Plan Overview

A Data Management Plan created using DeiC DMP

Title: HPC environment FMHAT

Creator: Sissal Erenbjerg

Principal Investigator: Sissal Erenbjerg

Data Manager: Sissal Erenbjerg

Affiliation: Københavns Universitet / University of Copenhagen

Template: DCC Template

ORCID iD: 0000-0001-7619-4049

Project abstract:

This project is to setup and continue a HPC center in the Faroe Islands. For this partuícular project the focus is on Faroe shelf 3d ocean modelleing with FarCoast and ROMS.

ID: 3624

Last modified: 30-06-2023

HPC environment FMHAT

Data Collection

What data will you collect or create?

Simulated physical ocean parameters and tracers from the Faroese shelf from 2010-2020. Including but not limited to current velocities, salinity, temperature etc.
This will be about 5-10 TB of data, and will be saved in netcdf format. The netcdf format is comprised and easy to store longterm.

How will the data be collected or created?

The data will be collected by running a longterm computer simulation of FarCoast160, on the local HPC environment at Fiskaaling.The data will be named by date and saved in a netcdf format with meta data. As in the standards used by the Regional Ocean Model System (ROMS). The entire dataset will be validated against excisting data at Fiskaaling.
The files will be located at offline servers at Fiskaaling, but available upon request for non-commersial use.

Documentation and Metadata

What documentation and metadata will accompany the data?

The data will be stored in netcdf format and the netcdf library is neccasary in order to read and use the dataset. The means that each file both contains the data and the metadata, each day will be saved in a seperate file including 24 one hour timestep averages. Each file will also contain the information of model version, period, time-step, and included model options, dimensions of the grid and a short description of all variables (including a short name, a long name, units and variable dimensions). This is the standard with in the Regeional Ocean Model System (ROMS).

Ethics and Legal Compliance

How will you manage any ethical issues?

Since this data only contains simulated physical variables, in a simulated environment, forced by puplically available data, at least for non-commersial use, there should not be any ethical concerns to our knowledge. The entire dataset can be shared with out restrictions, for non-commersial use. This project does not use any private or personal data (GDPR does not apply).

How will you manage copyright and Intellectual Property Rights (IPR) issues?

The team at Fiskaaling will have a two year priviligae for possible publication of the data, afterwards the data will be freely available for anyone that wishes to use the data for non-commersial use. The entire model (ROMS), forcing (WRF, Copernicus) and dataset is open source.

Storage and Backup

How will the data be stored and backed up during the research?

The data will be stored offline on servers at Fiskaaling. Where the data will also be backed up. The responsibility of backup is within the team of Fjord dynamics at Fiskaaling. The data will also be possible to recover by rerunning the model, potentially on another machine. The ability of creating restart files within FarCoast and ROMS ensures faster data recovery as well.

How will you manage access and security?

The access to the data will be at Fiskaaling since the data will mainly be stored offline, however the access can be gained with an ssh tunnel into the backup system.

Selection and Preservation

Which data are of long-term value and should be retained, shared, and/or preserved?

The model setup is self does not require a large space, >100GB (excluding forcing) wheras the dataset will become rather larger, dependent on number of tracers that will be included (at least 3). However the dataset will be kept off-line at fiskaaling.
The number of tracers will also depend on users in the general research environment of the Faroe Islands. Where the data set can be used as hydrodynamical input for sea-lice dispersion, wave and coastal simulation, resuspention of POM, fish welfare in fish farming cages and other.

What is the long-term preservation plan for the dataset?

The data repository will stay offline at Fiskaaling due to the large size of the data set. This is included in the budget of Fiskaaling.

Data Sharing

How will you share the data?

The data will be validated in scientific papers and thus be available for others to use in scientific prospects. The data will be available for others researchers 2 year after the data is created. Some parts of the data can be shared though gitlab, as well as parts of the code formulation can be shared at github.
This will seccure that the data set is findable, accessible, interoperable and re-usable (FAIR).

Are any restrictions on data sharing required?

We will require a data sharing agreement with any users, as we wish that the data set remains open and freely available and is not used for commercial.

Responsibilities and Resources

Who will be responsible for data management?

The main responsible for this DMP is Sissal Erenbjerg, PHD student at Fiskaaling, this DMD will be reviewed and revised anually by her, and the collagues at Fiskaaling including Trondur Tummason and Jóhannus Kritsmoundson at the Fjord ocenaography department and Gunnvør á Nórði and Birgitta Andreassen at ecology department. If funding for this project will be provided, this DMP will be publically available online at dmponline.deic.dk.

What resources will you require to deliver your plan?

For Backup 10 TB of offline storage will be purchased at a cost of about 7000 DKK.