

## APPENDIX 1: Making sense out of Akaike's Information Criterion (AIC): its use and interpretation in model selection and inference from ecological data



# APPENDIX 1: Making sense out of Akaike's Information Criterion (AIC): its use and interpretation in model selection and inference from ecological data

### Table des matières

- [Abstract](#)
- [OUT WITH THE OLD?](#)
- [FIRST THINGS FIRST](#)
- [MINIMIZING THE LOSS OF INFORMATION](#)
- [MODEL SELECTION](#)
- [COMPARING MODELS](#)
- [AIC VS HO](#)
- [AN EXAMPLE](#)
- [WHEN SEVERAL MODELS COMPETE FOR FIRST PLACE: MULTIMODEL INFERENCE](#)
- [SPECIAL ISSUES WITH COUNT DATA](#)
- [ADVANTAGES AND LIMITATIONS OF AIC](#)
- [A FLASH IN THE PAN OR HERE TO STAY?](#)
- [ACKNOWLEDGEMENTS](#)
- [REFERENCES](#)

MARC J. MAZEROLLE

*Centre de recherche en biologie forestière, Pavillon Abitibi-Price, Faculté de Foresterie et de Géomatique, Université Laval, Québec, Québec G1K 7P4, Canada.*

E-mail: [marc-j.mazerolle.1@agora.ulaval.ca](mailto:marc-j.mazerolle.1@agora.ulaval.ca)

## Abstract

A large component of ecology consists of observational studies where the goal is to explain a pattern, such as the number of individuals in a patch, with a series of explanatory (i.e., independent) variables. To do so, ecologists have long relied on hypothesis testing to include or exclude variables in models, although the conclusions often depend on the approach used (e.g., forward, backward, stepwise). In the mid 1970's, the advent of methods based on information theory, also known as information-theoretic approaches, has changed the way we look at model selection and inference. A few decades later, measures of information, such as the Akaike information criteria (AIC) and

associated measures of model uncertainty, have begun to surface in the ecological disciplines. Though still underutilized, these approaches provide a new framework on which to base both model selection and inference from ecological data, and are far superior to traditional hypothesis testing. In this paper, I illustrate the use of such approaches as well as the interpretation of results analysed in this framework. Hypothesis-testing is still useful in controlled experiments with very few parameters, but attention should strive away from mere declarations of significance and instead focus on the estimation of the effect and its precision. However, information-theoretic approaches, given their ease of use and interpretation, as well as their flexibility, should be readily adopted by ecologists in any exploratory data analysis, especially in issues of model selection.

*Key words* : data analysis; effect; hypothesis testing; model averaging; model selection; regression; precision.

## OUT WITH THE OLD?

When conducting statistical analyses, we often strive to estimate the effect (magnitude) of a given variable on a response variable and its precision. In certain instances, our objective is to go beyond and assess whether the effect is sufficiently important to include the parameter in the model in order to make predictions, an issue of model selection. This is often the case in observational studies, where a number of variables are believed to explain a given ecological process or pattern. Whereas classical techniques such as tests of null hypotheses are well-suited for manipulative experiments, their widespread use and abuse to tackle issues such as parameter estimation and model selection only reflects the slow migration of superior techniques from the distant world of statistics into ecological disciplines. Indeed, hypothesis testing is problematic as it indirectly addresses these issues (i.e., the effect is or is not significant), and it does not perform particularly well in model selection (e.g., variables selected by forward, backward, or stepwise approaches). Though this is debated by some (Robinson and Wainer 2002), better approaches do exist (Anderson et al. 2000, 2001, Guthery et al. 2001, Johnson 1999, 2002).

