Scale-Dependent Relationships between Population and Environment in Northeastern Thailand

Stephen J. Walsh, Tom P. Evans, William F. Welsh, Barbara Entwisle, and Ronald R. Rindfuss

Abstract

Social and biophysical data were collected, integrated, and analyzed to examine scale-dependent relationships between selected population and environmental variables for a study site in northeast Thailand. Data sets were generated through the use of remote sensing to characterize land-use/land-cover and plant biomass variation across the Nang Rong district; GIS to derive elevation, slope angle, and soil moisture potential; social survey data at the village level to categorize demographic variables; and a population distribution model to transform demographic data collected at discrete village locations to spatially continuous surfaces stratified by agricultural land uses. Statistical analysis employed multiple regression to estimate population density in relation to social and biophysical variables, and canonical analysis to relate population variables to environmental variables across a range of spatial scales extending from 30 to 1050 m. Findings indicate the importance of spatial scale in the study of population and the environment. Regression models reflect the scale dependence of the selected variables through plots of slope coefficients and R^2 values across nine scale steps. The variation in relationships among environment and population variables, evidenced through factor loadings associated with canonical correlation, suggest that relationships are not generalizeable across the sampled spatial scales.

Introduction

Fundamental to landscape ecology is the explicit interrelationship of scale, pattern, and process, and the realization that landscape form and function are linked through theory and practice. Implicit is the interaction of biophysical and social processes through a set of complex interrelationships that are composited and define spatial pattern across temporal and/or spatial scales. The landscape matrix reflects this imprint of human and environmental processes, feedback mechanisms between them, and the geographic site and situation of the landscape as viewed across space and time.

While prior research (e.g., Quattrochi and Goodchild, 1997) has sought to explore spatial patterns extending across a range of spatial scales associated with biophysical processes and "natural" landscapes, sufficient attention has not been paid to the importance of biophysical processes linked with social processes to examine the nature of patterns across spatial and/or temporal scales for landscapes having a pronounced "social" imprint. Scale dependent studies conducted in "human" environments have typically lacked the appropriate human data collected at fine spatial scales (e.g., household or village) and a sufficiently large sample. Moreover, researchers have been primarily biophysical scholars unaccustomed to working with social science data. In addition, the remote sensing and GIS research communities have only recently begun focusing attention on the interplay between social, biophysical, and geographic factors as drivers of land-use/land-cover change (LULCC).

Spatial scale is inherently involved in recognizing spatial patterns on the landscape and in estimating the relationships between landscape components and environmental and social processes deriving those patterns (Bian and Walsh, 1993; Allen and Walsh, 1996). Specific biotic, environmental, social, and historical processes function at various ranges of spatial scales. The ranges vary and overlap among processes. Spatial patterns of the landscape may be discernible at certain spatial scales and ranges of spatial scales (Meentemeyer and Box, 1987), and landscape may appear homogeneous at some scales but heterogeneous at others (Nellis and Briggs, 1989). Based upon the assumption that the spatial patterns of landscapes are formed by environmental and social processes whose effects vary with spatial scale, it is reasonable to hypothesize that the relationships between landscape pattern and the operative social and/or biophysical processes forming them are scale dependent.

Because of the inherent complexity of population-environment interactions, no one theoretical perspective is sufficient for analyzing and interpreting scale dependencies. Three theoretical foci are presented as a foundation and context for this research: (1) Landscape Ecology (e.g., Forman and Godron, 1986), (2) Human Ecology (e.g., Johnston *et al.*, 1995), and (3) Political Ecology (e.g., Blaikie and Brookfield, 1987). All of these paradigms share a common concern with interactions between human society and the natural environment, though each emphasizes different aspects of the relationship. Landscape Ecology examines the interrelationship of scale, pattern, and process; Human Ecology states that people are important actors on the landscape that shape and are shaped by the environment; and Political Ecology advocates the importance of exogenous environmental and social effects upon human behavior that may interact with current

---

S.J. Walsh and W.F. Welsh are with the Department of Geography, University of North Carolina, Chapel Hill, NC 27599-3220.

T.P. Evans is with Indiana University, ACT, Student Building 331, Bloomington, IN 47405-6101.

B. Entwisle and R.R. Rindfuss are with the Department of Sociology and Carolina Population Center, University of North Carolina, Chapel Hill, NC 27599-3997.

Photogrammetric Engineering & Remote Sensing,
Vol. 65, No. 1, January 1999, pp. 97–105.

0099-1112/99/6501–0105$3.00/0

LULC contexts but which may be beyond human control. The various concepts of scale are more or less embodied in the theories and methodologies of each of these areas of scholarly research.

The basic intent of this research is to examine the scale-dependent relationships between the social and biophysical environment in rural Thailand, a region that has experienced substantial deforestation and agricultural extensification over the last 40 years. Specific objectives include an examination of (1) how social data collected at discrete point locations through a population survey can be linked with remotely sensed data and derived GIS coverages represented continuously across the landscape, (2) how population and environment variables are related through multiple-regression analyses subject to concepts of scale dependence, and (3) how the assembled "group" of population and environment variables are related through canonical correlation techniques across the range of spatial scales, 30 to 1050 m. In so doing, we examine the general hypothesis that relationships among social and biophysical variables change as the scale at which they are examined changes.

Study Area: Nang Rong and the Northeast

The study area of our research is the Nang Rong district, located in northeastern Thailand (Figure 1). Northeast Thailand (or Isan) contains approximately one-third of the country's area, generates about one-fifth of the GNP, and has a population of more than 18 million people. The dominant occupation in the region is farming, and the majority of farm households own an average of three hectares of land (Ghassemi *et al.*, 1995). Deforestation associated with agricultural expansion has been underway in northeast Thailand for a century or more (Feeny, 1988). Prior to World War II, this expansion was accelerated with increased production of paddy rice. Since then, it has also included upland crops, particularly cassava and sugar cane. In the recent period, agricultural intensification occurred in the rice producing areas of the alluvial plains and lower terraces, while agricultural extensification occurred in the middle and high terraces and upland sites of this low relief, undulating, and marginalized environment.

