Progress in Physical Geography 31(5) (2007) pp. 471–479

The modifiable areal unit problem (*MAUP*) in physical geography

Shawna J. Dark[1*] and Danielle Bram[1]

[1]California State University, Northridge, CA 91330, USA

Abstract: Of particular importance to the study of large-scale phenomena in physical geography is the modifiable areal unit problem (*MAUP*). While often viewed as only a problem in human geography (particularly demographic studies), the *MAUP* is an issue for all quantitative studies in geography of spatial phenomena (Openshaw and Taylor, 1979). Increasingly, remote sensing and Geographic Information Systems (*GIS*) are being used to assess the distribution of phenomena from a large scale. These phenomena are modelled using areal units that can take any shape or size resulting in complications with statistical analysis related to both the scale and method used to create the areal units. In this paper, we define the modifiable areal unit problem, present examples of when it is a problem in physical geography studies, and review some potential solutions to the problem. Our aim is to increase awareness of this complicated issue and to promote further discussion and interest in this topic.

Key words: aggregation, GIS, hydrological modelling, modifiable areal unit problem (MAUP), scale

I Introduction

The study of physical geography naturally lends itself toward large spatial scale analyses. Over the past 15 years, the development of remote sensing techniques and sophisticated Geographic Information Systems (*GIS*) software has led to an increase in quantitative studies within physical geography conducted at large spatial scales such as the landscape and regional scales (Rosswell *et al.*, 1991; Cain *et al.*, 1997; Davis *et al.*, 1998; McDermid *et al.*, 2005).

Of particular importance to the study of large scale phenomena is the modifiable areal unit problem (*MAUP*). In geography, we use modifiable areal units in quantitative analysis (Openshaw and Taylor, 1979). These areal units can take any shape or size, resulting in complications with statistical analysis related to both scale and the method used to create the areal units. For example, in biogeographic studies, much of our knowledge of large-scale phenomena is derived from the aggregation of area-based information obtained from small areas (less than 1 km^2) or from data collected at specific predefined scales such as with remotely sensed imagery. This type of data and subsequent analyses may not be

*Author for correspondence. Email: shawna.dark@csun.edu

 DOI: 10.1177/0309133307083294

truly representative of the scale of the phenomena under examination (Burke *et al.*, 1991). In addition, the selected areal unit may have strong implications for statistical analysis (Openshaw and Taylor, 1979).

While the issue of scale has been widely examined in various aspects of physical geography, the *MAUP* has been largely ignored despite its presence in various types of large-scale spatial data analysis. The primary goal of this paper is to provide a comprehensive review of the *MAUP* in physical geography. As such, we will review the impacts of the *MAUP* particularly relative to the analysis of spatially explicit data, provide a review of the presence of the *MAUP* in key aspects of geographic research related to physical geography, and examine the potential solutions to the *MAUP*.

II The modifiable areal unit problem

For the purpose of analysis, a study area can be divided into non-overlapping areal units in a variety of ways, therefore the *MAUP* can exist in any aspect of large-scale, spatially explicit data (Openshaw, 1984; Marceau and Hay, 1999). There are two issues of concern related to the *MAUP*.

1. The scale effect is attributed to variation in numerical results owing strictly to the number of areal units used in the analysis of a given area (Openshaw and Taylor, 1979).
2. The zonation effect is attributed to changes in numerical results owing strictly to the manner in which a larger number of smaller areal units are grouped into a smaller number of larger areal units (Openshaw and Taylor, 1979).

