



Editorial

Biostatistics and the Editorial Process: a Caution against Fad Reviewing

Biostatistique et évaluation des manuscrits scientifiques : mise en garde contre les modes

Marc-André Villard and Thomas D. Nudds

In ecology and conservation biology, statistical analysis is essential for reliable inferences about cause-and-effect relationships. It is so essential that it is virtually impossible, and rightly so, to conduct and publish research in the absence of statistical inference. That said, there can be many different views about the appropriateness of particular statistical methods, even among statisticians. In such cases, during the review process, researchers can feel caught between reviewers from different schools of thought.

Perhaps many of us approach adoption of new statistical methods the same way we might approach adoption of new technology: long enough after it is first introduced, so as to have better assurance that the bugs are worked out, but once adopted, not so long as to be superseded by better technology. To exacerbate the situation, particularly for biologists who have a fear of numbers, the rate of proliferation of new, often more sophisticated statistical approaches seems only to accelerate once graduate school is over. Is the solution increasingly to turn to experts to analyze data? Is it even necessarily the case that we should take a Manichean view of statistical methods, that is, good or bad with no shades of grey?

In some university departments, statistical training, or even outright analysis as often observed in government agencies, is increasingly delegated to specialists. As a result, students and supervisors, and professional biologists and statisticians, may have different views on the statistical methods

appropriate to their goals. Are we facing a divorce between ecology and biostatistics, even in the academy, such that ecologists rely increasingly on a small number of statistical experts?

Although not on the same scale as “paradigm-swinging” debates about the relative value of frequentist versus Bayesian inference (e.g., Ellison 2004), or null hypothesis testing versus information-theoretic approaches for statistical inference (e.g., Stephens et al. 2005, Lukacs et al. 2007), another caution about being quick to jump to applications of particular approaches may be in order in the case of detectability.

Occupancy models that rely on reasonably readily collected presence-absence data have come to play a major role in ecology and conservation, and a recent debate about the best approach to account for variation in species detectability has sparked renewed activity in this area. Was a species actually absent from a given sampling unit during a specific survey period or did it simply go undetected? Obviously, the failure to detect species when they are actually present (false negatives) has major implications for habitat modeling and, eventually, conservation strategies. Assuming that a species occupies a site and simply fails to be detected during some surveys can also have implications if it turns out that it was actually absent! Detectability modeling (MacKenzie et al. 2006) is based on the assumption that a sampling unit that was occupied in at least one of repeated surveys was actually occupied, though the species was not detected in all



Sponsored by the Society of
Canadian Ornithologists and
Bird Studies Canada

*Parrainée par la Société des
ornithologistes du Canada et
Études d'oiseaux Canada*



**BIRD STUDIES
ÉTUDES D'OISEAUX CANADA**

other surveys (the “closure assumption”). Many factors may influence species detectability, some of which pertain to the probability that observers will see or hear individuals (Schieck 1997, MacKenzie et al. 2006) whereas others reflect movements of individuals in and out of the sampled unit, e.g., within-season dispersal after breeding failure (Jackson et al. 1989, Betts et al. 2008).

Through the careful analysis of occupancy patterns in two study areas, as well as simulated data, Rota et al. (2009) estimated that between 71 and 100% of species violated the closure assumption in surveys separated by eight days and three weeks, respectively. Apparent extinction or colonization events were best explained by changes in occupancy rather than by variable detectability. In a first case study, the occurrence of only 2 of 28 species was best predicted by “closed” models assuming site-visit-specific detection probability. The closure assumption was rejected for all 18 species in the other case study. Hence, applying detectability estimates without adequate consideration of violation of the closure assumption could reduce the performance of occupancy models for species and study systems where the assumption is routinely violated. This may be especially problematic in the case of species at risk, where violations of this assumption generally lead to an overestimation of the probability of occurrence (Rota et al. 2009).

Many papers have been rejected in the past several years because they did not account for variable detectability among species. However, as it turns out, broadly applying this approach to data sets without testing the validity of the closure assumption may actually have introduced a new type of uncertainty.

As in science generally, statistical techniques are in constant and rapid evolution. Hence, it should not be surprising to witness ebbs and flows of enthusiasm as new approaches are tested and gradually refined. For referees and editors, the lesson from the detectability example, as with other discussions about the merits of different approaches, is to be mindful of the larger context when requesting application of the latest statistical approaches, and not simply advise rejection on the basis that the latest technique is not in evidence. Just like new theories in other fields, time is required to judge merit which, in turn, may be conditional. New statistical tools emerge for good reasons; they are not fads. However, because those reasons may not

apply to all study systems, referees and editors have a duty not to simply submit to fad reviews.

