74th Meeting of the European working group " Multiple Criteria Decision Aiding"

Yverdon-les-Bains, Swisse, October 6-8, 2011

Implementation of Multicriteria Modules in a Geographic Information System

Antonio Boggia
University of Perugia, Italy
Salvatore Greco
University of Catania, Italy
Gianluca Massei
University of Perugia, Italy

Objective

The objective of this study is to present the implementation of five modules in an open source GIS system; four are based on MCDA:

-r.mcda.electre;

-r.mcda.fuzzy;

-r.mcda.regime;

-r.mcda.roughset;

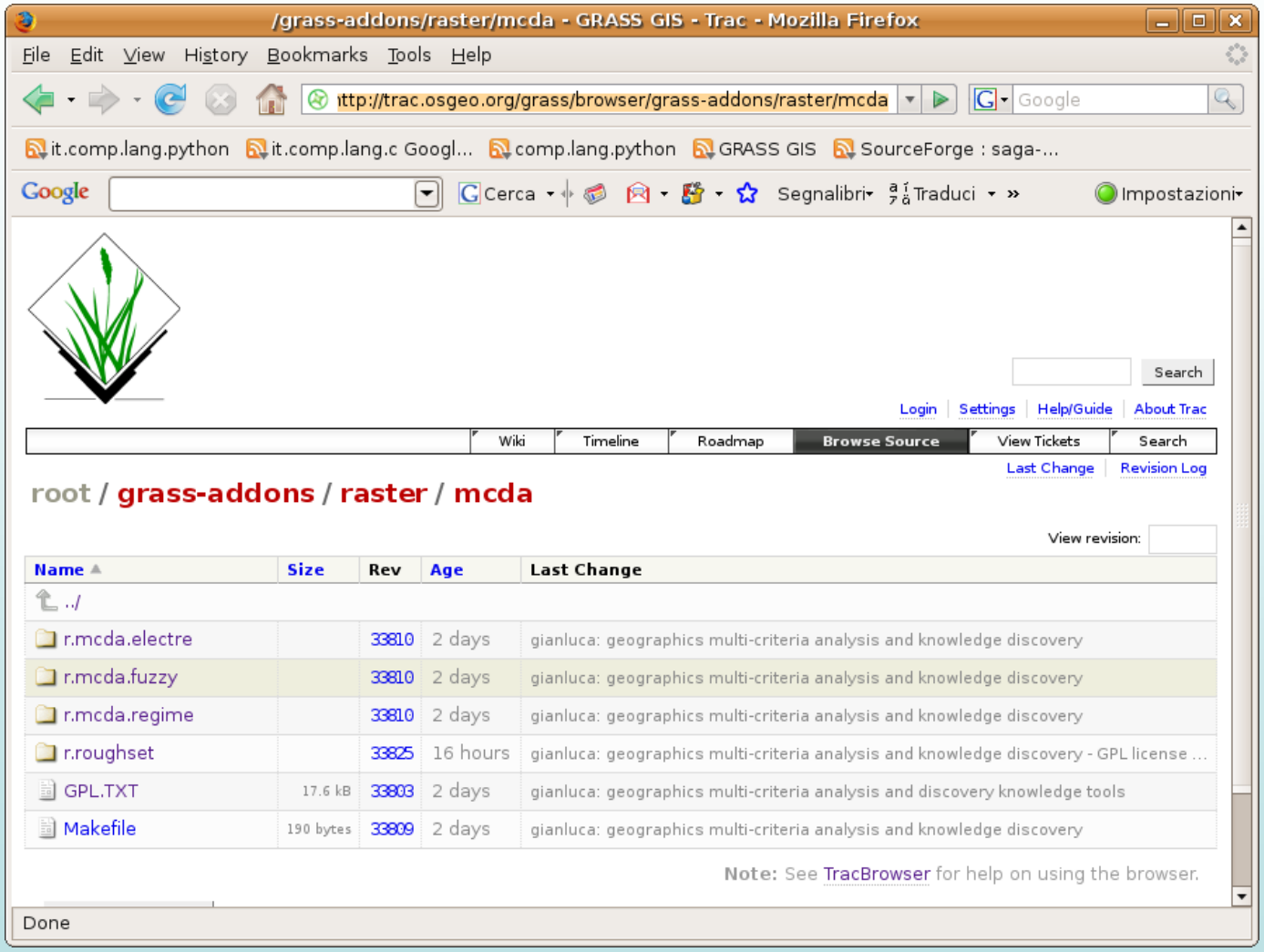
-r.roughset.

With particular focus on the r.mcda.roughset one, using a case study.

Plan

- MCDA modules in Grass GIS
- r.mcda.roughset
- Case Study
- Conclusion

MCDA modules in Grass GIS



MCDA modules in Grass GIS

GRASS 6.4 svn is a software GIS open source (GPL licence). It is advanced, well known and rich in functionalities (http://grass.osgeo.org).

GRASS *" is free Geographic Information System (GIS) software used for geospatial data management and analysis, image processing, graphics/maps production, spatial modelling, and visualization. GRASS is currently used in academic and commercial settings around the world, as well as by many governmental agencies and environmental consulting companies".*

It is written in C language and it is possible, thanks to its libraries and GPL licence, to develop new modules.

MCDA-GIS integration models

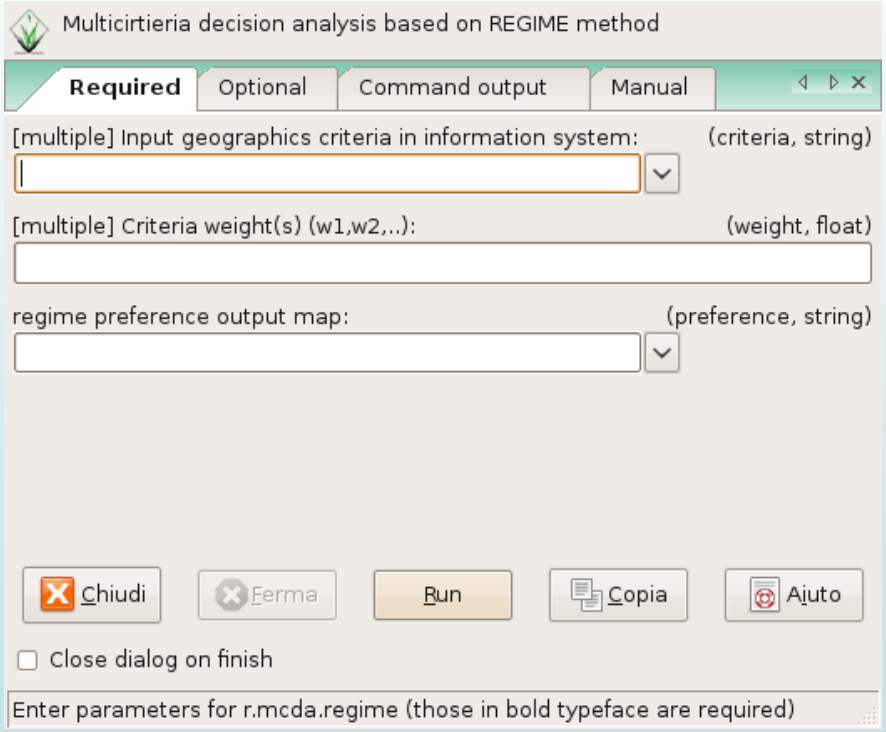
☑ **MCDA-GIS indirect integration**: MCDA and GIS models are separate and connected trough an *intermediate connection system*;

☑ ***Built-in* MCDA-GIS models**: MCDA model is a component integrated in the GIS system, but it is independent from the logical and functional point of view;

☑ **Complete MCDA-GIS integration**: MCDA and GIS systems use the same interface and the same database. The MCDA model is activated inside the GIS software just like any other analysis function.

