LUIZA MARIA PEREIRA PIERANGELI

GREEN TECH PROXIMAL SENSOR ANALYSES: ADVANCES FOR SOIL, PARENT MATERIAL AND GEOLOGICAL EXPLORATION

LAVRAS – MG
2024

LUIZA MARIA PEREIRA PIERANGELI

GREEN TECH PROXIMAL SENSOR ANALYSES: ADVANCES FOR SOIL, PARENT MATERIAL AND GEOLOGICAL EXPLORATION

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós- Graduação em Ciência do Solo, área de concentração em Recursos Ambientais e Uso da Terra, para a obtenção do título de Doutor.

Prof. Dr. Sérgio Henrique Godinho Silva
Orientador

LAVRAS – MG
2024

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Pierangeli, Luiza Maria Pereira.

Greentech proximal sensor analyses : advances for soil, parent material and geologicalexploration / Luiza Maria Pereira Pierangeli. - 2024.

118 p.

Orientador(a): Sérgio Henrique Godinho Silva.

Tese (doutorado) - Universidade Federal de Lavras, 2024.

Bibliografia.

1. Geoquímica do solo. 2. Sensores próximos. 3. Aprendizado de máquinas. I. Silva, Sérgio Henrique Godinho. II. Título.

LUIZA MARIA PEREIRA PIERANGELI

GREEN TECH PROXIMAL SENSOR ANALYSES: ADVANCES FOR SOIL, PARENT MATERIAL AND GEOLOGICAL EXPLORATION

ANÁLISE "VERDE" COM SENSORES PRÓXIMOS: AVANÇOS EXPLORAÇÃO DE SOLO, MATERIAL DE ORIGEM E GEOLÓGICA

Tese apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós- Graduação em Ciência do Solo, área de concentração em Recursos Ambientais e Uso da Terra, para a obtenção do título de Doutor.

APROVADA em 27 de março de 2024.
Dr. Sérgio Henrique Godinho Silva - UFLA
Dr. Michele Duarte de Menezes - UFLA
Dr. Renata Andrade - UFLA
Dr. Milson Evaldo Serafim - IFMT
Dr. Eduardo Guimarães Couto - UFMT

Prof. Dr. Sérgio Henrique Godinho Silva
Orientador

LAVRAS – MG
2024

AGRADECIMENTOS

À minha família, Dedezinha, Dreds, Lele, Brubs e tia Re, pelo amor incondicional, apoio, e incentivo a sempre continuar lutando.

Aos meus amigos que sempre me ajudaram e me acompanharam.

À Universidade Federal de Lavras e, principalmente, ao Departamento de Ciência do Solo por contribuírem para a minha formação acadêmica e profissional. O presente trabalho foi realizado com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), e Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG).

Ao Prof. Sérgio Henrique Godinho Silva pela orientação, atenção, paciência, confiança, ensinamentos e incentivos.

Aos professores do DCS/UFLA, em especial prof. Michele Duarte de Menezes e prof. Nilton Curi e a prof. Renata Andrade, pelos ensinamentos.

Aos colegas do departamento e laboratório, Geila, Mari, Livia, Álvaro José, Anita, Fernanda Magno, Fernanda Bocoli, Marcelo e Cristina.

A todos os funcionários do DCS, técnicos, secretarias, faxineiras, vocês também são uma parte fundamental desta jornada.

Ao Dr. David Weindorf, Dra. Mona Sirbescu e ao corpo docente e aos funcionários do departamento de Earth and Atmospheric Sciences da Central Michigan University por me receberem por um ano durante meu doutorado.

Obrigada a todos que contribuíram de alguma forma para que eu chegasse aqui!

Muito Obrigado!

RESUMO

O material de origem do solo surge como um fator crítico na compreensão da variabilidade do solo, com desafios em avaliações detalhadas devido à complexidade e inacessibilidade ao material de origem em solos profundos. Abordagens de sensores proximais, incluindo a espectroscopia de fluorescência de raios X portátil (pXRF) e a suscetibilidade magnética (SM), oferecem soluções práticas para predizer material de origem. A busca global por alternativas energéticas sustentáveis aumentou paralelamente, com o lítio (Li) emergindo como um elemento fundamental integrante das baterias recarregáveis de íons de Li. A procura de Li exige o desenvolvimento de métodos de exploração econômicos e rápidos para melhorar a identificação de novos depósitos, com pegmatitos de lítio-césio-tântalo (LCT) como fontes primárias de Li. Esta tese está dividida em quatro artigos, englobando: (I) um estudo piloto centrado na criação de modelos preditivos espaciais de material de origem para três tipos distintos de rochas (charnockito, argilito e sedimentos aluviais) em escala detalhada, usando algoritmo random forest (RF) combinado com pXRF e SM; (II) avaliação da eficácia dos pXRF e do algoritmo RF na predição da concentração de Li em amostras de solo e predição do PM de solo rico em Li, utilizando elementos *pathfinder* de Li; (III) (IV) avaliação do desempenho e a comparabilidade de dois sistemas pXRF, Alfa e Beta. Os resultados desta tese propõem métodos alternativos e econômicos para obtenção da variabilidade espacial do material de origem do solo. O artigo I apresentou acurácia na predição de material de origem (coeficiente Kappa = 0,85 e acurácia geral = 0,93), já o modelo de predição de material de origem do artigo (II) alcançou coeficiente Kappa de 0,77 e uma acurácia global de 0,85. O modelo de predição de Li artigo II) apresentou coeficiente de determinação (R^2) de 0,86, erro quadrático médio (RMSE) de 68,5 mg kg^{-1} e desvio do resíduo de predição (RPD) de 1,78. O artigo (III) mostrou que ambos os pXRFs alcançaram resultados semelhantes em comparação com as concentrações elementares em materiais de referência certificados, com tendências de superestimação e subestimação da concentração de elementos. Os sensores próximos aliado às técnicas de aprendizado de máquinas consistem em método alternativo para prospecção mais sustentáveis de material de origem do solo e exploração de conteúdo de Li. Além disso, auxiliam em decisões complexas e criteriosas e esclarece algumas dúvidas quanto à utilização e comparabilidade de resultados obtidos de uma mesma amostra com diferentes equipamentos de pXRF.

PALAVRAS-CHAVE: sensores próximos; aprendizado de máquinas; geoquímica do solo; material de origem do solo; pegmatitos; energia verde.

ABSTRACT

Soil parent material (PM) is proving to be a critical factor in understanding soil variability, with the complexity and inaccessibility of PM in deep soils posing a challenge for detailed assessments. Proximal sensor approaches, including portable X-ray fluorescence (pXRF) and magnetic susceptibility (MS), offer practical solutions for predicting soil PM. The global search for sustainable energy alternatives has increased significantly, with lithium (Li) emerging as a key element for rechargeable Li-ion batteries. The demand for Li requires the development of cost-effective and rapid exploration methods to improve the identification of new deposits, with lithium-cesium-tantalum (LCT) pegmatites as primary sources of Li. This dissertation is divided into four chapters: (I) a pilot study focused on the generation of spatial PM predictive models for three different rock types (charnockite, mudstone, and alluvial sediments) at the Palmital Experimental Farm (Brazil), using Random Forest (RF) algorithm combined with pXRF and MS data from A and B horizons; (II) the evaluation of the effectiveness of pXRF data and the RF algorithm in predicting Li content in soil samples and Li-rich soil PM using Li pathfinder elements; (III) the comparison pXRF vs. total digestion Inductively Coupled Plasma Optical Emission Spectroscopy (ICP-OES) for soils and rocks analysis. The findings in this dissertation propose alternative, cost-effective methods for assessing soil PM spatial variability. Paper (I) had a strong validation for PM prediction results (Kappa coefficient = 0.85 and overall accuracy = 0.93). Meanwhile, PM prediction model in paper (II) achieved a Kappa coefficient of 0.77 and an overall accuracy of 0.85. The Li prediction model tested in paper (II) achieved a coefficient of determination (R^2) of 0.86, root mean square error (RMSE) of 68.5 $mg \times kg^{-1}$, and residual prediction deviation (RPD) of 1.78. Paper (III) evaluated the performance and comparability of two pXRF systems, Alpha and Beta. Both pXRFs systems produced similar results compared to reported concentrations of certified reference materials between systems and methods, showing tendencies of overestimate or underestimate elements. They could provide an alternative, pXRF-based method for more sustainable prospecting methods for PM and Li content determination and exploration. Furthermore, they help in complex and careful decisions and clarify some doubts regarding the use and comparability of results obtained from the same sample with different pXRF models.

KEYWORDS: proximal sensors; machine learning; soil geochemistry; parent material; pegmatites; green-energy.

INDICADORES DE IMPACTO

Os impactos sociais incluem a potencial geração de empregos e o desenvolvimento de competências técnicas na operação e interpretação de dados de espectroscopia de fluorescência de raios X portátil (pXRF) e a suscetibilidade magnética (SM). Tecnologicamente, a pesquisa promove a inovação em métodos de prospecção mineral. Economicamente, a eficiência dos métodos propostos pode reduzir os custos de exploração mineral, influenciando o mercado de lítio. Culturalmente, a pesquisa contribui para a conscientização ambiental e valorização do conhecimento geológico local. Os impactos estão alinhados com os Objetivos de Desenvolvimento Sustentável da ONU, especialmente no que diz respeito à energia acessível e limpa (ODS 7), trabalho decente e crescimento econômico (ODS 8), indústria, inovação e infraestrutura (ODS 9), e ação contra a mudança global do clima (ODS 13).

IMPACT INDICATORS

Social impacts include potential job creation and the development of technical skills in the operation and interpretation of pXRF and SM data. Technologically, the research promotes innovation in mineral prospecting methods. Economically, the efficiency of the proposed methods can reduce mineral exploration costs, influencing the lithium market. Culturally, research contributes to environmental awareness and appreciation of local geological knowledge. The impacts are aligned with the UN Sustainable Development Goals, especially concerning affordable and clean energy (SDG 7), decent work and economic growth (SDG 8), industry, innovation and infrastructure (SDG 9), and action against global climate change (SDG 13).

SUMMARY

FIRST PART – GENERAL INTRODUCTION

GENERAL INTRODUCTION

Soil parent material (PM)serves as the foundation for pedogenesis processes shaping soils above (Schaetzl; Anderson, 2006). In Digital Soil Mapping (DSM), PM has been commonly utilized as a proxy predictor for several soil properties (Coelho *et al*., 2021; Ma *et al.*, 2019; McBratney; Mendonça Santos; Minasny, 2003). Yet this approach has limitations, mainly concerning mapping at high resolutions (Heung; Bulmer; Schmidt, 2014). Direct observation of soil PM over large areas is challenging, particularly in highly weathered soils, where soil depth often reaches several meters (Curi.; Franzmeier, 1987; Oliveira; Jacomine; Couto, 2017; Resende *et al*., 2014; Resende, Mauro *et al.*, 2019).

Geological surveys are used for various purposes, such as surveying the geology of a given area, studying potential regions for mineral exploration, mapping fossil fuels and their aggregates, to understand and mitigate possible geological “dangers”, obtaining knowledge to improve water management and the soil, among other purposes (Compton; Compton, 2017). However, countries worldwide face obstacles in soil mapping due to financial constraints and limited-scaled geological surveys. The traditional survey consists of the systematic investigation of the geology of a given area. It includes walking through the area to recognize the landscape and study rocky outcrops, use of drills or probes to collect deeper materials, aerial photography, and geophysical techniques such as seismic and electromagnetic methods. , ground-penetrating radar. As well as laboratory analyses, such as acid digestion, isotope analysis, and dating, for a more detailed characterization of the materials, which can be laborious, time-consuming, and expensive.

Conversely, high-resolution DSM techniques offer advantages by providing detailed variation maps of soil physical, chemical, and biological properties. However, since PM influences several soil properties, challenges remain regarding the creation of detailed soil PM, which could improve soil properties mapping. For example, soil PM maps could aid precision agriculture (Gonçalves *et al.*, 2022) and terroir traceability, where soil diversity and variability play pivotal roles. Thus, there is a growing interest in mapping soil parent materials using DSM techniques, leveraging machine learning algorithms such as random forest to create predictive modeling (Heung; Bulmer; Schmidt, 2014; Lacoste; Lemercier; Walter, 2011; Ma *et al*., 2019; Richter *et al.*, 2019).

Not only for soil PM prediction and mapping but also as an increasingly adopted tool to help soil characterization, the portable X-ray fluorescence (pXRF) spectrometer has been successful due to its speed, cost-effectiveness, and samples non-destructive nature (Andrade *et al.*, 2022; Lemière, 2018b; Mancini *et al.*, 2021b; Sá *et al.*, 2023; Silva, *et al.*, 2021; Trant *et*

al., 2021; Zhao *et al.*, 2022b). Similarly, geophysical techniques as magnetic susceptibility (MS) measurements have offered insights into soil mineralogy and, indirectly, soil PM (Barbosa *et al.*, 2021; Poggere *et al.*, 2018a). Combining data from these proximal sensors can aid in PM determination, particularly in highly weathered soils where traditional methods may fall short. Some Brazilian regions have made some efforts, but evaluations of this approach under different PM and soil scenarios are still lacking, motivating further investigations (Mancini *et al.*, 2019a, 2019b). Proximal sensors have also been helpful for soil chemical characterization. Accurately quantifying elemental composition of soils and other materials is vital for geochemical analysis (Bitencourt *et al.*, 2023; Gazulla *et al.*, 2022; Melaku; Dams; Moens, 2005; Sandroni; Smith; Donovan, 2003; Volpi *et al.*, 2022; Webb; Adeloju, 2013), but limitations accompany total digestion methods as those that involve hydrofluoric acid due to its aggressive nature and environmental toxicity (Hu; Qi, 2014; Navarro *et al.*, 2008; Pinna *et al.*, 2022; Potts, 1992; Zhang; Hu, 2019; Zimmermann *et al.*, 2020). Evaporation methods have emerged as solutions to mitigate hydrofluoric acid exposure (Dulski, 2001; Pinna *et al.*, 2022; Taylor; Toms; Longerich, 2002; Zimmermann *et al.*, 2020), although they have limitations such as sample contamination and loss of volatile elements. Therefore, it is imperative to test modifications on the currently used digestion methods to find more effective and safer digestion procedures for complex geological samples, along with assessing of proximal sensors capability to achieve such end.

In this context of geochemical characterization of geological samples, it is noticed that exploration for lithium-cesium-tantalum (LCT) pegmatites has been crucial for meeting the growing demand for lithium, especially for the production of rechargeable Li-ion batteries (USGS, 2023). Conventional exploration methods are expensive and time-consuming, prompting the search for faster and more cost-effective techniques as portable Laser Induced Breakdown Spectroscopy (pLIBS) and pXRF for soil analysis (Finnigan; Golitko, 2023; Silva, *et al.*, 2021; Sweetapple; Tassios, 2015; Wise *et al.*, 2022). These methods can potentially revolutionize exploration by providing rapid and environmentally friendly alternatives to traditional approaches.

Despite the widespread use of pXRF, challenges persist such as the need for specific calibration curves to improve instrument reliability. Overall, integrating advanced sensing technologies as pXRF and machine learning algorithms holds promise for improving soil mapping, geochemical analysis, and mineral exploration (Andrade *et al.*, 2023; Park *et al.*, 2022; Schnitzler; Ross; Gloaguen, 2019; Weindorf; Chakraborty, 2023; Zhao *et al.*, 2022).

However, continuous research is essential to tackle the challenges and improve the reliability and accuracy of these techniques across various environmental contexts.

This dissertation comprises four chapters. The first one is a pilot study that aimed to create predictive models for mapping PM variation at farm-scale using random forest algorithm combining with pXRF and MS data. The second chapter assessed different methods for total digestion of soil and rock samples using microwaved-assisted digestion for soil and rocks. In the third chapter, the accuracy of pXRF data combined with random forest algorithm to predict Li content in soil samples and Li-rich soil PM prediction were evaluated, utilizing Li pathfinder elements, since pXRF is not able to detect Li. Finally, the fourth chapter compared pXRF results with Certified Reference Materials (CRMs) and total digestion for soil and rocks analysis. In summary, this dissertation addresses specific challenges and contributes to the broader scientific community by advancing our understanding of environmental processes and analytical methodologies.

REFERENCES

ALVARES, C. A. et al. Köppen's climate classification map for Brazil. Meteorologische Zeitschrift, [s. l.], v. 22, n. 6, p. 711–728, 2013.

ANDRADE, R. et al. Prediction of soil fertility via portable X-ray fluorescence (pXRF) spectrometry and soil texture in the Brazilian Coastal Plains. Geoderma, [s. l.], v. 357, n. September 2019, p. 113960, 2020.

ANDRADE, R. et al. Proximal sensing provides clean, fast, and accurate quality control of organic and mineral fertilizers. Environmental Research, [s. l.], v. 236, p. 116753, 2023.

ANDRADE, R. et al. Proximal sensor data fusion and auxiliary information for tropical soil property prediction: Soil texture. Geoderma, [s. l.], v. 422, p. 115936, 2022.

AYUSO, R. A. et al. New U-Pb Zircon Ages for Rocks from the Granite-Gneiss Terrane in Northern Michigan: Evidence for Events at ~3750, 2750, and 1850 Ma. In: 64TH INSTITUTE ON LAKE SUPERIOR GEOLOGY PROCEEDINGS, 2018. Anais [...]. [S. l.: s. n.], 2018. p. 7–8.

BARBOSA, J. Z. et al. National-scale spatial variations of soil magnetic susceptibility in Brazil. Journal of South American Earth Sciences, [s. l.], v. 108, p. 103191, 2021.

BATT, G. E. et al. Agile management and long-term strategy in exploration: the 'lucky' discovery of the Sinclair Pollucite Deposit, Eastern Goldfields, Western Australia. Australian Journal of Earth Sciences, [s. l.], v. 68, n. 5, p. 684–696, 2021.

BITENCOURT, G. R. et al. Elemental determination in carbon nanotubes by inductively coupled plasma optical emission spectrometry after a greener and simple microwave-assisted

digestion method. Spectrochimica Acta Part B: Atomic Spectroscopy, [s. l.], v. 206, p. 106709, 2023.

BOELTER, J. M.; ELG, A. M. Soil Survey of Florence County, Wisconsin. [s. l.], 2004.

BRADLEY, D. C. et al. Lithium: Professional Paper. Reston, VA: [s. n.], 2017. Report. Disponível em: https://pubs.usgs.gov/publication/pp1802K. .

BRAND, N. W.; BRAND, C. J. Detecting the Undetectable: Lithium by Portable XRF. [s. l.], 2017.

BREIMAN, L. Random Forests. Machine Learning, [s. l.], v. 45, p. 5–32, 2001.

BURLEY, L. L. et al. Rapid mineralogical and geochemical characterisation of the Fisher East nickel sulphide prospects, Western Australia, using hyperspectral and pXRF data. Ore Geology Reviews, [s. l.], v. 90, p. 371–387, 2017.

CANNON, W. F. Part 2: Field Trip Guidebooks. In: 64TH INSTITUTE ON LAKE SUPERIOR GEOLOGY PROCEEDINGS, 2018. Anais [...]. [S. l.: s. n.], 2018.

CARDOSO-FERNANDES, J. et al. Tools for remote exploration: A lithium (li) dedicated spectral library of the fregeneda–almendra aplite–pegmatite field. Data, [s. l.], v. 6, n. 3, p. 1–10, 2021.

CEC. Guide to Drought Indices and Indicators Used in North America. Montreal, QC, CA: Commission for Environmental Cooperation, 2021. Disponível em: http://www.cec.org/files/documents/publications/11872-guide-drought-indices-and-indicators-used-in-north-america-en.pdf. .

CERNY, P.; ERCIT, T. S. The classification of granitic pegmatites revisited. The Canadian Mineralogist, [s. l.], v. 43, n. 6, p. 2005–2026, 2005.

CHANG, C.-W. et al. Near-Infrared Reflectance Spectroscopy-Principal Components Regression Analyses of Soil Properties. Soil Science Society of America Journal, [s. l.], v. 65, n. 2, p. 480–490, 2001.

CHAO, T. T.; SANZOLONE, R. F. Decomposition techniques. Geoanalysis, [s. l.], v. 44, n. 1, p. 65–106, 1992.

COBLINSKI, J. A. et al. Prediction of soil texture classes through different wavelength regions of reflectance spectroscopy at various soil depths. Catena, [s. l.], v. 189, 2020.

COELHO, F. F. et al. Digital soil class mapping in Brazil: a systematic review. Scientia Agricola, [s. l.], v. 78, n. 5, p. e20190227, 2021.

COMPTON, R. R.; COMPTON, J. S. Geology in the Field. Cape Town: Earthspun Books, 2017.

CONGALTON, Russell G. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, [s. l.], v. 37, n. 1, p. 35–46, 1991.

CONRAD, O. et al. System for automated geoscientific analyses (SAGA) v. 2.1.4. Geoscientific Model Development, [s. l.], v. 8, n. 7, p. 1991–2007, 2015.

COSTA, L. M. et al. Focused-microwave-assisted acid digestion: Evaluation of losses of volatile elements in marine invertebrate samples. Journal of Food Composition and Analysis, [s. l.], v. 22, n. 3, p. 238–241, 2009.

COX, T. Exploring for lithium-cesium-tantalum (LCT) pegmatites: Weathering and dispersion aureoles assessed through soil geochemistry. 2023. 34 f. - Central Michigan University, [s. l.], 2023.

CURI, N.; FRANZMEIER, D. P. Effect of parent rocks on chemical and mineralogical properties of some Oxisols in Brazil. Soil Science Society of America Journal, [s. l.], v. 51, n. 1, p. 153–158, 1987.

CURI, N; FRANZMEIER, D. P. Effect of Parent Rocks on Chemical and Mineralogical Properties of Some Oxisols in Brazil. Soil Science Society of America Journal, [s. l.], v. 51, n. 1, p. 153–158, 1987.

DANTAS, A. A. A.; CARVALHO, L. G. de; FERREIRA, E. Classificação e tendências climáticas em Lavras, MG. Ciência e Agrotecnologia, [s. l.], v. 31, n. 6, p. 1862–1866, 2007.

DEARING, J. A. Environmental magnetic susceptibility: using the Bartington MS2 system. 1. ed. [S. l.]: Chi Pub, 1994.

DOBOS, E. Digital soil mapping: as a support to production of functional maps. [S. l.]: Office for Official Publication of the European Communities, 2006.

DUDA, B. M. et al. Soil characterization across catenas via advanced proximal sensors. Geoderma, [s. l.], v. 298, p. 78–91, 2017.

DULSKI, P. Reference Materials for Geochemical Studies: New Analytical Data by ICP-MS and Critical Discussion of Reference Values. Geostandards Newsletter, [s. l.], v. 25, n. 1, p. 87–125, 2001.

DUTTON, C. E.; BRADLEY, R. E. Lithologic, geophysical, and mineral commodity maps of precambrian rocks in Wisconsin. [S. l.: s. n.], 1970. Disponível em: https://pubs.usgs.gov/publication/i631. Acesso em: 14 fev. 2024.

ENVIRONMENTAL PROTECTION AGENCY. Method 3052: Microwave Assisted Acid Digestion of Siliceous and Organically Based Matrices. Washington DC: [s. n.], 1996.

FALSTER, A.; SIMMONS, W.; WEBBER, K. Anatectic Origin of the post-Penokean Li-Cs-Ta-enriched Pegmatites in Florence County, Wisconsin, USA. In: XXII INTERNATIONAL MINERALOGICAL ASSOCIATION MEETING, 2018, Melbourne, Australia. Anais [...]. Melbourne, Australia: [s. n.], 2018.

FALSTER, A. U.; SIMMONS, W. B.; WEBBER, K. L. The mineralogy and geochemistry of the Animikie Red Ace pegmatite, Florence county, Wisconsin. [S. l.: s. n.], 1996. Disponível

em:
http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:The+mineralogy+and+geochemistry+of+the+Animikie+Red+Ace+pegmatite,+Florence+County,+Wisconsin#1.

FINNIGAN, L. D.; GOLITKO, M. A pXRF analysis of historic brick chemical data throughout the lake Michigan Morainic region, United States. Journal of Archaeological Science: Reports, [s. l.], v. 47, p. 103817, 2023.

FROST, B. R.; FROST, C. D. On charnockites. Gondwana Research, [s. l.], v. 13, n. 1, p. 30–44, 2008.

GAZULLA, M. F. et al. Determination of trace metals by ICP-OES in petroleum cokes using a novel microwave assisted digestion method. Talanta Open, [s. l.], v. 6, p. 100134, 2022.

GONÇALVES, M. G. M. et al. Soil and climate effects on winter wine produced under the tropical environmental conditions of southeastern Brazil. OENO One, [s. l.], v. 56, n. 2, p. 63–79, 2022.

GOZUKARA, G.; ZHANG, Y.; HARTEMINK, A. E. Using vis-NIR and pXRF data to distinguish soil parent materials – An example using 136 pedons from Wisconsin, USA. Geoderma, [s. l.], v. 396, n. February, p. 115091, 2021.

HAN, S. et al. Evaluation of various microwave-assisted acid digestion procedures for the determination of major and heavy metal elements in municipal solid waste incineration fly ash. Journal of Cleaner Production, [s. l.], v. 321, p. 128922, 2021.

HAN, Z. et al. Multi-scale exploration of giant Qulong porphyry deposit in a collisional setting. Ore Geology Reviews, [s. l.], v. 139, p. 104455, 2021.

HEUNG, B.; BULMER, C. E.; SCHMIDT, M. G. Predictive soil parent material mapping at a regional-scale: A Random Forest approach. Geoderma, [s. l.], v. 214–215, p. 141–154, 2014.

HORTA, A. et al. Integrating portable X-ray fluorescence (pXRF) measurement uncertainty for accurate soil contamination mapping. Geoderma, [s. l.], v. 382, 2021.

HU, Z.; QI, L. 15.5 - Sample Digestion Methods. In: HOLLAND, H. D.; TUREKIAN, K. K. (org.). Treatise on Geochemistry (Second Edition). Oxford: Elsevier, 2014. p. 87–109. Disponível em: https://www.sciencedirect.com/science/article/pii/B9780080959757014066.

JAMES, H. L. et al. Geology of central Dickinson County, Michigan: Professional Paper. [S. l.: s. n.], 1961. Report. Disponível em: https://pubs.usgs.gov/publication/pp310. .

JOHNSON, C. M.; DAALEN, C. M. V. Mineralogy and geochemistry of Late Archean and Paleoproterozoic granites and pegmatites in the Northern Penokean terrane of Marquette and Dickinson Counties, Michigan. 2015. - University of New Orleans Theses and Dissertations. 2088, [s. l.], 2015. Disponível em: https://scholarworks.uno.edu/td/2088.

KÄMPF, N.; MARQUES, J. J.; CURI, N. Mineralogia de solos brasileiros. In: KER, J. C. et al. Pedologia: Fundamentos. Viçosa, MG: SBCS, 2012. p. 81–145.

KHALEDIAN, Y.; MILLER, B. A. Selecting appropriate machine learning methods for digital soil mapping. Applied Mathematical Modelling, [s. l.], v. 81, p. 401–418, 2020.

KUHN, M. Building predictive models in R using the caret package. Journal of statistical software, [s. l.], v. 28, n. 5, p. 1–26, 2008.

LACOSTE, M.; LEMERCIER, B.; WALTER, C. Regional mapping of soil parent material by machine learning based on point data. Geomorphology, [s. l.], v. 133, n. 1, p. 90–99, 2011.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. Biometrics, [s. l.], v. 33, n. 1, p. 159, 1977.

LANGMYHR, F. J.; KRINGSTAD, K. An investigation of the composition of the precipitates formed by the decomposition of silicate rocks in 38-40% hydrofluoric acid. Analytica Chimica Acta, [s. l.], v. 35, p. 131–135, 1966.

LAWLEY, R.; SMITH, B. Digital Soil Mapping at a National Scale: A Knowledge and GIS Based Approach to Improving Parent Material and Property Information. In: HARTEMINK, A. E.; MCBRATNEY, A.; MENDONÇA-SANTOS, M. de L. (org.). Digital Soil Mapping with Limited Data. Dordrecht: Springer Netherlands, 2008. p. 173–182. Disponível em: https://doi.org/10.1007/978-1-4020-8592-5_14.

LEMIÈRE, B. A review of pXRF (field portable X-ray fluorescence) applications for applied geochemistry. Journal of Geochemical Exploration, [s. l.], v. 188, p. 350–363, 2018

LIAW, A.; WIENER, M. Classification and regression by random forest. R News, [s. l.], v. 2, p. 18–22, 2002.

LIMA, W. de et al. Tracing the origin of deposited sediments: A study applying proximal sensing in a drainage subbasin. Journal of South American Earth Sciences, [s. l.], v. 123, p. 104241, 2023.

LIU, X. M. et al. Processes controlling lithium isotopic distribution in contact aureoles: A case study of the Florence County pegmatites, Wisconsin. Geochemistry, Geophysics, Geosystems, [s. l.], v. 11, n. 8, p. 1–21, 2010.

LOISEAU, T. et al. Density of soil observations in digital soil mapping: A study in the Mayenne region, France. Geoderma Regional, [s. l.], v. 24, p. e00358, 2021.

LONDON, David. Pegmatites. [S. l.]: Mineralogical Association of Canada, 2008.

LONDON, D.; MORGAN, G. B. The pegmatite puzzle. Elements, [s. l.], v. 8, n. 4, p. 263–268, 2012.

MA, Y. et al. Pedology and digital soil mapping (DSM). European Journal of Soil Science, [s. l.], v. 70, n. 2, p. 216–235, 2019.

MANCINI, M. et al. Formation and variation of a 4.5 m deep Oxisol in southeastern Brazil. Catena, [s. l.], v. 206, 2021.

MANCINI, M. et al. Parent material distribution mapping from tropical soils data via machine learning and portable X-ray fluorescence (pXRF) spectrometry in Brazil. Geoderma, [s. l.], v. 354, p. 113885, 2019a.

MANCINI, M. et al. Soil parent material prediction for Brazil via proximal soil sensing. Geoderma Regional, [s. l.], v. 22, p. e00310, 2020.

MANCINI, M. et al. Tracing tropical soil parent material analysis via portable X-ray fluorescence (pXRF) spectrometry in Brazilian Cerrado. Geoderma, [s. l.], v. 337, p. 718–728, 2019b.

MCBRATNEY, A. B.; MENDONÇA SANTOS, M. L.; MINASNY, B. On digital soil mapping. Geoderma, [s. l.], v. 117, p. 3–52, 2003.

MELAKU, S.; DAMS, R.; MOENS, L. Determination of trace elements in agricultural soil samples by inductively coupled plasma-mass spectrometry: Microwave acid digestion versus aqua regia extraction. Analytica Chimica Acta, [s. l.], v. 543, n. 1–2, p. 117–123, 2005.

MELLO, F. A. O. et al. Soil parent material prediction through satellite multispectral analysis on a regional scale at the Western Paulista Plateau, Brazil. Geoderma Regional, [s. l.], v. 26, p. e00412, 2021.

MÜLLER, A. et al. GREENPEG–exploration for pegmatite minerals to feed the energy transition: first steps towards the Green Stone Age. Geological Society, London, Special Publications, [s. l.], v. 526, n. 1, p. SP526-2021, 2022.

MURATLI, J. M. et al. Dissolution of fluoride complexes following microwave-assisted hydrofluoric acid digestion of marine sediments. Talanta, [s. l.], v. 89, p. 195–200, 2012.

NAIMI, S. et al. Quantification of some intrinsic soil properties using proximal sensing in arid lands: Application of Vis-NIR, MIR, and pXRF spectroscopy. Geoderma Regional, [s. l.], v. 28, p. e00484, 2022.

NAVARRO, M. S. et al. The Direct Determination of Rare Earth Elements in Basaltic and Related Rocks using ICP-MS: Testing the Efficiency of Microwave Oven Sample Decomposition Procedures. Geostandards and Geoanalytical Research, [s. l.], v. 32, n. 2, p. 167–180, 2008.

NIEDER, R. et al. The geochemical signature of rare-metal pegmatites in the Central Africa Region: Soils, plants, water and stream sediments in the Gatumba tin-tantalum mining district, Rwanda. Journal of Geochemical Exploration, [s. l.], v. 144, n. PC, p. 539–551, 2014.

OLIVEIRA, V. A.; JACOMINE, P. T. K.; COUTO, E. G. Solos do bioma Cerrado. In: PEDOLOGIA - SOLOS DOS BIOMAS BRASILEIROS. N. Curi, J.C. Ker, R.F. Novais, P. Vidal-Torrado, C.E.G.R. Schaefered. Viçosa, MG: SBCS, 2017. p. 177–226.

PARK, C. et al. Detection of Titanium bearing Myeonsan Formation in the Joseon Supergroup based on Spectral Analysis and Machine Learning Techniques. ECONOMIC AND ENVIRONMENTAL GEOLOGY, [s. l.], v. 55, n. 2, p. 197–207, 2022.

PELEGRINO, M. H. P. et al. Synthesis of proximal sensing, terrain analysis, and parent material information for available micronutrient prediction in tropical soils. Precision Agriculture, [s. l.], v. 20, n. 4, p. 746–766, 2019.

PHELPS-BARBER, Z.; TRENCH, A.; GROVES, D. I. Recent pegmatite-hosted spodumene discoveries in Western Australia: insights for lithium exploration in Australia and globally. Applied Earth Science, [s. l.], v. 131, n. 2, p. 100–113, 2022.

PIERANGELI, L. M. et al. Prediction of li via geochemical data and machine learning algorithms in soils formed on pegmatite bedrock: a preliminary study. In: GSA CONNECTS 2022 MEETING IN DENVER, COLORADO, 2022. Anais [...]. [S. l.: s. n.], 2022. p. 382238. Disponível em: https://gsa.confex.com/gsa/2022AM/webprogram/Paper382238.html. Acesso em: 14 fev. 2024.

PIERANGELI, L. M. P. et al. Soil parent material spatial modeling at high resolution from proximal sensing and machine learning: A pilot study. Journal of South American Earth Sciences, [s. l.], v. 129, p. 104498, 2023a.

PIERANGELI, L. M. P. et al. Soil parent material spatial modeling at high resolution from proximal sensing and machine learning: A pilot study. Journal of South American Earth Sciences, [s. l.], v. 129, p. 104498, 2023b.

PINNA, M. et al. How to Clean and Safely Remove HF from Acid Digestion Solutions for Ultra-Trace Analysis: A Microwave-Assisted Vessel-Inside-Vessel Protocol. Methods and Protocols, [s. l.], v. 5, n. 2, p. 30, 2022.

POGGERE, G. C. et al. Maghemite quantification and magnetic signature of Brazilian soils with contrasting parent materials. Applied Clay Science, [s. l.], v. 161, p. 385–394, 2018.

POTTS, P. J. Classical and rapid methods of analysis. In: POTTS, P. J. (org.). A Handbook of Silicate Rock Analysis. Boston, MA: Springer US, 1992. p. 47–76. Disponível em: https://doi.org/10.1007/978-1-4615-3270-5_2.

