# ORIGINAL ARTICLE

# A GIS-based approach for gully erosion susceptibility modelling: a test in Sicily, Italy

Christian Conoscenti · Valerio Agnesi · Silvia Angileri · Chiara Cappadonia · Edoardo Rotigliano · Michael Märker

Received: 21 November 2011/Accepted: 17 December 2012/Published online: 12 February 2013 © Springer-Verlag Berlin Heidelberg 2013

Abstract The aim of this study is to analyze the susceptibility conditions to gully erosion phenomena in the Magazzolo River basin and to test a method that allows for driving the factors selection. The study area is one of the largest (225 km<sup>2</sup>) watershed of southern Sicily and it is mostly characterized by gentle slopes carved into clayey and evaporitic sediments, except for the northern sector where carbonatic rocks give rise to steep slopes. In order to obtain a quantitative evaluation of gully erosion susceptibility, statistical relationships between the spatial distributions of gullies affecting the area and a set of twelve environmental variables were analyzed. Stereoscopic analysis of aerial photographs dated 2000, and field surveys carried out in 2006, allowed us to map about a thousand landforms produced by linear water erosion processes, classifiable as ephemeral and permanent gullies. The linear density of the gullies, computed on each of the factors classes, was assumed as the function expressing the susceptibility level of the latter. A 40-m digital elevation model (DEM) prepared from 1:10,000-scale topographic maps was used to compute the values of nine topographic attributes (primary: slope, aspect, plan curvature, profile curvature, general curvature, tangential curvature; secondary: stream power index; topographic wetness index;

### M. Märker

LS-USLE factor); from available thematic maps and field checks three other physical attributes (lithology, soil texture, land use) were derived. For each of these variables, a 40-m grid layer was generated, reclassifying the topographic variables according to their standard deviation values. In order to evaluate the controlling role of the selected predictive variables, one-variable susceptibility models, based on the spatial relationships between each single factor and gullies, were produced and submitted to a validation procedure. The latter was carried out by evaluating the predictive performance of models trained on one half of the landform archive and tested on the other. Large differences of accuracy were verified by computing geometric indexes of the validation curves (prediction and success rate curves; ROC curves) drawn for each onevariable model; in particular, soil texture, general curvature and aspect demonstrated a weak or a null influence on the spatial distribution of gullies within the studied area, while, on the contrary, tangential curvature, stream power index and plan curvature showed high predictive skills. Hence, predictive models were produced on a multi-variable basis, by variously combining the one-variable models. The validation of the multi-variables models, which generally indicated quite satisfactory results, were used as a sensitivity analysis tool to evaluate differences in the prediction results produced by changing the set of combined physical attributes. The sensitivity analysis pointed out that by increasing the number of combined environmental variables, an improvement of the susceptibility assessment is produced; this is true with the exception of adding to the multi-variables models a variable, as slope aspect, not correlated to the target variable. The addition of this attribute produces effects on the validation curves that are not distinguishable from noise and, as a consequence, the slope aspect was excluded from the final multi-variables

C. Conoscenti  $(\boxtimes) \cdot V$ . Agnesi · S. Angileri · C. Cappadonia · E. Rotigliano

Dipartimento di Scienze della Terra e del Mare (DiSTeM), Università degli Studi di Palermo, Via Archirafi 22, 90123 Palermo, Italy e-mail: christian.conoscenti@unipa.it

Dipartimento di Scienze delle Produzioni Vegetali, del Suolo e dell'Ambiente Agroforestale (DIPSA), Università degli Studi di Firenze, Piazzale delle Cascine 14, 50144 Florence, Italy

model used to draw the gully erosion susceptibility map of the Magazzolo River basin. In conclusion, the research showed that the validation of one-variable models can be used as a tool for selecting factors to be combined to prepare the best performing multi-variables gully erosion susceptibility model.

**Keywords** Gully erosion · Susceptibility models · Validation test · GIS · Sicily

# Introduction

Erosion by water constitutes a relevant and worldwide problem that causes severe land degradation phenomena in semi-humid to arid areas (Bou Kheir et al. 2007; Buttafuoco et al. 2012). Climate changes leading to long dry periods followed by extreme rainstorm events, as well as the recent intensification of human activities and agricultural practices, make the Mediterranean region particularly susceptible to water erosion.

In the last decades a number of studies have been carried out with the aim of developing and applying models for the assessment of soil loss rates and water erosion risk. Methods allowing us to quantify the volumes of eroded soil are typically grouped into empirical and physically-based models; the popular universal soil loss equation (USLE; Wischmeier and Smith 1965) and the WEPP model (water erosion prediction project; Nearing et al. 1989) are good example of the first and the second group, respectively.

Most of the erosion models focus on the quantification of soil eroded by sheet and rill erosion at hillslope and test plot scale. Nevertheless, recent researches have shown the importance of gully erosion, revealing that the contribution of this phenomenon increases with the extension of the basin, ranging from 10 % up to 94 % of the total sediment yield caused by water erosion; moreover, it is through gullies that a large part of the soil eroded from slopes reaches the river networks (Poesen et al. 2003). These considerations suggest that ephemeral and permanent gully erosion phenomena need to be addressed for the evaluation of soil loss rates, especially at watershed scale; at present, there are few models capable of predicting soil loss produced by gully erosion, among which are Chemicals, Runoff and Erosion from Agricultural Management Systems (CREAMS; Knisel 1980), Ephemeral Gully Erosion Model (EGEM; Capra et al. 2005; Merkel et al. 1988; Woodward 1999), the method proposed by Sidorchuk (1999) and the channel erosion routine of the WEPP watershed model.

There are also methods that allow an investigator to assess soil erosion risk or produce erosion susceptibility maps, at various scales, by defining statistical relationships between a set of physical attributes and the spatial distribution of the landforms related to water erosion processes. The statistical methods, in general, do not provide volumes of soil loss due to water erosion, but allow for the production of susceptibility maps showing the expected spatial probability, defined in relative terms, of landform occurrence. These techniques often group the predicted landforms according to the type of process, or associate different typologies of erosion features in terrain units characterized by homogeneous erosion dynamics. The latter approach is used at the basis of the concept of the erosion response units (ERU), proposed by Märker et al. (1999); such a method allows an investigator is able to generate water erosion susceptibility maps, on a watershed scale, in which the various levels express aggregation of erosion features characterized by similar intensity (in terms of eroded volumes). Two categories of erosion landforms, which group features produced by diffused and linear processes, are used in the method proposed and tested in a Sicilian river basin by Conoscenti et al. (2008a), based on multivariate statistical approach applied to spatial unique conditions units (UCU; Carrara et al. 1995). A similar, but implemented by means of a bivariate classification procedure has been used in southern Italy by Conforti et al. (2011) and by Magliulo (2012). Recently, gully erosion susceptibility maps have been prepared also by exploiting advanced data mining techniques (Bou Kheir et al. 2007; Gutiérrez et al. 2009a, b; Lucà et al. 2011; Märker et al. 2011).

