

Project	MONOCLE H2020 (grant 776480)	Start / Duration	1 February 2018/ 48 Months
Dissemination	PUBLIC	Nature	OTHER
Date	20/11/2018	Version	1.1



Multiscale Observation Networks for Optical monitoring of Coastal waters, Lakes and Estuaries

Deliverable 9.3

Open Data Repositories

Project Description

Funded by EU H2020 MONOCLE creates sustainable *in situ* observation solutions for Earth Observation (EO) of optical water quality in inland and transitional waters. MONOCLE develops essential research and technology to lower the cost of acquisition, maintenance, and regular deployment of *in situ* sensors related to optical water quality. The MONOCLE sensor system includes handheld devices, smartphone applications, and piloted and autonomous drones, as well as automated observation systems for e.g. buoys and shipborne operation. The sensors are networked to establish interactive links between operational Earth Observation (EO) and essential environmental monitoring in inland and transitional water bodies, which are particularly vulnerable to environmental change.



Project	MONOCLE H2020 (grant 776480)	Start / Duration	1 February 2018/ 48 Months
Dissemination	PUBLIC	Nature	OTHER
Date	20/11/2018	Version	1.1

Deliverable Contributors:	Name	Organisation	Role / Title
Deliverable Leader	Oliver Clements	PML	WP5 developer
Contributing Author(s)	Stefan Simis	PML	Coordinator
	Jess Heard	PML	WP9 lead
	Kathrin Poser	Water Insight	WP5 lead
	Caitlin Riddick	U Stirling	Contributor
Reviewer(s)			
Final review and approval	Stefan Simis	PML	

Document History:

Release	Date	Reason for Change	Status	Distribution
1.0	27/6/2018	Initial document, reviewed	Completed	Public
1.1	20/11/2018	Digital Object Identifier added	Completed	Public

To cite this document:

Clements, O., Simis, S.G.H., Heard, J., Poser, K., Riddick, C. (2018). D9.3 Open Data Repositories. *Deliverable report 9.3 of project H2020 MONOCLE (grant 776480)*, doi: 10.5281/zenodo.1492132

© The authors



Project	MONOCLE H2020 (grant 776480)	Start / Duration	1 February 2018/ 48 Months
Dissemination	PUBLIC	Nature	OTHER
Date	20/11/2018	Version	1.1

TABLE OF CONTENTS

1. Executive Summary	4
2. Scope	4
3. Introduction	4
Data types	4
Repositories	5
4. Repository access	5
Website	5
Project management portal	5
Zenodo	6
OpenAire	6
5. Pre-existing Data Repositories	6
WISPCloud	6
LIMNADES	7
PML Thredds server	7
PML GISportal	7



Project	MONOCLE H2020 (grant 776480)	Start / Duration	1 February 2018/ 48 Months
Dissemination	PUBLIC	Nature	OTHER
Date	20/11/2018	Version	1.1

1. Executive Summary

MONOCLE is part of the H2020 Open Data Research pilot and committed to making in situ data from inland and coastal waterbodies available and accessible for satellite calibration and validation. The project further aims to fill current gaps in sensor technology and deployment protocols, designed to reduce the cost of operational satellite observation of coastal and inland water quality. Furthermore, MONOCLE develops low-cost sensor alternatives for measurement that significantly reduce the uncertainty of satellite observations, such as measurements describing the vertical distribution of matter in the water column. These technological improvements are accompanied with training materials for installation, maintenance and in some cases instructions for self-building and co-creation.

The materials and data created in the project are predominantly stored in open-access repositories. This document provides descriptions and access instructions for these locations. Regularly updated information on how MONOCLE data will be managed is provided in the Data Management Plan.

2. Scope

This document is intended for internal and external use. The instructions herein do not cover usage of externally managed data portals, but may point to externally provided documentation instead. For regularly updated information on how MONOCLE data will be managed, please refer to the Data Management Plan.

3. Introduction

Data types

Five data categories are identified for MONOCLE:

1. Reports, documents, protocols and scientific publications
2. Raw (uncalibrated or quality controlled) measurement data without metadata
3. Raw (uncalibrated or quality controlled) measurement data with metadata
4. Calibrated and/or quality controlled measurement data with metadata
5. 'Dummy' data produced solely to test interoperability mechanisms and data triggers
6. Simulated data for other than testing purposes

Each of the categories 2-6 can take the format of single data values of different (defined) types, array data obtained at a single point in space and time, image data, or image frames in a sequence (video).

Only categories 1, 3, 4 and 6 will under normal circumstances be considered of long-term value to external users. The open data repositories may store data from other categories but do not need to be set up for this purpose.

