The Auk
Ornithological Advances
AmericanOrnithology.org

Volume 134, 2017, pp. 783–801
DOI: 10.1642/AUK-16-195.1

RESEARCH ARTICLE

Logistic quantile regression provides improved estimates for bounded avian counts: A case study of California Spotted Owl fledgling production

Brian S. Cade,[1]* Barry R. Noon,[2] Rick D. Scherer,[2] and John J. Keane[3]

[1] Fort Collins Science Center, U.S. Geological Survey, Fort Collins, Colorado, USA
[2] Department of Fish and Wildlife Conservation, Colorado State University, Fort Collins, Colorado, USA
[3] Pacific Southwest Research Station, U.S. Forest Service, Davis, California, USA
* Corresponding author: cadeb@usgs.gov

Submitted September 19, 2016; Accepted April 19, 2017; Published July 12, 2017

ABSTRACT

Counts of avian fledglings, nestlings, or clutch size that are bounded below by zero and above by some small integer form a discrete random variable distribution that is not approximated well by conventional parametric count distributions such as the Poisson or negative binomial. We developed a logistic quantile regression model to provide estimates of the empirical conditional distribution of a bounded discrete random variable. The logistic quantile regression model requires that counts are randomly jittered to a continuous random variable, logit transformed to bound them between specified lower and upper values, then estimated in conventional linear quantile regression, repeating the 3 steps and averaging estimates. Back-transformation to the original discrete scale relies on the fact that quantiles are equivariant to monotonic transformations. We demonstrate this statistical procedure by modeling 20 years of California Spotted Owl fledgling production (0 3 per territory) on the Lassen National Forest, California, USA, as related to climate, demographic, and landscape habitat characteristics at territories. Spotted Owl fledgling counts increased nonlinearly with decreasing precipitation in the early nesting period, in the winter prior to nesting, and in the prior growing season; with increasing minimum temperatures in the early nesting period; with adult compared to subadult parents; when there was no fledgling production in the prior year; and when percentage of the landscape surrounding nesting sites (202 ha) with trees 25 m height increased. Changes in production were primarily driven by changes in the proportion of territories with 2 or 3 fledglings. Average variances of the discrete cumulative distributions of the estimated fledgling counts indicated that temporal changes in climate and parent age class explained 18% of the annual variance in owl fledgling production, which was 34% of the total variance. Prior fledgling production explained as much of the variance in the fledgling counts as climate, parent age class, and landscape habitat predictors. Our logistic quantile regression model can be used for any discrete response variables with fixed upper and lower bounds.

Keywords: bounded counts, cumulative ordinal logistic regression, fledgling production, logit transformation, quantile regression, Spotted Owl

La regresión logística por cuantiles provee mejores estimados de conteos de aves numéricamente limitados: un caso de estudio en una población de volantones de *Strix occidentalis occidentalis*

RESUMEN

Los conteos de volantones, pichones o nidadas de aves tienen números limitados entre cero y un número entero bajo de una variable discreta con distribución aleatoria que no puede ser bien estimada por distribuciones paramétricas de conteos convencionales como la de Poisson o la binomial negativa. Desarrollamos un modelo de regresión logística por cuantiles que provee estimados de la distribución condicional empírica de una variable discreta aleatoria con límites. El modelo de regresión logística por cuantiles requiere que a los conteos se les agregue ruido aleatoriamente para construir una variable continua aleatoria, que sean transformados con una función logit para asignar límites entre valores máximos y mínimos específicos, y que sean luego estimados en regresiones por cuantiles convencionales, repitiendo los tres pasos y haciendo un promedio de los estimados. Transformar los datos en reversa a la escala discreta original depende del hecho de que los cuantiles son equivariantes con las transformaciones monotónicas. Demostramos este procedimiento estadístico modelando 20 años de producción de volantones de *Strix o. occidentalis* (entre 0 y 3 por territorio) en el Parque Nacional Lassen, California, USA, y su relación con el clima, la demografía y las características del paisaje del hábitat en los territorios. Los conteos de volantones de *S. o. occidentalis* incrementaron de manera no lineal con disminuciones en la precipitación en el periodo temprano de la anidación, en el invierno anterior a la anidación y en la temporada de crecimiento anterior. Dichos conteos también disminuyeron cuando incrementaron las temperaturas mínimas durante el periodo temprano de anidación, cuando no hubo producción de

volantones en el año anterior y cuando incrementó el porcentaje del paisaje circundante a los sitios de anidación (202 ha) con árboles de más de 25m de altura. Además, los conteos fueron menores para padres adultos en comparación con padres subadultos. Los cambios en la producción se vieron afectados principalmente por cambios en la proporción de territorios con dos o tres volantones. Las varianzas promedio de las distribuciones discretas acumulativas de los conteos de volantones indicaron que los cambios temporales en el clima y en la clase de edad de los padres explicaron 18% de la varianza anual en la producción de volantones, la cual fue el 34% del total de la varianza. La producción previa de volantones explicó tanta varianza en los conteos de volantones como el clima, la clase de edad de los padres y los predictores que describían el paisaje del hábitat. Nuestro modelo de regresión logística por cuantiles puede ser usado para cualquier variable de respuesta discreta con límites máximos y mínimos fijos.

Palabras clave: conteos de rango limitado, regresión logística ordinal acumulativa, producción de volantones, transformación logit, regresión por cuantiles, *Strix occidentalis*

INTRODUCTION

Bounded small counts such as those associated with avian clutch size and number of nestlings or fledglings per nest or territory present several difficulties for statistical modeling. When empirical count distributions are bounded below by zero as well as above by some small integer (e.g., 0 3), typical parametric count distributions (e.g., Poisson, negative binomial, and their zero-inflated counterparts) that might be considered are unlikely to be reasonable approximations of the statistical distribution. The unrestricted higher counts associated with these parametric distributions and their mean–variance relationships do not reasonably approximate empirical distributions of bounded small counts. Simulations by McDonald and White (2010) suggested that using ordinary least squares (OLS) regression to model mean counts, even though it violates the assumption of homogeneous, normally distributed errors, was better than using inappropriate parametric count distributions for bounded small counts, especially with regard to power for detecting temporal trends. Ives (2015) made similar recommendations regarding use of log-transformed counts and OLS regression. Ordinal and multinomial logistic regression models for small counts also were considered by McDonald and White (2010) but often had lower power or computational difficulties associated with sparse counts in some region of the predictor space. While the power characteristics of OLS regression may be compelling, other issues such as incorrect variance estimates, predicted means and prediction intervals outside the range of the bounded counts, and an inability to decompose changes in mean counts to changes in proportions (quantiles) of the individual counts remain problematic.

Here, we develop an alternative logistic quantile regression model for bounded small counts that is applicable to statistical analyses of clutch, nestling, and fledgling counts for many avian species. Over half of 5,290 avian species have clutch sizes 3 and around 90% have clutch sizes 6 (Jetz et al. 2008), suggesting that small bounds on the upper counts of reproductive output are pervasive in ornithology. We demonstrate the method and advantages of modeling bound counts using logistic quantile regression with an analysis of California Spotted Owl (*Strix occidentalis occidentalis*) fledgling production on territories. Spotted Owl fledgling counts at a nesting territory are bounded between 0 and 3 (Franklin et al. 2004, Anthony et al. 2006, Blakesley et al. 2010), the range of counts that were the focus of the simulations in McDonald and White (2010). We compare estimates and interpretations with those provided by OLS mean regression, and also discuss similarities and differences between the logistic quantile count model and cumulative ordinal logistic regression without a proportional odds assumption, another statistical model that was not evaluated by McDonald and White (2010).

Quantile regression has been increasingly used in ecology as a general method for modeling heterogeneity in continuous outcome variables (Cade and Noon 2003, Cade et al. 2005, Lancaster and Belyea 2006, Simkin et al. 2016). Quantile regression provides a highly flexible statistical procedure to estimate heterogeneity and skewness in outcomes without requiring parametric distributional assumptions by modeling all or selected parts of the cumulative distribution of outcomes conditional on predictor variables through their inverse the quantiles (Cade and Noon 2003, Koenker 2005). Our logistic quantile regression model for bounded counts relies on another important property of quantile regression to extend estimates to bounded outcomes, equivariance to monotonic transformations (Machado and Santos Silva 2005, Cade and Dong 2008) including the nonlinear logit transformation. The logistic quantile regression model provides estimates and prediction intervals that are constrained to remain within the specified bounds without requiring any parametric assumption about the distributional form of the counts. Variances of the discrete cumulative distribution functions (cdf) of the counts estimated by logistic quantile regression for various models can be computed and compared.

Our development of the logistic quantile regression model for bounded counts was motivated by a desire to provide improved estimates and inferences for fledgling production of the Spotted Owl. The Spotted Owl has been intensively studied over the last 25 years with numerous statistical modeling endeavors intended to estimate population growth rates and other demographic parameters associated with various limiting factors. Populations of the various subspecies of this endangered/threatened species have been negatively impacted by loss of old-growth forest habitat due to logging, by changing weather and climate, and by competition with recently colonizing Barred Owls (*Strix varia*) (Franklin et al. 2000, 2004; Seamans et al. 2001; Olson et al. 2005; Anthony et al. 2006; Blakesley et al. 2010; Forsman et al. 2011; Glenn et al. 2011; Dugger et al. 2016). Reproductive output estimated from fledgling counts (i.e. the number of young successfully reared through leaving the nest) is a critical quantity required for most Spotted Owl population growth models (Franklin et al. 2000; Dugger et al. 2005; Blakesley et al. 2010; Glenn et al. 2010, 2011). Survival probabilities of adult California Spotted Owls are high with low temporal variation (Seamans et al. 2001, Franklin et al. 2004, Blakesley et al. 2010), while reproductive output is low and highly variable across years (North et al. 2000, Franklin et al. 2004, Blakesley et al. 2010). Previous studies of both Northern and California Spotted Owls found that years of high fledgling production are often followed by years of low production, suggesting a cost of current to future reproduction (Franklin et al. 2004, Anthony et al. 2006, Blakesley et al. 2010, Stoelting et al. 2015). This pattern may also reflect a bet-hedging strategy where adult survival is favored over current reproduction (Franklin et al. 2000, Noon and Franklin 2002, Glenn et al. 2010). This strategy may minimize costs of reproduction as owls may only reproduce in years when conditions favor current reproduction at no or relatively low cost to future survival and reproduction.

Here we first develop the logistic quantile regression model for bounded counts and then demonstrate its use by modeling changes in fledgling counts of California Spotted Owls in the Lassen National Forest, California, USA. We compare estimates of annual variation in fledgling production with estimates related to temporal variation in climate over 20 years, age class of parents, and landscape habitat surrounding nesting territories. We include terms for prior production of fledglings at territories in our models specifically to account for a possible cost of prior reproduction. The climate variables we used as predictors are intended to capture environmental conditions that are thought to provide favorable conditions for successful reproduction. Heterogeneity in the fledgling counts is inherently of substantive interest for evaluating the effects of various demographic, habitat, and climatic factors on fledgling production and would be ignored by regression modeling of means. Heterogeneous changes in proportions of territories with 1, 2, or 3 fledglings could potentially reflect responses to different physiological and environmental constraints. Finally, we provide additional suggestions on the use of logistic quantile regression for modeling bounded counts in other applications.

