

How to use the Madrigal database for atmospheric science

Bill Rideout

MIT Haystack Observatory

brideout@haystack.mit.edu

ISR workshop

Sodankylä Geophysical Observatory

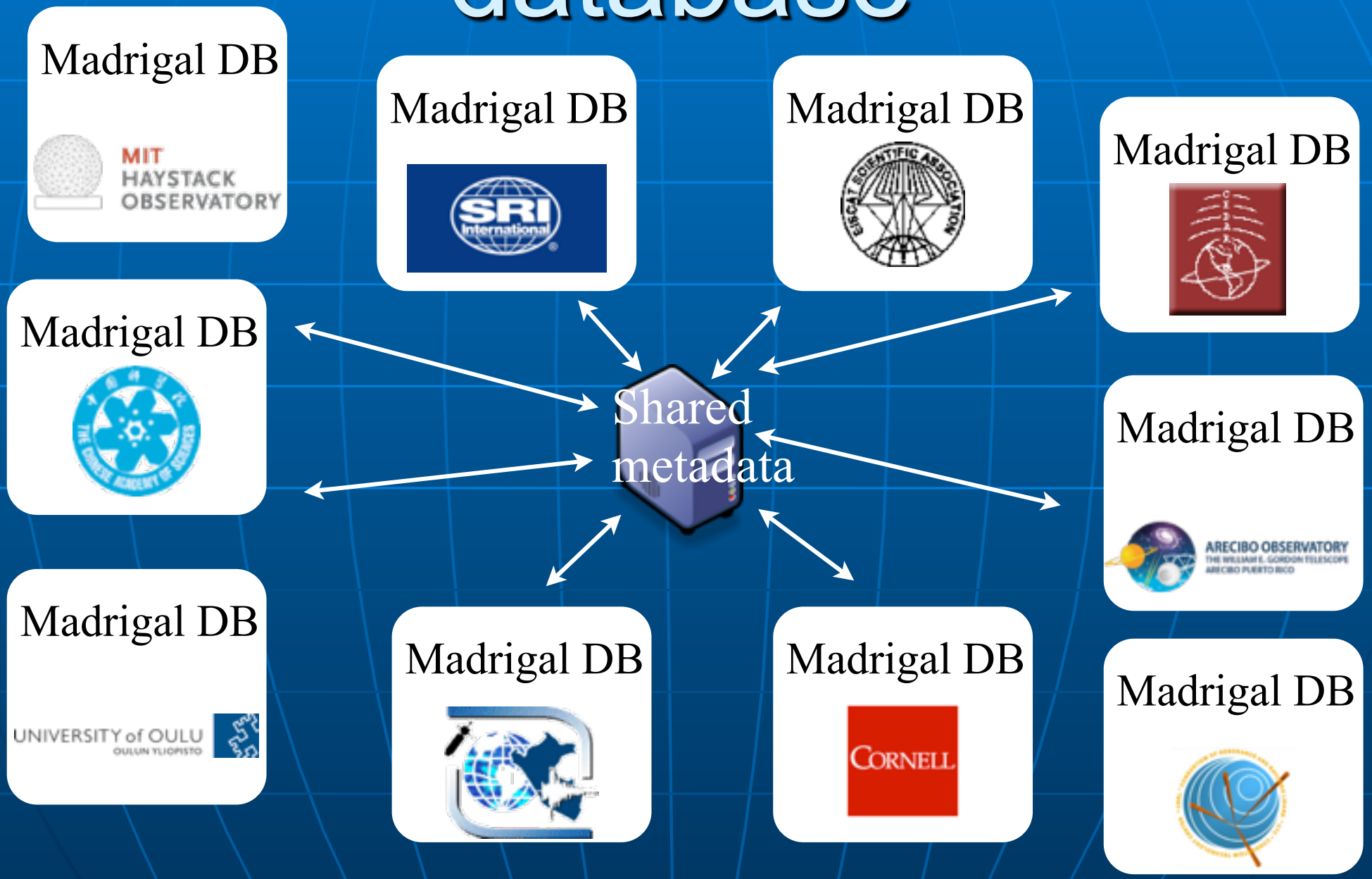
July 25, 2016

Outline

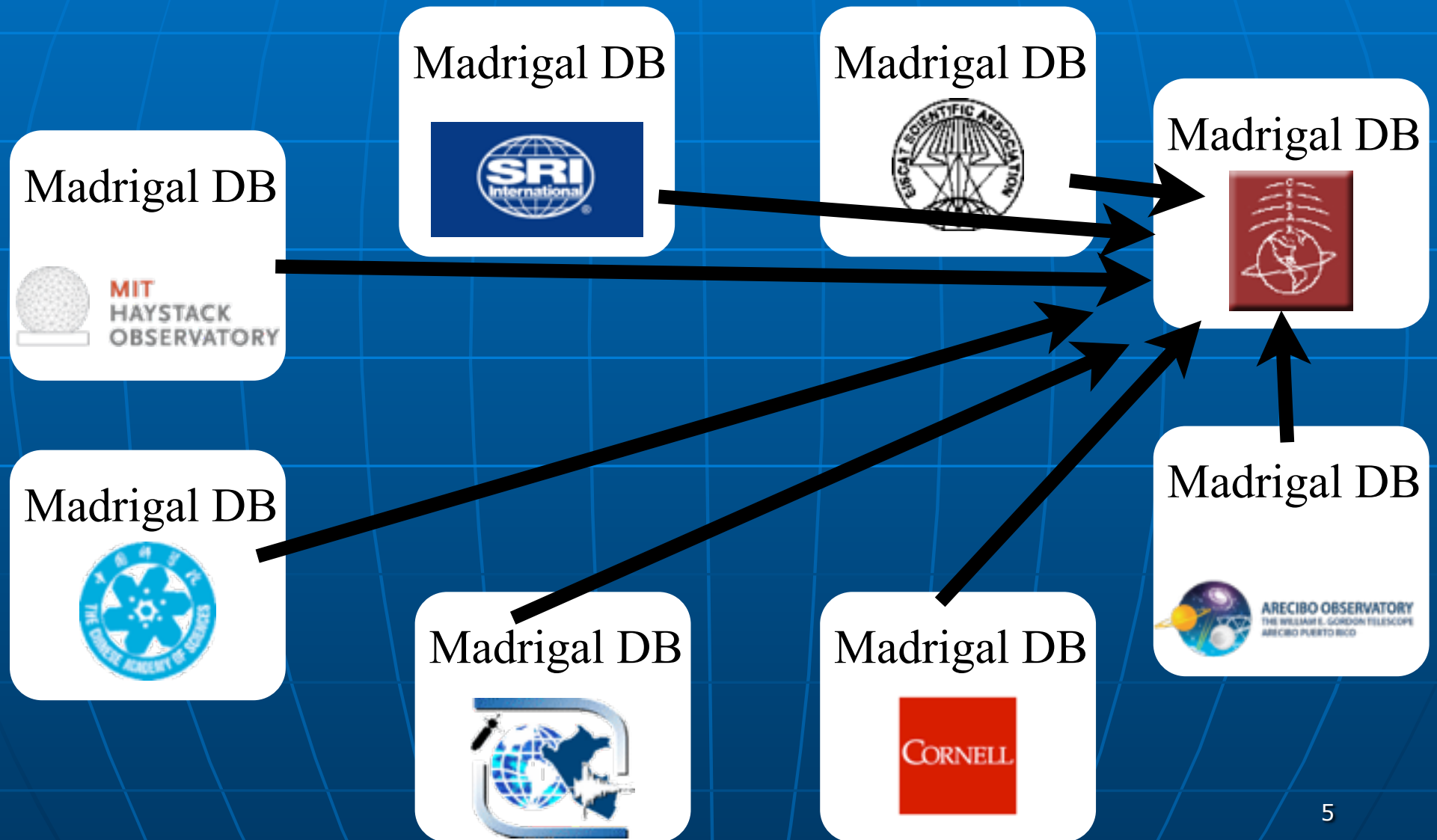
- What is Madrigal?
- What is the CEDAR database format?
- In what formats can I get Madrigal data?
- How do I use Madrigal?
 - Background
 - The website
 - Simple local data access
 - Full Access
 - Script data access
- Group exercises

What is Madrigal?

Madrigal is a distributed database



Cedar Madrigal archive imports all data weekly



The Madrigal database stores data from a wide variety of upper atmosphere research instruments

Incoherent Scatter Radar



TEC via GPS



MF Radar



Examples of number of instruments in Madrigal:

- Incoherent scatter radars: 22
- MST radars: 3
- MF radars: 16
- Meteor radars: 11
- FPI: 32
- Michelson Interferometers: 6
- Lidars: 9
- Photometers: 7

Other examples:

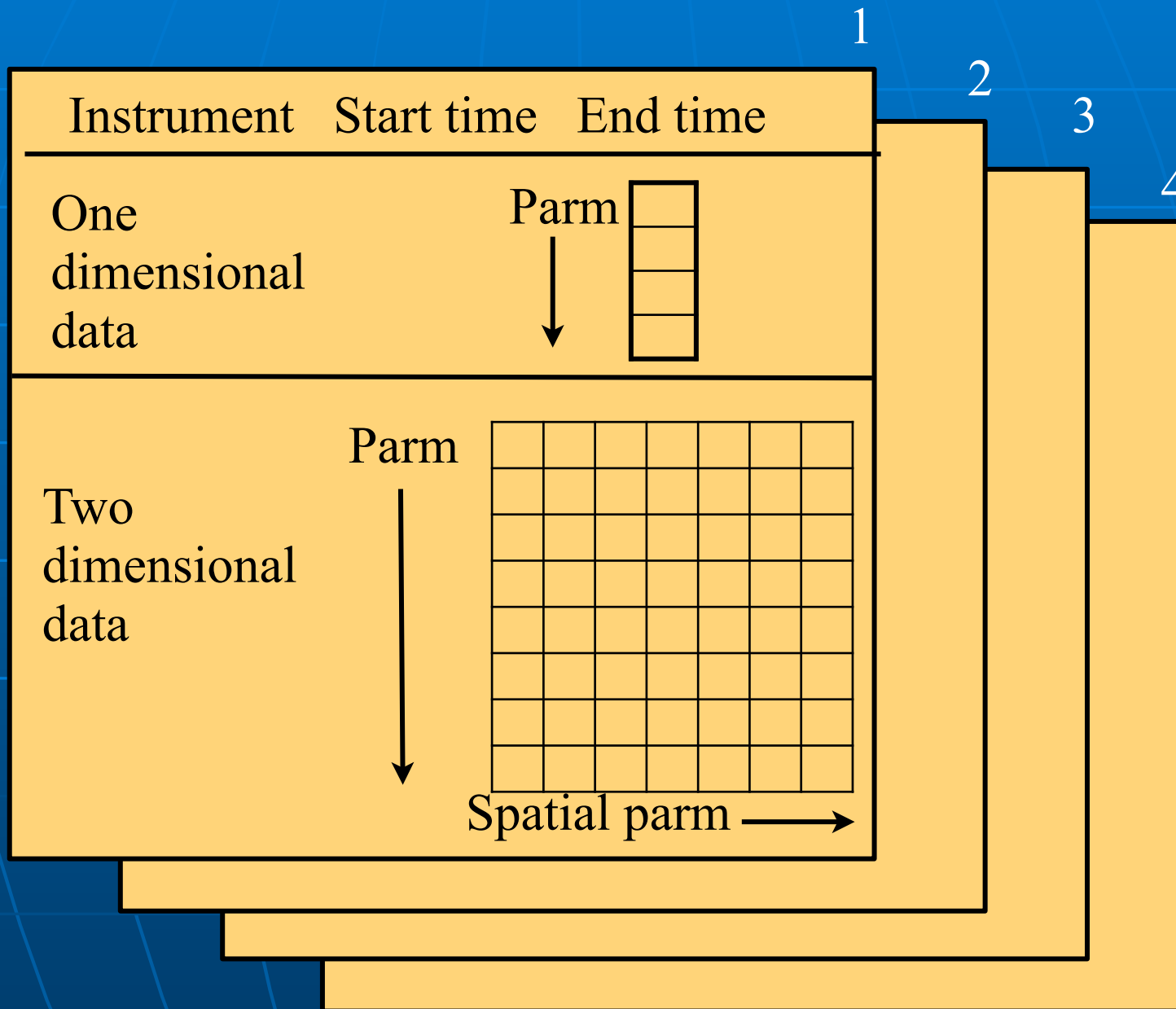
- GPS TEC
- DMSP

- What is the CEDAR database format?

Well defined parameters

- Standard descriptions of all parameters
- Allows the existence of derivation engine
- Madrigal allows extended descriptions
- All parameters have corresponding error parameters
- Missing, Assumed

