Data Management Plan (DMP)

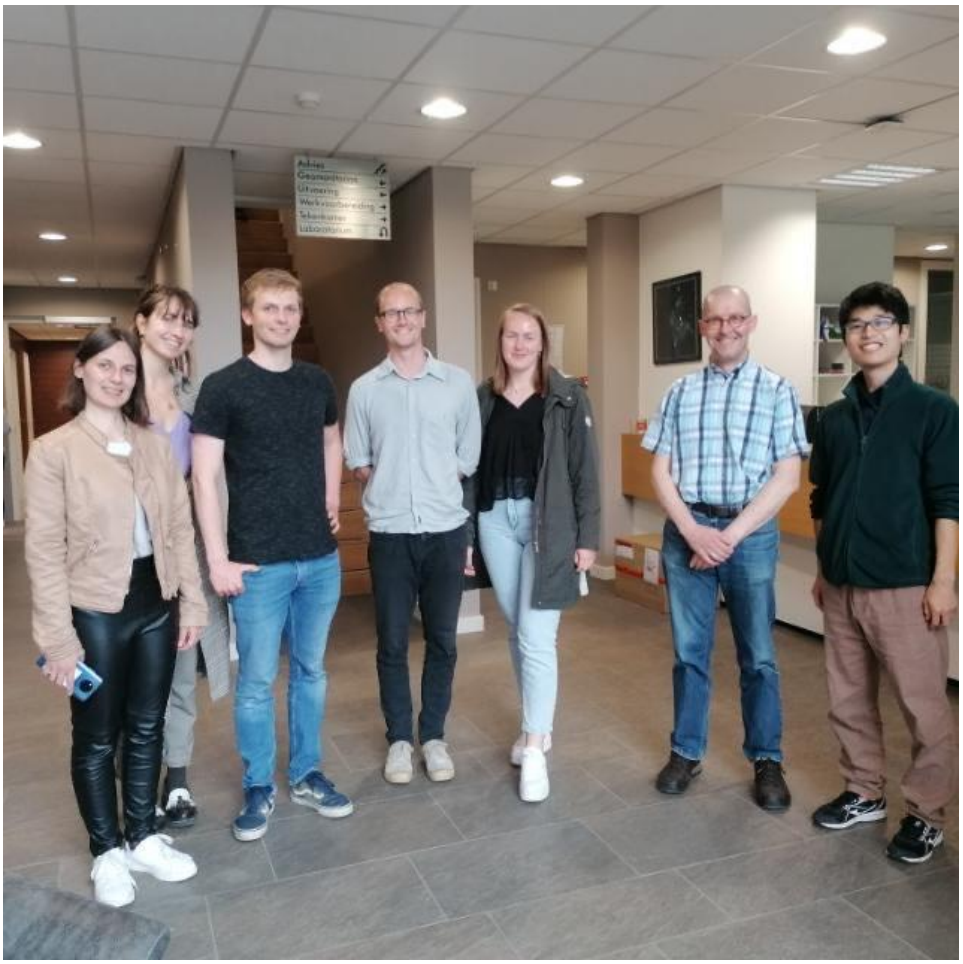
Remote sensing and GIS integration, Academic consultancy training

Group 6

Davey de Groot, Shoyo Nakamura, Kristie Swinkels,

Marieke Buuts, Persa Koutsouradi

28th June 2022



WAGENINGEN UR

For quality of life

Table of Contents

Organizational context

Name	Group 6 – DORA – for explorative research *Davey de Groot, Kristie Swinkels, Shoyo Nakamura, Persa Koutsouradi, Marieke Buuts*
Date	28-06-2022
Chair group	Geo-Information Science and Remote Sensing (GRS)
Commissioner/ GRS-supervisor	Wiertsma and Partners (Nico van der Moot, Kees-Jan van der Made, Johan Bergsma) Coach and expart in the GRS group (Nandika Tsendbazar, Magdalena Smigaj)
Start date of project	08-05-2022
File name of this DMP	Data_Management_Plan_ACT_Group6.pdf

Short description of your project

Title	Skimming the surface: Analysis of vegetation stress in Lauwersmeer using Remote Sensing techniques.
Summary	Wiertsema and Partners is a consultancy company specialized in monitoring of the subsurface. To enhance their monitoring skills, they are interested in the use of remote sensing as a tool to understand connections between below- and above-ground systems. One of their current projects is on monitoring the effects of salinization on vegetation stress in the Lauwersmeer area. Even though Wiertsema and Partners have little experience with remote sensing methods and data, they are interested in in exploring how remote sensing can help detect salinization stress in the Lauwersmeer area as monitoring below ground salinization is costly. Therefore, this project will serve as explorative research into the possibilities of how remote sensing can be useful for detecting belowground salinization effects by examining historical and present-day vegetation stress in the Lauwersmeer area. The objectives of the project are: 1) monitoring the trend in vegetation stress in the National Park Lauwersmeer from 1985 until 2021, 2) identifying highly stressed locations in the agricultural areas surrounding the Lauwersmeer, 3) evaluating the relationship between elevation and vegetation stress, and 4) making everything reproducible and easy to use for people in Wiertsema and Partners who have limited backgrounds on remote sensing.

