Environ Monit Assess (2009) 149:71–80
DOI 10.1007/s10661-008-0184-y

Use of space-filling curves to select sample locations in natural resource monitoring studies

Andrew J. Lister · Charles T. Scott

Received: 10 July 2007 / Accepted: 14 January 2008 / Published online: 12 February 2008

Abstract The establishment of several large area monitoring networks over the past few decades has led to increased research into ways to spatially balance sample locations across the landscape. Many of these methods are well documented and have been used in the past with great success. In this paper, we present a method using geographic information systems (GIS) and fractals to create a sampling frame, superimpose a tessellation and draw a sample. We present a case study that illustrates the technique and compares results to those from other methods using data from Voyageurs National Park in Minnesota. Our method compares favorably with results from a popular plot selection method, Generalized Random Tessellation Stratified Design, and offers several additional advantages, including ease of implementation, intuitive appeal, and the ability to maintain spatial balance by adding new plots in the event of an inaccessible plot encountered in the field.

Keywords GIS · Monitoring · Sample grid · Sample selection · Sampling frame · Space-filling curve

A. J. Lister (✉) · C. T. Scott
National Inventory and Monitoring Applications Center,
Northern Research Station Forest Inventory
and Analysis Unit,
11 Campus Blvd, Ste. 200,
Newtown Square, PA 19073, USA
e-mail: alister@fs.fed.us

C. T. Scott
e-mail: ctscott@fs.fed.us

Introduction

The design of networks of environmental monitoring sites has taken on a great deal of importance over the past several decades. The U.S. Environmental Protection Agency (EPA) has been particularly active in this area due to its legislative requirement to monitor compliance with the Clean Water Act in a scientifically defensible way (McDonald et al. 2002). EPA's Environmental Monitoring and Assessment Program (EMAP) was developed to determine the condition of and detect trends in the nation's ecosystems using a statistically valid monitoring framework (Palmer et al. 1992). Other regional and national scale monitoring networks have been developed with similar goals, including those of the U.S. Forest Service's Forest Inventory and Analysis (FIA) Program (U.S. Forest Service 1992; Gillespie 1999) and the USDA National Agricultural and Statistics Service's (NASS) Census of Agriculture (Cotter and Nealon 1987). A feature common to all of these monitoring networks is the distribution of sampling locations across the landscape in a manner that allows for the generation of statistically valid estimates of the attribute of interest.

In addition to the common requirements of statistical independence and unbiasedness (Cochran 1977), interpretability of an estimate of an environmental factor and its variance also is predicated on the distribution of the sample in space. For example, a single realization of a simple random spatial sample might not lead to a well distributed sampling network.

If the goal of the sampling network is to characterize the study area, then an approach that forces the samples to be distributed uniformly across the study area might be more appropriate. Many sampling designs have been created to force the even distribution of plots across the spatial domain of interest, most famously systematic tessellation (Shiver and Borders 1996; Olsen et al. 1998; Olea 1984). For example, FIA uses a regular hexagonal tessellation to distribute plots uniformly across the landscape with a degree of randomization inserted to avoid bias problems that co-occurring spatial periodicity of the environment and the sampling frame might introduce (Reams et al. 2005).

One disadvantage of using regular tessellations to distribute samples across the landscape is that partial cells (polygons) are created by study area boundaries. A decision rule must be constructed in this case—if one of the cells of a tessellation is split, how is it decided if a sample is located within the partial cell? If, for example, the position of the centroid of the polygon is chosen as a factor that decides if the partial cell is to be populated by a sample, then portions of cells within the study area that have their centroid outside the area have no chance of being measured. This unequal probability of selection for population elements found on the edges of study areas can bias estimates of population totals (Gregoire and Scott 2003).

To avoid problems with systematic or semi-systematic sampling networks, several alternative sample selection methods have been proposed (Olsen et al. 1999; Saalfeld 1998; Stevens 1997; Cotter and Nealon 1987; Olea 1984). One approach described by Stevens and Olson (1999, 2004) is to divide the study area recursively into quadrants, attach an ordered spatial address to each quadrant, and reorganize the spatial addresses such that samples can be randomly selected from them in a spatially uniform way. This procedure, known as Generalized Random Tessellation Stratified Design (GRTS), has several desirable properties, including applicability of standard design-based estimation procedures, spatial balance, ability to assign differential selection probabilities to elements in different areas, ability to create panels, and flexibility to add new points in a statistically valid way in the event of sampling frame imperfections or the occurrence of inaccessible plots. GRTS or GRTS-like methods have been used in the past (e.g., Henderson et al. 2005) and a variance estimation procedure for GRTS samples has been documented (Stevens and Olsen 2003). In addition, there are several software-based methods for generating a GRTS-like sample (West, Inc. 2006; Kincaid 2006; Theobald et al. 2006).