QUÉMÉNEUR, J. et al. Geologia da Folha Lavras 1:100.000. Geologia e Recursos Minerais do Sudeste Brasileiro. 1. ed. Belo Horizonte, Minas Gerais: Companhia Mineradora de Minas Gerais, Secretaria de Desenvolvimento Econômico, 2003.

R DEVELOPMENT CORE TEAM. R: A language and environmental for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2018.

RAVANSARI, R.; WILSON, S. C.; TIGHE, M. Portable X-ray fluorescence for environmental assessment of soils: Not just a point and shoot method. Environment International, [s. l.], v. 134, p. 105250, 2020.

REMETEIOVÁ, D. et al. Evaluation of US EPA Method 3052 Microwave Acid Digestion for Quantification of Majority Metals in Waste Printed Circuit Boards. Metals, [s. l.], v. 10, n. 11, p. 1511, 2020.

RESENDE, M. et al. Da rocha ao solo: enfoque ambiental. Lavras: Editora UFLA, 2019.

RESENDE, M. Mineralogia de solos brasileiros: interpretação e aplicações. [S. l.]: UFLA, 2005.

RESENDE, M. et al. Pedologia: base para distinção de ambientes. 6.ed.ed. Lavras, MG: Editora UFLA, 2014.

RICHTER, J. et al. Mapping parent material as part of a nested approach to soil mapping in the Arkansas River Valley. CATENA, [s. l.], v. 178, p. 100–108, 2019.

ROUILLON, M.; TAYLOR, M. P. Can field portable X-ray fluorescence (pXRF) produce high quality data for application in environmental contamination research?. Environmental Pollution, [s. l.], v. 214, p. 255–264, 2016.

SÁ, R. T. S. de et al. Detailed characterization of iron-rich tailings after the Fundão dam failure, Brazil, with inclusion of proximal sensors data, as a secure basis for environmental and agricultural restoration. Environmental Research, [s. l.], v. 228, p. 115858, 2023.

SANDRONI, V.; SMITH, C. M. M.; DONOVAN, A. Microwave digestion of sediment, soils and urban particulate matter for trace metal analysis. Talanta, [s. l.], v. 60, n. 4, p. 715–723, 2003.

SANTOS, H. G. dos et al. Sistema brasileiro de classificação de solos. 5, revista e ampliada. ed. Brasília: Embrapa Solos, 2018.

SCHAETZL, R.; ANDERSON, S. Soil-Genesis and Geomorphology. [S. l.: s. n.], 2006. v. 817

SCHNITZLER, N.; ROSS, P.; GLOAGUEN, E. Using machine learning to estimate a key missing geochemical variable in mining exploration: Application of the Random Forest algorithm to multisensor core logging data. JOURNAL OF GEOCHEMICAL EXPLORATION, [s. l.], v. 205, 2019.

SCHWERTMANN, U.; TAYLOR, R. M. Iron oxides. In: DIXON, J. B.; WEED, S. B. (org.). SSSA Book Series. Madison, WI, USA: Soil Science Society of America, 1989. p. 379–438. Disponível em: http://doi.wiley.com/10.2136/sssabookser1.2ed.c8. Acesso em: 7 ago. 2020.

SILVA, S. H. G. et al. Chapter One - pXRF in tropical soils: Methodology, applications, achievements and challenges. In: SPARKS, D. L. B. T.-A. in A. (org.). [S. l.]: Academic Press, 2021. v. 167, p. 1–62. Disponível em: https://www.sciencedirect.com/science/article/pii/S006521132030105X.

SILVA, F. M. et al. Chemical and mineralogical changes in the textural fractions of quartzite-derived tropical soils, along weathering, assessed by portable X-ray fluorescence spectrometry and X-ray diffraction. Journal of South American Earth Sciences, [s. l.], v. 112, p. 103634, 2021.

SILVA, S. H. G. et al. Proximal sensing and digital terrain models applied to digital soil mapping and modeling of Brazilian Latosols (Oxisols). Remote Sensing, [s. l.], v. 8, n. 8, p. 614, 2016.

SILVA, S. H. G. et al. Soil weathering analysis using a portable X-ray fluorescence (PXRF) spectrometer in an Inceptisol from the Brazilian Cerrado. Applied Clay Science, [s. l.], v. 162, p. 27–37, 2018.

SIMS, P. K. Boundary between Archean greenstone and gneiss terranes in northern Wisconsin and Michigan. In: GEOLOGICAL SOCIETY OF AMERICA SPECIAL PAPERS. [S. l.]: Geological Society of America, 1980. v. 182, p. 113–124. Disponível em: https://pubs.geoscienceworld.org/books/book/317/chapter/3795683/. Acesso em: 14 fev. 2024.

SIMS, P. K. Geologic map of Precambrian rocks, southern Lake Superior region, Wisconsin and northern Michigan. [S. l.: s. n.], 1992. Disponível em: https://pubs.usgs.gov/publication/i2185. Acesso em: 14 fev. 2024.

SIMS, P. K.; SCHULZ, K. J.; PETERMAN, Z. E. Geology and geochemistry of early Proterozoic rocks in the Dunbar area, northeastern Wisconsin: Professional Paper. [S. l.: s. n.], 1992. Report. Disponível em: https://pubs.usgs.gov/publication/pp1517. .

SIRBESCU, M. L. C.; HARTWICK, E. E.; STUDENT, J. J. Rapid crystallization of the Animikie Red Ace Pegmatite, Florence county, northeastern Wisconsin: Inclusion microthermometry and conductive-cooling modeling. Contributions to Mineralogy and Petrology, [s. l.], v. 156, n. 3, p. 289–305, 2008.

SOARES, C. C. V. et al. Mineralogical, micromorphological and geochemical transformations in the initial steps of the weathering process of charnockite from the Caparaó Range, southeastern Brazil. Journal of South American Earth Sciences, [s. l.], v. 56, p. 30–40, 2014.

SOIL SURVEY STAFF. Keys to Soil Taxonomy. 12. ed. Washington, DC: USDA, 2014.

SOIL SURVEY STAFF, NATURAL RESOURCES CONSERVATION SERVICE, UNITED STATES DEPARTMENT OF AGRICULTURE. Soil Survey Staff. [S. l.], 2023. Disponível em: https://websoilsurvey.sc.egov.usda.gov/app/. Acesso em: 16 out. 2023.

STEINER, B. Tools and Workflows for Grassroots Li–Cs–Ta (LCT) Pegmatite Exploration. Minerals, [s. l.], v. 9, n. 8, p. 499, 2019.

STOCKMANN, U. et al. Utilizing portable X-ray fluorescence spectrometry for in-field investigation of pedogenesis. Catena, [s. l.], v. 139, p. 220–231, 2016.

SWEETAPPLE, M. T.; TASSIOS, S. Laser-induced breakdown spectroscopy (LIBS) as a tool for in situ mapping and textural interpretation of lithium in pegmatite minerals. American Mineralogist, [s. l.], v. 100, n. 10, p. 2141–2151, 2015.

TAN, K. H. Principles of soil chemistry. [S. l.: s. n.], 1998. Disponível em: http://encore.keele.ac.uk/iii/encore/record/C__Rb1544139__Ssoil__P0,4__Orightresult__U__X6?lang=eng&suite=cobalt.

TANG, Y. et al. K-feldspar composition as an exploration tool for pegmatite-type rare metal deposits in Altay, NW China. Journal of Geochemical Exploration, [s. l.], v. 185, p. 130–138, 2018.

TAYLOR, V.; TOMS, A.; LONGERICH, H. Acid digestion of geological and environmental samples using open-vessel focused microwave digestion. Analytical and Bioanalytical Chemistry, [s. l.], v. 372, n. 2, p. 360–365, 2002.

THOMPSON, R.; OLDFIELD, F. Environmental magnetism. Dordrecht: Springer Netherlands, 1986. Disponível em: http://link.springer.com/10.1007/978-94-011-8036-8. Acesso em: 30 set. 2019.

TING, K. M. Confusion Matrix. In: , 2010. Encyclopedia of Machine Learning and Data Mining. [S. l.: s. n.], 2010.

TRANT, P. L. K. et al. Sampling density and spatial analysis: a methodological pXRF study of the geochemistry of a Viking-Age house in Ribe, Denmark. Archaeological and Anthropological Sciences, [s. l.], v. 13, n. 1, p. 21, 2021.

TRUEMAN, D. L. Exploration for rare-element granitic pegmatites. Short course in granitic pegmatites in science and industry, [s. l.], p. 463–493, 1982.

TWYMAN, R. M. SAMPLE DISSOLUTION FOR ELEMENTAL ANALYSIS | Wet Digestion. In: WORSFOLD, P.; TOWNSHEND, A.; POOLE, C. (org.). Encyclopedia of Analytical Science (Second Edition). Oxford: Elsevier, 2005. p. 146–153. Disponível em: https://www.sciencedirect.com/science/article/pii/B0123693977005392.

USGS. Mineral Commodity Summaries 2023: Mineral Commodity Summaries. [S. l.: s. n.], 2023. Disponível em: https://doi.org/10.3133/mcs2023. .

VAUDOUR, E. et al. An overview of the recent approaches to terroir functional modelling, footprinting and zoning. SOIL, [s. l.], v. 1, n. 1, p. 287–312, 2015.

VOLPI, M. et al. Microwave-assisted sample preparation of α-spodumene: A simple procedure for analysis of a complex sample. Minerals Engineering, [s. l.], v. 187, p. 107820, 2022.

WADOUX, A. M. J.-C.; MINASNY, B.; MCBRATNEY, A. B. Machine learning for digital soil mapping: Applications, challenges and suggested solutions. Earth-Science Reviews, [s. l.], v. 210, p. 103359, 2020.

WEBB, B.; ADELOJU, S. B. Evaluation of some wet digestions methods for reliable determination of total phosphorus in Australian soils. Microchemical Journal, [s. l.], v. 111, p. 47–52, 2013.

WEINDORF, D. C.; CHAKRABORTY, S. Balancing machine learning and artificial intelligence in soil science with human perspective and experience. Pedosphere, [s. l.], p. S1002016023001091, 2023.

WEINDORF, D. C.; CHAKRABORTY, S. Portable X-ray Fluorescence Spectrometry Analysis of Soils. Methods of Soil Analysis, [s. l.], v. 1, n. 1, p. 0, 2016.

WEINDORF, D. C.; CHAKRABORTY, S. Portable X-ray fluorescence spectrometry analysis of soils. Soil Science Society of America Journal, [s. l.], v. 84, n. 5, p. 1384–1392, 2020.

WHITTEN, D. G. A.; BROOKS, J. R. V. The Penguin Dictionary of Geology. Harmondsworth, England: Penguin Books, 1973.

WILSON, M. J. The importance of parent material in soil classification: A review in a historical context. CATENA, [s. l.], v. 182, p. 104131, 2019.

WISE, M. A. et al. Handheld LIBS for Li Exploration: An Example from the Carolina Tin-Spodumene Belt, USA. Minerals, [s. l.], v. 12, n. 1, 2022.

YUAN, Y. et al. Determination of Se and Te by hydride generation-inductively coupled plasma mass spectrometry after mixed-acid digestion of tungsten ores. Spectrochimica Acta Part B: Atomic Spectroscopy, [s. l.], v. 203, p. 106664, 2023.

YUAN, Z. et al. In situ monitoring of elemental losses and gains during weathering using the spatial element patterns obtained by portable XRF. Journal of Geochemical Exploration, [s. l.], v. 229, 2021.

ZHANG, W.; HU, Z. Recent advances in sample preparation methods for elemental and isotopic analysis of geological samples. Spectrochimica Acta Part B: Atomic Spectroscopy, [s. l.], v. 160, p. 105690, 2019.

ZHAO, L. et al. Pedogenic-weathering evolution and soil discrimination by sensor fusion combined with machine-learning-based spectral modeling. Geoderma, [s. l.], v. 409, p. 115648, 2022.

ZI, J.-W. et al. Refining the Paleoproterozoic tectonothermal history of the Penokean Orogen: New U-Pb age constraints from the Pembine-Wausau terrane, Wisconsin, USA. GSA Bulletin, [s. l.], v. 134, n. 3–4, p. 776–790, 2022.

ZIMMERMANN, T. et al. Substituting HF by HBF 4 – an optimized digestion method for multi-elemental sediment analysis via ICP-MS/MS. Analytical Methods, [s. l.], v. 12, n. 30, p. 3778–3787, 2020.

SECOND PART – ARTICLES

ARTICLE 1 - SOIL PARENT MATERIAL SPATIAL MODELING AT HIGH RESOLUTION FROM PROXIMAL SENSING AND MACHINE LEARNING: A PILOT STUDY

*Article published in the Journal of South American Earth Sciences in July 2023.
https://doi.org/10.1016/j.jsames.2023.104498

Luiza Maria Pereira Pierangeli[a], Sérgio Henrique Godinho Silva[a], Anita Fernanda dos Santos Teixeira[a], Marcelo Mancini[a], Renata Andrade [a], Michele Duarte de Menezes[a], Mona-Liza C. Sirbescu[b], João José Marques[a], David C. Weindorf[b], Nilton Curi[a*]

[a]Federal University of Lavras, Department of Soil Science. P.O. Box 3037. Zip Code 37200-900, Lavras, Minas Gerais, Brazil
[b]Department of Earth and Atmospheric Sciences, Central Michigan University, Mount Pleasant, MI, 48859, USA.

*Corresponding author: weind1dc@cmich.edu; Tel.: +1 989-774-6467

ABBREVIATIONS: pXRF – portable X-ray fluorescence; MS – magnetic susceptibility; GIS – geographic information system; DSM – Digital Soil Mapping; RMSE – root mean square error; R^2 – coefficient of determination; NIST – National Institute of Standards and Technology; RF – Random Forest; CV – coefficient of variation; Min – minimum value; Max – maximum value; SD – standard deviation; LOD – limit of detection.

ABSTRACT

Although parent material (PM) is one of the five soil formation factors providing key information on soil variability, the complexity of PM distributions and the difficulty of reaching PM in deep soils prevent its detailed assessment. Proximal sensor approaches such as portable X-ray fluorescence (pXRF) and magnetic susceptibility (MS) may be helpful in predicting soil PM in a more practical and accessible way. This pilot study aimed to create spatial PM predictive models for three distinct rock types (charnockite, mudstone, and alluvial sediments) of the Palmital Experimental Farm (Brazil) through random forest (RF) algorithm based on soil samples analyzed via pXRF and MS. Soils were sampled in A and B horizons following a regular-grid design covering the whole study area. The RF algorithm was calibrated to predict PMs using samples from the B horizon of soils with known PM. The prediction model was applied to the area for mapping PM across the whole farm. For validation, PM was identified at 15 different sites and compared with the predicted PM shown on the maps via overall accuracy, Kappa coefficient, producer's, and user's accuracies. Al, Fe, Si, Ti and MS proximal sensor data discriminated well among soils derived from charnockite, mudstone and alluvial sediments. The map built based on B horizon data showed greater accuracy (overall accuracy = 0.93, Kappa coefficient = 0.85, user's accuracy = 0.92, and producer's accuracy = 0.97) than the map built with A horizon samples (0.73, 0.48, 0.48, and 0.58). These results could represent alternative methods for reducing costs and accelerating the assessment of soil PM spatial variability, supporting optimized agronomic and environmental decision making.

Keywords: pedology, portable X-ray fluorescence spectrometer, magnetic susceptibilimeter, machine learning, tropical soils.

1. INTRODUCTION

Soil parent material (PM) constitutes the substrate on which the pedogenesis processes act to generate the upper living soils (Schaetzl and Anderson, 2006). Most of studies addressing Digital Soil Mapping (DSM) frameworks have largely applied bedrock lithology as a surrogate predictor of PM (Coelho et al., 2021; Ma et al., 2019; McBratney et al., 2003). Such approach has some limitations because: 1) it does not account for those soils that are formed from other type of geomorphic transportation, underrepresenting PM (Lacoste et al., 2011) and increasing bias in soil predictive models (Heung et al., 2014); and 2) the majority of information is generated in polygon format at a coarse scale (Lacoste et al., 2011). In addition, it is generally difficult and impractical to directly observe the soil PM over a large area (Ma et al., 2019),

especially in very weathered and leached soils because they may reach several meters in depth (Curi and Franzmeier, 1987; Oliveira et al., 2017; Resende et al., 2019, 2014). Furthermore, lithological characterizations in geological reports rely basically on 'fresh' material, containing only brief descriptions of weathered material and scarce information on soils, if any (Lawley and Smith, 2008).

The lack of financial support and small scale geological surveys represents an obstacle in developing countries. As alternatives to traditional geological surveys, high-resolution DSM techniques offer great advantages since they provide important environmental covariables (McBratney et al., 2003), related to several soil physical, chemical, mineralogical, and biological properties, besides soil PM (Heung et al., 2014; Lacoste et al., 2011; Wilson, 2019, Mancini et al., 2019). By addressing soil diversity and variability across the landscape, high resolution (5m) PM spatial prediction also consists of valuable information for precision agriculture management (Pelegrino et al., 2019), and terroir traceability of different agricultural products whose information is oftentimes linked to PM (Gonçalves et al., 2022; Vaudour et al., 2015).

Thus, there has been interest in mapping soil parent materials using the DSM techniques (Heung et al., 2014; Lacoste et al., 2011; Ma et al., 2019; Richter et al., 2019), especially by means of digital soil models. In this case, DSM are created based on a set of inferences from soil datasets as auxiliary spatial variables, describing the variations of the soil-forming factors over the mapped zone (Dobos, 2006). Machine learning techniques emerged as a powerful set of techniques for soil predictive models (Wadoux et al., 2020), specially Random Forest (RF), which has been used in a large number of successful applications. However, in general, final products at high-resolution demands higher density of information, decreasing the error of final spatial predictions (Khaledian and Miller, 2020; Loiseau et al., 2021). Thus, proximal sensing applied to soils to assess their geochemistry and geophysics becomes an important alternative strategy that allows rapid collection of information at low cost. In addition, soil A and B horizons are often easily accessible, and it is expected that soil could trace its PM (Mancini et al., 2020).

Portable X-ray fluorescence (pXRF) spectrometry has been increasingly adopted by the soil science community for geochemical characterizations (Lemière, 2018; Silva et al., 2021; Trant et al., 2021). It is capable of identifying and quantifying elements in soil (from Mg to U in periodic table) in a few seconds, at low cost and without generating analytical wastes. Conversely, geophysics is the application of physical principles to study the Earth. For instance, a magnetic susceptibilimeter has been used for decades to determine magnetism (Thompson

and Oldfield, 1986), which for soils is related to soil mineralogy and, indirectly, to soil PM (Barbosa et al., 2021; Poggere et al., 2018). Thus, combining these sensors may aid in the determination of soil PM, because the PM controls the amounts of magnetic minerals and of certain elements, that persist even in highly weathered/leached soils, in the form of Fe oxides (Curi and Franzmeier, 1987).

To further investigate the capability of proximal sensors to infer PM from their resulting soils under DSM framework, the objectives of this pilot study were to: 1) characterize soils developed from different PM's (charnockite, mudstone, and alluvial sediments) of the Palmital Experimental Farm, Ijaci, Minas Gerais, Brazil, through proximal sensors; 2) build a robust PM prediction model, point prediction, from soil samples with known PM, through RF algorithm; and 3) use pXRF and MS information to create a spatial PM predictive model of the soils in the study area and assess the accuracy of these models (maps) via external validation. We hypothesized that the elemental content obtained from pXRF coupled with MS data and random forest algorithm will be able to generate a powerful prediction model that will deliver an accurate PM map for the study area.

2. MATERIALS AND METHODS

2.1 General occurrence and features

This work was conducted using soil samples collected from Palmital Experimental Farm, located in the municipality of Ijaci, Minas Gerais state, Brazil, between UTM longitudes 506,793 and 508,882 m, and latitudes 7,659,470 and 7,660,685 m, zone 23K, datum SIRGAS 2000 (Figure 1). The farm occupies 117 ha and the collected samples covered three soil orders per the Brazilian Soil Classification System (Santos et al., 2018): Argisols, Cambisols, and Latosols. Such soils are equivalent to Ultisols, Inceptisols, and Oxisols, respectively, per Keys to US Soil Taxonomy (Soil Survey Staff, 2014).

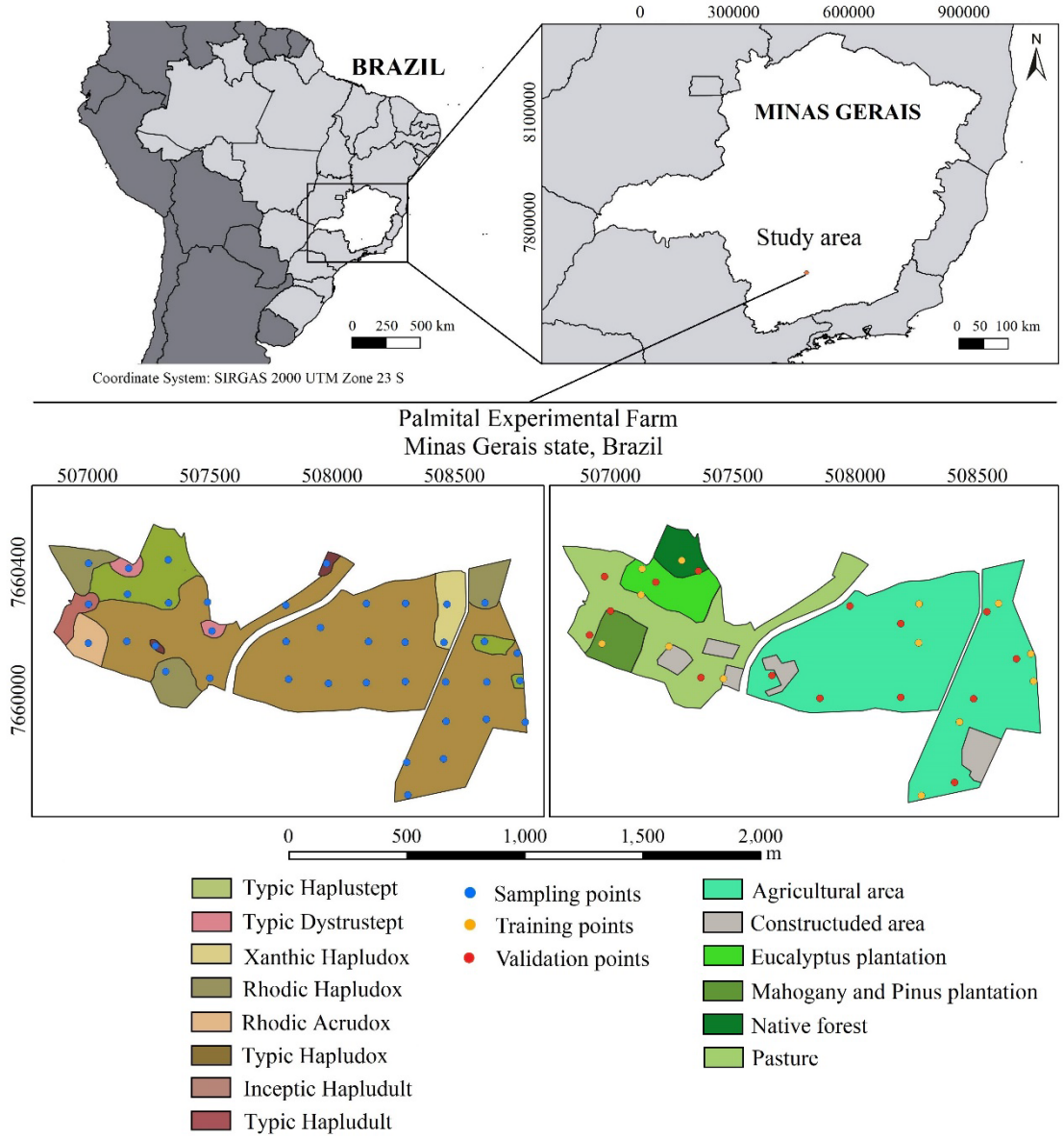

Figure 1. Study area location, land use and sampling locations for training, validation and spatial distribution of soil at the Palmital Experimental Farm, Minas Gerais state, Brazil.

Land uses include native vegetation (semiperennial tropical forest) and agricultural plantations (e.g., *Eucalyptus sp., Swietenia macrophylla, Pinus sp.*). The Köppen climate classification is humid semitropical with dry winters and rainy summers (Cwa), with annual mean temperature of 21ºC, annual mean precipitation of 1,500 mm; altitude ranges from 814 to 866 m (Alvares et al., 2013; Dantas et al., 2007). According to the *Geological Map of the Region of Lavras – 1:100,000* (Figure 2), charnockite, alluvial sediments, tonalite–trondhjemite–granodiorite (TTG) gneiss, metalimestone and calciphyllite are the parent materials of the study area (Quéméneur et al., 2003). Due to the small scale of the survey, different types of PM were

identified in the farm during the field work, by comparing with different soils and parent material of the region.

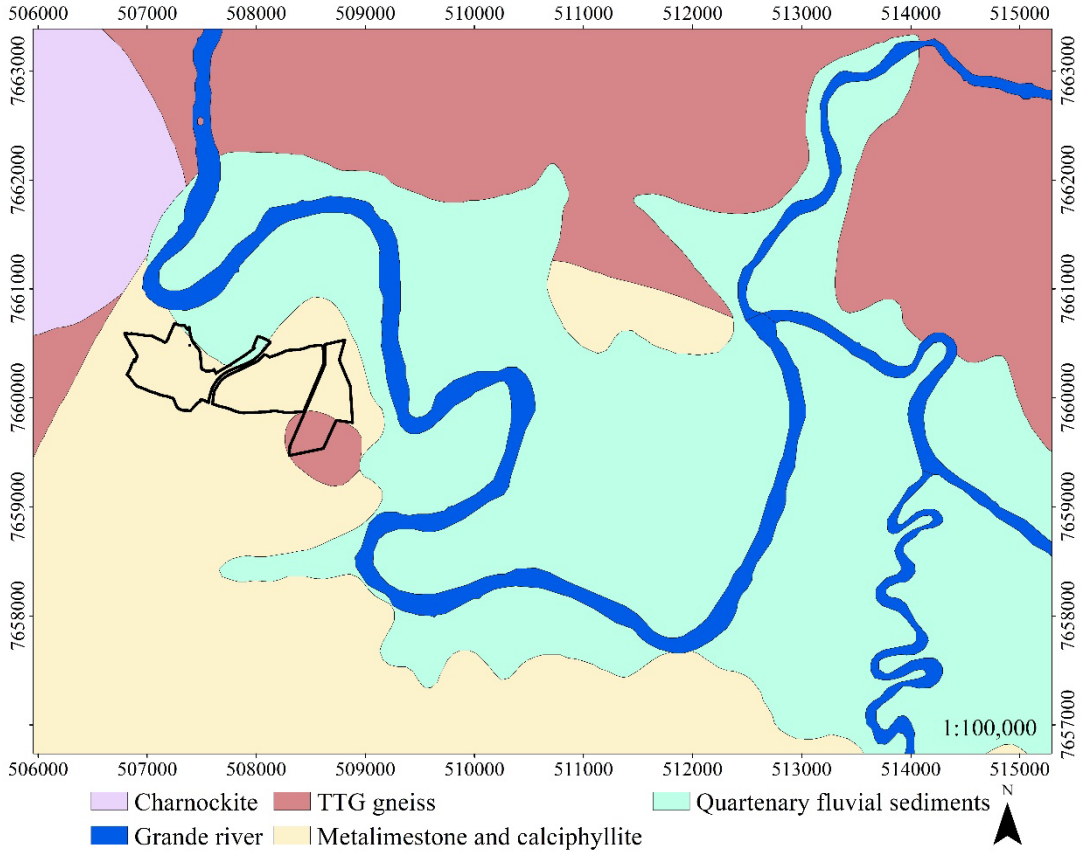

Figure 2. Geological survey adapted from *Geological Map of the Region of Lavras - 1:100,000 scale* (Quéméneur et al., 2003) at the Palmital Experimental Farm region, Minas Gerais state, Brazil.

2.2 Soil sampling and laboratory analyses

A total of 39 locations (small trenches) in a 173 m regular grid (Fig. 1) was designed throughout the study area for the calibration of models (13 locations), while other samples were collected at locations where soil PM was detectable for the validation of the generated maps (15 locations) (Figure 1). The farmland presented a great variation of weathering degree of soils, implying on variations in the B horizon thickness. The younger the soil, the thicker and further from underlying soil parent material is the B horizon (crescent order of weathering/age: Inceptisols → Ultisols → Oxisols). To standardize such effects, soils were sampled at depths of 0-0.20 m and 0.4-0.7, where A and B horizons, respectively, were located for all 78 samples. The soil samples were air-dried, disaggregated to pass a 2 mm sieve, and subjected to laboratory analysis. A pXRF spectrometer model Vanta M series (Olympus, Waltham, MA, USA) was

used to scan the samples according to Silva et al. (2021) and Weindorf and Chakraborty (2020). The equipment has a Rh X-ray tube operated at 50 kV and it was operated using the "soil mode" factory calibration to quantify the element concentrations based on two beams that operate sequentially. Each beam has a different intensity of energy, allowing for the detection of light and heavy elements according to this variable intensity. Each beam was configured to scan the samples for 30 s so that a full scan was completed in 60 s. All samples were scanned in triplicate and the average of the three scans was calculated to deliver the final elemental content.

To guarantee the quality of the generated data, materials certified by the National Institute of Standards and Technology (NIST) (2710a and 2711a) were analyzed as references to compare their elemental content with the pXRF results. The recovery values obtained by the equipment per element were calculated (recovery values = elemental content obtained by pXRF/certified elemental content). The recoveries were for Al; 0.61, 0.61, for Fe; 0.92, 0.96, for Si: 0.57, 0.61 and for Ti 0.90, 0.94 for NIST 2710a and 2711a, respectively.

Magnetic susceptibility (MS) is a measurement to determine how magnetized a material can become when a magnetic field is applied to it. MS per unit of mass (MS, x 10^{-7} m^3 kg^{-1}) was determined using a Bartington MS2B magnetic susceptibilimeter at a low frequency of 0.47 kHz (LF) (Dearing, 1994). The measurements were carried on 10 g of air-dried 2-mm sieved sample. Calculations were made via the expression $\chi_{lf} = (10 \times k)$ m^{-1}, where k is a dimensionless parameter, and m is the mass of the used sample.

2.3 Parent material prediction and spatial variability

Thirteen samples from soil B horizons with known PM were chosen to build the prediction model using the Random Forest (RF) algorithm. Those samples covered all three types of parent materials found in the study area during field work (charnockite, mudstone, and alluvial sediments). The independent variables of the prediction model were Al, Fe, Si, and Ti contents obtained using the pXRF in addition to MS data. These chemical elements are appropriate PM fingerprints in tropical soils due to high degree of weathering and leaching, elucidating differences when comparing soils developed from different types of PM (Curi and Franzmeier, 1987; Mancini et al., 2019a, 2019b; Mello et al., 2021; Silva et al., 2016).

Processing and modeling were performed using R software (R Development Core Team, 2018) through the 'randomForest' package (Breiman, 2001) with the following tuning parameters established: number of trees of the model (*ntrees*) = 1,000, *nodesize* was set to default, and *mtry* = 1/3 of the number of predictor variables as suggested by Liaw and Wiener (2002). Although RF does not provide a final equation, the importance of the predictor variables

can be assessed through the Mean Decrease Accuracy (the greater the importance of a variable, the greater the Mean Decrease Accuracy) as the less valuable variable is left out of the model, while other variables are maintained (Breiman, 2001; Liaw and Wiener, 2002).

For the spatial prediction of PM, the pXRF elemental contents of Al, Fe, Si, and Ti, and MS data collected at 39 locations on a regular 173 m grid were interpolated to the whole extension of the study area via Multilevel B-splines through the System for Automated Geoscientific Analysis (SAGA GIS) software (Conrad et al., 2015) with a spatial resolution of 10 m. This procedure was performed separately for samples collected from A and B horizons, thus creating maps of elemental contents and MS for each horizon. After the spatial rendering, the RF PM prediction model was applied to each pixel, creating the PM spatial-predicted maps (one per horizon, A and B).

For the validation of the predicted maps, 15 additional places where soil PM could be accessed (Figure 1) were adopted to compare the predicted PM type shown on each map to the real PM type. To assess the accuracy of the generated maps, the following parameters were used: Cohen's Kappa coefficient (Eq. 1), the overall accuracy (Eq. 2), the producer's accuracy (probability of a location containing a type of PM being correctly predicted on the map) (Eq. 3), and the user's accuracy (probability of a place classified as a PM type on the map to match the PM type found in the field) (Eq. 4) (Congalton, 1991). Eq. 1 follows as:

$$\text{Kappa Coefficient} = \frac{\text{Po - Pe}}{\text{1 - Pe}} \quad (1)$$

$$\text{Overall accuracy} = \frac{\text{number of correct predictions}}{\text{number of validation points}} \quad (2)$$

where *Po* is the proportion of correctly classified points and *Pe* is the probability of random agreement. The Kappa Coefficient results range from -1 to 1, indicating an increasing accuracy as the values approach 1 (Landis and Koch, 1977). Eqs. 3 and 34 follow as:

$$Producer\text{'s accuracy} = \frac{Xjj}{\sum_{j=1}^{r} Xij} \quad (3)$$

$$\text{User's accuracy} = \frac{\text{Xii}}{\sum_{i=1}^{r} \text{Xij}} \quad (4)$$

where *Xii* and *Xjj* indicate the number of correctly classified locations and *Xij* represents the sum of samples of a type of PM in a row (user's accuracy) or column (producer's accuracy) of a confusion matrix (Ting, 2010). The aforementioned approaches evaluated whether soil data provided by pXRF and MS can be used to predict PM and, if possible, which data (A or B horizon) provided optimal results for PM mapping. A flowchart was created to facilitate the comprehension of the methodological process adopted in this study (Figure 3).

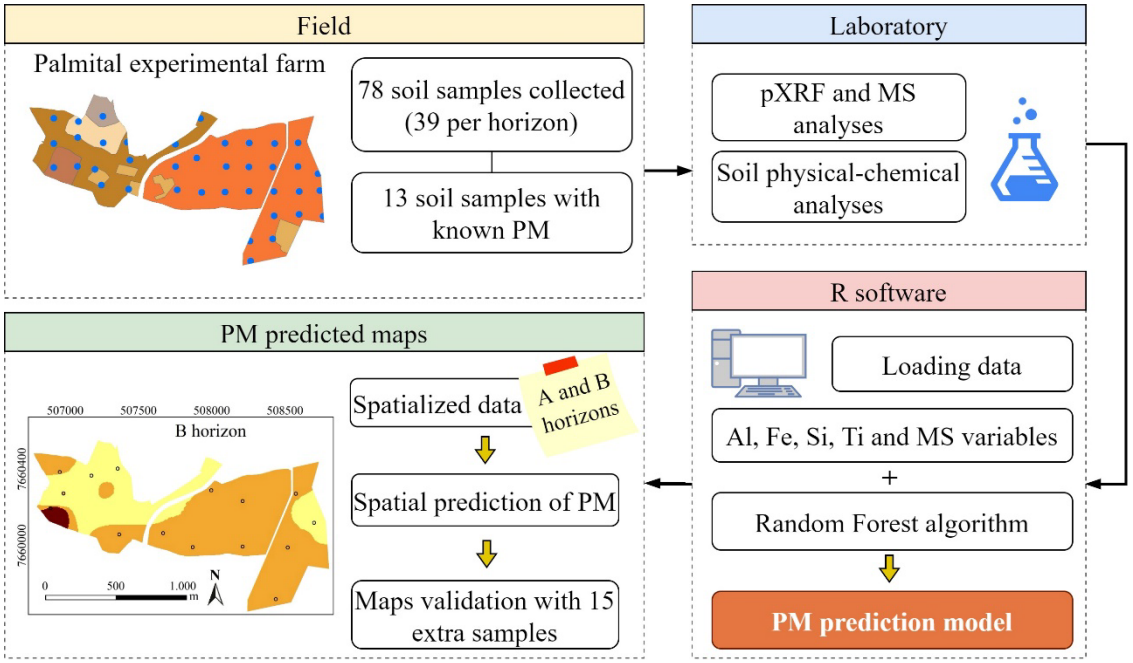

Figure 3. Flowchart illustrating the sequence of the methodological process for parent material (PM) prediction and mapping *via* portable X-ray fluorescence (pXRF), magnetic susceptibility (MS), and random forest (RF) algorithm at the Palmital Experimental Farm, Minas Gerais state, Brazil.