METHODS

Logistic Quantile Regression Model For Bounded Counts

Our logistic quantile regression model for bounded counts combines the approach of estimating quantiles for discrete counts by randomly jittering them into a continuous random variable (Machado and Santos Silva 2005, Cade and Dong 2008) with the logit transformation approach for estimating bounded responses (Bottai et al. 2010). This approach allows estimation and inference for quantile regression to be made in the conventional linear model formulation with continuous responses, conditions where the properties of the quantile regression estimates are well understood and inferential methods are well developed (Koenker 2005). Estimates made in the continuous linear scale are then back-transformed into the desired discrete count scale without bias because of the equivariance to monotonic transformation property of quantiles. The equivariance property of quantile regression implies that for a nonlinear, monotonically increasing transformation function $h(\ )$ on any random variable Y then $Q_{h(Y)}(\tau) = h(Q_Y(\tau))$, where $Q_Y(\tau)$ denotes the τth quantiles ($0 < \tau < 1$) of the conditional distributions of Y (Koenker 2005). For example, if $h(\ )$ is a logit or logarithmic transformation then we can estimate $Q_Y(\tau)$ without bias from an estimate on $Q_{h(Y)}(\tau)$ by applying the inverse transformation (h^{-1}). This property does not hold for means.

The logistic quantile model for bounded counts takes the following form where the response variable Y are the counts, $Z = Y + U[0,1)$ are continuous versions of the counts made by adding random uniform numbers in the interval [0, 1), $Q_Y(\tau|\mathbf{X})$ and $Q_Z(\tau|\mathbf{X})$ denote the τth quantiles ($0 < \tau < 1$) of the conditional distributions of Y and Z, respectively:

$$Q_z(\tau|\ \mathbf{X}) = \frac{\exp(\mathbf{X}\boldsymbol{\beta}(\tau))(zmax) + (zmin - \tau)}{\exp(\mathbf{X}\boldsymbol{\beta}(\tau)) + 1} \quad (1)$$

where $\mathbf{X}$ is an $n \times p$ matrix of predictors including a column of 1's for the intercept, $\boldsymbol{\beta}$ is a $p \times 1$ vector of parameters, *zmax* is the maximum possible value of Z, and *zmin* is the minimum possible value of Z. The additive term τ occurs with *zmin* because the $Q_Z(\tau|\mathbf{X})$ are bounded below by τ due to the addition of the random $U[0, 1)$

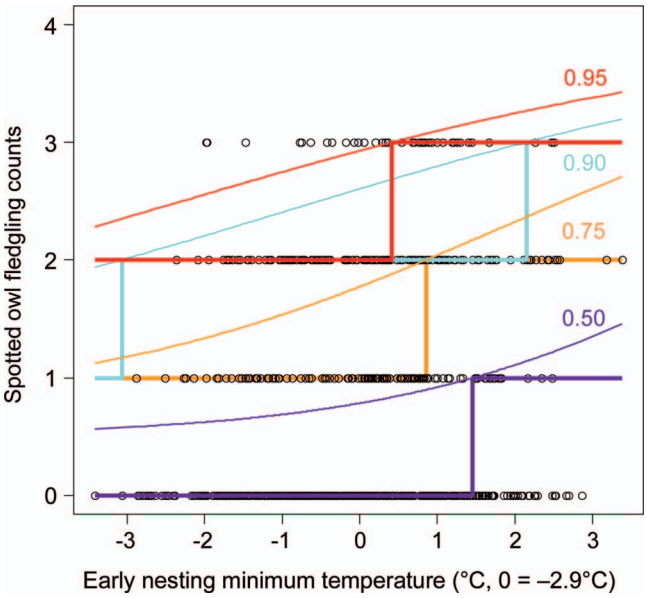


FIGURE 1. Example estimates for $\tau = 0.95$ (red), 0.90 (cyan), 0.75 (orange), and 0.50 (purple) for logistic quantile regression model in the randomly jittered continuous scale (thin curves) and interpolated to the discrete step functions (thick lines) for counts of California Spotted Owl fledglings as a linear function of early nesting minimum temperature ($n = 707$).

numbers, requiring that the lower bound of the logistic function is increased to *zmin* τ to keep $Q_Z(\tau|\mathbf{X})$ properly bounded. The transformation back to the original Y counts uses the ceiling function and returns the bounds to the minimum and maximum of Y:

$$Q_Y(\tau|\mathbf{X}) = \lceil Q_Z(\tau|\mathbf{X}) \rceil \quad (2)$$

where $\lceil a \rceil$ returns the smallest integer greater than or equal to a. The model for a sample of the continuous response z is estimated in its linear form by making the logit transformation of z (Bottai et al. 2010): logit $(z) = \log [(z \quad (zmin \quad \tau)) \div (zmax \quad z)] = \mathbf{X}\boldsymbol{\beta}(\tau)$. When logit (z) is undefined for $z \quad (zmin \quad \tau)$, we used the log of 0.00001 (any other reasonably small nonzero value would suffice). The utility of this transformation relies on the properties that quantiles are equivariant to monotonic transformations and invariant to censoring from below up to the quantile of interest (Machado and Santos Silva 2005, Cade and Dong 2008). Estimates of $\boldsymbol{\beta}(\tau)$ are obtained with conventional linear quantile regression for continuous outcomes by minimizing the asymmetrically weighted (τ and 1 τ) sum of absolute deviations (Cade and Noon 2003, Koenker 2005). Estimates can be obtained for any increments of τ that makes sense relative to the distribution of counts and sample sizes available for model estimation, where smaller increments are more likely to provide unique estimates with larger sample sizes. The objective is to provide reasonably precise estimates of proportions of the modeled cumulative distribution associated with different counts.

To remove the small source of extra variation imparted by adding random U [0, 1) numbers to y, we estimated the model (equation 1) $m = 500$ times using m random samples from the uniform distribution U[0, 1) and averaged the parameter estimates following Machado and Santos Silva (2005). Thus, our estimates of the y counts were based on the inverse logit transformation to:

$$\widehat{Q}_Y(\tau|\mathbf{X}) = \lceil \widehat{Q}_Z(\tau|\mathbf{X}) \rceil = \left\lceil \frac{\exp \quad \mathbf{X}\hat{\boldsymbol{\beta}}(\tau) \quad (\text{zmax}) + (\text{zmin} \quad \tau)}{\exp \quad \mathbf{X}\hat{\boldsymbol{\beta}}(\tau) \quad 1} \quad 1 \right\rceil \quad (3)$$

where $\hat{\boldsymbol{\beta}}(\tau)$ are the averages of the m estimated parameters for the m realizations of Z. The essence of this approach is that the smooth logistic quantile functions estimated in the continuous response scale interpolate the jumps in the step functions that define quantiles for the discrete counts, where the logistic functional form constrains estimates to always be between the lower and upper bounds (Figure 1). Unlike the smooth functions estimated for means in conventional parametric count models (e.g., Poisson, negative binomial), the quantile regression estimates are by definition step functions because of the discrete nature of the cdf for counts (Cade and Dong 2008). Confidence intervals were estimated in the continuous Z scale by averaging confidence interval endpoint estimates based on inverting the quantile rank score test across the m random iterations (Cade and Dong 2008). Type I error rates and power of the rank score test were investigated in Cade et al. (2006). Other inference procedures based on the asymptotic variance/covariance for quantile regression are possible and discussed by Machado and Santos Silva (2005).

We selected among candidate models based on average differences in Akaike Information Criterion (AIC) by (1) obtaining AIC for each of the candidate models and a base, reference model at each of the $m = 500$ simulations by τ; (2) computing differences in AIC (ΔAIC) for each candidate model from the base model at each of the m simulations by τ; and (3) then averaging ΔAIC for each candidate model across the m simulations by τ. This provides a function of average ΔAIC by τ for each candidate model that can be compared. The equivalence between the weighted sums of absolute deviations minimized in quantile regression estimation and maximum likelihood estimates required for computing AIC is based on assuming an asymmetric double exponential distribution (Koenker and Machado 1999, Yu and Moyeed 2001, Cade et al. 2005:appendix C).

Spotted Owl Fledgling Counts on the Lassen National Forest

We used fledgling counts made on surveys conducted from 1991 to 2010 on 88 California Spotted Owl territories on the Lassen National Forest in the southern Cascade Mountains in northern California, USA. Mixed-conifer forests composed of white fir (*Abies concolor*), ponderosa pine (*Pinus ponderosa*), sugar pine (*Pinus lambertiana*), Jeffrey pine (*Pinus jeffreyi*), incense cedar (*Calocedrus decurrens*), and Douglas-fir (*Pseudotsuga menziesii*) at elevations of 1,200–2,100 m were the dominant habitat types. Additional description of the owl habitat on the Lassen National Forest is provided in Franklin et al. (2004) and Blakesley et al. (2005). Field methods for repeated surveying of individual Spotted Owl territories to determine the number of young fledged, age of parents, and identification of banded individuals are described in detail in Franklin et al. (2004) and Blakesley et al. (2010). In general, these methods involved offering, on multiple occasions, live mice to male or female owls and observing if the mice were cached or eaten or taken to a parent on a nest or to juvenile owls (Lint et al. 1999). We only used observations ($n = 795$ territory-years) where these survey methods provided near certainty about the number of young fledged. The temporal sequence of the fledgling counts was slightly irregular with 88% of the territory-year observations in consecutive years and 12% 2 14 years apart. Fifty-one of the 88 territories had 7 years of fledgling counts.

Hypothesized Spotted Owl Relationships and Choice of Predictor Variables

The sampling units for assigning annual values to reproductive outcome, climate, and habitat predictor variables were based on each breeding pair's most frequently used nest site location, or locations averaged across years, or the average location of juvenile owls within their natal territory prior to dispersal (see Blakesley et al. 2005). Because Spotted Owl pairs show strong site fidelity (Blakesley et al. 2006), cumulative nest site or fledgling locations across years are generally closely aggregated and collectively defined the pair's territory.

Based on previous studies we expected most of the variation in fledgling counts to be temporal rather than spatial (Franklin et al. 2004, Blakesley et al. 2010, Glenn et al. 2011, Dugger et al. 2016). We estimated annual variation in fledgling counts in a model that included year as a categorical predictor, adjusting for the autocorrelated nature of the repeated surveys on territories with a 1st order lag-effect that accounted for the irregular intervals between some surveys. We used a 1st order lag-effect to account for reduced fledgling production associated with prior fledgling production on a territory, accounting for the cost of reproduction suggested by previous studies of Spotted Owls (Anthony et al. 2006, Forsman et al. 2011, Stoelting et al. 2015). Our statistical modeling focused on addressing how much of the annual variation in fledgling counts could be explained by temporal variation that was related to climate and demographic characteristics of the owls at territories. We also examined whether additional variation in fledgling production was related to spatial variation in large-scale habitat characteristics surrounding territories.

