Scientific exploitation of the KU Leuven Mercator telescope: addressing major shortcomings in our understanding of stellar structure and evolution of single and binary stars

A Data Management Plan created using DMPonline.be

Creator: Clio Gielen

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Grant number / URL: G0ABL24N

ID: 206161

End date: 31-12-2027

Project abstract:

Intermediate and high-mass stars are the chemical factories of our Universe. Their chemical and energetic feedback drives the evolution of galaxies while their end-of-life products serves as probed to study the distant universe or give rise to gravitational waves that can be detected by current and future facilities. Yet, our understanding of their stellar structure and evolution suffers from major shortcomings and yet the simplest part of their evolution, their main sequence, remain improperly modelled. These uncertainties grow larger once stars evolve and when effect of binary and mass transfer are taken into account.

The present project will leverage large observing programs at the Mercator telescope to address two important issues related to the stellar structure and evolution of single and binary stars, while using binaries as privilege laboratories: (i) the discrepant values of stellar mass estimates between different techniques, and (ii) the search for a large population of main sequence binary interaction products that can be used to calibrate the end product of stable mass transfer.

Last modified: 08-05-2024

Scientific exploitation of the KU Leuven Mercator telescope: addressing major shortcomings in our understanding of stellar structure and evolution of single and binary stars

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.

				Only for digital data	Only for digital data	Only for digital data	Only for physical data
Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)	Physical volume
		Please choose from the following options: • Generate new data • Reuse existing data	*Please choose from the following options:* • Digital • Physical	*Please choose from the following options:* • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA	*Please choose from the following options:* • .por, .xml, .tab, .csv,.pdf, .txt, .rtf, .dwg, .gml, … • NA	*Please choose from the following options:* • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA	
WP1-2	TESS spectroscopy	Reuse	Digital	Observational	.fits	<10TB	
	Mercator/Hermes spectroscopy	Reuse/new	Digital	Observational	.fits	<10TB	
WP1.1	Automatisation	Generate	Digital	Software	python scripts	<1GB	
WP1.2	Analytics	Generate	Digital	Software	python scripts	<1GB	
WP1.3	Code integration	Generate	Digital	Software	python scripts	<1GB	
	Atmospheric properties	Generate	Digital	Simulation data	.txt, .fits	<100GB	
	GSSP	Reuse	Digital	Software	software package	<1GB	
	TLUSTY	Reuse	Digital	Software	software package	<1GB	
WP1.4	Phoebe	Reuse	Digital	Software	software package	<1GB	
	Analytics	Generate	Digital	Software	python scripts	<1GB	
	Stellar properties	Generate	Digital	Simulation data	.txt, .fits	<1GB	
WP2.1	MESA	Reuse	Digital	Software	software package	<10GB	
	Evolutionary models	Generate	Digital	Simulation data	.txt, .fits	<100GB	
WP2.2	Analytics	Generate	Digital	Simulation data	.txt	<1GB	
WP2.3	Statistical analysis	Generate	Digital	Simulation data	.txt	<1GB	

WP2.4	Novel methodology	Generate	Digital	Software & numerical methods	python code, numerical methods	<1GB	
WP3.1	Statistical analysis	Generate	Digital	Simulation data	.txt	<1GB	
WP3.2	Bayesian-based bias- correction analysis	Generate	Digital	Simulation data	.txt	<1GB	
WP3.3	Post-interaction analysis	Generate	Digintal	Simulation data	.txt, .fits	<10GB	
WP3.4	Statistical analysis	Generate	Digital	Simulation data	.txt	<1GB	

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:

Software codes used:

- MESA: docs.mesastar.org
- GSSP: fys.kuleuven.be/ster/meetings/binary-2015/gssp-software-package
- TLUSTY: tlusty.oca.eu
- Phoebe: phoebe-project.org

Observational data used:

- TESS archival photometry: https://archive.stsci.edu/missions-and-data/tess
- Mercator HERMES archival spectroscopy: https://www.royer.se/melchiors.html

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.