Cedar file data model

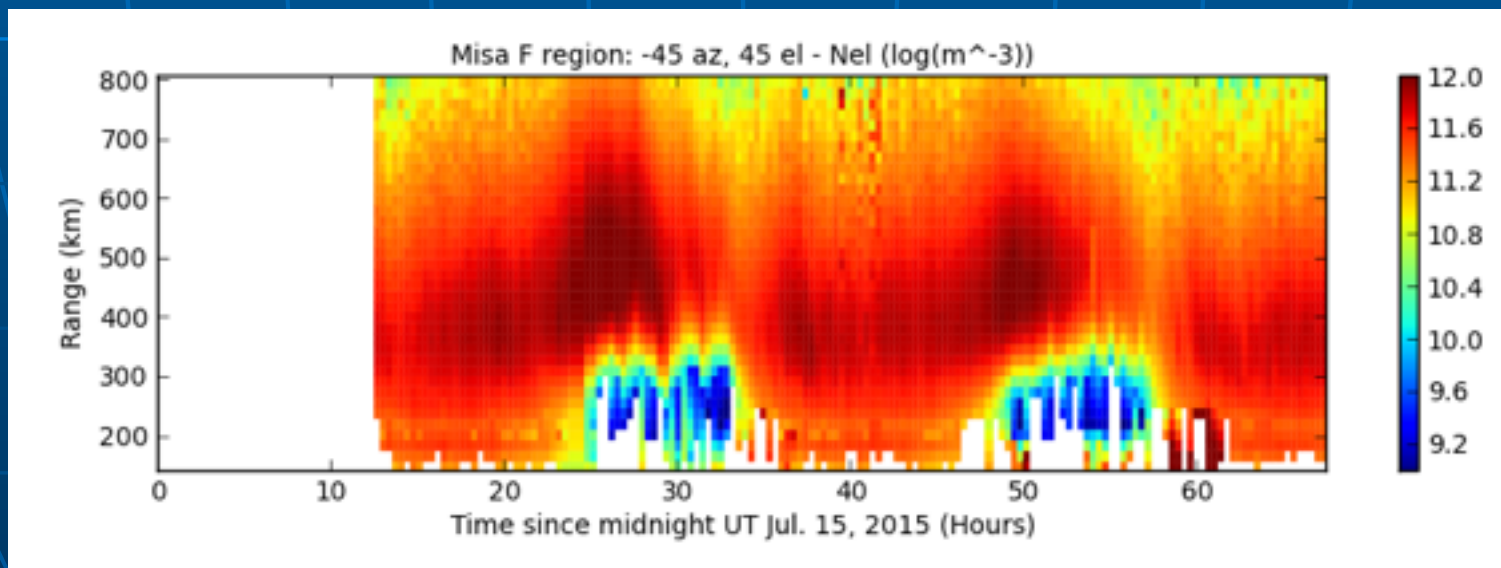


Modifications to CEDAR standard with Madrigal 3

- Independent spatial parameters now part of standard
 - Allows layout of 2D data as grid
- Parameters same from record to record
- Parameters can now be float, integer, or fixed length string

Madrigal 3 defines independent spatial parameters

- Based on Jicamarca input to Madrigal 2.6
- Allows trivial pcolor plotting
- Alternate layout in Hdf5
- Easy export to netCDF4



In what formats can I get
Madrigal data?

Cedar file format: past and near future

Cedar file format

- Developed in 1980
- 16 bit integer
 - Dynamic range problems

Hdf5

- Scientific standard
- Float based
- Flexible arrangement
- Table data, optional grid

Madrigal versions

| Release | Madrigal 2.6 (now) | Madrigal 3.0 (beta) |
|---|--------------------------------------|--------------------------------|
| Underlying format | Cedar file format | Hdf5 |
| Output formats | Ascii, Hdf5, Cedar file format | Ascii, Hdf5, netCDF4 |
| Formats with deriv parms | Ascii | Ascii, Hdf5, netCDF4 |

- What do I need to understand about Madrigal to use it?

Madrigal Data Model

Madrigal site
(typically a facility with scientists and a Madrigal installation)

Madrigal DB



Data shared among all Madrigal sites

Instruments
(ground-based, typically with a set location)



Experiments
(typically of limited duration, with a single contact)



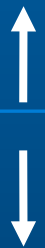
Experiment Files
(represents data from one analysis of the experiment)



Records
(measurement over one period of time)



Data unique to one Madrigal site



Madrigal Derivation Engine

- Derived parameters appear to be in file
- Engine determines all parameters that can be derived
- Easy to add new derived parameters using code written in C or Fortran



Classes of derived parameters

- Space, time
 - Examples: Local time, shadow height
- Geophysical
 - Examples: Kp, Dst, Imf, F10.7
- Magnetic
 - Examples: Bmag, Mag conjugate lat and long, Tsyganenko magnetic equatorial plane intercept
- Models
 - Examples: MSIS, IRI

- Using Madrigal

Rules of the road

- If you want to use data in Madrigal in a paper or talk, contact the PI
- The PI is listed on every web page



Use of data without informing PI may lead to seven years of bad luck in grant writing - don't let this happen to you!

How can the Madrigal database be accessed?



