USAGE WORKFLOW

CODE DEVELOPMENT SETUP

Step 1: Download and Install Java and Netbeans from the following link:

http://www.oracle.com/technetwork/java/javase/downloads/jdk-netbeans-jsp-142931.html

This will download latest Java JDK and the latest version of Netbeans.

Step 2: Open Netbeans, click on File -> **Open Project**, and select the source code location of this application. This will open up the project in NetBeans. User can re-build the project and test changes using this integrated development environment. It is necessary to open the project in Netbeans IDE, rather than other development environments like Eclipse. This is so because the directory structure and build configuration files are Netbeans specific, and won't be recognized by other environments.

Step 3: The project is configured in such a way that a successful clean and build creates an executable JAR file in the Dist directory to run.

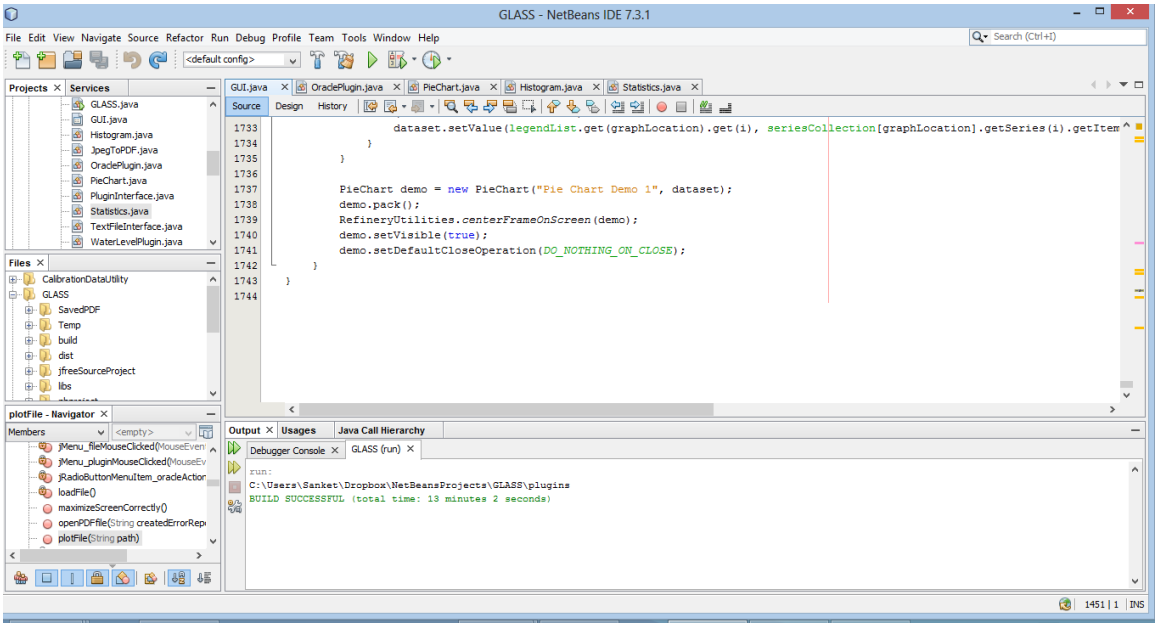


Figure 4.1. Netbeans IDE, Setup for Java Desktop Application Development

APPLICATION INSTALLATION WORKFLOW

Step 1: Update or Install latest Java JRE.

http://www.oracle.com/technetwork/java/javase/downloads/jre7-downloads-1880261.html

Use the above link to update or install Java.

Step 2: Unzip the Application package to any location.

Step 3: Double click .\Glass\Glass.bat to run. On loading the application the following screen should appear.

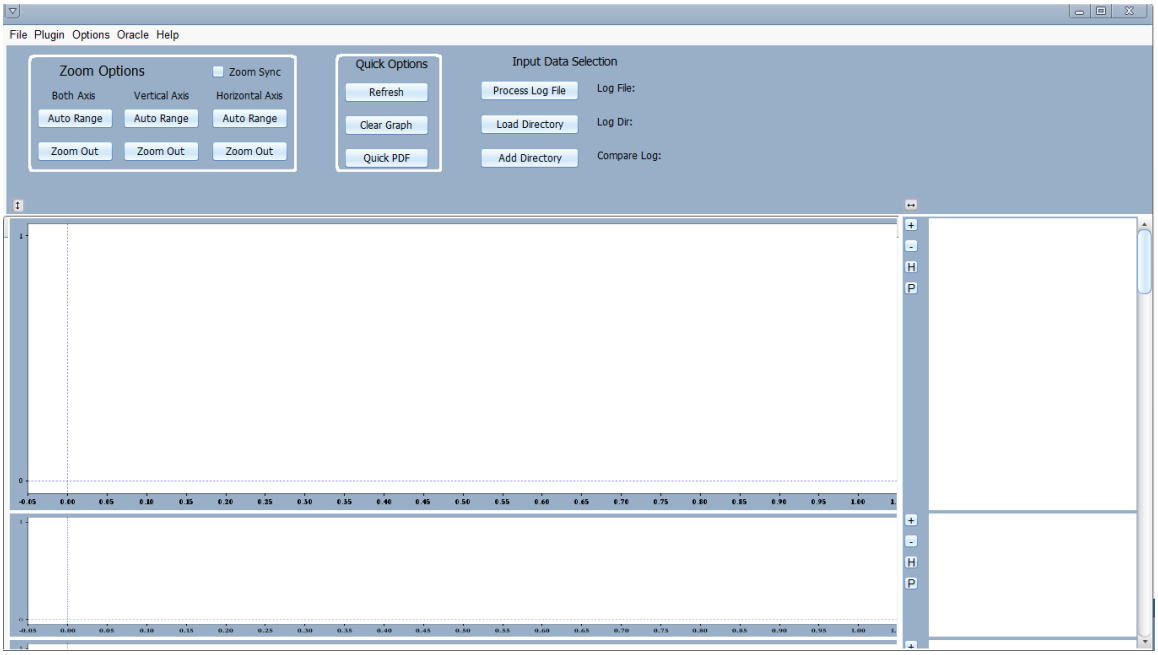


Figure 4.2. Screen shot of the application when a new instance is loaded.

ANALYSIS OF RAW LOG FILE

Use this workflow when analyzing a raw log file for the first time. The application will preprocess the log file and save the details in a temporary csv file to be used later. The application will remember the location of the temporary files and the user is just required to select the options as listed below.

Step 1: Open the GLASS application.

Step 2: Select appropriate plugin from the plugin meu.

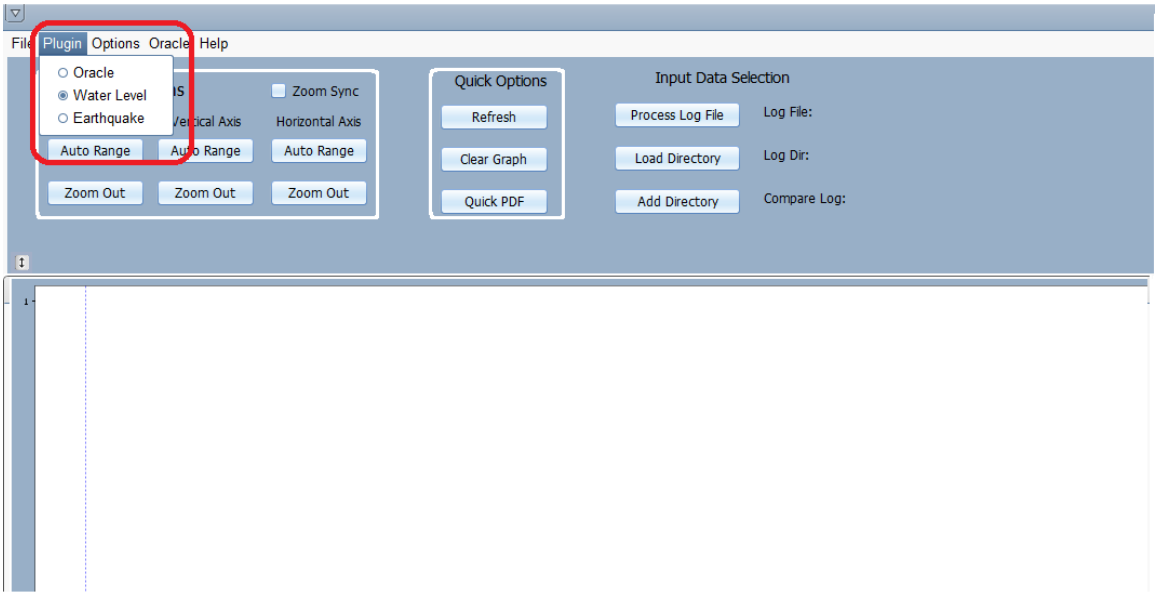


Figure 4.3. Step 2: Select appropriate plugin from the plugin meu.

Step 3: Click on Process Log File Button. This will open up a file selection menu. Select the correct log file to be processed. After selection the application pre-processes the raw log file to application understandable coordinate format, using the code present in the selected plugin.

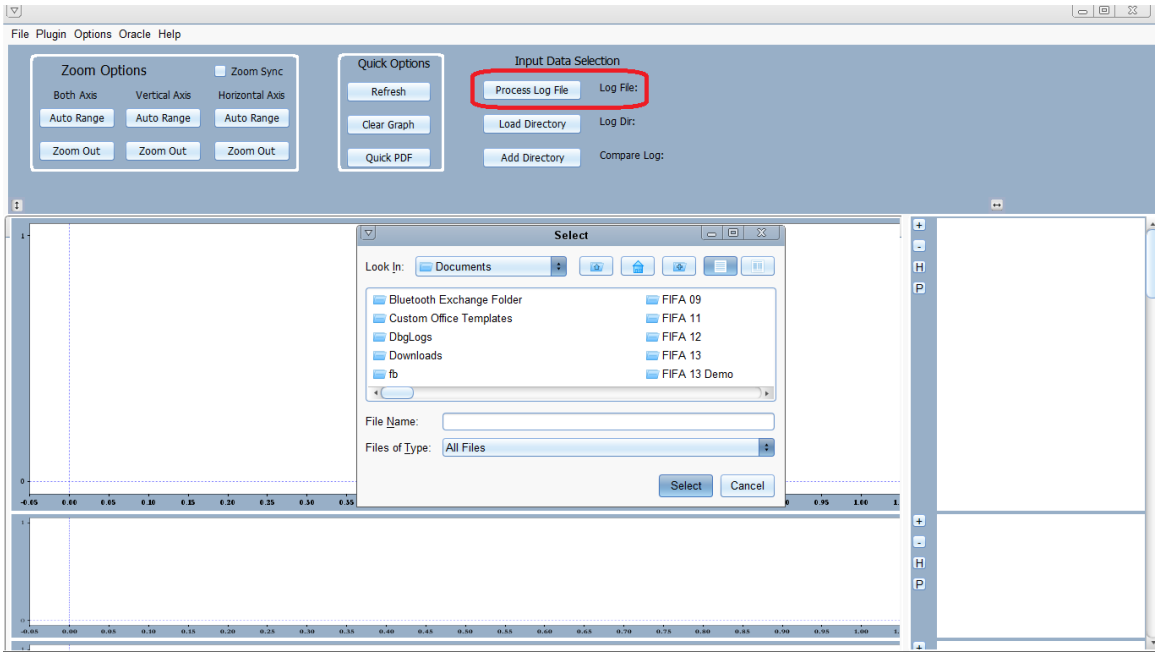


Figure 4.4. **Step 3: Click on Process Log File Button**

Step 4: Right click on the white graph area and select Add Graph Option, and then select the log file you want to plot. This will plot the selected log file as a graph, as shown in the figure.

