Conservation Biology

Contributed Paper

Using expert knowledge to support Endangered Species Act decision-making for data-deficient species

Daniel B. Fitzgerald [1*], David R. Smith,[1] David C. Culver,[2] Daniel Feller,[3] Daniel W. Fong,[4] Jeff Hajenga,[5] Matthew L. Niemiller,[6] Daniel C. Nolfi,[7] Wil D. Orndorff,[8] Barbara Douglas,[9] Kelly O. Maloney,[1] and John A. Young[1]

[1]U.S. Geological Survey, Leetown Science Center, 11649 Leetown Road, Kearneysville, WV 25430, U.S.A.
[2]Department of Environmental Science, American University, 4400 Massachusetts Avenue NW Washington, DC 20016, USA
[3]Maryland Department of Natural Resources, Appalachian Laboratory, 301 Braddock Rd, Frostburg, MD 21532, USA
[4]Department of Biology, American University, 4400 Massachusetts Avenue, NW, Washington, DC 20016, USA
[5]West Virginia Division of Natural Resources, PO Box 67, Ward Rd, Elkins, WV 26241, USA
[6]Department of Biology, University of Alabama Huntsville, SST 369, 301 Sparkman Drive, Huntsville, AL 35899, USA
[7]U.S. Fish and Wildlife Service, 4425 Burley Dr # A, Chubbuck, ID 83202, USA
[8]Virginia Department of Conservation and Recreation, 600 E. Main St., 24th Floor, Richmond, VA 23219, USA
[9]U.S. Fish and Wildlife Service, 90 Vance Drive, Elkins, WV 26241, USA

Abstract: Many questions relevant to conservation decision-making are characterized by extreme uncertainty due to lack of empirical data and complexity of the underlying ecologic processes, leading to a rapid increase in the use of structured protocols to elicit expert knowledge. Published ecologic applications often employ a modified Delphi method, where experts provide judgments anonymously and mathematical aggregation techniques are used to combine judgments. The Sheffield elicitation framework (SHELF) differs in its behavioral approach to synthesizing individual judgments into a fully specified probability distribution for an unknown quantity. We used the SHELF protocol remotely to assess extinction risk of three subterranean aquatic species that are being considered for listing under the U.S. Endangered Species Act. We provided experts an empirical threat assessment for each known locality over a video conference and recorded judgments on the probability of population persistence over four generations with online submission forms and R-shiny apps available through the SHELF package. Despite large uncertainty for all populations, there were key differences between species' risk of extirpation based on spatial variation in dominant threats, local land use and management practices, and species' microhabitat. The resulting probability distributions provided decision makers with a full picture of uncertainty that was consistent with the probabilistic nature of risk assessments. Discussion among experts during SHELF's behavioral aggregation stage clearly documented dominant threats (e.g., development, timber harvest, animal agriculture, and cave visitation) and their interactions with local cave geology and species' habitat. Our virtual implementation of the SHELF protocol demonstrated the flexibility of the approach for conservation applications operating on budgets and time lines that can limit in-person meetings of geographically dispersed experts.

Keywords: expert elicitation, extinction risk, remote elicitation, SHELF, species status assessment, *Stygobromus*

Uso del Conocimiento Experto para Respaldar la Toma de Decisiones del Acta de Especies en Peligro para Especies con Información Deficiente

Resumen: Muchas preguntas relevantes para la toma de decisiones de conservación se caracterizan por una incertidumbre extrema causada por la falta de información empírica y por la complejidad de los procesos ecológicos

*email dbfitzgerald@usgs.gov
Article Impact Statement: Remote expert elicitation can facilitate endangered species decision-making when available data, budgets, and time frames are limiting.
Paper submitted October 30, 2020; revised manuscript accepted January 12, 2021.

Conservation Biology, Volume 35, No. 5, 1627–1638

DOI: 10.1111/cobi.13694

subyacentes. Esto lleva a un rápido incremento en el uso de protocolos estructurados para obtener conocimiento de los expertos en el tema. Las aplicaciones ecológicas publicadas con frecuencia emplean un método Delphi modificado, en el cual los expertos proporcionan dictámenes anónimamente y luego se usan técnicas de agregación matemática para combinar estos dictámenes. El marco de trabajo de obtención Sheffield (SHELF) difiere en su enfoque conductual para sintetizar los dictámenes individuales en una distribución de probabilidad completamente especificada para una cantidad desconocida. Usamos el protocolo SHELF remotamente para evaluar el riesgo de extinción de tres especies acuáticas subterráneas que están siendo consideradas para ser incluidas en el Acta de Especies en Peligro de los E.U.A. Les proporcionamos a los expertos una evaluación empírica de la amenaza para cada localidad conocida durante una videoconferencia y registramos los dictámenes sobre la probabilidad de la persistencia poblacional durante cuatro generaciones por medio de formularios enviados en línea y las apps R-shiny disponibles a través del paquete SHELF. A pesar de la gran incertidumbre para todas las poblaciones, hubo diferencias importantes entre el riesgo de extirpación de las especies con base en la variación espacial en las amenazas dominantes, el uso del suelo local y las prácticas de manejo, y el microhábitat de las especies. Las distribuciones resultantes de la probabilidad proporcionaron al órgano decisorio un cuadro completo de la incertidumbre que fue consistente con la naturaleza probabilística de las evaluaciones de riesgo. Las discusiones entre los expertos durante la fase de agregación conductual de SHELF documentaron claramente las amenazas dominantes (p. ej.: desarrollo, extracción de madera, agricultura animal y visitas a las cuevas) y sus interacciones con la geología de las cuevas locales y el hábitat de la especie. Nuestra implementación virtual del protocolo SHELF demostró la flexibilidad del enfoque para las aplicaciones de la conservación que operan con presupuestos y líneas de tiempo que pueden limitar las reuniones en persona de expertos dispersados geográficamente.

Palabras Clave: evaluación del estado de la especie, obtención de expertos, obtención remota, riesgo de extinción, SHELF, *Stygobromus*

【摘要】: 由于经验数据的缺乏和潜在生态过程的复杂性, 许多与保护决策相关的问题具有极端不确定性, 这导致使用结构化研究方案来获得专家知识的实践迅速增加。在已有生态学应用中, 研究者通常采用一种改进的德尔菲法, 即用数学整合方法来综合专家匿名提供的判断。而谢菲尔德启发式框架 (SHELF) 的不同之处在于, 它的行为方法是将个人判断整合为一个未知量完全指定的概率分布。我们利用 SHELF 方法远程评估了正被考虑列入《美国濒危物种法案》的 3 种地下水生物种的灭绝风险。我们通过视频会议向专家提供了对每个已知分布点威胁的经验评估, 并通过在线提交表格和 SHELF 包的 R-shiny 应用程序记录了专家对这些物种 4 代以上续存概率的判断。虽然所有种群都存在很大的不确定性, 但由于其主要威胁、当地土地利用及管理措施和物种微生境的空间差异, 物种灭绝风险也存在明显差异。由此得出的概率分布为决策者提供了不确定性的整体情况, 这与风险评估的概率特性相一致。专家们在 SHELF 框架整合阶段的讨论清楚地记录了主要威胁 (如开发、木材采伐、畜牧业和探洞) 以及它们与当地洞穴地质和物种栖息地的相互作用。我们对 SHELF 框架的虚拟实现展示了保护应用在预算和时间上的灵活性, 从而可以减少分散在各地的专家的现场会议。**【翻译: 胡怡思; 审校: 聂永刚】**