3. RESULTS AND DISCUSSION

3.1 Characterization of soils from different parent material *via* proximal sensors

A boxplot of the 13 soil samples used to train the prediction model is given in Figure 4. The distribution of soil elemental contents and MS emphasizes the differences between soils formed from distinct PMs. Charnockite is a hypersthene-bearing granitoid, a felsic igneous rock (Frost and Frost, 2008; Soares et al., 2014). Mudstone, a fine-grained sedimentary rock, and alluvial sediments can have variable composition (depending on the sediment origins) (Resende et al., 2019). Thus, data delivered by pXRF and MS were able to detect differences in the chemical composition of the samples (Figure 4). Mancini et al. (2019), Silva et al. (2016) and Stockmann et al. (2016) also observed differences in elemental composition of soils derived from different PMs.

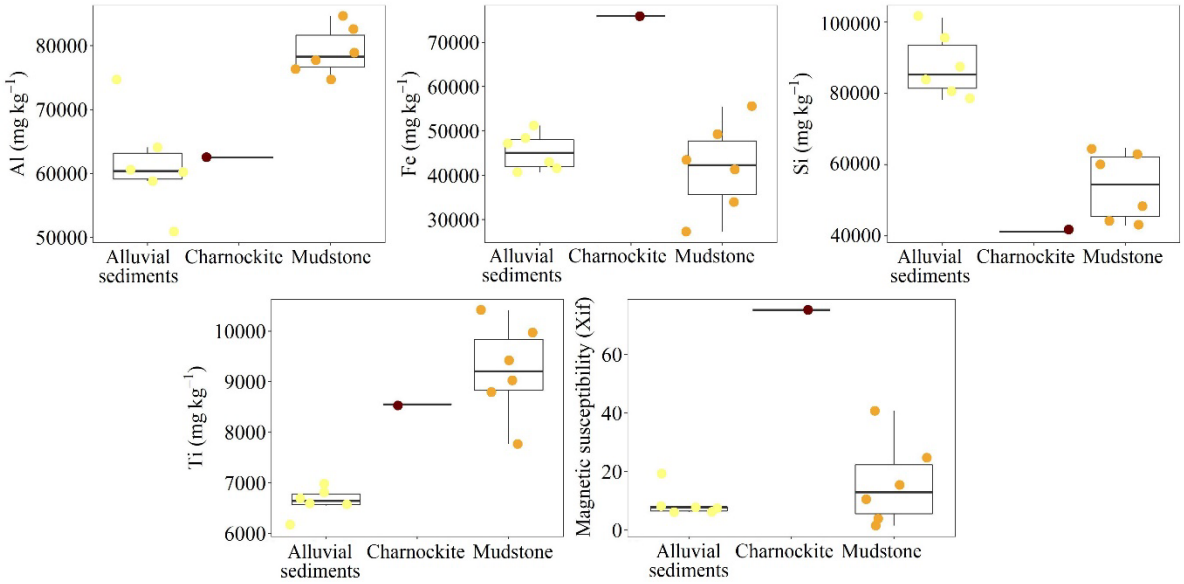

Figure 4. Training samples boxplot comparing elemental contents obtained *via* portable X-ray fluorescence (pXRF) and magnetic susceptibility (MS) of charnockite, mudstone, and alluvial sediments from the Palmital Experimental Farm, Minas Gerais state, Brazil.

The highest Fe and MS were found in the charnockite-derived soils (Figure 4). These soils may contain both maghemite (y-Fe2O3) and magnetite (Fe3O4) in the clay and sand fractions (Resende, 2005), respectively. Al and Ti had their highest contents in mudstone-derived soils. Kaolinite [(Al2Si2O5(OH)4] is the main mineral of mudstone (Whitten and Brooks, 1973) and it can also contain ilmenite (FeTiO3) and rutile (TiO2) (Resende, 2005). The highest contents of Si were found in alluvial sediments, which are dominated by silicate minerals such as quartz and kaolinite. The remarkably distinct chemical composition of soils formed from different PMs is a key factor enabling the random forest algorithm to make accurate predictions of PM type and indicate how relevant an element may be for the modeling process.

3.2 Assessing PM prediction maps

The performance of the prediction maps is illustrated in Figure 5. The map based on B horizon data achieved more reliable results than the map built with A horizon data. The B horizon-based map delivered an overall accuracy of 0.93 and Kappa coefficient of 0.85. Conversely, the A horizon-based map delivered an overall accuracy of 0.73 and Kappa coefficient of 0.48.

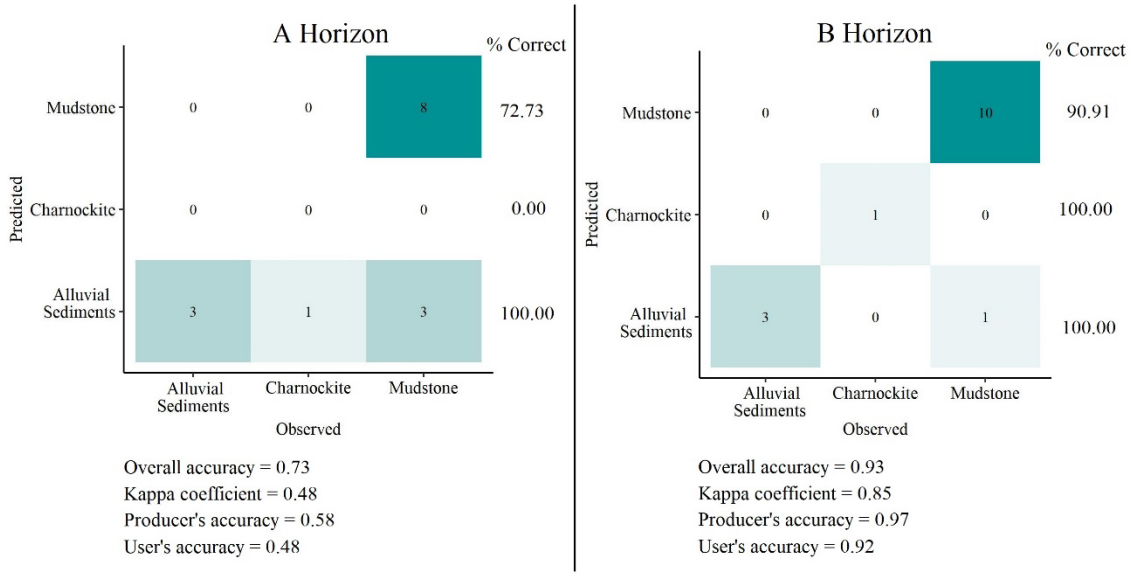

Figure 5. Confusion matrix and classification accuracy (% accuracy) for A and B horizons *via* portable X-ray fluorescence (pXRF), magnetic susceptibility (MS), and random forest (RF) algorithm for the Palmital Experimental Farm, Minas Gerais state, Brazil.

Per the confusion matrix, the A horizon-based map misclassified samples from mudstone and charnockite as alluvial sediments (Figure 5), and the predicted map showed that these misclassifications occurred in transitional areas between PMs (Figure 6). The producer's and user's accuracy did not achieve reasonable results either (0.58 and 0.48, respectively). The superficial soil horizon is more weathered and leached. Also, the soil chemical composition is susceptible to alterations caused by land use (Mancini et al., 2019a, 2019b). For these reasons, the two sedimentary formations (alluvial sediments and mudstone) in particular, could not be discriminated accurately in the A horizon. Cultivation practices disturbed the relations among the elemental contents in the superficial soil horizon compared to the natural elemental contents and their relations found in the soil PM, subsequently confusing the prediction model. In addition, soil surface contained higher amounts of organic matter, and the presence of carbon (light element) in large amounts may negatively affect the equipment measurements due to X-ray scattering and attenuation (Ravansari et al., 2020).

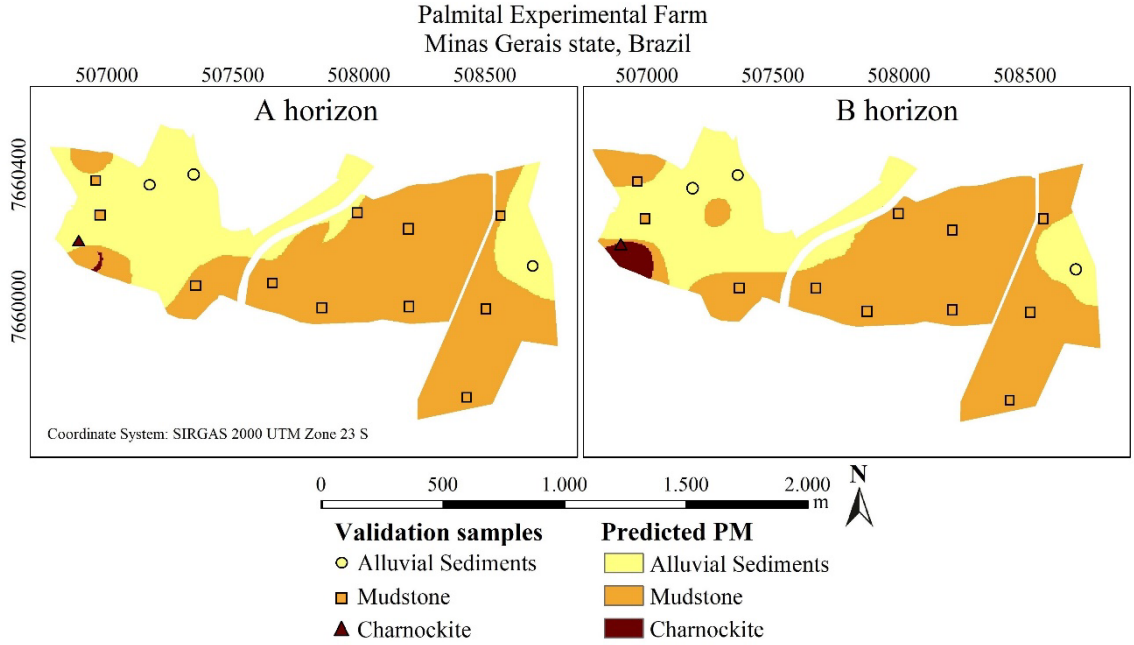

Figure 6. Validation points and parent material prediction maps for A and B horizons *via* portable X-ray fluorescence (pXRF), magnetic susceptibility (MS), and random forest (RF) algorithm for the Palmital Experimental Farm, Minas Gerais state, Brazil.

The PM maps predicted based on B horizon data were validated with a high accuracy. Alluvial sediments and charnockite achieved 100% correct sample classification and mudstone achieved 90.91% (1 out 11 samples was misclassified) (Figure 5). In tropical regions, Mancini et al. (2019a) attempted to predict and map variations of phyllite (Goethitic, Fe II and Hematitic) using pXRF data in Brazilian Cerrado. The authors reported that the most accurate prediction model (overall accuracy of 0.87 and Kappa coefficient of 0.79) was trained using PM fragment data obtained by pXRF, and the best PM predictions were made through the input of C horizon soil data. Mancini et al. (2019b) accurately predicted gabbro and gneiss parent materials (overall accuracy of 0.96 and Kappa coefficient of 0.91) using subsurficial soil data obtained by pXRF to train models and predict/map soil PM. Both studies concluded that models generated from samples collected in horizons closer to the parent material (B or C) produced reliable prediction models and yielded accurate PM maps. This trend was confirmed in this study, although it is the first time charnockite- and alluvial sediment-derived soils were predicted via this approach in Brazil.

3.3 Importance of variables

The distinct PM compositional patterns are reflected in the derived soils (section 3.1) This made the prediction model capable of differentiating soils derived from alluvial sediments,

mudstone, and charnockite. Figure 7 shows the variable importance for the prediction model. Si and Ti explanatory variables were the most important concerning prediction accuracy. Si was an important fingerprint, especially for the quartz-rich alluvial sediments. In Figure 7, the Si distribution in soil B horizons held valuable information regarding this particular PM; its distribution was almost identical to that of alluvial sediments in the predicted PM map (Figure 6).

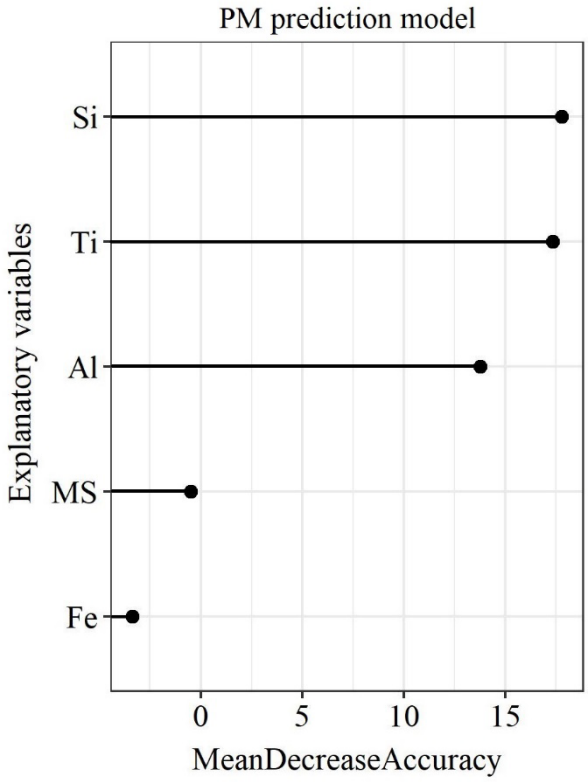

Figure 7. The importance of variables according to the random forest algorithm for predicting parent material *via* portable X-ray fluorescence (pXRF) and magnetic susceptibility (MS) for the Palmital Experimental Farm, Minas Gerais state, Brazil.

Ti was an important marker for mudstone-derived soils probably due to the presence of ilmenite and rutile in this parent material. Noteworthy, Ti had a good recovery using pXRF (0.90 and 0.94 for the NIST 2710a and 2711a standards, respectively). Thus, Ti is highly useful for predicting the bedrock because of its PM discriminating ability and accurate detection and quantification by pXRF. MS was an important marker for charnockite-derived soils, with

greater MS values where charnockite was present on the farm (Figure 8). Fe was the least important of the variables. Despite being a marker of PM's in general (Schwertmann and Taylor, 1989) and for the charnokite PM evaluated herein, it did not discriminate as well as the other variables between the sedimentary PM-derived soils that were dominant in this study (12 out of 13 control samples).

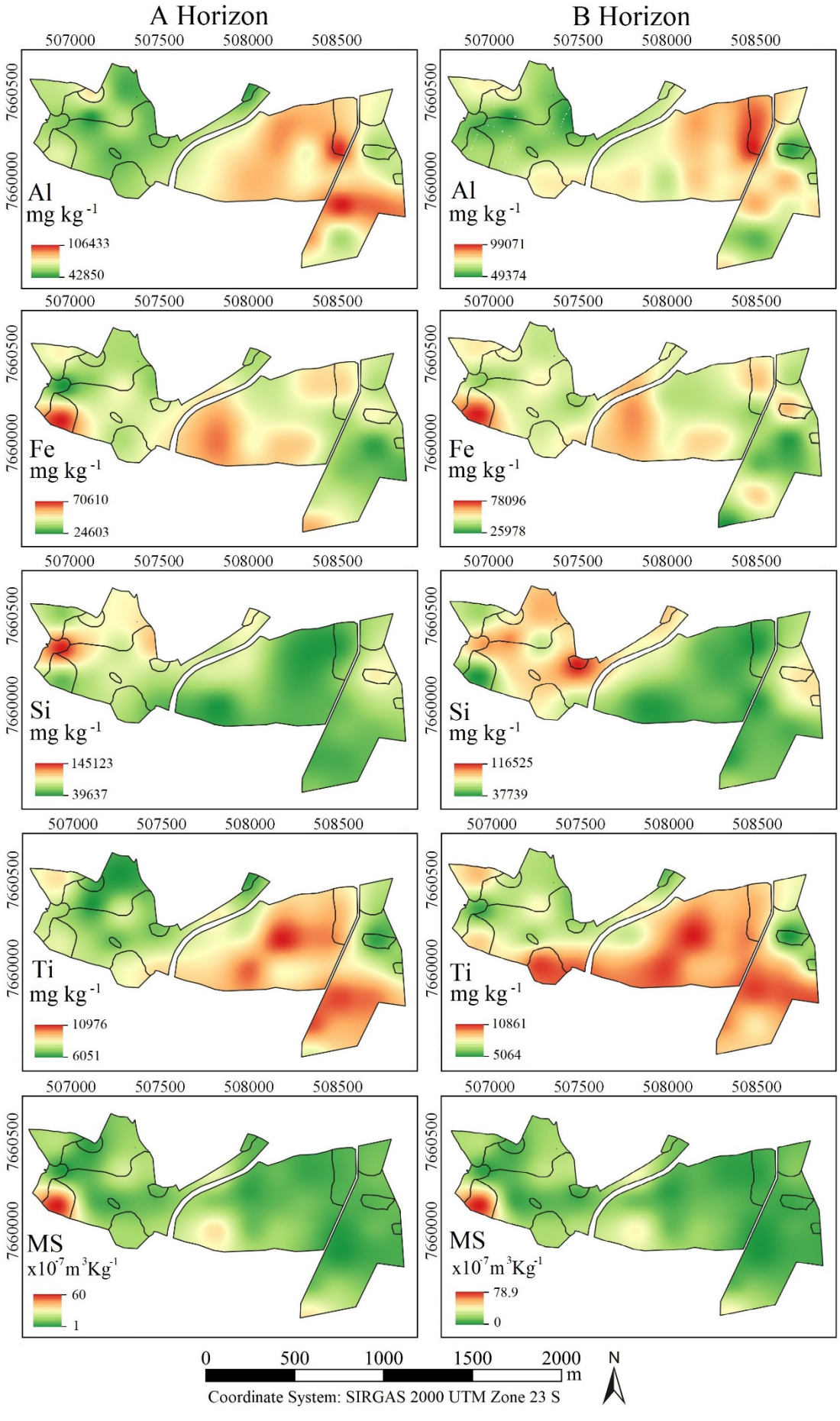
A Horizon
B Horizon
507000
507500
508000
508500
7660500
7660000
Al
mg kg -1
106433
42850
99071
49374
Fe
70610
24603
78096
25978
Si
145123
39637
116525
37739
Ti
10976
6051
10861
5064
MS
x10-7m3Kg-1
60
1
78.9
0
0
500
1000
1500
2000
m
N
Coordinate System: SIRGAS 2000 UTM Zone 23 S

Figure 8. Spatial variability of Al, Fe, Si, and Ti (mg kg^{-1}) obtained via portable X-ray fluorescence (pXRF) and magnetic susceptibility (MS) ($\times 10^{-7}$ m^3 kg^{-1}) for the A and B horizons based on 39 sampling locations the Palmital Experimental Farm, Minas Gerais state, Brazil.

3.4 Soil class and parent material relationships

The spatial maps for the five explanatory variables that were used to predict the soil PM are shown in Figure 8. MS and Ti concentrations did not vary significantly with depth (small differences between A and B horizons). The MS ranged from 1 to 60 ($\times 10^{-7}$ m^3 kg^{-1}) in A horizons and 0 to 79 ($\times 10^{-7}$ m^3 kg^{-1}) in B horizons, while Ti ranged from 6,051 to 10,976 (mg kg^{-1}) in A horizons and from 5,064 to 10,861 (mg kg^{-1}) in B horizons.

Fe, Al, and Si presented great variations with depth (Table 1) Tropical soils contain kaolinite and other Al-bearing minerals due to weathering of less stable aluminum silicate minerals (Kämpf et al., 2012), resulting in high soil Al contents.

Table 1. Descriptive statistics of elemental distribution with portable X-ray fluorescence (pXRF) spectrometer (mg kg^{-1}) and magnetic susceptibility (MS) ($\times 10^{-7}$ m^3 kg^{-1}) data per soil.

Attribute	Soil horizon	Min	Max	Mean	SD	CV (%)
Al	A	44,781	105,232	72,441	16,493.5	23
	B	49,650	98,170	70,835	12,003.1	17
Fe	A	24,863	68,014	43,567	9,604.5	22
	B	27,261	75,936	47,983	10,399.9	22
Si	A	41,149	144,397	70,634	21,794.5	31
	B	39,146	116,214	65,331	20,659.5	32
Ti	A	6,052	10,970	8,385	1,369.5	16
	B	5,144	10,829	8,369	1,598	19
MS	A	2	57	13	10.8	83
	B	1.5	75.2	15	13.8	92

The study area is comprised of three soil classes (Figure 1), which are well spread above the three predicted PMs (Figure 6). Due to different combinations of soil-forming factors, different soil classes were formed from the same parent material (e.g., Typic Haplustept, Inceptic Hapludult, and Typic Hapludox developed from alluvial sediments, but under varying

relief conditions). In addition to weathering of distinct PM's, another factor controlling the relations between chemical elements is their relative solubility as some elements are leached and others remain in soils (Schaetzl and Anderson, 2005). This factor also controls the soil classification (e.g. Rhodic Hapludox, Typic Hapludox, and Rhodic Hapludox). However, elemental variation in soils derived from the same PM coupled with varying land uses was not a constraint for the prediction model to correctly predict soil PM via pXRF and MS. This reinforces the great capability of proximal sensors to contribute to soil PM predictions even in areas with varied soil classes, weathering degree, and land uses.

Finally, the prediction model built only with B horizon data could accurately predict and map the distribution of the three different PM's found in the area, based on sufficient calibration soil samples of known PM. This approach can aid to detail not only PM maps, but also maps of soil classes and properties, due to the intricate relations between them, which is even more important in areas where reaching PM is hampered due to excessive soil thickness.

4. CONCLUSIONS

A random forest model trained using soil data provided by pXRF and MS from B horizons successfully predicted PM distribution, presenting strong validation results (Kappa coefficient = 0.85 and overall accuracy = 0.93). The hypothesis that pXRF and MS data obtained from soil samples can be used to predict the soil PM was confirmed. This can be especially important when a direct PM evaluation is not possible due to great thickness of the soil cover. Moreover, the PM prediction maps were more accurate when using B horizon data to predict soil PM as they preserve a closer soil-PM chemical relationship. In contrast, A horizon data were more prone to alterations due to land use and fertilizer management.

Relative differences of chemical/mineralogical traits in soils derived from diverse PMs were successfully detected by pXRF and MS, suggesting the potential application of these sensors in unveiling the links between distinct geological formations that underly the soils and key soil properties. Such results help to consolidate the feasibility of predicting PM using proximal sensor data in tandem with machine learning techniques, even for PM types investigated for the first time by this study.

The results in this pilot study could represent alternative methods for reducing costs and accelerating the assessment of soil PM spatial variability in tropical soils, supporting better agronomic management and sustainable decision making.

5. ACKNOWLEDGEMENTS

The authors would like to thank the National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES), and Foundation for Research of the State of Minas Gerais (FAPEMIG) for the financial support to develop this research.

6. REFERENCES

Alvares, C.A., Stape, J.L., Sentelhas, P.C., de Moraes Gonçalves, J.L., Sparovek, G., 2013. Köppen's climate classification map for Brazil. Meteorologische Zeitschrift 22, 711–728. https://doi.org/10.1127/0941-2948/2013/0507

Barbosa, J.Z., Poggere, G., Silva, S.H.G., Mancini, M., Motta, A.C.V., Marques, J.J.G. de S. e M., Curi, N., 2021. National-scale spatial variations of soil magnetic susceptibility in Brazil. Journal of South American Earth Sciences 108, 103191. https://doi.org/10.1016/j.jsames.2021.103191

Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32.

Coelho, F.F., Giasson, E., Campos, A.R., Tiecher, T., Costa, J.J.F., Coblinski, J.A., 2021. Digital soil class mapping in Brazil: a systematic review. Sci. agric. (Piracicaba, Braz.) 78, e20190227. https://doi.org/10.1590/1678-992x-2019-0227

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37, 35–46. https://doi.org/10.1016/0034-4257(91)90048-B

Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., Böhner, J., 2015. System for automated geoscientific analyses (SAGA) v. 2.1.4. Geosci. Model Dev. 8, 1991–2007. https://doi.org/10.5194/gmd-8-1991-2015

Curi, N., Franzmeier, D.P., 1987. Effect of parent rocks on chemical and mineralogical properties of some Oxisols in Brazil. Soil Science Society of America Journal 51, 153–158. https://doi.org/10.2136/sssaj1987.03615995005100010033x

Dantas, A.A.A., Carvalho, L.G. de, Ferreira, E., 2007. Classificação e tendências climáticas em Lavras, MG. Ciênc. agrotec. 31, 1862–1866. https://doi.org/10.1590/S1413-70542007000600039

Dearing, J.A., 1994. Environmental magnetic susceptibility: using the Bartington MS2 system, 1st ed. Chi Pub.

Dobos, E., 2006. Digital soil mapping: as a support to production of functional maps. Office for Official Publication of the European Communities.

Frost, B.R., Frost, C.D., 2008. On charnockites. Gondwana Research 13, 30–44. https://doi.org/10.1016/j.gr.2007.07.006

Gonçalves, M.G.M., Brant, L.A.C., Mota, R.V., Peregrino, I., Souza, C., Regina, M.A., Fruett, T., Inda, A.V., Curi, N., Menezes, M.D., 2022. Soil and climate effects on winter wine produced under the tropical environmental conditions of southeastern Brazil. OENO One 56, 63–79. https://doi.org/10.20870/oeno-one.2022.56.2.4617

Heung, B., Bulmer, C.E., Schmidt, M.G., 2014. Predictive soil parent material mapping at a regional-scale: A Random Forest approach. Geoderma 214–215, 141–154. https://doi.org/10.1016/j.geoderma.2013.09.016

Kämpf, N., Marques, J.J., Curi, N., 2012. Mineralogia de solos brasileiros, in: Pedologia: Fundamentos. SBCS, Viçosa, MG, pp. 81–145.

Khaledian, Y., Miller, B.A., 2020. Selecting appropriate machine learning methods for digital soil mapping. Applied Mathematical Modelling 81, 401–418. https://doi.org/10.1016/j.apm.2019.12.016

Lacoste, M., Lemercier, B., Walter, C., 2011. Regional mapping of soil parent material by machine learning based on point data. Geomorphology 133, 90–99. https://doi.org/10.1016/j.geomorph.2011.06.026

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159. https://doi.org/10.2307/2529310

Lawley, R., Smith, B., 2008. Digital Soil Mapping at a National Scale: A Knowledge and GIS Based Approach to Improving Parent Material and Property Information, in: Hartemink, A.E., McBratney, A., Mendonça-Santos, M. de L. (Eds.), Digital Soil Mapping with Limited Data. Springer Netherlands, Dordrecht, pp. 173–182. https://doi.org/10.1007/978-1-4020-8592-5_14

Lemière, B., 2018. A review of pXRF (field portable X-ray fluorescence) applications for applied geochemistry. Journal of Geochemical Exploration 188, 350–363. https://doi.org/10.1016/j.gexplo.2018.02.006

Liaw, A., Wiener, M., 2002. Classification and regression by random forest. R News 2, 18–22.

Loiseau, T., Arrouays, D., Richer-de-Forges, A.C., Lagacherie, P., Ducommun, C., Minasny, B., 2021. Density of soil observations in digital soil mapping: A study in the Mayenne region, France. Geoderma Regional 24, e00358. https://doi.org/10.1016/j.geodrs.2021.e00358

Ma, Y., Minasny, B., Malone, B.P., Mcbratney, A.B., 2019. Pedology and digital soil mapping (DSM). European Journal of Soil Science 70, 216–235. https://doi.org/10.1111/ejss.12790

Mancini, M., Silva, S.H.G., Teixeira, A.F. dos S., Guilherme, L.R.G., Curi, N., 2020. Soil parent material prediction for Brazil via proximal soil sensing. Geoderma Regional 22. https://doi.org/10.1016/j.geodrs.2020.e00310

Mancini, M., Weindorf, D.C., Chakraborty, S., Silva, S.H.G., dos Santos Teixeira, A.F., Guilherme, L.R.G., Curi, N., 2019a. Tracing tropical soil parent material analysis via portable X-ray fluorescence (pXRF) spectrometry in Brazilian Cerrado. Geoderma 337, 718–728. https://doi.org/10.1016/j.geoderma.2018.10.026

Mancini, M., Weindorf, D.C., Silva, S.H.G., Chakraborty, S., Teixeira, A.F. dos S., Guilherme, L.R.G., Curi, N., 2019b. Parent material distribution mapping from tropical soils data via machine learning and portable X-ray fluorescence (pXRF) spectrometry in Brazil. Geoderma 354, 113885. https://doi.org/10.1016/j.geoderma.2019.113885

McBratney, A.B., Mendonça Santos, M.L., Minasny, B., 2003. On digital soil mapping. Geoderma 117, 3–52. https://doi.org/10.1016/S0016-7061(03)00223-4

Mello, F.A.O., Bellinaso, H., Mello, D.C., Safanelli, J.L., Mendes, W.D.S., Amorim, M.T.A., Gomez, A.M.R., Poppiel, R.R., Silvero, N.E.Q., Gholizadeh, A., Silva, S.H.G., Curi, N., Demattê, J.A.M., 2021. Soil parent material prediction through satellite multispectral analysis on a regional scale at the Western Paulista Plateau, Brazil. Geoderma Regional 26, e00412. https://doi.org/10.1016/j.geodrs.2021.e00412

Oliveira, V.A., Jacomine, P.T.K., Couto, E.G., 2017. Solos do bioma Cerrado, in: Pedologia - Solos Dos Biomas Brasileiros. SBCS, Viçosa, MG, pp. 177–226.

Pelegrino, M.H.P., Weindorf, D.C., Silva, S.H.G., de Menezes, M.D., Poggere, G.C., Guilherme, L.R.G., Curi, N., 2019. Synthesis of proximal sensing, terrain analysis, and parent material information for available micronutrient prediction in tropical soils. Precision Agriculture 20, 746–766. https://doi.org/10.1007/s11119-018-9608-z

Poggere, G.C., Inda, A.V., Barrón, V., Kämpf, N., de Brito, A.D.B., Barbosa, J.Z., Curi, N., 2018. Maghemite quantification and magnetic signature of Brazilian soils with contrasting parent materials. Applied Clay Science 161, 385–394. https://doi.org/10.1016/j.clay.2018.05.014

Quéméneur, J., Ribeiro, A., Paciullo, F.V.P., Trouw, R.A.J., Heilbron, M., Valença, J.G., 2003. Geologia da Folha Lavras 1:100.000. Geologia e Recursos Minerais do Sudeste Brasileiro, 1st ed. Companhia Mineradora de Minas Gerais, Secretaria de Desenvolvimento Econômico, Belo Horizonte, Minas Gerais.

R Development Core Team, 2018. R: A language and environmental for statistical computing. R Foundation for Statistical Computing.

Ravansari, R., Wilson, S.C., Tighe, M., 2020. Portable X-ray fluorescence for environmental assessment of soils: Not just a point and shoot method. Environment International 134, 105250. https://doi.org/10.1016/j.envint.2019.105250

Resende, M., 2005. Mineralogia de solos brasileiros: interpretação e aplicações. UFLA.

Resende, M., Curi, N., Rezende, S.B. de, Silva, S.H.G., 2019. Da rocha ao solo: enfoque ambiental. Editora UFLA, Lavras.

Resende, M., Curi, N., Rezende, S.B., Corrêa, G.F., Ker, J.C., 2014. Pedologia: base para distinção de ambientes, 6.ed. ed. Editora UFLA, Lavras, MG.

Richter, J., Owens, P.R., Libohova, Z., Adhikari, K., Fuentes, B., 2019. Mapping parent material as part of a nested approach to soil mapping in the Arkansas River Valley. CATENA 178, 100–108. https://doi.org/10.1016/j.catena.2019.02.031

Santos, H.G. dos, Jacomine, P.K.T., Anjos, L.H.C. dos, Oliveira, V.Á. de, Lumbreras, J.F., Coelho, M.R., Almeida, J.A. de, Filho, J.C. de A., Oliveira, J.B. de, Cunha, T.J.F., 2018. Sistema brasileiro de classificação de solos, 5th, revista e ampliada ed. Embrapa Solos, Brasília.

Schaetzl, R., Anderson, S., 2006. Soil-Genesis and Geomorphology. https://doi.org/10.1017/CBO9780511815560

Schwertmann, U., Taylor, R.M., 1989. Iron oxides, in: Dixon, J.B., Weed, S.B. (Eds.), SSSA Book Series. Soil Science Society of America, Madison, WI, USA, pp. 379–438. https://doi.org/10.2136/sssabookser1.2ed.c8

Silva, S.H.G., Poggere, G.C., Menezes, M.D. de, Carvalho, G.S., Guilherme, L.R.G., Curi, N., Godinho Silva, S.H., Poggere, G.C., de Menezes, M.D., Carvalho, G.S., Guilherme, L.R.G., Curi, N., 2016. Proximal sensing and digital terrain models applied to digital soil mapping and modeling of Brazilian Latosols (Oxisols). Remote Sensing 8, 614. https://doi.org/10.3390/rs8080614

Silva, S.H.G., Ribeiro, B.T., Guerra, M.B.B., de Carvalho, H.W.P., Lopes, G., Carvalho, G.S., Guilherme, L.R.G., Resende, M., Mancini, M., Curi, N., Rafael, R.B.A., Cardelli, V., Cocco, S., Corti, G., Chakraborty, S., Li, B., Weindorf, D.C., 2021. Chapter One - pXRF

in tropical soils: Methodology, applications, achievements and challenges, in: Sparks, D.L.B.T.-A. in A. (Ed.), . Academic Press, pp. 1–62. https://doi.org/10.1016/bs.agron.2020.12.001

Soares, C.C.V., Varajão, A.F.D.C., Varajão, C.A.C., Boulangé, B., 2014. Mineralogical, micromorphological and geochemical transformations in the initial steps of the weathering process of charnockite from the Caparaó Range, southeastern Brazil. Journal of South American Earth Sciences 56, 30–40. https://doi.org/10.1016/j.jsames.2014.08.005

Soil Survey Staff, 2014. Keys to Soil Taxonomy, 12th ed. USDA, Washington, DC.

