Seth Spielman
Scale and the Decomposability Problem

Scale has been a thorn in the side of quantitative spatial analysis for decades. The most widely cited works on the "Scale Problem" in the discipline are largely descriptions of its existence (e.g. Fotheringham and Wong, work by Openshaw). However, describing a problem is very different from doing something about it, and in terms of "solutions" to the scale problem the discipline of geography has underperformed. Current practice in the discipline includes all sorts of statistically dodgy stuff - like doing the same analysis the same time at multiple scales and arguing that the right scale is the one highest model fit (see Spielman 2013, 2009). In the era of big data the scale problem is exacerbated because rather than working with precomputed area levels measures analysts have enormous flexibility in determining the scale of their analysis.

In this position paper I argue that the scale problem is related to the "decomposability" of systems and that many of the systems of interest to geographers are not decomposable.

Concerns about scale arise in the study of geography because the concepts under study can be operationalized at a variety of scales. In human geography, for example, the concept of racial segregation and vulnerability to natural hazards are vague regarding scale and there is a cottage industry of descriptive papers calculating social phenomena at multiple scales.

There is a tendency in geography to treat scale as "modifiable" – dial one turns up or down; or in a "multi-scale" approach a way of gathering multiple measures of the same phenomena. However, scale need not be arbitrary; in the era of large scale disaggregate data researchers can design geographic units of analysis. Given this flexibility, as Openshaw (1996) and King (1997) note, the scale becomes a theoretical question about the appropriate unit of analysis. The identification of "appropriate" units of analysis, for many geographic problems, is easier said than done.

In the late 30's and early 40's there was a debate about the appropriate unit of analysis for urban research. Writing in the American Sociological Review Quinn (1940) writes in response to Alihan's (1938) critical analysis of the human ecological movement which used the newly designed census zones as "natural areas" and unit of analysis,

> "Alihan sharply attacks the notion that zones are natural areas. Her arguments follow two main lines: (i) that radial gradients-which involve gradual increases or decreases of phenomena as one moves away from the center of dominance-make impossible the conception of zones demarcated by sharp breaks in criteria; and (2) that different criteria show different distributions so that no single set of zones can be drawn to include all of them. Alihan's arguments have both merits and weaknesses.
>
> The contention by Alihan that a gradual gradient makes impossible the existence of zones does not seem valid. In the field of physics, for example, the gradual change in the length of light rays throughout a rainbow spectrum may be taken as an example of a gradient. Nevertheless, distinct zones of red, yellow, and blue appear in the rainbow even though no sharp line of demarcation can be drawn between them. It seems possible, therefore, for distinct zones to appear even where gradients unquestionably exist.
>
> Alihan's second argument deserves careful checking. She points out that the zones drawn by persons studying domestic organization do not have boundaries identical with those drawn by

> persons investigating delinquency or mental disorders. Unless the basic criteria used in characterizing zones show essentially similar distributions, she declares that a single set of natural-area zones cannot be drawn… Existing studies do not clearly settle this point raised by Alihan. Further research is needed which (i) carefully describes the significant ecological criteria to be used and (2) determines whether they can be correlated to delimit natural-area ecological zones."

Almost 80 years on this debate is strikingly relevant, the identification of relevant levels of abstraction, scales at which the environment will be observed, measured, and studied is no less challenging a problem the better part of a century after Quinn's writing.

My position is that geographers have been doing scale wrong – we've approached it as a statistical problem, a descriptive problem, even as theoretical problem – when in fact scale is an ontological problem that is tightly coupled to the decomposibility of geographic systems.

In the "Architecture of Complexity" Simon (1962) outlines the foundations of hierarchy theory. Hierarchy theory, relates to the representation of a complex system using a series of levels (Ahl and Allen, 1996), Phillips (2004) notes that hierarchy theory is the basis of the physical geographic view of scale. Many physical, social, and symbolic systems are hierarchical in nature. Simon defines hierarchy not in terms of power relationships (super and subordinate levels) but in terms of information flow and interaction. According to Simon hierarchies have more interactions within levels than between levels, by simply looking at the interaction among parts of a system one can identify the hierarchical structure. University hierarchies can be identified within this framework, people within departments have stronger interactions than across departments. Departments communicate with executives (deans, provosts, presidents, etc.); these up-slope interactions will be less intense and less frequent than interactions between executives or peers in the same department.

Systems that behave this way, where interactions within levels are stronger than interactions across levels are "nearly decomposable." The term "near-decomposability" refers to the fact that because the links between levels are relatively weak one can consider the levels independently. Simon (1962, p. 475) indicates that the threshold for near-decomposability is a matter of judgment. Ahl and Allen (1996) extend this idea pointing out that sometimes identifying the levels of a system is not trivial. This is especially true in the study of human geography, a field in which David Harvey notes, "Everything is related to everything else…"

As the strength of the interaction between levels in a system increases the levels themselves become increasingly arbitrary. If a system is not "decomposable" it becomes difficult to sort out the relationships between levels. Geographic systems, especially human geographic systems, are difficult to decompose.

The decomposability problem throws a wrench into human geographic research becaue it makes it very difficult to determine appropriate units of analysis. I also believe complicates social measurement (but I don't have space to get into that in this position paper).

To sum up as I hit my two page limit. I do not believe that we can determine the spatial scale at which processes operate using purely statistical methods, the decomposability problem mucks up inference. The question we should be asking is not, "How should we (geographers) analytically deal with scale?" it is "How should we study non-decomposable geographic systems?"