Part 1: Responses to Topical Editor

Thank you for submitting a revised version of your manuscript. According to reviewer #1's and my own assessment, the concerns regarding clarity of the manuscript from the former review round were only partly addressed by your revision. Please have a close look at suggestions and examples in the referee report of reviewer #1. A few additional suggestions for improvements are also given below.

Please take some time to go thoroughly again through the whole manuscript and try to improve wording and explanations so that the text is easier to understand for the reader.

Response: Thanks a lot for handling our manuscript and for your comments. Combined with the comments from former review round and the referee report of reviewer #1, we further go thoroughly again through and revised our manuscript to make it clearer and easier to understand. We'd like to reply the comments one by one.

Specific comments:

1. Line 31: Sentence incomplete.

Response: We have rephrased the sentence. Please refer to P1L31~P2L36.

2. Line 42: "...Earth system models...".

Response: We have corrected it. Please refer to P2L42.

3. Line 75: "...is described...".

Response: We have corrected it. Please refer to P3L75.

4. Line 83: "...global communication with ...".

Response: We have revised the related statements. Please refer to P3L82 and P4L100.

5. Line 94: "communicators for the model, the filter and the coupling exchanges between them...". It is probably still not clear to the reader what you refer to here. Please explain in more detail.

Response: We have removed the description of PDAF communicators which is not easy to understand for readers. We then rephrased the statements. Please refer to P3L91~L94.

6. Line 98: "...waiting for the results...".

Response: We have corrected it. Please refer to P4L97.

7. Line 99-102: Sentence unclear. Please rephrase.

Response: We have rephrased the statements. Please refer to P4L98~L104.

8. Line 117: "The most significant..." or "A significant...".

Response: We have corrected it. Please refer to P4L118.

9. Line 118: "...or intra one...". Wrong wording. Use something like "...or within one...".

Response: We have revised the related statements. Please refer to P4L119 and P8L238.

Part 2: Responses to Anonymous Referee #1

My concerns about the general structure, especially the sections 2 and 3, of the paper have not been addressed. The text is still extremely difficult to understand. Below are my specific comments to the text of the section 2 and 3 to demonstrate the frustration of a reader. Besides, the WCDA implementation has not been properly introduced despite my previous concerns. I don't think that this article may be published in its current state in the GMD journal.

Response: Thanks a lot for your concerns and suggestions. In this revision, we have tried to improve the manuscript according to your comments.

Specific comments:

1. l.84: Maybe like this? Such an implementation maintains the independence between the DA and model modules. But the global communications are generally inefficient in the sequential DA systems because of idle processes due to sequential running of model and DA modules.

Response: We have rephrased the statements. Please refer to P3L83~L85.

2. L86: I don't understand these two sentences: In PDAF, a DA method is transformed into a native procedure that is called by the corresponding models via the PDAF application programing interfaces (APIs). Thus, a DA method can share the processes of the model ensemble.

Response: In EMPIRE, a DA method cannot share the processes of the model ensemble, because the DA method is compiled into a standalone executable and two different executables cannot share processes in MPI run. In PDAF, a DA method is transformed into a native procedure of a model, which also means

that the DA method and the model are compiled into the same executable. Thus, a DA method can share the processes of the model ensemble. Please refer to P3L85~L88.

3. L89: DA method shares all processes of the first ensemble member of the corresponding model… Not clear.

Response: We have rephrased the statements with more introductions. Please refer to P3L87~L91.

4. L101. …enables a DA method to share almost all the processes of the model ensemble… not clear

Response: We have rephrased the statements with more introductions. Please refer to P4L100~L102.

5. L102. When a DA algorithm uses processes different from a model ensemble member… not clear

Response: We have revised the statements. Please refer to P4L105~L107.

6. l.106. I would modify this sentence like this: Fortunately, such a challenge has already been overcome by most existing couplers (Craig et al., 2012; Valcke, 2012; Liu et al., 2014; Craig et al., 2017; Liu et al., 2018b). Each of these couplers (?) can transfer data between different process sets with different parallel decompositions without the global communications.

Response: Thanks a lot for your careful suggestions. We have rephrased the sentence. Please refer to P4L108~L110.

7. L116-118. Something shorter like: Combining multiple components into one MPI program is challenging because CC2 can handle exchanges between two components only.

Response: Thanks a lot for your suggestions. C-Coupler2.0 can handle coupling exchanges between two component models or within one component model no matter it is one MPI program or not. But C-Coupler2.0 cannot directly handle coupling exchanges between a DA algorithm and each model ensemble member as their process sets are partially overlapping. We have rephrased the sentence. Please refer to P4L118~L120.

8. L120. Please explain how these three new steps help to solve the mentioned problem.

Response: We have added more introductions about our solutions to the challenge. Please refer to P4L120~L126.

9. L121. The sentence 'These three steps enable all members in a model ensemble to use a DA algorithm cooperatively.' is not clear. Please reformulate it.

Response: We have rephrased the sentence. Please refer to P4L125~L126.

10. l.127 A DA algorithm can include a set of procedures such as observation operators and analysis modules, each of which can be called by the model separately. I would only say: Each module of the DA algorithm may be called separately in the PDAF system. Is this what you are saying here?