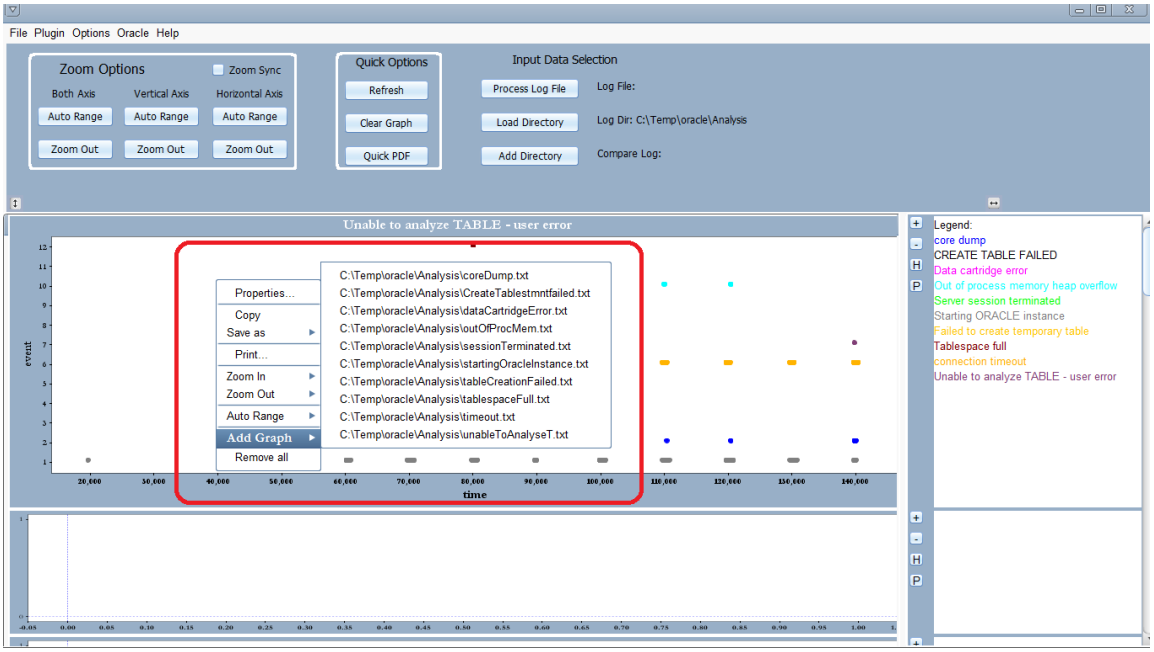


Figure 4.5. **Step 4: Right click on the white graph area and select Add Graph Option**

Step 5: Now the user can perform various functions like zooming, saving as image or PDF document, viewing the graph as a histogram or pie chart and so on. In the image below the buttons outlined in red can now be used for further analysis.

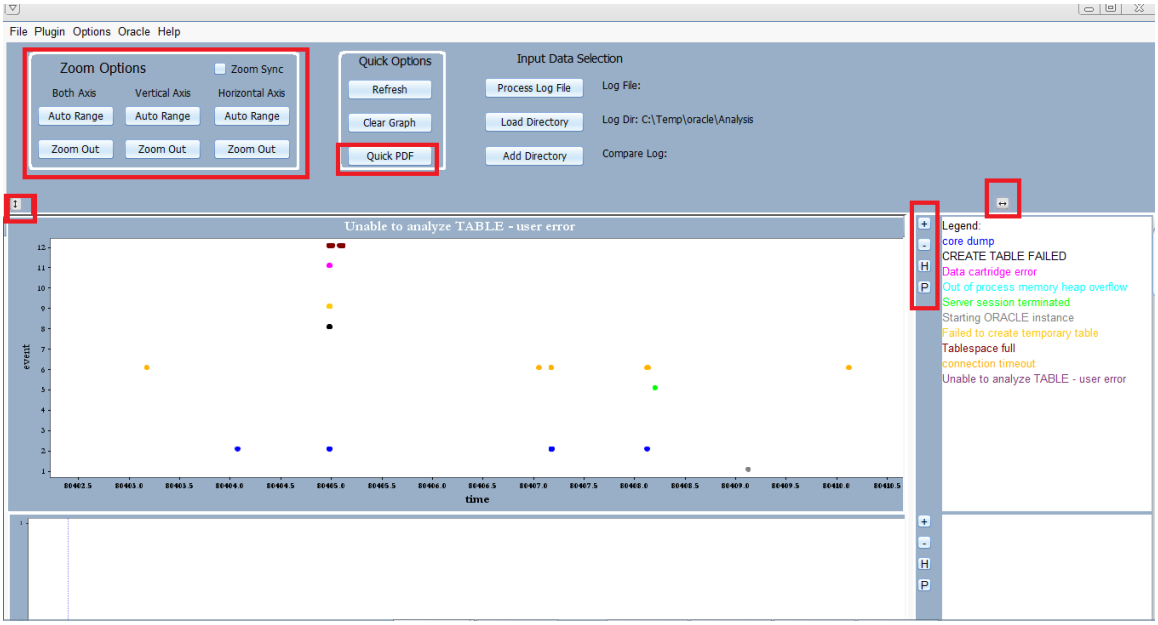


Figure 4.6.B**uttons outlined in red can now be used for further analysis.**

ANALYSIS OF PRE-PROCESSED LOG FILE

Use this workflow when analyzing already processed files. The application will load the directory where the files are located and make available all the graphs.

Step 1: Open the GLASS application.

Step 2: Click on load directory. Select the appropriate directory.

Step 3: Right click on the white graph area and select Add Graph Option, and then select the log file you want to plot. This will plot the selected log file as a graph, as shown in the figure. (Same as step 4 in Analysis of raw log file).

Step 4: Now the user can perform various functions like zooming, saving as image or PDF document, viewing the graph as a histogram or pie chart and so on. In the image below the buttons outlined in red can now be used for further analysis (Same as step 5 in Analysis of raw log file).

EXAMPLE: MYSQL SERVER LOGS ANALYSIS

The oracle database has a strong logging architecture, and logs for different processes and purposes are stored at different location. Over here we are most concerned about the user connection logs and trying to make sure new users are able to perform their task successfully. So we look at the BACKGROUND_DUMP_DEST or the bdump where the alert logs of the database server are stored. The code in the plugin(pre-processor) looks for all the error codes in the file and stores each different type of error in a new file. The error codes are dumped with their timestamp, so they are stored in a chronological order. After pre-processing the log we plot each error message file as a separate graph, one over the other. Each error message is represented by a color coded single dot on the graph.

Looking at the graph we can quickly conclude that most of the errors are related to each other, and a few of them always occur togethar. When the disc memory is full, errors like table space full, failed to create table, create table failed, heap over flow, core dump occur. Another instance is when the server session gets terminated, we can see errors like connection time out and core dump being thrown. Analysing the pie chart the user can quickly conclude that heap overflow is by far the largest occuring error, followed by connection time out and core dump. Amount of analysis the user can do is only limited by his imagination, the application allows the user to write powerful preprocessing functions to bring out the interesting details of the log files.

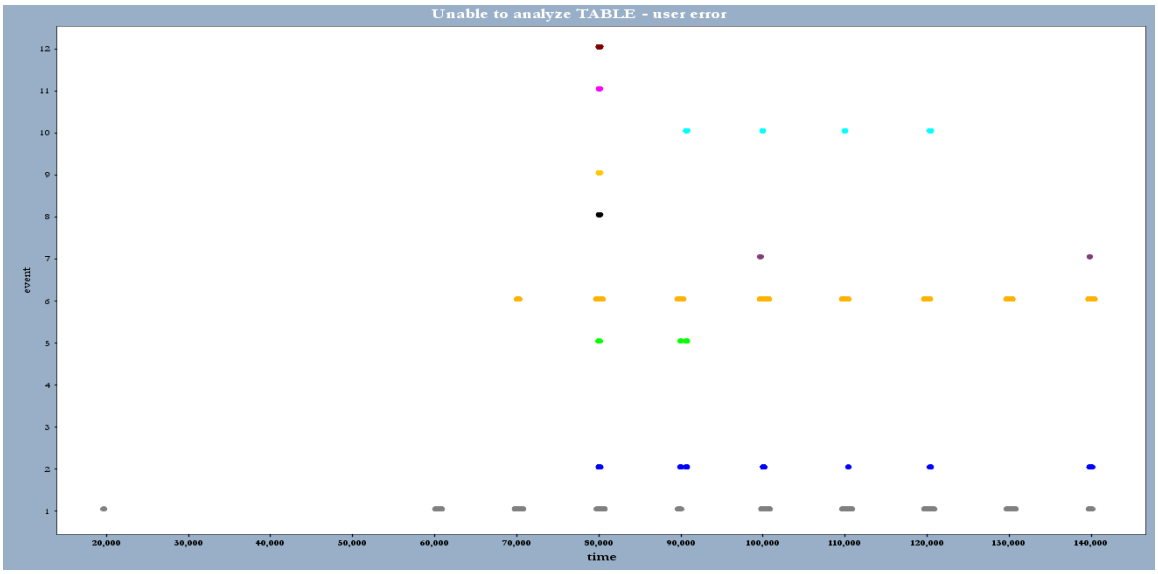


Figure 4.7. MySQL background dump trace log analysis (each dot represents an error)

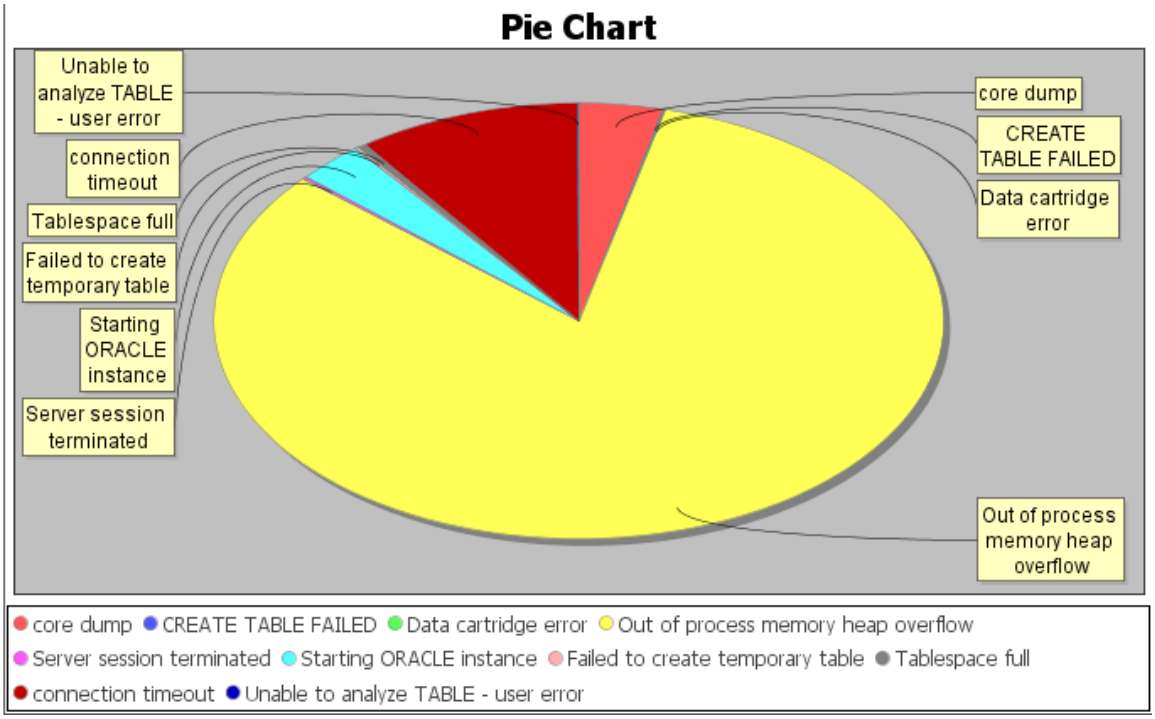


Figure 4.8. MySQL background dump trace log analysis (Pie chart)

SAN DIEGO COAST WATER LEVEL – PACIFIC TSUNAMI WARNING CENTER LOG ANALYSIS

Below is the data from the Pacific Tsunami warning center for San diego Coast water level[19]. The data is spanning over three days in each of the graphs. The histogram and statics panel displays information like min, max, standard deviation and mean, which turns out to be very help ful in this case.

