Reevaluating Europe's other debt with improved statistical tools

Frédéric GOSSELIN

Irstea, UR EFNO, F-45290 Nogent-sur-Vernisson, France.

Correspondence author. Telephone: +33 2 38 95 03 58. Fax: +33 2 38 95 03 59

E-mail: frederic.gosselin@irstea.fr

Accepted for publication in Biodiversity and Conservation.

The final publication is available at Springer via

http://dx.doi.org/10.1007/s10531-014-0799-y

Abstract:

Dullinger et al. (PNAS, 2013) showed that current proportions of threatened species in European countries are more correlated to socio-economic pressures calculated a century ago than to current socio-economic pressures. These results could have important impacts on future political decisions, yet the statistical analyses used may be strongly impacted by pseudo-replication and over-dispersion of data.

I reanalysed Dullinger et al.'s data using generalised linear mixed models accounting for pseudo-replication and over-dispersion.

These new statistical models indicated that socio-economic pressures had a much less clear impact on proportions of threatened species than indicated by Dullinger et al. In many cases, the effects of socio-economic pressures even vanished.

Dullinger et al. (2013)'s results are actually much less "considerable" than their paper implies and therefore should be viewed with caution. Ecologists should more often incorporate pseudo-replication and overdispersion in their statistical analyses.

Keywords: pseudo-replication; overdispersion; binomial distribution; statistical model; generalized linear mixed models; biodiversity; indicator; threatened species; socio-economic pressures

Introduction

Dullinger et al. (2013) showed that current proportions of threatened species in 22 European countries respond more to socio-economic pressures calculated a century ago than to current socio-economic pressures. Based on generalised linear mixed models (GLMMs), the authors found very clear discrepancies among the statistical models involving socio-economic pressures in 1900, 1950 and 2000 – with the 1900 model being the best. They also found fixed-effect estimators that were very significantly different from 0 for all three models (cf. their Table 1). They concluded that "irrespective of recent conservation actions, large-scale risks to biodiversity lag considerably behind contemporary levels of socioeconomic pressures".

These results could have important impacts on future political decisions, yet I have serious doubts about the way Dullinger et al. analysed their data. To be brief, they completely overlooked the possible influence of pseudoreplication and overdispersion. Regarding pseudo-replication, although the same socioeconomic variables were repeatedly applied to the seven taxonomic groups in each country, the statistical models did not include a random effect for country. This means that the data was analysed as if the socio-economic variables came from completely different countries. Concerning over-dispersion, after having taken into account the effect of socio-economic pressures and variations in the intercept among taxa, the proportion of threatened species might not be distributed like a classical binomial distribution but instead, may be over-dispersed compared to the binomial distribution. This is simply because substantial variations in the ratio of threat among countries – for single taxon analyses – or among taxa within countries – for multi-taxa analyses – might still remain even after having modeled these effects. Not taking into account pseudo-replication and overdispersion in the model is known to have deleterious effects on the standard errors of the estimates (Millar and Anderson 2004; Warton and Hui 2011), and thus on statistical tests. It also has some impact on the model comparison criteria.

Material and methods

Using the data published as supplementary material by Dullinger et al. (2013) complemented by data on the number of extinct species kindly supplied by S. Dullinger, I refitted the binomial GLMMs used by Dullinger et al. to analyze the link between the proportion of threatened species and the three socio-economic pressures – human population density, per capita GDP and human appropriation of net primary productivity –, quantified either in 1900, 1950 or 2000, and synthesized by a principal component analysis. I used R3.1.1 and the R libraries lme4 and AICcmodavg to fit and analyse both the multi-taxa and single taxon models. This was done in connection with S. Dullinger. While working with the Supplementary data, we realized that there was a mistake in the published version: the human appropriation of net primary productivity for Romania in 1900 was 0.279 and not 0.297. At the multi-taxa level, I first refitted Dullinger et al.'s models – except that I included grasshoppers, which Dullinger et al. had not done due to a coding error (Dullinger, personal communication). To allow for overdispersion and take pseudo-replication into account, I then fitted a new model by adding a random Country effect to the intercept and included a normally distributed random intercept (e.g. Warton and Hui 2011). At the single-taxon level, I similarly refitted Dullinger et al.'s models and also created new models incorporating a random Country effect to allow for over-dispersion. In addition, in both cases, I fitted a null model, with no fixed effect other than the intercept, to not only compare the socio-economic models among themselves but also with a model with no such effect. I interpreted the models using some of Dullinger et al.'s methods, by presenting the estimators and their confidence intervals of the fixed effects (as well as of the standard deviations of the random effects), the corrected Akaike Information Criterion (AICc) and the associated model weights (AW) for the four multi-taxa models and a graphical representation of the Akaike weights of the four models for each of the single-taxon model. AICc was calculated based on the degrees of freedom as estimated from the lmer object, as did Dullinger et al. This involves what a Bayesian would call hyperparameters, i.e. fixed effects and variance parameters. I acknowledge that this choice has not yet been stabilized for mixed models (e.g. Spiegelhalter et al. 2002). I therefore also report the log-likelihood of each model to allow the reader to compute other model selection criteria.

