GROUND REFERENCE FOR ASSESSMENT OF FORESTRY APPLICATIONS IN REMOTE SENSING

Kerry Van Siclen, William Stiteler IV and Dr. Paul Hopkins
State University of New York-College of Environmental Science and Forestry
Environmental Resources and Forest Engineering
312 Bray Hall
Syracuse, New York 13210
klvansic@syr.edu
wmstitel@syr.edu
phopkins@syr.edu

ABSTRACT

Accurate ground reference data are fundamental for using remote sensing techniques to produce accurate results in forestry applications. Having complete ground reference data also allows for valid assessment of modern remote sensing, while establishing consistent baseline information for evaluating or refining methods. Acquiring this data using survey grade Global Positioning Systems (GPS) equipment and forest inventory plots provides superior information for image processing relating to forest management. This type of detailed reference data can support varied remote sensing applications. Collecting such information as slope and aspect of the terrain, individual tree species and crown position and overall forest type within each plot provides the basis for deriving other forest measurements such as basal area or tree density. Using a high level of precision when locating plot centers allows application with both high-and low-resolution imagery. Establishing these plots in an unbiased fashion throughout the forest provides statistically sound reference for use with image processing. An accurate, complete reference data set is the key to getting remote sensing results that may be interpreted with confidence. This paper discusses the need and methodology for collecting accurate and detailed reference data as a means for consistently assessing the accuracy of modern remote sensing techniques in forestry applications.

INTRODUCTION

Since 2000, the National Aeronautics and Space Administration (NASA) has funded a project to help integrate remote sensing into the practices of State Forestry Organizations (SFOs). This project requires a complete and accurate set of ground reference data to aid in assessing current practices and exploring new applications. More broadly, the collection of ground reference is an important part of any classification and mapping project utilizing remote sensing (Congalton and Biging 1992). Therefore, it made sense to develop ground reference data for different areas within the Northeastern United States that would be useful for multiple projects in the future. Having a complete ground reference data set will support many modern remote sensing techniques. Because any application is only as good as the data used to develop it (Ma *et al*. 2001) compiling an accurate reference data set is a critical step in classifying forested areas as well as performing other remote sensing techniques. By allocating time and effort into this broader objective, it was anticipated that this reference data would be valid for a wide range of applications throughout the Northeastern United States.

Continuous Forest Inventory (CFI) plots, permanent plots for which information is recorded, were used for this reference. These plots are usually used by foresters to provide forest information over time, but can also provide detailed reference data for remote sensing. However, existing CFI plots are usually located generally in reference to known points, or at best with hand-held GPS receivers. More precise location information is needed for many remote sensing applications.

For a complete set of reference, it is imperative to acquire enough reference for each forest type. However, the Northeastern United States presents a difficult challenge because much of the forest in this region is a complex mixture of species with subtle variations over space. The region is especially dominated by mixtures of temperate deciduous species that are often consolidated into one category termed “Northern Hardwoods.” Interestingly, this area lacks the strong elevation gradients present in the Western United States, where it is easier to find particular forest types. When developing a plan to achieve reference data, this subtle variation of forest types became a factor. This paper discusses the problems encountered and lessons learned while trying to establish a complete set of reference data for remote sensing applications in forestry.

OBJECTIVES

When developing a system for collecting reference data in areas throughout the Northeastern United States, there were many objectives that needed to be met in order to efficiently collect complete and accurate data. The primary desire was to collect a complete set of reference data that could be used for a wide variety of remote sensing applications in the Northeastern United States. However, numerous secondary objectives and constraints affect the process of developing this type of data set.

There were many aspects and details that needed to be considered prior to commencing fieldwork. A goal of this project was to utilize both high and low spatial resolution imagery for different applications. This meant collecting data specific to this need while in the field. In order to use high spatial resolution imagery, tree-level information within the forestry plots was needed. At the same time, applications with lower resolution imagery could be developed by generalizing the tree data and collecting general information on the area surrounding each plot. The type of imagery used also dictated how accurately the reference data needed to be located. Plot centers needed to be located precisely for potential use with imagery of high spatial resolution.