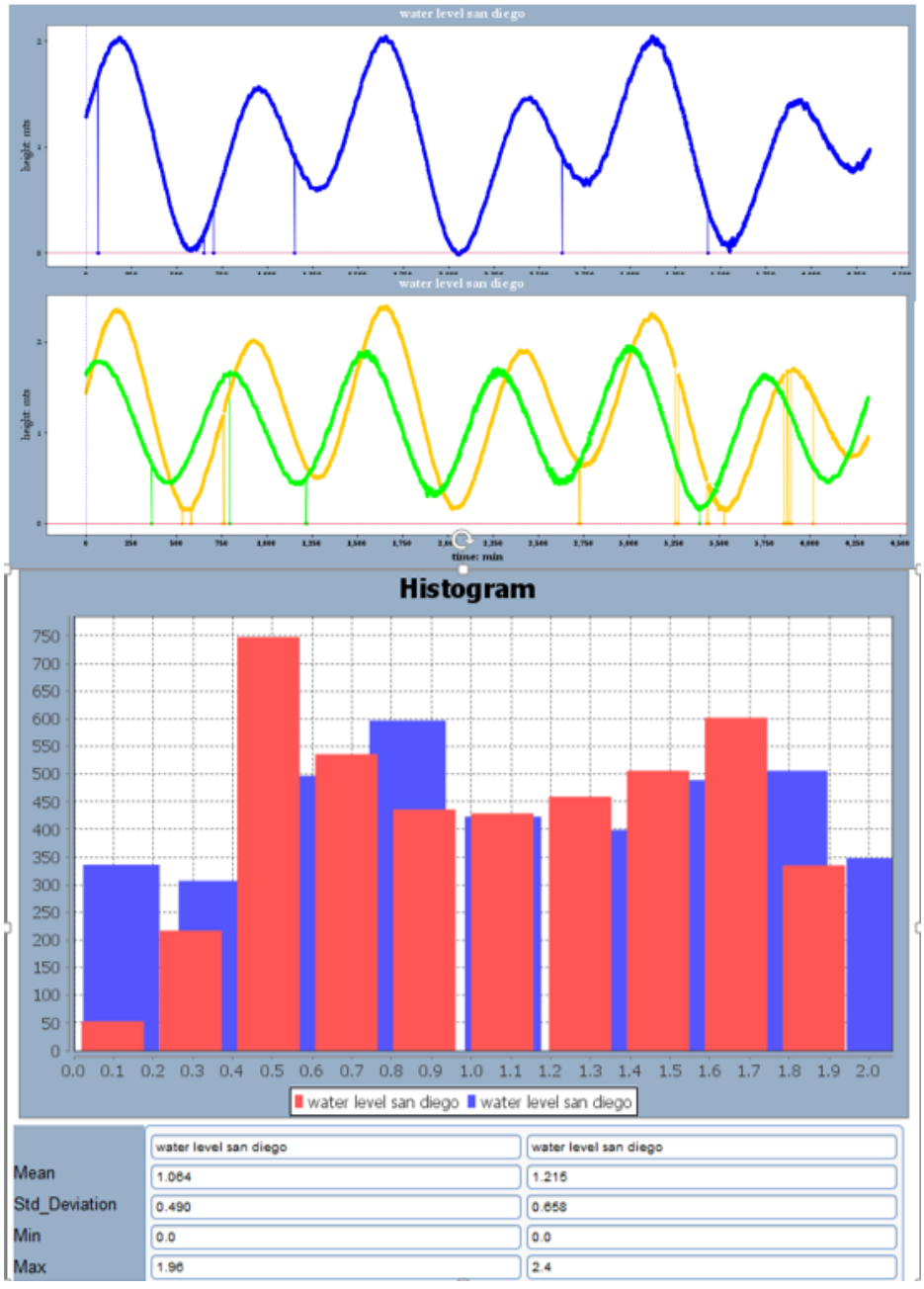
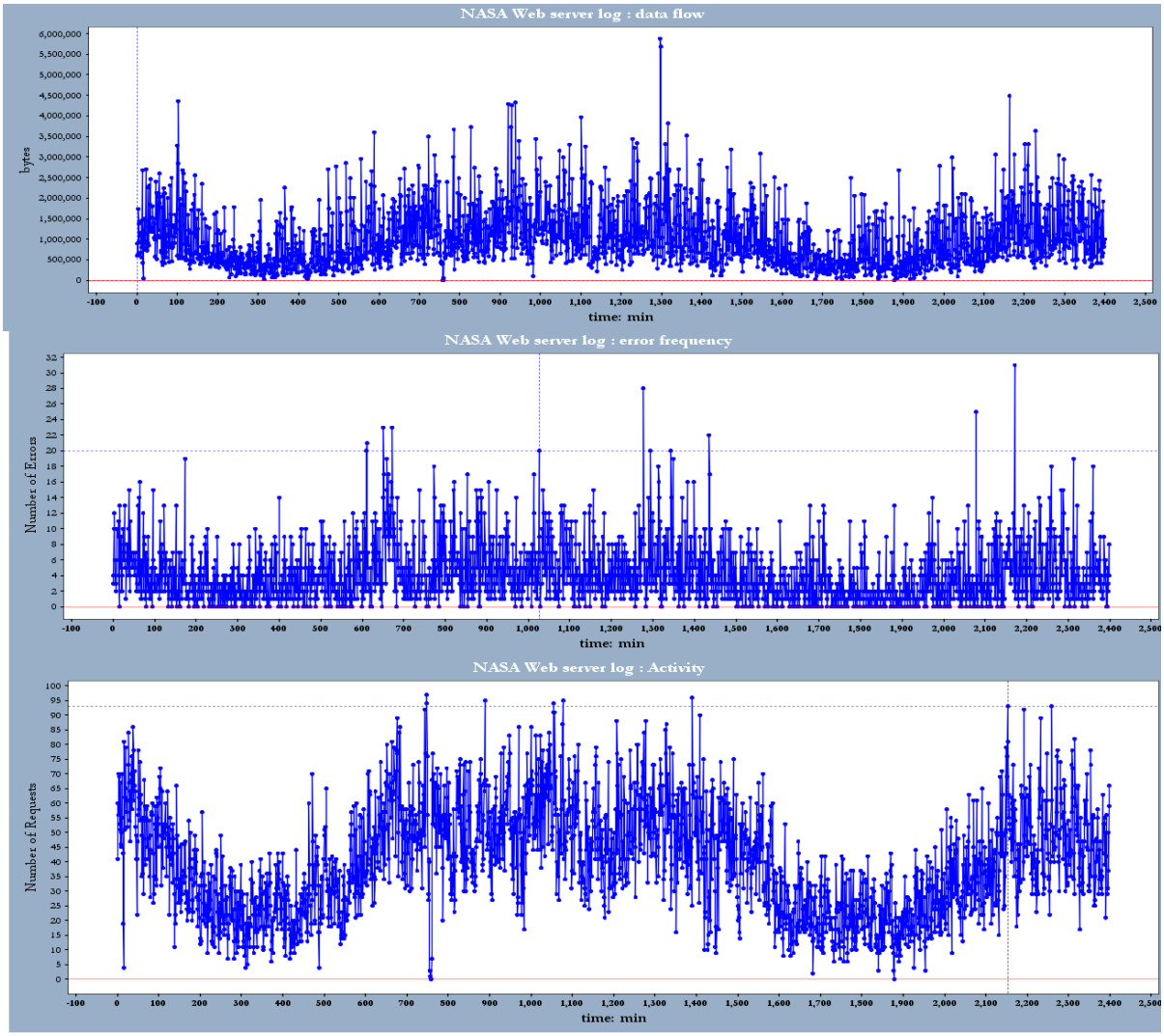


Figure 4.9. San Diego coast water level log analysis

NASA WEBSITE WEBSERVER LOGS 2005 JUNE 1

Below is the data from the NASA web server logs of 2005 June 1. The graphs show data flow, error frequency, and number of hits per minute.



	NASA Web server log : Activity	NASA Web server log : error frequency	NASA Web server log : data flow
Mean	40.181	4.085	1009492.061
Std_Deviation	18.203	3.719	706158.371
Min	0.0	0.0	0.0
Max	97.0	31.0	5878999.0

Figure 4.10. Webserver log analysis

TOOLBAR

The toolbar is an interface which quickly allows the user to perform required tasks, like zooming, saving results in pdf, loading input data and comparing results. The toolbar is divided into 3 sections, each of them concentrating on different tasks.

Figure 3.1. Screen shot of application, highlighting toolbar in a red box.

Zoom options toolbar

It focuses on zooming functions; the six different buttons allow a range of zooming features.

The **Auto Range** button performs automatic bounds calculation, so to set the graph to its default display area. It can be performed for either both the axis or for separate vertical or horizontal axis only.

Zooming out can be performed by clicking on one of the zoom out buttons, depending on if the user wants to zoom out vertically , horizontally, or both axis.

Zoom-in is easily performed by mouse drag on the selected area.

Panning can be performed by holding the Ctrl key and using the mouse drag.

Zoom Sync is a check box used to synchronize zooming between all the graphs active in the tool. If selected zooming actions will be mirrored in all the active graphs in the tool. When not selected, zooming is applied exclusively to the first graph.

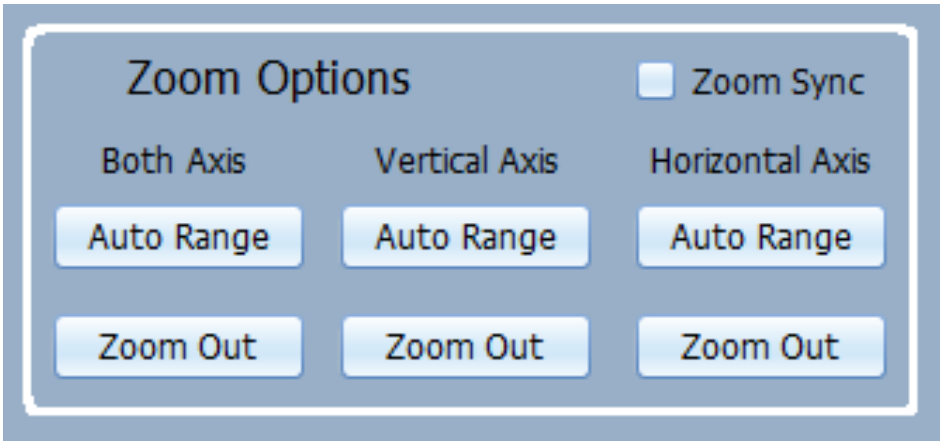


Figure 3.2. Zoom Tool bar

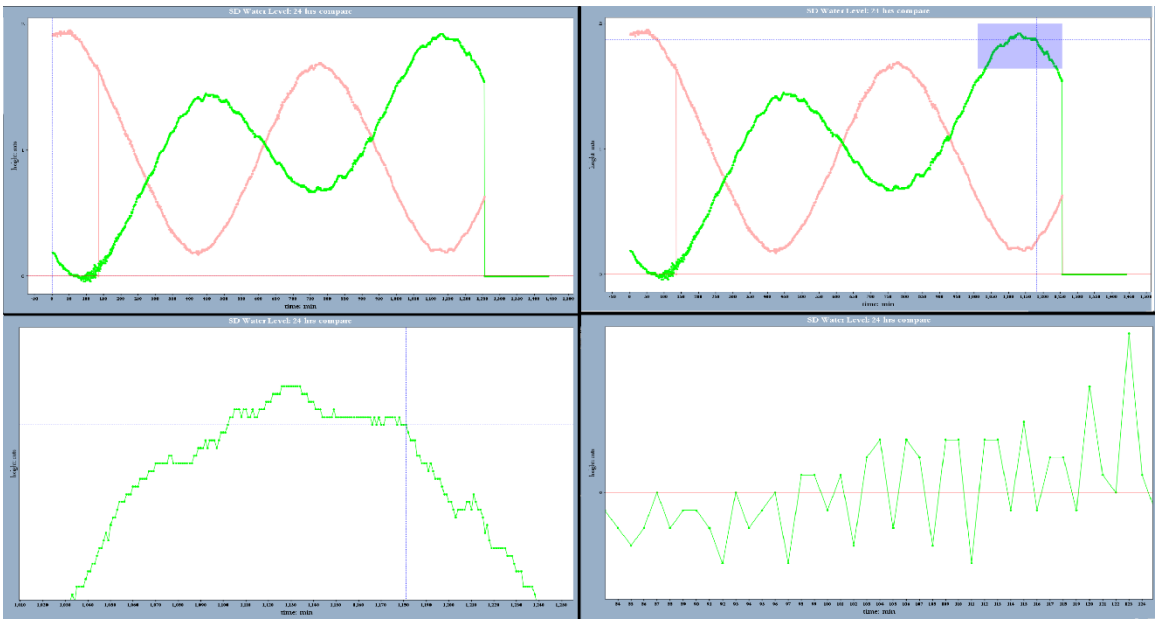


Figure 3.3. Screen shot of application graph panel showing zoom in feature.

