The Journal of Open Source Software

DAPPER: Data Assimilation with Python: a Package for Experimental Research

Patrick N. Raanes [1,2¶], **Yumeng Chen** [3], **and Colin Grudzien** [4,5]

1 NORCE, Bergen, Norway **2** NERSC, Bergen, Norway **3** Department of Meteorology and NCEO, University of Reading, Reading, UK **4** CW3E - Scripps Institution of Oceanography, University of California, San Diego, USA **5** Department of Mathematics and Statistics, University of Nevada, Reno, USA ¶ Corresponding author

DOI: 10.21105/joss.05150

Software

- Review
- Repository
- Archive

Editor: Fei Tao

Reviewers:

- @sara-02
- @Shreyas911

Submitted: 10 January 2023
Published: 29 February 2024

Summary

Data assimilation (DA) is the science of optimally combining sparse data and computational models, both of which are typically large and varied. Based on classical statistics, signal processing, and control systems theory, and increasingly, machine learning, DA was primarily developed in weather, climate, and oceanographic forecasting but is presently used all across the geosciences and beyond. DAPPER is a set of templates for benchmarking the performance of DA methods.

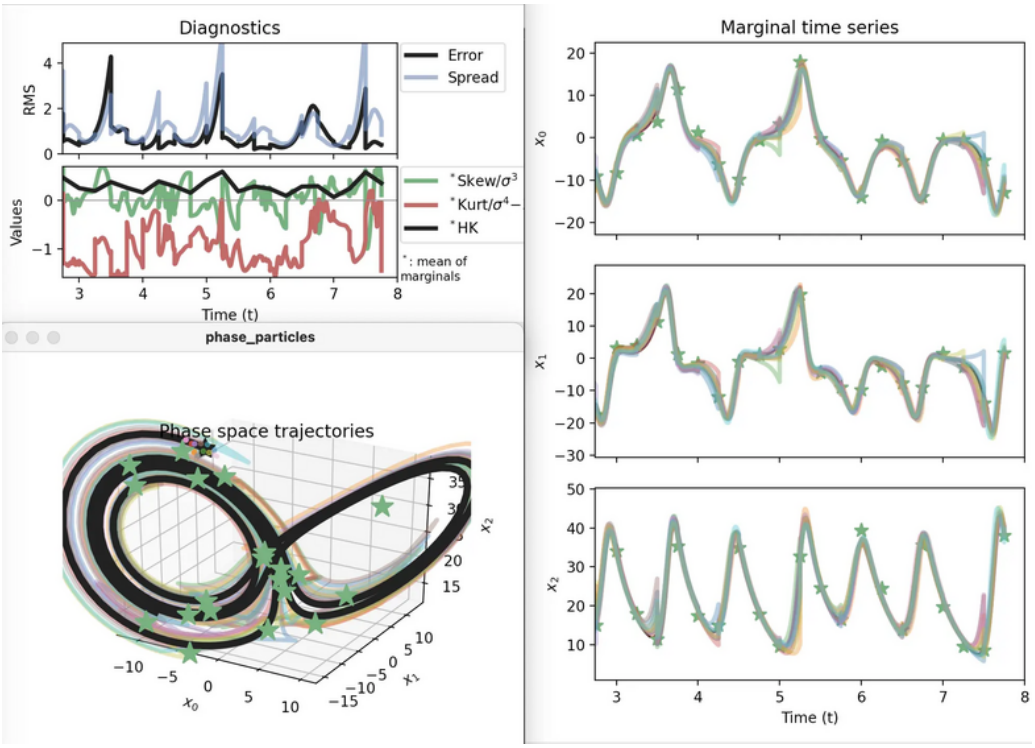

Figure 1: Screenshot from "liveplotting" in DAPPER

Statement of need

DAPPER provides experimental support and guidance for new developments in DA by facilitating numerical investigations through a variety of typical test cases and statistics. It reproduces numerical benchmarks results reported in the literature and facilitates comparative studies, thus promoting the reliability and relevance of the results. DAPPER is open source, written in Python, and focuses on readability; this promotes the reproduction and dissemination of the underlying science, and makes it easy to adapt and extend.

State of the field

The README contains a comprehensive list of related projects. Among projects aimed at research and teaching, DAPPER is probably the most mature, while it is small is size and complexity compared to those targeting real-world applications, such as DART, PDAF, JEDI, and OpenDA. The README also lists 5 publications (to date) that used DAPPER, and those results published in the literature that have been reproduced with DAPPER, of which we mention (Anderson, 2010; Asch et al., 2016; Bocquet & Sakov, 2014).

Outlook

The intention is for DAPPER to continue benchmarking and illustrating the latest methods in DA and beyond. For example, at the time of writing, there are 26 open ticket in the repository's issue tracker. Most have been opened by the principle investigator, and are tagged with *enhancement*. These are not necessarily minor, and have mainly been filed under the v2 milestone.

Acknowledgements

We acknowledge support from Alberto Carrassi and Laurent Bertino of NERSC in the genesis of this project, continued financial backing from NORCE, and contributions from Maxime Tondeur, and Remy Dubois.

References

Anderson, J. L. (2010). A non-Gaussian ensemble filter update for data assimilation. *Monthly Weather Review*, *138*(11), 4186–4198. https://doi.org/10.1175/2010MWR3253.1

Asch, M., Bocquet, M., & Nodet, M. (2016). *Data assimilation: Methods, algorithms, and applications* (p. xvii+295). SIAM. https://doi.org/10.1137/1.9781611974546

Bocquet, M., & Sakov, P. (2014). An iterative ensemble Kalman smoother. *Quarterly Journal of the Royal Meteorological Society*, *140*(682), 1521–1535. https://doi.org/10.1002/qj.2236