

Teaching FOSS-GIS at Technical University of Vienna – gvSIG vs. QGIS

Johannes Scholz, Paul Weiser, Amin Abdalla

Vienna University of Technology, Institute of Geoinformation and Cartography, 1040 Vienna

1. INTRODUCTION

The Institute of Geoinformation at the Technical University of Vienna offers a wide variety of GIS-courses at both Bachelor's and Master's level, traditionally embedded in the surveying curriculum. Currently, we offer four courses (lecture and lab) at Bachelor's level, directly related to Geoinformation Science and Technology, i.e., Introduction to Geoinformation, Advanced Geoinformation, Implementation of a GIS, and Feasibility Study for GIS. At Master's level, we also offer courses on the mathematical theories underlying GIS (GI-Theory), Ontology, Database Systems, and functional programming with Haskell.

We are anxious to provide our students not merely a training in a specific software product but to offer them an education that emphasizes fundamental GIS-principles. For a discussion of the differences between education and training in GIS see Moore (1998) and Longley et al. (2011). Most of the curriculum uses free and open-source software (FOSS) to achieve this goal. Open-source GIS-software allows students to experience such systems without the need to acquire costly licenses. Also, changes to the software can be made, should improvements or adoptions to a particular application be necessary.

In this article we discuss and evaluate two popular open source GIS-packages, namely gvSIG¹ and Quantum GIS² (QGIS). The goal is to compare and analyze both systems based on various geo-analysis tasks, typically taught in a first year's course on GIS. All presented examples and tasks were part of the lab section of an introductory GIS-course taught in the summer term 2011. This work discusses whether one product is to be favored over the other in terms of “ease of use” and suitability for teaching. Particularly emphasized is the fact that students had no prerequisites in using such a system.

The paper is organized as follows: First both software packages are introduced, including their add-ons and extensions. This is followed by a listing of exercises and tasks that were solved in the lab, utilizing both GIS-packages (Section 3). Section 4 talks about experiences made by both lecturers and students while working on the exercises. In the last section (5) we evaluate both software products in respect to their use in higher-level education.

2. UTILIZED OPEN-SOURCE GIS-SOFTWARE

This section describes the two GIS-software products used to carry out the evaluation. In addition, the add-ons and extensions used in this work are listed, ensuring repeatability of the results. All analyses were performed on a PC with an Intel Atom Processor (2 GB RAM), Windows 7 Starter N (32 bit), and Java VM 1.6.0 installed.

¹ gvSIG and extensions are available via <http://www.gvsig.org/>
² QGIS is available via <http://www.qgis.org/>

The open-source GIS-software gvSIG is written in Java and largely funded by the Spanish government (Acevedo, 2011). QGIS is written in C++ and based on the GRASS toolkit (Athan et al., 2011). Both systems provide support for raster and vector analysis and are widely extensible via plug-ins. Also, both software products are distributed as free software under the GNU distribution license, therefore free to study, develop, and improve. See Table 1 for details on the versions and extensions we used in our comparison.

Table 1: Utilized software, including extensions

gvSIG	Quantum GIS
gvSIG 1.10 (Build 1264) with extensions: - Sextante Toolbox 0.6.0-1232 - Network Analyst Redes 0.1.0-1238	QGIS 1.6.0 Copiapó with extensions: - fTools 0.5.10 - GDAL Tools: 1.2.16 - Raster-based terrain analysis - Georeferencer GDAL 3.1.9 - RasterCalc 0.2.1 (in development) - GRASS Toolbox

3. TASKS & LAB EXERCISES

The tasks listed here were utilized to compare the problem solving capabilities of open-source GIS-software. The exercises were part of the lab “Introduction to Geoinformation” (1st year) for which students used gvSIG. As part of an advanced course (2nd year), QGIS is used for a different set of exercises. For this paper, however, the first year's exercises were carried out in both gvSIG and QGIS to allow for a better comparison. The lab exercises consisted of four main topics:

- Geo-referencing of aerial images;
- Data capture (Digitizing);
- Spatial analysis (Raster and Vector analysis);
- Navigation and Routing (Network analysis).

