Characterization of Flood Plains Using Remotely Sensed Data

Progress Report on Year 3 of Project

James F. Cruise, Mohammad Al-Hamdan, and Fuad Musleh
Department of Civil and Environmental Engineering
University of Alabama in Huntsville

Introduction

The purpose of this component of the project was to determine the utility of remotely sensed data in characterizing the surface roughness characteristics of forested flood plain areas that might lie in highway corridors or be the object of highway crossings.

Flood Plain Roughness Characterization

Objective: To develop a methodology for determining roughness coefficients for flood plain corridors from remotely sensed data. The project includes both an experimental component and a remote sensing component. Specifically, the project seeks to:

- Determine relationships between landscape parameters amenable to measurement from remotely sensed data and corresponding roughness measures through laboratory experiments
- Develop landscape characterization techniques using remotely sensed data to estimate the parameters found to be relevant in the experimental analysis.

Background

The discharge capacity of flood plain cross sections can be expressed through the Manning equation:

$$Q = \frac{1.49}{n} A y^{2/3} S^{1/2}$$

where: Q = discharge (ft^3/s); A = cross sectional area (ft^2); y = flow depth (ft); S = stream slope (ft/ft); and n = roughness coefficient.

The retarding effects of the resistance to flow provided by landscape features such as trees, grass, boulders, etc. are encapsulated within the roughness coefficient, known as "Manning's *n*". This coefficient must be estimated from empirical sources for each area for which corridor crossings are to be designed. Therefore, the purpose of this project was to develop a methodology for providing guidance for the estimation of Manning *n* values from remotely sensed data sources.

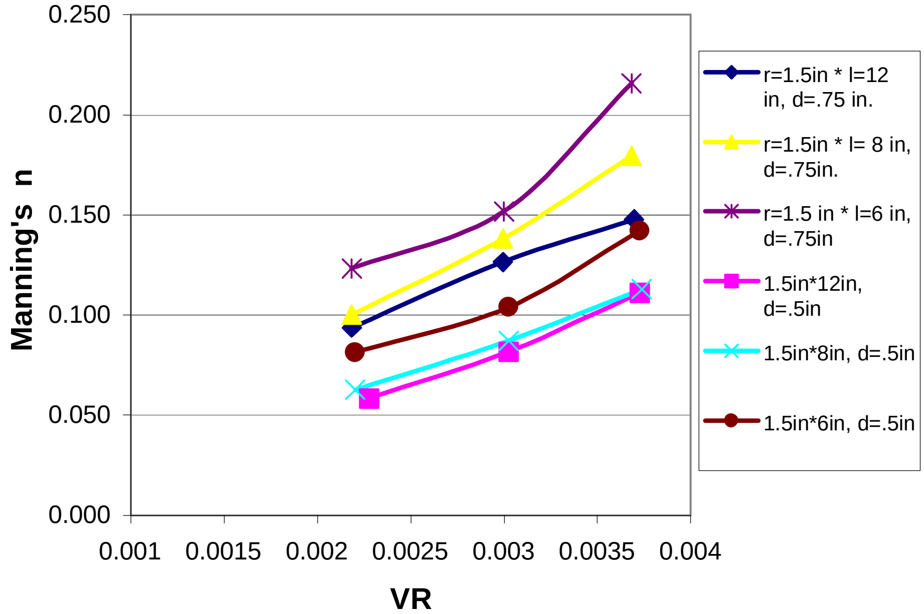
Experimental Component

The purpose of the experimental component of the project was to determine the factors that are important in the estimation of Manning's *n* values. In order to accomplish this a wide, shallow flume was constructed to represent a typical flood plain corridor. The flume dimensions are 4 ft wide by 1 ft high and 12 ft long (Figure 1). The roughness elements are represented by the rods mounted vertically from the cross members as shown in the figure. It is possible to use 6 longitudinal rows of rods, each with up to 30 rods attached, to represent a complete roughness field. The diameter and spacing of the rods on the member can be altered and the cross members can be set at various distances to change the longitudinal spacing of the elements. Twelve combinations of rod diameters and spacing were tested using three different discharge values for each configuration. The depth in the flume varied from 0.67 in (17 mm) to 1.6 in (41 mm), thus fulfilling the requirement that the flow width be much larger than the depth for wide, shallow flood plain crossings.

The results were analyzed in various ways. One traditional analysis technique is to plot the computed *n* values as a function of the product of the flow velocity and depth. A sample plot is shown in Figure 2. The first thing to note about this figure, is that the Manning *n* values are directly proportional to the flow parameters, which is the opposite of the relationship encountered in traditional analyses. Based on classical resistance theory, Manning's n has historically been considered to be inversely related to the depth-velocity product due to the fact that the relative roughness (y/k, k= height of roughness

Figure 1. Experimental Apparatus for Roughness Tests

Figure 2. Manning's n related to VR



element) decreases as y increases. However, in the case of rigid, non-submerged vegetation (such as trees) in the flow field, the roughness coefficient was found to most affected by the depth of flow. This can also be illustrated through the plots of n versus the dimensionless Froude number ($F = \frac{V}{\sqrt{gy}}$) as shown in Figure 3. Again, one can see the direct relationship between n and y for as y decreases, F increases and thus, the Manning n value also decreases with the decreasing depths. In fact, the research demonstrated that the most significant factor determining the flow resistance is the total cross sectional area of the obstructions that is available to absorb the momentum of the incoming flow. The most efficient way to express this relationship is to consider the relationship between the Manning n and the ratio of the cross sectional area of the obstructions to the total cross sectional area of the section (A/a), where A is the summation of the cross sectional areas of the rigid bodies (trees) and a is the total cross sectional area of the section. This relationship is plotted in Figure 4. It is important to note that the cross sectional area of any individual obstruction is given as yd, where d is the diameter of the body. Thus the direct relationship with the flow depth is made evident.

The derived relationship also shows that the most important vegetation property is the diameter of the obstructions (d). Thus, the ratio (A/a) would be the average stand thickness of vegetation in the flood plain. Past research has shown that this parameter can be effectively estimated from imaging radar (SAR) data, thus opening a strong possibility that Manning roughness coefficients can be estimated from remotely sensed data with an acceptable degree of confidence.

