Materials and Methods

Planning and Scoping

Several steps were taken to prepare for the vegetation classification and mapping of fire fuel models and vegetation at Delaware Water Gap National Recreation Area. Two planning and review meetings were held on June 4, 2001 and January 9, 2003 with ecologists from the Pennsylvania Natural Heritage Program and National Park Service (NPS) scientists, resource managers, fire specialists, and GIS specialists. The project timeline, mapping product specifications, access issues, natural resource management needs, available applicable GIS layers, and applicable previous and current research conducted at the park were discussed. A third meeting was held on April 25, 2003 with National Park Service Fire Management Program specialists to discuss the desired specifications for the fire fuel model mapping products.

Preliminary Data Collection and Review of Existing Information

Previous studies conducted at Delaware Water Gap National Recreation Area were obtained from the park's natural resource manager and reviewed for information pertinent to the park's vegetation. These reports included previous vegetation cover type mapping and rare species surveys (Myers and Irish 1981; Radis 1986). In addition, shapefiles containing digital information on bedrock geology, soils, streams, waterbodies, wetlands, agricultural fields, roads and trails, park boundaries, and parcel ownership were obtained from the park's GIS specialist.

Aerial Photography Acquisition and Processing

Color infrared, stereo pair 1:12,000 scale aerial photography for a digital orthophoto mosaic of Delaware Water Gap National Recreation Area was acquired from overflights on March 28, April 7, and April 11, 2002, during leaf-off conditions, by Kucera International. The photography, a total of 1,047 air photos that cover the park as well as a relatively large buffer area outside the park, was delivered to the National Park Service, quality checked, accepted as provided, and sent to North Carolina State University (NCSU). Upon receipt at NCSU, the air photos were counted to make sure that none were missing, scanned, and placed in the air photo archive maintained at NCSU for the NPS Northeast Region Inventory & Monitoring Program. Associated data and information provided by Kucera International, also stored in the air photo archive, included the airborne global positioning system and inertial mapping unit (GPS/IMU) data files, the camera calibration certificate for the camera, and the hardcopy flight report for the photography that crosswalks the airborne GPS/IMU data to the photo frame numbers.

The digital aerial photo mosaic was produced from the 151 color infrared air photos that cover the area within the park boundary, scanned at 600 dpi with 24-bit color depth. The scanned images of the air photos were imported into ERDAS Imagine (.img) format where a photo block was created using airborne GPS and IMU data that Kucera International supplied with the aerial photography. The photo block was manipulated until it could be triangulated with a root mean square error of less than 1. At this point, single frame orthophotos (one for each air photo) were generated within Imagine, the orthophotos were mosaicked in Imagine using an algorithm that contained color balancing and cutline information. Finally, the mosaic (.img image) was compressed using MrSID software with a 20:1 compression ratio.

A metadata record for the mosaic was prepared according to current Federal Geographic Data Committee standards (FGDC 1998a). Metadata were produced in notepad and parsed using the U.S. Geological Survey (USGS) metadata compiler program (MP) to locate errors and omissions (USGS 2004). After all errors and omissions were corrected MP was used to generate final TXT, HTML, and XML versions of the metadata record, which are stored in the air photo archive. Key information for the Delaware Water Gap National Recreation Area photo mosaic is summarized in Table 1.

Photointerpretation

After receiving the digital orthophoto mosaic and hard copy photographs from North Carolina State University, ecologists at the Pennsylvania Natural Heritage Program developed a vegetation cover type map that identified broad vegetation types. The vegetation cover types were differentiated based on vegetation structure, leaf phenology, and hydrologic regime because these characteristics can be easily identified on aerial photography. This map was developed as a guide over which the vegetation sampling efforts would be distributed.

The vegetation cover type map was created through aerial photograph interpretation. Aerial photograph interpretation is the act of examining aerial photographs in order to identify objects, in this case, vegetation types (Avery 1978). The diapositive photographs (color infrared, stereo pair, hard copy photographs) were examined through a mirror stereoscope type F-71 which provides the viewer with a three-dimensional view of the photographs. The digital orthophoto mosaic was also examined onscreen in ArcView 3.2 (ESRI 1992–2000a). In addition, digital topographic quad maps were examined in ArcView 3.2. Using information gathered from the diapositives, the mosaic, and the topographic maps, polygons representing different vegetation cover types and land uses were identified. These polygons were digitized onscreen using ArcView 3.2.

Polygons that represented vegetation were attributed with vegetation cover types (Table 2). Nine polygons were labeled as a mosaic of two vegetation cover types, Deciduous Forest and Talus or Scree (i.e., Deciduous Forest / Talus or Scree). These polygons were labeled as mosaics because both types were present in the polygon, but clear boundaries between the two types could not be delineated. Polygons that represented sparse vegetation over naturally occurring geologic features were labeled by substrate type, including Cliff, Cobble Bar, Ridgetop Bedrock Outcrop, and Talus or Scree. These vegetation cover types are not modified Anderson level II categories (see following paragraph) because they could represent sparse vegetation associations that are important to the park.

Polygons that represented other land uses, such as buildings and roads, were attributed with map labels modified from the Anderson level II categories (Anderson et al. 1976). Please note that the Anderson land use and land cover classification (1976) is unrelated to the Anderson fire behavior fuel models (1982) also used in this report. Table 3 lists the modified Anderson level II categories used in the cover type map and shows the relationship between these map labels and the Anderson level I and level II classifications. The Anderson classification was modified in

Table 1. Summary of key information for the Delaware Water Gap National Recreation Area photo mosaic.

Title of metadata record:	Delaware Water Gap National Recreation Area Color Infrared Orthorectified Photomosaic - Spring (leaf-off) (dewa_spring.img and dewa_spring.sid)
Publication date of mosaic (from metadata):	March 1, 2003
Date aerial photography was acquired:	March 28, April 7, and April 11, 2002
Vendor that provided aerial photography:	Kucera International
Scale of photography:	1:12,000
Type of photography:	Color infrared, stereo pairs
Number of air photos in mosaic:	151
Archive location of air photos, airborne GPS/IMU files and camera calibration certificate	North Carolina State University, Center for Earth Observation
Scanning specifications:	600 dpi, 24-bit color depth
Horizontal positional accuracy of mosaics: Accuracy Standard	2.448 meters, meets Class 1 National Map
Number of ground control points upon which estimated accuracy is based:	41
Method of calculating positional accuracy:	Root mean square error (RMSE)
Archive location of mosaics and metadata:	North Carolina State University, Center for Earth Observation
Format(s) of archived mosaics:	.img (uncompressed); MrSID (20:1 compression)

Table 2. Number of polygons, mapped hectares within the park boundary, and number of plots sampled in vegetation cover types and Anderson level II categories (modified) at Delaware Water Gap National Recreation Area.

