Biogeochemistry (2005) 74: 173–203
DOI 10.1007/s10533-004-3544-x

Multi-scale variability in tropical soil nutrients following land-cover change

KAREN W. HOLMES[1,2,*], PHAEDON C. KYRIAKIDIS[2], OLIVER A. CHADWICK[2], JOÃO VIANEI SOARES[3] and DAR A. ROBERTS[2]
*[1]School of Earth and Geographical Sciences, The University of Western Australia, 35 Stirling Highway Crawley, WA 6009, Australia; [2]Geography Department, 3611 Ellison Hall, University of California at Santa Barbara, CA 93106, USA; [3]Instituto de Pesquisas Espacias (INPE), Caixa Postal 515, Av. dos Astronautas, 1758 CEP, 12227-010 São José dos Campos, SP, Brazil; *Author for correspondence (e-mail: kholmes@segs.uwa.edu.au; fax: +61-8-6488-1037)*

Received 14 February 2004; accepted in revised form 19 September 2004

Key words: Geostatistics, Land-cover change, Scale, Soil nutrients, Spatial variability, Tropics

Abstract. The effects of land-cover change on soil properties have reached regional proportions in the southwestern Amazon, and can be detected in their patterns at the scale of hundreds of thousands of square kilometers. Spatial analysis of an extensive soil profile database revealed four nested spatial scales at which different patterns in soil properties occur: <3, <10, <68, and >68 km. The shortest scales account for the majority of the spatial variability in the biogeochemically important properties (phosphorus, nitrogen, organic carbon), while the longest scales explain the most for pedologic properties (pH, aluminum, percent clay and sand). The magnitude and sign of correlations among soil properties changed with scale, indicating major shifts in distribution and soil dynamics depending on the scale of observation and analysis. Precipitation, substrate composition, topography, and biological influences vary over large areas, leading to variable distribution of soil properties, and complicating the evaluation of the effect of deforestation on biogeochemical cycling. We modeled the relationships among scale-specific patterns in soil properties and the soil forming state factors to tease out the relative impact of changing land cover. Soil property short-scale variability corresponded with the distribution of land cover and terrain attributes, medium scale variation with geology and state soil map classifications, and long scale with geology and precipitation. The strength of these relationships may be partially attributable to the resolution of the maps used as proxies for the soil forming factors. Land-cover change has already left a discernable imprint on broad-scale soil nutrient patterns, although it is still not the dominant process.

Introduction

Deforestation in the Amazon affects regional and global ecosystem services such as sustainability of agriculture and carbon cycling dynamics (Nepstad et al. 1999; Houghton et al. 2000; Schneider et al. 2000). Various field investigations have shown that tropical forest clearing changes physical and chemical soil properties, but the regional impact of land-cover change is difficult to assess quantitatively because tropical soils are extremely spatially heterogeneous (Richter and Babbar 1991; Eswaran et al. 1992). Like soils

everywhere, tropical soils are formed through the interaction of multiple environmental factors that produce a complex mosaic of biological and chemical properties across Earth's terrestrial surface (Birkeland 1999; Brady and Weil 1999). The long history of landscape development in the terra firme (upland) areas of the Amazon basin and dense vegetation have led to complex spatial distributions of soil properties, now overprinted by recent dramatic changes in land cover (Holmes et al. 2004). These soil property patterns are not easy to map and interpret given the sparse soil profile data typically available.

Previous studies have taken a spatial approach to quantify and understand soil nutrient distributions (e.g. Uehara et al. 1985; Trangmar et al. 1986a, b; Oline and Grant 2002; Powers and Schlesinger 2002; Rodenburg et al. 2003), but that research addressed vastly different environmental conditions and study area extents. Not surprisingly, the results have been contradictory, and have not answered the fundamental question of what scales are appropriate for addressing regional research questions, nor which environmental processes are controlling these distributions and characteristic scales. For example, soil pH has been reported to display the highest variability among soil properties, with spatial autocorrelation up to 1 m (Oline and Grant 2002), and the lowest variability, with correlation up to 750 m (Sun et al. 2003). Likewise, soil phosphorus, generally considered the most limiting element in tropical, highly weathered ecosystems (Vitousek 1984; Dias-Filho et al. 2001), has been documented as spatially autocorrelated up to 639 m (Sun et al. 2003), and to 6300 m (Trangmar et al. 1986b). These results were interpreted as differences in landscape distribution of soil properties, but the discrepancy between the detected ranges of correlation may be an artifact of differences in sample distribution, sample density, and the extent of the field area (Western and Bloschl 1999). In addition, no robust conclusions were drawn concerning the processes responsible for that spatial distribution. In this paper, we address these issues by combining a geostatistical approach for detecting important scales from relationships among the data themselves, rather than imposing an arbitrary scale of analysis, and multivariate regression to investigate which processes are potentially the major forces determining large area soil nutrient distributions.

We used a well-populated regional soil profile database from the southwestern Amazon basin to address the ongoing debate concerning the effects of tropical deforestation for agriculture on regional biogeochemical cycling in the humid tropics, and provide a more robust basis for predicting the potential impacts of the continuing trend of deforestation. The database is georeferenced and contains a sufficiently large number of non-systematically sampled measurements. We applied geostatistical techniques to (1) quantify the magnitude of soil nutrient variability and identify the important scales at which patterns of nutrients emerge over a large region in Amazônia; (2) analyze relationships among soil properties at different spatial scales within the region; and (3) evaluate relationships at different scales among nutrients and the state factors driving soil formation (climate, geology, biota, topographic relief, time, and human activities), with a particular focus on the impact of recent (since the 1970s) forest clearing.

Data and methods

Field area

The Brazilian state of Rondônia covers an area of approximately 240,000 km^2, and is located in the arc of deforestation in the southwestern Amazon (Figure 1). Average annual temperature is 24 °C with 80–85% relative humidity (RADAM 1978). Average annual rainfall decreases from north to south (2600–1600 mm/year) (Dunne 1999), most of which falls between December and April. Average elevation increases from less than 100 m above sea level in the north to more than 500 m in the southeast, and topography varies, including flat river valleys, rolling hills, inselberg dotted plains, and rugged NNW–SSE oriented mountains/scarps in the central and eastern parts of the state. Rondônia is situated on the northwest edge of the Brazilian Shield, where lithology is dominated by granitoids and meta-supracrustal basement rocks, overlain in places by alluvium. Humans in Rondônia and the Amazon region as a whole have modified the landscape for thousands of years (Cleary 2001; Lima et al. 2002), but the most dramatic, widespread impacts have occurred since resource extraction booms in the 1970–1980s and government settlement projects attracted agrarian settlers who typically clear-cut the tropical forest to cultivate crops, most often cattle pasture (Fearnside 1993; Schneider et al. 2000). Deforestation rates in Rondônia from the 1970s to the present have fluctuated in response to regulatory policies and national economics (Walker and Homma 1996; Roberts et al. 2002), with annual averages ranging from 0.5 to 2.5% per year (Roberts et al. 2002). The mosaic of environmental conditions including variations in annual rainfall, topography, substrate composition, and land-use history suggests processes controlling soil formation occur at multiple spatial and temporal scales, the products of which jointly determine local soil properties in Rondônia.

Statistical methods

Geostatistical methods were originally developed to take advantage of the strong spatial correlations among field sampled data, which often confound the assumptions of sample independence in classical statistical techniques, to better model environmental variables and to optimally predict values in unsampled locations in the landscape. Geostatistics provide the tools to quantitatively determine the particular spatial scales at which distinctive patterns in soil properties emerge. Rather than producing a single, complex map for each soil property, we divide original laboratory measurements into maps of soil property estimates at multiple scales, which add together to produce the original complex map achieved by using standard spatial interpolation techniques. All measurement values are preserved at their original locations in the output maps.

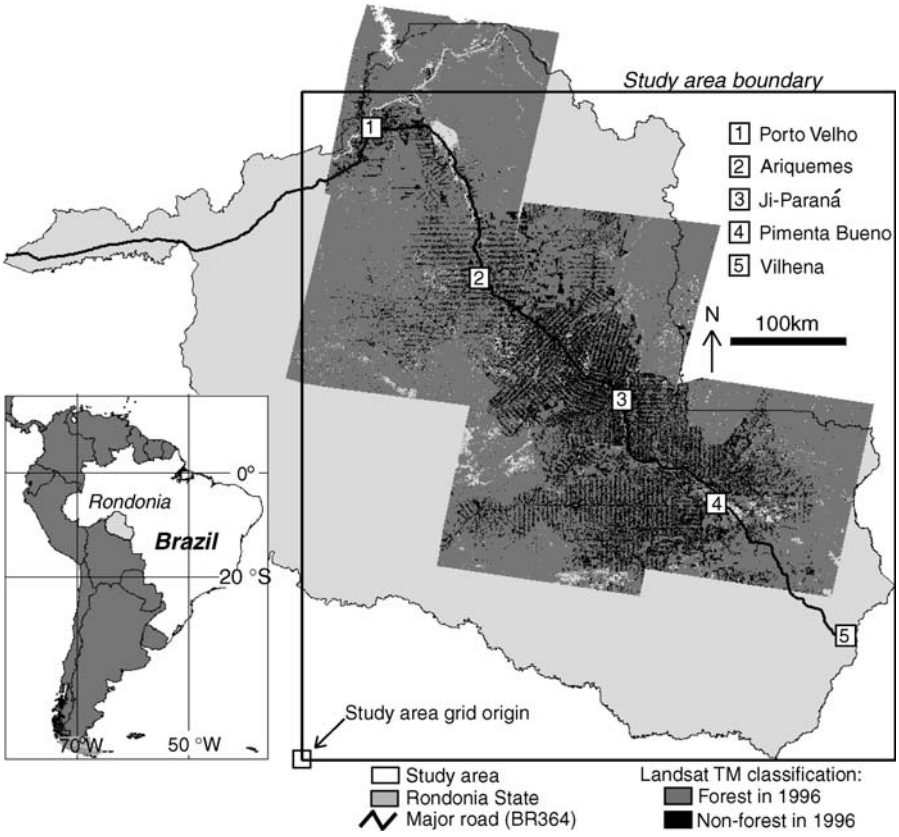

Figure 1. Location map for the study area, showing the extent of 5 Landsat TM scenes classified as forest and non-forest used as the land-cover map for our analysis. The major highway and large towns are included for reference. The lower left-hand corner of the box marking the study area is the coordinate system origin for all figures, which are measured in kilometers east and north of this point. The origin coordinates (UTM Zone 20), South American 1969 are (345935.5m, 8488501m).