Based on previous studies with Northern Spotted Owls (Rosenberg et al. 2003; Dugger et al. 2005; Glenn et al. 2010, 2011), we hypothesized that increased precipitation and decreased minimum temperatures in the winter (November–February) prior to nesting, during the early nesting period (March–April), and during late-nesting (May–June) to have negative impacts on fledgling production. We also examined whether increased precipitation in the previous growing season (May–October) might positively impact fledgling production because of an expected increase in vegetation supporting an increased rodent prey base. Climatic variables (temperature and precipitation) were calculated for territory locations using PRISM (PRISM Climate Group 2004). We used a reduced set of the survey data collected from territories occupied by banded owls of known age (adults 3 years age, subadults 1 2 years age), including 94% of males and 97% of females, to estimate effects of parent age class on fledgling production in conjunction with climatic and habitat variables. We anticipated a reduction in fledgling production with subadult parents (Franklin et al. 2004, Blakesley et al. 2010, Glenn et al. 2011, Stoelting et al. 2015, Dugger et al. 2016) and estimated the effect separately for males and females.

We used landscape-scale habitat measures estimated at 2 spatial scales (202 and 121 ha [500 and 300 acres, respectively]) surrounding territory centroids to characterize spatial variation in mature to old-growth forest habitat of owl territories. Scales of analysis were based on 2 estimates of core area—portions of an owl's breeding season home range that received disproportionately high use (Bingham and Noon 1997). The habitat variables percent of area in trees 25 m height, percent of area in trees 50 m height, average height of trees, standard deviation of height of trees, total canopy cover of trees, and average elevation were derived from the LANDFIRE database (LANDFIRE 2008).

The Logistic Quantile Regression Model for Spotted Owl Fledglings

Our logistic quantile regression estimates of changes in Spotted Owl fledgling counts (0 3) used $zmax = 4$ for the maximum possible value of z and $zmin = -\theta$ for the minimum possible value of z. We obtained estimates from $\tau = 0.05$ to $\tau = 0.98$ by increments of 0.01, which was adequate for providing estimates of the proportion of the

cdf associated with counts y {0, 1, 2, 3} to the nearest 1% where zeros dominated the lower quantiles. With larger sample sizes it would be possible to obtain estimates across finer scale increments of τ, e.g., by increments of 0.001.

The specific parameterization of our models for Spotted Owl fledgling production was designed to accommodate the irregular time-series of repeated counts of fledglings at a territory. We initially considered an autoregressive model that was linear in the time gap between successive counts similar to Wei et al. (2006). However, because few territories were observed at intervals >1 year (12% were 2–14 years apart), we simplified this 1st-order autoregressive effect to include a categorical predictor for whether the prior count was in the previous year or 2 years prior, a categorical predictor for whether the previous count was of 0 or 1 fledglings, and the interaction of these 2 categorical predictors; X_1 X_2 X_1X_2, where $X_1 = 0$ if prior production was 0 fledglings and 1 otherwise and $X_2 = 0$ if prior production was in the previous year and 1 otherwise. This parameterization allowed the majority of the observations (88%) that were in consecutive years to provide primary estimates of the effect of prior production while allowing other observations separated by 2 or more years also to be accommodated in the models. Incorporating the lagged fledgling counts in this autoregressive parameterization reduced sample size for our models to $n = 707$ territory-year observations. This 4-parameter model of the lagged effect of prior production served as our base, reference model.

Our estimate of annual variation in fledgling counts across 1992 to 2010 (1991 was absorbed by the 1st-order lag effect) was made by including year as a categorical predictor requiring 18 orthogonal contrast variables (1, 0, 1) to indicate deviation from the average (the intercept term β_0) across the 19 years given the 1st-order lag effects. The combination of the 1st-order prior fledgling production and annual variation constitutes our estimate of temporal variation. We then considered models with the lag effect of prior production and various combinations of the climate predictors (all centered on their means), first just using linear terms and then including quadratic terms. The variance attributed to these models was compared to the previous model of temporal variation to see how much of the annual variation in fledgling counts might be related to climate predictors. We then included age class of female (indicator variable with 0 if adult 3 years age and 1 otherwise) and male (indicator variable with 0 if adult 3 years age and 1 otherwise) parents for models using a reduced set of observations ($n = 639$) for territory-years with marked individuals of known age. The previous models for annual variation and climate predictors were re-estimated with this subset of observations for marked individuals so that comparisons could be made for the reduction in variances of the cdf of counts attributable to models with different combinations of predictor variables. Finally, we considered adding in landscape measures of habitat around nesting territories (mean centered predictors), with linear and quadratic terms, to see whether the spatial variation in habitat might explain any additional variance in the cdf of counts.

We graphed changes in estimates of the discrete cdf across years to display annual variation and across individual predictors in our best (largest average ΔAIC) climate, parent age class, and landscape habitat model to interpret patterns of partial effects. This was accomplished by accumulating all the predicted counts by quantile conditional on the predictor values and then graphing the quantiles (proportions of the partial cdfs) associated with shifts to the next highest count across each predictor variable. Because the step functions estimated by the logistic quantile regression model with a restricted range of counts result in considerable overlap in estimated quantiles by count (Figure 1), partial effects plots based on the entire cdf will be more readily interpreted than those based on a few selected quantiles as used in the quantile count models of Cade and Dong (2008).

We estimated variances in counts attributed to different models by computing means and variances of the estimated cdf of the counts for each observation in a model and then averaging the variances across the n observations. The variance for the discrete cdf of a model estimated with n observations is

$$\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} \widehat{\sigma}_i^2$$

where the variance of the discrete cdf for the ith observation is

$$\widehat{\sigma}_i^2 = (0 \quad \bar{Y}_i)^2 \times p_i(0) + (1 \quad \bar{Y}_i)^2 \times p_i(1) + (2 \quad \bar{Y}_i)^2 \times p_i(2) + (3 \quad \bar{Y}_i)^2 \times p_i(3),$$

$$\bar{Y}_i = 0 \times p_i(0) + 1 \times p_i(1) + 2 \times p_i(2) + 3 \times p_i(3)$$

and p_i (0) to p_i (3) are proportions of the estimated cdf for the ith observation with counts y {0, 1, 2, 3}. Our bounded logistic model (equation 3) provided us with estimates of the discrete cdf for the counts at all observations. However, because of estimation instability at the most extreme quantiles we assumed counts associated with $\tau > 0.98$ took the same value as estimated for $\tau = 0.98$. As this represents at worst a 0.02 proportionate difference in the computation of the variance of the discrete cdf if the count estimated for $\tau =$ 0.98 was other than 3, this inaccuracy is negligible. Scripts that use the quantreg package for R to perform our logistic quantile regression analyses are provided in the Supplemental Material Appendix B.

TABLE 1. Average variances (σ^2) of the discrete cumulative distribution function (cdf) of California Spotted Owl fledglings estimated from logistic quantile regression (QR) models for bounded (0–3) counts and estimated for ordinary least squares (OLS) mean regression for variance components decomposed by various models with p predictors for 1st-order autoregressive effects of prior production, annual variation (19 years), linear plus quadratic climate, age class (adult versus subadult) of parents, and linear plus quadratic landscape habitat surrounding territories, Lassen National Forest, California, 1991–2010. Sample size is $n = 639$ for all models that include only territory-years with marked individuals and with samples for 1991 absorbed in the 1st-order autoregressive effect for prior production.

Source of variance	Model	p	σ^2 logistic QR	σ^2 OLS
Total	Intercept	1	0.9571	0.9289
Prior production	Intercept autoregressive	4	0.9023	0.9090
Temporal	Prior production annual	22	0.6269	0.6273
Climate	Prior production linear plus quadratic climate	12	0.8750	0.8427
Climate parent age	Climate male/female parent age	14	0.8530	0.8243
Climate parent age landscape	Climate parent age linear plus quadratic landscape	16	0.8374	0.8145
% total σ^2 explained by temporal			34.5	32.5
% annual σ^2 explained by climate			9.9	23.5
% annual σ^2 explained by climate parent age			17.9	30.1
% total σ^2 explained by climate parent age landscape habitat			7.2	10.4

Cumulative Ordinal Logistic Regression Model without Proportional Odds

We also estimated our final fledgling counts models with cumulative ordinal logistic regression without a proportional odds assumption (Liu and Agresti 2005, Yee 2010, Agresti 2013) because of its similarity to our logistic quantile regression model. The cumulative ordinal logistic regression model without proportional odds is

$$\mathrm{P}(Y \quad k|\mathbf{X}) = \frac{\exp(\mathbf{X}\boldsymbol{\beta}_k)}{\exp(\mathbf{X}\boldsymbol{\beta}_k) + 1} = \mathrm{logit}[\mathrm{P}(Y \quad k|\mathbf{X}) = \mathbf{X}\boldsymbol{\beta}_k, k = 1, 2, \text{and } 3 \text{ fledglings} \qquad (4)$$

Unlike the proportional odds version of cumulative ordinal logistic regression, the slopes for all predictors in this model are allowed to vary by count category ($Y \quad 1$, $Y \quad 2$, and $Y \quad 3$), similar to what occurs with the logistic quantile regression. We initially estimated simple cumulative ordinal logistic regression models with the vector generalized linear modeling function vglm in the VGAM package for R (Yee 2010). However, numerous convergence issues with the vglm routine when using interactions and quadratic terms required that we estimate our final cumulative ordinal logistic regression models without proportional odds as a sequence of binary logistic regressions (Agresti 2013). We did not perform a model selection exercise on the cumulative ordinal logistic regression model because our intent was to compare these estimates with estimates obtained from a similar logistic quantile regression model. We present details of the cumulative ordinal logistic regression estimates in the Appendix and discuss similarities and differences with the logistic quantile regression estimates. Scripts that were used in R to estimate our cumulative ordinal logistic regression models are provided in the Supplemental Material Appendix C.

Ordinary Least Squares (OLS) Regression

We estimated the naïve OLS regression model of mean counts recommended by McDonald and White (2010) using exactly the same combinations of linear and quadratic terms for predictors as used in our logistic quantile regression models. The OLS regression model for mean counts is

$$\mathrm{E}[Y|\mathbf{X} = \mathbf{X}\boldsymbol{\beta} \quad \varepsilon, \qquad (5)$$

where ε are normally distributed $\mathrm{N}(0,\sigma 2)$.

The linear response of the mean in this model might reasonably approximate the central 80% of the response modeled by the logistic quantile regression for counts because the logistic model form is nearly linear over that portion of its range. However, estimates of mean counts and prediction intervals are not guaranteed to remain within the bounds of the counts due to heterogeneous variances, skewness, and the unbounded nature of the assumed continuous normal error distribution. We did not perform a model selection exercise on the OLS regression model because our intent was to compare these estimates with estimates obtained from a similar logistic quantile regression model. We present details of the OLS regression estimates in the Appendix and discuss similarities and differences with the logistic quantile regression estimates.

RESULTS

Logistic Quantile Regression Estimates of Spotted Owl Fledgling Production

We found that 34% of the total variance in the cdf of fledgling counts across 1992–2010 could be explained by prior production and annual variation (Table 1). More productive years were preceded by one or more much less productive years, with the possible exception of 2009 and 2010 when both years were similarly productive (Figure 2).