One such approach, developed in the early 1970's, rests on Akaike's information criterion (AIC) and its associated measures. This framework is also known as the information-theoretic approach, as it has arisen from information theory, a field encompassing a number of methods and theories pivotal to many of the sciences. Because information theory *per se* goes beyond the scope of the present paper, the reader should consult Kullback and Leibler (1951), Cover and Thomas (1991), and Burnham and Anderson (2002) for further discussions on the issue. In ecology, the AIC and its related measures were first applied almost exclusively in the context of model selection in capture-recapture analyses (Lebreton et al. 1992, Anderson et al. 1994), but have gained popularity since the last decade in more general situations (Johnson and Omland 2004). This trend becomes apparent by noting the number of published papers having adopted this approach in the leading ecological journals such as *Ecology*, *Ecological Applications*, *Oikos*, *Journal of Wildlife Management*, and

*Journal of Applied Ecology*. However, some fields such as herpetology, still seem reticent to use these techniques. In this paper, I illustrate with simple examples, the use and interpretation of results of herpetological data using information-theoretic approaches.

## FIRST THINGS FIRST

As pointed out by Burnham and Anderson (2001), three principles regulate our ability to make inferences in the sciences: 1) simplicity and parsimony, 2) several working hypotheses, and 3) strength of evidence. Simplicity and parsimony is a concept based on Occam's razor, which suggests that the simplest explanation is probably the most likely. This is a quality often strived for in science. Parsimony is particularly evident in issues of model building, where the investigator must make a compromise between model bias and variance. Here, bias corresponds to the difference between the estimated value and true unknown value of a parameter, whereas variance reflects the precision of these estimates; a common measure of precision is the SE of the estimate. Thus, a model with too many variables will have low precision whereas a model with too few variables will be biased (Burnham and Anderson 2002). The principle of multiple working hypotheses consists in testing a hypothesis from one experiment, then according to the results, formulating a new hypothesis to test with a new experiment (Chamberlin 1965). In model selection, this translates into testing, for the data at hand, a series of plausible models specified before conducting the analyses. Following the analyses, we then require an indication of which model is the best among those we considered and a measure of the strength of evidence for each model. Information-theoretic approaches adhere in part to all three principles, which makes them quite attractive.

## MINIMIZING THE LOSS OF INFORMATION

Before engaging in the construction of a model (e.g., a linear regression model or any generalized linear model), we must accept that there are no true models. Indeed, models only approximate reality. The question then is to find which model would best approximate reality given the data we have recorded. In other words, we are trying to minimize the loss of information. Kullback and Leibler (1951) addressed such issues and developed a measure, the Kullback-Leibler information, to represent the information lost when approximating reality (i.e., a good model minimizes the loss of information). A few decades later, Akaike (1973 cited by Burnham and Anderson 2001) proposed using Kullback-Leibler information for model selection. He established a relationship between the maximum likelihood, which is an estimation method used in many statistical analyses, and the Kullback-Leibler information. In essence, he developed an information criterion to estimate the Kullback-Leibler information, Akaike's information criterion (AIC), which he defined as

$$AIC = -2(\log\text{-likelihood}) + 2 K,$$

where  $K$  is the number of estimated parameters included in the model (i.e., number of variables + the intercept). The log-likelihood of the model given the data, is readily available in statistical output, and reflects the overall fit of the model (smaller values indicate worse fit).

In cases where analyses are based on more conventional least squares regression for normally distributed errors, one can compute readily the AIC with the following formula (where arbitrary constants have been deleted)

$$\text{AIC} = n \log(\hat{\sigma}^2) + 2K,$$

where  $\hat{\sigma}^2 = \frac{\text{Residual Sum of Squares}}{n}$

and  $n$  is the sample size. It is important to note here that because the variance is estimated, it must be included in the count of parameters ( $K$ ).

The AIC penalizes for the addition of parameters, and thus selects a model that fits well but has a minimum number of parameters (i.e., simplicity and parsimony). For small sample sizes (i.e.,  $n/K < \sim 40$ ), the second-order Akaike Information Criterion ( $\text{AIC}_c$ ) should be used instead

$$\text{AIC}_c = -2(\log\text{-likelihood}) + 2K + \frac{2K(K+1)}{(n-K-1)}$$

where  $n$  is the sample size. As sample size increases, the last term of the  $\text{AIC}_c$  approaches zero, and the  $\text{AIC}_c$  tends to yield the same conclusions as the AIC (Burnham and Anderson 2002).

