Invited Paper

INFORMATION THEORY IN WILDLIFE SCIENCE: CRITIQUE AND VIEWPOINT

FRED S. GUTHERY,[1] Department of Forestry, Oklahoma State University, Stillwater, OK 74078, USA
LEONARD A. BRENNAN, Caesar Kleberg Wildlife Research Institute, Texas A&M University-Kingsville, Kingsville, TX 78363, USA
MARKUS J. PETERSON, Department of Wildlife and Fisheries Sciences, Texas A&M University, College Station, TX 77843, USA
JEFFREY J. LUSK, Department of Forestry and Natural Resources, Purdue University, West Lafayette, IN 47097, USA

Abstract: We question whether the growing popularity of model selection based on information theory (IT) and using the Akaike's Information Criterion (AIC) represent a useful paradigm shift in data analysis or a substitution of 1 statistical ritual for another, which leaves in place long-standing problems in wildlife science. We discuss the relevance of model selection in science, problems in the IT-AIC algorithm, errors of commission and omission in IT-AIC-based studies, and the role of IT-AIC in knowledge accrual. Model selection is just another minor tool in the grand panorama of science. The human mind, not statistical methods, produces scientific breakthroughs. Although IT-AIC might include elements of hypothetico-deductive science, it is arguably a form of sensitivity analysis, magnitude of effects estimation, or simple description as currently applied. Accordingly, it is largely an inductive approach to knowledge accrual and, therefore, subject to the pitfalls of induction. The algorithm tends to over fit data (i.e, use too many variables), resulting in models that contain useless variables and that generalize poorly. Errors of commission in IT-AIC-based papers include hopelessly uninformative lists of encrypted models and imposition of the model-selection approach on studies better executed in a simple, descriptive format. The major error of omission is an almost universal failure to test selected models on independent data. From our perspective, IT-AIC is a harmless human construct that is being ritualistically applied and therefore cannot be expected to correct long-standing problems in the conduct of wildlife science, such as failure to apply the hypothetico-deductive method. We view the growing application of IT-AIC as problematic because that growth might discourage use of the full panoply of available methods of inquiry. Accordingly, we urge colleagues to avail themselves of the rich pageant of available analytical techniques that can be applied in wildlife research under the hypothetico-deductive method and to keep ecology, rather than statistics, in the forefront of wildlife science.

JOURNAL OF WILDLIFE MANAGEMENT 69(2):457–465; 2005

Key words: AIC, hypothetico-deductive methods, inference, information–theoretic models, model selection, philosophy of wildlife science.

[B]ecause most ecologists are inexperienced mathematically and theoretically, they are ill equipped to judge the validity of theoretical publications. The small contingent of theoreticians becomes a "priesthood" that, as the sole provider of theoretical guidance, can lead the others astray through the use of rhetorical devices (Ginzburg and Jensen 2004:123).

Since the appearance of Burnham and Anderson (1998), use of an information-theory-based, model-selection algorithm has become popular in ecological and natural resource sciences. We refer to this algorithm as IT-AIC, implying not only the basic theory but also the protocols associated with its application (e.g., a priori specification of global and submodels; presentation of log likelihoods, ΔAIC values, and Akaike weights; etc.; Burnham and Anderson 1998, 2002). Articles relying on IT-AIC model selection are appearing with increasing frequency (Fig. 1). About 20% of articles in recent issues of *The Journal of Wildlife Management*, and 7% of articles in *Ecology* use IT-AIC analytical methods. Johnson and Omland (2004:105) consider this a positive development; they write, "Wildlife biologists and molecular systematists have been at the forefront of bringing model selection to ecology and evolution…"

We question whether growing application of IT-AIC represents a useful paradigm shift in wildlife science or a simple substitution of 1 data-analytic ritual (IT-AIC) for another (frequentist statistics). Popularity does not necessarily imply legitimacy; frequentist statistics were popular for decades despite inherent, logical shortcomings (e.g., null hypothesis a priori false) and human abuse of the logic (e.g., naked *P*-values used to assess strength of evidence). We fear that many researchers do not (1) appreciate the place of