The analysis consisted of the following generalized steps: (1) the spatial scales were chosen by modeling spatial dependence of each property (auto-semivariograms), then every possible combination of properties (cross-semivariograms), and deriving the common scales (correlation ranges) at which detectable changes in spatial dependence (e.g. different patterns) occur for 11 soil properties; (2) the soil properties were estimated at every scale using a spatial filtering method (multivariate cokriging) both at the sample locations to be used in step (3), and on a 1-km × 1-km grid for visualization purposes; and (3) the relative power of the different soil forming factors for predicting soil property levels at each spatial scale (calculated using stepwise multiple regression) was ranked to evaluate potential environmental controls on quantitative soil properties at particular spatial scales. The geostatistical

methods have been described in detail in the soil science and geostatistical literature (see Goovaerts 1992, 1994, 1997; Dobermann et al. 1995), and the major steps are outlined in the following sections.

Defining critical scales

Measured soil properties at different locations may be similar, such as soil samples taken in particular geologic units, or samples taken at similar positions in different watersheds, indicating predictable spatial patterns. A semivariogram quantifies the dissimilarity between pairs of measurements separated by distance h, and is calculated as the average squared semi-difference between every possible pair of measurements separated by that distance (Goovaerts 1997):

$$\gamma(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [z(u_\alpha) - z(u_\alpha + h)]^2, \tag{1}$$

where $N(h)$ is the number of pairs at distance h, and z is the value of interest at a location with coordinates u_α. Fitting a model to the variogram is analogous to modeling spatial correlation at multiple scales (Myers 1997).

The type of variogram model and the model parameters elucidate the overall variability, size, orientation, and intensity of patterns in soil properties at each scale. Because we were interested in the interactions among soil variables, all properties were modeled with the same number and type of nested variograms and ranges. This entails that only one parameter was adjusted at each scale: the variance, or sill. The sill is the plateau in variance that is reached at each scale, where the slope of the variogram changes. The sum of the partial sills for a given variogram is the same as the total variance of the soil property. Therefore, a comparison of the scale-specific sills reveals the proportion of variance accounted for at each scale, which is an important clue toward understanding the optimal scale at which to study or sample a particular nutrient, and which soil forming factors dominate its distribution: for instance, those acting over very large scales cause large-scale patterns in soils such as regional climate, or highly localized forces such as land management or termite activity cause highly localized patterns.

The variance for each nested scale is formally expressed as (Goulard and Voltz 1992):

$$\gamma(h) = \sum_{l=1}^{N_s} b^l g^l(h), \tag{2}$$

where N_s is the number of common variogram structures $g(h)$, b is the sill, and l indicates a particular spatial scale (Dobermann et al. 1995; Goovaerts 1997). The same equation in matrix terms is

$$\Gamma(h) = \sum_{l=1}^{N_s} B^l g^l(h), \tag{3}$$

where $\Gamma(h)$ is the variogram matrix (all auto and cross-variograms) and B^l is a matrix of sills, called the coregionalization matrix (Goovaerts 1992). Each matrix of sill terms describes the relationships among the soil properties at spatial scale l, as defined by the variogram function $g^l(h)$.

A traditional correlation coefficient measures the global correlation between variables, but it can be adapted to measure scale-specific correlations to further our understanding of how soil property relationships change with scale. This is a unitless measure of correlation (ρ_{ij}^l) between two variables (i and j) at scale l, filtering out variation at other scales (Dobermann et al. 1995):

$$\rho_{ij}^l = \frac{b_{ij}^l}{\sqrt{b_{ii}^l \cdot b_{jj}^l}}. \tag{4}$$

If the relationships among soil properties change with spatial scale, then depending on the distance between sampling sites, different correlations will result, indicating that the global correlation coefficient is a poor predictor of relationships measured in the field; rather, the data are best analyzed by using the scale-specific correlations. An analysis of scale is wholly dependent on variogram models and the coregionalization matrix, and therefore contains certain assumptions, namely that the observed patterns result from a linear combination of independent processes (nested variograms) which act on distinct spatial scales. The assumption of independence is invoked due to lack of scale-specific measurements, but can be relaxed if inter-scale dependencies can be obtained through, say, physical models of soil formation/evolution.

Partitioning soil properties into multiple scales

Once the critical spatial scales are identified through variogram modeling, we decompose or filter the soil properties and map estimates of the soil properties for each spatial scale using cokriging; see Appendix A, and Wackernagel (1998) and Goovaerts (1997) for a more complete discussion. Cokriging estimates were computed at the sample locations for regression modeling and on a regularly spaced 1-km × 1-km grid for visualization purposes. Geostatistical exploratory analyses were performed using the public-domain 'GSLIB' geostatistical software library package (Deutsch and Journel 1998), and multivariate dual cokriging was performed in Matlab (The MathWorks, Inc., 2001,Version 6.1).

Regression analysis

Once the important scales for soil patterns were determined, and the soil property measurements were decomposed into their scale-specific contributions, we evaluated the relative importance of the soil forming factors at every scale. This was done using stepwise multiple regression and several specification tests to analyze the relationships between the spatial components at the sample locations and the environmental gradients theoretically controlling long

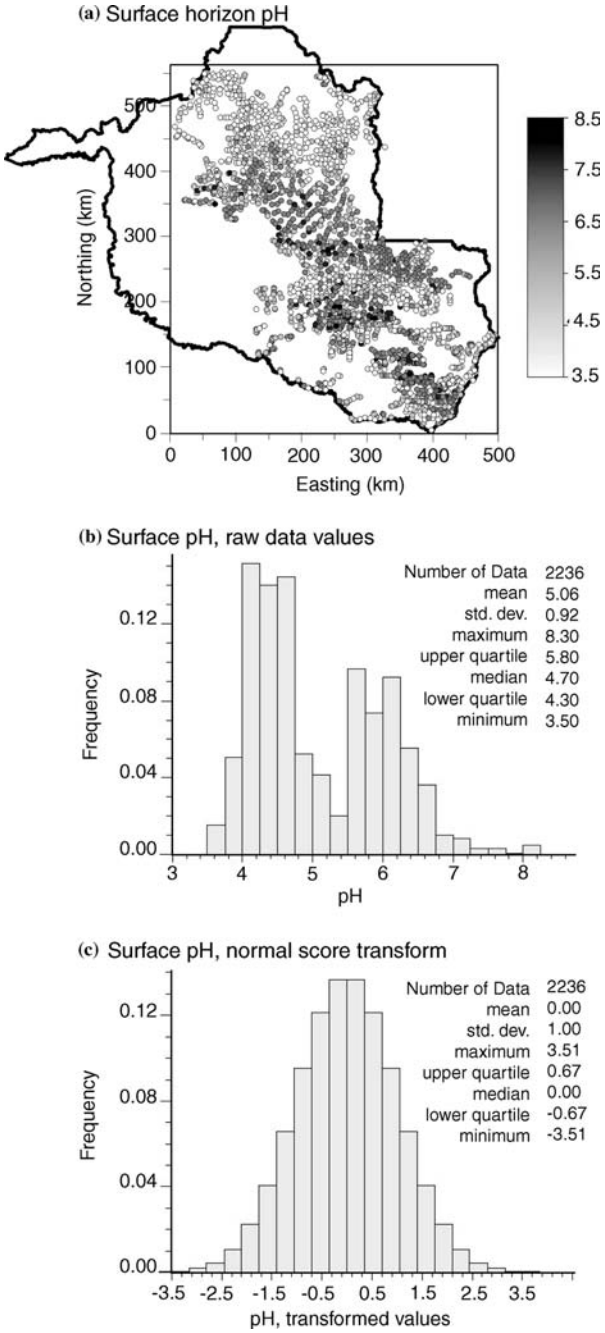

Figure 2. Map of sample locations and surface pH values (a), original histogram (b) and transformed histogram (c). Map units measured in kilometers from origin marked in Figure 1.

term soil development and soil nutrient availability. Regression modeling allowed simultaneous use of categorical and continuous scale explanatory variables, and tests of interaction effects. The predictor variables in the best models were ranked to evaluate the relative importance of the soil forming factors as controls on soil chemistry.

The best regression models were determined through a stepwise selection process, and reassessed at every step to identify the best model in terms of fit (explained variation) and parsimony. Specifically, independent variables were evaluated using either *t*-tests (continuous scale variables) or F-tests (groups of categorical variables) (Maddala 1992), and those with a *p*-value greater than 0.15 were omitted from the model. The 0.15 cutoff was chosen to mitigate Type II effects balanced against Type I effects. Once the best model was constructed, the relative contribution of each of the independent variables was determined and ranked based on the *p*-values from four specification tests (F-test, Likelihood Ratio, Wald Score Test, and the Lagrange Multiplier), by comparing a restricted model to an unrestricted model. Only the F-test is reported here for simplicity. S-plus was used for all regression calculations (MathSoft Inc., 2000 Professional Version 3).

Soil profile data

The soil data which made this study possible were collected between September, 1996, and November, 1997, for agricultural zoning (2^a ZEE-RO: Segunda Aproximação do Zoneamento Sócio-Econômico Ecológico do Estado de Rondônia) (Cochrane 1998). The full data set includes samples from approximately 3000 soil profiles, 2236 of which fell within the study area ($\sim$195,000 km^2) (see sample locations, Figure 2a). The average distance between samples was approximately 4 km, but many were separated from neighboring samples by much shorter distances. We assembled data for the soil surface and subsurface horizons, respectively defined as horizons with an upper boundary between 0 and 5 cm, and the deepest horizon sampled within each profile with an upper boundary deeper than 25 cm (Table 1).

We focused on measurements of key plant nutrients and soil nutrient supplying capacity: pH, exchangeable calcium (Ca), potassium (K), magnesium (Mg), aluminum (Al), available phosphorus (P), total nitrogen (N), organic carbon (OC), effective cation exchange capacity (ECEC), and particle size (clay and sand %) (Table 1). Few bulk density measurements were made, so we chose to work with soil nutrients as concentrations rather than calculate nutrient stocks, because the latter would increase measurement error. For general reference, the average bulk density for the surface and subsurface horizons was 1.26 ± 0.16 and 1.30 ± 0.17 g/cm^3, respectively. Land use, such as pasture for cattle, agroforestry, tree plantations, or urban, was noted at the sample locations by the surveyors, and for this analysis was simplified into forest and non-forest classes. Management practices such as degree of logging,

Table 1. General soil sample statistics for the surface and subsurface (deepest sampled) horizons in the Soteron dataset.

