ORIGINAL PAPER

Check for updates

Assessing invasion risks using EICAT-based expert elicitation: application to a conservation translocation

Maude Vernet · Amanda E. Trask · Caitlin E. Andrews · John G. Ewen · Suzanne Medina · Axel Moehrenschlager · Stefano Canessa

Abstract Conservation translocations are widely used to recover threatened species, but can pose risks to recipient ecosystems, particularly in the case of conservation introductions. Because of limited data and uncertainty, risk assessments for such projects often rely on extrapolated evidence and expert opinion, further complicating decision making. The Environmental Impact Classification for Alien Taxa (EICAT) serves to classify the realised impacts of invasive species. We developed a protocol combining EICAT and formal expert elicitation to predict these impacts. We applied our protocol to the extinct-in-the-wild sihek (Guam kingfisher; *Todiramphus cinnamominus*), for which introduction outside the known historical range is being considered. We elicited from multiple experts probability estimates of impact levels across four impact mechanisms and five candidate release sites. We aggregated estimates using simulation-based and Bayesian approaches, with and without accounting for expert confidence. Experts generally agreed that sihek introduction might impact the recipient ecosystem through predation, competition, and disease, although they disagreed about the likely impact levels. Releases to Palmyra Atoll were considered to pose the lowest risk across candidate sites, so this site was selected for further detailed ecological assessments and risk mitigation efforts. EICAT, with its standardized impact mechanisms and definitions, helped reduce

Maude Vernet and Amanda E. Trask have contributed equally to this manuscript.

M. Vernet (✉) · S. Canessa
Division of Conservation Biology, Institute of Ecology and Evolution, University of Bern, Bern, Switzerland
e-mail: maude.vernet@students.unibe.ch

A. E. Trask · C. E. Andrews · J. G. Ewen
Institute of Zoology, Zoological Society of London, London, UK

C. E. Andrews
The Nature Conservancy, Sacramento, CA, USA

S. Medina
Guam Department of Agriculture, Division of Aquatic and Wildlife Resources, Mangilao, Guam

A. Moehrenschlager
Conservation Translocation Specialist Group, IUCN Species Survival Commission, Calgary, AB, Canada

A. Moehrenschlager
Panthera, New York City, NY, USA

S. Canessa
Department of Environmental Science and Policy, Università Degli Studi Di Milano, Milan, Italy

the linguistic uncertainty and subjectivity common to expert-based assessments. Expressing judgments as probabilities allowed us to evaluate uncertainty transparently and to assess the weight of expert confidence on the overall risk estimation. Formal quantitative elicitation and aggregation then allowed a transparent evaluation of results, facilitating communication with stakeholders and decision-makers.

Introduction

Conservation translocations are increasingly used to restore or reinforce decreasing or locally extirpated populations (Seddon et al. 2007). However, such active interventions can entail risks, including potentially unwanted impacts on the recipient ecosystem (IUCN/SSC 2013). This concern is especially strong for conservation introductions (also known as assisted colonisation), the movement of species beyond their indigenous range, which are often proposed as a solution when ecosystems have been altered to the point where they are no longer habitable for their native species (Hoegh-Guldberg et al. 2008; Seddon 2010). In the face of such risks, translocation decisions require a balance of interests where costs and benefits must be carefully weighted (Canessa et al. 2016a, b, c). The stakes can be exceptionally high, with possible extinction on the one hand and damaging biological invasion risks on the other.

Careful risk assessments are a fundamental requirement of any such discussion (Mueller and Hellmann 2008), but they require knowledge about historical ranges and proposed release sites, including interactions among multiple species, the environment, and humans (Roy et al. 2020). The highly endangered species typically considered for conservation introduction are usually rare, and their ecology may be poorly known, even in their original range. Processes of decline, persistence at low numbers, and introduction outside historical range can, in themselves, induce changes in behaviour and ecology (Wilson et al. 2020); additional research to clarify these uncertainties is often limited by sample sizes and perceived risks (Martin et al. 2012; Tulloch et al. 2015;

Canessa et al. 2016a, b, c). Where empirical evidence is limited, risk assessments often require some level of expert judgment.

Where translocation decisions involve concerns of invasion risks, assessment should use insights and available methods from invasion ecology. For example, the Environmental Impact Classification for Alien Taxa (EICAT; Blackburn et al. 2014) has been adopted by the IUCN as a global standard to classify the impacts of species established beyond their range (IUCN 2020). This is done by identifying the mechanisms through which an invasive species impacts an ecosystem and by classifying the severity of impact based on five levels ranging from "Minimal" (negligible impact on native species) to "Massive" (irreversible impact resulting in the replacement or extinction of at least one native species) (Hawkins et al. 2015; Kumschick et al. 2020; Volery et al. 2020). The EICAT typically focuses on realised impacts by invasive species, it provides a standard classification system that might be extended to assessment of prospective impacts, including for conservation translocations of threatened species.

Shifting the focus from recorded to prospective impacts is likely to require some level of extrapolation. Although ideally evidence-based (Kesner and Kumschick 2018; Henry and Sorte 2022), EICAT assessments often complement empirical knowledge with expert opinion (Dehnen-Schmutz et al. 2022). A wide body of literature exists on how to carry out elicitation to minimise individual and group biases (Burgman et al. 2011; Martin et al. 2012; Sutherland and Burgman 2015). These methods have been used effectively in combination with EICAT before (Turbé et al. 2017; Vanderhoeven et al. 2017). However, assessments that use categorical ratings, like the impact levels in EICAT, remain challenging. Linguistic uncertainty, where people interpret different words in different ways (Wintle et al. 2019), is reduced by providing a standard set of definitions for all levels. However, even after definitions are clarified, experts may disagree in their judgments and differ in their levels of certainty. In these cases, seeking consensus by discussion may facilitate groupthink bias and reward overconfidence (Kuhnert et al. 2010).

Alternatively, variation across experts can be accounted for by aggregating judgments, although this is not straightforward for verbal expressions. For example, it is difficult to define an "average" rating

across a group of experts where judgments range from Minimal to Massive. It might be tempting to convert categories to numerical scores to aggregate them mathematically, for example by multiplying or averaging them (Evans 2018; Sohrabi et al. 2021), but such aggregation is often meaningless or misleading (Game et al. 2013; Canessa et al. 2021). Moreover, uncertainty is not limited to disagreements among experts but includes also different levels of confidence by individuals. This uncertainty, reflecting lack of knowledge or simply the challenge of predicting future events, must be accounted for in any analysis, both when considering individual estimates and when aggregating across groups. Previous studies have approached uncertainty in EICAT classification by including and comparing multiple experts and their confidence, but they still rely on verbal definitions and arbitrary categorisations (Probert et al. 2020; Clarke et al. 2021). There are several methods for more formal quantitative expert aggregation, for example using multinomial-Dirichlet conjugate priors to express probabilities of categorical events (Wilson et al. 2021), that would be useful in the assessment of invasion risks including from conservation translocations.

In this study, we combined EICAT impact definitions with formal expert elicitation and quantitative aggregation of expert judgments to carry out an assessment of translocation risks, maximising transparency and incorporating uncertainty within and across multiple experts. We used this process in the practical case study of an extinct-in-the-wild bird species being considered for a conservation introduction beyond its known historical range.

Methods

Risk context

The sihek (Guam kingfisher, *Todiramphus cinnamominus*) is a medium-sized endangered bird species endemic to the island of Guam in the North-Western Pacific (Jenkins 1983). Historically, they inhabited limestone and ravine forests, coconut groves, and strand vegetation and primarily fed on lizards and large insects, but were also observed to occasionally predate crabs, segmented worms, fish, and small birds (Beck et al. 1990 and references therein). Sihek

are territorial and lay one to three eggs by excavating cavities in softwood during the nesting season (December to July) (Beck et al. 1990 and references therein).

