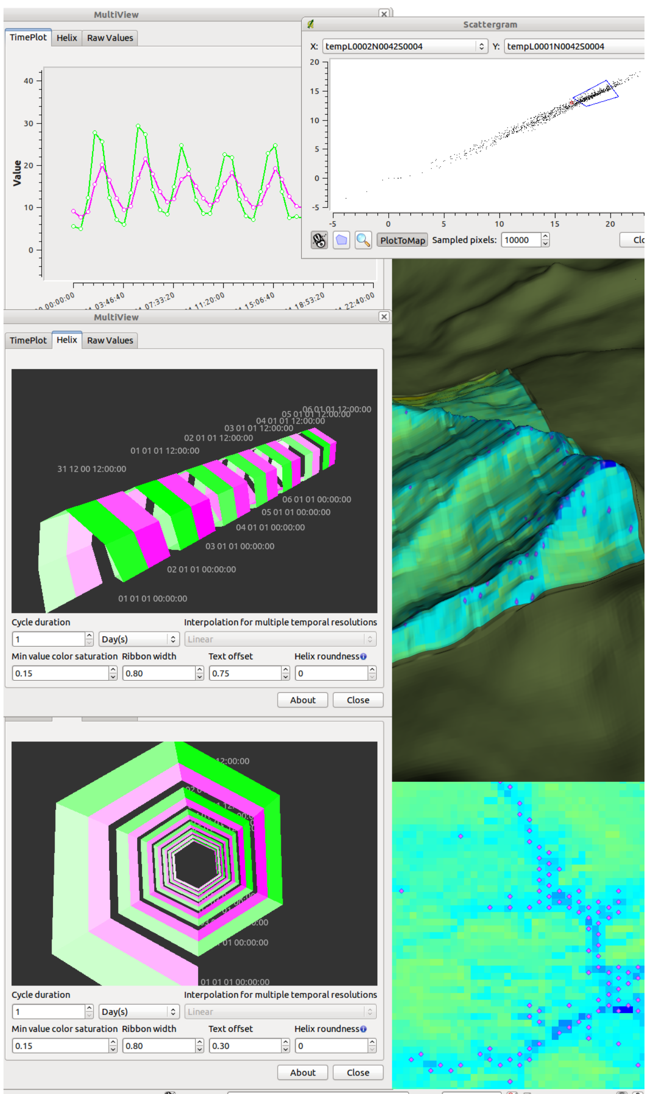
University of Zurich
Department of Geography

Visualising multivariate spatio-temporal data

Design and evaluation of an open source software environment based on 3D DEMs and multiple linked views



GIScience Center:
Geographic Information Visualization and Analysis (GIVA)

Master of Science Thesis [GEO511]
Delivery: 30.04.2011

Marco Bernasocchi
of Sagogn GR, Switzerland

Student-ID: 02-730-620
marco@bernawebdesign.ch

Advisor: **Dr. Arzu Çöltekin**
Co-Advisor: **Dr. Stephan Gruber**

Faculty Member:
Prof. Sara Fabrikant, PhD
GIScience Center
University of Zurich
http://www.geo.uzh.ch

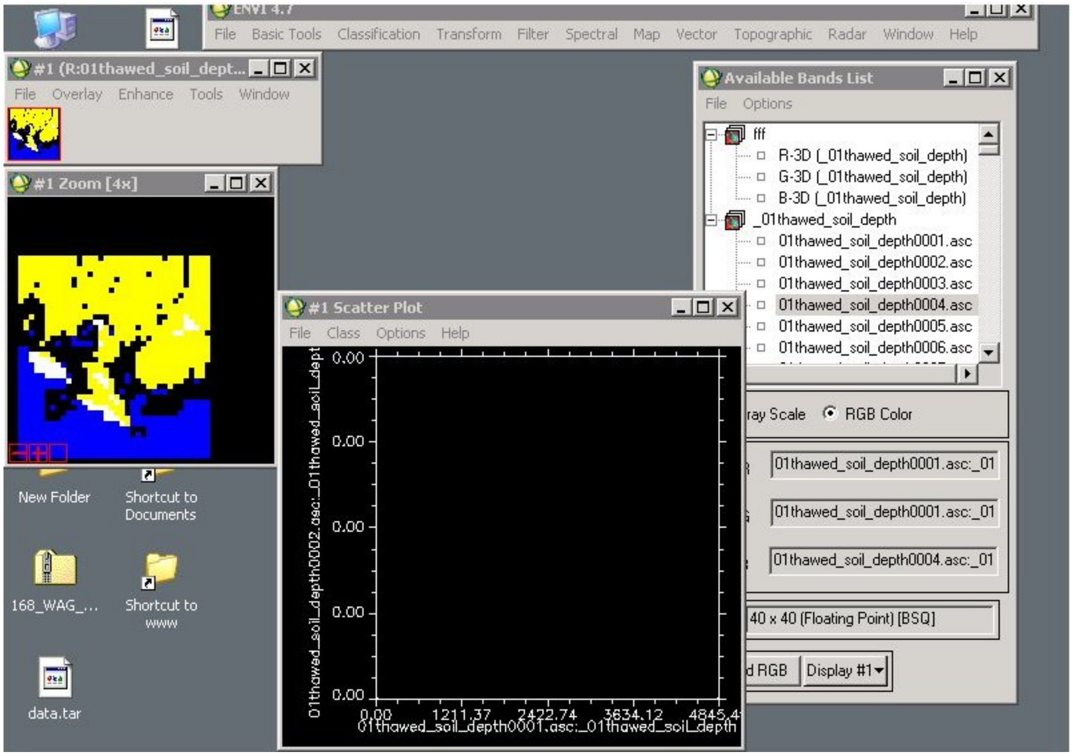


Figure 1.1: Envy, the main visualisation tool currently in use at GIUZ

map and a time-slider linked together. The analysis is then performed by moving the time cursor around different periods. Direct comparison is only available by opening two concurrent sessions of envy. The necessity of a better visualisation system to allow a thorough analysis of the data has risen with the growing of the dataset produced with the geoTOP model and is one of the main concerns addressed in this thesis.

