Use of Semivariances for Studies of Landsat TM Image Textural Properties of Loblolly Pine Forests

Jarek Zawadzki[1, 4], Chris J. Cieszewski[2], Roger C. Lowe[3], and Michal Zasada[5, 6]

Abstract.—We evaluate the applicability of Landsat TM imagery for analyzing textural information of pine forest images by exploring the spatial correlation between pixels measured by semivariances and cross-semivariances calculated from transects of the Landsat TM images. Then, we explore differences in semivariances associated with images of young, middle-aged, and old, and natural versus planted stands. Finally, we compare semivariances for loblolly pine (*Pinus taeda* L.) with those of longleaf pine (*Pinus palustris* Mill.) in Georgia, U.S.A. The results show that, in spite of the low Landsat TM resolution, the semivariances and cross-semivariances may provide useful additional information.

Remotely sensed data are inexpensive supplements to ground measurements and are frequently used in forest inventories of large areas due to the cost efficiency and the ability to provide a large amount of information in a short time **(**Campbell l994, Vogelmann *et al.* 1998). Most common methods for image classification of remotely sensed images are applied without considering potentially useful spatial information among various pixels. Semivariograms consider the spatial information and have proved useful in analyzing various spatial data (Curran 1988, Woodcock *et al.* 1988a, 1988b). So far, the semivariograms have been successfully used in forestry applications only with expensive high-resolution data (St.-Onge and Cavayas 1995, Treitz and Howarth 2000).

The objective of our study was to evaluate the applicability of the relatively inexpensive, low-resolution Landsat TM7 TM imagery for analyzing the textural information in images of loblolly pine forests (*Pinus taeda* L.) in Georgia, U.S.A., using geostatistical methods. We analyzed different ages and natural versus planted stands of loblolly pine using semivariograms and cross-semivariograms. To check if semivariograms can discriminate between different species, semivariograms for loblolly pine were compared with those of longleaf pine (*Pinus palustris* Mill.).

We analyzed data from the Thematic Mapper sensor of the Landsat TM7 satellite in combination with ground measurements. We used information from the visible red (RED), the near-infrared (NIR), and the middle-infrared (MIR) bands. The Normalized Difference Vegetation Index (NDVI) as well as the corrected NDVI ($NDVI_c$) and MIR/RED indices were studied.

Area Description, Methods, and Material Studied

Study Site and Data Description

We linked remote sensing images to vegetation data by using data collected in the field. The study area was located in western Georgia, U.S.A. The data collected contained stand information including stand-polygon GIS/GPS coordinates, vegetation type (e.g., species) as well as some quantitative data (e.g., age, basal area, density). We also used data from longleaf pine stands to compare their textural characteristics with another species. We differentiated between planted and natural stands, and divided all stands of both species into three age classes: young (6–11 years), medium (16–26 years), and old (older than 31 years).

Landsat TM data are appropriate for mapping and investigating broad vegetation types classified by the sensor's spectral and spatial characteristics. The important characteristics of the Landsat TM7 satellite are:

[1,5]Postdoctoral Fellow, [2]Assistant Professor, [3]GIS Analyst, D.B. Warnell School of Forest Resources, University of Georgia, Athens, GA 30602, USA, biomat@uga.edu.
[4]Assistant Professor, Environmental Engineering Department, Warsaw University of Technology, Nowowiejska 20, 00–61 Warsaw, Poland, jarek97@yahoo.com.
[6]Assistant Professor, Department of Forest Productivity, Faculty of Forestry, Warsaw Agricultural University, Rakowiecka 26/30, 02–528 Warsaw, Poland, zasada@delta.sggw.waw.pl.

