Combining Inferences from Models of Capture Efficiency, Detectability, and Suitable Habitat to Classify Landscapes for Conservation of Threatened Bull Trout

JAMES T. PETERSON* AND JASON DUNHAM†

*U.S. Geological Survey, Georgia Cooperative Fish and Wildlife Research Unit, University of Georgia, Athens, GA 30602, U.S.A., email peterson@smokey.forestry.uga.edu
†U.S. Department of Agriculture Forest Service, Rocky Mountain Research Station, Boise, ID 83702, U.S.A.

Abstract: *Effective conservation efforts for at-risk species require knowledge of the locations of existing populations. Species presence can be estimated directly by conducting field-sampling surveys or alternatively by developing predictive models. Direct surveys can be expensive and inefficient, particularly for rare and difficult-to-sample species, and models of species presence may produce biased predictions. We present a Bayesian approach that combines sampling and model-based inferences for estimating species presence. The accuracy and cost-effectiveness of this approach were compared to those of sampling surveys and predictive models for estimating the presence of the threatened bull trout (*Salvelinus confluentus*) via simulation with existing models and empirical sampling data. Simulations indicated that a sampling-only approach would be the most effective and would result in the lowest presence and absence misclassification error rates for three thresholds of detection probability. When sampling effort was considered, however, the combined approach resulted in the lowest error rates per unit of sampling effort. Hence, lower probability-of-detection thresholds can be specified with the combined approach, resulting in lower misclassification error rates and improved cost-effectiveness.*

Combinación de Inferencias de Modelos de Eficiencia de Captura, Detectabilidad y Hábitat Adecuado para Clasificar Paisajes para la Conservación de *Salvelinus confluentus*

Resumen: *Los esfuerzos de conservación efectivos para especies bajo riesgo exigen conocer la ubicación de las poblaciones existentes. La presencia de una especie se puede estimar directamente por medio de muestreos a campo o mediante modelos predictivos. Los muestreos directos pueden ser costosos e ineficientes, especialmente para especies raras y difíciles de muestrear y los modelos que predicen la presencia de especies pueden generar predicciones sesgadas. Presentamos una aproximación Bayesiana que combina el muestreo y las inferencias resultantes de modelos para estimar la presencia de especies. La precisión y rentabilidad de esta aproximación para estimar la presencia de la especie amenazada *Salvelinus confluentus* *se comparó con la del muestreo y de los modelos predictivos por medio de simulación con modelos existentes y con datos de muestreo empírico. Las simulaciones indicaron que el muestreo únicamente podría ser la más efectiva y tendría las menores tasas de error de clasificación de presencia y ausencia para los tres umbrales de probabilidad de detección. Sin embargo, cuando se consideró el esfuerzo de muestreo, la aproximación combinada resultó tener las menores tasas de error por unidad de muestreo. Por lo tanto, se pueden especificar umbrales menores de probabilidad de detección con la aproximación combinada, lo que resulta en menores tasas de error de clasificación y mayor rentabilidad.*

Paper submitted November 26, 2001; revised manuscript accepted October 16, 2002.

Introduction

Species conservation begins with two key questions: (1) Where does the species occur? and (2) Where could it occur? Answers to these superficially simple questions can be complex. The answer to the first question must come from field surveys that document the presence of a species. Determination of species presence is absolutely certain only when a species is detected or captured (assuming species are identified correctly). If a species is not detected in a survey, there are two alternatives: the species was truly absent or it was present but not detected during the survey. Species of concern typically exhibit some form of rarity (e.g., Rabinowitz et al. 1986), and many have cryptic coloration or behaviors that make them difficult to detect in field surveys. Accordingly, it is often difficult to determine if records of species "absence" represent true absence or a lack of detection.

An alternative to direct surveys is to develop models that predict species presence (Scott et al. 2002). The physical and biological characteristics of habitats or localities can be used to address the question of where a species could occur. If the presence of a particular species is consistently associated with a specific suite of physical and biological conditions, then it may be reasonable to assume it will occur when those conditions are met. Thus, the presence of a species may be inferred without direct sampling in the field. This indirect approach is appealing because field surveys to determine species presence can be expensive and time-consuming. There is one major disadvantage to this approach, however: models of species presence typically assume that species are detected with relatively high efficiency. This assumption is rarely validated and is likely violated with much greater frequency than is generally appreciated (Thompson et al. 1998).

