

Parcels v4: An Xarray-aligned, flexible lagrangian simulation framework for geoscience

Nick Hodgskin¹, Joe Schoonover², and Erik van Sebille¹

¹ Institute for Marine and Atmospheric Research, Utrecht University, the Netherlands ² Fluid Numerics, Hickory, NC, USA ¶ Corresponding author

DOI: [10.xxxxxx/draft](https://doi.org/10.xxxxxx/draft)

Software

- [Review](#)
- [Repository](#)
- [Archive](#)

Editor: [Open Journals](#)

Reviewers:

- [@openjournals](#)

Submitted: 01 January 1970

Published: unpublished

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](#))

Summary

Parcels ([Delandmeter & van Sebille, 2019](#); [Lange & van Sebille, 2017](#)) is a highly customisable Lagrangian simulation framework. Version 4 of the software is a major update which overhauls the software internals to natively leverage Xarray ([Hoyer & Hamman, 2017](#)) dataset objects. This makes Parcels compatible with many new data formats (e.g., Zarr, Icechunk) and execution modes (e.g., streaming data from cloud buckets or other data providers - such as the [Copernicus Marine Data Store](#)). With this update, Parcels also adds several new features including support for unstructured grid datasets (enabling simulations on (combinations of) different grid geometries), support for custom interpolators (surfacing to scientists even more control over the numerics of their simulation), and trajectory output in Parquet format.

Statement of need

The first release of Parcels was in July 2017. Since then, the geoscience modelling space has changed. A significant portion of the geospatial science community has shifted from creating their own scripts for manipulating NetCDF data, to using open source software - namely Xarray - for working with multidimensional climate data. Major benefits of this software include (a) providing a single in-memory, metadata-rich representation of a full NetCDF-like dataset, (b) its flexibility to manipulate datasets in other data-formats (e.g., Zarr, HDF5), and (c) its abstractions that facilitate easy data ingestion from network data sources.

Writing software that natively works from Xarray datasets provides a powerful abstraction layer allowing downstream developers to create software that works across data formats and data ingestion paradigms. This is particularly important as climate datasets are continually increasing in resolution and size, preventing local file based storage, and datasets are increasingly being stored in modern data formats such as Zarr.

Another interesting change is that climate modellers are increasingly providing model output on different grid geometries, which have attractive features compared to conventional structured grids. Running natively on these grid geometries, which can be represented as unstructured grids, without re-interpolation allows researchers to run simulations that fully capture the details from the original dataset.

Finally, an important point is the cross-domain generalisability of Lagrangian particle tracking tools within the geosciences. Many of these tools originate from a specific domain, and (as evidenced by documentation, API naming, and input data support) cater to a userbase within the same domain. This is despite the fact the numerics of the tools themselves is largely identical between the domains. This results in both fractured communities, and toolsets whose usage mirrors the blurry boundaries of the domains themselves. Having cross-domain tooling for Lagrangian particle tracking promotes cross-collaboration, software quality, building of

41 better standards, and the overall impact of the software in users scientific workflows.

42 State of the field

43 Parcels is not the only open-source Lagrangian particle tracking framework for research. Other
44 notable frameworks for oceanographic applications include OpenDrift (Dagestad et al., 2018),
45 TRACMASS (Döös et al., 2017), Ariane (Blanke & Raynaud, 1997), Drifters.jl, TrackMPD
46 (Jalón-Rojas et al., 2019), Ichthyop (Barrier et al., 2026), and the Connectivity Modelling
47 System (Paris et al., 2013) – each with their own strengths and limitations. OpenDrift is very
48 good for Search-and-Rescue simulations and operational forecasting, TRACMASS and Ariane
49 excel in deterministic (analytical) particle tracking on structured grids, Drifters.jl is useful for
50 advanced diagnostics, TrackMPD is designed specifically for marine plastic applications, and
51 Ichthyop and the Connectivity Modelling System are especially useful for biological applications.
52 Furthermore, many of the hydrodynamic models also have their (online) particle tracking
53 modules, but these are often tightly coupled to the specific model and lack the flexibility
54 offered by standalone frameworks like Parcels. In the atmosphere, FlexPart (Stohl et al., 1998)
55 is a widely used Lagrangian particle tracking framework. However, none of these packages
56 come with the flexibility in terms of structured and unstructured grid support, and custom
57 kernels and interpolators, that Parcels offers.

