Teaching Raster GIS Operations with Spreadsheets

Martin Raubal, Bernhard Gaupmann, Werner Kuhn

Department of Geoinformation
Technical University Vienna
Gusshausstr. 27-29
A-1040 Vienna, Austria
raubal@geoinfo.tuwien.ac.at
Phone: (+43 1) 58801-3790
Fax: (+43 1) 504-3535

ABSTRACT

Due to the rapid development and increasing use of Geographic Information Systems (GIS) more and more Universities have introduced basic GIS courses to teach students the principles of such systems. In addition to lectures, exercises with commercial GIS software are offered to show basic operations. Although students learn to execute such operations, the systems tend to hide their internal structure and logic. We propose to use a spreadsheet program as a teaching tool for basic raster operations such as filter and overlay. With this tool students are able to perform and visualize these operations as well as to see how the data are processed by the algorithms.

KEYWORDS

Spreadsheets, Raster model, Macro language, Visualization, Education.

1. INTRODUCTION AND MOTIVATION

In 1990 the National Center for Geographic Information and Analysis (NCGIA) released the revised core curriculum for teaching the fundamentals of GIS. In [Goodchild, & Kemp, 1992] the contents of this program are described and several problems in including laboratory exercises are brought forward. It is argued that processing data by hand will let students 'look behind the scenes' of operations. Also students need to get experience with commercial GIS software packages. Both of these arguments are very important for GIS education, but it is too time consuming to teach many operations by hand. Therefore we propose a commonly used spreadsheet program [Microsoft Excel, 1985-1994] to serve as a teaching tool. Such a tool is not only able to visualize the results of raster operations. It also makes the workings of the operations explicit which helps students to understand the logic of data manipulation. This is, in our experience, equal or superior to processing data by hand.

As part of the 'Geometric Data Management' course 1995, surveying engineering students at the Department of Geoinformation, TU Vienna, were confronted with an analog raster map of lake 'Retsar' and had to do the following exercises:

1. Compute and show the lake shore (by program).

2. Smooth the lake's shape by applying a convolution filter (by program).
   Experiment with various sizes and weights in the kernel.

There are different ways to find a solution for these problems: One can use any common programming languages (like PASCAL). This means writing a lot of code to get a user interface for the input, visualization and manipulation of data. Looking for an easier way to get the exercise done, one of the students came up with the idea of using a spreadsheet program. The results looked so promising that we tried to find the answer to an additional question: Does it make sense to use a spreadsheet program as a teaching tool for basic raster operations in introductory GIS courses?

The remainder of the paper is organized as follows: Section 2 discusses the role of spreadsheets in a raster model, describes a general approach based on spreadsheets and presents some applications of spreadsheet programs for the implementation of GIS operations. In section 3, we explain the macro code of some selected operations and visualize the results. Section 4 compares the spreadsheet to commercial GIS software. The last section assesses our achievement and sets forth conclusions and perspectives for future work.

2. APPROACH BASED ON SPREADSHEETS

Spreadsheets offer a practical way to demonstrate raster operations because the raster structure is captured in the form of rows and columns. There is a one-to-one mapping of raster space onto cells in the spreadsheet. Cells in a spreadsheet are arranged in a rectangular grid and contain either values (data) or mathematical formulae with references to the values of other cells. It is also possible to assign different colors and patterns to each cell. This is a useful way to represent attributes. It takes little effort to change cell values, as pointing to a cell and clicking allows immediate access to its value. One can also assign a certain value to a whole region by defining names for specific areas.

Spreadsheets have been used for doing simple spatial analyses, addressing various geographical problems such as natural resource management. Ali, Ross, and Beveridge [1991] have shown such analyses in a case study on carp in Pakistan where they have used a spreadsheet (Viewsheet spreadsheet package) for the implementation of GIS in aquaculture. Bossard and Zhang [1993] effectively combined a PC database with a spreadsheet and desktop mapping program and used census data to analyze the retail trade of shopping centers. They argue that this allows professionals with modest budgets and computer skills to undertake spatial analysis.
An interesting idea is using spreadsheets for image visualization. As described in [Levoy, 1994] spreadsheets for images can be used as a visual database browser, as a graphical user interface builder, as a smart clipboard for the desktop, and as a presentation tool. Cells contain not only formulae but also diverse graphical objects like images, volumes, or movies, which can be manipulated through operators.

