Review

A critical appraisal of population viability analysis

Vratika Chaudhary * and Madan K. Oli

Department of Wildlife Ecology and Conservation, Newins-Zeigler Hall, University of Florida, Gainesville, FL 32611, U.S.A.

Abstract: Population viability analysis (PVA) is useful in management of imperiled species. Applications range from research design, threat assessment, and development of management frameworks. Given the importance of PVAs, it is essential that they be rigorous and adhere to widely accepted guidelines; however, the quality of published PVAs is rarely assessed. We evaluated the quality of 160 PVAs of 144 species of birds and mammals published in peer-reviewed journals from 1990 to 2017. We hypothesized that PVA quality would be lower with generic programs than with custom-built programs; be higher for those developed for imperiled species; change over time; and be higher for those published in journals with high impact factors (IFs). Each included study was evaluated based on answers to an evaluation framework containing 32 questions reflecting whether and to what extent the PVA study adhered to published PVA guidelines or contained important PVA components. All measures of PVA quality were generally lower for studies based on generic programs. Conservation status of the species did not affect any measure of PVA quality, but PVAs published in high IF journals were of higher quality. Quality generally declined over time, suggesting the quantitative literacy of PVA practitioners has not increased over time or that PVAs developed by unskilled users are being published in peer-reviewed journals. Only 18.1% of studies were of high quality (score >75%), which is troubling because poor-quality PVAs could misinform conservation decisions. We call for increased scrutiny of PVAs by journal editors and reviewers. Our evaluation framework can be used for this purpose. Because poor-quality PVAs continue to be published, we recommend caution while using PVA results in conservation decision making without thoroughly assessing the PVA quality.

Keywords: demographic analysis, endangered species, extinction risk, IUCN, population viability analysis, probability of extinction

Una Evaluación Crítica del Análisis de Viabilidad Poblacional

Resumen: El análisis de viabilidad poblacional (AVP) es útil para el manejo de especies en peligro. La gama de aplicaciones incluye el diseño de la investigación, la valoración de amenazas y el desarrollo de marcos de trabajo para el manejo. Ya que los AVP son de suma importancia, es esencial que sean rigorosos y se adhieran a las directrices aceptadas por la mayoría; sin embargo, rara vez se examina. la calidad de los AVP publicados Evaluamos la calidad de 160 AVP para 144 especies de aves y mamíferos publicados en revistas con revisión por pares desde 1990 hasta 2017. Nuestra hipotesis consistió en que la calidad del AVP sería más baja con programas genéricos que con programas hechos a la medida; sería más alta para los programas desarrollados para especies en peligro; la calidad cambiaría con el tiempo; y la calidad sería más alta para los AVP publicados en revistas con un alto factor de impacto (VI). Cada estudio que incluimos fue evaluado con base en las respuestas a un marco de trabajo de evaluación que contenía 32 preguntas, las cuales reflejaban si y cuánto se adherían los AVP a las directrices publicadas para los AVP o si contenía componentes importantes de AVP. Todas las medidas de la calidad de los AVP fueron generalmente más bajas para los estudios basados en programas genéricos. El estado de conservación de las especies no afectó ninguna de las medidas de la calidad de los AVP, pero aquellos publicados en revistas con un VI alto tuvieron una mayor calidad. La calidad, en general, declinó con el tiempo, lo que sugiere que el alfabetismo cuantitativo de quienes practican los AVP no ha incrementado con el tiempo o que se están publicando AVP desarrollados por usuarios con poca práctica en revistas con revisión por pares. Sólo el 18.1% de los estudios fue de calidad alta (puntaje >75%), lo cual es preocupante porque los AVP de baja calidad podrían

*Address correspondence to V. Chaudhary, email chaudharyv@ufl.edu
Article impact statement: *Wildlife managers must exercise caution while using PVA results for conservation planning because poor quality PVA continue to be published in peer-reviewed journals.*
Paper submitted February 21, 2018; revised manuscript accepted June 4, 2019.

Conservation Biology, Volume 34, No. 1, 26–40

DOI: 10.1111/cobi.13414

mal informar las decisiones de conservación. Pedimos un incremento en el escrutinio de los AVP por parte de los editores y revisores. Nuestro marco de trabajo de evaluación puede usarse para este propósito. Ya que todavía se publican AVP con baja calidad, recomendamos que se tomen precauciones cuando se usen los resultados de un AVP en la toma de decisiones de conservación sin evaluar minuciosamente la calidad de dicho estudio.

Palabras Clave: análisis demográfico, análisis de viabilidad poblacional, especie en peligro, población, probabilidad de extinción, riesgo de extinción, UICN

摘要: 种群生存力分析 (population viability analysis, PVA) 是濒危物种管理的有效工具, 其应用范围包括研究设计、威胁评估及管理框架开发等。鉴于 PVA 分析的重要性, 它们应当确保严谨、遵守普遍接受的准则, 然而, 已发表的 PVA 分析的质量却很少得到评估。本研究评估了 1990 至 2017 年间在同行评议期刊上发表的针对144种鸟类及哺乳动物的 160 项 PVA 分析的质量。我们假设自行设定程序的 PVA 质量应比使用通用程序的更高;针对濒危物种的 PVA 质量更高; PVA 质量随时间变化;发表在高影响因子期刊上的 PVA 质量更高。我们基于包含三十二个问题的评估框架评估了每项纳入分析的研究, 这些问题反映了 PVA 研究是否及在多大程度上遵守了已发布的 PVA 指南或包含了重要的 PVA 组成部分。结果表明, 使用通用程序的研究的 PVA 质量在所有指标上都较低, 物种的濒危情况没有影响 PVA 质量, 而发表在高影响因子的期刊上的 PVA 质量更高。此外, PVA 质量普遍随时间推移而下降, 这表明 PVA 实践者的定量推理能力没有与时俱进, 或是同行评议期刊上发表了非熟练使用者开发的 PVA 研究。我们认为只有 18.1 %的研究属于高质量研究 (评分>75%), 这个结果十分令人担忧, 因为低质量的 PVA 分析可能会误导保护决策。因此, 我们呼吁期刊编辑和审稿人加强对 PVA 分析的审查, 而我们的评估框架就可以用于该目的。由于低质量的 PVA 还在持续发表, 我们建议在保护决策中应谨慎使用未经彻底评估质量的 PVA 结果。【**翻译: 胡怡思; 审校: 聂永刚**】

关键词: 种群动态分析, 濒危物种, 灭绝风险, IUCN, 种群生存力分析, 灭绝可能性

Introduction

Population viability analyses (PVAs)—an analysis that uses demographic data in analytical or simulation models to estimate risk faced by species or populations (Ralls et al. 2002)—have become an essential tool in conservation (Akçakaya & Sjögren-Gulve 2000; Beissinger & McCollough 2002; Burgman 2005). They are routinely used to assess conservation status of species and evaluate relative endangerment of species and populations for conservation prioritization (e.g., Wiegand et al. 1998). Additionally, PVAs are used to identify the proximate causes of population decline and potential future threats (e.g., Lunney et al. 2007), identify research priorities, evaluate the efficacy of alternative management strategies (e.g., Hostetler et al. 2013), and identify management actions needed to ensure long-term population persistence. Early recovery plans for threatened or endangered species listed under the U.S. Endangered Species Act (ESA) and listing criteria previously used by International Union for Conservation of Nature and Natural Resource's Red List (IUCN) have been criticized for the lack of objectivity and inconsistency across species (Tear et al. 2005; Neel et al. 2012), and PVAs have been recommended as a solution to this problem (Lindenmayer et al. 1993; Morris et al. 2002; Doak et al. 2015). The number of PVAs published in peer-reviewed journals (and presumably their use) has increased over the last 3 decades (Supporting Information). Recently, PVA results have been used as objective and SMART (specific, measurable, achievable, realistic, and time referenced) criteria for developing species recovery plans (e.g., red wolf [*Canis rufus*] [Faust et al. 2016]; West Indian manatee [*Trichechus manatus latirostris*] [Runge et al. 2017]).

The increase in use of PVAs to guide species recovery plans can be attributed partially to the availability of software packages (generic programs), such as Vortex (Lacy 1993), RAMAS series (Akçakaya & Root 2002), and ALEX (Possingham & Davies 1995). These programs offer a user-friendly interface and require little programing input from users. Furthermore, they can be used with minimal training in population ecology and mathematical or statistical modeling. Some generic programs (e.g., Vortex) are free and widely accessible. These features have made them popular, but in some cases they have facilitated inappropriate use of the generic programs (Reed et al. 2002; Beier et al. 2003).