The visualisation of multivariate spatio-temporal data is currently hot topic in visualisations research and different groups have published taxonomy works on how to categorise such visualisations based on the data that need to be analysed. A very good overview of all available techniques is given in Aigner et al. [2007, 2008], Andrienko et al. [2003] and Daassi et al. [2006].

Innovative visualisations such as Space-Time cube, Multi-variable-Time cube (MTC) and PCP-Time-Cube (PTC) [Li and Kraak, 2005], helix and pencil glyphs [Tominski et al., 2005], timeweel [Tominski et al., 2004], enhanced spiral display full. User controls for panning and zooming become in such cases indispensable [Grinstein et al., 2001].

2.1.2 Line graphs and multiple line graphs

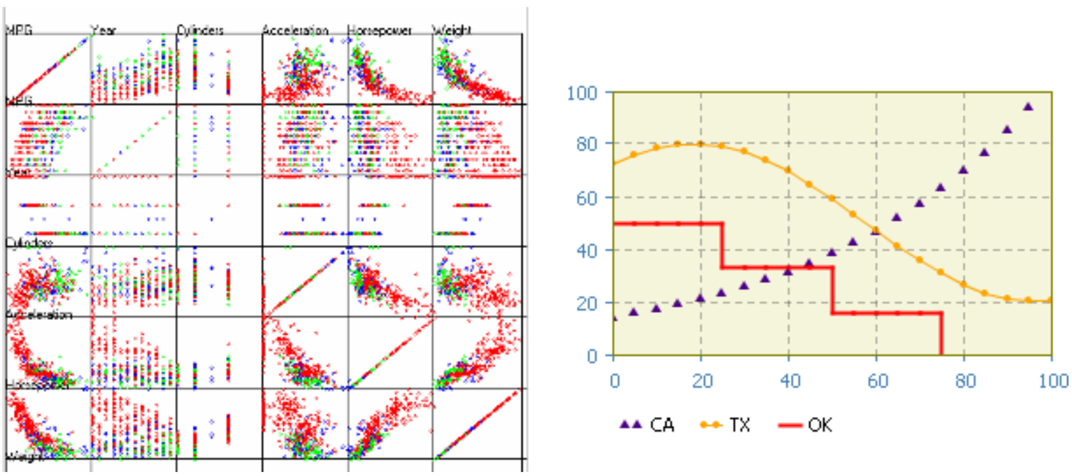


Figure 2.1: Scatter plot matrix and Time plot [sources: Grinstein et al. [2001]; `http://www.xjtek.com`]

According to Grinstein et al. [2001], multiple line graphs are used in multivariate analysis to display data that is dependent on one variable. Usually, the values of this variable are unique like (ID values or time-stamps). Despite the immediate readability of such visualisations, three main problems have been identified when working with many dimensions and need considerations when designing such tools. First, each variable needs to be displayed using a line with a different style. Second, the value scale of each variable needs to be present which might lead to confusion depending on the scale and whether or not an offset is used. If the absolute values are not relevant, this problem can easily be overcome by using data normalisation.

2.1.3 Parallel Coordinates Plots

Inselberg [1985] first described the plane with parallel coordinates, in his work he

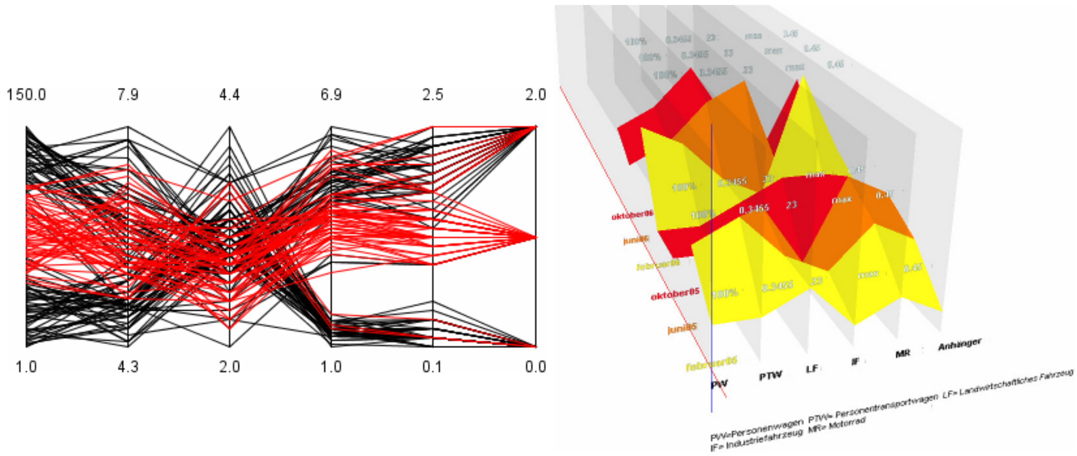


Figure 2.2: Parallel coordinates plot and Time extended parallel coordinates plot [source: Hautle [2010]]

presented a novel mathematical visualisations where high dimensional data can be plotted. Each N-dimensional point is no longer represented by a point but by a line. Each of the vertical parallel axes is the scale of a variable and where a line reaches it, it indicates the value of the sample for the specific variable. Twenty-five years after their formal introduction, PCPs have reached a well respected position among the multivariate visualisation and keep being a reference a base for novel visualisations [Hautle, 2010; Inselberg, 2009].

2.2 Multivariate temporal data visualisations

Temporal data visualisation is a strongly investigated topic in visualisation design. Many authors continuously develop and publish new approaches on the subject. So many different approaches have been published that different taxonomies, to help classify temporal visualisation methods, had to be created [Muller and Schumann, 2003; Daassi et al., 2006]. Not all of these methods are suitable for the analysis of multivariate temporal data or of interest for this project and therefore the selection of methods presented here is much more limited.