	Units	Surface (2236 samples)				Subsurface (1616 samples)				Lab method
		Mean	Stdev.	Min.	Max.	Mean	Stdev.	Min.	Max.	
Average sample depth	cm	7.15	2.04	1.50	35.00	112.74	8.45	82.50	202.50	Measured in field
pH		5.06	0.92	3.50	8.30	4.92	0.65	3.70	8.60	1:5 soil:water, colomel electrode
Phosphorus (P)	ppm	4.74	10.65	0.20	99.80	1.03	5.28	0.20	99.70	Mehlich method
Nitrogen (N)	%	0.14	0.08	0.00	0.69	0.00	0.00	0.00	0.08	Kjeldahl procedure
Organic carbon (OC)	%	1.77	1.13	0.17	13.25	0.30	0.25	0.02	4.83	Walkley–Black procedure
Effective cation exch. capacity	cmolc/kg soil	4.43	5.55	0.47	89.35	2.22	2.54	0.32	24.57	Calculated
Exch. calcium (CA)	cmolc/kg soil	2.75	5.36	0.00	87.69	0.68	1.39	0.01	22.61	NH4-Ac extract at pH 7, FAA
Exch. magnesium (MG)	cmolc/kg soil	0.44	0.83	0.02	18.88	0.32	1.02	0.02	14.30	NH4-Ac extract at pH 7, FAA
Exch. potassium (K)	ppm[a]	0.19	0.18	0.00	1.71	0.09	0.14	0.00	1.37	NH4-Ac extract at pH 7, FAA
Exch. aluminum (AL)	cmolc/kg soil	0.75	0.89	0.00	7.74	0.85	1.14	0.00	8.72	Exchangeable Al: extraction in 1 M KCl
Sand	%	56.50	20.79	1.80	92.00	43.07	23.09	1.00	91.00	Sedimentation cylinder
Clay	%	33.41	17.82	5.00	84.00	48.00	21.06	6.00	91.00	Sedimentation cylinder

[a]To convert to cmolc/kg soil, divide by 391.02.

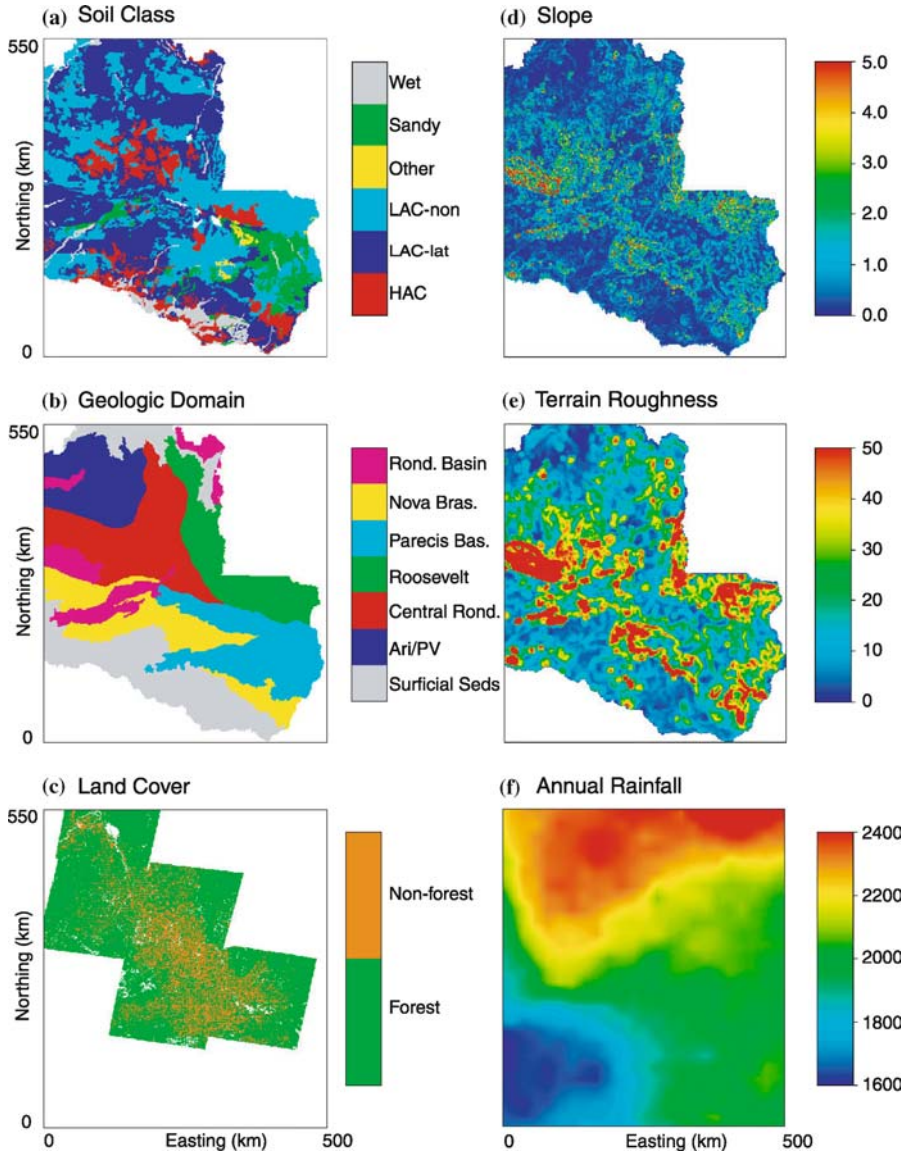

Figure 3. Environmental variables in the study area used as predictors for soil nutrient distributions: (a) soil types, (b) geologic domain, (c) land cover (forest/non-forest), (d) slope in degrees, (e) terrain roughness, (f) average annual precipitation. Map units measured from the origin marked in Figure 1. Geologic domain units listed in the legend are Rondonia Basin, Nova Brasilandia, Parecis Basin, Roosevelt, Central Rondonia, Ariquemes/PortoVelho, and Surficial Sediments. Please see the text for further explanation.

pasture fertilization, crop/animal rotation, and fire frequency were not recorded (Cochrane 1998).

Environmental drivers

Existing maps were used as proxies for the soil forming factors in Rondônia (Figure 3). A gridded map of average annual precipitation (Dunne 1999) was used to approximate the climate gradient, assuming temperature differences across the state to be negligible.

Slope, terrain 'roughness', and a soil wetness index were calculated from a digital elevation model (Moore et al. 1993; SEDAM 1999; Holmes et al. 2000) to assess the influence of landforms and water flow pathways on soil nutrient distributions. Roughness was measured as the standard deviation in elevation calculated on each pixel in the DEM within a 5-km radius circle. The soil wetness index (compound topographic index, CTI) was calculated from the DEM using the following equation (Moore et al. 1993):

$$\mathrm{CTI} = \ln(A_s / \tan\beta) \tag{7}$$

where A_s is the specific catchment area, or square meters per unit width orthogonal to the flow direction, and β is the slope angle in degrees.

As a proxy for parent material nutrient-supplying capacity we used the state geology map aggregated into stratigraphic-tectonic domains as defined by the state geological surveyors based on integrated petrologic, petrographic, geochronologic, geophysical, and structural data (Figure 3B) (Scandolara 1999). While these units are fairly coarse, they capture the differences in lithology from the highly weathered sediments around Porto Velho (Surficial sediments); intermediate composition granitoids in central Rondônia (Ariquemes/Porto Velho domain, Central Rondônia domain); and mixed granites and intrusive basic rocks in the south central/southeastern part of the state (Nova Brasilândia domain, Parecis Basin domain). The remaining domains are heterogeneous, consisting of a mixture of extrusive and intrusive rocks (Roosevelt domain), and basins receiving sediments from a variety of sources (Rondônia Basin).

We included the existing state soil map as a potential explanatory factor for soil chemistry in this analysis (EMBRAPA 1983), under the assumption that the map well summarizes local combinations of soil forming factors. When incorporating categorical variables into regression modeling, a large number of categories typically have only one or two samples, and can create matrix singularities. This leads to problems implementing and interpreting the model outputs. To avoid this problem, the state soil map was simplified into five categories which reflect major trends in tropical soil properties, namely clay type and drainage (Figure 3A): high activity clays (HAC), low activity clays in latossolos (LAC-lat), low activity clays on non-latossolos (LAC-non), sandy, and wet (e.g. IPCC/UNEP/OECD/IEA 1997; Bernoux et al. 2002). Soils with low activity clays are extremely common, so to better partition the samples, the

Table 2. Correlation coefficients among soil properties within the surface and subsurface samples.

Variable	pH	P	N	OC	ECEC	Ca	Mg	K	Al	Sand	Clay
(a) Surface horizons											
pH	1.00										
Phosphorus	0.30	1.00									
Nitrogen	0.04	0.20	1.00								
Org. carbon	0.04	0.20	0.91	1.00							
ECEC	0.53	0.35	0.33	0.32	1.00						
Calcium	0.64	0.37	0.26	0.25	0.97	1.00					
Magnesium	0.38	0.12	0.20	0.21	0.57	0.47	1.00				
Potassium	0.48	0.19	0.22	0.25	0.43	0.40	0.39	1.00			
Aluminum	−0.74	−0.14	0.14	0.15	−0.18	−0.36	−0.21	−0.25	1.00		
Sand %	0.02	0.04	−0.33	−0.33	−0.24	−0.15	−0.19	−0.24	−0.23	1.00	
Clay %	−0.11	−0.07	0.31	0.29	0.15	0.07	0.11	0.14	0.26	−0.96	1.00
(b) Subsurface horizons											
pH	1.00										
Phosphorus	0.14	1.00									
Nitrogen	−0.03	0.01	1.00								
Org. carbon	−0.07	0.04	0.07	1.00							
ECEC	0.10	0.07	−0.03	−0.05	1.00						
Calcium	0.58	0.17	−0.03	−0.02	0.63	1.00					
Magnesium	0.19	0.01	−0.02	−0.03	0.74	0.42	1.00				
Potassium	0.34	0.05	−0.04	−0.09	0.46	0.38	0.33	1.00			
Aluminum	−0.54	−0.06	−0.01	−0.02	0.52	−0.18	0.10	0.08	1.00		
Sand %	−0.06	0.04	0.12	−0.18	−0.23	−0.19	−0.13	−0.15	−0.11	1.00	
Clay %	0.04	−0.05	−0.11	0.21	0.13	0.14	0.05	0.06	0.04	−0.98	1.00

low activity clay category was divided into two groups: those belonging to the Brazilian soil order 'Latossolos', which cover approximately 45% of Rondônia (Cerri et al. 2000), and other soils. Latossolos are characterized as having highly weathered subsurface mineral horizons in which the clay fraction consists predominately of iron oxides, aluminum oxide, and kaolinite, virtually no primary minerals remaining, and rapid drainage (de Oliveira et al. 1992). They correspond with Oxisols and Ferralsols in the US and FAO-Unesco soil classification systems.

