Journal of Vegetation Science 12: 453-466, 2001

The distribution of heath balds in the Great Smoky Mountains, North Carolina and Tennessee

White, Peter S.[1*]; Wilds, S.P.[1,2] & Stratton, D.A.[3]

[1]*Department of Biology, University of North Carolina at Chapel Hill, Campus Box 3280, Chapel Hill, NC 27599-3280 USA;* [2]*P.O. Box 443, Ridgecrest, NC 28770-0443 USA;* [3]*Botany Department, Marsh Life Science Building, University of Vermont,85 South Prospect St., Burlington, Vermont 05405 USA;*
[*]*Corresponding author; Fax +19199626930; E-mail pswhite@unc.edu*

Abstract. We used remote sensing and a geographic information system to model the distribution of evergreen shrub communities, called 'heath balds', in the Great Smoky Mountains, North Carolina and Tennessee, USA. The 421 heath balds averaged 1.8 ha in size and covered 0.3% of the landscape. They reached their greatest importance on upper slopes (92% had relative slope positions > 80), convex topography (82% occurred on sites with a curvature greater than 2.6), and elevations between 1100 and 1600 m (94% of the balds). Although heath balds were found in old-growth watersheds, the two watersheds with the greatest number of balds burned extensively after logging in the early 1900s. Bald occurrence was positively correlated with burned sites, old growth condition, and a highly acidic rock type. Heath balds showed a striking geographic pattern, with 88.1% of the area of this community found in six watersheds comprising only 35.4% of the study area. Despite similar topography, geology, and history, the eleven other watersheds had only 11.9% of the bald area while comprising 64.4% of the study area. Multivariate models showed that this community occurs on only 0.4 to 9.0% of the seemingly appropriate sites. Once established, this shrub community, with its dense evergreen canopy and thick leaf litter, is resistant to tree invasion. Both forest and shrub communities are stable on sites that are seemingly ideal for heath bald occurrence.

Keywords: Evergreen shrub; Geographic Information System; Remote sensing; Watershed.

Nomenclature: White (1982).

Abbreviations: GIS = Geographic Information System; TN = Tennessee; NC = North Carolina; OG = old growth; DIST = second growth; ELEV = elevation; BEERS = Beers transformed slope aspect; CURVE = slope curvature; RSP = relative slope position; TCI = topographic convergence index; LFI = land form index; TSI = terrain shape index; SLOPE = slope steepness; GEOL = presence of Anakeesta bedrock; GROWTH = old or second growth condition; FIRE = burned or unburned condition; HBRA = heath bald-rich area; HBPA = heath bald-poor area.

Introduction

A fundamental goal in vegetation science is to predict the distribution of plant communities from environment and history. While early attempts to make such predictions were qualitative, geographic information systems, digital elevation models, and remote sensing have allowed vegetation scientists to quantitatively test explicit hypotheses about community (Brzeziecki et al. 1993; Walsh et al. 1994; van de Rijt et al. 1996; Zimmermann & Kienast 1999) and species distributions (Franklin 1998; Guisan et al. 1998; Leathwick 1998). In this paper we applied these tools to the distribution of evergreen shrub communities, locally called 'heath balds', in the Great Smoky Mountains, North Carolina and Tennessee, USA. Heath balds have attracted attention since the earliest ecological descriptions of this landscape (Cain 1930a) because they are stable evergreen shrublands in a heavily forested and predominantly broad-leaved deciduous landscape that lacks a climatic treeline (Cogbill et al. 1997).

Heath balds are distinctive in structure and composition: they have a 1-2 m tall evergreen canopy, deep leaf litter, very acidic A-horizons (Cain 1931), low species richness, and a mostly woody flora (ca. 15 woody species, 12 of which are in *Ericaceae*, and five herbaceous species; P. White unpubl.). Dominants include *Rhododendron maximum, R. catawbiense, R. minus*, *Kalmia latifolia*, and *Leiophyllum buxifolium.* Heath balds are abruptly bounded by much taller forests and thus are easily mapped. Comparison of selected heath balds on 1930s and 1980s aerial photographs showed no changes in area (P. White unpubl.); past studies have also treated these communities as stable (Whittaker 1956; Cain 1930b). There are almost no tree seedlings established in heath balds because of the dense evergreen canopy and the thick, acidic leaf litter. Productivity in heath balds is low (Whittaker 1961, 1962).

In 1956, R.H. Whittaker published an influential graphical model for our study area using elevation and a site moisture index to qualitatively describe the distribution of species and community types in watersheds that had not been logged or farmed. Whittaker's diagram depicted heath balds at moderate to high elevation on xeric sites (Whittaker 1956; see also Cain 1930a). No data on geology or other disturbances, such as fire, which is both natural and anthropogenic in this landscape, were included (see Harmon et al. 1983). A further problem was that Whittaker's site moisture index was a subjective scalar that was partly based on the vegetation itself. There is no way to repeat or verify this measure as an index of moisture availability.

We used a digital elevation model to produce objective indices of site environment. Our data also included information on geology and the history of fire and logging. We had two objectives in this work. First, we tested hypotheses about heath bald distribution. This analysis used Whittaker's basic approach to summarize community distribution but used more quantitative tools. Second, we sought to answer questions that go beyond the ones Whittaker was able to ask with his graphical model: What is the number, shape, size, geographical distribution, and arrangement of heath balds? What is the influence of geology and disturbance on heath bald distribution and characteristics? And finally, how predictable are heath balds from site variables? While Whittaker's model implies high predictability and strict environmental determinism, some species distributions and community patterns are unpredictable from site variables, implying a disequilibrium with climatic or other physical factors, a critical issue in vegetation science (Brown 1994; Leathwick 1998).

Methods

Study area

The Great Smoky Mountains comprise an area of 250,000 ha along the North Carolina and Tennessee border, USA (35° 45' N, 83° 30' W). The state line, which runs in a ENE-WSW direction, forms the high elevation backbone of the range, with Tennessee watersheds to the north of the ridge and North Carolina watersheds to the south. Elevations range from 250-2024 m. Annual precipitation increases with elevation from 150 to 250 cm. Mean temperatures decrease with elevation, the July mean temperature from 22° to 14 °C and the January mean temperature from 4° to –2 °C (Shanks 1954). Geology consists of metamorphosed sedimentary rock dominated by sandstones and phyllites. Some 212000 ha of the range are protected within Great Smoky Mountains National Park. Approximately one-third of the park landscape was never disturbed directly by farming and logging (Pyle 1988), though fire and exotic species invasions (chestnut blight and the balsam woolly adelgid) have affected some old growth areas (Harmon et al. 1983).

Data compilation

We defined heath balds as treeless plant communities (< 5% tree cover) with a continuous evergreen shrub cover, thus excluding other high elevation non-forested habitats: grassy balds (Wiser & White 1999; White & Sutter 1998), shrub balds (grassy balds with patches of shrubs and trees established during post-grazing succession, Ramseur 1960), communities of cliff faces and debris avalanche scars (Wiser 1994; Wiser et al. 1996), and seepage areas. All heath balds in Great Smoky Mountains National Park were mapped on 7.5 minute topographic maps from aerial photographs taken in April 1978 and November 1979 and the maps were then digitized. Environmental and historical information was assembled in a GIS, including geology (King et al. 1968), logging history (Pyle 1985), and fire history (Harmon 1980) at a 90-m pixel scale, and a digital elevation model at a 30-m pixel scale.