If, as seems likely, variability in the detectability of species is a function of survey conditions (e.g., environmental or habitat conditions, sampling methods), then observed patterns of species presence may not reflect actual patterns of presence. In fact, associations between patterns of species presence and habitat may be a function of both habitat preferences and variation in detectability (e.g., Bayley & Dowling 1993). Errors in habitat classification based on models with predictions biased by variation in detectability can compromise species conservation efforts. For example, some species may depend strongly on habitats where they are difficult to detect. In this case, a model based only on observed patterns of presence would not predict these habitats to be important. Consequently, key habitats may not receive adequate consideration in conservation planning.

Both approaches to determining species presence (direct surveys and prediction of presence with models) are subject to potentially significant shortcomings. Direct surveys alone may be expensive and inefficient, and models of species presence may produce biased predictions. Here, we show how information from both sampling (detectability) and habitat models can be combined in a third approach to produce more robust inferences about the probability of a species occurring in a given area (Fig. 1). To provide an example, we compare predictions from a landscape model of presence for threatened bull trout (*Salvelinus confluentus*; Dunham & Rieman 1999) to those from models of capture efficiency and detectability (Peterson et al. 2002) to show how each may be used to estimate the probability of presence. We then show, via simulation, how a combined approach using predictions from both models can be used to determine bull trout presence with greater efficiency and economy. In the combined approach, predictions from the landscape model serve as a prior probability of presence that is updated with new information from sampling surveys of bull trout.

Methods

Landscape Model of Bull Trout Presence

To provide a realistic basis for our simulations, we used predictions from existing models of presence for bull trout in the Boise River basin in southwestern Idaho (Rie-

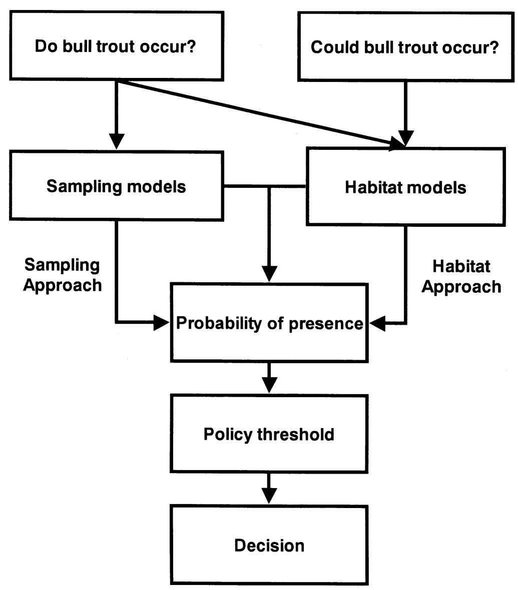

Figure 1. Three alternative approaches to determining probability of presence for bull trout. In the sampling-only approach, only information from sampling effort is used to estimate probability of presence. In the habitat approach, sampling is not conducted, and inferences about probability of presence are based on predictions from a habitat model alone. In the combined approach, predictions from the habitat model serve as a prior probability of presence that is sequentially updated with new information from sampling to produce predicted posterior probabilities of species presence.

man & McIntyre 1995; Dunham & Rieman 1999; Dunham et al. 2002). The distribution of suitable habitats in the basin was inferred from an elevation gradient that delineated the distribution of juvenile bull trout. "Patches" of suitable habitat in the basin were defined as watersheds above 1600 m elevation. The area of these individual patches, their isolation from other occupied patches, and road densities were significantly associated with presence of bull trout in a logistic regression model. We used predictions (predicted probabilities of presence) from this model as prior probabilities of presence in each patch.

For each simulation, we used 77 patches in the Boise River basin. In the original data set analyzed by Dunham and Rieman (1999), bull trout were known to be present in 29 patches and were believed to be absent in 52 (n = 81). We used only 77 of these 81 observations because information on fish densities was absent from 4 of the observations in the full data set. If bull trout were not observed in a patch (e.g., classified as "absent" by Dunham & Rieman [1999]), then we treated the patch as a true absence (zero fish density) in the simulations.

