ORIGINAL RESEARCH

Determining threatened species distributions in the face of limited data: Spatial conservation prioritization for the Chinese giant salamander (*Andrias davidianus*)

Shu Chen[1] | Andrew A. Cunningham[1] iD | Gang Wei[2] | Jian Yang[3] |
Zhiqiang Liang[4] | Jie Wang[5] | Minyao Wu[6] | Fang Yan[7] | Hanbin Xiao[8] |
Xavier A. Harrison[1] iD | Nathalie Pettorelli[1] | Samuel T. Turvey[1] iD

[1]Institute of Zoology, Zoological Society of London, London, UK

[2]Guiyang University, Guiyang, Guizhou, China

[3]Guangxi Teachers Education University, Nanning, Guangxi, China

[4]Fisheries Research Institute of Hunan Province, Changsha, Hunan, China

[5]Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, Sichuan, China

[6]Shaanxi Normal University, Xi'an, Shaanxi, China

[7]Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, Yunnan, China

[8]Yangtze River Fisheries Research Institute, Chinese Academy of Fisheries Science, Wuhan, Hubei, China

Correspondence
Andrew A. Cunningham and Samuel T. Turvey, Institute of Zoology, Zoological Society of London, London, UK.
Emails: a.cunningham@ioz.ac.uk;samuel.turvey@ioz.ac.uk

Abstract

The purpose of this study was to determine whether limited occurrence data for highly threatened species can provide useful spatial information to inform conservation. The study was conducted across central and southern China. We developed a habitat suitability model for the Critically Endangered Chinese giant salamander (*Andrias davidianus*) based on one biotic and three abiotic parameters from single-site locality records, which represent the only relevant environmental data available for this species. We then validated model quality by testing whether increased percentage of predicted suitable habitat at the county level correlated with independent data on giant salamander presence. We randomly selected 48 counties containing historical records which were distinct from, and independent of, the single-site records used to develop the model, and 47 additional counties containing >50% predicted suitable habitat. We interviewed 2,812 respondents near potential giant salamander habitat across these counties and tested for differences in respondent giant salamander reports between counties selected using each method. Our model predicts that suitable giant salamander habitat is found widely across central and southern China, with counties containing ≥50% predicted suitable habitat distributed in 13 provinces. Counties with historical records contain significantly more predicted suitable habitat than counties without historical records. There are no statistical differences in any patterns of respondent giant salamander reports in surveyed counties selected from our model compared with the areas of known historical giant salamander distribution. A Chinese giant salamander habitat suitability model with strong predictive power can be derived from the restricted range of environmental variables associated with limited available presence-only occurrence records, constituting a cost-effective strategy to guide spatial allocation of conservation planning. Few reported sightings were recent, however, with most being over 20 years old, so that identification of areas of suitable habitat does not necessarily indicate continued survival of the species at these locations.

1 | INTRODUCTION

Effective conservation management of threatened species requires a robust, evidence-based understanding of key population parameters such as geographic distribution and habitat requirements (Segan, Bottrill, Baxter, & Possingham, 2011; Stewart, Coles, & Pullin, 2005; Sutherland, Pullin, Dolman, & Knight, 2004). However, robust data are often unavailable for extremely rare species, which are most urgently in need of management action, as the very rarity of these species can make them difficult to study or even detect using standard field survey methods (Thompson, 2013). Assessing the extent to which limited available data can provide useful insights to inform management of cryptic or poorly known threatened species therefore represents an important conservation research goal.

Habitat suitability models are a group of mechanistic statistical models widely used in ecology, which relate the frequency of species occurrences to sets of environmental variables in order to generate predictions of locations where species are expected to occur (Franklin, 2009). Conservation effectiveness remains hindered by severe funding and other social resource limitations, especially when trying to support conservation interventions for species that occur across large geographic areas (Isaac, Redding, Meredith, & Safi, 2012; Marris, 2007), and so such models can potentially constitute an important cost-effective tool to optimize spatial prioritization of conservation activities. Considerable attention has been paid to factors that might affect the accuracy of occurrence probability and range prediction from habitat suitability models, including sample size, the use of presence–absence data versus presence-only data, data quality and representativeness (e.g., associated with variation in habitat use with life stage), and randomness of sampling (Aranda & Lobo, 2011; Feeley & Silman, 2011; Fei & Yu, 2016; Fithian, Elith, Hastie, & Keith, 2015; Hastie & Fithian, 2013; Lütolf, Kienast, & Guisan, 2006; Zajac, Stith, Bowling, Langtimm, & Swain, 2015). In practice, such models may be forced to rely on presence-only datasets comprising occurrence records that have been collected opportunistically rather than systematically, are of insufficient spatial resolution, and/or include bias in spatial search effort. For some threatened species, recent locality data might even be deliberately kept secret to reduce poaching risk (Meijaard & Nijman, 2014; Yang & Chan, 2016). As such incomplete and biased data often constitute the only information available for trying to determine potential geographic distributions for highly threatened species therefore, it is necessary to attempt to evaluate whether such data can provide a meaningful biogeographic signal.

The Chinese giant salamander (*Andrias davidianus*) (Figure 1), the world's largest amphibian, is a cryptobranchid salamander endemic to China, where it has been historically recorded from fast-flowing tributaries of the Yellow, Yangtze, and Pearl river systems (Fei, Hu, Ye,

FIGURE 1 Chinese giant salamander (*Andrias davidianus*)

& Huang, 2006; Wang et al., 2004). However, the species is severely threatened both by habitat loss and by unsustainable overexploitation of wild individuals, particularly for the recently developed domestic luxury food market, and the rapidly growing giant salamander farming industry might further threaten its survival in the wild (Cunningham et al., 2016; Huang, 1982; Wang et al., 2004); the species may therefore already be extirpated from areas of suitable remaining habitat. It is listed as Critically Endangered by the IUCN (2016), and it is a top priority for international conservation using prioritization metrics that incorporate evolutionary history, as it is one of only three extant species in the Cryptobranchidae, a lineage that diverged from other amphibians during the Jurassic (Isaac, Redding, Meredith, & Safi, 2012). Due to severe declines observed or inferred across its range, it is now extremely difficult to detect using standard ecological survey methods (Pierson, Yan, Wang, & Papenfuss, 2014; Tapley et al., 2015), and large-scale systematic surveys have been identified as a priority activity to inform spatial conservation planning for the species (Meredith, 2011).