Sihek were extirpated from Guam following the accidental introduction of the invasive brown tree snake (*Boiga irregularis*), likely through military activities (Savidge 1986; Engbring and Fritts 1988). Prior to extirpation, a conservation breeding population was established, which persists to date but is descended from only 16 individuals. The species is listed as Extinct in the Wild by the IUCN (BirdLife International 2017). Returning sihek to the wild has been recommended in part due to increasing genetic concerns arising from continued small population size and subsequent inbreeding (Trask et al. 2021), as well as potential adaptation to captive conditions. However, the continued presence of brown tree snakes in Guam remains a major challenge to reintroducing sihek in the immediate future. The establishment of a wild population outside the indigenous range could allow sihek to breed in a more natural environment. This, in turn, would reduce potential for adaptation to captive conditions and increase the global population size, thereby slowing the rate of increase in inbreeding and loss of genetic diversity (Trask et al. 2021), and providing valuable information about their behaviour and vital rates in the wild to inform future reintroduction efforts on Guam. Prior to any such action, the risks posed by sihek to recipient ecosystems should be thoroughly assessed; however, available information about sihek natural history is mainly limited to anecdotal descriptions and expert knowledge.

In 2019, we began evaluating potential release sites for sihek based on ecological suitability and logistical feasibility. From a prior longlist of sites potentially suitable for sihek throughout the Pacific Islands based on ecological criteria (Laws and Kesler 2012), a sihek working group identified five candidate release sites where sihek release would in principle be logistically feasible: Kosrae, Chuuk, and Yap in the Federated States of Micronesia; Palmyra Atoll in the Northern Line Islands; and Tinian as a representative island of the Commonwealth of the Northern Mariana Islands (Fig. 1) (Trask et al. 2019). These sites were selected for initial assessment to inform further engagement with local authorities and communities. To assess the risks posed by sihek introduction to these sites, we
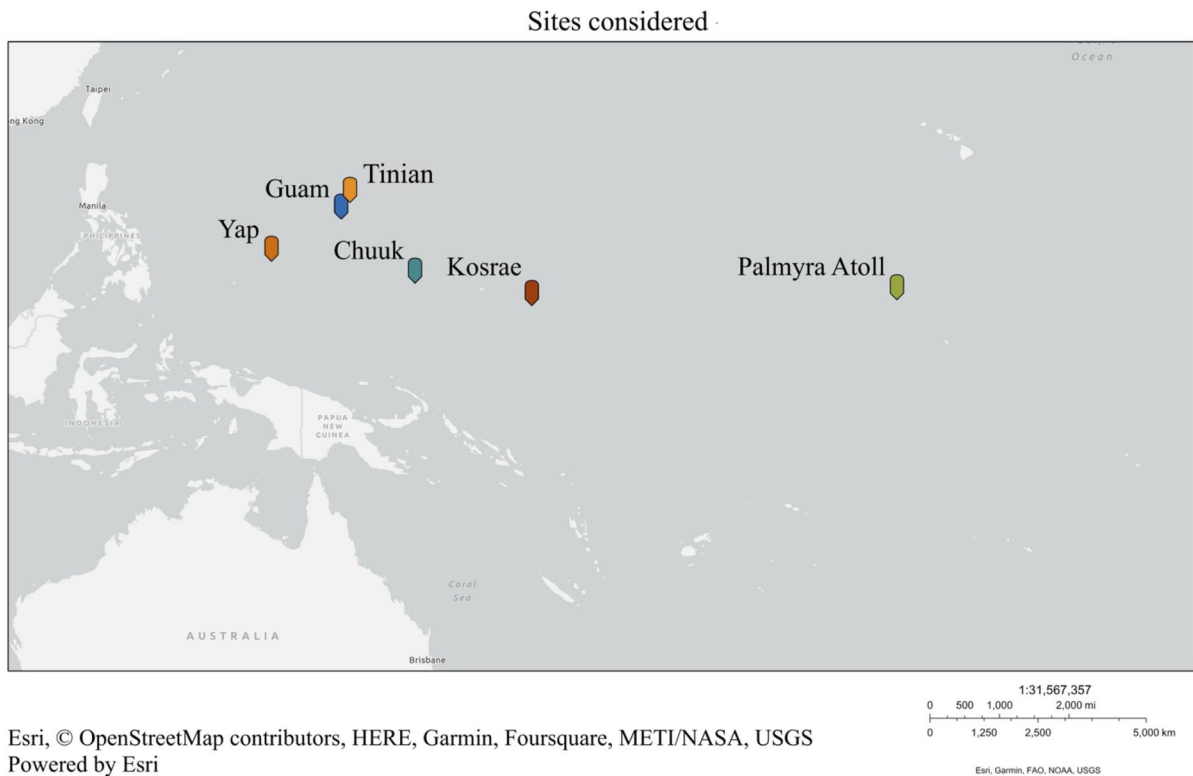
Fig. 1 Location of Guam and the five candidate introduction sites selected for this impact assessment study: Kosrae, Chuuk, and Yap in the Federated States of Micronesia, Palmyra Atoll in the Northern Line Islands, and Tinian in the Commonwealth of the Northern Mariana Islands (CNMI)

combined the EICAT framework with formal expert elicitation as described below.

EICAT-based risk assessment

To assess the risk of environmental impacts posed by sihek introduction to the candidate sites, we based our approach on the Environmental Impact Classification of Alien Taxa (EICAT; Blackburn et al. 2014), a well-developed framework widely used to classify the impacts of alien species based on twelve ecological impact mechanisms recognised by the IUCN SSC Invasive Species Specialist Group (IUCN 2020). The EICAT requires assessors to identify the relevant impact mechanisms and assign to them one of five impact levels: Minimal, Minor, Moderate, Major, and Massive (IUCN 2020; Table 1). Based on ecological knowledge of sihek and similar species, we were able to confidently exclude several impact mechanisms (e.g., chemical, or physical alteration of habitat). Out

of the twelve impact mechanisms described in the framework of the EICAT, we therefore concentrated this assessment only on competition, predation, disease, and hybridisation. Furthermore, because sihek are endemic to Guam and their extirpation is recent, their return would not represent a conservation introduction; therefore, we did not include Guam in the assessment for competition, predation, and hybridization impacts. We did, however, include Guam in the disease assessment, because birds brought from institutions in North America might accidentally involve release of pathogens novel to Guam.

Expert elicitation

To formally elicit expert judgment within the EICAT framework, we based our assessment on the IDEA protocol ("Investigate," "Discuss," "Estimate", and "Aggregate") for elicitation (Hemming et al. 2018). We selected a total of 40 experts based

Table 1 Top section: sihek-specific Description and Outcome for each impact mechanism as provided to experts. Bottom section: impact levels as defined in the EICAT (IUCN 2020), used by experts to quantify outcomes for each mechanism

	Competition	Predation	Hybridisation	Disease
Description of impact mechanisms				
	The alien taxon competes with native taxa for resources (e.g., food, water, space), leading to deleterious impact on native taxa	The alien taxon predates on native taxa, leading to deleterious impact on native taxa	The alien taxon hybridises with native taxa, leading to deleterious impact on native taxa	The alien taxon transmits diseases to native taxa, leading to deleterious impact on native taxa
Description of impact levels				
Minimal	Negligible impacts, and no reduction in performance of a native taxon's individuals			
Minor	Performance of individuals reduced, but no decrease in population size			
Moderate	Native taxon population decline			
Major	Native taxon local extinction (i.e. change in community structure), which is naturally reversible			
Massive	Naturally irreversible local or global extinction of a native taxon (i.e. change in community structure)			

on their knowledge about at least one of the following areas: the focal species, invasive species, one or more impact mechanisms, and source and release sites, and invited them by individual email to complete an assessment for the specific impact mechanism(s) relevant to their expertise. In total, 21 experts participated and completed 25 assessments (5–8 experts per mechanism; Table 2). Some experts contributed to evaluating more than one mechanism, and two collaborated on a joint assessment of the mechanisms predation and competition. The final expert cohort of respondents was composed of fifteen men and five women from Italy, New Zealand, the United Kingdom, and the USA, aged 20 to 60+, with a majority between 40 and 60 years old. Assessors were in majority research biologists except for disease impacts, where most contributors were wildlife veterinarians.