The digital elevation model was used to derive seven measures of topographic shape and position that are correlated with drainage and exposure. The method of Beers et al. (1966) was used to transform slope aspect so that the warmest (SW) exposure had the highest value and the coolest exposure (NE) had the lowest value. Slope steepness was measured as the angle of the slope relative to the horizontal in a perpendicular direction to slope contours. Slope curvature was computed by an algorithm that integrates two aspects of slope shape: profile curvature (convexity/concavity in the downslope direction) and planiform curviture (convexity/concavity in the direction perpendicular to the downslope direction; Wilds 1996). Relative slope position was calculated as the percentage of the distance from the nearest downhill stream to the nearest uphill ridge divide (Wilds 1996). The topographic convergence index, derived from hydrologic simulation models, was computed as a measure of slope shape that has been shown to be correlated with the steepness of water drainage (Beven & Kirkby 1979). Terrain shape index was computed according to algorithms of McNab (1989); this measure integrates slope curvature and slope position. Landform index, a measure of the degree of exposure, was computed as the mean angle from the site to the horizon (that is, to the highest elevation of nearby and potentially shading landforms; McNab 1993) at major compass points.

For each heath bald, the majority or most common value for each of the variables was selected to represent that bald. Thus, each heath bald was represented as an independent observation in the analyses and there was no bias caused by differences in bald size. There were 421 mapped heath balds, but the largest heath bald straddled a watershed boundary and was therefore divided into two sections for analysis (some topographic measures were computed based on position within a watershed), producing a total of 422 values for each variable. In addition, 422 random points were sampled from the GIS data set of the Park's landscape. The total of 844 data values for each variable were used to test whether heath bald distribution was random relative to the topography, logging history, fire history, and geology and to model heath bald on these interacting gradients. A very acidic rock, the Anakeesta formation, was the most important bedrock used in this analysis.

Because we discovered a striking geographic pattern in heath bald distribution (see Results), we separated park watersheds into a heath bald-rich area (351 heath balds) and a heath bald-poor area (71 heath balds). The heath bald-rich area consisted of the watersheds in which the heath balds area was higher than that expected based on the size of the watershed. Random points equal to the number of heath balds were sampled in each subarea.

Heath bald distribution as influenced by topography, logging history, fire history, and geology

Histograms were used to describe how heath balds sites differed from a random sample of sites (differences were tested with the t-test) and to select variables for use in a model predicting heath bald occurrence. Correlations among the variables were also computed.

Multivariate topographic models

Based on the histograms of heath bald occurrence, we developed a topographic screen in order to project potential heath bald area across the park landscape. We selected five topographic measures and the values of each for which heath bald frequency exceeded the random expectation: elevation (> 1100 to ≤ 1600 m), topographic convergence index (< 60), landform index (< 30), curvature (> 2.5), and relative slope position (> 90). We then plotted all sites (for the park as a whole and for various subsets of the landscape) whose topographic values met these criteria and compared the potential heath bald area to the actual distribution. In a second analysis, we compared locations 300 m (approximately twice the average heath bald length) of the heath balds to the predicted distribution to answer the question: did our models identify the general, if not the specific, location of heath balds?

Logistic regression analysis

Logistic regression (Anon. 1996), was used to build a multivariate model for heath bald distribution in each of three data sets (the park as a whole, the heath bald-rich area, and the heath bald-poor area). Other researchers have used logistic regression for similar questions and data sets (van de Rijt et al. 1996; Leathwick 1998; Franklin 1998; Zimmermann & Kienast 1999). We first identified factors having the most significant correlation with heath bald presence, and then chose a subset of variables that had the least amount of collinearity. Because heath balds at first increased and then decreased with elevation and because of the rapid acceleration of heath bald frequency at slope positions above 90, we also used the squares of these two variables in our analysis (Kleinbaum et al. 1988) to normalize the model. We then created logistic regression models predicting the probability of heath bald occurrence. The κ statistic was used to compare the relative strengths of each model (Goodchild 1994); this statistic compares the model with a random model at a specified probability level, selected by determining at which level model sensitivity (percentage of heath balds accurately predicted) and specificity (percentage of non-heath balds accurately predicted) were approximately equal.

The logistic model was applied to the whole park landscape using the GIS to create a surface of heath bald probability values ranging from 0-1. The probability surface was translated to a map of predicted heath bald occurrence by calculating the percentage of correct and incorrect predictions with changes in the probability criterion for heath bald presence. An optimal probability cutoff was determined by identifying the point at which misclassifications were minimized.

For both the topographic screen and the logistic regression, the predictions were evaluated by examining the frequency of the two correct predictions (predicted heath bald, actual heath bald; predicted non-heath bald, actual non-heath bald) and the two incorrect predictions (predicted non-heath bald, actual heath bald; predicted heath bald, actual non-heath bald). We then computed the percentages on row and column totals – that is, the percentage of the predicted classifications that were correct and the percentage of the actual vegetation that was correctly predicted.

Results

We mapped 421 heath balds totaling 754 ha or 0.4% of the park's landscape. The balds ranged in size from 0.05 ha to 30 ha (Brushy Mountain), with a mean of 1.8 ha and a skew towards smaller sizes (Fig. 1). With the Brushy Mountain bald divided into two sections, the total number of heath bald was 422.

Heath balds were narrow in shape: the average ratio of perimeter to area for heath balds was 0.061, corresponding roughly to a length:width ratio of 15:1. The narrow shape of heath balds reflects the shape of upper ridge topographic positions on which they are found.

Geographic patterns

Nearly half (45.2%) of the heath balds and 40.7% of the heath bald area occurred in two second growth watersheds, Little River and Big Creek, even though these two watersheds accounted for only 11% of the park's total area (Table 1 and Fig. 2). The six central and northeastern watersheds contained 83.2% of the heath balds and 88.1% of the heath bald area but only 35.4% of the park's total area. These watersheds were subsequently defined for further analysis as the 'heath bald-rich area' (Fig. 2). By contrast, the eleven southern and western watersheds contained 16.7% of the heath balds and 11.9% of the heath bald area, but 64.4% of the park's total area. These watersheds were subsequently defined for further analysis as the 'heath bald-poor area'. Most extreme were the seven western and south-western watersheds which contained 2.8% of the heath balds and 2.3% of the heath bald area but 37% of the park's total area (Fig. 2).

The absence of heath balds from the western Tennessee watersheds is probably because the maximum elevations reached in this area are marginal for heath bald development. The contrast between the North Carolina watersheds and the adjacent Tennessee watersheds must be due to other factors since the two sides of the park reach the same elevations, and have similar ranges of topography, geology, and disturbance history. The only obvious difference is that the heath bald-rich area is on the north side of the main mountain crest, whereas the heath bald-poor area is on the south side of that crest. Despite being more common on the part of the park with a gross aspect of north, heath balds tended to occur on xeric slope positions and were more frequent on the more southerly sides of ridges (see below).

Table 1. The geographic distribution of heath balds in the Great Smoky Mountains by watershed. The watersheds are listed from west to east within two categories: the heath bald-rich area and the heath bald-poor area (see Fig. 2). State abbreviations: TN = Tennessee, NC = North Carolina. Disturbance abbreviations: OG = predominantly old-growth watersheds, DIST = predominantly second growth watersheds (logged ca. 1880-1930). Letters after the watershed name correspond to abbreviations on Fig. 2. The percentage area does not sum to 100 because of small areas not included within these major watersheds. Area is in hectares.

