
Autoregressive Models for Estimating Phylogenetic and Environmental Effects: Accounting for Within-species Variations

Pierre-André Cornillon^a, Dominique Pontier^{b*}, and Marie-Joëlle Rochet^c

^aGroupe de biostatistiques et d'analyse des systèmes, CIRAD forêt, programme arbres et plantations, Campus International de Baillarguet, 5035, 34032, Montpellier Cedex 1, France

^bUMR CNRS 5558 "Biométrie et Biologie Evolutive" Université Claude Bernard LYON 1, 43, boulevard du 11 novembre 1918, 69622, Villeurbanne Cedex, France

^cLaboratoire MAERHA, IFREMER Nantes, 21105, 44311, NANTES Cedex 03, France

*dpontier@biomserv.univ-lyon1.fr

Abstract: Comparative methods are used to investigate the attributes of present species or higher taxa. Difficulties arise from the phylogenetic heritage: taxa are not independent and neglecting phylogenetic inertia can lead to inaccurate results. Within-species variations in life-history traits are also not negligible, but most comparative methods are not designed to take them into account. Taxa are generally described by a single value for each trait. We have developed a new model which permits the incorporation of both the phylogenetic relationships among populations and within-species variations. This is an extension of classical autoregressive models. This family of models was used to study the effect of fishing on six demographic traits measured on 77 populations of teleost fishes.

Keywords: Comparative method; fishing effect; intraspecific variation; autoregressive method

Abstract

Comparative approach in Biology is used to investigate the attributes of actual species or higher taxa. Difficulties arise from the phylogenetic heritage: taxa are not independent and neglecting phylogenetic inertia can lead to inaccurate results. Moreover intraspecific variations in life history traits are not negligible, most comparative methods are not designed to take them into account. Classically, taxa are described by a single value for each trait. In this paper, we propose a new modelling approach which permits to incorporate both the phylogenetic relationships among populations and intraspecific variations. This approach is an extension of classical autoregressive models. We applied this family of models to study the fishing effect on six demographic traits measured on 77 populations of Teleost fishes.

1 Introduction

Comparative methods are more and more used to test adaptive hypotheses in such diverse fields as ethology (reviewed in Foster & Cameron 1996), population dynamics and life history evolution (Stearns 1983; Gaillard *et al.* 1989; Promislow & Harvey 1990; Sinervo 1990; Pontier *et al.* 1993; Shine & Iverson 1995; Gaillard *et al.* 1997; Clobert *et al.* 1998), zoology (Gittleman & Luh 1992; Gordon & Bergstad 1992), parasitology (Skorping *et al.* 1991) and more recently in molecular biology (Bromham *et al.* 1996). Modern methods include phylogenetic information and are used to estimate or correct for phylogenetic inertia that might otherwise obscure the role of natural selection in trait variation (Wanntorp 1983; Harvey & Pagel 1991; Martins & Hansen 1997). Until now most of these methods have been based on between-species or higher order taxa data (e.g. Western 1979; Stearns 1983; Gaillard *et al.* 1989; Promislow & Harvey 1990; Winemiller & Rose 1992; Clobert *et al.* 1998). This implicitly assumes intraspecific variability in life history traits is negligible compared to inter-taxons variability. However a lot of studies has now demonstrated that traits may largely vary among populations of a same species (Leggett & Carscadden 1978; Brown 1983; Jennings & Beverton 1991; Sedinger & Flint 1995 Juanes *et al.* 1996; Reznick *et al.* 1996; Rose *et al.* 1998).

Despite this fact, most modern comparative methods are not designed to properly estimate the relationships between changes in trait and environment at the species level (Martins & Hansen 1997). Comparative studies of such variation, by enhancing the quality of higher-order comparisons, would greatly improve insight into evolutionary process (Gittleman & Kot 1990; Foster & Cameron 1996; Martins & Hansen 1997). The major problem is to combine different time-scales: historical (macroevolutionary processes) and present ecological time (microevolutionary processes).

Here we propose a new comparative method designed to incorporate the population level in interspecific comparative analyses. Our method is derived from autoregressive methods. These methods have been developed to statistically decompose trait variation into two parts: a phylogenetic autoregressive component based on the assumption that closer relatives will have more similar traits than less related species; and a residual specific component which includes the remaining variability (Cheverud *et al.* 1985; Gittleman & Kot

1990). We extended this framework by splitting the phylogenetic autoregressive component into two parts: the inter- and the intra-specific autoregressive components. We explicitly included environmental influence into the residual specific component.

Our method has been applied to analyze life history traits variation among 77 fish populations. These populations undergo various fishing pressures constituting novel environments for individuals. Therefore environmental effects can be estimated independently of phylogenetic influence. Our statistical approach allowed first to characterize covariation of fish life history characters when the effects of phylogeny and adult body size is removed; second to estimate and compare the effects of fishing on life history traits among species. We compared the results of our method with a non phylogenetic approach and with a classical autoregressive method.

The organization of this paper is as follows: in section 2, we introduce the classical autoregressive comparative model and explain the limits of this method. We propose a new family of models to take into account intraspecific variations. In section 3 we compare linear modeling of the effect of fishing to our family of models and discuss the differences. Finally section 4 contains our conclusions.

2 A family of autoregressive models

Comparative methods assume that a collection of p estimates of phenotypic means for a trait are available. Let denote $Y = (Y_1 \cdots Y_p)'$ the vector of the p measurements. As described in Cheverud *et al.* we assume that every trait for a given species i can be written as a linear function

$$Y_i = Z_i + \varepsilon_i + L_i\gamma,$$

where Z_i is the phylogenetic (or inherited) part, ε_i represents the adaptation of species i and $L_i\gamma$ a dependence with r covariates, typically size in an allometric context.