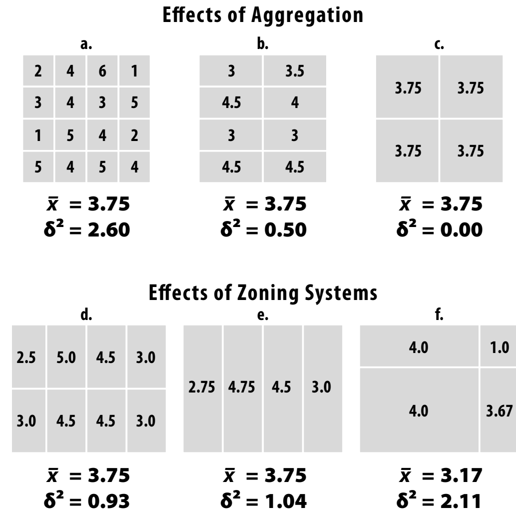
The first concern focuses on the issue of scale and variation. When areal units are aggregated into fewer, larger units for statistical analysis, values associated with the variation of the data decrease which will affect any associated statistical analysis. The second concern focuses on aggregation and the variation in results from statistical analysis as a result of alternative combinations of areal units at similar scales (Openshaw, 1984; Marceau and Hay, 1999). Several studies have confirmed that statistical results vary based on scale and aggregation which is a cause for concern for anyone conducting research with geographic data (Fotheringham, 1989; Marceau *et al.*, 1994; Amrhein and Reynolds, 1996; Arbia *et al.*, 1996; Goodchild & Quattrochi, 1997).

The *MAUP* is also associated with the concept of ecological fallacy. Ecological fallacy occurs when it is inferred that results based on aggregate zonal data can be applied to the individuals or specific sites (x, y coordinate locations) within the zone itself. The problem occurs because studies using areal data do not distinguish between spatial associations created by the aggregation of data and real associations possessed by the individual data prior to spatial aggregation (Openshaw, 1984). Any statistics or models, which are based on aggregated spatial datasets, may be valid at the scale of the dataset, but any attempts to infer to higher-resolution or lower-resolution data (such as the use of vegetation indices derived from remotely sensed imagery derived at a 30m scale being used to assess regional patterns in vegetation) may produce invalid results. The statistics and model parameters differ between the two levels of resolution, and we have no way to predict what they are at the higher level given the values at the lower level or *vice versa*. This is true for both the spatial and temporal scale. If ecological or social policies are based on such conclusions, there could be unforeseeable consequences.

Studies of the *MAUP* date back to the 1930s with the emphasis greatest in the late 1960s and 1970s. The results from studies of the *MAUP* have been highly variable and somewhat incomplete, thus making it difficult to make broad inferences about how the *MAUP* influences the performance of univariate, bivariate, and multivariate statistics.

However, some general patterns have arisen. In univariate statistics, when the *MAUP* is present the mean does not change and the variance declines with increasing aggregation (Gehlke and Biehl, 1934; Openshaw, 1984; Fotheringham and Wong, 1991). Essentially, there is a loss of information associated with a smoothing effect that occurs upon aggregation. This phenomenon has been recognized by most researchers (Gehlke and Biehl, 1934; Openshaw 1984; Fotheringham and Wong, 1991; Jelinski and Wu, 1996). Zoning effects have less predictable results for the mean and variance. Jelinski and Wu (1996) demonstrate the contrived affects of both scale and zonation in Figures 1 (a-c). In these figures the mean value does not change with aggregation, but the variance declines. In Figures 1 (d-f) the units have been aggregated into zones with varying orientations of the cardinal directions. For d and e there is no change in the mean, but the variance changes substantially. By comparing d-f one can see that even when the number of zones is held constant the mean and the variance are affected (Jelinski and Wu, 1996).

In the natural sciences, research has focused on the issue of scale and not aggregation. One of the major contributions in the field of natural sciences was to acknowledge the existence of natural scales at which ecological processes and physical characteristics occur within the landscape. This was revealed by a series of studies oriented toward the choice of an appropriate sampling unit size for analysing ecological phenomena, particularly to detect spatial patterns in plant communities (Kershaw, 1957; Mead, 1974; O'Neill *et al.*, 1986). Research suggested that because the scale of the study determines the range of patterns and processes that can be detected, an appropriate level of resolution for study of these processes should be identified. Because ecological and physical processes operate at different spatial scales, the need for appropriate scaling laws has been emphasized in current research in order to relate information across a wide range of scales.



