https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

Open Access Earth System Science Data Discussions

Digital soil mapping of lithium in Australia

Wartini Ng[1*], Budiman Minasny[1], Alex McBratney[1], Patrice de Caritat[2], John Wilford[2]

[1] Sydney Institute of Agriculture, School of Life and Environmental Sciences, The University of Sydney, NSW 2015, Australia
[2] Geoscience Australia, Canberra, ACT 2601, Australia

Correspondence to: Wartini Ng (wartini.ng@sydney.edu.au)

Highlights

- Machine learning can be used to relate soil data and environmental covariates
- The first Australian digital soil map of lithium content is presented
- The prediction map can be used to delineate potential areas for anomalous Li
- Elevated soil Li observed and modelled in a number of States/NT

Abstract. With a higher demand for lithium (Li), a better understanding of its concentration and spatial distribution is important to delineate potential anomalous areas. This study uses a digital soil mapping framework to combine data from recent geochemical surveys and environmental covariates to predict and map Li content across the 7.6 million km2 area of Australia. Soil samples were collected by the National Geochemical Survey of Australia at a total of 1315 sites, with both top (0–10 cm depth) and bottom (on average 60–80 cm depth) catchment outlet sediments sampled. We developed 50 bootstrap models using a Cubist regression tree algorithm for both depths. The spatial prediction models were validated on an independent Northern Australia Geochemical Survey dataset, showing a good prediction with a root mean square error of 3.82 mg kg-1 (which is 50.9 % of the inter-quartile range) for the top depth. The model for the bottom depth has yet to be validated. The variables of importance for the models indicated that the first three Landsat 30+ Barest Earth bands (blue, green, red) and gamma radiometric dose have a strong impact on Li prediction. The bootstrapped models were then used to generate digital soil Li prediction maps for both depths, which could select and delineate areas with anomalously high Li concentrations in the regolith. The map shows high Li concentration around existing mines and other potentially anomalous Li areas. The same mapping principles can potentially be applied to other elements. The Li geochemical data for calibration and validation are available at: (De Caritat and Cooper, 2011a; http://dx.doi.org/10.11636/Record.2011.020) and (Main et al., 2019; http://dx.doi.org/10.11636/Record.2019.002) respectively. The covariates data used for this study was sourced from Terrestrial Ecosystem Research Network (TERN) infrastructure, which is enabled by the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS) https://esoil.io/TERNLandscapes/Public/Products/TERN/Covariates/Mosaics/90m/ (TERN, 2019).

Keywords: digital soil mapping, lithium, machine learning, geochemical survey, mineral anomalous, critical element

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

Open Access Earth System Science Data Discussions

1 Introduction

Minerals have become essential commodities in modern human society. Many minerals are fundamental to technological and industrial advancement, particularly those utilised in renewable energy systems, electric vehicles, consumer electronics and telecommunications (Kabata-Pendias, 2010). These minerals can be considered critical, in the sense that they are of high importance and have a high risk of supply disruption. Methods for quantifying mineral criticality are discussed in detail in Graedel et al. (2012).

Lithium (Li) is an important chemical element as the world transitions towards a lower-carbon economy. It has been listed as one of the critical elements by various countries, including Australia, Canada, the European Union, Japan, the Republic of Korea and the United States (Mudd et al., 2018; D. Huston, Geoscience Australia, pers. comm. March 2022). Australia is endowed with significant resources of many of the critical elements and the critical minerals hosting them, including Li. Currently, Australia's ranking for economic resource of Li~~thium~~ was the second, but it ranked the first for its production (Senior, 2022), with potential of additional discoveries. According to recent survey (Senior, 2022), Australia produced 40 kilotons (kt) of Li (in terms of spodumene, $Li_2O.Al_2O_3.4SiO_2$, concentrates; assuming 6% of Li_2O in spodumene concentrates) in 2020, or 49% of the global production; a significant increase from 21.3 kt of Li in 2017 (Champion, 2019).

The two primary sources for Li are brine stores and mineral deposits, where Li is hosted mainly spodumene ($LiAlSi_2O_6$). A 2013 investigation by Geoscience Australia found that the potential of Li-rich salt lakes in Australia was relatively low in comparison to those, for instance, in the Americas (Jaireth et al., 2013; Mernagh et al., 2015; Mernagh et al., 2016). Most of the Li in Australia exists as mineral deposits (Champion, 2019). Despite Australia's current position as the world's leading supplier of Li, it has limited prospects for immediate expansion as the potential for spodumene similar deposits in Australia has not yet been fully investigated (Mudd et al., 2018). This study aims to contribute to filling this knowledge gap by providing the first digital map of Li content of Australian soils.

Lithium values ranges from <1 – 15 mg kg^{-1} in ultramafic rocks, 5.5 – 17 mg kg^{-1} in mafic rocks, while felsic rocks (granite, rhyolite and phonolite) contain higher Li concentrations, between 30 – 70 mg kg^{-1} (Foregs, 2006). Lithium concentration in clay minerals ranges between 7 – 6000 mg kg^{-1} (Starkey, 1982). With developments in technology, a process of extracting Li as Li-carbonate from certain minerals, other than spodumene, such as lepidolite ($KLi_2Al(Si_4O_{10})(F,OH)_2$) and petalite ($LiAlSi_4O_{10}$), has been identified (Sitando and Crouse, 2012; Vieceli et al., 2018). Lower Li concentration is found in salt lake brines (0.17 – 1.5 mg kg^{-1}) (Grosjean et al., 2012). Extraction of Li from salt lake brine is in the form of Li-chloride, which needs to undergo an energy-intensive process to be converted to Li-carbonate for use in batteries.

Lithium is found in trace amounts in all soil types, primarily in the clay fraction, with slightly smaller concentrations in the organic soil fraction (Kabata-Pendias, 2010). Possible means by which Li is bound to clay was explained elsewhere (Starkey, 1982). A typical background concentration of Li in the soil ranges from 7 – 200 mg kg^{-1} (Schrauzer, 2002). In New Zealand, a study of Li concentration in soil reported a range between 0.08 – 92 mg kg^{-1} (Robinson et al., 2018). While in southwestern Siberia, Gopp et al. (2018) reported soil-available Li content derived from ammonium acetate-buffered solutions ranged

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

from 0.24 – 0.68 mg kg^{-1}. The amount of soil-available Li is usually relatively low, about 3 – 5% of the total Li content in the surface layers (Gopp et al., 2018; Anderson et al., 1988). De Caritat and Reimann (2012) reported median Li concentrations (after aqua regia digestion) of 12 and 5.7 mg kg^{-1} in European agricultural topsoils and Australian surface sediments, respectively, both in the coarse (< 2 mm) fraction. Subsequently, Reimann and De Caritat (2017, Fig.2SM) published the first continental map of Li in Australian soils, based on National Geochemical Survey of Australia (NGSA) data, showing that regions of high and low concentrations are found across all states.

Higher concentrations of Li are often found in the deeper layers of soil profiles (Merian and Clarkson, 1991) because, typically, Li enters the soil column through the weathering of sedimentary minerals in the underlying saprolite and bedrock (Aral and Vecchio-Sadus, 2008). Because clay minerals predominantly drive the mineralisation and dissolution of Li, the clay mineral fraction will play a significant role in determining the Li concentration. The Li content of soil is controlled more by the soil formation conditions than by the composition of the parent materials (Kabata-Pendias, 2010).

Mineral exploration aims to find ore deposits for mining purposes. Therefore, delineating target areas for mineral exploration through a series of mapping activities is a crucial initial stage leading to discovery (Carranza, 2011). Mineral prospectivity mapping is a method to quantify the probability of mineralisation in a selected area for mineral exploration purposes. This prioritisation allows for the exploration of smaller, higher-potential areas for detailed prospecting to minimise exploration costs, e.g., the number of drillholes.

Two common paradigms for creating mineral prospectivity maps are knowledge-driven and data-driven models (Carranza, 2011). Knowledge-driven models do not require any data on mineral deposits, but rely on expert knowledge of spatial associations between mineral deposits and geological features, field experience and conceptual models to develop evidential maps that enables the discovery of mineral deposit (Carranza, 2008). Meanwhile, data-driven models utilise existing knowledge on the location of mineral occurrences, various survey data and spatial statistical methods to represent the likelihood of mineral occurrence within prospective areas (Carranza, 2008). With the development of machine learning and technology (computer hardware, software and geographic information system (GIS) technology), there have been growing applications of mineral prospectivity mapping in the recent decades (Carranza, 2011; Zuo, 2020).

Several studies have demonstrated the use of remote sensing to explore various deposit types, such as gold deposits (Cŕosta et al., 2003), copper deposits (Pour and Hashim, 2015) and iron ores (Ducart et al., 2016). Recently, the application of remote sensing for Li deposits has also emerged. Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) images were used to map Li content in the Vale do Jequitinhonha region of Brazil (Perrotta et al., 2005). Gopp et al. (2018) explored the use of Normalised Difference Vegetation Index (NDVI) to develop a predicted map of the plant available content of Li in southwestern Siberia soil. Cardoso-Fernandes et al. (2018) evaluated the potential use of Sentinel-2 in Li mapping in the Fregeneda-Almendra region across the Spain-Portugal border. Similarly, Köhler et al. (2021) further explored the use of combined geological data and remote sensing data for Li potential mapping.

In soil science, digital soil mapping (DSM) has been widely used to produce quantitative maps of soil attributes based on the known distributions of environmental covariates (i.e. rainfall, parent material, vegetation and landforms), that affect soil

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023


Open Access Earth System Science Data Discussions

formation. The DSM framework is derived from the conceptual model developed by Mcbratney et al. (2003) in which a certain soil attribute results from the interaction of soil-forming factors. These factors are modified from Jenny (1941) and include soil (*s*), climate (*c*), organisms (*o*), relief (*r*), parent material (*p*), age/time (*a*) and spatial position (*n*), or *scorpan*. The factors are measured or approximated from various data types, including point observations, maps (polygons), existing data, and remote sensing data and derivatives thereof (e.g., gradients, buffer distances, etc.); these can be numerical or categorical data types.

In this study, we attempt to model Li distribution in the surface and subsurface soils of Australia by invoking the NGSA soil geochemistry dataset and various environmental covariates commonly used in DSM related to soil formation in Australia. In detail, the objectives of this study are thus to:

(i) evaluate the use of ~~digital soil mapping~~ framework to predict Li concentration in Australian soils, and

(ii) delineate anomalous areas potentially attractive for Li exploration and discuss their interpretations.