The logistic regression model described in Dunham and Rieman (1999) had a 19.5% overall cross-validation error rate for presence and absence of bull trout. To more faithfully represent out-of-sample performance of the landscape model, we estimated presence probabilities for each patch via leave-one-out cross-validation (Efron 1983). In other words, we sequentially omitted each observation (e.g., a single patch) from the model and predicted presence by using the remaining observations. For our simulations, we classified each patch as "occupied" or "unoccupied" based on three pairs of cutoff values: 70-30%, 80-20%, and 90-10%. Patches with predicted probabilities greater than or equal to the higher and less than or equal to the lower cutoff value were classified as occupied or unoccupied, respectively (e.g., >70% was as predicted present and <30%, absent). Patches with predicted probabilities between each pair of cutoff values (e.g., <70% and >30%) were classified as "uncertain" and were not considered in the estimation of classification and prediction error rates.

For comparison, we also classified each patch based on a 50% cutoff value for probability of presence, which is commonly done in ecological studies (e.g., >50% = "present"). We computed classification and prediction error rates for each pair of cutoff values. Classification error rates were defined as the proportion of patches that were assigned the incorrect status (e.g., an occupied patch that was classified as unoccupied). Prediction error rates were the proportion of status-specific predictions that were incorrect (e.g., the proportion of patches predicted to be occupied were actually unoccupied).

Sampling Detection Probabilities

The sampling-only approach was based on the bull trout sampling protocols of Peterson et al. (2002). This protocol estimated fish-capture efficiency for a given method (e.g., day snorkeling, night snorkeling, and electrofishing) under a known range of sampling conditions (e.g., effects of habitat size, cover) in individual sampling units (stream reaches). The numbers of fish vulnerable to capture were similarly estimated using a model parameterized with bull trout sampling data from known occupied patches across the range of bull trout in the region. Then estimates of capture efficiency and expected numbers of fish vulnerable to capture were used to estimate bull trout detection probabilities in a single sampling unit. We estimated detection probabilities for patches by combining detection probabilities across units sampled within each patch with the methods detailed in Peterson et al. (2002). These estimates served as the basis for terminating sampling in a patch during our simulations of the sampling only approach.

With this approach, we assumed sampling units were randomly selected within each patch and sampled (sequentially) with single-pass, backpack electrofishing. If bull trout were not detected in the first sampling unit, additional units were sampled in each patch until bull trout were detected or a desired cutoff value for the probability of detection in the patch was reached. For example, a survey that failed to detect fish in a patch was deemed satisfactory if the estimated probability of detection was 90%. Patch-specific probabilities of detection, d, were computed by combining individual probabilities of sampling unit detection (hereafter, threshold detection probabilities) as

$$d = 1 - \prod_{j=1}^{k} (1 - d_j), \qquad (1)$$

where d_j are the sampling unit-specific threshold detection probabilities and k is the maximum number of samples collected in a patch.

Simulated Sampling-Only Approach

We conducted 1000 simulation runs for each of three patch-specific detection probability cutoff values: 70%, 80%, and 90%. During the simulations, we attempted to emulate the typical sampling process in which sites are randomly selected and sequentially sampled until bull trout are detected or the cutoff threshold detection probability is reached. To more faithfully represent typical sampling conditions, we randomly generated the abundance of bull trout and capture efficiency in each sampling unit by using statistical distributions fit to empirical sampling data from the Boise River basin (B. E. Rieman, unpublished data). We randomly assigned the average density of bull trout to each known occupied patch through a gamma distribution with shape and scale parameters 0.22 and 1.95, respectively, and the known unoccupied patches were assigned zero density. We then randomly generated the abundance of fish in each sampling unit within a patch assuming a Poisson distribution with a mean equal to the randomly assigned patch-specific density. We also randomly assigned capture efficiency for each unit from a beta distribution with shape and scale parameters 3.54 and 9.74, respectively. This resulted in simulated fish abundance and capture efficiency that varied, on average, 20% from those used to estimate threshold detection probabilities.

We simulated fish sampling by estimating a probability of detection (henceforth, sampling detection probability) for each sampling unit with the randomly generated fish abundance and randomly assigned capture efficiency. If the sampling detection probability for the unit exceeded a randomly generated uniform number (i.e., range 0–1), we assumed detection and ceased simulated sampling in the patch. Otherwise, we generated additional sampling units for each patch until the total threshold detection probability exceeded the pre-specified cutoff. For each simulation, we computed the error rate for those patches that were predicted to be unoccupied and were actually known to be occupied and estimated the average number of simulated samples per patch.