## MODEL SELECTION

In itself, the value of the AIC for a given data set has no meaning. It becomes interesting when it is compared to the AIC of a series of models specified a priori, the model with the lowest AIC being the « best » model among all models specified for the data at hand. If only poor models are considered, the AIC will select the best of the poor models. This highlights the importance of spending time to determine the set of candidate models based on previous investigations, as well as judgement and a knowledge of the system under study. After having specified the set of plausible models to explain the data and before conducting the analyses (e.g., linear regression), one should assess the fit of the global model, defined as the most complex model of the set. We generally assume that if the global model fits, simpler models also fit because they originate from the global model (Burnham and Anderson 2002, Cooch and White 2001).

Once the appropriate transformations have been conducted (if warranted) and the global model fits the data, one can run each of the models and compute the AIC (or  $AIC_c$ ). The models can then be ranked from best to worse (i.e., low to high AIC values). One should ensure that the same data set is used for each model, i.e., the same observations must be used for each analysis. Missing values for only certain variables in the data set can also lead to variations in the number of observations. Furthermore, the same response variable ( $y$ ) must be used for all models (i.e., it must be identical across models, consistently with or without transformation). Nonetheless, one may specify different link functions or distributions to compare different types of models (e.g., normal, Poisson, logistic; see McCullagh and Nelder 1989).

## COMPARING MODELS

Two measures associated with the AIC can be used to compare models: the delta AIC and Akaike weights. These are easy to compute, as calculations remain the same regardless of whether the AIC or  $AIC_c$  is used, and also have the advantage of being easy to interpret. The simplest, the delta AIC ( $\Delta_i$ ), is a measure of each model relative to the best model, and is calculated as

$$\text{Delta AIC} = \Delta_i = AIC_i - \min AIC$$

where  $AIC_i$  is the AIC value for model  $i$ , and  $\min AIC$  is the AIC value of the « best » model. As a rule of thumb, a  $\Delta_i < 2$  suggests substantial evidence for the model, values between 3 and 7 indicate that the model has considerably less support, whereas a  $\Delta_i > 10$  indicates that the model is very unlikely (Burnham and Anderson 2002).

Akaike weights ( $w_i$ ) provide another measure of the strength of evidence for each model, and represent the ratio of delta AIC ( $\Delta_i$ ) values for each model relative to the whole set of  $R$  candidate models:

Akaike weight = $w_i =$	$\frac{\exp(-\Delta_i/2)}{\sum_{r=1}^R \exp(-\Delta_r/2)}$
-------------------------	--

In effect, we are simply changing the scale of the  $\Delta_i$ 's to compare them on a scale of 1 (i.e., so that the sum of the  $w_i$  equals 1). The interpretation of Akaike weights ( $w_i$ ) is straightforward: they indicate the probability that the model is the best among the whole set of candidate models. For instance, an Akaike weight of 0.75 for a model, indicates that given the data, it has a 75% chance of being the best one among those considered in the set of candidate models. In addition, one can compare the Akaike weights of the « best » model and competing models to determine to what extent it is better than another. These are termed evidence ratios and are calculated as

Evidence ratio =	$\frac{w_j}{w_i}$
------------------	-------------------

where model  $j$  is compared against model  $i$ . For example, an evidence ratio of

$\frac{w_j}{w_i}$	=	$\frac{0.55}{0.40}$	=	1.375
-------------------	---	---------------------	---	-------

would indicate that model  $j$  is only 1.375 more likely than model  $i$  to be the best, given the set of  $R$  candidate models and the data. This suggests that the rank of model  $j$  might change if we were to take a series of independent samples of identical size (Burnham and Anderson 2002). In other words, there would be a high degree of uncertainty regarding the best model. Akaike weights are also useful to give a measure of the relative importance of a variable: one simply sums the  $w_i$  of the models including the variable and compares it to those that do not. However, a better approach is to obtain a model-averaged estimate for the variable across all models (see *multimodel inference* below).