2 Materials and methods

2.1 Li measurement

This study used two soil datasets, referred to as the calibration and validation datasets. The calibration dataset was used to build the spatial prediction model and the validation dataset was used to test the prediction quality of the calibrated model.

The calibration dataset data were generated as part of the NGSA project (www.ga.gov.au/ngsa), a collaborative project between Geoscience Australia and the States/NT between 2007 – 2011, which aimed to document the soil geochemical concentration levels and patterns across Australia. Details on the project, analysis, sampling methods and the measurement of other parameters can be found in De Caritat and Cooper (2011a) and De Caritat and Cooper (2016).

The NGSA collected samples at 1315 sites (including field duplicates) at or near the outlet of large catchments with a total area coverage of 6.17 million km^2 and an average sampling density of 1 site for every 5200 km^2 (De Caritat and Cooper, 2011a). The target sampling medium was floodplain sediments away from river channels, though in various places in Australia, an aeolian input can be important; thus, the medium was called ‘catchment outlet sediment’ rather than floodplain sediment. These geomorphological entities are typically vegetated and biologically active (plants, worms, ants, etc.), thereby making the collected materials true soils (e.g., Sssa, 2022), albeit soils developed on transported alluvium. Due to limitations in access, samples from some parts of South Australia and Western Australia could not be obtained.

Samples were collected from two depths, namely ‘top outlet sediment’ (TOS) from 0 – 10 cm depth, and ‘bottom outlet sediment’ (BOS) from, on average, 60–80 cm depth. All of the soil samples were air-dried, homogenised and dry sieved to <2 mm and <75 µm prior to various analyses for 60+ elements (see De Caritat et al. (2009) and De Caritat et al. (2010), for a full description of the NGSA sample preparation and analytical methods, respectively). A detailed quality assessment of the NGSA data is given in De Caritat and Cooper (2011b); for Li after aqua regia digestion, analytical precision (repeat analysis of certified reference material Till-1) of 12% and overall precision (based on field duplicates) of 39% were reported, whilst

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

accuracy could not be determined for lack of certified aqua regia Li data for Till-1. In this contribution, we use Li concentration after aqua regia digestion data for the NGSA <2 mm TOS and BOS samples analysed by inductively coupled plasma mass-spectrometry (ICP-MS) in a commercial laboratory. Any Li measurements that fell below the detection limit (0.1 mg kg^{-1}) were replaced with half the detection limit (0.05 mg kg^{-1}). The distribution of sampling sites and the concentration levels of Li are shown in Figure 1.

As an independent validation dataset, we used the geochemical dataset from the Northern Australia Geochemical Survey (NAGS) project (Main et al., 2019). This dataset contains 773 observations located in the Tennant Creek – Mt Isa region in the Northern Territory and Queensland, with an approximate sampling density of one sample every 500 km^2 and collected in 2017. The distribution of these samples is also shown in Figure 1. These samples were collected, prepared and analysed following the NGSA protocols (De Caritat and Cooper, 2011a), albeit at a higher sampling density. However, only TOS samples were collected in NAGS. Furthermore, these samples were collected at different times and /or laboratories. To address the analytical variation that could potentially arise, levelling method were utilized using the standards Certified Reference Materials (Main and Champion, 2022). In short, a correction factor based on the CRM measurements from the two datasets is calculated and applied as multiplier to relevel the data.

[Figure 1]

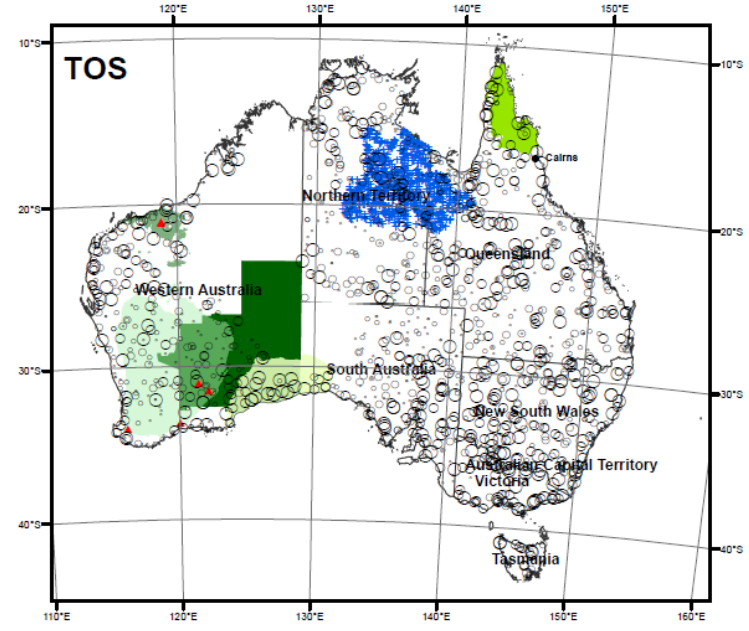
2.2 Environmental covariates

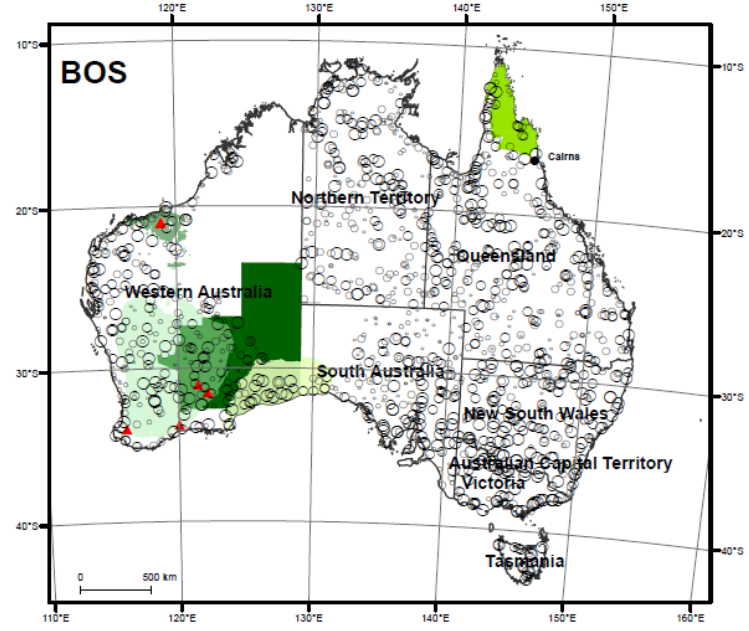
A total of 19 environmental covariates (Table 1) characterising factors of climate, parent material, soil, and topography, that contributes to soil formation were considered in this study.

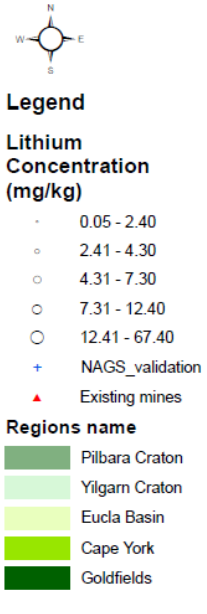
The first factor is climate. Water and temperature affect the rate of mineral weathering and thus soil formation. Hence, we included precipitation, evaporation and temperature data (Harwood, 2019), along with the topographic wetness index (TWI) data (Gallant and Austin, 2012a), informing the relative wetness within a landscape. In short, the TWI was derived from contributing area product, which was computed from Hydrologically enforced Digital Elevation Model, and from the percent slope product, which was computed from the Smoothed Digital Elevation Model (Gallant and Austin, 2012a).

The second factor is parent material (i.e. degree of weathering and mineralogical composition), including gamma-ray radiometric and total magnetic intensity. Gamma-ray radiometric surveys provide estimates for the concentrations of gamma-ray-emitting radioelements K, U and Th at/near the soil surface. The gamma-ray radiometric data was measured from airborne surveys throughout most of Australia (Poudjom Djomani et al. (2019). In this study we used a complete gamma-ray survey grid where gaps in the airborne coverage were filled in using covariate machine learning (Wilford and Kroll, 2020). Gamma-ray radiometric data have been found to be a useful covariate in identifying surface processes such as sediment transport and weathering (Wilford, 2012; Wilford et al., 1997) and detecting radioactive minerals deposits and occurrences (Alhumimidi et al., 2021; Wilford et al., 2009; Dickson et al., 1996; Dickson and Scott, 1997). Total magnetic intensity (TMI), which measures variations in the Earth's magnetic field intensity caused by the contrasting content of

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023







https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

Figure 1. Distribution of sampling sites from the National Geochemical Survey of Australia (NGSA, black circles) for both depths: top outlet sediment (TOS) 0–10 cm, top; and bottom outlet sediment (BOS) ~60–80 cm, bottom. Distribution of sampling sites from the Northern Australia Geochemical Survey (NAGS, blue plus signs) for TOS only, top. All data refer to the coarse fractions (<2mm). Aqua regia-soluble Li concentrations (mg kg^{-1}) are categorised in five quantile classes. Regions discussed in the text are highlighted in various shades of green. Projection: Australian Albers Equal Area (EPSG:3577). Data sources: De Caritat and Cooper (2011a), Hughes (2020).

The third factor is the soil itself, particularly, the relevant physical soil properties. Previous studies, e.g. by Kabata-Pendias (1995) and Robinson et al. (2018), highlighted the high correlation between Li and clay content of soil, soil texture was used as a covariate. The soil texture spatial information (sand and clay contents) was derived from Malone and Searle (2021), which contained updated information on soil texture map across Australia derived using a digital soil mapping approach. The sand and clay fractions were developed by integrating field morphological (n = 180,498) and laboratory measurements soil fractions (n = 17,367) from the Soil and Landscape Grid of Australia (SLGA). The SLGA is based on a comprehensive dataset of soil attributes across Australia including the NGSA dataset. These sand and clay content of Malone and Searle (2021) were for specific depth intervals of 0–5 cm, 5–15 cm, 15–30 cm, 30–60 cm, 60–100 cm, and 100–200 cm. They were converted to the depths corresponding to the NGSA Li measurement (0–10 cm and 60–80 cm) using the mass-preserving spline function, described in Bishop et al. (1999) and modified by Malone et al. (2009). Soil reflectance in the visible, near-infrared (NIR), and short-wave-infrared (SWIR) spectra captured by remote sensing images provides information on soil composition. However, the unprocessed images consist of a mixture of soil, bedrock, vegetation and clouds. By removing the influence of vegetation, Roberts et al. (2019) were able to document the 'barest' state of soil, so critical in mapping the characteristics of soil and rock. This was done by combining Landsat 5, 7, and 8 observations of the past 30 years to remove the contamination by vegetation, cloud cover, shadows, detector saturation and pixel saturation. The model used to develop Barest Earth was validated using the NGSA spectral archive (Lau et al., 2016).