Simulated Combined Approach

The combined approach builds on the Bayesian approach for estimating species presence detailed by Bayley and Peterson (2001). This approach used both a prior probability of presence and a probability of detection to estimate a posterior probability of species presence, given that it was not detected during sampling. Thus, it effectively combines the landscape model predictions (the empirical prior) and sampling-only approach (detection probabilities) to produce a posterior probability of patch occupancy.

The posterior probability of patch occupancy, given a species was not detected, $P(F|\mathrm{Co})$, is estimated as

$$P(F|Co) = \frac{P(Co|F) \cdot P(F)}{P(Co|F) \cdot P(F) + P(Co|{\sim}F) \cdot P({\sim}F)}, \quad (2)$$

where $P(F)$ is the prior probability of species presence, $P({\sim}F)$ is the prior probability of species absence, and $P(\mathrm{Co}|F)$ is the probability of not detecting a species when it occurs (Bayley and Peterson 2001).

Logically, the probability of species absence is the complement of presence, $P({\sim}F) = 1 - P(F)$, and the probability of not detecting a species when it is present is the complement of detection, $P(\mathrm{Co}|F) = 1 - d$, where d is the threshold probability of detection, and the probability of not detecting a species when it is absent is $P(\mathrm{Co}|{\sim}F) = 1$.

Equation 2 indicates that numerous combinations of prior probabilities of presence and detection can result in the same posterior probabilities (Fig. 2). Thus, by rearranging Eq. 2 we can estimate the probability of detection, required to achieve various posterior probabilities as

$$d' = 1 - \frac{P(F|Co) \cdot P({\sim}F)}{P({\sim}F|Co) \cdot P(F)}, \quad (3)$$

where $P({\sim}F|\mathrm{Co})$ is the complement of the desired posterior probability ($1 - P[F|\mathrm{Co}]$) and the remaining variables are as defined above. This provides a means to adjust the detection probability (and required sample size) for each patch, based on the value of the empirical prior derived from the landscape model and the desired cutoff value of the posterior probability (henceforth, posterior probability cutoff value) for estimating species presence or absence.

We conducted 1000 simulation runs for each of three pairs of posterior probability cutoff values identical to those used for landscape model evaluation. Prior to each sampling simulation, patches were classified as occupied, unoccupied, or uncertain based on the cutoff values and rule set described above. Required probabilities of detection, d', were then estimated for the uncertain patches with Eq. 3, the empirical landscape prior, and the (lower) cutoff values for defining absence (i.e., $P[F|\mathrm{Co}]$ equal to 30%, 20%, and 10%). These patches were then sampled (as outlined above) until a bull trout was detected or d' was exceeded by the total threshold detection probability. We then computed the prediction error rates, number of patches sampled, and average number of simulated samples collected per patch.

To put the sample size requirements estimated during the simulations in perspective, we contacted four research biologists responsible for planning the field sampling of bull trout and other salmonids in the northwestern United States. We asked each biologist to determine the size of a crew needed to sample bull trout with the bull trout protocol (Peterson et al. 2002) and the number of samples that could be collected per day. We then averaged these numbers and used the averages to estimate the number of sampling days required for each approach and probability cutoff value.

Results

In terms of prediction error rates from the alternative models (Table 1), landscape models were relatively accurate for the least-restrictive 50% probability cutoff value, and the accuracy of predictions increased with the more restrictive cutoff values. For example, for the landscape model at the 50% cutoff, prediction error rates for bull trout presences and absences were 0.250 and 0.163, respectively (Table 1). The corresponding prediction error rates for the more restrictive 70–30% cutoff were 0.074 and 0.038 lower for presences and absences, respectively, and more than 0.10 lower at the

