1. Does the paper address relevant scientific modelling questions within the scope of GMD?

Yes. This is an important and timely contribution, while a number of open source codes are now provided to the community, there is a distinct lacking of analysis tools. Another more general comment, it is kind of odd to call this open source since it is built on MatLab. Indeed what you provide is open source but it only works with MatLab and someone who does not have access to MatLab will not be able to use it. You do mention this in the abstract but it's an important point. Finally, extending StagLab beyond output from StagYY is important if it will be adopted as a community tool, since
of course, StagYY is not available to many/most geodynammicists. One might find it ironic that a suite of open source analysis tools are being made available for a code that is not generally available. Yet this is a step in the right direction. Fluidity also has a fairly small user base. It appears that StagLab is extensible, so perhaps the challenge will be taken on by other users to write a function to import other formats. There seems to be a matvtk plugin for MatLab, I don’t use MatLab anymore. I will note that vtk output is fairly standard and codes such as Citcom and Aspect support vtk output.

These are not a list of requirements to publish but things to think about. The idea of providing robust open-source analysis tools will greatly improve geodynamic modeling. Hopefully this will catch on and others will participate in the effort. One significant but time consuming step would be to translate this to python, which would free it from MatLab and make it really open source. It would be a significant effort but could be beneficial in the long run.

2. Does the paper present a model, advances in modelling science, or a modelling protocol that is suitable for addressing relevant scientific questions within the scope of EGU?

Yes. It is worth pointing out that the end the rainbow section is broader than geodynamic tools and this is an important contribution that could almost stand on it’s own. It is relevant to all readers of GMD. I recall 25+ years ago researchers in visualization came to an AGU meeting showing how rainbow type color palettes distort perception. Ironically this session was juxtaposed with a session that introduced several new tomographic models were presented, of course using a rainbow-like (redinium/blutonium) palette. The references here are important and I’m glad to see this is once again a topic and I hope that with the sharing of tools and availability of these new palettes, I hope that others will adopt them.

3. Does the paper present novel concepts, ideas, tools, or data?

Yes.
4. Does the paper represent a sufficiently substantial advance in modelling science?  
Yes.

5. Are the methods and assumptions valid and clearly outlined?  
Yes.

6. Are the results sufficient to support the interpretations and conclusions?  
Yes – not really relevant to this paper.

7. Is the description sufficiently complete and precise to allow their reproduction by fellow scientists (traceability of results)? In the case of model description papers, it should in theory be possible for an independent scientist to construct a model that, while not necessarily numerically identical, will produce scientifically equivalent results. Model development papers should be similarly reproducible. For MIP and benchmarking papers, it should be possible for the protocol to be precisely reproduced for an independent model. Descriptions of numerical advances should be precisely reproducible.  
Yes.

8. Do the authors give proper credit to related work and clearly indicate their own new/original contribution?  
Yes.

9. Does the title clearly reflect the contents of the paper? The model name and number should be included in papers that deal with only one model.  
Yes

10. Does the abstract provide a concise and complete summary?  
Yes.

11. Is the overall presentation well structured and clear?  
C3
Here there are some things to talk about. The output diagnostics begin with a list of specific diagnostics to subduction problems, which limits the usefulness of the tool. Beginning with more general output then moving to more specific cases would seem appropriate. In that sense starting with section 2.2 would seem more logical. What about Vrms and Nusselt number, and mean temperature as a function of depth? These all seem to be routine diagnostics of flow. As for plate-like flow, many people use plateness. Some of this seems to be embedded in sections 2.1.3 and 2.1.4 but they are more general than “plate tectonic diagnostics” and I would not expect to find stagnant lid diagnostics within a section labeled plate-tectonic diagnostics. Organization of this is a bit scattered and does not help the interested reader. I would suggest something like: Generic flow diagnostics (Vrms, Nu, heat flow, dynamic topography, geoid?, mean temperature, depth average temperature and velocities); then plate/slab diagnostics, then plume tracking, etc.

12. Is the language fluent and precise?
One comment here. "STAGLAB’s software design ensures an enjoyable user experience." Enjoyable? I’m not sure about this. How do you quantify that statement? After all what is enjoyable for one may not be for others. As the students in my geophysics class. I think you are trying to convey the idea that this is easy to use, flexible, and robust.

13. Are mathematical formulae, symbols, abbreviations, and units correctly defined and used?
Yes.

14. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated?
no

15. Are the number and quality of references appropriate?
Yes.

Is the amount and quality of supplementary material appropriate? For model description papers, authors are strongly encouraged to submit supplementary material containing the model code and a user manual. For development, technical, and benchmarking papers, the submission of code to perform calculations described in the text is strongly encouraged.

This is fine.