INTRODUCTION

Just like habitat fragmentation, the concept of an ecological threshold may act as an intellectual attractor (sensu Haila 2002). The idea that ecological relationships may be characterized by narrow transition zones in which system properties change dramatically is very appealing from both an intellectual and a practical perspective. From a conservation perspective, such thresholds may represent danger points where the risk of species loss increases dramatically. For example, overharvesting of cod (Gadus morhua) is thought to have changed zooplankton predator-prey dynamics, which in turn negatively influences cod population dynamics (Casini et al. 2009). Hence, ecological thresholds may offer critical insight into ecosystem functioning or species-specific requirements. They also offer useful guidelines when developing evidence-based conservation targets (Sutherland et al. 2004, Svancara et al. 2005, Rondinini and Chiozza 2010, Samhouri et al. 2010), for which birds have often been used as a focal group (Villard and Jonsson 2009). Perhaps for these reasons, the number of papers on this topic has increased exponentially since 1980.

Along with the rapid growth in interest about ecological thresholds, the number of methods used to determine if and where thresholds occur has also burgeoned. For example, Harper and Macdonald (2011) and Toms (2012) summarize 24 different methods that have been used to identify ecological thresholds, and many additional methods exist (Andersen et al. 2009). The term “ecological threshold” has also been used to refer to a wide range of properties (Muradian 2001, Briske et al. 2006).

We focus on two broad classes of ecological thresholds that differ in their conceptual basis and statistical properties. Breakpoint-based thresholds are characterized by relationships between response and explanatory variable that change at some point, termed a breakpoint (often abrupt, although there can be a smooth transition region around this point; Fig. 1). In contrast, classification thresholds correspond to a value of the explanatory variable that divides the response variable into two relatively homogeneous groups. Classification thresholds may occur even if there is a gradual change in the response. For example, a threshold may separate ecological conditions where a species is likely to be present versus those where it is likely to be absent (Fig.

Threshold detection: matching statistical methodology to ecological questions and conservation planning objectives

Judith D. Toms1 and Marc-André Villard2

1Eco-Logic Consulting, 2Département de biologie, Université de Moncton

ABSTRACT. Two types of ecological thresholds are now being widely used to develop conservation targets: breakpoint-based thresholds represent tipping points where system properties change dramatically, whereas classification thresholds identify groups of data points with contrasting properties. Both breakpoint-based and classification thresholds are useful tools in evidence-based conservation. However, it is critical that the type of threshold to be estimated corresponds with the question of interest and that appropriate statistical procedures are used to determine its location. On the basis of their statistical properties, we recommend using piecewise regression methods to identify breakpoint-based thresholds and discriminant analysis or classification and regression trees to identify classification thresholds.

Détection des seuils : choix de la méthode statistique en fonction des questions écologiques et des objectifs de planification de la conservation

RÉSUMÉ. À l’heure actuelle, deux types de seuils écologiques sont communément utilisés dans l’élaboration d’objectifs de conservation : alors que les seuils correspondant à un point de rupture représentent le point de basculement où les propriétés du système changent radicalement, les seuils de classification, eux, séparent les données en groupes ayant des propriétés divergentes. Tant les seuils fondés sur le point de rupture que ceux fondés sur la classification sont des outils utiles pour la conservation issue d’observations. Toutefois, le type de seuil que nous cherchons à estimer doit impérativement concorder avec la question qui nous préoccupe, tandis que la méthode statistique utilisée pour déterminer la localisation précise du seuil doit être appropriée. En nous fondant sur leurs propriétés statistiques, nous recommandons l’utilisation d’une méthode de régression par segments pour déterminer un seuil correspondant à un point de rupture et d’une analyse discriminante ou d’un arbre de classification ou de régression pour déterminer un seuil de classification.

Key Words: breakpoint-based threshold; classification and regression tree; classification threshold; ecological thresholds; logistic regression; occupancy threshold; piecewise regression; ROC analysis
Fig. 1. Simulated examples of different types of thresholds. Breakpoint-based ecological thresholds can be either continuous (a – b) or discontinuous (c – d); classification thresholds are typically estimated for presence – absence data (e – f), but can also be estimated for abundance data (g – h). In addition, all types of thresholds may be considered to occur at a single point (left column), or may exhibit a transition zone (right column).

1). Although classification thresholds are typically applied to presence / absence data, they can also be estimated for count or continuous responses.