Task 1: Georeferencing of aerial images

In this task, students were asked to geo-reference an aerial image based on a number of control points they had to determine from cadastral data (parcels and houses). The utilized dataset consisted of an orthoimage of the city of Eisenstadt (Austria) as well as buildings and parcels from the digital cadaster of Austria, provided by the national mapping agency. To verify the results, students had to provide a screenshot of the georeferenced image, the respective world file, and a table of control points.

Task 2: Data capture

This exercise dealt with the collection of spatial data, making use of on-screen digitizing based on an aerial image of the village Rastenberg in Austria (see Figure 1). Students were asked to generate a spatial data model that contained several layers with polygonal objects and appropriate attributes (e.g., several road classes). The objective of this exercise was to create a “complete” dataset without covering the whole test area. Also, students had to apply various generalization methods where necessary.



Figure 1: Aerial image of the village Rastenberg, Austria.

Task 3: Spatial analysis

Spatial analysis is one of the core functionalities any GIS offers. We defined three sub-tasks to cope with the wide variety of analyses possible. Exercises included analysis of both vector and raster data.

Sub-task 3a - Spatial analysis of raster data:

In this task raster data (wind speed and vegetation classes) for the continent of Africa were analysed. The main objective of this exercise was to extract geographic areas prone to erosion. Students had to retrieve “high-wind” areas (wind speeds > 3 m/s) and combine them with designated desert areas to identify the potential regions of interest. This task required students to utilize a raster buffer tool, reclassification (extraction of vegetation classes), and raster calculus functionalities (combining the layers).

Sub-task 3b: Spatial analysis of vector data – Geoprocessing:

Parts of this exercise were based on Richard E. Plant's qGIS labs, available on his Web site (Plant, 2011). The main task consisted of finding a suitable location for a new football stadium in the state of California (USA). The data-sets contained county borders, interstate highways, and urban areas. The location of the new stadium should satisfy the following criteria: (a) within Solano or Yolo County, (b) within a 3 kilometer radius of an urban area – but not inside an urban area, and (c) within 1 km of an interstate highway. The geoprocessing task consisted of the extraction of data, combination of several layers, as well as attribute queries.

Sub-task 3c: Spatial analysis of vector data – Classification:

Geo-demographical and election data of Austria were both mapped and visually analyzed. First, a map displaying the population below the age of 19 for each district was created, using equal intervals and natural breaks classification. Second, a map of the Austrian national election 2008 had to be created showing the winning party for each district. This task included data classification and table joins. In another exercise, students had to work with Viennese census data and identify areas based on various demographical factors (e.g., family-size, unemployment rate), as well as create population density maps for each district.

Task 4: Navigation and Routing

This exercise introduced students to spatial analyses based on navigation and routing methods, i.e., network analysis. It consisted of finding a solution to the traveling salesman problem and the creation of service areas.

Sub-task 4a: Traveling Salesman Problem:

Based on a data-set of the Austrian road network and various Austrian cities, students had to create an optimal shortest route of a virtual person visiting all provincial capital cities in Austria. The start and endpoints were fixed (Vienna) and cities could only be visited once. These tasks resembled a typical Traveling Salesman (TSP) problem.

Sub-task 4b: Service Area:

The calculation of Service Areas is an important tool for both business geographers and spatial planners alike. In this task, students should calculate and visualize potential Service Areas for public and private hospitals in Vienna.