## AIC VS HO

The AIC is not a hypothesis test, does not have an  $\alpha$ -value and does not use notions of significance. Instead, the AIC focuses on the strength of evidence (i.e.,  $\Delta_i$  and  $w_i$ ), and gives a measure of uncertainty for each model. In contrast, conventional model selection approaches such as backward, forward, or stepwise selection procedures are generally based on hypothesis tests, where at a certain  $P$ -value, a variable is included or excluded (Zar 1984, Hosmer and Lemeshow 1989, Afifi and Clark 1996, Kleinbaum et al. 1998). These techniques often yield different conclusions depending on the order in which the models are computed, whereas the AIC approach yields consistent results and is independent of the order in which the models are computed (Anderson et al. 2000, 2001, Burnham and Anderson 2002).

## AN EXAMPLE

Now, let's illustrate the use of the  $AIC_c$  in a real data set. In this example, we use the mass lost in water following dehydration of 126 green frogs (*Rana clamitans melanota*) on three different substrates (i.e., *Sphagnum* moss, soil, or peat) in or out of the shade (21 frogs for each combination of treatments). The initial mass in grams before dehydration was measured, as well as the snout-vent length of the individuals. The mass lost in water after 2 h was modeled with a linear regression fitted with maximum likelihood. Note that least-squares regression can also be used to compute AIC with a

simple formula (see above). Before the analyses, 5 cases with missing data were deleted (to avoid variations in the number of observations used in the analyses), and a logarithmic transformation (base  $e$ ) was applied to the dependent variable to homogenize variances. For the purpose of this example, I chose a set of 7 ( $R = 7$ ) candidate models (Table 1). The global model (model 1) suggested good fit, based on visual inspection of the residuals plotted against the predicted values.

The results in Table 1 indicate that model 4 is the best given the set of 7 candidate models with an Akaike weight of 0.61. However, model 1, which includes the additional variable initial mass, follows it rather closely. Indeed, model 1 has a  $\Delta_i$  of 0.95 and an Akaike weight of 0.38. Thus, model 1 is only 1.61 times more likely to be the best model than model 4 (evidence ratio =  $0.61/0.38$ ), and reveals a relatively high amount of uncertainty regarding the best model. Thus, both models are equally likely, whereas the other models in the set of candidate models are very unlikely (i.e.,  $\Delta_i > 10$ ). This reveals a common problem: when no single model is clearly the best, we cannot base predictions on the model ranked in first place. Fortunately, as highlighted in the next section, there are ways to address the issue.

**Table 1. Akaike's second-order information criterion (AIC<sub>c</sub>) of the regression models of mass lost by frogs after 2 h across three substrate types in or out of the shade. A total of 121 observations were retained for analysis.**

Model*	Model ID	Log-likelihood	Number of parameters (K)	AIC <sub>c</sub>	Delta AIC <sub>c</sub> ( $\Delta_i$ )	Akaike weight ( $w_i$ )
Shade subtype imass SVL	1	-3.54	6	19.81	0.95	0.38
Shade subtype imass	2	-9.30	5	29.12	10.25	0.00
Shade imass SVL	3	-16.74	4	41.82	22.96	0.00
Shade subtype SVL	4	-4.17	5	18.86	0.00	0.61
Shade imass	5	-22.60	3	51.40	32.54	0.00
Subtype imass SVL	6	-13.52	5	37.56	18.69	0.00
Substrate initmass	7	-17.99	4	44.33	25.47	0.00

\*Subtype: categorical variable with 3 levels (*Sphagnum* used as reference level); shade: categorical variable with 2 levels (i.e., no shade vs shade); SVL: snout-vent length (cm), imass: initial mass (g) of frogs before experiment.

## WHEN SEVERAL MODELS COMPETE FOR FIRST PLACE:

## MULTIMODEL INFERENCE

As noted above, in some instances, the « best » model may have competitors for the top rank (i.e.,  $i < 2$ , or equivalently, evidence ratios  $< 2.7$ ). A solution to this problem is to base the inference on the entire set of models, an approach termed multimodel inference or model averaging. Indeed, instead of relying solely on the estimates of the best model, we compute a weighted average of the estimates based on model uncertainty (i.e., the Akaike weights). In essence, we are using all the information available from the entire set of models to make inference and it is a very elegant way of tackling the problem.