关键词: 专家启发, 灭绝风险, 远程启动, 谢菲尔德启发式框架, 物种濒危状况评估, *Stygobromus* 属

Introduction

Many questions relevant to conservation decision-making are characterized by extreme uncertainty due to lack of empirical data and complexity of the underlying ecological processes (Kuhnert et al. 2010). Rare and at-risk species often lack the quantitative data needed to detect temporal trends in demographics (Bland et al. 2015; Kindsvater et al. 2018); yet, legislation such as the U.S. Endangered Species Act (ESA) may mandate that listing decisions be conducted over time frames not compatible with additional long-term data collection. Although some extinction-risk assessment frameworks provide mechanisms for classifying species as data deficient (e.g., International Union for Conservation of Nature [IUCN] Red List of Threatened Species), no equivalent category exists under the ESA once a substantial 90-day finding is made, and delayed decisions can result in costly legal actions (Stokstad 2005).

Expert knowledge is widely used to inform conservation decisions and may provide the only path forward for data-deficient species at the science–policy interface (Burgman 2004; Sutherland 2006). For example, the IUCN global assessments frequently rely on expert knowledge contributed through facilitated, regional workshops to comprehensively assess particular taxonomic groups (Lacher et al. 2012). The ubiquity of data gaps in species assessments and limited funding for new data collection can be linked to a rapid increase in the use of structured protocols to elicit expert knowledge in conservation (Martin et al. 2012; Drescher et al. 2013), including recent applications to assess extinction risk of birds and koalas, among many others (McBride et al. 2012; Adams-Hosking et al. 2016). In addition to reducing biases in probabilistic judgments, these protocols can provide a transparent and well-documented process for capturing uncertainty, which is critical for ecological applications that support policy decisions (Dias et al. 2018; O'Hagan 2019).

Several protocols exist for eliciting and combining judgments of multiple experts on an unknown quantity of interest, all with methods for capturing initial differences of opinion, recording uncertainty within judgments, and minimizing the influence of common cognitive biases (Dias et al. 2018; O'Hagan 2019). Many published ecological applications have employed a modified Delphi method in which experts provide judgments anonymously and mathematical aggregation (e.g., linear or weighted pooling) is required to combine judgments (Sutherland 2006; Kuhnert et al. 2010; McBride et al. 2012; Adams-Hosking et al. 2016; Hemming et al. 2018). The Sheffield elicitation framework (SHELF) differs from other leading protocols in its behavioral approach to synthesizing individual judgments into a fully specified probability distribution for an unknown quantity (Gosling 2018; O'Hagan 2019; Oakley & O'Hagan 2019). Following an initial anonymous judgment round, experts participate in open discussions focused on understanding the reasoning behind differing opinions before the group is asked to collectively provide judgments from the perspective of a rational impartial observer (RIO) (details in Methods). Although this open discussion requires careful facilitation to reduce certain cognitive biases (e.g., groupthink, overconfidence, and halo effects), the behavioral approach highlights key factors influencing uncertainty and clarifies that the final aggregate probability distribution represents a RIO's subjective beliefs (O'Hagan 2019). The SHELF protocol's readily accessible software and forms for documenting and recording the elicitation also meet the need for transparency in ESA assessments and other public decisions.

We applied the SHELF protocol to an extinction-risk assessment of 3 subterranean aquatic species petitioned for listing under the ESA: Cooper's cave amphipod (*Stygobromus cooperi*), minute cave amphipod (*Stygobromus parvus*), and Morrison's cave amphipod (*Stygobromus morrisoni*). Subterranean aquatic species (i.e., stygobionts) that occur in caves and shallow epikarst exemplify the need for expert opinion in assessments due to their rarity and the inaccessibility of their primary habitat (Pipan et al. 2010). For example, of the 33 stygobionts that occur in the state of West Virginia, 7 are known from fewer than 10 specimens (Fong et al. 2007). Although stygobionts are commonly thought to be K-selected species with delayed maturity, small population size, and low reproductive rates (Poulson & White 1969), specific demographic and life-history parameters are unavailable for most species. *S. parvus*, *S. cooperi*, and *S. morrisoni* are restricted to portions of Virginia and West Virginia (Lewis 2001; Fong et al. 2007; Holsinger et al. 2013); *S. parvus* localities extend across 1,467 km^2; and *S. morrisoni* localities extend across 2,266 km^2. *S. cooperi* is a single-site endemic known from only 3 specimens (Fong et al. 2007). Available data include opportunistic point localities of amphipod occurrence (Fig. 1), which cannot be reliably used to infer current population size, condition, or temporal trends in the occupied range and may date as far back as 1966.

The broad categories of threats to cave and karst biota are well known (Mammola et al. 2019). For example, *S. mackini* have shown occurrence patterns consistent with negative impacts of groundwater pollution by septic systems in Banner Cave, Virginia (Simon & Buikema 1997), and toxic pollutant spills represent a primary threat to many karst species (Loop & White 2001; Pipan et al. 2010). The magnitude of various stressors that *Stygobromus* populations can withstand, however, represents a critical source of uncertainty in the assessment process. Data from the Edwards Aquifer region of Texas suggest that the 10–15% impervious cover threshold for degradation of biological communities in surface waters represents a reasonable starting point for estimating impacts to subterranean species in the absence of karst-specific information (Veni 1999). Similar estimates for how other stressors may affect *Stygobromus* spp. resilience are lacking.

The recently implemented species status assessment (SSA) framework has shifted U.S. Fish & Wildlife Service (USFWS) assessments from a threats-focused analysis to one that explicitly considers species' responses to current and projected stressors (USFWS 2016b; Smith et al. 2018). Because few empirical data are available on these species' historical or current conditions, expert knowledge was used to obtain a scientific assessment of how populations may respond to current and projected levels of major stressors and to estimate the uncertainty in species' future viabilities. In addition to supporting the ESA decision-making process for 3 petitioned *Stygobromus* spp., we sought to provide an example of how the SHELF protocol can be applied to data-deficient species within the SSA framework.

