



Helmholtz-Zentrum für Ozeanforschung Kiel

**Guidelines for creating a README file
for publishing data on GEOMAR OPeNDAP Server**
<http://data.geomar.de>

Contact

Data Management

Address: GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel
Wischhofstr. 1-3
24148 Kiel | Germany

Phone: 0431 / 600 4025

E-Mail: datamanagement@geomar.de

All information about a dataset should be provided within a readme file to ensure that the data can be correctly interpreted by others when sharing or publishing data or even by yourself at a later date. A readme file must be submitted along with the publication of dataset file(s) on the GEOMAR OPeNDAP Server (Thredds)/ <http://data.geomar.de>.

Procedure to submit datasets for publication on GEOMAR OPeNDAP Server:

1. Fill in the data submission [form](#). By this, the data management team will get all necessary information to submit your data to the server.
2. **Create a readme file** for each dataset:
 - Name the file **README**; not read_me, readme etc.
 - Write your readme document as a plain text file and save it as README.txt or README.md.

The following information must be included within Readme files along submitted datasets.

RECOMMENDED CONTENT FOR README FILES

1. Introduction

- Dataset Title
- Authors contact information
- For each file or group of similar files, a short description of what data it contains
- Explanation of file naming convention and description of file relations (if there multiple files that relate to each other) or description of the file structure

2. Information about methodology

- Description of methodology used for generating the dataset and for processing data (if there are other than raw data contibuted)
- Description of the Software beeing used (including version number) to produce, render, analyse and /or needed to read the dataset
- Information about used standards and calibration
- Description of quality-assurance procedures performed on the data

3. Data specific information

- Full names and defintions (e.g. for column headings)
- Units of measurement
- Definitions for codes or symbols used to record missing data
- Formats and abbreviations used

4. Licence information

- Licences and restrictions placed on the data. Licenses allow you to specify the 'terms-of-use' for your data. The [public license selector](#) can help you to choose the right license for the use of your data.

Examples of a good README file which were submitted to the GEOMAR OPeNDAP Server for publication

Example 1:

January 2017, K. Kvale (kkvale@geomar.de)
This directory contains all files and links to files needed to reproduce Kvale et al. (2017), Evaluation of the Transport Matrix Method for simulation of ocean biogeochemical tracers, Geoscientific Model Development Discussions:

reproduce_the_paper/
This directory contains:

- 1) Model output used to make the figures found in the manuscript (Offline_output/, UW3_output/, FCT_output/)
- 2) Figure scripts used in the manuscript (figure_scripts/)

--

UVic_Kiel_increase_isopyc_diff/ This directory contains transport matrices and related files used in the offline simulations.

--

UVOK_Matrix1_test/
This directory contains the source code of UVOK, the TMM interface to the Kiel version of UVic ESCM's biogeochemical model; Matlab scripts to generate all necessary input data and load output; and an example run script. This is the code used in the paper. The latest version is always to be found at <https://github.com/samarkhatiwala/tmm/>, where you can also find more details on using the TMM.

To run this code requires a number of components that are not distributed here:

- 1) The full, updated UVic ESCM 2.9 model code available from:
UVic_ESCM.2.9.updated.tar.gz (2013-8-1) at
<http://wikyonos.seos.uvic.ca/model/>
- 2) Modifications to (1) involving extensive changes of the biogeochemical model and code implementing transport matrix extraction and interfacing of the biogeochemical model to the TMM:
https://thredds.geomar.de/thredds/catalog/open_access/Kiel_Jan_2017_updates_to_UVIC2.9/catalog.html OR
http://kelvin.earth.ox.ac.uk/spk/Research/TMM/Kiel_Jan_2017_updates_to_UVIC2.9.tar.gz
- 3) Matlab scripts to generate input data and read output:
http://kelvin.earth.ox.ac.uk/spk/Research/TMM/tmm_matlab_code.tar.gz
Unpack and add all the directories to your Matlab path.
- 4) A working installation of PETSc (<http://www.mcs.anl.gov/petsc/>) compiled with MPI and Fortran support. The TMM code is compatible with PETSc version 3.6.x.

