


Generalized additive models to analyze biomedical non-linear longitudinal data in R:

Beyond repeated measures ANOVA and Linear Mixed Models

Response to reviewers

Ariel I. Mundo 

Department of Biomedical Engineering, University of Arkansas, Fayetteville, AR, USA

John R. Tipton 

Department of Mathematical Sciences, University of Arkansas, Fayetteville, AR, USA

Timothy J. Muldoon*

Department of Biomedical Engineering, University of Arkansas, Fayetteville, AR, USA

tmuldoon@uark.edu

1 General Comments

2 Comments from reviewers

2.1 Missing data

The title for Section 3.4 has been changed from “Missing observations” to “Unbalanced data”. We have restructured this section in order to:

- Avoid confusion by using the term “missing data”. We now refer to different number of observations as “unbalanced data” (L201-204).
- Make clear that LMEMs also can work with missing observations (L209)
- Emphasize that GAMs are not immune to missing data problems (L217-220)

2.2 Unnecessary restriction of LMMs

2.3 GLMMs

2.4 GLMs, GAMs, and conditional distributions

2.5 Thin plate regression splines

The following changes have been made:

- The information on L289-294 pertaining splines has been changed. L296-307 now give a better description of cubic splines (CS), indicating their limitation due to knot placement. The lack of knot placement for thin plate regression splines (TPRS) is indicated as an advantage over CS.

- L297-300 no longer refer to “knots” when referring to the construction of the smooth, now indicating that the number of *basis functions* is what is specified when working with TPRS. In the updated manuscript this appears in L310-315.

2.6 Penalised splines

2.7 Bayesian

- We have clarified the concept “empirical Bayesian”, which appeared on L318-L319 in the original manuscript. In the revised manuscript, L335-336 now indicate that Stan, JAGS or other probabilistic programming language can be used to estimate GAMs using a full Bayesian approach.
- The sentence “Moreover, the use of the restricted maximum likelihood (REML) to estimate the smoothing parameters gives an empirical estimate of the smooth model”, which appears in L319 in the original manuscript has been removed from the text. Instead, the concept of REML has been moved to Section 6.2 L390-393, where we state some of the reasons indicated by Wood when choosing restricted maximum likelihood (REML) over the default general cross validation (GCV) method for smooth parameter estimation in *mgcv*.

2.8 Coverage of confidence intervals

The content in L320-329 regarding confidence intervals (CIs) in the original submission has been changed in the following manner:

- L339-L352 now contain a more detailed and accurate explanation on the differences and interpretation of “pointwise” CIs and “across-the-function” CIs. We also reference the work of Marra and Wood if the reader desires a more in-depth exploration.

2.9 Differences in smooths

2.10 Appendix