Vegetation Cover Type	Number of Polygons	Mapped Hectares within the Park Boundary	Number of Plots Sampled
Cliff	112	134.05	8
Cobble Bar	2	0.62	2
Deciduous Forest	1,163	12,072.77	67
Deciduous Forest / Talus or Scree	9	31.45	1
Deciduous Plantation	4	10.58	0
Deciduous Shrubland	169	534.82	7
Deciduous Woodland	63	131.45	5
Evergreen Forest	330	2,079.50	20
Evergreen Plantation	51	158.34	2
Evergreen Woodland	120	201.37	5
Grassland	229	520.18	11
Mixed Evergreen - Deciduous Forest	838	5,946.35	43
Mixed Evergreen - Deciduous Shrubland	164	583.53	11
Mixed Evergreen - Deciduous Woodland	127	330.95	6
Ridgetop Bedrock Outcrop	86	61.21	5
Saturated Herbaceous Vegetation	45	86.06	13
Seasonally Flooded Deciduous Forest	23	57.98	2
Seasonally Flooded Mixed Evergreen - Deciduous Forest	4	6.33	1
Seasonally Flooded Mixed Evergreen - Deciduous Shrubland	11	31.16	3
Semipermanently Flooded Deciduous Shrubland	10	20.05	3
Semipermanently Flooded Herbaceous Vegetation	20	26.04	5
Talus or Scree	28	28.19	0
Temporarily Flooded Deciduous Forest	150	882.03	13
Temporarily Flooded Deciduous Shrubland	16	58.12	7
Temporarily Flooded Deciduous Woodland	30	98.35	2
Temporarily Flooded Herbaceous Vegetation	54	65.72	7
Temporarily Flooded Mixed Evergreen - Deciduous Forest	24	194.00	2
Anderson level II category (modified)			
Built-up Land	131	351.97	0
Cropland	205	1,449.10	0
Pond	246	294.46	0
Right-of-way	22	56.94	0
River	12	1,184.97	0
Shale Pit	10	12.73	0
Transportation Corridor	12	130.98	0
Total	4,520	27,832.33	251

Table 3. Relationship between Anderson level II category (modified) map labels and the Anderson level I and II classifications.

Anderson level II (modified) map label	Anderson level I	Anderson level II	Anderson level II number
Built-up Land	Urban or Built-up Land	Residential	11
		Commercial and Services	12
		Industrial	13
		Industrial and Commercial Complexes	15
		Mixed Urban or Built-up Land	16
		Other Urban or Built-up Land	17
Cropland	Agricultural Land	Cropland and Pasture	21
Pond	Water	Lakes	52
		Reservoirs	53
Right-of-way	Urban or Built-up Land	Transportation, Communications, and Utilities	14
River	Water	Streams and Canals	51
Shale Pit	Barren Land	Strip Mines, Quarries, and Gravel Pits	75
Transportation Corridor	Urban or Built-up Land	Transportation, Communications, and Utilities	14

order to provide more specific information to the park and to avoid photo interpretation of human uses of various building types. For example, instead of Transportation, Communications, and Utilities, the map label Transportation Corridor was used for all roads and highways, while Right-of-way was used to label utility and communication corridors. Also, the use of specific buildings within the park was not distinguished by Residential, Commercial, Industrial, etc, but rather labeled with the generic map unit, Built-up Land. Given the focus on vegetation and fire fuel model mapping in this project, it was inappropriate to spend additional time mapping building uses and testing the accuracy of such map attributes.

To determine which vegetation cover types or modified Anderson level II categories should be assigned to each polygon, an aerial photography interpretation key (Appendix A) was used. The resulting vegetation cover type map (Figure 4) identified 4,520 map polygons, each labeled with a vegetation cover type or a modified Anderson level II category (Table 2). This vegetation cover type map was used to guide vegetation plot sampling in the park.

Field Data Collection and Classification

All vegetation plot sampling followed the USGS/NPS Vegetation Mapping Program protocols (TNC and ESRI 1994b). The Delaware Water Gap National Recreation Area is considered a large park (278 km^2 [107 mi^2]) for which the gradient directed transect sampling (gradsect) approach should be used. The gradsect approach is a variant of a stratified random sampling strategy that intends to efficiently describe the full range of vegetation by sampling along the full range of environmental variability (Gillison and Brewer 1985; Austin and Adomeit 1991). Gradsects are areas of the park selected for sampling that contain the strongest environmental gradients and that are reasonably accessible. This strategy optimizes the amount of data collected relative to the time and effort spent during vegetation sampling. The total area of the gradsects should include at least 15% of the park area (TNC and ESRI 1994b).

In order to select gradsect areas several environmental gradients and factors were examined, including elevation, slope, aspect, and geology. Due to the orientation of the Delaware Water Gap National Recreation Area around the Delaware River these factors form strong environmental gradients from the river to the ridgetops in both Pennsylvania and New Jersey. Thirteen potential gradsects, stretching from rivershore to the park boundary, were identified. These gradsects were distributed along the length of the park and on both sides of the river in order to capture environmental gradients between the southern and northern ends of the park, as well as the eastern and western sides of the river. For each of these potential gradsects, land ownership, accessibility, bedrock geology, and diversity of vegetation cover types were then assessed. Seven gradsects that contained publicall owned land, reasonable access, and representative diversity of bedrock geology and vegetation cover types were selected for sampling (Figure 5). Three gradsects were located in Pennsylvania and four occurred in New Jersey. These gradsects encompassed 5,377 ha (13,281 ac) in total, covering 19% of the park area.

