Forest Ecology and Management 358 (2015) 365–368

Contents lists available at ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

Correspondence

Range wide analysis of northern spotted owl habitat relations: Reply to comments

Craig Loehle [a,*], Larry Irwin [b]

[a] National Council for Air and Stream Improvement, Inc., 552 S Washington Street, Suite 224, Naperville, IL 60540, USA
[b] PO Box 68, Stevensville, MT 59870, USA

ARTICLE INFO

Article history:
Received 17 July 2015
Accepted 17 July 2015

ABSTRACT

Bell et al. (2015) and Dunk et al. (2015) comment on our appraisal (Loehle et al., 2015) of biological insights from the US Fish and Wildlife Service models for northern spotted owl critical habitat. We here respond to those comments. We argue that while the low predictability of vegetation plot data by the gradient nearest neighbor (GNN) models may average out at very large scales and thus be useful in that context, errors at the site-specific scale may confound the modeling used to develop critical habitat designations. We further found that GNN errors violate statistical assumptions and are not propagated through the modeling exercise. We found multiple lines of evidence for habitat model instability, which may result from GNN uncertainty. We believe our evidence for lack of demographic predictability from the MaxEnt RHS values remains relevant to judicious use of these models for conservation. We similarly respond to other particular concerns with our analysis and conclude with suggestions.

1. Introduction

In this reply we respond to two comments (Bell et al., 2015; Dunk et al., 2015) on our assessment (Loehle et al., 2015) of the critical habitat modeling of the northern spotted owl (NSO) by the US Fish and Wildlife Service (FWS).

As part of critical habitat designation for the NSO, FWS (2011) developed models for NSO nest site occurrence based on gradient nearest neighbor (GNN) vegetation and other data. The FWS used the MaxEnt statistical tool and a large database of owl nest site locations to construct habitat models for 11 regions. This exercise has considerable impact on land management and the economy across the Pacific Northwest and northern California. There were several reasons for our study of the modeling effort. First, we wondered about biological interpretability. For example, Dugger and Davis (2011) noted that the effects of coarse scale measures of suitable habitat and barred owls on NSO population performance are just beginning to be understood. With a limited link to demography, it seems challenging to design effective reserves or management plans. Second, we had questions about the underlying GNN database and the ability of tools such as MaxEnt to identify NSO habitat.

Loehle et al. (2015), therefore, aimed to evaluate the FWS modeling effort using independent modeling methods and test data. We found that relative habitat suitability (RHS) from MaxEnt for our test sites was not statistically related to owl reproductive performance. Our independent surveys found both false negatives and false positives that were disconcertingly large. We found evidence for GNN data deficiencies. Our independent modeling exercises led us to conclude that the FWS habitat models were unstable, probably due to the GNN issues.

The two comments on our paper raise a number of issues. Some are statistical. Some involve opinions about how accurate, meaningful, and useful the results of such an exercise are. Here we respond to these criticisms, beginning with Bell et al. (2015).

2. GNN data adequacy

The base maps for the FWS analysis are the GNN vegetation maps, plus GIS-map layers for abiotic variables. GNN (Ohmann and Gregory, 2002) uses Forest Service inventory plots to calibrate and spatially interpolate satellite-sensed data for attributes such as tree species, basal area, and tree density. Bell et al. (2015) defend the GNN approach and make several claims about our test of GNN.

In Loehle et al. (2015), we utilized standard variable radius inventory plots measured within the 95% utilization polygon determined from radio-tracking data for multiple owls. Bell et al. argue

* Corresponding author.
E-mail addresses: cloehle@ncasi.org (C. Loehle), llirwin@blackfoot.net (L. Irwin).

http://dx.doi.org/10.1016/j.foreco.2015.09.033
0378-1127/

first that the average plot size, being smaller than the 30 × 30 m GNN predicted pixels, could produce higher variability and thus lower correlations. Because our plot data were not taken with testing GNN in mind, this plot size discrepancy may have biased our correlations down somewhat. However, the LiDAR plot data test correlations (around 0.3–0.5) in Bell et al. (their Table 1) are also quite low.

A second criticism was that our use of samples only within the home range produced an incomplete assessment of GNN. Their assessment of this hypothesis in Table 1 actually shows LiDAR plots (both scales) for all data vs. those only within 800 m to be as likely to have better as worse correlation, so their data do not support this claim. The larger LiDAR plots alone give modest support. We also note that only within 125 m from nest sites did we find that about 40–50% of plots were sampled in older stands (e.g., Irwin et al., 2012). This percentage declined rapidly, such that >75% of plots beyond 1.5 km were not in old forests, where conditions approached landscape condition. Thus, only a small proportion of samples was in old forest and the majority was sampled in non-old forest, providing a broad range of forest conditions. In any case, we were only testing the ability of GNN to predict plot characteristics within the home range.

Finally they raise the question of scale and the use of the data. They argue that larger scale (e.g., landscape) estimates of forest attributes are likely to be more accurate than fine scale measures, and therefore adequate for informing land managers. While this seems possible, they admit that no test data exist at the 200-ha scale used by FWS for MaxEnt analyses nor at larger extents.

