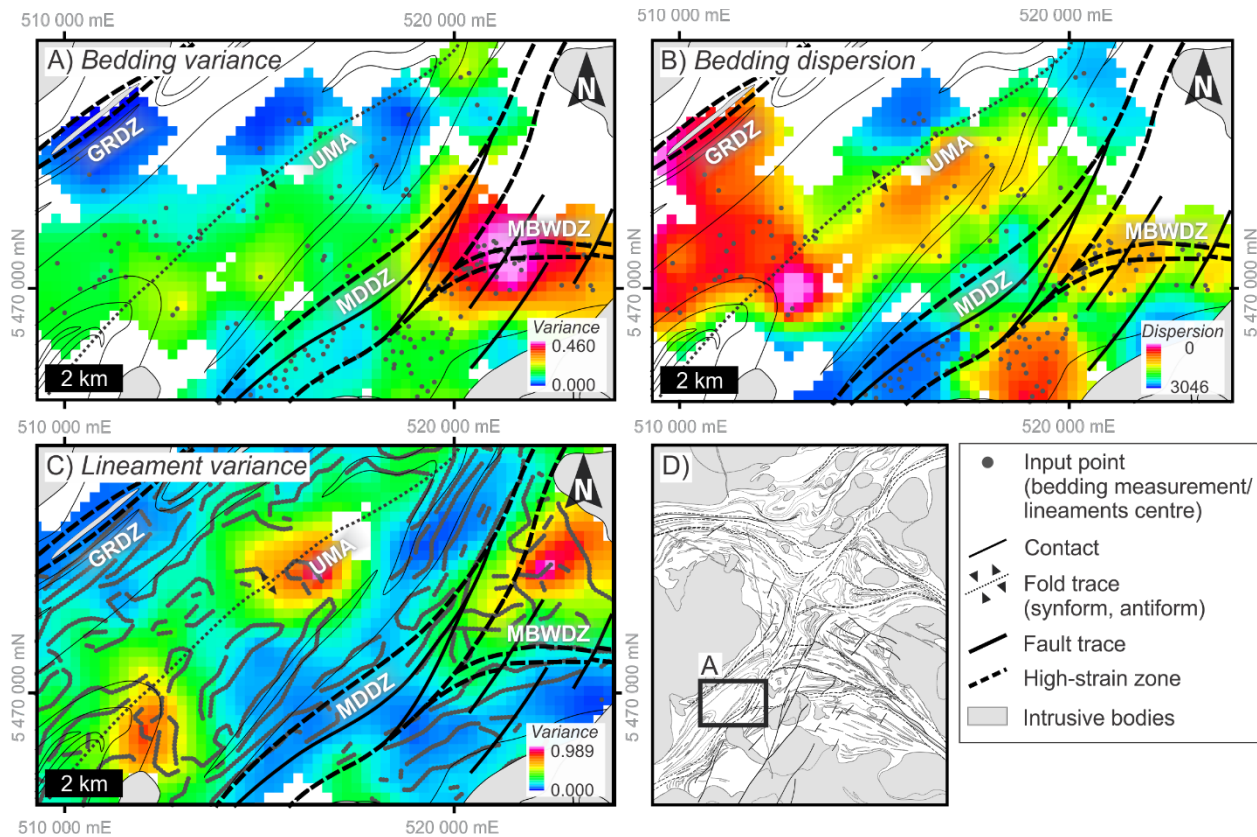
Figure 6-8. Comparison of structural complexity results for A) variance of bedding strike, B) dispersion of bedding strike/dip, and C) variance 200 m lineament segments; variance and dispersion calculations were performed using optimal parameters reported in Table 6-3; Grided color range used a standard deviation (n = 2.5) rescaling with a 0.5 gamma stretch (see online version for color); Coordinates in NAD 83 UTM zone 15N.

6.3.2.2. Dispersion

Dispersion calculations (Davis, 2002; Middleton, 2000) were performed using Fisher statistics (Fisher, 1959; Fisher et al., 1987) to test how significantly a second orientation component (i.e., dip) may affect the quantification of structural complexity. Dispersion was calculated in two stages. First, the strike (S) measurements were converted to dip direction (i.e., azimuth: $A = S + \pi/2$). Second, the down-dip vector that lies within the bedding plane was quantified using the orientation of the vector is defined by its north (x), south (y), and vertical (z) coordinates with the length of the vector defined by its vertex. These variables are expressed using Equations 2, 3, and 4, where D is dip of bedding and A is dip direction.

$$x = cos \, cos \, D \, cos \, cos \, A \qquad \text{(EQ 2)}$$
$$y = cos \, cos \, D \, sin \, sin \, A \qquad \text{(EQ 3)}$$
$$z = sin \, sin \, D \qquad \text{(EQ 4)}$$

Only bedding measurements with a recorded strike and dip were used to calculate spherical dispersion. As with variance, a range of neighborhood radii from 250 m to 70000 m and a minimum population of n > 3 were used. From each neighborhood population, the spherical mean direction is defined by the length of the resultant vector® using Equation 5. The dispersion (k) relative to nearest neighbour R vectors can then be compared using Equation 6.

$$R = \sqrt{\left(\sum x_i \right)^2 + \left(\sum y_i \right)^2 + \left(\sum z_i \right)^2} \qquad \text{(EQ 5)}$$

$$\kappa = \frac{(n-2)}{(n-R)} \qquad \text{(EQ 6)}$$

Similar to circular variance, spherical dispersion was calculated for each point in the dataset and the resulting values were individually assigned to those points. Dispersion values were interpolated using the minimum curvature method with a 250 m grid cell size and 1000 m blanking distance in Oasis Montaj™ (Figure 6-8B).

6.3.3. Validation using the Kolmogorov-Smirnov test

The two-sample Kolmogorov-Smirnov (K-S) statistical test in Python (Kolmogorov, 1933; Smirnov, 1948; Davis, 2002) was used to verify that results reflect real phenomenon rather than a random distribution of values. This was done by quantifying the similarity between results and a randomly generated, normally distributed synthetic population of the same size and range. For results from each permutation of input parameters, the maximum difference (Δ) between the cumulative distribution function of the result and a randomly generated dataset were compared. According to the null hypothesis, if Δ is greater than the critical value (Δ α), the null hypothesis can be rejected meaning the variance or dispersion result does not resemble a random distribution. With this logic, the null hypothesis was rejected for all results using the 0.01% critical value, indicating a very small probability (0.01%) that variance and dispersion observed was random (See Table 6-2). A rejected hypothesis means all datasets passed the K-S test in Table 6-2. The critical value (α) was calculated using Equation 7 (Knuth, 1998) where α denotes the confidence level, c(α) is the corresponding coefficient for α, and n1 and n2 represent the population size of the calculated and random populations, respectively.

$$\Delta \alpha = c(\alpha) \sqrt{\frac{n_1 + n_2}{n_1 n_2}}$$

Table 6-2. Results of performing the Kolmogorov-Smirnov (K-S) test on variance and dispersion results for the Dryden and Timmins map areas; If the K-S value (Dα) is greater than the critical value for a given confidence interval (95% and 99%), then the null hypothesis is rejected and the dataset passes the K-S test; 'Pass' results indicate that the variance or dispersion result does not resemble a random normal population of the same size and range.

Location	Input	Calculation	Δα	95% Critical value	95% Confidence	99% Critical value	99% Confidence
Dryden	Bedding	Variance	0.17	0.06	PASS	0.07	PASS
	Bedding	Dispersion	0.33	0.06	PASS	0.07	PASS
	Lineament	Variance	0.03	0.01	PASS	0.01	PASS
Timmins	Lineament	Variance	0.06	0.02	PASS	0.02	PASS

6.3.4. Parameter sensitivity testing

When performing variance and dispersion calculations on both bedding measurements and magnetic lineament trends, several parameters were tested to determine which were most appropriate given available data and geological constraints. These include varying the maximum neighbourhood population size (Figure 6-9), neighborhood radius (supplementary material Figure 6-1.2), segment length of lineaments (supplementary material Figure 6-1.3), high pass wavelengths from the PPDRC filtering technique (supplementary material Figure 6-1.4), input grid resolution (supplementary material Figure 6-1.5), minimum number of points required in a neighborhood, and minimum number of points with respect to data density (Figure 6-9). The values tested and optimal results for each parameter are listed in Table 6-3.
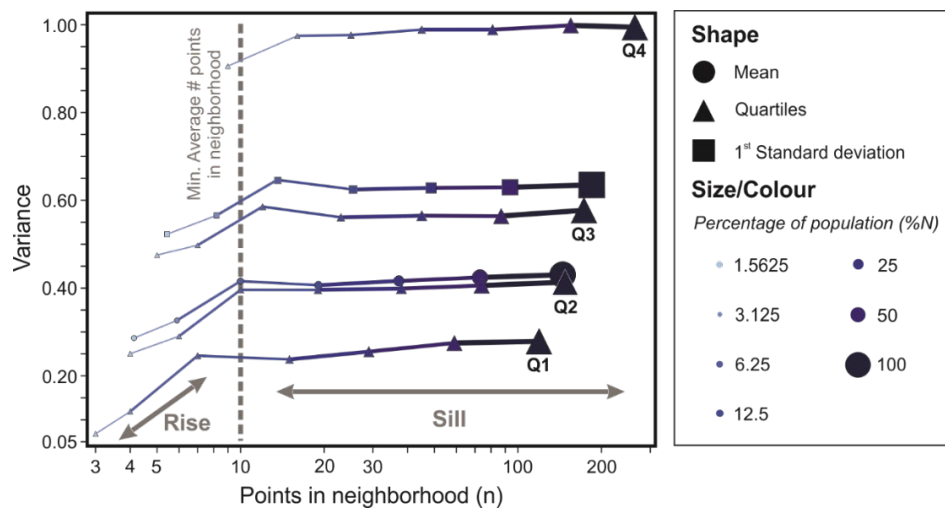
Figure 6-9.Cross plot of neighborhood population size versus variance for subsets of 100 m segments of autodetected lineaments. The inflection point between the rise and sill of summary statistics such as the mean (circle) quartiles (triangle), and first standard deviation (square) indicate the minimum average number of points within the 1750 m neighborhood radius; The global population of lineament segments was sequentially reduced by 50% to test minimum point density for successful structural complexity analyses. Color and size of symbol indicates the percentage of points from the original global population used for structural complexity analysis.

6.3.4.1. Neighborhood radii

Neighborhood radii for variance and dispersion calculations from 250 m to 70000 m were tested. A density cross-plot and summary statistics from 250 m to 10000 m radii results are plotted against variance in Figure 6-9 to determine the optimal neighborhood radius. From 10000 m to 70000 m radii, changes are negligible. Summary statistics such as quartiles (triangles), mean (circles), and standard deviations (squares) for lineament results are plotted in Figure 6-10 to highlight data trends and illustrate how the optimal radius value was determined. The distribution of these variables follows an exponential function. This likely reflects variance in small neighborhood radii which is commonly low since included data points are spatially and structurally related. As distance increases, data points from adjacent structures and, perhaps, different structural domains were included in the variance calculation. However, at some point, values converge along a sill since strike cannot exceed 360° (and dip cannot exceed 90°). Due to the heterogeneous spatial distribution of field measurements, in some cases, it was challenging to graphically determine which was the optimal neighborhood radius for bedding measurement results. Instead, the optimal search radius was visually determined based on a balance between resolution, spatial association with interpreted geological structures, and noise (see supplementary material Figure 6-1.1). For lineament variance results where data was homogeneously distributed, an optimum of 1750 m was identified. This optimal distance was selected as the inflection point between the rise and sill of summary statistics in Figure 6-10 (i.e., where data trends flatten). The sill represents the point where increasing the radius no longer statistically changes the data distribution, it will simply smooth results. Since the inflection points between the rise and sill are not exactly the same for all trendlines, a range of neighborhood radii (1250 – 2000 m) could be considered optimal and the median value (1750 m) was used for further calculations.

Table 6-3. Parameters tested throughout the structural complexity analysis workflow organized by step according to Figure6- 6. For each parameter, associated values and selected optimal ranges are also reported.

Step	Sub-step	Parameter	Tested values	Optimal
0 - Input		Grid resolution	40 m, 200 m	40 m
1 - Gridding	Filtering with PPDRC	Wavelength	400, 760, 1 480, 2880, 5560, 10720, 20720, and 40000 m	2880 m (Dryden) and 1480 m (Timmins)
2 - Lineament extraction	Edge detection and thresholding	See Table 6-1	See Table 6-1	See Table 6-1
	Post-processing	Segment length	100, 200, 500, 1000, 2000 m	200 m
3 – Calculate	Variance/Dispersion	Minimum neighborhood population size	3, 5, 7, 9, 11, 13, 15	3
3 – Calculate	Variance/Dispersion	Neighborhood radius	250 to 3000 (250 m intervals) 3000 to 5000 (500 m intervals) 5000 to 10000 (1000 m intervals) 10000 to 40000 (5000 m intervals) 40000 70000 (10000 m intervals)	1750 m
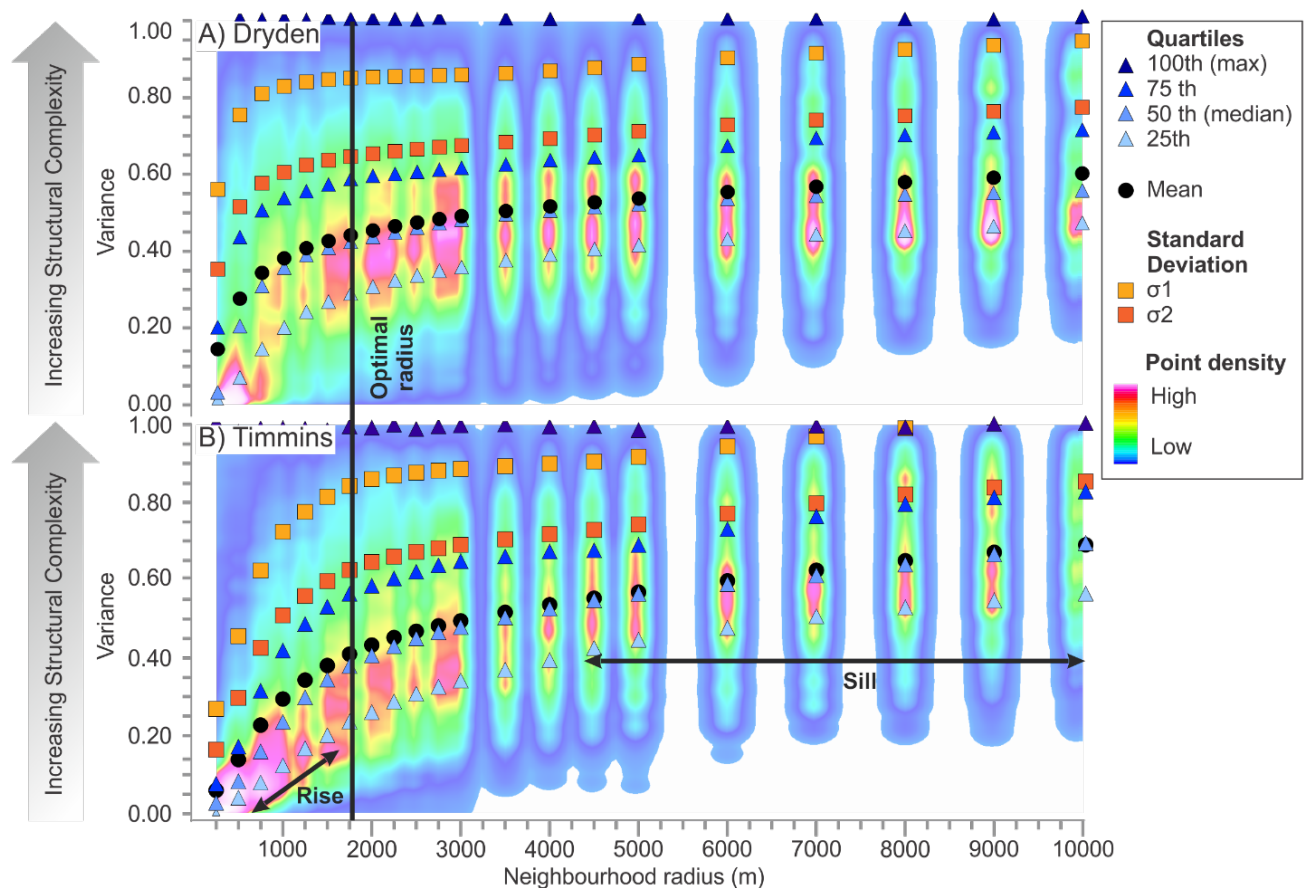
Figure 6-10. Cross plots displaying the point density for neighborhood radius versus lineament variance for the A) Dryden and B) Timmins study areas; Solid line indicates the optimal neighborhood radius (1750 m), selected based on the inflection point between rise and sill of trends in summary statistics (e.g., quartiles, mean, standard deviations); Statistical parameters plotted include quartiles (triangles), mean (circles), and 1-sigma and 2- sigma standard deviations (squares).

6.3.4.2. Segment lengths

Increasing the segment lengths of lineaments (100, 200, 500, 1000, 2000 m) had minimal effect on results (supplementary material Figure 6-1.3). Results from all tests displayed variance anomalies in the same location; however, there was some minor smoothing and decrease in resolution with increasing segment lengths. To balance resolution with computational expenditure, 200 m segment lengths were selected as optimal; however, the difference in computational time was in the order of a few minutes.

6.3.4.3. PPDRC wavelengths

The wavelength used in PPDRC filtering of aeromagnetic grids is a significant parameter for characterizing geologic trends. A range of wavelengths were tested (400, 760, 1480, 2880, 5560, 10720, 20720, and 40000 m; supplementary material Figure 6-1.4) and compared to mapped structures as well as geological interpretations. The range of tested wavelengths were chosen by selecting eight equal

intervals between a minimum (400 m) and maximum (40000 m) range. Based on these results it was concluded that smaller wavelengths (ideally 1000 and 3000 m) better characterize geological features, provide the highest resolution, and minimize the effects of intrusions; however, extremely small wavelengths (<1000 m) introduce a significant amount of noise. For Dryden, where there are few diabase dikes and moderate surface exposure, a wavelength of 2880 m was selected as an optimal balance between resolution and noise. For Timmins, where diabase dike swarms are prevalent and thick glacial overburden attenuates geophysical signals, a smaller wavelength of 1480 m was selected to minimize the effect of intrusions and to amplify subtle trends.

6.3.4.4. Grid resolution

Anomalies and trends resulting from varied grid resolutions (40 m and 200 m line spacing; supplementary material Figure 6-1.5) were generally co-located; however, results from the higher resolution survey (40 m) display stronger spatial links to mapped geology, especially in fold hinges, while low resolution results display a larger number of broad anomalies without geologic links at this scale. The input grid resolution parameter is significant in that higher resolution data captures and characterizes smaller trends while low resolution data averages signals. Thus, any results coming from the high resolution data are expected to have increased complexity whereas low resolution results should be smoothed and possibly offset as a function of the averaging nature of regional geophysical surveys. Since 40 m surveys were available in both map areas, they were used for structural complexity analysis.

6.3.4.5. Neighborhood population

A range in minimum neighborhood population (n) values (3, 5, 7, 9, 11, 13, 15) were also tested. In all results, variance anomalies were co-located and similar in form. However, increasing minimum neighborhood populations essentially down-sampled the global population such that fewer points were retained and gridded, thereby limiting surface coverage. To maximize coverage, a minimum neighborhood population value of 3 was selected as optimal.

6.4. Results

Anomalies from regional structural complexity grids (Figure 6-11) are co-located with interpreted geological features such as shear zones, faults, and fold axes (Figure 6-12). In general, variance results in both study areas display high values along fold axes and broad low variance anomalies spatially associated with ductile deformation zones. Brittle structures such as (late) faults mark boundaries between adjacent high and low anomalies (Figures 6-11 and 6-12).
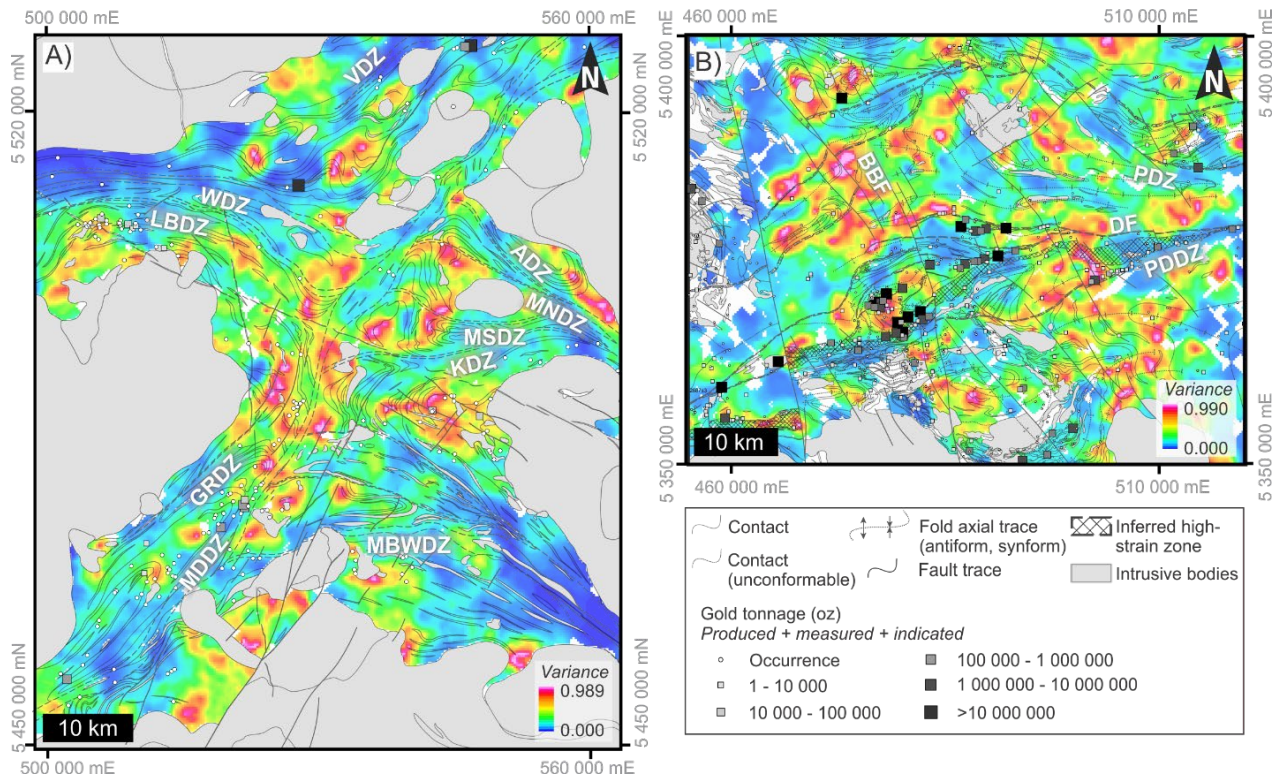
Figure 6-11. Interpolated variance results of auto-detected (200 m) magnetic lineaments with superimposed geological interpretations and known gold occurrences by tonnage for the A) Dryden and B) Timmins study areas; Variance is symbolized using a standard deviation (n=2.5) rescaling with a 0.5 gamma stretch (see online version for color); Coordinates in NAD 83 UTM zone 15N for Dryden and 17N for Timmins.

6.4.1. Kilometer-scale folds

Variance analysis of both the bedding and magnetic lineament data shows a systematic spatial correlation with kilometer-scale folds. Nearly every mapped fold axial trace is associated with high variance anomalies (variance > 0.6). Representative results from domains with folds in both study areas are displayed on Figure 6-12. These include the Upper Manitou antiform (Figure 6-12A), Long Lake River antiform (Figure 6-12B), and Thunder Lake antiform/synform (Figure 6-12E) that each correlate to high variance results. Some of these also represent refolded folds ("1" in Figure 6-12C) and/or drag folds ("2" in Figure 6-12A), which are also spatially associated with high variance.

6.4.2. Shear zones

Broad, linear areas of low variance are spatially associated with mapped shear zones. These zones differ from fold limbs as they display low variance anomalies along a laterally extensive trend (> 10 km long and >2 km wide). Examples of low variance related to known shear zones are the Manitou-Dinorwic deformation zone (Figure 6-12A), Kawashegamuk deformation zone (Figure 6-12B), Vermillion

deformation zone (Figure 6-12E), and the PDDZ (Figure 6-12C and 6-12F). Additionally, a broad, linear-shaped low variance trend is observed below the northern extent of the Porcupine basin ("3" in Figure 6-12I).

Commonly, deformation (shear) zones define boundaries between structural domains and, thus, separate domains of high and low variance. For example, the high variance associated with folds south of the Melgund North deformation zone ("4" in Figure 6-12D) is truncated by a low variance zone associated with a shear zone to the north.

6.4.3. Faults

Faults and other brittle structures are indirectly identified as they often act as boundaries between high and low variance anomalies or are observed to offset anomalies. The best examples of this are observed along northwest-trending faults in the Long Lake River antiform ("5" in Figure 6-12B), where high variance associated with fold axes are contained and offset by the fault traces. Another example is the north-northwest trending Burrows Benedict Fault (BBF; Figure 6-12C), which separates high variance spatially associated with fold traces in the west and low variance in the east.

6.4.4. Unidentified Dikes

As a part of lineament extraction, several previously unidentified and/or unexposed likely Proterozoic, diabase dikes were interpreted from the aeromagnetic grids in Timmins (Figure 6-7). These were identified as parallel lineaments that crosscut or were at a high angle to mapped volcanic or sedimentary bedding that are also spatially correlated to magnetic lineaments.

6.4.5. Au and variance

The spatial association between known Au mineralization and variance anomalies was quantitatively assessed (Figure 6-13) to determine if variance could be used as a proxy for prospective fluid pathways and traps in mineral exploration applications. The distance between high variance anomalies ($>\mu + 1\sigma$; Dryden: 0.598; Timmins: 0.613), which are co-located with interpreted fold hinges, and mineralization was calculated.

The distance between Au mineralization and low variance ($<\mu - 1\sigma$; Dryden: 0.206; Timmins: 0.189), which are interpreted to be deformation zones and fluid pathways where structures and rock units are transposed into sub-parallel orientations, was also calculated.

Gold deposits and prospects are all located at less than ~1500 m from a high or less than ~1700 m from a low variance anomaly (Figure 6-13). When the distance to nearest high and low anomaly are plotted against one another for each deposit and occurrence, two parallel trends emerge, offset by 1500 m along both axes (e.g., grey polygons in Figure 6-13). Point density for deposits and occurrences highlight these trends for minor deposits and occurrences from the Dryden area and major deposits in Timmins. The largest deposits (Dome, Hollinger, McIntryre, Porcupine, Pamour) are all located adjacent (<1000 m) to either a high or low variance anomaly.

Along both trends, Au grade increases towards the x-axis (closer proximity to variance high anomalies and structural traps): i.e., the highest-grade deposits (>7 g/T) such as Hollinger, McIntyre, and Hoyle Pond occur within or directly adjacent to (<500 m) high variance anomalies (e.g., within regional fold hinges); the lowest grade deposits (<3 g/T) such as Dome and Porcupine are within or adjacent to (<500 m) low variance anomalies (e.g., the PDDZ).

Figure 6-12 (next page). Variance grids with superimposed geological interpretations for representative regions in the Dryden (A, B, D, E) and Timmins (C and F) study areas. Locations indicated on simplified maps for G) Dryden and H) Timmins; All grids display calculated variance of magnetic lineaments (200 m) using a 1750 m neighborhood radius; These representative regions demonstrate a spatial association between high variance and fold hinges (A, B, C, and E), low variance and deformation zones (A, B, C, D, E, F, and I), faults as boundaries between anomalies (B, C, and D); Structural complexity grids are symbolized using a standard deviation (n=2.5) rescaling with a 0.5 gamma stretch (see online version for color); Coordinates in NAD 83 UTM zone 15N for Dryden and 17N for Timmins; (ADZ) Aiabewatik deformation zone, (BBF) Burrows Benedict Fault, (DF) Dome Fault, (FLS) Finlayson Lake Synform, (GRDZ) Gold Rock deformation zone, (KDZ) Kawashegamuk deformation zone, (LLRA) Long Lake River antiform, (MDDZ) Manitou-Dinorwic deformation zone, (MNDZ) Melgund north deformation zone, (MSDZ) Melgund south deformation zone, (PDZ) Pipestone deformation zone, (PDDZ) Porcupine-Destor deformation zone, (SLA) Suzanne Lake antiform, (TLA) Thunder Lake antiform, (TLS) Thunder Lake Synform, (UMA) Upper Manitou antiform, (VDZ) Vermillion deformation zone.
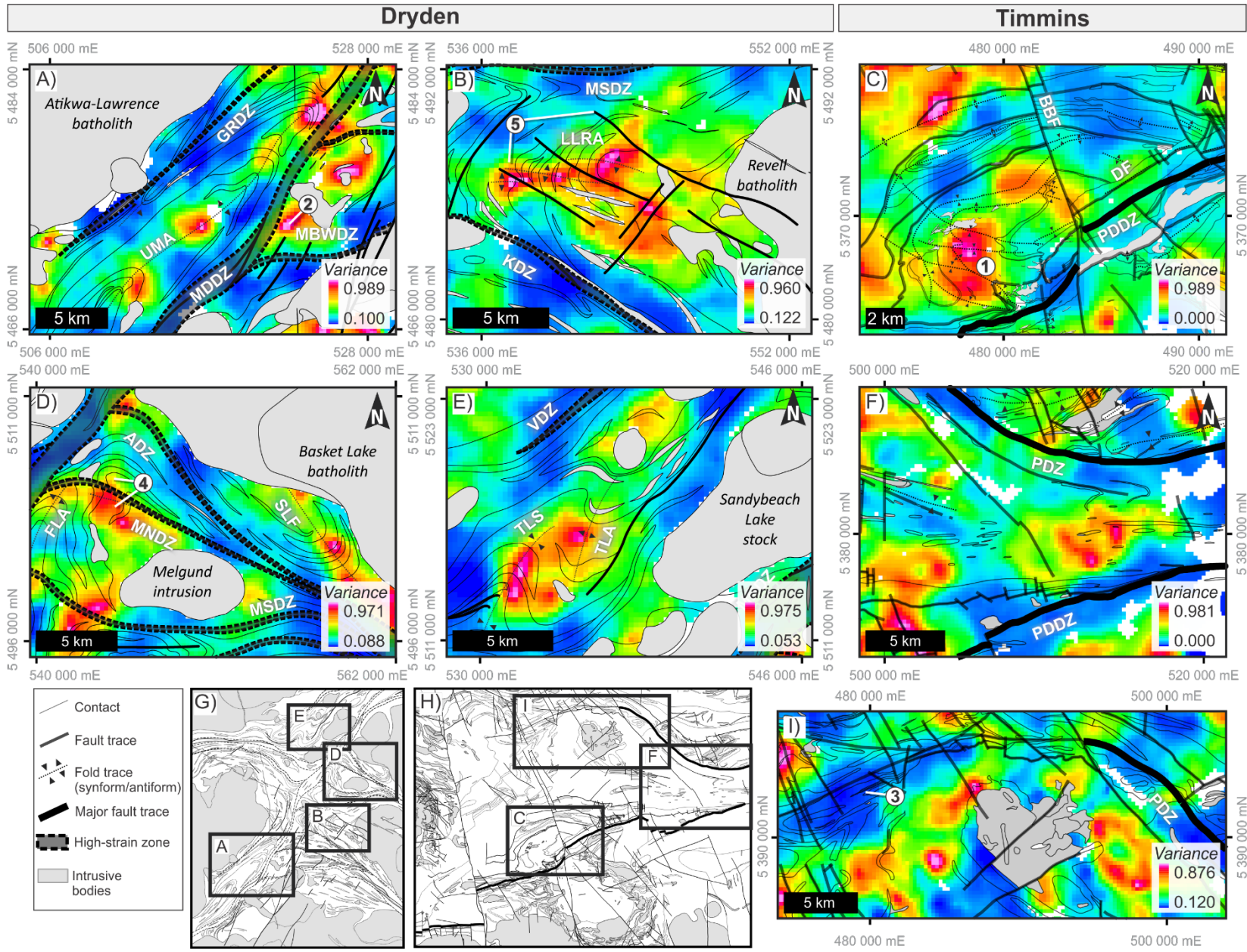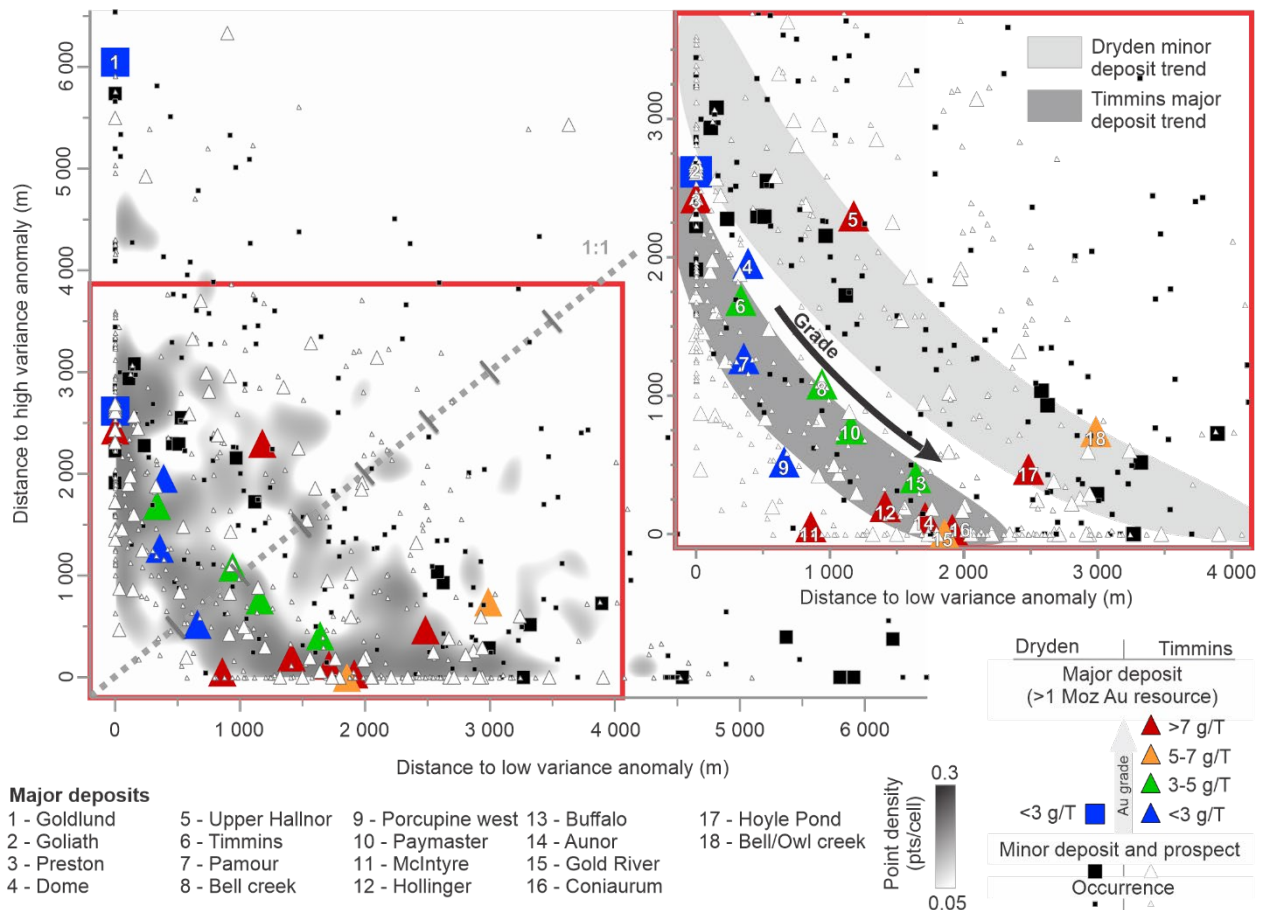
Figure 6-13. Scatter plot of distance to nearest low aeromagnetic lineament variance anomaly (< μ - 1σ) versus distance to nearest high variance anomaly (> μ + 1σ) for major deposits (Au resources > 1 Moz), minor deposits (deposits and prospects < 1 Moz Au resource), and occurrences. Gold mineralization associated with known deposits for the Dryden (squares) and Timmins (triangles) area are displayed. Mineralization data are overlain on a point density plot that highlights two trends, one for each map area that are offset by ~1500 m relative to one another (grey polygons in inset diagram). The gold grade of major deposits increases towards the x-axis along these trends.

6.5. Discussion

6.5.1. Geological significance

Statistical analyses and interpretation of remotely sensed data (e.g., geophysics, spectral imagery, digital elevation model, etc.) is becoming common practice (e.g., Gunn et al., 1997; Eaton and Vasudevan, 2004; Feumoe et al., 2012; Coulter et al., 2017; Baratoux, 2019). However, linking anomalies and results to geological features can be challenging (e.g., Aitken and Betts, 2009; Blundell et al., 2019). As stated previously, environmental, anthropogenic, geological, and many other factors can obscure geological, geophysical, and geochemical signals. The results of the structural complexity analysis presented here display a robust spatial correlation to geological features, indicating that circular variance of aeromagnetic lineaments and dispersion of bedding measurements provide quantitative tools to characterize structure in regional scale mapping

programs. These methods allow for the calculation of additional quantitative metrics that relate structure to anomalous concentrations of precious metals and are discussed in detail below.

6.5.1.1. Kilometer-scale folds

Refined interpretations of fold axial traces were constrained by high variance anomalies, especially in the Dryden map area. Given the isoclinal geometry of folds, bedding along fold limbs is sub-parallel and variance increases with proximity to hinges. Therefore, high variance values in fold hinges would be expected and is commonly observed (Figure 6-11). In some cases, high variance also correlates to fold interference patterns (e.g., Perrouty et al., 2017). For example, the anomalously high zone in Timmins ("1" in Figure 6-12C) highlights where early isoclinal folds are refolded by later z-folds and drag folding related to the PDDZ (Ayer et al., 2005; Bateman et al., 2008). High variance anomalies (>0.6) are also associated with hinges of drag folds, particularly where early folds are modified by obliquely cross-cutting, younger shear zones in the Dryden area ("2" in Figure 6-12A).

Structural complexity analysis also provides further geological insight into structures poorly defined by field mapping. For example, the location of the Long Lake River antiform axial trace in Dryden is well defined by the structural complexity results. This structure was originally interpreted by Kresz et al. (1982a, b) and later refined by Montsion et al. (2019); however, uncertainty in the location and geometry of the structure persisted, in part, due to dense vegetation and glacial overburden coincident with the feature. Using lineament variance results, the axial trace of the fold was extended to the west until it is offset by late northwest-trending faults.

Dispersion results, which capture 3D orientation variance of bedding, can also be used to enhance structural characterization of folds. For example, a linear low variance anomaly, parallel to the Upper Manitou antiform (Figure 6-8B), indicates dip direction of the fold based on its position southeast of the fold's axial trace. This interpretation of dip directions is supported by our field observations.

6.5.1.2. Shear zones and faults

Anomalous variance zones are sometimes bounded or offset by shear zones and/or faults. In some cases, the relative offsetting relationship can provide relative timing constraints on deformation features. For example, the Long Lake River antiform is offset by northwest-trending faults (#5 in Figure 6-12B), indicating that folding pre-dates faulting. Additionally, shear zones and/or faults form boundaries between adjacent structural domains. Linear variance tends across these boundaries differ as a result of offsetting (laterally or vertically) continuous stratigraphic packages or juxtaposing distal and unrelated structural blocks.

The width of linear-shaped low variance anomalies may also aid in defining the extent of high-strain zones associated with these structures. Traditionally, shear zones are represented as finite lines in geological maps; however, this practice oversimplifies the complex 3D nature of shear zones. This is due to the fact that shear zones affect volumes of rock where progressive strain may be localized into networks of anastomosing fault/slip planes (Sibson, 1977; Fossen and Cavalcante, 2018; Parsons et al., 2018). Shear zones, especially at crustal-scale, commonly record protracted phases of brittle-ductile deformation. Complexity, represented by variance or dispersion, within these deformation corridors is likely masked by intense transpressive strain along strike, as lineaments would have been passively rotated and/or sheared into sub-parallel orientations. The sub-parallel orientation of these strained structures appear as linear and low variance anomalies and the intensity of the low anomaly may be indicative of strain intensity. Numerical modeling, which is out of the scope of this contribution, would be required to fully test this relationship.

New geological insight and improved map interpretations related to low variance and deformation (shear) zones includes straightening and narrowing the southern extent of the MDDZ, based on low variance trends as well as the location of high variance in adjacent drag fold axis. The Gold Rock deformation zone (GRDZ) was also identified by a linear and low variance trend along a northeast-trending fold axis to the south of the Atikwa-Lawrence batholith (Figure 6-12A). Similarly, in the Timmins area, the Pipestone deformation zone (PDZ) is hypothesized to continue below the northern boundary of the Porcupine sedimentary basin. This continuation of the basal fault zone is indicated by a broad, linear, low-variance trend ("3" in Figure 6-12I). Due to sparse outcrop exposure, mapping the continuation of this structure has been challenging. Thus, application of the structural complexity method described may provide a means of tracking the fault below glacial cover outside of the present study area.

6.5.1.3. Implications for Au exploration

For both map areas, the spatial relationship between known mineralization and variance anomalies was tested (Figure 6-13). Commonly, primary, first-order faults/shears and associated higher-order faults or splays are prospective in an orogenic system as conduits for mineralized fluids ascending through the crust (Dubé and Gosselin, 2007; Goldfarb et al., 2005; Groves and Santosh, 2015; Robert et al., 2005; Robert and Poulsen, 2001). As a result, identification of crustal-scale structures and nearby higher-order splays is central to many prospectivity studies (e.g., Carranza and Hale, 2001; Bierlein et al., 2006; Joly et al., 2015; Zhang and Zhou, 2015). For both Dryden and Timmins, low variance anomalies are spatially related to known orogenic Au mineralization (Figure 6-13) and, therefore, this correlation may represent identification of major fluid pathways along which Au and related metals were deposited.

Gold mineralization is also commonly observed adjacent to high variance zones, which are interpreted to represent structural traps for fluids such as fold hinges (Figure 6-13). The close proximity (<2 km) between high variance anomalies and mineralization is expected as fluids and precipitated metals may be emplaced in structural traps such as fold hinges (Cox et al., 1991; Goldfarb et al., 2005). Furthermore, the grade of major deposits appears to increase with proximity to structurally complex zones, which may have experienced a protracted structural and hydrothermal history that may have included Au remobilization and, therefore, higher observed grade. However, economic parameters such as grade and tonnage are biased by date and method of mine development (parameter measurement) and will vary as new sections of deposits are exposed and extracted.

Patterns in deposit, prospect and occurrence density (Figure 6-13) highlight similarly oriented trends between both map areas, although an offset of 1500 m is observed between them. In Timmins, deposits, prospects, and occurrences are uniformly distributed between high and low variance anomalies. In contrast, the distribution of prospects and occurrences in the Dryden area is polarized with clusters proximal to either high or low variance anomalies. Since major deposits are located adjacent to both high and low variance anomalies as opposed to a strong correlation in Timmins, the Dryden area may have experienced a different structural and hydrothermal evolution. Therefore, further research into these aspects of the Dryden region is recommended.

6.5.2. Limitations and recommendations

Applications of this technique to improve geological and geophysical mapping should consider the influence of input data, method, analytical parameters, and geologic setting on results. Below, the major influences of these parameters are discussed in detail.

6.5.2.1. Data availability, density, and quality

Several issues affect data availability and quality of field-based geological maps and their derivative products. These include vegetation cover or severe weathering, which frequently limit outcrop exposure, resulting in locally disparate field measurements, and undersampling of regional-scale structures. Geophysical data such as aeromagnetic grids are biased by survey parameters, environment, anthropogenic effects, and geological settings (e.g., sedimentary cover, primary magnetite content, hydrothermal alteration, etc.). Thus, a balance must be struck between the noise associated with high resolution surveys and the averaging or undersampling effect of coarse surveys.

Along with data availability, data density is a significant consideration. To identify a desirable minimum data density for the structural complexity analysis outlined herein, 100 m lineament segments were sequentially subsampled by 50%, with variance calculated using optimal parameters

in Table 6-3, and derivation of summary statistics for each subsample (Figure 6-9). These results yield quantitative trends for quartiles, mean, and standard deviation, which increase towards a sill defined by a plateau. The inflection point between the rise and sill is interpreted as the minimum average number of data points for a given neighbourhood radius (here 1750 m). Therefore, assuming a roughly uniform distribution of points, a minimum of two points per km2 is considered as the minimum data density for the application of this structural complexity mapping technique. Dryden bedding measurements have an average point density of 1.8 points per km2, which approaches this minimum threshold. However, the Timmins measurements yield an average density of 0.4 points per km2. In part, this relatively low point density is a result of thick glacial overburden that limits outcrop exposure in portions of the study area. In comparison, Perrouty et al., (2017) applied a similar technique near Malartic, Québec using a much higher density dataset of bedding measurements (average density of 7.7 points per km2) and determined that variance of bedding strike was spatially related to structurally complex zones and gold mineralization. The higher data density allowed for higher resolution correlations, as opposed to the larger regional view presented for Dryden and Timmins.

6.5.2.2. Petrophysical considerations

Lineament detection depends on magnetic properties, which are linked to mineralogy of lithologic units and magnetic susceptibility contrast between adjacent layers. If there is insufficient contrast between volcanic and/or sedimentary layers, strong magnetic remanence (e.g., Koenigsberger Ratio>1; Enkin et al., 2012), or interference from intrusions (i.e., dikes), bedding-related signals may be obscured during lineament extraction. In part, the Timmins study area was selected to test how the presence of dense diabase dike swarms would affect derivative results. During sensitivity testing, a range in PPDRC wavelengths were applied as well as a range in filter radii during lineament extraction. Smaller wavelengths and filter radii were found to be most effective in diminishing the effect of dikes that crosscut stratigraphy. With these parameters and the removal of lineaments within 300 m of interpreted dike polygons, structural complexity analysis in the variance of lineaments within the Timmins map area is able to characterize regional structural complexity.

6.5.2.3. Method selection
6.5.2.3.1. Variance

Structural complexity mapped by circular variance of bedding are generally co-located with interpreted regional structures (Figure 6-8A and supplementary material Figure 6-1.1a). When used as an input for variance, field bedding measurements are the most accurate and precise means of capturing primary lithological layering. However, bedding strike variance is disadvantaged by heterogeneously distributed and/or clustered data that is biased toward preserved outcrop

distributions as discussed above (section 6.3), which affects both statistical calculations and later grid interpolation as well as the time investment needed to create structural databases.

Sampling bias is prevalent in field-based data sets and, commonly, results in undersampling of regional structure. For example, if only the limbs of an isoclinal fold are observed and characterized, resultant variance will be low since isoclinal fold limbs are sub-parallel. To properly capture the fold, both limbs and hinge should be measured with sufficient data density (see above), or, alternatively, the spherical dispersion calculation may be preferable.

Variance of magnetic lineament trends (Figure 6-8C) provides a quantitative means to map structural complexity at regional scales. While aeromagnetic data are affected by intrusions or alteration, they are a continuous, homogeneously distributed dataset that minimizes sampling bias. At local scales (<1:30000), variance of aeromagnetic lineaments provides nearly complete coverage but lacks the definition and precision provided by bedding dispersion.

6.5.2.3.2.　　　Dispersion

Spherical dispersion of the down dip vectors of bedding planes (Figure 6-8B and supplementary material Figure 6-1.1b) was calculated to constrain how significant a second orientation component (i.e., dip) affects resultant structural complexity calculations. To evaluate the strengths and weaknesses of dispersion, a visual comparison to results from variance of bedding and lineaments yield was performed. In variance-based results, local anomalies and trends diverge from the dispersion results, particularly at high resolution, indicating that dispersion calculations provide the best representation of structural complexity at local (<1:30000) scales. Furthermore, these anomalies were the most strongly correlated with interpreted geological and geophysical features, especially in hinges of regional folds (Figure 6-8B). Thus, in specific instances, the dispersion results provided additional information not obtained by the 2D variance calculations such as dip direction of a fold's axial plane (e.g., to the southeast for the Upper Manitou antiform; Figure 6-8B).

At a deposit-scale or brownfields setting where data is sufficiently dense and representative of relevant structures, spherical dispersion calculation will likely provide the most robust input for highlighting structural complexity as it provides the best results at higher resolutions. Conversely, the dispersion calculations may under-represent structure at the regional scale where input data is irregularly sampled, sparse, and/or clustered.

6.5.2.4. Benefits and potential applications

The technique discussed here has applications beyond mapping structural complexity using bedding measurements and aeromagnetic data. For example, lineament detection and structural complexity

can be used to identify areas where further investigation is required to better constrain interpretations, fill gaps where field measurements are sparse (e.g., greenfield settings), to highlight buried or unidentified intrusions (e.g., dikes or clusters of porphyries), and can characterize subtle intrusive phases in large plutonic bodies. The application of structural complexity mapping may also have greater success in Phanerozoic terranes where shallowly dipping geological features are common (Bally et al., 2012), as opposed to characteristic steeply dipping Archean structures observed in Dryden and Timmins study areas.

Other input datasets such as digital elevation models (DEM), seismic or gravity surveys (in either 2D and 3D), electromagnetic surveys, and spectral data may also provide a means of capturing structure and complexity that leverage constraints on additional physical rock properties; however, each dataset will have specific use cases.

Most importantly, this technique has potential as a mapping tool for fluid pathways and traps in mineral prospectivity modeling. Frequently, inputs for prospectivity studies rely heavily on interpretations or knowledge because they integrate geological expertise with observations while providing the spatial continuity needed (e.g., Carranza and Hale, 2001; Bierlein et al., 2006; Joly et al., 2015; Zhang and Zhou, 2015). However, inputs constrained by knowledge-based features are influenced by compounding and unquantified biases that decrease confidence and accuracy of results. In orogenic gold exploration, structural complexity is often represented by overlaying buffers around surface traces of relevant linear structures (e.g., faults, folds) or in structure density maps (Knox-Robinson and Wyborn, 1997; Carranza and Hale, 2001; Bierlein et al., 2006; Kreuzer et al., 2015; Zhang and Zhou, 2015; Campos et al., 2017; Ford, 2019). While these inputs are useful, they lack the rigor of a purely data-driven, quantitative approach such as mapping variance and dispersion of structural measurements.