This decomposition into 2 or 3 parts (presence/absence of covariates) is the same as those used by Cheverud *et al.* (1985) or Lynch (1991) in two different models. That is what we can call the basic phylogenetic assumption.

It follows from this equation that we need to make more assumptions to know the precise form of Z_i , the phylogenetic value. Using Cheverud *et al.*'s (1985) or Gittleman and Kot's (1990) assumption, Z_i the inherited value of population i can be reconstructed from the other observations of the trait Y . Because phylogenetic neighbors are thought to have close values of the trait, the phylogenetic part is written as a weighted mean of Y

$$Z_i = \sum_{j=1, j \neq i}^p \rho W_{ij} Y_j.$$

W is an $p \times p$ weighting matrix where W_{ij} are close to 1 for species i and j who have a recent common ancestor (who are neighbors on the phylogenetic tree) and close to 0 otherwise. The weights are chosen as the inverse of the distance between species i and j .

If no precise information on the distance between ancestors are known, which is usually the case, every edge of the phylogenetic tree is set equal to 1 in order to capture only the shape of the tree. The parameter ρ is unknown and has to be estimated. It represents the intensity of the phylogenetic part in the analyzed trait Y . We can afford this simplification because autoregressive methods are known to be robust to inaccuracies in branch length (Gittleman & Luh, 1992).

To fit the environmental residual part we can use again the classical assumption of Cheverud *et al.* (1985) or Gittleman & Kot (1990), that is for each species the environmental part follows an identically independent distributed normal law with mean μ and variance σ^2

$$\varepsilon_i \equiv \mathcal{N}(\mu, \sigma^2). \quad (1)$$

In other words for each species the adaptation of species is the same in mean and fluctuate around μ with the same magnitude σ^2 . Thus this hypothesis is tenable only if each species is suspected to have the same environment. When it is not the case this hypothesis has to be relaxed. Moreover recall that for classical comparative studies, each measure Y_i is a phenotypic mean in the species i .

To take into account intra-specific variations, which is our goal, we have to work on n measures of the trait Y *on different populations*. The hypothesis (1) of the same environment for each population is too restrictive : we want to model intra-specific variation and we suppose with (1) that there is the same intra-specific contribution μ for all the n populations. The best formulation is to make an intra-specific contribution μ_i different for each population i but it will be impossible to estimate all these coefficients. Thus we have to suppose the existence of G groups of homogeneous environments. We relax hypothesis (1) as

$$\varepsilon_i \equiv \mathcal{N}(\mu_{g(i)}, \sigma^2), \quad (2)$$

where $g(i) \in \{1 \cdots G\}$ denote a group of homogeneous environment to which the population i belongs. G is chosen small to have enough populations in each group for the coefficients to be estimable. We also can do a more complex assumption on the environmental effect

$$\varepsilon_i \equiv \mathcal{N}(\mu_{g(i)}, \sigma_{g(i)}^2). \quad (3)$$

Thus we have now 3 possibilities to model the environmental part. When the most complex one (3) is true, there exists G groups of homogeneous environmental variation around the mean $\mu_{g(i)}$ with magnitude of $\sigma_{g(i)}^2$. Difference due to adaptation and thus intra-specific variation is now accounted by these G groups. The model (2) is just an intermediary level between the simplest hypothesis of Cheverud *et al.* and the complex one.

Using population instead of species dictates adaptations for the phylogenetic effect. Because there still exists some links between populations of the same species we can use traits of population in the same species to estimate the phylogenetic part. But we cannot use the same scale for the inter-specific evolution and these intra-specific links. Thus we split

the inherited part into 2 terms, the intra-specific (within) autoregression and inter-specific (between) autoregression

$$Z_i = \sum_{j=1, j \neq i}^n \rho^{(w)} W_{ij}^{(w)} Y_j + \sum_{j=1, j \neq i}^n \rho^{(b)} W_{ij}^{(b)} Y_j.$$

$W_{ij}^{(b)}$ is the inverse of path's length on the phylogenetic tree between the species of population i and the species of population j (Cheverud et al., 1985). $W_{ij}^{(w)}$ are the same everywhere and chosen without loss of generality equal to 1.

Finally we have a family of models defined by

$$\begin{aligned} Y &= \rho^{(w)} W^{(w)} Y + \rho^{(b)} W^{(b)} Y + \varepsilon + L\gamma \\ \text{and } \varepsilon_i &\equiv \mathcal{N}(\mu, \sigma^2) \text{ model 1, same environment} \\ \varepsilon_i &\equiv \mathcal{N}(\mu_{g(i)}, \sigma^2) \text{ model 2, equal variance in all groups of environment} \\ \varepsilon_i &\equiv \mathcal{N}(\mu_{g(i)}, \sigma_{g(i)}^2) \text{ model 3.} \end{aligned}$$

From these equations the log-likelihood of models can be found. For model 3:

$$\begin{aligned} \mathcal{L} &= \ln |B| - \frac{1}{2} \ln |\Sigma| - \frac{1}{2} [BY - \theta]' \Sigma^{-1} [BY - \theta] \\ &\text{where } B = I_n - \rho^{(w)} W^{(w)} + \rho^{(b)} W^{(b)} \\ &\Sigma = \text{diag}(\dots \sigma_{g(i)}^2 \dots) \\ &\text{and } \theta' = (\dots \mu_{g(i)} \dots) + (L\gamma)'. \end{aligned}$$

We use maximum likelihood estimates of the parameters of the three models. Maximization of \mathcal{L} implies an iterative method. To decrease the number of parameters involved in the maximization we can compute the concentrated likelihood. For example in the model 3 (or in the model 1), we first set the partial derivatives of \mathcal{L} with respect to $\{\sigma_{g(i)}\}$ equal to zero to obtain the estimate of all $\sigma_{g(i)}$

$$\sigma_{g(i)}^2 = \frac{1}{\#(g)} [BY - \theta]' D_g [BY - \theta] \quad (4)$$

where $\#(g)$ denote the cardinal of group g

$D_g = \text{diag}(\dots 1_{g(i)} \dots)$ and $1_{g(i)}$ indicator function of group $g(i)$.

