
Making Consistent IUCN Classifications under Uncertainty

H. REŞİT AKÇAKAYA,*†† SCOTT FERSON,* MARK A. BURGMAN,†
DAVID A. KEITH,‡ GEORGINA M. MACE,§ AND CHARLES R. TODD**

*Applied Biomathematics, 100 North Country Road, Setauket, NY 11733, U.S.A.

†Schools of Botany and Forestry, University of Melbourne, Parkville 3052, Australia

‡New South Wales National Parks and Wildlife Service, Hurstville, New South Wales 2220, Australia

§Institute of Zoology, Zoological Society of London, London NW1 4RY, United Kingdom

**Arthur Rylah Institute for Environmental Research, 123 Brown Street, Heidelberg, Victoria 3084, Australia

Abstract: *The World Conservation Union (IUCN) defined a set of categories for conservation status supported by decision rules based on thresholds of parameters such as distributional range, population size, population history, and risk of extinction. These rules have received international acceptance and have become one of the most important decision tools in conservation biology because of their wide applicability, objectivity, and simplicity of use. The input data for these rules are often estimated with considerable uncertainty due to measurement error, natural variation, and vagueness in definitions of parameters used in the rules. Currently, no specific guidelines exist for dealing with uncertainty. Interpretation of uncertain data by different assessors may lead to inconsistent classifications because attitudes toward uncertainty and risk may have an important influence on the classification of threatened species. We propose a method of dealing with uncertainty that can be applied to the current IUCN criteria without altering the rules, thresholds, or intent of these criteria. Our method propagates the uncertainty in the input parameters and assigns the evaluated species either to a single category (as the current criteria do) or to a range of plausible categories, depending on the nature and extent of uncertainties.*

Establecimiento de Clasificaciones Consistentes de IUCN bajo Incertidumbre

Resumen: *La Unión Mundial para la Conservación (IUCN) definió un grupo de categorías referentes a estados de conservación sustentadas en reglas de decisión basadas en umbrales de parámetros como son el rango de distribución, el tamaño poblacional, la historia poblacional y el riesgo de extinción. Estas reglas han recibido aceptación internacional y se han convertido en una de las herramientas más importantes para la toma de decisiones en biología de la conservación debido a su amplia aplicabilidad, objetividad y simplicidad de uso. Los datos requeridos para estas reglas son frecuentemente estimados con una incertidumbre considerable debido a errores de medición, variación natural y vaguedad en la definición de los parámetros usados en las reglas. Actualmente no existen lineamientos específicos para enfrentar la incertidumbre. La interpretación de datos inciertos por diferentes estimadores puede conducir a clasificaciones inconsistentes debido a que ciertas actitudes hacia la incertidumbre y el riesgo pueden tener una influencia importante en la clasificación de especies amenazadas. Proponemos un método para enfrentar a la incertidumbre que puede ser aplicado a los criterios actuales de IUCN sin alterar las reglas, los umbrales, o la intención de estos criterios. Nuestro método propaga la incertidumbre en los parámetros usados y asigna a la especie evaluada a una sola categoría (a como lo hace el criterio actual) o a un rango de categorías plausibles, dependiendo de la naturaleza y la extensión de las incertidumbres.*

††email: resit@ramas.com

Paper submitted December 1, 1999; revised manuscript accepted March 3, 2000.

Introduction

When resources for conservation are limited, there is an imperative to rank species according to the risks they face. These ranks are used to set priorities for management action at local and national scales (e.g., Czech & Krausman 1997; Breininger et al. 1998), and they are an important part of national and international reporting on the state of the environment. Through legislative and administrative mechanisms, many countries have developed approaches to setting priorities within a context of political and social constraints.

Some threat assessment schemes make use of thresholds to assign scores and sum these scores over a number of attributes to indicate overall conservation status or priority (e.g., Millsap et al. 1990; Lunney et al. 1996). Other schemes use qualitative criteria (e.g., U.S. Fish and Wildlife Service 1983) or a mixture of qualitative and quantitative criteria (e.g., Master 1991; The Nature Conservancy 1994). In the early 1970s, The World Conservation Union (IUCN) adopted a set of qualitative criteria for the classification of conservation status (Fitter & Fitter 1987). Mace and Lande (1991) suggested that status should be assessed quantitatively, and they defined critically endangered species as those facing a 50% probability of extinction within 5 years. The classification of risk involves three parameters: time, probability of decline, and amount of decline (Akçakaya 1992). Threat may then be seen as a combination of the magnitude of the impending decline within some time frame and the probability that a decline of that magnitude will occur. Assessing the threat level is the main goal of the IUCN criteria, which are necessary but probably not sufficient for setting conservation priorities.

The IUCN (1994) defined a set of categories for conservation status supported by decision rules based on thresholds of parameters such as distributional range, population size, and population history, as well as risk of extinction. The IUCN (1994) rules have received international acceptance and have become one of the most important decision tools in conservation biology. Decision rules are attractive because of their wide applicability, objectivity, and simplicity of use (Mace & Lande 1991). By necessity, the choices of thresholds that delimit categories of risk are somewhat arbitrary (Regan et al. 2000).

Existing methods do not explicitly consider the amount and quality of the data, despite the fact the data for different species vary markedly. Like the choice of a method, the choice of the way in which uncertainty is handled can change the resulting classification of threat (Burgman et al. 1999). Moreover, despite differences in the kinds and quality of information from which inferences may be drawn, there is no guidance on how to interpret such variation, although the IUCN (1994) ex-

presses the intent of precaution in the face of uncertainty. This is especially important because our attitude to this uncertainty may have an important influence on the ranks assigned to different species. In particular, there is no consensus regarding the problem of how to rank species when data are missing (Mace 1995). Some methods ignore the issue, effectively relegating the species to a "safe" category whenever data are absent (e.g., Millsap et al. 1990), whereas others let missing data induce a conservation status toward the middle of the range of threat (e.g., Lunney et al. 1996). Our goal is to describe how a new method of dealing with uncertainty can be applied to the current IUCN criteria without altering the rules, thresholds, or intent of the criteria.

Sources of Uncertainty

Any method for classification of conservation status involves several kinds of uncertainty, which may be categorized as semantic uncertainty, measurement error, and natural variability. For example, the IUCN (1994) decision rules require the user to specify the number of adult individuals, the area of occupancy, and the level of fluctuations in these parameters. Each of these parameters is affected to an extent by at least one of these sources of uncertainty.

Semantic uncertainty arises from the use of inexact definitions. For example, the IUCN (1994) asks whether there have been extreme fluctuations in the number of individuals of a species. *Extreme* is defined as the situation in which "population size or distribution area varies widely, rapidly and frequently, typically with variation greater than one order of magnitude . . ." The definition is such that variation in reporting among different people will not be eliminated, even if they are provided with exact information about past fluctuations. The terms *widely*, *rapidly*, and *frequently* may mean different things to different people, and the time horizon over which to evaluate changes of an order of magnitude is not specified. It might be possible to reduce or eliminate variation in responses to this question by making the definition more exact, but only with some loss of generality. This means replacing a vague but inclusive definition by a somewhat arbitrary numerical threshold (Regan et al. 2000).

Some definitions may defy any arbitrary attempt to make them precise. For example, the IUCN (1994) recommends that extent of occurrence be measured by a minimum convex polygon that encloses all known, inferred, or projected occurrences. Different definitions of the minimum convex polygon are possible, however, if different attributes are used to define occurrences. For example, people may choose to include or exclude old records, records that are not substantiated by a mu-

seum or herbarium voucher, records made by “unreliable” collectors, the results of one or another statistical method for extrapolating or interpolating distributions, records that include only juveniles, records of the ranges of floaters in bird populations, records that include only single individuals, records from unusual vegetation types, outliers, and so on. Each of these subjective decisions will affect the estimate of extent of occurrence, and different people will produce different results from the same data. It would be impossible to create a sensible set of rules that would apply to all taxa and that could be interpreted unambiguously. We must conclude that some components of semantic uncertainty are irreducible and that we must learn to live with them. But this does not mean that we should ignore them.