6.6.   Conclusions

The increasing influence of computational techniques on geological workflows has fueled the development of tools that improve geological knowledge and interpretations. The spatial distribution of structural complexity was mapped by applying Fisher statistics such as circular variance and spherical dispersion to bedding measurements and auto-detected magnetic lineaments. In general, anomalous results (high or low variance) and interpreted geological structures are co-located. Linear, low variance trends occur along shear zones, which likely reflect parallel and transposed structure in shear zones. These observations were used to refine interpretations, define zones of influence, and track shear and/or fault zones under cover. Zones of high variance and low dispersion correlate to areas with well-defined folds as constrained from field

investigations. Sharp boundaries between high and low variance anomalies are also interpreted to represent shear and/or fault zones that juxtapose structural blocks.

Gold resources are also spatially associated with adjacent low (within <~1700 m) and high (within <~1500 m; highest Au grade within <500 m) variance anomalies. In the Timmins study area, deposits, prospects, and occurrences are uniformly distributed between high and low variance anomalies; however, in the Dryden area, prospects and occurrences are polarized with clusters proximal to either a high or low variance anomaly. The difference in distribution may imply different structural or hydrothermal evolution, which may have led to contrasting mineral endowment.

Based on sensitivity testing, variance of magnetic lineaments was determined to be most effective in mapping structural complexity at regional scales. Aeromagnetic datasets provide uniformly distributed data points that more completely capture trends along regional structures. In contrast, field-based measurements that are commonly clustered tend to undersample regional structures (e.g., both limbs and hinge of folds). Conversely, given sufficiently dense data (e.g., 2 measurements per km2 in this study), most likely at a deposit scale, spherical dispersion calculations that integrate the down-dip vector of bedding planes may provide the most comprehensive means of capturing structural complexity and related variance calculations. This may also be applied to mapping structural complexity in 3D geological modeling

Key considerations when applying the methods presented here are scale, data availability/quality, data density, and optimization of input parameters (e.g., neighborhood radius, lineament segment length, minimum number of points in neighborhood). Each of these parameters should be considered on a case-by-case basis as they will strongly impact iterative decisions and the resultant structural complexity maps. Additionally, one must take caution since interpretation of variance is dependent on the data used to generate estimates. As a result, inferred zones of anomalous values may simply be due to high frequency noise (high variance) or suppressed signal (low variance) due to interference from local environments.

Aeromagnetic surveys of greenstone belts near Dryden and Timmins, Ontario has allowed us to map structural complexity using a range of parameters and input datasets. When compared to existing geological interpretations, results were used to refine and confirm existing geological interpretations and provide new insights into relative timing and significance of geological features. Ultimately, these results were used to identify potential traps for mineralizing fluids, which can be integrated into subsequent prospectivity analyses for orogenic gold systems.

Chapter 7. Development and application of feature engineered geological layers for ranking magmatic, volcanogenic, and orogenic system components in Archean greenstone belts

Key words

Abstract

Geologically representative feature engineering is a crucial component in geoscientific applications of machine learning. Many commonly applied feature engineering techniques used to produce input variables for machine learning apply geological knowledge to generic data science techniques, which can lead to ambiguity, geological oversimplification, and/or compounding subjective bias. Workflows that utilize minimally processed input variables attempt to overcome these issues, but often lead to convoluted and uninterpretable results. To address these challenges, new and enhanced feature engineering methods were developed by combining geological knowledge, understanding of data limitations, and a variety of data science techniques. These include non-Euclidean fluid pre-deformation path distance, rheological and chemical contrast, geologically constrained interpolation of characteristic host rock geochemistry, interpolation of mobile element gain/loss, assemblages, magnetic intensity, structural complexity, host rock physical properties. These methods were applied to compiled open-source and new field observations from Archean greenstone terranes in the Abitibi and western Wabigoon subprovinces of the Superior Province near Timmins and Dryden, Ontario, respectively. Resulting feature maps represent conceptually significant components in magmatic, volcanogenic, and orogenic mineral systems. A comparison of ranked feature importance from random forests to conceptual mineral system models show that the feature maps adequately represent system components, with a few exceptions attributed to biased training data or limited constraint data. The study also highlights the shared importance of several highly ranked features for the three mineral systems, indicating that spatially related mineral systems exploit the same features when available. Comparing feature importance when classifying orogenic Au mineralization

in Timmins and Dryden provides insights into the possible cause of contrasting endowment being related to fluid source. The study demonstrates that integrative studies leveraging multi-disciplinary data and methodology have the potential to advance geological understanding, maximize data utility, and generate robust exploration targets.

Acknowledgements

7.1.   Introduction

Geoscientific applications of machine learning mainly focus on model efficiency/sensitivity while applying undeveloped feature engineering methodologies to represent relevant geological features and processes. While improving modeling capabilities is necessary, geological plausibility of inputs is equally critical to generate robust results and gain meaningful information. Without geologically appropriate feature maps, results may be spurious or challenging to interpret. Herein, we demonstrate existing and new geologically representative feature engineering techniques designed to comprehensively capture components in magmatic, volcanogenic, and orogenic mineral systems controlling metal endowment. Comparisons of ranked feature lists generated by random forest models to each conceptual mineral system model are used to validate geological representativeness of feature engineering methods applied to two greenstone belts in the southern Abitibi (near Timmins, Ontario) and western Wabigoon (near Dryden, Ontario) subprovinces of the Superior Province.

Broadly, machine learning applications to geoscience consists of 1) posing a geological question, 2) gathering data, 3) applying geological knowledge to select relevant data identifying features/processes relevant to the given question and can be adequately captured by available data, 4) applying various methods to derive representations (i.e., feature engineering explanatory variables) from data, 5) selecting typical locations for classes (e.g., mineralized target vs background), 6) training a model to identify the combined characteristic signature from explanatory variables, 7/9) collect performance metrics and tools and evaluate model, and 8/9) predict and evaluate targets (prediction not performed herein (Figure 7-1).

Feature engineering methods that derive representations from integrated geological knowledge/data are sometimes oversimplifications of complex geology, thus embedding subjective biases that compound throughout iterative workflows. Common examples are distance/density of fault traces, fold axes, favorable lithological units (Knox-Robinson and Wyborn, 1997; Carranza and Hale, 2001; Bierlein et al., 2006; Kreuzer et al., 2015; Zhang and Zhou, 2015; Campos et al., 2017; Ford, 2019). In some cases, these derivative products are the only available option due to limited local information or data availability; however, combined data science and expert knowledge may provide a means to improve representations of relevant geology.

At the other end of the spectrum, purely data driven inputs for machine learning are used to train predictive models, especially unsupervised techniques. While using directly analysed/measured input features such as geophysical and remotely sensed hyperspectral data incorporates the least amount of human-related bias, outputs are often challenging to evaluate and interpret. Some level

of expert interaction is needed to filter convoluted signals and select relevant signatures. Most prospectivity studies use a combination of both ends of the spectrum (e.g., Brown et al., 2000; Rigol-Sanchez et al., 2003; Feltrin et al., 2016), but could benefit from further integration of geological knowledge and data science.
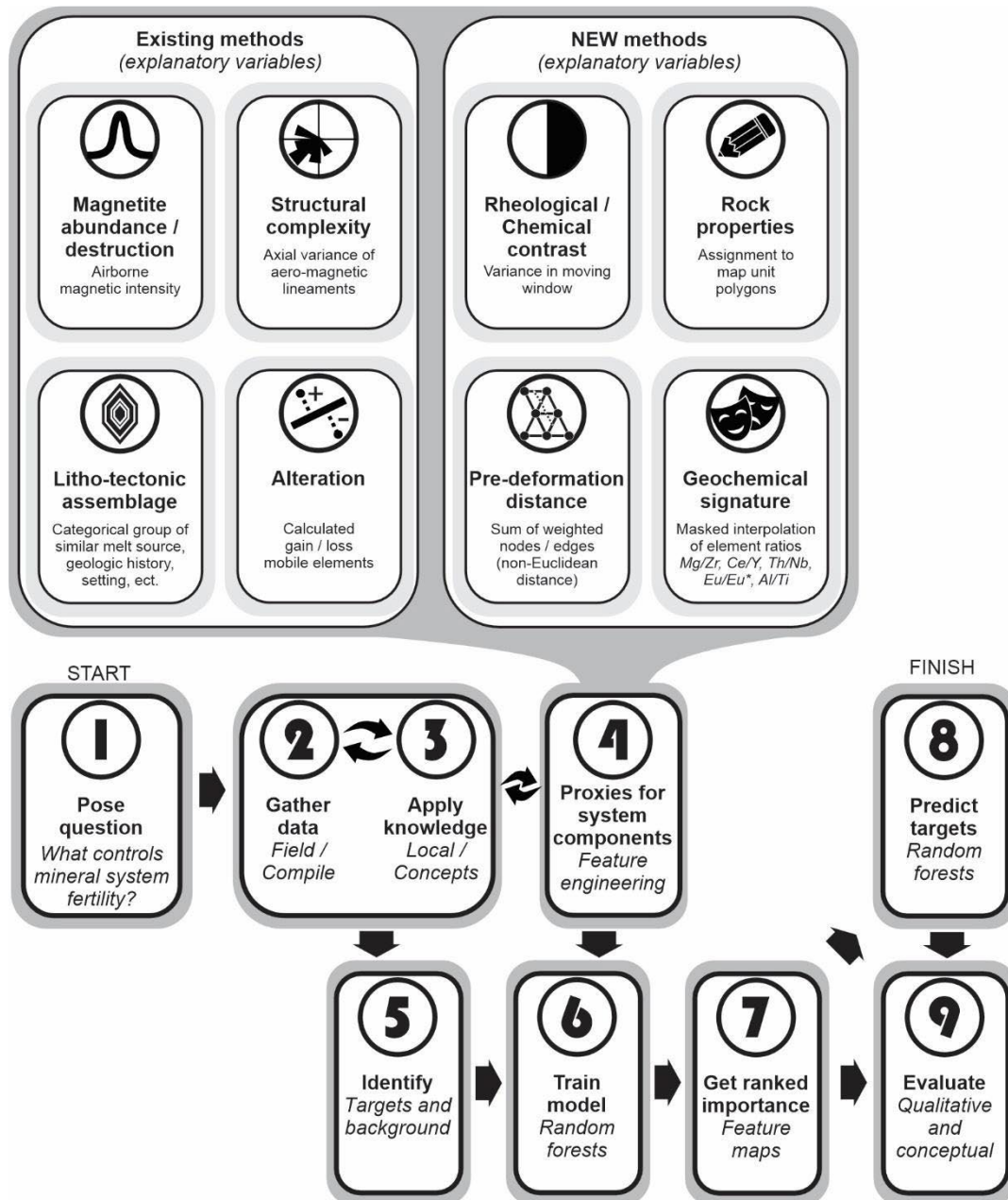
Figure 7-1. Workflow diagram of broad methodological phases and random forests application proposed here. Initial phases involve assessing the geological problem, gathering data, and using geological knowledge to attribute value to each dataset. Combined, this information is used to design relevant feature engineering techniques that isolate signals or generate representations of mineral system components. Eight new and existing feature engineering methods are indicated as an inset to the diagram. Target and background locations are synchronously identified and used to constrain training random forests classification and generate a ranked list of feature importance.

This study attempts to bridge the divide between oversimplified and pure-data input variables for geoscientific applications of machine learning. Special consideration of features/processes controlling mineral system fertility, their natural expression, data able/available to capture their expression, and data science techniques that approximate relationships/function constrained the design of feature engineering methods. Two Archean greenstone belts in the southern Superior Province near Timmins and Dryden, Ontario, Canada are compared to test the engineered variables. Geologic representativeness of the variables is evaluated by comparing conceptual models to importance ranking of features from random forests (Lindsay et al., 2022). Deviations from expected ranking may then guide knowledge and/or methodological development.

7.1.1. Mineral system models

Conceptual models are used to anticipate necessary data and selection of appropriate methods. Below, the components in magmatic, volcanogenic, and orogenic systems are described in terms of their dominant role (e.g., source, pathway, trap/deposit site/host, and footprint). A generalized summary of components and proposed methods of representing them are provided in Table 7-1.

7.1.1.1. Magmatic Ni-Cu-PGE

Current conceptual models for magmatic systems suggest that 1) metal-enriched magma source(s), 2) sufficiently strong flux to drive magma ascent, 3) connected and extensive plumbing systems (i.e., crustal-scale deformation zones), and 4) sulfur sources along magma pathways are the main controlling components to generate large Ni-Cu-PGE deposits (Naldrett, 1999; Barnes et al., 2007; Houlé and Lesher, 2011; Naldrett 2011; Barnes and Fiorentini, 2012; Le Vaillant et al., 2016; Lesher, 2019). While conceptual and practical understanding of these deposit types are continually evolving with new observations, the characteristics of existing components can be used to guide design of feature engineering methods.

Large deposits in the Yilgarn, Superior, and other Archean cratons are often cumulate-rich, magnesian komatiites that are slightly denser than nearby ultramafic/mafic volcanic magmas (Sproule et al., 2002; Naldrett 2011). In exploration, these prospective ultramafic rocks are identified geochemically by their relative abundance of incompatible trace elements (e.g., Th, U, Y), which indicate cumulate-rich fertile vs cumulate-poor infertile melt sources (Sproule et al., 2002; Davidson et al., 2007; Otamendi et al., 2016). Mechanisms that aid the ascent of these magmas may involve a stronger upflow where there is sufficient energy to sustain and transport large volume of magma (Campbell, 1989; Sproule et al., 2002), seismic pumping (Maier et al., 2016), roof foundering (Saumur et al. 2015), or mixing/assimilation of buoyant material (Wilson and Head 1981; Ripley, 1986; Lister and Kerr 1991).

The availability and incorporation of crustal sulfur during ascent is also considered to be a critical characteristic of fertile magmas (Nesbitt et al., 1979; Begg et al., 2010; Barnes and Fiorentini, 2012; Le Vaillant et al., 2016; Barnes et al., 2019; Lesher, 2019). Moderate crustal interaction/fractional crystallization signatures in ultramafic rocks is associated with the most fertile Ni-Cu-PGE camps, which indicates moderate rather than fast flow is preferable (Sproule et al., 2002; Begg et al., 2010; Barnes et al., 2012; Barnes and Fiorentini, 2012; Le Vaillant et al., 2016). Frequently, sulfur isotopes and enrichment in incompatible elements (i.e., Zr, Th, LREEs) are used as proxies for crustal contamination (Lesher et al., 2001; Le Vaillant et al., 2016) and thus opportunity for assimilation. Herein, Th/Nb is used as a proxy for crustal interaction/fractional crystallization (after Pearce, 2008; Montsion et al., submitted b/Chapter 5).

7.1.1.2. Volcanogenic Massive Sulfide (VMS) Cu-Zn-Pb-Ag(-Au)

Conceptual models for Volcanogenic Massive Sulfide (VMS) systems are variable based on age and geodynamic setting. In Archean greenstone belts of the Abitibi, volcanogenic deposits are typically observed in thin, juvenile bi-modal (mafic-felsic) volcanic crust and influenced by 1) a rift or back-arc geodynamic setting (extension/crustal thinning), 2) long-lived and shallow heat source(s), and 3) spatially extensive and connected fluid pathways (Franklin et al., 1981; Galley, 2003; Franklin et al., 2005; Hannington et al., 2005; Gibson et al., 2007; Piercy, 2010).

Characteristic Archean mafic volcanic rocks are low-Ti island arc tholeiites and back-arc basin basalts derived from depleted mantle source(s) with limited crustal contamination/fractional crystallization signatures (low Th/Nb and flat REE patterns; Pearcy, 2010). Common felsic magmas associated with volcanogenic deposits in Archean bi-modal volcanic settings are FIII rhyolites (Lesher et al., 1986), which are characterized by flat REE curves, negative Eu-anomalies, low Zr/Y, and intermediate to high abundances of high field strength elements. Geochemical signatures in these felsic rocks are indicative of modification in high-level magma chambers, which likely supplied some of the heat that contributed to hydrothermal convection in the system (Lesher et al., 1986). Other heat sources in fertile volcanogenic systems vary in composition and form; however, gabbro-diorite-tonalite-trondhjemite-granodiorite intrusion complexes are commonly associated with clusters of Archean volcanogenic deposits (Franklin et al., 2005; Galley, 2003; Galley et al., 2007).

Regardless of age or geodynamic setting, crustal-scale deformation zones are interpreted as critical features to focus large volumes of magma and fluid during convection and seafloor venting. Such structures are favorable if they were active for a prolonged period during volcanism, allowing for the steady accumulation of metal precipitates in the shallow crust and on the ocean floor (Campbell et al., 1981; Franklin et al., 2005; Galley, 2003; Lesher et al., 1986).

Table 7-1. Conceptual model of mineral systems used to design feature extraction. Models are defined in terms of feature used to represent a given system component, what primary data constrained feature extraction, method of extraction, and the feature's relevance to the given system. System components are grouped based on their principal but not exclusive role. Interm: Intermediate; SM: Supplementary Material

Feature			Primary data	Method	System component (s) and exploration criteria		
					Magmatic	**Volcanogenic**	**Orogenic**
Distance to	Deformation zone	Early E-W (crustal) / Early E-W (splay) / Mid-deformation NE-SW	Field observations, historical interpretations, aeromagnetic maps	Non-Euclidean distance from map interpretation	**Pathways:** Trans-crustal structures and splays	**Pathways:** Possible syn-volcanic trans-crustal faults	**Pathways:** Trans-crustal structures and splays **Traps:** Splays ± faults
	Basement	Orthogneiss / Para-gneiss			N/A	N/A	Variable based on P-T conditions
	Igneous — Volcanic Intrusion	Ultramafic / Mafic to interm. / Interm. to felsic / Ultramafic to mafic / Mafic to interm. / TTG to felsic			**Source:** Ultramafic magma	**Sources** (heat): Shallow syn-volcanic intrusive bodies **Trap/Host:** Bi-modal (mafic and felsic) volcanism	**Sources** (metal, fluid, heat): Variable **Traps** (chemical and structural): Varies by rock characteristics **Host:** Variable metamorphic protolith
	Sedimentary	Clastic / Chemical			Possible sulfur **source**		**Source** (metals and S) **Physical trap:** Permeable/Porous material **Chemical trap:** Fe-rich material
Structural complexity		Calculated value	Magnetic intensity + Map units	Variance of lineaments (excluding intrusions)	**Pathways:** Low complexity (corridors of transposed structures)	**Footprint feature:** Volcanogenic alteration promotes later deformation	**Pathways:** Low complexity (transposed structures) **Traps:** High complexity (fold hinges)
Assemblages			Map units and geochronological ages	Grid of assemblages	**Sources:** Lower/Older stratigraphy typically contain ultramafic vol. **Traps:** Upper/Younger assemblages overlay sulfur sources	**Traps/Host:** Assemblages with bi-modal (mafic and felsic) stratigraphy	Variable
Primary chemical signature		Mg/Zr / Ce/Y / Th/Nb / Al/Ti / Eu/Eu*	Compiled geochemical databases	Semi-discrete gridding: Minimum curvature interpolation for each map unit polygon	**Sources** (metal): Magnesian ultramafic (High Mg/Zr), cumulate-rich (low Ce/Y) melts with crustal interaction/fractional crystallization (low to intermediate Th/Nb)	**Sources** (metal)/ **Geodynamic setting:** Mafic (mod. Mg/Zr) and FIII (v. low Eu/Eu* & high Al/Ti) felsic (low Mg/Zr & Ce/Y) magmas with limited crustal interaction/ fractional crystallization (low Th/Nb)	Variable
Alteration/ Hydrothermal haloes		CO_2 / SiO_2 / K_2O / Na_2O	Compiled geochemical databases	Empirical Bayesian kriging of gain/loss normalized to least altered Mg/Zr (igneous) and Al_2O_3 (sed)	N/A	**Geodynamic setting:** CO_2 seafloor alteration **System Core: Acidic fluids leach** SiO_2 and/or diluted by S or metals **System Periphery:** K_2O gain & Na_2O loss with sericitization	**Deposit footprint:** CO_2 and SiO_2 gain from hydrothermal alteration
Physical properties		Mag. intensity	Airborne collection	Dynamic range compression of RTP grid filtered by cosine roll off	**Source:** High Mag	**Pathways:** Low from destruction of magnetite by hydrothermal fluids	**Pathways:** Low from destruction of magnetite by hydrothermal fluids

				Sources: Characteristic values for ultramafic lithologies	Sources: Characteristic values for mafic and felsic volcanic rocks	Primary rock signatures variable/non-discriminatory (exception = BIF)
	Petrophysical *(total porosity, grain density, mag. susceptibility)* Ductile *(shear, bulk, compression moduli)* Brittle *(friction coefficient, cohesion, UCS)* Electrical *(resistivity)*	Average values from literature for each lithologic type (SM – Table 7-2.1)	1) Discrete assignment of average values to each lithology polygon 2) Individual properties scaled 0 to 1 before combining (sum)			
Contrast	Rheological	Maps of brittle and ductile properties	1) Calculate σ in moving 1 x 1 km windows on individual grids 2) Calculate variance (σ^2) 3) Individually scaled 0 to 1 4) Sum of grids	Pathway: Structures and zones of mechanical weakness	Traps: Pressure interfaces between lithologies with contrasting mechanical properties	Traps: Pressure interfaces between lithologies with contrasting mechanical properties
	Chemical	Element ratio maps (see above)		Non-discriminatory	Variable and relative based on properties of neighboring rock types	Variable and relative based on properties of neighboring rock types

7.1.1.3. Orogenic Au

Conceptual models for orogenic Au deposits are based on fluid/metal/heat source(s) in variable geologic settings. However, it is widely interpreted that fertile orogenic systems are the result of 1) shortening driving deformation, magmatism, and metamorphism to generate fluids and leach metals to be later 2) transported along pathways and 3) concentrated in traps (e.g., Robert and Poulsen, 2001; Robert et al., 2005; Dubé and Gosselin, 2007; Groves and Santosh, 2015; Groves et al., 2018).

Orogenic deposits are hosted by many metamorphosed rocks, including volcanic (e.g., Bateman et al., 2008; Dubé and Mercier-Langevin, 2020), intrusive (e.g., Robert et al., 2001; Katz et al. 2017), and sedimentary (e.g., De Souza et al., 2020; Pitcairn et al., 2021), meaning there are no globally characteristic host rock characteristics to direct exploration strategies. Locally, at camp scales, some rock types with unique characteristics may be associated with mineralization and used for targeting and prospectivity. Additionally, the broad consensus is that globally, metal and fluid sources in orogenic systems vary from metamorphic dehydration/devolatilization (Henley et al., 1976; Phillips, 1993; Goldfarb et al., 2005), magmatic-hydrothermal fluids (Burrows et al., 1986), fluids released directly from the mantle (Cameron, 1989, 1988; Colvine, 1989), and fluids expelled from external sources such as subducting oceanic plates (Pitcairn et al., 2014; Pitcairn et al., 2021). It is also possible that fertile orogenic belts have one or a combination of these sources contributing to deposit formation; however, direct linkages are difficult to unravel as orogenic belts are extensive (sometimes 100s of kilometers) and resolving these issues is out of scope for most exploration campaigns.

Mechanical and/or chemical traps for precipitating metals are often the most successful exploration criteria (Goldfarb et al., 2005; Dubé et al., 2020). Common mechanical traps such as antiform fold hinges (Cox et al., 1991; Goldfarb et al., 2005; Groves et al., 2018) and zones with strong rheological contrast (e.g., contacts between volcanic stratigraphy and syn-volcanic intrusions; Colvine, 1989; Witt, 1992; Robert, 2001) are interpreted as features that affect fluid flow rates, thus promoting

metal-sulfide precipitation. Chemical traps are typically lithologies with high Fe/(Fe+Mg) ratios (e.g., banded iron formations, iron-rich tholeiites, ferruginous shales, and some felsic igneous rock types; Cox et al., 1991; Goldfarb et al., 2011). The iron in such lithologies is interpreted to react with sulfur in fluids, causing desulfidation and precipitation of metals (Colvine, 1989; Goldfarb et al., 2005; Robert et al., 2005).

7.1.2. Geological setting

The Superior Province is the largest preserved Archean craton and consists of Eo- to Neoarchean granite-greenstone, gneissic, and metasedimentary domains that amalgamated as a series of east-trending belts between ~2720 and ~2660 Ma (Percival et al., 2012; Frieman et al., 2017). Mineral deposits are typically hosted in greenstone belts, with the largest and most abundant deposits in the Abitibi Subprovince. The regions near Timmins (southern Abitibi) and Dryden (western Wabigoon), Ontario (Figure 7-2), selected for this comparative study have similar geologic histories in terms of geodynamic setting, volcanism, deformation, and metamorphism (Beakhouse et al., 2011; Zammit, 2020; summary in Chapter 2) and benefit from similar quality/coverage of pre-competitive data. However, their contrasting reported endowment of base and precious metals suggest some underlying difference. Since all rock units in both areas are, with the exception of Proterozoic diabase dikes, metamorphosed to greenschist to amphibolite facies, the term 'meta' will not be used in lithologic descriptions.

Both study areas (geological maps in Montsion et al., 2021a/Chapter 3) are comprised of bi-modal volcanic assemblages that are unconformably overlain by two sedimentary successions. Disconformities at volcanic succession boundaries are locally marked by iron formation and cherts (Davis et al., 1989; Thurston and Chivers, 1990; Percival et al., 2004; Tomlinson et al., 2004; Davis et al., 2005; Thurston et al., 2008; Frieman et al., 2017). Late-volcanic basins dominated by marine turbiditic successions (Porcupine assemblage in Timmins and Thunder Lake group in Dryden) unconformably overlaying the volcanic stratigraphy. Coarse- to medium-grained syn-deformation molasse basins dominated by conglomerates and minor alkalic volcanic layers (the Timiskaming assemblage in Timmins and Stormy Basin group in Dryden) overlay older assemblages adjacent to trans-crustal deformation zones (Beakhouse et al. 1995; Dostal et al., 2004; Ayer et al., 2005; Bateman et al., 2005; Corcoran and Mueller, 2007; Bleeker, 2015; Montsion et al., 2021a/Chapter 3).

Syn-volcanic and syn-deformation batholiths, plutonic suites, dikes, and sills variably intrude volcanic and sedimentary assemblages, and are commonly deformed at their margin and proximal to shear zones. Similar to their volcanic counterparts, intrusions display ultramafic, gabbroic to granodioritic, and tonalitic to "potassic" (i.e., granite, monzonite, and syenite) compositions. Late intrusive S-type

felsic granitoids, spatially associated with pegmatite dikes post-dates the bulk of the deformation. Several swarms of diabase dike intruded all the above units during the Proterozoic (Blackburn et al., 1981; Ayer et al., 2002; Davis et al., 2005; Duguet and Beakhouse, 2012, Montsion et al., 2021a/Chapter 3).

Structural histories in the two map areas differ subtly (Kresz et al., 1982a, 1982b; Ayer et al., 2002, 2005; Bateman, 2008; Bleeker, 2015; Monecke et al., 2017; Montsion et al., 2021a/Chapter 3; Zammit 2020). Two exceptions to this are the surface expression of gneissic basement and development of a NE/SW trans-crustal deformation zones in the Dryden region; the latter being locally associated with gold occurrences and deposits.

Economically, the two greenstone belts differ greatly in abundance and grade of magmatic, volcanogenic, and orogenic mineral occurrences. The Timmins mining camp hosts several world-class deposits, including the Kidd-Creek Volcanogenic Massive Sulfide (VMS; 9.6 Mt Zn, 3.7 Mt Cu), Dome (27.1 Moz Au), Hollinger (20.5 Moz Au), McIntyre (10.8 Moz Au) and 10+ smaller orogenic deposits (>1 Moz Au), and a few magmatic deposits including Hart (1.9 Mt Ni and Cu), Langmuir (North, 1, and 2; 12 Mt), Alexo (0.6 Mt), and Redstone (0.5 Mt). Total endowment represents over 2700 t of gold, 10 Mt of zinc, 4 Mt of copper, and 60000 t of nickel (measured and indicated, Ontario Geological Survey, 2019). In contrast, Dryden hosts rare magmatic occurrences that are associated with cumulates within mafic intrusions north of the Atikwa-Lawrence batholith, several massive sulfide occurrences north of the Wabigoon deformation zone, and several moderately-sized orogenic Au deposits, prospects, and occurrences including Goliath (1.2 Moz Au), Goldlund (0.8 Moz Au), Gold Rock (0.15 Moz) and Van Horne (0.002 Moz, past production plus measured and indicated resources; Ontario Geological Survey, 2019).

7.1.3. Data source/availability

Data was compiled from pre-competitive datasets through the Ontario Geological Survey web portal (https://www.geologyontario.mndm.gov.on.ca/index.html) and new observations from four field seasons (2018-2021). Geological measurements and interpretations as well as re-processed geophysical data were released by Montsion et al., (2021a)/Chapter 3. Recompiled and new geochemical data will be released as a comprehensive database by Montsion et al. (submitted a)/Chapter 4. These data were overlayed, integrated, and interrogated to generate 34 new explanatory variable maps.

7.1.4. Training points

Five unique training sets were generated for each mineral system (magmatic, volcanogenic, orogenic) in each map area. Reported mineral deposits, prospects, and occurrences were used as

'targets' and given a training value of 1. Background points with no known association to mineralization were randomly generated in areas greater than two kilometers from target points and were required to be at least two kilometres away from each other. Background values were assigned a training value of 0 and their population size is equal to that of each system's occurrences/deposits.

During training, random forests generate a list that ranks input variables (i.e., maps/grids) using a value that represents the number of times a variable is responsible for a split (i.e., a decision within a tree) divided by the number of trees. This ranking value is called the Gini coefficient (Gini, 1912; Ceriani and Verme, 2012) and can be generated many times over for all variables in a given model as a sensitivity measure. During training, each model was validated 100 times for 5 different training sets in each of the three mineral systems. Mineralized targets are constant for each mineral system, independent from one another, and the location of barren background is randomly selected 5 times anywhere >2 km from mineralization. Gini coefficients were converted to percentage importance for interpretability

Insufficient training points for magmatic and targets in the western Wabigoon (see Table 7-2, geological map in Figure 3-2, and supplementary material Figure 7-4.2), and thus the Dryden area, limited the application of random forests to the southern Abitibi (i.e., Timmins area). In contrast, orogenic deposits and occurrences are abundant in both areas and thus supported random forests and a comparison of ranked feature lists.

Table 7-2. Count and brief summary of the occurrences/deposits reported for each mineral system in the Timmins and Dryden areas.

System	Timmins		Dryden	
	Count	Description	Count	Description
Magmatic	81	Mainly clustered around Alexo/Dundonald deposits north of the Pipestone fault and Hart/Langmuir mines along the periphery of the Shaw Dome	24	Small clusters around ultramafic to mafic intrusions/sills
Volcanogenic	97	Mainly clustered around 1) the Kidd Creek mine, 2) around the large TTG intrusion west of the Mattagami Fault, and 3) the Cross Lake discovery in the upper Tisdale Assemblage	13	Very rare and often modified by orogenic overprint
Orogenic	295	Mainly located adjacent to the Porcupine Destor deformation zone and its splays	212	Mainly located in hinges of regional-scale folds

7.2. Feature engineering input layers

In subsequent sections, a brief summary of each proposed feature engineering technique and its link to conceptual models are presented. Detailed methodology for feature engineering techniques is beyond the scope of this comparative study and are instead presented in supplementary material 7-

1 to 7-5. The applicability of individual methods to other geological regions or mineral systems are dependent on the presence and quality of primary datasets as well as geological relevance to the scientific objective. A correlation matrix (supplementary material 7-6) was used to confirm that feature maps do not duplicate information and bias the training phase of random forests. Broadly, the low correlation values in the correlation matrix indicate sufficient independence between feature maps and appropriateness for inclusion during training.

Explanatory variables were generated as grids clipped to map extents indicated in Figure 7-2 using a cell size of 250 m and NAD 83 UTM 17N (Timmins) and 15N (Dryden) projection. An exception to this cell size is made for the petrophysical grids, which use a 50 m cell size to increase resolution on thin rock units.

7.2.1.1. Assemblage maps

Assemblage maps are well established in terms of relevance to geological studies and methodology. Certain stratigraphic levels or positions in many mineral systems are associated with known mineralization, particularly in volcanogenic systems (e.g., McNicoll et al., 2014). Additionally, chemical and physical characteristics of assemblages evolve as new melt sources/compositions are exhausted/tapped. It is therefore relevant to include an assemblage map as an explanatory variable to act as a comprehensive proxy for geological settings and magmatic evolution, factors that are challenging to represent quantitatively but are significant in concept models. Some subsequent feature layers are partially constrained by the assemblage map leading to some redundancy during feature engineering. However, a correlation matrix of feature layers (supplementary material Table 7-6.1) indicates sufficient independence to support training a machine learning model.

Figure 7-3. (next page) A) A bedrock geological map with interpretations and compiled mineral deposit locations in the Timmins area (from Montsion et al., 2021a/Chapter 3) provides context for examples of new explanatory variables (maps) used to represent components in mineral systems including B) empirical Bayesian Kriging interpolation of estimated gain/loss factor of SiO2 normalized to mafic compatibility, C) Minimum Curvature interpolation of mafic compatibility (Mg/Zr) masked for each lithology type and assemblage, D) Rheological contrast resulting from the summed variance of brittle (cohesion, uniaxial compressive strength, friction coefficient) and ductile (bulk, shear, compression moduli) property maps, as well as non-Euclidean distance from E) Clastic sedimentary map units and F) Trans-crustal deformation zones calculated from the sum of weighted nodes and edges in a triangulated mesh using graph analysis.
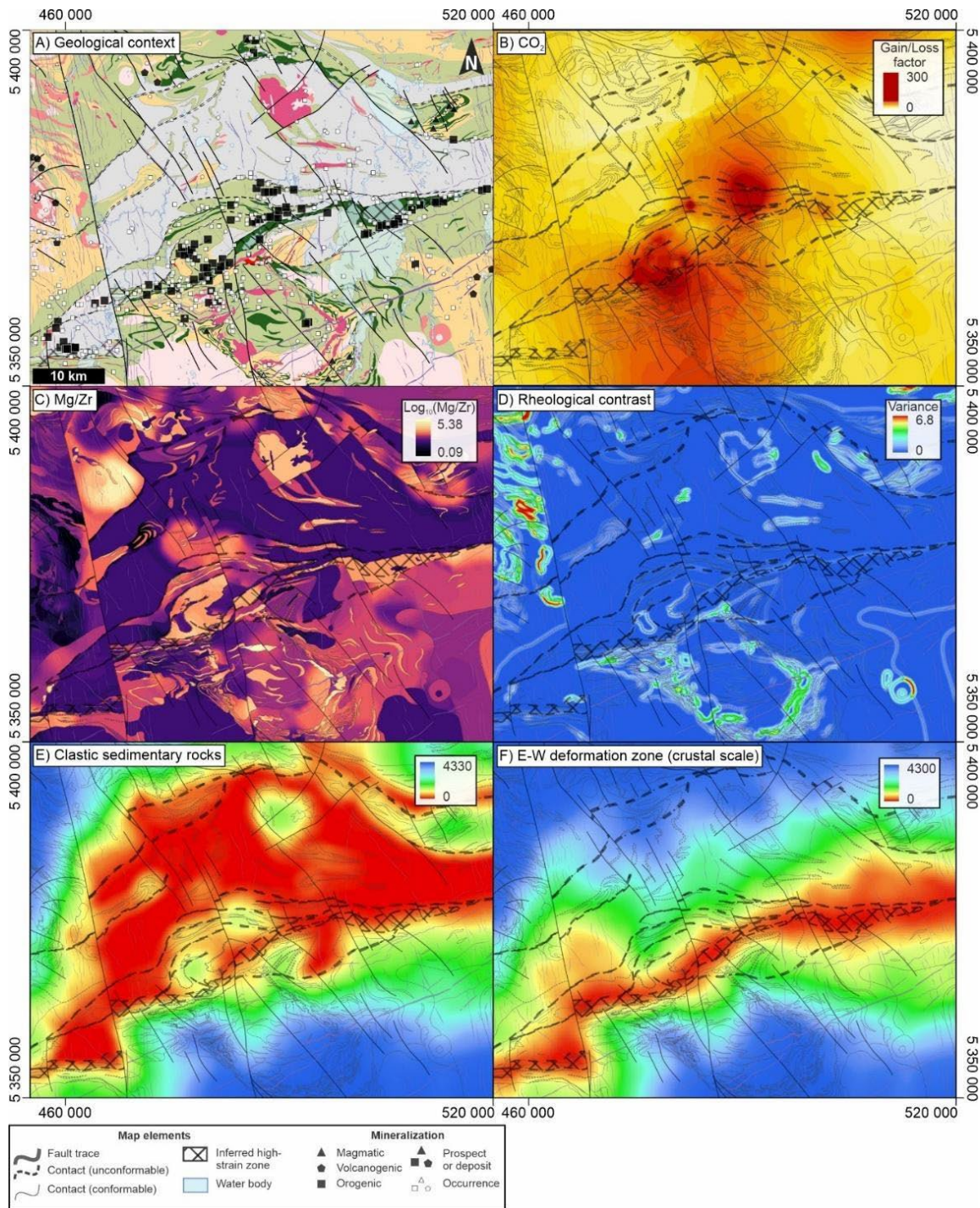
Assemblage maps were interpreted using isotopic ages, field cross-cutting relationships, and rock types. Since assemblage age ranges can overlap or have gaps, the assemblage explanatory variable was treated as 'categorical'. This differs from all subsequent layers that involve quantifiable continuous values.

Limitations in this feature map arise from 1) using map units/structures that contain subjective interpretation biases as well as 2) generalization from using sparse geochronological samples and 3) discretizing eruption/deposition ages into map unit polygons.

7.2.1.2. Interpolated mobile element anomalies

7.2.1.2.1. Description of method

Fluid chemistry and mobility (i.e., proxies for alteration) are a critical component in mineral exploration as they carry metals to deposition sites. Mass balance calculations are common practise proxies to track fluid path, chemistry, and volume (Gresens, 1967; MacLean, 1990; MacLean and Barrett, 1993). In essence, mass balance quantifies the difference between the concentration of a mobile element (or element group) in a sample (a single point on the map) compared to 'least altered' sample(s) (Poldervaart, 1953; Bogolepov, 1963; Appleyard, 1980, 1990). In practice, this can be time consuming, necessitates density data, and requires lithologies to be interrogated independently. It also assumes there are enough non-altered/minimally altered samples to be statistically representative. The point-based nature of these data and the previously described requirements make inclusion of mass balance feature maps difficult (e.g., Hood et al., 2019; Perrouty et al., 2019).

The transport path of metalliferous fluids can be tracked by mapping alteration resulting from fluid-wall rock interaction. Herein we present a 'pseudo-mass balance' strategy for capturing gain/loss of mobile elements. In this strategy, a suite of less altered samples is identified (<2% LOI), their primary composition (ultramafic to felsic) was characterized geochemically according to a lithology normalization term, and average mobile element concentrations were identified by fitting a linear function. This step normalizes mobile elements to a presumed un-altered concentration typical for a given lithology and mitigates reported issues (Mathieu, 2018) when applying popular alteration indices to compositionally variable datasets. The concentrations of mobile elements representing carbonatization (CO_2), silicification (SiO_2), albitization (Na_2O) and/or sericitization (K_2O) are then quantitively compared with the less-mobile average.

Since least altered samples from all lithologies were not available for this regional scale study, an extrapolation from a linear function that approximates the average of less altered samples was used. All samples with < 2.0 LOI % were considered 'less altered' and a mass loss/gain factor was determined by taking the difference between mobile element concentrations and approximated linear functions (supplementary material Figure 7-5.1). These functions were constrained by trends in immobile (igneous: $\log(Mg/Zr)$; sedimentary: A_2O_3 wt%) vs mobile (CO_2, SiO_2, K_2O, Na_2O) cross-plots with less altered samples. A mass gain/loss factor = 1 implies no change, 0 implies 100% loss, and 2 implies 100% gain in the mobile element.

In addition to LOI, Empirical Bayesian Kriging (EBK; Krivoruchko, 2012) was used to interpolate gain/loss factors on samples within the map area and a 15 km buffer zone to generate a spatially

continuous representation with minimal edge effects. EBK is an evolution of kriging and uses hundreds to thousands of variograms (herein n = 1000) to select the optimal appropriate range sill parameters. The spatial and statistical distribution of available samples varied and necessitated independent tuning of hyperparameters (reported in supplementary material Figure 7-3. And 7-5.4) to select the result with the fewest artifacts.

7.2.1.2.2.   Limitations

Limitations for this feature engineering method lay within three groups 1) common sample bias, 2) normalization and calculation of mobile element gain/loss and 3) generating spatial representations.

Frequently, rocks sampled in the field are preferentially biased to a given study's focus. Regional studies often sample bulk rock geochemistry and target material as far away from veins as is reasonable. Conversely, economic studies may exclusively sample veins. In either case, the resulting data will be biased towards the absence/presence of fluid chemistry from veins.

Normalization and calculation limitations include: 1) Incomplete or biased 'least altered' material not fully capturing background rocks. Since many mineral systems are spatially distributed, finding samples that are wholly unaffected is challenging. Additionally, the 'least altered' material selection assumes metamorphism to be isochemical. 2) Functions use to approximate trends in least altered samples and predict 'unaltered' mobile element concentrations probably oversimplify true trends. Segmented linear functions are used to capture variability in point clouds; however, functions with higher degrees of freedom may be better suited. Additionally, a density-based envelope around data clouds rather than function may be more reasonable when representing possible least altered values; 3) Lithological normalization terms are a necessary oversimplification and the comprehensibility provided by multi-variate normalization would be more appropriate. However, implementation of higher dimensional control on lithology normalization is challenging due to limited diversity in elemental analyses and interpretation complexity (Montsion et al., submitted a/Chapter 4 and Montsion et al., submitted b/Chapter 5). 4) The crystallization/depositional processes and thus geochemical differences between igneous and sedimentary rocks necessitate different lithological normalization terms. As a consequence, analytical and natural geological biases heterogeneously affect results. A focussed assessment with a significantly larger database would be required to address this.

Limitations with generating maps of gain/loss of mobile elements are the same as any other method where spatially continuous grids are derived from discrete points. The spatial distribution of samples is sub-optimal as accessibility to rocks, collection/analytical budgets, and natural variability of rock properties obscure meaningful information.

7.2.1.3. Structural complexity

Structural complexity is defined here by increased variance of planar feature orientations (e.g., bedding) in response to deformation. It can be used to map fluid pathways such as high strain zones where shortening and/or shearing rotates existing structural grain into subparallel orientations (linear low anomalies) and structural fluid traps such as fold hinges (high anomalies; Montsion et al., 2021b/Chapter 6) where bedding is non-parallel. This data-driven method provides a less subjective alternative to popular distance to/density of fault traces and fold axes (e.g., fault traces, fold axes, favorable lithological units; Knox-Robinson and Wyborn, 1997; Carranza and Hale, 2001; Bierlein et al., 2006; Kreuzer et al., 2015; Zhang and Zhou, 2015; Campos et al., 2017; Ford, 2019).

Structural complexity (i.e., axial variance, Montsion et al., 2021b/Chapter 6) is calculated in neighbourhoods of segmented auto-detected aeromagnetic lineaments and interpolated across the map area. Aeromagnetic lineaments are interpreted to reflect variations in the concentration of magnetic mineral in differing stratigraphic units; however, any dataset that captures linear stratigraphic characteristics would also be applicable (e.g., gravity, radiometric surveys, elevation models). Limitations with this feature engineering method arise from 1) variable signal-to-noise ratio for magnetic stratigraphy; 2) interference from poly-phase intrusive suites (which can be muted); 3) and availability of geophysical data.

7.2.1.4. Non-Euclidean distance maps

In many prospectivity studies, distance/ proximity maps are used to capture spatial relationships between features (e.g., fluid pathways and mineralization) and overcome a spatial blindness inherent in many machine learning algorithms (e.g., linear regression, random forests). However, such maps represent the linear (Euclidean) distance in today's map configuration from one location to another. They do not account for the spatial and topological relationships that existed when the mineral system was active or before deformation. As an example, the distance between two points on a fold using today's Euclidean distance is not equal to the distance between points before folding. It is, therefore, important to acknowledge, if not overcome, the effect of using Euclidean distance maps as explanatory variables in geological applications of machine learning.

In an exploration context, proximity to a wide variety of features can be favorable. Proximity to fluid pathways (e.g., deformation zones, faults, more permeable rocks), structural (e.g., folds, faults, step-over zones), chemical (e.g., Fe-rich lithologies causing desulfidation) or physical traps (e.g., impermeable, or rheologically competent lithologies), heat or fluid sources (e.g., plutons), sulfur sources (e.g., banded iron formation, shales) are some of the many exploration criteria for mineralized targets.

Herein, we apply network graph analysis (e.g., Curriero, 2006; Lu et al., 2011; Rizzo & de Barrow, 2017; Tang et al., 2021) to calculate a more geologically representative pre- to syn- deformation distance, termed 'non-Euclidean'. In essence, the graph analysis takes the sum of weighted nodes in a triangulated mesh of map, thus respecting spatial and topological relationships and can be viewed as a pseudo-restoration when traditional methods are not viable. This technique can also be used to investigate the spatial/temporal impacts of differing topological models. Detailed descriptions of this methodology, limitations, and its application to the Timmins and Dryden map areas are provided in supplementary material Figure 7-1.1.

Broadly, limitations or bias related to this feature engineering technique arise from 1) using interpreted map units/structures as the starting point for distance calculations; 2) using a single interpreted topological model that dictates connectivity; 3) ambiguity or generalization from using idealised weighting on map units and structures; 4) ability of triangulated meshing technique to capture irregular map unit polygons.

7.2.1.5. Discrete assignment

Petrophysical properties (e.g., magnetic susceptibility, total porosity, grain density), physical properties (e.g., p- and s-wave velocities, shear modulus, bulk modulus, compression modulus, frictional coefficient, cohesion, uniaxial compressive strength), and electrical properties (e.g., rock resistivity, conductivity) are often underutilized but valuable when understanding physical and chemical processes involved in mineral deposit formation. For example, the porosity of a lithology may be used to identify permeable zones (i.e., sinks) or impermeable caps (i.e., traps) for fluids. Also, the ductile strength captured by seismic velocities and derived moduli indicate stress focal points and perhaps likely structural fluid traps. In contrast with the petroleum industry, mineral exploration has neglected systematic standardized methods recording rock properties, thus limiting opportunities for methods development that would enhance interpretability/reconstruction of ore systems. Since a centralized rock property database is absent, our application used available data from several sources with varying degrees of quality and completeness (see supplementary material Table 7-2.1 for values and Table 7-2.2 for references).

Herein, average values from local petrophysical databases and available published geotechnical reports are assigned uniformly to all polygons (i.e., map units) of the same lithology. Values used for this assignment are used to represent available data for fresh, unaltered rocks to approximate values for these rocks during magmatism/volcanism. This form of assignment is described as 'discrete' because of its uniform nature and sharp constraint to map unit boundaries.

Limitations and bias related to this feature engineering technique arise from 1) using interpreted map units to constrain property distribution; 2) generalization from uniformly assigning rock properties based on lithology and 3) limited availability of rock property values.

7.2.1.6. Semi-discrete assignment

The chemical composition of a given lithology evolves throughout depositional/eruption cycles as new sources are exhausted, tapped, and/or fractionated (e.g., Sproule et al., 2002). In an exploration setting, compositional changes are important to document as mineralization may be favoured in specific strategic levels with optimal chemical/mineralogical/physical characteristics.

Logarithmic (log10) grids of Mg/Zr, Ce/Y, Th/Nb, Eu/Eu* and Al/Ti element ratios were generated using masked minimum curvature interpolations for each lithology type in each assemblage. These ratios are used to characterise lithological composition, nature of melt/magma sources, crustal interaction/fractional crystallization, and plagioclase/titanite-related igneous processes (after Montsion et al., submitted b/Chapter 5). Map unit polygons were used to constrain the interpolation as masks and thus, this technique is considered 'semi-discrete'.

Limitations and bias related to this feature engineering method arise from 1) using interpreted map units, structures, and assemblages to constrain interpolation; 2) irregularly distributed samples analysed for relevant elements; 3) oversimplification in interpolation with minimum curvature, and 4) interpolating using Euclidean space with the contemporary configuration of geological features

7.2.1.7. Rheological/Chemical contrast mapping

Areas of rheological contrasts are common traps for orogenic Au (Goldfarb et al., 2001; Joly et al., 2012) as focused strain/stress may create zones of weakness and voids for fluid to precipitate metals. Chemical contrast is also a valuable exploration vector as strong contrast between one rock unit and its neighbor may cause precipitation of metal from hydrothermal fluids. As with many other property-related explanatory variables, these layers are rarely included as targeting criteria and data to support feature engineering is limited. However, the system components they represent in the conceptual model are considered important and warrant attention.

Rheological contrast represents the sum of variance grids for ductile (i.e., bulk, shear, compression moduli) and brittle (i.e., cohesion, frictional coefficient, uniaxial compressive strength) rock properties. Similarly, geochemical contrast represents the sum of variance for each of the five semi-discrete element ratio grids describe above. Variance for individual property maps was calculated as the square of the standard deviation of the property within a 1 km diameter circular moving window.

In addition to limitations and biases related to the discrete assignment of their parent features, it should be noted that contrast grids are affected by the size of the moving window from which statistics are calculated.

7.3. Random forests

7.3.1. Random Forest methodology

Forest-based classification, here random forests, uses a collection of explanatory variables (layers/grids) to classify pre-defined training points based on similarities. Random forests apply an ensemble approach where a pre-defined number of trees are generated and each randomly samples the collection of explanatory variables to classify a random subset of training points. As a consequence of its versatile and interpretable nature, random forests have been applied to a wide range of geological questions including lithologic classification (Kuhn et al., 2020), prospectivity (Ford, 2020), classifying deposit types (Zhan et al., 2023), and capturing magmatic intensity in a porphyry system (Zhou et al., 2022). Cracknell and Reading (2014) compared several machine learning methodologies (neural networks, support vector machines, naïve Bayes, k-nearest neighbours) and determined that random forests was best suited for geoscientific supervised classification as it is intellectually accessible and generally stable given a range of hyperparameters (detailed in the next paragraph). Here, the 'Forest-based classification and Regression' tool available in ArcGIS Pro v. 2.9 (documentation: https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-statistics/forestbasedclassificationregression.htm), which is an adaptation of Breiman's random forest algorithm (Breiman, 1996, 2001), is used.

