VOPlot: A Toolkit for Scientific Discovery using VOTables

Sonali Kale, T.M. Vijayaraman, P.R. Krishnan, Amey Navelkar,
Hrishikesh Hegde, K.D. Balaji

Persistent Systems Private Limited, Pune, India[1]

Ajit Kembhavi, Pallavi Kulkarni

Inter University Center for Astronomy and Astrophysics, Pune, India[2]

Abstract. Graphical interfaces are extremely useful tools in data mining as well as statistical analysis of huge datasets. We describe a graphical visualization tool, VOPlot, which can be interfaced with data in VOTable format to form a valuable aid in scientific discovery. The VOPlot toolkit, developed by us, is available as either a stand-alone tool or as a web-based applet fully integrated with the VizieR catalogue service. The VOPlot tool is very versatile, and can be used to perform a number of graphical tasks and simple statistical estimations. VOPlot can produce output in postscript and VOTable formats and can be easily integrated with packages that require visualization support. It has full two-way communication with Aladin, and has been adopted in packages like LEDAS. We describe VOPlot and recent enhancements to it, and demonstrate its usefulness in science applications. This paper provides a brief description of VOPlot. The toolkit was developed by Persistent Systems[1] and IUCAA[2] in collaboration with CDS, Strasbourg as part of the Virtual Observatory India initiative.

1. VOTable Basics

The VOTable format is an XML (Extensible Markup Language) standard for exchange of tabular data, with particular emphasis on astronomical tables. A table is an unordered set of rows, each of a uniform format, as specified in the

---

[1]Persistent Systems Pvt. Ltd., located in Pune, India, is an 12-year old, 700-person, software development company. Persistent specializes in developing data infrastructure software. Persistent has developed data management components such as bit-vector indexes, query optimizers, ETL Tools, OLAP Tools etc. It has worked for companies like Microsoft, Hewlett-Packard, Agilent Technologies, Informix, i2 Technologies, Engage Technologies etc. - http://www.persistent.co.in

[2]The Inter-University Centre for Astronomy and Astrophysics (IUCAA) is located in Pune, India. It has a vigorous programme of research in theoretical astrophysics, observational astronomy and instrumentation development. IUCAA has a well-developed data center and hosts mirror sites of various astronomical databases and scientific literature archives - http://www.iucaa.ernet.in/.

table metadata. Each row is a sequence of table cells, and each of these is either a primitive data type, or an array of such primitives. VOPlot uses data in TABLEDATA section for visualization.

2. VOPlot

The motivation behind the VOPlot was to provide a web-based visualization tool for astronomical data in VOTable format. A typical astronomer downloads catalogue data on his personal computer (PC) and uses various plotting tools for visualization. This method is inconvenient and hence the requirement of a web-based tool that can plot data online. VOPlot is a java applet that can be integrated with any web-based catalogue service that generates data in VOTable format.

The stand-alone version of VOPlot can be started using the command 'java -jar voplot.jar' on the command prompt. Java Runtime Environment (JRE) 1.3 or later is required to be installed on user's computer. The web-based version on VizieR can be started by querying a catalogue and selecting the output format as VOPlot. This requires a java plugin version 1.3.1 or later installed on the user's computer.

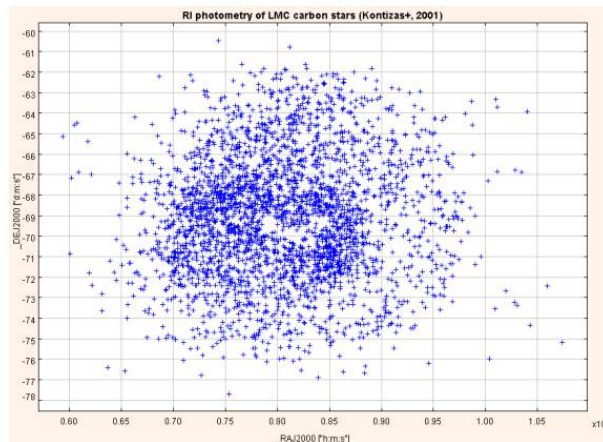## 3. Salient features of VOPlot

3.1. Scatter Plots

Figure 1.   Sample Scatter Plot

VOPlot can plot columns against each other as a scatter plot or a connected plot. The user can select columns to be plotted from list of columns displayed. The plots can be overlaid for comparison. Overlaying is simultaneously viewing multiple plots with similar range on the same axes system. The user can plot data points on a log scale by setting the option appropriately using the "Log" checkbox. User can plot error bars by specifying upper and lower error limits for the Y-column. Plot properties such as X and Y axis labels, title of the plot,

color and marker style can be changed dynamically from the 'Plot Properties' dialog.

VOPlot supports reading from multiple VOTable documents. Data from multiple VOTables can be overlaid for comparison. Data distribution can be viewed in the form of histograms.

3.2.   Create data subsets

Users can create new data subsets by defining filters on them. A condition with relational operators and logical operators can be used to create new data subsets. For example data subsets can be created using expression '$\$1 + log(\$2) > 12'$ which can be used for plotting
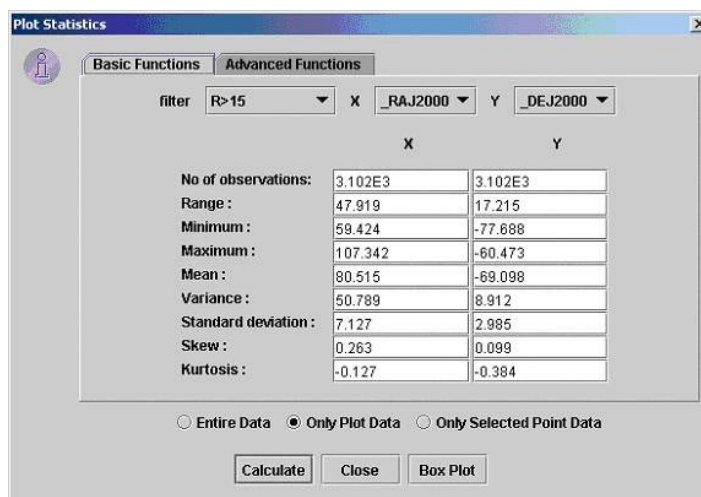
3.3.   Statistical Analysis

Figure 2.   Sample screen-shot showing statistical functions

VOPlot supports uni-variate and multivariate functions such as range, mean, variance, standard deviation, skew, kurtosis, linear correlation, significance, rank correlation and partial correlation.

3.4.   Interoperation with Aladin

VOPlot has been successfully integrated with Aladin. Aladin, developed at CDS, is an interactive software sky atlas allowing the user to visualize digitized images of any part of the sky. An application of this synergy is a simultaneous visualization of some graph in VOPlot together with a representation as a sky plot in Aladin. Selecting some region in the graph highlights the corresponding points in the sky plot and vice versa.

3.5.   Relevant websites

1. Aladin - http://aladin.u-strasbg.fr/aladin.gml
2. CDS - Centre de Donnes astronomiques de Strasbourg
   (http://cdsweb.u-strasbg.fr/)

3. IUCAA - Inter University Center for Astronomy and Astrophysics
   (http://iucaa.ernet.in)
4. LEDAS - LEicester Database and Archive Service
   (http://ledas-www.star.le.ac.uk/)
5. Persistent Systems Private Limited
   http://www.persistent.co.in
6. Vizier Catalogue Service
   http://urania.iucaa.ernet.in/
7. VOTable - http://vizier.u-strasbg.fr/doc/VOTable/