A more fundamental and common type of uncertainty arises from measurement error, the lack of precise information for some or most of the variables used in the rules. It is unlikely, for instance, that the exact number of mature individuals and the exact rate of decline over the past 10 years are known for any species. All numerical data, as well as qualitative information such as whether there are extreme fluctuations, are uncertain.

Measurement error is distinguished from semantic uncertainty by the fact that it may be reduced or eliminated, at least in theory, by acquiring additional data. Estimates of population size rest on the assumption that there is a fact of the matter, a true number of adult individuals, but because we cannot count them all, sampling theory provides the means to estimate the parameter and its error. Taylor (1995) suggested that we interpret uncertainty in the parameters cautiously. Burgman et al. (1999) suggested dealing with this kind of uncertainty within the IUCN (1994) rules by using percentiles other than the median to rank species, depending on the attitude of the interpreter to the risk. Like semantic uncertainty, measurement error can strongly affect the final classification.

Natural temporal and spatial variation in population size and distribution affects our ability to report on these parameters with certainty. Natural variation exists because populations vary in space and time in response to environmental variation and the vagaries of demographic processes (Burgman et al. 1993). For this reason, our expectation of population size over the next few years cannot ever be entirely certain, even in the absence of measurement error or knowledge of the exact details of deterministic trends because chance birth and death events will determine the exact outcome. This uncertainty requires that viability be expressed in probabilistic terms, such as the risk of extinction. In some of the IUCN criteria, natural temporal and spatial variation are not relevant because, for example, population sizes are specified for a particular point in time. In other IUCN criteria, natural variation is accounted for explicitly by

the use of variables that refer to extreme fluctuations, fragmentation, and risk of extinction.

Characterizing Uncertainty

The IUCN rules for classification of threatened species consist of three sets, which define the classes critically endangered (CR), endangered (EN), and vulnerable (VU). The evaluation of a given taxon under each of these sets of rules results in a boolean (true or false) outcome. The rule sets are nested, so that there are only four possible classifications (Table 1).

Each boolean (Table 1) is evaluated based on five criteria, as follows: A or B or C or D or E (expression 1). The five criteria consider population decline (A), restricted distribution and decline, fluctuation, and/or fragmentation (B), small population size and decline (C), very small population size (D), and high risk of extinction (E). Each criterion (A to E) is, in turn, evaluated by comparing the data for the taxon to threshold values. For example, criterion D for CR is number of mature individuals < 50 (expression 2). Several such comparisons may be combined with logical operators AND and OR within each criterion. For example, criterion A for CR can be summarized as (past reduction \geq 80%) or (future reduction \geq 80%) (expression 3), where both past and future reductions are measured over a time period of 10 years or three generations, whichever is longer.

Specifying Uncertain Data

There are various ways to specify uncertainty in a quantity such as number of mature individuals. When natural variability is the main source of uncertainty and the sampling program has been sufficiently rigorous, we may specify confidence intervals. Most of the criteria in the IUCN (1994) rules include at least some semantic uncertainty. Even those variables amenable to measurement sometimes are not measured but are the subject of expert opinion or guesswork informed by a few samples.

Table 1. Possible outcomes of evaluation of a given taxon under IUCN rules to classify the taxon into a threat category.

| VU | Rule sets* | | | Classification of the taxon |
|-------|------------|-------|--|-----------------------------|
| | EN | CR | | |
| true | true | true | | critically endangered |
| true | true | false | | endangered |
| true | false | false | | vulnerable |
| false | false | false | | lower risk |

*Three rules sets of IUCN (1994): VU, vulnerable; EN, endangered; and CR, critically endangered.

When different kinds of uncertainty are confounded, one of the simplest ways to represent uncertainty is to specify a best estimate and a range of plausible values. For example, we may represent an adult population size with a best estimate of 90 and bounds of 70 and 120. The bounds may be based on a variety of measures ranging from statistically estimated likelihood distributions to the various opinions of a group of experts.

The resulting information can be represented as a triangular fuzzy number (Kaufmann & Gupta 1985; Ferson et al. 1998), as in Fig. 1a. The x -axis is the number of mature individuals. The y -axis of this and all other fuzzy numbers we use in this paper gives the possibility level. At each possibility level, there is an interval defined by the left and right sides of the fuzzy number. This level inversely measures the surety that the parameter is within the interval at that level: as the scale on the y -axis increases, the surety decreases that the parameter is within the corresponding interval.

In cases when consensus about the best estimate is not possible, it may be necessary to pool different estimates from different experts or from several assessment methods. In such cases the best estimate itself may be an interval such as 85–95 instead of a point estimate. In this case, the resulting object can be represented as a trapezoidal fuzzy number (Fig. 1b). This number implies that the true value of the parameter is around 85 to 95 (the interval at the top of the fuzzy number) and that it is surely within the range of 70–120 (at the bottom of the fuzzy number). If it was obtained by pooling expert opinion, it may imply that everyone agrees that the values 85–95 are plausible, whereas the interval [70, 120] is the sum (union) of all opinions about plausible values.

In some cases, the data to be pooled may be contradictory (for example, [70, 90] and [100, 120]). In such cases, automatic rules for pooling are not appropriate. A better approach to the problem would be communication between the experts who produced the contradictory data, with the purpose of understanding where the disagreement is coming from. It is most likely that the cause of the difference is use of different information. When such information is shared, the problem may disappear.

When data are very uncertain, the range for the best estimate may coincide with the range of plausible values. In this case, the resulting object can be represented as a rectangular fuzzy number (i.e., an interval; Fig. 1c).

Some of the data for the IUCN rules are required in the form of yes/no or true/false—for example, whether there are extreme fluctuations in numbers and whether all individuals are in a single population. It is not always possible to be certain about responses to these questions for the same reasons it is not possible to be certain about quantitative information. Measurement error may make the response to questions about the magnitude of fluctuations or the number of populations unreliable.

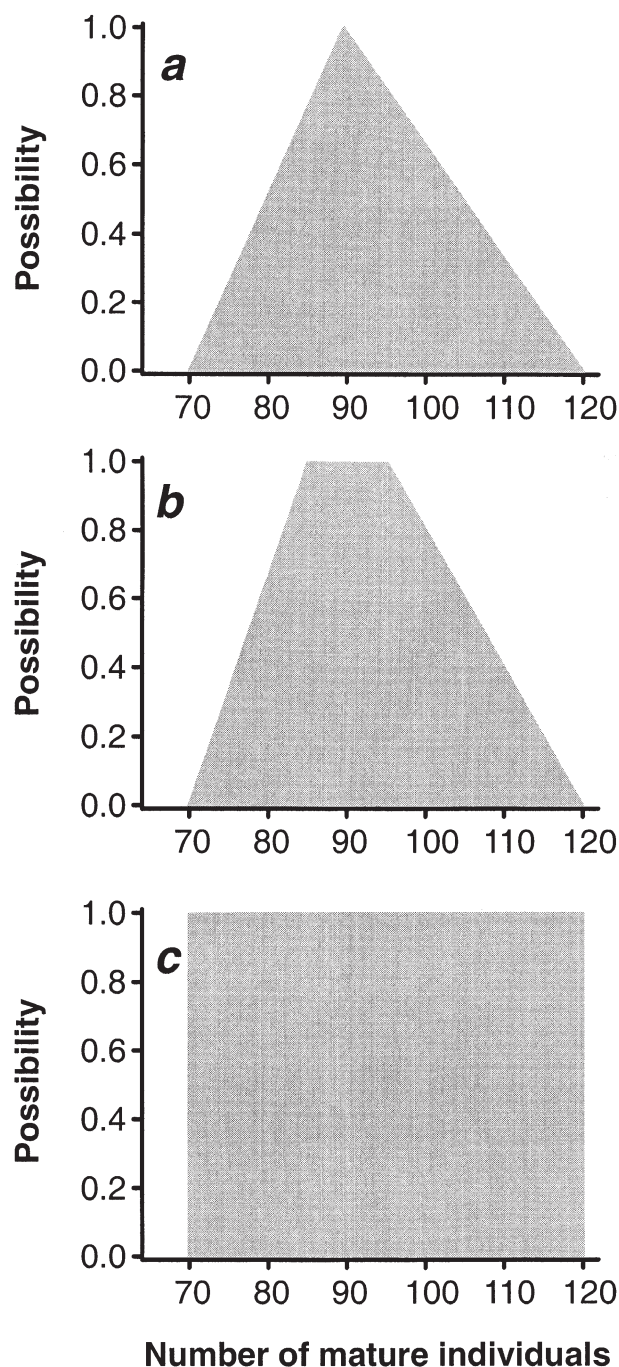


Figure 1. Examples of (a) triangular, (b) trapezoidal, and (c) rectangular fuzzy numbers for the number of mature individuals of a species.

