Autoplot: a Browser For Science Data on the Web

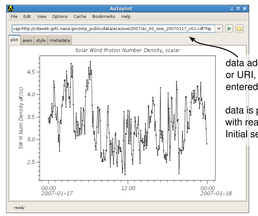
Faden, J (1,3) faden@cottagesystems.com
Weigel, R S (2)
West, E E (3)
Merka, J (4,5)

1 Cottage Systems, 1141 E. Court St., Iowa City, IA 52240, United States
2 George Mason University, 4400 University Drive, Fairfax, VA 22030, United States
3 The University of Iowa, 210 Van Allen Hall, Iowa City, IA 52242, United States

Abstract

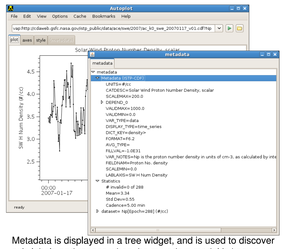
Autoplot (www.autoplot.org) is software for plotting data from many different sources and in many different file formats. Data from CDF, CEF, Fits, NetCDF, and OpenDAP can be plotted, along with many other sources such as ASCII tables and Excel spreadsheets. This is done by adapting these various data formats and APIs into a common data model that borrows from the netCDF and CDF data models. Autoplot uses a web browser metaphor to simplify use. The user specifies a parameter URL, for example a CDF file accessible via http with a parameter name appended, and the file resource is downloaded and the parameter is rendered in a scientifically meaningful way. When data span multiple files, the user can use a file name template in the URL to aggregate (combine) a set of remote files. So the problem of aggregating data across file boundaries is handled on the client side, allowing simple web servers to be used. The das2 graphics library provides rich controls for exploring the data. Scripting is supported through Python, providing not just programmatic control, but for calculating new parameters in a language that will look familiar to IDL and Matlab users. Autoplot is Java-based software, and will run on most computers without a burdensome installation process. It can also used as an applet or as a servlet that serves static images. Autoplot was developed as part of the Virtual Radiation Belt Observatory (ViRBO) project, and is also being used for the Virtual Magnetospheric Observatory (VMO). It is expected that this flexible, general-purpose plotting tool will be useful for allowing a data provider to add instant visualization capabilities to a directory of files or for general use in the Virtual Observatory environment.

Autoplot plots data from many different sources using an interface that is familiar and simple.

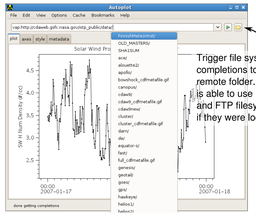


data address, or URI, is entered here

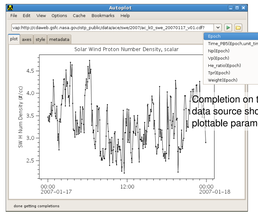
data is plotted with reasonable Initial settings



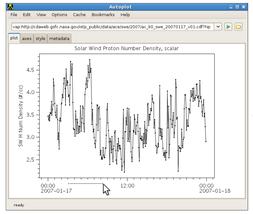
Metadata is displayed in a tree widget, and is used to discover axis labels and ranges when the metadata model is known. ISTP-CDF is the model recognized here.

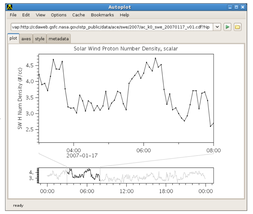


Trigger file system completions to list the remote folder. Autoplot is able to use HTTP and FTP filesystems as if they were local.



Completion on the CDF data source shows plottable parameters



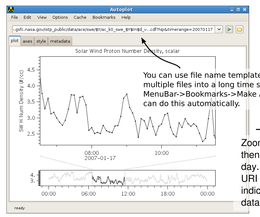


Browsing the data is fast and intuitive. Drag out a range on an axis to zoom in on the axis. The mouse wheel works as in Google maps: press the mouse wheel and drag to pan the axis. Spin the mouse wheel to zoom in and out.

A context overview plot shows the new setting in context.

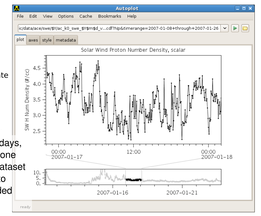
Autoplot is different from other analysis systems

* simple, general-purpose interface increases acceptance and broadens utility
* Java Webstart makes installation and maintenence trivial
* remote filesystem layer plus aggregation templates provide excellent browsing using a simple server
* handles data at many levels of abstraction, from array of numbers to timeseries with physical units
* all data is converted to standard, uniform model
* anti-aliased fonts and graphics improves ledgibility and appearance
* highly interactive graphics components
* extensible plug-in data sources
* plug-ins output digital data to Excel, CDF, many others to come
* Jython scripting allows derivation of new datasets, complex applications, and programmatic control
* separation of Model and Controller code from View allows for use on the server-side as well as client-side
* open-source, and building procedure documented from SVN checkout to invocation

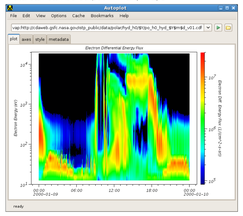


You can use file name templates to aggregate multiple files into a long time series. MenuBar->Bookmarks->Make Aggregation can do this automatically.

