Susceptibility assessment of landslides: A comparison of two GIS-based methods in the region of Al Hoceima (Eastern Rif, Morocco).

A. N. EL FAHCHOUCH [1], L. AIT BRAHIM[2], O. RAJI[3]

[1,2]Mohammed V University, faculty of science, department of earth sciences, GEORISK research unit: Geological risks, remote sensing and sustainable development. Ibn Battuta street, PO Box 1014, RABAT, MOROCCO.
[3]Mohammed VI Polytechnic University, Geology & Sustainable Mining, BENGUERIR, MOROCCO.

Abstract. The evaluation of the degree of susceptibility to landslides has become a major concern in mountainous areas, it is a key component of manager policies efforts in disaster prevention, mitigate risk and manage the consequences. The region of Al Hoceima is one of most mountain regions in Morocco, and is highly exposed landslides events. For this reason, the area was selected in order to determine its susceptibility to landslides using two methods. The purpose of this study is to evaluate and to compare the results of multivariate (logical regression) and bivariate (landslide susceptibility) methods in Geographical Information System (GIS) based landslide susceptibility assessment procedures. In order to achieve this goal, the selected methods were compared by the Seed Cell Area Indexes (SCAI) and by the spatial locations of the resultant susceptibility pixels. We found that both of the methods converge in 80% of the area; however, the weighting algorithm in the bivariate technique (landslide susceptibility method) had some severe deficiencies, as the resultant hazard classes in overweighed areas did not converge with the factual landslide inventory map. The result of the multivariate technique (logical regression) was more sensitive to the different local features of the test zone and it resulted in more accurate and homogeneous susceptibility maps. This information may have direct applications in landslides susceptibility research programs and can assist decision-makers in the implementation of preventive management strategies in the most sensitive areas.
Keywords: Landslide susceptibility, Multivariate, Bivariate, Al Hoceima, Morocco.

1 Introduction

Mass movements, In particular the landslides are very destructive factors, which cause every year lots of financial losses in mountainous area [1]. They can interact in very different ways and scattered in different places [2]. They are in general a manifestation of landmasses gravity displacement destabilized because of natural conditions.

Landslide susceptibility expresses the spatial probability that a landslides, which exist or may potentially occurs in an area for different local environmental conditions [3]. The most important conditions for a landslide susceptibility zonation include geomorphology, geology, geotechnical parameters, hydrology, topography, meteorology and botany [1].

The susceptibility of an area may be determined through very diverse, their comparison has largely fueled the scientific literature [4]. GIS and remote sensing has contributed largely to the development of these methods. Among the most recent methods used, there is the bivariate analysis based on Bayes' theorem named theory

of evidence [5-7]. Actually, statistical methods seem to be more applicable and the preferred by the scientific community [8].

In Morocco, the identification, description and mapping of landslide hazard is a relatively new field of study. In the 1960's, [9] Avenard (1965) have mapped out the erosion in the Sebou basin. Since then, several studies have followed [10-15]. Most of these studies were conducted in Rif region, which is highly exposed to landslides hazards. However, the evaluation of the susceptibility to landslides in this area using probabilistic methods remains insufficient.

The objective of this research is to: (i) systematically inventory region, characterize and map the landslides in the Al Hoceima region, (ii) Build a geographic information systems (GIS) database of these hazard events and its causative factors, (iii) Assess the susceptibility to landslides through a bivariate and multivariate probability model. (iv) Compare the results of the two selected calculation methods

2 Study area

The study area (Al Hoceima region) is located in the central part of the (Rif northern Morocco), between latitudes 34°00" to 35°20" N and longitudes 4°20" to 3°50" W "figure 1". The total surface area of the selected area is 414 km2. This Rif in general and the Al Hoceima region is located in an area geologically unstable, subject to earthquakes [16] and is considered as the area most affected by different ground instabilities [15].