When establishing or locating existing continuous forest inventory plots, the statistical validity needed to be addressed. The need for a complete set of reference meant trying to include enough plots of each forest type commonly found in the northeast. However, it was also desirable to have the reference plots confined to somewhat confined geographic areas. This would reduce the expense of the imagery needed as well as allow for a reasonable amount of reference data on each image used. Therefore, it was necessary to develop a scheme that clustered the plots (to encompass these ideas) while proving to be statistically valid.

Certain areas throughout the northeast were selected as study areas in order to account for the range of forest types found in this region. For example, the Allegheny Plateau is known for having an abundance of black cherry trees. Other forested areas have a slightly different composition of forest types. Visiting a number of areas within this region allowed for a diverse range of forest types across the northeast. Study areas were also chosen based on their accessibility.

A complete reference set for the Northeastern U.S. is a very valuable resource for both SFOs and for other researchers. Therefore, this reference is being shared through a web page and it is important for the reference data to be free of restrictions on use.

STUDY AREA

The need for particular forest types and accessibility of forest areas dictated the study areas for this project. In the Northeastern United States in particular, it is difficult to acquire a complete set of reference data for each tree type. There are certain forest types that dominate and others that are sparse; however it is important to acquire enough reference for each type of forest cover. This project's affiliation with the State University of New York's College of Environmental Science and Forestry meant that forests owned by the college were readily available as study sites. Additional public lands were utilized to supplement campus properties. Use of college and public lands helped to avoid privacy issues associated with specific plot information and the objective of sharing this collected data with other projects and researchers.

With these factors in mind, project personnel visited six main areas of interest (AOIs): Central New York, Central Adirondacks, Eastern Adirondacks, Catskills and Hudson River Valley, Allegheny Plateau and Northern New Hampshire. For the Central New York area, existing plots within Heiberg Memorial Forest were located and supplemented with new plots in the nearby Kettlebail and Morgan Hill State Forests as well as the Labrador Hollow Unique Area. In the Central Adirondacks region, crews located existing CFI plots in Huntington Wildlife Forest and established new plots in the nearby Vanderwacker Forest. Pack Forest in the Eastern Adirondacks provided already established CFI plots for field crews to locate. New forestry plots were established in Allegany State Forest in New York and the adjacent Allegheny National Forest in Pennsylvania. Project areas in the Catskills and Hudson Valley in New York consisted of new plots located in the Mongaup, Peekamoose and Overlook Mountain regions of the Catskill Park and existing plots in the Lennox and Ninham Mountain State Forests. The project established new plots in Nash Stream Forest and Bunnell Forest Preserve in Northern New Hampshire. These locations encompass many of the forest types found in the Northeastern United States. However, there still exists a need for more of certain forest types in order to have a complete reference set. Figure 1 shows the location of these areas of interest.

