Applied Vegetation Science ■■ (2014)

Discriminant analysis reveals limited association between forest habitat types and the environment in western United States land classification

Antonín Kusbach, John D. Shaw & James N. Long

Keywords
Discriminant analysis; Habitat type; Land classification; Plant indicator species; Random Forests; Series

Abbreviations
FIA = Forest Inventory and Analysis; HT = habitat type; NRCS = Natural Resources Conservation Service; PC = principal component; PCA = Principal Components Analysis; PRISM = Parameter-elevation Regression on Independent Slopes Model; RF = Random Forests; USDA = United States Department of Agriculture

Nomenclature
USDA NRCS (2006)

Received 6 June 2013
Accepted 2 May 2014
Co-ordinating Editor: Ladislav Mucina

Kusbach, A. (corresponding author, kusbach@mendelu.cz): Department of Forest Botany, Dendrology and Geobiocoenology, Mendel University in Brno, Zemědělská 3, CZ-613 00, Brno, Czech Republic
Shaw, J.D. (jdshaw@fs.fed.us): USDA Forest Service Rocky Mountain Research Station FIA, 507 25th Street, Ogden, UT, 84401, USA
Long, J.N. (james.long@usu.edu): Quinney College of Natural Resources and Ecology Center, Utah State University, 5230 Old Main Hill, Logan, UT, 84322-5230, USA

Abstract

Aims: Critical assessment of the connection between units of the habitat type system and physiographic, climatic and soil factors in interior western United States land classifications.

Location: Interior western United States including Utah, SE Idaho, W Wyoming and Colorado, N Arizona, NW New Mexico and E Nevada, representing 11 Bailey's ecoregion sections and covering 389 519 km^2.

Methods: We analysed 2754 plots from the extensive FIA database, representing 185 different habitat types. We used two techniques: discriminant analysis represented by Random Forests and ordination represented by principal components analysis to discriminate among habitat type classification units, and assessed their relationships with the physical environment.

Results: Neither habitat types nor series correspond well to environmental (climatic, physiographic, edaphic) differences. Plant indicator species (*sensu* Daubenmire) representing the habitat types generally failed in differentiating between important factors of the physical environment.

Conclusions: The failure of the habitat type classifications in discriminating between key environmental factors calls into question the basic premise that habitat types, as used in much of the western United States, are representative of basic ecological units of land. Given its broad acceptance and importance in land management, a fundamental reexamination of the habitat type concept is warranted.

Introduction

There is a rich history in western North America of using vegetation as indicators of the environment, beginning with Merriam's (1890) description of broad life zones on San Francisco Peak in Arizona and progressing to the biogeoclimatic ecosystem classifications developed in British Columbia (Pojar et al. 1987). In the western United States, the habitat type system pioneered by Rexford Daubenmire (1952) has been used extensively in natural resource management for more than 50 yr (Pfister 1976; Kotar 1988; Wellner 1989; Kusbach et al. 2012).

The habitat type system consists of two general levels of classification: series and habitat type. A series is represented by a single overstorey tree species, e,g. *Abies lasiocarpa*, ABLA. A habitat type is represented by combination of the overstorey species and understorey species, e.g. *Abies lasiocarpa/Vaccinium scoparium*, ABLA/VASC (Pfister & Arno 1980), as a subunit of series. The basis of the habitat type concept and related classifications is the premise that

Doi: 10.1111/avsc.12121

a relatively few species, i.e. one overstorey and one understorey species, indicative of late-successional natural vegetation (Pfister & Arno 1980), reflect 'the algebraic sum of all environmental factors important to plants' (Daubenmire 1976, p. 119). It follows that a given habitat type is indicative of all land units (e.g. forest stands) with the same unique set of important environmental factors (Pfister 1976; Pfister & Arno 1980). Each of the levels, identified by indicator species, selected on the basis of abundance (cover) and frequency (constancy), is indicative of a land segment in which these species would be part of the potential natural vegetation. Therefore, the habitat type system is fundamentally intended to be a land classification rather than a community (vegetation) classification. Frequently, the potential natural vegetation is difficult to assess due to frequent disturbances in western North America (Pfister & Arno 1980). In particular, some communities may not reach maturity or a near-climax stage due to frequent fires, e.g. aspen (*Populus tremuloides*) and lodgepole pine (*Pinus contorta*; Kusbach 2010). Therefore, community types reflecting existing vegetation have been recognized as a part of the habitat type system (e.g. Mauk & Henderson 1984; Mueggler 1988).

Forest habitat types have been widely used for fine-scale land classification since the early 1970s. The system has been used in support of a broad spectrum of theoretical and management applications including: site productivity associated with silvicultural manipulations (e.g. Daubenmire 1976; Monserud 1984; Stage 1989; Fassnacht & Gower 1998), tree regeneration and succession (e.g. Mathiasen et al. 1987; Stansfield & McTague 1992; Sterba & Monserud 1995), genetic variations (Kramer & Johnson 1986) and seed-zone classification (Campbell & Franklin 1981), wildlife distribution and management and forage/browse production (e.g. Kashian et al. 2003), disturbances and biotic hazards, e.g. fire (e.g. Hall et al. 2003), insects (e.g. Eisen et al. 2003) and disease (e.g. Smith & Hoffman 2001), exotic plant species distributions (Pauchard et al. 2003), site properties such as physical and chemical soil factors (e.g. Fosberg et al. 1989; Fassnacht & Gower 1998), hydrology and erosion (e.g. Buckhouse & Mattison 1980), plant physiology (e.g. Myszewski et al. 2002) and recreation and aesthetic considerations (e.g. Layser 1974; Daubenmire 1976; Pfister 1976; Pfister & Arno 1980). The system has also provided a framework for organizing a broad range of ecological and resource management considerations and facilitated interdisciplinary communication (Kotar 1988).

The habitat type concept is illustrated in Fig. 1 following Pfister & Arno 1980, p. 62. Figure 1a illustrates a hypothetical distribution of plots within three different series. Figure 1b illustrates a hypothetical distribution of habitat types within a series. Both series and habitat types are associated with hypothetical gradients in temperature and moisture. The limited overlap between different series and habitat types within environmental space reflects the postulate that both series and habitat types are influenced by, and indicative of, specific sets of environmental factors.