A disadvantage of GRTS, however, is that the algorithms behind its execution tend to obscure the process by which samples are selected, i.e., the process tends to be a black box. Notwithstanding its published documentation, it is difficult for the practitioner to visualize the complete sample selection process—outputs of the existing software packages tend to be ordered lists of coordinates. Furthermore, when adding new points to a survey due to frame imperfections, denied access, or dangerous conditions, current GRTS-based implementations tend to add them in a spatially disjoint manner due to the randomization procedure used (Stevens and Olsen 1999, 2004; Theobald et al. 2006). To maintain the greatest amount of integrity in the spatial balance of the sampling frame and to minimize travel time of field workers, it would be more desirable to randomly choose a replacement sample from a set of pre-chosen options located near the original (although this slightly alters selection probabilities, possibly affecting decisions to use classical statistical estimators). Finally, conceptual clarity and ease of implementation are important to many practitioners, but the theory behind GRTS is difficult to describe to the lay person.

GRTS-based approaches come from the concept of space-filling curves (SFCs). An SFC is a curve or shape that completely occupies an area of interest (Bartholdi and Platzman 1988). A specific type of SFC is a Peano curve fractal (Peano 1890; Mandelbrot 1982), which is a type of repeating, self-similar shape that, if repeatedly recursed, fills a planar surface (Fig. 1). In effect, each point on the Peano curve creates a one-dimensional spatial reference for each point in two dimensions, thus representing locations in two dimensional space as locations along a line. This partitioning of space creates a de facto tessellation that can divide the sampling frame into an infinite number of regions of equal size.

Peano curves and other SFCs have been used in the past to condense multidimensional problems into one dimension. For example, Jin and Mellor-Crummey (2005) showed how exploitation of SFCs in the design of computing systems can increase computational efficiency. Saalfeld et al. (1992) recommended using SFCs to condense multidimensional survey data to one dimension for nearest neighbor calculations for the U.S. Census Bureau. In a geographic context,

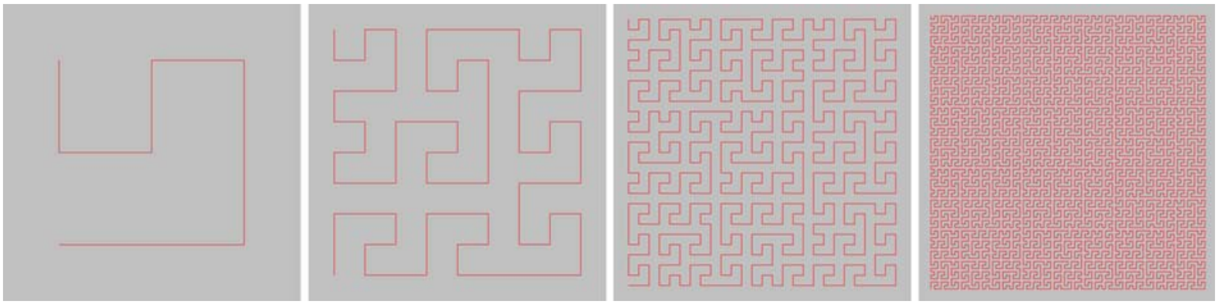

Fig. 1 An example of one type of Peano curve. Four levels of recursion (left–right) are shown, demonstrating its self-similar, space-filling properties. The intensity of recursion affects the density of the curve within the sampling frame. Figures generated by a java applet constructed by V. B. Balayoghan, retrieved 12/1/2006, from http://www.cs.utexas.edu/users/vbb/misc/sfc/index.html

Saalfeld (1998) gave an excellent overview of the use of SFCs and other geographic ordering schemes in sampling. Johnson et al. (1993) described using Peano curves to order the primary sampling units used to derive the U.S. Bureau of Labor Statistics' Consumer Price Index. Lam and Liu (1996) used Peano-ordering to create spatially contiguous groupings of counties for a medical study. Finally, Stevens and Olsen (1999, 2004) and Theobald et al. (2006) used an SFC-based approach for GRTS.

In the current study, we build on this work by presenting an SFC-based approach to translating the spatial location of each element in a two-dimensional sampling frame to a one-dimensional spatial address, grouping these addresses into contiguous or semi-contiguous groups of potential sample locations, and randomly selecting samples from within these groups. This approach offers most of the benefits that the previously described methods offer and has the added advantage of relative transparency and intuitive clarity. Furthermore, it can easily be implemented with the native functionality of a commonly used geographic information system (GIS). Sample networks were constructed for a test area using simple random spatial sampling, GRTS and our method, and the methods were compared in various ways.

Methods

The study area is in and around Voyageurs National Park in northern Minnesota (Fig. 2). The study area

