CHAPTER 4

Spatial Dimension of Scaling Up and Out

*Simon E. Cook** and *Sam Fujisaka***

Introduction: The Concept of Space and Scale

The conceptualization of space has preoccupied philosophers and scientists since Aristotle's *Physics* (Couclelis, 1998), which expressed concepts to help understand inert entities that exist in space, and human interactions with them. A succession of ideas has enlarged this understanding, from the positivist *absolute space* of Newton to constructivist positions proposed by Werlen (1993).

Scale is a concept used to manage information about the real world and to summarize observations about complex phenomena that vary within space, time, or other dimensions. The ordering of phenomena according to scale enables human beings to store, recall, and analyze information about features that would otherwise be impossible to evaluate. The concept is essential to researchers of agricultural development processes because scale organizes our understanding of complex socioeconomic and biophysical processes that interact in space (valley, region, continent), time (daily, annual), and institutions (household, community, nation). Scale is especially useful where variation is essentially "lumpy".

Scale is perceived differently by the respective disciplines that attempt to deal with it (Marceau, 1999). Social or economic systems and biophysical systems tend to be referenced internally. That is, social networks are described without reference to biophysical characteristics, and biophysical with weak reference to socioeconomic. Since both overlap in space and time, the distinction between social and biophysical systems is to some extent arbitrary. Such systems can be modeled explicitly in space by expressing interactions formally in a way that can be observed.

---

* Project Manager, Land Use Project, Centro Internacional de Agricultura Tropical (CIAT), Cali, Colombia.
** Consultant.

Most studies that use geographic information systems (GIS) to represent specific scale-dependent entities do so to represent biophysical processes. The concept of scale in space and time has been a major preoccupation with biophysical sciences for some time because of the obvious connotations scale has on process, on research domains, and on the validity of extrapolation and interpolation. Scale is essential to understand fundamental biophysical phenomena that are too big, too small, too fast, or too slow to be observed directly. Notwithstanding the search for scale invariance, predominantly in the natural and information sciences (e.g., see Burrough, 1981; Barabasi and Albert, 1999; Gisiger, 2001), processes are generally assumed to be scale dependent and to operate within predominant *domains*. Moving up or down scale from the domain at which the concept was developed introduces additional uncertainties because of phenomena referred to as ecological or atomistic fallacies.

Gimblett (2002) points out that biophysical studies of scale dependence tend to neglect the human dimensions of such systems, specifically through the use of social science data and modern intelligent simulation techniques. Increasing attention has been directed in recent years to spatial and temporal dimensions of social interactions. Giddens (1984) sees space as both a medium of social relations and a material product that influences interactions. Raedeke and Rikoon (1997) identify time and space as fundamental categories of human experience.

Social scientists use scaling concepts to describe variations of purely human phenomena such as institutions and policy (Gibson et al., 2000). Indeed, the concepts of scaling up and scaling out, as described in companion chapters to this, are applied equally to institutional dimensions and the biophysical environment in which they exist. Institutions, however, exist in real space and are inevitably influenced to some degree by variation in spatial dimensions. It is essential to describe spatial characteristics where spatial variation is significant to the processes being examined.

Variation in the spatial dimensions of scale should be described if such variation significantly influences the validity of representation; that is, if representation of the location, size, or spatial proximity between entities helps identify the process. However, broadening the concept of scaling to include spatial dimensions increases the complexity of analysis, and few would willingly embark on this process if it cannot be shown to be necessary. Therefore, our first objective is to clarify when the spatial dimension is significant to scaling up and scaling out in the socioeconomic sense (as is used predominantly by other chapters in this book).

Acknowledging that variation within real space may be significant, the second objective of the chapter is to identify the various modeling approaches that can be used to describe spatial variation.

When Is the Spatial Dimension Significant to Scaling Up and Out?

Basic concepts

Scaling out means extending scope by repeating a process at one scale to other individuals of about the same scale. This process is influenced by the variation among individuals.