Experts contributing to assessments of competition and predation impacts had expertise in Pacific birds, as well as potential competitor or prey taxa, respectively. Contributors assessing hybridisation impacts had expertise in conservation genetics, hybridisation, and speciation. Invasive species experts also contributed to the assessment of these mechanisms. For the assessment of disease impacts, contributors had expertise in wildlife health, with most experts also familiar with the IUCN wildlife disease risk analysis process. Experts were also provided with scientific literature relevant to their assessment. This comprised of a list of species that could potentially interact with sihek at each site through the mechanism assessed, as well as background information on sihek or related species and the potential release sites (summarised in Table 2), and description of the impact mechanisms and levels (summarised in Table 1). The material provided to experts in its entirety (full list of species and complete background information) is available in the Supplementary Material (Appendix 1).

To avoid the linguistic uncertainty associated with verbal definitions of risk, avoid arbitrary categorisation, and facilitate comparison and aggregation, we asked each expert to estimate the probability that a given impact would reach each of the five levels at a given release site (summing to 1 for each site). Experts were asked to provide a written justification for their impact estimates (Supplementary Material, Appendix 3). We also asked experts

Table 2 Summary of the information provided to experts for their assessment. Top: Summary information about sihek or sihek-related species (drawn from the literature on sihek, related species, and the release sites). Bottom: Summary list of species potentially relevant to the assessment of each impact mechanism. Numbers in parentheses indicate the number of assessments and of contributing experts

Summary of background information provided to experts

		Competition ($n=6/7$*)	Predation ($n=8/9$*)	Hybridisation ($n=6/7$*)	Disease ($n=5$)	
Sihek/related species	All sites	Habitat requirements Diet Predation habits Breeding characteristics	Habitat requirements Diet Predation habits	Phylogenetic relatedness of *Todiramphus spp.* morphology	Diseases identified in the captive sihek population Proposed translocation pathway (i.e. flight path, stopovers) and quarantine procedures	
Summary of potentially threatened endemic species at release sites	Kosrae	Boettger's skink (endangered) Kosrae White-eye (endemic)		Sacred kingfisher (potential migrant)	Avian malaria vectors present	
	Chuuk	Teardrop white-eye (endemic & endangered)		Few susceptible species	Chuuk monarch (endemic, endangered)	
		Boettger's skink (endangered)	Micronesian speckle-bellied gecko (endangered)			
	Yap	Many endemic passerines			Many endemic passerines	
		Giant Micronesian geckos (declining)	Micronesian speckle-bellied gecko (endangered)			
	Palmyra	Nocturnal geckos Coconut and terrestrial crabs (nocturnal)		Few susceptible species		
	Tinian	Tinian monarch (endemic)				
		Collared kingfisher (endemic and endangered)				
	Guam	Not Applicable				

Total number of assessments and experts $=25(21$**$)$

*Two experts collaborated on a single assessment

**Seven experts contributed to the assessment of more than one impact mechanism

to estimate confidence in their knowledge, ranging from 0% (no confidence) to 100% (total confidence). We repeated the elicitation for all impact mechanisms across all sites. All elicited estimates, in anonymous format, can be found in Supplementary Material.

Aggregation of estimates and uncertainty

After obtaining quantitative estimates from each expert about the probability of different impact levels for each mechanism, we sought to summarise that information quantitatively, incorporating both the dispersion across the group and the confidence of individual experts. We developed, used and compared

three methods for aggregation that provided different but complementary ways to account for this uncertainty.

First, we aggregated the expert estimates without considering their expressed confidence (hereafter, we refer to this as "basic method"). For each expert, we defined a Dirichlet distribution, using the expressed probabilities of different impact levels as shape parameters (intuitively corresponding to "votes" for a given impact level: each expert allocated their votes according to their beliefs), and sampled that distribution 100 times. We then aggregated those votes into a single distribution (with a length equal to 100 times the number of experts for the given impact mechanism) and assessed it using

summary statistics (mean, median and standard deviation) and plots.

Second, we repeated the same process but accounted for uncertainty by weighting experts proportionally to their expressed confidence (hereafter, "bootstrapping method"). For each expert, the shape parameters of the Dirichlet distribution were proportional to that expert's expressed confidence:

$$v_e = \frac{c_e EV}{\sum_{e=1}^{N} c_e}$$

where c_e is the confidence expressed by expert e, E is the total number of experts, and V is the number of votes per expert in the basic method. For example, for an assessment with four experts who had expressed uncertainties c of 100, 80, 60 and 40, their proportional share of 400 votes ($V = 100$ each in the basic method, times four experts) would be 0.36, 0.29, 0.21 and 0.14, corresponding to v_e of 143, 114, 86 and 57 votes respectively. Again, we aggregated the draws into a single distribution and assessed it using summary statistics and plots.

Third, we used a Bayesian approach to estimate a posterior distribution for the mean probabilities across the expert group, accounting for uncertainty. For each expert, we created deterministically a vector of votes of length equal to their expressed confidence, subdivided proportionally to the expressed probabilities for each impact level. We then bundled all expert vectors together and used a Bayesian sampler to estimate the underlying probabilities, using an uninformative Dirichlet prior (conjugate to the categorical distribution and a common choice for such processes). We drew N samples from the posterior distribution, where N was calculated to provide the same total number of draws as obtained for the basic and bootstrap aggregation after a burn-in of 2000 ($N = V\,E + burn\text{-}in$), over three Markov chains with a thinning rate of 3. We assessed chain convergence by visual inspection and the R-hat statistic (Brooks and Gelman 1998).

In summary, the basic method represents uncertainty across the group but ignores uncertainty in individual judgments; the bootstrapping method includes both group and individual uncertainty by weighting; and the Bayesian method uses the weighted estimates to derive a hypothetical underlying posterior distribution ("consensus" group mean)

for each impact probability. We consider none of the methods to be inherently superior to the others, but to represent different approaches to treating uncertainty (ignoring it, fully accepting it, and seeking consensus) that should be evaluated jointly.

All analyses described above were conducted using R (R Core Team 2021) and JAGS (Plummer 2003), and R packages *readxl* (Wickham et al. 2023), *tidyverse* (Wickham et al. 2019), *rBeta2009* (Cheng et al. 2014), *jagsUI* (Plummer 2003), *ggplot2* (Wickham 2011), and *gridExtra* (Auguie and Antonov 2017). Reproducible code can be accessed from GitHub (https://github.com/maude-v/invasion_risk_EICAT).

Results

As expected, the mean estimates from bootstrapping and Bayesian methods were equal within two decimal digits. Unless otherwise stated, we report the latter here for consistency. Detailed estimates for all three methods (basic, bootstrapping and Bayesian) can be found in Supplementary Material, Appendix 2, Tables S1–S2.