In the present paper, we investigate susceptibility conditions to gully erosion in a watershed of southern Sicily, the Magazzolo River basin, by following a statistical approach aimed at mathematically defining spatial relationships between gully occurrence and variability of twelve physical attributes. The main goal of the research is to draw up a gully erosion susceptibility map able to well reproduce the distribution of ephemeral and permanent gullies and to indicate those portions of the basin, at present not hosting erosion landforms, where new gullies are more likely to occur in the future. The best performing multi-variables susceptibility model, chosen after testing various combinations of the twelve environmental variables, is used to prepare the susceptibility map. To accomplish this task, one of the key points to be faced, is to define a procedure to estimate the contribution in the predictive performance of multivariate models, which is provided by a single predictor variable. Another important objective of this research is to verify whether the controlling role of each of the selected physical attributes can be evaluated by assessing the predictive skill of susceptibility models derived from each single variable. To this aim, the accuracy of the latter is evaluated and compared to that of multi-variables models prepared by differently combining the independent variables, so to assess how including a variable affects or not the performance of the model.

# Materials and methods

# Study area

The Magazzolo River flows in the southern portion of Sicily, draining a basin of  $225 \text{ km}^2$ ; the main fluvial axis has a general NE-SW orientation and runs for about 36 km from the southern slopes of the Sicani Mounts to the Sicilian Channel (Fig. 1). The area experiences a Mediterranean climate, with mild wet winter periods and dry, hot summer times; mean annual rainfall is about 700 mm and concentrates in a few rainy days of the winter period, while summers are almost dry.

The basin is located in the mildly folded foredeep foreland sector of the Sicilian collisional complex (Catalano et al. 1993). The outcropping rocks are: limestones (lower liassic-upper trias), dolomitic limestones (lowermiddle Jurassic), pelagic marly limestones and marls (upper Cretaceous-Eocene) pertaining to the Sicanian basinal succession; marls and limestones (Oligocene) of the Trapanese Platform; conglomerates, clayey sandstones and marls (upper Tortonian-lower Messinian) of the Terravecchia formation; carbonates, gypsum rocks and marls of the Messinian evaporitic succession (upper Messinian); pelagic marly calcilutites (lower Pliocene) of the Trubi formation; present day beach, fluvial and slope deposits.

The watershed extends from NE to SW with an elongated shape that narrows in the middle and in the coastal sector. The northern portion, where carbonate rocks largely outcrop, is the highest sector of the basin and is characterized by steep slopes and scarps, affected by debris and rock falls. A hilly area can be recognized in the centralnorthern portion of the basin; this sector, characterized by gentle slopes carved into evaporitic and clayey sediments of the evaporitic succession, is affected by landslides and severe water erosion phenomena. The coastal zone is dominated by a wide alluvial plain and by the outcropping of marls, calcarenites and clays.

## Gully erosion landforms and variables

Remote and field surveys carried out in the Magazzolo River basin allowed us to recognize about a thousand ephemeral and permanent gullies (Agnesi et al. 2007). In particular, by means of stereoscopic analysis of aerial photographs (scale 1:10,000) taken in 2000, a map representing the spatial distribution of gullies was produced; field surveys were carried out in 2006 in order to verify the reliability of the remote analysis and to improve the gully map in critical zones. This map, that was turned into a



Fig. 1 Magazzolo River basin location and hillshaded DEM

geo-referenced GIS vector layer (Fig. 2), shows all sites of the investigated basin where gully erosion processes are able to produce landforms, with the exception of mapping errors and of those ephemeral gullies that were filled by farmers at the time of the aerial and field surveys. These obliterated gullies could be a source of error for the susceptibility models, especially considering that, in the studied area, the use of this agricultural practice is quite frequent even if new incisions tend to reappear soon where ephemeral channels were filled.

The development of gully erosion susceptibility models requires the selection of environmental variables able to reproduce the geographical variability of the main factors potentially controlling the phenomenon; for this research, the selection was based on geomorphological knowledge of gully erosion phenomena and on the availability, for the area, of environmental data related to erosion processes. The occurrence of this phenomenon depends on climate, topography, lithology, soil characteristics and land use (Poesen et al. 2003; Gutiérrez et al. 2009a); thus, susceptibility to gully erosion is a function of erodibility of outcropping materials and erosivity of runoff waters (Conoscenti et al. 2008a). To the aim of reproducing proneness to erosion of rocks/soils and erosive power of runoff, three erodibility and nine erosivity variables were selected: bedrock lithology (LTL), soil texture (TXT) and land use (USE), as erodibility variables; slope angle (SLO) and aspect (ASP), plan curvature (PLC), profile curvature (PRC), general curvature (CRV), tangential curvature (TNC), stream power index (SPI), topographic wetness index (TWI) and length-slope USLE factor (LSF), as erosivity variables. All the physical attributes were spatially defined as 40-m square grids, by combining information derived from available thematic maps and field surveys (erodibility variables) and by processing a digital elevation model (erosivity variables).

The bedrock lithology grid (Fig. 3a) was prepared in accordance with the expected erodibility of outcropping materials. The limits of nine LTL classes were derived from existing geological maps and field surveys; the most diffused lithological types are clays (37.3 %), evaporitic rocks (24.5 %) and marls (13.5 %).

Soil information was obtained from the soil map of Sicily (Fierotti 1988); among the five classes of soil texture defined in the studied area (Fig. 3b), the fine-medium (61.2 %) and the medium (22.2 %) classes are the most frequent. A regional soil use map (ARTA Sicilia 1994a) was used to derive the grid of land use (Fig. 3c), which shows the spatial pattern of sixteen Corine legend classes of land cover.

The frequency distributions of the classes of the three erodibility variables are plotted as bar diagrams in Fig. 4.