Theory and application of PVA have recently received increased scrutiny concomitant with an increased recognition of the role of PVAs in conservation decision making. With the broader goal of identifying when and how PVAs should be used, attempts have been made to unambiguously establish what constitutes a valid PVA. What components should a study contain and what analyses should be performed for it to be considered a PVA? What outputs should a study report so that PVA results can be compared across studies or species? Reviews of the existing PVA literature identify several problems with the use of PVA results in conservation decision making (e.g., Beissinger & Westphal 1998; Burgman & Possingham 2000; Morris et al. 2002; Pe'er et al. 2013), for example, inadequate background information about study species or articulation of study objectives; inadequate description of data sources, period, and methods of data collection;

failure to estimate demographic parameters with statistically robust methods or to clearly describe how demographic parameters were estimated; use of model structures that inadequately capture the life history of study species (e.g., age-structured models sometimes used to model dynamics of stage-structured populations); failure to incorporate parametric uncertainty (which can substantially affect model results); inadequate justification of the choice of demographic model structure and modeling platform; failure to consider factors that can profoundly influence population dynamics, such as stochasticity, density dependence, extrinsic threats, and management actions, or to adequately justify their exclusion; failure to perform perturbation analyses involving population growth rate and extinction parameters; inadequate discussion of extinction threshold (e.g., critical population size for quasi extinction) and projection intervals relative to the study species' generation time; failure to report variances (or other measures of precision) of population growth rate and extinction parameters; failure to test predictive accuracy of PVA models (with new field data or through parsing data); failure to adequately discuss model or data limitations; and management recommendations not supported by PVA results (e.g., management recommendations based on studies with purely heuristic objectives).

A persistent challenge to the application of PVAs in conservation is the lack of consistency across studies in terms of how PVA models are formulated or implemented and how the results are interpreted. These inconsistencies have led to substantial variation in PVA quality (defined in Methods), with some PVA studies lacking vital PVA components. Because of an increasingly important role of PVAs in conservation decision making, the quality of PVA studies needs evaluation so that conservation professionals can make informed decisions regarding the usefulness of PVA results in the formulation or implementation of conservation policies. Yet, such an assessment has not been made, and a comprehensive assessment framework has not been developed.

We evaluated the quality of PVAs for mammals and birds published from 1990 to 2017 by assessing the extent to which authors incorporated important PVA components (Table 1) and followed published PVA guidelines. We developed an assessment framework to objectively assess the background-information, model, and analysis quality and overall PVA quality (Supporting Information). We expected that the PVA quality would be higher for threatened species (listed as threatened, vulnerable, endangered, or critically endangered by IUCN or comparable national lists [e.g., under ESA]) than nonthreatened species due to the research and conservation focus afforded them; would change over time because of temporal changes in ecological and quantitative literacy; would be higher for studies published in high impact than low impact factor (IF) journals because manuscripts submitted to high IF journals presumably receive more rigorous reviews; and would be lower with generic programs than custom-built programs because generic programs can facilitate inappropriate program use by users with limited PVA experience.

Methods

Literature Search and Study Selection and Evaluation

To locate relevant studies, we searched for published articles on Web of Science (www.clarivate.com) and Google Scholar (www.scholar.google.com) databases with ≥ 1 of the following keywords: *demographic models, extinction probability, time to extinction, persistence probability, PVA, population viability analysis, population extinction,* and *stochastic population model*. Additional literature was sourced from the references cited in these publications. We selected an article if the study species was a bird or mammal; was published in a peer-reviewed journal from 1990 to 2017; and at least one of the key extinction parameters (probability of extinction, mean or median time to extinction, or distribution of extinction time) was reported. For the 160 published studies that met these criteria, we recorded year of publication, IF of journal, modeling platform used (generic or custom built), and whether the study species was listed as threatened on the IUCN Red List (IUCN 2017) or a similar national list. Journal IFs were obtained from the Clarivate Analytics website for 2016 (www.clarivate.com). In cases where the journal was discontinued, IF of the last year of publication was used. It is possible that we missed some relevant studies despite our best efforts.

Study evaluation was based on answers to 32 questions reflecting whether and to what extent a PVA study adhered to published PVA guidelines or contained essential PVA components (Supporting Information). We divided these questions into 3 categories—quality of background information (4 questions), model (13), and analysis (15) (details below)—to assess how each study addressed key PVA components. Each question could be answered with *yes* (score 1) or *no* (score 0). A *yes* response indicated that either the essential component was included in the study or its exclusion was addressed and adequately justified. A *no* response indicated the study neither included the PVA component nor adequately justified its exclusion.

Background-Information Quality

Background-information quality (hereafter, background quality) was evaluated based on the proportion of *yes* responses to questions 1–4 (Supporting Information) designed to evaluate whether the study provided sufficient background information about the life history of the study species because the reader would not be able to objectively evaluate PVA results without this information;

Table 1. Essential components of population viability analysis (PVA) and logic behind their inclusion.

Component	*Logic*	*Reference*
Objectives	PVAs address specific questions or objectives that must be described clearly so that results can be interpreted in the context of the study objectives.	Grant 1986; Boyce 1992
Demographic data	Adequate data are necessary to estimate demographic variables with sufficient accuracy and precision and encompass the range of environmental variation. However, PVAs of data-sparse species conducted with the objectives of guiding future research have intrinsic value.	Beissinger & Westphal 1998; Coulson et al. 2001; Reed et al. 2002
Knowledge of biology and life history	Sufficient information on the species biology and life history is necessary for meaningful interpretation of PVA modeling results.	Boyce 1992; Beissinger & Westphal 1998
Model structure	PVA model structure should adequately capture life history of the study species and be described clearly; model structure should be appropriate for best-available data.	Beissinger & Westphal 1998; Beissinger & McCollough 2002; Reed et al. 2002
Stochasticity	Demographic, environmental, and genetic stochasticity and catastrophic events can affect population dynamics and persistence and should be incorporated in PVA.	Boyce 1992; Ralls et al. 2002; Reed et al. 2002; Pe'er et al. 2013
Density dependence	Density-dependence generally reduces extinction risk by stabilizing the population and should be included when possible.	Boyce 1992; Ralls et al. 2002; Henle et al. 2004
Extrinsic factors	Factors, such as habitat loss, poaching, and disease, can affect population dynamics and persistence and drive populations to extinction. Simulation of management scenarios is important to assess the efficacy of alternative management scenarios and should be applied when possible.	Boyce 1992; Reed et al. 2002
Definition of extinction parameters	Clear definition of probability of extinction or quasiextinction threshold is essential for informed use of PVA results.	Ralls et al. 2002; Pe'er et al. 2013
Time horizon	Choice of endpoint of time (time interval) for simulation-based studies can affect PVA results. Choice of time horizon should be appropriate for the generation time of the study species and should be reported. Reporting results based on multiple time horizons may be helpful in some cases.	Ralls et al. 2002; Pe'er et al. 2013
Means and variances of extinction parameters	Point estimates and measures of precision (CI and variances) of population growth rate and extinction parameters should be reported.	Boyce 1992; Ellner et al. 2002
Perturbation analysis	Understanding the absolute and proportional sensitivity of population growth rate and extinction parameters to vital demographic rates is essential for planning future research and management. Therefore, perturbation analysis should be an integral part of PVAs.	Boyce 1992; Ralls et al. 2002; Reed et al. 2002
Validation	Model validation by comparing PVA-predicted population sizes with observed sizes should be conducted to evaluate predictive accuracy of the PVA model when possible.	Beissinger & Westphal 1998; Ralls et al. 2002

clearly specified the study objectives, which is essential because PVA results must be interpreted relative to study objectives; and clearly explained quantity and quality of data and period of data collection because PVA results depend on the data used to parameterize the model. When input parameters were sourced from other publications, we reviewed the source publications to answer the questions related to the description of data quantity and quality and parameter-estimation methods.

Model Quality

Model quality was evaluated based on the proportion of *yes* responses to questions 5–17 (Supporting Information) designed to assess whether the model structure adequately reflected the life history of the study species; input parameters were estimated using statistically robust analysis of the best available data set; parametric uncertainty was appropriately incorporated (White 2000); influence of environmental, demographic, and genetic stochasticity was included (Lande 1993); influence of density dependence was included (many populations may be density regulated, and ignoring density dependence when it is operating in a population can lead to incorrect estimates of extinction parameters); and extrinsic threats or management scenarios were included. These factors are important because extrinsic threats (e.g., habitat loss and poaching) can influence population dynamics and persistence and objective evaluation of alternative management scenarios (e.g., population supplementation) can inform conservation planning.