User



Web interface

Web services API

- From anywhere on internet
- Python API
- Matlab API
- IDL API
- Other could be written

Typical use - data discovery

Typical use - data downloading of more than one experiment - *zero clicks to data*

Live demo of Madrigal web page

- Madrigal 2
 - <http://cedar.openmadrigal.org> or <http://www.eiscat.se/madrigal/>)
- Madrigal 3 beta
 - <http://madrigal3.haystack.mit.edu>

Web interface review

The screenshot shows a web browser window titled "Madrigal Database Access" with the URL "madrigal.haystack.mit.edu/cgi-bin/madrigal/accessData.cgi". The page content includes a navigation bar with "Back to [Millstone homepage](#)", a main heading "Access Madrigal Data", and a dropdown menu for "Go to a different Madrigal site: Millstone". The main content is divided into four sections, each with a blue callout bubble:

- Simple Local Data Access**: "Print or download local Millstone data files, or view existing plots. Click [here](#) for a tutorial. Use [Browse for Individual Madrigal Experiments](#) to search all Madrigal sites at once or for advanced features like derived parameters." A callout bubble points to this section with the text: "Simpler to use, only local data, no derived parameters".
- Browse for Individual Madrigal Experiments**: "Search available experiments on either *all* Madrigal databases, or just the local Millstone database. You can choose which parameters to print, including derived the data using any parameter." A callout bubble points to this section with the text: "Search all Madrigal sites, any parameter, more complex".
- Global Madrigal Search**: "Generate a command that runs on your local computer that can search multiple local Millstone experiments at once. The data can be filtered, derived parameters are allowed, and data matching your criteria. Click [here](#) for a tutorial." A callout bubble points to this section with the text: "Generates scripts - to be discussed".
- Plot Data from Instruments**: "Create new plots from one or more instruments and/or Madrigal experiments versus time on a single web page. The data comes from the local Millstone database. Click [here](#) for a tutorial." A callout bubble points to this section with the text: "Generates scripts - to be discussed".

Group exercise 1

- See tinyurl.com/2016ISR -> Madrigal Exercise 1

Remote Access to Madrigal Data

- Built on web services
- Like the web, available from anywhere on any platform
- Read only API
- Complete Python, Matlab, and IDL APIs written
- More APIs available on request or via contribution

Remote Access to Madrigal Data

Simple

Most complex



I just want to download lots of files to my PC



globalDownload



I want to only download filtered data with derived parameters



globalIsprint

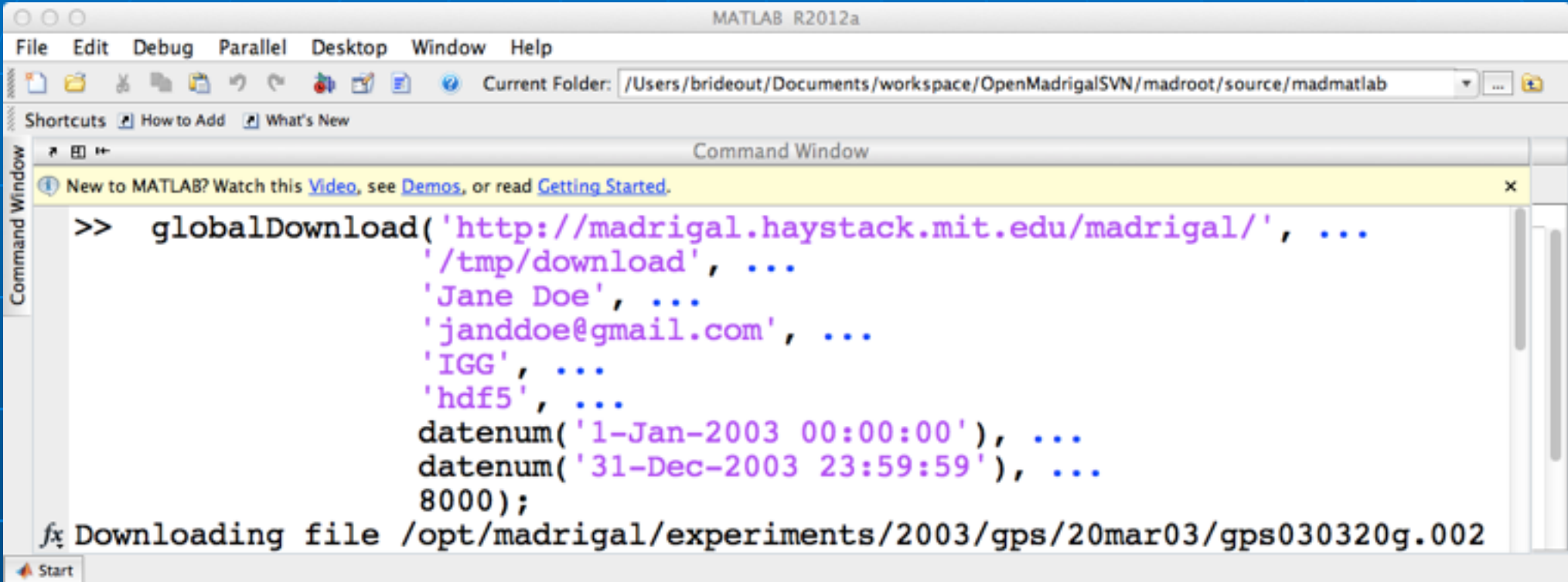


I want to write my own script



Work with API's

Matlab globalDownload does it with one command!



