From Statistical Data to Spatial Knowledge – informing decision-making in Vietnam

The display and analysis of spatial information is indispensable to generate knowledge about the location of objects, about spatial clusters, and relationships that informs decision-makers and researchers in Vietnam.

Michael Epprecht, Andreas Heinimann, Nicholas Minot, Daniel Müller and Tim Robinson

INTRODUCTION

Over the past fifteen years, Vietnam has made great progress in the field of socio-economic and agricultural data collection. In addition to a large number of other surveys, a total of five national living standards surveys, a population and housing census and three agricultural censuses have been conducted. While these data have been analysed statistically and made partially available in aggregated tabular form, few attempts have been made at spatial analysis or representation, or to make them more readily accessible to users. The potential of modern desktop geographic information systems (GIS) to analyse these tabular data, and to translate them into more readily accessible information, has thus far been largely untapped. Visualizing geographical variation in statistical data tremendously enhances their value, especially in heterogeneous regions such as Vietnam: the long narrow coastline, over 2,000 km, with two large river deltas, contrasting with a vast mountainous hinterland. This diversity is reflected in its socio-cultural, socio-economic and agricultural characteristics, which can be much better comprehended when viewed as maps. Socio-economic data in Vietnam is, as in many countries in a similar state of development, rarely available as spatially explicit information, which prevents the implementation of spatial analytical approaches that link the socio-economic realities with the local environment. The authors have therefore been involved in a number of initiatives to translate national statistics into spatially highly disaggregated maps, and to add value through spatial analysis. The availability of detailed socio-economic spatial data sets allows exploiting the integrative nature of GIS to explore relations between people and the environment at high spatial resolutions. Spatial analysis of such human-environment data potentially helps to detect and unravel various patterns and processes that may contribute to better decision-making towards sustainable development.

Researchers from the International Food Policy Research Institute (IFPRI) and the Institute of Development Studies (IDS), in close collaboration with numerous local partner institutions, mainly from the Ministry of Agriculture and Rural Development (MARD) and the General Statistics Office (GSO), have been involved in a number of research projects in the country. Data from the 1999 Population and Housing Census and from various living standards surveys have been used extensively in several research projects, in particular linking survey and census data to develop and map estimates of poverty and inequality at a high level of spatial disaggregation.

Researchers from the Swiss National Center of Competence in Research (NCCR) 'North-South', in collaboration with GSO and MARD, have recently produced a comprehensive *Socio-Economic Atlas of Vietnam* (Epprecht and Heinimann, 2004), which has proven to be a valuable source of information for a wide variety of users.

The Pro-Poor Livestock Policy Initiative (PPLPI) of the Food and Agriculture Organisation of the United Nations (FAO) is involved in a number of research projects in the country, collaborating with a variety of local partner institutions. Agricultural census data and living standards survey data have been used in several of PPLPI's research projects, and the 2001 Rural, Agricultural and Fishery Census serves as the main source of information for the new *Agricultural Atlas of Vietnam*.

The outputs of those efforts and selected analytical applications presented here, aim to demonstrate how statistical data can be translated into more valuable and readily understood spatial information that can contribute significantly to informed decision-making.

MAPPING STATISTICAL DATA

Methodology

Geographic information systems are powerful tools with the ability to analyze and display large amounts of data in formats that are readily assimilated, thus supporting the transformation of data into information, and of information into knowledge that is accessible to policy-makers and other non-specialist audiences. Whilst modern computing technology facilitates spatial visualization of data, there are many different ways in which, for instance, socio-economic characteristics can be represented spatially, and these need briefly to be reviewed.

Spatial representation in GIS is always an abstraction of reality: in the present case, socio-economic or agricultural attribute data being linked to spatial objects for representation. Gatrell (1991) defines such spatial objects as entities with both spatial location and geographically independent attribute characteristics. In making such abstractions it is important to decide what exactly should be represented, and for what purpose, and, moreover, who this information is targeted towards and to what end. The relevance of these issues is illustrated by a spatial visualization of the geographic distribution of poverty in Vietnam (Figures 1 and 2), taken from Minot, Baulch and Epprecht (2006). While both maps, Figure 1 the poverty incidence, and Figure 2 the poverty density, are based on the same estimates of commune-level poverty (derived using small area estimate techniques that combine household survey and census data), each representation shows more or less the inverse distribution of the location of geographical 'hotspots' of poverty, compared with the other. Whilst the reason for this apparent contradiction lies in the underlying population distribution, the example illustrates the importance of deciding what should be visualized, and for what purpose. Should 'poverty density' be mapped to identify areas where most poor people live, or do we want to identify 'poor areas' through a geographic depiction of the share of poor individuals of a population within a certain area?

While socio-economic phenomena vary across space, their exact 'values' can typically not be measured and attributed to exact locations: it is generally not clear precisely to which spatial units they relate. This raises an important question about how a particular socio-economic data set should be geo-referenced, that is, with what spatial entities it should be associated. This difficulty in exactly geo-referencing individual records means that indirect referencing in one way or another is always necessary. This is usually done through aggregation of the individual records by geographic areas, such as administrative entities, to which the aggregates are then linked. While the administrative unit is often the areal unit for which socio-economic data are reported, visual representation of such data at administrative level is not always without limitations. Unwin (1981) notes that the administrative unit as the spatial object for geo-referencing socio-economic