Using this estimation in the log-likelihood equation leads to the concentrated likelihood (for models 1 and 3)

$$\mathcal{M} = \frac{1}{2} \sum_{g=1}^G \#(g) \ln \{ [BY - \theta]' D_g [BY - \theta] \} - \ln |B|. \quad (5)$$

The parameter estimators are the minimum of \mathcal{M} . It can be seen easily that a minimization of such a function is not possible in a formal way and we use an iterative algorithm:

the steepest descent algorithm, see for instance Hager *et al.* 1993 (Appendix B). Using concentrated likelihood for model 3 allows to decrease the number of parameters to be calculated by an iterative method of optimization by G : $\{\sigma_g\}$ are no longer present in the definition of \mathcal{M} . We could use (5) for model 2 too, but it eliminates only one parameter: σ . Therefore we use another concentrated likelihood which does not involve the $\{\mu_g\}$ (see appendix A) and decrease the number of parameters by G .

Finally we have a family of models among which we have to select the best one. As these models are nested, we can apply a likelihood ratio test procedure. For instance if we want to test a classical model (model 1, $\mu_g = \mu, \sigma_g = \sigma, \forall g \in \{1, \dots, G\}$) with $q_0 = 4 + r$ parameters, H_0 , against the environmental model 2 ($\mu_g, \sigma_g = \sigma, \forall g \in \{1, \dots, G\}$) with $q_1 = G + 3 + r$ parameters, H_1 , we use twice the difference between the log-likelihood estimated under H_1 ($\widehat{\mathcal{L}}_{H_1}$) and the log-likelihood estimated under H_0 ($\widehat{\mathcal{L}}_{H_0}$). This statistic follows a Chi-square distribution with $q_1 - q_0 = G - 1$ degrees of freedom

$$\lambda = 2 \left(\widehat{\mathcal{L}}_{H_1} - \widehat{\mathcal{L}}_{H_0} \right) \cong \chi^2(q_1 - q_0).$$

If observed λ is greater than $\chi_{1-\alpha}^2(q_1 - q_0)$, the $(1 - \alpha)^{th}$ quantile, then the model H_0 is rejected. The size of test α is chosen by the user and usually equal to 5%.

To guarantee a familywise error rate equal to α chosen by the user the first test is model 1 (H_0) against model 2 (H_1), and if H_0 is rejected then we perform a second test: model 2 (H_0) against model 3 (H_1) (see for instance Marcus & Peritz, 1976)(Marcus, Peritz, & Reitmeir, 1976). We can apply the same procedure to test whether a phylogenetic effect exists or not, we just need to estimate the parameters with $\rho^{(w)} = \rho^{(b)} = 0$. The corresponding models are also nested within their counterpart with phylogenetic effect. We can build an intermediary family of models with $\rho^{(w)} = 0$, close to the classical model. A comparison of these models with informative criterion such as Akaike Informative Criterion (Akaike, 1974) is also possible and gives usually the same results. For sake of completeness the maximum likelihood theory allows us to compute asymptotic confidence intervals for parameters from the Fisher information matrix.

3 Example

We are interested in measuring the effect of fishing among 77 diversely exploited fish populations from 49 species. Numerous examples of influence of fishing on fishes are discussed in the literature (e.g. Schaffer and Elson 1975; de Veen 1976; Ricker 1981; Nelson and Soulé 1987; Jorgensen 1990; Rijnsdorp 1991; Trippel 1995; Trippel *et al.* 1997). Monostock studies seem a natural way to investigate this effect on a given population, but it is often mixed with other, generally unknown variations. To circumvent this difficulty we are led to comparative approach on populations. To take into account both intraspecific variation due to fishing and phylogenetic links between these 77 populations, we use the method of the previous section. We compare the results to an linear model fitting including effect of fishing and a phylogenetic effect.

We analyze fishing pressure, thus we build $G = 3$ homogeneous groups of environments with low, moderate and high fishing pressure. Each population was assigned to a fishing pressure group on the basis of the ratio of fishing mortality F to natural mortality M : low fishing pressure group ($F/M < 1$), intermediate ($1 < F/M < 2$) and high fishing pressure group ($F/M > 2$).

The constitution of these groups is independent of other possible variation : phylogenetic part and other intraspecific variation. For instance fishing pressure is independent of taxa, and uniformly spread along the phylogenetic tree. Thus we can state that the mean of these fluctuations, in one given group g is zero, and they just contribute to the fluctuation around the mean (i.e. σ_g). This is obviously our environmental hypothesis. If there is some effect of fishing it would be reflected in the coefficients $\{\mu_g\}$ and the selected model have to be one with different means (hypothesis (2) or (3)). A contrario, if no effect of fishing on the analyzed trait is present then the chosen model is (1).

