Meffe GK and Carroll CR (1999) *Principles of Conservation Biology*, 2nd edn. Sunderland, MA: Sinauer.
Naveh Z and Liebermann AS (1984) *Landscape Ecology: Theory and Application*. New York: Springer-Verlag.
Peterken GF (2002) *Reversing Fragmentation: Habitat Networks as a Basis for Woodland Creation*. Edinburgh, UK: Forestry Commission Practice Note.
Woodland Trust (2002) *Space for Nature: Landscape-Scale Action for Woodland Biodiversity*. Grantham, UK: The Woodland Trust.

Spatial Information

T Tokola, University of Helsinki, Helsinki, Finland

Introduction

A common representation of forest characteristics within spatial analysis and geographic information systems (GIS) can often be found in native objects whose interactions are based on simple distance and connectivity relationships. The spatial description of forest objects can be understood as a continuous two-dimensional process as an intensity field, or a collection of discrete locations of spatial objects. Geometrical features, such as points, lines, polygons, and raster cells, are commonly used to describe real-world objects and their characteristics in computerized mapping systems. Data modeling is a process that simplifies and defines real-world objects as database objects. Further spatial analysis may be engaged in when database objects have sufficient characteristics for spatial analysis.

The quantification of heterogeneity in forest areas has long been an objective of forest inventory and management. Heterogeneity depends highly on scale. The spatial and temporal variation of the property that can be detected will often depend on the spatial and temporal scale at which the property was sampled, and the size of the mapping unit. The information levels used in forestry reporting are hierarchically divided into: (1) tree level; (2) stand level; (3) farm level; (4) region level; and (5) country level. The data collection is normally based on measured sample units or subjective field observations that come from reporting units. The spatial pattern of reporting units can be mapped using remote sensing techniques or field observations.

The relative spatial distribution of forests and trees varies, because of changing land use practices, differences in the fertility of soil, and the hydrology, competition, and size distribution of trees. It is well known that the spatial distribution of seedlings in stands of natural generation depends highly on the location of mother trees and soil preparation affects the probability of survival of seedlings. Spatial information is used in forest inventory planning, and the construction of growth models and problems relating to forest regeneration and thinning. For example, the predictors of a spatial growth model for drained peatlands normally include variables such as the distance between the tree and the nearest ditch. The optimal sampling design of forest inventory can be defined if a spatial pattern of large variation is known and the size of a sample unit can be determined when the probability of tree occurrence can be modeled. Different indices and techniques have traditionally been applied to seedling surveys, in order to find out if the spatial distribution of seedling and saplings is regular. In addition, the effectivity of thinning and stand growth estimates depends highly on spatial regularity.

There are many forestry variables that are spatially sparse and scattered. This is often the case when one is assessing coarse woody debris in managed forests, or surveying threatened species. The spatial description of sparse populations can also be problematic. On the landscape level, information about spatial distribution of different key habitats and areas with a high ecological value has also been used to assess the probability of existing rare species. Field data about indicator species and remote sensing data about landscape features are valuable a priori information for estimating the presence/absence probability and for stratifying areas of interest.

Spatiality of Trees

The simplest point process model that can be used for the spatial pattern of trees is the Poisson process, which is typically used to produce random Poisson forests and when there is no interaction between the locations of trees. There are several modifications that stem from the basic model, such as the inhomogeneous Poisson processes, the Poisson cluster processes, and the doubly stochastic Poisson processes. The location of seedlings after natural regeneration is often generated using the Poisson cluster processes. Lattice-based processes are suitable models for spatial patterns of trees in plantations. Pair correlation processes produce patterns in which points either 'reject' (regular) or 'attract' (clustered) other trees to each other. Hard-core processes reject other trees with such a high intensity that other trees cannot exist closer than the radius of the core area. The Markov point processes and the Gibbs process are often used as well, because interactions between trees can be sufficiently modeled and empirical data can be used to create a probability model.