Watershed	State	Disturbance	No. heath balds	% of heath balds	Total heath bald area	% of heath bald area	% of study area
Heath bald-rich area (HBRA):							
Middle Prong Little River (MP)	TN	DIST	61	14.5	94	12.4	6.9
Little River (LR)	TN	DIST	114	27.0	127	16.8	6.6
West Prong Little Pigeon (WP)	TN	OG	32	7.6	112	14.9	6.4
Middle Prong Little Pigeon (LP)	TN	OG	38	9.0	94	12.4	5.8
Cosby Creek (CO)	TN	OG	32	7.6	58	7.7	5.3
Big Creek (BC)	NC	DIST	74	17.5	180	23.9	4.4
Subtotal			351	17.5	664	88.1	35.4
Heath bald-poor area (HBPA):							
Panther Creek (PC)	TN	DIST	0	0	0	0	2.0
Abrams Creek (AC)	TN	DIST	0	0	0	0	9.9
Hesse Creek (HE)	TN	DIST	0	0	0	0	1.8
Twenty Mile (TW)	NC	DIST	0	0	0	0	2.8
Eagle Creek (EC)	NC	DIST	6	1.4	8	1.1	4.8
Hazel Creek (HC)	NC	DIST	6	1.4	10	1.3	7.0
Forney/Noland Creek (FN)	NC	DIST	0	0	0	0	8.9
Deep Creek (DC)	NC	DIST	12	2.8	15	2.0	5.4
Oconoluftee/Bradley Fork (OB)	NC	DIST	16	3.8	16	2.1	7.0
Raven Fork (RF)	NC	OG	14	3.3	13	1.7	7.0
Cataloochee (CA)	NC	OG	17	4.0	28	3.7	7.8
Subtotal			71	16.7	90	11.9	64.4
Total			422	100	754	100	99.8

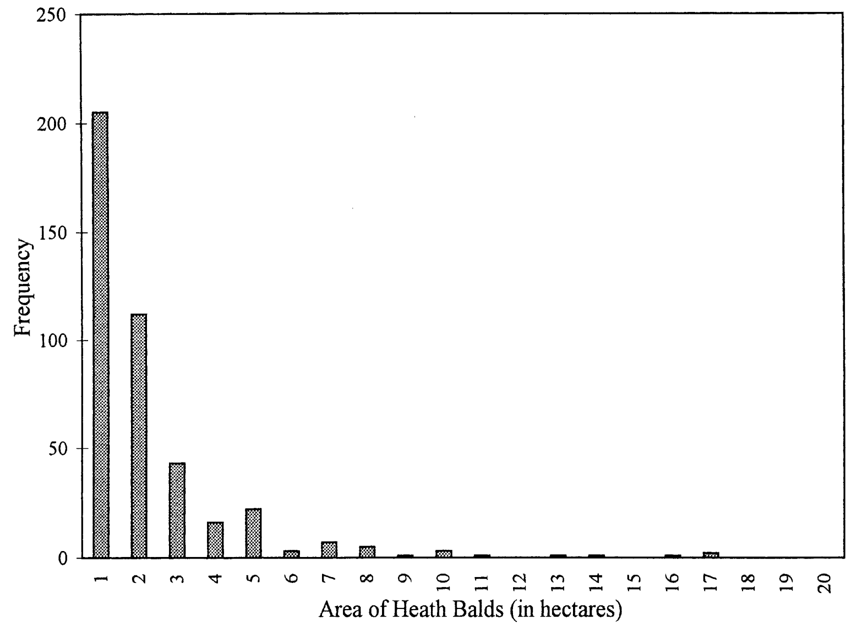

Fig. 1. The size distribution of heath balds in the Great Smoky Mountains.

Topography, disturbance history, and geology

Heath balds occupied a highly non-random assortment of sites. The elevation of heath balds ranged from 940 to 1730 m (3100 to 5710 ft) with a mean value of 1359 m (4491 ft) (Fig. 3a). This was significantly higher than the average elevation of the park (1014 m or 3346 ft) ($p < 0.0001$). 94% of the balds were between 1100 and 1600 m (3610-5252 ft) in elevation. Heath balds were more frequent for low values of transformed aspect than the random expectation, which corresponds to slope aspects from SE to SW (Fig. 3b). Generally, heath balds tended to be on moderately steep slopes between 20 and 35°; however, this was not significantly different from the slope steepness of average sites (Fig. 3c). Balds at the top of some ridges had very low slopes, while those on the steeper sides had slopes up to 43 ° and higher on a local scale.

As is evident from casual observations, the heath balds were most common on ridge tops and convex slopes and differed greatly from random expectations for variables that measure topographic shape (Fig. 3d-h). 82% of the balds occurred at curvatures above 2.6 (mean = 10.3; Fig. 3d). 92% of heath balds occurred at relative slope position equal to or greater than 80 (mean = 91; Fig. 3e). 94% occurred at topographic convergence index values less than 60 (mean = 42; Fig. 3f). 87% occurred at landform index values less than 30 (mean = 19; Fig. 3g). Heath balds occurred at low values of the terrain shape index (< – 20; Fig. 3h). These

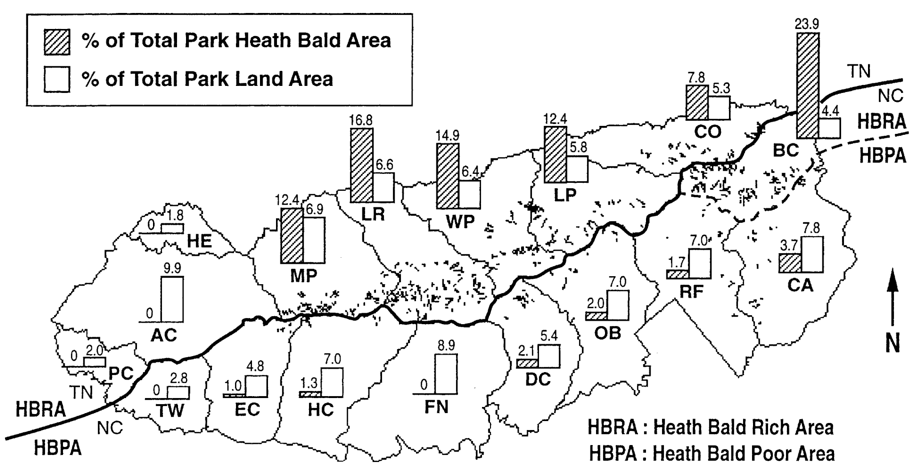

Fig. 2. Distribution of heath bald area by watershed in the Great Smoky Mountains. The joint boundary of Tennessee (TN) and North Carolina (NC) forms the high ridge of the mountain range, with watersheds draining to the north and south. The heath bald-rich area (HBRA) and heath bald-poor area (HBPA) are defined in the text.

