SOIL

Quantifying soil and critical zone variability in a forested catchment through digital soil mapping

M. Holleran[1,*], **M. Levi**[2], **and C. Rasmussen**[1]

[1]Department of Soil, Water and Environmental Science, Univ. of Arizona, Tucson, Arizona, USA
[2]USDA-ARS Jornada Experimental Range, New Mexico State Univ., Las Cruces, New Mexico, USA
[*]now at: Geosyntec Consultants, San Francisco, California, USA

Correspondence to: C. Rasmussen (crasmuss@cals.arizona.edu)

Abstract. Quantifying catchment-scale soil property variation yields insights into critical zone evolution and function. The objective of this study was to quantify and predict the spatial distribution of soil properties within a high-elevation forested catchment in southern Arizona, USA, using a combined set of digital soil mapping (DSM) and sampling design techniques to quantify catchment-scale soil spatial variability that would inform interpretation of soil-forming processes. The study focused on a 6 ha catchment on granitic parent materials under mixed-conifer forest, with a mean elevation of 2400 m a.s.l, mean annual temperature of 10 °C, and mean annual precipitation of ~ 85 cm yr^{-1}. The sample design was developed using a unique combination of iterative principal component analysis (iPCA) of environmental covariates derived from remotely sensed imagery and topography, and a conditioned Latin hypercube sampling (cLHS) scheme. Samples were collected by genetic horizon from 24 soil profiles excavated to the depth of refusal and characterized for soil mineral assemblage, geochemical composition, and general soil physical and chemical properties. Soil properties were extrapolated across the entire catchment using a combination of least-squares linear regression between soil properties and selected environmental covariates, and spatial interpolation or regression residual using inverse distance weighting (IDW). Model results indicated that convergent portions of the landscape contained deeper soils, higher clay and carbon content, and greater Na mass loss relative to adjacent slopes and divergent ridgelines. The results of this study indicated that (i) the coupled application of iPCA and cLHS produced a sampling scheme that captured the greater part of catchment-scale soil variability; (ii) application of relatively simple regression models and IDW interpolation of residuals described well the variance in measured soil properties and predicted spatial correlation of soil properties to landscape structure; and (iii) at this scale of observation, 6 ha catchment, topographic covariates explained more variation in soil properties than vegetation covariates. The DSM techniques applied here provide a framework for interpreting catchment-scale variation in critical zone process and evolution. Future work will focus on coupling results from this coupled empirical–statistical approach to output from mechanistic, process-based numerical models of critical zone process and evolution.

1 Introduction

The spatial complexity of soils presents a significant hurdle to predicting and modeling critical zone (CZ) processes and characteristics, where the CZ is defined as the earth surface system that extends from the top of the canopy down to groundwater (NRC, 2001). Developing robust data-driven methods that provide accurate, reliable, and high-resolution characterization of soil properties is a major challenge to earth scientists and is needed for better understanding and quantification of CZ process and function such as soil erosion, hydrologic cycling, and carbon cycling (NRC, 2010). Here we address this challenge in a forested catchment using a combination of digital soil mapping (DSM) and statistical approaches to quantify soil physical and chemical properties

at high spatial resolution (~ 2 m pixels). The analyses described herein provide one means to unravel the catchment-scale soil complexity that is central to CZ function and evolution.

The coupled use of DSM and statistically based sampling designs has greatly improved the quality and resolution of predicted soil variability (McBratney et al., 2000, 2003; Park and Vlek, 2002; Kempen et al., 2012; Goovaerts, 2000). These approaches build from traditional methods of conceptualizing soil variability (Dokuchaev, 1967; Jenny, 1941) and soil survey (Soil Survey Staff, 1999) using ancillary data, such as remotely sensed imagery and digital elevation data, as spatially extensive measures and proxies of soil-forming factors that may be used to develop predictive models of soil properties and spatial variability (Buchanan et al., 2012; Scull et al., 2005). A core principle of DSM may be derived from the classic statement of soil-forming factors (Jenny, 1941) restated as SCORPAN (McBratney et al., 2003): $S = f(c, o, r, p, a, n)$, where a soil property or soil type (S) is a function of the external factors of climate (c), organisms (o), relief (r), parent material (p), age (a), and its location in space (n). Recent advances in environmental sensing, ancillary data production, and modeling facilitate quantifying these external factors, or what have been termed "environmental covariates", as digital spatial data sets. For example, remotely sensed data provide proxies for organisms and density of vegetative cover in addition to mineralogy of soils and parent materials (Sullivan et al., 2005; Saadat et al., 2008), whereas digital elevation data provide proxies for relief and local-scale variation in climate, namely solar radiation and surface water redistribution (Ziadat, 2005; Moore et al., 1993).

These quantitative environmental covariates may then be used to predict soil property spatial distribution using spatial soil prediction models. For example, remotely sensed reflectance has been quantitatively related to soil properties such as particle size (Dematte and Nanni, 2003; Salisbury and Daria, 1992; Ben-Dor et al., 2002), mineralogy (Ben-Dor et al., 2003; Dematte et al., 2004; Galvao et al., 2008), soluble salts (Howari et al., 2002), organic matter, and extractable iron (Ben-Dor, 2002) in both laboratory and field settings. Similarly, digital terrain models and terrain attributes, such as slope, aspect, surface curvature and roughness, and wetness indices, have been used to predict surface redistribution of water and sediment as well as variation in solar energy inputs to the soil surface (Moore et al., 1991; Irvin et al., 1997; Florinsky, 1998). These models include a wide range of methods, including but not limited to regression analyses, principal component analyses, supervised and unsupervised classification techniques, and geostatistical methods (Eldeiry and Garcia, 2010; Hengl et al., 2007a; McKenzie and Ryan, 1999).

Developing sampling schemes from a set of geospatial data that captures the greater part of landscape variance can improve the efficiency and effectiveness of the sampling and modeling process (Brus and Heuvelink, 2007; Vasat et al., 2010). Here we apply a data-driven approach that combines an iterative principal component analysis (iPCA) data reduction with a conditioned Latin hypercube sampling (cLHS) sampling scheme to develop a sample and model framework for soil property prediction. The iPCA approach selects the covariate data that account for the greatest range of landscape variability, whereas the cLHS approach is used to design a sample set that effectively captures the geographic space and the feature space of covariate layers (Mulder et al., 2013; Hansen et al., 2012; Van Camp and Walraevens, 2009; Minasny and McBratney, 2006). Levi and Rasmussen (2014) applied similar iPCA–cLHS techniques to characterize soil–landscape variability and predict soil physical properties for a 6250 ha area in southern Arizona. The iPCA reduced an initial set of 13 environmental covariates to 4 that accounted 95 % of the variance in covariate data. Soils were then sampled in real geographic space based on the cLHS approach. Field data were coupled with covariate data using a regression kriging approach to predict soil physical properties which exhibited good agreement with vector-based soil survey maps and captured internal map unit spatial variability. Regression kriging includes using regression models to extract information from sampled locations using covariate layers and then interpolating model residuals with ordinary kriging (Odeh et al., 1994) or other geostatistical approaches.

The objective of this study was to apply a coupled iPCA–cLHS routine to build robust statistical models for predicting high-resolution soil physical and chemical properties as measures of critical zone structure and function for a forested catchment in southern Arizona, USA. Similar methods have been successfully applied over larger areas with diverse soils, parent materials, relief, and vegetation (Levi and Rasmussen, 2014); however, these techniques have not yet been tested on smaller areas, such as a single first-order catchment with limited variability in soil-forming factors. We hypothesized that (i) the coupled method of iPCA and cLHS modeling would efficiently quantify landscape variability in a small-forested catchment and (ii) that it would facilitate the use of relatively simple statistical methods to predict soil chemical and physical properties throughout the catchment.

2 Methods

2.1 Study site

This study was based in Marshall Gulch (MG), a forested catchment located in the Santa Catalina Mountains, just outside of Tucson, AZ ($111°16'41''$ W and $33°25'46''$ N) (Fig. 1). This study focused on one small ~ 6 ha north-facing catchment with an elevation range of 2300–2500 m a.s.l. The catchment was occupied by mixed conifer vegetation, primarily *Pinus ponderosa*, with lesser amounts of *Pseudotsuga menziesii* and *Abies concolor* (Whittaker and Niering, 1968). The climate of MG includes an ustic soil moisture
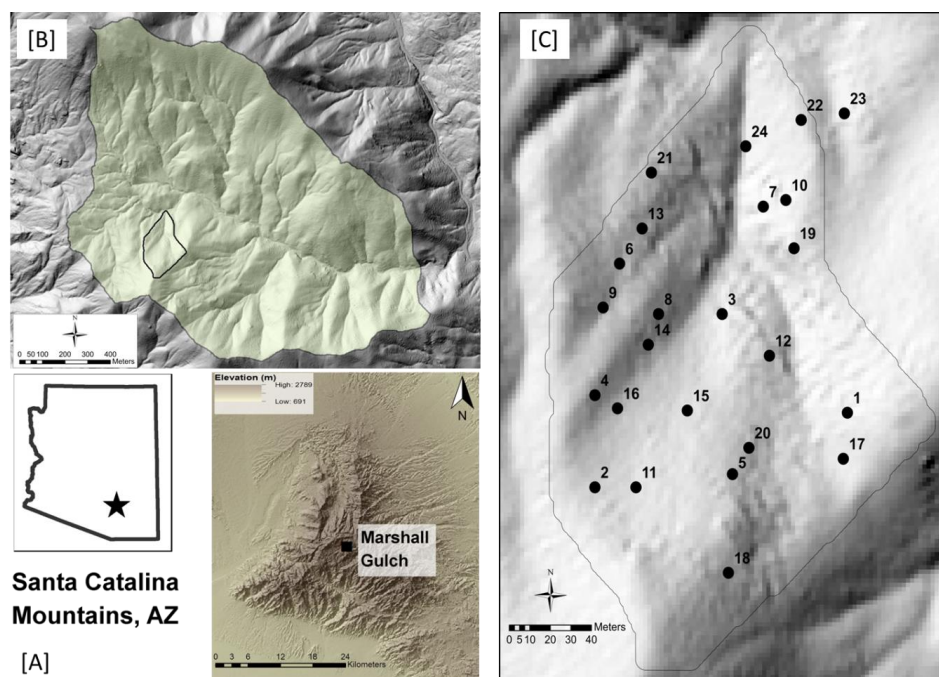
Figure 1. Location of study catchment in the **(a)** Santa Catalina Mountains in southern Arizona, USA, the **(b)** Marshall Gulch watershed, and **(c)** sample locations in the study catchment located in Marshall Gulch.

regime (Soil Survey Staff, 1999) that receives an average of ∼ 85–90 cm of annual precipitation split evenly between a combination of cold rain and snow during winter months and convective summer thunderstorm rainfall. Mean annual temperature at MG is 10 °C, with a mesic soil temperature regime (Soil Survey Staff, 1999). The catchment is bounded by ridgelines to the east and west, with very steep slopes, exceeding 45 % slope gradient in some areas, and several small ephemeral drainages that only flow following extreme precipitation events or during snowmelt.

The Santa Catalina Mountains encompass a metamorphic core complex system (Arca et al., 2010) that yields a complex array of bedrock materials proximal to the MG field site, including Tertiary-aged granitic rocks and Paleozoic-aged meta-sedimentary (Dickinson, 1992). The 1 : 250 000 geologic map of the area indicates the study catchment is situated on the Wilderness granite suite that consists of an Eocene-aged two-mica granite; however, detailed field investigation indicated the roughly 20 % of the catchment was underlain by a combination of a hornblende-rich amphibolite, in addition to areas underlain by quartzite (Figs. 2 and 3).

