vegetation distribution using ancillary and remote sensing data: a case study

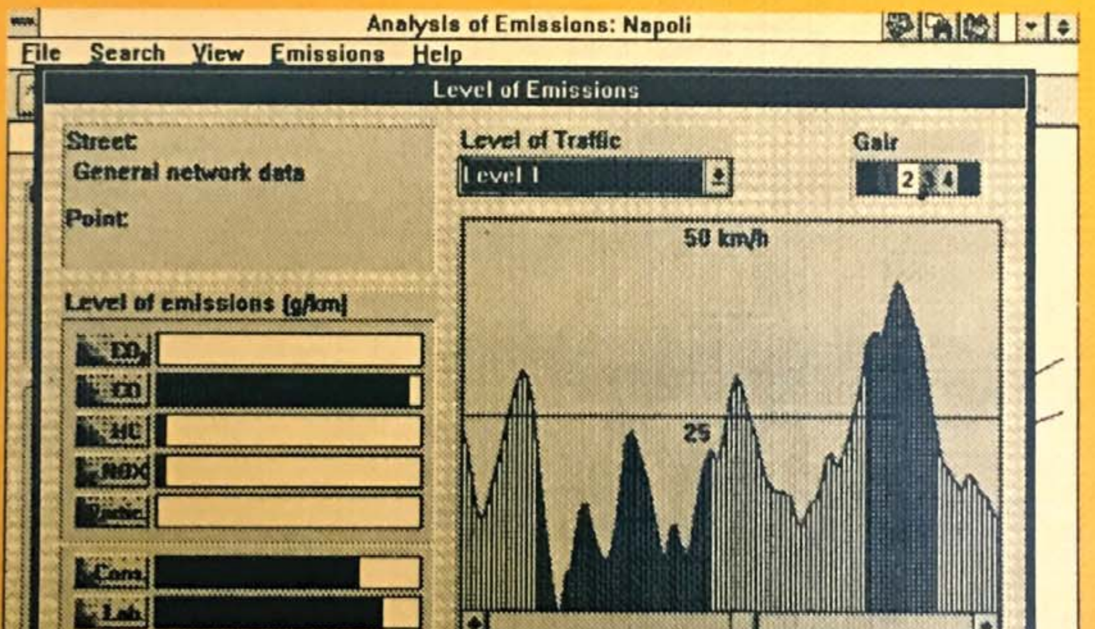
Conference Paper · September 1996

1 author:

Mohammed Taleb Obaidat
Jordan University of Science and Technology; Jadara University

Development and Application of Computer Techniques to Environmental Studies VI

Editors: P. Zannetti, C.A. Brebbia



Street: General network data

Level of Traffic: Level 1

Computational Mechanics Publications

Studying vegetation distribution using ancillary and remote sensing data: a case study

M.T. Obaidat
Civil Engineering Department, Jordan University of Science and Technology (J.U.S.T.), Irbid, P.O.Box 3030, Jordan

Abstract

A Landsat Thematic Mapper image (TM) was used, in combination with ancillary topographic and topoclimatic data, to study the distribution of vegetation classes in the Niwot Ridge-Colorado, U.S.A. A logical channel approach; i.e. spectral and ancillary data, for digital classification of remote sensing data was used. The analysis was performed using the SPSS statistical package. The vegetation class was dependent on nine selected topoclimatic and topographic data variables. These variables include: topographic slope, aspect, albedo with and without slope/aspect consideration, Normalized Difference (ND) with and without slope/aspect consideration, convexity, Potential Solar Insolation (PSI), and Slope Aspect Index (SAI). Three random sampling techniques, to select vegetation classes samples from the map, were used: random samples from all the regions in the study area, samples from hilly areas only, and samples using a strip area along the map profile. The results of this study showed that the combination of TM data with the topographic and topoclimatic data variables is an efficient way to study the distribution of wet and dry vegetation classes. Minor effects were found for samples locations on the discrimination analysis process.

