European Commission

September 1st, 2022

To Geoscientific Model Development (GMD)

Dear Editor and Reviewers,

It is a great pleasure to submit a revised version of our manuscript entitled "Matrix representation of lateral soil movements: scaling and calibrating CE-DYNAM (v2) at a continental level". The first version was asked for modifications on July 26th, 2022.

We want to thank the Editor and the Reviewers for the comments and the time spent reading and evaluating our work. Please find in this response letter the following content:



On behalf of the other authors, I thank you again for your time and consideration. Please feel free to contact me during the revision process if necessary.

Yours sincerely,


*Arthur Nicolaus Fendrich*

*Université Paris-Saclay*
*Joint Research Centre*

# 1. Comments from Referees

## 1.1. Yuanyuan Huang

"The authors developed the process-based model that coupling carbon, erosion, transport and deposition processes. They brought in the lateral movement of carbon into process-based land carbon modelling. The community has been expecting this advancement, especially in large scale studies, for long, but with limited progresses partly due to the high computational cost. The authors presented a detailed and exciting case. The well-illustrated their model formulations, how they tackled on the computational bottleneck, the calibration and validation, with adequate discussion of limitations and future improvements. The study is well-designed and informative. The writing is generally clear despite some part might be a little lengthy. My criticises are between major and minor. Please check below.

It is not clear to me how the matrix-relevant techniques helped the current study. The equations for the lateral carbon fluxes are presented mostly in carbon balance equations (i.e., no need for the matrix form). Are the matrix techniques only used for constructing the ORCHIDEE emulator? If the ORCHIDEE output only an input to CE-DYNAM, or any parameter changes that require a re-do of model spin-up that requires computation resources?

Parameter values for soil discretization are optimized. How about parameter values for other parts of the model? I might miss some part, but is there any table or supplementary information that documents values of relevant parameters used in this study?

**Minor comments**:

Line 1-11. Is the background for abstract a little lengthy?

Line 24. I would suggest DayCent. For people without background, they don't know what DayCent refers to. DayCent is not mentioned in texts other than here. Besides, the CENTURY vs. DayCent is another layer that needs background. So DayCent here is not necessary to add more information to this already complex manuscript.

Line 88 add in before Table 1

Line 92. Is it better to write ORCHIDEE CENTURY as ORCHIDEE (CENTURY-Carbon type) or other better rewordings. ORCHIDEE CENTURY is confusing.

Line 93. Is the first order kinetics necessary? Current application is with linear-models due to computational cost. In theory, for example, if our studying region is small, the coupling with an nonlinear carbon model is possible, right?

Line 125. Could we write it as " induced by the terrain slope (S[x,y]) and the flow accumulation (w[(x,y)])" to reduce confusing?

Table 1. Could you use other symbols for ω[(x,y)] vs w[(x,y)] (?)? They looks the same

Equation 6. Lines 200-205, and across the manuscript. By erosion b – t, you mean from lower soil layer to the upper soil layer and the flux is there between adjacent soil layers? Please clarify, by b – t, it refers specifically from the third to the first layer, in your context. So it is not clear to me why "losses from the layers below must be added to the layers above"

Line 525. Is the breakdown of aggregates a transport process? Please clarify. Aggregates breakdown could happen without the transport process.

Line 525-530. If the "halfway between…."a cause of the difficulty in finding the optimal resolution, or the computation cost and applications?"

### 1.2. Holger Metzler

"The authors provide a spatially upscaled version of an eriosion, transport and deposition model (CE-DYNAM) to European scale, while sticking to a high spatial resolution. This is an important scientific contribution because lateral C transport has largely been ignored so far in process-based models because usually models consider different grid cells as independent. Such lateral transports can have a significant effect and should be investigated because nonlinear effects could potentially lead to drastically new insights and better understanding compared to existing models.