We analyzed five life history traits : time-to-5%-survival (T_{05}), the time elapsed from sexual maturity until 95% of a cohort is dead, which is an index of life-span as affected by fishing mortality; length-at-5%-survival (L_{05}) was used as adult size parameter; median age and length at sexual maturity (Am and Lm); and Fb , the slope of the log-log fecundity-length relationship. Only the topology of the phylogenetic tree is known, based on morpho-anatomical characters, and branch lengths were set equal to one. For further detail and data sources refer to Rochet (1998). Prior to all the studies, we use a log transformation on each trait except for the coefficient Fb .

In order to estimate effect of fishing on these traits, we performed a non-phylogenetic analysis by fitting a linear model to log-transformed data, with adult size and fishing pressure group as explanatory variables. Fishing has a significant effect on two traits, length at maturity and time-to-5%-survival (Tab. 1). We can also incorporate a rough phylogenetic constraint in the linear model: a categorical variable which represents the order to which a population belongs (3-factor model with or without interaction). Whereas the latter model shows significant effect of phylogeny on three traits (Am , T_{05} , Fb), there is no significant interaction between phylogeny at the order-level and fishing; moreover, this rough phylogenetic approach does not change the results, compared to the 2-factor model (Table 1).

By contrast, our models reveal both a significant phylogenetic effect on each trait except length at maturity, and a significant effect of fishing on all traits except adult size (Table 2). The fishing effect on Am and Fb is concealed when phylogenetic constraints are not properly taken into account.

The phylogenetic part for Length at Maturity (Lm) is not significant in any of the models. This may be explained by the regression performed on Length at 5% survival, which removes most of the phylogenetic constraint. Our family of models provides a test to know whether a phylogenetic component is significant, and which links are important (within or between species). A formal examination of phylogenetic part, if present, can help to understand evolutionary patterns (Rochet *et al.* 2000). This type of analysis is beyond the scope of this paper.

To interpret fishing pressure we have to look at the parameters $\{\mu_g\}$ when present (all

traits except L_{05}). Thus we can notice on Fig. 1 that, unsurprisingly, fishing decreases time-to-5%-survival, but also affects other traits: increasing fishing pressure decreases age at maturity and increases length at maturity and the slope of the fecundity-length relationship. Populations compensate for the high adult mortality caused by fishing by maturing earlier, growing faster and by a higher fecundity.

4 Conclusion

The comparative method presented here estimates the relative contribution to trait variability of phylogeny, allometry and environment. This method analyses patterns of variability and does not aim to infer the processes underlying these patterns. Among the existing comparative methods, many assume an evolutionary process underlying phylogenetic similarity, e.g. a Brownian motion model (Felsenstein 1985; Grafen 1989) or stabilizing selection (Hansen 1997). A phylogeny with known branch lengths is necessary to estimate the parameters of these models (Felsenstein 1985; Hansen 1997). On the other hand, phylogenetic autoregression has been criticized because the interpretation of the estimated parameters is not clear in evolutionary terms (Gittleman & Luh 1992; Martins 1996). However, as we do not have any information about the history of the selective regime on the phenotypic traits analyzed, a method accurately describing existing patterns seems more reasonable than inferring untestable evolutionary processes.

Our method extends the previous autoregressive method developed by Cheverud *et al.* (1985) and Gittleman & Kot (1990) in two manners. First by splitting the autoregressive component into two parts, one inter-specific and one intraspecific. This additional complexity is required because branch lengths are not known. When inter and intra-specific variations are not specified in the model, the difference of time scale between evolutionary and ecological processes has to be described by arbitrary differences of branch lengths between or within species (Rochet 1998). The second extension relative to classical autoregression is the estimation of an environmental component. In the classical autoregression, the influence of environment can be sought by examining correlation of environmental variables with the residual 'specific' component after it has been estimated (Gittleman & Kot 1990; Rochet 1998). Our method both estimates and tests for environmental effects while fitting the autoregressive model. The simultaneous incorporation of all effects in a single model allows a better estimation of each component. This is exemplified by the fact that, for some traits, effect of fishing are significant only when phylogenetic effects are described in the model. That the significance of effects of fishing differs when phylogenetic effects are taken into account or not is one more plea to incorporate phylogenetic information into the analysis of comparative data.

A further advantage of the likelihood procedures used here is that they allow to test for each separate effect by likelihood ratio test. It has been suggested that comparative methods should not be used unless phylogenetic effects are evidenced by various tests (Gittleman & Kot 1990; Björklund 1997). However, to detect whether a phylogenetic pattern exists in the data (H_1) or not (H_0), we need a statistic which is a function of the

observations. Because data are distributed in a random unknown fashion, this function also has an unknown distribution. The assumptions needed for deriving the distribution of the statistics are directly linked to the models supposed to be true under H_0 (Cliff & Ord 1981). Moreover, to have a good control on the power of the test, the probability distribution under H_1 needs to be known (Lehmann 1986). Here H_1 is not entirely specified: we just know that independence of observations is not satisfied, this is the second difficulty with testing independence. As these informations required to conduct a classical test are not available, we prefer an alternative to this approach: we first define models and then conduct a statistical choice between them.

A strong assumption of the present approach is that there is no interaction between the various components (phylogenetic, allometric, environmental). Each component is considered as a constraint which cannot be influenced by the other components. This clearly may be partially wrong, and for example we can reasonably assume that fish from different taxa will respond differently to fishing. Hence interaction between phylogeny and environment may exist, which has not been modeled here. The estimation of interaction terms requires a much larger sample than the one used here. The present model could easily be extended to account for interactions between components. The analyses performed in this paper are univariate; an extension to multivariate modeling would also be possible. As done by Lynch (1991) simplification for the covariance structure is necessary to limit the increase in the number of parameters; further work is needed on this topic to obtain tractable models.

