Remote Sensing of Environment 165 (2015) 86–99

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Multiple influences on regolith characteristics from continental-scale geophysical and mineralogical remote sensing data using Self-Organizing Maps

M.J. Cracknell [a,b,*], A.M. Reading [a,b], P. de Caritat [c,d]

[a] School of Physical Sciences (Earth Sciences), University of Tasmania, Hobart, TAS 7001, Australia
[b] Centre of Excellence in Ore Deposits (CODES), University of Tasmania, Hobart, TAS 7001, Australia
[c] Geoscience Australia, Canberra, ACT 2601, Australia
[d] Research School of Earth Sciences, Australian National University, Canberra, ACT 2601, Australia

ARTICLE INFO

Article history:
Received 6 January 2015
Received in revised form 17 April 2015
Accepted 25 April 2015
Available online xxxx

Keywords:
Self-Organizing Maps
Unsupervised clustering
Geophysics
Regolith
Mineral prospectivity
Australia

ABSTRACT

We investigate the characteristics of regolith through the application of statistical learning to diverse layers of terrestrial, continental-scale remote sensing data. This combination allows us to explore the multiple influences of bedrock, climate, biota, landscape and time on regolith development and properties: an interdisciplinary geoscience modeling problem. From a wide variety of available data for Australia, we select remotely sensed geophysical, geomorphological and mineralogical inputs with good spatial coverage. We use Self-Organizing Maps (SOM), a topologically constrained unsupervised statistical learning algorithm, to characterize the geophysical and mineralogical signatures of regolith and bedrock. Regolith materials cover more than 80% of the Australian continent, range in age from Precambrian to Quaternary and vary in thickness from less than a meter to more than a kilometer. The diversity of regolith cover type and character across Australia provides an opportunity to demonstrate knowledge discovery from remote sensing data. The outputs of our SOM analysis are combined with ground observations from locations showing naturally occurring anomalous concentrations of nickel, tin and uranium. We identify a minimum number of natural clusters indicating subtle but significant differences in regolith and bedrock mineralization characteristics. Our results show that SOM identifies spatially contiguous regions representing unique regolith and bedrock materials. In the Yilgarn Craton we observe key differences in landscape character, density of the crust, and relative abundance of radioactive elements and alumino-silicate and ferric oxide minerals. These properties discriminate between nickel-prospective residual deeply weathered regolith formed on mafic and/or ultramafic bedrock and uranium-prospective Cainozoic paleochannels containing felsic bedrock source materials. National-scale data are publicly available for many continental regions, as in the Australian example, and our approach has general applicability. We demonstrate that remote sensing data may be used to understand the regolith, revealing the interplay between environmental history and bedrock character at regional scales, and differences between residual and transported regolith, provenance of source materials and their relative ages.

1. Introduction

Regolith is the mechanically and/or chemically weathered residual (*in situ*) or transported unconsolidated or secondarily re-cemented materials covering fresh rock (Scott & Pain, 2008; Taylor & Butt, 1998). Regolith is found on most terrestrial and extra-terrestrial surfaces (Clarke, 2003) and is the product of interactions between the lithosphere, biosphere, hydrosphere and atmosphere (Taylor & Eggleton, 2001). Its physical and chemical properties are influenced by climate, parent material, landscape and biological activity through time. Hence, the geochemical, mineralogical and morphological characteristics of regolith can be used to infer the parent rock from which it was derived and to unravel the prevailing environmental conditions contributing to its development (Taylor & Butt, 1998).

Australia provides researchers with unique opportunities and challenges with respect to understanding the origins of, and influences on, regolith. Approximately 80% of the Australian continent is mantled by regolith (Pain, Pillans, Roach, Worrall, & Wilford, 2012) extending to depths of less than a meter to more than a kilometer (Scott & Pain, 2008) and ranging in age from the Proterozoic-Late Proterozoic to the Quaternary/Holocene (Taylor & Butt, 1998). Modern landforms and regolith can be superimposed on ancient, preserved landscapes (and regolith) formed under different conditions to those currently observed (Pillans, 2008).

* Corresponding author at: Private Bag 79, Hobart, TAS, Australia, 7001.
E-mail address: m.j.cracknell@utas.edu.au (M.J. Cracknell).

http://dx.doi.org/10.1016/j.rse.2015.04.029

An unprecedented volume and variety of publicly available pre-competitive geoscience remote sensing data covering the Australian continent is now publically available. For example, Geoscience Australia (http://www.ga.gov.au/) provides access to interpolated and leveled gravity, airborne geophysical data, e.g., Total Magnetic Intensity (TMI) and Gamma-Ray Spectrometry (GRS), and Digital Elevation Models (DEM), and along with CSIRO (http://www.csiro.au/) distributes Advanced Spaceborne Thermal Emission and Reflection (ASTER) geoscience products representing surface mineralogical characteristics. CSIRO also provides a range of DEM derivatives, e.g., Topographic Wetness Index (TWI), Multi-resolutional Valley-Bottom Flatness (MrVBF) or Ridge-Top Flatness (MrRTF) indices, slope and aspect. These data provide opportunities with which to gain an understanding of ancient climatic conditions, biological activity, landscape evolution and the nature and characteristics of large but subtle geophysical and geochemical mineralization footprints (Taylor & Butt, 1998). Nevertheless, unraveling the contribution of bedrock, climate, landscape and biota to regolith development and mineralization potential by integrating of large volume of multivariate geoscience data is a challenging task.

