

The Second Global Open Science Cloud (GOSC) workshop

Incoherent scatter radar data fusion and
computation case study

*Ingemar Häggström/EISCAT Scientific Association
Team EGI/CAS/CNIC/CoData
22 October 2021*



GOSC Case 1 Space Physics

Incoherent scatter radar data fusion and computation

EISCAT-3D radar, EISCAT Scientific association.

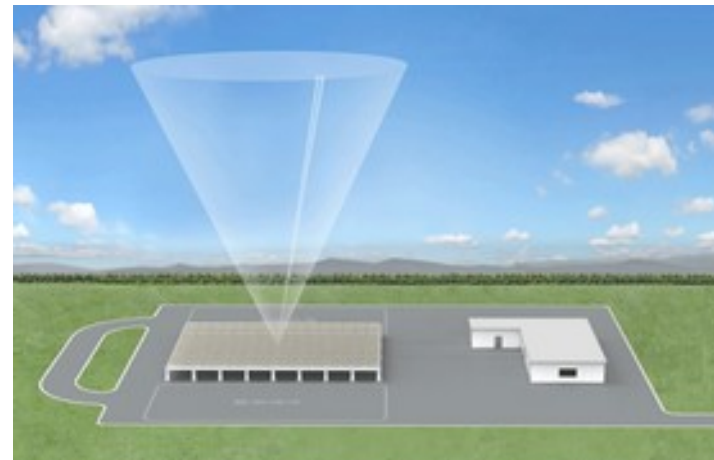
Sanya Incoherent Scatter radar (SYISR), Institute of Geology and Geophysics.



Menu: HQ Operations Schedule, June 2021

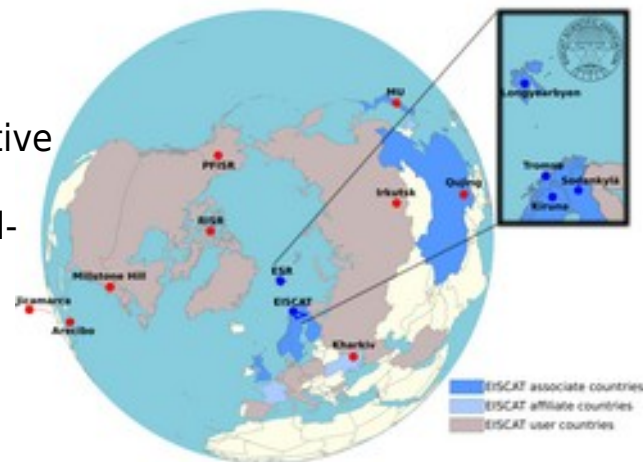
Order	Order Date and Month	Status	Order ID
1	2021-06-01	Completed	123456789
2	2021-06-02	In Progress	987654321
3	2021-06-03	Planned	567890123
4	2021-06-04	Planned	234567890
5	2021-06-05	Planned	890123456
6	2021-06-06	Planned	456789012
7	2021-06-07	Planned	012345678
8	2021-06-08	Planned	678901234
9	2021-06-09	Planned	345678901
10	2021-06-10	Planned	901234567

- Next generation of incoherent scatter radar systems
 - HPLA HighPowerLargeAperture
- Similar hardware
 - Systems of distributed phased array radars
 - Enable comprehensive three-dimensional observations of the atmosphere and ionosphere.
 - Increased temporal and spatial resolution
 - Continuous measurement capabilities
 - Inclusion of detailed incoherent scatter radar data into climate and Earth system modeling.
 - Large VVVV
 - Velocity, Volume, Versatility, Variability

