Modeling relationships between landscape-level attributes and snorkel counts of chinook salmon and steelhead parr in Idaho

William L. Thompson and Danny C. Lee

Abstract: Knowledge of environmental factors impacting anadromous salmonids in their freshwater habitats, particularly at large spatial scales, may be important for restoring them to previously recorded levels in the northwestern United States. Consequently, we used existing data sets and an information-theoretic approach to model landscape-level attributes and snorkel count categories of spring–summer chinook salmon (*Oncorhynchus tshawytscha*) and steelhead (*Oncorhynchus mykiss*) parr within index areas in Idaho. Count categories of chinook salmon parr were negatively related to geometric mean road density and positively related to mean annual precipitation, whereas those for steelhead parr were negatively related to percent unconsolidated lithology. Our models predicted that chinook salmon parr would be in low count categories within subwatersheds with >1 $km \cdot km^{-2}$ geometric mean road densities and (or) <700 mm mean annual precipitation. Similarly, steelhead parr were predicted to be in low count categories in subwatersheds with >30% unconsolidated lithology. These results provide a starting point for fish biologists and managers attempting to map approximate status and quality of rearing habitats for chinook salmon and steelhead at large spatial scales.

Résumé : La connaissance des facteurs environnementaux influant sur les salmonidés anadromes dans leurs habitats dulcicoles, particulièrement aux grandes échelles spatiales, peut être importante pour le rétablissement des populations aux niveaux observés dans le passé dans le nord-ouest des États-Unis. Ainsi, nous avons utilisé des ensembles de données déjà existants et une approche basée sur la théorie de l'information pour relier des attributs du paysage avec l'abondance des tacons de saumon quinnat (*Oncorhynchus tshawytscha*) et de saumon arc-en-ciel (*Oncorhynchus mykiss*), dénombrés dans l'eau par plongée au tuba, dans des secteurs témoins de l'Idaho. Les catégories d'abondance des tacons de saumon quinnat étaient corrélées négativement avec la moyenne géométrique de la densité des routes et corrélées positivement avec les précipitations anuelles moyennes, tandis que celles des tacons de saumon arc-en-ciel étaient corrélées négativement avec le pourcentage de matière non consolidée. Nos modèles prévoient que les tacons de saumon quinnat seraient peu abondants dans les bassins secondaires où les moyennes géométriques de la densité des routes sont >1 $km \cdot km^{-2}$ et (ou) les précipitations annuelles moyennes sont <700 mm. De même, ils prévoient que les tacons de saumon arc-en-ciel seraient peu abondants dans les bassins secondaires où le pourcentage de matière non consolidée est >30%. Ces résultats peuvent servir de point de départ aux biologistes et aux gestionnaires responsables de la faune ichthyenne qui veulent établir des cartes représentant l'état et la qualité approximatifs des habitats où se développent le saumon quinnat et le saumon arc-en-ciel couvrant de grandes régions.

[Traduit par la Rédaction]

Introduction

Numbers of anadromous salmonids have greatly decreased from previously recorded levels for many stocks in the northwestern United States (Nehlsen et al. 1991). For instance, numbers of salmon and steelhead in the Columbia Basin have decreased sharply from an estimated 10–16 million adults to about 1.5–4.0 million adults during this century (Northwest Power Planning Council 1986). One factor thought to be influencing these declines is loss or degradation of freshwater spawning and rearing habitats (Nehlsen et al. 1991). Unfortunately, empirical data supporting this assertion at the landscape or basinwide scale are lacking in the published literature because fishery research has traditionally been conducted on smaller spatial scales (Schlosser 1991; but see Dunham and Rieman 1999; Torgersen et al. 1999). Thus, there is a need for empirically based models to investigate relationships between large-scale habitat and land management attributes and numbers of anadromous salmonids in their rearing environments (e.g., Bradford et al. 1997). These models could be used to predict status and quality of salmon spawning and rearing habitats across an area of interest as well as serve to highlight possible factors affecting population status and trends.

Availability of broadscale habitat and land management data generated by the recent interior Columbia Basin assessment (Quigley and Arbelbide 1997) and a 10-year data set of

Received May 20, 1999. Accepted May 31, 2000.
J15149

W.L. Thompson[1] and D.C. Lee.[2] USDA Forest Service, Rocky Mountain Research Station, 316 East Myrtle Street, Boise, ID 83702, U.S.A.

[1]Author to whom all correspondence should be addressed. e-mail: bthompson02@fs.fed.us

[2]Present address: USDA Forest Service, Sierra Nevada Conservation Framework, 801 I Street, Sacramento, CA 95814, U.S.A.

spring–summer chinook salmon (*Oncorhynchus tshawytscha*) and steelhead (*Oncorhynchus mykiss*) parr counts obtained from streams across Idaho (Hall-Griswold and Petrosky 1996) provided an opportunity to build large-scale predictive models based on empirical data. Consequently, we applied the latest information-theoretic modeling techniques (Buckland et al. 1997; Burnham and Anderson 1998) to investigate possible relationships between broadscale habitat and land management attributes and snorkel counts of spring–summer chinook salmon and steelhead parr within index streams in the Snake River drainage in Idaho. This geographical area is of particular importance because the indigenous stocks of spring–summer chinook salmon and steelhead have been listed as threatened under the Endangered Species Act (Federal Register 1997, 1998*a*). Due to various shortcomings with the snorkel count data, emphasis of this paper is as much on the approach to extract information from this broadscale but problematic data set as it is on interpretation of model results. We emphasize that the information-theoretic approach to model building, model selection, and model averaging applied in this paper is relevant to any study requiring a statistically based modeling approach.

