

How to use the Madrigal database for atmospheric science

Bill Rideout

MIT Haystack Observatory

brideout@haystack.mit.edu

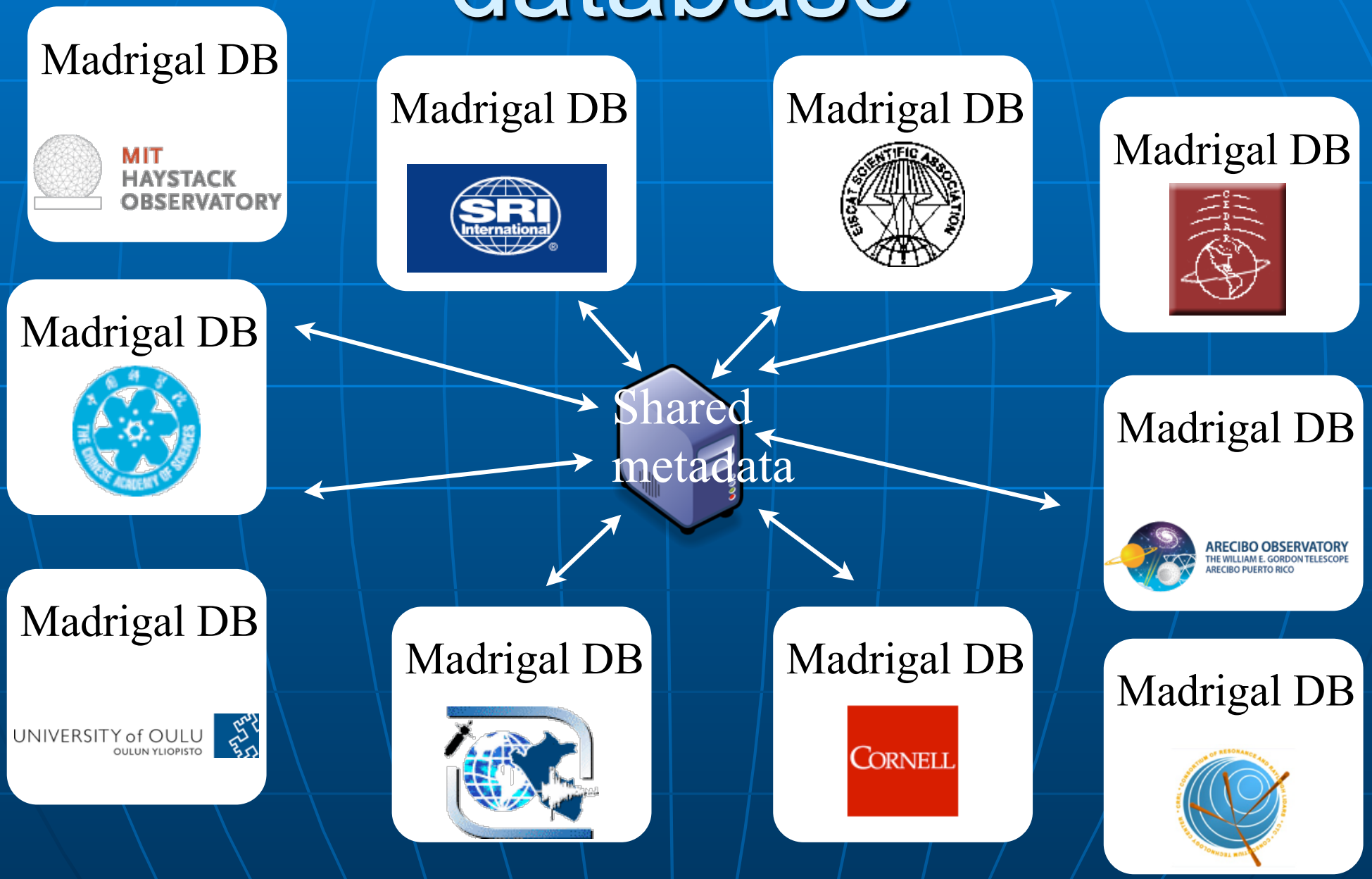
ISR workshop
Arecibo Observatory
July 24, 2017

