Spatial characteristics of vegetation index map in urban area derived by variogram analysis

Keiji Osaki
International Christian University
Tokyo Japan

It is important to keep monitoring the characteristics of vegetation in urban areas by remotely sensed data from an environmental viewpoint. Recent sensors on earth observation satellites have excellent and very high spatial resolution of a few meters on the surface of the earth. However, the higher the resolution of sensor becomes, the more difficult the analysis of land surface seems to be due to too fine aspects in the satellite imagery. The primary objective of the research presented here is to clarify the spatial characteristics of vegetation distribution in urban areas by analyzing properties of variograms . To depict the spatial patterns of the observed scenes, quantity measured by second-order statistics has been used in applications such as mining exploration and other engineering fields. Related practitioners call the field of spatial statistics 'geostatistics'. An important concept of the quantity inherent to the scenes is spatial continuity and is measured as covariance and semivariance. The semivariance plays a very important role in the analysis of data's spatial statistics in the present research. Among many vegetation indices, NDVI(normalized difference vegetation index) has been most widely used for investigation of environmental assessment. The NDVI data used for the current analysis is derived from the multi-spectral data of 'QuickBird' (earth observation satellite) with the resolution of 2.8m. The NDVI lies in its characteristics that can reduce the multidimensional data yielded from multi-spectral sensor systems to a single index which is sensitive to various characteristics related to vegetation activities such as biomass, productivity, leaf area, amount of photo-synthetically active radiation, and percent vegetative ground-cover etc.

The semivariogram is defined as the following equation:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1,N} (z_i - z_{i+h})^2$$

A useful measure of spatial variation in the values of a variable z is the semivariance, which is half the average squared difference in z values between pairs of sample points. The key to investigation of the semivariance is the construction of a semivariogram, which is a plot of the semivariance, as a function of distance h. At a distance referred to as "range", the semivariance levels off to a relatively constant value, referred to as the "sill". This implies that beyond this range, z values are no longer spatially correlated.

The variogram can provide the spatial structure or patterns of observed objects on the earth quantifying dissimilarity as a function of separation and direction. Here we skip the effects of anisotropy for simplified analysis.

As is shown in the following equation, we introduce two sets of parameters which represents the characteristics of variograms, that is, two 'sills' and 'ranges'. By nonlinear least squares regression method 'nls' in 'R', we can derive two sets of characteristic parameters of variograms which can let variograms be fitted to 'nested spherical model' (Hans Wackernagel, 2003)[1] :

$$\gamma(h) = \begin{cases} \sum_{k=1}^{2} \mathrm{Sill}_k \left\{ \frac{3}{2}\frac{h}{\mathrm{Range}_k} - \frac{1}{2}\left(\frac{h}{\mathrm{Range}_k}\right)^3 \right\} & \text{for } 0<h<\mathrm{Range}_k \\ \sum_{k=1}^{2} \mathrm{Sill}_k & \text{for } h > \mathrm{Range}_k \end{cases}$$

where sill1, range1, sill2 and range2 are defined by nonlinear least squares regression fitting. We can identify the difference between two areas i.e. "range" is larger for area, which contains much natural objects such as vegetation than for area like urban area with short-range which contains many artifacts. Since it demands vast computation cost for calculation of semivariances for remotely sensed scenes, we use ordinary simple 'random sampling ' method to select sampling pixels from the target areas.

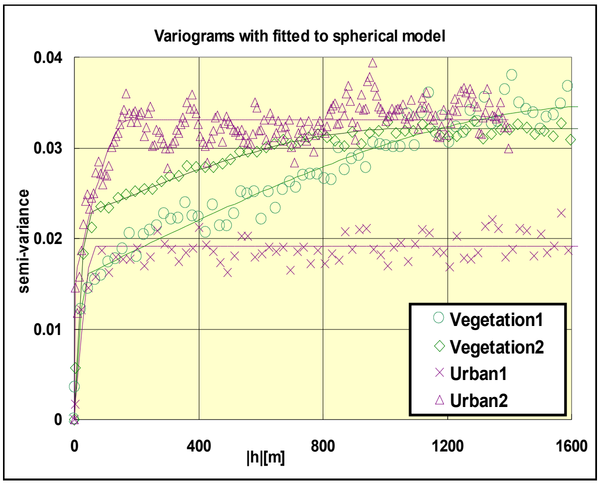

Figure Variograms for vegetation and urban areas with lines fitted to nested spherical model.

To extract interesting and important features of variograms in certain range of 'lag's where variograms would show rich vegetation characteristics, variograms and their fitted curves are calculated by nonlinear least squares regression fit to nested spherical model in the above equation. It is interesting shown in Figure that two types of rich and poor vegetation areas show their intrinsic spatial pattern in variograms. Variograms in urban areas show almost zero slope in the range of 'lag' (500m < |h| <1000m), while in richly vegetated areas have slower increase of variogram in that range.

References

[1] Hans Wackernagel, *Multivariate Geostatistics*, Springer, 2003.