Based on a non-spatial analysis of all the available soil measurements for the area, the major distinction in regional soil biogeochemistry in this landscape is between natural forest and human-altered areas (Holmes et al. 2004). Our intent here is to assess whether this relationship holds true at all sampling or analysis scales, and which state factors exert the strongest influence on soil property distributions at different scales. To determine land cover at every location across the majority of the field area, we used a mosaic of five Landsat Thematic Mapper (TM) images from 1996 classified as forest or non-forest using spectral mixture decomposition (Roberts et al. 2002). The mosaic provided information for 66% of the sample locations, but regrettably full coverage for Rondônia of sufficient quality for accurate classification was not available. We coarsened the resolution of the images from 30-m to 90-m pixels to improve the agreement

between the remotely sensed land-cover classification and survey reported land cover. Each cell in the new grid was assigned the percent of forest found in the nine co-located original 30-m cells. There was 89% agreement between soil-surveyor land cover and classified land cover at sample locations, with all mixed pixels classified as forest. We infer that soil samples labeled as forest were collected in forest margins, which is likely due to logistical difficulties for sample collection in the often thick tropical vegetation. Samples collected near boundaries have a higher likelihood of misclassification due to georeferencing error when combining the soil pit locations with the classified images. For consistency in the analysis of soil forming factors, only those soil samples were included for which Landsat TM classified land cover was available, reducing sample size by 30% (see extent of Landsat coverage, Figure 1).

Results and discussion

Non-spatial soil property variability

Each soil property had a relatively wide range of values and a corresponding large standard deviation (Table 1). The single exception was nitrogen in the subsurface, where it was practically absent. None of the soil properties fit a Gaussian distribution so prior to analysis each was transformed using normal scores, a non-linear, rank preserving transform that remaps any distribution to a normal distribution, with zero mean and a standard deviation of one (Figure 2B, C) (Goovaerts 1997).

There were strong positive correlations among several surface horizon properties ($r > 0.5$: pH and K, Ca; OC and N; ECEC and Ca, Mg), strong negative correlations were found among others ($r < -0.7$: pH and Al; sand and clay), and some properties generally showed weak correlations with the other variables (P, N, OC, with the exception of OC:N) (Table 2). The strong correlation between pH and some of the other properties is due to the role of pH in regulating soil chemical reactions. Soil pH is considered a master variable, and is commonly measured and used for a quick practical assessment of soil conditions. Much of our discussion of patterns in soil properties is focused on pH for efficiency.

In general, the correlations in the subsurface were weaker than at the surface, except for those among ECEC and Ca, Mg, K, and Al, which were stronger. The pool of exchangeable base cations in the surface horizons was dominated by Ca, while at depth, K and Mg were more abundant, as was Al, indicating an increased likelihood of Al-toxicity for plants. Average values for most variables (pH, P, N, OC, ECEC, Ca, Mg, K, sand) decreased with depth reflecting the importance of organic contributions to the surface soil properties, with the exception of exchangeable aluminum and percent clay, which are more prominent at depth for pedological reasons (Table 1). Soil property correlations between the surface and subsurface were typically higher than 0.5

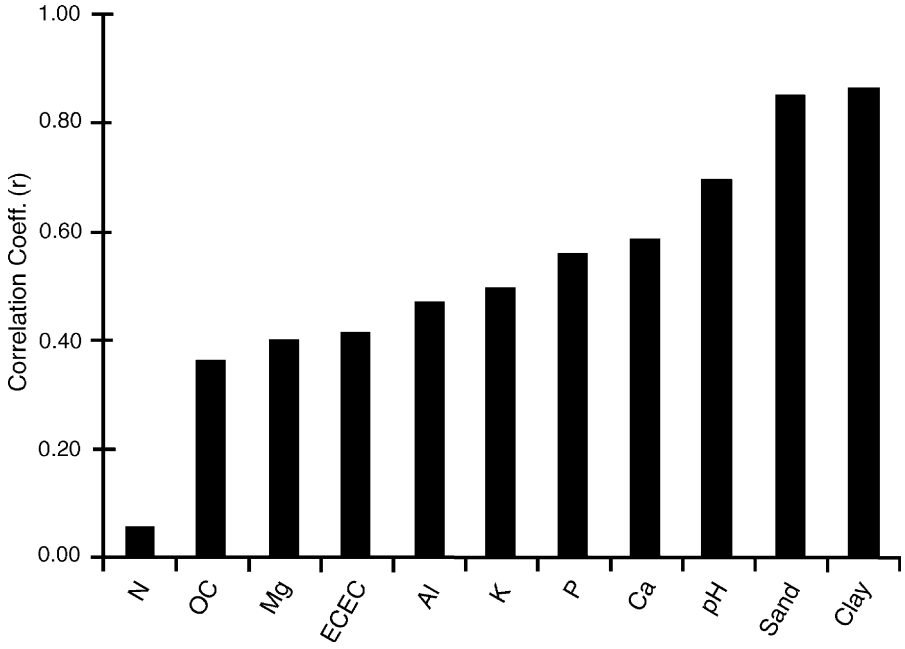

Figure 4. Correlation coefficients calculated between soil properties in the surface horizons versus the subsurface horizons.

(Figure 4). Nitrogen, OC, and Mg had low surface-to-subsurface correlations, whereas Ca, Sand, and Clay had high correlations. The lack of correlation between N and OC in the surface and subsurface reflects shallow rooting and tight nutrient cycling by plants and/or recent changes in land management practices (Neill and Davidson 2000).

Nested spatial scales

Several scales of soil property spatial variability were visually identifiable from the auto-variogram (comparing how one property changes over increasing distances) and the cross-variogram (comparing the correlation between two properties over increasing distances) analysis, and were quantifiable using co-regionalization modeling (Figure 5). In theory, the scales of soil property variation are determined by the combined action of environmental processes. The scale, in a geostatistical context, refers to the distance (x-axis of variogram plot) at which changes in the slope (sill) of the variogram occur. We jointly modeled the 11 auto- and 55 cross-variograms for the 11 soil properties using an automated iterative weighted least squares fitting program (Goulard and Voltz 1992). The final models consisted of four nested spatial scales: a nugget (< 3 km), short (< 10 km), medium (< 68 km), and long (> 68 km) scales (Table 3).

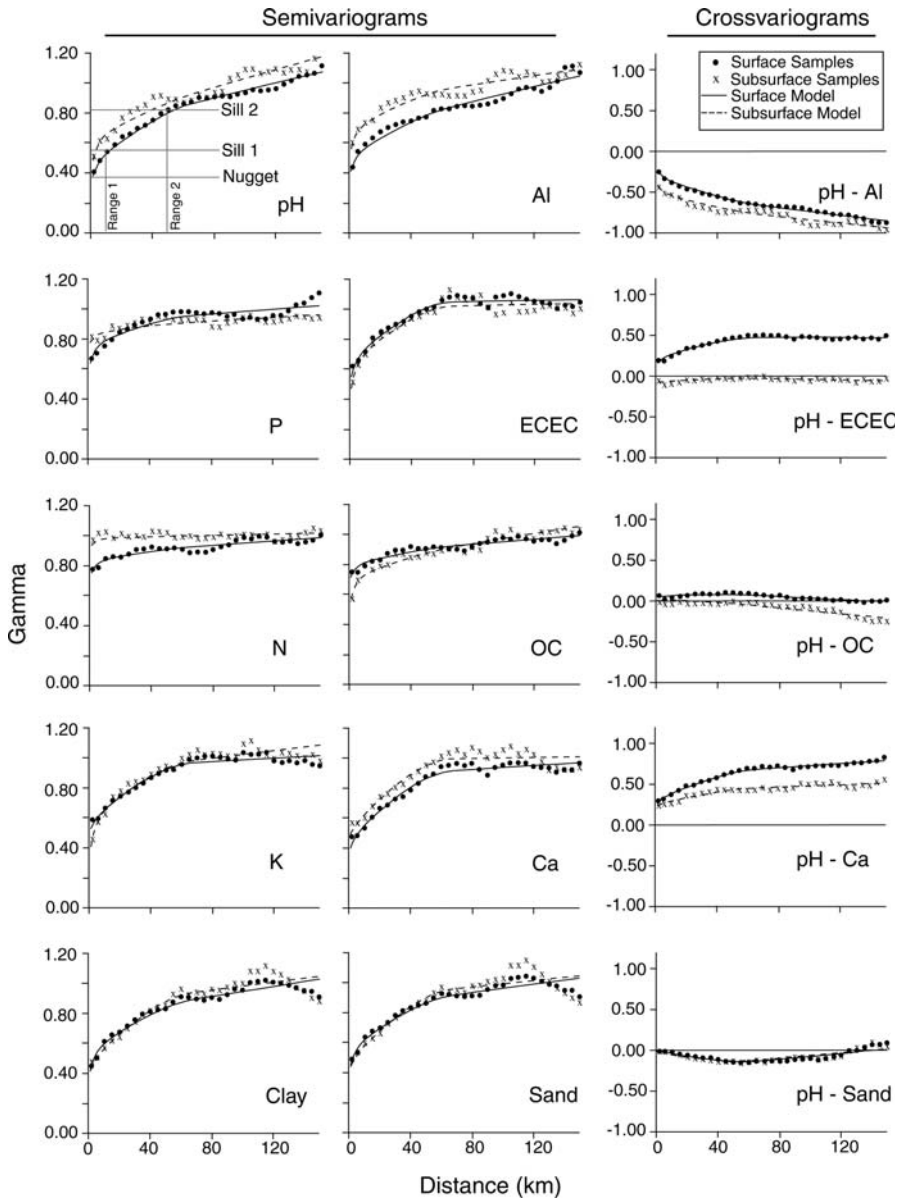

Figure 5. The first two columns show the variograms (points) and models (lines) for all soil properties in the surface and subsurface horizons. Cross-variograms between pH and 5 other properties are presented in the right-hand column. The nugget, sill, and range are illustrated for the surface horizon pH variogram.

Table 3. Types of nested variogram functions and their ranges used in coregionalization modeling for surface and subsurface horizons, inferred over all directions for 30 5-km lags, tolerance set to 3 km.

Scales	Range (km)	Model type
Micro (nugget)	<3	Nugget
Short (structure 1)	<10	Exponential
Medium (structure 2)	<68	Spherical
Long (structure 3)	>68	Spherical

The nugget of the variogram (γ value at extremely small distance) accounts for both micro-scale variability in soil properties (e.g. localized disturbances such as tree fall) and laboratory and other procedural errors. A variable with no spatial dependence would have a variogram that is pure nugget (a horizontal line), in which case, a comparison of measurements separated by any distance returns the global variance, i.e., the spatial distribution is completely random, and the geostatistical approach yields the same results as those obtained through classical statistical analysis. Any errors associated with laboratory measurements or georeferencing are likely to be independent and cancel each other out, so the nugget in the cross-variogram better represents the true micro-scale variability (Goovaerts 1994). For these data, the cross-variogram nugget values were low and the auto-variogram nugget values were high, suggesting there is some inherent short distance variability, but there is also a large amount of uncertainty in the soil property measurements. The extremely large nugget for several of the properties suggested that the sample density was too low to adequately analyze local patterns (N, OC, P). Only a subset of the 55 crossvariograms could be shown due to space limitations so crossvariograms including pH were selected to supply examples of how well-studied correlations between soil properties change with increasing distance in this environment (Figure 5).

