



A Galaxy based webportal:
openIFS use case

Outline

- UIO geoportal:
 - What & Why?
 - How?
- OpenIFS at UIO:
 - What & Why?
 - How?
- Example:
 - Generation of meteorological fields for OsloCTM3
- And then what next?

UIO Geoportal

What?

- **Geoportal:** Geosciences Portal
- A **platform independent**, open, web-based platform for running scientific applications in the domain of Geosciences
- A **workflow-centric Galaxy:** Geoportal will host a set of tools that can be chained to form complex workflows
- Galaxy portals pass jobs to a **High Performance Computing (HPC)** Linux cluster (abel.uio.no)

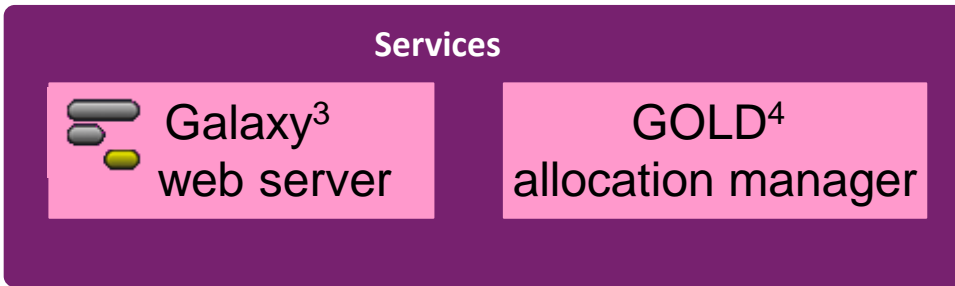
Why?

- Improve the **availability** and **usability** of large/complex Models
- Provide an **online learning** environment to students and teachers
- **Foster collaboration** and interdisciplinary research
- Address the “**reproducibility challenge**”
- **Share** scientific research to increase the **visibility** of researcher’s work and the recognition for it

How?

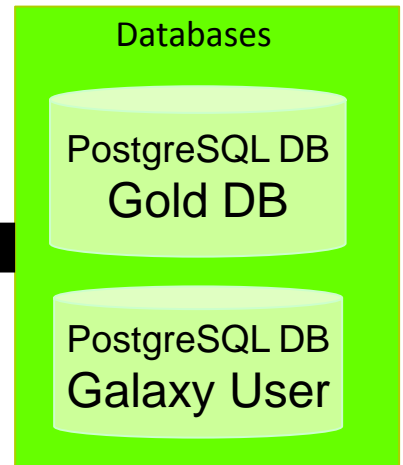
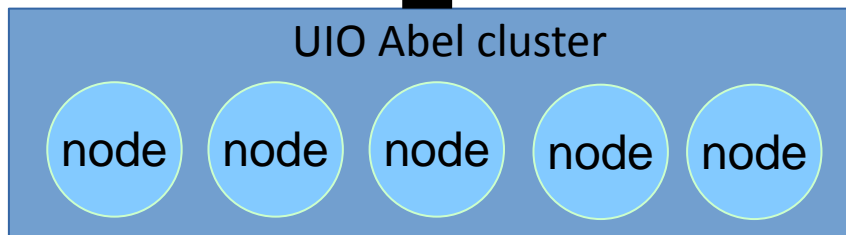
Transfer of knowledge from other Galaxy portals maintained at the University of Oslo, such as the lifeportal and the Language Analysis Portal

Autentication
(FEIDE¹, Ne²)