Some recent field data on local giant salamander presence or absence are available at the single-site level (Fellowes, Chan, Lau, Ng, & Siu, 2003; Pan et al., 2016; Pierson et al., 2014; Tapley et al., 2015; Wang et al., 2004), but such data are often too limited either to evaluate the continued occurrence of giant salamanders more widely across China, or to extrapolate likely habitat suitability across the species' former range for survey planning. In addition, some recent giant salamander records might represent introduced individuals from government release programs, which may occur in unsuitable habitat outside the species' natural geographic range and/or habitat, rather than representing surviving native populations (Cunningham et al., 2016; Zhang, Dearing, Tong, & Hughes, 2016; Zhang, Jiang, et al., 2016). A small set of single-site giant salamander records is available with a series of habitat parameters with local presence of the species (Table 1). Wider-scale

TABLE 1 Available data on environmental variables associated with presence of Chinese giant salamander (*Andrias davidianus*), from four locations in four Chinese provinces

Province	County	Location	Mean annual temperature (°C)	Mean annual rainfall (mm)	Elevation (m)	References
Guizhou	Guiding	Yanxia	13.9	—	1,100	Li, Yu, and Ma (2009)
Henan	Lushi	Lushi County Nature Reserve	12.7	732.6	300–800	Zheng (2006)
Hunan	Dayong	Zhangjiajie	16	—	190–500	Luo, Liu, Liu, Luo, and Tang (2007)
Hunan	Dayong	Zhangjiajie	13.4–16.8	1,500	250–650	Luo, Liu, and Zhang (2009)
Hunan	Dayong	Zhangjiajie	—	—	276–470	Luo (2009)
Hunan	Dayong	Zhangjiajie (Golden Whip Stream)	12.8	1,200–1,600	491	Luo and Kang (2009)
Hunan	Sangzhi	Zhangjiajie	16	1,400	200–650	Luo, Zhang, Liu, Chen, and Gan (2009)
Shanxi	Yuanqu	Lishan National Nature Reserve	14	780	490–1,330	Guo (2011)

historical locality data, largely dating from the 1980s and 1990s shortly before the giant salamander's major population decline due to overexploitation (Cunningham et al., 2016; Wang et al., 2004) have been compiled from local gazetteer records (Fei et al., 2006). These data do not report specific localities, but are instead recorded at a relatively low spatial resolution (e.g., at the level of local municipality, county, or mountain range), with records representing relatively large areas across which key environmental parameters may show considerable variation. Analysis of these ecological and occurrence data is therefore required to assess whether they contain enough information to identify potentially suitable giant salamander habitats that could be targeted by novel field surveys.

Thus, two limited and independent, but complementary, sources of data are available to assess the geographic distribution and habitat requirements of the Chinese giant salamander: single-site locality records associated with specific environmental variables, and wide-scale locality records from historical gazetteers with no associated environmental data. In order to provide an improved baseline for prioritizing field conservation activities for the Chinese giant salamander, we use the environmental data associated with single-site locality records together with open-source ecological data to develop the first predictive habitat suitability model for this highly threatened species. We then assess the likely accuracy of this model both using the historical gazetteer record as an independent comparative data source and by ground-truthing model predictions with data from a new large-scale questionnaire survey conducted across central and southern China, to determine the extent to which incomplete occurrence data for highly threatened species can provide useful information on their geographic distribution.

2 | METHODS

2.1 | Habitat suitability model

Eight studies documenting environmental data associated with Chinese giant salamander locality records are available in the

Chinese literature, from four provinces (Hunan, *n* = 5; Guizhou, *n* = 1; Henan, *n* = 1; Shanxi, *n* = 1) (Table 1). In addition to specific water quality or microhabitat parameters for which country-level spatial mapping data are unavailable (e.g., flow rate, dissolved oxygen, nitrogen, water hardness, substrate, bank gradient), these locality records provide site-specific information on elevation, mean annual temperature, and mean annual precipitation. We included these three predictor variables in our habitat suitability model, using the following ranges from the literature to define suitable giant salamander habitat: 190–1,330 m a.s.l. elevation, 12.7–16.8°C mean annual temperature, ≥732.6 mm mean annual precipitation (Table 1). We also included vegetation cover as a fourth predictor variable, as available evidence suggests that Chinese giant salamanders require extensive bankside vegetation and do not occur in extensively human-modified landscapes such as cropland, bare ground, or urban environments, unlike Japanese giant salamanders (*Andrias japonicus*) (Browne et al., 2014).

We carried out spatial analyses using ArcGIS 10.1 (ESRI, 2014). We downloaded maps of Chinese administrative areas, elevation, and land use from DIVA-GIS (available at http://www.diva-gis.org/Data); we extracted the three vegetation categories of tree cover, shrub cover, and tree cover/other natural vegetation mosaic from the land-use map and grouped them into a single forest cover category for analysis. This coarse scale was used due to a lack of more detailed habitat data for CGS, other than an association with vegetation cover (Table 1). We downloaded mean monthly temperature and precipitation data from the WorldClim global climate database (available at http://www.worldclim.org/current) at a resolution of 30 arc seconds and averaged these data to generate measures of mean annual temperature and precipitation. We produced a habitat suitability model for the Chinese giant salamander based on predicted habitat suitability for China using the Raster Calculator in the Spatial Analyst Tools in ArcMap 10.1 by intersecting the selected ranges of the four environmental predictor variables. We calculated the percentage of suitable giant salamander

habitat in each Chinese county (n = 2,852) using the Zonal Statistics in the Spatial Analyst Tools in ArcMap 10.1.

All Chinese counties with historical giant salamander gazetteer records listed in Fei et al. (2006) were given the value 1 (presence). Some historical gazetteer localities were recorded at the coarser resolution of municipality or mountain area; in these cases, all counties belonging to the relevant administrative or geographic areas were categorized as 1. All other counties lacking historical gazetteer records were given the value 0 (absence). We used these presence/absence values to produce a historical distribution map for the Chinese giant salamander using ArcMap 10.1.

We used a randomization approach to test the significance of the observed relationship between historical giant salamander records and the proportion of predicted suitable salamander habitat. We did not logit-transform the proportion data as they contained a high frequency of zeroes. At each iteration, for a total of 10,000 iterations, we randomly resampled the binary indicator variable of historical giant salamander presence and fitted a linear model specifying the resampled habitat variable as a predictor of proportion of suitable habitat. We stored the F statistic from each of these models to derive a null distribution of effects. We calculated a two-tailed p value as the proportion of iterations yielding F statistics greater than the true F calculated from a linear model fitted to the raw data.