The screenshot shows the MATLAB R2012a Command Window. The current folder is `/Users/brideout/Documents/workspace/OpenMadrigalSVN/madroot/source/madmatlab`. The Command Window contains the following command and output:

```
>> globalDownload('http://madrigal.haystack.mit.edu/madrigal/', ...  
                '/tmp/download', ...  
                'Jane Doe', ...  
                'janddoe@gmail.com', ...  
                'IGG', ...  
                'hdf5', ...  
                datenum('1-Jan-2003 00:00:00'), ...  
                datenum('31-Dec-2003 23:59:59'), ...  
                8000);  
f: Downloading file /opt/madrigal/experiments/2003/gps/20mar03/gps030320g.002
```

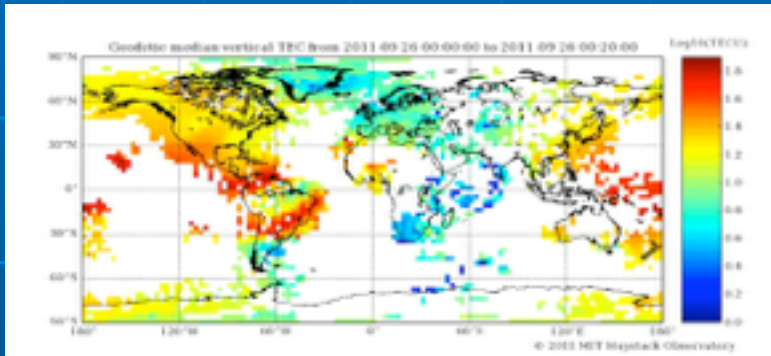
- New release as of May 3, 2014
- Uses `wget` if available, unreliable `matlab urlread` if not
- Full documentation: http://madrigal.haystack.mit.edu/madrigal/rr_matlab.html#globalDownload

Python version

```
>> globalDownload.py --url= http://madrigal.iggcas.ac.cn/madrigal/ \  
--outputDir=/tmp --user_fullname="Jane Doe" \  
--user_email=janedoe@gmail.com --user_affiliation=IGG \  
--format=hdf5 --startDate=01/01/2003 \  
--endDate=-01/30/2003 --inst=8000
```

- No need to understand python
- Install from <http://madrigal.haystack.edu/madrigal/madDownload.html>
- Formats: 'hdf5', 'ascii' - 'hdf5' faster
- Full documentation: <http://madrigal.haystack.mit.edu/madrigal/madpyDoc/remotePythonAPI/madrigalWeb/globalDownload.py.html>

More advanced globallsprint scripts

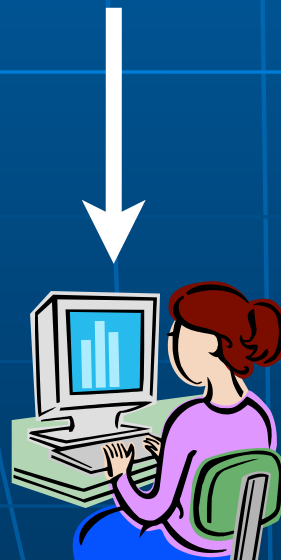


TEC data: instrument id = 8000

Madrigal DB



<http://madrigal.iggeas.ac.cn/madrigal/>



Show me only data
where $TEC > 100TEC_u$
when $Kp > 7$ in 2003

The web interface will generate the script for you

Back to [Millstone homepage](#)

Access Madrigal Data

Go to a different Madrigal site:

Simple Local Data Access

Print or download local Millstone data files, or view existing plots. Click [here](#) for a tutorial. Use [Browse for Individual Madrigal Experiments](#) to search all Madrigal sites at once or for advanced features like derived parameters.

Global Madrigal Search

Generate a command that runs on your local computer that can search multiple local Millstone experiments at once. The data can be filtered, derived parameters are allowed, and data matching your criteria will be returned in a single ascii report. Click [here](#) for a tutorial.