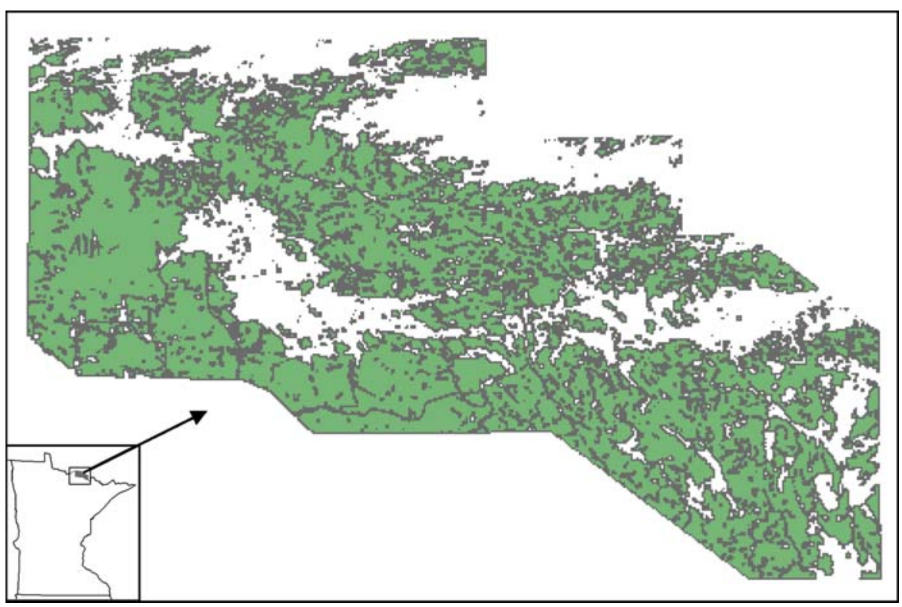

Fig. 2 The Voyageurs National Park, MN, USA

has roughly 95,000 ha, is mostly forested, and has numerous lakes. For this study, the intent was to construct a spatially balanced sampling network where each sample has an equal probability of selection, i.e., inclusion probability.

To create the sampling frame, the area was first discretized into elements (30-m pixels) by converting a vector GIS coverage of the land in the study area to a raster using ESRI's ArcINFO GIS software. After creating the raster sampling frame, ArcINFO again was used to convert the raster to a set of points, where each point occupies the center of what was each pixel. This discretization process creates the set of possible samples, spaced at 30 m, from which a subset was drawn. The choice of resolution for the rasterization of the study area was based on a tradeoff between computational efficiency, GIS file size, and a desire to establish a nearly continuous set of elements across the sampling frame from which to draw samples. Generally, the finest possible resolution that can be processed in an efficient manner should be chosen. In the case of our study, 1,064,467 potential sample locations were generated.

Next, ArcINFO's "spatialorder" function was used to construct what amounts to a Peano curve passing through the set of points. The ArcINFO algorithm creates a spatial address index ranging from 0–1 (Fig. 3). By clustering these spatial address indices into n groups of (mostly) contiguous values, n clusters of elements can be created, from which a subset of n samples can be drawn. For this study, a spatially balanced sample of 139 forested plots was chosen because of previous studies that have used this number in the park, so 139 clusters of 7,658 elements (1,064,467/139=7,658) were constructed using the ArcINFO "collocate" function. From each of these clusters, one sample was randomly chosen (Fig. 4). To remove any periodicity introduced by the 30-m point spacing, each of the chosen points was located randomly within the 30×30 m square centered on the point.

To compare the degree of spatial balance achieved by this procedure to that from other methods, the above steps were first performed 1,000 times. Next, the r-GRTS implementation of the GRTS procedure (Kincaid 2006), with an equal probability sample

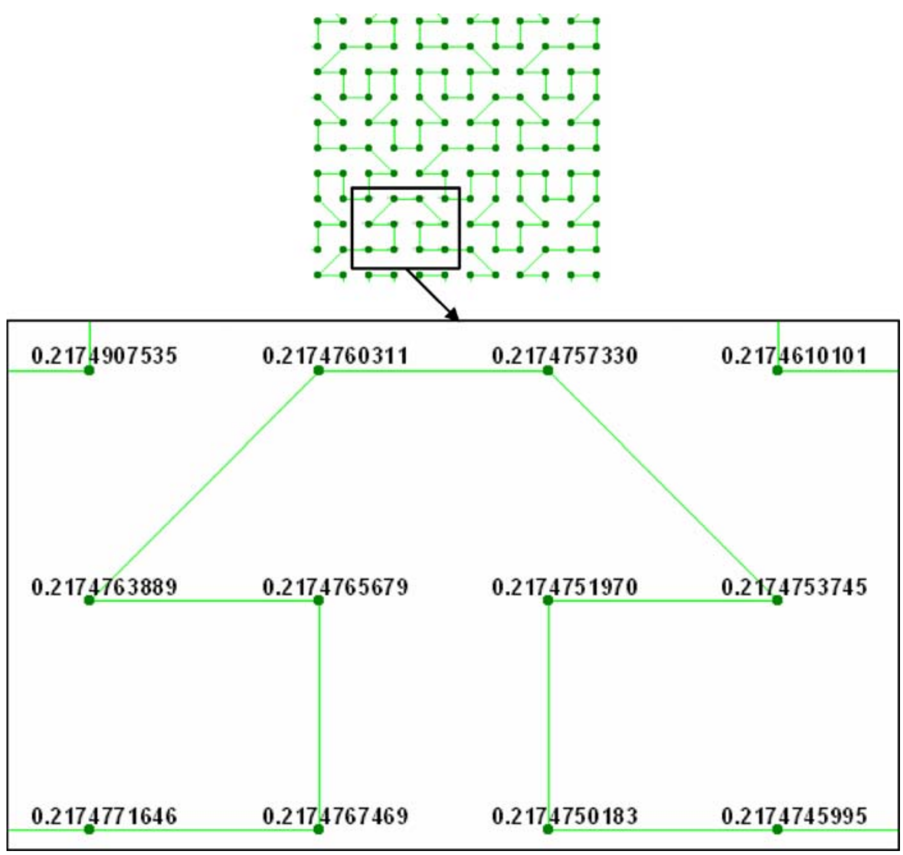

Fig. 3 Subset of the Peano curve drawn through the elements of the sampling frame. A continuous line is drawn through the study area. Each vertex of this Peano curve is given a spatial address that describes its position on the line. By grouping these spatial addresses successively, relatively contiguous groupings of elements from which samples can be drawn are assembled