Geologically, the Al Hoceima area is part of the internal domain of the Rif Belt related to the Alpine Orogeny. The area is formed by a range of diverse facies and is characterised by the stacking of several structural units that extends from Paleozoic to Cenozoic. Among these units, there is the external dorsal limestone, which supports, as a tectonic klippe, the terrains of the internal dorsal limestone and those of the Palaeozoic Ghomarides nappes. These units overlay the terrains of the pre-dorsal zone (Tertiary sole) "figure 1". From a geomorphologic point of view, Al Hoceima region is dominated by rugged and mountainous terrain characterized by very steep slopes and by diverse topography. Elevations range from a few meters to 1200m [34-35]. The Al Hoceima region has a large forest area and under Mediterranean climate influences, it receives an annual rainfall of between 360 and 700 mm.
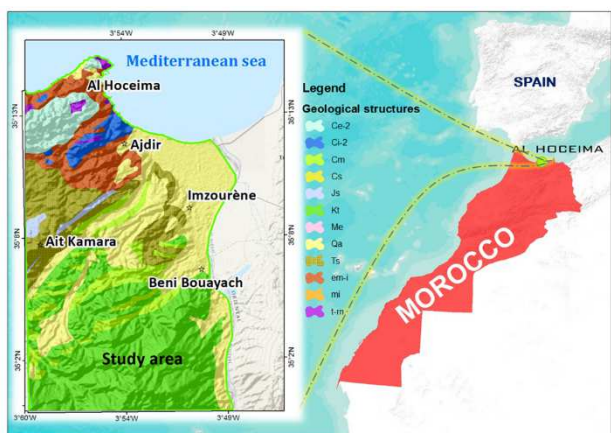
Fig.1. Location map of the study area, Ce-2: Lias White limestone; Ci-2: Marl-sandstone from Priabonian-Oligocene; Cm: Cretaceous Early Marl; Cs: Cretaceous late Marl; Js: Jurassic limestone; Kt: Schist from ketama unit; Me: Flysch siltstone; Qa: Quaternary; Ts: Marl of Tisirene unit; em-i: Ultrarif Jurassic Marl; mi: Triasic marl-clay of mezzo-rif unit; t-rn: Marl-dolomitic.

3 Data and methodological approach

This study is based on a comparison of two methods of susceptibility assessment using GIS tools. The framework, which was adopted to achieve this goal, consists of three steps "figure 2": (i) Data sets and database creation (ii) Landslide susceptibility assessment by bivariate and multivariate methods and (iii) Comparison of methods.
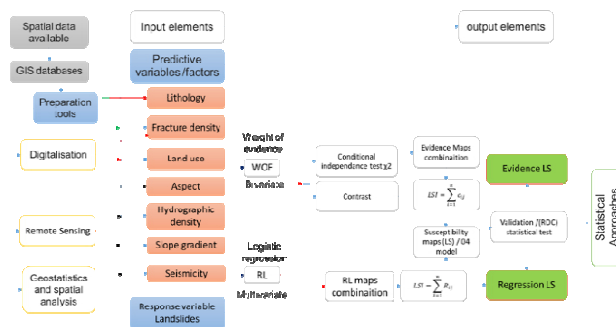
Fig.2. Methodological procedure adopted to assess landslind succptibility

3.1 Data sets and database creation

The selection of causal factors is crucial to determine the susceptibility to landslides events. In this context, we reviewed multiple physical parameters used throughout the world for the development of susceptibility mapping [17-19]. A preliminary analysis was necessary to reduce the number of parameters and refine these accordingly. The choice of variables is based on two vital factors: i) applicability at the scale of the study area; and ii) data availability. Consequently, the seven following quantifiable variables were chosen to evaluate the susceptibility of study area to landslides:

Land cover: The spatial distribution of landslides is also related to land uses/land cover. Indeed, it is recognized that bare soils favor terrain instability [20]. In our case, the land-cover map was produced by the analysis of a Landsat ETM+ image (year 2000).

Fracture density: lineation has an effective role in a landslide occurrence [21], to obtain the fracture density map, we used the results of the analysis of satellite images supplemented by data from geological maps and fields data; Which allowed us to establish a fracturing density map.

Lithology: Represents a major conditioning factor, the lithological map is based on the main lithological units obtained by the segmentation of the landscape into homogeneous macro-areas closely associated with sediment facies [22].

Slope: An important parameter in stability considerations and, more generally, slopes greater than 5°, are the most able to produce instabilities. However, it has to be taken into account that high slope angles do not always is synonym of largest landslide occurrence [23].

