Reviewer comments – Automated Geological Map Deconstruction for 3d Model Construction

R.M. Harrap, Queen's University, Kingston, Ontario.
harrap@queensu.ca

General Comments

This is an important contribution to an area of growing importance and relevance. The paper thoroughly describes a tool which is significant both in terms of general approach and highlighting areas of map to model automation that are or are not challenging, and also exposing the areas in which geological mapping and the thinking process that geologists go through show significant ambiguity that plagues later use of that data. While the paper is quite focused on the tool, the deconstruction process and the discussion of areas in which the tool can (or perhaps cannot) be improved are of general interest.

My only general objection to the form of the paper as it now stands is that there is no description of how the tool is actually run – is it an iterative process with a human intervening, or a batch process where a model run that fails or succeeds can be rerun, or… exactly what? Many tools we use in geology, and perhaps those that are the most influential, are used as what-if thinking tools in a way that becomes second nature to an experienced user. I'd like to see some discussion of how the tool fits into the workflow of a geologist/modeler 'as a tool.'

I highly recommend the paper be published with minor corrections. Thank you for the opportunity to be involved in the process.

Scientific Specific Comments and Requests for Clarification

Some of these are quite minor points, some may reflect me misunderstanding the paper (which says something) and some are fairly technical. The format is line number or range followed by the point, however large or trivial.

30: Why say deformed? Doesn't it work on flat-lying strata?

35: The point of this work (I'm the author) was consistency checking both *during* and *after* map creation and especially when large, complex compilation maps were being created and to focus on areas where a legend contradicts map relationships. This was implemented at one point in an unpublished MSc thesis and it worked. The point here is that workers such as Brodaric and Harrap were always interested in tools used *during* mapping as well as *during compilation* and not simply in analyzing others maps. Not sure this requires any change in the paper, but I thought it interesting to point out given the long and slow history of developments of tools in this area!

43: I'm not sure what irreproducible means here.

46: Is that transformation unique? Or is it an interpretation especially in cases where legends are being adapted? Brodaric, and Coleman-Sadd (iirc) did some work on reconciling maps where sub-maps had different authors and different emphasis during mapping.

51: As well as for use in other studies?

54: … lack stratigraphic information. This reads oddly since drill holes in non-hardrock terrain would never lack stratigraphic information? Perhaps I'm misunderstanding here.

55: The 3 types specified here – might be expanded with one sentence of more detail on each? And why (just) these three? As you'll see below I have comments about others though you address that very well in your Discussion in the end.

58: Combine direct with conceptual?

a) What is knowledge from nearby areas?
b) Tectonic history (okay, you cover this)
c) What about base scientific knowledge? How granites behave?

(you might just note that you discuss this in the discussions but… I think it is important to point out what forms of conceptual knowledge that geologists use are NOT in your model here…)

60: A cross-section is an artefact o f us working with a-c reasoning for a good enough (the A.I. term is satisficing) 3d model. It feels like you are stumbling over 2 fundamentally different *kinds* of constraints: this unit has "this place and this geometry" versus "this unit is *younger* than that unit." The kinds of inference / constraints used in geological reasoning are well studied e.g. for the case of how people map and how they interpret data (e.g. Bond and students, etc. etc.). I don't expect a huge change here but flag *something* about what is in and what is not for your discussion later asper the previous comment?

63: Do you really want to reproduce them? Or get the same or better resulting outputs with reproducibility?

67: Are you capturing stuff people don't "bother" to do or stuff that is intractable to do manually at all?

70: Is the summary human readable? Further to my general comment, would that be desirable for human AI model development? Mostly addressed, but…

94: Given the point about modeler 'in the head' knowledge and 3d Models I might say 'un-related to 3d Modelling as currently practiced'? Not sure…

100-104: Was there any co-evolution of these libraries and your stuff as your project went on or are you totally arms-length?

114: I think you need to cite something on the geology of the area? A report?

118: Confusing. I read this as you saying the Jupyter notebook hits those layers and then provides a human-in-the-loop UI/Configuration file writer / process? Consider putting the "there are currently" sentence first for clarity?

122: It feels like you are saying this is a legend-language style depiction but I think you are actually saying it is a recoded _map_ in a unique chronostratigraphic order. So … someone has to 'do' that first…?