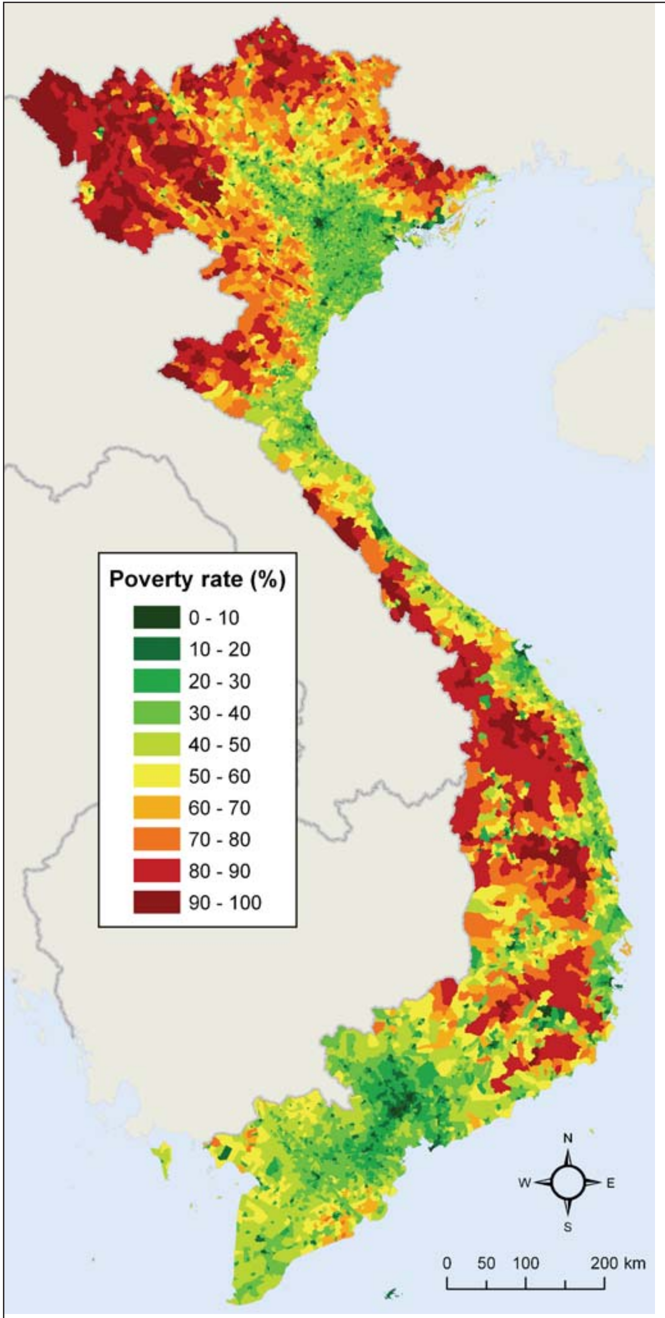

Figure 1. Poverty rate. Colour versions of figures are available in the online version, see Editorial for details.

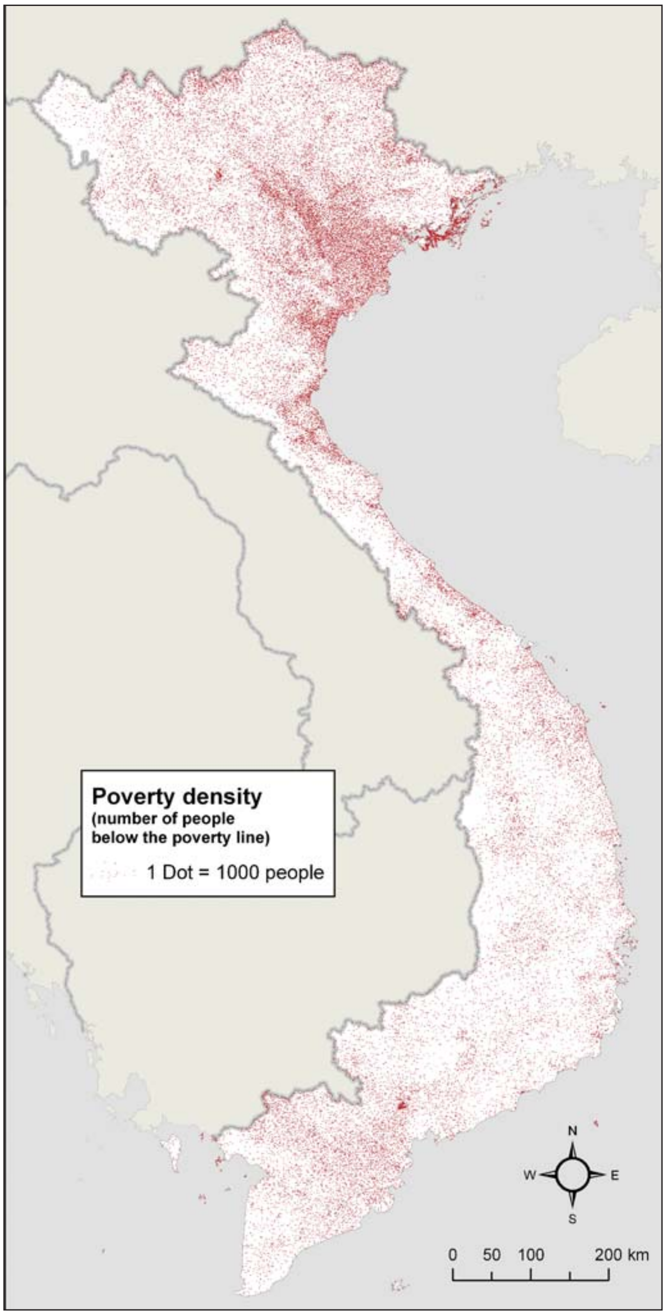

Figure 2. Poverty density.

data is problematic because such geographic units are imposed rather than natural units: the social phenomena might not at all be related to the location of the administrative boundaries.

A further problem is the so-called 'ecological fallacy' (Blalock, 1964), where observed relationships between two variables at one level of aggregation are not valid at another level of aggregation. Related to the ecological fallacy are issues of scale in representing spatial data. Aggregation and visualization of socio-economic data at different administrative levels, for instance, may result in quite different pictures of 'reality' being portrayed by the different maps. An example of this problem can be found in Minot, Baulch and Epprecht (2006), where the geographic depiction of poverty rates at provincial level implies only moderate poverty rates over large parts of the country, while the depiction of those poverty rates at commune level reveals, in fact, that high poverty rates persist in many parts of the country. Related both to scale issues and the ecological fallacy is the 'modifiable areal unit problem' (MAUP), extensively discussed by Openshaw and Taylor (1981) and Openshaw (1984). This problem was known well before the development of modern GIS, and describes the implications of alternative choices of areal units for aggregation, or zoning and geo-referencing of data. An example of MAUP that became known as 'gerrymandering' (e.g. Bivand, 1998), refers to the deliberate redesigning of administrative units to benefit a particular political party, making reference to the re-districting of the state of Massachusetts in 1812 by the state's governor E. Gerry in such a way as to maximize the number of representatives from the Republican Party in the US House of Representatives. These issues illustrate potential technical, conceptual and possibly ethical difficulties in representing data spatially. They also demonstrate the considerable responsibility of users to understand the implications of a chosen mapping approach that is, inherently, an abstraction of a socio-economic or ecological reality. The degree of abstraction depends on the quality and accuracy of the data, as well as on the spatial disaggregation. While there is generally not only one 'correct' mapping approach, it is important to consider carefully the nature of the data and the purpose of a visual representation when choosing how to present them spatially.