2.2 Environmental covariates

Covariate layers included remotely sensed 1 m pixel resolution four-band aerial imagery from the National Agriculture Imagery Program (NAIP) collected in June 2010, and topographic covariates derived from a 2 m resolution lidar-derived bare earth digital elevation model. Derived covariate layers included NAIP band ratios as well as topographically derived annual solar radiation, slope, regolith depth, and wetness. All covariate layers were clipped using 5 m buffer polygon around the selected catchment, resampled to a 2 m pixel resolution to match the lidar data, and projected to a common datum, NAD83 UTM zone 12N.

The NAIP imagery includes four bands that span red (1), blue (2), green (3), and near-infrared (4) bands. All possible band ratios were derived from these data, including B3 : B1, B2 : B1, B3 : B2, B4 : B1, and B4 : B2. Additionally, the normalized difference vegetation index (NDVI) was derived from these data as an index of surface greenness: $\text{NDVI} = (\text{NIR} - R) / (\text{NIR} + R)$, where NIR is the near-infrared band and R is the red band (Huete et al., 1985).

The System for Automated Geoscientific Analysis (SAGA) version 2.0.4 (Conrad, 2006) was used to calculate annual solar radiation [W m^{-2} yr^{-1}], slope [degrees], and the SAGA wetness index [unitless]. Terrain analysis was performed with parallel processing module using a multiple flow direction algorithm (Freeman, 1991) to compute slope, contributing area, and SAGA wetness index (Boehner et al., 2002). Total incoming solar radiation (direct and diffuse) was calculated with the incoming solar radiation module in SAGA for 1 year on a 14-day time step using (Wilson and Gallant, 2000). Additionally, a modeled soil depth layer was included as a topographic variable. Modeled soil depth was generated and validated previously for this catchment using a geomorphic framework implementing a nonlinear depth- and slope-dependent sediment transport model
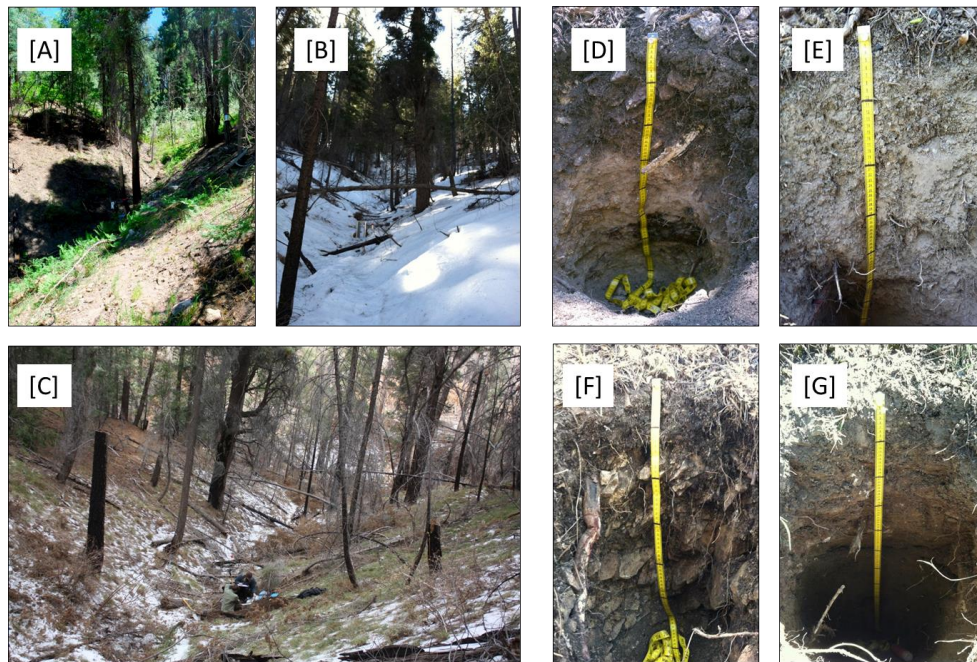
Figure 2. Study catchment seasonal cover and vegetation for **(a)** summer, **(b)** winter, and **(c)** early spring and typical soil profiles on **(d)** a granite hillslope, **(e)** granite divergent summit, **(f)** quartzite profile, and **(g)** amphibolite profile.
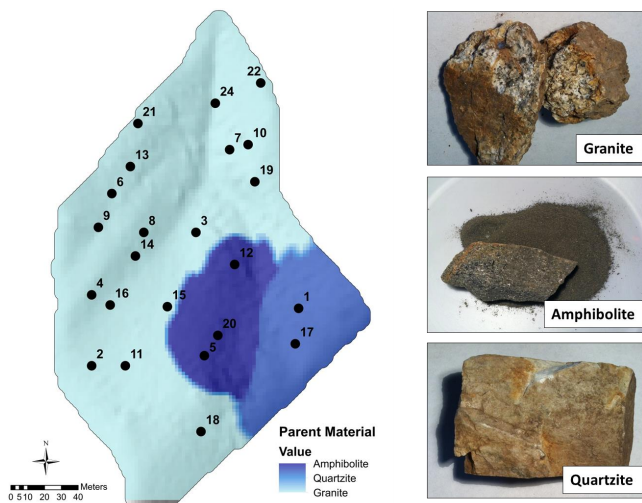

Figure 3. Modeled spatial distribution of parent materials in the study catchment and examples of typical rocks sampled from the soil profiles sampled on each parent material.

with an exponential soil production function (Pelletier and Rasmussen, 2009).

2.3 Data reduction

A data-driven iterative principal component analysis (iPCA) was used to determine those layers dominating soil–landscape variance based on Nauman (2009) and Levi and Rasmussen (2014). Similar methods have been used with multivariate soil prediction mapping (Hengl et al., 2007b; Vasat et al., 2010) to optimize sample locations and to ensure landscape variability is captured in the sampling scheme. Here we used iPCA coupled with a factor loading analysis to select the final set of covariate layers used for soil prediction models. All covariate layers were standardized using a z score prior to iPCA:

$$Z_{ij} = \frac{x_{ij} - \mu_j}{\sigma_j}, \tag{1}$$

where Z_{ij} is the z score of pixel i in layer j, x_{ij} is the untransformed value of pixel i of layer j, μ_j is the mean of layer j, and σ_j is the standard deviation of layer j prior to iPCA. The standardized data were grouped into topographic and NAIP indices, and each group was handled separately for the initial step of the data reduction. The iPCA eigenmatrix and eigenvalues were used to calculate loading factors (R_{kp}) of each input band using the degree of correlation:

$$R_{kp} = \frac{a_{kp} \cdot \sqrt{\lambda_p}}{\sqrt{\text{Var}_k}}, \tag{2}$$

where a_{kp} is the eigenvector for band k and component p, λ_p is pth eigenvalue, and Var_k is the variance of band k in the covariance matrix (Jensen, 2005). The absolute values of loading factors for each covariate layer were summed and ranked from greatest to lowest, providing a quantitative metric of the total contribution of each covariate layer to the overall

variance of the data set. The number of principal components required to reach 95 % cumulative explained variance in the data set determined the number of covariate layers to retain for subsequent iterations. The covariate layers retained were those with the greatest absolute summed loading factors. This was repeated until all principal components were needed to achieve 95 % of cumulative variance. After processing topographic parameters and NAIP reflectance ratios separately, the final layers from each group were merged and reduced in the same manner. The covariate layers determined in the final PCA included NAIP B3 : B2, NAIP NDVI, solar radiation, SAGA wetness index, slope, and modeled regolith depth. This set of covariates was used for field sample design and modeling of soil properties.

2.4 Sample design

A conditioned Latin hypercube sampling (cLHS) scheme was used to develop a sampling design to sample locations that were randomly distributed in geographic space throughout the catchment and that captured the distribution of the selected covariate layers (Minasny and McBratney, 2006). We implemented cLHS using the z-scored layers of the six covariate layers determined from the iPCA to ensure landscape variability was accounted for with the chosen sample sites. The cLHS was performed using open source code from Minasny and McBratney (http://www.iamg.org/CGEditor/index.htm, downloaded 16 February 2011) using MATLAB version 7.11.0 (The MathWorks Inc. 2010). We ran the cLHS model using 40 000 iterations to randomly determine sample site locations (n) within the study site using a $n = 5$, $n = 10$, $n = 15$, $n = 20$, and $n = 30$. It was concluded that $n = 20$ best captured the distribution of the environmental covariates using the least number of sample locations based on comparison of sample site covariate statistics (mean, range, skewness, etc.) with the original covariate layers. Four supplemental sample points (sample ID 21–24) were also included from Lybrand et al. (2011). The additional samples were incorporated where it was subjectively determined that the cLHS had missed key landscape positions, such as in the middle of the main drainage flowing out of the catchment and divergent summit positions.

2.5 Field methods and sample characterization

A rugged Trimble Yuma outdoor tablet and GPS unit was used to locate the sample sites in the field to sub-meter accuracy. At each of the 24 sample sites, soil pits were dug with a spade to the depth of refusal that coincided with the saprolite–saprock interface. Each pit was described and sampled by genetic horizons following standard protocols and soil morphological and physical properties including color, structure, consistence, root abundance, and rock fragment content were described in the field (Schoeneberger et al., 2002) (Fig. 2). In general, five horizons per pedon were described, resulting in a total of 100 collected samples. Following field sampling, it was determined that five locations were underlain by different parent materials. Specifically, pits 1 and 17 were underlain by quartzite; pits 5, 12, and 20 were underlain by a hornblende-rich amphibolite; and all remaining sample locations were underlain by granite (Fig. 3).

Soil bulk density measurements were collected at each pedon using a simple core and hammer method (Blake and Hartge, 1986). Assuming parent material homogeneity within the soil profile, representative samples of the parent material was collected at the saprolite–saprock boundary at each pedon for chemical and mineralogical analyses. All soil samples were air-dried and sieved to isolate the < 2 mm fine-earth fraction and all soil characterization conducted on this fraction unless otherwise stated. Soil pH was measured on all samples at weight to volume ratios of 1 : 2 (soil : water) solution, 1 : 2 (soil : 1 M KCl) solution, and 1 : 4 (soil : 0.02 M CaCl$_2$) solution (Soil Survey Staff, 2004). Soil electrical conductivity (μS cm^{-1}) was determined for all samples using a 1 : 2 (soil : water) extract (Burt, 2004).

Soil organic matter content was determined by loss on ignition (Konen et al., 2002) for all collected samples. Samples (20 g) were placed in a furnace at 105 °C for 24 h, weighed, and then placed in a muffle furnace at 360 °C for 2 h, with the change in weight as a proxy for organic matter. Total carbon and nitrogen content (wt/wt %) and stable isotope signature (δ^{15}N and δ^{13}C) were measured for all samples on a continuous-flow gas-ratio mass spectrometer (Finnigan Delta PlusXL) coupled to an elemental analyzer (Costech) at the University of Arizona, Environmental Isotope Laboratory. The samples did not contain carbonates and it was assumed that total carbon was equivalent to total organic carbon.