Different people may hold different opinions. In addition, semantic uncertainty may make the uncertainty irreducible. For example, estimate of the number of populations depends on the interpretation of the term population, a concept fraught with subjectivity (Wells & Richmond 1995).

When the information about these aspects of the status of a taxon are uncertain, the answer can be ex-

pressed as a number between 0 (false) and 1 (true), representing the truth value of the statement. In some cases, this number may represent the frequency with which the statement is true. More generally, it represents the degree of plausibility of the statement or the reliability of the information about that aspect of the population (Fig. 2a). When assessments made by a number of experts are pooled, the result can be expressed as an interval (Fig. 2b). If the procedure for pooling expert opinion includes assigning different weights to different opinions (based on expertise, experience, or performance), the result may be expressed as a triangular or trapezoidal fuzzy number (Fig. 2c).

It is important to note that fuzzy numbers are not probability distributions; they are simply stacks of intervals nested so that those below include those above. The calculations described here are extensions of simple interval arithmetic and can be applied to the two intervals (the best estimate at the top and the plausible range at the bottom) and to any other interval in between (Kaufmann & Gupta 1985; for a synopsis of fuzzy calculations, see the Appendix).

Using Uncertain Data in IUCN Rules

When the parameters are uncertain and are specified as described above, expressions 1, 2, and 3 must be generalized to allow fuzzy numbers. With uncertain input data specified as fuzzy numbers (e.g., as in Figs. 1 & 2), the results of expressions 1, 2, and 3 are themselves fuzzy logical values ranging between 0 and 1 that are derived from methods discussed in the Appendix (these methods generalize operators such as AND, OR, $>$, $<$, \geq , \leq , and others).

Similarly, the result of the three rule sets, CR, EN, and VU, are also fuzzy numbers between 0 and 1, representing the truth values for membership of a taxon to each rule set. Thus, a taxon has a separate truth value for membership to each category. In this case, the possible outcomes (Table 2 lists some) are no longer limited to the four cases listed in Table 1. *True* is represented by the graph with a spike at the right end, which corresponds to the logical value of 1; *false* is represented by the graph with a spike at the left end, which corresponds to the logical value of 0; and *unknown* is represented by a solid rectangle, which corresponds to the logical interval of [0, 1] (Table 2). Other fuzzy numbers represent gradations among these cases.

Once the three fuzzy numbers representing CR, EN, and VU are calculated for a taxon, based on uncertain input data and the calculations described in the Appendix, the next step is to combine these into a single classification of threat. Several word descriptions of classification of threat are given as examples in Table 2; the same set of uncertainties may be described in different ways by different people. The step of combining the three fuzzy

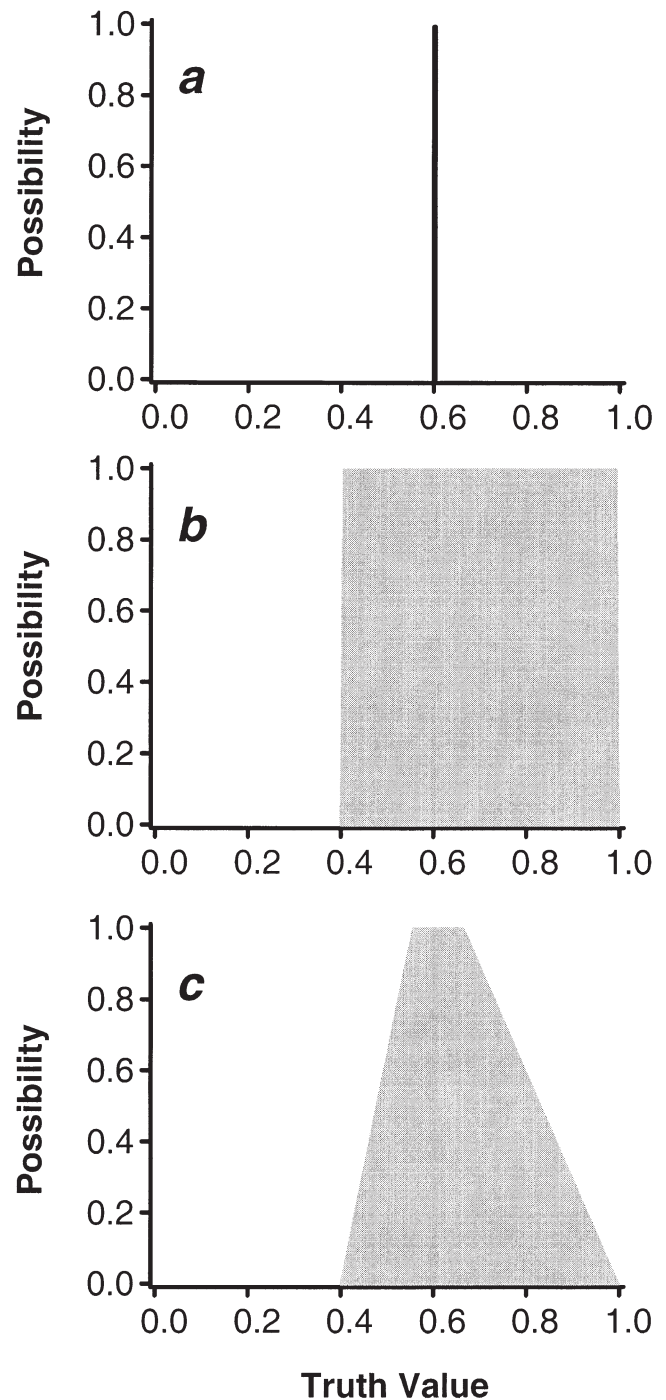


Figure 2. Examples of fuzzy numbers representing a logical value, such as whether there are extreme fluctuations in a population. A value of 0 represents false; 1 represents true. Intermediate values, such as 0.6 (a) represent uncertainty and can be estimated, for example, as the proportion of cases (such as populations) for which the statement is true. Further uncertainty (e.g., due to range of opinion) can be represented using (b) an interval or (c) a trapezoidal fuzzy number.

Table 2. Sample outcomes of evaluation of a given taxon with uncertain data under IUCN rules.

| Rule sets* | | | Classification of the taxon |
|------------|----|----|---|
| VU | EN | CR | |
| □ | □ | □ | critically endangered |
| □ | □ | □ | endangered |
| □ | □ | □ | vulnerable |
| □ | □ | □ | lower risk |
| □ | ■ | ■ | at least endangered, maybe critically |
| □ | ■ | □ | at least vulnerable, maybe endangered, but not critically |
| □ | ■ | ▲ | at least vulnerable, likely endangered, possibly critically |

*Three rules sets of IUCN (1994): VU, vulnerable; EN, endangered; and CR, critically endangered.

numbers into a single classification of threat can be done in different ways, because it depends on preferences and attitudes toward risk and uncertainty.

