Full title: ASSESSMENT OF SOIL FACTORS CONTROLLING EPHEMERAL GULLY EROSION ON AGRICULTURAL FIELDS

Short title: SOIL FACTORS CONTROLLING EPHEMERAL GULLY EROSION

P. Ollobarren [1, 2, 3, *], M.A. Campo-Bescós [1, 2, 3], R. Giménez [1, 2, 3], J. Casalí [1, 2, 3]

[1] Department of Projects and Rural Engineering, Public University of Navarre (UPNA), Campus de Arrosadía s/n, 31006 Pamplona, Navarre, Spain.

[2] THERRAE research group (Remote Sensing, Hydrology, Erosion, Irrigation and Structural Analysis) UPNA.

[3] IS-FOOD Institute (Innovation & Sustainable Development in Food Chain) UPNA.

[*] Corresponding author email: paul.ollobarren@unavarra.es (P. Ollobarren)

Abstract

The soil factor is crucial in controlling and properly modelling the initiation and development of ephemeral gullies (EGs). Usually, EG initiation has been related to various soil properties (i.e. sealing, critical shear stress, moisture, texture, etc.); meanwhile, the total growth of each EG (erosion rate) has been linked with proper soil erodibility. But, despite the studies to determine the influence of soil erodibility on (ephemeral) gully erosion, a universal approach is still lacking. This is due to the complex relationship and interactions between soil properties and the erosive process. A feasible soil characterization of EG erosion prediction at large scale should be based on simple, quick and inexpensive tests to perform. The objective

of this study was to identify and assess the soil properties –easily and quickly to determine– which best reflect soil erodibility on EG erosion. Forty-nine different physical-chemical soil properties that may participate in establishing soil erodibility were determined on agricultural soils affected by the formation of EGs in Spain and Italy. Experiments were conducted in the laboratory and in the field (in the vicinity of the erosion paths). Because of its importance in controlling EG erosion, 5 variables related to antecedent moisture prior to the event that generated the gullies and 2 properties related to landscape topography were obtained for each situation. The most relevant variables were detected using multivariate analysis. The results defined 13 key variables: water content before the initiation of EGs, organic matter content, cation exchange capacity, relative sealing index, 2 granulometric and organic matter indices, seal permeability, aggregates stability (3 index), crust penetration resistance, shear strength and an erodibility index obtained from the jet test erosion apparatus. The latter is proposed as a useful technique to evaluate and predict soil loss caused by EG erosion.

Keywords**:** ephemeral gully erosion, physico-chemical soil properties, soil erodibility, multivariate statistical analysis, jet erosion test.

Introduction

Ephemeral gullies (EGs) are concentrated channels that form mainly in agricultural thalwegs when vegetation cover is minimal (Bennett et al., 2000; Casalí et al., 1999) and the accumulation, intensity or duration of rainfall is sufficient to generate a rate of runoff which exceeds the soil resistance to

detachment (Dong et al., 2015; Foster, 1986). In fact, rainfall velocity and its erosive energy are mostly controlled by landscape shape (Daggupati et al., 2014). Therefore, the occurrence of an EG will mainly depend on topographical attributes, such as upstream drainage area or terrain slope (for example, Casalí et al., 1999; Desmet et al., 1999; Nachtergaele et al., 2001; Svoray et al., 2012; Thorne and Zevenbergen, 1984; Vandekerckhove et al., 1998; Vandaele et al., 1996). The erosion models simulating the appearance and subsequent growth of EGs are thus usually based (only) on geomorphological parameters (e.g. AnnAGNPS model, Bingner et al. 2015).

However, if we take into account that EGs are typical of agricultural fields and that the latter frequently have a barely marked relief, the soil factor would also be an important conditioning element in the erosion process in these cases (e.g. Bryan, 2000; Bryan, 2004; Knapen and Poesen, 2010; Li et al., 2004; Nachtergaele and Poesen, 2002; Valentin et al., 2005).