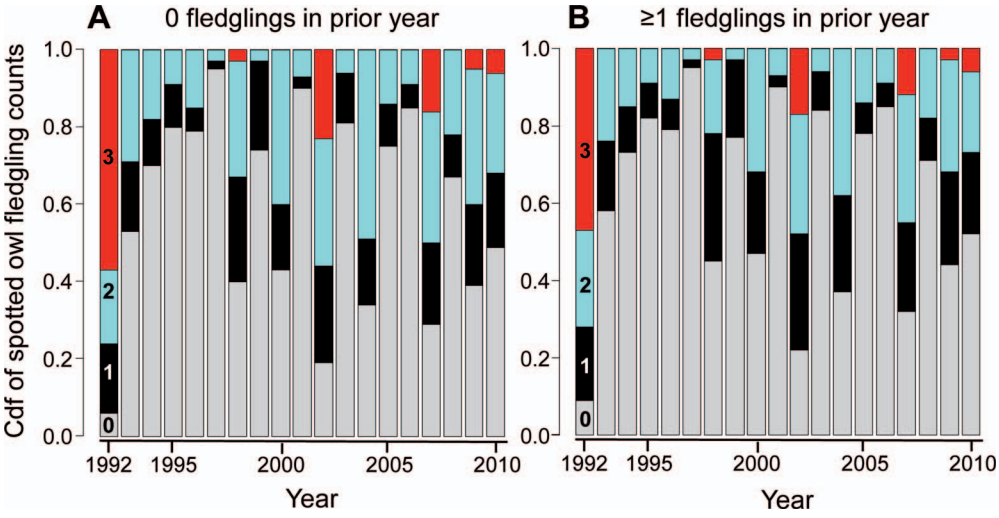


FIGURE 2. Annual variation in cumulative distributions of California Spotted Owl fledgling counts (0–3) estimated with logistic quantile count model ($n = 707$) for territories with prior year counts of 0 or 1 fledglings in 1992 to 2010. The other 2 categories for counts 2 years prior follow a similar pattern. Red portions are for counts of 3, cyan for counts of 2, black for counts of 1, and gray for counts of 0 fledglings.

The proportion of territories producing 2 and 3 fledglings (0.03–0.76) varied annually much more than the proportion (0.02–0.33) producing a single fledgling (Figure 2). Thus, when more territories produced fledglings this generally reflected more territories producing 2 or 3 fledglings. The effect of prior production only contributed about 17% to the temporal variation and was a weak, consistent negative effect when prior production occurred in the previous year. Estimates in the logit scale varied from 0.20 to 0.40 for most τ and can be interpreted as odds ratios, exp(0.20) = 0.819 and exp(0.40) = 0.670. This indicates fledgling production in the continuous logit scale when there was prior production was 82% to 67% of production when there was no production in the prior year. Returning these estimates to discrete cdfs indicated 0% to 5% fewer territories produced any fledglings but 0% to 11% fewer territories produced 2 or 3 fledglings when there was fledgling production in the prior year (Figure 2).

Our models of fledgling counts using climate predictors found little support for including winter minimum temperature or late nesting precipitation based on average ΔAIC. There was some support for including late nesting minimum temperature, but because it was strongly correlated ($r = -0.71$) with early nesting minimum temperature, we included only the latter variable. There was more support for models with linear and quadratic terms for predictor variables at higher quantiles associated with counts of 2 and 3 fledglings (average ΔAIC 2). A linear model was better supported only at lower quantiles associated with counts of a single fledgling. The climate model we selected included linear and quadratic terms for winter precipitation, for early nesting precipitation and minimum temperature, and for precipitation in the previous growing season ($n = 707$). This model was re-estimated with the reduced set of observations ($n = 639$) that only included parents of known age class. This model explained an average of 10% of the variance in the cdf of counts attributed to annual variation (Table 1). Adding the age class of male and female parents to the model was strongly supported (Figure 3) and increased the amount of annual variation explained to 18% (Table 1).

The models we then considered that included variation in landscape habitat structure among owl territories indicated that percentages of the forest cover in trees 25 m in height in 202 or 121 ha (500 or 300 ac, respectively) areas surrounding nesting sites were the best candidate predictors. Although both spatial scales were supported, we focused our modeling on the percentage of area with trees 25 m in height in the 202 ha surrounding territories. Adding the linear and quadratic terms for this landscape habitat predictor to the model was well supported across most quantiles (Figure 3). This full model with prior production, climate variables, age class of parents, and landscape habitat accounted for 7% of the total variance in the cdf of fledgling production (Table 1). There was some support for including the quadratic effect of average elevation across the 121 ha core area as evidenced by average ΔAIC of 2 7 at quantiles >0.88. However, including elevation in the model only reduced the modeled variance of the cdf by 0.003. We, thus, chose not to include elevation in our final model. Other habitat predictors considered were not included in the final model either because they were strongly correlated ($r = 0.96$ for average height of trees and $r = 0.87$ for standard deviation

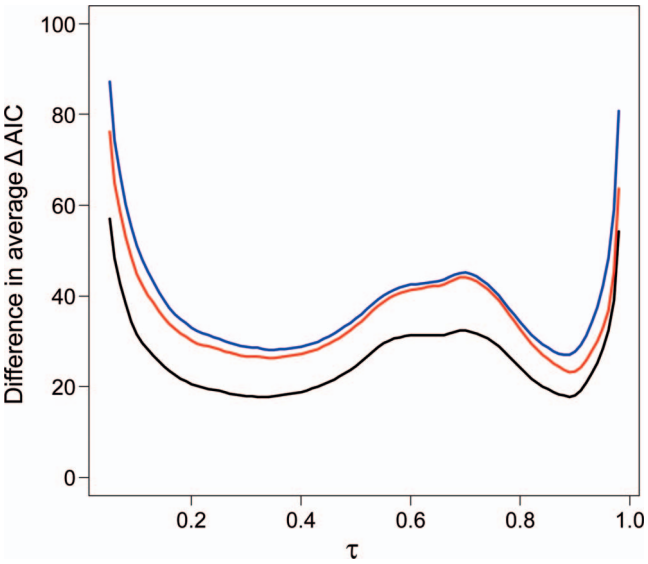


FIGURE 3. Average differences in AIC (ΔAIC) by τ (0.05–0.98) for logistic quantile models of California Spotted Owl fledgling counts ($n = 639$) as (1) quadratic functions of early nesting precipitation, early nesting minimum temperature, winter precipitation, previous growing season precipitation, and prior production (black, number of parameters [p] = 12); (2) the same predictors as in (1) plus age class of male and female parents (red, $p = 14$); and (3) the same predictors as in (2) plus a quadratic function of percentage of area in trees 25 m height (blue, p =16). ΔAIC were computed relative to the base model of prior production that had just an intercept and the lagged autoregressive terms ($p = 4$).

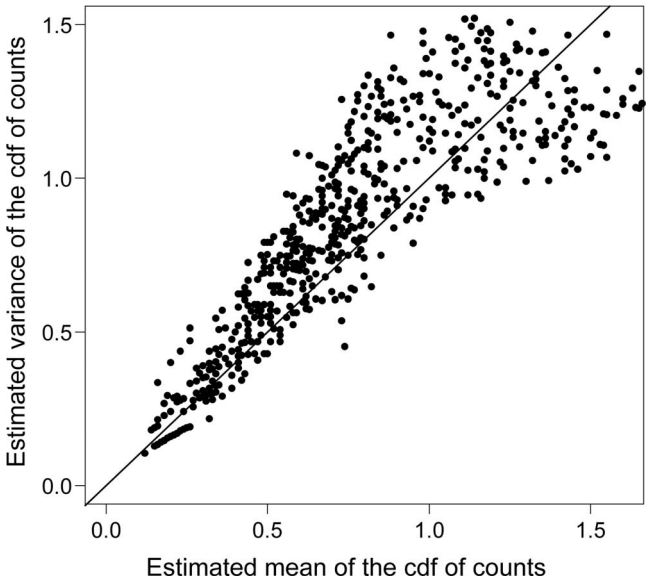


FIGURE 4. Means and variances of the estimated cumulative distribution functions (cdfs) for each observation ($n = 639$) from the logistic quantile regression model that included prior production, climate, parent age, and landscape habitat predictors (Table 1). Solid line is 1:1 relationship.

of height of trees) with percentage of area in trees 25 m height or because they had too restricted a range of values (80% of territories had 80 100% total canopy cover of trees and 0% percentage area in trees 50 m height) with insufficient support for inclusion.

It is important to recognize that the variances of the cdf attributed to a model were averages across all the observations and that there was considerable heterogeneity in variances of the cdf conditional on the predictors. For example, the average estimated variance of the cdf $\hat{\sigma}^2 =$ 0.837 in our full model (Table 1) but the variances of the cdfs ($\hat{\sigma}_i^2$) and variation among them across the predictor space increased with the mean of the cdf of counts from a low near 0 to a range of 0.8–1.5 as means of the cdf approached and exceeded 1.0 (Figure 4). The variance–mean relationships associated with the cdfs indicated that as mean counts increased from 0.5 to 1.0 the proportion of observations with variances greater than the mean and the size of those variances increased, but as means increased from 1.25 to 1.5 the variances decreased to less than the means (Figure 4). This variance pattern would be difficult to model with parametric count distributions and certainly was inconsistent with equality of means and variances associated with a Poisson distribution.

Quantile estimates and 95% confidence intervals for all parameters by τ for the full model are in Figure 5. Note that confidence interval estimates often increase greatly at the more extreme τ. Estimates for the autoregressive terms associated with prior production indicate that the negative effect of fledgling production in the prior year (Figure 5B) was considerably stronger than in the model for temporal variation. Estimates range from 0.80 to 1.20 from lower to higher quantiles. The odds ratios interpretation of these estimates indicate 45% [exp(0.80)] to 30% [exp(1.20)] reduction of fledglings from lower to higher quantiles in the continuous logit scale associated with producing any fledglings in the prior year. The partial effect of fledgling production in the prior year in terms of discrete counts is available by comparing the cdf for adults with 0 fledglings in the prior year (Figure 6F), which is the cdf for the intercept term, with the cdf for adults with 1 fledgling in the prior year (Figure 6F). These cdfs are conditional on the climate and habitat predictors at their mean values. Comparing these two cdfs indicates that the partial effects were to increase the proportion of territories producing zero fledglings from 40% to 58% when there was production in the previous year, with a concomitant small decrease in the proportion producing a single fledgling (25% to 20%) and less than a third as many (16% to 5%) producing 3 fledglings (Figure 6F).