Methods

Structuring the Quantities of Interest

Empirical data are available to estimate levels of several major threats based on proxy variables (e.g., percentage of various land-use classes, number of mining operations [Appendix S1]). The missing quantities of interest are species' responses to various stressor levels. Experts discussed multiple approaches for structuring the elicited quantity and determined that estimating the probability of persistence based on the empirical habitat conditions for each locality was most tractable. Specifically, future viability was assessed as the probability of persistence

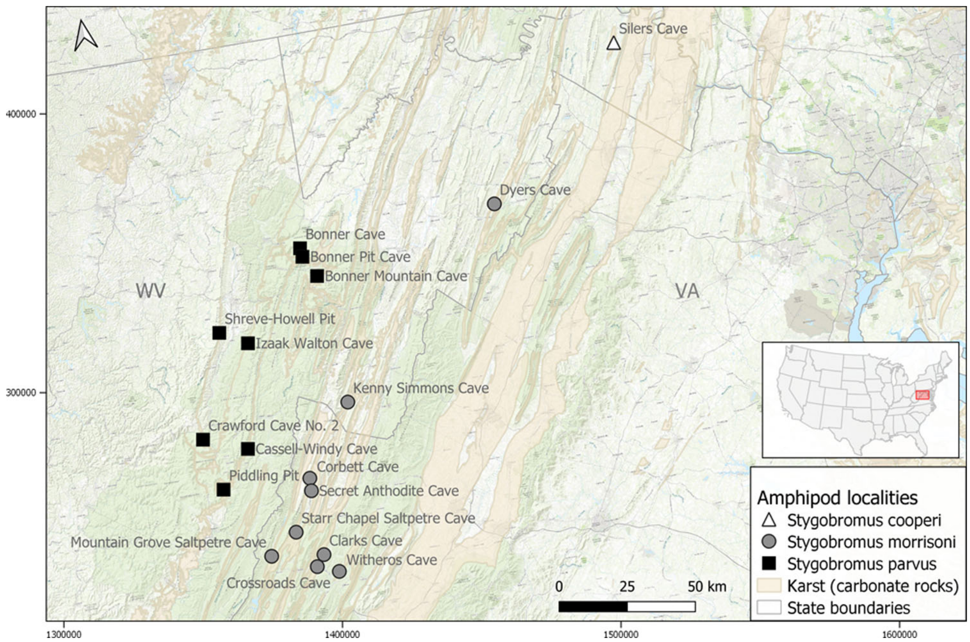

Figure 1. Known localities for 3 Stygobromus *spp. in relation to exposed karst areas. Data are from Fong et al. (2007) and Holsinger et al. (2013) and are referenced to the Albers Equal Area North American Datum 1983 Coordinate Reference System (EPSG = 42303).*

of each metapopulation (i.e., locality) over 4 generations (roughly 10−20 years). Because so little is known about current population conditions, no attempt was made to further classify a populations' future status beyond persistence or extirpation.

We assumed each locality represented a distinct metapopulation. *S. morrisoni* is currently known from 9 localities, *S. cooperi* from 1 and *S. parvus* from 8. Data from epikarst copepods suggest that populations generally extend <1 km along a cave passage (Pipan & Culver 2007) and that genetic differentiation or metapopulation structure can be detectable at scales as small as tens of meters (Sbordoni et al. 2000). The minimum nearest neighbor distance between known localities was 3.1 and 4.3 km for *S. parvus* and *S. morrisoni*, respectively; values ranged up to 88 km. Karst areas between known localities have not been sampled adequately, and the true extent of each metapopulation is a major source of uncertainty. This uncertainty will affect estimates of the number of populations, the potential for genetic connectivity, and the likelihood of persistence for each population.

The appropriate spatial extent for considering threats to population resilience represents another major source of uncertainty. Experts were not aware of dye-tracing studies from the identified caves that could aid delineation of groundwater influence zones. Although the boundaries of groundwater basins frequently deviate from surface basins, impacts to surface waters in karst areas affect groundwater quality. For example, cave streams in West Virginia have elevated nitrate and pesticide levels in agricultural areas (Boyer & Pasquarell 1995; Pasquarell & Boyer 1996). Protection of surface areas is critical for the conservation of subterranean fauna, particularly for epikarst specialists, such as *S. parvus* and *S. cooperi* (Culver et al. 2000; Pipan et al. 2010).

Available information on historical and current conditions and individual site threat assessments that provided visual and numeric summaries of past, current, and projected stressor proxies were used as the basis for expert judgments. Information on threat proxies was displayed at several spatial scales due to experts' beliefs that the appropriate scale depends on the specific threat being assessed. At the smallest spatial extent, the catchment area of individual epikarst drips are generally less than a few hundred square meters (Pipan & Culver 2013). Therefore, a 1-km^2 closeup of each locality showing aerial imagery from 2019 (U.S. National Agriculture Imagery Program) was used to assess current land use in the immediate vicinity (Appendix S1). Based on data from epikarst copepods (Pipan et al. 2010), a 1-km area around sampling localities was assumed to represent the potential area occupied by each metapopulation. Local catchments intersecting this area and their upstream watersheds represented other potentially relevant scales because surface water can act as a vector for contaminants moving downstream and laterally through karst environments. Aerial imagery was shown at the local catchment scale. Land-use statistics were quantified at the upstream watershed scale in 2006 and 2016 based on National Landcover Data (NLCD 2016) and projected to 2030 based on the Intergovernmental Panel on Climate Change emission scenarios A1B, A2, B1, and B2 (Sohl et al. 2018). Experts were also provided regional land-use projections out to 2040, due to uncertainty in generation time, and regional projections of precipitation based on 20 climate models (Abatzoglou 2013). Surface catchments were defined based on the U.S. National

Hydrography Plus (NHDPlus Version 2) data set. The largest spatial extent provided a 2016 land-use model (NLCD 2016) and the locations of mining operations (EIA 2020), dams (USACE 2020), and impaired surface waters (EPA 2020) at least 10 km away from known localities due to high uncertainty in subsurface movement of water through karst environments and the possibility for metapopulation dynamics between surrounding karst regions not sampled directly.

SHELF Elicitation Workshop

A critical step in expert knowledge elicitation (EKE) is identifying and recruiting the appropriate expert panel, which typically includes 4–8 participants in the SHELF protocol (O'Hagan 2019). The involvement of multiple experts provides decision makers with a diversity of perspectives and helps reduce the risk of overconfidence in judgments by any single expert. Potential experts were identified based on experience sampling the species, peer-reviewed publications on the genus *Stygobromus*, and professional involvement in conservation and management of karst biota. Out of 14 experts initially contacted, 7 were available for participation (coauthors 3–9 [Appendix S2]). Experts were provided training in quantifying personal beliefs through materials from the SHELF protocol (Oakley & O'Hagan 2019) and an example quantity of interest formulated as the probability that the maximum age of *S. cooperi* is >6 years.

The EKE workshop was carried out remotely in a series of 2-hour video conference calls (8 hours total) from 9 June to 2 July 2020. For each locality, experts were first provided an evidence dossier summarizing the current conditions and threats assessment in order to reduce the availability bias (Oakley & O'Hagan 2019). During the individual judgment round, experts used the quartile method to provide private judgments for the probability that the metapopulation surrounding the locality would persist over 4 generations. This approach begins by specifying an upper and lower plausible limit to counter overconfidence and anchoring effects (O'Hagan 2019), followed by sequential implementations of the bisection method (Raiffa 1968) to provide a median, upper quartile, and lower quartile. Although uncertainty about an event can be described by a single probability with SHELF methods for discrete quantities, experts may be unwilling to provide single estimates. The quartile method allowed experts to express uncertainty in their probabilities and provided decision makers with an indication of how robust the expected values may be to new information. Judgments were submitted privately via an online form and probability distributions were fit by minimizing the sum of squared differences between elicited and fitted probabilities along the cumulative distribution function with the SHELF package in R (Oakley 2019). Certain populations (Corbett and Secret Anthodite Caves, Mountain Grove and Starr Chapel Saltpetre Caves, and Bonner Mountain and Bonner Pit Caves) were assessed simultaneously due to similarities in observed land use identified through *k*-means clustering and their geographic proximity within connected regions of karst.