	In order to meet the objectives, trends of the Normalized Difference Vegetation Index (NDVI), which is often used as a proxy for vegetation stress, will be derived from historical satellite data for both nature and agricultural areas. The potential monitoring locations will be detected by identifying the intersection of low elevation and high stressed areas. In addition, correlation analysis will be conducted with the Algemeen Hoogtebestand Nederland and NDVI trend data. The deliverables of the project will include maps of stressed locations, graphs of NDVI trends, and an interactive application including a step-by-step manual to achieve reproducibility.

Data management roles

Roles	
Who is **collecting** the data?	Ourselves (Davey de Groot, Persa Koutsouradi, Shoyo Nakamura, Kristie Swinkels, Marieke Buuts) Wiertsema and Partners
Who is **analyzing** the data?	Ourselves (Davey de Groot, Persa Koutsouradi, Shoyo Nakamura, Kristie Swinkels, Marieke Buuts)
Other (Is there a person in the project group with a specific responsibility for data management? Do other persons contribute, for example by writing code?)	Dr. Magdalena Smigaj, teacher and researcher at Wageningen University and Research Andrei Mîrt, PhD candidate at Wageningen University and Research
What is the role of your **supervisor?**	Supervision of project, provision of data, no analysis/storage

Expected type of project data, software choices, data size

Data stage	Specification of type of project data	Software choice	Data size
Source data	Landsat satellite imagery (public)	Google Earth Engine (GEE)	See below*
	Algemeen Hoogtebestand Nederland (AHN)	QGIS	35 MB
	Area of Interest nature area	GEE/QGIS (.shp)	400 KB
Result data	LandTrendr application for time series analysis using Landsat NDVI data from 1985 until 2021	GEE	See below**
	NDVI trend application (with Theil Sen regression) showing the slope of the NDVI change over time from 1985 until 2021	GEE	See below**
	Overlay map of stressed locations and elevation (AHN)	QGIS	350 KB
	Map showing suggested locations of new monitoring wells	QGIS	350 KB
Models/code	LandTrendr_APP	GEE (.js)	60.4 KB
	Library	GEE (.js)	66.2 KB
	UI	GEE (.js)	19.5 KB
	Theil Sen Regression APP	GEE (.js)	6.66 KB

** All the satellite data are collected, processed and stored on Google Earth Engine. Therefore, we do not know the exact data size and it does not matter.*
*** The applications are stored on the cloud, so we do not know the exact data size and it does not matter. Instead, the script file sizes matter and you can see them in the table above.*

Short term storage solutions

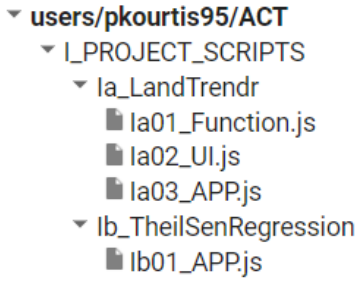
Describe where the data will be stored physically and how the back-up is organized.

Data stage	**Storage location**	**Backup procedures** (storage medium and location/ how often?)
Source data	Google Earth Engine	Satellite images will be imported in the code on Google Earth Engine and will be saved online when necessary.
	Github	The area of interest files will be on Github. We will store them once at the end of the project.
Result data	Google Earth Engine	The finalized scripts will be stored online on Google Earth Engine and the results will also be shown in an online application. The scripts and accordingly the application will be stored on Google Earth Engine when necessary.
	Microsoft Teams	Intermediate results and reports will be stored on Microsoft Teams (saved when needed).
	Github	All final products, including links to applications, manual, report, license and scripts, will be stored in the Github repository. We will store them once at the end of the project.
Models/code	Google Earth Engine	Google Earth Engine stores scripts and changing history online. We will save our scripts regularly to avoid losing our changes.
	Github	Copies of the code are stored on the Github repository. We will store them once at the end of the project.
Meeting minutes	Microsoft Teams	Minutes are stored on Microsoft Teams and are saved automatically.
Report	Microsoft Teams	The reports including previous versions are backed-up on Microsoft Teams and they are saved automatically.
	Github	Final report will be stored on the Github repository once at the end of the project.
Manual	Microsoft Teams	Manual (and previous versions) are backed-up on Microsoft Teams (saved automatically).
	Github	The final manual will be stored in the Github repository (stored once at the end of the project).

Structure of the data and information

Give a visual representation of the system for the directory- and filenames you intend to use. See the examples for inspiration.

We will make use of 2 data storage solutions:
For managing the satellite data, Google Earth Engine itself will be used. There code will be stored in a shared project in Google Earth Engine. The system will look something like this:



Besides, the component data will also be stored in Google Earth Engine under the same names as the codes described above.

For storing output data, OneDrive will be used, as OneDrive allows all members to have direct access to the same data.

Does your workflow provide for version control? If not, describe how you intend to keep versions apart.

For the LandTrendr application, there are three JavaScripts which are all linked with each other. Google Earth Engine has a built-in revision history function which are used for version control. Furthermore, adjustments in the scripts are made in a cloned version and only copied to the original version when working properly.