Results

The standard deviation of the added random effects in the multi-taxa models turned out to be between 0.34 and 0.53. These added levels of variation had a considerable effect on the results (cf. Table 2), which proved to be much less clearcut than those given by Dullinger et al. (cf. Table 1). Even if the best model remained the one with the 1900 socio-economic indicators, the Akaike weights were much more evenly spread among models and the Akaike weight of the null model (0.26) was non negligible. Also, fewer estimators were significantly different from zero.

For the single-taxon models, there was a systematic shift between Dullinger et al.'s best model, which included one set of socio-economic indicators, and the best model when over-dispersion was incorporated, which included no socio-economic indicators (Figure 1).

Discussion

Some readers might mainly view the new multi-taxa results as being qualitatively the same as in the original publication. This line of interpretation would be acceptable if the main message to scientists and to decision makers in our scientific studies were merely qualitative. However, such an attitude could relegate science to the story telling realm. Quantitative aspects are also very important. Granted, with the original analyses, we have no doubt about which model is the best of the four fitted models, nor do we doubt the significance of the effects in these models. Yet, as I explained in the introduction, these analyses are not correct. There is other evidence to indicate that the original models have quantitative problems. Firstly, when we simulate data with levels of over-dispersion and pseudo-replication as estimated with the new methods and with no effect of the socio-economic pressures, and when we analyse these data with Dullinger et al.'s methods, we get results that are very much in line with those we obtained in Table 1 in terms of significance of the effects and of the distribution of Akaike weights. Another clue indicating there are some statistical problems with the original models is that the error in the data for net primary productivity for Romania in 1900 indicated in the material and methods section (0.297 instead of 0.279) – which is a rather minor error with respect to the range of variation for this variable – yielded an AICc shift of more than 30 units – a huge difference!

The new statistical analyses, taking into account both pseudo-replication and over-dispersion indicated that the results are much less clear cut, leaving greater room for doubt as to which one is the best model. This doubt increases further when we consider the single-taxon analyses, all of which indicated that the best model of the four was the one without any socio-economic indicator.

I therefore conclude from the re-analyses of Dullinger et al's data that their results are actually much less "considerable" than their paper implies and should be viewed with caution. I would also like to remind ecologists of a a classical warning (Hurlbert 1984; White and Bennetts 1996; Millar and Anderson 2004; Noe et al. 2010; O'Hara and Kotze 2010; Warton and Hui 2011): great care should be taken to incorporate sources of variation corresponding to pseudo-replication and over-dispersion into statistical analyses of count data.

Acknowledgements. I am grateful to Vicki Moore for having checked the English, Frédéric Archaux for re-reading a first version of this manuscript, an anonymous reviewer for helping improve the ppaper and Stefan Dullinger for exchanges about data analysis. This research was supported by the French ministry in charge of Ecology through the DEB-Irstea convention (Action S).

References:

Dullinger S, Essl F, Rabitsch W, Erb KH, Gingrich S, Haberl H, Hülber K, Jarošík V, Krausmann F, Kuhn I, Pergl J, Pyšek P, Hulme PE (2013) Europe's other debt crisis caused by the long legacy of future extinctions. Proc Nat Acad Sciences USA 110:7342-7347. doi: 10.1073/pnas.1216303110

Hurlbert SH (1984) Pseudoreplication and the design of ecological field experiments. Ecol Monogr 54:187-211.

Millar RB, Anderson MJ (2004) Remedies for pseudoreplication. Fish Research 70:397-407. doi: 10.1016/j.fishres.2004.08.016

Noe DA, Bailer AJ, Noble RB (2010) Comparing methods for analyzing overdispersed count data in aquatic toxicology. Envir Toxicol Chemis 29:212-219. doi: 10.1002/etc.2

O'Hara RB, Kotze DJ (2010) Do not log-transform count data. Meth Ecol Evol 1:118-122. doi: 10.1111/j.2041-210X.2010.00021.x

Spiegelhalter DJ, Best NG, Carlin JB, van der Linde A (2002) Bayesian measures of model complexity and fit (with Discussion). J Roy Statist Soc B, 64:583-616. doi: 10.1111/1467-9868.00353

Warton DI, Hui F (2011) The arcsine is asinine: the analysis of proportions in ecology. Ecology 92:3-10. doi: 10.1890/10-0340.1

White GC, Bennetts RE (1996) Analysis of frequency count data using the negative binomial distribution. Ecology 77:2549-2557.