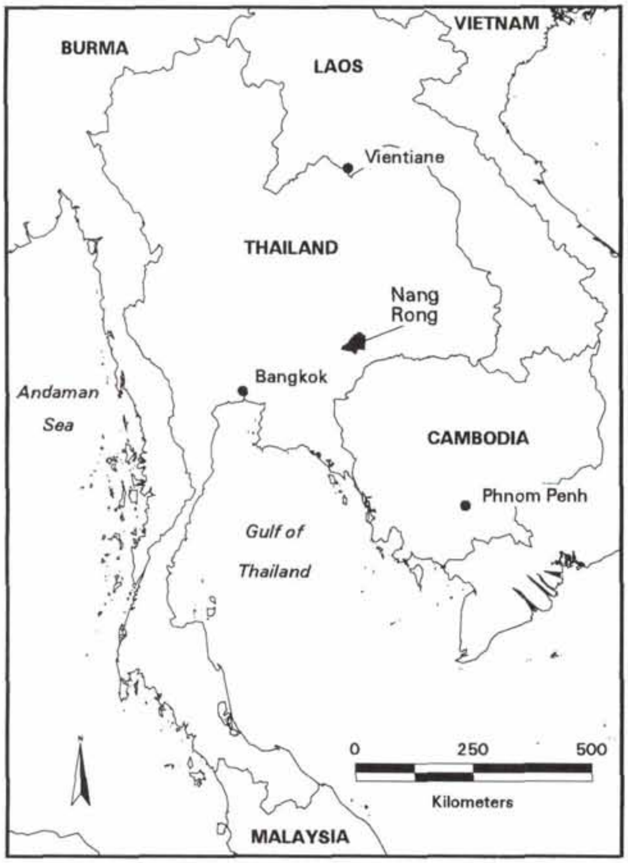

Figure 1. Study area location, Nang Rong district, northeast Thailand.

The Southeast Asian monsoon climate prevails in the northeast, with over 80 percent of the average annual precipitation occurring as unevenly distributed torrential rains during April to November (Kaida and Surarerks, 1984; Rigg, 1991). For the rest of the year, soil moisture deficits are common (Ghassemi *et al.*, 1995). Floods and droughts are a constant threat, adding even more risk to agriculture (Fukui, 1993).

The Khorat Plateau, within which Isan broadly coincides, is a wide, shallow basin underlain by Cretaceous sandstones, shales, and siltstones, though intruded in places by Tertiary basalts. Layers of rock salt and other salt-bearing strata are common. Soil types and their physical characteristics tend to be closely associated with specific landforms (e.g., terraces, hills, floodplains), and, hence, a relationship between topography and agricultural suitability exists within the region (Dixon, 1976; Parnwell, 1988; Fukui, 1993). Heavily leached fine sandy loams (e.g., the Khorat series) spatially predominate. Many areas have soils with poor drainage and/or low innate fertility. The natural vegetation of Isaan consists of a dry monsoon forest predominated by dwarf dipterocarp trees, and contains areas of grassland, thorny shrubs, and bamboo thickets (Parnwell, 1988). Except under spatially limited hydroclimatological conditions (e.g., along perennial streams), vegetation phenology is largely drought-controlled (Ghassemi *et al.*, 1995; Rundel and Boonpragop, 1995).

Data Sets and Methods

The data sets used in this research included (1) selected population variables from a comprehensive social survey of 310 villages in the Nang Rong district, northeast Thailand, conducted in 1994 at the community level; (2) a February 1993 Landsat Thematic Mapper (TM) digital data set; and (3) a constructed digital elevation model of the study area and derived biophysical coverages (i.e., elevation, slope angle, and soil moisture potential). The primary methods used included (1) processing of Landsat data to generate a Level-1 LULC classification and plant biomass levels for 1993 through the Normalized Difference Vegetation Index (NDVI); (2) spatial aggregation procedures to rescale the LULC, NDVI, and social and biophysical variables across nine scale steps extending from 30 to 1050 m; (3) development of a generalized model of population distribution from nuclear village centroids to the surrounding landscape by transforming discrete data to continuous data for subsequent analyses; (4) multiple-regression analysis (linear models) to examine the relationships between population and environment across the sampled spatial scales; and (5) canonical correlation to relate a "group" of population variables to a "group" of environment variables across the range of spatial scales.

Aggregation/Scale Levels

Nine scale steps (i.e., cell dimensions) were selected for analysis: 30, 90, 150, 210, 300, 450, 600, 900, and 1050 m. They represent the spatial resolution of the Landsat TM sensor at the finest spatial scale and the NOAA Advanced Very High Resolution Radiometer (AVHRR) sensor at the coarsest scale. The selected spatial scales also reflect landscape pattern and process relationships observed throughout the study area. For example, individual land clearings generated at the household level to expand agriculture at the expense of forests tend to form a fragmented landscape matrix extending across spatial scales. Clearings generated by individual households for upland agriculture (approximately 100 by 100 m in size) are surrounded by expanses of forest of varying densities, whereas individual rice paddies (approximately 10 by 20 m in size) coalescence into large tracts of cleared land with sparse forest remnants doting the landscape, a pattern that typifies lowland rice cultivation within the region. This pattern of deforestation and agricultural extensification, mediated by topography and access to water, produces a landscape characterized by local diversity and regional homogeneity. Rice paddies form the regional backdrop to the landscape, while forests and upland crops serve as counterbalances to the areal transition. Social processes related to the capability of a household to clear land and cultivate crops were hypothesized to function at fine spatial scales, biophysical processes influenced by environmental gradients were hypothesized to operate at coarse spatial scales, and social and biophysical interactions related to resource use were hypothesized to function at intermediate spatial scales.

Aggregations across the nine scale steps were generated by returning to the 30- by 30-m base surfaces to initiate all computations. For the LULC scaling, a grid-based plurality procedure was used where the most frequently occurring sub-grid cover type was used to code each grid cell at degraded spatial resolutions. The aggregated LULC data were used to stratify the scene for exploring the scale dependence associated with only specific LULC types. Pixels of NDVI were aggregated to generate small blocks of a finer resolution image, with their own degrees of variability and spatial pattern. The NDVI data were calculated from corrected spectral radiance values for each aggregation level by assuming that each TM (or resampled MSS) pixel located within aggregated pixels contributed equally to the areally averaged value (i.e., flat-field aggregation), without probability values or other weights assigned to pixel locations or LULC types. Each portion of the image that was defined by the boundaries of an aggregated pixel were treated as a separate block. Elevation data were aggregated in the same manner.