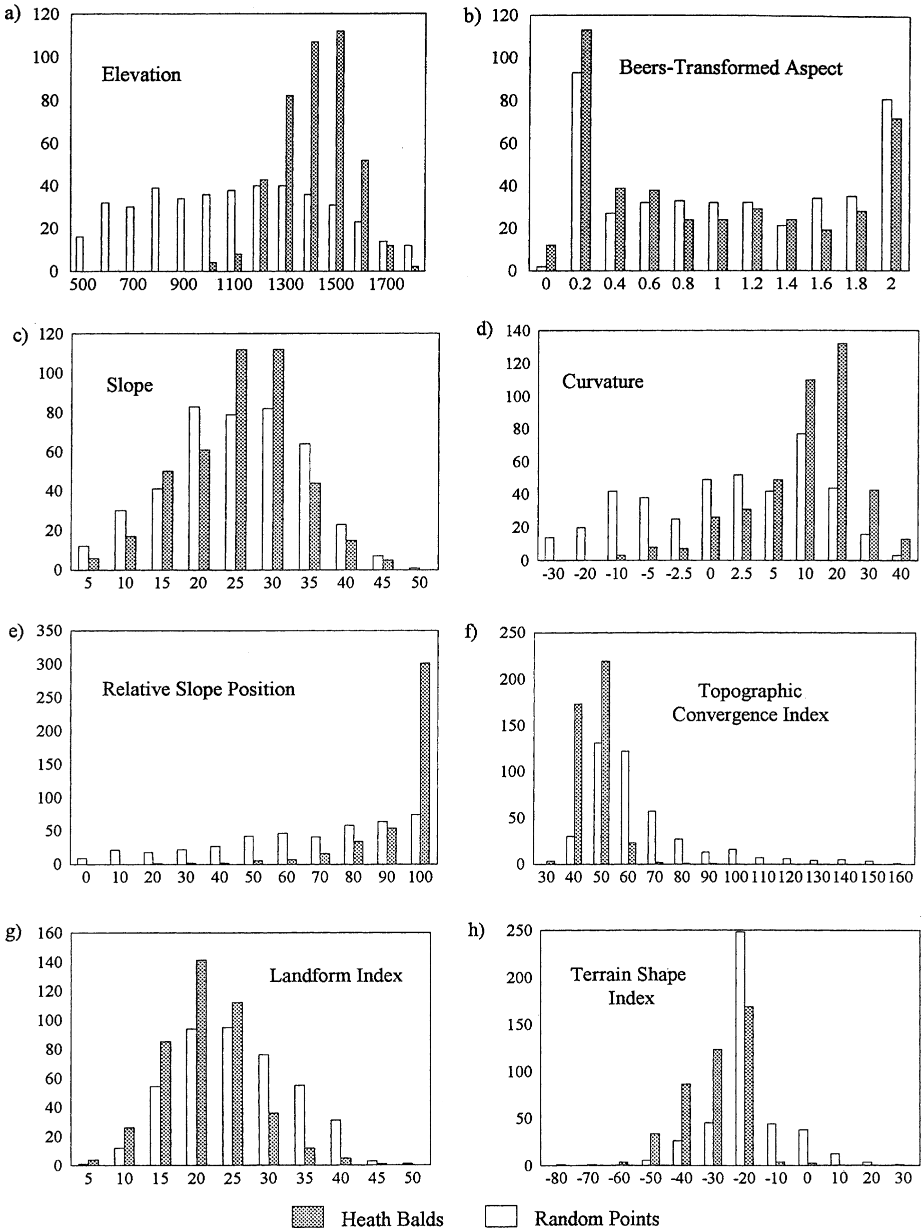

Fig. 3. Distribution of heath balds and random points by topographic variables in the Great Smoky Mountains.

distributions were significantly different from the expectation based on random samples of the landscape ($p < 0.0001$).

64% of the balds were approximately symmetrical on the ridges on which they occurred and contained no spurs or extensions onto adjacent slopes. The other 46% showed two interesting patterns related to aspect: balds on south facing ridges tended to have symmetrical extensions (the heath balds, the average aspect of these extensions, and the ridge of which the heath bald occurred all had the similar slope aspects), but balds on other slope aspects were asymmetrical, with their greatest development on aspects that were more southerly than the major ridge on which they occurred, with the degree of displacement increasing through east and west to north facing ridges (Fig. 4).

Heath balds tended to be more frequent on Anakeesta bedrock, compared to a random expectation (Fig. 5a-c). Heath balds were more frequent than the random expectation on old-growth sites for the park as a whole and for the heath bald-rich area, but were more frequent than the random expectation for second-growth sites for the heath bald-poor area (Fig. 5d-f). Heath balds were more frequent than the random model suggested in burned areas (Fig. 5g-i).

As might be expected from these results, heath balds

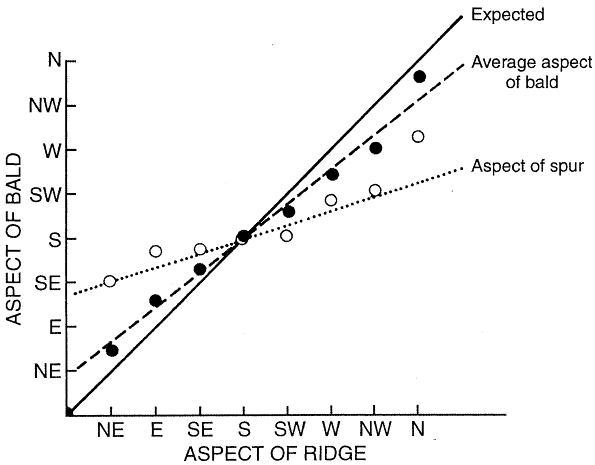

Fig. 4. The compass aspect of the major heath bald axis compared to the aspect of the ridge on which the heath bald was found. If heath balds were symmetrical on the ridges on which they occurred, they would be expected to lie along the solid line (heath bald aspect = ridge aspect). However, moving away from south both the heath balds themselves and major side branches of the balds ('spurs') tend to be larger on warmer slope faces than the ridge on which they occur.

showed strong correlations with elevation, relative slope position, topographic convergence aspect, curvature, terrain shape index, and landform index (Table 2). Relative slope position had the highest correlation coefficient of the topographic measures and strong correlations with all other measures of site topography, among which there was a high degree of colinearity. Heath bald occurrence was also correlated with the presence of Anakeesta bedrock. Although burned areas were negatively correlated with old-growth conditions, heath balds were positively correlated with both (note, however, that heath bald correlation with Anakeesta bedrock, old growth conditions, and burned sites were several fold weaker than with topographic characteristics; Table 2). The occurrence of Anakeesta bedrock and old growth conditions showed the same correlations. These were positively correlated with elevation, slope steepness, slope curvature, and relative slope position, and negatively correlated with topographic convergence index and terrain shape index. Anakeesta bedrock was also positively correlated with old growth conditions.

A similar analysis was carried out for heath bald size, as for heath bald occurrence. Nearly all factors which were strongly correlated with heath bald occurrence were also correlated with heath bald size. Mean heath bald size increased from 1.6 ha at 1200 m elevation (n = 55), to 1.8 ha at 1200-1600 m elevation (n = 353), and 2.8 ha at elevations greater than 1600 m (but note the small sample size at these elevations, n = 14; none of the sets of numbers were affected by excluding the largest bald which was several times the size of the second largest bald). Mean heath bald size increased from 1.4 ha at relative slope positions 60-80 (n = 50) to 2.3 ha at relative slope positions greater than 95 (n = 224; these numbers were unaffected by excluding the largest bald). Heath balds were also larger in burned areas (mean size was 2.2 ha, n = 139; if the largest bald is excluded, the mean was 1.9 ha) than in unburned areas (1.6 ha, n = 283). Although more frequent on logged sites, mean heath bald size tended to be higher in old growth (2.0 ha, n = 173; if the largest bald is excluded, the average is 1.8 ha) than in second growth (1.6 ha, n = 249).

Table 2. Correlations among heath bald occurrences, environmental, and historical variables 1, 2. Lower case letters show significance levels (a < 0.05, b < 0.01, c < 0.001, d < 0.0001, n.s. = not signficant).