Particle size was determined by laser diffraction using a Beckman Coulter LS 13 320 laser diffraction particle size analyzer at the University of Arizona, Center for Environmental Physics and Mineralogy. Following pretreatment to remove organics using NaOCl adjusted to pH 9.5, roughly 0.2 and 0.1 g of sample were weighed and put into tubes and then mixed for 24 h with 5 mL of deionized water using a Thermo Scientific Labquake® shaker/rotator, followed by the addition of 5 mL of 5 % sodium hexametaphosphate solution for an additional 24 h to ensure dispersion of soil particles prior to particle size analysis.

2.6 Soil elemental and mineralogical characterization

Elemental concentrations for major and trace elements were determined for all soil and rock samples by X-ray fluorescence (XRF). Soil and rock samples were ground by ball milling ~ 3.5 g of sample in a plastic scintillation vial containing three tungsten carbide bearings for 10 min and then pressed into pellets at a pressure of 25 t for 120 s bound with a layer of cellulose wax (3642 cellulose binder – SPEX SamplePrep PrepAid®) and analyzed using a polarized energy-dispersive X-ray fluorescence spectrometer

(EDXRF – SPECTRO XEPOS, Kleve – Germany) at the University of Arizona, Arizona Laboratory for Emerging Contaminants.

For mineralogical analyses, samples were pretreated to remove organic matter (Jackson, 2005) with a 100 mL solution of 6 % NaOCl adjusted to a pH of 9.5, rinsed with deionized water, centrifuged, dried, and mixed. Mineral phases for soil and rock samples were identified by quantitative X-ray diffraction. A known amount of internal standard (corundum) was added to each sample to allow for quantitative interpretation of diffraction peaks. Samples were ground using a McCrone micronizing mill (Eberl, 2003). All sample preparation steps were intended to maximize random orientation and to increase the exposed surface area of the included minerals. Samples were run as random powder mounts, measured from 5 to 65° 2θ, with a step size of 0.02° 2θ at the University of Arizona, Center for Environmental Physics and Mineralogy, using a PANalytical X'Pert PRO-MPD X-ray diffraction system (PANalytical, Almelo, AA, the Netherlands) generating Cu–Kα radiation at an accelerating potential of 45 kV and current of 40 mA. The resulting diffractograms were analyzed using Rietveld analyses to identify mineral phases and mineral abundance (Moore and Reynolds, 1997).

2.7 Elemental mass transfer and elemental mass flux

Soil chemical denudation was determined using the loss of Na relative to the parent material. Sodium serves as a proxy of plagioclase feldspar weathering in granitic terrain as Na is for the most part biologically inert, and plagioclase minerals are often the first minerals to chemically decompose in granite bedrock weathering to regolith (Brantley and White, 2009). We used a dimensionless mass transfer coefficient, τ, as a measure of soil Na loss and chemical denudation relative to the parent material (Chadwick et al., 1990):

$$\tau_{j,\mathrm{w}} = \left(\frac{C_{j,\mathrm{w}} C_{i,\mathrm{p}}}{C_{i,\mathrm{w}} C_{j,\mathrm{p}}} \right) - 1, \tag{3}$$

where C is the concentration of an immobile element i, in this case the element zirconium, and a mobile element j, here Na, in the parent bedrock p and the weathered soil w to calculate relative elemental loss or gain. Values of $\tau = 0$ indicate no change from parent material, whereas $\tau > 1$ is equivalent to elemental gain, and $\tau < 1$ equivalent to elemental loss.

In addition, the volumetric strain, ε, associated with soil formation was calculated as (Brimhall and Dietrich, 1987):

$$\varepsilon_{i,\mathrm{w}} = \left(\frac{\rho_{\mathrm{p}} C_{i,\mathrm{p}}}{\rho_{\mathrm{w}} C_{i,\mathrm{w}}} \right) - 1, \tag{4}$$

where ρ_{p} is the bulk density of the parent bedrock and ρ_{w} is the bulk density of the soil. The total mass flux of Na [$m_{j,\mathrm{flux}}$; M L^{-2}] was calculated for individual soil horizons (Egli and

Fitze, 2000; Heckman and Rasmussen, 2011):

$$m_{j,\mathrm{flux}} = \rho_{\mathrm{p}} C_{j,\mathrm{p}} \tau_{\mathrm{w}} [z_{\mathrm{w}} (1 - \eta_{\mathrm{w}})] \left(\frac{1}{\varepsilon_{i,\mathrm{w}} + 1} \right), \tag{5}$$

where z_{w} is horizon thickness and η_{w} is the fractional volumetric rock fragment content. The summation of $m_{j,\mathrm{flux}}$ for all horizons in a single pedon yields the total Na elemental flux from the pedon that has occurred during pedogenesis:

$$m_{\mathrm{flux,total}} = \sum_{w=1}^{k} \sum_{j=1}^{n} m_{j,\mathrm{flux}}, \tag{6}$$

where n is the total number of elements of interest, in this case simply Na, and k is the total number of horizons in the profile.

2.8 Statistical analyses and parent material spatial modeling

Summary statistics of measured soil properties were performed on all collected soil horizons ($n = 103$) and for pedon summed values reported on a mass per area basis ($n = 24$) (Table 1). Simple means comparison of soil physicochemical properties among the different parent materials were conducted using unequal variance t tests given the unequal sample number per parent material: quartzite $n = 7$, amphibolite $n = 10$, and granite $n = 86$, where n is the number of horizons per parent material (Table 2).

Analysis of parent rock geochemistry and mineral composition indicated substantial variation of several key parameters among the three parent materials, including Ti : Zr, [Ca + Mg + Fe], and weight percent hornblende and quartz (Table 3). To define the spatial extent of the various parent materials, the lowermost C or Cr horizon values for these parameters from each sample pit were interpolated across the catchment using an inverse distance weighting (IDW) routine. Specifically, the IDW power was set equal to 2, with a circular search radius of 250 m, a maximum of 23 neighbors, and 1 search sector. The interpolated values for each parameter were then exported, standardized, and categorized into three classes using a hierarchical clustering routine using Ward's minimum variance method for computing the distance between clusters (Milligan, 1979). The resulting three clusters summarized the variation in C and Cr horizon geochemistry and mineral composition of the three parent materials and provided a spatial estimation of the extent and location of the different parent materials (Fig. 3).

2.9 Regression modeling

Soil prediction models were developed using all 24 sample locations. Target variables selected to model were clay [%], KCl pH, τ_{Na}, soil depth [cm], organic carbon content [kg m^{-2}], clay content [kg m^{-2}], and Na mass flux [kg m^{-2}]; note that modeled soil depth was not used as a predictor in the

Table 1. Summary statistics of soil variables.

	Mean	SD	Median	Minimum	Maximum	Variance	Skewness	Kurtosis	CV
				All soil horizons $n = 103$					
Clay [%]	10.9	3.7	10.2	4.8	26.9	13.6	1.3	3.3	34
C [%]	2.1	1.6	1.5	0.3	6.7	2.7	1.2	0.5	79
KCl pH	4.9	1.0	5.0	3.0	7.0	1.0	0.1	−1.1	20
τ_{Na}	−0.1	0.3	−0.2	−0.4	1.5	0.1	4.0	16.7	−242
Quartz [%]	35.3	8.5	36.7	4.1	68.6	72.7	−0.4	4.2	24
Hornblende [%]	4.6	11.5	0.0	0.0	59.1	133	3.2	10.7	253
Na [%]	1.8	0.3	1.8	0.6	2.5	0.1	−1.0	1.3	19
Al [%]	8.2	0.7	8.0	6.6	10.0	0.5	0.6	0.2	9
Si [%]	30.5	3.4	31.9	19.9	34.8	11.9	−1.7	2.0	11
[Ca + Mg + Fe] [%]	4.6	5.4	2.3	0.7	22.0	28.8	2.1	2.9	116
Ti : Zr	30.2	63.2	5.9	0.7	351	39 993	3.4	12.5	209
				Pedon sum $n = 24$					
Depth [cm]	71.5	24.9	69.5	40.0	130.0	617.7	0.7	−0.2	34.7
Clay [kg m^{-2}]	55.4	34.3	46.8	12.3	142.6	1176.2	1.0	0.4	61.9
Carbon [kg m^{-2}]	8.0	3.4	7.7	4.2	16.8	11.5	1.4	1.9	42.3
Na mass flux [kg m^{-2}]	−269	527	−249	−931	1848	277 795	3	12	−196

Table 2. Unequal variance t tests – all horizons with parent material as the main effect.

	Granite ($n = 86$)	Amphibolite ($n = 10$)	Quartzite ($n = 7$)	F ratio	P value
Clay (%)	10.83 ± 3.73	12.34 ± 3.88	9.03 ± 1.88	3.41	0.063
C (%)	2.1 ± 1.57	1.53 ± 1.59	2.86 ± 2.46	0.91	0.432
KCl pH	4.87 ± 0.96	5.15 ± 1.17	4.98 ± 1.03	0.27	0.770
Rock fragment (wt %)	50.24 ± 14.93	33.81 ± 9.32	58.18 ± 13.02	13.85	0.001
τ_{Na}	-0.19 ± 0.07	-0.19 ± 0.14	0.83 ± 0.55	11.58	0.003
Quartz (%)	36.29 ± 4.66	17.87 ± 6.81	47.6 ± 12.71	35.54	0.000
Hornblende (%)	0.98 ± 3.48	30.53 ± 16.35	1.15 ± 1.47	15.53	0.000
Na (%)	1.88 ± 0.25	1.2 ± 0.16	1.45 ± 0.56	68.61	0.000
[Ca+Mg+Fe] (%)	2.85 ± 2.81	16.79 ± 4.15	4.2 ± 0.87	51.30	0.000
Ti : Zr	11.99 ± 24.42	158.92 ± 101.57	13.33 ± 3.81	10.05	0.001

F ratio and P value are result of unequal variance t tests for each soil variable by parent material.

models for measured soil depth. These target variables were chosen to represent a range of important chemical, physical, and biological soil properties. The variables clay %, KCl pH, and τ_{Na} were modeled as depth-weighted averages for each profile where individual horizon values were weighted according to their relative fraction of total soil and saprolite depth, whereas soil depth, clay content, organic carbon content, and sodium mass flux data were calculated as profile sums. Logit transformations were performed on all values prior to regression modeling (Hengl et al., 2004):

$$z^+ = \frac{z - z_{min}}{z_{max} - z_{min}}; \ z_{min} < z < z_{max}, \tag{7}$$

where the raw values z were standardized to the min and max of the raw data z^+. Logit transformations normalize distributions of non-normal data sets and ensure that predicted soil values remained within the range of the target variable values, making prediction more efficient and statistically robust. The z^+ values were then standardized from 0 to 1 (Hengl et al., 2004):

$$z^{++} = \ln\left(\frac{z^+}{1 - z^+}\right); \ 0 < z^+ < 1, \tag{8}$$

where z^{++} are the logit values used in soil property prediction models. After the prediction models were constructed and interpolated residuals were added to the regression predictions (see below) using logit z^{++} values, predicted z^{++} values were back-transformed (Hengl et al., 2004):

$$z_{bt} = \left(\frac{e^{z^{++}}}{1 + e^{z^{++}}}\right)(z_{max} - z_{min}) + z_{min}, \tag{9}$$

Table 3. Summary of key parent material mineral and geochemical parameters.