Satellite Digital Data

Following standard preprocessing to correct for geometric, atmospheric, and radiometric corrections, the TM (01 February 1993) data set was classified through a supervised approach into generalized Level-1 categories: forest, water, rice, and upland agriculture. A 20 December 1988 Landsat TM image was integrated into the classification process to confirm the distinction between rice and pasture in semi-permanent rice paddies, because of the image's representation of the landscape during an alternate phenological stage associated with the cultivation of rice. NDVI values were calculated from corrected spectral radiance values for each aggregation level from the February 1993 TM scene. NDVI reduces multiple channels of spectral data to a single number per pixel (after Rouse *et al.* (1973)) that estimates a number of canopy characteristics (e.g., plant biomass and leaf area). NDVI is defined through an empirically derived relationship between vegetation characteristics and spectral responses in the visible, near-infrared, and middle-infrared spectral regions (Jensen, 1996; Schott, 1997).

Persistent and extensive cloud coverage associated with monsoonal rains precluded the use of TM images during the rice growing period. This February image date has implications for subsequent analyses as well as for interpretations that should be noted. In February, the rice has been harvested and stubble remains in the field. Depending upon the nearness to water, harvested rice paddies may contain grass of varying densities. While the phenology and hence its spectral response varies by crop type, rice areally dominants the landscape with particular bias on the alluvial plains and low terraces. At this time of year, the landscape is dry, visually brown in color, and agricultural endeavors have become more focused on the upland sites where primarily cassava and sugar cane are grown. The cassava uplands may be vegetated or completely barren, if just harvested or planted. The native forests that remain are only remnants of past extents primarily located along streams, surrounding nuclear village compounds, sparsely distributed (dotting) in rice paddies, and in upland sites that are geographically remote from villages, on marginal sites, or yet unattended for agricultural purposes. Such forested areas, along with some areas of conservation forests, serve to interdigitate the rice-dominated landscape with alternate cover types, creating a fragmented landscape.

From a biomass gradient perspective, February offers a landscape view where rice paddies are relatively low in NDVI values, upland agriculture relatively high in NDVI values (low if harvested), and forests intermediate in value. As part of the culture in northeast Thailand, fire is often used to reduce rice paddy stubble and to aid in site fertilization, among other objectives. Therefore, the rice areas, normally rich in chlorophyll and biomass during the growing season, are in February typical of a senescent landscape, thereby creating strong negative relationships of biomass to rice paddies or low correlations between total land under cultivation (as reported by the social survey) and the calculated NDVI values for dominant agricultural (rice) villages. Within the study area (and during this time of year), NDVI values likely increase with elevation, because senescent rice paddies are at lower elevations and forests and upland agriculture are primarily at the higher elevations. Contrary to that general relationship is the existence of isolated and sparsely clustered trees growing within and adjacent to rice paddies, thereby elevating the NDVI values in the post-harvest, dry period characterized by the February TM image. Therefore, sign reversals indicating the relationship between rice (a normally high biomass cover type during the growing and pre-harvest period) and NDVI are a consequence of the then limited availability of TM images assembled for this study.

Because of the pronounced wet and dry periods throughout the study area, the landscape is continually undergoing a cycle of change, from a lush and vigorous rice-dominated landscape in November and December, to a dry, stubbled landscape following the rice harvest and extending into the early spring, and returning to a rice producing landscape with the arrival of the monsoonal rains typically in May or June. While the February 1994 Landsat TM image reflects a post-rice harvest landscape, a simple reversal of relationships between Landsat TM spectral values and land-cover types is noted to reflect the realization that the land-cover classification may indicate rice but the land-use implication is for a stubble field of harvested rice and/or a dry pasture during the February image date.

Biophysical Data

Four biophysical variables (their codes are also indicated) were generated for this analysis through a combination of remote sensing and GIS techniques: plant biomass (ndvi), elevation (elev), slope angle (slope), and soil moisture potential (wet). Plant biomass was measured through the calculation of the NDVI for the February 1993 TM image, and elevation and slope angle values were generated through the construction of a digital elevation model (DEM) of the study area from 1:50,000-scale (contour interval of 20 m), 1984 topographic base maps. The contour lines and point elevations from the base maps were scan-digitized, attributed, edge-matched, and processed to yield a DEM. The hydrography coverage, previously digitized from the same set of base maps, was used for hydrologic enforcement in the generation of the DEM. Elevation and slope angle values were transformed (1/square root) to reduce multicollinearity in the analyses, following their initial calculation. Soil moisture potential was estimated through the DEM using an index of saturation potential developed by Beven and Kirkby (1979), and calculated by using flow directions and accumulation grids based on eight directions for pixel-to-pixel flow routing after Moore *et al.* (1991). Topographic wetness potential provides a simple but effective model that characterizes channels, gullies, and moisture sinks (Allen and Walsh, 1996). Townsend and Walsh (1996; 1998) used such an approach to derive and compare soil moisture potential values generated through a variety of terrain and soil measurement approaches. A digital hydrography coverage was used to support the calculation of the index. The wetness index used calculations of slope and upslope contributing area to derive a dimensionless index of potential wetness.

Social Survey Data

The data used in this study were collected in 1994 for all 310 villages that comprised the district. While an extensive array of variables were collected, only a small subset were used for this preliminary study. The variables and their codes used in subsequent analyses are as follows: total population (pop), total land under cultivation (land), total land under cultivation/total population (lapo), sex ratio (sexr), and the number of households (hous). The information was collected as part of a group interview of the headman, members of the village community, and other knowledgeable persons and was conducted in each village. The population data represented reasonably accurate counts, as headmen are asked to keep track of this information for other administrative purposes.

The population variables were selected to represent various supply and demands considerations as they relate to labor and the capacity to alter the landscape through work: (1) population is a consumption sink in a subsistence environment and a production reservoir where extensification and intensification of agriculture occur; (2) total land under cultivation is an indication of the resource base used in agricultural production; (3) ratio of total land under cultivation and total population is a density indicator of population to land under cultivation; (4) sex ratio reflects gender-based job sorting, primarily the predominance of men in the use of mechanization; and (5) number of households is the landscape work unit where local decision-making at the household level affects LULC patterns. To reduce multicollinearity in the analyses, the ratio of total land under cultivation and total population was transformed (1/square root) after ratioing.