Scaling up occurs when the dimension of a process is increased, for example, when a social interaction process increases from village to municipality; or a hydrologic process from first- to third-order catchments. For reasons discussed below, changing the dimensions of a process from that at which it has been observed almost always introduces new sources of uncertainty.

Scaling up tries to represent a process of interaction among individuals that becomes bigger or more distant. The process can be social, such as cooperation between two people within a community, or economic, such as trade between two countries. It can be biophysical, for example gas exchange between trees and the atmosphere, or water flow between an irrigation plot and a river. The essential feature is that exchange or flow occurs between two or more individuals. The objective of the study is to understand the nature of the interaction. Having understood the interaction process better, the objective of intervention is to improve the overall result.

Modeling the scaling process

Since a major purpose of this chapter is to examine the effects of space on processes of scaling out (dissemination), we need to define the concept of interactions as they occur in physical space. We do this by describing a basic model of interaction between two individuals that accounts for spatial characteristics such as size, location, distance, or direction.

Intuitively, these spatial characteristics seem important to the scaling processes. Few would doubt that interaction between neighbors is more likely than between individuals in different continents, or that groups of similar size interact more easily, or that germplasm is more likely to flow between sites with similar environments. But the question is whether the effects can be described in a form that can be analyzed. To do this, the scaling process is rephrased in a more analyzable form by considering it as a quantifiable process of attraction between pairs of individuals. Space influences two broad types of characteristics significant to the interaction process. First, the strength of attraction is determined by the suitability of individuals for interaction to occur, for which information about their location and size is useful. Second, the interaction may be influenced by resistance or loss of signal that might result from distance, friction of the surface over which interaction is attempted, or interruption from intervening processes. We illustrate below a selection of spatial modeling processes that attempt to quantify these effects.

Assessing the significance of the spatial dimension

Failure to acknowledge the influence of spatial processes reduces the accuracy of statements that can be made about the process of interaction. The question is, "How significant is the variation, and how can it be predicted?" The significance of the spatial dimension to the scaling process can be assessed from the following questions:

Does the process of interaction change significantly if:

(1) It is larger or smaller, that is, if the process applies to more or fewer individuals?
(2) Individuals are in particular locations?
(3) It occurs over different distances?
(4) It occurs in different directions?
(5) If individuals are formed in different configurations?

Question 1: Does size matter? More individuals are involved as processes are scaled up or scaled out (i.e., extended to reach individuals further away). Larger processes encounter more cumulative variation among individuals simply because individuals differ, and according to the standard tenets of probability theory (albeit that individual variation is not distributed randomly). Processes that are highly adapted to a particular group of individuals become less and less suitable as individuals become more dissimilar over space. Conversely, processes that are described for large-scale phenomena encounter a reverse problem when applied to subpopulations. Aggregate solutions that exist for a whole population become increasingly at variance with individuals as the process is scaled down.

Scientists tend to adopt a pragmatic approach to scale by predefining the object of study—for example, *global* climate change, a study of *catchment* process, or development of *community* preferences. Experienced practitioners make reasonable assumptions about the limitations of scaling up or down from the definition. Few hydrologists would contemplate applying models developed from measurements at field plot level to large regions, even fewer would attempt to predict field conditions from global-scale models: Social scientists have developed methods of analysis that are specific to individuals, families, communities, persons affiliated via different kinship systems, and to aggregates of people tied together through political units such as municipalities, states, or countries.

Methods of numerical spatial analysis to describe the deviations of individuals from purely random patterns include measures of spatial autocorrelation, clustering, or geographically weighted regressions (see Cliff and Ord, 1981; Diggle, 2003).

Question 2: Is location significant? Information about the non-random spatial variation of individuals can be used to assess the likely degree of interaction between them. A site that has a suitable condition increases the likelihood of adoption during scaling up or out. For example, a crop that is bred at a dryland research station is more likely to succeed in other dryland sites than in humid ones, even if such areas are closer. A technological innovation to ease cultivation that is suited for men is unlikely to be successful in locations where most farmers are women.