Even though they are conceptually quite distinct, these two types of thresholds are often treated interchangeably in the literature (Guénette and Villard 2004, King et al. 2007, Cuffney et al. 2011, Daily et al. 2012, Baker and King 2013). This can sometimes result in inappropriate statistical methods being used to estimate a particular type of ecological threshold. For example, in a series of simulations, Ficetola and Denoël (2009) tested whether logistic regression could identify a breakpoint-based threshold even though standard logistic regression can only be used to identify classification thresholds. Similarly, Baker and King (2010) implied that their approach could identify breakpoint-based thresholds, and was therefore tested for its ability to do so by Cuffney and Qian (2013), although it actually identifies a community-level analogue of a classification threshold. The two types of thresholds will tend to differ in location because they must be identified using different statistical methods. Such confusion and application of invalid statistical methods could easily lead to poor outcomes, e.g., if used for conservation planning. Because ecological thresholds are more commonly being used for this purpose, it is timely to review the differences between breakpoint-based and classification thresholds and outline statistical methods that are appropriate for estimating each.

**BREAKPOINT-BASED THRESHOLDS VERSUS CLASSIFICATION THRESHOLDS**

In statistical terms, the usual definition of an ecological threshold is that of a breakpoint-based threshold: the relationship between the response and explanatory variables is not linear, but changes at some point, the threshold (Muradian 2001). This change is often assumed to be a change in slope, so the relationship is
continuous and piecewise linear (Toms and Lesperance 2003; Fig. 1a – b). However, the change can also be discontinuous, such as a step function (Fig. 1c – d), and the relationships on either side of the threshold need not be restricted to linear forms. In addition, the threshold need not change at a single point, but could instead occur over a (relatively) narrow transition zone (Fig. 1b and d). This transition zone might reflect measurement error or variability in the phenomenon being measured, or could be an inherent property of the threshold relationship. Distinguishing between these alternatives may be impossible in most situations, but identifying the width of the transition zone is likely more important for conservation or management purposes than determining the reason why an ecological threshold has a zonal transition.

In contrast, a classification threshold does not necessarily involve a nonlinear relationship between the response and explanatory variables, but instead splits the response variable into two groups. The classification threshold is then defined as the value of the explanatory variable corresponding to the optimal division between these groups. In its most common usage with binary responses, the two groups correspond to portions in an ecological gradient where the species is likely to be present versus those where it is likely to be absent (Figure 1e – f). For count data, we might instead identify a threshold in abundance that splits the data into groups where the species is common versus uncommon or absent (Figure 1g – h).

Ficetola and Dinoël (2009) proposed the term “abrupt” threshold for a breakpoint-based threshold. We avoid this terminology because it can be confusing when there is a smooth transition around the breakpoint. Similarly, they use the term “smooth” threshold to refer to a classification threshold, but this suffers from similar problems because transition zones are not always present. The term classification threshold is also applicable to different types of response variables (e.g., presence / absence, continuous and count data) and commonly used in the statistical literature.

Both breakpoint-based thresholds and classification thresholds can be insightful for management and conservation purposes. Breakpoint-based thresholds represent “tipping points,” i.e., a value along a gradient where the focal ecosystem property changes dramatically. For example, when nutrient levels exceed critical thresholds, freshwater lakes and coral reef ecosystems enter eutrophied conditions (Knowlton 1992, Scheffer et al. 1993, 2001). Classification thresholds, in contrast, identify values of the explanatory variable corresponding to different groups of the response. For example, Betts and Villard (2009) determined the amount of mature forest at the scale of landscapes that was associated with the presence or absence of several songbird species.

It is critically important to decide which type of threshold corresponds to a particular question of conservation concern so that the correct statistical methods can be used to estimate its location. Breakpoint-based thresholds apply in cases where a change in the relationship between the response and explanatory variable is of concern: For example, does the relationship between distance from the habitat edge and density of desert riparian birds change at some distance (Brand et al. 2006)? Classification thresholds apply in cases where you want to partition the response variable based on the explanatory variable: For example, does the occurrence of Bobolinks (Dolichonyx oryzivorus) within a landscape depend on habitat openness (Keyel et al. 2012)?

Although the two types of thresholds may happen to coincide, they more commonly will not (e.g., Betts and Villard 2009, Ficetola and Dinoël 2009) because they are measuring different properties. For example, consider the relationship between the abundance of a species, e.g., a woodpecker, and the amount of a resource, e.g., dead wood, a critical foraging substrate. A breakpoint-based threshold would be expected if woodpecker territory size was related to the amount of available dead wood; then woodpecker density would decline linearly with the density of dead wood available, until dead wood was so rare that the territory size needed to maintain an individual woodpecker or breeding pair exceeded a size where movement costs are lower than energetic gains (Rolstad 1991, Bütler et al. 2004a, Warren et al. 2005). In contrast, a classification threshold would help managers identify ranges of dead wood amounts corresponding to a high frequency of occurrence of the focal woodpecker species (e.g., Bütler et al. 2004b, Roberge et al. 2008, Müller et al. 2009, Müller and Bütler 2010).