While there has been some questioning of the habitat type concept (e.g. Hall 1985; Cook 1996; Spribille et al. 2001; Kusbach et al. 2012), it has been extensively used in forestland management in the Interior west of the United States for over 50 yr, and remains broadly accepted (e.g. Wellner 1989; Kusbach et al. 2012). However, the fundamental ecological premise of the habitat type concept has never been rigorously tested, and an explicit link between habitat types and the environment remains largely hypothetical (Pfister & Arno 1980). Some studies, especially assessments of the ability of classifications to differentiate site productivity of the series and habitat type levels, have produced ambiguous or negative results (Campbell & Franklin 1981; Kramer & Johnson 1986; Mathiasen et al. 1987; Verbyla & Fisher 1989; Stansfield & McTague 1992). A few studies examining the links between habitat types and specific environmental factors focusing mostly on a limited number of major habitat types within one or two series have had more positive results (e.g. Lawton 1979; Monserud 1984; Mathiasen et al. 1987; Stansfield & McTague 1992). A broad assessment of the basic concept involving e.g. common and rare habitat types (e.g. Mauk & Henderson 1984) or habitat generalists and specialists (Kusbach et al. 2012) and comparisons within and between series has never been done for large geographic areas.

The geographic scope of the study includes a major portion of the western United States, and an exceedingly broad range of environmental conditions. Our objective was to critically assess the connection between series, habitat types and important climatic, physiographic and soil factors in the interior western United States.

Methods

Study area

Bailey (1998) classified large portions of North America according to broad environmental and vegetation characteristics into units he referred to as ecoregions; the ecoregions were further delineated into sections. Eleven Bailey's ecoregion sections covering Utah and extending into adjacent states (ID, WY, CO, NM, AZ and NV; 389 519 km^2 in total) were selected to cover a large portion of the Interior US west. This study area was geographically stratified to the mountainous subarea (M, with steep altitudinal development and prevailing presence of high conifer forest; sections M331D; M331E; M341A, B and M341C) and non-mountainous subarea (non-M,

Applied Vegetation Science
Doi: 10.1111/avsc.12121

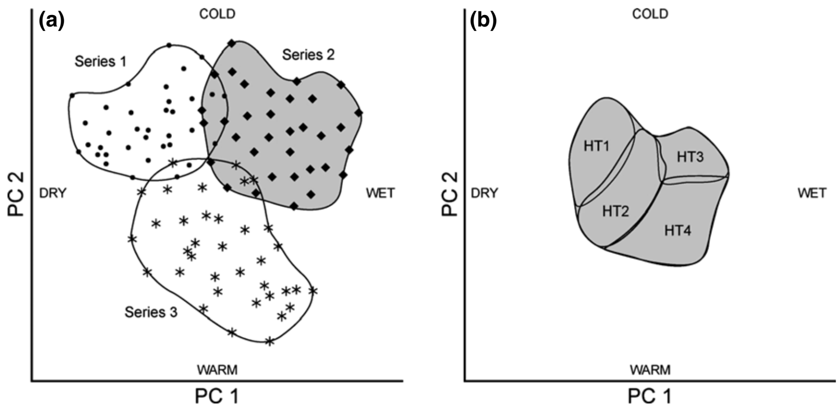

Fig. 1. Hypothetical distribution of plots within three series: (**a**) and habitat types (HT) within series 2, (**b**) with respect to temperature and moisture gradients (after Pfister & Arno 1980).

relatively flat, without excessive altitudinal development and prevailing presence of woodland – low piñyon–juniper conifers; 313A; 341A; 341B; Bailey 1998) which generally lack habitat type classification (e.g. Wellner 1989; Fig. 2).

Data

Primary data were drawn from Forest Inventory and Analysis (FIA) plots measured between 1981 and 2010

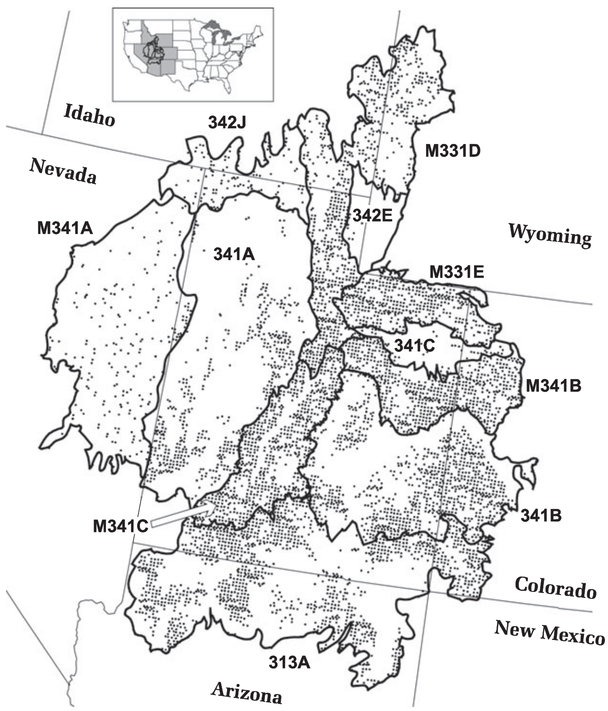

Fig. 2. The study area and distribution of FIA plots (black dots, N = 9754) considered for analysis with sections (labelled after Bailey 1998) and US state boundaries.

(three inventory cycles). FIA plot locations were intersected with the most recent update to sections (Bailey 1998; Cleland et al. 2005). Plot data occurring in the sections of interest fall into three groups: (1) FIA periodic inventory data collected during the period spanning 1981–2002; (2) FIA Phase 2 annual inventory data collected using the current four subplot mapped plot design; and (3) FIA Phase 3 (Forest Health Monitoring) plots that are co-located with approximately 1/16 of the Phase 2 plots, on which additional data are collected.

From Phase 2 plots, we used habitat type and physiographic designation (elevation, slope, aspect and topographic position). Data collected on Phase 3 plots included all Phase 2 data and additional soil variables including properties performed on field-collected samples (O'Neill et al. 2005; Amacher & Perry 2010; Table 1). Because of the lower intensity of Phase 3 plots, the number of plots with soil data (N = 363) was not sufficient for some analysis units – e.g. underrepresented or rare HTs. In order to augment soil data, we sampled soils using Phase 3 field protocols on 115 Phase 2 plots that were scheduled for a regular field visit by FIA crews in 2010 and 2011, bringing the number of plots with soil analysis to 478.

Parameter-elevation Regression on Independent Slopes Model (PRISM)

We assigned climate data to each FIA plot location by intersecting the coordinates for each FIA plot with gridded PRISM (Parameter-elevation Regression on Independent Slopes Model; Daly et al. 2008) data of 1981–2010 monthly and annual average precipitation and temperature. The PRISM data sets were created at 30-arc sec (~800-m) grid resolution (Daly et al. 2008). PRISM uses a two-layer altitudinal atmospheric model to stratify stations of measurements into those below and above inversion

Doi: 10.1111/avsc.12121

Table 1. Types of environmental factor used in the analysis.

