Response to the comments of the referees

Dear Dr. Arndat and referees,

We would like to thank the two referees for their careful and constructive comments for improving our manuscript. In the following, we address all comments and questions from the referees. The comments by the referees are given in blue color, our responses are in black. We have considered all comments for the preparation of a revised version of the manuscript.

Response to comments from Dr. C.A. Sierra

General comments

This manuscript presents a framework to implement new biogeochemical processes within the Community Land Model (CLM). The approach is interesting and can potentially help developers of other large models to include new processes. This type of manuscripts are of great value and deserve publication in Geoscientific Model Development. However, I have concerns about the reproducibility of the framework presented here. The manuscript lacks detailed descriptions of the mathematical framework and its implementation. In my opinion, this manuscript may only be of value for people already familiar with CLM, unless the authors make an extra effort and present the framework in more detail and perhaps with more explicit mathematics.

We thank Dr. Sierra for the positive comments.

The framework we presented includes two parts: 1) a reaction database for the users to flexibly extract processes of interest in a particular application to form a reaction network, and 2) using the reaction network approach to derive ODEs to be solved. The reaction database is captured in an ASCII file. It can be created by the user from scratch using known reaction pathways, or expand from the database in the supplementary material. In this model, the ODEs were solved using an explicit method.

The main contribution of this approach is to derive a system of ODEs automatically using matrix decomposition. The model reads in the stoichiometries of the reaction network and stores them in Matrix [A] of Eq (2). The creation of the stoichiometries is automated by running a perl script to parse the reaction network. If a system is composed of only slow reactions, the ODEs can be derived directly by expanding the matrix form of Eq (2). When there are fast reactions, a substance in a fast reaction is selected as a pivot element to perform column reduction by using a matrix row operation to convert the column in the reaction matrix that contains the pivot element to a unit column. We used a simple system with one slow reaction and one fast (equilibrium) reaction, and three state variables to illustrate the matrix decomposition in the revised manuscript on pages 9 to 11. The method presented in the manuscript is not only applicable to CLM, but also to other land surface models. It is a standalone module and can be implemented in other codes. Any terrestrial ecosystem modeler can make use of our approach in their models.

We have added a diagram to illustrate the framework of the model and added the details in the revised manuscript (page 12).

In particular, I found hard to understand the reaction pathways depicted in Tables 2 and 3. Perhaps it would help the reader if the reactions in this table are also presented in terms of systems of ODEs that more explicitly represent the process. The terms R in these formulas are ambiguous and it is not completely clear to me if they are calculated as a first-order linear reaction (e.g., R = kC), or whether the term R generically represents a flux calculated either as linear or non-linear interaction.

We apologize for the confusion. We define the flux of a substance as the derivative of that substance with respect to time, and has the dimensions of amount of substance per unit volume transformed per unit time. Reaction rate for a single reaction differs from the rate of increase of mass of a substance by the reciprocal of its stoichiometric number in that reaction. Using the following reaction as an example:

$aA \leftrightarrow bB$
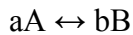
where a and b are the stoichiometric number of A and B, respectively. The rate of this reaction is $R = -\frac{1}{a}\frac{dA}{dt} = \frac{1}{b}\frac{dB}{dt}$, while the flux of A is $\frac{dA}{dt} = -aR$ and the flux of B is $\frac{dB}{dt} = bR$.

The rate of a reaction is always positive, and it can be calculated either as a linear or non-linear function. We made the clarifications in the revised manuscript on page 9.

We have added the systems of ODEs derived from the reactions in Table 2 and 3, and the reaction kinetics (first order kinetics in this model) in the supplementary material.

In this regard, the use of the word 'reaction' in the manuscript may be confusing. In geosciences it is common to discriminate between fluxes (units of mass per time) and rates (units of per time). The authors use the word reaction rate for R, but my impression is that they are referring to a flux. I know chemists prefer the term reaction for what could be considered a flux, so I recommend the authors to define better their terminology and the exact meaning of R.

We clarified the terminology in the revised manuscript on page 9.

Technical/minor comments

• Introduction. Although the discussions here on sulphur and microbial dynamics are interesting, it seems disconnected from the rest of the paper. There are no references at the end of the manuscript on how sulphur or microbial dynamics can be implemented in the framework. I suggest the authors to either make the connection more explicitly in the discussion or remove this part from the introduction.

The reason to discuss sulfur and microbial dynamics is to point out how complex a system can be and the efforts required to incorporate them in the system. We added a conceptual diagram of the microbial model and used it as an example showing how the model can be implemented in the database in the revised manuscript on page 15. We also added the following references for the sulfur and microbial model:

Mitchell, M.J. and Fuller, R.D., 1988. Models of Sulfur Dynamics in Forest and Grassland Ecosystems with Emphasis on Soil Processes. Biogeochemistry, 5(1): 133-163.
Wieder, W.R., Bonan, G.B. and Allison, S.D., 2013. Global soil carbon projections are improved by modelling microbial processes. doi:10.1038/NCLIMATE1951. Nature Climate Change

• Equation 2. If [I] is the identity matrix, why do you need it there to multiply the vector of differential equations? If you multiply by the identity matrix you get the same vector back. Am I missing something? Do you mean something different when you say 'unit' matrix? Please explain this better.

The identity matrix is useful in performing matrix decomposition on the equation, especially when there are fast reactions. This is illustrated by the simple system presented in the revised manuscript on pages 9-11.

• Page 3217, line 11. Can you give more detail about how the ODEs are implemented and solved in CLM? Is there a standard library called to solve the ODEs or are they solved explicitly within the model?

The ODEs in CLM are solved using explicit time integration, no standard library is needed. This is clarified in the CLM introduction section in the revised manuscript on page 6.

• Page 3219, line 21. Can you point out where the user manual can be found? Can you add it as supplementary material for this manuscript? Please see GMD submission guidelines and be aware that manuscripts in this journal usually include user manual and access to source code.

We removed the sentence and added the format in the supplementary material.

• Page 3020, line 2. What are p and c?

p represents plant functional type, c represents column. We reworded the sentence to better define the two symbols as "For each process in the database, state variables are separated into plant functional type (p) or column averaged type (c)".

• Page 3223, line 4. What do you mean by 'recent technology'? I guess you probably want to say recent 'advances' or 'contributions'. By the way, are you familiar with other P cycling models? You may want to look at Goll et al. (2012, Biogeosciences 9: 3547), and Buendia et al. (2010, Biogeosciences 7: 2025).

Modified to "recent contributions" as suggested.

Thanks for pointing out other P cycling models. We have added them in the references.
It would be useful for future modeling comparison studies.

• Supplementary material. Please add a commented heading explaining what this
code does. If possible, also add inline comments within the code.

The database is a text file, listing the reaction pathways. It can be expanded when more
processes become available. We added the comments and format in the database.