According to Bou Kheir et al. (2007), i.e., the occurrence of gullies is primary controlled by the topographic



Fig. 2 Gully erosion landforms map and examples of ephemeral and permanent gullies affecting the area



Fig. 3 Spatial variability of the erodibility variables: LTL (a), TEX (b), USE (c)

features, six primary and three secondary topographic attributes (Wilson and Gallant 2000) were selected as predictive variables to generate gully erosion susceptibility maps. Since climatic variables can be considered as being fairly homogeneous in the studied area (Ferro et al. 1991), the spatial variability of runoff erosive power is here assumed to be expressed simply by means of topographic attributes.

The nine topographic attributes were extracted by processing a 40-m grid digital elevation model (DEM), using freeware ArcView GIS 3.2 (ESRI 1999) spatial analysis tools (Sinmap, Demat and Topocrop). The DEM was obtained by digitizing contour lines and elevation points from 1:10,000 scale topographic maps, with 10-m contour interval (ARTA Sicilia 1994b), and by interpolating elevation data by means of the Topo to Raster tool of ArcGIS 9.3 (ESRI 2008). Although a better resolution would have been possible, we chose a pixel of 40-m in order to avoid that the topographic variables were altered by the presence of the incisions, the latter of a width largely smaller than the size of a cell; in fact, we felt it was more useful to assign high susceptibility to the general topography of the slopes where gullies occur, rather than identify as highly susceptible values of topographic attributes calculated on cells with a size comparable to that of the hosted erosional landforms. Moreover, this selection agrees with the optimum cell size computed by following the theory that information content grows with entropy of data (Shannon and Weaver 1949). According to Sharma et al. (2011), the spatial variability of elevation, measured by calculating the entropy of a DEM, is low both for oversampled and for undersampled DEMs; in fact, redundant elevation values or loss of micro relief information could be responsible for low entropy values. In order to calculate the optimum cell size, DEMs of four different resolution (10, 20, 30, 40-m) were interpolated; for each of them, entropy (H) was computed using the following equation:

$$H = -\sum_{i=1}^{n} \left( P_i \times \log_2 P_i \right)$$

where  $P_i$  is the probability of a cell being classified as class *i* and *n* is the number of classes. Since entropy of a grid is influenced by its number of pixels (*N*), the entropy values were normalized by dividing them with 2ln (*N*) (Sharma et al. 2011). According to the normalized entropy values, showed in Table 1, the relative information content of the



**Fig. 4** Relative frequency distributions of the erodibility variables classes (*grey bars*) and linear density of gullies computed on each class (*empty bars*): LTL (**a**), TEX (**b**), USE (**c**). *Bottom X* axis: relative frequency (%); top X axis: gully linear density (km/km<sup>2</sup>)

interpolated DEMs increases when increasing the size of the cells, with a variation that gradually reduces from 20 to 40-m pixels. These results support the selection of the DEM with a pixel 40-m, being the latter adequately representative of the topographic heterogeneity of the studied area.

The primary topographic attributes—slope, aspect and plan curvature—are directly derived by investigating, using the D8 algorithm (O'Callaghan and Mark 1984), the relationships among cells, within a neighbourhood in the DEM. The slope angle is the maximum first derivative of the altitude, whose compass direction represents the slope aspect. Plan and profile curvature are the second derivative of elevation measured orthogonal and parallel to the aspect direction, respectively; the difference between the latter gives the value of the general curvature while tangential curvature is calculated along the line orthogonal to the line of steepest gradient. The second topographic attributes used in the present research are derived from slope angle

Cell size (m)	Number of cells	Entropy	Normalized entropy	Change	Change (%)
10	2,207,360	10.028	0.343		
20	551,823	10.018	0.379	0.036	10.4 %
30	245,277	10.016	0.404	0.025	6.5 %
40	137,961	10.010	0.423	0.019	4.8 %

Table 1 Entropy and normalized entropy of DEMs at different resolution

and specific catchment area, the latter corresponding to the upslope area per unit width of contour lines (Wilson and Gallant 2000). In particular, the stream power index is calculated as SPI =  $\ln[A_s \tan(\text{SLO})]$ , where  $A_s$  is the specific catchment area; the topographic wetness index is computed as TWI =  $\ln[A_s/\tan(\text{SLO})]$ , while the length-slope USLE factor as LSF =  $(((A_s/22.13)^{\wedge}0.4) \times 1.4((\sin(\text{SLO})/0.0896)^{\wedge}1.3))$ . All the topographic attributes, except for slope aspect, were reclassified in quarter standard deviation intervals, in order to give weight to their relative variability rather than to their absolute values. Figure 5 shows the frequency distributions of the classes of the nine topographic attributes.

The morphodynamic significance of the selected topographic attributes, in terms of runoff erosive power, can be resumed by the following considerations: SLO can control the overland and subsurface flow velocity and runoff rate; ASP influences solar insolation and vegetation distribution on slopes and, also, could indirectly express (proxy role) the influence of the structural setting; the four computed curvature attributes measure the convergence or divergence of runoff water; moreover, if we assume that discharge is proportional to  $A_s$  and that overland flow velocity is proportional to SLO, then SPI increases as a function of the runoff erosive power, TWI is proportional to soil saturation and LSF can be considered a measure of sediment transport capacity.

# Erosion model and validation procedure

In this research, gully erosion susceptibility is assessed by adopting the conditional analysis (Davis 1973; Carrara and Guzzetti 1995) according to which the susceptibility level of a homogenous domain corresponds to the density of water erosion landform computed. Homogenous units are here spatially defined by grouping cells pertaining to a single variable class, in univariate analysis, or by identifying sets of cells characterized by unique combinations of a set variable classes, in the multivariable analysis. The homogenous domains, prepared by applying this procedure, correspond to the concept of UCU, which is widely adopted in landslide hazard studies (Carrara et al. 1995; Clerici et al. 2002; Conoscenti et al. 2008b) and more recently in water erosion assessment (Conoscenti et al. 2008a).