1 Introduction

Phillipson (1980) defined the remote sensing task as: it is used to collect, analyze, and convert remotely sensed data to useful information, for problem solving or decision making. Among other approaches, remote sensing data has been extensively used in the last two decades to study the classification of vegetation distributions and plant communities in various areas in the world. Becking (1959), and Jensen and Estes (1978) advocated the use of aerial photograph in vegetation and forestry in many fields such as mapping, interpretation of forest conditions, reconnaissance and management planning, and crop classification using digital Landsat data which proved superior to analysis of digitized high-altitude photography. Practically various techniques and strategies were developed to deal with digital classification of remote sensing data. Three approaches may be used for this purpose: pre-classification scene stratification, post-classification class sorting, and logical channel approach -classification modification- through increasing the number of observation channels (Hutchinson, 1982). Thus, ancillary data such as digitized maps and terrain data if combined with Landsat data is expected to improve the digital classification of vegetation classes. The resulted classification might be used in natural resource inventory, with condition that the analyst has a detailed understanding of the objects of interest and their relation with the ancillary data before using them. From this perspective, Cibula and Nyquist (1987) found that applying geographic models and topographic data such as solar insolation, snow cover, and wind besides spectral bands assisted the classification process, and high accuracy level was maintained.

Thematic Mapper (TM) data is normally used in classification studies, since there is a great sensitivity of a forest's apparent reflectance to the acquisition geometry in the middle infrared part of the spectrum of the vegetation (Leprieur *et al.*, 1988). Previous research indicated that using TM data besides topoclimatic and topographic variables could be used successfully to map dominant vegetation communities (Frank, 1988). Frank and Isard (1986) stated that accuracy measurement of topographic setting is necessary to distinguish between various alpine vegetation types in the rocky mountain, because the local topographic site factors influence the plant distribution.

This paper explored the vegetation distribution of Colorado Rocky Mountain Front Range according to the altitude and various topographic and topoclimatic data, and then compared the resulted classification to the map classification. In order to extract the samples of vegetation classes from the map, the study took three random sampling methods into consideration: samples from all locations of the map, samples from hilly areas only, and strip samples (profile). A discriminative statistical analysis to check the significance of the used variables and sampling methods was also performed

2 Study Site and Vegetation Distribution

The study area is located in the Niwot Ridge at Colorado Rocky Mountain enclosed within the Ward, which is part of the backbone of the Americas - a backbone that extends more than 10,000 miles from Alaska to Patagonia. The study area is extended from 105° 32' 30"W to 105° 37' 30"W, and from 40° 02' 30"N to 40° 04' 00"N. The area is characterized of high peaks (above 14,000ft). The difference between its summits and bases is about 5,000-7,000ft. The forests are extensive and great vistas showing spectacular landform with colorful rocks and forests. The vegetation of the Rocky Mountain is altitudinally distributed and zoned with a relationship with topography (Hunt, 1974). The vegetation classes on the mountain, starting from the top of the summits, include: Alpine Zone (A), Canadian Zone with Spruce-fir Forest (SF), Transition Zone (TZ), Mostly Yellow Pine (YP), Lodgepole Pine (LP), Upper Sonaran Zone with Pinyon-Juniper Woodland (PJ), Short Grass (SG), and Sagebrush (S).

The reasons behind the selection of this study site are due to the availability of the classified vegetation map (US/IBP Tundra Biome Program and A

Contribution to the US/MAB 6 Mountains and Tundra Project, 1976; and Frank, 1988), as well as the availability of the Landsat image of the area. Frank and Isard (1986) listed the NODA classification of vegetation for the Niwot Ridge area using six groups: dry meadow which includes other six classes (1 to 6), dry fellfield (2 to 5, and 7), moist shrub tundra (20), moist meadow (8-13), snowbed (14-17), and wet meadow (18-19). These classes were mainly used in the map to check the accuracy of our study. Other classifications, using hierarchical braun-blanquet classification system based on floristic-sociological principles, and dominant vegetation communities, were also used.

2.1 Factors Affecting the Vegetation Distribution

Many factors affect the distribution of the vegetation in the selected area including:

1. Snow cover and moisture content of soil.
2. Interactive relationship between insolation, snow, wind, soil moisture, relief, and vegetation distribution .
3. Atmospheric control which is presented by solar radiation.

The classification of vegetation communities was based on the fact that there is a correlation between these factors and the vegetation classes.

3 Material and Methodology

The image used in this study was taken by Landsat 5 for Niwot Ridge- Colorado on July 29, 1984 at 10 o'clock in the morning with a sun angle of 60 degrees. The image contains digital TM data which consists of seven bands including: Band 1 (Blue-Green) 0.45-0.52 µm, Band 2 (Green) 0.52-0.60 µm, Band 3 (Red) 0.63-0.69 µm, Band 4 (Near Infrared) 0.76-0.90 µm, Band 5 (Middle Infrared) 1.55-1.75 µm, Band 6 (Thermal) 10.3-12.5 µm, and Band 7 (Middle Infrared) 2.08-2.36 µm (Campbell, 1987). The image was rectified using ERDAS software with nearest neighbor resampling approach.