PRISM Climatic Factors	Abbreviation	Units/Values	N*	Transformation	Variability†
Mean Monthly Precipitation	ppt_01-12	mm	2754	log	0.99
Mean Minimum Monthly Temperature	tmin_01-12	°C	2754	log	0.98
Mean Maximum Monthly Temperature	tmax_01-12	°C	2754	log	0.96–0.98
Physiographic Factors					
Elevation	elev	Meters	2754	power 2	0.994
Slope Gradient	sl	%	2754	log	1.001
Slope Aspect Value	av	Values 0–1 (Roberts & Cooper 1989)	2754	log	0.997
Physiographic Class	topos	Classes 1–39 (O'Connell et al. 2012)	2754	NA	1
Soil Factors					
Carbon Organic Percentage	C_ORG_PCT	%	478	NA	NA
Carbon Total Percentage	C_TOTAL_PCT	%	478	NA	NA
Carbon Inorganic Percentage	C_INORG_PCT	%	478	NA	NA
Nitrogen Total Percentage	N_TOTAL_PCT	%	478	NA	NA
Average Forest Floor Thickness	FORFLTHK	cm	478	NA	NA
Average Litter Layer Thickness	LTRLRTHK	cm	478	NA	NA
Texture	TXTRLYR	Classes (O'Connell et al. 2012)	478	NA	NA
Coarse Fraction Percent	COARSE_FRACTION_PCT	%	478	NA	NA
pH	PH	NA	478	NA	NA
Exchangeable Sodium	EXCHNG_NA	$mg{\cdot}kg^{-1}$	478	NA	NA
Exchangeable Potassium	EXCHNG_K	$mg{\cdot}kg^{-1}$	478	NA	NA
Exchangeable Magnesium	EXCHNG_MG	$mg{\cdot}kg^{-1}$	478	NA	NA
Exchangeable Calcium	EXCHNG_CA	$mg{\cdot}kg^{-1}$	478	NA	NA
Exchangeable Aluminium	EXCHNG_AL	$mg{\cdot}kg^{-1}$	478	NA	NA
Effective Cation Exchange Capacity	ECEC	$cmolc{\cdot}kg^{-1}$	478	NA	NA
Olsen Phosphorus	OLSEN_P	$mg{\cdot}kg^{-1}$	478	NA	NA
Depth to a Restricted Layer	DPTHSBSL	cm	478	NA	NA

Soil factors except forest floor and litter thickness were considered for layers specified in O'Connell et al. (2012).
*Number of plots in the analysis where the factor was recorded.
†SD of important factors used in PCA.

layers at regional scales (Dobrowolski et al. 2009). PRISM thus captures altitudinal climatic changes together with inversion topoclimatic irregularities (e.g. Lookingbill & Urban 2003). In this study, we used PRISM estimates of mean monthly and annual temperature minimums and maximums, and mean monthly and annual precipitation.

A list of all environmental factors (climatic, physiographic and soil) with important information is provided in Appendix S1 and generally in Table 1.

Data analysis

For plots that had been measured in more than one FIA cycle, we deleted all but the most recent measurement. We also deleted plots without a clear habitat type designation, for example, plots which were only designated at the series level. We also eliminated very rare HTs represented by less than three plots, and plots without climatic (PRISM) data. After all preparation steps, the total number of 9754 FIA plots measured in three inventory cycles available for the analysis representing PRISM, physiographic and soil data (Table 1) and 185 HTs was reduced to 2754 for the final data set. These data were then subjected to discrimination and ordination analysis.

Discrimination

Random Forest (RF) classification (Breiman 2001; Liaw & Wiener 2002; Cutler et al. 2007) was used to: (1) discriminate among units of the habitat type system, and (2) identify those factors of the physical environment most strongly associated with the units of land classification. These units (classes) included broad mountainous and non-mountainous subareas, sections (Bailey 1998) and fine series and habitat types. RF is an efficient method for classification of large data sets containing an array of many qualitatively and functionally different factors. The most influential factors were ranked in the RF variable importance analysis according to Mean Decrease Accuracy. Associations between environmental factors and the classes were based on estimates of' 'out-of-bag' misclassification errors (Chen et al. 2004; Breiman & Cutler 2005). R (v 2.7.2, the randomForest package; http://www.rproject.org/; The R Project for Statistical Computing, Vienna, AT) was used for the RF analysis.

Because of the very high number of PRISM factors (36), an initial RF analysis (2754 plots) was conducted to identify a subset of climatic factors most influential in differentiating the land classifications. We ignored extreme

Applied Vegetation Science
Doi: 10.1111/avsc.12121

climatic outliers with SD >4 (e.g. McCune et al. 2002) since, due their extreme irregularity, we did not expect them to have an important influence on classification. The resulting set of important climatic factors was used in next RF and ordination analyses in combination with the physiographic factors (elevation, slope, aspect, physiographic position). Aspect in degrees was recalculated for aspect values (*Av*) 0–1, where 0 indicated warm and dry sites and 1 the opposite (Roberts & Cooper 1989).

In order to assess the potential effect of soil properties on land classifications, we conducted RF analysis on a subset of 478 plots for which there were soil factors in addition to climatic and physiographic factors. This step was repeated with and without soils for a baseline comparison of the potential improvement of classifications with the addition of soil factors.

The frequency distribution of different habitat types represented in the mountain subarea was used to characterize common vs relatively rare habitat types. We included plots represented by actual habitat type and excluded plots represented by community type (*Pinus contorta* and *Populus tremuloides*). RF analyses on the two resulting data sets were used to compare how common and less abundant habitat types were discriminated by climatic and physiographic factors.

Ordination

Principal components analysis (PCA) was calculated based on a plot x environment (climate and physiographic important factors from RF) matrix in order to detect gradients in environmental space and to assess the association of series and habitat types with these gradients. In order to examine this association in different spatial settings, we conducted PCA for common series and habitat types (for both the same overstorey and understorey spp.) in the entire M subarea (880 plots) and then for section M331D (378 plots) representing SE Idaho, W Wyoming and N Utah (Fig. 2).

The software PC-ORD (v 6.0; MjM Software Design, Gleneden Beach, OR, US) was used for the PCA ordinations. Orthogonal rotations and correlation type of a cross-products matrix were used to identify independent, mutually uncorrelated principal components (PCs; Lattin et al. 2003). We transformed the factors/variables with $|$skewness$|$ >1 to be close to multivariate normality (Table 1, Appendix S1), standardized the data by adjustment to standard deviate (z-scores), and checked the data set for outliers (either factors or plots) using a cut-off of 2.0 SD from the grand mean (McCune et al. 2002). Significance of PCs was tested by Monte Carlo randomization tests based on proportion-based *P*-values and the broken-stick eigenvalue for each PC. To determine the relationship of the environmental variables with the PCs, we calculated correlation coefficients (loadings) with each ordination axis and the linear (parametric Pearson's *r)* and rank (non-parametric Kendall's τ) relationships between the ordination scores and the observed variables in order to assess statistical significance of loadings (McCune et al. 2002).