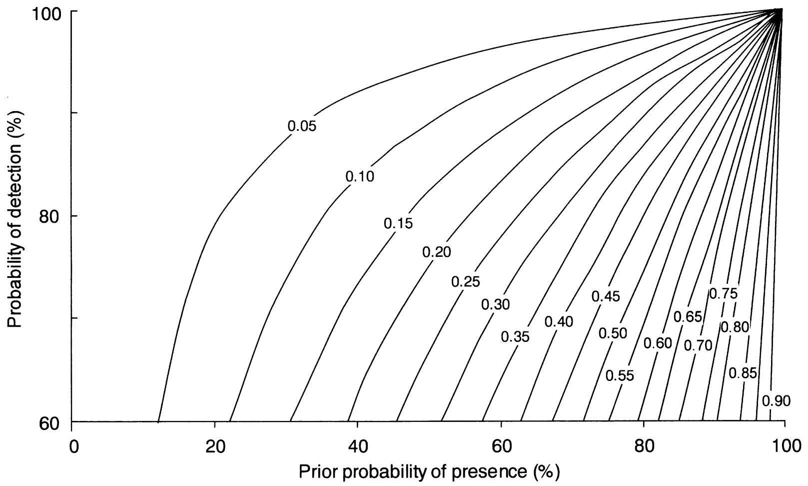

Figure 2. Response surface of predicted posterior probabilities of species presence, given no detection, for various combinations of prior probabilities of presence and detection for bull trout. Lines represent posterior probabilities (numbers) for specific prior probability and detection combinations. Power of detection is an index of the amount of sampling effort required to obtain a given posterior probability of presence, given no detection. For example, if a posterior probability of presence of 0.05 is desired, lower power (effort) is required when prior probabilities of presence are lower. Thus, a given amount of sampling effort is more informative when prior probabilities of presence are lower.

80–20% and 90–10% cutoff values. However, the proportion of patches classified as uncertain increased from 0.260 at the 70–30% probability cutoff to 0.636 at the 90–10% cutoff (Table 1).

For all approaches, we found prediction error rates were lowest for the most restrictive cutoff values 90–10% (Table 1). Among approaches, error rates were lowest for the sampling-only approach for all probability cutoff levels (Table 1). On average, error rates for the sampling-only approach were 0.071, 0.037, and 0.040 lower than the landscape model and combined approach at the 70–30%, 80–20%, and 90–10% cutoff values, respectively (Table 1). Much of these differences, however, were due to the zero error for predicting presence (i.e., false presence) with the sampling-only approach. When these values were not considered, error rates for the sampling-only approach were only 0.037, 0.022, and 0.021 lower than the landscape model and combined approaches at the three pairs of cutoff values, respectively. Expected error rates of the combined approach were the next best among the approaches considered and were, on average, 0.020 lower than the landscape-model across probability cutoff values. Additionally, no patches were classified as uncertain with the sampling effort and combined approach.

Although the simulations indicated the greatest accuracy for the sampling-only approach, it also required greater numbers of patches sampled and average number of samples per patch compared with the combined approach (Table 2). On average, the number of patches sampled and the number of samples per patch for the combined approach were 55.4% and 33.2%, respectively, lower than the sampling-only approach. Based on the judgment of four research biologists, we estimated that a four-person crew could collect an average of 2.77 samples per day (range 2.5–4). Thus, we estimate that

Table 1. Comparison of the proportion of uncertain patches ($n = 77$) and the mean prediction and classification errors for the three approaches to predicting bull trout patch occupancy at three probability cut-off levels.

Probability cut-off levels and approach	*Uncertain patches*	*Prediction error[a,b]*		*Classification error[b]*	
		present	*absent*	*present*	*absent*
70%, 30%					
landscape model	0.260	0.176 (0.250)	0.125 (0.163)	0.263 (0.276)	0.079 (0.146)
sampling only	0.000	0.000	0.114	0.217	0.000
combined	0.000	0.128	0.138	0.255	0.063
80%, 20%					
landscape model	0.442	0.111	0.059	0.200	0.030
sampling only	0.000	0.000	0.082	0.150	0.000
combined	0.000	0.044	0.107	0.187	0.021
90%, 10%					
landscape model	0.636	0.125	0.050	0.125	0.050
sampling only	0.000	0.000	0.048	0.084	0.000
combined	0.000	0.039	0.063	0.110	0.021

[a] *Mean prediction errors for sampling effort and combined approaches based on 1000 simulations.*
[b] *Category-wise errors for landscape model with 50% probability cut-off level in parentheses.*

Table 2. Comparison of the level of bull trout sampling effort and number of days required for a four-person crew to complete the sampling-only approach and combined approach to predicting bull trout patch occupancy at three probability cut-off levels.[a]