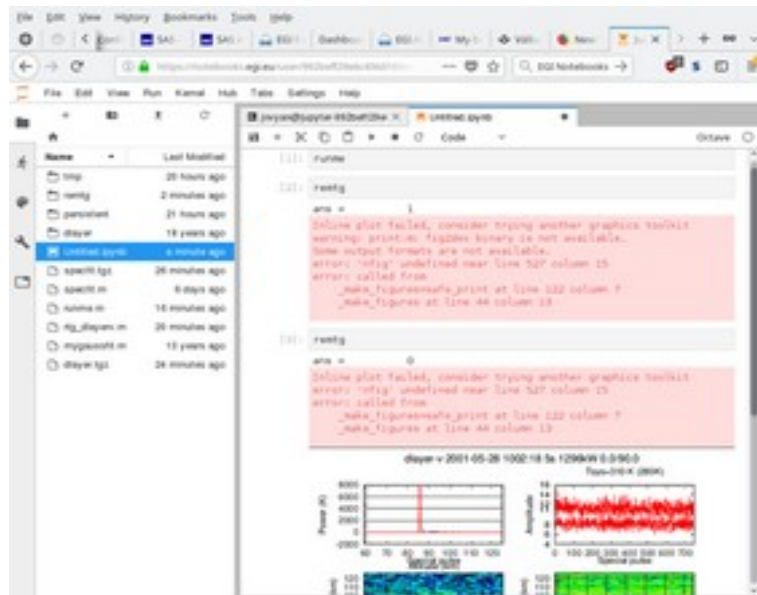


Use Case Mission

- Objectives
 - To explore data and metadata models for collaborative production and sharing.
 - To develop enhanced algorithms, methods for cloud-based radar data processing and federated data learning.
 - Other issues to address for data interoperability.
-
- Key deliverables
 - Exploration of technical solutions for EISCAT and SYISR (Meta)data federation.
 - Deployment of GOSC technical solutions supporting the cloud federated data processing and on-demand data movement.
 - Lessons and good practices for analogous research infrastructures in OSCs.



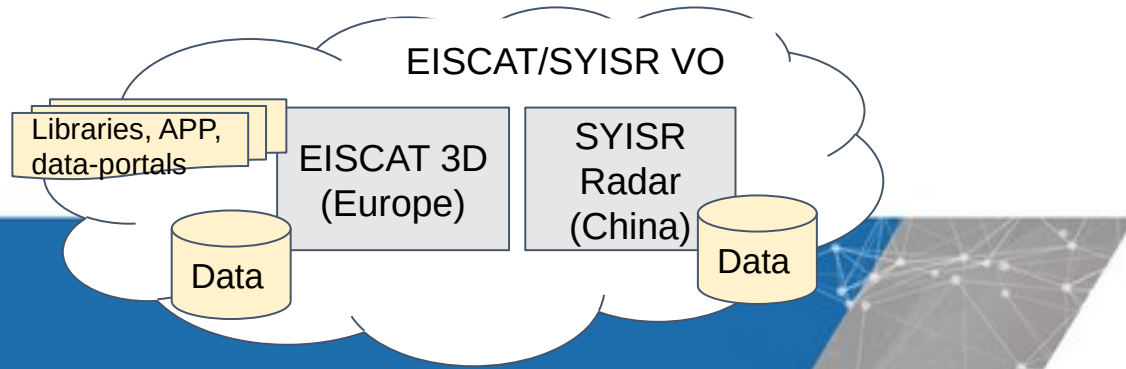
- Metadata and data federation based on radar data production
 - Real-time radar data sharing policies and techniques.
 - Metadata and data standards for radar data production.
- A joint radar data campaign
 - Carry out joint a data campaign with cross-domain radar data as byproducts.
 - Carry out tests for data quality validation and the technical performance of large-scale data movement between regions.
 - Workflow specification that is adaptive to multiple sources of data capture, real-time data sharing, and federated data learning.
- Monthly WG meetings, regular science communication, training, other social activities for publicity and potential alignment.




Job definition file

```
[
  Executable = "run_rtg_docker.sh";
  Arguments = "";
  JobName = "my_job_name";
  Site = "Cloud.CSC.fi";
  CPUTime = 86400;
  InputSandbox = {
    "run_rtg_docker.sh",
    "LFN:/eiscat.se/archive/"
  }
  2016/...
  OutputSandbox = {"output/*"};
]
```

Integration Chinese CSTCloud with EGI



30 cores /100GB
to EISCAT VO
from CSTCloud

Check-in 

OpenStack Dashboard
<https://federation.cstcloud.cn/horizon/auth/login/?next=/horizon/>