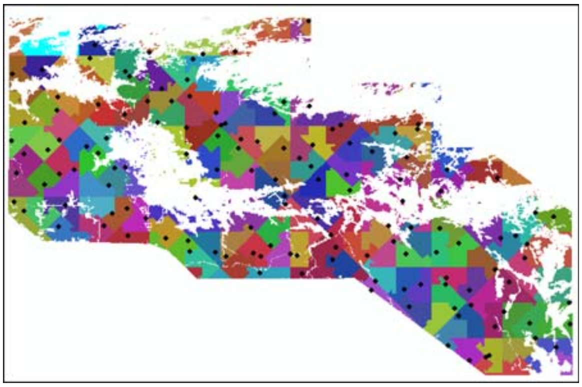

Fig. 4 Each element in the sampling frame is grouped into a class based on its spatial address assigned by the Peano curve. One sample is drawn at random from each class to create the final sample network

chosen as the sampling method, was used to create 1,000 sets of 139 points, as was a simple random sampling procedure (SIM) implemented in ESRI's ArcMap GIS. The degree of spatial balance achieved by the three methods was determined by comparing frequency distributions of variances of the areas of the rasterized Thiessen polygons surrounding the plots. A Thiessen polygon is a polygonal area surrounding each point, within which every location is closer to that point than to any other point. This method is analogous to the comparison conducted by Stevens and Olsen (2004) in which the variance of a set of Voronoi polygons was used as a spatial balance index. For each realization of the simulation, each pixel in a rasterized map of the study area was assigned the identifier of the sample plot closest to it. By labeling all of the pixels in this manner, contiguous areas of pixels labeled with the same plot identifier were created. These areas can be thought of as the spatial domain represented by each sample. The variance of the areas of these spatial domains was calculated, creating a value which is directly related to the degree of spatial balance of the plots (Stevens and Olsen 2004). Larger variance values indicate clumping of plots, whereas smaller values indicate equal spacing.

To assess each method with respect to its ability to precisely estimate the true areas of land-cover classes, we used ArcINFO GIS to sample a 66 class land cover map of the Voyageurs Park (U.S. Geologic Survey 2001) with each of the 1,000 realizations of each of the three methods. For each method's set of 1,000 realizations, we calculated the coefficient of variation (CV) for each land cover class as an index of precision. For each class and each realization, we also calculated exact binomial confidence intervals around estimates of class proportions. We then determined for each class and each method the number of times out of 1,000 that the confidence intervals contained the true land-cover class proportion.

Finally, to address the concern we had that areas close to the study area boundaries might have an anomalous sampling intensity due to an artifact in either the GRTS or the SFC methods, we computed the average density of plots within 40 five-meter-wide

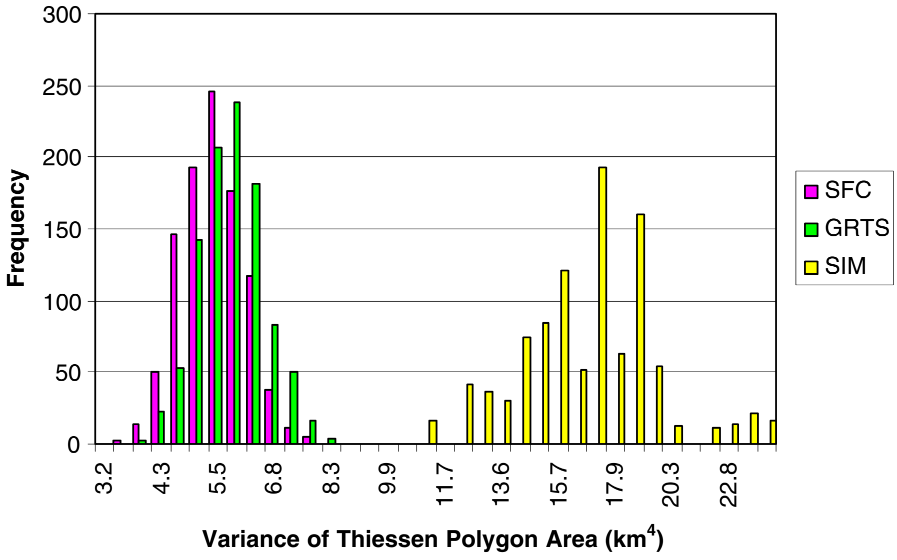

Fig. 5 Frequency distributions of the variances of the Thiessen polygons (spatial domains, or areas represented by the samples) surrounding plots from 1,000 iterations of the *SFC*, *GRTS*, and *SIM* sample selection methods. For each method, 139 samples were chosen, Thiessen polygons were drawn around the plots, and the variance of the areas of these 139 domains was calculated. This process was repeated 1,000 times for each method and histograms of variance values were built

distance bands that radiate inward from study area boundaries to determine if study area boundary regions appear to be sampled with an intensity similar to that found in interior regions.

