Draft version
Not for citation

SCALE OF QUESTIONS & SCALE OF ANSWERS:
THE ROLE OF REMOTE SENSING FOR PUBLIC HEALTH POLICY

Marcia Caldas de Castro
Department of Population and International Health
Harvard School of Public Health

March, 2007

Paper prepared for presentation at the 2007 Annual Meeting of the Population Association of America, New York, NY.

SCALE OF QUESTIONS & SCALE OF ANSWERS: THE ROLE OF REMOTE SENSING FOR PUBLIC HEALTH POLICY

Introduction

Recent advancements and improvements in technology, data storage, computer software, and data analysis opened innumerous possibilities for the use of GIS, remote sensing and spatial modeling. The incorporation of these techniques into social sciences is yet to become fully recognized, broadened, and well understood (Castro, 2007). In the public health arena, however, innumerous applications using these techniques have been made, resulting in a large literature available in articles, special editions of journals, and books.

Although the contributions of GIS, remote sensing and spatial modeling to public health are enormous they involve two challenges: (i) confidentiality, and (ii) modifiable areal unit problem (MAUP). The first concerns the protection of the identity of interviewees when collecting, analyzing, and disseminating survey data (Fox, Suryanata and Hershock, 2005, VanWey, Rindfuss, Gutmann, Entwisle and Balk, 2005). The second is a potential source of error in studies that analyze aggregated spatial data. Results found at one level of aggregation (e.g. census tract) are not necessarily the same when a different level of aggregation is used (e.g. county) (Pryor, 1984). In a loose sense, MAUP is a spatial ecological fallacy. As Openshaw (1984) stated: "*the areal units (zonal objects) used in many geographical studies are arbitrary, modifiable, and subject to the whims and fancies of whoever is doing, or did, the aggregating.*" Although individual level data analysis may partly avoid this issue, scaling-up the results could also lead to ambiguous statements.

MAUP has two components: (i) a scale effect – spatial data analysis at different scales may produce different results, and (ii) a zoning effect – regrouping zones at a given scale may produce different results (Fotheringham, Brunsdon and Charlton, 2000). This paper focuses on the first component, scale, particularly related to the use of remotely sensed information. Although available imagery, at all levels of resolution, may be suitable for public health studies, we argue that the spatial resolution should be chosen based on the area being studied, the purpose of the research, the future use of the results, and the geographical coverage of the study. Additionally, we suggest that multiscale studies, when possible, are important for the understanding of disease transmission processes (which are likely to operate at multiple scales). Ultimately, we propose guidelines on the most appropriate imagery to use for purposes of planning, program evaluation, surveillance and research of public health related subjects.

Scale & Remote Sensing

Dungan (2001) provides a comprehensive review of the different definitions that can be associated to the word scale. In this paper, scale is defined as the size of a certain geographical area, represented in measurement units (e.g. meters, kilometers) or general concepts (e.g. national, local). The latter is referred to in discussions about policy making, while the former is used when relating to remotely sensed imagery, and therefore it would be similar to spatial resolution – the pixel size of the image. We consider high spatial resolution imagery those that have a pixel size lower or equal to 10 m, medium resolution if the size is between 10 m and 1 km, and low resolution when the pixel size is greater or equal to 1 km. These cutoffs were arbitrarily determined based on the current availability of imagery, and therefore do not represent a widely accepted rule.

Statistical models – e.g. semivariogram and local variance – have been proposed to facilitate the choice of spatial resolution (Woodcock and Strahler, 1987, Curran, 1988, Woodcock, Strahler and Jupp, 1988, Atkinson and Curran, 1997). The basic idea is to describe the spatial structure in a particular scene as a function of its spatial resolution. Results offer tangible indication of the scale needed to better characterize specific features of the studied area. One of the advantages is that these methods can be applied to a simple photograph scanned at high resolution, providing guidelines for the ideal remotely sensed data to acquire.