Quick Options toolbar

It has three buttons, Refresh, Clear graph, and Quick PDF.

All of the above functionalities can also be accessed via the menu bar, or by right clicking on the tool.

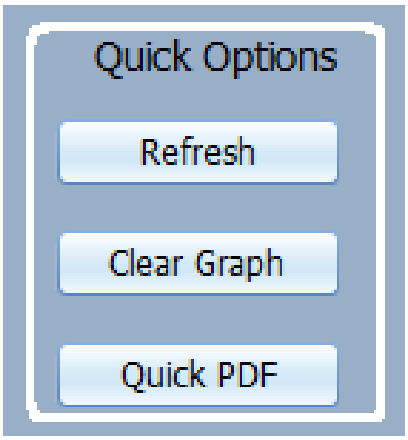


Figure 3.4. Quick action tool bar buttons

Refresh button performs actions like the windows refresh action. It cleans the GUI, if the different components of the tool are stuck or the GUI has crashed at some point. It can be accessed via a short cut F5 key press. If we want to discuss more in detail about what the code does during the refresh action, we can say that it calculates the bounds of the screen,

checks if the tool is correctly sitting in the screen area; if not, it changes the maximum bounds of the tool. Then it goes forward and recalculates the position of each and every component in the GUI and sets its bounds. A refresh button is also available in the menu bar.

Clear Graph button just clears the contents of the graph area plotted on the screen.
So if the user wants to clean his work area and start analysis from scratch he can use this feature. It gives a similar feel of reopening the tool from scratch. The code basically clears all the variables values acquired during the runtime.

Quick PDF creates a PDF document of the graphs along with analysis data and opens it up for the user to view and save. It is a nice feature to quickly save results for future reference. The PDF generated maintains the zoomed bounds, and also adds the right side text box which displays the graph stats and the legends. Highest resolution is selected to generate the PDF documents which allow the user to zoom in to greater depths in the PDF viewer. This option is also available through file menu.

Input data selection toolbar

It helps the user quickly select the input file or directory for analyzing using the application. Here the important thing to note is that the user first selects the plugin type and then the data file to be processed. Selection of the plugin type defines the preprocessing which is to be applied to the data to generate text files with comma separated values to plot using this application.

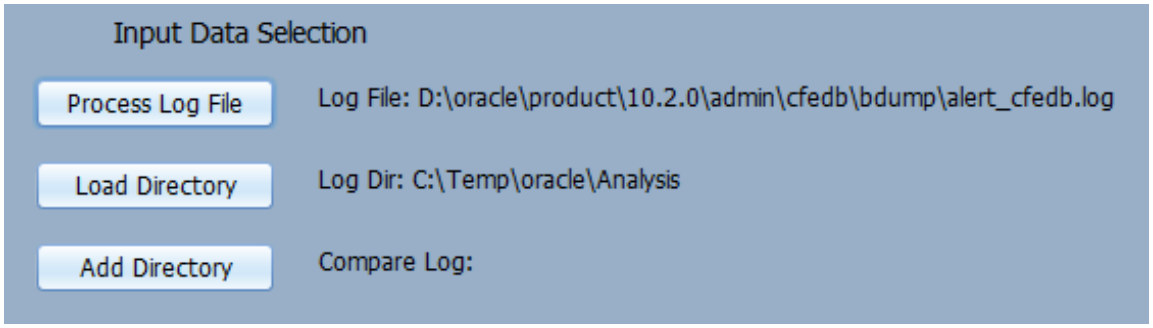


Figure 3.5. Input data selection tool bar buttons

Process Log File button opens a file dialog box, which allows the user to select a data file which he wants the selected plugin to process. After selection, the tool automatically processes the raw data into time series data which can be directly used to plot and analyze data. The tool stores processed data in a temporary location and loads the available graphs options. This option is also available in the file menu.

Load Directory button is used to open already preprocessed data files, which contain configuration files along with time series data. This skips preprocessing of data, and searches for available graph configuration files.

Add directory (compare) button works similarly to the load directory button but is used for comparing similar data sets. After loading a default directory, the user can also load another directory to compare graphs by using the add directory button. When the user selects one graph type, the tool loads plots from both the directories together in the same chart, overlapping each other, allowing accurate visual comparison.

MENU BAR

Like all desktop applications, this application also comes with a fully loaded menu bar. There are different options aggregated under different Menus like File, Plugin, Options, Oracle and Help.

File Menu

Load Data File, Load Log Dir, and Save as PDF are the same as the ones covered in the toolbar, and they perform the same functionality. So we will cover the remaining two here.

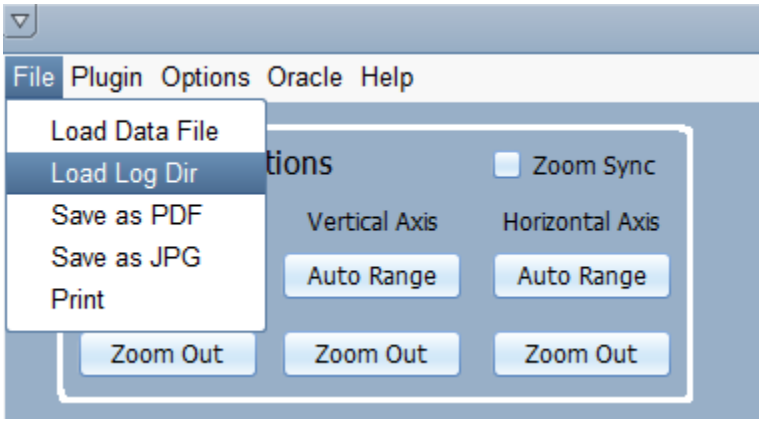


Figure 3.6. File Menu screen shot

3.2.1.1 SAVE AS JPG

This option opens a file saving popup menu to save the first graph in JPG/PNG format. This option works by saving only the first graph. To save other graphs, right click on the respective graph and click on the save option. This is covered in more detail in the graph options.

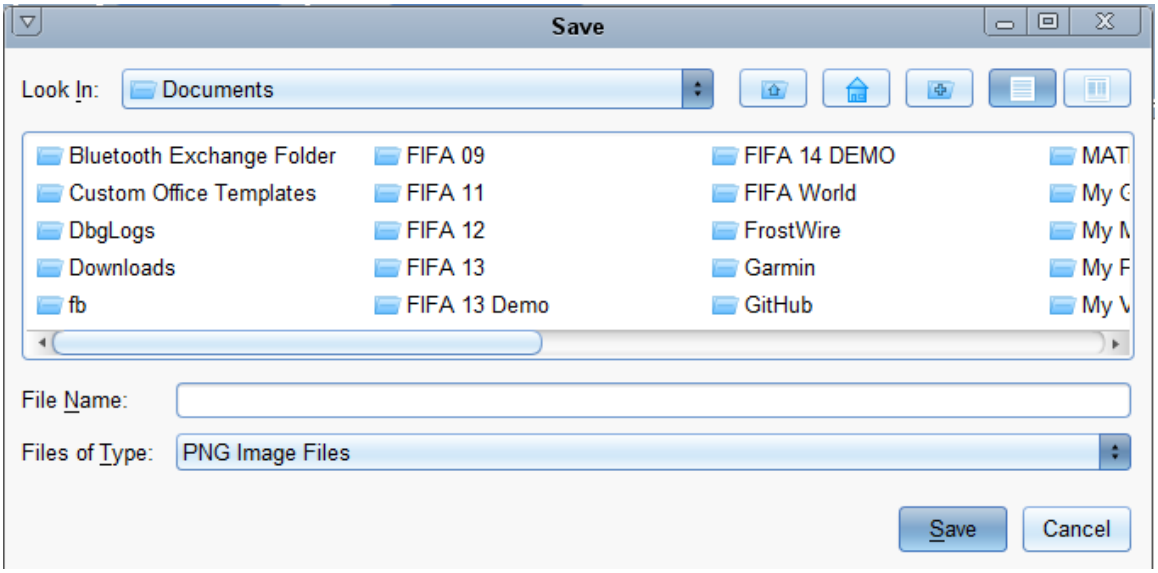


Figure 3.7. File Selection Menu screen shot

3.2.1.2 PRINT

This option opens a print request and a page setup popup as shown in the figure below. The first graph along with labels is sent to the printer; this is similar to sending the saved image to print.

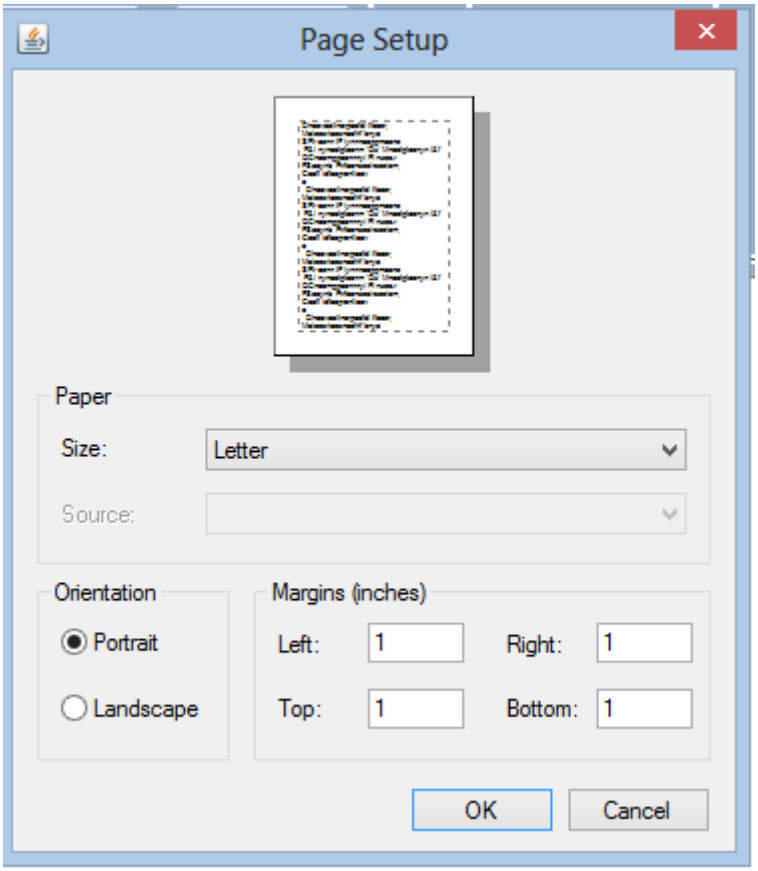


Figure 3.8. page setup Menu screen shot

Plugin menu

This can have different options in different installations of the tool. The plugin menu is used to select the plugin to be used to preprocess the log files. By default an oracle plugin is installed in all the versions of the tool, and by default this is also selected at the start of the application.

The input to the processFile() plugin function is the path of file to be processed, and the output is the path of the processed file. Here the output is expected to be a file or a directory

containing multiple such files with x, y coordinates in each line. The application also expects a configuration file with [output filename]_config.txt describing some basic behavior of the plot. The configuration file will be discussed later in the a separate chapter

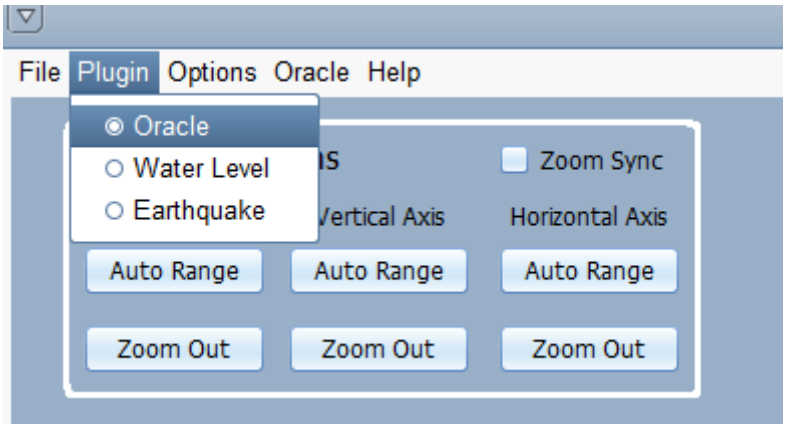


Figure 3.9. Plugin Menu screen shot

Options Menu

Options menu has four options out of which the first two are covered in the Quick options buttons description. The toggle collages toolbar and toggle collapse right panel will be covered here.

