available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/ecolinf

Predictive modeling and mapping sage grouse (*Centrocercus urophasianus*) nesting habitat using Maximum Entropy and a long-term dataset from Southern Oregon

Andrew C. Yost[a,*], *Steven L. Petersen*[b], *Michael Gregg*[c], *Richard Miller*[d]

[a]*Oregon Department of Forestry, 2600 State St., Salem OR 97310, United States*
[b]*Department of Plant and Wildlife Sciences, 275 WIDB, Brigham Young University, Provo, UT 84602, United States*
[c]*U.S. Fish and Wildlife Service, Mid-Columbia River NWR Complex, 64 Maple St., Burbank, WA 99323, United States*
[d]*Department of Rangeland Ecology and Management, 202 Strand Ag Hall, Oregon State University, Corvallis, OR 97331-2218, United States*

ARTICLE DATA

Article history:
Received 5 March 2008
Received in revised form
12 August 2008
Accepted 18 August 2008

Keywords:
Sage grouse
Predictive modeling and mapping
Maximum Entropy
Hart Mountain National Antelope Refuge

ABSTRACT

Predictive modeling and mapping based on the quantitative relationships between a species and the biophysical features (predictor variables) of the ecosystem in which it occurs can provide fundamental information for developing sustainable resource management policies for species and ecosystems. To create management strategies with the goal of sustaining a species such as sage grouse (*Centrocercus urophasianus*), whose distribution throughout North America has declined by approximately 50%, land management agencies need to know what attributes of the range they now inhabit will keep populations sustainable and which attributes attract disproportionate levels of use within a home range. The objectives of this study were to 1) quantify the relationships between sage grouse nest-site locations and a set of associated biophysical attributes using Maximum Entropy, 2) find the best subset of predictor variables that explain the data adequately, 3) create quantitative sage grouse distribution maps representing the relative likelihood of nest-site habitat based on those relationships, and 3) evaluate the implications of the results for future management of sage grouse. Nest-site location data from 1995 to 2003 were collected as part of a long-term research program on sage grouse reproductive ecology at Hart Mountain National Antelope Refuge. Two types of models were created: 1) with a set of predictor variables derived from digital elevation models, a field-validated vegetation classification, and UTM coordinates and 2) with the same predictors and UTM coordinates excluded. East UTM emerged as the most important predictor variable in the first type of model followed by the vegetation classification which was the most important predictor in the second type of model. The average training gain from ten modeling runs using all presence records and randomized background points was used to select the best subset of predictors. A predictive map of sage grouse nest-site habitat created from the application of the model to the study area showed strong overlap between model predictions and nest-site locations.

1. Introduction

Understanding the quantitative relationships between a species and the biophysical features of the ecosystem in which it occurs is fundamental when developing a sustainable resource management policy for that species and ecosystem. Predictive modeling and mapping that is based on these relationships forms an analytical foundation for informed

* *Corresponding author*. Tel.: +1 503 945 7410.
E-mail address: ayost@odf.state.or.us (A.C. Yost).

1574-9541/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecoinf.2008.08.004

conservation planning, mapping patterns of biodiversity, detecting distributional changes from monitoring data, and quantifying how variation in species performance relates to one or more controlling factors (Guisan and Hofer, 2003; McCune, 2006; Phillips et al., 2006). For example, to create management strategies with the goal of sustaining a species such as sage grouse (*Centrocercus urophasianus*) whose distribution throughout North America has declined by approximately 50% (Aldridge and Brigham, 2002), land management agencies need to know what portions of their former range they now inhabit and what attributes make these areas capable of population sustainability and attract disproportionate levels of use within a home range. This study tests the application of predictive modeling and mapping to sage grouse nesting habitat as a method for generating information valuable to a sustainable wildlife management policy.