Outline

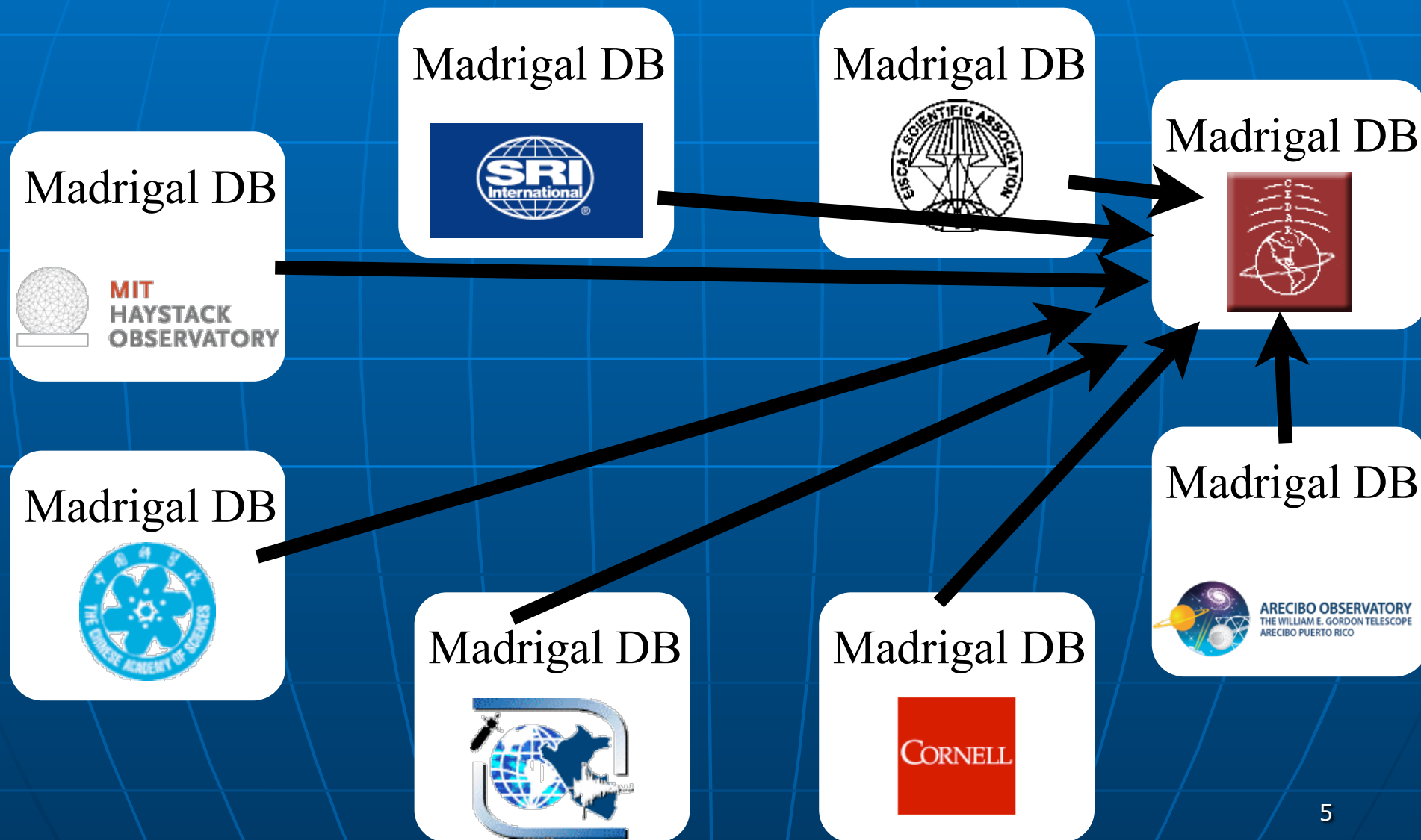
- What is Madrigal?
- What is the CEDAR database format?
- In what formats can I get Madrigal data?
- How do I use Madrigal?
 - Background
 - The website
 - Madrigal 2 (production)
 - Madrigal 3 (being released)
 - Script data access
- Group exercises