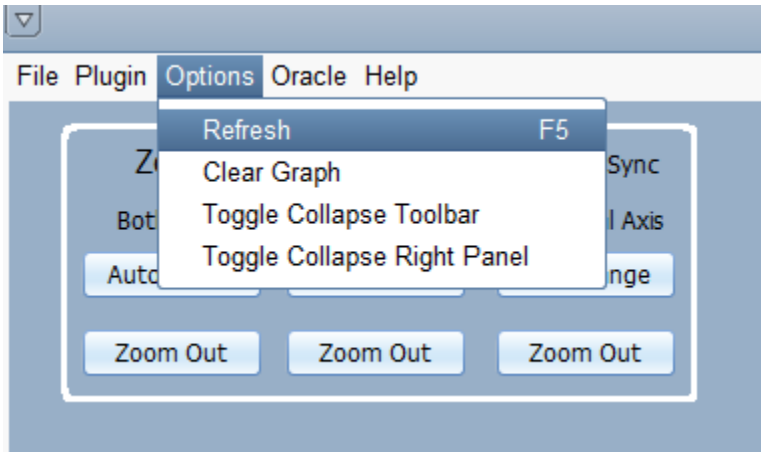


Figure 3.10. Option Menu screen shot

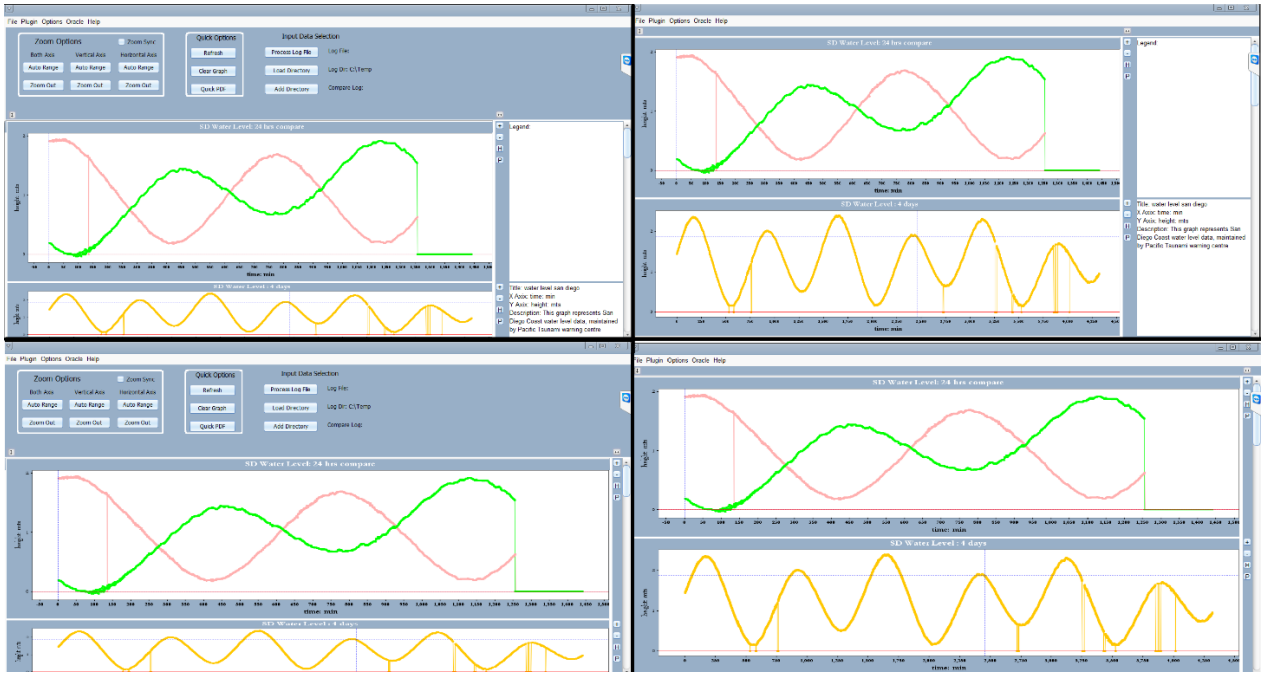


Figure 3.11. Application screen shots displaying the use of toggle collapse buttons. In the first screenshot both the areas are un-collapsed. In the last screen shot both the areas are collapsed.

TOGGLE COLLAPSE TOOLBAR

Selecting this button removes/adds the toolbar from/to the GUI. This allows more space on the screen for the user to zoom in and analyze the graphs. The tool bar can be toggled back in and out easily without making any changes to the basic graph area.

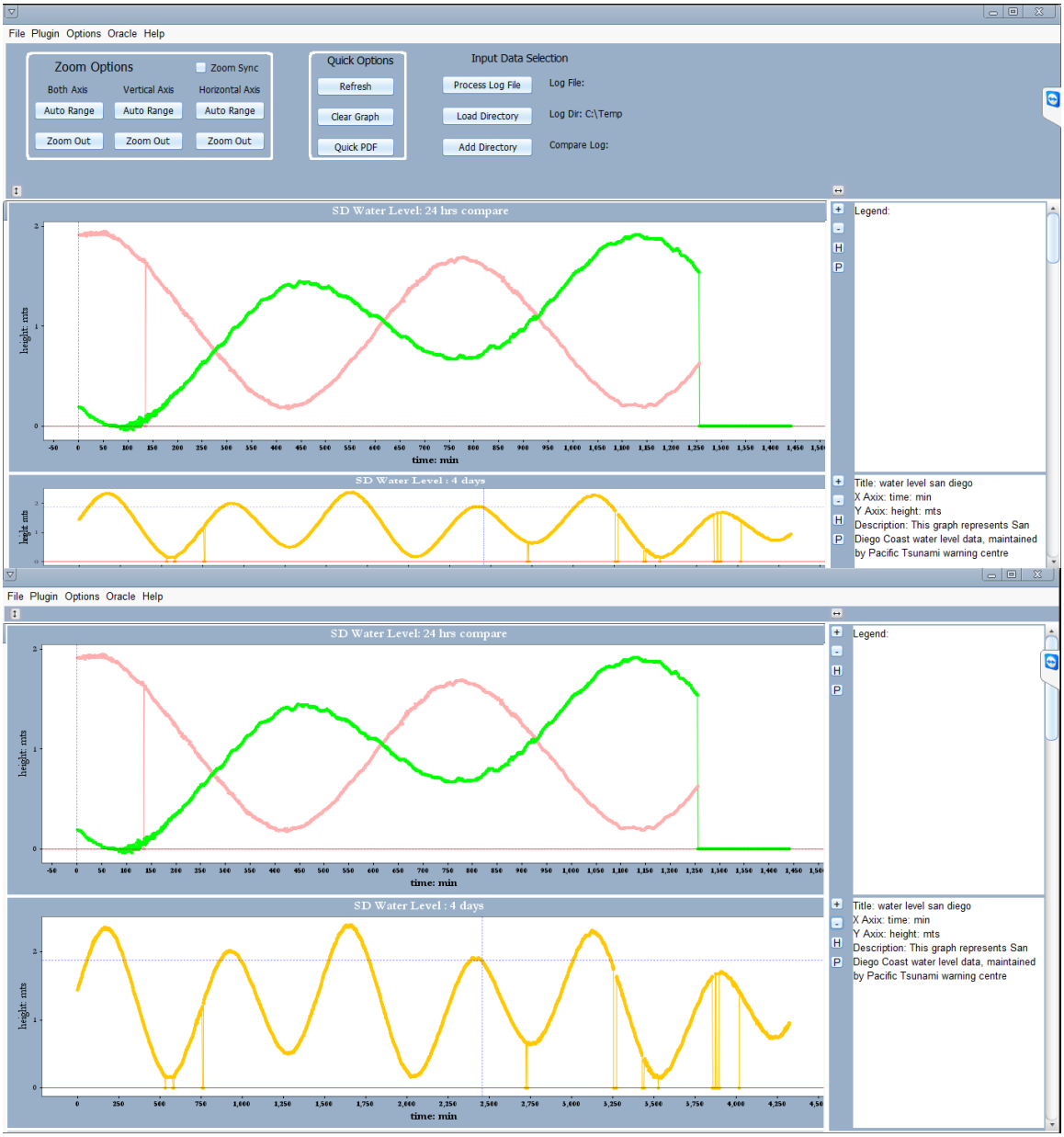


Figure 3.12. Application screen shots displaying the use of toggle toolbar collapse button. In the upper screenshot toolbar is un-collapsed. In the last screen shot toolbar is collapsed.

TOGGLE COLLAPSE RIGHT PANEL

Selecting this button removes/adds the right panel from/to the GUI. This allows more space on the screen for the user to zoom in and analyze the graphs. The right panel can be toggled back in and out easily without making any changes to the basic graph area.

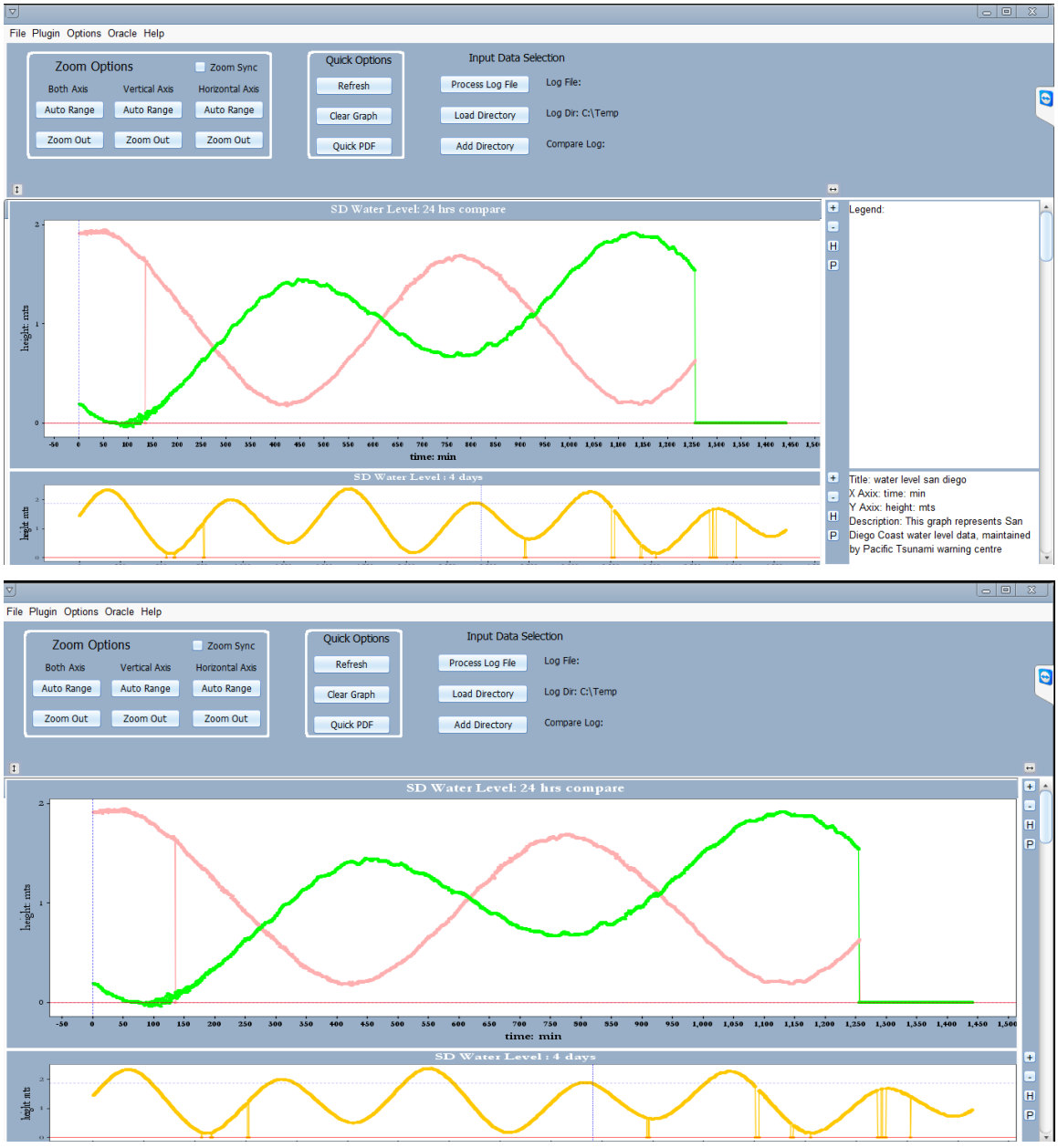


Figure 3.13. Application screen shots displaying the use of toggle panel collapse button. In the upper screenshot panel is un-collapsed. In the last screen shot panel is collapsed.