In these cases, information about location is used to carry knowledge about site conditions that determine whether interaction takes place. Classifying sites quantitatively according to similarity improves definition. An example is the exchange process provided by the assessment of site suitability for germplasm transfer in models such as FloraMap (Jones and Gladkov, 1999). FloraMap is premised on the assumption that climate at a site strongly influences the regional distribution of germplasm.

Question 3: Is process influenced by distance? In addition to the effect of increasing variance that occurs as populations expand (documented well by geostatistical theory [see Isaaks and Srivanasta, 1989]), distance will decrease the strength of interactions among individuals. As distance increases, the chances that an intervening process will occur that influences individuals in a different way also increases. Sociologists have for some time referred to path length between two individuals as a key factor in determining networking (Newman et al., 2002).

The effect of distance formed the basis of classic spatially sensitive theories of geo-economic development from von Thunen and Christaller in the last century. Recent additions to economic theory include Vickerman et al. (1999). Inadequate infrastructure hinders people's access of people to markets, services, and one another, leading to the so-called "spatial poverty trap" (Ravallion, 1997). Because accessibility and its inverse, isolation, are considered significant factors in development (Deichmann, 2001), tools have been developed to model explicitly these factors over space (e.g., Farrow and Nelson, 2001). In this model, accessibility is defined by the shortest travel cost distance, accounting for the cumulative distance over which exchange occurs and the friction of the surface. Factors not explicitly modeled by Farrow and Nelson (2001) include the cost of transport (roughly equivalent to surface friction) and the opportunity costs to individuals.

Question 4: Is direction significant? This question concerns spatial anisotropy of *process*, that is, variation that is introduced solely by a change in direction of process, such as dispersion across- or down-stream.

This is in addition to effects that are caused by variation of site characteristics such as landscape obstacles.

Anisotropy can be represented quantitatively using geostatistical models of anisotropic spatial dependence (see Isaaks and Srivastava, 1989). The vector-dependent processes have been modeled successfully to reveal spread along transport networks (Deichman, 2001) and spread of diseases over two-dimensional grids using a process of Eulerization (Colville and Briggs, 2000).

Question 5: Is the configuration of multiple individuals significant? The sections above describe how size, location, distance, or direction influences a single process of interaction between two individuals or nodes. In reality, of course, interactions are not restricted to two individuals or one process, but occur simultaneously among many nodes with multiple processes. Therefore we now move on to outline some concepts that are used to analyze the more complex spatial characteristics of *multiple* interacting nodes. We describe three approaches to illustrate the development of analytical models: Network analysis, analysis of cellular automata (CA), and the multi-agent system (MAS).

Network analysis has a long history in the social and economic sciences to describe interactions among multiple individuals through descriptors such as connectivity, accessibility, or path length analysis. This use is almost exclusively aspatial, that is, analysis is concerned primarily with network topology rather than the spatial characteristics that influence the state of individuals or the interactions among them.

Recent additions to the social science literature provide examples of explicit modeling of social networks. Newman et al. (2002) describe network activity quantitatively within different groups of social actors and test the model by comparing predicted with actual measures of network function. Burt (2002) evaluates network function within a commercial organization and identifies clear relationships between an individual's position within a network and his/her apparent activity. Several other examples exist of models that attempt to quantify social or economic behavior, while assuming biophysical variation to be insignificant.

In the late 1980s and early 1990s, spatial scientists realized that complex socioeconomic processes of land use change could not be predicted exactly by rigid mathematical models, regardless of how elaborate they might be. In parallel, the expansion of insights into ecological processes revealed that existing models based on simple deterministic functions could not describe behaviors such as complexity, self-organization, chaos, and multi-scale functionality (Heylighen, 1999). Janowski and Richard (1994) note the shortcomings of trying to oversimplify social processes in space: "Current GIS analysis is based on simple spatial geometric processing operations such as overlay comparison, proximity measures, and buffering. It does not provide optimization, iterative equation solving, and simulation capabilities necessary in planning".

