Conservation Biology

Contributed Paper

Inferring the Nature of Anthropogenic Threats from Long-Term Abundance Records

KEVIN T. SHOEMAKER * AND H. RESIT AKÇAKAYA

Department of Ecology & Evolution, Stony Brook University, Stony Brook, NY, 11794, U.S.A.

Abstract: *Diagnosing the processes that threaten species persistence is critical for recovery planning and risk forecasting. Dominant threats are typically inferred by experts on the basis of a patchwork of informal methods. Transparent, quantitative diagnostic tools would contribute much-needed consistency, objectivity, and rigor to the process of diagnosing anthropogenic threats. Long-term census records, available for an increasingly large and diverse set of taxa, may exhibit characteristic signatures of specific threatening processes and thereby provide information for threat diagnosis. We developed a flexible Bayesian framework for diagnosing threats on the basis of long-term census records and diverse ancillary sources of information. We tested this framework with simulated data from artificial populations subjected to varying degrees of exploitation and habitat loss and several real-world abundance time series for which threatening processes are relatively well understood: bluefin tuna (*Thunnus maccoyii*) and Atlantic cod (*Gadus morhua*) (exploitation) and Red Grouse (*Lagopus lagopus scotica*) and Eurasian Skylark (*Alauda arvensis*) (habitat loss). Our method correctly identified the process driving population decline for over 90% of time series simulated under moderate to severe threat scenarios. Successful identification of threats approached 100% for severe exploitation and habitat loss scenarios. Our method identified threats less successfully when threatening processes were weak and when populations were simultaneously affected by multiple threats. Our method selected the presumed true threat model for all real-world case studies, although results were somewhat ambiguous in the case of the Eurasian Skylark. In the latter case, incorporation of an ancillary source of information (records of land-use change) increased the weight assigned to the presumed true model from 70% to 92%, illustrating the value of the proposed framework in bringing diverse sources of information into a common rigorous framework. Ultimately, our framework may greatly assist conservation organizations in documenting threatening processes and planning species recovery.*

Keywords: Bayesian model selection, conservation planning, declining-population paradigm, long-term monitoring, population viability analysis

Inferencia la Naturaleza de las Amenazas Antropogénicas para los Registros de Abundancia a Largo Plazo

Resumen: *Diagnosticar los procesos que amenazan la permanencia de las especies es crítico para la planeación de la recuperación y la predicción de riesgos. Las amenazas dominantes se infieren comúnmente por expertos con base en un collage de métodos informales. Las herramientas de diagnóstico transparentes y cuantitativas podrían contribuir con la tan necesitada consistencia, objetividad y rigor para el proceso de diagnosticar amenazas antropogénicas. Los registros de censos a largo plazo, disponibles para un creciente y diverso conjunto de taxa, pueden exhibir rasgos característicos de procesos específicos de amenaza y así proporcionar información para la diagnosis de amenazas. Desarrollamos un marco de trabajo Bayesiano y flexible para diagnosticar amenazas con base en los registros de censos a largo plazo y diversas fuentes subsidiarias de información. Probamos este marco de trabajo con datos simulados de poblaciones artificiales sujetas a diferentes grados de explotación y pérdida de hábitat y varias series de tiempos de abundancia reales para los cuales están bien entendidos los procesos de amenaza: Thunnus maccoyii y Gadus morhua para la explotación; Lagopus lagopus scotica y Alauda arvensis para la pérdida de hábitat. Nuestro método identificó correctamente el proceso conductor de la declinación poblacional para más del 90% de las series de tiempo simuladas bajo escenarios moderados y severos de amenaza. Nuestro método identificó las amenazas*

*Address correspondence to K. T. Shoemaker, email kevintshoemaker@gmail.com
Paper submitted January 2, 2014; revised manuscript accepted April 27, 2014.

Conservation Biology, Volume 00, No. 0, 1–12

DOI: 10.1111/cobi.12353

con menos éxito cuando los procesos de amenaza eran débiles o cuando las poblaciones estaban afectadas simultáneamente por amenazas múltiples. Nuestro método seleccionó el modelo de la presunta verdadera amenaza para todos los estudios de caso reales, aunque los resultados fueron algo ambiguos en el caso de Alauda arvensis. En el último caso, la incorporación de una fuente subsidiaria de información (registros de cambio en el uso de suelo) incrementaron la fuerza asignada al supuesto modelo verdadero del 70% al 92%, ilustrando el valor del marco de trabajo propuesto en la contribución de diversas fuentes de información para un marco de trabajo común y riguroso. Finalmente, nuestro marco de trabajo puede asistir enormemente a las organizaciones de la conservación en la documentación de procesos de amenaza y en la planeación de la recuperación de especies.

Palabras Clave: análisis de viabilidad poblacional, monitoreo a largo plazo, paradigma de población en declinación, planeación de la conservación, selección de modelo Bayesiano

Introduction

In an effort to improve the link between conservation science and practice, Caughley (1994) famously proposed a "declining population paradigm" for conservation biology that prioritized research on the deterministic processes governing population declines. Diagnosing the causes of population decline (e.g., habitat loss, exploitation, climate change, and invasive species) remains fundamental for species-level conservation (Caughley & Gunn 1996; Beaudry et al. 2008), including conservation prioritization, recovery planning, and projection of species susceptibility to anthropogenic environmental change (Clark et al. 2002; van de Pol et al. 2010). For instance, the effectiveness and rigor of recovery plans for threatened species could be vastly improved if specific drivers of population decline could be identified and targeted via management interventions (Clark et al. 2002; Leidner & Neel 2011). Furthermore, assessment of the species-level consequences of global change, including projections under no-analog conditions, will require a mechanistic understanding of the processes governing population growth and decline (Bellard et al. 2012). Consequently, conservation organizations such as the International Union for Conservation of Nature (IUCN) document putative species-level threats to support conservation status assessments, management recommendations, and risk forecasts (e.g., Baillie et al. 2004).

Peery et al. (2004) identified 6 strategies for diagnosing the drivers of population decline: (1) direct experimentation, (2) comparison of hypothetical (modeled) population responses to threats with observed population dynamics, (3) demographic comparison of geographically isolated populations experiencing a range of hypothesized risk factors, (4) inferring susceptibility on the basis of life history characteristics, (5) identification of external factors (ancillary data) that exhibit strong correlation with abundance dynamics (e.g., correspondence of population decline with the onset date of a particular threat), and (6) use of a "multiple competing hypotheses" approach (an extension of strategies 2 and 5, by which multiple sources of data are compared with expectations under a candidate set of plausible threats). Some of these proposed methods are impractical in the context of threatened or rare species (e.g., experimentation and comparison of multiple populations [Green 1995; Peery et al. 2004]) and others are inconclusive (e.g., use of life history traits to infer susceptibility [Peery et al. 2004]). However, data limitation appears to be the primary factor preventing widespread application of the declining population paradigm. In practice, dominant threats are typically inferred by experts on the basis of a patchwork of informal methods. However, as wildlife monitoring programs become more widespread (e.g., National Ecological Observatory Network [Keller et al. 2008]) and as these data sources are made publicly available, methodological limitations could quickly supplant data limitations. We posit that Bayesian methods are especially well suited to the task of threat diagnosis because they can be used to consider multiple sources of information within a common analytical framework and to specify a broad range of models (e.g., models linking threatening processes with demographic processes) and sources of uncertainty (Clark 2005).