Figure 3. Manning's n Vs Froude Number

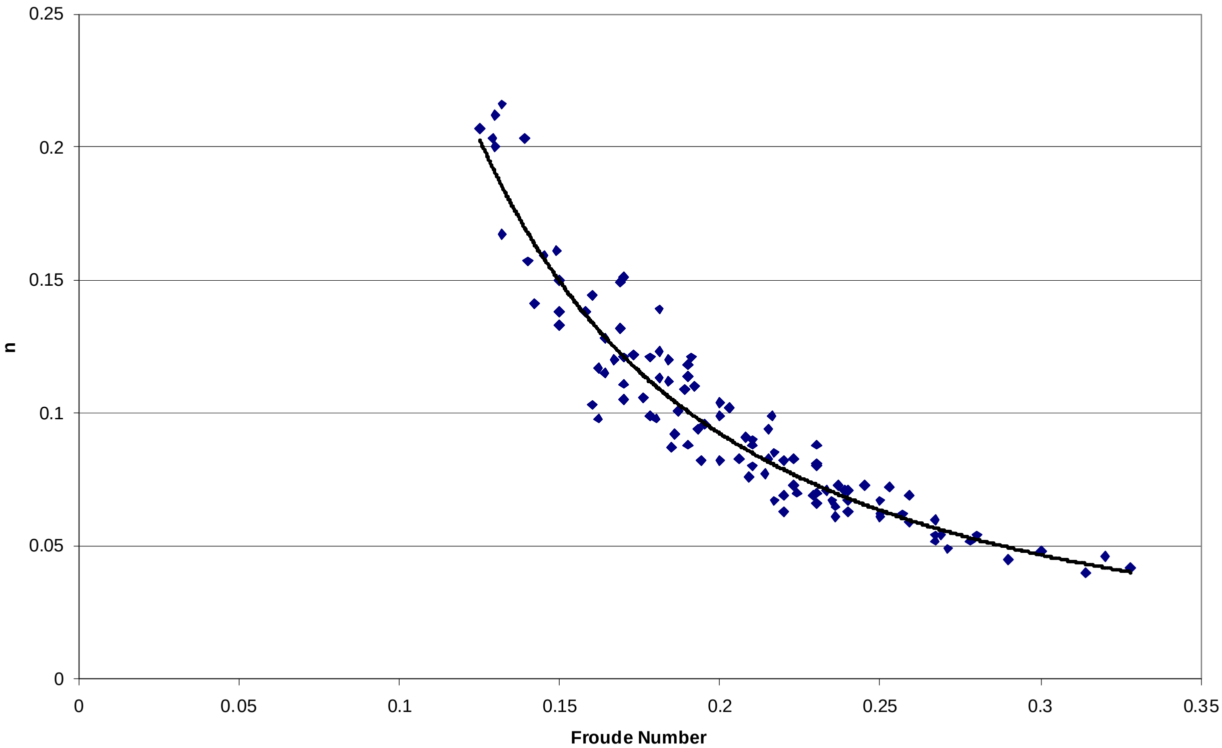
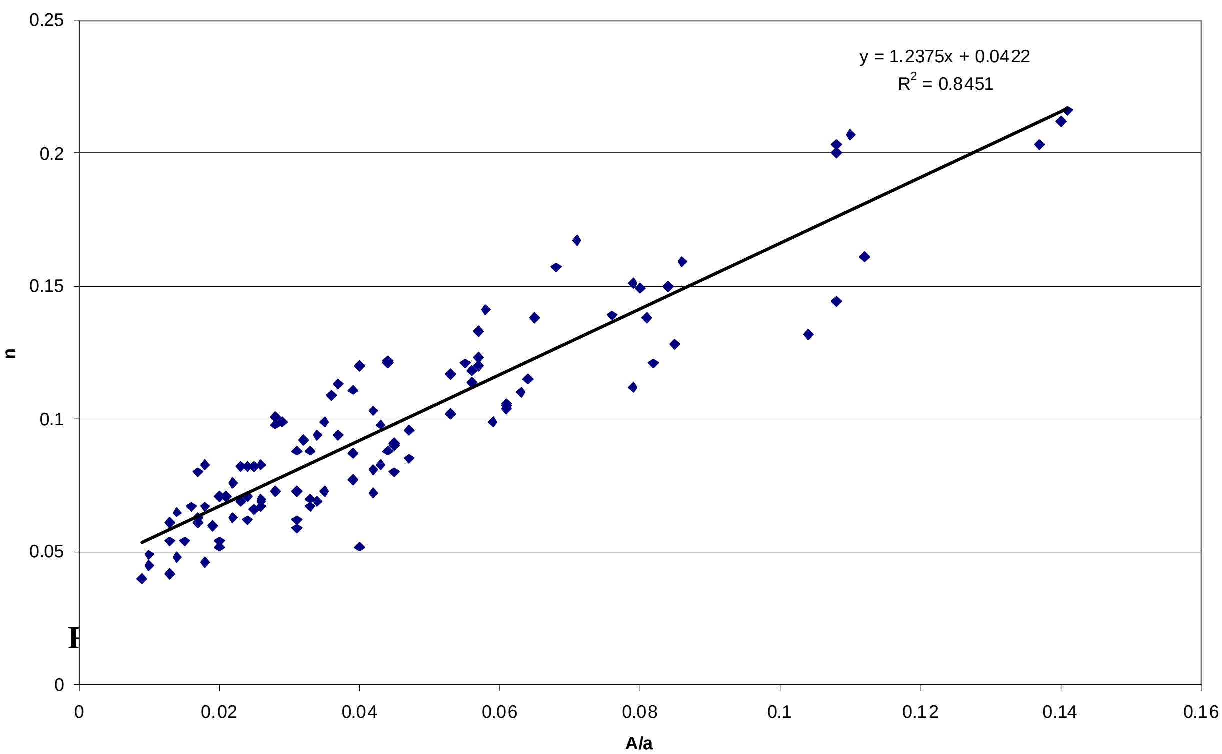


Figure 4. Manning n vs A/a



Remote Sensing Component

The purpose of this component of the project was to determine whether or not important measures of surface roughness (including average stand thickness) can be effectively estimated from remotely sensed data. Because passive (radiometric) data (e.g., Landsat TM) are more accessible and cost effective than SAR data, radiometric sources are utilized in this study. The study is designed to answer several questions:

- What level of detail about surface roughness can be obtained from radiometric remote sensing data? (i.e., can stand thickness be estimated?)
- Can flood plain areas be effectively identified and delineated from passive RS data?
- What is the effect of spatial and spectral resolution of RS data on the estimates of indices used in determination of surface roughness and complexity?

In order to answer these questions, several RS images of different spatial and spectral resolution were obtained. Several of these images cover areas with known forest stand characteristics (i.e., national forests), while others cover specific flood plain areas of interest. Spatial analysis techniques were used to characterize these images in terms of image complexity and roughness. Two statistical indices were employed: fractals and Moran's I. Fractals are measures of the self- similarity of a landscape and thus ultimately measure the degree of complexity of the surface. The fractal dimension can be visualized as related to the relationship between the total length of a surface feature and the step size used to measure that length. A linear relationship in log space is hypothesized; thus:

$$\text{Log}(L) = C + B\log(S)$$

Where, L = feature length; S = step size, and C and B are coefficients. Then the fractal dimension, D, is simply given by D= 1-B. Although theoretically D could vary from 1 to 3, in practice the normal range of variation when employing RS data is 2.5 for a homogeneous surface (water body for example) to 3 for a very heterogeneous surface.

Moran's I is simply a classical measure of spatial correlation, i.e.,

$$I(d) = \frac{\sum_i^n \sum_j^n w_{ij} z_i z_j}{W \sum_i^n z_i^2}$$

Where, w_{ij} = weight at distance d so that $w_{ij} = 1$ if point j is within distance d of point i, and $w_{ij} = 0$ otherwise; z_i and z_j are deviations from the mean (i.e. $z_i = x_i - x_{mean}$); and W is the sum of all the weights where $i \neq j$.

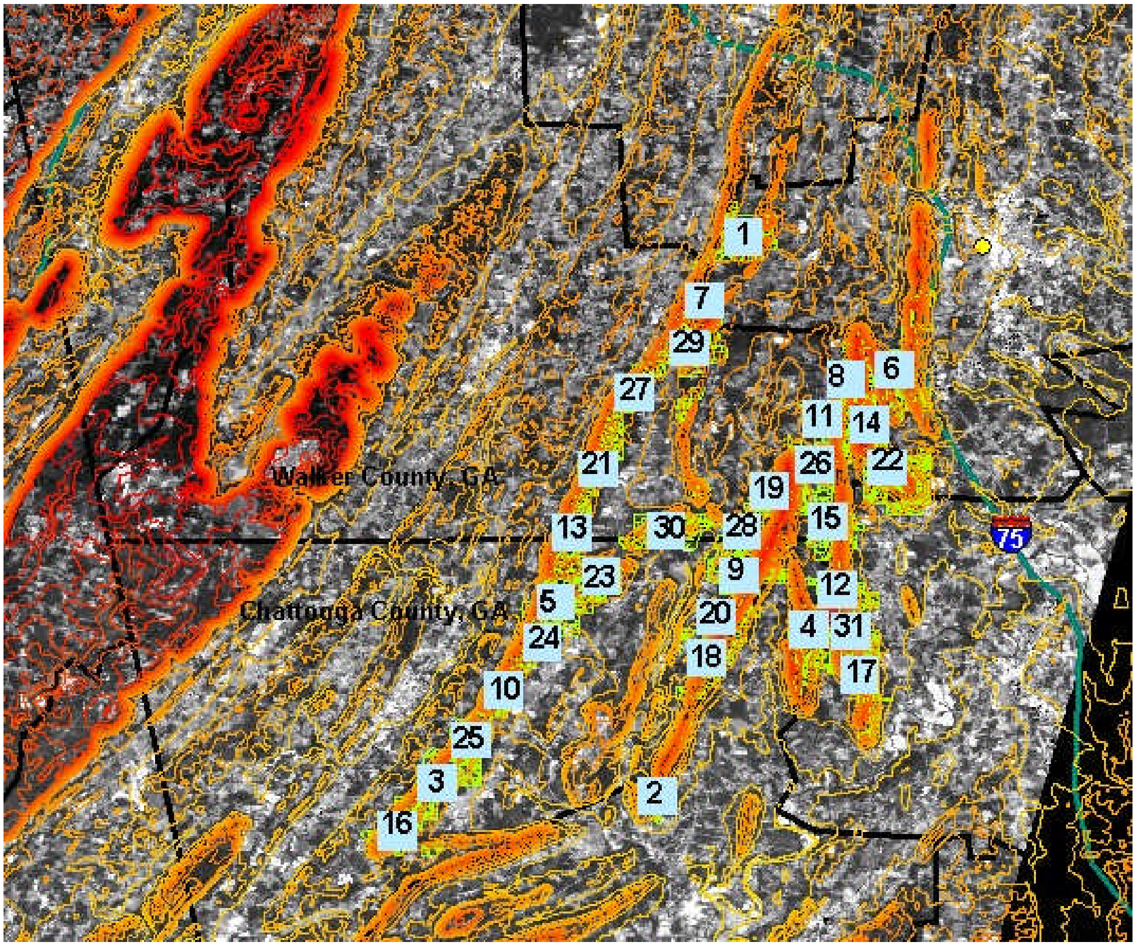
In order to examine issue 1 above, *i.e.*, the degree of detail relative to landscape roughness that can be determined from RS data, Landsat TM images were obtained that covered four national forest areas wherein the forest stand characteristics (trunk size, distribution, species, *etc*) are known with a good degree of precision and spatial detail. Topographic data were also obtained from USGS DEM data sets in order that it might also be used in the analysis if appropriate. The data were obtained for Chattahoochee National Forest (GA), Talladega National Forest (AL), Oakmulgee National Forest (AL), Bankhead National Forest (AL). A data set was constructed in Arc-View that consisted of the spatial distribution of species attributes, digital elevation files and the Landsat scenes. The satellite data are composed of leaf-on scenes since forest canopies reflect energy more efficiently than do bare tree stomata. A review of the forestry literature revealed that the relationship between tree canopy characteristics and trunk size has been extensively studied and quantified. Not surprisingly, these studies have shown that canopy size and trunk diameter are strongly related and thus it should be possible to indirectly estimate trunk diameters from observations of the canopies.