DETECTING BREAKPOINT-BASED THRESHOLDS

The simplest method consists of visually estimating the location of the threshold from plotted data, perhaps using a smoother or generalized additive model to clarify the trends in noisy data. Visual estimation can be surprisingly accurate, especially when the threshold is pronounced (Ficetola and Dinoël 2009), but this is not always the case (Swift and Hannon 2010). Other tests compare response values along a gradient of values of the explanatory variable (Qian et al. 2003, Zeileis et al. 2003, Andersen et al. 2009, Harper and Macdonald 2011), typically indicating a range of probable values rather than estimating the location of the threshold directly. Because these methods cannot estimate precision in the threshold location, their usefulness for practical applications is limited.

Therefore, we believe that models that directly incorporate a parameter specifying the location of the threshold are preferable, if the model assumptions can be met. In the case of a continuous threshold, i.e., a threshold where the two regression segments join at a point, piecewise regression (Toms and Lesperance 2003) has been proven to be the most efficient method of statistical inference (Chen et al. 2011). The simplest form of the piecewise linear model joins two straight lines at the threshold, whereas other forms incorporate a smooth zonal transition around the threshold and may be easier to fit if there is limited data near the threshold or if the estimated threshold converges on one of the observed data points (Toms and Lesperance 2003, Toms 2012). Piecewise regression can also be used if the threshold is discontinuous, i.e., follows a step function, by first transforming the data using the cumulative sum (Muggeo and Adelfio 2011). Of great practical importance, it is possible to construct confidence intervals for the estimated threshold in this model (Toms and Lesperance 2003).

Piecewise regression models are fitted using software to fit nonlinear models, or using other specialized software such as the segmented package for R (Muggeo 2008). Fitting a series of models with thresholds fixed at points in a plausible range of
values and then using AIC or other methods to choose the best fitting model, although common, should be avoided because this approach does not produce a confidence interval for the threshold location and inflates the Type I error rate (Benedetti et al. 2009). Instead, statistical tests can be used to determine if a piecewise regression model fits significantly better than a linear model (Liu and Qian 2009). Model estimates will not converge if a threshold does not exist, and so errors of nonconvergence are likely meaningful if the initial parameter estimates are good. In truth, the fact that a piecewise regression model will not estimate a threshold if none exists is a strength of this method.

The piecewise regression model can also be extended to binary and count data with generalized linear models, producing what may be termed piecewise logistic, also called segmented logistic, or piecewise Poisson models. For example, Zuckerberg and Porter (2010) used piecewise logistic models to look for thresholds in occurrence, colonization, extinction, and persistence of 25 forest bird species over a 20-year period. Bayesian versions of piecewise regressions have also been developed (Beckage et al. 2007), as have models that incorporate thresholds into time-series models (Muggeo 2008b). It is also possible to construct piecewise nonlinear models, or models that include more than one threshold. If multiple thresholds are suspected, an algorithm developed by Muggeo and Adelfio (2011) may be useful in determining the number and location of multiple thresholds.

**DETECTING CLASSIFICATION THRESHOLDS**

Classification thresholds have been most commonly estimated for binary data, where they are sometimes called occurrence thresholds (Guénette and Villard 2004, 2005). In this context, a logistic regression model is typically fit to the data. The data are then split into two groups (points where the species is likely to be present, versus points where it is likely to be absent) using a specified or estimated probability of occurrence. The fitted logistic regression function is then used to find the value of the explanatory variable corresponding to this probability of occurrence, the classification threshold. Some studies have used an arbitrary probability of occurrence, such as 50% (e.g., Bisson and Stutchbury 2000, Reunanen et al. 2002). Alternatively, a value that is optimal in some sense can be derived from the receiver operating characteristic (ROC) curve, i.e., the value that minimizes the sum of false positives and false negatives (Guénette and Villard 2004, 2005). Classification thresholds identified using ROC analysis are remarkably stable as false negatives and false positives are randomly added to a perfect binary response, even though the associated logistic regression curve quickly loses its “S” shape (Villard 2009). Choosing the probability of occurrence to be equal to the observed proportion of occurrences (Cramer 1999) is a simpler approach that appears to be at least as effective (Li et al. 2005). However, the sensitivity of this method to the rate of false negatives and false positives has not been determined.