Probability cutoff and approach	*Patches sampled*	*Samples per patch*	*Sampling days per patch*[b]	*Sampling days for basin-wide survey*[c]
70%, 30%				
sampling only	77	12.1	4.4	337.5
combined	20	8.2	3.0	59.2
80%, 20%				
sampling only	77	15.6	5.6	433.9
combined	34	11.2	4.1	137.8
90%, 10%				
sampling only	77	21.6	7.8	601.2
combined	49	13.2	4.8	233.2

[a] *Values are means of 1000 simulations.*
[b] *Based on average 2.77 samples per days estimated by field biologists.*
[c] *Basin-wide estimates include only the 77 patches considered.*

the sampling-only approach would take an average of 70.6% more days for a single crew, across probability cutoff values, to complete a basin survey (assuming 77 patches/basin) compared with the combined approach

Discussion

Several factors must be considered when developing a sampling or monitoring strategy, including the effectiveness and cost of data collection and the interpretability of the data. Of these, our simulations suggest that a sampling-only approach would be the most effective in terms of minimizing error rates for a given probability cutoff. However, the combined approach results in the lowest error rates per unit of effort when sampling effort is considered. For example, we estimated that the average number of sampling days required for a single crew to complete a basinwide survey with the sampling-only approach and a 70–30% cutoff was 104 days greater than the combined approach using a 90–10% cutoff. The false absent error rate for the combined approach with a 90–10% cutoff also was 0.051 lower than for the sampling-only approach with a 70–30% cutoff. Hence, the combined approach with a 90–10% cutoff required fewer samples and resulted in a lower false-absent error rate compared with the sampling-only approach with a 70–30% cutoff. In practice, the lower cost of the combined approach may be appealing when resources for field surveys are limited.

Another potential difficulty with the sampling-only approach is how to interpret an event when a species is not detected in a sampling frame (e.g., a patch in this example). Previous approaches have used the complement of the detection probability (i.e., 1 − detection) as an estimate of the probability that a sampling frame (e.g., patch) was occupied, generally in reference to some pre-specified threshold density (Green and Young 1993; Watson & Hillman 1997). Because all detection probabilities depend on the sampling frame being occupied or occupied at some threshold density, the interpretation of an event when a species is not detected requires the consideration of total probability (Bayley & Peterson 2001). To illustrate, assume that 100 patches known to contain bull trout are sampled with an 80% threshold detection probability. On average, bull trout would not be detected in 20 of the patches, yet all of the patches (100%) contained bull trout. This 100% would be the posterior probability of bull trout presence given no detection. In contrast, assume a situation in which 1 of the 100 patches did not contain bull trout. Consequently, on average, bull trout would not be detected in the known absent patch and in 20 others (20% of 99) that actually contained bull trout. Thus, 20 of the 21 patches in which bull trout were not detected actually contained bull trout and the posterior probability of presence (given no detection) would be 95%. In both examples, the posterior probability of presence depended on the prior knowledge of the actual number of patches containing bull trout. Similarly, proper interpretation of an event when a species was not detected required a detection probability and a prior estimate of the probability of species presence. We used a landscape model to provide a prior estimate of presence for the combined approach to estimating occurrence. This prior estimate could be assumed to be unknown (i.e., an uninformative prior; Gelman et al. 1995) or could be based on expert opinion (Henrion et al. 1991). However, a more sound and defensible approach would be to develop priors based on empirical models (Bayley and Peterson 2001), as we did here.

As with all statistical estimation methods, the accuracy of the combined approach is significantly influenced by the accuracy of the prior probability of presence and the species detection probability estimates. The prior probability of presence has the greatest influence on the posterior probability of presence at high values (Fig. 2). In other words, if a habitat model predicts a high prior probability of occurrence for a species, more sampling effort will be required to attain a "low" (e.g., 10–30%) posterior probability of occurrence, given no detection. Hence, empirical models that overestimate the probability of presence would substantially affect the accuracy of the combined approach. However, because most models of presence are based on incomplete determinations of true presence it is more likely by far that "underprediction" is the case for "habitat" or "landscape" models.