1. scene coverage–115 miles by 115 miles
2. spectral resolution–three bands in the visible portion of the spectrum, three bands in the reflective-infrared portion of the spectrum, one band in the thermal portion of the spectrum, and a panchromatic (black and white) band
3. spatial resolution–30 meters for the visible band
4. temporal resolution–16 days

We used digital numbers (DN) from the RED band (red, 0.63-0.69$_{\mu}$m), the NIR band (reflective-infrared, 0.76-0.90$_{\mu}$m), and the MIR band (mid-infrared 1.55-1.75$_{\mu}$m). The RED band is sensitive enough for discriminating between many plant species. The NIR band is especially sensitive to the amount of vegetation biomass present in a scene. The MIR band is sensitive to the amount of water in plants (ERDAS Field Guide 1990). Finally, we also studied the geostatistical characteristics of the Normalized Difference Vegetation Index (NDVI) by Rouse (1973), and the corrected $NDVI_c$ as well as the MIR to RED ratio vegetation index (MIR/RED) by Jordan (1969). The NDVI was calculated according to the following formula:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

where RED and MIR denotes the red and the near-infrared reflectance. The $NDVI_c$ vegetation index is a NDVI modified index, especially designed for distinguishing coniferous forests (Nemani *et al.* 1993). $NDVI_c$ is given:

$$NDVI_c = \frac{MIR - RED}{MIR + RED}\left(1 - \frac{NIR - NIR_{min}}{NIR_{max} - NIR_{min}}\right)$$

where the first factor in the equation is the NDVI and the second factor is a correction of the NDVI. The NIR_{min} is the reflectance value of pixels corresponding to field plots with the lowest tree canopy, and NIR_{max} is the reflectance value of pixels with the highest canopy cover.

All remotely sensed images were analyzed using ERDAS Imagine 8.5 Software.

Methods

Geostatistics comprises many methods for evaluating the autocorrelation that commonly exists in spatial data. The main tool of geostatistics is the semivariogram (semivariance), which is a measure of spatial continuity. The experimental semivariogram is derived by calculating half the average squared difference in data values for every pair of data locations along a specified direction:

$$\gamma = \frac{1}{2N}\sum_{i=1}^{N}\left[Z(x_i) - Z(x_i + h)\right]^2$$

where x_i is a data location, h is a lag vector, $Z(x_i)$ is the data value at location x_i, N is the number of data pairs spaced a distance and direction h units apart. These values are then plotted against the distances between data pairs. Such a plot is commonly referred to as a variogram and has a classic form shown in figure 1.

Semivariograms are roughly defined by three characteristics:

1. sill–the plateau that the semivariogram reaches. The sill is the amount of variation explained by the spatial structure.
2. range of the influence (correlation)–the distance at which the semivariogram reaches the sill.
3. nugget effect–the vertical discontinuity at the origin. The nugget effect is a combination of sampling error and short-scale variations that occur at a scale smaller than the closest sample spacing. The sum of the nugget effect and the sill is equal to the variance of the sample.