In a socio-economic or agricultural atlas the objective is to provide a graphic visualization of attribute data in space, in the form of a thematic map. (Topographic maps, conversely, show how objects are distributed in space, where special attention is given to the accurate representation of proportional relationships between objects). Different types of thematic maps have evolved over the years, each of them appropriate for specific types of information. Typically, five types of thematic maps can be differentiated.

In 'dot maps', each dot represents an equal number or quantity of the attribute being displayed. Geographical distributions and relative densities of absolute numbers, such as populations, can meaningfully be represented with dot maps. The dots are typically distributed randomly within the enumeration area (e.g. the province) for which the attribute exists. The poverty density in Figure 2 is a typical example of a dot map.

In 'choropleth maps', the entire area is divided into discrete regions such as administrative entities, for which attribute data exist. Political administrative maps are typical examples, where each country or province is depicted in a distinctive colour, and colours change along the boundaries only. Though this wrongly implies uniformity within each entity, and sharp changes at the borders, choropleth maps are frequently used to depict socio-economic data by administrative units. Choropleth maps typically depict relative numbers (i.e. percentages and ratios rather than totals), for example the poverty rate map reproduced in Figure 1.

'Isarithmic maps' represent trends in continuous data (actually observed, or interpolated from discrete data) through lines of equal values (isolines). Typical examples include meteorological maps of atmospheric pressure (isobars) or temperature (isotherms) and elevation maps depicting contour lines of equal elevation (isohypses). It is not common for socio-economic data to be depicted as isarithmic maps.

On 'symbol maps' the attributes are represented by symbols (e.g. circles) and the size of the symbols varies according to their attribute value. A special case of symbol maps is the graph map, where statistical graphs are used to show the values of multiple attributes in space. The *Digital Atlas of California* is an example of the use of symbol maps for geographic depiction of socio-economic data (Bowen, 2000), while an application of graph maps is provided, e.g., in Minot, et al. (2006).

'Trend surface maps' depict continuous surfaces as a raster grid and are used to visualize individual values for any point in space at the given spatial resolution. Such surfaces are generally obtained either through remotely sensed data (e.g. elevation data), modelled data – such as, for instance, 'accessibility' of certain service locations (e.g. Epprecht and Heinimann, 2004) – or through interpolation of attribute data of measurement points (e.g. air pollution).

Dot and choropleth thematic maps are generally the most useful and widely used to represent socio-economic and agricultural data. The applications presented in this paper show examples of such mapping approaches at the commune level for Vietnam. These are often converted to trend surface raster images, however, for purposes of analysis and for integration with other datasets, such as remotely sensed imagery, that are not obtained at the same unit of administrative level.

SOCIO-ECONOMIC ATLAS OF VIETNAM

Policy-makers and researchers are increasingly confronted with environmental and socio-economic problems at multiple temporal and spatial scales such as, for example, land use/land cover change and its implications and the impacts of related policies across socio-economic, biophysical and geographic domains (Crews-Meyer, 2002). Due to the advances in earth observation, computing power, as well as in GIS technologies and applications over the last two decades, the availability of biophysical data is rapidly augmenting across the entire scale continuum and at ever increasing resolutions. This technological progress offers new avenues for the presentation and analysis of spatial data.

In a development context, socio-economic data, however, are still typically either available at very local scales (e.g. household) for small pilot areas only, or on highly aggregated levels (e.g. province or district). This may partly be attributed to the fact that researchers are often not aware of the value of spatial reference of their data or available respective methods (Rindfuss and Stern, 1998). The consequences, in the specific case of Vietnam, are threefold. First, spatial representations of key socio-economic variables are only available at an aggregated level and consequently there is little knowledge existing about eventual emerging development-relevant spatial patterns. Secondly, the scale gap between the available biophysical and the socio-economic data is just too wide to overcome the involved methodological problems to link these disparate data types appropriately for analytical purposes. Hence we lack insights into eventual causal links or relations between the biophysical conditions and socio-economic realities, beyond the site specific contexts of local case studies, at a policy relevant meso-scale. Finally, as ultimate consequence, policy- and decision-makers at various levels are not equipped with the necessary information to ensure informed decision-making or to allow for cost effective targeting of certain population groups or problem contexts.

Against this background, the Swiss National Centre of Competence in Research (NCCR) 'North-South', in collaboration with the Vietnamese General Statistics Office (GSO) and the Vietnamese Ministry of Agriculture and Rural Development (MARD), decided in 2004 to generate geographically highly disaggregated socio-economic spatial datasets for Vietnam. Based on the 1999 Population and Housing Census of Vietnam, a set of 57 socio-economic maps and respective interpretations were elaborated at commune level (10,474 communes) (Epprecht and Heinimann, 2004). The resulting bi-lingual (English and Vietnamese) *Socio-Economic Atlas of Vietnam* is available as A3 hardcopy for easy reference as well as online to ensure a distribution to the wider public (see note, p. 201).

The *Atlas* is divided into seven main chapters. The maps included in the first chapter provide a general overview of geographical aspects of Vietnam, such as topography, main infrastructural elements, forest cover, and the administrative division. In the second chapter, maps of general demographic characteristics are presented, while the third chapter encompasses maps on aspects related to literacy and education. Chapter Four provides a series of maps of economic activity. The maps presented in Chapter Five deal with aspects of society related to religion and ethnicity, before several variables on household characteristics and living conditions in Vietnam are presented on the maps in Chapter Six. Maps revealing the spatial distribution of poverty in Vietnam (Minot, Baulch and Epprecht, 2006), along with a map on accessibility, are included in Chapter Seven.