Considerable work has been conducted evaluating habitat attributes at the site level for sage grouse nesting (Aldridge and Brigham, 2002; Gregg et al., 1994; Holloran et al., 2005; Sveum et al., 1998), however, few studies have evaluated the attributes at the landscape level. Gregg (2006) studied the nutritional ecology of sage grouse productivity and chick survival in Southern Oregon and northern Nevada. He found forb consumption and high insect availability in Spring were important for brood production and chick survival. The quality and availability of nutritional resources, however, are not distributed homogeneously across the landscape nor are the optimal locations that sage grouse select for nesting. Even if food resources were abundant and of high quality, selection of a nest site that increases the chances of exposure to predation or lethal climate conditions would have negative effects on grouse productivity. Moreover, how the spatial distribution of these and other attributes (i.e. topography) influence site selection and distribution during pre-nesting, nesting, rearing, or wintering at the landscape level is unknown. Gregg's (2006) study produced an extensive database of sage grouse nest-site locations over an eight-year time period sufficient for creating predictive models of the relationships between nest-site location and the biophysical attributes that might be important in sage grouse productivity. Therefore, the purpose of this study was to quantify the relative importance of the relationships between nest-site locations with the biophysical features that accompany those locations and then map the spatial distribution of sage grouse nesting habitat.

Advancements in computer technology, statistical modeling, and Geographic Information Systems (GIS) software allow the knowledge of animal/habitat relationships to be used for predicting the geographic distribution of individual populations of wildlife species. Predictive species mapping was defined by Franklin (1995) as predicting the distribution of a particular species across a landscape from mapped environmental variables. Predictive species mapping is founded in ecological niche theory and gradient analysis and rests on the premise that species distributions can be predicted from the spatial distributions of environmental variables that correlate with or control the occurrence of a plant or animal. Environmental conditions at occurrence localities constitute samples from a species' realized niche which is smaller than its fundamental niche (Hutchinson, 1957; Phillips et al., 2006). There are three major steps involved with predictive modeling and mapping: 1) collect species-level occurrence data and associated biophysical attributes of the landscape, preferably with a randomized sampling design, 2) build the models to determine the best subset of predictors and their parameter coefficients, and 3) application of the models to GIS data or new sites to forecast probability of occurrence for unsampled locations within the range of the study area.

Unlike vegetation monitoring datasets that typically contain some sampling sites with a particular species present and some where it was absent (Yost, 2008), wildlife sampling datasets often consist of "presence-only" data. General purpose statistical methods such as generalized linear models can be used for presence/absence datasets but there are a limited number of options available for presence-only datasets. Recently, Phillips et al. (2004, 2006) introduced the use of the Maximum Entropy (Maxent) method for modeling species geographic distributions with presence-only data. Maxent is a general purpose machine learning method for making predictions or inferences from incomplete information. The method estimates a target probability distribution across a study area by finding the probability distribution that is closest to uniform, or spread-out, subject to a set of constraints that represent our incomplete information about the target distribution. The information available about the target distribution presents itself as a set of real-valued variables, or "features" and the constraints are that the expected value of each feature should match its empirical average (average value for a set of sample points taken from the target distribution). When applied to presence-only distribution modeling, the pixels of the study area make up the space on which the probability distribution is defined, pixels with known species occurrence records constitute the sample points, and the features are the predictor variables that have digital geographic representation.

In addition to creating quantitative probability maps, the shape of response function and strength of predictability for each predictor variable can be graphically and quantitatively evaluated. This provides the capability to discover which gradients are most influential in predicting the likely occurrence of a particular species given they can be represented in a geographic database. Knowledge of the strength and functional response of species occurrences with each predictor provides valuable information for identifying which landscape features should be the focus of management for habitat sustainability.

The specific objectives of this study were to 1) quantify the relationships between sage grouse nest-site locations and a set of associated biophysical attributes with Maxent, 2) find the best subset of predictor variables that explain the data adequately, 3) create quantitative sage grouse distribution maps representing the relative likelihood of nest-site habitat based on those relationships, and 4) evaluate the implications of the results for future management of sage grouse.

2. Materials and methods

2.1. Data and study area

Locations of sage grouse nest sites from 1995 to 2003 were collected as part of a long-term research program on sage