133-43: How do you / can you deal with adjacent map sheets with different levels of detail? Or are you 'one map only' or would a prelim recode be required? Perhaps in discussion or not at all.

145: Was this map chrono… suitable or did you recode?

A general question about this point. In field geology we use 'established, inferred, assumed' language for contacts. I'm assuming that you just have one contact type. This becomes VERY important in your discussion and conclusions when you talk about issues because not all contacts are fixed. Some are topologically required but can be 'safely' moved as they are under cover / underconstrained. So… is the 'fixist' logic HERE a problem later when you might want to reconcile some of your issues by shifting a boundary in the map because it is 'unbound' or 'unconstrained?'

191-3: I find this confusing. The way it is worded L167 filtered by… implies that it is done FIRST as data prep. But is it happening along the way? During the run?

It feels like you should discriminate things that are done 'totally before' form those that happen during… Yes, I'm being a bit picky here but I'm trying to understand the workflow (e.g. general comment about what the experience of using it is like).

(3)

217: Not sure this is meaningful because I'm not reading your code, but would it be possible to modify Figure 4 to show what is map2model and what is map2loop, or is the back and forth too complex to make that meaningful?

220: DTM is online only whereas other data can be local?

231: What about lenses in stratigraphic sequences that pinch out? Can you handle that topology?

255: Is the unique fault name auto-generated if the fault was not named in the source data?

325: 3.2 Need to remind us what these are, how they are stored? Sentence here feels like it should actually start 3.2.1 and not be a general section intro?

330: Overrides – implies you either would do this before, or during, or after and rerun, or??? Clarify the workflow please!

380: This is an area where I hoped in the discussion there would be more … discussion … about how you validated your tool as you worked on it? As in, more discussion of running with different map types, geometries etc. It is a bit too succinct there for my liking but perhaps you are at a length limit for the paper…

390: At this point I think a high level description would help. Something about 'first we do the stratigraphy, then the intrusions, then the faults, then the…. I struggled reading this trying to decide whether you were using stratigraphy as a catch all term or just for… stratigraphy. Your language is a bit ambiguous here, and a very short intro on the process would help to say what is covered and what is not.

414: It feels like this could be a parallel combinatorial approach; mostly addressed in the discussion.

427: is the graph exportable/ usable?

439: Again, the workflow. Is this process interactive? Iterative? Black box? Is the human in the loop? Especially given the statement below (geoscientist in minutes etc). Expose your workflow more clearly!!!

449: (perhaps) one sentence on what geophysics would add?

514: No map… Actually as you hint at later, there are MANY situations where a human can make a map using regional knowledge, stated assumptions, theory, potential field data… The interesting thing is the gap between formally decidable maps and what a geoscientist would be comfortable to handle. There is abundant literature on this e.g. in field mapping and in interpretation (Bond, as you cite). Is the gap large? What are those humans DOING? Are they nuts? Brodaric has talked a lot about this as have the cognitive scientists who study the mapping process. There is a fairly interesting philosophy of geology topic you're going to hit if you continue in this direction (as an aside, this has been examined in the case of highly constrained geotechnical 3d modeling, but only at the very local scale and for a few unconstrained parameters). **There is a huge gap between a provable and a plausible model.**

550: Fascinating topic. Could something like a generative adversarial network approach be used (build a procedural simulator, like Noddy on steroids, that then subsets it's outputs to feed into map2loop which can then be verified since … the full inputs were known).

565: This is the combinatorics I was referring to. Specifically, the work that has been done on combinatorial possible worlds (this is just an aside…)

A general comment, I was surprised that Varnes' paper doesn't show up in your early intro as it was the source of a lot of early thinking. Not a requirement, but it is an interesting read!

Technical Corrections

50: … and to extract …?

52: Commonly… often… two in the same sentence seems a little wishy-washy.

53: or, if available, logged well data…?

55-56: Feels like 34d category should be in same sentence? Feels like an edit error here.

67: Not previously available. (missing .)

68: … from GIS layers stored locally or online servers?

80: One too many) ?

81: these packages?

84: and so this necessary toolset will not be discussed further here?

361: A single item list?

366: Compare reference at 335 to that here and make your style consistent with journal standards?

389 / 834: Perrin reference is incomplete? I didn't go through them all very carefully but…