TM data besides ancillary data variables (including topographic and topoclimatic variables) were used to predict vegetation distribution of the study area. A combination of these variables was selected to discriminate between six vegetation communities using channel approach method; i.e. combination of spectral and ancillary data. This approach is used because it is expected to be efficient in the vegetation classification regions which have premafrost, snow accumulation, soil moisture growing season, and solar insolation effect.

All computer work was done using ERDAS software, whereas TERRAMAR software was used for three-dimensional (3-D) purposes.

3.1 Study Variables

The TM data was calibrated geometrically before quantifying the indexes in order to correct for latitude/longitude geographic location and time factor. Two cases were studied: calibration with and without slope and aspect consideration.

The following topoclimatic and topographic variables were used:

1. Slope: the magnitude of the elevation gradient.
2. Aspect (ASP): the direction of the slope (0°-360°).
3. Albedo (ALB): the outgoing radiation over incoming radiation, which is a measure of the biomass. This depends on the calibration of the TM data at a given time, the sun angle, the atmosphere, the slope and aspect of terrain, and the surface cover (Robinove, 1982). Two cases of albedo were studied: with and without slope-aspect effect (abbreviated as ALBT and ALB, respectively).
4. Normalized Difference Index (ND): a function of the near infrared band (NIR) and the red band (R). Equation 1 shows the ND formula (Frank, 1988):

$$ND=(NIR-R)/(NIR+R) \qquad (1)$$

Normalized Difference Index was considered with and without slope-aspect effect (abbreviated as NDT and ND, respectively). Vegetation Index (VI) was also considered for preliminary study, and finally ignored due to similar characteristics with ND.

5. Convexity (CON): the second derivative of the slope.
6. Potential Solar Insolation (PSI): the direct beam potential solar insolation was computed from March to December for 24 hours a day.
7. Slope Aspect Index (SAI): a topoclimatic index used to study wind effect which is function of slope and aspect. Equation 2 shows the SAI formula (Frank, 1988):

$$SAI=(Sin(Slope)*Aspect)/(MaxSAI*k) \qquad (2)$$

where: *max. SAI* is the maximum index value, and *k* is a constant to convert to eight bit value.

The data variables, including the listed seven variables besides ALBT and NDT, were classified into nine variables, where each one of them was dependent on the Digital Elevation Model (DEM).

ERDAS software was used to produce the images of SAI, Slope, albedo, and ND. The important characteristics of the generated images are shown in Table 1. Layer combinations of bands, representing the combination between topographic and topoclimatic variables, using a SUBSET program were also generated. Combinations of these images include: convexity, ND, and aspect bands; slope, SAI, and ALB; PSI, SAI, and ND; Slope, aspect, and albedo; and convexity, PSI, and SAI.

TERRAMAR software was also used to construct 3-D images from various directions for the nine studied variables of ancillary data. The images represent a set of combination of 3-D images for: slope, PSI, and SAI; slope, aspect, and albedo; albedo, ND, and aspect; and ND, PSI, and convexity. The images clearly showed the accumulation of snow cover in East face which was associated with low SAI, while the west face was free of snow due to blow of the wind.

Table 1: Important characteristics of the image for the used nine variables.

Features	Characteristics			
	Minimum	Maximum	Mean	Standard Deviation
Slope	0	53	14.8	8.4
Aspect	0	359	156.9	106.4
ALB	15	98	39.3	21.8
ALBT	15	100	39.9	21.5
ND	0	255	127.5	63.3
NDT	0	255	127.5	63.3
Convexity	-10	96	10	13
PSI	0	1580	1315.1	203.2
SAI	0	132	27.2	18.9

3.2 Random Sampling

Three sampling techniques were used to represent the vegetation distribution on the used map. Around 30 points were digitized for each class from the map. This leads to at least 180 data points for each sampling case for the whole classes. The sampling points were kept approximately equally distributed for the specified samples. The three sampling cases used include:

1. All regions sampling including flat, hilly, and in between terrains.
2. Hilly areas sampling.
3. Strip (profile) sampling of about one tenth the width of the area.