There was little effect associated with fledgling production 2 years prior as all 95% confidence intervals strongly

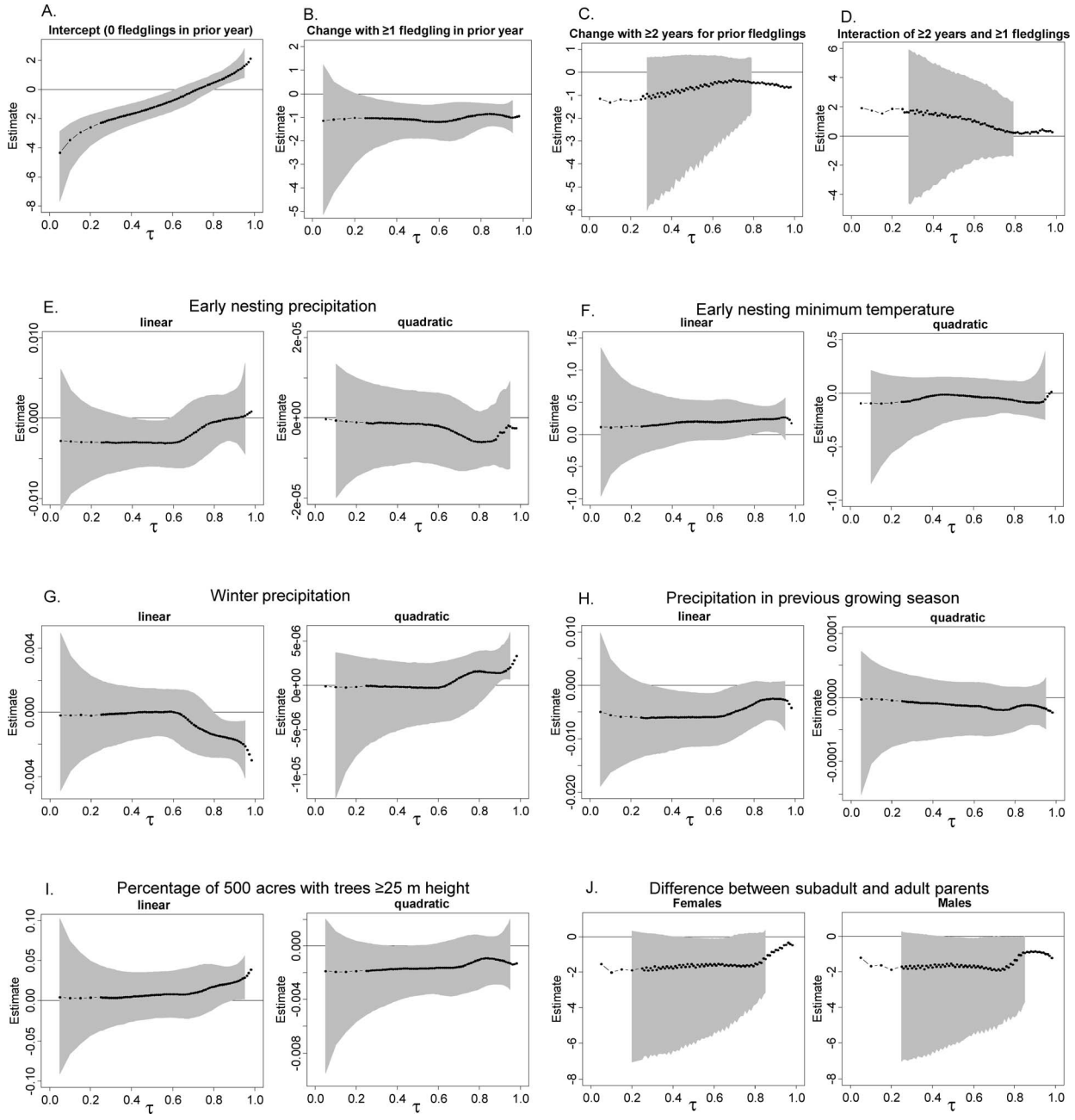


FIGURE 5. Averaged ($m = 500$) parameter estimates and 95% confidence intervals (gray band) by $\tau = (0.05–0.98)$ in the continuous logit scale for the logistic quantile regression model of California Spotted Owl fledgling counts ($n = 639$) that included an (**A**) intercept; (**B–D**) three terms for the lagged 1st-order autoregressive effect of prior production; linear and quadratic terms for partial effects of (**E**) early nesting precipitation, (**F**) early nesting minimum temperature, (**G**) winter precipitation, (**H**) precipitation in previous growing season, (**I**) percent of 500 ac around territory with trees 25 m height; and (**J**) age class of male and female parents. Confidence intervals are not shown for all quantiles estimated because they became too wide to graph at more extreme τ.

overlapped zero (Figure 5C and 5D). To investigate whether inclusion of the 12% of observations where prior production at a territory was observed 2 years previously impacted our estimates of partial effects for climate, parent age, and landscape predictors, we estimated the same model but without the terms for prior production 2 years previously and related observations ($n = 571$). Estimates of partial effects for all predictor variables were very similar to those estimated in the model where we included prior production 2 years previously, with slightly larger confidence interval widths at more extreme quantiles for some predictors.

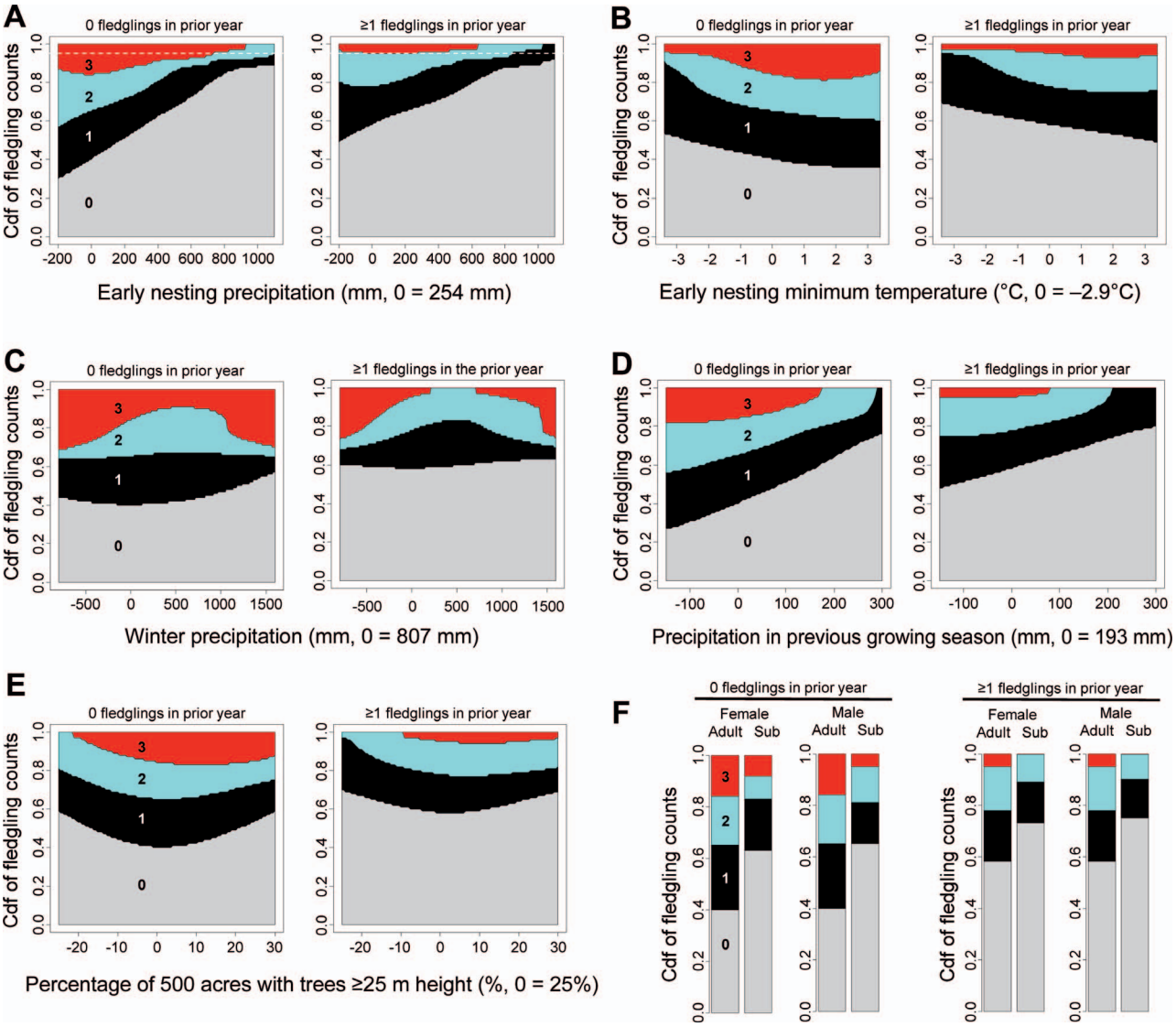


FIGURE 6. Estimated cumulative distribution functions (cdf) for California Spotted Owl fledgling counts for partial effects of (**A**) early nesting precipitation, (**B**) early nesting minimum temperature, (**C**) winter precipitation, (**D**) precipitation in previous growing season, (**E**) percent of 500 ac around territory with trees 25 m height, and (**F**) age class of male and female parents in logistic quantile model that included prior fledgling production, climate, parent age, and landscape habitat ($n = 639$). Cdfs were estimated by holding other continuous predictors at their mean values (all centered on zero) for adult parents, with estimates for territories with 0 (left panel) or 1 (right panel) fledglings in the prior year. Red portions of cdf are for counts of 3, cyan for counts of 2, black for counts of 1, and gray for counts of 0 fledglings. Dashed white lines in (**A**) are 0.95 quantiles of partial cdf that can be interpreted as a prediction interval.

The estimated partial effects of the climate predictors in our full model were consistent with hypothesized relationships gleaned from previous literature for early nesting precipitation and minimum temperatures (Figures 5E and 5F). Productivity decreased with increasing early nesting precipitation (Figure 6A) and increased with increasing minimum temperature in the early nesting period (Figure 6B), where the effect of the quadratic terms was to capture some nonlinearity in those effects. The partial effect of winter precipitation was more complicated with an increase in territories producing 3 rather than 2 fledglings as winter precipitation decreased below approximately 1,200 mm and increased above approximately 1,800 mm when no fledglings were produced in the previous year, with minimal effect on the proportion of territories producing a single fledgling (Figure 6C). When fledglings had been produced in the previous year, the effect of winter precipitation was similar except the trade-off in

increasing proportion of territories with 3 fledglings was for a decrease in the proportion producing both one and two fledglings. The partial effect of precipitation in the previous growing season was counter to our expectation in that there was decreased fledgling production with increasing precipitation in the previous growing season (Figure 6D).

The partial effects of age class (adult vs. subadult) for both female and male parents (Figure 5J) indicated that most of the effect could be described as a reduction in the proportion of territories producing 2 or 3 fledglings with a concomitant increase in the proportion of territories producing no fledglings for subadults. There was 15% to 23% less production with subadult females and 17% to 25% less production with subadult males (Figure 6F), where smaller differences were associated with territories that had production in the prior year. There was less difference in the proportion of territories producing a single fledgling associated with age class of parents, a decrease of 4% to 9% for subadults.

The spatial variation in fledgling production explained by landscape habitat structure was small. The partial effects of percentage of the forest cover with trees 25 m in height surrounding territories indicated decreased fledgling production for all counts as the percentage decreased below 25% or increased above 35% (Figure 6E). The effect of fledgling production in the previous year was to primarily lower the overall production of fledglings while maintaining a similar pattern of changing production with percentage of the forest cover with trees 25 m in height.

Prediction intervals for a new observation of fledgling counts on a territory for a single predictor varying in value while holding the others constant at their mean values for adult parents are equivalent to placing a horizontal line on the partial effects graphs (Figure 6) corresponding to a selected quantile of the cdf associated with a desired prediction interval level. For example, the 0.95 quantile estimates for early nesting season precipitation (Figure 6A) is the upper endpoint for a two-tailed 90% or upper one-tailed 95% prediction interval for a single new outcome of fledgling counts associated with the domain of precipitation values. This horizontal line at the 0.95 quantile of the cdf when there is no production of fledglings in the previous year indicates a prediction interval that includes {0, 1, 2, 3} fledglings for early nesting season precipitation until precipitation exceeds 700 mm above the average (254 mm) where the prediction interval changes to {0, 1, 2} fledglings (Figure 6A). When 1 fledglings were produced in the prior year, the corresponding prediction intervals would be {0, 1, 2} fledglings, changing to {0, 1} fledglings when early nesting season precipitation exceeds 850 mm above the average. Similar intervals can be obtained for the other predictor variables or at any specified combination of predictor values.