7.3.1.1. Hyperparameters

All random forest models were trained using five thousand trees (i.e., for each of the 5 training sets characterizing each of the three mineral systems). Hyperparameters such as percentage of training points used in each tree (70%), percentage of the subset training points kept for validation (20%), and number of variable grids (1/3 of all grids) were intended to diminish training set biases. Model fit was controlled by the leaf size and maximum tree depth; however, settings vary and are tuned to balance model fit and performance (distribution of explanatory variable values; Reported in supplementary material Table 7-7.1). Hyper-parameter selections were qualitatively made during extensive sensitivity testing to balance fit using common performance metrics. Validation of training results and feature ranking list was run 100 times and provides a large dataset (total 500 for each mineral system) to statistically examine consistency in feature importance.

Performance metrics compare model success and assist in 'exploitability' (i.e., understanding success/failure and influence of parameters on results). Four performance metrics are relevant when assessing and comparing random forest models: 1) Out of Bag error (OOB) indicates model accuracy by measuring how well each tree predicts training points that were set aside for validation. These points played no part in generating decision trees and is reported as a percentage of each category being misclassified, thus high values are undesirable. All models used 20% of training data for validation (a.k.a. five-fold cross-validation). 2) Mean-Squared Error (MSE) indicates model accuracy by indicating the OOB over all categories. 3) F1-Score and 4) Matthews Correlation Coefficient (MCC) are metrics that summarize the confusion matrix in terms of precision, recall, and specificity. An F1-Score of 0 is undesirable as precision or recall would be zero. An F1-Score of 1 indicates perfect precision and recall, but likely reflects an overfit model. In practise, a high F1-Score (close to but not actually 1) is desired. In cases where both training classes are equally important (i.e., in mineral exploration targeting), MCC is considered to be more robust (Chicco et al., 2021). MCC scores range from -1 (very poor classification) to +1 (perfect classification).

Broadly, all models performed within reasonable ranges (see ranges in Table 7-3 as well as detailed report in supplementary material Table 7-8.1 and Table 7-8.2) and display a tolerably balanced fit. In the case of magmatic and volcanogenic models, results and performance metrics may affected by limited number of training points and err on the side of overfitting.

		Timmins			Dryden
		Magmatic	**Volcanogenic**	**Orogenic**	
Out of bag errors	MSE	9.6-14.9	22.4-32.8	12.1-17.8	15.7-20.0
	0	6.2-11.5	22.8-38.0	8.7-25.5	17.8-23.7
	1	11.9-19.9	22.6-34.1	10.4-15.4	12.8-18.4
Training	FI - 0	0.96-0.99	0.98-1.00	0.90-0.98	0.94-0.98
	MCC - 0	0.92-0.98	0.96-1.00	0.82-0.95	0.89-0.96
	Sensitivity	0.98-1.00	0.97-1.00	0.84-1.00	0.91-0.98
	Accuracy	0.96-0.99	0.98-1.00	0.91-0.98	0.95-0.98
	FI - 1	0.96-0.99	0.98-1.00	0.92-0.98	0.95-0.98
	MCC - 1	0.92-0.98	0.96-1.00	0.82-0.95	0.89-0.96
	Sensitivity	0.92-0.98	0.99-1.00	0.95-0.98	0.97-0.99
	Accuracy	0.96-0.99	0.98-1.00	0.91-0.98	0.95-0.98
Validation	FI - 0	0.88-0.9	0.71-0.80	0.81-0.88	0.79-0.85
	MCC - 0	0.75-0.83	0.44-0.53	0.70-0.76	0.62-0.66
	Sensitivity	0.82-1.00	0.65-0.82	0.85-0.93	0.76-0.87
	Accuracy	0.87-0.91	0.71-0.76	0.85-0.88	0.81-0.83
	FI - 1	0.87-0.91	0.69-0.76	0.84-0.88	0.80-0.84
	MCC - 1	0.75-0.83	0.44-0.53	0.70-0.76	0.62-0.66
	Sensitivity	0.76-0.93	0.69-0.8	0.83-0.89	0.78-0.87
	Accuracy	0.87-0.91	0.71-0.76	0.85-0.88	0.81-0.83

Table 7-3. Summery of performance metrics reported by random forests for classification of magmatic, volcanogenic, and orogenic targets using five unique training sets. Metrics are reported as minimum and maximum values. '0' denotes results for classifying background locations and '1' denotes mineralized targets. MSE: Mean Squared Error, F1: F1-score, MCC: Matthews Correlation Coefficient

7.3.2. Ranked system components

Feature importance is a common performance metric used to assess the sensibility of trained models. Importance (i.e., Gini index) is the sum of all Gini impurity scores associated with each split a feature map is responsible for, over the number of all splits in the tree (Gini, 1912; Ceriani and Verme, 2012). Gini impurity governs which map is selected to define the next split and quantifies the likelihood of a training point being misclassified if given a random label (i.e., target or background. Essentially, this metric quantifies how often and how reliable a feature map splits training points according to their labeled classes, making it an ideal metric for assessing model stability.

In this application of random forests, relative ranking and interpreted importance in conceptual mineral system models are compared to validate the geological representativeness of feature maps and thus appropriateness of proposed feature engineering methodology. Agreement between feature ranking and conceptual models indicates that mineral system components were adequately captured and feature engineering methods are appropriate. Deviations from expected rankings may indicate areas for methodological improvement or knowledge gaps in conceptual models.

7.3.2.1. Confidence of feature importance

To assess confidence in Gini indices, values were calculated for 100 validation runs using each of the 5 training sets for each mineral system, meaning 500 Gini indices were returned and converted to percentages. The distribution of the 500 percent importance values was used to assess confidence (i.e., consistency). Feature maps with variable importance are represented as broad box and whiskers in Figure 7-4 (i.e., lower confidence). Those that were consistently useful/not useful have narrow box and whiskers (i.e., higher confidence). Stability of entire models and representatives of training points can also be assessed from trends based on all feature importance distributions. For example, models trained using limited or biased training sets generally have consistently broader (lower confidence) distributions compared to those trained using larger point populations or contain less bias. In general, high-ranking features have broad distributions compared to those at the bottom (Figure 7-4).

7.3.2.2. Model stability and training points

When assessing model stability for each mineral system, the association between training population sizes and width of box and whiskers (Figure 7-4) is obvious. Models trained using magmatic (n = 162) have the highest variability, volcanogenic (n = 194) display intermediate distribution width, and orogenic (n = 590 for Timmins and n = 424 for Dryden) training points are most consistent.

However, distributions may also be affected by poor representation of mineralization. Magmatic targets in the Timmins area are clustered around the Mc Watters/Langmuir deposits on the southern flank of the Shaw Dome and Alexo/Dundonald deposits north of the Pipestone fault (see map in Figure 3-3). The observed clustering likely biases training to classify points based on local geological criteria in addition to the underlying system controls responsible for mineralization. For that reason, Timmins is not an ideal natural laboratory for assessing fertile magmatic system criteria. Instead, we should perhaps look to other settings that host major deposits and have stronger signal-to-noise geochemical characterizations (e.g., Kambalda in the western Yilgarn Craton, Barnes et al., 2007; Barnes and Fiorentini, 2012; Le Vaillant et al., 2016). Nevertheless, performance metrics (supplementary material 7-8) and the emergence of trends in importance ranking distributions indicate that the training model is sufficiently stable. Magmatic occurrences in Dryden are rare and only found adjacent to mafic intrusions at the margins of the Atikwa-Lawrence batholith and were, therefore, not used for training.

Volcanogenic training points are similarly affected by spatial clustering. In the Timmins area, target points are associated with the Kidd Creek deposit north of the Porcupine Basin, Kam Kotia/Jamieson deposits adjacent to TTG plutons west of the Mattagami River fault, or a poorly constrained Tisdale felsic volcanic unit in the southeast corner of the map area (see map in Figure 3-3). The narrower distribution, however, suggests that there is better representation of mineralization controls rather than local geological characteristics and the model is stable. Near Dryden, massive sulfide occurrences have been reported hosted by volcanic and sedimentary rocks across the Thunder Lake group north of the Wabigoon deformation zone; however, that population is too small to reliably train a random forest model.

Orogenic targets are the most abundant and widely distributed across both map areas. Clusters around major deposits are balanced by many smaller occurrences outside of orebody footprints. The Timmins mining camp is globally significant for orogenic Au and exploration has been extensive, which is reflected in nearly complete coverage and excellent representativeness of mineralization in training points, leading to a stable and, perhaps, meaningful classification model. Dryden has garnered less exploration interest and thus the representativeness of training points is diminished. However, occurrences are well distributed across the map area and support the training of a stable model.

7.3.2.3. Ranking position and application

Validation of proposed feature engineering methods was performed by comparing feature ranking to conceptual models (Table 7-1) to box and whisker plots ordered by mean importance for each of

the random forest models (Figure 7-4) as well as a ternary diagram comparing proportional importance (Figure 7-5). The scatter plots are used as a visual comparison to determine shared and unique controls on magmatic and volcanogenic system prospectivity. Additionally, the scatter plot comparison of Timmins and Dryden orogenic rankings facilitates discussion of controls governing contrasting Au mineralization. Ranked features are qualitatively divided into 3 importance categories ('High,' 'Medium,', 'Low' in Figure 7-5).

Generally, ranked feature lists support conceptual mineral system models and confirm that the applied feature engineering methods adequately represent relevant geology. However, there were a few exceptions that provided insight about methodological limitations or knowledge gaps. These exceptions will be discussed in subsequent sections.
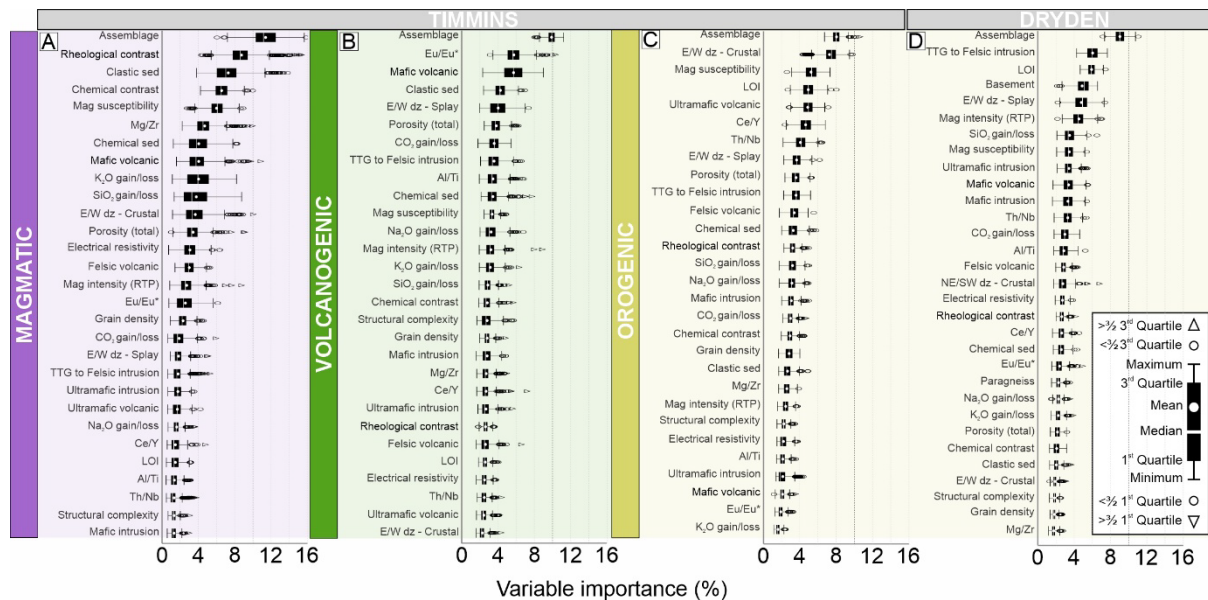
Figure 7-4. Box and whisker plots of ranked explanatory variables capture the statistical spread of 100 validation runs for each of the five training datasets (i.e., 500 validation results) used to classify A) magmatic, B) volcanogenic, and C/D) orogenic mineralization near Timmins (A-C) and Dryden (D).

7.3.2.3.1.   Magmatic

7.3.2.3.1.1.   Confirmation of conceptual model

In the conceptual model for magmatic systems the parameters assumed to govern system fertility are 1) metal-enriched magma source(s), 2) sufficiently strong flux to drive magma ascent, 3) connected and extensive plumbing systems (i.e., crustal-scale deformation zones), and 4) sulfur sources along magma pathways. With the exception of absent proxies for metal enriched sources, the ranked feature list of variables more or less supports the conceptual model and thus confirms that new feature maps reasonably capture key components in the system. A detailed description of

high and low ranked features that were both expected and unanticipated given the conceptual mineral system model is provided in Table 7-4.

Proxies for high temperature (magnesian) ultramafic host rocks (Mg/Zr), proximity to trans-crustal magma conduits (deformation zones), and proximity to S-rich lithologies (i.e., clastic and chemical sedimentary rocks) comprise those in the high to medium importance categories (Figures 7-4A, 7-5A).

As a confirmation of the conceptual model, variables such as 1) hydrothermal alteration proxies (e.g., LOI, CO_2, Na_2O), structural trap indicators (e.g., high structural complexity), 2) spatial relationships with syn-volcanic felsic intrusions as well as syn-tectonic ultramafic to felsic intrusions, and 3) non-discriminatory rock properties (i.e., total porosity, resistivity) were expected to rank low (i.e., unimportant).

7.3.2.3.1.2.   Deviation from conceptual model

Variables that ranked low, despite their hypothesized importance in fertile magmatic systems, include 1) chemical characteristics of favorable host rocks, 2) proximity to ultramafic rocks (intrusive and volcanic), and 3) characteristic rock properties of ultramafic rocks (e.g., grain density, resistivity). Specific explanations are presented in Table 7-4
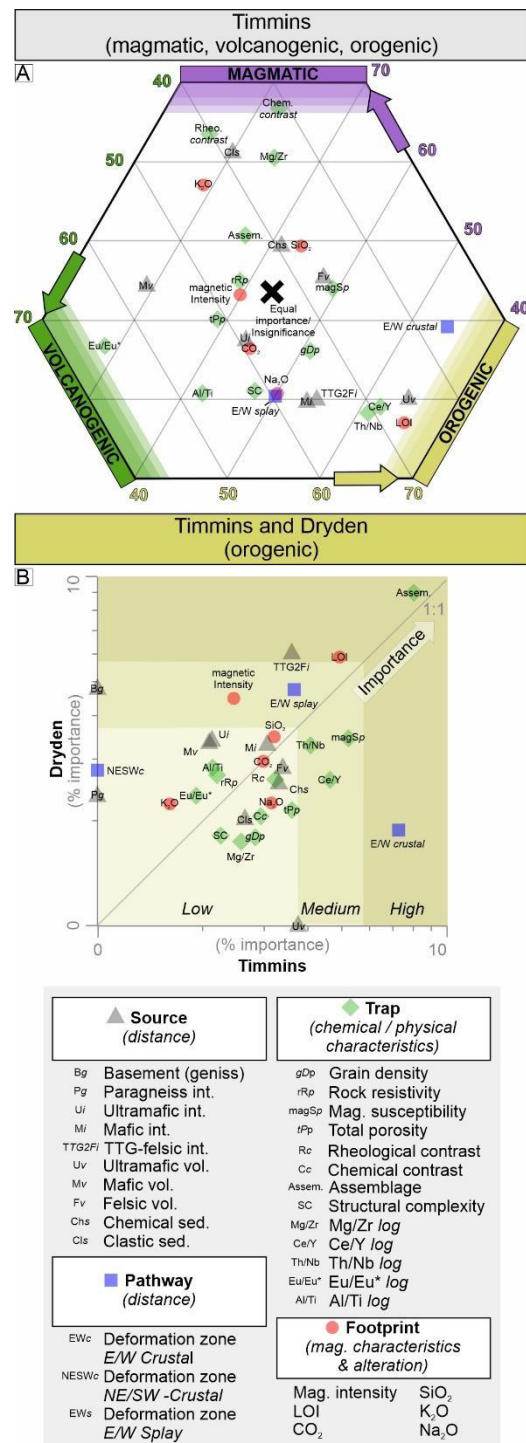
Figure 7-5 (above right). Comparison of explanatory variable (i.e., map) importance rankings trained to classify known mineralized targets from randomized background locations. (A) Relative variable rankings for magmatic, volcanogenic, and orogenic mineralization in Timmins (are compared using a ternary diagram to indicate shared and differing importance of exploration criteria. (B) Variable rankings for orogenic mineralization near Timmins and Dryden (are compared to investigate contrasting Au endowment between two similar geological terranes. Variables are symbolized based on dominant role in conceptual mineral systems models and qualitatively subdivided into groups of low, medium, and high importance to facilitate discussion.

Table 7-4. Detailed description of feature ranking and favorable signatures (within 1 km of known mineralization) that both support and were unanticipated from expected results given the conceptual magmatic mineral system model. Explanations of ranking results are speculative and require further investigation

	Feature	Geological rationale	Favorable signature	Feature	Explanation
	High rank			**Low rank**	
Predicted by conceptual model	Assemblage	Unique combination of melt source(s), ascent and eruption conditions, fluid-rock interactions	Kidd-Munroe and Tisdale	CO_2 (gain/loss)	Proxies for hydrothermal alteration that post-date volcanism/magmatism and mineralization
	Magnetic susceptibility	Proxy for higher temp./dense magnesian melt	High (0.016 SI)	LOI	
	Mafic compatibility (Mg/Zr)		High (~1000)	Na_2O (gain/loss)	
	Clastic and chemical sed (distance)	Possible S-source	Distal (~320)	Structural complexity	Syn to post deformation structural trap unrelated to mineralization
	Mafic volcanic (distance)		Proximal (~45)		
	E/W trans-crustal deformation zone (distance)	Path for melt ascent	Proximal (~500)	TTG to felsic intrusion (distance)	
	Rheological contrast	Indication of favorable conditions for magma emplacement (volcanic flow/sedimentary stratigraphy in contact with poorly consolidated volcaniclastic/sedimentary material). Generally marking contact between ultramafic rocks and felsic volcanic stratigraphy	High (~1.45)	Plag./Titanite fractionation (Eu/Eu*)	Not characteristics of common host rocks
	Chemical contrast		High (~0.24)	Melt partitioning (Al/Ti)	
	Felsic volcanic (distance)	Control height and volume of magma ascent	Proximal (~40)	Porosity (total)	
	K_2O (gain/loss)	Inverse relationship with MgO	High/Gain (~1.5)	Resistivity	
Unanticipated with conceptual model	SiO_2 (gain/loss)	**Methodological:** Usually associated with hydrothermal alteration but characteristic low concentrations in ultramafic igneous rocks	Less loss than other areas (~0.97)	E/W deformation zone splay (distance)	**Methodological:** Classifying the basal Pipestone fault (associated with a large cluster of targets) as 'splay' for non-Euclidean distance
				Melt evolution (Th/Nb)	**Methodological:** Favorable signatures within compositional group obscured by trends across full range
				Magnetic intensity	**Geological:** Not all ultramafic/mafic rocks are favorable, only those erupted during optimal conditions are associated with mineralization
				Density (grain)	
				Ultramafic volcanic and intrusive rocks (distance)	
				Mafic intrusion (distance)	

7.3.2.3.1.3. Limitations

Components not captured during feature engineering are 1) mechanisms for ascent (excluding high temperature) and 2) ascent vs sulfur assimilation rates. Tools that capture presence and interaction with volatile sources to counteract settling of dense metal-rich sulfur compounds in melts may greatly benefit system characterization. Unfortunately, this area of magmatic research is ongoing

and more detailed case studies are needed. Ascent/Assimilation rates are also challenging components to capture as favorability is variable based on melt and country rock chemistry/conditions (pressure/temperature). Frequently, sulfur isotopes and enrichment in incompatible elements (i.e., Zr, Th, LREEs) are used as proxies for crustal contamination (Lesher et al., 2001; Le Vaillant et al., 2016) and may be the best tools for characterizing favorable assimilation.

7.3.2.3.2.   Volcanogenic

7.3.2.3.2.1.   Confirmation of conceptual model

In the volcanogenic conceptual model, mineralization is focused in thin, juvenile bi-modal (mafic-felsic) volcanic crust and depend on the 1) geodynamic setting, 2) long-lived and shallow heat source(s), and 3) spatially extensive and connected fluid pathways. Generally, the ranked feature list (Figures 7-4B, 7-5A) supports the conceptual model and confirms that key components are adequately captured. A detailed description of high and low ranked features that were both expected and unanticipated given the conceptual mineral system model is provided in Table 7-5.

Proxies for a favorable geodynamic setting rank highly and include 1) assemblages deposited during later-stage volcanogenic and are dominated by bi-modal mafic to fractionated felsic melts (i.e., Tisdale and Blake River, supplementary material Figure 7-9.1 and 7-9.2), 2) non-Euclidean proximity to mafic volcanic rocks, and 3) low Eu/Eu* (, supplementary material Figure 7-9.1 and 7-9.2).

A medium ranking for non-Euclidean proximity to TTG to felsic intrusions indicates a moderate linkage between mineralization and a favorable heat source and supports the observation that gabbro-diorite-tonalite-trondhjemite intrusion complexes, such as those found in the Godfrey Township west of the Mattagami River Fault, are commonly associated with clusters of Archean volcanogenic deposits (Franklin et al., 2005; Galley, 2003; Galley et al., 2007).

Additionally, representations/indicators for fluid pathways are ranked with medium importance. Non-Euclidean proximity E-W deformation zone splays to distribute fluids from the main trans-crustal conduit (supplementary material Figure 7-9.2) and indicators of fluid interaction such as low magnetic intensity, from the destruction of magnetite, as well as low CO_2 and high Na_2O, K_2O, and SiO_2 support commonly applied exploration criteria (reviewed in Piercey, 2010).

As a negative confirmation of the conceptual model, layers that represent spatial relationships and proxies for non-prospective host rocks rank low in the feature importance list. Specifically, non-Euclidean distance to ultramafic rocks and mafic intrusions. Characteristic rock properties (e.g., resistivity, grain density, magnetic susceptibility) of host rock were also anticipated to rank low since mineralization is hosted in bi-modal volcanic stratigraphy. Additionally, trap-related features that

were not expected to be related to mineralization that forms from accumulated precipitates on the sea floor rank low. These include structural complexity and areas of chemical/rheological contrast.

7.3.2.3.2.2. Deviation from conceptual model

Deviations from the volcanogenic conceptual model are mainly low-ranking features thought to be significant in the conceptual model and include 1) LOI, non-Euclidean proximity to 2) trans-crustal deformation zones as well as 3) felsic volcanic rocks, and 4) geochemical characteristics of host rocks.

Table 7-5 (this page and next). Detailed description of feature ranking and favorable signatures (within 1 km of known mineralization) that both support and deviate from expected results given the conceptual volcanogenic mineral system model. Explanations of ranking results are speculative and require further investigation

	Feature	Geological rationale	Favorable signature	Feature	Explanation
		High rank			**Low rank**
Predicted by conceptual model	Assemblage	Unique combination of melt source(s), ascent and eruption conditions, fluid-rock interactions	Later-stage, dominated mafic to fractionated felsic melts (i.e., Tisdale and Blake River)	Ultramafic volcanic (distance)	Rarely associated with mineralization
	Plag./Titanite fractionation (Eu/Eu*)	Plag./Titanite fractionation in shallow felsic plutonic source	Low (~0)	Ultramafic intrusion (distance)	
	Mafic intrusion (distance)	Commonly hosts mineralization and dominates greenstone belt	Proximal (~150)	Resistivity	Non-discriminatory rock properties for common host rocks
	TTG to felsic intrusion (distance)	Heat/melt source during convection of fluids through shallow crust	Proximal (100)	Density (grain)	
	E/W deformation zone splay (distance)	Pathway for convecting and venting fluid	Proximal (~150)	Magnetic susceptibility	
	CO$_2$ (gain/loss)	Indicator of fluid – rock interaction	Lower gains than background (~26)	Rheological contrast	Unnecessary fluid traps as metals often precipitate during seawater interactions
	SiO$_2$ (gain/loss)		Less loss than background (~0.97)	Chemical contrast	
	Na$_2$O (gain/loss)		Moderate gain (~1.05) compared to areas 2-3 km from targets	Structural complexity (lineament variance)	
	K$_2$O (gain/loss)		Moderate gain (~1.3) compared to areas 2-3 km from targets	Mafic compatibility (Mg/Zr)	Bi-modal host rocks vary
	Porosity (total)	High porosity of Timiskaming conglomeratic on the hanging wall of the Porcupine Destor shear zone	High (~1.05)	Melt signature (Ce/Y)	
	Clastic sed (distance)	Kidd Creek mine on hanging wall of the basal Pipestone Fault & adjacent to the Porcupine Basin	Proximal (~125)		

	Feature	Explanation	Value	Feature	Explanation
Unanticipated with conceptual model	Melt partitioning (Al/Ti)	**Geological:** Blake River and Kidd Munroe assemblages (host the majority of targets) have lower Al/Ti, which may reflect differing melt sources or underlying mantle plume conditions (Sproule et al., 2002)	Low (~20)	LOI	**Geological:** Some uncaptured volatile not relevant to volcanogenic mineralization may dominate LOI signal
				Trans-crustal deformation zone (distance)	**Geological:** Main branch for fluid ascent may be less favorable than splays linked to vents
				Melt evolution (Th/Nb)	**Methodological:** Low values where anticipated to reflect a thin crustal environment but these values may be obscured by the large range of values calculated for the entire dataset
				Felsic volcanic (distance)	**Geological:** Not all bi-modal volcanic rocks are favorable
	Chemical sed (distance)	**Methodological:** BIF covered by thick till and poorly mapped in regions with volcanogenic mineralization	Distal (~425)	Magnetic intensity	**Geological:** Destruction of magnetite from fluid interaction expected but signal obscured by many highly magnetic diabase dikes

7.3.2.3.2.3.   Limitations

Exploration for volcanogenic deposits depends greatly on mapping characteristic alteration haloes created as fluids convect through the upper crust. However, analytical and methodological limitations make generating reliable spatially continuous representation (i.e., maps) challenging. Methodological limitations have already been discussed (Section 7.2); however, it is important to add that many alteration indices and mass balance calculations are strongly affected by initial rock mineralogy/chemistry. Tools that be applied to regional datasets are underdeveloped and improvements with related feature engineering would be valuable.

7.3.2.3.3.   Orogenic

7.3.2.3.3.1.   Confirmation of conceptual model

The natural expression of orogenic conceptual modes can vary based on local geological setting. However, central themes for orogenic mineralization include 1) shortening driving deformation, magmatism, and metamorphism to generate 2) fluids and leach 3) metals to be later transported along 4) pathways and concentrated in 5) traps (e.g., Robert and Poulsen, 2001; Robert et al., 2005; Dubé and Gosselin, 2007; Groves and Santosh, 2015; Groves et al., 2018). Overall, ranked feature lists (Figures 7-4C, 7-4D, 7-5B) for classifying targets near Timmins and Dryden agree with priority exploration criteria and the conceptual model. A detailed description of high and low ranked features that were both expected and unanticipated given the conceptual mineral system model is provided in Table 7-6.

In the Timmins area, features that represent fluid mobility (transport/interaction) such as non-Euclidean proximity to trans-crustal deformation zones and their splays, porous host rocks, and LOI rank high. These features emphasize the importance of fluid-related controls over other system components such as sources (fluid/metal) and traps. Proximity to TTG – felsic intrusions also ranks moderately high; however, a clear interpretation of that importance and the role such intrusions

play is unclear. Some studies have linked granitoid intrusions as a source for fluid, heat, and/or metal during later orogenesis (Sillitoe and Thompson, 1998; Robert, 2001). Others suggest that these intrusive boundaries, where rheological contrast is high, are fluid pathways/traps.

In the Dryden area, the most important features ('High' in Figure 7-5B) when classifying orogenic targets also support the conceptual model and are similar to Timmins. For a direct comparison, see Figure 7-5B, where features with similar importance lay along the 1:1 line. Features that are preferred by the orogenic system in either area plot towards that locality's axis. Notable differences to Timmins are the higher rank of non-Euclidean distance to TTG – felsic intrusions (orogenic mineralization is generally found adjacent to intrusions), basement gneisses (far from mineralization), LOI (~3%), and magnetic intensity (low, magnetite destruction).

Table 7-6 (this page and next). Detailed description of feature ranking and favorable signatures (within 1 km of known orogenic Au mineralization) for the Timmins and Dryden areas that were anticipated and unanticipated in the conceptual orogenic mineral system model. Explanations of ranking results are speculative and require further investigation

		Feature	Geological rationale	Favorable signature	Feature	Explanation
		High rank			**Low rank**	
Predicted by conceptual model	TIMMINS	E/W trans-crustal deformation zone (distance)	Common pathways that connect fluids from source to traps	Proximal (~100)	Volcanic, intrusive, and clastic sed. rocks (distance)	Mineralization is hosted by a wide range of metamorphic protoliths whose compositions and physical characteristics vary
		E/W splay (distance)		Proximal (~60)	Primary chemical characteristics (Mg/Zr, Ce/Y, Th/Nb, Eu/Eu*, Al/Ti)	
		Porosity (total)		High (~1.1)		
		LOI	Indicator of fluid - rock interaction that mobilized volatiles	Moderate (~7.25 wt%)		
		TTG to felsic intrusion (distance)	Heat/Fluid source	Proximal (~80)	Primary rock properties (density, resistivity)	
	DRYDEN	E/W deformation zone splay (distance)	Common pathway that connects fluids from trans-crustal structure to traps	Proximal (~130)	Para-gneiss, volcanic, intrusive, and clastic sed. rocks (distance)	Mineralization is hosted by a wide range of metamorphic protoliths whose compositions and physical characteristics vary
		LOI	Indication of fluid - rock interaction	High (~3.5%)	Primary chemical characteristics (Mg/Zr, Ce/Y, Th/Nb, Eu/Eu*, Al/Ti)	
		SiO_2 (gain/loss)		Small loss (~0.93)		
		TTG to felsic intrusion (distance)	Heat/Fluid source	Proximal (~60)		
		Basement (distance)	Mineralization occurs in the upper crust as fluids ascend away from basement	Distal (~250)	Primary rock properties (porosity, density, resistivity)	

		Criterion	Geological	Value	Feature	Explanation
Unanticipated with conceptual model	**TIMMINS**	Assemblage	**Geological:** Favorable combination of melt source(s), ascent and eruption conditions, country rock interactions	Tisdale	CO_2, SiO_2, K_2O, Na_2O (gain/loss)	**Geological:** Only reactions with Auriferous fluids are prospective. **Methodological:** Poor capture of small/local anomalies associated with footprints
					Magnetic intensity	
		Melt signature (Ce/Y)	**Geological:** Characteristic signatures of tholeiitic mafic volcanic host rocks with optimal Fe/(Fe+Mg) ratios to cause desulfidation and metal precipitation	Depleted (~ 0.7)	Chemical sed (distance)	**Geological:** S-sourced elsewhere
		Melt evolution (Th/Nb)		Minimal crustal interaction/fractional crystallization (~0.2)	Rheological & Chemical contrast	**Geological:** These trap types are not significant for mineralization in this area
		Magnetic susceptibility		0.013 SI	Structural complexity	**Methodological:** Inappropriate scale/Poor capture of significant features
	DRYDEN	Assemblage	**Geological:** Favorable combination of melt source(s), ascent and eruption conditions, country rock interactions	Wabigoon	CO_2, SiO_2, K_2O, Na_2O (gain/loss)	**Geological:** Only reactions with Auriferous fluids are prospective. **Methodological:** Poor capture of small/local anomalies associated with footprints
					Magnetic intensity	
		Magnetic susceptibility	Characteristic signatures of tholeiitic mafic volcanic host rocks with optimal Fe/(Fe+Mg) ratios to cause desulfidation and metal precipitation	High (0.014 SI)	Chemical sed (distance)	**Geological:** S-sourced elsewhere
					Rheological & Chemical contrast	**Geological:** These trap types are not significant for mineralization in this area
		Magnetic intensity		High (-75 nT/m)	Structural complexity	**Methodological:** Inappropriate scale/Poor capture of significant features
					All trans-crustal deformation zones (distance)	Geological: The main conduit for fluids may not as prospective as splays in the Dryden area

7.3.2.3.3.2. Deviation from conceptual model

Structural and/or chemical traps adjacent to crustal-scale deformation zones are common exploration criteria that rank lower than anticipated. As fluids ascend along pathways, nearby traps such as zones of high structural complexity (e.g., antiform hinges) and interfaces between rocks with contrasting mechanical and/or chemistry (e.g., contacts between volcanic stratigraphy and syn-volcanic gabbro or tonalite-trondhjemite-granodiorite intrusions) are opportunely situated to encourage deposition of metals. However, such trap-related feature maps rank low for both areas and may signal a lesser dependence on the presence of traps in favor of source (fluid or metal) characteristics. It could also indicate a scale-related limitation where maps represent coarse regional features rather than smaller structures or localized trends trapping fluids and contributing to metal precipitation. Additionally, non-Euclidean distance to clastic sedimentary rocks ranks low for Timmins, and non-Euclidian distance to the NE-SW striking trans-crustal deformation zone (Manitou-Dinorwic) ranks low for Dryden, which was unexpected.

7.3.2.3.3.3. Limitations

No clear consensus exists on critical orogenic model components and their characterization beyond a vague need for a source of fluids/metals, a pathway for ascent, and a fluid trap where metals

precipitate. Despite being one of the most investigated deposit types and centuries of study, development of a consistent conceptual model is limited by the conditional and variable nature of orogenic systems. The distance between mineralization and metal/fluid source(s) can vary by orders of magnitude, leading to a challenging exploration environment where there may be direct linkages or a completely disconnected relationship to deposits. The nature of sources differs and will significantly impact geodynamic criteria for exploration (e.g., metamorphic dehydration/devolatilization, magmatic-hydrothermal fluids, direct release from mantle, expelled from external sources such as subducting oceanic plates). The nature of favorable traps widely varies and are inconsistently influenced by local conditions and other system components (e.g., connectivity in fluid pathways, chemistry of fluids vs host rock, pressure-temperature conditions).

Challenges that must be overcome to fully characterize orogenic models are: 1) Multi-scale features where the mappable expression of traps is smaller than other system components (i.e., sources and fluid pathways). In many cases, resolution of data, and consequently scale, are dependent on the coarsest dataset and features of interest causing sensitivity to be sacrificed for completeness. 2) Natural dependency of features and processes controlling mineralization is challenging to capture as mineralization if the result of many components working together in the correct way at the correct time. For example, the perfect trap in a map area is meaningless if it is not adequately connected to a favorable source under optimal conditions.

7.4.   Limitations in mineral systems modeling

While every reasonable effort is made to mitigate biases in training data and feature engineering techniques, the reality of modeling nature is that perfection is impossible. In such scenarios, clearly defining limitations is a critical step for result interpretability, confidence, and most especially, future learning.

The process of capturing a meaningful sliver of reality and making it useful for rigorous statistical interrogation is a daunting task requiring uncountable decisions and assumptions. During development of the presented feature engineering techniques and feature ranking workflow, several themes surrounding limitations emerged. It is uncommon to have limitations presented in detail; however, advancement requires open discourse. Therefore, subsequent sections will discuss limitations related to this application of feature engineering techniques and random forests to aid future prospectivity modeling and geologically focussed data science.

7.4.1. Training sets

Like all machine learning techniques, results from random forests are highly sensitive to the location, abundance, and nature of training points. Using 5 unique training sets for each mineral system as well as randomly subsampling during training attempted to overcome some bias.

7.4.1.1. Abundance of mineralized targets

The Abitibi, especially the Timmins region, has been integral to understanding fertile magmatic systems (Houlé and Lesher, 2011). Although they are not considered 'world class' in terms of endowment, deposits near Timmins(e.g., Alexo-Kelex, Langmuir, and Mc Watters, Redstone) represent some of the earliest global discoveries of Ni-Cu-PGE associated with komatiites. The abundance of magmatic-related deposits and occurrences (n=81 or 0.025 points/km2) near Timmins is sufficient to support training in random forests; however, they are not as numerous as volcanogenic and especially orogenic targets. In practice the fewer training examples likely lead to higher variability in feature ranking (longer boxes and whiskers in Figure 7-4).

The Abitibi is considered a world-class region hosting classic examples of volcanogenic deposits with variably high grade and tonnage. Notable deposits near Kamiskotia in Ontario and at Joutel, Mattagami Lake, Noranda, Bousquet-LaRonde, Val d'Or, and Chibougamau in Québec have been used to define fertile volcanogenic system models. Whilst the Abitibi is essentially the poster child for Archean volcanogenic mineralization, its reported targets are sparse in the Timmins area (n = 97 or 0.030 points/km2) and mainly dominated by the giant Kidd Creek deposit, which is itself an enigma compared to deposits in the eastern Abitibi (e.g., Noranda). Therefore, training sets for volcanogenic systems are sub-optimal here, but are acceptable for this method demonstration and knowledge gathering exercise. Future improvements should involve similar/comparative studies in other areas (prospective and not-prospective) to seek patterns in each of the mineral systems.

7.4.1.2. Nature of targets

Target points representing mineralization are not created equal. There is inconsistency in how target databases are generated, sometimes within the same mining district. A cluster of points are sometimes used to indicate a mineralized zone in a small area (i.e., occurrences) while a single point indicates the center of a major deposit and all its neighboring prospects/occurrences (e.g., Kidd Creek). In other areas, each prospect and occurrence outside of the main ore zone is indicated by individual points (e.g., Langmuir 1, 2, W2, W4). The inconsistent representation of mineralization variably imbalances targets to some extent and should, therefore, be acknowledged. Ideally, a consistent scheme would be used when compiling mineralized targets and points (e.g., deposit/mine, prospect, occurrence) would be classified based on endowment calculated/measured

at equivalent times (preferably up-to-date/current); however, this extra step can be time-consuming, requires the aggregation of information, and may imbalance (significantly more of one class versus another) the training set.

7.4.2. Feature engineering system components

Variable maps are continuous representations of some relevant feature that is defined by expert knowledge, derived from information (e.g., data, knowledge, interpretation), and captured using some process deemed to be 'appropriate' (e.g., gridding, statistics, value assignment). Several aspects of the design, generation, and interpretation of such maps introduce variably compounding bias and ambiguity from oversimplification. In these cases, documentation is the only option to move forwards; however, some methods have in-built uncertainty functions that should be preferred.

7.4.2.1. Bias

Subjective bias arising from using geological interpretations (e.g., bedrock map, geological history) is introduced in the early stages of every workflow and is compounded as investigations progress. For example, geological map interpretations are often the foundation from which variable maps are built upon. Their easy implementation, interrogation, and flexibility make them ideal for capturing concepts not otherwise possible. However, they themselves are the product of a complex process of iteratively interpolating /extrapolating from disparate datasets and integrating expert knowledge. While they may be the only dataset available, they should be used with caution. Herein, map unit polygons and interpreted structural features are used to constrain multiple variable maps (i.e., assemblage, structural complexity, non-Euclidean distance, rock properties, geochemical characteristics, and rheological/chemical contrast). Where possible, information captured by geological interpretations should be combined with independent datasets.

Subjective bias may also be introduced during early-stage decision making. For example, the calculation of non-Euclidean distance is the sum of weighted nodes and edges in a triangulated mesh (i.e., representation of knowledge graph). The assignment of the weights based on association to lithology or structure is relative and is subjectively defined by expert interpretation of which features will conduct/impede fluid. Many feature maps result from similar processes when minor, but impactful, decisions govern results. This type of bias is inevitable and should be reported sufficiently to facilitate repeatable outcomes.

Sample bias is, perhaps, the most recognized and bemoaned issue in any scientific discipline. Common themes in geoscientific sample distribution are: 1) Samples cannot be gathered in ideal locations for a host of reasons (e.g., logistics, political, environmental, below cover); 2) Samples are

not collected in currently 'insignificant' locations (i.e., in the middle of a batholith, basin, the center of a marsh). Neglected sample locations during field campaigns may later become important for several reasons but generally center on addressing new knowledge/ideas or balancing statistical analyses (e.g., interpolation, representing background/least altered); 3) Samples are clustered over 'interesting' areas and cause statistical imbalance (e.g., geochemical database herein where veins are not included in samples submitted for bulk geochemistry analyses, giving an apparent loss of SiO_2); and 4) Purpose-driven sampling (e.g., veins or thin map units vs least altered material with minimal veins representative lithologies) are not versicle in application, especially when compiling regional datasets to make broad inferences. The most appropriate example from the presented feature engineering methods are the semi-discrete gridding maps that attempt to characterise geochemical characterises of each lithology in each assemblage. The database used consists of samples of representative and rare/volumetrically small lithologies unevenly distributed across the map. When trying to capture geochemical variability in a given lithologic unit, several interpolation strategies and methods were tested. In the end, no solution emerged as a clear winner and a simplified but statistically applicable method was implemented. The main struggle in characterizing such layers is caused by spatially inconsistent sampling where some rock units are not represented in the database and must either be assigned a realistic value or interpolation coverage must be extended. A similar issue is present in the pseudo-mass balance maps. The absence of samples/analyses in certain areas results in unlikely interpolation results (e.g., Gold Rock area south of the Atikwa-Lawrence batholith near Dryden; see supplementary material Figure 7-5.4). This is significant where sample gaps are located along the edges of the map or sample set, especially if they represent anomalous values or differ from the nearby data points. There is no one-size-fits-all solution to these biases; however, adopting flexible sampling strategies and anticipating short and long-term requirements for data should guide sampling programs. If possible, samples should be added outside of the desired map extent to minimize edge effects and new sampling performed around anomalous data points. Herein, we use a 5km buffer of points beyond map extents to perform Empirical Bayesian Kriging on mobile major element gain/loss maps (Section 7.2.1.2)

7.4.2.2. Oversimplification and generalized assumptions

To make scientific investigations manageable, oversimplification or generalization are often necessary. To a point, geology can be viewed as fractal and up-scaling/down-scaling information is performed to make a dataset fit for purpose. In doing so, there is the possibility that features could be missed or eliminated. Herein, complex interactions between variably dependent or coupled geological processes are generalized to be representative of components in the conceptual mineral system models at a scale appropriate for a complete system view (i.e., regional). The issue is, natural

processes are non-linear and are governed by complex relationships and feedback loops. To make sense of such systems, linear approaches that relate an output of one process to the input of another are applied. However, applying linear approaches to non-linear problems results in oversimplification. The best example to illustrate this oversimplification are the pseudo mass balance (i.e., 'mobile element gain/loss factors'). In this feature engineering method, linear functions average and extrapolate concentrations of mobile elements in 'least-altered' samples. The functions are constrained by mobile element concentrations in those samples and a lithology normalization variable (here igneous = Mg/Zr and sedimentary = Al_2O_3). Certainly, a single value cannot fully capture the complex interdependencies and assemblages of minerals making up a given lithology, and a linear function is a rigid fit to naturally varying data. The nature of a rock type is oversimplified in that the single term herein does not explicitly capture petrologic heterogeneity in primary rock types from initial melt conditions/sources, melt evolution during ascent (e.g., fractional crystallization, variable degrees of partial melting, crustal contamination/interaction), fluid chemistry, as well as pressure-temperature conditions during fluid-rock interaction. However, alteration is a key component for exploration and increasing dependence on machine learning techniques requires spatially continuous inputs. To that end, a compromise between reasonability and ease of implementation was made. In future, perhaps envelopes around least altered samples, functions with higher degrees of freedom, multi-dimensional normalization for lithology (i.e., more than one element), and vast geochemical/petrophysical databases could progress this technique.

Other oversimplifications or poor representation may also arise from methodological limitations. For example, the meshing technique used to generate the basis for graph analysis (i.e., network of nodes and edges of map unit polygons) is irregular and mis-represents thin vs large units. For a thin lithology, especially those that are structurally complex (i.e., folded), a higher density of nodes is required to constrain the shape of the polygon. In contrast, a large, rounded pluton will require few nodes to constrain its shape. Since the meshing technique creates triangles based on nodes rather than surface area, the thin lithology will have many more triangles than the pluton. When summing weighted nodes and edges to calculate non-Euclidean distances, both map units will be poorly represented. In future, new meshing techniques or regularization of polygon constraints may improve upon the current method. Despite limitations, methods that capture geological knowledge and represent challenging features are needed for machine learning. While oversimplification is inevitable when capturing features/processes in real world settings, it is important to clearly describe assumptions and limitations.

Another example of oversimplification is the semi-discrete element ratio and pseudo-mass balance maps. When interpolating and extrapolating values, a regular gid in Euclidean space is used rather

than one that is representative of deformation and stratigraphic relationships. To improve accuracy, a tool that constrains interpolations to a geologically representative grid (i.e., deformed) would be needed; however, its generation would require significant effort and knowledge to be done correctly.

7.4.3. Dimension vs doable

The reality of modeling three-dimensional features using two-dimensional methods is that information is sacrificed for utility. By eliminating the depth dimension, relationships to underlying features and processes that involve vertical transport (e.g., fluids, energy) are lost. In a mineral systems application, this loss is significant as some of the most important components are poorly captured or absent.

Geophysical inversion and explicit/implicit 3D modelling are the most common tools to capture sub-surface architecture and provide the basis for feature engineering; however, much like their 2D counterparts (i.e., geological maps), they introduce uncertainty and ambiguity that compounds with each stage in the workflow. Whether using inversions or projections from surface/drill hole data, the amount of constraint information is significantly less in a domain that is orders of magnitude larger. Additionally, traditional 2D mapping has the advantage of access where interpretations or projections can be verified in the field or by drilling.

Despite these drawbacks, it is logical that making efforts to include 3D constraints in prospectivity is still useful. In the end, a reasonably constrained and representative feature model is better than complete neglect of underlying features and vertically controlled processes.

7.5.   Geological outcomes

While this application of random forests is used as a validation of geological representativeness of a wide range of feature engineering methods, importance rankings can be used to support existing geological knowledge and provide new insight.

7.5.1. Mineral systems

It is widely accepted that different mineral systems can exploit the same features when available, which should lead to similar importance rankings. In Figure 7-5A, the proportional importance for each of the three mineral systems are presented in a ternary diagram. Features with similar importance in all models are located near the center and those that are more dominant in one system over another move towards the edges.

Classification using magmatic targets for training rank mafic compatibility (Mg/Zr), K_2O, distance to clastic sedimentary map units, and rheological/chemical contrast highly compared to the other

systems. As discussed previously (Section 7.1.1.1), high magnesian igneous rocks are characteristic host rocks for magmatic deposits and rarely associated with the other deposit types. The higher importance of K_2O can be explained by (an inverse relationship of very low K_2O to high MgO in ultramafic rocks (see supplementary material Figure 7-5.1). The high rank of rheological contrast may be related to favorable conditions for higher and more voluminous magma ascent that dictate emplacement of mineralized magma (e.g., flows versus invasive flows versus sills and/or dikes; Houlé et al., 2008, Lesher, 2019). Throughout the Abitibi, characteristically high rheological contrast mark zones where volcanic flows or sedimentary stratigraphy are in contact with unconsolidated volcaniclastic rocks (herein classified as 'felsic') or sediments (Houlé et al., 2008). The remaining uniquely dominant features are puzzling, and no clear geological or methodological explanation emerges. Bearing in mind that these features have some of the largest variance in Figure 7-4, it is reasonable to assume that they are artifacts of limited/biased training points that broadly represent three clusters of moderately endowed deposits. Comparable studies in regions (e.g., Yilgarn Craton in Western Australia) with more abundant and endowed deposits could resolve this ambiguity.

The most important features for classifying volcanogenic targets are proxies for chemical FIII rhyolites (Eu/Eu* and Al/Ti) and distance to mafic volcanic map units. FIII rhyolites that are fractionated and indicate a shallow felsic melt source in a thin crustal geodynamic setting are characteristic for volcanogenic deposits and are rarely significant components in other mineral systems (Lesher et al., 1986; Montsion et al., submitted b/Chapter 5). Distance to mafic volcanic map units at first glance support the observation that volcanogenic deposits develop in regions with bi-modal volcanism. Contrary to initial assumptions, signatures within 1 km of targets (supplementary material Figure 7-9.2) indicate that being moderately far away from mafic map units is favorable. This can be explained as a methodological limitation and reflects map-scale observations that volcanogenic mineralization is mainly spatially associated with felsic volcanic rocks in the Timmins area (see map in Figure 3-3).

Dominant features used to classify orogenic targets are 1) fluid-related features such as LOI, Na_2O, distance to mafic to felsic intrusions (possible fluid source from dewatering), as well as distance to E/W trans-crustal structures and their splays. In the conceptual model, the hydrothermal nature of the system and strong genetic connection to fluids predicted that features related to tracking fluid path and interactions would rank highly. 2) Features representing Fe-rich chemical traps such as distance to ultramafic volcanic map units, as well as tholeiitic basalt chemical signatures (Ce/Y, Th/Nb).

7.5.2. Controls on contrasting Au in Timmins and Dryden

The greenstone belts near Timmins and Dryden, Ontario have similar geologic histories yet host contrasting amounts of reported orogenic Au mineralization. By comparing feature rankings for models trained using orogenic targets in each area (Figure 7-4C and D; Figure 7-5B), several geologically relevant observations can be made. Overall, features rank similarly (plot along 1:1 line) and support the conceptual model. Assemblage ranks highly for both areas, but favorable signatures in supplementary material Figure 7-9.1 indicate that mid-stratigraphic levels (mainly Tisdale assemblage) host deposits near Timmins while lower crustal levels (mainly Wabigoon assemblage) host mineralization near Dryden.