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

Madrigal is open-source

Madrigal Database

http://www.openmadrigal.org/

The Open Madrigal Initiative

- [What is Madrigal?](#)
- [Download/update Madrigal](#) - includes Madrigal server and client APIs
- [Documentation](#)
 - [Web access](#)
 - [Script access](#)
- [Empirical Ionospheric Models](#)
- [Subversion Source Control](#)
- [Mailing Lists](#)
- [Administering OpenMadrigal](#)

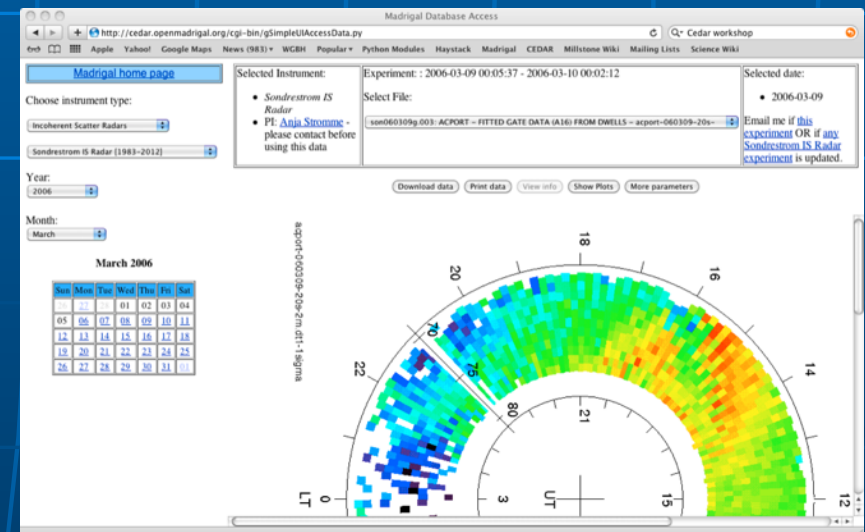
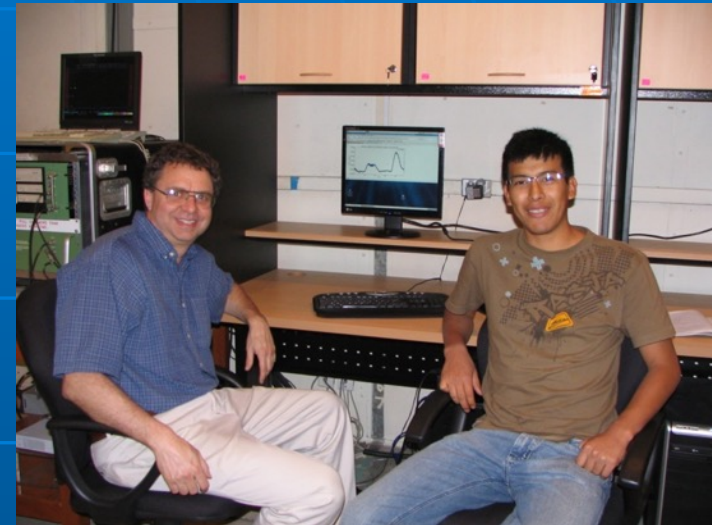
The OpenMadrigal project seeks to develop and support an on-line database for geospace data. The project has been led by [MIT Haystack Observatory](#) since 1980, but now has active support from [Jicamarca Observatory](#) and other community members. Madrigal is a robust, World Wide Web based system capable of managing and serving archival and real-time data, in a variety of formats, from a wide range of ground-based instruments. Madrigal is installed at a number of sites around the world. Data at each Madrigal site is locally controlled and can be updated at any time, but shared metadata between Madrigal sites allow searching of all Madrigal sites at once from any Madrigal site.

Madrigal is a robust, World Wide Web based system capable of managing and serving archival and real-time data, in a variety of formats, from a wide range of instruments. Data can be accessed from the Madrigal sites at [Millstone Hill, USA](#), [EISCAT, Norway](#), [SRI International, USA](#), [Arecibo, Puerto Rico](#), [Cornell University, USA](#), [Jicamarca, Peru](#), the [Institute of Geology and Geophysics](#), the Chinese Academy of Sciences, and the [CEDAR Madrigal archive](#) using standard Web browsers; and directly, using APIs which are available for python, Matlab, and IDL.