The objective of this research was to determine if forest areas with specific growth parameters exhibit a consistent pattern in the spatial complexity indices given above. In other words, can the spatial statistics indices be used to distinguish growth classes in forest environments? Samples were collected from each forest area making sure to obtain equal coverage of all parts of the forest. A total of at least 150 samples were collected from the forests for statistical analysis. The average stand characteristics and distributions of each forest are summarized below along with the average value of each spatial index for each forest. The average was computed as the average of all seven bands from the TM image for the entire forest. These data are given for illustrative purposes only.

	Growth Stage (*in*)							
	Saplings 1-5	**Poletimber** 5-9	**Sawtimber** >9	Hard	Soft	FD	I	C
Oakmulgee (%)	26	6	68	62.7	30.5	2.88	.82	.18
Talladega (%)	15	6	79	41.3	57.5	2.85	.82	.17
Chattahoochee (%)	25	7	68	73.5	15.3	2.88	.76	.23
Bankhead (%)	29	15	56	72.5	15.7	2.88	.89	.11

32 samples were collected from the Bankhead National Forest, 52 samples were taken from the Oakmulgee National Forest, 36 samples were collected from the Talladega National Forest and 31 samples have been taken from Chattahoochee National Forest (as shown in Figure 5). The results appear to demonstrate that the growth classes given above can be distinguished using the spatial statistics discussed previously. Two way analysis of variance tests showed that the both the fractal dimension and Moran's I were able to differentiate the growth classes at a significance level of less that .01 in all cases.

Figure 5. Sampling Locations in Chattahoochee National Forest



The ability to predict within the classes also shows promise although there is quite a bit of variation in the strength of the relationships among the different forests and among the classes. Linear regression analyses were run on the FD vs class % and I vs class % for all the forests and individual classes. Comparisons of the results of the regression on each class for the four National Forests are given for the fractal dimension in Figures 6-8 and for the Moran's I in Figures 9-11. The figures illustrate that the FD and I vs class relationships appear to be fairly strong for the Oakmulgee, Bankhead, and Chattahoochee National Forests in the cases of the Sawtimber and Saplings classes. However, only in the Oakmulgee case is the relationship strong for the Poletimber class and these results are biased due to the number of samples from this forest that had no trees in this class. In addition, for the Talladega Forest, none of the data sets yield an acceptable result in terms of fit to the assumed linear relationship.

Additional analyses using the topographic data revealed that the Talladega Forest samples exhibited significantly greater differences in topographic relief than did the other forests. In fact, the standard deviation of the sample elevations was over three times as great for the Talladega data than for any other forest. The Talladega samples exhibited a sample standard deviation in the elevation data of 94.7 m (N=36) compared to values of 22.7 m for Oakmulgee (N=52), 23.3 m for Bankhead (N=32), and 29.9 m for Chattahochee (N=31). Thus it appears that the landscape slope may have distorted the signal from the Talladega canopy and masked the true relationship between the indices and the forest stand characteristics. Regression of the FD against elevation revealed a significant relationship for the Talladega samples as shown in Figure 12 while the relationship was not significantly strong for the other forests.