Browse for Individual Madrigal Experiments

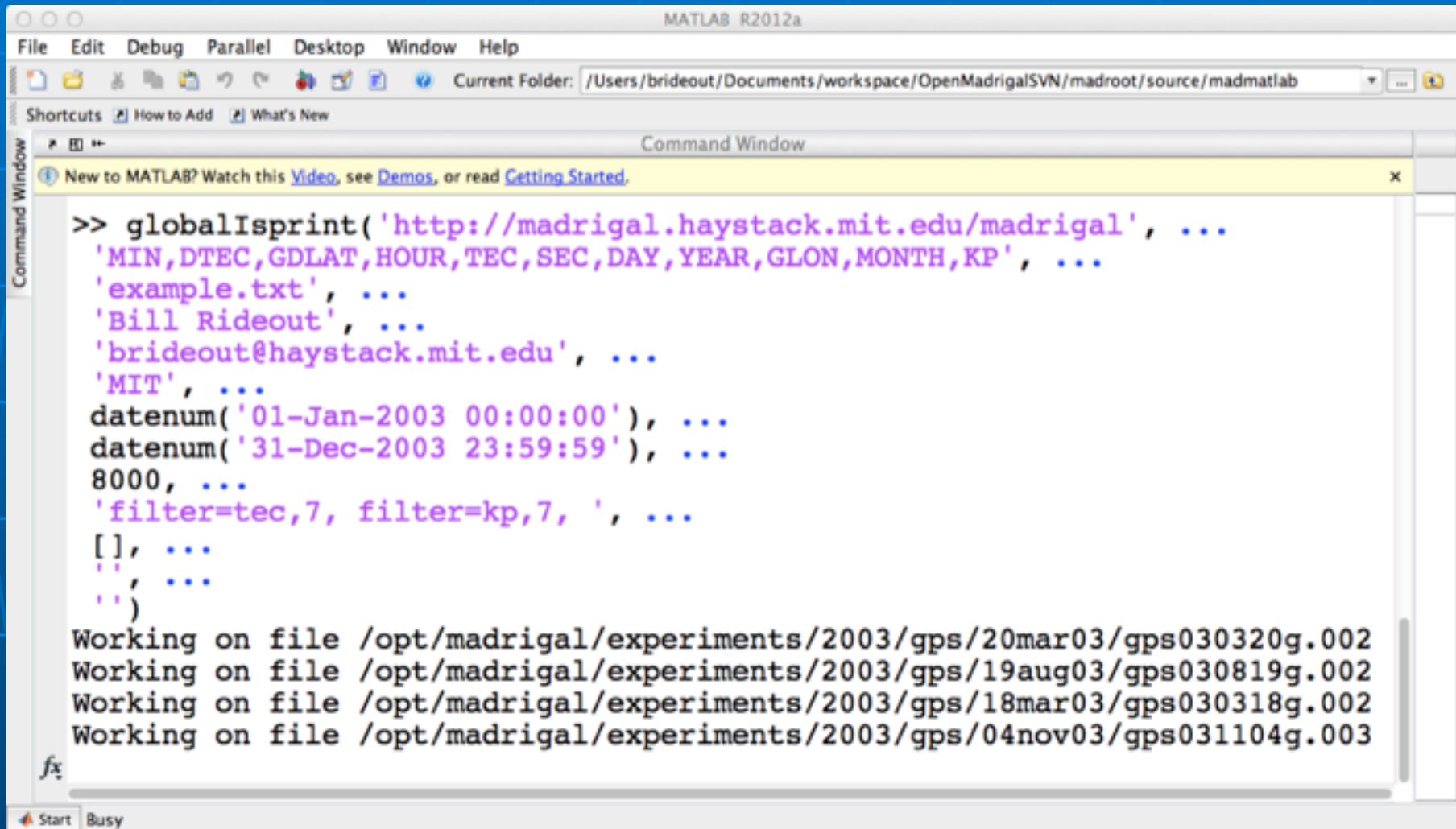
Search available experiments on either *all* Madrigal databases, or just the local Millstone database. You can choose which parameters to print, including derived parameters, and can filter the data using any parameter. Click [here](#) for a tutorial.

Plot Data from Instruments

Create new plots from one or more instruments and/or Madrigal experiments versus time on a single web page. The data comes from the local Millstone database. Click [here](#) for a tutorial.

Demo for example above

Matlab globalIsprint example



The screenshot shows the MATLAB R2012a Command Window. The current folder is `/Users/brideout/Documents/workspace/OpenMadrigalSVN/madroot/source/madmatlab`. The Command Window displays the following code and output:

```
>> globalIsprint('http://madrigal.haystack.mit.edu/madrigal', ...  
'MIN,DTEC,GDLAT,HOUR,TEC,SEC,DAY,YEAR,GLON,MONTH,KP', ...  
'example.txt', ...  
'Bill Rideout', ...  
'brideout@haystack.mit.edu', ...  
'MIT', ...  
datenum('01-Jan-2003 00:00:00'), ...  
datenum('31-Dec-2003 23:59:59'), ...  
8000, ...  
'filter=tec,7, filter=kp,7, ', ...  
[], ...  
...  
...  
)  
Working on file /opt/madrigal/experiments/2003/gps/20mar03/gps030320g.002  
Working on file /opt/madrigal/experiments/2003/gps/19aug03/gps030819g.002  
Working on file /opt/madrigal/experiments/2003/gps/18mar03/gps030318g.002  
Working on file /opt/madrigal/experiments/2003/gps/04nov03/gps031104g.002
```