The pattern of decline in abundance may provide clues about the dominant process that has caused a population to decline (Mace et al. 2008; Di Fonzo et al. 2013). Therefore, long-term census records may serve as an important source of information for threat diagnosis. Census records (time series of abundance estimates or indexes) are among the most widely available sources of data for assessing the status of wild populations (Traill et al. 2007). Thousands of long-term census records, hundreds of which span more than 30 years, are readily accessible via databases such as the Global Population Dynamics Database (GPDD) (NERC 2010). However, conservation biologists have rarely used long-term abundance records to infer the nature of the processes underlying population declines (but see Mac Nally et al. 2010; Di Fonzo et al. 2013).

We developed a flexible Bayesian analytical framework for diagnosing the nature of anthropogenic threats. This method can simultaneously incorporate 4 of the 6 threat

diagnosis strategies identified by Peery et al. (2004; strategies 2, 3, 5, and 6). We investigated the performance of this technique through the use of simulated and real-world census records. Ultimately, this technique may greatly assist conservation organizations in documenting threatening processes and planning species recovery.

Methods

We developed and tested a Bayesian model selection method for diagnosing a single driver of population decline on the basis of long-term abundance records (see "Inferring Threats"). We tested this method in 2 ways. First, we simulated census records from demographic models that incorporated specific known threats and then we tested the ability of our method to diagnose the correct threat process. Second, we applied our method to long-term abundance records from 4 well-studied wild populations for which the presumed true threatening process is well understood.

Inferring Threats

To discriminate among a priori hypotheses regarding the processes governing population decline, we first specified likelihood functions corresponding to each of 3 candidate threatening processes: null (no threat), exploitation (overharvest), and habitat loss. In general, any threatening process that can be expressed as a likelihood function could be included in the candidate model set.

For the null model (Eqs. 1 and 2), we used a discrete logistic (i.e., Ricker) population growth model with stable carrying capacity (K). Carrying capacity (K) for the null model (allowing for minor or inconsequential changes in K; see below) was

$$K_t = K_0 + \beta \cdot t, \tag{1}$$

where β is a linear rate of habitat change, which provided an additional model degree of freedom to fit minor trends in abundance whereby K_t was constrained to remain between 110% and 90% of initial abundance (Supporting Information), K_0 is the initial carrying capacity, and t is time in years. The likelihood function for the null model was

$$\log(N_t) = \log(N_{t-1}) + r_{\max}(K_t - N_{t-1})/K_t + e_t, \tag{2}$$

where log (N_t) represents the natural log of abundance (N) at time t, $r_{\max}$ represents the intrinsic maximum rate of growth (and specifies the strength of compensatory density dependence), and e_t represents *iid* lognormal temporal process variation.

We assumed that the exploitation threat could be expressed either as removal of a constant fraction of the population each year (fractional harvest model, Eq. 3) or as removal of a constant number of individuals each year (constant harvest model, Eq. 4). The likelihood function for the fractional harvest model was

$$\log(N_t) = \log(N_{t-1}) + \log(1 - h \cdot \nu_t) + r_{\max}(K - N_{t-1})/K + e_t, \tag{3}$$

and the constant harvest model was specified as

$$\log(N_t) = \log(N_{t-1} - h \cdot \nu_t) + r_{\max}(K - N_{t-1})/K + e_t, \tag{4}$$

where ν_t is a binary indicator of whether the threat is in effect (dummy variable) and h is the harvest rate. Habitat loss was specified as a linear decline in K (Eq. 5), and the population-level response took either a discrete logistic formulation (Eq. 6) or a ceiling model formulation (Eq. 7), whereby suppression of growth occurs only if abundance exceeds K (ceiling model, Eq. 6). The carrying capacity for the habitat loss model was

$$K_t = K_0 + \beta \cdot \nu_t \cdot t, \tag{5}$$

the likelihood function for the Ricker formulation model was

$$\log(N_t) = \log(N_{t-1}) + r_{\max}(K_t - N_{t-1})/K_t + e_t, \tag{6}$$

and the likelihood function for the ceiling formulation was

$$\log(N_t) = \log(N_{t-1}) + (1 - \tau_t) \cdot r_1 + \tau_t \cdot \left[r_2(K_t - N_{t-1})/K_t\right] + e_t, \tag{7}$$

where the r terms represent intrinsic rates of growth (in the ceiling model, r_1 represents the constant growth rate below K, while r_2 is analogous to an $r_{\max}$ term when abundance exceeds K) and τ indicates if abundance exceeds K (dummy variable). If external ancillary data are available to characterize temporal trends for any of the threatening processes considered (e.g., data on changes in habitat availability), these data may be used in place of the simplistic assumption of a stationary threat (e.g., Eq. 5 could be replaced by a linear function of habitat availability or an index thereof). All model parameters were assigned uniform prior distributions to circumscribe a plausible parameter space (see Supporting Information for details of our rationale for assigning prior distributions).

Using a standard Bayesian model selection procedure (Bijak 2006), we selected the most probable threat from among the 3 candidate threats: null, exploitation, or habitat loss (i.e., the model with the greatest posterior probability of generating the observed data was selected as the best model [Wasserman 2000]) and stored the posterior probabilities (posterior model weights) assigned to each model for further analysis. In Bayesian model selection, each candidate model must be assigned a prior probability representing a weight of evidence for each model

before any data are observed. To ensure that the null model was selected for census records exhibiting little or no decline, we assigned 50% of prior model weight to the null model (Eq. 2) and distributed the remaining prior model weight equally among the exploitation (Eq. 3 or Eq. 4; equal prior model weight assigned to each alternative exploitation process) and habitat loss models (Eq. 6 or Eq. 7; equal prior model weight assigned to the each alternative habitat loss process). Alternative prior model weights could be assigned on the basis of prior research, expert opinion, or stakeholder risk tolerance.

Candidate models were fit to time-series records with Markov Chain Monte Carlo (MCMC) in WinBUGS (Lunn et al. 2000), which was called from the R environment with the R2WinBUGS package (Sturtz et al. 2005). We ran 20,000 MCMC iterations for each time series ($n = 960$). The first 10,000 iterations were discarded as burn-in, and every 10th iteration was stored for further analysis. The R and WinBUGS code for running the analyses are provided in Supporting Information.

Test with Simulated Data

We used RAMAS Metapop 6.0 (Akçakaya & Root 2013) to generate multiple replicate time series for populations unaffected by a threatening process (null scenarios) and for populations threatened by exploitation or habitat loss. Threatening processes were specified at 3 distinct severity levels—weak, moderate, and severe—corresponding to expected 45-year declines of 10%, 50%, and 80%, respectively. To assess diagnostic power across a wide range of realistic conditions, we varied the strength of density-dependent population growth (r_{max}), threat duration (time-series length), and temporal process variation. In total, we generated 14,100 replicate time series from 282 unique data-generating models (single-threat scenarios; see Supporting Information for details). In addition to single-threat scenarios, we also simulated census records affected simultaneously by both exploitation and habitat loss (combined threat scenarios; see Supporting Information).