Results

Climatic factors did not discriminate among series or habitat types. Only broad levels of land classification (subareas and sections) were associated with climate represented by PRISM. The optimal climatic land stratification was the split between the mountain and non-mountain subareas (M and non-M) with the lowest classification error (Table 2). Over multiple runs of the RF analysis of PRISM data, the ranking of the important climatic factors was quite stable for solutions with 6–12 variables randomly used at each split (*mtry* argument in R) and number of trees 500–1000 (*ntree* argument in R) used to grow a 'forest' in the machine-learning process (Liaw & Wiener 2002; Table 2). Similar results were obtained when RF analyses of PRISM data were conducted separately for the mountain and non-mountain subareas (not shown).

Running RF for each subarea when the environment was represented by both climate and physiographic factors (elevation, slope, aspect, physiographic position) gave essentially the same results. This combination of environmental variables discriminated between sections, but again, not between, series or habitat types. For the mountain subarea, the most influential factors in differentiating between sections were: June precipitation, *ppt6* (0.93), August precipitation, *ppt8* (0.93), May min temperature, *tmin5* (0.93), December, January max temperature, *tmax12, 1* (0.90), and elevation, *elev* (0.89). For the non-mountain subarea the most influential climatic factors were: February precipitation, *ppt2* (1.02), June precipitation, *ppt6* (0.99), January max temperature, *tmax1* (0.94), May precipitation, *ppt5* (0.93), December max temperature, *tmax12* (0.86) and elevation, *elev* (0.85). Mountain subarea results are presented in Table 3; non-mountain subarea results are in Appendix S2.

Physical and chemical soil properties were available for 254 mountain and 224 non-mountain subarea plots. RF analysis revealed that the climatic and physiographic factors failed to differentiate any of the classification levels (Table 4). The addition of soil variables resulted in a modest improvement in the misclassification error (Table 5). For this subset of plots, the combination of climatic, physiographic and soil variables did not discriminate between sections, series or habitat types. Similar results for non-mountain plots are in Appendix S3.

Table 2. Random Forests analysis of climatic data for the entire data set. Explanation of the column headers is in the text, for abbreviations see Table 1.

Class/Unit	*mtry/ntree*	Error (%)	No. of Classes	(+)	(−)	Important Factors (in Order of Importance)
Subareas	12/1000	3.4	2	2	0	tmax12,1,2; ppt7; tmax3
Sections	12/1000	6.3	11	7	4	ppt6; tmax12; ppt5; tmax1; ppt7
Series	12/1000	44.7	24	2	22	NA
Habitat Types	6/1000	74.7	110	0	110	NA

(+) number of 'good' units, error <25%; (−) 'poor' units, error >25%.
Number of factors-predictors = 39; N = 2754.

Table 3. Random Forests analysis of climatic and physiographic data for the M subarea. Explanation of the column headers is in the text, for abbreviations see Table 1.

Class/Unit	*mtry/ntree*	Error (%)	No. of Classes	(+)	(−)	Important Factors (in Order of Importance)
Sections	6/1000	10.6	5	5	0	ppt6; ppt8; tmin5; tmax12,1; elev
Series	3/1000	43.2	24	2	22	NA
Habitat Types	3/500	72.7	134	0	134	NA

(+) number of 'good' units, error <25%; (−) 'poor' units, error >25%.
Number of factors-predictors - 9; N = 1533.

Table 4. Random Forests analysis of climatic and physiographic data for the mountain subarea.

Class/Unit	*mtry/ntree*	Error (%)	No. of Classes	(+)	(−)
Sections	14/1000	30.7	5	1	4
Series	7/1000	48.2	15	1	14
Habitat Types	25/1000	64.0	35	3	32

(+) number of 'good' units, error <25%; (−) 'poor' units, error >25%.
Number of factors-predictors - 45; N = 254.

Table 5. Random Forests analysis of climatic, physiographic and soil data for the mountain subarea.

Class/Unit	*mtry/ntree*	Error (%)	No. of Classes	(+)	(−)
Sections	3/1000	28.6	5	2	3
Series	3/1000	54.0	15	0	15
Habitat types	3–6/1000	68.3	35	3	32

(+) number of 'good' units, error <25%; (−) 'poor' units, error >25%.
Number of factors-predictors - 10; N = 254.

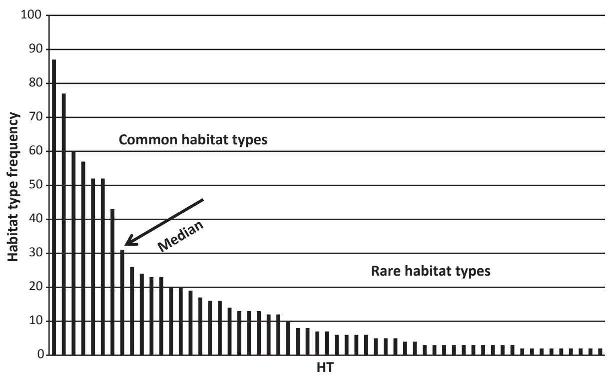

Fig. 3. Frequency distribution of distinct habitat types for the mountain subarea with 880 plots corresponding to 57 different habitat types. The median is suggested as a break between common and rare habitat types.

In the frequency distribution of habitat types for the mountain subarea, the median frequency separated the seven most common (represented by 428 plots) and 50 rarer (represented by 452 plots) habitat types (Fig. 3). We suggest this division is an approximation of major vs minor or incidental habitat types *sensu* Steele et al. (1983), Mauk & Henderson (1984) and Youngblood & Mauk (1985). RF analysis revealed that both common and less abundant habitat types were poorly discriminated by climatic and physiographic factors. The misclassification error was 44.2% (*mtry* = 3, *ntree* = 500) for the major habitat types and 66.2% (*mtry* = 3, *ntree* = 1000) for the minor ones.