Aspect: is referred to as orientation from north of the terrain surface slopes. It is determined by measuring, in a clockwise direction, the angle between the north and the horizontal projection of the maximum slope. Both slope and aspect maps were prepared in ArcGIS from DEM.

Hydrographic network density: The disturbance of the slopes by the torrents is considered as a criterion of

instability [1]. The erosion of stream banks can cause the triggering of a landslides or its reactivation.

Seismicity: Using a global catalog of landslides, [24] Keefer had identified that the area where landslides are likely to occur, increases with magnitude of earthquakes and that a minimum magnitude of 4 is required to trigger a landslide.

3.2 Bivariate method

The weight of evidence method is based on a statistical bayesian bivariate approach. Actually is being used in several geology fields, and the interest for this method extended in the field of landslide hazard assessment is constantly increasing [25, 6, 7].

A detailed description of the mathematical formulation of the method is available in [26] and mathematical relationships for landslide hazard mapping is available in [27]. The method calculates the weight for each predictive factor based on the presence (positive) or absence (negative) of the training point theme units (D) within the area of each binary predictor theme (B), as indicated in [28]:

$$W_i^+ = \ln \frac{P\{B/Vd\}}{P\{B/\overline{V}d\}} \qquad (1)$$

$$W_i^- = \ln \frac{P\{\overline{B}/Vd\}}{P\{\overline{B}/\overline{V}d\}}$$
(2)

A positive weight (Wi+) indicates that the causative factor is present at the landslide location and there is a positive correlation between presence of the causative factor and landslides. A negative weight (Wi-) indicates an absence of the causative factor and shows the level of negative correlation. The difference between the two weights is known as the weight contrast, Wf (W+ - W-), and the magnitude of contrast reflects the overall spatial association between the causative factor and landslides. In weights-of-evidence model, the combination of causative factors assumes that the factors are conditionally independent of one another with respect to the landslides [26].

The calculation of weights of evidence assumes conditional independence among the causative factor input to the model. One technique to assess the conditional independence between pairs of binary causative factor is to calculate a Chi-squared statistic (X2) to assess the variation between the expected and observed occurrences on and off the patterns in the two factor.

The pairwise test between two evidential factors involves a contingency table calculation, applicable only to locations at which landslides occur. The rows of the contingency table are the classes of one evidential factor, and the columns of the contingency table are the classes of the second factor. Each cell (i,j) of the table records the

number of landslides occurring for a specific overlap of the i-th class of factor 1 and the j-th class of factor 2.

The null hypothesis of conditional independence is tested by determining if the measured chi-squared value exceeds a theoretical chi-squared value, given the level of significance. The (α) is the number of rows – 1 times the number of columns - 1. So for binary themes, the (α) =1. The level of significance for most tests is taken as 95%, equal to (1-probability) or (α) =0.05 [5, 29] Probability values (α) <0.05 indicates some conditional probability, or the failure of the conditional independence test at the 95% level. Low values of probability indicate conditional independence and the lower the value the greater the indication of conditional independence. The main problem of this test is a sensitivity in the number of landslides taken into account, the results varying by 30% [30, 31].

3.3 Multivariate (logical regression) method

As with the weight of evidence method, the logical regression has been largely used to study the susceptibility to landslides. Indeed, it allows forming a multivariate regression relation between causative factors that might affect the probability of the landslide occurrence [32]. The predicted values, which range from 0 to 1, can be defined by the following formulas [33], 32]:

p =ez/(1+ ez)   *(3)*

z= b0+ b1 x1 + b2 x2 + …..+bn xn   *(4)*

where P is the probability of landslide occurrence (landslide susceptibility index), z is the linear logistic model, b0 is the intercept of the model, n is the number of landslide–causative factors, b is the weight of the each factor, and x is the landslide conditional factors.

Logistic regressions have several advantages, namely [34]:

1. They work with quantitative and / or categorical variables

2. No analysis of factor conditional independence (X2 test) is necessary to obtain consistent results;

3. The method of calculation does not necessarily require a partition of the study area by homogeneous units

3.4 Comparison of methods

The results susceptibility maps produced from bivariate analysis and logical regression analyses. Both of them produced acceptable results will be compared depending on how their high susceptibility classes differ or converge. Therefore, the seed cell area index (SCAI) was used which is the area percent values divided with the landslide seed cell percent values. The high and very

high susceptibility classes should have very small SCAI values and low and very low susceptibility classes to have higher SCAI values.

