Estimation of vegetation coverage in urban area by variogram analysis

Keiji Osaki

Natural Sciences, International Christian University, 3-10-2 Osawa, Mitaka, Tokyo, 181-8585, Japan – keiji@icu.ac.jp

Abstract –There exist several problems in estimate of vegetation coverage derived from earth observation satellite data of various sensors with different spatial resolutions. The seamless use of vegetation indices is to be guaranteed for recent very high-resolution sensors. We have investigated the dependence of NDVI (Normalized Difference Vegetation Index) on sensor's spatial resolution. It has been attempted to investigate interesting characteristics of semi-variance of QuickBird scenes in urban areas with highest spatial resolution. The characteristics of variograms in NDVI maps for rich vegetated and/or poor vegetated areas show apparent different features. In order to make it clear that the vegetation coverage of urban area has some correlation with the variogram's characteristics, we have investigated the NDVI distribution map with simulated different resolutions derived by discrete wavelet transform method. It is manifested that the properties of variograms in NDVI map show a correlation with the vegetation coverage property.

Keywords: vegetation coverage, variogram, sensor resolution, QuickBird.

1. INTRODUCTION

It is possible to monitor environmental circumstances of densely populated area by assessment of vegetation coverage in target area. One of environmental issues in urban areas in metropolitan cities is so-called 'heat island' effect that is pertinent to shrink of vegetation area and its activity. It is important to monitor the characteristics of vegetation in urban areas by remotely sensed data. To achieve monitoring its temporal changes regularly, various earth observation satellite data is very useful with its advantages for wide and periodical coverage of target area. However, since urban areas contain various constituent objects in a rather small area, they are often observed not as a homogeneous area but a heterogeneous area as mixed pixels by sensors with an insufficient ground spatial resolution. Since the mixed pixels are not representation of a single land cover type, it is desirable to use sensors with higher ground spatial resolution for the analysis of urban areas. For the mean while, there arises a trade-off among higher spatial resolution of the sensors, its narrower coverage and difficulty in grasping global context of the scenes due to their excessive variation of objects. Further, it is important to consider the continuity of values of vegetation indices to monitor vegetation over a long period. It is also necessary to compare .

Table 1. Specification of used products acquired by OuickBird and ASTER.

No	Sensor	Spatial Resolution[m]	Acquisition Date	Bands for NDVI	Products
1	QuickBird	2.8	2003/4/9	Band3,4	Standard
2	QuickBird	2.8	2003/9/3	Band3,4	Standard
3	ASTER	15.0	2003/9/5	VNIR(2,3N)	ASTL01B

Acknowledgements: ERSDAC and NASA have provided ASTER data. DigitalGlobe provides all QuickBird data

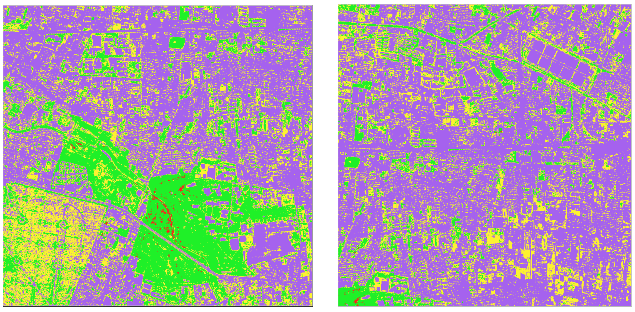

Figure 1. NDVI(Normalized Difference Vegetation Index) distribution map for rich vegetation area (Left) and poor vegetation area (Right) in western part of Tokyo.

differences in values of vegetation indices among new various s ensors with different spatial resolutions. In spite of several defects of Normalized difference vegetation index (NDVI) as a vegetation index, its continuity as obtained from NOAA's AVHRR over twenty years shows its advantages (S.O. Los, 1994).
The advent of sensors with high spatial resolution has offered us lots of advantages for the analysis of vegetation coverage in urban areas. It has also presented some difficulties and challenges in incorporation of vast accumulated data from sensors with lower spatial resolutions. The higher the resolution of sensors becomes, the more difficult the analysis of land surface seems to be due to too fine aspects in the satellite imagery. It has been pointed out that sensed data with an appropriate spatial resolution should be applied to observed objects with their intrinsic spatial patterns on the earth (Aaron Moody, 1995). Their results indicate significant relationships between the spatial characteristics of cover types and scale-dependent proportion errors. A related trend was found earlier by Turner et.al. (Turner, M., 1989). In a somewhat different context, Marceau et.al. investigated the impact of measurement scale and aggregation level on the information content of images and classification accuracy (Marceau, D., 1994). Key concept of scaling-dependency lies in an important relationship between spatial properties of scenes and spatial resolution of the sensor, which observes the scenes. And besides, we should take into account the concepts of aggregation and/or scale-reduction from images with finer resolution to coarser one. In order to extract appropriate features from the images with high resolution, it might be necessary to make use of global context of rather coarser images.