The PCA of plots representing common conifer series and important environmental factors (analysed in RF) within the M subarea resulted in two significant PCs ($P < 0.001$), explaining, respectively, 41%, and 19% of the total variance in climatic and physiographic factors (Fig. 4). PC1 was associated with temperatures (*tmax8*, $r = -0.87$, $\tau = -0.65$; *tmin5*, $r = -0.86$, $\tau = -0.65$; *tmax1*, $r = -0.75$, $\tau = -0.55$; *tmax12*, $r = -0.74$, $\tau = -0.56$; *elev*, $r = 0.68$, $\tau = 0.47$); we interpreted this as a temperature gradient, with temperatures decreasing with elevation. PC2 was associated with temperature, precipitation and elevation (*ppt6*, $r = 0.65$, $\tau = 0.40$; *tmax12*, $r = -0.62$, $\tau = -0.40$; *elev*, $r = -0.58$, $\tau = -0.42$; *tmax1*, $r = -0.59$, $\tau = -0.39$); we interpreted this as a complex temperature/moisture gradient, with precipitation increasing and temperature decreasing with increasing elevation.

Distribution of plots in ordination space did not reflect clear separation of the series. While some series, e.g. *Pinus flexilis* and *Picea engelmannii*, seem to occupy fairly discrete

Applied Vegetation Science
Doi: 10.1111/avsc.12121

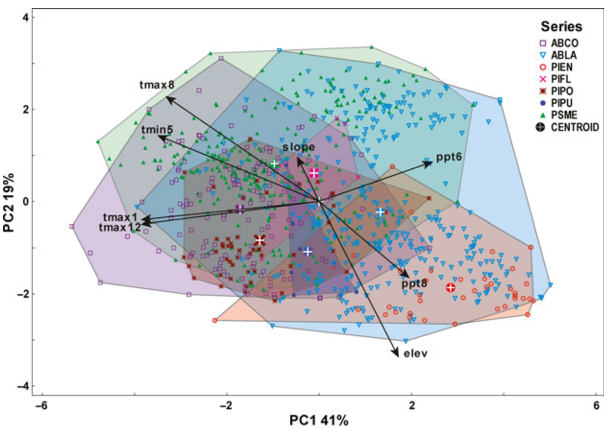

Fig. 4. PCA of major conifer series represented by important environmental factors within the mountain subarea. ABCO, *Abies concolor*; ABLA, *Abies lasiocarpa*; PIEN, *Picea engelmannii*; PIFL, *Pinus flexilis*; PIPO, *Pinus ponderosa*; PIPU, *Picea pungens*; PSME, *Pseudotsuga menziesii*. For factor abbreviations see Table 1.

'niches' with respect to temperature and moisture, overall the pattern is not consistent with what had been postulated (Fig. 1a). For example, the *Abies lasiocarpa* series spans almost the entire range of environmental gradients represented by both PCs. Also noteworthy is that the group centroids of the *Abies concolor*, *Pinus ponderosa* and *Pseudotsuga menziesii* series are very close.

The PCA of plots representing two very common habitat types within the *Pseudotsuga menziesii* series resulted in two significant PCs ($P < 0.001$), explaining, respectively, 32% and 21% of the total variance in climatic and physiographic factors (Fig. 5a). PC1 represents a temperature gradient (*tmax12*, $r = -0.76$, $\tau = -0.58$; *tmin5*, $r = -0.77$, $\tau = -0.54$; *tmax8*, $r = -0.74$, $\tau = -0.53$; *tmax1*, $r = -0.67$, $\tau = -0.48$). PC2 was associated with temperature, precipitation and elevation (*elev*, $r = 0.72$, $\tau = 0.57$; *tmax1*, $r = 0.64$, $\tau = 0.35$; *tmax8*, $r = -0.49$, $\tau = -0.37$; *ppt8*, $r = 0.43$, $\tau = 0.37$; *tmax12*, $r = 0.57$, $\tau = 0.30$). As with the series ordination (Fig. 4), we interpreted this as a complex temperature/moisture gradient, with precipitation increasing and temperature decreasing with increasing elevation.

Distribution of the plots in the environmental ordination space of these two *Pseudotsuga menziesii* series habitat types is not consistent with that which had been postulated (Fig. 1b). Rather, the distributions almost completely overlap and their group centroids are very close (Fig. 5a). When the PCA was repeated for the distribution for series and habitat types within a more geographically restricted area (i.e. the M331D section), the results were similarly negative (Appendices S4 and S5).

The PCA of plots representing two very common habitat types representing the same understorey indicator species in two different series resulted in two significant PCs ($P < 0.001$), explaining, respectively, 44% and 18% of the total variance in climatic and physiographic factors (Fig. 5b). PC1 was associated with temperature and elevation (*tmax8*, $r = -0.94$, $\tau = -0.81$; *tmin5*, $r = -0.92$, $\tau = -0.75$; *tmax1*, $r = -0.84$, $\tau = -0.63$; *elev*, $r = 0.73$, $\tau = 0.58$; *ppt8*, $r = 0.73$, $\tau = 0.56$; *tmax12*, $r = -0.73$, $\tau = -0.48$) and was interpreted as a temperature gradient, with temperature decreasing with elevation. PC2 was associated with aspect and slope (*Av*, $r = -0.71$, $\tau = -0.50$; *slope*, $r = 0.56$, $\tau = 0.35$), perhaps reflecting a mesoclimatic gradient influenced by local topography.

Vaccinium scoparium is a widespread shrub that is used as an understorey indicator species for habitat types in both the *Abies lasciocarpa* and *Picea englemannii* series (Steele et al. 1983; Mauk & Henderson 1984; Youngblood & Mauk 1985). The considerable overlap in ordination space of these two common habitat types from different series is similar to the pattern observed for habitat types in the same series (Fig. 5a). Neither Fig. 5a nor 5b reflects the

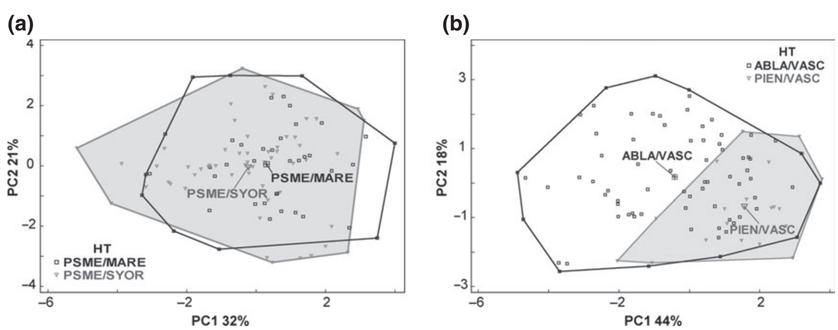

Fig. 5. (a) PCA of two common habitat types (PSME/MARE, *Pseudotsuga menziesii/Mahonia repens*; PSME/SYOR, *Pseudotsuga menziesii/Symphoricarpos oreophilus*) represented by important environmental factors within a single series for the mountain subarea. (b) PCA of two common habitat types (ABLA/VASC, *Abies lasiocarpa/Vaccinium scoparium*; PIEN/VASC, *Picea engelmannii/Vaccinium scoparium*) representing the same understorey indicator species in two different series for the mountain subarea. For factor abbreviations see Table 1.