Figures 1 (a-c) In these figures the mean value does not change with aggregation, but the variance declines. In Figures 1 (d-f) the units have been aggregated into zones with varying orientations of the cardinal directions. For d and e there is no change in the mean, but the variance changes substantially. By comparing d-f one can see that even when the number of zones is held constant the mean and the variance are affected (Jelinski and Wu, 1996)

Wiens (1989) and Levin (1993) both argue that a variety of statistical and mathematical tools, such as correlation and extrapolation, can be used for scaling. However, they concluded that these techniques are appropriate only when applied for short-term or small-scale predictions or, in other words, within the relevant domain of scale for the phenomenon under investigation. Extension across scale thresholds may be hazardous due to the instability of the dynamics of the transition zone between two domains of scale.

While studies relating to the issue of scale continue to be prolific in biogeography and other sub-disciplines within physical geography, there has been little concern about the issue of aggregation. With the increased use of satellite imagery (which is inherently aggregated and further aggregated in order to create *GIS* data layers) concerns about aggregation should become increasingly important.

III Remote sensing applications

Remotely sensed imagery has been used for the development of large-scale ecosystem models, forest resource assessment, analysis of ecosystem health, vegetation mapping, and assessment of large-scale biogeographic patterns for a variety of taxa (Gillespie, 2001; Nightingale *et al.*, 2004; Boyd and Danson, 2005). While the use of remotely sensed imagery has clearly increased our understanding of the spatial components of the natural environment, it is important to recognize the limitations of this data (Hay *et al.*, 2001).

Marceau (1992) was among the first to recognize the relationship between the *MAUP* and remotely sensed imagery. One key limitation of remotely sensed data is the issue of spatial scale. Remotely sensed data is collected at a predefined scale without regard to the specific phenomena that may be examined with such data. The issue of scale has been examined by scientists in different disciplines and it has been widely recognized that the complexity of natural systems requires a multiscale approach (Hay *et al.*, 2001).

As stated by Hay *et al.* (2001), 'remotely sensed imagery represents a largely unrecognized case of the *MAUP*.' The resolution of a raster data set corresponds to the scale effect of the *MAUP*. The choice of the scale for analysis is predetermined with satellite imagery by the resolution of the data (Jelinski and Wu, 1996). Remotely sensed imagery essentially represents an arbitrary sampling grid superimposed over the surface of the earth. The areal unit used for sampling is determined by the mechanics of the satellite rather than an ecologically or scientifically significant scale. Therefore, remotely sensed data is largely prone to the problems associated with the *MAUP* because the real world data is aggregated based on the size of the grid being superimposed and not underlying natural processes. Likewise, remotely sensed imagery is often re-sampled in a process where neighbourhood raster cell values are averaged or mathematically combined to smooth or filter data (Lillisand *et al.*, 2004). This re-sampling corresponds to the aggregation effect which can result in erroneous data and results when extracted for statistical analyses.

Several studies assessing the impact of the *MAUP* on remotely sensed data have found that this type of data induces the scale and aggregation effects that define the *MAUP*. Marceau *et al.* (1994) conducted an empirical investigation to verify the impact of spatial resolution and aggregation level on the classification accuracy of forest data. Their results indicated that per-class accuracies were considerably affected by changing scale and aggregation level, which led to the conclusion that remotely sensed data are not independent of the *MAUP*. In addition, Arbia *et al.* (1996) confirmed the unpredictable effects of the *MAUP* on the accuracy of maximum-likelihood image classification. Still, there is little discussion or research with regard to the *MAUP* and its impact on large-scale remotely sensed biogeographic data such as measurements of sea ice, spectral reflectance, and sea-surface temperatures.