Materials and methods

Snorkel count data set

The Idaho Department of Fish and Game (IDFG) and several cooperating agencies conducted snorkel counts of juvenile chinook salmon and steelhead (i.e., parr) in the Salmon River, Clearwater River, and lower Snake River drainages in Idaho during 1986–1995 (Hall-Griswold and Petrosky 1996) (Fig. 1). Abundance indices were obtained via snorkel counts by divers swimming approximately 100 m upstream within stream sections. One to five divers were used depending on stream size (Petrosky and Holubetz 1986). Stream sections were chosen based on a variety of criteria such as access, existence of previous counts, and perceived quality of rearing habitat (J. Hall-Griswold, IDFG, Stanley, Idaho, personal communication). Thus, selection of stream sections was nonrandom, but these sections represented a spectrum of habitats, stocks, and production types (i.e., wild (native) and natural (having a previous hatchery influence); Rich and Petrosky 1994). Although an attempt was made to survey the same sections over time, location and size (length and width) of snorkeled sections often varied among years mainly due to loss of previous section boundary markers, difficulties in relocating inadequately described sections, loss of access, and annual differences in stream flows. Further, not all sections were surveyed every year because of personnel, funding, and logistical constraints (J. Hall-Griswold, IDFG, Stanley, Idaho, personal communication). Finally, some stream sections were stocked with hatchery fish to better evaluate population responses of parr to mitigation measures (Petrosky and Holubetz 1986).

Subsetting the snorkel count data set

We only analyzed counts from stream sections where mitigation measures and stocking were not applied (see Rich and Petrosky 1994, their appendix B) because of confounding effects of those factors on the relationship between landscape-level variables and fish abundance. In addition, we limited our analyses to counts conducted when the water temperature exceeded 9°C because of the low detectability of fish below this temperature (Thurow 1994), which also would have had a confounding effect on the relationship between landscape attributes and snorkel counts.

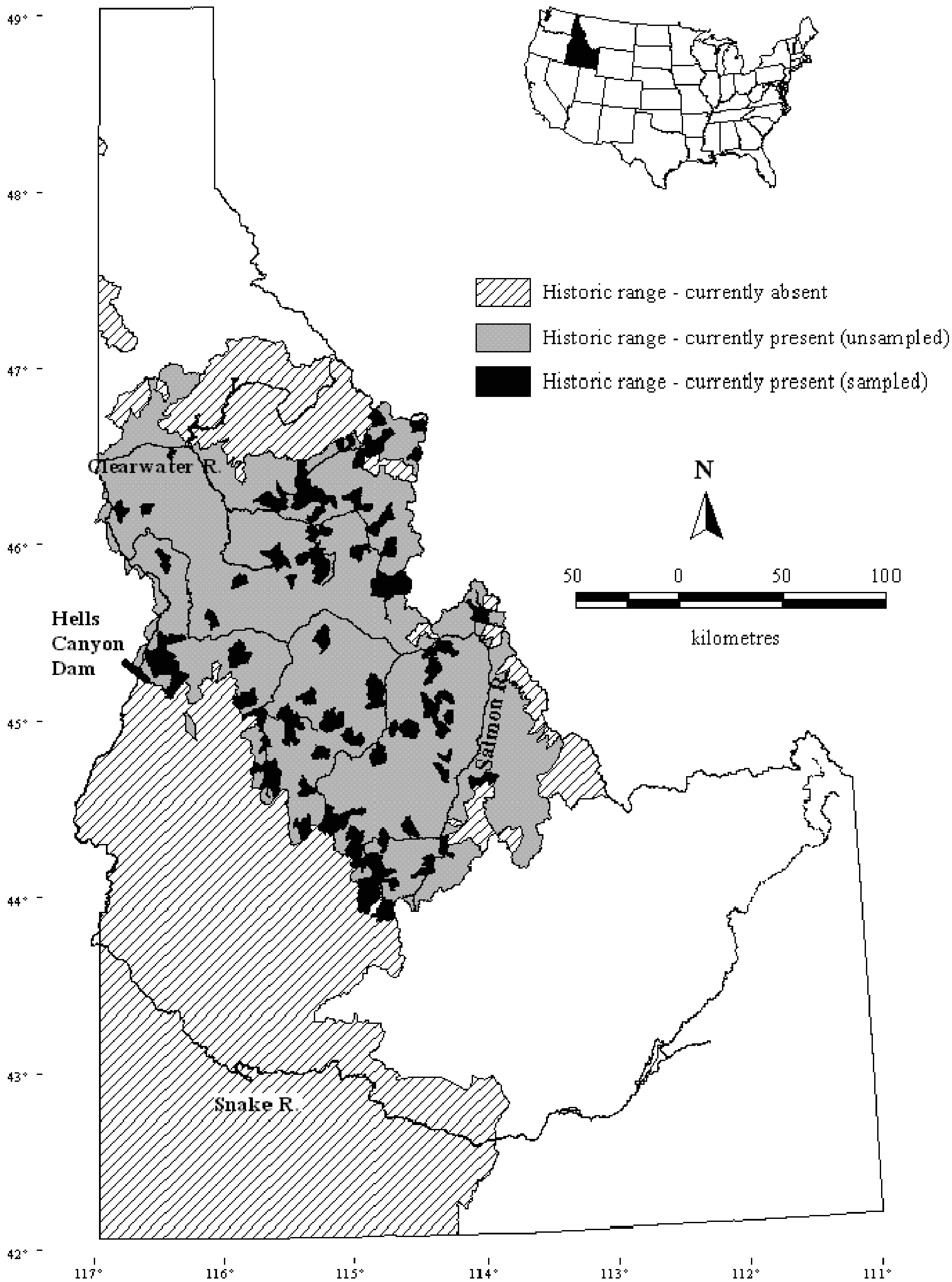
Because snorkel counts were uncorrected for incomplete detectability of fish within sections and therefore contained an unknown amount of bias (e.g., Rodgers et al. (1992) reported that only 40% of fish were detected during their snorkel counts), we pooled them into two categories in an attempt to alleviate detrimental effects of this bias on interpretation of model results. Categories were defined based on fish density indices and habitat ratings used by IDFG to categorize quality of rearing habitat; these values were 0.12 parr·m^{-2} for chinook salmon and 0.06 parr·m^{-2} for steelhead (Hall-Griswold and Petrosky 1996). Counts were divided by estimated area of each snorkeled stream section to provide a common unit of comparison with the IDFG ratings. Snorkel counts per unit area at or below 0.12 parr·m^{-2} for chinook salmon or 0.06 parr·m^{-2} for steelhead were placed into category 1, whereas higher counts per unit area were placed into category 2. This approach may lessen effects of bias, for instance, when two stream sections have the same actual densities of chinook salmon parr (e.g., 0.06 parr·m^{-2}) but different detection rates of individuals (e.g., 40 and 80%). In this case, if observed counts were used, one section would be improperly modeled with twice the observed count than the other section, which could lead to spurious model results. Conversely, results of both counts would be placed into the same count category under the categorization approach described above. Note, however, that categorization still could lead to misclassification of stream sections (and spurious results), depending on the detectability of fish within a given section and how close the count per unit area was to the cutoff value used in the categorization.