Cumulative Ordinal Logistic Regression Estimates of Spotted Owl Fledgling Production

The cumulative ordinal logistic regression model for prior production, climate, age class of parents, and landscape habitat had estimates (Appendix Table 2) and partial effects (Appendix Figure 7) similar to the logistic quantile regression model (Figure 6). However, the partial effects of the cumulative ordinal logistic regression model tended to be smoother and more symmetric, principally because the logistic form in this model not only constrains estimates to remain within the bounds of 0–3 but also links the quantiles of the conditional cdfs in a parametric logistic form. In contrast, the logistic quantile regression model has no parametric form linking the estimates among the quantiles. The flat, step function appearance of the logistic quantile regression partial effects for continuous predictors (Figure 6) is partly due to the discrete nature of the quantiles for counts and partly an artifact of only estimating the cdfs by increments of 0.01 quantiles. The cumulative ordinal logistic regression model had less flexibility to estimate nonlinear partial effects of the cdf for winter precipitation (Appendix Figure 7C) compared to that estimated by the logistic quantile regression (Figure 6C). This lack of flexibility also was responsible for crossing of some estimates at extreme regions of the predictor space for winter precipitation and landscape habitat (Appendix Figure 7C, E). Conceptually, we could have computed an average variance of the cdfs associated with cumulative ordinal logistic regression models but did not. Crossing of the estimates in some regions of predictor space created issues for this computation.

OLS Regression Estimates of Mean Spotted Owl Fledgling Production

The OLS regression estimates (Appendix Table 3) for the full model including prior production, climate, age class of parents, and landscape habitat were in direction and magnitude similar to those from the logistic quantile regression model although not directly comparable because they were for a model linear in the fledgling counts. The linearity of the OLS regression model in the fledgling count space resulted in estimates of mean counts that tended to underestimate the comparable means of the cdfs estimated by logistic quantile regression when they were <0.5 or >1 (Appendix Figure 8). Estimates of mean counts were more comparable when means were 0.5–1.0. Nineteen of the OLS estimates of mean counts were <0 and, thus, outside the bounds of 0 3 (Appendix Figure 8). Estimates of 95% prediction intervals for a new observation based on the OLS model had lower bounds <0 at predictor values associated with all 639 observations and

upper bounds >3 for predictor values associated with 60 observations. This is a consequence of the unbounded nature of the assumed normal distribution, linearity, and failing to account for heterogeneous variances associated with these bounded counts (Figure 4).

Average variances of the temporal models for OLS regression and logistic quantile regression were similar (Table 1). This is not unexpected because this model only includes categorical predictors where the logit model form is of no real advantage in constraining estimates to remain within the 0 3 bounds (Bottai et al. 2010). However, when the continuous climate and landscape habitat predictors are included in models, the OLS estimates of variance tend to be smaller than those for the empirical cdfs estimated by logistic quantile regression (Table 1). The difference in absolute magnitude of the variance estimates for OLS estimates compared to logistic quantile regression estimates may be small, but when expressed as variance components, 2 to 3 times the percentage of the annual variance is attributed to the OLS regression estimates compared to the logistic quantile regression estimates (Table 1).

DISCUSSION

Our logistic quantile regression model of the discrete cdf of small bounded counts provided a more nuanced interpretation of Spotted Owl fledgling production than was provided by OLS regression estimates of mean counts. The logistic quantile regression model by definition provided prediction intervals for an individual outcome at any predictor values that are constrained between 0 and 3. This is not the case for OLS regression. Although not done here, it is possible to estimate discrete confidence interval endpoints for an interval of quantiles at selected values of the predictors to form tolerance intervals for specified proportions of the populations (equivalent to prediction intervals for any number of new outcomes) as done for quantile count models (Cade and Dong 2008). The logistic quantile regression model provides a very flexible approach to accommodate heterogeneity and skewness in statistical models for small bounded counts that is not readily accomplished with mean regression models. Our logistic quantile regression model estimates emphasized that both the effects of prior fledgling production and subadult parents had more negative impacts on the production of 2 or 3 Spotted Owl fledglings than on the production of single fledglings. This would not be obvious from estimates of mean counts.

Our comparisons between variance estimates from the logistic quantile regression and OLS mean regression models indicate that the latter is likely to provide underestimates of variance attributed to models with continuous predictor variables such as the climate and habitat predictor variables used in our Spotted Owl models. These underestimates of variance although small can translate into much larger estimates of proportions of variance explained relative to annual or total variance in fledgling production (i.e. inflated variance component estimates). We suspect this inflation of explained variances has occurred in many previous Spotted Owl analyses that relied on mean regression models with a homogeneous, normal error distribution. Many people analyzing Spotted Owl fledgling production actually used the recommendations of McDonald and White (2010) to justify using more complicated mixed-effects mean regression models (e.g., Blakesley et al. 2010, Dugger et al. 2016). There is little reason to think that these more complicated implementations of the normal distribution with homogeneous variances will appropriately account for variances in small bounded counts. Furthermore, the simulations by McDonald and White (2010) only provided support for using mean regression models to estimate rates of change (the regression coefficients) in small bounded counts associated with predictor variables (e.g., estimated trends over time) and not for the use of variance estimates from these models.

The cumulative ordinal logistic regression model without a proportional odds assumption was a reasonable alternative statistical model for small bounded counts. However, the cumulative ordinal logistic model had greater difficulty accommodating more complex models that included quadratic terms compared to our logistic quantile regression models. The increased propensity for crossing of estimates from the cumulative ordinal logistic regression model without a proportional odds assumption was likely due to its categorical representation of counts and sparsity of those categories for more extreme values of predictor variables. This was not an issue with our logistic quantile regression model of small bounded counts although linear quantile regression is not immune to issues with crossing of quantile estimates. The qualitative interpretations of partial effects for our Spotted Owl models were similar between the 2 approaches. The cumulative ordinal logistic regression model does not require as much post-processing of estimates to provide estimated changes in the partial cdfs of predictors as is required for logistic quantile regression. However, effective computer code to facilitate post-processing of logistic quantile regression estimates could be developed easily to make the procedure less time-intensive to implement.

Our logistic quantile regression model of California Spotted Owl fledgling production on the Lassen National Forest in California found that a relatively small percentage (18%) of the annual variation in fledgling production across 20 years could be attributed to changes in climate and age class of parents. The direction of effects for most of our climate predictors and effects of parent age class were similar to those from other investigations of

Northern and California Spotted Owls. The percentage of variance in fledgling production accounted for by our models tended to be on the low end of that found in other studies (Dugger et al. 2005; Blakesley et al. 2010; Glenn et al. 2010, 2011), although their estimates are likely inflated due to their use of mean regression models. Our estimated cdfs of annual variation in fledgling production showed that consecutive years of high productivity are rare, consistent with patterns found in other Spotted Owl studies. It has been referred to as an even–odd year effect where higher production occurs in even numbered years (Dugger et al. 2005; Glenn et al. 2010, 2011; Stoelting et al. 2015; Dugger et al. 2016). Our estimates for California Spotted Owl productivity in the Lassen National Forest are not consistent with an even–odd year effect. We found that more productive years only occurred after one or more unproductive years, but more productive years were not always even-numbered years (e.g., 2007 and 2009) and many even-numbered years (1994, 1996, 2006, and 2008) had lower fledgling productivity. Thus, the fledgling production cycle at Lassen National Forest is not reasonably characterized as biennial.

Similar to Stoelting et al. (2015), we found a strong, consistent statistical effect of prior production on California Spotted Owl fledgling production but where the effect on fledgling production over time was not substantial, with up to 5% fewer territories producing any fledglings and up to 11% fewer territories producing multiple fledglings. This cost of reproduction was associated with $<$20% of the temporal variation in California Spotted Owl fledgling production at the Lassen National Forest. It remains unclear whether this cost of reproduction is tied more directly to a physiological cost to individual owls (e.g., through delayed molting and production of new feathers), to individual territories (e.g., through reduction in the prey base available to feed young), or to both. There clearly was much more annual variation in fledgling production on the Lassen National Forest population than was modeled by prior production, climate, and age class of parents. The spatial variation we incorporated in our model associated with the percentage of large, mature trees in landscapes surrounding territories only explained a small proportion of additional variance in fledgling production. This is perhaps not too surprising as it might reasonably be expected that the selection of nesting territories within a declining population of Spotted Owls (Connor et al. 2013) would result in little among-territory variation in habitat structure or composition as only the higher-quality habitats were occupied.

We included a climate predictor, precipitation in the previous growing season, that we thought would be related to increased prey abundance through increased vegetation used by small rodents, but the relationship with California Spotted Owl fledgling production was counter to our expectation. However, we believe that prey availability for nesting California Spotted Owls likely is a key driver of fledgling production as has been suggested for Northern Spotted Owls (Rosenberg et al. 2003). It is possible that important prey items for Spotted Owls such as deer mice (*Peromyscus maniculatus*) may respond to food sources that are not directly tied to prior year precipitation (e.g., conifer seed production) and that there are complicated interactions between prey abundance and weather conditions that ultimately impact prey availability for nesting Spotted Owls (Rosenberg et al. 2003). A further complication is that epizootic diseases such as plague that may impact rodent prey populations also likely involve complicated interactions of prey populations and weather. Plague appears to be common in small rodents in the Lassen National Forest and surrounding area in California (Smith et al. 2010). Additional research on prey populations and Spotted Owl fecundity is clearly warranted.

Our logistic quantile regression model for bounded counts can be readily adapted for other applications in ecology involving responses that are discrete random variables on a restricted range. Bounded counts are a common characteristic of the reproductive output per parents or territory for many avian species and also for many mammals (e.g., Rosenberry et al. 2011, Peacock et al. 2013). Our model for Spotted Owl fledgling production was of moderate complexity, involving polynomial terms on predictor variables and 1st-order lagged effects. Both simpler and more complex models can be accommodated because the linear quantile regression estimator can include any parameterization of predictor variables that might be used in other linear models (Koenker 2005). Although the logistic quantile regression model for bounded counts is appropriate for any upper bound on the counts, it is most likely to provide improved estimates over the more conventional exponential model form (log transformation of counts) used with counts (Cade and Dong 2008) when there are many observations that have counts near the upper bound. When there are few observations near the upper bound of the counts, the logistic quantile regression estimates will rarely achieve an asymptote near the upper bound and a simpler, exponential model form may be adequate. When a bounded response distribution includes a sufficient range of values with minimal tied values (e.g., 0 100% canopy cover of plants), then the logistic quantile regression procedure for continuous responses of Bottai et al. (2010) can be used directly without the random jittering simulations and ceiling functions required for discrete random variables.

ACKNOWLEDGMENTS

We thank G. Jehle for providing support in data preparation and J. Blakesley, M. Bottai, and T. McDonald for reviewing

drafts of the manuscript. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Funding statement: Funding for the demographic field studies of California Spotted Owls was provided by the USDA Forest Service. Funding for the analyses, interpretation, and write up of this publication was provided by the USDA Forest Service and USDI Geological Survey.

