Hydrogeological feedback mechanisms during juvenile dune development

A Data Management Plan created using DMPonline.be

Creator: Jadon Beerlandt

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Grant number / **URL:** 1SG3423N

ID: 199226

Start date: 01-11-2022

End date: 31-10-2026

Project abstract:

Today, the threat of rising sea levels is not limited to the beach and dunes; we must also resist the intrusion of saltwater under the sandy coastline. Traditionally, freshwater lenses beneath our existing dunes have protected us from the infiltration of saltwater. However, our coastline and climate are changing, resulting in a decline in the extent of freshwater lenses. Understanding the hydrogeological feedback mechanisms during dune development is crucial for predicting the development of freshwater lenses, particularly as coastal management increasingly focuses on natural solutions for coastal protection, such as dune-for-dike approaches. These feedback mechanisms strongly depend on the interactions between groundwater fluctuations, tidal action, infiltration, and dune development.

Despite decades of research, we still cannot reliably predict the growth of freshwater lenses due to a lack of process knowledge and the inherent complexity of dynamic systems. During this project, we aim to monitor and model the development of a freshwater lens using modern techniques in a large and unique artificial dune area in Raversijde, Belgium. In doing so, we aim to reintroduce the dune to the coast as it was traditionally and work towards an integrated coastal model that not only protects us against flooding but also protects our hinterland and economy against salinization and contamination of our groundwater resources.

Last modified: 28-04-2023

Hydrogeological feedback mechanisms during juvenile dune development
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
ERT Profiles	Geophysical ERT profiles that are going to be taken periodically during the project.	☒ Generate new data ☐ Reuse existing data	☒ Digital	☒ Observational ☒ Experimental	☒ .py	☒ < 1 TB	/
EM Profiles	Geophysical EM profiles that are going to be taken periodically during the project.	☒ Generate new data	☒ Digital	☒ Observational ☒ Experimental	☒ .py	☒ < 100 GB	/
GPR Profiles	Geophysical GPR profiles that are going to be taken periodically during the project.	☒ Generate new data	☒ Digital	☒ Observational ☒ Experimental	☒ .py	☒ < 100 GB	/
Monitoring well dataset	Data from the clustered monitoring wells placed on the different cross section	☒ Generate new data	☒ Digital	☒ Observational ☒ Experimental	☒ .csv	☒ < 100 GB	/
Drone profiling	Monthly drone profiling that we receive from MDK	☒ Reuse existing data	☒ Digital	☒ Observational ☒ Experimental	☒ .tiff	☒ < 1 TB	/
RTK GPS dataset	RTK GPS data that we take at certain time to compare against drone data or want a newer image.	☒ Generate new data	☒ Digital	☒ Observational ☒ Experimental	☒ .csv	☒ < 100 GB	/
Simulation data SEAWAT	Simulation results/scrips and steps generated using SEAWAT	☒ Generate new data	☒ Digital	☒ Simulation data	☒ .SWT	☒ < 100 GB	/
Python scripts	Scripts that are written in Python to process geophysical data and inversions.	☒ Generate new data	☒ Digital	☒ Software	☒ .py	☒ < 1 GB	/
MATLAB scripts	Scripts that are written in Python to process geophysical data and inversions.	☒ Generate new data	☒ Digital	☒ Software	☒ .mat	☒ < 1 GB	/
Pictures	Pictures that are taken during the project to visualize changes in the dune area.	☒ Generate new data	☒ Digital	☒ Observational	☒ .png	☒ < 100 GB	/

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:

We are able to use the monthly available drone data from MDK that are generated from the test site.
This is governmental data that we are able to use for our research.

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).

To manage and document my different types of data, I follow the following approach:

- For geophysical and hydrogeological data, I ensure that each dataset is well-organized and labelled with a clear name. I define the variables in a codebook and include metadata such as the location, date, and time of the data collection. I also document any processing or filtering steps applied to the data.
- For MATLAB and Python scripts, I store them in a version control system such as Git, along with documentation describing their purpose and any necessary input/output data. I make sure that the scripts are well-commented and follow good programming practices.
- Pictures are stored in a directory with a clear and consistent naming convention. I add metadata to the pictures, such as the date, location, and any relevant details about the picture. Additionally, if pictures are used in analysis, I document which pictures were used and how they were analysed.
- For simulation data, I make sure that it is clearly labelled and organized. I create a codebook to describe the variables in the dataset. I also document any scripts or code used to run the simulation, including details such as the simulation parameters and any assumptions made.
- For drone profiling data, I label it with information such as the location, date, and time of the data collection. I create a codebook to describe the variables in the 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

If no, please specify (where appropriate per dataset or data type) which metadata will be created:

- We will use the convention that is also used during the project and organization of the living lab. For now, this is still in discussion.
- At this moment, we accompany all available metadata (location, time, date, boundary factors) that are available with every dataset.