- Yes

Software codes and archival observations are free to use but need to be acknowledged and/or cited in publications following the instructions of the original authors/source.

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).

Combining a human-readable ascii format (like a .csv/txt file) with appropriate column names ensures simple reuse of model output data. All newly-developed source code materials for simulations and numerical models will be described in detail in open-access journal publications.
All these computational tools will be open-source om GitHub with accompanying readme.txt files with detailed description of the input and output parameters. Software code will include readme files on Github and will be commentated inline in the code (ReadMe files on BitBucket and/or Github and documented inline within the code using, e.g. docstrings in Python code.).
Journal publications will use dedicated journal standards (e.g. keywords) to identify research context.

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.

- Yes

Numerical models and simulations will follow the standard for astronomical data and metadata, the most commonly-used being FITS. A description of astronomical metadata standards can be found the Strasbourg Astronomical Data Center https://cds.unistra.fr/data-publication/

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

Where will the data be stored?

All material and codes in process will be stored on the shared internal netword of the Institute of Astronomy of KU Leuven.
All final new source code will be stored and shared on Github.
All fill numerical models and metadata will be stored long-term on public data repositories of the CDS

How will the data be backed up?

Continuous data preservation (beyond the 10-year requirement) is possible thanks to the well-developed in-house storage and backup system of the Institute of Astronomy of KU Leuven.

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

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

The Institute of Astronomy of KU Leuven has an automatic periodic backup (daily and off-site) on the network server of raw and processed data according to KU Leuven and in-house security standards: all data are secured with access restrictions on file-system level.

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

Data storage needs are estimated at ~20TB,including backups. This storage is currently available in the in-house server and storage system of the host institute and thus comes at no additional cost.

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...).

Continuous data preservation (even beyond the 10-year requirement) is possible thanks to the well-developed in-house storage and backup system of the Institute of Astronomy and the continuous data preservation of the CDS.

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

All data will be stored long-term on the backup system (onsite and offline) of the Institute of Astronomy and on the servers of the CDS.

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

The costs for back-up and storage are included in the general working budget of the Institute of Astronomy.

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 an Open Access repository

All computational tools will be open-source on Github for the benefit of the scientific community at large, in particular for those astrophysical projects that rely on modern numerical modelling.
Furthermore, all relevant tables, databases etc. given in the publications will be given in the open-acces publications and if applicable downloadable from the Strasbourg Astronomical Data Center (CDS). The CDS is a data center dedicated to the collection and worldwide distribution of astronomical data and related information (https://cds.unistra.fr/about/cds/).
All journal publications will be open access via the ArXiv.org and Lirias servers.

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

NA

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.

All computational tools will be open-source on Github (GitHub.com)
The output of all models will be available in open-access publications and the CDS (https://cds.unistra.fr/data-publication/).
All journal publications will be open access via the ArXiv.org and KU Leuven Lirias servers.

When will the data be made available?

Upon publication of the research results.

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

Data: CC-BY 4.0
Software/code: GNU GPL-3.0

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.

- Yes

Journal publications will receive a DOI. Large data sets that are published on the CDS data center will also receive a dedicated DOI, complementary to the article DOI.

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

Publication costs are budgeted in the project. There are no associated costs with making the data available on the CDS servers.

6. Responsibilities

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

The PI and team members will be responsible for the garanteeing data quality, documentation, storage and backup, in collaboration with the IT team of the Institute of Astronomy of KU Leuven.

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

The PI and team members will be responsible for the garanteeing data quality, documentation, storage and backup, in collaboration with the IT team of the Institute of Astronomy of KU Leuven.

Who will manage data preservation and sharing?

The PI and team members will responsible for uploading the data to the relevant back-up and online sites. The IT team of the Institute of Astronomy of KU Leuven will be responsible for the continued back-up after the project.

Who will update and implement this DMP?

The PI, with assistance from the institute's project coordinator, bears the end responsibility of updating & implementing this DMP.