We pooled seasonal runs for both species because sample sizes were inadequate to model these data separately. Although peaks in average counts per unit area differed slightly between runs for both species in two years during 1986–1990, their 90% confidence intervals broadly overlapped. We also included both wild and natural populations in our analyses. Further, we concentrated on counts from C channels for chinook salmon and B channels for steelhead because these were their optimal habitats (Hall-Griswold and Petrosky 1996) and therefore should have supported higher densities of parr. C channels occurred in low-gradient (<2% slope) terrain, whereas B channels were those in moderate-gradient (2–4% slope) terrain (Rosgen 1985, 1996). Finally, data from the year with the highest average counts per unit area for each species were used in our investigation of landscape linkages. We did this to maximize our ability to detect a difference between better and lower quality sites, where quality was defined in terms of fish counts per unit area during years of high fish numbers. Preliminary analyses suggested that average counts per unit area of fish were similar between better quality sites and lower quality sites during years of low counts, whereas these sites were much more distinct during years of high counts (J. Peterson, Rocky Mountain Research Station, Boise, Idaho, unpublished data). Difficulties associated with site identification, nonrandom site selection, incomplete time series of surveys, and counting bias precluded use of typical methods for modeling time series data.

Landscape habitat and anthropogenic data

Landscape-level data were compiled by Lee et al. (1997) and defined at the subwatershed scale, which is about 7800 ha on average within the Columbia Basin. These variables were categorized as either physiographic and geophysical or as anthropogenic (Table 1). One of the anthropogenic variables, management cluster, was categorical (i.e., each subwatershed was assigned the predominant category) and was generated by Lee et al. (1997) from results of a cluster analysis of variables representing land-type classification, management classification, ownership, percent grazed, and percent wilderness (for details, see Lee at al. 1997). We further pooled these results into four broad categories for simplicity (Ta-

Fig. 1. Mapping of historical and current range of chinook salmon and steelhead in Idaho, U.S.A. (Lee et al. 1997). Dark areas are subwatersheds containing stream sections that were sampled for chinook salmon parr during 1987 and steelhead parr during 1990.



ble 1). This variable is an index to potential effects of land use and land management practices on adjacent streams and stream fish populations.

Modeling approach

We employed the information-theoretic approach to model building and selection suggested by Akaike (1973) and extended by Burnham and Anderson (1998). First, a global (i.e., overall; Burnham and Anderson 1998) logistic regression model was constructed with count category as the dichotomous response and landscape-level habitat and anthropogenic covariates that were deemed ecologically most relevant as predictors. Choice of predictors was based on results from Lee et al. (1997) and subject area experts familiar with the study area. We then assessed the fit of the global model via the Hosmer–Lemeshow goodness-of-fit (GOF) test and checked the Pearson χ^2 residuals for obvious outliers (i.e., >2; Hosmer and Lemeshow 1989). An outlier was dropped from

Table 1. Category and description of covariates used in modeling landscape (subwatershed scale) habitat and land management attributes (Lee et al. 1997) with count categories of chinook salmon and steelhead parr.

Category	Model covariate	Description
Physiographic and geophysical	Precip	Mean annual precipitation (mm) based on the PRISM model (Daly et al. 1994)
	Sumtemp	Mean annual maximum summer temperature (°C)
	Slope	% of subwatershed with slopes >50%
	Mafic	% of subwatershed with mafic lithology
	Unconsol	% of subwatershed with unconsolidated lithology
Anthropogenic	Georoad	Geometric mean road density (km·km^{-2})
	Mngclus	Management cluster variable containing four land use and ownership categories: (1) HIF (high impact forest): high impact, grazed USDA Forest Service forest (2) MF (managed forest): moderate to high impact, ungrazed USDA Forest Service forest (3) W (wilderness): USDA Forest Service wilderness (4) R (rangeland): USDI BLM rangelands and moderate impact, grazed USDA Forest Service rangeland

analysis if its inclusion caused serious model lack of fit (see below). The Hosmer–Lemeshow GOF statistic was generated by ordering observations by their event probabilities, grouping them into a 2 × g table (where g is number of groups; for the grouping procedure, see Hosmer and Lemeshow 1989), and calculating a Pearson χ^2 GOF statistic for this table. Low P values ($P < 0.10$) indicated model lack of fit. If the global model adequately fitted the data, we constructed a subset of candidate models from it that represented ecologically meaningful combinations of the landscape covariates. Each subsetted model was assumed to provide an adequate fit if the global model did so (Burnham and Anderson 1998).

Model selection was performed using a modification of Akaike's information criterion (AIC) (Akaike 1973; Burnham and Anderson 1998). An extension of likelihood theory, AIC is an estimate of the relative distance between model pairs (Burnham and Anderson 1998), where distance refers to the Kullback–Leibler distance of information theory (Kullback and Leibler 1951). The Kullback–Leibler distance is a measure of the degree of information loss when a model is used to approximate reality (Cover and Thomas 1991; Burnham and Anderson 1998). Specifically, AIC is defined as

$$\mathrm{AIC} = -2\ \ln(L(\hat{\theta}\,|\,\mathrm{data})) + 2k$$

where $\ln(L(\hat{\theta}\,|\,\mathrm{data}))$ is the maximized log-likelihood over the unknown model parameters (θ) given the data and k is the number of estimable parameters in the model (Buckland et al. 1997; Burnham and Anderson 1998). We used the small sample adjustment to the AIC that also corrects for overdispersion in count data, called QAICc. This statistic is calculated as

$$\mathrm{QAICc} = \frac{-2\ln(L(\hat{\theta}\,|\,\mathrm{data}))}{\hat{c}} + 2k + \frac{2k(k+1)}{n-k-1}$$

where $\hat{c}$ is the χ^2 GOF statistic for the global model and n is sample size (Burnham and Anderson 1998). Overdispersion refers to instances where sampling (observed) variance exceeds the theoretical variance of the underlying model (e.g., binomial model) and is commonly present in count data (Burnham and Anderson 1998). We used $\hat{c}$ to adjust for overdispersion in parameter estimates for each candidate model as well.