4 Discriminant Analysis

SPSS statistical package was used to make discriminant statistical analysis to study the dependence of the vegetation classes on the nine used variables, using the three mentioned sampling methods. The variance-covariance matrices showed a strong correlation between ND and albedo, and slope and SAI, therefore, albedo and slope were not included in the final classification process. Thus, only five variables which represent the solar, topoclimatic, and topographic variables were used for the purpose of statistical analysis to discriminate between vegetation distribution for the study area. These variables include: PSI, convexity, normalized difference with topography consideration (NDT), aspect, and SAI. Minor correlation exists between these variables, but the SAI and convexity.

Only five vegetation classes were also used in the statistical analysis. Dry fellfield class was combined with dry meadow class to be referred as class 1, due to difficulties to discriminate between the two classes. The other four classes include: moist meadow (class 2), wet meadow (class 3), moist shrub tundra (class 4), and snowbed (class 5).

The discrimination analysis results for the three sampling cases are shown in Table 2. While the canonical discriminant functions evaluated at group means (group centroids) are shown in Table 3. Variables in the table are ordered by size of correlation within function. Four functions, for each sampling case, representing the discriminate function between vegetation distribution, were obtained as functions of the five used variables (NDT, aspect, convexity, PSI, and SAI). The functions were representing albedo (irradiance/radiance), snow and wind, solar insolation, and fourth functions with common characteristics of most of the variables. In the random sampling case the four function represent: albedo, topography, albedo, and potential solar energy, respectively. The strip sampling functions represent albedo, topography, solar energy, and a fourth function of common characteristics of the previous three functions, respectively. While, the hilly area sampling method functions represent albedo, potential solar energy combined with other factors, potential solar energy, and snow and wind, respectively.

From the previous functions, it is obvious that albedo, solar insolation, and snow accumulation and wind direction play as the major factors in the discrimination and distribution of the vegetation communities in the Niwot Ridge-Colorado. Albedo differentiated class 2 from other classes, potential solar insolation differentiate class 3 and class 5 from other classes, and snow and wind factors discriminate class 1 and 4 from other classes.

The trend for vegetation distribution is dependent on PSI by which the vegetation classes are going from dry to moist; i.e. from class 1 to class 5, as the PSI value increases. Except for classes 3 and 4, the vegetation distribution is becoming wet as the convexity increases; i.e. going down the crest. Similar observations could be drawn for ND, aspect, and SAI. SAI is a good discriminant of alpine vegetation, since the decrease of SAI is an indication of wind blowing and snow free. Thus, the vegetation distribution is mainly dependent on both topoclimatic and topographic variables.

Comparing the mean values of the three sampling cases for each variable, variances of less than 20% exist. This is due to the little effect of the way of selection of samples from the used map.

The classified percentages correct for the strip and random sampling cases were better than the hilly area sampling. They were 36.6%, 42.1%, and 42.5% for hilly areas, strip, and random sampling, respectively. Classes 2 and 4 were correctly classified up to 44.4% and 58.5%, respectively, using strip sampling. While hilly area sampling gave better classification results for Class 1 with 44.9% correct. Random sampling was so effective to discriminate Class 4 and Class 5 with percentages correct reached 73.3% and 63.3%, respectively. Whereas, Class

3 (wet meadow) was hardly discriminated using the three sampling techniques. It reached a maximum percentage correct of about 37.2% using hilly area sampling. Figure 1 depicts the correct percentages of classification for the five vegetation classes of the three sampling cases. As expected, taking comprehensive sampling for the hilly areas as well as flat areas; i.e. random sampling, is the best way to have a representative sample especially if it is well distributed all over the map area.

Table 2: Classification analysis results for the three sampling cases.