	Granite ($n = 4$)	Amphibolite ($n = 2$)	Quartzite ($n = 4$)
Quartz (%)	41 ± 5	2 ± 2	79 ± 26
Hornblende (%)	0 ± 0	72 ± 14	0 ± 0
[Ca+Mg+Fe] (%)	0.65 ± 0.33	35 ± 10	0.67 ± 0.04
Ti : Zr	3.13 ± 0.29	118 ± 30	2.33 ± 0.97

for displaying results and to ensure that back-transformed values, z_{bt}, represent the scale of the original data, z.

Two sets of multiple linear regressions were developed using either the z-scored environmental covariate layers or the principal components (PCs) resulting from the final iPCA iteration as the independent variables, and the logit-transformed soil data as the dependent variables. The total number of data points for developing regression models was $n = 24$. This represents a relatively limited number of data points for model development. We set a minimum cutoff of 17 degrees of freedom for each model, with the goal of maximizing model degrees of freedom and minimizing the number of independent variables used in each model. We further limited interaction terms among independent variables to a maximum of two-way interactions to reduce model complexity and the risk of over-fitting the data. For regression models developed using the environmental covariate data we calculated the variance inflation factor (VIF) for each model component to determine whether any possible independent variable colinearity exerted undue influence on model fit and parameter estimation. A VIF value of > 5 indicates colinearity (Freund et al., 2003); all measured VIF values were < 3, indicating colinearity did not affect model fit and parameter estimation (Table 6).

Multivariate stepwise linear regression models were constructed using JMP (JMP v11, SAS Institute Inc., Cary, NC, 2012). Regression modeling was performed using a combination of forward and backward stepwise least-squares regression including all independent variables and their interaction terms. Model and independent variable parameter selection was based on a combination of metrics including maximizing the k-fold cross validation R^2, where k was set to 24 and therefore equivalent to a leave-one-out cross validation (LOOCV). We also calculated the prediction residual sum of squares (PRESS) statistic (Freund et al., 2003): PRESS $= \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$, where y_i is the value at the ith location, and $\hat{y}_i$ is the value predicted for the ith location when removed from the model, providing a measure of model fit based on LOOCV. From this parameter, we calculated the PRESS root-mean-square error (PRMSE) as

$$\text{PRMSE} = \sqrt[2]{\text{PRESS} / n}, \tag{10}$$

which may be compared to model RMSE; large differences between the two indicates the model is sensitive to the presence/absence of specific observations. In addition to calculating regression model R^2 and P value, we also calculated the PRESS R^2 or

$$R^2_{\text{Predicted}} = 1 - \text{PRESS} / \text{SST}, \tag{11}$$

where SST is total sums of squares. The $R^2_{\text{Predicted}}$ is similar to model R^2 in that it ranges from 0 to 1, with values of 1 indicating that the model explains 100 % of the variance of the predicted data. These values were calculated for all models, and the final model selection for soil property prediction was based on comparison of these values among regression models using the set of z-scored covariate data or the PC data.

During model selection we also calculated Cook's D for each location to determine whether any point was exerting undue influence on model parameters (Freund et al., 2003). A Cook's D value > 1 indicates a specific point exerts undue influence on model fit. Only one soil variable, Na mass flux [kg Na m^{-2}], exhibited locations with Cook's D values > 1 for regression models developed using either the environmental covariates or the PCs. Specifically, one location, pit 17, underlain by quartzite, exerted undue influence on model fit. This point was excluded from model development and parameter estimation for the Na mass flux model.

Finally, we calculated the normalized mean square error (NMSE) to compare the relative goodness of model fits among models with different units and to deal with comparing models developed using logit-transformed variables as dependent variables:

$$\text{NMSE} = \frac{\frac{1}{n} \sum_{i=1}^{n} (p_i - o_i)^2}{s^2}, \tag{12}$$

where n is the number of observations, p_i is the predicted value at location i, o_i is the observed value at location i, and s^2 is the variance of the observed samples (Park and Vlek, 2002; Hengl et al., 2004).

2.10 Residual spatial analysis

Logit-transformed regression model residuals were spatially interpolated using an optimized IDW algorithm, IDW$_{\text{opt}}$, in ArcGIS 10.0 (ESRI, Redlands, CA) following Molotch et al. (2005). Due to the small size of the study area and only having 24 sample points, kriging was not deemed an appropriate technique for interpolating residuals. The IDW$_{\text{opt}}$ routine is a conservative interpolation, where predicted values are not over- or underestimated relative to the maximum or minimum values of the data set, generating a smooth surface with peaks and valleys between data points. The IDW$_{\text{opt}}$ option allows for the optimization of the power, search radius, and maximum number of neighbors used in the weighted distance calculation using LOOCV with the goal of minimizing the RMSE of the residual prediction. The IDW$_{\text{opt}}$ was

performed using a search radius ranging from 75 to 250 m and varying number of neighbors and final parameter selection based on the combination that best minimized RMSE of LOOCV and provided an error mean that approached zero. The interpolated values were then added back to the values predicted using the regression model and the resulting sum back-transformed from logit units to the units of the predicted dependent variable for "final" maps of predicted soil properties.

3 Results

3.1 General soil properties

The observed soil profiles exhibited consistent taxonomic and morphologic soil properties, despite differing parent material and a relatively broad range of soil physicochemical properties (Table 1). Taxonomically, all MG soils were classified as Ustorthents or Haplustolls, with umbric or mollic epipedons and no subsurface diagnostic horizons (Holleran, 2013). The average soil profile included a surface Oi horizon consisting of pine litter and partially decomposed organic materials and several A horizons darkened by organic matter accumulation that transition to multiple AC and C horizons grading towards saprolite and saprock (Fig. 2).

Soil depth ranged from 40 to 130 cm, with an average of 72 ± 25 cm. Soils averaged 2.1 ± 1.6 % organic C with enrichment of C in surface horizons. Surface A horizons averaged organic C content of 4.3 ± 1.6 %, which decreased roughly linearly with soil depth to organic C values averaging 0.9 ± 0.4 % for C and Cr horizons. Soil carbon stocks averaged 8.0 ± 3.4 kg m^{-2}, with the maximum and minimum values of 16.8 and 4.2 kg m^{-2} located in convergent and divergent portions of the landscape, respectively. Soil pH in general was moderately acidic, with KCl pH values averaging 4.9 ± 1.0, ranging from acidic at KCl pH of 3.0 to a near-neutral pH of 7.0 (Table 1). Higher pH values were generally associated with the organic-matter-rich surface horizons and the convergent portions of the landscape, with the most acidic values in saprolite and saprock horizons. The majority of MG soils were texturally classified as sandy loams, with some loams and sandy clay loams. Clay and sand weight percentages averaged 11 ± 4 and 58 ± 8 %, respectively. Despite substantial variation in parent material composition, none of these soil physicochemical properties exhibited significant variation among the three parent materials (Table 2; all P values > 0.05). Rock fragment abundance increased with depth, from a range of 2–10 % in surface A horizons to > 50 % in the deepest C and Cr horizons. Rock fragment content was significantly greater in the quartzite soil horizons ($P < 0.001$, Table 2).

3.2 Elemental composition and mineral assemblage

Elemental and mineral composition of parent material samples from each profile indicated the catchment was dominantly underlain by granitic materials with small intrusions of quartzite and a mafic, hornblende-enriched amphibolite (Fig. 3). There was large variation in quartz content, ranging from ~ 2 % by weight in the amphibolite to nearly 80 % in the quartzite (Table 3). The amphibolite was dominantly hornblende, averaging > 70 % hornblende by weight, with ferropargasite as the other dominant amphibole. The mineralogical differences were also expressed in variation in rock geochemical composition, with much greater values for [Ca + Mg + Fe] and Ti : Zr in the amphibolite (Table 3).

The elemental and mineral composition of the MG soils was relatively heterogeneous, following the variation in parent material (Table 2). The 21 profiles on granitic and quartzitic bedrock were relatively uniform in chemical composition, with a limited range of values for each element. In contrast, profiles located on the amphibolite exhibited significant differences in elemental and mineral composition relative to the other soils. Specifically, the granite and quartzite soils exhibited Si concentrations that averaged 32 ± 1.3 g Si kg^{-1}, relative to the mafic soils that averaged 24 ± 2.9 g Si kg^{-1} ($P < 0.0001$). The mafic soils were enriched in Ca, Mg, and Fe, with [Ca + Mg + Fe] values that averaged 17 ± 4 %, compared to the quartzite and granite soils that averaged [Ca + Mg + Fe] of 4.2 ± 0.9 and 2.9 ± 2.8 %, respectively ($P < 0.001$). Additionally, the soils on amphibolite exhibited Ti : Zr of 159 ± 101 relative to granite and quartzite soils where values were < 14. Quantitative XRD analyses indicated hornblende content of 31 ± 16 % in the amphibolite soils, relative to the granite and quartzite soils that averaged < 1.0 % ($P < 0.001$). The quartzite soils exhibited a greater concentration of quartz, with an average of 48 ± 13 % compared to the granite soils that averaged 36 ± 5.7 %, and the mafic soils with average quartz concentration of 18 ± 7 % ($P < 0.001$).

Sodium mass transfer, τ, and total Na mass flux values, $m_{\text{flux,total}}$, indicated Na loss in all soil profiles (Table 1), with the exception of the two profiles on quartzite, pit 17 and pit 1. All the granite and amphibolite soils exhibited similar Na depletion profiles, with τ_{Na} values averages of -0.19 ± 0.07 and -0.19 ± 0.12, respectively (Table 2). This equates to an average of ~ 20 % loss of Na due to chemical weathering. In contrast, the quartzite profiles averaged > 80 % enrichment in Na. Analysis of parent rock and soil geochemistry in the quartzite profiles indicated that much of the soil material in the quartzite profiles exhibited a geochemical signature more similar to that of soil materials derived from granite rather than the underlying quartzite parent material (Fig. 4); thus Na mass transfer and mass flux calculations for these locations were strongly influenced by the discrepancy in soil and parent material geochemistry used in Eq. (3). The quartzite parent material contained little Na, on the order of 0.4 %,
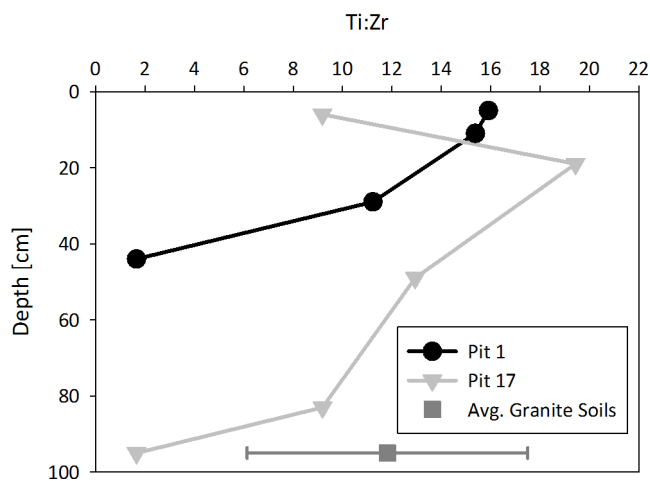
Figure 4. Depth profiles of fine-earth fraction Ti : Zr in the profiles location on quartzite relative to average Ti : Zr of soils derived from granite in the study catchment.

compared to an average of 1.4 % in the soils putatively derived from quartzite. Greater Na concentration of Na in the quartzite profile and the variation in Ti : Zr indicates the addition of granitic and/or mafic materials via colluvial transport from neighboring soils. This supposition is supported in that the relative Na concentration in the quartzite soils was equivalent to the Na concentrations noted in the surrounding granite and mafic soils.