Resource management
DRMAA

Job scheduler
SLURM



SSL connection

¹ Federated National Identity

² Norwegian einfratructure

³ <https://galaxyproject.org/>

⁴ <http://www.adaptivecomputing.com/products/open-source/gold/>

Galaxy portal



Tool menu

Task banner

The screenshot shows the Geoportal web interface. At the top, there's a navigation bar with "Apply for a project", "Analyze Data", "Workflow", "Shared Data", "Visualization", "Admin", "Help", and "User". Below this is a "Tools" sidebar on the left with a search bar and various tool categories like "Get Data", "Send Data", "Tests", and "Workflows". The main content area features a large "GEO PORTAL" logo with several scientific data visualizations (maps, charts, and 3D models). Below the logo is a green status bar with a checkmark icon and the text "Geoportal-dev is up and running...". Underneath this, a red banner reads "The Geoportal is NOT OPERATIONAL yet!". At the bottom, a paragraph states: "Geoportal will give you easy access to the High Performance Computing cluster Abel at the University of Oslo. The Galaxy based Geoportal is under development and may be shutdown at anytime". On the right side, there's a "History" panel with a search bar and a list of recent tasks, including "71: OpenIFS InterpolationTest.log", "70: OpenIFS InterpolationTest.html", "58: OpenIFS EIApril2009.html.zip", "46: OpenIFS EIApril2009.html", "32: ifsnam_fc_template.nml", "31: ifsnam fpos template.nml", "25: Test.html", "15: options.tar", "8: Abel handshake", and "4: Abel handshake".

Result frame

History

OpenIFS at UIO

- What?

- **A modified version of openIFS** for generating meteorological fields for Oslo CTM3 model
- Visualization: Metview, magics++, python, matlab
- A **workflow-centric Galaxy**: python scripts to run openIFS
- Installed on Notur HPCs:
abel.uio.no (Sandy bridge, FDR interconnect, 10000+ cores, FhGFS), vilje.hpc.ntnu.no (SGI Altix ICE X, Sandy bridge, 22464 cores, FhGFS), hexagon.bccs.uib.no (Cray XE6, 22272 cores, Lustre)

- Why?

- Computation of **convective mass fluxes and detrainment rates** (variables not available in MARS) for OsloCTM model
- Full control on the IFS model version used
- Allow master Students to use one of the best forecast model
- And hopefully more once we are familiar with OpenIFS!

How?

- OpenIFS modified version kept and maintained in **git**
- Use **ECMWF Web API** to retrieve ERA-Interim fields from MARS
- **Python** & **bash** scripts to control openIFS **workflow** (get reanalysis data, interpolate input fields to model grid, run forecast, convert to Oslo CTM3 binary format)
- Write **new set of galaxy tools** for OpenIFS

Example: Generate Oslo CTM3 meteorological fields

Horizontal resolution: T42 (2.8 x 2.8 degrees) / T159 (1.125 x 1.125)

Vertical resolution: 60-layers (surface – 0.1hPa)

Transport: Secondary order moments scheme
(Prather et al., 2008), with improved
polar cap transport (Søvde et al., 2012)

Meteorology: ECMWF IFS / OpenIFS
3hourly forecasts, starting from ERA-Interim.
Uses convective mass fluxes and
detrainment rates.

Photolysis: Fast-JX version 6.7

Chemistry: Tropospheric (incl. sulphur), stratospheric

Aerosols: Tropospheric sulphate, sea salt, mineral dust, nitrate,
black and organic carbon, secondary organic aerosols.

Scavenging: Neu and Prather, 2012.

Emissions: anthropogenic and natural emissions, biomass burning from e.g.
GFEDv3, and surface concentrations of long-lived species.

Lightning: Horizontal dist.: Søvde et al (2012), vertical dist.: Ott et al (2010).

References

Ott et al. (JGR 2010), 10.1029/2009JD011880

Prather et al. (PNAS 2008), doi:10.1073/pnas.0806541106

Søvde et al. (accepted for GMD, 2012), doi:10.5194/gmdd-5-1561-2012

Galaxy tool: OpenIFS workflow



Get ERA Interim input data ✕

listfiles (html)

Interpolate and run OpenIFS ✕

namelist template file for interpolation

namelist template file for openIFS forecast run

Input GRIB data retrieved from MARS

Input GRIB directory

listfiles (html)

logfile (txt)

oifs2ctm ✕

tarball containing openIFS data

listfiles (html)

OsloCTM

OsloCTM

namelist

- 1- Get ERA Interim input data
- 2- Interpolate and run OpenIFS
- 3- Convert to Oslo CTM 3 binary format
- 4- Run Oslo CTM 3 (*not implemented yet*)