Figure 6. Comparison of Relationship for FD and Sawtimber %

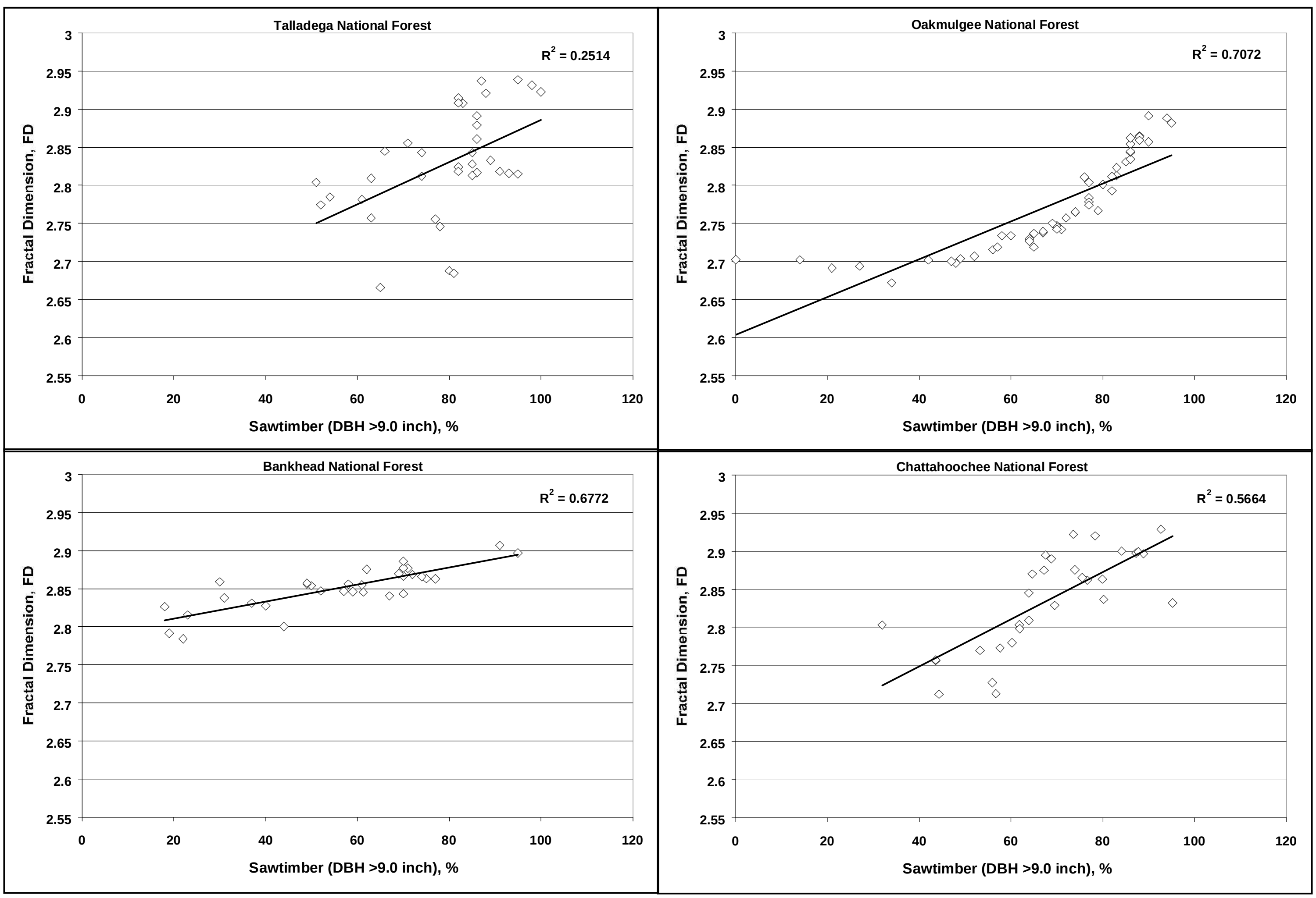


Figure 7. Comparison of Relationship for FD and Poletimber %

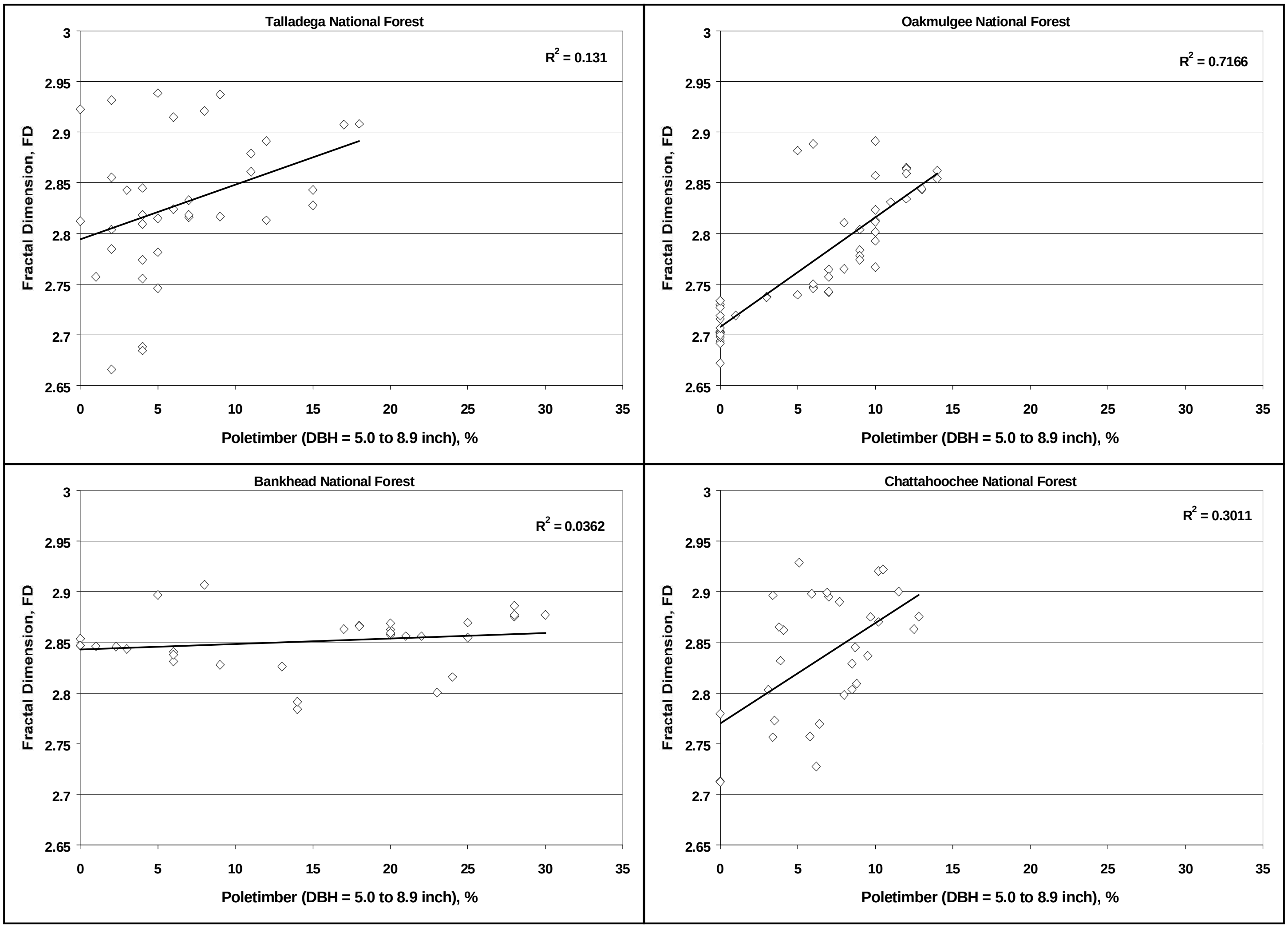


Figure 8. Comparison of Relationship for FD and Saplings %

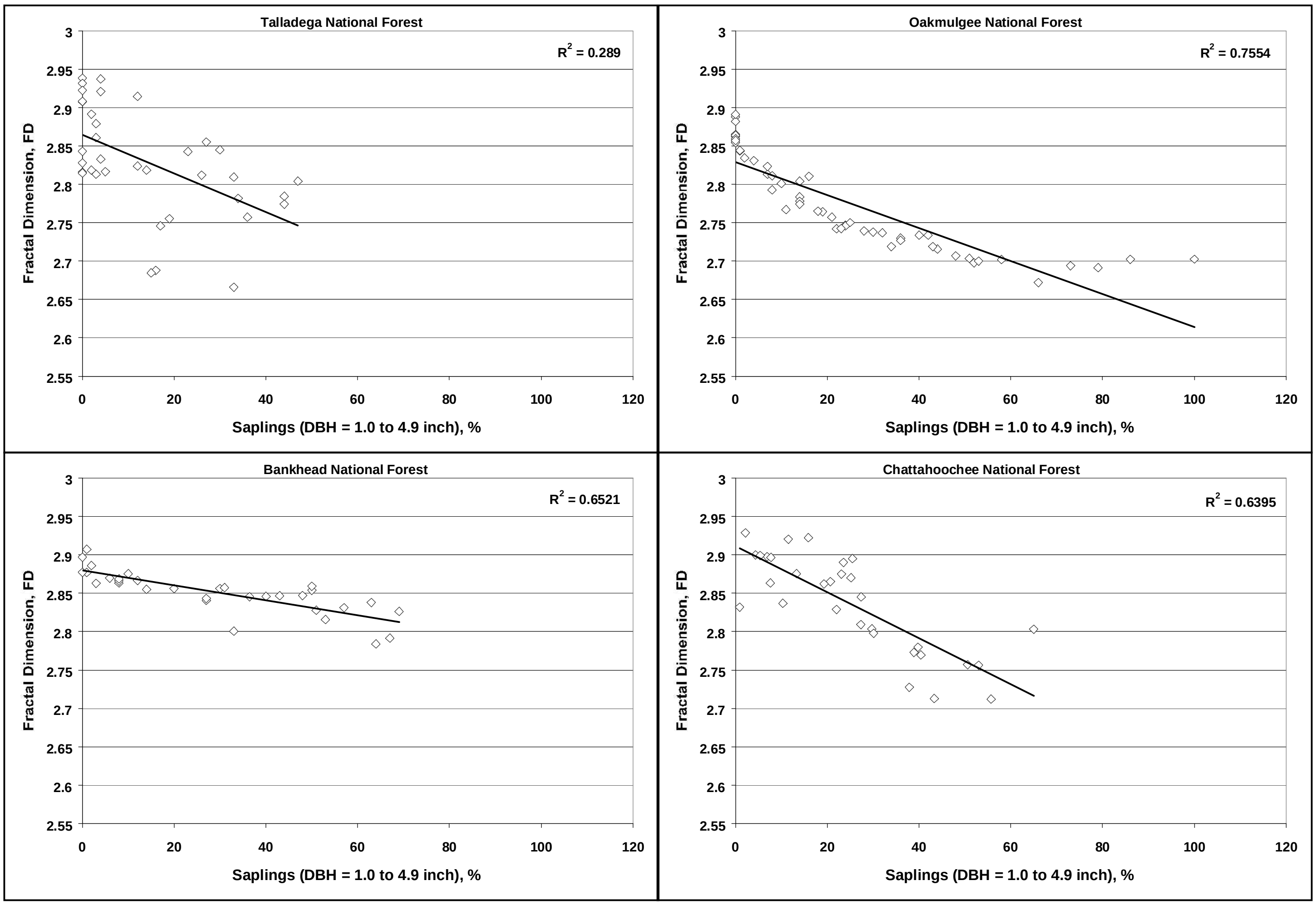


Figure 9. **Comparison of Relationship for Moran's I and Sawtimber %**

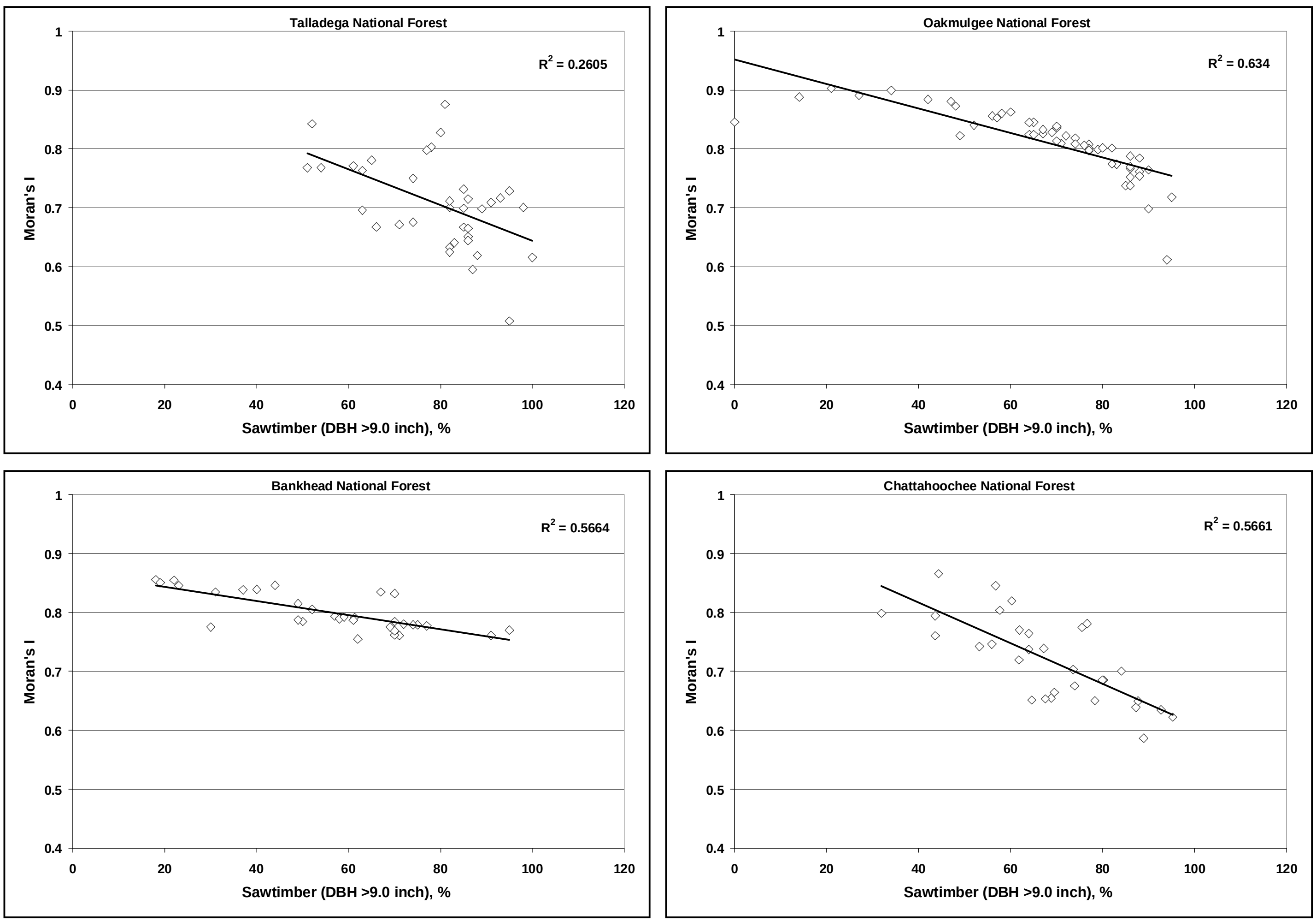