References

- Akaike (1974). A new look at the statistical identification model. *IEEE Transactions on automatic control*, **19**, 716–723.
- Björklund, M. (1997). Are 'comparative methods' always necessary?. *Oikos*, **80**, 607–612.
- Bromham, L., Rambaut, A. & Harvey, P. H. (1996). Determinants of rate variation in mammalian DNA sequence evolution. *Journal of molecular evolution*, **43**, 610–621.
- Brown, K. M. (1996). Do life history tactics exist at the intraspecific level? Data from freshwater snails. *The American Naturalist*, **121**, 871–879.
- Cheverud, J. M., Dow, M. M. & Leutenegger, W. (1985). The quantitative assessment of phylogenetic constraints in comparative analyses : Sexual dimorphism in body weights among primates. *Evolution*, **39**, 113–121.
- Clobert, J., Garland, T. & Barbault, R. (1998). The evolution of demographic tactics in lizards: a test of some hypotheses concerning life history evolution. *Journal of Evolutionary Biology*, **11**, 329–364.
- de Veen, J. F. (1976). On changes in some biological parameters in the North Sea sole. *Journal du Conseil international pour l'Exploration de la Mer*, **37**, 60–90.

- Felsenstein, J. (1985). Phylogenies and the comparative method. *The american naturalist*, **125**, 1–15.
- Foster, S. A., & Cameron, S. A. (1996). Geographic variation in behavior: a phylogenetic framework for comparative studies. In: *Phylogenies and the comparative method in animal behaviour* (Martins, E. P., ed). Oxford: Oxford University Press.
- Gaillard, J. M., D. Pontier, D. A., Loison, A., Hervé, J.-C. & Heizmann, A. (1997). Variation in growth form and precocity at birth in eutherian mammals. *Proceedings of the Royal Society of London B, Biological Sciences*, **264**, 859–868.
- Gittleman, J. L., & Kot, M. (1990). Adaptation : statistics and a null model for estimating phylogenetic effects. *Systematic Zoology*, **39**, 227–241.
- Gittleman, J. L., & Luh, H. K. (1992). On comparing comparative methods. *Annual review of ecology and systematics*, **39**, 383–404.
- Gordon, J. D. M., & Bergstad, O. A. (1992). Species composition of demersal fish in the Rockall Trough, North- Eastern Atlantic, as determined by different trawls. *Journal of the marine biology Association of the U.K.*, **72**, 213–230.
- Grafen, A. (1989). The phylogenetic regression. *Philosophical transaction of the royal society, London, serie B*, **326**, 119–157.
- Hager, W. W., Horst, R. & Pardalos, P. (1993). Mathematical programming - a computational perspective. In: *Handbook of statistics* (Rao, C. R., ed), pp. 201–278. Amsterdam: Elsevier Science.
- Hansen, T. F. (1997). Stabilizing selection and the comparative analysis of adaptation. *Evolution*, **51**, 1341–1351.
- Harvey, P. H., & Pagel, M. D. (1991). *The comparative method in evolutionary biology*. Oxford: Oxford university press.
- Jennings, S., & Beverton, R. J. H. (1991). Intraspecific variation in the life history tactics of Atlantic herring stocks. *ICES Journal of marine Science*, **48**, 117–125.
- Jorgensen, T. (1990). Long-term changes in age at sexual maturity of Northeast Arctic cod (*Gadus morhua* L.). *Journal du Conseil international pour l'Exploration de la Mer*, **46**, 235–248.
- Juanes, F., Hare, J. A. & Miskiewicz, A. G. (1996). Comparing early life history strategies of *Pomatomus saltatrix*: a global approach. *Marine and Freshwater Research*, **47**, 365–379.
- Leggett, W. C., & Carscadden, J. E. (1978). Latitudinal variation in reproductive characteristics of American shad (*Alosa sapidissima*): evidence for population specific life history strategies in fish. *Journal of the Fishery Research Board of Canada*, **35**, 1469–1478.

- Lehmann, E. L. (1959). *Testing statistical hypotheses*. New-York: John Wiley and Sons.
- Lynch, M. (1991). Methods for the analysis of comparative data in evolutionary biology. *Evolution*, **45**, 1065–1080.
- Marcus, R., Peritz, E. & Reitmeir, P. (1976). On closed testing procedures with special reference to ordered analysis of variance. *Biometrika*, **63**, 655–660.
- Martins, E. (1996). Phylogenies, spatial autoregression, and the comparative method : a computer simulation test. *Evolution*, **50**, 1750–1765.
- Martins, E. P., & Hansen, T. F. (1997). Phylogenies and the comparative method: A general approach to incorporating phylogenetic information into the analysis of interspecific data. *The American Naturalist*, **149**, 646–667.
- Nelson, K., & Soulé, M. (1987). Genetical conservation of exploited fishes. In: *Population genetics and fisheries management* (Ryman, N., & Utter, F., eds). Washington: Washington Sea Grant.
- Pontier, D., Gaillard, J. M. & Allainé, D. (1993). Maternal investment per offspring and demographic tactics in placental mammals. *Oikos*, **66**, 424–430.
- Promislow, D. E. L., & Harvey, P. H. (1990). Living fast and dying young: a comparative analysis of life-history variation among mammals. *Journal of Zoology*, **220**, 417–437.
- Reznick, D. N., Rodd, F. H. & Cardenas, M. (1996). Life-history evolution in guppies (*Poecilia reticulata*: Poeciliidae). IV. Parallelism in life-history phenotypes. *The american naturalist*, **147**, 319–338.
- Ricker, W. E. (1981). Changes in the average size and average age of Pacific Salmon. *Canadian Journal of Fisheries and Aquatic Sciences*, **38**, 1636–1656.
- Rijnsdorp, A. D. (1991). Changes in fecundity of female North Sea plaice (*Pleuronectes platessa* L.) between three periods since 1900. *ICES Journal of marine Science*, **48**, 253–280.
- Rochet, M. J., Cornillon, P. A., Sabatier, R. & Pontier, D. (2000). Comparative analysis of phylogenetic and fishing effects in life history patterns of teleost fishes. *Oikos*, **91**, 255–270.
- Rose, K. E., Clutton-Brock, T. H. & Guinness, F. E. (1998). Cohort variation in male survival and lifetime breeding success in red deer. *J. Anim. Ecol.*, **67**, 979–986.
- Schaffer, W. M., & Elson, P. F. (1975). The adaptive significance of variations in life history among local populations of atlantic salmon in North America. *Ecology*, **56**, 577–590.