Detection of temporal change in vegetation in urban areas is one of important issues for an environmental monitoring. The seamless exploitation of accumulated data of vegetation index by many sensors for many years is required for precise analysis of vegetation coverage. In other words, it is required to use remotely sensed data for analysis of vegetation coverage and land classification in urban areas with fitted sensor's spatial resolution.

The primary objective of the research presented here is to clarify the relationship between vegetation coverage properties in the NDVI map and the spatial structures of the objects in the scenes.

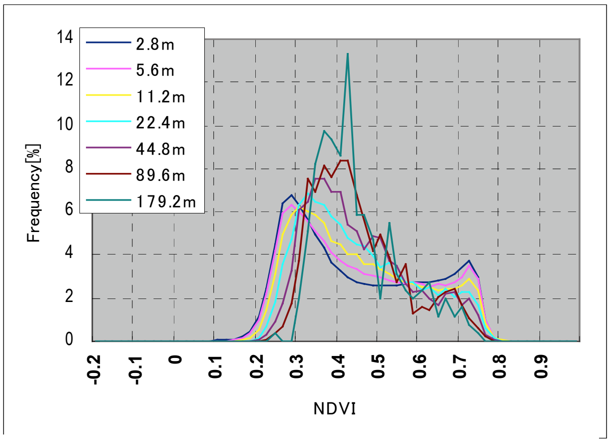

Figure 2. Comparison of NDVI histograms for rich vegetation area derived by QuickBird scene and its resolution-reduced scenes.

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 (Journel, A.G., 1978). 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. Then the dependency of NDVI on the spatial resolution of the sensor is possible to be speculated by the results based on the above relationship.

So it has been expected to evaluate the scale or resolution dependence of vegetation index by several scale reduction methods. Scale reduction methods such as simple down sampling or pixel averaging often give insufficient results for aggregation of the images with lower resolution from those with high resolution. To avoid the deficiency, it is utilized to use discrete wavelet transform for reduction of the spatial resolution of the images properly. Adopted method to achieve the above purpose is 'Multiple Resolution Analysis' by discrete wavelet transform because of several advantages especially in reversibility of resolution change (Jean-Paul Donnay, 2001).

Initial NDVI data with the high resolution is derived from the multi-spectral data of 'QuickBird' with the resolution of 2.8m. Multi-spectral data from ASTER with the resolution of 15m on the platform TERRA is also employed for the comparison with the simulated results of the current analysis of variograms. Through the investigation of the properties of semivariances for various scenes that are entire scenes or sub-scenes sampled with respect to several NDVI ranges which contain natural objects such as vegetation or artifacts, it is revealed that quite interesting characteristics of the relation between the objects' spatial scale and the sensor's spatial resolution.

2. METHODOLOGY, DATA AND RESULTS

Most widely exploited vegetation index is Normalized Difference Vegetation Index since it has several advantages such that ratio-based indices can reduce topographic effects on spectral response and other several noise influences. NDVI has been extensively used to assess vegetation productivity and its coverage (Groten, S.M., 1993). One of the reason of the usefulness of 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 canopy characteristics such as biomass, productivity, leaf area, amount of photo-synthetically active radiation, and percent vegetative groundcover (Dale A., 1997).If an appropriate atmospheric correction for satellite data is not readily available, we can approximately use at-sensor radiance instead of surface reflectance.

In order to investigate closely the resolution dependent properties of NDVI, it is preferable to derive resolution-reduced data from the finest resolution data. Among several methods of resolution reduction, we have chosen and utilized the discrete wavelet transform as the fitted way of technique to the present research.

The wavelet transform can lead to a time-frequency representation of the data under investigation: in the case of images, the wavelet transform leads to scale-space representation (Jean-Paul Donnay, 2001). In the approach of multi-resolution analysis, the size of a pixel is defined as a resolution of reference to allow a measure of local variation in the image. The definition and detailed properties of the wavelet transform are skipped here due to the limitation of space.