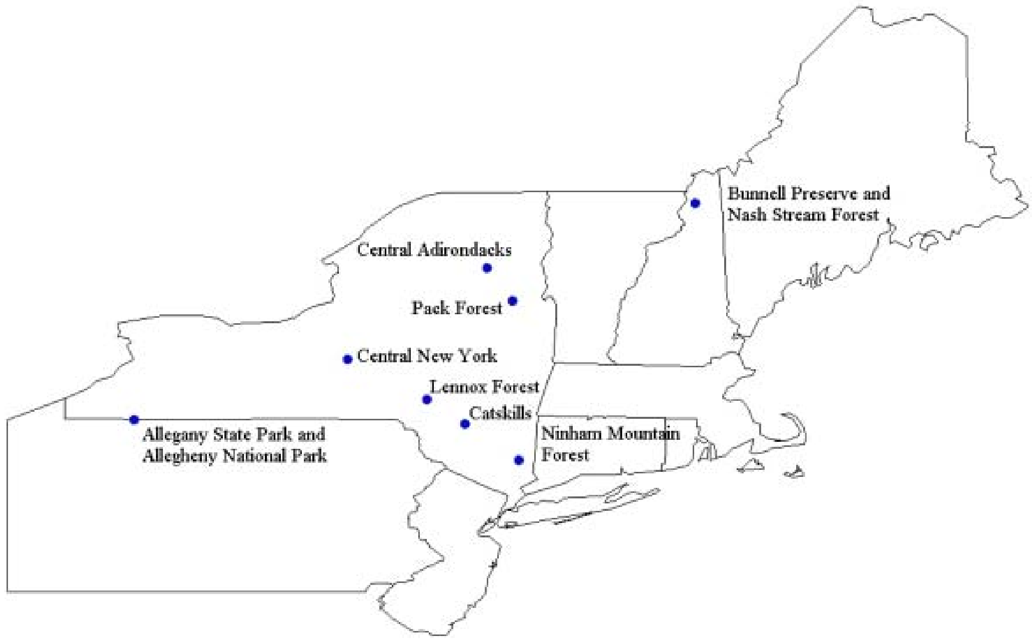

Figure 1. Forest Locations throughout the Northeast.

METHODS AND MATERIALS

Establishing New CFI Plots

Establishing new continuous forest inventory plots was an essential part of collecting ground reference data for use with remote sensing techniques. Supplementing existing plots with new plots provided a better representation of forest types. In trying to reduce expenses and have better ground reference data, the forestry plots were clustered near each other within each region so that a reasonable number of images would cover the reference sites. For each region or area of interest, we used a grid with 15- chain (roughly 300 meter) spacing. Each grid intersection was a possible plot location for a particular area. Time dictated how many grid intersections field crews could cover throughout the field season. Two or three person crews paced this grid from a predetermined origin that was designated after talking to regional foresters. The general local knowledge helped to select plots that represented common forest types that were previously not sufficiently covered.

Figure 2. Typical fieldwork procedure.

Once at a plot center, crews marked it with a 2-foot piece of iron bar. This mark was left under the ground surface to be unobtrusive to the forest while providing a means for returning with a magnetic field locator. At each plot, tree measurements were recorded with a TDS Ranger electronic data recorder. These measurements included:

- Species
- Position within the plot (distance and direction from plot center)
- Crown position (dominant, co-dominant, intermediate or understory)
- Vigor (dead, low, average or above average)
- Diameter at breast height (DBH) for trees that have a DBH larger than 7.5 cm
- Tree height for tallest tree in each of the cardinal directions

General plot measurements were also recorded for use with lower spatial resolution imagery. These plot measurements included:

- Forest type designation
- A count of the total number of small trees (< 7.5 cm DBH)
- An estimate of the mix of species for the smaller trees
- Slope of the terrain (%)
- Aspect of the terrain (flat, or one of 8 cardinal directions)
- Any other general plot information that may be useful (e.g. located near a stream or wetland)

To collect a more complete dataset for lower spatial resolution imagery, crews also stopped every 5 chains (roughly 100 meters) between consecutive plots, and measured information for "subplots." Information recorded for subplots included general forest type, tree density, and average DBH of non-understory trees.

Locating Forestry Plots

Locating both existing and newly established forest inventory plots was a fundamental part of collecting the ground reference data. To use this reference data with high spatial resolution imagery required the use of survey grade global positioning systems (GPS) equipment. GPS equipment of this nature has centimeter-level accuracy for open ground. However, the signal can be degraded under forest canopy to have a significantly lower accuracy. This project used Leica GPS 300 series units including two roving receivers, antennas and processors as well as a local base station.

Crews set up the GPS rover unit on the plot center, attached the antenna to a tripod, and extended the antenna approximately 3 meters to remove some of the forest canopy effects. The GPS unit was set to record for 180 epochs at a 5 second interval (approximately 15 minutes). Recording for this length of time proved to adequately balance locational precision with time efficiency. During the GPS data collection period, the crew would commonly acquire three pictures with a digital camera; one of the canopy directly above the GPS antenna and two side shots of the plot to give an idea of what the plot and its surroundings look like. In addition, the crew would write down information about tree type within the plot and its surrounding area and any information that would help to further describe the plot. The crews paid particular attention to whether the forest cover was similar inside the plot to the area just outside the plot.