[1] E-mail: forfsg@okstate.edu

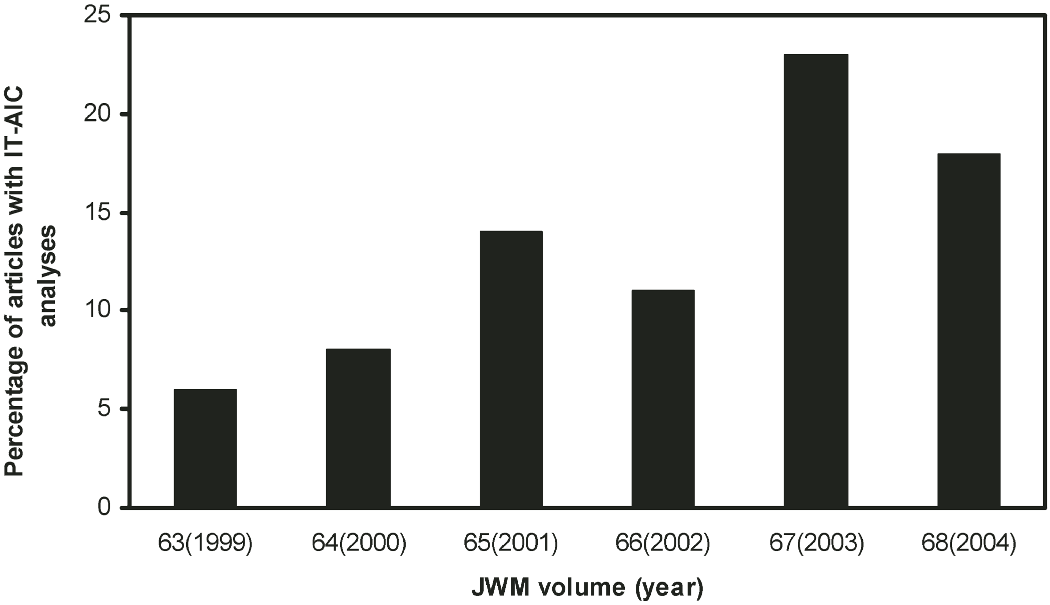

Fig. 1. Percentage of papers in *The Journal of Wildlife Management* (*JWM*) that used Information Theoretic-AIC analyses, 1999 (volume 63) through 2004 (volume 68).

model-selection exercises in the grand panorama of science, (2) recognize the dangers of rote application of IT-AIC, and (3) realize that viable and sometimes more appropriate alternatives to IT-AIC are available for use when research requires model selection. Moreover, we see errors of commission and omission in the research products of IT-AIC analyses that appear in the technical literature, despite published guidelines for avoiding some such errors (Anderson and Burnham 2002). Our perception is that growing popularity of a statistical tool might be smothering good practice in wildlife science, or, alternatively, preserving bad practice. Such behavior has ample precedent in human endeavor (Brothers 1997).

Our purpose is to evaluate and critique the role of IT-AIC in wildlife science. Specifically, we situate IT-AIC in the context of science, identify some of its known shortcomings, point out alternatives for model selection, discuss errors of omission and commission in published articles based on the IT-AIC algorithm, and evaluate the role of IT-AIC in knowledge accrual. We conclude with an appeal to keep ecology, not statistics, in the forefront of wildlife science.

PLACE IN SCIENCE

"Science … is the *organized, systematic enterprise that gathers knowledge about the world and condenses the knowledge with testable laws and principles*" (Wilson 1998:53, emphasis in original). The enterprise envisioned by Wilson is diverse, complex, and multifaceted. It includes description at microscopic to macroscopic scales and hypothesis formulation and testing within this range of scales. Model selection is but a minor, methodological tool of the great phenomenon, science; it certainly deserves no singling out as *the* way to gather and interpret information about nature. Breakthroughs in understanding that reverberate across natural resource science, and change paradigms with powerful new ideas, remain a product of the human mind, not of the tools used to analyze data (Guthery et al. 2001). Of course, IT-AIC might foster such breakthroughs, but given the way that the algorithm is typically applied, such instances will be rare. There are some examples in the literature where the IT-AIC approach was used to an advantage for analyses of complex, long-term data (Franklin et al. 2000, LaHaye et al. 2004) and meta-analyses (Noon and Franklin 2002, Franklin et al. 2004). However, such publications are the exception, rather than the rule, when it comes to typical applications of IT-AIC in wildlife research.

The place of a statistical tool in knowledge accrual depends on the nature of its application (e.g., simple description, magnitude of effects estimation, retroduction, induction, deduction).

As applied, IT-AIC might sometimes be weakly hypothetico-deductive (i.e., deduce outcomes from hypotheses and test for their occurrence; Romesburg 1981). Multinomial probability, logistic regression, multiple regression, and other types of models, which are a form of hypothesis (Guthery et al. 2004), ostensibly are advanced before the fact of data collection, and then they are evaluated as to plausibility (i.e., the models may be retroductions, or after-the-fact hypotheses, advanced to explain patterns in existing knowledge). Occasionally, researchers posit a model that explicitly tests previous research findings (Brennan et al. 1987). We view this practice as a legitimate and desirable form of inductive science (Maurer 2004) because previous findings should be challenged for validity.

The trouble is, the a priori models in most IT-AIC analyses often are too numerous (Anderson and Burnham 2002) to be credible as retroductive hypotheses. We have seen 50 models advanced with a sample size of 47 (Broders and Forbes 2004) in what could easily be interpreted as model dredging. Moreover, IT-AIC studies (e.g., Broders and Forbes 2004) often are justified on the basis of "little is known about…" (this is an omnibus though paltry justification for ecological studies, whether based on IT-AIC or some other approach). It is mysterious how one can generate a plethora of retroductively inspired models about something for which one claims ignorance. Eberhardt (2003) remarked on the difficulty of positing a suite of viable models for some research questions.