Two additional resolutions are important for remotely sensed imagery: spectral and temporal. The first refers to the number and width of wavelengths that can be detected at the electromagnetic spectrum (Jensen, 2007). Images with one single band are considered to have low spectral resolution; multispectral images have medium resolution, and hyperspectral images, with 100-200 bands, have high resolution. The advantage of multiple bands is that it allows different physical features to be identified more accurately. For example, in Landsat TM images band 4 (0.76-0.90 μm) is particularly useful for crop identification, while band 7 (2.08-2.35 μm) is most appropriate to assess rock formations (Jensen, 1996).

Temporal resolution refers to the time interval it takes for a sensor to cover the same area of the Earth’s surface. Time intervals greater or equal to 15 days are associated to low temporal resolution, those between 15 days to one day have medium resolution, and in those that are lower or equal to one day have high resolution. Although high temporal resolution imagery can play a crucial role for monitoring activities, usually it is inversely related to the spatial resolution. For example, Landsat 5 and 7 have a spatial resolution of 30 m and a temporal resolution of 16 days. Therefore, each area is captured no more than twice every month. In tropical areas that may result in very few cloud-free images every year (Hay, 2000). In contrast, AVHRR – Advanced

Very High Resolution Radiometer – has a spatial resolution of 1090 m, but a temporal resolution of only 12 hours.

Public Health & Remote Sensing

The first use of a remotely sensed image (an aerial photograph) to study the potential breeding sites of a mosquito was done in 1971 by NASA, and by the end of the 1970s satellite images were being applied for similar studies (Hay, Snow and Rogers, 1998). Almost four decades after the initial study, an extensive literature shows the varied uses of remotely sensed data for public health assessments. Applications used aerial photographs, satellite images, and radar images – the latter has the important advantage of being an all-weather sensor, and therefore cloud coverage is not a problem (Jensen, 2007) . A significant number of these applications are for vector-borne diseases, given the likely environmental influence on risk and transmission. Applications of remote sensing have included: (i) assessment of environmental characteristics likely to be related with vectors, such as patterns of land use, deforestation, identification of water bodies (rivers, lakes, canals and temporary water pools), soil moisture, landscape structure, calculation of vegetation indices (e.g. normalized difference vegetation index - NDVI), and amount of rainfall – for a review, check Beck, Lobitz and Wood (2000), Hay, Randolph and Rogers (2000), and Tatem and Hay (2004); (ii) generation of variables that characterize risk factors not collected on regular field surveys due to financial limitations, access constraints, or sampling design (Castro, Monte-Mór, Sawyer and Singer, 2006); (iii) construction of a spatial database containing roads, houses, hospitals and other key features – a recent example is shown by Kitron, Clennon, Cecere, Gurtler, King and Vazquez-Prokopec (2006).

The choice of remotely sensed imagery is not trivial. Currently there are more than 200 products to choose from past and current missions. While sometimes the choice is constrained by availability (e.g. multitemporal images for a retrospective study and lack of usable products due to cloud-coverage), often investigators may have multiple options to choose from. First, the choice depends on the disease under investigation and its associated risk factors, as well as the location where it occurs, among other factors intrinsic to the purpose and uses of the study (Beck, Lobitz and Wood, 2000). Second, the final choice may be determined by technical and financial issues (Perera and Tateishi, 1995). The latter concern, specially for developing countries, is likely to become less important over time, based on on-going initiatives that are progressively removing older barriers to the use of remotely sensed data – for a comprehensive review, check the National Research Council (2002) report. Meanwhile, the former concern becomes more complex over time as new high resolution products are made available, expanding even more the range of choices. A useful tool, particularly for beginners, is a search engine made available by CHAART – Center for Health Applications of Aerospace Related Technologies (http://geo.arc.nasa.gov/sge/health/sensor/senchar.html), which provides guidance on the best sensors based on the final use of the study (research, surveillance, risk assessment, prevention, or control of disease transmission). Although it does not point to a single product, it narrows down the options.