For competition, there was general agreement across most experts that Major or Massive impacts were unlikely. The mean estimates for these two levels ranged from 0 to 0.2 in all sites, with the exception of Tinian, where experts did not rule out a higher probability of Major or Massive impacts, albeit with some disagreements (Fig. 2). Palmyra Atoll was considered the lowest-risk site for this impact mechanism, with Minimal impacts receiving a mean probability greater than 0.8 regardless of the aggregation method used (Fig. 2; Tables S1–S2). For Chuuk, Kosrae and Yap, estimates were broadly distributed across Minimal to Minor impacts, with greater uncertainty (wider confidence intervals), reflecting divergences of opinion among experts (Fig. 2; Table S2). The presence of the Mariana kingfisher *Todiramphus albicilla*, congeneric to sihek, and several other bird species also led some experts to increase their estimates for competition impacts on Tinian. Uncertainty around this mechanism was also high, because candidate sites host many insufficiently surveyed diurnal skink and gecko species with poorly known systematics and unknown densities.

Fig. 2 Aggregation
of expert estimates for
competition impacts for
all sites. Panel (Basic)
represents the basic method
(no consideration of expert
confidence), panel (Boot-
strapping) the bootstrapping
method (experts weighted
by confidence), Panel
(bayesian) the Bayesian
method (posterior distribu-
tion of group estimates).
In each panel, black dots
indicate means, black lines
indicate the 25th and 75th
percentiles, violin plots
indicate the probability dis-
tributions by impact level,
increasing in severity from
left to right
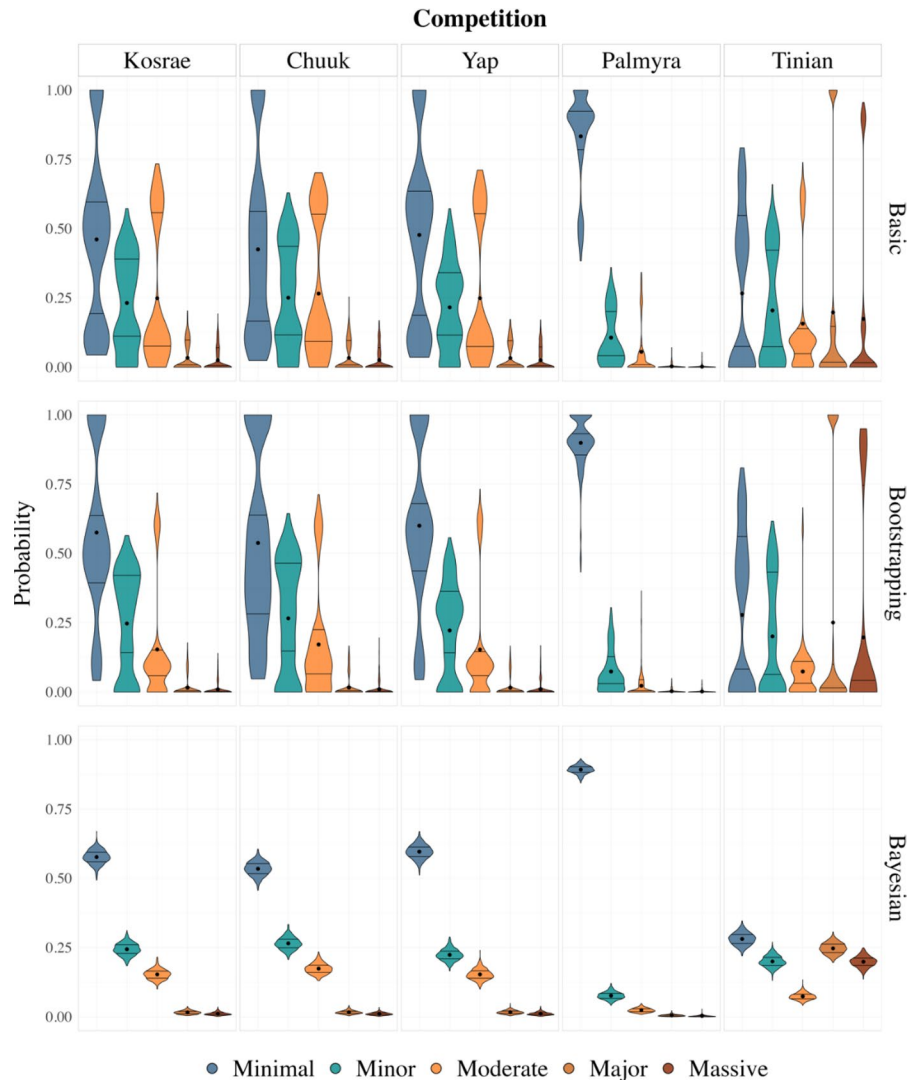
Overall, predation was considered the most severe potential mechanism of impact: for all sites except Palmyra Atoll, the probability of Massive impacts (non-reversible loss of at least one species, Table 1) was higher than 0.15. Sihek have a relatively wide range of possible prey (Beck et al. 1990). Probabilities were estimated more evenly across levels and sites than for other impact mechanisms (mean ~ 0.2, Table S1). Estimates of Minor and Moderate impacts showed a relatively clear binomial split in the group, with the probability of these impacts being expected around 0.15 or 0.5 by different experts (Fig. 3). Across the group, Moderate impacts were considered the most probable for all sites; the probability of Massive impacts was lowest at Palmyra Atoll

(0.03; Table S1), both in terms of the mean estimate across the group and of the estimate by the most pessimistic expert. All other islands were rated similarly, with Yap having the highest probability of Massive impacts (0.22; Table S1). Experts expected released sihek to feed opportunistically on many taxa, possibly reducing the risk of substantial impacts on any one species. However, some experts reasoned that the sihek's hunting habits and known historical prey preferences could indicate a particular risk to small passerines and diurnal lizard species. In particular, Yap and Kosrae host endemic passerine species (the Kosrae white-eye *Zosterops cinereus*, the Yap olive white-eye *Z. oleaginus* and the plain white-eye

Fig. 3 Aggregation of expert estimates for predation impacts for all sites. Panel (Basic) represents the basic method (no consideration of expert confidence), panel (Bootstrapping) the bootstrapping method (experts weighted by confidence), Panel (Bayesian) the Bayesian method (posterior distribution of group estimates). In each panel, black dots indicate means, black lines indicate the 25th and 75th percentiles, violin plots indicate the probability distributions by impact level, increasing in severity from left to right
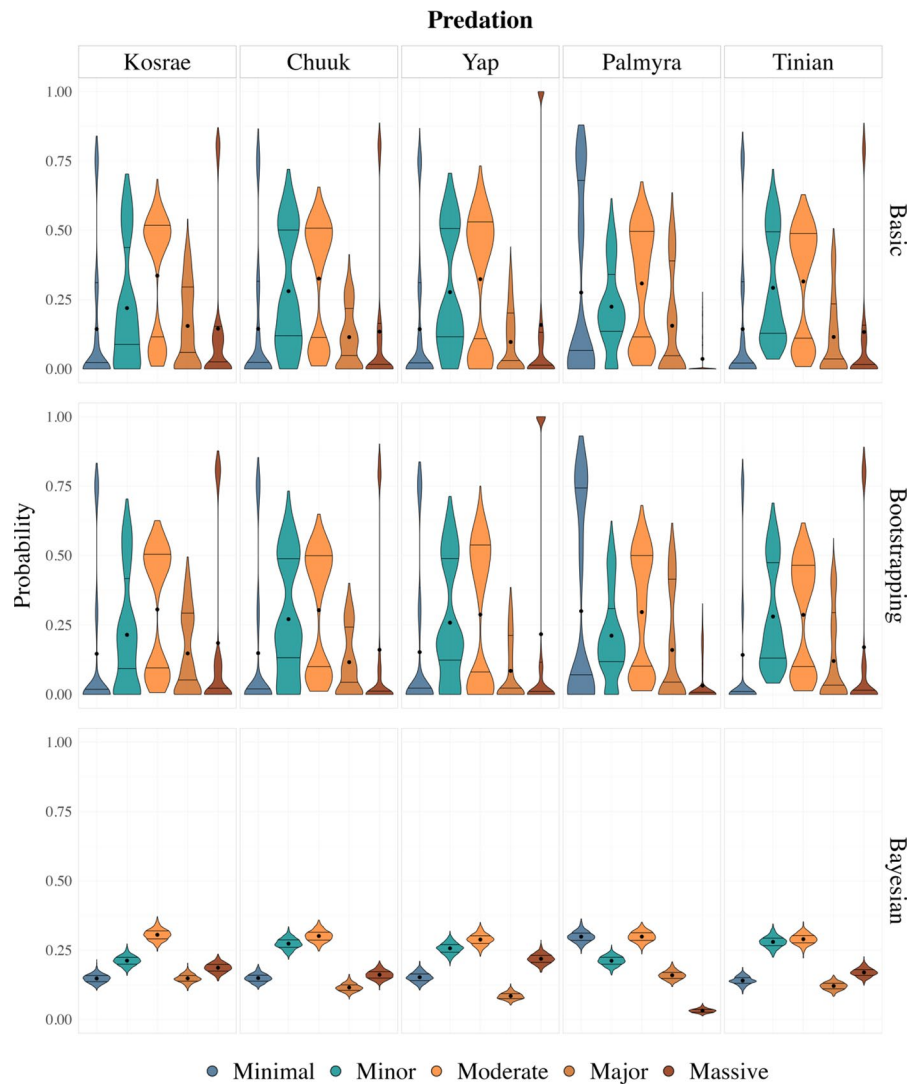