Figure 1.—*The "classic" form of semivariance.*

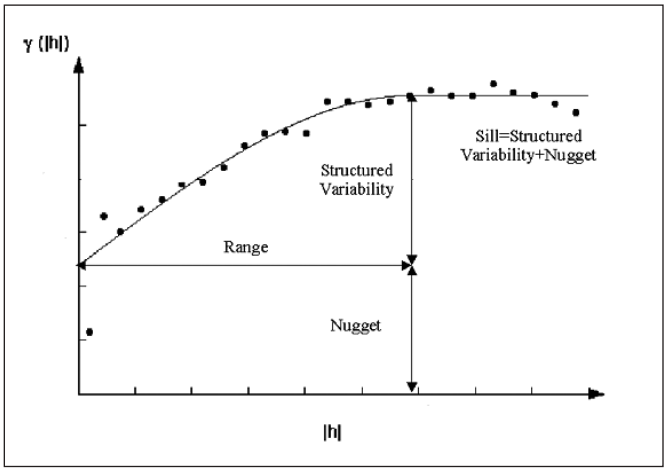
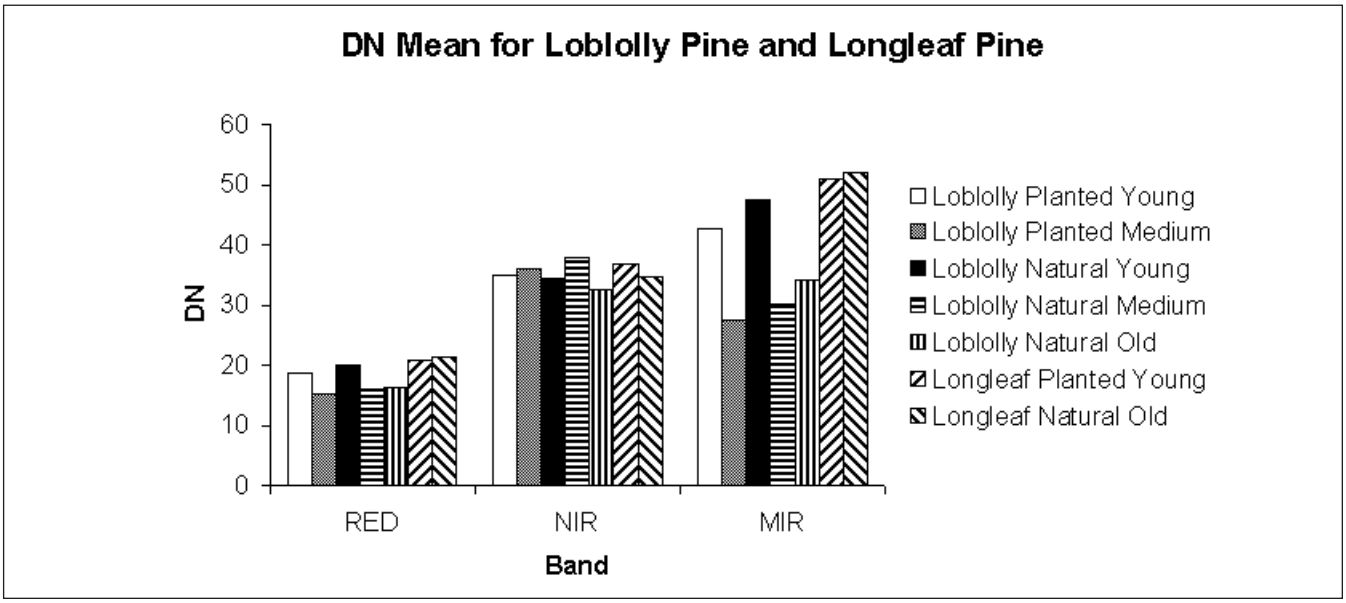


Figure 2.—*Mean of DN for loblolly pine and longleaf pine calculated from Landsat TM image, channels RED, NIR and MIR, stand origin 1988.*



The obtained experimental semivariogram is used to fit an appropriate theoretical model, as e.g., spherical, exponential, etc., and can be used in other geostatistical analyses, e.g. kriging.

Remotely sensed images can be also used in semivariogram calculations. The semivariogram is calculated from the transects running across a remotely sensed image using digital numbers as data values $Z(x_i)$.

Another important measure of spatial correlation is the cross-semivariogram:

$$\gamma_{WZ} = \frac{1}{2N}\sum_{i=1}^{N}\left[W(x_i) - W(x_i + h)\right]\left[Z(x_i) - Z(x_i + h)\right]$$

where x_i is a data location, h is a lag vector, $Z(x_i)$ and $W(x_i)$ are the DN values at location x for different bands, N is the number of data pairs spaced a distance and direction h units apart. The cross-semivariogram quantifies the joint spatial variability (cross correlation) between two radiometric bands.

Semivariograms can be a useful tool in classification, but there are some important difficulties in applying semivariograms to forest classification. First of all, often in forested areas semivariograms are much more complicated than the "classic" ones. For example, some periodic and aspatial variations of the classic semivariogram were often observed for forested areas. The first type of semivariogram appears when a repetitive pattern is studied, and the second one appears when random patterns are investigated. There were also "unbounded" forms of semivariograms observed in the study. The unbounded semivariogram may represent a situation in which a trend or many spatially correlated phenomena exist. These nonclassic semivariograms are much more difficult to model and interpret.