Experts were then led through a facilitated group discussion where they provided the reasoning behind judgments, including which stressors were of greatest concern and which factors generated the most uncertainty. During the group judgment round, experts were asked to provide new quartiles from the perspective of a RIO who had listened to their discussion and understood their arguments (O'Hagan 2019). Due to concerns over potential dominance or halo effects within the group (i.e., discussions conform to ideas of a forceful or esteemed member), combined with the remote nature of the workshop, each expert first privately provided a RIO judgment by using the procedure described above. These judgments and their linear pool were then used as feedback during a facilitated discussion to select the final RIO quartiles. This slight modification to the SHELF protocol required each expert to provide twice as many judgments, but ensured that no single expert dominated selection of the group quartiles. Finally, a scaled beta distribution was fit to the final RIO quartiles to represent the collective uncertainty in the probability of persistence.

Results

The experts had low confidence in persistence for the single *S. cooperi* population, reflected by both the low median value and large degree of uncertainty in estimates (Table 1, Fig. 2, Appendix S2). Low confidence in persistence stemmed from the high number of developed areas within the region and an anticipated increase in visitation rates due to recent changes in cave ownership (Appendix S2). The isolated nature of this locality constrains the potential area the metapopulation could occupy, resulting in a high risk of extirpation from stochastic events. Although population estimates were not available, sampling experience suggests *S. cooperi* is consistently collected at lower densities than other *Stygobromus* spp. in the region and that even moderate increases in threats could have severe consequences for the population (Appendix S2). The higher median probability of persistence provided by expert F (Fig. 2) reflects experience from other regions where epikarst *Stygobromus* species were among the last taxa to persist in caves affected by similar threats. However, all experts agreed that this population had the lowest probability of persistence among assessed caves, as reflected by the lower 90% credible interval of 0.19 for the final RIO distribution.

Table 1. Summary statistics for the probability of persistence of the *Stygobromus* metapopulation surrounding each occurrence.

Species	*Locality*	*Median*	*90% CI*	*p*(x > 0.5)	*Confidence in persistence**
S. cooperi	Silers Cave	0.48	0.19–0.77	0.47	very low
S. morrisoni	Dyers Cave	0.60	0.33–0.83	0.71	medium
	Kenny Simmons Cave	0.47	0.21 - 0.71	0.44	very low
	Corbett Cave	0.67	0.37 - 0.90	0.81	medium
	Secret Anthodite Cave	0.67	0.37- 0.90	0.81	medium
	Mountain Grove Saltpetre Cave	0.73	0.46- 0.91	0.92	high
	Starr Chapel Saltpetre Cave	0.73	0.46–0.91	0.92	high
	Clarks Cave	0.55	0.27–0.78	0.61	low
	Crossroads Cave	0.60	0.35–0.83	0.74	medium
	Witheros Cave	0.76	0.46–0.95	0.92	high
S. parvus	Bonner Cave	0.76	0.48–0.94	0.93	high
	Bonner Mountain Cave	0.72	0.44–0.90	0.91	high
	Bonner Pit Cave	0.72	0.44–0.90	0.91	high
	Shreve-Howell Pit	0.67	0.34–0.91	0.78	medium
	Izaak Walton Cave	0.78	0.49–0.94	0.95	high
	Crawford Cave No. 2	0.71	0.32–0.95	0.80	medium
	Cassell-Windy Cave	0.73	0.37–0.96	0.85	medium
	Piddling Pit	0.81	0.58–0.94	0.99	high

**Classified based on the p*(x > 0.5) *as follows: very low, < 0.5; low, 0.5–0.7; medium, 0.7–0.9; high, > 0.9.*

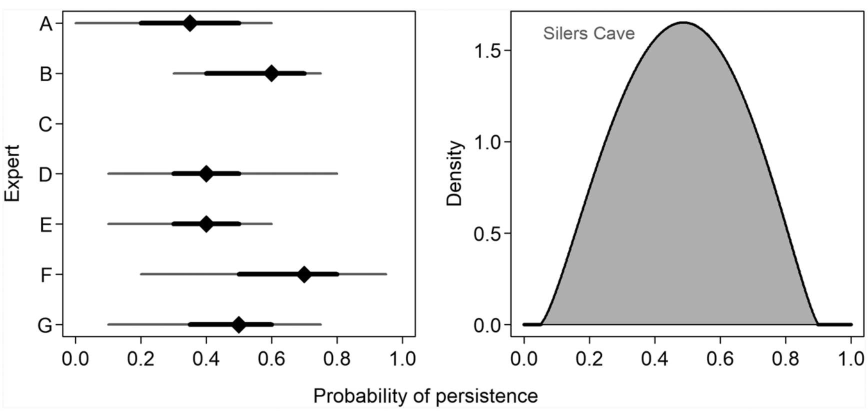

Figure 2. Expert judgments on the probability of persistence of Stygobromus cooperi *at the single known locality where it occurs: (left) individual expert judgments for the median (diamond), 50% credible interval (thick black line), and plausible limits (thin gray line) and (right) final probability distribution from the group judgment round in the Sheffield elicitation framework. Expert C was unable to provide judgments for this locality due to video conference connectivity issues, but agreed that the final distribution captures the risk and uncertainty for this locality.*

Expert judgments generally expressed high confidence in future persistence for *S. parvus* populations (Table 1 & Fig. 3). Indeed, the Piddling Pit metapopulation had the highest confidence in persistence among any locality assessed; 90% credible intervals ranged from 0.58 to 0.94 based primarily on land ownership by a conservation organization and restricted cave access. Although several localities are surrounded by mostly intact forests, threats from agricultural land use and mining were present in the region and contributed to differences in uncertainty among localities (Appendix S2). For example, the presence of past limestone mining and agriculture in the vicinity of Crawford Cave No. 2 led experts to judge that the probability of persistence could credibly be as low as 0.32, despite a high median value. Lower 90% credible intervals were similarly low for Shreve-Howell Pit and Cassell-Windy Cave. These uncertainties are critical to consider when assessing potential risk. Logging, mining, and agriculture have occurred throughout the range historically, and observations of the species within the last 30 years were taken as evidence that populations can persist amidst historical levels of disturbance. During group discussions, however, experts frequently pointed to differences in contemporary agricultural practices, such as the increase in large-scale poultry farming, the application of poultry manure as fertilizers, and use of industrial herbicides and pesticides. Other species in the genus *Stygobromus* have been negatively affected by concentrated waste products in septic systems in Virginia (Simon & Buikema 1997), and decreased certainty in estimates for several *S. parvus* populations reflected concern that similar effects could arise from