Models with lower QAICc values are considered better approximating models than those with higher values. However, QAICc is a relative statistic. The meaningful quantity for comparing candidate models is the difference between a particular model's QAICc value and the lowest QAICc value from all models; this difference is referred to as ΔQAICc (Burnham and Anderson 1998). The relative plausibility or weight of evidence of each model, given the data (w_i), can then be computed as

$$w_i = \frac{\mathrm{e}^{(-\Delta\mathrm{QAICc}_i/2)}}{\sum_{r=1}^{R}\mathrm{e}^{(-\Delta\mathrm{QAICc}_r/2)}}$$

where $\Delta\mathrm{QAICc}_i$ is the ΔQAICc value for the ith model in a set of R candidate models (Buckland et al. 1997). These w_i, or model weights, also can be used in model averaging. Instead of assuming a single "best" model and using its parameter estimates to make inferences, we based our inferences and predictions on a composite model generated from the w_i weighted average of parameter estimates for each landscape covariate from the set of candidate models (for details on model averaging, see Burnham and Anderson 1998). Model averaging incorporates both uncertainty related to model selection and uncertainty associated with parameter estimates within each candidate model. Inference based on a single model will lead to underestimates of variance and hence poor confidence interval coverage for parameter estimates unless its w_i is much higher (see below) than that of all other competing models (Burnham and Anderson 1998). Our composite models (one each for chinook salmon and steelhead data) only contained landscape covariates within candidate models whose w_i were at least one tenth of the maximum w_i, which is comparable with the minimum cutoff point (i.e., 8 or 1/8) suggested by Royall (1997) as a general rule-of-thumb for evaluating strength of evidence.

Interpreting model results

Data for landscape covariates were standardized so that their coefficients could be interpreted on a common scale. We also computed an odds ratio for each covariate by using its unstandardized coefficient, e.g., raising the coefficient to base "e" or $\mathrm{e}^{\hat{\beta}_1}$, to facilitate interpretation of the magnitude of its effect on parr densities. As given, these odds ratios are based on a single unit change, whereas larger (or smaller) units of change may be more ecologically interpretable. Therefore, we multiplied relevant unstandardized coefficients by a constant (C) whose magnitude reflected a more meaningful interpretation than a single unit change (e.g., $\mathrm{e}^{C\hat{\beta}_1}$; Hosmer and Lemeshow 1989). We obtained an initial estimate of the magnitude of the constant for each covariate based on the difference represented in two standard deviations from its mean as computed from the database compiled by the interior Columbia Basin assessment (Lee et al. 1997). Then, we consulted with subject area experts familiar with the study area to fine-tune these estimates. For example, the model coefficient for percentage of subwatershed containing >50% slopes (Slope) was multiplied by 10 because a 10% change in Slope from one subwatershed to another had more meaning, in terms of physical processes potentially affecting the streams and fish therein, than a single unit (1%) change in Slope. However, we also present unstandardized coeffi-

Table 2. Model selection results for logistic regression models containing landscape habitat predictor variables and count categories of chinook salmon parr sampled during 1987 (n = 37 subwatersheds (72 stream sections)).

Candidate model	QAICc	ΔQAICc	ΔQAICc weight	% of maximum ΔQAICc weight
Precip, Slope, Georoad	46.50	0	0.379	100
Precip, Mafic	48.48	1.98	0.141	37.2
Precip	48.77	2.27	0.122	32.2
Precip, Slope, Mafic	49.15	2.65	0.101	26.6
Precip, Slope, Unconsol, Georoad	49.16	2.66	0.101	26.6
Precip, Unconsol	50.07	3.57	0.064	16.9
Georoad	51.72	5.22	0.028	7.4
Precip, Slope, Unconsol	51.98	5.48	0.024	6.3
Sumtemp	53.50	7.00	0.011	2.9
Slope, Mafic	53.87	7.37	0.010	2.6
Sumtemp, Mafic	54.19	7.69	0.008	2.1
Global Model	54.94	8.44	0.006	1.6
Slope, Georoad, Mngclus	56.87	10.37	0.002	0.5
Mngclus	58.05	11.55	0.001	0.3
Slope, Mngclus	58.44	11.94	0.001	0.3
Georoad, Mngclus	58.60	12.10	0.001	0.3
Unconsol, Georoad, Mngclus	61.42	14.92	<0.001	<0.1

cients and their standard errors for those interested in interpreting odds ratios based on a single unit change.

We did not simply rely on statistical significance to interpret model results because an odds ratio could be small enough to be considered ecologically unimportant but still be statistically significant (Yoccoz 1991). Note that statistical significance can be construed if the confidence interval for an odds ratio does not include 1; this is equivalent to testing, say, $\hat{\beta}_1 = 0$, which can be respecified in terms of an odds ratio, $e^{\hat{\beta}_1} = e^0 = 1$. We evaluated ecological importance of each covariate in the composite model by computing 90% confidence intervals for the scaled odds ratios (e.g., $e^{C\hat{\beta} \pm 1.64 C \widehat{se}(\hat{\beta}1)}$, where $z_{0.95}$ = 1.64; Hosmer and Lemeshow 1989) and interpreting magnitudes of the values contained within these intervals (Gerard et al. 1998). A confidence interval that only contained values whose sizes were considered meaningful indicated an ecologically important relationship between the covariate and parr count categories. Conversely, an interval that only contained values whose magnitudes were considered of minimal importance indicated a covariate exhibiting a weak relationship with parr count categories. Finally, a confidence interval that contained values for odds ratios either on both sides of 1 or whose range included both ecologically important and unimportant magnitudes indicated inconclusive results due to imprecision from inadequate sample sizes.