Stockmann, U., Cattle, S.R.R., Minasny, B., McBratney, A.B., 2016. Utilizing portable X-ray fluorescence spectrometry for in-field investigation of pedogenesis. Catena 139, 220–231. https://doi.org/10.1016/j.catena.2016.01.007

Thompson, R., Oldfield, F., 1986. Environmental magnetism. Springer Netherlands, Dordrecht. https://doi.org/10.1007/978-94-011-8036-8

Ting, K.M., 2010. Confusion Matrix, in: Encyclopedia of Machine Learning and Data Mining.

Trant, P.L.K., Kristiansen, S.M., Christiansen, A.V., Wouters, B., Sindbæk, S.M., 2021. Sampling density and spatial analysis: a methodological pXRF study of the geochemistry of a Viking-Age house in Ribe, Denmark. Archaeological and Anthropological Sciences 13, 21. https://doi.org/10.1007/s12520-020-01243-7

Vaudour, E., Costantini, E., Jones, G.V., Mocali, S., 2015. An overview of the recent approaches to terroir functional modelling, footprinting and zoning. SOIL 1, 287–312. https://doi.org/10.5194/soil-1-287-2015

Wadoux, A.M.J.-C., Minasny, B., McBratney, A.B., 2020. Machine learning for digital soil mapping: Applications, challenges and suggested solutions. Earth-Science Reviews 210, 103359. https://doi.org/10.1016/j.earscirev.2020.103359

Weindorf, D.C., Chakraborty, S., 2020. Portable X-ray fluorescence spectrometry analysis of soils. Soil Science Society of America Journal 84, 1384–1392. https://doi.org/10.1002/saj2.20151

Whitten, D.G.A., Brooks, J.R.V., 1973. The Penguin Dictionary of Geology. Penguin Books, Harmondsworth, England.

Wilson, M.J., 2019. The importance of parent material in soil classification: A review in a historical context. Catena 182, 104131. https://doi.org/10.1016/j.catena.2019.104131

ARTICLE 2 - SOIL GEOCHEMISTRY TOWARDS LITHIUM PEGMATITE EXPLORATION: BUILDING A PREDICTIVE MODEL FROM MACHINE LEARNING VIA PORTABLE-XRF*

*This article was submitted to Journal of Economic Geology

Luiza Maria Pereira Pierangeli[a], Mona-Liza C. Sirbescu[b,*], Sérgio Henrique Godinho Silva[a], David C. Weindorf[a,c], Thomas R. Benson[d,e], Nilton Curi[a]

Universidade Federal de Lavras, Department of Soil Science. P.O. Box 3037. Zip Code 37200-900, Lavras, Minas Gerais, Brazil

[b]Department of Earth and Atmospheric Sciences, Central Michigan University, Mount Pleasant, MI, 48859, USA.

[c]School of Earth, Environment, and Sustainability, Georgia Southern University, Statesboro, GA, USA

[d]Lithium Americas (Argentina) Corp., Vancouver, BC, Canada

[e]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA

*Corresponding author: sirbe1mc@cmich.edu; Tel.: +1 989.774.4497

ABSTRACT

As the demand for lithium (Li) increases, less costly, more sustainable, and faster exploration methods are needed for the identification and characterization of new Li deposits. Li-rich pegmatites are major sources of Li, but their exploration is often hindered by soil cover. Portable X-ray fluorescence (pXRF) can rapidly and accurately quantify the chemical composition of soils to determine the economic potential of bedrock, but unfortunately, Li is nondetectable via pXRF. Herein, pXRF data and *Random Forest* algorithms were used to predict both Li contents in soil samples and Li-rich soil parent material based on concentrations of 15 predictors (K:Rb, Al, Ba, Ca, etc.). 112 soil samples were collected overlying spodumene-rich pegmatites, barren granitic pegmatites, peraluminous granite, and metamorphic host rocks from forested, glaciated northern Wisconsin and Michigan, USA. Li concentrations were independently measured using ICP-OES. The best Li prediction based on the predictors quantified via pXRF achieved accurate results, yielding a coefficient of determination (R^2) = 0.86, a root mean square error = 68.5 mg·kg^{-1}, and residual prediction deviation = 1.78. The parent material prediction model achieved a Kappa coefficient of 0.77 and an overall accuracy of 0.85. The pXRF analysis was able to delineate differences among soil samples formed on bedrock with distinct mineralogy. Using pXRF to determine the Li contents in the soil and the type of underlying bedrock could become an alternative, more sustainable exploration method. Further systematic studies are necessary to better constrain the factors controlling soil geochemistry and the relationships to their parent material.

Keywords: Principal component analysis; *Random Forest*; LCT pegmatites; lithium pathfinders; lithium anomalies; exploration; Florence County pegmatites, Wisconsin, USA.

1. INTRODUCTION

With increasing dependence on energy storage in support of electric transportation, the importance of Li has grown due to the exponential increase in the production of rechargeable Li-ion batteries. Finding new lithium-cesium-tantalum (LCT) pegmatites is vital for current demands and to ensure coverage of future needs for Li and other critical metals, such as Cs, Rb, Sn, Ta, and Nb (Linnen et al., 2012; USGS, 2023). Exploration for LCT-pegmatites is expensive and time-consuming, especially given their relatively small size and masking by overburden materials. For example, exploration in vegetated, glaciated terrains from wet temperate climates is difficult because of poor bedrock exposure. In addition, pegmatite properties (color, reflectivity, density, magnetism, electric conductivity, etc.) do not typically bear sufficient

contrast relative to their host rocks and between mineralized and barren bodies, hindering remote-sensing and geophysical exploration methods (Cardoso-Fernandes et al., 2021; Köhler et al., 2021; Müller et al., 2022).

Prior investigations targeting soil, laterite, glacial till, or stream-sediment geochemistry were moderately successful in the discovery of buried mineralized pegmatites (Luecke, 1984; Smith et al., 1987; Selway, 2005; Galeschuk and Vanstone, 2005; Turner and Young, 2008; Steiner, 2019a; Xu et al., 2019). However, conventional bulk analysis of digested samples using benchtop instrumentation such as ICP (Inductively Coupled Plasma) – Mass Spectrometry (MS) or Optical Emission Spectroscopy (OES) can be time-consuming and cost prohibitive. The use of corrosive acids or fluxes for acid digestion could also be banned by some institutions or governments for safety or environmental reasons. Thus, developing inexpensive, environmentally friendly, and faster exploration methods is crucial for improving the identification and characterization of new Li-rich pegmatite deposits. Portable Laser Induced Breakdown Spectroscopy (pLIBS) has gained popularity for its ability to detect and quantify elements of interest for LCT pegmatite exploration (including Li) in single minerals and whole rock (e.g., Sweetapple and Tassios, 2015; Wise et al., 2024). However, few attempts to apply LIBS to soils as a Li exploration have been made (Wise et al., 2022), in part because the samples would have to be dried, pulverized or sieved, and pelletized prior to LIBS testing, and specialized factory calibrations would be required (Harmon and Senesi, 2021).

Portable X-ray fluorescence (pXRF) has been widely used as a rapid and low-cost analytical technique in different fields of science and for various purposes (Burley et al., 2017; Finnigan and Golitko, 2023; Han et al., 2021; Lima et al., 2023; Pierangeli et al., 2023), including the rapid and nondestructive analysis of soils and rocks (Weindorf and Chakraborty, 2020; Lemiere, 2018). Although numerous studies used pXRF to better understand pedological processes, weathering, and contamination in soils (Stockmann et al., 2016; Duda et al., 2017; Gozukara et al., 2021; Silva et al., 2021; Naimi et al., 2022; Zhao et al., 2022), only a few studies have been published involving pXRF analysis of soils applied to Li-pegmatite exploration (Steiner, 2019a,b; Batt et al., 2021). This is in part because pXRF cannot detect Li; it can detect and quantify pathfinder elements heavier than Mg associated with Li in LCT pegmatites, such as Rb, K/Rb ratio, Cs, Ba, Sr, Cu, Zn, Pb, Ga, Sn, Ta, ± W (Brand and Brand, 2017; Steiner, 2019b; Phelps-Barber et al., 2022).

However, proximal-sensor analysis of soils derived on LCT pegmatites and host rocks within their metasomatic aureole (e.g., Barros et al., 2022) has potential to fill-in gaps in understanding of the paths followed by Li and Li-pathfinder elements during weathering and

soil formation. Indeed, fundamental and applied aspects of the formation of a geochemical dispersion halo in the soils formed from a LCT pegmatite parent material could be derived from soil chemical composition. Scientifically, understanding weathering processes that disperse/concentrate Li and Li-pathfinder elements within the soil profile constitutes and essential step in constraining the mass transfer of rare metals from their hard-rock sources to surficial environments (e.g., Gloaguen et al., 2023). Economically, a pathfinder-based Li prediction using proximal sensors constitutes a more sustainable and rapid exploration approach that could be of interest to exploring and mining companies (Müller et al., 2022).

The aim of this study is to develop a pilot exploration methodology applicable to specific weathering patterns of forested terrains in a wet-temperate climate. We anticipate that Li and its pathfinders are present in immature, shallow soils from North America and other glaciated terrains as detrital mineral fragments, mobile species, or species fixed to clays or organic matter. Furthermore, we hypothesize that those elements and other major components in soils quantifiable via pXRF (Al, Si, K, Ca, and Fe) can serve as adequate predictors of Li contents and Li-rich parent materials. Specifically, we assess the use of machine learning, Random Forest algorithms to predict 1) Li contents in soil samples with diverse parental materials and 2) their Li-rich parental materials. Three models were tested to evaluate the efficacy of Li prediction using pXRF data from soils on top and in the vicinity of simple and mineralized pegmatites. A fourth model was created to discriminate between Li-rich and Li-poor parental materials, based on the same input data.

2. MATERIALS AND METHODS

2.1 Study Areas

Soil and rock samples were collected from 1) the Li-rich Florence County pegmatite field (Falster et al., 1996, Sirbescu et al., 2008, Liu et al., 2010) situated near the northernmost edge of the Paleoproterozic Wisconsin Magmatic Terrane in northeastern Wisconsin, USA (Zi et al., 2022) and 2) barren granitic pegmatites hosted by Paleoproterozoic Dickinson Group rocks, at the southernmost edge of the Superior Province (Sims, 1980; Ayuso et al., 2018; Cannon, 2018), in central Dickinson County, Michigan (Fig. 1). Although located in a distinct tectonic province, the Dickinson County pegmatites (DCP) were chosen as a control for this study, as the closest field of mineralogically simple, non-mineralized pegmatites located ~30 km NE from the mineralized Florence County pegmatites (FCP). The FCP and DCP soils are optimal for this study because the pegmatites have never been mined and were only minimally affected by anthropogenic activity (logging and mineral collecting). The region was affected by

prolong glaciation, starting at least ~780,000 years ago, with the Wisconsin event ending in the area ~10,000 years BP, leaving behind a characteristic hummocky terrain with glacial till and abundant erratics. When present, the rock outcrops have typical glacial polish, striations, and grooves. Characteristic of postglacial, young soils, those in north-eastern Wisconsin and north-western Michigan underwent minimal weathering and closely resemble their parent materials (Schaetzl, 2009). The humid continental climate in the region is classified as Dfb (Köppen), featuring the "warmest" month with an average temperature below 22 °C and the coldest month with an average temperature below freezing, 0 °C (CEC, 2021).

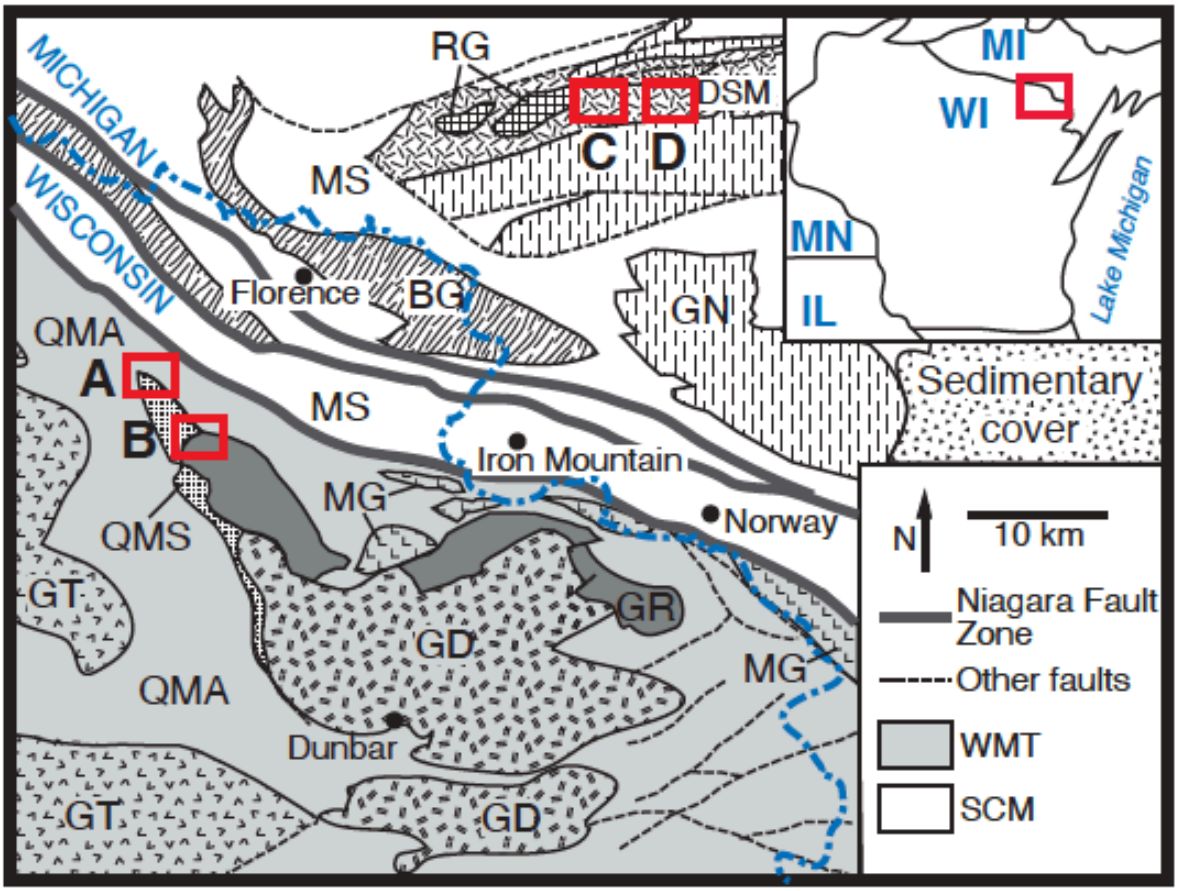

Figure 1. Simplified regional geology of the study areas near the Wisconsin (WI) - Michigan (MI) border (blue dash-dotted line and top right corner inset). Other states in the inset: Minnesota (MN) and Illinois (IL). Sampled areas (red rectangles): A: Florence County Pegmatite field, B: Bush Lake Granite, and C: Dickinson County Pegmatite field. Major tectonic provinces separated by the Niagara Fault Zone (thick gray lines): Wisconsin Magmatic

Terrane (WMT) and Superior Craton Continental Margin (SCM). Other major lithologic units: QMA – Quinnesec mafic amphibolite; QMS – Quinnesec micaschist; MG – metagabbro; GR - Hoskin Lake and Spikehorn Creek granites; GD – other granitoids: quartz diorite, tonalite, and gneiss; GT – granites and tonalites; GN – gneisses; MS – Michigamme Schist; DSM – Dickinson Six-Mile amphibolite; RG – porphyritic red granite; BG – Badwater Greenstone. List sources (simplified after Sims, 1992).

2.2 Local Geology, Mineralogy, and Geochemistry

The local geology and mineralogy of the rocks and brief soil descriptions are given in Table 1. The FCPs are subparallel thin dikes, typically <5 m wide; with one exception reaching 8 m in width. The dikes trend NS to NW-SE, and dip W to SW, intruding discordantly into mafic and felsic metavolcanics and metasediments of the Quinnesec formation (Dutton and Bradley, 1970; Sims et al., 1992). Four mineralized (spodumene-bearing) pegmatites were included in this study; the Animikie Red Ace (ARA), King's-X (KX), King's-X2 (KX2), and Price Lake (PL). Based on the consistent presence of spodumene, the FCPs have a lithium-cesium-tantalum (LCT) petrogenetic affiliation (Černý and Ercit, 2005; Bradley et al., 2017). A two-mica Bush Lake (BL) granite (Sims et al., 1992), located < 3 km southeast of the exposed pegmatites was also included in the study, as a potential parental granite of the FCPs (Fig. 1).

Table 1. Location, abbreviation, coordinates, rock and host rock types, soil type, and soil horizon collected with n, their respective number of soil samples. Host rock abbreviations: QMS – Quinnesec quartz micaschist; QMA – Quinnesec mafic amphibolite; DSM – Dickinson Six-Mile Lake amphibolite. Soil Series: RmC and RmD – Metonga-Sarona soil complex formed on rock outcrops with 1 to 15% slope and 15-35% slope, respectively; SaB – Sarona fine sandy loam, formed on glacio-fluvial sediment; PeB – Padus-Pence sandy loams; H69D – Emmet soil on rock outcrops; and 13D – Pemene fine sandy loam soil.

Sampling locations	Coordinates East North	Magmatic bedrock (n)	Host bedrock (n)	Soil order	Soil series	Soil horizon
Florence County Pegmatites (FCP)						
Animikie Red Ace (ARA)	3950305078334	LCT-peg (6)	QMS (24)	Spodosol	RmD	A and E
Kings-X2 (KX2)	3947575078458	LCT-peg (7)	QMA (19)	Spodosol	RmC	A and E
Kings-X (KX)	3946805078660	LCT-peg (4)	QMA (10)	Spodosol	PeB	A and E
Price Lake (PL)	3924955079234	LCT-peg (6)	QMA (16)	Spodosol	RmD	O and A
Bush Lake (BL)	3955755073770	Granite (3)	---	Spodosol	SaB	O and E
Dickinson County Pegmatites (DCP)						
Six-mile Lake (6MI)	4297495096470	Barren peg (3)	DSM (5)	Alfisol	H69D	O and A
Sturgeon River Quarry (SRQ)	4200335096098	Barren peg (4)	DSM? (5)	Alfisol	13D	A and B

The Animikie Red Ace (ARA) pegmatite system consists of a main dike with thinner offshoots emplaced discordantly in a fine-grained quartz mica schist (QMS), belonging to the Paleoproterozoic Quinnesec formation (Sims, 1992). At its widest known exposure, the main ARA dike is <3 m in width. Although prior researchers reported a length of ~600 m (Falster et al., 1996), the ARA dike could be traced with certainty along strike for only ~100 m. The main dike thins out to less than 0.5 m in width at both southern and northern ends (Fig. 2A), before being completely covered by soil. Mineralogically, the ARA is a highly fractionated, zoned, granitic pegmatite with abundant raspberry-colored elbaite tourmaline (rubellite var.; Fig. 2B),

lepidolite, and clusters of spodumene partially to completely replaced by albite-mica (Falster et al. 1996; Sirbescu et al. 2008; Sirbescu et al., 2009). The K-feldspar (microcline) in the ARA pegmatite contains 5,000 to 10,000 $mg \cdot kg^{-1}$ Rb and extremely low K:Rb ratios ranging from ~11 to 22 (Falster et al., 1996). Lepidolite mica forms discontinuous bands and masses 20 x 1200 cm in size and contains ~10,000 to 15,000 $mg \cdot kg^{-1}$ of Rb reaching a K:Rb ratio as low as 5 (Falster et al., 1996). Accessory amounts of cobalt-blue manganoan apatite-(CaF) with up to 12.7 wt% MnO (Sirbescu et al., 2009) and small phosphate pods rich in lithiophilite-triphyllite - $Li(Mn,Fe)[PO_4]$ (Falster et al., 1996) highlight the unusually high Mn contents of the ARA pegmatite. Biotite, quartz, and plagioclase are the main constituents of the host QMS, with minor calcite, muscovite, and amphibole. Abundant tourmaline, apatite, and microcline visible within ~5 cm from contact with pegmatite are a consequence of pegmatite fluids infiltrating the schist (Sirbescu et al., 2008). Liu et al. (2010) demonstrated that Li infiltrated as far as 25 m from the main dike using Li and Li-isotope analyses.

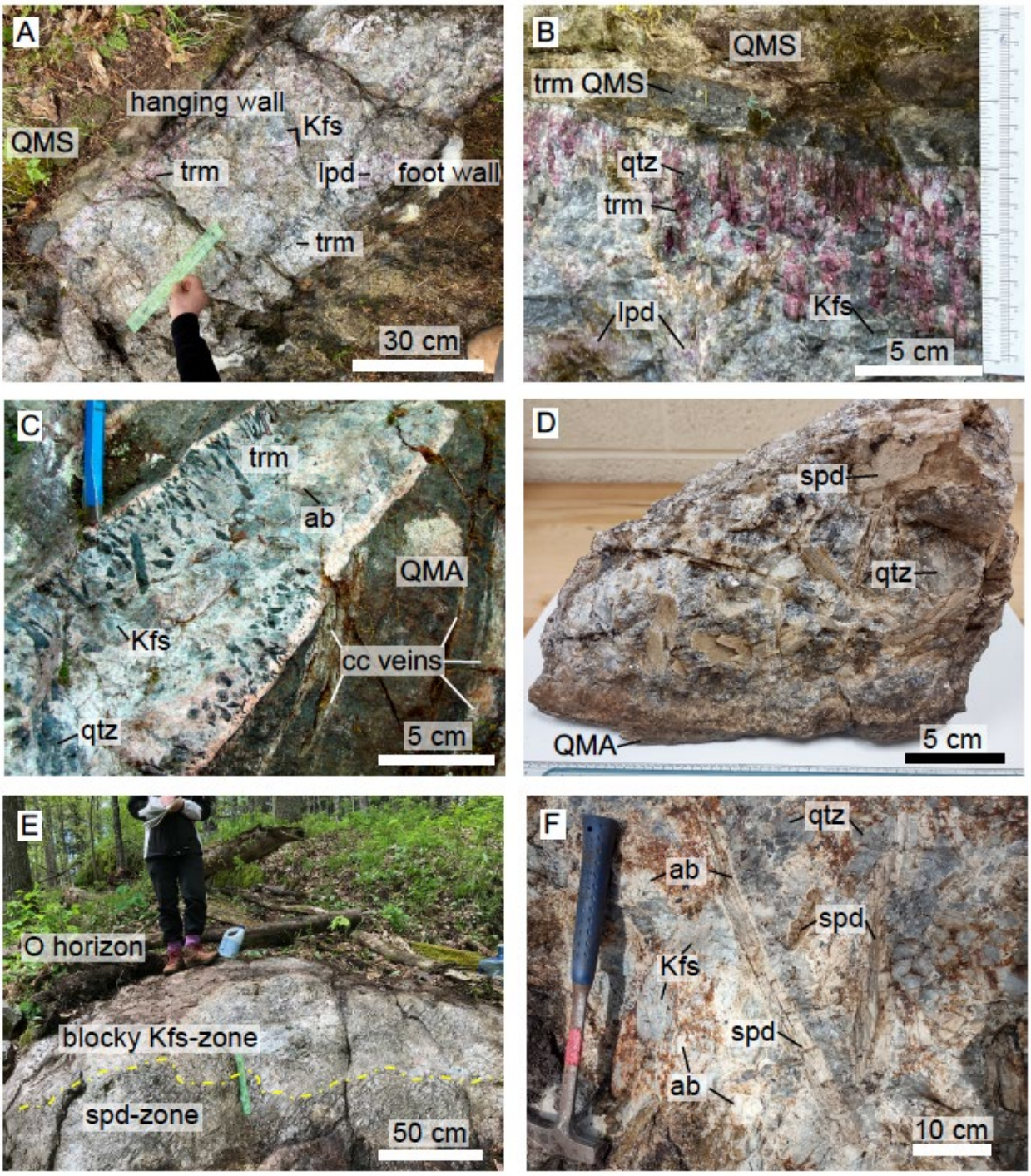

Figure 2. Examples of LCT pegmatite collection sites and mineral assemblages. Mineral abbreviations following Whitney and Evans (2010): ab – albite, Ksp – K-feldspar, lpd – lepidolite, qtz – quartz, spd – spodumene, and trm – tourmaline. A. Overview of the Animikie Red Ace (ARA) dike, narrowed to ~50 cm in width, near its northern end in contact with the Quinnesec quartz micaschist (QMS), covered by thin spodisol, B. ARA comb elbaite tourmaline (var. rubellite), < 7 cm long, nucleated on the sub-concordant contact with the QMS host rock. Despite weathering, a dark band of tourmalinization (trm QMS), ~2 cm wide, is present in the exocontact. C. King's X2 (KX-2) pegmatite finger (~20 cm wide) with feldspathic

fine-grain border and comb schorl tourmaline on the contact with mafic amphibolite (QMA). Notice the calcite lenses or veins (cc), partially dissolved by weathering. D. KX-2 spodumene-rich assemblage at the contact with hanging wall QMA. E. Overview of Price Lake (PL) pegmatite outcrop and soil cover. F. Spodumene in the footwall of the PL dike. Long-prismatic crystals of spodumene rimmed by thin white albite-muscovite and surrounded by dark-gray quartz, light gray K-feldspar, and white albitite masses and stringers (in part covered by iron stain). Contact with QMA (not shown) located ~30 cm to the left of the image.

The King's-X (KX) is a 3 m wide pegmatite dike exposed in a road cut (Falster et al., 2005; Liu et al., 2010). The contacts with the host amphibolite and biotite schists (QMA), a more mafic member of the Quinnesec formation, are sharp and discordant. KX is a spodumene – amblygonite-montebrasite $LiAl[PO_4](F,OH)$ bearing pegmatite dominated by feldspar and quartz, with accessory muscovite and dark blue tourmaline. Spodumene is relatively fine-grained and partially replaced. The dike could not be followed along the strike because of soil coverage (agricultural field to the north) and debris from recent logging to the south. Liu et al. (2010) described the QMA amphibolites to consist of amphibole, plagioclase, quartz, biotite, chlorite, and calcite. The QMA biotite schist lacked the amphibole, but contained chlorite, tourmaline, and epidote. Calcite layers and quartz veins were common. Tourmaline and minor fluorite veinlets were found within 0.5 m of the contact with KX, and Li traveled at least 30 m away from the contact into the host rock (Liu et al., 2010).

The King's-X2 (KX2) is a 0.5 to 4 m wide dike hosted by QMA with similar mineralogy but better exposures than KX, hypothesized to be a continuation of the KX outcrop (Hockemeyer et al., 2010). Based on recent studies, KX2 crops out discontinuously for ~35 m (Pierangeli et al., 2022; Cox et al., 2023). The KX2 is also highly fractionated and internally zoned (Fig. 2C), with irregular mineral zones distributed discontinuously along the strike, but, unlike the ARA pegmatite, KX2 lacks elbaite tourmaline and lepidolite. Clusters of coarse, bladed spodumene occur locally in the footwall zone, in a coarse matrix of dark gray K-feldspar, gray quartz, and coarse albite (cleavelandite) (Fig. 2D), with accessory schorl tourmaline. Oval-shaped pods of amblygonite-montebrassite are also commonly found in the footwall zone. Granular to short prismatic spodumene is present in meter-size zones occurring discontinuously along the core of the pegmatite. Accessory white to purple mica and beryl were revealed in the core. The spodumene from KX2 suffered only minor alteration and rimming by albite-mica ± quartz intergrowth (Meldrum et al., 2023), suggesting that the initial magmatic Li was less re-distributed by secondary processes than at the ARA. The KX2 is hosted by QMA with similar

lithology as at KX (see description for the KX pegmatite). Preferentially dissolved calcite layers/veinlets in the QMA amphibolite display are abundant (Fig. 2D).

Finally, the Price Lake (PL) pegmatite is a newly discovered spodumene-bearing dike (Pierangeli et al., 2022) located east of Price Lake, Florence Co., WI, also hosted by QMA. Falster et al. (2018) mentioned the presence of spodumene and/or petalite pegmatites in the area north of Price and Patten lakes, but the conference abstract does not include the geographic coordinates to confirm whether PL was part of Falster et al. (2018)'s study. The PL dike has an average of 2 m width but can reach up to 8 m. Similar to the KX2, PL is dominated by K-feldspar, quartz, and albite forming a blocky K-feldspar zone with accessory spodumene, black to green tourmaline, dark mica (zinnwaldite), and white beryl (Fig. 2E). A zone rich in granular to short prismatic spodumene in a matrix of quartz, K-feldspar, and albite occurs discontinuously towards the footwall of the dike (Fig. 2E). Sporadic clusters of slender spodumene crystals, up to 0.5 m long, partially replaced by thin albite-muscovite rims were found in the wall zone (Fig. 2F), < 20 cm from the hanging contact with the host QMA (Cox et al., 2023).

The BL granite is a poorly exposed ~3 x 10 km body located < 3 km away from the FCP pegmatites, in Florence County, northern Wisconsin. It is a biotite-muscovite peraluminous granite to alkali-feldspar granite and the most fractionated granite in the area (Sims et al., 1992). Accessory minerals include zircon and apatite associated with biotite. It has an aluminum saturation index (ASI) value of 1.12 (data from Sims et al., 1992), with ASI defined as Al/(Ca - 1.67P + Na + K) (Frost and Frost, 2008). Soils formed on BL granite were sampled in this study as the most probable parental granite to the FCP pegmatites, based on its distinct composition, slightly enriched in SiO_2 and more fertile composition compared to the other granitoids in the area (e.g., Hoskin Lake and Spikehorn Creek granites; Sims et al., 1992). The BL granite has the highest Rb:Sr (avg. of 2.35) and the lowest K:Rb and K/Cs ratios, on average 80 and 3,214, respectively (data from Sims et al., 1992), that are well within the range of Precambrian fertile granites (Selway et al., 2005; and references therein).

The Dickinson County pegmatites (DCP) sampled in this study include the Sturgeon River Quarry (SRQ) pegmatite and a thinner pegmatite body that was named Six-Mile Lake (6MI), after the nearby lake (Fig. 1). The SRQ is a large, elongated pegmatite body trending NE, ~140 m wide and at least 1000 m long. It is very coarse, lacks any clear internal zoning, and is mineralogically simple, consisting of pink to orange feldspar, quartz, muscovite, biotite, with accessory black tourmaline and garnet. Van Daalen (2015) reported beryl, columbite, and uraninite at SRQ and classified it as a rare-element, niobium-yttrium-fluorine (NYF) family

pegmatite, based on its REE accessory minerals. The SRQ pegmatite is located only a few hundreds of meters from (and possibly related to) an unnamed, 2.099 B.y. large 'porphyritic red granite' (Ayuso et al., 2018; Cannon, 2018). The contacts with the surrounding metamorphic rock (undifferentiated Dickinson group; Cannon, 2018) could not be observed.

Samples were also collected from the 6MI barren pegmatite intruding discordantly the Dickinson Six Mile Lake amphibolite (DSM), a member of the Dickinson group (James et al., 1961; Cannon, 2018), situated ~10 km east of SRQ. The dike is ~80 m long and at least ~3 m wide, although its contacts with the host rocks were poorly exposed. The rock samples collected at this site suggest a very simple granitic mineralogy with quartz, feldspar, and micas. Accessory schorl tourmaline was reported in a different dike from the same pegmatite field (Robinson and Carlson, 2013). The DSM amphibolite is a dark, fine-grained to medium-grained, massive to layered amphibolite (James et al., 1961). It is frequently cut by younger pods or irregular bodies of simple granitic pegmatites, or larger dikes, such as the 6MI pegmatites (Cannon et al., 2018).

2.3 Soil Samples

Soils in the study area belong to two soil orders, Spodosols and Alfisols. Only soils formed directly on bedrock were selected in this study, except for the KX pegmatite location, where the fluvioglacial sediment cover could not be avoided. The Spodosols sampled in Florence County belong to four mapping units (Table 1, Fig. 3): the RmC and RmD – Metonga-Sarona soil samples formed on rock outcrops with 1 to 15% slope and 15-35% slope, respectively, were sampled at ARA, KX2, and PL; the PeB – Padus-Pence sandy loams, represented by Padus and Pence soils, formed on a layered glacio-fluvial sediment overlying the KX pegmatite (Boelter and Elg, 2004). The BL location was part of a broader mapping unit (SaB) – Sarona fine sandy loam formed on glacial till (Boelter and Elg, 2004), however the BL soil samples collected for this study were formed directly on a BL granite outcrop. In Dickinson County, the Alfisols sampled represented map units: 13D – Pemene fine sandy loam, represented by Pemene soil and minor components overlying the SRQ pegmatite, and H69D – Emmet-Rock outcrop complex, represented by Emmet soil, rock outcrops, and minor components, overlying the 6MI pegmatite (Soil Survey Staff, 2023).

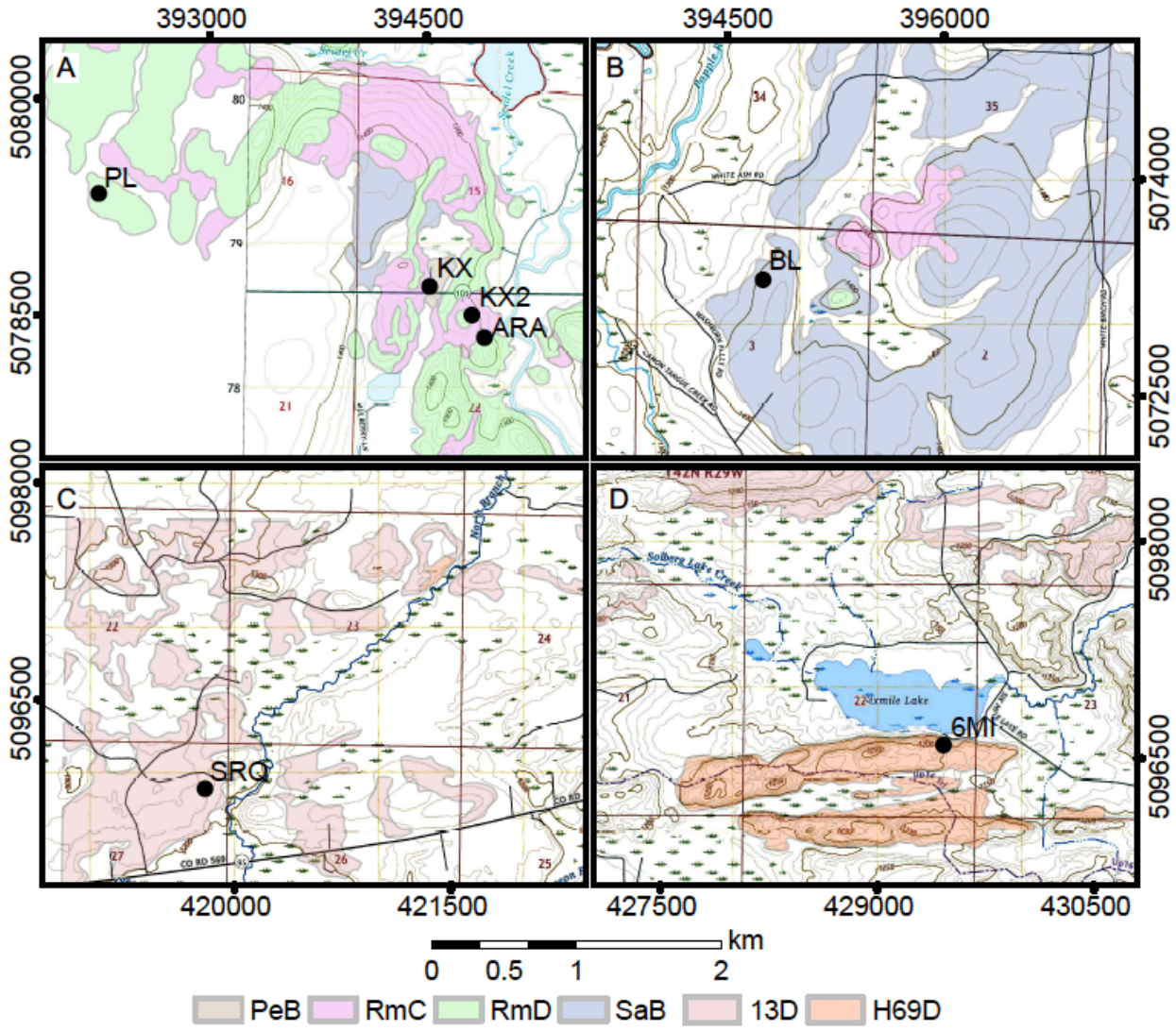

Figure 3. Soil unit maps with sampling locations. Top maps, Florence County, Wisconsin: A. ARA: Animikie Red Ace, KX2: King's-X2, KX: King's-X, PL: Price Lake, and B. BL: Bush Lake granite. Bottom maps: Dickinson County, Michigan: C. SRQ: Sturgeon River Quarry and D. 6MI: Six-mile Lake. The topographic base from U. S. Geological Survey, with elevations in units of feet. Soil map units: RmC and RmD – Metonga-Sarona soil complex formed on rock outcrops with 1 to 15% slope and 15-35% slope, respectively; SaB – Sarona fine sandy loam formed on glacial till; PeB – Padus-Pence sandy loams, formed on layered glacio-fluvial sediment; H69D – Emmet soil on rock outcrops; and 13D – Pemene fine sandy loam soil. Only soil series pertinent to this study are shown (Soil Survey Staff, 2023).