	BALDS	ELEV	BEERS	SLOPE	CURVE	RSP2	TCI	LFI	TSI	GEOL	GROWTH
ELEV	0.491 d	-----									
BEERS	–0.096 b	n.s.	-----								
SLOPE	n.s.	0.250 d	n.s.	-----							
CURVE	0.419 d	0.296 d	–0.089 b	n.s.	-----						
RSP2	0.592 d	0.378 d	–0.151 d	n.s.	0.538 d	-----					
TCI	–0.500 d	–0.429 d	0.076 a	–0.308 d	–0.682 d	–0.666 d	-----				
LFI	–0.311 d	–0.143 d	n.s.	0.405 d	–0.253 d	–0.541 d	0.266 d	-----			
TSI	–0.458 d	–0.344 d	0.129 c	n.s	–0.768 d	–0.547 d	0.714 d	0.235 d	-----		
GEOL	0.157 d	0.244 d	n.s.	0.167 d	0.152 d	0.123 c	–0.158 d	n.s.	–0.165 d	-----	
GROWTH	0.076 a	0.377 d	–0.137 d	0.160 d	0.099 b	0.100 b	–0.159 d	n.s.	–0.119 c	0.238 d	-----
FIRE	0.186 d	0.141 d	0.140 d	0.071 a	n.s.	0.109 b	–0.092 b	n.s.	n.s.	n.s.	–0.285 d

1Continuous variable abbreviations: ELEV (Elevation), BEERS (Beers transformation of slope aspect), SLOPE (Slope steepness), CURVE (Slope curvature), RSP2 (Relative slope position squared), TCI (Topographic convergence index), LFI (Landform index), TSI (Terrain shape index).
2Categorical variable abbreviations and values: BALDS (1 = heath bald present, 0 = heath bald absent), GEOL (1 = Anakeesta bedrock, 0 = other bedrock types), GROWTH (1 = old-growth, 0 = second growth), FIRE (1 = burned, 0 = unburned).

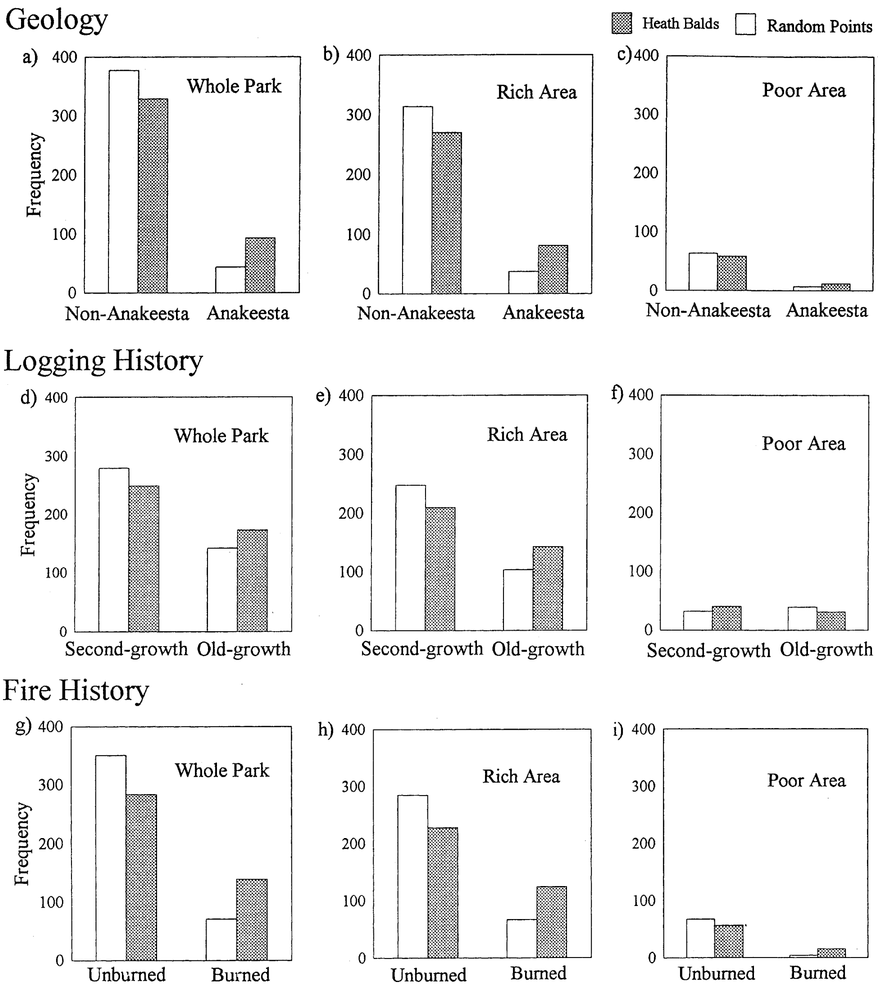

Fig. 5. Heath bald distribution by geology, logging history, and fire history in the Great Smoky Mountains.

Topographic models

The topographic screen predicted that heath balds would occur on 6.7-14.3% of the study area, depending on the data subset that was analysed (Table 3). These values were much higher than the actual values for this vegetation type. Predicted heath bald occurrence was 9.1% of the whole study area (actual was 0.4%), 8.4% of the heath bald-rich area (actual was 0.9%), and 9.4% of the heath bald-poor area (actual was 0.1%) (Table 3). However, this model predicted the greatest heath bald occurrence on sites in which actual values were also relatively high, although the actual percentages were much lower than the predicted percentages. Sites with high predicted and actual values were: old growth sites (14.3% predicted, 0.7% actual), Anakeesta bedrock in the heath bald-rich area (14.1% predicted, 1.8% actual), Anakeesta bedrock throughout the study area (13.3% predicted, 1.3 actual), burned sites in the heath bald-rich area (13.3% predicted, 2.3 actual), and old growth in the heath bald-rich area (12.4% predicted, 1.4 actual).

Summarizing across the models, predicted area was 6-12 × higher than actual area for the best heath bald sites, while for the park as a whole, the predicted area was 22 × the actual area and was 94 × the actual area for the heath bald-poor area.

We evaluated the model further by computing the percentage of the predicted area that was correct (that is, was actually dominated by the appropriate vegetation) and by computing the percentage of the actual vegetation that was correctly predicted (Table 3).

Of the total area predicted to be heath balds, only 2.4% supported this vegetation (Table 3). This figure more than doubled to 6.5% for the heath bald-rich region, while it dropped to 0.4% in the remainder of the park. Both park-wide and within the heath bald-rich area, this percentage remained low, rising to a maximum of 9.0% in burned areas within the heath bald-rich area (Table 3). Actual heath bald area as a percentage of potential area was also relatively high in old growth areas and areas with Anakeesta bedrock. The model did much better for non-heath bald sites: 98.7-100% of the

Table 3. Predicted and actual heath bald sites in the Great Smoky Mountains. The potential heath bald sites were defined by five topographic variables (see text; HBRA = heath bald-rich area, HBPA = heath bald-poor area). In the first two columns the numbers in parentheses are the percentages of the landscape subarea that are predicted and actual heath balds.

Landscape Subarea	Area of Predicted Heath balds	Area of Actual Heath balds ha (%)	Area of Intersection Correct ha (%)	% Predicted Heath bald Correct ha	% Predicted Non-bald Correct	% Actual Heath bald Correct	% Actual Non-bald
Whole park	18658.7 (9.1)	754.1 (0.4)	452.1	2.4	99.8	60.0	91.1
HBRA	6152.5 (8.4)	665.0 (0.9)	398.2	6.5	99.6	59.9	92.0
HBPA	12498.2 (9.4)	89.1 (0.1)	54.0	0.4	100	60.6	90.6
Whole park:							
Old-growth	6795 (14.3)	346.7 (0.7)	219.0	3.2	99.7	63.2	86.1
Second-growth	11865.1 (7.5)	425.9 (0.3)	229.9	1.9	99.9	54.0	92.6
Burned	2569.5 (10.7)	295.2 (1.2)	153.0	6.0	99.3	51.8	89.9
Unburned	16091.9 (8.8)	450.6 (0.2)	292.9	1.8	99.9	65.0	91.3
Anakeesta	1665.7 (13.3)	167.5 (1.3)	91.6	5.8	99.3	54.7	87.3
Non-Anakeesta	16990.5 (8.8)	580.1 (0.3)	355.1	2.1	99.9	61.2	91.4
HBRA:							
Old-growth	2,715.1 (12.4)	312.7 (1.4)	197.1	7.3	99.4	63.0	88.4
Second-growth	3440.3 (6.7)	367.0 (0.7)	196.1	5.7	99.6	53.4	93.6
Burned	1610.2 (13.3)	282.1 (2.3)	144.9	9.0	98.7	51.4	87.7
Unburned	4547.9 (7.5)	376.4 (0.6)	245.6	5.4	99.8	65.2	92.9
Anakeesta	1202.8 (14.1)	153.8 (1.8)	84.4	7.0	99.1	54.9	86.6
Non-Anakeesta	4950.9 (7.7)	504.0 (0.8)	308.8	6.2	99.7	61.3	92.7

predicted area was correctly classified and, thus, a very low percentage of the predicted non-heath bald areas was actually dominated by heath balds.