MCDA modules in Grass GIS

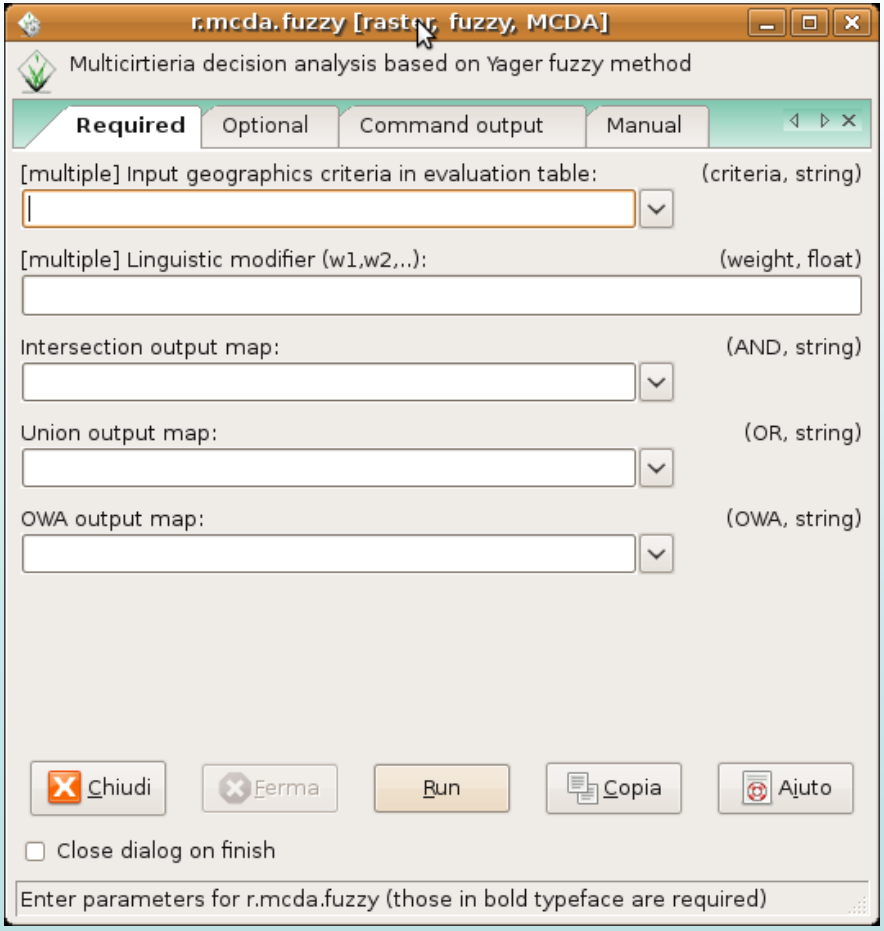
r.mcda.regime



It is the implementation of the REGIME multicriteria algorithm in GRASS GIS environment. The input required is the list of raster representing the criteria to be assessed in the multicriteria evaluation and the vector of weights to be assigned. Every single cell of the GRASS region is considered as one of the possible alternatives to evaluate and it is described with the value assumed for the same cell by the raster used as criteria.

MCDA modules in Grass GIS

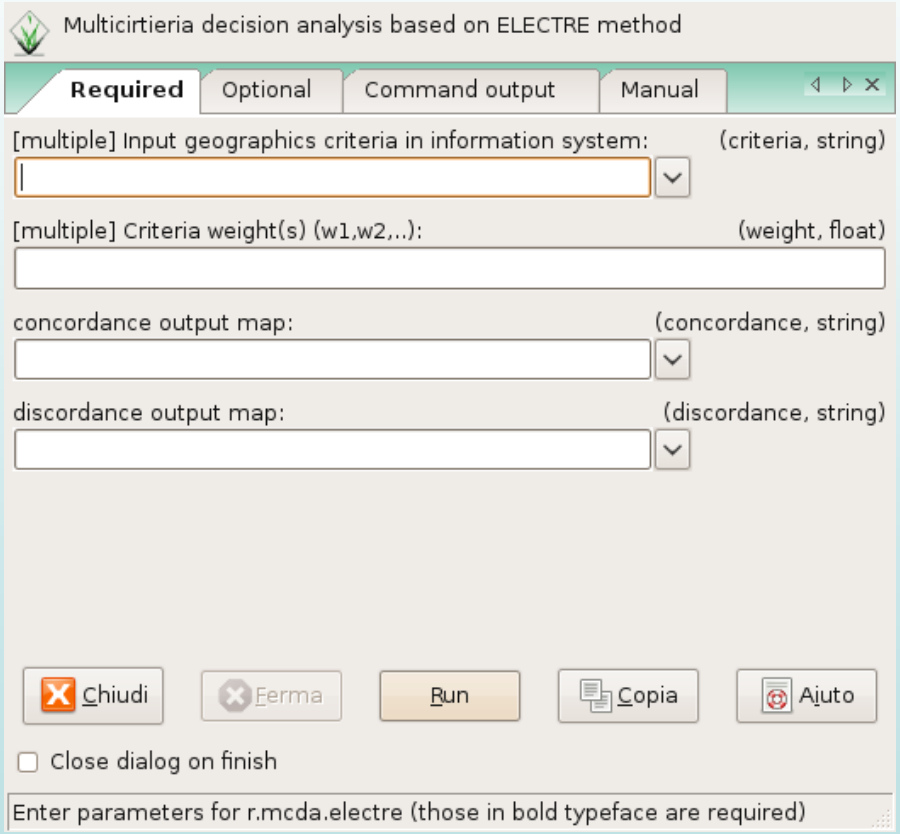
r.mcda.fuzzy



It is the implementation of the FUZZY multicriteria algorithm proposed by Yager R., in GRASS GIS environment. The input required is the list of raster representing the criteria to be assessed in the multicriteria evaluation and the vector of linguistic modifiers to be assigned. Every single cell of the GRASS region is considered as one of the possible alternatives to evaluate and it is described with the value assumed for the same cell by the raster used as criteria. It is possible to get three different output files as the result of the intersection operator, the union operator and the ordered weighted averaging (OWA) operator.

MCDA modules in Grass GIS

r.mcda.electre



It is the implementation of the ELECTRE multicriteria algorithm in GRASS GIS environment.