Finally, topography is represented by elevation and slope. These factors also play an important role, as they affect how water is added to and/or lost from soil. The elevation was derived from the smoothed ~~Digital Elevation Model (~~DEM-S~~)~~ which was obtained from the 1 arc-second resolution Shuttle Radar Topography Mission (SRTM) data acquired by NASA in February 2000 (Gallant, 2011). The slope covariate was also calculated from DEM-S using the finite difference method (Wilson and Gallant, 2000). The different spacing in the E-W and N-S directions due to the geographic projection of the data was accounted for by using the actual spacing in metres of the grid points calculated from the latitude.

All covariates were reprojected to EPSG:3577 (GDA94 datum; Australian Albers equal area projection) and resampled to a common spatial resolution of 3 km prior to any analysis. All the environmental covariates used are shown in Table 1.

[Table 1**]**

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

Open Access Earth System Science Data Discussions

Table 1. Environmental covariates used for digital soil mapping of Li.

Covariate	Description	Source	Original resolution
PTA	Annual precipitation (mm)	Harwood (2019)	90 m
EPA	Annual potential evaporation (mm)	Harwood (2019)	90 m
TRA	Annual temperature range (°C)	Harwood (2019)	90 m
Dose	Radiometrics: filtered dose (nGy/h)	Wilford and Kroll (2020)	0.001 degree
K	Radiometrics: filtered K element concentrations (%)	Wilford and Kroll (2020)	0.001 degree
Th	Radiometrics: filtered Th element concentrations (ppm)	Wilford and Kroll (2020)	0.001 degree
Th/K	Radiometrics: derived Th to K ratio (ppm/%)	Wilford and Kroll (2020)	0.001 degree
TMI	Total magnetic intensity (nT/m)	Poudjom Djomani et al. (2019)	90m
Sand	Sand content (%)	Malone and Searle (2021)	90 m
Clay	Clay content (%)	Malone and Searle (2021)	90 m
Landsat band 1*	Blue (450–510 nm)	Wilford and Roberts (2019)	25 m
Landsat band 2*	Green (530–590 nm)	Wilford and Roberts (2019)	25 m
Landsat band 3*	Red (640–670 nm)	Wilford and Roberts (2019)	25 m
Landsat band 4*	Near infrared NIR (850–880 nm)	Wilford and Roberts (2019)	25 m
Landsat band 5*	Shortwave infrared SWIR1 (1570–1650 nm)	Wilford and Roberts (2019)	25 m
Landsat band 6*	Shortwave infrared SWIR2 (2110–2290 nm)	Wilford and Roberts (2019)	25 m
Elevation	3 Second DEM - Shuttle Radar Topography Mission (m asl)	Gallant (2011)	1 arc-second
Slope	Elevation gradient (%)	Gallant and Austin (2012b)	90 m
TWI	Topographic wetness index (dimensionless)	Gallant and Austin (2012a)	30 m

*All Landsat bands referred here are from the Landsat 30 + Barest Earth products

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

2.3 Modelling

Here we used a machine learning model Cubist to relate soil observations to the environmental covariates. Cubist is a tree-based regression algorithm based on the M5 theory (Quinlan, 1993). This algorithm creates partitions of data with similar spectral characteristics and creates one or more rules for each partition. If the partition rules are satisfied, then the linear regression of that partition is used to create the prediction (Eq. 1). Each rule can be defined as:

If [condition is true], then [regression], else [apply next rule] (Eq. 1)

The Cubist model has two tuning parameters: committees (number of sequential models included in the ensemble) and neighbours (number of training instances that are used to adjust the model-based prediction). A full combination of committees (5,10, 20, 30, 40, 50) and neighbours (0, 1, 5, 9) were tested to tune the Cubist model. To obtain the best estimates of optimum parameters, a 10-fold cross-validation approach was utilised. Based on the optimum parameters, 50 bootstrap models ('sampling with replacement') were trained.

The performances of the prediction models were then evaluated on both internal evaluation and on the independent validation dataset. An internal evaluation of the model was conducted using "out of bag" samples, which were not used during the development of the bootstrap models. The NAGS dataset was used to evaluate the performance on the independent dataset (top depth only). The following metrics were used: adjusted coefficient of determination (R^2_{adj}), Lin's concordance correlation coefficient (LCCC), root mean square error (RMSE), bias, and ratio of performance to interquartile distance (RPIQ). R^2_{adj} is a measure of the linear association between observed and predicted values; LCCC measures the agreement between the observed and predicted values in relation to the 1:1 line; RMSE is a measure of the differences between the observed and predicted values; bias is the measure of the difference between the mean of the observed and the mean of the predicted values; and RPIQ is a measure of performance that takes into account the distribution of the values, and can be calculated as a fraction of the interquartile range of the observed values (Q_3–Q_1) and the RMSE (RPIQ = (Q_3–Q_1)/RMSE) (Bellon-Maurel et al., 2010).

Variable importance analysis was also conducted to evaluate the contributions of each covariate in the Li prediction. The relative variable importance is measured as the percentage of times the environmental covariate is either used as a condition or a rule within the Cubist model. These bootstrap models were then used to generate output maps with the same extent and resolution. The final map output was derived based on the mean prediction of the bootstrap models; similarly, the standard deviation was obtained based on the standard deviation of the prediction from the bootstrap models.

2.4 Data processing and statistical computing

All the data analytics, modelling, and mapping procedures in this study were conducted in R statistical open-source software (R Core Team, 2021). Besides the base R functionality, the R packages used in this study included "Cubist" (Kuhn and Quinlan, 2021) for fitting cubist models; "caret" (Kuhn, 2021) for tuning the hyperparameter of the Cubist model; and

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

"raster" (Hijmans, 2021) for handling raster layers and generating soil map predictions. All soil maps were produced in ArcGIS version 10.8 (ESRI 2019) using the Albers equal area projection (EPSG:3577).

3 Results and discussion

3.1 Descriptive analysis

The distribution of Li concentrations (NGSA dataset, De Caritat and Cooper, 2011a) was positively skewed (Figure 2) with concentrations ranging from 0.1 – 67.4 and 0.1 – 56 mg kg^{-1}, for TOS and BOS respectively. Only limited observations above 20 mg kg^{-1} of Li concentrations were found in this study for both TOS (n = 76) and BOS (n = 95). The median concentration of TOS (5.7 mg kg^{-1}) was slightly lower than that of BOS (7.0 mg kg^{-1}). These concentrations were lower than those observed in Négrel et al. (2019) for mean Li concentration in European soil at 11.3 mg kg^{-1}, and across the background concentrations of Li in the world (7 – 200 mg kg^{-1}), according to Schrauzer (2002). The Li concentration at TOS was strongly correlated with BOS (r = 0.75, p <0.0001).

[Figure 2]

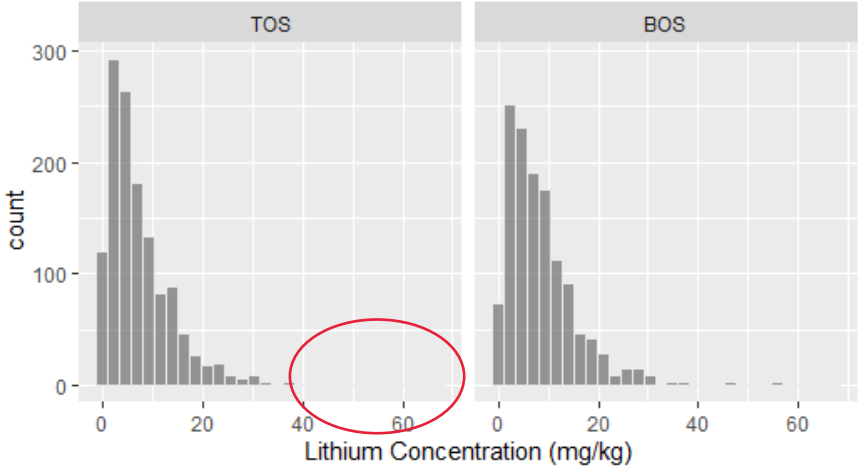


Figure 2. Histograms of Li concentrations for both NGSA depths: top outlet sediment (TOS) 0–10 cm, left; and bottom outlet sediment (BOS) ~60–80 cm, right. Data source: De Caritat and Cooper (2011a).

Based on the data collected by the NGSA project, the highest concentration of Li for both TOS and BOS was found in northernmost Queensland (Cape York Peninsula), as shown in Figure 1. Other regions that have significant quantities of Li were located in the Goldfields-Esperance region in Western Australia, which has been recognised as one of the most resource-rich areas on the planet (Champion, 2019), and the region around the Victoria-New South Wales border (Figure 1). Some of the findings correlate well with the existing mine sites in Australia (red triangles in Figure 1). The largest deposit of Li found in Australia is the Greenbushes deposit, south of Perth. Other regions include Mount Marion and Earl Grey in the

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

Yilgarn Craton, and Pilgangoora in the Pilbara Craton (Champion, 2019). In July 2019, Strategic Metals Australia (SMA) found a new Li exploration target near Cairns, in the Georgetown province of north Queensland (Gluyas, 2019). However, this discovery has not been updated in the data collected by Geoscience Australia because considerable work such as drilling, modelling, resource calculation and feasibility studies are needed to bring the discovery to the feasibility stage.