Ethics statement: All Spotted Owl data were collected following protocols under animal care and use permits overseen by the USDA Forest Service. These protocols were followed and all relevant state and federal permits were acquired and maintained for the Lassen National Forest.

Author contributions: (1) BRN, RDS, and JJK formulated the questions; (2) BRN and JJK collected data and supervised field research; (3) BSC developed the statistical methods and analyzed the data; (4) BSC, BRN, and RDS wrote the paper; and (5) JJK contributed substantial materials, resources, and funding.

Data accessibility: The data file with Spotted Owl fledgling counts and predictor variables is available on ScienceBase (https://www.sciencebase.gov/catalog/). https://doi.org/10.5066/F7DR2SZR

LITERATURE CITED

Agresti, A. (2013). Categorical Data Analysis, 3rd edition. John Wiley & Sons, Hoboken, NJ, USA.

Anthony, R. G., E. D. Forsman, A. B. Franklin, D. R. Anderson, K. P. Burnham, G. C. White, C. J. Schwarz, J. D. Nichols, J. E. Hines, G. S. Olson, S. H. Ackers, et al. (2006). Status and trends in demography of Northern Spotted Owls, 1985 2003. Wildlife Monographs 163:1–48.

Bingham, B. B., and B. R. Noon. (1997). Mitigation of habitat "take": Application to habitat conservation planning. Conservation Biology 11:127–139.

Blakesley, J. A., D. R. Anderson, and B. R. Noon. (2006). Breeding dispersal in the California Spotted Owl. The Condor 108:71–81.

Blakesley, J. A., B. R. Noon, and D. R. Anderson. (2005). Site occupancy, apparent survival, and reproduction of California Spotted Owls in relation to forest stand characteristics. Journal of Wildlife Management 69:1554–1564.

Blakesley, J. A., M. E. Seamans, M. M. Conner, A. B. Franklin, G. C. White, R. J. Gutiérrez, J. E. Hines, J. D. Nichols, T. E. Munton, D. W. H. Shaw, J. J. Keane, et al. (2010). Population dynamics of Spotted Owls in the Sierra Nevada, California. Wildlife Monographs 174:1–36.

Bottai, M., B. Cai, and R. E. McKeown. (2010). Logistic quantile regression for bounded outcomes. Statistics in Medicine 29: 309–317.

Cade, B. S., and B. R. Noon. (2003). A gentle introduction to quantile regression for ecologists. Frontiers in Ecology and the Environment 1:412–420.

Cade, B. S., B. R. Noon, and C. H. Flather. (2005). Quantile regression reveals hidden bias and uncertainty in habitat models. Ecology 86:786–800.

Cade, B. S., J. D. Richards, and P. W. Mielke, Jr. (2006). Rank score and permutation testing alternatives for regression quantile estimates. Journal of Statistical Computation and Simulation 76:331–355.

Cade, B. S., and Q. Dong. (2008). A quantile count model of water depth constraints on Cape Sable Seaside Sparrows. Journal of Animal Ecology 77:47–56.

Connor, M. M., J. J. Keane, C. V. Gallagher, G. Jehle, T. E. Munton, P. A. Shaklee, and R. A. Gerrard. (2013). Realized population change for long-term monitoring: California Spotted Owl case study. Journal of Wildlife Management 77:1449–1458.

Dugger, K. M., E. D. Forsman, A. B. Franklin, R. J. Davis, G. C. White, C. J. Schwarz, K. P. Burnham, J. D. Nichols, J. E. Hines, C. B. Yackulic, P. F. Doherty, Jr., et al. (2016). The effect of habitat, climate, and Barred Owls on long-term demography of Northern Spotted Owls. The Condor: Ornithological Applications 118:57–116.

Dugger, K. M., F. Wagner, R. G. Anthony, and G. S. Olson. (2005). The relationship between habitat characteristics and demographic performance of Northern Spotted Owls in southern Oregon. The Condor 107:863–878.

Forsman, E. D., R. G. Anthony, K. M. Dugger, E. M. Glenn, A. B. Franklin, G. C. White, C. J. Schwarz, K. P. Burnham, D. R. Anderson, J. D. Nichols, J. E. Hines, et al. (2011). Population demography of Northern Spotted Owls. Studies in Avian Biology, no. 40.

Franklin, A. B., D. R. Anderson, R. J. Gutiérrez, and K. P. Burnham. (2000). Climate, habitat quality, and fitness in Northern Spotted Owl populations in northwestern California. Ecological Monographs 70:539–590.

Franklin, A. B., R. J. Gutiérrez, J. D. Nichols, M. E. Seamans, G. C. White, G. S. Zimmerman, J. E. Hines, T. E. Munton, W. S. LaHaye, J. A. Blakesley, C. N. Steger, et al. (2004). Population dynamics of the California Spotted Owl (*Strix occidentalis occidentalis*): A meta-analysis. Ornithological Monographs, no. 54.

Glenn, E. M., R. G. Anthony, and E. D. Forsman. (2010). Population trends in Northern Spotted Owls: Associations with climate in the Pacific Northwest. Biological Conservation 143:2543–2552.

Glenn, E. M., R. G. Anthony, E. D. Forsman, and G. S. Olson. (2011). Reproduction of Northern Spotted Owls: The role of local weather and regional climate. Journal of Wildlife Management 75:1279–1294.

Ives, A. R. (2015). For testing the significance of regression coefficients, go ahead and log-transform count data. Methods in Ecology & Evolution 6:828–835.

Jetz, W., C. H. Sekercioglu, and K. Böhning-Gaese. (2008). The worldwide variation in avian clutch size across species and space. PLOS-Biology 6:2650–2657.

Koenker, R. (2005). Quantile Regression. Econometric Society Monographs No. 38, Cambridge University Press, New York, NY, USA.

Koenker, R., and J. A. F. Machado. (1999). Goodness of fit and related inference processes for quantile regression. Journal of the American Statistical Association 94:1296–1310.

Lancaster, J., and L. R. Belyea. (2006). Defining the limits to local density: Alternative views of abundance–environment relationships. Freshwater Biology 51:783–796.

LANDFIRE. (2008). Existing Vegetation Type Layer, *LANDFIRE* 1.1.0, U.S. Department of the Interior, Geological Survey. http://landfire.cr.usgs.gov/viewer/

Lint, J., B. R. Noon, R. Anthony, E. Forsman, M. Raphael, M. Collopy, and E. Starkey. (1999). Northern Spotted Owl effectiveness monitoring plan for the Northwest Forest Plan. USDA Forest Service PNW-GTR-440.

Liu, I., and A. Agresti. (2005). An analysis of ordered categorical data: An overview and a survey of recent developments. Test 14:1–73.

Machado, J. A. F., and J. M. C. Santos Silva. (2005). Quantiles for counts. Journal of the American Statistical Association 100: 1226–1237.

McDonald, T. L., and G. C. White. (2010). A comparison of regression models for small counts. Journal of Wildlife Management 74:514–521.

Noon, B. R., and A. B. Franklin. (2002). Scientific research and the Spotted Owl (*Strix occidentalis*): Opportunities for major contributions to avian population ecology. The Auk 119:311–320.

North, M., G. Steger, R. Denton, G. Eberlein, T. Munton, and K. Johnson. (2000). Association of weather and nest-site structure with reproductive success in California Spotted Owls. Journal of Wildlife Management 64:797–807.

Olson, G. S., R. G. Anthony, E. D. Forsman, S. H. Ackers, P. J. Loschl, R. A. Reid, K. M. Dugger, E. M. Glenn, and W. J. Ripple. (2005). Modeling of site occupancy dynamics for Northern Spotted Owls with emphasis on the effects of Barred Owls. Journal of Wildlife Management 69:918–932.

Peacock, E., M. K. Taylor, J. Laake, and I. Stirling. (2013). Population ecology of polar bears in Davis Strait, Canada and Greenland. Journal of Wildlife Management 77:463–476.

PRISM Climate Group. (2004). Oregon State University, Corvallis, Oregon, USA. http://prism.oregonstate.edu

Rosenberg, D. K., K. A. Swindle, and R. G. Anthony. (2003). Influence of prey abundance on Northern Spotted Owl reproductive success in western Oregon. Canadian Journal of Zoology 81:1715–1725.

Rosenberry, C. S., A. S. Norton, D. R. Diffenbach, J. T. Fleegle, and B. D. Wallingford. (2011). White-tailed deer age ratios as herd management and predator impact measures in Pennsylvania. Wildlife Society Bulletin 35:461–468.

Seamans, M. E., R. J. Gutiérrez, C. A. Moen, and M. Z. Peery. (2001). Spotted Owl demography in the central Sierra Nevada. Journal of Wildlife Management 65:425–431.

Simkin, S. M., E. B. Allen, W. D. Bowman, C. M. Clark, J. Belnap, M. L. Brooks, B. S. Cade, S. L. Collins, L. H. Geiser, F. S. Gilliam, S. E. Jovan, et al. 2016. Conditional vulnerability of plant diversity to atmospheric nitrogen deposition across the United States. Proceedings of the National Academy of Sciences USA 113: 4086–4091.

Smith, C. R., J. R. Tucker, B. A. Wilson, and J. R. Clover. (2010). Plague studies in California: A review of long-term disease activity, flea–host relationships and plague ecology in the coniferous forests of the Southern Cascades and northern Sierra Nevada mountains. Journal of Vector Ecology 35:1–12.

Stoelting, R. E., R. J. Gutiérrez, W. L. Kendall, and M. Z. Peery. (2015). Life-history tradeoffs and reproductive cycles in Spotted Owls. The Auk: Ornithological Advances 132:46–64.