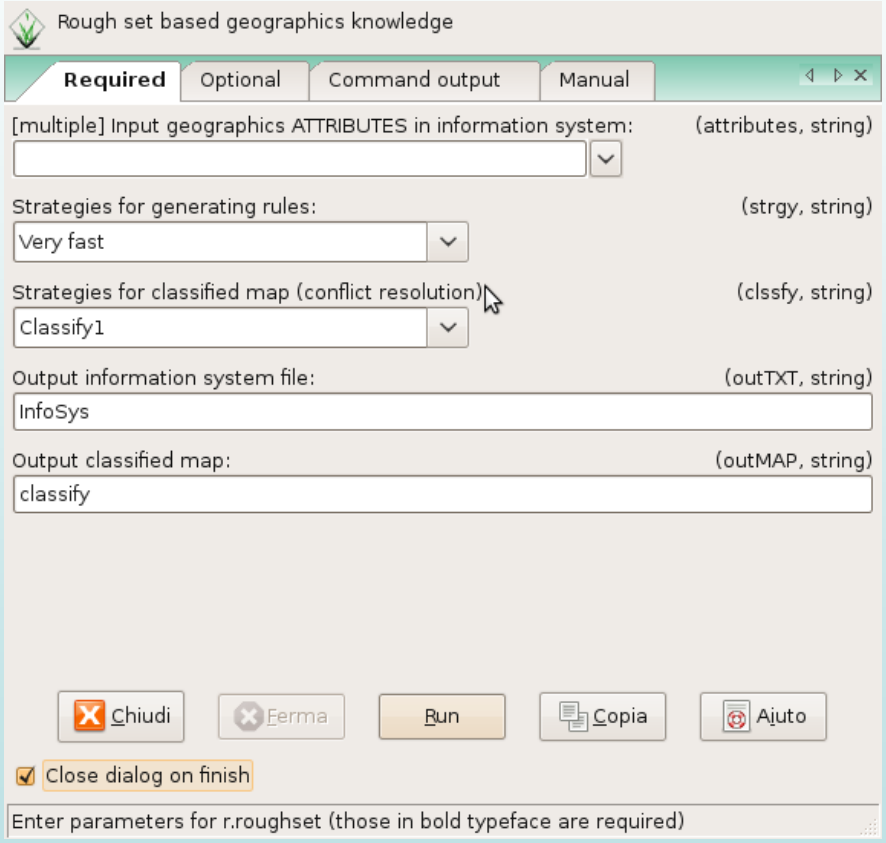
Input: the list of raster representing the criteria to be assessed in the multicriteria evaluation and the vector of weights to be assigned.

Alternatives: Every single cell of the GRASS region is considered as one of the possible alternatives to evaluate and it is described with the value assumed for the same cell by the raster used as criteria.

Output: There are two output files. One represents the spatial distribution of the concordance index, the other one of the discordance index. The optimal solution is the one presenting the maximum concordance value and the minimum discordance value at the same time.

r.rough set module in Grass GIS

r.rough set



It is the implementation of the rough set theory in GRASS GIS environment. It requires the following input:

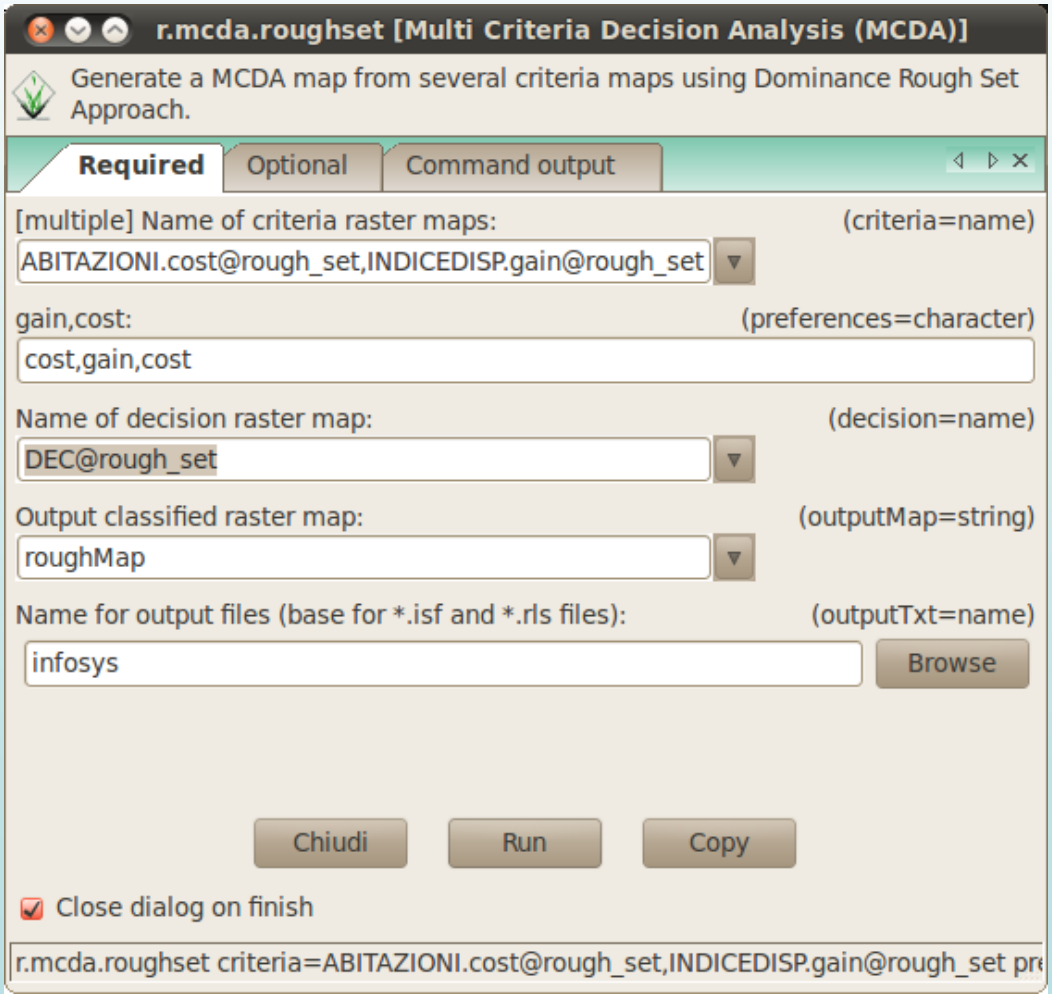
1. the geographical attributes constituting the information system for the rough set analysis; they have to describe environmental, economic or social issues
2. the theme in which areas with the issues to be studied are identified.

An information system is generated and the functions of version 2 of rough set library (RSL, ver. 2.0) are applied, following one of the decision strategies available with the option

The choice of the classification strategy, is necessary to generate a map based on the attributes and the decision rules.

r.mcda.roughset

Implementing Dominace based Rough Set Approach (Greco, S., Matarazzo, B., Slowinski, R., 1999)



Dominance-based Rough Set Approach (DRSA)

What type of Rough Set Approach?

