Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-272-AC1, 2019

Geoscientific
Model Development
Discussions

Interactive comment on "Bayesian Inference and Predictive Performance of Soil Respiration Models in the Presence of Model Discrepancy" *by* Ahmed S. Elshall et al.

Ahmed S. Elshall et al.

mye@fsu.edu

Received and published: 14 March 2019

Anonymous Referee #1

Comment: The paper evaluates the impacts of statistical data assumptions in soil microbial respiration modeling on estimated model parameters and on model predictions. Inference is done using various soil respiration models and various likelihood functions, using half hourly CO2 flux data from a field site. It's an interesting study, but I suggest additional effort to clarify and increase contribution of the work.

Response: We are very thankful for the reviewer for talking the time to evaluate the

manuscript, and for providing constructive comments.

Comment: 1. Contribution: the authors should more clearly spell out the explicit contributions of the paper. On the one hand, the methodology is not new and has been developed and applied in hydrological studies. On the other hand, the application to CO2 modeling may also not be entirely new since the likelihood approach used here has already been applied to ecological modeling (including carbon flux modeling); a recent example is Scholz, K., Hammerle, A., Hiltbrunner, E. et al. Ecosystems (2018) 21: 982. https://doi.org/10.1007/s10021-017-0201-5.

Response: We explicitly spelled out the novel contrition of this paper, which is the systematic evaluation of the impact of data model selection on Bayesian inference and predictive performance of soil respiration modeling with different degrees of model fidelity. We did a systematic review of Bayesian inference for soil respiration modeling. Most studies assume independent, Gaussian, and homoscedastic residuals. Few studies have relaxed these assumptions(e.g. Elshall et al., 2018; Scholz et al. 2018). However, only very few studies have focused on investigating the impacts of these assumptions for soil respiration modeling by relaxing the independent residuals assumption (Ricciuto et al., 2011) and the Gaussian residuals assumption (Ricciuto et al., 2011; van Wijk et al., 2008). By relaxing these three assumptions step-wise resulting in eight data models, to our knowledge this is the first study that systematically evaluates the impact of data models on Bayesian inference and predictive performance of soil respiration modeling.

The revised manuscript reads: "Bayesian inference of soil respiration models often adopts the assumption of independent, normally distributed and homoscedastic residuals (e.g. Ahrens et al., 2014; Bagnara et al., 2015, 2018; Barr et al., 2013; Barron-gafford et al., 2014; Braakhekke et al., 2014; Braswell et al., 2015; Correia et al., 2012; Du et al., 2015, 2017; Hararuk et al., 2014; Hashimoto et al., 2011; He et al., 2018; Klemedtsson et al., 2008; Menichetti et al., 2016; Raich et al., 2002; Ren et al., 2013; Richardson and Hollinger, 2005; Steinacher and Joos, 2016; Tucker et al., 2014; Tuomi et al., 2008; Xu et al., 2006; Yeluripati et al., 2009; Yuan et al., 2012, 2016; Zhang et al., 2014; Zhou et al., 2010). These assumptions are conveniently adopted since the requirement of using an unknown probability model in Bayesian statistics is called "a basic dilemma" by Box and Tiao (1992). Postulating the data models is always based on assumptions about residual statistics, and the most widely used assumptions are paired as follows: (i) independent vs. correlated residuals, (ii) homoscedastic vs. heteroscedastic residuals, and (iii) Gaussian vs. non-Gaussian residuals. For soil respiration modeling few studies have relaxed the independent residuals assumption (e.g. Cable et al., 2008, 2011; Li et al., 2016b), the homoscedasticity assumption (e.g. Berryman et al., 2018; Elshall et al., 2018; Ogle et al., 2016; Tucker et al., 2013), and the non-Gaussian and homoscedasticity assumptions (e.g. Elshall et al., 2018; Ishikura et al., 2017; Kim et al., 2014). A recent study (Scholz et al., 2018) relaxed these three assumptions using the generalized likelihood function (Schoups and Vrugt, 2010). However, few studies have focused on investigating appropriateness and impact of these assumptions for soil respiration modeling. This was performed by relaxing the independent residuals assumption (Ricciuto et al., 2011) and the Gaussian residuals assumption (Ricciuto et al., 2011; van Wijk et al., 2008). By relaxing these three assumptions stepwise resulting in eight data models, to our knowledge this is the first study that systematically evaluates the impact of data model selection on Bayesian inference and predictive performance of soil respiration modeling. In addition, to our knowledge this is the first soil respiration modeling study that investigates the impact of data models in relation to model fidelity." In the first paragraph of the introduction we also stated "While a large number of data models have been used (e.g. Elshall et al., 2018; Scholz et al., 2018) to our knowledge comprehensive and systematic evaluation of data models for soil respiration modeling has not been reported in literature."