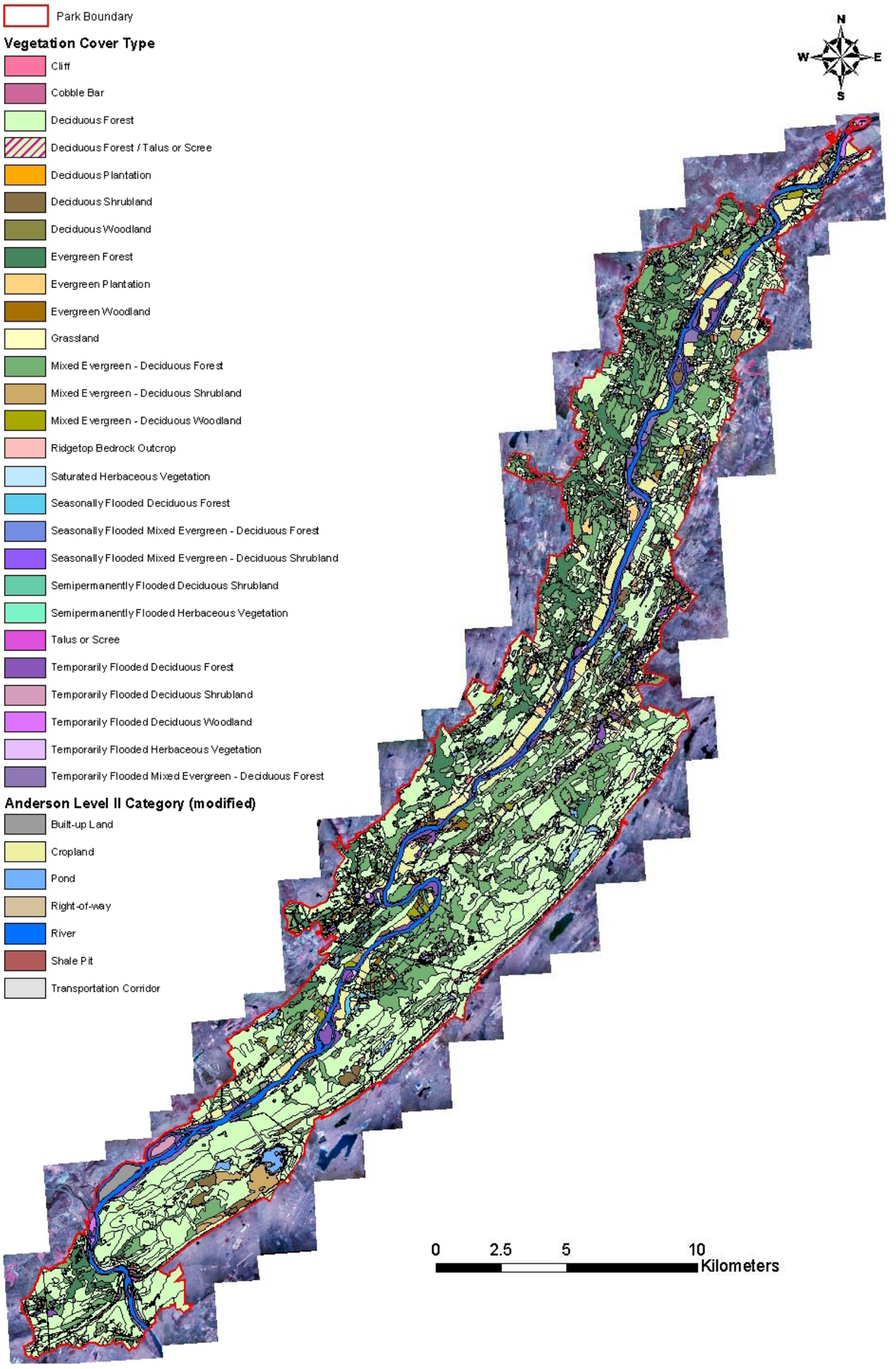

Figure 4. Vegetation cover types and Anderson level II categories (modified) delineated in Delaware Water Gap National Recreation Area in order to guide vegetation plot sampling.

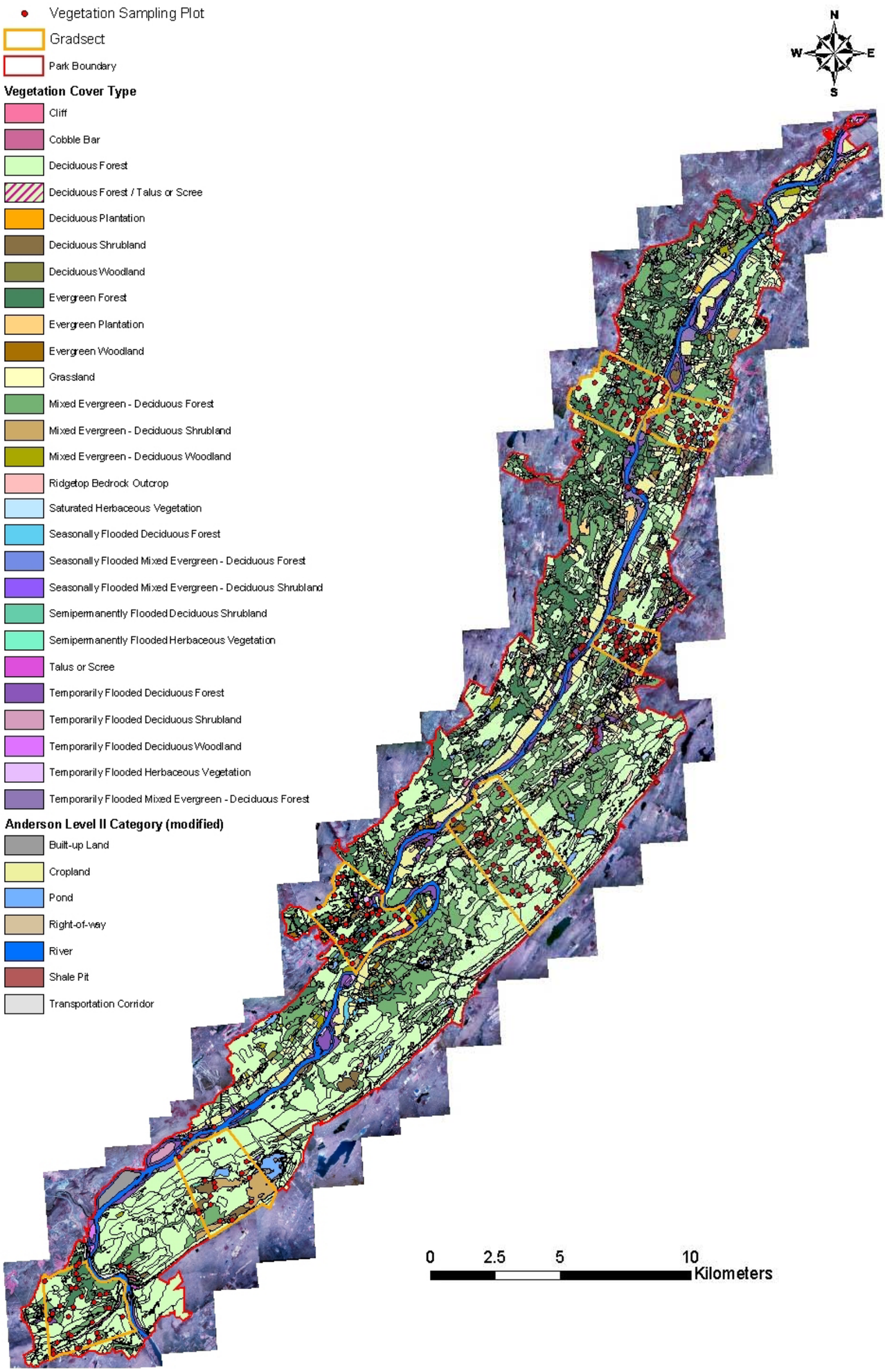

Figure 5. Locations of gradsects and vegetation plots sampled in Delaware Water Gap National Recreation Area for vegetation classification and mapping.

To determine the number of plots that should have been sampled in each gradsect the vegetation cover type map was used as a guide. The number of vegetation cover types and the number of polygons per cover type in each gradsect was examined. The USGS/NPS Vegetation Mapping Protocol suggests that each vegetation association should have been sampled at least three times in order to capture the naturally occurring variation within the park (TNC and ESRI 1994b). Therefore, in each gradsect, at least three polygons of each vegetation cover type were selected for sampling. For vegetation cover types with less than three polygons, each polygon was selected for sampling. For cover types with numerous polygons, or for types that covered large areas, polygons were selected that represented different aerial photograph signatures and environmental settings in order to capture the natural range of variability of the vegetation. Polygons that represented modified Anderson level II categories were not sampled.

In addition to the gradsect sampling, vegetation cover types not adequately sampled by the gradsect method were targeted. Temporarily flooded and semipermanently flooded vegetation cover types on Delaware River islands and shoreline were accessed by canoe and sampled over a three-day period. Rare wetland and ridgetop communities identified in previous surveys were visited and sampled during six additional days. Towards the end of the second field season of plot sampling the distribution of vegetation cover types sampled was compared to the distribution of all vegetation cover types in the park. Cover types that were underrepresented by the plot sampling were targeted for additional field visits. During these field visits it was determined that most of the vegetation encountered were already adequately sampled; however, some additional plots were sampled in vegetation that differed from previously sampled plots.