The sill of the variogram, in combination with the distance (range) at which that sill is reached, represents the scale threshold between different relationships in the data. The nugget, equivalent to the first sill of the variogram, explained a larger percentage of variance in the surface horizons than the other scales for N, OC, P, ECEC, Mg, and K (Table 4). The longest scale was dominant for pH, Al, sand, and clay. Only Ca was best described by the medium scale, and the short scale explained the least variability for any of the soil properties with the exception of N, OC, and ECEC. The long-scale variation reflects changes in climate and parent material that most strongly influence basic pedological properties such as clay and sand. The short-scale variability is controlled by plant distribution and biocycling of carbon and nutrients. In these low activity soils, ECEC is imparted by organic matter and therefore correlates with OC and N. The fact that Ca scale dependence is at medium scale may well relate to the fact that it combines a pedological component (long scale) with a biocycling component (short scale).

Table 4. Percent of total variance accounted for by each scale of analysis in the surface horizons, determined through coregionalization modeling.

	Surface				Subsurface			
	Nugget < 3 km, %	Str. 1 < 10 km, %	Str. 2 < 68 km, %	Str. 3 > 68 km, %	Nugget < 3 km, %	Str. 1 < 10 km, %	Str. 2 < 68 km, %	Str. 3 > 68 km, %
Al	23	9	10	58	38	12	14	36
pH	20	8	14	58	26	11	6	57
Clay	26	12	17	45	30	7	30	34
Sand	30	11	20	40	34	6	33	28
OC	57	9	6	27	37	15	7	41
P	49	13	13	25	70	8	3	18
N	62	9	4	24	87	6	0	7
Ca	33	8	36	22	45	5	45	5
K	43	7	30	19	26	19	24	31
Mg	43	9	37	11	44	14	36	6
ECEC	45	15	35	6	37	22	35	6

Crossvariograms for all possible combinations of soil properties showed significant spatial dependence (e.g. pH vs. selected properties, right-hand column Figure 5). The correlation between pH in the surface and subsurface is high ($r = 0.7$), therefore the changing relationships between pH and other properties with depth is dominated by changes in the measured amounts of other properties. Aluminum and pH were strongly negatively correlated at the surface and at depth. Although pH and ECEC were reasonably well correlated at the surface, at depth they showed virtually no correlation. Organic matter supplies a large percentage of cation exchange sites in the surface, but is far less important in the subsurface with a consequent reduction in the pH–ECEC correlation (Table 2). Similarly, the pH–OC cross-variogram, although virtually zero, changes sign between the surface and subsurface. The variance between pH and Ca increased with distance, and had a long distance trend similar to that of the pH variogram, although it was less pronounced in the subsurface, a pattern that reflects the shift from biological control of Ca at short distance to lithological control at long distances. Finally, pH and Sand showed low negative correlation over most of the range, and low positive correlation at distances close to 150 km for both the surface and subsurface. Sand does not contribute positively or negatively to the pH dynamics in soil.

Variograms of the soil forming factors were also computed for a qualitative comparison with the soil property variograms discussed above (Figure 6A and B). Variograms for the categorical datasets, geology, soil class, and land cover, could not be constructed using the same methods as for continuous data. Instead, the indicator variogram was calculated, which shows the likelihood that points separated by a given distance belong to a

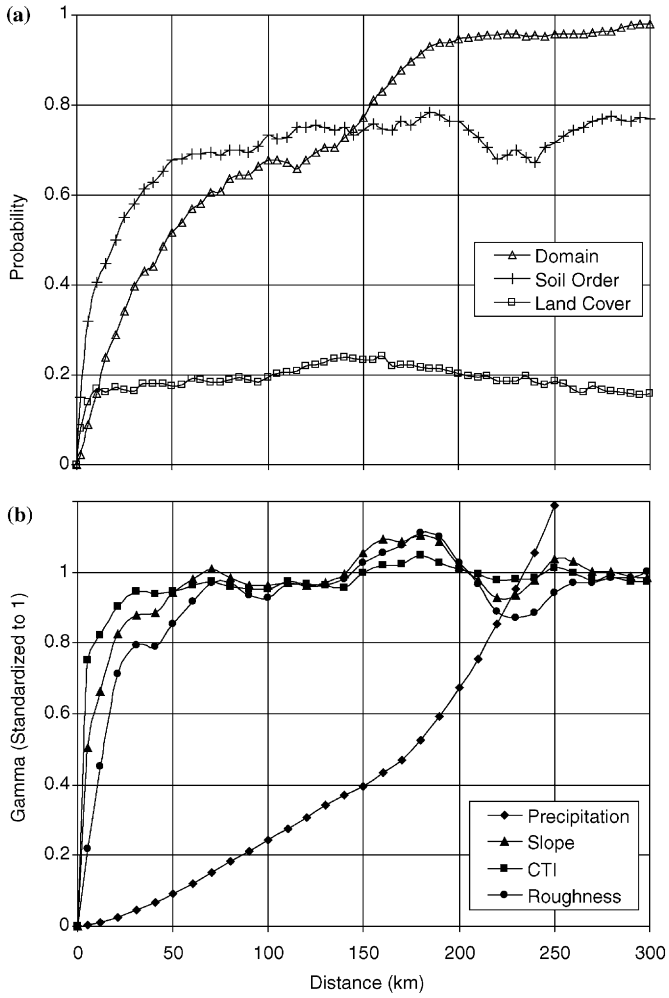

Figure 6. Variograms of the environmental variables included in the study. (a) The continuous variable variograms (precipitation, slope, CTI (wetness index), roughness) were standardized to 1. (b) The categorical variable variograms (geologic domain, soil order, and land cover) were calculated as indicator variograms, and then summed to represent the percent likelihood of any two samples falling within the same category when sampled a given distance apart.

different unit or class (Goovaerts 1994). The shape and ranges of the resulting indicator variograms can be compared in a qualitative way with variograms of the soil forming factors measured on continuous scales and the individual soil properties.

A comparison of the soil forming factor variograms with those of the soil properties suggests that soil nutrients with high variability at short distances (< 10 km) may be most closely tied to changes in land cover, slope, CTI, and

Roughness. Roughness had a longer range than the other terrain variables, but was still highly variable compared to the other soil forming factors considered (Figure 6B). The soil type and geology variograms had more gentle slopes, reaching a sill between 50 and 70 km which corresponds well with the sills identified at < 68 km in the soil chemistry samples. Both soil type and geologic domain continued to increase with increasing distance, as did precipitation which roughly followed a parabolic trend. The increasing variability with distance of both substrate characteristics and precipitation are likely causes of the long distance trend (> 68 km) identified in the variograms of pH, Al, and sand and clay.

Scale-specific correlations among soil variables

The linear correlation coefficient can be used to determine relationships among soil properties at the different spatial scales by analyzing the variogram partial sills. The scale-specific correlation coefficients for pH vs. the other 10 properties illustrate changing correlation with spatial scale (Table 5); discussion of the other variables is omitted because they do not provide added insight into scale dependent properties.

All variables showed stronger correlations at one or more discrete scales than was revealed by the traditional, global correlation analysis (see Results and discussion: *Non-spatial soil property variability*, and Table 5, column 1). The sign of correlations between pH and some variables remained constant regardless of spatial scale (pH with ECEC, Ca, Mg, K, and Al), while others switched sign with change of scale (pH with P, N, OC, sand, clay). Variables with a constant relationship to pH are those that are directly responsible for

Table 5. Structural correlation coefficients for pH and all other soil properties in the surface and subsurface horizons.

		Surface					Subsurface				
		Global	Nugget < 3 km	Short < 10 km	Medium < 68 km	Long > 68 km	Global	Nugget < 3 km	Short < 10 km	Medium < 68 km	Long > 68 km
P	pH	0.3	0.4	−0.1	0.1	0.8	0.1	0.2	−0.1	0.0	0.9
N	pH	0.0	0.1	0.0	0.6	−0.6	0.0	0.1	−0.1	0.5	−0.2
OC	pH	0.0	0.1	0.2	0.6	−0.7	−0.1	−0.1	0.4	0.4	−0.8
ECEC	pH	0.5	0.3	0.6	0.9	0.0	0.1	−0.2	0.0	0.4	−0.4
Ca	pH	0.6	0.7	0.6	0.9	1.0	0.6	0.5	0.4	0.7	1.0
Mg	pH	0.4	0.4	0.5	0.9	1.0	0.2	0.3	0.1	0.7	0.8
K	pH	0.5	0.2	0.4	0.9	0.8	0.3	0.1	0.0	0.6	0.9
Al	pH	−0.7	−0.6	−0.8	−1.0	−0.9	−0.5	−0.8	−0.6	−1.0	−1.0
Sand	pH	0.0	0.0	−0.1	−1.0	0.9	−0.1	0.1	−0.5	−1.0	0.9
Clay	pH	−0.1	0.0	0.0	1.0	−0.9	0.0	−0.1	0.6	1.0	−0.9

determining pH whereas those with changing scale relationships to pH are ones that have little or no direct effect on pH. The dramatic change in sign and magnitude of correlation between pH and clay at the longest scale could result from two possible scenarios. For areas with high activity clays, we would expect high clay content to correspond with high pH (positive correlation), and for areas with low activity clays, we would expect high clay content with low pH (negative correlation). The correlations measured for pH and clay are negative at short and long distances. The positive correlation between pH and clay at medium distances suggests either: (1) there is a characteristic spacing between geologic or geomorphic units dominated by high activity clays, and this is expressed in the variogram, which is a possibility because major river valleys with large amounts of low activity clays are found at the northern and southern boundaries of the study area; or (2) the central study area has experienced extensive deforestation, and the short-term influx of cations from burned vegetation to the soil has significantly increased soil pH, thus affecting the medium scale correlations.

The subsurface horizons showed similar patterns of correlation with scale. Stronger correlations between pH and Ca, Mg, Al, and texture at the medium and long scale suggested vertical translocation of solutes and clay. The global correlations were generally lower at depth than at the surface, but this is misleading because strong correlations were found at specific scales, and those with opposite signs cancelled out in the calculation of the global correlations.

Partitioning and mapping soil property measurements by spatial scale

Maps of the soil properties at each scale were produced from multivariate factorial kriging based on the nugget, sill, and range parameters determined through variogram modeling (Figures 7 and 8). The study area was divided into smaller blocks for processing due to computational demands, and slight differences in block means caused the checkerboard effect seen in the long-scale maps. The maps are presented in transformed and standardized units (mean = 0, standard deviation = 1) to enhance comparisons among properties with different measurement units or ranges. The sum of the maps across scales (from left to right in Figures 7 and 8) is equivalent to a map produced from the original data by standard spatial interpolation (ordinary kriging).