Suggestions and comments should be directed to madrigal@haystack

Link to Subversion (source code)

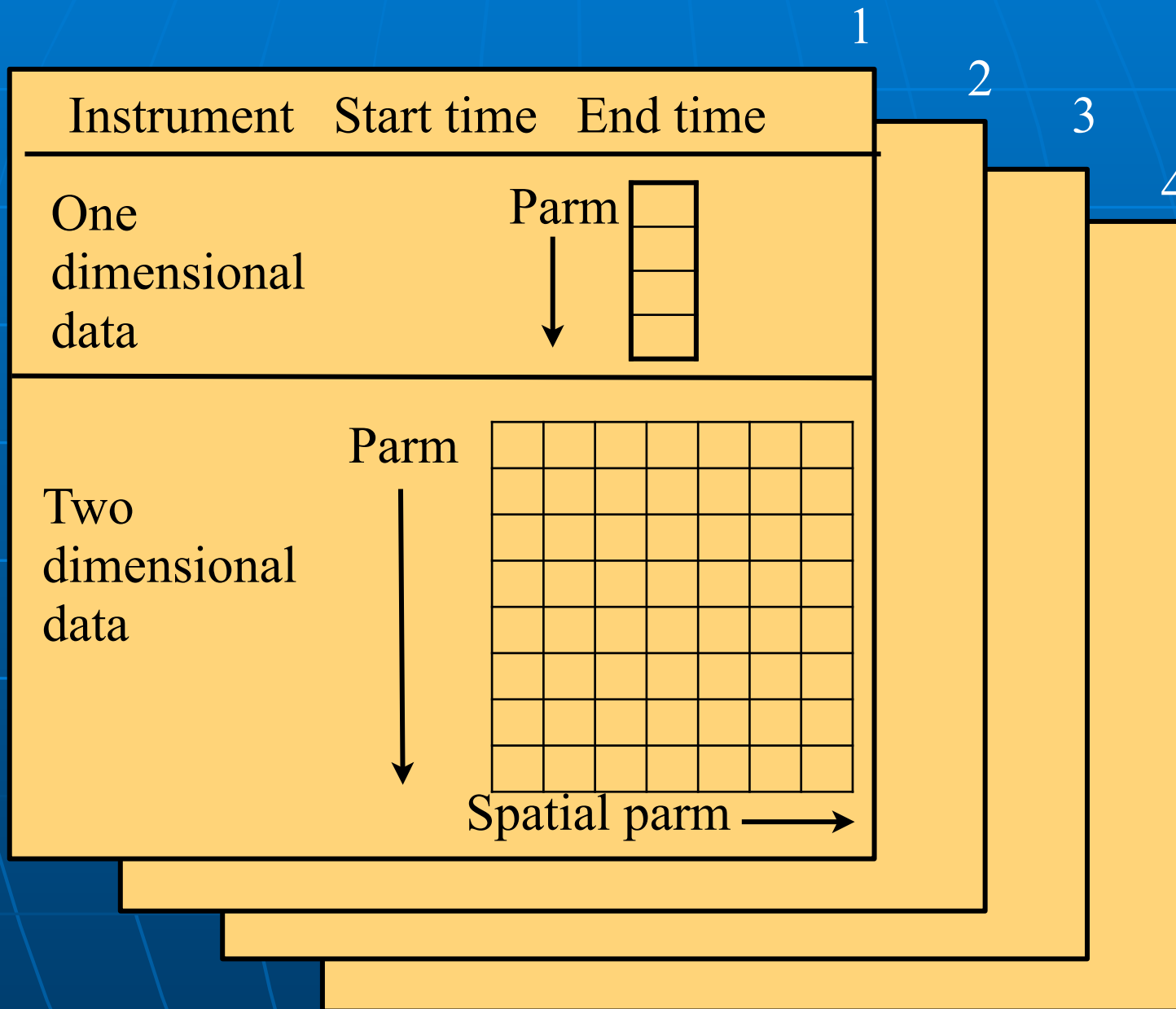


- What is the CEDAR database format?