The main differences between the two areas are 1) significantly higher importance for proximity to E/W striking trans-crustal structures in Timmins, and 2) proximity to TTG - felsic intrusions has higher importance in Dryden. When combined, these observations may suggest differing sources of metalliferous fluid such that magmatic-hydrothermal systems and marine/turbiditic sedimentary metamorphic sources are the dominant fluid supply during orogenesis in well endowed greenstone belts (e.g., in the Abitibi, Pitcairn et al., 2021) while magmatic-hydrothermal systems are the dominant fluid supply in poorly endowed greenstone belts (e.g., in the Swayze, Katz et al., 2017)

The strong spatial association between orogenic Au, marked by a variety of alteration signatures, and vertically extensive fluid pathways in Timmins is well documented and has been used to develop broadly accepted conceptual models (Thompson, 2005; Dubé et al., 2020). Trans-crustal structures in Dryden are marked by less pronounced alteration, are significantly less associated with mineralization, and do not have large proximal marine/turbiditic sedimentary basin, possibly indicating smaller and more localized fluid and metal sources.

Orogenic Au association to TTG to felsic plutons is strong in both areas, though more dominant in Dryden. The spatial association, paired with low importance of rheological contrast that would have indicated favorable structural/chemical traps, argues toward magmatic-hydrothermal/intrusion-related orogenic model (Thompson, 2005; Stromberg, 2017; Dubé et al., 2020) being common to both areas.

7.6.   Implications for exploration strategies

Data scientific techniques used during feature engineering were repurposed from applications in other scientific disciplines. Selected statistical techniques were interpreted to best approximate or highlight relevant geological features; however, many more statistical tools have been developed to solve a myriad of scientific problems. Machine learning and statistical analyses have great potential to empower geoscientists so they may generate meaningful representations that both reduce

subjective biases and leverage domain knowledge. Concepts and processes that are challenging to represent can be embedded, which enhances the utility of datasets. However, caution should be used when selecting appropriate techniques. The fundamental concepts and limitations of a technique should be well understood before its application to solve a geological problem.

Based on the agreement of conceptual mineral system models and random forests' ranked feature lists, this study demonstrates the potential for generating a more comprehensive assessment of mineral systems, including capturing processes or features that may have been overlooked using traditional approaches. By incorporating both local and general geological knowledge, a foundational understanding of how data reflects physical rocks, as well as statistical techniques, we can improve the utility of existing data. Additionally, the iterative stages of decision-making, interrogation, and sensitivity testing in this study also resulted in a more detailed understanding of the features being represented. Jointly considering geological questions and appropriateness of data/statistical methods challenged interpretations and drove innovation to meet newly identified gaps. Ultimately, as exploration workflows move towards quantitative approaches, the need to give old data new life, generate new geological ideas, and reduce ambiguity in knowledge-driven methods will emphasize the importance of such interdisciplinary methods.

7.7. Conclusion

Geologically appropriate feature engineering is often underdeveloped in geoscientific applications of machine learning. Geological plausibility of inputs is critical to generate robust results and gain meaningful information from machine learning. New and existing geologically representative feature engineering techniques were designed and demonstrated to comprehensively capture magmatic, volcanogenic, and orogenic mineral systems. Comparisons of ranked feature lists generated by random forests to each conceptual mineral system model were used to validate geological representativeness of feature engineering methods applied to two greenstone belts in the southern Abitibi (near Timmins, Ontario) and western Wabigoon (near Dryden, Ontario) subprovinces of the Superior Province.

Structural complexity, identifying zones of rheological and chemical contrast, and geologically constrained interpolation of characteristic element ratios (Mg/Zr, Ce/Y, Th/Nb, Eu/Eu*, Al/Ti), calculation and interpolation of mobile element gain/loss (CO_2, SiO_2, Na_2O, K_2O), interpolation of loss on ignition analyses, assemblage maps, filtered magnetic intensity, and rock property maps (magnetic susceptibility, grain density, total porosity, rock resistivity) were used to generate geologically plausible representations of favorable components identified in conceptual mineral system models.

Geologic representativeness in feature maps was qualitatively assessed by comparing ranked feature lists produced by random forests to conceptual mineral system models. Broadly, importance lists are consistent with expectations from the conceptual models and local geological knowledge. This strong agreement suggests that relevant features adequately represent their intended subjects and that feature engineering was useful. A few exceptions were observed and are mainly attributed to biased (low abundance and clustered) training data or feature maps affected by limited constraint data; however, further testing is required. While assessing importance rankings, models suffering from training bias were indicated by wide distributions of importance distributions during 500 validation runs. In these cases, feature maps related to characteristics of the location where training points were clustered obscure underlying system components.

Comparison of feature importance for each of the targeted systems in the Timmins area reaffirms that many geological features exploit the same features when available. System-specific characteristics include:

Magmatic

- Confirmation: Proxies for high stratigraphic level ultramafic host rocks located adjacent to trans-crustal structures or in areas with high rheological contrast
- Deviation: Low importance of ultramafic characteristics

Volcanogenic

- Confirmation: Maps/Proxies of high stratigraphic level comprising bi-modal mafic/felsic volcanic stratigraphy characterized by fractionated geochemical signatures. Favorable host rocks display common volcanogenic alteration signatures for CO_2, SiO_2, Na_2O, K_2O and are proximal to TTG – felsic intrusions as well as splays of spatially extensive deformation zones
- Deviation: Minorly affected by biased training data causing maps related to sedimentary rocks (i.e., total porosity and distance to clastic/chemical sediments) to rank highly. LOI, non-Euclidean distance to trans-crustal deformation zones, geochemical melt evolution signature (Th/Nb), and magnetic intensity ranked lower than anticipated and require further investigation

Orogenic - Timmins

- Confirmation: Proxies/Maps indicating potential fluid mobility (i.e., structural pathways, porosity, LOI, TTG – felsic intrusions), especially proximity to E/W trans-crustal deformation zones

- Deviation: Proxies for tholeiitic volcanic stratigraphy (Ce/Y, Th/Nb, magnetic susceptibility) suggests the importance of Fe-rich host rocks is higher than anticipated. Indications of alteration (CO_2, SiO_2, Na_2O, K_2O) and mechanical traps rank low and may indicate a scale-related methodological limitation or that traps are less important than other components in the system.

Orogenic – Dryden

- Confirmation: Proxies/Maps indicating potential fluid mobility (i.e., structural pathways, porosity, LOI, SiO_2), especially TTG – felsic intrusions
- Deviation: Non-Euclidean distances away from the surface expression of basement litho-tectonic blocks rank highly. Indications of alteration (CO_2, Na_2O, K_2O) and mechanical traps rank low and may indicate a scale-related methodological limitation or that traps are less important than other components in the system.

The main differences between the greenstone belts near Timmins and Dryden, Ontario (described above) may indicate fluid source is the controlling factor on metal endowment. Near Timmins, ranking results suggest that both magmatic-hydrothermal systems and marine/turbiditic sedimentary metamorphic sources (possibly in inverted back-arc-like basin settings) are the dominant fluid supply during orogenesis (e.g., Pitcairn et al., 2021). Near Dryden, only magmatic-hydrothermal systems (possibly in continental-arc-like setting overlying gneissic basement settings) are the dominant fluid supply (e.g., Swayze greenstone belt in the western Abitibi, Katz et al., 2017)

Integrative studies in other areas/mineral systems that leverage multi-disciplinary data and methodology have potential to advance geological understanding, maximize data utility, and generate robust exploration targets. As exploration workflows adopt quantitative approaches, the need to incorporate existing data, expand databases, and project beyond constrained data will drive innovation towards interdisciplinary approaches.

Chapter 8. Conclusion

Geologically appropriate feature engineering is underdeveloped in geoscientific applications of machine learning. Poor representation by features derived from map interpretations and purely measured/analysed datasets lack the comprehensibility needed to capture complex geological systems. By combining local and general geological knowledge, fundamental understanding of how natural features are represented by data, and using appropriate data science techniques, we demonstrate a means to capture mineral system components and gain new geological insight.

Compiled open-source and new field observations from Archean greenstone belts near Timmins and Dryden, Ontario in the southern Superior Province were used to demonstrate existing and new geologically appropriate feature engineering methods. Data resources include reprocessed geophysical grids, newly digitized outcrop lithologies, structural measurements, geochronological data, whole rock geochemistry, mineral deposit indices. These constraints were overlayed with existing interpretations to generate regionally seamless geological maps in both areas, and provide the basis for feature engineering. Four field seasons (2018-2021) provided a means to validate interpretations, fill in gaps, enhance existing knowledge, and guide design of relevant feature engineering methods.

Using the newly compiled lithogeochemistry database, a new classification scheme was designed to characterize and classify Precambrian igneous rocks. Using the new 5D classification scheme, rocks from the full compositional range (ultramafic to felsic + TTG + lamprophyre) and variably affected by metasomatism can be compared, which facilitates a regional comparison of individual samples and broad comparison between differing geological domains. Using principal component analysis, five geologically relevant element ratios that statistically maximize the spread of data highlight variations in mantle/crustal sources as well as relative degrees of crustal interaction/assimilation, fractional crystallization, and accumulation. Additionally, samples proximal to magmatic, volcanogenic, and orogenic deposits in the southern Superior Province were further used to define three new Igneous Rock Favorability (IRF) indices near Timmins and Dryden, Ontario, which should be applicable to other Archean cratons. Their implementation into future mineral exploration studies may ultimately lead to more efficient targeting.

Geologically focussed feature engineering methods demonstrated in these two Archean greenstone belts include non-Euclidean distance to map units and structures, rheological and chemical contrast, structural complexity, semi-discrete interpolation of characteristic element ratios, calculation and interpolation of mobile element gain/loss, as well as utilization of rock properties. These methods

were designed and selected to represent favorable components in conceptual models for magmatic, volcanogenic, and orogenic mineral systems.

Geologic representativeness of the feature maps was qualitatively assessed by comparing ranked feature lists produced by random forests to conceptual models. Broadly, importance lists are consistent with expectations from the conceptual models and local geological knowledge. This strong agreement suggests that relevant features adequately represent their intended subjects and that feature engineering was useful. A few exceptions were observed and are mainly attributed to biased (low abundance and clustered) training data or feature maps affected by limited constraint data. While assessing importance rankings, models suffering from training bias were indicated by wide distributions of importance distributions during 500 validation runs. In these cases, feature maps related to characteristics of the location where training points were clustered obscure underlying system components.

Comparison of feature importance for each of the targeted systems in the Timmins area reaffirms that many geological features exploit the same features when available. A comparison of features important when classifying orogenic Au mineralization in Timmins and Dryden may provide some insight as to the cause of contrasting endowment. Notable, a high ranking for proximity to E/W trans-crustal deformation zones in Timmins supports the long established and important control on focusing large volumes of fluid ascending from depth. Conversely, a high ranking for proximity to TTG plutons in Dryden and a low ranking for E/W trans-crustal deformation zones may indicate that metalliferous fluids were primarily sourced from nearby intrusions rather than deep crustal levels.

Participation in the entire workflow from data collection, conceptualization, methodological selection/design, application, and interpretation of results afforded a wider view of methodological capabilities and highlighted areas for improvement. Biases related to imperfect distribution of constraint or training data, limitations in technique capability to isolate relevant signals, necessary oversimplification and generalized assumptions, as well as reduced comprehensibility of 2D datasets ultimately affect modeling capabilities. While many of these limitations are inevitable when interacting with the natural world, their influence on results could be better described.

Ultimately, the integration of geological knowledge and data science is a developing domain of geoscience. Integrative studies leveraging multi-disciplinary data and methodology have potential to advance geological understanding, maximize data utility, and generate robust exploration targets. Careful thought experiments that engage a range of expert knowledge during project design will benefit immediate geological problems and have far reaching impacts on exploration approaches.

References

Acosta, I.C.C., Khodadadzadeh, M., Tusa, L., Ghamisi, P., Gloaguen, R., 2019. A Machine learning framework for drill-core mineral mapping using hyperspectral and high-resolution mineralogical data fusion. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 12, 4829–4842. https://doi.org/10.1109/JSTARS.2019.2924292

Aitken, A.R.A., Betts, P.G., 2009. Multi-scale integrated structural and aeromagnetic analysis to guide tectonic models: An example from the eastern Musgrave Province, Central Australia. Tectonophysics 476, 418–435. https://doi.org/10.1016/j.tecto.2009.07.007

Almasi, A., Yousefi, M., Carranza, E.J.M., 2017. Prospectivity analysis of orogenic gold deposits in Saqez-Sardasht Goldfield, Zagros Orogen, Iran. Ore Geol. Rev. 91, 1066–1080. https://doi.org/10.1016/j.oregeorev.2017.11.001

Appleyard E.C., 1980. A preliminary metasomatic assessment of laminated siltstones at the Silverfields mine, Cobalt, Ontario, in: Pye E.G. (Ed.), Grant 70 Geoscience Research Grant Program Summary of Research 1979–1980, Ontario Geological Survey Miscellaneous Paper (MP093), 262p. http://www.geologyontario.mndm.gov.on.ca/mndmfiles/pub/data/imaging/MP093/MP093.pdf

Appleyard E.C., 1990. Mass balance corrections applied to lithogeochemical data in mineral exploration, in: Beck L.S. and Harper C.T. (Eds.), Modern exploration techniques, Saskatchewan Geological Society Special Publication 10, pp. 27–40. https://archives.datapages.com/data/sgs/1990/027/027.htm

Ayer, J.A., Amelin, Y., Corfu, F., Kamo, S., Ketchum, J., Kwok, K., Trowell, N., 2002. Evolution of the southern Abitibi greenstone belt based on U–Pb geochronology: autochthonous volcanic construction followed by plutonism, regional deformation and sedimentation. Precambrian Res. 115, 63–95. https://doi.org/10.1016/0730-725X(91)90535-T

[Dataset] Ayer, J.A. and Chartrand, J.E., 2011. Geological compilation of the Abitibi greenstone belt. Ontario Geological Survey Miscellaneous Release Data (MRD 282). ISBN 978-1-4435-6834-0 (CD); ISBN 978-1-4435-6835-7. http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=MRD282

[Dataset] Ayer, J.A., Trowell, N.F. and Josey, S., Geological compilation of the Abitibi greenstone belt, 2004. Ontario Geological Survey, Miscellaneous Release Data (MRD 143). http://www.geologyontario.mndmf.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=MRD143

Ayer, J.A., Thurston, P.C., Bateman, R., Dubé, B., Gibson, H.L., Hamilton, M.A., Hathway, B., Hocker, S.M., Houlé, M.G., Hudak, G.J., Ispolatov, V., Lafrance, B., Lesher, C.M., MacDonald, P.J., Péloquin, A.S., Piercey, S.J., Reed, L.E., Thompson, P.H., 2005. Overview of results from the Greenstone Architecture Project: Discover Abitibi Initiative, Ontario Geological Survey. Open File Report (OFR 6154). 177 p. http://www.geologyontario.mndm.gov.on.ca/mndmfiles/pub/data/records/OFR6154.html

[Dataset] Ayer, J.A., Thurston, P.C., Bateman, R., Gibson, H.L., Hamilton, M.A., Hathway, B., Hocker, S.M., Hudak, G., Lafrance, B., Ispolatov, V., MacDonald, P.J., Péloquin, A.S., Piercey, S.J., Reed, L.E., Thompson, P.H. and Izumi, H., 2005. Digital Compilation of Maps and Data from the Greenstone Architecture Project in the Timmins-Kirkland Lake Region: Discover Abitibi Initiative. Ontario Geological Survey, Miscellaneous Release Data (MRD 155). http://www.geologyontario.mndmf.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=MRD155

Baratoux, D., 2019. A multi-scale roughness map of the Guiana shield, in: SAXI- XI Inter Guiana Geological Conference 2019. Paramaribo, Suriname, pp. 15–18.

Barker, F., Arth, J.G., 1976. Generation of trondhjemitic-tonalitic liquids and Archean bimodal trondhjemite-basalt suites. Geol. 4, 596–600. https://doi.org/10.1130/0091-7613(1976)4<596:GOTLAA>2.0.CO;2

Barnes, S.J., Fiorentini, M.L., 2012. Komatiite magmas and sulfide nickel deposits; a comparison of variably endowed Archean terranes. Econ. Geol. 107, 755–780. https://doi.org/10.2113/econgeo.107.5.755

Barnes, S.J., Robertson, J.C., 2019. Time scales and length scales in magma flow pathways and the origin of magmatic Ni-Cu-PGE ore deposits. Di xue qian yuan. 10, 77–87. https://doi.org/10.1016/j.gsf.2018.02.006

Barnes, S.J., Fiorentini, M.L., 2012. Komatiite magmas and sulfide nickel deposits; a comparison of variably endowed Archean terranes. Econ. Geol. 107, 755–780. https://doi.org/10.2113/econgeo.107.5.755

Barnes, S.J., Lesher, C.M., Sproule, R.A., 2007. Geochemistry of komatiites in the Eastern Goldfields Superterrane, Western Australia and the Abitibi greenstone belt, Canada, and implications for the distribution of associated Ni-Cu-PGE deposits. Inst. Min. Metall. Trans. Sect. B Appl. Earth Sci. 116, 167–187. https://doi.org/10.1179/174327507X271996

[Map] Bateman, R., 2005. Precambrian geology of Tisdale Township and parts of Deloro, Mountjoy and Ogden townships, Ontario Geological Survey Preliminary Map (P 3555). scale 1:10 000. http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=P3555

Bateman, R., Ayer, J.A., Dubé, B., Hamilton, M.A., 2005. The Timmins–Porcupine gold camp, northern Ontario: the anatomy of an Archaean greenstone belt and its gold mineralization: Discover Abitibi Initiative; Ontario Geological Survey Open File Report (OFR 6158). https://dx.doi.org/10.2113/gsecongeo.103.6.1285

Bateman, R., Ayer, J.A., Dubé, B., 2008. The Timmins-Porcupine gold camp, Ontario: Anatomy of an Archean greenstone belt and ontogeny of gold mineralization. Econ. Geol. 103, 1285–1308. https://doi.org/10.2113/gsecongeo.103.6.1285

Baxter, M.J., 1995. Standardization and Transformation in Principal Component Analysis, with Applications to Archaeometry. J. R. Stat. Soc. Ser. C (Applied Stat. 44, 513–527. https://doi.org/10.2307/2986142

Beakhouse, G. P., Blackburn, C. E., Breaks, F. W., Ayer, J. A., Stone, D., & Stott, G. M., 1995. Western Superior Province fieldtrip guidebook. Ontario Geological Survey Open File Report (OFR 5924). http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=OFR5924

[Dataset] Beakhouse, G. P., Webb, J. L., Rainsford, D. R. B., Stone, D., & Josey, S. D., 2011. Western Wabigoon GIS synthesis-2011. Ontario Geological Survey Miscellaneous Release—Data (MRD 280). http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=MRD280

Begg, G.C., Hronsky, J.A.M., Arndt, N.T., Griffin, W.L., O'Reilly, S.Y., Hayward, N., 2010. Lithospheric, cratonic, and geodynamic setting of Ni-Cu-PGE sulfide deposits. Econ. Geol. 105, 1057–1070. https://doi.org/10.2113/econgeo.105.6.1057

[Map] Berger, B., 2010. Precambrian Geology of the Kidd-Munro assemblage. Ontario Geological Survey Preliminary Map (P 3614). scale 1:100 000. http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=P3614

[Dataset] Berger, B.R., Chapman, J.B., Peter, J.M., Layton-Matthews, D. and Gemmell, J.B., 2011. Geochemical data from the Targeted Geoscience Initiative III Kidd-Munro Project, Ontario Geological Survey Miscellaneous Release Data (RMD 277). ISBN 978-1-4435-6312-3. http://www.geologyontario.mndmf.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=MRD277

Bérubé, C.L., Olivo, G.R., Chouteau, M., Perrouty, S., Shamsipour, P., Enkin, R.J., Morris, W.A., Feltrin, L., Thiémonge, R., 2018. Predicting rock type and detecting hydrothermal alteration using machine learning and petrophysical properties of the Canadian Malartic ore and host rocks, Pontiac Subprovince, Québec, Canada. Ore Geol. Rev. 96, 130–145. https://doi.org/10.1016/j.oregeorev.2018.04.011

Bierlein, F.P., Murphy, F.C., Weinberg, R.F., Lees, T., 2006. Distribution of orogenic gold deposits in relation to fault zones and gravity gradients: Targeting tools applied to the Eastern Goldfields, Yilgarn Craton, Western Australia. Miner. Depos. 41, 107–126. https://doi.org/10.1007/s00126-005-0044-4

[Dataset] Biswas, S., 2019. Ontario Precambrian bedrock magnetic susceptibility geodatabase for 2001 to 2017. Ontario Geological Survey Miscellaneous Release—Data (MRD 273–Revision 2). http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=MRD273-REV2

[Dataset] Bjorkman, K.E. and Hollings, P., 2018. Geochemical analyses of rocks in the central Wabigoon superterrane, northwestern Ontario, Ontario Geological Survey Miscellaneous Release—Data (MRD 370). http://www.geologyontario.mndmf.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=MRD370

[Map] Blackburn, C. E., Beard, R. C., and Rivett, A. S., 1981. Geological compilation, Kenora–Fort Frances, Kenora and Rainy River districts. Ontario Geological Survey Preliminary Map (P 2443). Scale 1:253 440. http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=M2443

[Map] Blackburn, C.E., 1978. Upper Manitou Lake, Kenora District, Ontario Geological Survey Preliminary Map (P 2409). Scale 1:31 680. http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=P2409

Bleeker, W., 2015a. Contributions to the understanding of Precambrian lode gold deposits and implications for exploration, in: Dubé, B., Mercier-Langevin, P. (Eds.), Targeted Geoscience Initiative 4: Contributions to the Understanding of Precambrian Lode Gold Deposits and Implications for Exploration. Geological Survey of Canada, Ottawa, Canada, pp. 1883–1884. https://doi.org/10.4095/296626

Bleeker, W., 2015b. Synorogenic gold mineralization in granite-greenstone terranes: the deep connection between extension, major faults, synorogenic elastic basins, magmatism, thrust inversion, and long-term preservation, in: Dubé, B., Mercier-Langevin, P. (Eds.), Targeted Geoscience Initiative 4: Contributions to the Understanding of Precambrian Lode Gold Deposits and Implications for Exploration. Geological Survey of Canada, pp. 25–47. https://doi.org/10.4095/296626

Blundell, C.C., Armit, R., Aillères, L., Micklethwaite, S., Martin, A., Betts, P.G., 2019. Interpreting geology from geophysics in poly-deformed and mineralised terranes; the Otago Schist and the Hyde-Macraes Shear Zone. New Zeal. J. Geol. Geophys. 62, 550–572. https://doi.org/10.1080/00288306.2019.1579741

Bogolepov, V.G., 1963. The recomputation of the chemical analyses ol rocks in studying metasomatic processes. Int. Geol. Rev. 5, 1585–1592. https://doi.org/10.1080/00206816309473899

Boisvert, J.B., Deutsch, C. V, 2011. Programs for kriging and sequential Gaussian simulation with locally varying anisotropy using non-Euclidean distances. Comput. Geosci. 37, 495–510. https://doi.org/https://doi.org/10.1016/j.cageo.2010.03.021

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24, 123–140. https://doi.org/10.1007/BF00058655

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324

Brown, W.M., Gedeon, T.D., Groves, D.I., Barnes, R.G., 2000. Artificial neural networks: A new method for mineral prospectivity mapping. Aust. J. Earth Sci. 47, 757–770. https://doi.org/10.1046/j.1440-0952.2000.00807.x

Burkin, J.N., Lindsay, M.D., Occhipinti, S.A., Holden, E.-J., 2019. Incorporating conceptual and interpretation uncertainty to mineral prospectivity modelling. Geosci. Front. 10, 1383–1396. https://doi.org/10.1016/j.gsf.2019.01.009

Burns, K.L., 1975. Analysis of geological events. J. Int. Assoc. Math. Geol. 7, 295–321. https://doi.org/10.1007/BF02081703

Burrows, D.R., Spooner, E.T.C., 1989. Relationships between Archean Gold Quartz Vein-Shear Zone Mineralization and Igneous Intrusions in the Val d'Or and Timmins Areas, Abitibi Subprovince, Canada. Geol. Gold Depos. Perspect. 1988. https://doi.org/10.5382/Mono.06.33

Burrows, D.R., Wood, P.C., Spooner, E.T.C., 1986. Carbon isotope evidence for a magmatic origin for Archaean gold-quartz vein ore deposits. Nature 321, 851–854. https://doi.org/10.1038/321851a0

Cameron, E.M., 1988. Archean gold: relation to granulite formation and redox zoning in the crust. Geology 16, 109–112. https://doi.org/10.1130/0091-7613(1988)016<0109:AGRTGF>2.3.CO;2

Cameron, E.M., 1989. Scouring of gold from the lower crust. Geology 17, 26–29. https://doi.org/10.1130/0091-7613(1989)017<0026:SOGFTL>2.3.CO;2

Campbell, I.H., Franklin, J.M., Gorton, M.P., Hart, T.R., Scott, S.D., 1981. The Role of subvolcanic sills in the generation of massive sulfide deposits. Econ. Geol. 2248–2253. https://doi.org/10.2113/gsecongeo.76.8.2248

Campos, L.D., de Souza, S.M., de Sordi, D.A., Tavares, F.M., Klein, E.L., Lopes, E.C. dos S., 2017. Predictive Mapping of Prospectivity in the Gurupi Orogenic Gold Belt, North–Northeast Brazil: An Example of District-Scale Mineral System Approach to Exploration Targeting. Nat. Resour. Res. 26, 509–534. https://doi.org/10.1007/s11053-016-9320-5

Canny, J., 1986. A Computational Approach to Edge Detection. IEEE Trans. Pattern Anal. Mach. Intell. PAMI-8, 679–698. https://doi.org/10.1109/TPAMI.1986.4767851

Carranza, E.J.M., Hale, M., 2001. Geologically Constrained Fuzzy Mapping of Gold Mineralization Potential, Baguio District , Philippines. Nat. Resour. Res. 10, 125–136. https://doi.org/10.1023/A:1011500826411

Caté, A., Schetselaar, E., Mercier-Langevin, P., Ross, P.S., 2018. Classification of lithostratigraphic and alteration units from drillhole lithogeochemical data using machine learning: A case study from the Lalor volcanogenic massive sulphide deposit, Snow Lake, Manitoba, Canada. J. Geochemical Explor. 188, 216–228. https://doi.org/10.1016/j.gexplo.2018.01.019

Ceriani, L., Verme, P., 2012. The origins of the Gini index: extracts from Variabilità e Mutabilità (1912) by Corrado Gini. J. Econ. Inequal. 10, 421–443. https://doi.org/10.1007/s10888-011-9188-x

Chen, Y., Wu, W., 2017. Mapping mineral prospectivity using an extreme learning machine regression. Ore Geol. Rev. 80, 200–213. https://doi.org/https://doi.org/10.1016/j.oregeorev.2016.06.033

Chicco, D., Tötsch, N., Jurman, G., 2021. The matthews correlation coefficient (Mcc) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. BioData Min. 14, 1–22. https://doi.org/10.1186/s13040-021-00244-z

Chicco, D., Tötsch, N., Jurman, G., 2021. The matthews correlation coefficient (Mcc) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. BioData Min. 14, 1–22. https://doi.org/10.1186/s13040-021-00244-z

Colvine, A.C., 1989. An empirical model for the formation of Archean gold deposits: products of final cratonization of the Superior province, Canada, The Geology of Gold Deposits: The Perspective in 1988. Economic Geology Monograph No. 6. https://doi.org/10.5382/Mono.06.03

Condie, K.C., 1989. Geochemical changes in basalts and andesites across the Archean-Proterozoic boundary: Identification and significance. Lithos 23, 1–18. https://doi.org/10.1016/0024-4937(89)90020-0

Condie, K.C., 2015. Changing tectonic settings through time: Indiscriminate use of geochemical discriminant diagrams. Precambrian Res. 266, 587–591. https://doi.org/https://doi.org/10.1016/j.precamres.2015.05.004

Corcoran, P.L., Mueller, W.U., 2007. Time-Transgressive Archean Unconformities Underlying Molasse Basin-Fill Successions of Dissected Oceanic Arcs, Superior Province, Canada. J. Geol. 115, 655–674. https://doi.org/10.1086/521609

Corfu, F., 1993. The evolution of the southern Abitibi Greenstone Belt in light of precise U-Pb geochronology. Econ. Geol. 88, 1323–1340. https://doi.org/10.2113/gsecongeo.88.6.1323

Coulter, D.W., Zhou, X., Wickert, L.M., Harris, P.D., 2017. Advances in Spectral Geology and Remote Sensing: 2008 – 2017, in: Tschirhart, V., Thomas, M.D. (Eds.), Exploration 17. Decennial Mineral Exploration Conferences, pp. 23–50.

Cox, K.G., Bell, J.D., Pankhurst, R.J., 1979. The interpretation of igneous rocks. Springer Dordrecht, London, Great Britain. https://doi.org/10.1007/978-94-017-3373-1

Cox, S.F., Wall, V.J., Etheridge, M.A., Potter, T.F., 1991. Deformational and metamorphic processes in the formation of mesothermal vein-hosted gold deposits - examples from the Lachlan Fold Belt in central Victoria, Australia. Ore Geol. Rev. 6, 391–423. https://doi.org/10.1016/0169-1368(91)90038-9

Cracknell, M.J., Reading, A.M., 2014. Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. Comput. Geosci. 63, 22–33. https://doi.org/https://doi.org/10.1016/j.cageo.2013.10.008

Curriero, F.C., 2006. On the use of non-euclidean distance measures in geostatistics. Math. Geol. 38, 907–926. https://doi.org/10.1007/s11004-006-9055-7

Davidson, J., Turner, S., Handley, H., Macpherson, C., Dosseto, A., 2007. Amphibole "sponge" in arc crust? Geology 35, 787–790. https://doi.org/10.1130/G23637A.1

Davis, D.W., Amelin, Y., Nowell, G.M., Parrish, R.R., 2005. Hf isotopes in zircon from the western Superior province, Canada: Implications for Archean crustal development and evolution of the depleted mantle reservoir. Precambrian Res. 140, 132–156. https://doi.org/10.1016/j.precamres.2005.07.005

Davis, D.W., Hamilton, M.A., 2010. OGS geochronology report: Atikokan region, unpublished report for the Ontario Geological Survey. Toronto, ON.

Davis, D.W., Poulsen, K.H., Kamo, S.L., 1989. New Insights into Archean Crustal Development from Geochronology in the Rainy Lake Area, Superior Province, Canada. J. Geol. 97, 379–398. https://doi.org/10.1086/629318

Davis, J.C., 2002. Statistics and data analysis in geology, 3rd ed. Wiley, New York. ISBN: 978-0-471-17275-8

De Souza, S., Dubé, B., Mercier-Langevin, P., McNicoll, V., Dupuis, C., Kjarsgaard, I.M., 2019. Hydrothermal alteration mineralogy and geochemistry of the Archean world-class Canadian Malartic disseminated-stockwork gold deposit, southern Abitibi greenstone belt, Quebec, Canada. Econ. Geol. 114, 1057–1094. https://doi.org/10.5382/econgeo.4674

De Souza, S., Perrouty, S., Dubé, B., Mercier-Langevin, P., Linnen, R.L., Olivo, G.R., 2020. Metallogeny of the Neoarchean Malartic Gold Camp, Québec, Canada, in: Sillitoe, R.H., Goldfarb, R.J., Robert, F., Simmons, S.F. (Eds.), Geology of the World's Major Gold Deposits and Provinces. Society of Economic Geologists Special Publication 23, p. 29-52. https://doi.org/10.5382/SP.23.02

Debon, F., Le Fort, P., 1983. A chemical–mineralogical classification of common plutonic rocks and associations. Trans. R. Soc. Edinb. Earth Sci. 73, 135–149. https://doi.org/10.1017/S0263593300010117

Dostal, J., Mueller, W.U., Murphy, J.B., 2004. Archean Molasse Basin Evolution and Magmatism, Wabigoon Subprovince , Canada. J. Geol. 112, 435–454. https://doi.org/https://doi.org/10.1086/421073

Drake, M.J., Weill, D.F., 1975. Partition of Sr, Ba, Ca, Y, Eu2+, Eu3+, and other REE between plagioclase feldspar and magmatic liquid: an experimental study. Geochim. Cosmochim. Acta 39, 689–712. https://doi.org/10.1016/0016-7037(75)90011-3

Dubé, B., Gosselin, P., Hannington, M., Galley, A.G., 2007. Gold-rich valcanogenic massive sulphide deposits, in: Goodfellow, W.D. (Ed.), Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration Methods, Special Publication No. 5. Geological Association of Canada, Mineral Deposits Division, pp. 75–94. https://gac.ca/product/mdd-sp-5-mineral-deposits-of-canada-a-synthesis-of-major-deposit-types-district-metallogeny-the-evolution-of-geological-provinces-exploration-methods/

Dubé, B., Gosselin, P., 2007. Greenstone-Hosted Quartz-Carbonate Vein Deposits, in: Goodfellow, W.D. (Ed.), Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration Methods, Special Publication No. 5. Geological Association of Canada, Mineral Deposits Division, pp. 49–73. https://gac.ca/product/mdd-sp-5-mineral-deposits-of-canada-a-

synthesis-of-major-deposit-types-district-metallogeny-the-evolution-of-geological-provinces-exploration-methods/

Dubé, B., Mercier-Langevin, P., 2020. Gold deposits of the Archean Abitibi greenstone belt, Canada. Rev. Econ. Geol. 23, 669–708. https://doi.org/10.5382/SP.23.32

Dubé, B., Mercier-Langevin, P., Ayer, J.A., Pilote, J.-L., Monecke, T., 2020. Gold deposits of the world-class Timmins-Porcupine Camp, Abitibi greenstone belt, Canada. Rev. Econ. Geol. 23, 53–80. https://doi.org/10.5382/SP.23.03

Duguet, M., Beakhouse, G.P., 2012. Metamorphism in the western Wabigoon Subprovince: Insights from the Bending Lake area, in: Beakhouse, G.P., Dyer, R.D., Easton, R.M., Burnham, O.M., Berger, B.R., Bajc, A.F., Préfontaine, S., Hamilton, S.M., Brunton, F.R., Parker, J.R., Debicki, E.J. (Eds.), Summary of Field Work and Other Activities 2012. Ontario Geological Survey, Sudbury, ON, Open File Report (OFR 6280) pp. 8–1 to 8–12. http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=OFR6280

[Dataset] Heaman, L.M., and Easton, R.M., 2006. Preliminary U/Pb Geochronology Results: Lake Nipigon Region Geoscience Initiative. Ontario Geological Survey Miscellaneous Release Data (MRD 191) 86p. http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=MRD191

Holden, E.-J., Kovesi, P., Wedge, D., Aitken, A.R.A., 2013. Effective methods to highlight and delineate anomalies from geophysical images, in: ASEG Extended Abstracts. pp. 1–4. https://doi.org/10.1071/ASEG2013AB147 .

Eaton, D., Vasudevan, K., 2004. Skeletonization of aeromagnetic data. Geophysics 69, 478–488. https://doi.org/10.1190/1.1707068

[Dataset] Enkin, R.J., Cowan, D., Tigner, J., Severide, A., Gilmour, D., Tkachyk, A., Kilduff, M., Baker, J., 2012. Physical property measurements at the GSC paleomagnetism and petrophysics laboratory, including Electric Impedance Spectrum methodology and analysis, Geological Survey of Canada Open File (OF7227). https://doi.org/10.4095/291564

[Dataset] Ernst, R.E., Buchan, K.L., 2010. Geochemical database of Proterozic intraplate mafic magmatism in Canada, Geological Survey of Canada Open File (OF 6016). Ottawa, Canada. https://doi.org/10.4095/261831

Feltrin, L., Motta, J.G., Al-Obeidat, F., Marir, F., Bertelli, M., 2016. Combining Weights of Evidence Analysis with Feature Extraction - A Case Study from the Hauraki Goldfield, New Zealand. Procedia Comput. Sci. 83, 1262–1267. https://doi.org/10.1016/j.procs.2016.04.262

Feltrin, L., Motta, J.G., Al-Obeidat, F., Marir, F., Bertelli, M., 2016. Combining Weights of Evidence Analysis with Feature Extraction - A Case Study from the Hauraki Goldfield, New Zealand. Procedia Comput. Sci. 83, 1262–1267. https://doi.org/10.1016/j.procs.2016.04.262

Feumoe, A.N.S., Ndougsa-Mbarga, T., Manguelle-Dicoum, E., Fairhead, J.D., 2012. Delineation of tectonic lineaments using aeromagnetic data for the south-east Cameroon area. Geofizika 29, 175–192.

Fisher, N.I., 1995. Statistical Analysis of Circular Data. Cambridge University Press, Cambridge, Great Britain. ISBN: 978-0-521-56890-6

Fisher, R.A., 1956. Statistical Methods and Scientific Inference, 1st ed. Hafner Publishing Co. ISBN: 978-1-013-56097-2

Fleet, M.E., Pan, Y., 1995. Crystal chemistry of rare earth elements in fluorapatite and some calc-silicates. Eur. J. Mineral. 7, 591–605. https://doi.org/10.1127/ejm/7/3/0591

Ford, A., 2019. Practical Implementation of Random Forest-Based Mineral Potential Mapping for Porphyry Cu – Au Mineralization in the Eastern Lachlan Orogen, NSW, Australia. Nat. Resour. Res. https://doi.org/10.1007/s11053-019-09598-y

Ford, A., McCuaig, T.C., 2010. The effect of map scale on geological complexity for computer-aided exploration targeting. Ore Geol. Rev. 38, 156–167. https://doi.org/10.1016/j.oregeorev.2010.03.008

Franklin, J.M., Gibson, H.L., Jonasson, I.R., Galley, A.G., 2005. Volcanogenic massive sulfide deposits, in: Hedenquist, J.W., Thompson, J.F.H., Goldfarb, R.J., Richards, J.P. (Eds.), Economic Geology 100th Anniversary Volume 1905-2005. Society of Economic Geologists, Littleton, Colorado, pp. 523–560. https://www.segweb.org/store_info/AV/AV100/AV100-Additional-Product-Info.pdf

Franklin, J.M., Lydon, J.W., Sangster, D.F., 1981. Volcanic-Associated Massive Sulfide Deposits. Seventy-Fifth Anniv. Vol. https://doi.org/10.5382/AV75.15

Frieman, B.M., Kuiper, Y.D., Kelly, N.M., Monecke, T., Kylander-Clark, A., 2017. Constraints on the geodynamic evolution of the southern Superior Province: U-Pb LA-ICP-MS analysis of detrital zircon in successor basins of the Archean Abitibi and Pontiac subprovinces of Ontario and Quebec, Canada. Precambrian Res. 292, 398–416. https://doi.org/10.1016/j.precamres.2017.01.027

Galley, A.G., 2003. Composite synvolcanic intrusions associated with Precambrian VMS-related hydrothermal systems. Miner. Depos. 38, 443–473. https://doi.org/10.1007/s00126-002-0300-9

Galley, A.G., Hannington, M.D., Jonasson, I.R., 2007. Volcanogenic massive sulphide deposits, in: Goodfellow, W.D. (Ed.), Mineral Deposits of Canada: A Synthesis of Major Deposit-Types, District Metallogeny, the Evolution of Geological Provinces, and Exploration Methods. Geological Association of Canada, Mineral Deposits Division, pp. 141–161. https://gac.ca/product/mdd-sp-5-mineral-deposits-of-canada-a-synthesis-of-major-deposit-types-district-metallogeny-the-evolution-of-geological-provinces-exploration-methods/

Galley, A.G., van Breemen, O., Franklin, J.M., 2000. The relationship between intrusion-hosted Cu-Mo mineralization and VMS deposits of the Archean Sturgeon Lake mining camp, northwestern Ontario. Econ. Geol. 95, 1543–1550. https://doi.org/10.2113/gsecongeo.95.7.1543

Ghaffari, H.O., Nasseri, M.H.B., Young, R.P., 2011. Fluid Flow Complexity in Fracture Networks: Analysis with Graph Theory and LBM. https://doi.org/10.48550/arxiv.1107.4918

Gibson, H.L., Allen, R.L., Riverin, G., Lane, T.E., 2007. The VMS Mode : Advances and Application to Exploration Targeting, in: Fifth Decennial International Conference on Mineral Exploration - Ore Deposits and Exploration Technology. Toronto, ON, pp. 713-730 pp. http://www.dmec.ca/ex07-dvd/E07/pdfs/49.pdf

Gini, C., 1912. Variabilità e Mutuabilità. Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche. C. Cuppini, Bologna.

Goldfarb, R.J., Baker, T., Dubé, B., Groves, D.I., Hart, C.J.R., Gosselin, P., 2005. Distribution, character, and genesis of gold deposits in metamorphic terranes, in: Economic Geology 100th Anniversary Volume 1905-2005. pp. 407–450. https://www.segweb.org/store_info/AV/AV100/AV100-Additional-Product-Info.pdf

Goldfarb, R.J., Groves, D.I., Gardoll, S., 2001. Orogenic gold and geologic time: a global synthesis. Ore Geol. Rev. 18, 1–75. https://doi.org/10.1016/S0169-1368(01)00016-6

Granek, J., 2016. Application of Machine Learning Algorithms to Mineral Prospectivity Mapping. University of British Columbia. https://doi.org/10.14288/1.0340340

Greenacre, M., Grunsky, E.C., Bacon-Shone, J., 2021. A comparison of isometric and amalgamation logratio balances in compositional data analysis. Comput. Geosci. 148, 104621. https://doi.org/https://doi.org/10.1016/j.cageo.2020.104621

Gresens, R.L., 1967. Composition-volume relationships of metasomatism. Chem. Geol. 2, 47–65. https://doi.org/10.1016/0009-2541(67)90004-6

Groves, D.I., Santosh, M., 2015. Province-scale commonalities of some world-class gold deposits: Implications for mineral exploration. Geosci. Front. 6, 389–399. https://doi.org/10.1016/j.gsf.2014.12.007

Groves, D.I., Santosh, M., Goldfarb, R.J., Zhang, L., 2018. Structural geometry of orogenic gold deposits: Implications for exploration of world-class and giant deposits. Geosci. Front. 9, 1163–1177. https://doi.org/10.1016/j.gsf.2018.01.006

Grunsky, E.C., 2010. The interpretation of geochemical survey data. Geochemistry Explor. Environ. Anal. 10, 27–74. https://doi.org/10.1144/1467-7873/09-210

Grunsky, E.C., de Caritat, P., 2019. State-of-the-art analysis of geochemical data for mineral exploration. Geochemistry Explor. Environ. Anal. 20, 217–232. https://doi.org/10.1144/geochem2019-031

Gunn, P.J., Maidment, D., Milligan, P.R., 1997. Interpreting aeromagnetic data in areas of limited outcrop. AGSO J. Aust. Geol. Geophys. 17, 175–185. https://doi.org/10.1071/EG995227

Hagberg, A., Swart, P., S Chult, D., 2008. Exploring network structure, dynamics, and function using networkx, in: Varoquaux, G., Vaught, T., Millman, J. (Eds.), 7th Python in Science Conference (SciPy2008). Pasadena, CA USA, pp. 11–15. https://www.osti.gov/biblio/960616

Hart, T.R., Gibson, H.L., Lesher, C.M., 2004. Trace element geochemistry and petrogenesis of felsic volcanic rocks associated with volcanogenic massive Cu-Zn-Pb sulfide deposits. Econ. Geol. 99, 1003–1013. https://doi.org/10.2113/gsecongeo.99.5.1003

[Map] Hathway, B., 2005. Precambrian geology, parts of Godfrey, Robb, Jamieson, Loveland, Macdiarmid and Thorburn townships. Ontario Geological Survey Preliminary Map (P 3556). Scale 1:20 000. http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=P3556

[Dataset] Haus, M. and Pauk, T., 1993, Data from the PETROCH Lithogeochemical Database. Ontario Geological Survey Miscellaneous Release Data (MRD 250) http://www.geologyontario.mndmf.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=MRD250

Henley, R.W., Norris, R.J., Paterson, C.J., 1976. Multistage ore genesis in the New Zealand geosyncline a history of post-metamorphic lode emplacement. Miner. Depos. 11, 180–196. https://doi.org/10.1007/BF00204480

[Dataset] Hillary, E.M. and Grunsky, E.G., 2010. Compilation of reanalyses and new analyses of lithogeochemistry - Abitibi greenstone belt, Ontario and Quebec. Geological Survey of Canada Open File (OF 6623) https://doi.org/10.4095/287144

[Dataset] Hillary, E.M., Grunsky, E.C., Adcock, S.W., 2008. Compilation of lithogeochemistry - Abitibi greenstone belt - Ontario portion. Geological Survey of Canada Open File (OF 5510) https://doi.org/10.4095/224833

Hannington, M.D., de Ronde, C.E.J., Peterson, S., 2005. Sea-Floor Tectonics and Submarine Hydrothermal Systems, in: Hedenquist, J.W., Thompson, J.F.H., Goldfarb, R.J., Richards, J.P. (Eds.), 100th Anniversary Volume (1905-2005). Society of Economic Geologists (SEG), Littleton, Colorado, pp. 111–141. https://www.segweb.org/Store/detail.aspx?id=EDOC100EEP05

Holden, E.-J., Kovesi, P., Wedge, D., Aitken, A.R.A., 2013. Effective methods to highlight and delineate anomalies from geophysical images, in: ASEG Extended Abstracts. pp. 1–4. https://doi.org/10.1071/ASEG2013AB147

Hood, S.B., Cracknell, M.J., Gazley, M.F., Reading, A.M., 2019. Element mobility and spatial zonation associated with the Archean Hamlet orogenic Au deposit, Western Australia: Implications for fluid pathways in shear zones. Chem. Geol. 514, 10–26. https://doi.org/10.1016/j.chemgeo.2019.03.022

Hotelling, H., 1933. Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. 24, 417–441. https://doi.org/10.1037/h0071325

Houlé, M.G., Gibson, H.L., Lesher, C.M., Davis, P.C., Cas, R.A.F., Beresford, S.W., Arndt, N.T., 2008. Komatiitic Sills and Multigenerational Peperite at Dundonald Beach, Abitibi Greenstone Belt, Ontario: Volcanic Architecture and Nickel Sulfide Distribution. Econ. Geol. 103, 1269–1284. https://doi.org/10.2113/gsecongeo.103.6.1269

[Map] Houlé, M.G. and Hall, L.A.F., 2007. Geological compilation of the Shaw Dome area, northeastern Ontario. Ontario Geological Survey Preliminary Map (P 3595). Scale 1:50 000. http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=P3595

Houlé, M.G., Lesher, C.M., 2011. Komatiite-Associated Ni-Cu-(PGE) Deposits, Abitibi Greenstone Belt, Superior Province, Canada. Magmat. Ni-Cu PGE Depos. Geol. Geochemistry, Genes. https://doi.org/10.5382/Rev.17.03

Hronsky, J.M.A., 2020. Deposit-scale structural controls on orogenic gold deposits: an integrated, physical process–based hypothesis and practical targeting implications. Miner. Depos. 55, 197–216. https://doi.org/10.1007/s00126-019-00918-z

Hronsky, J.M.A., Kreuzer, O.P., 2019. Applying spatial prospectivity mapping to exploration targeting: Fundamental practical issues and suggested solutions for the future. Ore Geol. Rev. 107, 647–653. https://doi.org/10.1016/j.oregeorev.2019.03.016

Hyslop, N.P., White, W.H., 2009. Estimating Precision Using Duplicate Measurements. J. Air Waste Manage. Assoc. 59, 1032–1039. https://doi.org/10.3155/1047-3289.59.9.1032

Irvine, T.N., Baragar, W.R.A., 1971. A Guide to the Chemical Classification of the Common Volcanic Rocks. Can. J. Earth Sci. 8, 523–548. https://doi.org/10.1139/e71-055

Ishikawa, Y., Sawaguchi, T., Iwaya, S., Horiuchi, M., 1976. elineation of Prospecting Targets for Kuroko Deposits Based on Modes of Volcanism of Underlying Dacite and Alteration Halos. Min. Geol. 26, 105–117.