IV *GIS* applications

A geographic information system (*GIS*) is a powerful tool that is increasingly utilized to perform spatial analysis in physical geography as well as many other disciplines. Researchers can use *GIS* to perform hydrologic modelling, climate modelling, predictive mapping for plant function types, land use evaluation modelling, digital terrain analysis modelling, and more (Clarke *et al.*, 2002). Whereas the scale and aggregation standards of many projects are often based on remotely sensed data, a *GIS* may also rely on data collected from field surveying procedures, and manual delineation of areal features from pre-existing analog maps or aerial photography. Although *GIS* can provide incredibly robust and extensive information to geographers in a highly efficient manner, it is not by any means immune to the problematic influences of the *MAUP*.

Issues related to hydrologic modelling provide an excellent example of how the *MAUP* can be manifested in physical geography through the application of *GIS*. Hydrological datasets such as flow accumulation, flow direction, and watershed boundaries (Hellweger and Maidment, 1999) are most commonly derived from the processing of raster Digital Elevation Models with a *GIS* (Skidmore, 2002). More current DEMs may be produced from high resolution remotely sensed imagery (Skidmore, 2002) and earlier versions with 30m elevation postings were generated from either low resolution aerial imagery or the digitization of topographic maps (Clarke *et al.*, 2002). In either case, the topography of the landscape was transferred to a raster grid by superimposing a pre-defined spatial grid on the landscape. Models then generated from this data are likewise subjected to the *MAUP*.

Flow accumulation models provide an explicit example of the *MAUP*. These models estimate the accumulation of water at a given location based on a statistical calculation derived from a DEM. Flow accumulation models are then used to extract linear features such as a stream drainage network within a watershed (Maidment and Djokic, 2000). While providing an appropriate general perspective on the flow of water across the landscape, the variation in flow accumulation values is dependent upon the scale of the DEM from which it is derived. Likewise, real world data are aggregated to the sampling scale of the DEM. The extraction of a flow line (stream feature) from the accumulation grid further reduces the fine scale variation in the accumulation of water across the landscape and results in an inaccurate spatial placement of the flow line. Figure 2 demonstrates this phenomenon where a 30m resolution DEM is used to extract a flow line using the described process, and is then compared to a medium resolution (10m) DEM and a high resolution (1m) true colour image.

Slope calculations may also suffer from the consequences of the *MAUP*. Slope values are derived from DEMs and are used for a variety of purposes such as predicting erosion, stream nutrient loads, and evapotranspiration rates (Evans *et al.*, 2002; Lu *et al.*, 2003). Slope models are increasingly being used in a *GIS* to identify riparian zones for land management protection. The goal is to

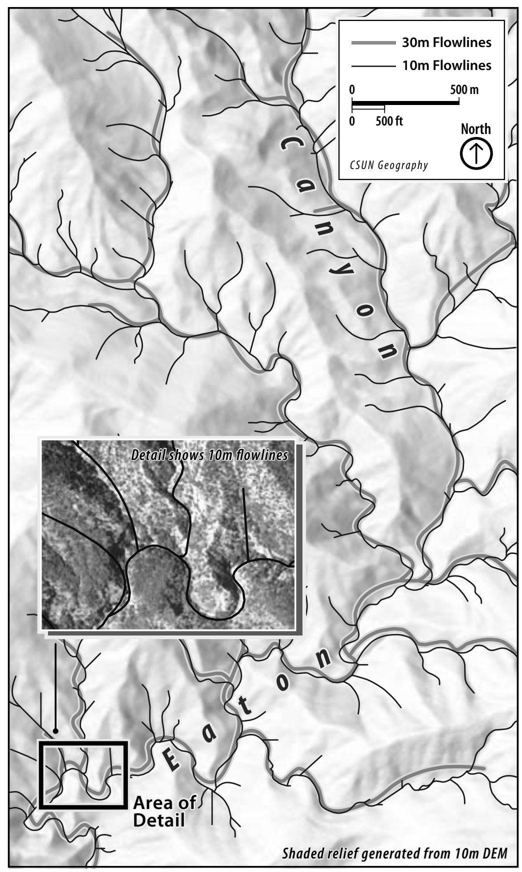


Figure 2 Flow line models extracted from a 30 metre resolution DEM and a 10 metre resolution DEM. The inset shows both models overlaid onto a 1 metre resolution true colour image