3.3 Principal component analysis and regression modeling

The iPCA and factor loading analyses indicated that six environmental covariates accounted for 95 % of the variance in the landscape (Table 4). The selected environmental covariates included solar radiation (Sol_Rad), SAGA wetness index (Wet_Ind), slope, NDVI, NAIP B3 : B2, and numerically modeled soil depth (Mod_Depth) from Pelletier and Rasmussen (2009), with the largest amount of variance attributed to slope, solar radiation, SAGA wetness index, and B3 : B2 (Table 4).

The final set of PCs from the final iPCA were strongly correlated with distinct sets of environmental covariates (Table 5). Specifically, PC1 and PC2, which accounted for 56 % of cumulative variance (Table 4), were strongly correlated to topographic variables where PC1 was most strongly related to modeled soil depth, slope, and wetness index, and PC2 was strongly correlated with wetness index, modeled soil depth, and solar radiation. In contrast, PC3 and PC4 were dominantly weighted towards NDVI and B3 : B2, respectively. PC5 exhibited significant correlation with all of the environmental covariates, whereas PC6 was only significantly correlated to wetness index and slope. Environmental covariates exhibited relatively few significant correlations

with each other, with the strongest correlation of $r = 0.73$ between modeled soil depth and wetness index.

The final set of soil prediction models for the soil variables of KCl pH, clay [%], τ_{Na}, soil depth [cm], carbon and clay stocks [kg m^{-2}], and Na mass flux [kg m^{-2}] varied in terms of prediction parameters, the relative strength of model fit, and model predictive capability (Table 6). The variables of KCl pH, τ_{Na}, and Na mass flux exhibited the best prediction models using PCs. Specifically, prediction parameters for KCl pH included PC2 and PC4 that were strongly correlated to topographic parameters and vegetation indices (Table 5). Similarly, prediction parameters for Na mass flux included topography-related components PC2 and PC6, and the vegetation-related component PC3. The prediction model for τ_{Na} was comprised of topography-related components of PC1, PC2, and PC6. In contrast, the prediction models for clay %, soil depth, and carbon and clay stocks all exhibited the strongest fits with environmental covariate layers related to topography including wetness index, solar radiation, and modeled soil depth (note that modeled soil depth was excluded from models developed for measured soil depth) (Table 5).

All of the prediction models, except for clay %, included one two-way interaction term and all models had 19–20 degrees of freedom for the model error term. The exception was the model for Na mass flux that only had 17 error term degrees of freedom. As indicated previously, the Na mass flux models were developed excluding pit 17 as Cook's D values indicated this location exerted a large effect on model parameter estimation and model fit. All variance inflation factors were < 5, indicating that colinearity among independent variables was not an issue, even for those models that included both modeled soil depth and wetness index that were significantly correlated.

The best regression models based on model R^2, NMSE, and R_p^2 were the models for clay and carbon stocks, KCl pH, and τ_{Na} (Table 6). These models exhibited the lowest relative NMSE values, ranging from 0.32 to 0.44, and explained between 56 and 68 % of variance in the dependent variable values, and between 42 and 52 % of variance in predicted values based on LOOCV. In comparison, while the models developed for clay %, soil depth, and Na mass flux exhibited a reasonable fit to all dependent variables used in the model, explaining 41 to 46 % of variable variation, they explained a relatively small fraction, only 16 to 21 %, of the variance for predicted variables using LOOCV. The worst model fit was for Na mass flux, with the greatest NMSE of 0.87 and model P value of only 0.028. All models exhibited near-equivalent values for RMSE and PRMSE.

3.4 Residual interpolation

Logit-transformed regression model residuals were interpolated using IDW$_{opt}$. The IDW$_{opt}$ yielded variable power values, with most values equal to 1 and only IDW models for

Table 4. Eigenvectors, variance, and cumulative variance for final PCA.

	PC1	PC2	PC3	PC4	PC5	PC6
Sol_Rad	0.531	0.350	0.094	0.174	−0.596	0.449
Wet_Ind	−0.388	0.611	−0.070	0.178	0.433	0.502
Slope	−0.489	−0.363	0.036	−0.447	−0.366	0.542
Mod_Depth	−0.488	0.479	−0.023	−0.019	−0.535	−0.495
NDVI	−0.187	−0.127	0.901	0.371	0.007	−0.003
B3 : B2	−0.235	−0.356	−0.416	0.774	−0.194	0.086
Eigenvalues	2.03	1.31	0.96	0.90	0.46	0.34
Var. (%)	34	22	16	15	8	6
Cum. var. (%)	34	56	72	87	94	100

τ_{Na} and carbon stock with power values roughly equivalent to 2 (Table 7). The LOOCV mean errors approached zero for all variables. Further, we observed no consistent correlation of model residuals to landscape position (Fig. 5). The IDW$_{opt}$ results indicated the coupled iPCA and cLHS generated randomly distributed model errors. Interpolated model residuals were then added back to values predicted using the regression models.

3.5 Predicted soil properties

The final sets of values from the regression modeling and residual interpolation were back-transformed to the relative scale of each soil variable and expressed as maps of soil physicochemical properties (Fig. 6). The final maps indicated clear spatial variation in soil properties that were largely related to topographic variation, with a general trend of greater depth, clay and carbon content, and Na loss in the convergent drainages. Final predicted value summary statistics were comparable to those of the locations used to develop the prediction models (Table 8). Measured and predicted values generally exhibited similar median and quartile values and summary statistics. The largest deviations between measured and predicted values were observed for τ_{Na} and Na mass flux. This deviation was driven by the area of the catchment underlain by quartzite parent material (Fig. 3) that exhibited positive values for τ_{Na} and Na mass flux (Fig. 5) due to the variation in chemistry between the quartzite rock and the overlying soil material (Fig. 4).

4 Discussion

4.1 Factors controlling soil property spatial variation

The observed and modeled data indicated that water and energy, represented by the topographically derived variables of wetness index and solar radiation, were central factors for predicting soil properties across the catchment. Solar radiation and wetness index were included directly as environmental covariates or principal components in all of the regression models developed for soil property prediction

(Table 6). The significance of solar radiation and wetness index is also apparent in the final PCA, where the dominant eigenvectors of PC1 and PC2, which accounted for a combined 56 % of landscape variability, were strongly correlated to solar radiation and wetness index (Table 4). These findings agree with other studies at this scale of observation that note significant correlations of soil properties to solar radiation, wetness index, and slope (Moore et al., 1993; Gessler et al., 2000; Ashtekar and Owens, 2013). Solar radiation is a primary control on soil water availability, serving as a metric of radiative energy transfer to the soil system that drives soil heating, evaporative transpiration, photosynthesis, and soil water availability (Beaudette and O'Geen, 2009). The degree of soil wetness drives primary production, weathering rates, and biological activity, and has previously been reported as an important topographic index for predicting soil properties at scales ranging from 10^0 to 10^3 ha (Bodaghabadi et al., 2011; Levi and Rasmussen, 2014; Browning and Duniway, 2011).

Remotely sensed NDVI and NAIP bands B3 : B2 characterizing vegetation variability were only included in the models for KCl pH and Na mass flux, and in those cases only indirectly through their contribution to PC3 and PC4. In the final output from the iPCA, vegetation layers only accounted for ∼ 14 % of landscape variance (Table 4). The lack of stronger relationships of soil properties to vegetation was most likely a function of the relative homogeneity of the vegetation assemblage in the observed catchment and the relatively small catchment area, 10^0 ha, encompassed by this study. These results contrast with those of previous DSM studies at larger spatial resolutions spanning areas on the order of 10^4 to 10^6 ha that identified strong relationships between soil variation and remotely sensed vegetation indices (Ballabio et al., 2012; Dobos, 2003).

4.2 Soil property variation with topography

The predicted spatial patterns of soil properties indicated strong correspondence to local topography and landscape structure (Fig. 5); however, this was expected given the
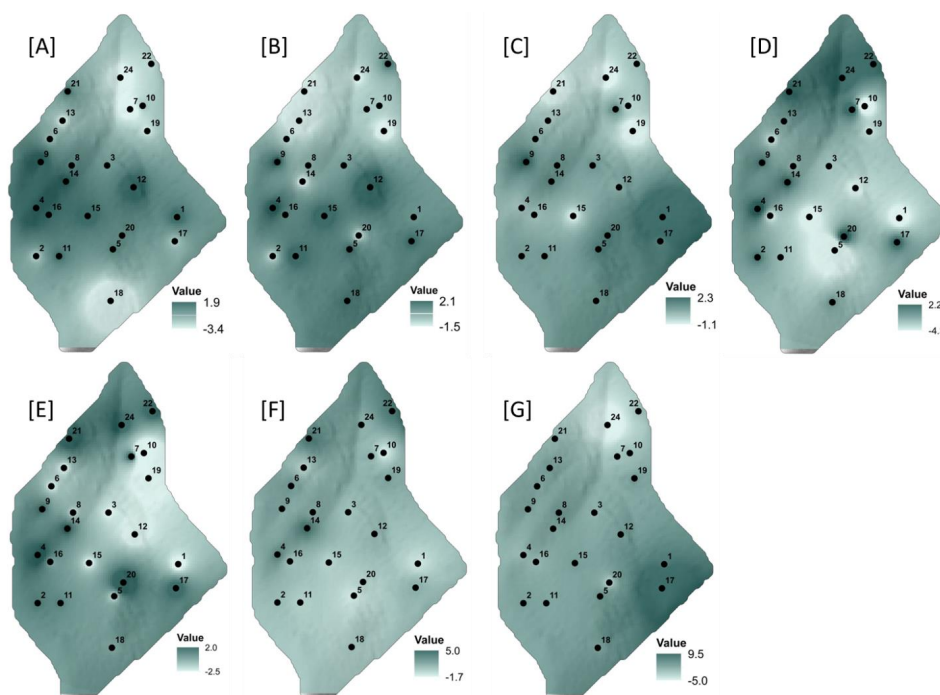
Figure 5. Spatially interpolated logit-transformed regression model residuals for **(a)** clay percent, **(b)** KCl pH, **(c)** τ_{Na}, **(d)** soil depth, **(e)** clay stocks, **(f)** carbon stocks, and **(g)** Na mass flux. Residuals were interpolated using an optimized inverse distance weighting.

Figure 6. Final maps of predicted soil properties including **(a)** clay percent, **(b)** KCl pH, **(c)** τ_{Na}, **(d)** soil depth, **(e)** clay stocks, **(f)** carbon stocks, and **(g)** Na mass flux for the study catchment. Final maps are the back-transformed summation of logit-transformed regression estimations and interpolated residuals.

Table 5. Correlation matrix of environmental covariates and principal components.