Results

The GRTS and SFC methods performed similarly with respect to the attainment of spatial balance, and both performed much better than SIM (Fig. 5). The distributions of the SFC and GRTS spatial variance values not only had lower means, but also had shapes that were more symmetrical and narrower than that of the SIM.

SFC yielded the highest or tied for the highest CV for only 14 of the 66 land-cover classes, compared to 12 times for GRTS and 39 for SIM (Fig. 6). For all three methods and all 66 classes, at least 94% of the 95% binomial confidence intervals around the land cover proportion estimates captured the true values (a condition hereafter referred to as a success). When considering the set of 1,000 realizations, SFC had the highest frequency of successes (or tied for most frequent) for 31 of the classes, followed by GRTS (29 classes) and SIM (21 classes). When considering estimates of all classes on the landscape simultaneously, SFC yielded success for all classes in a single realization 443 out of 1,000 times, followed by GRTS (422) and SIM (381).

The assessment of plot densities within different distance bands around edges showed that none of the methods appear to undersample the boundary regions (Fig. 7).

Discussion

Our SFC approach has a number of desirable properties and compares favorably with a popular, well documented, commonly used sample selection procedure (GRTS). We expected to find that GRTS and SFC performed similarly with respect to attaining spatial balance and that both would be better than SIM. SIM created networks that were generally clumpier than the other methods, leading to higher variance of Thiessen polygon areas. Nonetheless, SIM could be useful for many applications, particularly where large numbers of inexpensive samples can be taken. On the other hand, GRTS and SFC were designed to spread points relatively evenly across the landscape, allowing for as complete a representation of an area as possible with a limited number of plots.

Similar levels of spatial balance could have been attained via a randomization heuristic (e.g., locate random points in the study area such that no point is within 1,000 m of another point), but this method is not deterministic, leading to less predictable spatial patterns in the results. Similarly, regular tessellation of the study area, like FIA uses (Reams et al. 2005),

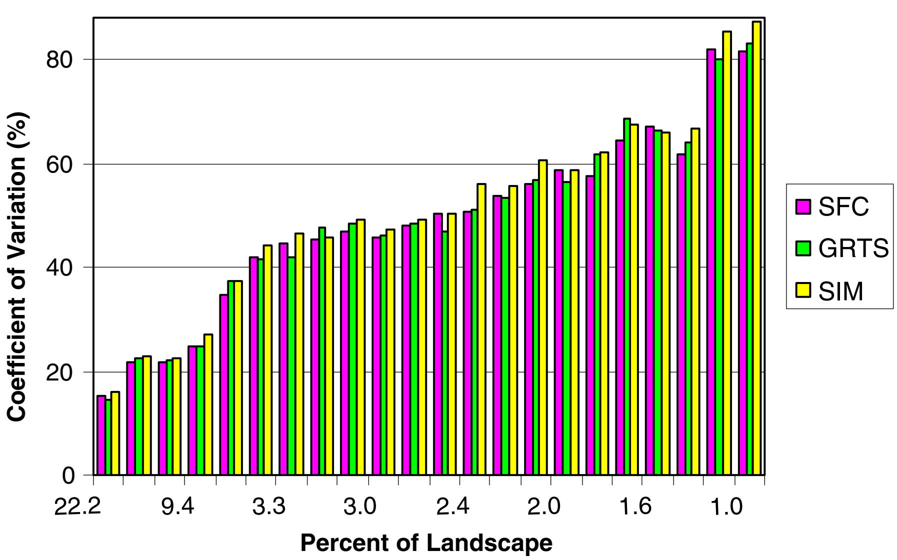

Fig. 6 *Coefficients of variation* (CV) of each of the distributions of landcover estimates for each landcover class. For display purposes, all landcover classes that were less than 1% of the landscape were grouped together into a single class, representing 12.3% of the landscape. Landcover classes are arranged in order (left to right) of decreasing landcover proportion. SIM had the highest CV most often

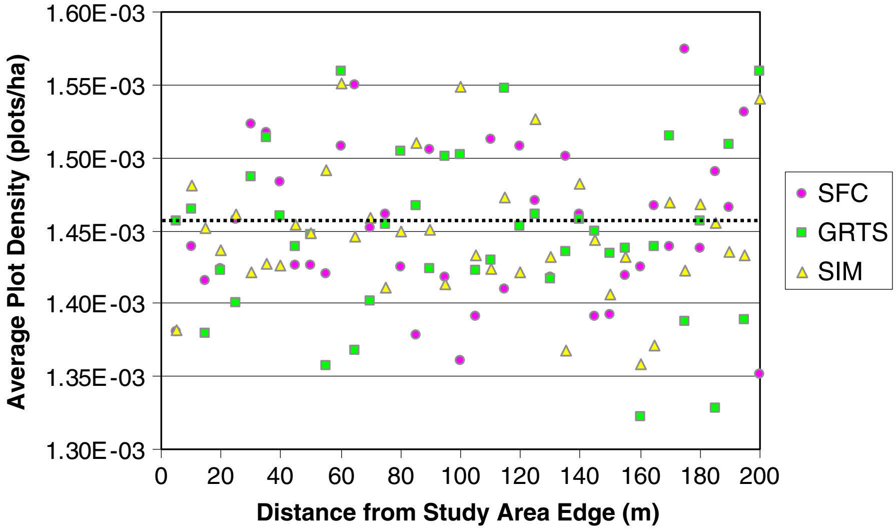

Fig. 7 Plot densities within concentric, 5-m wide distance bands radiating inward for 200 m from the study area boundary. *Dotted line* is the expected plot density for the entire study area