In this thesis temporal data is subdivided into two main groups, *linear* and *cyclical* data. We will consider linear data, datasets where cyclicity might be

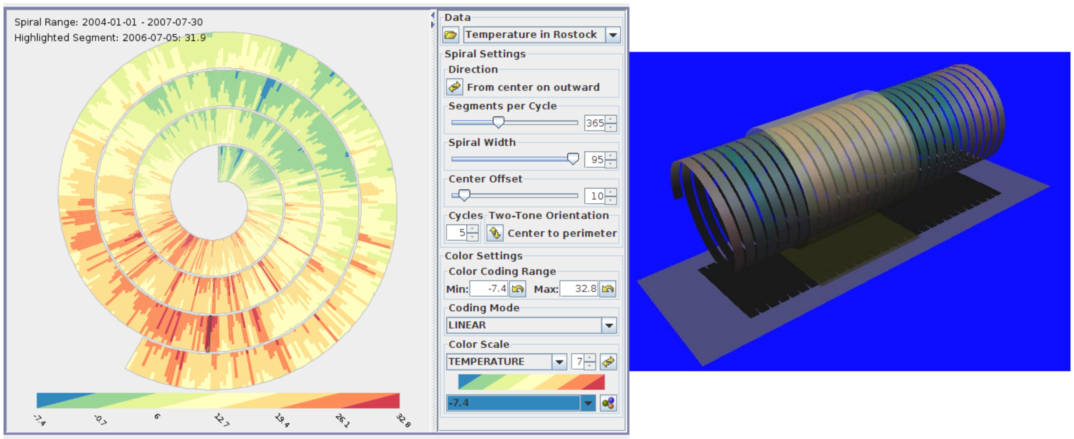


Figure 2.3: Enhanced Interactive Spiral Display and 3D Helix [sources: Tominski et al. [2004]; Weber et al. [2001]]

dimensional helix Weber et al. [2001]; Tominski et al. [2005] that allows direct interaction with the data. Weber et al. [2001] note, that the helix visualisation can be difficult to interpret when the helix is seen from the side since not all the values are directly visible.

2.2.3 Axes-Based Visualizations with Radial Layouts

Two further methods basing on radial axes are proposed in Tominski et al. [2004], the TimeWheel and the MultiComb [see figure 2.4]. The idea behind the TimeWheel is to present the time axis at the centre of the visualisation and to arrange the other variables' axes around it and to assign each axis a unique colour. The points on each of the axes are then connected to the time axes to show the relation between the axes similarly to PCP. The two parallel axes are the easiest ones to interpret and the two perpendicular ones the hardest, thus the TimeWheel allows an interactive rotating of the axes. Further techniques such as colour fading and axes length adjustment are used to avoid over-clogging the visualisation.

The idea behind the MultiComb is to align a series of plots in a circular form. Each of the plots (i.e. timeplot) are so comparable to each other by analysing the expressiveness of each plot. Both approaches have been extended to use the three dimensional space [see figure 2.5] in form of the 3D TimeWheel Aigner et al.

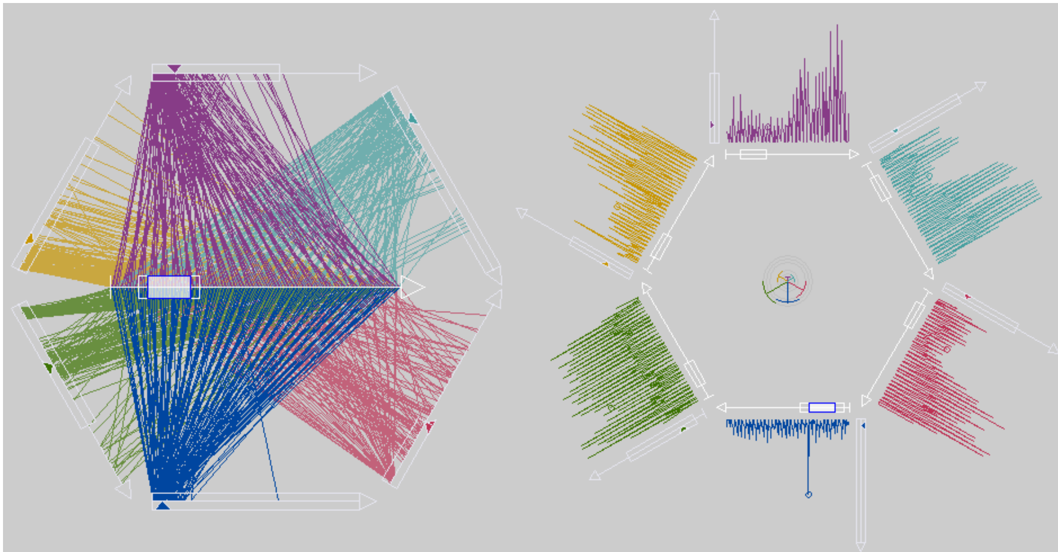


Figure 2.4: TimeWheel and MultiComb [source: Tominski et al. [2004]]

[2008] where the axes are all placed parallel to the time axis and the Stardiagramm Noirhomme-Fraiture [2002] where each plot is perpendicular to the time axis.

2.3 Spatio-temporal data visualisations

Spatio-temporal data can be separated in two main categories, movement data and static data. While in movement data the sensors move over time (such as in wildlife tracking), in static data the sensor do not move. This type of data is the typical outcome of environmental monitoring projects.