Figure 3. Typical canopy photograph of a plot center.

The base station unit was setup on a nearby point of known coordinates so that the plot locations could be differentially corrected. The use of a local base station helped ensure that the base station and the roving receivers were recording the same satellites. Differential correction adjusts the collected locations for atmospheric distortion or other environmental distortions that could be present at the time of data collection. The base station unit was set over control points of the highest order that could be found in the vicinity of each forest. (Information on control points can be found at the Federal Base Network web site.[1]) The new Federal Geodetic Control Committee (FGCC) accuracy standards for GPS positioning techniques, define higher order control points of order B, A and AA and these points can be respectively 10, 100 and 1000 times more accurate than the old first order points (Van Sickle 1996). Use of these higher order control points would allow our field crews to attain more accurate GPS locations. Commonly, crews set up the base station unit on B or first order control points.

RESULTS

In this project's two-year duration, there has been substantial progress towards collecting ground reference for the Northeastern United States. While there is still more to be achieved, two field seasons have provided roughly 600 forestry plots of differing forest types. In the summer of 2000, field crews collected nearly 250 already established forestry plots. The specific breakdown of the forest types and locations can be seen in Table 1. In the summer of 2001, field crews located nearly 400 existing and newly established plots. Table 2 provides a breakdown of specific forest information for the plots visited in 2001.

Table 1. Continuous Forest Inventory Plots Visited in 2000.

	Forest Properties (New York State)				
Forest Type	Central NY	Central Adir., NY	Eastern Adir., NY	Catskills and Hudson Valley, NY	**Total Obtained in NYS**
Birch				3	3
Hemlock	2		5	1	8
Maple	1			25	26
Maple/Beech	1	6		3	10
Maple/Beech/Birch	34	35	7	17	93
Maple/Black Cherry	2			3	5
Mixed Conifer		1	3	0	4
Mixed Hardwood	1			10	11
Mixed Hardwood/Conifer	6	10	22	8	46
Mixed Conifer Plantation	1		1	0	2
White Pine			7	0	7
Norway Spruce Plant.	1		1	0	2
Oak			2	3	5
Open Field Brush	2			0	2
Other Forest Type	1			0	1
Plantation Other	6		1	0	7
Red Pine Plant.	5		1	0	6
Scotch Pine Plant.	1		1	0	2
Spruce/Fir		2		0	2
White Pine Plant.			1	0	1
Total In Property	64	54	51	73	242

[1] The FBN site can be accessed through http://www.ngs.noaa.gov/PROJECTS/FBN/.

Table 2. Continuous Forest Inventory Plots Located in 2001.

	Forest Locations (North-Eastern United States) 2001 collection					
Forest Type	**Allegheny Plateau**	**Catskills and Hudson Valley, NY**	**Central NY**	**Central Adir., NY**	**Northern NH**	**Total**
Alder					1	1
Ash/Sugar Maple	3	2	2			7
Aspen					1	1
Beech/Maple/Oak	2	1				3
Black Cherry	8	1				9
Bog/Wetland					7	7
Cherry/Ash	1	1				2
Cherry/Beech	1					1
Cherry/Beech/Maple	1	2				3
Fir/Birch					7	7
Hemlock/Hardwood	16	9	3	6		34
Maple	2	1	3		2	8
Maple/Beech	5	6	3	9		22
Maple/Beech/Birch	1	3	20	46	5	75
Maple/Birch	1			2	10	13
Maple/Black Cherry	17	1	3	2	1	24
Mixed Conifer			8	1		9
Mixed Conifer Plantation			5			5
Mixed Hardwood	17	5	8	7	6	43
Mixed Hardwood/Conifer			8	21	11	40
Norway Spruce			9			9
Oak	6					6
Oak/Maple	6	1				7
Open Field Brush	3		7	2	1	13
Pine/Hardwood		1	5	1		7
Plantation Other			4			1
Red Pine			4			4
Scotch Pine			1			1
Spruce/Fir				4	3	7
Spruce/Hardwood			7	8	2	17
Total	90	34	96	109	57	**386**