4 Results and discussion

4.1 Landslides inventory and predictive factors mapping

At a scale of 1:50,000 to 1:30,000 through interpretation of multiple data sources (satellite image, geological and geomorphological maps) combined with field data, a landslide inventory maps was constructed "figure 3". It shows the spatial distribution of existing landslides that have moved in the past. 187 landslides of various dimensions have been distinguished, they cover an area of 30 km2 accounting for 7 % of the study area. The rotational and translational landslides represents the majority with 80% of the total number of observed landslides, followed by flows (12%) and falls (8%).
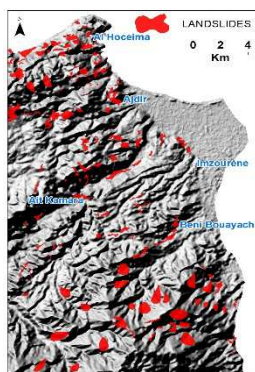

Fig.3. Inventery landslides map

Figure 4 shows the predictive variables maps for the study area: (i) the land use "figure 4a", shows a dominance of bare soils with about 57% followed by agricultural areas and urban area with about 33%. It is noted that the majority of inventoried landslides (60%) occur in areas with bare soil. (ii) the orientation of the lineaments is dominated by the NE-SW direction. The highest fracturing density is found in the NW of the study area "figure.4b". The main lithological units of the study area have been grouped into four categories "figure 4c", which according to their percentages are: 43.05% of quartzitic sandstone, 37.78% of shale, 14.05% of land Marl-schist, and finally 5.09% of the limestones. In this sense, it should be noted that most of inventoried landslides in our study area are related to the presence of clays, marls and shales. With regard to slope, the majority of the territory has a slope higher than 6% "figure 4d". The maximum values are naturally found in mountain areas where the slope usually exceeds 20%. The Aspect map shows that the most exposed areas are in the mountainous areas on the western side of the study area "figure 4e". The study area is traversed by a very dense hydrographic network but the rivers are non-permanent limited to periods of high floods "figure 4e". It has a strong seismicity at the NW but the magnitudes do not exceed 5 "figure 4g".

Generally, in case of landslides, evaluation of a prediction result is based on predictive ability of the prediction result. The best validation method is to see if the prediction result may be identical with inventory landslides.

Table 1. Shi-sequar result for each predictive factor; Probability values depend on chi-squared values (number of degrees of freedomX2) and Values α < .05 indicate some conditional dependence.

Parameters	Test	Fracture density	Seismicity	Lithology	Hydrographic network density	Aspect	Land use
Slope	X^2	109	**6**	30	43	117	19
	α	0.00	**0.99**	0.00	0.00	0.00	0.00
Fracture density	X^2		34	87	47	**12**	**12**
	α		0.00	0.00	0.00	**0.76**	**0.71**
Seismicity	X^2			25	24	**10**	**18**
	α			0.00	0.00	**0.84**	**0.35**
Lithology	X^2				**14**	18	27
	α				**0.59**	0.33	0.00
Hydrographic network density	X^2					6	15
	α					0.99	0.56
Aspect	X^2						26
	α						0.00

In trying to determine the best combination of predictive factors and landslide occurrence, the conditional independence was tested for each parameter. The table 1 shows chi-square for all possible pairs. As a result, for the degree of dependency, which exceeds the 76 per cent, it evidence a dependence between Hydrographic network density-Aspect (99%) and between Slope-Seismicity (99%) and Aspect-Seismicity (84%). Based on these results, we selected four models: a benchmark model combining all parameters and three other models by combining only independent factors "table 2".

