Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2018-295-AC1, 2019

Geoscientific Model Development Discussions

Interactive comment on "MAgPIE 4 – A modular open source framework for modeling global land-systems" *by* Jan Philipp Dietrich et al.

Jan Philipp Dietrich et al.

dietrich@pik-potsdam.de

Received and published: 20 February 2019

We would like to thank the Referee for the time spent reviewing the paper and the given remarks which in our opinion were without exception useful and helped to significantly improve the paper.

[REFEREE COMMENT 1]: The manuscript reflects an impressive effort: taking an existing model and turning it into a framework while meeting the demanding requirements of open sourcing it (licensing, distribution, documentation, and so on). The manuscript is a well-structured overview of the MAgPIE 4 framework. In places, the manuscript can benefit from clarification and polish: page1_line8: The abstract lists "flexible detail in process dynamics" as a feature. In the main

text this phrasing does not recur, and it is unclear what it refers to: adjustable temporal resolution? Otherwise? Modify to bring the abstract in harmony with the content.

[AUTHORS RESPONSE 1]: "flexible detail in process dynamics" was meant to refer to exchangeability of different realizations of a module as this is not necessarily the case just because code is modular. We rephrased it to "modular structure with exchangeable module implementations" avoiding the term "realization" as its definition is coming later in the text.

[REFEREE COMMENT 2]: page2_lines29-30 "It also means that the complexity of a module realization can be chosen based on the importance of this component for the given question". I presume this refers to the freedom of choosing between different realizations of a module, picking one with a degree of complexity sufficient for the task at hand. If so, the phrasing is incorrect since "the complexity of a module realization" is fixed and hence can not be chosen. Rephrase.

[AUTHORS RESPONSE 2]: We rephrased the paragraph:

[TEXT EDIT 2] "Flexibility in the level of detail means adjusting the temporal and spatial resolution. It also means that module realizations can be chosen based on the research question and thereby adjusting the model complexity appropriately."

[REFEREE COMMENT 3]: page3_lines6 The sentence starting with "An output" is confusing. Suggestion: The main text is completed by an output section – showing some select model output and a specific use case of the spatial flexibility provided by the framework – as well as a conclusions and outlook section.

[AUTHORS RESPONSE 4]: We modified the text as suggested.

[REFEREE COMMENT 5]: page8_lines17-18 Imply that the modularity is implemented in GAMS: "The inner layer written in GAMS (...) including the code modularity implementation". As explained in appendix A, the modularity is in part

enabled by a naming convention as GAMS lacks name spaces, and in part by R code to check that the naming convention is adhered to. This extends beyond what GAMS provides. Moreover, it is the reviewers understanding that further R functionality is used to compose the chosen module realizations written in GAMS to a single GAMS source file. As such, it is inaccurate to imply that the modularity is implemented in GAMS. Rather the modularity results from extending GAMS with a naming convention and R helper code. Please reflect this in the text. Some words on the composition would also be welcome: much emphasis is put on the modularity of the framework, so the text should reflect it accurately and completely.

[AUTHORS RESPONSE 5]: Strictly speaking, the statement that the modularity implementation is part of the inner layer is not wrong. All rules developed to achieve the modular structure are applied to the GAMS code itself and can be handled without additional support from external functions. However, it is correct that in MAgPIE the outer layer is adding functionality to the modular structure, especially monitoring the proper application of modularization rules to the GAMS code. Without the outer layer, it would be rather complicated to detect code violations. Besides it we also have some convenience functions in R making it easier to manipulate the GAMS code in a compliant way to the given modularization rules. To make it more transparent we kept the sentence about the inner layer as it was, but added some information about the supporting extensions coming from the outer layer:

[TEXT EDIT 5] "The inner layer written in GAMS contains the optimization model with all its equations and constraints, including the code modularity implementation. The latter is assisted by the outer layer which is monitoring code compliance and providing convenience functions for easier code manipulation in compliance with the modular structure (lucode)."

[AUTHORS RESPONSE 5]: The run composition to a single GAMS file mentioned by the reviewer is another feature coming from the outer layer which is attached to the

modular structure. However, the modular structure also works properly if the runs are not composed to a single file. Hence, this feature is an independent feature focussing on parallelization and reproducibility of model runs. As suggested by the reviewer, we extended the description of the model run composition by adding a dedicated subsection "Model run composition" to the framework description:

[TEXT EDIT 5] "To allow for parallel execution of model runs and to improve reproducibility MAgPIE performs a model run composition. Purpose of the composition is to isolate the current model run before execution. Isolation is achieved by creating a separate output folder for each run in which all relevant data is copied. The main component of each output folder is a single GAMS file containing the full GAMS model and all inputs. This file is created by replacing all include statements in the original GAMS model code with corresponding input files or code segments. In case of conditional inclusions (e.g. realization selection) only the active inclusion is considered (e.g. the chosen realization). This approach leads to a fully self-contained GAMS file which can be shared and runs standalone. All other files in the output folder are supplementary and either used for run post-processing or provide additional information about the run setup (e.g. the run configuration file). For archiving it is recommended to store the whole output folder as an image of the respective run."