2.2 | Study area

We tested model predictions as part of a large-scale survey to assess the current status of giant salamanders across China (see Cunningham & Chen, 2016 for details of all fieldwork activities). We randomly selected 50 counties containing historical giant salamander records and 50 further counties from a sample of all counties that contained >50% predicted suitable giant salamander habitat and which did not have a historical record (using the RANDBETWEEN function in Microsoft Excel v. 14.4.8). We were able to conduct interview surveys in 95 of the selected 100 target counties (48 counties containing historical records and 47 counties containing >50% predicted suitable habitat). These sites represent 15 Chinese provinces or equivalent administrative units (Anhui = 6, Chongqing = 3, Fujian = 1, Gansu = 4, Guangdong = 1, Guangxi = 10, Guizhou = 32, Henan = 3, Hubei = 4, Hunan = 12, Jiangxi = 2, Shaanxi = 3, Sichuan = 9, Yunnan = 1, Zhejiang = 4). As the cost and effort of surveying sites were high and associated logistics were challenging, it was impractical to survey additional sites where we did not think giant salamanders were likely to occur.

2.3 | Questionnaire survey

The use of community-based interview surveys has recently been shown to constitute an effective survey method for detecting Chinese giant salamanders (Pan et al., 2016), and so fieldwork to investigate giant salamander status was carried out by conducting interviews in villages within 1 km of a 1-km target stretch of potentially suitable giant salamander habitat in each selected county (fast-flowing rocky tributaries within or adjacent to forest; Browne et al., 2014; Fei et al., 2006) as identified by county-level fisheries and/or forestry bureau

officials. We aimed to conduct 30 interviews per county, either as 10 interviews each in three villages or more interviews in fewer villages, depending on the number of communities available for sampling within the survey region; a minimum target number of 10 interviews per village will likely capture most or all existing variation in relevant experiences for many respondent groups (Guest, 2006). Respondent selection criteria/methods and interview protocols are given in Pan et al. (2016). Project design was approved by the Zoological Society of London's Ethics Committee (ref. WLE569).

We used a standard questionnaire for all interviews, which took c. 20–30 min to complete, and which contained a series of descriptive, contrast, and structured questions (Appendix S1). Following an initial study around three national nature reserves in Guizhou in 2013 to trial interview methods (Pan et al., 2016), we conducted interviews between May 2013 and June 2016. Interviews were conducted and recorded in Chinese by Chinese field teams led by the authors and who received training in standardized interview techniques before fieldwork commenced. As part of a wider series of questions, we asked respondents whether they knew what giant salamanders were, to describe their appearance and to identify, without any prompting, the giant salamander from illustrations of a range of salamander species found in China, taken from Fei, Hu, Ye, & Huang (2006). If respondents could correctly identify and describe the Chinese giant salamander, we asked them whether they had seen the species, and if so how recently. Respondents reported sighting records using a variety of different methods for describing the timing of past events, and we converted alternative formats to direct calendar years for analysis using the approach described by Turvey et al. (2016). As interviews were conducted across a period of 3 years due to the logistical demands of fieldwork, we then converted sighting records to number of years before the date on which interviews were conducted, to allow comparison between sites.

We investigated differences in the pattern of respondent reports of giant salamanders between surveyed counties containing historical giant salamander records and surveyed counties selected from our habitat suitability model using R version 3.4.1 (R Core Team, 2017). Differences in the overall proportion of each county type (historical record vs. habitat suitability) from which we obtained giant salamander sighting reports were investigated using chi-squared tests. We investigated other potential effects of county selection method (0 for selection using habitat suitability model, or 1 for selection based on the existence of historical giant salamander records) on respondent reports of giant salamanders across the surveyed counties using mixed effects models, as we wanted to include the Chinese province in which each county is located as a random effect on the model intercept to control for potential variation in survey effort (as different provinces were investigated by different Chinese survey teams). We used our observed data structure (replication of records within and among grouping levels of random effects) to investigate the probability of detecting a true difference between our two datasets (historical records vs habitat suitability) given our baseline probability in the reference category of model history, and a range of true differences in reporting probabilities (Appendix S2).

We investigated differences in proportions of respondents reporting giant salamander sightings per county by fitting a binomial mixed effects model using the "glmer" function in the R package "lme4" (Bates, Maechler, Bolker, & Walker, 2015), with county selection method as a fixed effect and province as a random intercept. We calculated overdispersion for this model following guidelines in Harrison (2014). We detected strong overdispersion, with model residuals demonstrating approximately 10 times the expected variance, so we fitted an observation-level random effect (OLRE) to control for overdispersion following Harrison (2015). The OLRE model significantly improved model fit relative to the overdispersed model ($\chi^2_1 = 759.86$, $p < .001$). We investigated differences in time (years) since the most recently reported giant salamander sighting per county by fitting a Poisson mixed effects model using the "glmer" function in the R package "lme4" (Bates et al., 2015), with county selection method as a fixed effect and province as a random intercept. As with the proportion models, this model exhibited strong overdispersion (variance inflation factor = 10.42), and so we refitted the model with a negative binomial error structure using the glmmADMB package (Fournier et al., 2012; Skaug, Fournier, Bolker, Magnusson, & Nielsen, 2016). The negative binomial model resulted in a significant improvement in fit over the Poisson model ($\chi^2_1 = 375.80$, $p < .001$). Finally, we investigated differences in mean time (years) since all respondents reported giant salamander sightings per county. As this response contained a mix of zeroes and noninteger values, we fitted a mixed effects model with a Tweedie error structure using the "cpglmm" function in the R package "cplm" (Zhang, 2013). Unlike Poisson count models, Tweedie models automatically model overdispersion in the data by estimating a dispersion parameter. As for the previous models, the Tweedie model contained county selection method as a fixed effect and a random intercept for province. We derived p values from all mixed effects models by comparison of nested models using a likelihood ratio test.

3 | RESULTS

Historical giant salamander gazetteer records are documented from 145 counties (5.1% of the total number of Chinese counties) across 18 Chinese provinces or equivalent administrative areas (Anhui, Chongqing, Fujian, Gansu, Guangdong, Guangxi, Guizhou, Henan, Hubei, Hunan, Jiangsu, Jiangxi, Qinghai, Shaanxi, Shanxi, Sichuan, Yunnan, Zhejiang) (Figure 2). Guizhou has the highest number of gazetteer records (30 counties), followed by Gansu (13 counties), Guangxi (12 counties), Henan (12 counties), Hubei (11 counties), Sichuan (11 counties), and Hunan (9 counties).

Our habitat suitability model predicts that suitable habitat for the Chinese giant salamander is found widely across central and southern China (Figure 3). The percentage of predicted suitable giant salamander habitat present in different Chinese counties across the entire country ranges from 0% to 95.7%. Counties with ≥50% predicted suitable habitat ($n = 156$, representing 5.5% of the total number of Chinese counties) are distributed in 13 provinces or equivalent administrative areas (Anhui, Chongqing, Fujian, Guangxi, Guizhou, Henan, Hubei, Hunan, Jiangxi, Shaanxi, Sichuan, Yunnan, Zhejiang). Almost

FIGURE 2 Distribution (red) of counties containing historical gazetteer records of Chinese giant salamander (*Andrias davidianus*). Data from Fei et al. (2006)

half of these counties are in Guizhou (n = 44, including 22 counties with ≥80% suitable habitat) and Hunan (n = 32, including 7 counties with ≥80% suitable habitat); other counties with ≥50% suitable habitat are mainly in Zhejiang (n = 25), Hubei (n = 14), Anhui (n = 12), and Sichuan (n = 8).