Analysis Quality

Analysis quality was evaluated based on the proportion of *yes* responses to questions 18–32 (Supporting Information) designed to assess whether perturbation (or sensitivity) analyses involving population growth or extinction parameters were conducted. Proportional or absolute sensitivity analyses provide important insights into the factors that affect population growth or extinction parameters. These results can guide future research or management efforts (McCarthy et al. 1995). They were also used to assess whether alternative scenarios were evaluated and results of all scenarios were reported; a clear definition of extinction threshold was provided because the usefulness of PVA studies in species conservation planning is difficult to evaluate without this information; estimates of population growth rate and ≥1 extinction parameter were reported; variances of population growth rate and extinction parameters (or other measures of precision) were reported because this information is necessary to gauge the precision of these estimates; predictive accuracy of PVA model was tested by parsing data or using time-series data from the subsequent years or similar population; and data and model limitations were discussed clearly because data are rarely sufficient to permit perfect analyses with adequate precision. A PVA model only outlines population processes and can never capture the complexity of nature perfectly. Thus, it is necessary to delineate these limitations so that readers can make informed decisions.

Overall Quality

Overall quality was assessed based on the proportion of *yes* responses to all 32 questions (Supporting Information). We defined *quality* as the proportion of *yes* responses in each category. Thus, quality ranged from 0.0 (poor-quality study) to 1.0 (perfect study; answers to all questions were *yes*). We considered studies scoring >0.75 to be of high quality and studies scoring <0.50 of poor quality. The studies with scores of 0.50−0.75 were considered of average quality.

Questions in the evaluation framework (Supporting Information) were designed such that they were weighted equally. If a study did not consider a PVA component discussed above but provided an adequate justification for its exclusion, it was not penalized. For example, if authors explained that they ignored demographic stochasticity because the starting population size was in the 1000s, the study received a score of 1 for that question. Likewise, if a study contained statistical evidence that the size of the study population was too small for the manifestation of density-dependent effects and adequately addressed the issue, it was not penalized for excluding density dependence.

Statistical Analyses

We used generalized linear mixed models with binomial distribution (Zuur et al. 2009; Agresti 2015) to test for the effect on PVA quality of the following covariates: conservation status of study species (threatened vs. not), year of publication (continuous variable, 1990–2017), journal IF (continuous variable, 0.31−37.02), and modeling platform (generic vs. custom built). We tested for singular and additive effects of these covariates and the relevant 2-way interactive effects. Our sample included instances of ≥1 PVAs of the same species and ≥1 studies by the same authors (Supporting Information). To account for potential lack of independence of studies by the same authors or on the same species, we included the random effect of study species and the first author of the study. We used an information-theoretic approach Akaike information criterion to perform model selection and for statistical inference (Burnham & Anderson 1998). Covariate effects were assessed by comparing models with and without the covariate and determining whether 90% CI for the slope parameter overlapped 0 (in case of categorical variables, the odds ratio overlapped 1.0). We used Hosmer–Lemeshow test (Hosmer & Lemeshow 1980) to assess goodness of fit of the saturated model (Burnham & Anderson 1998). Marginal R^2, which quantifies the proportion of variance explained by fixed effects (Nakagawa & Schielzeth 2013; Johnson 2014), was used as a measure of the fit of the top model.

We fitted generalized linear mixed effect models with the lme4 package (Bates et al. 2015), implemented the Hosmer–Lemeshow goodness-of-fittest with generalhoslem package (Matthew 2017), and calculated marginal R^2 with MuMIn package (Barton 2013) in R computing environment (R Core Team 2017). To account for model-selection uncertainties, we conducted model averaging to estimate model-averaged parameters and their unconditional variances (Burnham & Anderson 1998) with MuMIn package (Barton 2013).

Out of 160 total studies evaluated, 155 claimed to be PVAs, whereas the rest were designed to assess population-level impact of climate change (Hunter et al. 2010), population dynamics under alternative management scenarios (Taylor et al. 2008), or population dynamics of the study species (Caswell et al. 1999). These 5 studies were included here because they met our selection criteria and results of some of these studies have been used in conservation planning (e.g., Hunter et al. 2010). Because they did not claim to be PVAs, we evaluated the ranking of these studies by modifying guiding questions that were specific to PVAs but did not necessarily apply to general demographic studies (Supporting Information).

Results

We evaluated 61 studies of birds and 99 studies of mammals (144 species) published in 154 peer-reviewed articles. These studies reported results for 56 birds and 88 mammal species (Supporting Information). Most studies (73.8%, all percentages are based on $n = 160$) focused on threatened species, clearly articulated the study objectives (66.2%), and described the period and methods of data collection (73.1%). The majority of studies used an individual-based modeling (IBM) framework; unstructured model based on count data was the least commonly used modeling approach (IBM 45.6%; structured matrix population models 45.0% and unstructured models 9.4%). The majority of the custom-built models were structured matrix population models (64.0%); 21.8% used diffusion-approximation models and 14.2% used individual-based models. Vortex was the most commonly used (73.9%), and unified life model (0.01%) was the least commonly used generic program (Supporting Information). The majority of studies incorporated environmental (83.8%) and demographic (84.4%) stochasticity, but few incorporated genetic stochasticity (25.0%). The percentage of studies that included the influence of (or adequately justified exclusion of) catastrophes, density dependence, potential management scenarios, and extrinsic threats was 42.5%, 50.6%, 35.6%, and 56.2%, respectively. The majority of the studies reported the population growth rate (82.5%), probability of extinction (85.6%), and mean time to extinction (45.6%). Variance (or other measures of precision) of population growth rate, probability of extinction, and mean time to extinction were reported by 63.7%, 50.6%, and 28.1% of the studies, respectively. Most studies (91.8%) included sensitivity analysis involving at least one extinction parameter.

Background Quality

The background quality was lower for studies based on generic programs (odds ratio [OR] = 0.45, 90% CI 0.30–0.58). It was higher for studies published in high IF journals (slope parameter: β [SE] = 0.09 [0.04], 90% CI 0.01–0.6) and it decreased over time ($\beta = -0.03$ [0.01], 90% CI −0.05 to −0.01). The background quality was not affected by conservation status of the study species (OR = 1.10, 90% CI 0.73–1.66) (Supporting Information).

Hosmer–Lemeshow test provided no evidence of lack of fit of the saturated model ($\chi^2 = 4.81$, df = 8, $p = 0.78$). The most parsimonious model (Table 2, background quality) included an interactive effect of the journal IF and the use of generic programs (Fig. 1). According to this model, background quality of the PVAs based on generic programs was higher for studies published in high IF journals ($R^2 = 0.10$).

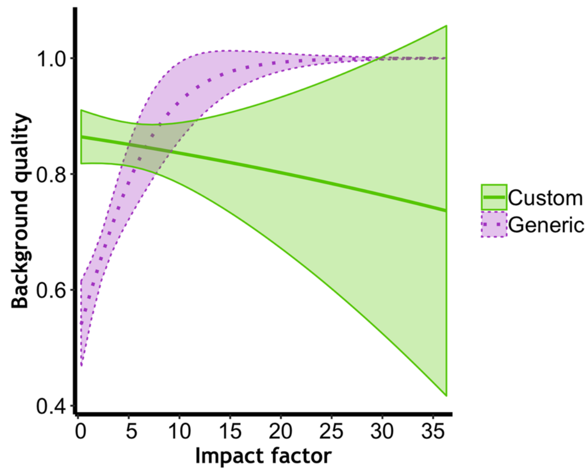

Figure 1. Interactive effect of journal impact factor (IF) and modeling platform (generic or custom built) on background-information quality as inferred from the most parsimonious model (background-quality model, Table 3). Journal IFs range from 0.31 (Great Basin Naturalist renamed Western North American Naturalist) to 37.02 (Science). The highest IF just below Science was 9.661 (Proceedings of National Academy of Sciences of the United States of America).

Model Quality

Model quality was lower for studies based on generic programs (OR = 0.78, 90% CI 0.65–0.94). It was not affected by the IF of the journal (β [SE] = 0.01 [0.01], 90% CI −0.01 to 0.04) or conservation status of the study species (OR = 1.1.2, 90% CI 0.91–1.38) and did not change substantially over time ($\beta = -0.004$ [0.006], 90% CI −0.01 to 0.007) (Supporting Information).

Hosmer–Lemeshow test provided no evidence of lack of fit of the saturated model ($\chi^2 = 3.41$, df = 8, $p = 0.91$). The most parsimonious model (Table 2, model quality) included the use of generic programs and random effect of author ($R^2 = 0.05$). Model quality was low for studies based on generic programs (OR = 0.76, 90% CI 0.62–0.93). A competing model (Table 2, model quality) included an interactive effect of year of publication and modeling platform (Table 2, model quality); however, 90% CIs for all model parameters included 0 (Table 3).