Table 2. The variables combinations

Model	Combination
M1	Slope **(S)** + Fracture density **(Frd)** + Seismicity **(Ss)** + Lithology **(Lth)** + Hydrographic network density **(Hyd)** + Aspect **(As)** + Land use **(Lu)**
M2	Frd + Lth + Hyd + Lu + Ss
M3	Frd + Lth + Hyd + Lu + S
M4	Frd + Lth + As + Lu + S

4.2 Application of weights of evidence mode and logistic Regression

Fig.4. parameters maps

Fig.5. Wof adopted models
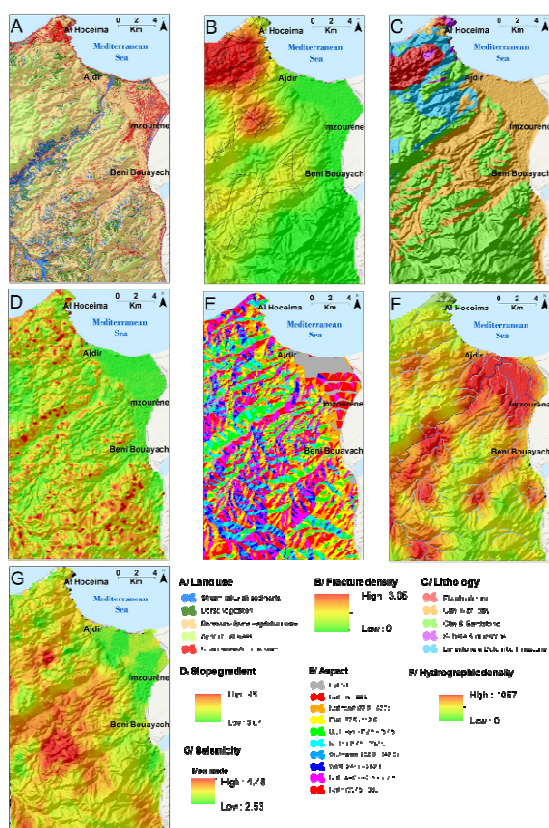
The maps in figure 5 show spatial variability of susceptibility to landslides from model. To test the success and prediction rates of models, ROC curve was used "figure 7". Each resulting map was reclassified into four susceptibility classes based on the quartile ranges.

According to the results, all selected models provide generally good performance but the map produced by the model 4 exhibited the better one with 0.7395 (73.95 %) of predictive accuracy. The north of the study area is highly susceptible to landslides, whilst the Al Hoceima plain have low susceptibility to landslides.
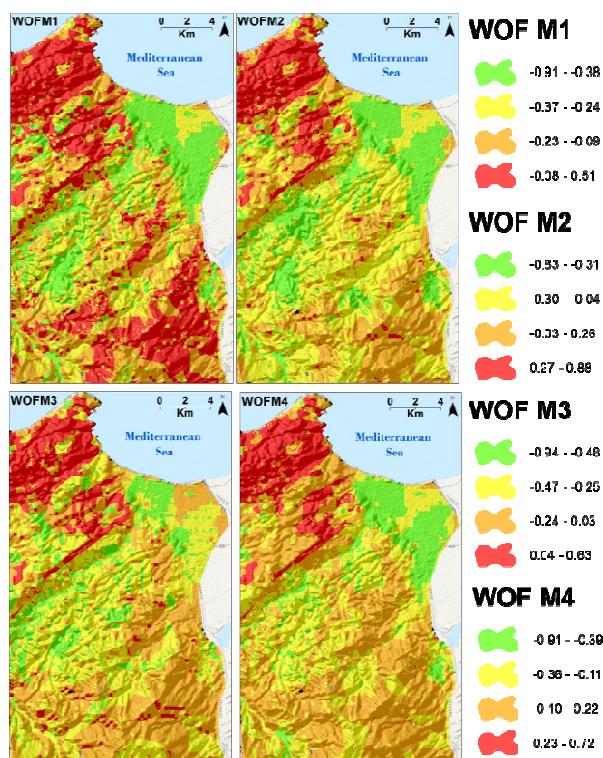
LR model was employed using the same models as for weight of evidence method. The generated Landslide susceptibility map was also reclassified into four susceptibility classes "figure 6". Compared to WOF susceptibility maps, more area seem to be susceptible to landslides and only Al Hoceima plain show the lowest susceptibility values. Unlike WOF results, models 1 and 3 posters the best prediction performance of the landslide susceptibility maps produced with respectively 0.78 and 0.75 "figure 7".