Model Selection or Sensitivity Analysis?

We also note that IT-AIC may be viewed either as a model-selection routine or a form of sensitivity analysis. Under the algorithm, one retroductively advances a global model based on existing knowledge, resulting in a presumably valid model if the existing knowledge is reliable. Then one pares and rearranges variables in the global model for apparently arbitrary reasons. We say the reasons are arbitrary because if the global model presumably is valid, why should it be pared? If a pared model presumably is valid, why would one advance a presumably invalid global model? Why not simply test the mettle of the pared model instead?

Accordingly, the pared "best" models make little sense in relation to the global model unless the pared "best" models are taken as indications that (1) the a priori knowledge used to construct the global model was reliable, but the data used to test it were faulty (i.e., not representative of the sampled population); (2) the data used to test the global model were reliable, but the a priori knowledge used to construct it was unreliable; or (3) the knowledge used to construct the global model and the data used to test it were reliable. In situation (1), we learned nothing from model selection and we possibly injected illegitimate findings into the literature. In situation (2), we made apparent progress by exposing false knowledge but we do not know whether we worked with faulty data. In situation (3), we did a sensitivity analysis and found that the global model contained weak predictors that may be eliminated. To the extent that situation (3) holds, pared "best" models are the residue of sensitivity analysis. The resulting models remain inductions about nature until tested; thus, they retain the epistemological weaknesses of induction in general (Popper 1959, Romesburg 1981, Howson 2000).

Alternatively, IT-AIC as sensitivity analysis could be viewed as magnitude of effects estimation. As Edwards (1992:2) observed, we often know that effects are real, and the research problem is estimation of their magnitudes. If researchers can advance a global model based on known relationships, we would seem to have an exercise in sensitivity analysis or magnitude of effects estimation. These are legitimate though non-creative modes of scientific endeavor.

In the final analysis, IT-AIC studies tend to be simple descriptive and observational. This statement is supported by the general failure of authors to set up deductions in objectives statements, and to justify models based on existing knowledge or hypothesized relationships. In fact, while authors might provide ecological rationales for inclusion of variables in the global model, they rarely provide ecological justification for subset models. The notion that these studies represent good science (Burnham and Anderson 2002:441) because of multiple, competing hypotheses (Chamberlin 1890) is somewhat of a sham, especially if you view IT-AIC as sensitivity analysis or magnitude of effects estimation.

Philosophical and Operational Weaknesses

Even if we assume that IT-AIC is other than sensitivity analysis, the IT-AIC algorithm remains subject to philosophical and operational weaknesses. Likelihood, the foundation of IT-AIC and a clear improvement over significance testing (Edwards 1992), is based on assumptions about probability distributions. Failure to meet the assumptions that underlie these techniques might render inferences suspect (Guthery et al. 2001). An IT-AIC "best" model is based on a supposedly optimal tradeoff between bias and variance, which involves just the right number of independent variables in a model. Burnham and Anderson (2002:31) call this optimal tradeoff the principle of parsimony; we call it the Goldilocks Principle because it is metaphysical (or mystical) to suppose that a condition (optimal tradeoff) arbitrarily imposed under statistical theory translates to ecological reality in the laboratory or field. We illustrate what we mean with a nonsequitur (i.e., a statement wherein the conclusion does not follow from the premise): "If we have optimized the bias-variance tradeoff according to the assumptions underlying statistical theory, *then we have come closer to ecological reality*." It does not follow as a logical result that optimal bias-variance tradeoff implies ecological reality. This proposition might be true in specific cases, but it certainly is not true in general. The point is that statistical assumptions serve the artificial world of statistical theory; real-world ecological processes operate largely independently of statistical assumptions and theory.

The statistical principle of parsimony is not the same as the one put forth by William of Ockham, a fourteenth-century monk (Silver 1998:169). The former is based on optimizing the bias-variance tradeoff, the latter (Ockham's Razor) on the metaphysical conjecture that a hypothesis with the lowest tally of assumptions is more likely to be true than alternative hypotheses. Although overparameterized models may be of no value (Ginzburg and Jensen 2004), and there is a modicum of logic in Ockham's Razor (Guthery 2004), counter examples to the razor are known: "Fifty years ago [ca. 1855] physicists considered, other things being equal, a simple law as more probable than a complicated law. But this belief is now repudiated; and yet, how many times are we compelled to act as though we still held it!" (Poincarè 1952:206, originally published in 1905). Contrary to popular opinion, the statistical principle of parsimony borrows no legitimacy from Ockham's Razor, which, in any case, is epistemically ambiguous.