	Mod_Depth	Wet_Ind	Sol_Rad	Slope	NDVI	B3:B2	PC1	PC2	PC3	PC4	PC5	PC6	KCl-pH	Clay (%)	τ_{Na}	Soil depth (cm)	Carbon (kg m^{-2})	Clay (kg m^{-2})	Na mass flux (kg m^{-2})
Mod_Depth	1.00																		
Wet_Ind	**0.73*****	1.00																	
Sol_Rad	−0.13	0.09	1.00																
Slope	0.28	−0.03	**−0.59****	1.00															
NDVI	−0.01	−0.10	−0.19	0.24	1.00														
B3:B2	0.07	0.17	−0.04	−0.07	0.17	1.00													
PC1	**0.76******	**0.54******	**−0.58****	**0.67******	0.35	0.31	1.00												
PC2	**0.66*****	**0.77******	**0.41***	−0.32	−0.37	−0.28	0.10	1.00											
PC3	−0.09	−0.20	−0.11	0.23	**0.93******	−0.19	0.17	−0.30	1.00										
PC4	0.02	0.23	0.22	−0.39	0.35	**0.89******	0.12	−0.11	0.05	1.00									
PC5	**0.66*****	0.24	0.16	**0.44***	0.11	0.28	**0.54****	0.19	0.00	0.14	1.00								
PC6	−0.31	0.30	0.01	0.12	0.00	−0.03	−0.02	−0.04	0.01	−0.02	**−0.43***	1.00							
KCl-pH	0.38	0.16	−0.18	**0.52****	0.19	−0.07	**0.46***	0.05	0.19	−0.18	**0.43***	−0.03	1.00						
Clay (%)	0.11	**0.45***	−0.24	−0.03	0.01	0.01	0.24	0.19	−0.03	0.03	−0.32	**0.49***	0.22	1.00					
τ_{Na}	0.02	−0.25	0.19	−0.30	−0.03	−0.16	−0.26	0.09	0.04	−0.03	0.07	**−0.56****	−0.14	−0.27	1.00				
Soil depth (cm)	0.24	0.34	−0.23	0.05	−0.30	−0.10	0.19	0.28	−0.30	−0.18	−0.13	0.16	−0.17	0.16	−0.14	1.00			
Carbon (kg m^{-2})	0.01	0.14	−0.36	0.05	−0.24	0.06	0.13	−0.01	−0.28	−0.06	−0.26	0.19	0.06	**0.32***	**−0.66*****	**0.89******	1.00		
Clay (kg m^{-2})	0.23	**0.45***	−0.26	0.13	−0.19	0.09	0.31	0.21	−0.26	−0.01	−0.11	0.36	0.03	0.37	−0.25	**0.89******	**0.73******	1.00	
Na mass flux (kg m^{-2})	0.03	−0.16	0.32	**−0.44***	−0.01	−0.02	−0.28	0.15	0.02	0.15	0.09	**−0.52****	−0.07	−0.27	**0.62****	−0.38	**−0.43***	**−0.49***	1.00

Pairwise correlation coefficients. Bold values are significant at $P < 0.05$*, $P < 0.01$**, $P < 0.001$***, $P < 0.0001$****.

amount of landscape variance accounted for by topographic variables and the inclusion of these parameters in all of the derived soil prediction models (Table 5). At the scale of the catchment studied here, variation in relief/topography was the primary variable accounting for soil variability, with minimal variation in other soil-forming parameters, including climate, organisms, landscape age, and parent material. The observed significant variation in parent material appeared to exert minimal control on spatial soil property variation and model residuals, further highlighting the dominant role of topographic control on soil variation at this scale. In particular, the geochemical signature of the soils located on the quartzite parent material indicated mixing with adjacent granitic and mafic parent materials, likely via topographically driven processes of colluviation and bioturbation. All modeled soil properties exhibited strong trends related to the convergent portions of the landscape. Specifically, in the convergent areas we observed deeper soils, greater concentration of clay and clay stocks, increased Na depletion and Na mass flux, and greater soil organic C stocks (Fig. 5); much of this variation may be attributed to topographic controls on water and sediment redistribution.

In terms of water redistribution, the convergent areas of the landscape collect both surface and subsurface flow from adjacent water-shedding landscape positions (Beven and Kirkby, 1979). Wetness index values for the observed catchment ranged from ~ 0.9 on summit and ridge landscape positions to over 14 in convergent positions. These data suggest that the convergent positions may receive an amount of water up to an order of magnitude greater than other portions of the landscape. Increased hydrologic inputs facilitate greater physical and chemical weathering owing to a combination of (i) greater water available to serve as both reactant in weathering reactions as well as a transport agent to remove soluble weathering products from the reaction zone (White and Brantley, 1995) and (ii) through the input of surface and subsurface waters enriched in dissolved organic matter, including organic acids and ligands, sourced from adjacent landscape positions (e.g., Vazquez-Ortega et al., 2014). We observed maximum Na depletion and Na mass flux, on the order of 30 % Na and -900 kg Na m^{-2}, in the convergent areas relative to depletion and mass flux values of ~ 5–10 % and -150 to -250 kg Na m^{-2} in ridge and summit positions (Fig. 5). Greater Na loss may be attributed to the greater potential for chemical weathering, and likely higher weathering rates of plagioclase feldspar and amphibole in these portions of the catchment. Additionally, an unknown portion of the greater Na depletion and mass loss in the convergent areas may be attributed to the accumulation of previously weathered materials transported to the convergent areas from adjacent portions of the landscape via colluviation and other mass transport processes (Yoo et al., 2007; Yoo and Mudd, 2008).

Table 6. Summary statistics of regression models for predicting logit-transformed soil physicochemical properties.

Dependent variable	Independent variables	Coefficients	SD error	P value	VIF	df_E	P value	R^2	RMSE++	NMSE	PRMSE++	R_P^2
		Model components				Model summary statistics						
KCl-pH++	Intercept	−0.68	0.28	0.02	–	20	0.001	0.56	1.16	0.44	1.23	0.46
	PC2	0.08	0.28	0.77	1.01							
	PC4	−0.99	0.29	0.00	1.32							
	(PC2+0.462) · (PC4+0.153)	−1.85	0.38	< 0.0001	1.31							
Clay (%)++	Intercept	0.17	0.40	0.67	–	20	0.006	0.46	1.48	0.55	1.71	0.21
	Mod_Depth	−1.38	0.55	0.02	2.34							
	Wet_Ind	2.75	0.73	0.00	2.32							
	Sol_Rad	−1.41	0.60	0.03	1.09							
τ_{Na}++	Intercept	−0.93	0.18	< 0.0001	–	19	0.002	0.57	0.75	0.43	0.88	0.34
	Mod_Depth	1.14	0.30	0.00	2.81							
	Wet_Ind	−1.55	0.39	0.00	2.54							
	Slope	−0.92	0.24	0.00	1.32							
	(Mod_Depth+0.071) · (slope−0.313)	−0.71	0.25	0.01	1.07							
Soil depth (cm)++	Intercept	−0.96	0.59	0.12	–	20	0.012	0.41	2.16	0.59	2.44	0.19
	Wet_Ind	1.66	0.70	0.03	1.02							
	Sol_Rad	−0.96	0.85	0.27	1.02							
	(Wet_Ind+0.191) · (Sol_Rad+0.442)	−3.95	1.39	0.01	1.03							
Carbon (kg m^{-2})++	Intercept	−1.26	0.42	0.01	–	20	0.001	0.56	1.54	0.44	1.57	0.50
	Wet_Ind	0.81	0.50	0.12	1.02							
	Sol_Rad	−1.17	0.60	0.07	1.02							
	(Wet_Ind+0.191) · (Sol_Rad+0.442)	−4.22	0.99	0.00	1.03							
Clay (kg m^{-2})++	Intercept	−0.84	0.41	0.05	–	19	0.000	0.68	1.48	0.32	1.72	0.52
	Mod_Depth	−1.19	0.55	0.04	2.35							
	Wet_Ind	3.16	0.73	0.00	2.35							
	Sol_Rad	−1.38	0.60	0.03	1.10							
	(Wet_Ind+0.191) · (Sol_Rad+0.442)	−4.30	0.95	0.00	1.03							
Na mass flux (kg m^{-2})++ *excludes pit 17	Intercept	−1.31	0.17	< 0.0001	–	17	0.028	0.46	0.65	0.87	0.76	0.16
	PC2	0.20	0.17	0.27	1.18							
	PC3	0.15	0.16	0.36	1.10							
	PC6	−1.56	0.42	0.00	1.58							
	(PC2+0.525) · (PC6−0.018)	0.88	0.32	0.01	1.68							

Table 7. IDW of residuals – optimized modeling parameters.

Variable	Power	Search radius	# of sectors	Mean++	RMSE++
Clay %	1.00	250	1	0.07	1.39
KCl pH	1.00	250	1	0.00	1.10
τNa	1.62	250	1	0.00	0.64
Depth	1.00	250	8	−0.05	2.26
Carbon	1.99	150	1	−0.12	1.29
Na−flux	1.00	250	1	−0.06	2.85
Clay	1.00	250	1	−0.03	1.46

IDW performed with a minimum of 10 neighbors and maximum of 24. Circular search radius.

Convergent portions of the landscape typically exhibit deeper soils than adjacent ridges as a result of the localized accumulation of soil and sediment produced in upslope positions (Dietrich et al., 2003), consistent with the patterns observed here (Fig. 5). Based on observed soil morphology in convergent areas, these soils represent cumulic profiles with a series of deep A and AC horizons that exhibit some particle size stratification, suggesting individual packets of sediment accumulated from episodic physical deposition events (Dickinson, 1992). Soils located on summit/ridgetop positions were generally shallow and rocky, with morphology indicative of in situ soil production, whereas hillslope soils contained colluvial surface horizons darkened by organic matter over soil derived from bedrock weathering in place

(Fig. 2). Soil production rates, i.e., the physical conversion of bedrock to potentially mobile soil material, typically present an inverse relation with soil depth in that rates decrease exponentially with increasing soil depth (Heimsath et al., 1997). For the catchment observed here, this translates into localized variation in soil production, with the lowest potential rates in convergent areas where soil production rates would be limited due to the deep accumulation of sediment from adjacent landscape positions.

The convergent areas also contained a greater magnitude of clay and carbon stocks (Fig. 5e, f), driven in part by the greater soil depth observed in these areas. Clay percentages were also substantially greater in convergent areas, with up to 14 % clay by weight on average, relative to summit and

Table 8. Comparison of measured and modeled soil property summary statistics.

Property	Type	Quartiles			Mean	SD	CV	Skewness	Kurtosis
		25 %	50 %	75 %					
KCl pH	Measured	5.20	4.50	4.05	4.66	0.69	15	0.8	−0.7
	Modeled	4.97	4.59	4.30	4.66	0.58	12	0.5	−0.3
Clay (%)	Measured	12.33	9.61	8.48	10.12	2.35	23	0.1	−1.0
	Modeled	11.56	9.49	7.29	9.54	2.59	27	0.1	−1.1
τ_{Na}	Measured	−0.13	−0.17	−0.22	−0.12	0.22	−182	2.7	6.1
	Modeled	0.03	−0.08	−0.15	−0.04	0.18	−473	1.3	2.0
Soil depth (cm)	Measured	90.0	69.5	51.0	71.5	24.3	34	0.6	−0.4
	Modeled	91.9	62.8	51.6	73.2	28.7	39	0.8	−0.7
Carbon (kg m^{-2})	Measured	8.68	7.52	5.34	8.03	3.33	41	1.3	1.3
	Modeled	11.80	7.80	5.89	8.99	3.99	44	0.7	−0.8
Clay (kg m^{-2})	Measured	72.3	46.8	27.7	55.4	33.6	61	1.0	0.1
	Modeled	88.1	46.9	26.7	60.7	42.4	70	0.7	−0.8
Na mass flux (kg m^{-2})	Measured	−148	−249	−575	−269	516	−192	2.6	9.1
	Modeled	−252	−420	−559	−315	475	−151	2.1	5.1

n for measured is 24; *n* for modeled is 20 910.

ridgetop positions that contained only ∼ 5–6 % clay on average (Fig. 5a). Thus greater clay stocks in the convergent areas were not only a function of deeper soils but also increased concentration of fine-grained materials. The enrichment of clay in these areas is likely due to a combination of sediment transport, particularly surface wash processes that preferentially transport fine-grained materials (Pelletier, 2008), and enhanced weathering and in situ transformation of primary minerals to secondary phases (Maher, 2010).

