Storing data in EISCAT HDF5 format

Rikard Slapak

February 26, 2021

Contents

1 Introduction

The purpose of this document is to describe the procedure of generating EISCAT HDF5 files from data obtained by the current radar system at EISCAT. It can be done directly when running the analysis using the GUISDAP software, but there are also procedures to handle old and already analysed data that has been stored and formated for Madrigal.

The routines developed to handle this are MATLAB functions that has been implemented into GUISDAP:

`EISCAT_hdf5.m`
`mat2hdf5.m`
`metacompl.m`
`cedar2hdf5.m`
`store_image2hdf5.m`
`strds2hdf5.m`

and their features will be described shortly in this document.

2 Generating HDF5 files

2.1 New analysis

After analysis and calibration of an experiment an HDF5 file with all data can be made by calling the function `mat2hdf5`:

>> `mat2hdf5(datapath,resultpath,addfigs,addnotes)`.

`datapath` points to the data (MAT-files), metadata, figures and notes from the analysis and calibration of a certain experiment. `resultpath` set the path where a folder – containing the generated HDF5 file, PDF:s of figures and a TAR file conaining all inputs – is stored. The folder is named the same as HDF5 file itself (but without the extension obviously). For example, a file `EISCAT_yyyy_mm_dd_expr_intper@ant.hdf5` will be stored in

`/resultpath/EISCAT_yyyy_mm_dd_expr_intper@ant/`

together with PDF(s) and the TAR file. The inputs `addfigs` and `addnotes` must be set non-empty in order for the routine to store any figures and notes*.txt in the HDF5 file.

For multistatic experiments `mat2hdf5` will call the function `mat2hdf5_vel` to create the corresponding EISCAT HDF5.

2.2 Old analysis

There are additional routines to generate EISCAT HDF5 files from older analysed data, already stored in Madrigal-2. These data can be retrieved at the EISCATHQ server at `eiscathq:/opt/madrigal/experiments`. The data are then sorted under year, site and date. Experiments analysed by On-An have their data stored in older HDF5 in correspondence to the Madrigal-2 required

format and structure, whereas GUISDAP-analysed data are in a TAR-file in the form av MAT-files. For a given date several experiments and/or analyses can be stored, so that the figures need to be sorted when generating new EISCAT HDF5s.

In order to handle all experiments for a given date the function `EISCAT_hdf5` is used. It requires an input folder `datapath` (for example `eiscathq:/opt/madrigal/experiments/year/site/date`) and an output folder `resultpath` where the result folders, containing the generated EISCAT HDF5 files and PDFs for example, are stored.

```
>> EISCAT_hdf5(datapath,resultpath,showfigs)
```

The third input argument determines if the figure(s) in the datapath folder will be displayed (`showfigs` non-empty) or not.

The function looks for TAR.GZ files containing MAT files generated by Guisdap and for HDF5s with On-An analysed data. A TAR.GZ file will be unpacked and uncompressed and its content is put in a temporal folder `untaredpath`. which then work as the data input for `mat2hdf5` to generate a new EISCAT HDF5 file.

For On-An analysed data the function `cedar2hdf5` (or `cedar2hdf5_vel` for multistatic analyses) is called to generate an EISCAT HDF5 from the Madrigal-2 HDF5.

```
>> cedar2hdf5(madHDF5file,resultpath)
```

One year of old data

A function `EISCAT_hdf5_year` takes care of a whole year of Madrigal-2 data. Instead of a date as input a specific year is given (e.g. `eiscathq:/opt/madrigal/experiments/year`).

```
>> EISCAT_hdf5_year(yearpath,resultpath,listmatfile)
```

If list-matfile is left out or empty all sites and dates will be considered and a new EISCAT HDF5 will be created for each analysis and experiment. For each analysis a matfile containing three lists will be updated. *List_HandledOK*: all dates that have been handled without issues (without crashing somewhere), *List_2BHandled*: all dates not yet handled, and *List_Crashed*: all dates where the routine crashed somewhere. By having the list-matfile (if exists) as input, the routine will go on with dates in *List_2BHandled* or *List_Crashed*.

3 Functions

Below follow a summary of the functions in the routine of generating EISCAT HDF5s.

3.1 Main functions

`mat2hdf5:` Generating an EISCAT HDF5 from a Guisdap analysed experiment.
`mat2hdf5:` Generating an EISCAT HDF5 from a Guisdap analysed multistatic

experiment.
`cedar2hdf5:` Generating an EISCAT HDF5 from a On-An analysed experiment stored in Madrigal-2.
`cedar2hdf5_vel:` Generating an EISCAT HDF5 from a On-An analysed multistatic experiment stored in Madrigal-2.
`EISCAT_hdf5`: Generating EISCAT HDF5s for all Madrigal-2 stored experiment/analyses for a specific date, by calling above functions.
`EISCAT_hdf5_year`: Generating EISCAT HDF5s for all Madrigal-2 stored experiment/analyses for a specific year, by calling above functions.

3.2 Some additional functions

`note2hdf5`: writes notes/comments to an EISCAT HDF5.
`image2hdf5`: writes figures to an EISCAT HDF5.
`strds2hdf5`: writes string datasets (structured in cells) to an EISCAT HDF5.
`metacompl`: when Guisdap analysed experiments miss setup data, metadata from the corresponding Madrigal-2 HDF5 is taken.

4 EISCAT HDF5 file structure and content

The format of an HDF5 is of hierarchical type, where the data and metadata are sorted in different groups based on the data characteristics. The structure and contents of the EISCAT HDF5 files are described in the document "EISCAT Level 3 data storage: the HDF5 files".