- Sedinger, J. S., & Flint, P. L. (1995). Environmental influence on life-history traits: growth, survival and fecundity in black brant (*Branta bernicla*). *Ecology*, **76**, 2404–2414.
- Shine, R., & Iverson, J. B. (1995). Patterns of survival, growth, and maturation in turtles. *Oikos*, **72**, 343–348.
- Sinervo, B. (1990). The evolution of maternal investment in lizards: an experimental and comparative analysis of egg size and its effects on offspring performance. *Evolution*, **44**, 279–294.
- Skorping, A., Read, A. F. & Keymer, A. E. (1991). Life history covariation in intestinal nematods of mammals. *Oikos*, **60**, 365–372.
- Stearns, S. C. (1983). The influence of size and phylogeny on patterns of covariation among life-history traits in the mammals. *Oikos*, **41**, 173–187.
- Trippel, E. A. (1995). Age at maturity as a stress indicator in fisheries. *Bioscience*, **45**, 759–771.
- Trippel, E. A., Kjesbu, O. S. & Solemdal, P. (1997). Effects of adult age and size structure on reproductive output in marine fishes. In: *Early life history and recruitment in fish populations*, Vol. 21 of *Fish and Fisheries Series*, pp. 29–62. London: Chapman and Hall.
- Wanntorp, H. E. (1983). Historical constraints in adaptation theory: traits and non-traits. *Oikos*, **41**, 157–160.
- Western, D. (1979). Size, life history and ecology in mammals. *Journal of African Ecology*, **17**, 185–204.
- Winemiller, K. O., & Rose, A. (1992). Patterns of life-history diversification in North American fishes: implications for population regulation. *Canadian Journal of Fisheries and Aquatic Sciences*, **49**, 2196–2218.

A concentrated likelihoods

Recall that for models 1 and 3 the concentrated likelihood is

$$\mathcal{M} = \frac{1}{2} \sum_{g=1}^G \#(g) \ln \{ [BY - \theta]' D_g [BY - \theta] \} - \ln |B|. \quad (6)$$

The first partial derivatives are then

$$\frac{\partial \mathcal{M}}{\partial \mu_g} = -\text{card}(I) \frac{\mathbb{1}'_g D_g (BY - \theta)}{(BY - \theta)' D_g (BY - \theta)}, \quad (7)$$

$$\text{where } \mathbb{1}_g = (\cdots 1_{g(i)} \cdots)' \quad (8)$$

$$\frac{\partial \mathcal{M}}{\partial \rho^{(b)}} = - \sum_{g=1}^G \#(g) \frac{Y' W^{(b)'} D_g (BY - \theta)}{(BY - \theta)' D_g (BY - \theta)} + \text{tr} \left\{ (BY - \theta)^{-1} W^{(b)} \right\} \quad (9)$$

$$\frac{\partial \mathcal{M}}{\partial \gamma} = -\#(g) \frac{L' D_g (BY - \theta)}{(BY - \theta)' D_g (BY - \theta)}. \quad (10)$$

The partial derivative with respect to $\rho^{(w)}$ is obviously obtained from equation (9) by exchanging superscript (w) for (b) . The gradient of \mathcal{M} is then

$$\nabla \mathcal{M} = \left(\frac{\partial \mathcal{M}}{\partial \rho^{(b)}}, \frac{\partial \mathcal{M}}{\partial \rho^{(w)}}, \cdots, \frac{\partial \mathcal{M}}{\partial \mu_g}, \cdots, \frac{\partial \mathcal{M}}{\partial \gamma} \right). \quad (11)$$

For the model 2 we first set partial derivatives of the log-likelihood with respect to μ_g equal to zero leading to

$$\tilde{\mu}_g = \frac{1}{\#(g)} [Y' B' - \gamma' L'] \mathbb{1}_g. \quad (12)$$

Thus we obtain the concentrated likelihood for model 2

$$\mathcal{M} = -\ln |B| + n \ln \sigma + \frac{1}{2\sigma^2} [BY - (\tilde{\mu} + L\gamma)]' [BY - (\tilde{\mu} + L\gamma)],$$

where $\tilde{\mu}$ is a vector of n coordinates where the i^{th} is $\tilde{\mu}_{g(i)}$.