2.4 Sampling Strategies

A total of 112 soil samples were collected from the surface horizons (O, A, and E) down to a maximum of 40 cm depth. All soil sampling sites included rock outcrops. Three types of

soil samples were collected from all LCT pegmatite localities: a) soil formed on known pegmatite dikes, collected along their strike: 4 to 7 soil samples for each pegmatite, spaced from 3 to 30 m apart depending on the length of the pegmatite outcrop and soil availability; b) soil formed on the host rock sampled along traverses cutting across the strike of outcropping dikes: 5 to 14 soil samples at each location; and c) soil sampled along 'blind' traverses cutting perpendicular to the hypothetical extensions of dikes concealed under thicker overburden: 11 to 20 soil samples for each FCP pegmatite. The traverses across exposed pegmatites extended 5 to 25 m away from the contact with the pegmatite outcrop, as a function of soil availability and sampling feasibility (e.g. dense vegetation, steeply sloping terrain, or large outcrops). The blind traverses were located 10 to 20 m from the last visible pegmatite outcrop.

The distance between soil traverse points varied from 1 to 4 m, depending on the soil availability, to a total length of 7 to 40 m which overlaps, and slightly exceeds the size of the metasomatic halo produced by the pegmatite magma within the host rocks at ARA and KX (Liu et al., 2010). Although the distance from the contact to the pegmatite was recorded, the compositional variation as a function of distance from the source is not the focus of this study. At the KX pegmatite and its host rock, PeB sandy loams soils formed on glacio-fluvial sediment rather than directly on bedrock. Therefore, the soils at KX location were sampled as a representative of till-covered mineralized pegmatite. Except at the KX pegmatite, occasional soils that formed on glacial till or fluvial sediment were recognized by subangular to rounded pebbles of various lithologies and were avoided because they were not representative of local bedrock.

For the control location at BL, a large granite pluton, only soil formed on the granite rock was collected, because the contact with the host rock was not exposed. For the control location at SRQ the host rock is only tentatively assigned to DSM Lake amphibolite because of the lack of host rock outcrops at the sampling site (Table 1).

2.5 Sample Preparation

Soil aliquots were dried in an oven at ~65 °C, disintegrated using a mortar and pestle, and sieved to eliminate organic fragments >2 mm. A sample splitter was used to divide the sample into four equal aliquots. One aliquot was pulverized in a shatter box for complete homogenization and further split for pXRF analysis and ICP-OES analysis. Prior to the pXRF analysis, one homogenized soil aliquot was placed in a preassembled Spectrocup© XRF sample cup (Chemplex Industries, Inc.) covered with a Prolene® 4.0 μm thin film.

An ETHOS X (Milestone MLS, Bergamo, Italy) microwave digestion system equipped with a high-pressure segmented SK-15 rotor and 11 Teflon vessels (internal volume 100 mL) was used in this study. One vessel is used for internal temperature control. To estimate the digestion recoveries for various elements, NIST 7010A or OREAS 47a standard soils were included in each soil digestion batch. Approximately 250 mg of samples were placed inside each vessel. 6 mL of Optima-grade HNO_3 (67%, Fisher) was added to each vessel and left overnight under the corrosive-acid fume hood. This step was added to the method to guarantee that carbon of soil organic matter was released as CO_2 during the reaction with HNO_3 (Chao and Sanzolone, 1992). After this step, 3 mL of HF (70%) and 2 mL of HCl (37%) were added to each vessel. All vessels were tightly sealed and placed inside the microwave. The method used was SK-Environmental 27 – Siliceous soil with organic matter (Milestone), which consisted of a 20-minute ramp to reach 210 °C and a 15 min hold at 210 °C.

To ensure total recovery of the digestion sample, the solution was transferred to 60 mL PFA vessels (Savillex, Minneapolis, MN, USA), and the sample droplets remaining on the microwave vessel were rinsed and recovered with 10 mL nanopure H_2O. To evaporate any remaining HF, the 60 mL PFA vessels were transferred to a hot plate at 95ºC under the hood. The liquid evaporated and the sample was brought to a gel in <48 hours. Following evaporation, the samples were diluted in an Optima-grade HNO_3 (3%) solution, passed through 0.45 μm, Titan3 nylon filters (ThermoFisher), 20 mm in diameter, to eliminate any residual particles, and stored in 15 mL polypropylene tubes before analysis by ICP-OES.

2.6 pXRF Analysis

A pXRF spectrometer model TRACER 5i (Bruker Corp, Billerica, Massachusetts, United States) was used to scan the samples in the lab, following methods of Weindorf and Chakraborty (2020) and Silva et al. (2021). The equipment has an Rh X-ray tube operated at adjustable voltage ranging from 6 to 50 kV, and current 4.5-195 μA. The TRACER 5i was operated in lab conditions, under a normal air atmosphere, using the *GeoExploration* factory calibration (Bruker Corp). The method included three beams with three sequential operating conditions, to optimize the quantitative analysis of elements from Mg to U. Beams 1 and 2 were configured to scan the samples for 40 s, and beam 3 scan time was 90 s; the full scan was completed in 170 s.

To guarantee data quality, thirteen certified reference materials (CRMs) were prepared, stored, and analyzed in identical settings as the samples. The CRMs consisted of powdered and homogenized soils (OREAS 25a, NIST 2709, 2710a, 2711); tills (OREAS 47, SRM TILL 2 and

4); and Li or Li-Nb-Sn pegmatites (OREAS 147, 148, 750, 751, 752, 753) (Ore Research and Exploration, Bayswater North, Victoria, Australia; NIST, Gaithersburg, Maryland, USA; CCRMP, Ottawa, Ontario, Canada). The pXRF concentrations were compared against their elemental concentrations. The recovery values for each element defined as (elemental content obtained by pXRF)/(certified elemental content)·100 are reported in Table 2. Correction factors obtained through linear regressions (R^2 ranging between 0.78 to 1.00) using the 13 standards were applied to the raw results (Table 2). Bruker reported the detection limits (in $mg \cdot kg^{-1}$) for pXRF measurements for the *GeoExploration* calibration as follows Al = 360, Si = 921, P = 21, K = 24, Ca = 30, Ti = 40, Fe = 15, and between 1 and 10 ppm for analyzed trace elements Mn, Ni, Cu, Zn, Rb, Ba, and Sr (Silva et al., 2023).

2.7 Lithium Analysis Using ICP-OES

A Thermoscientific ICAP 7400 ICP-OES was used to analyze the Li concentrations of the acid-digested bulk soils. The Li was calibrated and quantified in axial mode, using a 2 $mg \cdot L^{-1}$ Y solution as internal standard. Li contents of all digested samples exceeded the estimated detection limit of Li of 0.01 $mg \cdot L^{-1}$. Digestion-related issues were corrected based on the average recovery of Li concentration of NIST 2710a soil standard (24.57 $mg \cdot kg^{-1}$) that was processed and analyzed together with each of the 11 batches of unknowns.

Table 2. Corrections to the "GeoExploration" factory calibration (Bruker Corp.) that were applied to the raw results. Average recovery values in percentages; linear correction parameters a and m according to equation: $y = a + mx$ in $mg \cdot kg^{-3}$; R^2 – coefficient of determination; and N – number of certified reference materials used for each element.

	Al	Ba	Ca	Cu	Fe	K	Mn	Ni	P	Rb	Si
Avg recovery (%)	87	85	96	122	101	94	111	160	55	96	82
m	1.01	1.43	1.07	0.93	0.88	1.37	0.89	0.88	1.42	1.10	0.80
a	8316	-8	62.68	-2.98	2802.3	1606	7.28	-7.16	211.94	-18.81	111772
R^2	0.87	0.99	0.99	1.00	0.99	0.97	1.00	0.99	0.78	1.00	0.81
N	12	13	13	13	13	11	13	11	13	10	13

2.8 Machine Learning: Lithium content and parent material predictions

One hundred and twelve samples with distinct parent materials were used to build the Li content prediction model with the *Random Forest* algorithm (Fig. 4). The Li concentrations in the unknown samples ranged from 5.7 to 1,354.1 $mg \cdot kg^{-1}$. The independent variables of the prediction model obtained via pXRF consisted of 14 predictors (Al, Ba, Ca, Cu, Fe, K, Mn, Ni, P, Rb, Si, Sr, Ti, Zn) that yielded concentrations significantly higher than the analytical detection limits. The absolute analytical errors reported by the pXRF unit were converted into relative (%) errors for each element. To improve the PCA correlation and the machine learning (*Random Forest*) regressions for the Li prediction, the list of variable predictors was further reduced to 10 elements: Ba, Ca, Cu, Fe, K, Mn, Ni, P, Rb, and Zn. Due to the distinct geochemistry of pegmatites and host rocks (Trueman, 1982; London, 2008; Tang et al., 2018), it was anticipated that several elements in soils formed on LCT-pegmatite bedrock could be used to fingerprint their geochemistry (such as K, Rb, and P) whereas other elements (e.g., Ba, Ca, Cu, Fe, Ni, Sr, and Ti) were expected to be depleted in the highly fractionated pegmatites relative to the barren granite and pegmatites, and enriched in the host metamorphic rocks. The K:Rb ratio, which typically correlates negatively with the degree of fractionation of granitic pegmatites and their Li prospectivity (e.g., Trueman, 1982; Tang et al., 2018) was used as an additional predictor.

The *Random Forest* models built on data obtained from pXRF were trained to predict Li concentration using R software (R Development Core Team, 2018). Training and validation were composed of 75% and 25% of samples, respectively, with random separation (Table 3). Three Li-predictive models were created: Model A, using all 14 pXRF predictors listed above; Model B, created by eliminating Al, Ni, Sr, and Ti, resulting in a list of 10 predictors (Ba, Ca, Cu, Fe, K, Mn, P, Rb, Si, Zn); and Model C, with all 14 predictors plus the K:Rb ratio.

Table 3. Prediction models for Li and parental material (PM) and n – number of samples. See text for model description.

Model	Prediction type	Variables (n)
A	Li	Al, Ba, Ca, Cu, Fe, K, Mn, Ni, P, Rb, Si, Sr, Ti, Zn (14)
B	Li	Ba, Ca, Cu, Fe, K, Mn, P, Rb, Si, Zn (10)
C	Li	Al, Ba, Ca, Cu, Fe, K, Mn, Ni, P, Rb, Si, Sr, Ti, Zn + K/Rb (15)
D	PM type	Al, Ba, Ca, Cu, Fe, K, Mn, Ni, P, Rb, Si, Sr, Ti, Zn + K/Rb (15)

Processing and modeling were performed using R software through the 'caret' package (Kuhn, 2008) with the following tuning parameters: number of trees of the model (*ntrees*) = 1,000, *nodesize* was set to default, and *mtry* = 1/3 of the number of predictor variables as suggested by Liaw and Wiener (2002). The importance of variables is an important output parameter indicating which variables have more impact on the *Random Forest* model (Liaw and Wiener, 2002). Models were assessed through three parameters: the coefficient of determination (R^2), the root mean square error (RMSE) (Eq. 1), and the residual prediction deviation (RPD) (Eq. 2). Models with greater R^2 and smaller RMSE were considered the best for predicting Li concentration in soil. RPD is an indicator divided into three classes: RPD > 2 accurate predictions, 1.4 ≤ RPD ≤ 2 moderately accurate prediction, and RPD < 1.4 no prediction ability (Chang et al., 2001). Eqs 1 and 2 follow as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - x_i)^2}{N}} \text{ (in mg·kg}^{-3}\text{ of Li)} \quad (1)$$

$$\text{RPD} = \text{SD}/\text{RMSE} \quad \text{(unitless)} \quad (2)$$

where N is the number of observations, y_i is the value estimated by the model, x_i is the value measured through independent laboratory measurements, and SD (standard deviation; Eq. 3) is a measure of how dispersed the model data is in relation to the mean of observed data, in mg·kg^{-3} of Li. Eq. 3 follows as:

$$SD = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \overline{x})^2}{N-1}} \quad (3)$$

where x_i is the observed values of a sample item, $\overline{x}$ the mean value of the observations.

For the parental material prediction (Model D), samples were randomly split with 85 samples for modeling and 27 for validation. The samples used to train Model D were classified based on three main parent materials: PM1 included the soils formed on the control granite and other felsic lithologies; PM2 were the soils formed on mafic amphibolites; and PM3 were the soils formed on Li-rich pegmatites. Fifteen predictors were used for training purposes, i.e., the same predictors as for Li concentration Model C prediction (Table 3). Processing and modeling were performed using R software through the 'caret' package with the same tuning parameters used for the Li concentration prediction models. To assess the accuracy of the prediction model, the following parameters were used: Cohen's Kappa coefficient (Eq. 4), overall accuracy (Eq. 5), producer's accuracy (Eq. 6), and user's accuracy (Eq. 7) (Congalton, 1991):

$$\text{Kappa Coefficient} = \frac{\text{Po} - \text{Pe}}{1 - \text{Pe}} \quad (4)$$

$$\text{Overall accuracy} = \frac{\text{number of correct predictions}}{\text{number of validation points}} \quad (5)$$

$$\text{Producer's accuracy} = \frac{Xjj}{\sum_{j=1}^{r} Xij} \quad (6)$$

$$\text{User's accuracy} = \frac{Xii}{\sum_{i=1}^{r} Xij} \quad (7)$$

where *Po* is the proportion of correctly classified points and *Pe* is the probability of random agreement; the Kappa Coefficient ranges from -1 to 1, indicating an increasing accuracy as the values approach 1 (Landis and Koch, 1977); *Xii* and *Xjj* indicate the number of correctly classified samples; and represents the sum of samples of a type of parental material in a row (user's accuracy) or column (producer's accuracy) of a confusion matrix (Ting, 2010).

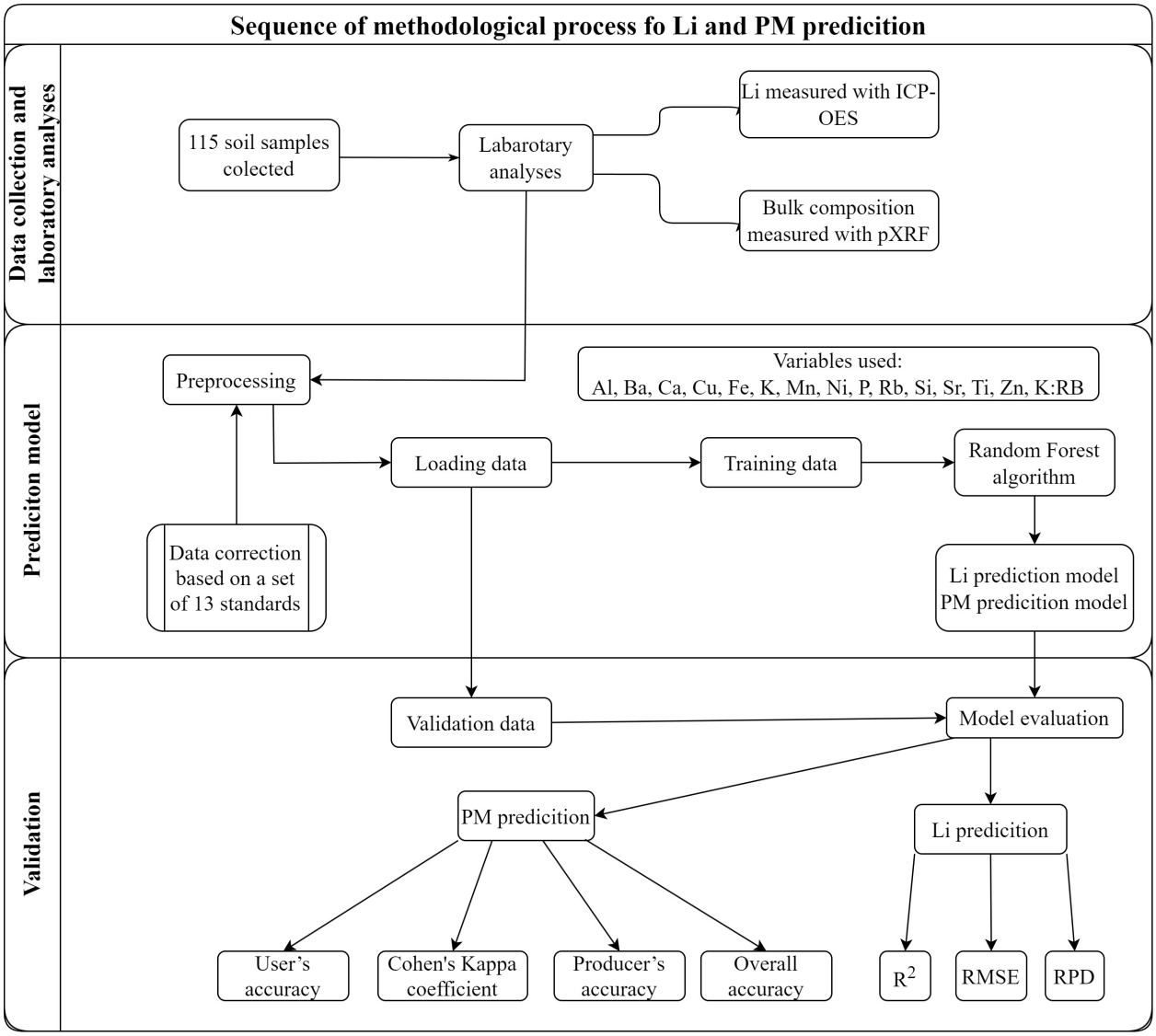

Figure 4. Flowchart illustrating the method sequence for the Li content and parent material predictions *via* portable X-ray fluorescence (pXRF) and *Random Forest* algorithm.

3. RESULTS AND DISCUSSION

3.1 Soil Geochemistry

Compositional boxplots for selected elements show the differences between soils of distinct parent materials (Fig. 5) and suggest the potential for the pXRF analysis to differentiate the Li-rich pegmatites from nonmineralized bedrock. Additional boxplots with the remaining elements are shown in Appendix Figure A1. As expected, the ARA, KX2, and PL soils had the highest Li and Rb contents, averaging 410 $mg \cdot kg^{-1}$ Li and 860 $mg \cdot kg^{-1}$ Li. In comparison, the soils formed on granite (BL), compositionally simple pegmatite (SRQ, 6MI), and their host rocks (DSM), averaged cumulatively only 30 $mg \cdot kg^{-1}$ Li and 80 $mg \cdot kg^{-1}$ Rb (Figs. 5A, B). The dataset included also transitional Li and Rb values for soils formed on the host rock within the metasomatic aureoles of the mineralized pegmatites. The Li contents of these soils were highly variable. In the quartz micaschist (QMS), the host rock of the ARA, Li contents ranged from 11 to 87 $mg \cdot kg^{-1}$, and in the amphibolite (QMA), the host of the KX2 and PL, Li ranged from 17 to 652 $mg \cdot kg^{-1}$ (Table 4, Fig. 5A). This variation ensured the continuity in the dataset, as required for constructing an accurate Li predictive algorithm, from 5 $mg \cdot kg^{-1}$ Li, the minimum in a control DSM amphibolite soil, to 1,582 $mg \cdot kg^{-1}$ Li, the maximum concentration reached in KX2 spodumene-pegmatite soil.

Although the KX is also spodumene bearing, the KX sandy-loam soil samples averaged only 32 $mg \cdot kg^{-1}$ of Li and 50 $mg \cdot kg^{-1}$ of Rb, with the latter falling below the control average. In addition, although the KX pegmatite is surrounded by an extensive Li halo of at least 30 m in its QMA host rock (Liu et al., 2010), the mean Li concentration in the soil traverses across QMA was also low (34.5 ppm). These low concentrations suggest that the glaciofluvial sediment cover at KX, a transported mixture of materials sourced from a broad region, effectively masked the chemical signature of the local bedrock (Table 4, Fig. 5, Fig. A1).

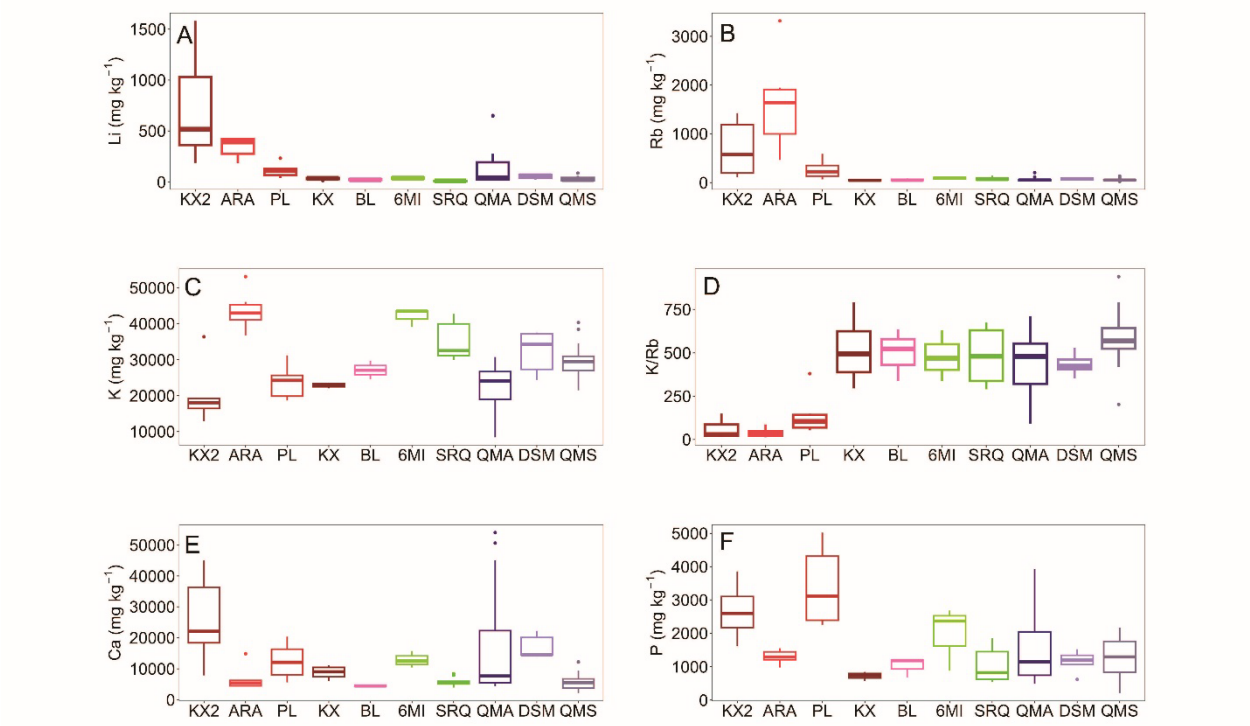

Figure 5. Box plots of selected soil elemental concentrations in mg·kg^{-1} resulted from inductively coupled plasma optical emission spectroscopy (ICP-OES) analysis (Li) and portable X-ray fluorescence (pXRF) analysis (all other elements). Box plots for additional elements in Fig. A1. Parental material abbreviations: KX2: King's-X2; ARA: Animikie Red Ace; KX: King's-X; PL: Price Lake; BL: Bush Lake granite; 6MI: Six-mile Lake; SRQ: Sturgeon River Quarry; QMA: Quinnesec mafic amphibolite; DSM: Dickinson Six-mile mafic amphibolite; QMS: Quinnesec mica schist.

Unlike Rb, the K contents do not correlate with Li (Figs. 5C, 6A,B). The ARA soils are the most enriched in K (mean of 43,730 mg kg^{-1}) but the KX2 soils are the poorest in K (mean of 19,792 mg·kg^{-1}) falling below the background value of barren granite-pegmatite and their hosts rocks (mean of 34,091 mg·kg^{-1}). However, the K:Rb ratios show a strong inverse correlation with the Li contents across the entire dataset (Figs. 5D, 6C) and, with a few exceptions, effectively discriminate between soils formed on mineralized pegmatites and host rocks in the close proximity to mineralized pegmatites (at values of K:Rb < 275), vs. soils formed on barren bedrock and glacial till (K:Rb > 275).

Phosphorous contents are in general higher in the Li-rich soils covering the KX2 and PL pegmatites (mean of 3,000 mg·kg^{-1}) and host rock in their metasomatic aureoles (max of 3,929 mg·kg^{-1}, near the contact with KX2), consistent with the common occurrence of Li

phosphates montebrasite-amblygonite in these pegmatites and abundant metasomatic apatite in the nearby QMA host rock. The P concentrations of the ARA soils are lower (mean of 1,295 $mg \cdot kg^{-1}$) overlapping with those of the surrounding host QMS, and only slightly exceeding the mean P concentrations in the soils formed on the BL granite and nonmineralized SRQ and 6MI pegmatites (Table 4, Fig. 5F). The P contents in KX soils are the lowest (mean of 776 $mg \cdot kg^{-1}$) reflecting their fluvio-glacial parent material, that most likely lost its P via weathering. Overall, there is a poor positive correlation between Li and P (Fig. 6D).

Table 4. The average (avg), minimum (min), maximum (max), and standard deviation (std) of elements used in the model were measured with portable X-ray fluorescence (pXRF) spectrometer ($mg \cdot kg^{-1}$) and Li (ICP-OES) ($mg \cdot kg^{-1}$) data.

		Li	Al	Ba	Ca	Cu	Fe	K	Mn	Ni	P	Rb	Si
ARA	Avg	344	63443	318	6791	16	12136	43730	2006	8	1295	1639	31
	Min	182	58995	145	4326	11	9768	36741	965	5	979	464	30
	Max	430	69043	491	14825	23	15233	53129	4618	11	1548	3318	32
	Std	106	4102	119	4021	5	1775	5556	1334	2	206	999	63
KX2	Avg	724	51857	339	26340	48	43855	19792	4084	22	2658	696	27
	Min	186	33799	149	7889	29	19915	12868	666	11	1611	111	24
	Max	1582	60454	565	45008	102	72260	36413	9320	34	3858	1422	32
	Std	534	8740	133	13183	27	23005	7646	3358	10	788	565	29
KX	Avg	32	50073	320	8832	37	33948	22894	776	31	723	50	34
	Min	13	47667	266	6118	32	28705	22052	452	23	576	29	33
	Max	43	52924	415	11119	47	37821	23629	974	41	847	77	35
	Std	13	2700	67	2279	7	3884	701	233	7	117	20	81
PL	Avg	114	52010	478	12427	33	35535	23817	7179	12	3395	265	30
	Min	39	40294	390	5647	23	17738	18660	693	6	2254	67	29
	Max	233	64075	554	20338	41	65541	31215	13611	18	5023	586	31
	Std	69	8879	62	5713	7	18884	4761	5145	5	1209	192	87
BL	Avg	26	54416	455	4492	26	23568	27133	1014	13	1024	59	33
	Min	19	42416	310	4041	23	18891	24596	285	9	684	43	32
	Max	38	63702	555	4974	29	26985	29716	1704	16	1205	88	33
	Std	10	10899	129	468	3	4192	2560	710	4	295	25	69
6MI	Avg	39	61756	720	12882	21	58877	42085	3627	26	1974	93	27
	Min	26	58608	599	10443	19	56697	39066	1412	23	878	69	25
	Max	52	65162	803	15724	25	63103	43696	5385	29	2680	116	28
	Std	13	3285	107	2663	4	3660	2617	2025	3	962	24	13
SRQ	Avg	12	61589	325	5683	21	21338	34071	790	11	1306	93	33
	Min	8	57482	152	3957	19	12928	29981	220	8	624	56	31
	Max	14	65612	401	7924	23	26778	42730	1410	13	1853	141	34

	Std	3	3464	118	1653	2	6162	5857	647	2	511	41	14
QMA	Avg	112	58374	426	15848	41	44816	22254	2172	27	1453	57	31
	Min	17	38274	220	4429	19	22244	8392	304	9	487	32	24
	Max	652	73632	833	54024	106	84154	30758	16794	56	3929	204	37
	Std	147	6810	104	15008	21	18016	6188	3381	10	940	28	36
DSM	Avg	31	58860	534	11461	19	41748	33788	1329	24	974	74	30
	Min	5	50640	152	4132	13	13241	24378	245	7	539	46	25
	Max	71	68344	812	22211	22	76404	40656	3430	48	1523	106	36
	Std	27	5593	184	6480	3	25247	5290	1190	16	394	24	44
QMS	Avg	31	54362	480	5730	23	22842	29437	1276	13	1259	54	31
	Min	11	37166	0	2184	14	9647	21494	148	4	212	23	23
	Max	87	64699	646	12196	35	29265	40337	9355	26	2170	134	37
	Std	19	8240	138	2278	4	6069	4397	1856	7	536	21	29

Other major elements quantified with the pXRF such as Ca and Fe have a general negative correlation with K and Si (Figs. 6E, F), consistent with their ‘mafic’ affiliation. The high concentrations of majors (Ca and Fe) and minor elements (Ti, Mn, Cu, Ni, and Mn) in soils with a pegmatite parent hosted by a mafic amphibolite (Table 4, Fig. 5, App. Fig. A1), namely KX2 and PL from Florence Co. field and 6MI from Dickinson Co. field hosted by amphibolites QMA and DSM, respectively, appear to be strongly influenced by the pegmatite’s host. The extreme Ca values in QMA soils are likely derived from their abundant calcite veins (Fig. 2C). In contrast, the soils formed on the ARA pegmatite, which is hosted by QMS (felsic metavolcanics and/or metasediments), have relatively low Ca and Fe values, slightly depleted compared to soils with a biotite-bearing granite (BL) or pegmatite (SRQ) parent material (Table 4, Figs. 5F, 6E, F). Additionally, the low Ca values coupled with high Si values of KX soils and their QMA host rock (Fig. 6E, F), despite the mafic composition and numerous calcite veins of the QMA visible in the roadcut, reinforce their glacio-fluvial sediment parent material.

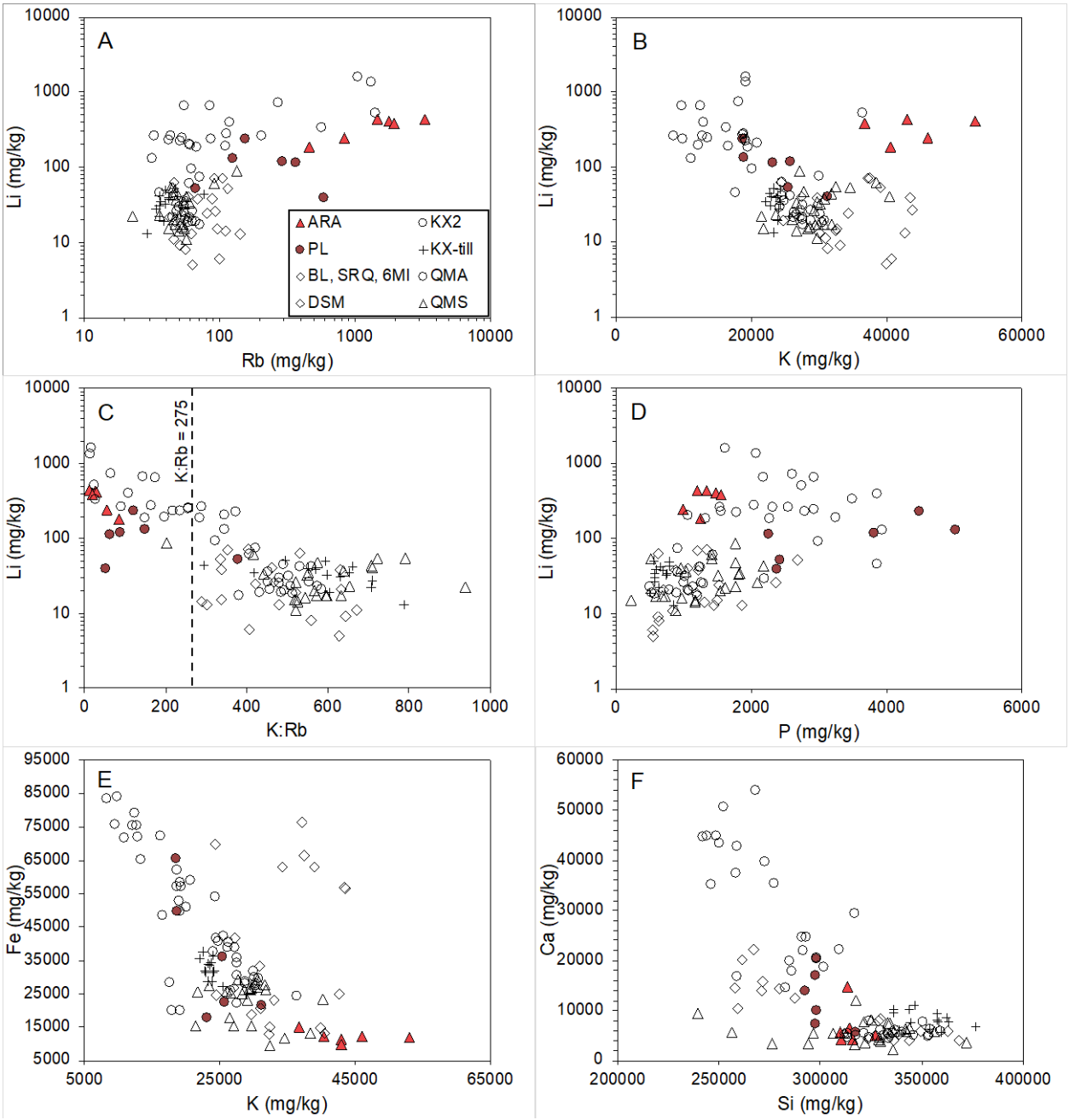

Figure 6. Elemental binary plots, with all soil elemental concentrations in mg·kg^{-1} from inductively coupled plasma optical emission spectroscopy (ICP-OES) analysis (Li) and portable X-ray fluorescence (pXRF) analysis (all other elements). A. Plot of Li vs. Rb; B. Li vs. K; C. Li vs. K:Rb ratio, dash line – line of K:Rb ratio of 275 corresponding to the limit between Li-rich soils and Li-poor soils; D. Li vs. P; E. Fe vs. K; and F. Ca vs. Si. Symbol legend placed in figure A indicates soil parent materials: ARA: Animikie Red Ace pegmatite, KX2: King's-X2 pegmatite, PL: Price Lake pegmatite, KX-till: King's-X pegmatite covered by till, BL: Bush Lake granite, SRQ: Sturgeon River Quarry pegmatite, 6MI: Six-mile Lake pegmatite, QMA: Quinneseq mafic amphibolite, DSM: Dickinson Six-mile mafic amphibolite; QMS : Quinneseq quartz micaschist.