The linear density of ephemeral and permanent gullies is so here exploited as the gully erosion susceptibility function; the density values, computed for each class of the selected conditioning factors, were considered as expressing the susceptibility levels and were used to prepare gully erosion susceptibility models both from single variables and combinations of them. According to the Bayes' Theorem, the conditional probability for a mapping unit to host a gully, under the condition that its physical attributes would be those represented by the value assumed by the independent variable or variables, can be expressed by adapting from Davis (1973), as

$$P(\text{gully}|\text{UCU}^*) = \frac{P(\text{UCU}^*|\text{gully}) \times P(\text{gully})}{P(\text{UCU}^*)}$$

The probabilities can be computed in terms of count of cells, so that

$$P(\text{gully}|\text{UCU}^*) = \frac{\begin{pmatrix} \text{UCU}_{\text{gully}}^* \\ \text{UCU}_{\text{gully}} \end{pmatrix} \times \begin{pmatrix} \text{UCU}_{\text{gully}} \\ \text{UCU}_{\text{ALL}} \end{pmatrix}}{\begin{pmatrix} \text{UCU}^* \\ \text{UCU}_{\text{ALL}} \end{pmatrix}} = \begin{pmatrix} \text{UCU}_{\text{gully}}^* \\ \text{UCU}^* \end{pmatrix}$$
$$= \delta_{\text{UCU}^*}.$$

The density function  $(\delta_{UCU^*})$  is, for the cells having a UCU\* value, the ratio between eroded (gully) and total (ALL) counts of cells. GRID layers for single or multiple variables and gullies are spatially intersected, so that for each of the UCUs, the area hosting gullies is obtained. The susceptibility value for each UCUs is finally obtained by dividing the extension of the intersection and the total area.

Twelve one-variable models (OVM) were derived directly by computing the gully density values for each of the factor classes. The OVMs were then submitted to a validation procedure in order to test and quantify their predictive skill; the results of the validation were assumed as a quantitative assessment of the spatial correlation between the environmental variables and the distribution of gullies within the studied area. To the aim of verifying the relationships, in terms of predictive performance, between individual variables and combinations of them, the same validation procedure was applied to multi-variable models (MVM); the comparison of the validation results indicate how including a single variable affects the performance of the MVMs. A number of MVM were prepared by differently combining the physical attribute grids; the susceptibility level of each individual combination was then defined by computing the arithmetic mean of the density values of the combined factors classes and by reclassifying the average values into ten intervals according to an equal area criterion.

Both the one-variable and the multi-variables susceptibility models were submitted to a validation procedure based on a random partition of the erosion landforms in a training and a test subset (Chung and Fabbri 2003); training and test gullies were singled out by using a geostatistical tool of ArcGIS 9.3 (ESRI 2008) that allowed us to randomly split the gully database into two numerically and spatially balanced subsets. In accordance with the adopted validation strategy, the susceptibility models were derived



◄ Fig. 5 Relative frequency distributions of the erosivity variables classes (*grey bars*) and linear density of gullies computed on each class (*empty bars*): SLO (a), ASP (b), PLC (c), PRC (d), CRV (e), TNC (f), SPI (g), TWI (h) and LSF (i). *Bottom X* axis: relative frequency (%); *top X* axis: gully linear density (km/km<sup>2</sup>)

from the training gullies and were compared with the spatial distributions of both the subsets of landforms, in order to draw prediction and success rate curves (Chung and Fabbri 2003; Conoscenti et al. 2008a, b). These are cumulative curves that plot the fraction of landforms (y axis) against the fraction of the study area, measured within the susceptibility classes, the latter arranged in decreasing order along the x axis. Prediction rate curves, which were drawn using the lengths of test gullies intersecting the susceptibility levels, quantify the prediction skill of the models, while success rate curves assess the models fit, since they are derived from the same landforms used to train the models. The more accurate the model is, the more the test landforms concentrate on the highest susceptibility levels; hence, a good performance of the model produces prediction curves with high steepness in the first part and far from the diagonal trend, the latter representing a model not correlated with the target pattern. Also, a good model fit is testified by a prediction rate curve very close to the success rate curve.

In order to quantitatively assess the prediction skill and the fit of the models, three geometric indexes of the curves were computed: ARPA, SHIFT and EFR. ARPA is the area between the validation curves and the diagonal of the graph, while SHIFT is the area between prediction and success rate curves (Rotigliano et al. 2011; Costanzo et al. 2012; Rotigliano et al. 2012). Since the diagonal trend attests for a not-effective prediction, a high performance produces high values of ARPA; a good fit of the model is testified by low SHIFT results. EFR stands for the effectiveness ratio (Chung and Fabbri 2003; Guzzetti et al. 2006), i.e., the ratio between the fraction of landforms intersected by each susceptibility class and the proportion of the latter in the study area; an effective classification should produce susceptibility levels with EFR values distant from 1 (i.e., same fraction of landforms and area for each class), at least 1.5 for more susceptible classes and at most 0.5 for less susceptible classes, according to Guzzetti et al. (2006). By drawing a theoretical validation curve respecting these threshold values, Rotigliano et al. (2012) indicate 0.12 as the lower limit of ARPA for an effective susceptibility model.

A more detailed discrimination of the curves was achieved by calculating two other shape indexes: ARPA20 and EFR20; these correspond to the values of ARPA and EFR computed for the 20 % more susceptible portion of the basin and express the concentration of gullies in this area. In accordance with the threshold values proposed by Guzzetti et al. (2006) for a reliable prediction, and reproducing the same curve used by Rotigliano et al. (2012), EFR20 and ARPA20 should be at least equal to 1.5 and 0.01, respectively.

To further investigate the reliability of the one-variable and multi-variable gully erosion susceptibility models, a second procedure was adopted. This was performed by comparing the spatial pattern of the susceptibility levels with the presence/absence of test gullies, verified within each cell. To this aim, the shapefile of the test gullies was converted into a 40-m grid layer, where each cell indicates the presence or absence of gullies; hence, a spatial analysis tool of ArcView GIS 3.2 was used to count the number of positive cases within each UCU. The frequency of cells hosting gullies was compared with binary classifications of the UCUs into units predicted as positive (susceptible) and units predicted as negative (not-susceptible), obtained by means of cut-off values that split the range of susceptibility levels into two parts. For each of the classifications derived from the susceptibility models, contingency tables counting the number of true positives, true negative, false positive and false negative, were produced over the whole range of the possible cut-offs; these data, obtained by means of the software TANAGRA (Rakotomalala 2005), an open-source data mining tool, were used to compute true positive (TP) and false positive (FP) rates for all the gully erosion susceptibility models. TP and FP rates correspond to "sensitivity" and 1-"specificity", where sensitivity refers to the fraction of cells containing test gullies that were correctly classified as susceptible and specificity is the fraction of cells free of gullies that were correctly classified as not-susceptible. The predictive performance of the models was assessed by drawing their receiver operating characteristic (ROC) curves (Goodenough et al. 1974; Lasko et al. 2005) that plot TP rates against FP rates. The ROC curves have been recently adopted to evaluate model performances in landslide susceptibility researches (Begueria 2006; Van Den Eeckhaut et al. 2009; Frattini et al. 2010) and gully erosion susceptibility mapping (Gutiérrez et al. 2009a, b). The shape of the ROC curves, which can be quantitatively described by the area under a ROC curve (AUC; Hanley and McNeil 1982), shows how much a model correctly reproduce the occurrence of positive and negative cases: the larger the area, the best the predictive skill of the model. The AUC was exploited as a further index to assess the predictive performance of the one-variable and multi-variable susceptibility models; a susceptibility model that respects the limits of EFR suggested by Guzzetti et al. (2006) and adopted by Rotigliano et al. (2012), produce a ROC curve with AUC at least equal to 0.63.