Partial derivatives of \mathcal{M} are then

$$\begin{aligned} \frac{\partial \mathcal{M}}{\partial \sigma} &= \frac{n}{\sigma} - \frac{1}{\sigma^3} (BY - \theta)' (BY - \theta), \\ \frac{\partial \mathcal{M}}{\partial \rho^{(b)}} &= \text{tr}(B^{-1} W^{(b)}) - \frac{1}{\sigma^2} \left(W^{(b)} Y + \frac{\partial \tilde{\mu}}{\partial \rho} \right)' (BY - \theta), \\ \frac{\partial \mathcal{M}}{\partial \gamma} &= -\frac{1}{\sigma^2} \left[L + \frac{\partial \tilde{\mu}}{\partial \gamma} \right]' [BY - \theta]. \end{aligned}$$

We can exchange the superscript $\rho^{(w)}$ for $\rho^{(b)}$ to get the missing partial derivative. The gradient is defined by

$$\nabla \mathcal{M} = \left(\frac{\partial \mathcal{M}}{\partial \rho^{(b)}}, \frac{\partial \mathcal{M}}{\partial \rho^{(w)}}, \frac{\partial \mathcal{M}}{\partial \sigma}, \frac{\partial \mathcal{M}}{\partial \gamma} \right). \quad (13)$$

B steepest descent algorithm

To implement our method we need to estimate for each model the parameters Φ . For model 1, Φ is $(\rho^{(b)}, \rho^{(w)}, \mu, \sigma, \gamma)$; for model 2, Φ is equal to $(\rho^{(b)}, \rho^{(w)}, \dots, \mu_{g(i)} \dots, \sigma, \gamma)$ and for model 3 it is $(\rho^{(b)}, \rho^{(w)}, \dots, \mu_{g(i)} \dots, \dots, \sigma_{g(i)} \dots, \gamma)$. To obtain this estimation we minimize the function $\mathcal{M}(\phi)$ (equation (6) for models 1 and 3 or equation (13) for model 2). The value $\hat{\phi}$ which achieves the minimum is the estimation of ϕ . To get the estimation of the missing parameters (σ for model 1, $\dots, \mu_{g(i)} \dots$ for model 2 and $\dots, \sigma_{g(i)} \dots$ we use equation (4) with $g = 1$, (12) or (4) respectively. This procedure is summarized in the table (3). To obtain $\hat{\phi}$, we use the following minimization iterative algorithm :

steepest descent algorithm for model 1 or 3

Require: $I_n - \rho^{(b)(k)} W^b - \rho^{(w)(k)} W^w$ non singular $\forall k$ {That algorithm stop whenever this condition is not satisfied anymore}

Search: $\hat{\phi} = \arg \max_{\phi} \mathcal{M}(\phi)$ {the estimation of the coefficient of the chosen model}

$\phi^{(0)} \leftarrow (\rho^{(b)(0)}, \rho^{(w)(0)}, \dots, \mu_{g(i)}^{(0)} \dots, \gamma^{(0)})$ { $2 + g + r$ initial values, chosen by the user}

$H^{(0)} \leftarrow I_{2+g+r}$ {Common choice for the starting approximation of the Hessian}

δ and ν {absolute function and parameters convergence tolerances}

while $|\mathcal{M}(\phi^{(k+1)}) - \mathcal{M}(\phi^{(k)})| \leq \delta$ or $\max |\phi^{(k+1)} - \phi^{(k)}| \leq \nu$ **do**

$\phi^{(k+1)} \leftarrow \phi^{(k)} - s^{(k)} H^{(k)-1} \nabla \mathcal{M}'$ {updating the parameters for the steepest descent of \mathcal{M} , where $\nabla \mathcal{M}$ is the gradient of \mathcal{M} , see appendix A, $s^{(k)}$ is the step size chosen (with Armijo's rule for instance) and $H^{(k)}$ is an approximation of the Hessian}

$H^{(k+1)} \leftarrow H^{(k)} - \frac{\nabla \mathcal{M}' \nabla \mathcal{M}}{\nabla \mathcal{M} H^{(k)-1} \nabla \mathcal{M}}$ {updating the approximation of the Hessian}

end while

$\hat{\sigma}_g^2 = [By - \theta]' D_g [BY - \theta] / \#(g)$ {equation (4)}

Calculus of $\widehat{Var}(\hat{\phi}) \approx I(\hat{\phi})^{-1}$ {where $I(\hat{\phi})$ is the Fisher information matrix}

The algorithm is the same for model 2 except that ϕ is now $(\rho^{(b)}, \rho^{(w)}, \sigma, \gamma)$ and $\mathcal{M}(\phi)$ is defined by equation (13). In this case, the computation of $\hat{\sigma}_g^2$ at the end of the while loop is replaced by the computation of $\hat{\mu}_g$ using the equation (12). We do not show the equation of the Fisher information matrix because of its length. The step size can be chosen as constant $s^{(k)} = s \forall k$, or as optimal step size $s^{(k)} = \arg \min_{s>0} \mathcal{M}(\phi^{(k)}) - \mathcal{M}(\phi^{(k)} - sH^{(k)-1} \nabla \mathcal{M}')$ or by using a method to select it : Armijo's rule, Goldstein or Wolfe methods (Hager et al., 1993). Recall that the implementation of a steepest gradient algorithm is straightforward in mathematical or statistical softwares such as `Splus`, `matlab`, `mathematica` etc... `Splus` programs are available from the first author.