Attitudes toward Risk and Uncertainty

Uncertain data may lead to inconsistent classifications because different people emphasize different parts of a range or different aspects of the uncertainty. These emphases can be summarized as differences in attitude toward risk and uncertainty. Different attitudes are unavoidable, but this need not make rule-based classifications (such as the IUCN criteria) inconsistent as long as the underlying attitude is explicitly stated. We recognize two fundamental aspects of attitude about risk and uncertainty: dispute tolerance and risk tolerance.

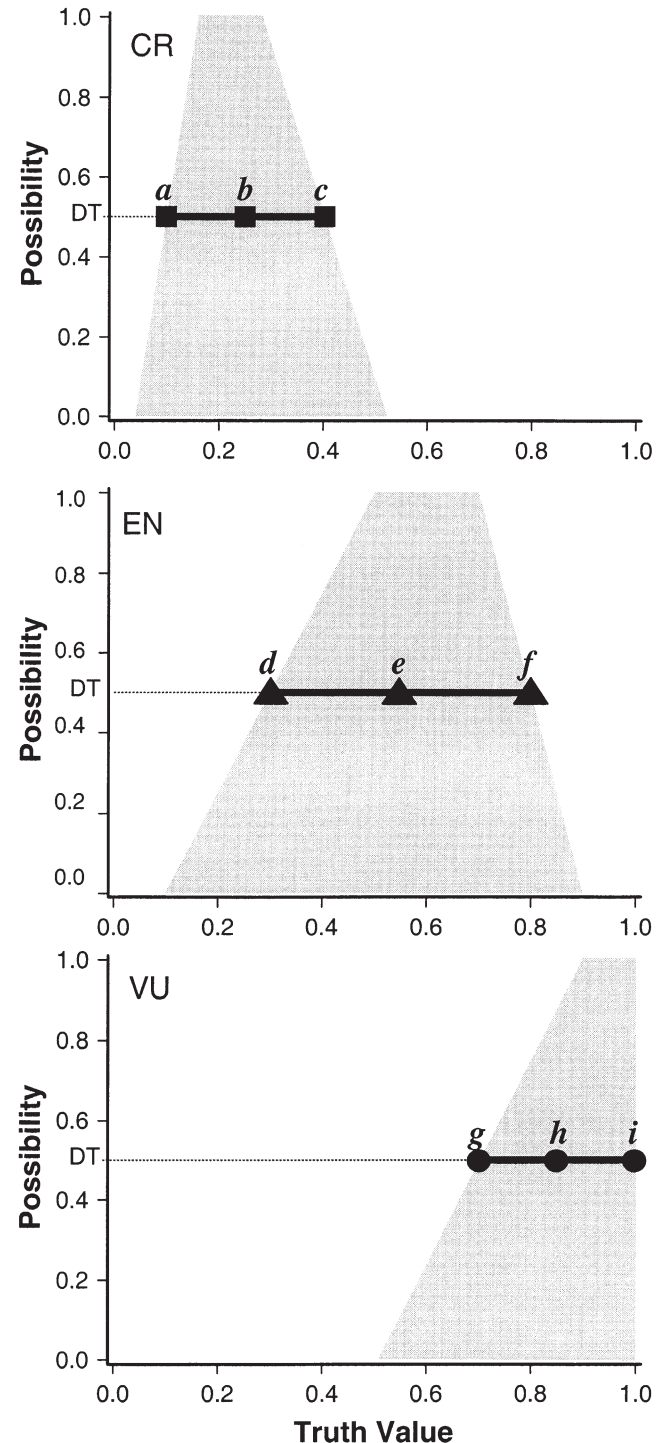
Dispute Tolerance

When input data are certain, the three logical values for CR, EN, and VU are either 0 (false) or 1 (true). In other words, they are booleans. When the input data are uncertain, each of these may be a fuzzy number (as in Table 2).

Figure 3. An example of the three fuzzy numbers, which are the result of the rule sets for classifications of critically endangered (CR), endangered (EN), and vulnerable (VU), calculated according to the IUCN rules and methods described in the Appendix. Dispute tolerance (DT) determines the level (represented by the horizontal line) to be used in determining the degree of threat. The bottom of each fuzzy number represents the case that includes all possibilities (DT = 0). The top of each fuzzy number represents the case when dispute tolerance is highest (DT = 1). The points of intersection (e.g., a and c for CR) with the DT line and the middle point (e.g., b for CR) are used in Fig. 4.

The gray polygons in Fig. 3 are examples of the three fuzzy numbers for CR, EN, and VU, calculated according to the IUCN rules with methods described in the Appendix (each cell of Table 2 contains one such fuzzy number).

The x-axis represents the truth value (e.g., the truth value of the rule set EN). The y-axis represents the possibility levels, as in all other fuzzy numbers. The uncertainty in the data is reflected in the width of the fuzzy number. For example, the bottom of the fuzzy number



might enclose the result of the rule set EN using the full range of opinions from a number of experts. The top of the fuzzy number might represent the result of the same rule set based on the best estimates for each parameter from these same experts.

Dispute tolerance (DT) determines the level (represented by the horizontal line) to use in determining the degree of threat. The bottom of the fuzzy number represents the case that includes all possibilities (DT = 0). The top of the fuzzy number represents the case when dispute tolerance is highest (consensus; DT = 1). Dispute tolerance represents how those interpreting the data feel about uncertainty. If the interpreters want to be sure of encompassing all possibilities, thus avoiding disagreements and dispute, they would use a DT close to 0. If they want to be restrictive in data use and rely only on the experts' best estimates, they would use a DT close to 1. A high DT would mean disregarding the more extreme opinions, which might lead to dispute. If the tails of the distribution seem too extreme but the best estimates are too small to bound uncertainty, the interpreter may choose an intermediate value of, say, DT = 0.5.

Risk Tolerance

Risk tolerance (RT) ranges from 0 for risk-averse, precautionary; through 0.5 for risk neutral; to 1 for risk prone, evidentiary. A precautionary attitude would accept that a species is safer than endangered only if it is certainly not endangered, and RT would be closer to 0 than to 1. An evidentiary attitude would demand substantial evidence of endangerment before allowing such a classification, and RT would be closer to 1. As when dealing with dispute tolerance, interpreters may choose to take an intermediate stance with, say, RT = 0.5 or to vary risk tolerance from case to case, depending on what may be lost or gained by a classification that might be either too high (risk overestimated) or too low (risk underestimated). Of course, mistakes of both kinds carry consequences in terms of the inefficient use of scarce conservation resources, and the possibility of inappropriate priorities for action that lead to a species' demise that might otherwise have been avoided.

Combining Attitudes with Uncertain Data

A given level of dispute tolerance reduces the three fuzzy numbers for CR, EN, and VU to three intervals with midpoints, such as those in Fig. 3:

$$\text{CR: } [a, b, c] = [0.10, 0.25, 0.40];$$

$$\text{EN: } [d, e, f] = [0.30, 0.55, 0.80];$$

$$\text{VU: } [g, h, i] = [0.70, 0.85, 1.00].$$

These intervals can be ordered as in Fig. 4. The categories corresponding to CR, EN, and VU are on the x-axis,

and the lower, mid, and upper values of the intervals are plotted on the y-axis. Thus, each interval is represented by a vertical stack of up to three points (square for CR, triangle for EN, and circle for VU). The width of these intervals is a function of dispute tolerance attitude. The middle points b , e , and h are needed for making a "most plausible" classification under uncertain data. If they are omitted, and only the lower (a , d , g) and upper (c , f , i) values are used, some species with uncertain data will be classified in a range of categories (such as EN to VU) and will not be assigned to a single most plausible category. These ranges need not be in the middle, however. Their location can be defined as a third attitude option.

Fig. 4 shows three lines, each starting at the lower left corner of the graph and ending in the upper right corner. The lower line connects the lower bounds of the three intervals, the mid line (dotted) connects the midpoints of the three intervals, and the upper line connects the upper bounds of the three intervals.