Table 1. Results of Dullinger et al. (2013)'s statistical models, not accounting for pseudo-replication among countries and over-dispersion, of proportions of threatened species from seven taxonomic groups and 22 European countries as explained –or not – by socioeconomic indicators calculated at different dates.

	Null model	1900 model	1950 model	2000 model
Estimate for PC1	-	0.206 ([0.193;0.218]) ****	0.159 ([0.146;0.174]) ****	0.110 ([0.095;0.122]) ****
Estimate for PC2	-	0.139 ([0.116;0.163) ****	0.040 ([0.022;0.059]) ***	-0.058 ([-0.075;-0.040]) ***
Estimate for PC3	-	-0.346 ([-0.381;-0.311]) ****	0.181 ([0.143;0.219]) ****	0.124 ([0.075;0.173]) ***
Standard deviation of Taxon random effect	0.276 ([0.095;0.404])	0.306 ([0.113;0.429])	0.284 ([0.113;0.399])	0.280 ([0.104;0.422])
Log-likelihood	-3463.9	-2770.5	-3173.9	-3328.4
AICc	6931.8	5551.4	6358.3	6667.3
AW	$<10^{-9}$	1	$<10^{-9}$	$<10^{-9}$

Fixed effects (omitting the intercept), standard deviation of random effects, log-likelihood, corrected Akaike information criterion (AICc), and Akaike weights (AW) for binomial generalized mixed-effects models not allowing for pseudoreplication among countries and overdispersion relating the variation in the proportion of red-listed species to either no effect or historical (1900 and 1950) or contemporary (2000) values of socioeconomic indicators. These indicators – human population density, per capita GDP and human appropriation of net primary productivity – were subjected to a principal component analysis before fitting the regression models, yielding three uncorrelated variables (PC1, PC2 and PC3). Estimates include the mean estimate, the 95% confidence interval estimated with the standard bootstrap method (confint function in library lme4) in parentheses and the level of significance of difference from 0 at the end (* for $P<0.05$, ** for $P<0.01$, *** for $P<0.001$ and **** for $P<10^{-9}$). AICc and AW were calculated with the aictable function in the AICcmodavg R library. The results are slightly different from those in Table 1 in Dullinger *et al.* (2013) because the analyses here include the data for grasshoppers, which were left out of the analyses published in Dullinger et al. (2013) (S. Dullinger, personal communication).

Table 2. Results of statistical models accounting for pseudo-replication among countries and over-dispersion, of proportions of threatened species from seven taxonomic groups and 22 European countries as explained –or not – by socioeconomic indicators calculated at different dates, in.

	Null model	1900 model	1950 model	2000 model
Estimate for PC1	-	0.209 ([0.086;0.327]) **	0.187 ([0.042;0.335]) **	0.169 ([0.029;0.308]) *
Estimate for PC2	-	-0.007 ([-0.242;0.243])	-0.054 ([-0.279;0.155])	-0.064 ([-0.261;0.150])
Estimate for PC3	-	0.116 ([-0.196;0.456])	-0.208 ([-0.608;0.169])	-0.222 ([-0.710;0.297])
Standard deviation of Taxon random effect	0.342 ([0.000;0.468])	0.306 ([0.000;0.472])	0.303 ([0.000;0.446])	0.302 ([0.000;0.465])
Standard deviation of Country random effect	0.460 ([0.274;0.617])	0.346 ([0.021;0.469])	0.368 ([0.090;0.465])	0.386 ([0.074;0.494])
Standard deviation of Country*Taxon random effect	0.525 ([0.427;0.633])	0.527 ([0.423;0.630])	0.529 ([0.421;0.627])	0.529 ([0.431;0.624])
Log-likelihood	-554.9	-550.4	-551.3	-552.0
AICc	1118.1	1115.6	1119.0	1117.5

AW	0.156	0.536	0.207	0.101

Fixed effects (omitting the intercept), standard deviations of random effects, log-likelihood, corrected Akaike information criterion (AICc), and Akaike weights (AW) for binomial generalized mixed-effects models allowing for pseudo-replication and over-dispersion relating the variation in the proportion of red-listed species to either no effect or historical (1900 and 1950) or contemporary (2000) values of socioeconomic indicators. Notations and methods are as in Table 1. AICc values should be compared only among models in Table 2, not with models from Table 1 which have different types of random effects. Akaike Weights were much more evenly spread among models and fixed effects were much less significant than in Table 1.