Predicted suitable habitats are present in 116 of the 145 counties (80.0%) that contain historical giant salamander gazetteer records. The percentage of predicted suitable habitat in these 116 counties ranges from <1% to 95.5%, with 51 containing ≥50% predicted suitable habitat; these 51 high-suitability counties are mainly distributed in Guizhou (n = 23), Zhejiang (n = 7), and Hunan (n = 6). Counties with historical giant salamander gazetteer records contain a significantly higher percentage of predicted suitable giant salamander habitat (34.6%) compared to counties without historical records (6.1%) (permutation test p_{RAND} < .001).

We interviewed 2,812 respondents in 95 counties (mean age = 47.11, age range = 15–89, SD = 14.56; male = 69.6%, female = 30.4%; mean number of interviews/county = 29.6, range = 11–36, SD = 2.57) (Figure 4). In total, 1,299 respondents (46.2%) reported having seen wild giant salamanders, with 1,146 respondents (40.8%) providing a last-sighting date (Table S1). There were no statistical differences between surveyed counties containing historical giant salamander records and surveyed counties selected from our habitat suitability model, either in the overall proportion of counties from which we obtained giant salamander sighting reports (historical: 43/48; model: 40/47; χ^2 = 0.12, df = 1, p = .73), in the proportion of respondents who had seen giant salamanders (historical: mean = 0.47; model: mean = 0.45; χ_1^2 = 2.37, p = .12), in time (years ago) since the most recent reported giant salamander last-sighting date/county (historical: mean = 5.37 years ago; model: mean = 8.03 years ago; χ_1^2 = 1.62, p = .20), or in time (years ago) since mean giant salamander last-sighting date/county (historical: mean = 20.59 years ago; model: mean = 20.72 years ago; χ_1^2 = 0.002, p = .96). Our analysis of statistical power revealed that for our data structure, we would have >80% power to detect a difference of 10% or more in reporting probabilities (Appendix S2).

4 | DISCUSSION

The Chinese giant salamander is one of a large number of highly threatened but poorly known species for which only relatively limited and geographically unevenly sampled, presence-only ecological data are available to inform conservation assessment and management. However, we demonstrate that the restricted available ecological knowledge-base for the Chinese giant salamander is still sufficient to develop a habitat suitability model that shows close spatial congruence with independently derived historical distribution data for the species, and predicts occurrence of the species at geographical localities where local observers report statistically similar levels and patterns of giant salamander encounters compared to areas of known giant salamander distribution in China. These two independent approaches for assessing and validating the quality of our habitat
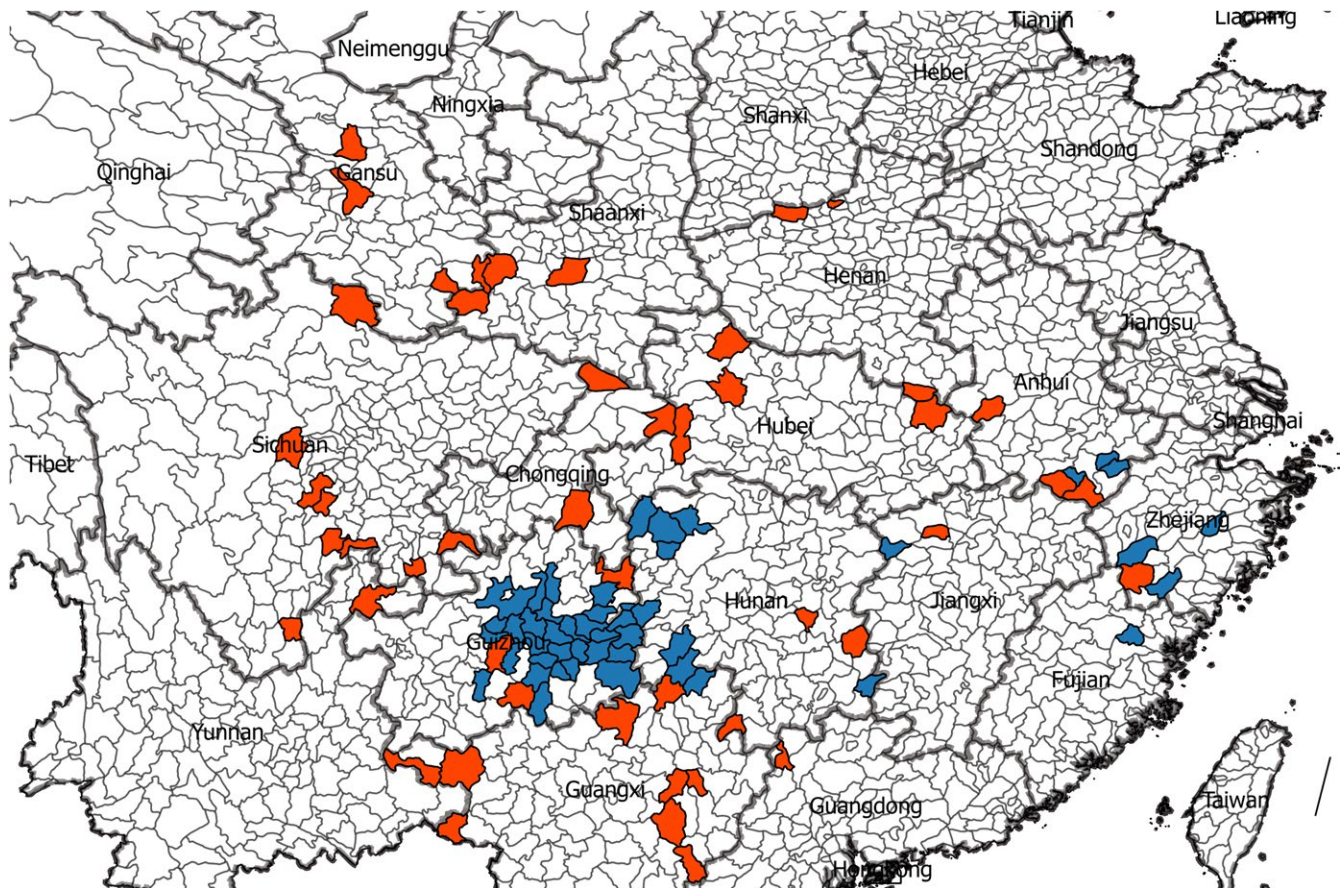
suitability model demonstrate that even data for a limited series of environmental parameters can provide a robust baseline on species ecology that can be used to understand likely geographic distributions and to guide spatial allocation of conservation resources and planning.