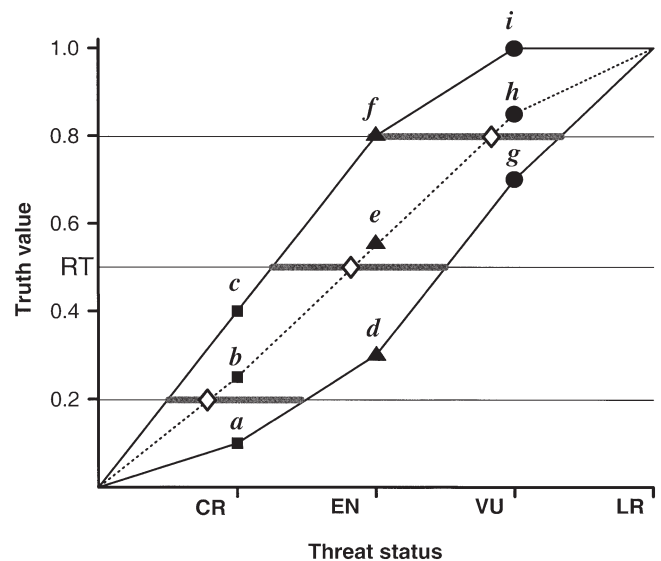


Figure 4. Graphical representation of the method for combining attitudes and uncertain results to obtain threat status based on IUCN criteria. The set of three numbers for each category is derived from Fig. 3: squares mark points [a, b, c] for critically endangered (CR), triangles mark points [d, e, f] for endangered (EN), and circles mark points [g, h, i] for vulnerable (VU). These points correspond to the points in Fig. 3. The LR represents the nonthreatened category of "lower risk." The intersection of the middle curve with the horizontal lines (RT, risk tolerance) gives the threat status (marked with a diamond). The intersection of upper and lower curves with RT gives the range of plausible categories (thick line). This threat status and the range of plausible categories are used to classify the species in Fig. 5.

These three lines are intersected by horizontal lines that represent the risk tolerance attitude. This intersection (the thick horizontal line) is the range of categories that describes the threat status of this taxon, and the intersection with the dotted middle line (diamond-shaped markers) gives the threat status of this taxon. Thus, this taxon is classified as VU if $RT = 0.8$, EN if $RT = 0.5$, and CR if $RT = 0.2$. The status of the taxon under these three RT values is given in Fig. 5, which is similar to Fig. 4 but with less detail. The range of plausible categories are CR to EN with $RT = 0.2$, EN to VU with $RT = 0.5$, and EN to LR with $RT = 0.8$. Thus, the effect of lower risk tolerance is a higher threat category.

Grevillea caleyi: a Case Study

We demonstrate our proposed method of propagating uncertainty in IUCN rules with a case study of *Grevillea caleyi*, a shrub that grows to about 2.5 m in dry sclerophyll Eucalypt forest on the northern outskirts of Sydney, Australia. A few scattered populations remain, and they face changed fire regimes and the pressures of land clearance and changes in seed predation.

The population size of *Grevillea caleyi* is estimated as a triangular fuzzy number, [800, 2000, 3000]. This estimate is based on field samples and extrapolations over all potential habitat and over known and suspected populations. We take the lower bound to be the lower estimates of population size and apply them to known populations. We take the upper bound to be high estimates of population size based on field samples and apply them to known populations as well as potential populations in unexplored habitat.

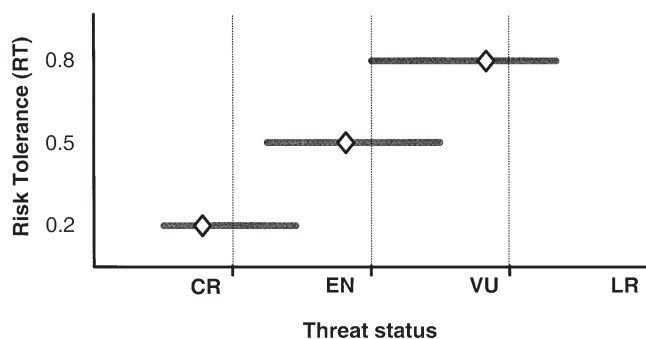


Figure 5. Threat classification based on Fig. 4. Each line indicates the range of plausible categories, and the diamond marker indicates the most plausible threat status. The lines and markers are determined in Fig. 4. With $RT = 0.2$, the status is CR, with the plausible range of CR to EN; with $RT = 0.5$, the status is EN, with the plausible range of EN to VU; with $RT = 0.8$, the status is VU, with the plausible range of EN to LR.

For extreme fluctuations in population size, area, extent, and number of populations, we specified point estimates of 1 (true). The populations experience extreme fluctuations because they are periodically burned and standing plants killed. Seedlings regenerate after fire, but the seed store is depleted until the new germinants grow to maturity and replenish it. Populations are at risk because there is a possibility of repeated fires at intervals that preclude the development of adult plants (Bradstock et al. 1998). Area and extent fluctuate for the same reasons that populations fluctuate in size at any given location, because periodic planned and unplanned fires kill adult plants. Populations may be eliminated and ranges changed when fire frequencies become too high. In addition, there may be trends in area and extent that result from habitat loss from urban development.

For continuing declines in population size, extent, area, and quality of habitat, we specified a value of 0.5 (midpoint between true and false; suspected). This reflects the fact that we are uncertain about whether, in fact, there is a continuing decline. Any deterministic trends are masked by natural variation and sampling error, making any such judgements uncertain, no matter how reliable the data. Maximum uncertainty could be specified by a uniform distribution between 0 and 1 (no knowledge).

Past population reduction (specified as [0, 5, 25]%) is uncertain because no exact records of population sizes exist for the last 20 years. Future population reduction (specified as [5, 10, 30]%) is uncertain because it depends on natural variation and on how fire regimes are managed. These parameters were determined in part by economic and property risks and by risks to human life (Gill & McCarthy 1998), and they are inherently unpredictable. The bounds were derived from expert judgement of the plausible extremes.

The extent of occurrence of [4, 6, 7] km² was based on minimum convex polygons of known locations. The extent is uncertain because of the concern for the appropriateness of including a small outlying population that might have been introduced. Area of occupancy of [0.15, 0.25, 1.00] km² was based on field surveys of known populations and extrapolation to potential habitat. The bounds were derived from expert judgement of the plausible extremes, supported by maps of potential habitat.

The number of populations and locations was estimated as [3, 5, 7]. Populations are more or less isolated from one another by roads and other human infrastructure. There are no obvious dispersal mechanisms to carry seeds between patches. Thus, we conclude that the species' populations might be severely fragmented, and we specified the truth value as [0.5, 1.0]. The size of the largest population is [200, 700, 900], based on field sampling and the assumption that the known populations include the largest.

An individual-based model was written to calculate the likelihood of extinction of populations, taking into account the dynamics of fires within patches, the dynamics of seed banks, local competition for space between plants, predation of seeds by small mammals, and several other factors (Regan et al. 1998). If we assume that all individuals of a species are in a single population, thus, the probability of extinction ranges from [0.1, 0.5, 1.0] in 100 years to [0, 0.1, 0.5] in 21 years. The upper and lower bounds depend on the level of optimism selected for the parameters and structures in the model. The model was designed to predict extinction risks in single populations. The probability of extinction of the species as a whole is not the primary focus of this modeling exercise. The integration of the single-population results to estimate extinction risks for the entire species involves assumptions about the spatial distribution of fire events, creating a degree of subjective uncertainty in the predictions.

When we used point estimates for each of the parameters, *G. caleyi* was classified as CR. Criterion B determined this status because the extent of occurrence is <100 km², the area of occupancy is <10 km², decline is continuing, and several population attributes show extreme fluctuations.

