

Geosci. Model Dev. Discuss., referee comment RC1
<https://doi.org/10.5194/gmd-2021-66-RC1>, 2021
© Author(s) 2021. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2021-66

Anonymous Referee #1

Referee comment on "Spatial agents for geological surface modelling" by Eric A. de Kemp,
Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2021-66-RC1>, 2021

General comments:

I'll start with a caveat: in retrospect, I am probably not be an ideal reviewer for this paper, as I have no experience in the application of solid modeling techniques to infer and visualize subsurface geological structures. However, a silver lining maybe is that I can provide a general geoscience perspective on this paper.

(One terminology note: the phrase 'geological modelling' is used here to mean creating digital representations of 3D sub-surface geological structures, but the same term also refers to the use of numerical models to study the dynamics of geological and geophysical systems. These are very different things, so a definition early in the paper would be helpful.)

This paper makes two valuable contributions. First, it evaluates the use of agent-based modeling techniques, which are widely used to study system dynamics in fields like ecology and sociology, for inferring and graphically representing 3D geological structures from sparse data. Second, it presents a new open-source software package to carry out this kind of modeling, using the NetLogo package. To the best of my (admittedly limited) knowledge, both of these represent novel, interesting, and valuable contributions. And clearly a tremendous amount of effort has gone into developing the ideas and the accompanying software implementation. For these reasons, I feel that the manuscript is appropriate for Geoscientific Model Development, and should find an audience among geologists who are interested in inferring subsurface geological structures from limited observations.

One of the challenges with this manuscript is that it assumes a lot of background knowledge on the part of the reader, which risks limiting its impact. One recommendation therefore is to do some fairly thorough editing to make it more widely accessible. One thing that I believe would help is some re-ordering of the information presented,

especially in the introduction. Such re-ordering could make this manuscript accessible to a wider audience, and therefore draw more attention to this important work. Here's one potential re-organization of the introductory material:

a - Explain briefly what is meant by geological modelling in this paper, and why it matters (a few sentences or a paragraph, with general references for those unfamiliar with the field/area).

b - Explain the major challenges that this paper and the techniques and software it describes are meant to address (e.g., sparse information about the subsurface, uncertainty in 3D location and/or properties, long computation times using standard algorithms, or whatever)

c - Explain why current methods are limited (i.e., why do we need a new and different approach?), and thus why it's worth trying an alternative approach.

d - Give a quick background summary (2-3 sentences) on agent-based modeling that gives readers a basic sense of what it is, e.g., that it's a technique used in simulation modeling of complex systems in which individual entities such as animals or households interact with an environment and with one another.

e - wrap up the intro with a quick summary of what this paper does (maybe this is a good place to say you're doing this in the context of Loop 3D, and explain briefly what that is).

This way, by the end of your Intro section, readers will know what problem you are trying to solve, why it matters, why it is a problem, and what (basically) your proposed approach is. You have all of the pieces already, but they are currently presented in an order that risks leaving readers drowning in a sea of jargon before they have a chance to get to the cool new ideas and techniques.

A related challenge is that the manuscript does not really articulate (at least not until deep into the details) what is wrong with current modeling algorithms and why an agent-based approach might offer a better alternative. A couple of sentences articulating this somewhere in the introduction would help motivate the rest of the paper.

One general grammatical note: the manuscript contains quite a few incomplete sentences, which need to be fixed before final publication. I have flagged some of these below with 'inc'.

Specific comments (by page and line numbers):

1, 13-15 this opening statement of the abstract is a nice example of the challenge readers face with this piece (but there is a potential fix). For a reader to care about the topic in this sentence, they have to know what is meant by 'spatial and property interrogation functions' and 'estimations and construction operations'. It's only at the end of the sentence that we get to the heart of the matter: 3D geological surfaces. You have an opportunity to invert this in a way that could make it more accessible. Start off with a statement about how important it is build 3D-rendered models of geologic structures, and how the sparsity of data makes this difficult. Then state that agent-based modeling provides a potential solution, and that the contribution of this paper is to test it out. By laying out the ideas in this sequence, right away you have presented an interesting research challenge, and followed up with a solution. Now the reader has a reason to care about technical things like property interrogation functions.

1, 18-19 geologic modelling has multiple meanings - please define what you mean by it in this context.

2, 3-22 As noted above, the introduction is hard to follow. It could benefit from a clearer articulation of the nature of the goal and the need for solutions. In particular, it sounds like the core problem is one of limited data, so why does the answer lie in software architecture? How do those two things relate?

2, 23 - 3, 1: I have some familiarity with agent-based modelling but I worry that readers who have never heard of it would struggle to get the basic idea from this somewhat abstract description. Maybe it would help to start off not by identifying different categories of ABMs, but instead to convey the gist of the technique first and what fields it is used in (which you address later in the MS; e.g., to simulate individual entities in computer games, animals in ecological simulations, households or sectors in economic simulations, etc.) I do like the carpenter example (examples are always helpful) but in that particular example it's not clear why 'single agent' applies (is the carpenter one agent, and the house another? is there only one carpenter?)

3, 8 Surely ABMs can themselves be computationally expensive. It is not clear why they would be expected to reduce computation time relative to whatever the alternatives might be (presumably some kind of 'global' algorithm?).