Field validation can sometimes reveal poor performance of habitat suitability models, even when numerous parameters are used to populate the model (e.g., Anderson et al., 2016). Our predictive model was only able to include data for four environmental parameters associated with giant salamander presence. These comprised three abiotic parameters (elevation, mean annual temperature, mean annual precipitation) and one biotic parameter (vegetation cover), all of which have previously been demonstrated to constitute important predictors of distribution for amphibian species (Buckley & Jetz, 2007; Chen, 2013). Other environmental parameters, such as fish presence, water depth, and speed of water flow, also have been found to be predictors of the distribution of some amphibian species (Manenti & Pennati, 2016). Unfortunately, information about how these (or other) parameters affect the distribution of CGS is unknown, so they could not be included in our predictive model. Attempts to develop habitat suitability models for other poorly known species with limited associated ecological data should assess which environmental factors are known to limit the distribution of better-studied related species, as exclusion of such parameters would be expected to reduce model performance.

Although Chinese counties with historical giant salamander records contain a significantly higher percentage of predicted suitable giant salamander habitat compared to counties without historical records, the congruence between counties identified by our predictive habitat suitability model and by our descriptive historical species distribution model is not complete. In addition to the possibility for error introduced by using only a small number of environmental parameters, and/or a potentially incomplete understanding of variation in giant salamander environmental tolerance associated with these parameters, this incomplete spatial agreement in inferred giant salamander distribution across China might be caused by several additional factors. Counties selected on the basis of habitat suitability might lack existing salamander records due to incomplete past survey effort across the large historical geographic distribution of this species, with many local populations potentially scientifically undetected or unreported in areas of suitable habitat within their range (Edwards, Cutler, Zimmermann, Geiser, & Alegria, 2005; Engler, Guisan, & Rechsteiner, 2004; Guisan et al., 2006). Indeed, our large-scale questionnaire survey provides indirect evidence, based on local reports, of recent or past giant salamander occurrence in 42 Chinese counties for which no historical records are known to exist, but for which likely giant salamander occurrence was predicted by our model. Habitat suitability models have successfully detected previously unknown populations of many other rare, cryptic, or otherwise poorly studied species (Cleve, Perrine, Holzman, & Hines, 2011; Ferreira de Siqueira, Durigan, de Marco Júnior, & Peterson, 2009; Menon, Choudhury, Khan, & Peterson, 2010; Raxworthy et al., 2003; Rebelo & Jones, 2010).

Counties with historical giant salamander records might not have been identified on the basis of habitat suitability for a variety of

FIGURE 3 Distribution (red) of suitable Chinese giant salamander (*Andrias davidianus*) habitat across China according to the output of our habitat suitability model

reasons. As some historical gazetteer localities were only recorded at the coarse resolution of municipality or mountain area, it is possible that some counties included within these broader administrative or geographic areas, which were interpreted as having contained giant salamanders in our historical dataset, might not actually contain suitable environmental conditions for salamanders, and might never have been home to wild populations. In addition, China has experienced extensive habitat loss in recent decades, and national afforestation and reforestation statistics mask the ongoing degradation of native forest biodiversity, including within protected areas (Hua et al., 2016; Zhang & Song, 2006; Zhang et al., 2010). It is therefore likely that suitable vegetation cover has been lost from some historical giant salamander localities, leading to their exclusion from our predictive habitat suitability model based on recent DIVA-GIS land-use data.

Furthermore, although giant salamander populations across China are currently interpreted as conspecific, considerable genetic variation and phylogeographic structuring have been detected between populations occupying different river drainages (Murphy, Fu, Upton, De Lama, & Zhao, 2000). Environmental data associated with giant salamander presence are available from only four provinces, with nearly all available data from Hunan (Table 1), and it is possible that this geographically restricted baseline fails to capture true levels of variation in environmental tolerances shown by different giant salamander populations across China. Indeed, our model notably fails to predict the potential occurrence of giant salamanders on the high-elevation

Qinghai–Tibet Plateau. A giant salamander specimen was reportedly collected in 1966 from the headwaters of the Yangtze River in Qumalai County, Qinghai Province (Figure 2), potentially representing a disjunct, isolated salamander population occurring at an elevation >2,000 m higher than any other known population (Chen, 2011; Pierson et al., 2014). The existence of this population has not been confirmed (Pierson et al., 2014), and we were unable to include data for this record in our predictive model due to uncertainty over its exact provenance or associated environmental conditions. If this constitutes a true giant salamander population, it is likely to be genetically and ecophenotypically distinct with different patterns of environmental tolerance to lower-elevation giant salamander populations and could even represent a cryptic species.

Despite this minor variation in historical versus predicted salamander distribution across China, the general accuracy of our habitat suitability model is supported by the close statistical congruence shown by the pattern and timing of giant salamander reports made by local respondents across surveyed counties irrespective of which method was used for county selection, indicating that regions identified using our predictive model show a similar signal of giant salamander detectability based on community-based interviews as regions where giant salamanders are reported to have occurred in the past. Interview data collected from untrained local respondents do not represent direct observations of a target species made by scientific experts and therefore include the potential for both error and bias when inferring

FIGURE 4 Distribution across China of surveyed counties containing historical giant salamander records (red) and surveyed counties selected from our habitat suitability model (blue)

species presence (McKelvey, Aubry, & Schwartz, 2008). However, we consider it highly unlikely that this statistical congruence represents an artefactual "false-positive" signal, as the Chinese giant salamander has cultural and economic importance in China (Cunningham et al., 2016; Pan et al., 2016) and our interview design aimed to minimize the potential for inaccuracy by requiring respondents to identify and describe the species correctly. Other community-based interview surveys conducted in China have shown that patterns of local ecological knowledge on the local status of charismatic freshwater vertebrates and other rarely encountered species match independently derived scientific field data on spatiotemporal population trends for these taxa (Turvey et al., 2013, 2016).

We note, however, that respondent experience of past giant salamander sightings, supporting our prediction of local habitat suitability, does not necessarily indicate continued survival of the species across the survey region, as very few reported sightings had been made within the past decade, and most were over 20 years old. While our predictive habitat suitability model is therefore a robust indicator of former salamander occurrence, intensive overexploitation of giant salamander populations has recently occurred across China (Cunningham et al., 2016; Wang et al., 2004). Our large dataset of respondent reports of past giant salamander sightings made during recent decades therefore cannot be used to confirm the continued occurrence of the species anywhere

across its range. Giant salamanders are vulnerable both to overexploitation and to habitat destruction through loss of riparian vegetation cover (from agricultural conversion and urbanization) and aquatic habitats (from water development projects and pollution), which have modified Chinese natural landscapes dramatically and present substantial challenges for future conservation of giant salamanders and many other species (Zhang, Dearing, et al., 2016; Zhang, Jiang, et al., 2016). However, the absence of recent giant salamander reports from remaining areas of suitable habitat as revealed by this study suggests that overexploitation is likely to be a more serious threat to the species. Any surviving giant salamander populations across our study area are clearly at high risk of continued exploitation; however, reporting the outputs of our habitat suitability model for the species at the broad country-wide scale presented here is unlikely to pose an additional threat.