Operationally, IT-AIC has certain troubling properties, most of which can be counteracted by researchers aware of their presence and implications. The algorithm sometimes leads to ambiguous and internally inconsistent, "plausible" models (ΔAIC < 2; Norman et al. 2004). The AIC tends to overestimate the number of parameters in a model (Kadane and Lazar 2004). These 2 problems arise in part from the structure of the information criterion or its various forms:

$$\mathrm{AIC} = -2\ln L + 2K\,,$$

where L = likelihood and K = the number of estimable parameters. Consider, as an example, a situation where 2 models differ by 1 predictor variable. The added variable could add absolutely nothing to the fit of the models so that $-2\ln L$ is the same for both models, and the penalty term for the additional variable is 2; thus, $\Delta\mathrm{AIC} = 2$ and the larger model is "plausible" despite containing a meaningless variable. It is possible under similar arguments to obtain $\Delta\mathrm{AIC} < 2$ with the addition of a useless variable.

Models with uninformative or useless variables tend to generalize poorly when exposed to new data (i.e., they produce poor approximations of nature). For example, Wellendorf et al. (2004) included a useless variable in an "Akaike best" logistic regression model. In a model predicting the probability that a northern bobwhite (*Colinus virginianus*) covey calls in the morning, the logistic regression coefficient for percent cloud cover (–0.002) had essentially no effect on model predictions because the coefficient was near 0.0.

The most profound shortcoming of the IT-AIC algorithm is that AIC and similar model-selection tools are relative statistics. This means that the algorithm will identify "best" or "plausible" model(s) in a set of outrageously bad models. It is, of course, incumbent upon authors to show whether "best" or "plausible" models have any ecological substance, but this does not always happen. Accordingly, Pygmalion models enter the published literature dressed up in "model-selection rigor" in much the same way the results of unnecessary or misleading tests of null hypothesis significance did previously (Cherry 1998, Johnson 1999, Guthery et al. 2001).

Alternatives to AIC

Given research that is solely a model-selection exercise, are there alternatives to IT-AIC? There certainly are alternatives to AIC and AIC_c that could be used within an IT framework. These include generalizations of AIC such as Takeuchi's Information Criterion (TIC; Takeuchi 1976) and the Network Information Criterion (NIC; Murata et al. 1994). There are also comparable Bayesian approaches to model selection, including the

Bayesian Information Criterion (BIC; Schwarz 1978) and Deviance Information Criterion (DIC; Spiegelhalter et al. 2002). The DIC, particularly, has considerable potential as an alternative to AIC for ecologists conducting model selection (Spiegelhalter et al. 2002, Ellison 2004). There are also other model-selection techniques commonly used with linear regression (e.g., Mallow's C_p). Clearly, statisticians working on model selection are not yet convinced that a single, universally "best" information criterion for model selection has been found (Spiegelhalter et al. 2002, Claeskens and Hjort 2003, Ellison 2004). For example, Ellison (2004:517) argued that, while the Focused Information Criterion (FIC; Claeskens and Hjort 2003) had certain promising characteristics, "it has not developed yet to the extent that it can be applied to even basic ecological problems." Taper (2004) argued that model selection criteria should allow for adaptive penalization for increased parameterization based on the size of the model set. Currently, applied ecologists are faced with an expanding array of information criteria for model selection, not with the narrowing of this field to a single best approach.

Statisticians will probably sort out the pros and cons of currently viable model-selection tools in time, but new tools undoubtedly will become available. How should wildlife scientists address this dizzying array of information criteria and other model-selection approaches? Kadane and Lazar (2004:286) observed, concerning the structure of various model-selection protocols, that, "all are matters of opinion on which conscientious statisticians and users of statistics can legitimately disagree without making a provable logical error." They also maintained that common sense could (and we believe should) prevail to some degree in model selection. In other words, get rid of models that predict poorly and select simpler, ecologically plausible models from among those that predict well. Curtis and Jensen (2004) provided a good example of the common-sense approach to model selection. We add, from the perspective of common sense, "Posit only ecologically plausible models so that you will not end up with ecologically implausible 'best' models after the IT-AIC routine."

ERRORS OF COMMISSION AND OMISSION

Thus far, we have blended the blemishes of IT-AIC per se with general problems in application of the algorithm. For example, the relative nature of AIC is a quirk of IT-AIC; the failure of ecologists to recognize and react to the troubling properties of IT-AIC is an issue of application. We now turn to more specific problems in the application of IT-AIC.

First, we see errors of commission. Some IT-AIC articles are hopelessly uninformative. These articles might consist of a long list of encrypted models, most of which are "implausible." In fact, it may be the first time in the history of scientific publication that authors are encouraged and, in some cases, required to publish information on bad or failed models.