Different kinds of indices can also be used to measure the deviation of the given tree pattern from the Poisson forest. These methods of analysis can be divided into three groups depending on the measurements done regarding the population:

1. Number of trees within plots are counted.
2. The distance from tree (or random point) to closest tree is measured.
3. All trees in the forest are mapped.

The methods of the first two groups are more suitable for field work, while the latter group gives detailed information about the underlying process, and allows for the estimation of parameters necessary to begin a selected point process. The known field methods include the index of dispersion and the distance-based indices. Ripley's K, Moran's I, and Geary's C are commonly used for testing randomness or clusterness of tree patterns, although Moran's I and Geary's C have mostly been used for characterizing the autocorrelation. When permanent sample plots of Finland were analyzed, 57% of the plots had a regular tree pattern, while 25% were random, and 18% were clustered. When the basic pattern of trees is identified, it can be utilized directly to determine the sampling unit and design.

One common example in forestry would be the stand representation of discrete forest patches, and the partition of forests into distinct classes or strata. The discrete model is usually adopted when the boundaries of units can be unambiguously delineated. This happens, for example, when there are sharp discontinuities in attribute values. The delineation of stands is typically guided by three criteria: (1) the forest characteristics in different parts of the stand should be similar; (2) the stand should be a practical management and harvesting unit; and (3) the stand should be identifiable to allow for the monitoring and updating of information. Because these criteria can be contradictory, one or more of them are often compromised, and an estimation of stand values is often based on the simple summation of sample data, an approach which sometimes masks substantial variations found within discrete forest stands. This spatial autocorrelation can be studied using a dense network of sample plots within an area of interests, and correlograms/semivariance of variable of interest can be estimated using localized sample plot data. The spatial autocorrelation within forest stands is larger in forests with regular tree patterns than it is in forests with random or clustered tree patterns. In Finland, some studies indicate that within stands, the autocorrelation for the basal area and growing stock volume of trees only exists within a 20–30 m distance, while 5–10 m distance intervals are used. The continuous description of forest characteristics has not received much attention in the forest resource inventory.

Spatiality of Landscapes

Habitat mosaics have been found to affect diversity and dynamics in both pristine and managed boreal forests, and many important processes have also been identified as being driven or affected by landscape heterogeneity. Efforts to quantify the spatial heterogeneity of landscapes began in the 1980s, but have accelerated in recent years, so that at the present there are hundreds of indicators that allow for some sort of quantification of various aspects of spatial heterogeneity at a landscape level. Thematic maps and satellite image-based products have also been used to estimate landscape indicators on a regional level.

Composition

Composition is typically indicated by the number of categories or classes in the map, the proportion of each class in the map, and the presence of diversity. Diversity measures typically combine two components of diversity: richness, which refers to the number of classes present, and evenness, which refers to the distribution of objects that are among the classes. Typical diversity indices are Shannon's and Simpson's.

Spatial Configuration

Spatial configuration of properties attempts to describe the spatial characteristics of individual patches, and the spatial relationship among multiple patches. Patch-based measures of pattern include size, number, and the density of patches. Useful edge information includes the perimeter of individual patches and various edgemetrics that incorporate the contrast between the patch and its neighbors. For example, there is less contrast between a mature forest stand and a young stand than there is between a mature stand and clear-cut areas. Patches can take a variety of shapes, making shape difficult to quantify. Most shape indices use a perimeter–area relationship. A widely used index related to both patch size and shape is the 'core area,' which is the proportion of a patch that is further than the specified distance from an edge. Patch cohesion has been proposed to quantify the connectivity of habitat as perceived by organisms dispersed among binary landscapes. Some pattern indices examine spatial neighborhoods, patch orientation, and isolation. These were developed primarily to predict the relative connectivity of habitat islands.

Contagion

Contagion is designed to quantify both composition and configuration. It measures the extent to which cells of similar classes are aggregated. It is calculated using the frequencies with which different pairs of classes occur as adjacent pixels on a map. This appears to summarize the overall clumpiness of areas of interest.