The choice of the spatial resolution is further complicated by the MAUP, and therefore studies are not scale independent. Relationships found at one scale may disappear or change in magnitude and direction when analyzed at another (Foody and Curran, 1994). Additionally, scaling-up or scaling-down the results of a study may generate misleading conclusions. That poses challenges especially for policy makers (Oliver, 2001). In that regard, multiscale studies present an opportunity to appraise both global patterns and individual processes of disease risk and transmission (Wilson, 2002). For a review of potential ways to overcome the MAUP in the context of remote sensing check Marceau and Hay (1999).

Deciding on the ideal spatial resolution may face yet another challenge – the lack of ground reliable data related to the disease under study at the same scale as the imagery (Wilson, 2002). Specifically for predictive models of disease transmission/prevalence, the lack of disease-related data undermines assessing the validity of the model (Atkinson and Graham, 2006). Although some countries may have administrative data reporting health events, the quality of the data is critical (De Savigny and Binka, 2004). Local level studies can benefit from specific surveys that collect reliable information at high spatial resolution – e.g. Demographic Surveillance Systems – DSS (Byass, Berhane, Emmelin, Kebede, Andersson, Hogberg and Wall, 2002).

Scale of questions & Scale of answers

Remote sensing has been and will continue to be an important tool for public health studies. Recent improvements in imagery resolution open new possibilities (Beck, Lobitz and Wood, 2000, Wilson, 2002) but also brings new challenges regarding imagery selection. Is there one particular sensor that is suitable for studies in urban areas? Is there an ideal scale that allows for the construction of predictive models of disease transmission? While a control effort requires detailed information of issues that impact on the levels and intensity of disease transmission, a research project may seek to investigate the patterns of disease prevalence given agriculture practices. Each study requires a different spatial resolution – finer in the former, coarser in the latter.

We argue that recommendations about spatial resolution depend not only on the area being studied (e.g. urban x rural) but also on the purpose of the research (e.g. identification of a pattern or explanation of a process), how the results will be used (e.g. budget allocation, research, program evaluation), and at what level (e.g. continental, national, local). Additionally, we argue that the use of multiple scales can facilitate the understanding of the underlying processes of transmission. Next we address each one of these issues.

The area being studied creates specific resolution demands. An urban context often requires very high spatial resolution (Welch, 1982). It tends to be heterogeneous, with many different elements of varied size on the ground. Therefore, images with low or medium spatial resolution are likely to present high pixel mixture (Tatem and Hay, 2004). Additionally, houses may be built with materials found on the ground, imposing additional challenges for image classification. Jensen and Cowen (1999) offer a comprehensive description of attributes that can be measured in urban contexts, and the required spatial, spectral and temporal resolutions. With the exception of meteorological data, finer spatial resolution is more appropriate. In contrast, rural areas may present a landscape dominated by large open fields, crops and pasture, scattered villages, or any combination of these. Medium spatial resolution imagery is likely to suit the needs of analysis, unless the goal is to precisely characterize the structural organization of ground elements (e.g. location of houses and road network). The spectral resolution in both areas is important. Multispectral or hyperspectral images are recommended in order to better characterize the different elements and physical/biological attributes that compose the scene.

With respect to the purpose of the research, public health studies may seek the identification of a pattern, the understanding of processes determining disease transmission, or the construction of a predictive model. The first type reveals how a certain disease is spatially distributed, but does not explain why. Since overall patterns do not seek to explain individual or localized variability, low spatial resolution data is suitable for this study (downscale generalizations, however, should be avoided). The second type seeks to identify factors determining the transmission of the disease. In this case, the individual/local variability is of crucial importance, and high spatial resolution is more appropriate (Levin, 1992). Nevertheless, other imagery could also be used in order to assess the study at a multiscale approach, identifying factors that determine transmission at different levels of scale. Finally, the third type is the ultimate goal: combining an observed pattern with identified determinants in a modeling approach that results in disease prediction, facilitating the task to establish a disease early warning system (Myers, Rogers, Cox, Flahault and Hay, 2000). As highlighted before, this is conditioned on the availability of reliable disease-related data.