would achieve a high degree of spatial balance, but at the cost of being tied to a fixed sampling grid and the associated difficulties with intensifying the sample.

We wanted to compare the results of our implementations of the methods with respect to the accuracy and precision of land cover proportion estimates. As expected, SFC and GRTS performed similarly to one another and much better than SIM in terms of precision. SFC and GRTS had success for all 66 classes within a single realization at a rate that was about 16% and 11% higher (respectively) than that of SIM, further indicating that, as expected, estimates yielded by SIM are less precise than those yielded by methods that distribute the samples relatively uniformly across the landscape.

Cochran (1977), when comparing random and systematic sampling designs, points out that systematic samples generally yield more precise estimates than do random samples. However, he goes on to state that phenomena like hidden periodicity in the data could alter this outcome. In our experiment, SFC and GRTS forced the plots to be distributed somewhat evenly across the landscape, and we potentially created a situation in which some classes were less precisely measured due to the unique configuration of those rare classes on the landscape. SIM, on the other hand, had the potential to capture these unique classes due to a lack of restrictions on plot configuration.

Several authors have pointed out problems associated with analyzing field plot data collected from areas where spatial autocorrelation exists in the variable being measured, the environment, or both (e.g., Fortin et al. 1989; Legendre 1993; Legendre et al. 2002). The method presented here is intended to produce a sampling network that can be used in conjunction with classical estimators such as those described in Cochran (1977) to generate design-unbiased estimates of average or total amounts of environmental attributes. Samples are thus drawn independently of one another, so spatial autocorrelation is not a factor in sample-based estimation. If the sampling network is to be used for other purposes, however, we recommend an alternative sampling approach, or the adoption of methods described in Legendre (1993) to address spatial autocorrelation in field surveys.

We were very sensitive to concerns that our discretization procedure might lead to an anomaly in the sampling intensity close to borders. Figure 7 indicates that all three methods produce sample intensities in the areas immediately adjacent to boundaries that are close to the mean of those found in the interior of the study area. In earlier studies, we found that with repeated samples the periodicity of the 30-m grid point spacing of SFC created zones of high and low plot densities, spaced at approximately 30 m. The addition of the randomization of the plot location within the 30-m pixel centered on the plot eliminated this phenomenon.

The main advantage of our SFC method over GRTS and other strategies is that it is transparent to the user and is performed with a few lines of code in a common GIS. Each element from the sampling frame can be visualized, and the actual SFC can be drawn to connect the elements (Figs. 3 and 4). Another benefit of our SFC approach is that in the event of frame imperfections, new samples can be chosen randomly from the set of elements in the vicinity of the inaccessible element, instead of in a spatially disjoint manner (i.e., selecting a replacement randomly from the entire population). We acknowledge that our approach slightly alters inclusion probabilities for new plots chosen in this manner, but in practice, we feel that the added efficiency our method provides in the field outweighs this minor anomaly. Furthermore, other approaches that force a new plot to be chosen in a spatially disjoint manner might spatially bias new samples away from the vicinity of an inaccessible sample. Finally, another approach to dealing with inaccessible plots that does not require relocating a plot is simply using the inaccessible plot as an estimate of "area inaccessible to the survey", making the inaccessible plot issue moot.

Situations commonly arise in natural resource monitoring studies conducted by local, state, and federal agencies in which the sample network might need to be adjusted. For example, when land holdings are acquired or lost, or if funding shifts occur and fewer or more samples can be collected, decisions must be made about either dropping or adding samples to adjust the spatial balance and the sampling intensity. Overton and Stehman (1996) state that a desirable trait of a sampling design is the ability to restructure a sample to adapt to changes in objectives or the sampling frame. In the case of SFC, it is a simple matter to re-tessellate the study area (regroup the elements along the SFC), identify which of the newly formed clusters already contain samples, and either drop redundant plots or add new plots to empty clusters. If new territory is acquired and the original sampling intensity is desired, then new plots can simply be chosen by continuing the SFC into the newly acquired area and selecting plots from newly formed clusters as before.

Our SFC approach has the added advantage of representing the sampling frame as a set of discrete points that can easily be counted, grouped, intersected with ancillary GIS layers, and managed using a variety of common GIS, spreadsheet, or database procedures. These characteristics allow for complete control over how to visualize the sampling frame, add or subtract samples, panelize, and pre- or post stratify. For example, to create a sample with optimal (disproportionate) allocation, we could easily intersect our network of potential samples (elements) with a GIS layer that partitions the study area into strata, each of which could receive a different sampling intensity. We could then select a user-defined number of points within each stratum, combining all selected points into a single file at the end. Similarly, panelizing with our method is a simple spreadsheet exercise of sorting the plots by their spatial address and labeling them with panel identifiers. The known implementations of GRTS allow for unequal probability sampling, panelization, and stratification, but once the GRTS sample is chosen, the relationship between it and the set of unchosen elements is lost, making it more of a "black box".

