

## ***r.finder***

by Augusto Palombini

*r.finder* is a GRASS-GIS script created for the intelligent geographic survey, in order to check the analogies of the places where some known items are located and show the most similar areas in the region. *r.finder* outs a map containing all cells matching the same combination of thematic maps parameters of the non-null/non-zero cells in the input map. May out as well a statistical report and graphs of the cell values.

---

## ***r.finder*: a short tutorial**

### **1- Purpose**

*r.finder* is a GRASS-GIS script created for the intelligent geographic survey. The approach is based on the hypothesis that the elements to be found on a territory, lie on portions of land which are somehow similar, according to specific thematic features expressed by different maps.

Such thematic maps are compared, on the basis of the already known elements, and the analogies of their places are be used to check the region for areas with the same features, in order to focus the survey on the most appropriate zones.

*r.finder* reads an 'input' map consisting in the location of the known elements (sites, occurrences), and a series of thematic maps showing particular features of the region analyzed. *r.finder* distinguishes two kinds of maps: qualitative and quantitative. For qualitative maps the analysis is performed checking for any single category value of the thematic maps in which the occurrences fall (if the the occurrences fall on cells with value of 3, 24, 394 in the qualitative map 'soil type', only such values are considered). For quantitative maps, the analysis is performed checking for the whole range between the lowest and the highest value (if the the occurrences fall on cells with value of 3, 24, 394 in the quantitative map 'terrain height', the whole range between 3 and 394 are considered, or – if the flag -d is given – the standard deviation).

The final result is a map resulting for the sum of the single analysis, showing the higher or lesser concentration of analogies with the cells on which elements fall.

*r.finder* may also outs statistical reports and graphs, to allow the evaluation of any thematic map taken into account for the final result. Moreover *r.finder* allows as well to store the analysis results in a rules file, to be used on different areas.

### **2- Requirements**

From a technical point of view, *r.finder* is a GRASS-GIS script using unix bash-shell commands, and the specific syntax of GRASS-GIS modules. Thus, it can only run inside GRASS-GIS software on Unix based operating systems.

- Technical requirements:

Linux platform with Bash-shell terminal

GRASS-GIS (rel. 6.4 or later)

unix sed

unix tr  
unix bc  
unix awk

### 3- GUI structure

As for the most part of GRASS-GIS modules, *r.finder* may be used both on command line and by Graphical user Interface. The GUI is structured in two tabs: *basic* and *advanced*. The first contains all the required input items, and allows a complete use of the script in a basic way. So, for the basic user, there's no need to turn on the *advanced* tab.

#### 3.1- Basic usage

In its basic tab (fig.5) *r.finder* contains all the fields whose content is needed for a basic use, to make as simple as possible the script functions.

The fields are:

- Name of input map (an existing GRASS raster map): insert here the name of the raster map containing the valid areas (elements, sites) whose location is already known.
- Name of qualitative thematic raster map(s) (one or more existing GRASS raster maps): insert here the name of the thematic raster map(s) to be analyzed following the qualitative approach (see below)
- Name of quantitative thematic raster map(s) (one or more existing GRASS raster maps): insert here the name of the thematic raster map(s) to be analyzed following the quantitative approach (see below)
- Name of output raster map containing results: insert here the name you want to give to the raster map containing the analysis results.
- Overwrite flag: check to overwrite existing map with the same name.
- Use standard deviation instead of range for quantitative map(s) flag

#### 3.1.1- Qualitative approach

For each qualitative map the analysis is performed checking for cells with any single category value of the thematic maps in which the occurrences fall (if the the occurrences fall on cells with value of 3, 24, 394 in the qualitative map 'soil type', only such values are considered). The related validity map consists in a region map where valid areas cells (= cells whose value category contains at least one occurrence) have value 1 and non valid areas cells are set to 0.

#### 3.1.2- Quantitative approach

For each quantitative map the analysis is performed checking for cells inside the whole range between the lowest and the highest value (if the the occurrences fall on cells with value of 3, 24, 394 in the quantitative map 'terrain height', the whole range between 3 and 394 are considered valid). The related validity map consists in a region map where valid areas cells (= cells whose value category is contained in the range of categories containing at least one occurrence) have value 1 and non valid areas are set to 0.

It is also possible to use standard deviation instead of the whole range: dealing with the analysis of quantitative maps, the whole range of value between *min* and *max* may be sometimes misleading, due to the isolated values on the distribution tail. *r.finder* allow to use standard deviation instead, checking the -d flag (“Use standard deviation instead of range for quantitative map(s)”).