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

Where will the data be stored?

- Internal storage work PC
- Local mass network storage at our campus
- KU Leuven OneDrive
- Private NAS at my home

See next section for more details.

How will the data be backed up?

To ensure that I do not lose any important files, I have set up a backup system using multiple storage locations.

· Firstly, I will use the internal storage location on my PC to store my files, as it is readily accessible and does not require an internet connection. However, I am aware that if my PC fails or experiences a hardware failure, I may lose all of my data stored on the internal storage location.

· Secondly, I will utilize a local mass drive storage at our campus to store my files (Large Volume Storage of the KU Leuven ICTS). This will provide an additional layer of security and ensure that my data is still accessible even if my PC fails. This storage option can also be accessed from anywhere on the campus network, making it easy to share files with colleagues.

· Thirdly, I will back up my data to our university's Microsoft OneDrive, which provides secure cloud storage. This ensures that my files are stored remotely on servers managed by Microsoft, and that my data is accessible from any device with an internet connection. In the event of a PC failure, my files will still be backed up and accessible. Daily back-up of newly generated data is guaranteed with the KU Leuven One Drive service, while the Large Volume Storage of the KU Leuven ICTS secures daily backup of the large files/datasets directly uploaded there.

· Finally, I will use a local Network Attached Storage (NAS) device at my home to back up my data. This will provide an additional layer of security for my data and ensure that my files are still accessible even if all other backup options fail. The NAS can be accessed remotely, making it easy to retrieve files even when I am away from home.

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 OneDrive has a limitation in storage capacity (2TB) which will be sufficient for the small datasets, scrips and reports generated in this project. The research group provides sufficient data storage for all larger data files in the Large Volume Storage of the KU Leuven ICTS.

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

Only the PhD student, Jadon Beerlandt, and the research group/living lab will have access to the data.
To prevent unauthorized access or modification of my data, I will implement appropriate security measures. For example, I will use a strong password or passcode to protect my PC and NAS device from unauthorized access. I will also ensure that the storage device at my campus is only accessible to authorized personnel and that my OneDrive account has a strong password and two-factor authentication.
Our internal KU Leuven accounts also have two-factor authentication.
Data that needs to be transferred will be done using our universities protected sharing client (Belnet).
I will regularly update my security measures and review my backup system to ensure that my data remains secure.

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

No additional costs. Costs for large volume on ICTS servers are covered by our research group. KU Leuven OneDrive is without charge.

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 (the data mentioned in beforementioned table) will be preserved for at least 5 years after ending of the project on:

- The large volume of our research group ICTS data storage.
- The data storage of the living lab

These will also be split in different version of the data (periods in time or stages in PhD)

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

- The large volume of our research group ICTS data storage.
- The data storage of the living lab

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

Our research group will cover the cost of the Large Volume Storage of the KU Leuven ICTS. There are no additional costs for storing on the storage of the living lab project.

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, …)

Hydrogeological and geophysical datasets will be made available through the repository of the living lab (still under decision) and the repository of the KU Leuven.
Python and MATLAB scripts will be uploaded on GitHub, which is a standard hosting service to share scripts in the scientific community.

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

/

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.

Hydrogeological and geophysical datasets will be made available through the repository of the living lab and the repository of the KU Leuven RDR.
Python and MATLAB scripts will be uploaded on GitHub.

When will the data be made available?

Once the related research papers are accepted for publication, the data will be made available.

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

The dataset will be licensed under the licensing agreements of the living lab (not decided yet).

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

Once the related research papers are accepted for publication, the data DOI will be provided.

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

No additional costs are expected.
Uploading and sharing on GitHub is free of share.

6. Responsibilities

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

Jadon Beerlandt, the PhD candidate, will manage the data documentation and metadata during the research project.

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

Jadon Beerlandt, the PhD candidate, will manage data storage and backup during the research project.

Who will manage data preservation and sharing?

Jadon Beerlandt, the PhD candidate, will manage data preservation and sharing.

Who will update and implement this DMP?

Jadon Beerlandt, the PhD candidate, will update and implement this DMP.