In some cases, model-selection verbiage occupies numerous pages, whereas the only useful biological information might be a model-averaged estimate of some demographic variable. We suspect, therefore, that IT-AIC has increased the ratio of statistics to biology in the pages of ecological journals, which we view as unfortunate. We also encounter what could be construed as naked AIC models (i.e., models unadorned with biological information or interpretation). We see hypotheses in the form of model statements that match the null hypothesis in vacuity (e.g., approximate analogues of distance equals rate × time). Studies of the obvious (i.e., trivial hypotheses) can be justified in terms of the magnitude of obvious effects (Edwards 1992), and in our judgment, they should be. Finally, we surmise that IT-AIC has become so trendy that it is being forced on research projects like Procrustes forced his victims to fit into a bed by stretching or amputating their limbs. For example, projects best executed under a simple, descriptive format transmogrify into model-selection ritual. Griffin et al. (2003) provide an example of gratuitous use of IT-AIC for a simple, descriptive problem. This is not desirable because it unnecessarily complicates that which is fundamentally simple. Model selection is not necessarily relevant to every, or even most, ecological studies.

Second, we see 2 major errors of omission. The first is a general failure to fully interpret "best" or "plausible" models. Interpretation could involve actually using the "best" models (e.g., Curtis and Jensen 2004) to determine trends or relationships at different values of independent variables and to identify values of independent variables for which a "best" model generates absurd predictions, such as the presence of beavers (*Castor canadensis*) in the absence of water (Curtis and Jensen 2004). Such modeling also would assist in identifying useless variables in "best" models, thus testing the IT-AIC bias for over fit. After all, it is these biological and ecological relationships that should be of paramount importance to an ecologist.

The second omission, and the most serious one, is a general failure to test the validity of "best" models with independent data, although Wellendorf et al. (2004) provide a good counter example. Testing model validity is important because of the shortcomings of IT-AIC (i.e., relative nature, tendency to over fit and possibly select useless variables); these problems can be identified with independent data, or with data held back in model development, and used to evaluate "best" models. Because non-global "best" models appear to be inductive, such models are in clear need of testing; inductive findings remain forever tentative. Model testing is also important because any data set might be idiosyncratic and thereby lead to the development of models that do not hold when applied to new data. For example, "Akaike best" models suggest faulty data whenever the "best" model does not contain the trivially obvious, such as year effects on the annual survival of an *r*-selected species. Because of the possibility of faulty data, the results of randomized clinical trials in medicine are viewed with prejudice until evaluated in an independent study (Sackett et al. 2000). We readily acknowledge, however, that tests of model validity are generally lacking in even conventional regression analysis, sans AIC, but this does not impact our argument.

PARADIGM SHIFT OR RITUAL SUBSTITUTION?

At the outset, we questioned whether the growing application of IT-AIC represented a useful paradigm shift or the simple substitution of 1 data-analytic ritual for another in wildlife science. The algorithm certainly has changed the way many wildlife scientists approach data analysis, as evidenced by its growing prevalence in the wildlife literature (Fig. 1). This growing prevalence could be a boon to wildlife science if knowledge developed using the IT-AIC approach is somehow better than, or accrues at a faster rate than, knowledge acquired under alternative approaches to research. It could be a bane to wildlife science if it results in over reliance on a single analytical tool. In order to evaluate these questions, one must first consider the knowledge-development (epistemic) pathways important to wildlife science. After all, "the existence of multiple stories for the same 'facts' makes the process of being human inherently problematic. The human condition is that of being simultaneously, variably enmeshed in multiple social realities, each with its own logic of meaning and action" (Pearce 1989:21).

Epistemologists maintain that knowledge in late, modern human society, of which scientists are part, is arrived at through ≥2 approaches: post-positivist, and critical-constructivist (Hall 1990, Peat 2002, Peterson et al. 2002). Within a post-positivist perspective, knowledge is a set of non-falsified hypotheses (Popper 1959, 1962). The hypothetico-deductive method of science (Romesburg 1981) is a prime example of post-positivist thinking, and its merits related to the acquisition of reliable scientific knowledge are well known to wildlife scientists. Post-positivism, however, is far less suitable for introspective critiques of wildlife science or determining how best to approach social and political controversies associated with natural resource management (Jasanoff 1996, Peterson et al. 2002).

From a critical-constructivist perspective, knowledge is a societal construction about which intersubjective consensus is reached among individuals deemed competent (trusted) to interpret the substance of the construction for society (Jasanoff 1996, Schwandt 1998). Persons may be considered competent by the society in question for a myriad of reasons (e.g., experience, status as a scientist, wisdom, charisma). Much of the knowledge derived from case studies, which allow us to learn vicariously, is constructivist. Critical analysis contributes to knowledge through a series of structural or historical insights that transform as time progresses. This approach to knowledge makes use of dialectic approaches (i.e., discussion and reasoning by logical dialogue) as a method of intellectual investigation. It is exemplified by the Socratic tradition of exposing false beliefs and eliciting truth. Recent essays published in The Wildlife Society journals pointing out logical problems with analytical methods and offering alternatives, such as Cherry (1998), Johnson (1999, 2002*a*, *b*), Anderson et al. (2000, 2001), Anderson (2001, 2003), Guthery et al. (2001), and Eberhardt (2003), are informed by this tradition. Further, Peterson et al. (2002, 2004) critically analyzed qualitative data to suggest how wildlife managers might more effectively address conflicts associated with endangered species.