Anthropologists and archeologists have more recently examined fractal patterns of human settlements, linking the emergence of higher fractal dimensions (a measure of how quickly self-similarity patterns scale up) to higher efficiency and, possibly, eventually to social collapse (e.g., the prehistoric lowland Maya). Analysis of both systems efficiency and eventual collapse is based on spatial analysis of the simultaneous scaling out and "scaling in" (a sort of urban intensification) of such self-organizing complex systems (Ravilious, 2004).

Out of such thinking arose a new breed of models that treated complex socioeconomic systems as self-organizing, partially predictable systems, but for which models became tools with which to visualize complex dynamics rather than the basis on which to make definitive statements.

A different model emerged, based on the concepts of CA developed by von Neuman over 60 years ago, in which complex dynamic behavior is described as the result of relatively simple transition "rules", which govern the rate and direction of change amongst individual "cells". While Wu and Marceau (2002) observe that the concepts of self-organization, emergence, and order date back to even earlier ideas of ecologists such as Clements (1916), the realization of such theories has been strengthened considerably through the ability to model such hypotheses in GIS.

The basic principle includes concepts of site suitability and access. As explained by Engelen et al. (1997), however, CA models also include a neighborhood function, to account for intrinsically spatial features such as agglomeration, dispersion, or other pattern-creating processes. Cells are allowed to take a number of states (z), and are expected to allocate themselves into whichever state seems most probable, according to the general expressions:

$$P(z) = v(S_z A_z N_z)$$

Where $P(z)$ is the potential for transition into state z; S signifies suitability of the cell for state z; A accessibility to acquire state z; and N the neighborhood effect on z. Parameter v is simply a stochastic disturbance term, which can be adjusted to accommodate random effects. Cells are shown to change to the state that acquires the highest transition potential (P_Z).

Rules in CA models can be quantitative or qualitative, and deterministic or stochastic. Some CA models attempt to condense complex behavior into two or three rules. The number of rules can be increased to represent the richness of dynamic processes that are believed to operate within a region.

While CA theory has successfully demonstrated a few patterns of behavior, some modelers regard the technique as inadequate to describe complex human systems because of its dependence on rigid spatial structure and synchronicity between processes. In response, increasing attention is being directed towards methods of agent-based modeling (ABM), which can represent dynamic patterns of individual behavior within complex biophysical and social systems. Whereas behavior in CA models is generalized, processes in ABMs are object oriented and can perform asynchronously to one another. This distinctly "bottom-up" approach of individual models has the advantage of recognizing individual complexity (Judson, 1994). As perceived by Wu and Marceau (2002), complex systems are so because they are not completely reducible to components. The challenge remains to determine rigorous theories of behavior that are also comprehensive enough to be realistic. Many examples exist of the use of ABMs in ecological and land use studies that purport to represent dynamic processes of diffusion and change (Parker et al., 2002).

Discussion

This chapter reviews methods that enable social and biological scientists to represent processes of interactions between individuals with explicit definition of spatial attributes such as size, distance, and direction. Through these methods it is possible to show how space influences the diffusion processes as they involve more individuals (scaling out) or individuals at higher levels of organization (scaling up). The advantage of developing explicit models is that predictions of interactions can be tested against observation and used to reveal obstacles to beneficial diffusion.

The steady improvement in the ease of use of GIS, coupled with the availability of better coverage of spatial data, has enabled scientists to analyze social behavior within a biophysical setting more realistically, and to create models to reflect their understanding of complex processes. Indeed, the understanding that a priori oversimplification can actually obstruct accurate modeling of complex processes stimulated the development of ABMs. The question is, "What is the gain of more complex models, and in what circumstances are such models essential for reasonable representation of complex processes of human interaction?"