Several visualisations exist for movement data [Kjellin et al., 2008; Andrienko et al., 2003] including: maps with timestamped objects, Space-Time cube [Kraak, 2003], Multi-variable- and PCP-Time-Cube [Li and Kraak, 2005] as well as animation. All these techniques but animation are, due to their intrinsic connection to point data, useful only for movement visualisation and are therefore irrelevant to this project. Animation is an exception since it can be used to visualise both movement and static temporal data [Andrienko et al., 2003].

More relevant to this thesis are visualisations like pencil and helix glyphs

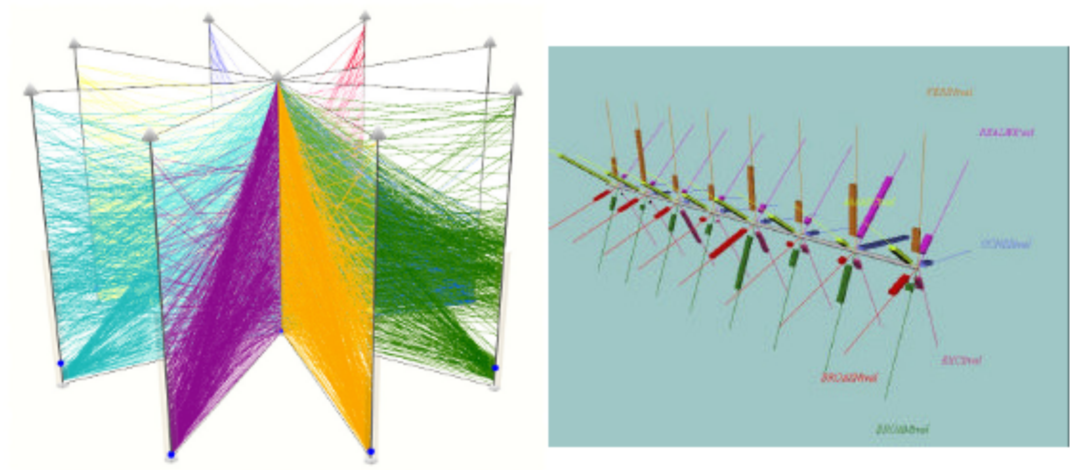


Figure 2.5: 3D TimeWheel and Stardiagramm [source: Aigner et al. [2008]; Noirhomme-Fraiture [2002]]

[Tominski et al., 2005] placed as objects on a map, change maps [Andrienko et al., 2003] or all the multivariate temporal data visualisations discussed in section 2.2 if used in a linked views context as explained in section 2.5. All these visualisations allow displaying changes over time for specific locations. On the one hand, multiple glyphs placed directly on a map as in Tominski et al. [2005], allow direct comparison of temporal data at different locations. On the other hand, the use of the tools described in section 2.2, in combination with a location input view, allows the exploration of the dataset of one location at a time. Further, even more complex visualisations tools that integrate data mining capabilities are being studied and developed [Compieta et al., 2007].

2.4 3D digital terrain model visualisations

The evolution of digital globes since the first released tools [Riedl, 2005] has been exponential. The release of tools such as Google Earth and NASA World Wind have made digital globes become increasingly popular. Nowadays, it has become a routine to check vacations destinations and personal whereabouts in such tools. When ten years ago explaining what GIS was to a layman involved long expla-

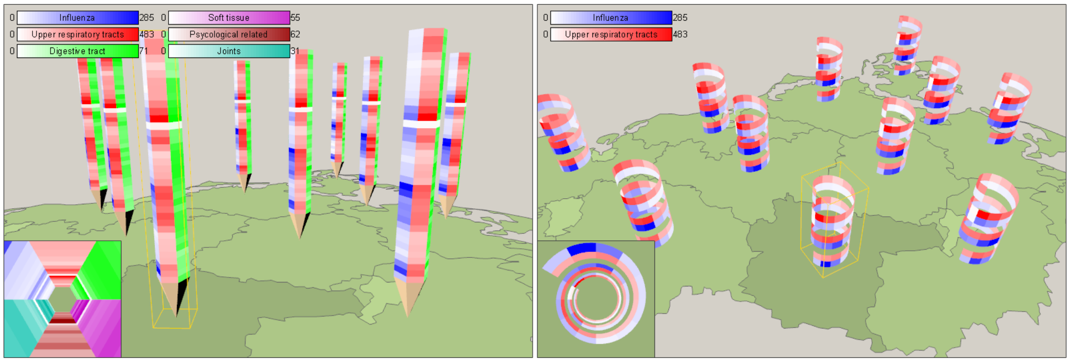


Figure 2.6: Pencil and helix glyphs places on a 3D map[source: Tominski et al. [2005]]

nations, it is now possible to give almost anybody in the western world a quick idea of what GIS is all about by mentioning one of these tools. A further, much more scientifically relevant point is the fact that many researchers started using these tools to display their research results [Bleisch and Nebiker, 2008; Butler, 2006]. Several studies have investigated the usefulness of 3D globes as part of linked view tools. The results are controversial. On the one hand researchers say that the 3D environment enhances the readability of the observed landscape and its features [Bleisch and Dykes, 2008; Jones et al., 2009]. On the other hand, the augmented cognitive load is cited as the main negative factor in navigating such environments [Rase, 2003]. In a study regarding 2D/3D integration, Bleisch and Nebiker [2008, p:1] conclude that

> "It seems that connecting 2D data displays to the 3D views and updating or changing them dynamically allows overcoming some of the shortcomings of using stand-alone 3D views of information. This includes difficulties with navigation, occlusion of content, lower information density or projective distortion of the display. We assume, supported by combined 2D/3D displays that have already been tested as mentioned in the introduction and by our own experiences with the prototypal implementation described in this paper, that combining 2D displays with data representations in 3D virtual environments will certainly lead to new possibilities and ways for explorative analysis of spatial information. This may especially be true for data sets