The patterns of nutrients at short scales are highly localized, corresponding to small areas around data points or small groups of points. The patches begin to merge at the medium scale, showing strong variability both north to south and east to west. At the long scale, the simplified pattern shows a strong north–south trend. The subsurface horizon maps are more homogeneous than those for the surface horizons. Soil pH and P have opposite trends in the surface and subsurface: pH appears lower on average in the subsurface than in the surface,

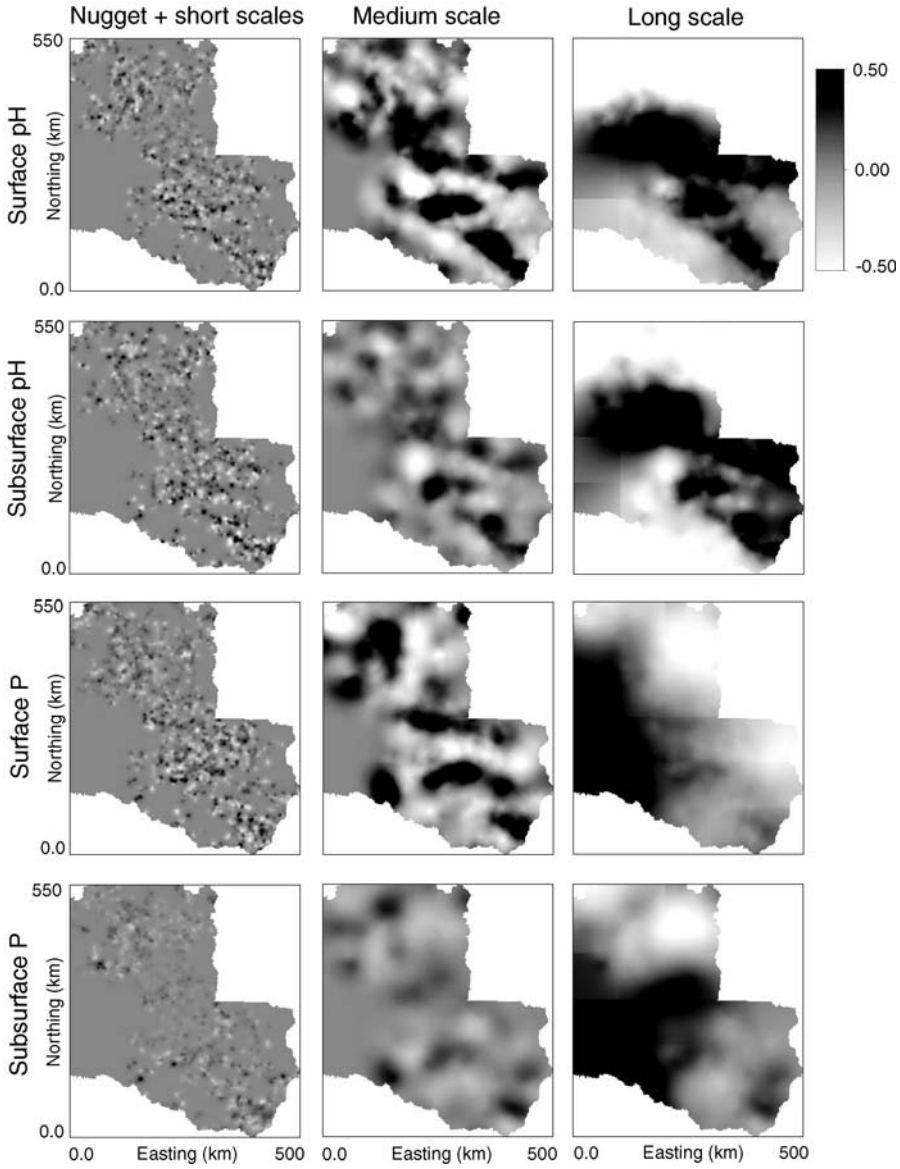

Figure 7. Spatial components of surface and subsurface pH and P, calculated using the coregionalization models and multivariate factorial kriging. All variables are presented in normal score transformed units to permit visual comparison of variables with different units and ranges of values. Map units measured in kilometers from origin marked in Figure 1.

while P values increase with depth. The overall smoothing effect in the subsurface makes sense from a physical standpoint: the surface layer is directly affected by localized processes such as tree fall, termite mounds, or

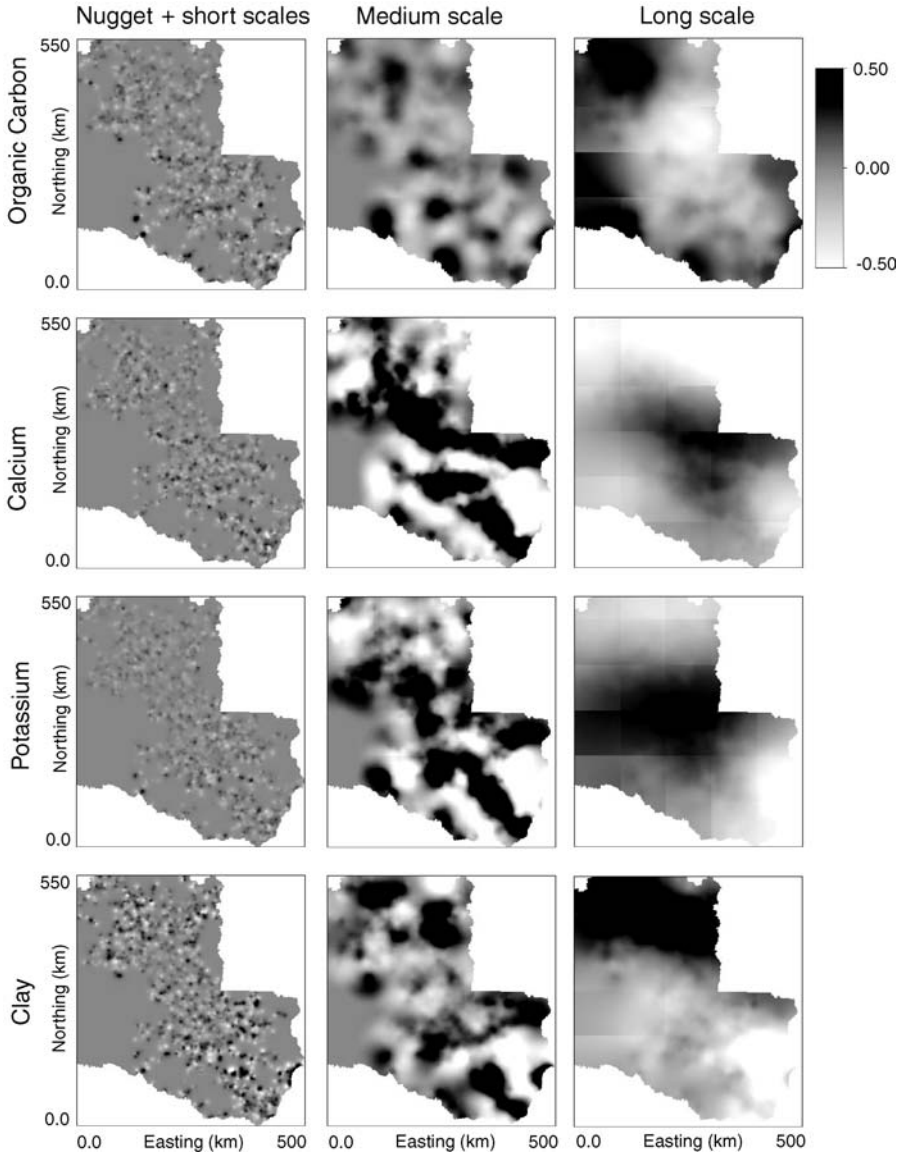

Figure 8. Spatial components of surface horizon OC, Ca, K, and Clay, calculated using the coregionalization models and multivariate factorial kriging. All variables are presented in normal score transformed units to permit visual comparison of variables with different units and ranges of values. Map units measured in kilometers from origin marked in Figure 1.

accumulation of ash or organic matter, which alter surface soil chemistry levels. However, it may take many years of decomposition and translocation of materials for these surface changes to affect the subsurface horizons.

Environmental factors driving soil variability

We used multiple linear regression to identify relationships among the scale-specific soil property estimates and environmental drivers. We were specifically interested in the effects of land-cover change relative to the rest of the soil forming factors, so we included the interactions of land-cover change and the other predictors in regression modeling.

The regression model fit for the surface horizon improved with increasing scale (Tables 6 (panel a) and 7 (panel a)). The nugget and short-scale were modeled both individually and combined. We used the combined value as representative of short-scale behavior because the models were very similar, no supplemental information was available for interpreting the nugget, and short-scale variation explained little of the overall property variance (Table 4). The surface horizon regression models for medium- and long-scales explained respectively 20–42% and 34–75% of the sample variance. The regression model fits for the subsurface horizon are also best for the long scale (Tables 6 (panel b) and 7 (panel b)). As expected from the variogram inspection, we found that terrain variables and land cover were the best predictors in the surface horizon at short scales, followed by geologic domain at medium scales, and precipitation at the longest scale. Calcium and pH were best predicted by land cover at short scales, while terrain Roughness was the best predictor of K at every scale.