There are numerous studies that estimate soil vulnerability to concentrated flow erosion through normally empirical techniques and procedures, given the complexity of the erosion process. It should be noted that each of those studies usually addresses only a reduced number of the, notwithstanding, many properties of the soil involved in the erosion process. These properties could be grouped –following our criteria and only for their presentation– as in the following: (i) topsoil texture (e.g. Sheridan et al., 2000; Lentz et al., 1993), (ii) topsoil stoniness (e.g. Poesen et al., 1999; Rieke-Zapp et al., 2007), (iii) aggregate stability (e.g. Chaplot et al., 2013; Geng et al., 2015), (iv) resistance to penetration (e.g. Bouma and Imeson, 2000; Verachtert et al., 2013), (v) resistance to shear stress (e.g. Léonard and Richard, 2004; Knapen et al., 2007), (vi) susceptibility

77 to surface sealing and crusting (e.g. Martínez-Casasnovas et al., 2003; Valentin
78 et al., 2005), and (vii) physico-chemical properties (e.g. Rienks et al., 2000; Van
79 Zijl et al., 2014).
80 However, the empirical nature of these measuring techniques, together with
81 the limited number of soil properties analyzed simultaneously, would largely
82 explain the fact that current knowledge about the role of the soil during
83 concentrated flow erosion processes –particularly for EGs– is still limited. On the
84 other hand, evaluation of large-scale soil erodibility, e.g. catchment scale, would
85 only be feasible through simple, rapid and economical determinations of soil
86 properties (Le Bissonnais et al., 2005)
87 This study aimed at identifying and assessing these soil properties that are
88 easily and quickly determined and that best reflect soil vulnerability to EG erosion
89 in arable lands. The results obtained in this study are expected to introduce
90 changes into current erosion models, with the ultimate goal of improving
91 (ephemeral gully) erosion simulation.
92 The study was conducted on diverse agricultural soils of Navarre, León (Spain)
93 and Sicily (Italy) affected by EG erosion. Experiments were performed in situ – in
94 microplots and with rainfall simulators– and in the laboratory. A total of 56
95 variables were evaluated, mostly edaphic but also some topographic and rainfall
96 ones. Data were analyzed using multivariate statistics approaches.
97
98 **Material and methods**
99 *Description of the study area*
100 A total of 20 agricultural soils affected by EG erosion were assessed. These soils
101 were located in 3 large study areas: (i) León (NW Spain), (ii) Navarre (N Spain),

and (iii) center of Sicily (South of Italy). In these areas, several soil losses caused by EG dynamics have been reported (e.g. Casalí et al., 1999; Casalí et al., 2008; Capra et al., 2009; Capra and LaSpada, 2015) (Figure 1, Table 1). The dominant crop in the 20 soils is winter cereal (e.g. wheat, rye), so that soil management is similar in all studied areas. Namely, field sowing is done between September and October, after preparing the seeding bed with moldboard plough and chisel, while harvesting takes place in June. All soils presented a medium-fine granulometry texture (Table 1). Also, the studied EGs were formed in areas under a typically Mediterranean climate (Table 1). Thus, the mean annual rainfall range is approximately between 450 and 1310 mm, and is concentrated (ca. 75%) in the period comprised from October to May (Table 1). The 20 EGs selected were developed in different time periods during the years 2012 to 2014, on landscapes with a slope of approximately between 3 and 25% (Table 1) and under different rainfall events (Table 2).

Determination of soil losses due to ephemeral gullies

For each EG, a digital elevation model (DEM) of 1x1 m was created after mapping both erosive flow path and their drainage area, using a total station (Leica TPS1200). The drainage area was carefully surveyed every meter. Meanwhile, a variable number of cross sections were delimited across the EG reach (Table 3) by measuring points (between 5 and 10) in the cross-sectional profiles, depending on their complexity following Castillo et al. (2012). In order to adjust any possible pit and spike, the original point cloud data were analyzed with the Leica Geo Office software (Leica Geosystems, 2006). Then, a DEM was built with a 1 m cell size, using the ArcView 3.2 software (E.S.R.I., 2000). Finally, DEM was

126 corrected by means of the filling sink function included in the Hydro Tools 1.0
127 extension for ArcView 3.2 (Schäuble, 2003)

128 Drainage area, length and volume of each EG were determined (Table 3) from
129 adjusted DEM information. The volume was obtained by, first, dividing up the EG
130 channel into homogeneous reaches –normally of between 1 to 5 m in length–
131 whose cross-sectional area was assumed to be equal to the average of the cross-
132 sections delimiting that reach (Casalí et al., 2006; De Santisteban et al., 2006).
133 Then, the volume of each reach was defined as the product of its cross-sectional
134 area and its length. Finally, the sum of these volumes defined the total volume
135 eroded by the EG (V_T, Table 3).