Village Territories

The villages within the Nang Rong district are spatially concentrated or nuclear in structure. Households are areally clustered and agricultural lands are arrayed in a generally circular pattern surrounding the village centroid depending upon village adjacencies, ownership/use patterns within the village and between villages, and hydrographic and transportation network characteristics, among other factors. The general conceptual model used in the setting of village territories is for farmers to leave the village compound each day and walk/ride to their fields to perform various activities. They return each day to the village household. During various periods of the year (e.g., rice planting and harvesting), however, a small percentage of farmers may temporarily live in their fields to reduce time/cost associated with commuting between the primary household and work sites.

With land title unclear and political boundaries relatively unimportant within this context, approaches for estimating village territories have been examined. Evans (1998) explored the partitioning of space through region growing techniques that involved fuzzy logic and transitional boundaries. In addition, radial buffers of varying dimensions (Entwisle *et al.*, 1998; Walsh *et al.*, 1998) as well as Thiessen polygons (Evans, 1998) have been used in our research to set village territories for addressing various research questions. Here, a 5-km radial buffer was selected to define the boundary around the 310 nuclear villages. The buffer dimension was set through knowledge of land ownership patterns and modes of travel within the district as of 1994. An objective of the territory-setting process was to assign as much land as possible in the district to one of the 310 villages while minimizing the amount of land assigned to district villages outside the district boundary. Iterative partitioning of space through different buffer dimensions was tested and descriptive statistics were computed before the 5-km boundary dimension was selected. At that size buffer, approximately 98 percent of the district land was associated to one of the 310 villages. A 10-km buffer around the district was used to subset the satellite data and other data sets so that village territories could extend across district boundaries.

Population Distribution Model

Population data collected as part of the 1994 community survey for the district reported household information at the village level. Recalling that the villages are arranged in some form of a nuclear pattern, the survey data were geocoded at the centroid of each village compound. Therefore, the social data were all represented at discrete point locations, whereas the biophysical data, collected by satellite sensors at pixel locations (e.g., LULC and NDVI values) and represented in the GIS in a raster model (e.g., topographic variables), were represented as a continuous distribution across the study area in a regular matrix format. As a result, the social data had to be cartographically transformed to be spatially compatible with the spectral and biophysical data for subsequent aggregation and scaling.

A simple population distribution model was used to link people to the land in an areal fashion. Using the LULC classification for 1993, the landscape was stratified into agriculture (i.e., rice and upland agriculture, primarily cassava and sugar cane) and non-agriculture categories. In designated agricultural portions of the district, the selected population variables were distributed on a per-pixel basis (depending upon the scale step) based upon an equal value spread function. Operationally, the number of agriculture cells within the derived village territories was defined and the available total population (reported by the social survey), for example, was equally distributed on a village by village basis within the defined village territories. Villages "claiming" identical areas of the landscape within their assigned territories generated higher population totals at the cellular level as a consequence of the equal value decision rule.

Multiple-Regression Analysis

The objective of the statistical analysis was to examine, through multiple-regression analysis, the relationships between the above-described set of population and environment variables as descriptors of "lapo" (ratio of total land under cultivation and total population) across nine scale steps ranging from 30 to 1050 m. Table 1 summarizes the regression statistics for "lapo" (population density relative to cultivated land; the population dependent variable) against

Table 1. Regression Summary Across Scales: Land/Pop (Lapo) Dependent Variable

Scale (m)	n	Var	Slope	Std. Error	t-value	P > |t|	Adj. R^2
30	10,532						
		elev	0.00925	0.00047	19.63	0.0001	0.444
		slope	0.01468	0.01647	0.89	0.3731	
		wet	−0.00132	0.00041	−3.25	0.0011	
		ndvi	1.01975	0.08933	11.42	0.0001	
		sexr	9.94295	0.22138	44.91	0.0001	
		hous	−40.06327	1.24442	−32.19	0.0001	
90	10,918						
		elev	0.00889	0.00048	18.71	0.0001	0.456
		slope	0.06316	0.02445	2.58	0.0098	
		wet	−0.00114	0.00044	−2.57	0.0102	
		ndvi	1.49646	0.09083	16.47	0.0001	
		sexr	9.55451	0.21712	44.01	0.0001	
		hous	−4.82307	0.13675	−35.27	0.0001	
150	11,456						
		elev	0.00801	0.00048	16.69	0.0001	0.456
		slope	0.09759	0.02979	3.28	0.0011	
		wet	−0.00129	0.00048	−2.71	0.0067	
		ndvi	1.85586	0.09731	19.07	0.0001	
		sexr	9.37788	0.20894	44.88	0.0001	
		hous	−1.81487	0.05027	−36.11	0.0001	
210	11,757						
		elev	0.00798	0.00049	16.29	0.0001	0.459
		slope	0.17069	0.03587	4.76	0.0001	
		wet	−0.00144	0.00050	−2.86	0.0043	
		ndvi	1.79976	0.09875	18.23	0.0001	
		sexr	9.38728	0.21024	44.65	0.0001	
		hous	−0.92587	0.02615	−35.59	0.0001	
300	8,362						
		elev	0.00683	0.00061	11.26	0.0001	0.460
		slope	0.29425	0.05024	5.86	0.0001	
		wet	−0.00133	0.00063	−2.10	0.0355	
		ndvi	2.37762	0.13292	17.89	0.0001	
		sexr	8.91536	0.25020	35.63	0.0001	
		hous	−0.45625	0.01543	−29.56	0.0001	
450	3,823						
		elev	0.00530	0.00093	5.69	0.0001	0.469
		slope	0.54203	0.09222	5.88	0.0001	
		wet	−0.00018	0.00099	−0.18	0.8572	
		ndvi	3.01379	0.20856	14.45	0.0001	
		sexr	8.67822	0.36547	23.75	0.0001	
		hous	−0.19976	0.01032	−19.36	0.0001	
600	2,204						
		elev	0.00335	0.00131	2.55	0.0107	0.475
		slope	0.82158	0.13953	5.89	0.0001	
		wet	0.00037	0.00135	0.28	0.7821	
		ndvi	3.53473	0.29930	11.81	0.0001	
		sexr	8.50133	0.46504	18.28	0.0001	
		hous	−0.11217	0.00762	−14.72	0.0001	
900	1,020						
		elev	0.00039	0.00201	0.20	0.8436	0.504
		slope	1.36684	0.24587	5.56	0.0001	
		wet	−0.00070	0.00213	−0.33	0.7419	
		ndvi	4.19461	0.48152	8.71	0.0001	
		sexr	8.55094	0.70804	12.08	0.0001	
		hous	−0.05163	0.00527	−9.80	0.0001	
1050	758						
		elev	−0.00194	0.00243	−0.80	0.4236	0.503
		slope	1.51691	0.30102	5.04	0.0001	
		wet	−0.00186	0.00251	−0.74	0.4581	
		ndvi	4.83686	0.60673	7.97	0.0001	
		sexr	7.86407	0.83377	9.43	0.0001	
		hous	−0.03789	0.00451	−8.40	0.0001	