Figure 1.a

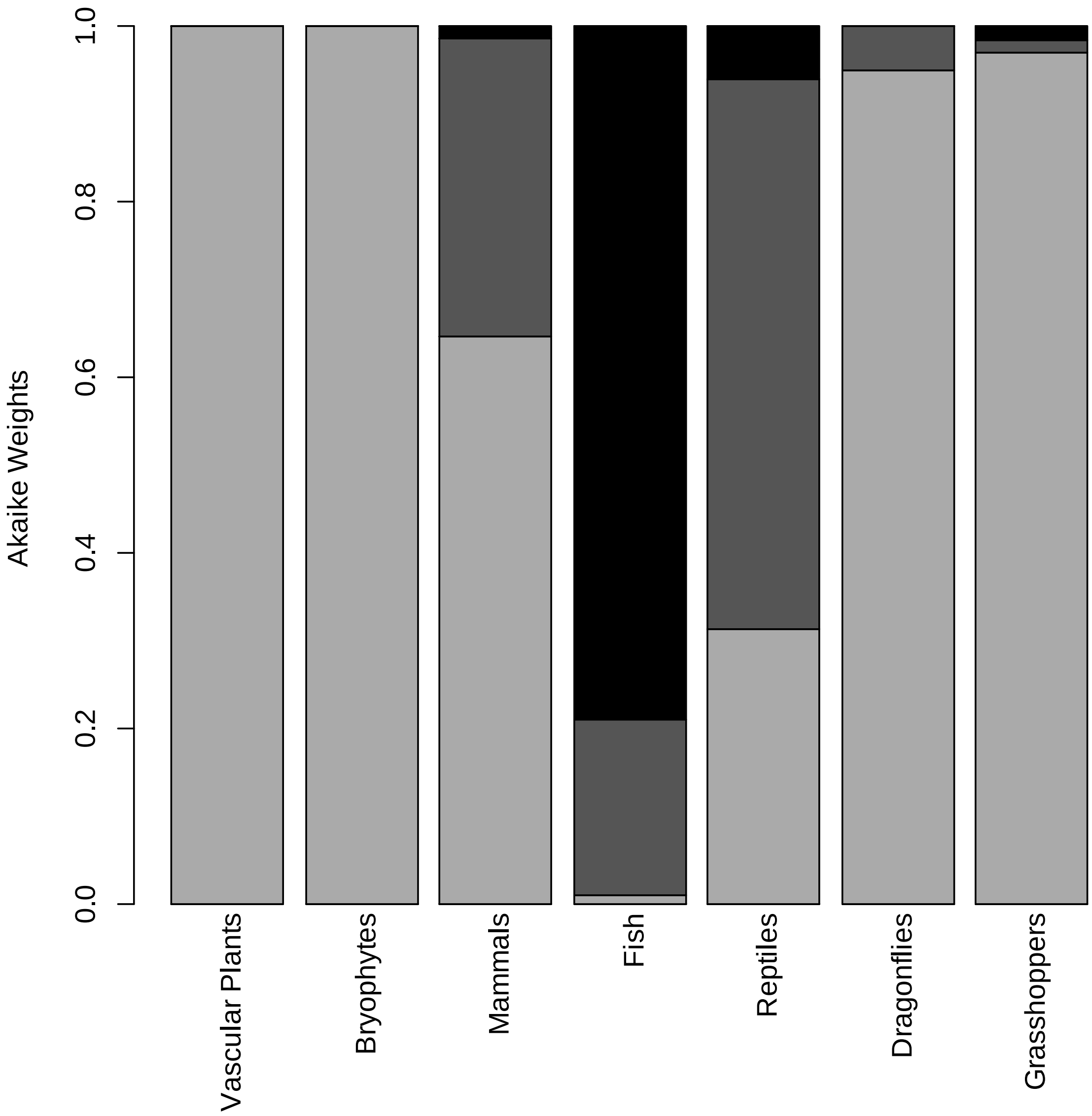
Akaike Weights
0.0
0.2
0.4
0.6
0.8
1.0
Vascular Plants
Bryophytes
Mammals
Fish
Reptiles
Dragonflies
Grasshoppers