Models	Am	Lm	T ₀₅	Fb
A) 2 factor model: $Y_i = \gamma L_{05} + \mu_{g(i)}$				
L ₀₅	89.32***	1361.18***	70.57***	
Fishing	2.29(NS)	6.14**	21.40***	1.86(NS)
B) 3 factor model: $Y_i = \gamma L_{05} + \mu_{g(i)} + \text{Order}_i$				
L ₀₅	130.98***	1537.40***	87.29***	
Order	5.64***	1.68(NS)	4.79***	3.26**
Fishing	1.81(NS)	8.93***	19.93***	3.00(NS)
C) 3 factor model with interaction: $Y_i = \gamma L_{05} + \mu_{g(i)} + \text{Order}_i + \mu_{g(i)} * \text{Order}_i$				
L ₀₅	137.18***	1818.99***	89.80***	
Order	5.91**	1.99(NS)	4.93***	3.17**
Fishing	1.89(NS)	10.57***	20.50***	2.92(NS)
Interaction	1.34(NS)	2.32(NS)	1.21(NS)	0.79(NS)

Table 1: F-values for linear models fitted to life-history traits: A) with two factors (non phylogenetic approach). B) with three factors (rough phylogenetic approach). C) with three factors with interaction (rough phylogenetic approach with interaction). NS: $P > 0.05$; *: $0.05 > P > 0.01$; **: $0.001 < P < 0.01$; ***: $P < 0.001$.

Models	Am	Lm	T ₀₅	L ₀₅	Fb
"Between-Within" Models: $\rho^{(b)}$ and $\rho^{(w)}$					
(1)	15.6146	98.5432	8.6155	3.6635	-29.9463
(2)	23.7192	104.5312	27.3791	5.0536	-25.3097
(3)	26.4237	113.4313	27.6976	7.9291	-20.1064
"Between": $\rho^{(w)} = 0$					
(1)	18.3908	98.5432	6.2175	-10.9749	-31.6265
(2)	20.7376	104.5312	23.9823	-9.5576	-29.2779
(3)	24.1296	113.4313	23.5953	-6.5538	-21.3669
"Independent observations": $\rho^{(b)} = 0$ and $\rho^{(w)} = 0$					
(2)	20.7376	104.5312	23.9823	-12.2362	-30.5127
(3)	24.1296	113.4313	24.4058	-8.4790	-22.9660

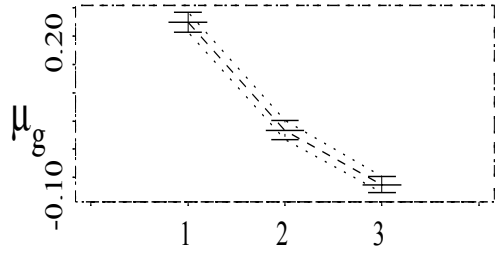
Table 2: Log-likelihood of the nested autoregressive models, (3) full model: $\mu_1 \cdots \mu_G$, $\sigma_1 \cdots \sigma_G$; (2) $\sigma_I = \sigma \forall I$; (3) $\mu_I = \mu \forall I$ and $\sigma_I = \sigma \forall I$; threshold value for an increase of 2 parameters $\chi_{0.95}^2(2) \approx 3$ or 1 parameter $\frac{1}{2}\chi_{0.95}^2(1) \approx 1.95$. Selected model in **bold** font.

	Model 1	Model 2	Model 3
Models Parameters Φ (all the parameters)	$(\rho^{(b)}, \rho^{(w)}, \mu, \sigma, \gamma)$	$(\rho^{(b)}, \rho^{(w)}, \dots, \mu_{g(i)} \dots, \sigma, \gamma)$	$(\rho^{(b)}, \rho^{(w)}, \dots, \mu_{g(i)} \dots, \dots, \sigma_{g(i)} \dots, \gamma)$
Parameters ϕ of the concentrated likelihood	$(\rho^{(b)}, \rho^{(w)}, \mu, \gamma)$	$(\rho^{(b)}, \rho^{(w)}, \sigma, \gamma)$	$(\rho^{(b)}, \rho^{(w)}, \dots, \mu_{g(i)} \dots, \gamma)$
Concentrated Likelihood $\mathcal{M}(\phi)$	equation (6) with $G = 1$	equation (13)	equation (6)
Estimation of ϕ Gradient $\nabla \mathcal{M}$	Steepest descent equation (11)	Steepest descent equation (13)	Steepest descent equation (11)
Calculus of missing parameters	equation (4) with $G = 1$	equation (12)	equation (4)

Table 3: Parameters to be estimated and equation number for models 1,2 and 3.

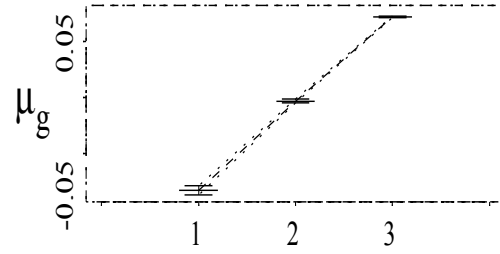
Figure 1: Fishing effect on demographic traits: estimated μ_g and its asymptotic confidence interval for the groups of fishing pressure (1: low fishing mortality, 2: moderate fishing pressure, 3: high fishing pressure).

Am



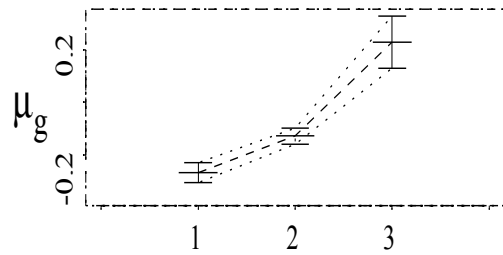
Environmental groups: g

Lm



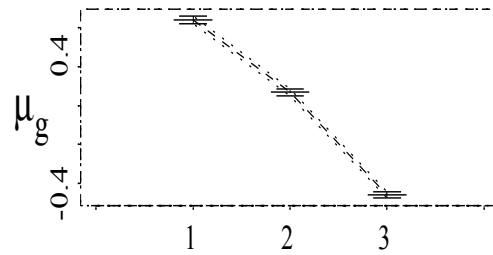
Environmental groups: g

Fb



Environmental groups: g

T₀₅



Environmental groups: g