

Geosci. Model Dev. Discuss., referee comment RC1
<https://doi.org/10.5194/gmd-2021-169-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2021-169

Anonymous Referee #1

Referee comment on "C-LLAMA 1.0: a traceable model for food, agriculture, and land use"
by Thomas S. Ball et al., Geosci. Model Dev. Discuss.,
<https://doi.org/10.5194/gmd-2021-169-RC1>, 2021

This paper presents a relatively simple non-economic model representing the food system and its relation to agricultural production and land use based on empirical data from the FAO. While the ambition is interesting and laudable the manuscript needs major improvement before publication. Also, I have major concerns with several key assumptions made in the modelling.

General comments

- The introduction is very poor. The specific models that are referred to as 'opaque' are not discussed in any detail. These models are the agro-economic components of the IAMs, specifically GLOBIOM, MAgPIE, GCAM, AIM, IMAGE. While these models might be considered relatively complex, they are described in high detail in numerous publications and model intercomparisons and are based on fairly basic economic principles. Moreover, some of these models by now are open source available (GCAM <https://github.com/JGCRI/gcam-core/releases> and MAgPIE <https://github.com/magpiemodel>) or have detailed model descriptions online (IMAGE <https://models.pbl.nl/image>), and for all models large amount of results are available for public use (<https://data.ene.iiasa.ac.at/iamc-1.5c-explorer>). The JULES and LPJmL models that are discussed explicitly are a very different type of model that do not consider the food system and are therefore irrelevant for this discussion. A key improvement for this paper needs to be a proper discussion of why the type of model presented here is of relevance and interest, especially in relation to models describing the food system in the literature. Currently the only explanation for the endeavour undertaken here is that these are opaque, which is not sufficient. In addition, an argumentation is needed why a move from a simple global approach to a country-based approach is an improvement as it inevitably increases the complexity of the model. Currently there is only an explanation of the advantage of moving from excel to python which is not so relevant for this manuscript in my view (though definitely an improvement).
- Throughout the manuscript, there is extensive reference to the FALAFEL model which is

not described itself. Also, in the results section comparisons to results from the FALAFEL model are presented. If a model comparison is made, the model to which it is compared also needs to be presented and explained in detail. I think however that it is preferable and more clear if the paper only discusses the C-LLAMA model and does not lean too much on results from an older model. On the other hand, the paper would benefit from more extensive discussion of the results in light of food system trends in the recent past: do the projected developments of land use, food and feed production and per capita consumption make sense compared to real-world trends of the last twenty to thirty years?

- Some of the key assumptions made in the modelling are surprising to me and might require revision:
 - The agricultural industrialisation parameter is based on food energy consumption per capita. This is surprising as this variable is very indirectly related to agricultural industrialisation. Indeed, highly modernized countries such as Korea and Japan show very different consumption for cultural and physical reasons that are completely unrelated to farming practices. On the other hand, middle income countries such as Morocco, Algeria, Turkey and Romania have higher consumption than high income countries such as Netherlands and Australia, while the former definitely have less industrialized farming than the latter. There are variables available that are more directly linked to industrialization. I would recommend to look into yield gap data, for example from the yield gap atlas (<https://www.yieldgap.org/>, no global coverage unfortunately) or from analyses such as Mueller et al 2012 (<https://www-nature-com/articles/nature11420>, supplementary data available per country). Also, the industrialisation parameters should definitely be shown in the appendix as it seems to be the dominant parameter as shown by the sensitivity analysis.
 - The model does not take into account other use of crops. Most notably bioenergy is crucial, with nearly half of all maize production in the USA and nearly half of sugarcane in Brazil used for bioenergy. Also cotton production for clothing is important.
 - The assumption that all countries converge to 3200 kcal per capita by 2100 is crucial for total production estimates. While such an assumption is reasonable in stylized scenarios assessing the effects of dietary change, I don't believe it makes sense for business-as-usual projections. In the recent past consumption in countries above the 3200 kcal thresholds have shown no substantial decreases in consumption, so this would be a clear break of the historical trend. On the other hand, in countries that for non-economic reasons have relatively low consumption (e.g. earlier discussed Japan and Korea) there is no reason to assume an increase to 3200 kcals.
 - In the second sentence of section 3.5.1 it is explained that the land use area results are projected based from the most recent FAOSTAT value. I find this very odd and it is unclear to me how much effect it could have. Are resulting land use areas from the model very different from the FAO reported values in some regions? This could indicate important mismatches in the modelling and I would like to see these data. And why don't you choose to use a calibration factor? In my view this would not violate the purpose of the model to explore sensitivities (a goal which should be mentioned in the introduction if it is indeed a key purpose of the model).
- The results section only presents data on per capita consumption and land use. I would be especially interested in total required production of food and feed crops, fodder, yields and livestock efficiencies. To be able to understand the model dynamics it is quite important to add these. Also all results should be presented at the minimum at the regional level (and if possible even country level in appendix).

