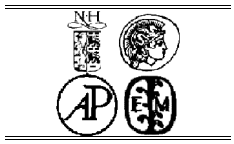


GumTree: Data Reduction

Hugh Rayner*, Paul Hathaway, Nick Hauser, Yang Fei, Ferdi Franceschini, Tony Lam

Bragg Institute, Australian Nuclear Science and Technology Organisation, Lucas Heights, NSW, Australia

Elsevier use only: Received date here; revised date here; accepted date here

Abstract

Access to software tools for interactive data reduction, visualisation and analysis during a neutron scattering experiment enables instrument users to make informed decisions regarding the direction and success of their experiment. ANSTO aims to enhance the experiment experience of its facility's users by integrating these data reduction tools with the instrument control interface for immediate feedback.

GumTree is a software framework and application designed to support an Integrated Scientific Experimental Environment, for concurrent access to instrument control, data acquisition, visualisation and analysis software. The Data Reduction and Analysis (DRA) module is a component of the GumTree framework that allows users to perform data reduction, correction and basic analysis within GumTree while an experiment is running. It is highly integrated with GumTree, able to pull experiment data and metadata directly from the instrument control and data acquisition components.

The DRA itself uses components common to all instruments at the facility, providing a consistent interface. It features familiar ISAW-based one and two dimensional plotting, an OpenGL-based three dimensional plotter and peak fitting performed by fityk This paper covers the benefits of integration, the flexibility of the DRA module, ease-of-use for the interface and audit trail generation.

Data Analysis; Data Reduction; GumTree; ANSTO; OPAL

1. Introduction

This paper describes the data reduction and online analysis facilities being created for the GumTree Integrated Scientific Experiment Environment. These tools are primarily for online use -whilst the experiment is in progress - but can be used offline on acquired data from completed experiments, or for comparison of new and historical data sets.

The GumTree DRA (Data Reduction and Analysis) module has three major goals: to provide tools to enable users to make informed decisions while performing experiments; to provide a familiar interface to users who use multiple instruments; and to maintain an audit trail of all data sets in use.

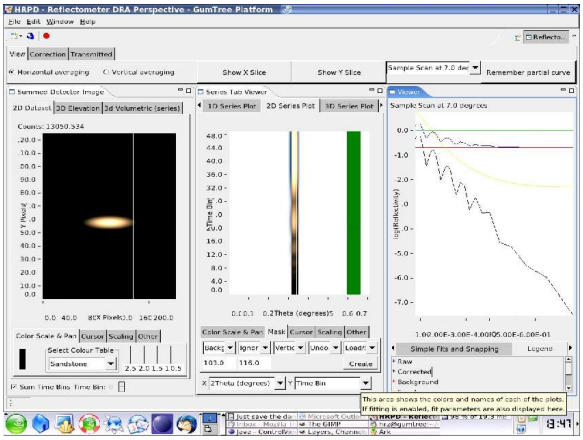
This module is integrated with the GumTree instrument control package and is designed to be flexible, consistent, intuitive and traceable.

By integrating the data reduction tools with the data acquisition package, data is made available for visualisation as it is acquired, complete with any automatic data corrections and reduction steps. This allows users to monitor their data statistics, giving them informed control over the progress of the experiment. Experiment, instrument and sample metadata, such as beam wavelengths, motor positions and sample orientation are also made available for data reduction and can be stored with reduced data.

* Corresponding author. Tel.: 9717-7026; e-mail: hrz@ansto.gov.au.

2. Flexibility

One of the key objectives for the data visualisation and reduction in GumTree is flexibility. Firstly the data can be acquired from a number of sources and treated in the same way. These sources typically include the data acquisition histogram memory, and data files from previous sessions. This allows the user to re-enact the reduction steps performed during the experiment in order to verify the results, or compare the results from two separate experiments, either off-line or while one experiment is running.



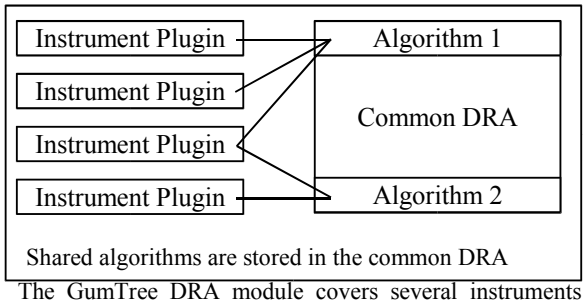
Screen 1: Reflectometer. From left to right: Detector image, averaged time bins and calculated reflectivity.

The GumTree application is based on the Eclipse platform which provides for interface views that can be stacked, tabbed and detached freely during the session while ensuring that none of the key views are lost. An extensive set of masking tools is provided, allowing users to quickly and interactively select a subset of the acquired data image to process in real-time.

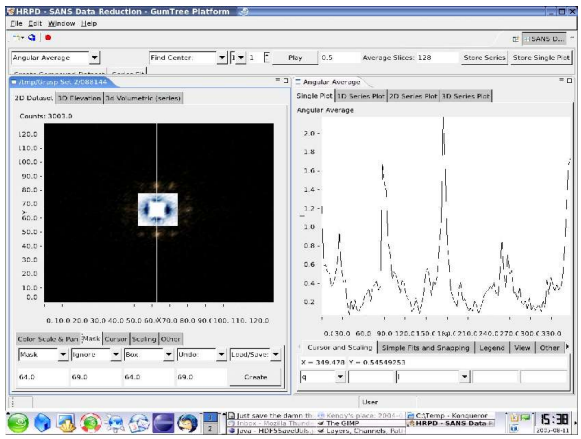
A key feature is the ability to view the data in several different but consistent ways. Two standard views used for most instruments are the “Series Tab Viewer” and the “2D Tab Viewer”, designed to display either a series of 1D data sets or a single 2D data set in a variety of ways. For example, the 2D Tab Viewer can display a data set as a 2D colour map or a 3D height map, or it can display a series of datasets as a volumetric 3D plot, with intensity shown by colour and low intensity pixels transparent. For the case of a 2D reflectometer pattern binned by time of flight, the intensity can be replotted against any two axes from Qx, Qz, wavelength, vertical position on detector, time bin and reflected angle.