Zoom out to 15 days, then zoom in to one day. Note the dataset URI is updated to indicate the loaded dataset



Data can can be output in many ways



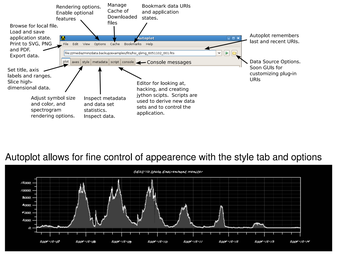
Render to SVG, PNG, and PDF

Save Autoplot configuration to .vap file

Copy and Paste image into email or PowerPoint

Export Data to ASCII table, Excel Spreadsheet, CDF, Binary table

Autoplot with all of its features enabled. Tabs can be undocked allowing the data and metadata to be viewed at the same time.



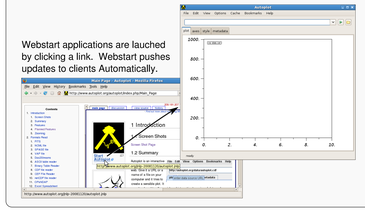
Autoplot allows for fine control of appearence with the style tab and options

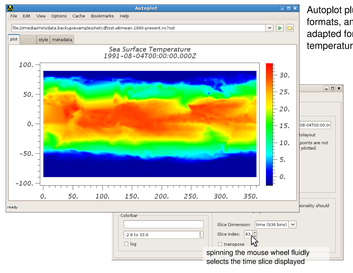
Launching Autoplot

Autoplot is released using Sun Microsystems' Java Webstart and runs on Java 1.5. When webstart is properly configured, Autoplot is launched by clicking a link on a web site. Most desktops have Java installed already and can run webstart applications. The production version of Autoplot is linked at the top of www.autoplot.org.

Java can be downloaded freely from Sun (or Apple), and instructions for configuring webstart can be found at www.autoplot.org.

Webstart applications are lauched by clicking a link. Webstart pushes updates to clients Automatically.

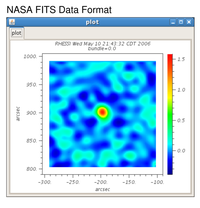




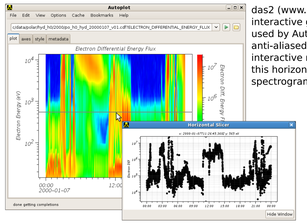
Autoplot plug-ins allow it to access many different forms of data. Plug-ins adapt file formats, and also interfaces like OpenDAP. Here, the Java NetCDF code is adapted for Autoplot so that NetCDF files can be read in. (Note the sea surface temperature is interpolated over the continent, this will be fixed soon!)

spinning the mouse wheel fluidly selects the time slice displayed

NASA FITS Data Format



das2 (www.das2.org) is the interactive graphics library used by Autoplot. It provides anti-aliased graphics and interactive mouse actions, like this horizontal slice of the spectrogram.



Integration with the VxOs

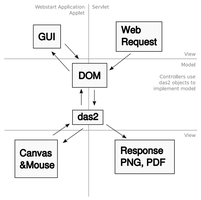
ViRBO: TSDS Server
* Specialized plugin written to request data at appropriate resolution from the TSDS server
* uses server side Autoplot to generate preview graphics
* Lightweight applet contains just one plugin to browse TSDS

ViRBO: VxOWare Server
* netCDF files are staged on the server, then Autoplot is launched with dynamic jnlp server so that webstart launches Autoplot and the resource is plotted.

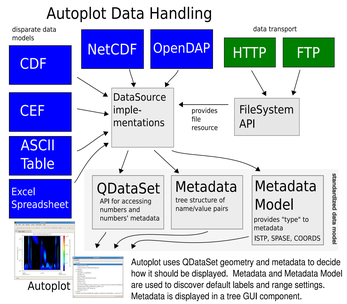
VMO:
* references to CDF files are plugged in by dynamic jnlp server so that webstart launches Autoplot with the CDF URI pre-loaded. CDF completions display the list of plottable parameters

DataShop:
* Specialized plugin quickly developed to use data from the server.

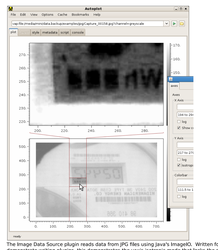
Model-View-Controller Architecture makes Autoplot well-suited for server side applications as well. An example servlet is provided that shows how Autoplot and Tomcat make a high-performance graphics server.