◄ Fig. 6 Prediction and success rate curves derived from the validation of the one-variable models: LIT (a), TEX (b), USE (c), SLO (d), ASP (e), PLC (f), PRC (g), CRV (h), TNC (i), SPI (j), TWI (k) and LSF (l). X axis: fraction of the studied area; Y axis: fraction of gully length

#### **Results and discussion**

#### One-variable susceptibility models

The values of linear density of training gullies computed for each class of the physical attributes are plotted in Figs. 4 and 5, together with the frequency distributions of the classes. The density values indicate which classes, and, as a consequence, which portions of the basin, are more frequently associated with the presence of ephemeral and permanent gullies.

The graphs of the erodibility variables (Fig. 4a–c) show that the most susceptible classes are: evaporitic rocks and clays, among the bedrock lithology types; fine-medium and medium soil texture classes; not-irrigated arable lands and annual and permanent crops among land cover Corine classes. As regards the primary topographic attributes (Fig. 5a–f), gullies concentrate on slopes inclined between 5 and 15°, characterized by concave plan and tangential curvature and convex profile curvature, while density values of ASP classes seem poorly differentiated. The diagrams of gully density values computed with respect to secondary topographic attributes (Fig. 5g–i) indicate that ephemeral and permanent gullies are to be more expected where cells have medium–high SPI and medium TWI values, while LSF intervals do not show very differentiated density values.

Prediction and success rate curves were produced for each of the twelve OVM (Fig. 6); the shape of the curves depends on the correlation degree between the geographic variability of the susceptibility levels, which were derived from the training gullies subset, and the spatial occurrence of the test gullies. A visual comparison among the validation curves and between the curves and the random trend (diagonal), provides a qualitative and relative assessment of the models' effectiveness: TNC, SPI and PLC models show the best predictive skill since they produce curves that are clearly above the others; ASP curves are not very far away from the diagonal trend attesting for a very weak correlation between the variable classes and the gullies spatial distribution; the models of the other attributes draw validation curves which stay in a middle zone, indicating medium to weak (TEX) predictive skills.

In order to quantitatively discriminate between the onevariable validation curves, ARPA, ARPA20, EFR20 and SHIFT indexes were calculated (Table 2). With the exception of profile curvature, soil texture and slope aspect, all the factors are characterized by ARPA values that are above the limit of effectiveness (0.12): ARPA20 and EFR20 confirm, for the most susceptible zones, an unsatisfactory predictive skill for TEX and ASP, since their values are quite below the thresholds (0.01 and 1.5). On the other hand, the geometric indexes of the validation curves attest for good performance and fit for TNC, SPI and PLC models that correctly predict about 40 % of the landforms within the 15 % most susceptible portion of the basin. General curvature and length-slope USLE factor produce a more powerful prediction respect to the susceptibility models derived from SLO, USE, LTL and TWI; the latter provide susceptibility classifications characterized by similar predictive skill, even if all the geometric indexes show that TWI validation curves have the weakest correlation with gullies. The SHIFT index attests for moderate differences of ARPA between prediction and success rate curves (not more than 10 % of their average value of ARPA) for all the environmental variables with the exception of slope aspect whose validation curves have a SHIFT value comparable with their ARPA values.

The comparison between susceptibility levels of each cell of the basin, which were derived from the density of training gullies, and the presence/absence of gully test, allowed us to draw a ROC curve for each of the OVM (Fig. 7). By computing AUC values, we are able to quantitatively characterize the shape of the curves and to evaluate, in relative terms, the correlation degree between the geographical variability of the variables and the spatial distribution of ephemeral and permanent gullies on the studied area.

ROC curves and values of AUC (Table 2) confirm the assessment of the models predictive skill provided by the

 Table 2 Geometric indexes of prediction (pred) and success (succ)

 rate curves, and AUC values of the ROC curves, computed for all the

 one-variable models

OVM	ARPA		ARPA	ARPA20		SHIFT	AUC
	Pred	Succ	Pred	Succ			
LTL	0.124	0.129	0.012	0.011	1.580	0.005	0.623
TEX	0.095	0.097	0.006	0.006	1.302	0.002	0.588
USE	0.126	0.137	0.013	0.014	1.643	0.012	0.624
SLO	0.135	0.125	0.012	0.010	1.582	-0.010	0.630
ASP	0.010	0.036	0.001	0.003	1.029	0.026	0.512
PLC	0.205	0.197	0.031	0.033	2.359	-0.009	0.678
PRC	0.105	0.103	0.009	0.010	1.457	-0.002	0.592
CRV	0.183	0.174	0.024	0.024	2.069	-0.009	0.658
TNC	0.214	0.205	0.032	0.034	2.435	-0.009	0.691
SPI	0.206	0.193	0.038	0.032	2.522	-0.013	0.676
TWI	0.123	0.125	0.010	0.011	1.520	0.001	0.612
LSF	0.162	0.149	0.015	0.016	1.840	-0.013	0.640
Limit	0.120	0.120	0.010	0.010	1.500	-	0.630



Fig. 7 ROC curves derived from the validation of the one-variable models

quantitative analysis of the validation curves. TNC, SPI and PLC are the independent variables showing the best predictive results, as attested by values of AUC equal to 0.691, 0.676 and 0.678, respectively; also models of CRV (AUC = 0.658), LSF (AUC = 0.640) and SLO (AUC = 0.630) produce ROC curves respecting the threshold of an effective susceptibility model (AUC  $\geq$  0.63). OVMs of USE, LTL and TWI draw ROC curves just below the limit, while PRC (AUC = 0.592) and TEX (AUC = 0.588) predictive performances are quite distant; finally, as attested from prediction and success rate curves, slope aspect does not demonstrate to be correlated with the spatial distribution of gullies in the basin, since the ROC curve derived from ASP model is very similar to the random trend.