The results of the study may have different uses, such as guide the definition of health priorities and budget allocation, surveillance, improve the design of alternative strategies for disease control, program evaluation, or pure research. With the exception of the latter, these applications involve decision making. Considering the importance of spatially targeted interventions for disease control (Carter, Mendis and Roberts, 2000), these decisions are likely to be made at the local level. Therefore, finer spatial resolution proves to be more appropriate for control efforts (Kitron, Clennon, Cecere, Gurtler, King and Vazquez-Prokopec, 2006). In a few instances, however, priority setting and budget allocation may be made at a national level. In this case, a coarser spatial resolution should fulfill the requirements (Randolph, 2000). Especially for surveillance and planning activities a medium temporal resolution is recommended in order to capture idiosyncratic short-term seasonal variability.

Finally, the geographical extent of the study can vary from worldwide to local. The larger the extent, the less important is the local variability, and therefore high spatial resolution is not needed. Studies with a broad geographical coverage most often portray patterns and do not seek for causes of that pattern, as highlighted above. These studies may be of limited use for disease control at the local level. Nevertheless, when combined with evaluations performed at a finer scale, they may contribute to the elaboration of an early warning system.

Table 1 summarizes the discussion presented above. When more than one resolution was selected for the same purpose, the choice depends on the type of area and the purpose of the research. We suggest that all levels of resolution are useful for research purposes – research questions may range from characterizing the continental pattern of diseases based on a climate model, or explaining the processes through which disease transmission takes place at a particular area. For any other purpose that involves decision-making, analyses must rely on finer resolutions. Mostly important, the choice of any spatial resolution needs to be supported by matching data related to the disease under study.

Table 1 – Recommended resolution of remotely sensed data for public health applications

Purpose of the study	Temporal resolution			Spectral resolution			Spatial resolution		
	Low	Medium	High	Low	Medium	High	Low	Medium	High
Research	X	X	X	X	X	X	X	X	X
Surveillance		X			X			X	X
Planning (1)		X			X			X	X
Program evaluation (2)	X				X			X	X

(1) Includes budget allocation and the definition of strategies for prevention or control of disease transmission. (2) Assumed to be always performed at the local level

It is clear that there is not a unique spatial resolution that is appropriate for public health applications. Above all, it is important to recognize that disease risk patterns operate at different spatial and temporal scales. As a result, an ideal study, yet labor and cost intensive, should adopt a multiscale approach (Levin, 1992, Foody and Curran, 1994), seeking to explain the determinants of diseases at each scale, but also assessing how they change as one moves from the local to the global scale. Additionally, studies may combine multiple imagery – e.g. Landsat

TM (30 m of spatial resolution) to characterize land use patters, Ikonos (1-4 m) or Quickbird (0.61 cm to 2.88 m) images to obtain road network and house location, and radar images (SARs - Synthetic Aperture Radars) to identify wet soils. They should also combine different sources of data – e.g. disease-related, demographic, ecological, political, and economic. Ultimately, such studies would be multiscale and multidisciplinary, constituting a major contribution to all purposes described in table 1.

Discussion

As proposed by Levin (1992), scale is a fundamental conceptual problem in science. Nonetheless, a debate around the usefulness of scale recently emerged among geographers (Marston, Jones III and Woodward, 2005, Hoefle, 2006). In the public health arena the importance of scale is beyond discussion. Determinants of disease risk and transmission operate at different scales. Characterizing malaria risk in a rapidly transforming environment such as the Amazon, for example, requires a consideration of biological and ecological phenomena acting at multiple spatial scales, juxtaposed with behavioral and economic conditions. In this regard, *frontier malaria* operates at three spatial scales: micro/individual, community, and national (Castro, Monte-Mór, Sawyer and Singer, 2006). This is just one example, among many, that substantiates the need for a multiscale/multidisciplinary analysis.