Our assessment of the information content associated with ecological data available for the Chinese giant salamander reveals that even a restricted range of environmental correlates derived from a limited sample of presence-only occurrence records can, at least in some cases, be used to develop robust models with strong predictive power. In the case of the Chinese giant salamander, this improved understanding of the likely distribution of suitable habitats can be used further to investigate continued survival of local populations in high-suitability sites, especially at sites where local respondents have reported more

recent giant salamander sightings, to assess habitat suitability within existing protected areas that have already been established for the species, and to inform site selection for other conservation activities such as reintroduction and restocking (Cunningham et al., 2016; Zhang, Dearing, et al., 2016; Zhang, Jiang, et al., 2016). A larger series of static and dynamic environmental variables would undoubtedly refine our habitat suitability model, and confirmation of continued salamander existence across areas of high habitat suitability requires additional direct field investigation and systematic collection of presence–absence data. However, available independent spatial and survey data indicate that even the simple model we have developed statistically matches independent available data and accurately describes the species' recent geographic distribution. Our study therefore supports the potential applicability of similarly limited occurrence data for setting cost-effective yet meaningful conservation baselines for other poorly known threatened species, and for evaluating potential responses to future environmental and climatic change (e.g., Duan, Kong, Huang, Varela, & Ji, 2016). The modern conservation toolkit will have to draw upon different complementary and often limited, incomplete, or biased types of data in order to prevent future extinctions of highly threatened species in China and elsewhere. Realistically, we have no choice but to utilize whatever information is available on such species, and to continue to develop approaches to critically assess the extent to which imperfect data are useful and can be used to inform conservation planning (Hirzel, Le Lay, Helfer, Randin, & Guisan, 2006; Zajac et al., 2015).

AUTHOR CONTRIBUTIONS

A.A.C., S.T.T., and S.C. conceived the ideas; A.A.C. obtained the funding; S.C., G.W., J.Y., Z.L., J.W., M.W., F.Y., and H.X. collected the data; S.C., S.T.T., X.H., and N.P. analyzed the data; and S.T.T. and S.C. led the writing; all authors contributed to the writing of the manuscript.

CONFLICT OF INTEREST

None declared.

ORCID

Andrew A. Cunningham http://orcid.org/0000-0002-3543-6504

Xavier A. Harrison http://orcid.org/0000-0002-2004-3601

Samuel T. Turvey http://orcid.org/0000-0002-3717-4800

REFERENCES

Anderson, O. F., Guinotte, J. M., Rowden, A. A., Clark, M. R., Mormede, S., Davies, A. J., & Bowden, D. A. (2016). Field validation of habitat suitability models for vulnerable marine ecosystems in the South Pacific Ocean: Implications for the use of broad-scale models in fisheries management. *Ocean & Coastal Management*, *120*, 110–126. https://doi.org/10.1016/j.ocecoaman.2015.11.025

Aranda, S. C., & Lobo, J. M. (2011). How well does presence-only-based species distribution modelling predict assemblage diversity? A case study of the Tenerife flora. *Ecography*, *34*, 31–38. https://doi.org/10.1111/j.1600-0587.2010.06134.x

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.

Browne, R. K., Li, H., Wang, Z., Okada, S., Hime, P., McMillan, A., … Briggler, J. T. (2014). The giant salamanders (Cryptobranchidae): Part B. Biogeography, ecology and reproduction. *Amphibian and Reptile Conservation*, *5*, 30–50.

Buckley, L. B., & Jetz, W. (2007). Environmental and historical constraints on global patterns of amphibian richness. *Proceedings of the Royal Society B*, *274*, 1167–1173. https://doi.org/10.1098/rspb.2006.0436

Chen, X. (2011). Amphibia and reptilia. In D. Li (Ed.), *Economic fauna of Qinghai* (pp. 173–227). Xining, China: Northwest Plateau Institute of Biology.

Chen, Y. (2013). Habitat suitability modeling of amphibian species in southern and central China: Environmental correlates and potential richness mapping. *Science China: Life Sciences*, *56*, 476–484. https://doi.org/10.1007/s11427-013-4475-3

Cleve, C., Perrine, J., Holzman, B., & Hines, E. (2011). Addressing biased occurrence data in predicting potential Sierra Nevada red fox habitat for survey prioritization. *Endangered Species Research*, *14*, 179–191. https://doi.org/10.3354/esr00348

Cunningham, A. A., & Chen, S. (2016) *19-003: A sustainable future for Chinese salamanders*. Final report to Darwin Initiative. Retrieved from http://www.darwininitiative.org.uk/documents/19003/23771/19-003%20FR%20-%20edited.pdf

Cunningham, A. A., Turvey, S. T., Zhou, F., Meredith, H. M. R., Wei, G., Liu, X., … Wu, M. (2016). Development of the Chinese giant salamander (*Andrias davidianus*) farming industry in Shaanxi Province, China: Conservation threats and opportunities. *Oryx*, *50*, 265–273. https://doi.org/10.1017/S0030605314000842

Duan, R., Kong, X., Huang, M., Varela, S., & Ji, X. (2016). The potential effects of climate change on amphibian distribution, range fragmentation and turnover in China. *PeerJ*, *4*, e2185. https://doi.org/10.7717/peerj.2185

Edwards, T. C., Cutler, D. R., Zimmermann, N. E., Geiser, L., & Alegria, J. (2005). Model-based stratifications for enhancing the detection of rare ecological events. *Ecology*, *86*, 1081–1090. https://doi.org/10.1890/04-0608

Engler, R., Guisan, A., & Rechsteiner, L. (2004). An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *Journal of Applied Ecology*, *41*, 263–274. https://doi.org/10.1111/j.0021-8901.2004.00881.x

ESRI (2014) *ArcMap, version 10.1*. Redlands, CA: Environmental Systems Research Institute.

Feeley, K. J., & Silman, M. R. (2011). Keep collecting: Accurate species distribution modelling requires more collections than previously thought. *Diversity and Distributions*, *17*, 1132–1140. https://doi.org/10.1111/j.1472-4642.2011.00813.x

Fei, L. (1999). *Atlas of amphibians of China*. Zhengzhou, Henan, China: Henan Science and Technology Press.