58 Software design

59 When working with output from a single circulation model, version 4 of Parcels assumes data
60 (i.e., the field data and mesh data) is all contained in a single Xarray Dataset object that
61 has appropriate metadata. This required metadata includes certain CF-convention metadata,
62 providing information about important variables/coordinates and their units, as well SGRID or
63 UGRID metadata, providing grid geometry information for structured grid data and unstructured
64 grid data respectively. This metadata rich object can either be opened directly from disk/a data
65 store, or is constructed by the user from the component input/output files from their particular
66 circulation model. Parcels provides “converter” functions for various circulation models which
67 understand model-specific conventions, that then attach this metadata appropriately. This
68 overall approach allows the internals of Parcels to assume a certain dataset structure and
69 metadata richness, while still allowing the software to transparently facilitate users who come
70 with metadata-poor model data. The diagram in Figure 1 illustrates potential code paths when
71 working from model data to a fully constructed FieldSet. If a user wants to run a simulation
72 with fields from different models, they load each model data into its own FieldSet and then
73 combine the FieldSets together into a single FieldSet simply with an addition.

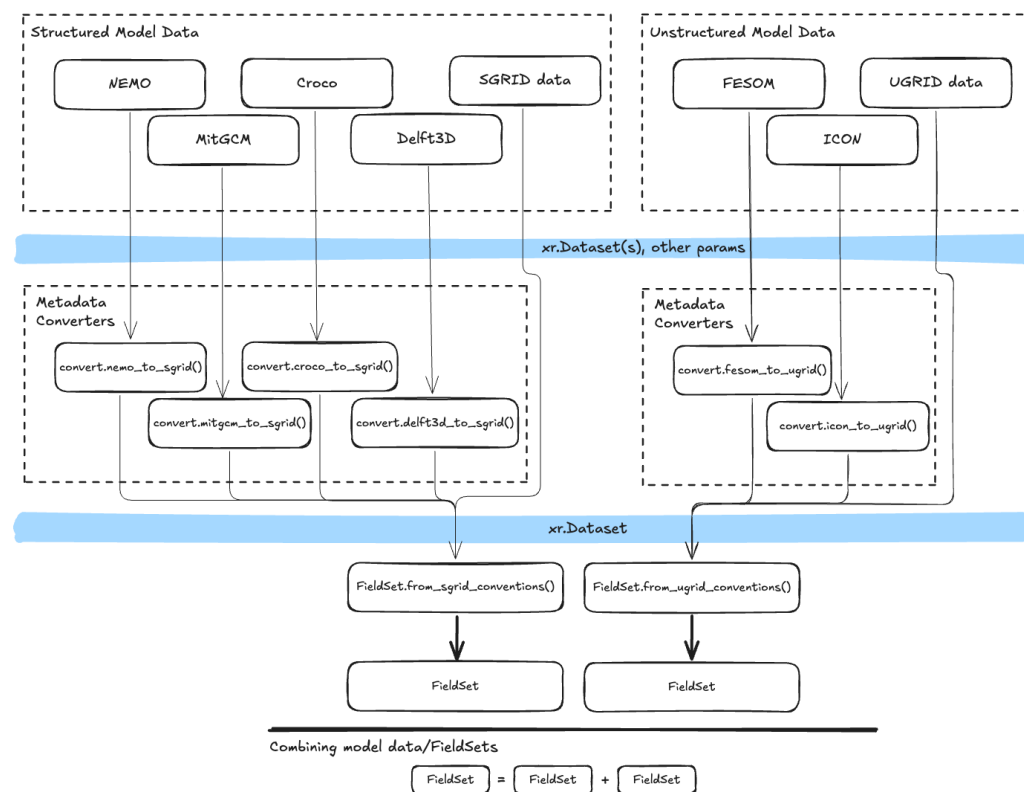


Figure 1: Code path for the new FieldSet ingestion in v4 of Parcels

A strength of earlier Parcels versions has been the ability for users to write custom “kernels” which encode particle actions over the course of a simulation. This flexibility has enabled users to model a wide range of physical phenomena, from plastic pollution to plankton and fish larvae. Parcels version 4 adds custom interpolators, allowing users to also have control on how field data is interpolated at particle positions. Users can use a range of pre-packaged interpolators, or write their own, and set them on a field-by-field basis overriding the default linear interpolators.