For a given parameter, the first step consists in rearranging the AIC table with the models containing the parameter of interest. Delta AIC and Akaike weights are then recomputed for this subset of the models. To conduct model averaging, the estimate of the parameter for each model is then weighted by the Akaike weights, as follows

$$\text{Model-averaged estimate} = \hat{\theta} = \sum_{i=1}^R w_i \hat{\theta}_i$$

where  $\hat{\theta}_i$  denotes the estimate for model  $i$ .

Similarly, we can also compute the precision (SE) of the model-averaged estimate, termed as the unconditional SE (i.e., a SE not restricted to a single "best" model)

$$\text{Unconditional SE} = \sqrt{\sum_{i=1}^R w_i \left[ \text{var}(\hat{\theta}_i | g_i) + (\hat{\theta}_i - \hat{\theta})^2 \right]}$$

where  $\text{var}(\hat{\theta}_i | g_i)$  represents the variance of the estimate  $\hat{\theta}_i$  given model  $g_i$ . Note that  $\text{var}(\hat{\theta}_i | g_i)$  equates to squaring the SE of  $\hat{\theta}_i$ . Returning to our example with dehydrated frogs, we can easily compute the model-averaged estimate. Table 2 illustrates this approach for the effect of the shade variable.

In many cases, model averaging reduces bias and increases precision, which are very desirable properties (Burham and Anderson 2002). Once the model-averaged estimates and SE are calculated, we can use confidence intervals to assess the magnitude of the effect. For a 95% confidence interval,

$$\text{Upper 95\% confidence limit} = \text{estimate} + (1.96)\text{SE}$$

and

$$\text{Lower 95\% confidence limit} = \text{estimate} - (1.96)\text{SE}$$

We conclude that the estimate is different from 0 (i.e., there is an effect) when the confidence interval excludes 0. In our dehydrated frog example, the 95% confidence interval for the model-averaged estimate of shade would be (0.1233, 0.3019). The confidence interval excludes 0 and indicates that frogs out of the shade lost more water than those in the shade.

**Table 2. Akaike's second-order information criterion (AIC<sub>c</sub>) and associated measures recomputed to obtain the model-averaging estimate and precision (i.e., unconditional SE) of the shade variable, based on normal regression models of mass lost by frogs after 2 h across three substrate types in or out of the shade. Analyses are based on 121 observations.**

Model*	Model ID	AIC <sub>c</sub>	Delta AIC <sub>c</sub> (Δ <sub>i</sub> )	Akaike weight (w <sub>i</sub> )	Regression estimate of shade	SE of estimate of shade
Shade subtype imass SVL	1	19.81	0.95	0.38	0.2114	0.0454
Shade subtype imass	2	29.12	10.25	0.00	0.2057	0.0476
Shade imass SVL	3	41.82	22.96	0.00	0.2104	0.0506
Shade subtype SVL	4	18.86	0.00	0.61	0.2134	0.0456
Shade imass	5	51.40	32.54	0.00	0.2035	0.0531
Model-averaged estimate					0.2126	
Unconditional SE						0.0455

\*Subtype: categorical variable with 3 levels (*Sphagnum* used as reference level); shade: categorical variable with 2 levels (i.e., no shade vs shade); SVL: snout-vent length (cm); imass: initial mass (g) of frogs before experiment.