Because GumTree is built on a plugin based architecture and has publicly available java source code, it is easier for a user to add their own algorithms in java to the DRA module. This is achieved by either creating a new plug-in for their instrument, extending the views of an existing instrument, or by patching existing plug-ins with the new functionality.

1.3. Consistency and Reuse



The GumTree DRA module covers several instruments with similar needs by reusing many visual and algorithmic components.

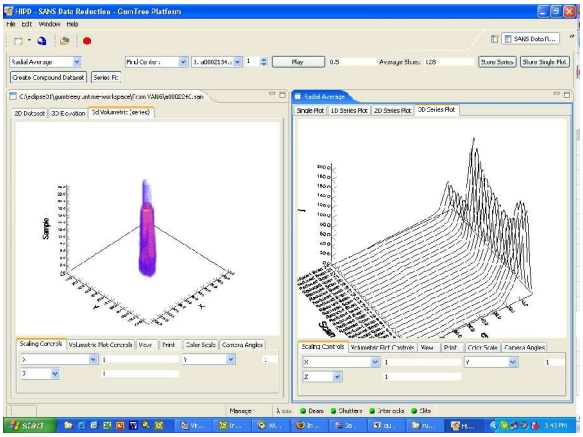


Screen 2: SANS. The left view shows the detector image (with a mask applied over the centre), and the right view shows an angular average applied over the rest of the image. The masking tools are the same as those available to the reflectometer.

Algorithms and techniques useful to more than one instrument are made accessible to all the appropriate instruments through GumTree's plugin based architecture. Tools primarily developed for one instrument can easily be made available to others. Improvements made to one algorithm are automatically made available to all other tools using it.

The plotting tools represent a layer of functionality that is common to all the instruments. On any plot where a 2D data set or a series of 1D data sets are available, the user can view the plot as a 2D colour map or a 3D height map (as lines or as a surface). All 1D and 3D plots can have

their axes scaled by function, so a log X vs Y^2 vs 1/Z plot can be displayed by selecting the appropriate functions from pop-up menus and providing any required parameters. Peak fitting is common between all instruments, allowing various curves to be fitted to any single 1D data set and some 1D series data sets.



Screen 3: SANS and 3D. The left view shows a volumetric plot, and the right view shows a series reduction as a 3D line plot.

The DRA module allows a user to make use of their favourite supported plotting system by being visualisation-agnostic. That is, the analysis does not depend upon the particular way a visualisation is rendered to screen, as long as the chosen visualiser code provides a required interface. There are some visualisation features that are optional. The visualiser will still provide their available capability whether or not these optional features are provided. Consequently, different types of plots can be simultaneously displayed using independent visualisation tools. For example, an OpenGL-based 3D plot can be displayed at the same time as an ISAW 2D plot while a 1D plot is rendered by PTPlot.

The masking tools used to select parts of a diffraction pattern for analysis are available to both SANS and the reflectometer users. The 'slicer' is used to visually select a band of interest from a diffraction pattern is common amongst High Intensity and High Resolution Powder Diffraction instruments, the residual stress instrument, and any other instruments requiring this functionality.

2.4. Responsiveness and Intuitiveness

The DRA module is likely to employed by users while performing their experiments. It is clear that it should be easy to learn and use. To this end, the interface is driven mainly by on-screen buttons, with on-screen descriptions of button function available when delaying the movement of the mouse over the button.

A responsive interface, with responsive controls, is essential to making the DRA module easier to use. Whenever the user action triggers a change, the results of that change is displayed as quickly as possible. This is demonstrated as the user is creating a mask over the diffraction pattern updating the results on other views in real-time. The user can also smoothly scroll between angles of averaging or data sets, with immediate effects in several views of intermediate results.

6.5. Audit Trail

When a user creates a reduced data set , the relevant metadata describing the reduction, with references to the source data set, is added to the new data set. Because a NeXus-based data format is used, precise values of all metadata is associated with the processed data and the original data sets, creating an effective audit trail, describing all the transformations that have been applied to the data.

3.

4.The user is able to retrace any of the reduction steps based on this stored data.

5.6. Integrated Scientific Experiment Environment

The GumTree DRA module functions as part of an Integrated Scientific Experiment Environment or ISEE. The main benefits of this are that it allows data to be transferred transparently and seamlessly between the data acquisition and the data reduction and analysis tools while allowing simultaneous views from both acquisition and analysis.

6.7. Possibilities for The Future

The flexibility of GumTree and the capability to integrate many tools for data analysis makes the package open to adoption by other facilities, further disseminating a common look and feel. It is designed to leverage other distributed analysis initiatives, such as the DANSE framework, under development by CalTech for the SNS and collaborating facilities.

7.8. Conclusion

8.The GumTree data reduction module provides an effective tool for users to view the progress of their experiments and prepare the data for further analysis. The

software has been tested on supplied and simulated data sets. Development of features continues via integration into the data analysis software suite being made available for the ANSTO OPAL facility..