VirtualShip for simulating oceanographic fieldwork anywhere in the global ocean

Jamie R. C. Atkins[1¶], **Emma Daniels**[1], **Nick Hodgskin**[1], **Aart Stuurman**[1], **Iury Simoes-Sousa**[2], **and van van Sebille**[1]

1 Institute for Marine and Atmospheric Research, Utrecht University, the Netherlands **2** Woods Hole Oceanographic Institution, Falmouth, MA, USA ¶ Corresponding author

DOI: 10.xxxxxx/draft

Software

- Review
- Repository
- Archive

Editor: Open Journals

Reviewers:

- @openjournals

Submitted: 01 January 1970
Published: unpublished

Summary

VirtualShip is a Python-based package which exploits the customisability of the open-source Parcels Lagrangian simulation framework (Delandmeter & van Sebille, 2019; Lange & van Sebille, 2017) to simulate measurements as if they were coming from real-life oceanographic instruments. The software builds a virtual ocean world by streaming data from the Copernicus Marine Data Store on-the-fly, facilitating virtual expeditions anywhere on the globe.

Statement of need

Marine science relies on fieldwork for data collection, yet sea-going opportunities are limited due to financial costs, logistical constraints, and environmental burdens. We present an alternative means, namely VirtualShip, for training scientists to conduct oceanographic fieldwork in an authentic manner, planning future expeditions and deployments, and comparing directly observational strategies with model data.

VirtualShip goes beyond simply extracting grid-cell averaged values from climate model output. Instead, it uses sophisticated interpolation techniques (with Parcels underpinnings) to access data in *exact* locations and timings, as if they were being collected by real-world instruments. VirtualShip shares some functionality with existing ocean model data analysis tools, such as OceanSpy (Almansi et al., 2019), but extends existing capabilities to mesh diverse instrument deployments, with programmable behaviours, into the same expedition simulation. Moreover, VirtualShip exploits readily available, streamable data, via the Copernicus Marine Data Store, removing the need for users to download and manage large datasets locally and/or arrange for access to remote servers.

Functionality

VirtualShip simulates the deployment of virtual instruments commonly used in oceanographic fieldwork, with empahsis on realism in how users plan and execute expeditions. For example, users must consider ship speed and instrument deployment/recovery times to ensure their expedition is feasible within given time constraints. Possible instrument implementations include surface Drifter, CTD (Conductivity-Temperature-Depth), Argo float, XBT (Expendable Bathythermograph), underway ADCP (Acoustic Doppler Current Profiler) and underway Underwater_temperature/salinity probes. More detail on each instrument is available in the documentation.

Figure 1 shows an example expedition around the Agulhas Current and South Eastern Atlantic, deploying a suite of instruments to sample physical and biogeochemical properties. Notable

The Journal of Open Source Software

oceanographic features, such as the strong Agulhas Current and Agulhas Retroflection (drifters retroflecting back into the Indian Ocean), are clearly visible via the underway ADCP measurements and drifter releases in the early waypoints (Figure 1b and c). CTD profiles (Figure 1d and e) also capture the vertical structure of temperature and oxygen across the expedition route, including the warmer surface waters of the Agulhas region (early waypoints) and the Oxygen Minimum Zone in the South Eastern Atlantic (final waypoints).