It has been shown that there are only five independent factors among 55 different landscape metrics. These are:

1. Number of classes.
2. Dominance.
3. Contagion.
4. Fractal dimension from perimeter/area.
5. Average patch perimeter/area ratio.

In Finnish forests, landscapes are transformed into a mosaic of managed forest stands of small size, in which species composition has become more homogeneous, and the age distribution of stands more even. Low values of contagion characterize the Finnish landscape, which is dominated by many small size-dispersed patches. The development of species occurrence and landscape level habitat models, based on National Forest Inventory (NFI) data and environmental information at the national level, highlight small variations among the indicators. This seems to have important implications for the spatial distributions of species. Landscape metrics has been proven to be an efficient method of monitoring forest characteristics, which might have important applications for management planning in order to improve forest biodiversity.

Spatiality of Regions

On a national level, both correlation and autocorrelation functions have been employed to study large structures of forest area, and volume of forests. In Scandinavian forests there is a slightly increasing autocorrelation almost until a 200 m distance, when larger distances are applied. This information is used to estimate the optimal distance between inventory tracts, the overall shape of the tract, and the distance between sample plots within tracts. To obtain a sufficient sampling set-up for the entire nation of Sweden, the country is divided into five regions with marginally diverse correlation functions for different variables. The final sampling design is based on their spatial characteristics and other practical considerations.

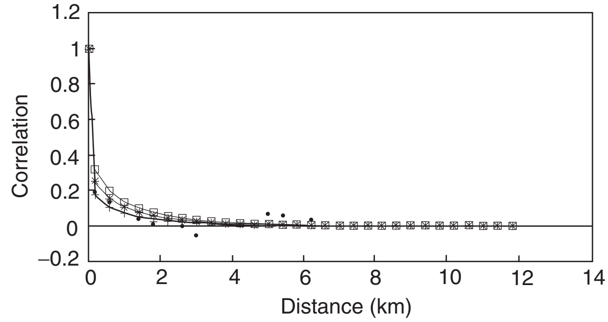

Figure 1 The estimated correlation function and empirical correlations for growing stock volume, as a function of distance between sample plots in three different areas. Reproduced with permission from Kangas A (1993) Estimating parameters of systematic cluster sampling by model based inference. *Scandinavian Journal of Forestry Research* 8: 571–582.

The correlograms (**Figure 1**) have also been used to estimate the standard error as regards large areas. The standard errors of systematic cluster sampling can be estimated using model-based estimators, which utilize the parameters of correlation functions. That way, the information about spatial dependency can be utilized in error estimation and spatial structure of forests is taken into account.

Border Effects in Spatiality

Spatial analysis is always connected to edge effects. In addition, edge effect correction methods are often used in inventory procedures, as well as in the analysis of limited empirical data sets. Four main methods are typically applied:

1. Plus sampling: additional data is measured, so that entire neighborhoods can be covered.
2. Minus sampling: buffer zones with a plot radius are generated for neighborhood calculation, and sampling is made in the core area only.
3. Toleroid edge correction, or mirror-sampling: an edge zone with plot radius is copied to the edge buffer, and the neighborhood is calculated from 'duplicated data.'
4. Weighting of observations: neigborhood data are weighted according to the probability of information existence.

Spatial Information in GIS and Remote Sensing

A forest information system contains information for decision-making in forestry, and forest inventory yields data for such a system. The initial and most