## SPECIAL ISSUES WITH COUNT DATA

Discrete data (i.e., data occurring as integers) such as the number of individuals in a trap can be modeled using Poisson regression (McCullagh and Nelder 1989, Agresti 1996). However, it is common to encounter overdispersion in such data. In other words, data vary more than expected from data following a Poisson distribution (McCullagh and Nelder 1989). Poisson-distributed data have a mean equal to the variance (i.e.,  $\mu = \sigma^2$ ), whereas overdispersion occurs when the mean exceeds the variance (i.e.,  $\mu > \sigma^2$ ). Overdispersion may arise due to biological phenomenon such as aggregation, or may be a sign of inadequacy of a model. To detect whether overdispersion occurs in a data set subjected to Poisson regression, we can estimate the dispersion parameter, *c-hat*, with the

ratio of the deviance over the residual degrees of freedom (McCullagh and Nelder 1989),

$$c\text{-hat} = \frac{\text{Residual deviance}}{\text{Residual df}}$$

If  $c\text{-hat} = 1$ , then no overdispersion occurs. If  $c\text{-hat}$  exceeds 1, then there is indication of overdispersion; values  $< 1$  may suggest underdispersion but often hint inadequate model structure. Regardless, a model with  $c\text{-hat} \ll 1$  or  $c\text{-hat} \gg 4$  suggests that a Poisson model is probably not adequate. Alternatively, a negative binomial model could be used to account for overdispersion (McCullagh and Nelder 1989). We can account for overdispersion in the AIC as follows,

$$\text{QAIC} = \frac{-2(\log\text{-likelihood}) + 2K}{c\text{-hat}}$$

Similarly, the  $\text{AIC}_c$  can also be adjusted for overdispersion:

$$\text{QAIC}_c = \frac{\text{QAIC} + \frac{2K(K+1)}{n-K-1}}{c\text{-hat}}$$

Note that  $c\text{-hat}$  is an additional parameter to estimate. Thus, it must be included in the count of parameters. As the estimated  $c\text{-hat}$  will vary from model to model, it is advised to use the  $c\text{-hat}$  of the global model (i.e., the most complex model) and use it consistently for the other models. The logic being that the most complex model will yield the best estimate for  $c\text{-hat}$ . Refer to Burnham and Anderson (2002) for further issues in overdispersion, especially regarding model averaging and estimating  $c\text{-hat}$  when no global model exists.

## ADVANTAGES AND LIMITATIONS OF AIC

The AIC provides an objective way of determining which model among a set of models is most parsimonious, as we do not rely on  $\alpha$ . It is rigorous, founded on solid statistical principles (i.e., maximum likelihood), yet easy to calculate and interpret. Indeed, all elements can be obtained from most statistical analysis software, such as SAS, R, S-PLUS, or SPSS. In addition, the measures associated with the AIC, such as delta AIC and Akaike weights, supply information on the strength of evidence for each model. Thus, the concept of significance becomes superfluous with the AIC. Anderson et al. (2001) suggest using it to solve conflicts in the applied sciences. The greatest strength of the AIC is its potential in model selection (i.e., variable selection), because it is independent of the order in which models are computed. In the case where there are many models ranked highly based on the AIC, we can incorporate model uncertainty to obtain robust and precise estimates, and confidence intervals.

Nonetheless, the AIC approach is not a panacea. For instance, a model is only good as the data which have generated it. In addition, the conclusions will depend on the set of candidate models specified before the analyses are conducted: we will never know if a better model exists unless it is specified in the candidate set. To the great joy of ANOVA aficionados, in certain cases, it is still preferable to use hypothesis testing instead of the AIC. This is particularly true for controlled experiments, randomized, and replicated with few independent variables. Indeed, using the AIC for a three-way ANOVA, or whenever only few models, say 1-3, appear in the candidate set will not be more instructive than hypothesis tests. However, even in these cases, investigators should routinely report the estimate and SE of the parameters they have estimated as they are much more instructive than  $P$ -values and simple declarations of « significant » or « not significant ». The presentation of estimates and associated SE greatly improves the value of a study, and it becomes especially useful for investigators subsequently conducting meta-analyses.

## A FLASH IN THE PAN OR HERE TO STAY?

In conclusion, the information-theoretic approach revolving around the AIC shows great promise for various applications in ecology, conservation biology, behavioral ecology, and physiology. Its strength is particularly in model selection, for situations generated by observational studies conducted in the field, where regressions are sought to model a given pattern or process as a function of a number of independent variables. As it is rather straightforward in computation and interpretation, the AIC is a useful tool that should be seriously considered by biologists faced with the task of analysing empirical data.

