Plan Overview

A Data Management Plan created using DMPonline

Title: Vici: Tracing Marine Macroplastics by Unraveling the Ocean’s Multiscale Transport Processes

Creator: Erik van Sebille

Principal Investigator: Erik van Sebille

Data Manager: Erik van Sebille

Affiliation: Utrecht University

Funder: Netherlands Organisation for Scientific Research (NWO)

Template: Data Management Plan NWO (September 2020)

ORCID iD: 0000-0003-2041-0704

Project abstract:

Ocean currents transport plastic around the globe. However, plastic concentrations vary by orders of magnitude within mere kilometers. It is unclear why the distribution of oceanic plastic is so heterogeneous. The aim of this project is to unravel the oceanographic transport, fragmentation and beaching processes underlying this spatiotemporal heterogeneity of macroplastic particles and to predict the accumulation of macroplastic on the sea surface and on coastlines of Northwest Europe.

Macroplastics (>5cm) have significantly different properties than the more studied microplastics. This project will be the first to integrate all relevant physical oceanic processes that act on macroplastics – from small-scale surface waves to large-scale currents – into one mechanistic framework, which will be offered to the community as open-source toolbox.

Recent computational advances will allow my team and me to simulate these processes on scales from centimeters to thousands of kilometers, and for the first time also include the fragmentation of macroplastics due to biological, physical, and chemical weathering.

We will validate these simulations by tracking novel open-hardware drifters that represent floating macroplastics. By combining simulations and observations, we will create a tool to predict plastic hotspot emergence at sub-kilometer resolution.

These innovations will allow us to identify under which conditions the different processes are important for the transport, fragmentation, and beaching of macroplastic items; and use that information to predict when and where macroplastic accumulates to optimize clean-up strategies.

The plastic polluting our ocean is atrocious, and this project will provide practitioners with a tool of unprecedented accuracy to guide cleanup-efforts. At the same time, tracing the floating plastic is a unique opportunity to improve our fundamental understanding of ocean transport processes. This Vici project thus helps to combat an environmental problem while advancing scientific knowledge.

ID: 118639

Start date: 01-09-2023

End date: 31-08-2028

Last modified: 02-06-2023

Grant number / URL: VI.C.222.025

Vici: Tracing Marine Macroplastics by Unraveling the Ocean’s Multiscale Transport Processes

General Information

Name applicant and project number

Erik van Sebille
VI.C.222.025

Name of data management support staff consulted during the preparation of this plan and date of consultation.

Lena Thöle
RDM Suppor

1. What data will be collected or produced, and what existing data will be re-used?

1.1 Will you re-use existing data for this research?

If yes: explain which existing data you will re-use and under which terms of use.

- Yes

Hydrodynamic data will be re-used from the Copernicus Marine Service (https://marine.copernicus.eu/).
The Copernicus Marine Service provides free, open, regular and systematic reference information on the blue (physical), white (sea ice), and green (biogeochemical) ocean state, variability and dynamics across the global ocean and European regional seas.

1.2 If new data will be produced: describe the data you expect your research will generate and the format and volumes to be collected or produced.

Two types of data will be produced: from observations and from simulations
1. Observational data will be produced in WP3 (Mapping the coast-ocean-coast pathway), from the drifter release experiments. The drifters that we build will generate time series of their position after they are released. This data will be saved in json files (as we have done before in e.g. https://galapagosplasticfree.nl/drifters/). We expect this to be max 1MB per drifter, so approx 1GB for the full dataset
2. Simulation data will be produced in all work packages. These are time series of trajectory data of virtual macroplastic particles. These are stored in zarr format (and can be converted to netcdf for longer-term storage) and are typically 10GB per simulation. Since we expect 100s of simulations

throughout the course of the project, the total amount of data generated will be a few TB

1.3. How much data storage will your project require in total?

- >1000 GB

See answer to 1.2 above

2. What metadata and documentation will accompany the data?

2.1 Indicate what documentation will accompany the data.

The json files from the observational campaign will be accompanied by a peer-reviewed article detailing the experiment design and execution
All simulation data will be accompanied by a GitHub repository with the full code that was used to run the simulation; as well as described in the methods section of peer-reviewed articles

2.2 Indicate which metadata will be provided to help others identify and discover the data.

The observational drifter data will be featured on an interactive public website (following the framework we used in https://galapagosplasticfree.nl/drifters/), where the json file can also be downloaded
All simulation data will be linked in the peer-reviewed articles in which the simulations are described. Furthermore, all these simulation data will also be accompanied by easy readme files, that provide a short description of how to use the data.

3. How will data and metadata be stored and backed up during the research?

3.1 Describe where the data and metadata will be stored and backed up during the project.

- Institution networked research storage

All data will be stored on Yoda (https://www.uu.nl/en/research/yoda)

3.2 How will data security and protection of sensitive data be taken care of during the research?

- Not applicable (no sensitive data)

There is no sensitive data

4. How will you handle issues regarding the processing of personal information and intellectual property rights and ownership?

4.1 Will you process and/or store personal data during your project?

If yes, how will compliance with legislation and (institutional) regulation on personal data be ensured?

- No

4.2 How will ownership of the data and intellectual property rights to the data be managed?

All data will be made available via CreateCommons and/or MIT licenses. The owner will be UU

5. How and when will data be shared and preserved for the long term?

5.1 How will data be selected for long-term preservation?

- All data resulting from the project will be preserved for at least 10 years

5.2 Are there any (legal, IP, privacy related, security related) reasons to restrict access to the data once made publicly available, to limit which data will be made publicly available, or to not make part of the data publicly available?

If yes, please explain.

- No

5.3 What data will be made available for re-use?

- All data resulting from the project will be made available

5.4 When will the data be available for re-use, and for how long will the data be available?

- Data available as soon as article is published

5.5 In which repository will the data be archived and made available for re-use, and under which license?

Data will be preserved via the Yoda platform (https://www.uu.nl/en/research/yoda)

5.6 Describe your strategy for publishing the analysis software that will be generated in this project.

All software that we write will be published open at http://github.com/oceanParcels, following the RDM at https://github.com/OceanParcels/UtrechtTeam/wiki/Research-Data-Management-Plan-for-OceanParcels

6. Data management costs

6.1 What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?

The project includes budget for a 0.3 FTE dedicated software engineering will support the team with the practicalities of code development, including installation on the SURFsara Snellius High Performance Computing cluster, proactive Research Data Management, open-source licensing, and diligent version control of the code