We note that samples chosen in a systematic or semi-systematic way, such as those of SFC and GRTS, violate the equal joint probabilities of selection assumption, making variance estimation using standard equations problematic. Stevens and Olsen (2003) developed a variance estimator for samples chosen with a GRTS approach. This approach could be adapted to use with samples chosen with SFC. Similarly, an adaptation of Cochran's (1977) equations for estimation of variance with one unit per stratum could be used since SFC, in effect, creates equally sized spatial strata with one sample located in each stratum. We developed this SFC technique to be a more local, flexible extension of the hexagonal tessellation approach used by FIA, however, and feel that using the same assumptions that FIA uses and the variance estimators documented in Scott et al. (2005) leads to estimates of precision that can be used for monitoring decisions, with the understanding that the variance estimated in this way is likely slightly overestimated. If the assumptions that FIA uses are unacceptable to the practitioner, however, we recommend using the aforementioned alternative estimation methods.

A disadvantage of our SFC approach is that GIS-based storage and processing of large sampling frames can be less efficient than that achieved by a programming language that does not require disk storage of large numeric matrices. This limitation can be rectified easily, however, by basic GIS procedures that partition the study area into subregions that can be independently processed. A similar disadvantage occurs due to the very nature of discretizing an area into pixels—at boundaries, a ragged edge can be created, leading to a minor generalization of the study area boundary. This irregularity occurs in any sample selection procedure that relies on study area discretization, including at least some GRTS-like software implementations (e.g., S-draw (West, Inc. 2006) and RQRR (Theobald et al. 2006)). If there is concern about the practical impact of any irregularity of the sampling frame at sample boundaries, simple GIS procedures could be used to assuage this concern, or a finer discretization can be applied to further minimize boundary generalization.

In conclusion, the SFC method we propose shares most of the advantages of GRTS, produces a sample that is nearly identical (in our study, slightly better) in terms of spatial balance to that of GRTS, and provides the benefits of transparency, the ability to easily change the size of the sampling frame, and the ability to locate replacement samples in the vicinity of chosen elements that for some reason can not be sampled. Our SFC approach can be implemented with minimal code in a commonly used GIS and can be easily adapted to allow for unequal probability sampling and panelization. Disadvantages include problems with alterations of inclusion probabilities when samples are added or dropped (compared to the ordered list approach taken by GRTS), although they are likely of minimal practical impact. The SFC approach can serve as a useful tool for extending FIA-like methodology into smaller areas and establishing new monitoring networks for natural resource applications.

Acknowledgements We would like to thank Tony Olsen, John Stanovick, Ray Czaplewski, Susan King, Susan Wright, and an anonymous reviewer for their valuable reviews of this paper.

References

Bartholdi, J. J., & Platzman, L. K. (1988). Heuristics based on spacefilling curves for combinatorial problems in Euclidean space. *Management Science*, *34*(3), 291–305.

Cochran, W. G. (1977). *Sampling techniques*. New York: Wiley.

Cotter, J., & Nealon, J. (1987). *Area frame design for agricultural surveys*. National Agricultural Statistics Services, U. S. Department of Agriculture.

Fortin, M. J., Drapeau, P., & Legendre, P. (1989). Spatial autocorrelation and sampling design in plant ecology. *Vegetatio*, *83*, 209–222.

Gillespie, A. (1999). Rationale for a national annual forest inventory program. *Journal of Forestry*, *97*(12), 16–20.

Gregoire, T. G., & Scott, C. T. (2003). Altered selection probabilities caused by avoiding the edge in field surveys. *Journal of Agricultural, Biological and Environmental Statistics*, *8*, 36–47.

Henderson, R. C., Archer, E. K., Bouwes, B. A., Coles-Ritchie, M. S., & Kershner, J. L. (2005). PACFISH/INFISH Biological Opinion (PIBO): effectiveness monitoring program seven-year status report, 1998 through 2004. Gen. Tech. Rep. RMRS-GTR-162, U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station, Fort Collins, CO.

Jin, G., & Mellor-Crummey, J. (2005). Using space-filling curves for computation reordering. Proceedings of the Los Alamos Computer Science Institute Sixth Annual Symposium (published on CD). Los Alamos National Laboratory, Santa Fe, NM.

Johnson, W., Rhee, Y., & Shoemaker, O. (1993). Redesigning the consumer price index area sample. BLS Statistical Survey Papers, Bureau of Labor Statistics, Washington, DC.

Kincaid, T. (2006). User guide for spsurvey, version 1.4, probability survey design and analysis. Retrieved 12/5/2006, from http://www.epa.gov/nheerl/arm/analysispages/software.htm.

Lam, N., & Liu, N. (1996). Use of space-filling curves in generating a national rural sampling frame for HIV/AIDS research. *The Professional Geographer*, *48*(3), 321–332.