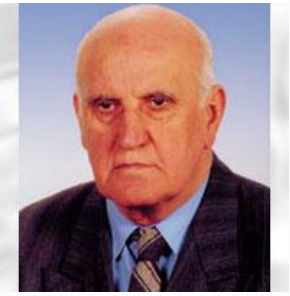
- In the last decade, an extension of Classical Rough Set Theory permitting to deal with decision problems has been proposed,
  applied, and thoroughly investigated

Greco, S., Matarazzo, B., Słowiński, R.: Rough sets theory for multicriteria decision analysis.
European J. of Operational Research, 129 (2001) no.1, 1-47

- It is called **Dominance-based Rough Set Approach** (**DRSA**)

- DRSA proved to be useful in many real world applications
  ranging from investment analysis to credit risk evaluation, from customer satisfaction analysis to technical and medical diagnoses

Classical Rough Set Approach (Z.Pawlak)



- The granules of indiscernible objects are used to approximate classes

Student	Mathematics (M)	Physics (Ph)	Literature (L)	Overall class
S1	good	medium	bad	bad
S2	medium	medium	bad	medium
S3	medium	medium	medium	medium
S4	medium	medium	medium	good
S5	good	medium	good	good
S6	good	good	good	good
S7	bad	bad	medium	bad
S8	bad	bad	medium	bad

Classical Rough Set Approach (Z.Pawlak)

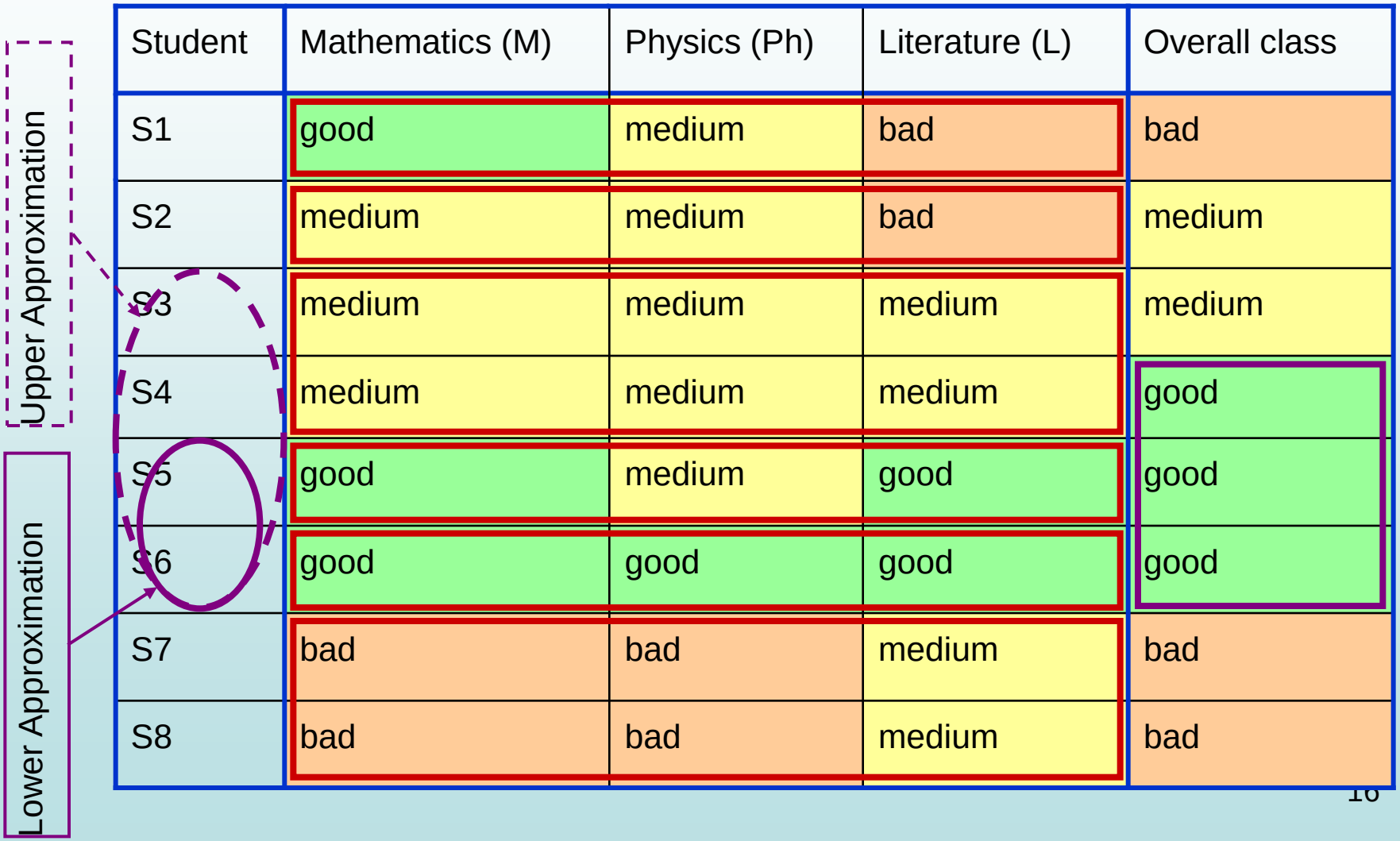
- Lower approximation of class „good”

Student	Mathematics (M)	Physics (Ph)	Literature (L)	Overall class
S1	good	medium	bad	bad
S2	medium	medium	bad	medium
S3	medium	medium	medium	medium
S4	medium	medium	medium	good
S5	good	medium	good	good
S6	good	good	good	good
S7	bad	bad	medium	bad
S8	bad	bad	medium	bad

Lower Approximation

Classical Rough Set Approach (Z.Pawlak)

- Lower and upper approximation of class „good”



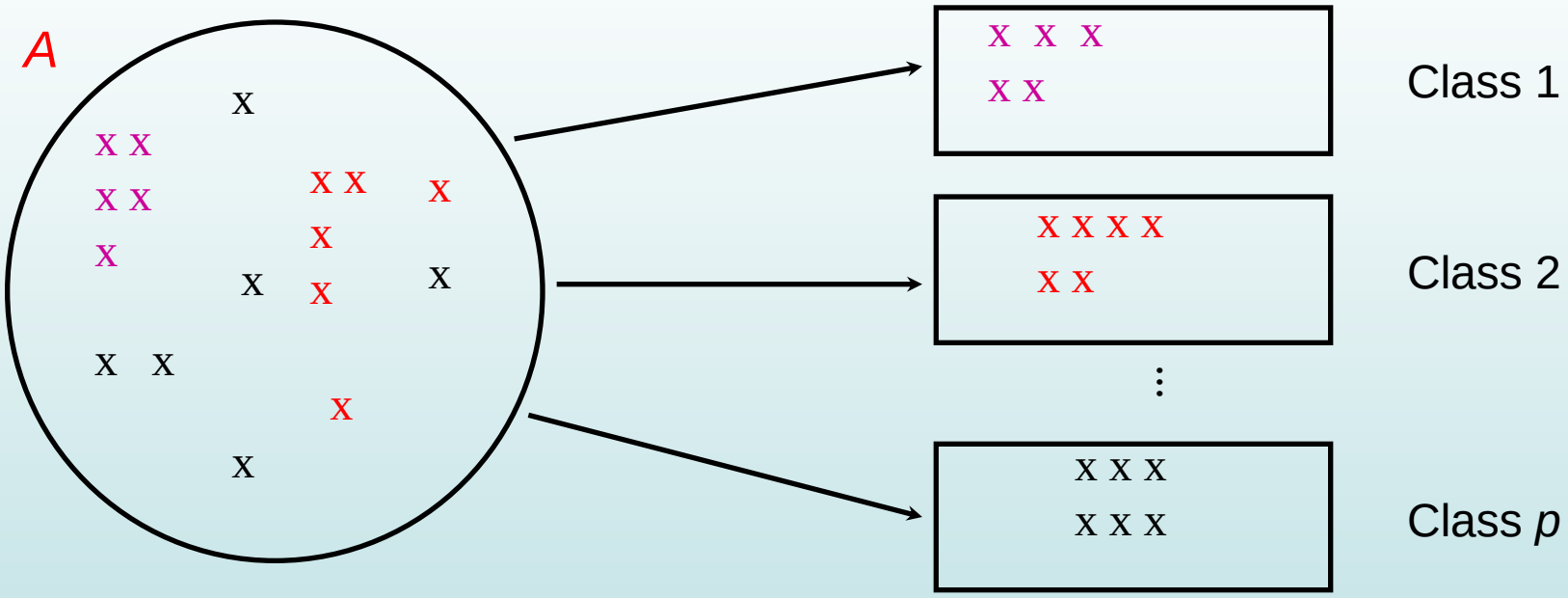
Student	Mathematics (M)	Physics (Ph)	Literature (L)	Overall class
S1	good	medium	bad	bad
S2	medium	medium	bad	medium
S3	medium	medium	medium	medium
S4	medium	medium	medium	good
S5	good	medium	good	good
S6	good	good	good	good
S7	bad	bad	medium	bad
S8	bad	bad	medium	bad