Spreadsheets have also been used as teaching tools in various disciplines, i.e., economics, statistics and also geography. In [Campbell, 1991] a new approach of introducing students to the principles and practices surrounding GIS was taken. Although spreadsheets are much less sophisticated than commercial GIS software packages, they offer many pedagogic advantages. Campbell used the spreadsheet VPPlanner Plus for mapping classified information, but he focused on spatial analysis, not on the functions and operations themselves. In addition, the visualization quality of results in the form of a scattergram was poor. Showing classifications and maps is a nice way to explain attribute data and map features, but it is not sufficient to get an idea of what basic GIS operations do and how they do it.

Our approach of using spreadsheets as teaching tools is new in that we concentrate on the algorithms of operations. We make explicit which functions GIS are actually evaluating when performing a particular operation. Of course, all of the advantages of spreadsheets which have been pointed out in the literature are still valid when using them in the way we do. Low cost of spreadsheets and ease of use were main arguments in both [Campbell, 1991] and [Bossard, & Zhang, 1993]. The latter also referred to ease of learning and operation. It is a fact that spreadsheet programs are familiar to many students who use them to perform various tasks in other subjects.

3. OPERATIONS

In this section we look for an answer to the question whether it makes sense to use a spreadsheet program as a teaching tool for basic raster operations in introductory GIS courses. We tried to execute some raster operations using MS Excel for Windows which is the commonly used spreadsheet program at the TU Vienna. Its user interface allows for easy manipulation and management of spatially arranged data. All operations are implemented by MS Excel 4.0 macros. This is a useful programming environment, because the macro language is interpreted rather than compiled, providing interactivity during the running of macros. Also one can avoid writing explicit code using a macro recorder keeping track of all steps performed on a sheet.

The basic raster operations are similar to those which students had to execute with commercial GIS software in the preceding 'Introduction to GIS' course. The following operations are discussed in more detail in this paper:

- Overlay
- Filter
- Border

3.1. Overlay

For our demonstration we defined a region which is an approximation of Austria. To determine suitable climatic conditions for firtrees one needs maps, showing the average rainfall per year and Austria's elevation model. Suitable climatic conditions could be:

- rainfall between 800 and 1500 mm per year (Figure 1)
- elevation lower than 1500 m above sea level (Figure 2)

These figures are used for demonstration only and represent a crude approximation of reality.

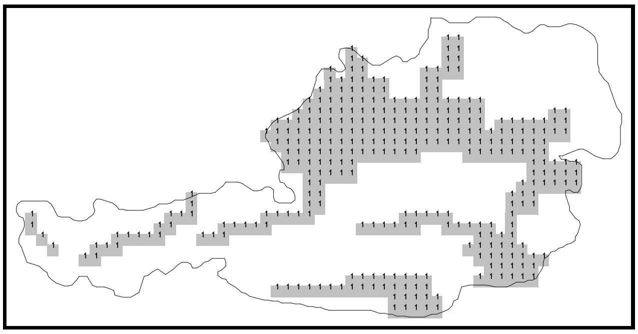

Figure 1: Rainfall between 800 and 1500m.

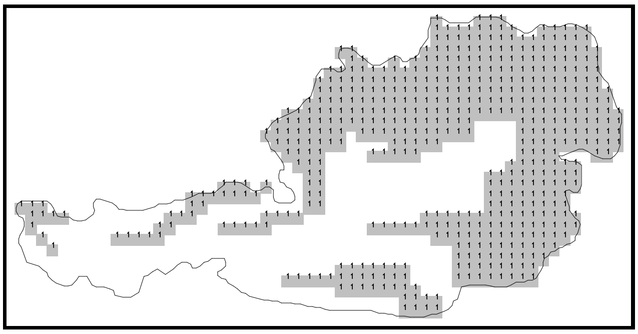

Figure 2: Elevation lower than 1500m.