Our method of decomposition of the images with high resolution is a conventional discrete wavelet transform based on Haar wavelet mother functions since the Haar wavelet function is very simple and has good computational cost. Besides, it has an excellent property of reversibility of multi-resolution images. It enables us to reconstruct the images with hierarchically different resolution by combining wavelet components and scaling components in the discrete wavelet transform. By discrete wavelet transform , two kinds of resolution-reduced images are generated, one is the averaged image with coarser resolution and others are three wavelet component images which consist of wavelet coefficients. Images with reduced spatial resolutions are generated by the discrete wavelet transform from an initial image with the finest resolution. Its size is confined to 2 to the power in the number of pixels of row or column due to the restriction of discrete wavelet transform. The specification of currently used satellite data is shown in Table 1.

The NDVI distribution maps of item No.1 in Table 1 are shown in Fig. 1 as an color mapped image of 1024 x 1024 pixel where red, green, yellow and purple correspond to the following NDVI ranges : over 0.6, 0.6~0.4, 0.4~0.2 and below 0.2, respectively. Two regions are selected as 'rich vegetation' and 'poor vegetation' for the further detailed variogram analysis. The threshold value 0.2 of NDVI for vegetation area is determined by an inspection of two-dimensional scatter plot of two bands of near

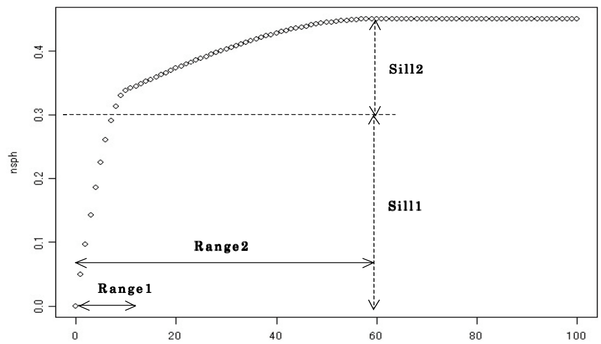

Figure 3. Example of two nested spherical variogram with sill1 = 0.3, range1 = 10, sill2 = 0.15 and range2 = 60.

infrared and visible red.
The dependency of frequency of NDVI is calculated by multi-resolution analysis and is shown in Fig. 2. Among resolution reduced data simulated by discrete wavelet transform, an interesting characteristics is found that there is an unchanged mean value and almost unchanged frequency value near two ranges of NDVI that are 0.3 and 0.6 .

There exists a limitation to investigate the spatial patterns or characteristics of two-dimensional digital images by the first-order statistics such as mean, standard deviation, median and histogram. Though the histogram is closely related to the probability density function in the scene, it cannot describe the characteristics of measures of pixels associated with their locations and distance of their pairs. Semivariance that is one of the second-order statistics characterizes the spatial structure of data sets, i.e. continuity or roughness. It is useful quantitative statistic measure even in the case common descriptive statistics and the histograms fail to identify the difference between similar data sets in terms of first-order statistics.

In geostatistics, the covariance function of a one-dimensional data set is defined as (Isaaks, E.H., 1989),

$$C(h) = \frac{1}{N(h)} \sum_{(i,j)|h_{ij}=h} z_i z_j - \mu_{-h}\mu_{+h} \qquad (1).$$

The parameter, h_{ij}, is called the lag and is the distance between the pairs of data points located at i and j. The quantities are the mean of z of all data points at a distance of -h or +h from some other data point; N(h) is the number of data point pairs separated by h.
The semivariogram is given by the following equation

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

The semivariogram usually exhibits a characteristic shape, increasing from small lags to larger lags. The plateau where becomes more or less constant is the sill. The distance from zero lag to the onset of the sill is the range.

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 Fig. 3, we introduce two sets of parameters which represents the characteristics of variograms, that is, two 'sills' and 'ranges'. By non-linear least squares method, we can derive two sets of characteristic parameters of variograms which can let variograms be fitted to 'nested spherical model' (Hans Wackernagel, 2003) :

where sill1, range1, sill2 and range2 are defined by non-linear

$$\gamma(h) = \begin{cases} \sum_{k=1}^{2} Sill_k \{ \frac{3}{2}\frac{h}{Range_k} - \frac{1}{2}(\frac{h}{Range_k})^3 \} & for\ 0 < h < Range_k \\ \sum_{k=1}^{2} Sill_k & for\ h > Range_k \end{cases} \qquad (3)$$