Doi: 10.1111/avsc.12121

expected separation of habitat types along environmental gradients (Fig. 1b). The results are, however, consistent with the RF analysis (Tables 2–5).

Discussion

Even though it has long been used as a fundamental tool in land management in the western US, the basic premise of the habitat type concept (Fig. 1a,b) and its implications have never been explicitly and thoroughly tested. The few attempts at critical assessment have been limited in scope and the results have been ambiguous (Mathiasen et al. 1987; Spribille et al. 2001; Kusbach et al. 2012).

Our results were not consistent with the fundamental premise that series and habitat types represented by indicator species *sensu* Daubenmire are indicative of the physical environment. Random Forest and PCA revealed the fundamental influence of climate, as represented by PRISM data, on discrimination of broad land units (Table 2). Here, PRISM factors acted as a macroclimatic model strongly associated with elevation. The PRISM factors were not, however, mesoclimatic in the sense of reflecting local, topographically influenced (i.e. aspect and slope) climate (Major 1951). Importantly, neither habitat types nor even series were differentiated by the climatic and physiographic factors (Table 4). We had speculated that the putative relationship between habitat types and the environmental factors was perhaps masked by a few very common and widely distributed habitat types. However, even the less common habitat types were not differentiated by the climatic and physiographic factors any better than were the common habitat types. For a subset of plots for which we had soils data, neither habitat types nor series were differentiated by the whole suite of climatic, physiographic and soil factors (Table 5).

Although ordinations of plots represented by series and habitat types and important environmental factors provided significant results, the distribution of plots within the various classes was not at all consistent with those which had been postulated (Fig. 1a,b). Rather, there was considerable overlap of plot distributions and in some cases the overlap was complete. There were some series, e.g. *Picea engelmannii* and *Pinus flexilis*, that may occupy mostly non-overlapping environmental spaces (Fig. 4). More typically, however, there was considerable, and in some cases near total, overlap of the series. The considerable ecological amplitude of *Abies lasiocarpa* means that its series is represented on all but the warmest sites (Fig. 4). The ordination of plots representing two common habitat types in the *Pseudotsuga menziesii* series (*P. menziesii*/*Mahonia repens* and *P. menziesii*/*Symphoricarpos oreophilus*) illustrate the general failure of the habitat types as indicators of the climatic and physiographic factors (Fig. 5a). A similar result was obtained from the ordination of two other common habitat types. *Vaccinium scoparium* is a widely distributed species in the Rocky Mountains. Indeed, it is so common that it is the understorey indicator species for habitat types in both the *Abies lasiocarpa* and *Picea engelmannii* series. The modest separation in environmental space associated with these two habitat types appears to be the result of series differences (Fig. 5b).

Clearly, these results raise serious questions about the usefulness of habitat types as the basis for ecological land classification. The considerable overlap between series and habitat types within the environmental ordination spaces are counter to the premise that habitat types are strongly influenced by, and are indicative of, specific sets of environmental factors. This, in turn, suggests that habitat types may have fairly limited utility in land management, e.g. development of silviculture prescriptions, fire and fuels management, and invasive plant control.

There are, of course, potential limitations in our analysis of the habitat type concept. The scale of the climatic data may not be entirely appropriate. On the other hand, the broad climate characterization represented by the PRISM data did not even effectively discriminate between series (Table 2) and finer-scale mesoclimatic data including extrapolation of climatic meso-scale extremes (e.g. cold air drainage, late snow and frost) were not available for such a large area. It is also noteworthy that the suite of environmental factors used in the analysis included slope, aspect, physiographic class and physical soil factors, which in combination should reflect important differences in mesoclimate. Nevertheless, the climatic, physiographic and soil factors together failed to discriminate between habitat types or even series (Tables 3–5). It is possible that some critical physiographic or soil factors were missing from the broad suite of environmental factors used in our analysis. For example, digital elevation model techniques, which can calculate topography-based indices (wetness, heat, topo exposure, etc.), could be promising proxies of cardinal topography-dependent factors (soil moisture and temperature). It is also possible that additional spatial/functional stratification of either subareas or sections is necessary to capture potential mesoclimatic and soil differences (Kusbach et al. 2014). Additionally, our essentially negative results could stem from common misidentification of series and habitat types in the field. We think this is unlikely since the FIA crews are well trained in the use of habitat typing procedures and keys.

We believe the most likely explanation for our results is ultimately the way that habitat types are initially characterized. We suggest that the use of a single overstorey species and one or two common understorey species is almost certainly an inappropriate basis for an effective ecological land classification system. The choice of indicator species in the habitat type system is abundance-based and favours

Applied Vegetation Science
Doi: 10.1111/avsc.12121

generalists (Kusbach et al. 2012). It is perhaps not surprising that these common species generally fail as indicators and discriminators of important factors of the physical environment since they are, almost by definition, species with fairly broad ecological amplitude.

These results should not be taken as evidence against the potential for plant community classifications to provide considerable insight into the physical environment. We propose that a more effective system of land classification should begin with explicit recognition and classification of climatic and edaphic environmental factors, and subsequent coupling of that product with an enhanced vegetation classification (Kusbach et al. 2012). The added superstructure represented by broader units to the system would reflect broad overstorey species distributions and overall climatic patterns as represented, e.g. by biogeoclimatic zones (Pojar et al. 1987) or forest vegetation tiers/belts (Zlatník 1976). The basic units of land classification (analogous to series and habitat types) would be 'hung' on this superstructure. The inclusion of soil descriptors in such a comprehensive classification would undoubtedly improve its overall utility (Kusbach et al. 2014). Preliminary study will be needed to determine the specific chemical and physical soil properties representing an appropriate balance of cost and effective land classification. Equally important, the vegetation component of the land classification system should not be based on a few common indicator species, but on a suite of diagnostic species (Chytrý et al. 2002; Jennings et al. 2009).

Conclusions

This geographically extensive study revealed a serious limitation of the habitat type concept. The results were not consistent with the fundamental premise that series and habitat types represented by indicator species *sensu* Daubenmire are indicative of the physical environment. Common indicator species are not an appropriate basis for land classification. The failure of the fundamental conceptual premise calls into question that habitat types, as have long been used as a basic tool in land management in much of the western US, are representative of basic ecological units of land. Given its broad acceptance and importance in land management, a fundamental re-examination of the habitat type concept is warranted.