The principal component biplot (Fig. 7) and Pearson correlation (Fig. 8) were constructed to better constrain the best predictors for Li mineralization. Elements that correlate positively with Li at are Rb, P, Zn, and Mn, whereas K:Rb correlates negatively with Li (-0.63) (Figs. 7, 8), consistent with the binary plot (Fig. 6C). The soils formed on Li-pegmatite bedrock separate towards negative values of PC2 consistent with their high Li contents, high Rb, but low K:Rb ratios (Fig. 7) whereas the mean values of soils with nonmineralized parental material separate towards high K:Rb ratios. The soils formed on barren granite and pegmatite parental material's (BL and SRQ), quartz micaschists (QMS), and glacial till/fluvial sediment (KX) cluster away at negative PC1, corresponding to high contents in the 'felsic' components Si and K, whereas the ones formed on barren 6MI pegmatite, its amphibolite host (DSM), and the average of amphibolite QMA cluster at positive PC1 and PC2, corresponding to high contents in the 'mafic' components, Fe, Ni, Ti, Cu, and Ca. Remarkably, several QMA datapoints stretch along the positive PC1 axis, at low PC2 values, at high Ca values, consistent with the abundant calcite lenses and veins in their parental material. Significant differences between the soils formed on distinct Li-pegmatites (ARA vs. and KX2 and PL; Fig. 7) stem from elevated Ca, Fe, Mn, P, and Zn in the KX2 and PL, consistent with their distinct calcite-rich amphibolite host rocks. Elements that discriminate poorly between Li-rich and Li-poor soils (Fig. 8) included Al, Sr, Ni, and Ti. Separate Li predictive algorithms were designed with (model A and C) or without these elements (model B).

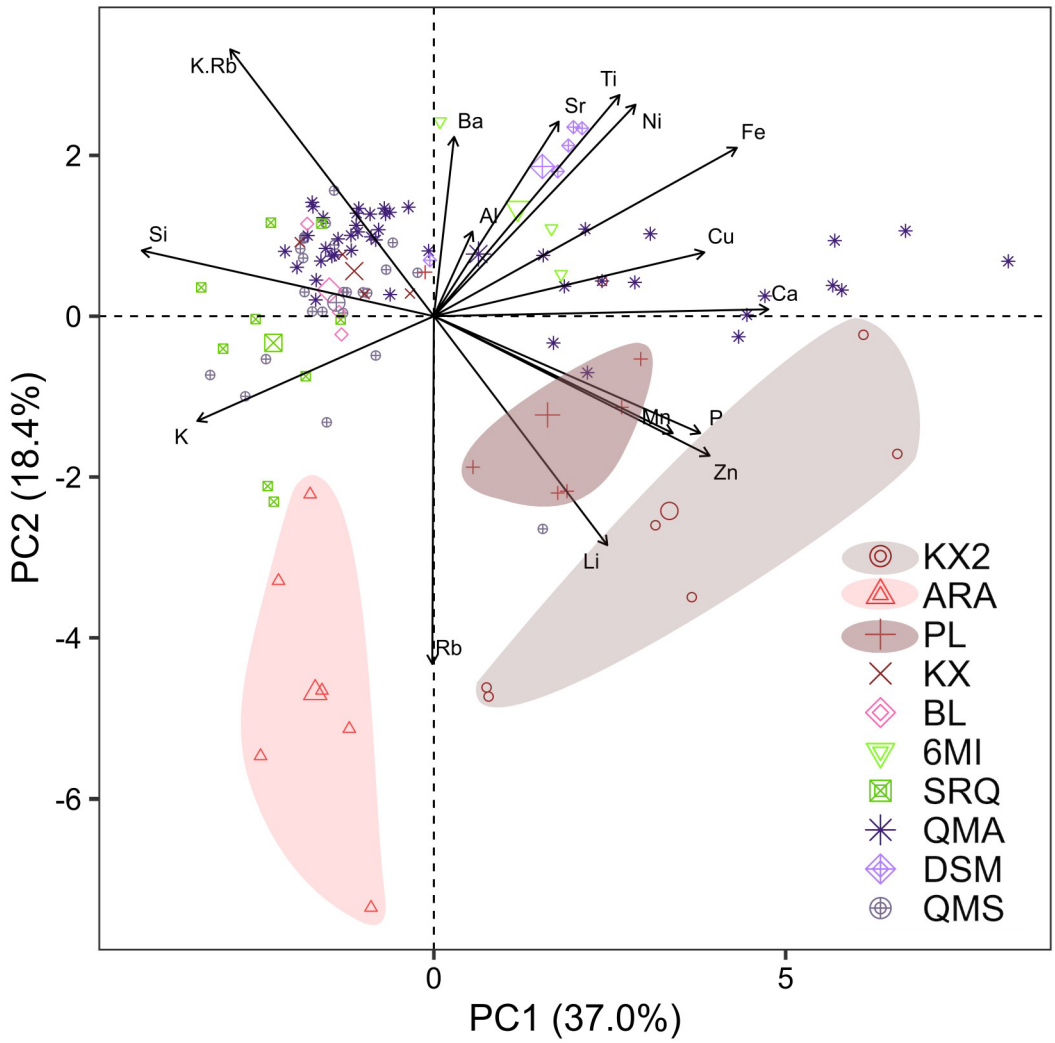

Figure 7. PC1 vs PC2 biplot showing the principal component analysis for Li (via inductively coupled plasma optical emission spectroscopy) and all portable X-ray fluorescence (pXRF) variables ($mg \cdot kg^{-1}$) including the K:Rb ratios for the soil samples from Florence County, WI, and Dickinson County, MI.

3.2 Discussion of geochemical data

Prior to introducing the results of *Random Forest* predictive algorithms for Li and the parental material, the first order interpretations based on soil geochemical data presented above are highlighted. The primary factor controlling the soil geochemistry is parent material. Within the group of magmatic parent materials, the two-mica granite, two-mica pegmatite, beryl-tourmaline bearing pegmatite, and Li-pegmatite produce soils with distinct chemical signatures, mimicking the magmatic fractionation trends of increasing Li with increasing Rb and decreasing K:Rb ratio (Trueman, 1982; Selway et al., 2005; Tang et al., 2018).

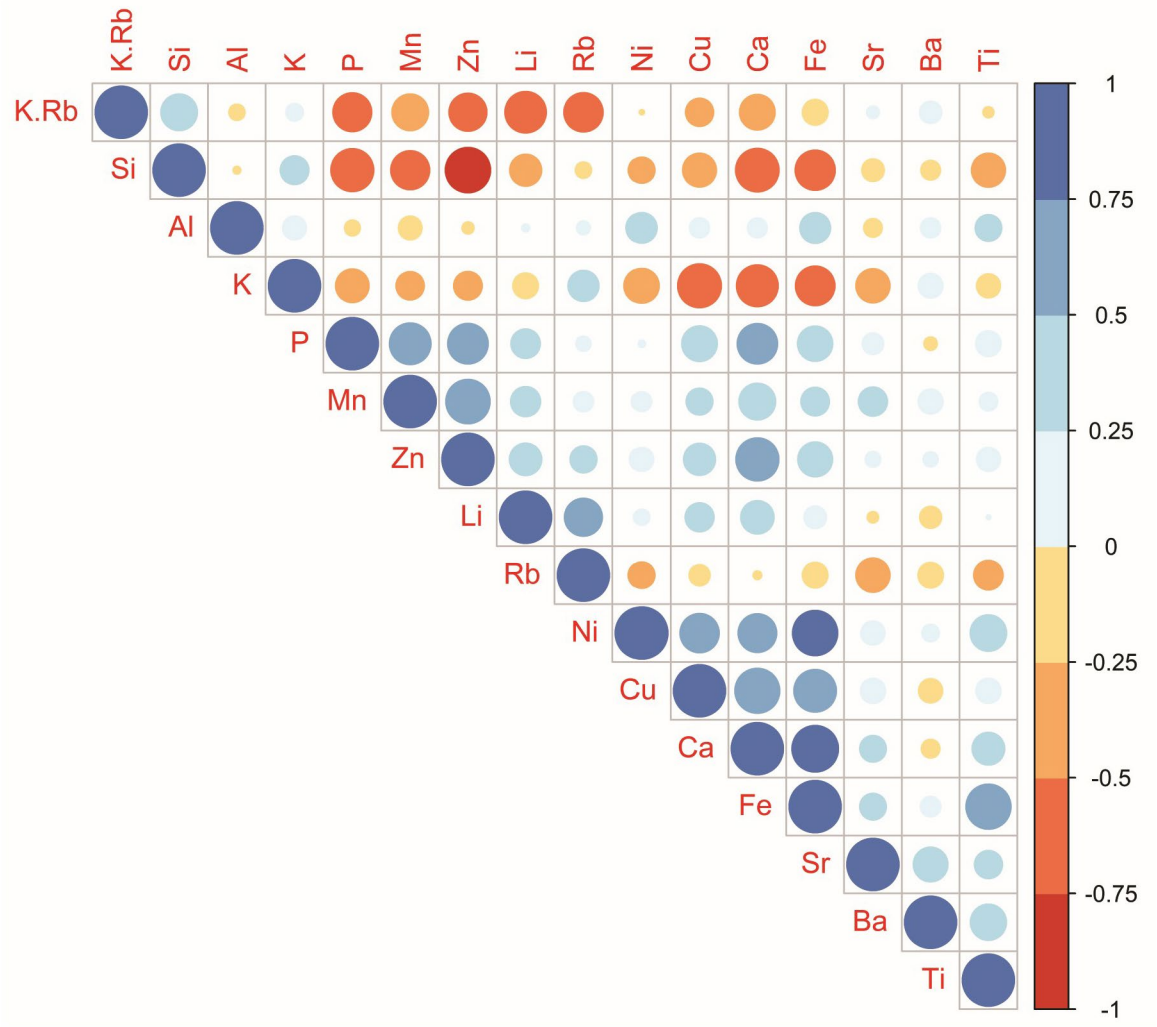

Figure 8. Pearson correlation of Li (inductively coupled plasma optical emission spectroscopy) and the portable X-ray fluorescence (pXRF) variables ($mg \cdot kg^{-1}$) including the K:Rb ratio of the soil samples collected in Florence County, WI and Dickinson County, MI.

Important cross-contamination occurs between soils formed on pegmatites and those on the surrounding host rock, in particular for mafic ± calcite-bearing host rock as noticed in the KX2-QMA soil traverses (e.g., Figs. 6, 7). The distance between the sampling site and the pegmatite (although not the focus of this study) seems to be linked to the soil compositional variations. Metasomatic fluid infiltration during pegmatite emplacement explains the elevated Li in the QMA soils proximal to the KX2 pegmatite (within ~< 20 m from the contact with pegmatite or its concealed extensions). This original distribution of Li and Li-pathfinder was likely kept intact since the emplacement of the pegmatites. Despite the proximity to the Niagara fault system, the pegmatites and the host rocks are not pervasively affected by shearing or post-

emplacement metamorphism. The latest glaciation eroded the area down to fresh bedrock, as demonstrated by numerous “roches moutonnées”, polished outcrops with glacial striations and grooves. Therefore, any effects of surficial processes older than Pleistocene were effectively removed. However, postglaciation surficial processes may have resulted in the contamination of the pegmatite soil, when surrounded by soil formed on host-rock with contrasting composition. For example, because Ca released by chemical weathering from calcite veins and Fe from mafic minerals are soluble in the presence of low pH meteoric water (due to humic acid), it is likely that the soil formed on relatively narrow pegmatites is enriched in Ca and Fe (e.g. Figs. 6E, F) due to supergene influx from surrounding soils with QMA parent, not because of primary processes.

Other surficial factors leading to variability in the data set may include partitioning of different elements in detrital minerals with various resistance to chemical and mechanical weathering, differential solubility/mobility of various elements (such as Li vs. Rb), and the topographic slope in relation to the contact with the pegmatite. It is likely that the latter would influence the soil contamination across lithologic boundaries via both the dissolved load (through meteoric water runoff and infiltration) and detrital load (i.e., particles carried by runoff).

Internally-zoned LCT pegmatites such as the ARA and KX2 are very heterogeneous, with discontinuous mineral zones and a broad variation of accessory minerals, giving them unique compositional signatures (London, 2008; London and Morgan, 2012). These variations, together with the heterogeneous distribution of Li-mineral rich clusters along the pegmatite strike can explain the large spread in data points (Figs. 6, 7). In addition, the remarkable clustering of Li-rich datapoints illustrated by the PCA results (Fig. 7) may suggest that soil geochemistry can discriminate between pegmatites with fresh spodumene (i.e., the KX2 and PL pegmatites) and lepidolite-rich pegmatites with hydrothermally altered spodumene (i.e., the ARA pegmatite). Although soils formed on the sedimentary cover were avoided, it is possible that localized, undetected small masses of loess may have been present, blocking the bedrock signature.

3.3 Validation of Li content prediction models

The accuracy for Li content prediction is shown in Fig. 9 and summarized in Table 4. All three models can be used for Li content prediction with good predictive power. Models A, B, and C (defined in Table 2) achieved R^2 values of 0.81, 0.86, and 0.82, respectively. Among the three models tested, model B, constructed by removing four elements with the lowest

correlation with Li (Al, Ni, Sr, and Ti; Fig. 8), achieved the best results, with the smallest RMSE (68.5 mg kg^{-1}). For model C, the addition of K/Rb to all 14 variables used in model A did not improve R^2 and RMSE significantly. When comparing the prediction power, the models A, B, and C achieved moderate prediction accuracy with RPDs 1.72, 1.78, and 1.69, respectively (Table 4).

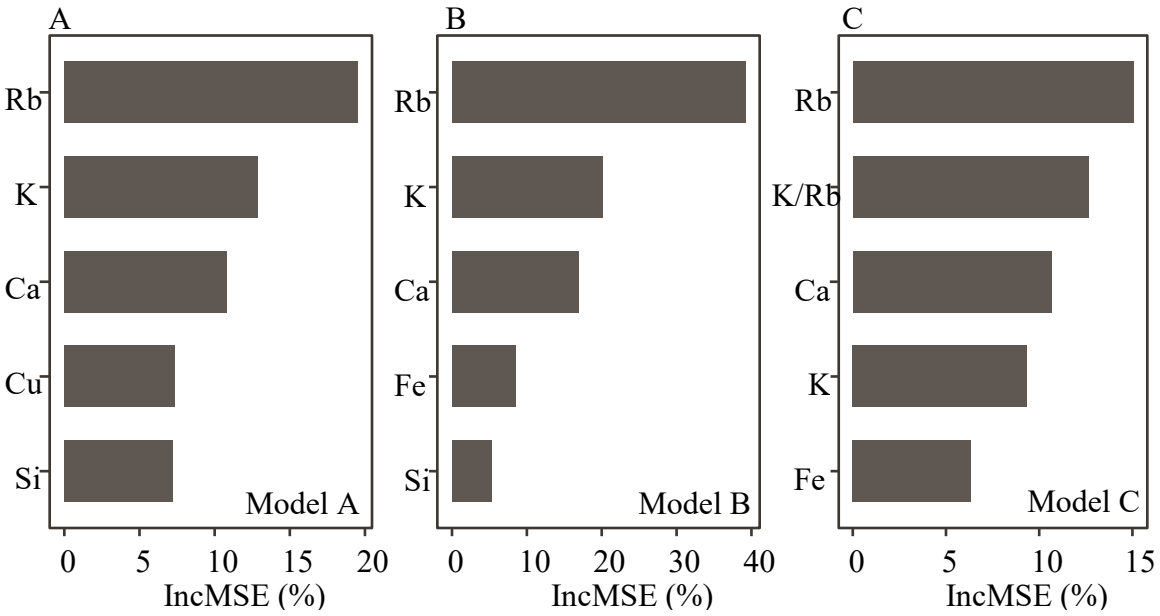

Figure 9. The importance of portable X-ray fluorescence (pXRF) variables according to the *Random Forest* algorithms for the three best Li prediction content for soils from Florence County, WI and Dickinson County, MI.

The most important *Random Forest* variables were determined using the increase in mean squared error (%IncMSE), a parameter that calculates the increase in MSE of predictions when a variable is randomly permutated (Fig. 9) (Liaw and Wiener, 2002). With the exception of Rb, a trace element characteristic of LCT pegmatites, and Cu, a trace element of mafic affiliation, all the other elements that were ranked top five in importance were major constituents in the Earth's crust: Si, K, Ca, and Fe. Notably, Al, another major element included in models A and C was not selected as a top 5 predictor, consistent with its relatively poor discrimination power (Fig. 7) and lack of correlation to Li (Fig. 8). All models agreed that Rb, K, and Ca are relevant, and ranked Rb as the most important predictor of Li content. Model C, differentiated from model A by the addition of K/Rb as a predictor, performed slightly better than model A (Table 4), and K/Rb ratio became the second most important variable. Although soils have a diluted elemental content compared to rock due to weathering (Schaetzl and

Anderson, 2005), most of the elements used in the study had contents well above the pXRF limits of detection and outstanding recovery values when compared to 11 standard materials. The robustness of the pXRF analysis is consistent with the *Random Forest* model performance, indicating that pXRF can be an alternative to indirectly quantify Li in soils formed on pegmatites rich in Li.

Table 5. Coefficient of determination (R^2), root mean square error (RMSE) in $mg \cdot kg^{-1}$, and residual prediction deviation (RPD) for the Li-prediction validation for Random Forest models A, B, and C (defined in Table 3).

Model	R^2	RMSE	RPD
A	0.81	71.80	1.72
B	0.86	68.48	1.78
C	0.82	75.37	1.69

3.4 Validation of the Parental Material Prediction Model

The performance of parental material prediction model D (Table 3) is shown in Figure 10 along with the importance of the top 5 variables. The parental material classification categories were PM1 representing soil formed on felsic lithology and glacio-fluvial sediment; PM2 representing soil formed on mafic amphibolites; and PM3 representing soil formed on Li-rich pegmatites and their host rock within <20 m from contact with pegmatite, with a K:Rb < 275 (Fig. 10A). The model delivered an overall accuracy of 0.85 by correctly classifying 23 of the 27 validation data points, and a Kappa coefficient of 0.71. The producer's and user's accuracy achieved reasonable results (0.88 and 0.82, respectively; Fig. 10B). Per the confusion matrix (Fig. 10A), the model misclassified a total of four samples from all three parental material categories. However, when observing Table 4 and Figures 5 and 6, it is noticeable that those misclassified outliers have intermediate elemental contents, in part caused by 'cross-contamination' factors discussed above or Li-poor segments of otherwise mineralized pegmatites. A previous version of this model (D0; Fig. A2) was constructed by allocating only Li-rich bedrock to the PM3 training dataset. This model achieved an overall accuracy of only 0.81. By reclassifying the parental material datapoints with K:Rb ratio maximum <275 (corresponding to a distance to the pegmatite < 20 m) not only did model D achieve better statistics (Fig. 10), but also the user could gain further useful information for Li exploration,

i.e. that the unknown data point has a high chance to be located either directly on a mineralized pegmatite or in close proximity to one, if the data point was classified in the PM3 category. In general, Rb and Zn were the most important variables for model D (Fig. 10B). Throughout this study, Rb emerged as a significant fingerprint for soils originating from LCT pegmatites, consistent with findings from numerous other studies (e.g., Nieder et al., 2014; Cardoso-Fernandes et al., 2021; Müller et al., 2022). Calcium and Fe exhibited distinct patterns among parental material materials, displaying a wide range in both pegmatites and their host rocks.

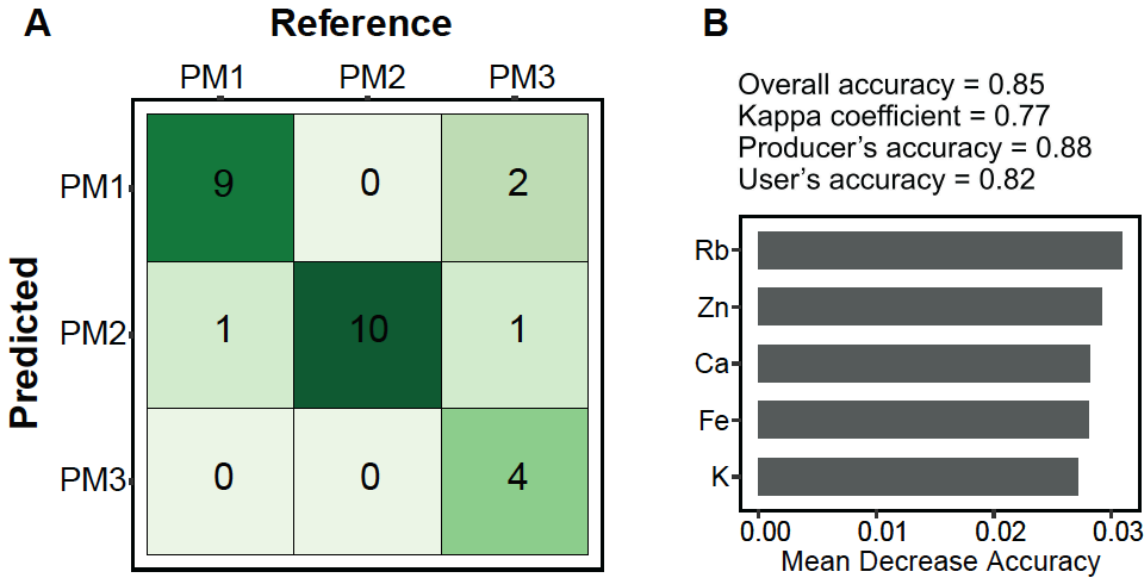

Figure 10. Model D, the parental material prediction model *via* portable X-ray fluorescence (pXRF) and *Random Forest* algorithm at Florence County, WI and Dickinson County, MI. A. Confusion matrix, where PM1: number of validation soils formed on Li-poor felsic lithologies: granite (BL), Sturgeon River Pegmatite (SRQ), Sixmile pegmatite (6MI), Quinnesec felsic metavolcanics and metasediments (QMS), and soils formed on glacio-fluvial sediment (KX); PM2: validation soils formed on Li-poor mafic lithologies: mafic amphibolites (QMA and DSM); and PM3: validation soils formed on Li-rich pegmatites: Animikie Red Ace (ARA), King's X2 (KX2), and Price Lake (PL) and any QMA and QMS soil samples situated within <20 m visible contact with Li-rich pegmatite with a K:Rb ratio < 275. B (Fig. 6C). Classification accuracy parameters, and top-5 variable importance for parent material prediction.

3.5 Discussion and recommendations for Li exploration studies

The reasonable parental material assignment to Li-mineralized pegmatites or to host-rock within proximal distance from Li-mineralized pegmatites, and their reasonably-successful discrimination from non-mineralized bedrock suggest that this machine learning approach is a promising tool for preliminary Li exploration studies, at least for relatively young (Holocene) and shallow Spodosols and Alfisols in glaciated terrains of the Northern hemisphere. However, certain issues may hinder the success of the Li and parental material prediction based on pXRF analysis and recommendations for future application of machine learning approaches to Li exploration.

The samples included in this pilot study encompassed superficial soil horizons, selected in areas with exposed bedrock, so that they can be properly classified for training of the parental material model without the use of deep-sampling tools. The O and A horizons tend to be more weathered, leached, and more prone to anthropogenic influence (Mancini et al., 2020). Contrariwise, a factor that may affect elemental contents in soils is topographic relief. As most locations in this study are in relatively steep terrain, contamination of downslope soils with soluble elements and/or detrital particles may occur. Future studies should investigate compositional variations with depth within thicker soil profiles, to improve the accuracy of the parental material prediction by avoiding surficial cross-contamination and anthropogenic effects, and to extend the applicability of the method to broader areas, including those of uncertain mineralization potential with thicker soils and poor outcrops.

Studies using soil elemental contents measured via pXRF to distinguish or predict parental materials have succeeded in temperate and tropical regions (Mancini et al., 2020; Gozukara et al., 2021; Yuan et al., 2021; Pierangeli et al., 2023). When models created based on superficial and deeper soil horizons were compared, the parental material models that used deeper horizons achieved better results due to the proximity to bedrock and less influence of anthropogenic processes. However, the soils of this study are shallow (< 3 m), because the sampling strategically targeted the residual soil units formed on scoured outcrops. Because they post-date the end of the Wisconsin glacial episode at ~10,000 BP (Larson and Kincare, 2009), they experienced only minimum weathering, thus their geochemistry is largely influenced by mineral and lithic fragments derived from the bedrock (Schaetzl, 2009). These features ensure that the concentrations of Li and the predictor elements exceed the detection limits of ICP-OES and pXRF analyses, respectively. In essence, this relatively rapid method, based on bulk soil geochemistry, is robust enough for machine learning models to perform adequately, but only when dealing with thin soil veneer, directly formed on bedrock. When soils formed on a glacial/fluvial sediment, such as those collected on top of the KX spodumene pegmatite

location, the chemical signature of the underlying bedrock was effectively obscured, at least within the analytical resolution of this study.

For deeper bedrock covered by overburden ranging from tens to hundreds of meters thick, exploration for Li-pegmatite via the B-horizon geochemistry and machine learning remains feasible in areas unaffected by mining and other anthropogenic influences, utilizing a selective leaching methodology (e.g., proprietary *Enzyme Leach*^{SM} methods; Galeschuck and Vanstone, 2005). However, unlike the rapid pXRF method proposed in this pilot study, the analysis of the selective leachates would require ICP-MS to detect low $\mu g \cdot kg^{-1}$ (ppb level) Li and pathfinder anomalies generated by bedrock deeply concealed pegmatites (Xu et al., 2023), which is costly and time prohibitive. To implement this approach to Li-exploration projects from other climates, the machine learning algorithm would need to be recalibrated, especially in places with thicker, much older, soil cover (Curi and Franzmeier, 1987; Oliveira et al., 2017; Mancini et al., 2021).

Despite its inability to explore for deeper pegmatites, the method has the potential to rapidly direct field prospecting in areas with thick vegetation and highly weathered outcrops near the surface that would hinder rapid mineral identification, upon visual inspection and/or using proximal sensors. For example, as a continuation of this pilot study, although it took several months to complete the analytical work in the lab, soils with high Rb and low K:Rb ratio collected along a blind traverse cutting the concealed extension of KX2 pointed successfully to shallow pegmatite (< 0.5 m) that otherwise would have been missed. The Li-rich pegmatite body and its holmquistite (Li-amphibole) bearing metasomatic aureole were uncovered during the following field season (Cox et al., 2023). To enable *in-situ*, rapid screening of potential targets for further trenching or coring, custom calibration algorithms for the Li prediction and Li-rich parental material recognition would need to be developed and embedded within the pXRF unit. Achieving this would necessitate partnerships with analytical instrumentation companies and blind tests to validate/optimize the algorithm for specific geological and pedological conditions. Nonetheless, this pilot study shows promise constraining the relationships between bulk soil geochemistry and their Li-rich or Li-poor parental materials and suggests that, with further improvements, pXRF can support rapid, cost-effective exploration for Li pegmatites in the field.

4. CONCLUSIONS

This pilot exploration study suggests that predictive machine learning algorithms can be used successfully to assess the Li contents and the parental material of soils based on

concentrations of elements heavier than Al measured with a Bruker TRACER 5i pXRF with the *Geo-exploration* factory calibration, after a correction based on a set of 11 soils and Li pegmatite-ore standards. The only Li-pathfinders of true LCT signature were Rb, the K:Rb ratio, and, for Li-phosphate-bearing pegmatites, P, whereas major constituents of both felsic (e.g., Si, K) and mafic affiliation (e.g., Ca, Fe) consistently ranked within the top 5 in terms of importance in the predictive models.

Despite some overlaps, significant compositional differences in superficial soils derived from different parental materials were detected by pXRF, enabling separation between spodumene-bearing LCT pegmatites and their Li-rich contact aureoles (K:Rb < 275 and Li > 100 $mg \cdot Kg^{-1}$) vs. simple granites and pegmatites. Li concentration prediction using selected major, minor, and trace element concentrations acquired rapidly and accurately with a pXRF.

Due to pegmatites' complexity and diversity of pedologic processes, further studies are required to improve this new exploration approach and better understand the relationship between soil and parental material. If the machine learning calibration of Li concentrations in soil could be integrated with pXRF equipment, this study could become a prototype method for reducing costs and accelerating the prospecting for Li via soils. The approach devised in this pilot study, targeting glaciated terrains preselected from available pedologic databases to ensure shallow outcrops and avoid sedimentary cover, holds potential for application in various climates and soil types. This approach could contribute to improved management and sustainable decision-making for exploration and mining companies worldwide.

5. ACKNOWLEDGEMENTS

The authors would like to thank the National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES), and Foundation for Research of the State of Minas Gerais (FAPEMIG) for the financial support to develop this research. We are thankful for generous support from the Department of Earth and Atmospheric Science and the Office of Research and Graduate Studies at Central Michigan University (CMU) for hosting Luiza and supporting their 'sandwich' PhD research between Federal University of Lavras (UFLA), Brazil and CMU, Michigan. Thanks to Lithium Americas Corporation, Geological Society of America (GSA), and Mineralogy Geochemistry Petrology and Volcanology (MGPV) division of GSA for sponsoring this research, which allowed the use of soil geochemistry as an exploration tool for elements critical to green technologies.

6. REFERENCES

Ayuso, R.A., Schulz, K.J., Cannon, W.F., Woodruff, L.G., Vazquez, J.A., Foley, N.K., Jackson, J., 2018. New U-Pb Zircon Ages for Rocks from the Granite-Gneiss Terrane in Northern Michigan: Evidence for Events at ~3750, 2750, and 1850 Ma. Presented at the 64th Institute on Lake Superior Geology Proceedings, pp. 7–8.

Barros, R., Kaeter, D., Menuge, J.F., Fegan, T., Harrop, J., 2022. Rare element enrichment in lithium pegmatite exomorphic halos and implications for exploration: evidence from the Leinster albite-spodumene pegmatite belt, Southeast Ireland. Minerals. 12: 981. https://doi.org/10.3390/min12080981

Batt, G.E., Crook, D., Brand, N., Kerr, S., 2021. Agile management and long-term strategy in exploration: the 'lucky' discovery of the Sinclair Pollucite Deposit, Eastern Goldfields, Western Australia. Australian Journal of Earth Sciences 68, 684–696. https://doi.org/10.1080/08120099.2021.1848922

Boelter, J.M., Elg, A.M., 2004. Soil Survey of Florence County, Wisconsin.

Bradley, D.C., Stillings, L.L., Jaskula, B.W., Munk, L., McCauley, A.D., 2017. Lithium, U.S. Geological Survey Professional Paper 1802K. Reston, VA. https://doi.org/10.3133/pp1802K

Brand, N.W., Brand, C.J., 2017. Detecting the Undetectable: Lithium by Portable XRF.

Burley, L.L., Barnes, S.J., Laukamp, C., Mole, D.R., Le Vaillant, M., Fiorentini, M.L., 2017. Rapid mineralogical and geochemical characterisation of the Fisher East nickel sulphide prospects, Western Australia, using hyperspectral and pXRF data. Ore Geology Reviews 90, 371–387. https://doi.org/10.1016/j.oregeorev.2017.04.032

Cannon, W.F., 2018. Part 2: Field Trip Guidebooks. Presented at the 64th Institute on Lake Superior Geology Proceedings.

Cardoso-Fernandes, J., Silva, J., Dias, F., Lima, A., Teodoro, A.C., Barrès, O., Cauzid, J., Perrotta, M., Roda-robles, E., Ribeiro, M.A., 2021. Tools for remote exploration: A

lithium (Li) dedicated spectral library of the Fregeneda–Almendra aplite–pegmatite field. Data 6, 1–10. https://doi.org/10.3390/data6030033

CEC, 2021, Guide to Drought Indices and Indicators Used in North America: Commission for Environmental Cooperation, 62 p., accessed at http://www.cec.org/files/documents/publications/11872-guide-drought-indices-and-indicators-used-in-north-america-en.pdf.

Černý, P., Ercit, T.S., 2005. The classification of granitic pegmatites revisited. The Canadian Mineralogist 43, 2005–2026. https://doi.org/10.2113/gscanmin.43.6.2005

Chang, C.-W., Laird, D.A., Mausbach, M.J., Hurburgh, C.R., 2001. Near-Infrared Reflectance Spectroscopy-Principal Components Regression Analyses of Soil Properties. Soil Science Society of America Journal 65, 480–490. https://doi.org/10.2136/sssaj2001.652480x

Chao, T.T., Sanzolone, R.F., 1992. Decomposition techniques. Journal of Geochemical Exploration 44, 65–106. https://doi.org/10.1016/0375-6742(92)90048-D

Congalton, R.G., 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37, 35–46. https://doi.org/10.1016/0034-4257(91)90048-B

Cox, T., Pierangeli, L.M.P.*, Brennan, C.*, Sirbescu, M.-L.C. (2023) Lithium-cesium-tantalum (LCT) pegmatite exploration in Florence County, Wisconsin, USA: Following the lithium from bedrock to soil. Geological Society of America Abstracts with Programs. 55(6) https://doi.org/10.1130/abs/2023AM-393437

Curi, N., Franzmeier, D.P., 1987. Effect of Parent Rocks on Chemical and Mineralogical Properties of Some Oxisols in Brazil. Soil Science Society of America Journal 51, 153–158. https://doi.org/10.2136/sssaj1987.03615995005100010033x

Da Silva, A.C., Triantafyllou, A. and Delmelle, N., 2023. Portable x-ray fluorescence calibrations: Workflow and guidelines for optimizing the analysis of geological

samples. Chemical Geology, 623, 121395. https://doi.org/10.1016/j.chemgeo.2023.121395

Duda, B.M., Weindorf, D.C., Chakraborty, S., Li, B., Man, T., Paulette, L., Deb, S., 2017. Soil characterization across catenas via advanced proximal sensors. Geoderma 298, 78–91. https://doi.org/10.1016/j.geoderma.2017.03.017

Dutton, C.E., Bradley, R.E., 1970. Lithologic, geophysical, and mineral commodity maps of precambrian rocks in Wisconsin. https://doi.org/10.3133/i631

Falster, A.U., Simmons, W.B., Webber, K.L., 1996. The mineralogy and geochemistry of the Animikie Red Ace pegmatite, Florence county, Wisconsin. Recent research developments in mineralogy, vol 1.