One danger of critical–constructivist approaches is that the individual or small cadre of individuals deemed competent to interpret the substance of the construction might lead society astray. This problem explains in part how hypotheses in wildlife science "gain credence and the status of laws through rhetoric, taste, authority, and verbal repetition" without benefit of sufficient testing (Romesburg 1981:295).

While both of these epistemic approaches can contribute to wildlife ecology and management, both can produce unreliable knowledge. The fact that alternatives to AIC (e.g., BIC, DIC, NIC, TIC) are rarely used by wildlife scientists—even though they often perform as well and are sometimes preferable (Burnham and Anderson 2002:268–351, Spiegelhalter et al. 2002, Kadane and Lazar 2004, Ellison 2004)—speaks volumes about how recent knowledge of analytical approaches in wildlife science has developed. Surely much of the reason these alternatives are uncommonly used is that biometricians and other applied statisticians who champion these approaches (e.g., Ripley 1996, Spiegelhalter et al. 2002, Ellison 2004) typically do not publish in outlets read by wildlife scientists. This fact alone, however, does not necessarily have anything to do with the merit of these model-selection procedures.

Hurlbert (1984:188) argued that, "it is the elementary principles of experimental design, not advanced or esoteric ones, which are most frequently and severely violated by ecologists." Neither AIC nor any other information criterion will protect researchers from continuing with the same general class of design and analysis mistakes and problems decried by Cherry (1998), Johnson (1999), and Anderson et al. (2000). None of these model-selection approaches necessarily prevent naked *P*-values from becoming naked models based on information criterion values (no biological explanation provided), vacuous null hypotheses from becoming vacuous model statements, or an obsession with inferential statistics from becoming an obsession with IT-AIC or some other approach. Description using any algorithm remains description. The same errors of commission and omission can prevail, and an obsession with analysis can supersede an obsession with ecology. More than 2 decades since Romesburg (1981), wildlife science still struggles to elevate the formation of meaningful research hypotheses and sound experimental designs in relation to the analytic procedures (Guthery et al. 2001). After all, as Hurlbert (1984:188–189) noted, of the 5 components of an experiment—research hypotheses, experimental design, experimental execution, statistical analysis, interpretation—"statistical analysis and interpretation are the least critical …, in that if purely statistical or interpretive errors are made, the data can be reanalyzed. On the other hand, the only complete remedy for design or execution errors is repetition of the experiment." No analytical technique can escape the "garbage-in-garbage-out" syndrome.

From our perspective, IT-AIC has not necessarily improved wildlife science, but this may be more an issue of its application (a human problem) than any faults inherent in the rather harmless algorithm. However, we find growing reliance on a single methodology problematic; just as the null hypothesis became "so catholic and ritualized as to seemingly impede clear thinking and alternative analysis approaches," (Anderson et al. 2001:315), IT-AIC is now similarly becoming catholic and ritualized. We concur with Stump (1992:459) that, "using only one method where several may apply amounts to a reduction and distorts" science by discouraging us from using other methods of inquiry. Thus we suggest that wildlife science should not only remain open to alternative model selection methodologies, but also "place logical argument, effect estimation, descriptive statistics, graphic approaches, qualitative analysis, and other methods in the realm of scientific orthodoxy" (Guthery et al. 2001:384). As wildlife ecologists, we understand the dangers of relying on overly reductionistic experimental designs; we must become equally aware of the dangers associated with reducing our analytic approaches to essentially a single method.

Finally, we urge wildlife scientists to keep ecology, not statistics, in the forefront of wildlife science. Statistics are messy tools we use because time and money constraints force insufficient sampling, and we must therefore seek logical support from theories of uncertainty. Regardless, ecology should remain paramount in our efforts because the measure of wildlife scientists is not their facility with statistical algorithms, but rather their contributions to knowledge, understanding, and innovation in the realm of ecology. This assertion is easily confirmed by juxtaposing phrases such as "niche gestalt," "fragmentation and metapopulations," or "the theory of natural selection" with phrases such as "*P*-value," "*t*-test," or "IT-AIC algorithm." Statistical algorithms, after all, are simply 1 set of means (tools) that can be used in our attempts to reach the end of reliable knowledge in wildlife ecology and management; they are not ends in and of themselves.