identify the area in which hillslope processes contribute to sediment accumulation (Schoorl *et al.*, 2000; Wilson *et al.*, 2001). These areas are then mapped and classified as riparian and consequently given importance in maintaining the overall health of the larger watershed. To determine the contribution of hillslope processes to riparian areas, slope values are averaged across a buffer distance generated from a stream flow line (Naiman and Decamps, 1997). These buffers, represented as polygons, are then used to map the riparian zones to be protected. In this case, values are summarized both from the DEM and across the buffer distance, potentially resulting in a loss of information. This loss of information may be manifested in a lack of inclusion of hillslope processes that are indeed important aspects of riparian areas.

Project design and the data collection associated with field-based vegetation mapping provide another example of how the *MAUP* is manifested in *GIS*. In cases where landcover/vegetation layers are generated using data collected in the field, it is necessary to simplify and aggregate vegetation types due to time and resource limitations in data collection. Because of this, superior sampling techniques are critical to ensure an accurate and comprehensive collection of information is obtained. As a result, it is inevitable that subtle variations in vegetation or unique occurrences of rare species may be ignored. The *GIS* layer ultimately created and the conclusions drawn as a result of the simplification of this data could be incomplete and erroneous. A potential solution to this is to increase the amount of field data available. Murguia and Villasenor (2000) proposed methods to assess the error associated with data in presence–absence matrices. Their quantitative analysis illustrates that since the quality of data included in the study is directly related to the level of resolution and attention to detail of the biogeographical analysis, the degree of error diminishes as data availability increases. However, the feasibility of collecting sufficient field data is often limited by logistics and funding.

The *MAUP* can also result from the derivation of spatial data and/or its associated attribute values. One manner in which this can happen is the conversion of vector data (ie, a point, line or polygon) into a raster format. The information and values that may originally be linked to a discrete point location can be resampled and associated with a raster cell that represents an area covering up to 30m (Skidmore, 2002). Additionally, merging either adjacent polygons with the same value or distributed polygons with similar values that are simplified (Skidmore, 2002) could result in effects associated with the *MAUP*. The development of complex spatial models has become a mainstay for *GIS* (Goodchild and Quattrochi, 1997). However, the representation of the individual entities being modeled has rarely been addressed. As data are aggregated, this lack of knowledge can lead to problems associated with the *MAUP* and ecological fallacy.

V Some potential solutions

While considered largely intractable, there are some potential solutions to the *MAUP* (Openshaw, 1984). These are discussed extensively in several papers (Openshaw, 1977, 1984; Tobler, 1979; Fotheringham, 1989; Jelinski and Wu, 1996; Marceau and Hay, 1999). Even with the numerous options suggested, none provide a comprehensive solution that is capable of easily and accurately quantifying the effects of the *MAUP*. We will briefly review some of these approaches.

Openshaw (1984) proposed four initial solutions which were used as base models and expanded upon by other geographers within the discipline of human geography. The first solution simply suggests that one could ignore the problem of the *MAUP* and hope that the outcome of the research was still significant. Although this is the easy way out of the problem, it has one significant flaw. Results using this technique may appear to be significant, but because of the *MAUP* these results may not be accurate. Being unaware of the degree to which the conclusions have been affected by the *MAUP* may have serious consequences for analyses impacted by this phenomenon. This solution is almost certainly used more often than others due to the fact that calculating the effects of the *MAUP* is such a challenging process.

The second solution proposed by Openshaw (1984) acknowledges but diminishes the importance of the *MAUP* by examining and addressing the significance of spatial entities. Openshaw (1984) suggests the *MAUP* exists due to the ambiguity of which spatial entities are being examined. Therefore, if geographers agree to the 'objects of geographical enquiry' (Openshaw, 1984) or analysis is performed only with basic entities (Hay *et al.*, 2001), a resolution may exist for this predicament. In other words, a focus on identifying the appropriate spatial scale of analysis may limit the impacts of the *MAUP*. However, identification of the appropriate scale for spatial analysis continues to be a complicated and unresolved issue for natural phenomena studied within the discipline of physical geography.