Z. hipolais), which experts highlighted as potential sihek prey species.

For hybridisation, substantial impacts were expected only on Tinian, where Minimal and Minor impacts were considered most likely (0.41 and 0.35, respectively; Fig. 4, Table S1), but Moderate and even Major impacts could not be ruled out (0.17 and 0.05, respectively; Fig. 4). For all other sites, the most likely impacts through hybridisation were generally considered Minimal or Minor (0.93, and 0.06 to 0.07, respectively; Fig. 4, Table S1), with high confidence that impacts of greater magnitude were unlikely (0.00 mean and third quartile; Table S1). Higher estimates for Tinian reflected the experts' concern that sihek introduction to northern Marianas islands might have

deleterious effects on the endemic Mariana kingfisher. Although reproductive isolation between sihek and Mariana kingfisher was considered likely (Andersen et al. 2015), a risk of impact through hybridisation could not be ruled out with certainty.

Disease impacts through sihek introduction were also prominent in the judgment of experts, with probabilities evenly spread between Minimal and Major impacts and the lowest disagreement among experts of all mechanisms (Fig. 5). The probability of Minor to Moderate impacts was estimated around 0.3 at all sites (Table S1). Tinian had the highest probability of Major and Massive impacts (0.19 and 0.09, respectively; Fig. 5, Table S1), whereas Palmyra Atoll had the highest probability of Minimal

Fig. 4 Aggregation
of expert estimates for
hybridisation impacts for
all sites. Panel (Basic)
represents the basic method
(no consideration of expert
confidence), panel (Boot-
strapping) the bootstrapping
method (experts weighted
by confidence), Panel
(Bayesian) the Bayesian
method (posterior distribu-
tion of group estimates).
In each panel, black dots
indicate means, black lines
indicate the 25th and 75th
percentiles, violin plots
indicate the probability dis-
tributions by impact level,
increasing in severity from
left to right
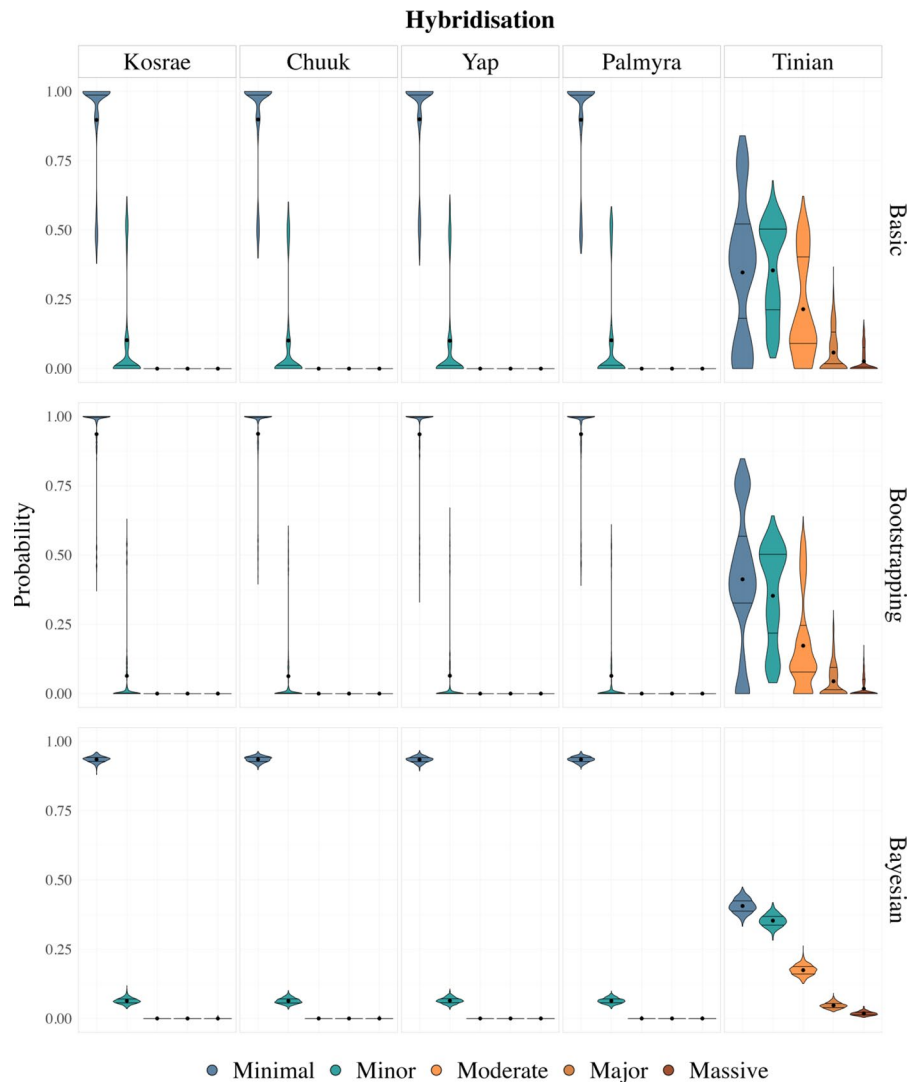
impacts and the lowest probability of Major to Mas-
sive impacts (0.04; Fig. 5, Table S1). While there
were slight disagreements among experts about
the estimated probabilities of Minor to Moderate
impacts, there was general agreement about the
probability of Major impacts at all sites, estimated
around or above 0.1 except for Palmyra Atoll. At
Palmyra Atoll, experts judged the risks of disease
impacts to be generally lower (Fig. 5), particularly
because this site does not host local passerines, but
they could not rule out impacts on other taxa.

When we compared estimates across the three
methods, we found only minor differences between
the probability distributions of expected impacts
under the basic and bootstrapping methods,

suggesting uncertainty by experts (and conversely
their self-assessed confidence) did not significantly
change results. For competition, at most sites, incor-
porating uncertainty slightly changed individual
distributions, but not the ranking of impact levels
or of sites. For hybridisation, including uncertainty
increased the confidence in minimal impacts. The
Bayesian approach returned the same mean estimates
as the bootstrapping method (as expected), but much
narrower uncertainty ranges, suggesting that deriving
a posterior "consensus" mean in this case might be
misleading.