For each unique data-generating model, we evaluated the strength and consistency with which the true threatening process was selected from among the candidate models on the basis of the frequency (out of 50 replicate time series) with which the true threatening process was assigned the highest posterior model weight and the posterior model weight (average model weight across all 50 replicates) assigned to the true threatening process. We also evaluated the frequency by which the wrong threat was detected, the frequency by which the diagnosis was strongly correct (more than 50% of model weight assigned to the true model), and the frequency by which the diagnosis was highly misleading (more than 50% of model weight assigned to the wrong threat). For combined threat scenarios, we evaluated the frequency with which each candidate model was selected and the extent to which the posterior model weights for each candidate threat reflected the relative severities of the true threatening processes acting on the simulated population.

To evaluate the conditions under which our method was most likely to succeed or to provide indeterminate (or potentially misleading) results, we modeled diagnostic performance (for each unique data generating model, the proportion of the 50 replicate time series assigned to the correct threatening process) as a function of the true threatening process, threat severity, level of temporal environmental variability, strength of density dependence, and time-series length. These analyses were performed using random forest (to assess the relative importance of each variable for threat diagnosis [Breiman 2001]) and conditional inference tree analyses (a distribution-free method for recursive partitioning [Hothorn et al. 2006]), which were implemented using the 'party' package in R (Hothorn et al. 2006). Relative importance values were computed as the relative loss in predictive performance incurred after scrambling the values of each predictor variable (Breiman 2001). Univariate relationships were visualized using standard box plots.

We evaluated the accuracy of threat diagnosis by treating individual simulated census data sets (replicate time series) as the fundamental unit of observation (analogous to census records from a single real-world population). In these analyses, the response variable was diagnostic success (a binary variable indicating whether or not the highest posterior weight was assigned to the correct model) and the predictor variables were 2 summary statistics computed for each replicate time series: realized decline trend (signal) and variability (noise). We used univariate logistic regression (each predictor variable was considered separately) to perform these analyses (glm function in R). The decline signal was computed as the fractional decline in expected abundance over the study period, determined on the basis of an ordinary linear regression. The variability or noise associated with each replicate time series was determined as the root mean squared error (RMSE) from a locally weighted regression (loess function in R with standard smoothing parameter) expressed as a percentage of initial abundance. This function accommodates nonlinear relationships; therefore, the residuals more closely represent noise, or variation around the expected value for each year.

Test with Real-World Case Studies

We obtained long-term abundance records for several iconic examples of population declines that are widely understood to be caused by exploitation or habitat loss. We selected time series for Atlantic cod (*Gadus morhua*) and southern bluefin tuna (*Thunnus maccoyii*) to

Table 1. Performance metrics (proportion correct[a] and average model weight[b]) illustrating the success of our proposed Bayesian method in diagnosing the cause of population decline on the basis of census records simulated under known threatening processes (either null or exploitation or habitat loss).[c]

	Proportion selected (n = 50)[a]				*Average model weight*[b]		
Threatening process	*correct*	*strongly correct*	*wrong threat*	*strongly misleading*[d]	*none (null)*	*exploitation*	*habitat loss*
Null	0.82	0.80	0.18	0.08	0.659	0.113	0.228
Weak exploitation (3%/year or 150 individuals/year)	0.36	0.16	0.64	0.38	0.261	0.338	0.401
Moderate exploitation (4.5% or 275/year)	0.72	0.70	0.28	0.24	0.134	0.588	0.278
Severe exploitation (6% or 400/year)	0.88	0.88	0.12	0.12	0.043	0.802	0.155
Weak habitat loss (K reduced by 50/year)	0.30	0.26	0.70	0.58	0.538	0.120	0.341
Moderate habitat loss (125/year)	0.80	0.66	0.20	0.16	0.233	0.188	0.579
Severe habitat loss (200/year)	0.92	0.90	0.08	0.08	0.043	0.134	0.823

[a]*Proportion of replicate time series (n = 50) in which the true threatening process was correctly diagnosed.*
[b]*Mean posterior model weight (n = 50) assigned to each of 3 candidate threats: null, exploitation, and habitat loss.*
[c]*This table does not summarize the results from all 282 unique data-generating models used to generate simulated time-series records. Rather, each row of this table presents results from a single-threat scenario (n = 50 replicate time series) at moderate settings for density dependence, environmental process variance, and time-series length. See text for details.*
[d]*Results from the threat diagnosis algorithm for a given replicate were characterized as strongly misleading if ≥50% of model weight was assigned to an incorrect hypothesis.*

represent exploitation (Walters & Maguire 1996; Collette et al. 2011), and we selected the red grouse in Scotland (*Lagopus lagopus scotica*) and the Eurasian Skylark (*Alauda arvensis*) in England to represent loss of habitat (Chamberlain et al. 2000; Thirgood et al. 2000). Data sets were obtained from the GPDD (tuna and cod; NERC 2010), published accounts (grouse) (Thirgood et al. 2000), and British Trust for Ornithology (BTO) Breeding Bird Survey (BBS), and the BTO Common Birds Census (CBC) (skylark) (Williamson & Homes 1964). For the skylark case study, we also obtained an index of changing farming practices in Great Britain (Chamberlain et al. 2000; axis 1 of detrended correspondence analysis summarizing 32 agricultural variables) as ancillary data for enhancing diagnostic power. This index has been linked to the decline of farmland-associated birds in Great Britain (Chamberlain et al. 2000). To model the effect of farmland habitat alteration on skylark annual growth rates, we modeled carrying capacity as a linear function of the farmland habitat index. For the final 15 years of the skylark time series, for which the Chamberlain habitat index data were unavailable, the rate of habitat alteration was assumed constant at the rate characteristic of the final 10 years for which the index was available. As another preliminary test of the use of ancillary data to improve threat diagnosis, we incorporated the year (1974) at which strict take quotas were enacted for the Greenland cod fishery (Horsted 2000) to represent the effective end date for the exploitation threat in this population.

Results

Test with Simulated Data

For time series simulated under conditions of moderate density dependence ($r_{max} = 1.1$), threat duration (30 years), and process variation (SD 0.05; Table 1), our method correctly diagnosed the true threatening process for 88% of severe exploitation scenarios, 92% of severe habitat loss scenarios, and 82% of null (no threat) scenarios. Diagnostic power remained >70% at moderate severity levels but was substantially reduced (<40%) when threatening processes were weak (Table 1). For exploitation scenarios simulated under a constant harvest regime (150–400 individuals harvested per year depending on threat severity), the ability of our algorithm to correctly identify exploitation as the threat decreased such that the proportion of correct diagnosis was reduced by 20–26%. Simulated data sets were nested within the posterior predictive distribution (PPD; cloud of expected trajectories) for threat models assigned high posterior weight, suggesting adequate goodness of fit (Fig. 1). This result was confirmatory yet unsurprising given that a close approximation of the true data-generating process was included in the candidate model set.