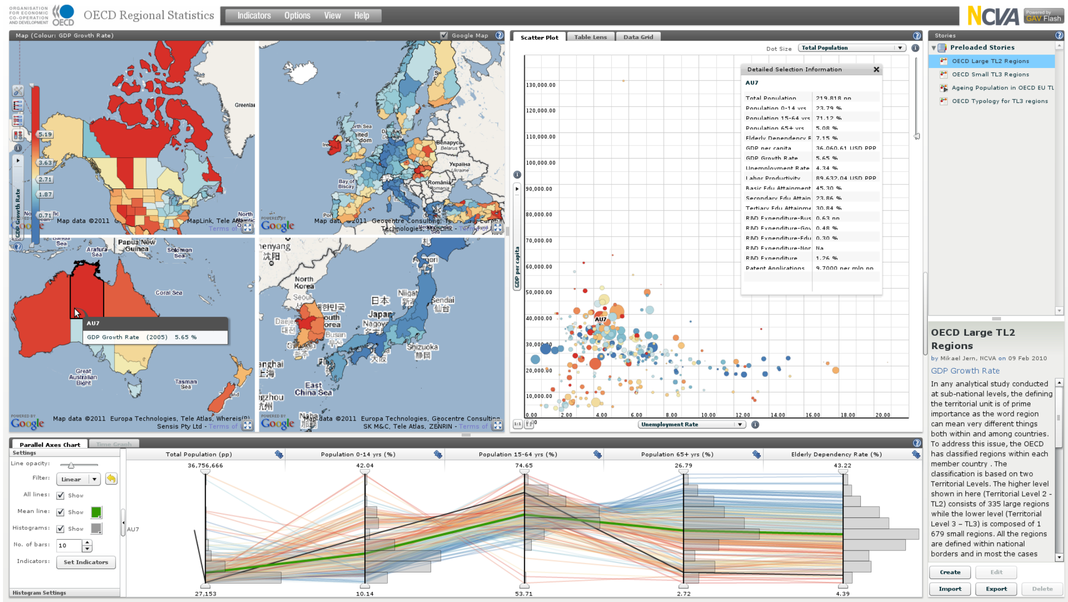


Figure 2.7: Linked views example [source: OEC [2011]]

the necessary amount of interaction, however, is not."

To help develop applications using multiple linked windows, Wang Baldonado et al. [2000] developed a set of eight guidelines that should allow a developer to assess if a multiple linked windows implementation is a sensible solution to his/her problem. Table 2.1 summarizes the guidelines giving a short description of each one and the main pros and cons of following or disregarding the indication.

2.6 Available tools

Many multivariate temporal data visualisation tools are available nowadays, and some of them deal with spatio temporal data as well. Most packages are focused towards visualisations and less toward mature GIS integration. The work-flow to use such tools requests there fore the separate usage of a GIS to prepare the data, an export/import step and finally the visualisation and analysis of the data with the chosen tool. Below a list can be found, added for reference only, of some

4.2 System architecture

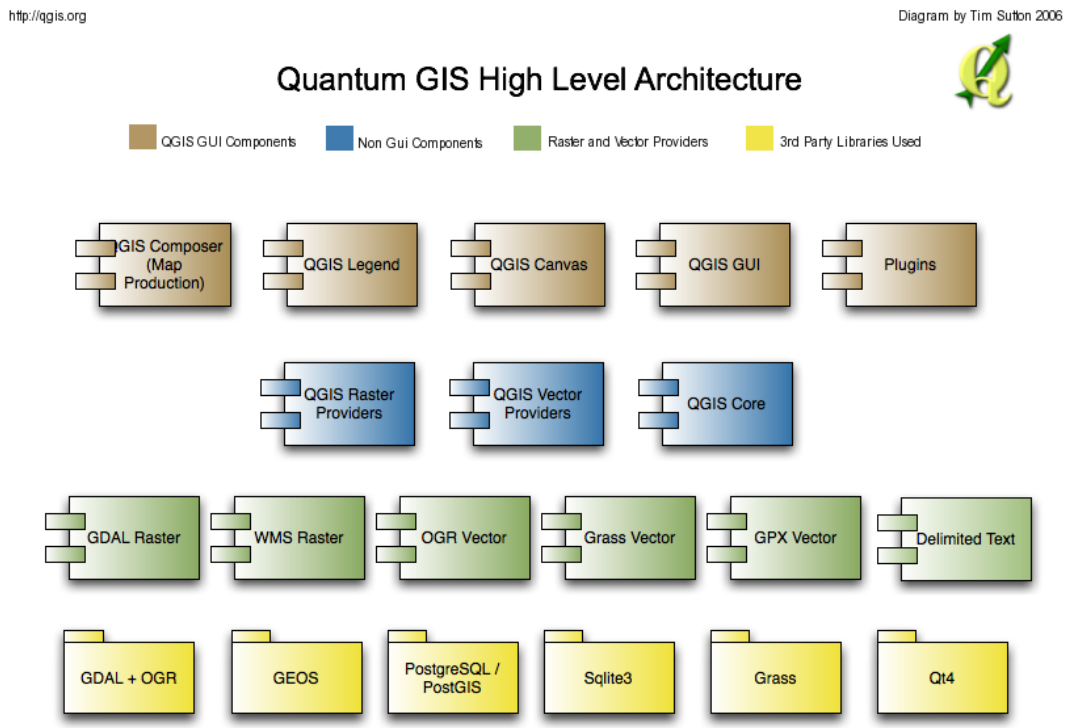


Figure 4.2: QGis high level archiutecture [source: `http://www.qgis.org`]

QGis leverages many libraries and frameworks to achieve its full functionality. In this section I will first explain the high level QGis architecture and then I will explain for what each component is necessary and how it relates to the others.

As figure 4.2 clearly shows, QGis is build of multiple modules that deliver specific functionality to the whole system. Five main GUI components manage the whole application: Map composer, Legend, Canvas, GUI and plugins. Map composer is responsible for the creation and printing of maps, the legend and canvas components are used to maintain the state of the data while the GUI manages the layout of the application. Finally a vast ecosystem of plugins contributes to QGis flexibility by allowing everyone to program custom plugins to suit their needs.