Warning: if the script runs on more than one quantitative map, the option is automatically applied to all the maps considered.

### 3.2- Advanced use

In its advanced tab (fig.6) *r.finder* contains the fields related to an advanced use of the script.

The fields are:

- -k option flag: allows to keep current region settings.
- -s option flag: allows to save statistical reports and histograms on data distribution.
- -r option flag: allows to save a rules file storing informations on data distribution.
- -t option flag: allows to save a template rules file.
- Optional *r.finder* rules path: allows to indicate a previously created/edited rules files to be used for the analysis instead of input file (in the basic tab)
- Optional comments for rules file: allows to add a comment text for rules file. Unless the flag -r is checked, this field is ignored.

#### 3.2.1- Region setting

By default, *r.finder* sets the working region on the size of the input raster map, that seems to be the most correct default approach to the analysis for non-expert use. The option -k, if checked, allows to skip the setting and keep the current region settings. Thus, it is possible to apply the input map parameters to a wider area in which the input map region is contained. For using the input map parameters to analyze a different area (=not containing the input map region), the use of rules files is needed (see below).

Dealing with raster analysis, it is important to note that the original data resolution is deeply correlated to the reliability of final result. For this reason the option to convert a vector sites map to a raster one (*v.to.rast*) to be used for input has not been included in the script. The decision on which resolution to be set converting the vectors is crucial, for the comparison with thematic maps, and affect the result reliability.

#### 3.2.2- Working with statistical reports

Checking the option -s (Save statistical reports) makes *r.finder* create a folder in your home directory named as “[name of the output]-data” containing, for each thematic map used, a textual report (*r.report*) and an histogram bar graph (*d.histogram*) of cell number for each raster value both for the whole thematic map and for the only cells corresponding to known occurrences in the input map. Thus, each thematic map will produce 4 files.

This can be helpful to test the distribution of occurrence cells in comparison to the general one of the map. In statistical terms, similarity between the two distributions (all the cells VS occurrence cells) should indicate randomness distribution and non correlation of the occurrence distribution to the map theme; otherwise, differences should be index of preference criteria. Moreover, the distribution histogram may evidence particular features in distribution shapes, such as multi-modal trends, that can be used for rules file formulation.

### 3.2.3- Working with rules files

Once identified some thematic maps whose combination is worth considering to focus the areas of interest, it may be useful to store such informations to be used also in regions far from the one where our known sites are located. For instance, if we check that all the elements lie on places between 300 and 400 mt asl, we may try to use such a filter also to analyze a different territory.

*r.finder* may store the resulting filter informations of any analysis performed in rules files, checking the -r flag (“save rules file”). Thus, a rule text file is created (.rul) in a specific folder in your home directory (named as your output file).

The following is a rules file example:

```
Title of the rules file
-----
[text (You can add, edit or remove whichever comment here)]
-----

MAP RULES BOX: THIS IS THE PART OF THE TEXT FILE READ AT RUNTIME:

#(qual)soil,categories:34;21;210;456#
#(quant)slope,ranges:0-30#
#(qual)landuse,categories:3;7;19;23;35;37#
#(quant)dem,ranges:0-1000#

-----
```

The dashed lines separate different sections: title, comment, map rules data. Each part may be manually edited to perform specific analysis. The data format is easily readable and self explaining. Note that – although the rules file produced by the script has a single range for each quant map used – editing manually the file it is possible to set for each quantitative map more data ranges, in order to study multimodal distributions (that is not possible in basic usage).

To use a rule file instead of an input map, you may just insert the path and name of the rules file browsing your pc, in the “optional *r.finder* rules file to be used instead valid areas (.rul)” field. Moreover, a rule file can be compiled completely manual, to check research hypothesis. For helping in such approach a specific function (flag f “Save a Template rules file and exit”) allow to create (in a specific folder in your home directory) a neutral template as follows, to be manually edited.

```
[title]
-----
[text (You can add, edit or remove whichever comment here)]
-----

MAP RULES BOX: THIS IS THE PART OF THE TEXT FILE READ AT RUNTIME:

#(qual)mapnameA,categories:0;0;0#
#(quant)mapnameB,ranges:0-0#
#(qual)mapnameC,categories:0;0;0;0;0;0#
#(quant)mapnameE,ranges:0-0;0-0;0-0#

-----
```

## 4. Examples

### 4.1. Basic usage

Some occurrences lie on a territory, they are mapped in the map 'sites' (Fig. 1). For the same area are available: the height map ('dem'), the 'slope' map, the 'soil' type map (Fig. 2,3,4).