We computed the predicted probability ($\hat{p}$) that a subwatershed had a low count category of parr (category 1) or a moderate to high count category of parr (category 2) using the formula $\hat{p} = \dfrac{1}{1 + e^{(-\beta_0 - \tilde{\beta}\tilde{X})}}$, where β_0 is the model intercept, $\tilde{B}$ is the vector of slope estimates, and $\tilde{X}$ is the vector of predictor variables (Hosmer and Lemeshow 1989). A Pearson correlation (r) was then calculated between predicted probability and predictor variable(s) with an ecologically important relationship with parr count category in both the chinook salmon and the steelhead composite models. If more than one predictor variable was ecologically important, we used the additional predictors as a basis for stratification for the correlation analysis. For instance, if composite model results indicated that both geometric mean road density and mean annual precipitation had ecologically important relationships with chinook salmon parr count categories, correlations were computed between predicted probability and mean annual precipitation for subwatersheds with both low and medium to high geometric mean road densities.

The SAS statistical package (SAS Institute Inc. 1996) was used for all of our analyses. Both the type I error rate (α) for GOF tests and the confidence coefficient for confidence intervals were set at 0.10 prior to analyses.

Results

We used data from 1987 for chinook salmon and from 1990 for steelhead because these years contained both the highest average parr counts per unit area and the narrowest confidence intervals of these estimates for each species. After removing one obvious outlier whose inclusion caused a serious model lack of fit, the global model for chinook salmon adequately fitted the data (Hosmer–Lemeshow GOF statistic = 6.06, 7 df, P = 0.53). The global model for steelhead also adequately fitted the data (Hosmer–Lemeshow GOF statistic = 9.81, 7 df, P = 0.20) and had no obvious outliers.

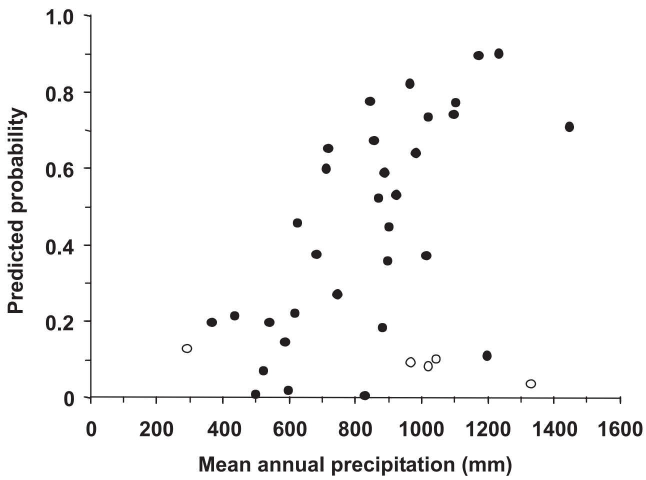
For the chinook salmon parr data, the candidate model containing mean annual precipitation, percentage of subwatershed containing >50% slopes, and geometric mean road density was nearly three times more plausible than the next best approximating model (Table 2). The composite habitat model for the chinook salmon data contained three covariates that were statistically significant, two of which had a fairly strong relationship with parr count categories (Table 3). Geometric mean road density exhibited a negative relationship with chinook salmon parr count categories in that moderate to high counts of parr were 1.33 (1/0.750) times less likely to occur in subwatersheds with every increase in 1 km·km^{-2} road densities. Conversely, moderate to high counts of chinook salmon parr were at least 1.29 times more likely to occur in subwatersheds with every 200-mm increase in mean annual precipitation. The lower bound of the odds ratio for percent slope >50% in a subwatershed was statistically significant but of trivial magnitude (Table 3).

Table 3. Model-averaged results of composite models for chinook salmon and steelhead.

Species	Model parameter	Estimated coefficient (SE)	Standardized coefficient	OR unit change	Estimated OR	90% CI for OR	
						Lower	Upper
Chinook salmon	Intercept	–2.336 (1.531)	—	—	—	—	—
	Precip	0.004 (0.002)	0.658	200	2.164	1.293	3.622
	Slope	–0.125 (0.073)	–0.598	10	0.286	0.086	0.948
	Mafic	–1.490 (1.114)	–0.512	10	<0.001	<0.001	28.992
	Unconsol	0.008 (0.024)	0.079	10	1.080	0.723	1.612
	Georoad	–1.023 (0.448)	–0.624	1	0.360	0.172	0.750
Steelhead	Intercept	–1.357 (1.728)	—	—	—	—	—
	Precip	0.002 (0.001)	0.300	200	1.350	0.922	1.975
	Sumtemp	0.274 (0.140)	0.259	2	1.728	1.092	2.735
	Slope	–0.018 (0.028)	–0.085	10	0.833	0.525	1.324
	Mafic	0.026 (0.009)	0.390	10	1.300	1.120	1.509
	Unconsol	–0.107 (0.044)	–0.645	10	0.342	0.167	0.701
	Georoad	0.191 (0.250)	0.131	1	1.210	0.803	1.824

Note: Model parameters whose 90% confidence intervals (CI) for their estimated odds ratios (OR) do not include 1 are statistically significant.

Fig. 2. Relationship between mean annual precipitation and geometric mean road density (open circles indicate low density and solid circles indicate medium to high density) and predicted probability ($\hat{p}$) of a subwatershed containing either low (e.g., $\hat{p} \leq 0.5$) or moderate to high (e.g., $\hat{p} > 0.5$) count categories of chinook salmon parr.



Evidence was inconclusive for the remaining covariates in the composite habitat model for chinook salmon.

Subwatersheds with low geometric mean road density (<1 km·km^{-2}; adapted from Lee et al. 1997) had a strong positive correlation (r = 0.643, n = 32) between mean annual precipitation and predicted probability of chinook salmon parr count category, whereas those with medium to high road density (≥1 km·km^{-2}) had a very strong negative correlation (r = –0.874, n = 5) between these variables. The five subwatersheds containing medium to high geometric mean road densities had predicted probabilities close to 0, which indicated that these subwatersheds were classified as containing low count categories of chinook salmon parr regardless of mean annual precipitation levels (Fig. 2).

For the steelhead parr data, the model containing mean annual precipitation and percent unconsolidated lithology was the best approximating model but was only slightly more plausible than the next highest ranked model (Table 4). The composite habitat model contained three covariates whose odds ratios were statistically significant (Table 3) but only one (percent unconsolidated lithology) had a fairly strong relationship with parr count categories. That is, moderate to high counts of steelhead parr were at least 1.43 (1/0.701) times less likely to occur in subwatersheds with every 10% increase in unconsolidated lithology than low densities. Thus, there was a negative relationship between steelhead parr counts and unconsolidated lithology.