The topographic screen correctly classified 51.4 - 65.2% of the actual bald area depending on data subset (Table 3), even though this type occupies only 0.1-2.3 of the landscape and the random expectation from the model would be a correct classification rate of 6.7-14.3%, based on the percentage of the study area and its subsets predicted to be heath balds (Table 3). The topographic screen correctly classified 86.1-93.6% of the actual non-bald area, with the remaining incorrectly predicted to be heath bald area.

We can conclude from these figures that the major characteristics of the models were an accurate assessment of non-heath bald area (almost all the predicted non-heath bald area was correct), a reasonably definition of heath bald sites (over 60% of the heath bald area correctly predicted on the best heath bald sites), and an overprediction of heath bald area.

We used the same model to ask what percentage of the predicted area fell within 300 m (near the average length of the balds) of the actual heath balds (Table 4). For the heath bald-rich area, 42.3% of the predicted area was within 300 m of a bald. For the park as a whole, this number fell to 16.7% and was only 4.2% for the heath bald-poor area.

Logistic regression

In our multivariate analysis, we used elevation, relative slope position, transformed aspect, and slope curvature and three dichotomously coded variables (old-growth/second-growth, burned/unburned, Anakeesta/non-Anakeesta substrate). These variables are highly successful in distinguishing heath bald and non-heath bald data points in the 844 observation data set (422 heath bald data points, 422 random data points). As might be expected, the model for the heath bald-rich area had the highest κ statistic (0.74, within the ‘very good to excellent range’ for model performance of Zimmermann & Kienast 1999) and lowest percentages of false positives (13.8%) and false negatives (14.2%) among the three models, although differences in the three models were not very pronounced (Table 5). As indicated by the Wald χ^2 statistic, elevation and relative slope position were consistently the most significant variables across all three models. Fire was also significant in all three models (heath bald probability was increased on burned sites). Logging history was important in two models: heath balds probability was higher on second growth sites for the whole data set and for the heath bald-poor area. Transformed aspect and geology were insignificant in all three models.

We used the logistic regression model to predict heath bald sites for the entire park landscape. We examined the model’s classification error rate by varying the critical

Table 4. Percentage of predicted heath bald sites that is within 300 m of a heath bald in the Great Smoky Mountains. The potential heath bald sites were defined by five topographic variables (see text; HBRA = heath bald rich area, HBPA = heath bald poor area).

Landscape Subarea	Area of Predicted Heath balds (ha)	Area of Actual Heath balds + 300 m (ha)	Area of Intersection (ha)	% Predicted Heath bald Correct	% Predicted Non-bald Correct	Heath bald Correct	% Actual Non-bald Correct
Whole park	18658.7	14980.3	31109.9	16.7	93.7	20.8	91.9
HBRA	6152.5	12221.2	2604.1	42.3	85.6	21.3	94.2
HBPA	12498.2	2687.2	507.0	4.1	98.2	18.9	90.8

probability for predicting heath bald presence from 0.95 to 1.0. At the 0.9985 level, misclassifications were minimized. Out of a total of 301451 heath bald cells, this model predicted 266716 or 88.5% (compared to 51.4-65.2% in the topographic screen), with the other 11.5% predicted to be non-heath balds. However, just as with the topographic model developed above, the logistic model overpredicted heath balds; when applied to the whole park landscape, only 1.3% of the predicted area was dominated by heath balds, a similar finding to that of the topographic model.

Table 5. Summary of logistic regression models for (**a**) the whole park, (**b**) the heath bald rich area, and (**c**) the heath bald poor area.

a. The whole park data set (*n* = 844) κ statistic 0.72
Probability value for approximate equality of false positives and false negatives 0.62
False positives 14%, False negatives 15%

Variable	DF	Parameter estimate	Standard error	Wald χ^2	Pr > χ^2	Standardized estimate
INTERCPT	1	–46.4151	5.2036	79.5627	0.0001	.
ELEV	1	0.0650	0.00777	69.9145	0.0001	10.826664
ELEV2	1	–0.00002	2.9E-6	66.6255	0.0001	–8.736711
RSP2	1	0.000426	0.000045	88.7610	0.0001	0.764181
CURVE	1	0.00377	0.00122	9.5811	0.0020	0.256908
GROWTH	1	–0.5864	0.2485	5.5662	0.0183	–0.156448
FIRE	1	0.9508	0.2801	11.5201	0.0007	0.226757

b. Heath bald rich area (*n* = 702) κ statistic 0.74
Probability value for approximate equality of false positives and false negatives 0.64
False positives 13.8%, false negatives 14.2%

Variable	DF	Parameter estimate	Standard error	Wald χ^2	Pr > χ^2	Standardized estimate
INTERCPT	1	–46.1689	5.6207	67.4701	0.0001	.
ELEV	1	0.0649	0.00848	58.6111	0.0001	11.181226
ELEV2	1	–0.00002	3.202E-6	55.2846	0.0001	–8.957543
RSP2	1	0.000421	0.00005	70.4714	0.0001	0.755453
CURVE	1	0.00415	0.00141	8.6338	0.0033	0.277853
FIRE	1	0.7181	0.3047	5.5519	0.0185	0.176308

c. Heath bald poor area (*n* = 142) κ statistic 0.68
Probability value for approximate equality of false positives and false negatives 0.56
False positives 18.3%, false negatives 18.3%

Variable	DF	Parameter estimate	Standard error	Wald χ^2	Pr > χ^2	Standardized estimate
INTERCPT	1	–63.3923	17.4275	13.2313	0.0003	.
ELEV	1	0.0875	0.0246	12.6150	0.0004	9.128416
ELEV2	1	–0.00003	8.714E-6	12.7203	0.0004	–8.436669
RSP2	1	0.000503	0.00012	17.6669	0.0001	0.902806
GROWTH	1	–1.5936	0.6499	6.0134	0.0142	–0.440810
FIRE	1	1.7363	0.8615	4.0619	0.0439	0.327057

Discussion

From one perspective, our analyses supported Whittaker's diagram – heath balds occupied the highly non-random assortment of sites that his diagram showed. However, while multivariate models were successful at identifying heath bald sites, they greatly overpredicted the amount of heath bald vegetation. From this perspective, our results challenge the determinism implied by Whittaker's model. The questions suggested by this finding are these: Why is the percentage of the predicted area actually dominated by heath balds low? Why do the same kinds of sites support such physiognomically different vegetation types as tall forests or shrub thickets? What accounts for the wide differences in heath bald occurrence in the rich area (predominantly on the north side of the mountains in Tennessee) and poor area (predominantly on the south side in North Carolina)? The answers to these questions likely involve both disturbance and environment.

