Responses to reviewers

Ms. No. ESPR-D-15-00741R1
submitted to
Environmental Science and Pollution Research

Eduard Szöcs and Ralf B. Schäfer

April 17, 2015

Dear editor Dr. Schulz and reviewers,

We are thankful for reviewing our manuscript a second time and the comments that helped to improve the paper. We revised the manuscript accordingly and are re-submitting the manuscript for consideration for publication in *Environmental Science and Pollution Research*.

In the remainder of this document, we describe the changes that we have made to the paper for resubmission. To assist the assessment of our changes we have submitted two versions of the revised manuscript: one with highlighted changes (compared to revision 1) and another without any highlighting. Note, that we did not highlight changes in citations and figures.

Kind regards,
Eduard Szöcs and Ralf B. Schäfer

Response to Reviewer 2

Comment 1: *"One additional point is that Tony Ives has an in press paper at Methods in Ecolgoy and Evolution on a similar topic - arguing that LM more reliably maintains nominal Type I error levels than GLM for count data, and that this is an argument in defence of transform-LM (similar to ter Braak and Smilauer 2014). This should probably get a mention."*

Response: We are thankful for pointing to this recently accepted paper. We picked it up in the discussion. See also comments 17-19.

Comment 2: *"p1 col1 l27 allow one to directly model(?)"*

Response: We fixed this sentence. It now reads :
"Generalised Linear Models (GLM) allow directly model such data, without the need for transformation".

Comment 3: *"p1 col1 l46 extremely; p2 col2 l1 for more than 40; p2 col1 l7 Warton 2005 was about counts not proportions."*

Response: We fixed these typos.

Comment 4: *"p2 col1 l25 may enhance..., when appropriately used [to reflect the change of emphasis requested to caution about misuse, suggested by reviewers 1 and 3]"*

Response: We agree and changed accordingly.

Comment 5: *"equation 1: superscript T is not the best choice, this is standard notation for a matrix transpose. y_new might be worth a shot..."*

Response: We agree and changed accordingly to $y_{new\ i}$.

Comment 6: *"equation 2-3: β_Treatment_i is awkward notation."*

Response: We agree and changed notation to βx_i.

Comment 7: *"p2 col2 l21 Poisson not poisson"*

Response: We fixed this typo.

Comment 8: *"Section 2.2.2 Reviewer 3 requested a statement of the underlying assump-*

tion that each of the units being counted is iid, which I could not see in this section. This connects to the topic of overdispersion (which arises when not iid)"

Response:

Comment 9: *"p4 bottom col1: Type I error and power at what significance level."*

Response: We added *"at a significance level of* $\alpha = 0.05$".

Comment 10: *"p4 col2 l18: considerably higher; p4 col2 l21: led to; p4 col2 l50 the parameteric bootstrap"*

Response: We fixed these typos.

Comment 11: *"Fig 2: Type I error off the scale is undesirable, the point that Type I error is poor is harder to see when you can't see it. Maybe use a log-scale for Type I error and power?"*

Response:

Comment 12: *"p5 col1 l47: Type I not Type 1, happens elsewhere too"*

Response: We fixed this throughout the manuscript.

Comment 13: *"p7 col1 line 60: residual vs fits plots can also be very informative (e.g. Wang et al 2012)"*

Response: We added residual vs. fitted values plot.

Comment 14: *"p8 col 1 line 7 delete "to"; p8 col2 line 1 add space after 2002)"*

Response: We fixed both typos.

Response to Reviewer 3

Comment 15: *"1. take the sentence in the abstract: "Generalised Linear Models (GLM) allow directly model distributions fitting such data." which cannot be understood, neither by an ecologists nor by a statistician (see further under Language) and "*

Response: We agree and rephrased. See also comment 2.

Comment 16: *"2. eqs 2-5 & 7 where the authors cannot get their math right. The paper need a lot of editing both linguistically and statistically."*

Response: We edited the equations, see also comments 5 and 6.

Comment 17: *"3. Also the paper fails to indicate the trade-off between model and computational complexity, the potential gain in, for example, power and (loss/gain) in control of the type I error. For example, what is the gain of using the npb (where does this abbreviation come from??) over the much simpler qp method, and of the qp method over LM on transformed data? "*

Response: We agree and compared the gains. See also comment 1 + 19.

Comment 18: *"4. Some summary measures of gain should be included and "*

Response: See comments 1, 17, 19.

Comment 19: *"5. an overall conclusion in favour of the qp method should be drawn. "*

Response: We agree and after comparison of gains we draw an overall conclusion on GLM_{qp}. However, this is only valid for one-factorial designs - as GLM_{qp} showed increased Type I errors in multiple regression (Ives, 2015). See also comment 1.