Field Survey

Within each polygon selected for sampling a plot was established in an area that was most representative of the existing vegetation association (Mueller-Dombois and Ellenberg 1974). All vegetation data were collected following NatureServe's accepted natural heritage sampling protocols (Strakosch-Walz 2000), with 20-m × 20-m plots in forests and woodlands, 10-m × 10-m plots in shrublands, and 5-m × 5-m plots in herbaceous vegetation. The plot sampling data form used in this project is shown in Appendix B. The vegetation was visually divided into eight strata: emergent trees (variable height), tree canopy (variable height), tree subcanopy (>5m in height), tall shrub (2–5m), short shrub (<2m), herbaceous, non-vascular, and vines. The percent cover was estimated for each species in each stratum using modified Braun - Blanquet cover classes (Strakosch-Walz 2000). Specimens of species that were not identifiable in the field were collected for later identification. The diameter at breast height (dbh 1.3 m) was measured with a Biltmore stick or a diameter tape for all trees larger than 10 cm in diameter that were rooted in the plot. The diameters were recorded by species and strata.

In addition to floristic information, other environmental variables recorded at each plot included slope, aspect, topographic position, hydrologic regime, soil stoniness, average soil texture, and soil drainage. Any unvegetated area of the plot was characterized by the exposed substrate. Notes were taken on the plot representativeness of the surrounding vegetation and any other significant environmental information, such as landscape context, herbivory, stand health, recent disturbance, or evidence of historic disturbance. The vegetation profile and topographic position were sketched in cross-section to represent the location and setting of the plot.

The vegetation in each plot was assigned one of the Anderson (1982) fire fuel models and justification for the assignment was written on the field form. Canopy cover in the plot was assigned to one of four canopy cover classes (1=1–20%; 2=21–50%; 3=51–80%; 4=81–100%) required by the FARSITE software. In forest stands with sufficient ladder fuels to ignite a crown fire, the average crown height, average height to live crown, and average crown width were measured with a Biltmore stick or meter tape and recorded. However, only one of the 251 plots contained forest with sufficient ladder fuels to ignite a crown fire.

A digital photograph of each plot was also taken. The location of each plot was recorded with one of the following global positioning system (GPS) units: Garmin III+ in conjunction with CSI MBX-3 2 channel Automatic Differential Beacon Receiver, or Trimble Pocket Pathfinder in conjunction with Beacon-on-a-Belt and a Compaq Ipaq Pocket PC interface. The datum on the GPS unit was set to North American 1983 (Conus) and the coordinate system was set to Universal Trans-Mercator (UTM) zone 18.

Plot sampling was conducted from May through September in 2003 and 2004. In total, 251 plots were sampled throughout the park (Figure 5). The distribution of plots by vegetation cover type is shown in Table 2. No polygons labeled Talus or Scree were sampled because the majority of these polygons are inacessable and/or dangerous. However, sandstone talus included in this type was sampled in a mosaic polygon, Deciduous Forest / Talus or Scree. No polygons labeled Deciduous Plantation were sampled; however, these larch (*Larix* sp.) plantations were captured in some of the Mixed Evergreen - Deciduous Plantation polygons that were sampled.

Vegetation Classification and Characterization

Data from the 251 vegetation plots were entered into the NatureServe PLOTS 2.0 Database System on a Microsoft Access platform during winter 2003–2004 and winter 2004–2005. In the PLOTS 2.0 database, species were assigned standardized codes based on the *PLANTS Database, Version 3.5* developed by the Natural Resource Conservation Service in cooperation with the Biota of North America Program (USDA, NRCS 2004). For this report, some common names listed in the PLANTS Database were changed to reflect the common names typically used by ecologists and resource managers in this region. The common and scientific names of plants observed during the vegetation plot sampling are listed in Appendix C. Some tree and shrub seedlings and immature herbaceous plants could only be identified to the genus level and are therefore listed in the appendix as such. Environmental variables and species percent cover data were exported from the PLOTS database into Excel in order to be manipulated into a format compatible with PC-ORD version 4.0 Multivariate Analysis software (McCune and Mefford 1999).

The vegetation plot data were analyzed using several multivariate statistical techniques available in the PC-ORD software. In general, the analyses were designed to progressively fragment the dataset into more workable subsets using cluster analysis, two-way indicator species analysis, and non-metric multidimensional ordination analysis. Different techniques were employed to provide multiple lines of evidence from which to interpret the results. For a detailed discussion of the statistical techniques used in this study, please refer to McCune and Grace (2002). The data analysis and interpretation process was iterative in order to identify and analyze increasingly finer groups until vegetation associations were characterized.

Cluster analyses were performed using the percent cover of species data. This agglomerative analysis produces a hierarchical classification of the plots based on the similarity in their species composition. Euclidean distance measure and Ward's group linkage method were used in this analysis. Two-way indicator species analyses (TWINSPAN) were also performed using the percent cover of species data. This divisive analysis successively divides the plots into groups that are similar in species composition (Hill and Gauch 1979). Non-metric multidimensional ordination analyses (NMS) were performed using both the percent cover of species and the environmental variables from the plots. NMS is an ordination technique well suited to non-normal data sets (Kruskgal and Wish 1978). In this analysis Sorensen distance measure, a random starting configuration, and a stability criterion of 0.005 were employed. Forty runs were performed with the real data, with a maximum of 400 iterations.

The results of the classification analyses were then evaluated and compared to the Terrestrial and Palustrine Plant Communities of Pennsylvania (Fike 1999), Draft New Jersey Ecological Community Crosswalk (Walz et al. 2006), vegetation classifications at other national parks, and the ecologists' field experiences at the Delaware Water Gap National Recreation Area. Based on this evaluation, preliminary vegetation associations were identified. Compositional statistics were calculated to evaluate the consistency of associations and to assist in naming and describing the community types. Total and relatve cover, total and relative frequency, and relative importance value of each species were calculated by association. To compute total and relative cover the cover classes used in the field were converted to the midpoints of their respective percent ranges. Relative importance value of each species was calculated by averaging the relative cover and the relative frequency of that species.

After local, park-specific descriptions had been developed for the preliminary vegetation associations, based on the analyses described above, these associations were then crosswalked to the National Vegetation Classification System (NVCS). The NVCS was developed by ecologists of the Natural Heritage Program network and The Nature Conservancy after many years of literature review, data collection, and data anlysis. This collaborative effort culminated in the publication of *International Classification of Ecological Communities: Terrestrial Vegetation of the United States* (Grossman et al. 1998). The International Classification of Ecological Communities, now known as the International Vegetation Classification, of which the NVCS is a subset, has been revised and refined since 1998 and is now managed by NatureServe in continued collaboration with the network of Natural Heritage Programs. The classification is housed in the Biotics database and is updated regularly. The upper levels of the NVCS were adopted as a standard by the Federal Geographic Data Committee to support the production of uniform statistics on vegetation at the national level (FGDC 1996). The Vegetation Mapping Program of the National Park Service adopted the alliance level and, where possible, the association level, as the mapping unit for national parks.

The crosswalk to the NVCS for each association was determined through qualitative comparison of the preliminary park-specific local vegetation associations to existing associations in the National Vegetation Classification System by searching for alliances sharing similar dominant species, as well as physiognomy and environmental setting. Total floristic composition was used to determine the appropriate association within the alliance. Global information on the associations from the NVCS was then appended to the local descriptions to provide resource managers with a broader context for the vegetation in the park.

Each preliminary vegetation association was assigned a common name based on the Terrestrial and Palustrine Plant Communities of Pennsylvania (Fike 1999) or Draft New Jersey Ecological Community Crosswalk (Walz et al. 2006). If no appropriate name existed in Fike (1999) or Walz et al. (2006), the National Vegetation Classification System common name was used, or a park-specific common name was created in the case of successional and cultural vegetation types not easily handled by Fike (1999), Walz et al. (2006), or the NVCS.