Classical Rough Set Approach (Z.Pawlak)

- The information in the decision table can be represented by the following “if ..., then ...” decision rules
- 1) If Literature=“good”, then the student is good (s5,s6)
- 2) If Mathematics=“bad”, then the student is bad (s7,s8)
- 3) If Mathematics=“good” and Literature=“bad”, then the student is bad (s1)
- 4) If Mathematics=“medium” and Literature=“bad”, then the student is medium (s2)
- 5) If Physics=“medium” and Literature=“medium”, then the student is medium or good (s3,s4)

Classification in strict sense and
ordinal classification (sorting)

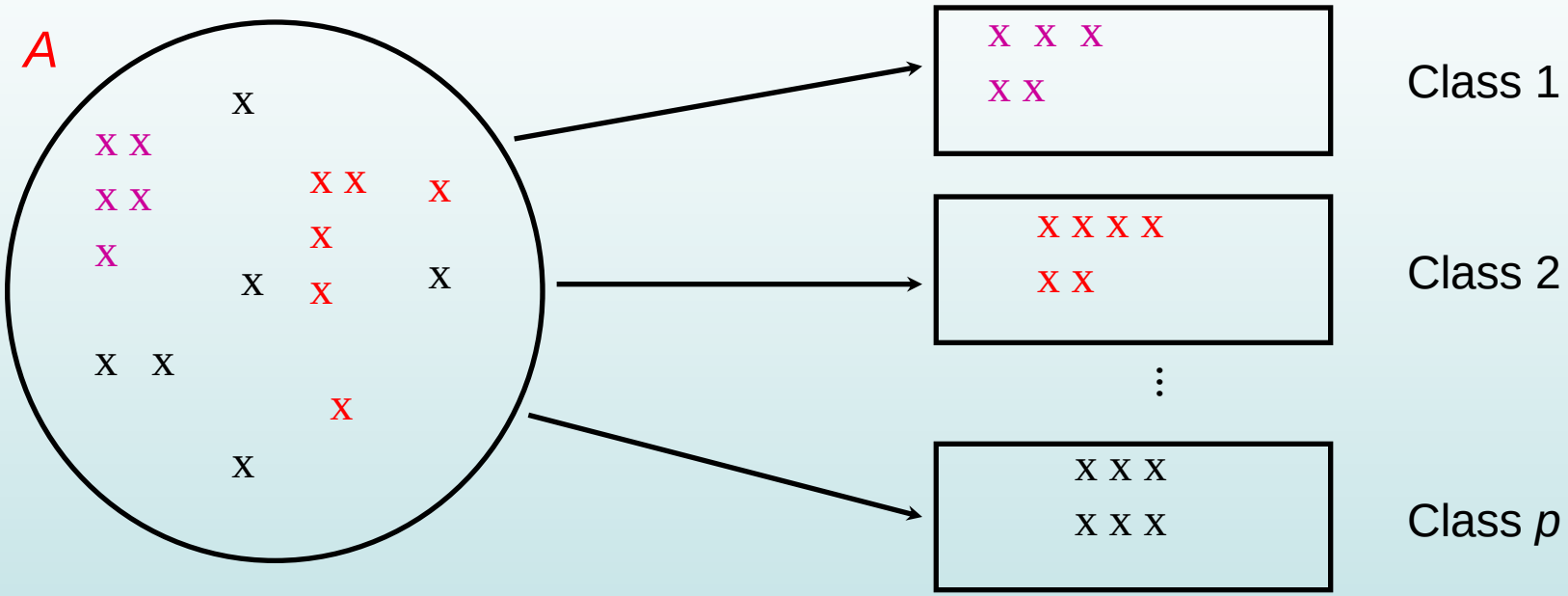
Classification to preferentially non-ordered classes (classification in the strict sense)



Classification in the strict sense – medical diagnosis

Patients	Temperature	Dry-Cough	Headache	Muscle pain	Flu
1	normal	absent	absent	absent	absent
2	normal	absent	present	present	absent
3	subfeb.	absent	present	present	present
4	subfeb.	present	absent	absent	absent
5	subfeb.	present	absent	absent	present
6	high	absent	absent	absent	absent
7	high	present	absent	absent	absent
8	high	present	absent	absent	present
9	high	present	present	absent	present

Ordinal classification to preferentially ordered classes (sorting)



Class 1 $\succ$ Class 2 $\succ$ ... $\succ$ Class *p*

Ordinal classification – example of multiple criteria sorting of students

Student	Mathematics (M)	Physics (Ph)	Literature (L)	Overall class
S1	good	medium	bad	bad
S2	medium	medium	bad	medium
S3	medium	medium	medium	medium
S4	medium	medium	medium	good
S5	good	medium	good	good
S6	good	good	good	good
S7	bad	bad	bad	bad
S8	bad	bad	medium	bad

Rough set approach and decision problems

- **The only class of decision problems which can be dealt with Classical Rough Set Theory is classification in the strict sense**
- This class is rather limited (even if it is practically the only class of decision problems considered within data mining and knowledge discovery!)
- Why classical rough set approach is not useful to deal with choice problems, sorting problems and ranking problems?
- **Because Classical Rough Set Approach does not consider preferences**

Why Classical Rough Set Approach has to be adapted to MCDM?