The largest carbon stocks, on the order of 17 kg m^{-2}, were observed in the convergent landscape positions (Fig. 5f). Topographic variation in C stocks may be related to localized variation in primary production and C inputs, greater frequency of soil saturation and anaerobic conditions, and accumulation of C-rich surface material from adjacent landscape positions via physical transport (Berhe and Kleber, 2013; Harden et al., 2008). We did not observe a significant difference in canopy cover and vegetation density in the convergent areas relative to other portions of the catchment, so local variation in primary production likely does not account for increased C stocks. The convergent areas do tend to have longer and more frequent occurrence of saturated conditions (Heidbuchel et al., 2013) such that localized and temporally variably saturation and water table rise may limit aerobic decomposition of C inputs in these areas. However, in contrast to clay percent, the average profile percent C did not exhibit any significant relationship to topography or other environmental covariates that would be expected if localized chemical processes were enhancing C accumulation. Thus, given the cumulic nature of the convergent area soil profiles, and

the lack of topographic variation in C percent, enhanced C storage in the convergent areas is most likely the product of physical transport and accumulation of C-rich sediments from upslope positions.

5 Conclusions

Digital soil mapping methods were applied to model soil property spatial variation in a 6 ha forested catchment in southern Arizona. The process included use of an iterative data reduction technique to isolate those environmental covariates that accounted for the greater part of landscape variance and that were then coupled with a conditioned Latin hypercube sample design focused on capturing the overall numeric and spatial distribution of the selected environmental covariates. The key conclusions drawn from this work include the following:

- The coupled iPCA and cLHS routine produced a sampling scheme that captured much of the soil variability in the catchment, demonstrating the utility of this approach for establishing robust sample designs for soil and geomorphic studies.

- Simple linear regression models and IDW interpolation of residuals captured the variance in measured soil properties and predicted spatial correlation of soil properties to landscape structure.

- For the catchment-scale observation studied here, topographic covariates explained more variation in soil

properties than reflectance variables, a function of the relatively small scale of observation and uniform vegetation cover in this forested system.

– The modeled spatial distribution of soil properties provides a framework for interpreting catchment-scale variation in critical zone process and evolution.

Future work will focus on coupling results from this coupled empirical–statistical approach to output from mechanistic, process-based numerical models of critical process and evolution.

References

Ashtekar, J. M. and Owens, P. R.: Remembering Knowledge: An Expert Knowledge Based Approach to Digital Soil Mapping, Soil Horizons, 54, doi:10.2136/sh13-01-0007, 2013.

Ballabio, C., Fava, F., and Rosenmund, A.: A plant ecology approach to digital soil mapping, improving the prediction of soil organic carbon content in alpine grasslands, Geoderma, 187, 102–116, doi:10.1016/j.geoderma.2012.04.002, 2012.

Beaudette, D. E. and O'Geen, A. I.: Quantifying the Aspect Effect: An Application of Solar Radiation Modeling for Soil Survey (vol 73, pg 1345, 2009), Soil Sci. Soc. Am. J., 73, 1755–1755, doi:10.2136/sssaj2008.0229er, 2009.

Ben-Dor, E.: Quantitative remote sensing of soil properties, Adv. Agronomy, 75, 173–243, doi:10.1016/S0065-2113(02)75005-0, 2002.

Ben-Dor, E., Patkin, K., Banin, A., and Karnieli, A.: Mapping of several soil properties using DAIS-7915 hyperspectral scanner data – a case study over clayey soils in Israel, Int. J. Remote Sens., 23, 1043–1062, doi:10.1080/01431160010006962, 2002.

Ben-Dor, E., Goldlshleger, N., Benyamini, Y., Agassi, M., and Blumberg, D. G.: The spectral reflectance properties of soil structural crusts in the 1.2- to 2.5-mu m spectral region, Soil Sci. Soc. Am. J., 67, 289–299, 2003.

Berhe, A. A. and Kleber, M.: Erosion, deposition, and the persistence of soil organic matter: mechanistic considerations and problems with terminology, Earth Surf. Process. Landf., 38, 908–912, doi:10.1002/Esp.3408, 2013.

Beven, K. J. and Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology/Un modèle à base physique de zone d'appel variable de l'hydrologie du bassin versant, Hydrol. Sci. Bull., 24, 43–69, doi:10.1080/02626667909491834, 1979.

Blake, G. R. and Hartge, K. H.: Bulk Density, in: Methods of soil analysis. Part 1- Physical and Mineralogical Methods, edited by: Klute, A., ASA and SSSA, Madison, WI, 1986.

Bodaghabadi, M. B., Salehi, M. H., Martinez-Casasnovas, J. A., Mohammadi, J., Toomanian, N., and Borujeni, I. E.: Using Canonical Correspondence Analysis (CCA) to identify the most important DEM attributes for digital soil mapping applications, Catena, 86, 66–74, doi:10.1016/j.catena.2011.02.009, 2011.

Boehner, J., Koethe, R., Conrad, O., Gross, J., Ringeler, A., and Selige, T.: Soil Regionalisation by Means of Terrain Analysis and Process Parameterisation, in: Soil Classification 2001, edited by: Micheli, E., Nachtergaele, F., and Montanarella, L., European Soil Bureau, Research Report No. 7, EUR 20398 EN, Luxembourg, 213–222, 2002.

Brantley, S. L. and White, A. F.: Approaches to Modeling Weathered Regolith, in: Thermodynamics and Kinetics of Water-Rock Interaction, edited by: Oelkers, E. H. and Schott, J., Rev. Mineralogy Geochem., 70, 435–484, 2009.

Brimhall, G. H. and Dietrich, W. E.: Constitutive Mass Balance Relations between Chemical-Composition, Volume, Density, Porosity, and Strain in Metasomatic Hydrochemical Systems – Results on Weathering and Pedogenesis, Geochim. Cosmochim. Acta, 51, 567–587, 1987.

Browning, D. M. and Duniway, M. C.: Digital soil mapping in the absence of field training data: A case study using terrain attributes and semiautomated soil signature derivation to distinguish ecological potential, Appl. Environ. Soil Sci., 2011, 421904, doi:10.1155/2011/421904, 2011.

Brus, D. J. and Heuvelink, G. B. M.: Optimization of sample patterns for universal kriging of environmental variables, Geoderma, 138, 86–95, doi:10.1016/j.geoderma.2006.10.016, 2007.

Buchanan, S., Triantafilis, J., Odeh, I. O. A., and Subansinghe, R.: Digital soil mapping of compositional particle-size fractions using proximal and remotely sensed ancillary data, Geophysics, 77, Wb201–Wb211, doi:10.1190/Geo2012-0053.1, 2012.

Burt, R.: Soil Survey Laboratory Methods Manual, Soil Survey Investigations Report No. 42, Version 4.0, United States Department of Agriculture, Natural Resources Conservation Service, National Soil Survey Center, Lincoln, NE, 2004.

Chadwick, O. A., Brimhall, G. H., and Hendricks, D. M.: From a black to a gray box – a mass balance interpretation of pedogenesis, Geomorphology, 3, 369–390, 1990.

Conrad, O.: SAGA – Program Structure and Current State of Implementation, in: SAGA – Analysis and Modeling Applications, edited by: Böhner, J., McCloy, K. R., and Strobl, J., Verlag Erich Goltze GmbH, 39–52, 2006.

Dematte, J. A. M. and Nanni, M. R.: Weathering sequence of soils developed from basalt as evaluated by laboratory (IRIS), airborne (AVIRIS) and orbital (TM) sensors, Int. J. Remote Sens., 24, 4715–4738, 2003.

Dematte, J. A. M., Genu, A. M., Fiorio, P. R., Ortiz, J. L., Mazza, J. A., and Leonardo, H. C. L.: Comparison between soil maps obtained by spectral remote sensing and conventional method, Pesqui Agropecu Bras, 39, 1219–1229, doi:10.1590/S0100-204x2004001200009, 2004.

Dickinson, W. R.: Geologic Map of Catalina Core Complex and San Pedro Trough: Arizona Geological Survey Contributed Map CM-92-C, Arizona Geological Survey, 1992.

Dietrich, W., Bellugi, D. G., Sklar, L., Stock, J. D., Heimsath, A. M., and Roering, J. J.: Geomorphic transport laws for predicting landscape form and dynamics, in: Prediction in geomorphology,

edited by: Wilcock, P. R., and Iverson, R. M., American Geophysical Union, Washington, DC, 103–132, 2003.

Dobos, E.: The application of remote sensing and terrain modeling to soil characterization, Innovative Soil-Plant Systems for Sustainable Agricultural Practices, 328–348, 2003.

Dokuchaev, V. V.: Selected works of V. V. Dokuchaev, Israel Program for Scientific Translations, available from the U.S. Dept. of Commerce, Clearinghouse for Federal Scientific and Technical Information, Jerusalem, Springfield, Va., 1967.

Eberl, D. D.: User Guide to RockJock – A program for determining quantitative mineralogy from x-ray diffraction data, USGS Open-File Report: 2003-78, 2003.

Egli, M. and Fitze, P.: Formulation of pedologic mass balance based on immobile elements: A revision, Soil Science, 165, 437–443, doi:10.1097/00010694-200005000-00008, 2000.

Eldeiry, A. A. and Garcia, L. A.: Comparison of Ordinary Kriging, Regression Kriging, and Cokriging Techniques to Estimate Soil Salinity Using LANDSAT Images, J. Irrigat. Drain. Eng.-Asce, 136, 355–364, doi:10.1061/(asce)ir.1943-4774.0000208, 2010.

Florinsky, I. V.: Combined analysis of digital terrain models and remotely sensed data in landscape investigations, Prog. Phys. Geog., 22, 33–60, doi:10.1177/030913339802200102, 1998.

Freeman, T. G.: Caclculating catchment-area with divergent flow based on a regular grid, Comput. Geosci., 17, 413–422, doi:10.1016/0098-3004(91)90048-i, 1991.

Freund, R. J., Littell, R. C., and Creighton, L.: Regression Using JMP, SAS Institute Inc., Cary, NC, 2003.

Galvao, L. S., Formaggio, A. R., Couto, E. G., and Roberts, D. A.: Relationships between the mineralogical and chemical composition of tropical soils and topography from hyperspectral remote sensing data, Isprs J. Photogramm, 63, 259–271, doi:10.1016/j.isprsjprs.2007.09.006, 2008.

Gessler, P. E., Chadwick, O. A., Chamran, F., Althouse, L., and Holmes, K.: Modeling soil-landscape and ecosystem properties using terrain attributes, Soil Sci. Soc. Am. J., 64, 2046–2056, 2000.

Goovaerts, P.: Estimation or simulation of soil properties? An optimization problem with conflicting criteria, Geoderma, 97, 165–186, doi:10.1016/S0016-7061(00)00037-9, 2000.