Results and Analyses

The basic descriptive statistics of the analyzed forest types (fig. 2) reveal some distinctions between the different stands but do not provide any textural information. To explore the textural continuity of studied stands, we calculated and analyzed the semivariograms for RED, MIR, and NIR bands, as well as the cross-semivariograms between these bands. The semivariograms for the above mentioned vegetation indices were also calculated.

To understand better the factors that influence the semivariograms, we calculated them in large and potentially homogeneous areas, changing for comparison only one essential stand feature, e.g., age (young, medium, old) or type of stand (planted, natural). Figure 3 shows typical, standardized (divided by theirs variances)

Figure 3.—*Isotropic semivariances for a planted loblolly pine stand calculated from Landsat TM images; stand origin 1988.*

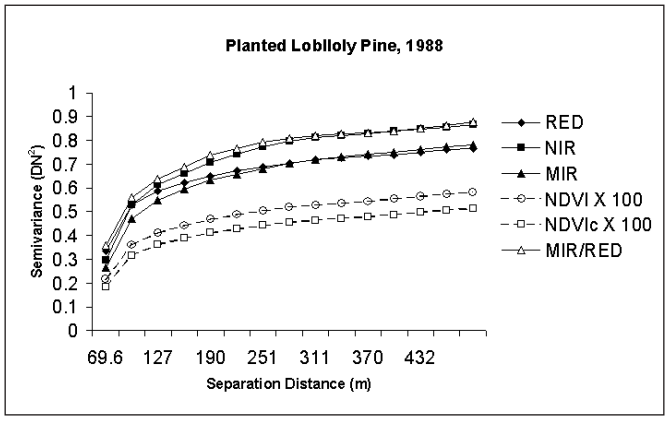
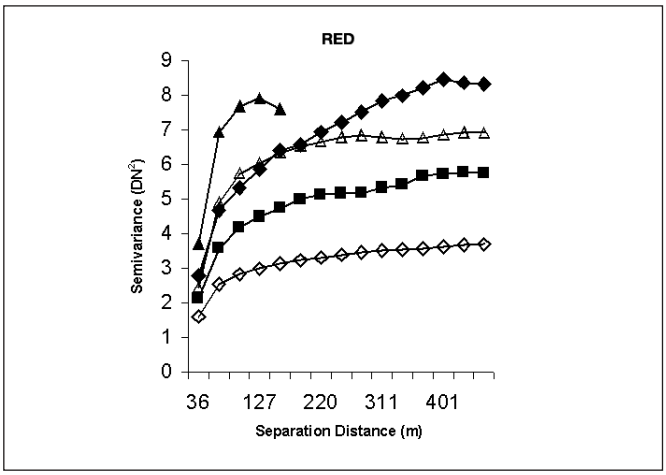


Figure 4.—*Isotropic semivariances for different types of loblolly pine stands, calculated from Landsat TM image using RED band.*



semivariograms, calculated for 12-year-old planted loblolly pine stands. These semivariograms have a typical "unbounded" shape showing many spatially correlated phenomena.

For the small separation distances (a few lags) the semivariogram curve rises relatively fast. Then, at the greater distances it exhibits a gentle sloping and becomes almost linear. The initial increase of the semivariogram curve results from the fast decrease of spatial continuity at the distances of a few lags (1 lag = 30 meters). This means that the spatial correlations between pixels decreases rapidly for short distances.

At longer distances, the semivariograms do not reach saturation but increase almost linearly. This means that many sizes and shapes of the forest stands are present in the scene. As was already mentioned, these semivariograms are difficult to use for classification purposes. For example, it is clear that the range and the sill cannot be distinctive parameters for different vegetation communities (fig. 3).