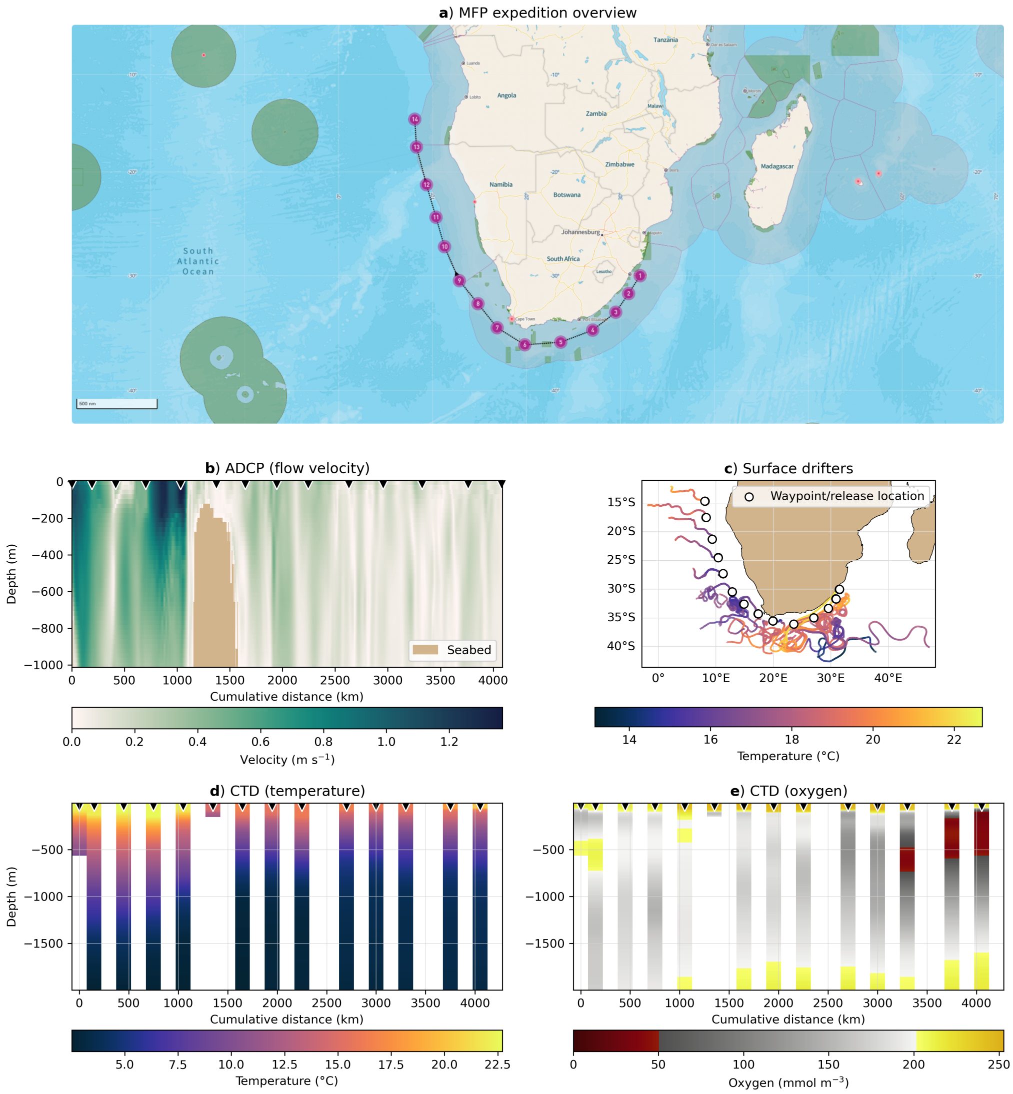

Figure 1: Example VirtualShip expedition simulated in July/August 2023. Expedition waypoints (a), Underway ADCP measurements across the expedition route (b), Surface drifter releases per waypoint (c; 90-day lifetime per drifter), and CTD vertical profiles per waypoint (as a function of cumulative distance across the expedition) for temperature (d) and oxygen (e). Black triangles in b), d) and e) mark locations of waypoints across expedition route, corresponding to the purple markers in a).

The software is designed to be highly accessible to the user. It is wrapped into three high-level command line interface commands (using Click):

1. `virtualship init`: Initialises the expedition directory structure and a `expedition.yaml` configuration file, which controls the expedition route, instrument choices and deployment timings. A common workflow is for users to first define expedition waypoint locations

via the external NIOZ Marine Facilities Planning (MFP) mapping tool. The coordinates can be exported and fed into `init` via the `--from-mfp` CLI flag.

2. `virtualship plan`: Launches a user-friendly Terminal-based expedition planning User Interface (UI), built using `Textual`. This allows users to intuitively modify their expedition waypoint locations, timings and instrument selections.
3. `virtualship run`: Executes the virtual expedition according the planned configuration. This includes streaming data via the Copernicus Marine Data Store, simulating the instrument beahviours and sampling, and saving the resulting data in `Zarr` format.