The validation results of the OVM, which we assume as expressing the spatial correlation between physical attributes and gullies, are in general in accordance with the geomorphological meaning of the selected environmental variables and with the quality and resolution of the input data. Stream power index, in addition to tangential and plan curvatures constitute a measure of runoff convergence and, as a consequence, of linear erosive power of flowing water; hence, these environmental attributes were expected to be strongly correlated with the spatial distribution of ephemeral and permanent gullies. Also the good predictive performances of LSF and SLO models agree with the geomorphological significance of these variables, since they are related to flow velocity and volume, and, therefore, should express sediment transport capacity. The weak spatial correlation between TWI susceptibility classification and test gullies, could also be expected since the wetness index is inversely related to slope steepness and its spatial variability could be rather more related to the position of gullies head-cuts; on the other hand, profile curvature, which is supposed to control runoff erosivity, does not contribute to explain the spatial occurrence of gullies in the studied area. Moreover, ASP model validation results point out that factors connected to slope aspect (i.e., solar insolation, vegetation distribution, structural setting), which were defined on a single cell scale, do not control gully erosion phenomena in the studied basin. In general, the better performances of models derived from erosivity attributes compared to those derived from erodibility variables demonstrate a strong control of topographic setting on gully erosion phenomena, even though the poor predictive skill of LTL, USE and TEX could also be explained considering the low resolution of the input layers. In particular, the five classes of TEX, which were derived from a regional map at a scale of 1:250,000, probably do not represent adequately the spatial heterogeneity of soil texture; furthermore, the discriminant ability of this attribute is negatively affected by the high spatial frequency of the most susceptible texture class (fine-medium), covering more than 60 % of the entire basin. Finally,

Table 3 Environmental variables combined to prepare the tested multi-variables models (MVM)

MVM	Environmental variables											
	LTL	TEX	USE	SLO	ASP	PLC	PRC	CRV	TNC	SPI	TWI	LSF
HIGH						Х			Х	Х		
MOD				Х				Х				Х
LOW	Х		Х								Х	
HM				Х		Х		Х	Х	Х		Х
HML	Х		Х	Х		Х		Х	Х	Х	Х	Х
HMLV	Х	Х	Х	Х		Х	Х	Х	Х	Х	Х	Х
ALL	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х	Х
VLN		Х			Х		Х					

considering the controlling role of soil texture on erosion phenomena, a higher resolution of this attribute on the investigated basin should be addressed in future studies to improve the quality of susceptibility models.

# Multi-variables susceptibility models

A great number of multi-variables susceptibility models can be obtained by differently selecting and combining the layers of the environmental attributes. Several different combinations of the latter were exploited to prepare MVMs that were, successively, tested in order to verify their accuracy. In this paper we present the validation results of some of the examined MVMs that are of particular interest in light of the combined variables and the achieved results; this choice was aimed at highlighting how the predictive performance of the MVM is affected by the addition of the physical attributes, which were classified according to the efficiency of their OVMs in predicting the spatial distribution of gullies. In light of the validation results of the



Fig. 8 Prediction and success rate curves derived from the validation of the multi-variable models: HIGH (a), MOD (b), LOW (c), HM (d), HML (e), HMLV (f), ALL (g) and VLN (h). X axis: fraction of the studied area; Y axis: fraction of gully length

 Table 4 Geometric indexes of prediction (pred) and success (succ)

 rate curves, and AUC values of the ROC curves, computed for all the multi-variables models

MPM	ARPA		ARPA20		ER20	SHIFT	AUC
	Pred	Succ	Pred	Succ			
HIGH	0.225	0.214	0.041	0.038	2.617	-0.011	0.703
MOD	0.226	0.214	0.041	0.037	2.605	-0.012	0.703
LOW	0.184	0.196	0.027	0.028	2.120	0.012	0.679
HM	0.240	0.230	0.046	0.043	2.733	-0.010	0.711
HML	0.269	0.267	0.054	0.052	2.981	-0.002	0.739
HMLV	0.271	0.270	0.056	0.053	3.023	-0.001	0.742
ALL	0.270	0.271	0.056	0.054	3.011	0.001	0.742
VLN	0.094	0.111	0.006	0.008	1.350	0.017	0.642
Limit	0.120	0.120	0.010	0.010	1.500	-	0.630



Fig. 9 ROC curves derived from the validation of the multi-variable models

OVM, the independent variables can be grouped in accordance with their controlling role as follows:

- High: TNC, SPI, PLC
- Moderate: CRV, LSF, SLO
- Low: USE, LTL, TWI
- Very low: TEX, PRC
- None: ASP

Prediction and success rate curves of eight MVM (Table 3) are plotted in Fig. 8; the values of the geometric indexes (ARPA, ARPA20, EFR20 and SHIFT) of the curves are reported in Table 4, together with the AUC values computed on the ROC curves (Fig. 9) that were prepared for each of the examined MVMs.

The validation results attest for predictive skills that are above the threshold of an effective susceptibility model for all the MVMs, except the model VLN, whose poor predictive performance was expected since it derives from the combination of the two not-effective OVMs. The MVM HIGH, MOD and LOW show better performances if compared to the single source OVMs; even LOW, which is derived from low performing OVMs, has a predictive capability higher than the best one-variable model. The validation results of the others MVM indicate that the predictive skill increases together with the number of the combined attributes, provided that any of the latter show a slight correlation with the target variable; they could in fact be ranked, starting from the best performing, as HMLV (AUC = 0.742),HML (AUC = 0.739)and HM (AUC = 0.711). ALL performance (AUC = 0.742) is not distinguishable from the one of HMLV, indicating that adding a not correlated variable (i.e., ASP) to the multivariable model does not affect (negatively or positively) the predictive skill of the MVM. Finally, the very low correlation of profile curvature and soil texture with the target variable is also confirmed by the imperceptible change of predictive performance produced when they are aggregated to the multi-variable model.