Fig. 5 Aggregation of expert estimates for disease impacts for all sites. Panel (Basic) represents the basic method (no consideration of expert confidence), panel (Bootstrapping) the boot-strapping method (experts weighted by confidence), Panel (Bayesian) the Bayesian method (poste-rior distribution of group estimates). In each panel, black dots indicate means, black lines indicate the 25th and 75th percentiles, violin plots indicate the probabil-ity distributions by impact level, increasing in severity from left to right
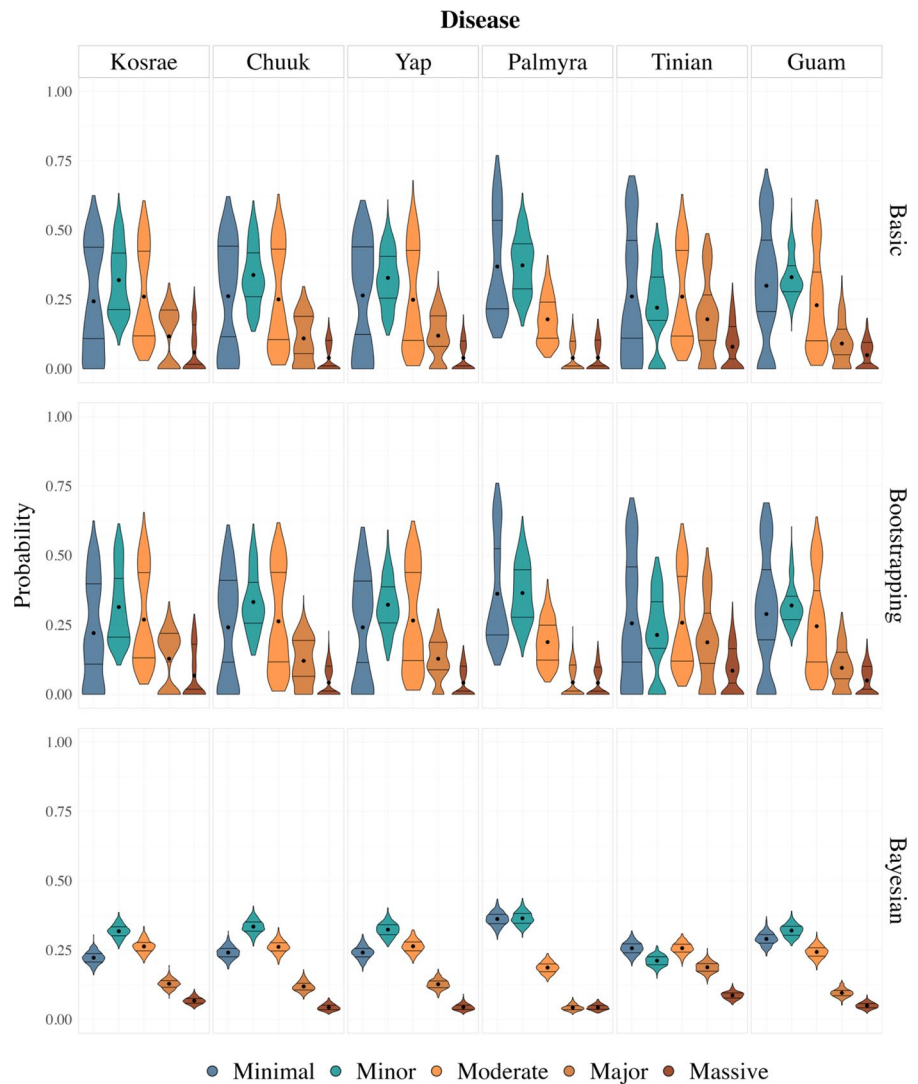
Discussion

Across our expert group, a sihek introduction was considered most likely to affect candidate release sites through predation, disease, and, to a lesser extent, competition. Overall, Tinian was deemed least favourable to sihek introduction, partly due to the presence of many potential bird competi-tors (Table 2). Experts suggested a potential sihek introduction would pose the lowest risk at Palmyra Atoll relative to other sites. However, competition, predation and disease risks for Palmyra Atoll were not considered negligible in absolute terms. There-fore, the Team decided to undertake a quantitative

assessment of competition and predation risks at Palmyra Atoll, by collecting more information about the Atoll's ecological community and pre-dicting impacts via ensemble ecosystem modelling (Baker 2017), and to obtain a detailed disease risk analysis based on surveillance and expert veterinar-ian advice (Sainsbury and Vaughan-Higgins 2012). The results of these further assessments will then inform the ultimate decision to proceed or not with releases, and the extent and type of risk mitigation actions such as pre-and post-release quarantine and monitoring.

Our approach to assessing risks of sihek trans-location also provides lessons for future similar

conservation decisions. Formal elicitation and treatment of expert judgments are especially important when expanding a framework like EICAT–intended to assess realised impacts based on evidence–to prediction of future impacts and their likelihood. In our case, there was general agreement among experts, particularly in the overall ranking of sites, but some key differences could still be seen. For example, on Tinian, experts disagreed about the probability of major and massive impacts. These differences became more marked when including expert confidence using our bootstrapping approach, because confidence was generally high for this site. As another example, for competition impacts on Palmyra Atoll, weighting estimates by confidence increased the mean probability of minimal impacts and decreased the mean probability of moderate impacts. This suggests that experts expecting lower impacts were more confident than those expecting greater impacts.

Research on expert judgment has shown that groups generally outperform individuals, reducing error and improving decisions (Burgman et al. 2011). However, when consulting multiple experts, the challenge is what to do with disagreements and confidence. Creating consensus through discussion may be tempting, but it can be easily biased by power dynamics, groupthink, cascade effects, and simply "discussion fatigue" (Kuhnert et al. 2010). Most elicitation methods indeed do not regard consensus as an objective; rather, they seek to define uncertainty and reduce overall noise and error (Kuhnert et al. 2010). The form in which judgments are elicited then defines how they can be aggregated or compared and how confidence can be accounted for. Verbal definitions are vulnerable to linguistic uncertainty and prevent aggregation beyond simple vote counting (Game et al. 2013; Canessa et al. 2021). These limitations apply both across experts when opinions diverge and within each expert, when one feels a range of impacts could occur because they lack certain knowledge or because the ultimate outcome depends on stochastic circumstances and thus may vary. In such cases, asking an expert to indicate only one impact level ignores uncertainty altogether and causes a substantial loss of information.

Our approach circumvented these challenges by asking experts to estimate numerical probabilities directly. Experts were able not only to select all the impact levels they considered realistic but also to indicate precisely how likely they considered each level, which allowed intuitive representation and treatment of confidence both within and across experts. Numerical probabilities were superior to verbal or constructed scales in this sense because we could follow well-defined rules for aggregation, estimation, and visual representation. On their natural 0–1 scale, numerical probabilities are also easier and more transparent to directly compare among experts, communicate with partners and stakeholders, and easier to update with additional information in an adaptive framework (Runge 2011). This also applies to the expression of confidence, which we quantified as the probability of being correct (0–100), avoiding further discretization, such as "low-medium–high" confidence scales assigned to arbitrary probability levels (Probert et al. 2020; Clarke et al. 2021). Since risk is a product of outcome and probability, the estimates we obtained could be used, for example, in decision-support methods that represent uncertainty, like decision trees (Rout et al. 2013), or to analyse risk attitudes more formally, for example by using the principle of stochastic dominance (Game et al. 2013; Canessa et al. 2016a, b, c). Separating probabilities, impacts, and confidence also helps avoid hidden value judgments and improves decision-making (Game et al. 2013).