DISCUSSION

General Issues

The fieldwork performed for this project was difficult due to the nature of forests in the northeast combined with the desire to develop a database of representative ground reference for the region. Northeastern forests contain a large abundance of maple, beech and birch while having other tree species spread throughout the region. This made collecting a complete reference dataset more difficult for certain forest types. Additional effort and locations will be needed to adequately represent forest conditions in the northeast.

In order to achieve this desired dataset, it was necessary to collect detailed information in the field. Collecting data regarding tree height, tree DBH and tree spacing, allows the calculation of tree volume and basal area measurements. The reference data can be expanded to meet other applications of future work without returning to the field. This allows applications at using imagery at varying scales to utilize this fieldwork effort. The data can be used for image processing of high-and low-resolution imagery because of the detail of the data collected in the field.

Electronic data recorders were used to record field measurement data. These recorders were essential for reducing the time required to get the data in a useable form for remote sensing applications. However, it was still necessary to spend significant time and effort in cleaning up the data and preparing it for use. Therefore, for future work, the file type used for export and the setup of the interface used for data recording will be designed to reduce the amount of manipulation needed following fieldwork. Recording plot information manually in field books requires even more time and effort in transcription and clean up.

Ground Reference Collection vs. Aerial Collection

There are many methods for collecting and acquiring remote sensing reference data. Reference data are often collected from photo interpretation of large-scale aerial photographs or by flying over the area and recording observations (Congalton and Biging 1992). However, many geographic data collection methods are still observed through ground surveys (DeMers 2000). The high resolution of modern imagery now demands data that are better than can usually be collected by photo interpretation or aerial sketching. For this project, ground surveying involving field crews visiting each forest seemed to be the most viable and thorough method for collecting ground reference data. Collecting ground reference through fieldwork was necessary due to the need for tree-level information to support using high spatial resolution imagery and a broad array of analysis procedures. Canopy measurements were recorded, but other ground measurements were also desired that would not be possible to record through other means.

Locating Pre-existing Plots in the Field

A portion of the collection of ground reference required locating previously established continuous forest inventory plots. Since field measurements had been recorded previously, this step only requires locating the plot center with a GPS roving unit. This task may seem trivial but it can be difficult to locate these plots in the field. Pacing from a determined landmark to the plot center usually was the method used to first establish the plot center. Repeating the pacing usually allows crews to get in the vicinity of the plot; however even when close, the plot center is not always obvious. The more recently the plots were measured and reestablished by forest personnel, the easier it was to visually locate each plot due to flagging or other tagging methods included in reestablishment. In certain instances, it was not possible to locate a plot or plot center due to changes in the forest since the last time forest personnel visited. If the plot itself was found, the center could be re-established using tree distance and azimuth information.

Sampling Strategy for New Plots

Collecting reference data in a statistically valid sampling method provides the mathematical foundation for developing scientifically based inferences (Stehman 1992). This project designed field procedures to ensure the sampling was performed in a statistically valid manner. The plots sampled throughout each forested area were established in an unbiased fashion by choosing areas within the forests that are known to have desired general forest types and systematically positioning the forestry plots.

Two basic sampling methods prove to be statistically valid: simple random and systematic. The field procedures designed for this project most closely represented a systematic sampling method using a grid with a randomly defined origin. Simple random sampling results in a statistically sound sample, but is not feasible due to expense of imagery and the time needed to invest with such a sampling strategy. Systematic sampling results in a spatially well-distributed sample and therefore usually has better precision (Stehman and Czaplewski 1998).