3.1.1 Correlation between Li with other measured soil properties

Despite other studies (Robinson et al., 2018; Kashin, 2019) reporting strong correlations between Li and Mg, and other elements elsewhere, including Al, B, Fe, K, Mn and Zn, the NGSA data only show strong correlations between Li and Al (Pearson's correlation coefficient r = 0.74), Ga (r = 0.69), Cs (r = 0.68), and Rb (r = 0.66) for TOS, and slightly lower correlation for BOS: Al (r= 0.69), Ga (r = 0.64), Cs (r = 0.62), Rb (r = 0.61). Correlations between Li with K and Mg were only moderate for both TOS (r = 0.48 and 0.43) and BOS (r = 0.46 and 0.33). Similarly, Foregs (2006) also observed good correlations (r > 0.4) of Li with Al, Ga and Rb within the floodplain sediment samples.

The Li concentration in soil was negatively correlated with measured sand content from the NGSA dataset (r = -0.55), and positively correlated with clay content (r = 0.44). This is consistent with the findings of Kabata-Pendias (2010) and Robinson et al. (2018), who noted the tendency of clay minerals to concentrate Li. It has been suggested that Li may be located internally within clay minerals, mainly kaolinite, illites, smectites including hectorite, palygorskite and sepiolites, in ditrigonal cavities via isomorphous substitution, rather than on exchange sites (Anderson et al., 1988; Starkey, 1982).

3.1.2 Correlation with environmental covariates

Overall, the correlation of Li concentration with the environmental covariates was relatively low (Figure 3). The correlation with sand and clay content derived from digital soil maps was lower in comparison to the measured (NGSA) values discussed above, with r = -0.28 and 0.25, respectively, for TOS; and r = -0.23 and 0.22, respectively, for BOS.

For TOS, the Landsat bands 3 (Red), 5 (SWIR1) and 6 (SWIR2) had similar correlations with Li content (r = -0.15 to -0.17). For gamma-ray radiometric data, total dose and K content had correlations with Li of r = 0.10 to 0.14. These positive correlations are expected as the associations of Li deposits and felsic rocks (high in both total dose and K) due to the observed incompatibility in mineral structures (Benson et al., 2017). Precipitation has positive correlation (r = 0.12), while both temperature and elevation had negative correlations (r = -0.12) with Li content. Topographic variables such as slope had negligible correlation with Li content (r = 0.05).

For BOS, similar observations on the correlations between Li content and environmental covariates were found, except for the following differences. Landsat bands 3, 5 and 6 had stronger negative correlations (r = -0.14 to -0.16) compared to other bands. Potassium (r = 0.10) and dose (r = 0.09) had higher correlation with Li compared to the Th/K ratio (r = -0.02). Both temperature (r = -0.11) and elevation (r = -0.08) had negative correlations with Li content, while slope (r = 0.05) and precipitation (r = 0.08) had low positive correlations.

[Figure 3]

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

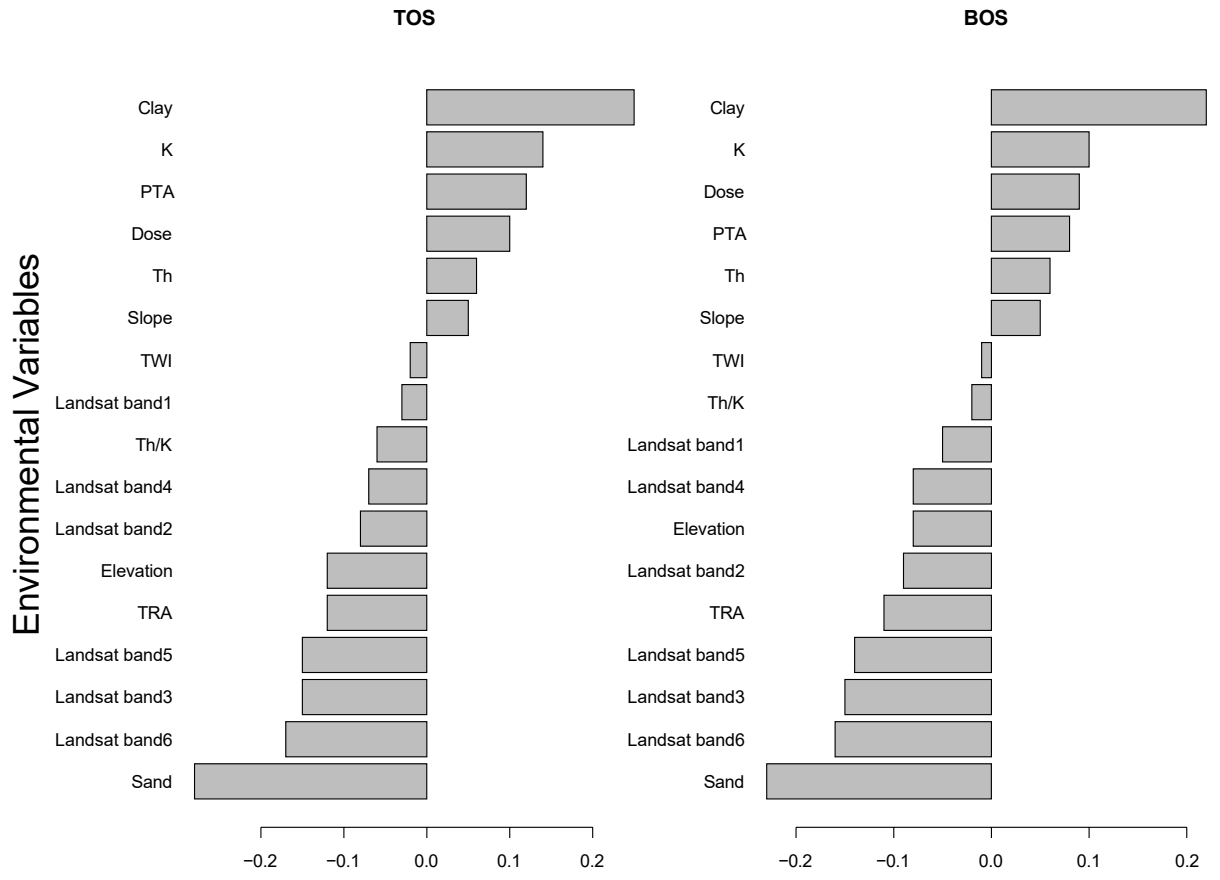


Figure 3. Pearson's correlation coefficient (r) between Li content and the environmental covariates (*scorpan*) for both NGSA depths: top outlet sediment (TOS) 0–10 cm, left; and bottom outlet sediment (BOS) ~60–80 cm, right. Data sources: de Caritat and Cooper, 2011b; Gallant, 2012a; Harwood, 2019; Wilford, 2019; Wilford and Kroll, 2020; Malone and Searle, 2021. See Table 1 for abbreviations.

3.2 Model evaluation

The final Cubist model is tuned with committees of 20 and neighbours of 9, which resulted in the lowest value of RMSE compared to the other combinations of hyperparameters, indicating an optimised Cubist model.

3.2.1 Internal evaluation

Validation statistics based on internal evaluation using the out-of-bag data for the Li predictions are presented in Table 2. Higher accuracy was observed in TOS (R^2_{adj} = 0.20; LCCC = 0.36) compared to BOS (R^2_{adj} = 0.12; LCCC = 0.29). There was a slightly lower RMSE on the prediction for TOS (RMSE = 6.29 mg kg^{-1}) compared to BOS (RMSE = 7.28 mg kg^{-1}).

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

This is expected as most of the environmental covariates reflected soil surface conditions. To the best of our knowledge, the machine learning models developed in most mineral exploration studies were assessed based on classification accuracy, (i.e. presence or absence of specific minerals in sample) instead of regression accuracy (Jooshaki et al., 2021). In addition, remote sensing studies on mapping Li minerals are rarely validated (e.g. Cardoso-Fernandes et al. (2019)). Hence, no comparison can be made with other studies.

[Table 2]

Table 2. Model evaluation and validation results for the prediction of Li concentrations using Cubist model for both NGSA depths: top outlet sediment (TOS) 0–10 cm; and bottom outlet sediment (BOS) ~60–80 cm. The independent validation is based on comparing predictions to the NAGS dataset Li concentrations.

Depth	R^2_{adj}	LCCC	RMSE	bias	RPIQ
TOS (0–10 cm)	0.20	0.36	6.29	-0.80	1.20
BOS (~60–80 cm)	0.12	0.29	7.28	-0.76	1.14
Independent Validation: TOS (0-10cm)	0.32	0.44	3.67	2.27	1.08

3.2.2 Independent validation dataset

The predictive model performance was also evaluated using an independent dataset (NAGS, TOS only) that was not part of the calibration dataset. Upon relevelling, the median Li concentration of this validation dataset (3.46 mg kg^{-1}) was lower than that observed in the calibration dataset (5.7 mg kg^{-1}), with a range of values between 0.1 to 22.5 mg kg^{-1}. A comparison of the subset of NGSA within the extent of the NAGS dataset also showed similar result, with slightly higher concentrations observed within the local NGSA dataset, which ranges between 0.1 and 28.7 mg kg^{-1} and has a median of 4.1 mg kg^{-1}. However, the samples from the two datasets were deemed to have similar distribution with the two-sample Kolmogorov-Smirnov test (D = 0.18, p-value = 0.012).

We reported the performance of model validation the same way the model evaluation was conducted (Table 2; Figure 4). The model validation resulted in higher accuracy (R^2 = 0.44; LCCC = 0.59). The RMSE was also slightly lower than those observed in the TOS model evaluation, most likely due to lower observation values within the NAGS validation dataset. The model overestimated the concentration with a mean error of 1.46 mg kg^{-1}.

[Figure 4]

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

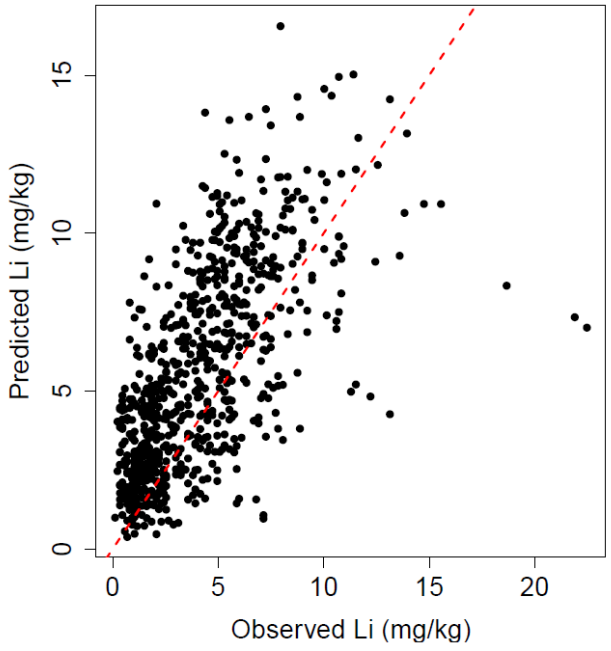


Figure 4. Goodness-of-fit plot showing observed vs predicted Li concentrations based on the independent validation dataset (NAGS, TOS only). Red dashed line is the 1:1 line.