a set of descriptor variables (elevation (elev), slope angle (slope), soil moisture potential (wet), plant biomass (ndvi), sex ratio (sexr), and total number of households (hous); total land under cultivation (land) and total population (pop) were excluded from the analysis because "lapo" was derived through their ratio). The table shows the sample size of cells used to compute the regressions, slope coefficients for each variable in the model, standard error, t-values, p-values, and the adjusted R^2 for each of the nine scale steps. Recall that some variables were transformed to reduce multicollinearity, but correlation between variables still exists. Also, note that the sample size changes with scale and that affects standard

Table 2. Canonical Coefficients for Population Variables: Association of Population Variables to Population Axes

(a) Canonical Coefficients (30 by 30 m)

Variable	Axis 1	Axis 2	Axis 3
land	−0.0924	1.9742	0.2690
lapo	0.5798	−1.4504	−1.3019
pop	−2.5506	−0.3573	−1.8045
sexr	0.4429	−0.1708	0.4731
hous	2.7585	−1.7839	0.6913

(b) Canonical Coefficients (150 by 150 m)

Var	Axis 1	Axis 2
land	0.2147	−1.8202
lapo	0.3726	1.2873
pop	2.7338	−0.1389
sexr	0.3664	0.2957
hous	2.6568	2.1912

(c) Canonical Coefficients (1050 by 1050 m)

Var	Axis 1	Axis 2	Axis 3
land	0.7839	−1.2210	−1.4972
lapo	−0.0498	0.9273	1.9552
pop	−2.4210	−1.7370	1.3213
sexr	0.2590	0.5312	−0.8629
hous	1.6362	3.4231	−0.3748

errors, t-values, and p-values. These statistics are still presented, but they should be interpreted as general indicators. A random sample of cells was instituted at the fine grain scales to reduce computational intensity of the regression runs by capping the sample size below 12,000 cells, or approximately 5 percent of the total available cells.

The descriptor variables used in the "population" model runs measure some similar effects. For example, slope angle and soil moisture potential are associated in that they reflect the gravity flow potential of materials: i.e., water and soil. In this environment, water is the most important "flow" resource, but soil creep and redistribution of sediment along topographic gradients occur through related biophysical processes. Also, the social variables are similarly related as to effect. Total population and number of households occurring within a village are related in that they reflect labor supply and resource demands, but total population is distributed across the non-water landscape and the number of households are used to describe local decision-making relative to LULCC. In short, a number of the descriptor variables are statistically related and associated by effects to some degree. Transformations involving factor analysis and principal components analysis could be used to transform such variables, but interpretation becomes less direct and focused. Also, other variables will be selected in subsequent scale-dependent analyses to reduce the similarity of effects represented through them, but the variables included here offer an initial view of scaling properties to suggest the importance of the scale-pattern-process paradigm applied to an integrated social and biophysical landscape through a focus on scaling approaches applicable to population-environment studies.

Canonical Correlation Analysis

Canonical correlation analysis is used here to examine the relationship between a group of population variables (total population (pop), total land under cultivation (land), ratio of total land under cultivation and total population (lapo), sex ratio (sexr), and total number of households (hous)) and a group of environment variables (elevation (elev), slope angle (slope), plant biomass (ndvi), and soil moisture potential (wet)) across the range of spatial scales sampled. A subset of the nine scale steps are presented in the output from the canonical analyses to conserve space, but analyses were conducted for the entire range of scales, 30 to 1050 m.

Tables 2 and 3 show the association of the population variables and the environment variables to the derived population and environment axes (statistically significant at the 0.001 level) at the 30-, 150-, and 1050-m scale steps, respectively. The underlined canonical coefficients indicate the dominant environment and population variables for the 30-, 150-, and 1050-m scales. Table 2 indicates, for example, that the number of households (hous) is most associated with axis 1 at the 30-m scale step, total land under cultivation (land) is most associated with axis 2, and the ratio of total land under cultivation and total population (lapo) is most associated with axis 3. Table 3 shows, for example, that elevation (elev) is most associated with axis 1 at the 30-m scale step, plant biomass (ndvi) is most associated with axis 2, and soil moisture potential (wet) is most associated with axis 3. The variation in variable-axis relationships as a consequence of scale variation are readily apparent in Tables 2 and 3.

Results and Discussion

The following describes some of the results achieved through the analyses described above, and interpretations are made only for selected items. This section is centered around two areas—multiple-regression analyses for a selected population variable (ratio of total land under cultivation and total population (lapo)) across the nine scale steps, and canonical analysis for the group of environment and population variables across the entire range of sampled scales, but reported here only for a scale subset.

Multiple-Regression Models

The multiple-regression models for the population variable, "lapo," at each of the nine scale steps are presented in Table 1. While the sign of the slope coefficient is related to the February 1993 date of the Landsat TM data set (a senescent rice landscape at the time of imaging, but an actively growing upland crop and forested landscape) used to characterize NDVI values across the landscape, the magnitude of the slope coefficients suggests the relative importance of each variable to the derived regression model at each of the sampled scales. Recall that the scale ranges were selected to represent hypothesized fine grain social processes and coarse grain biophysical processes affecting plant biomass values. Figure 2 shows plots of the slope coefficients for each of the six de-

Table 3. Canonical Coefficients for Environmental Variables: Association of Environmental Variables to Environmental Axes

(a) Canonical Coefficients (30 by 30 m)

Var	Axis 1	Axis 2	Axis 3
ndvi	0.2208	0.8236	0.1317
elev	0.9451	−0.5915	−0.0594
slope	0.0349	0.3784	0.7682
wet	0.0898	−0.3727	1.0618

(b) Canonical Coefficients (150 by 150 m)

Var	Axis 1	Axis 2
ndvi	0.3870	−0.7530
elev	0.8266	0.7246
slope	0.1018	−0.5000
wet	0.1245	0.2385

(c) Canonical Coefficients (1050 by 1050 m)