Stronger density dependence (higher r_{max} values) tended to increase diagnostic performance (proportion of simulated time series successfully diagnosed) for habitat loss and null (no threat) scenarios but generally reduced diagnostic performance for exploitation scenarios (Figs. 2a–2c). Higher levels of process variation tended to

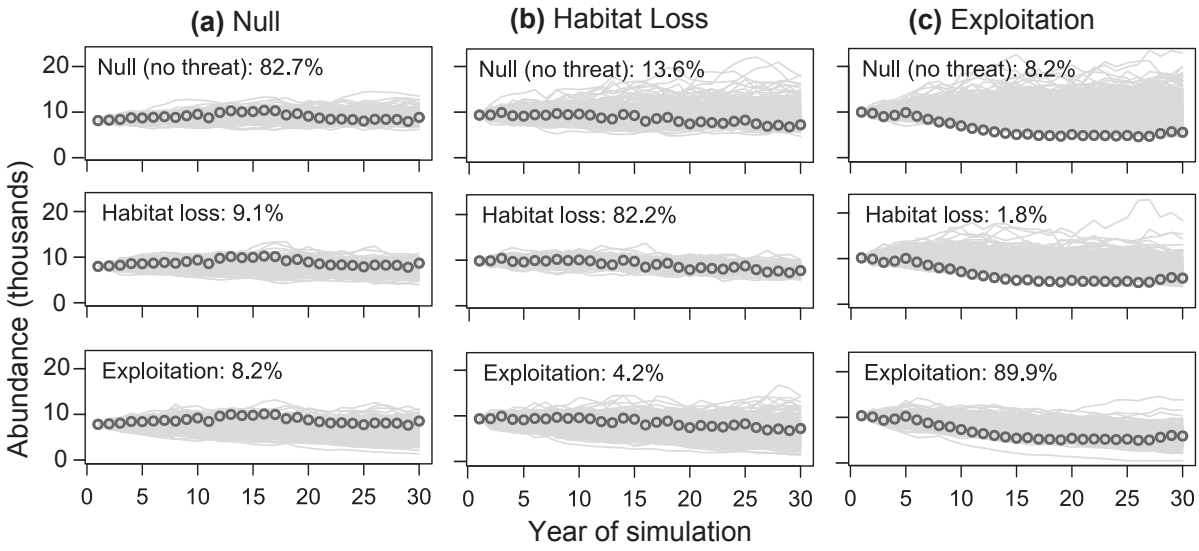

Figure 1. Illustration of model selection and goodness-of-fit for 3 replicate abundance trajectories for a generic organism with annual, nonoverlapping generations simulated under (a) null (no threat), (b) moderate habitat loss, and (c) moderate exploitation scenarios, respectively. Simulated census data (open circles, a single simulated abundance trajectory) are superimposed on the posterior predictive distribution (cloud of trajectories expected to be produced under each candidate model) for each of 3 candidate threatening processes (null, habitat loss, and exploitation). Posterior model weights (representing the relative support for each candidate model given the available census data) are listed as percentages.

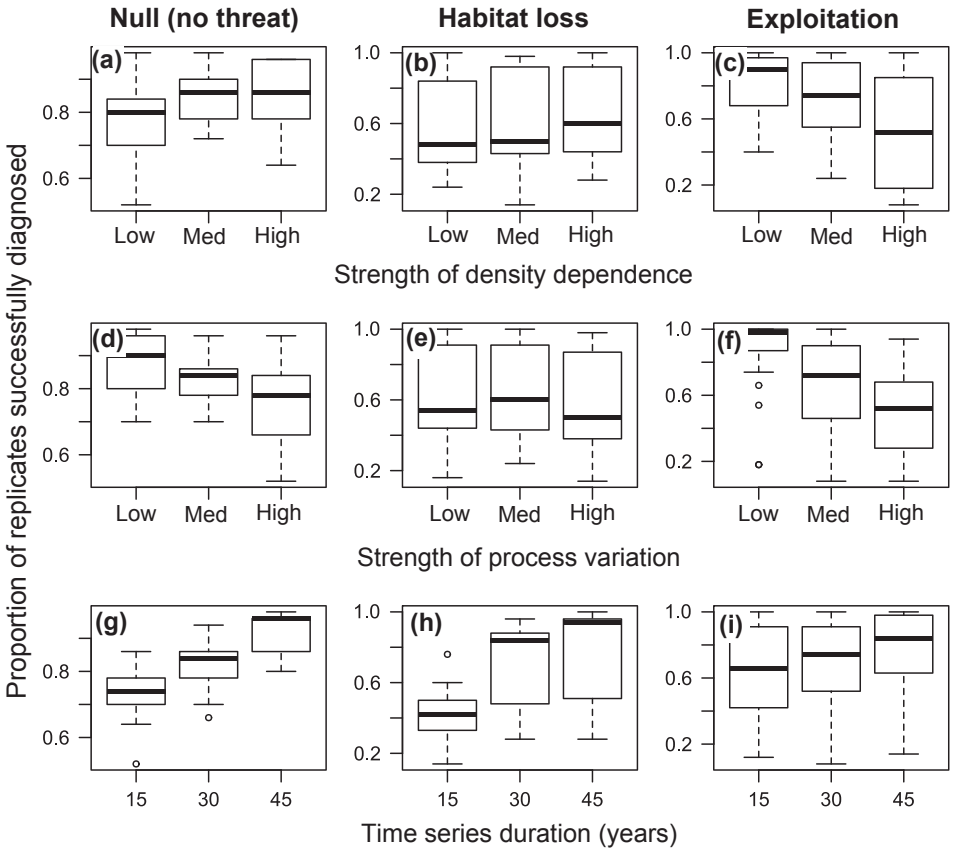

*Figure 2. Diagnostic performance (proportion of replicate simulations for which the true threatening process was correctly diagnosed on the basis of census records) as a function of (a–c) strength of density dependence, (d–f) strength of environmental process variation, and (g–i) time-series duration (columns, 3 threatening processes under which abundance trajectories were simulated; solid lines, median proportion; box, interquartile range [IQR]; whiskers, range of data; circles, outliers [points located more than 1.5*IQR from the nearest box edge]). The y-axis scale for the no-threat scenario differs from those for habitat loss and exploitation.*

reduce diagnostic performance for time series simulated under null and exploitation scenarios but did not exert a strong effect for habitat loss scenarios (Figs. 2d–2f). Longer threat duration was associated with a substantial improvement in predictive power for time series' simulated under null and habitat loss scenarios, but this effect was weak for time series simulated under an exploitation scenario (Figs. 2g–2i).

Discriminatory power was highest for scenarios with moderate to severe threat severity level, time-series duration ≥30 years (over which threatening processes were stationary), and moderate to low process variability, whereby correct diagnosis rate reached 95% (Fig. 3a). In contrast, the algorithm performed poorly (26% correct diagnosis) for weak habitat loss scenarios with 15 year threat duration and (29% correct diagnosis) for weak exploitation or habitat loss scenarios with strong density dependence (Fig. 3a) (in both cases, the more parsimonious null model was generally selected). A random forest algorithm indicated that threat severity was the most important factor determining diagnostic performance, followed by time-series length (threat duration), threat scenario (null, exploitation, and habitat loss), and the magnitude of process variation. The strength of density dependence was relatively unimportant as a determinant of diagnostic performance (Fig. 3b).