Oracle Menu

The oracle menu has five options, all related to Oracle database management. This options help the user to get more detailed information about the oracle database easily.

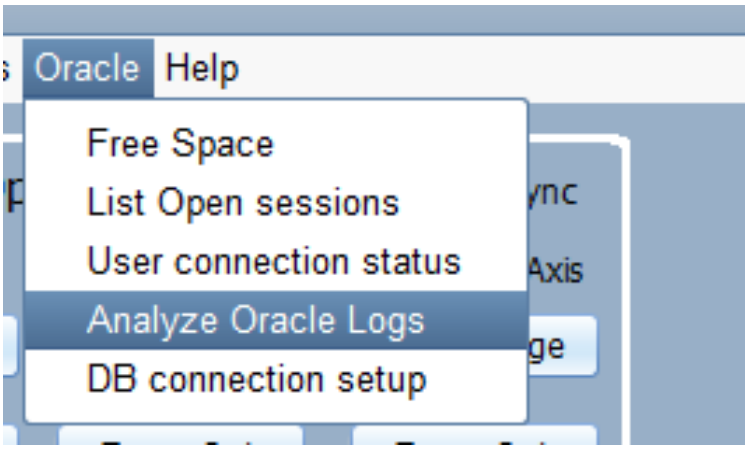


Figure 3.14. Oracle Menu screen shot

FREE SPACE

Selecting this menu option opens up popup or a jOptionPane describing free space in the C: and D: drive of the users system and the space occupied by the Oracle database in the system. This helps the user to identify if the database has session logs occupying more space than expected. It gives the user a clear picture of the memory status of the drive where Oracle is installed.

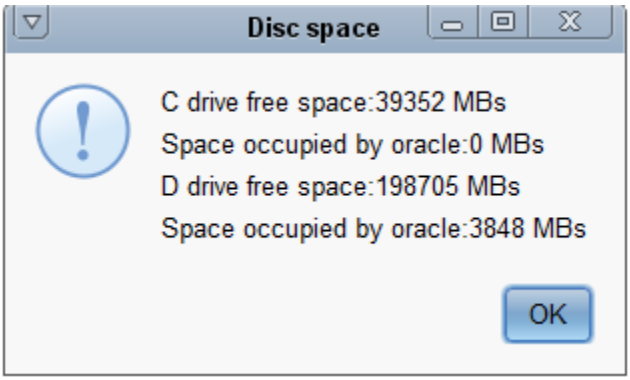


Figure 3.15. Disk space Option Pane

LIST OPEN SESSIONS

On selecting this option, the application makes a JDBC connection to the database specified in the DB connection setup menu. Then it executes a query to list open sessions or processes using the database. The results are displayed in a new window. The PROCESSLIST table provides information about which threads are running [16]. This statement is very useful if you get the "too many connections" error message and want to find out what is going on. MySQL reserves one extra connection to be used by accounts that have the SUPER privilege, to ensure that administrators should always be able to connect and check the system (assuming that you are not giving this privilege to all your users). Threads can be killed with the KILL statement

The following statements are equivalent:

SELECT * FROM INFORMATION_SCHEMA.PROCESSLIST

SHOW FULL PROCESSLIST

The columns produced by SHOW PROCESSLIST have the following meanings:

INFORMATION_SCHEMA Name	SHOW Name	Remarks
ID	Id	MySQL extension
USER	User	MySQL extension
HOST	Host	MySQL extension
DB	db	MySQL extension
COMMAND	Command	MySQL extension
TIME	Time	MySQL extension
STATE	State	MySQL extension
INFO	Info	MySQL extension

Figure 3.16. Informatio_schema table description

Id: The connection identifier.

User: The MySQL user who issued the statement. If this is the system user, it refers to a non-client thread spawned by the server to handle tasks internally. This could be the I/O or SQL thread used on replication slaves or a delayed-row handler. Unauthenticated user refers to a thread that has become associated with a client connection but for which authentication of the client user has not yet been done.

Host: The host name of the client issuing the statement (except for system user where there is no host).

Db: The default database, if one is selected, otherwise NULL.

Command: The type of command the thread is executing.

Time: The time in seconds that the thread has been in its current state.

State: An action, event, or state that indicates what the thread is doing.

Info: The statement the thread is executing, or NULL if it is not executing any statement.

Id	User	Host	db	Command	Time	State
313888	irnadmin	172.19.0.239:40136	v3	Sleep	0	
314075	irnadmin	172.19.0.239:41113	v3	Sleep	0	
314118	irnadmin	172.19.0.239:41282	v3	Query	34978	freeing
314686	irnadmin	172.19.0.239:43251	v3	Sleep	0	
314732	irnadmin	172.19.0.239:43436	v3	Query	34978	freeing
314984	irnadmin	172.19.0.239:44366	v3	Sleep	2	
315051	irnadmin	172.19.0.239:44713	v3	Query	0	NULL
315198	irnadmin	172.19.0.239:51569	v3	Sleep	2	
315280	irnadmin	172.19.0.239:51849	v3	Query	34978	freeing
315320	irnadmin	172.19.0.239:52045	v3	Query	34978	freeing
315384	irnadmin	172.19.0.239:52463	v3	Sleep	1	
452248	irnadmin	172.19.0.28:54899	v3	Query	34978	freeing
452291	irnadmin	172.19.0.28:55045	v3	Sleep	1	
452316	irnadmin	172.19.0.28:55144	v3	Sleep	0	
452353	irnadmin	172.19.0.28:55278	v3	Sleep	0	
452382	irnadmin	172.19.0.28:55371	v3	Query	34978	freeing
452413	irnadmin	172.19.0.28:55479	v3	Sleep	1	
452541	irnadmin	172.19.0.28:55946	v3	Query	34978	freeing
452626	irnadmin	172.19.0.28:56215	v3	Sleep	2	
452711	irnadmin	172.19.0.28:39916	v3	Sleep	0	
452781	irnadmin	172.19.0.28:40161	v3	Sleep	1	
452904	irnadmin	172.19.0.28:40955	v3	Query	34978	freeing
453014	irnadmin	172.19.0.28:41291	v3	Query	34978	freeing
453057	irnadmin	172.19.0.28:41377	v3	Query	34978	freeing

Figure 3.17. Informatio_schema table example

USER CONNECTION STATUS

On selecting this option, the application makes a JDBC connection to the database specified in the DB connection setup menu. Then it executes a query to show connection status. This helps the user evaluate the connection status of the SQL server as a quick summary, and helps the user to quickly analyze any heavy connection problems.

SHOW STATUS provides server status information. This information also can be obtained using the mysqladmin extended-status command. The LIKE clause, if present, indicates which variable names to match. The WHERE clause can be given to select rows using more general conditions

```
+--------------------------+------------+
| Variable_name            | Value      |
+--------------------------+------------+
| Aborted_clients          | 0          |
| Aborted_connects         | 0          |
| Bytes_received           | 155372598  |
| Bytes_sent               | 1176560426 |
| Connections              | 30023      |
| Created_tmp_disk_tables  | 0          |
| Created_tmp_tables       | 8340       |
| Created_tmp_files        | 60         |
...
| Open_tables              | 1          |
| Open_files               | 2          |
| Open_streams             | 0          |
| Opened_tables            | 44600      |
| Questions                | 2026873    |
...
| Table_locks_immediate    | 1920382    |
| Table_locks_waited       | 0          |
| Threads_cached           | 0          |
| Threads_created          | 30022      |
| Threads_connected        | 1          |
| Threads_running          | 1          |
| Uptime                   | 80380      |
+--------------------------+------------+
```

Figure 3.18. Show status sample output

DB CONNECTION SETUP

This menu option allows the users to specify details that are required for JDBC connection. The parameters like user name, password and database URL are required for successful JDBC connection.

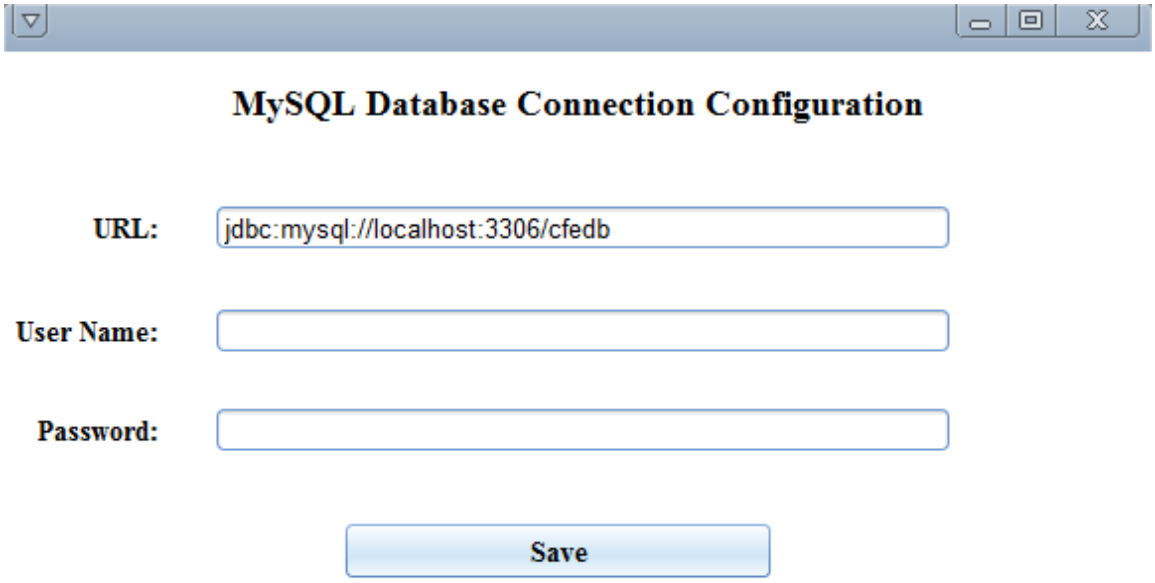


Figure 3.19. Database connection setup screen shot

Help Menu

This menu contains two submenus, about menu item and Doc and Support menu item.

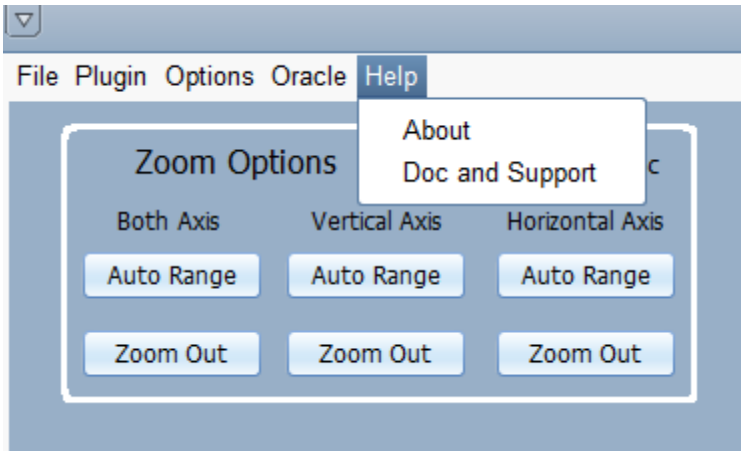


Figure 3.20. Help Menu screen shot

ABOUT

About menu item when clicked opens a popup dialog box.

The dialog box displays the application name, version, help and support contact details.