Falster, A.U., Simmons, W.B., Webber, K.L., 2005. Origin of the pegmatites in the Hoskin Lake pegmatite field, Florence Co., Wisconsin. In: Pezzotta, F. (Editor), Crystallization Processes in Granitic Pegmatites, International Meeting. Mineralogical Society of America, Cavoli, Elba Island, Italy.

Falster, A., Simmons, W., Webber, K., 2018, Anatectic Origin of the post-Penokean Li-Cs-Ta-enriched Pegmatites in Florence County, Wisconsin, USA, in XXII International Mineralogical Association meeting – Melbourne, Australia.

Finnigan, L.D., Golitko, M., 2023. A pXRF analysis of historic brick chemical data throughout the lake Michigan Morainic region, United States. Journal of Archaeological Science: Reports 47, 103817. https://doi.org/10.1016/j.jasrep.2022.103817

Frost, B.R., and Frost, C.D., 2008, A Geochemical Classification for Feldspathic Igneous Rocks: Journal of Petrology, 49(11), 1955–1969.

Galeschuk, C., and Vanstone, P., 2005. Exploration for buried rare element pegmatites in the Bernic Lake region of southeastern Manitoba, in Linnen, R.L., and Samson, I.M.

(Editors), Rare-element geochemistry and mineral deposits: Geological Association of Canada Short Course Notes 17, 159–173.

Gloaguen E., Melleton, J., Bourcerol, B., Millot, R., 2023. Lithium Mineralization, Contributions of Paleoclimates and Orogens. In Decree, S. (Ed), Metallic Resources 2. Geodynamic Framework and Remarkable Examples in the World. ISTE Ltd and John Wiley and Sons, Inc. 1–51. www.iste.co.uk/decree/metallic2.zip

Gozukara, G., Zhang, Y., Hartemink, A.E., 2021. Using vis-NIR and pXRF data to distinguish soil parent materials – An example using 136 pedons from Wisconsin, USA. Geoderma 396, 115091. https://doi.org/10.1016/j.geoderma.2021.115091

Han, Z., Wang, R., Tong, X., Sun, F., Li, Y., Liu, S., Xue, Q., 2021. Multi-scale exploration of giant Qulong porphyry deposit in a collisional setting. Ore Geology Reviews 139, 104455. https://doi.org/10.1016/j.oregeorev.2021.104455

Harmon, R.S. and Senesi, G.S., 2021. Laser-induced breakdown spectroscopy–a geochemical tool for the 21st century. Applied Geochemistry, 128, p.104929. https://doi.org/10.1016/j.apgeochem.2021.104929

Hockemeyer, J.A., Paliewicz, C.C., Sirbescu, M.-L.C., 2010. Textural and mineralogical changes during crystallization of a lithium pegmatite, Florence County, Wisconsin, USA. Geological Society of America Abstracts with Programs 42(5), 286

James, H.L., Clark, L.D., Lamey, C.A., Pettijohn, F.J., 1961. Geology of central Dickinson County, Michigan, U.S. Geological Survey Professional Paper 310, 1–171. https://doi.org/10.3133/pp310

Kuhn, M., 2008. Building predictive models in R using the caret package. Journal of statistical software 28, 1–26.

Landis, J.R., Koch, G.G., 1977. The measurement of observer agreement for categorical data. Biometrics 33, 159. https://doi.org/10.2307/2529310

Larson, G.J., Kincare, K., 2009. Late Quaternary history of the Eastern mid-continent region, USA. In Schaetzl, R., Darden, J., and Brandt, D. (Editors), Michigan Geography and Geology, Pearson Custom Publishing. 69–90

Liaw, A., Wiener, M., 2002. Classification and regression by random forest. R News 2, 18–22.

Lima, W. de, Mancini, M., Avanzi, J.C., Silva, S.H.G., Acuña-Guzman, S.F., Demattê, J.A.M., Curi, N., 2023. Tracing the origin of deposited sediments: A study applying proximal sensing in a drainage subbasin. Journal of South American Earth Sciences 123, 104241. https://doi.org/10.1016/j.jsames.2023.104241

Linnen, R., Van Lichtervelde, M., Černý, P., 2012. Granitic Pegmatites as Sources of Strategic Metals, 8, 275-280 pp. https://doi.org/10.2113/gselements.8.4.275

Liu, X.M., Rudnick, R.L., Hier-Majumder, S., Sirbescu, M.L.C., 2010. Processes controlling lithium isotopic distribution in contact aureoles: A case study of the Florence County pegmatites, Wisconsin. Geochemistry, Geophysics, Geosystems 11, 1–21. https://doi.org/10.1029/2010GC003063

London, D., 2008. Pegmatites. Canadian Mineralogist Special Publication, Mineralogical Association of Canada, 368 pp.

London, D., Morgan, G.B., 2012. The pegmatite puzzle. Elements 8, 263–268. https://doi.org/10.2113/gselements.8.4.263

Luecke, W., 1984, Soil Geochemistry above a Lithium Pegmatite Dyke at Aclare, Southeast Ireland: Irish Journal of Earth Sciences, v. 6, no. 2, p. 205–211.

Mancini, M., Silva, S.H.G., Teixeira, A.F. dos S., Guilherme, L.R.G., Curi, N., 2020. Soil parent material prediction for Brazil via proximal soil sensing. Geoderma Regional 22. https://doi.org/10.1016/j.geodrs.2020.e00310

Mancini, M., Silva, S.H.G., Hartemink, A.E., Zhang, Y., de Faria, Á.J.G., Silva, F.M., Inda, A.V., Demattê, J.A.M., Curi, N., 2021. Formation and variation of a 4.5 m deep Oxisol in southeastern Brazil. Catena 206. https://doi.org/10.1016/j.catena.2021.105492

Meldrum, J., Brennan, C., Cox, T., Sirbescu, M.-L.C. (2023) Exploring textural variations of spodumene in lithium-cesium-tantalum (LCT) pegmatites from Florence County, Wisconsin, USA. Geological Society of America Abstracts with Programs. 55(6) https://doi.org/10.1130/abs/2023AM-394073

Müller, A., Reimer, W., Wall, F., Williamson, B., Menuge, J., Brönner, M., Haase, C., Brauch, K., Pohl, C., Lima, A., 2022. GREENPEG–exploration for pegmatite minerals to feed the energy transition: first steps towards the Green Stone Age. Geological Society, London, Special Publications 526, SP526-2021.

Naimi, S., Ayoubi, S., Di Raimo, L.A.D.L., Dematte, J.A.M., 2022. Quantification of some intrinsic soil properties using proximal sensing in arid lands: Application of Vis-NIR, MIR, and pXRF spectroscopy. Geoderma Regional 28, e00484. https://doi.org/10.1016/j.geodrs.2022.e00484

Nieder, R., Weber, T.K.D., Paulmann, I., Muwanga, A., Owor, M., Naramabuye, F.X., Gakwerere, F., Biryabarema, M., Biester, H., Pohl, W., 2014. The geochemical signature of rare-metal pegmatites in the Central Africa Region: Soils, plants, water and stream sediments in the Gatumba tin-tantalum mining district, Rwanda. Journal of Geochemical Exploration 144, 539–551. https://doi.org/10.1016/j.gexplo.2014.01.025

Oliveira, V.A., Jacomine, P.T.K., Couto, E.G., 2017. Solos do bioma Cerrado, in: Pedologia - Solos Dos Biomas Brasileiros. SBCS, Viçosa, MG, pp. 177–226.

Phelps-Barber, Z., Trench, A., Groves, D.I., 2022. Recent pegmatite-hosted spodumene discoveries in Western Australia: insights for lithium exploration in Australia and globally. Applied Earth Science 131, 100–113. https://doi.org/10.1080/25726838.2022.2065450

Pierangeli, L.M., Cox, T., Silva, S.H.G., Sirbescu, M.-L.C., 2022. Prediction of Li via geochemical data and machine learning algorithms in soils formed on pegmatite bedrock: a preliminary study. Geological Society of America Abstracts with Programs, 54(4) https://doi.org/10.1130/abs/2022AM-382238

Pierangeli, L.M.P., Silva, S.H.G., Teixeira, A.F. dos S., Mancini, M., Andrade, R., Menezes, M.D., Sirbescu, M.-L.C., Marques, J.J., Weindorf, D.C., Curi, N., 2023. Soil parent material spatial modeling at high resolution from proximal sensing and machine learning: A pilot study. Journal of South American Earth Sciences 129, 104498. https://doi.org/10.1016/j.jsames.2023.104498

R Development Core Team, 2018. R: A language and environmental for statistical computing. R Foundation for Statistical Computing.

Robinson, G.W., Carlson, S.M., 2013. Mineralogy of Michigan Update. A.E. Seaman Mineral Museum, Michigan Technological University, Houghton, Michigan, 47 pp.

Schaetzl, R., Anderson, S., 2006. Soil-Genesis and Geomorphology. Cambridge University Press, 817 pp. https://doi.org/10.1017/CBO9780511815560

Schaetzl, R., 2009. Soils. In Schaetzl, R., Darden, J., and Brandt, D. (Editors), Michigan Geography and Geology, Pearson Custom Publishing, 315-329

Selway, J.B., Breaks, F.W., Tindle, A.G., 2005. A review of rare-element (Li-Cs-Ta) pegmatite exploration techniques for the Superior Province, Canada, and large worldwide tantalum deposits. Exploration and Mining Geology, 14: 1-30.

Silva, F.M., Silva, S.H.G., Acuña-Guzman, S.F., Silva, E.A., Ribeiro, B.T., Fruett, T., Inda, A.V., Teixeira, A.F. dos S., Mancini, M., Guilherme, L.R.G., Curi, N., 2021. Chemical and mineralogical changes in the textural fractions of quartzite-derived tropical soils, along weathering, assessed by portable X-ray fluorescence spectrometry and X-ray diffraction. Journal of South American Earth Sciences 112, 103634. https://doi.org/10.1016/j.jsames.2021.103634

Silva, S.H.G., Ribeiro, B.T., Guerra, M.B.B., de Carvalho, H.W.P., Lopes, G., Carvalho, G.S., Guilherme, L.R.G., Resende, M., Mancini, M., Curi, N., Rafael, R.B.A., Cardelli, V., Cocco, S., Corti, G., Chakraborty, S., Li, B., Weindorf, D.C., 2021. Chapter One - pXRF in tropical soils: Methodology, applications, achievements and challenges, in: Sparks, D.L.B.T.-A. in A. (Ed.), Academic Press, pp. 1–62. https://doi.org/10.1016/bs.agron.2020.12.001

Sims, P.K., 1980. Boundary between Archean greenstone and gneiss terranes in northern Wisconsin and Michigan, in: Geological Society of America Special Papers. Geological Society of America, pp. 113–124. https://doi.org/10.1130/SPE182-p113

Sims, P.K, 1992. Geologic map of Precambrian rocks, southern Lake Superior region, Wisconsin and northern Michigan, 1:500.000. USGS Miscellaneous Investigations Series Map I-2185. https://doi.org/10.3133/i2185

Sims, P.K., Schulz, K.J., Peterman, Z.E., 1992. Geology and geochemistry of early Proterozoic rocks in the Dunbar area, northeastern Wisconsin, U.S. Geological Survey Professional Paper 1517. https://doi.org/10.3133/pp1517

Sirbescu, M.L.C., Hartwick, E.E., Student, J.J., 2008. Rapid crystallization of the Animikie Red Ace Pegmatite, Florence county, northeastern Wisconsin: Inclusion microthermometry and conductive-cooling modeling. Contributions to Mineralogy and Petrology 156, 289–305. https://doi.org/10.1007/s00410-008-0286-0

Sirbescu, M.L.C., Leatherman, M.A., Student, J.J., Beehr, A.R., 2009. Apatite Textures and Compositions as Records of Crystallization Processes in the Animikie Red Ace Pegmatite Dike, Wisconsin, USA. Canadian Mineralogist, 47: 725-743. 10.3749/canmin.47.4.725

Smith, R.E., Perdrix, J.L., and Davis, J.M., 1987, Dispersion into pisolitic laterite from the greenbushes mineralized Sn-Ta pegmatite system, Western Australia: Journal of Geochemical Exploration, v. 28, no. 1, p. 251–265.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey. accessed at https://websoilsurvey.sc.egov.usda.gov/app/ (accessed 10.16.23).

Steiner, B., 2019a. W and Li-Cs-Ta geochemical signatures in I-type granites – A case study from the Vosges Mountains, NE France. Journal of Geochemical Exploration. 197: 238-150. https://doi.org/10.1016/j.gexplo.2018.12.009

Steiner, B., 2019b. Tools and Workflows for Grassroots Li–Cs–Ta (LCT) Pegmatite Exploration. Minerals 9, 499. https://doi.org/10.3390/min9080499

Stockmann, U., Cattle, S.R.R., Minasny, B., McBratney, A.B., 2016. Utilizing portable X-ray fluorescence spectrometry for in-field investigation of pedogenesis. Catena 139, 220–231. https://doi.org/10.1016/j.catena.2016.01.007

Sweetapple, M.T., Tassios, S., 2015. Laser-induced breakdown spectroscopy (LIBS) as a tool for in situ mapping and textural interpretation of lithium in pegmatite minerals. American Mineralogist. 100, p. 2141–2151. https://doi.org/10.2138/am-2015-5165.

Tang, Y., Wang, H., Zhang, H., Lv, Z.-H., 2018. K-feldspar composition as an exploration tool for pegmatite-type rare metal deposits in Altay, NW China. Journal of Geochemical Exploration 185, 130–138. https://doi.org/10.1016/j.gexplo.2017.11.015

Ting, K.M., 2010. Confusion Matrix, in: Encyclopedia of Machine Learning and Data Mining.

Trueman, D.L., 1982. Exploration for rare-element granitic pegmatites. Short course in granitic pegmatites in science and industry 463–493.

Turner, D.J., and Young, I., 2008, Geological assessment report on the SELWYN 1-10 claims: Victoria, British Columbia, Canada, War Eagle Mining Company, accessed at http://yma.gov.yk.ca/095100.pdf.

USGS, 2023, Mineral Commodity Summaries 2023, accessed at https://doi.org/10.3133/mcs2023 (accessed 10.16.23).

Van Daalen, C.M., 2015. Mineralogy and geochemistry of Late Archean and Paleoproterozoic granites and pegmatites in the Northern Penokean terrane of Marquette and Dickinson Counties, Michigan. University of New Orleans Theses and Dissertations. 2088, 385 pp. https://scholarworks.uno.edu/td/2088

Weindorf, D.C., Chakraborty, S., 2020. Portable X-ray fluorescence spectrometry analysis of soils. Soil Science Society of America Journal 84, 1384–1392. https://doi.org/10.1002/saj2.20151

Wise, M.A., Harmon, R.S., Curry, A., Jennings, M., Grimac, Z. and Khashchevskaya, D., 2022. Handheld LIBS for Li exploration: an example from the carolina tin-spodumene belt, USA. Minerals, 12(1), p. 77. https://doi.org/10.3390/min12010077

Wise, M.A., Curry, A.C. and Harmon, R.S., 2024. Reevaluation of the K/Rb-Li Systematics in Muscovite as a Potential Exploration Tool for Identifying Li Mineralization in Granitic Pegmatites. Minerals, 14(1), p.117. https://doi.org/10.3390/min14010117

Whitney, D.L., Evans, B.W., 2010. Abbreviations for names of rock-forming minerals. American Mineralogist 95, 185-187. https://doi.org/10.2138/am.2010.3371

Xu, Z.Q., Liang, B., Geng, Y., Liu, T., Wang, Q.B., 2019. Extraction of soils above concealed lithium deposits for rare metal exploration in Jiajika area: A pilot study. Applied Geochemistry, 107: 142-151. https://doi.org/10.1016/j.apgeochem.2019.05.018

Yuan, Z., Chang, H., Zhou, S., Zhang, Z., Cheng, Q., Xia, Q., Zuo, R., Zhang, S., Wang, H., 2021. In situ monitoring of elemental losses and gains during weathering using the spatial element patterns obtained by portable XRF. Journal of Geochemical Exploration 229. https://doi.org/10.1016/j.gexplo.2021.106842

Zhao, L., Fang, Q., Hong, H., Algeo, T.J., Lu, A., Yin, K., Wang, C., Liu, C., Chen, L., Xie, S., 2022. Pedogenic-weathering evolution and soil discrimination by sensor fusion combined with machine-learning-based spectral modeling. Geoderma 409, 115648. https://doi.org/10.1016/j.geoderma.2021.115648

Zi, J.-W., Sheppard, S., Muhling, J.R., Rasmussen, B., 2022. Refining the Paleoproterozoic tectonothermal history of the Penokean Orogen: New U-Pb age constraints from the Pembine-Wausau terrane, Wisconsin, USA. GSA Bulletin 134, 776–790. https://doi.org/10.1130/B36114.1

APPENDIX 1

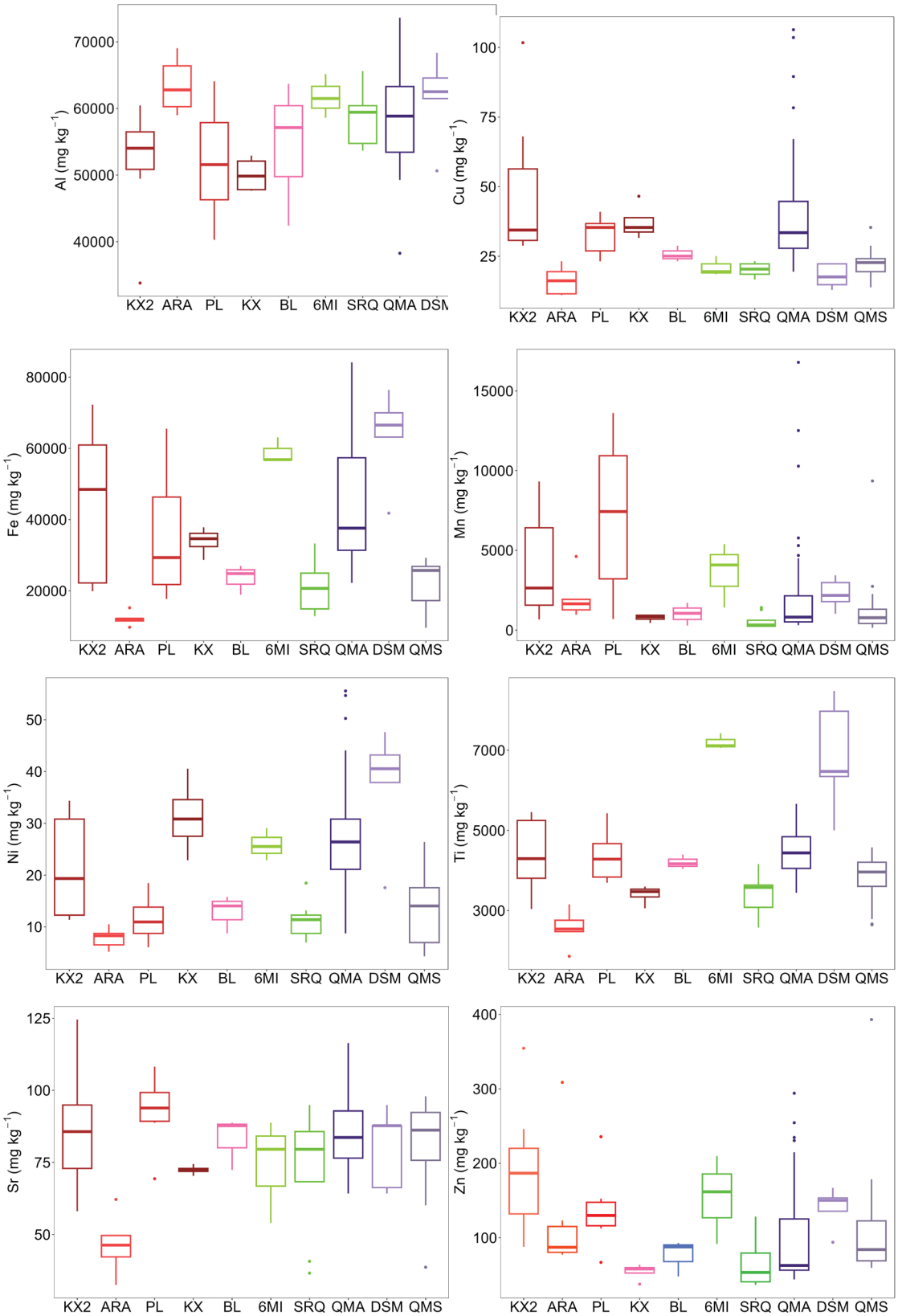

Figure A1. Supplementary box plots of soil elemental concentrations (in mg·kg-1) resulted from inductively coupled plasma optical emission spectroscopy (ICP-OES) analysis (Li) and portable X-ray fluorescence (pXRF) analysis (all other elements) not plotted in Fig. 5. Parental

material abbreviations: KX2: King's-X2; ARA: Animikie Red Ace; KX: King's-X; PL: Price Lake; BL: Bush Lake granite; 6MI: Six-mile Lake; SRQ: Sturgeon River Quarry; QMA: Quinnesec mafic amphibolite; DSM: Dickinson Six-mile mafic amphibolite; QMS: Quinnesec mica schist.

APPENDIX 2

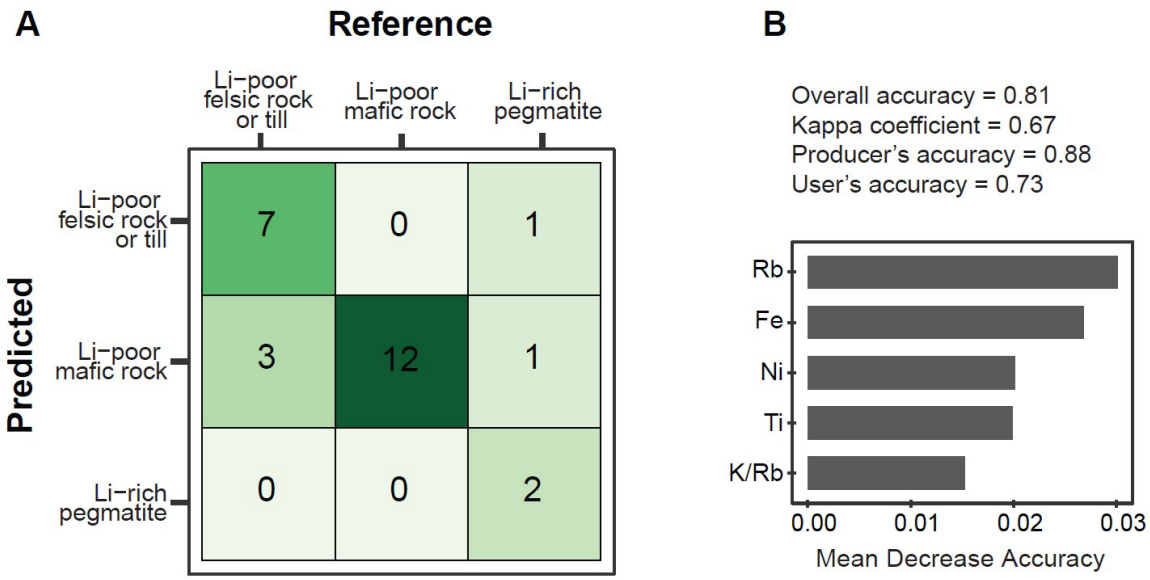

Figure A2. Model D0, a parental material prediction model via portable X-ray fluorescence and *Random Forest* algorithm at Florence County, WI and Dickinson County, MI.

A. Confusion matrix, where "felsic rock or till" is the number of validation soils formed on granite (BL), Sturgeon River Pegmatite (SRQ), Six-mile Lake Pegmatite (6MI), Quinnesec felsic metavolcanics and metasediments (QMS), and soils formed on glacio-fluvial sediment (KX); "mafic rock" represents the validation soils formed on Li-poor mafic lithologies: mafic amphibolites (QMA and DSM); and "Li rich pegmatite" represents the validation soils formed on: Animikie Red Ace (ARA), King's X2 (KX2), and Price Lake (PL).

In contrast to Model D (see text), all the soil samples formed on metamorphic bedrock were included in first or second category, regardless of their distance from the visible contact with the Li-rich pegmatites or their K:Rb ratio < 275.

B. Classification accuracy parameters, and top-5 variable importance for parent material prediction.

ARTICLE 3 – ARE DATA FROM THE SAME SET OF SAMPLES OBTAINED BY DIFFERENT PORTABLE X-RAY FLUORESCENCE SYSTEMS COMPARABLE?

*This article was prepared following the format style of Geoderma Regional Journal

ABSTRACT

Although portable X-ray fluorescence (pXRF) has been widely used worldwide for fast chemical characterization of multiple materials, a question that remains concerning whether pXRF data obtained from different models and brands generates different results due to internal calibration. For this reason, this study investigates the performance of two pXRF models and their factory calibrations to assess the effectiveness of these instruments in determining K, Ca, Fe, Ti, Rb, Mn, Al, P, and Si concentrations across various sample types, including 13 certified reference materials (CRMs, soils, tills, and pegmatite ores) and 115 unknown soil samples. The models used were Bruker Tracer 5i, with Soil mode and GeoExploration mode, and Olympus Vanta "M" series Geochem mode. The objectives were to compare CRMs' reported values with elemental concentrations obtained with both pXRFs, to compare both pXRFs models and their factory setup calibrations, to compare similar methods between both models, and create a correction factor (CF) that allows the use of data from different pXRFs together for different research projects. Results indicated an overall good correlation ($R^2 \geq 0.70$) between element concentration from CRMs and pXRF results. Comparisons of different pXRF models show that the results for some elements for the same samples are comparable, indicating that using data from different equipment may be possible. The use of CFs for data correction has shown promise for comparing different pXRF models. In general, pXRF obtained closer results to CRMs-reported concentrations than results from digestion. Moreover, the use of total digestion may not bring the best results for the analysis of elemental content compared to pXRF. However, further investigations into pXRF performance are still encouraged.

Keywords: pXRF; evaluation; data quality; geochemistry; correction factor

1. INTRODUCTION

In geochemical analysis, selecting appropriate instrumentation and methodologies plays a pivotal role in ensuring accurate and reliable results. Portable X-ray fluorescence (pXRF) spectrometry has been widely used in recent decades, especially for its ease of use, reliability, fast, and environmentally friendly method. For this reason, pXRF has been used in several areas of knowledge, including geochemistry. Several works have been carried out in recent years with different purposes, such as predicting soil parent material (Gozukara; Zhang; Hartemink,

2021a; Mancini *et al.*, 2020a; Pierangeli *et al.*, 2023b), prediction of soil properties (Andrade *et al.*, 2020; Coblinski *et al.*, 2020; Naimi *et al.*, 2022), contamination monitoring (Horta *et al.*, 2021; Sá *et al.*, 2023), and assessment of pedogenesis and weathering (Silva *et al.*, 2018; Stockmann *et al.*, 2016; Zhao *et al.*, 2022b), among others.

Although pXRF is widely used worldwide and offers promising results overall, some challenges still exist. One of them is the fact that each pXRF comes with an installed inbuilt specific factory configuration. However, developing specific calibration curves for different matrices is recommended to optimize the quality of the data obtained (Rouillon; Taylor, 2016), which can improve the results obtained. Conversely, developing calibration curves is the most recommended in the literature since the curve created will be stored in the equipment's software. However, this requires time, financial resources, and deeper knowledge of the equipment. The process is not simple, as it involves using samples with known and certified elemental content and understanding how each element can interfere with the reading of others (Acquah et al., 2022; Da Silva; Triantafyllou; Delmelle, 2023).. Furthermore, it is recommended to create calibration curves for a set of samples with similar matrices, which is not common in soils. For this reason, factory calibration is often adopted when analyzing different matrices, which can generate inconsistency in the results. Furthermore, X-ray tubes have a “shelf life” and equipment is subject to damage, whether due to misuse or falls. Thus, a method capable of converting results from one pXRF equipment to another is of utmost importance.

Specifically, some questions arise in this context. For instance, can results from different pXRF models be comparable between research groups? In case a pXRF models is broken, needing to be replaced, it would be necessary to reanalyze a dataset of samples? These are interesting questions, especially given the current scenario of data sharing and interdisciplinarity. We hypothesize that results from the same set of samples are not comparable when analyzed by different pXRF models, but that correction factors may be created between pXRF models. For this reason, this study delves into the performance and comparability of two pXRF models, Bruker Tracer 5i and Olympus Vanta “M” series, focusing on their analytical capabilities and applications. Employing each pXRF factory-made calibration, this study aims to investigate the effectiveness of these instruments in determining elemental concentrations across various sample types, including certified reference materials (CRMs, soils, tills, and pegmatite ores) and unknown soil samples. Specifically, the objectives were: I) to compare CRMs reported values with elemental concentrations obtained with both pXRFs; II) to assess the results of two setup calibrations of pXRF Bruker Tracer 5i, Soil mode and GeoExploration

modes; and III) to confront results between Bruker Tracer 5i and Olympus Vanta "M" series obtained using similar configurations, GeoExploration and Geochem, respectively.

2. MATERIAL AND METHODS

2.1 pXRF Instrumentation

This study investigated the performance and comparability of two pXRF systems, Bruker Tracer 5i and Olympus Vanta "M" series (Table 1). The analysis was conducted in lab conditions under air atmosphere. pXRF factory-made calibrations were used in the study. Bruker Tracer 5i was run in two modes (Soil mode and GeoExploration method) and Olympus Vanta "M" series in one mode (Geochem).

Soil mode method is specifically for soil analysis and includes two beams with two sequential operational conditions. For this study, each beam was run for 60 s, totaling 120 s of analysis. Each beam mode utilizes a different voltage to enhance the fluorescence for specific elements. Both devices offer a third beam mode designed for light elements. GeoExploration and Geochem were subject to the same analysis time to investigate instrument variation. The analysis time greatly influences the results. Longer analysis time can improve detection limits for elements and enhance the signal-to-noise ratio, leading to more precise readings. Each model has a similar application mode used for geochemical analysis including three sequential beams to optimize the quantitative analysis of elements from Mg to U. Each beam has a specific energy range, prioritizing certain elements (Table 1).)

Table 1. Parameters and conditions used with Bruker Tracer 5i and Olympus Vanta "M" series

Method	Current and voltage	Application	Factory calibration modes	Beam time
Bruker Tracer 5i	4.5-195 µA 50 kV	Soil	Soil mode	Beam 1: 60 s Beam 2: 60 s

Bruker Tracer 5i	4.5-195 µA 50 kV	Geochemical	GeoExploration	Beam 1: 40 s Beam 2: 40 s Beam 3: 90 s
Olympus Vanta "M" Series	5-200 µA 50 kV	Geochemical	Geochem	Beam 1: 90 s Beam 2: 40 s sBeam 3: 40 s

2.2. Certified reference materials and soil samples

Thirteen CRMs were scanned from pXRF models to assess data quality (Table 2). The CRMs consisted of powdered and homogenized soils (OREAS 25a, NIST 2709, 2710a, 2711), tills (OREAS 47, SRM TILL 2 and 4), and Li or Li-Nb-Sn pegmatites (OREAS 147, 148, 750, 751, 752, 753) (Ore Research and Exploration, Bayswater North, Victoria, Australia; NIST, Gaithersburg, Maryland, USA; CCRMP, Ottawa, Ontario, Canada).

Table 2. Summary of comparison test using different portable X-ray fluorescence.

Comparison test	Number of samples	Type of sample
Soil mode vs. GeoExploration vs. Geochem	13	CRMs
Soil mode vs. GeoExploration	115	Unknown soil samples
GeoExploration vs. Geochem	31	Unknown soil samples

Soil and rock samples were collected from two locations. The first location was in Florence County, near the northernmost edge of the Paleoproterozic Wisconsin Magmatic Terrane in northeastern Wisconsin, USA (Zi *et al.*, 2022). The second location was at the southernmost edge of the Superior Province, in central Dickinson County, Michigan (Ayuso *et al.*, 2018; Cannon, 2018; Sims, 1980). The region was affected by prolonged glaciation starting at least approximately 780,000 years ago, with the Wisconsin event ending in the area approximately 10,000 years BP, leaving behind a characteristic hummocky terrain with glacial till and abundant erratics. The soils in both locations are young and have undergone minimal weathering, resembling their parent materials (Schaetzl, 2009). The humid continental climate in the region is classified as Dfb (Köppen), with the "warmest" month having an average temperature below 22°C and the coldest month having an average temperature below freezing, 0°C (CEC, 2021). Overall, 115 soil samples were collected from soil horizons (O, A, and E)

down to a maximum of 40 cm depth from Spodosols and Alfisols (Soil Survey Staff, 2023), with the Spodosols found in Florence County and the Alfisols found in Dickinson County.

The soil samples were first dried in an oven at approximately 65°C. They were then crushed into fine particles using a mortar and pestle, and any organic fragments larger than 2mm were removed through sieving. The samples were then divided into four equal parts using a sample splitter. One of these parts was pulverized in a shatter box to ensure complete mixing. The homogenized soil sample was placed in a Spectrocup© XRF sample cup pre-covered with a 4.0 μm thin film made of Prolene® for pXRF analysis.

2.3. Data processing

The results were compared via linear regressions. The results from the pXRF models were compared to each other, as well as the concentrations reported by the CRMs and the digestion results. The coefficient of determination (R^2) and 1:1 line were used to evaluate the performance of the pXRF models and whether the data generated is reliable.

2.4 Correction factor

A simple correction factor (CF) was created to resolve variation in results obtained with pXRF. Two CRMs (NIST2701 and OREAS25a) were selected to generate this factor. Both CRMs consist of soil samples, with NIST2710a (Montana I Soil) being soil with high concentrations of trace elements and OREAS25a (ferruginous soil) being a weathered soil developed from basalt. To create the CF, these CRMs were analyzed, and the recovery (elemental content obtained by pXRF)/(certified elemental content) of each element was calculated (Table 3). After, the CF obtained was applied to the pXRF data (Equation 1).

$$x_a = y_a * CF_b \quad (1)$$

where x_a is the pXRF data transformed, y_a is the data obtained with pXRF A and, CF_b is the correction factor obtained for pXRF B.

Table 3. Average recovery values of NIST2710a and OREAS25a in percentages for Soil mode, GeoExploration, and Geochem factory calibration.