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

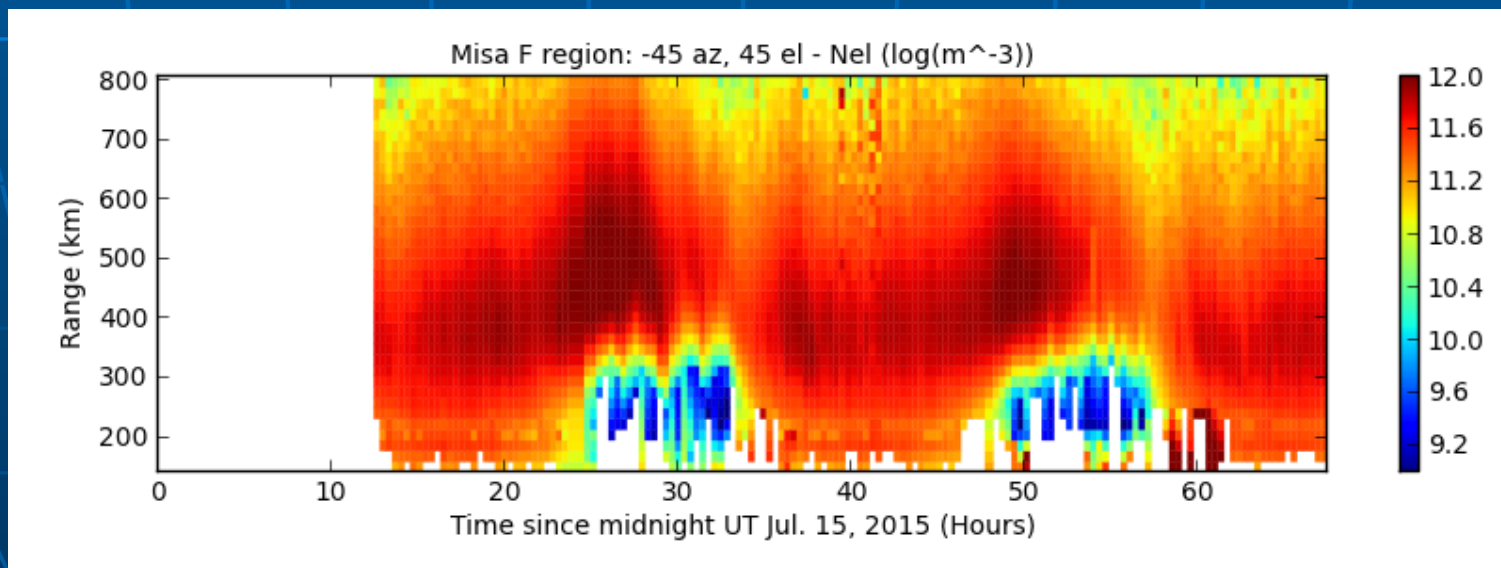


Modifications to CEDAR standard with Madrigal 3

- 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
- Automatic export to netCDF4



In what formats can I get
Madrigal data?

Cedar file format: past and very near future

Cedar file format

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

Hdf5

- Scientific standard
- Float based
- Flexible arrangement
- Table data, grid if any ind parms
- Self-describing

Madrigal versions

Release	Madrigal 2.6 (now)	Madrigal 3.0
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 many experiments at a time - *zero clicks to data*

Live demo of Madrigal web page

- Madrigal 2

- <http://cedar.openmadrigal.org> or <http://www.eiscat.se/madrigal/>)

- Madrigal 3

- <http://madrigal3.haystack.mit.edu>

Group exercise 1

- See tinyurl.com/2017ISR -> 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 (python 2 and 3)
- 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