OpenIFS: Setup experiment

Galaxy / Geoportal x

https://geoportal-dev.hpc.uio.no

Galaxy Geoportal Apply for a project Analyze Data Workflow Shared Data Visualization Admin Help User Using 893.8 MB

Tools

search tools

Get Data

Send Data

Lift-Over

Tests

OsloCTM Model

[oifs2ctm](#) Generate OsloCTM data from OpenIFS outputs

[OsloCTM](#) Run OsloCTM model

OpenIFS Model

[Get ERA Interim input data](#)

[Fetch ERA Interim data from ECMWF](#)

Interpolate and run OpenIFS
interpolate data to your model domain and run openIFS

WRF Model

MITgcm Model

SPEEDY Global Model

ROMS Ocean Model

Flexpart Model

Filter and Sort

Join, Subtract and Group

Convert Formats

Statistics

Graph/Display Data

Interpolate and run OpenIFS interpolate data to your model domain and run openIFS (Galaxy Tool Version 1.0.0) Options

Remaining CPU hrs in your default project **gx_default** are : 198.23

OpenIFS version

38r1v04

Experiment Identifier

Test

Starting date YYYYMMDD

20090415

Ending date YYYYMMDD

20090415

times to retrieve HH separated by slash to specify for more than one time

00

Input filename for surface data available on a reduced gaussian grid

iniggsfc.grb

Input filename for model level data available on a reduced gaussian grid

ggml.grb

Input filename for spherical harmonics model level data

shml.grb

output resolution

42

grid type

46: OpenIFS_EIApril2009.html

openIFS configuration

osloCTM

Job Resource Parameters

Specify job resource parameters

Processors (processes)

1

History

search datasets

Unnamed history

12 shown, 63 deleted

1.8 GB

75: OpenIFS InterpolationTest.log

185 lines

format: **txt**, database: ?

Starting job 11446318 (") on c2-14 at sø. 07. juni 13:50:43 +0200 2015

total 1

```
-rwxr-xr-x 1 galaxy galaxy 1005 2015-06-07 13:50 galaxy_101.sh
-rw-r--r-- 1 galaxy galaxy 73 2015-06-07 13:50 galaxy_101.o
-rw-r--r-- 1 galaxy galaxy 0 2015-06-07 13:50
```

Code 174 Date/time= 200

Interpolated between month 4 with

Code 15 Date/time= 200

Interpolated between month 4 with

Code 16 Date/time= 200

OpenIFS: Setup experiment

Galaxy / Geoportal x

https://geoportal-dev.hpc.uio.no

Galaxy Geoportal Apply for a project Analyze Data Workflow Shared Data Visualization Admin Help User Using 893.8 MB

Tools

search tools

Get Data

Send Data

Lift-Over

Tests

OsloCTM Model

[oifs2ctm](#) Generate OsloCTM data from OpenIFS outputs

[OsloCTM](#) Run OsloCTM model

OpenIFS Model

[Get ERA Interim input data](#)

[Fetch ERA Interim data from ECMWF](#)

[Interpolate and run OpenIFS](#) interpolate data to your model domain and run openIFS

WRF Model

MITgcm Model

SPEEDY Global Model

ROMS Ocean Model

Flexpart Model

Filter and Sort

Join, Subtract and Group

Convert Formats

Statistics

Graph/Display Data

46: OpenIFS_EIApril2009.html

openIFS configuration

osloCTM

Job Resource Parameters

Specify job resource parameters

Processors (processes)

16

Number of processing cores, 'ppn' value (1-64). Leave blank to use default value.

Processors (threads)

1

Number of processing cores, 'ppn' value (1-16). Leave blank to use default value.

Memory

1

Memory size in gigabytes, 'pmem' value (1-900). Leave blank to use default value.

Time

1

Maximum job time in hours, 'walltime' value (1-672) or 28 days. Leave blank to use default value.

Project

Default Geoportal 200hrs project

Project to assign resource allocation to. Leave blank to use default value.