A park-specific dichotomous key was created for the vegetation associations to guide accuracy assessment and for use by park natural resource managers and others (Appendix D). A dichotomous key is a tool for identifying unknown entities—in this case, vegetation associations. It is structured by a series of couplets, two statements that describe different, mutually exclusive characteristics of the associations. Choosing the statement that best fits the association in question leads the user to the correct association. The dichotomous key should be used in conjunction with the detailed vegetation association descriptions to confirm that the association selected with the key is appropriate.

Map Preparation

Following the vegetation data analysis, the vegetation cover-type map was edited and refined to develop a preliminary association-level vegetation map. Using ArcView 3.2, polygon boundaries were revised onscreen based on the plot data, field observations, classification analyses, aerial photography signatures, and topographic maps. Each polygon was assigned the name of a preliminary vegetation association based on the five information sources listed above. A mirror stereoscope type F-71 and a Bausch and Lomb zoom stereoscope were used to interpret the aerial photography signatures. An aerial photograph interpretation key for the vegetation associations and Anderson level II categories (modified) is located in Appendix A. Numerous polygons were labeled as mosaics of two associations because both types were present in the polygons, but clear boundaries between the two associations could not be delineated (i.e., Calcareous Riverside Outcrop / Calcareous Riverside Seep). These mosaics are not unique associations themselves, but rather indicate that two associations occur and intergrade within the polygon. Polygons labeled with Anderson level II categories (modified) retained their attributes in the vegetation association map. The only exceptions were the polygons labeled Right-of-way, which were assigned the names of preliminary vegetation associations in the association map.

In addition to preliminary vegetation association names, each polygon was assigned a fire behavior fuel model and a canopy cover class. Fuel model and canopy cover class were assigned based on the plot data, field observations, aerial photography signatures, topographic maps, and preliminary vegetation associations. The Anderson fire fuel behavior models that were most appropriate for the vegetation in the park were used. Park-specific descriptions of the fire fuel models, the relationship between the fuel models and the vegetation associations, and the original fuel model descriptions from Anderson (1982) are provided in Appendix E. The standard canopy cover classes defined by the FARSITE model were used to create the canopy cover class map. Canopy cover is defined as the horizontal percentages of the ground surface that is covered by tree crowns (Finney 1998). The cover classes are: Class 1: 1–20% of horizontal area covered by tree canopy; Class 2: 21–50% of horizontal area covered by tree canopy; Class 3: 51–80% of horizontal area covered by tree canopy; and Class 4: 81–100% of horizontal area covered by tree canopy.

Accuracy Assessment

Two sources of potential error in the vegetation map include: 1) horizontal positional accuracy, in which a location on the photomosaic does not accurately align with the same location on the ground due to errors in orthorectification or triangulation; and 2) thematic accuracy, in which the vegetation type assigned to a particular location on the map does not correctly represent the vegetation at the same location in the park due to mapping error. The USGS/NPS Vegetation Mapping Program protocols (TNC and ESRI 1994c) were followed to assess the positional accuracy of the digital photo mosaic and the thematic accuracy of the vegetation association, fire behavior fuel model, and canopy cover class maps.

Positional Accuracy Assessment

Well-defined positional accuracy ground control points, spaced throughout all quadrants of the mosaic, were placed on the final mosaic in ArcMap. Ground control points and zoomed-in screenshots of each point were plotted on hard copy maps with the mosaic as a background. These maps and plots were used to locate the ground control points in the field. For each plotted ground control point, field staff noted any alterations to the location in the field, and then recorded the coordinates with a real-time differentially corrected Garmin III+ in conjunction with CSI MBX-3 2 channel Automatic Differential Beacon Receiver. Mapped ground control points that were physically inaccessible were also noted. The field crew correctly located and collected accuracy assessment data at 42 ground control points. Prior to calculating accuracy, one ground control point was identified as an outlier with SAS's JMP program and removed. The field-collected "true" or "reference" GPS coordinates for the remaining 41 points were compared to the coordinates obtained from the mosaic viewed in ArcMap. Both pairs of coordinates for each point were entered into a spreadsheet in order to calculate horizontal accuracy (in meters). The accuracy calculation formula is based on root mean square error (FGDC 1998b; Minnesota Governor's Council on Geographic Information and Minnesota Land Management Information Center 1999). Figure 6 shows the distribution of the 41 ground control points within the park and surrounding area.

Thematic Accuracy Assessment

The thematic accuracies of the preliminary vegetation association map, the fire behavior fuel model map, and the canopy cover class map were assessed by the Pennsylvania Natural Heritage Program. A stratified random sampling approach was used, distributing the sampling effort across the preliminary vegetation association map. Since the vegetation association map contained many more types to be assessed than the fire fuel model or the canopy cover class map, one set of random sampling points could be used to assess the accuracy of all three maps. This combined sampling strategy was inherently more efficient and resulted in adequate sampling of the vegetation map and over sampling of the fire fuel model and canopy cover class map.