As the value of the prior probability of presence decreases, detection probability estimates have a greater influence on the posterior probability of presence (Fig. 2). If habitat models predict a low prior probability of occurrence, then less sampling effort is required to attain a low posterior probability of occurrence, given no detection. Underestimates of detection probabilities would have the least impact on the accuracy of the combined approach because a species would likely be detected when present, but it would likely increase costs by requiring too many samples in locations where a species is absent. However as we have shown, incorporating prior information would minimize the increased costs associated with underestimated detection probabilities. In contrast, overestimates of detection probabilities would likely result in an insufficient amount of sampling effort and increased incidence of falsely concluding absence. To minimize such errors, we strongly recommend that researchers and managers examine the accuracy of all their models via an unbiased validation procedure, such as cross validation.

The accuracy of the bull trout presence and absence estimates is also significantly influenced by the choice of probability cutoff. For all three approaches considered, the highest error rates were associated with the least restrictive 70–30% cutoff. Although we arbitrarily selected three symmetric pairs of cutoff values to use during the simulations, most practical applications will have to select cutoff values based on economic and ecological considerations. For example, falsely predicting the presence of an endangered species may result in unnecessary land-use restrictions and an associated increase in management costs. Conversely, falsely predicting the absence of a species could lead to its extirpation through the failure to enact the proper protection strategies. Additionally, the impact of inaccurate prior probability of presence models and uncertainty regarding estimates of detection probabilities can be partially compensated by adjusting the cutoff probability. Thus, the choice of cutoff should be based on the accuracy of the prior probability of presence, the accuracy of the probability of detection estimates, the effect of land management actions on populations, and the costs associated with management actions. For example, in cases where the cost of falsely predicting species absence is a concern (e.g., unknowingly harming an endangered species), it may be prudent to use a more conservative lower cutoff (e.g., 10%) for assuming "absence," or low probability of presence, and a less conservative upper cutoff for "presence" (e.g., 70%) to ensure that all potentially occupied habitats are protected.

During our simulations, we did not consider the initial effort and cost associated with developing the landscape model necessary for estimating the prior probability of presence. Although this is certainly a substantial outlay, we believe that in many instances initial models can be developed using existing data for the target or ecologically similar species. Alternatively, empirical priors can be estimated by combining data for several ecologically similar species via Bayesian meta-analysis (Tufto et al. 2000). These initial models can then be examined through sensitivity analysis to identify the major sources of uncertainty and develop monitoring designs that can collect the most useful data for improving the models. More efficient approaches also may be possible through the use of adaptive sampling techniques (Thompson & Seber 1996) and Bayesian statistical methods for model updating (Spiegelhalter et al. 1993). Thus, existing survey and monitoring efforts can be designed to simultaneously determine the distribution and status of target species and improve the prior models.

Management Implications

Bull trout is an excellent example of the management dilemmas posed by uncertainty regarding patterns of species occurrence. Habitat requirements for bull trout are specific and sensitive to land-use impacts (e.g., Rieman & McIntyre 1993). Therefore, efforts to restore and protect habitat for bull trout could have substantial effects on land management throughout the broad range of this species (e.g., Rieman et al. 1997). While many details of the habitat requirements of bull trout are unknown, it is clear there are many potentially suitable habitats where bull trout are unlikely to occur (Dunham & Rieman 1999). In developing a management strategy, an obvious first step is to identify habitats where occurrence of bull trout is known or believed to be likely. As illustrated above, this is not a simple matter, and expert opinion on habitat relationships for bull trout is highly variable (Rieman et al. 2001).

For several reasons, we believe the combined approach to determining occurrence is most likely to provide useful information for bull trout conservation efforts. Because bull trout are difficult to sample, it is unlikely that a sampling-only approach will suffice for all cases. Another drawback of a sampling-only approach is that it does not provide useful information on habitat requirements for bull trout. Finally, without good information on the distribution of habitats on the landscape (e.g., Dunham et al. 2002), it will be difficult to delineate where sampling should occur. A habitat-only approach would be less costly to develop, but uncertainty about actual patterns of occurrence would likely lead to more conservative habitat protection. The costs of restrictions on land use resulting from these protections could be considerable and controversial. The combined approach to determination of bull trout occurrence requires greater initial investment, but it provides managers with much greater flexibility and information. As such, the longer-term benefits of the combined approach should far outweigh the initial costs.