Execute

SYNTAX

History

search datasets

Unnamed history

12 shown, 63 deleted

1.8 GB

75: OpenIFS InterpolationTest.log

185 lines

format: **txt**, database: ?

```
Starting job 11446318 ("") on c2-14 at sø. 07. juni 13:50:43 +0200 2015
total 1
-rwxr-xr-x 1 galaxy galaxy 1005 2015-06-07 13:50 galaxy_101.sh
-rw-r--r-- 1 galaxy galaxy 73 2015-06-07 13:50 galaxy_101.o
-rw-r--r-- 1 galaxy galaxy 0 2015-06-07 13:50
```

Code 174 Date/time= 200
Interpolated between month 4 with
Code 15 Date/time= 200
Interpolated between month 4 with
Code 16 Date/time= 200

Click on Execute to start running

OpenIFS: Run experiment

The screenshot displays the Galaxy/Geoportal web interface. The browser address bar shows <https://geoportal-dev.hpc.uio.no>. The top navigation bar includes links for 'Analyze Data', 'Workflow', 'Shared Data', 'Visualization', 'Admin', 'Help', and 'User'. The main content area features a green notification box with a checkmark icon, stating: '1 job has been successfully added to the queue - resulting in the following datasets:'. Below this, two datasets are listed: '70: OpenIFS_InterpolationTest.html' and '71: OpenIFS_InterpolationTest.log'. A paragraph of text explains that users can check the status of queued jobs and view the resulting data by refreshing the History pane. On the right side, the 'History' panel shows a list of datasets, including '71: OpenIFS InterpolationTest.log' and '70: OpenIFS InterpolationTest.html'. A 'View data' button is overlaid on the '71: OpenIFS InterpolationTest.log' entry. The bottom part of the history panel shows details for job 11437724, including the start time and a JSON configuration snippet.

Galaxy / Geoportal x

<https://geoportal-dev.hpc.uio.no>

Galaxy **Geoportal** Apply for a project Analyze Data Workflow Shared Data Visualization Admin Help User Using 893.8 MB

Tools

search tools

Get Data

Send Data

Lift-Over

Tests

OsloCTM Model

[oifs2ctm](#) Generate OsloCTM data from OpenIFS outputs

[OsloCTM](#) Run OsloCTM model

OpenIFS Model

[Get ERA Interim input data](#)

[Fetch ERA Interim data from ECMWF](#)

[Interpolate and run OpenIFS](#) interpolate data to your model domain and run openIFS

WRF Model

MITgcm Model

SPEEDY Global Model

ROMS Ocean Model

Flexpart Model

Filter and Sort

Join, Subtract and Group

Convert Formats

Statistics

Graph/Display Data

History

search datasets

Unnamed history

10 shown, 61 [deleted](#)

893.8 MB

71: OpenIFS InterpolationTest.log

70: OpenIFS InterpolationTest.html

58: OpenIFS EIApril 2009.html.zip

46: OpenIFS EIApril 2009.html

255 bytes

format: **html**, database: ?

Starting job 11437724 ("") on c1-17 at fr. 05. juni 22:22:59 +0200 2015

```
{
  "url" : "https://api.ecmwf.int/v1",
  "key" :
    "49a2868cd38100727ed2e394495"

  "email" : "annefou@geo.uio.no"
}
```

Code 174 Date/time=

View data

Galaxy tool: Explore outputs

The screenshot displays the Galaxy web interface. The main content area shows a list of output files, each with a blue underlined link. The files are:

- [./20090415/00/ICMSHTest+000108](#)
- [./20090415/00/NODE.001_01](#)
- [./20090415/00/ICMGGTest+000099.128](#)
- [./20090415/00/ICMGGTest+000036.128](#)
- [./20090415/00/ICMGGTest+000018](#)
- [./20090415/00/ICMGGTest+000000.128](#)
- [./20090415/00/ICMSHTest+000000](#)
- [./20090415/00/ICMGGTest+000000](#)
- [./20090415/00/ICMGGTest+000045.128](#)
- [./20090415/00/ICMSHTest+000018](#)
- [./20090415/00/ICMSHTest+000009](#)