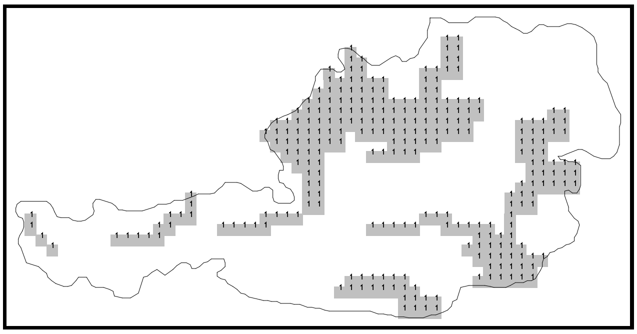

Figure 3: Result of the AND-operation.

The values in Figure 3 show the result of an overlay operation using the logical operation AND on each cell. This operation is very helpful for certain analyses. Truth values are taken from the standard truth-table for Conjunction [Piff, 1991] (Table 1).

P	Q	P ∧ Q
T	T	T
T	F	F
F	T	F
F	F	F

Table 1: Truth-table for Conjunction.

```
Overlay_And
FOR.CELL("Overlay";Output_Map;)                              {Overlay is the reference to
SELECT(OFFSET(Overlay;0;0))                                   the Output_Map which is a defined
IF(AND(OFFSET(Overlay;-33;-61)=1;OFFSET(Overlay;-33;0)=1);   name for a certain array of cells}
  AND(PATTERNS(1;;15;TRUE);FORMULA("1"));                    {color gray; value 1}
                           FORMULA(""))                      {no color; no value}
NEXT()
RETURN()
```

'Overlay_And' runs over each cell of the Output_Map and sets a new value which is a combination of the corresponding cell values of two different original maps. OFFSET sets the cursor to the cells referenced in the parentheses. In our example the original maps are located 33 rows above the output map and 0 or 61 columns displaced (OFFSET(Overlay;-33;-61)). Viewing this macro is a simple but effective way to show the logic behind the 'Overlay_And'- operation. Moreover the macro can be easily changed e.g. into an OR-operation.

3.2. Filter

The filter operation computes a new value for every location as a function of its neighborhood [Tomlin, 1990]. We use the 8-neighborhood, which consists of the four horizontal and vertical and the four diagonal neighbors [Rosenfeld, 1979]. To show how a smoothing filter works, Figure 4 represents Austria with the National Park 'Hohe Tauern' region. The aim of applying the filter operation is to smooth both the boundary of Austria and of the National Park. For this filter operation we take a kernel (Table 2) which severely smoothes the spatial variation on a layer [Goodchild, & Kemp, 1990].

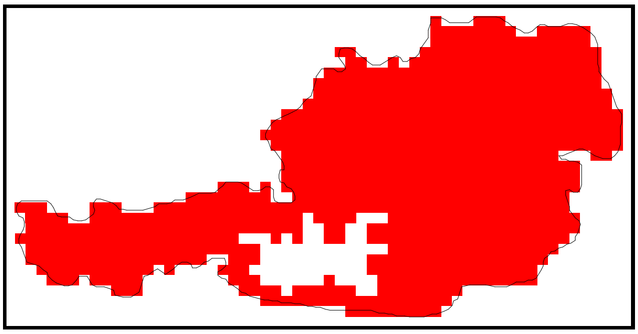

Figure 4: Original Map with National Park 'Hohe Tauern' region.

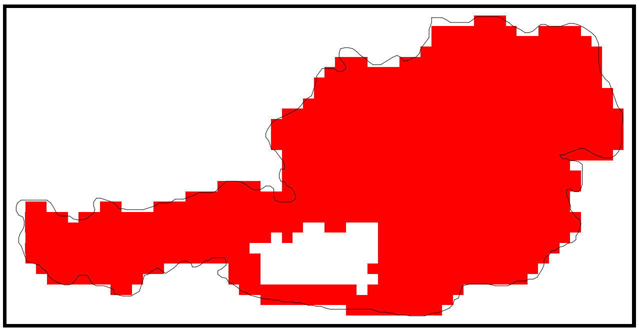

Figure 5: Smoothed Map.