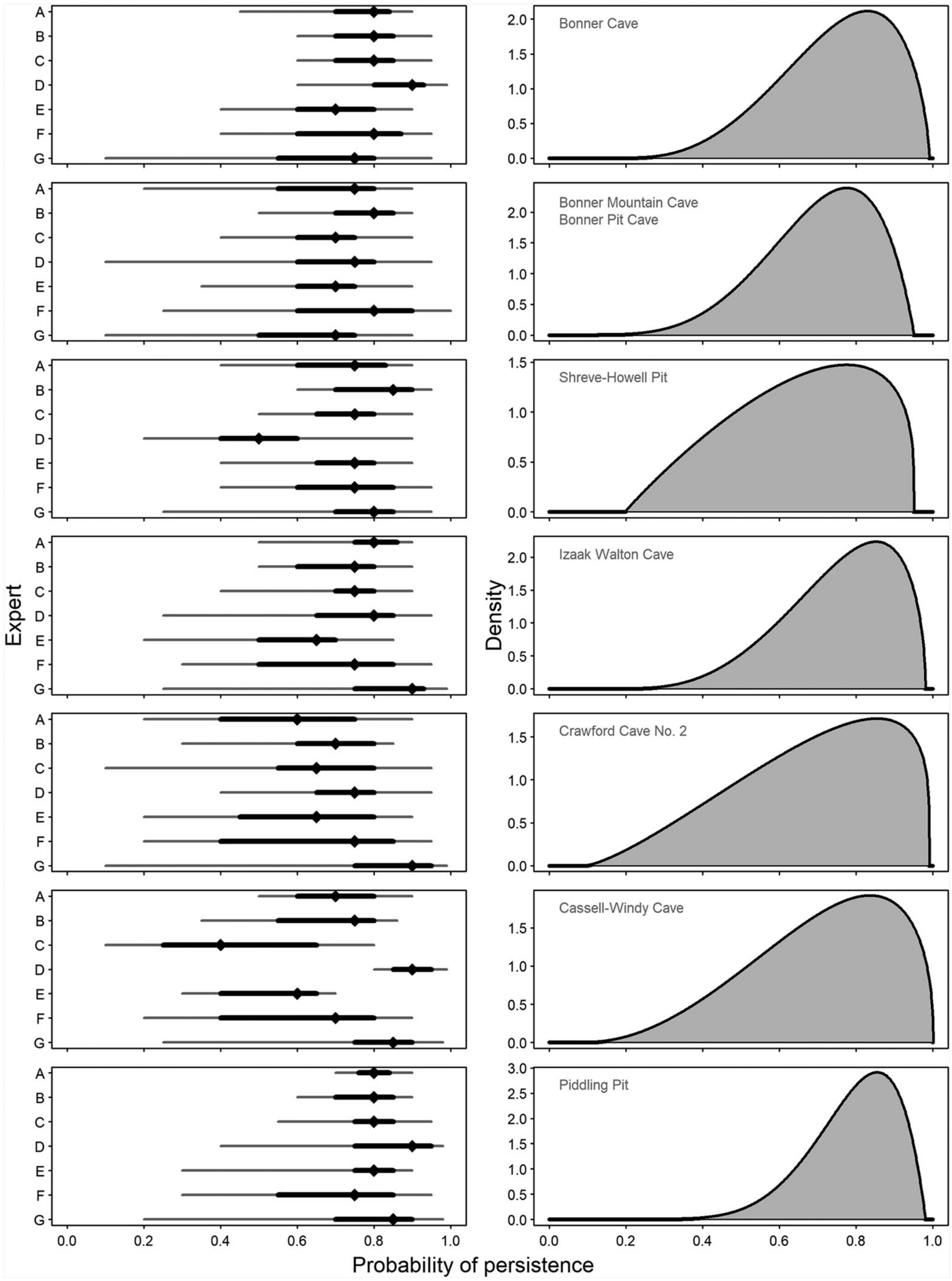

Figure 3. Expert judgments on the probability of persistence of Stygobromus parvus *at localities where it occurs: (left) individual expert judgments for the median (diamond), 50% credible interval (thick black line), and plausible limits (thin gray line) and (right) final probability distributions from the group judgment round in the Sheffield elicitation framework.*

increased animal agriculture. The possibility for emerging threats, such as disease and an unknown likelihood of stochastic events, also contributed to uncertainty in estimates and should be considered when assessing species future viability. The 2 localities with the lowest median values and lowest certainty (Crawford Cave No. 2 and Shreve Howell Pit) occurred in a distinct western band of karst, suggesting that the greatest risk to persistence occurs in the western portion of the *S. parvus* range (Fig. 1).

For *S. morrisoni*, expected risk to persistence varied widely among localities (Table 1 & Fig. 4), based primarily on differences in land ownership and agricultural intensity. Expert judgments suggested that populations