Hay *et al.* (2001) are among the first to offer an approach for 'analysing and upscaling remotely sensed data' by applying an 'object-specific approach' (OSA). This approach represents a multiscale technique that defines unique spatial measures of objects composing a remotely sensed image. These spatial measures are then used as weighting functions for upscaling an image to a more coarse resolution. This process reduces the effect of the *MAUP* by incorporating object-specific measures throughout the analysis of upscaled data. Even with the application of this solution, two problems remain. First, the 'objects' in question will be different based on what is being studied and the level of functional geographic knowledge of the researcher. Second, although the aggregation effect will be removed, other remnants of the *MAUP* will remain.

Jelinski and Wu (1996) propose another potential solution to the *MAUP* by identifying and analysing all individual entities in a project in order to avoid the *MAUP* entirely. While this would undoubtedly provide the most accurate results, it is entirely unrealistic considering the incredible time and resource demands. Additionally, Jelinski and Wu (1996) recommend building on this concept with the 'sensitivity analysis approach'. This solution aims at obtaining a sense of project scope and magnitude through detailed assessments, and then determines how sensitive the different variables are to variations in scale and zoning configurations. This approach is akin to the recommendations of Fotheringham (1989), in which the variables and relationships that become unpredictable due to changes in scale are examined in detail. Methods such as fractal dimension analysis and spatial autocorrelation can be used to identify these characteristic scales so geographers could gain greater insight into the instability of their data and focus their studies accordingly. Although comprehensive and intensive solutions such as these are beneficial due to the illumination of the project data and design, they are simultaneously impractical for projects which contain large amounts of data and numerous variables.

Another solution proposes a new methodology which defies the normal science paradigm by including a hypothesis in the data set-up design of the spatial analysis project. A hypothesis is created based on the expected result for an analysis, and aerial units are aggregated to the point where the target result is attained. In this case, spatial units are defined with awareness of the entity and particular analysis and they are also geographically meaningful. While contrary to the normal science paradigm, Openshaw (1984) makes the point that since the initial unit design of a project is subjective and adjustable anyway, why not mould the project to produce a predicted outcome? According to Openshaw, if the 'statistical assumptions or geographical factors are violated, the hypothesis must be rejected'. In this case, the theory would be confirmed and the *MAUP* avoided.

A variety of other options have been discussed (Fotheringham, 1989; Tobler, 1979; Jelinski and Wu, 1996; Marceau and Hay, 1999). However, which solution is used is based on the project topic, type of analysis, and degree to which the *MAUP* affects the results. Certain solutions may not be logistically feasible or are inappropriate for most types of analyses. As geographers become more aware of the *MAUP* and include the appropriate analyses in their projects, potential solutions may either become more viable or new ones suggested.

VI Conclusion

The purpose of this paper is to present a review of the *MAUP* and its potential implications in physical geography. The problems associated with the *MAUP* were identified by social scientists and economic geographers more than 40 years ago (Gehlke and Biehl, 1934; Yule and Kendall, 1950; Blalock, 1964; Openshaw, 1977). However, since then the topic has been given little attention, particularly by physical geographers. The recent emergence of the use of remotely sensed imagery in physical geography, has renewed some interest in the *MAUP*, resulting in studies that focus on the development of new scaling techniques (Jelinski and Wu, 1996; Cain *et al.*, 1997; Marceau and Hay, 1999). Nonetheless, the *MAUP* is still considered an unresolved problem in spatial analysis and a major conceptual challenge for any geographic study (Openshaw, 1984). As such, physical geographers working with remotely sensed data and *GIS* should continue to be aware of this issue and attempt to address it when possible.

References

Amrhein, C.G. and **Reynolds, H.** 1996: Using spatial statistics to assess aggregation effects. *Geographical Systems* 3, 143–58.

Arbia, G., Benedetti, R. and **Espa, G.** 1996: Effects of the *MAUP* on image classification. *Geographical Systems* 3,123–41.