0.11	0.11	0.11
0.11	0.11	0.11
0.11	0.11	0.11

Table 2: Smoothing Filter.

```
Filter
FOR.CELL ("Filter";Output_Map;)
SELECT (OFFSET(Filter;0;0))
IF ((        (OFFSET (Filter;-34;-1)* $K$67 +          {$K$67...upper left
              OFFSET (Filter;-34;0)* $L$67 +             position of the filter}
              OFFSET (Filter;-34;1)* $M$67 +
              OFFSET (Filter;-33;-1)* $K$68 +
              OFFSET (Filter;-33;0)* $L$68 +
              OFFSET (Filter;-33;1)* $M$68 +
              OFFSET (Filter;-32;-1)* $K$69 +
              OFFSET (Filter;-32;0)* $L$69 +
              OFFSET(Filter;-32;1)* $M$69) / SUM ($K$67:$M$69))>=0,5;
  AND (PATTERNS (1;;3;TRUE); FORMULA("1"));            {color gray, value 1}
  FORMULA(""))                                          {no color, no value}
NEXT()
RETURN()
```

The kernel is represented by an area below the input and output maps on the spreadsheet. Using the OFFSET-function the kernel runs over the original map which is located in the EXCEL sheet 33 rows above the output map. The kernel covers 9 cells of the original map. All of them are multiplied by the corresponding kernel values, the results added, and then divided by the sum of kernel values. The final rounded results (0 or 1) are written to the output map. A spreadsheet offers the possibility of using different filters very flexibly. Students only have to change the kernel's values. Afterwards they can run the macro again and see the effect of the changed filter.

3.3. Border

In our example of the border-operation, we use a combination of two definitions:
1. "The *border* of the image is the outer boundary of the square corresponding to the array." [Samet, 1989, p.2] Implementing this definition requires the possibility to show vector-graphics in the output map.
2. In addition to the vector boundary we also visualize a second form of border in the same output map, namely the closed chain of neighbor cells outside the area.

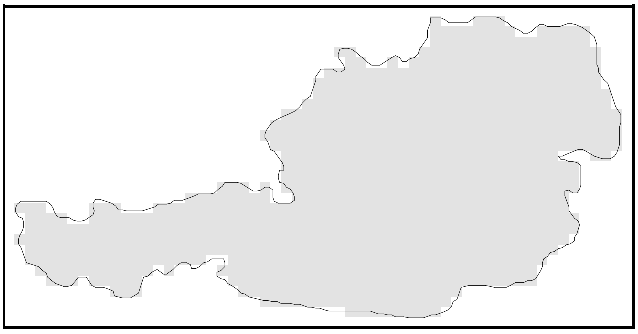

Figure 6: Original Map. Gray cells represent Austria.

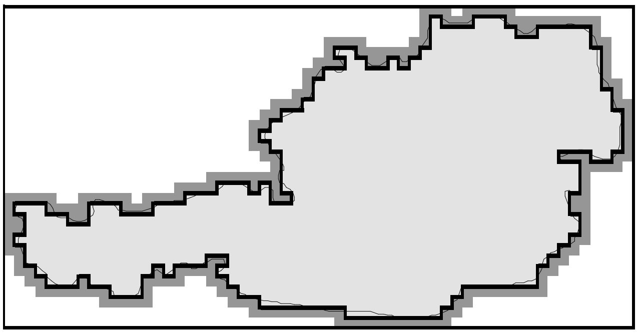

Figure 7: Output Map. Gray: Austria, dark gray: border, white: other cells. Black: Vector boundary

Border
```
FOR.CELL ("G", Original_Map, )
  IF (COUNTA(G)=1, COLOR_GRAY(), COLOR_DARK_GRAY())

  IF (AND (COUNTA(G)=1, OFFSET(G,0,-1)=0, OFFSET(G,-1,0)<>0, OFFSET(G,0,1)<>0, OFFSET(G,1,0)<>0),
        BORDER (0,2,0,0,0,,,0))
  IF (AND (COUNTA(G)=1, OFFSET(G,0,-1)=0, OFFSET(G,-1,0)<>0, OFFSET(G,0,1)<>0, OFFSET(G,1,0)=0),
        BORDER(0,2,0,0,2,,,0))
  IF (AND (COUNTA(G)=1, OFFSET(G,0,-1)=0, OFFSET(G,-1,0)<>0, OFFSET(G,0,1)=0, OFFSET(G,1,0)<>0),
        BORDER(0,2,2,0,0,,,0))
                .
                .
                .
  IF (AND (COUNTA(G)=1, OFFSET(G,0,-1)<>0, OFFSET(G,-1,0)=0, OFFSET(G,0,1)=0, OFFSET(G,1,0)=0),
        BORDER(0,0,2,2,2,,,0))
NEXT ()
RETURN ()
```

