Geosci. Model Dev. Discuss.,
https://doi.org/10.5194/gmd-2019-82-AC2, 2019

Geoscientific Model Development Discussions
Open Access
EGU

Interactive comment on "pygeodyn 1.0.0: a Python package for geomagnetic data assimilation" *by* Loïc Huder et al.

Loïc Huder et al.

loic.huder@univ-grenoble-alpes.fr

Received and published: 5 July 2019

The manuscript presents a python software package/library for the simulation and data assimilation of geomagnetic models. The packages provides a surface dynamic model, a reduced order model based on autoregressive processes, geomagnetic observations, and an data assimilation method (the augmented Kalman filter) in a single package. All of the results in the paper are easily reproduced by downloading pygeodyn and the plotting package webgeodyn also developed by the same group. Although there are some significant deficiencies in the software package itself, as well as a lack of accessible user manual, the paper present a comprehensive description of the default features and data. I recommend accepting the paper after some minor revisions are done. The software on the other hand, is far from ready for widespread user adoption.

Printer-friendly version

Discussion paper

If the main purpose is to make this package accessible, I strongly urge the developers to provide a human-readable user manual, tutorial, and customization examples.

Major Comments:

1. My major concern is that the software is far from ready for use by non-advanced python expert. The code is written is such a way that it's all but impossible to read and understand, much less modify to include new data, models, assimilation techniques. Worst of all, there is no proper documentation detailing the structure of the package, objects being used, and organization of the assimilation system. These are indispensable elements for customization and none are present. If the developers really want a widespread adoption of pygeodyn, then they need to work hard on making the software accessible and well documented. To be completely sincere, I wouldn't recommend this package to anyone in the geosciences community.

   ⇒ **In fact, such information existed under the form of a README in the root folder, an advanced guide in the doc folder and the online developer documentation. Still, we agree that is was not easy enough to find nor organised enough. Following the referee's comments, we have totally reworked our documentation in the new version (1.1.0) in order to present more clearly all the needed information (Now online at https://geodynamo.gricad-pages.univ-grenoble-alpes.fr/pygeodyn/ and offline in the docs folder after HTML generation). Namely:**

   - **The package READMEs were broken down in several sections that were also expanded.**
   - **These sections are now navigable online and comprise:**
     * **Installation instructions**

- **A brief scientific overview of the algorithm**
- **An expanded description of the `run_algo` script (including structure)**
- **Tutorials on the definition of new types and on the reuse of forecast/analysis steps**
- **The developer API that was originally online**

We are grateful to the referee for triggering this documentation rework that should be a big step towards the accessibity of pygeodyn.

2. The git repository should only be used for the python software and not for the data. I strongly suggest that the data be stored on a separate repository or server since it is over one gigabyte of data. It makes no sense to store the data together with the python scripts.

   ⇒ **We followed the referee's suggestion by separating the package sources from the data, each having now their own repository. We provide the user the commands to either fetch only the sources or the complete package.**

Minor Comments:

1. On the Introduction (page 2, lines 3–6) the authors mention that there are two main families: sequential and variational. This might be an oversimplification since the 3D-Var is a variational method that is sequential, and ensemble Kalman smoother is not a variational that is a smoother. I suggest the authors rework this sentence since it's misleading.

   ⇒ **We agree with the referee, and now present the two main families of DA tools as being the variational (minimizing a cost function) and sta-**

tistical (based on Bayes rule) avenues, with references to the books by Kalnay (2003) and Evensen (2009).

2. In section 2.2, the authors further classify the type of users for pygeodyn. As stated above, the software is far from ready for customization so I would suggest the authors rework or remove this section since it would be misleading to claim that the python package is accessible, it is not.

   ⇒ **The end of the section was rewritten to integrate the rework of the documentation. We hope that this will improve the accessibility of the package.**

3. The proper websites of where to download pygeodyn is buried at the very end of the paper. I strongly suggest this be moved in the forefront, maybe at the end of the introduction.

   ⇒ **Actually the recommendation of the referee goes against the journal guidelines asking to put the links for download in the *Code and data availability* section.**

4. increase the font size on the axis and labels for Figure 4

   ⇒ **This figure was generated directly from the webtool as a demonstration and is therefore not easily manipulated. We increased the size of the figure to accommodate this but the customization of these plots is part of future development of the webtool. Note however that the raw data can be exported to do a plot with other plotting softwares.**

Interactive comment on Geosci. Model Dev. Discuss., https://doi.org/10.5194/gmd-2019-82, 2019.