The spatial upscaling was computationally feasible only because the authors emulated the original model by a matrix approach. This allows the application of sparse matrix models as well as an improved application of parallel computation methods. Another advantage of matrix models is that they allow a rigorous mathematical analysis, something the authors did not do in this manuscript (it was not their goal) but can be done in the future based on the matrix reformulation. This was not possible with the original implementation. Furthermore, I appreciate all the effort the authors showed in reimplementing an existing model with a matrix approach, emulating the original simulation results very vell. In this regard I like also the explanation of the matrix shape in Section 3.3. Nevertheless, I wished I could have "seen" a matrix, at least as a schematic block matrix after the authors speaking so much about matrices.

The presentation and the writing is clear, explaining the model calibration and simulation results as well as model limitiations and future opportunities very well. Sometimes though the text appears too lengthy in my opinion. This starts with the abstract and continues with quite some overlap in the sections about calibration, results, simulations and limitations. Furthermore, in particular in later sections I was overwhelmed by an extensive use of potentially unnecessary numbers.

I furthermore do have some issues about an easy reproducibility of the method, because to me it seems that in Section 2 some formulas are incorrect and notation is not precise. Well, either the formulas are incorrect or I understood them wrong, neither option is preferrable. In particular I consider the statement (p. 3, l. 71): "We expect our mathematical development of CE-DYNAM to facilitate its reproduction and incorporation in LSMs such as ORCHIDEE, DayCent, and others." to be pretty bold. Under this point of view formulas and notation should show no flaws.

This starts with Table 1, which is in general very nice, but it is incomplete, some symbols that are used later on do not appear here. This was sometimes annoying for me while reading. For example I* was never properly introduced. On p.5, l. 132 it is stated that inputs are disturbed along D, which I do not understand.

**Here my issues with notation and formulas, which should be thoroughly checked:**

p. 8. l. 1: Should it be $\Delta_i$? Is it used to compute $d_0=0$, $d_1=d_0+\Delta_0$ and so forth?

p. 8. l. 183: $k_e$ is missing in Table 1. The "Total mass of soil" under the bracket refers to total mass of soild in hill slopes? Why is it decreasing with h? To me this looks like we have an infinite erosion if we do not have any hill slopes, isn't this going in the wrong direction?

p. 10, l. 202. If there is going to be a new input flux, should there also be a new output flux? The input must come from somewhere. Please also indicate where the new input flux goes, I do not want to guess here.

p. 10, l. 213. The way $k_T$ (missing in Table 1) is defined two lines later, it seems wrong to me to call it a flux. It is a rate (dimension 1/time). It will only become a flux once it is multiplied with a pool content. This leads to major confusions for me later on.

p. 11, l. 235: Please use other notation than (a, b) for the indexing of the sum, the two letters are already taken. It becomes very confusing this way. Again, I think that if $k_T$ is a rate rather than a flux, then $k_s$ will be as well because all the other factors in Eq. 11 are dimensionless.

p. 10, l. 241 "all PFTs" should rather be "all PFTs but $p_0$", right?

p.12, l. 247: "Such input flux". I disagree again, same problem. It is not yet a flux because it is not yet multiplied with a carbon stock, which should be $S_{[(a,\, b),\, \ldots]}$ here?

p. 12, l. 257: The "P" here looks different from the ones introduced on p. 10, l. 221. Furthermore, I am not sure whether it belongs here in the first place, well the absolute value is unecessary in any case. But why multiply by the number of non-zero PFTs in cell (a, b)? Shouldn't this be already included in $k_T$ already? Could you write it down explictly for yourself without the P but a second sum instead and check whether it is correcct this way?

p. 12, l. 260: In $k_s$ I think that source and target are confused.

p.13, l. 265: What is $k_t^*$?

p. 13. l. 266. Should $k_s$ with source (a, b) be multiplied with some stock indexed by (a, b) instead (x, y).

I obviously do have some confusions about the firs-order description, where sources and targets seem not to match, at least in my head. So would like to encourgae the authors to carefully check the notatian and the formulas again, along with their implementation.