Isles, D.J., Rankin, L.R., 2013. Geological Interpretation of Aeromagnetic Data, Science. CSIRO, Collingwood, VIC. https://doi.org/10.1126/science.147.3661.1052

Iwamori, H., Yoshida, K., Nakamura, H., Kuwatani, T., Hamada, M., Haraguchi, S., Ueki, K., 2017. Classification of geochemical data based on multivariate statistical analyses: Complementary roles of cluster, principal component, and independent component analyses: CLASSIFICATION OF GEOCHEMICAL DATA. Geochemistry, Geophys. geosystems G3 18, 994–1012. https://doi.org/10.1002/2016GC006663

Jensen, L.S., 1976. A New Cation Plot for Classifying Subalkalic Volcanic Rocks, Ministry of Natural Resources Miscellaneous Paper (MP66). http://www.geologyontario.mndmf.gov.on.ca/mndmfiles/pub/data/imaging/MP066/MP066.pdf

Joly, A., Porwal, A., McCuaig, T.C., 2012. Exploration targeting for orogenic gold deposits in the Granites-Tanami Orogen: Mineral system analysis, targeting model and prospectivity analysis. Ore Geol. Rev. 48, 349–383. https://doi.org/10.1016/j.oregeorev.2012.05.004

Joly, A., Porwal, A., McCuaig, T.C., Chudasama, B., Dentith, M.C., Aitken, A.R.A., 2015. Mineral systems approach applied to GIS-based 2D-prospectivity modelling of geological regions: Insights from Western Australia. Ore Geol. Rev. 71, 673–702. https://doi.org/10.1016/j.oregeorev.2015.06.007

Karpatne, A., Ebert-Uphoff, I., Ravela, S., Babaie, H.A., Kumar, V., 2019. Machine Learning for the Geosciences: Challenges and Opportunities. IEEE Trans. Knowl. Data Eng. 31, 1544–1554. https://doi.org/10.1109/TKDE.2018.2861006

Katz, L.R., Kontak, D.J., Dubé, B., McNicoll, V., 2017. The geology, petrology, and geochronology of the Archean Cote gold large-tonnage, low-grade intrusion-related Au(-Cu) deposit, Swayze greenstone belt, Ontario, Canada. Can. J. Earth Sci. 54, 173–202. https://doi.org/10.1139/cjes-2016-0007

Knox-Robinson, C.M., Wyborn, L.A.I., 1997. Towards a holistic exploration strategy: Using Geographic Information Systems as a tool to enhance exploration. Aust. J. Earth Sci. 44, 453–463. https://doi.org/10.1080/08120099708728326

Knuth, D.E., 1997. The art of computer programming, 3rd ed. ed. Addison-Wesley, Reading, Mass. ISBN: 0201896834

Kolmogorov A.N., 1933. Sulla determinazione empirica di una legge di distribuzione. Giornale dell' Instituto Italiano degli Attuari 4:83-91

Kovesi, P., 2012. Phase Preserving Tone Mapping of Non-Photographic High Dynamic Range Images, in: 2012 International Conference on Digital Image Computing Techniques and Applications (DICTA). pp. 1–8. https://doi.org/10.1109/DICTA.2012.6411698

[Map] Kresz, D.U., 1987. Stratigraphic map of the Kawashegamuk Lake area, with sample locations, District of Kenora, Ontario Geological Survey Preliminary Map (P 3100). Scale 1:31 680. http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=P3100

[Map] Kresz, D. U., Blackburn, C. E., & Fraser, F. B., 1997. Precambrian geology of the Kawashegamuk Lake area, eastern part, Kenora District. Ontario Geological Survey Preliminary Map (P.2570). Scale 1:15 840. http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=P2570

[Map] Kresz, D. U., Blackburn, C. E., & Fraser, F. B., 1997. Precambrian geology of the Kawashegamuk Lake area, western part, Kenora District. Ontario Geological Survey Preliminary Map (P.2569). Scale 1:15 840. http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=P2569

Kreuzer, O.P., Miller, A.V.M., Peters, K.J., Payne, C.E., Wildman, C., Partington, G.A., Puccioni, E., Mcmahon, M.E., Etheridge, M.A., 2015. Comparing prospectivity modelling results and past exploration data : A case study of porphyry Cu – Au mineral systems in the Macquarie Arc , Lachlan Fold Belt , New South Wales. Ore Geol. Rev. 71, 516–544. https://doi.org/10.1016/j.oregeorev.2014.09.001

Krivoruchko, K., 2012. Empirical Bayesian kriging; implemented in ArcGIS geostatistical analyst. ARCuser (Redlands, Calif.) 15, 6–10. https://pro.arcgis.com/en/pro-app/latest/help/analysis/geostatistical-analyst/what-is-empirical-bayesian-kriging-.htm

Kuhn, S., Cracknell, M.J., Reading, A.M., Sykora, S., 2020. Identification of intrusive lithologies in volcanic terrains in British Columbia by machine learning using random forests: The value of using a soft classifier. Geophysics 85, B249–B258. https://doi.org/10.1190/geo2019-0461.1

Kuno, H., 1968. Differentiation of basalt magmas, in: Hess, H.H., Poldervaart, A. (Eds.), Basalts. Wiley Interscience, New York, NY, pp. 623–688.

Large, R.R., Gemmell, J.B., Paulick, H., 2001. The alternation box plot: A simple approach to understanding the relationship between alteration mineralogy and lithogeochemistry associated with volcanic-hosted massive sulfide deposits. Econ. Geol. 96, 957–971. https://doi.org/10.2113/gsecongeo.96.5.957

Le Bas, M.J., 2000. IUGS Reclassification of the High-Mg and Picritic Volcanic Rocks. J. Petrol. 41, 1467–1470. https://doi.org/10.1093/petrology/41.10.1467

Le Bas, M.J., Le Maitre, R.W., Streckeisen, A., Zanettin, B.A., 1986. Chemical classification of volcanic rocks based on the total alkali-silica diagram. J. Petrol. 27, 745–750. https://doi.org/10.1093/petrology/27.3.745

Le Maitre, R.W., 1989. A Classification of igneous rocks and glossary of terms. Blackwell Scientific, Oxford.

Le Vaillant, M., Fiorentini, M.L., Barnes, S.J., 2016. Review of lithogeochemical exploration tools for komatiite-hosted Ni-Cu-(PGE) deposits. J. geochemical Explor. 168, 1–19. https://doi.org/10.1016/j.gexplo.2016.05.010

Lesher, C.M., 2019. Up, down, or sideways: emplacement of magmatic Fe-Ni-Cu-PGE sulfide melts in large igneous provinces. Can. J. Earth Sci. 56, 756–773. https://doi.org/10.1139/cjes-2018-0177

Lesher, C.M., 2017. Roles of xenomelts, xenoliths, xenocrysts, xenovolatiles, residues, and skarns in the genesis, transport, and localization of magmatic Fe-Ni-Cu-PGE sulfides and chromite. Ore Geol. Rev. 90, 465–484. https://doi.org/https://doi.org/10.1016/j.oregeorev.2017.08.008

Lesher, C.M., 1989. Komatiite-Associated Nickel Sulfide Deposits. Ore Depos. Assoc. with Magmas. https://doi.org/10.5382/Rev.04.05

Lesher, C.M., Burnham, O.M., Keays, R.R., Barnes, S.J., Hulbert, L., 2001. Trace-element geochemistry and petrogenesis of barren and ore-associated komatiites. Can. Mineral. 39, 673–696. https://doi.org/10.2113/gscanmin.39.2.673

Lesher, C.M., Goodwin, A.M., Campbell, I.H., Gorton, M.P., 1986. Trace-element geochemistry of ore-associated and barren, felsic metavolcanic rocks in the Superior Province, Canada. Can. J. Earth Sci. 23, 222–237. https://doi.org/10.1139/e86-025

Liang, Y., Meng, X., Shujiang, A., 2011. Canny edge detection method and it's application. Appl. Mech. Mater. 50–51, 483–487. https://doi.org/10.4028/www.scientific.net/AMM.50-51.483

Lin, N., Chen, Y., Liu, Haiqi, Liu, Hanlin, 2021. A Comparative Study of Machine Learning Models with Hyperparameter Optimization Algorithm for Mapping Mineral Prospectivity. Minerals 11, 159. https://doi.org/https://doi.org/10.3390/min11020159

Lindsay, M.D., Aitken, A.R.A., Ford, A., Dentith, M.C., Hollis, J., Tyler, I., 2016. Reducing subjectivity in multi-commodity mineral prospectivity analyses: Modelling the west Kimberley, Australia. Ore Geol. Rev. 76, 395–413. https://doi.org/10.1016/j.oregeorev.2015.03.022

Lindsay, M.D., Piechocka, A.M., Jessell, M.W., Scalzo, R., Giraud, J., Pirot, G., Cripps, E., 2022. Assessing the impact of conceptual mineral systems uncertainty on prospectivity predictions. Geosci. Front. 13, 101435. https://doi.org/https://doi.org/10.1016/j.gsf.2022.101435

Lister, J.R., Kerr, R.C., 1991. Fluid-mechanical models of crack propagation and their application to magma transport in dykes. J. Geophys. Res. Solid Earth Planets 96, 10049–10077. https://doi.org/10.1029/91JB00600

Loughlin, W.P., 1991. Principal Component Analysis for Alteration Mapping. Photogramm. Eng. Remote Sensing 57, 1163–1169.

Liu, L., Cao, W., Liu, H., Ord, A., Qin, Y., Zhou, F., Bi, C., 2022. Applying benefits and avoiding pitfalls of 3D computational modeling-based machine learning prediction for exploration targeting: Lessons from two mines in the Tongling-Anqing district, eastern China. Ore Geol. Rev. 142, 104712. https://doi.org/https://doi.org/10.1016/j.oregeorev.2022.104712

Lu, B., Charlton, M., Harris, P., Fotheringham, A.S., 2014. Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data. Int. J. Geogr. Inf. Sci. 28, 660–681. https://doi.org/10.1080/13658816.2013.865739

MacDonald, P.J., Piercey, S.J., 2019. Geology, lithogeochemistry, and significance of porphyry intrusions associated with gold mineralization within the Timmins–Porcupine gold camp, Canada. Can. J. Earth Sci. 56, 399–418. https://doi.org/10.1139/cjes-2018-0091

MacLean, W.H., Barrett, T.J., 1993. Lithogeochemical techniques using immobile elements. J. Geochemical Explor. 48, 109–133. https://doi.org/10.1016/0375-6742(93)90002-4

MacQueen, L., 1967. Some methods for classification and analysis of multivariate observations, in: Le Cam, L.M., Neyman, J. (Eds.), Berkeley Symposium on Mathematical Statistics and Probability. University of California, pp. 281–297.

Maier, W.D., Smithies, R.H., Spaggiari, C. V, Barnes, S.J., Kirkland, C.L., Yang, S., Lahaye, Y., Kiddie, O., MacRae, C., 2016. Petrogenesis and Ni–Cu sulphide potential of mafic–ultramafic rocks in the Mesoproterozoic Fraser Zone within the Albany–Fraser Orogen, Western Australia. Precambrian Res. 281, 27–46. https://doi.org/10.1016/j.precamres.2016.05.004

Mathieu, L., 2018. Quantifying Hydrothermal Alteration: A Review of Methods. Geosciences. https://doi.org/10.3390/geosciences8070245

McNicoll, V., Goutier, J., Dube, B., Mercier-Langevin, P., Ross, P.-S., Dion, C., Monecke, T., Legault, M., Percival, J., Gibson, H., 2014. U-Pb geochronology of the Blake River Group, Abitibi greenstone belt, Quebec, and implications for base metal exploration. Econ. Geol. 109, 27–59. https://doi.org/10.2113/econgeo.109.1.27

[Dataset] Meek, D. M., Mole, D. R., & Fernandes, B. A., 2020. Metal Earth Geochronology Compilation: Superior Craton and surrounding area. Laurentian University Mineral Exploration Research Center Publication MERCME-2020-073. https://merc.laurentian.ca/news-standard/metal-earth-geochronology-compilation-superior-craton-and-surrounding-area

Melling, D.R., Blackburn, C.E., Watkinson, D.H., Parker, J., 1988. Geological setting of gold, western Wabigoon Subprovince, Canadian Shield; exploration targets in mixed volcanic successions. Can. J. Earth Sci. 25, 2075–2088. https://doi.org/10.1139/e88-192

Mercier-Langevin, P., Gibson, H.L., Hannington, M.D., Goutier, J., Monecke, T., Dubé, B., Houlé, M.G., 2014. A special issue on Archean magmatism, volcanism, and ore deposits: Part 2. Volcanogenic massive sulfide deposits: Preface. Econ. Geol. 109, 1–9. https://doi.org/10.2113/econgeo.109.1.1

Metelka, V., Baratoux, L., Naba, S., Jessell, M.W., 2011. A geophysically constrained litho-structural analysis of the Eburnean greenstone belts and associated granitoid domains, Burkina Faso, West Africa. Precambrian Res. 190, 48–69. https://doi.org/10.1016/j.precamres.2011.08.002

Middlemost, E.A.K., 1994. Naming materials in the magma/igneous rock system. Earth Sci. Rev. 37, 215–224. https://doi.org/10.1016/0012-8252(94)90029-9

Middleton, G. V, 2000. Data Analysis in the Earth Sciences Using Matlab. Prentice Hall.

Miles, R.E., 1972. The Random Division of Space. Adv. Appl. Probab. 4, 243–266. https://doi.org/10.2307/1425985

Miller, H.G., Singh, V., 1994. Potential field tilt-a new concept for location of potential field sources. J. Appl. Geophys. 32, 213–217. https://doi.org/10.1016/0926-9851(94)90022-1

[Dataset] Mining Innovation, Rehabilitation and Applied Research Corporation, 2005. Integrated GIS compilation of geospatial data from the Abitibi greenstone belt, northeastern Ontario: Discover Abitibi Initiative. Ontario Geological Survey Miscellaneous Release—Data (MRD 186). http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=MRD186

Monecke, T., Mercier-Langevin, P., Dubé, B., Frieman, B.M., 2017. Geology of the Abitibi Greenstone Belt, in: Reviews in Economic Geology. pp. 7–49. https://doi.org/10.5382/Rev.19.01

[Dataset] Montsion, R.M., Perrouty, S., Lesher, C.M., submitted a. Geochemical data compilation for the western Wabigoon and southern Abitibi subprovinces of the Superior Province, Ontario, Canada.

Montsion, R.M., Perrouty, S., Lesher, C.M., Lindsay, M.D., submitted b. Statistical optimization for 5D classification of Archean igneous rock and application for exploration targeting in the Superior Province, Canada.

[Dataset] Montsion, R.M., Perrouty, S., Frieman, B.M., 2021a. Geological and geophysical data compilation for the western Wabigoon and southern Abitibi subprovinces of the Superior Province, Ontario, Canada. Data Br. 37, 107159. https://doi.org/10.1016/J.DIB.2021.107159

Montsion, R.M., Perrouty, S., Lindsay, M.D., Jessell, M.W., Frieman, B.M., 2021b. Mapping structural complexity using geophysics: A new geostatistical approach applied to greenstone belts of the southern Superior Province, Canada. Tectonophysics 228889. https://doi.org/10.1016/j.tecto.2021.228889

Montsion, R.M, Perrouty, S., Frieman, B.M., 2019. Field Constraints on Geophysical Investigations of a Regional Fold Structure in the Western Wabigoon Subprovince, in Easton, R.M., Préfontaine, S., Bajc, A.F., Hamilton, S.M., Rainsford, D.R.B., Burnham, O.M., Duguet, M., Brunton, F.R., Hechler, J.H., Dyer, R.D. (Ed.), Summary of Field Work and Other Activities 2019. Ontario Geological Survey Open File Report (OFR 6360), Sudbury, ON, pp. 38-1-1 to 38-11

[Map] Montsion, R.M., Thurston, P., Ayer, J., 2018. 1:2 000 000 Scale Geological Compilation of the Superior Craton - Version 1. Mineral Exploration Research Centre, Harquail School of Earth Sciences, Laurentian

University Document Number MERC-ME-2018-017.
https://www.arcgis.com/home/item.html?id=8cf2b0aabee54953b6be3a081ae15bb5

Mortensen, J.K., 1993. U-Pb geochronology of the eastern Abitibi Subprovince. Part 1: Chibougamau-Matagami-Joutel region. Can. J. Earth Sci. 30, 11–28. https://doi.org/10.1139/e93-002

Mortensen, J.K., 1993. U-Pb geochronology of the eastern Abitibi Subprovince. Part 2: Noranda - Kirkland Lake area. Can. J. Earth Sci. 30, 29–41. https://doi.org/10.1139/e93-003

Moyen, J.-F., Martin, H., 2012. Forty years of TTG research. Lithos 148, 312–336. https://doi.org/https://doi.org/10.1016/j.lithos.2012.06.010

Naldrett, A.J., 1999. World-class Ni-Cu-PGE deposits; key factors in their genesis. Miner. Depos. 34, 227–240. https://doi.org/10.1007/s001260050200

Naldrett, A.J., 2011. Fundamentals of Magmatic Sulfide Deposits. Magmat. Ni-Cu PGE Depos. Geol. Geochemistry, Genes. https://doi.org/10.5382/Rev.17

Nesbitt, R.W., Sun, S.S., Purvis, A.C., 1979. Komatiites; geochemistry and genesis. Can. Mineral. 17, 165–186. https://pubs.geoscienceworld.org/canmin/article/17/2/165/11334/Komatiites-geochemistry-and-genesis?searchresult=1

[Dataset] Ontario Geological Survey, 2003. Ontario airborne geophysical surveys, magnetic data, grid data, Geosoft® format, magnetic supergrids. Ontario Geological Survey Geophysical Data Set (GDS 1037b). http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=GDS1037B

[Dataset] Ontario Geological Survey, 2011. Ontario airborne geophysical surveys, magnetic and electromagnetic data, grid and vector data, Geosoft® format, Stormy Lake area. Ontario Geological Survey Geophysical Data Set (GDS 1107b—Revision 1). http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=GDS1107B-REV1

Ontario Geological Survey, 2018. Mineral Deposit Category Definitions, https://www.mndm.gov.on.ca/en/mines-and-minerals/applications/ogsearth/mineral-deposits-mdi.

[Dataset] Ontario Geological Survey, 2019. Mineral Deposit Inventory. Ontario Geological Survey, https://www.mndm.gov.on.ca/en/mines-and-minerals/applications/ogsearth/mineral-deposits-mdi

Ordóñez-Calderón, J.C., Gelcich, S., 2018. Machine learning strategies for classification and prediction of alteration facies: Examples from the Rosemont Cu-Mo-Ag skarn deposit, SE Tucson Arizona. J. Geochemical Explor. 194, 167–188. https://doi.org/10.1016/j.gexplo.2018.07.020

Otamendi, J.E., Tiepolo, M., Walker, B.A., Cristofolini, E.A., Tibaldi, A.M., 2016. Trace elements in minerals from mafic and ultramafic cumulates of the central Sierra de Valle Fertil, Famatinian Arc, Argentina. Lithos 240–243, 355–370. https://doi.org/10.1016/j.lithos.2015.11.009

Parker, J.R., 1989. Geology, Gold Mineralization and Property Visits in the Area Investigated by the Dryden-Ignace Economic Geologist. Ontario Geological Survey Open File Report (OFR 5723) pp. 339. http://www.geologyontario.mndm.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=OFR5723

Parsons, A.J., Coleman, M.J., Ryan, J.J., Zagorevski, A., Joyce, N.L., Gibson, H.D., Larson, K.P., 2018. Structural evolution of a crustal-scale shear zone through a decreasing temperature regime: The Yukon River shear zone, Yukon-Tanana terrane, Northern Cordillera. Lithosphere 10, 760–782. https://doi.org/10.1130/L724.1

Parzen, E., 1962. On Estimation of a Probability Density Function and Mode. Ann. Math. Stat. 33, 1065–1076. https://doi.org/10.1214/aoms/1177704472

PCI Geomatics Enterprises Inc., 2018. Geomatica Banff Edition, Document Number: 007-013339-001, Revision D, https://www.pcigeomatics.com/geomatica-help/references/pciFunction_r/python/P_line.html

PCI Geomatics Enterprises Inc., 2019. Geomatica Banff Edition: Focus, Software for object -based image analysis, https://www.pcigeomatics.com/software/geomatica/professional

Pearce, J.A., 2008. Geochemical fingerprinting of oceanic basalts with applications to ophiolite classification and the search for Archean oceanic crust. Lithos 100, 14–48. https://doi.org/10.1016/j.lithos.2007.06.016

Pearce, J.A., 1996. A user's guide to basalt discrimination diagrams, Trace Element Geochemistry of Volcanic Rocks: Applications for Massive Sulphide Exploration. Geological Association of Canada.

Pearson, K., 1901. LIII. On lines and planes of closest fit to systems of points in space. London, Edinburgh, Dublin Philos. Mag. J. Sci. 2, 559–572. https://doi.org/10.1080/14786440109462720

Percival, J.A., McNicoll, V., Brown, J.L., Whalen, J.B., 2004. Convergent margin tectonics, central Wabigoon subprovince, Superior Province, Canada. Precambrian Res. 132, 213–244. https://doi.org/10.1016/j.precamres.2003.12.016

Percival, J.A., Skulski, T., Sanborn-barrie, M., Stott, G.M., Leclair, A.D., Corkery, M.T., Boily, M., 2012. Geology and tectonic evolution of the Superior Province , Canada, in: Percival, J.A., Cook, F.A., Clowes, R.M. (Eds.), Tectonic Styles in Canada: The LITHOPROBE Perspective. Geological Association of Canada, pp. 321–378. https://gac.ca/product/tectonic-styles-in-canada/

Percival, J.A., Helmstaedt, H., 2004. Insights on Archean continent - Ocean assembly, western Superior Province, from new structural, geochemical and geochronological observations: Introduction and summary. Precambrian Res. 132, 209–212. https://doi.org/10.1016/j.precamres.2003.11.006

Perrouty, S., Aillères, L., Jessell, M.W., Baratoux, L., Bourassa, Y., Crawford, B., 2012. Revised Eburnean geodynamic evolution of the gold-rich southern Ashanti Belt, Ghana, with new field and geophysical evidence of pre-Tarkwaian deformations. Precambrian Res. 204–205, 12–39. https://doi.org/10.1016/j.precamres.2012.01.003

Perrouty, S., Gaillard, N., Piette-Lauzière, N., Mir, R., Bardoux, M., Olivo, G.R., Linnen, R.L., Bérubé, C.L., Lypaczewski, P., Guilmette, C., Feltrin, L., Morris, W.A., 2017. Structural setting for Canadian Malartic style of gold mineralization in the Pontiac Subprovince, south of the Cadillac Larder Lake Deformation Zone, Québec, Canada. Ore Geol. Rev. 84, 185–201. https://doi.org/10.1016/j.oregeorev.2017.01.009

Perrouty, S., Linnen, R.L., Lesher, C.M., Olivo, G., Piercey, S.J., Gaillard, N., Clark, J., Enkin, R.J., 2019. Expanding the size of multi-parameter metasomatic footprints in gold exploration: utilization of mafic dykes in the Canadian Malartic district, Québec, Canada. Miner. Depos. 54, 761–786. https://doi.org/10.1007/s00126-018-0829-x

Phillips, G.N., 1993. Metamorphic fluids and gold. Mineral. Mag. 57, 365–374. https://doi.org/10.1180/minmag.1993.057.388.02

Piercey, S.J., 2010. An overview of petrochemistry in the regional exploration for volcanogenic massive sulphide (VMS) deposits. Geochemistry Explor. Environ. Anal. 10, 119–136. https://doi.org/10.1144/1467-7873/09-221

Pitcairn, I.K., Craw, D., Teagle, D.A.H., 2014. The gold conveyor belt: Large-scale gold mobility in an active orogen. Ore Geol. Rev. 62, 129–142. https://doi.org/10.1016/j.oregeorev.2014.03.006

Pitcairn, I.K., Leventis, N., Beaudoin, G., Faure, S., Guilmette, C., Dubé, B., 2021. A metasedimentary source of gold in Archean orogenic gold deposits. Geology 49, 862–866. https://doi.org/10.1130/G48587.1

Poldervaart, A., 1953. Petrological calculations in metasomatic processes. Am. J. Sci. 251, 481–504. https://doi.org/10.2475/ajs.251.7.481

Ragan, D.M., 2009. Structural Geology: An Introduction to Geometrical Techniques. Cambridge University Press, Cambridge. https://doi.org/10.1017/CBO9780511816109

Rigol-Sanchez, J.P., Chica-Olmo, M., Abarca-Hernandez, F., 2003. Artificial neural networks as a tool for mineral potential mapping with GIS. Int. J. Remote Sens. 24, 1151–1156. https://doi.org/10.1080/0143116021000031791

Ripley, E.M., 1986. Application of Stable Isotopic Studies to Problems of Magmatic Sulfide Ore Genesis With Special Reference to the Duluth Complex, Minnesota BT - Geology and Metallogeny of Copper Deposits, in: Friedrich, G.H., Genkin, A.D., Naldrett, A.J., Ridge, J.D., Sillitoe, R.H., Vokes, F.M. (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 25–42. https://link.springer.com/chapter/10.1007/978-3-642-70902-9_3

Rizzo, C.B., de Barros, F.P.J., 2017. Minimum Hydraulic Resistance and Least Resistance Path in Heterogeneous Porous Media: MINIMUM HYDRAULIC RESISTANCE. Water Resour. Res. 53, 8596–8613. https://doi.org/10.1002/2017WR020418

Robert, F., 2001. Syenite-associated disseminated gold deposits in the Abitibi greenstone belt, Canada. Miner. Depos. 36, 503–516. https://doi.org/10.1007/s001260100186

Robert, F., Poulsen, K.H., 2001. Vein Formation and Deformation in Greenstone Gold Deposits, in: Richards, J.P., Tosdal, R.M. (Eds.), Structural Controls on Ore Genesis. Society of Economic Geologists. https://doi.org/10.5382/Rev.14.05

Robert, F., Poulsen, H.K., Cassidy, K.F., Hodgson, J.C., 2005. Gold Metallogeny of the Superior and Yilgarn Cratons. Econ. Geol. 1001–1033. https://doi.org/10.1523/JNEUROSCI.0613-14.2015

Rock, N.M.S., 1977. The nature and origin of lamprophyres: some definitions, distinctions, and derivations. Earth-Science Rev. 13, 123–169. https://doi.org/10.1016/0012-8252(77)90020-4

Rodriguez-Galiano, V., Sanchez-Castillo, M., Chica-Olmo, M., Chica-Rivas, M., 2015. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geol. Rev. 71, 804–818. https://doi.org/10.1016/j.oregeorev.2015.01.001

Rosenblatt, M., 1956. Remarks on Some Nonparametric Estimates of a Density Function. Ann. Math. Stat. 27, 832–837. https://doi.org/10.1214/aoms/1177728190

Salui, C.L., 2018. Methodological Validation for Automated Lineament Extraction by LINE Method in PCI Geomatica and MATLAB based Hough Transformation. J. Geol. Soc. India 92, 321–328. https://doi.org/10.1007/s12594-018-1015-6

[Dataset] Sarbas, B., Nohl, U., 2009. The GEOROC database - a decade of "online geochemistry." Geochim. Cosmochim. Acta 73. http://georoc.mpch-mainz.gwdg.de/georoc/

Saumur, B.M., Cruden, A.R., Boutelier, D., 2015. Sulfide Liquid Entrainment by Silicate Magma: Implications for the Dynamics and Petrogenesis of Magmatic Sulfide Deposits. J. Petrol. 56, 2473–2490. https://doi.org/10.1093/petrology/egv080

Schaeben, H., Kost, S., Semmler, G., 2019. Popular Raster-Based Methods of Prospectivity Modeling and Their Relationships. Math. Geosci. 51, 945–971. https://doi.org/10.1007/s11004-019-09808-6

Scott, D.W., 1992. Multivariate density estimation theory, practice, and visualization , Wiley series in probability and mathematical statistics. Wiley, New York. https://doi.org/10.1002/9780470316849

Seequent, 2019. Oasis Montaj, Software to visualize, analyze, and integrate geophysical and geoscientific data, https://www.seequent.com/products-solutions/geosoft-oasis-montaj/

Shannon, R.D., 1976. Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. Acta Crystallogr. Sect. A 32, 751–767. https://doi.org/10.1107/S0567739476001551

Sibson, R.H., 1977. Fault rocks and fault mechanisms. J. Geol. Soc. London. 133, 191–213. https://doi.org/10.1144/gsjgs.133.3.0191

Sibson, R.H., Robert, F., Poulsen, K.H., 1988. and Mesothermal Gold-Quartz Deposits. Geology 16, 551–555. https://doi.org/https://doi.org/10.1130/0091-7613(1988)016<0551:HARFFP>2.3.CO;2

Silverman, B.W., 1986. Density estimation for statistics and data analysis , Monographs on statistics and applied probability. Chapman and Hall, London.

Sleep, N.H., Ebinger, C.J., Kendall, J.-M., 2002. Deflection of mantle plume material by cratonic keels. Early Earth Phys. Chem. Biol. Dev. https://doi.org/10.1144/GSL.SP.2002.199.01.08

Smirnov, N., 1948. Table for Estimating the Goodness of Fit of Empirical Distributions. Ann. Math. Stat. 19, 279–281. https://doi.org/10.1214/aoms/1177730256

Smithies, R.H., Lowrey, J.R., 2020. Eastern Goldfields greenstone geochemical barcoding project, Record 202. ed. Geological Survey of Western Australia, Perth, Australia. https://nla.gov.au/nla.obj-2965151348

Sproule, R.A., Lesher, C.M., Ayer, J.A., Thurston, P.C., Herzberg, C.T., 2002. Spatial and temporal variations in the geochemistry of komatiites and komatiitic basalts in the Abitibi greenstone belt. Precambrian Res. 115, 153–186. https://doi.org/10.1016/S0301-9268(02)00009-8

[Dataset] Sproule, R.A., Lesher, C.M., Ayer, J.A., and Thurston, P.C., 2003. Compilation of Geochemical Data from Komatiitic Rocks in the Abitibi Greenstone Belt, Ontario. Ontario Geological Survey Miscellaneous Release Data (MRD 120) http://www.geologyontario.mndmf.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=MRD120

Stone, D., Davis, D.W., Hamilton, M.A. and Falcon, A., 2010. Interpretation of 2009 geochronology in the Central Wabigoon Subprovince and Bending Lake areas, northwestern Ontario in Ayer, J.A., Easton, R.M., Beakhouse, G.P., Stott, G.M., Kelly, R.I., Debicki, E.J., Parker, J.R. and Brown, T. (Ed.), Summary of Field Work and Other Activities 2010. Ontario Geological Survey Open File Report (OFR 6260), pp.14-1 to 14-13

[Dataset] Stone, D., 2010. Geochemical analyses of rocks, minerals and soil in the central Wabigoon Subprovince area, northwestern Ontario. Ontario Geological Survey Miscellaneous Release—Data (MRD 242) http://www.geologyontario.mndmf.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=MRD242

Stromberg, J.M., 2017. Geochemistry of the Dome Mine Ankerite Veins: Insights into the Multi-Stage Enrichment of a World-Class Orogenic Gold Deposit. Western University Department of Earth Sciences. Electronic Thesis and Dissertation Repository. https://ir.lib.uwo.ca/etd/5090

Sun, S.S., McDonough, W.F., 1989. Chemical and isotopic systematics of oceanic basalts: Implications for mantle composition and processes. Geol. Soc. Spec. Publ. 42, 313–345. https://doi.org/10.1144/GSL.SP.1989.042.01.19

Sun, T., Li, H., Wu, K., Chen, F., Zhu, Z., Hu, Z., 2020. Data-Driven Predictive Modelling of Mineral Prospectivity Using Machine Learning and Deep Learning Methods: A Case Study from Southern Jiangxi Province, China. Minerals. https://doi.org/10.3390/min10020102

Tang, Y.B., Zhao, J.Z., Bernabe, Y., Li, M., 2021. Fluid flow concentration on preferential paths in heterogeneous porous media; application of graph theory. J. Geophys. Res. Solid earth 126, n/a. https://doi.org/10.1029/2021JB023164

Taylor, S.R., McLennan, S.M., 1985. The continental crust: Its composition and evolution. Blackwell Scientific Pub.,Palo Alto, CA, United States.

Taylor, W.R., Rock, N.M.S., Groves, D.I., Perring, C.S., Golding, S.D., 1994. Geochemistry of Archean shoshonitic lamprophyres from the Yilgarn Block, Western Australia: Au abundance and association with gold mineralization. Appl. Geochemistry 9, 197–222. https://doi.org/10.1016/0883-2927(94)90007-8

Thompson, P.H., 2005. A new metamorphic framework for gold exploration in the Timmins-Kirkland Lake area, western Abitibi greenstone belt (Discover Abitibi Initiative), in: Open File Report. Ontario Geological Survey, Sudbury, ON, p. 104.

Thompson, M., Howarth, R.J., 1973. The rapid estimation and control of precision by duplicate determinations. Analyst 98, 153–160. https://doi.org/10.1039/AN9739800153

Thompson, M., Howarth, R.J., 1976. Duplicate analysis in geochemical practice. Part I. Theoretical approach and estimation of analytical reproducibility. Analyst 101, 690–698. https://doi.org/10.1039/AN9760100690

Thurston, J.B., Smith, R.S., 1997. Automatic conversion of magnetic data to depth, dip, and susceptibility contrast using the SPI (TM) method. Geophysics 62, 807–813. https://doi.org/10.1190/1.1444190

Thurston, P.C., Chivers, K.M., 1990. Secular variation in greenstone sequence development emphasizing Superior Province, Canada. Precambrian Res. 46, 21–58. https://doi.org/10.1016/0301-9268(90)90065-X

Thurston, P.C., Ayer, J.A., Goutier, J., Hamilton, M.A., 2008. Depositional Gaps in Abitibi Greenstone Belt Stratigraphy: A Key to Exploration for Syngenetic Mineralization. Econ. Geol. 103, 1097–1134.

Tomlinson, K.Y., Stott, G.M., Percival, J.A., Stone, D., 2004. Basement terrane correlations and crustal recycling in the western Superior Province: Nd isotopic character of granitoid and felsic volcanic rocks in the Wabigoon subprovince, N. Ontario, Canada. Precambrian Res. 132, 245–274. https://doi.org/10.1016/j.precamres.2003.12.017

Ueki, K., Hino, H., Kuwatani, T., 2018. Geochemical Discrimination and Characteristics of Magmatic Tectonic Settings: A Machine-Learning-Based Approach. Geochemistry, Geophys. geosystems G3 19, 1327–1347. https://doi.org/10.1029/2017GC007401

[Dataset] Vaillancourt, V. and Hall, L., 2003. Lithogeochemical data for the Timmins Wester area: Carscallen, Denton, Briston, Ogden and Deloro townships. Ontario Geological Survey Miscellaneous Release Data (MRD 123) http://www.geologyontario.mndmf.gov.on.ca/mndmaccess/mndm_dir.asp?type=pub&id=MRD123

Vallée, M.A., Morris, W.A., Perrouty, S., Lee, R.G., Wasyliuk, K., King, J.J., Ansdell, K., Mir, R., Shamsipour, P., Farquharson, C.G., Chouteau, M., Enkin, R.J., Smith, R.S., 2019. Geophysical inversion contributions to mineral exploration: lessons from the Footprints project1. Can. J. Earth Sci. 56, 525–543. https://doi.org/10.1139/cjes-2019-0009

Verduzco, B., Fairhead, J.D., Green, C.M., MacKenzie, C., 2004. New insights into magnetic derivatives for structural mapping. Lead. Edge 23, 116–119. https://doi.org/10.1190/1.1651454

Vermeesch, P., 2006. Tectonic discrimination diagrams revisited. Geochemistry, Geophys. geosystems G3 7, Q06017-n/a. https://doi.org/10.1029/2005GC001092

Waskom, M., 2022. Seaborn: Statistical data visualization [WWW Document]. Seaborn. https://seaborn.pydata.org/index.html

Wilke, M., Behrens, H., 1999. The dependence of the partitioning of iron and europium between plagioclase and hydrous tonalitic melt on oxygen fugacity. Contrib. to Mineral. Petrol. 137, 102–114. https://doi.org/10.1007/s004100050585

Williams, M., Schoneveld, L., Mao, Y., Klump, J., Gosses, J., Dalton, H., Bath, A., Barnes, S.J., 2020. pyrolite: Python for geochemistry. J. Open Source Softw. 5, 2314. https://doi.org/10.21105/joss.02314

Wilson, L., W, H.J., 1981. Ascent and eruption of basaltic magma on the earth and moon. J. Geophys. Res. Solid Earth Planets 86, 2971–3001. https://doi.org/10.1029/JB086iB04p02971

Winchester, J.A., Floyd, P.A., 1977. Geochemical discrimination of different magma series and their differentiation products using immobile elements. Chem. Geol. 20, 325–343. https://doi.org/10.1016/0009-2541(77)90057-2

Witt, W.K., 1992. Porphyry intrusions and albitites in the Bardoc-Kalgoorlie area, Western Australia, and their role in Archean epigenetic gold mineralization. Can. J. Earth Sci. 29, 1609–1622. https://doi.org/10.1139/e92-127

Wong, J. C., Holden, E. J., Dentith, M., Kovesi, P., Fu, S.C., 2006. About CET Grid Analysis plugins: Center for Exploration Targeting, 38 pp.

Yousefi, M., Carranza, E.J.M., 2016. Data-Driven Index Overlay and Boolean Logic Mineral Prospectivity Modeling in Greenfields Exploration. Nat. Resour. Res. 25, 3–18. https://doi.org/10.1007/s11053-014-9261-9

Yousefi, M., Nykänen, V., 2016. Data-driven logistic-based weighting of geochemical and geological evidence layers in mineral prospectivity mapping. J. Geochemical Explor. 164, 94–106. https://doi.org/10.1016/j.gexplo.2015.10.008

Yu, L., Porwal, A., Holden, E.-J., Dentith, M.C., 2012. Towards automatic lithological classification from remote sensing data using support vector machines. Comput. Geosci. 45, 229–239. https://doi.org/10.1016/j.cageo.2011.11.019

Zammit, K., Perrouty, S., Frieman, B.M., Marsh, J.H., Holt, K.A., 2022. Structural and geochronological constraints on orogenic gold mineralization in the western Wabigoon subprovince, Canada. Can. J. Earth Sci. 59, 278–299. https://doi.org/10.1139/cjes-2021-0042

Zanettin, B., 1984. Proposed New Chemical Classification of Volcanic Rocks. Episodes 7, 19. https://doi.org/10.18814/epiiugs/1984/v7i4/003

Zekri, H., Cohen, D.R., Mokhtari, A.R., Esmaeili, A., 2019. Geochemical Prospectivity Mapping Through a Feature Extraction–Selection Classification Scheme. Nat. Resour. Res. 28, 849–865. https://doi.org/10.1007/s11053-018-9422-3

Zhang, S.E., Bourdeau, J.E., Nwaila, G.T., Ghorbani, Y., 2022. Advanced geochemical exploration knowledge using machine learning: Prediction of unknown elemental concentrations and operational prioritization of Re-analysis campaigns. Artif. Intell. Geosci. 3, 86–100. https://doi.org/https://doi.org/10.1016/j.aiig.2022.10.003

Zhang, J., Lin, S., Linnen, R.L., Martin, R., 2014. Structural setting of the Young-Davidson syenite-hosted gold deposit in the Western Cadillac-Larder Lake Deformation Zone, Abitibi Greenstone Belt, Superior Province, Ontario. Precambrian Res. 248, 39–59. https://doi.org/10.1016/j.precamres.2014.04.007

Zhang, N., Zhou, K., 2015. Mineral prospectivity mapping with weights of evidence and fuzzy logic methods. J. Intell. Fuzzy Syst. 29, 2639–2651. https://doi.org/10.3233/IFS-151967

Zhang, P., Zhang, Z., Yang, J., Cheng, Q., 2023. Machine Learning Prediction of Ore Deposit Genetic Type Using Magnetite Geochemistry. Nat. Resour. Res. (New York, N.Y.) 32, 99–116. https://doi.org/10.1007/s11053-022-10146-4

Supplementary material (Chapter 5)

Supplementary material 5-1: Supplementary figures
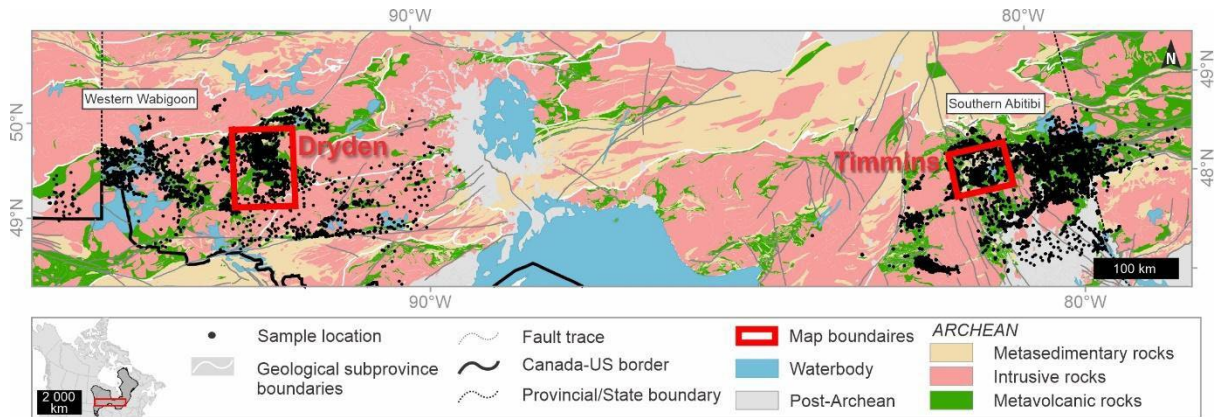
SM Figure 5-1.1. Location of lithogeochemical samples compiled in Montsion et al. (submitted s)/Chapter 4 overlying a simplified geological map of the southern Superior Province in Canada (after Montsion et al., 2018). An inset map of the North American landmass (light grey) and the surface exposure of the Superior Province (dark grey) provide geographic context for the simplified geological map and sample locations. The extent of the Timmins and Dryden areas of interest used to define favorability are indicated in red boxes.
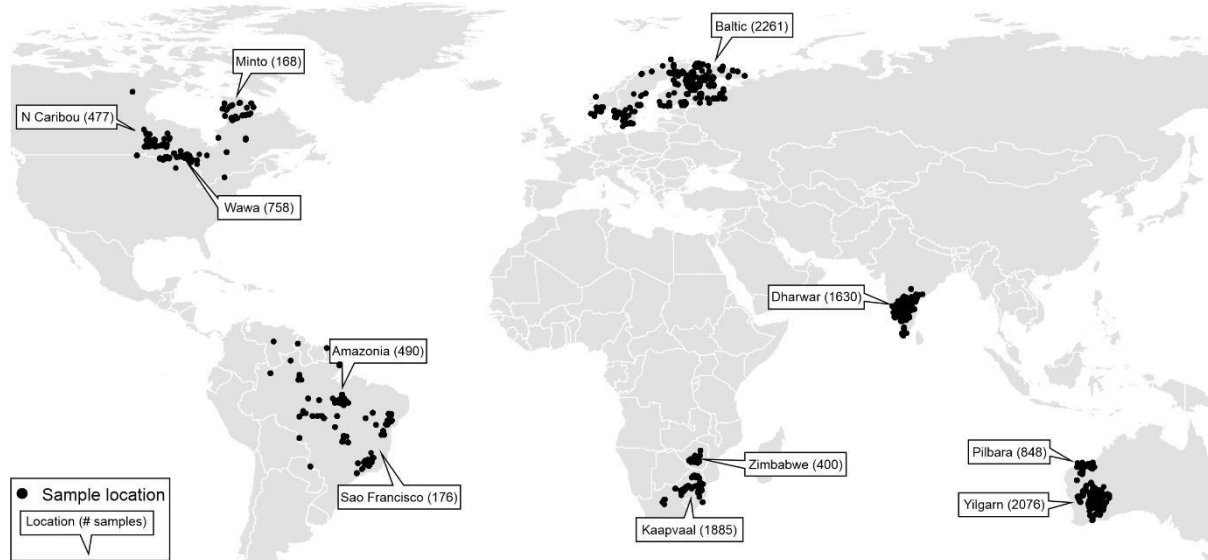
SM Figure 5-1.2. Location of lithogeochemical samples compiled in GEOROC (2021) overlain on a simplified map of Earth's landmass (light grey). The number of available samples from each location/craton is labeled in white boxes.
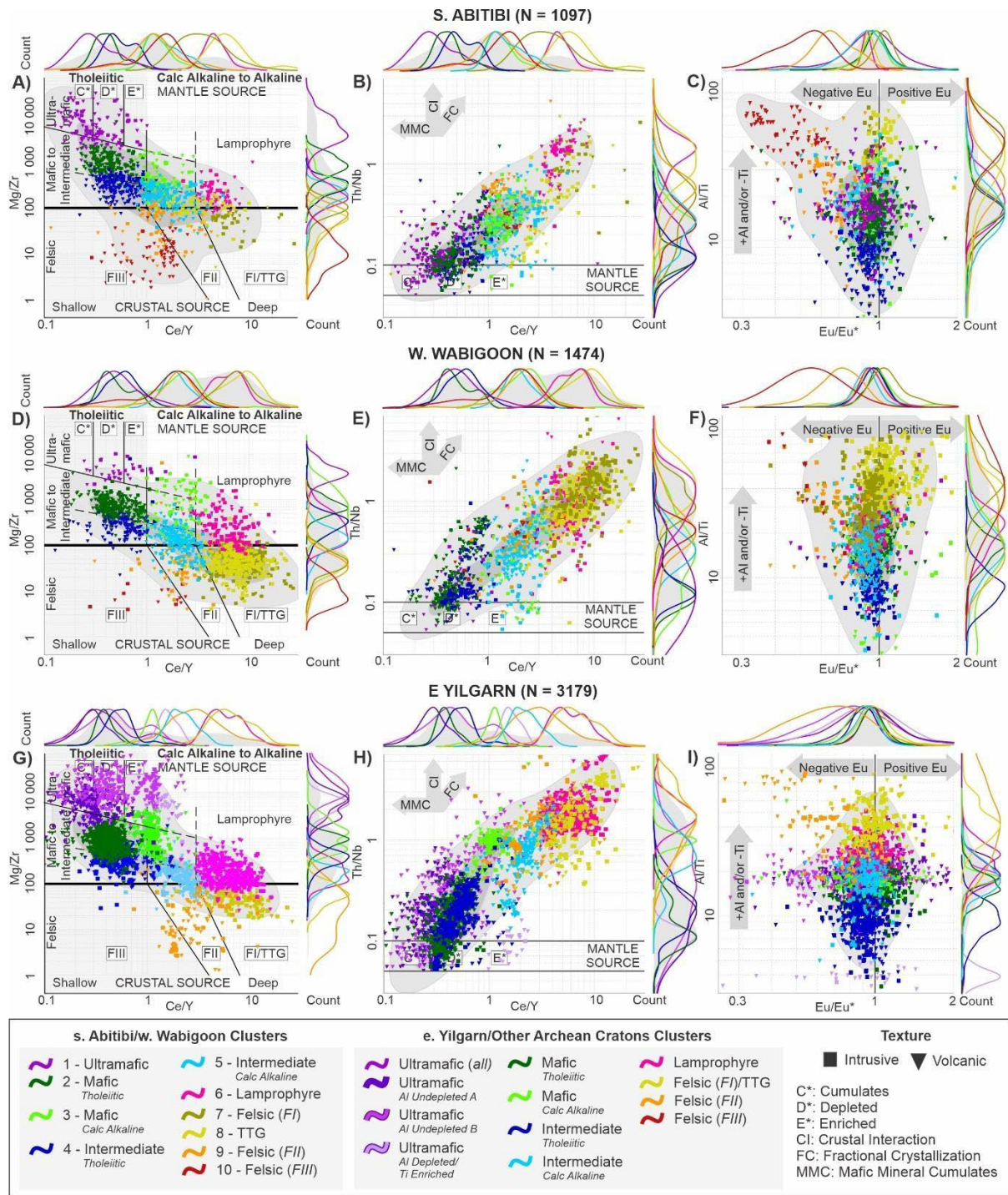
SM Figure 5-1.3. Interpreted compositional clusters plotted on the proposed classification scheme for the southern Abitibi (A, B, C), western Wabigoon (D, E, F), and eastern Yilgarn (G, H, I). A normalized histogram of each compositional group is indicated for each axis.
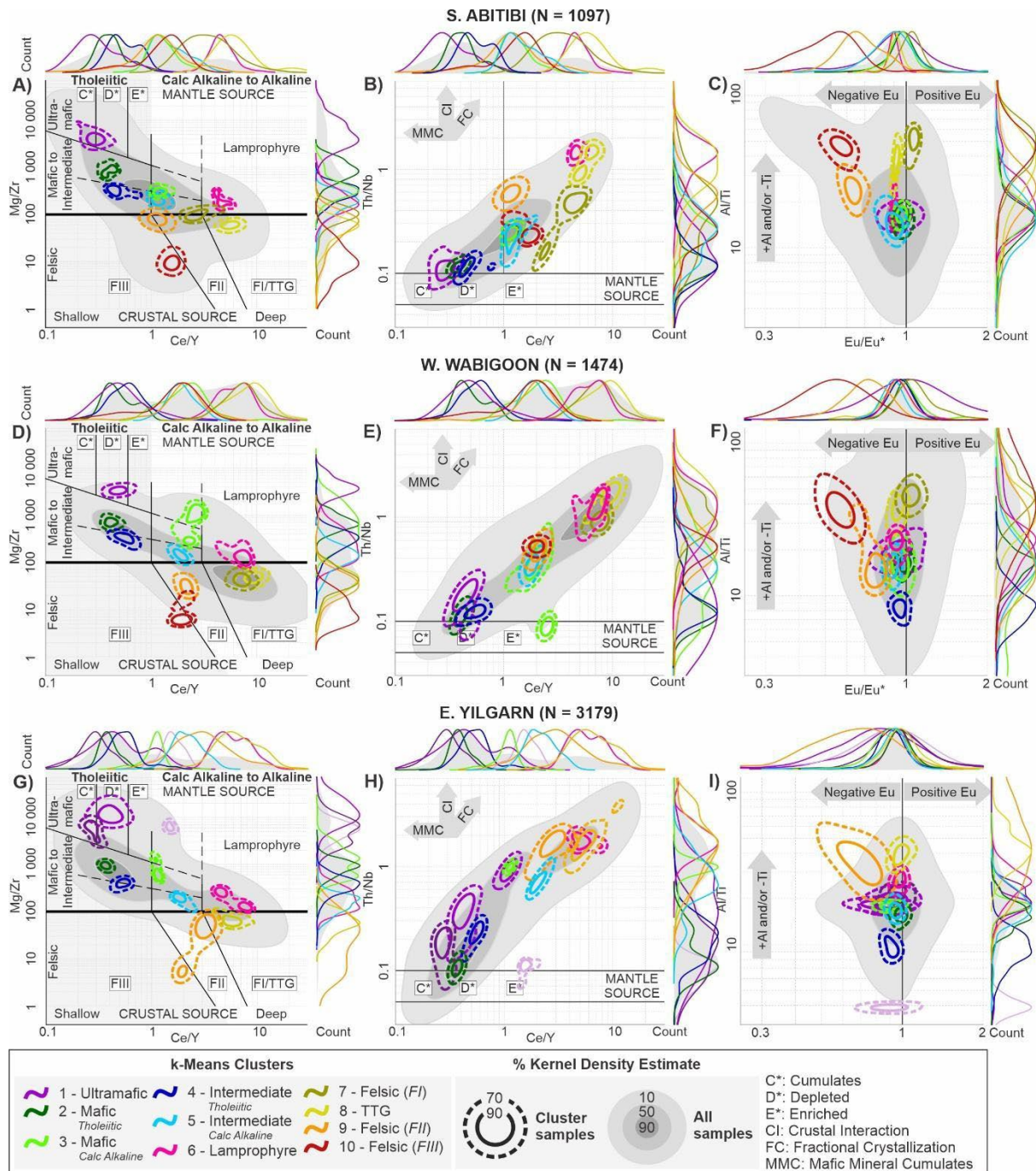
SM Figure 5-1.4. Kernel Density Estimate (KDE) of interpreted compositional clusters plotted on the proposed classification scheme for the southern Abitibi (A, B, C), western Wabigoon (D, E, F), and eastern Yilgarn (G, H, I). A normalized histogram of each compositional group is indicated for each axis.