In addition to the uncertainty represented by the probabilities of different impact levels, we incorporated expert confidence by weighting experts, whereby the most confident had the greatest weight on the aggregated estimates. More commonly, uncertainty is expressed in Dirichlet-multinomial assessments by directly eliciting quantiles from experts (Zapata-Vázquez et al. 2014). In our case, this would have increased complexity beyond what we could require of our volunteer experts for a rapid assessment. Either way, confidence is self-expressed: this approach is intuitive and routinely used in the aggregation of elicited values (Hanea et al. 2018). However, it may reward overconfidence, a common cognitive bias of experts, particularly those with higher perceived expertise status (Burgman et al. 2011). Formal expert elicitation methods are designed specifically to reduce this overconfidence through consultation of multiple experts, anonymity, and explicit elicitation of confidence and uncertainty (Speirs-Bridge et al. 2010). Further, we still recommend including all methods (ignoring confidence, sampling

the weighted group distribution, and evaluating a consensus mean) for comparison, as we did in our study, or carrying out sensitivity analysis on weights to identify key shifts in impact or site rankings.

The impact levels themselves remain categorical, and thus, our approach falls short of a fully quantitative analysis, such as a complete epidemiological model for specific disease threats. However, the EICAT definitions ensure greater consistency than simply using ad hoc verbal definitions and allow a faster assessment than a quantitative one could ever be. Pre-assessment training helps further reduce inconsistencies, although it increases the time and effort required by facilitators and experts. In our case, a thorough check of all responses showed no evidence of misinterpretation or misclassification of impact mechanisms or levels. Regardless, for future assessments, especially those concerning more complex ecological communities, we will consider including formal pre-assessment training in EICAT. Our approach of using a fast, broad EICAT-based screening to guide where detailed assessments of high-risk impacts and management options are required is ideally suited for urgent and high-stakes decisions like conservation introductions of species at extreme risk of extinction.

Ultimately, risk assessments cannot make decisions about translocations, only inform them. The same expected risks and benefits could be acceptable to some, unacceptable to others; the same level of confidence may be regarded by some as sufficient, by others as unreliable (Tulloch et al. 2015). This conflict is recognized for invasive species in general (Vimercati et al. 2022), but dealing with endangered species is likely to add more complex emotional layers, potentially conducive to irrational decision making (Wintle et al. 2022). Fully understanding this decision space is difficult but will help make better decisions with more satisfying outcomes in the long term. Decision makers might adopt multi-criteria decision methods to balance different objectives (Adem Esmail and Geneletti 2018). Uncertainty and risk attitudes could be treated by adopting a general precautionary principle, by seeking to minimise worst-case outcomes, or by eliciting and analysing fully defined utility functions for all stakeholders (McCarthy 2014). These solutions will require both institutional support, such as better legislation and decision structures, and local effort for specific decisions.

Our approach flagged Palmyra Atoll as the site most suitable for sihek introduction from an environmental risk perspective, and highlighted predation and disease as the impact mechanisms to prioritize for further assessment and mitigation. More generally, the EICAT was a useful framework to assess translocation-related impacts across candidate sites, especially when combined with formal expert elicitation of quantitative probabilities. Using the predefined EICAT categories reduced linguistic uncertainty; the formal elicitation process allowed multiple opinions and sharing of knowledge; and the quantitative expression and aggregation helped us compare opinions and assess discrepancies and confidence levels. Risk assessments for biological invasions, including but not limited to those that might result from conservation translocations, seek to predict future events in novel systems and places, so expert judgment will always be necessary. Simple and rigorous ways of obtaining and summarising those judgments can increase reliability and improve subsequent decision-making.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

References

Adem EB, Geneletti D (2018) Multi-criteria decision analysis for nature conservation: a review of 20 years of applications. Methods Ecol Evolut 9(1):42–53. https://doi.org/10.1111/2041-210X.12899

Andersen MJ et al (2015) Rapid diversification and secondary sympatry in Australo-Pacific kingfishers (aves: alcedinidae: todiramphus). Royal Soc Open Sci 2(2):140375. https://doi.org/10.1098/rsos.140375

Auguie B, Antonov A (2017) *Miscellaneous Functions for "Grid'' Graphics*, *CRAN*. https://CRAN.R-project.org/package=gridExtra

Baker CM, Gordon A, Bode M (2017) Ensemble ecosystem modelling for predicting ecosystem response to predator reintroduction. Conserv Biol 31(2):376–384. https://doi.org/10.1111/cobi.12798

Beck RE, Savidge JA and Martin WE (1990) Native forest birds of Guam and rota of the Commonwealth of the Northern Mariana Islands recovery plan, U.S. Fish and Wildlife Service. https://ecos.fws.gov/docs/recovery_plan/900928b.pdf.

BirdLife International (2017) The IUCN Red List of Threatened Species^TM, *IUCN Red List* [Preprint]. Doi https://doi.org/10.2305/IUCN.UK.2017

Blackburn TM et al (2014) A unified classification of alien species based on the magnitude of their environmental impacts. PLoS Biol 12(5):e1001850. https://doi.org/10.1371/journal.pbio.1001850

Brooks SP, Gelman A (1998) 'General methods for monitoring convergence of iterative simulations)? J Computat Graph Stat 7(4):434–455. https://doi.org/10.1080/10618600.1998.10474787

Burgman MA et al (2011) Expert status and performance. PLoS ONE 6(7):e22998. https://doi.org/10.1371/journal.pone.0022998

Canessa S, Guillera-Arroita G et al (2016) Adaptive management for improving species conservation across the captive-wild spectrum. Biol Conserv 199:123–131. https://doi.org/10.1016/j.biocon.2016.04.026

Canessa S, Converse SJ et al (2016) Planning for ex situ conservation in the face of uncertainty. Conserv Biol 30(3):599–609. https://doi.org/10.1111/cobi.12613

Canessa S, Ewen JG et al (2016) Stochastic dominance to account for uncertainty and risk in conservation decisions. Conserv Lett 9:260–266. https://doi.org/10.1111/conl.12218

Canessa S, Trask AE, Ewen JG (2021) Mind the gap (between assessing risks and prioritizing management). NeoBiota 68:1–4. https://doi.org/10.3897/neobiota.68.60816

Cheng CW, Hung YC, Balakrishnan N (2014) Generating beta random numbers and Dirichlet random vectors in R: the package rBeta2009. Computat Stat Data Anal 71:1011–1020. https://doi.org/10.1016/j.csda.2013.02.019

Clarke DA et al (2021) Options for reducing uncertainty in impact classification for alien species. Ecosphere 12(4):e03461. https://doi.org/10.1002/ecs2.3461

Dehnen-Schmutz K et al (2022) Integrating expert knowledge at regional and national scales improves impact assessments of non-native species. NeoBiota 77:79–100. https://doi.org/10.3897/neobiota.77.89448

Engbring J, Fritts TH (1988) Demise of an insular avifauna: the brown tree snake on Guam. Trans West Sect Wildl Soc 24:31–37

Evans T (2018) Quantifying and categorising the environmental impacts of alien birds, PhD University College London

Game ET, Kareiva P, Possingham HP (2013) Six common mistakes in conservation priority setting. Conserv Biol 27(3):480–485. https://doi.org/10.1111/cobi.12051

Hanea AM et al (2018) The value of performance weights and discussion in aggregated expert judgments. Risk Anal 38(9):1781–1794. https://doi.org/10.1111/risa.12992

Hawkins CL et al (2015) Framework and guidelines for implementing the proposed IUCN environmental impact classification for alien taxa (EICAT). Diver Distribut 21(11):1360–1363. https://doi.org/10.1111/ddi.12379

Hemming V et al (2018) A practical guide to structured expert elicitation using the IDEA protocol. Methods Ecol Evolut 9(1):169–180. https://doi.org/10.1111/2041-210X.12857

Henry AK, Sorte CJB (2022) Impact assessment of coastal marine range shifts to support proactive management. Front Ecol Environ 20:161–169. https://doi.org/10.1002/fee.2447

Hoegh-Guldberg O et al (2008) Ecology: Assisted colonization and rapid climate change. Science 321:345–346. https://doi.org/10.1126/science.1157897

IUCN (2020) IUCN EICAT Categories and Criteria. The environmental impact classification for alien taxa (EICAT), 1. https://doi.org/10.2305/IUCN.CH.2020.05.en

IUCN/SSC (2013) *Guidelines for Reintroductions and Other Conservation Translocations*, *Species Survival Commission*. Gland, Switzerland.