The main GUI components are independent and do not directly provide GIS

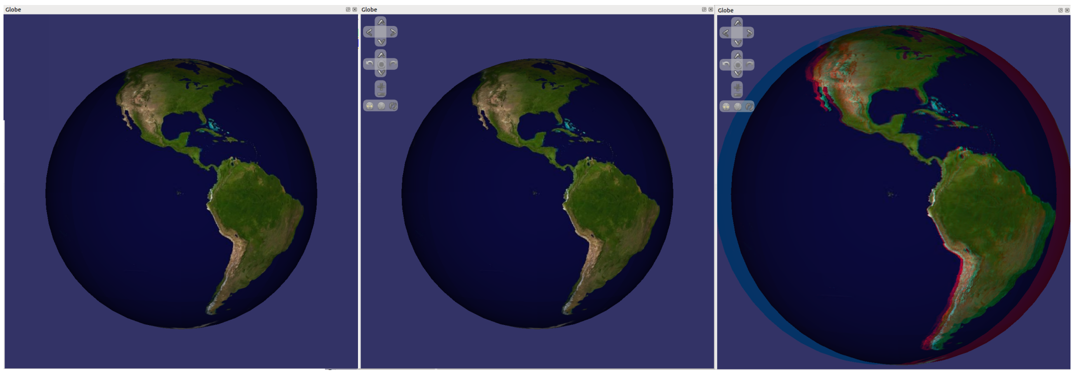


Figure 4.5: Evolution of the Globe plugin. Basic globe, globe with GUI, globe with GUI and stereo 3D

the high level architecture of the plugin highlighting where adapter classes had to be programmed to allow the OSGEarth globe to communicate with QGis. The used architecture bases on the wish to have the globe running as a Qt widget directly inside QGis and not to call it as an external program. This approach allows us leverage the QGis rendering engine for the globe as well which results in the same layer symbology as on map canvas being used automatically. The main drawback of this approach was the relative difficulty of implementing the adapter layers, but once this was done, the benefits were worth the effort.

As it can be seen in figure A.2 al the features listed early have been implemented in two main groups, on one hand on one hand all the ControlHandlers structs[25] take care of the on screen GUI controls that allow zooming, panning, tilting, rotating and syncing of the globe [see figure 4.5]. The control handlers all extend either the class NavigationControlHandler or ControlEventHandler from OSGEarth to take advantage of the built-in capabilities of the library. On the other hand, the QgsGlobePluginDialog deals with the runtime configuration of the plugin allowing the user to add or remove DEMs of various sources, formats[26] and resolutions, set stereo 3D settings, and as a still experimental feature, add 3D models for objects. The OSGEarth toolkit takes internally care of the generation of the terrain model, the textures and, if activated, the stereo pairs according to the user settings [see figure 4.7]. All stereo settings are saved in QGis settings and

[25] in C++ a struct is the same as a class except that its members are public by default
[26] http://osgearth.org/wiki/TileSourcePlugins

Figure 5.2: Dialog presented to users when an header file is found in the data import folder

As a direct result of this test task, the data loader has been improved by putting an explanatory text in the log area when the loader is opened. This should allow future users to easily understand how the loader works and make them more efficient. Also, two users noted that this is a minor issue since it only poses a problem during the first usage and that even without my input they would have, by default of other options, clicked on the correct button. Furthermore, the loader icon has been changed to a more meaningful icon as can seen in figure 5.3.

Figure 5.3: MultiView plugin, old dataloader and new dataloader icons

5.2.2 QGis layer visibility

When trying to identify temporal patterns using only the standard QGis layer visibility options the participants started switching on and off subsequent layers but mostly gave up within 3 minutes complaining that the method was very inefficient and that it was very hard to remember one time-step and compare it to the next or previous. Furthermore all the participants said that it was almost impossible to remember and compare more than two steps at a time. No user could identify any temporal pattern but the could correlate the variations of the variables values to the topography.