Soil pH, which is representative of general soil solution conditions and potential Al-toxicity, exhibits strong, regional patterns, and is less variable than most other soil property measurements. Phosphorus, on the other hand, is highly variable at short distances (nugget = 45%) with a long-range component (long scale = 25%) in the surface horizons, but in the subsurface is dominated by micro scale variability (nugget = 70%). Phosphorus is minimally soluble and likely the pattern reflects areas of concentration around specific plants as it is biocycled and concentrated. Most variability in soil OC is documented in the nugget (57%), but there is a longer scale component (27% at >68 km) which is probably conditioned by the effect of rainfall on forest type and rates of carbon cycling. Our analysis has shown that K is more variable at short scales than Ca, and that it is more highly concentrated in the surface soil (Figure 4). The Ca and K models are very similar at short scales, and Ca follows the general trend of geologic domain and precipitation as dominant predictors at medium and long scales. However, K is dominated by local variations in topography (roughness) at every scale. Potassium is a highly mobile ion, and its distribution reflects local transport pathways as well as biocycling. Clay concentration could not be modeled at short scales, suggesting that the maps of lithology and geomorphic units we used were not adequate to represent realistic short scale variation in the study area. Lithology in Rondônia is extremely heterogeneous, and the state geology map depicts bedrock lithology as necessary for mineral exploration, which is not typically the parent material for soils. This is a common occurrence, due to

Table 6. Ranked importance of environmental predictors (1 is highest) used in the best models for each spatial component of the soil properties (pH, P, OC) in the (panel a) surface and (panel b) subsurface horizons.

	pH			Phosphorous			Organic carbon		
	Nug/Sh.	Med	Long	Nug/Sh.	Med	Long	Nug/Sh.	Med	Long
Panel a: Surface									
Precipitation		4	1		4	1	6	3	1
Geol. domain	3	1	2		1	2	7	1	2
Soils	5	2	4	5	2	5		2	3
Roughness			5		8	3		6	
CTI				2		8	1	9	
Slope	4			1			5	8	
Land cover (LC)	1	3	3	4	3	4	3	5	4
LC: precipitation		7			5	7		4	
LC: geol. domain	2	6	6			6	4		5
LC: soils		5	7	3	6			7	
LC: roughness					7				
LC: CTI							2		
LC: slope									
Resid SE	0.60	0.24	0.31	0.79	0.25	0.21	0.79	0.13	0.14
DF	1435	1429	1429	1443	1433	1432	1443	1432	1433
R-squared	0.08	0.38	0.66	0.02	0.21	0.35	0.02	0.24	0.68
Adj. *R*-sq.	*0.06*	*0.37*	*0.66*	*0.01*	*0.20*	*0.34*	*0.01*	*0.23*	*0.68*
F-test	6.21	34.47	112.28	2.22	18.24	35.32	2.22	20.43	148.26
Pr(*F*)	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00
Panel b: Subsurface									
Precipitation	8	3	1	2	4	1		3	1
Geol. domain	7	1	2		1	2		1	2
Soils		2	3		2	5		2	5
Roughness	9		4					5	9
CTI	1	5		5	6	6	1		10
Slope	3			4		9	2	6	
Land cover (LC)	2	4	5	3	3	3		4	3
LC: precipitation	4			1	7	7			8
LC: geol. domain	6				5	4			4
LC: soils						8			
LC: roughness			6					7	6
LC: CTI		6							7
LC: slope	5							8	
Resid SE	0.67	0.14	0.31	0.82	0.10	0.18	0.74	0.14	0.26
DF	986	1002	1002	1027	996	990	1030	1000	993
R-squared	0.08	0.31	0.59	0.02	0.19	0.47	0.01	0.12	0.52
Adj. *R*-sq.	*0.05*	*0.30*	*0.59*	*0.02*	*0.17*	*0.45*	*0.01*	*0.11*	*0.51*
F-test	2.79	29.59	96.86	4.58	10.99	32.22	6.54	8.23	45.39
Pr(*F*)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

differences between the scale at which the geology is mapped, the lack of distinction made between geomorphic surfaces, and the practical necessity of sampling soil profiles at points in the landscape. At the medium scale, some

Table 7. Ranked importance of environmental predictors (1 is highest) used in the best models for each spatial component of the soil properties (Ca, K, Clay) in the (panel a) surface and (panel b) subsurface horizons.

	Calcium			Potassium			Clay		
	Nug/Sh.	Med	Long	Nug/Sh.	Med	Long	Nug/Sh.	Med	Long
Panel a: Surface									
Precipitation		7	1		6	6		5	1
Geol. domain		1	2		3	3		1	2
Soils		3	4		2	2		2	3
Roughness	2	4	6	1	1	1			6
CTI					7	7			
Slope									
Land cover (LC)	1	2	3	2	4	4		3	4
LC: precipitation		8			9	9			
LC: geol. domain		6	5		8	8		4	5
LC: soils		5	7		5	5			7
LC: roughness	3								
LC: CTI									
LC: slope									
Resid SE	0.57	0.34	0.16	0.66	0.28	0.15	NA	0.25	0.20
DF	1472	1428	1429	1473	1427	1427		1435	1429
R-squared	0.03	0.43	0.66	0.01	0.42	0.52		0.23	0.76
Adj. *R*-sq.	*0.03*	*0.42*	*0.66*	*0.01*	*0.40*	*0.51*		*0.22*	*0.75*
F-test	15.33	41.17	112.48	8.48	37.61	56.51		22.12	179.94
Pr(*F*)	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00
Panel b: Subsurface									
Precipitation		6	1			1			1
Geol. domain		2	2		3	2		1	2
Soils		1	3		1	3		2	3
Roughness	1	3	6	1	2	5			5
CTI				2		9			
Slope									
Land cover (LC)	2	4	4	3	4	4		3	4
LC: precipitation									9
LC: geol. domain			7			6		4	7
LC: soils									6
LC: roughness		5	5			7			8
LC: CTI						8			
LC: slope									
Resid SE	0.60	0.45	0.06	0.63	0.27	0.19	NA	0.3098	0.1469
DF	1030	1002	996	1029	1004	994		999	990
R-squared	0.01	0.32	0.75	0.02	0.39	0.62		0.36	0.81
Adj. *R*-sq.	*0.01*	*0.31*	*0.74*	*0.01*	*0.38*	*0.61*		*0.35*	*0.81*
F-test	7.45	31.94	141.72	5.26	49.15	70.33		31.89	158.72
Pr(*F*)	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00

variance in clay % was explained by the regression model (adjusted $R^2 = 0.22$). The longest scale explained about 75% of the clay variance, representing broad changes in climate and parent material across the study area.

The regression modeling makes the scale dependence of soil forming factors explicit. There is however the possibility that the scale dependence is prejudiced by the resolution of the predictor variables: if we had high resolution maps of precipitation and parent material, we might resolve short-scale correlation among soil properties and these components. Scale matching is a critical problem in environmental modeling, particularly in remote regions with few data collection stations.

Conclusions

In this study, we utilized a large database that enabled the decomposition of large area soil nutrient distributions into their scale-specific patterns. Large database analysis is growing in importance for biogeochemical research because of the need to address the implications of regional perturbations at the appropriate scale. The analysis of large areas requires spatial treatment to avoid developing global results that overly smooth important environmental variability. Using geostatistical techniques on this extensive soil database we have determined the critical spatial scales that allow us to evaluate differing patterns of soil and ecosystem properties. It is evident that across Rondônia soil properties that are most strongly influenced by plant growth and vegetation decay show strong variation over very short spatial scales (scales that were not well captured by the regional database), whereas soil properties that are more strongly related to climate and substrate vary over broader spatial scales. The analysis provides a quantitative view of the patterns of short-scale variation in nutrient cycling overlaid on a more slowly changing background. Ultimately, these scales combine to determine the soil properties measured at individual field sites, and our approach allows disaggregation of the measured signal. The properties associated with biocycling of nutrients vary over short distances and they are the ones that are most strongly perturbed by land-use change. However, those nutrients with extremely high microscale variability are as well characterized by global statistics as using a spatial approach, and the additional effort required for taking a geostatistical approach is probably not justified.

Two major sources of error are evident in the work. The first is that large databases are seldom collected with a specific analytical intent and their generality often ensures that the wrong things have been measured. The second is that spatial interpretation of the database requires correlation with other geographical databases to link information on the distribution of climate, geology, etc. These data may be severely limited in spatial resolution and thus act to degrade the analytical power of spatial statistical techniques. On the other hand, if such high resolution data are available, the analytical power of the spatial approach could be dramatically improved. Still, an understanding of the scale-dependence of biogeochemical properties is critical to evaluating regional response to land clearing and other types of changes in land use.

Acknowledgements

This work was supported by NASA Headquarters under Earth System Science Fellowship Grant NGT5-30357 and LBA Ecology Grant NCC5-282. We thank SEPLAN/PNUD, Porto Velho, Rondônia, for making the soil profile database available to researchers, and several anonymous reviewers whose insightful comments greatly improved this manuscript.

Appendix

Estimates of scale-specific soil properties were computed using an optimal linear unbiased estimator (kriging):

$$z(u) = \sum_{l=1}^{L} z^{l}(u) + m. \qquad (A.1)$$

The spatial components $z^{l}(u)$ represent the value of variable $z(u)$ at the spatial scales l defined by the ranges determined through coregionalization modeling. Each spatial component is estimated as a linear combination of the observations and the stationary mean m:

$$z^{l*}(u) = \sum_{\alpha=1}^{n} d_{\alpha} C^{l}(u_{\alpha} - u) + m, \qquad (A.2)$$

where $d_{\alpha}(u)$ is the dual weight associated with the scale l specific covariance $C^{l}(u_{\alpha} - u)$, estimated through a system of n linear equations to ensure unbiased estimation and minimum error variance (Goovaerts 1994). We used dual multivariate cokriging with a stationary mean because it reduced the computational costs for the large number of data and estimates (Goovaerts 1997; Wackernagel 1998). The global mean m was added to the longest range component $z^{l*}(u)$ (Jaquet 1989).

Dual simple kriging

This discussion closely follows Goovaerts (1997). Dual kriging is an alternative (but equivalent) kriging expression in which the estimates are determined as linear combinations of covariance functions rather than data values (Goovaerts 1997). The standard simple kriging algorithm is as follows:

$$z^{*}(u) = \sum_{\alpha=1}^{n(u)} \lambda_{\alpha}(u)[z(u_{\alpha}) - m] + m, \qquad (A.3)$$

where $z^{*}(u)$ is the estimate at location u, estimated from neighboring data z measured at locations u_{α}, and a known global mean m. The kriging weights λ are

calculated from the covariances among known data values by solving the system of n-linear equations:

$$\sum_{\beta=1}^{n} \lambda_\beta(u) C(u_\alpha - u_\beta) = C(u_\alpha - u), \quad \alpha = 1, \ldots, n. \tag{A.4}$$

Based on this relationship, the simple kriging estimate can be represented as a linear combination of these covariance values plus the stationary mean m:

$$z^*(u) = \sum_{\alpha=1}^{n} d_\alpha C(u_\alpha - u) + m, \tag{A.5}$$

where d_α is the dual weight associated with the covariance $C(u_\alpha - u)$. This is the 'dual' form of the traditional simple kriging estimate. The dual kriging system does not result from minimizing error variance, as does the traditional simple kriging estimate, but rather it is derived from the exactitude property of the kriging estimate. The dual weights $d_\alpha(u)$ are determined per solution of the following system of equations:

$$z^*(u_\alpha) = \sum_{\beta=1}^{n} d_\beta C(u_\beta - u_\alpha) + m = z(u_\alpha), \quad \alpha = 1, \ldots, n. \tag{A.6}$$

Dual factorial kriging

Standard kriging techniques are used to estimate the point or block value of a continuous variable, such as soil pH at given locations. Soil pH can be affected by very localized processes, such as fertilizer applications, and by much broader scale processes, such as regional patterns of precipitation. If the scales at which the factors influence pH are very different from one another, they should result in different structures in the pH variogram. In the case presented in this paper, common ranges or scales were determined jointly from 11 soil chemical and physical properties: a nugget effect, short (<10 km), medium (<68 km), and long (>68 km) range structures, which can be related to reasonable physical controlling factors, such as micro-scale disturbances and measurement error, land management practices, lithology, and a regional precipitation gradient. Factorial kriging consists of partitioning the variance of the random variable of interest into several zero mean, independent spatial components (l), or factors, based on the variogram model.