Jenkins JM (1983) The native forest birds of Guam. Ornithol Monogr 31:61

Kesner D, Kumschick S (2018) Gastropods alien to South Africa cause severe environmental harm in their global alien ranges across habitats. Int J Bus Innov Res 17(3):8273–8285. https://doi.org/10.1002/ece3.4385

Kuhnert PM, Martin TG, Griffiths SP (2010) A guide to eliciting and using expert knowledge in Bayesian ecological models. Ecol Lett 13:900–914. https://doi.org/10.1111/j.1461-0248.2010.01477.x

Kumschick S et al (2020) Appropriate uses of EICAT protocol, data and classifications. NeoBiota 62:193–212. https://doi.org/10.3897/neobiota.62.51574

Laws RJ, Kesler DC (2012) A Bayesian network approach for selecting translocation sites for endangered island birds. Biolog Conserv 155:178–185. https://doi.org/10.1016/j.biocon.2012.05.016

Martin TG et al (2012) Eliciting expert knowledge in conservation science. Conserv Biol 26(1):29–38. https://doi.org/10.1111/j.1523-1739.2011.01806.x

McCarthy MA (2014) Contending with uncertainty in conservation management decisions. Ann New York Acad Sci 1322(1):77–91. https://doi.org/10.1111/nyas.12507

Mueller JM, Hellmann JJ (2008) 'An assessment of invasion risk from assisted migration. Conserv Biol 22(3):562–567. https://doi.org/10.1111/j.1523-1739.2008.00952.x

Plummer, M (2003) 'JAGS: just another Gibbs sampler', proceedings of the 3rd international workshop on distributed statistical computing (DSC 2003) [Preprint]. http://www.R-project.org/conferences/DSC-2003/

Probert AF et al (2020) Understanding uncertainty in the impact classification for alien taxa (ICAT) assessments. NeoBiota 62:387–405. https://doi.org/10.3897/neobiota.62.52010

R Core Team (2021) 'R: a language and environment for statistical computing. Vienne, Austria: R Foundation for Statistical Computing

Rout TM et al (2013) How to decide whether to move species threatened by climate change. PLoS ONE 8(10):e75814. https://doi.org/10.1371/journal.pone.0075814

Roy HE, Peyton JM, Booy O (2020) Guiding principles for utilizing social influence within expert-elicitation to inform conservation decision-making. Global Change Biol 26:3181–3184. https://doi.org/10.1111/gcb.15062

Runge MC (2011) An introduction to adaptive management for threatened and endangered species. J Fish Wildl Manag 2(2):220–233. https://doi.org/10.3996/082011-JFWM-045

Sainsbury AW, Vaughan-Higgins RJ (2012) Analyzing disease risks associated with translocations. Conserv Biol 26(3):442–452. https://doi.org/10.1111/j.1523-1739.2012.01839.x

Savidge JA (1986) The role of disease and predation in the decline of Guam's avifauna. PhD University of Illinois at Urbana-Champaign. https://hdl.handle.net/2142/77672

Seddon PJ (2010) From reintroduction to assisted colonization: moving along the conservation translocation spectrum. Restorat Ecol 18(6):796–802. https://doi.org/10.1111/j.1526-100X.2010.00724.x

Seddon PJ, Armstrong DP, Maloney RF (2007) Developing the science of reintroduction biology. Conserv Biol 21:303–312. https://doi.org/10.1111/j.1523-1739.2006.00627.x

Sohrabi S et al (2021) Quantifying the potential impact of alien plants of Iran using the Generic Impact Scoring System (GISS) and environmental impact classification for alien taxa (EICAT). Biolog Invas 23(8):2435–2449. https://doi.org/10.1007/s10530-021-02515-6

Speirs-Bridge A et al (2010) Reducing overconfidence in the interval judgments of experts. Risk Anal 30(3):512–523. https://doi.org/10.1111/j.1539-6924.2009.01337.x

Sutherland WJ (ed) (2022) Transforming conservation: a practical guide to evidence and decision making. Open Book Publishers, Cambridge, UK. https://doi.org/10.11647/OBP.0321

Sutherland WJ, Burgman MA (2015) Policy advice: use experts wisely. Nature 526(7573):317–318. https://doi.org/10.1038/526317a

Trask AE et al (2021) 'Multiple life-stage inbreeding depression impacts demography and extinction risk in an extinct-in-the-wild species. Sci Rep 11(1):682. https://doi.org/10.1038/s41598-020-79979-4

Trask AE, Canessa S, Ewen JG (2019) Sihek conservation translocation planning workshop 1: summary report. Guam, 13–16 February 2019. Unpublished report to Guam Department of Agriculture

Tulloch AIT et al (2015) Effect of risk aversion on prioritizing conservation projects. Conserv Biol 29(2):513–524. https://doi.org/10.1111/cobi.12386

Turbé A et al (2017) Assessing the assessments: evaluation of four impact assessment protocols for invasive alien species. Diver Distrib 23(3):297–307. https://doi.org/10.1111/ddi.12528

Vanderhoeven S et al (2017) Beyond protocols: improving the reliability of expert-based risk analysis underpinning invasive species policies. Biolog Invas 19(9):2507–2517. https://doi.org/10.1007/s10530-017-1434-0

Vimercati G et al (2022) The EICAT+ framework enables classification of positive impacts of alien taxa on native biodiversity. PLoS Biol 20(8):e3001729. https://doi.org/10.1371/journal.pbio.3001729

Volery L et al (2020) Improving the environmental impact classification for Alien Taxa (EICAT): a summary of revisions to the framework and guidelines. NeoBiota 62:547–567. https://doi.org/10.3897/neobiota.62.52723

Wickham H (2011) ggplot2. Interdisci Rev Computat Statist 3(2):180–185. https://doi.org/10.1002/wics.147

Wickham H et al (2019) Welcome to the Tidyverse. J Open Sour Softw 4(43):1686. https://doi.org/10.21105/joss.01686

Wickham H, Bryan J (2023) readxl: Read excel files, CRAN. https://readxl.tidyverse.org, https://github.com/tidyverse/readxl.

Wilson MW et al (2020) Ecological impacts of human-induced animal behaviour change. Ecol Lett 23(10):1522–1536. https://doi.org/10.1111/ele.13571

Wilson KJ, Elfadaly FG, Garthwaite PH, Oakley JE (2021) Recent advances in the elicitation of uncertainty distributions from experts for multinomial probabilities. In: Hanea AM, Nane GF, Bedford T, French S (eds) Expert judgement in risk and decision analysis. Springer International Publishing, Cham, pp 19–51. https://doi.org/10.1007/978-3-030-46474-5_2

Wintle BC et al (2019) Verbal probabilities: very likely to be somewhat more confusing than numbers. PLoS ONE 14(4):e0213522. https://doi.org/10.1371/journal.pone.0213522

Zapata-Vázquez RE, O'Hagan A, Soares Bastos L (2014) Eliciting expert judgements about a set of proportions. J Appl Statist 41(9):1919–1933. https://doi.org/10.1080/02664763.2014.898131