SM Figure 5-1.5. Detailed comparisons of ultramafic clusters from the southern Abitibi and western Wabigoon (A, B, G), eastern Yilgarn (C, D, G), and other Archean cratons (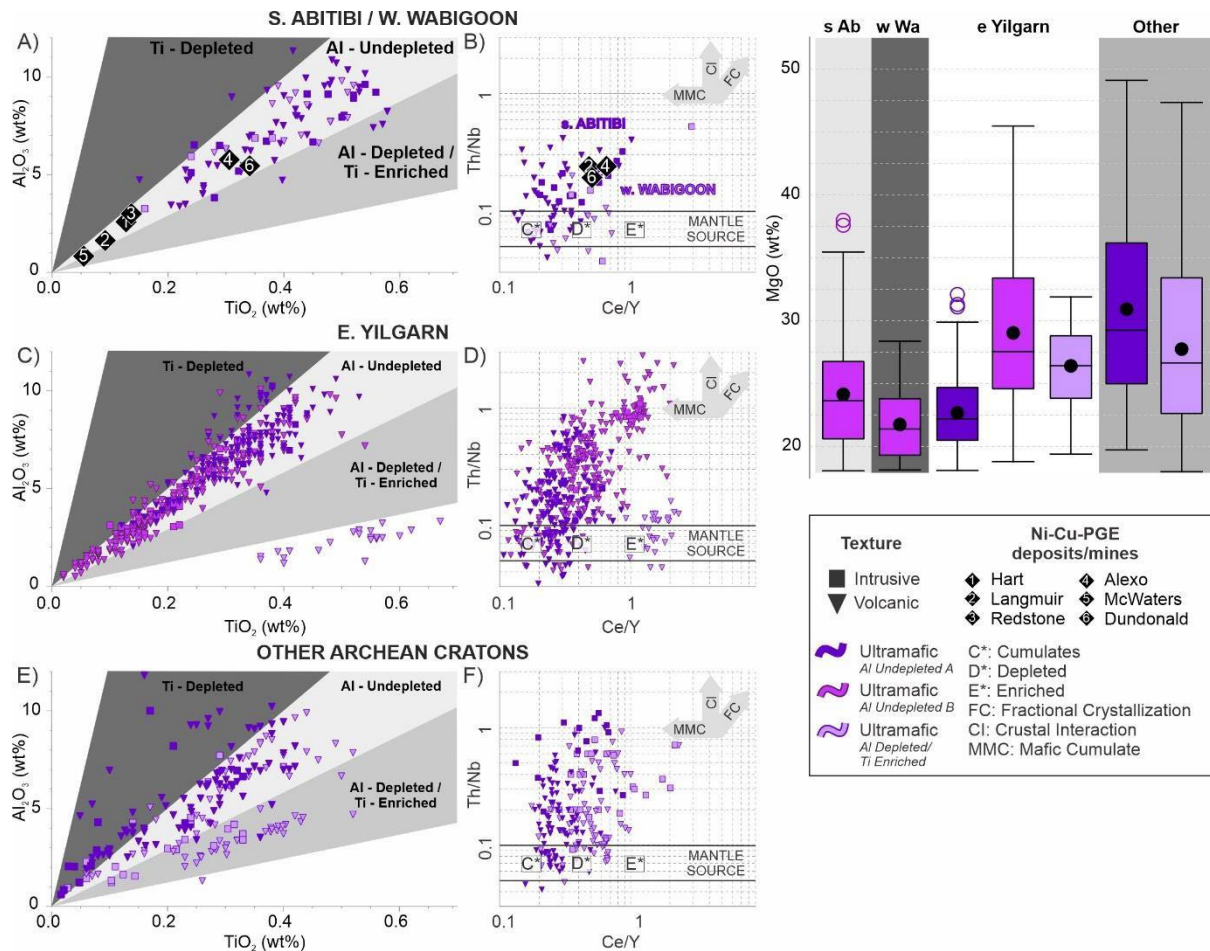E, F, G). Clusters are classified using Al2O3 and TiO2 concentrations according to Nesbitt et al. (1979) and Sproule et al. (2002)'s criteria for classification of ultramafic rocks by mantle source composition and degrees of partial melting (A, C, E). Ultramafic clusters are further classified (B, D, F) using mafic cumulate/mantle source composition (Ce/Y) vs crustal interaction and fractional crystallization (Th/Nb). Samples nearby known magmatic deposits in the southern Abitibi are plotted as numbered diamonds to indicate favorable geochemical signatures. Some deposit-related samples did not contain the full range of analyses and could not be plotted on both diagrams. G) Finally, the MgO concentrations of each cluster is compared and used a proxy for plume temperature and ability to take up crustal S. A normalized histogram of each distance group is indicated for each axis.
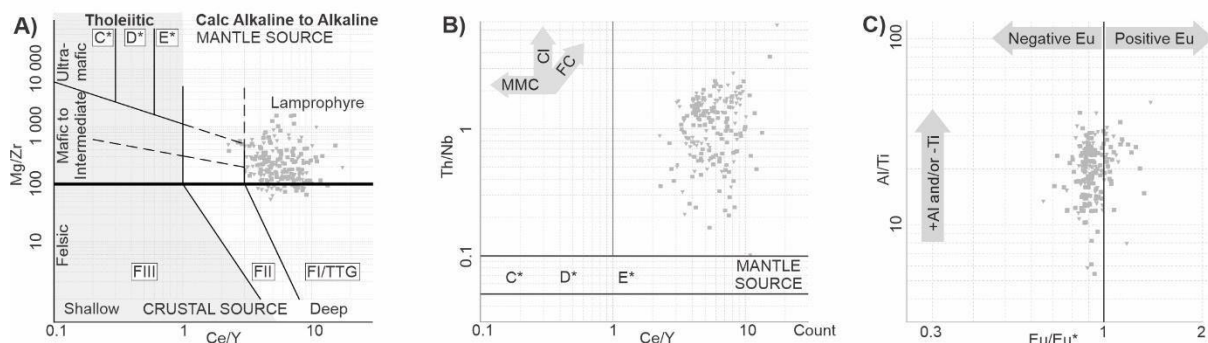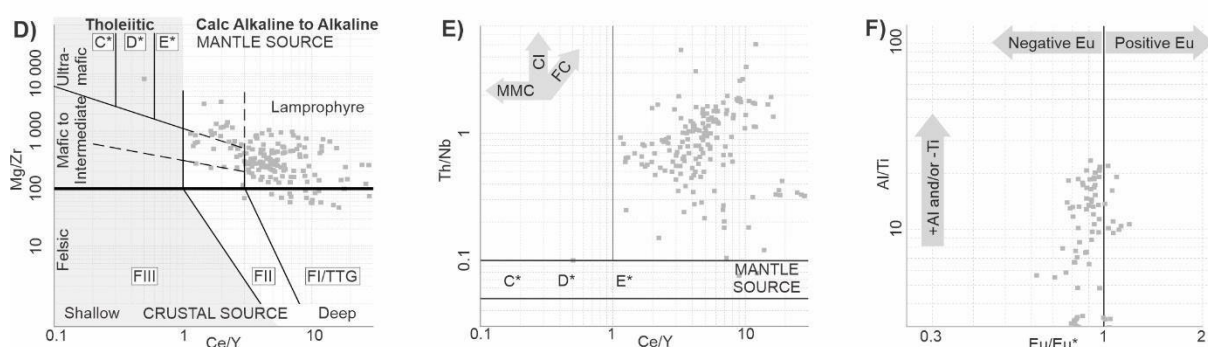
SM Figure 5-1.6. Geochemical signature of hand samples identified as 'lamprophyre' collected in the southern Abitibi and western Wabigoon (A, B, C) and other Archean cratons Yilgarn (D, E, F) in the context of the proposed classification
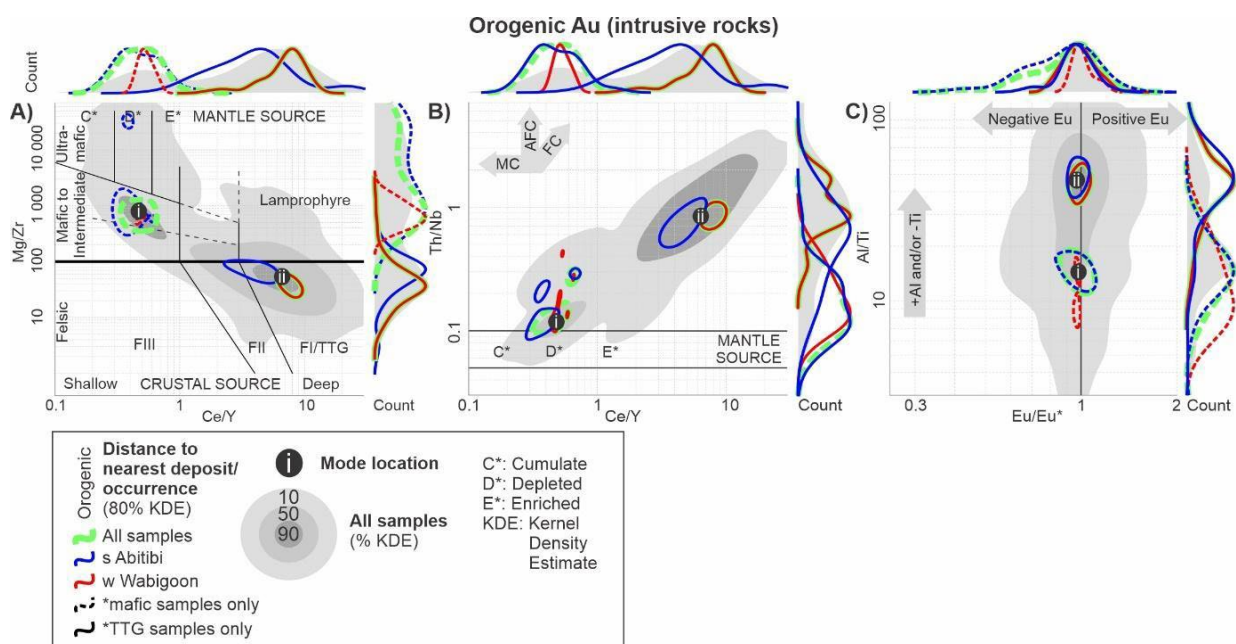
SM Figure 5-1.7 (previous page). Geochemical signature of samples represented as kernel density estimates for intrusive samples spatially associated with orogenic Au deposits/occurrences. The characteristics of samples from each area and dominant compositional class have been modeled separately to compare differences.
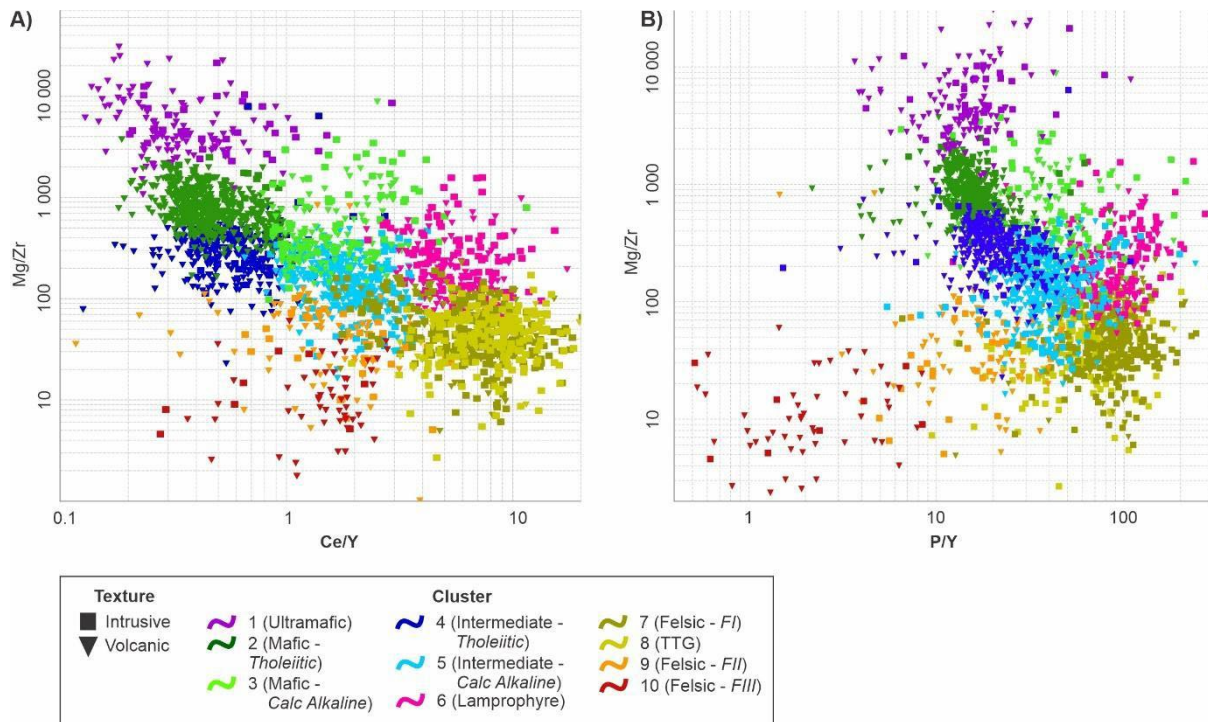
SM Figure 5-1.8. Comparison of trends and discrimination between clusters using the proposed Ce/Y that represent variance along PC-2 and P/Y that is geological significant for felsic and lamprophyric rocks.

Supplementary material 5-2: Directed study presentation

A set of PowerPoint slides were submitted as material for a directed study and present geochemical analytical background information and results of PCA and clustering using an early iteration of the geochemical compilation database. The PDF can be accessed at supplementary material with the submitted *Data in Brief* publication and is stored on Mendeley Data (https://data.mendeley.com/datasets/f493ctkr38)

Supplementary material 5-3: Supplementary data

The lithogeochemical data in Mendeley Data (https://data.mendeley.com/datasets/f493ctkr38) are available as Excel spreadsheets separated by location (western Wabigoon and southern Abitibi subprovinces) and in a complete format where data from both areas are combined (sample sites with geological context indicated in Figure 4-1). Laboratory brochures that describe packages and methods used to generate each dataset are also included.

Supplementary material (Chapter 6)
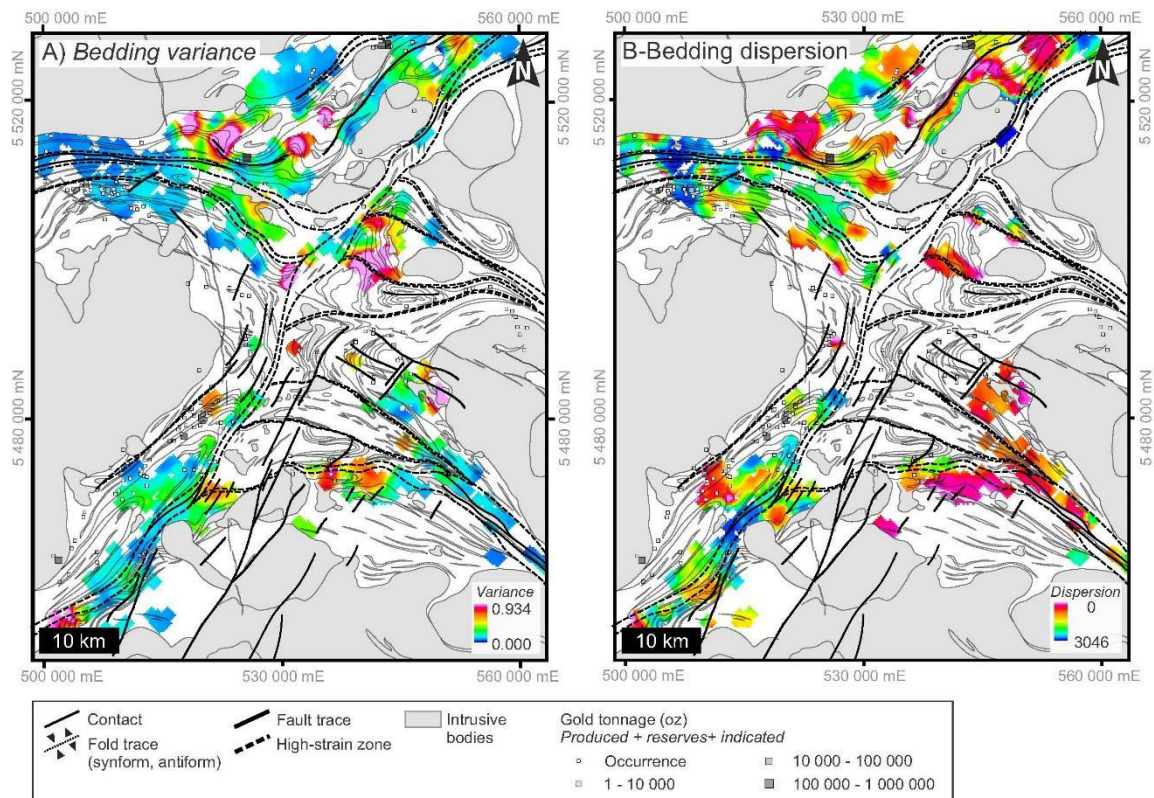
SM Figure 6-1.1. Regional structural complexity results from A) variance and B) dispersion of bedding measurements; Both maps display poor coverage for the Dryden map area; Grid of variance of strike displays broad high and low variance anomalies that are loosely associated with interpreted structures; Trends for dispersion, however, are chaotic and at a regional scale do not correlate well with geological interpretations; Color ramp of variance result rescaled using percent clip technique in ArcMap (0.5 minimum and 0.5 maximum) and the histogram equalize color rescaling method for dispersion; Coordinates in NAD 83 UTM zone 15N.
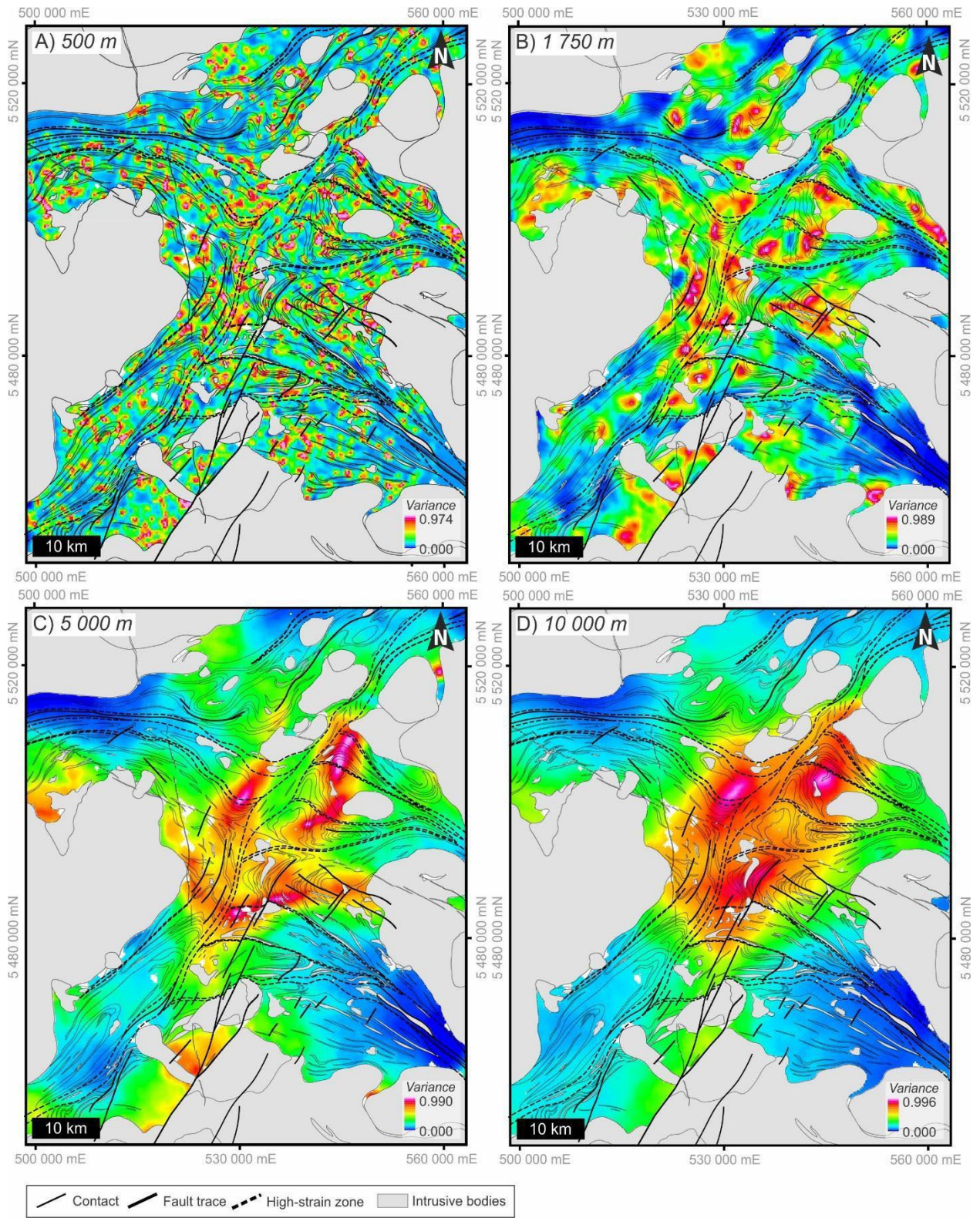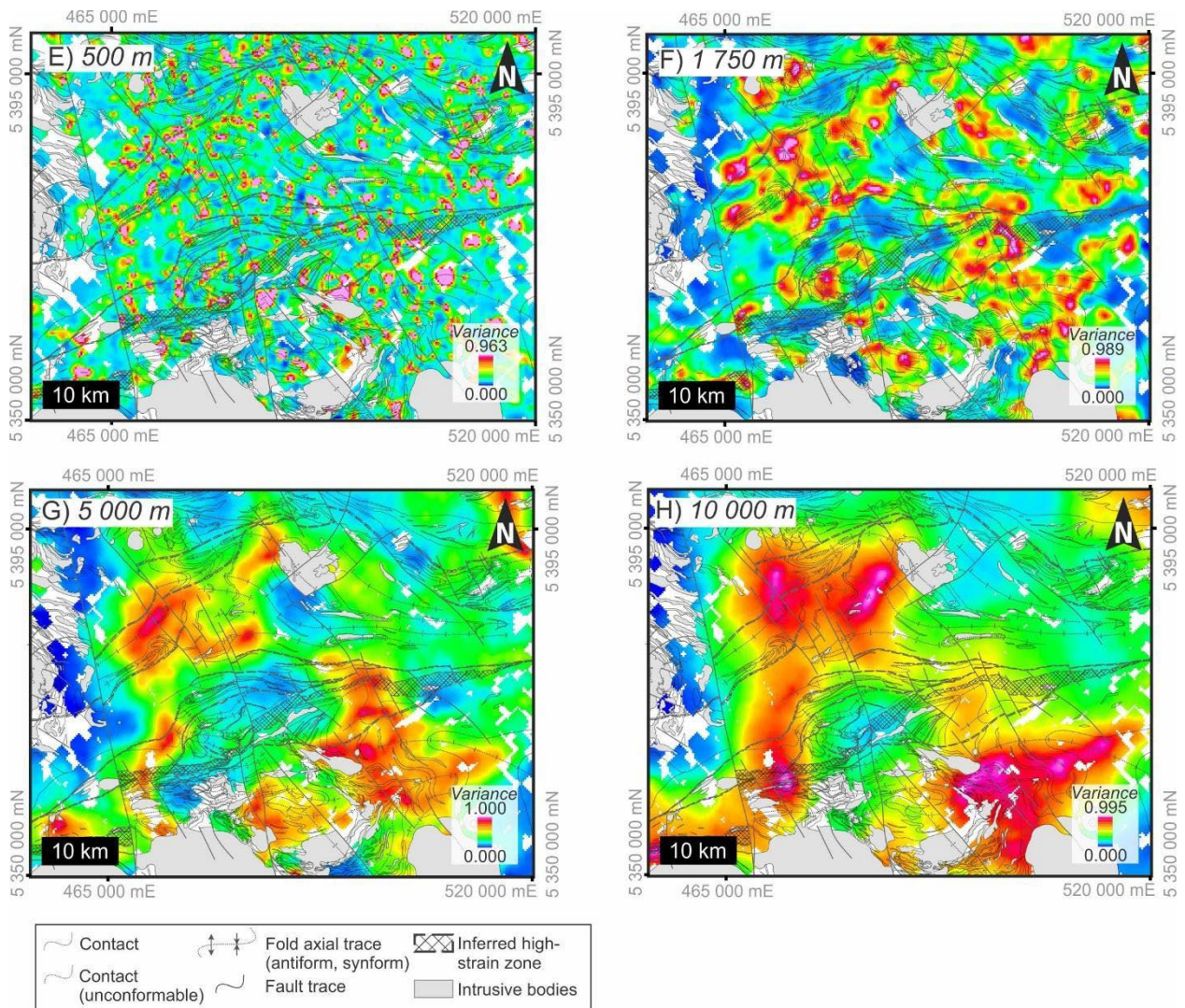
SM Figure 6-1.2 (this and previous page). 200 m lineament segments from (A – D) Dryden and (E - H) Timmins were used to test the effect of neighborhood radius on variance results. Neighborhoods with A) 100 m, B) 500 m, C) 1000 m, and D) 2000 m radii were tested and grids from Dryden and Timmins show that neighborhood radius significantly influences results. Generally, as neighborhood radius increases, high and low variance anomalies progressively broaden and eventually merge; Grids are symbolized using standard deviation (n=2.5) rescaling with a 0.5 gamma stretch; Coordinates in NAD 83 UTM zone 15N in Dryden and Zone 17N in Timmins.
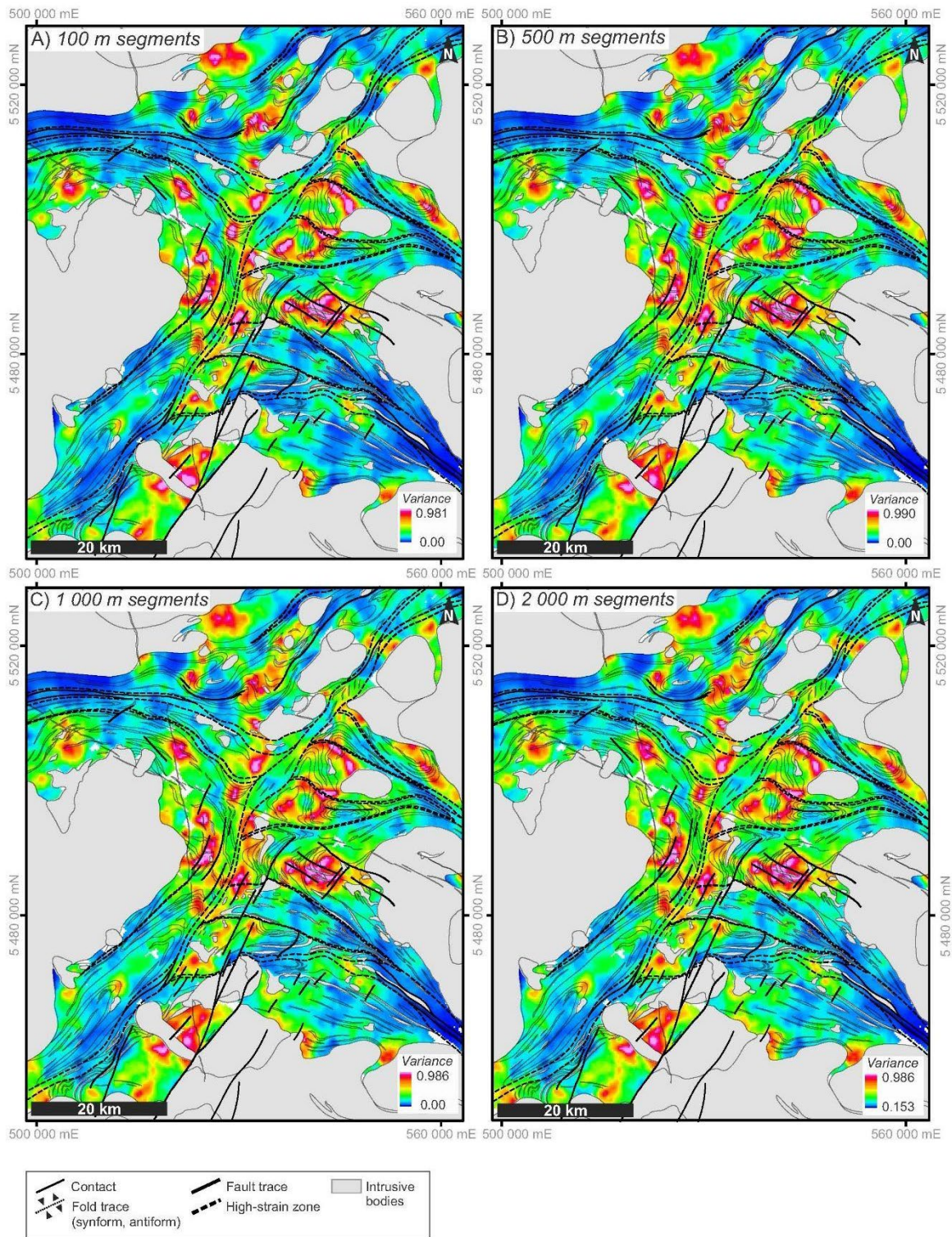
SM Figure 6-1.3. To test the effect of lineament segment length on variance results, lineaments segments of A) 100 m, B) 500 m, C) 1000 m, and D) 2000 m, were used as inputs for variance calculations using all other input parameters described in Table 3. All grids display co-located high and low variance anomalies and increasing segment length only subtly smooths anomalies; Grids are symbolized using standard deviation (n=2.5) rescaling with a 0.5 gamma stretch; Coordinates in NAD 83 UTM zone 15N.
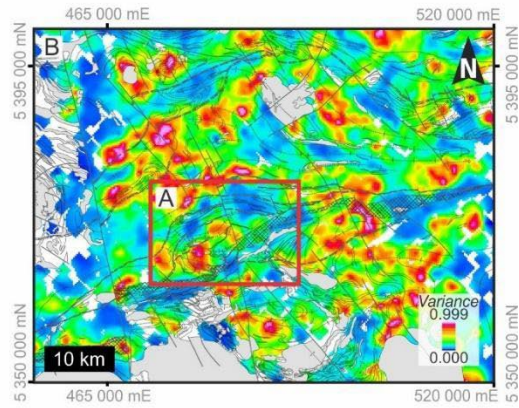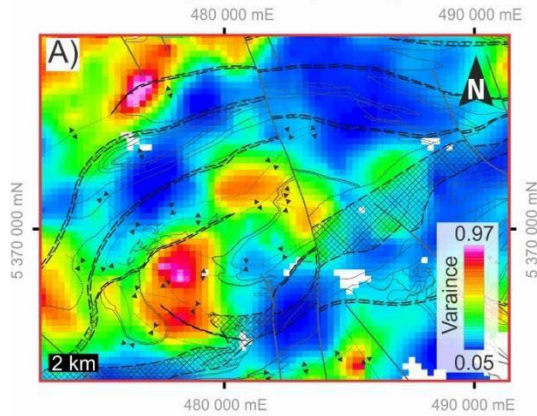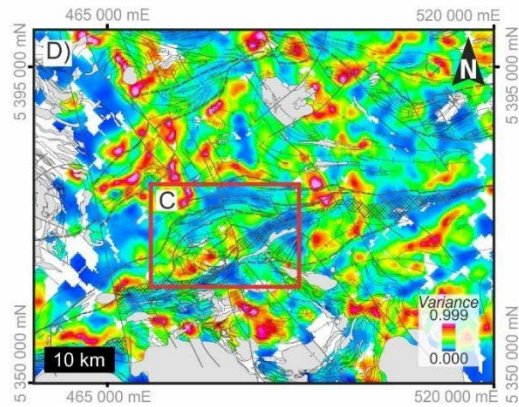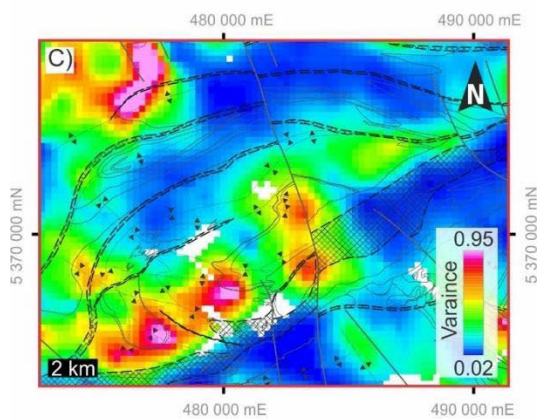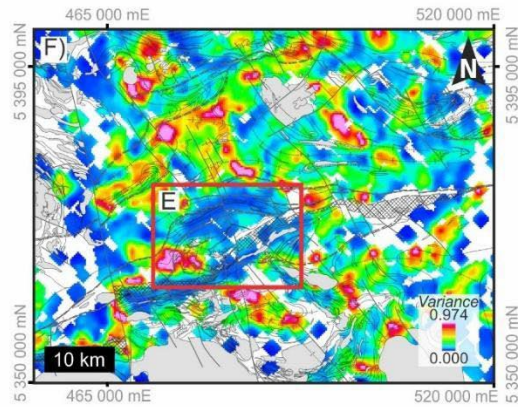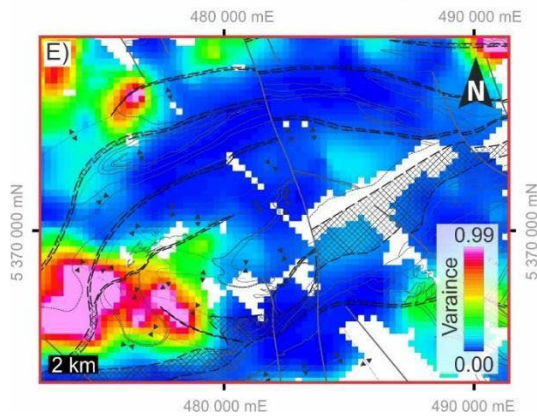
SM Figure 6-1.4. While testing the effect of maximum wavelength in PPDRF filtering of aeromagnetic grids on lineament extraction and variance results, three representative grids were plotted for Timmins; 200 m lineament segments were extracted from grids with maximum wavelengths A) 37 cells (1480 m), B) 72 cells (2880 m), and C) 139 cells (5560 m); Each regional scale grid had an associated inset centered on the Timmins gold cap; At a regional scale, high and low variance anomalies are co-located; however, at local scales, anomalies differ in location and intensity; See Figure 9 for a comparison of input PPDRC grids and lineaments; Variance calculations used all other optimal parameters presented in Table 3; Variance grids are symbolized using standard deviation (n=2.5) rescaling with a 0.5 gamma stretch; Coordinates in NAD 83 UTM Zone 17N.
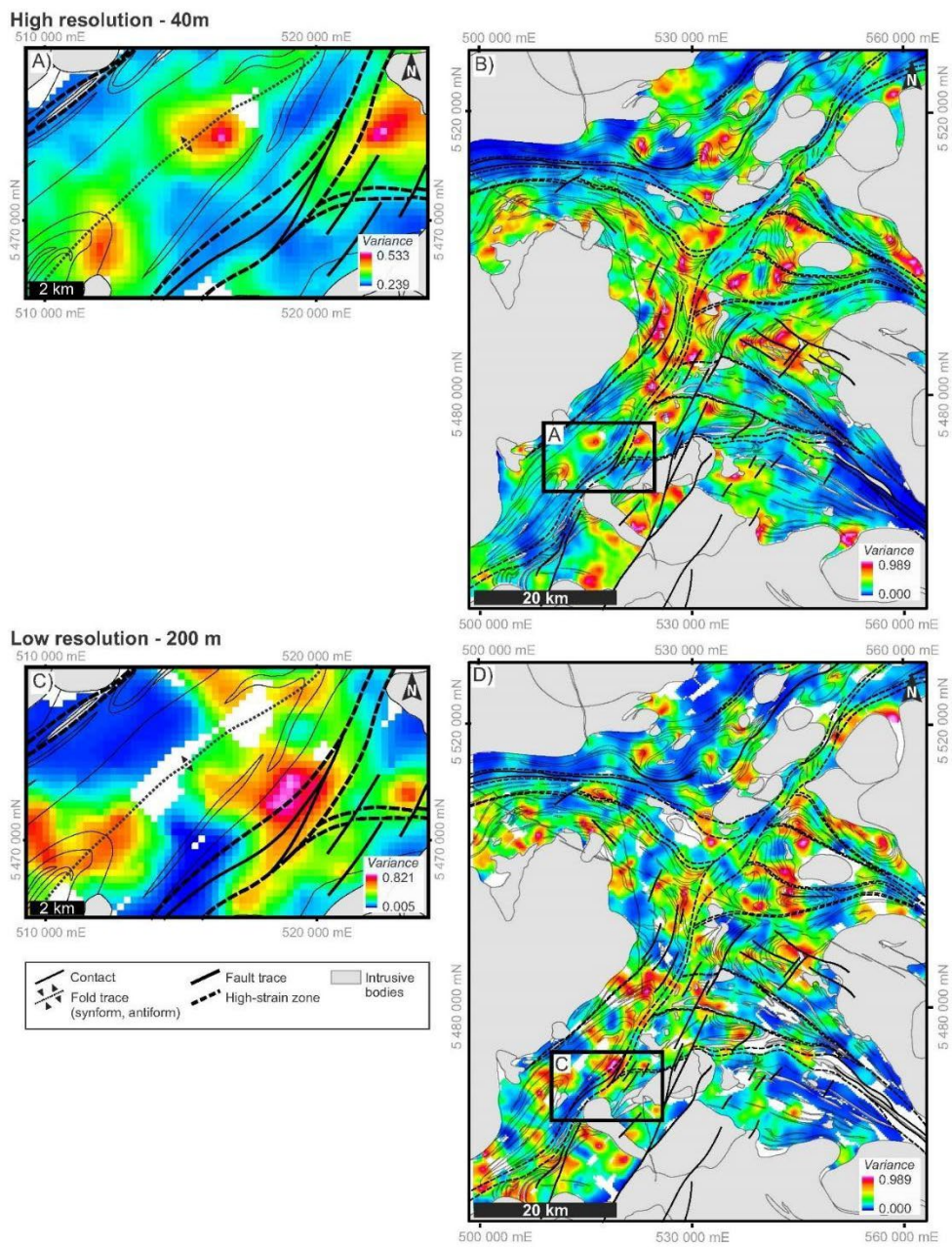
SM Figure 6-1.5. Local and regional variance grids from 200 m lineament segments using two different input magnetic grid resolution for the Dryden map area. Line spacing in aeromagnetic survey used to generate these results are A) 40 m and B) 200 m. Variance calculations used a 1750 m neighborhood radius, 200 m lineament segment lengths, and are symbolized using standard deviation (n=2.5) rescaling with a 0.5 gamma stretch. The high-resolution maps (A) used a PPDRC maximum wavelength of 72 cells. The low-resolution grids (B) used a PPDRC maximum wavelength of 19 cells; Coordinates in NAD 83 UTM zone 15N.

Supplementary material (Chapter 7)

Supplementary material 7-1: Description of non-Euclidean distance method and presentation of map results

Graph-based methodology is used to capture topology/connectivity of features across a geological map area and calculate a non-Euclidean distance of fluid paths prior to deformation.

Method

The description of geology as graphs pre-dates routine geospatial analysis and facilitates interrogation of stratigraphic and fault relationships in a single representation (Burns, 1975). This can be extended to analysing potential preferred fluid pathways in fault networks (Ghaffari et al., 2011) and anisotropic distance metrics in folded strata (Boisvert & Deutsch, 2011). A simplified schematic of the workflow is presented as four stages in SM2 – Figure 1.

Stage one: To begin, the nodes of map features (e.g., fault trace, map unit polygon) constrain Delaunay triangulation, which involves that utilizes Voronoi tessellation and natural neighbours of node points to generate a triangulated mesh (Miles, 1972). During this process, duplicate points are removed along faults and other features that cross-cur map units within a buffer zone. The resulting triangulated mesh will have internally consistent topology where nodes representing edges and triangle centroids are tabulated and assigned to a geological feature. These tabulated and labelled nodes are used as inputs for later graph analysis.

Stage two: Graph edges represent the connectivity between fault segments, between adjacent stratigraphic volumes, and between fault surfaces and stratigraphic volumes (SM2 – Figure 1B to D) are calculated and encode topology of geological features.

SM Table 7-1.1. Qualitative ranked weights assigned to map units for use in non-Euclidean distance calculation using graph analysis

Weight	Feature
1	Shear zones (crustal-scale)
2	Shear zones (secondary splays)
4	Fold axial planes
5	Unconformities
5	Intrusive contacts
6	Conformable volcanic contacts
6	Chemical sediments (aka BIF/cherts)
7	Clastic sediments
7	Felsic volcanic units
8	Ultramafic/Mafic volcanic units
10	Intrusive bodies (all compositions)

Stage three: Both nodes and edges were qualitatively assigned relative weights based on inferred fluid impedance ranking (reported in SM Table7-1.1).

Stage four: Once the weights have been assigned, standard graph analytic tools (Hagberg et al., 2008) were used to calculate the minimum 'geological distance' between a vectorized map feature (e.g.,

fault trace, map unit polygon) and every node in the network (i.e., node on mesh). Distance represents the sum weighted nodes and edges.

This approach provides new avenues for the interpretation of fluid-related processes as well as interpretation of alteration or mineralization patterns in complexly deformed terranes. Additionally, this strategy of reducing spatial relationships to a graph or network acts akin to a structural restoration and may facilitate more comprehensive geological investigations where traditional restoration is not feasible.
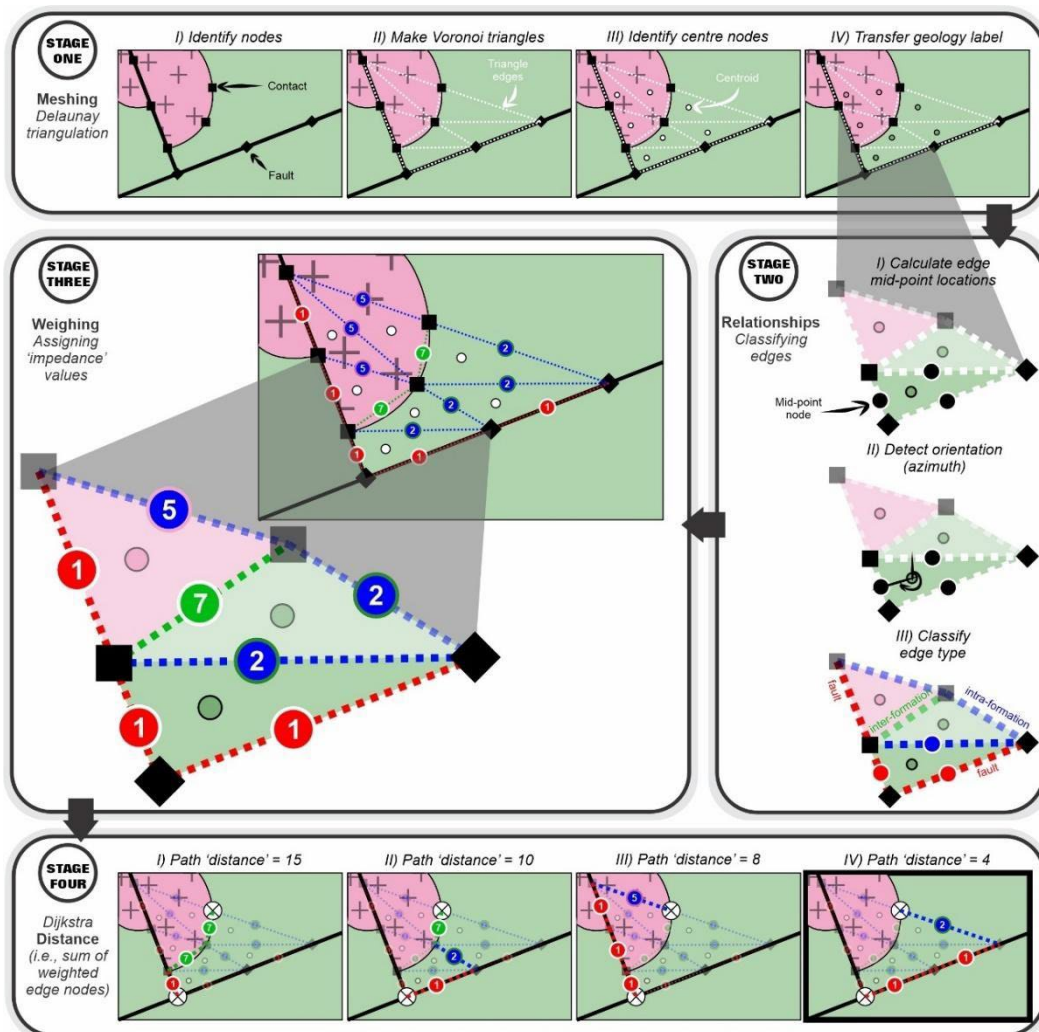
SM Figure 7-1.1. Workflow stages of non-Euclidean distance calculation

Results

Results of interpolated minimum non-Euclidean distance within the graph are presented in the following maps. Each result is labelled based on location and feature and symbolized using a 'histogram equalize' transformed colour ramp from ArcGIS ProTM. A geologic map for each map area is also provided to give context.

SM Figure 7-1.2 & 7-1.3 (next 5 pages). Map results for non-Euclidean distance maps for Timmins and Dryden with superimposed geology and mineral deposits/occurrences for context

A) Geological context

B) Ultramafic volcanic

C) Mafic volcanic

D) Felsic volcanic

E) Ultramafic intrusion

F) Mafic intrusion

10 km

633 / 0

4300 / 0

4370 / 0

5240 / 0

4430 / 0

Inset map

Political boundary
Sub-province
Superior Craton
Continental landmass

Map elements

Fault trace
Contact (unconformable)
Inferred high-strain zone
Water body

500 km

Mineralization

Prospect or deposit / Occurrence

▲ △ Magmatic
⬠ ○ Volcanogenic
■ □ Orogenic

Lithology

Felsic to intermediate intrusive rocks
Intermediate to mafic intrusive rocks
Ultramafic intrusive rocks
Chemical sedimentary rocks
Clastic sedimentary rocks
Felsic to intermediate volcanic rocks
Intermediate to mafic volcanic rocks
Ultramafic volcanic rocks

M) NE-SW deformation zone (trans-crustal)

Supplementary material 7-2: Description of discrete gridding map results

A single value, representing the most reasonable representation of fresh Archean rock types, was assigned by compiling values from a wide range of geophysical and geomechanical studies. Values used in this application of rock properties are reported in SM Table 7-2.1. In some cases, values could not be obtained from literature and were necessarily inferred or calculated from similar rock types. These include total porosity of ultramafic volcanic rocks, which was inferred from reported values of ultramafic intrusions. Magnetic susceptibility, resistivity, and chargeability of ultramafic intrusion and volcanic were inferred from mafic intrusion and volcanic, respectively. Mafic gneiss values inferred from amphibolite samples wherever available. Finally, the average of all gneiss values was used to represent basement properties

Proxies for ductile properties (i.e., bulk, shear, and compression moduli) were calculated using reported Poisson ratio and Young's modulus. Value source references are indicated by superscript Greek symbols for each property or individualized by value where relevant. A table of associated references to each symbol are reported in SM Table 7-2.2.

		Intrusion				Volcanic			Sedimentary				Gneisses
		Ultramafic	Mafic to int.	TTG	Int. to felsic	Ultramafic	Mafic to int.	Int. to felsic	Conglomerate (molasse basin)	Silt to sandstone (rift basin)	Chemical (BIF & chert)	Chemical (BIF)	Gneisses
Independently gridded	Grain density$^\Sigma$ (kg/m³)	2893	2816	2689	2679	2891	2824	2740	2796	2736	3217		2844
	Total porosity$^\Sigma$ (%)	0.10	0.80	1.15	0.91	0.10*	1.75	1.12	2.03	0.33	0.74		0.96
	Rock resistivity$^\Sigma$ (Ω/m)	39637	38475	5986	42565	107596	635664	259825	222830	40004	355782		29396
	Magnetic susc.$^\Sigma$ (SI)	0.0210	0.0205	0.0132	0.0102	0.0345	0.0159	0.0113	0.0047	0.0042	0.0454		0.0148
	P-wave$^\beta$ (m/s)	7991$^\theta$	5865	4678	5038	5594	4814	3186	3014	3408			
	S-wave$^\beta$ (m/s)	4539$^\theta$	3428	2625	2755	3120	2810	2045		2102			
	ν - poisson ratio$^\beta$	0.24^i	0.24	0.25	0.25	0.28	0.22	0.26	0.19	0.22		0.19$^\lambda$	0.21
	E - Youngs modulus$^\beta$ (Gpa)	82$^\rho$	87	48	43	71	55	29	59	30		44$^\lambda$	74
Calculated ductile properties	K - Bulk modulus$^{\beta C}$	52.6	55.8	32.5	27.9	54.1	32.5	20.5	31.6	17.7		23.6$^{\lambda C}$	42.7
	G - Shear modulus$^{\beta C}$	33.1	35.3	19.3	17.1	27.6	22.7	11.6	24.9	12.3		18.5$^{\lambda C}$	30.4
	M - Compression modulus$^{\beta *}$	96.7	102.9	58.2	50.7	90.9	62.8	35.9	64.8	34.2		48.3$^{\lambda C}$	83.3
Brittle properties	Cohesion$^\beta$ (Mpa)	27.23$^\theta$	31.9	25.5	31.4	14.9$^\theta$	17.8	14.6	8.56^X	39.65$^\Psi$	0.35$^\mu$		29.6
	Friction coefficient$^\beta$	50.54$^\theta$	58.0	53.3	44.2	41.54$^\theta$	34.2	29.0	37^X	47.45$^\Psi$	30.28$^\mu$		51.8
		146.03$^\theta$	218	145	127	59.89$^\theta$	145	74	52	88	96$^\mu$		163.2

SM Table 7-2.1 (previous page). Averaged/ Idealized/Reported/ Calculated physical properties (e.g., magnetic susceptibility, total porosity, grain density, bulk thermal conductivity), mechanical properties (e.g., p- and s-wave velocities, shear modulus, bulk modulus, compression modulus, frictional coefficient, cohesion, uniaxial compressive strength), and electrical properties (e.g., rock resistivity, conductivity) of each lithologic type found in Archean greenstone belts near Timmins and Dryden, Ontario. Modifier to indicate processing step for missing values derived from (c) calculations, or (i) inferred from similar rock types. References for values reported in SM3 – Table 2

SM Table 7-2.2. References for values used in rock property table. Greek letters indicate specific references for values where applicable (superscript for individual cells) or groups (superscript for group headings)

β	Bär, K., Reinsch, T., Bott, J. (2020) The PetroPhysical Property Database (P3) – a global compilation of lab-measured rock properties. Earth System Science Data, 12(4), 2485–2515. https://doi.org/10.5194/essd-12-2485-2020
θ	Diamantis, K. , Exarhakos, G. , Migiros, G. and Gartzos, E. (2016) Evaluating the Triaxial Characteristics of Ultamafic Rocks from Central Greece Using the Physical, Dynamic and Mechanical Properties. Open Access Library Journal, 3, 1-20. doi: 10.4236/oalib.1103214
Σ	Enkin, R J; Cowan, D; Tigner, J; Severide, A; Gilmour, D; Tkachyk, A; Kilduff, M; Baker, J (2012) Physical property measurements at the GSC paleomagnetism and petrophysics laboratory, including Electric Impedance Spectrum methodology and analysis. Geological Survey of Canada, https://doi.org/10.4095/291564
μ	Joyal, G., Esmaeilli, K., Garand, P. (2012) Joyce Lake and area DSO Project Geotechnical engineering Feasibility Study - Open pit. Labec Century Iron Ore Inc. 267pp
λ	Lanaro, F., Bygg Konsult, B., Öhman, J., and Fredriksson, A., 2006, Rock mechanics modelling of rock mass properties – summary of primary data. Swedish Nuclear Fuel and Waste Management Co. 68 pp, https://www.osti.gov/etdeweb/biblio/20843898
ψ	Li, D., Wong, L. N. Y., Liu, G., and Zhang, X. (2012). Influence of water content and anisotropy on the strength and deformability of low porosity meta-sedimentary rocks under triaxial compression. Engineering Geology, 126, p46–66. https://doi.org/https://doi.org/10.1016/j.enggeo.2011.12.009
ρ	Palmström, A. and Singh, R. (2001). The deformation modulus of rock masses — comparisons between in situ tests and indirect estimates. Tunnelling and Underground Space Technology, 16(2), 115–131. https://doi.org/https://doi.org/10.1016/S0886-7798(01)00038-4
ꭙ	Zhu, Q., Wang, B., Zhao, X., Liu, C., Yu, Q., Zhang, L., Hou, T., and Sang, G. (2020). Effect of Coalbed Methane Well Fracturing on Slope Stability of Open-Pit Coal Mine: A Case Study of Shengli East No. 2 Open-Pit Coal Mine. Advances in Civil Engineering, 2020, 4708274. https://doi.org/10.1155/2020/4708274

Results

Results of assigned rock properties to map unit polygons are presented in the following maps. Each map is labelled based on rock property and is symbolized using a 'stretched standard deviation' transformed colour ramp from ArcGIS ProTM. A geologic map for each map area is also provided to give context.