Equipment	K	Ca	Fe	Ti	Mn	P	Al	Si
NIST2710a								

Soil mode	88	79	96	82	87	27	76	82
GeoExploration	80	86	98	86	112	58	90	75
Geochem	89	74	104	103	100	71	73	65
				OREAS25a				
Soil mode	6	100	107	131	120	33	101	97
GeoExploration	96	103	114	128	128	43	118	86
Geochem	84	68	112	124	107	66	83	70

3. RESULTS AND DISCUSSIONS

3.1 Comparison between pXRF methods analyzing 13 CRMs

The effectiveness of two pXRF units and three application modes for measuring elemental concentrations was assessed by comparing their recovery rates to 13 CRMs, as shown in Figure 1. Correlations between the measured K, Ca, Fe, Ti, Rb, Mn, P, Al, and Si content with pXRF and the reported standard reference showing R^2 were used to compare measurements obtained with the pXRFs. Except for P, all the other modes obtained promising results for the elements assessed. pXRF GeoExploration presented overall superior results, followed by Soil mode and Geochem. Most regression models presented excellent correlations ($R^2 \geq 0.95$), except for Al and Si, which showed a good correlation ($R^2 \geq 0.70$), and P which showed different behaviors for each pXRF and method. Both GeoExploration and Geochem have the addition of a third beam light and silicon drift detectors (SDD) technology, which is optimized for the detection of light elements such as P, Si, and Al (Adams et al., 2020). However, improvement in detection of P was not observed (Table 3).

Depending on the model and method, pXRF can overestimate or underestimate the elemental content. Bruker soil mode and GeoExploration methods were compared in regard of their detection performance (refer to section 2.1). Those modes present different conditions and calibration curves, so different results are expected. Each application has a set of matrix-matched samples for the calibration curve (Acquah et al., 2022; Johnson et al., 2021; Lu et al., 2022). The Soil mode uses soil matrix samples for the calibration curve. Meanwhile, a geochemical application mode uses samples from different matrices for calibration (i.e. mudrock, limestone, soil, till). For example, the Geochem method overestimated Al, and the linear regression presented an R^2 of 0.94. Meanwhile, for Al different performances were obtained depending on the mode (Soil mode and GeoExploration). Most elements determined with the different pXRFs showed close agreement ($R^2 \geq 0.95$) with the reported reference;

however, most did not follow 1:1 correspondence with intercept equal to zero and slope almost equal to one. All pXRFs tested tended to underestimate the concentration for K, P, and Si. Meanwhile, all equipments tested tended to overestimate Fe and Ti. On the other hand, the pXRFs presented different behaviors regarding Rb, Ca, and Mn. The sample's composition (matrix) influences the accuracy and precision of element measurements, mainly reflected by the absorption and enhancement or overlap in the spectral peaks (Lu et al., 2022).

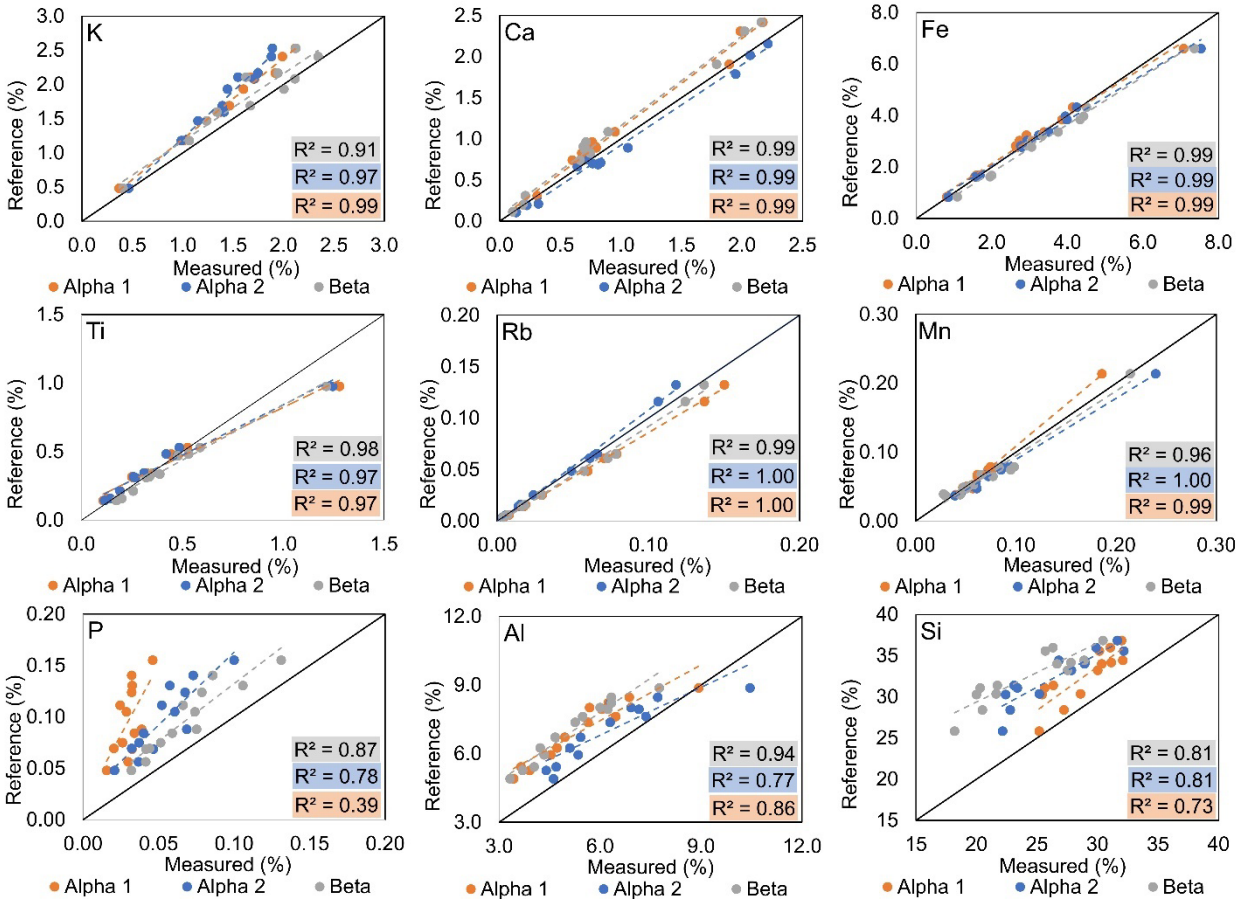

Figure 1. Comparison between results for 13 CRMs samples using Soil mode, GeoExploration and Geochem modes for K, Ca, Fe, Ti, Rb, Mn, P, Al, and Si.

The detection and measurement of elements using different models and methods are subject to varying calibration and detection limits. As a result, the list of elements that can be measured is also affected. Major elements, including Al, Ca, Fe, K, and Mg, as well as base metals, can be reported by all modes. Additionally, minor elements, such as Cd, Sb, Sn, and Rb, can also be detected. However, there are some challenges in detecting light energy elements like Al, Si, and K, as observed in Figure 1. Conversely, heavier energy elements such as Fe, Ca, Rb, Mn, and Ba are detected more accurately. To ensure better detection, beam 3 in pXRF

GeoExploration and beam 1 in Geochem, which are optimized for a wide range of elements (40 kV), were prioritized in this study, leading to a longer analysis time. This prioritization may be the reason for the better agreement with the 1:1 line and high R^2 values for Rb and Mn using pXRF GeoExploration and Geochem compared to the reference value.

3.2 Comparison between pXRF models analyzing unknown soil samples

Figure 2 shows the correlation between Soil mode and GeoExploration mode, where 115 unknown soil samples were compared. Most elements had excellent correlations (R2 ≥ 0.90) between modes, except Si, which presented a good correlation, and Al, which showed poor R2 = 0.25. Fe and Ca (R2 of 0.99 and 1.00, respectively) presented a close 1:1 correspondence between modes, with no tendencies of overestimating or underestimating results. Elements Si, K, Rb, and Ti tended to underestimate results when using GeoExploration, especially at high concentrations, as observed for K, Rb, and Ti. For Si, all concentrations were underestimated using GeoExploration. On the other hand, Mn, P, and Al overestimated the concentration in GeoExploration compared to Soil mode.

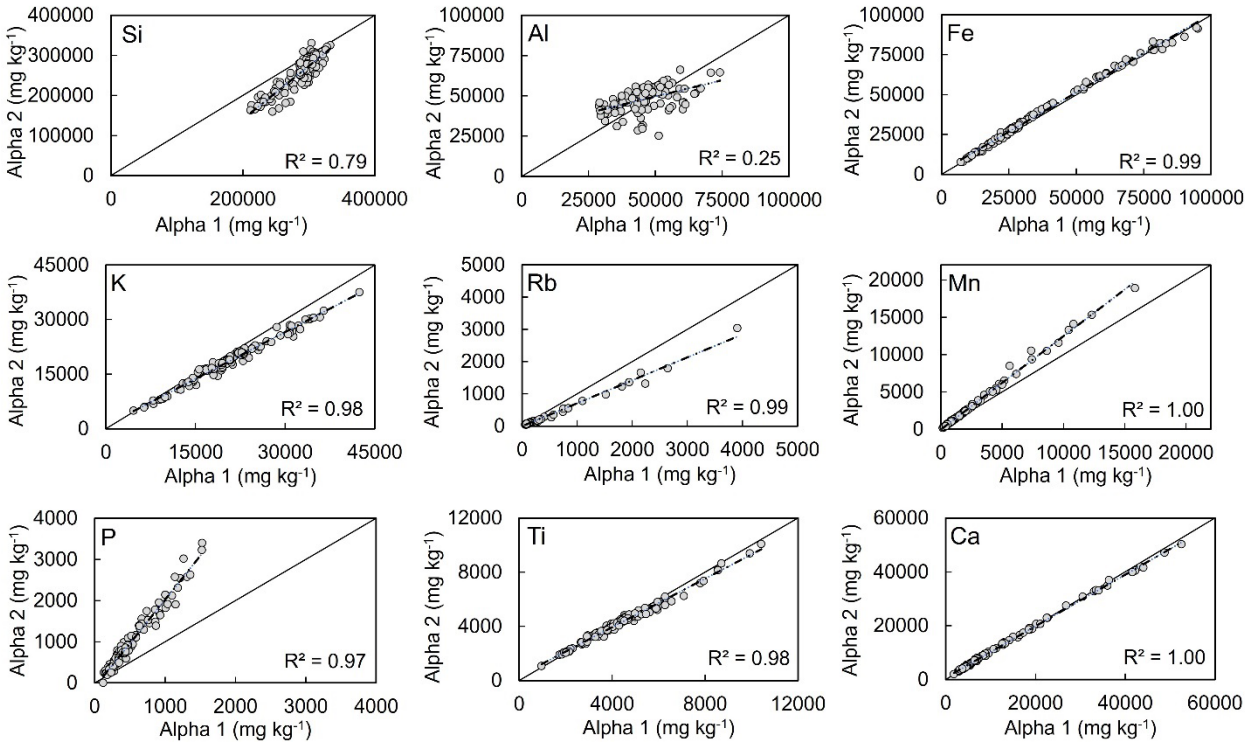

Figure 2. Comparison between results for 115 unknown soil samples using Soil mode and GeoExploration modes for Si, Al, Fe, K, Rb, Mn, P, Ti, and Ca.

To compare pXRFs GeoExploration and Geochem (Figure 3), 31 unknown samples were analyzed with similar factory-made calibrations. All elements presented R^2 greater than 0.90 when comparing different pXRF modes, showing outstanding correlation. On the other hand, most elements did not follow 1:1 correspondence, except for K, Rb, and Ti, whose slopes were closer to the line. pXRF Geochem overestimated Si, Al, and Ca concentrations compared to pXRF GeoExploration. However, Fe, P, and Mn were overestimated at higher concentrations with Geochem. Conversely, Rb and Ti concentrations tended to be overestimated using pXRF GeoExploration.

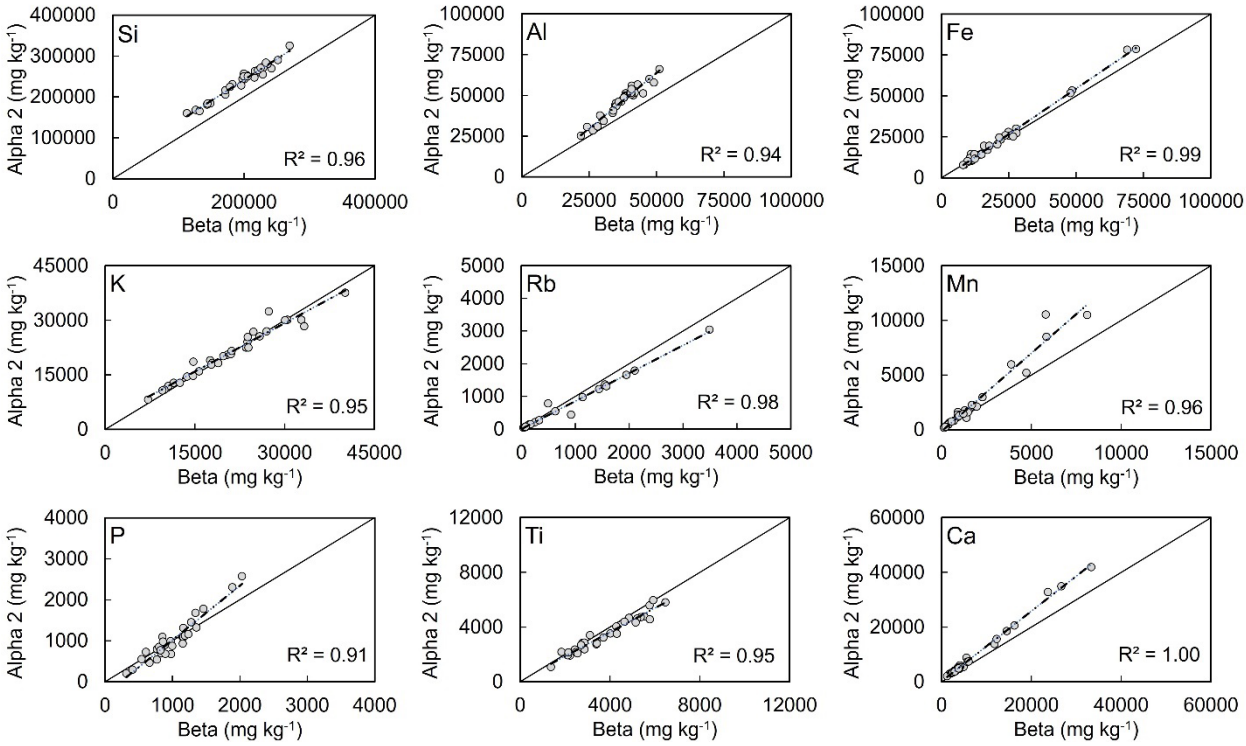

Figure 3. Comparison between results for 31 unknown soil samples using GeoExploration and Geochem modes for Si, Al, Fe, K, Rb, Mn, P, Ti, and Ca.

Although most elements showed great correlations and similarities between pXRFs modes. It is clear the difference when analyzing the correlation between Al concentration with Geochem and GeoExploration, and Soil mode and GeoExploration. Both applicaitons had a poor correlation (R^2 of 0.25). However, when compared to pXRF Soil mode and Geochem, Al had an R^2 of 0.94. This difference may be due to Al being a light element, and pXRFs GeoExploration and Geochem using a three-beam light scanning set, where each beam focuses

on a different range of energy. It is also possible that the smaller number of samples used for comparison between GeoExploration and Geochem influenced the dispersion of the points.

3.4 Comparison between pXRF modes analyzing unknown soil samples after a CF was applied

The initial assessment of the effectiveness of the two pXRF modes showed great correlation for most of the elements when analyzed with the different modes (previous sections). However, it tended not to follow a 1:1 correspondence line, which means that they correlated (R^2) but the concentrations are far from similar. For this reason, we proposed the use of a CF to aid with this challenge. Figure 6 shows the original (orange color) data without the CF approach and the same data after the CF was applied to the data of pXRF Soil mode and 2. As observed, R^2 remains constant since the data with the CF only increases in proportion compared to the original data since the data was multiplied by the CF. On the other hand, the correspondence with the 1:1 line increased in some cases, such as Si, Rb, and P. It is important to note that the sample matrix can influence the analyses; both CRMs are soil samples, although they are different in matrix.

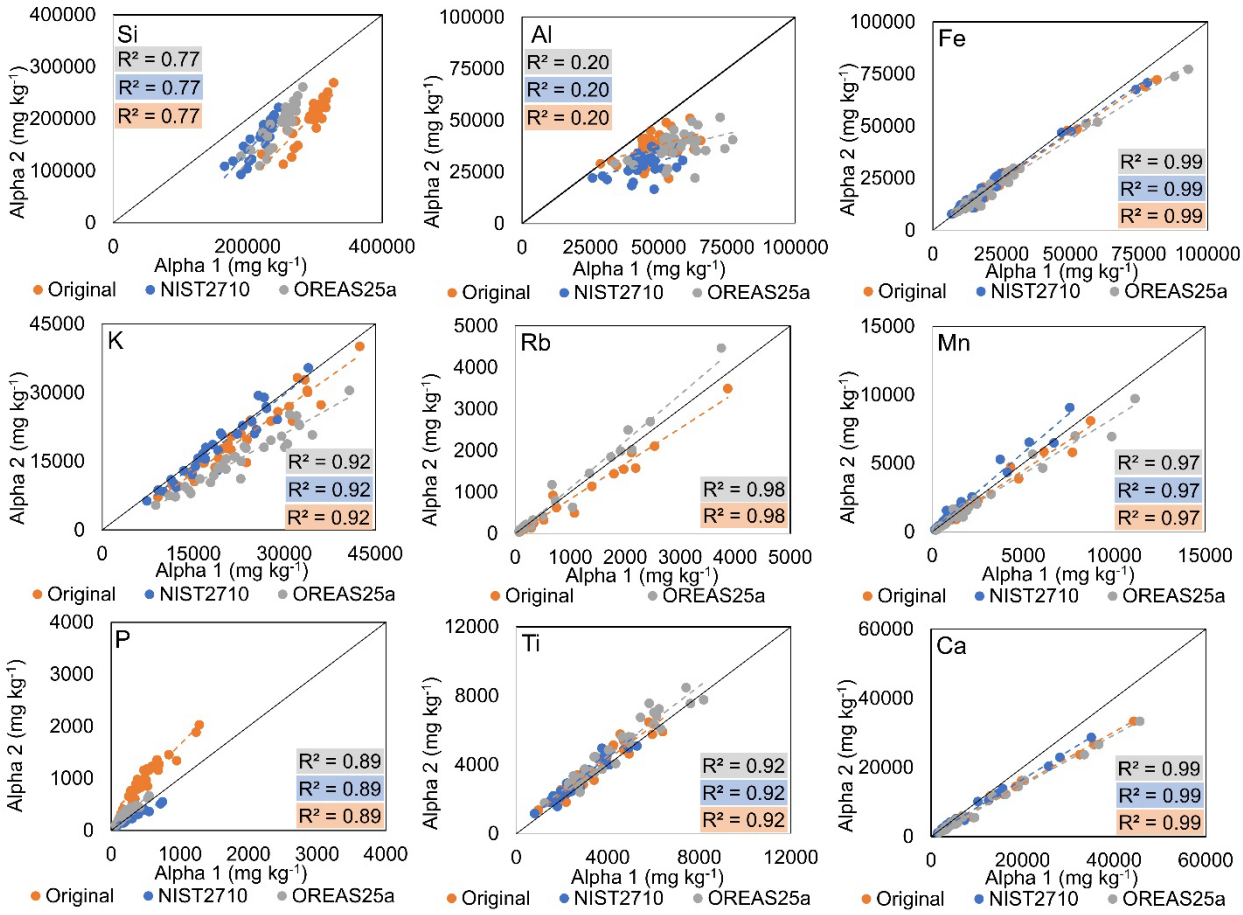

Figure 6. Comparison results of 31 unknown samples with and without a Correction factor using CRMs NIST2710 and OREAS25a between pXRFs Soil mode and GeoExploration modes for K, Rb, Mn, Fe, Ti, and P.

Figure 7 shows the comparison between model Geochem and GeoExploration, where the CF was applied to 31 unknown soil samples. The same trend was observed where CF data presented a close slope with the 1:1 line. For elements such as Si, Rb, Al, Fe, and Ca, the comparison between GeoExploration and Geochem got significantly better, with a slope close to 1:1. For P, the CF made the overestimated tendencies of pXRF Geochem more accentuated. In some cases, the CF showed no improvement in comparison with the original data, as seen with Al and Ca with CF from NIST2710 and Fe with OREAS25a CF.

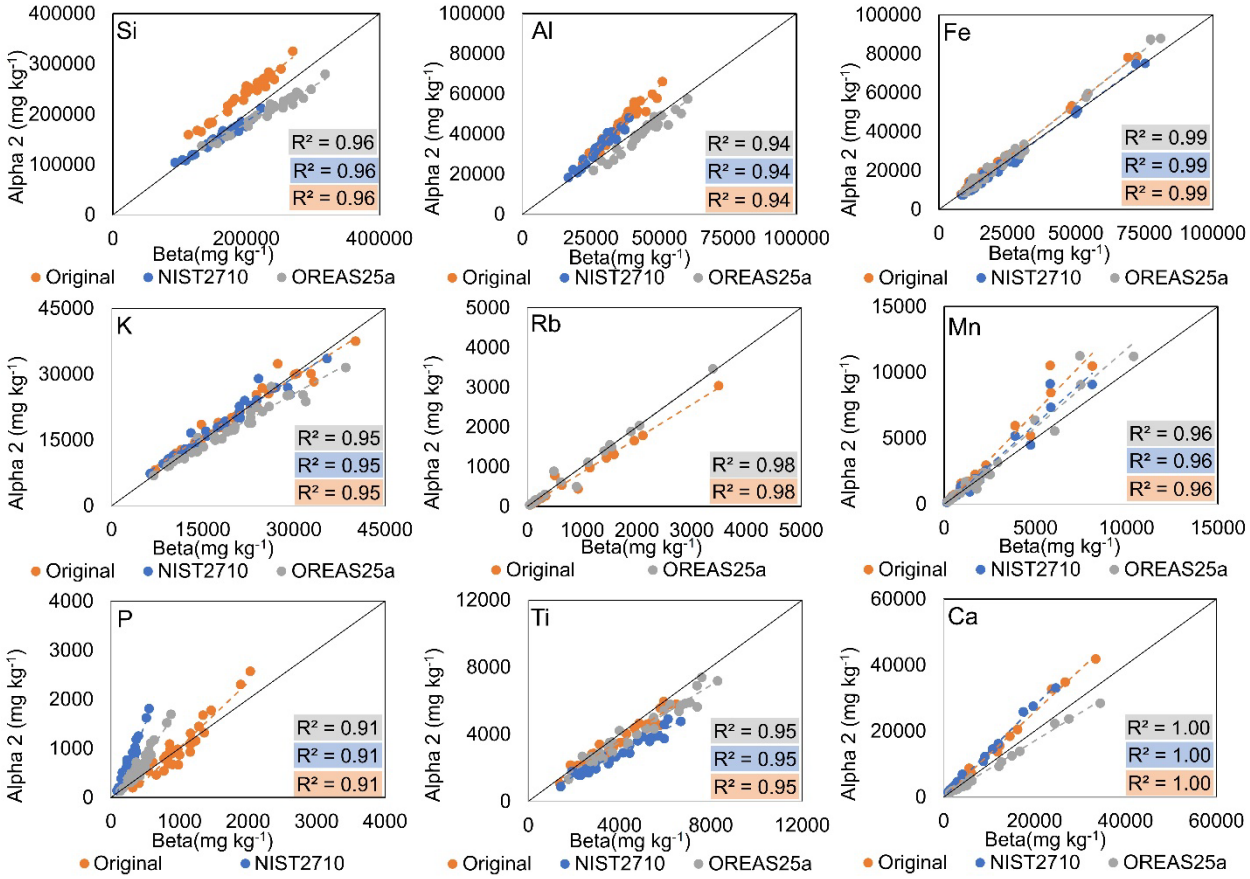

Figure 7. Comparison results of 31 unknown samples with and without a Correction factor using CRMs NIST2710 and OREAS25a between pXRFs GeoExploration and Geochem modes for K, Rb, Mn, Fe, Ti, and P.

Although pXRF technology has encountered certain obstacles, it remains a dependable and useful method for elemental analysis of soil samples, making it a crucial asset in geochemical research and analysis. To tackle the issue of inconsistent data output across various pXRF modes, a CF could be established to standardize data and simplify comparison. It's worth noting, however, that the choice of pXRF configuration and CRM may impact the results achieved. Using a CF could enable the sharing of data gathered from diverse sources, enabling researchers to sharpen their findings. Nevertheless, additional research is necessary to refine this concept.

4. CONCLUSIONS

Different elements (K, Ca, Fe, Ti, Rb, Mn, Al, P, and Si) were determined in different samples (soil, till, and pegmatite ores) with different compositions. The analytical performance of the application modes was assessed by CRM recoveries. In general, all pXRFs had a great outcome overall. It is not possible to choose the best one due to each pXRF having a calibration curve based on different samples. Tests comparing the results obtained with different pXRFs show that the results for some elements are comparable, indicating that using data from different pXRF modes may be possible. The comparative assessment of different pXRF methods for determining elemental concentrations in geochemical samples underscores the nuanced performance variations across various elements and methodologies. The use of a CF to correct data from different pXRFs modes and make them comparable is a nice approach to overcome challenges with data from different sources. These findings confirm the importance and power of pXRF for the rapid and accurate analysis of the concentration of different chemical elements.

5. ACKNOWLEDGMENTS

The authors would like to thank the National Council for Scientific and Technological Development (CNPq), Coordination for the Improvement of Higher Education Personnel (CAPES), and Foundation for Research of the State of Minas Gerais (FAPEMIG) for the financial support to develop this research. We are thankful for generous support from the Department of Earth and Atmospheric Science and the Office of Research and Graduate Studies at Central Michigan University (CMU) for hosting Luiza and supporting their 'sandwich' PhD research between Federal University of Lavras (UFLA), Brazil and CMU, Michigan. Thanks to Lithium Americas Corporation, Geological Society of America (GSA), and Mineralogy Geochemistry Petrology and Volcanology (MGPV) division of GSA for sponsoring this research, which allowed the use of soil geochemistry as an exploration tool for elements critical to green technologies.

6. REFERENCES

Acquah, G.E., Hernandez-Allica, J., Thomas, C.L., Dunham, S.J., Towett, E.K., Drake, L.B., Shepherd, K.D., McGrath, S.P., Haefele, S.M., 2022. Portable X-ray fluorescence (pXRF) calibration for analysis of nutrient concentrations and trace element contaminants in fertilisers. PLoS ONE 17, e0262460. https://doi.org/10.1371/journal.pone.0262460

Adams, C., Brand, C., Dentith, M., Fiorentini |, M., Caruso, S., Mehta, M., 2020. The use of pXRF for light element geochemical analysis: a review of hardware design limitations and an empirical investigation of air, vacuum, helium flush and detector window technologies. Geochemistry: Exploration, Environment, Analysis 20, 366–380. https://doi.org/10.1144/geochem2019-076

Andrade, R., Faria, W.M., Silva, S.H.G., Chakraborty, S., Weindorf, D.C., Mesquita, L.F., Guilherme, L.R.G., Curi, N., Missina, W., Henrique, S., Silva, G., Chakraborty, S., Weindorf, D.C., Felipe, L., Roberto, L., Guilherme, G., Curi, N., 2020. Prediction of soil fertility via portable X-ray fluorescence (pXRF) spectrometry and soil texture in the Brazilian Coastal Plains. Geoderma 357, 113960. https://doi.org/10.1016/j.geoderma.2019.113960

Ayuso, R.A., Schulz, K.J., Cannon, W.F., Woodruff, L.G., Vazquez, J.A., Foley, N.K., Jackson, J., 2018. New U-Pb Zircon Ages for Rocks from the Granite-Gneiss Terrane in Northern Michigan: Evidence for Events at ~3750, 2750, and 1850 Ma. Presented at the 64th Institute on Lake Superior Geology Proceedings, pp. 7–8.

Cannon, W.F., 2018. Part 2: Field Trip Guidebooks. Presented at the 64th Institute on Lake Superior Geology Proceedings.

CEC, 2021. Guide to Drought Indices and Indicators Used in North America. Commission for Environmental Cooperation, Montreal, QC, CA.

Coblinski, J.A., Giasson, É., Demattê, J.A.M., Dotto, A.C., Costa, J.J.F., Vašát, R., 2020. Prediction of soil texture classes through different wavelength regions of reflectance spectroscopy at various soil depths. Catena 189. https://doi.org/10.1016/j.catena.2020.104485

Gozukara, G., Zhang, Y., Hartemink, A.E., 2021. Using vis-NIR and pXRF data to distinguish soil parent materials – An example using 136 pedons from Wisconsin, USA. Geoderma 396, 115091. https://doi.org/10.1016/j.geoderma.2021.115091

Horta, A., Azevedo, L., Neves, J., Soares, A., Pozza, L., 2021. Integrating portable X-ray fluorescence (pXRF) measurement uncertainty for accurate soil contamination mapping. Geoderma 382. https://doi.org/10.1016/j.geoderma.2020.114712

Johnson, L.R., Ferguson, J.R., Freund, K.P., Drake, L., Duke, D., 2021. Evaluating obsidian calibration sets with portable X-Ray fluorescence (ED-XRF) instruments. Journal of Archaeological Science: Reports 39, 103126. https://doi.org/10.1016/j.jasrep.2021.103126

Lu, J., Guo, J., Wei, Q., Tang, X., Lan, T., Hou, Y., Zhao, X., 2022. A Matrix Effect Correction Method for Portable X-ray Fluorescence Data. Applied Sciences 12, 568. https://doi.org/10.3390/app12020568

Mancini, M., Silva, S.H.G., Teixeira, A.F. dos S., Guilherme, L.R.G., Curi, N., 2020. Soil parent material prediction for Brazil via proximal soil sensing. Geoderma Regional 22, e00310. https://doi.org/10.1016/j.geodrs.2020.e00310

Naimi, S., Ayoubi, S., Di Raimo, L.A.D.L., Dematte, J.A.M., 2022. Quantification of some intrinsic soil properties using proximal sensing in arid lands: Application of Vis-NIR, MIR, and pXRF spectroscopy. Geoderma Regional 28, e00484. https://doi.org/10.1016/j.geodrs.2022.e00484

Pierangeli, L.M.P., Silva, S.H.G., Teixeira, A.F. dos S., Mancini, M., Andrade, R., Menezes, M.D. de, Sirbescu, M.-L.C., Marques, J.J., Weindorf, D.C., Curi, N., 2023. Soil parent material spatial modeling at high resolution from proximal sensing and machine learning: A pilot study. Journal of South American Earth Sciences 129, 104498. https://doi.org/10.1016/j.jsames.2023.104498

Rouillon, M., Taylor, M.P., 2016. Can field portable X-ray fluorescence (pXRF) produce high quality data for application in environmental contamination research? Environmental Pollution 214, 255–264. https://doi.org/10.1016/j.envpol.2016.03.055

Sá, R.T.S. de, Prianti, M.T.A., Andrade, R., Silva, A.O., Batista, É.R., Santos, J.V.D., Silva, F.M., Carneiro, M.A.C., Guilherme, L.R.G., Chakraborty, S., Weindorf, D.C., Curi, N.,

Silva, S.H.G., Ribeiro, B.T., 2023. Detailed characterization of iron-rich tailings after the Fundão dam failure, Brazil, with inclusion of proximal sensors data, as a secure basis for environmental and agricultural restoration. Environmental Research 228, 115858. https://doi.org/10.1016/j.envres.2023.115858

Schaetzl, R.J., 2009. Michigan geography and geology. Pearson Learning Solutions.

Silva, A.C., Triantafyllou, A., Delmelle, N., 2023. Portable x-ray fluorescence calibrations: Workflow and guidelines for optimizing the analysis of geological samples. Chemical Geology 623, 121395. https://doi.org/10.1016/j.chemgeo.2023.121395

Silva, S.H.G., Hartemink, A.E., Teixeira, A.F. dos S., Inda, A.V., Guilherme, L.R.G., Curi, N., 2018. Soil weathering analysis using a portable X-ray fluorescence (PXRF) spectrometer in an Inceptisol from the Brazilian Cerrado. Applied Clay Science 162, 27–37. https://doi.org/10.1016/j.clay.2018.05.028

Sims, P.K., 1980. Boundary between Archean greenstone and gneiss terranes in northern Wisconsin and Michigan, in: Geological Society of America Special Papers. Geological Society of America, pp. 113–124. https://doi.org/10.1130/SPE182-p113

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture, 2023. Soil Survey Staff [WWW Document]. Web Soil Survey. URL https://websoilsurvey.sc.egov.usda.gov/app/ (accessed 10.16.23).

Stockmann, U., Cattle, S.R.R., Minasny, B., McBratney, A.B., 2016. Utilizing portable X-ray fluorescence spectrometry for in-field investigation of pedogenesis. Catena 139, 220–231. https://doi.org/10.1016/j.catena.2016.01.007

Zhao, L., Fang, Q., Hong, H., Algeo, T.J., Lu, A., Yin, K., Wang, C., Liu, C., Chen, L., Xie, S., 2022. Pedogenic-weathering evolution and soil discrimination by sensor fusion combined with machine-learning-based spectral modeling. Geoderma 409, 115648. https://doi.org/10.1016/j.geoderma.2021.115648

Zi, J.-W., Sheppard, S., Muhling, J.R., Rasmussen, B., 2022. Refining the Paleoproterozoic tectonothermal history of the Penokean Orogen: New U-Pb age constraints from the

Pembine-Wausau terrane, Wisconsin, USA. GSA Bulletin 134, 776–790. https://doi.org/10.1130/B36114.1

THIRD PART – FINAL CONSIDERATIONS

FINAL CONSIDERATIONS

The successful prediction of soil PM distribution using a random forest model trained on data obtained from pXRF and MS of B horizons highlights the potential of proximal sensor data in soil mapping, especially in areas where direct PM evaluation is challenging due to thick soil cover. This finding underscores the importance of leveraging advanced sensing technologies and machine learning techniques to improve soil mapping accuracy and efficiency, particularly in tropical soils where spatial variability is high.

The ability of pXRF and MS to detect relative differences in chemical/mineralogical traits across soils derived from diverse PMs suggests their potential application in uncovering links between geological formations and key soil properties. These results strengthen the case for using proximal sensor data alongside machine learning algorithms for PM prediction, even for PM types investigated for the first time. This approach holds promise for reducing costs and accelerating the assessment of soil PM spatial variability, supporting better agronomic management and sustainable decision-making in tropical environments.

The successful development of predictive machine learning algorithms for assessing Li contents and parental material of soils based on pXRF data demonstrates the potential of this approach for rapid and cost-effective exploration of Li-rich pegmatite deposits. Despite some overlaps, significant compositional differences in superficial soils derived from different parental materials were detected, enabling accurate prediction of Li concentrations and parental material types. This pilot exploration study lays the groundwork for further research to refine and expand this approach, with potential applications in various climates and soil types globally.

The findings highlight the significance and power of pXRF technology in providing rapid and accurate analysis of the concentration of different chemical elements. This confirms the versatility and reliability of pXRF as a valuable tool in geochemical analysis, offering opportunities for efficient and cost-effective elemental determination across diverse sample types. As advancements continue to improve pXRF technology and methodologies, its role in geochemical exploration, environmental monitoring, and resource assessment is expected to expand further, contributing to enhanced understanding and management of geological and environmental systems.