**Small issues:**

- In general units are sometimes italic and sometimes not, sometimes with a space between the number and the unit, sometimes without.
- The use of singular and plurar gets mixed up quite often.
- p. 3, l. 86: "approach" --> "an approach"?
- p. 3, l. 88: Table 1 --> (Table 1)?
- Table 1:
- - Description sometimes ends with a period, sometimes not.
- - The adimensional respiration rates of carbon $k_r$ actually do have a dimension: 1/day.
- - $I^{[c\_j,\ c\_j]}$: One of the j's should be an i. Probably the first one, then please also adapt the description, to make it consistent with $k_t$.
- - Then it is a little unfortunate to use ω and w for the depth to bedrock and the flow accumulation, respectively.
- p. 7, l. 145: I am not sure if such a procedure necessarily converges to a pullback attractor (given there is one in the first place). The pullback attractor is reached when starting the simulation earlier and earlier, basically moving toward an infinite simulation history.
- p.7, l. 152: Why now change the notation from S, I, and k to x, B, and A?
- p. 10, l. 243: "soi"
- p.14, l. 280: The seventh-largest what?
- p. 15, l. 308: $u_i$ instead of $p_i$?
- p. 15, l. 312: "P is Panagos", what does this mean?
- p.15. unnamed formula: what is r(y, m)?
- p. 18, l. 398, 399: What are C-factor and "R factor"? Please also note the different way of writing them."

## 2. Authors' response

### 2.1. To Yuanyuan Huang

"Dear Dr. Huang,

First, thank you for the thoughtful review, which will certainly improve our manuscript. We answer the major questions below and will address the minor comments in the next round of manuscript revision.

**Question**: "It is not clear to me how the matrix-relevant techniques helped the current study. The equations for the lateral carbon fluxes are presented mostly in carbon balance equations (i.e., no need for the matrix form). Are the matrix techniques only used for constructing the ORCHIDEE emulator? If the ORCHIDEE output only an input to CE-DYNAM, or any parameter changes that require a re-do of model spin-up that requires computation resources?"

**Answer**: The matrix techniques are used only for constructing the emulator, as the Reviewer asked. The ORCHIDEE output is used as an input to CE-DYNAM, in such a way that if we run CE-DYNAM without enabling erosion, transport, and deposition (ETD) modules, we recover the original ORCHIDEE results. However, when we enable these modules, the results change, and spin-up calculations must be re-done because a new equilibrium state is obtained when new processes are included. In CE-DYNAM, this is done with the emulator by using the rates extracted from the original ORCHIDEE plus the new rates of ETD dynamics presented in the manuscript. Regarding how the matrix technique helped the study, we agree that all equations refer to carbon balance. However, the dependence between adjacent cells significantly impacts the time demanded to calculate the new equilibrium state and the model dynamics. As we mentioned in the paragraph L.480-501, the matrix approach allowed us to overcome the existing barriers to the implementation of CE-DYNAM at a continental scale, as we could precalculate the matrices for every simulation month (L.478-479) and increase the number of parallel threads compared to the previous implementation (L. 485). In the next revision, we will improve this part of the text to clarify for the readers.

**Question**: "Parameter values for soil discretization are optimized. How about parameter values for other parts of the model? I might miss some part, but is there any table or supplementary information that documents values of relevant parameters used in this study?"

**Answer**: In this study, we intend to isolate the effects of ETD on the carbon cycle. Therefore, the parameters presented refer only to those introduced by the current formulation. We opted not to modify any of ORCHIDEE's default values, which can be found in other publications such as Krinner et al. (2005) [10.1029/2003GB002199].”

### 2.2. To Holger Metzler

“Dear Dr. Metzler,
We would like to thank you for the very relevant comments about our work.

We answer below the comments, and we will fix all the issues in the next round of manuscript revision.

**Reviewer**: "In this regard I like also the explanation of the matrix shape in Section 3.3. Nevertheless, I wished I could have "seen" a matrix, at least as a schematic block matrix after the authors speaking so much about matrices."