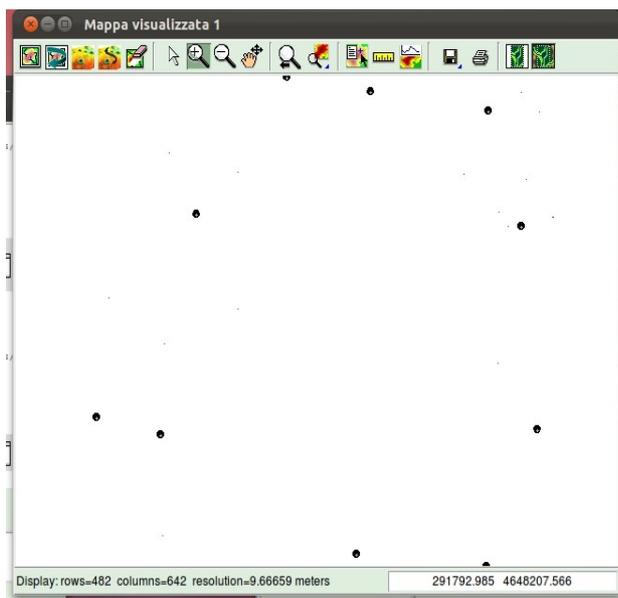


Fig. 1: Site map

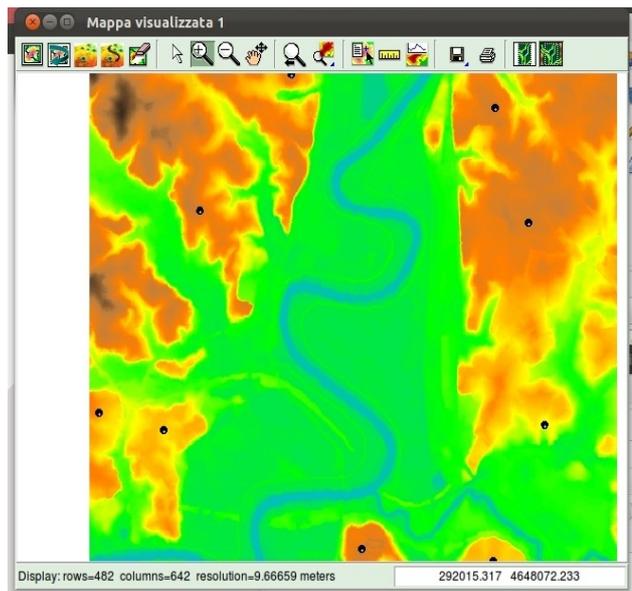


Fig. 2: Digital Elevation Model

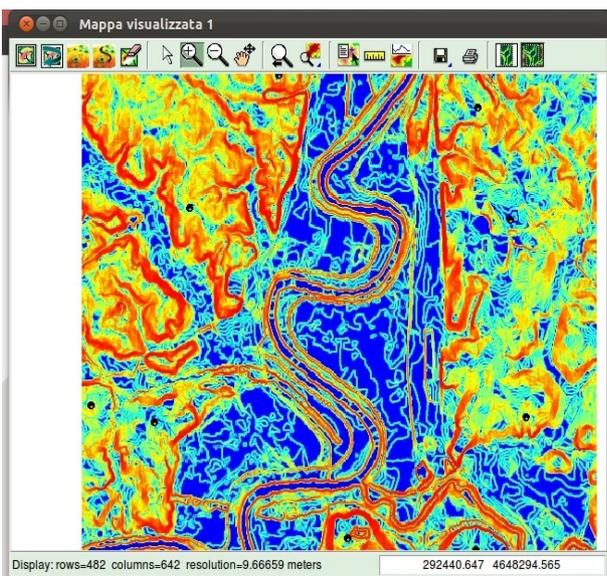


Fig. 3: Slope map

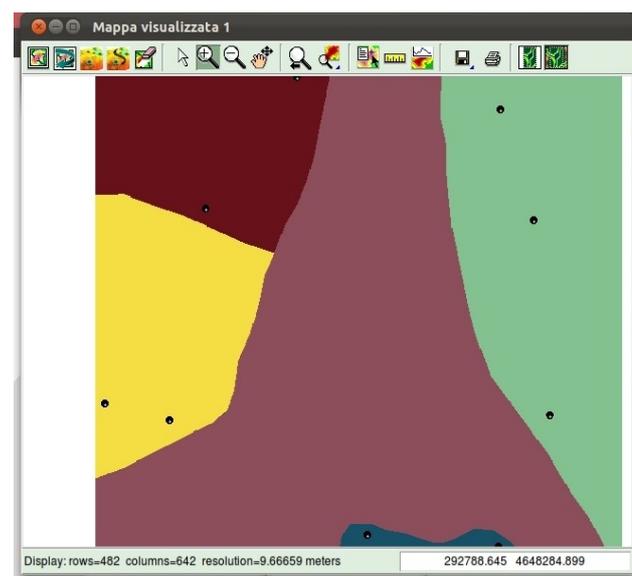


Fig. 4: Soil type map

We run `r.finder` to check on which categories of the soil map (qual maps: 'soil' ) and in which

category range (quant maps: 'dem', 'slope') items are present.