Response: We have rephrased the sentence. Please refer to P5L131~L132.

11. L128-130. Instead of 'The framework uses the dynamic link library (DLL) technique for the connection of a DA algorithm program to a model, where a DA algorithm program is compiled into a DLL that is dynamically linked to a model when an instance of the DA algorithm is initialized.' I would say: 'The dynamic link library (DLL) technique is used to connect a DA algorithm program to a model program. The DA algorithm program is compiled into a DLL that is dynamically linked to a model program.'

Response: Thanks a lot for your careful suggestions. We have rephrased the statements. Please refer to P5L132~L134.

12. L130-133. Instead of 'With the DLL technique, a new DA algorithm can be used by a model without modifying and recompiling the model codes, and the original configuration and compilation systems of a DA algorithm can generally be preserved for greater independence of the DA algorithms from the models and for less work in integrating a DA algorithm.' I would say 'Using the DLL technique allows us to couple DA algorithm and model codes without modifying the codes.'

Response: Thanks a lot for your careful suggestions. We have rephrased the statements. Please refer to P5L134~L136.

13. L133-134. The ensemble component manager governs the communicators of ensemble members.

Response: We have rephrased the sentence. Please refer to P5L136~L137.

14. L134-135. The online DA procedure manager provides several APIs that enable the ensemble members of a component model to initialize, run and finalize a DA instance cooperatively. Not clear. And this part of the sentence may be a separate sentence 'automatically handles the data exchanges between the ensemble members and the DA algorithm'. You can delete 'with a set of operations'.

Response: We have rephrased the statements. Please refer to P5L137~L140.

15. L136-138. This is not clear: The ensemble DA configuration manager enables the user to flexibly specify the DA algorithm, DA frequency and the operations for the data exchange in a DA

simulation through a configuration file. What is DA frequency? What is 'to specify the DA algorithm', what are 'the operations for the data exchange'?

Response: We have rephrased the statements. Please refer to P5L140~L141.

16. L139. Not clear: Guided by the architecture in Fig. 2, we implemented the new framework

Response: We have rephrased the statements. Please refer to P5L142~L143.

17. L143. DA algorithms at different frequencies… What do you mean?

Response: Here, different frequencies refer to different time periods to calling DA algorithm program. We have rephrased the statements. Please refer to P5L145~L147.

18. L143. while component 3 does not use DA. Not clear.

Response: We have rephrased the statements. Please refer to P5L147.

19. L158-161. Not clear.

Response: We have removed some difficult-to-understand terms about C-Coupler2.0 and rephrased the related statements. Please refer to P6L161~L163.

20. L169-170. Given an ensemble run of a coupled model, all ensemble members of the component models of the coupled model can be organized as one level of models. Not clear.

Response: We have rephrased the related statements. Please refer to P6L171~L173.

21. L170-174 although we recommend constructing two hierarchical levels of models with 170 the first level corresponding to all ensemble members of the coupled model and each ensemble member including the component models at the second level (Fig. 5b), because the hierarchical organization retains the original architecture of the coupled model through a simple additional registration of the coupled model to C-Coupler2.0.. Not clear.

Response: We have rephrased the related statements. Please refer to P6L172~L176.

22. L174-176. As a DA algorithm that handles ensemble fields can run on the MPI processes of all ensemble members of a component model (Fig. 3), a special C-Coupler2.0 component model that covers all ensemble members of the component model (this 175 special component model is called ensemble-set component model hereafter) is required for using the DA algorithm (Fig. 5b). Not clear.

Response: We have rephrased the related statements. Please refer to P6L177~L179. For the definition of ensemble-set component model, please refer to P4L120~L121.

23. L177-178. The ensemble component manager provides the capability to generate an ensemble-set component model, which does not introduce global synchronization and only involves the ensemble members of the corresponding component model. Not clear.

Response: We have rephrased the related statements. Please refer to P6L180~L183.

24. L180. A pair of a model and a DA algorithm have essentially the relationship between a caller and a callee in a program. Why is so?

Response: We have rephrased the related statements. Please refer to P6L185~L186.

25. L183 what is native code?

Response: We have removed the statement. Please refer to P6L188~L189.

26. L183 a corresponding compiler… not clear

Response: We have rephrased the related statements. Please refer to P6L188~L189.

27. L185. What is host model?

Response: A model that calls a DA algorithm is called the host model of the DA algorithm. Please refer to P6L187~L188.

28. L186. To address the above challenge... The challenge is not clear to me.

Response: Here, the challenge is that when a DA algorithm is enclosed in a DLL and dynamically linked to the host model, compilers cannot guarantee the consistency of the argument list between the host model and the DA algorithm. We have rephrased the related statements. Please refer to P6L188~L191.

29. I stopped reading here. This article would have been an excellent reference for the future users of such complicated system. However, I don't see how it can be helpful for them due to such unclear text.

Response: Besides the revisions of above specific comments, we further go thoroughly again through our manuscript, especially the sections 2 and 3, and we have revised most long difficult-to-understand sentences and added more introductions about some acronyms or special terms to make it clearer and easier to understand. Please refer to sections 2 and 3 for more details.