## ACKNOWLEDGEMENTS

Discussions with M. Bélisle, Y. Turcotte, and A. Desrochers have improved this manuscript. This work was funded by NSERC and FCAR to M. J. Mazerolle, A. Desrochers, and L. Rochefort.

## REFERENCES

- Afifi, A. A., and V. Clark. 1996. Computer-aided Multivariate Analysis, 3rd edition. Chapman & Hall/CRC, New York.
- Agresti, A. 1996. An introduction to categorical data analysis. John Wiley & Sons, New York, USA.
- Anderson, D. R., K. P. Burnham, and G. C. White. 1994. AIC model selection in overdispersed capture-recapture data. *Ecology* **75** :1780-1793.

- Anderson, D. R., K. P. Burnham, and W. L. Thompson. 2000. Null hypothesis testing: problems, prevalence, and an alternative. *Journal of Wildlife Management* **64** :912-923.
- Anderson, D. R., K. P. Burnham, and G. C. White. 2001. Kullback-Leibler information in resolving natural resource conflicts when definitive data exist. *Wildlife Society Bulletin* **29** :1260-1270.
- Anderson, D. R., K. P. Burnham, W. R. Gould, and S. Cherry. 2001. Concerns about finding effects that are actually spurious. *Wildlife Society Bulletin* **29** :311-316.
- Anderson, D. R., W. A. Link, D. H. Johnson, and K. P. Burnham. 2001. Suggestions for presenting the results of data analyses. *Journal of Wildlife Management* **65** :373-378.
- Burnham, K. P., and D. R. Anderson. 2001. Kullback-Leibler information as a basis for strong inference in ecological studies. *Wildlife Research* **28** :111-119.
- Burnham, K. P., and D. R. Anderson. 2002. *Model Selection and Multimodel Inference: a practical information-theoretic approach*, 2nd edition. Springer-Verlag, New York.
- Chamberlin, T. C. 1965. The method of multiple working hypotheses. *Science* **148** :754-759. (reprint of 1890 paper in *Science* **15** :92).
- Cooch, E., and G. White. 2001. Program MARK: Analysis of data from marked individuals, "a gentle introduction", 2nd edition. <http://www.cnr.colostate.edu/~gwhite/mark/mark.html>.
- Cover, T. M., and Thomas, J. A. 1991. *Elements of Information Theory*. John Wiley and Sons, New York, New York.
- Guthery, F. S., J. J. Lusk, and M. J. Peterson. 2001. The fall of the null hypothesis: liabilities and opportunities. *Journal of Wildlife Management* **65** :379-384.
- Hosmer, D. W., and S. Lemeshow. 1989. *Applied Logistic Regression*. John Wiley & Sons, Inc., New York, USA.
- Johnson, D.H. 1999. The insignificance of statistical significance testing. *Journal of Wildlife Management* **63**, 763-772.
- Johnson, D. H. 2002. 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 & Evolution* **19** :101-108.
- Kleinbaum, D. G., L. L. Kupper, K. E. Muller, and A. Nizam. 1998. *Applied regression analysis and other multivariable methods*, 3rd edition. Duxbury Press, Toronto.

Kullback, S., and R. A. Leibler. 1951. On information and sufficiency. *Annals of Mathematical Statistics* **22** :79-86.

Lebreton, J.-D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals: a unified approach with case-studies. *Ecological Monographs* **62** :67-118.

McCullagh, P., and J. A. Nelder. 1989. *Generalized Linear Models*, 2nd edition. Chapman & Hall, New York, USA.

Robinson, D. H., and H. Wainer. 2002. On the past and future of null hypothesis significance testing. *Journal of Wildlife Management* **66** :263-271.

Zar, J. H. 1984. *Biostatistical Analysis*, 2nd edition. Prentice Hall, Englewood Cliffs, New Jersey, USA.



Bibliographie