This editorial marks our last contribution as Coeditors-in-Chief of ACE-ÉCO. We welcome the newly appointed Coeditors, Keith Hobson and Ryan Norris, who will take over our duties in 2012. The foundation of this journal has been an exciting adventure and with our recent inclusion into Thomson Reuters' Web of Science, we feel that we are leaving our young journal on solid footing.

Dans le domaine de l'écologie et de la biologie de conservation, les analyses statistiques sont essentielles pour tirer des conclusions fiables sur les relations de cause à effet. Leur importance est telle qu'il est pratiquement impossible, et avec raison, d'effectuer et de publier une étude sans inférence statistique. Cela dit, différents points de vue sur la pertinence de certaines méthodes statistiques peuvent se confronter, même parmi les statisticiens. Dans ces cas, durant le processus d'évaluation des manuscrits, les chercheurs peuvent se sentir pris entre des examinateurs provenant de différentes écoles de pensée.

Il est possible que bon nombre d'entre nous approchions l'adoption d'une nouvelle méthode statistique de la même façon que nous approchons l'adoption d'une nouvelle technologie : suffisamment longtemps après qu'elle a été introduite, de façon à ce que nous ayons l'assurance que les problèmes ont été résolus, mais pas trop longtemps pour qu'elle soit alors supplantée par une meilleure technologie. Cette situation est exacerbée, particulièrement pour les biologistes qui ont une peur bleue des chiffres, puisque le taux de prolifération des nouvelles approches statistiques, souvent plus sophistiquées, semble indubitablement s'accélérer après que les études graduées sont complétées. Est-ce que l'option, de plus en plus courante, de se tourner vers les experts pour analyser les données est la solution? Est-il même nécessaire d'adopter une perspective manichéenne des méthodes statistiques, soit bonnes ou mauvaises, sans nuance?

Dans certains départements universitaires, la formation statistique des étudiants – voire l'analyse statistique proprement dite, comme c'est souvent le cas dans les organismes gouvernementaux – est de plus en plus confiée aux spécialistes. Il en résulte

que les étudiants et les professeurs, ainsi que les biologistes et les statisticiens professionnels, peuvent avoir des points de vue différents sur les méthodes statistiques appropriées pour les objectifs visés. Faisons-nous face à un divorce entre l'écologie et la biostatistique, même dans le milieu universitaire, de sorte que les écologistes dépendent de plus en plus d'un petit nombre d'experts en statistique?

Bien que le débat récent sur la détectabilité ne soit pas de l'ordre de ceux opposant les paradigmes de l'inférence fréquentiste et celle de l'inférence bayésienne (p. ex. Ellison 2004), ou encore les approches du test de l'hypothèse nulle et de la théorie de l'information en inférence statistique (p. ex. Stephens et al. 2005, Lukacs et al. 2007), il mérite une mise en garde à l'intention de ceux qui sont prompts à appliquer des approches particulières.

Les modèles de présence qui reposent sur des données de présence-absence relativement faciles à récolter jouent maintenant un rôle important en écologie et en conservation, et un débat récent portant sur la meilleure approche pour tenir compte de la détectabilité des espèces a provoqué un regain d'intérêt dans ce domaine. Une espèce était-elle réellement absente d'une parcelle d'échantillonnage au moment précis de son inventaire ou n'a-t-elle tout simplement pas été détectée? Évidemment, l'incapacité de détecter une espèce alors qu'elle est pourtant présente (faux négatifs) a des répercussions majeures sur la modélisation de l'habitat et, éventuellement, sur les stratégies de conservation. La supposition qu'une espèce est présente dans un site mais qu'on ne l'a tout simplement pas détectée durant un inventaire peut aussi avoir des répercussions si cette espèce était réellement absente! La modélisation de la détectabilité (MacKenzie et al. 2006) est fondée sur la prémisse qu'une parcelle d'échantillonnage occupée lors d'au moins une visite parmi une série d'inventaires répétés était réellement occupée, même si l'espèce n'a pas été détectée à chacun de ces inventaires (la condition « de fermeture »). Parmi les nombreux facteurs qui peuvent influencer la détectabilité d'une espèce, certains sont liés à la probabilité que les observateurs puissent voir ou entendre les individus (Schieck 1997, MacKenzie et al. 2006) tandis que d'autres se rapportent au déplacement des individus à l'intérieur ou à l'extérieur de la parcelle d'échantillonnage, par exemple la dispersion intrasaisonnière suite à l'échec de la nidification (Jackson et al. 1989, Betts et al. 2008).