Figure 10. Comparison of Relationship for Moran's I and Poletimber %

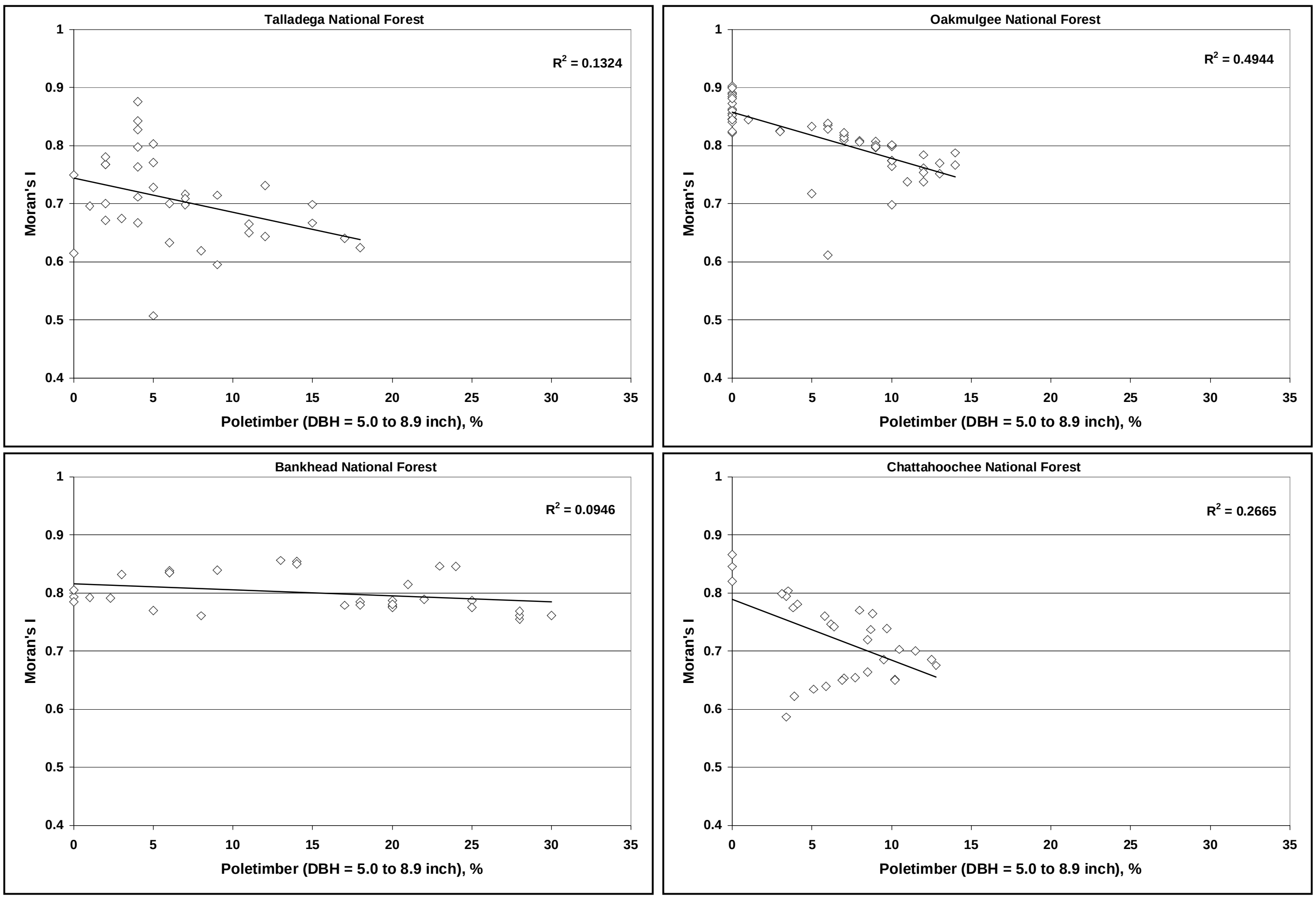


Figure 11. **Comparison of Relationship for Moran's I and Saplings %**

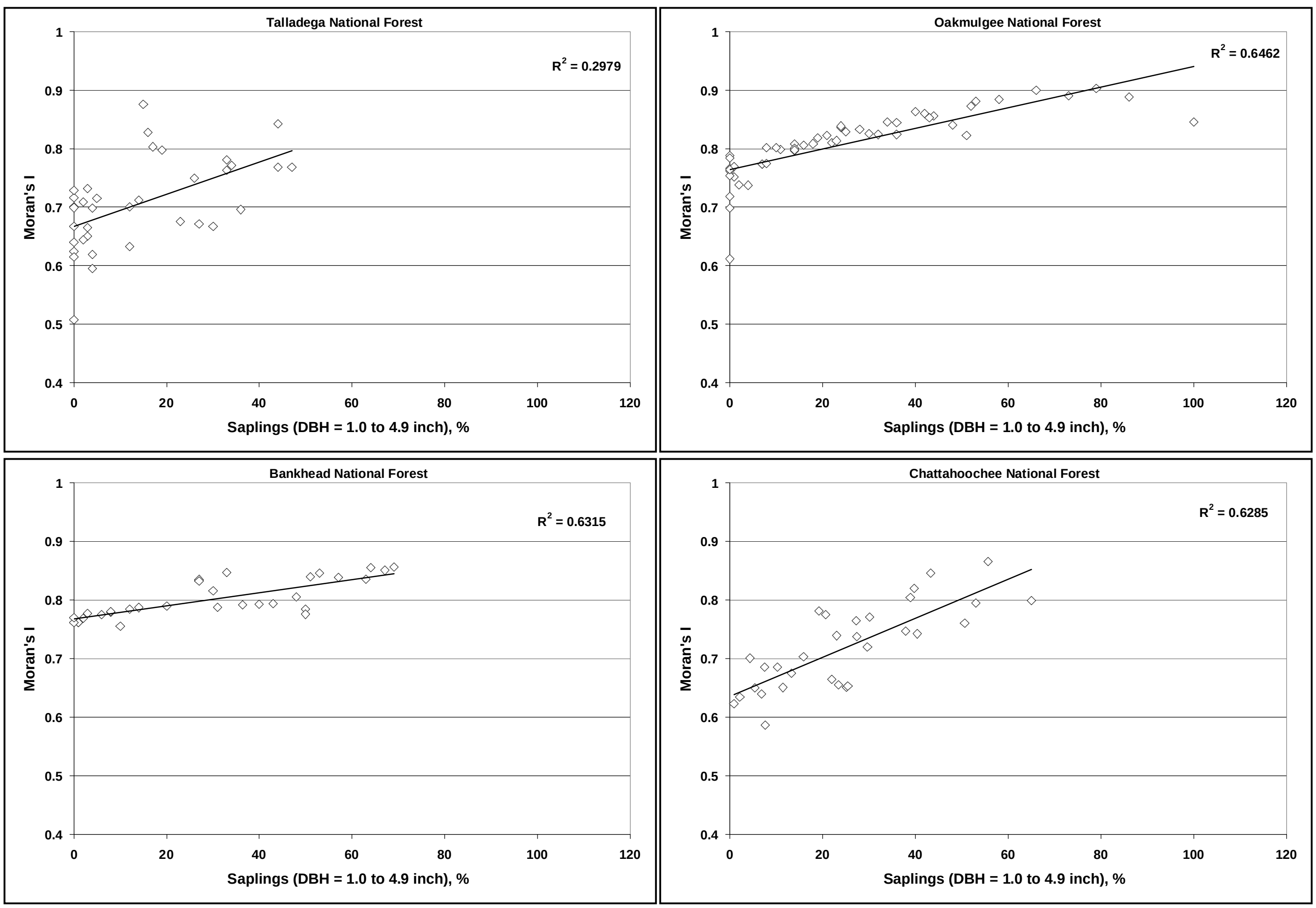
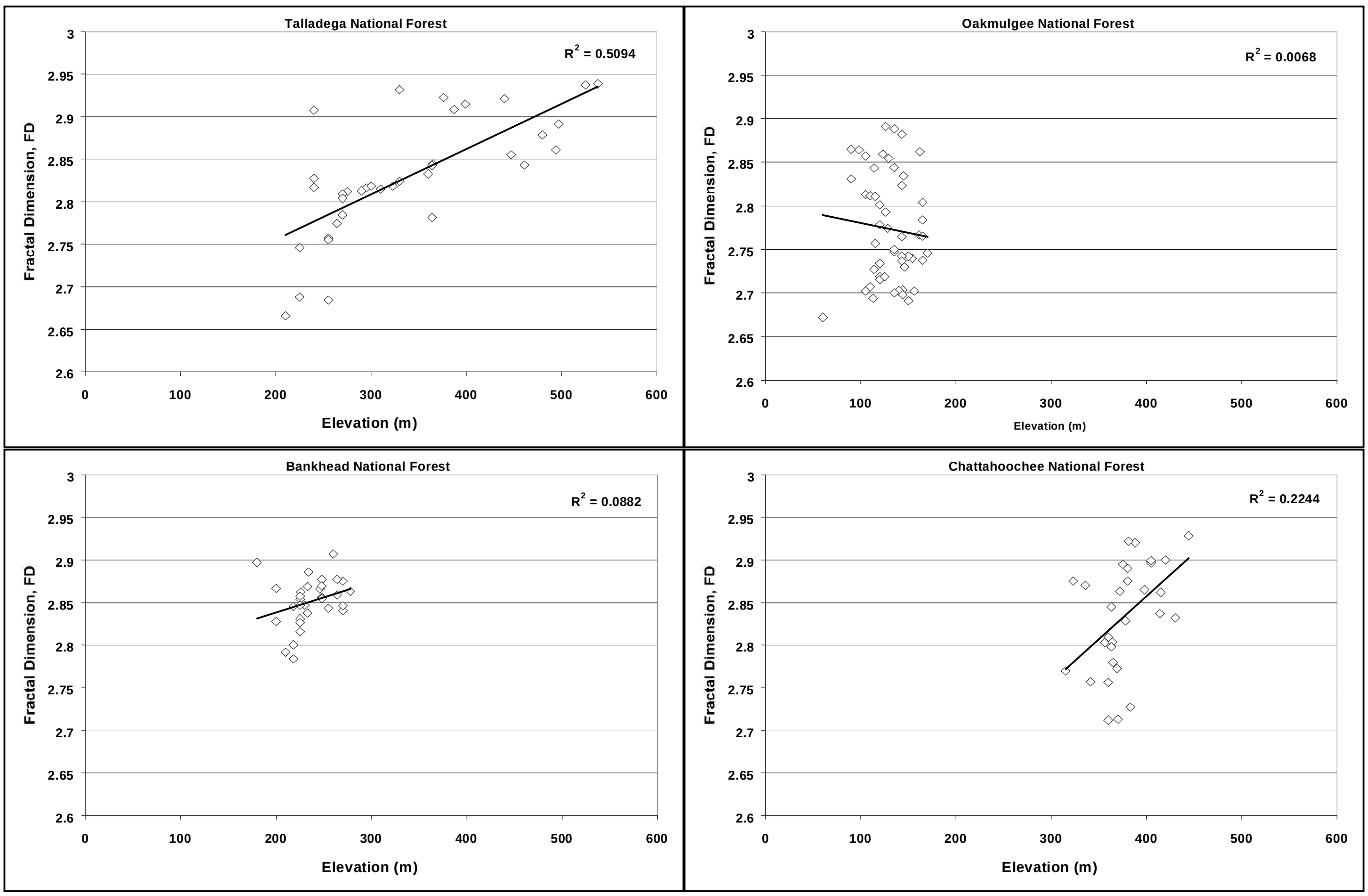


Figure 12. Comparison of Relationship for FD and Mean Sample Elevation



Analysis of the results given above appear to lead to the general conclusion that fairly strong linear relationships exist between FD (and I) and class size for the sawtimber and saplings classes so that prediction might be possible in these cases. However, in general, there does not appear to be a sufficient linear relationship to develop a prediction model in the case of poletimber. In addition, the addition of elevation as an independent variable in the analysis might be expected to strengthen the prediction capability in the case of the Talladega forest, but probably not in the other cases.