r.finder input=sites qual=soil quant=dem,slope output=result (Fig.5)

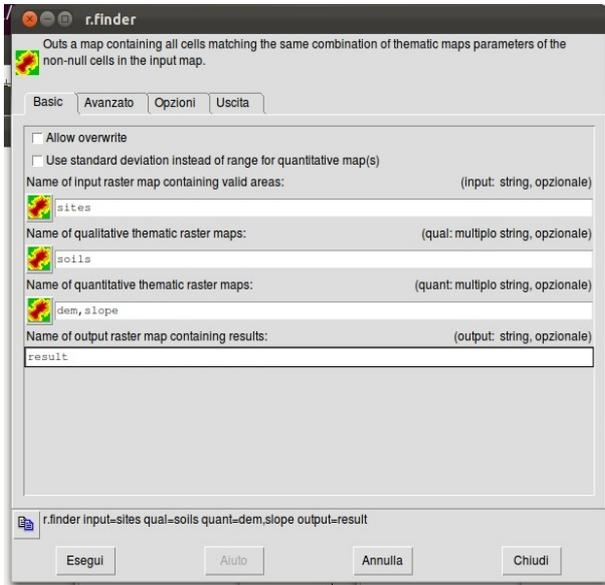


Fig. 5: GUI “Basic” tab

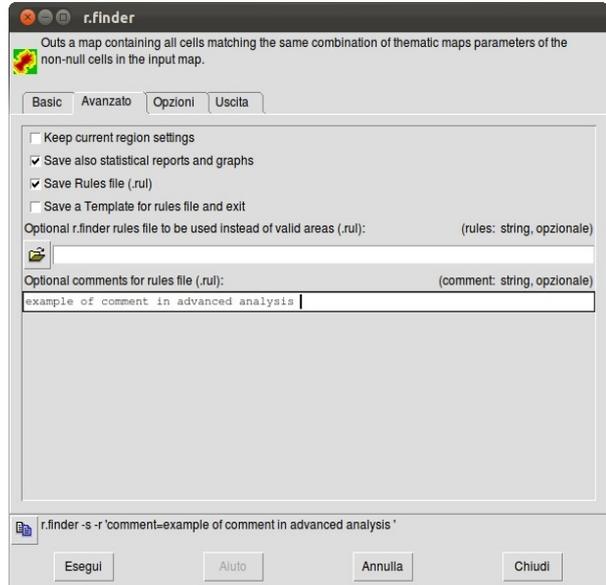


Fig.6: GUI “Advanced” tab

We obtain the map showing the higher or lesser concentration of analogies with the cells on which elements fall (Fig. 7) or, using the standard deviation instead of whole ranges for quantitative maps (option -d), (Fig. 8)