ACKNOWLEDGMENTS

We thank B. A. Collier, T. R. Peterson, S. D. Fuhlendorf, R. W. DeYoung, R. L. Bingham, and W. E. Palmer for commenting on earlier drafts of this manuscript. The Oklahoma Agricultural Experiment Station supported F. S. Guthery and approved our paper for publication. The Texas Agricultural Experiment Station and Texas A&M University System supported M. J. Peterson's contributions to this project. This is contribution number 05-108 from the Caesar Kleberg Wildlife Research Institute.

LITERATURE CITED

ANDERSON, D. R. 2001. The need to get the basics right in wildlife field studies. Wildlife Society Bulletin 29:1294–1297.

———. 2003. Response to Engeman: index values rarely constitute reliable information. Wildlife Society Bulletin 31:288–291.

———, AND K. P. BURNHAM. 2002. Avoiding pitfalls when using information–theoretic methods. Journal of Wildlife Management 66:912–918.

———, ———, W. R. GOULD, AND S. CHERRY. 2001. Concerns about finding effects that are actually spurious. Wildlife Society Bulletin 29:311–316.

———, ———, AND W. L. THOMPSON. 2000. Null hypothesis testing: problems, prevalence, and an alternative. Journal of Wildlife Management 64:912–923.

BRENNAN, L. A., W. M. BLOCK, AND R. J. GUTIÉRREZ. 1987. Habitat use by mountain quail in northern California. Condor 89:66–74.

BRODERS, H. G., AND G. J. FORBES. 2004. Interspecific and intersexual variation in roost-site selection of northern long-eared and little brown bats in the Greater Fundy National Park Ecosystem. Journal of Wildlife Management 68:602–610.

BROTHERS, L. 1997. Friday's footprint: how society shapes the human mind. Oxford University Press, New York, USA.

BURNHAM, K. P., AND D. R. ANDERSON. 1998. Model selection and inference. Springer, New York, USA.

———, AND ———. 2002. Model selection and multimodel inference. Springer, New York, USA.

CHAMBERLIN, T. C. 1890. The method of multiple working hypotheses. Science 148:754–759.

CHERRY, S. 1998. Statistical tests in publications of The Wildlife Society. Wildlife Society Bulletin 26:947–953.

CLAESKENS, G., AND N. L. HJORT. 2003. The focused information criterion. Journal of the American Statistical Association 98:900–916.

CURTIS, P. D., AND P. G. JENSEN. 2004. Habitat features affecting beaver occupancy along roadsides in New York state. Journal of Wildlife Management 68:278–287.

EBERHARDT, L. L. 2003. What should we do about hypothesis testing? Journal of Wildlife Management 67:241–247.

EDWARDS, A. W. F. 1992. Likelihood. Johns Hopkins University Press, Baltimore, Maryland, USA.

ELLISON, A. M. 2004. Bayesian inference in ecology. Ecology Letters 7:509–520.

FRANKLIN, A. B., D. R. ANDERSON, R. J. GUTIÉRREZ, AND K. B. BURNHAM. 2000. Climate, habitat quality, and fitness in northern spotted owl populations in northwestern California. Ecological Monographs 70:539–590.

———, R. J. GUTIÉRREZ, J. D. NICHOLS, M. E. SEAMANS, G. C. WHITE, G. S. ZIMMERMAN, J. E. HINES, T. E. MUNTON, W. S. LAHAYE, J. A. BLAKESLEY, G. N. STEIGER, B. R. NOON, D. W. H. SHAW, J. J. KEANE, T. L. MCDONALD, AND S. BRITTING. 2004. Population dynamics of the California spotted owl (*Strix occidentalis occidentalis*): a meta-analysis. Ornithological Monographs 54:1–54.

GINZBURG, L. R., AND C. X. J. JENSEN. 2004. Rules of thumb for judging ecological theories. Trends in Ecology and Evolution 19:121–126.

GRIFFIN, P. C., L. BIENEN, C. M. GILLIN, AND L. S. MILLS. 2003. Estimating pregnancy rates and litter size in snowshoe hares using ultrasound. Wildlife Society Bulletin 31:1066–1072.

GUTHERY, F. S. 2004. Commentary: the flavors and colors of facts in wildlife science. Wildlife Society Bulletin 32:288–297.

———, J. J. LUSK, AND M. J. PETERSON. 2001. The fall of the null hypothesis: liabilities and opportunities. Journal of Wildlife Management 65:379–384.

———, ———, AND ———. 2004. In my opinion: hypotheses in wildlife science. Wildlife Society Bulletin 32:1325–1332.

HALL, J. R. 1990. Epistemology and sociohistorical inquiry. Annual Review of Sociology 16:329–351.

HOWSON, C. 2000. Induction and the uniformity of nature. Pages 181–183 *in* W. H. Newton-Smith, editor. A companion to the philosophy of science. Blackwell Publishers, Malden, Massachusetts, USA.

HURLBERT, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Monographs 54:187–211.

JASANOFF, S. 1996. Beyond epistemology: relativism and engagement in the politics of science. Social Studies of Science 26:393–418.