- Classical rough set approach does not detect inconsistency w.r.t. dominance (Pareto principle)

Student	Mathematics (M)	Physics (Ph)	Literature (L)	Overall class
S1	good	medium	bad	bad
S2	medium	medium	bad	medium
S3	medium	medium	medium	medium
S4	medium	medium	medium	good
S5	good	medium	good	good
S6	good	good	good	good
S7	bad	bad	bad	bad
S8	bad	bad	medium	bad

Classical Rough Set Theory vs. Dominance-based Rough Set Theory from indiscernibility principle to dominance principle

Classical Rough Set Theory

⇓

Indiscernibility principle

If *x* and *y* are indiscernible with respect to all relevant **attributes**,

then *x* should be classified to the same class as *y*

Dominace-based Rough Set Theory

⇓

Dominance principle

If *x* is at least as good as *y* with respect to all relevant **criteria**,

then *x* should be classified at least as good as *y*

S.Greco, B.Matarazzo, R.Słowiński: Rough sets theory for multicriteria decision analysis. *European J. of Operational Research*, 129 (2001) no.1, 1-47

Rough Set approach to multiple-criteria sorting

- Set of decision rules in terms of {M,Ph,L} representing preferences:

If L $\succeq$ good, *then* student $\succeq$ good {S5,S6}

If M $\succeq$ medium & L $\succeq$ medium, *then* student $\succeq$ medium {S3,S4,S5,S6}

If M $\succeq$ medium & L $\preceq$ bad, *then* student is bad or medium {S1,S2}

If M $\preceq$ bad, *then* student $\preceq$ bad {S7,S8}

If L $\preceq$ bad, *then* student $\preceq$ medium {S1,S2,S7}

Greco, S., Matarazzo, B., Słowiński, R.: Axiomatic characterization of a general utility function and its particular cases in terms of conjoint measurement and rough-set decision rules. *European J. of Operational Research*, 158 (2004) no. 2, 271-292

Using DRSA rules as a decision model

- New student to be evaluated

Student	Mathematics	Physics	Literature
S9	Medium	Medium	Good

- Set of activated decision rules:

If L $\succeq$ good, *then* student $\succeq$ good {S5,S6}

If M $\succeq$ medium & L $\succeq$ medium, *then* student $\succeq$ medium {S3,S4,S5,S6}

- Set of non-activated decision rules:

If M $\succeq$ medium & L $\preceq$ bad, *then* student is bad or medium {S1,S2}

If M $\preceq$ bad, *then* student $\preceq$ bad {S7,S8}

If L $\preceq$ bad, *then* student $\preceq$ medium {S1,S2,S7}

Using DRSA rules as a decision model: relevant rules

- New student to be evaluated

Student	Mathematics	Physics	Literature
S9	Medium	Medium	Good

- Set of activated decision rules:

If L $\succeq$ good, *then* student $\succeq$ good {S5,S6}

If M $\succeq$ medium & L $\succeq$ medium, *then* student $\succeq$ medium {S3,S4,S5,S6}

- Set of non-activated decision rules:

If M $\succeq$ medium & L $\preceq$ bad, *then* student is bad or medium {S1,S2}

If M $\preceq$ bad, *then* student $\preceq$ bad {S7,S8}

If L $\preceq$ bad, *then* student $\preceq$ medium {S1,S2,S7}

Using DRSA rules as a decision model

- New student to be evaluated

Student	Mathematics	Physics	Literature
S9	Medium	Medium	Good

- Set of relevant decision rules:

If L $\succeq$ good, *then* student $\succeq$ good {S5,S6}

If M $\preceq$ bad, *then* student $\preceq$ bad {S7,S8}

Conclusion: S9 is a good student because he is good in Literature and he is not bad in Mathematics

Good properties of DRSA

- DRSA has a lot of good properties:
  - Transparence
  - Use of qualitative evaluations
  - Decision analysis based on ordinal properties only
  - Traceability
- DM gives preference information by answering easy questions, and obtains transparent feedback („glass box”) in a learning oriented perspective

DRSA – preference modeling by decision rules

- A set of ($D_{\geq}$ $D_{\leq}$ $D_{\geq\leq}$)-rules induced from rough approximations represents a preference model of a Decision Maker

- Traditional preference models:
  - utility function (e.g. additive, multiplicative, associative, Choquet integral, Sugeno integral),
  - binary relation (e.g. outranking relation, fuzzy relation)

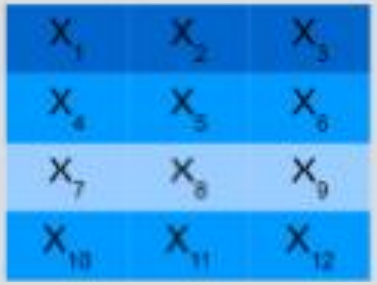
- Decision rule model is the most general model of preferences:
  a general utility function, or outranking relation exists if and only if there exists the decision rule model

Greco, S., Matarazzo, B., Słowiński, R.: Axiomatic characterization of a general utility function and its particular cases in terms of conjoint measurement and rough-set decision rules. *European J. of Operational Research*, 158 (2004) no. 2, 271-292

r.mcda.roughset

The conceptual model

Geographic Information System

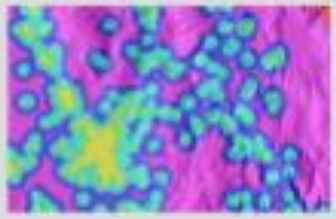


Part of land
N x M cells table

C

N x M	S	LU	E	F	S
X1	1	0	1	0	2
X2	2	1	0	1	0
X3	0	1	0	0	2
X4	2	0	0	1	0
X5	0	1	2	0	1
X6	0	1	0	0	0
X7	2	1	0	1	0
X8	1	0	1	0	0
X9	0	1	2	0	1
X10	0	1	1	0	1
X11	2	1	0	1	0
X12	2	1	0	1	0

Rough set based information system



Output map
classified according
to DRSA rules

r.mcda.roughset

Input data:

- Name of criteria raster maps
- Preferences, in term of gain and cost
- Name of decision raster map
- Name of output classified raster map
- Name for output f txt files

Output:

1. Graphic: raster map classifying the whole study area according to the rules extracted

2. Text file: containing all the rules "*if… and… and… then*"

Case Study

r.mcda.roughset applied to forest fire risk assessment

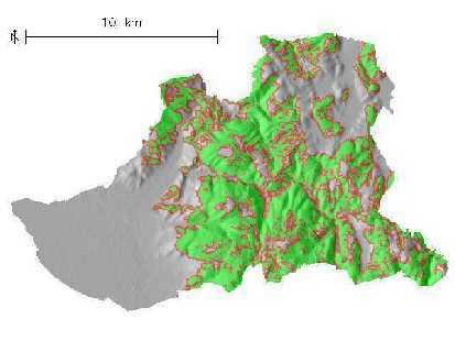
Forest fire risk assessment is suitable for experiencing the DRSA because some of the parameters certainty play an important role in the occurrence of the phenomenon, but there are others whose role is not clear and defined.

The dominance-based rough set approach allows to handle imprecise and vague information. It seems appropriate for classifying a land context in terms of risk of fire, given a plurality of thematic geographic data into a database.

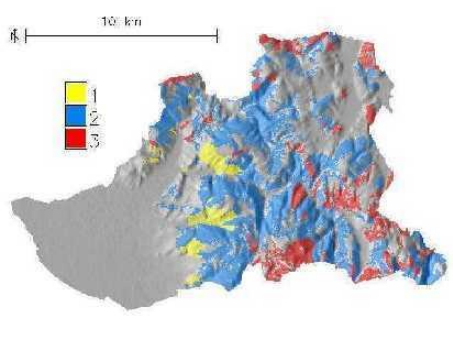
The combination of the benefits arising from the DRSA with the potential of geographic information systems further extends the interest and possibilities of application for land use and environmental assessment and management.