The processing and visualization of socio-economic data of Vietnam at a spatially highly disaggregated level has great potentials to facilitate informed decision-making at various levels: by granting easy access to individual socio-economic variables via countrywide maps; by conveying the communication of specific socio-economic realities to policy-makers; by increasing the awareness of the general public about specific socio-economic realities; by laying the basis for discovering specific patterns and relations between different socio-economic variables; by narrowing the scale gap between biophysical and socio-economic data and thereby furthering the integration of different data sources for further analysis; by supporting the efficiency of specific development efforts and investments by improved targeting (see Figures 1 and 2); and eventually possibly even by empowering marginal groups, as maps grant easy access to information to a wide public (compared to the complex tabular statistical information).

This *Atlas* is now widely appreciated among many national and international players in Vietnam. It is used by the private as well as by the governmental sector as a source of information on socio-economic characteristics of the Vietnamese society.

AGRICULTURAL ATLAS OF VIETNAM

In order to provide reliable and contemporary information to feed into research programs and policy processes, the Government of Vietnam's General Statistics Office (GSO) conducted Rural, Agricultural and Fishery Censuses in 1994 and 2001. The 2001 census collected data from more than 13.9 million households; 31.3 million rural laborers; 61,017 farms; 8,934 communes; 7,171 cooperatives; 710 handicraft villages; and 3,599 private enterprises. The data provide a comprehensive picture of Vietnam's rural population and agriculture 15 years after the '*Doi Moi*' reforms – decentralization and liberalization efforts that allow agricultural producers a much greater degree of autonomy – were initiated. In order to encourage widespread use of the information collected, the GSO, with technical and financial support from FAO, and with additional funding from the Swiss Agency for Development and Cooperation (SDC), is compiling an *Agricultural Atlas of Vietnam*, based on this census.

The *Atlas* essentially links commune-level summaries of the census data to a digital file of commune-level administrative boundaries. There are 10,108 rural communes and urban wards with agricultural activities in Vietnam and the median land area of a commune is of the order of 18 km^2; so this results in very detailed digital maps that can be used to present the data in an easily understood manner, and also to facilitate analysis of the data, using GIS technology, both within the census dataset and with spatially referenced data from other sources.

The 2001 *Agricultural Atlas of Vietnam* will potentially provide a valuable resource for researchers, policy-makers, educational institutions, development agencies and other international organizations, as well as for the reader who has a general interest in the agriculture and rural development of Vietnam. Particularly for researchers, the availability of spatially highly disaggregated agricultural data that is based on actually enumerated numbers opens doors to tap the potential of GIS-based agricultural analyses that go beyond the typical analysis of land cover maps based on remote sensing.

The *Agricultural Atlas* is set out in five broad sections, covering:

1. the basic demography and infrastructure
2. arable agriculture, aquaculture and forestry
3. livestock
4. commercial farms and cooperatives
5. the agro-ecological environment.

The first four sections merely describe the census data and present them in appropriate map formats. Livestock, for example, are shown as density maps based on the total area in a commune or the designated agricultural area. This is an important distinction in communes where only a small proportion of land is dedicated to agriculture. Livestock are also shown in terms of the average number of animals kept by livestock-keeping households; pertaining to the typical scale of livestock operations in that commune. In the final section of the atlas the emphasis is more analytical, in which some derived maps are presented. These are discussed in more detail in the next section.

ANALYTICAL APPLICATIONS

Mapping Poverty and Inequality

In Vietnam, where regional disparities in human welfare are basically large, information about the spatial distribution of poverty is especially important because it supports policy-makers and program designers in gaining insights to the causes of poverty and targeting development activities to the poorest regions. A study conducted by the International Food Policy Research Institute (IFPRI), in collaboration with the Institute of Development Studies (IDS) and the Vietnamese inter-ministerial poverty mapping task force, provided small-area estimates of poverty and inequality for Vietnam and explored the geographic determinants of poverty (Minot, Baulch and Epprecht, 2006).

Living standards surveys are typically conducted regularly, and provide insights into differences in living standards among different regions of a country. They fail, however, to reveal welfare differences at a higher level of geographical disaggregation. The IFPRI study applied the method of 'small-area estimation' (see e.g. Ghosh and Rao, 1994; Pfefferman, 2002) to combine information from the 1997–98 Vietnam Living Standards Survey (VLSS) and the 1999 Population and Housing Census, to estimate various measures of poverty and inequality for small areas. First, the VLSS data were used to evaluate the relationship between per capita expenditure and various household characteristics. The derived model was then applied to data on those same characteristics from the Census, thus generating poverty estimates for each household. These results were then aggregated to generate estimates of poverty and inequality for each of 61 provinces, 614 districts, and 10,474 communes.

The results, shown in Figure 1, indicated that the poverty rate was highest in the remote areas of the northeast and northwest regions, the upland areas of the north central coast, and the northern part of the central highlands. Poverty rates are intermediate in the Red River and Mekong River deltas. The lowest poverty rates were found in the principal cities of Hanoi and Ho Chi Minh City, in other urban areas, and in the southeast region. An analysis of the density of poverty, Figure 2, revealed, however, that most of the rural poor live in the lowland deltas; areas where the incidence of poverty is relatively low but the absolute numbers of poor people are high, due to the high population density.

Comparing these results with the official district-level numbers of poverty from the Vietnamese Ministry of Labor, Invalids, and Social Affairs MoLISA, we find very little correlation. The most likely reason for this is variation in the methods used by MoLISA from one district to another.

The poverty analysis revealed that the level of inequality in Vietnam was relatively low by international standards, in spite of the wide disparities in poverty rates from one region to another. Inequality was greatest in the large cities and, perhaps surprisingly, in some rural upland areas. Inequality was lowest in the Red River Delta, followed by the Mekong River Delta. A decomposition analysis indicated that just one-third of the inequality was found among districts and two-thirds within districts, suggesting that targeting of anti-poverty programs should be done at the commune level, rather than at the district level.