GOCDDB

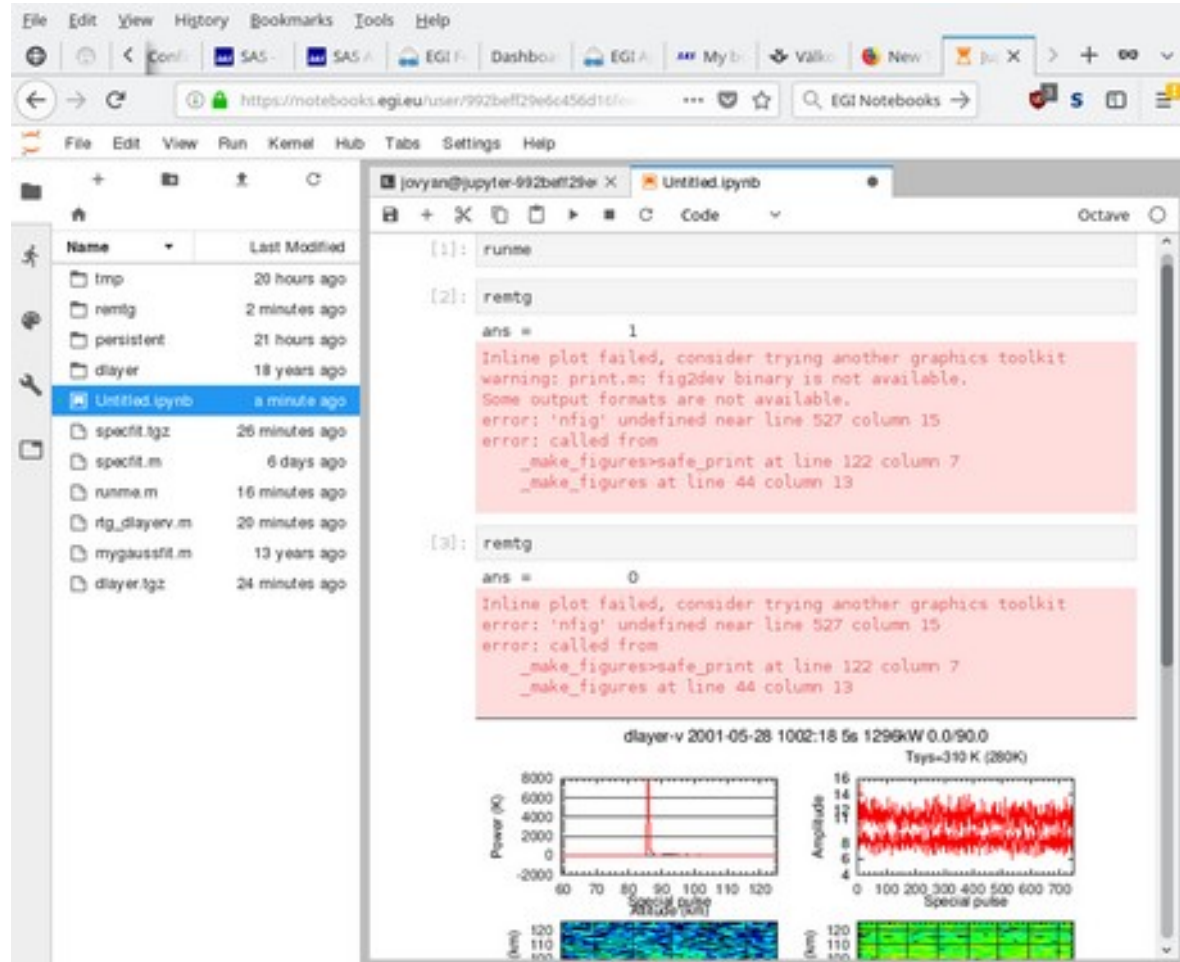
A contribution to
the **Global Open
Science Cloud**

The bottom section features the EGI logo on the left and the China Science & Technology Cloud logo on the right. The China Science & Technology Cloud logo includes the Chinese characters "中國科技雲" and the English text "China Science & Technology Cloud". An upward-pointing arrow is positioned between the two logos.

Refer to
[EGI News](#) for
details

Task 1 Harmonisation work...