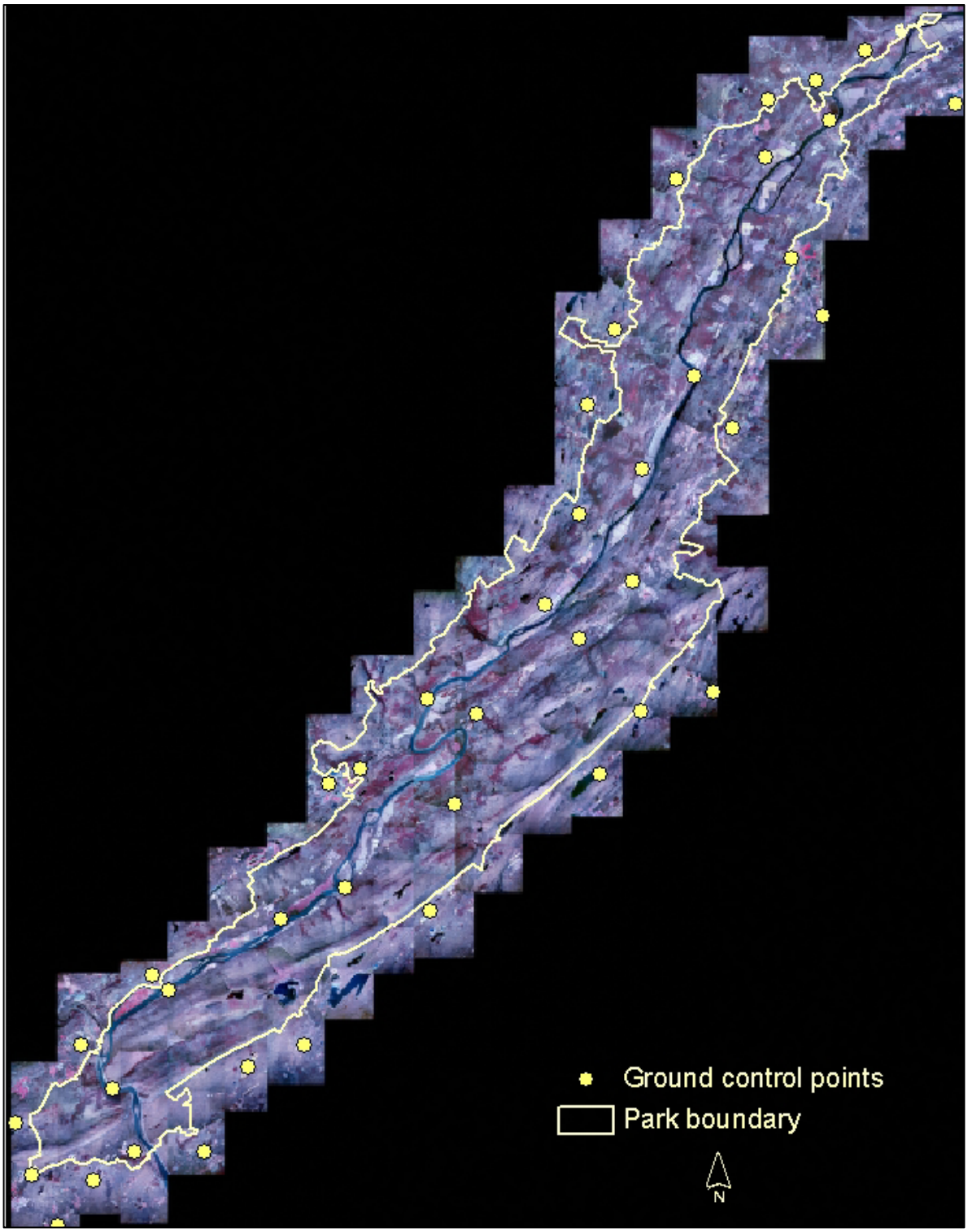

Figure 6. Ground control points used to calculate horizontal positional accuracy of the Delaware Water Gap National Recreation Area mosaic.

In this stratified random sampling design, the sampling effort was distributed across the preliminary vegetation associations (Table 4). Polygons labeled with modified Anderson level II categories were not included in the thematic accuracy assessment sampling. Two associations, Calcareous Seep and Water-willow Emergent Bed, were excluded from the accuracy assessment because they could not be mapped due to small patch size and the timing of the aerial photography, respectively.

The names of vegetation associations listed in Table 4 are from the preliminary classification determined before thematic accuracy assessment. Numerous changes were made to the classification after accuracy assessment; therefore, the preliminary association names in Table 4 do not fully match the final association descriptions presented in this report. For the thematic accuracy assessment sampling strategy, polygons labeled as a mosaic of two vegetation associations were lumped with the least abundant of the component associations. Since mosaics are not unique associations themselves, they should not be sampled separately in the accuracy assessment. Lumping the mosaics with the least abundant association favored additional sampling in less abundant associations; thus, increasing the sampling power.

In the stratified random sampling approach, the number of samples per association varied according to the rarity of the vegetation association, both in terms of number of polygons and polygon size. The following rules were used to determine the number of points assigned to each association (TNC and ESRI 1994c):

Scenario A: The class is abundant. It covers more than 50 hectares in total area and consists of at least 30 polygons. In this case, it is recommended that 30 polygons be selected at random from the set of the association's polygons. One sampling point will be assigned to each of the 30 selected polygons.

Scenario B: The association is relatively abundant. It covers more than 50 hectares in total area but consists of fewer than 30 polygons. In this case, it is recommended that 20 polygons be selected at random from the set of the association's polygons, and that one sampling point be assigned to each of the 20 selected polygons. If the association contains less than 20 polygons, some polygons will contain multiple sampling points. The number of sampling points assigned to each polygon is determined by the relative area of that polygon compared with the other polygon in that association.

Scenario C: The association is relatively rare. It covers less than 50 hectares in total area but consists of more than 30 polygons. In this case, it is recommended that 20 polygons be selected at random from the set of the association's polygons. One sampling point will be assigned to each of the 20 selected polygons.

Scenario D: The class is rare. It has 5–30 polygons and covers less than 50 hectares. In this case, it is recommended that five polygons be selected at random from the set of the association's polygons. One sampling point will be assigned to each of the five selected polygons.

Scenario E: The association is very rare. It has fewer than five polygons and occupies less than 50 hectares of the total area. In this case, it is recommended that one sampling point be assigned to each polygon.

Table 4. Thematic accuracy assessment (AA) sampling strategy for Delaware Water Gap National Recreation Area.

Preliminary Vegetation Association[1]	Number of Polygons	Mapped Hectares within Park Boundary	Number of AA Points Sampled[2]
Acidic Seep	1	0.09	1
Alder Wetland	2	2.21	2
Bear Oak - Wavy Hairgrass Shrubland	18	15.21	5
Bear Oak - Wavy Hairgrass Shrubland / Dry Oak - Heath Woodland	2	8.82	
Big Bluestem - Indian Grass River Grassland	5	2.43	5
Bitternut Hickory Lowland Forest	3	15.64	3
Bottomland Mixed Hardwood Palustrine Forest	25	99.84	20
Bottomland Mixed Hardwood Palustrine Forest / Silky Dogwood - Swamp Rose Wetland	3	6.55	
Bottomland Oak Palustrine Forest	9	49.98	5
Buttonbush Wetland	4	7.12	4
Calcareous Fen	3	3.61	3
Calcareous Riverside Outcrop / Calcareous Riverside Seep	4	2.90	4
Conifer Plantation	202	463.61	30
Dry Eastern Hemlock - Oak Forest	91	831.61	30
Dry Eastern White Pine - Oak Forest	250	1,659.59	30
Dry Hickory Ridgetop Forest	13	144.89	20
Dry Oak - Heath Forest	222	4,015.13	30
Dry Oak - Heath Woodland	43	159.20	30
Dry Oak - Mixed Hardwood Forest	215	2,877.19	30
Eastern Hemlock - Mixed Hardwood Palustrine Forest	4	20.56	4
Eastern Hemlock - Northern Hardwood Forest	161	1,166.83	30
Eastern Hemlock Forest	66	1,315.10	30
Eastern Red-cedar (Pitch Pine) - Prickly Pear Shale Woodland	81	68.30	30
Eastern Red-cedar Forest	16	23.37	5
Eastern White Pine - Successional Hardwood Forest	69	614.23	30
Eastern White Pine Forest	99	375.52	30
Hickory - Eastern red-cedar Rocky Woodland	22	15.16	5
Hickory - Eastern red-cedar Rocky Woodland / Dry Oak - Heath Woodland	1	18.29	
Hickory - Eastern red-cedar Rocky Woodland / Sandstone Talus	1	4.02	
Highbush Blueberry - Steeplebush Wetland	16	14.51	5
Japanese Knotweed Herbaceous Vegetation	11	6.29	5
Reed Canarygrass Riverine Grassland / Japanese Knotweed Herbaceous Vegetation	1	0.64	
Leatherleaf Peatland	1	0.62	3
Leatherleaf Peatland / Red Maple - Black Spruce - Highbush Blueberry Palustrine Woodland	2	6.99	
Little Bluestem Grassland	10	10.80	5
Marl Fen	1	0.19	1
Mixed Forb Marsh	29	52.14	30
Mixed Forb Marsh / Wet Meadow	1	1.91	
Pond / Mixed Forb Marsh	1	0.45	
Cattail Marsh	1	0.55	
Northeastern Modified Successional Forest	486	1,998.80	30
Northern Red Oak - Mixed Hardwood Forest	79	871.24	30
Oak Talus Forest	45	390.28	30
Dry Oak - Heath Forest / Oak Talus Forest	8	167.39	
Dry Oak - Mixed Hardwood Forest / Oak Talus Forest	3	144.16	
Old Field	253	543.29	30
Modified Successional Forest / Old Field	1	1.83	
Successional Shrubland / Old Field	1	0.52	

Table 4. Thematic accuracy assessment (AA) sampling strategy for Delaware Water Gap National Recreation Area (continued).