Both mean annual maximum summer temperature and percent mafic lithology had a small positive relationship with moderate to high counts of steelhead parr. Moderate to high steelhead parr counts were at least 1.09 times more likely to occur in subwatersheds with every increase in 2°C mean annual maximum summer temperature and at least 1.12 times more likely to occur in subwatersheds with every increase in 10% mafic lithology (Table 3). Information on all other covariates in the composite habitat model was inconclusive.

Subwatersheds with <10% unconsolidated lithology had a weakly to moderately negative correlation (r = –0.343, n = 69) with predicted probability of steelhead parr count category, whereas those with >10% unconsolidated lithology exhibited an extremely strong negative correlation (r = –0.996, n = 10) with these predicted probabilities. Using the typical 0.5 cutoff for categorizing predicted probabilities (Hosmer and Lemeshow 1989), the five subwatersheds with >30% unconsolidated lithology would be classified as containing low count categories of steelhead parr (Fig. 3).

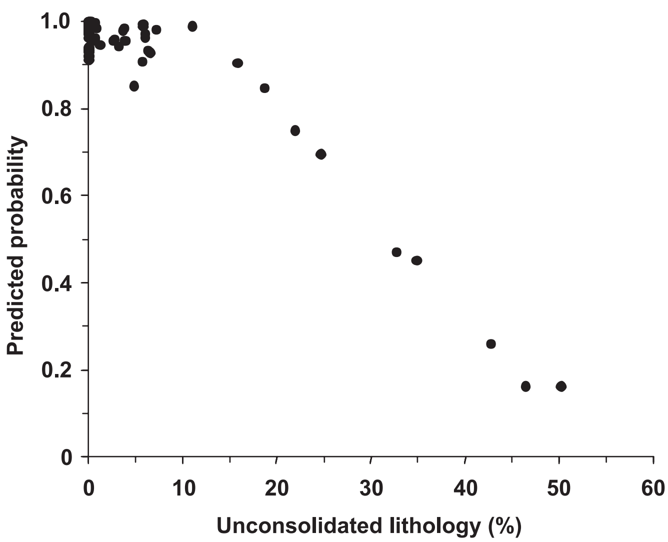
Discussion

Results generated from our analyses must be viewed within the limitations of the parr monitoring and landscape habitat data sets. First, our analyses suffered from the fact that our objective differed from the one originally set forth in the parr monitoring project, and therefore, we subsetted the data accordingly. Second, problems with study design, particularly the unknown impact of bias generated from the nonrandom selection of stream sections and counts uncor-

Table 4. Model selection results for logistic regression models containing landscape habitat predictor variables and count categories of steelhead parr sampled during 1990 (n = 79 subwatersheds (155 stream sections)).

Candidate model	QAICc	ΔQAICc	ΔQAICc weight	% of maximum ΔQAICc weight
Precip, Unconsol	81.51	0	0.280	100
Precip, Mafic	82.10	0.59	0.208	74.3
Precip, Slope, Unconsol	82.96	1.45	0.136	48.6
Sumtemp, Mafic	83.23	1.72	0.119	42.5
Precip, Slope, Mafic	84.28	2.77	0.070	25.0
Precip, Slope, Unconsol, Georoad	84.90	3.39	0.051	18.2
Unconsol, Georoad, Mngclus	85.76	4.25	0.034	12.1
Slope, Mafic	85.82	4.31	0.033	11.8
Sumtemp	86.39	4.88	0.024	8.6
Georoad	87.05	5.54	0.018	6.4
Precip	88.30	6.79	0.009	3.2
Global Model	88.58	7.07	0.008	2.9
Precip, Slope, Georoad	89.44	7.93	0.005	1.8
Mngclus	91.89	10.38	0.002	0.7
Georoad, Mngclus	92.76	11.25	0.001	0.4
Slope, Mngclus	94.13	12.62	<0.001	0.2
Sumtemp, Mngclus	94.73	13.22	<0.001	0.1

Fig. 3. Relationship between percent unconsolidated lithology and predicted probability ($\hat{p}$) of a subwatershed containing either low (e.g., $\hat{p} \leq 0.5$) or moderate to high (e.g., $\hat{p} > 0.5$) count categories of steelhead parr.



rected for incomplete detectability of individuals within sections, compelled us to further subset and pool the data. In the latter case, simply modeling raw counts with covariates thought to influence detectability of fish within sampled sections will not correct for sampling bias but will only reflect how well the covariates relate to the biased counts. The matter of confounding still exists. Such a modeling approach would only be valid if (*i*) the nature and magnitude of the counting bias were known for single or repeated counts or (*ii*) repeated counts were conducted on each stream section and the true abundance did not change among counts. Changes in both abundance and covariate values across repeated counts produce confounding between biased counts and covariates. Third, by scaling up to the subwatershed level, we assumed that sampled stream sections were an adequate representation of chinook salmon or steelhead populations for all relevant stream sections within their respective subwatersheds.

Because of various difficulties inherent in the data, in this paper we placed as much emphasis on our analytic approach as we did on interpretation of results. Our procedure for subsetting and modeling a problematic data set should be of interest to fishery biologists, especially because snorkel counts are so commonly used in stream fish studies. We stress, however, that there is no substitute for proper study design and statistically sound sampling methods. It is more preferable to model counts directly than to lose information by pooling data. Nonetheless, we deemed the potential for spurious results due to biased counts to be far more serious than loss of information due to pooling data.

The modeling component of our analyses, in particular, has applications well beyond those used in this paper. AIC-based model selection has a strong theoretical basis (for details, see Burnham and Anderson 1998) and, as such, represents a fundamental departure from traditional methods of model building and variable selection based on null hypothesis testing (e.g., various stepwise and all subset selection procedures). Further, model averaging explicitly incorporates model selection uncertainty into model parameter estimates and also provides a statistically rigorous means to handle the common situation where there is no single model that is clearly better than other models. Ideally, construction of the global and candidate set of models would occur during the design stage of a study and be dictated by the research or management questions being addressed as well as existing information from previous studies. It is important to remember that AIC-based model selection will only choose the best approximating model in the candidate set; it will not correct for poor data or model choice. No analytical methods exist that can completely rescue a data set generated from an inadequately designed study.