Project	MONOCLE H2020 (grant 776480)	Start / Duration	1 February 2018/ 48 Months
Dissemination	PUBLIC	Nature	OTHER
Date	20/11/2018	Version	1.1

Repositories

Thus far, seven repositories have been identified to host MONOCLE data. These are:

- Project website (data category 1)
- Project management portal (category 1)
- Zenodo (all categories)
- OpenAire (publications in category 1)
- WispWeb (categories 3, 4)
- LIMNADES (category 4)
- PML Thredds server (categories 2-6)

Access to these sites is described in the next section.

4. Repository access

Website

The project website is registered at <http://www.monocle-h2020.eu> and has no access restrictions. Documents that are exclusively hosted on the website do not receive a digital object identifier (DOI). The website can link to documents hosted on other open repositories providing DOI registration. The advantage of hosting documents on the website is that stories can be added. Website documents will be grouped under the 'Materials' tab located in the top level menu (Figure 1).

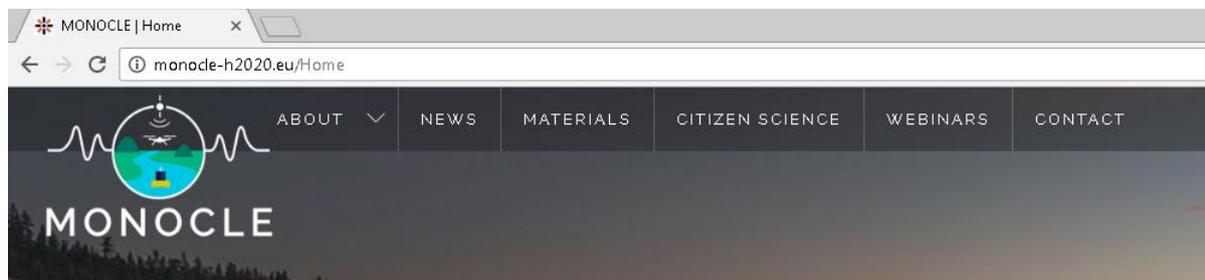


Figure 1 Project website top level menu

Project management portal

The project management portal is a project-internal document repository held on a secure server at PML, only accessible to registered users. The portal is described in project deliverable D1.2 (confidential). The exception to private access is for documents that are explicitly made openly accessible, but this use is not foreseen nor encouraged except for e.g. event management, linking to announcements made on the project website.

The project management portal may be used to temporarily share data between project partners under category 5, particularly while dedicated data sharing services are still being developed.



Project	MONOCLE H2020 (grant 776480)	Start / Duration	1 February 2018/ 48 Months
Dissemination	PUBLIC	Nature	OTHER
Date	20/11/2018	Version	1.1

Zenodo

Zenodo is the preferred document repository for public dissemination of category 1 data and discrete data sets ('frozen data') in categories 4 and 6. Data sets on Zenodo can be registered with a DOI and are therefore traceable in future publications and project impact reports.

A Zenodo 'community' repository has been set up at:

https://zenodo.org/communities/monocle_h2020/

To go directly to the community data upload page use:

https://zenodo.org/deposit/new?c=monocle_h2020

The above address will automatically ensure that those who use it have their record added to the MONOCLE community collection.

Finally, an OAI-PMH feed is available to be used by other digital repositories to harvest this community:

https://zenodo.org/oai2d?verb=ListRecords&set=user-monocle_h2020&metadataPrefix=oai_dc

OpenAire

OpenAire is hosted on Zenodo and automatically lists public documents that are registered in the H2020 participant portal (once they have been accepted), to allow free access.

OpenAire is available at <https://www.openaire.eu>

The direct link to MONOCLE on OpenAire is

www.openaire.eu/search/project?projectId=corda_h2020::c8657863c298e99e1843c21f1c884102

5. Pre-existing Data Repositories

WISPcloud

WISPcloud is the online database system of Water Insight for spectral measurements. The core of WISPcloud is a database that holds raw and processed spectral measurements as well as metadata on the measurements (such as date and time, location, weather conditions) and instruments (such as calibrations). Data are made available via a REST API. While WISPcloud is not an open access system, fine grained access control on the measurement level allows to make data available to the public as well as to share them within a project consortium. Currently, the database contains only measurements by WISPstations, but it is designed to be flexible enough to include spectral data from other instruments as well.