Analysis Quality

Analysis quality was lower for studies based on generic programs (OR = 0.67, 90% CI 0.56–0.79) than custom-built programs; decreased over time (β [SE] = -0.03 [0.006], 90% CI -0.04 to -0.02); and was higher for studies published in high IF journals ($\beta = 0.05$ [0.01], 90% CI 0.02–0.08) than for studies published in low IF journals.

Table 2. Results of model comparison to test the effect of journal impact factor (IF), year of the publication (year), use of generic programs versus custom programs, and conservation status of the species (threatened vs. not threatened) on the probability of high-quality background information, model, analysis, and overall population viability analysis (PVA) quality.[a]

Model	*K*[b]	*AICc*[c]	*ΔAICc*[d]	*Weight*[e]	*Log likelihood*
Background quality					
IF * generic + (1 ∣ species)	5	395.70	0.00	0.20	−192.60
IF * generic + (1 ∣ author)	5	395.90	0.20	0.20	−192.70
year + IF * generic + (1 ∣ species)	6	396.50	0.90	0.10	−192.00
year + IF * generic + (1 ∣ author)	6	396.60	0.90	0.10	−192.00
threatened + IF * generic + (1 ∣ species)	6	397.60	2.00	0.10	−192.50
threatened + IF * generic + (1 ∣ author)	6	397.80	2.20	0.10	−192.60
year + threatened + IF * generic + (1 ∣ species)	7	398.60	2.90	0.00	−191.90
year + threatened + IF * generic + (1 ∣ author)	7	398.60	2.90	0.00	−191.90
generic + year + (1 ∣ author)	4	399.70	4.00	0.00	−195.70
generic + year + (1 ∣ species)	4	400.10	4.40	0.00	−195.90
Model quality					
year * generic + (1 ∣ author)	5	662.90	0.00	0.20	−326.20
generic + (1 ∣ author)	3	662.90	0.00	0.20	−328.40
threatened + year * generic + (1 ∣ author)	6	664.40	1.50	0.10	−325.90
threatened + generic + (1 ∣ author)	4	664.50	1.60	0.10	−328.10
IF + year * generic + (1 ∣ author)	6	664.60	1.70	0.10	−326.00
generic + IF + (1 ∣ author)	4	664.60	1.70	0.10	−328.20
generic + year + (1 ∣ author)	4	664.70	1.80	0.10	−328.20
IF * generic + (1 ∣ author)	5	665.60	2.70	0.10	−327.60
IF + (1 ∣ author)	3	665.70	2.90	0.00	−329.80
IF + generic + threatened + (1 ∣ author)	5	666.10	3.20	0.00	−327.80
Analysis quality					
generic + year + (1 ∣ author)	4	754.50	0.00	0.20	−373.10
year + generic + threatened + (1 ∣ author)	5	755.70	1.30	0.10	−372.70
year * generic + (1 ∣ author)	5	756.00	1.50	0.10	−372.80
IF + generic + year + (1 ∣ author)	5	756.20	1.80	0.10	−372.90
year + (1 ∣ author)	3	756.80	2.40	0.10	−375.30
generic + (1 ∣ author)	3	757.00	2.50	0.10	−375.40
threatened + year * generic + (1 ∣ author)	6	757.30	2.90	0.10	−372.40
IF + generic + threatened + year + (1 ∣ author)	6	757.50	3.00	0.00	−372.50
IF + year * generic + (1 ∣ author)	6	757.90	3.40	0.00	−372.70
year + IF + (1 ∣ author)	4	757.90	3.50	0.00	−374.80
Overall quality					
generic + year + (1 ∣ author)	4	874.10	0.00	0.20	−432.90
year + generic + threatened + (1 ∣ author)	5	874.90	0.80	0.10	−432.20
IF + generic + year + (1 ∣ author)	5	875.10	1.00	0.10	−432.40
year + IF * generic + (1 ∣ author)	6	875.70	1.60	0.10	−431.60
IF + generic + threatened + year + (1 ∣ author)	6	875.70	1.70	0.10	−431.60
year * generic + (1 ∣ author)	5	875.80	1.70	0.10	−432.70
threatened + year * generic + (1 ∣ author)	6	876.50	2.40	0.00	−432.00
year + threatened + IF * generic + (1 ∣ author)	7	876.60	2.50	0.00	−430.90
IF + year * generic + (1 ∣ author)	6	876.70	2.60	0.00	−432.10
generic + (1 ∣ author)	3	876.70	2.60	0.00	−435.30

[a] *Ten best-supported models presented: +, additive effect; *, additive and interactive effect; (1 ∣ species), random effect of species; (1 ∣ author), random effect of the author.*
[b] *Number of parameters.*
[c] *Akaike information criterion corrected for small sample size.*
[d] *Measure of each model compared with the best model; calculated as a value of AICc – minimum AICc.*
[e] *Weight of the model.*

Conservation status of the study species did not influence the analysis quality (OR = 1.15, 90% CI 0.95–1.40).

Hosmer–Lemeshow test provided no evidence of lack of fit of the saturated model ($\chi^2 = 4.67$, df = 8, $p = 0.79$). The most parsimonious model (Table 2, analysis quality) included an additive effect of the year of publication and modeling platform (Table 2, analysis quality). This model suggested that the overall PVA quality was generally lower for studies based on generic programs and it decreased over time, irrespective of modeling platform (Fig. 2) ($R^2 = 0.09$).

Overall Quality

Overall quality (Table 2, overall quality) increased as journal IF increased (β [SE] = 0.03 [0.01], 90% CI 0.02–0.05)

Table 3. Estimates of regression coefficients (β) based on well-supported models (i.e., models with $\Delta AIC_c < 2$) for quality of background information, model, analysis and, overall PVA quality to assess the effect of covariates impact factor of journal (IF), year of publication (year), use of generic versus custom programs (generic), and conservation status of the species (threatened or not threatened) and the interactive effect (colon) between IF and generic and year and generic.

Model		β^a	*SE*[b]	*p*	*LCI*[c]	*UCI*[d]
Background quality						
1	intercept	1.62	0.18	0.00	1.33	1.94
	IF	−0.05	0.11	0.63	−0.22	0.15
	generic (yes)	−0.65	0.23	0.00	−1.04	−0.28
	IF: generic (yes)	0.88	0.31	0.01	0.37	1.41
	random effect (species) SD	0.24	NA	NA	0.00	0.57
2	intercept	1.61	0.19	0.00	1.32	1.93
	IF	−0.05	0.11	0.67	−0.21	0.17
	generic (yes)	−0.65	0.23	0.00	−1.04	−0.28
	IF: generic (yes)	0.86	0.31	0.01	0.35	1.37
	random effect (author) SD	0.16	NA	NA	0.00	0.58
3	intercept	1.59	0.18	0.00	1.30	1.91
	year	−0.12	0.10	0.25	−0.28	0.05
	IF	−0.06	0.11	0.55	−0.23	0.13
	generic (yes)	−0.62	0.23	0.01	−1.01	−0.24
	IF: generic (yes)	0.83	0.31	0.01	0.33	1.36
	random effect (species) SD	0.20	NA	NA	0.00	0.56
4	intercept	1.59	0.19	0.00	1.30	1.91
	year	−0.12	0.10	0.24	−0.30	0.04
	IF	−0.06	0.11	0.60	−0.22	0.15
	generic (yes)	−0.62	0.23	0.01	−1.02	−0.25
	IF: generic (yes)	0.82	0.31	0.01	0.31	1.33
	random effect (author) SD	0.19	NA	NA	0.00	0.58
5	intercept	1.54	0.25	0.00	1.15	1.97
	threatened (yes)	0.10	0.22	0.66	−0.27	0.46
	IF	−0.05	0.11	0.67	−0.22	0.16
	generic (yes)	−0.65	0.23	0.00	−1.04	−0.28
	IF: generic (yes)	0.87	0.31	0.01	0.36	1.40
	random effect (species) SD	0.24	NA	NA	0.00	0.58
6	intercept	1.54	0.25	0.00	1.15	1.97
	threatened (yes)	0.10	0.22	0.66	−0.27	0.46
	IF	−0.05	0.11	0.67	−0.22	0.16
	generic (yes)	−0.65	0.23	0.00	−1.04	−0.28
	IF: generic (yes)	0.87	0.31	0.01	0.36	1.40
	random effect (species) SD	0.24	NA	NA	0.00	0.58
Model quality						
1	intercept	0.77	0.10	0.00	0.59	0.92
	year	0.11	0.10	0.25	−0.09	0.13
	generic (yes)	−0.25	0.13	0.05	−0.42	0.03
	year: generic (yes)	−0.26	0.13	0.04	−0.11	0.52
	random effect (author) SD	0.40	NA	NA	0.30	0.54
2	intercept	0.76	0.10	0.00	0.60	0.93
	generic (yes)	−0.26	0.13	0.04	−0.47	−0.06
	random effect (author) SD	0.43	NA	NA	0.31	0.55
3	intercept	0.69	0.14	0.00	0.46	0.93
	threatened (yes)	0.11	0.13	0.42	−0.11	0.33
	year	0.12	0.10	0.24	−0.05	0.28
	generic (yes)	−0.25	0.12	0.05	−0.46	−0.04
	year: generic (yes)	−0.26	0.13	0.04	−0.48	−0.05
	random effect (author) SD	0.39	NA	NA	0.27	0.52
4	intercept	0.68	0.14	0.00	0.45	0.92
	threatened (yes)	0.10	0.14	0.45	−0.12	0.33
	generic (yes)	−0.26	0.13	0.04	−0.47	−0.06
	random effect (author) SD	0.42	NA	NA	0.30	0.55
5	intercept	0.76	0.10	0.00	0.60	0.93
	IF	0.04	0.06	0.50	−0.06	0.16
	year	0.12	0.10	0.22	−0.04	0.29
	generic (yes)	−0.22	0.13	0.08	−0.44	−0.01
	year: generic (yes)	−0.27	0.13	0.04	−0.48	−0.05
	random effect (author) SD	0.40	NA	NA	0.27	0.53