3.3 Variable importance analysis

From the Cubist model, we can infer the relative importance of the covariates by calculating the percentage of times a covariate is being used in the model. The variables used by Cubist model can be further split in terms of "importance in the conditions" and "frequency of predictor usage in models".

For Li prediction in TOS, the variables TWI and Landsat bands 2, 3 and 4 are of higher importance in the conditions than other variables (Figure 5). This implies that the model separates out prediction values based on its spectral response of vegetation and Fe-bearing minerals related to Landsat bands 2 to 4 and the wetness index. However, within the regression models, the top five variables most frequently used in the regression were the Landsat band 2, band 6, band 1, band 3 and gamma-ray radiometric total dose. The first three Landsat bands (red, green, and blue) and band 6 (SWIR2) have been commonly used to predict soil properties and delineate geological boundaries, as well as discriminate and differentiate vegetation zones (Khorram et al., 2012). While the gamma radiometric dose discriminated the various soil types and their mineral makeup. The next set of covariates was annual precipitation and clay and sand content which bound the Li in the soil, indicating they have lower importance as predictors. As indicated in the correlation analysis, slope was not significant.

For the BOS model, Th/K had the highest importance in the conditions of the model (Figure 5), separating high and low values, but it does not affect the regression. Landsat bands 2 and 6, and temperature range also affect model conditions. Overall, parameters that influenced the BOS regression model were similar to those for TOS, i.e. the top-five are Landsat

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

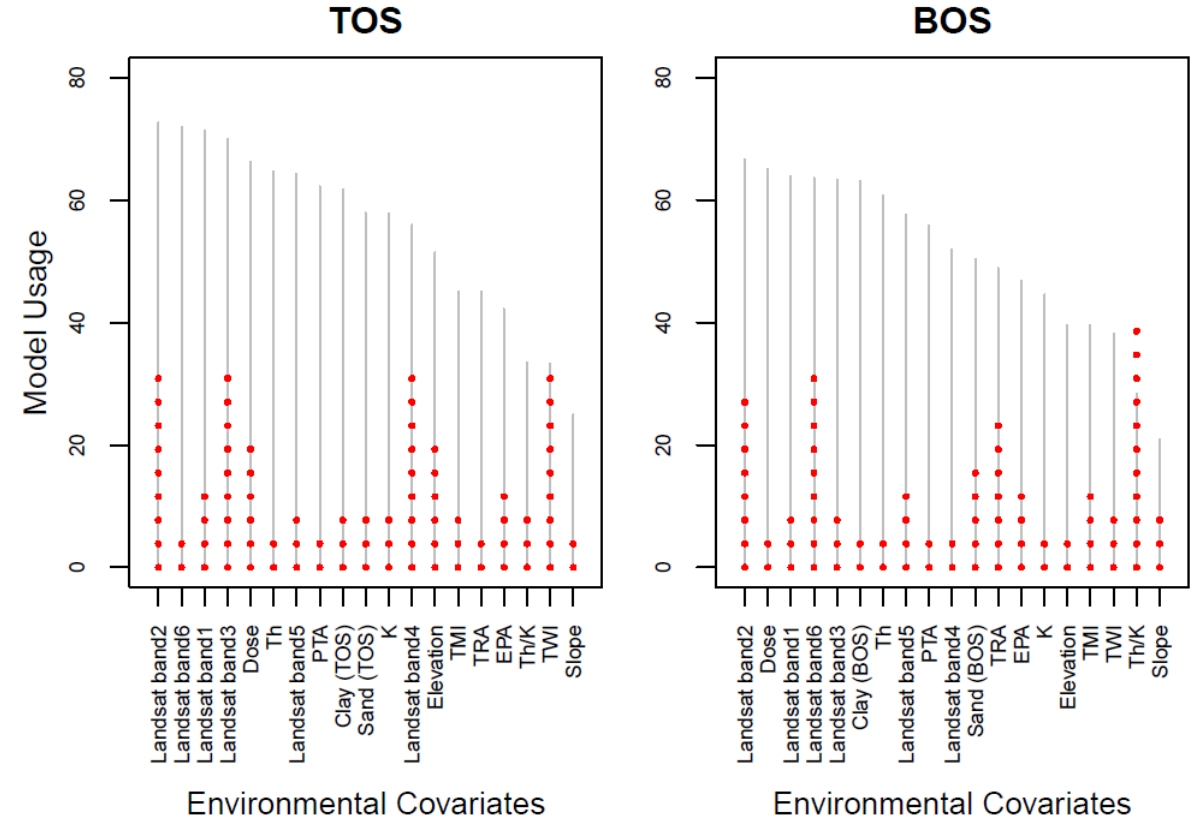


Figure 5. Variable importance of covariates in terms of importance as conditions (red dotted lines) and frequency of predictor usage (grey lines) by the Cubist algorithm for both NGSA depths: top outlet sediment (TOS) 0–10 cm, left; and bottom outlet sediment (BOS) ~60–80 cm, right. Covariates are sorted in order of decreasing frequency of usage.

3.4 Li prediction maps

The Cubist model led to the generation of spatial predictions of Li concentration in fluvial sediment-derived soils across Australia at two depths (Figure 6). So far, there are only five known Li mines in Australia (mostly in Western Australia), all of which are located within areas that were predicted by the model developed here to have a higher concentration of soil Li, especially for the BOS model (>~8 mg kg^{-1}) (Figure 7).

[Figure 6] [Figure 7]

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

360

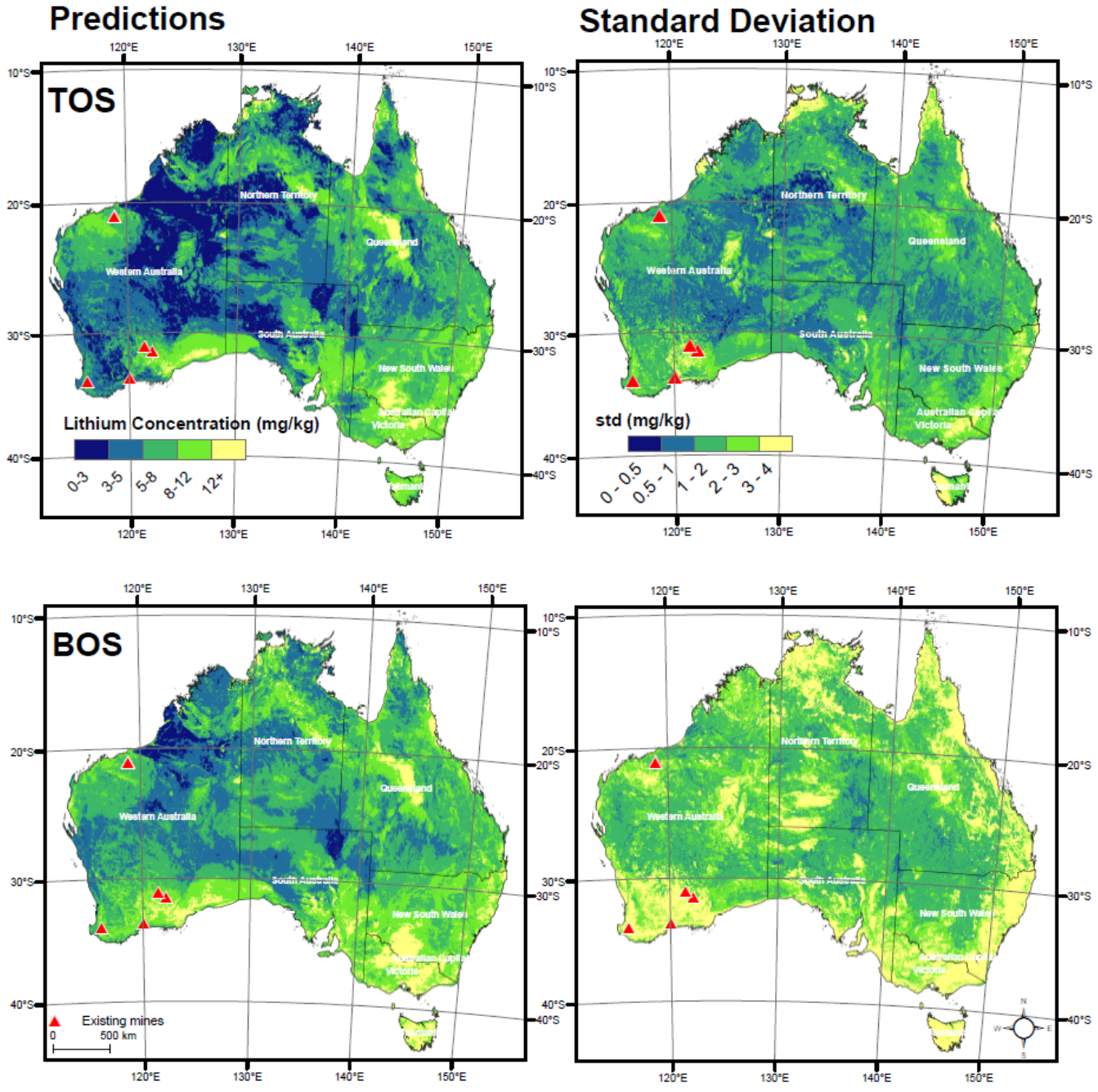


Figure 6. Spatial distributions of predicted aqua regia-soluble Li concentrations (mg kg^{-1}) in coarse fraction (<2 mm) alluvial soils across Australia, left, and standard deviations (mg kg^{-1}), right, for both National Geochemical Survey of Australia (NGSA) depths: top outlet sediment (TOS) 0–10 cm, top; and bottom outlet sediment (BOS) ~60–80 cm, bottom.