SM Figure 7-2.1 & 7-2.2 (next two pages, respectively). Map results for discrete gridding of rock properties assigned to map units near Timmins and Dryden with superimposed geology and mineral deposits/occurrences for context
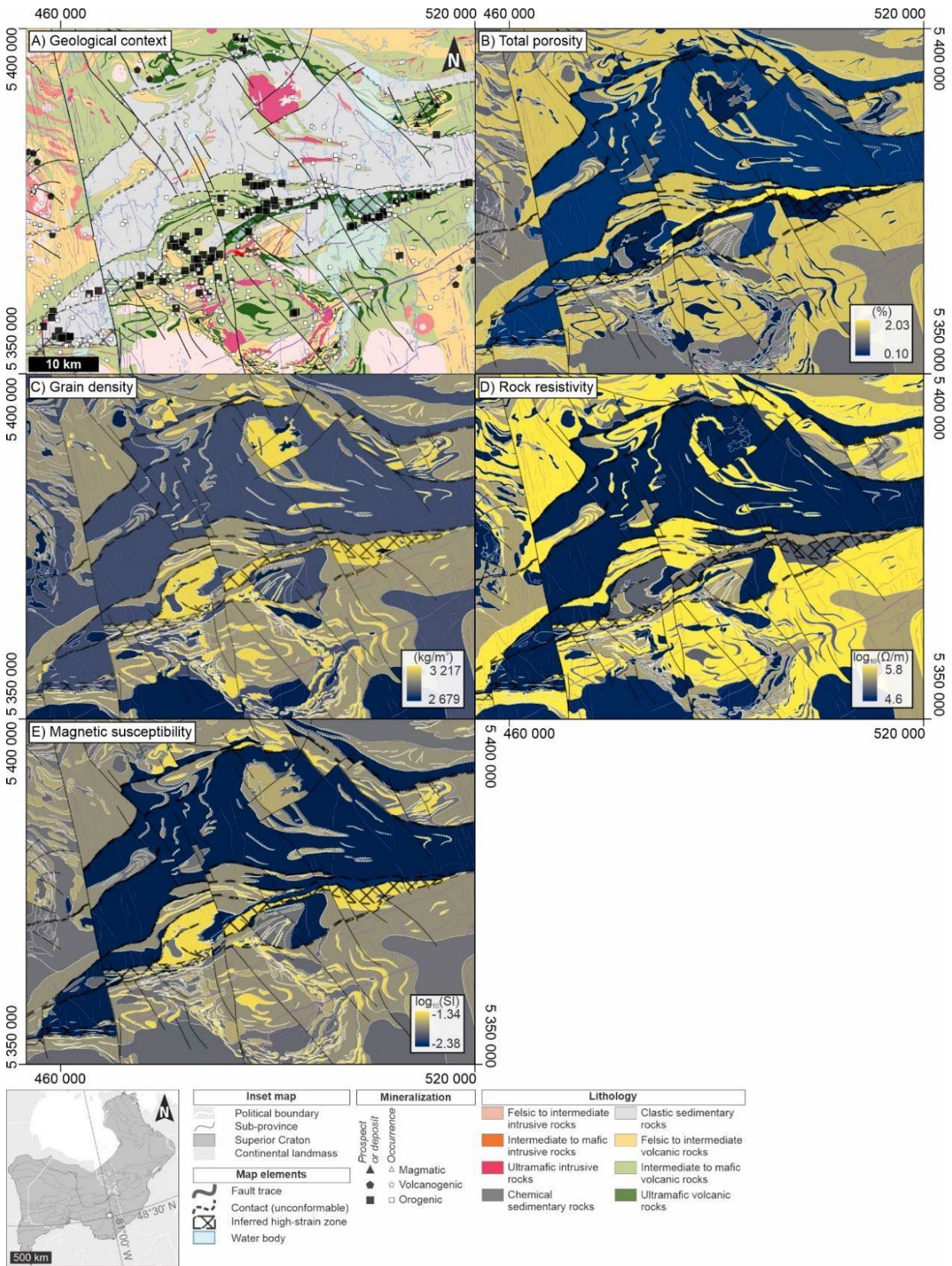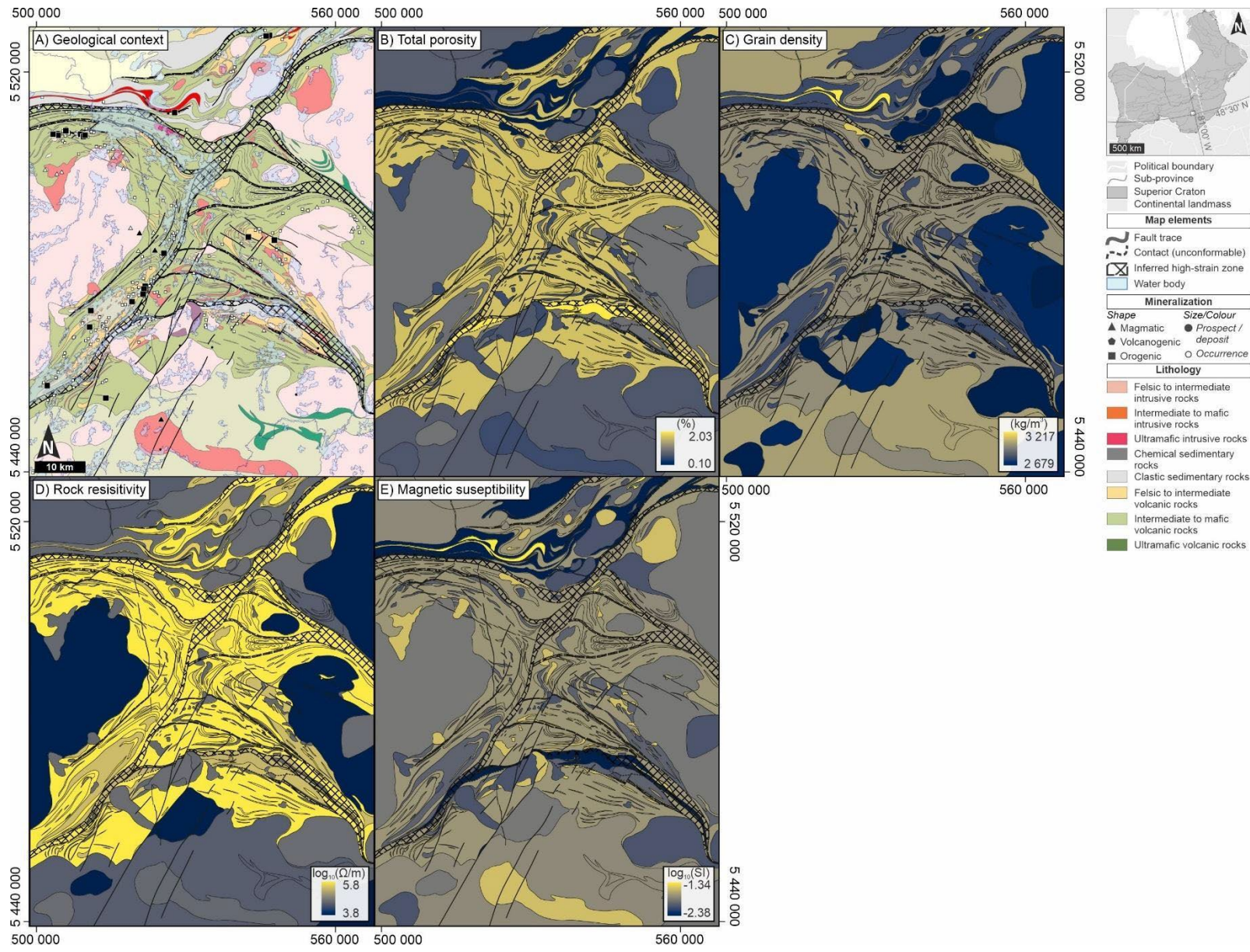
Supplementary material 7-3: Description of semi-discrete gridding results

Element ratios grids log10(Mg/Zr, Ce/Y, Th/Nb, Eu/Eu*, Al/Ti) were generated using masked minimum curvature interpolations for each lithology type in each assemblage. These ratios are used to characterise lithological composition, nature of melt/magma sources, crustal interaction/fractional crystallization, and plagioclase/titanite-related igneous processes (after Montsion et al., submitted b/Chapter 5). Map unit polygons were used to constrain the interpolation as masks and thus, thus technique is considered 'semi-discrete'.

Results

Results of masked minimum curvature interpolated element ratios are presented in the following maps. Each map is labelled based on element ratio and is symbolized using a '30 geometrical interval classes' transformed colour ramp from ArcGIS ProTM. A geologic map for each map area is also provided to give context.

SM Figure 7-3.1 & 7-3.2 (this page and next). Map results for semi-discrete gridding of geochemical ratios maps for Timmins and Dryden with superimposed geology and mineral deposits/occurrences for context

A) Geological context

B) Mg/Zr

Log₁₀(Mg/Zr)
$Log_{10}(Mg/Zr)$
4.19
0.60

C) Ce/Y

$Log_{10}(Ce/Y)$
1.71
-0.73

D) Th/Nb

$Log_{10}(Th/Nb)$
1.65
-1.86

E) Eu/Eu*

$Log_{10}(Eu/Ei^*)$
0.39
-0.80

F) Al/Ti

$Log_{10}(Al/Ti)$
2.73
0.45

Political boundary
Sub-province
Superior Craton
Continental landmass

Map elements
Fault trace
Contact (unconformable)
Inferred high-strain zone
Water body

Mineralization
Shape
▲ Magmatic
⬠ Volcanogenic
■ Orogenic

Size/Colour
● Prospect / deposit
○ Occurrence

Lithology
Felsic to intermediate intrusive rocks
Intermediate to mafic intrusive rocks
Ultramafic intrusive rocks
Chemical sedimentary rocks
Clastic sedimentary rocks
Felsic to intermediate volcanic rocks
Intermediate to mafic volcanic rocks
Ultramafic volcanic rocks

SM Figure 7-3.3 to 7-3.8 (This and the next three pages). Screenshots of element ratio plots of samples from each litho-tectonic assemblage in the Timmins and Dryden areas. These ratio plots indicate the unique geochemical characteristics of each assemblage and facilitate discussion of geodynamic setting during emplacement/eruption as well as melt composition/evolution Montsion et al. (submitted b)/Chapter 5. Samples for a given assemblage are indicated in as large red symbols while those from the remainder of the area are small, black symbols. Volcanic samples are upside-down triangles and intrusive samples are squares.

Blake river
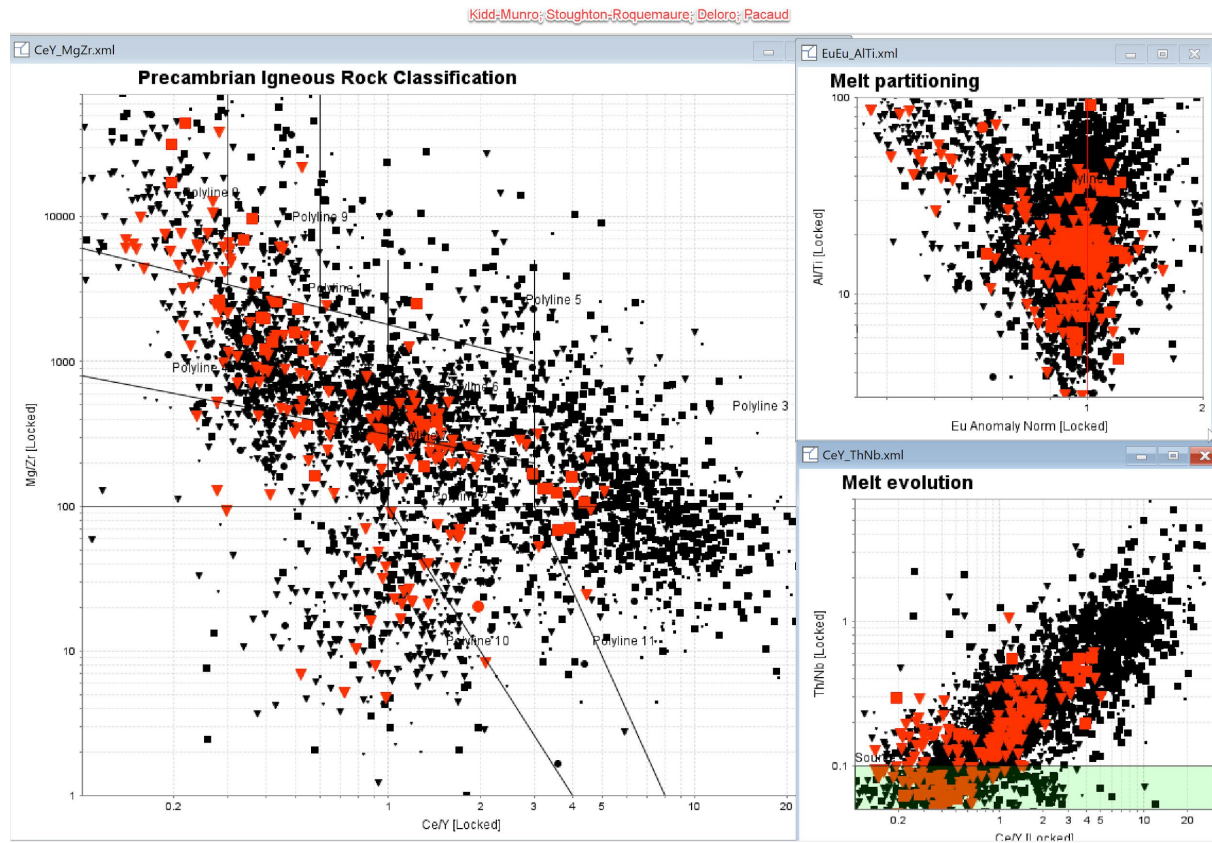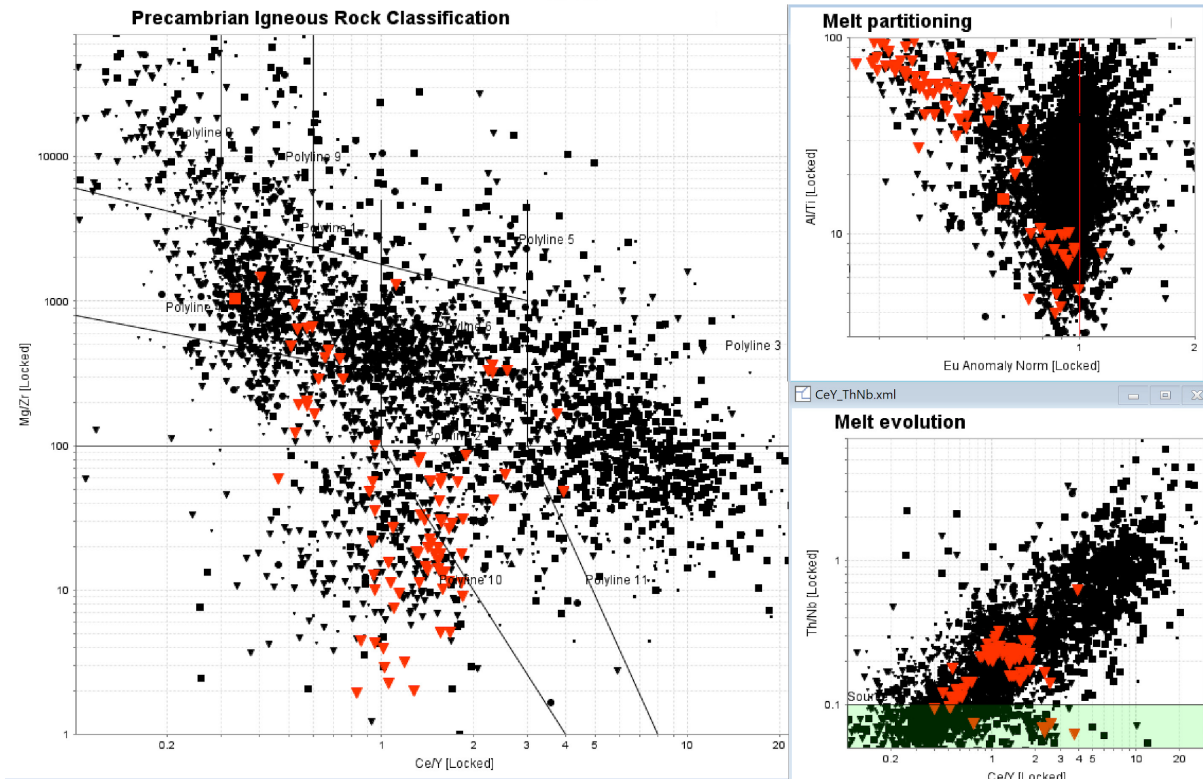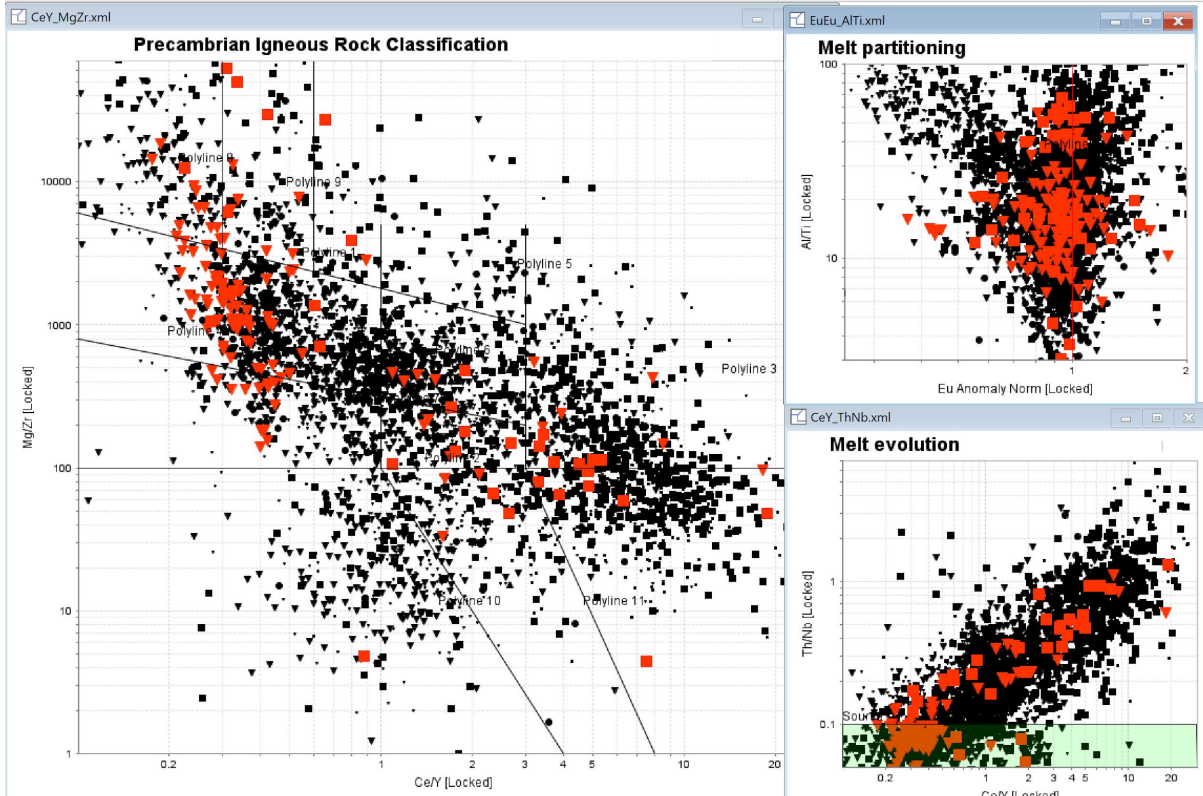
Wabigoon Assemblage
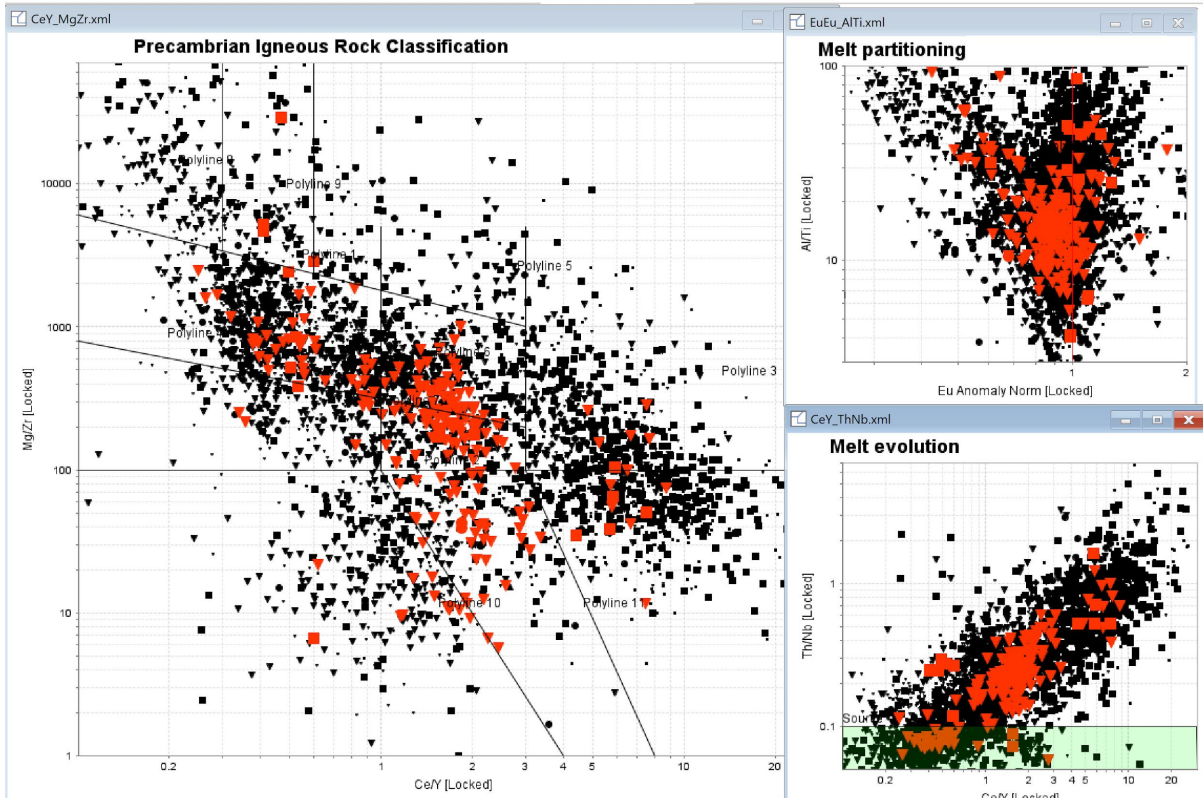
Kawashegamu and Wapageisi

CeY_MgZr.xml

Precambrian Igneous Rock Classification

EuEu_AlTi.xml

Melt partitioning

CeY_ThNb.xml

Melt evolution
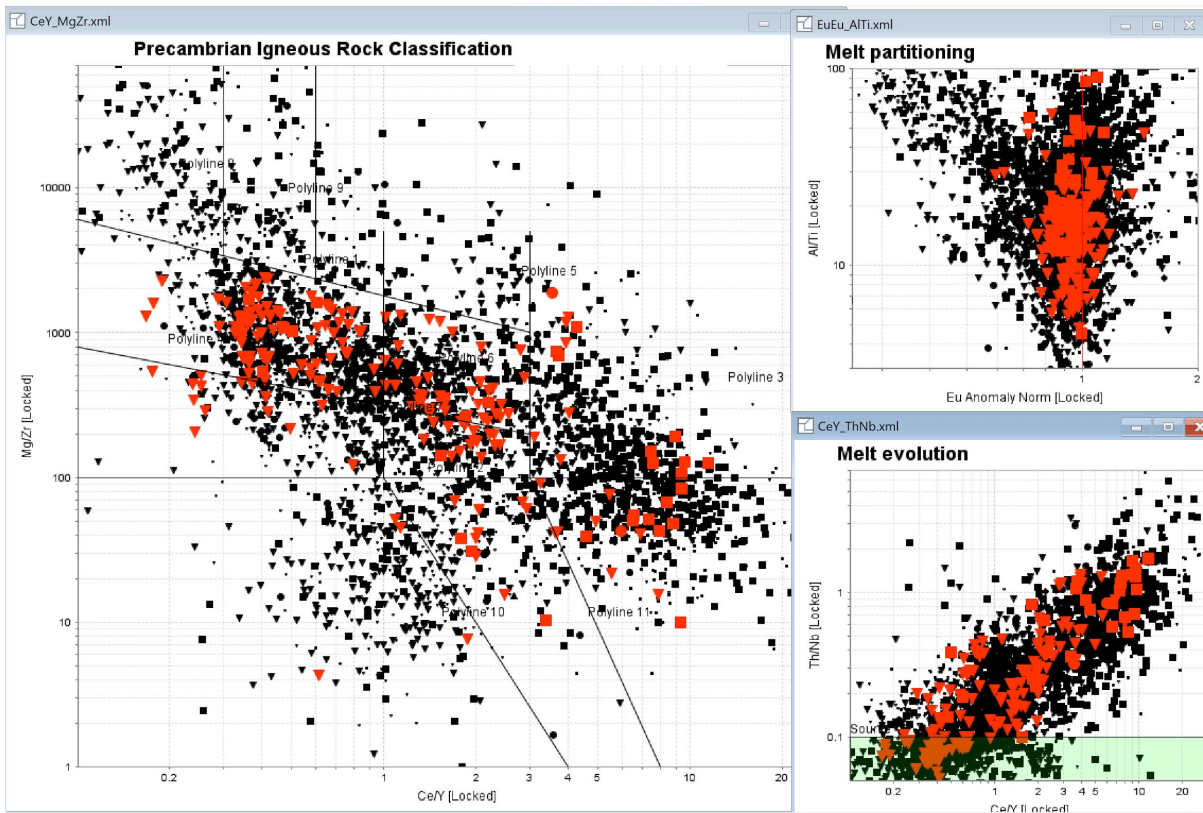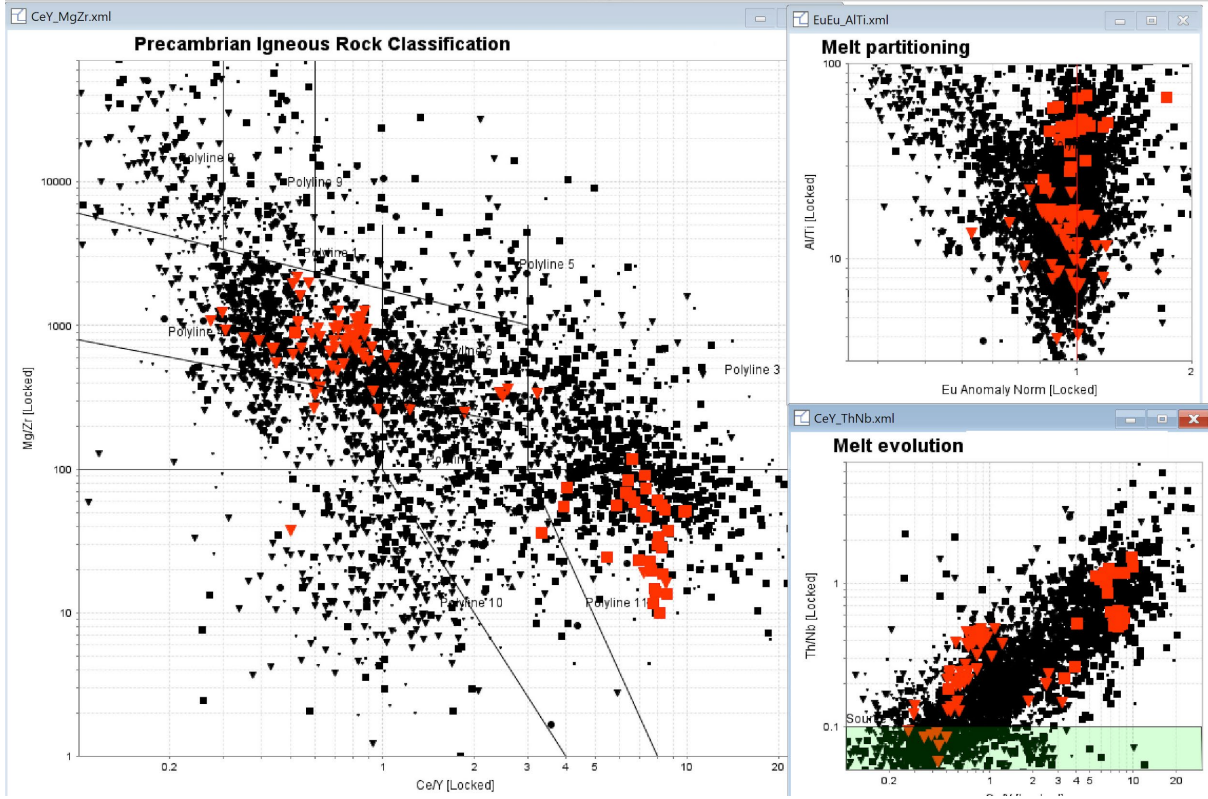
Supplementary material 7-4: Description of contrast maps

The sum variance grids of ductile (i.e., bulk, shear, compression moduli), brittle (i.e., cohesion, frictional coefficient, uniaxial compressive strength) and geochemical (element ratio grids) properties to identify areas of contrast (high variance) and uniformity (low variance).

Results

Results of rheological and chemical contrast calculations are presented in the following maps. Both maps are symbolized using a 'stretched min/max' transformed colour ramp from ArcGIS ProTM. A geologic map for each map area is also provided to give context.

SM Figure 7-4.1 & 7-4.2 (the next two pages). A) Geological context of the Timmins and Dryden areas; B) Litho-tectonic assemblage maps (after Montsion et al., 2021a/Chapter 3) with supporting eruption and maximum deposition geochronological ages (Meek et al., 2020); C) Coloured Reduced to Pole (RTP) magnetic intensity aeromagnetic data filtered by cosine roll-off superimposed on hill shaded first-vertical derivative RTP (from Montsion et al., 2021a/Chapter 3); D) Calculated structural complexity map using segmented aeromagnetic lineaments (from Montsion et al., 2021b/Chapter 6). Intrusions are assigned an average complexity; E) Map of chemical contrast indicated by the sum of squared standard deviation (i.e., variance) in 1 km circular moving window for each element ratio map (SM 3 – Figures 1 and 2); F) Map of

rheological contrast indicated by the sum of squared standard deviation (i.e., variance) in 1 km circular moving window for each ductile and brittle property map
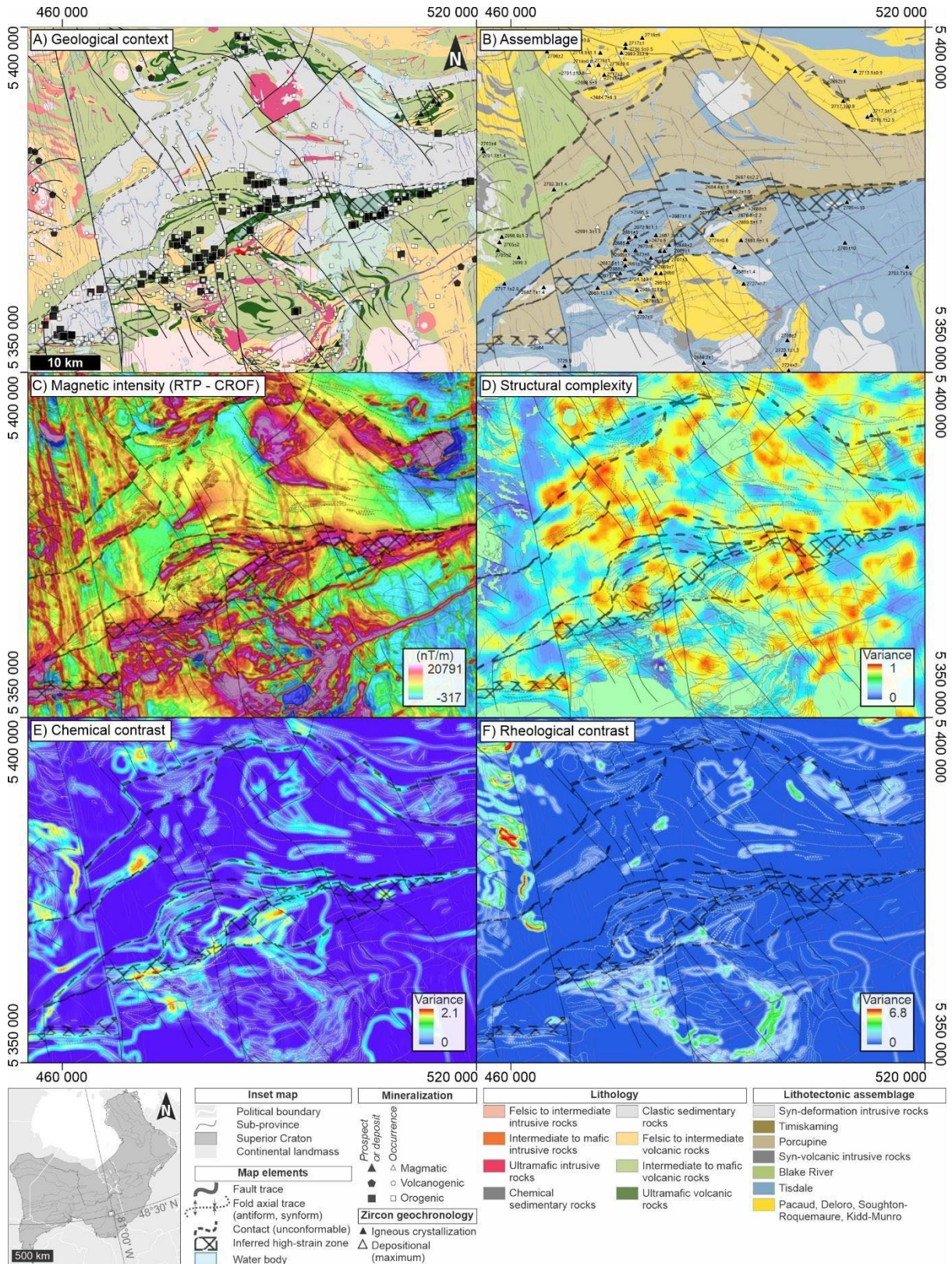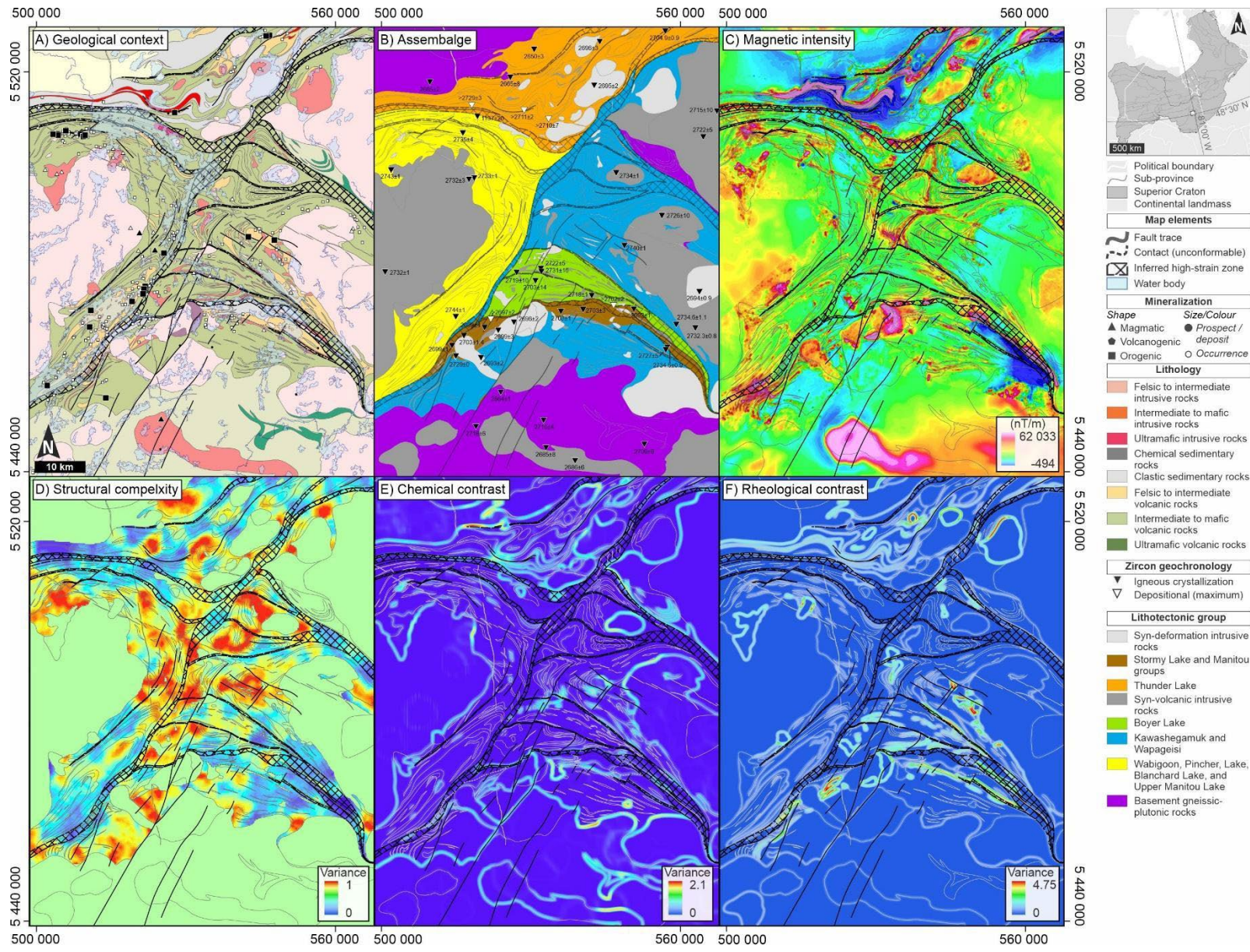
Supplementary material 7-5: Interpolated mass balance and alteration

Method

Calculation of mobile element gain/loss

In this strategy, a suite of less altered samples is identified (<2% LOI; SM6 – Figure 1), their primary composition (ultramafic to felsic) was characterized geochemically according to mafic compatibility (Mg/Zr for igneous samples and Al_2O_3 for sedimentary samples), and average mobile element concentrations were identified by fitting a linear function. Since least altered samples from all lithologies were not available for this regional scale study, an extrapolation from a linear function that approximates the global average of less altered samples was used. All samples with < 2.0 LOI % were considered 'less altered' and a mass loss/gain factor was determined by taking the difference between CO_2, SiO_2, K_2O, Na_2O concentrations and approximated linear functions. These functions were constrained by trends in immobile (igneous: log(Mg/Zr); sedimentary: A_2O_3 wt%) vs mobile elements representing carbonatization (CO_2), silicification (SiO_2), albitization (Na_2O) and/or sericitization (K_2O) (cross-plots with less altered samples (SM6 – Figure 2). A mass gain/loss factor of 1 implies no change, 0 implies 100% loss, and 2 implies 100% gain in the mobile element.

Samples from banded iron formation and a couple extreme outliers (described as deformed, altered, or metamorphosed) were removed from the low LOI population and their elemental concentrations do not affect linear trend functions.
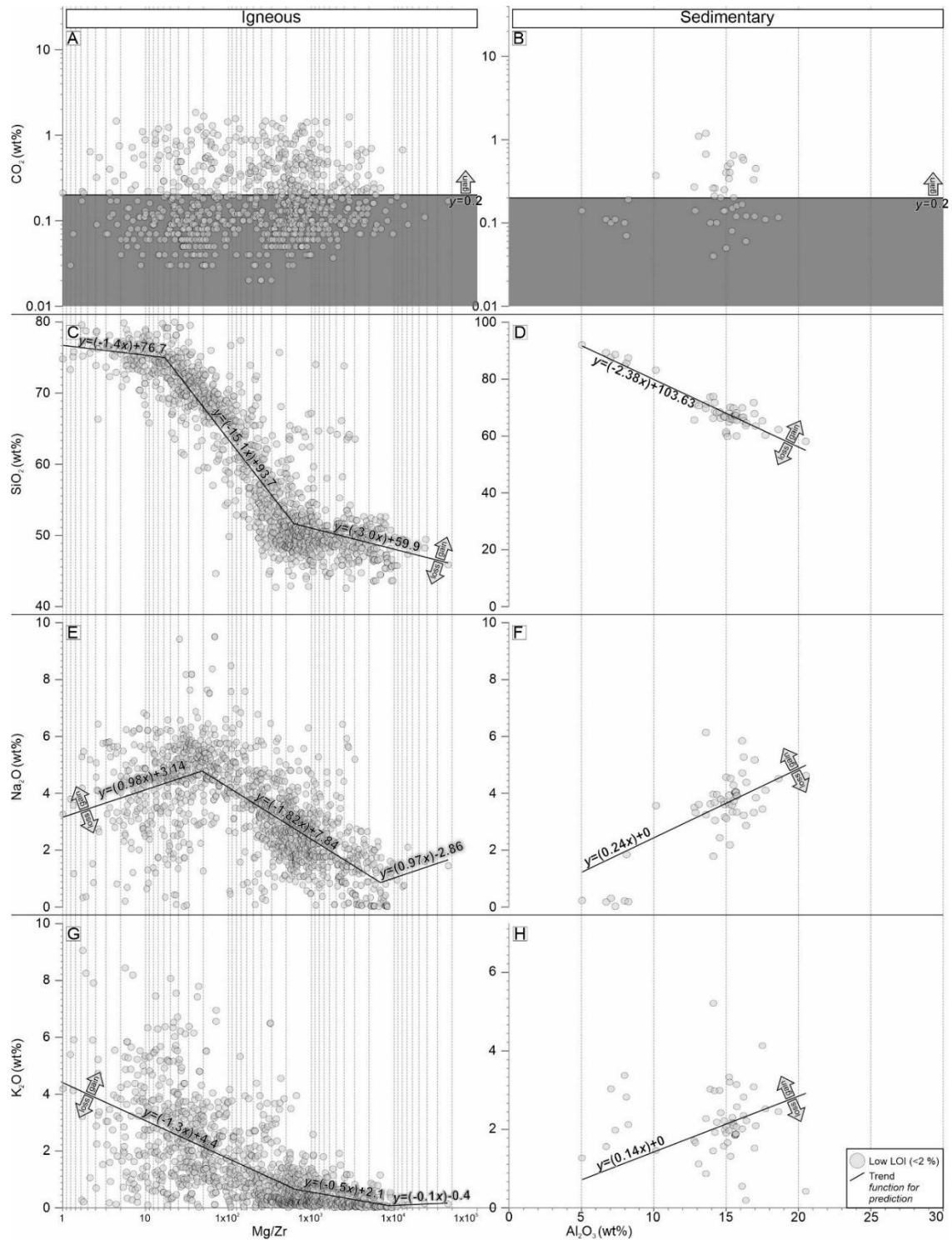
SM Figure 7-5.1. Mobile element concentrations such as A/B) CO_2, C/D) SiO_2, E/F) K_2O, G/H) Na_2O of samples with low LOI (<2%) are plotted against a lithology normalizing variable (Mg/Zr for igneous samples and Al_2O_3 for sedimentary samples) and a segmented linear trendline is fitted along trends. The linear functions are used to estimate/predict a 'least altered mobile element concentration for the full compositional range of samples.
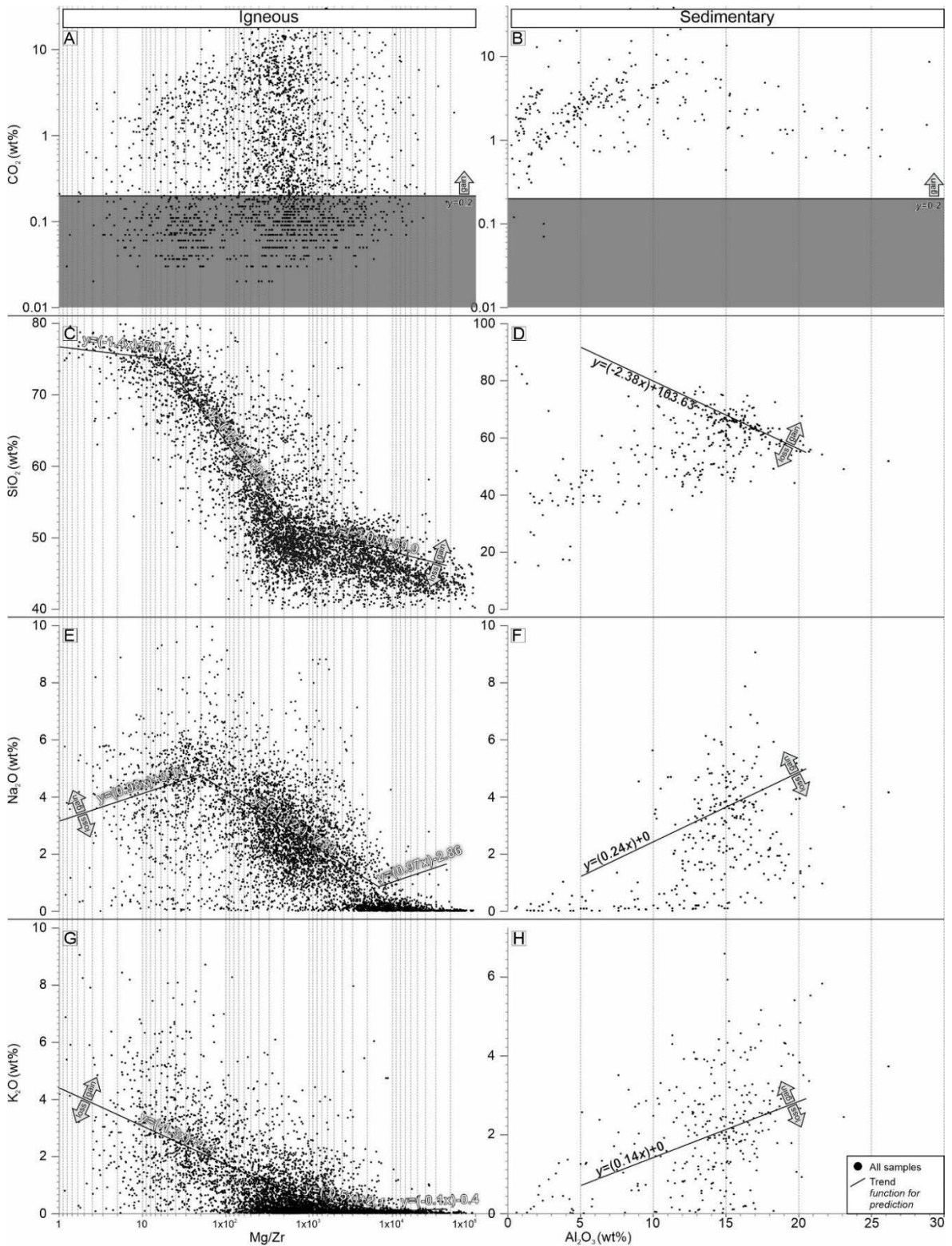
SM Figure 7-5.2. Mobile element concentrations such as A/B) CO_2, C/D) SiO_2, E/F) Na_2O, G/H) K_2O of all samples are plotted against a lithology normalizing variable (Mg/Zr for igneous samples and Al_2O_3 for sedimentary samples) and a segmented linear trendline is fitted along trends. The linear functions are used to estimate/predict a 'least altered mobile element concentration for the full compositional range of samples.