Although there is not a unique scale that is appropriate to address public health issues, there are some recommendations that can be put forth based on the targeted area, and the purposes and uses of the study. In that regard, in this paper we indicate a few suggestions towards the choice of temporal, spectral and spatial resolutions, with an emphasis on the latter.

References

Atkinson, P. & Curran, P. (1997). Choosing an appropriate spatial resolution for remote sensing investigations. *Photogrammetric Engineering and Remote Sensing*, 63(12), 1345-1351

Atkinson, P. M. & Graham, A. J. (2006). Issues of scale and uncertainty in the global remote sensing of disease. (In S. I. Hay, A. J. Graham & D. J. Rogers (Eds.), *Global mapping of infectious diseases: methods, examples and emerging applications* (pp. 79-118). Amsterdam; Boston: Elsevier.)

Beck, L. R., Lobitz, B. M. & Wood, B. L. (2000). Remote sensing and human health: new sensors and new opportunities. *Emerging Infectious Diseases*, 6(3), 217-226

Byass, P., Berhane, Y., Emmelin, A., Kebede, D., Andersson, T., Hogberg, U. & Wall, S. (2002). The role of demographic surveillance systems (DSS) in assessing the health of communities: an example from rural Ethiopia. *Public Health*, 116(3), 145-150

Carter, R., Mendis, K. N. & Roberts, D. (2000). Spatial targeting of interventions against malaria. *Bulletin of the World Health Organization*, 78(12), 1401-1411

Castro, M. C. (2007). Spatial Demography: an opportunity to improve policy making at diverse decision levels. *Population Research and Policy Review*, Forthcoming,

Castro, M. C., Monte-Mór, R. L., Sawyer, D. O. & Singer, B. H. (2006). Malaria Risk on the Amazon Frontier. *Proceedings of the National Academy of Sciences*, 103(7), 2452-2457

Curran, P. J. (1988). The semivariogram in remote sensing: an introduction. *Remote Sensing of Environment*, 24, 493-507

De Savigny, D. & Binka, F. N. (2004). Monitoring future impact on malaria burden in sub-Saharan Africa. *American Journal of Tropical Medicine and Hygiene*, 71(Suppl 2), 224-231

Dungan, J. L. (2001). Scaling up and scaling down: the relevance of the support effect on remote sensing of vegetation. (In N. J. Tate & P. M. Atkinson (Eds.), *Modelling Scale in Geographic Information Science* (pp. 221-235). Chichester: John Wiley & Sons, Ltd.)

Foody, G. M. & Curran, P. J. (1994). Scale and environmental remote sensing. (In G. M. Foody & P. J. Curran (Eds.), *Environmental remote sensing from regional to global scales* (pp. 223-232). Chichester; New York: John Wiley & Sons.)

Fotheringham, A. S., Brunsdon, C. & Charlton, M. (2000). *Quantitative geography: perspectives on spatial data analysis.* (London; Thousand Oaks, Calif.: Sage Publications)

Fox, J., Suryanata, K. & Hershock, P. (2005). *Mapping Communities: Ethics, Values, Practices.* (Honolulu, HI: East-West Center)

Hay, S. I. (2000). An overview of remote sensing and geodesy for epidemiology and public health application. (In S. I. Hay, S. E. Randolph & D. J. Rogers (Eds.), *Remote sensing and geographical information systems in epidemiology* (pp. 1-35). San Diego, CA; London: Academic Press.)