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

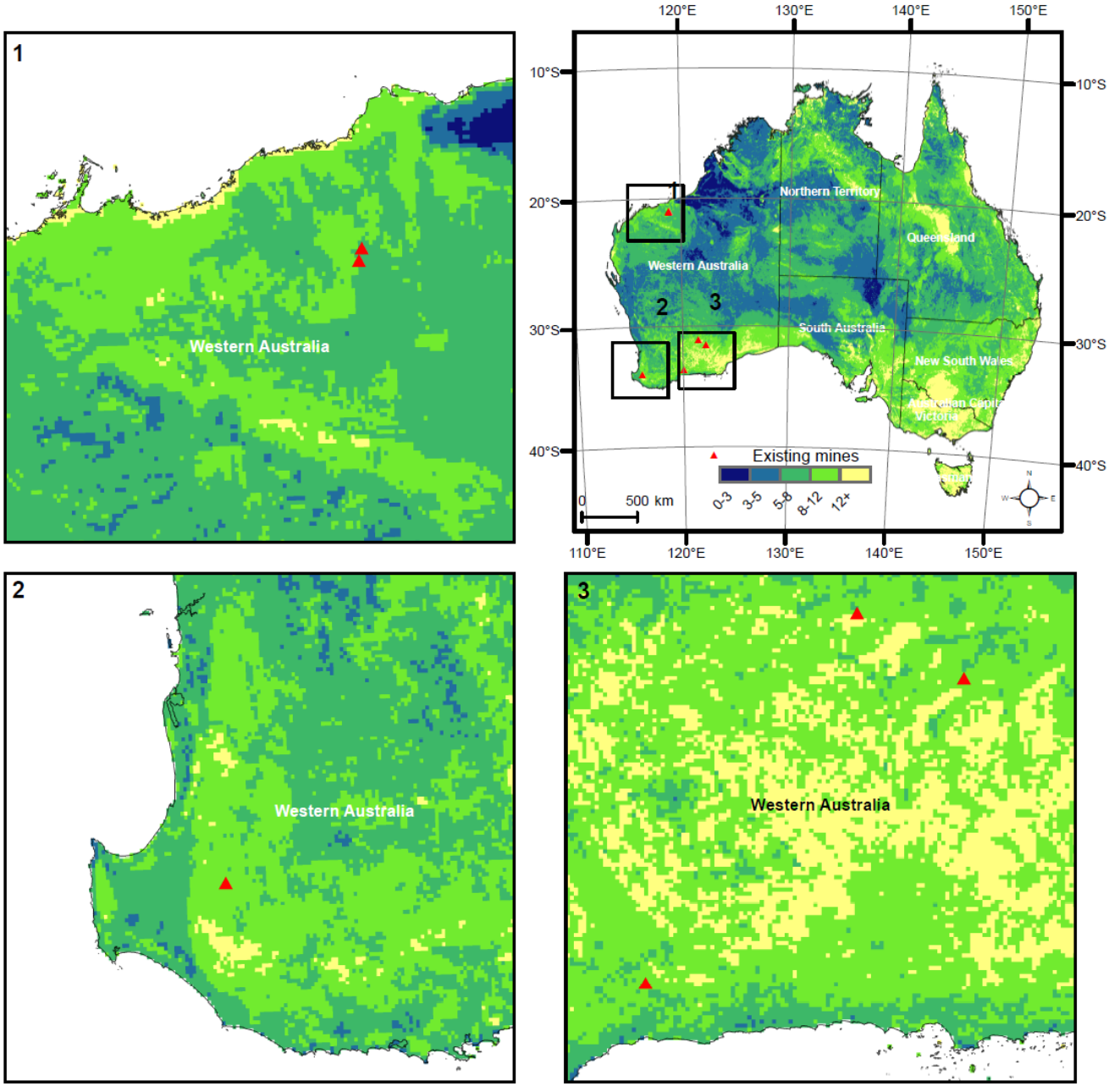


Figure 7. Distribution of Li mines in the digital soil map of Li in Australia for bottom outlet sediment (BOS) ~60–80 cm depth.

In Australia, the largest producer of spodumene is the Greenbushes Li operation, approximately 250 km south-southeast of Perth. In the most recent public report, the company reported combined measured and indicated resources of 118.4 million tons (Mt) of 2.4% Li_2O containing proved and probable reserves of 61.5 Mt grading 2.8% Li_2O (Champion, 2019). Other locations explored for Li include Mount Cattlin and Mount Marion in the Goldfields-Esperance region, and Pilgangoora of

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

East Pilbara. In a recent review (Champion, 2019), these projects' report estimated Li resources ranging from 11.8 to 71.3 Mt at 1.01 to 1.37% Li_2O.

The predicted soil Li concentrations at the known Li mine sites range from 4.5 to 7.3 mg kg^{-1} for TOS and from 7.1 to 12.6 mg kg^{-1} for BOS. The highest TOS and BOS concentrations of Li proximal to a known mine site are for the Mount Marion deposit in Western Australia.

Although most Li exploration to date has been conducted in Western Australia, our map indicates that other regions in Australia are anomalous in Li (Figure 6). These areas are located for instance within the central west region of Queensland and visually correspond to areas of widespread black cracking (smectite-rich) soils, or vertosols (Isbell and Ncst, 2021). Elevated concentration of Li was also observed over parts of Eucla Basin, which has a widespread distribution of iron-oxide rich regolith with carbonate accumulations (Johnson, 2015). The sources of carbonate include weathered Proterozoic and Palaeozoic carbonate bedrock, vast marine sediments that extend across the low-lying and offshore areas associated with Cenozoic sedimentary basins and abundant widespread pedogenic carbonates (Johnson, 2015). This is in line with Foregs (2006) observations, where higher Li concentration of up to 56 mg kg^{-1} were identified in calcareous soil (high carbonates accumulation) in comparison to those of organic soil (1.3 mg kg^{-1}). The Fe in iron oxides and oxyhydroxides that help retaining Li may be released from oxidation of primary minerals during weathering (Kabata-Pendias, 2010). The ultimate origin of Li within these clay-, iron- and carbonate-rich soils remains to be established in the case of Australia. Other regions of potential interest occurring on different soil types are located in southern New South Wales and parts of Victoria.

We further explored the correlation of Li concentration against soil orders (Searle, 2021). Figure 8 shows the range of Li concentration across various soil types identified within the sampling locations. The Li concentration tended to be slightly higher on Vertosols, Calcarasols, as well as Dermosols. These observations indicate Li accumulated in a more uniform soil profiles with less differentiation between top and subsoils. In addition, clay soils (Vertosols) and soils with high $CaCO_3$ (calcarosols) appeared to have larger Li concentrations. These observations supported the anomalous map prediction on various parts of Australia mentioned earlier.

[Figure 8]

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

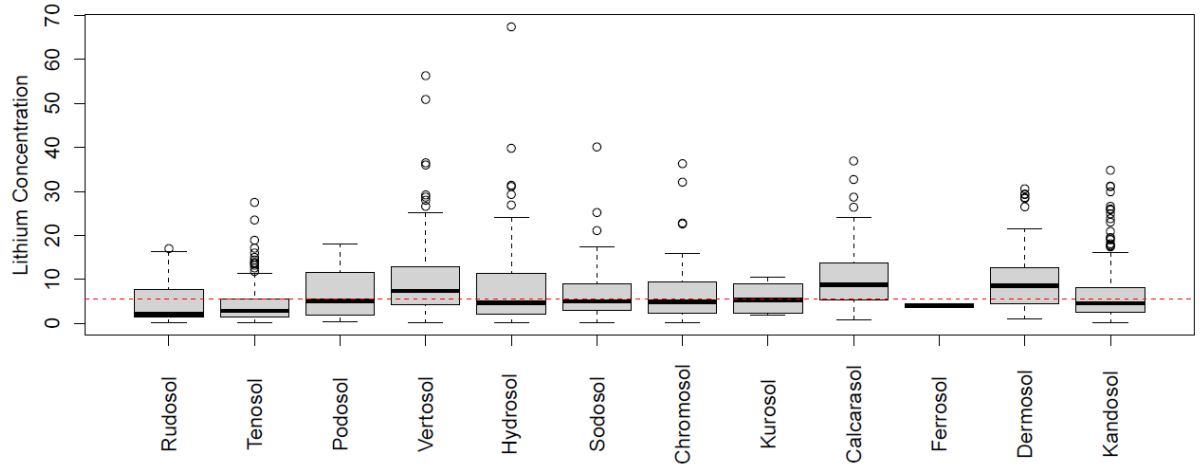


Figure 8. Boxplots of ~~lithium~~ concentration across various soil orders based on the Australian Soil Classification (ASC) system. Red dashed line represented the median values of Li~~thium~~ across both TOS and BOS depth.

The highest predicted values on the Li digital soil maps are 28 mg kg^{-1} and 22 mg kg^{-1} in TOS and BOS, respectively. Although higher Li concentration was expected to be observed in the deeper layer, the model used in this study was not able to support such predictions yet. This is most likely because the covariates used within the model represent observations from TOS instead of BOS. The variance of covariates within BOS was not obtained, and hence yielding lower accuracy predictions.

3.5 Study limitations

While we have successfully modelled soil Li distribution in Australia and validated it using an independent sample dataset, we recognise that there are limitations to this study's approach. (1) The NGSA data used apply to catchment outlet sediment representing the local accumulation of mainly detrital minerals. Therefore, strictly speaking, the predictions developed herein apply only to similar alluvial soils. (2) Despite the large amount and spread of data, the NGSA does not cover the whole of Australia. Notably, there is a data gap in parts of Western Australia and South Australia. (3) The environmental covariates used in the study were selected based on our understanding of relevant soil-forming processes. (4) There is also limited information on how the covariates vary with depth, except for the soil texture (sand and clay content) data. The inclusion of more environmental covariates related to depth and geological information may improve the predictive capability of these machine learning models. The final product was only validated in one area within Australia (Tennant Creek – Mt Isa region in the Northern Territory and Queensland). Despite our predictions of elevated soil Li in parts of Queensland, New South Wales and Victoria, ground-truthing is required to confirm them and further work is necessary to determine the origin of the contained Li.