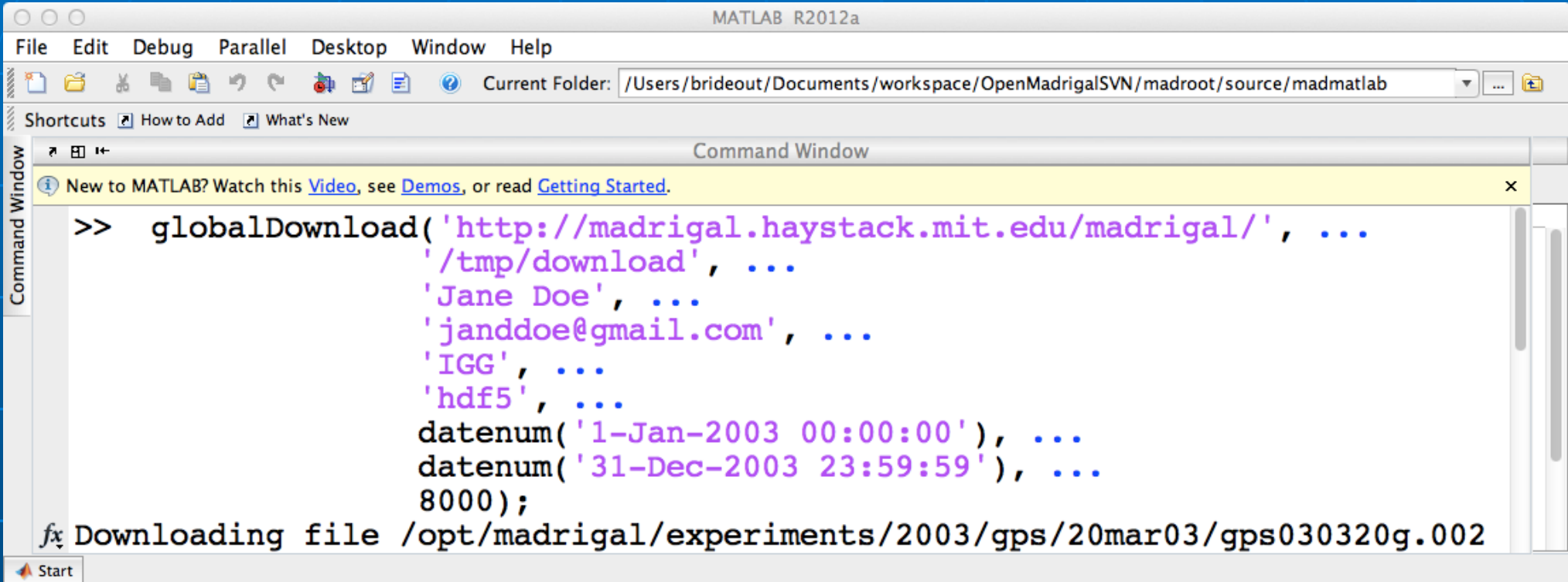


I want to write my own script



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);  
fx Downloading file /opt/madrigal/experiments/2003/gps/20mar03/gps030320g.002
```

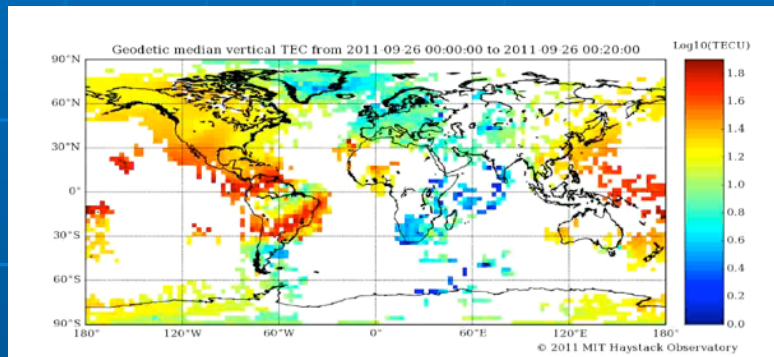
- New release as of May 3, 2014
- Uses `wget` if available, unreliable `matlab urlread` if not
- Full documentation: http://madrigal.haystack.mit.edu/madrigal/rr_matlab.html#globalDownload

Python version

```
>> globalDownload.py --url= http://madrigal.igggcas.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