When replicate time series, rather than unique data-generating models, were treated as the observation unit, the strength of the realized trend was a strong determinant of successful threat diagnosis (Figs. 4a–4c). Diagnoses were generally successful for exploitation and habitat loss trajectories that exhibited declining trends greater than or equal to approximately 50% over the study duration, and threat diagnoses were generally successful for realized null scenarios that exhibited declining trends less than or equal to approximately 20%. Furthermore, the amount of temporal variation in realized time series tended to exert opposing effects on diagnostic success (Figs. 4d–4f). Time series with high temporal variation were associated with improved probability of successful diagnosis under a habitat loss scenario, whereas lower temporal variation was associated with higher probability of correct diagnosis for trajectories generated under an exploitation scenario (much of the temporal variation exhibited by populations undergoing habitat loss was related to compensatory density-dependent growth and therefore represented signal rather than noise).

For simulated time series simultaneously affected by habitat loss and exploitation, diagnostic performance was generally poor (Supporting Information). Because trajectories simulated under habitat loss scenarios tended to exhibit convex trajectories (with strongest effects manifesting in the early years of a time series) while exploitation scenarios tended to produce concave trajectories (e.g., Fig. 1), exploitation effects tended to manifest earlier than habitat loss effects. Even weak exploitation processes tended to force abundance below carrying capacity for the majority of the simulation duration, masking much of the habitat loss signal. Consequently, habitat loss was almost never (<1% of replicate time series) diagnosed as the cause of decline for populations subjected to severe exploitation coupled with weak habitat loss. Furthermore, exploitation remained the most common diagnosis (57% of replicate time series) for artificial populations experiencing weak exploitation coupled with severe habitat loss (Supporting Information).

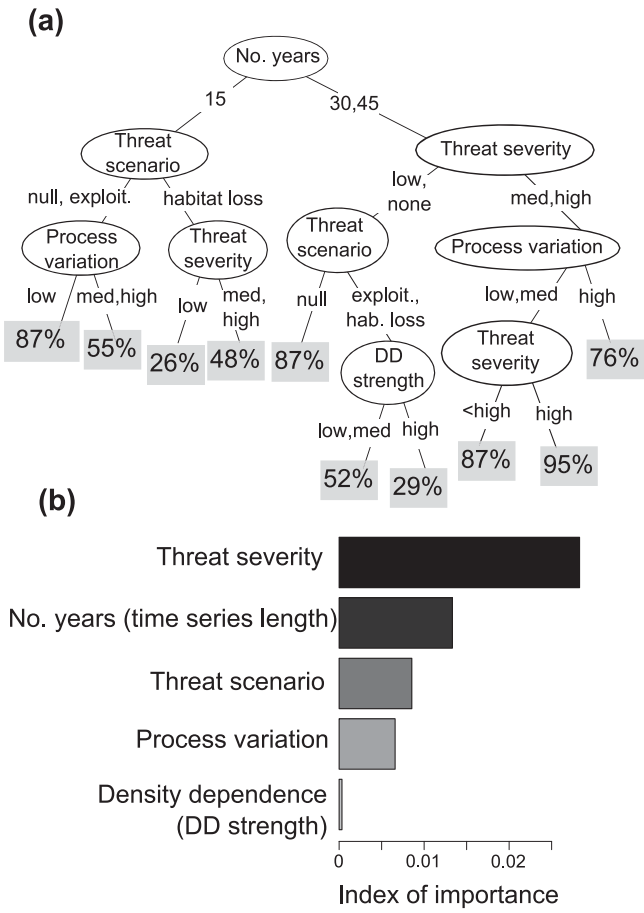

Figure 3. Sensitivity of diagnostic performance (proportion of replicate simulations for which the true threatening process was correctly diagnosed on the basis of census records) to 5 variables hypothesized to influence the power to infer threatening processes: threat severity, time-series length, threat scenario, magnitude of process variation, and strength of density dependence. (a) Results from a conditional inference tree (ovals, predictor variables; lines, individually labeled splitting rules; gray, percent correct diagnosis). (b) Relative importance of each variable for predicting diagnostic performance (derived from a random forest algorithm).