Project	MONOCLE H2020 (grant 776480)	Start / Duration	1 February 2018/ 48 Months
Dissemination	PUBLIC	Nature	OTHER
Date	20/11/2018	Version	1.1

LIMNADES

LIMNADES is a database which comprises global lake and inland water bio-optical measurements and matchup data for remote sensing (www.limnades.org). It was developed by and is held in trust and maintained by the UK GloboLakes project (www.globolakes.ac.uk). This database presently comprises in situ data from over 250 lakes, including coincident Inherent Optical Properties (IOPs), radiometric quantities and biogeochemical data.

Account registration for the LIMNADES database can be done through the website (<https://www.limnades.org/register.psp>). Once registered, data must be first submitted to the repository in order to view and download any publicly available data. A request must be made in order to access restricted datasets.

PML Thredds server

Plymouth Marine Laboratory (PML) hosts an operational cluster of data servers providing standardised access through the Open Geospatial Consortium Web Coverage Service (WCS) and Web Map Service (WMS). The system holds an archive of existing remote sensed earth observation data. The services offered provide open access to the data (WCS) and to generated imagery (WMS). WCS provides a set of operations that allow users to interrogate the service and retrieve metadata about a particular dataset or coverage. Using this metadata, e.g. name of dimensions, a user is able to specify the spatial and temporal subset required. The subsets are used in a “GetCoverage” request that returns the actual data file matching the boundaries given by the user. The format of the returned data can be specified from a list of available formats. Where WCS provides access to the raw data WMS provides access to imagery that represents the data. This can be very useful for quick assessment of data trends or specific localised features. WMS provides a set of operations that allow the user to find out what time steps are available for time-series data.

The services can be accessed through a web browser, <https://vortices.npm.ac.uk>, or integrated into existing web interfaces that follow the published standards. One such example is the GISportal provided by PML.

PML GISportal

The Geographic Information Service (GIS) portal, <https://monocle.eofrom.space>, provides a web based platform to discover, visualise and analyse data. Using the OGC services discussed above it is possible to quickly visualise and review data.



Project	MONOCLE H2020 (grant 776480)	Start / Duration	1 February 2018/ 48 Months
Dissemination	PUBLIC	Nature	OTHER
Date	20/11/2018	Version	1.1

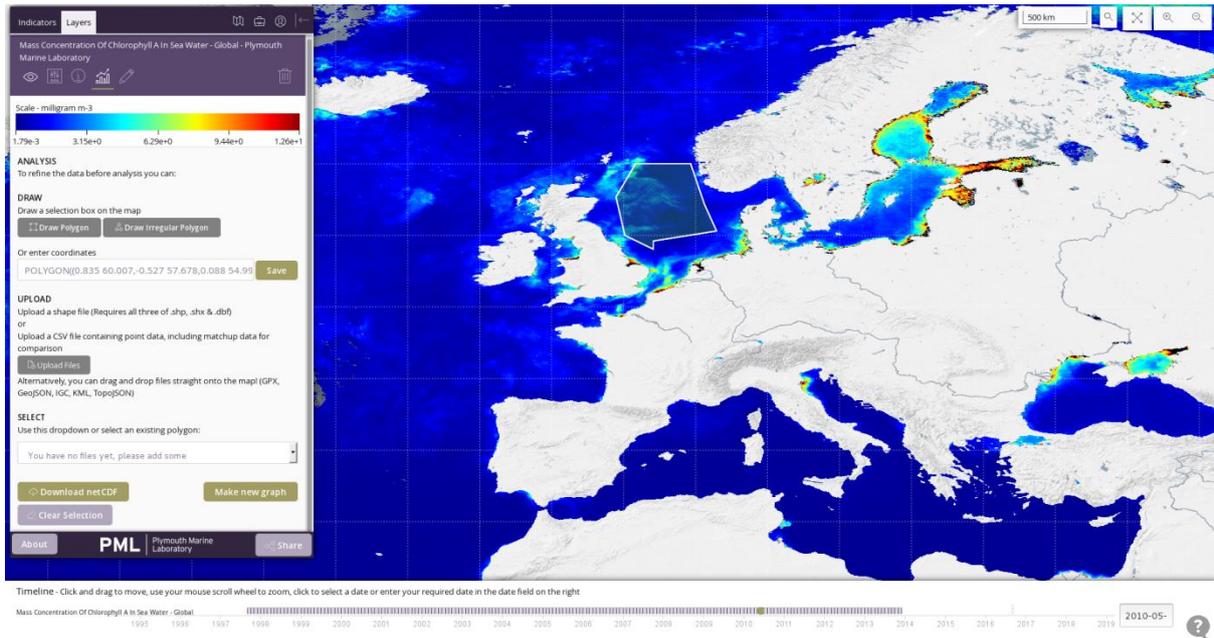


Figure 2 Screenshot showing GISportal being used for data analysis