Prediction of Flood Plain Forest Characteristics

A model to predict forest characteristics of vegetated flood plains was developed using three of the forest data sets in this study while keeping the fourth data set in reserve for verification purposes. The model was developed to predict the relative percentages of the sawtimber (d>9 in) and saplings (d<5 in) classes. The percent of poletimber would then be given by 100 – (sawtimber %+saplings %). The data for Talladega, Oakmulgee, and Bankhead National forests were combined for model development purposes. Two way ANOVA tests were conducted to determine if these data sets could be assumed to arise from the same population and thus can be combined. The tests were significant at the 0.05 level. Since the analysis of the individual data sets had revealed that elevation was a significant explanatory variable in the spatial data for at least one forest, multivariable regression models were developed using the combined data from the three forests with FD and Elevation as independent variables in one model and with I and Elevation as the independent variables in the other. In both models, the percent of the data in each of the two forest stand classes was the dependent variable. The pertinent data for the models are given in the table below.

Stand Size Prediction Model*	R^2	Adjusted R^2
$Sawtimber(\%) = -340.51 + 145.45 * FD$	0.224	0.217
$Sawtimber(\%) = -458.82 + 189.84 * FD - 0.0449 * Elevation$	0.224	0.211
$Saplings(\%) = 593.02 - 202.73 * FD$	0.360	0.354
$Saplings(\%) = 746.87 - 260.91 * FD + 0.0587 * Elevation$	0.366	0.355
$Sawtimber(\%) = 209.48 - 181.88 * I$	0.423	0.418
$Sawtimber(\%) = 297.75 - 270.35 * I - 0.0886 * Elevation$	0.525	0.516
$Saplings(\%) = -131.6696 + 199.51 * I$	0.421	0.416
$Saplings(\%) = -219.08 + 287.13 * I + 0.0878 * Elevation$	0.504	0.495

* Sawtimber: diameter at breast height (DBH) > 9 inch, Poletimber: DBH = 5-8.9 inch, and Saplings: DBH = 1-4.9 inch

Some interesting observations can be made with respect to these results. First, the regression model using Moran's I as the chief independent variable appears to be significantly stronger (in terms of R^2) overall than does the model that utilized the FD as the chief independent variable. Another interesting observation is that the addition of elevation as an explanatory variable appears to have very little value in the fractal based model, while it adds substantially to the Moran's I model. Overall one might conclude that Moran's I is a more robust indicator of forest growth characteristics than is the fractal dimension as the R^2 values were generally higher for that model.

The models were then applied to the Chattahoochee data set as a test case. The statistics of these results are shown in the following table:

Class	**Mean Error (%)**		**MAE (%)**		**RMSE (%)**		**N-S R^2**	
	FD	I	FD	I	FD	I	FD	I
Sawtimber	3	2	8	11	10	13	0.25	0.22
Saplings	-4	-5	8	11	11	14	0.50	0.32

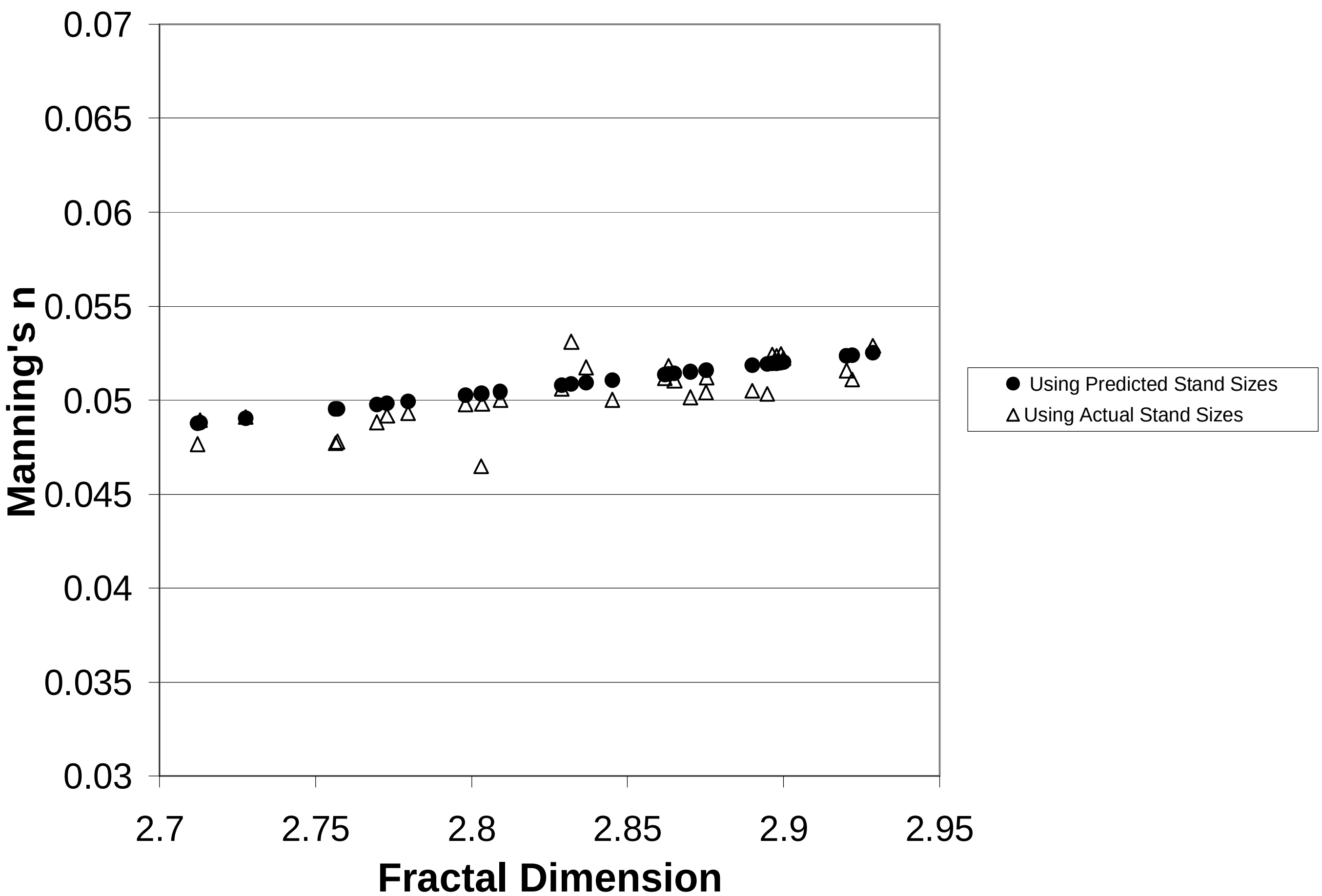
The analysis of these results must be done while keeping in mind the mean percentages of each class in the Chattahoochee forest, *i.e.,* Sawtimber (68%) and Saplings (25.3%). Thus, the Mean Error percentages look very good for both classes while the MAE percentages are good for Sawtimber but not so good for the Saplings class. The most important statistics in this table are the Root Mean Square Error (RMSE) percentage and the Nash-Sutcliffe Efficiency Statistic (N-S R^2). The NS-R^2 is computed as:

$R^2 = 1 - \frac{ModelError}{ObservedVariance}$ and thus can vary from -4 to +1 . Any positive value of this parameter indicates some predictive value in the model above random selection and thus both classes can be predicted with some level of skill by the models. The values obtained here (in the 0.25-0.50 range) are actually on the high side of values given in the literature for models using remotely sensed data to predict other physical processes. It is interesting to note that in this case the fractal-based model appears to be superior to the I-based model in terms of all error measurement indices. This is in spite of the fact that the Moran's I model had the highest conventional R^2 values during the fit to the original data.

Prediction of Manning Roughness Coefficients From Fractals

An analysis was performed to examine the sensitivity of the Manning roughness coefficients computed using the stand characteristics for the Chattahoochee forest predicted by the model. In making this analysis, it was first determined from FIA data that the mean number of trees per acre in the Chattahoochee National Forest was 598 and that the forest area was 2225 acres. Thus, there would be 1,330,550 total trees in the forest. Then, the median diameters of each stand class was employed, i.e., d = 3 in for saplings, d = 7 in for poletimber, and d = 12 in for sawtimber with the predicted size of each class from the model to determine the total cross sectional area of each sample that would be covered by trees. In this way, the (A/a) ratio was determined. The Manning *n* value was then computed for each sample using the equation given in Figure 4 above. The computations were carried out for a constant flow depth of 1 ft. The results are shown in Figure 13.

