DMP title

Project Name FWO DMP for SB PhD fellowship - DMP title

Grant Title 1S24022N

Principal Investigator / Researcher Pietro Parodi

Description An emerging concept called Air-Breathing Electric Propulsion (ABEP) could allow to fly spacecraft in the currently unexploited Very Low Earth Orbits, using air collected from the atmosphere as propellant for an electric thruster. A small-scale hypersonic low density facility sponsored by ESA is under development at the von Karman Institute for Fluid Dynamics (VKI) for the experimental testing of an ABEP intake-collector system. The stream of fast particles required in the facility to replicate the orbital flow will be accomplished through electrostatic acceleration of plasma. The objective of this numerical project is to provide a predictive model of the ABEP intake in the ground testing configuration. We will develop and use a hybrid PIC-DSMC code to simulate the collisional, nonequilibrium plasma flow involved. Numerical results will be produced during our project. The purpose is to analyze this data by comparing it to existing literature and with experiments in the VKI facility, in order to have a validation of our models and methods.

Institution KU Leuven

1. General Information

Name applicant

Pietro Parodi

FWO Project Number & Title

1S24022N – Advancing modeling and particle methods for plasma simulation applied to Air-Breathing Electric Propulsion

Affiliation

- KU Leuven

2. Data description

Will you generate/collect new data and/or make use of existing data?

- Generate new data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

The project will produce mainly three types of data: source code, simulation results, and post-processed data also in the form of elaborated manuscripts.

1. Source code will consist of the Fortran code base for the PANTERA PIC/DSMC software that we will develop. Numerous miscellaneous scripts for preliminary analysis and data post-processing will be written, mainly in the Python language. Estimated volume: 10 GB
2. Results from simulations using particle methods will include averaged flow field data and particle data dumps. We will make use of the open-source VTK format (described at https://kitware.github.io/vtk-examples/site/VTKFileFormats/) as much as possible for ease of storage and exchange of these data. Large volumes may be temporarily produced, such as from test or failed simulation runs, but we estimate the final volume of data worthy of being saved to be around 500 GB
3. Post-processed data will include various data formats such as VTK, MS Excel, MS Word, and LaTeX. Estimated volume: 100 GB

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- No

Privacy Registry Reference:
Short description of the kind of personal data that will be used:

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s)

- No

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- No

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- No

4. Documentation and metadata

What documentation will be provided to enable reuse of the data collected/generated in this project?

1. Source code will be adequately commented. It will be constantly synchronized with a GitLab repository, which allows for the creation of a Wiki page to descrive and document the software. We plan to write Wiki pages for the parts of the most complex parts of the code that may be not clear to future users. Python pre- and post-processing scripts will be adequately commented.
2. Simulation data will be orgainized in folders with meaningful names. Results will be accompanied by the input script of the simulation, where the numerical problem is setup. The output to terminal they produced and a brief description of the case will be included.
3. Post-processed data will be in common formats and self-describing.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- No

5. Data storage and backup during the FWO project

Where will the data be stored?

The simulation data will be stored on our personal laptop, which has a 1 TB hard drive. In addition, it will be periodically selected and stored on two additional and separate portable hard drives also with a capacity of 1 TB. We also 2 TB of OneDrive cloud storage provided by KU Leuven, which we will use as storage for smaller volume data. Source code for the PIC/DSMC code will be subjected to version control on a GitLab repository.

How is backup of the data provided?

The backup will be guaranteed by the three independent copies of the data (personal laptop + 2 external HDDs). New data will be added to the hard drives at intervals of about two weeks. Additionally, the two portable drives will be kept physically separated at all times: one at the office, one at home, to prevent contemporaneous damage/loss of these devices. Data of lower volume will also be periodically copied on the OneDrive cloud storage provided by KU Leuven. The GitLab repository provides another backup for the code.

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

We have 1 TB of capacity for backed-up data, which is less than the total data volume foreseen.

What are the expected costs for data storage and back up during the project? How

will these costs be covered?

None: we have already the hard drives at our disposal, and currently the storage capacity is sufficient. The 2TB OneDrive cloud storage is provided by KU Leuven free of charge for all employees and students.

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

We will not be working with sensitive data. Nonetheless, the data on the personal laptop is password-protected, and physical access is limited for the portable hard drives.

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

All the data listed at point 2, after deletion of duplicate or useless files, will be stored for the mandated 5 years period after the end of the project.

Where will the data be archived (= stored for the longer term)?

After deleting useless data, it will be stored on the same hard drives used during the project. The code will be additionally stored on the GitLab repository.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

Should one of the hard drives fail, a new one (~100€) will be purchased using the bench fee at our disposal.

7. Data sharing and reuse

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- No

Which data will be made available after the end of the project?

We plan to release the PIC/DSMC code as open source software. Therefore, the software will be licensed and access to the relative GitLab repository, accompanied by the Wiki pages, will be opened to the public. Simulation data of particular interest can be made available upon request by email after the related research results are published.

Where/how will the data be made available for reuse?

- In an Open Access repository
- Upon request by mail

PIC/DSMC code on the GitLab repository. Simulation result files upon request through the OneDrive storage provided by KU Leuven.

When will the data be made available?

- Immediately after the end of the project
- Upon publication of the research results

The GitLab repository will be opened to the public once the code is sufficiently stable and documented, which is in any case planned before the end of the project. Simulation data can be shared with requesting researchers as soon as the relative research results are published.

Who will be able to access the data and under what conditions?

Anyone with a GitLab account will have access to our code under the conditions set by the open source software license (GNU GPL v3). Simulation data will be made available upon request by email after the research results are published, under the condition that it is properly referenced if it is used in publications by other authors.

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

No cost is foreseen: the GitLab repository is free, and large simulation results files can be shared using the OneDrive storage provided by KU Leuven.

8. Responsibilities

Who will be responsible for data documentation & metadata?

Pietro Parodi

Who will be responsible for data storage & back up during the project?

Pietro Parodi

Who will be responsible for ensuring data preservation and reuse ?

Pietro Parodi

Who bears the end responsibility for updating & implementing this DMP?

The PI bears the end responsibility of updating & implementing this DMP.