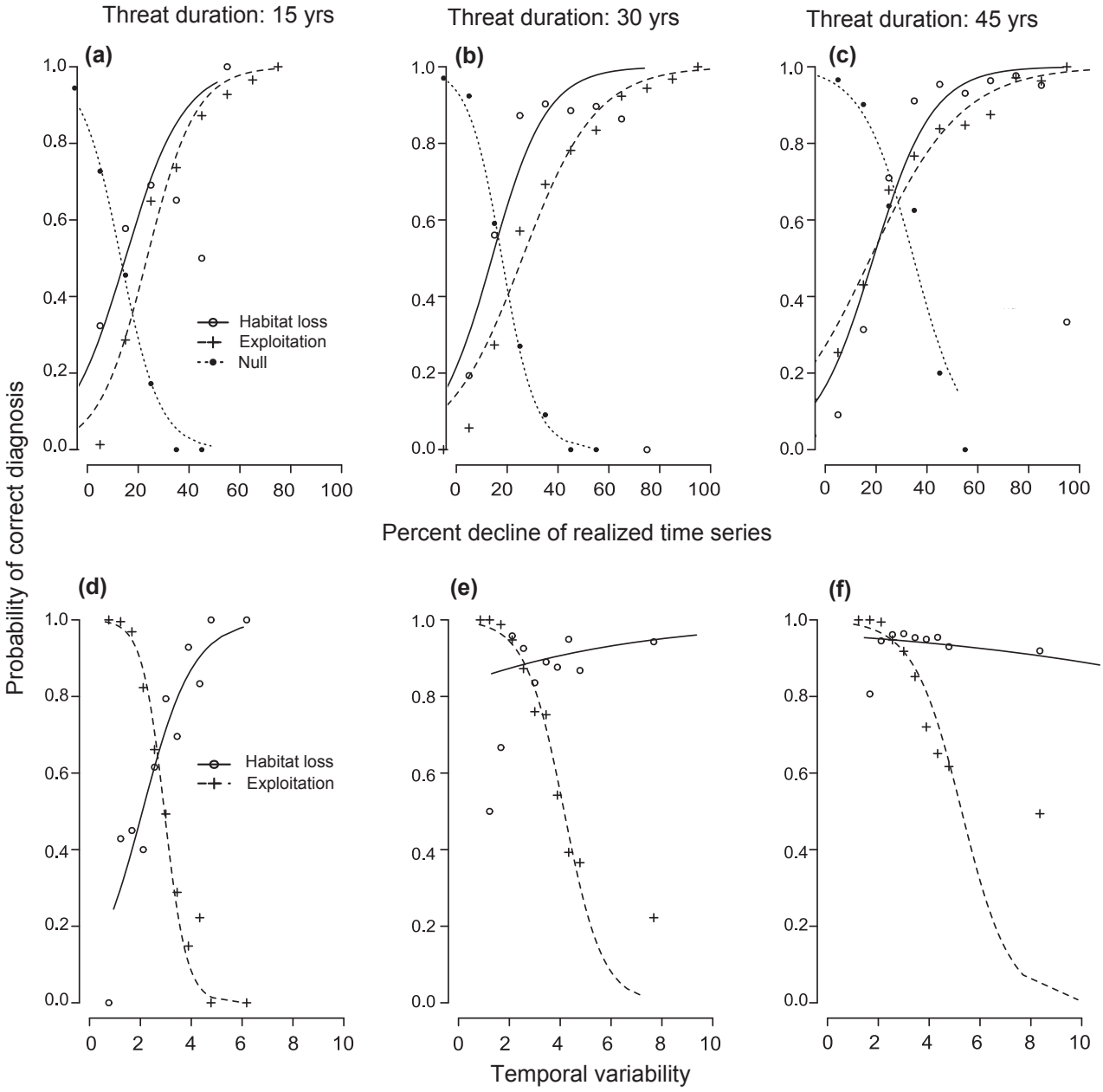

Figure 4. Relationship between the probability of correct threat diagnosis and (a–c) the severity of the realized decline and (d–f) the temporal variability (noise of realized time series, expressed as the root mean squared residual error from a locally weighted regression model) exhibited by simulated abundance trajectories with time frames of (a, d) 15, (b, e) 30, and (c, f) 45 years (points, mean diagnostic performance across 10 evenly spaced bins spanning the range of each predictor variable; closed circles, null model [no threat]; open circles, habitat loss; crosses, harvest processes; curves, fitted using logistic regression [each replicate time series treated as an independent data point]). Only time series exhibiting realized declines ≥25% (corresponding to populations likely to be flagged as declining by conservationists) were included in results depicted in (d–f).

Test with Real-World Case Studies

Our threat diagnosis approach correctly identified the presumed true threat model for all 4 real-world case studies. For the southern bluefin tuna, the exploitation model received 100% of the posterior model weight (Fig. 5a). Similarly, the exploitation model received 97% of posterior model weight for the Atlantic cod case study (Fig. 5b) when the date of enactment of strict harvest quotas (and thereby much lower harvest rates) was used to model the effective termination of the exploitation threat for this fishery. However, the model weight in favor of the exploitation model was negligible under a stationary harvest

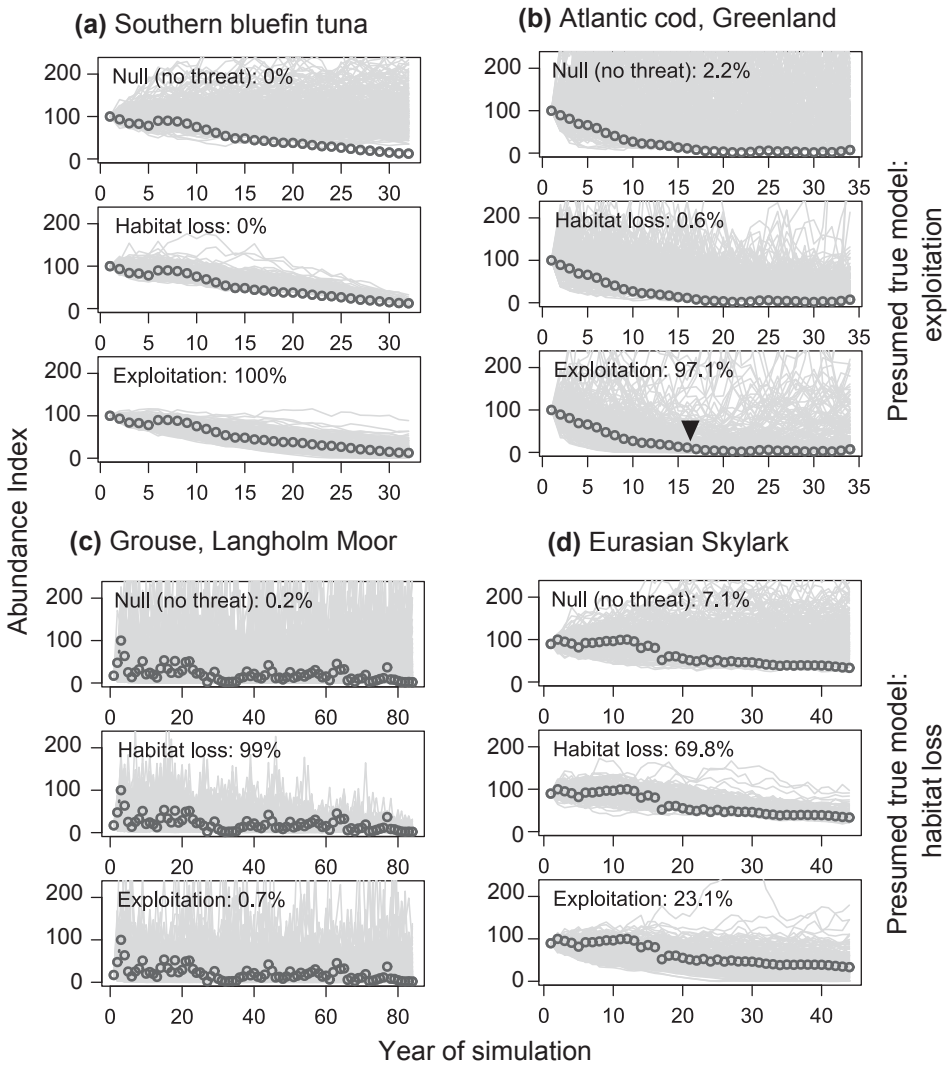

Figure 5. Model selection and goodness-of-fit for real-world time series drawn from the published literature for which declines in abundance are understood to be driven by (a, b) exploitation (southern bluefin tuna and Atlantic cod in Greenland) or (c, d) habitat loss (Red Grouse at Langholm moor in Scotland and Eurasian Skylark in Great Britain). Observed census data (open circles) are superimposed on the posterior predictive distribution (cloud [gray] of expected trajectories) for each of 3 candidate threatening processes. Posterior model weights (representing the relative support for each candidate model given the available census data) are listed as percentages. For the Atlantic cod, the black triangle indicates the harvest termination date (whereby strict harvest quotas were enacted).

model (i.e., harvest process assumed constant across the study duration, in which case a majority of model weight was assigned to the habitat loss model). The decline of the red grouse in Langholm Moor, Scotland, presumed to be a result of habitat loss, was diagnosed as such with a high degree of certainty (99% posterior model weight; Fig. 5c). The diagnostic results were somewhat more ambiguous in the case of the Eurasian Skylark; 70% of posterior weight was assigned to the habitat loss model (Fig. 5d). However, when data on the increasing mechanization of farming practices in Great Britain were used to model changes in carrying capacity for Eurasian Skylarks, the posterior model weight in favor of the habitat loss model increased from 70% to 92%.

