wpg-mb methods

All statistical analyses were performed in R (version 4.0.4, R Core Team, 2021). The tidyr and dplyr packages Wickham *et al.* (2021) were used for data wrangling, while plots were created using the ggplot2 and cowplot packages Wilke (2020). The code is available on GitHub at <https://github.com/simpson-lab/wpg-mb-lakes>.

Pigment concentrations were modelled with a hierarchical generalized additive model (HGAMs) using the mgcv package (Wood, 2011, 2017; Simpson, 2018; Pedersen *et al.*, 2019). The HGAM used a Gamma location-scale family which allowed to estimate the trends in the concentrations’ mean and variance (Wood *et al.*, 2016). The mean concentration was modelled using factor smooths for each combination of the 2 lakes and 5 pigments, for a total of 10 factors (model *S* sensu Pedersen *et al.* (2019)). The shape term of the distribution was modelled with a similar factor smooth and a smooth of each slice’s log-transformed interval, which accounted for the differences in information between slices, since slices which averaged more years were expected to be less variable (i.e. more stable). Observations were weighted by the (standardized) interval of the sample. The smoothness parameter of the model was estimated using restricted maximum likelihood (*REML*, see Simpson (2018)). The estimated variance was then obtained by calculating the product of the mean and shape estimates, and 95% Bayesian credible intervals for the variance were obtained by running 10,000 simulations and taking the 0.025% and 97.5% quantiles of the posterior distribution.

The pigment concentrations in the four cores from Lake Manitoba were estimated in a similar way, but the factor smooths for each combination of the 4 cores and 5 pigments, for a total of 20 factors.

The phosphorous concentrations in the first two cores of Lake Manitoba were also modelled using a GAM. However, since the trends were fairly simple and no complex inference was required, the models were created using the geom\_smooth() function from the ggplot2 package with arguments formula = y ~ s(x) and method = 'gam', rather than fitting models with the mgcv package.

To identify periods of significant change, the first derivative of the estimated smooth trend was evaluated for each proxy from the relevant model (Bennion, Simpson & Goldsmith, 2015; Simpson, 2018). The first derivative was estimated using functions built with the gratia package (Simpson, 2021). Periods of significant change were identified as periods where the credible interval of the estimated derivative excluded 0 at the level of significance.

***ask Gavin to add some text regarding derivatives***

# References

Bennion H., Simpson G.L. & Goldsmith B.J. (2015). Assessing degradation and recovery pathways in lakes impacted by eutrophication using the sediment record. *Frontiers in Ecology and Evolution* **3**. <https://doi.org/10.3389/fevo.2015.00094>

Pedersen E.J., Miller D.L., Simpson G.L. & Ross N. (2019). Hierarchical generalized additive models in ecology: An introduction with mgcv. *PeerJ* **7**, e6876. <https://doi.org/10.7717/peerj.6876>

R Core Team (2021). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Simpson G.L. (2021). *Gratia: Graceful ’ggplot’-Based Graphics and Other Functions for GAMs Fitted Using ’mgcv’*.

Simpson G.L. (2018). Modelling Palaeoecological Time Series Using Generalised Additive Models. *Frontiers in Ecology and Evolution* **6**, 149. <https://doi.org/10.3389/fevo.2018.00149>

Wickham H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York.

Wickham H. (2021). *Tidyr: Tidy Messy Data*.

Wickham H., François R., Henry L. & Müller K. (2021). *Dplyr: A Grammar of Data Manipulation*.

Wilke C.O. (2020). *Cowplot: Streamlined Plot Theme and Plot Annotations for ’ggplot2’*.

Wood S.N. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)* **73**, 3–36

Wood S.N. (2017). *Generalized Additive Models: An Introduction with R*, 2nd edn. Chapman; Hall/CRC.

Wood S.N., N., Pya & S"afken B. (2016). Smoothing parameter and model selection for general smooth models (with discussion). *Journal of the American Statistical Association* **111**, 1548–1575