We are uncertain about the fact that there is a continuing decline. If we had decided instead that there was no continuing decline, the species would have been classified as EN, again on the basis of criterion B, not for the reasons listed above but because it exists at five or fewer locations (criterion B1) and populations are fluctuating (B3) rather than exhibiting a continuing decline (B2). This emphasizes the importance of dealing explicitly with uncertainty. In this case, the classification of the species as CR or EN rests on the belief that there is a continuing decline. Continuing declines are masked by natural variation, sampling error, and uncertainty in the definition of terms and parameters. We are reasonably certain, however, that the species is at least endangered (at least, if we ignore uncertainty in the other parameters).

With the methods described above, and with attitude options DT = 0.5 and RT = 0.5, the species is again classified as CR, with plausible categories including EN and VU. This classification depends on attitude to risk (Fig.

6). In this case, if those classifying the species demand that evidence be presented to establish that a higher classification is warranted (RT = 0.8), then the species would be classified as VU. In this case, the classification was not very sensitive to different attitudes toward uncertainty (compare middle three lines with different values of dispute tolerance, DT). Figures like this one can be used as sensitivity analyses to reveal the weight of choices about attitude.

Discussion

Types of Uncertainty

Our proposed approach provides a general framework for incorporating all three types of uncertainty we described in the introduction. Nevertheless, it is most appropriate to interpret the uncertainties in terms of measurement error. Some types of semantic uncertainty, such as vagueness, can also be represented with fuzzy numbers for input data. A more efficient way to deal with these uncertainties might be to make the definitions more explicit and to describe the method of measurement of each variable in detail. As we discussed above, it might be possible to reduce or eliminate semantic uncertainty in some cases, but only at some cost of loss of generality. In other cases, it may not be possible to reduce it. When it is not desirable or possible to reduce semantic uncertainty, as when flexibility is a desired characteristic of rule sets, the approach we have outlined can incorporate such uncertainties.

Natural variation in input variables can also be represented with fuzzy numbers, but two better ways of incorporating natural variability are the use of probability distributions and the inclusion of information on natural variation in the input data. Using probability distributions instead of fuzzy numbers is a more efficient way to incorporate natural variability. This is because probability distributions make it easier to add more information about the variable, including the shape and variance of the distribution and the exact dependencies, when these are known.

We believe the most efficient method of accounting for natural variability in rule sets is to explicitly add in-

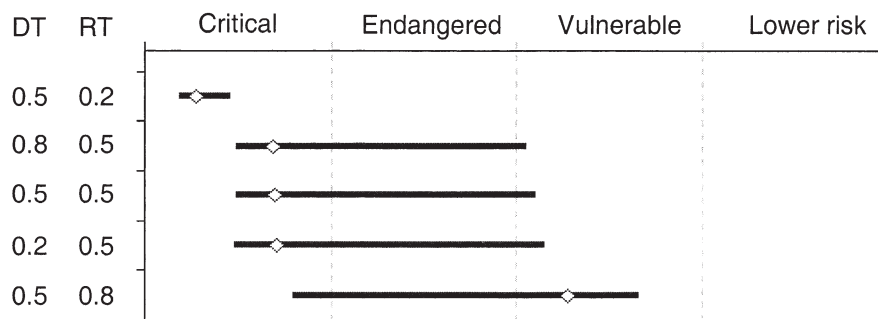


Figure 6. Threat classification of *Grevillea caleyi* according to IUCN rules under uncertainty and with different attitudes toward dispute tolerance and risk tolerance. Each line indicates the range of plausible categories, and the diamond marker indicates the threat status.

formation on natural variability. This is already (partly) done in the IUCN rules. For example, the rules ask for information about current total abundance (thus explicitly excluding spatial and temporal variability) and for other information on temporal variability (e.g., whether there are fluctuations in the number of populations and locations and the risk of extinction) and spatial variability (e.g., whether all individuals are in the same population).

Meaning of Unknown

One of the problems with using uncertain data in classifications relates to quantities for which there are no data at all. As the IUCN rules are applied now, one is allowed to omit a criterion for which there are no data. Thus, if data for one or more of the criteria A to E are unknown, that criterion is omitted from expression 1. Operationally, this is equivalent to answering the question for that criterion as “no.” For example, if the number of mature individuals is unknown, criterion D (expression 2) would be assumed to be unmet, which gives the same result as assuming it to be false. There are three problems with this approach. First, when uncertainty is considered explicitly, “no” and “don’t know” should be distinct answers. In this example, the answer (to the question “Is the number of mature individuals <50 ?”) should be the interval $[0, 1]$, representing complete ignorance about this logical value. Second, setting D equal to false gives the least conservative answer because it is equivalent to assuming that the number of mature individuals is ≥ 1000 (the threshold for VU, the least threatened category). Third, setting D equal to false gives the narrowest bounds on the answer. If the data had suggested that the number of mature individuals was, for example, $[40, 60]$ the classification would have been both more threatened (including CR) and more uncertain (including more than one category). Thus, assuming that “unknown” is the same as “false” rewards those with the least amount of data. Those that have more data are punished, in the sense that they receive a more uncertain classification than those who did not even try to collect the necessary information.

One solution to the problem is to set “unknown” to the total range of possible and relevant answers. For example, if the number of mature individuals is unknown, it can be set to the interval $[0, \infty]$. If it is not known whether the taxon experiences extreme fluctuations, the answer can be set to the interval $[0, 1]$. The methods we describe will then propagate the uncertainty in a proper and consistent way. The results will be truthful, and the procedure will encourage collection of data. The problem with this approach is that the classification of many species may include the category CR. Although this may well be justified based on the available data, it may also be rejected on practical grounds. Our sugges-

tion is to leave this decision to the interpreters of data and ask them to explicitly state whether any criterion should be ignored (omitted from calculations).

Advantages of the Proposed Approach

There may be various ways of using uncertain data in the classification of species, according to the IUCN criteria. The particular approach we propose has a number of advantages:

- It does not alter the rules, thresholds, or intent of the IUCN criteria.
- When there is no uncertainty, the classification is the same as the current application of the IUCN criteria.
- When there is only minimal uncertainty, the classification does not change substantially or qualitatively from the classification without uncertainty.
- The classification does not change when the uncertainty is irrelevant—for example, when the range for number of mature individuals does not include the threshold for any of the categories.
- In other cases, the width of the range of resulting categories, such as CR to EN, increases with increasing uncertainty.
- The proposed method is simpler than many alternatives. Asking for only a best estimate and an interval (a range of plausible values) is much simpler than asking for probability distributions or other representations of uncertainty.
- The method is based on intervals, which, unlike probability distributions, make minimal assumptions about mass; they are only bounds. This approach uses the most elementary method of uncertainty propagation.
- The method discourages ignoring any of the criteria.
- It recognizes that different people may have different attitudes to risk and uncertainty.
- It leaves arbitrary decisions, such as those about attitude, to the interpreters, but it requires an explicit statement of these decisions. When different interpreters reach different conclusions based on the same data, the differences may be traced to specific decisions. Given a certain and explicitly stated set of attitudes, the classifications are consistent.
- When the effect of uncertainty on the listing is made transparent, it becomes impossible to use uncertainty as an excuse for promoting a certain agenda of either under- or overprotecting species. If parties disagree with the listing, they have the option of either arguing for different priorities through the choice of the attitude parameters or reducing the range of plausible listing categories by gathering better data.
- It distinguishes between and provides acceptable so-

lutions in circumstances in which (1) we are ignorant of required information, (2) we wish to ignore a piece of information, or (3) we are indifferent to the risks posed by a particular circumstance.

These benefits are not achieved without some costs. Decisions about where to set DT and RT levels need to be made by some process that links them to values that people have for biodiversity conservation. It may be that conservation biologists and resource managers typically have contrasting attitudes (e.g., Mace & Hudson 1999), and the setting of these levels will become the focus for disagreements rather than the status of particular species. Unfortunately, despite the fact that it is precisely these attitudes and values that underlie many disagreements, few biologists are trained to tackle such debates.