Gully erosion susceptibility map

The validation procedures applied to both the one-variable and MVM allowed us to identify the one more accurately fitting the spatial distribution of actual gullies in the investigated area. The HMLV model, trained using all the mapped gullies, was so used to draw the gully erosion susceptibility map of the Magazzolo River basin (Fig. 10). The latter depicts how the proneness to gully erosion phenomena spatially changes on the investigated area. Two large susceptible zones are highlighted on the maps: a northern and a central one, where high susceptibility conditions are assigned also to slopes where linear erosion landforms were not recognized on the field and remote surveys; the central susceptible zone, where the two main tributaries converge and give rise to the Magazzolo river, is characterized by almost undifferentiated high susceptibility conditions.

# **Concluding remarks**

The spatial occurrence of erosional landforms and its relationships with the variability of twelve physical attributes were explored in the studied area in order to assess gully erosion susceptibility. Linear density of gullies computed on homogenous terrain units was selected as susceptibility function to prepare twelve one-variable and a





number of multi-variable gully erosion susceptibility models. The validation results of these models allowed us to achieve the main objectives of the research and provided useful information on spatial relationships between gullies and environmental attributes, when taken as an individual predictive variable or combined to prepare MVM.

The gully density values demonstrated a general congruence between the spatial distribution of landforms and what was expected for linear water erosion phenomena. This congruence is more evident for topographic attributes that showed higher discriminant ability respect to the erodibility variables; gullies are in fact more associated with cells where topography favours linear erosive power and sediment transport capacity of water flow. Moreover, the accuracy statistics of the OVM generally agree with the meaning of the attributes and with the quality and resolution of the input layers; the better predictive skills of erosivity variables in respect to erodibility variables, in fact, reflect a strong control of topography on gully erosion phenomena and a more detailed resolution of DEM-derived attributes.

The validation results of MVM indicate that: (i) predictive skills of MVMs are in accordance with OVMs performances of the combined variables (i.e., MVMs made up of "good" predictive variables are better than those made up of "bad" predictive variables); (ii) the accuracy of the susceptibility models generally grows up when we increase the number of combined variables with an acceptable predictive power. The best performing MVM showed good predictive skills, demonstrating that a statistical approach based on the computation of linear density of gullies within attributes classes could be used to prepare reliable gully erosion susceptibility models and maps; as stated from the validation of the HMLV model, the final susceptibility map (Fig. 10) well represents the spatial distribution of the recognized gullies and individuates those portions of the territory, at present not hosting erosion landforms, more likely to be affected by gullies in the future.

In conclusion, the research carried out allowed us to prepare a reliable gully erosion susceptibility map for the Magazzolo River basin starting from data already available and derived from field and remote surveys; moreover, the general congruence of predictive skill between OVM and MVM indicates that preparing and validating OVMs could be used as a tool to relatively quantify the controlling role of the predictive variables and select those that have to be combined to prepare the best performing MVMs. In fact, procedures of forward selection of variables have been applied for logistic regression and discriminant analysis models (e.g., Carrara et al. 2008; Van Den Eeckhaut et al. 2009). In the present paper a similar approach is proposed for models based on conditional analysis, which is applicable to unique condition units method. Finally, the accuracy and simplicity of the method encourage further applications to other areas where the availability of detailed

thematic maps and DEMs allow for deriving good-quality layers of the predictive variables.

Acknowledgments The research herein described was funded by the Italian Ministry of Education, University and Research, in the framework of two projects: the COFIN 2004 and the COFIN 2007, nationally coordinated by Prof. Giuliano Rodolfi and Dr. Michael Märker, respectively; Prof. Valerio Agnesi was the local coordinator for both the editions of the project. Prof. V. Agnesi has participated to the final interpretation of the data; the other authors commonly shared all the parts of the research. Clare Hampton has linguistically edited the final version of this text. The authors thank the anonymous referees for their critical review and useful comments and suggestions.

#### References

- Agnesi V, Cappadonia C, Conoscenti C, Di Maggio C, Rotigliano E (2007) Gully erosion susceptibility assessment: a case study in the Magazzolo River basin, Sicily. Italy 12th Belgium–France– Italy–Romania geomorphological meeting "Climate change and related landscapes", 26th–29th September 2007, programme and abstracts, Savona, Italy, pp 31–32
- Begueria S (2006) Validation and evaluation of predictive models in hazard assessment and risk management. Nat Hazards 37(3): 315–329
- Bou Kheir R, Wilson J, Deng Y (2007) Use of terrain variables for mapping gully erosion susceptibility in Lebanon. Earth Surf Proc Land 32:1770–1782
- Buttafuoco G, Conforti M, Aucelli PPC, Robustelli G, Scarciglia F (2012) Assessing spatial uncertainty in mapping soil erodibility factor using geostatistical stochastic simulation. Environ Earth Sci 66:1111–1125
- Capra A, Mazzara LM, Scicolone B (2005) Application of the EGEM model to predict ephemeral gully erosion in Sicily, Italy. Catena 59(2):133–146
- Carrara A, Guzzetti F (1995) Geographical information systems in assessing natural hazards. Kluwer Academic Publishers, Dordrecht
- Carrara A, Cardinali M, Guzzetti F (1995) GIS technology in mapping landslide hazard. In: Carrara A, Guzzetti F (eds) Geographical information systems in assessing natural hazards. Kluwer Academic Publisher, Dordrecht, pp 135–175
- Carrara A, Crosta G, Frattini P (2008) Comparing models of debrisflow susceptibility in the alpine environment. Geomorphology 94:353–378
- Catalano R, Di Stefano P, Nigro F, Vitale FP (1993) Sicily mainland and its offshore: a structural comparison. In: Max MD, Colantoni P (eds) Geological development of the Sicilian– Tunisian platform. UNESCO Report in Marine Science 58, pp 19–24
- Chung CF, Fabbri AG (2003) Validation of spatial prediction models for landslide hazard mapping. Nat Hazards 30:451–472
- Clerici A, Perego S, Tellini C, Vescovi P (2002) A procedure for landslide susceptibility zonation by the conditional analysis method. Geomorphology 48:349–364
- Conforti M, Aucelli PPC, Robustelli G, Scarciglia F (2011) Geomorphology and GIS analysis for mapping gully erosion susceptibility in the Turbolo stream catchment (Northern Calabria, Italy). Nat Hazards 56(3):881–898
- Conoscenti C, Di Maggio C, Rotigliano E (2008a) Soil erosion susceptibility assessment and validation using a geostatistical multivariate approach: a test in Southern Sicily. Nat Hazards 46:287–305. doi:101007/s11069-007-9188-0