**Answer**: We agree that it would be interesting to visualize the matrix structure. We will add this image and information in the next manuscript version.

**Reviewer**: "Sometimes though the text appears too lengthy in my opinion. This starts with the abstract and continues with quite some overlap in the sections about calibration, results, simulations and limitations. Furthermore, in particular in later sections I was overwhelmed by an extensive use of potentially unnecessary numbers."

**Answer**: It was a common point in both Reviewer's comments that some text sections are lengthy. In order to fix this problem and improve readability, we will rewrite several passages in the next version of the manuscript, with extra attention to the aforementioned unnecessary numbers.

**Reviewer**: "I furthermore do have some issues about an easy reproducibility of the method, because to me it seems that in Section 2 some formulas are incorrect and notation is not precise. Well, either the formulas are incorrect or I understood them wrong, neither option is preferrable (...) Under this point of view formulas and notation should show no flaws. (...) This starts with Table 1, which is in general very nice, but it is incomplete (...). So would like to encourgae the authors to carefully check the notatian and the formulas again, along with their implementation."

**Answer**: We thank very much the Reviewer for pointing out all the problems with the notation adopted. We strongly agree that this is critical for adopting and disseminating the matrix approach and for future developments of the method. In order to fix all issues, all the formulas of the manuscript will be carefully reviewed and rewritten to prioritize precision, conciseness, and clarity. We hope the new notation will be precise and resolve any confusion that may have arisen.

**Reviewer**: "This allows the application of sparse matrix models as well as an improved application of parallel computation methods. Another advantage of matrix models is that they allow a rigorous mathematical analysis  something the authors did not do in this manuscript (it was not their goal) but can be done in the future based on the matrix reformulation."

**Answer**: We truly appreciate this suggestion and agree that it would be an interesting work to develop in the future.“

## 3. Authors' changes in the manuscript

Dear Editors and Reviewers,

A common point among the Reviewers' comments was that the manuscript seemed lengthy, which we understand was largely caused by the mathematical notation adopted. As seen in both Reviewers’ comments, such a problem created an unnecessary confusion in the understanding of the manuscript, which should be resolved. For this reason, we have revised and reworked all the mathematical notation in the article in order to make it more succinct and easily understandable. We detail the main changes below.

- Subsection 2.1.1: In this section, there was ambiguous terminology in the previous version of the text. For example, the index $m$ in Equation 2, which is fundamental to the presentation of the matrix method, was not properly defined, making it impossible for the reader to understand the dimension of the matrices. In addition, Table 1 presented contradictory information, as mentioned by the Reviewers. All these problems were duly solved and changed into a new, clearer and more precise notation.
- Subsection 2.1.2: Similar to the previous subsection, subsection 2.1.2 had elements that could confuse the readers. As mentioned by both Reviewers, the variable $\Delta$ was not properly introduced, and the variable $\omega$ was likely to be confused with the other variable $w$. To address these issues, we redefined the variables to clarify any ambiguities that might appear. We also chose different letters for the output variables: respiration, erosion, and lateral transfer. By doing this, we make it easier for the user to identify these respective rates and fluxes in the following sections of the text.

- Subsections 2.1.3 and 2.1.4: In these subsections, we apply all the aforementioned changes to the core CE-DYNAM formulas. As a consequence, it was possible to reduce their lengthy aspect and clarify for readers the role of each flux and rate in the model. In addition, throughout this subsection and the text, we standardized the use of the terms "flux" and "rate," as well as avoided the use of the nomenclature "a, b" for neighborhood relations. According to one of the reviewers, both points could confuse the reader and should be changed.
- Subsection 2.2.1: In this subsection, we were asked by the Reviewer to clarify what was the source of the LS-factor and the P-factor of the RUSLE equation (Equation 16), and how the monthly erosivity was derived for the historical reconstructions. Both points were changed in the new version of the manuscript.

In our understanding, the new version of the article, with more accurate and succinct representations, has greatly improved compared to the past version, and we hope that such changes will please the Reviewers and the readers of GMD.