Color_ Gray
```
SELECT (OFFSET(G,33,0))
PATTERNS (1,,3,TRUE)                                    {color gray}
RETURN ()
```

Color_ Dark_Gray
```
IF (OR (OFFSET(G,-1,-1)<>0, OFFSET(G,-1,0)<>0, OFFSET(G,-1,1)<>0,
        OFFSET(G,0,-1)<>0, OFFSET(G,0,1)<>0, OFFSET(G,1,-1)<>0,
        OFFSET(G,1,0)<>0, OFFSET(G,1,1)<>0),
  AND (SELECT(OFFSET(G,33,0)), PATTERNS(1,,6,TRUE)),    {color dark gray}
  AND (SELECT(OFFSET(G,33,0)),,PATTERNS(1,,4,TRUE)))    {color white}
RETURN()
```

In our example the Original_Map contains area 'Austria' whose cells are set to 1. All other cells have the value 0. The macro 'Border' checks for each cell of the Original_Map if its value equals 1. If so, then the corresponding output cells become gray, indicating that the cells belong to area 'Austria'. Otherwise the macro 'Color_Dark_Gray' checks if every cell has at least one neighbor containing value 1. If this is the case, the output cells are assigned the color dark gray and are part of the closed boundary chain. The white cells in the output map show the area which does not belong to Austria. Their original value is 0 and they do not have any neighbors containing value 1. In order to show the vector boundary, the macro 'Border' evaluates the following condition: A cell of 'Original_Map' has to show the value 1 and one of its 4-neighbors has to be 0. If these conditions are met, the macro draws a line (syntax BORDER(0,2,0,0,0,,,0)) between the corresponding output cells.

4. COMPARISON WITH OTHER TEACHING SOFTWARE

We looked at two commercial GIS software packages that are commonly used for teaching, to see how these three basic raster operations are performed and explained there:

- GISTutor is a guide to GIS, covering concepts and technical issues. It contains a variety of information sources. Raper and Green [1992] give a good overview of the GISTutor project. Unlike some other tutorials GISTutor explains the concepts behind both the raster- and vector data model. It is designed to be used by clicking on screen icons. Although it presents a good overview of GIS functions, users are not able to change the input values of operations, e.g. there are good demonstrations of a low- and high-pass filter, but there is no possibility of creating and testing a kernel with arbitrary values. Some GIS functions lack complete explanations, e.g. for overlay operations the results are well explained, but there is no description of the path from input to output data. Our approach of demonstrating basic raster operations with a spreadsheet forms an ideal complement to compensate for these shortcomings of GISTutor.

- IDRISI is a raster GIS which offers different modules to perform geographical analyses. The 'collective geographic analysis project' is described in [Eastman, & Warren, 1987] and has been used for many commercial applications as well as for teaching students the principles of GIS. Performing simple raster operations with IDRISI is much more complex than doing it with a spreadsheet. Considerable previous knowledge from the manual is necessary to go through the many steps which lead from input to output data. With a spreadsheet these steps can be combined in a macro, so operations become easier to observe. In addition to this, with IDRISI one cannot see the result of an operation immediately, but has to change into a display module. With a spreadsheet it is possible to view input and output data as well as other elements (like a kernel) on one screen. This is a big advantage if one wants to compare the original map with the resulting map of an operation.