Acknowledgements

This research was supported by USDA Forest Service Rocky Mountain Research Station FIA (Agreement No. 09-JV-11221638-260) and the Utah Agricultural Experiment Station, Utah State University, and approved as journal paper no. 8555. We thank Robert Justin DeRose, USDA Forest Service, for helpful advice and Wanda Lindquist for invaluable support with the figures. This paper was prepared in part by an employee of the US Forest Service as part of official duties and therefore is in the public domain.

References

Amacher, M.C. & Perry, C.H. 2010. *The soil indicator of forest health in the forest inventory and analysis program*. United States Department of Agriculture Forest Service Intermountain Research Station [Proceedings RMRS-P-59], Ogden, UT, US.

Bailey, R.G. 1998. *Ecoregions map of North America*. United States Department of Agriculture, Forest Service [Miscellaneous Publication no. 1548], Washington, DC, US.

Breiman, L. 2001. Random forests. *Machine Learning* 45: 5–32.

Breiman, L. & Cutler, A. 2005. *Random forests*. Utah State University, Department of Mathematics and Statistics, Logan, UT, US.

Buckhouse, J.C. & Mattison, J.F. 1980. Potential soil erosion of selected habitat types in the high desert region of central Oregon. *Journal of Range Management* 33: 282–285.

Campbell, R.K. & Franklin, J.F. 1981. A comparison of habitat type and elevation for seed-zone classification of Douglas-fir in western Oregon. *Forest Science* 27: 49–59.

Chen, C., Liaw, A. & Breiman, L. 2004. *Using random forest to learn imbalanced data*. Statistics Department of University of California, Berkeley, CA, US, [Technical Report 666].

Chytrý, M., Exner, A., Hrivnák, R., Ujházy, K., Valachovič, M. & Willner, W. 2002. Context-dependence of diagnostic species: a case study of the central European spruce forests. *Folia Geobotanica* 37: 403–441.

Cleland, D.T., Freeouf, J.A., Keys, J.E., Nowacki, G.J., Carpenter, C.A. & McNab, W.H. 2005. Ecological Subregions: Sections And Subsections for the Conterminous United States, Presentation Scale 1:3,500,000; Colored. United States Department of Agriculture Forest Service, Washington, DC, US.

Cook, J.E. 1996. Implications of modern successional theory for habitat typing: a review. *Forest Science* 42: 67–75.

Cutler, D.R., Edwards, T.C. Jr, Beard, K.H., Cutler, A., Hess, K.T., Gibson, J. & Lawler, J.J. 2007. Random forests for classification in ecology. *Ecology* 88: 2783–2792.

Daly, C., Halbleib, M., Smith, J.I., Gibson, W.P., Dogget, M.K., Taylor, G.H., Curtis, J. & Pasteris, P.P. 2008. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology* 28: 2031–2064.

Daubenmire, R.F. 1952. Forest vegetation of northern Idaho and adjacent Washington, and its bearing on concept of vegetation classification. *Ecological Monographs* 22: 301–330.

Daubenmire, R.F. 1976. The use of vegetation in assessing the productivity of forest lands. *The Botanical Review* 42: 115–143.

Dobrowolski, S.Z., Abatzoglou, J.T., Greenberg, J.A. & Schladow, S.G. 2009. How much influence does landscape-scale physiography have on air temperature in a mountain environment? *Agricultural and Forest Meteorology* 149: 1751–1758.

Eisen, R.J., Eisen, L., Castro, M.B. & Lane, R.S. 2003. Environmentally related variability in risk of exposure to lyme disease spirochetes in northern California: effect of climatic conditions and habitat type. *Environmental Entomology* 32: 1010–1018.

Fassnacht, K.S. & Gower, S.T. 1998. Comparison of soil and vegetation characteristics of six upland forest habitat types in North Central Wisconsin. *Northern Journal of Applied Forestry* 15: 69–76.

Fosberg, M.A., Hironaka, M. & Houston, K.E. 1989. The soil–habitat type relationship. In: Ferguson, D.E., Morgan, P. & Johnson, F.D. (eds.) *Proceedings – land classifications based on vegetation: applications for resource management*, pp. 277–278. United States Department of Agriculture Forest Service Intermountain Research Station [GTR INT-257], Ogden, UT, US.

Hall, F.C. 1985. The habitat type controversy; two common concepts. *Rangelands* 7: 170–171.

Hall, W.L., Zuuring, H.R., Hardy, C.C. & Wakimoto, R.D. 2003. Applying logistic regression to determine regeneration risk to stand replacement fire on the Kootenai National Forest, Montana. *Western Journal of Applied Forestry* 18: 155–162.

Jennings, M.D., Faber-Langendoen, D., Loucks, O.L., Peet, R.K. & Roberts, D. 2009. Standards for associations and alliances of the U.S. National Vegetation Classification. *Ecological Monographs* 79: 173–199.

Kashian, D.M., Barnes, B.V. & Walker, W.S. 2003. Landscape ecosystems of northern lower Michigan and the occurrence and management of the Kirtland's Warbler. *Forest Science* 49: 140–159.

Kotar, J. 1988. Ecological land classification – the first step in forest resource management. In: Johnson, J.E. (ed.) *Managing North Central forests for non-timber values*, pp. 43–52. Proceedings of the Fourth Society of American Foresters Region V Technical Conference, Nov 29–Dec 1, 1988, Duluth, MI, US. SAF publication 88-04, Bethesda, MA, US.

Kramer, N.B. & Johnson, F.D. 1986. Mature forest seed banks of three habitat types in central Idaho. *Canadian Journal of Botany* 65: 1961–1966.

Kusbach, A. 2010. *Vegetation geo-climatic zonation in the Rocky Mountains, northern Utah*. PhD dissertation, Utah State University, Logan, UT, US.

Kusbach, A., Long, J.N. & Van Miegroet, H. 2012. Fidelity and diagnostic species concepts in vegetation classification in the Rocky Mountains, northern Utah, USA. *Botany* 90: 1–16.

Kusbach, A., Van Miegroet, H., Boettinger, J.L. & Long, J.N. 2014. Vegetation geo-climatic zonation in the Rocky Mountains, northern Utah, USA. *Journal of Mountain Science* 11: 656–673.

Lattin, J.M., Carrol, J.D. & Green, P.E. 2003. *Analyzing multivariate data*. Brooks/Cole, Thompson Learning, Pacific Grove, CA, US.

Lawton, P.M. 1979. *An investigation of the environmental relationships of selected forest habitat types in northern Utah*. MS thesis, Utah State University, Logan, UT, US.

Layser, E.F. 1974. Vegetative classification: its application to forestry in the Northern Rocky Mountains. *Journal of Forestry* 72: 354–357.

Liaw, A. & Wiener, M. 2002. Classification and regression by Random forests. *R News: The Newsletter of the R Project 2*. pp. 18–22.