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

4 Conclusions

Spatial prediction models have been increasingly utilised to help minimise risk and thus cost of mineral exploration. In this study, digital soil mapping for Li concentrations at two different depths (TOS: 0–10 cm, BOS: ~60–80 cm) based on the Cubist model was carried out across Australia using the National Geochemical Survey of Australia data and publicly available environmental covariates. Geology and mineralogy are of high importance in predicting soil Li anomalies, as demonstrated by the reliance of the model on the Landsat and gamma-ray radiometric covariates. Despite most mineral exploration for Li being conducted in Western Australia, other regions (such as Queensland, New South Wales and Victoria) have elevated predicted Li concentration and could become potential areas of interest with anomalous Li concentration. The model accuracy tested on the independent Northern Australia Geochemical Survey (TOS only) was reasonable compared to the calibration model performance. Overall, the model performance was on the low side and inclusion of the results into a prospectivity framework needs to consider the model uncertainties. This approach provides an estimate of the environmental background concentration of Li which is reflecting a range of processes including source rock geochemistry from which the sediments were derived, weathering (including pedogenic) and geomorphic processes. The work provides a framework to better understand the process (as revealed through the covariate relationships) controlling Li concentration at the surface and the modelling effectively delineates regions with locally higher Li source potential. Future work should include other relevant environmental covariates, which could further improve model performance, ground-truthing of anomalous regions, and investigation of ultimate Li sources. As more survey data are collected, the use of more complex models can also be explored including the use of Li concentrations in bedrock materials.

CRediT authorship contribution statement

WN: conceptualisation, data curation, analysis, writing – original draft, review and editing. **BM**: conceptualisation, methodology, writing. **AM**: conceptualisation, methodology, writing. **PdeC**: conceptualisation, data curation, writing. **JW**: conceptualisation, data curation, writing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The Li geochemical data for calibration and validation are available at: http://dx.doi.org/10.11636/Record.2011.020 (De Caritat and Cooper, 2011a) and http://dx.doi.org/10.11636/Record.2019.002 (Main et al., 2019) respectively. The covariates data used for this study was sourced from Terrestrial Ecosystem Research Network (TERN) infrastructure, which is enabled by the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS) https://esoil.io/TERNLandscapes/Public/Products/TERN/Covariates/Mosaics/90m/ (TERN, 2019).

Acknowledgments

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

BM is supported by the ARC Discovery project Forecasting soil conditions (DP200102542). AMcB is supported by the ARC Laureate Fellowship 'A calculable approach to securing Australia's soil' (FL210100054). The National Geochemical Survey of Australia (NGSA) project (http://www.ga.gov.au/ngsa) was funded by the Australian Government's Onshore Energy Security Program (OESP 2007-2011). The Geochemical Grids of Australia (GGA) project is funded by the Australian Government's Exploring for the Future (EFTF 2020-2024) initiative. We acknowledge the traditional custodians of the lands on which these samples were collected and thank all landowners for granting access to the NGSA sampling sites. We are grateful to the Geoscience Australia laboratory staff for their assistance with sample preparation. We thank Geoscience Australia reviewers for their detailed and constructive critique of our work. PdeC and JW publish with the permission from the Chief Executive Officer, Geoscience Australia.

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

5 References

Alhumimidi, M. S., Aboud, E., Alqahtani, F., Al-Battahien, A., Saud, R., Alqahtani, H. H., Aljuhani, N., Alyousif, M. M., and Alyousef, K. A.: Gamma-ray spectrometric survey for mineral exploration at Baljurashi area, Saudi Arabia, Journal of Radiation Research and Applied Sciences, 14, 82-90, https://doi.org/10.1080/16878507.2020.1856600, 2021.

Anderson, M. A., Bertsch, P. M., and Miller, W. P.: The distribution of lithium in selected soils and surface waters of the southeastern USA, Appl Geochem, 3, 205-212, 1988.

Aral, H. and Vecchio-Sadus, A.: Toxicity of Lithium to humans and the environment—A literature review, Ecotoxicology and Environmental Safety, 70, 349-356, https://doi.org/10.1016/j.ecoenv.2008.02.026, 2008.

Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, B., Roger, J. M., and McBratney, A.: Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy, Trac-Trend Anal Chem, 29, 1073-1081, https://doi.org/10.1016/j.trac.2010.05.006, 2010.

Benson, T. R., Coble, M. A., Rytuba, J. J., and Mahood, G. A.: Lithium enrichment in intracontinental rhyolite magmas leads to Li deposits in caldera basins, Nat Commun, 8, https://doi.org/10.1038/s41467-017-00234-y, 2017.

Bishop, T. F. A., McBratney, A. B., and Laslett, G. M.: Modelling soil attribute depth functions with equal-area quadratic smoothing splines, Geoderma, 91, 27-45, https://doi.org/10.1016/S0016-7061(99)00003-8, 1999.

Cardoso-Fernandes, J., Lima, A., and Teodoro, A.: Potential of Sentinel-2 data in the detection of Lithium (Li)-bearing pegmatites: a study case, SPIE Remote Sensing, SPIE Remote Sensing, Berlin, Germany, https://doi.org/10.1117/12.2326285, 2018.

Cardoso-Fernandes, J., Teodoro, A. C., and Lima, A.: Remote sensing data in lithium (Li) exploration: A new approach for the detection of Li-bearing pegmatites, Int J Appl Earth Obs, 76, 10-25, https://doi.org/10.1016/j.jag.2018.11.001, 2019.

Carranza, E. J. M.: Handbook of Exploration and Environmental Geochemistry, Elsevier Science B.V., 189-247 pp., https://doi.org/10.1016/S1874-2734(09)70011-7, 2008.

Carranza, E. J. M.: Geocomputation of mineral exploration targets, Computers & Geosciences, 37, 1907-1916, https://doi.org/10.1016/j.cageo.2011.11.009, 2011.

Champion, D.: Australian Resource Reviews: Lithium 2018, 2019.

Cŕosta, A. P., De Souza Filho, C. R., Azevedo, F., and Brodie, C.: Targeting key alteration minerals in epithermal deposits in Patagonia, Argentina, using ASTER imagery and principal component analysis, Int J Remote Sens, 24, 4233-4240, https://doi.org/10.1080/0143116031000152291, 2003.

de Caritat, P. and Cooper, M.: National Geochemical Survey of Australia: The Geochemical Atlas of Australia. Record 2011/020, Geoscience Australia [dataset], http://dx.doi.org/10.11636/Record.2011.020, 2011a.

de Caritat, P. and Cooper, M.: National Geochemical Survey of Australia: Data Quality Assessment. Record 2011/021, 2011b.

de Caritat, P. and Cooper, M.: A continental-scale geochemical atlas for resource exploration and environmental management: the National Geochemical Survey of Australia, Geochemistry: Exploration, Environment, Analysis, 16, 3-13, https://doi.org/10.1144/geochem2014-322, 2016.

de Caritat, P. and Reimann, C.: Comparing results from two continental geochemical surveys to world soil composition and deriving Predicted Empirical Global Soil (PEGS2) reference values, Earth and Planetary Science Letters, 319-320, 269-276, https://doi.org/10.1016/j.epsl.2011.12.033, 2012.

de Caritat, P., Cooper, M., Lech, M., McPherson, A., and Thun, C.: National Geochemical Survey of Australia: Sample Preparation Manual. Record 2009/08, 2009.

de Caritat, P., Cooper, M., Pappas, W., Thun, C., and Webber, E.: National Geochemical Survey of Australia: Analytical Methods Manual. Record 2010/15, 2010.

Dickson, B. L. and Scott, K. M.: Interpretation of aerial gamma-ray surveys - adding the geochemical factors, AGSO Journal of Australian Geology and Geophysics, 17, 187-200, 1997.

Dickson, B. L., Fraser, S. J., and Kinsey-Henderson, A.: Interpreting aerial gamma-ray surveys utilising geomorphological and weathering models, Journal of Geochemical Exploration, 57, 75-88, https://doi.org/10.1016/S0375-6742(96)00017-9, 1996.