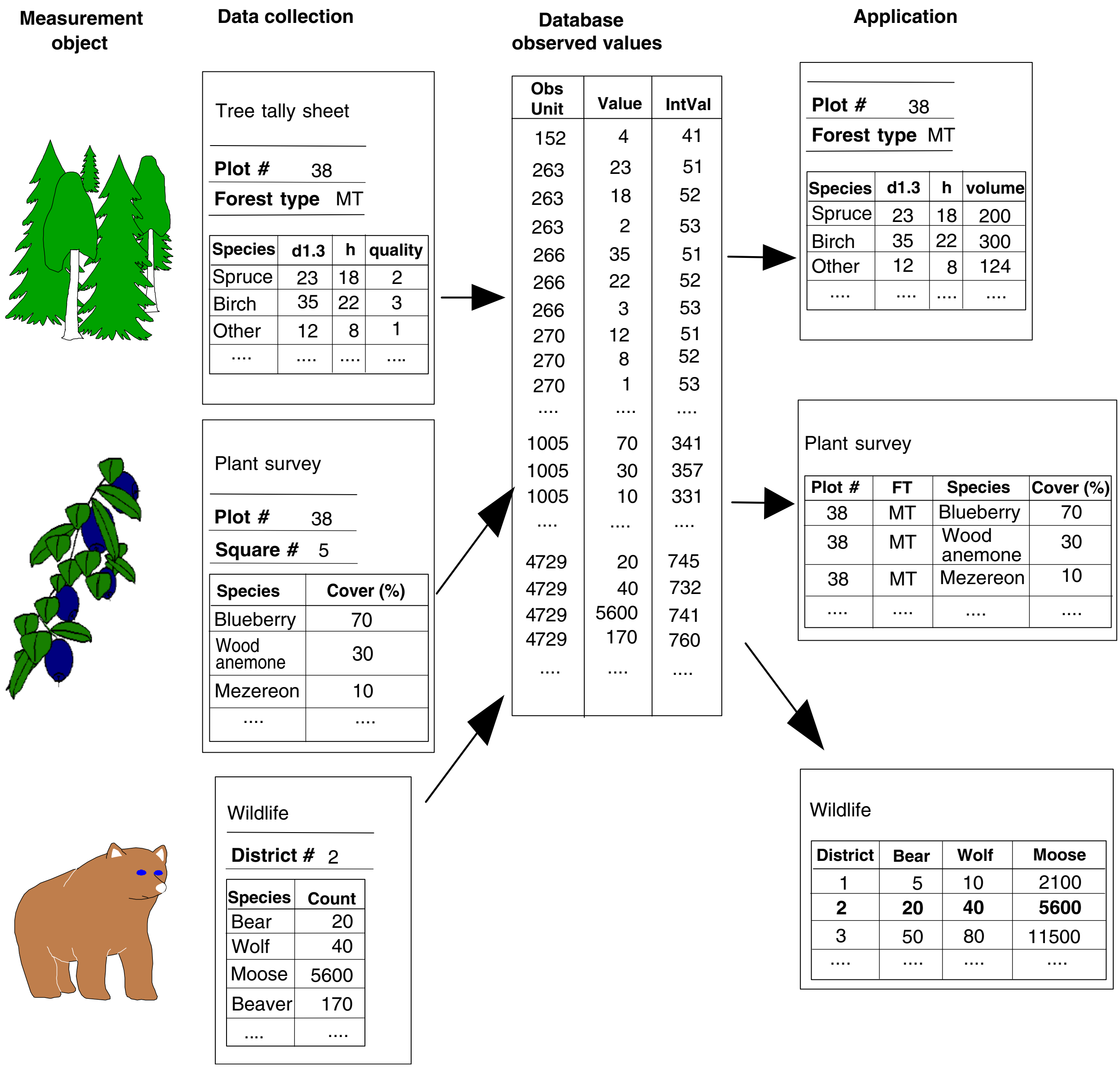

Figure 2 Storage structure for forest inventory attribute data. Reproduced with permission from Tokola T, Turkia A, Sarkeala J, and Soimasuo J (1997) Entity-relationship model for forest inventory. *Canadian Journal of Forest Research* 27: 1586–1594.

important phase of the forest database system design is the construction of a data model, which is primarily used to perceive, organize, and describe data in a conceptual schema. Real-world objects are defined in the data model in such a way that they can be described in databases (**Figure 2**).