Preliminary Vegetation Association[1]	Number of Polygons	Mapped Hectares within Park Boundary	Number of AA Points Sampled[2]
Pitch Pine - Mixed Hardwood Rocky Summit	62	32.41	30
Dry Oak - Mixed Hardwood Forest / Pitch Pine - Mixed Hardwood Rocky Summit	1	19.74	
Dry Oak - Heath Forest / Pitch Pine - Mixed Hardwood Rocky Summit	4	43.15	
Red Maple - Black Spruce - Highbush Blueberry Palustrine Woodland	8	21.81	5
Red Maple - Highbush Blueberry Palustrine Forest	29	34.63	5
Red Maple - Sweet Birch Hardwood Forest	146	1,556.53	30
Red Maple Palustrine Forest	54	115.80	30
Reed Canarygrass Riverine Grassland	12	28.56	5
Riverine Scour Vegetation	30	20.91	20
Riverine Scour Vegetation / Modified Successional Forest	1	7.16	
Sandstone Talus	26	18.62	30
Oak Talus Forest / Sandstone Talus	7	40.67	
Shale Scree Slope	13	16.60	5
Shale Scree Slope / Eastern Red-cedar (Pitch Pine) - Prickly Pear Shale Woodland	1	1.32	
Silky Dogwood Successional Palustrine Shrubland	28	47.36	30
Silky Dogwood Successional Palustrine Shrubland / Wet Meadow	7	15.08	
Silver Maple Floodplain Forest	76	538.84	30
Sparsely Vegetated Cliff	14	19.15	30
Hickory - Eastern Red-cedar Rocky Woodland / Sparsely Vegetated Cliff	24	36.64	
Oak Talus Forest / Sparsely Vegetated Cliff	1	0.60	
Successional Eastern Red-cedar (Eastern White Pine) Woodland	133	216.02	30
Successional Bear Oak - Heath Shrubland	30	285.18	30
Successional Bear Oak - Heath Shrubland / Dry Oak - Heath Woodland	1	10.24	
Successional Mixed Hardwood Forest	43	250.23	30
Successional Shrubland	364	904.45	30
Modified Successional Forest / Successional Shrubland	12	51.17	
Sugar Maple - American Beech - Sweet Birch Forest	113	1,632.95	30
Sugar Maple - American Basswood Forest	7	66.56	20
Sugar Maple Floodplain Forest	14	128.17	20
Sycamore (Willow) - Mixed Hardwood Riverine Dwarf Shrubland	12	7.56	5
Sycamore - Mixed Hardwood Floodplain Forest	30	65.78	30
Sycamore - Mixed Hardwood Riverine Shrubland	13	10.04	5
Sycamore - Mixed Hardwood Riverine Shrubland / Riverine Scour Vegetation	1	4.14	
Sycamore Floodplain Forest	5	9.77	5
Tuliptree - Beech - Maple Forest	12	156.26	20
Tussock Sedge Marsh	15	10.57	5
Silky Dogwood Successional Palustrine Shrubland / Tussock Sedge Marsh	1	0.93	
Red Maple Palustrine Forest / Tussock Sedge Marsh	1	0.47	
Wavy Hairgrass - Common Sheep Sorrell Rock Outcrop	27	6.75	5
Wet Meadow	49	60.76	30
Wet Meadow / Old Field Vegetation	1	1.98	
Old Field Vegetation / Wet Meadow	5	19.78	
Successional Shrubland / Wet Meadow	1	2.13	
Total	4,013	24,695.11	1,130

[1] Names of vegetation associations listed in the table are from the preliminary classification determined before thematic accuracy assessment. Numerous changes were made to the classification after accuracy assessment, therefore the preliminary association names in this table do not fully match the final association descriptions presented in this report.
[2] Number of points determined by USGS/NPS Vegetation Mapping Program protocol (The Nature Conservancy and Environmental Systems Research Institute 1994c).

In order to randomly select the polygons in Scenarios A, B, C, and D, the Create Random Selection tool in Hawth's Analysis Tools was used in ArcGIS (Beyer 2004; ESRI 1999–2000b). For all of the scenarios, the Generate Random Points tool in Hawth's Analysis Tools was used to randomly determine the location of the sampling points in the polygon. If the randomly selected polygon or point fell on inaccessible privately owned lands the point was reassigned to a randomly selected polygon of the same association that fell on publicly owned land. The resulting 1,130 thematic accuracy assessment sampling points are shown in Figure 7.

These sampling points were also used to assess the thematic accuracy of the fire behavior fuel model map and the canopy cover class map. Because the vegetation association map contained many more types than either the fuel model or the cover class maps, the sampling design for the vegetation association map was also sufficient for the fire behavior fuel model map and the canopy cover class map. Even though the sampling design was created for the vegetation association map, the design distributed the sampling points across the fuel model and cover class map units proportionally to the number of polygons and the total mapped hectares of the units (Tables 5 and 6). For the fire behavior fuel models, five of the 12 models were oversampled, while seven of the 12 models were undersampled, according to the USGS/NPS Vegetation Mapping Program protocol (TNC and ESRI 1994c). For the canopy cover class map, all four classes were oversampled, yielding approximately 3–15 times the number of points recommended by the USGS/NPS Vegetation Mapping Program protocol (TNC and ESRI 1994c).

From July through September 2005 and April through July 2006, each accuracy assessment point was located in the field using one of the following GPS units: Trimble Pocket Pathfinder attached to a Beacon-on-a-Belt with a Compaq Ipaq Pocket PC interface, or Garmin GPS map 76 WAAS enabled. The datum on the GPS units was set to North America 1983 (Conus) and the coordinate system was set to Universal Trans-Mercator (UTM) zone 18. At the accuracy assessment point, the vegetation association at that location was determined using the dichotomous key and the detailed vegetation descriptions. The fire behavior fuel model appropriate for the area's vegetation was recorded. The percent canopy cover and the canopy cover class of the area were also noted. The minimum area of observation around the sampling point was a circle with a radius of 50 m. The accuracy assessment data form used in this study is shown in Appendix F.

Data from the 1,130 accuracy assessment points were entered into the NatureServe PLOTS 2.0 Database System on a Microsoft Access platform from November 2005 through July 2006. In the PLOTS database, species were assigned standardized codes based on the *PLANTS Database, Version 3.5* (USDA, NRCS 2004). For this report, some common names listed in the PLANTS Database were changed to reflect the common names typically used by ecologists and resource managers in this region. The common and scientific names of plants observed during thematic accuracy assessment sampling are listed in Appendix C. Some tree and shrub seedlings and immature herbaceous plants could only be identified to the genus level and are therefore listed in the appendix as such.