[REFEREE COMMENT 6]: page8_line21 "a physical separation of the respective model code". Presumably this is meant to reflect the organization of the model code in directories and files. If so, using the word physical here obfuscates the matter, and is not accurate given the many layers of indirection between logical and physical storage in modern computing systems. Suggestion: "a hierarchical organization of the respective model code"

[AUTHORS RESPONSE 6]: We removed the sentence as it was indeed misleading and not providing any additional insight.

[REFEREE COMMENT 7]: page9_line3 "Physically a module..." Similar concern

as above. Suggestion: "A module in MAgPIE is represented as a folder..."

[AUTHORS RESPONSE 7]: We modified the text as suggested.

[REFEREE COMMENT 8]: page10_section3.5 discusses the model evaluation. Specifically, line 9 mentions "The automatized model evaluation documents currently validate". As written, this suggests that the documents are automatized and perform validation. From the preceding text, it is clear that instead the PDF evaluation documents are automatically generated, in principle allowing for human evaluation, though, at 2000 pages, practice is unlikely to reflect principle. Rephrase. Suggestion: "The automatically generated model evaluation documents currently allow comparison of about 1,000 output variables with reference data".

[AUTHORS RESPONSE 8]: This is a valid objection. We corrected the description of evaluation documents and improved the wording in the whole paragraph.

[REFEREE COMMENT 9]: page13_line4- The paragraph discusses a revised setup emphasizing Brazil, but reducing the number of clusters elsewhere. It seems implied, but is not explicitly stated, that this serves to keep resource usage tractable or constant. This paragraph can benefit from clarification and more lucid phrasing.

[AUTHORS RESPONSE 9]: The assumption is correct. In our specific case the applied solver limits the complexity because with higher complexity it will be not able to compute a solution to the problem. We made this limitation now explicit in the text:

[TEXT EDIT 9] "Detail gained for Brazil has to be bought with reduced detail for the rest of the world to keep the model complexity manageable for the applied solver."

[REFEREE COMMENT 10]: page15_fig5/page13_lines11- Some words on the causative mechanisms for the marked outcome difference between the default and Brazil setup would be welcome.

[AUTHORS RESPONSE 10]: This is an important remark! Checking the estimates again and looking for an explanation it turned out that our initial analysis was incomplete, ignoring some model internal dynamics. A deeper look revealed that the global deforestation estimate in the Brazil setup is actually unreliable as it is caused by unrealistic production shifts within the Rest of the world (ROW) region. We corrected that in the revised version and added some discussion for the figure:

[TEXT EDIT 10] "Comparison with historical data sets as well as projections on forest cover show that the differences between mappings are rather small compared to the overall uncertainty in these numbers. Nevertheless, a deeper look into the simulations uncovers that the global numbers of the Brazil-centric setup are unreliable as the reduced deforestation rate compared to the default setup is a consequence of the applied mapping. As the ROW region basically acts as a huge free trade region it can fulfill strong demand pressure coming from Sub-Saharan Africa with production from elsewhere, while trade limitations in the default setup limit this exchange and trigger deforestation within Sub-Saharan Africa Africa (Dietrich, 2018, compare m4p_default_validation.pdf p1558 and m4p_brazil_validation.pdf p1465). In the case of LAM both runs show a rather similar picture in the aggregated forest cover projections for the region and it is not possible to clearly reject one of them. This is particular important as the regional aggregates in LAM are in the scope of both mappings and therefore should be sound. When choosing between them, one has to decide whether spatial details in Brazil or global trade patterns are the more decisive factor for accurate estimates of regional forest cover in LAM."

[REFEREE COMMENT 11]: Correction suggestions for spelling/syntax/punctuation:

- **page1_line5 computationally intensive**
- **page3_line27 region-specific**

- **page5_line1 choose a regional aggregation, with the country level () as the highest....**
- **page5_lines45 food-demand**
- **page8_line28 realizations**
- **page9_line25 The model outcomes at the cluster level**
- **page9_line27 data pre-processing at ISO country or 0.5 degree level**
- **page13_line4 less -> fewer**
- **page16_line11 and other research institutions, as enabled by**

[AUTHORS RESPONSE 11]: We corrected the text as suggested.

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2018-295, 2018.