Discussion

Our results suggest that the patterns of population decline embodied in long-term census records can provide valuable information for threat diagnosis. The approach we propose for inferring the cause of a population's decline has the potential to benefit the recovery planning process by putting the identification of threats on a robust, objective basis. It would be especially beneficial for populations wherein the most effective conservation or recovery action requires knowing which threat (from among a set of several plausible threats) is the dominant driver of the population's decline. The red grouse case study, for which the population-level benefits of habitat restoration would likely far exceed that of establishing and enforcing harvest quotas, provides a case in point.

Mace et al. (2008) suggest that the shape of the decline–rate curve depends on the threatening process involved. Testing this suggestion, Di Fonzo et al. (2013) found that the shape (especially the concavity) of population decline curves can be used to discern between broad categories of pressure or threat types, but they were not able to associate shape of decline with specific threats. We used annual realized growth rates as the unit of observation, focusing on fine-grained patterns within the decline curve instead of general shape characteristics.

Decline curves helped identify causes of population decline, especially when analyzed in conjunction with ancillary data relevant to specific threatening processes or events (e.g., dates of threat onset or cessation). Our approach integrates the "modeled population response," "multiple competing hypotheses," and "timing of decline" methods discussed by Peery et al. (2004) (with potential for integrating the "population comparisons" method) and can be implemented with a source of data (long-term census records) that has been underutilized for purposes of diagnosing population declines.

For our proposed methods to be successful in correctly diagnosing the causes underlying population decline for a wide range of taxa and threatening processes, population growth patterns should exhibit unique statistical signatures for each candidate threat. Threats that produce similar decline patterns cannot be discriminated solely on the basis of these patterns. In our simulation study, proportional harvest (constant proportion of the population harvested annually) and habitat loss mechanisms exhibited unique statistical signatures such that information on the underlying mechanism could generally be recovered. However, we found that the constant harvest model (constant absolute rate of take) yielded a pattern of population growth that was difficult to distinguish from a habitat loss model, resulting in somewhat lower proportion of correct diagnoses. In such cases, incorporating ancillary sources of information may substantially improve the reliability of predictions. This point was demonstrated in the Eurasian Skylark case study, in which information on habitat degradation and changing farming practices in Great Britain helped to lend additional weight to the habitat loss hypothesis. The value of ancillary data for improving discrimination power was also demonstrated in the Atlantic cod case study, in which the presumed true threatening process (exploitation) was correctly diagnosed only when the effective end date of the exploitation threat was modeled explicitly (Fig. 5b).

Sources of ancillary data that may improve threat diagnoses include time series of climatic variables, harvest records, dates of introduction and spread of invasive species, estimated temporal changes in habitat suitability (e.g., based on historical land use change), and timing of catastrophic events (e.g., Esler et al. 2002). Improvement in diagnostic power is likely to be particularly strong if ancillary data sources are available for all candidate threats (e.g., Wolf & Mangel 2008) or if spatially explicit ancillary data can be associated with census data from multiple spatially distinct populations. For example, invasions by introduced predators, competitors, and disease tend to begin as localized, low-abundance colonies, to proceed as abundance and geographic range expansion, and end in a new stable state (Sakai et al. 2001). The expected timing of invasive species impact for each of several focal populations, perhaps informed by a spatial spread model (Hastings et al. 2005), may provide a distinct spatiotemporal signature for diagnosing threats due to invasive predators, competitors, or disease, even if the statistical signature from individual time series was weak or not unique (Peery et al. 2004).

Challenges for our threat diagnosis method, as currently implemented, include weak threats, multiple threats, temporal lags, and measurement error. Although our method frequently failed to identify weak threats, we do not consider this a major issue because diagnosing the causes of population decline is less urgent in those cases. Multiple threats present a more serious challenge. Many species and populations are threatened by multiple interacting threats rather than a single dominant stressor, and synergistic interactions among population-level stressors may increase under climate change (Brook et al. 2008). In our combined-threat scenarios, exploitation was more frequently diagnosed as the primary driver of decline (Supporting Information). This result was not unexpected because the exploitation process tended to force populations below carrying capacity (K), effectively suppressing the habitat loss signature despite simulated declines in K. However, this result may be problematic from a conservation standpoint. In particular, the identification of exploitation as the primary threat may prompt managers to implement harvest quotas and to ignore important threats for which statistical signals may be masked. In cases where the objective of threat diagnosis is to identify the contribution of multiple interacting stressors rather than to identify a single dominant stressor, a more effective strategy may be to combine all threatening processes into a single likelihood function such that different parameters may be interpreted as the relative contribution from alternative threats (e.g., Wolf & Mangel 2008). However, further work is necessary to develop and test such a framework for cases where census records represent the only available data source.

Measurement (observational) error can overwhelm process error (Maxwell & Jennings 2005; Dennis et al. 2006) and lead to difficulties in differentiating signal from noise and potentially biasing model selection (Nadeem & Lele 2012). Although we have not yet challenged our framework against artificial census data simulated with common observation error processes, hierarchical Bayesian analytic methods can greatly simplify the formal treatment of measurement error (Clark 2005). In future efforts, we will focus on building and testing methods that are robust to measurement error. Finally, temporal lags in population response to an anthropogenic threat may result when stage-structured populations are exposed to threatening processes that differentially affect various life stages (Solberg et al. 1999). In many cases (e.g., small numbers of stage classes or where census data are themselves stage structured), stage structure may be modeled explicitly (i.e., represented in likelihood functions). In some cases, it may be impractical to specify explicit likelihood functions for all candidate threats, as

may be the case for stage-structured, spatially explicit metapopulations or complex multiparameter threat models. To accommodate such cases, we plan to implement and test our approach in an approximate Bayesian computation framework (ABC) (Beaumont 2010).

Our approach shows potential to provide an objective, robust method for diagnosing causes of population decline, especially when combined with case-specific hypotheses, use of ancillary data when available, simultaneous analysis of time series of multiple populations or species, other forms of diagnosis when possible (e.g., compare different populations or species), and experimentation when feasible.

Acknowledgments

This research was funded in part by NSF grant DEB-1146198. K. Risely of the British Trust for Ornithology provided the raw count data for the Eurasian Skylark case study. We thank C. Carroll and 2 anonymous reviewers for helpful comments and discussion.

Supporting Information

Details on the specification of prior distributions (Appendix S1), a table of diagnostic results for multithreat scenarios (Appendix S2), annotated WinBUGS code (Appendix S3) and R functions and scripts (Appendix S4) for performing the analysis described herein, and details of the demographic simulation model for generating artificial census records (Appendix S5) are available online. The authors are solely responsible for the content and functionality of these materials. Queries (other than absence of the material) should be directed to the corresponding author.

Literature Cited

Akçakaya, H. R., and W. Root. 2013. RAMAS Metapop: viability analysis for stage-structured metapopulations (version 6.0). Applied Biomathematics, Setauket, New York.

Baillie, J. E. M., C. Hilton-Taylor, and S. N. Stuart, editors. 2004. 2004 IUCN red list of threatened species. A global species assessment. IUCN, Gland, Switzerland.