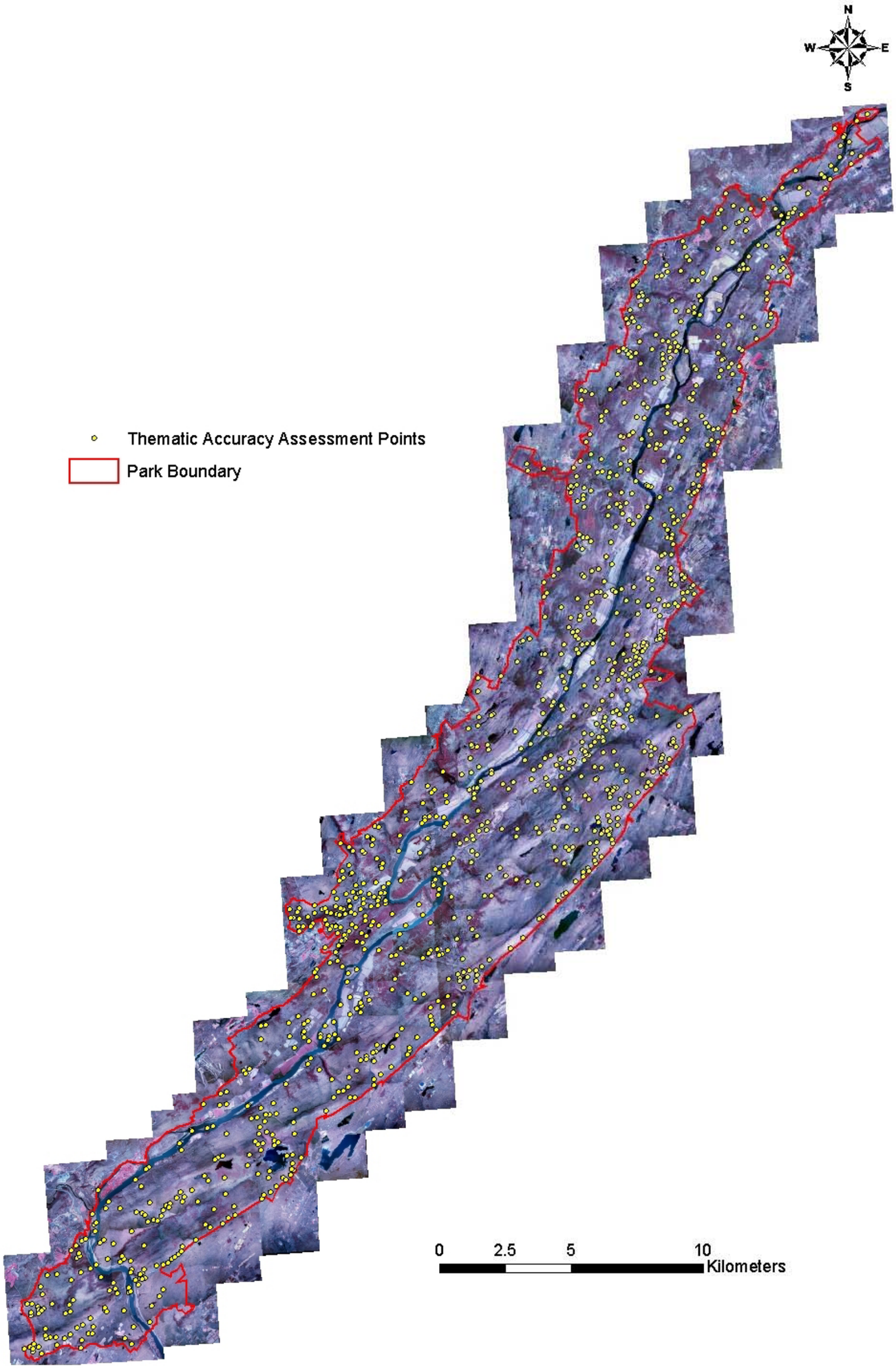

Figure 7. Locations of thematic accuracy assessment sampling points in Delaware Water Gap National Recreation Area.

Table 5. Distribution of thematic accuracy assessment sampling points across fire behavior fuel models in Delaware Water Gap National Recreation Area.

Fire Behavior Fuel Model	Number of Polygons	Mapped Hectares within Park Boundary	Number of Points Recommended by Protocol[1]	Number of AA Points Sampled
0	1,090	4,127.77	30	383
1	141	152.62	30	23
1/3	669	1,474.60	30	81
3	48	144.12	30	20
4	82	496.05	30	10
5	113	317.44	30	12
6	306	3,757.91	30	77
6/7	63	1,498.89	30	17
7	76	462.13	30	24
8	64	383.36	30	26
8/9	1,661	14,355.15	30	402
9	266	748.08	30	55
Total	4,579	27,918	360	1,130

[1] The Nature Conservancy and Environmental Systems Research Institute. 1994 (c). NBS/NPS Vegetation Mapping Program: Accuracy Assessment Procedures. 71pp. Report to the National Biological Survey and the National Park Service. Arlington, VA and Redlands, CA. <http://biology.usgs.gov/npsveg/standards.html> Last accessed 17 March 2005.

Table 6. Distribution of thematic accuracy assessment sampling points across canopy cover class map units in Delaware Water Gap National Recreation Area.

Canopy Cover Class	Number of Polygons	Mapped Hectares within Park Boundary	Number of Points Recommended by Protocol[1]	Number of AA Points Sampled
1	1,632	5,120.25	30	344
2	289	802.16	30	94
3	948	4,787.12	30	237
4	1,710	17,208.58	30	455
Total	4,579	27,918	120	1,130

[1] The Nature Conservancy and Environmental Systems Research Institute. 1994 (c). NBS/NPS Vegetation Mapping Program: Accuracy Assessment Procedures. 71 pp. Report to the National Biological Survey and the National Park Service. Arlington, VA and Redlands, CA. <http://biology.usgs.gov/npsveg/standards.html> Last accessed 17 March 2005.

The thematic accuracy of the vegetation association map was then tabulated using a contingency matrix that compared the mapped vegetation associations with the actual vegetation associations observed in the field. Contingency matrices were also tabulated for the fire fuel model and the canopy cover class maps. Overall percent accuracy and Kappa indices were calculated for the three maps (TNC and ESRI 1994c). Overall percent accuracy was calculated by dividing the number of correctly classified accuracy assessment points by the total number of accuracy assessment points. The Kappa index is the preferred method of reporting overall thematic accuracy because it takes into account that a certain number of correct classifications will occur by chance (Foody 1992). The USGS/NPS vegetation mapping protocol requires that the Kappa index of vegetation associations maps exceed 80% (TNC and ESRI 1994c).

Producer's accuracy and user's accuracy were also calculated in the contingency matrix. The producer's and user's accuracies for the vegetation association map should exceeded 80%, according to the USGS/NPS Vegetation Mapping Protocol (TNC and ESRI 1994c). These statistics are not independent, such that one incorrect observation point influences both the producer's and the user's accuracy.

Producer's accuracy represents the percentage of a given association that is correctly identified on the map. In other words, from the perspective of the map, what are the chances that this mapped association correctly matches the vegetation on the ground? Producer's accuracy is calculated by dividing the number of correct observation points in one mapped association class by the total number of observation points in that mapped association class. Error of omission is calculated as 1 - producer's accuracy. This error indicates the percentage of observation points that should have been mapped a given association, but were omitted.

User's accuracy represents the probability that a given location on the ground is classified correctly on the map. In other words, from the perspective of the real world vegetation, what are the chances this association on the ground correctly matches the mapped association? User's accuracy is calculated by dividing the number of correct observation points in one observed association by the total number of points at which that association was observed. Error of commission is calculated as 1 - user's accuracy. This error indicates the percentage of the observation points for a given association that were shown on the map as other associations.