Figure 1.b

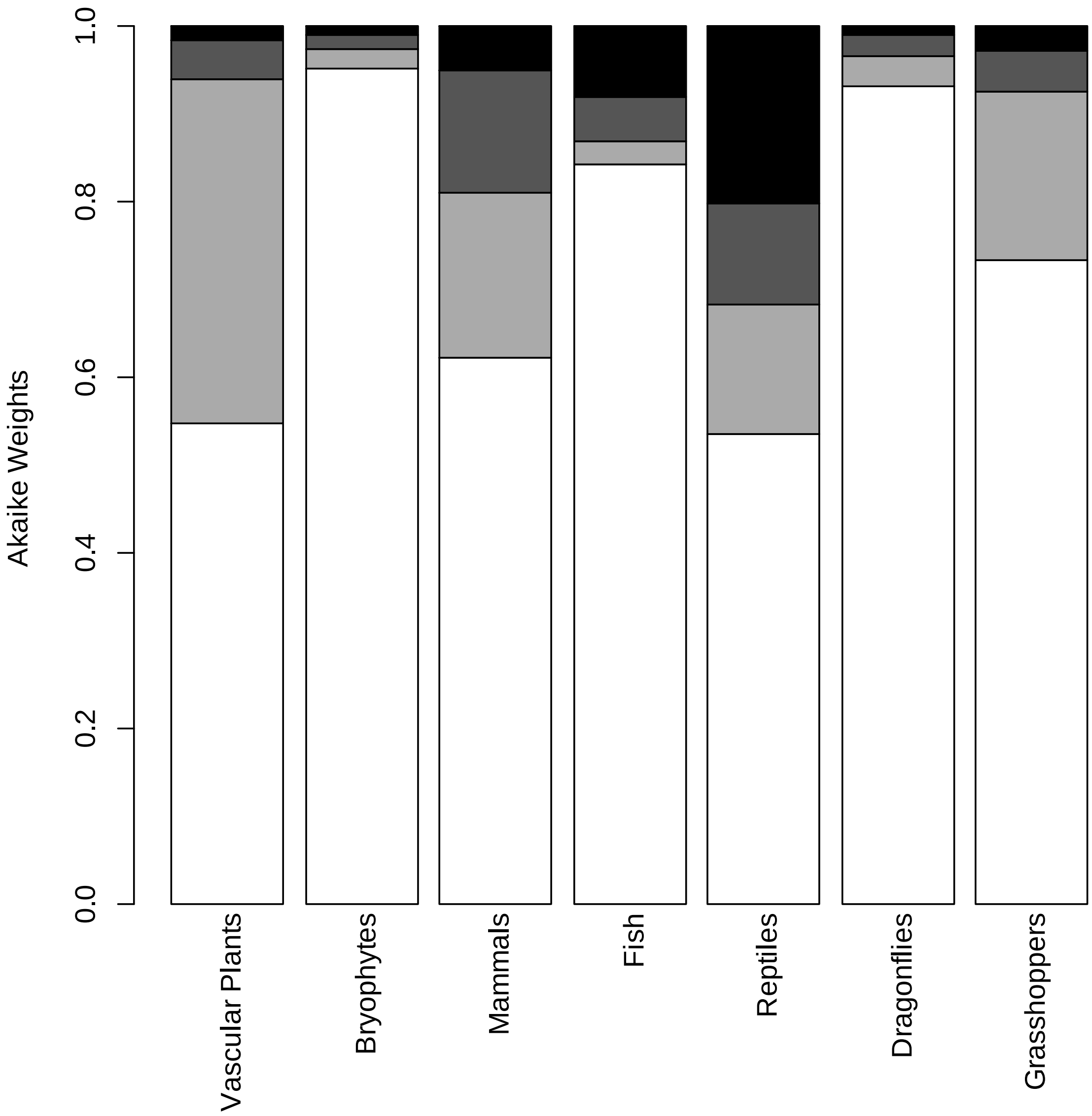
Akaike Weights
0.0
0.2
0.4
0.6
0.8
1.0
Vascular Plants
Bryophytes
Mammals
Fish
Reptiles
Dragonflies
Grasshoppers

Figure 1. Relative support (Akaike Weights) for models explaining the proportion of red-listed species in each of the seven taxonomic groups in 22 European countries. The models incorporated either only an intercept (white coulor), or historical (1900, light gray; 1950, dark gray) or contemporary (2000, black) levels of socio-economic impact indicators (human population density, per capita GDP, human appropriation of net primary productivity). (a) Statistical models that did not model overdispersion, as in Dullinger et al. (2013); (b) Statistical models including over-dispersion.