Multi-scale modeling of solar wind expansion, acceleration, and turbulence

A Data Management Plan created using DMPonline.be

Creator: Fabio Bacchini

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Fabio Bacchini

Data Manager: Fabio Bacchini

Project Administrator: Fabio Bacchini

Grant number / URL: G020224N

ID: 203657

Start date: 01-01-2024

End date: 31-12-2027

Project abstract:

The solar wind (SW), a continuous stream of charged particles released from the Sun, fills the interplanetary space with "collisionless" plasma where charged particles interact only weakly. Recently launched space missions provide measurements for the SW plasma (e.g. density and temperature) from distances unprecedentedly close to the solar surface. In particular, Parker Solar Probe (PSP) and Solar Orbiter (SolO) are now reporting intriguing results that are poorly understood by current theoretical models, mainly concerning plasma energization. For example, it is unclear why and how the SW accelerates from the solar surface to the high speeds measured at large distances; how large temperature anisotropy (with respect to the local magnetic field) affects plasma heating; and how ions and electrons are energized differentially. It is also conjectured that turbulence and expansion, ubiquitously occurring in the SW, can alter particle acceleration and overall plasma dynamics as the SW propagates outward from the Sun. In this proposal, we will address several of the open questions in the physics of SW plasmas with first-principles simulations. We will model expansion, acceleration, and turbulence in the collisionless SW to correctly capture the interplay of these and other mechanisms (e.g. instabilities) that determine the plasma energization. Our results will be fundamental to explain and guide current and future observations from high-profile missions such as PSP and SolO.

Last modified: 18-12-2023

Created using DMPonline.be. Last modified 18 December 2023

1 of 6

Multi-scale modeling of solar wind expansion, acceleration, and turbulence
Application DMP

Questionnaire

Describe the datatypes (surveys, sequences, manuscripts, objects …) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

Simulation files and codes: text files (txt, f90, cpp), data files (h5, vtu, vtk). These will be new data from simulation results, containing raw numbers to analyze physical quantities such as magnetic fields. The estimated total size will not exceed a few TBs. None of the data is confidential.

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

1. Designation of responsible person (If already designated, please fill in his/her name.)
2. Storage capacity/repository
   - during the research
   - after the research

Fabio Bacchini will be responsible for the storage and management of data. All data will be stored on local hard drives, the VSC clusters, and the KUL's RDR, ensuring data is stored unlimitedly.

What's the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

N/A.

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

N/A.

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

N/A.

Multi-scale modeling of solar wind expansion, acceleration, and turbulence
FWO DMP (Flemish Standard DMP)

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)
Codes	Numerical codes to conduct simulations	Generate new data	Digital	Software	.txt, .f90, .cpp	<100MB
Simulations	Files containing results from simulations	Generate new data	Digital	Simulation data	.h5, .vtu, .vtk	<50 TB

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

N/A.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.

- No

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

README.txt files will be created for each code and for each simulation dataset.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

3. Data storage & back-up during the research project

Where will the data be stored?

Local hard drives, VSC facilities, KUL RDR

How will the data be backed up?

Multiple copies will be created on each storage device/server.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely.
If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.

- Yes

The indicated facilities have more than enough storage for unlimited time.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The facilities indicated are secured by ssh restricted access and not accessible by external, unauthorized people, unless specifically allowed by the PIs.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

N/A.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

All data will be retained for five years. After that, they may be deleted or kept indefinitely depending on the need.

Where will these data be archived (stored and curated for the long-term)?

Software is archived in online repositories for unlimited duration. Simulation data may be kept on physical hard drives if needed.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

N/A.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in a restricted access repository (after approval, institutional access only, …)

Codes will be on online repositories, accessible upon request. Data will be stored on the KUL RDR in the long term, and accessible upon request.

If access is restricted, please specify who will be able to access the data and under what conditions.

The PIs will have access to the data. External members (hirees, collaborators) may be given access if needed for the project and/or upon reasonable request.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- No

Where will the data be made available? If already known, please provide a repository per dataset or data type.

GitHub for codes (to be created).
KU Leuven RDR for data.

When will the data be made available?

Immediately and while it is being produced.

Which data usage licenses are you going to provide? If none, please explain why.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- No

What are the expected costs for data sharing? How will these costs be covered?

N/A.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

Fabio Bacchini (supervisor-spokesperson)

Who will manage data storage and backup during the research project?

Fabio Bacchini (supervisor-spokesperson)

Who will manage data preservation and sharing?

Fabio Bacchini (supervisor-spokesperson)

Who will update and implement this DMP?

Fabio Bacchini (supervisor-spokesperson)