Case Study

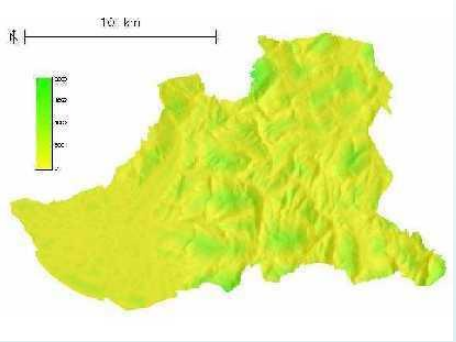
The territorial information system



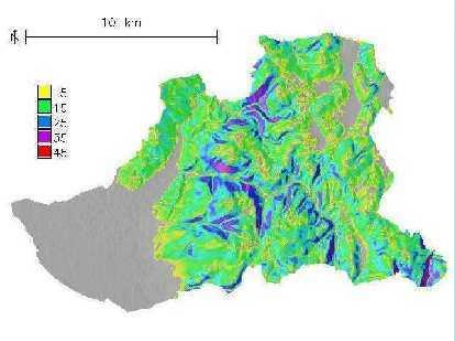
Forests and distance from the boundary (100 m)



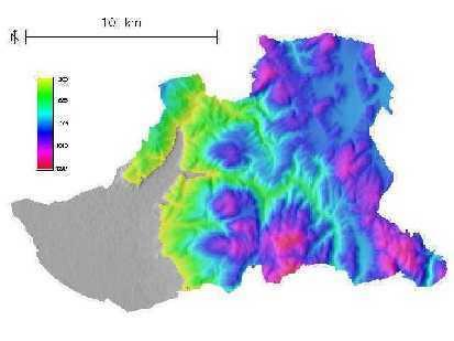
Vegetation (combustion classes)



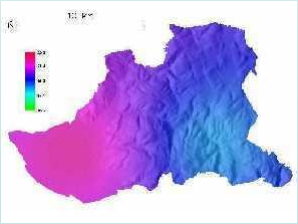
Distance from roads (m)

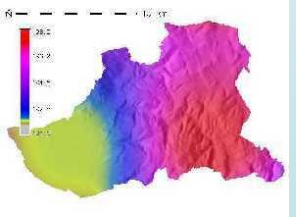


Slope (degree)



Altitude (m)

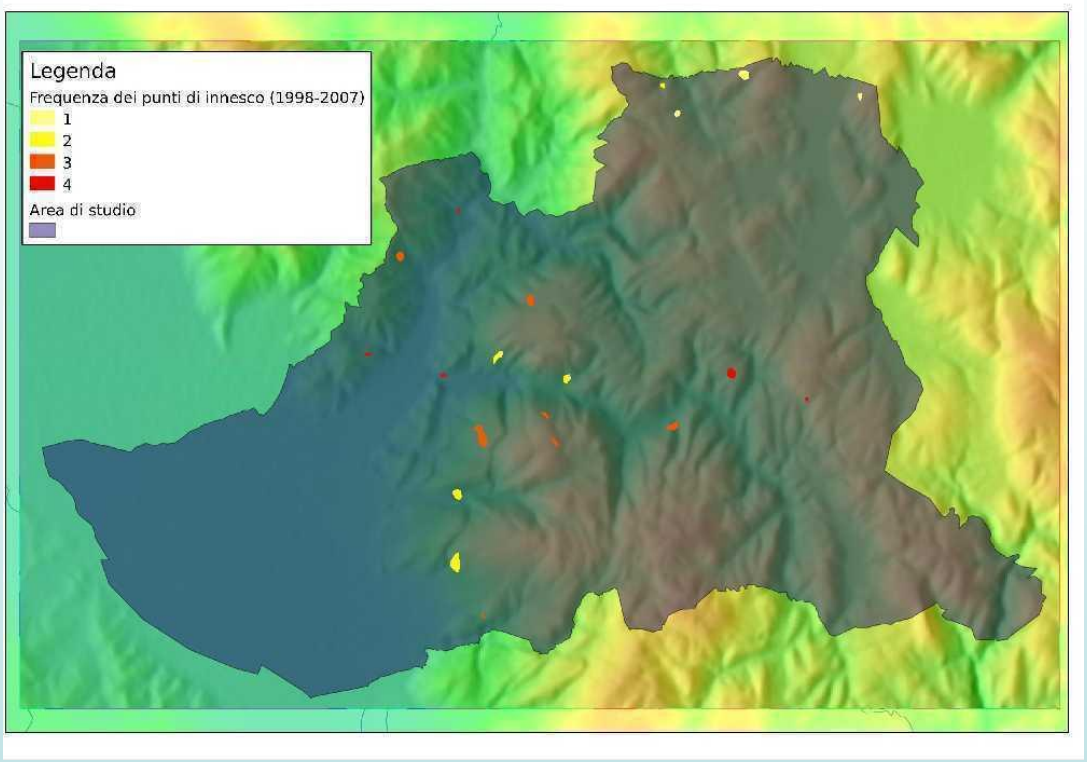




Climate parameters: summer rain (mm) and summer average temperature (*C)

Case Study

The territorial information system



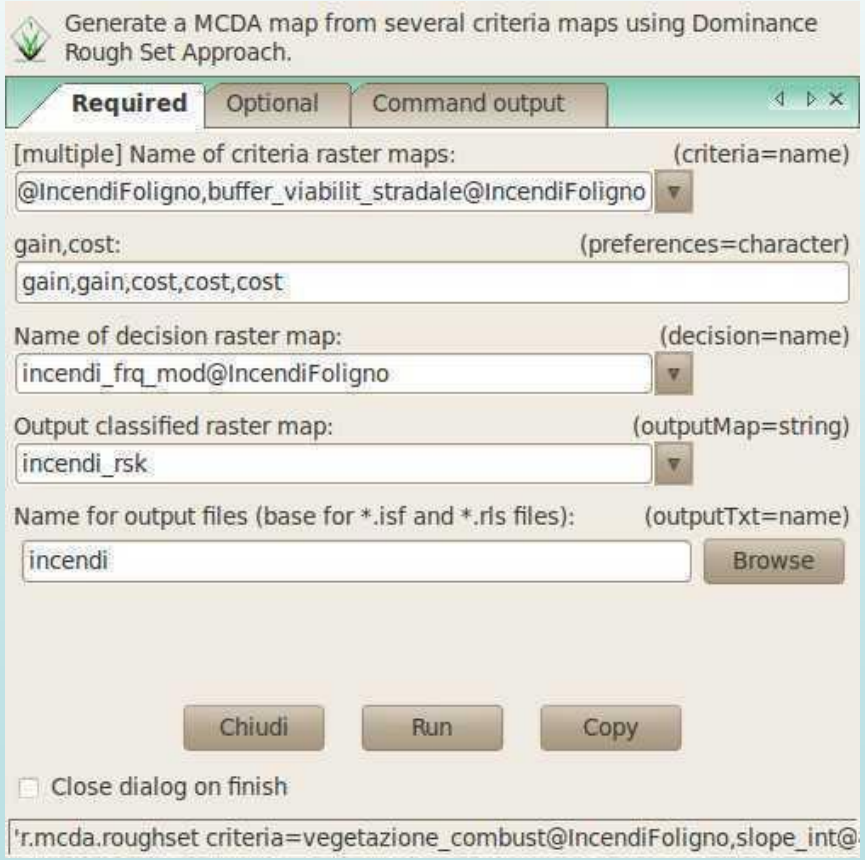
Exemplary case: areas for which the fire inception frequence is known from a 10 years observation period