Var	Axis 1	Axis 2	Axis 3
ndvi	0.7337	−0.5441	−0.3840
elev	0.0923	1.2758	−0.9401
slope	0.4838	−0.2601	1.5440
wet	0.0450	0.4845	0.2519

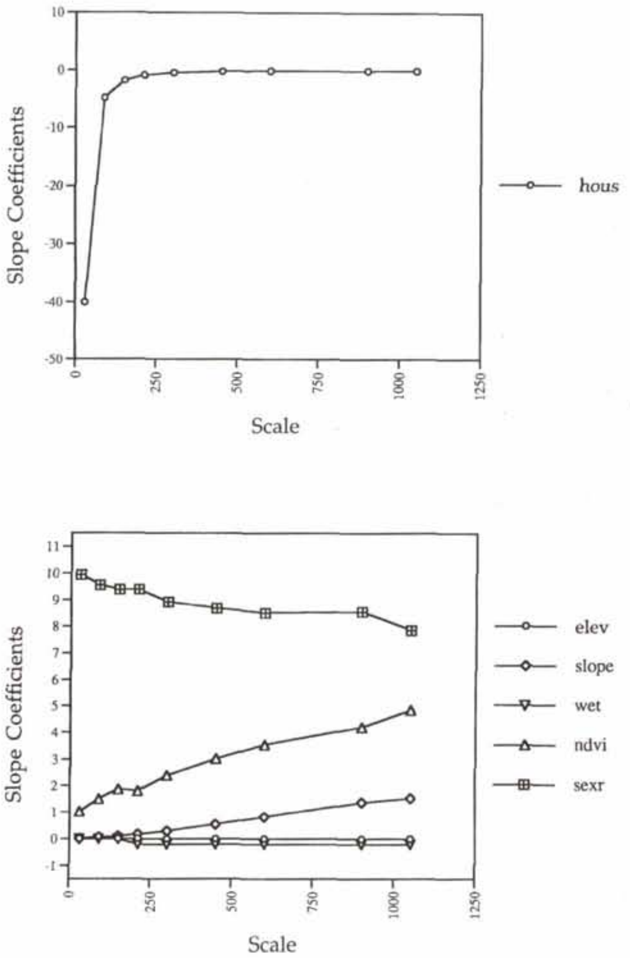

Figure 2. Plots of slope coefficients for the population and environment variables.

scriptor variables used in the "lapo" models at each of the nine scale steps, and Figure 3 shows the variation in the adjusted R^2 values across scales. The derived models indicate (1) the influence of plant biomass (ndvi) increases with scale; (2) the influence of slope angle (slope) also increases with scale; (3) number of households (hous) is important at fine scales, but decreases in importance at coarser scales; and (4) the sex ratio (sexr) is consistently strong across scales, and is the strongest variable beyond the finest scales sampled; the greater number of males at the village level are related to the larger amount of land under cultivation. The adjusted R^2 values indicate an increasing upward trend with scale. The nature of the distances that separate adjacent villages likely affects the higher R^2 values with scale, because the transitional boundaries between the villages are smoothed as a result of the spatial aggregation associated with the coarser scale steps.

Canonical Correlation

Canonical structure shows the interpretation of effects and association of variables through loadings. Tables 4 and 5 show the canonical structure for the 30-, 150-, and 1050-m scales associated with the generated axes for the environment variables and the population variables, respectively. The tables show the association of environment variables to population variables through loadings. The highest loading and the associated variable is underlined on each of the tables. According to Table 4, the ratio of the total land under cultivation and total population (lapo) is most associated with elevation (elev) (axis 1), the total number of households (hous) is most associated with plant biomass (ndvi) (axis 2), and the total land under cultivation (land) is most associated with soil moisture potential (wet) (axis 3) at the 30-m scale. At the 150-m scale, the ratio of total land under cultivation and total population (lapo) is most associated with elevation (elev) (axis 1), and the total number of households (hous) is most associated with plant productivity (ndvi) (axis 2). At the 1050-m scale, the ratio of total land under cultivation and total population (lapo) is most associated with plant productivity (ndvi) (axis 1), the total number of households (hous) is most associated with elevation (elev) (axis 2), and the sex ratio is most associated with slope angle (slope) (axis 3). According to Table 5, at the 30-m scale, elevation (elev) is most associated with the ratio of total land under cultivation and total population (lapo) (axis 1), plant productivity (ndvi) is most associated with the total number of households (hous) (axis 2), and soil moisture potential (wet) is most associated with the total land under cultivation (land) (axis 3). At the 150-m scale, elevation is most associated with the ratio of total land under cultivation and total population (lapo) (axis 1), and plant productivity (ndvi) is most associated with the total number of households (hous) (axis 2). At the 1050-m scale, plant productivity (ndvi) is most associated with the ratio of total land under cultivation and total population (lapo) (axis 1), elevation (elev) is most associated with the total number of households (hous) (axis 2), and slope angle (slope) is most associated with the sex ratio (sexr) (axis 3).

The statistical results indicate an "upland" effect associated with the site suitability for cassava, a crop grown as animal fodder for European markets. Cassava (and sugar cane to a lesser degree) was the primary crop linked to deforested land within the district in the past half century. It is generally restricted to the higher elevations and the drier sites. Depending upon the time of year and market conditions, upland crops are in various stages of development or the fields are being prepared for re-planting. Plant productivity levels track relative to growth cycles, and social variables, particularly the sex ratio, are associated with villages with a significant amount of their prescribed village territory growing upland crops. The bias towards male-dominated use of mechanization for cassava cultivation is suggested in the importance of the sex ratio variable in villages of such predominance. Soil moisture potential, a variable of immense

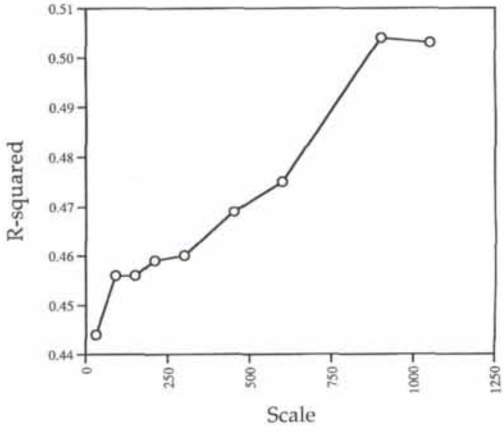

Figure 3. Plots of the R^2 values for the lapo model.