Version 4 of Parcels also improves generalisability to other domains, such as atmospheric or cryospheric particle tracking. It allows the encoding of various input datasets, relaxes domain specific requirements, replaces domain specific terminology, and updates our overall branding.

Parcels version 4 also changes the output format from Zarr to Parquet, aligning better with that tabular nature of particle trajectory output.

Example use-case

A key use-case of the Parcels version 4 is the combining of various model data which not only have different sources, but also very different grid geometries. In this section we present an example simulation focused on the Dutch coast.

We combine flow data from Deltares’ 3D DCSM-FM model, SWAN Wave model data from KNMI, and wind model data from Copernicusmarine. The 3D DCSM-FM model is an unstructured model, SWAN has a structured curvilinear grid, and Copernicusmarine provides data on a structured rectilinear grid geometry. We seed particles along the Dutch coast, and within estuaries facing the North Sea. We advect the particles for a total of a month (from 2025-11-01 to 2025-12-01) with a timestep of 10 minutes. Figure 2 shows the varied types of

96 data and grid geometries that we combine, along with the resulting particle tracks.

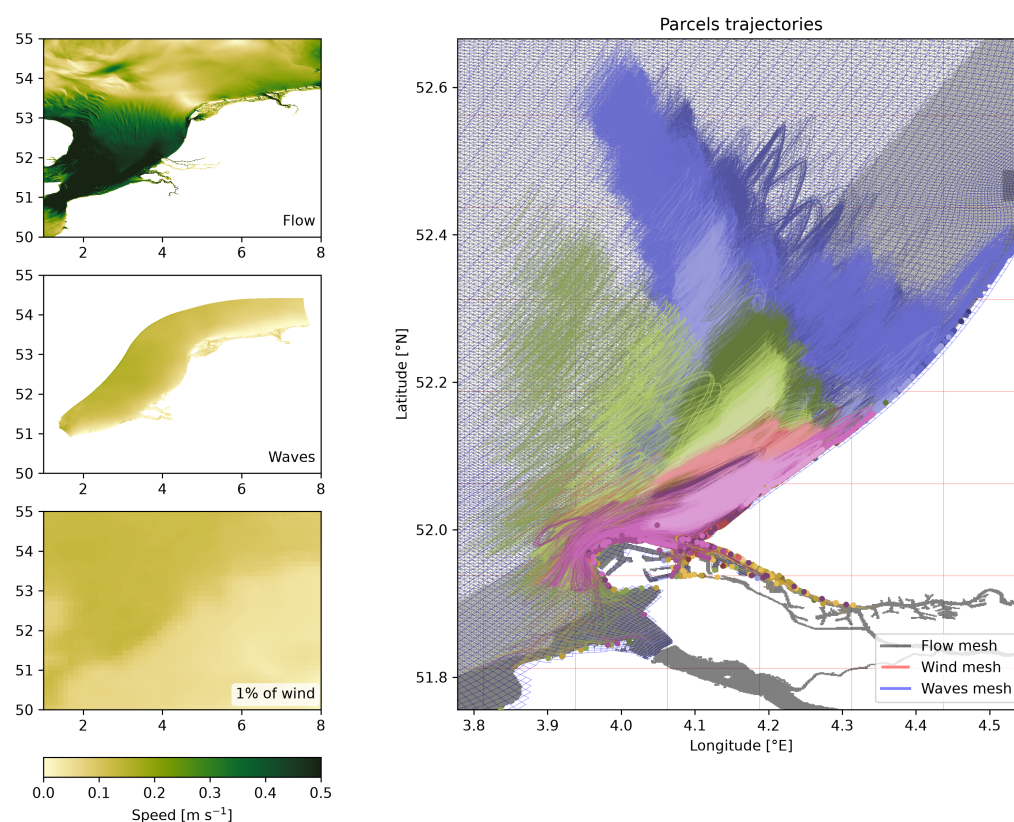


Figure 2: Figure 2: Trajectories of almost 5,000 particles seeded near the port of Rotterdam in The Netherlands. Trajectories are coloured by start date, and final locations of the particles are marked by circles. The mesh of the unstructured flow model (DCSM-FM) is shown in grey, the mesh of the rectilinear Copernicusmarine wind model is shown in red, and the mesh of the curvilinear KNMI wave model is shown in blue.