Actual Group	% Correct of Predicted Group Membership for Different Classes (Numbers in parentheses are the correct predicted classified samples)					
	No. of Cases	Class 1 (%)	Class 2 (%)	Class 3 (%)	Class 4 (%)	Class 5 (%)
A. Strip Sampling (% of classes correctly classified= 42.1%)						
Class1	54	20.4% (11)	16.7% (9)	27.8% (15)	25.9% (14)	9.3% (5)
Class 2	27	3.7% (1)	44.4% (12)	18.5% (5)	3.7% (1)	29.6% (8)
Class 3	27	22.2% (6)	3.7% (1)	29.6% (8)	18.5% (5)	25.9% (7)
Class 4	41	4.9% (2)	2.4% (1)	24.4% (10)	58.5% (24)	9.8% (4)
Class 5	41	4.9% (2)	17.1% (7)	17.1% (7)	0% (0)	61.0%(25)
B. Hilly Area Sampling (% of classes correctly classified= 36.6%)						
Class 1	98	44.9% (44)	14.3% (14)	23.5% (23)	10.2% (10)	7.1% (7)
Class 2	55	12.7% (7)	38.2% (21)	16.4% (9)	9.1% (5)	23.6% (13)
Class 3	43	18.6% (8)	18.6% (8)	37.2% (16)	23.3% (10)	2.3% (1)
Class 4	32	31.3% (10)	9.4% (3)	25.0% (8)	25.0% (8)	9.4% (3)
Class 5	51	9.8% (5)	33.3% (17)	15.7% (8)	15.7% (8)	25.5%(13)
C. Random Sampling (% of classes correctly classified= 42.5%)						
Class 1	62	40.3% (25)	17.7% (11)	11.3% (7)	14.5% (9)	16.1% (10)
Class 2	29	20.7% (6)	31.0% (9)	0% (0)	6.9% (2)	41.4% (12)
Class 3	30	16.7% (5)	13.3% (4)	6.7% (2)	16.7% (5)	46.7% (14)
Class 4	30	6.7% (2)	16.7% (5)	0% (0)	73.3% (22)	3.3% (1)
Class 5	30	3.3% (1)	13.3% (4)	3.3% (1)	16.7% (5)	63.3%(19)

Table 3: Canonical discriminant functions evaluated at class centroids for the three sampling cases.

Class	Functions			
	Function 1	Function 2	Function 3	Function 4
Strip Sampling				
Class 1	0.24915	0.45501	0.01167	0.07238
Class 2	-0.56103	-0.00372	0.33797	-0.08824
Class 3	0.37043	0.14705	-0.20003	-0.17618
Class 4	0.82845	-0.48495	0.03244	0.03650
Class 5	-1.03108	-0.20871	-0.13865	0.0423
Hilly Area Sampling				
Class 1	0.35959	-0.35786	0.00665	-0.00346
Class 2	-0.55841	-0.01422	-0.23960	0.07627
Class 3	0.40840	0.50035	-0.21306	-0.07675
Class 4	0.28217	0.42499	0.32855	0.12162
Class 5	-0.61016	0.01446	0.21910	-0.08721
Random Sampling				
Class 1	0.02773	0.46304	-0.05657	-0.01214
Class 2	-0.53406	-0.25594	-0.38017	0.01462
Class 3	-0.07360	0.25102	0.28692	0.02652
Class 4	1.13765	-0.41396	-0.03007	0.00044
Class 5	-0.60511	-0.54660	0.22757	-0.01601

5 Discussion and Suggestions

Although the discrimination results, between the vegetation classes in the Niwot Ridge-Colorado using topoclimatic and topographic variables besides the TM data, were not absolutely effective, the study shows that using such variables is promising for discrimination purposes. According to NODA classification, it was found that it is extremely difficult to discriminate between dry fellfield and dry meadow. Other variables might be useful for the purpose of vegetation distribution prediction, such as: vegetation index (VI), elevation, reflectance absorption index

(R/A), relief, ND, aspect, and SAI.

Snow and wind, potential solar energy, and albedo had significant impact on the classification process. This is an indication of observed correlation between variables which had some sharing of common factors. Based on these indicators, as a generally speaking, the dry vegetation is located on the high altitude, while the moist vegetation is down the slope.

The effect of sample location on the classification process of vegetation distribution was minor. It might be due to the use of the same variables in the three cases and minimum number of points (about 30 for each class). Anyhow, the classification was enhanced of about 6% when using strip, or random sampling if compared to hilly area sampling.