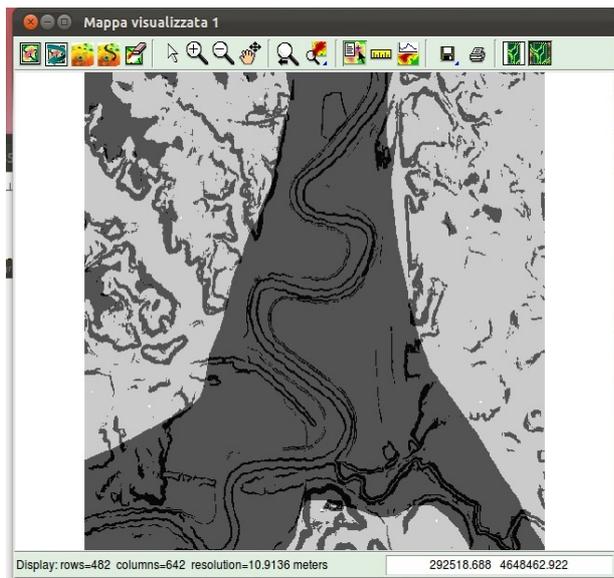


Fig. 7: Final result map

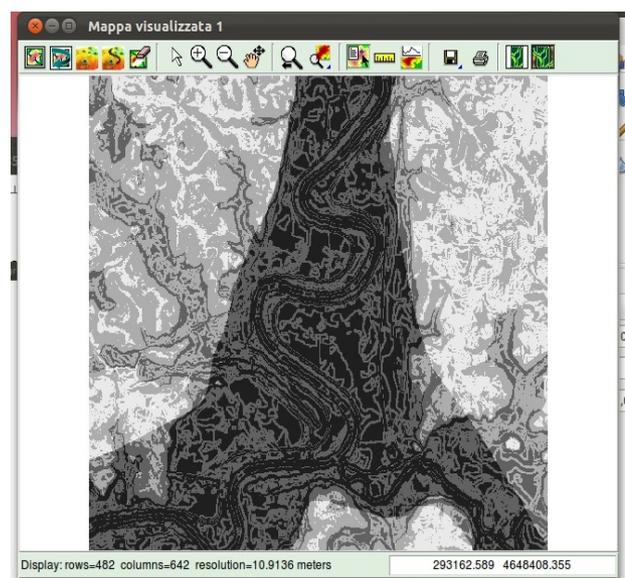


Fig.8: Final result map using standard deviation

## 4.2 Advanced use

We may perform the previous operation adding the option `-r` (Save Rules file) and `-s` (Save also statistical data) (fig.6). Then we will have a rules file containing the rules (optionally containing some comment text), saved in the subfolder `/result` (=named from the output map) of our home directory:

```
/home/user/result-data/result-rules.rul
```

the rules file will look as follows:

```
rules file
-----

[...]

NOTES:
-----

MAP RULES BOX: THIS IS THE PART OF THE TEXT FILE READ AT RUNTIME:

#(qual)Litostratigrafia,categories:0;26;48;80;87#
#(quant)DemNuovoLod210mAnticoLug07,ranges:0-78#
#(quant)DemNuovoLod210mAnticoLug07SLOPE,ranges:0-21.3588809967041#
-----
```

The statistical function will produce two histograms for each thematic map, respectively showing the general distribution of cells in the whole map and the distribution of cells corresponding to the presences. For the DEM, in this situation, we will have the histograms in fig. 9 and 10

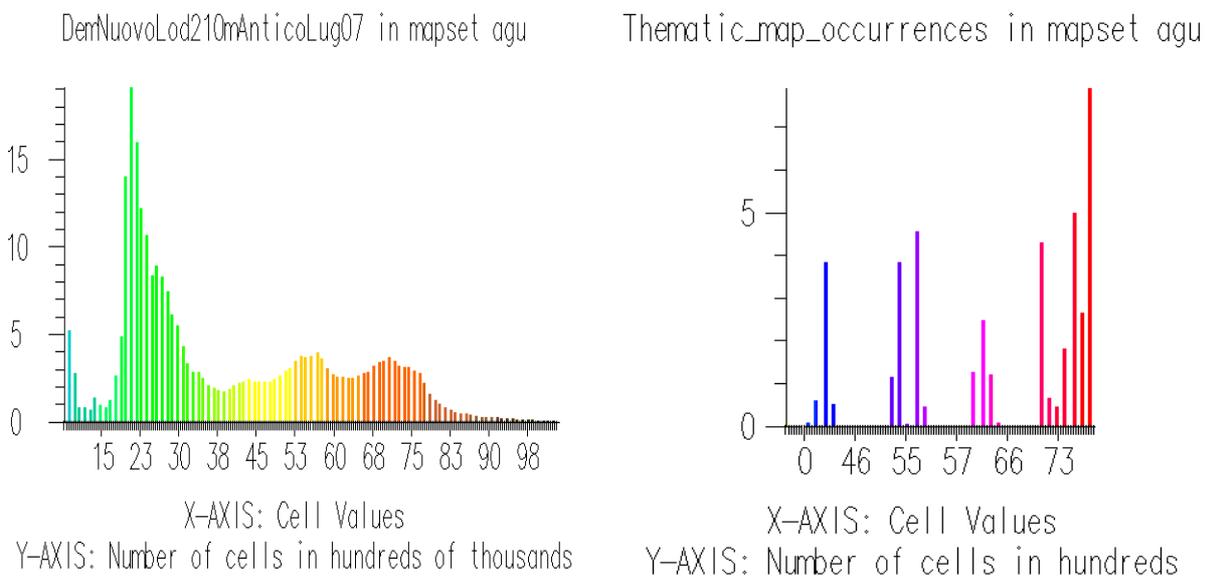


Fig. 9: DEM cell distribution

Fig.10: Presence cells distribution in DEM map

Checking the statistical distribution of sites (Fig.10) we may have the idea that they have a multimodal distribution (0-40 mt, 50-65 mt, 70-78 mt).