In this context, the logistic regression model described above effectively acts as a logistic discriminant function that uses the explanatory variable to classify the points into two groups, rather than a regression model per se, and other methods of discriminant analysis could also be used for this purpose. One flexible alternative are classification and regression trees (De’ath and Fabricius 2000, Müller and Bütler 2010), which can be used with binary, count, or continuous data. Newer variants of these trees, such as random forests and bagged or boosted trees, typically have reduced bias and variation compared to simple classification and regression trees (Bauer and Kohavi 1999, Prasad et al. 2006, Cutler et al. 2007, Elith et al. 2008), although simple trees should produce similar results in the case of a single explanatory variable. A similar approach has been developed for analyzing community composition data, which combines multivariate regression trees and indicator species analysis (Baker and King 2010).

**DISCUSSION**

Conservation planners and land managers regularly ask ecologists to provide concrete advice to help them develop meaningful strategies to protect biodiversity. Both breakpoint-based and classification thresholds are useful tools in evidence-based conservation (Sutherland et al. 2004, Svancara et al. 2005, Rondinini and Chiozza 2010). Thresholds in response variables may be expected when the ecological gradient reflects the amount of a critical ecological resource, e.g., the amount of live or dead wood for a woodpecker or the area of suitable habitat within a certain landscape context (Jansson and Angelstam 1999, Betts et al. 2007), or the negative impacts of anthropogenic changes, e.g., the abundance of Sprague’s Pipit (Anthus spragueii) is reduced near roads or crops (Koper et al. 2009). Taken in isolation, individual variables may help predict species response, or perhaps species richness for a specific guild, but do not cover all the critical resources required for this or these species. In addition, they do not necessarily take into account the complex relationships underlying species responses to habitat amount. For example, the amount of suitable habitat in the landscape may increase the probability of occupancy of a given species at a given location (e.g., Venier and Fihrig 1996), but this relationship can be complicated by the parallel response of a variety of other sympatric species acting as predators, parasites, or competitors (Evans 2004). Hence, in most ecological systems, a noisy relationship will be the rule rather than the exception. If so, abrupt shifts in the rate of change along an ecological gradient can be difficult to detect, even if suitable methods are used.

However, there are many examples where ecological thresholds have been successfully identified. Unfortunately, the wide range of methods available and numerous definitions of ecological thresholds have led to confusion in some instances. Estimated thresholds can be misleading and potentially dangerous when used to define management objectives if arbitrary, subjective, or statistically inappropriate methods are used (Guénette and Villard 2004, Lindenmayer et al. 2005, Betts and Villard 2009, Ficetola and Denoël 2010). For instance, as we have shown, the distinction between breakpoint-based and classification thresholds must be clearly understood when deciding which to estimate, and appropriate methods must be used to identify each type of threshold.

Once a threshold is identified, developing conservation targets requires consideration of additional factors. First, thresholds are, by definition, values where the response, e.g., species abundance or occupancy, rapidly changes into a less desirable
state. However, ecological degradation can occur before the threshold is reached, and a single threshold value may not be valid across all landscapes (Lindenmayer and Luck 2005, Johnson 2013, van der Hoek et al. 2013). Therefore, conservation targets should include a safe buffer away from these critical values (Bennett and Radford 2009). Second, we rarely manage for one species in isolation, and different species are likely to respond at different threshold values. Therefore, conservation targets should ideally be based on the responses of multiple species (Guénette and Villard 2005, Lindenmayer and Luck 2005, Betts and Villard 2009, Müller and Bütler 2010, Zuckerberg and Porter 2010, Johnson 2013), with due consideration of potential interactions, e.g., predation or competition, among species whenever possible. When managing for suites of species, it may also be useful to look for thresholds in community composition (e.g., Toms and Lesperance 2003, Radford et al. 2005, Baker and King 2010, Müller and Bütler 2010), which are indicative of substantial change in multiple species. Third, the location of ecological thresholds may depend on the time scale examined, e.g., the threshold for long-term persistence in a given habitat may be greater than the threshold for short-term occurrence (Zuckerberg and Porter 2010, van der Hoek et al. 2013).

Throughout the world, ecosystems are rapidly being altered through the intensification of anthropogenic land use (Lindenmayer et al. 2012). Managing individual species and ecosystems in the face of such change is challenging, and the wise planning of land use requires broad ecological knowledge informed by specific conservation targets. Empirical thresholds represent an interesting avenue to influence the planning process through the identification of critical levels of required resources and ecological pressures (Angelstam et al. 2004, Villard and Jonsson 2009, Samhouri et al. 2010, Johnson 2013). However, such empirical targets are only useful if appropriate statistical methods are used and the nature of the threshold, i.e., breakpoint-based versus classification threshold, is clearly understood.

**Responses to this article can be read online at:**

---

**Acknowledgments:**

We thank two anonymous reviewers for their helpful comments on the manuscript.

**LITERATURE CITED**


Bütler, R., P. Angelstam, P. Eklund, and R. Schlaepfer. 2004b. Dead wood threshold values for the Three-toed Woodpecker...