Lookingbill, T.R. & Urban, D.L. 2003. Spatial estimation of air temperature differences for landscape scale studies in montane environments. *Agricultural and Forest Meteorology* 114: 141–151.

Major, J. 1951. A functional, factorial approach to plant ecology. *Ecology* 32: 392–412.

Mathiasen, R.L., Blake, E.A. & Edminster, C.B. 1987. Estimates of site potential for ponderosa pine based on site index for several southwestern habitat types. *Great Basin Naturalist* 47: 467–472.

Mauk, R.L. & Henderson, J.A. 1984. *Coniferous forest habitat types of Northern Utah*. United States Department of Agriculture Forest Service Intermountain Forest and Range Experiment Station [GTR INT-170], Ogden, UT, US.

McCune, B., Grace, J.B. & Urban, D.L. 2002. *Analysis of ecological communities*, 2nd ed. MjM Software Design, Gleneden Beach, OR, US.

Merriam, C.H. 1890. Results of a biological survey of the San Francisco Mountain region and the desert of the Little Colorado, Arizona. *USDA, North American Fauna* 3: 1–136.

Monserud, R.A. 1984. Height growth and site index curves for Inland Douglas-fir based on stem analysis data and forest habitat type. *Forest Science* 30: 943–965.

Mueggler, W.F. 1988. *Aspen community types of the Intermountain Region*. United States Department of Agriculture Forest Service Intermountain Research Station [GTR INT-250], Ogden, UT, US.

Myszewski, J.H., Fins, L., Moore, J.A., Rust, M. & Mika, P.G. 2002. Variation in the root bark phenolics/sugar ratio of Douglas-fir grown in two plantations in northern Idaho. *Canadian Journal of Forest Research* 32: 556–560.

O'Connell, B.M., LaPoint, E.B., Turner, J.A., Ridley, T., Boyer, D., Wilson, A.M., Waddell, K.L., Christensen, G. & Conkling, B.L. 2012. *The forest inventory and analysis database: database description and users manual. Version 5.1.2 for phase 2*. United States Department of Agriculture Forest Service Forest Inventory and Analysis, Arlington, VA, US.

O'Neill, K.P., Amacher, M.C. & Perry, C.H. 2005. *Soils as an indicator of forest health: a guide to the collection, analysis, and interpretation of soil indicator data in the Forest Inventory and Analysis program*. United States Department of Agriculture Forest Service North Central Research Station [GTR NC-258], St. Paul, MN, US.

Pauchard, A., Alaback, P.B. & Edlund, E.D. 2003. Plant invasions in protected areas at multiple scales: *Linaria vulgaris*

(Scrophulariaceae) in the West Yellowstone area. *Western North American Naturalist* 63: 416–428.

Pfister, R.D. 1976. *Forest habitat type classification*. Mapping, and ecoregions in the northern Rocky Mountains. Discussion paper, XVI IUFRO World Congress, June 20-July 2, Oslo, NO.

Pfister, R.D. & Arno, S.F. 1980. Classifying forest habitat types based on potential climax vegetation. *Forest Science* 26: 52–70.

Pojar, J., Klinka, K. & Meidinger, D.V. 1987. Biogeoclimatic classification in British Columbia. *Forest Ecology and Management* 22: 119–154.

Roberts, D.W. & Cooper, S.V. 1989. Concepts and techniques in vegetation mapping. In: Ferguson, D.E., Morgan, P. & Johnson, F.D. (eds.) *Proceedings – land classifications based on vegetation: applications for resource management*, pp. 90–96. United States Department of Agriculture Forest Service Intermountain Research Station [GTR INT-257], Ogden, UT, US.

Smith, J.P. & Hoffman, J.T. 2001. Site and stand characteristics related to white pine blister rust in high-elevation forests of southern Idaho and western Wyoming. *Western North American Naturalist* 61: 409–416.

Spribille, T., Stroh, H.G. & Triepke, F.J. 2001. Are habitat types compatible with floristically-derived associations? *Journal of Vegetation Science* 12: 791–796.

Stage, A.R. 1989. Utility of vegetation-based land classes for predicting forest regeneration and growth. In: Ferguson, D.E., Morgan, P. & Johnson, F.D. (eds.) *Proceedings – land classifications based on vegetation: applications for resource management*, pp. 180–185. United States Department of Agriculture Forest Service Intermountain Research Station [GTR INT-257], Ogden, UT, US.

Stansfield, W.F. & McTague, J.P. 1992. Dominant height and site index equations for Douglas-fir and Engelmann spruce in East-Central Arizona. *Western Journal of Applied Forestry* 7: 40–44.

Steele, R., Cooper, S.V., Ondov, D.M., Roberts, D.W. & Pfister, R.D. 1983. *Forest habitat types of eastern Idaho-western Wyoming*. USDA Forest Service Intermountain Forest and Range Experiment Station [GTR INT-114], Ogden, UT, US.

Sterba, H. & Monserud, R.A. 1995. Potential volume yield for mixed-species Douglas-fir stands in the northern Rocky Mountains. *Forest Science* 41: 531–545.

USDA NRCS, National Resources Conservation Service & Skinner, M.W. 2006. *The PLANTS database. Version 3.5*. National Plant Data Center, Baton Rouge, LA, US.

Verbyla, D.L. & Fisher, R.F. 1989. Ponderosa pine habitat types as an indicator of site quality in the Dixie National Forest, Utah. *Western Journal of Applied Forestry* 4: 52–54.

Wellner, C.A. 1989. Classification of habitat types in the western United States. In: Ferguson, D.E., Morgan, P. & Johnson, F.D. (eds.) *Proceedings – land classifications based on vegetation: application for resource management*, pp. 7–21. United States Department of Agriculture Forest Service Intermountain Research Station [GTR-INT-257], Ogden, UT, US.

Youngblood, A.P. & Mauk, R.L. 1985. *Coniferous forest habitat types of central and southern Utah*. United States Department of Agriculture Forest Service Intermountain Research Station [GTR-INT-187], Ogden, UT, US.

Zlatník, A. 1976. Přehled skupin typů geobiocénů původně lesních a křovinných. [Forest and shrub geobiocoenosis type groups: an overview]. *Zprávy Geografického ústavu ČSAV v Brně* 13: 55–60.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Complete list of climatic, physiographic and soil factors considered in the study analysis.
Appendix S2. Random Forests analysis of PRISM and physiographic data for the non-M subarea.
Appendix S3. Random Forests analysis of PRISM, physiographic and soil data for the non-M subarea.
Appendix S4. PCA of major conifer series/cover types within the M331D section.
Appendix S5. PCA of common habitat types within the PSME series for the M331D section.