Each observation and measurement in a system is linked to the geometry of GIS. The geometry of forestry objects can be presented using vector (point, line, polygon), raster, and dynamic segment-based models (**Figure 3**). Vector-based description is typically used for discrete phenomena (i.e., stand border maps) and raster-based systems (i.e., stratification of volume, elevation) do their best when continuous spatial surfaces are presented. Dynamic segmentation is used for objects which locate near line features, and can be located using the distance measure from known location along the line. Such a feature could be a hydrographic measurement unit along a river that has a particular distance from a known crossing.

The requirements for interoperable, computationally scalable software tools suggest the need for developing open software standards such as the Open GIS Consortium (www.opengis.org). The current Open GIS data model supports geometrical primitives, but requires application-specific definitions from the user side, as well as support for advanced analysis, which needs to be built externally.

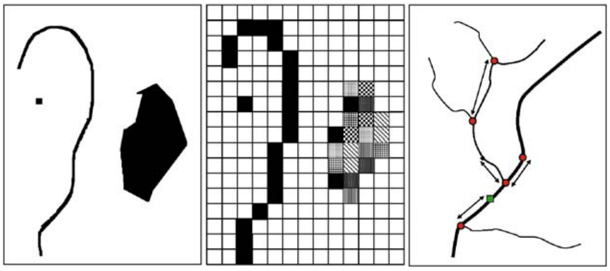

Figure 3 Vector-, raster-, and dynamic segmentation-based models are used to describe the geometry of real-world objects within global information systems. Reproduced from Tokola T and Kalliovirta J (2003) *Paikkatietoanalyysi*. Publications 34. Helsinki, Finland: Department of Forest Resource Management.

Remote sensing is considered to be an efficient tool for data acquisition and the updating of information into forestry GIS system. Yet, the remote sensing sensors can vary a lot, and it is important to realize that the spatial dependency of forest objects also affects the data collections phase.

When the pixel size is larger than the forest objects, information is perceived to be lost because the spatial resolution of images is so low. In contrast, when very-high-resolution images are used, adjacent pixels give the same information, because it is highly probable that they are taken from the same object. Another drawback of high-resolution data is that the amount of data soon becomes too enormous to compute. A worst-case scenario occurs when the use of unsuitable resolution results in erroneous image interpretation, due to the spatial autocorrelation of neighboring pixels. Good examples are the Landsat MSS images which have strong positive autocorrelation between neighboring pixels. It has been found that there are three reasons for this phenomenon: (1) a natural continuity of land cover, compared to the spatial resolution of the imaging system; (2) a positive correlation caused by the imaging system itself; and (3) image processing algorithms such as resampling. Studies with Landsat MSS images have indicated that the radiance of one pixel can affect the radiance of surrounding pixels that are 4–6 pixels apart. With higher-resolution systems like Landsat TM and SPOT, the positive autocorrelation can be even higher. The autocorrelation of pixel gray values in a forest environment is usually clear, and depends on the crown sizes and crown cover proportions of the trees. When semivariograms derived from aerial photographs have been studied in boreal areas, spatial resolution in which the variance was maximized was 2–4 m. Principally, such results have indicated that the local variance curve maximum is dependent on the object size, and a maximum is reached when the resolution is somewhat smaller than the object size. Nowadays, single tree-based digital delineation techniques are often used in various remote sensing materials, and 0.5 m resolution image material is generally used during these interpretation processes.

Satellite image-based surveys often utilize distant field sample plots from target areas. Normally, the margin of error increases when distant field sample plots from existing field samples are used. In one study, the best results were achieved when plots were within a 20-km range. Stand margin areas are also critical in remote sensing. The accuracy of growing stock volume estimation near a stand edge can be much lower than it will be inside the stands.

Simulated Forests

Simulation models of landscape levels are needed to understand large-scale variation, as well as predict the development of forests under different management schemes.