1.1. Statistical learning

Statistical learning offers analysts knowledge-driven (supervised) and data-driven (unsupervised) approaches for understanding integrated data. In the data-driven case, observations representing a random vector X, having joint probability density $\Pr(X)$, are presented to the unsupervised learning algorithm. The aim is to infer the properties of this probability density without prior conditions on the arrangement of data (Hastie, Tibshirani, & Friedman, 2009; Witten & Frank, 2005). Unsupervised learning algorithms are useful for finding natural groups or clusters of similar samples as indicated by regions of high density within $\Pr(X)$ (Hastie et al., 2009; Marsland, 2009; Xu & Wunsch, 2005). Unlike supervised learning, the validity of the results from unsupervised learning cannot be evaluated against reference information representing sample associations. Therefore, success is usually assessed by interrogating the characteristics of input variables in conjunction with ancillary data and/or knowledge of the phenomena under investigation (Hastie et al., 2009; Ripley, 1996).

1.2. Self-Organizing Maps

Self-Organizing Maps (SOM; Kohonen, 1982, 2001) are a type of an unsupervised learning algorithm based on the principles of non-linear statistical models known as Neural Networks (Penn, 2005; Ripley, 1996). SOM employs competitive learning to reduce n-dimensional (nD) multivariate data to two dimensions (Bação, Lobo, & Painho, 2008; Hastie et al., 2009; Klose, 2006). This is achieved using vector quantization and measures of vector similarity as a means of "mapping" nD samples to cells (nodes) arranged in a 2D topological space (Bação et al., 2008; Hastie et al., 2009; Kohonen, 2001). The topology of neighboring nodes in 2D space indicates their relative similarities in nD space. After training, SOM nodes are represented by an nD vector (code-vector) of the same dimensionality as the input data (Bação et al., 2008). This code-vector summarizes the characteristics of samples associated with a particular node. Interrogation of the SOM component planes aids visualization of dominant patterns and structures present within multivariate data (Bierlein, Fraser, Brown, & Lees, 2008; Penn, 2005; Ultsch & Herrmann, 2005; Wehrens & Buydens, 2007).

To implement SOM an $m \times n$ dimension matrix is created from the input data with m rows of samples and n columns of variables $X = (x_1, x_2, \ldots, x_n) \in \Re^n$. SOM nodes are trained from randomly sampled reference (seed-) vectors M_i of equal length to n, via an iterative two-stage process. Firstly, seed-vectors are shown to the network and compared to all x_n that fall within a predefined radius of assessment, commonly using a Euclidean distance metric:

$$\|X - M\| = \sqrt{\sum_{i=1}^{n} (x_i - m_i)^2}. \tag{1}$$

The closest seed-vector, deemed most similar to the considered x_n, is the best-matching unit (node) M_c according to:

$$\|X - M_c\| = \min_i \{\|X - M_i\|\}. \tag{2}$$

Secondly, the weights of M_c and its neighboring M_i within the search neighborhood $N(t)$ of radius r, are adjusted to more closely correspond to the properties of x_n. The learning rate factor $a(t)$ controls the rate of change of M_i during the adjustment process. These steps are repeated while reducing $N(t)$ and $a(t)$ for a given iteration t. In this way, M_i become trained nodes linked to code-vectors that summarize the characteristics of associated input samples (Bierlein et al., 2008; Fraser & Dickson, 2007; Peeters, Lobo, & Dassargues, 2007).

SOM has been shown to be useful for identifying, visualizing and analyzing coherent groups within multivariate geoscience data. For example, Penn (2005) employed SOM to efficiently classify major element geochemical data and hyperspectral reflectance data to interpret the petrogenetic characteristics of igneous rocks. Peeters et al. (2007) used SOM to investigate the hydrochemical properties of confined aquifers. Bierlein et al. (2008) applied SOM to databases of mineral deposits and major faults using SOM. Their comparisons suggested that major differences between deposit type, host-rock mineralization style could be identified. Bedini (2009, 2012) used SOM to identify exposed lithologies and alteration halos in sparsely vegetated high-latitude regions from HyMap™ imagery. Iwashita, Friedel, Filho, C. R. de S, and Fraser (2011) integrated airborne geophysical and soil geochemical data with SOM to investigate geochemical weathering processes. Carneiro, Fraser, Croacutesta, Silva, and Barros (2012) mapped geological units from airborne geophysical data in a heavily vegetated region of the Amazon Basin using SOM. Abedi, Norouzi, and Torabi (2013) successfully employed SOM to generate clusters representing varying degrees of mineral prospectivity for a copper (Cu) deposit in Iran. Cracknell, Reading, and McNeill (2014) applied SOM to soil geochemical data linked to samples representing discrete volcanic units as a means of mapping volcanic-hosted massive sulfide alteration halos and subtle variations in their primary composition.

1.3. Study aims

This study builds on the analysis documented in Cracknell and Reading (2014), which demonstrated the use of publicly available continental-scale geophysical and mineralogical remote sensing data in conjunction with SOM. The results of that study identified and visualized patterns in multivariate geophysical and mineralogical remote sensing data that relate to regolith and bedrock geological materials. In this contribution, we extend our analysis to characterize spectral and spatial correlations in these data. Our analysis provides a context for mapping regolith, bedrock mineralization footprints and multiple influences on landscape evolution and environment.

Mineral commodities provide spatially extensive reference information, and we use nickel (Ni), tin (Sn) and uranium (U) as case study examples for our SOM analyses. Mineral deposits of Ni, Sn and U form in contrasting geological environments and are mined from both bedrock and regolith materials. Additional information on typical tectonic settings and geological characteristics of Ni, Sn and U mineralization in Australia is given in *Appendix A — Mineral deposit characteristics*.