5 Conclusion

Landslides are very tragic disasters that can occurs regularly in a variety of places. During the last years, many quantitative methodologies have been developed to study the relationship between landslides and causative factors. In this paper, two methods were selected and tested in the Al Hoceima area: The weight of evidence and the logistic regression.

Seven predictive factors were identified in order to apply these statistic methods, some of them were derived from DEM and the rest from available data of the study area. This study shows that the approach taken using GIS is appropriate for study area and that all selected models present good performances. However, the landslide susceptibility map produced using the logistic regression method predicts higher percentages of landslides, especially in very high zones than the logistic regression method. Furthermore, it should be noted that the best model (best combination of predictive factor) is different from one method to the other.

Generally, all the models tested show that the northern zone of the study area and the high-altitude

zones at the south are the areas which have the high susceptibility values in comparison with the Al Hoceima plain which have the lowest values.
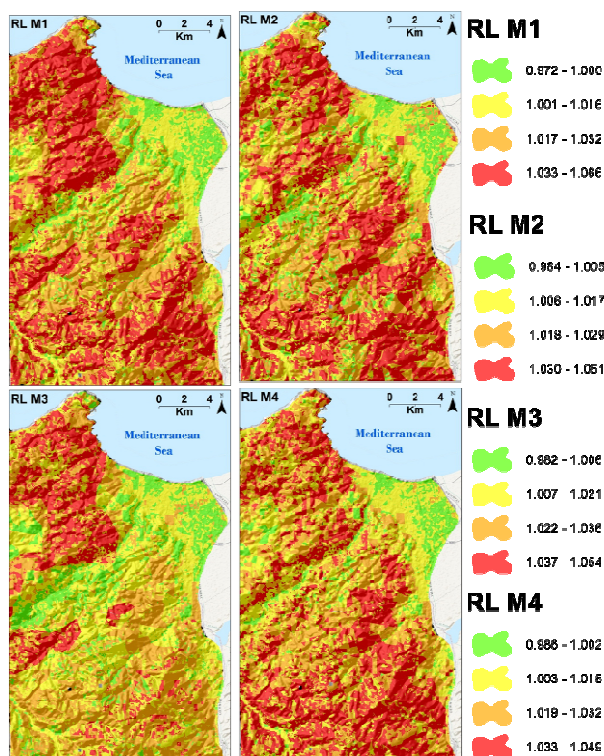

Fig.6. Regression logistic adopted models
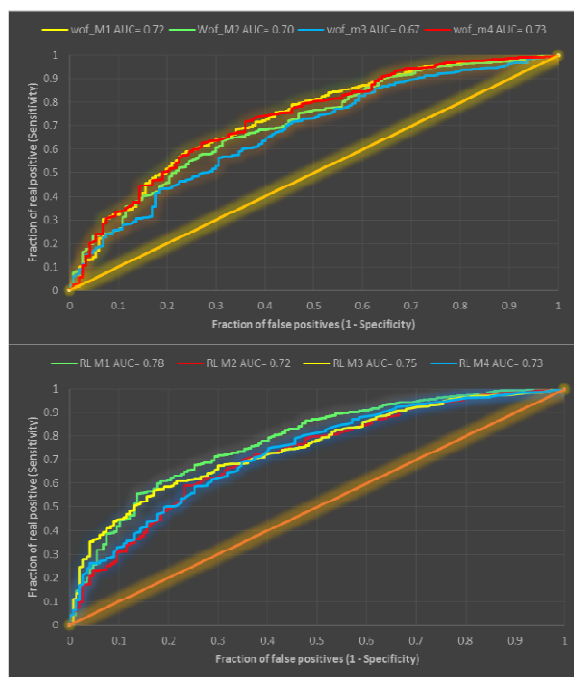

Fig.7. ROC compared graphs

References

1. Y. Thiery, J. P.Malet, S. Sterlacchini, A. Puissant, O. Maquaire, *Revue Internationale de Géomatique*, 15(2), **227-245** (2005)

2. S. Gruber, C. Huggel, R. Pike, *Developments in Soil Science*, 33, **527-550** (2009)

3. S. Lee, J. S. Won, S. W. Jeon, I. Park, M. J. Lee, *Mathematical Geosciences*, 47(5), **565-589**, (2015)

4. C.J. Van Westen, T.W.J. Van Asch, R. Soeters, *Bulletin of Engineering Geology and the Environment*, 65, **167–184**. (2006)