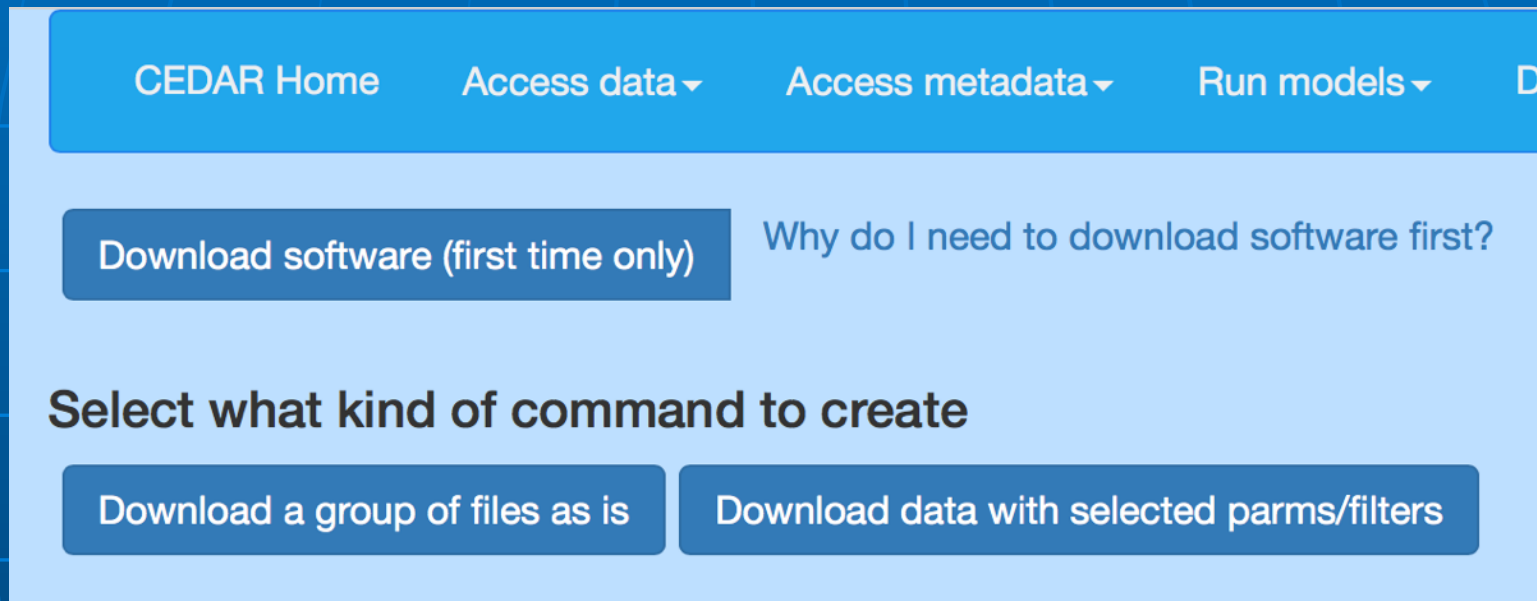


<http://madrigal.iggeas.ac.cn/madrigal/>



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

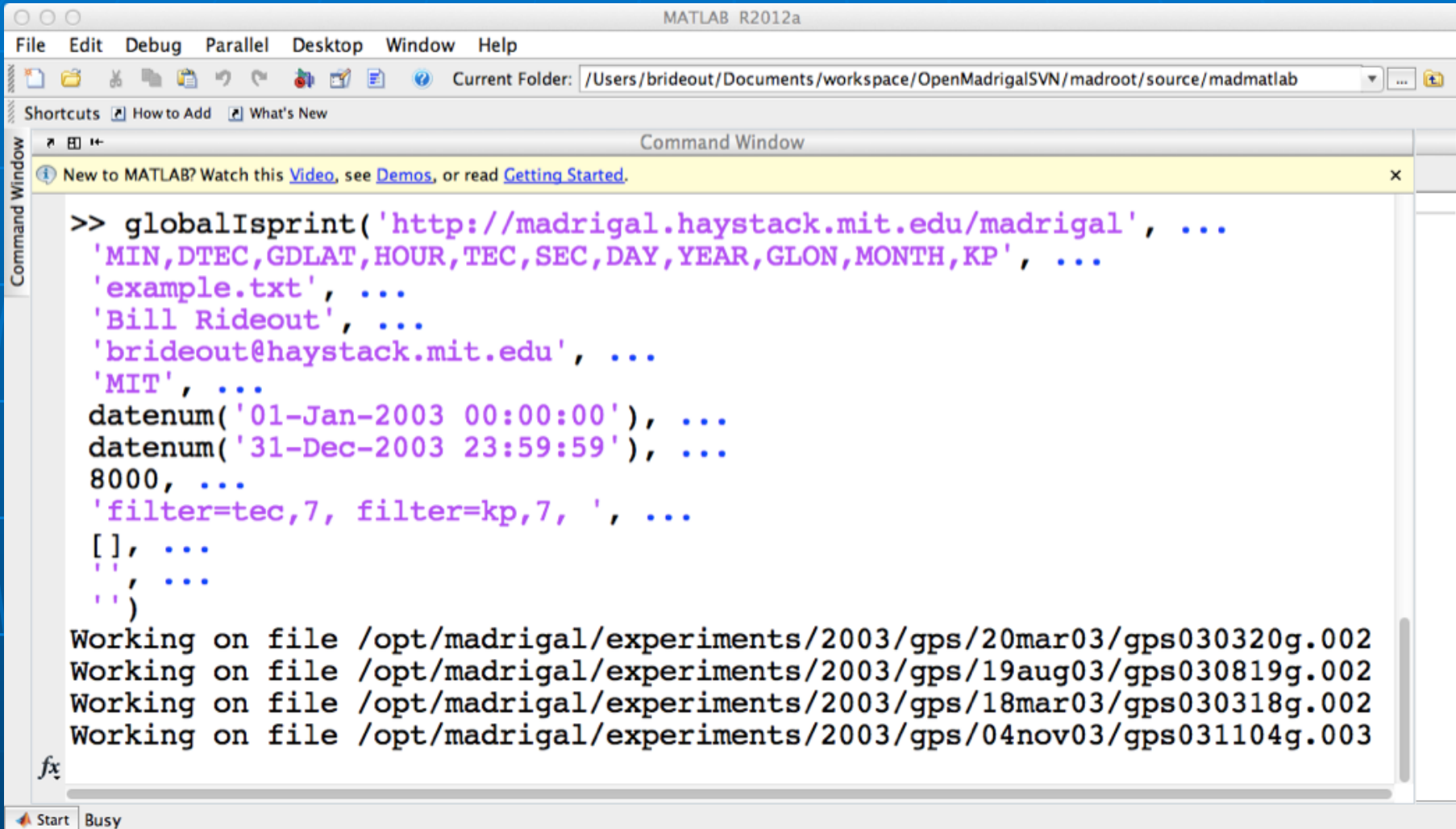
The web interface will generate the script for you



The screenshot shows a web interface with a blue header containing navigation links: CEDAR Home, Access data (with a dropdown arrow), Access metadata (with a dropdown arrow), Run models (with a dropdown arrow), and a partially visible 'D'. Below the header, there is a dark blue button labeled 'Download software (first time only)' and a link 'Why do I need to download software first?'. A section titled 'Select what kind of command to create' contains two buttons: 'Download a group of files as is' and 'Download data with selected parms/filters'.

Demo for example above

Matlab globalIsprint example



The screenshot shows the MATLAB R2012a Command Window. The title bar reads "MATLAB R2012a". The menu bar includes "File", "Edit", "Debug", "Parallel", "Desktop", "Window", and "Help". The current folder is "/Users/brideout/Documents/workspace/OpenMadrigalSVN/madroot/source/madmatlab". A yellow banner at the top of the Command Window says "New to MATLAB? Watch this [Video](#), see [Demos](#), or read [Getting Started](#)." The Command Window contains the following code and output:

```
>> globalIsprint('http://madrigal.haystack.mit.edu/madrigal', ...  
'MIN,DTEC,GDLAT,HOURL,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

Most complex - write your own Matlab script

- Methods

- getInstrumentsWeb
- getExperimentsWeb
- getExperimentFilesWeb
- getParametersWeb
- isprintWeb
- madDownloadFile
- madCalculatorWeb

- Methods match Madrigal model

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

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

How can I put my instrument's data on Madrigal?



Send data to CEDAR Madrigal

Set up your own Madrigal site

Madrigal DB



- Send data to MIT Haystack in your own format
- Loading program written by MIT Haystack, verified by you
- Add new data in batch or via automated upload (eg, sftp, web access, etc)

- MIT Haystack will help with installation and writing needed loading programs
- You control when data uploaded
- Automated backup to central CEDAR Madrigal site³⁶