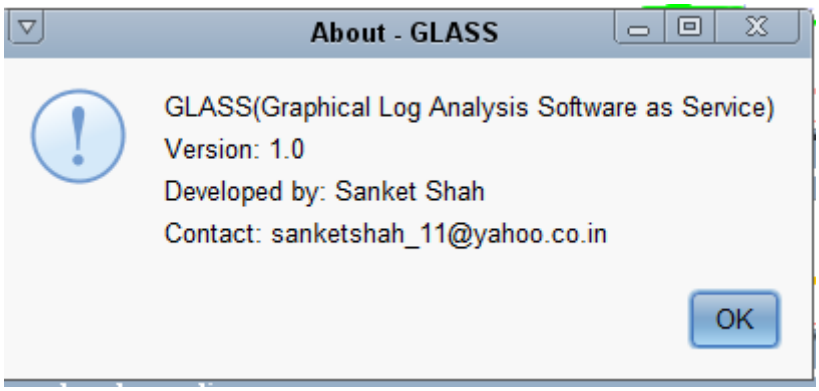


Figure 3.21. About menu option pane screen shot

DOC AND SUPPORT

Clicking on this button opens up a PDF describing how to use this application.

GRAPHING TOOL BAR

There are four graph icons provided for use with each graph. They are “+”, “-”, “H”, and “P”, with names Increase Graph Size, Decrease Graph Size, Histogram and statistics, and Pie Chart. These icons are present for each Graph Area populated, and map to the corresponding graph area. They perform specific functions on the graphs and its data.

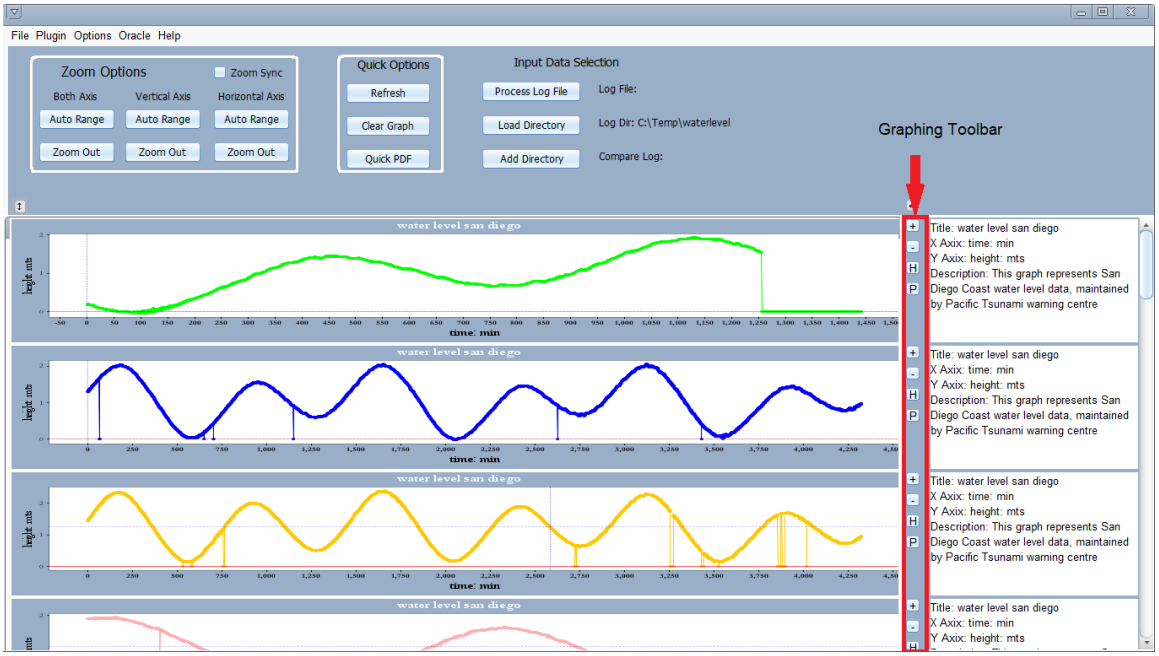


Figure 3.22. Application screen shot highlighting the graphing toolbar in red box

Increase Graph Area or the “+” icon

Not to be confused with zoom in and zoom out buttons, this button is used to increase the vertical component size of the selected graph by 50 pixels. It can be used a multiple number of times and there is no maximum limit for the size of a graph. The best advantage it provides is increased visual clarity of the plotted graph.

Decrease Graph Area or the “-” icon

Not to be confused with zoom in and zoom out buttons, this button is used to decrease the vertical component size of the selected graph by 50 pixels. It can be used a multiple number of times but there is a restriction on the minimum size of a graph to be 100 pixels. So after the graph height is reduced to 100 pixels, it cannot be reduced any more. The best advantage

it provides is to compare multiple graphs plotted one below another. It helps the user to view multiple graphs besides each other.

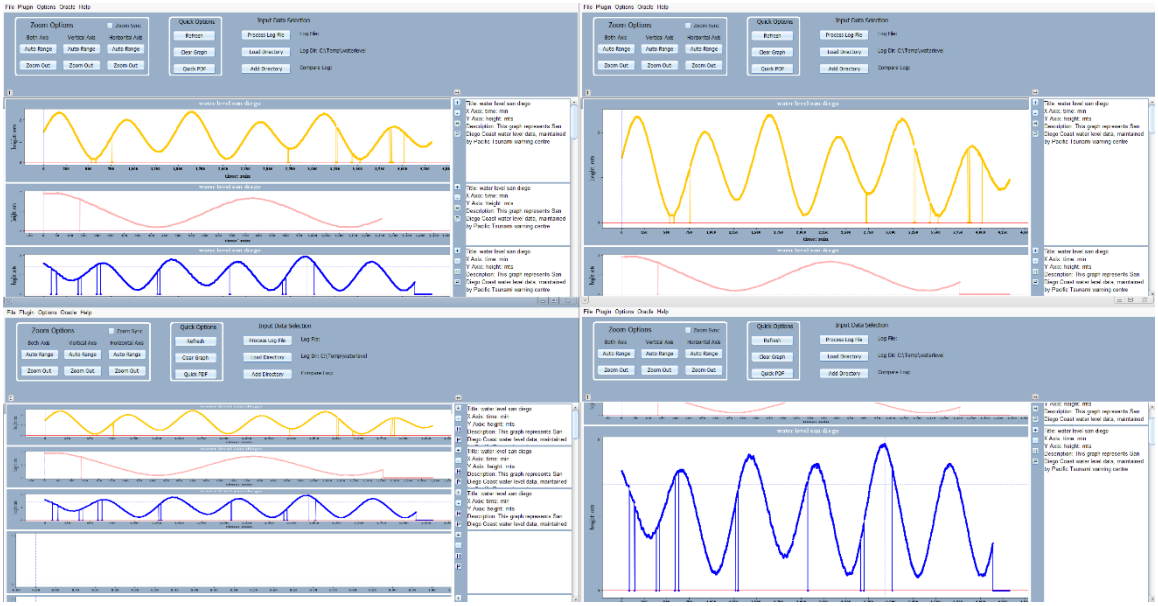


Figure 3.23. Multiple screen shots displaying the working of increase and decrease graph area button

Histogram and Statistics or the “H” icon

This button when clicked generates a Histogram out of the graph plotted and opens it up in a new window. It uses the coordinate data present in the graph to convert it to a histogram depending upon the type of the graph involved. For a single XY line graph, the data is divided into ten separate bins and plotted as a histogram. The histogram window also displays information like Maximum, Minimum, Mean, and Standard Deviation, calculated on the values of the plotted graph. More information about the histogram window will be discussed in a separate section later. The advantage of a histogram is in the situation where the visual distribution of the set of data is required. That is when the log analysis requires knowledge of frequency of occurrence rather than sequence of occurrence.

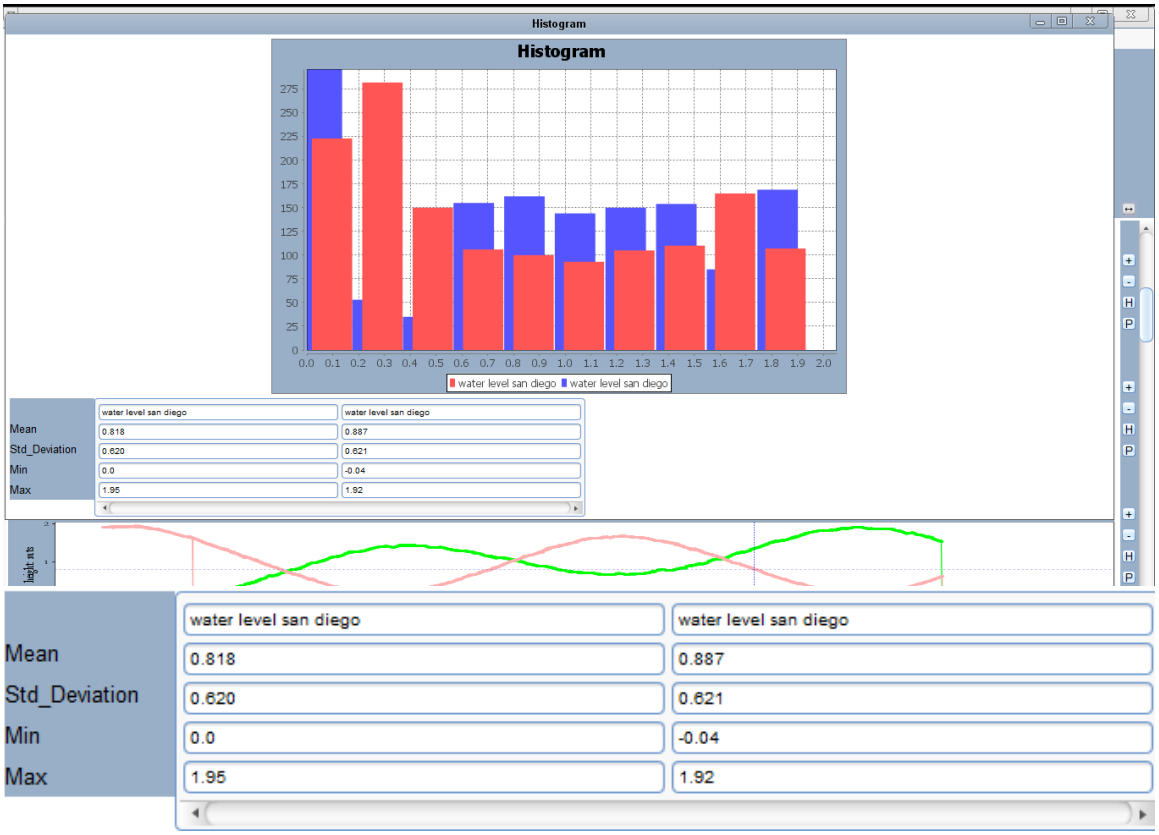


Figure 3.24. Histogram and statistics output when used for multiple overlaid plots.

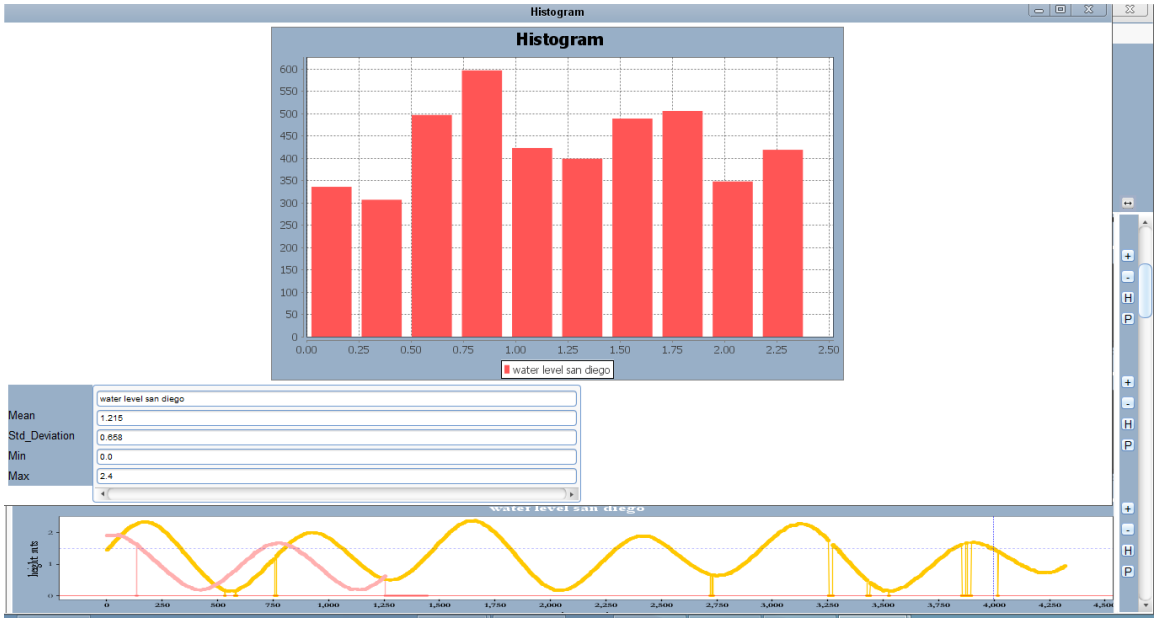


Figure 3.25. Histogram and statistics output when used for a single plot.