Continued

Table 3. Continued.

Model		*β[a]*	*SE[b]*	*p*	*LCI[c]*	*UCI[d]*
6	intercept	0.76	0.10	0.00	0.60	0.93
	IF	0.04	0.06	0.50	−0.06	0.16
	year	0.12	0.10	0.22	−0.04	0.29
	generic (yes)	−0.22	0.13	0.08	−0.44	−0.01
	year: generic (yes)	−0.27	0.13	0.04	−0.48	−0.05
	random effect (author) SD	0.40	NA	NA	0.27	0.53
7	intercept	0.75	0.10	0.00	0.58	0.92
	generic (yes)	−0.24	0.13	0.07	−0.45	−0.02
	IF	0.04	0.06	0.51	−0.06	0.15
	random effect (author) SD	0.43	NA	NA	0.31	0.55
8	intercept	0.75	0.10	0.00	0.59	0.92
	generic (yes)	−0.25	0.13	0.05	−0.46	−0.04
	year	−0.04	0.07	0.55	−0.15	0.07
	random effect (author) SD	0.43	NA	NA	0.31	0.55
9	intercept	0.75	0.10	0.00	0.59	0.92
	IF	0.02	0.07	0.80	−0.09	0.13
	generic (yes)	−0.19	0.14	0.16	−0.42	0.03
	IF: generic (yes)	0.20	0.19	0.29	−0.11	0.52
	random effect (author) SD	0.42	NA	NA	0.30	0.54
Analysis quality						
1	intercept	0.62	0.11	0.00	0.45	0.80
	generic (yes)	−0.28	0.13	0.04	−0.50	−0.06
	year	−0.15	0.07	0.03	−0.27	−0.04
	random effect (author) SD	0.56	NA	NA	0.45	0.67
2	intercept	0.52	0.15	0.00	0.28	0.77
	year	−0.15	0.07	0.03	−0.27	−0.04
	generic (yes)	−0.28	0.13	0.03	−0.50	−0.06
	threatened (yes)	0.14	0.14	0.34	−0.10	0.38
	random effect (author) SD	0.55	NA	NA	0.45	0.67
3	intercept	0.62	0.11	0.00	0.45	0.79
	year	−0.21	0.11	0.05	−0.40	−0.03
	generic (yes)	−0.28	0.13	0.04	−0.50	−0.06
	year: generic (yes)	0.10	0.14	0.47	−0.13	0.34
	random effect (author) SD	0.55	NA	NA	0.45	0.67
4	intercept	0.62	0.11	0.00	0.44	0.79
	IF	0.03	0.06	0.59	−0.07	0.13
	generic (yes)	−0.26	0.14	0.06	−0.48	−0.04
	year	−0.15	0.07	0.04	−0.27	−0.03
	random effect (author) SD	0.56	NA	NA	0.45	0.67
5	intercept	0.45	0.07	0.00	0.35	0.56
	year	−0.18	0.07	0.01	−0.29	−0.06
	random effect (author) SD	0.58	NA	NA	0.47	0.69
Overall quality						
1	intercept	0.77	0.07	0.00	0.65	0.89
	generic (yes)	−0.32	0.09	0.00	−0.47	−0.16
	year	−0.11	0.05	0.03	−0.19	−0.03
	random effect (author) SD	0.40	NA	NA	0.33	0.48
2	intercept	0.68	0.11	0.00	0.51	0.86
	year	−0.11	0.05	0.03	−0.19	−0.03
	generic (yes)	−0.32	0.09	0.00	−0.48	−0.17
	threatened (yes)	0.12	0.10	0.24	−0.05	0.29
	random effect (author) SD	0.40	NA	NA	0.33	0.47
3	intercept	0.76	0.08	0.00	0.63	0.88
	IF	0.05	0.04	0.29	−0.02	0.12
	generic (yes)	−0.29	0.10	0.00	−0.45	−0.14
	year	−0.10	0.05	0.04	−0.18	−0.02
	random effect (author) SD	0.40	NA	NA	0.33	0.48

Continued

Table 3. Continued.

Model		β^a	SE^b	p	LCI^c	UCI^d
4	intercept	0.76	0.08	0.00	0.64	0.89
	year	−0.09	0.05	0.06	−0.18	−0.01
	IF	0.03	0.04	0.53	−0.04	0.10
	generic (yes)	−0.25	0.10	0.01	−0.42	−0.09
	IF: generic (yes)	0.18	0.14	0.21	−0.06	0.41
	random effect (author) SD	0.40	NA	NA	0.32	0.47
5	intercept	0.66	0.11	0.00	0.49	0.84
	IF	0.05	0.04	0.26	−0.02	0.12
	generic (yes)	−0.30	0.10	0.00	−0.46	−0.14
	threatened (yes)	0.13	0.10	0.22	−0.04	0.29
	year	−0.10	0.05	0.04	−0.18	−0.02
	random effect (author) SD	0.39	NA	NA	0.32	0.47
6	intercept	0.66	0.11	0.00	0.49	0.84
	IF	0.05	0.04	0.26	−0.02	0.12
	generic (yes)	−0.30	0.10	0.00	−0.46	−0.14
	threatened (yes)	0.13	0.10	0.22	−0.04	0.29
	year	−0.10	0.05	0.04	−0.18	−0.02
	random effect (author) SD	0.39	NA	NA	0.32	0.47
7	intercept	0.69	0.11	0.00	0.51	0.86
	threatened (yes)	0.12	0.10	0.23	−0.05	0.29
	year	−0.07	0.08	0.38	−0.19	0.06
	generic (yes)	−0.32	0.09	0.00	−0.47	−0.16
	year: generic (yes)	−0.07	0.10	0.48	−0.23	0.09
	random effect (author) SD	0.40	NA	NA	0.32	0.47

[a] *Slope parameter estimate.*
[b] *Standard error of slope parameter.*
[c] *Lower value of 90% CI of the slope parameter.*
[d] *Upper value of 90% CI of the slope parameter.*

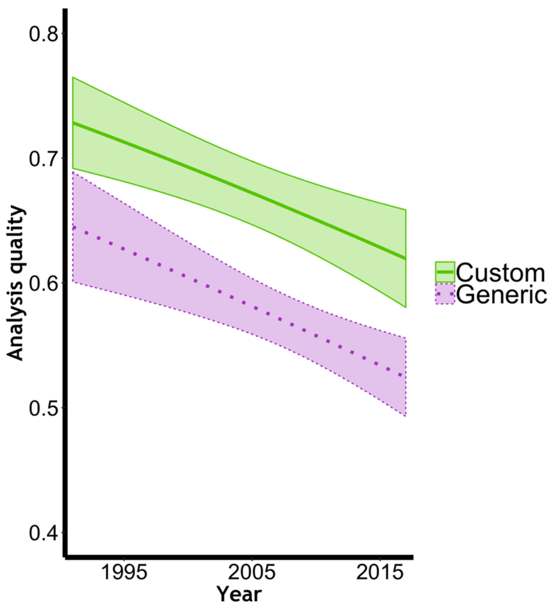

Figure 2. Additive effect of the year of publication and modeling platform (generic vs. custom built) on analysis quality as inferred from the most parsimonious model (analysis quality, Table 3).

and was lower for studies developed using generic programs than custom-built programs (OR = 0.69, 90% CI 0.61–0.77). Overall quality was unaffected by the conservation status of the study species (OR = 1.13, 90% CI 0.99–1.30) and decreased over time (β = -0.02 [0.001], 90% CI -0.02 to -0.01).