Hansen, C. W., Helton, J. C., and Sallaberry, C. J.: Use of replicated Latin hypercube sampling to estimate sampling variance in uncertainty and sensitivity analysis results for the geologic disposal of radioactive waste, Reliab. Eng. Syst. Safe, 107, 139–148, doi:10.1016/j.ress.2011.12.006, 2012.

Harden, J. W., Berhe, A. A., Torn, M., Harte, J., Liu, S., and Stallard, R. F.: Soil erosion: Data say C sink, Science, 320, 178–179, doi:10.1126/science.320.5873.178, 2008.

Heckman, K. and Rasmussen, C.: Lithologic controls on regolith weathering and mass flux in forested ecosystems of the southwestern USA, Geoderma, 164, 99–111, doi:10.1016/j.geoderma.2011.05.003, 2011.

Heidbuchel, I., Troch, P. A., and Lyon, S. W.: Separating physical and meteorological controls of variable transit times in zero-order catchments, Water Resour. Res., 49, 7644–7657, doi:10.1002/2012wr013149, 2013.

Heimsath, A., Dietrich, W., Nishiizumi, K., and Finkel, R. C.: The soil production function and landscape equilibrium, Nature, 388, 358–361, 1997.

Hengl, T., Heuvelink, G. B. M., and Stein, A.: A generic framework for spatial prediction of soil variables based on regression-kriging, Geoderma, 120, 75–93, 2004.

Hengl, T., Heuvelink, G. B. M., and Rossiter, D. G.: About regression-kriging: From equations to case studies, Comput. Geosci., 33, 1301–1315, doi:10.1016/j.cageo.2007.05.001, 2007a.

Hengl, T., Toomanian, N., Reuter, H. I., and Malakouti, M. J.: Methods to interpolate soil categorical variables from profile observations: Lessons from Iran, Geoderma, 140, 417–427, doi:10.1016/j.geoderma.2007.04.022, 2007b.

Holleran, M. E.: Quantifying catchment scale soil variability in Marshall Gulch, Santa Catalina Mountains Critical Zone Observatory, University of Arizona, University of Arizona, 101 pp., 2013.

Howari, F. M., Goodell, P. C., and Miyamoto, S.: Spectral properties of salt crusts formed on saline soils, J. Environ. Qual., 31, 1453–1461, 2002.

Huete, A. R., Jackson, R. D., and Post, D. F.: Spectral response of a plant canopy with different soil backgkrounds, Remote Sens. Environ., 17, 37–53, doi:10.1016/0034-4257(85)90111-7, 1985.

Irvin, B. J., Ventura, S. J., and Slater, B. K.: Fuzzy and isodata classification of landform elements from digital terrain data in Pleasant Valley, Wisconsin, Geoderma, 77, 137–154, doi:10.1016/S0016-7061(97)00019-0, 1997.

Jackson, M. L.: Soil Chemical Analysis: Advanced Course, 2nd Edn., UW-Madison Libraries Parallel Press, Madison, WI, USA, 2005.

Jenny, H.: Factors of soil formation; a system of quantitative pedology, 1st Edn., McGraw-Hill book company, inc., New York, London, 281 pp., 1941.

Jensen, J. R.: Introductory digital image processing: A remote sensing perspective, 3rd Edn., Prentice Hall, Upper Sadder River, NJ, 2005.

Kempen, B., Brus, D. J., Stoorvogel, J. J., Heuvelink, G. B. M., and de Vries, F.: Efficiency Comparison of Conventional and Digital Soil Mapping for Updating Soil Maps, Soil Sci. Soc. Am. J., 76, 2097–2115, doi:10.2136/sssaj2011.0424, 2012.

Konen, M. E., Jacobs, P. M., Burras, C. L., Talaga, B. J., and Mason, J. A.: Equations for predicting soil organic carbon using loss-on-ignition for north central US soils, Soil Science Society of America Journal, 66, 1878-1881, 2002.

Levi, M. R. and Rasmussen, C.: Covariate selection with iterative principal component analysis for predicting physical soil properties, Geoderma, 219, 46–57, 2014.

Lybrand, R., Rasmussen, C., Jardine, A., Troch, P., and Chorover, J.: The effects of climate and landscape position on chemical denudation and mineral transformation in the Santa Catalina mountain critical zone observatory, Appl. Geochem., 26, S80–S84, 2011.

Maher, K.: The dependence of chemical weathering rates on fluid residence time, Earth Planet. Sci. Lett., 294, 101–110, 2010.

McBratney, A. B., Odeh, I. O. A., Bishop, T. F. A., Dunbar, M. S., and Shatar, T. M.: An overview of pedometric techniques for use in soil survey, Geoderma, 97, 293–327, 2000.

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

McKenzie, N. J. and Ryan, P. J.: Spatial prediction of soil properties using environmental correlation, Geoderma, 89, 67–94, doi:10.1016/s0016-7061(98)00137-2, 1999.

Milligan, G. W.: Ultrametric Hierarchical Clustering Algorithms, Psychometrika, 44, 343–346, doi:10.1007/Bf02294699, 1979.

Minasny, B. and McBratney, A. B.: A conditioned Latin hypercube method for sampling in the presence of ancillary information, Comput. Geosci., 32, 1378–1388, doi:10.1016/j.cageo.2005.12.009, 2006.

Molotch, N. P., Colee, M. T., Bales, R. C., and Dozier, J.: Estimating the spatial distribution of snow water equivalent in an alpine basin using binary regression tree models: the impact of digital elevation data and independent variable selection, Hydrol. Process., 19, 1459–1479, doi:10.1002/Hyp.5586, 2005.

Moore, D. M. and Reynolds, R. C.: X-ray diffraction and the identification and analysis of clay minerals, 2nd Edn., Oxford University Press, Oxford, New York, xviii, 378 pp., 1997.

Moore, I. D., Grayson, R. B., and Ladson, A. R.: Digtial terrain modeling – a review of hydrological, geomorphological and biological applications, Hydrol. Process., 5, 3–30, doi:10.1002/hyp.3360050103, 1991.

Moore, I. D., Gessler, P. E., Nielsen, G. A., and Peterson, G. A.: Soil Attribute Prediction Using Terrain Analysis, Soil Sci. Soc. Am. J., 57, 443–452, 1993.

Mulder, V. L., de Bruin, S., and Schaepman, M. E.: Representing major soil variability at regional scale by constrained Latin Hypercube Sampling of remote sensing data, Int. J. Appl. Earth Obs., 21, 301–310, doi:10.1016/j.jag.2012.07.004, 2013.

Nauman, T.: Digital soil-landscape classification for soil survey using ASTER satellite and digital elevation data in Organ Pipe Cactus National Monument, Arizona, MS Thesis, Univ. of Arizona, 169 pp., 2009.

NRC: Basic Research Opportunities in Earth Sciences, National Research Council, Washington, DC, 2001.

NRC: Landscapes on the Edge: New Horizons for Research on Earth's Surface, The National Academies Press, New York, NY, 163 pp., 2010.

Odeh, I. O. A., McBratney, A. B., and Chittleborough, D. J.: Spatial Prediction of Soil Properties from Landform Attributes Derived from a Digital Elevation Model, Geoderma, 63, 197–214, 1994.

Park, S. J. and Vlek, P. L. G.: Environmental correlation of three-dimensional soil spatial variability: a comparison of three adaptive techniques, Geoderma, 109, 117–140, doi:10.1016/S0016-7061(02)00146-5, 2002.

Pelletier, J. D.: Quantitative Modeling of Earth Surface Processes, Cambridge University Press, New York, 2008.

Pelletier, J. D. and Rasmussen, C.: Geomorphically based predictive mapping of soil thickness in upland watersheds, Water Resour. Res., 45, W09417, doi:10.1029/2008WR007319, 2009.

Saadat, H., Bonnell, R., Sharifi, F., Mehuys, G., Namdar, M., and Ale-Ebrahim, S.: Landform classification from a digital elevation model and satellite imagery, Geomorphology, 100, 453–464, doi:10.1016/j.geomorph.2008.01.011, 2008.

Salisbury, J. W. and Daria, D. M.: Infrared (8-14 Mu-M) Remote-Sensing of Soil Particle-Size, Remote Sens. Environ., 42, 157–165, doi:10.1016/0034-4257(92)90099-6, 1992.

Schoeneberger, P. J., Wysocki, D. A., Benham, E. C., and Broderson, W. D. E.: Field book for describing and sampling soils, Version 2.0, Natural Resources Conservation Service, National Soil Survey Center, Lincoln, NE, 2002.

Scull, P., Okin, G., Chadwick, O. A., and Franklin, J.: A comparison of methods to predict soil surface texture in an alluvial basin, Professional Geographer, 57, 423–437, doi:10.1111/j.0033-0124.2005.00488.x, 2005.

Soil Taxonomy: a basic system of soil classification for making and interpreting soil surveys, Agriculture handbook no. 436, US Dept. of Agriculture, Natural Resources Conservation Service, Washington, DC, Soil Survey Staff, 1999.

Sullivan, D. G., Shaw, J. N., and Rickman, D.: IKONOS imagery to estimate surface soil property variability in two Alabama physiographies, Soil Sci. Soc. Am. J., 69, 1789–1798, doi:10.2136/sssaj2005.0071, 2005.

Van Camp, M. and Walraevens, K.: Pumping test interpretation by combination of Latin hypercube parameter sampling and analytical models, Comput. Geosci., 35, 2065–2073, doi:10.1016/j.cageo.2008.12.011, 2009.

Vasat, R., Heuvelink, G. B. M., and Boruvka, L.: Sampling design optimization for multivariate soil mapping, Geoderma, 155, 147–153, doi:10.1016/j.geoderma.2009.07.005, 2010.

Vazquez-Ortega, A., Hernandez-Ruiz, S., Amistadi, M. K., Rasmussen, C., and Chorover, J.: Fractionation of Dissolved Organic Matter by (Oxy)hydroxide-Coated Sands: Competitive Sorbate Displacement during Reactive Transport, Vadose Zone J., 13, doi:10.2136/vzj2013.10.0179, 2014.

Whittaker, R. H. and Niering, W. A.: Vegetation of Santa Catalina Mountains, Arizona .5. Biomass, Production, and Diversity Along Elevation Gradient, Ecology, 56, 771–790, 1975.

White, A. F. and Brantley, S. L.: Chemical weathering rates of silicate minerals: An overview, in: Chemical Weathering Rates of Silicate Minerals, Rev. Mineral., 31, 1–22, 1995.

Wilson, J. P. and Gallant, J. C.: Terrain analysis: Principles and applications, John Wiley and Sons, New York, 479 pp., 2000.

Yoo, K. and Mudd, S. M.: Discrepancy between mineral residence time and soil age: Implications for the interpretation of chemical weathering rates, Geology, 36, 35–38, 2008.

Yoo, K., Amundson, R., Heimsath, A. M., Dietrich, W. E., and Brimhall, G. H.: Integration of geochemical mass balance with sediment transport to calculate rates of soil chemical weathering and transport on hillslopes, J. Geophys. Res.-Earth Surf., 112, F02013, doi:10.1029/2005JF000402, 2007.

Ziadat, F. M.: Analyzing digital terrain attributes to predict soil attributes for a relatively large area, Soil Sci. Soc. Am. J., 69, 1590–1599, doi:10.2136/sssaj2003.0264, 2005.