3, 15-18 This is a useful statement of the paper's objective. It sets up readers to expect to learn next what are these graphical tasks, and why agent-based techniques makes them easier or more efficient or more effective, etc.

3, 20 - 4, 2 This is a nice overview of ABM applications. Consider moving it before the more abstract discussion of agent properties. It gives readers a sense for how and why the technique is used.

4, 3-15 It is not totally clear what you mean by anisotropy and gradient-type information. Consider leaving that bit for later, when the application examples will make it easier to understand in context. More generally, there is a lot of jargon in this paragraph: to make sense of it, a reader has to know what you mean by multi-scalar environment, model element interactions, multi-source physical dependencies, non-centralized control structures, global partitioned data structures, etc. Consider either adding text to define these various terms, or deleting them, or replacing them with more accessible descriptions. That said, the examples are great; in general, more examples and fewer jargon terms would help.

section 1.3 this is a good description of the problem of interpreting a 3D subsurface geologic structure given sparse data. Consider moving it, or a suitably edited version of it, up near the beginning of the paper, where you are framing the fundamental problem to be solved.

section 1.4: It is great that there are example codes provided.

section 2, 2-5 This sentence or something like it would be helpful in the first paragraph of the paper to educate the reader on what is meant in this context by 'geological modelling'.

6, 11-12 and fig 1: nice illustration of the challenge, and how different algorithms can come up with very different solutions given the same data. Suggest moving this figure up closer to the opening of the manuscript.

8, 19 I'm not familiar with the terms 'line and fabric densification' or 'contact estimation' - some definitions would help here.

section 2.1 This section says a bit about what tasks agents are used to achieve (e.g., interpolation) but so far one thing that is missing is an explanation of why agent-based algorithms might be expected to do a better job than the more conventional alternatives. For example, is the idea that a large number of agents, each implementing a fairly simple set of rules, would be easier to program, or faster to run, or less costly in memory resources, than a running a single but presumably more complicated 'global' algorithm? Or that an agent based approach makes it easier to adapt sampling density according to local features in the data, because agents can 'signal' one another or spawn new agents in response to finding something 'interesting' (by whatever criteria) in the data? Whatever

the case, this paper would really benefit from a concise statement about how and why agents might be expected to provide a better solution (and along the way, to tell us a bit about what the standard, non-agent algorithms and methods look like, and why they are problematic).

8, 21-23 The idea that agents construct triangular meshes from point data is alluded to here, and shown in Fig 4, but the description is relegated to an appendix. Yet this actually seems important enough to consider pulling into the main text. At this point in the manuscript, the end of page 8, the text has not yet offered much description of what it is these spatial agents actually do. It would help to have a description of at least one example of a particular agent algorithm, to help readers picture how this is meant to work. For example, does an agent 'move' around the space at random until it finds a point, and then create a triangular mesh element? A verbal description of the algorithm, even if it is just one out of several algorithms that are used in the accompanying software, would really help a lot in understanding how this approach works. We learn later, in section 3, about how some of these agents operate, but by giving an example earlier in the manuscript you can give the reader a more intuitive sense of what this kind of agent-based approach involves.

11, 6 what is a 'free' agent and how does it differ from any other type(s) of agent?

11, 12-13 can you say more about how an agent would perform interpolation? Would it be limited to a particular neighborhood of points, and if so how is that neighborhood defined? Again, why would this be preferable to just doing a global IDW or SLERP?

Section 2.3: now we are finally getting a look at advantages of agent-based methods over traditional approaches. It would be really helpful to give a brief summary of this information in the introduction, so readers understand the motivation for this alternative approach.

13, 6 Would you not have issues with inter-agent communication in a parallel distributed system? It certainly seems worth experimenting on, but I know that the somewhat similar method of discrete-event simulation has challenges with parallel operation because you can't predict in advance when initially independent operations on different parts of the model space will end up triggering simultaneous or conflicting modifications to the same data (still, the comment about 'yet to be tested' is fair enough, and I agree).

13, 18 This is interesting - can you say more about 'interrogators or observation sets'? Does this mean for example that you use agents both to represent observational features (such as a known dip at a particular (x,y,z) location) and to perform actions like interpolation?

Section 3: it is great that there are codes provided for each of these examples.

15, 2-5 This specific example of an agent's behaviour is really helpful. Consider presenting this or another example much earlier in the manuscript, not as a comprehensive description of the different agent types and their behaviours (you already have that material down here in section 3), but just to give the reader a general idea of what you mean when you refer to 'spatial agents' in the paper's introduction. (Something like: 'For example, to help interpolate the surface of a dipping geologic unit, one could define an agent that moves randomly through the data volume until it encounters a dip observation, at which point it uses a local interpolation algorithm to spawn new dip markers in the area around the observation'...or something along those lines, though actually Mesh agents might be an easier example to understand).

Technical corrections:

2, 7 17 inc

4, 6 inc

5, 2 inc

12, 10 inc (also, not obvious what 'contributions from multi-scalar and deep multi-property data' means)

12, 12 inc

12, 17 inc

13, 7 inc

15, 13 criterion

19, 7 inc

20, 3 inc

20, 9 was

20, 19 inc

21, 3 inc