This study also explored the geographic determinants of poverty using two types of model. In the first model, a simple regression analysis to 'predict' district-level poverty in urban and rural areas, geographic determinants, including agro-climatic variables and market access, explained about three-quarters of the variation in district-level rural poverty. Poverty was higher in districts with sloping land; bare and rocky land; poor-quality soils (sandy, saline, or acid sulphate); and far from towns. However, these variables did not explain urban poverty well. The second model was a spatially-weighted local regression model of rural poverty in which coefficients were allowed to vary over space (see e.g. Brunsdon, Fotheringham and Charlton, 1996). This model revealed that flat land and high road densities were associated with lower poverty rates throughout Vietnam, but that other variables, such as rainfall and forest cover, were positively associated with poverty in some areas and negatively associated in others. Overall, the relationship between agro-climatic variables and poverty varies significantly from one area of Vietnam to another.

Many anti-poverty programs are geographically targeted in Vietnam. Minot, Baulch and Epprecht (2006), however, highlighted the difference between areas where the incidence of poverty is high and areas where the density of poverty is high (Figure 1). Programs that concentrate exclusively on areas with high poverty rates will not reach the majority of the poor. Since three-quarters of the variation in district-level rural poverty can be explained by variables linked to low agricultural potential and lack of market access, improving market access (through road improvements, for example) seems important. Poor agro-climatic conditions are not easily remedied, and in some cases migration out of disadvantaged regions may the best strategy. Current policies to restrict migration may prevent families in these regions from pursuing this strategy.

Considering that the demand for spatially disaggregated information on poverty in Vietnam was great among national and international players in the development field, interest in the new poverty maps produced was accordingly high. An important objective of this poverty mapping analysis was to promote well-informed debate on poverty and welfare inequalities in Vietnam, based on consistent and reliable data on poverty. The exercise showed that it was possible to produce, based on reliable survey and census data, estimates of poverty, including confidence intervals, at a high level of spatial disaggregation.

The reception of the poverty maps (and the respective data) in Vietnam, and its influence on policy making, was, however, rather two-fold: while many national and international agencies in Vietnam found the maps to closely agree with their fieldwork, and therefore to be rather credible (Swinkels and Turk, 2007), most officials were initially reluctant to use data that deviated from the official MoLISA data (*ibid.*). Most international organizations, on the other hand, were more comfortable in using those new poverty data that were generated based on an internationally recognized transparent methodology. A growing number of national agencies are now also interested in using poverty numbers generated via the small-area estimation technique.

POVERTY, REMOTENESS, AND ETHNICITY

Poverty and inequality maps such as the ones developed for Vietnam are increasingly being used as planning and targeting aids, and are powerful tools to reveal and communicate information. However, the interpretation of such maps is not always straightforward, and such maps can be prone to misinterpretation. Epprecht, Müller and Minot (2007) demonstrated the risks of biased inferences involved in interpreting spatial patterns of poverty and inequality estimates. They developed poverty and inequality maps disaggregated by population subgroups and by spatial units. Using the small-area estimate techniques described above, they created separate poverty incidence and inequality indices for population sub-groups at a high level of spatial disaggregation.

Epprecht, Müller and Minot (2007) showed that the urban-rural welfare divide was the most important single contributor to overall inequality, closely followed by a welfare divide between ethnic minorities and the ethnic majority populations. The study highlighted that in rural areas, both of these sub-populations demonstrated similar levels of inequality, but that welfare levels between the two differed by factor two: the ethnic minority populations being twice as poor as the rest. In urban areas, higher poverty rates were statistically related to higher inequality levels. These patterns were more pronounced for the ethnic majority population, but were also present for the ethnic minority sub-group. Given Vietnam's large agriculturally-based rural population, increasing rural-urban migration may reduce rural-urban income differences, although possibly at the cost of increasing inequality in urban areas. Improving the legal status of migrants and easing registration procedures in the cities may be one policy option through which the burden for urban newcomers can be eased.

In Vietnam local poverty rates are strongly positively related to remoteness (Figures 1 and 3), and local measures of accessibility are more important than access to larger cities (Epprecht, Müller and Minot, 2007). This indicates the important role of local infrastructure, such as local markets, health care facilities and schools, in poverty reduction. The more local a measure of access, the less important the actual physical distances become, and the more important socio-cultural distances, such as language barriers or cultural differences grow. Language barriers, for instance, might prevent an ethnic family from completing the paper work required to obtain credit, or from purchasing appropriate drugs at a pharmacy, even if the travel time to the bank branch or a pharmacy is minimal.

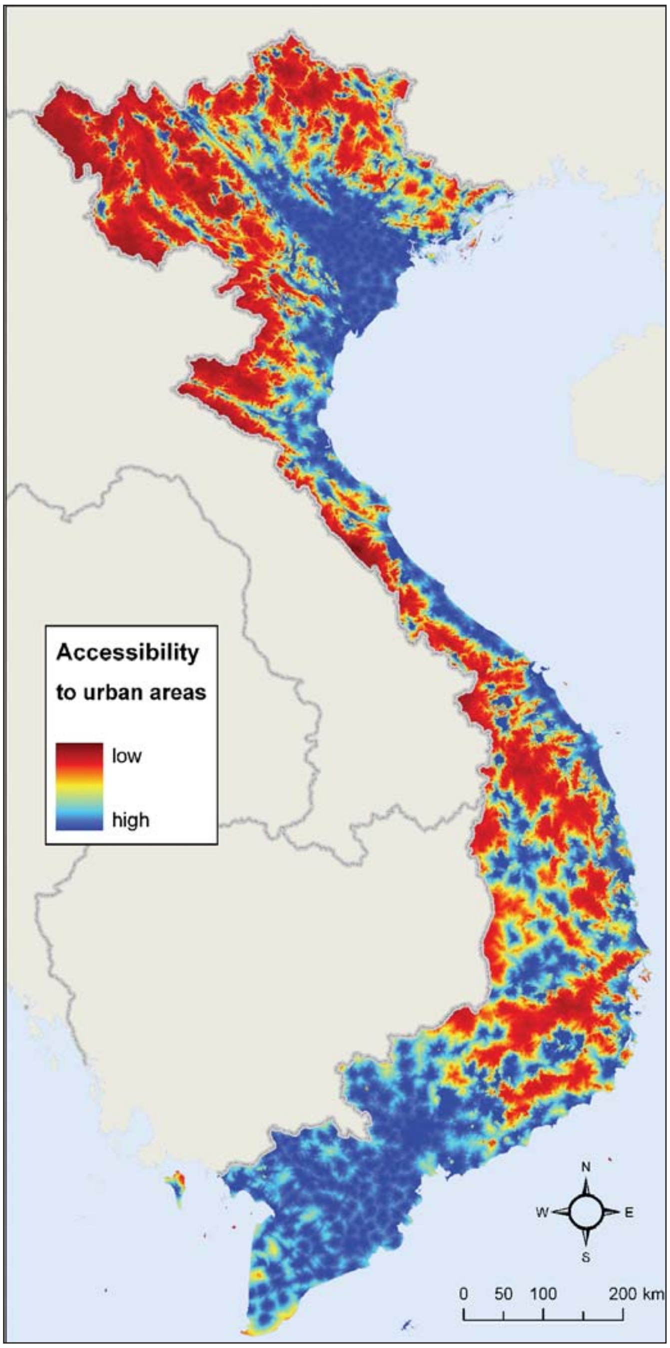

Figure 3. Accessibility to urban areas.