Then, we can modify the rules file about the data range of this map in the relative row:

```
# (quant) DemNuovoLod210mAnticoLug07, ranges:0-78#
```

changing in:

```
# (quant) DemNuovoLod210mAnticoLug07, ranges:0-40;50-65;70-78#
```

Thus, we can refer to another area using the same filter rules. We change our working region and then tell the script to use rules file and keep the current region (options -rk):

```
r.finder -rk qual=soil quant=dem,slope output=result2 rules=/home/user/result/result-rules.rul
```

This is intended as a simple example, just to show the script capabilities.

## 5. Warnings and known bugs

Due to the math library used (bc), *r.finder* in "quant" mode works properly only with double decimal numbers. In case of cell values with more than two decimals, only the first two are evaluated in order to set the range. Otherwise, for "qual" maps, values are considered as strings, then whichever number can be used. See 3.2.1 section for warnings related to map resolution settings.

Another warning is connected to map extension. If the extension of any of the thematic map used is smaller than the presence map, the default use of the script (without -k option) will make occurrences result as on null cells, strongly influencing final result.

## 6. Credits and bug report

Released under the GPL by Augusto Palombini. Refer to [augusto.palombini@itabc.cnr.it](mailto:augusto.palombini@itabc.cnr.it)

## 7. Appendix: error messages

Error message	cause	solution
"You must be in GRASS GIS to run this program."	The script can run only inside GRASS environment	Start GRASS and then run the script in the GRASS terminal
"sed [or: awk, bc, tr] required, please install sed first"	Unix <i>sed</i> , <i>awk</i> , <i>bc</i> , <i>tr</i> , are requested to run the script	Install the lacking program and re-start GRASS and r.finder
"You must provide an output file name or create a template rules file with flag -t. Exiting"	The script can't work if the output file name is not specified, unless you only want to create a template rules file (flag -t)	Specify an output file name in the basic tab, or check the "save template rules file" (-t) flag in the advanced tab
"XXX is not a valid rules file. See manual pages. Exiting"	You specified a rules file not compliant with the correct format	A rules file must follow a particular format. See sections 3.2.3 and 4.2 to edit your file.
"WARNING! File XXX exists. You can use -o flag to overwrite. Exiting."	You specified an already existing file as output file name.	Change the output file name or check the "allow overwrite" (-o) flag if you want to overwrite an existing file
"WARNING! Input raster file or rules text file must be entered. Exiting."	You omitted to specify an input raster file or a rules file.	Specify an input raster file or a rules file name.
"WARNING! Either input raster file or rules text file must be entered. Exiting."	You specified both an input raster file and a rules file. Both files are not admitted.	Specify an input raster file <b>or</b> a rules file name.
"WARNING! Both reading and writing rules file: operation not permitted. Exiting."	The script can not read and write a rules file at the same time.	Specify a rules file to be read or check the "save rules file" (-r) option in the advanced tab
"WARNING! Both QUAL input raster file(s) and rules text file have been entered. Only rules file will be considered."	You specified both qualitative input raster file(s) and a rules file.	The script runs correctly but the user-specified qual file is ignored and only the rules file indications are considered
"WARNING! Both QUANT input raster file(s) and rules text file have been entered. Only rules file will be considered."	You specified both quantitative input raster file(s) and a rules file.	The script runs correctly but the user-specified quant file is ignored and only the rules file indications are considered
"WARNING! Use of standard deviation for rules file is not possible. Exiting."	You both specified a rules file and checked the standard deviation flag: if quantitative intervals are specified by the rules file, stdev flag has no sense.	Specify an input rules file <b>or</b> specify quant file(s) name checking the stdev flag
"WARNING! File XXX seems to be floating. You can't use a floating point raster as 'qual' map"	Qualitative raster maps are related to qualitative terrain categories: floating values maps are not allowed	Use integer rasters as qual maps