TABLE 4. CANONICAL STRUCTURE FOR ENVIRONMENTAL VARIABLES: ASSOCIATION OF ENVIRONMENT TO POPULATION THROUGH LOADINGS

(a) Loadings (30 by 30 m)			
Var	Axis 1	Axis 2	Axis 3
ndvi	0.4132	0.7502 (hous)	0.1256
elev	0.9728 (lapo)	−0.1532	−0.1082
slope	0.3646	0.3914	0.2708
wet	−0.2606	−0.3847	0.7243 (land)
(b) Loadings (150 by 150 m)			
Var	Axis 1	Axis 2	
ndvi	0.5751	−0.6699 (hous)	
elev	0.9197 (lapo)	0.3076	
slope	0.5155	−0.3344	
wet	−0.2831	0.3132	
(c) Loadings (1050 by 1050 m)			
Var	Axis 1	Axis 2	Axis 3
ndvi	0.8552 (lapo)	−0.3028	−0.3796
elev	0.6606	0.6782 (hous)	0.0149
slope	0.6823	0.3197	0.5869 (sexr)
wet	−0.4129	0.1098	−0.1505

importance in the region, reflects the paucity of moisture in upland sites and the juxtaposition of hydrologic features to rice paddies. The slope angle variable indicates the relevance of topography in discriminating between the relatively flat lowland and upland agriculture and the low-middle-high terraces that exist as the transitional landscape. Plant biomass indicates the separation of landscapes on the basis of regional crop phenologies and the location of relatively dense forests in upland locations and sparse and discontinuous forests distributed across the rice-producing lowlands. The social variables interact with the biophysical variables through their ability to alter the landscape through deforestation and agricultural extensification. The density of population on cultivated land, the number of households in a village making decisions about LULC, and the topologic relationship between villages and the resource endowments of their territories affect the complex interactions between population and the environment that are scale dependent.

Conclusions

Many landscapes are the product of social and biophysical processes that are related to spatial (and/or temporal) patterns through scale-dependent relationships. The grain and extent of landscape measurement and observation are central to the interrelationship of scale, pattern, and process. In northeast Thailand, social and biophysical processes associated with agricultural extensification and intensification are altering the landscape through deforestation, road building, and hydromodifications in response to regional, national, and global economies, and regional to national migration patterns. Biophysical, geographic, and social factors are combining in complex ways to alter LULC patterns as a consequence of resource endowments and environmental gradients, accessibility and site marginalization, and labor and kinship relationships, among other factors. Such processes may influence the composition and spatial organization of the landscape across time and space and link across domains in interesting and confounding ways.

The nature of the study area landscape offers some interesting insights into the effects of spatial scaling. In the rice growing areas, riparian forests exist along many water courses, some conservation or protected forests exist in the lowlands, and sparsely distributed forests are scattered throughout the paddies. The backdrop to the lowlands is clearly and definitely rice, and depending upon the time of year and actual site conditions, the forests become less significant as a results of a larger cell size. Therefore, their impact within the integrated pixel concept becomes diminished with scale. In the upland sites, cassava fields are large and rather homogeneous (more so than rice because of the scattered trees and differential access to water). They are, however, surrounded by other upland crops, eucalyptus forest, and secondary growth forests generally restricted to hilly, infertile, or remote sites. Therefore, cell grading from finer to coarser dimensions reflect an increasingly vegetated landscape, particularly during the February time period, because of the areal extent of forests relative to upland crops sites that may be non-vegetated as of the February time period.

With respect to our guiding hypotheses, the data suggest that social drivers are associated with finer-scale spatial patterns and that biophysical drivers are associated with coarser-scale spatial patterns. Relationships between sets of biophysical and social variables change as a function of spatial scale. Much remains to be accomplished in our study of scale dependence as applied to an environment strongly impacted by human and biophysical processes. Determining important relationships across scale and defining the effects to be scaled as inputs to local, regional, and global models of LULCC are significant challenges. Initial steps include the integration of social and biophysical variables to represent drivers of LULCC, deriving statistical relationships across space and time scales, and interpreting results relative to existing theory and through techniques documented in associated areas of scientific inquiry. GIS, remote sensing, social surveys, and spatial and statistical analyses offer a research synergism suitable to addressing questions related to population and environment interactions in numerous environments, including our site in northeast Thailand.

The results suggest that the relationships between biophysical and social variables change as the scale at which one examines these relationships changes. This is an important result in light of the large volume of LULC research currently underway by a variety of research teams around the world. Based on preliminary research reported at meetings, it is clear that different teams are operating at different spatial scales. Sometimes this is the result of using data from different sensors (AVHRR versus TM, for example) or sometimes this is the result of the social data being available at different

TABLE 5. CANONICAL STRUCTURE FOR POPULATION VARIABLES: ASSOCIATION OF POPULATION TO ENVIRONMENT THROUGH LOADINGS.

(a) Loadings (30 to 30 m)			
Var	Axis 1	Axis 2	Axis 3
land	0.2624	0.3511	−0.8049 (wet)
lapo	0.7078 (elev)	0.4943	−0.3244
pop	−0.4178	−0.4061	−0.5797
sexr	0.5806	0.3857	0.2489
hous	−0.2570	−0.5295 (ndvi)	−0.5347
(b) Loadings (150 to 150 m)			
Var	Axis 1	Axis 2	
land	0.3297	−0.3230	
lapo	0.7520 (elev)	−0.4514	
pop	−0.4251	0.3760	
sexr	0.5808	−0.2694	
hous	−0.2732	0.5135 (ndvi)	
(c) Loadings (1050 to 1050 m)			
Var	Axis 1	Axis 2	Axis 3
land	0.3796	−0.0972	−0.3045
lapo	0.8732 (ndvi)	−0.1708	0.3011
pop	−0.5708	0.2699	−0.3562
sexr	0.6600	0.0037	−0.3568 (slope)
hous	−0.4932	0.4401 (elev)	−0.3155

scales. Whatever the reason dictating why one research team might operate at a fairly fine scale and another team at a coarser scale, the results of the present paper suggest that research teams operating at different scales will likely reach different conclusions.