- Conoscenti C, Di Maggio C, Rotigliano E (2008b) GIS analysis to assess landslide susceptibility in a fluvial basin of NW Sicily (Italy). Geomorphology 94:325–339
- Costanzo D, Cappadonia C, Conoscenti C, Rotigliano E (2012) Exporting a Google Earth<sup>TM</sup> aided earth-flow susceptibility model: a test in central Sicily. Nat Hazards 61:103–114
- Davis JC (1973) Statistics and data analysis in Geology. Wiley, New York
- ESRI (1999) ArcView 3.2 user manuals. Environmental Systems Research Institute Inc, Redlands California
- ESRI (2008) ArcGIS Version 9.3. ESRI, Redlands, USA
- Ferro V, Giordano G, Iovino M (1991) Isoerosivity and erosion risk map for Sicily. J Hydrol Sci 36(6):549–564
- Fierotti G (1988) Carta dei Suoli della Sicilia. Istituto di Agronomia, Università di Palermo e Regione Sicilia, Assessorato Territorio ed Ambiente, Palermo
- Frattini P, Crosta G, Carrara A (2010) Techniques for evaluating the performance of landslide susceptibility models. Eng Geol 111(1–4):62–72
- Goodenough DJ, Rossmann K, Lusted LB (1974) Radiographic applications of receiver operating characteristic (ROC) analysis. Radiology 110:89–95
- Gutiérrez ÁG, Schnabel S, Felicísimo ÁM (2009a) Modelling the occurrence of gullies in rangelands of southwest Spain. Earth Surf Proc Land 34(14):1894–1902
- Gutiérrez ÁG, Schnabel S, Lavado Contador JF (2009b) Using and comparing two nonparametric methods (CART and MARS) to model the potential distribution of gullies. Ecol Model 220(24):3630–3637
- Guzzetti F, Reichenbach P, Ardizzone F, Cardinali M, Galli M (2006) Estimating the quality of landslide susceptibility models. Geomorphology 81:166–184
- Hanley JA, McNeil BJ (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143(1):29–36
- Knisel WG (1980) CREAMS: a field scale model for chemicals, runoff, and erosion from agricultural management systems. US department of agriculture, conservation report 26
- Lasko TA, Bhagwat JG, Zou KH, Ohno-Machado L (2005) The use of receiver operating characteristic curves in biomedical informatics. J Biomed Inform 38(5):404–415
- Lucà F, Conforti M, Robustelli G (2011) Comparison of GIS-based gullying susceptibility mapping using bivariate and multivariate statistics: northern Calabria, south Italy. Geomorphology 134:297–308
- Magliulo P (2012) Assessing the susceptibility to water-induced soil erosion using a geomorphological, bivariate statistics-based approach. Environ Earth Sci, doi:10.1007/s12665-012-1634-y
- Märker M, Flügel WA, Rodolfi G (1999) Das Konzept der "Erosions Response Units" (ERU) und seine Anwendung am Beispiel des semi-ariden Mkomazi-Einzugsgebietes in der Provinz Kwazulu/ Natal, Südafrika. In: Tübinger Geowissenschaftliche Studien, Reihe D: Geoökologie und Quartaerforschung Angewandte Studien zu Massenverlagerungen, Tübingen
- Märker M, Pelacani S, Schroder B (2011) A functional entity approach to predict soil erosion processes in a small Plio-Pleistocene Mediterranean catchment in northern Chianti, Italy. Geomorphology 125(4):530–540
- Merkel WH, Woodward DE, Clarke CD (1988) Ephemeral gully erosion model (EGEM). In: Agricultural, forest, and Rangeland hydrology. American Society of Agricultural Engineers Publication 07-88:315–323
- Nearing MA, Foster GR, Lane LJ, Finkner SC (1989) A processbased soil erosion model for USDA—water erosion prediction project. Technology of the American Society of Agricultural Engineering 32:1587–1593

- O'Callaghan JF, Mark DM (1984) The extraction of drainage network from digital elevation data. Comput Vis Graph Image Process 28:323–344
- Poesen J, Nachtergaele J, Verstraeten G, Valentin C (2003) Gully erosion and environmental change: importance and research needs. Catena 50(2–4):91–133
- Rakotomalala R (2005) "TANAGRA: un logiciel gratuit pour l'enseignement et la recherche". In: Proceedings of the EGC'2005, RNTI-E-3, vol 2, pp 697–702
- Rotigliano R, Agnesi V, Cappadonia C, Conoscenti C (2011) The role of the diagnostic landforms in the assessement of landslide susceptibility models: a test in the sicilian chain. Nat Hazards 58:981–999
- Rotigliano E, Cappadonia C, Conoscenti C, Costanzo D, Agnesi V (2012) Slope units based landslide susceptibility model: using validation tests to select controlling factors. Nat Hazards 61:143–153
- Shannon CE, Weaver W (1949) The mathematical theory of information. Urbana University of Illinois Press, Champaign
- Sharma A, Tiwari KN, Bhadoria PBS (2011) Determining the optimum cell size of digital elevation model for hydrologic application. J Earth Syst Sci 120:573–582

- ARTA Sicilia (1994a) Carta dell'Uso del Suolo (scala 1:250000). Assessorato Territorio ed Ambiente della Regione Sicilia, Palermo
- Sicilia ARTA (1994b) Sezioni: 619160, 620090, 620100, 620130, 620140, 620150, 628040, 628080, 628110, 628120, 628150, 628160, 629010, 629020, 629050, 629090. Assessorato Territorio ed Ambiente della Regione Sicilia, Palermo
- Sidorchuk A (1999) Dynamic and static models of gully erosion. Catena 37:401–414
- Van Den Eeckhaut M, Reichenbach P, Guzzetti F, Rossi M, Poesen J (2009) Combined landslide inventory and susceptibility assessment based on different mapping units: an example from the Flemish Ardennes, Belgium. Nat Hazards Earth Syst Sci 9(2):507–521
- Wilson JP, Gallant JC (2000) Digital terrain analysis. In: Wilson JP, Gallant JC (eds) Terrain analysis: principles and applications. Wiley, New York, pp 1–27
- Wischmeier WH, Smith DD (1965) Predicting rainfall erosion losses from cropland east of the rocky mountains. Handbook No 282. US department of agriculture, Washington
- Woodward DE (1999) Method to predict cropland ephemeral gully erosion. Catena 37:393–399