Blalock, H. 1964: *Causal inferences in non-experimental research*. Chapel Hill, NC: University of North Carolina Press.

Boyd, D.S. and **Danson, F.M.** 2005: Satellite remote sensing of forest resources: three decades of research development. *Progress in Physical Geography* 29(1), 1–26.

Burke, I.C., Kittel, T.G.F., Lauenroth, W.K., Snook, P., Yonker, C.M. and **Parton, W.J.** 1991: Regional analysis of biomass, primary productivity, and net CO_2 flux. *Vegetation* 80, 71–89.

Cain, D.H., Riitters, K. and **Orvis, K.** 1997: A multi-scale analysis of landscape statistics. *Landscape Ecology* 12, 199–212.

Clarke, K.C., Parks, B.O. and Crane, M.P. 2002: *Geographic information systems and environmental modeling*. Upper Saddle River, NJ: Prentice-Hall, Inc.

Davis, F.W., Stoms, D.M., Hollander, A.D., Thomas, K.A., Stine, P.A., Odion, D., Borchert, M.I., Thorne, J.H., Gray, M.V., Walker, R.E., Warner, K. and **Graae. J.** 1998: *The California Gap Analysis Project – Final Report*. Santa Barbara, CA: University of California Press.

Evans, B.M., Lehning, D.W., Corradini, K.J., Petersen, G.W., Nizeyiman, E., Hamlett, J.M., Robillard, P.D. and **Day, R.L.** 2002: A comprehensive GIS-based modeling approach for predicting nutrient loads in watersheds. *Spatial Hydrology* 2, 1–18.

Fotheringham, A.S. 1989: Scale-independent spatial analysis. In Goodchild, M.F. and Gopal, S., editors, *Accuracy of Spatial Databases*, London: Taylor and Francis, 221–8.

Fotheringham, A.S. and **Wong, D.W.** 1991: *Statistical geography: problems in analyzing areal data*. Glencoe, IL: The Free Press.

Gehlke, C.E. and **Biehl, H.** 1934: Certain effects of grouping upon the size of the correlation coefficient in census tract material. *Journal of the American Statistical Association Supplement* 29, 169–70.

Gillespie, T.W. 2001: Remote sensing of animals. *Progress in Physical Geography* 25, 355–62.

Goodchild, M.F. and **Quattrochi, D.A.** 1997: *Scale, multiscaling, remote sensing, and GIS*. New York: Lewis Publishers.

Haverkamp, S., Fohrer, N. and **Frede, H.G.** 2005: Assessment of the effect of land use patterns on hydrologic landscape functions: a comprehensive GIS-based tool to minimize model uncertainty resulting from spatial aggregation. *Hydrological Processes* 19, 715–27.

Hay G.J., Marceau, D.J., Dube, P. and **Bouchard, A.** 2001: A multiscale framework for landscape analysis: object-specific analysis and upscaling. *Landscape Ecology* 16, 471–90.

Hellweger, F.L. and Maidment, D.R. 1999: Definition and connection of hydrologic elements using geographic data. *Journal of Hydrologic Engineering* 4, 10–18.

Jelinski, D.E. and **Wu, J.** 1996: The modifiable areal unit problem and implications for landscape ecology. *Landscape Ecology* 11, 129–40.

Jelinski, D.E., Goodchild, M.F. and **Steyaert, L.** 1994: Multiple roles for GIS in global change research: towards a research agenda. In Michener, W.K., Brunt, J.W. and Stafford, S.G., editors, *Environmental Information Management and Analysis: Ecosystem to Global Scales*, London: Taylor and Francis, 41–58.

Kershaw, K.A. 1957: The use of cover and frequency in the detection of pattern in plant communities. *Ecology* 38, 291–9.

Levin, S.A. 1993: Concepts of scale at the local level. In Ehleringer, J.R. and Field, C.B., editors, *Scaling Physiological Processes: Leaf to Globe*, San Francisco, CA: Academic Press, 7–19.