A full example workflow is outlined in the Quickstart Guide documentation.

Implementation

Under the hood, `VirtualShip` is modular and extensible. The workflows are designed around `Instrument` base classes and instrument-specific subclasses and methods. This means the platform can be easily extended to add new instrument types. Instrument behaviours are encoded as `Parcels` kernels, which allows for extensive customisability. For example, a `Drifter` advects passively with ocean currents, a `CTD` performs vertical profiling in the water column and an `ArgoFloat` cycles between ascent, descent and drift phases, all whilst sampling physical and/or biogeochemical fields at their respective locations and times.

Moreover, the data ingestion system, relies on streaming Analysis-Ready and Cloud-Optimized (ARCO; (Stern et al., 2022), (Abernathey et al., 2021)) data directly from the Copernicus Marine Data Store, via the `copernicusmarine` Python toolbox. This means users can simulate expeditions anywhere in the global ocean without downloading large datasets by default. Leveraging the suite of physics and biogeochemical products available on the Copernicus plaform, expeditions are possible from 1993 to present day and forecasted two weeks into the future. There is also an option for the user to specify local `NetCDF` files for data ingestion, if preferred, with the necessary file stuctures and naming conventions outlined in the relevant documentation.

Applications and future outlook

`VirtualShip` has already been extensvely applied in Master's teaching settings at Utrecht University as part of the "VirtualShip Classroom" initiative. Educational assignments and tutorials have been developed alongside to integrate the tool into coursework, including projects where students design their own research question(s) and execute their fieldwork and analysis using `VirtualShip`. Its application has been shown to be successful, with students reporting increased self-efficacy and knowledge in executing oceanographic fieldwork (Daniels et al., 2025).

We encourage researchers to use `VirtualShip` a cost-effective means to plan future expeditions, as tool for accessing ocean model data in a realistic manner, and to compare models to observations in a like-for-like manner. Both the customisability of the `VirtualShip` platform and the exciting potential for new ARCO-based data hosting services in domains beyond oceanography (e.g., atmospheric science) means there is potential to extend VirtualShip (or "VirtualShip-like" tools) to other domains in the future.

As the `Parcels` underpinnings in themselves continue to evolve, with a future (at time of writing) v4.0 release focussed on alignment with Pangeo standards and `xarray` data structures, `VirtualShip` will also benefit from these improvements, further enhancing its capabilities, extensibility and compatability with modern cloud-based data pipelines.

Acknowledgements

The VirtualShip project is funded through the Utrecht University-NIOZ (Royal Netherlands Institute for Sea Research) collaboration.

References

Abernathey, R. P., Augspurger, T., Banihirwe, A., Blackmon-Luca, C. C., Crone, T. J., Gentemann, C. L., Hamman, J. J., Henderson, N., Lepore, C., McCaie, T. A., Robinson, N. H., & Signell, R. P. (2021). Cloud-native repositories for big scientific data. *Computing in Science & Engineering*, *23*(2), 26–35. https://doi.org/10.1109/MCSE.2021.3059437

Almansi, M., Gelderloos, R., Haine, T. W. n., Saberi, A., & Siddiqui, A. H. (2019). OceanSpy: A python package to facilitate ocean model data analysis and visualization. *Journal of Open Source Software*, *4*(39), 1506. https://doi.org/10.21105/joss.01506

Daniels, E., Chytas, C., & Sebille, E. van. (2025). The virtual ship classroom: Developing virtual fieldwork as an authentic learning environment for physical oceanography. *Current: The Journal of Marine Education*. https://doi.org/10.5334/cjme.121

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

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

Stern, C., Abernathey, R., Hamman, J., Wegener, R., Lepore, C., Harkins, S., & Merose, A. (2022). Pangeo forge: Crowdsourcing analysis-ready, cloud optimized data production. *Frontiers in Climate*, *Volume 3 - 2021*. https://doi.org/10.3389/fclim.2021.782909