The need for a complete set of data for each forest type also dictated how to best sample in the field. When developing the grid for each forested area, a general idea of forest type was known and the areas visited by field crews were often those that contained under-represented forest types. In order to provide some randomness, the beginning of the grid was setup to be quasi-random in an area easily located, such as a road intersection or other marker within the forest. In using this sampling strategy, less common forest types still tended to be missed or sparsely sampled, while the abundant forest types were continually sampled. Stratification can be used with systematic sampling to ensure that certain forest cover types are prevalent in the data collected. The method in place for sampling to date touched on this idea but did not fully stratify the forest types prior to fieldwork. Future work will address the best method to better locate less common forest types, possibly through stratification.

GPS Accuracy

Survey grade GPS units were essential to meet the goals of the project. In an effort to achieve the highest accuracy possible, a local base station was also used to differentially correct the locations collected by the roving units. Using the survey grade GPS equipment in this manner is known to achieve centimeter-level accuracy in open areas. However, the forest canopy substantially degrades the precision and accuracy of collected points. Under the canopy, the number of satellites the GPS receiver can record a signal from is reduced causing the geometric dilution of precision to increase and the signal to noise ratio to decrease. The effect canopy has on accuracy, as opposed to precision, is less well understood. Other studies, as seen reported by Naesset (2001), have shown that characteristics related to the canopy and the observation period had the most significant effect on accuracy. However, there still exists a need to estimate accuracy under different canopy cover using survey grade GPS equipment. This project is presently developing a traverse with a total station under different forest densities to test the accuracy of the GPS equipment against a sample of known points.

SUMMARY AND RECOMMENDATIONS

Being involved with a project that focuses on further integrating remote sensing into State Forestry Organizations prompted the necessity of a complete and accurate ground reference data set. This data will prove to be viable for a multitude of applications and techniques. The need for tree-level detail of each plot made field work an essential part of this project. Using other reference collection methods would only provide canopy information.

The Northeastern United States proved to be a difficult area in which to acquire a complete reference set for each forest type. Because the northeast is such a challenging area, methods developed here should be widely applicable for many other regions. To date, a large dataset has already been acquired, but there is still more to be done. For the next field season, this project will visit areas to collect data for the sparsely represented forest types. Trying to create a statistically valid clustering of forestry plots has made it difficult to acquire an adequate number of reference points for all forest types. Future work may involve stratifying forested areas by forest type once the grid is laid out.

Additional Information

The project referred to in this paper is discussed in detail on the NASA FOReST (Forestry Organization Remote Sensing Technology) Project’s web site. These web pages be accessed through http://www.esf.edu/forest. This web site discusses the project in its entirety and distributes a complete set of information of previously existing and newly located plots.

REFERENCE

Congalton, Russell G. and Greg S. Niging (1992). A Pilot Study Evaluating Ground Reference Data Collection Efforts for Use in Forest Inventory. *Photogrammetric Engineering and Remote Sensing*, Vol.58, No.12, pp.1669-1671.

DeMers, Michael N. (2000). *Fundamentals of Geographic Information Systems-Second Edition*. John Wiley & Sons, Inc., New York.

Ma, Zhenkui, Melissa M. Hart and Roland L. Redmond (2001). Mapping Vegetation across Large Geographic Areas: Integration of Remote Sensing and GIS to Classify Multisource Data. *Photogrammetric Engineering and Remote Sensing*, Vol. 67. No.3, pp.295-307.

Naesset, Erik (2001). Effects of Differential Single- and Dual-Frequency GPS and GLONASS Observations on Point Accuracy under Forest Canopies. *Photogrammetric Engineering and Remote Sensing*, Vol. 67, No. 9, pp. 1021-1026.

Stehman, Steven V. (1992) Comparison of Systematic and Random Sampling for Estimating the Accuracy of Maps Generated from Remotely Sensed Data. *Photogrammetric Engineering and Remote Sensing,* Vol. 58, No.9, pp. 1343-1350.

Stehman, Stephen V. and Raymond l. Czaplewski (1998). Design and Analysis for Thematic Map Accuracy Assessment: Fundamental Principles. *Remote Sensing of Environment*. 64:331-334.

Van Sickle, Jan (1996). *GPS for Land Surveyors*. Ann Arbor Press, Inc. Chelsea, Michigan.