4. RESULTS

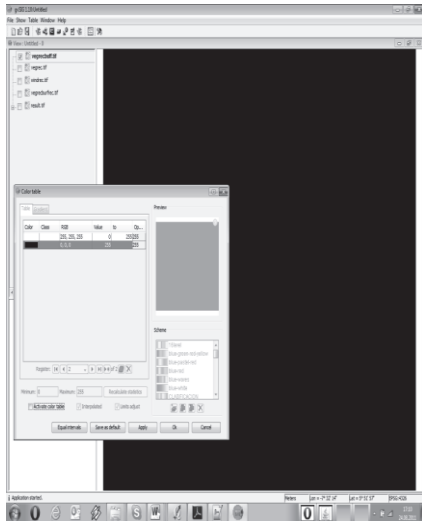
In this section, the results of the comparison between gvSIG and QGIS are discussed. Based on the exercises, carried out with both software products, we evaluate and discuss the performance accordingly. The following subsections elaborate on the results of the tasks achieved as well as possible errors, and bugs.

4.1 Evaluation of gvSIG

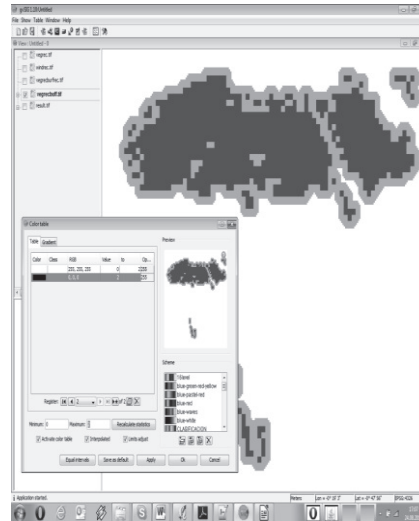
Task 1 can be achieved by utilizing gvSIG's built-in geo-referencing tool. Control points can be selected from the image to be geo-referenced as well as the reference data set, both displayed side by side. This makes the usual layer transparency control and repeatedly switching between layers unnecessary. The interface is consistent, straight-forward, and easy to use. With some assistance from the lecturers, most students had little problems in carrying out this task.

The "data capture" exercise was also completed using the built-in functions of gvSIG. It was necessary, however, to create a new shapefile with corresponding attributes beforehand. In addition, the on-screen digitizing of polygons is possible with the help of the snapping function, similar to the ones known from various proprietary software products. Unfortunately, students reported a number of unexpected errors and bugs – especially null pointer exceptions, making the digitizing process difficult and frustrating. Often, the instructors could not find the cause of the problem and tasks that worked on one student's computer failed on another, despite the fact that the exact same steps were carried out.

The raster analysis task (task 3a), can be completed using gvSIG's Sextante Toolbox. The toolbox is fully integrated into gvSIG's user interface (UI) and offers 286 spatial data analysis functions to date. For this task the following functions were used: "Reclassify", "Raster Buffer", and "Raster Calculator". Basically, the functions worked as expected, but some drawbacks concerning the overall handling were encountered. Below is an example of using the raster buffer that frustrated students in the lab. The creation of a buffer in a raster layer results in a black map window, making the user believe the buffer was specified incorrectly. The generated color table does not consider the values of the layer and resulting in a black map window. Consequently, the correct results were not visible due to an automatically assigned color table (see Figure 2).



(a)

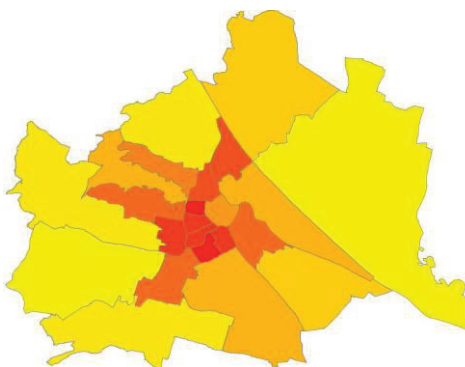


(b)

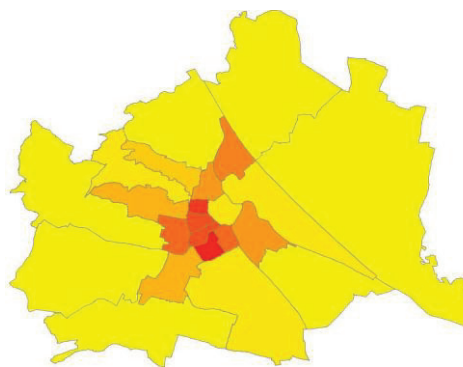
Figure 2: Misleading results of Sextante Toolbox – buffering example in (a). Adjusting the color table in (b) reveals the results that were correctly calculated.