Figure 13. Manning *n* Predicted from Fractal Dimensions



The figure demonstrates the expected linear relationship between FD and *n*. This is due to the linear equations that were used to estimate both the forest class sizes and the Manning *n* from the FD. The *n* values in this particular experiment range from about 0.049 to 0.053 while the FD varied from about 2.71 to 2.94. Thus, a variation of about 8.5% in FD resulted in about a 8.2% change in the Manning *n* value. This result would appear to indicate that the roughness coefficient is not particularly sensitive to the estimated fractal dimension.

In order to test this thesis, the analysis was run again with an increased stand thickness assumed. The percentage of each forest class in the Chattahoochee was kept to the predicted values, however the number of trees per acre was varied. In order to estimate the total cross sectional area of flow (a) it was assumed that the radius of influence of an individual tree in each class would be 2 ft (sapling), 3 ft (poletimber), and 5 ft (sawtimber). These values were used along with the predicted class sizes for each sample in the Chattahoochee forest. The computations were carried out this time for a hypothetical area of 100 acres in the forest. The Manning *n* value was computed in the same manner as before. The results for a constant flow depth of 1 ft are shown in Figure 14 below.

Figure 14. Manning's n vs Fractal Dimension

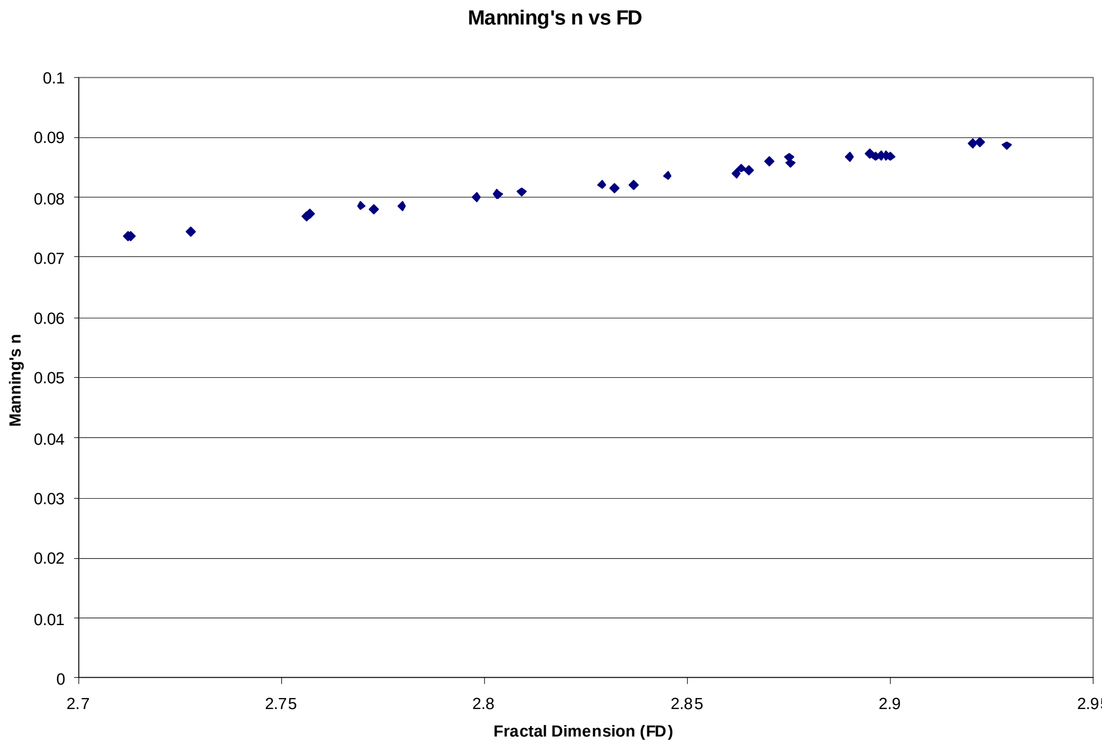


Figure 14 shows that the roughness coefficients tended to be higher than previously due to the increased stand thickness. The *n* values ranged from about .073 - .09. The magnitude of these numbers is due to the assumed radii of influence of each tree stand as well as the overall area of 100 acres, both of which lead to the value of a in the A/a ratio. If different values had been assumed, then different magnitudes of the *n* factor would have resulted. This time the data reveal that for a variation in FD of 8%, the corresponding variation in the roughness coefficient was 23% indicating that the sensitivity of the roughness coefficient estimates increases as the stand density increases. Thus, it appears that FD can be an effective measure of the roughness of forest surfaces and can encapsulate hydraulically important changes in the roughness of forest landscapes.

Effects of Spectral Resolution

In this component of the study, the fractal and auto-correlation responses from each Landsat TM band were examined to determine if significant differences existed between bands. First, the mean responses were examined to determine if there was reason to suspect differences in spectral response. Visual inspection of the results obtained from the samples of each forest indicated that the means of the FD and I from the three visible bands (1, 2, and 3) appeared to be consistently different than those from the mid-infrared bands, particularly bands 4 and 5. Band 6 was excluded from the analysis due to its larger spatial resolution (250 m compared to 30 m). The mean FD of the visible bands was consistently higher than that of the mid-infrared bands while the reverse was true of the Moran's I values. This result simply indicates that the data from visible bands represent a more complex, or rough surface, than do the data from the near IR bands.

Of more interest is the variation (variance) of the indices among samples within each band. These data indicate the sensitivity of each band to changes in spectral response from the individual canopies of each sample. The average standard deviations of the samples in each forest are given for each band in the following table:

	Band						
	1	2	3	4	5	6	7
FD							
Talladega	0.077	0.071	0.101	0.051	0.090	0.075	0.104
Oakmulgee	0.084	0.080	0.088	0.044	0.064	0.066	0.079
Bankhead	0.029	0.037	0.029	0.045	0.033	0.054	0.038
Chattahoochee	0.084	0.085	0.096	0.074	0.070	0.071	0.098

	Band						
	1	2	3	4	5	6	7
Moran's I							
Talladega	0.161	0.078	0.096	0.034	0.064	0.056	0.085
Oakmulgee	0.101	0.070	0.059	0.033	0.043	0.042	0.056
Bankhead	0.059	0.046	0.023	0.036	0.027	0.0009	0.029
Chattahoochee	0.167	0.078	0.088	0.041	0.042	0.050	0.085

Examination of these results again leads to the conclusion that the variation in the indices in the visible bands may be greater than that in the mid-IR spectral range, at least for some forests. c^2 tests on the variance (s^2) of these data lead to interesting, and sometimes contradictory, conclusions. Two tailed tests, at the 5% level of significance, were performed comparing the variance of each visible band to the variance of each of the mid-IR bands. The results revealed that there were much more likely to be differences between the variances of the Moran's I statistic than for the fractal dimension variances. The FD variances of the visible bands were found to be significantly higher than those in the mid-IR bands in all cases for the Oakmulgee forest and for all except two comparisons (bands 2 to 5, and bands 1 to 5) for the Talladega National Forest.

Conclusions

In this study remotely sensed data were used to evaluate the flow resistance characteristics of forested flood plain areas. An experimental component of the project developed a relationship between a commonly employed resistance coefficient (Manning's n) and forest stand characteristics. Another component then developed a relationship between remotely sensed indices of surface roughness or complexity (FD and I) and the forest stand characteristics. Finally, the two components were joined to directly estimate the Manning n coefficient from the remotely sensed fractal dimension (FD). The results revealed that the spatial complexity index (FD) is sensitive to changes in hydraulic roughness and that the sensitivity increases as the vegetation thickness increases. Thus, the FD index is capable of capturing hydraulically significant surface roughness indicating that it can be an effective measure of flow resistance in forested flood plain areas. It was also convincingly demonstrated that the FD (and I) can be effectively estimated from Landsat TM data, and thus the Manning roughness coefficient itself can be fairly estimated from the remotely sensed data.