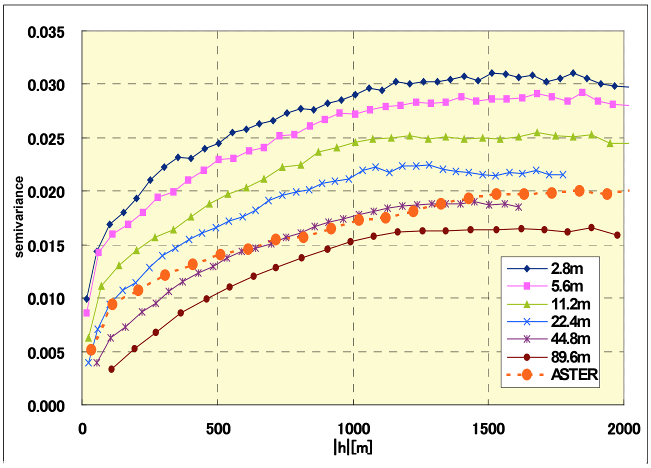

Figure 4. Dependence of variograms on sensor spatial resolution for simulated scenes with reduced resolution derived from 1024 x 1024 QuickBird scene including richly vegetated areas. Variogram derived from observed data by ASTER lies near simulated 44.8m variogram.

least squares 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.

In Fig. 4 and 5, the different characteristics of variograms between rich vegetated areas and poor ones are apparently shown. The values of 'Range' are fairly smaller in urban (poor vegetation) areas than in rich vegetation areas. The values of 'Sill' decrease according as the sensor spatial resolution becomes coarser. Simulated results of 'Sill' and 'Range' dependency on the sensor resolution are summarized in Fig. 6 and 7. Especially in Fig. 7, the different aspects of 'Range2' between rich and poor vegetation areas with respect to the sensor resolution are seen. To extract interesting and important features of variogram's in certain range of 'lag's where variograms would show rich vegetation characteristics, variograms and their fitted curves are calculated by non-linear least squares fit to nested spherical model in eq. (3). The interesting results are shown in Fig. 8. Two types of rich and poor vegetation areas show their intrinsic spatial pattern in

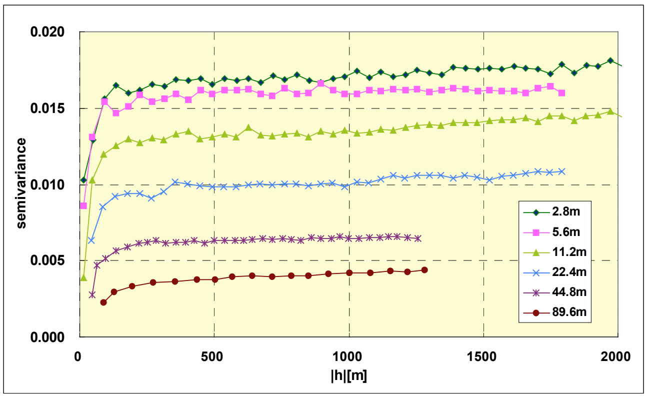

Figure 5. Dependence of variograms on sensor spatial resolution for simulated scenes with reduced resolution derived from 1024 x 1024 QuickBird scene including poorly vegetated area.

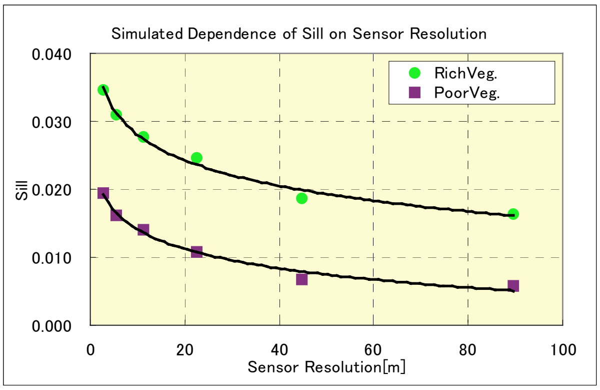

Figure 6. Simulated dependence of 'sill' on sensor resolution.

variograms. Variograms in urban areas show zero slope in the range of 'lag' (500m < |h| <1000m), while in vegetated areas have slower increase of semi-variance in that range. The

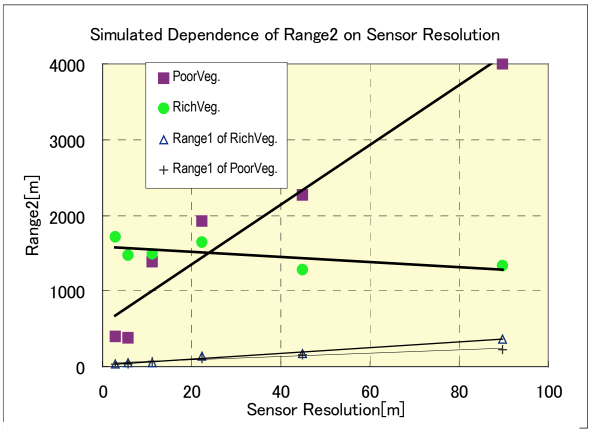

Figure 7. Simulated dependence of 'range1,2' on sensor resolution.

vegetation coverage percentile in Fig. 9 means the ratio of area in the range of NDVI(0.2~1.0) to the entire area.