Fei, L., Hu, S., Ye, S., & Huang, Y. (2006). *Fauna Sinica (Amphibia I)*. Beijing, China: Science Press.

Fei, S., & Yu, F. (2016). Quality of presence data determines species distribution model performance: A novel index to evaluate data

quality. *Landscape Ecology*, *31*, 31–42. https://doi.org/10.1007/s10980-015-0272-7

Fellowes, J. R., Chan, B. P. L., Lau, M. W. N., Ng, S. C., & Siu, G. L. P. (2003). *Report of rapid biodiversity assessments at Cenwanglaoshan Nature Reserve, northwest Guangxi, China, 1999 and 2002*. South China Forest Biodiversity Survey Report Series 27. Hong Kong, China: Kadoorie Farm & Botanic Garden.

Ferreira de Sequeira, M., Durigan, G., de Marco Júnior, P., & Peterson, A. T. (2009). Something from nothing: Using landscape similarity and ecological niche modeling to find rare plant species. *Journal for Nature Conservation*, *17*, 25–32. https://doi.org/10.1016/j.jnc.2008.11.001

Fithian, W., Elith, J., Hastie, T., & Keith, D. A. (2015). Bias correction in species distribution models: Pooling survey and collection data for multiple species. *Methods in Ecology and Evolution*, *6*, 424–438. https://doi.org/10.1111/2041-210X.12242

Fournier, D. A., Skaug, H. J., Ancheta, J., Ianelli, J., Magnusson, A., Maunder, M., ... Sibert, J. (2012). AD Model Builder: Using automatic differentiation for statistical inference of highly parameterized complex nonlinear models. *Optimization Methods & Software*, *27*, 233–249. https://doi.org/10.1080/10556788.2011.597854

Franklin, J. (2009). *Mapping species distributions: Spatial inference and predictions*. New York, NY: Cambridge University Press.

Guest, G. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, *18*, 59–82. https://doi.org/10.1177/1525822X05279903

Guisan, A., Broenniamann, O., Engler, R., Vust, M., Yoccoz, N., Lehmann, A., & Zimmermann, N. (2006). Using niche-based models to improve the sampling of rare species. *Conservation Biology*, *20*, 501–511. https://doi.org/10.1111/j.1523-1739.2006.00354.x

Guo, J. (2011). *Resources and characteristics of habitat of wild Chinese giant salamander (Andrias davidianus) in Shanxi*. Master's thesis, Department of Life Science, Shanxi University, Taiyuan, China.

Harrison, X. A. (2014). Using observation-level random effects to model overdispersion in count data in ecology and evolution. *PeerJ*, *2*, e616. https://doi.org/10.7717/peerj.616

Harrison, X. A. (2015). A comparison of observation-level random effect and beta-binomial models for modeling overdispersion in binomial data in ecology and evolution. *PeerJ*, *3*, e1114. https://doi.org/10.7717/peerj.1114

Hastie, T., & Fithian, W. (2013). Inference from presence-only data; the ongoing controversy. *Ecography*, *36*, 864–867. https://doi.org/10.1111/j.1600-0587.2013.00321.x

Hirzel, A. H., Le Lay, G., Helfer, V., Randin, C., & Guisan, A. (2006). Evaluating the ability of habitat suitability models to predict species presences. *Ecological Modelling*, *199*, 142–152. https://doi.org/10.1016/j.ecolmodel.2006.05.017

Hua, F., Wang, X., Zheng, X., Fisher, B., Wang, L., Zhu, J., ... Wilcove, D. S. (2016). Opportunities for biodiversity gains under the world's largest reforestation programme. *Nature Communications*, *7*, 12717. https://doi.org/10.1038/ncomms12717

Huang, Z. (1982). The Chinese salamander. *Oryx*, *16*, 272–273.

Isaac, N. J. B., Redding, D. W., Meredith, H. M. R., & Safi, K. (2012). Phylogenetically-informed priorities for amphibian conservation. *PLoS ONE*, *7*, e43912. https://doi.org/10.1371/journal.pone.0043912

IUCN (2016). *The IUCN Red List of Threatened Species. Version 2016-2*. Retrieved from www.iucnredlist.org

Li, H., Yu, L., & Ma, J. (2009). Population status and history dynamics of wild Chinese giant salamander (*Andrias davidianus*) in Yanxia Natural Reserve in Guizhou Province, China. *Resources and Environment in the Yangtze Basin*, *18*, 654–657.

Luo, Q. (2009). Habitat characteristics of *Andrias davidianus* in Zhangjiajie of China. *Chinese Journal of Applied Ecology*, *20*, 1723–1730.

Luo, Q., & Kang, L. (2009). Habitat characteristics of Chinese giant salamander in Golden Whip Stream of Zhangjiajie National Forest Park, China. *Chinese Journal of Ecology*, *28*, 1857–1861.

Luo, Q., Liu, Q., Liu, Y., Luo, H., & Tang, C. (2007). Preliminary study on ecological conditions in breeding den of Chinese giant salamanders. *Chinese Journal of Zoology*, *42*, 114–119.

Luo, Q., Liu, Y., & Zhang, L. (2009). The status and countermeasure of protection and augment for Chinese giant salamander resources in Zhangjiajie City. *Journal of Anhui Agricultural Science*, *37*, 9023–9052.

Luo, Q., Zhang, L., Liu, Y., Chen, G., & Gan, M. (2009). Investigation on resources of Chinese giant salamander in Sangzhi County. *Resources and Environment in the Yangtze Basin*, *18*, 727–731.

Lütolf, M., Kienast, F., & Guisan, A. (2006). The ghost of past species occurrence: Improving species distribution models for presence-only data. *Journal of Applied Ecology*, *43*, 802–815. https://doi.org/10.1111/j.1365-2664.2006.01191.x

Manenti, R., & Pennati, R. (2016). Environmental factors associated with amphibian breeding in streams and springs: Effects of habitat and fish occurrence. *Amphibia-Reptilia*, *37*, 237–242. https://doi.org/10.1163/15685381-00003040

Marris, E. (2007). Conservation priorities: What to let go. *Nature*, *450*, 152–155. https://doi.org/10.1038/450152a

McKelvey, K. S., Aubry, K. B., & Schwartz, M. K. (2008). Using anecdotal occurrence data for rare or elusive species: The illusion of reality and the call for evidentiary standards. *BioScience*, *58*, 549–555. https://doi.org/10.1641/B580611

Meijaard, E., & Nijman, V. (2014). Secrecy considerations for conserving Lazarus species. *Biological Conservation*, *175*, 21–24. https://doi.org/10.1016/j.biocon.2014.03.021

Menon, S., Choudhury, B., Khan, M. L., & Peterson, A. T. (2010). Ecological niche modeling and local knowledge predict new populations of *Gymnocladus assamicus* a critically endangered tree species. *Endangered Species Research*, *11*, 175–181. https://doi.org/10.3354/esr00275

Meredith, H. M. R. (2011). *International Conservation Workshop for the Chinese Giant Salamander. Workshop Report (2010)*. Unpublished report to Ocean Park Conservation Foundation, Hong Kong, China.