Alternative sampling designs for forest inventories can be evaluated through several means. One way involves carrying out actual manual inventories in the target forest. Computer-simulated samplings offer a cheaper and more flexible option for this. Following the first inventory round, it is possible to use remote sensing data to create a computerized depiction of the selected target area. The main advantages of such a simulation are that by manipulating input variables, a process of controlled experimentation and sensitivity analysis can take place. In spite of certain disadvantages, a simulation approach provides a flexible method for experimentation with prospective sampling designs, and is free from the restrictions associated with the analytical error propagation methods.

Using simulated sampling designs, changes in the sampling error of the estimate can be examined, and one can discover the best cluster design. In the simulation approach, all local small-scale characteristics of spatial variation can also be included in the analysis. In numerous other studies, entire spatial variation has been described by using the average covariance function. Certain simulation models use satellite image-based simulated models of forests. These models normally show a moderate fit with the field data. This type of model can also be used to test different sampling designs for carrying out forest inventory, and seemed to be representative in terms of the sampling error of estimate when compared to the overall structure of the spatial distribution of a given forest.

The standwise interpretation result of small-scale remote sensing materials does not vary within forest stands, and the correlation between neighboring satellite image pixels is relatively high. This correlation is caused by observation characteristics of the satellite sensor. While the small-scale accuracy of Landsat TM interpretation is low, and the tree-level information is missing, the main focus of this type of inventory planning concentrates on the shape of clusters and distance between plots. If small-scale variation was available, it could also perhaps determine the size and shape of plots. Unfortunately, fairly detailed aerial photographs would be needed, if we wished to see a true realization of stand structure taken using remote sensing material.

Most approaches to creating spatially explicit simulations of forest landscapes have been developed in North America, yet a large amount of simulators are available for specific purposes. They have been used to demonstrate the effects of different potential disturbance regimes, and for planning alternative forestry management schemes. Trying to make predictions as regards landscape scale, one has to deal with all of the ecological site types and tree species as well as all possible stand development scenarios. In addition, landscape-level processes have to be incorporated. One solution to such difficulties has been to simplify the description of forest stands by discarding most of the quantitative attributes, such as stand basal area or size of trees, and to use a semiquantitative approach to describe tree stands by the age structure of each tree species. It is important to note that the purpose of simulation models is not necessarily to predict reality directly, but rather to reveal the logical consequences of the assumptions incorporated in the structure of application-specific computer models and parameter values. Complex spatial models are indeed hard to evaluate, because it is difficult to find sufficient empirical data sets, as well as to compare exactly which aspects of spatiotemporal patterns are crucial for either a correct simulation, or a future model application.

See also: **Inventory**: Forest Measurements; GIS and Remote Sensing; Large-scale Forest Inventory and Scenario Modeling; Multipurpose Resource Inventories.

Further Reading

Congalton RG (1988) Using spatial autocorrelation analysis to explore the errors in maps generated from remotely sensed data. *Photogrammetric Engineering and Remote Sensing* 54: 587–592.

Hyppänen H (1996) Spatial autocorrelation and optimal spatial resolution of optical remote sensing data in boreal forest environment. *International Journal of Remote Sensing* 17(17): 3441–3452.

Matern B (1960) Spatial variation. *Meddelanden fron Statens Skogsforskningsinstitutr* 49(5): 1–144.

Korhonen K and Maltamo M (1991) The evaluation of forest inventory designs using correlation functions. *Silva Fennica* 25(2): 77–83.

Riitters KH, O'Neill RV, Hunsaker CT, *et al.* (1995) A factor analysis of landscape pattern and structure metrics. *Landscape Ecology* 10: 23–39.

Ripley BD (1981) *Spatial Statistics*. New York: John Wiley.

Tomppo E (1986) Models and methods for analysing spatial pattern of trees. *Comunicationes Instituti Forestalis Fenniae* 138: 65.

Uuttera J, Haara A, Tokola T, and Maltamo M (1998) Determination of spatial distribution of trees from digital aerial photographs. *Forest Ecology and Management* 110: 275–282.