Pie Chart or the “P” icon

This button when clicked generates a Pie chart out of the graph plotted and opens it up in a new window. It uses the coordinate data present in the graph to convert it to a pie chart depending upon the type of the graph involved. For a single XY line graph, the data is divided into six bins and plotted as a pie chart. Here five bins consist of the first five modes present in the data and the sixth bin is the remaining points, named others. That means the first bin is the percentage of the value that occurs the maximum number of times in the data, and the second bin signifies the second most occurring value in the data, and so on till five. The sixth bin displays the percentage of the remaining values.

For a graph that contains multiple data lines plotted in parallel, the pie chart displays the respective percentage of the number of data points contained in each graph. Of course individual distribution of each set of data can be viewed by plotting it separately and generating a pie chart out of it. The advantage of a pie chart is in the situation where the visual distribution of the set of data is required. That is when the log analysis requires knowledge of frequency of occurrence rather than sequence of occurrence.

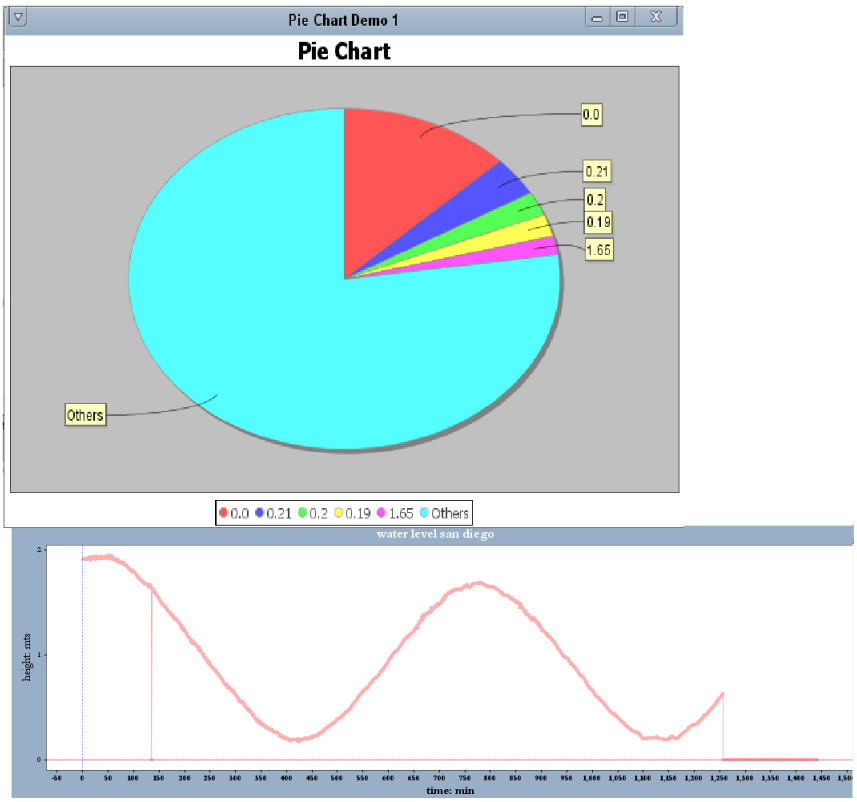


Figure 3.26. Pie chart output when used for a single plot.

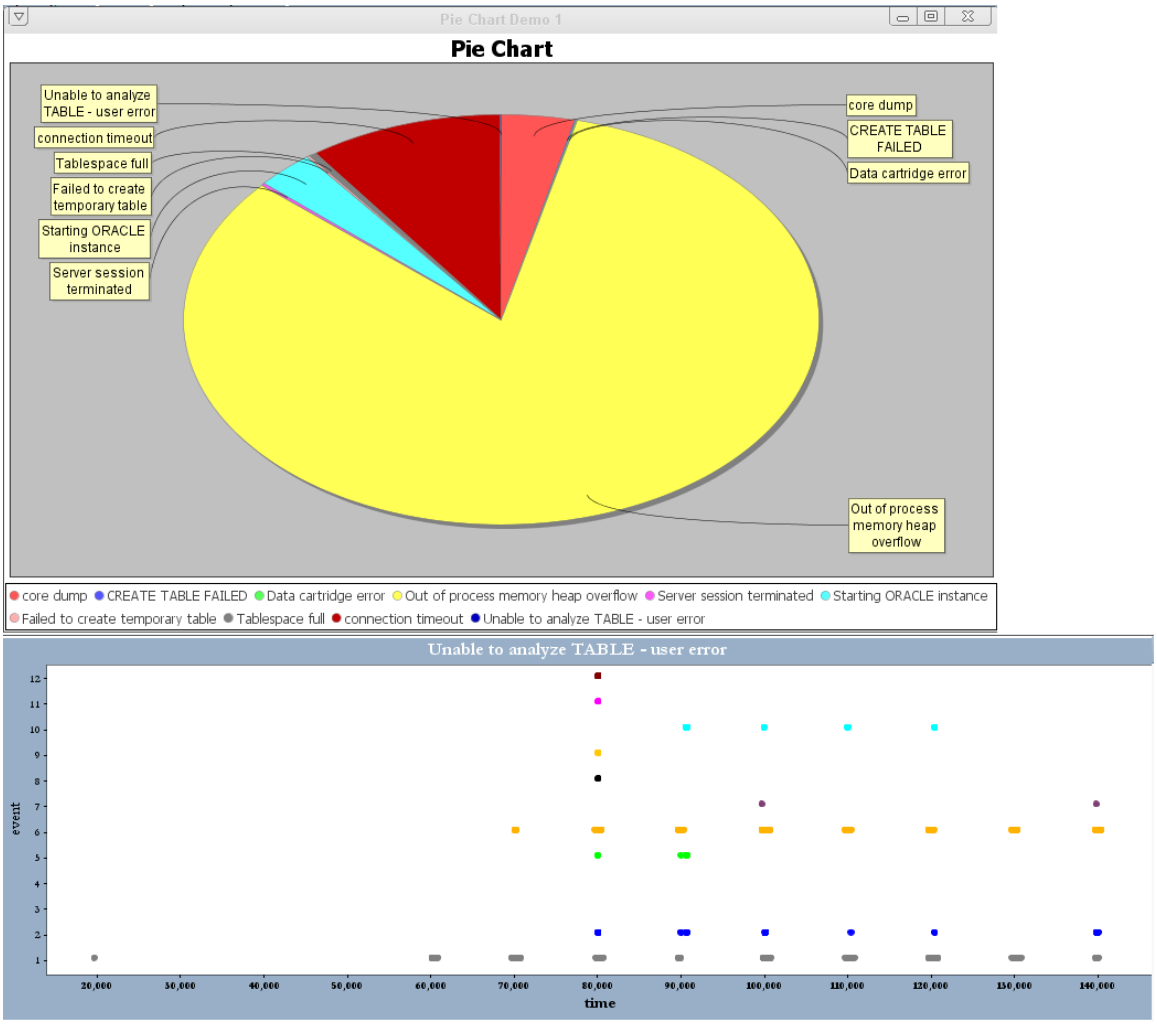


Figure 3.27. Pie Chart output when used for multiple overlaid plots.

3.4 GRAPH POPUP MENU

The graph popup menu is accessed via right clicking on any graph.

It contains the following options.

1. Properties
2. Copy
3. Save As
4. Print
5. Zoom in

6. Zoom out
7. Auto Range
8. Add Graph
9. Remove All

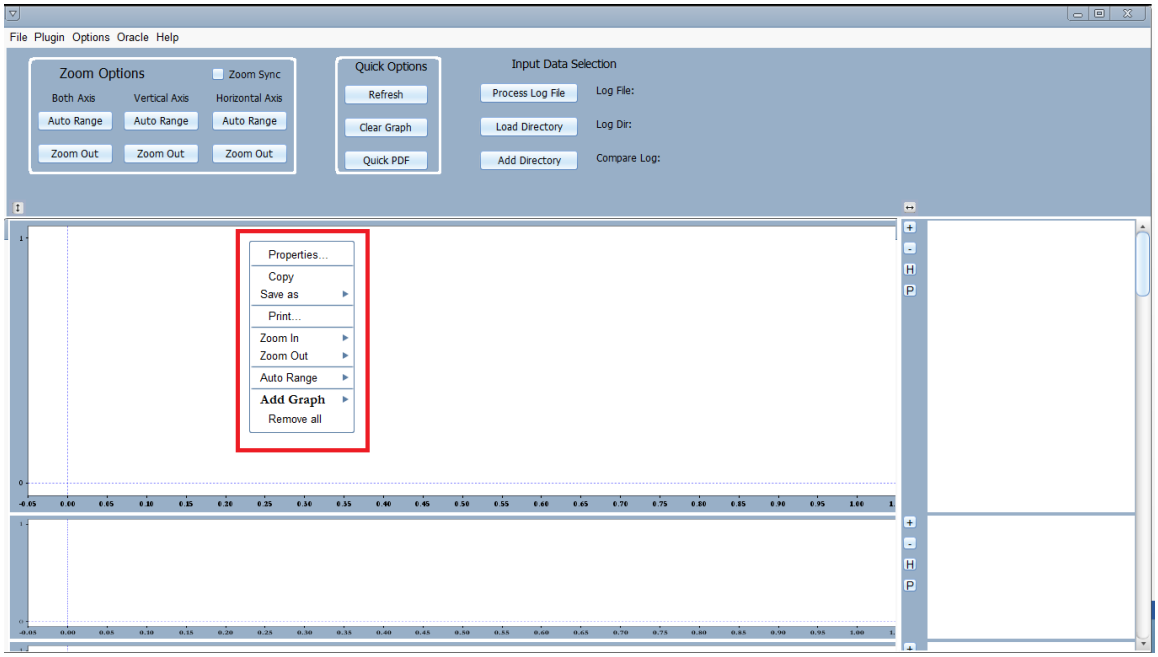


Figure 3.28. Application screen highlighting popup menu in red box.

3.4.1 Properties

Selecting this option opens up a menu to customize the currently plotted graph.

There are options available to change font type, font size, and color of the legend labels, and title. Background paint, series color grid color and other similar features can be altered for better presentation purposes. There are a lot of options to play with to completely customize the look and feel of the graph without altering the graph values.

3.4.2 Copy

Copies the current graph as an image to the clip board. It is similar to taking the screenshot of the selected graph.

3.4.3 Save as PNG

Opens a file saving popup menu to save the selected graph in PNG format.

3.4.4 Print

Opens a print request and a page setup popup as shown in the figure below. The selected graph along with labels is sent to the printer; this is similar to sending the saved image to print.

3.4.5 Add Graph

Add Graph is the most important option of all, in fact the most important GUI feature. This allows the user to select the graph he wants to plot in the selected area. Users can dynamically add multiple series to the same graph, allowing them to analyze the data the way they want. The same series can be plotted at multiple locations simultaneously with different series layered upon it, allowing the user to compare the same graph with multiple different graphs. There is a restriction on the total number of series that can be plotted in one particular area to 10, and there are 20 such areas available in parallel. This allows users to plot 200 different graphs at a single time, simultaneously. The graphs are differentiated by the path of the file they are being plotted from. This allows the users to plot and compare the graphs or files with same names and similar data, providing no collision during runtime and allowing seamless comparison.

3.4.6 Remove All

Remove all option allows the user to remove all the currently plotted lines/plots on the selected graph.