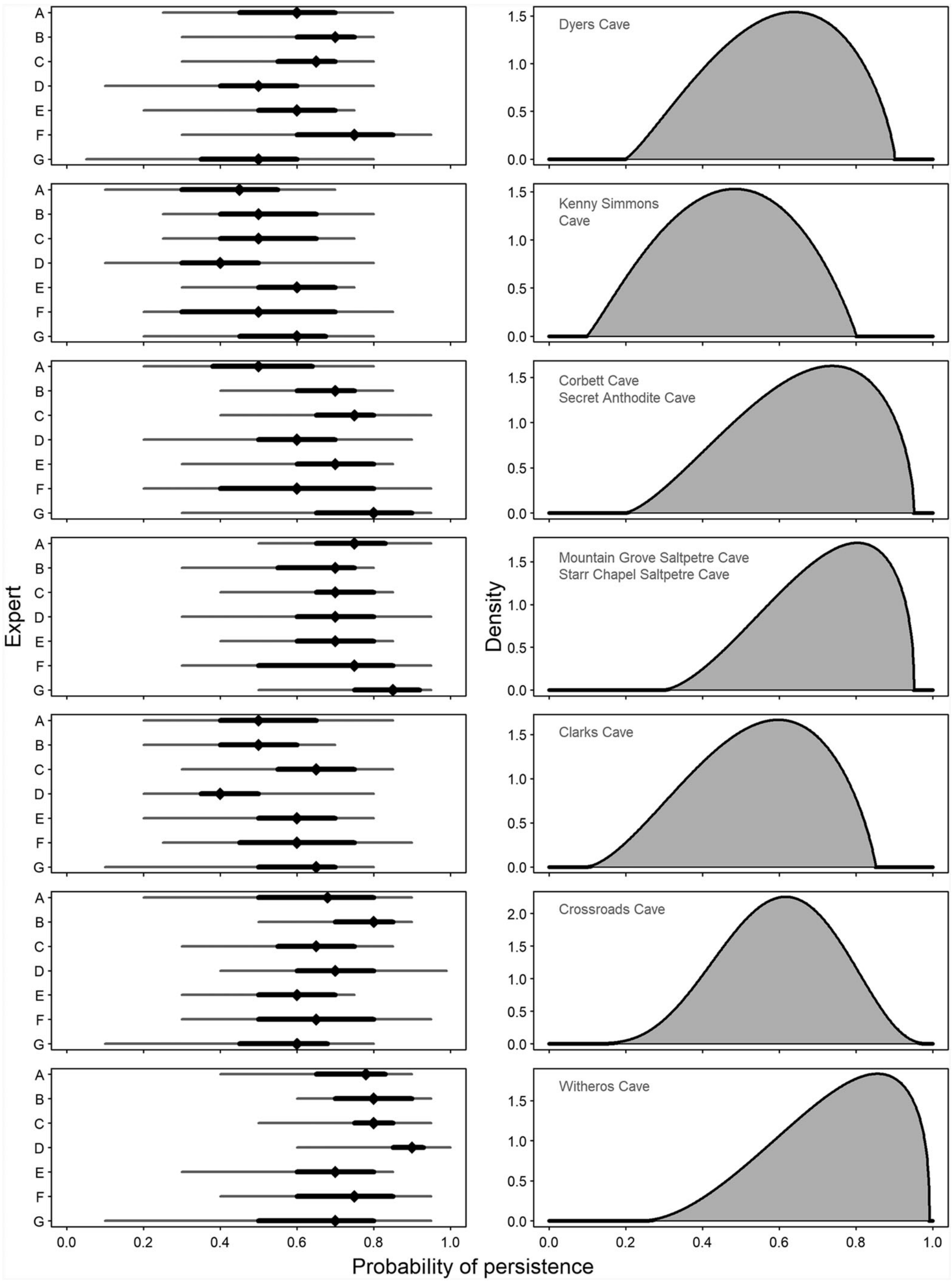

Figure 4. Expert judgments on the probability of persistence of Stygobromus morrisoni *at localities where it occurs: (left) individual expert judgments for the median (diamond), 50% credible interval (thick black line), and plausible limits (thin gray line) and (right) the final probability distributions from the group judgment round in the Sheffield elicitation framework.*

in 3 out of the 9 known localities could be characterized as having high confidence in persistence, whereas populations in 2 localities faced much greater risks to persistence (Table 1). For example, the estimates for both Mountain Grove and Starr Chapel Saltpetre Caves suggested populations in these localities are likely to persist despite uncertainty in current conditions. The lack of information on historical conditions at these sites was assumed to reflect sampling deficiencies rather than the historic or current condition of the populations. In contrast, the population in Kenny Simmons Cave had among the lowest confidence of persistence due to the lack of available survey data and proximity to high-intensity animal agriculture. Importantly, the projected risk was not distributed evenly across the range of *S. morrisoni*; 3 of the populations with the greatest risk comprised all known occurrences in a central karst belt (Clarks and Crossroads Caves) and a disjunct northern portion of the range (Dyers Cave). Populations in all 3 localities had high levels of uncertainty in persistence reflected by wide confidence intervals and median estimates of the probability of persistence from 0.55 to 0.60 (Table 1), suggesting a future risk of increased fragmentation between extant populations.

Discussion

Although uncertainty exists in the estimated risk of extirpation for all populations, expert judgments provided useful information on the relative magnitude of uncertainty that can aid decision-making. For example, the inherently greater risk of extinction for the narrowly endemic *S. cooperi* is compounded by the fact that the single metapopulation was judged to have the lowest confidence in persistence of any locality assessed (Table 1 & Fig. 2). This reflected expert judgments that proximity to development was among the greatest threats to persistence, including associated increases in cave visitation, surface alteration, and increased risk of pollution and contaminants. Conversely, expert judgments for *S. parvus* localities generally suggested high confidence in persistence despite considerable uncertainty in estimates for several metapopulations (Fig. 3). The greatest threat to persistence for most *S. parvus* populations related to forest management and timber harvest practices by landowners. Uncertainty in the probability of persistence was generally much higher for populations of *S. morrisoni*, based on suspected differences in habitat use and variation in the characteristics of occupied caves.

All 3 species can be considered groundwater habitat specialists, but important differences in microhabitat use may affect species' vulnerability to common threats. *Stygobromus parvus* and *S. cooperi* are epikarstic species, occupying water percolating through the uppermost layer of karst at the rock-soil interface. Epikarst is often characterized by greater organic matter inputs and environmental variation relative to deeper subterranean habitats (Culver et al. 2010) and is more susceptible to environmental degradation on the surface (Pipan et al. 2010). Although *S. parvus* and *S. cooperi* are often collected from drip pools in caves, these areas may not represent primary habitat and are likely operating as sink populations that are highly dependent on immigration from the epikarst (Pipan et al. 2010). *S. morrisoni* may rely more on cave streams and pools, based on known collection sites and its larger body size (e.g., Culver et al. 2010). The proximity to development and roads, intensity of agricultural practices, and levels of cave visitation also represent major threats to *S. morrisoni* populations. However, differences in local geology, including cave depth, size, and susceptibility to flooding, appeared to have a greater influence on judgments (Appendix S2). There was less certainty about microhabitat use of *S. morrisoni*, resulting in more variability between populations and generally wider distributions compared with the other species assessed. The fact that these potential ecological differences were reflected in judgments of population persistence suggests that the SHELF protocol can effectively provide decision makers with useful summaries of experts' understanding of both risk and uncertainty.

One way that structured expert elicitation may aid conservation decision-making is by highlighting situations where improved empirical data may have the greatest impact on perceived risk to persistence. For example, it has been suggested that the northern-most occurrence of *S. morrisoni* in Dyers Cave, West Virginia, may represent a distinct species (Holsinger 1978; Fong et al. 2007). Similarly, *S. morrisoni* individuals collected from Mountain Grove Saltpetre Cave in Virginia display morphological differences that warrant further genetic study (Appendix S2). Cryptic diversity is likely widespread in groundwater amphipods due to strong morphological convergence in subterranean habitats (e.g., Trontelj et al. 2009), which may have a greater impact on understanding of species-level risk for *S. morrisoni* due to greater variability among localities compared with *S. parvus*. In particular, the potential that the metapopulation near Dyers Cave could represent a single-site endemic species warrants further study given its wide credible interval for the probability of persistence (0.33−0.83).

Sampling of stygobionts is generally limited, and uncertainty in the fine-scale distributions of all 3 species affects interpretation of the EKE results. Although our analysis focused on known localities, long-term viability requires protecting unsampled karst regions to maintain connectivity and the potential for recolonization following localized stochastic events (Pipan et al. 2010). Christman et al. (2016) compiled over 11,000 records of cave species spanning the Appalachian region and found no records of these 3 *Stygobromus* spp. outside of the ranges displayed in Fig. 1. Data from European stygobionts suggest that species ranges of >200 km are extremely rare (Trontelj et al. 2009), and nearly half (44%) of U.S. species are known from a single county (Culver et al. 2000). This suggests that the known localities provide a reasonable estimate of the species' extent; however, the available data and approach used in the EKE do not capture potential changes in historical occurrence patterns throughout the range. Several methods exist for mapping relative differences in the intrinsic vulnerability of groundwater to contamination (e.g., Doerfliger et al. 1999). Although karst vulnerability mapping was not used in the present assessment because it does not incorporate species response to threats, this approach may prove useful for extrapolating expert judgments on relative risk to unsampled karst regions.