Comment: 2. The authors find some problems with the estimation of autocorrelation and suggest an alternative approach (Evin et al.). Why not test this approach as well? I'm not sure this would warrant a separate publication. Including it here would enhance novelty of the paper in my opinion. Note also that the high temporal resolu-

tion (half hourly) of the data used by the authors may be a complicating factor; see the following paper that discusses this: https://www.hydrol-earth-syst-sci-discuss.net/hess-2018-406/.

Response: Thank you very much for bring our attention to this recent article of Ammann et al. (2018).

This manuscript provides a systematic evaluation of the impact of data model selection on Bayesian inference and predictive performance of soil respiration modeling. Figure 10 for example shows specific trends that would occur when relaxing the three assumptions of non-correlation, normality, and homoscedasticity using joint inversion approach, which has never been reported before in literature.

Autocorrelation is a complicated problem that we are currently working on. Joint inversion of heteroscedasticity and autocorrelation parameters can lead to poor predictive performance (Evin et al., 2013, 2014; Ammann et al. 2018; and this study). To address this problem a two-step procedure (e.g. Lu et al., 2013; Evin et al., 2013, 2014) was proposed. Our preliminary results show that using the sequential approach of Evin et al. (2013; 2014) by estimating the autoregressive parameters sequentially (after estimating the soil respiration model parameters and data-model parameters) did not solve this problem. Ammann et al. (2018) even states that the joint inversion is still preferred, and understanding the conditions where accounting for auto-correlation can be achieved remain poorly understood. The problem of autocorrelation has several interlinked aspects that we would like to address in another manuscript. Auto-correlated errors might be attributed to a systematic error in the soil respiration model. The most obvious solution is to improve the soil respiration model. Otherwise, we can improve our data model. Our hypothesis that we would like to test is that omitting autocorrelation error through a filter approach (e.g. Schoups and Vrugt, 2010; Evin et al., 2013; 2014; this study) could be tricky as this leads to a loss of information content. Thus, joint approach may lead to biased parameter estimation (Figure 5) and poor predictive performance (Figure 10). While sequential approach would avoid the biased parameter estimation, but would still lead a poor predicative performance.

Our current understanding is that this problem could emerge from several interlinked factors: âĂć Non-stationarity due to wet-dry periods as proposed by Ammann et al. (2018) could be a reason for this problem and thus accounting for non-stationarity (Smith et al., 2010b, Ammann et al. 2018) could alleviate this problem. âĂć The method for accounting for autocorrelation could have an impact. Autocorrelation could be addressed using a likelihood function based on covariance matrix of residuals L(e) (e.g. Lu et al., 2013) with transformed residuals, and likelihood function of normalized residuals L(a) (e.g. Schoups and Vrugt, 2010; Evin et al., 2013; 2014; this study) with autoregressive model that filter out autocorrelation. Note that "e" is a vector of transformed residuals, while "a" is a vector independent and identically distributed random errors with zero mean and unit standard deviation. The impact of method selection is still unclear and needs investigation. âĂć Joint versus sequential inversion for autocorrelation could also have an impact. Ammann et al. (2018) suggests that the joint inversion is still preferred over sequential inversion. This will be investigated under both L(e) and L(a) approaches. In addition, we would like to test novel joint inversion procedure that combines the L(a) and L(e) approaches as follows. First, the parameters of the linear heteroscedastic model will be estimated similar to Schoups and Vrugt (2010) to remove heteroscedasticity. For each MCMC sample, after applying the linear heteroscedasticity model, the auto-correlation parameters can be deterministically calculated as internal variables of the data model similar to Lu et al. (2013) and not as calibration parameters(e.g. Schoups and Vrugt, 2010; Evin et al., 2013;2014). This is mainly to avoid interaction between heteroscedasticity and autocorrelation parameters. The auto-correlation parameters can be calculated following Lu et al. (2013). All these interlinked factors require careful consideration, and this is warranted in another manuscript.