Case Study

Input parameters

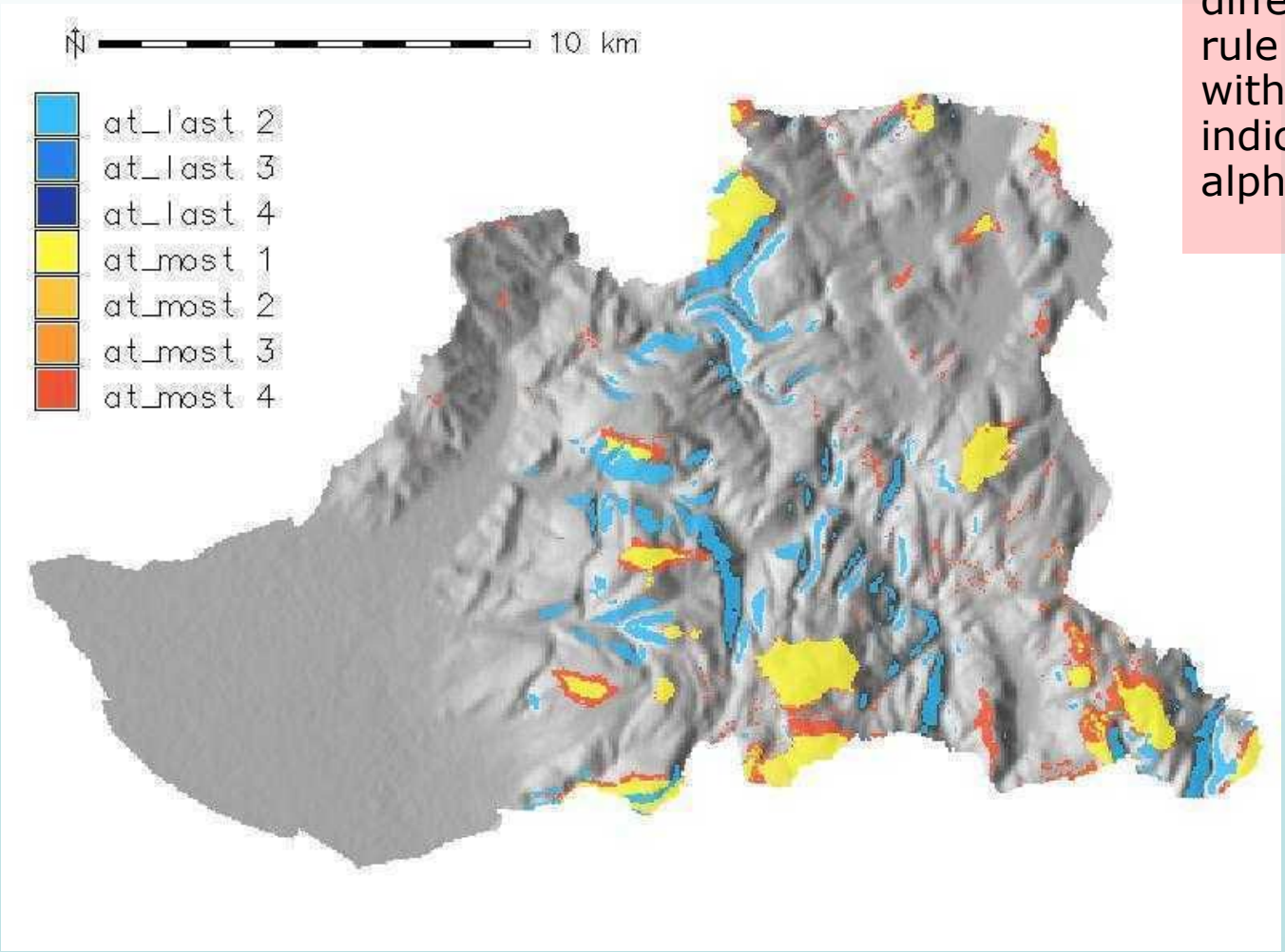
r.mcda.roughset *criteria*=name[,name,...] *preferences*=character *decision*=name *outputMap*=string *outputTxt*=name.

Criterio geografico	Preference
Vegetation (combustion classes)	cost
Slope	gain
Altitude (m s.l.m.)	cost
Forests and distance from the boundary (100 m)	cost
Distance from roads	cost
Summer rain	cost
Average summer temperature	gain



Case Study

Graphic outputs



Areas having "sure rules" have in the map different colours. Each rule is represented with a number indicating the alphanumeric id

Alphanumeric output: .rls file

The main alphanumeric output is a file with .rls extension, containing the description of the rules got from the module and used to draw the map The general syntax is: *if …. [and] …. then ...*

RULES:
[...]
4: (road_distance_ <= 0.000) & (altitude <= 834.000) & (buff_forest <= 1.000) & (slope >= 23.000)=> (class at_least , 3)
5: (slope >= 28.000)=> (class at_least , 4)
[...]
9: (altitude >= 1063.000) & (road_distance >= 335.000)=> (class at_most , 1)
10: (road_distance >= 873.000) & (vegetation >= 3.000) & (slope <= 18.000)=> (class at_most , 1)
[...]
32: (slope <= 5.000) & (altitude >= 1028.000) & (vegetation >= 3.000)=> (class at_most , 2)
35: (slope <= 6.000) & (altitude >= 1040.000) & (vegetation >= 3.000)=> (class at_most , 2)
[...]

Alphanumeric output: .rls file

How to read (e.g.: rule 4):

4: (road_distance <= 0.000) & (altitude <= 834.000) & (buff_forest <= 1.000) & (slope >= 23.000)=> (class at_last , 3)

Road_distance <= 0.000	The cell is close to a road
&	and
quota <= 834.000	Altitude is <= 834 m
&	and
buff_forest <= 1.000	Land use is forest
&	and
slope >= 23.000	Slope is >= 23 degrees
=>	then
Class at last 3	The cell is classified at least in forest fire risk class 3

Conclusion

MCDA is an useful approach in environmental and land assessment and evaluation processes.

Traditional numeric MCDA cannot manage the geographic dimension often useful to better understand and to find sound solutions for the assessment and evaluation problems. Instead, to manage the geographic dimension is the role of GIS.

The integration of the analytical capabilities of MCDA with the geographic data management of GIS allows an improvement of the analysis, assessment and evaluation of environmental, social and economic aspects.

Conclusion

The MCDA suite is available in the **GRASS addons** repository (http://trac.osgeo.org/grass/browser/grass-addons/raster/mcda) with GNU GPL license. All modules are actively being tested using several case studies.

The final product of the work consists of a "suite" that can efficiently integrate the main methods of MCDA in a full GIS GRASS environment.

- DRSA has a lot of good properties in this context:
  - Transparence
  - Use of qualitative evaluations
  - Decision analysis based on ordinal properties only
  - Traceability (it means that from score we can go back to rules and from rules to objects in the data set.
- DM gives preference information by answering easy questions, and obtains transparent feedback („glass box") in a learning oriented perspective