The third major feature of relevance to scaling up and out is the dramatic increase in availability of spatial data, against which such complex models can be tested. These data describe both the "*y*"s (population, income, adoption, etc.) and "*x*"s (environmental attributes that modify the influences). Whilst data are more available for remotely sensed biophysical attributes, this also influences the resolution of insight for social phenomena (Geores, 2000).

Conclusions

Processes of scaling up or scaling down involve a change in multiple interactions between pairs of individuals. The change due to scaling up or out occurs through extension or diffusion to actors that are more distant or closer, or as repetition within new pairs of actors. The spatial dimension is significant to these processes when the likelihood of interactions is influenced by spatial attributes of size, location, distance, and direction. This influence is exerted in a number of ways—through the size-related variation of actors, their location-determined suitability for interaction, and the distance- or direction- determined cost of interaction.

The effects of space on scaling processes can be modeled in GIS in a number of ways. The effect of changes in size and location on likely interactions is modeled through its known effect on variation of individuals; for example, a community-scale institute will prove incapable of handling the uncertainty of national-scale problems; germplasm spreads more easily to locations with similar characteristics. Distance and direction effects are more effectively modeled as repetitive transfers over a variable surface. More complex processes of self-organization of individuals can also be represented through rule-based or agent-based models that use these spatial attributes to modify individual transactions within the overall system.

These techniques exist to describe spatial effects on scaling processes, but they require significant effort and data to achieve a useful accuracy. So, are the benefits adequate? The first major advantage of formally modeling spatial influence on scaling processes is that the processes can be understood more completely through visualization of effects over "real" areas, as represented by maps. Patterns and associations with cultural or biophysical variables may become evident only after the process is represented spatially. A second advantage is that spatial analysis of processes of diffusion can identify and quantify constraints to scaling processes that may not be evident before the information is assembled within a spatial context. Spatial epidemiology provides the most obvious example where vectors have been identified only *after* spatial representation, but many other examples exist in social science or economic literature where diffusion processes could be explained most easily in relation to spatial features such as roads or geographic clusters. Finally, through spatial modeling it is becoming increasing possible to predict complex diffusion processes realistically to identify the likely influence of changes in policy, markets, or biophysical change. Through rule-based and agent-based models it is possible to represent increasingly complex social and biophysical effects on individual-to-interactions as they scale up and out within an uncertain world.

References

Barabasi, A.; Albert, R. 1999. Emergence of scaling in random networks. Science 286:509-512.

Burrough, P.A. 1981. Fractal dimensions of landscapes and other environmental data. Nature 294:240-243.

Burt, R.S. 2002. The social capital of structural holes. *In*: Guillen, M.F.; Collins, R.; England, P.; Meyer, M. (eds.). New economic sociology: Developments in an emerging field. Russell Sage Foundation, NY, USA. p. 202-247. Available in: http://gsbwww.uchicago.edu/fac/ronald.burt/research/SCSH.pdf

Clements, F.E. 1916. Plant succession: An analysis of the development of vegetation. Carnegie Institute of Washington, USA. 512 p.

Cliff, A.D.; Ord, J.K. 1981. Spatial processes: Models and applications. Pion, London, GB. 266 p.

Colville, R.; Briggs, D.J. 2000. Dispersion modelling. *In:* Elliott, P.; Wakefield, J.C.; Best, N.G.; Briggs, D.J. (eds.). Spatial epidemiology: Methods and applications. Oxford, NY, USA. p. 375-392.

Couclelis, H. 1998. Aristotle and spatial dynamics in the age of geographic information systems. *In*: Egenhofer, M.J.; Golledge, R.G. (eds.). Spatial and temporal reasoning in geographic systems. Oxford University Press, GB. p. 109-118.

Deichmann, U. 2001. African population database documentations. Available in: http://grid2.cr.usgs.gov/globalpop/africa/

Diggle, P.J. 2003. Statistical analysis of spatial point patterns. 2nd edition. Oxford U.P., NY, USA. 288 p.