Results of an EKE depend critically on the experts involved, and questions inevitably arise regarding the accuracy of results. Other researchers have used calibration variables designed to test experts' statistical accuracy on quantities that will be known in the near future (Wittmann et al. 2015). However, this approach is rare in ecological applications because of the additional time required of the experts and the difficulty of identifying relevant test quantities (Hemming et al. 2020). Although quantities such as annual biomass of commercial fishes may provide useful seed questions for some management problems (Wittmann et al. 2015), it is difficult to identify questions that reasonably test an expert's ability to judge population or species persistence over several generations. The behavioral aggregation approach of SHELF does not require additional calibration quantities (Gosling 2018; O'Hagan 2019), a feature that may prove useful for many conservation applications. Results are considered accurate in the sense that the final probability distributions represent the experts' subjective beliefs and collective uncertainty in a quantitative way consistent with probability theory and the available evidence (O'Hagan 2019). Although results are unavoidably subjective, it is important to emphasize that decision makers should only turn to expert judgment after all empirical data have been exhausted (Burgman 2016) and that the primary role of EKE in this context is to help experts express their knowledge in a coherent framework that can directly support decision-making.

The same resource constraints that limit empirical data for conservation assessments may also affect the application of structured elicitation processes, which require substantial time and effort by both facilitators and experts. Although the benefits of in-person workshops are clear, their costs and logistics have led many to seek options for remote elicitation (Kuhnert et al. 2010; Hemming et al. 2018). For example, the use of web- or email-based surveys have allowed for elicitation processes on national and international scales that would be otherwise prohibitive (Donlan et al. 2010; McBride et al. 2012). Reduced levels of communication have been reported as a key drawback of remote EKE (McBride et al. 2012). However, we found video conferencing and online judgment submission forms were highly compatible with SHELF's behavioral aggregation techniques. Facilitation methods such as round-robin formats, directly calling on specific experts, and a slight modification requiring private RIO judgments as a starting point for the group judgment round ensured adequate discussion occurred and all views were captured. In addition, a manageable group size and history of collaboration among experts likely contributed to the overall success of the remote process. The greatest concern with applying SHELF remotely is connectivity because the group RIO judgments require all experts to be present and communicating in real time. Only 1 expert lost connectivity for 1 locality during the entire EKE (Fig. 2). Although this expert was provided an opportunity to comment on the final RIO distribution afterwards (Appendix S2), it is not possible to directly incorporate views post hoc. Eliciting experts' individual judgments to capture any divergence of opinion before allowing comment on the final RIO distribution could serve as a reasonable compromise, provided the number of experts experiencing connectivity problems is small.

S. parvus, *S. cooperi*, and *S. morrisoni* were assessed using a common framework because all 3 species share similar data availability, are highly specialized for subterranean habitats, face common threats to persistence, and are on similar ESA assessment time lines. Expert elicitation is time intensive, and the need to use volunteered expertise efficiently is underscored by the large backlog of ESA candidate species (USFWS 2016a). Conducting the elicitation as part of a multispecies assessment makes efficient use of experts' time, provides for consistent methods across species, and may promote enhanced understanding of the factors affecting persistence through discussion of contrasting ecological needs. Assessing the probability of persistence of a population requires the experts to consider interactions between potentially synergistic threats and a range of demographic processes. Although these are difficult judgments with unknowable values, our methods captured experts' uncertainty in a scientifically rigorous manner that can support decision-making in the absence of a data-deficient classification option and highlight research needs that could improve empirical understanding of the extinction risk for assessed species.

Acknowledgments

This work was supported by the U.S. Geological Survey (USGS) Ecosystems Mission Area and Mendenhall Research Fellowship Program (D.B.F.). The findings and conclusions in this article are those of the authors and USGS and do not necessarily represent the views of the U.S. Fish and Wildlife Service. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Supporting Information

Additional information is available online in the Supporting Information section at the end of the online article. The authors are solely responsible for the content and functionality of these materials. Queries (other than absence of the material) should be directed to the corresponding author.

Literature Cited

Abatzoglou JT. 2013. Development of gridded surface meteorological data for ecological applications and modelling. International Journal of Climatology **33:**121–131.

Adams-Hosking C, et al. 2016. Use of expert knowledge to elicit population trends for the koala (Phascolarctos cinereus). Diversity and Distributions **22:**249–262.

Bland LM, Collen B, Orme CDL, Bielby J. 2015. Predicting the conservation status of data-deficient species. Conservation Biology **29:**250–259.

Boyer DG, Pasquarell GC. 1995. Nitrate concentrations in Karst Springs in an extensively grazed area. Water Resources Bulletin **31:**729–736.

Burgman M. 2004. Expert frailties in conservation risk assessment and listing decisions. Pages 20–29 in Hutchings P, Lunney D, and Dickman C (editors). Threatened species legislation: is it just an act?. Royal Zoological Society of New South Wales, Mosman.

Burgman MA. 2016. Trusting judgement: how to get the best out of experts. Cambridge University Press, Cambridge, United Kingdom.

Christman MC, Doctor DH, Niemiller ML, Weary DJ, Young JA, Zigler KS, Culver DC. 2016. Predicting the occurrence of cave-inhabiting fauna based on features of the earth surface environment. PLoS ONE **11**. https://doi.org/10.1371/journal.pone.0160408.

Culver DC, Holsinger JR, Christman MC, Pipan T. 2010. Morphological differences among eyeless amphipods in the genus stygobromus dwelling in different subterranean habitats. Journal of Crustacean Biology **30:**68–74.

Culver DC, Master LL, Christman MC, Hobbs HH. 2000. Obligate cave fauna of the 48 contiguous United States. Conservation Biology **14:**386–401.

Dias LC, Morton A, Quigley J (editors). 2018. Elicitation: the science and art of structuring judgement. Springer International, Cham, Switzerland.

Doerfliger N, Jeannin PY, Zwahlen F. 1999. Water vulnerability assessment in karst environments: a new method of defining protection areas using a multi-attribute approach and GIS tools (EPIK method). Environmental Geology **39:**165–176.

Donlan CJ, Wingfield DK, Crowder LB, Wilcox C. 2010. Using expert opinion surveys to rank threats to endangered species: a case study with sea turtles. Conservation Biology **24:**1586–1595.

Drescher M, Perera AH, Johnson CJ, Buse LJ, Drew CA, Burgman MA. 2013. Toward rigorous use of expert knowledge in ecological research. Ecosphere **4**. https://doi.org/10.1890/Es12-00415.1.