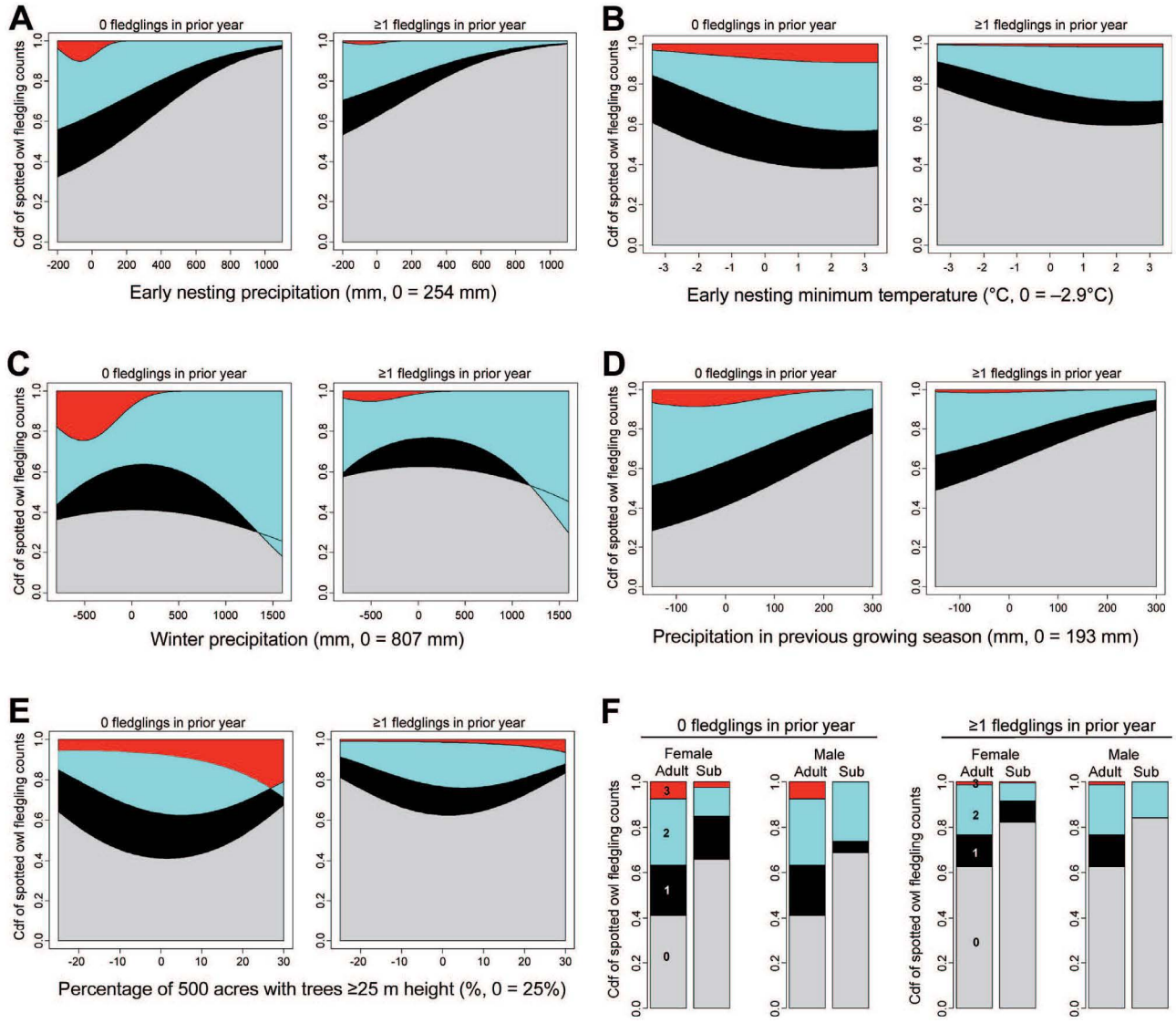
Wei, Y., A. Pere, R. Koenker, and X. He. (2006). Quantile regression methods for reference growth charts. Statistics in Medicine 25:1369–1382.

Yee, T. W. (2010). VGLMs and VGAMs: An overview for applications in fisheries research. Fisheries Research 101: 116–126.

Yu, K., and R. A. Moyeed. (2001). Bayesian quantile regression. Statistics & Probability Letters 54:437–447.

APPENDIX TABLE 2. Parameter estimates for cumulative ordinal logistic regression models without proportional odds assumptions for California Spotted Owl fledgling counts (0–3) using the same predictor terms as in the full logistic quantile regression model of fledgling counts ($n = 639$); autoregressive prior production, quadratic functions of climate, age of parents, and quadratic function of landscape habitat surrounding territories.

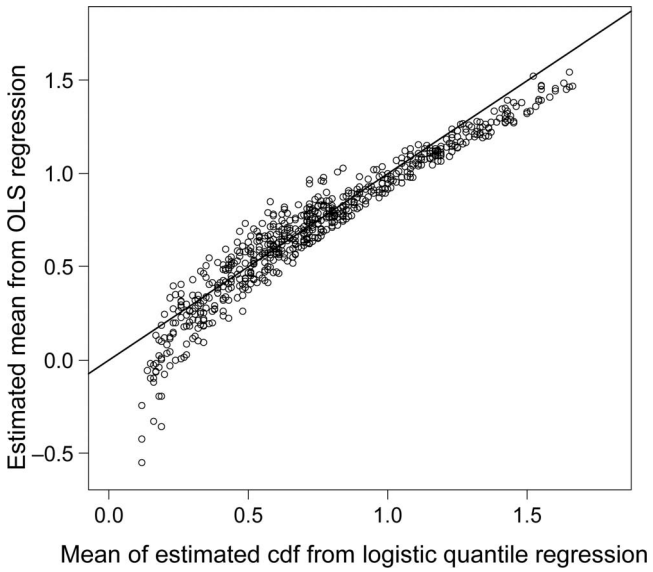
Prob(*Y* 1)		
Predictor	Estimate	95% confidence interval
Intercept	0.358	[0.014, 0.735]
1 fledglings in prior year	0.870	[1.257, 0.489]
2 years for prior production	0.557	[1.313, 0.166]
Interaction of above 2 terms	1.313	[0.112, 2.520]
Early nesting precipitation - linear	0.002	[0.004, 0.0004]
Early nesting precipitation - quadratic	1.1e 06	[4.9e-06, 2.4e 06]
Early nesting min. temperature - linear	0.128	[0.036, 0.294]
Early nesting min. temperature - quadratic	0.032	[0.139, 0.076]
Winter precipitation - linear	2.4e 05	[8.2e-04,7.7e 04]
Winter precipitation - quadratic	2.9e 07	[1.06e-06, 1.4e 06]
Precipitation in previous growing season - linear	0.004	[0.006, 0.002]
Precipitation in previous growing season - quadratic	3.7e 06	[2.1e-05, 1.3e 05]
Percentage of 500 acres with trees 25 m height - linear	0.004	[0.011, 0.019]
Percentage of 500 acres with trees 25 m height - quadratic	0.001	[0.002, 0.0004]
Difference in subadult from adult female parents	1.009	[1.795, 0.294]
Difference in subadult from adult male parents	1.141	[1.950, 0.413]
Prob(*Y* 2)		
Predictor	Estimate	95% confidence interval
Intercept	0.544	[0.950, 0.142]
1 fledglings in prior year	0.642	[1.072, 0.223]
2 years for prior production	0.124	[0.935, 0.631]
Interaction of above 2 terms	0.326	[1.995, 1.110]
Early nesting precipitation - linear	0.002	[0.004, 0.0001]
Early nesting precipitation - quadratic	1.1e 06	[5.6e-06, 2.7e 06]
Early nesting min. temperature - linear	0.206	[0.023, 0.396]
Early nesting min. temperature - quadratic	0.039	[0.161, 0.079]
Winter precipitation - linear	2.4e 05	[0.001, 0.0006]
Winter precipitation - quadratic	9.5e 07	[4.2e-07, 2.2e 06]
Precipitation in previous growing season - linear	0.004	[0.007, 0.002]
Precipitation in previous growing season - quadratic	5.4e 06	[2.6e-05, 1.3e 05]
Percentage of 500 acres with trees 25 m height - linear	0.014	[0.003, 0.030]
Percentage of 500 acres with trees 25 m height - quadratic	0.001	[0.002, 0.0003]
Difference in subadult from adult female parents	1.184	[2.285, 0.292]
Difference in subadult from adult male parents	0.485	[1.331, 0.267]
Prob(*Y* 3)		
Predictor	Estimate	95% confidence interval
Intercept	2.506	[3.664, 1.554]
1 fledglings in prior year	1.790	[3.068, 0.759]
2 years for prior production	1.341	[4.267, 0.343]
Interaction of above 2 terms	13.980	[565.6, 79.7]
Early nesting precipitation - linear	0.009	[0.029, 0.003]
Early nesting precipitation - quadratic	6.4e 05	[1.6e-04, 2.1e 06]
Early nesting min. temperature - linear	0.168	[0.230, 0.630]
Early nesting min. temperature - quadratic	0.031	[0.315, 0.206]
Winter precipitation - linear	0.005	[0.010, 0.002]
Winter precipitation - quadratic	5.2e 06	[1.6e-05, 3.4e 06]
Precipitation in previous growing season - linear	0.004	[0.014, 0.002]
Precipitation in previous growing season - quadratic	3.7e 05	[1.2e-04, 2.1e 05]
Percentage of 500 acres with trees 25 m height - linear	0.031	[0.0002, 0.0644]
Percentage of 500 acres with trees 25 m height - quadratic	7.3e 04	[0.001, 0.003]
Difference in subadult from adult female parents	1.144	[4.120, 0.632]
Difference in subadult from adult male parents	16.79	[465.4, 37.5]



APPENDIX FIGURE 7. Partial effects of (**A**) early nesting precipitation, (**B**) early nesting minimum temperature, (**C**) winter precipitation, (**D**) precipitation in previous growing season, (**E**) percent of 500 ac around territory with trees 25 m height, and (**F**) age of male and female parents in cumulative ordinal logistic regression model without a proportional odds assumption corresponding to logistic quantile model of California Spotted Owl fledgling counts that included lagged fledgling production, climate, parent age, and landscape habitat ($n = 639$). Each panel is the estimated cumulative distribution functions (cdf) for fledgling counts for partial effects of predictors (all continuous predictors are centered on their means) made by holding other continuous predictors at their mean values for adult parents, with estimates for territories with 0 (left) or 1 (right) fledglings in the prior year. Red portions of cdf are for counts of 3, cyan for counts of 2, black for counts of 1, and gray for counts of 0 fledglings. Note the crossing of estimates at large values of the predictors in **C** and **E**.

APPENDIX TABLE 3. Parameter estimates for ordinary least squares (OLS) regression of mean California Spotted Owl fledgling counts using the same predictor terms as in the full logistic quantile regression model of fledgling counts ($n = 639$); autoregressive prior production, quadratic functions of climate, age of parents, and quadratic function of landscape habitat surrounding territories.

Predictor	Estimate	95% confidence interval
Intercept	0.9895	[0.839, 1.139]
1 fledglings in prior year	0.3657	[0.521, 0.210]
2 years for prior production	0.1756	[0.467, 0.116]
Interaction of above 2 terms	0.2637	[0.214, 0.742]
Early nesting precipitation - linear	7.606e 04	[1.437e 03, 8.442e 05]
Early nesting precipitation - quadratic	1.257e 07	[1.300e 06, 1.049e 06]
Early nesting min. temperature - linear	0.0686	[0.0009, 0.1363]
Early nesting min. temperature - quadratic	0.0049	[0.0474, 0.0376]
Winter precipitation - linear	0.0002	[0.0006, 0.0001]
Winter precipitation - quadratic	2.946e 07	[1.653e 07, 7.544e 07]
Precipitation in previous growing season - linear	0.0017	[0.0026, 0.0008]
Precipitation in previous growing season - quadratic	5.511e 07	[6.909e 06, 5.807e 06]
Percentage of 500 acres with trees 25 m height - linear	0.0058	[0.0002, 0.0119]
Percentage of 500 acres with trees 25 m height - quadratic	0.0005	[0.0009, 0.0001]
Difference in subadult from adult female parents	0.3719	[0.6433, 0.1006]
Difference in subadult from adult male parents	0.3563	[0.6272, 0.0854]



APPENDIX FIGURE 8. Means of the estimated cumulative distribution functions (cdf) for each observation ($n = 639$) from the logistic quantile regression model and for the OLS regression model that included prior production, climate, parent age class, and landscape habitat predictors (Table 1). Solid line is the 1:1 relationship between estimates. OLS regression estimates falling below the line indicated substantial under prediction relative to the logistic quantile regression estimates.