97 Research impact statement

98 Parcels has been cited in over 330 peer reviewed scientific papers so far mostly within the field
99 of oceanography.

100 AI usage disclosure

101 Large Language Models (Claude Opus 4.6, ChatGPT ...) have been used in a guided manner
102 for code generation, documentation, refactoring, and testing. All LLM written code and
103 documentation has been verified by the authors. This manuscript was fully written by humans.

104 Acknowledgements

105 This work has been funded via Nederlandse Organisatie voor Wetenschappelijk Onderzoek,
106 Exacte en Natuurwetenschappen (VI.C.222.025) as part of the project “Tracing Marine
107 Macroplastics by Unraveling the Ocean’s Multiscale Transport Processes” and via the
108 Warmworld ELPHE project, and compute on the SURF Research Cloud via project
109 EINF-15719.

References

- Barrier, N., Verley, P., Lett, C., & Mullon, C. (2026). *Ichthyop: A Lagrangian tool for modelling ichthyoplankton dynamics*. Zenodo. <https://doi.org/10.5281/ZENODO.4243813>
- Blanke, B., & Raynaud, S. (1997). Kinematics of the Pacific Equatorial Undercurrent: An Eulerian and Lagrangian approach from GCM results. *Journal of Physical Oceanography*, 27(6), 1038–1053.
- Dagestad, K.-F., Röhrs, J., Breivik, Ø., & Ådlandsvik, B. (2018). OpenDrift v1.0: A generic framework for trajectory modelling. *Geoscientific Model Development*, 11(4), 1405–1420. <https://doi.org/10.5194/gmd-11-1405-2018>
- Delandmeter, P., & van Sebille, E. (2019). The Parcels v2.0 Lagrangian framework: new field interpolation schemes. *Geoscientific Model Development*, 12(8), 3571–3584. <https://doi.org/10.5194/gmd-12-3571-2019>
- Döös, K., Jönsson, B., & Kjellsson, J. (2017). Evaluation of oceanic and atmospheric trajectory schemes in the TRACMASS trajectory model v6. 0. *Geoscientific Model Development*, 10, 1733–1749. <https://doi.org/10.5194/gmd-10-1733-2017>
- Hoyer, S., & Hamman, J. (2017). Xarray: N-D labeled arrays and datasets in python. *J. Open Res. Softw.*, 5(1), 10.
- Jalón-Rojas, I., Wang, X. H., & Fredj, E. (2019). A 3D numerical model to Track Marine Plastic Debris (TrackMPD): Sensitivity of microplastic trajectories and fates to particle dynamical properties and physical processes. *Marine Pollution Bulletin*, 141, 256–272. <https://doi.org/10.1016/j.marpolbul.2019.02.052>
- Lange, M., & van Sebille, E. (2017). Parcels v0.9: prototyping a Lagrangian ocean analysis framework for the petascale age. *Geoscientific Model Development*, 10(11), 4175–4186. <https://doi.org/10.5194/gmd-10-4175-2017>
- Paris, C. B., Helgers, J., Sebille, E. van, & Srinivasan, A. (2013). Connectivity Modeling System: A probabilistic modeling tool for the multi-scale tracking of biotic and abiotic variability in the ocean. *Environmental Modelling & Software*, 42(C), 47–54. <https://doi.org/10.1016/j.envsoft.2012.12.006>
- Stohl, A., Hittenberger, M., & Wotawa, G. (1998). Validation of the lagrangian particle dispersion model FLEXPART against large-scale tracer experiment data. *Atmospheric Environment*, 32(24), 4245–4264. [https://doi.org/10.1016/S1352-2310\(98\)00184-8](https://doi.org/10.1016/S1352-2310(98)00184-8)