Despite such difficulties, our approach can provide a basis for constructive methods to resolve disputes over species listings. For example, most of the difficulty and controversy involved in listing decisions at present is the result of uncertainty of various kinds. Commonly, a species might change from one category to another depending upon how the evaluator chooses to extrapolate information from one part of a species range to the entire species, or whether or not the evaluator believes natural fluctuations can explain current trends. Without additional data, such issues are essentially unresolvable, and so the debate becomes a polarized one based on different attitudes to risk and uncertainty. A procedure such as the one we have described can formalize and structure such a debate. In addition, if the operational levels for DT and RT are determined beforehand, a conclusion can be reached based on accepted principles rather than personal perspectives.

Our approach also clarifies the different threat processes that evaluators may wish to represent in prospective classifications. For example, under criterion A in the IUCN system, species may be listed as threatened if the population is expected to decline by a specified rate in the future. This expected decline may be deduced from quite deterministic processes, such as continuing habitat loss from commercial harvesting that proceeds at a predictable rate. Alternatively, future declines may be expected as a result of stochastic events of high impact but low predictability, such as changes in the area that may influence key breeding sites and affect juvenile survival rates. In the former case, there is a decline rate that can be measured more or less precisely, whereas in the latter the expected decline rate has more uncertainty associated with it. Under the simple implementation of the IUCN system, these two cases might be given the same expected decline rate. Under the system we have described, the larger uncertainty of the second case contributes differently to the outcome. The transparency of dealing with these difficult listing decisions is an important byproduct of this system.

Acknowledgments

Discussions with H. Regan and M. Colyvan helped clarify our ideas. We also thank B. Taylor, A. Punt, and two other reviewers for helpful comments and B. Root for implementing the proposed method as computer software. H.R.A. and S.F. were supported by grants from the National Science Foundation, the National Institutes of Health, and the Electric Power Research Institute.

Literature Cited

- Akçakaya, H. R. 1992. Population viability analysis and risk assessment. Pages 148–157 in D. R. McCullough and R. H. Barrett, editors. *Wildlife 2001: populations*. Elsevier, Amsterdam.
- Akçakaya, H. R., and S. Ferson. 1999. RAMAS red list: threatened species classifications under uncertainty. User manual for Version 1.0. Applied Biomathematics, Setauket, New York.
- Bradstock, R. A., M. Bedward, B. J. Kenny, and J. Scott. 1998. Spatially-explicit simulation of the effect of prescribed burning on fire regimes and plant extinctions in shrublands typical of south-eastern Australia. *Biological Conservation* **86**:83–95.
- Breining, D. R., M. J. Barkaszi, R. B. Smith, D. M. Oddy, and J. A. Provanca. 1998. Prioritizing wildlife taxa for biological diversity conservation at the local scale. *Environmental Management* **22**: 315–321.
- Burgman, M. A., S. Ferson, and H. R. Akçakaya. 1993. Risk assessment in conservation biology. Chapman and Hall, London.
- Burgman, M.A., D.A. Keith, and T.V. Walshe. 1999. Uncertainty in comparative risk analysis of threatened Australian plant species. *Risk Analysis* **19**:585–598.
- Czech, B., and P. R. Krausman. 1997. Distribution and causation of species endangerment in the United States. *Science* **277**:1116–1117.
- Ferson, S., W. Root, and R. Kuhn. 1998. RAMAS risk calc: risk assessment with uncertain numbers. Applied Biomathematics, Setauket, New York.
- Fitter, R., and M. Fitter, editors. 1987. *The road to extinction*. World Conservation Union and United Nations Environmental Programme, Gland, Switzerland.
- Fréchet, M. 1935. Généralisations du théorème des probabilités totales. *Fundamenta Mathematica* **25**:379–387.
- Gill, A. M. and M. A. McCarthy. 1998. Intervals between prescribed fires in Australia: what intrinsic variation should apply? *Biological Conservation* **85**:161–169.
- Hailperin, T. 1986. *Boole's logic and probability*. North-Holland, Amsterdam.
- IUCN (Species Survival Commission). 1994. *International Union for the Conservation of Nature Red List Categories*. Gland, Switzerland. Also available from <http://www.iucn.org/themes/ssc/redlists/ssc-rl-c.htm>.
- Kaufmann, A., and M. M. Gupta. 1985. *Introduction to fuzzy arithmetic: theory and applications*. Van Nostrand Reinhold, New York.
- Klir, G. J., and B. Yuan. 1995. *Fuzzy sets and fuzzy logic: theory and applications*. Prentice Hall PTR, Saddle River, New Jersey.
- Lunney, D., A. Curtin, D. Ayers, H. G. Cogger, and C. R. Dickman. 1996. An ecological approach to identifying the endangered fauna of New South Wales. *Pacific Conservation Biology* **2**:212–231.
- Mace, G. M., 1995. Classification of threatened species and its role in conservation planning. Pages 197–213 in J. H. Lawton and R. M. May, editors. *Extinction rates*. Oxford University Press, Oxford, United Kingdom.
- Mace, G. M., and E. J. Hudson. 1999. Attitudes toward sustainability and extinction. *Conservation Biology* **13**:1–6.
- Mace, G. M., and R. Lande. 1991. Assessing extinction threats: towards

a re-evaluation of IUCN threatened species categories. *Conservation Biology* 5:148-157.

Master, L. L. 1991. Assessing threats and setting priorities for conservation. *Conservation Biology* 5:559-563.

Millsap, B. A., J. A. Gore, D. E. Runde, and S. I. Cerulean. 1990. Setting the priorities for the conservation of fish and wildlife species in Florida. *Wildlife Monographs. Supplement Journal of Wildlife Management* 54:5-57.

Regan, H. M., A. Rawlinson, D. A. Keith, T. Auld, and M. Burgman. 1998. Population viability analysis for *Grevillea caleyi*, *Epacris stuartii* and *E. barbata*. Pages 53-71 in *Plant population viability analysis and case studies*. Project FN-37. Environment Australia, University of Melbourne, Melbourne.

Regan, H. M., M. Colyvan, and M. A. Burgman. 2000. A proposal for fuzzy IUCN categories and criteria. *Biological Conservation* 92: 101-108.

Taylor, B. L. 1995. The reliability of using population viability analysis for risk classification of species. *Conservation Biology* 9:551-558.

The Nature Conservancy. 1994. The conservation of biodiversity in the Great Lakes ecosystem: issues and opportunities. Great Lakes Program, Chicago.

U.S. Fish and Wildlife Service. 1983. Endangered and threatened species listing and recovery priority guidelines. *Federal Register* 43(184): 43098-43105.

Wells, J. V., and M. E. Richmond. 1995. Populations, metapopulations and species populations: what are they and who should care? *Wildlife Society Bulletin* 23:458-462.

Wise, B. P., and M. Henrion. 1986. A framework for comparing uncertain inference systems to probability. Pages 69-83 in L. N. Kanal and J. F. Lemmer, editors. *Uncertainty in artificial intelligence*. Elsevier Science Publishers, Amsterdam.

Appendix

Uncertainty Propagation

A fuzzy number can be considered a set of nested intervals $[a_{li}, a_{ri}]$ at each of many levels i from zero (bottom of the fuzzy number) to 1 (top). These levels correspond to the vertical axes in Figures 1 and 2. The subscripts l and r refer to left (lower) and right (upper) bounds of each interval, so that $a_{li} \leq a_{ri}$ for all i . The intervals are nested so that those below are included in those above: $a_{li} \leq a_{lj}$ and $a_{rj} \leq a_{ri}$ whenever $i \leq j$.

An ordinary scalar number is a special case of a fuzzy number for which $a_{li} = a_{ri}$ for all i . A fuzzy number is an interval if $a_{li} = a_l$ and $a_{ri} = a_r$ for all i . All the arithmetic operations we use on fuzzy numbers are defined as level-wise operations on the intervals that comprise the fuzzy numbers. In the definitions below, we omit the subscripted i s for the sake of simplicity, but that the interval operations are repeated at all levels to define the fuzzy operations.