The results of Epprecht, Müller and Minot (2007) provide strong empirical evidence that remoteness is a considerably weaker determinant of poverty compared to the ethnicity of a household. The broad spatial patterns of poverty are actually due largely to the spatial distribution of the ethnic minority populations, and to a much lesser extent due to spatial patterns in physical accessibility. Furthermore, inequality levels were found not to be related to remoteness; rather, the higher levels of inequality found in the poor upland areas seem to be due to the spatial coexistence of poorer ethnic minority people with the better-off ethnic majority.

Present pro-poor development policies are to some degree targeted at ethnic minority areas but not specifically towards ethnic minority households or the local ethnic minority population in general (van de Walle and Gunewardena, 2001). With socio-cultural distances emerging potentially as being more important than geographic distances in determining inequality, development policies aimed at balancing welfare levels across sub-populations require increased emphasis on the targeting of specific population segments. Such targeted approaches need to take into account the specific natural, physical, and human endowments of a subgroup, as well as the potential economic returns of specific intervention strategies to particular sub-populations. Recent efforts in Vietnam as part of the on-going public administration reform, and the 'rolling-out' of the Comprehensive Poverty Reduction and Growth Strategy (CPRGS), signal a shift in the planning process in this direction. But poor incentives for local decision-makers to improve the situation of disadvantaged population segments, coupled with a lack of the necessary capacities, render its implementation at the local level a big challenge.

Epprecht, Müller and Minot (2007) have demonstrated that maps of poverty and inequality can be prone to biased inferences, and that further analyses may reveal important insights not captured by an overall visualization of poverty and inequality. Furthermore, mapping of additional socio-economic characteristics of a population, such as ethnicity (Figure 4), can help point towards factors that shape the landscape of poverty and inequality.

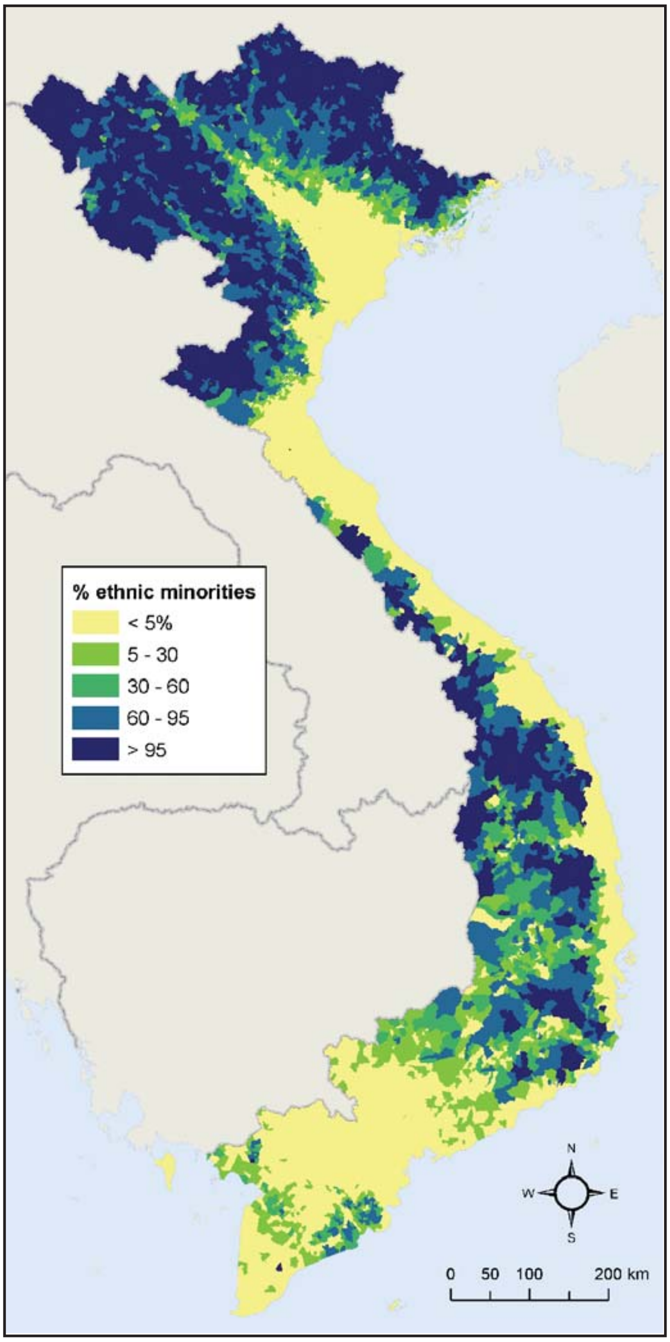

Figure 4. Ethnic minorities (%).

POVERTY, POPULATION, AND THE FOREST

The Government of the Socialist Republic of Vietnam formulated precise economic and environmental targets in the Vietnamese Development Goals (VDG) in 2002. The progress towards the VDG targets of poverty reduction and environmental sustainability is monitored using, among several other indicators, the poverty incidence and the percentage of forest cover (Socialist Republic of Vietnam, 2002: p. 120). Efforts towards effective allocation of investments into rural development aiming at poverty reduction and concurrent environmental protection will benefit from spatial insights into the nature of relationships between poverty and the state and extent of forests.