The dual form for the lth spatial component is as follows:

$$z^{l*}(u) = \sum_{\alpha=1}^{n} d_\alpha^l C^l(u_\alpha - u) + m, \tag{A.7}$$

where $C^l(u_\alpha - u)$ is the covariance function of the spatial component $z^l(u)$. The dual weights are determined by

$$z^*(u_\alpha) = \sum_{\beta=1}^{n} d_\beta^l C(u_\beta - u_\alpha) + m = z(u_\alpha), \quad \alpha = 1, \ldots, n. \tag{A.8}$$

The dual weights only have to be solved for once, because they do not depend on the spatial component being estimated. Therefore, the dual factorial kriging estimate can be rewritten as:

$$z^{l*}(u) = \sum_{\alpha=1}^{n} d_\alpha^l C^l(u_\alpha - u) + m, \quad l = 0, \ldots, L. \tag{A.9}$$

In the case of traditional simple factorial kriging, the weights λ and kriging system would have to be solved $L + 1$ times at each location u.

In the case that the distance between data $(u_\alpha - u)$ is larger than the range of correlation of a specific covariance component $C^l(h)$, then all covariance terms in the above expression disappear, and the estimated value of the corresponding spatial component is zero, as expected.

References

Bernoux M., Carvalho M.C.S., Volkoff B. and Cerri C.C. 2002. Brazil's soil carbon stocks. Soil Sci. Soc. Am. J. 66: 888–896.

Birkeland P.W. 1999. Soils and Geomorphology. Oxford University Press, Inc., New York, 430 pp.

Brady N.C. and Weil R.R. 1999. The Nature and Properties of Soils. Prentice-Hall, Inc., Upper Saddle River, NJ.

Cerri C.C., Bernoux M., Arrouays D., Feigl B.J. and Piccolo M.C. 2000. Carbon stocks in soils of the Brazilian Amazon. In: Lal R., Kimble J.M. and Stewart B.A. (eds), Global Climate Change and Tropical Ecosystems. CRC Press LLD, New York, pp. 33–50.

Cleary D. 2001. Towards an environmental history of the Amazon: from prehistory to the nineteenth century. Latin Am. Res. Rev. 36: 65–96.

Cochrane T.T. 1998. SIGTERON: Sistema de informação geográfica para os terrenos e solos do estado de Rondônia, Brasil, completed by the consultancy consortium TECNOSOLO-DHV. Porto Velho, Rondônia.

de Oliveira J.B., Jacomine P.K.T. and Camargo M.N. 1992. Classes gerais de solos do Brasil: guia auxiliar para seu reconhecimento. FUNEP/UNESP, Jaboticabal, Sao Paulo, 201 pp.

Deutsch C.V. and Journel A.G. 1998. GSLIB: geostatistical software library and user's guide. Oxford University Press Inc., New York.

Dias-Filho M.B., Davidson E.A. and de Carvalho C.J.R. 2001. Linking biogeochemical cycles to cattle pasture management and sustainability in the Amazon Basin. In: McClain M.E., Victoria R.L. and Richey J.E. (eds), The Biogeochemistry of the Amazon Basin. Oxford University Press, Oxford, pp. 85–105.

Dobermann A., Goovaerts P. and George T. 1995. Sources of soil variation in an acid Ultisol of the Philippines. Geoderma 68: 173–191.

Dunne T. 1999. Amazon River Basin Precipitation 1972–1992. Available on-line [http://www.daac.ornl.gov] from Oak Ridge National Laboratory Distributed Active Archive Center, Oak Ridge, Tennessee, USA (Data set).

EMBRAPA 1983. Mapa de levantamento de reconhecimento de meia intensidade dos solos do estado de Rondônia, 1:500,000. Comissão Estadual de Planejamento Agrícola, Porto Velho.

Eswaran H., Kimble J., Cook T. and Beinroth F.H. 1992. Soil diversity in the tropics: implications for agricultural development. In: Lal R. and Sanchez P.A. (eds), Myths and Science of Soils of the Tropics. Soil Science Society of America, Madison, WI, pp. 1–16.

Fearnside P.M. 1993. Deforestation in Brazilian Amazônia: the effect of population and land tenure. Ambio 22: 537–545.

Goovaerts P. 1992. Factorial kriging analysis: a useful tool for exploring the structure of multivariate spatial soil information. J. Soil Sci. 43: 597–619.

Goovaerts P. 1997. Geostatistics for Natural Resources Evaluation. Oxford University Press Inc., New York, 483 pp.

Goovaerts P. 1994. Study of spatial relationships between two sets of variables using multivariate geostatistics. Geoderma 62: 93–107.

Goulard M. and Voltz M. 1992. Linear coregionalization model: tools for estimation and choice of cross-variogram matrix. Math. Geol. 24: 269–286.

Holmes K.W., Chadwick O.A. and Kyriakidis P.C. 2000. Error in a USGS 30-meter digital elevation model and its impact on terrain modeling. J. Hydrol. 233: 154–173.

Holmes K.W., Roberts D.A., Sweeney S.H., Numata I., Matricardi E., Biggs T.W., Batista G. and Chadwick O.A. 2004. Soil databases and the problem of establishing regional biogeochemical trends. Global Change Biol. 10: 796–814.

Houghton R.A., Skole D.L., Nobre C.A., Hackler J.L., Lawrence K.T. and Chomentowski W.H. 2000. Annual fluxes of carbon from deforestation and regrowth in the Brazilian Amazon. Nature 403: 301–304.

IPCC/UNEP/OECD/IEA 1997. Revised 1996 IPCC Guidelines for National Greenhouse Gas Inventories: Reference Manual (Vol. 3). Intergovernmental Panel on Climate Change, United Nations Environment Programme, Organization for Economic Co-Operation and Development, International Energy Agency, Paris.

Jaquet O. 1989. Factorial kriging analysis applied to geological data from petroleum exploration. Math. Geol. 21: 683–691.

Lima H.N., Schaefer C.E.R., Mello J.W.V., Gilkes R.J. and Ker J.C. 2002. Pedogenesis and pre-Colombian land use of ''Terra Preta Anthrosols'' (''Indian black earth'') of Western Amazonia. Geoderma 110: 1–17.

Maddala G.S. 1992. Introduction to Econometrics. MacMillan Publishing Company, New York, 631 pp.

Moore I.D., Gessler P.E., Nielsen G.A. and Peterson G.A. 1993. Soil attribute prediction using terrain analysis. Soil Sci. Soc. Am. J. 57(2): 443–452.

Myers D.E. 1997. Statistical models for multiple-scaled analysis. In: Quattrochi D. and Goodchild M.F. (eds), Scale in Remote Sensing and GIS. CRC Press, Inc., New York, pp. 273–293.

Neill C. and Davidson E.A. 2000. Soil carbon accumulation or loss following deforestation for pasture. In: Lal R., Kimble J.M. and Stewart B.A. (eds), Global Climate Change and Tropical Ecosystems. CRC Press, Boca Raton, FL, USA, pp. 197–211.

Nepstad D.C., Verissimo A., Alencar A., Nobre C., Lima E., Lefebvre P., Schlesinger P., Potter C., Moutinho P., Mendoza E., Cochrane M. and Brooks V. 1999. Large-scale impoverishment of Amazonian forests by logging and fire. Nature 398: 505–508.

Oline D.K. and Grant M.C. 2002. Scaling patterns of biomass and soil properties: an empirical analysis. Landscape Ecol. 17: 13–26.

Powers J.S. and Schlesinger W.H. 2002. Relationships among soil carbon distributions and biophysical factors at nested spatial scales in rain forests of northeastern Costa Rica. Geoderma 109: 165–190.

RADAM 1978. Mapa exploratório de solos, Projeto Radambrasil: Programa de Integração Nacional, 1:1,000,000. Ministerio das Minas e Energia, Departamento Nacional de Produção, Rio de Janeiro.

Richter D.D. and Babbar L.I. 1991. Soil diversity in the tropics. Adv. Ecol. Res. 21: 315–389.

Roberts D.A., Numata I., Holmes K.W., Batista G., Krug T., Monteiro A., Powell B. and Chadwick O.A. 2002. Large area mapping of land-cover change in Rondônia using multitemporal spectral mixture analysis and decision tree classifiers. JGR-C. Atmospheres 107(D20): LBA 1–18.

Rodenburg J., Stein A., van Noodwijk M. and Ketterings Q.M. 2003. Spatial variability of soil pH and phosphorus in relation to soil run-off following slash-and-burn land clearing in Sumatra, Indonesia. Soil Till. Res. 71: 1–14.

Scandolara J. 1999. Geologia e Recursos Minerais do Estado de Rondônia: Texto Explicativo e Mapa Geológico do Estado de Rondônia. CPRM, Brasília.

Schneider R.R., Arima E., Verissimo A., Barreto P. and Souza C. Jr. 2000. Sustainable Amazon: Limitations and Opportunities for Rural Development. World Bank and Imazon, Brasília, 61 pp.

SEDAM 1999. Digital Terrain Model. Secretaria do Meio Ambiente, Porto Velho, RO.

Sun B., Zhou S. and Zhao Q. 2003. Evaluation of spatial and temporal changes of soil quality based on geostatistical analysis in the hill region of subtropical China. Geoderma 115: 85–99.

Trangmar B.B., Yost R.S. and Uehara G. 1986a. Spatial dependence and interpolation of soil properties in west Sumatra, Indonesia: I. Anisotropic Variation. Soil Sci. Soc. Am. J. 50: 1391–1395.

Trangmar B.B., Yost R.S. and Uehara G. 1986b. Spatial dependence and interpolation of soil properties in west Sumatra, Indonesia: II. Co-regionalization and Co-kriging. Soil Sci. Soc. Am. J. 50: 1396–1400.

Uehara G., Trangmar B.B. and Yost R.S. 1985. Spatial variability of soil properties. In: Nielsen D.R. and Bouma J. (eds), Soil Spatial Variability: Proceedings of a Workshop of the ISSS and the SSSA, Las Vegas, USA, November 1984. PUDOC, Wageningen, The Netherlands, pp. 61–95.

Vitousek P.M. 1984. Litterfall, nutrient cycling, and nutrient limitation in tropical forests. Ecology 65: 285–298.

Wackernagel H. 1998. Multivariate Geostatistics: An Introduction with Applications. Springer-Verlag, New York, 291 pp.

Walker R. and Homma A.K.O. 1996. Land use and land cover dynamics in the Brazilian Amazon: an overview. Ecol. Econ. 18: 67–80.

Western A.W. and Bloschl G. 1999. On the spatial scaling of soil moisture. J. Hydrol. 217: 203–224.