For the NDVI trend application with the Theil Sen regression, no adjustments have to be made so no version control is necessary.

Documentation and metadata

Describe how you are going to document your data collection process, what the resulting data files comprise and how they will be processed further. Think about documenting the:

1. content (what does your dataset contain?)

2. context (who, what, why, where and how will the data be collected and analyzed)

3. process (are there specific processes and does it make sense to organize notes by the process?)

1. Satellite images, the Algemeen Hoogtebestand Nederland (AHN), shapefile of the Area Of Interest (AOI).

2. All analysis will be performed by ourselves on Google Earth Engine as the satellite images are easily imported there. A shapefile of the AOI is provided by Wiertsema and Partners uploaded to Google Earth Engine. The application output can be visually assessed and downloaded. An overlay with the AHN map and identification of stressed locations in surrounding agricultural areas were made in QGIS. Correlation analysis between NDVI slope trends and elevation was performed in Excel.

3. Clear and descriptive notes will be added to every step in the scripts. Furthermore, a manual containing the workflow of how to use the application will be kept making the application understandable and easy to use for Wiertsema and Partners.
Besides, all project outputs will be shared via a Github repository including a README file which gives an overview of what every file contains and where the data and applications can be found.

Sharing and ownership*

Sharing and ownership	**(With) who(m), what and how?**
Data **sharing** - Do you expect that others may be interested in re-using your data? Do you have plans to share your data with these parties? - How are you going to make sure your data files will be accessible once you finish this project? Who will take care of your data?	-Yes, Wiertsema and Partners (commissioner) will be re-using the data via the provided applications and scripts. All outputs of the project, including the scripts and a link to the applications, will be shared in a Github repository. In addition, a Google account with access to Google Earth Engine was created for Wiertsema and Partners in order to modify the scripts. Furthermore, an elaborate manual on how to use the applications was provided. Note: There are some limitations regarding commercial use of Earth Engine. All terms and conditions can be found on https://earthengine.google.com/terms/. - We created our own group folders on Microsoft Teams, Google Earth Engine and Github. We will keep the folders and their contents, such that people can have access to our data even after the course. We all will take care of the data as we have our own agreement that we share all the data on Teams, Google Earth Engine, and Github and take all the member's consent when sharing them outside.

Data **ownership** - Any funder's requirements to share your data, or to impose an embargo? - Are there agreements on how the data will be used and shared within your group or with other parties involved in this project? (outside your group or outside Wageningen UR)	-A link to the Github repository will be shared. All the outputs of the project can be found and downloaded there. Furthermore, the scripts will be shared in a repository created on the provided Google Earth Engine account. -All scripts and data will be shared under the Apache License version 2.0. All the data produced during this project are owned by this group and we have our agreement shown above.
Privacy - Are there privacy or security issues, and if there are, how are you dealing with them?	There are no privacy or security issues involved. All data used (Landsat satellite images and AHN) is public data.

Long term storage

Which part of your project data has value for long term storage? Do you intend to preserve these data for the long term?

Yes or no?	Argumentation
Yes	The applications will be used by Wiertsema and Partners. Furthermore, the scripts will be provided and can be modified by them. Therefore, the scripts, the application and all documents describing and explaining them should be stored for a long term. Downloading and uploading data and/or maps from or to Google Earth Engine happens on Google Drive. In case Wiertsema and Partners would like to use the applications and scripts extensively, data should be downloaded from Google Drive and stored on a local drive on a regular basis as Google Drive has limited storage capacity.

Which data archive do you intend to use?

All project outputs will be shared on a Github repository. Links to the applications can also be found there. All scripts will also be stored online in a Google Earth Engine repository for which Wiertsema and Partners will have access with the provided Google and Earth Engine account.

Agreement between project group and WU-supervisor.

We, the undersigned, agree upon the following points:

1) The project **group 6** provides a Data Management Plan (DMP) as part of the Project Proposal. If no Project Proposal has to be written (anymore), the project group provides a DMP by **28th June**, as agreed with the WU-supervisor.

2) The project group provides core data sets to the daily supervisor (or commissioner) at the end of the project or earlier as agreed upon with the WU-supervisor (or commissioner).

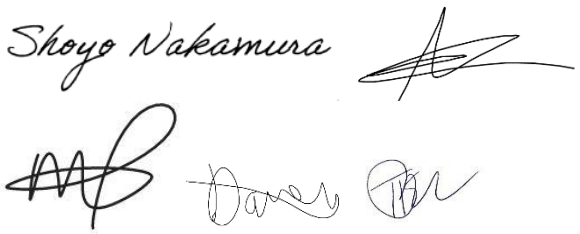
3) All data, scripts, products and results are owned by Wageningen University and custodianship is with Wageningen University. The project group has the right to use them.

As agreed, upon,

Names project group: 6 Date: 28th of June

1. Davey de Groot
2. Kristie Swinkels
3. Shoyo Nakamura
4. Persa Koutsouradi
5. Marieke Buuts

Signatures:



Name(s) WU-supervisor/Commissioner: Date:

Signature(s):