Task 3b can be achieved by either using the build-in Geoprocessing Toolbox or the Sextante library. Students encountered that the buffer algorithm of Sextante does not offer the possibility to dissolve the buffer zones automatically after buffers were created. Therefore, this process must be carried out manually. A general problem of vector analysis with gvSIG concerns the homogeneous coloring (light blue) of resulting layers. This is not intuitive and facilitates problems when visually inspecting the result.

The classification assignment (task 3c) can also be achieved using the built in functions of gvSIG. First, election data (table) had to be merged with spatial data in order to generate a spatial dataset, using the join function. Second, results were analyzed using different classification methods, e.g., natural breaks and equal interval. Sample results displaying the population density for the city of Vienna are visualized in Figure 3. We encountered no problems with this exercises.



(a)



(b)

Figure 3: Maps showing the population density of Vienna: Natural breaks (a) and Equal interval (b).

In order to perform network analysis in gvSIG, the extension “Network Analyst Redes” needs to be installed separately. This allows the generation of a network topology, based on polyline layers. In addition, the calculation of shortest paths between nodes is possible. Hence, the provincial capitals of Austria were defined as stops in the Network Analyst. Furthermore, the calculation of a round-trip, with start and destination in Vienna, was required. The result of such a simple TSP calculation, as carried out by our students is shown in Figure 4.



Figure 4: One solution to the Traveling Salesman Problem.

Task 4b required students to calculate Service Areas for both public and private hospitals in the city of Vienna. This can be accomplished using the functionalities of the Network Analyst in gvSIG. Using the extension allows the modeling of simple Service Areas based on travel costs of the road network, i.e., time or distance. Calculated Service areas were defined as regions within 3000 meters (network distance) from any hospital.

In general, gvSIG's network extension offers a wide variety of features on networks analysis. We experienced, however, some minor inconsistencies and unexpected crashes. Students seemed to appreciate the functionalities but were a little frustrated with the overall handling.

4.2 Evaluation of QGIS

The geo-referencing task can be completed with the extension “Georeferencer GDAL 3.1.9”. This add-on enables the geo-referencing of an aerial image by selecting control points interactively or manually (direct coordinate input). Additionally, the reference points can also be selected from the main map window. Nevertheless, the aerial image to be referenced cannot be viewed side by side with the main map, so only one of them is fully visible. In this regard, gvSIG provides a much better user interface and handling making the process of geo-referencing easier.

“Data capture” requires no additional extensions, thus, on-screen digitizing of the aerial image of Rastendorf could be completed with the built-in functionalities of QGIS. Especially useful was the snapping tool, facilitating the generation of “clean” geometries.

The raster analysis of task 3a can be achieved using the GRASS Toolbox available as extension for QGIS. Unfortunately, built-in functionalities of QGIS do not provide appropriate tools to analyze raster data accordingly. The GRASS toolbox offers a variety of tools integrated into QGIS. In this task

the following tools were used: “r.mapcalculator”, “r.reclass”, and “r.buffer”. Fortunately, we did not experience any errors or bugs while working with these extensions.

Geoprocessing (task 3b) can only be completed using the extension fTools, allowing the user to carry out basic geoprocessing, such as buffering, intersection, and difference. In general, the extension worked well and is very intuitive to use. One encountered drawback concerned the deletion of duplicate field names. fTools cannot deal with duplicate field names – which may occur with two intersecting features. However, the extension “Table Manager” can be used to fix duplicate field names before running the spatial analysis.