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

505 Ducart, D. F., Silva, A. M., Toledo, C. L. B., and de Assis, L. M.: Mapping iron oxides with Landsat-8/OLI and EO-1/Hyperion imagery from the Serra Norte iron deposits in the Carajas Mineral Province, Brazil, Braz J Geol, 46, 331-349, https://doi.org/10.1590/2317-4889201620160023, 2016.
Geochemical Atlas of Europe: Li-Lithium: http://weppi.gtk.fi/publ/foregsatlas/text/Li.pdf, last access: 08/09.
Gallant, J. and Austin, J.: Topographic Wetness Index derived from 1" SRTM DEM-H. v2. CSIRO. Data Collection, 2012a.
510 Gallant, J. and Austin, J.: Slope derived from 1" SRTM DEM-S. v4. CSIRO. Data Collection, 2012b.
Gallant, J., Wilson, N., Dowling, T., Read, A., Inskeep, C: SRTM-derived 1 Second Digital Elevation Models Version 1.0. Record 1, 2011.
Explorer makes significant Lithium discovery in North Queensland: https://www.australianmining.com.au/news/explorer-makes-significant-lithium-discovery-in-north-queensland/, last access: 14 March.
515 Gopp, N. V., Savenkov, O. A., Nechaeva, T. V., and Smirnova, N. V.: The Use of NDVI in Digital Mapping of the Content of Available Lithium in the Arable Horizon of Soils in Southwestern Siberia, Izvestiya - Atmospheric and Ocean Physics, 54, 1152-1157, https://doi.org/10.1134/S0001433818090165, 2018.
Graedel, T. E., Barr, R., Chandler, C., Chase, T., Choi, J., Christoffersen, L., Friedlander, E., Henly, C., Jun, C., Nassar, N., Schechner, D., Warren, S., Yang, M.-y., and Zhu, C.: Methodology of Metal Criticality Determination, Environmental
520 Science & Technology, 46, 1063-1070, https://doi.org/10.1021/es203534z, 2012.
Grosjean, C., Miranda, P. H., Perrin, M., and Poggi, P.: Assessment of world Lithium resources and consequences of their geographic distribution on the expected development of the electric vehicle industry, Renewable and Sustainable Energy Reviews, 16, 1735-1744, https://doi.org/10.1016/j.rser.2011.11.023, 2012.
Harwood, T.: 9s climatology for continental Australia 1976-2005: BIOCLIM variable suite. v1. CSIRO. Data Collection,
525 2019.
Hijmans, R. J.: raster: Geographic Data Analysis and Modeling. R package version 3.5-2 [code], 2021.
Hughes, A.: Australian Operating Mines Map 2019, 2020.
Isbell, R. F. and NCST: The Australian Soil Classification, CSIRO Publishing, Melbourne, Victoria, 192 pp., https://doi.org/10.1071/9781486314782, 2021.
530 Jaireth, S., Bastrakov, E. N., Wilford, J., English, P., Magee, J., Clarke, J., Caritat, P. d., Mernagh, T. P., McPherson, A., and Thomas, M.: Map of Salt Lake Systems Prospective for Lithium Deposits, 2013.
Jenny, H.: Factors of Soil Formation: A System of Quantitative Pedology, McGraw-Hill, New York, NY1941.
Johnson, A. K.: Regolith and associated mineral systems of the Eucla Basin, South Australia, Department of Geology ad Geophysics, Adelaide University, 2015.
535 Jooshaki, M., Nad, A., and Michaux, S.: A Systematic Review on the Application of Machine Learning in Exploiting Mineralogical Data in Mining and Mineral Industry, Minerals, 11, 816, 2021.
Kabata-Pendias, A.: Biogeochemistry of Lithium, Proc. Int. Symp. Lithium in the Trophic Chain SoilPlant-Animal-Man, Warsaw, September 13–14,
Kabata-Pendias, A.: Trace elements in soils and plants (4th ed.), CRC press2010.
540 Kashin, V. K.: Lithium in Soils and Plants of Western Transbaikalia, Eurasian Soil Science, 52, 359-369, https://doi.org/10.1134/S1064229319040094, 2019.
Khorram, S., Koch, F. H., van der Wiele, C. F., and Nelson, S. A. C.: Remote Sensing, Springer, New York2012.
Köhler, M., Hanelli, D., Schaefer, S., Barth, A., Knobloch, A., Hielscher, P., Cardoso-Fernandes, J., Lima, A., and Teodoro, A. C.: Lithium Potential Mapping Using Artificial Neural Networks: A Case Study from Central Portugal, 11, 1046, 2021.
545 Kuhn, M.: caret: Classification and Regression Training. R package version 6.0-90 [code], 2021.
Kuhn, M. and Quinlan, R.: Cubist: Rule- And Instance-Based Regression Modeling. R package version 0.3.0 [code], 2021.
Lau, I., Bateman, R., Beattie, E., de Caritat, P., Thomas, M., Ong, C., Laukamp, C., Caccetta, M., Wang, R., and Cudahy, T.: National Geochemical Survey of Australia reflectance spectroscopy measurements. v4. CSIRO. Data Collection., https://doi.org/10.25919/5cdba18939c29, 2016.
550 Main, P. T. and Champion, D. C.: Levelling of multi-generational and spatially isolated geochemical surveys, Journal of Geochemical Exploration, 240, https://doi.org/10.1016/j.gexplo.2022.107028, 2022.
Main, P. T., Bastrakov, E. N., Wygralak, A. S., and Khan, M.: Northern Australia Geochemical Survey: Data Release 2 – Total (coarse fraction), Aqua Regia (coarse and fine fraction), and Fire Assay (coarse and fine fraction) element contents, Geoscience Australia, Canberra. [dataset], http://dx.doi.org/10.11636/Record.2019.002, 2019.

555 Malone, B. and Searle, R.: Updating the Australian digital soil texture mapping (Part 2 *): spatial modelling of merged field and lab measurements, Soil Research, 59, 435-451, https://doi.org/10.1071/SR20284, 2021.
Malone, B. P., McBratney, A. B., Minasny, B., and Laslett, G. M.: Mapping continuous depth functions of soil carbon storage and available water capacity, Geoderma, 154, 138-152, https://doi.org/10.1016/j.geoderma.2009.10.007, 2009.
McBratney, A. B., Mendonça Santos, M. L., and Minasny, B.: On digital soil mapping, Geoderma, 117, 3-52,
560 https://doi.org/10.1016/S0016-7061(03)00223-4, 2003.
Merian, E. and Clarkson, T. W.: Metals and their compounds in the environment, VCH, Weinheim1991.
Mernagh, T. P., Bastrakov, E. N., Jaireth, S., Caritat, P. d., English, P. M., and Clarke, J. D. A.: A review of Australian salt lakes and associated mineral systems, 2015.
Mernagh, T. P., Bastrakov, E. N., Jaireth, S., de Caritat, P., English, P. M., and Clarke, J. D. A.: A review of Australian salt
565 lakes and associated mineral systems, Australian Journal of Earth Sciences, 63, 131-157, https://doi.org/10.1080/08120099.2016.1149517, 2016.
Mudd, G. M., Werner, T. T., Weng, Z.-H., Yellishetty, M., Yuan, Y., McAlpine, S. R. B., Skirrow, R., and Czarnota, K.: Critical Minerals in Australia: A Review of Opportunities and Research Needs. Record 2018/51, 2018.
Négrel, P., Ladenberger, A., Reimann, C., Birke, M., Demetriades, A., and Sadeghi, M.: GEMAS: Geochemical background
570 and mineral potential of emerging tech-critical elements in Europe revealed from low-sampling density geochemical mapping, Appl Geochem, 111, 104425, https://doi.org/10.1016/j.apgeochem.2019.104425, 2019.
Perrotta, M. M., Souza Filho, C. R., and Leite, C. A. S.: Mapeamento espectral de intrusões pegmatíticas relacionadas a mineralizações de lítio, gemas e minerais industriais na região do Vale do Jequitinhonha (MG) a partir de imagens ASTER, Proceedings of the Anais do XII Simpósio Brasileiro de Sensoriamento Remoto, Goiânia, Brasil, 16–21 April 2005, 1855–
575 1862,
Poudjom Djomani, Y., Minty, B. R. S., Hutchens, M., and Lane, R. J. L.: Total Magnetic Intensity (TMI) Grid of Australia 2019 - seventh edition - 80 m cell size, 2019.
Pour, A. B. and Hashim, M.: Hydrothermal alteration mapping from Landsat-8 data, Sar Cheshmeh copper mining district, south-eastern Islamic Republic of Iran, Journal of Taibah University for Science, 9, 155-166,
580 https://doi.org/10.1016/j.jtusci.2014.11.008, 2015.
Quinlan, J. R.: C4.5: Programs for Machine Learning, Morgan Kaufmann Publishers Inc., San Mateo, California1993.
Reimann, C. and de Caritat, P.: Establishing geochemical background variation and threshold values for 59 elements in Australian surface soil, Sci Total Environ, 578, 633-648, https://doi.org/10.1016/j.scitotenv.2016.11.010, 2017.
Roberts, D., Wilford, J., and Ghattas, O.: Exposed soil and mineral map of the Australian continent revealing the land at its
585 barest, Nat Commun, 10, 5297, https://doi.org/10.1038/s41467-019-13276-1, 2019.
Robinson, B. H., Yalamanchali, R., Reiser, R., and Dickinson, N. M.: Lithium as an emerging environmental contaminant: Mobility in the soil-plant system, Chemosphere, 197, 1-6, https://doi.org/10.1016/j.chemosphere.2018.01.012, 2018.
Schrauzer, G. N.: Lithium: Occurrence, Dietary Intakes, Nutritional Essentiality, Journal of the American College of Nutrition, 21, 14-21, https://doi.org/10.1080/07315724.2002.10719188, 2002.
590 Searle, R.: Australian Soil Classification Map. Version 1.0.0. Terrestrial Ecosystem Research Network [dataset], https://doi.org/10.25901/edyr-wg85, 2021.
Senior, A., Britt, A.F., Summerfield, D., Hughes, A., Hitchman, A., Cross, A., Sexton, M., Pheeney, J., Teh, M., Hill, J., Cooper, M.: Australia's Identified Mineral Resources 2021, Geoscience Australia, Canberra, http://dx.doi.org/10.11636/1327-1466.2021, 2022.
595 Sitando, O. and Crouse, P. L.: Processing of a Zimbabwean petalite to obtain lithium carbonate, International Journal of Mineral Processing, 102-103, 45-50, https://doi.org/10.1016/j.minpro.2011.09.014, 2012.
Soil Science Society of America Glossary. Available at: https://www.soils.org/publications/soils-glossary, last
Starkey, H. C.: The Role of Clays in Fixing Lithium, Report 1278F, https://doi.org/10.3133/b1278F, 1982.
TERN: TERN Landscape Covariates 90m, Terrestrial Ecosystem Research Network [dataset],
600 https://esoil.io/TERNLandscapes/Public/Products/TERN/Covariates/Mosaics/90m/, 2019.
Vieceli, N., Nogueira, C. A., Pereira, M. F. C., Durão, F. O., Guimarães, C., and Margarido, F.: Recovery of lithium carbonate by acid digestion and hydrometallurgical processing from mechanically activated lepidolite, Hydrometallurgy, 175, 1-10, https://doi.org/10.1016/j.hydromet.2017.10.022, 2018.

https://doi.org/10.5194/essd-2022-418
Preprint. Discussion started: 13 January 2023

Wilford, J.: A weathering intensity index for the Australian continent using airborne gamma-ray spectrometry and digital terrain analysis, Geoderma, 183-184, 124-142, https://doi.org/10.1016/j.geoderma.2010.12.022, 2012.
Wilford, J. and Roberts, D.: Landsat 30+ Barest Earth, 2019.
Wilford, J., Worrall, L., and Minty, B.: Radiometric map of Australia provides new insights into uranium prospectivity, Ausgeo News, 95, 1-4, 2009.
Wilford, J. R. and Kroll, A.: Complete Radiometric Grid of Australia (Radmap) v4 2019 with modelled infill, 2020.
Wilford, J. R., Bierwirth, P. N., and Craig, M. A.: Application of airborne gamma-ray spectrometry in soil/regolith mapping and applied geomorphology, AGSO Journal of Australian Geology and Geophysics, 17, 201-216, 1997.
Wilson, J. and Gallant, J.: Primary topographic attributes, in: Terrain Analysis: Principles and Applications, edited by: Wilson, J. P., and Gallant, J. C., John Wiley & Sons, 51-85, 2000.
Zuo, R.: Geodata Science-Based Mineral Prospectivity Mapping: A Review, Natural Resources Research, 29, 3415-3424, https://doi.org/10.1007/s11053-020-09700-9, 2020.