Hosmer–Lemeshow test provided no evidence of lack of fit of the saturated model (χ^2 = 2.19, df = 8, p = 0.97). The most parsimonious model describing the overall quality included an additive effect of the use of generic program and year of publication (Table 2). This model suggested that the overall quality of studies based on generic programs was lower than for custom-built programs and declined over time, irrespective of the modeling platform (Fig. 3) (R^2 = 0.05).

There was substantial model-selection uncertainty; several models were within ΔAIC<2 for all measures of PVA quality (Table 2). Marginal R^2 values were low even for the best-supported models. Model-averaged results generally agreed with those based on the respective top-ranked models (Supporting Information).

Ranking of PVA Studies Based on Quality

The percentage of studies with a quality score >0.75 was 31.8%, 30.0%, and 25.0% (n = 160) for background, model, and analysis quality, respectively. Only 18.1% (n = 160) of the studies had overall quality score > 0.75 (Supporting Information). The percentage of studies that scored < 0.50 was 8.7%, 18.7%, 25.0%, and 14.3%

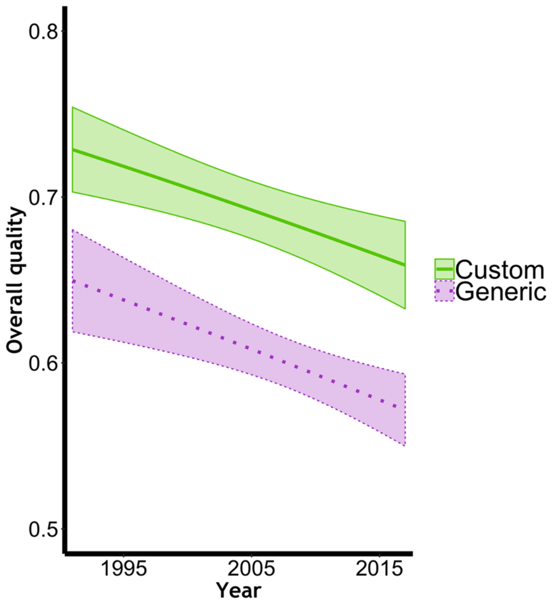

Figure 3. Additive effect of the year of publication and modeling platform (generic vs. custom built) on overall PVA quality inferred from the most parsimonious model (overall quality, Table 3).

for background, model, analysis, and overall quality, respectively.

The top-performing study in overall quality was Wiegand et al. (1998), and the second-best score was shared among Gaona et al. (1998), Bakker et al. (2009), and Hostetler et al. (2013). Among studies based on generic programs, Forys and Humphrey (1999) and Slotta-Bachmayr et al. (2004) were the top performers. The top 5 studies (based on overall quality score) used custom-built models, whereas the 5 lowest-scoring studies used generic programs.

Discussion

Limited resources available for the management of threatened species necessitate that conservation decisions be made based on sound quantitative assessment of the relative or absolute risk faced by the species (Doak et al. 2015). Population viability analysis can be used to objectively assess the absolute or relative viability of species or populations, to identify proximate causes of population declines, compare management alternatives, and to optimize further research (Himes Boor 2014; Doak et al. 2015). Therefore, it is essential for readers to be able to assess the quality and reliability of results provided by PVAs in light of the broadly accepted PVA guidelines. Our goal was to objectively evaluate the quality of PVAs of birds and mammals published since 1990 and to test for the influence of modeling platform, year of publication, publication journal IF, and conservation status of the study species on the PVA quality.

Quality of PVA Studies Based on Generic Programs

Generic PVA programs, such as Vortex and RAMAS, were developed by scientists who are experts in the theory and application of PVA. These programs are powerful tools when used by modelers with adequate knowledge of study species' life history, population modeling, and PVA guidelines. Recent versions of generic programs, such as Vortex (http://www.vortex10.org/Vortex10.aspx), now provide sufficient flexibility to allow customization for some aspects of PVA. However, concerns have been raised regarding the quality of PVA studies based on generic programs (Reed et al. 2002; Beier et al. 2003) because even users with little experience of demographic modeling can easily use these programs. Therefore, we expected the PVA quality to be lower for studies based on generic programs. Consistent with this expectation, background, model, analysis, and overall quality of PVA studies based on generic programs were lower compared with those that were custom-built. The lower quality of PVAs based on generic programs may be a consequence of several factors, including their relative ease of use. Performing PVAs with generic programs requires little ecological or statistical knowledge or programming experience. This lack of relevant knowledge can adversely affect PVA quality (Burgman & Possingham 2000; Reed et al. 2002). Generic programs are sometimes used for PVAs even when the structure of the model or models offered by the software is inconsistent with study species' life history or when the model cannot be adequately parameterized with available data. Generic programs, such as Vortex, provide default parameter values, which permit analyses even when data are insufficient for a meaningful PVA.

One potential advantage of generic programs is that the studies developed using these programs should be reproducible. However, Morrison et al. (2016) found that 40% of the PVAs based on generic programs they examined are nonreproducible because authors often failed to report values of input parameters or to adequately explain how the demographic parameters were estimated. We found that PVAs developed using generic programs fared poorly in analysis quality. This result is troubling because the use of generic programs to conduct PVAs has increased over time (Supporting Information). The aforementioned issues are primarily a consequence of inappropriate use of these tools or a result of users' lack of knowledge of population dynamics or study species' life histories. However, generic programs can be used to produce high-quality PVAs when used appropriately and nonreproducibility was not a problem specific to PVAs developed using generic programs.

Change in PVA Quality Over Time

Since the 1990s, computational power has increased exponentially, and several PVA guidelines have been published (e.g., Beissinger & McCollough 2002; Pe'er et al. 2013; Zeigler et al. 2013). Also, there have been concerted efforts to improve quantitative skills of life-science students over time (Thompson et al. 2013). Thus, we expected PVA quality to increase over time. Interestingly, model quality did not change over time, but background, analysis, and overall quality declined. Furthermore, the consideration of parametric uncertainty in input parameters also declined over time (Supporting Information).The decline in overall PVA quality over time is troubling and perhaps reflects the fact that some researchers are unaware of, or fail to adhere to, widely accepted PVA guidelines. Even though overall and analysis quality declined for PVAs developed using both custom built and generic models, it remained lower for PVAs developed using generic platform compared with custom-built models. Thus, it appears that the number of PVA studies, and presumably their use in conservation decision making, has increased over time at the cost of quality (Supporting Information). Because most PVA guidelines were published before 2003, we expected studies published after 2005 to be of higher quality. Contrary to our expectation, the overall PVA quality was lower for studies published after 2005 (Supporting Information).

Effect of Journal IF on PVA Quality

We expected the PVA quality to be higher for studies published in high IF journals because high IF is often thought to reflect better study quality (Saha et al. 2003), which can be a result of more rigorous review process. Consistently, background, analysis, and overall qualities were positively associated with the IF of the journal. This suggests that high IF journals scrutinized PVA manuscripts more rigorously with respect to the quality and communication of background information and results. The background quality of PVAs developed using generic programs increased with journal IF, whereas journal IF had no influence on the background quality of PVAs developed using custom-built models. However, journal IF had no effect on the model quality, suggesting that there was no difference between high-profile journals and journals of lower IF in scrutinizing PVA manuscript with respect to many important aspects of PVA (e.g., robust estimation of input parameters and incorporation of stochasticities). Highest IF journal included in this study was Science (IF = 37.02) (Frick et al. 2010). The IF for this journal was substantially higher than the other journals included in our study. However, reanalysis of data after excluding this study did not substantially alter the relationship between IF and PVA quality or conclusions regarding the relationship between IF and PVA quality.

Quality of Threatened Species PVA

We expected the quality of PVAs of threatened species to be higher because more resources are devoted to studying imperiled species (Taylor et al. 2005; Rodrigues et al. 2006) and because of the expectation that researchers would be more thorough due to the potentially high cost of erroneous results. Contrary to our expectation, conservation status of the study species had no discernible effect on PVA quality. Many threatened species occur in low numbers and are difficult to study, which sometimes lead to insufficient data for rigorous estimation of demographic parameters and can negatively influence PVA quality.