Murphy, R. W., Fu, J., Upton, D. E., De Lama, T., & Zhao, E. (2000). Genetic variability among endangered Chinese giant salamanders, *Andrias davidianus*. *Molecular Ecology*, *9*, 1539–1547. https://doi.org/10.1046/j.1365-294x.2000.01036.x

Pan, Y., Wei, G., Cunningham, A. A., Li, S., Chen, S., Milner-Gulland, E. J., & Turvey, S. T. (2016). Using local ecological knowledge to assess the status of the Chinese giant salamander (*Andrias davidianus*) in Guizhou Province, China. *Oryx*, *50*, 257–264. https://doi.org/10.1017/S0030605314000830

Pierson, T. W., Yan, F., Wang, Y., & Papenfuss, T. (2014). A survey for the Chinese giant salamander (*Andrias davidianus*) in the Qinghai Province. *Amphibian & Reptile Conservation*, *8*, 1–6.

R Core Team (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raxworthy, C. J., Martinez-Meyer, E., Horning, N., Nussbaum, R. A., Schneider, G. E., Ortega-Huerta, M. A., & Peterson, A. T. (2003). Predicting distributions of known and unknown reptile species in Madagascar. *Nature*, *426*, 837–841. https://doi.org/10.1038/nature02205

Rebelo, H., & Jones, G. (2010). Ground validation of presence-only modelling with rare species: A case study of barbastelles *Barbastella barbastellus* (Chiroptera: Vespertilionidae). *Journal of Applied Ecology*, *47*, 410–420. https://doi.org/10.1111/(ISSN)1365-2664

Segan, D. B., Bottrill, M. C., Baxter, P. W. J., & Possingham, H. P. (2011). Using conservation evidence to guide management. *Conservation Biology*, *25*, 200–202. https://doi.org/10.1111/j.1523-1739.2010.01582.x

Skaug, H., Fournier, D., Bolker, B., Magnusson, A., & Nielsen, A. (2016). *Generalized linear mixed models using 'AD Model Builder'. R package version 0.8.3.3*.

Stewart, G. B., Coles, C. F., & Pullin, A. S. (2005). Applying evidence-based practice in conservation management: Lessons from the first systematic review and dissemination projects. *Biological Conservation*, *126*, 270–278. https://doi.org/10.1016/j.biocon.2005.06.003

Sutherland, W. J., Pullin, A. S., Dolman, P. M., & Knight, T. M. (2004). The need for evidence-based conservation. *Trends in Ecology and Evolution*, *19*, 305–308. https://doi.org/10.1016/j.tree.2004.03.018

Tapley, B., Okada, S., Redbond, J., Turvey, S. T., Chen, S., Lü, J., ... Cunningham, A. A. (2015). Failure to detect the Chinese giant salamander (*Andrias davidianus*) in Fanjingshan National Nature Reserve, Guizhou Province, China. *Salamandra*, *51*, 206–208.

Thompson, W. L. (Ed.). (2013). *Sampling rare or elusive species: Concepts, designs, and techniques for estimating population parameters*. Washington, DC: Island Press.

Turvey, S. T., Bryant, J. V., Duncan, C., Wong, M. H. G., Guan, Z., Fei, H., ... Fan, P. (2016). How many remnant gibbon populations are left on Hainan? Testing the use of local ecological knowledge to detect cryptic threatened primates. *American Journal of Primatology*, *79*, 1–13. https://doi.org/10.1002/ajp.22593

Turvey, S. T., Risley, C. L., Moore, J. E., Barrett, L. A., Hao, Y., Zhao, X., ... Wang, D. (2013). Can local ecological knowledge be used to assess status and extinction drivers in a threatened freshwater cetacean? *Biological Conservation*, *157*, 352–360. https://doi.org/10.1016/j.biocon.2012.07.016

Wang, X., Zhang, K., Wang, Z., Ding, Y., Wu, W., & Huang, S. (2004). The decline of the Chinese giant salamander *Andrias davidianus* and implications for its conservation. *Oryx*, *38*, 197–202.

Yang, J. H., & Chan, B. P. (2016). Two new species of the genus *Goniurosaurus* (Squamata: Sauria: Eublepharidae) from southern China. *Zootaxa*, *3980*, 67–80.

Zajac, Z., Stith, B., Bowling, A. C., Langtimm, C. A., & Swain, E. D. (2015). Evaluation of habitat suitability index models by global sensitivity and uncertainty analyses: A case study for submerged aquatic vegetation. *Ecology and Evolution*, *5*, 2503–2517. https://doi.org/10.1002/ece3.1520

Zhang, Y. (2013). Likelihood-based and Bayesian methods for Tweedie compound Poisson linear mixed models. *Statistics and Computing*, *23*, 743–757. https://doi.org/10.1007/s11222-012-9343-7

Zhang, K., Dearing, J. A., Tong, S. L., & Hughes, T. P. (2016). China's degraded environment enters a new normal. *Trends in Ecology and Evolution*, *31*, 175–177. https://doi.org/10.1016/j.tree.2015.12.002

Zhang, M., Fellowes, J. R., Jiang, X., Wang, W., Chan, B. P. L., Ren, G., & Zhu, J. (2010). Degradation of tropical forest in Hainan, China, 1991-2008: Conservation implications for Hainan gibbon (*Nomascus hainanus*). *Biological Conservation*, *143*, 1397–1404. https://doi.org/10.1016/j.biocon.2010.03.014

Zhang, L., Jiang, W., Wang, Q., Zhao, H., Zhang, H., Marcec, R. M., ... Kouba, A. J. (2016). Reintroduction and post-release survival of a living fossil: The Chinese giant salamander. *PLoS ONE*, *11*, e0156715. https://doi.org/10.1371/journal.pone.0156715

Zhang, Y., & Song, C. (2006). Impacts of afforestation, deforestation, and reforestation on forest cover in China from 1949 to 2003. *Journal of Forestry*, *104*, 383–387.

Zheng, H. (2006). *Population ecology and environmental adaptability of Chinese giant salamander in Lushi County, Henan Province*. PhD thesis, East China Normal University, Shanghai, China.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Chen S, Cunningham AA, Wei G, et al. Determining threatened species distributions in the face of limited data: Spatial conservation prioritization for the Chinese giant salamander (*Andrias davidianus*). *Ecol Evol*. 2018;8:3098–3108. https://doi.org/10.1002/ece3.3862