To check whether the semivariograms can be treated as "spatial signatures" of different type of coniferous forests, we calculated them for different types of loblolly pine stands. We calculated semivariograms using the DNs from RED, NIR, and MIR bands as well as NDVI and MIR/RED indices. The largest differences between semivariograms calculated for the investigated loblolly pine stands were obtained from the RED and MIR bands. The results of the calculations for the RED band are shown in figure 4.

Distinctly smaller differences were observed between semivariograms calculated for DN from the NIR band as well as between semivariograms calculated from the vegetation indices NDVI and MIR/RED. This somewhat surprising behavior of semivariograms from vegetation indices can be explained by the smoothing effect; these indices are the ratios of DNs coming from different bands.

Natural stands have higher semivariogram values than even-aged planted stands (fig. 4). This is because natural stands' have a higher textural variability than planted stands.

We also compared semivariograms for different species of pine by calculating semivariograms for planted and natural stands of longleaf pine. The exemplary semivariograms of loblolly pine and longleaf pine calculated from the DN for the MIR band are shown in figure 5.

Large differences exist between semivariograms calculated from loblolly pine and longleaf pine stands. The semivariogram values for longleaf pine are much higher than those of loblolly pine, calculated for the stands of similar type and age. The values of semivariograms at the distance of a few lags can be also used as a discriminative parameter.

The cross-semivariograms quantify the joint spatial variability between two bands. Therefore, they can be also used for texture-based classification adding new spatial information. So, at the end of our analysis we calculated also cross-semivariograms between bands RED, MIR, and NIR for planted and natural, medium-aged stands of loblolly pine. The largest cross-correlations were between the RED and MIR bands both for the planted and for natural stands. The cross-correlations between bands RED and NIR as well as between MIR and NIR were substantially smaller. The values of the cross-semivariogram for natural stands were much higher than for planted stands. The cross-cor-

relations between the RED and MIR bands calculated for studied loblolly pine stands are shown in figure 6. Clearly, all age classes are well separated. The largest cross-semivariogram values were obtained for young stands and the smallest for old stands, both for planted and natural stands.

Conclusions

In spite of the low-resolution of the remote imagery, distinct differences were found in semivariograms of images for the studied forests. This means that such semivariograms can be treated as "spatial signatures" for the studied forest stands. The classical semivariogram's parameters, such as range and sill, are not appropriate as differentiated parameters because of the low-resolution of the remote imagery and the nonclassic, unbounded type of observed semivariograms. However, there are important differences for semivariogram and cross-semivariogram values at the distances of several lags. Our study suggests that the semivariogram values for such separation distances (e.g., from the 4th to the 7th lags) are appropriate for these purposes. The observed differences between semivariograms at distances of several lags arise from different spatial correlations existing in the studied forest stands at distances of a few tens to a few hundred meters. The low-resolution of Landsat TM7 remote imagery does not allow distinguishing separate trees. The observed spatial correlations can be attributed to the similarity in arrangements of bigger objects as groups of trees (or stands), areas with similar underbrush, etc. The largest differences in semivariograms were obtained for RED and MIR bands.

Figure 5.—*Semivariograms calculated from remote images of loblolly pine and longleaf pine stands using MIR band.*