Legendre, P. (1993). Spatial autocorrelation: trouble or new paradigm. *Ecology*, *74*, 1659–1673.

Legendre, P., Dale, M. R. T., Fortin, M. J., Gurevitch, J., Hohn, M., & Myers, D. (2002). The consequences of spatial structure for the design and analysis of ecological field surveys. *Ecography*, *25*, 601–615.

Mandelbrot, B. (1982). *The fractal geometry of nature*. New York: Freeman.

McDonald, M. E., Paulsen, S., Blair, R., Dlugosz, J., Hale, S., Hedtke, S., et al. (2002). Environmental Monitoring and Assessment Program research strategy. EPA 620/R-02/002, US Environmental Protection Agency, Washington, D.C.

Olea, R. A. (1984). Sampling design optimization for spatial functions. *Mathematical Geology*, *16*(4), 369–392.

Olsen, A. R., Sedransk, J., Edwards, D., Gotway, C. A., Liggett, W., Rathbun, S. L., et al. (1999). Statistical issues for monitoring ecological and natural resources in the United States. *Environmental Monitoring and Assessment*, *54*(1), 1–45.

Olsen, A. R., Stevens, D. L., Jr., & White, D. (1998). Application of global grids in environmental sampling. *Computing Science and Statistics*, *30*, 279–284.

Overton, W. S., & Stehman, S. V. (1996). Desirable design characteristics for long-term monitoring of ecological variables. *Environmental and Ecological Statistics*, *3*, 349–361.

Palmer, C., Riitters, K., Strickland, T., Cassell, D., Byers, G., Papp, M., et al. (1992). Monitoring and research strategy for forests—Environmental Monitoring and Assessment Program (EMAP). EPA/600/4-91/012, U.S. Environmental Protection Agency, Office of Research and Development, Washington, DC.

Peano, G. (1890). Sur une courbe, qui remplit une aire plane. *Mathematische Annalen*, *36*, 157–160.

Reams, G. A., Smith, W. D., Hansen, W. H., Bechtold, W. A., Roesch, F. A. & Moisen, G. G. (2005). The forest inventory and analysis sampling frame. In W. A. Bechtold & P. L. Patterson (Eds.), *The enhanced forest inventory and analysis program—national sampling design and estimation procedures* (pp. 11–26)). Asheville, NC: Gen. Tech. Rep. SRS-80, U.S. Department of Agriculture Forest Service, Southern Research Station.

Saalfeld, A. (1998). Sorting spatial data for sampling and other geographic applications. *Geoinformatica*, *2*, 37–57.

Saalfeld, A., Zayatz, L., & Hoel, E. (1992). *Contextual variables via geographic sorting: a moving averages approach. Proceedings of the section on survey research methods*. Alexandria, VA: American Statistical Association.

Scott, C. T., Bechtold, W. A., Reams, G. A., Smith, W. D., Westfall, J. A., Hansen, W. H., et al. (2005). Sample-based estimators used by the Forest Inventory and Analysis National Information Management System.. In W. A. Bechtold & P. L. Patterson (Eds.), *The enhanced forest inventory and analysis program—national sampling design and estimation procedures* (pp. 43–67). Asheville, NC: Gen. Tech. Rep. SRS-80: U.S. Department of Agriculture Forest Service, Southern Research Station.

Shiver, B. D., & Borders, B. E. (1996). *Sampling techniques for forest resource inventory*. New York: Wiley.

Stevens, D. L., Jr. (1997). Variable density grid-based sampling designs for continuous spatial populations. *Environmetrics*, *8*, 167–195.

Stevens, D. L., Jr., & Olsen, A. R. (1999). Spatially restricted surveys over time for aquatic resources. *Journal of Agricultural, Biological, and Environmental Statistics*, *4* (4), 415–428.

Stevens, D. L., Jr., & Olsen, A. R. (2003). Variance estimation for spatially balanced samples of environmental resources. *Environmetrics*, *14*, 593–610.

Stevens, D. L., Jr., & Olsen, A. R. (2004). Spatially balanced sampling of natural resources in the presence of frame imperfections. *Journal of the American Statistical Association*, *99*, 262–278.

Theobald, D. M., Stevens, D. L., Jr., White, D., Urquhart, N. S., & Olsen, A. R. (2006). Using GIS to generate spatially-balanced random survey designs for natural resource applications. Retrieved 12/5/2006, from http://www.nrel.colostate.edu/projects/starmap/rrqrr_index.htm.

U.S. Forest Service (1992). *Forest Service resource inventories: an overview*. Washington, DC: U.S. Department of Agriculture, Forest Service, Forest Inventory, Economics, and Recreation Research.

U.S. Geological Survey. (2001). Vegetation spatial database coverage for the Voyageurs National Park Vegetation Mapping Project. USGS-NPS Vegetation Mapping Program, U.S. Geological Survey, Center for Biological Informatics, Denver, CO. Retrieved 1/17/2007, from http://biology.usgs.gov/npsveg/voya/metavoyaspatial.html.

West, Inc. (2006). S-Draw help file. Retrieved 12/5/2006, from http://www.west-inc.com/computer.php.