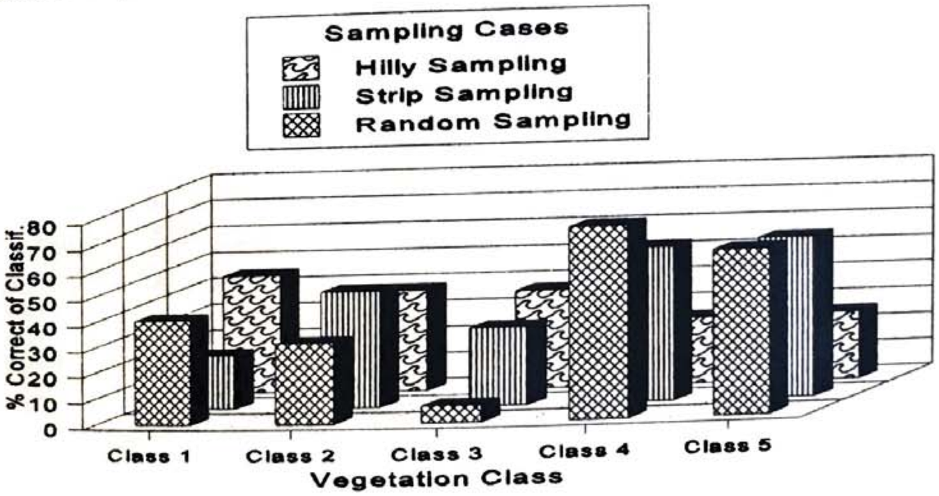


Figure 1: Classification percentage correct for five classes of all sampling cases.

For further future research, in order to enhance the classification process efficiency, it is recommended to do the following:

1. Add new variables to discriminate between class 1 and class 2. These variables might include VI, R/A, relief, and elevation.
2. Increase the number of samples for each class to assure that the sample will follow a normal distribution.
3. Use other classified map to include more detailed vegetation communities distribution.
4. Use approaches other than the logical channel approach to study the potential of remote sensing data on the vegetation distribution.

6 Conclusions

The logical channel approach is proven to be useful in feature characterization of vegetation communities. The classification accuracy of different types of vegetation was around 45%. This is due to the well known fact that spectral reflectance patterns are associated with certain variables such as slope, aspect, SAI, ND, PSI, albedo, and any combination of these variables. Using topoclimatic and topographic indexes besides the remote sensing data have the advantage of combining both spectral and information category classification. Samples for the vegetation classes should also be comprehensive and well distributed all over the vegetation classes in order to enhance the accuracy of the classification process of vegetation communities. Further, other parameters should be considered in order to practically adopt this approach of classification.

References

1. Becking, R. W. "Forestry Applications of Aerial Color Photography", *Photogrammetric Engineering and Remote Sensing (PE&RS)*, 1959, 25: 559-565.

2. Campbell, J. B. "Introduction to Remote Sensing", *The Guilford Press, New York*, 1987.

3. Cibula, W. G.; and Nyquist, M. O. "Use of Topographic and Climatological Models in A Geographical Data Base to Improve Landsat MSS Classification for Olympic National Park", *Photogrammetric Engineering and Remote Sensing (PE&RS)*, 1987, 53 (1), pp. 67-75.

4. Frank, T. D. "Mapping Dominant Vegetation Communities in the Colorado Rocky Mountain Front Range with Landsat Thematic Mapper and Digital Terrain Data", *Photogrammetric Engineering and Remote Sensing (PE&RS)*, 1988, 54 (12), pp. 1727-1734.

5. Frank, T. D.; and Isard, S. A. "Alpine Vegetation Classification using High Resolution Aerial Imagery and Topoclimatic Index Values", *Photogrammetric Engineering and Remote Sensing (PE&RS)*, 1986, 52 (3), pp. 381-388.

6. Hunt, C. B. "National Regions of the United States and Canada", *W. H. Freeman and Company, San Francisco*, 1974.

7. Hutchinson, C. F. "Techniques for Combining Landsat and Ancillary Data for Digital Classification Improvement", *Photogrammetric Engineering and Remote Sensing (PE&RS)*, 1982, 48 (1), pp. 123-130.

8. Jensen, J. R.; and Estes, J. E. "High-Altitude Versus Landsat Imagery for Digital Crop Identification", *Photogrammetric Engineering and Remote Sensing (PE&RS)*, 1978, 44 (6), pp. 723-733.

9. Leprieur, C. E.; Durand, J. M. ; and Peyron, I. L. "Influence of Topography on Forest Reflectance Using Landsat Thematic Mapper and Digital Terrain Data", *Photogrammetric Engineering and Remote Sensing (PE&RS)*, 1988, 54 (4), pp. 491-496.

10. Philipson, W. R "Problem-Solving with Remote Sensing", *Photogrammetric Engineering and Remote Sensing (PE&RS)*, 1980, 46 (10), pp. 1335-1338.

11. Robinove, J. C. "Computation with Physical Values from Landsat Digital Data", *Photogrammetric Engineering and Remote Sensing (PE&RS)*, 1982, 48 (5), pp. 781-784.