Grâce à une analyse rigoureuse des profils de présence dans deux aires d'étude, ainsi qu'à des données simulées, Rota et al. (2009) ont estimé qu'entre 71 et 100 % des espèces n'ont pas respecté la condition de fermeture dans le cas d'inventaires séparés par huit jours et par trois semaines, respectivement. Les cas apparents de disparition ou de colonisation étaient mieux expliqués par des changements dans la présence que par une détectabilité variable. Dans un premier cas d'étude, seule la présence de 2 des 28 espèces était mieux expliquée par des modèles « fermés » supposant une probabilité de détection relative à la visite d'un site spécifique. La condition de fermeture a été réfutée pour l'ensemble des 18 espèces du second cas d'étude. Ainsi, l'utilisation d'estimations de la détectabilité sans considérer la réfutabilité de la condition de fermeture peut diminuer la performance des modèles de présence chez les espèces et les systèmes d'étude pour lesquels la condition de fermeture est couramment réfutée. Cette pratique peut être particulièrement problématique dans les cas d'espèces en péril, pour lesquelles la réfutabilité de la condition conduit généralement à une surestimation de la probabilité de présence (Rota et al. 2009).

De nombreux articles ont été refusés ces dernières années parce qu'ils ne tenaient pas compte de la détectabilité variable des espèces. Toutefois, en fin de compte, l'application élargie de cette approche aux données sans tester la validité de la condition de fermeture a peut-être introduit un nouveau type d'incertitude.

Tout comme en science en général, les techniques statistiques évoluent rapidement et constamment. Par conséquent, il ne devrait pas être surprenant de constater qu'il y a des hauts et des bas d'enthousiasme à mesure que les nouvelles approches sont testées et progressivement améliorées. Lorsque vient le temps de demander aux chercheurs d'utiliser les démarches statistiques les plus récentes, les examinateurs et les rédacteurs en chef devraient retenir la leçon tirée de l'exemple de la détectabilité – tout comme celles tirées d'autres discussions sur les mérites des différentes approches – plutôt que de recommander le rejet de l'étude sur le simple fait qu'elle n'utilise pas la plus récente approche. Tout comme pour les nouvelles théories dans d'autres domaines, il faut du temps pour juger du mérite, qui peut être conditionnel. Les nouveaux outils statistiques surgissent pour de bonnes raisons; ils ne constituent pas que des modes passagers. Cependant, parce que ces raisons ne

s'appliquent pas nécessairement à tous les systèmes d'étude, les examinateurs et les rédacteurs en chef se doivent de ne pas simplement suivre les modes.

Cet éditorial marque notre dernière contribution comme co-rédacteurs en chef d'ACE-ÉCO. Nous souhaitons la bienvenue aux nouveaux co-rédacteurs, Keith Hobson et Ryan Norris, qui prendront le relais en 2012. La création de cette revue scientifique s'est avérée une aventure excitante et avec notre récent catalogage dans le « Web of Science » de Thomson Reuters, nous avons le sentiment de quitter notre jeune revue sur des bases solides.

RESPONSES TO THIS ARTICLE Responses to this article are invited. If accepted for publication, your response will be hyperlinked to the article. [To submit a response, follow this link.](#) [To read responses already accepted, follow this link.](#)

LITERATURE CITED

Betts, M. G., N. L. Rodenhouse, T. S. Sillett, P. J. Doran, and R. T. Holmes. 2008. Dynamic occupancy models reveal within-breeding season movement up a habitat quality gradient by a migratory songbird. *Ecography* 31:592-600. <http://dx.doi.org/10.1111/j.0906-7590.2008.05490.x>

Ellison, A. M. 2004. Bayesian inference in ecology. *Ecology Letters* 7:509-520. <http://dx.doi.org/10.1111/j.1461-0248.2004.00603.x>

Jackson, W. M., S. Rohwer, and V. Nolan. 1989. Within-season breeding dispersal in prairie warblers and other passerines. *Condor* 91:233-241. <http://dx.doi.org/10.2307/1368300>

Lukacs, P. M., W. L. Thompson, W. L. Kendall, W. R. Gould, P. F. Doherty, K. P. Burnham, and D. R. Anderson. 2007. Concerns regarding a call for pluralism of information theory and hypothesis testing. *Journal of Applied Ecology* 44:456-460. <http://dx.doi.org/10.1111/j.1365-2664.2006.01267.x>

MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2006. *Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence*. Elsevier, Amsterdam, The Netherlands.

Rota, C. T., R. J. Fletcher, R. M. Dorazio, and M. G. Betts. 2009. Occupancy estimation and the closure assumption. *Journal of Applied Ecology* 46:1173-1181. <http://dx.doi.org/10.1111/j.1365-2664.2009.01734.x>

Schieck, J. 1997. Biased detection of bird vocalizations affects comparisons of bird abundance among forested habitats. *Condor* 99:179-190. <http://dx.doi.org/10.2307/1370236>

Stephens, P. A., S. W. Buskirk, G. D. Hayward, and C. Martinez del Rio. 2005. Information theory and hypothesis testing: a call for pluralism. *Journal of Applied Ecology* 42:4-12. <http://dx.doi.org/10.1111/j.1365-2664.2005.01002.x>