- Uses wget if available, unreliable matlab urlread if not
- Full documentation: http://madrigal.haystack.mit.edu/madrigal/rr_matlab.html#globalIsprint

Python example

```
>> globalIsprint.py --url=http://madrigan.haystack.mit.edu/madrigan \  
--parms=year,month,day,hour,min,sec,gdlat,glon,tec,dtec,kp \  
--output=/tmp/gps.txt --user_fullname="Jane Doe" \  
--user_email=janedoe@gmail.com --user_affiliation=IGG --inst=8000 \  
--startDate=01/01/2003 --endDate=12/31/2003 \  
--filter=kp,7, --filter=tec,100,
```

- You pick the output parameters
- Install from <http://madrigan.haystack.edu/madrigan/madDownload.html>
- Only ascii output - all to one file
- Full documentation: <http://madrigan.haystack.mit.edu/madrigan/madpyDoc/remotePythonAPI/madriganWeb/globalIsprint.py.html>

Most complex - write your own Matlab script

- Methods
 - getInstrumentsWeb
 - getExperimentsWeb
 - getExperimentFilesWeb
 - getParametersWeb
 - isprintWeb
 - madDownloadFile
 - madCalculatorWeb
- Methods match Madrigal model

Simple Matlab example

```
filename = '/usr/local/madroot/experiments
           /2003/tro/05jun03/NCAR_2003-06-05_tau2pl_60_uhf.bin';

eiscat_cgi_url = 'http://www.eiscat.se/madrigal/cgi-bin/';

% download the following parameters from the above file: ut, gdalt, ti

parms = 'ut,gdalt,ti';

filterStr = 'filter=gdalt,200,600 filter=ti,0,5000';

% returns a three dimensional array of double with the dimensions:
%
% [Number of rows, number of parameters requested, number of records]
%
% If error or no data returned, will return error explanation string instead.
data = isprintWeb(eiscat_cgi_url, filename, parms, filterStr);
```

**Matlab
Madrigal
API call**



■ Matlab API Links

- Reference: http://madrigal.haystack.mit.edu/madrigal/rr_matlab.html
- Tutorial: http://madrigal.haystack.mit.edu/madrigal/rt_matlab.html

Example python script

```
In [1]: import madrigalWeb.madrigalWeb
```

The next command connects to one particular Madrigal site, and creates an object we will use for the rest of the session. Because there are some new features in Madrigal 3, we connect to that site.

```
In [12]: madDB = madrigalWeb.madrigalWeb.MadrigalData('http://madrigal3.haystack.mit.edu')
```

The top level of metadata at any site is the instruments it has available. This next method lists the available instruments.

```
In [13]: instList = madDB.getAllInstruments()
```

```
In [14]: print(instList[7])
```

```
name: Millstone Hill IS Radar
code: 30
mnemonic: mlh
latitude: 42.619
longitude: 288.51
altitude: 0.146
category: Incoherent Scatter Radars
```

See <http://tinyurl.com/2016ISR> to download full notebook

Python Remote API

- Can run on any platform with python (PC, Unix, Mac, etc)
- Fully documented with examples
- Links
 - **Reference:** <http://madrigan.haystack.mit.edu/madrigan/madpyDoc/remotePythonAPI/index.html>
 - **Tutorial:** http://madrigan.haystack.mit.edu/madrigan/rt_python.html

IDL Remote API

- Methods
 - madGetAllInstruments
 - madGetExperiments
 - madGetExperimentFiles
 - madGetExperimentFileParameters
 - madSimplePrint
 - madPrint
 - madDownloadFile
 - madCalculator
- Methods again match Madrigal model

Madrigal Web Services

- Simple delimited output via CGI scripts
- Not based on SOAP or XmlRpc since no support in languages such as Matlab
- CGI arguments and output fully documented at <http://www.haystack.edu/madrigal/remoteAPIs.html>

Extending/contributing to Madrigal

- Madrigal is completely open source
- See www.openmadrigal.org for CVS
- All new code is Python or C. Imported derivation methods sometimes in Fortran.
- I appreciate all contributions
 - Suggestions and ideas
 - Finding bugs
 - Code

Group exercises

- Break into your groups
- See tinyurl.com/2016ISR
 - Web interface - Exercise 1
 - Script interface - Exercise 2