JOHNSON, D. H. 1999. The insignificance of statistical significance testing. Journal of Wildlife Management 63:763–772.

———. 2002*a*. The importance of replication in wildlife research. Journal of Wildlife Management 66:919–932.

———. 2002*b*. The role of hypothesis testing in wildlife science. Journal of Wildlife Management 66:272–276.

JOHNSON, J. B., AND K. S. OMLAND. 2004. Model selection in ecology and evolution. Trends in Ecology and Evolution 19:101–108.

KADANE, J. B., AND N. A. LAZAR. 2004. Methods and criteria for model selection. Journal of the American Statistical Association 99:279–290.

LAHAYE, W. S., G. S. ZIMMERMAN, AND R. J. GUTIÉRREZ. 2004. Temporal variation in vital rates of an insular population of spotted owls (*Strix occidentalis occidentalis*): contrasting effects of weather. Auk 121:1056–1069.

MAURER, B. A. 2004. Models of scientific inquiry and statistical practice: implications for the structure of scientific knowledge. Pages 17–50 *in* M. L. Taper and S. R. Lele, editors. The nature of scientific evidence: statistical, philosophical, and empirical considerations. University of Chicago Press, Chicago, Illinois, USA.

MURATA, N., S. YOSHIZAWA, AND S. AMARI. 1994. Network information criterion—determining the number of hidden units for artificial neural network models. IEEE Transactions on Neural Networks 5:865–872.

NOON, B. R., AND A. B. FRANKLIN. 2002. Scientific research and the spotted owl: opportunities for major contributions to avian population demography. Auk 119:311–320.

NORMAN, G. W., M. M. CONNER, J. C. PACK, AND G. C. WHITE. 2004. Effects of fall hunting on survival of male wild turkeys in Virginia and West Virginia. Journal of Wildlife Management 68:393–404.

PEARCE, W. B. 1989. Communication and the human condition. Southern Illinois University Press, Carbondale, USA.

PEAT, D. F. 2002. From certainty to uncertainty: the story of science and ideas in the twentieth century. Joseph Henry Press, Washington, D.C., USA.

PETERSON, M. N., S. A. ALLISON, M. J. PETERSON, T. R. PETERSON, AND R. R. LOPEZ. 2004. A tale of two species: habitat conservation plans as bounded conflict. Journal of Wildlife Management 68:743–761.

———, T. R. PETERSON, M. J. PETERSON, R. R. LOPEZ, AND N. J. SILVY. 2002. Cultural conflict and the endangered Florida Key deer. Journal of Wildlife Management 66:947–968.

POINCARÉ, H. 1952. Science and hypothesis. Dover Publications, New York, USA.

POPPER, K. R. 1959. The logic of scientific discovery. Hutchinson, United Kingdom.

———. 1962. Conjectures and refutations: the growth of scientific knowledge. Basic Books, New York, USA.

RIPLEY, B. D. 1996. Pattern recognition and neural networks. Cambridge University Press, New York, USA.

ROMESBURG, H. C. 1981. Wildlife science: gaining reliable knowledge. Journal of Wildlife Management 45:293–313.

SACKETT, D. L., S. E. STRATUS, W. S. RICHARDSON, W. ROSENBERG, AND R. B. HAYNES. 2000. Evidence-based medicine. Churchill Livingston, New York, USA.

SCHWANDT, T. A. 1998. Constructivist, interpretivist approaches to human inquiry. Pages 221–259 *in* N. K. Denzin and Y. S. Lincoln, editors. The landscape of qualitative research: theories and issues. Sage Publications, Thousand Oaks, California, USA.

SCHWARZ, G. 1978. Estimating the dimension of a model. Annals of Statistics 6:461–464.

SILVER, B. L. 1998. The ascent of science. Oxford University Press, Oxford, United Kingdom.

SPIEGELHALTER, D. J., N. G. BEST, B. P. CARLIN, AND A. VAN DER LINDE. 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society Series B-Statistical Methodology 64:583–639.

STUMP, D. 1992. Naturalized philosophy of science with a plurality of methods. Philosophy of Science 59:456–460.

TAKEUCHI, K. 1976. Distribution of informational statistics and a criterion for model fitting. Suri-Kagaku 153:12–18. [In Japanese.]

TAPER, M. L. 2004. Model identification from many candidates. Pages 488–501 *in* M. L. Taper and S. R. Lele, editors. The nature of scientific evidence: statistical, philosophical, and empirical considerations. University of Chicago Press, Chicago, Illinois, USA.

WELLENDORF, S. D., W. E. PALMER, AND P. T. BROMLEY. 2004. Estimating calling rates of northern bobwhite coveys and measuring abundance. Journal of Wildlife Management 68:672–682.

WILSON, E. O. 1998. Consilience: the unity of knowledge. Alfred A. Knopf, New York, USA.