Characteristics of Top-Performing PVAs

The highest ranking PVAs received identical overall quality scores (Supporting Information). These studies shared many characteristics because they all thoroughly described study species' life histories and clearly stated study objectives; estimated model parameters with statistically robust analysis of best available data sets and adequately described methods of data collection and analytical approaches used to estimate model parameters; developed the model tailored to the study species' life histories; rigorously estimated demographic parameters using the best available data; adequately addressed stochasticities, density dependence, and external factors that may influence population dynamics and persistence or adequately addressed their exclusion; conducted perturbation analysis; reported estimates of population growth and extinction parameters, which were defined unambiguously; and discussed model limitations. All 4 studies used custom-built modeling platform, and all except Hostetler et al. (2013) were published in the same high IF journal (*Ecological Monographs*).

Among studies based on generic programs, Forys and Humphrey (1999) and Slotta-Bachmayr et al. (2004) were the 2 top performers and both were conducted using Vortex. These studies clearly stated study objectives; adequately described life history of the study species, methods, and period of data collection; incorporated stochasticites and threats; and modeled alternative management scenarios. Furthermore, they conducted perturbation analysis, clearly defined extinction parameters, and reported means and variances of extinction parameters and discussed model limitations. These and other features of top-performing studies, however, were lacking for many studies that were based on generic programs. Demographic studies that did not claim to be PVAs but were included in our study (e.g., Hunter et al. 2010) tended to be of high quality, especially with respect to background and model quality (Supporting Information).

Although the use of PVA in conservation planning is often emphasized, there are many other potential uses of PVA (Beissinger & McCollough 2002; Morris & Doak 2002). For example, PVAs that do not directly deal with conservation planning may still contribute to the body of knowledge via collation of information and analyses of monitoring data, identification of key life-history stages as management targets, research prioritization, or provision of answers to what-if questions (i.e., heuristic studies). Given the many uses of PVA, it is not always clear how to evaluate their quality. We assumed the primary use of PVA is to inform the conservation decision making and evaluated all PVAs similarly based on criteria identified as important in the PVA literature. Although we believe that our approach to PVA evaluation is reasonable and unbiased, we acknowledge that it may sometimes unfairly down rank PVA studies of data-deficient species or those that occur in numbers too low to permit sufficient data for rigorous estimation of demographic parameters.

Although predictive accuracy is considered an important component of PVA (Brook et al. 2000; McCarthy et al. 2001), it is not always clear how best to test predictive accuracy or validate PVA models (Beissinger & Westphal 1998). Furthermore, many PVA studies may not have sufficient data to perform model validation or test predictive accuracy. Well-conducted PVA studies guided by specific questions have intrinsic values, even if they lack long-term data or do not test for predictive accuracy. Also, because our evaluation depended solely on the information provided by the authors, we could not distinguish, for example, between a poor description of a high-quality study or simply a poor-quality PVA study. Thus, a study that used detailed data but failed to adequately describe data collection or parameter estimation methods would be ranked similarly to one that lacked adequate data. Finally, we tested for covariate effects on measures of PVA quality based on the statistical significance of the slope parameter defining the relationship between the 2. Statistical significance does not necessarily imply biological significance (Yoccoz 1991), and our results should be interpreted with caution.

The use of PVA results in conservation decision making has been recommended partly because they are thought to be objective and repeatable (Doak et al. 2015). Given the importance of PVAs in imperiled species management, many authors have insisted on rigor and consistency in PVA analysis and reporting of results (Beissinger & McCollough 2002). Yet, we found that many published PVAs failed to adequately describe data sources and failed to report estimates of model parameters, making many such studies nonreproducible (Morrison et al. 2016). Only 18.1% of the studies we evaluated were scored as high quality (score > 0.75); 14.3% of the studies were of poor quality (score < 0.50). Management recommendations based on poor-quality PVAs can mislead wildlife managers and thus adversely affect the persistence of imperiled populations.

Although widely accepted PVA guidelines have been available for more than a decade, our results show that many authors are either unaware of these guidelines or simply ignore them. Even more troubling is the fact that poor-quality PVAs continue to be published in peer-reviewed journals. The ultimate responsibility to ensure accuracy and reproducibility of PVAs lies with authors, but the responsibility to ensure that poor-quality PVAs are not published rests with the journal editors and reviewers. The PVAs published in peer-reviewed journals are considered reliable by the general public and wildlife managers. Thus, we call for an increased scrutiny of PVAs, especially of imperiled species, by journal editors and reviewers because of the potentially high cost of faulty conservation decisions based on unreliable PVA results. Because poor-quality PVAs continue to be published, we recommend caution while using PVA results in conservation decision making without assessing their quality. Our evaluation framework can be used to evaluate PVA quality by journal editors and reviewers and conservation decision makers.

Acknowledgments

We thank the Department of Wildlife Ecology and Conservation, the Center for Tropical Conservation and Development (University of Florida), WildLandscapes International, and University of Florida Biodiversity Institute for supporting V.C.'s graduate studies. We express our gratitude to E. Morton, M. Burgman, F. Jarrad, C. Rondinini, J. Baum, and 3 anonymous reviewers for many helpful comments that improved the quality of our manuscript.

Supporting Information

The number of PVA studies published every year from 1990 to 2017 (Appendix S1), criteria used for evaluating PVA studies (Appendix S2), model averaged estimates of the regression parameters defining the relationship between covariates and measure of PVA quality (Appendix S3), ranking of PVAs evaluated in this study (Appendix S4), overall quality of PVA studies before and after availability of PVA guidelines (Appendix S5), number of PVA studies developed using custom-built and generic programs 1990–2017 (Appendix S6), results of models testing influence of individual covariates on measures of PVA quality (Appendix S7), number of PVA studies developed using various generic programs (Appendix S8), and proportion of studies that incorporated uncertainty in input parameter from 1990 to 2017 (Appendix S9) are available online. The authors are solely responsible for the content and functionality of these materials. Queries (other than the absence of the material) should be directed to the corresponding author.

Literature Cited

Agresti A. 2015. Foundations of linear and generalized linear models. John Wiley & Sons, Hoboken, New Jersey.

Akçakaya HR, Root W. 2002. RAMAS GIS: linking landscape data with population viability analysis. Version 4.0. Applied Biomathematics. Setauket, New York.

Akçakaya HR, Sjögren-Gulve P. 2000. Population viability analyses in conservation planning: an overview. Ecological Bulletins **1:**9–21.

Bakker VJ, Doak DF, Roemer GW, Garcelon DK, Coonan TJ, Morrison SA, Lynch C, Ralls K, Shaw R. 2009. Incorporating ecological drivers and uncertainty into a demographic population viability analysis for the island fox. Ecological Monographs **79:**77–108.

Barton K. 2013. MuMIn: multi-model inference, R package version 1.9. 13. R Foundation for Statistical Computing, Vienna.

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software **67**. https://doi.org/10.18637/jss.v067.i01.

Beier P, Vaughan MR, Conroy MJ, Quigley H. 2003. An analysis of scientific literature related to the Florida panther. Final report. Florida Fish and Wildlife Conservation Commission, Tallahassee. Available from http://mountainlion.org/us/fl/FL-A-FFWCC-Bier-2003-Analysis-of-Scientific-Literatrure-Related-to-Florida-Panther.pdf (accessed January 1 2017).

Beissinger SR, McCollough DR. 2002. Population viability analysis. University of Chicago Press, Chicago, Illinois.

Beissinger SR, Westphal MI. 1998. On the use of demographic models of population viability in endangered species management. Journal of Wildlife Management **62:**821–841.

Boyce MS. 1992. Population viability analysis. Annual Review of Ecology and Systematics **23:**481–506.

Brook BW, Burgman MA, Frankham R. 2000. Differences and congruencies between PVA packages: the importance of sex ratio for predictions of extinction risk. Conservation Ecology **4:** http://www.ecologyandsociety.org/vol4/iss1/art6/.

Burgman M. 2005. Risks and decisions for conservation and environmental management. Cambridge University Press, Cambridge, United Kingdom.

Burgman M, Possingham H. 2000. Population viability analysis for conservation: the good, the bad and the undescribed. Pages in 97–112 in Young AG, Clarke GM, editors. Genetics, demography and viability of fragmented populations. Cambridge University Press, Cambridge, United Kingdom.

Burnham KP, Anderson DR. 1998. Model selection and inference: a practical information-theoretic approach. Springer-Verlag, New York.

Caswell H, Fujiwara M, Brault S. 1999. Declining survival probability threatens the North Atlantic right whale. Proceedings of the National Academy of Sciences of the United States of America **96:**3308–3313.