We discuss two contrasting scenarios for heath bald occurrence. First, heath balds may form on upper slopes with rocky substrates and thin soils as part of primary succession. We think this scenario is unlikely for reasons discussed below. Second, heath balds may form from closed forests after wind or fire disturbance that removes trees but not understory heath shrubs. The evergreen shrubs that dominate heath balds are shade tolerant and are found in forest understories as well as on exposed heath balds. In mesic areas near streams, this shrub layer has been shown to drastically reduce tree seedling establishment. Barden (1979) argued that one function of nurse logs in the old growth hemlock-hardwood forests is to flatten the shrub layer and provide an elevated site for seedling establishment. Evergreen shrubs have also been found to suppress forest regrowth after logging; the removal of these layers has been a management objective in southern Appalachian National Forests (Della-Bianca & McGee 1972).

The first scenario for heath bald origin is that they are part of a primary successional sere that ends in forest and during which soils become deeper and more developed. This scenario requires that forests on heath bald sites developed from heath balds, which does not seem to occur, and/or that downslope erosion of soil keeps pace with soil development, thus preventing succession from occurring. There is no evidence that heath balds are successional or that soil development and erosion are in competition under heath balds, although down slope leaching of plant nutrients could be ongoing. Heath shrubs do occur, with other species, on rock outcrops (Wiser 1994). Thin soils and downhill movement of soil and boulders may be an explanation for the persistence of these non-forested habitats, but these communities have little compositional or structural similarity to heath balds except when the cliff face terminates at its upper edge in a typical heath bald topographic situation.

We think it is more likely that heath balds result from disturbance to forest-heath communities. Cain (1930b) interpreted them as post-climax communities that developed from forest heaths after death of the trees. Regardless of origin, however, once a dense shrub layer is formed, it is very resistant to tree establishment so that whatever combination of historical factors leads to heath bald formation, the imprint of that history is a long lasting one.

Given an understory of evergreen shrubs, any disturbance that removes trees may allow the understory shrub layer to become dominant and to prevent or slow the reestablishment of trees. The evergreen shrubs cast a deep year-round shade and produce a thick, slowly decomposing, and acidic litter layer. Cain (1931) was the first to document low pH under heath balds. One soil scientist suggested that pH was so low under heath balds that aluminum levels were toxic to the roots of *Picea rubens* (M. O. Springer unpubl.). Gant (1978) reported an allelochemic effect on seed germination of test plants (which, however, did not include tree seedlings). This scenario suggests that trees established before or with the heath shrubs on these sites because an initial colonization by heath shrubs alone would have produced a stable shrub community and excluded tree establishment (see also below).

Although heath balds are found in old-growth, unburned sites, fires and logging increased the frequency and size of heath balds in some areas. The two watersheds with the most heath balds were Little River, which burned in the early 1900s and in 1925, and Big Creek, which burned in 1924 (Lambert 1958). These were severe fires on logging slash after whole watersheds were logged. Such fires, in this region of very heavy rainfall, often resulted in severe soil erosion as well. Although not in a logged watershed, the largest heath bald, on Brushy Mountain, burned three times in the 1920s (Lambert 1958). If heath shrubs survived these fires, they could have expanded onto the adjacent eroded slopes, since these shrubs are tolerant of acid, low nutrient conditions. The removal of trees followed by coalescing of understory shrubs would also result in heath bald dominance in areas formerly dominated for closed forest with a heath understory. These shrubs do not colonize the sites by seed over significant distances and it is a necessary condition that they be present on the site before the fire and survive the fire that subsequently occurred. Further, there are many burned sites that are not dominated by heath balds (see Lindsay & Bratton 1979) and a number of old growth watersheds that have heath balds that did not experience fire. The elevations of the heath balds put them in an environment that has very low natural fire frequency (Harmon et al. 1983). Further, disturbances to the forests surrounding and down slope from heath balds do not result in permanent heath shrub dominance. Thus, the disturbances themselves must interact with topography and preexisting vegetation to produce this community type.

In both the treefall and fire scenarios, a necessary precondition for heath bald formation is the presence of the evergreen shrubs in forest understories. Understory heath distributions are patchy in the landscape we studied (Fig. 6), leading to another source of variation in heath bald presence, regardless of site or history. Further, understory heaths on mid-slopes and flat or concave topography do not coalesce to form balds after tree disturbance. Thus, there must be environmental conditions of the heath bald sites that lead first to heath understories and then to heath balds.

Because there are so many more heath balds on the Tennessee side of the Great Smokies, there may be unmeasured environmental and geological factors that are correlated with geographical position. The Tennessee watersheds face generally north from the stateline ridge, although the heath balds in these watersheds tend to face slope aspects other than north. Prevailing winds and clouds generally come from west and northwest, striking the Tennessee side of the high ridge before moving across to the North Carolina side. The north facing watersheds might then have higher cloud cover (adding to other causes of low productivity) and a higher throughflow of moisture (however, run-off is steep and soil moisture low on the heath bald sites despite this potentially higher rainfall) which could result in leaching of soil nutrients. These factors alone or in combination might result in reduced productivity compared to the south facing watersheds, factors which might benefit evergreen species. In addition, lower humidity, higher insolation, warmth, and drought may limit development of both heath understories and heath balds on the predominantly south-facing side of the study area. However, no data exist with which to examine the effect of gross differences in aspect on these mountain environments.

Experimental work on establishment, survivorship, and competition for heath shrubs and tree seedlings would allow better understanding of the role of soil acidity and nutrients, deep leaf litter, drought, and deep shade in preventing tree seedling invasion and succession on the heath balds, both in old growth and disturbed watersheds. Such information is needed not just for better understanding of these communities, but also for possible restoration of sites influenced by logging slash fires and soil erosion. Field studies of topography at a smaller scale than that used here (see also discussion in Brown 1994) would also help refine our understanding of the topographic conditions under which heath balds develop.

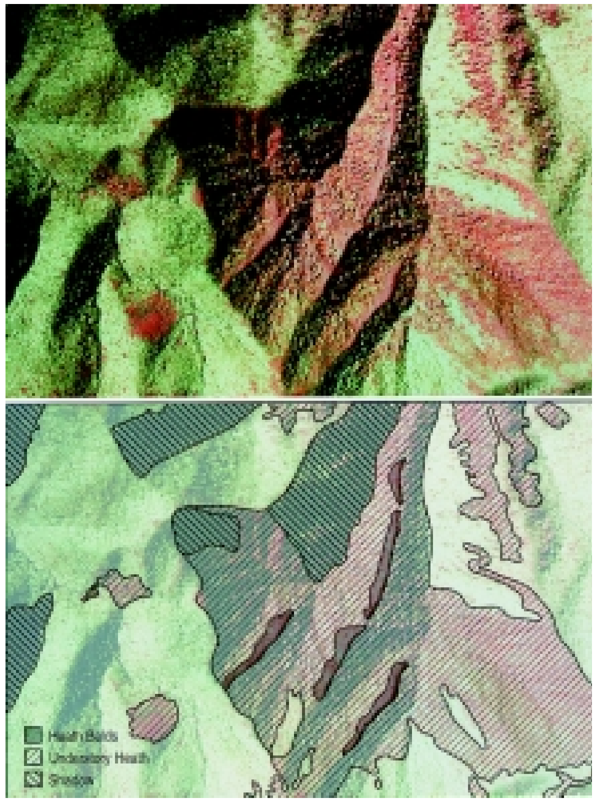

Fig. 6. Color infrared photography taken in November (leaf off) condition showing the patchy distribution of understory heath shrubs (red color).

Computer mapping and new statistical tools help us phrase explicit questions that deepen our understanding about vegetation-environment relations. These tools are essential to predicting the response of vegetation to climate change (Guisan et al. 1998; Walsh et al. 1994; Brzeziecki et al. 1993). While our logistic regression performed very well in comparison with other published models in discriminating sites used for model building (e.g., Brown 1994; Leathwick 1998; Zimmermann & Kienast 1999), the overprediction of heath balds suggested both a disequilibrium with site factors and the need for additional field data to refine models of site environment. In particular, disturbances or events like past climate variation (Leathwick 1998) that alter species distributions, while often difficult to quantify, are often implicated by model failure (Brown 1994; Brzeziecki et al. 1993). Remote sensing and computer mapping can thus be used to identify further research questions.