3. CONCLUSIONS

The dependency of NDVI histograms on the sensor spatial resolution is assessed directly by data from QuickBird with 2.8m

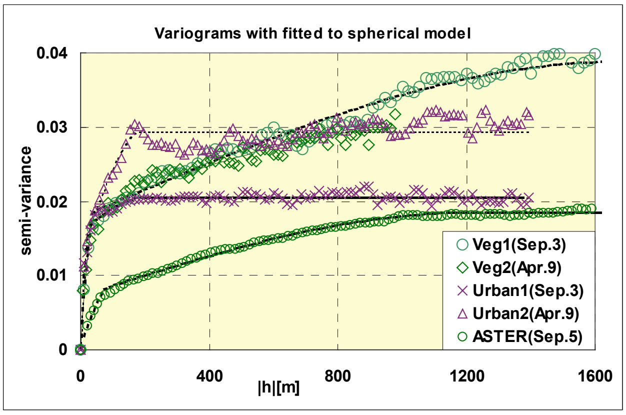

Figure 8. Variograms for vegetation and urban areas with lines fitted to nested spherical model.

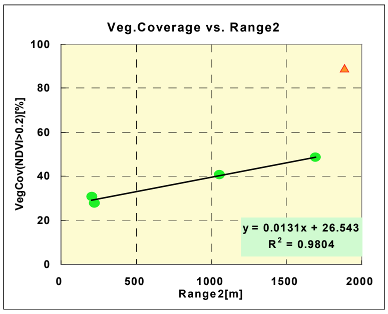

Figure9. Relationship between percentile of vegetation coverage and a 'range' of variogram . Circles and triangle are derived from QuickBird and ASTER, respectively.

resolutions, their derived data with reduced spatial resolutions and ASTER with 15m. Through the analysis of the variograms for the entire scenes and both rich and poor vegetation areas, it is manifested the linear relation between the ground spatial correlation length, "range" and the sensor spatial resolution. There exists an apparent difference in the ground spatial correlation range between areas containing much natural objects and those including many artifacts. It is interesting that there exists the relationship between percentile of vegetation coverage and the 'range' of variogram and increasing tendency of 'Range2' on the increase of NDVI for rich vegetation areas. It is also shown that the slope of variograms in the range of 'lag' between 500m and 1000m for vegetated areas would become practically useful index for describing global vegetation context of .the intrinsic spatial properties on the surface.

4. REFERNCES

Jean-Paul Donnay, Michael J.Barnsley and Paul A. Longley, *Remote Sensing and Urban Analysis*, Chapter 2, Taylor & Francis, Landon and New York, 2001.

Groten, S. M., "NDVI crop monitoring and early warning yield assessment of Burkina Faso", *International Journal of Remote Sensing*, **14(8),** 1495,1993.

S. J. Hooka, J. J. Myersb, K. J. Thomec, M. Fitzgeraldb, A. B. Kahlea, "The MODIS/ASTER airborne simulator (MASTER) - a new instrument for earth science studies," *Remote Sensing of Environment*, **76**, 93-102, 2001.

Isaaks, E.H. and R.M. Srivastava, *An Introduction to Applied Geostatistics*, Oxford University Press, New York, 1989.

Journel, A.G. and C.J.Hujibregts, *Mining Geostatistics*, Academic Press, London, 1978.

S.O. Los, C.O. Justice, C.J. Tucker, "A global 1 by 1 degree NDVI data set for climate studies derived from the GIMMS continental NDVI data", *International Journal of Remote Sensing*, **15(17),** 3493- 3518,1994.

Marceau, D., Howarth, P. and Gratton, D.,"Remote sensing and the measurement of geographical entities in a forested environment. 1. The scale and spatial aggregation problem", *Remote Sensing Environment*, **vol. 49**,93-104, 1994.

Aaron Moody and Curtis E. Woodcock, "The influence of scale and the spatial characteristics of landscapes on land-cover mapping using remote sensing", *Landscape Ecology,* **vol. 10**, no.6,363-379,1995.

Dale A. Quattrochi and Michael F.Goodchild, *Scale in Remote Sensing and GIS*, Lewis Publishers, Boca Raton, 1997

Turner, M., O'Neill, R., Gardner, R. and Milne, B., "Effects of changing spatial scale on the analysis of landscape pattern", *Landscape Ecology,* **vol. 3**,153-162,1989.

Hans Wackernagel, *Multivariate Geostatistics*, Springer, 2003.