Acknowledgments

The work reported in this paper is supported by the National Aeronautics and Space Administration under its Land-Cover and Land-Use Change program (NAG5-6002). This paper is part of a larger set of interrelated projects funded by the National Institute of Child Health and Human Development (R01-HD33570 and R01-HD25482), the National Science Foundation (SBR 93-10366), the EVALUATION Project (USAID Contract #DPE-3060-C-00-1054), and the MacArthur Foundation (95-31576A-POP). The larger set of projects involves various collaborations between investigators at the University of North Carolina, Carolina Population Center and the Department of Geography and investigators at the Institute for Population and Social Research (IPSR), Mahidol University, Bangkok, Thailand. For this research, Phil Page, John Vogler, Tom Crawford, Rick O'Hara, and Erika Stone at the University of North Carolina and Thomas R. Allen, Old Dominion University, provided programming, statistical analysis, and general assistance with elements of this research.

References

Allen, T.R., and S.J. Walsh, 1996. Spatial and compositional pattern of alpine treeline, Glacier National Park, Montana, *Photogrammetric Engineering & Remote Sensing*, 62(11):1261–1268.

Beven, K.J., and M.J. Kirkby, 1979. A physically-based variable contributing area model of basin hydrology, *Hydrological Sciences Bulletin*, 24(1):43–69.

Bian, L., and S.J. Walsh, 1993. Scale dependencies of vegetation and topography in a mountainous environment in Montana, *The Professional Geographer*, 45(1):1–11.

Blaikie, P., and H. Brookfield, 1987. *Land Degradation and Society*, Methuen, London, 296 p.

Dixon, C.J., 1976. Settlement and environment in northeast Thailand, *The Journal of Tropical Geography*, 46:1–10.

Entwisle, B., S.J. Walsh, R.R. Rindfuss, and A. Chamratrithirong, 1998. Landuse/landcover (LULC) and population dynamics, Nang Rong, Thailand, *Proceedings, People and Pixels*, National Academy of Science/National Research Council, pp. 121–144.

Evans, T.P., 1998. *Integration of Community-Level Social and Environmental Data: Spatial Modeling of Community Boundaries in Northeast Thailand*, doctoral dissertation, Department of Geography, University of North Carolina, Chapel Hill.

Feeny, D., 1988. Agricultural expansion and forest depletion in Thailand, 1900–1975, *World Deforestation in the Twentieth Century* (J.F. Richards and R.P. Tucker, editors), Duke University Press, Durham, North Carolina, pp. 112–143.

Forman, R.T.T., and M. Godron, 1986. *Landscape Ecology*, John Wiley & Sons, New York, 620 p.

Fukui, H., 1993. *Food and Population in a Northeast Thai Village*, Monographs of the Center for Southeast Asian Studies, Kyoto University, English-Language Series, No. 19, University of Hawaii Press, Honolulu, 421 p.

Ghassemi, F., A.J. Jakeman, and H.A. Nix, 1995. *Salinisation of Land and Water Resources: Human Causes, Extent, Management and Case Studies*, University of New South Wales Press Ltd., Sydney, 526 p.

Jensen, J.R., 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective, Second Edition*, Prentice-Hall, Englewood Cliffs, New Jersey, 316 p.

Johnston, R.J., D. Gregory, and D.M. Smith (editors), 1994. *The Dictionary of Human Geography, Third Edition*, Blackwell Publishers Ltd., Oxford, 724 p.

Johnston, R.J., P.J. Taylor, and M.J. Watts (editors), 1995. *Geographies of Global Change: Remapping the World in the Late Twentieth Century*, Blackwell, Oxford, 462 p.

Kaida, Y., and V. Surarerks, 1984. Climate and Agricultural Land Use in Thailand, *Climate and Agricultural Land Use in Monsoon Asia* (M.M. Yoshino, editor), University of Tokyo Press, Tokyo, pp. 231–254.

Meentemeyer, V., and E.O. Box, 1987. Scale Effects in Landscape Studies, *Landscape Heterogeneity and Disturbance* (M.G. Turner, editor), Springer-Verlag, New York, pp. 15–36.

Moore, I.D., R.B. Grayson, and A.R. Landson, 1991. Digital terrain modeling: A review of hydrological, geomorphological, and biological applications, *Hydrological Processes*, 5:3–30.

Nellis, M.D., and J.M. Briggs, 1989. The effect of spatial scale on Konza landscape classification using textural analysis, *Landscape Ecology*, 2(2):93–100.

Parnwell, M.J.G., 1988. Rural poverty, development and the environment: the case of North-East Thailand, *Journal of Biogeography*, 15:199–313.

Quattrochi, D.A., and M.F. Goodchild (editors), 1997. *Scale in Remote Sensing and GIS*, Lewis Publishers, New York, 406 p.

Rigg, J., 1991. Homogeneity and heterogeneity: An analysis of the nature of variation in northeastern Thailand, *Malaysian Journal of Tropical Geography*, 22(1):63–72.

Rouse, J.W., R.H. Haas, J.A. Schell, and D.W. Deering, 1973. Monitoring vegetation systems in the Great Plains with third ERTS, *ERTS Symposium*, NASA No. SP-351, pp. 309–317.

Rundel, P.W., and K. Boonpragop, 1995. Dry forest ecosystems of Thailand, *Seasonally Dry Tropical Forests* (S.H. Bullock, H.A. Mooney, and E. Medina, editors), Cambridge University Press, New York, pp. 93–123.

Schott, J.R., 1997. *Remote Sensing*, Oxford University Press, New York, 394 p.

Townsend, P.A., and S.J. Walsh, 1996. Spatial variability of a wetness model to soil parameter estimation approaches, *Earth Surface Processes and Landforms*, 21(4):307–326.

———, 1998. Floodplain inundation modeling using an integrated GIS with radar and optical remote sensing, *Geomorphology*, 21(3–4):295–312.

Walsh, S.J., T.P. Evans, W.F. Welsh, R.R. Rindfuss, and B. Entwisle, 1998. Population and environmental characteristics associated with village boundaries and landuse/landcover patterns in Nang Rong district, Thailand, *Proceedings, PECORA 13 Symposium, Human Interactions with the Environment: Perspectives from Space*, pp. 395–404.

Walsh, S.J., A. Moody, T.R. Allen, and D.G. Brown, 1997. Scale dependence of NDVI and its relationship to mountainous terrain, *Scale in Remote Sensing and GIS* (D.A. Quattrochi and M.F. Goodchild, editors), CRC Lewis Publishers, Boca Raton, Florida, pp. 27–55.