Engelen, G.; White, R.; Uljee, I. 1997. Integrating constrained cellular automata models, GIS and decision support tools for urban planning and policy making. *In:* Timmermans, H.P.J. (ed.). Decision support systems for urban planning. E&FN Spon., London, GB. p. 125-155.

Farrow, A.; Nelson, A. 2001. Accessibility modelling in ArcView. An extension for computing travel time and market catchment information. Available in: http://www.ciat.cgiar.org/access/pdf/ciat_access.pdf

Geores, M. 2000. Considering scale in resource definition. Presented at “Constituting the commons: Crafting sustainable commons in the new millennium”, 8th Conference of International Association for Study of Common Property, 31 May-4 June, Bloomington, IN, USA.

Gibson, C.C.; Orstom, E.; Ahn, T. 2000. The concept of scale and the human dimensions of global change—A survey. Ecol. Econ. 32:217-239.

Giddens, A. 1984. The constitution of society: Outline of the theory of structuration. Polity press, Cambridge, GB. 402 p.

Gimblett, H.R. 2002. Integrating geographic information systems and agent-based technologies for modelling and simulating social and ecological phenomena. *In*: Gimblett, H.R. (ed.). Integrating geographic information systems and agent-based modelling techniques for simulating social and ecological processes. Oxford Univ. Press, GB. p. 1-21.

Gisiger, T. 2001. Scale variance in biology: Coincidence and footprint of a universal mechanism? Biol. Rev. 76:161-209.

Heylighen, F. 1999. The science of self-organization and adaptivity. *In*: Heylighen, F. (ed.). The encyclopedia of life support systems. Collective intelligence and its implementation on the Web: Algorithms to develop a collective mental map. Comput. Math. Theory 5(3):253-280.

Isaaks, E.H.; Srivastava, R.M. 1989. Applied geostatistics. Oxford University Press, NY, USA. 561 p.

Jankowski, P.; Richard L. 1994. Integration of GIS-based suitability analysis and multicriteria evaluation in a spatial decision support system for route selection. Environ. Planning B 21:323-340.

Jones, P.G.; Gladkov, A. 1999. FloraMap: A computer tool for predicting the distribution of plants and other organisms in the wild; version 1. CIAT CD-ROM Series. Centro Internacional de Agricultura Tropical (CIAT), Cali, CO. 1 CD-ROM with 87 p indexed manual.

Judson, O. 1994. The rise of the individual-based model in ecology. Trends Ecol. & Evol. 9(1):9-14.

Marceau, D.J. 1999. The scale issue in the social and natural sciences. Entropy, information, and evolution: New perspectives on scaling of biophysical processes. Can. J. Remote Sensing 25:347-356.

Newman, M.E.J.; Watts, D.J.; Strogatz, S.H. 2002. Random graph models of social networks. Proc. Nat. Acad. Sci. 99(1):2566-2572.

Parker, D.C.; Berger, T.G.; Aquino, C.M. 2002. Agent-based models of land-use and land-cover change: Report and review of an international workshop, 4-7 October 2001, Irvine, CA. Indiana University, USA. 124 p.

Raedeke, A.H.; Rikoon, J.S. 1997. Temporal and spatial dimensions of knowledge: Implications for sustainable agriculture. Agric. Hum. Values 14:145-158.

Ravallion, M. 1997. Good and bad growth: The human development reports. World Dev. 25(5):631-638.

Ravilious, K. 2004. Destined for destruction. New Scientist 10 January 2004:42-45.

Vickerman, R.K.; Spiekermann, K.; Wegener, M. 1999. Accessibility and economic development in Europe. Reg. Stud. 33(1):1-15.

Werlen, B. 1993. Society, action and space. An alternative human geography. Routledge, NY, USA. 249 p.

Wu, J.; Marceau, D. 2002. Modelling complex ecological systems: An introduction. Ecol. Model. 153:1-6.