We have revised the manuscript to further clarify these issues as follows: “This study confirms the empirical findings and theoretical analysis (Evin et al., 2013; 2014; Am-

mann et al. 2018) that separate accounting for autocorrelation or joint inversion of correlation and heteroscedasticity can be problematic. By drawing on similarity from surface hydrology, the study of Ammann et al. (2018) suggests that this might be attributed to non-stationarity due to wet-dry periods with half-hourly data. Accounting for non-stationarity (Smith et al., 2010b, Ammann et al. 2018) could address this problem. Relatively poor performance with respect to autocorrelation can be also attributed to the implementation scheme. The inference scheme such as joint inference as in this study, post-processing inference approach for autocorrelation (Evin et al., 2013; 2014), residuals transformation approach (e.g. Lu et al., 2013) or other strategies (Li et al., 2015, 2016a) could have an impact. Yet Ammann et al., (2018) study states that the joint inversion is still preferred, and understanding the conditions where accounting for auto-correlation can be achieved remain poorly understood. Further investigation of this point is warranted in a future study.”

Comment: 3. The paper should be checked for various grammatical errors and typos. One example is "heteroscedasticity", which is spelled in multiple creative ways throughout the paper.

Response: Thank you very for pointing this out and we have corrected "heteroscedasticity" at eight different locations throughout the manuscript. We corrected several other grammatical errors and typos.

Comment: 4. Description of the various evaluation metrics seems better placed in the methods than results section.

Response: We moved the description of the various evaluation metrics from the results to the methods section.

Comment: 5. Terminology: the distinction between model fidelity and discrepancy is not clear

Response: We clarified these two terms as follows: “We use the terms model fidelity and model discrepancy interchangeably. Model fidelity refers to the degree of realism of representing our scientific knowledge with respect to the real world system. That is a high fidelity model has less discrepancy."

Comment: 6. Line 305, "discrete proposal distribution": I don't think the proposal is discrete, it is a proposal distribution over a continuous parameter space.

Response: We revised "discrete proposal distribution" to "adaptive proposal distribution."

Comment: 7. Line 477: please rephrase; I don't think it's "expected" that accounting for autocorrelation leads to biased parameter values. I would expect the opposite, since autocorrelation provides a (simple) way to account for model errors.

Response: We rephrased this sentence to "First, we obtained biased parameter estimates that is out the reasonable physical range."

Comment: 8. Eq. 23: is index i an index over time or is it an ensemble index? Please clarify.

Response: Thank you very much for point this out. We clarified that this is an ensemble prediction Yij where i is index over time, and revised other parts of the manuscript accordingly. The new sentence read "the ensemble prediction Yij is similar to Yi above where is index over time and specific to the j-th combination."

Comment: 9. Line 598: approaches that use "total residual error" typically still separate out parametric uncertainty, so the residual error includes measurement, model input, and model structure uncertainty, but not parameter uncertainty.

Response: That is true. We rephrased that sentence to "total residuals that separates out parametric uncertainty, so the residual error includes measurement, model input, and model structure uncertainty."

Thank you very much for your constructive comments.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-272, 2018.