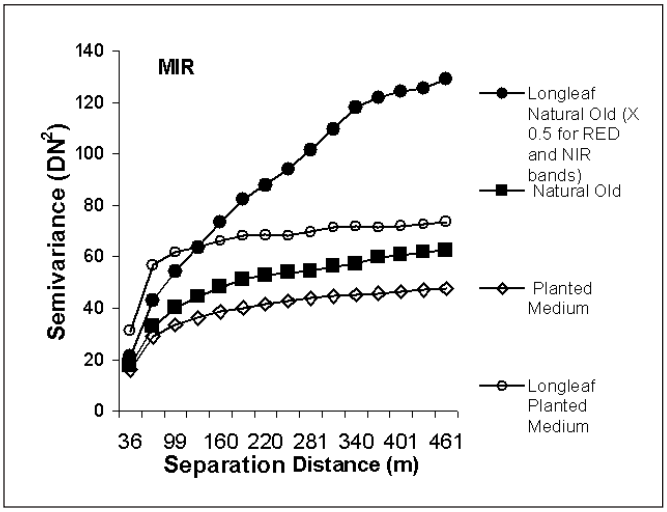
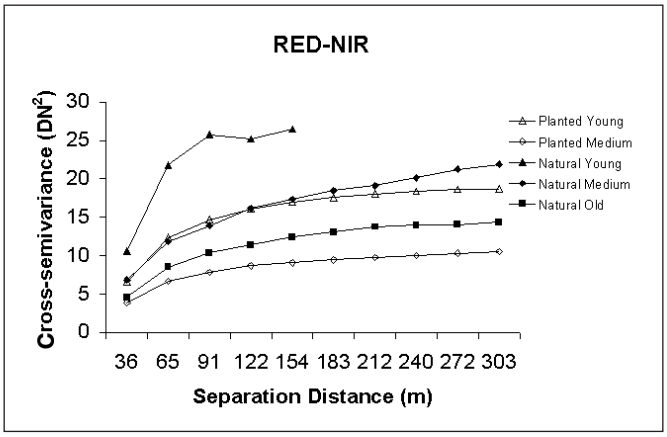


Figure 6.—*The cross-correlations between the RED and NIR bands calculated for studied loblolly pine stands.*



Acknowledgments

This study was sponsored by the School of Forest Resources, University of Georgia, and the Traditional Industry Program 3 (TIP3) funding. We are much indebted for this support. Jaroslaw Zawadzki thanks the Warnell School of Forest Resources, University of Georgia, for a creative and hospitable environment; Ingvar Elle for help with English; and Tripp Lowe for technical assistance.

Literature Cited

Campbell, J.B. 1994. Introduction to remote sensing. New York, NY: Guilford Press.

Curran, P.J. 1988. The semivariances in remote sensing: an introduction. Remote Sensing of Environment, 24: 493–507.

ERDAS, Inc. 1990. ERDAS field guide, 5th ed. 1990. Atlanta, GA: ERDAS, Inc.

Jordan, C.F. 1969. Derivation of leaf area index from quality of light on the forest floor. Ecology. 50: 663–666.

Nemani, R.; Pierce, L.; Running, S. 1993. Forest ecosystem process at the watershed scale: Sensitivity to remotely-sensed leaf area index estimates. International Journal of Remote Sensing. 14: 2519–2539.

Rouse, J.W.; Hass. R.H.; Schell, J.A.; Deering, D.W. 1973. Monitoring vegetation systems in the great plains with ERTS. In: Proceedings, 3d ERTS symposium; [Dates held unknown]; [where held unknown]. NASA SP–351. Washington, DC; NASA: 309–317.

St.-Onge, B.A.; Cavayas, F. 1995. Estimating forest stand structure from high-resolution imagery using the directional semivariogram. International Journal of Remote Sensing. 16: 1999–2001.

Treitz, P.; Howarth, P. 2000. High spatial remote sensing data for forest ecosystem classification: an examination of spatial scale. Remote Sensing of Environment. 72: 268–289.

Vogelmann, J.E.; Sohl, T.L.; Campbell, P.V.; Shaw, D.M. 1998. Regional land cover characterization using Landsat Thematic Mapper data and ancillary data sources. Environmental Monitoring and Assessment. 51: 415–428.

Woodcock, C.E.; Strahler, A.H.; Jupp, D.B. 1988a. The use of semivariograms in remote sensing: I: scene models and simulated images. Remote Sensing of Environment. 25: 323–348.

Woodcock, C.E.; Strahler, A.H.; Jupp, D.B. 1988b. The use of semivariograms in remote sensing: II: real digital images. Remote Sensing of Environment. 25: 349–379.