Within the boundaries of inference allowed by the data set, there were some notable patterns that emerged between parr count categories and various landscape attributes. For instance, the negative relationship between geometric mean road density and count categories of chinook salmon may be of particular interest to land managers who are charged with ensuring the persistence of anadromous salmonid populations. Particular attention should be paid to those subwatersheds with >1 km·km^{-2} geometric mean road densities. Lee et al. (1997) also reported a negative relationship between road densities and fish population status in the Columbia Basin. Unfortunately, the correlative nature of the data is insufficient for identifying the important drivers behind this relationship. Nevertheless, these findings are noteworthy with respect to the recent road closure policy proposed by the USDA Forest Service (Federal Register 1998*b*).

The fairly strong positive influence of mean annual precipitation on count categories of chinook salmon parr may be related to the positive impact that stream discharge typically has on survival rates of anadromous salmonids (Gibson and Myers 1988; Bradford 1994; Fukushima and Smoker 1997). However, other factors related to high stream flows may be influencing chinook salmon parr numbers as well, such as lower predation rates (Bradford 1994), increased rearing habitat (Bradford 1994), and decreased egg mortality due to freezing (Gibson and Myers 1988).

Model results also infer that surrounding lithology may be especially important to steelhead parr numbers, even on a landscape scale. The fairly strong negative relationship between unconsolidated lithology and steelhead parr count categories could be related to sedimentation. An unconsolidated lithology is one that tends to slough off more than other more consolidated lithologies and hence would contribute more sediment inputs into surrounding streams, which could adversely affect parr survival (Crouse et al. 1981; Waters 1995). Conversely, a mafic lithology contains a strong alkaline component, and hence, its inputs may be tied to higher alkalinity in streams, which has been previously related to increased fish productivity (Scarnecchia and Bergersen 1987; Waters et al. 1993; Kwak and Waters 1997). This idea is consistent with the positive relationship between average maximum summer temperature (which was within the range of tolerance for steelhead) and steelhead parr count categories, where elevated summer temperature may increase primary production in a stream or parr metabolism and growth rates.

Our composite model results represent an initial approximation for fishery biologists and managers interested in mapping approximate status and quality of rearing habitats for chinook salmon and steelhead in relevant areas of Idaho. Assuming that our count categories provide an adequate index of density, subwatersheds with medium to high (>1 km·km^{-2}) geometric mean road densities and (or) low (<700 mm) mean annual precipitation levels may indicate low densities of chinook salmon parr, whereas subwatersheds with high percentages (>30%) of unconsolidated lithologies may indicate low densities of steelhead parr. These models could be updated and refined as more and better information became available and then used to help evaluate possible factors affecting salmonid population status and trends. If additional population data are collected, an effort should be made to collect them at the same spatial scale as the predictor variables.

There are probably factors unrelated to habitat that may be affecting status and distribution of these two species in Idaho. For example, deleterious effects of dams on access to spawning and rearing areas, stock productivity, and survival rates could be the overriding factors influencing parr numbers (or even presence) (Schaller et al. 1999). A number of subwatersheds may have an inherent capacity to support high parr densities, based on landscape-level habitat attributes, but may lack proper access for anadromous salmonids (e.g., blockage of the upper Snake River drainage by Hells Canyon Dam). In any event, identifying cause and effect relationships between anthropogenic variables (e.g., road density, land management practices, and dams) and parr numbers will require carefully planned, well-funded, large-scale field experiments.

Acknowledgments

R. King, M. Jones, C. Paulsen, C. Petrosky, and two anonymous reviewers provided many helpful comments on earlier drafts of this manuscript. We thank J. Hall-Griswold and C. Petrosky of the IDFG for allowing access to the parr monitoring data set as well as providing background on the General Parr Monitoring Program. J. Peterson conducted a preliminary descriptive analysis of the data set and offered helpful suggestions for its more detailed analysis. We gratefully acknowledge D. Myers for generating the study area map. This work was funded by the U.S. Department of Energy, Bonneville Power Administration, under Project 92-32 and the USDA Forest Service, Rocky Mountain Research Station.

References

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. *In* Second International Symposium on Information Theory. *Edited by* B.N. Petrov and F. Csaki. Akademiai Kiado, Budapest, Hungary. pp. 267–281.

Bradford, M.J. 1994. Trends in the abundance of chinook (*Oncorhynchus tshawytscha*) of the Nechako River, British Columbia. Can. J. Fish. Aquat. Sci. **51**: 965–973.

Bradford, M.J., Taylor, G.C., and Allan, J.A. 1997. Empirical review of coho salmon smolt abundance and the prediction of smolt production at the regional level. Trans. Am. Fish. Soc. **126**: 49–64.

Buckland, S.T., Burnham, K.P., and Augustin, N.H. 1997. Model selection: an integral part of inference. Biometrics, **53**: 603–618.

Burnham, K.P., and Anderson, D.R. 1998. Model selection and inference: a practical information-theoretic approach. Springer-Verlag, New York.

Cover, T.M., and Thomas, J.A. 1991. Elements of information theory. John Wiley & Sons, New York.

Crouse, M.R., Callahan, C.A., Malueg, K.W., and Dominguez, S.E. 1981. Effects of fine sediments on growth of juvenile coho salmon in laboratory streams. Trans. Am. Fish. Soc. **110**: 281–286.

Daly, C., Neilson, R.P., and Phillips, D.L. 1994. A statistical–topographic model for mapping climatological precipitation over mountainous terrain. J. Appl. Meteorol. **33**: 140–158.

Dunham, J.B., and Rieman, B.E. 1999. Metapopulation structure of bull trout: influences of physical, biotic, and geometrical landscape characteristics. Ecol. Appl. **9**: 642–655.

Federal Register. 1997. Endangered and threatened species: listing of several evolutionary significant units (ESUs) of west coast steelhead. Fed. Regist. **62**(159): 43 937 – 43 954.

Federal Register. 1998*a*. Endangered and threatened species: withdrawal of proposed rule to list Snake River spring/summer chinook and fall chinook salmon as endangered. Fed. Regist. **63**(7): 1807–1811.