Autoplot uses Data Source Plug-ins to access new data sources. The plug-ins standardize the data and metadata, but also provide the schema for creating valid URIs. Data sources are registered by extension and optionally mime-type. A URI can explicitly request a data source by using "vap+ext:" as the uri scheme. Some plug-ins also provide formatting, adding another option to the export data dialog. You can look at the list of loaded plug-ins by entering "about:plugins" in the address bar.

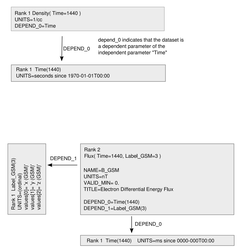


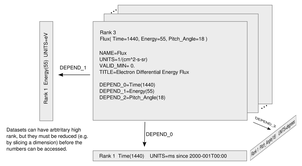
Autoplot uses QDataSet geometry and metadata to decide how it should be displayed. Metadata and Metadata Model are used to discover default labels and range settings. Metadata is displayed in a tree GUI component.



The Image Data Source plugin reads data from JPG files using Java's ImageIO. Written for fun and to demonstrate writing plugins, this demonstrates the y axis isotropic mode that locks the scaling. The context overview check box causes the context overview to be displayed. Dragging in the context overview controls zoom window of the top plot. Jpeg EXIF Metadata is displayed in the metadata tab (which is hidden).

"QDataset" is the working name for the internal data model. It borrows good ideas from NetCDF, CDF, and CEF. It uses a thin syntax layer and semmantics to achieve abstraction. QDataSets can be serialized to a "QStream" without any information loss.





Jython is the Python language in Java. Autoplot presents a programming interface so that Jython can be used to programmatically control Autoplot. QDataSets can be used with operator overloading to form an expressive language with rich operators, such as unit-aware addition, slicing and subsetting.

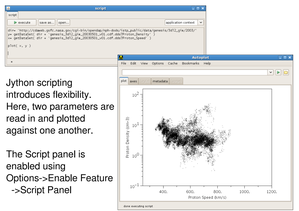
New datasets can be derived using Jython scripts. This script loads B-field magnitude data and converts it to the electron cyclotron frequency (fce):

```
=== file:///home/jbf/data/mag/fce.jy ===
yyjjjj= params['yyjjjj' ]
asciiTableParams= 'time=field0&timeFormat=YY+DJ+DH+DM+DS&column=field5'
file= 'file:///home/jbf/data/mag/'+yyjjjj+'.dat'
bmag= getDataSet( file + '?' + asciiTableParams )
fce= bmag * 28
fce.putProperty( QDataSet.UNITS, Units.hertz )
fce.putProperty( QDataSet.TITLE, 'electron cyclotron frequency' )
fce.putProperty( QDataSet.LABEL, 'fce' )
```

Then the Autoplot URI is file:///home/jbf/data/mag/fce.jy?yyjjjj=96292&fce

There are more example scripts on the Autoplot website.

Autoplot can also provide data to IDL and Matlab. This is documented on the Autoplot website.



Jython scripting introduces flexibility. Here, two parameters are read in and plotted against one another.

The Script panel is enabled using Options->Enable Feature ->Script Panel

Here are some things that can be done with scripts:
* batch plotting of data
* batch reformatting of data
* batch data processing, such as data reduction
* add a feature, such as a digitizer that records events into a file
* add a mouse action, such as run analysis on selected interval

There are many example scripts on the Autoplot web site.

In the Virtual Observatory Environment, data is produced and consumed in many ways. Autoplot is well-positioned to adapt to many applications. These are all working operational modes:

The Autoplot application allows the scientist to trivially access data in many forms.

URIs provide compact way to refer to data.

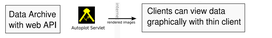
The scientist can "just get the numbers" for analysis without writing more code to read data.

Anyone can create a rich data service without complex server-side code.

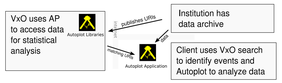
Data visualization is easily added to archives by creating plug-in.

Autoplot can run as an unsigned applet embedded in a web page.

Java Servlets can use Autoplot to serve plot images or reformatted data.

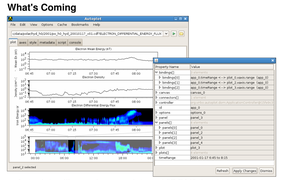


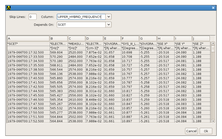
Autoplot allows VxOs to access data and add value to archives.



Autoplot is open-source and can be used freely.

What's Coming



Autoplot will soon have an internal "DOM" that is a tree containing the entire application state. This allows for multiple panel plots and improves support for scripting.

URI Customizer GUIs will allow clients to graphically specify dataset URIs. This will improve data discovery because metadata and caveats can be displayed. Some data providers may display logos for data source branding.