The classification assignment, (task 3c) can be completed using standard functionalities QGIS offers in combination with fTools. In order to join datasets, i.e., combining plain text (table) with spatial data the extension fTools is needed. For analysis and visualization the standard functions of QGIS are adequate. In addition, QGIS offers a great range of symbology possibilities (in comparison to gvSIG).

In theory, solving an advanced spatial analysis task (e.g., TSP) should be possible in QGIS utilizing the GRASS toolbox. It offers functions to create a topologically clean road network from Shape-data using the modules “v.clean” and “v.build”. They would serve as basis for further calculations of the TSP. In the experiments conducted for this paper, however, the TSP algorithm in GRASS (“v.net.salesman”) failed. Results could not be computed because the module crashed unexpectedly. Unfortunately, the error messages do not help in finding the reason for the module crash (see Figure 5). Generally, using the GRASS extension and GRASS layers results in a slower performance of QGIS compared to using layers directly loaded into QGIS. Especially, when working with large data-sets (e.g. road network of Austria), the user is able to notice the decreased performance.

Task 4b required the calculation of service areas for all hospitals in Vienna. In order to achieve this task we used the GRASS toolbox (module “v.net.alloc”). It offers the possibility to calculate subnets that are closest to a given point (in our case: hospitals). Unfortunately, it was not possible to define a threshold value for allocating subnets, e.g., travel time. Thus, the results depicted in Figure 6 differ from the definition of “real” service areas.

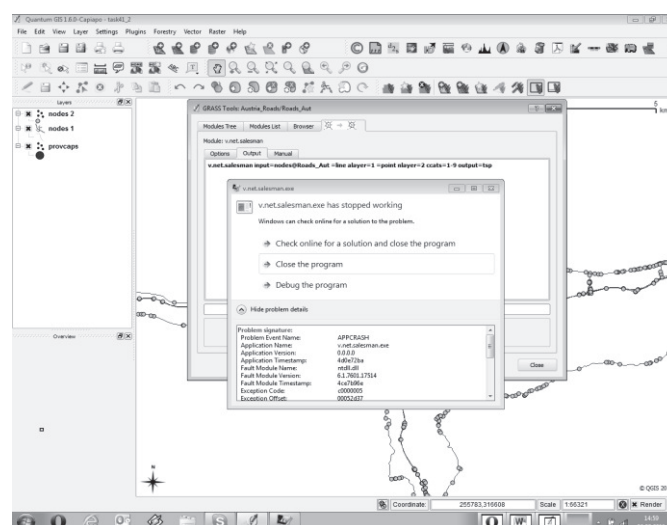


Figure 5: Error message indicating the crash of the TSP calculation in QGIS.

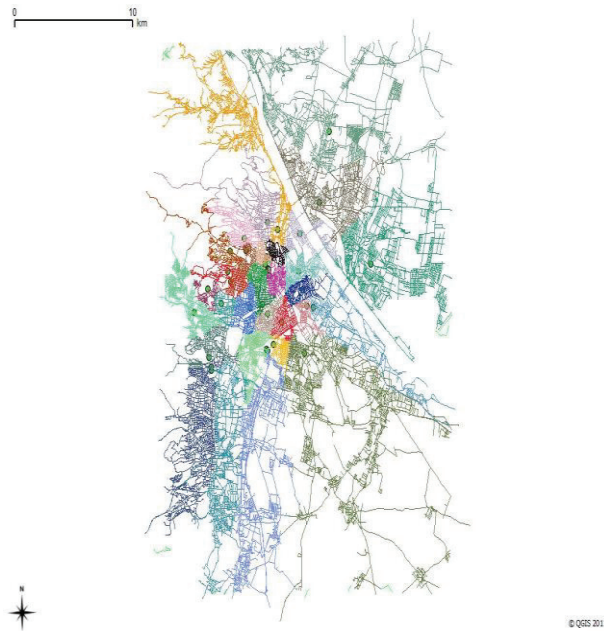


Figure 6: Service Areas for Vienna's hospitals (green dots), calculated with the QGIS "v.alloc module".

