MODELLING THE POTENTIAL DISTRIBUTION OF GULLIES IN SPANISH DEHESAS

Álvaro GOMEZ GUTIERREZ, Susanne SCHNABEL, Francisco LAVADO CONTADOR

Avda. University s/n, Fac. Filosofía y letras, Dpto. A. and Territory Sciences, 10071, Cáceres, alvgo@unex.es

ABSTRACT

Gully erosion represents an important soil degradation process on Spanish rangelands. Its main negative consequences are soil loss, reduction of water content on soil, pollution of flow, flooding of channels and reservoirs increasing inundation risk, etc. Therefore is important to know where it is going to happen on landscape, to take preventive or control measures and to reduce its environmental damages and economical costs.

The main objectives of this work are: 1) modelling spatial distribution of gullying on rangelands in the southwest of Spain, 2) comparing two non-parametric schemes to construct predictive models 3) analyze the role of prevalence (different proportion of presence/absence data on the dataset) and scale for susceptibility models on geomorphology, 4) evaluate the importance of the factors implicated and 5) implementing and mapping the results with the help of a Geographical Information System (GIS).

Two methods were used to model the response of a dependent variable (gullying) from a set of independent variables: Classification And Regression Trees [CART; Breiman et al., 1984] and Multivariate Adaptive Regression Splines [MARS; Friedman, 1991]. CART can be considered as one of the most important methods of supervised learning (Berzal Galiano, 2002), while MARS is a relatively new method based on classical regression, recursive binary partitioning and brute force intelligent algorithms. Three different datasets were used; the first one in order to construct the model (training dataset) and the others to validate it (external datasets). These datasets are formed by a target variable (present or absence of gullies) and a set of independent variables. Dependent variable was obtained mapping gully positions with the help of a GPS and high resolution aerial ortophotographs. We used 32 independent variables reflecting topography, lithology, edafology, climate and land use and vegetation cover conditions of each area.

To evaluate the performance of the models we used a non-dependent threshold method: the Receiver Operating Characteristic curve [ROC; Deleo, 1993] and the Area Under the Curve (AUC). AUC represents the ability of a predictive model to discriminate between presences and absences of the target variable.

The results show a better performance of MARS for predicting gullying with an AUC of 0.98 and 0.97 for validation datasets while CART presents values of AUC of 0.96 and 0.70. With regard to the importance of factors implicated, a different outcome was obtained with each model.

Finally, we generated cartography with the results of each model which can represent an important tool for farmers or regional managers in the fight against erosion and desertification.

REFERENCES

Berzal Galiano, ART: an alternative method for decision trees, PhD Thesis, University of Granada, 2002.

Breiman, L., Friedman, J. H., Olsen, R.A. and Stone, C.J., Classification and Regression Trees, CA, Belmont, 1984.

Deleo, J.M., Receiver Operating Characteristic Laboratory (ROCLAB): Software for developing decision strategies that account for uncertainly, Proceedings of the 2nd International Symposium on Uncertainly Modelling and Analysis, College Park, Computer Society Press, 318-325, 1993.

Friedman, J.H., Multivariate Adaptive Regression Splines, Ann. Statist., 19, 1-141, 1991.