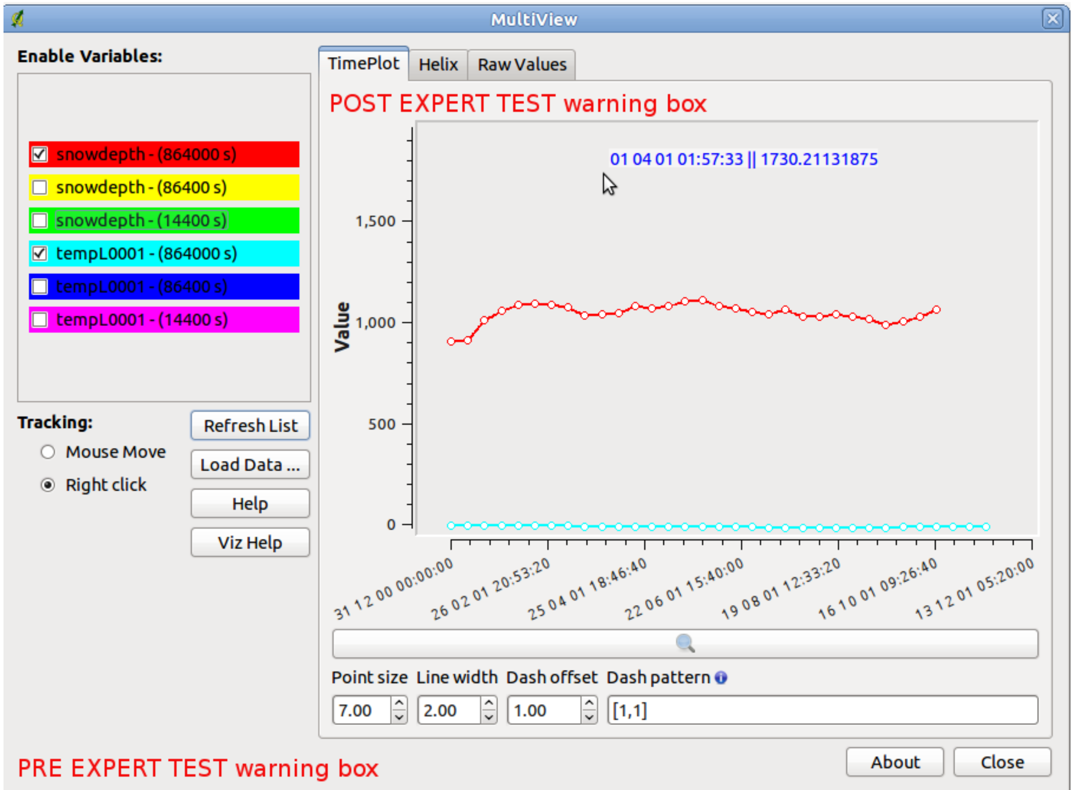


Figure 5.4: TimePlot visualisation as used during the expert testing with added pre and post expert testing warning box

of the plot. This axis only shows one value scale for all the active variables and thus if the range of a variable is for example 50-3000 and for another it is 0-10, the differences in the second variable will be flattened out by the big combined range (0-3000). This was a known problem and already a top development priority. Thanks to the user feedbacks some alternative solutions to this problem have been found and will be implemented in the near future. Two users proposed to use multiple axes (one for each variable type) but then agreed that having more than two-three Y axes on a plot would render it very difficult to read. The solution that will be implemented will allow the user to toggle Y value normalization so that the different variables can be better analysed for scale independent trends. Further, the tool should allow to display the real value on demand using the cursor tooltip.

Subsequent locations' analysis: Currently the time plot tool allows the analysis of one location at a time (or multiple location by moving the mouse on the map canvas) for multiple time steps and multiple variables. It was suggested by the participants to allow spatial analysis by leaving the last visualised graph's lines on the plot and setting their appearance to dashed or dotted lines [see figure 5.5] so that a direct comparison of two different locations could be performed much more easily. This feature still requires a substantial amount of work and is therefore planned for future development.

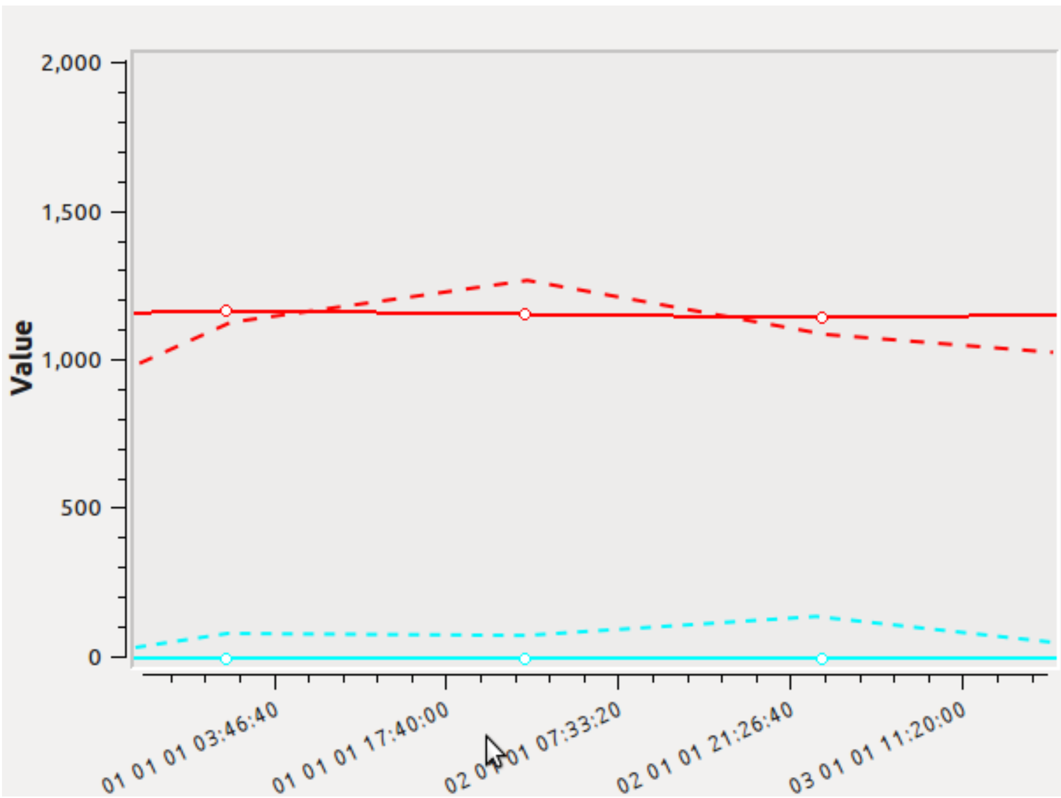


Figure 5.5: TimePlot visualisation sketch with added last selected values lines

Time axis labels reduction: The amount of date labels on the time axis was found to be too high by most participants. Furthermore, the experts suggested to put the labels only where data was present instead of at regular intervals. This feature was considered very important by the participants and he work load for the implementation was estimated to be acceptable so the feature was imple-