5. CONCLUSION

Unfortunately, gvSIG proved to be quite unstable. Also, user interface and input forms tend to be largely unclear and inconsistent. Even though the Sextante library comes with over 250 useful functions, most of them are poorly documented. Therefore, an exact knowledge of each algorithm is required, making it difficult for students to "just play around". Furthermore, we experienced several major problems with Sextante's raster functionality. Sometimes the results even appeared to be different depending on the PC's operating system and/or system's specifications.

We found most students forced to concentrate on workarounds to a great extent, in order to achieve reasonable results. Thus, they could not spend much time on trying out and working with algorithms and data. This, however, would have been a crucial process to intensify the knowledge gained in the lectures. While we experienced these problems mostly when working with raster data-sets, vector analysis worked reasonably well. In general, we found that one should avoid working with temporary layers. Every time some geo-processing task is completed gvSIG asks the user to either store the result as a temporary (RAM) or permanent (Disk) layer. Working with temporary layers often lead to Null-Pointer exceptions and similar unpleasant error messages and crashes.

Initially, QGIS comes with a very small set of functionalities, thus, at a first glimpse it simply appears to be a data-viewer (although with rich functionality in comparison to other viewers like uDig). Unfortunately, any form of spatial analysis is impossible before installing some of the extensions mentioned in this article. There are, however, plenty of plug-ins available, helping to overcome this issue. The GRASS toolbox provides a considerable number of algorithms, but the user interface of the toolbox remains an issue of improvement. Also, the slow computation speed of the GRASS toolbox can be annoying when dealing with big datasets. As with gvSIG, we witnessed unexpected crashes, leading to situations where we were unable to open GRASS mapsets. Apparently, this was caused by a ".gislock", a file created by GRASS, not automatically deleted during the program's shut down. Also, QGIS's documentation seems to be more extensive and reliable compared to gvSIG's.

Students spent a considerable amount of time on finding certain functions and/or figuring out how to use them. This was mainly due to the lack of a comprehensive documentation or support. This is particularly true for gvSIG since some parts of the manual are written in poor English, others are only available in Spanish. Such resources are crucial for any GIS-user, particularly for beginner's. This makes it especially difficult for students who want to understand the theory behind the functions and make sense of the data. From an instructor's perspective it proved quite difficult to work with both systems. Preparation of lab exercises was tedious and little satisfying. Some subjects had to be omitted due to unexpected program behaviour, others because the software would not allow to carry out specific functions.

In conclusion, we believe that the current state of both systems is suited for advanced GIS-courses only. That is, courses where the task is not to understand the basics, but to enhance the applied skills of the students. In our opinion, courses on an introductory level are better off with established, stable and well documented GIS-software products. This ensures a motivating teaching environment, such that students can focus on experimenting with data and functionalities, rather than having to struggle with workarounds and ambiguous interfaces. We believe, however, that the two compared GIS-software products can, under certain constraints, be utilized for teaching. Currently, however, they are not mature enough to completely replace proprietary systems.

BIBLIOGRAPHY

- Acevedo, V., gvSIG-Desktop 1.11. User Manual (PDF Version). Available online: http://forge.osor.eu/docman/view.php/89/706/gvSIG-1_11-man-v1-en.pdf (Accessed September 2011).
- Athan, T., Blazek, R., Contreras, G., et al. QGIS Manual (PDF Version). Available online: http://download.osgeo.org/qgis/doc/manual/qgis-1.7.0_user_guide_en.pdf (Accessed September 2011).
- Longley, P.A., Goodchild, M.F., Maguire, D.J., Rhind, D.W., Geographic Information Systems & Science. Wiley 2011.
- Moore, J.W., Education versus Training. University of Wisconsin at Madison, Madison, WI. Journal of Chemical Education. 1998.
- Plant, R.E., QGIS Labs. Department of Plant Sciences. University of California at Davis. Available online: <http://www.plantsciences.ucdavis.edu/plant/qgislabs.htm> (Accessed September 2011).