Detailed comments

Introduction

Line 16-18: these statements need references. I would maybe refrain from discussing mitigation strategies such as afforestation and BECCS in the introduction. This article is about the food system and resulting land use: the GHG emissions of the agriculture itself and land-use change, as well as other environmental impacts are more relevant in that context.

Line 25-26: 'United Nations 2015' reference is not available in the reference list, if this is about the Paris agreement I think it should refer to the UNFCCC.

Line 26: I was taught that 'land-use' only needs a dash when used as adverb, e.g. in land-use change, but not when it is used as a noun. However I am not a native speaker so not sure, but please check.

Line 27: the statement 'somewhat opaque' is very vague and not sufficient as underpinning of the added value of this model (see also general comments).

Line 37-40: I don't know the exact policy of the GMD journal, but in my view the programming language used does not require lengthy discussion in a scientific article. Also it is quite self-evident that python is advantageous over Microsoft excel. The fact that your model is open-source is more relevant and can be mentioned as an argument for the transparency of your approach.

Model overview

Line 50: a normal ref to the FAOSTAT website is sufficient here.

Table 1: the names of the program modules is not so relevant for a journal in my view. It is important to provide the model open-source including a documentation of the different modules and how it can be used. You might consider adding it to the appendix, but I don't think it should be part of the main text.

Line 72 and figure 1: it is quite key to know which processes operate at what spatial

aggregation. Maybe this can be included in the model structure overview.

Line 73: 'global food production' - I assume you are talking about total agricultural production here, including all crops and animal product consumption. This is definitely not all directly used for food, with crops produced for bioenergy, animal feed, clothing, etc. Please reconsider your terminology.

Line 77: 'discounted' - the term discounting in my experience is used in economics to take into account that anything in the future is 'worth less' (simply put), it might be more clear if you use the term 'excluded' here.

Lines 73-95 and figures 2 and 3: a lot of text and figure space is used to explain why a few countries are not considered in the model. In my view this is more appropriate for the appendix and can do with a short description in the main text. The space that is saved on these figures can be used to add additional figures with model results.

Model components

Line 118: a linear regression is made, but it is unclear of what. What are different variables used? Also it is unclear to me what the effect of this is although it is very important whether people eat relatively more animal or vegetal products. Please present the regressions in the appendix.

Line 239-241: these yield assumption could be quite important for the model and are unclear. Does this imply yields in for example Africa can never exceed yields that have already occurred somewhere within the region before, even though most of the continent has relatively high yield gaps? And at what resolution is this analysis done: at the country-level and for each crop? Quite important to add these maximum yields to the appendix.

Model output

Line 355-356: it is irrelevant to explain that you write out data as a CSV.

Line 360-361: also the discussion on storage space is irrelevant in my opinion, but maybe GMD has a policy on this that I am not aware of.

Figure 4: I am quite surprised about certain trends: the Americas show a strong increase in both crop and pasture land from 2015-2050 while this has not been the case in the last 30 years. The strong increase in pasture in Asia is also odd, as pasture areas have been very stable in the last decades. Can you explain this? If the trends are so different from the recent past, is the model then actually up to the task? Also you might consider excluding the same countries for the historical period as you did in the modelling to make the graphic more clear.

Figure 5: it would be great to combine this graph with historical data similar to the land use, to see how the trends compare to the recent past.

Figure 6: the difference in land use expansion between C-LLAMA and FALAFAL is very large, most notably for pasture. Please discuss why this is the case. Also it is surprising that the data are so different in 2015. Why is this?

Figure 7: this figure is very difficult to interpret as the changes are so small. You could consider showing changes over the 2015-2050 period instead.

Discussion

The discussion requires an additional section that compares the results from C-LLAMA to the wider literature: a lot of models are available that project land use. Please discuss both the type of model and the results in this context.