Once you have all the above components:

- 1) Edit Makefile to set the variables BASES1 to point to the toplevel of Kiel_Jan_2017_updates_to_UVIC2.9; BASE2 to the version 2.9 updates directory; and BASE3 to the main 2.9 source tree.
- 2) Compile with: make smallf make tmmuvok
- 3) Generate input data by executing 'make_input_files_for_uvok_model.m' in Matlab.
- 4) Run the model using 'runscript'. 'tmmuvok' is a parallel program
- 5) which you can execute with mpirun/mpiexec or whatever the appropriate wrapper command is on your machine.
- 6) Load the output in Matlab with load_output_time_avg.m and load_diagnostics.m. This will also write the output in netcdf format.

Example 2:

http://data.geomar.de/thredds/catalog/open_access/ruehs_et_al_2017_jpo/catalog.html/

Siren Rhs (November, 2017)

Data and code access for Eddy diffusivity estimates from Lagrangian trajectories simulated with ocean models and surface drifter data - a case study for the greater Agulhas system by Rhs et al. (2017) published at JPO (doi: 10.1175/JPO-D-17-0048.1)

Summary: This repository provides the original set of simulated trajectories from the Lagrangian reference experiment (SIMeddy-5d) needed for reproducing all major results of the above mentioned study (by applying the methods thoroughly described in the manuscript). It further contains all inferred data as well as a bundle of scripts necessary to repeat the production of Figures 2-5, which visualize the detailed evaluation of experiment SIMeddy-5d and its comparison with drifter derived diffusivities (OBS). All further data from additionally performed Lagrangian experiments, e.g., to assess the sensitivity of the diffusivity estimates to the resolution of the Eulerian model fields as well as to smoothing choices, can be provided on request from the corresponding author (sruehs@geomar.de or siren.ruehs@gmail.com).

Repository content:

/Ruehs2017_Data

Data folder with:

- 1) Set of simulated trajectories for the Lagrangian reference experiment SIMeddy-5d/SIMeddy5d_ARIANE_TrajQual.nc [1.7G]
- 2) Diffusivity estimates derived from the simulated drifter tracks 2D spatial pattern of maximum and asymptotic near-surface eddy diffusivities as well as pseudo-Eulerian mean currents and EKE (in overlapping 5deg bins):
 /SIMeddy5d_Diffu_5degbins.nc [22K]
 /SIMeddy5d_EKE_5degbins.nc [19K]
- Time-lag dependent diffusivity and dispersion estimates for selected bins
 /SIMeddy5d_Diffu_Disp_5degbin_AC.nc [3K]
 /SIMeddy5d_Diffu_Disp_5degbin_ARC.nc [3K]
 /SIMeddy5d_Diffu_Disp_5degbin_AR.nc [3K]
 /SIMeddy5d_Diffu_Disp_5degbin_eSAG.nc [3K]
- 3) Diffusivity estimates derived from observed drifter tracks 2D spatial pattern of maximum and asymptotic near-surface eddy diffusivities as well as pseudo-Eulerian mean currents and EKE (in overlapping 5deg bins):

/OBS_Diffu_5degbins.nc [22K]
/OBS_EKE_5degbins.nc [19K]

- 4) Additional masks needed to reproduce Figures 2 to 5 with the data provided by this repository
/INALT01_topo_k2.nc [7.2M]
/iffu_topo_mask.nc [8.5 K]

Please refer to ./Ruehs2017_Data/README.md for more detailed information on the individual datasets.

/Ruehs2017_Scripts

Script folder with:

- 1) PyFerret script to replot Figure 2
/Plot_Ruehs2017_Figure2.jnl
- 2) PyFerret script to replot Figures 3 to 5
/Plot_Ruehs2017_Figure3to5.jnl
- 3) Customized colors/colorbars needed to reproduce Figures 2 to 5
/myred.spk
/mycolbar.spk

Please refer to ./Ruehs2017_Scripts/README.md for more detailed information on how to run the scripts. Feel free to use and distribute these scripts under the License agreement ./Ruehs2017_Scripts/LICENSE.