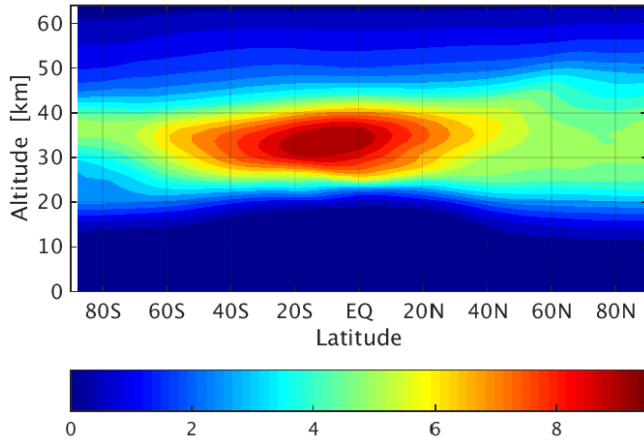
On the right side, the History panel shows a list of datasets. Two red callout boxes provide instructions:

- A red box with a pointer to the file link [./20090415/00/ICMGGTest+000018](#) contains the text: "Click on a file to download it".
- A red box with a pointer to the file `74: OpenIFS InterpolationTest.html` in the history panel contains the text: "Click to download all the output directory".

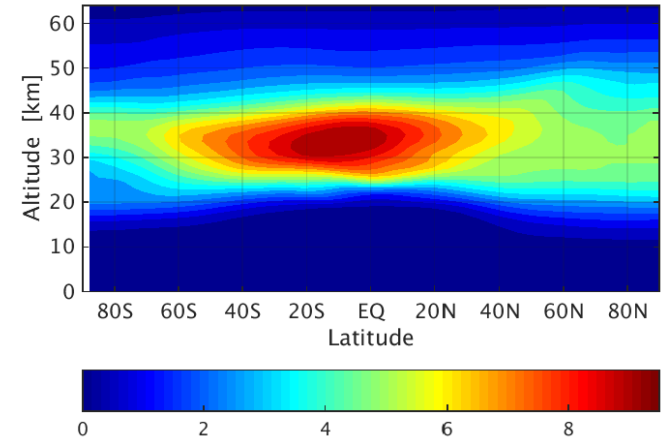
The History panel also shows details for the selected dataset, including its size (4.0 KB), format (html), and a log of the job execution.

OsloCTM: visualize & validate

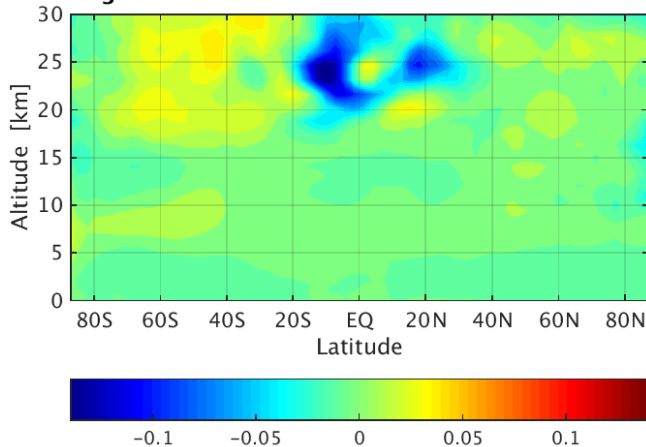
O₃ January 2013 cy38



O₃ January 2013 openIFS



O₃ (ppmv) January 2013 openIFS-cy38



Comparison between OpenIFS & IFS CY38R2:

The figures show that there are slightly no differences between a run with open IFS and a run using the IFS model CY38R2 at ECMWF. The plots are zonal mean of ozone for of January 2013

And then what next?

- New tools and improve existing workflows: visualization, archiving in the Norwegian National Archive (NorStore)
- Set-up tutorials and organize trainings once operational.