Comment 20: *"6. The analysis of LOEC is very inconsistent and should be redone/reconsidered. The reason is that authors claim that the Williams test is easily applied in GLM context (p7,44-48,l), but not used at all. So why is the Williams not used in the simulations? It likely gives a much higher increase in power than any of model comparison performed in the paper."*

Response: We added reference to Hothorn et al (2008) for a Williams-type multiple contrast test in a GLM framework. Moreover, we added justification for the use of Dunnett contrasts. The section now reads:
"The choice of transformation contributed only little to the differences. If the assumptions of Williams test are met it has strictly greater power than Dunnett contrasts (Jaki and Hothorn, 2013), which explains the differences in the case study. A generalisation of the Williams test as multiple contrast test (MCT) can be used in a GLM framework (Hothorn et al, 2008). Nevertheless, such a Williams-type MCT is not a panacea (Hothorn, 2014) and our simulated semi-concave dose-response relationship is a situation where it fails and underestimates the LOEC (Kuiper et al, 2014). "

Comment 21: *"The real reason why GLMs are great is beyond the scope of experiments analysed in this paper. The real advantage of GLMs is that they allow separate specification of the distribution of the response variable and of the scale on which effects are additive. Because they are just simple means in the models in the paper and nothing what requires additivity or linearity on some scale, this key advantage falls outside the scope of the paper. Please tell something of this sort in the intro or the discussion!"*

Response:

Comment 22: *"Please also mention that the quasi-likelihood approach to GLMs in which it are not the distribution of the response variable that is key to the method, but the mean-variance relationship (this relates to comments 7 and 26)."*

Response: This is already mentioned in eqn. 4 and accompanying text.

Comment 23: *"Language: There is a tendency of stenography: applying least-squares methods (by the way, a term not used!!) after data transformation is described as data transformation or as transform the data (in the abstract on 44L and 25L). Brevity is nice but it should remain understandable. Another example:"Nevertheless, they are often analysed using methods assuming a normal distribution and variance homogeneity". Who assumes what in this sentence. A method does not assume anything (the user does, and the method is guaranteed to have some properties when the assumptions hold true.) and"They" refers to data which cannot assume anything either. There are many of these misconstructions. "*

Response:

Comment 24: *"My previous comment (in comment 39): "(3) Without the use of a GLM equivalent of the Williams test all the advantage of the use of GLM in terms of power are gone. See the example. Discuss this ambiguity. You can perhaps use a bootstrap test based on (GLM?) monotonic regression or similar. I know some cues/leads in this direction." has led to (unverified) statements on the Williams test without implementing the test. See general, point 6."*

Response: See response to comment 20.

Comment 25: *"P3,49l. Add (y_i) after number of occurrences, otherwise y_i undefined (or number of occurrences?!)."*

Response: We agree and clarified this section.

Comment 26: *"P3,58L Delete: However. "*

Response: We agree and changed accordingly.

Comment 27: *"P3,58L where can I see the beta is "parameters""*

Response: We fixed this typo and changed to *"coefficients"*.

Comment 28: *"P4,L Rephrase sentences with "kept equal""*

Response: We agree and rephrased these two sentences.

Comment 29: *"P4, 55, R. qp is not mention in remarks onType I error. Why not?"*

Response: We added LM and GLM_{qp} to this section.

Comment 30: *"Legend Fig2. Add inbetween "error are" (GLM_p and GLM_nb)"*

Response: We agree and changed accordingly.

Comment 31: *"Fig.3 Is it explained why npb is not in this figure?"*

Response: As stated in the methods section, we applied the parametric bootstrap only to the LR test.

Comment 32: *"P4,20,R And what is the estimated value of k for the case study. Now it cannot be verified that the simulations loosely mimic the case study."*

Response: We added the value of $\kappa = 3.91$.

Comment 33: *"P4,29R. Say here or in the discussion that this LR test turned out to be invalid as it has inflated Type I error. "*

Response: We agree and added this to the discussion.

References

Hothorn LA (2014) Statistical evaluation of toxicological bioassays – a review. Toxicol Res 3(6):418–432, DOI 10.1039/C4TX00047A, URL http://xlink.rsc.org/?DOI=C4TX00047A

Hothorn T, Bretz F, Westfall P (2008) Simultaneous Inference in General Parametric Models. Biometrical Journal 50(3):346–363

Ives AR (2015) For testing the significance of regression coefficients, go ahead and log-transform count data. Methods in Ecology and Evolution pp n/a–n/a, DOI 10.1111/2041-210X.12386, URL http://onlinelibrary.wiley.com/doi/10.1111/2041-210X.12386/abstract

Jaki T, Hothorn LA (2013) Statistical evaluation of toxicological assays: Dunnett or Williams test—take both. Archives of Toxicology 87(11):1901–1910

Kuiper RM, Gerhard D, Hothorn LA (2014) Identification of the Minimum Effective Dose for Normally Distributed Endpoints Using a Model Selection Approach. Statistics in Biopharmaceutical Research 6(1):55–66, DOI 10.1080/19466315.2013.847384, URL http://www.tandfonline.com/doi/abs/10.1080/19466315.2013.847384