Interpolation

Empirical Bayesian Kriging (EBK) was used to interpolate calculated mobile element anomalies using a 15 km buffer zone to generate a spatially continuous representation to minimize edge effects. Hyperparameters for EBK are 10000 power variogram models performed on 250 m cells. Other hyperparameters such as variogram extent (radius of neighbourhood), the number of points randomly sub-selected from the neighbourhood, amount of allowed overlap (duplication) between sub-selected samples, and a smoothing factor vary for each element and location (reported in SM Table 7-5.1).

SM Table 7-5.1. Summary of hyperparameters used to generate Empirical Bayesian Kriging interpolation for LOI and each of the mobile elements' loss/gain factors

Oxide	Variogram model	# Points in each local model	Local model area overlap	Smoothing factor	Radius
LOI	Power	TM: 200	TM: 0.25	TM: 0.25	TM: 7 000
		DY: 100	DY: 0.25	DY: 0.75	DY: 15 000
CO_2		TM: 50	TM: 0.25	TM: 0.75	TM: 17 000
		DY: 200	DY: 5	DY: 0.25	DY: 20 000
SiO_2		TM: 500	TM: 0.25	TM: 0.75	TM: 6 000
		DY: 200	DY: 1	DY: 0.25	DY: 15 000
Na_2O		TM: 500	TM: 0.25	TM: 0.75	TM: 6 000
		DY: 200	DY: 1	DY: 0.25	DY: 15 000
K_2O		TM: 500	TM: 0.25	TM: 0.75	TM: 7 000
		DY: 200	DY: 1	DY: 0.25	DY: 15 000

SM Figure 7-5.3 & 7-5.4 (next two pages). Interpolated map results for B) LOI and mobile element C) CO_2, D) SiO_2, E) Na_2O, and F) K_2O gain/loss factor maps for 3) Timmins and 4) Dryden with superimposed geology and mineral deposits/occurrences for context. Distribution of samples used to constrain interpolation and grid results are compared in the legend to assess reprehensively of geochemical trends in data captured by the interpolation.

A) Geological context

10 km

B) LOI

C) CO₂

No data

No data

D) SiO₂

E) Na₂O

F) K₂O

Map elements

Geochemical sample
Fault trace
Contact (unconformable)
Inferred high-strain zone
Water body

Mineralization

Prospect or deposit

Occurrence

▲ △ Magmatic
⬟ ⬠ Volcanogenic
■ □ Orogenic
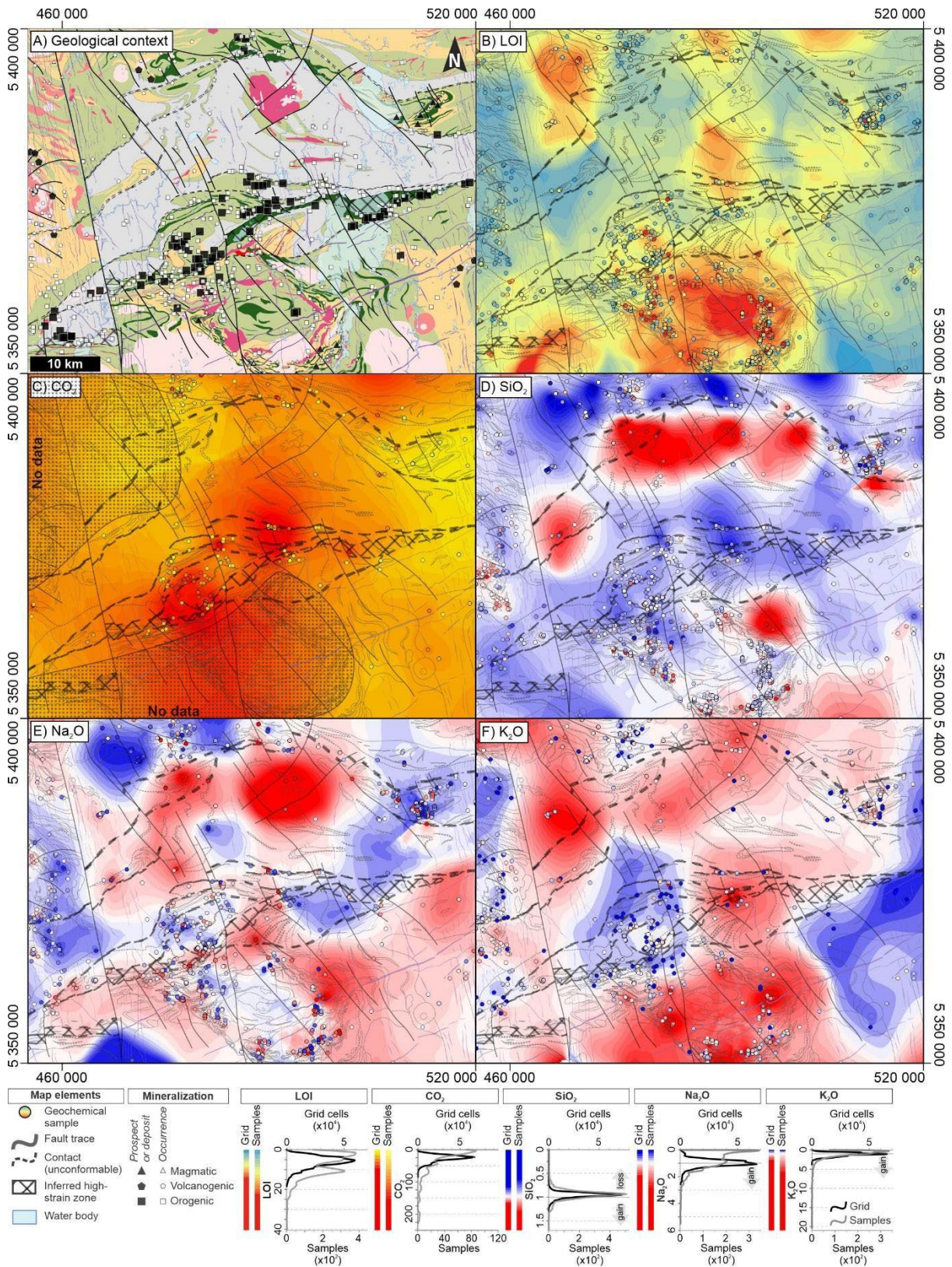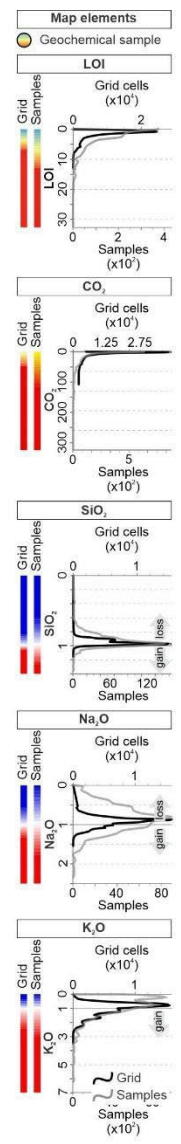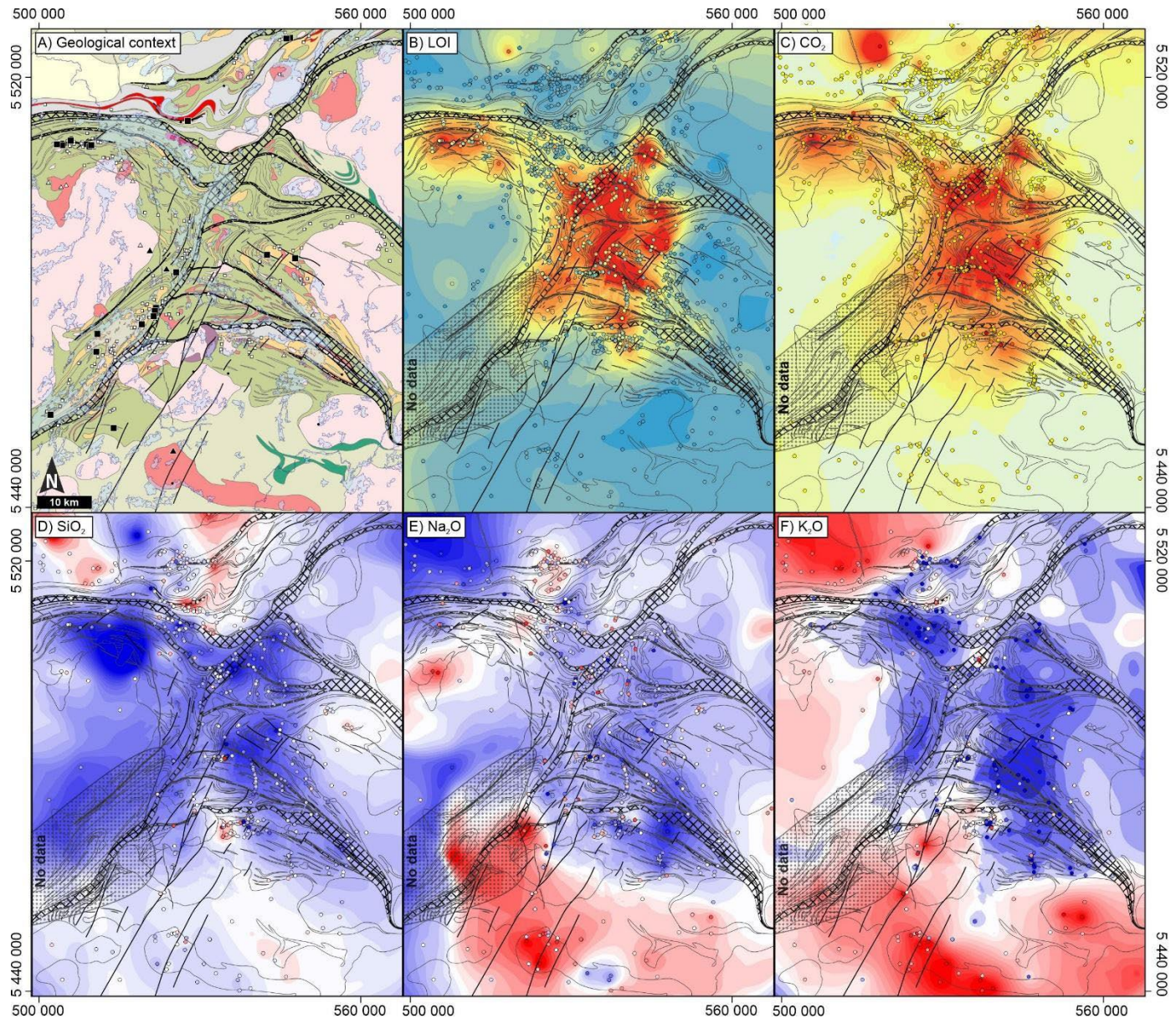
Supplementary material 7-6: Explanatory variable correlation matrix

SM- Table 7-6.1. Correlation matrix of all explanatory variables for both map areas indicate most are uncorrelated and are valid for use in machine learning/classification applications

	Mag. Intensity	Structural complexity	LOI	CO₂	SiO₂	Na₂O	K₂O	Mag Sus.	Porosity	Density	Resistivity	Rheological	Chemical	Mg/Zr	Ce/Y	Th/Nb	Eu/Eu*	Al/Ti	Basement	Para	Ultramafic (Intr.)	Mafic (Intr.)	TTG to felsic	Ultramafic (Volc.)	Mafic (Volc.)	Felsic	Clastic	Chemical (Sed.)	E-W (crustal)	E-W (splay)	NE-SW (crustal)
Mag. Intensity		-0.07	0.11	0.07	0.02	0.08	0.09	0.12	-0.06	0.21	0.04	0.14	0.07	0.14	-0.10	0.08	0.03	-0.08	0.04	0.05	0.04	0.01	0.00	0.04	0.01	0.00	0.00	-0.02	0.02	0.03	-0.04
Structural complexity	-0.07		0.08	0.08	-0.05	0.06	-0.01	-0.05	-0.01	0.02	0.01	0.11	0.08	0.03	-0.03	-0.03	0.03	-0.03	0.06	0.02	0.06	0.01	0.03	-0.11	-0.02	0.01	0.01	0.00	0.00	-0.04	-0.03
LOI	0.11	0.08		0.83	-0.25	0.06	0.12	-0.08	0.02	0.13	0.40	0.22	0.12	0.44	0.50	0.47	0.16	0.33	0.18	0.61	0.60	0.00	0.09	0.24	-0.18	-0.27	0.13	0.23	0.03	0.20	0.52
CO₂	0.07	0.08	0.83		-0.27	0.04	-0.03	0.07	0.05	0.15	0.41	0.22	0.13	0.36	0.48	0.44	0.15	0.33	0.14	0.59	0.58	-0.04	0.10	0.30	-0.19	-0.27	0.15	0.29	0.11	0.24	0.49
SiO₂	0.02	-0.05	-0.25	-0.27		0.44	0.30	-0.14	-0.26	0.02	-0.19	0.01	0.00	-0.15	0.15	0.19	0.02	0.20	0.27	0.03	0.12	0.10	0.11	0.14	0.11	0.15	0.06	0.10	0.19	0.21	0.09
Na₂O	0.08	0.06	0.06	0.04	0.44		0.17	-0.06	-0.16	0.03	-0.03	0.06	0.04	0.12	0.06	0.09	0.14	0.00	0.16	0.03	0.05	0.06	-0.12	0.13	-0.04	0.01	0.03	0.12	0.10	0.11	-0.12
K₂O	0.09	-0.01	0.12	-0.03	0.30	0.17		-0.01	-0.26	0.07	-0.19	0.01	0.00	-0.08	0.21	0.24	-0.12	0.18	0.18	0.00	0.12	0.21	0.28	0.07	0.25	0.28	0.14	0.24	0.33	0.38	0.13
Mag Sus.	0.12	-0.05	-0.08	0.07	-0.14	-0.06	-0.01		0.47	0.54	0.26	0.06	0.01	0.35	0.18	0.30	0.17	0.21	0.10	0.21	0.18	0.00	-0.03	0.14	0.04	0.09	0.23	0.12	0.11	0.17	0.15
Porosity	-0.06	-0.01	0.02	0.05	-0.26	-0.16	-0.26	0.47		0.27	0.64	-0.10	-0.05	0.31	-0.30	0.43	0.10	-0.40	0.18	0.11	0.04	-0.06	-0.09	0.10	0.09	0.10	0.07	-0.02	0.07	-0.06	0.01
Density	0.21	0.02	0.13	0.15	0.02	0.03	0.07	0.54	0.27		0.55	0.13	0.08	0.49	0.42	0.38	0.06	0.40	0.19	0.10	0.09	0.02	0.29	0.05	0.09	0.01	0.11	0.08	0.05	0.02	0.08
Resistivity	0.04	0.01	0.40	0.41	-0.19	-0.03	-0.19	0.26	0.64	0.55		0.15	0.10	0.66	0.71	0.72	0.08	0.68	0.11	0.31	0.35	0.12	0.11	0.12	-0.28	-0.32	0.23	0.21	0.20	0.26	0.36
Rheological	0.14	0.11	0.22	0.22	-0.01	0.06	0.01	0.06	-0.10	0.13	0.15		0.48	0.24	0.27	0.21	0.09	0.21	0.08	0.22	0.21	0.08	0.05	0.23	0.10	0.16	0.08	0.08	0.01	0.06	0.20
Chemical	0.07	0.08	0.12	0.13	0.00	0.04	0.00	0.01	-0.05	0.08	0.10	0.48		0.11	0.14	0.12	0.13	0.07	0.11	0.15	0.14	0.03	0.00	0.10	0.07	0.08	0.10	0.11	0.05	0.09	0.12
Mg/Zr	0.14	0.03	0.44	0.36	-0.15	0.12	-0.08	0.35	0.31	0.49	0.66	0.24	0.11		0.75	0.70	0.06	0.69	0.08	0.30	0.34	0.10	0.05	0.10	0.21	0.25	0.14	0.13	0.08	0.17	0.34
Ce/Y	-0.10	-0.03	0.50	0.48	0.15	0.06	0.21	0.18	-0.30	0.42	0.71	0.27	0.14	0.75		0.84	0.13	0.72	0.14	0.45	0.46	0.11	-0.08	0.22	0.27	0.33	0.24	0.23	0.17	0.31	0.44
Th/Nb	0.08	-0.03	0.47	0.44	0.19	0.09	0.24	0.30	0.43	0.38	0.72	0.21	0.12	0.70	0.84		0.00	0.71	0.10	0.35	0.40	0.16	0.04	0.22	0.27	0.33	0.18	0.14	0.11	0.25	0.42
Eu/Eu*	0.03	0.03	-0.16	-0.15	0.02	0.14	-0.12	0.17	0.10	0.06	0.08	0.09	-0.13	0.06	0.13	0.00		0.01	0.10	0.35	0.25	-0.09	-0.21	0.21	0.02	0.08	0.11	0.06	0.04	0.13	0.17
Al/Ti	-0.08	-0.03	0.33	0.33	0.20	0.00	0.18	0.21	-0.40	0.40	0.68	0.21	0.07	0.69	0.72	0.71	0.01		0.00	0.23	0.31	0.16	-0.01	0.13	0.24	0.27	0.17	0.14	0.11	0.24	0.31
Basement	0.04	0.06	0.18	0.14	0.27	0.16	0.18	0.10	0.18	0.19	0.11	0.08	0.11	0.08	0.14	0.10	0.10	0.00		0.46	0.44	0.46	-0.23	0.23	0.59	0.51	0.57	0.42	0.37	0.37	0.51
Para	0.05	0.02	0.61	0.59	0.03	0.03	0.00	0.21	0.11	0.10	0.31	0.22	0.15	0.30	0.45	0.35	0.35	0.23	0.46		0.92	0.19	-0.32	0.41	0.45	0.53	0.47	0.43	0.28	0.49	0.76
Ultramafic (Intr.)	0.04	0.06	0.60	0.58	0.12	0.05	0.12	0.18	0.04	0.09	0.35	0.21	0.14	0.34	0.46	0.40	0.25	0.31	0.44	0.92		0.41	-0.10	0.29	0.65	0.70	0.60	0.56	0.45	0.65	0.88
Mafic (Intr.)	0.01	0.01	0.00	-0.04	0.10	0.06	0.21	0.00	-0.06	0.02	0.12	0.08	0.03	0.10	0.11	0.16	-0.09	0.16	0.46	0.19	0.41		0.30	-0.08	0.86	0.79	0.65	0.50	0.59	0.62	0.57
TTG to felsic	0.00	0.03	0.09	0.10	0.11	-0.12	0.28	-0.03	-0.09	0.29	0.11	0.05	0.00	0.05	-0.08	0.04	-0.21	-0.01	0.23	-0.32	-0.10	0.30		0.12	0.22	0.23	0.02	0.01	0.20	0.11	0.07
Ultramafic (Volc.)	0.04	-0.11	0.24	0.30	0.14	0.13	0.07	0.14	0.10	0.05	0.12	0.23	0.10	0.10	0.22	0.22	0.21	0.13	0.23	0.41	0.29	-0.08	0.12		0.07	-0.15	0.00	0.15	0.23	0.02	-0.33
Mafic (Volc.)	0.01	0.02	-0.18	-0.19	0.11	0.04	0.25	0.04	0.09	0.09	-0.28	0.10	0.07	0.21	0.27	0.27	0.02	0.24	0.59	0.45	0.65	0.86	0.22	0.07		0.94	0.79	0.65	0.72	0.80	0.74
Felsic	0.00	0.01	-0.27	-0.27	0.15	0.01	0.28	0.09	0.10	0.01	-0.32	0.16	0.08	0.25	0.33	0.33	0.08	0.27	0.51	0.53	0.70	0.79	0.23	-0.15	0.94		0.76	0.63	0.69	0.80	0.79
Clastic	0.00	-0.01	0.13	0.15	0.06	0.03	0.14	0.23	0.07	0.11	0.23	0.08	0.10	0.14	0.24	0.18	0.11	0.17	0.57	0.47	0.60	0.65	0.02	0.00	0.79	0.76		0.70	0.82	0.81	0.66
Chemical (Sed.)	0.02	0.00	0.23	0.29	0.10	0.12	0.24	0.12	-0.02	0.08	0.21	0.08	0.11	0.13	0.23	0.14	0.06	0.14	0.42	0.43	0.56	0.50	0.01	0.15	0.65	0.63	0.70		0.82	0.76	0.42

Def. zone																																
E-W (crustal)	0.02	0.00	-0.03	-0.11	0.19	0.10	0.33	0.11	-0.07	-0.05	-0.20	0.01	-0.05	-0.08	0.17	0.11	0.04	0.11	0.37	0.28	0.45	0.59	0.20	0.23	0.72	0.69	0.82	0.82		0.78	0.49
E-W (splay)	0.03	-0.04	-0.20	-0.24	0.21	0.11	0.38	0.17	-0.06	-0.02	-0.26	0.06	-0.09	-0.17	0.31	0.25	0.13	0.24	0.37	0.49	0.65	0.62	0.11	0.02	0.80	0.80	0.81	0.76	0.78		0.59
NE-SW (crustal)	-0.04	-0.03	-0.52	-0.49	0.09	-0.12	0.13	0.15	0.01	-0.08	-0.36	-0.20	-0.12	-0.34	0.44	0.42	0.17	0.31	0.51	0.76	0.88	0.57	0.07	-0.33	0.74	0.79	0.66	0.42	0.49	0.59	

Supplementary material 7-7: Hyperparameters tables

SM Table 7-7.1. Summary of hyperparameters used to train all random forest models

System	Training set	# Trees	Min leaf size	Max tree depth	Mean tree depth	Data available per tree (%)	# Sampled variables	% Data for validation	# Validation runs
					Timmins				
Magmatic	1	5000	5	10	6	70	10	20	100
	2	5000	5	10	6	70	10	20	100
	3	5000	5	10	6	70	10	20	100
	4	5000	5	10	6	70	10	20	100
	5	5000	5	10	6	70	10	20	100
Volcanogenic	1	5000	5	10	9	100	10	20	100
	2	5000	5	10	9	100	10	20	100
	3	5000	5	10	9	100	10	20	100
	4	5000	5	10	9	100	10	20	100
	5	5000	5	10	9	100	10	20	100
Orogenic	1	5000	20	15	14	70	10	20	100
	2	5000	20	15	14	70	10	20	100
	3	5000	20	15	14	70	10	20	100
	4	5000	20	15	14	70	10	20	100
	5	5000	20	15	14	70	10	20	100
	1	5000	1	16	15	100	18	20	100
	2	5000	1	16	15	100	20	20	100
	3	5000	1	15	14	100	20	20	100
	4	5000	1	15	14	100	22	20	100
	5	5000	1	15	14	100	22	20	100

System	Training set	# Trees	Min leaf size	Max tree depth	Mean tree depth	Data available per tree (%)	# Sampled variables	% Data for validation	# Validation runs
					Dryden				
Orogenic	1	5000	5	15	14	70	11	20	100
	2	5000	5	15	14	70	11	20	100
	3	5000	5	15	14	70	11	20	100
	4	5000	5	15	14	70	11	20	100
	5	5000	5	15	14	70	11	20	100

Supplementary material 7-8: Performance metrics tables

Model performance metric (Out of Bag, MCC, F-Score)

*SM Table 7-8.1. Summary of performance metrics reported by random forests for each system model and training set. Blue cells indicate best results and red indicate poor performance of random forest models. Performance metrics are described in section 7.3.1.2. *S = Sensitivity; *A = Accuracy; 0 = Background; 1 = Mineralized target*

| | | TIMMINS | | | | | | | | | | | | | | | DRYDEN | | | | |
| | | Magmatic | | | | | VMS | | | | | Orogenic | | | | | | | | | |
Training set		1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Out of bag	MSE	14.93	12.89	14.32	11.08	9.60	28.49	32.79	22.38	24.14	28.12	14.51	13.48	12.63	12.08	17.79	19.32	19.47	19.99	16.86	15.65
	0	10.08	9.27	11.54	6.18	7.42	23.13	37.97	22.18	24.91	24.22	20.04	12.09	11.38	8.72	25.51	22.25	23.73	21.53	17.76	18.78
	1	19.89	16.79	16.89	16.23	11.91	34.12	28.08	22.58	23.37	32.09	10.40	14.75	13.84	15.40	10.71	16.35	15.30	18.44	15.98	12.77
Training	FI - 0	0.98	0.97	0.99	0.97	0.96	0.98	0.99	1.00	0.99	0.99	0.97	0.98	0.97	0.96	0.90	0.94	0.97	0.98	0.95	0.95
	MCC - 0	0.97	0.94	0.98	0.94	0.92	0.96	0.97	1.00	0.97	0.99	0.95	0.95	0.93	0.92	0.82	0.89	0.93	0.96	0.91	0.91
	*S	1.00	1.00	1.00	0.98	1.00	0.97	0.99	1.00	0.99	1.00	0.97	1.00	0.98	0.97	0.84	0.92	0.96	0.98	0.94	0.91
	*A	0.98	0.97	0.99	0.97	0.96	0.98	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.96	0.91	0.95	0.97	0.98	0.95	0.95
	FI - 1	0.98	0.97	0.99	0.97	0.96	0.98	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.96	0.92	0.95	0.97	0.98	0.96	0.96
	MCC - 1	0.97	0.94	0.98	0.94	0.92	0.96	0.97	1.00	0.97	0.99	0.95	0.95	0.93	0.92	0.82	0.89	0.93	0.96	0.91	0.91
	*S	0.97	0.94	0.98	0.95	0.92	0.99	0.99	1.00	0.99	0.99	0.98	0.96	0.95	0.95	0.97	0.98	0.98	0.98	0.97	0.99
	*A	0.98	0.97	0.99	0.97	0.96	0.98	0.99	1.00	0.99	0.99	0.98	0.98	0.97	0.96	0.91	0.95	0.97	0.98	0.95	0.95
Validation	FI - 0	0.88	0.90	0.88	0.90	0.90	0.71	0.73	0.80	0.77	0.74	0.81	0.87	0.88	0.87	0.86	0.79	0.82	0.79	0.82	0.85
	MCC - 0	0.78	0.83	0.75	0.81	0.83	0.47	0.44	0.51	0.53	0.47	0.70	0.74	0.76	0.74	0.70	0.62	0.66	0.63	0.66	0.66
	*S	1.00	1.00	0.82	0.93	1.00	0.71	0.65	0.82	0.75	0.74	0.89	0.85	0.93	0.89	0.85	0.84	0.79	0.76	0.85	0.87
	*A	0.88	0.91	0.87	0.91	0.91	0.74	0.71	0.76	0.76	0.74	0.85	0.87	0.88	0.87	0.85	0.81	0.83	0.82	0.83	0.83
	FI - 1	0.87	0.91	0.87	0.91	0.91	0.76	0.69	0.71	0.76	0.74	0.87	0.86	0.88	0.87	0.84	0.82	0.84	0.84	0.83	0.80
	MCC - 1	0.78	0.83	0.75	0.81	0.83	0.47	0.44	0.51	0.53	0.47	0.70	0.74	0.76	0.74	0.70	0.62	0.66	0.63	0.66	0.66
	*S	0.76	0.83	0.93	0.88	0.83	0.76	0.80	0.69	0.78	0.74	0.83	0.89	0.84	0.85	0.85	0.78	0.86	0.87	0.81	0.78
	*A	0.88	0.91	0.87	0.91	0.91	0.74	0.71	0.76	0.76	0.74	0.85	0.87	0.88	0.87	0.85	0.81	0.83	0.82	0.83	0.83

*S = Sensitivity; *A = Accuracy

0 = Background; 1 = Mineralized target

Confusion matrix

SM Table 7-8.2 (next page). Confusion matrices reported by random forests for each model as exact number of classified/misclassified training points or as a percentage. Red cells indicate poor performance where a high percentage of points were misclassified.

Timmins

Number of training points

Magmatic

		Set 1 Background	Set 1 Target	Set 2 Background	Set 2 Target	Set 3 Background	Set 3 Target	Set 4 Background	Set 4 Target	Set 5 Background	Set 5 Target
Set 1	Background	15	0								
Set 1	Target	4	13								
Set 2	Background			14	0						
Set 2	Target			3	15						
Set 3	Background					14	3				
Set 3	Target					1	13				
Set 4	Background							14	1		
Set 4	Target							2	15		
Set 5	Background									14	0
Set 5	Target									3	15

Volcanogenic

		Set 1 Background	Set 1 Target	Set 2 Background	Set 2 Target	Set 3 Background	Set 3 Target	Set 4 Background	Set 4 Target	Set 5 Background	Set 5 Target
Set 1	Background	12	5								
Set 1	Target	5	16								
Set 2	Background			15	8						
Set 2	Target			3	12						
Set 3	Background					18	4				
Set 3	Target					5	11				
Set 4	Background							15	5		
Set 4	Target							4	14		
Set 5	Background									14	5
Set 5	Target									5	14

Orogenic

		Set 1 Background	Set 1 Target	Set 2 Background	Set 2 Target	Set 3 Background	Set 3 Target	Set 4 Background	Set 4 Target	Set 5 Background	Set 5 Target
Set 1	Background	33	4								
Set 1	Target	11	52								
Set 2	Background			30	11						
Set 2	Target			8	51						
Set 3	Background					50	4				
Set 3	Target					10	51				
Set 4	Background							51	6		
Set 4	Target							9	51		
Set 5	Background									52	9
Set 5	Target									8	46

% of training points

Magmatic

		Set 1 Background	Set 1 Target	Set 2 Background	Set 2 Target	Set 3 Background	Set 3 Target	Set 4 Background	Set 4 Target	Set 5 Background	Set 5 Target
Set 1	Background	100	0								
Set 1	Target	24	76								
Set 2	Background			100	0						
Set 2	Target			17	83						
Set 3	Background					82	18				
Set 3	Target					7	93				
Set 4	Background							93	7		
Set 4	Target							12	88		
Set 5	Background									100	0
Set 5	Target									17	83

Volcanogenic

		Set 1 Background	Set 1 Target	Set 2 Background	Set 2 Target	Set 3 Background	Set 3 Target	Set 4 Background	Set 4 Target	Set 5 Background	Set 5 Target
Set 1	Background	71	29								
Set 1	Target	24	76								
Set 2	Background			65	35						
Set 2	Target			20	80						
Set 3	Background					82	18				
Set 3	Target					31	69				
Set 4	Background							75	25		
Set 4	Target							22	78		
Set 5	Background									74	26
Set 5	Target									26	74

Orogenic

		Set 1 Background	Set 1 Target	Set 2 Background	Set 2 Target	Set 3 Background	Set 3 Target	Set 4 Background	Set 4 Target	Set 5 Background	Set 5 Target
Set 1	Background	89	11								
Set 1	Target	17	83								
Set 2	Background			73	27						
Set 2	Target			14	86						
Set 3	Background					93	7				
Set 3	Target					16	84				
Set 4	Background							89	11		
Set 4	Target							15	85		
Set 5	Background									85	15
Set 5	Target									15	85

Dryden

Number of training points

Orogenic

		Set 1 Background	Set 1 Target	Set 2 Background	Set 2 Target	Set 3 Background	Set 3 Target	Set 4 Background	Set 4 Target	Set 5 Background	Set 5 Target
Set 1	Background	31	6								
Set 1	Target	10	36								
Set 2	Background			31	8						
Set 2	Target			6	37						
Set 3	Background					28	9				
Set 3	Target					6	39				
Set 4	Background							33	6		
Set 4	Target							8	35		
Set 5	Background									41	6
Set 5	Target									8	28

% of training points

Orogenic

		Set 1 Background	Set 1 Target	Set 2 Background	Set 2 Target	Set 3 Background	Set 3 Target	Set 4 Background	Set 4 Target	Set 5 Background	Set 5 Target
Set 1	Background	84	16								
Set 1	Target	22	78								
Set 2	Background			79	21						
Set 2	Target			14	86						
Set 3	Background					76	24				
Set 3	Target					13	87				
Set 4	Background							85	15		
Set 4	Target							19	81		
Set 5	Background									87	13
Set 5	Target									22	78

Supplementary material 7-9: Variable importance distributions

SM Figure 7-9.1. Prospective signatures (i.e., proximal to known mineralization) in top ranked variables for each mineral system in order of descending importance. Pie charts for assemblages hosting mineralization (A, E, I, M) are arranged from oldest to youngest in an anti-clockwise fashion and indicate mineralization is mainly hosted by volcanic assemblages. Box and whisker plots of quantitative variables indicate criteria when using rheological contrast (B), distance to clastic sedimentary rocks (C), and Mg/Zr (D) when exploring for magmatic targets. Criteria for using Eu/Eu* (F), distance to mafic volcanic rocks (G), and CO$_2$ (H) also presented for volcanogenic targets. Finally, criteria and signatures related to distance to E-W trans-crustal deformation zones (J,N), LOI, and distance to TTG – felsic intrusions (L, P) are compared for orogenic mineralization near Timmins and Dryden.
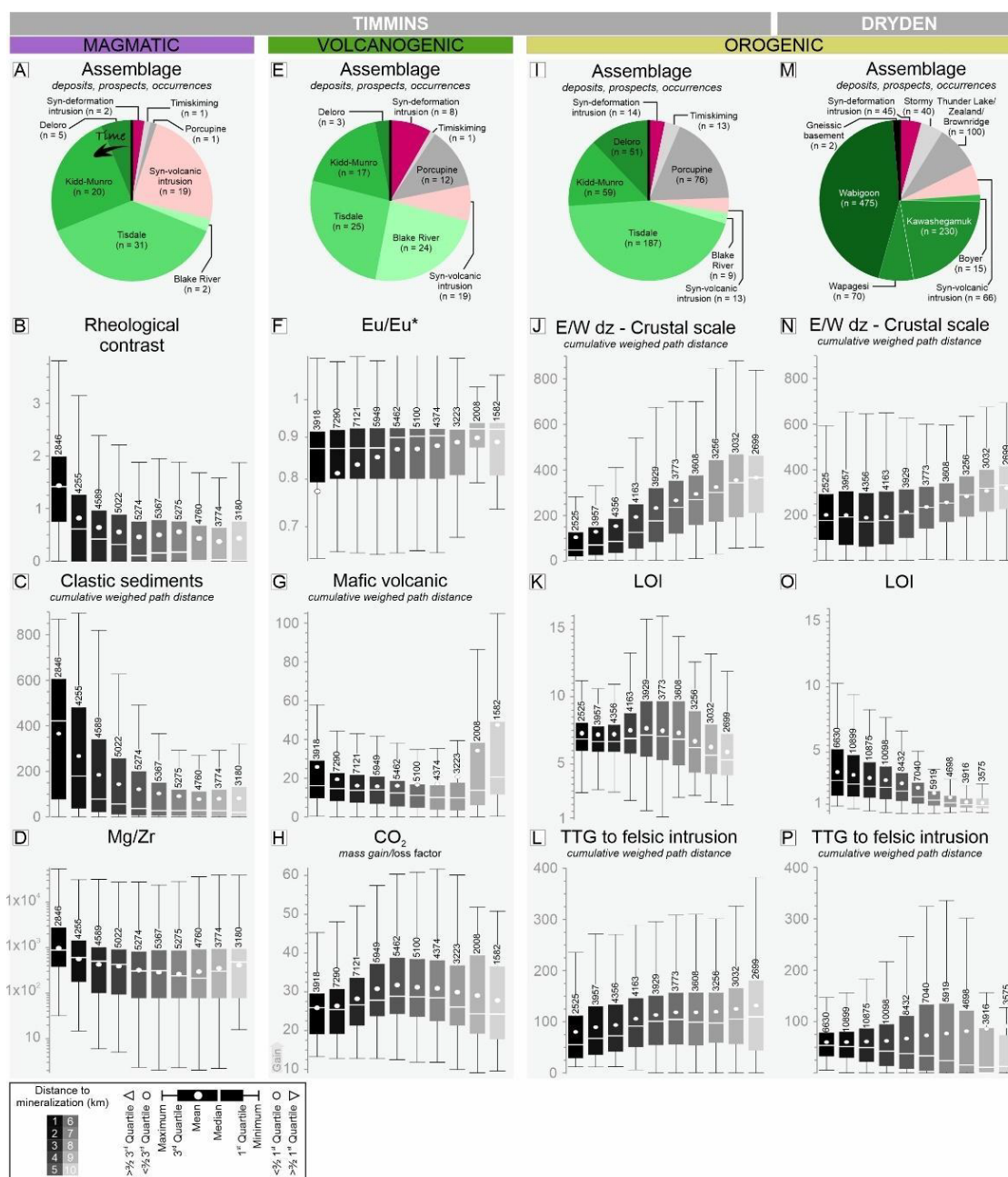
SM Figure 7-9.2 (next ten pages). Box and whisker plots representing distribution of all explanatory variable values for samples proximal to known mineralization. Similar to SM 10 Figure 1, each box/whisker unit captures the distribution of grid cells within 1 km increments are areas around targets.

TIMMINS - MAGMATIC

K) Log₁₀(Mg/Zr)
L) Log₁₀(Ce/Y)
M) Log₁₀(Th/Nb)
N) Log₁₀(Eu/Eu*)
O) Log₁₀(Al/Ti)
P) Chemical contrast
Q) LOI
R) CO₂ (gain/loss)
S) SiO₂ (gain/loss)
T) Na₂O (gain/loss)
U) K₂O (gain/loss)
V) Magnetic intensity (nT/m)
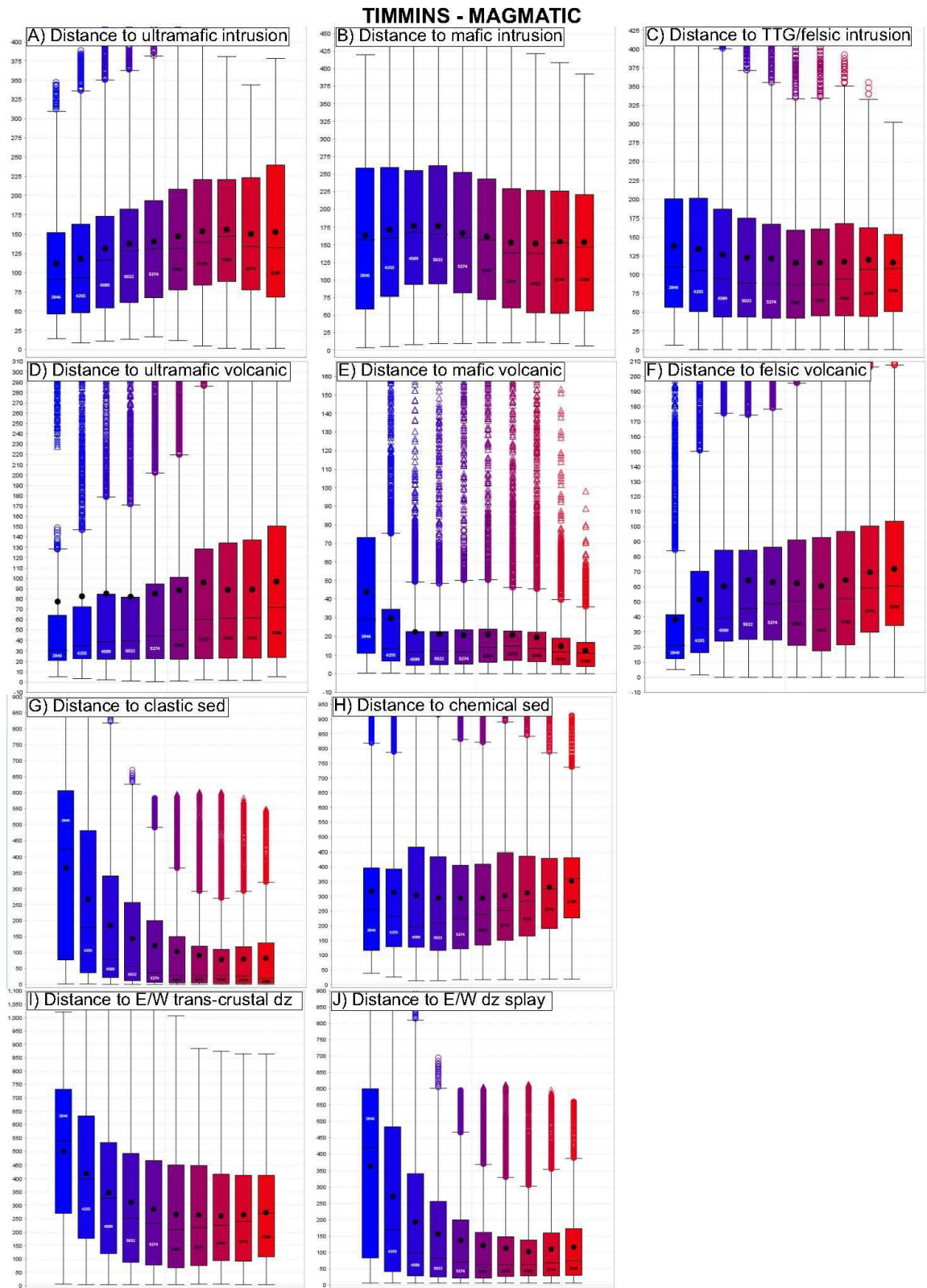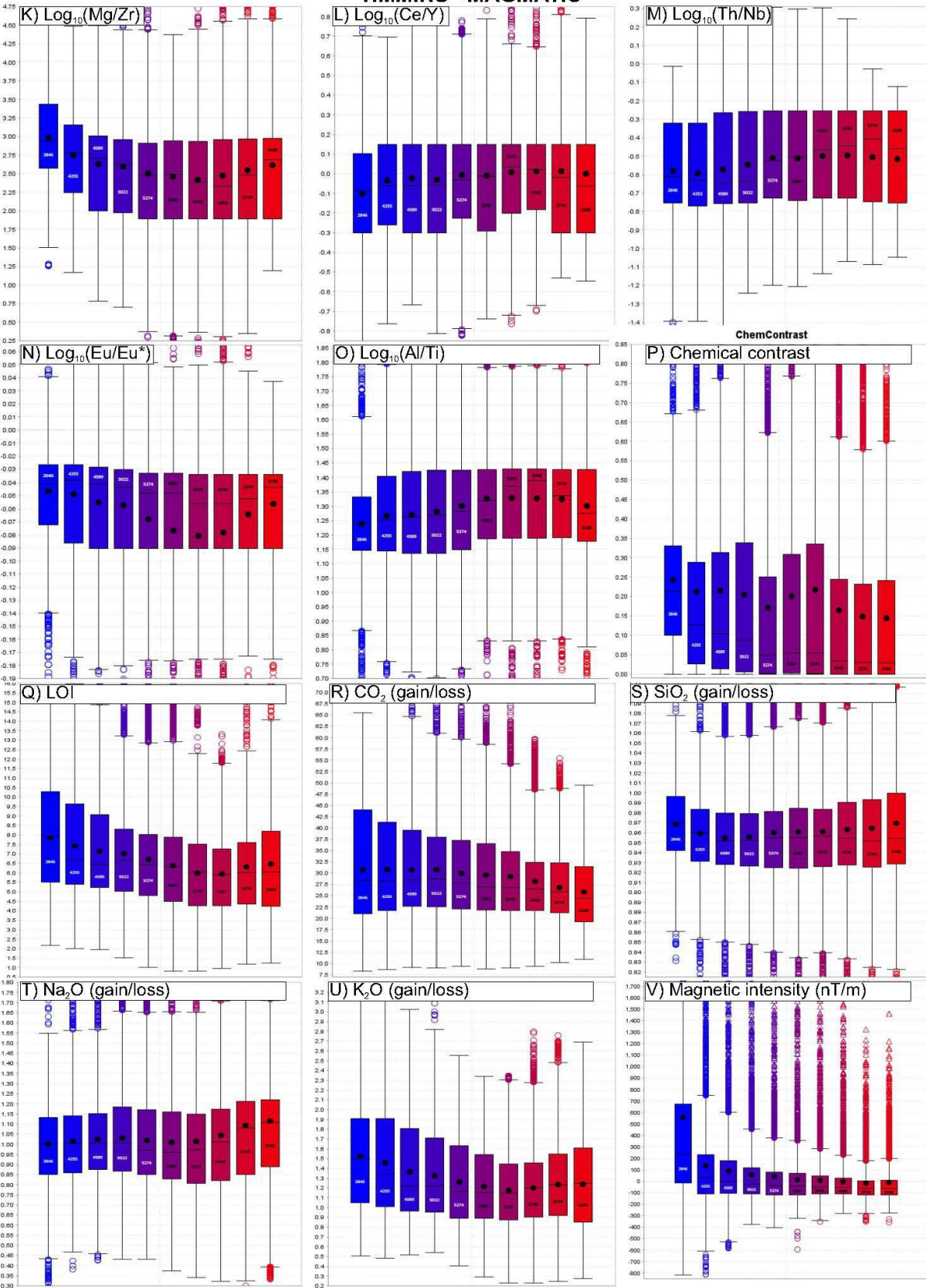
W) Log$_{10}$(Magnetic susceptibility)
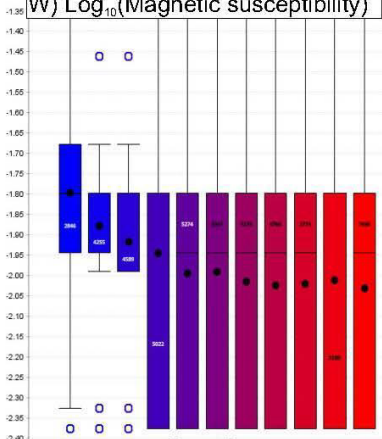
X) Total porosity
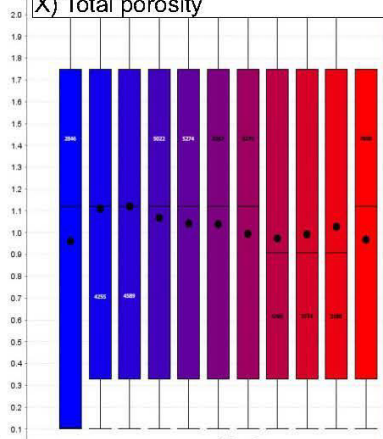
Y) Rheological contrast
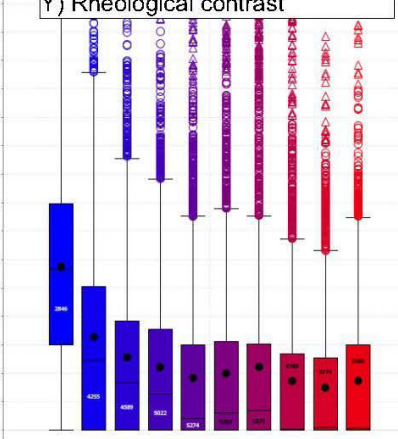
Z) Grain density
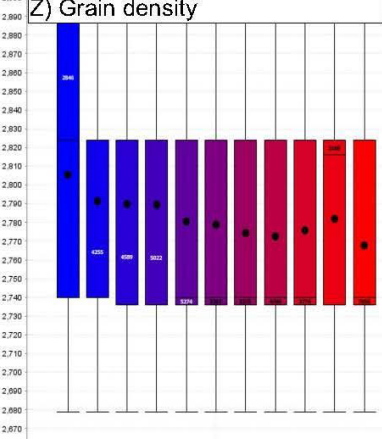
AA) Log$_{10}$(Resisitivity)

AB) Structural complexity

AC) Distance to ultramafic intrusion

AD) Distance to mafic intrusion

AE) Distance to TTG/felsic intrusion

AF) Distance to ultramafic volcanic

AG) Distance to mafic volcanic

AH) Distance to felsic volcanic

AI) Distance to clastic sed

AJ) Distance to chemical sed

AK) Distance to E/W trans-crustal dz

AL) Distance to E/W dz splay

AM) $Log_{10}(Mg/Zr)$

AN) $Log_{10}(Ce/Y)$

AO) $Log_{10}(Th/Nb)$

AP) $Log_{10}(Eu/Eu^*)$

AQ) $Log_{10}(Al/Ti)$

AR) Chemical contrast

AS) LOI

AT) CO_2 (gain/loss)

AU) SiO_2 (gain/loss)

AV) Na_2O (gain/loss)

AW) K_2O (gain/loss)

AX) Magnetic intensity (nT/m)

AY) Log$_{10}$(Magnetic susceptibility)

AZ) Total porosity

BA) Rheological contrast

BB) Grain density

BC) Log$_{10}$(Resisitivity)

BD) Structural complexity

BE) Distance to ultramafic intrusion

BF) Distance to mafic intrusion

BG) Distance to TTG/felsic intrusion

BH) Distance to ultramafic volcanic

BI) Distance to mafic volcanic

BJ) Distance to felsic volcanic

BK) Distance to clastic sed

BL) Distance to chemical sed

BM) Distance to E/W trans-crustal dz

BN) Distance to E/W dz splay

BO) Log₁₀(Mg/Zr)

BP) Log₁₀(Ce/Y)

BQ) Log₁₀(Th/Nb)

BR) Log₁₀(Eu/Eu*)

BS) Log₁₀(Al/Ti)

BT) Chemical contrast

BU) LOI

BV) CO₂ (gain/loss)

BW) SiO₂ (gain/loss)

BX) Na₂O (gain/loss)

BY) K₂O (gain/loss)

BZ) Magnetic intensity (nT/m)

CH) Distance to basement gneiss

CI) Distance to para-gneiss

CJ)Distance to ultramafic intrusion

CK)Distance to mafic intrusion

CL) Distant to TTG/Felsic intrusion

CM) Distance to E/W trans-crustal dz

CN) Distance to mafic volcanic

CO) Distance to felsic volcanic

CP) Distance to E/W dz splay

CQ) Distance to clastic sed

CR Distance to chemical sed

CS) Distance to NE/SW trans-crustal dz

CT) Log₁₀(Mg/Zr)

CU) Log₁₀(Ce/Y)

CV) Log₁₀(Th/Nb)

CW) Log₁₀(Eu/Eu*)

CX) Log₁₀(Al/Ti)

CY) Chemical contrast

CZ) LOI

DA) CO₂ (gain/loss)

DB) SiO₂ (gain/loss)

DC) Na₂O (gain/loss)

DD) K₂O (gain/loss)

DE) Magnetic intensity (nT/m)

DF) Log$_{10}$(Magnetic susceptibility)

DG) Total porosity

DH) Rheological contrast

DI) Grain density

DJ) Log$_{10}$(Resisitivity)

DK) Structural complexity