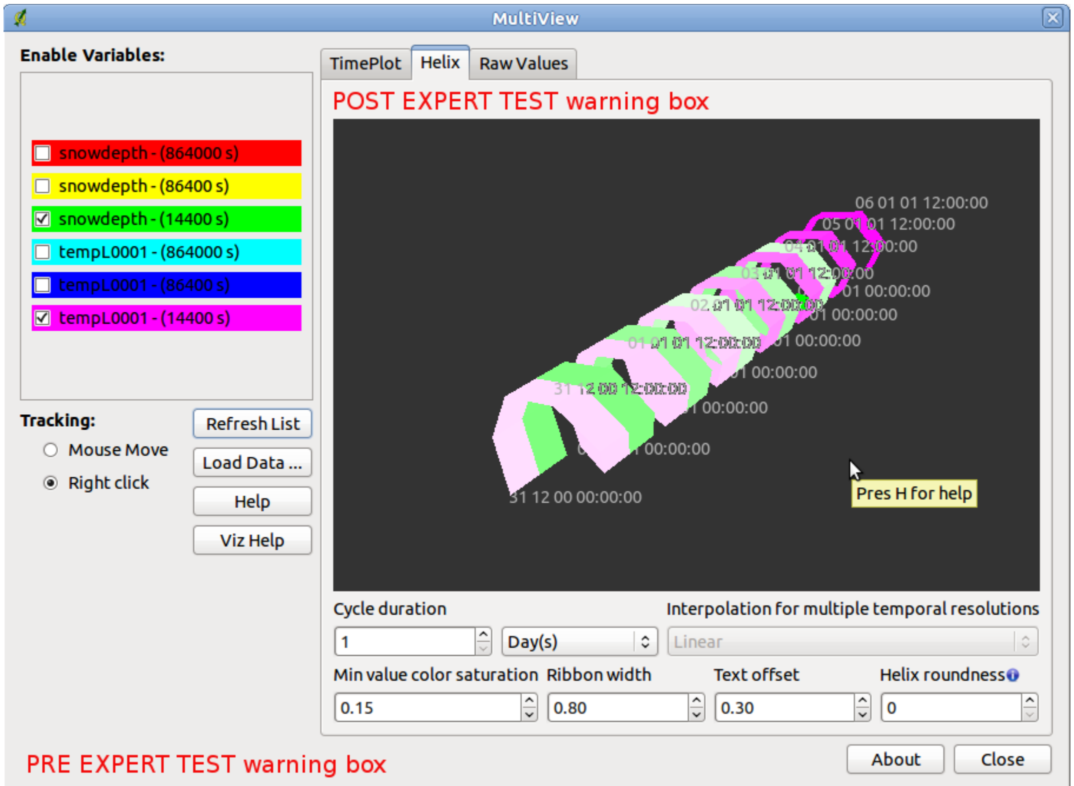


Figure 5.6: HelixView visualisation as used during the expert testing with added pre and post expert testing warning box

a very intuitive way of displaying data and that it was very easy to identify patterns in the data. Indeed, the participants performed very well when asked to identify patterns using this visualisation. This is a very satisfactory result since one of the aims of this project was to create a tool that would allow to replace the currently used visualisations.

The general opinion about HelixView was that it is difficult to understand how it works at the beginning but might be very helpful for cyclic data. The users could not easily identify complex patterns immediately, but easier patterns like the diurnal cycle were easily discovered. All testers stated that TimePlot was their preferred tool but that they would like to try using HelixView with data they know so to bee able to exactly understand how it works and what it's potential is.

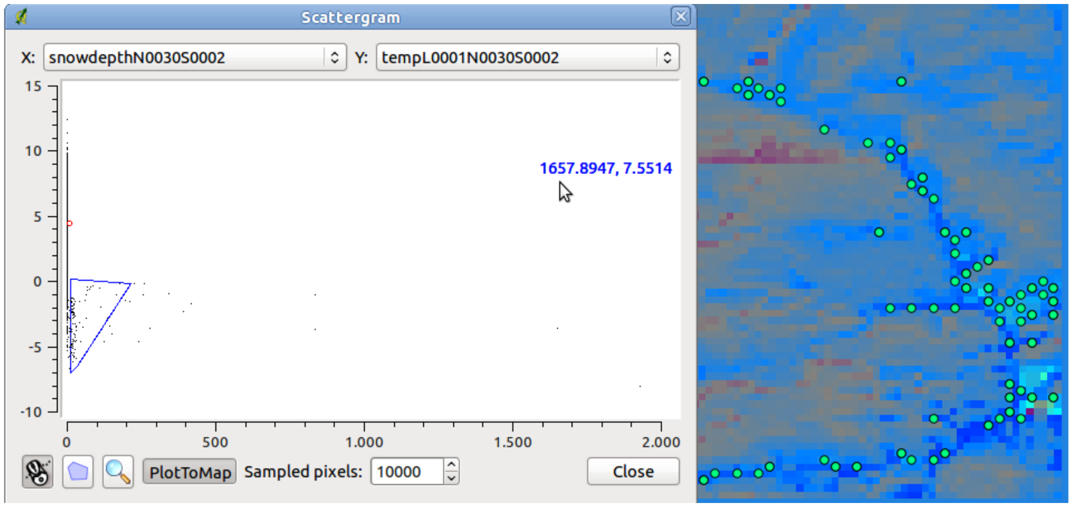


Figure 5.7: Scattergram visualisation used during the expert testing

5.2.7 Globe

The last task the users were asked to perform was a superficial evaluation of the UI of the Globe plugin, by telling what they thought each button of the UI would do. This task had to be kept minimal since at the moment of the test, due to versions incompatibilities, the plugin was running incredibly slow and only on an old branch of QGis. As seen in sections 4.3, 4.4 and 4.5 the globe needs a multi-threaded version of QGis which nowadays is a branch of QGis 1.4 but MultiView and some functionalities of Scattergram require at least QGis 1.7.

Most users found all the buttons to work as expected. The only button that was confusing the users was the one to reset rotation and tilting to perpendicular view with north above. As it can be seen in figure 5.8 this button was the same as the tilt down button and thus it was hard to understand if it was a bug or actually a button with functionality. The button has been changed to reflect the testers suggestions.

It was very unfortunate that the participants could not test the interaction with globe, since previous undocumented quick tests on friends had shown positive results.