Coulson T, Mace GM, Hudson E, Possingham HP. 2001. The use and abuse of population viability analysis. Trends in Ecology & Evolution **16:**219–221.

Doak DF, Himes Boor GK, Bakker VJ, Morris WF, Louthan A, Morrison SA, Stanley A, Crowder LB. 2015. Recommendations for improving recovery criteria under the US Endangered Species Act. BioScience **65:**189–199.

Ellner SP, Fieberg J, Ludwig D, Wilcox C. 2002. Precision of population viability analysis. Conservation Biology **16:**258–261.

Faust LJ, Simonis JS, Harrison R, Waddell W, Long S. 2016. Red wolf (*Canis rufus*) population viability analysis report to U.S. Fish and Wildlife Service. Lincoln Park Zoo, Chicago, Illinois.

Forys EA, Humphrey SR. 1999. Use of population viability analysis to evaluate management options for the endangered Lower Keys marsh rabbit. Journal of Wildlife Management **63:**251–260.

Frick WF, Pollock JF, Hicks AC, Langwig KE, Reynolds DS, Turner GG, Butchkoski CM, Kunz TH. 2010. An emerging disease causes regional population collapse of a common North American bat species. Science **329:**679–682.

Gaona P, Ferreras P, Delibes M. 1998. Dynamics and viability of a metapopulation of the endangered Iberian lynx (*Lynx pardinus*). Ecological Monographs **68:**349–370.

Grant WE. 1986. Systems analysis and simulation in wildlife and fisheries sciences. Wiley, New York.

Henle K, Sarre S, Wiegand K. 2004. The role of density regulation in extinction processes and population viability analysis. Biodiversity and Conservation **13:**9–52.

Himes Boor GK. 2014. A framework for developing objective and measurable recovery criteria for threatened and endangered species. Conservation Biology **28:**33–43.

Hosmer DW, Lemeshow S. 1980. Goodness of fit tests for the multiple logistic regression model. Communications in Statistics-Theory and Methods **9:**1043–1069.

Hostetler JA, Onorato DP, Jansen D, Oli MK. 2013. A cat's tale: the impact of genetic restoration on Florida panther population dynamics and persistence. Journal of Animal Ecology **82:**608–620.

Hunter CM, Caswell H, Runge MC, Regehr EV, Amstrup SC, Stirling I. 2010. Climate change threatens polar bear populations: a stochastic demographic analysis. Ecology **91:**2883–2897.

IUCN (International Union for the Conservation of Nature), UN Environment Programme (UNEP). 2017. Red list of threatened species. Version 2017-1. World database on protected areas. IUCN, Gland, Switzerland, UNEP, Paris. Available from http://www.iucnredlist.org (accessed July 2017).

Johnson PC. 2014. Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. Methods in Ecology and Evolution **5:**944–946.

Lacy RC. 1993. Vortex—a computer-simulation model for population viability analysis. Wildlife Research **20:**45–65.

Lande R. 1993. Risks of population extinction from demographic and environmental stochasticity and random catastrophes. The American Naturalist **142:**911–927.

Lindenmayer DB, Clark TW, Lacy RC, Thomas VC. 1993. Population viability analysis as a tool in wildlife conservation policy: with reference to Australia. Environmental Management **17:**745–758.

Lunney D, Grasser S, O'Neill LE, Matthews A, Rhodes J. 2007. The impact of fire and dogs on Koalas at Port Stephens, New South Wales, using population viability analysis. Pacific Conservation Biology **13:**189–201.

Matthew J. 2017. Generalhoslem: goodness of fit tests for logistic regression models. R package version 1.3.2. Available from https://CRAN.R-project.org/package=generalhoslem. (accessed on February 20, 2018).

McCarthy MA, Burgman MA, Ferson S. 1995. Sensitivity analysis for models of population viability. Biological Conservation **73:**93–100.

McCarthy MA, Possingham HP, Day JR, Tyre AJ. 2001. Testing the accuracy of population viability analysis. Conservation Biology **15:**1030–1038.

Morris WF, Bloch PL, Hudgens BR, Moyle LC, Stinchcombe JR. 2002. Population viability analysis in endangered species recovery plans: past use and future improvements. Ecological Applications **12:**708–712.

Morris WF, Doak DF. 2002. Quantitative conservation biology: theory and practice of population viability analysis. Sinauer Associates, Sunderland, Massachusetts, USA.

Morrison C, Wardle C, Castley J. 2016. Repeatability and reproducibility of population viability analysis (PVA) and the implications for threatened species management. Frontiers in Ecology and Evolution **4:**98. https://doi.org/10.3389/fevo.2016.00098.

Nakagawa S, Schielzeth H. 2013. A general and simple method for obtaining R2 from generalized linear mixed-effects models. Methods in Ecology and Evolution **4:**133–142.

Neel MC, Leidner AK, Haines A, Goble DD, Scott JM. 2012. By the numbers: How is recovery defined by the US Endangered Species Act? BioScience **62:**646–657.

Pe'er G, et al. 2013. A protocol for better design, application, and communication of population viability analyses. Conservation Biology **27:**644–656.

Possingham HP, Davies I. 1995. ALEX—a model for the viability analysis of spatially structured populations. Biological Conservation **73:**143–150.

R Core Team. 2017. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna.

Ralls K, Beissinger SR, Cochrane JF. 2002. Guidelines for using population viability analysis in endangered-species management. Pages 521–550 in Beissinger SR, McCullough DR, editors. Population viability analysis. University of Chicago Press, Chicago, Illinois.

Reed JM, Mills LS, Dunning JB, Menges ES, McKelvey KS, Frye R, Beissinger SR, Anstett MC, Miller P. 2002. Emerging issues in population viability analysis. Conservation Biology **16:**7–19.

Rodrigues AS, Pilgrim JD, Lamoreux JF, Hoffmann M, Brooks TM. 2006. The value of the IUCN Red List for conservation. Trends in Ecology & Evolution **21:**71–76.

Runge MC, Sanders-Reed CA, Langtimm CA, Hostetler JA, Martin J, Deutsch CJ, Ward-Geiger LI, Mahon GL. 2017. Status and threats analysis for the Florida manatee (*Trichechus manatus latirostris*), 2016. U.S. Geological Survey Scientific Investigation Report 2017-5030. U.S. Geological Survey, Washington, D.C.

Saha S, Saint S, Christakis DA. 2003. Impact factor: A valid measure of journal quality? Journal of the Medical Library Association **91:** 42–46.

Slotta-Bachmayr L, Boegel R, Kaczensky P, Stauffer C, Walzer C. 2004. Use of population viability analysis to identify management priorities and success in reintroducing Przewalski's horses to southwestern Mongolia. Journal of Wildlife Management **68:**790–798.

Taylor MF, Suckling KF, Rachlinski JJ. 2005. The effectiveness of the Endangered Species Act: a quantitative analysis. AIBS Bulletin **55:**360–367.

Taylor MK, Laake J, McLoughlin PD, Cluff HD, Born EW, Rosing-Asvid A, Messier F. 2008. Population parameters and harvest risks for polar bears (*Ursus maritimus*) of Kane Basin, Canada and Greenland. Polar Biology **31:**491–499.

Tear TH, et al. 2005. How much is enough? The recurrent problem of setting measurable objectives in conservation. BioScience **55:**835–849.

Thompson KV, Cooke TJ, Fagan WF, Gulick D, Levy D, Nelson KC, Redish EF, Smith RF, Presson J. 2013. Infusing quantitative approaches throughout the biological sciences curriculum. International Journal of Mathematical Education in Science and Technology **44:**817–833.

White GC. 2000. Population viability analysis: data requirements and essential analyses. Pages 288–331 in Boitani L, Fuller TK, editors. Research techniques in animal ecology: controversies and consequences. Columbia University Press, New York.

Wiegand T, Naves J, Stephan T, Fernandez A. 1998. Assessing the risk of extinction for the brown bear (*Ursus arctos*) in the Cordillera Cantabrica, Spain. Ecological Monographs **68:**539–570.

Yoccoz NG. 1991. Use, overuse, and misuse of significance tests in evolutionary biology and ecology. Bulletin of the Ecological Society of America **72:**106–111.

Zeigler SL, Che-Castaldo JP, Neel MC. 2013. Actual and potential use of population viability analyses in recovery of plant species listed under the U.S. Endangered Species Act. Conservation Biology **27:**1265–1278.

Zuur, AF, Leno EN, Walker NJ, Saveliev AA, Smith GM. 2009. Mixed Effects Models and Extensions in Ecology with R, Springer Science and Business Media, New York.