Interactive comment on "Natural stochasticity vs. management effort: use of year-to-year variance for disentangling significance of two mutually confounding factors affecting water quality of a Norwegian cold dimictic lake" *by* A. T. Romarheim et al.

Anonymous Referee #3

Received and published: 16 December 2014

The paper describes an interesting simulation experiment for disentangling natural fluctuations of water quality variables from management-induced effects.

In a first step, a dynamic model (MyLake) is fitted to observed data using a Bayesian MCMC procedure. In the simulation experiments, selected groups of forcing variables are replaced by a "pseudo-repeated average year", i.e. repeated series of an annually averaged data. The model outcome is analyzed with parametric, p-value-based (fixed-

effects) ANOVA and with multivariate analysis (PCA).

The paper is very well written, the methods are well explained and the outcome is plausible. However, there is still room for improvement.

=== Major points ===

1) Style: The figures are difficult to read, both on printed paper and on electronic devices, especially the combination of tiny axis annotations and the large number of sub-figures (esp. 4x7 for Fig. 2), but also the unnecessary small axis annotations of Fig. 3 and 4.

2) I wonder, why a p-value based ANOVA was used and not a variance components analysis (cf. Crawley 2012, p 475). In my opinion, it would be preferred to measure the variance contribution of the different factors instead of (or at least in addition to) significance testing. Additional arguments can be found for example in Nakagawa & Cuthill (2007) and many other recent papers.

3) The approach of an average year as a reference sounds plausible at a first look, but averaging several years does not produce an "average year", because intra-annual stochasticity is lost. This well-known effect can lead to systematic bias, e.g. less turbulence due to the cut-off of extreme wind events. Therefore, hydrodynamic models are usually driven by stochastically generated time series (cf. Semenov et al. 1998 or Schlabing et al. 2012). How does this influence the results (-> Discussion)?

4) Replacing single (or groups of) forcing variables by others may lead to unrealistic cases, e.g. blue sky with strong rain. I admit that the proper treatment of such dependencies and cross-correlations is difficult, but it should at least be considered in the Discussion.

=== Minor ===

p 12492 L 09: 400 mg mˆ-3 is indeed very high (comment only)

L 17: please mention name of model here (first occurrence)

L 22: the average of years is not an "average year"; averaging reduces intra-annual variability and autocorrelation structure of the time series

p 12493 dito

p 12495 L 15 Ruttner (no umlaut)

L 19... MCMC - I like this approach very much - which algorithm / software / package was used (citation?) - which priors have been used (non-informative?)

p 12496 L 25: meteorological

p 12497 L 11: why statistical tests and not variance decomposition?

L 25: "shallower", "deeper": not completely clear if this concerns vertical layers or the horizontal zones (littoral, pelagic zone) of the lake

p 12498 L 16: _as expected_, "global radiation varied most during summer months" - because this is completely obvious.

p 12499 L 8: "some extent" - How much is "some"? Measurement instead of p-values!

p 12501 L 30: yes, dependency/covariance is not considered, see major point (4)

In general: methodological deficits should be discussed more in detail, especially how they influence the conclusions. Simplifications are unavoidable, of course, but it should be discussed if the conclusions are still on the safe side, are "best case" resp. "worst case" approximations.

Fig. 2: much too small

Fig. 3, 4: axis annotations too small

Fig. 5: please add %of total variance to the PC axis

=== Additional References ===

Crawley, M. J. (2012) The R Book. Wiley & Sons, p 475.

Nakagawa, S. & Cuthill, I. C. (2007) Effect size, confidence interval and statistical significance: a practical guide for biologists Biological Reviews, Wiley Online Library, 2007, 82, 591-605, doi:10.1111/j.1469-185X.2007.00027.x

Nuzzo, R. (2014) Statistical errors. Nature 506, 150-152.

Schlabing, D., Frassl, M.A., Eder, M.M., Rinke, K., Bárdossy, A. (2014) Use of a weather generator for simulating climate change effects on ecosystems: A case study on Lake Constance. Environmental Modelling & Software 61, 326-338. doi:10.1016/j.envsoft.2014.06.028

Semenov, M.; Brooks, R.; Barrow, E. & Richardson, C. (1998) Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates Climate Research 10, 95-107