Lillisand, T.M., Keifer, R.W. and **Chipman, J.W.** 2004: *Remote sensing and image interpretation*. Hoboken, NJ: John Wiley and Sons.

Lu, J.G., Sun, G., McNulty, S.G. and **Amaytya, D.M.** 2003: Modeling actual evapotranspiration from forested watershed across the southeastern United States. *Journal of the American Water Resources Association* 39, 886–96.

Maidment, D. and **Djokic, D.** 2000: *Hydrologic and hydraulic modeling support with geographic information systems*. Redlands, CA: ESRI Press.

Marceau, D.J. 1992: *The problem of scale and spatial aggregation in remote sensing: an empirical investigation using forestry data*. Unpublished PhD thesis. Ontario, Canada: Department of Geography, University of Waterloo, 180.

Marceau, D.J. and **Hay, G.J.** 1999: Remote sensing contributions to the scale issue. *Canadian Journal of Remote Sensing* 25, 357–66.

Marceau D.J., Howarth, P.J. and **Gratton, D.J.** 1994: Remote sensing and the measurement of geographical entities in a forested environment. Part I: The scale and spatial aggregation problem. *Remote Sensing and Environment* 49, 93–104.

McDermid, G.J., Franklin, S.E. and **LeDrew, E.F.** 2005: Remote sensing for large-area habitat mapping. *Progress in Physical Geography* 29, 449–74.

Mead, R. 1974: A test for spatial pattern at several scales using data from a grid of contiguous quadrats. *Biometrics* 30, 295–307.

Murguia, M. and **Villasenor, J.L.** 2000: Estimating the quality of the records used in quantitative biogeography with presence–absence matrices. *Ann. Bot. Fennici* 37, 289–96.

Naiman, R.J. and **Decamps, H.** 1997: The ecology of interfaces: riparian zones. *Annual Review of Ecology and Systematics* 28, 621–58.

Nightingale, J.M., Phinn, S.R. and **Held, A.A.** 2004: Ecosystem process models at multiple scales for mapping tropical forest productivity. *Progress in Physical Geography* 28, 241–81.

Openshaw, S. 1977: A geographical solution to scale and aggregation problems in region-building, partitioning and spatial modelling. *Institute of British Geographers, Transactions, New Series* 2, 459–72.

Openshaw, S. 1984: Ecological fallacies and the analysis of areal census data. *Environment and Planning A* 16, 17–31.

Openshaw, S. and **Taylor, P.J.** 1979: A million or so correlation coefficients: three experiments on the modifiable areal unit problem. In Wrigley, N., editor, *Statistical applications in spatial sciences*, London: Pion, 127–44.

O'Neill, R.V., DeAngelis, D.L., Waide, J.B. and **Allen, T.F.H.** 1986: *A hierarchical concept of ecosystems*. Princeton, NJ: Princeton University Press.

Rosswell, T., Woodmansee, R.G. and **Matson, P.A.** 1991: What does remote sensing do for ecology? *Ecology* 72, 45–54.

Schoorl, J.M., Sonneveld, P.W. and **Veldkamp, A.** 2000: Three-dimensional landscape process modelling: the effect of DEM resolution. *Earth Surface Processes and Landforms* 25, 1025–34.

Skidmore, A. 2002: *Environmental modelling with GIS and remote sensing*. London: Taylor and Francis.

Tobler, W. 1979: Cellular geography. In Gale, S. and Olsson, G., editors, *Philosophy in Geography*, Dordrecht: Reidel, 379–86.

Wiens, J.A. 1989: Spatial scaling in ecology. *Functional Ecology* 3, 385–97.

Wilson, C.J., Cartey, J.W., Beeson, P.C., Gard, M.O. and **Lane, L.J.** 2001: A GIS-based hillslope erosion and sediment delivery model and its application in the Cerro Grande burn area. *Hydrological Processes* 15, 2995–3010.

Yule, G. and **Kendall, M.** 1950: *An introduction to the theory of statistics*. New York: Charles Griffin and Company Limited.