5. G. Bonham-Carter, *Pergamon Press, Oxford*, **398**, (1994)

6. C-J. Van Westen, G.F. *Lulie Geomorphology* 54 (1–2), **77–89**, (2003)

7. Y. Thiery, Strasbourg, **16**, (2006)

8. E. Yesilnacar, T. Topal, *Engineering Geology*, 79(3), **251-266**, (2005)

9. J.M. Avenard, INRA, **114** (1965)

10. G. Maurer, *Université de Paris*, **500**. (1968)

11. A. El Gharbaoui, *série Géologie et Géographie Physique*. n° 15. **439**, (1981)

12. Kh. Margaa, *Université de Franche-Comté*, **196**, (1994)

13. A. Fares, *Université de Franche Comté*. **177**, (1994)

14. A. Sossey, *Université Mohamed V, Faculté des Sciences*, **175**. (2005)

15. A.N. El-Fahchouch, L. Ait-Brahim, O. Raji, A. Khouakhi, *Afrique science* 11(2), **44-57**, (2015)

16. A. Talhaoui, M. Aberkan, I. Brahim, A. El Mouraouah, Pangea, 43, **3-18** (2005)

17. E. Yesilncar, T. Topal, *Engineering Geology, Vol. 79*, **251-266**(2005)

18. L. J. Wang, K. Sawada, S. Moriguchi, *Computers & Geosciences*, 57, **81-92**, (2013)

19. E. K. Sangchini, S. N. Emami, N. Tahmasebipour, H. R. Pourghasemi, S. A. Naghibi, S. A. Arami, B. Pradhan, *Arabian Journal of Geosciences*, 9(3), **1-15** (2016)

20. P. Reichenbach, C. Busca, A.C. Mondini, M. Rossi *Environmental Management*, 54(6):**1372-1384**. (2014)

21. H. Shahabi, B. B. Ahmad, S. Khezri, *Arabian Journal of Geosciences,* 6(10), **3885-3907** (2013)

22. VAN WESTEN, C. Johannes. (1993).

23. A. Donnarumma, P. Revellino, G. Grelle, F. M. Guadagno, *Landslide Science and Practice*. **425-433** (2013)

24. D. K. Keefer, *Surveys in geophysics*, 23(6), **473-510**, (2002*)*

25. S. Lee, K. Min, *Environmental Geology*, 43: **120-131**. (2002)

26. I. M. Kettles, , G. F. Bonham-Carter, *Environment, Analysis,* 2(2), **99-110** (2002)

27. R. K. Dahal, S. Hasegawa, A. Nonomura, M. Yamanaka, T. Masuda, K. Nishino, *Environmental Geology*, 54(2), **311-324** (2008)

28. G. F. Bonham-Carter, F. P. Agterberg, D. F. Wright, *Statistical applications in the earth sciences*, 89(9), **171-183** (1989)

29. M. Davis, *Neuron*, 36(3), **340-343**, (2002).

30. A. Pistocchi, L. Luzi, P. Napolitano, *Environmental geology*, **41**(7), (2002).

31. P. Dumolard, N. Dubus, L. Charleux, *Les statistiques en géographie, coll. Atouts, Belin, Paris.* (2003)

32. L. Ayalew, H. Yamagishi, *Geomorphology*, 65(1), **15-31** (2005).

33. S. Menard, *Sage University Paper Series on Quantitative Applications in Social Sciences, vol. 106. Thousand Oaks, California.* **98**. (1995)

34. Y.Thiery, J. P.Malet, S. Sterlacchini, A. Puissant, O. Maquaire, *Geomorphology,* 92(1), **38-59**, (2007).

35. L. Ait Brahim, C. Nakhcha, B. Deffontaines, H. Sami, P. Chotin (2004)- Revue Télédétection- photo interprétation. Editions ESKA, France N° 4, pp.**13-49.**

36. L. Ait Brahim, B. Tadili, C. Nakhcha, I. Mouayn, M. Ramdani, M. Limouri, (2002) - Pure and Applied Geophysics,: Volume 161, Numbers 5-6, pp. **1081-1091**