Hay, S. I., Randolph, S. E. & Rogers, D. J. (2000). *Remote sensing and geographical information systems in epidemiology.* (San Diego, CA; London: Academic Press)

Hay, S. I., Snow, R. W. & Rogers, D. J. (1998). From predicting mosquito habitat to malaria seasons using remotely sensed data: practice, problems and perspectives. *Parasitology Today*, 14, 306-313

Hoefle, S. W. (2006). Eliminating scale and killing the goose that laid the golden egg? *Transactions of the Institute of British Geographers*, 31, 238-243

Jensen, J. & Cowen, D. (1999). Remote sensing of urban suburban infrastructure and socio-economic attributes. *Photogrammetric Engineering and Remote Sensing*, 65(5), 611-622

Jensen, J. R. (1996). *Introductory digital image processing: a remote sensing perspective.* (Upper Saddle River, N.J.: Prentice Hall)

Jensen, J. R. (2007). *Remote sensing of the environment: an earth resource perspective.* (Upper Saddle River, NJ: Pearson Prentice Hall)

Kitron, U., Clennon, J. A., Cecere, M. C., Gurtler, R. E., King, C. H. & Vazquez-Prokopec, G. (2006). Upscale or downscale: applications of fine scale remotely sensed data to Chagas disease in Argentina and schistosomiasis in Kenya. *Geospatial Health*, 1, 49-58

Levin, S. A. (1992). The problem of pattern and scale in ecology. *Ecology*, 73(6), 1943-1967

Marceau, D. J. & Hay, G. J. (1999). Remote sensing contributions to the scale issue. *Canadian Journal of Remote Sensing*, 25(4), 357-366

Marston, S. A., Jones III, J. P. & Woodward, K. (2005). Human geography without scale. *Transactions of the Institute of British Geographers*, 30, 416-32

Myers, M. F., Rogers, D. J., Cox, J., Flahault, A. & Hay, S. I. (2000). Forecasting disease risk for increased epidemic preparedness in public health. (In S. I. Hay, S. E. Randolph & D. J. Rogers (Eds.), *Remote sensing and geographical information systems in epidemiology* (pp. 309-330). San Diego, CA; London: Academic Press.)

National Research Council (U.S.) (2002). *Down to earth: geographic information for sustainable development in Africa.* (Washington, D.C.: National Academies Press)

Oliver, M. A. (2001). Determining the spatial scale of variation in environmental properties using the variogram. (In N. J. Tate & P. M. Atkinson (Eds.), *Modelling scale in geographical information science* (pp. 193-219). Chichester; New York: Wiley.)

Openshaw, S. (1984). *The modifiable areal unit problem.* (Norwick, UK: GeoBooks)

Perera, L. K. & Tateishi, R. (1995). Do remote sensing and GIS have a practical applicability in developing countries? (including some Sri Lankan experiences). *International Journal of Remote Sensing*, 16(1), 35-51

Pryor, R. J. (1984). Methodological problems in Population Geography. (In J. I. Clarke (Ed.), *Geography and population: approaches and applications* (pp. 25-34). 1st ed. Oxford; New York: Pergamon Press.)

Randolph, S. E. (2000). Ticks and tick-borne disease systems in space and from space. (In S. I. Hay, S. E. Randolph & D. J. Rogers (Eds.), *Remote sensing and geographical information systems in epidemiology* (pp. 217-243). San Diego, CA; London: Academic Press.)

Tatem, A. J. & Hay, S. I. (2004). Measuring urbanization pattern and extent for malaria research: a review of remote sensing approaches. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, 81(3), 363-376

VanWey, L. K., Rindfuss, R. R., Gutmann, M. P., Entwisle, B. & Balk, D. L. (2005). Confidentiality and spatially explicit data: concerns and challenges. *Proceedings of the National Academy of Sciences of the United States of America*, 102(43), 15337-15342

Welch, R. (1982). Spatial resolution requirements for urban studies. *International Journal of Remote Sensing*, 3(2), 139-146

Wilson, M. L. (2002). Emerging and vector-borne diseases: role of high spatial resolution and hyperspectral images in analyses and forecasts. *Journal of Geographical Systems*, 4, 31-42

Woodcock, C. E. & Strahler, A. H. (1987). The factor of scale in remote sensing. *Remote Sensing of Environment*, 21, 311-332

Woodcock, C. E., Strahler, A. H. & Jupp, D. L. B. (1988). The use of variograms in remote sensing: II. Real digital images. *Remote Sensing of Environment*, 25, 349-379