2.9 Accuracy assessment

Once the aerial photo interpretation team had delineated polygons, an assessment of the thematic accuracy was conducted. Accuracy requirements for the USGS-NPS Vegetation Mapping Program specify 80% accuracy for each map unit. The number of sites visited per map unit depended on how common the map unit was, from 30 polygons for widespread map units to 5 for rare map units. Points were stratified as much as possible by the 3 major areas. Accuracy assessment data were collected in the field during the 1997 and 1998 field seasons.

During 1997 and early 1998 field seasons, the focus was in the northern one-third of the project area. A digital form of the polygon database was not yet available, so the accuracy assessment teams relied on copies of interpreted overlays and the aerial photos to select sites. Data for 552 accuracy assessment sites were collected using this procedure (Figure 10). The accuracy assessment team chose polygons and then hiked to the polygons to determine the community type present. The points were also distributed as evenly as possible across the northern one-third of the project area; e.g. not all of the Aspen-Birch points were taken from the same locale. Teams recorded locational information and briefly described the vegetation, noting canopy structure, dominant species by strata, and environmental features, including topographic position, slope, and aspect. The nearest vegetation communities, if known, within 50 m of the point were also recorded. Rationale for the classification type chosen was explained (Appendix C).

By the second year, much of the polygon data had been digitized, enabling the accuracy assessment team to have a computer-generated selection of sampling points. The majority of the accuracy assessment sites were selected in the bottom two-thirds of the project area. Three times the number of sites needed were generated so that points that were inaccessible or falling near polygon lines could be deleted from the selection. It was recognized that points falling on or near polygon lines could create positional errors in the field because of inherent error in the coordinate readings; PLGR units often have 10 m or more error associated with the readings. Thus, UTM coordinates selected within 10 m of a polygon boundary could actually fall within a neighboring polygon. To help avoid this potential problem, points near polygon lines were eliminated from the sampling scheme. However, some points within 10 m of polygon lines were kept so that narrow polygons would be sampled.

Hard copy 1:12,000-scale orthophoto quadrangle maps, produced from USGS 3.75-minute DOQ's, were plotted with the polygons and final accuracy assessment points overlaying the maps. Each point had a corresponding UTM coordinate that was uploaded into a PLGR GPS unit. The PLGR was used to navigate to each point. The orthophoto maps were used with the GPS unit to help navigate across the terrain. Once the sampling site was reached, the accuracy assessment team assessed the plant community within a 0.5 h radius (the minimum mapping unit) and assigned a provisional community name. Dominant species, environmental data, and rationale for classification were recorded for each site (Appendix C). Data for 736 sites were collected using this method (Figure 9).

The accuracy assessment data (1,288 points) were entered into the PLOTS database and subsequently reviewed for data entry and false errors. Incomplete data on the field sheets, including missing GPS coordinates, resulted in dropping 37 sites from further analysis. Five of these 37 sites were time factor errors. Time factor errors were due to changes in the vegetation community between the time the polygon was mapped and the accuracy assessment was conducted (i.e. exposed mud flats due to lake drawdown, or a recent beaver flooding)

The remaining 1251 points were plotted by their UTM coordinates over the spatial coverage so that all points could be viewed in relation to the vegetation map coverage. USNVC Community Element Global database codes (CEGL) for each accuracy assessment point were compared to corresponding

polygon map unit codes. An output file was created that listed all the accuracy assessment points and their corresponding map codes from the coverage and subsequently run through a SAS routine (SAS Institute 1996) that reported all the matches and mismatches. All mismatches were reviewed to see if there were any "false errors." Mismatches that were deemed "false errors" were corrected, resulting in either a match or a true error.

A false error is defined as a mismatch between a polygon and an accuracy assessment call if the disagreement was caused by one of the following problems:

- A GPS error;
- An accuracy assessment point occurring in a zone of transition between two types (an ecotone);
- An accuracy assessment point that was classified differently than the polygon but was clearly too small to map (an inclusion)

GPS errors occurred when the recorded coordinates were inaccurate and placed points in the wrong polygons. For example, accuracy assessment point VOYA.275, when plotted on the spatial coverage, fell within a Black Ash Forest polygon. However, the data sheet claimed the point fell within a Black Spruce / Alder Rich Swamp. Also, the point fell 2 m from a Black Spruce / Alder Rich Swamp polygon, well within the reported GPS error for this site of plus or minus 7.9 m. Thus, it was revealed that the cause of the error was inaccuracy inherent in the GPS reading (Figure 11). GPS errors were corrected after comparisons between the data sheet, the point's location, and the spatial coverage. Some GPS errors, once corrected, still resulted in a true error.

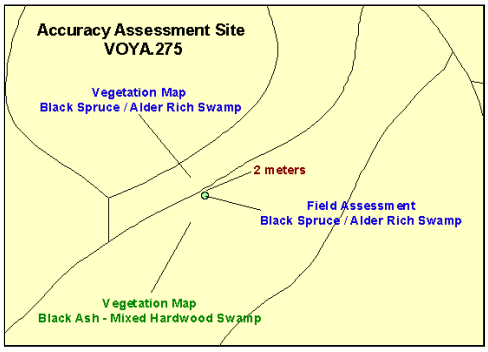

Figure 11. Example of a GPS location error

Ecotone errors occurred when sites fell in transition zones (typically near a polygon boundary) between 2 associations and the accuracy assessment crew acknowledged the presence or mixture of more than one association. When the field crew was uncertain of the correct vegetation type and one of their calls matched the polygon, the mismatch was corrected so that the polygon and accuracy assessment call matched one another.

Inclusion errors were sites where the vegetation associations recorded on the data sheets were different than the associations represented by the mapped polygons AND the accuracy assessment call was determined to represent a small area (< the minimum mapping unit) within the polygon. If it was easily determined as an area too small to map, accuracy assessment site was considered an inclusion and it was assumed that the polygon call was correct.

Map units that represented a single association were assessed together. Map units JPOM, MPHW, and OW were all considered a single association according to the classification: Northern Pin Oak - Bur Oak - (Jack Pine) Rocky Woodland Association. The 3 map units represent different phases of this association.

Three other map unit phase situations were also combined together for the assessment: JPW and JPM map units, representing evergreen and mixed phases of the Boreal Pine Rocky Woodland Association, BSL and BST map units, representing evergreen and mixed phases of Black Spruce / Labrador Tea Poor Swamp Association, and WCS and WCT map units, representing the White Cedar - (Mixed Conifer) / Alder Swamp Association.

PROC FREQ (SAS Institute 1996) was used to compare the final accuracy assessment database with the corresponding polygons. The SAS output was transferred into a contingency table, or error matrix (Appendix D). The error matrix is an array of numbers set in rows and columns which express the number of sample units (polygons) assigned to a particular category relative to the actual category as verified on the ground. The columns represent the accuracy assessment data while the rows represent the mapped polygons. Map accuracy of each category is expressed as a percentage of correctly classified polygons compared to the accuracy assessment results.