Federal Register. 1998*b*. Administration of the forest development transportation system. Fed. Regist. **63**(18): 4350–4351.

Fukushima, M., and Smoker, W.W. 1997. Determinants of stream life, spawning efficiency, and spawning habitat in pink salmon in the Auke Lake system. Can. J. Fish. Aquat. Sci. **54**: 96–104.

Gerard, P.D., Smith, D.R., and Weerakkody, G. 1998. Limits of retrospective power analysis. J. Wildl. Manage. **62**: 801–807.

Gibson, R.J., and Myers, R.A. 1988. Influence of seasonal river discharge on survival of juvenile Atlantic salmon, *Salmo salar*. Can. J. Fish. Aquat. Sci. **45**: 344–348.

Hall-Griswold, J.A., and Petrosky, C.E. 1996. Idaho habitat/natural production monitoring. Part 1. U.S. Dep. Energy Bonneville Power Admin. Annu. Rep. 1995, Project 91-73. Available on the Internet at <http://www.efw.bpa.gov/Environment/EW/EWP/DOCS/REPORTS/HABITAT/withpdf.htm>

Hosmer, D.W., Jr., and Lemeshow, S. 1989. Applied logistic regression. John Wiley & Sons, New York.

Kullback, S., and Leibler, R.A. 1951. On information and sufficiency. Ann. Math. Stat. **22**: 79–86.

Kwak, T.J., and Waters, T.F. 1997. Trout production dynamics and water quality in Minnesota streams. Trans. Am. Fish. Soc. **126**: 35–48.

Lee, D.C., Sedell, J.R., Rieman, B.E., Thurow, R.F., Williams, J.E., Burns, D., Clayton, J., Decker, L., Gresswell, R., House, R., Howell, P., Lee, K.M., Macdonald, K., McIntyre, J., McKinney, S., Noel, T., O'Connor, J.E., Overton, C.K., Perkinson, D., Tu, K., and Van Eimeren, P. 1997. Broadscale assessment of aquatic species and habitats. *In* An assessment of ecosystem components in the interior Columbia Basin and portions of the Klamath and Great basins. Vol. III. *Edited by* T.M. Quigley and S.J. Arbelbide. USDA For. Serv. Gen. Tech. Rep. PNW-GTR-405, Portland, Oreg. pp. 1057–1713. Available on the Internet at <http://www.fs.fed.us/pnw/int_col.htm>

Nehlsen, W., Williams, J.E., and Lichatowich, J.A. 1991. Pacific salmon at the crossroads: stocks at risk from California, Oregon, Washington, and Idaho. Fisheries (Bethesda), **16**(2): 4–21.

Northwest Power Planning Council. 1986. Compilation of information on salmon and steelhead losses in the Columbia River Basin. Northwest Power Planning Council, Portland, Oreg.

Petrosky, C.E., and Holubetz, T.B. 1986. Idaho habitat evaluation for off-site mitigation record. U.S. Dep. Energy Bonneville Power Admin. Annu. Rep. 1985, Project 83-7, Portland, Oreg.

Quigley, T.M., and Arbelbide, S.J. (*Technical Editors*). 1997. An assessment of ecosystem components in the interior Columbia Basin and portions of the Klamath and Great basins. USDA For. Serv. Gen. Tech. Rep. PNW-GTR-405, Portland, Oreg. Available on the Internet at <http://www.fs.fed.us/pnw/int_col.htm>

Rich, B.A., and Petrosky, C.E. 1994. Idaho habitat and natural production monitoring. Part 1. U.S. Dep. Energy Bonneville Power Admin. Annu. Rep. 1992, Project 91-73, Portland, Oreg.

Rodgers, J.F., Solazzi, M.F., Johnson, S.L., and Buckman, M.A. 1992. Comparison of three techniques to estimate juvenile coho salmon populations in small streams. N. Am. J. Fish. Manage. **12**: 79–86.

Rosgen, D.L. 1985. A stream classification system. *In* Riparian ecosystems and their management: reconciling conflicting uses. *Edited by* R.R. Johnson, C.D. Ziebell, D.R. Palton, P.F. Ffolliott, and R.H. Hamre. USDA For. Serv. Gen. Tech. Rep. RM-120, Fort Collins, Colo. pp. 91–95.

Rosgen, D.L. 1996. Applied river morphology. Wildland Hydrology, Pagosa Springs, Colo.

Royall, R.M. 1997. Statistical evidence: a likelihood paradigm. Chapman and Hall, New York.

SAS Institute Inc. 1996. SAS/STAT software: changes and enhancements, through version 6.11. SAS Institute Inc., Cary, N.C.

Scarnecchia, D.L., and Bergersen, E.P. 1987. Trout production and standing crop in Colorado's small streams, as related to environmental features. N. Am. J. Fish. Manage. **7**: 315–330.

Schaller, H.A., Petrosky, C.E., and Langness, O.P. 1999. Contrasting patterns of productivity and survival rates for stream-type chinook salmon (*Oncorhynchus tshawytscha*) populations of the Snake and Columbia rivers. Can. J. Fish. Aquat. Sci. **56**: 1031–1045.

Schlosser, I.J. 1991. Stream fish ecology: a landscape perspective. BioScience, **41**: 704–712.

Thurow, R.F. 1994. Underwater methods for study of salmonids in the Intermountain West. USDA For. Serv. Gen. Tech. Rep. INT-GTR-307, Ogden, Utah.

Torgersen, C.E., Price, D.M., Li, H.W., and McIntosh, B.A. 1999. Multiscale thermal refugia and stream habitat associations of chinook salmon in northeastern Oregon. Ecol. Appl. **9**: 301–319.

Waters, T.F. 1995. Sediment in streams: sources, biological effects and control. Am. Fish. Soc. Monogr. No. 7.

Waters, T.F., Kaehler, J.P., Polomis, J.T., and Kwak, T.J. 1993. Production dynamics of smallmouth bass in a small Minnesota stream. Trans. Am. Fish. Soc. **122**: 588–598.

Yoccoz, N.G. 1991. Use, overuse, and misuse of significance tests in evolutionary biology and ecology. Bull. Ecol. Soc. Am. **72**: 106–111.