5. CONCLUSIONS AND FUTURE WORK

In this paper we presented reasons and examples for using a common spreadsheet program as a teaching tool for basic raster operations in introductory GIS courses. The most important reasons are:

- use in class
- simple user interface
- familiarity to students
- low cost
- flexibility of changing cell values
- changeability of parameters
- easy programming environment (macro-language)
- possibility to 'look behind the scenes of operations' by viewing the code

Despite these advantages there are some obvious drawbacks. Calculation is relatively slow, which becomes an even bigger problem when using a lot of data. The limited resolution is not very important for demonstrating operations, but has to be considered when using the spreadsheet for real world applications. Students need to have some understanding of the macro language to be able to look behind the operation's code. But only a few macro functions are necessary for demonstrating basic raster operations. The MS Excel 4.0 macro language is logically structured (e.g. syntax IF(logical_test, value_if_true, value_if_false)) and similar to other programming languages, e.g. BASIC (FOR...NEXT command). Therefore lecturers are able to teach basic commands within a few minutes. Higher resolution of input areas (in our examples we use about 1800 cells) leads to a linear increase in calculation time. However during frontal instructions using an overhead it gives teachers time to explain the mode of operation, while macros are executed cell by cell in front of the class.

In order to use this teaching tool in introductory GIS courses, it will be necessary to implement additional raster operations. We have already implemented the area-, buffer- and reclassify operation, and also a visualization of a quadtree. An interesting extension would be the implementation of such a tool for the visualization of operations in GIS courses offered via the World Wide Web. Another possibility for future work is to simplify macro language programming to help in the development and understanding of raster operations.

Related future work could concentrate on the use of spreadsheet programs for building cellular automata models. Engelen, White, & Uljee [1995] present such a model as the core of a Decision Support System to help urban planners and policy makers in simulating and analyzing urban layouts, land uses and growth patterns.

ACKNOLEDGEMENTS

Discussions with Hermann Müllner were very helpful at the beginning of the project. Andrew Frank provided support and comments.

REFERENCES

Ali, C. Q., Ross, L. G., & Beveridge, M. C. M. 1991. Microcomputer spreadsheets for the implementation of geographic information systems in aquaculture: a case study on carp in Pakistan. Aquaculture 92 (2-3):199-205.

Bossard, E. G., & Zhang. 1993. Loosely coupled PC programs as a framework for spatial analysis. URISA Proceedings:13-24.

Campbell, N. 1991. The political economy of GIS: The spreadsheet as a low cost GIS. Cartographica 28 (3: GIS Education and Training):65-74.

Eastman, J. R., & Warren, S. 1987. IDRISI: a collective geographic analysis system project. (Chrisman, N. R., Eds.). Auto-Carto 8, Baltimore, MA. ASPRS & ACSM. Eighth International Symposium on Computer-Assisted Cartography, pp: 421-430.

Engelen, G., White, R., & Uljee, I. 1995. Integrating Constrained Cellular Automata Models, GIS and Decision tools to Support Urban Planning and Policy Making. *9th European Colloquium on Theoretical and Quantitative Geography*, Spa, Belgium. To be published.

Goodchild, M. F., & Kemp, K. K., eds. 1990. Introduction to GIS - NCGIA Core Curriculum: National Center for Geographic Information and Analysis, University of California, Santa Barbara.

Goodchild, M. F., & Kemp, K. K. 1992. NCGIA education activities: the core curriculum and beyond. International Journal of Geographical Information Systems 6 (4):309-320.

Levoy, M. 1994. Spreadsheets for Images. (Mair, S. G., Eds.). *SIGGRAPH 94*, Orlando, Florida. ACM SIGGRAPH. Annual Conference Series, pp: 139-146.

Microsoft Excel for Windows NT 5.0 . Microsoft Corporation.

Piff, M. 1991. Discrete Mathematics - An introduction for software engineers: Cambridge University Press.

Raper, J., & Green, N. 1992. Teaching the principles of GIS: lessons from the GISTutor project. International Journal of Geographical Information Systems 6 (4):279-290.

Rosenfeld, A. 1979. Digital Topology. *American Mathematical Monthly* 86:621-630.

Samet, H. 1989. Applications of Spatial Data Structures: Computer Graphics, Image Processing and GIS. Reading, MA: Addison-Wesley.

Tomlin, C. D. 1990. Geographic Information Systems and Cartographic Modeling. Englewood Cliffs, N.J.: Prentice Hall.