This brings up the question of GNN adequacy for the purpose of habitat modeling. During GNN data development, a leave-one-out (LOO) procedure was used for testing. Most of their LiDAR test data correlations are lower than the LOO correlations (their Table 1). These correlations need a context. In a typical statistical analysis, the independent variable such as tree diameter is assumed to be measured with very little error. The variability in the dependent variable is assumed to mostly represent unmeasured factors (e.g., genetics, weather) with some measurement error. In this case, the low correlations mean that the independent variables being used in the MaxEnt analysis are not estimated by GNN very well at all, in contrast to the assumptions of most statistical methods. As an example, in accuracy assessment reports (http://lemma.forestry.oregonstate.edu) the predicted GNN vs. observed correlations for some variables are less than 0.4. This means that very little of the variance in the data is explained (<16%), or put another way the actual state of a plot is unlikely to have been predicted correctly. While the accuracy reports find that overall landscape statistics are closer than plot-level values, this is not clearly useful when trying to evaluate habitat patches within 200-ha areas for owl conservation. For such use of secondary data, one must also consider bias, not just correlation. Data vs. predicted should fall along a 45-degree line. A correlation of 1.0 can be obtained by any line deviating from 45 degrees, but such biased data are not useful as input for secondary modeling or management. The plots in Ohmann and Gregory (2002, Fig. 10) and at the lemma website show visual evidence of bias as well as very low values for R^2 sometimes near zero. These violations of statistical assumptions are serious. It is taking as input independent variables such as conifer basal area that have huge error terms and treating them as if they were accurately measured. There is also no formal propagation of error from the highly uncertain GNN results through to MaxEnt output. For example, in regression when the independent variable has measurement error it is recommended to either use Reduced Major Axis regression or adjust the regression slope (e.g., Smith, 2009). For spatial data used for estimating hydrologic response, effective methods for assessing input data error have been suggested (Hong et al., 2006). But in the MaxEnt analysis, which is essentially a regression approach, there is no consideration of large errors in the independent variables. All of this is not to deny that GNN output might have some descriptive utility at the landscape scale, but the practice of treating highly uncertain independent variables as if they were accurately measured for the purpose of habitat modeling is questionable. It is one thing to state that variable *X*, accurately measured, contributes only 10% to predictions of nest site location, but another entirely to state that *X* was only measured with 10% accuracy. In the latter case, the door is wide open for spurious correlation and large uncertainty.

3. MaxEnt modeling

In the second comment (Dunk et al., 2015), a variety of criticisms are made. We here address these points in order.

Dunk et al. (2015) criticize our recent paper (Loehle et al., 2015) by suggesting that we mischaracterized the literature on NSOs. In fact, we agreed with Dunk et al. (2015) that mature and old forest have repeatedly been shown to be strongly associated with habitat selection among NSOs, along with other factors: "... numerous studies have repeatedly demonstrated the importance of vegetative structures found most often in late-seral and old-growth forest to nest-site and foraging habitat selection." Thus, we agree that various measures of those features are justified for use in habitat modeling as conducted by the FWS (2011, 2012). We also agreed with Dunk et al. (2015) that the predictive relationships between owl demography and a variety of habitat measures has not been proved as strong or prevalent for all demographic parameters or among all physiographic provinces (e.g., Dugger and Davis, 2011:17). Dunk et al. argue that poor predictability of owl reproductive performance occurs because climatic factors override habitat factors. That may occur chiefly because coarse-scale habitat-type characterizations are insufficient in capturing habitat features that influence reproduction. It is important to note that other studies demonstrating close associations with habitat selection were based on estimating habitat features using aerial photos (e.g., Meyer et al., 1998). Thus, we were concerned with whether the satellite-based GNN database accurately and reliably incorporated those factors. In addition, Zabel et al. (2003) used pseudo-threshold and quadratic relations in demonstrating the virtues of nesting and roosting (NR) and foraging habitats (F), evidently also estimated via aerial photos, to predict owl locations. The FWS modeling did not include such transforms. In fact, NR and F did not always rank high in some model regions in the FWS models, possibly as a result of less-accurate GNN data in those regions. There is also reason to suggest that habitats used during the non-breeding season may differ from old-growth types around nest sites (e.g., Wiens et al., 2014) and thereby have unaccounted carry-over demographic effects. Finally, many of the habitat variables used in the MaxEnt models have not been clearly implicated in either nest site selection or demography. Thus the agreement that big trees and old forest are important does not mean that these factors alone are sufficient to predict either owl presence or demographic performance.

Dunk et al. next asserted that we mischaracterized the literature on MaxEnt. We noted some literature that criticizes MaxEnt but in the end identified the GNN data as the cause of model problems. MaxEnt is not a magic elixir, nor is any method. If the input data are compromised, or have varying levels of accuracy among model regions, no statistical tool can fix it.

We evaluated the Ackers et al. (2015) model for habitat modeling and found poor agreement between GNN-based, LiDAR-based, and aerial photo-based habitat areas. We used state of the art image analysis software to evaluate classification consistency,