Going from fuzzy numbers that represent particular parameters (Figs. 1 & 2) to a fuzzy number representing a risk category (Fig. 3) requires a logical system that includes comparison of the numerical value of these parameters (e.g., expression 2 in discussion of characterizing uncertainty). In the proposed system, numerical magnitude comparisons are defined according to the following four rules:

$$\begin{aligned}
 a < b &= \begin{cases} 1 & \text{if } a_r < b_l, \\ 0 & \text{if } b_r \leq a_l, \\ [0, 1] & \text{otherwise.} \end{cases} & a > b &= \begin{cases} 1 & \text{if } a_l > b_r, \\ 0 & \text{if } b_l \geq a_r, \\ [0, 1] & \text{otherwise.} \end{cases} \\
 a \leq b &= \begin{cases} 1 & \text{if } a_r \leq b_l, \\ 0 & \text{if } b_r < a_l, \\ [0, 1] & \text{otherwise.} \end{cases} & a \geq b &= \begin{cases} 1 & \text{if } a_l \geq b_r, \\ 0 & \text{if } b_l > a_r, \\ [0, 1] & \text{otherwise.} \end{cases}
 \end{aligned}$$

The IUCN criteria also utilize AND and OR operations (e.g., expression 3 in discussion of characterizing uncertainty). Many generalizations of the AND and OR operations have been proposed in the literature on uncertainty propagation, and fuzzy interpretations of the logical operations have been debated throughout the history of fuzzy logic (Klir and Yuan 1995). Originally, the min and max functions were proposed to model conjunction (AND) and disjunction (OR). This meant, for instance, that when a had the value 0.3 and b had the value 0.6, the conjunction (a AND b) would have the value 0.3. These functions generalize the traditional Boolean operators, but they are not the only possible functions that do. In the last three decades, many different operators have been suggested for fuzzy logic, although each is arbitrary in its own way. The difference among these operators reduces to an assumption each makes about the dependence between a and b .

As Wise and Henrion (1986) explain, the original min and max functions effectively assume that the operands are perfectly dependent. This is a rather strong assumption, yet it seems better than assuming independence among the quantities mentioned in the IUCN criteria, which we believe will rarely if ever be a tenable assumption. As an inclusive compromise, we follow Wise and Henrion and compute bounds on the conjunction and disjunction without making any assumption about the dependence between the operands except that it is positive. We use the operations

$$\begin{aligned}
 a \text{ AND } b &= \text{env}(a \times b, \min(a, b)), \\
 a \text{ OR } b &= \text{env}(\max(a, b), 1 - (1 - a) \times (1 - b)).
 \end{aligned}$$

In these two operations, there are three functions (env, min, max) and two operators ($-$ and \times). These are defined as follows:

$$\begin{aligned}
 \min(a, b) &= [\min(a_l, b_l), \min(a_r, b_r)], \\
 \max(a, b) &= [\max(a_l, b_l), \max(a_r, b_r)], \\
 \text{env}(a, b) &= [\min(a_l, b_l), \max(a_r, b_r)], \\
 a + b &= [a_l + b_l, a_r + b_r], \\
 a - b &= [a_l - b_r, a_r - b_l], \\
 a \times b &= [a_l \times b_l, a_r \times b_r].
 \end{aligned}$$

The a and b are two intervals, and the subscripts l and r refer to the left and right (lower and upper) ends of these intervals. The definitions are demonstrated as follows, with the following two fuzzy numbers as examples: $A = [0.35, 0.6, 0.7, 0.95]$ and $B = [0.15, 0.4, 0.525]$.

The result of the operation A or B is given in Fig. 7. The result is obtained by repeating the OR operation defined above for each level of the two fuzzy numbers A and B . Thus, there are many calculations, which correspond to levels of possibility. The following numerical example demonstrates the operations defined above for one level of these two fuzzy numbers. At possibility = 0.2, we have the intervals $a = [0.2, 0.5]$ and $b = [0.4, 0.9]$. With this level as an example, the operations min, max, $1 - a$, $a \times b$, a AND b , a OR b are as follows:

$$\begin{aligned}
 \min(a, b) &= [\min(0.2, 0.4), \min(0.5, 0.9)] = [0.2, 0.5] \\
 \max(a, b) &= [\max(0.2, 0.4), \max(0.5, 0.9)] = [0.4, 0.9] \\
 1 - a &= [1 - a_r, 1 - a_l] = [(1 - 0.5), (1 - 0.2)] = [0.5, 0.8] \\
 1 - b &= [1 - b_r, 1 - b_l] = [(1 - 0.9), (1 - 0.4)] = [0.1, 0.6] \\
 a \times b &= [a_l \times b_l, a_r \times b_r] = [0.2 \times 0.4, 0.5 \times 0.9] = [0.08, 0.45] \\
 a \text{ AND } b &= \text{env}(a \times b, \min(a, b)) \\
 &= \text{env}([0.08, 0.45], [0.2, 0.5]) \\
 &= [0.08, 0.5] \\
 a \text{ OR } b &= \text{env}(\max(a, b), 1 - (1 - a) \times (1 - b)) \\
 &= \text{env}([0.4, 0.9], 1 - ([0.5, 0.8] \times [0.1, 0.6])) \\
 &= \text{env}([0.4, 0.9], 1 - ([0.05, 0.48])) \\
 &= \text{env}([0.4, 0.9], [0.52, 0.95]) \\
 &= [0.4, 0.95].
 \end{aligned}$$

These definitions of AND and OR are implemented in the software used for the calculations in this paper (Akçakaya and Ferson 1999). They account for the fact that it is often unclear what kind of depen-

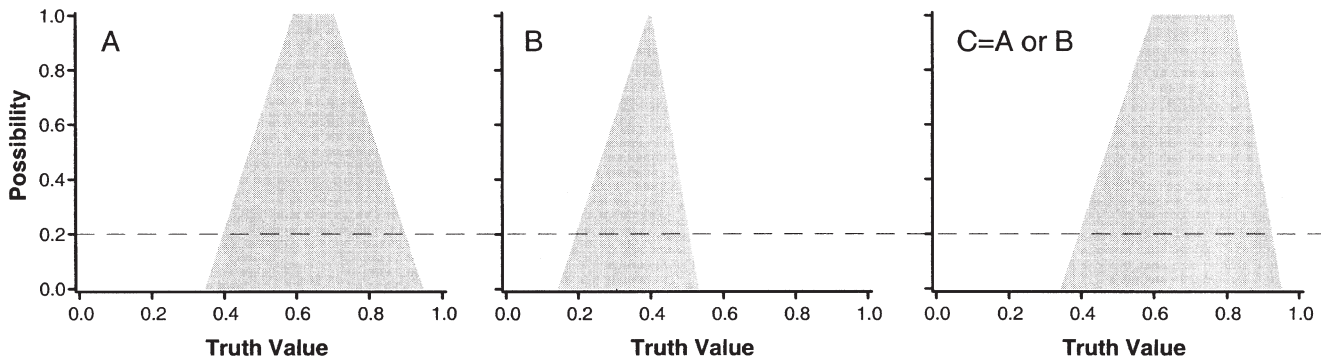


Figure 7. Demonstrations of the fuzzy arithmetic rule we used for disjunction (OR) operation: $A = [0.35, 0.6, 0.7, 0.95]$, $B = [0.15, 0.4, 0.525]$, $C = A \text{ or } B = [0.35, 0.6, 0.82, 0.97625]$. The numerical example in the text corresponds to the level indicated by the horizontal dashed line.

dence assumption should be made among the variables when empirical information is lacking. It would also be possible to compute bounds on the conjunction and disjunction that make no assumption

whatever about the dependence between a and b (Fréchet 1935; Hailperin 1986; Ferson et al. 1998). The resulting bounds would be roughly twice as wide as those we compute.