EIA (Energy Information Administration). 2020. Appalachian Basin oil and gas fields. U.S. EIA, Washington, D.C. Available from https://www.eia.gov/maps/maps.htm (accessed April 2020).

EPA (Environmental Protection Agency). 2020. WATERS geospatial data downloads. U.S. EPA, Washington, D.C. Available from https://www.epa.gov/waterdata/waters-geospatial-data-downloads (accessed April 2020).

Fong DW, Culver DC, Hobbs HH, Pipan T. 2007. The invertebrate cave fauna of West Virginia, 2nd edition. West Virginia Speleological Survey Bulletin **13:**1–163.

Gosling JP. 2018. SHELF: the Sheffield elicitation framework. Pages61–94 in Dias LC, Quigley J, and Morton A (editors). Elicitation: the science and art of structuring judgement. Springer International, Cham, Switzerland.

Hemming V, Burgman MA, Hanea AM, McBride MF, Wintle BC. 2018. A practical guide to structured expert elicitation using the IDEA protocol. Methods in Ecology and Evolution **9:**169–180.

Hemming V, Hanea AM, Walshe T, Burgman MA. 2020. Weighting and aggregating expert ecological judgments. Ecological Applications **30**. https://doi.org/10.1002/eap.2075.

Holsinger JR. 1978. Systematics of the Subterranean Amphipod Genus Stygobromus (Crangonyctidae). Part II: Species of the Eastern United States. Smithsonian Contributions to Zoology **160:**1–152.

Holsinger JR, Culver DC, Hubbard DA, Orndorff WD, Hobson CS. 2013. The invertebrate cave fauna of Virginia. Banisteria **42:**9–56.

Kindsvater HK, Dulvy NK, Horswill C, J.-J. MJ, Mangel M, Matthiopoulos J. 2018. Overcoming the data crisis in biodiversity conservation. Trends in Ecology & Evolution **33:**676–688.

Kuhnert PM, Martin TG, Griffiths SP. 2010. A guide to eliciting and using expert knowledge in Bayesian ecological models. Ecology Letters **13:**900–914.

Lacher TE, Boitani L, da Fonseca GAB. 2012. The IUCN global assessments: partnerships, collaboration and data sharing for biodiversity science and policy. Conservation Letters **5:**327–333.

Lewis JJ. 2001. Conservation Assessment for Minute Cave Amphipod (*Stygobromus parvus*). U.S. Department of Agriculture, Forest Service, Eastern Region, Milwaukee, Wisconsin.

Loop CM, White WB. 2001. A conceptual model for DNAPL transport in karst ground water basins. Ground Water **39:**119–127.

Mammola S, et al. 2019. Scientists' warning on the conservation of subterranean ecosystems. Bioscience **69:**641–650.

Martin TG, Burgman MA, Fidler F, Kuhnert PM, Low-Choy S, Mcbride M, Mengersen K. 2012. Eliciting expert knowledge in conservation science. Conservation Biology **26:**29–38.

McBride MF, et al. 2012. Structured elicitation of expert judgments for threatened species assessment: a case study on a continental scale using email. Methods in Ecology and Evolution **3:**906–920.

NLCD (National Landcover Database). 2016. National Landcover Database. Multi-Resolution Land Characteristics Consortium. Available from https://www.mrlc.gov/data (accessed May 2020).

O'Hagan A. 2019. Expert knowledge elicitation: subjective but scientific. American Statistician **73:**69–81.

Oakley JE. 2019. SHELF: tools to support the Sheffield elicitation framework. R package version 1.6.0. Available from https://CRAN.R-project.org/package=SHELF (accessed February 2020).

Oakley JE, O'Hagan A. 2019. SHELF: the Sheffield elicitation framework. Version 4. School of Mathematics and Statistics, University of Sheffield, United Kingdom. Available from http://tonyohagan.co.uk/shelf (accessed February 2020).

Pasquarell GC, Boyer DG. 1996. Herbicides in karst groundwater in southeast West Virginia. Journal of Environmental Quality **25:**755–765.

Pipan T, Culver DC. 2007. Copepod distribution as an indicator of epikarst system connectivity. Hydrogeology Journal **15:**817–822.

Pipan T, Culver DC. 2013. Forty years of epikarst: what biology have we learned? International Journal of Speleology **42:**215–223.

Pipan T, Holt N, Culver DC. 2010. How to protect a diverse, poorly known, inaccessible fauna: identification and protection of source and sink habitats in the epikarst. Aquatic Conservation-Marine and Freshwater Ecosystems **20:**748–755.

Poulson TL, White WB. 1969. The cave environment. Science **165:**971–981.

Raiffa H. 1968. Decision analysis: introductory lectures on choice under uncertainty. Addison-Wesley, Reading, Massachusetts.

Sbordoni V, Alegrucci G, Cesaroni D. 2000. Population genetic structure. speciation and evolutionary rates in cave dwelling organisms. Pages 453–477 in Wilkens H, Culver DC, and Humphreys W (editors). Subterranean ecosystems. Elsevier, Oxford, United Kingdom.

Simon KS, Buikema AL. 1997. Effects of organic pollution on an Appalachian cave: changes in macroinvertebrate populations and food supplies. American Midland Naturalist **138:**387–401.

Smith DR, Allan NL, McGowan CP, Szymanski JA, Oetker SR, Bell HM. 2018. Development of a species status assessment process for decisions under the US Endangered Species Act. Journal of Fish and Wildlife Management **9:**302–320.

Sohl TL, et al. 2018. Conterminous United States Land Cover Projections - 1992 to 2100, Available from https://doi.org/10.5066/P95AK9HP (accessed May 2019).

Stokstad E. 2005. What's wrong with the Endangered Species Act? Science **309:**2150–2152.

Sutherland WJ. 2006. Predicting the ecological consequences of environmental change: a review of the methods. Journal of Applied Ecology **43:**599–616.

Trontelj P, Douady CJ, Fiser C, Gibert J, Goricki S, Lefebure T, Sket B, Zaksek V. 2009. A molecular test for cryptic diversity in ground water: how large are the ranges of macro-stygobionts? Freshwater Biology **54:**727–744.

USACE (U.S. Army Corps of Engineers). 2020. National inventory of dams. USACE, Washington, DC. Available from https://nid.sec.usace.army.mil/ords/f (accessed April 2020).

USFWS (U.S. Fish and Wildlife Service). 2016*a*. Methodology for prioritizing status reviews and accompanying 12-month findings on petitions for listing under the endangered species act. Federal Register **81:**49248–49255.

USFWS (U.S. Fish and Wildlife Service). 2016*b*. USFWS species status assessment framework: an integrated analytical framework for conservation. Version 3.4. U.S. Fish and Wildlife Service, Falls Church, Virginia.

Veni G. 1999. A geomorphological strategy for conducting environmental impact assessments in Karst areas. Geomorphology **31:**151–180.

Wittmann ME, Cooke RM, Rothlisberger JD, Rutherford ES, Zhang H, Mason DM, Lodge DM. 2015. Use of structured expert judgment to forecast invasions by bighead and silver carp in Lake Erie. Conservation Biology **29:**187–197.