Beaudry, F., P. G. deMaynadier, and M. L. Hunter. 2008. Identifying road mortality threat at multiple spatial scales for semi-aquatic turtles. Biological Conservation **141:**2250–2563.

Beaumont, M. A. 2010. Approximate Bayesian computation in evolution and ecology. Annual Review of Ecology Evolution and Systematics **41:**379–406.

Bellard, C., C. Bertelsmeier, P. Leadley, W. Thuiller, and F. Courchamp. 2012. Impacts of climate change on the future of biodiversity. Ecology Letters **15:**365–377.

Bijak, J. 2006. Bayesian model averaging in forecasting international migration. European Population Conference. Liverpool, United Kingdom, 21–24 June 2006. Session 904: Projections.30 pp. Available at http://www.infostat.sk/vdc/epc2006/papers/epc2006s60094.pdf.

Breiman, L. 2001. Random forest. Machine Learning **45:**5–32.

Brook, B. W., N. S. Sodhi, and C. J. Bradshaw. 2008. Synergies among extinction drivers under global change. Trends in Ecology & Evolution **23:**453–460.

Caughley, G. 1994. Directions in conservation biology. Journal of Animal Ecology **63:**215–244.

Caughley, G., and Gunn, A. 1996. Conservation biology in theory and practice (Vol. 459). Blackwell Science, Cambridge.

Chamberlain, D. E., R. J. Fuller, R. G. H. Bunce, J. C. Duckworth, and M. Shrubb. 2000. Changes in the abundance of farmland birds in relation to the timing of agricultural intensification in England and Wales. Journal of Applied Ecology **37:**771–788.

Clark, J. A., J. M. Hoekstra, P. D. Boersma, and P. Kareiva. 2002. Improving U.S. Endangered Species Act recovery plans: key findings and recommendations of the SCB recovery plan project. Conservation Biology **16:**1510–1519.

Clark, J. S. 2005. Why environmental scientists are becoming Bayesians. Ecology Letters **8:**2–14.

Collette, B., S.-K. Chang, A. Di Natale, W. Fox, M. J. Juan-Jorda, N. Miyabe, R. Nelson, Y. Uozumi, and S. Wang. 2011. "*Thunnus maccoyii*". IUCN Red List of Threatened Species. Version 2011.2. International Union for Conservation of Nature. Retrieved 19 December 2012.

Dennis, B., J. M. Ponciano, S. R. Lele, M. L. Taper, and D. F. Staples. 2006. Estimating density dependence, process noise, and observation error. Ecological Monographs **76:**323–341.

Di Fonzo, M., B. Collen, and G. M. Mace. 2013. A new method for identifying rapid decline dynamics in wild vertebrate populations. Ecology and Evolution **3:**2378–2391.

Esler, D., T. D. Bowman, K. A. Trust, B. E. Ballachey, T. A. Dean, S. C. Jewett, and C. E. O'Clair. 2002. Harlequin duck population recovery following the '*Exxon Valdez*' oil spill: progress, process and constraints. Marine Ecology Progress Series **241:**271–286.

Green, R. E. 1995. Diagnosing causes of bird population declines. Ibis **137:**S47–S55.

Hastings, A., et al. 2005. The spatial spread of invasions: new developments in theory and evidence. Ecology Letters **8:**91–101.

Horsted, S. A. 2000. A review of the cod fisheries at Greenland, 1910–1995. Journal of Northwest Atlantic Fishery Science **28:**1–112.

Hothorn, T., K. Hornik, and A. Zeileis. 2006. Unbiased recursive partitioning: a conditional inference framework. Journal of Computational and Graphical Statistics **15:**651–674.

Keller, M., D. S. Schimel, W. W. Hargrove, and F. M. Hoffman. 2008. A continental strategy for the National Ecological Observatory Network. Frontiers in Ecology and the Environment **6:**282–284.

Leidner, A. K., and M. C. Neel. 2011. Taxonomic and geographic patterns of decline for threatened and endangered species in the United States. Conservation Biology **25:**716–725.

Lunn, D. J., A. Thomas, N. Best, and D. Spiegelhalter. 2000. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing **10:**325–337.

Mace, G. M., et al. 2008. Quantification of extinction risk: IUCN's system for classifying threatened species. Conservation Biology **22:**1424–1442.

Mac Nally, R., et al. 2010. Analysis of pelagic species decline in the upper San Francisco Estuary using multivariate autoregressive modeling (MAR). Ecological Applications **20:**1417–1430.

Maxwell, D., and S. Jennings. 2005. Power of monitoring programmes to detect decline and recovery of rare and vulnerable fish. Journal of Applied Ecology **42:**25–37.

Nadeem, H., and S. R. Lele. 2012. Likelihood based population viability analysis in the presence of observation error. Oikos **121:**1656–1664.

NERC Centre for Population Biology, Imperial College. 2010. The global population dynamics database. Version 2. Available

from http://www.sw.ic.ac.uk/cpb/cpb/gpdd.html (accessed 20 July 2012).

Peery, M. Z., S. R. Beissinger, S. H. Newman, E. B. Burkett, and T. D. Williams. 2004. Applying the declining population paradigm: diagnosing causes of poor reproduction in the marbled murrelet. Conservation Biology **18:**1088–1098.

Sakai, A. K., et al. 2001. The population biology of invasive species. Annual Review of Ecology and Systematics **2001:**305–332.

Solberg, E. J., B. E. Saether, O. Strand, and A. Loison. 1999. Dynamics of a harvested moose population in a variable environment. Journal of Animal Ecology **68:**186–204.

Sturtz, S., U. Ligges, and A. Gelman. 2005. R2WinBUGS: a package for running WinBUGS from R. Journal of Statistical Software **12:**1–16.

Thirgood, S. J., S. M. Redpath, D. T. Haydon, P. Rothery, I. Newton, and P. J. Hudson. 2000. Habitat loss and raptor predation: disentangling long-and short-term causes of red grouse declines. Proceedings of the Royal Society of London. Series B: Biological Sciences **267:**651–656.

Traill, L. W., C. J. Bradshaw, and B. W. Brook. 2007. Minimum viable population size: A meta-analysis of 30 years of published estimates. Biological Conservation **139:**159–166.

Van De Pol, M., Y. Vindenes, B.-E. Sæther, S. Engen, B. J. Ens, K. Oosterbeek, and J. M. Tinbergen. 2010. Effects of climate change and variability on population dynamics in a long-lived shorebird. Ecology **91:**1192–1204.

Walters, C., and J. J. Maguire. 1996. Lessons for stock assessment from the northern cod collapse. Reviews in Fish Biology and Fisheries **6:**125–137.

Wasserman, L. 2000. Bayesian model selection and model averaging. Journal of Mathematical Psychology **44:**92–107.

Williamson, K., and R. C. Homes. 1964. Methods and preliminary results of the Common Bird Census, 1962–63. Bird Study **11:**240–256.

Wolf, N., and M. Mangel. 2008. Multiple hypothesis testing and the declining-population paradigm in Steller sea lions. Ecological Applications **18:**1932–1955.