- Data levels
 - L1, Amplitude data
 - L2, Spectral data
 - L3, Data products
- Processing
 - L1→L2
 - Lag profiling
 - L2
 - **Visualisation**
 - L2→L3
 - **Analysis**
- Tools
 - **Jupyter**
 - Matlab for jupyter
 - HDF
 - Metadata



The screenshot shows a Jupyter Notebook interface with a file browser on the left and a code editor on the right. The file browser displays a list of files and folders, including 'tmp', 'remtg', 'persistent', 'dlayer', 'specfit.tgz', 'specfit.m', 'runma.m', 'rtg_dlayer.v', 'mygaussfit.m', and 'dlayer.tgz'. The code editor shows three code cells:

```
[1]: runme
```

```
[2]: remtg
```

```
ans =
```

```
1
```

Two error messages are displayed in red boxes:

```
Inline plot failed, consider trying another graphics toolkit
warning: print.m: fig2dev binary is not available.
Some output formats are not available.
error: 'nfig' undefined near line 527 column 15
error: called from
_make_figures>safe_print at line 122 column 7
_make_figures at line 44 column 13
```

```
[3]: remtg
```

```
ans =
```

```
0
```

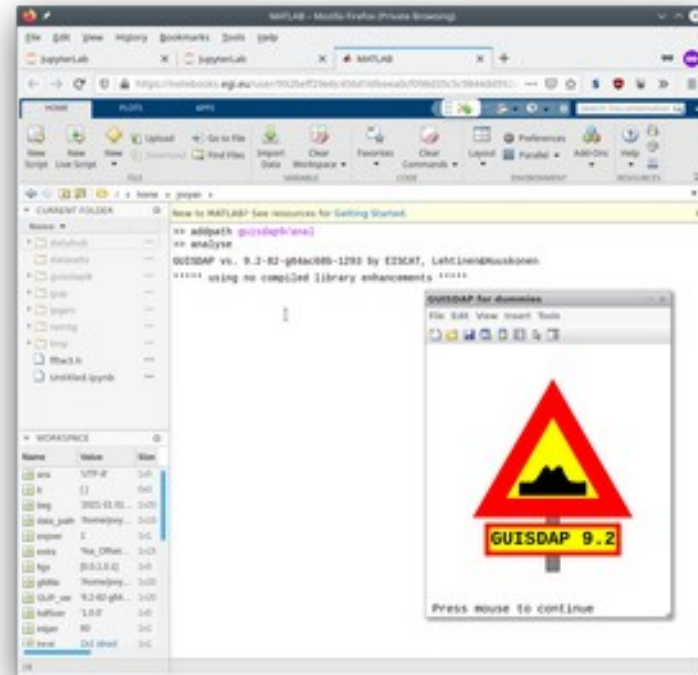
Another error message is displayed in a red box:

```
Inline plot failed, consider trying another graphics toolkit
error: 'nfig' undefined near line 527 column 15
error: called from
_make_figures>safe_print at line 122 column 7
_make_figures at line 44 column 13
```

Below the error messages, there are two plots and two heatmaps. The first plot is titled 'dlayer-v 2001-05-28 1002:18 5s 1296kW 0.0/90.0' and shows Power (K) vs. Special pulse. The second plot is titled 'Tsys=310 K (280K)' and shows Amplitude vs. Special pulse. The two heatmaps show data for 'Special pulse' vs. 'Time (s)'.

Task 1 Harmonisation work...

- Data levels
 - L1, Amplitude data
 - L2, Spectral data
 - L3, Data products
- Processing
 - L1→L2
 - Lag profiling
 - L2
 - Visualisation
 - L2→L3
 - **Analysis**
- Tools
 - Jupyter
 - **Matlab for jupyter**
 - HDF
 - Metadata



- **Oct21 - Mar22**
 - Elaborated radar data campaign plan, and study on data and metadata sharing policies and **techniques**.
- Apr22 - Sep22
 - Fulfill the joint radar data campaign plan, and carry out the two experimental tests for data quality validation and large-scale data movement between regions.
- Oct22 - Mar23
 - Exploration of interoperable data standards and workflows.
- Apr23 - Sep23
 - Training, workshops, and community outreach for larger engagement and impacts
- Meetings/Reporting
 - Monthly co-chair calls
 - Quarterly CS member calls
 - **Annual EGI Conference, 19 - 21 October 2021**
 - SciDataCon 2021, 18 - 28 October 2021. GOSC work plans presentation on 27 October
 - International Data Week 2022, 20 - 23 June 2022, Seoul.
 - FDO Conference and the FAIR Convergence Symposium, October, 2022. Leiden.
 - International Data Week 2023, October, Salzburg.