Müller, Epprecht and Sunderlin (2006) investigated the spatial linkages in the forest–poverty nexus in Vietnam using the official forest cover data set from 1999 for Vietnam, produced by the Vietnamese government's Forest Inventory Program 286, and the countrywide small-area estimations of poverty from Minot, Baulch and Epprecht (2006). The analysis linked the commune-level poverty estimates with the share of forest cover by forest quality within each commune. The forest cover data, which is differentiated by forest quality, represent the most recent and accessible official data set to date, and was used by the Vietnamese Government to assess progress towards the VDGs. Descriptive empirical insights were drawn from map comparisons, visual descriptive statistics, augmented with insights from exploratory spatial data analysis (ESDA) (Anselin, 1995; Haining, 2003).

Map comparisons show that most of the Vietnamese poor live in densely populated river deltas and cities, while remote upland areas have the highest poverty incidences, gaps, and severities (Minot, Baulch and Epprecht, 2006). Forests of high value for both local livelihoods and the regional and global environment are located in areas where relatively few poor people live, but where the incidence, gap, and severity of poverty are highest, and where the livelihood strategies are largely based on agricultural and forest activities (Müller, Epprecht and Sunderlin, 2006).

Extensive forest areas coincide geographically with areas exhibiting high poverty incidences. The higher the poverty incidence in a commune, the higher is the likelihood of high forest cover percentage. Conversely, the share of forest cover within a commune is negatively related with poverty density: communes with the highest poverty density tend to have little forest, while communes in the sparsely populated upland areas, with low poverty density (Minot, Baulch and Epprecht, 2006), tend to have a high share of forest cover at the scale of analysis.

The data allowed a spatially-explicit investigation of the degree of association of forest quality, as an environmental indicator, and poverty incidence as well as poverty density. Spatial correlations and regional linkages were visualized between forest cover and the small-area poverty estimates to improve the understanding of the relationship between poverty and forest quality.

Policies that aim at poverty alleviation and concurrent forest conservation may benefit from such analytical integration of socio-economic and environmental data. Assuming data quality is adequate for the areas under investigation, such results may prove useful for targeting investments aimed at concurrent poverty alleviation and forest protection. Spatial correlations help identifying areas where benefits may arise on both fronts and, in that way, could serve achieving the long-term policy objectives of the VDGs.

CONCLUSION

We have presented and evaluated a number of approaches to convert statistical and spatial data into information to generate additional knowledge that informs decision-makers and researchers in Vietnam. These included the mapping of data from the Population and Housing Census and from the Agricultural Census for the smallest administrative units, an analysis of living standards survey and population census data to generate small-area estimates of poverty measures, disaggregation of poverty measures for population and spatial subgroups, and the linkage of poverty measures to forest cover to exemplify the spatial relationships in the human–environment nexus.

These efforts aim at providing additional insights into the spatial distribution and spatial relationships of variables and to make such information more readily accessible. Mapping, or geovisualization, of data and of analysis results are efficient ways to display the spatial variation of large amounts of socio-economic, biophysical, and infrastructural data. Such maps can be easily and efficiently conveyed to decision-makers and researchers alike; in that way, the formulation and targeting of programs and investments can be improved.

Furthermore, geovisualization combined with spatially explicit analysis may generate complementary insights into relationships and processes that cannot be gleaned from the analysis of non-spatial data. The heterogeneous distribution of such relationships over an entire country may reveal specific patterns of problems and potentials related to sustainable development. In addition, the integration of data contained in censuses and surveys facilitates a range of possibilities for spatial analyses that incorporate spatial processes and dependencies of relevance to rural development and environmental management. Spatial analysis therefore provides ample opportunities for development research and may add to an improved understanding and a better subsequent targeting of investments.

The provided examples all showed that location matters. Yet, other attributes like ethnicity or history are sometimes more important and easily overlooked by a simple inspection of maps or by simple correlations of spatial or non-spatial data. This underlines that great care is required in deriving conclusions from the sheer interpretation of maps that, used in isolation, may disguise underlying factors.

There is a great untapped potential in deriving insights from the integration of socio-economic data with environmental data. The example of linking poverty to forest cover and exploring the spatial relationships in the poverty–forest nexus, presented an approach that may contribute to an improved geographic targeting of investments that aim at concurrent poverty alleviation and forest protection. A large body of additional application along these lines can be imagined, e.g., the interaction between people’s livelihoods and water pollution, soil degradation, or the state of and changes in biodiversity.

The display and analysis of spatial information is indispensable to produce knowledge about the location of objects, about spatial clusters, and relationships. The amounts of spatial data will continue to increase in the future, the quality of data is likely to improve, and computing technology will advance (Goodchild et al., 2000). Such trends allow for a wide range of additional and promising future applications that hopefully lead to better management and decisions. The applications presented in this paper have exemplified ways for the display and analysis of human-environment data that serve both development policy and research by creating complementary information.

Note

The *Socio-economic Atlas of Vietnam* is available online at http://www.north-south.unibe.ch/content.php/publication/id/1712

References

Anselin, L. (1995) Local Indicators of Spatial Association: LISA. *Geographical Analysis,* 27 (2), 93–115.

Bivand, R. (1998) *A review of spatial statistical techniques for location studies.* Norwegian School of Economics and Business Administration, Bergen.

Blalock, M. (1964) *Causal inferences in nonexperimental research.* Chapel Hill, University of North Carolina.

Bowen, W. (2000) *Digital Atlas of California.* California State University, Northridge. (http://130.166.124.2/CApage1.html)

Brunsdon, C.F.; Fotheringham, A.S. and Charlton, M.E. (1996) Geographically weighted regression: A method for exploring spatial nonstationarity. *Geographical Analysis,* 28: 281–298.

Crews-Meyer, K.A. (2002) Challenges for GIscience: Assessment of policy relevant human-environment interactions. In: S. Walsh and K. A. Crews-Meyer, eds. *Linking people, place and policy a GIScience approach.* Boston, Dordrecht, London: Kluwer Academic Publishers, pp. 1–6.

Epprecht, M. and Heinimann, A., eds. (2004) *Socio-economic Atlas of Vietnam – a depiction of the 1999 Population and Housing Census.* Swiss National Centre of Competence in Research (NCCR) North-South, University of Berne, Berne. 168pp.