Acknowledgments

We thank the many individuals who assisted during the field portion of this research. We are especially grateful to B. E. Rieman for use of his data on bull trout. We also thank J. Guzevich, D. Horan, D. Myers, and D. Weigel for their assistance with the estimates of sampling effort. The manuscript was improved with suggestions from M. Conroy, C. Moore, B. Rieman, D. Stauffer, and an anonymous reviewer.

Literature Cited

Bayley, P. B., and D. C. Dowling. 1993. The effects of habitat in biasing fish abundance and species richness estimates when using various sampling methods in streams. Polskie Archiwum Hydrobiologii **40:** 5–14.

Bayley, P. B., and J. T. Peterson. 2001. Species presence for zero observations: an approach and an application to estimate probability of presence of fish species and species richness. Transactions of the American Fisheries Society **130:**620–633.

Dunham, J. B., and B. E. Rieman. 1999. Metapopulation structure of bull trout: influences of habitat size, isolation, and human disturbance. Ecological Applications **9:**642–655.

Dunham, J. B., B. E. Rieman, and J. T. Peterson. 2002. Patch-based models of species presence: lessons from salmonid fishes in streams. Pages 327–334 in J. M. Scott, P. J. Heglund, F. Samson, J. Haufler, M. Morrison, M. Raphael, and B. Wall, editors. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California.

Efron, B. 1983. Estimating the error rate of a prediction rule: improvement on cross-validation. Journal of the American Statistical Association **78:**316–331.

Gelman, A. B., J. S. Carlin, H. S. Stern, and D. B. Rubin. 1995. Bayesian data analysis. Chapman and Hall, Boca Raton, Florida.

Green, R. H., and R. C. Young. 1993. Sampling to detect rare species. Ecological Applications **3:**351–356.

Henrion, M., J. S. Breese, and E. J. Horvitz. 1991. Decision analysis and expert systems. Artificial Intelligence **4:**64–91.

Peterson, J., J. Dunham, P. Howell, R. Thurow, and S. Bonar. 2002. Protocol for detecting bull trout presence. Western Division of the American Fisheries Society. Available from http://www.fisheries.org/wd/committee/bull_trout/bull_trout_committee.htm (accessed 26 August 2002).

Rabinowitz, D., S. Cairns, and T. Dillon. 1986. Seven forms of rarity and their frequency in the flora of the British Isles. Pages 182–204 in M. E. Soulé, editor. Conservation biology: the science of scarcity and diversity. Sinauer Associates, Sunderland, Massachusetts.

Rieman, B. E., and J. D. McIntyre. 1993. Demographic and habitat requirements for conservation of bull trout. U.S. Forest Service, Pacific Northwest Research Station General Technical Report INT-302, Fort Collins, Colorado.

Rieman, B. E., and J. D. McIntyre. 1995. Presence of bull trout in naturally fragmented habitat patches of varied size. Transactions of the American Fisheries Society **124:**285–296.

Rieman, B. E., D. C. Lee, and R. F. Thurow. 1997. Distribution, status, and likely future trends of bull trout within the Columbia River and Klamath basins. North American Journal of Fisheries Management **17:**1111–1125.

Rieman, B. E., et al. 2001. Evaluation of potential effects of federal land management alternatives on trends of salmonids and their habitats in the interior Columbia River basin. Forest Ecology and Management **53:**43–62.

Scott, J. M., P. J. Heglund, F. Samson, J. Haufler, M. Morrison, M. Raphael, and B. Wall, editors. 2002. Predicting species occurrences: issues of accuracy and scale. Island Press, Covelo, California.

Spiegelhalter, D. J., A. P. Dawid, S. L. Lauritzen, and R. G. Cowell. 1993. Bayesian analysis in expert systems. Statistical Science **8:** 219–283.

Thompson, S. K., and G. A. F. Seber. 1996. Adaptive sampling. Wiley, New York.

Thompson, W. L., G. C. White, and C. Gowan. 1998. Monitoring vertebrate populations. Academic Press, New York.

Tufto, J., B. E. Saether, S. Engen, P. Arcese, K. Jerstad, O. W. Roestad, and J. N. M. Smith. 2000. Bayesian meta-analysis of demographic parameters in three small, temperate passerines. Oikos **88:**273–281.

Watson, G., and T. W. Hillman. 1997. Factors affecting the distribution and abundance of bull trout: an investigation at hierarchical scales. North American Journal of Fisheries Management **17:**237–246.