References

Anon. 1996. *SAS User's Guide: Statistics, Version 5 ed.* SAS Institute, Inc., Cary, NC.

Barden, L.S. 1979. Tree replacement in small canopy gaps of a *Tsuga canadensis* forest in the southern Appalachians, Tennessee. *Oecologia* (*Berl.*) 44: 141-142.

Beers, T.W., Dress, P.E. & Wensel, L.C. 1966. Aspect transformation in site productivity research. *J. For.* 64: 691-692.

Beven, K.J. & Kirkby, M.J. 1979. A physically based, variable contribution area model of basin hydrology. *Hydrol. Sci. Bull.* 24: 43-69.

Brown, D.G. 1994. Predicting vegetation types at treeline using topography and biophysical disturbance variables. *J. Veg. Sci.* 5: 641-656.

Brzeziecki, B., Kienast, F. & Wildi, O. 1993. A simulated map of the potential natural forest vegetation of Switzerland. *J. Veg. Sci.* 4: 499-508.

Cain, S.A. 1930a. *The vegetation of the Great Smoky Mountains: an ecological study*. Ph.D. Diss., Univ. of Chicago. Chicago, IL.

Cain, S.A. 1930b. An ecological study of the heath balds of the Great Smoky Mountains. *Butler Univ. Bot. Stud.* 1: 177-207.

Cain, S.A. 1931. Ecological studies of the vegetation of the Great Smoky Mountains of North Carolina and Tennessee. I. Soil reaction and plant distribution. *Bot. Gaz.* 91: 22-41.

Cogbill, C.V., White, P.S. & Wiser, S.K. 1997. Predicting treeline elevation in the southern Appalachians. *Castanea* 62: 137-146.

Della-Bianca, L. & McGee, C.E. 1972. Reaction of natural *Rhododendron maximum* to liming in the southern Appalachians. *J. Elisha Mitchell Sci. Soc.* 88: 109-112.

Gant, R.E. 1978. *The role of allelopathic interference in the maintenance of southern Appalachian heath balds.* Ph.D. Thesis, University of Tennessee, Knoxville, TN.

Goodchild, M.F. 1994. Integrating GIS and remote sensing for vegetation analysis and modeling: methodological issues. *J. Veg. Sci.* 5: 615-626.

Guisan, A., Theurillat, J.-P. & Kienast, F. 1998. Predicting the potential distribution of plant species in an alpine environment. *J. Veg. Sci.* 9: 65-74.

Franklin, J. 1998. Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. *J. Veg. Sci.* 9: 733-748.

Harmon, M.E. 1982. The fire history of the westernmost portion of Great Smoky Mountains National Park. *Bull. Torrey Bot. Club* 109: 74-79.

Harmon, M.E., Bratton, S.P. & White, P.S. 1983. Disturbance and vegetation response in relation to environmental gradients in the Great Smoky Mountains. *Vegetatio* 55: 129-139.

King, P.B., Newman, R.B. & Hadley, J.B. 1968. *Geology of the Great Smoky Mountains National Park, Tennessee and North Carolina.* USDI Geological Survey Professional Paper 587, Washington, DC.

Kleinbaum, D.G., Kupper, L.L. & Muller, K.E. 1988. *Applied regression analysis and other multivariable methods.* Duxbury Press, Belmont, CA.

Lambert, R.S. 1958. *Logging in the Great Smoky Mountains National Park.* Unpubl. Report to the Superintendant, Great Smoky Mountains National Park, Gatlinburg, TN.

Leathwick, J.R. 1998. Are New-Zealand's *Nothofagus* species in equilibrium with their environment? *J. Veg. Sci.* 9: 719-732.

Lindsay, M.M. & Bratton, S.P. 1979. The vegetation of grassy balds and other high elevation disturbed areas in Great Smoky Mountains National Park. *Bull. Torrey Bot. Club* 106: 264-275.

McNab, W.H. 1989. Terrain Shape Index: quantifying effect of minor landforms on tree height. *For. Sci.* 35: 91-104.

McNab, W.H. 1993. A topographic index to quantify the effect of mesoscale landform on site productivity. *Can. J. For. Res.* 23: 1100-1179.

Pyle, C. 1985. *Vegetation disturbance history of Great Smoky Mountains National Park: an analysis of archival maps and records.* USDI National Park Service Research/Resources Management Report SER-77, Atlanta, GA.

Pyle, C. 1988. The type and extent of anthropogenic vegetation disturbance in the Great Smoky Mountains before National Park Service acquisition. *Castanea* 53: 225-235.

Ramseur, G.S. 1960. The vascular flora of high mountain communities of the southern Appalachians. *J. Elisha Mitchell Sci. Soc.* 76: 82-112.

Shanks, R.E. 1954. Climates of the Great Smoky Mountains. *Ecology* 35: 354-361.

van de Rijt, C.W.C.J., Hazelhoff, L. & Blom, C.W.P.M. 1996. Vegetation zonation in a former tidal area: a vegetation-type response model based on DCA and logistic regression using GIS. *J. Veg. Sci.* 7: 505-518.

Walsh, S.J., Davis, F.W. & Peet, R.K. 1994. Applications of remote sensing and geograhic information systems in vegetation science. *J. Veg. Sci.* 5: 609-750.

White, P.S. 1982. *The flora of Great Smoky Mountains National Park: an annotated checklist of the vascular plants and a review of previous floristic work.* USDI, National Park Service, Southeast Regional Office, Research/Resource Management Report SER-55, Atlanta, GA.

White, P.S. & Sutter, R.D. 1998. Southern Appalachian grassy balds: lessons for management and regional conservation. In: Peine, J.D. (ed.) *Ecosystem management: principles and practices illustrated by a regional biosphere cooperative*, pp. 375-396. St. Lucie Press, Delray Beach, FL.

Whittaker, R.H. 1956. Vegetation of the Great Smoky Mountains. *Ecol. Monogr.* 26: 1-80.

Whittaker, R.H. 1961. Estimation of net primary production of forest and shrub communities. *Ecology* 42: 173-180.

Whittaker, R.H. 1962. Net production relations of shrubs in the Great Smoky Mountains. *Ecology* 43: 357-377.

Wilds, S. 1996. *Gradient analysis of the distribution of flowering dogwood (Cornus florida L.) and dogwood anthracnose (Discula destructiva Redlin.) in western Great Smoky Mountains National Park*. Master of Arts Thesis, University of North Carolina, Chapel Hill, NC.

Wiser, S.K. 1994. High elevation cliffs and outcrops of the southern Appalachians: vascular plants and biogeography. *Castanea* 59: 85-116.

Wiser, S.K., Peet, R.K. & White, P.S. 1996. High elevation rock outcrop vegetation of the Southern Appalachian Mountains. *J. Veg. Sci.* 7: 703-722.

Wiser, S.K. & White, P.S. 1999. High elevation outcrops and barrens of the southern Appalachian mountains. In: Anderson, R.C., Fralish, J.S. & Baskin, J.M. (eds.) *Savannas, barrens, and rock outcrop plant communities of North America,* pp. 119-132. Cambridge University Press, Cambridge.

Zimmermann, N.E. & Kienast, F. 1999. Predictive mapping of alpine grasslands in Switzerland: species versus community approach. *J. Veg. Sci.* 10: 469-482.

Received 9 March 1999;
Revision received 10 July 2000;
Accepted 29 December 2000.
Coordinating Editor: C. Leuschner.