Epprecht, M.; Müller, D. and Minot, N. (2007) How remote are Vietnam's poor? An analysis of spatial patterns of poverty and inequality in Vietnam's society. IFPRI discussion paper series (submitted). Washington DC: International Food Policy Research Institute.

Gatrell, A. (1991) Concepts of space and geographical data. In: Maguire, D.J., Goodchild, M.F. and Rhind, D.W. (eds) *Geographical information systems: principles and applications.* Harlow, Longman, New York, John Wiley Sons Inc. Vol 1: 119–34.

Ghosh, M. and Rao, J.N.K. (1994) Small area estimation: An appraisal. *Statistical Science,* vol 9, no.1 (1994), 55–76.

Goodchild, M.F.; Anselin, L.; Appelbaum, R.P. and Harthorn, B.H. (2000) Toward spatially integrated social science. *International Regional Science Review,* 23, 139–159.

Haining, R. (2003) *Spatial data analysis: Theory and practice.* Cambridge. Cambridge University Press.

Minot, N.; Baulch, B. and Epprecht, M. (2006) Poverty and inequality in Vietnam – Spatial Patterns and geographic determinants. IFPRI research report #148. Washington DC: International Food Policy Research Institute.

Minot, N.; Epprecht, M.; Tran Thi Tram Anh and Le Quang Trung. (2006) Income diversification and poverty in the Northern Uplands of Vietnam – Patterns, trends, and policy implications. IFPRI research report #145. Washington DC: International Food Policy Research Institute.

Müller, D.; Epprecht, M.and Sunderlin, W.D. (2006) Where are the poor and where are the trees? Targeting of poverty reduction and forest conservation in Vietnam. Bogor, Indonesia: Working Paper No. 34, Center for International Forestry Research (CIFOR).

Openshaw, S. (1984) The modifiable areal unit problem. *Concepts and Techniques in Modern Geography,* 38. Norwich. Geo-Books.

Openshaw, S. and Taylor, P.J. (1981) The modifiable areal unit problem. In: Wrigley, N. and Bennett, R.J. (eds) *Quantitative geography: a British view.* London, Routledge: 60–70.

Pfefferman, D. (2002) Small area estimation – New developments and directions. *International Statistical Review,* (2002), 70, 1, 125–143.

Rindfuss, R.R. and Stern, P. (1998) Linking remote sensing and social science: The needs and challenges. In: D. Livermann et al., editors. *People and pixels linking remote sensing and social science.* Washington: National Academy Press, pp. 1–28.

Socialist Republic of Vietnam. (2002) *The Comprehensive Poverty Reduction and Growth Strategy (CPRGS).* The Socialist Republic of Vietnam (Approved by the Prime Minister), Hanoi.

Swinkels, R. and Turk, C. (2007) Poverty mapping: Experience from Vietnam. In: Coudouel, A. and Bedi, T. eds. *More than a pretty picture: Using poverty maps to design better policies and interventions.* World Bank Publications, Washington DC.

Unwin, D.J. (1981) *Introductory spatial analysis.* London, Methuen.

van de Walle, D. and Gunewardena, D. (2001) Sources of ethnic inequality in Viet Nam. *Journal of Development Economics.* 65: 177–207.

Abstract

Vietnam has developed rapidly over the past 15 years. However, progress was not uniformly distributed across the country. Availability, adequate visualization and analysis of spatially explicit data on socio-economic and environmental aspects can support both research and policy towards sustainable development. Applying appropriate mapping techniques allows gleaning important information from tabular socio-economic data. Spatial analysis of socio-economic phenomena can yield insights into locally-specific patterns and processes that cannot be generated by non-spatial applications. This paper presents techniques and applications that develop and analyze spatially highly disaggregated socio-economic datasets. A number of examples show how such information can support informed decision-making and research in Vietnam.

FROM STATISTICAL DATA TO SPATIAL KNOWLEDGE

Keywords: decision support; spatial data; rural development; poverty; geovisualization; spatial analysis; Vietnam

Michael Epprecht is a senior researcher with the Swiss National Center of Competence in Research (NCCR) 'North-South', Centre for Development and Environment (CDE), University of Berne, Steigerhubelstrasse 3, CH–3008 Berne, Switzerland. His research focuses on spatial analysis of socioeconomic rural development. He holds an MSc and a PhD in Geography from the University of Berne. E-mail: michael@epprecht.org

Andreas Heinimann is a program coordinator and senior research fellow at the Swiss National Centre of Competence (NCCR) 'North-South', Centre for Development and Environment (CDE), University of Berne, Steigerhubelstrasse 3, CH–3008 Berne, Switzerland. His research interests include land use/land cover change and question of data aggregation and generalization. He received his PhD in Geography from the University of Berne, and his MSc in Environmental Science from the Swiss Federal Institute of Technology in Zürich. E-mail: andreas.heinimann@cde.unibe.ch

Nicholas Minot is a senior research fellow in the Markets, Trade, and Institutions Division of the International Food Policy Research Institute (IFPRI), 2033 K Street, NW, Washington, DC 20006–1002, USA. He leads a research program on strategies to promote smallholder participation in high-value agricultural commodity supply chains. His research interests include agricultural market reform, fruit and vegetable markets, and the spatial patterns in poverty. He received his PhD and MSc in agricultural economics from Michigan State University and his BA in international development from Brown University. E-mail: n.minot@cgiar.org

Daniel Müller is a Senior Researcher at the Center for Development Research (ZEF) of the University of Bonn, Walter-Flex-Straße 3, 53113 Bonn, Germany. His research focuses on spatially explicit analysis of land use and rural development in Eastern Europe, Southeast Asia, and East Africa. He is an agricultural economist by training and received his MSc from Humboldt University Berlin and his PhD from the Georg-August-University Göttingen. E-mail: d.mueller@rz.hu-berlin.de

Tim Robinson is the Livestock Information Officer of the Pro-Poor Livestock Policy Initiative (PPLPI) of the Food and Agriculture Organization of the United Nations (FAO), Viale delle Terme di Caracalla, 00100, Rome, Italy. He graduated from the University of Oxford with a degree in pure and applied biology in 1988. His PhD, at the University of Reading, was on the ecology of the African armyworm, and provided his introduction to the world of remote sensing, geographic information systems and spatial analysis. E-mail: tim.robinson@fao.org