136 The erosion rate for each EG (variable TSL, Table 3) was determined applying
137 equation 1 (Casalí et al., 1999). However, these erosion rates yielded by the
138 studied EG were produced under rainfalls with different characteristics (Table 2).
139 This fact was due both to the different geographical location of the study areas
140 (Figure 1, Table1) and to the fact that experimental data were obtained in 3
141 different years. Several studies have remarked the importance of different rainfall
142 erosivity parameters on the final value of the TSL recorded (e.g. Archibold et al.,
143 2003; Capra et al., 2009; Han et al., 2017; Hoober et al., 2017; Poesen et al.,
144 2003; Valentin et al 2005), among others. In order to make the erosion rates for
145 each EG comparable, soil loss was quantified by normalizing the variable TSL
146 through equation 2 (Yoshimura et al., 2015).

147

148 $$TSL = \left(\frac{V_T \cdot PH_1}{A}\right) \cdot 10 \quad (1)$$

149 $$TSL_n = \frac{TSL}{R_{TOT}} \quad (2)$$

where TSL is the total soil loss per surface area per year (t ha^{-1} yr^{-1}), V_T is the total volume of the eroded soil (m^3), PH_1 is the bulk density of the soil (kg m^{-3}), A is the total EG drainage area at the mouth of the gully (m^2), TSLs is the total normalized soil loss (t MJ^{-1} mm^{-1} h yr^{-1}), and R_{TOT} is the sum of the R factor of all the events identified as being erosive (i.e., volume of rain > 12.7 mm; Renard et al., 1997) as from the formation of the EG up to the experimentation date (MJ mm ha^{-1} h^{-1}).

Edaphic, topographic and rainfall variables

Forty-nine soil variables proposed in the literature as potential drivers of soil erodibility by EGs were determined in each situation (Table 4). A first set of variables was measured directly in the field in areas close to the EG channel, where some variables were measured on a microplot (0.0625 m^2) after the action of a controlled rainfall (F_{IN}, Table 4) and others were measured outside that plot (F_{OUT}, Table 4). In the former, the following properties were determined: (i) hydraulic conductivity of the crust formed after rainfall simulation, and (ii) soil and surface crust resistance to the penetration. In the latter, the variables measured were: (i) soil bulk density, and (ii) soil shear strength.

After oven-drying and sieving at (< 2 mm), a composed sample of topsoil (0-15 cm) close to the EG channel was used to carry out several tests in the laboratory (L, Table 4): (I) physico-chemical variables (e.g. organic matter, stoniness, structural stability indices, etc.), (ii) soil susceptibility to sealing and crusting, and (iii) aggregate stability. In addition, undisturbed soil samples were extracted in 15 cm high metal cylinders to determine the critical shear stress and the erodibility coefficient of the soil. For this purpose, a Jet Test apparatus was

used under controlled laboratory conditions (Hanson and Cook, 2004; Hanson and Hunt, 2007). The Jet Test apparatus was the same as the one used by Hanson and Hunt (2007). This device consisted of the following parts: a jet tube, nozzle, point gage, and jet submergence tank where the soil samples were placed. Ten soil samples were taken from the vicinity areas across each EG path. Before starting the Jet Test, the soil material was saturated to their field capacity by the absorption of water by capillarity. Then, soil samples were stored for 48 h to give time for the soil particles to hydrate (Al-Madhhachi et al, 2013; Hanson and Hunt, 2007). This procedure allowed the following: (1) to begin all Jet Tests with the same initial soil moisture conditions, and (2) to avoid the slaking of soil aggregates caused by rapid water uptake at the beginning of the Jet Test. This type of soil breakdown is not caused by the direct effect of the impinging jet and could disturb the Jet Test results. The scour data generated by the Jet Test were analyzed using a spreadsheet routine developed by Hanson and Cook (2004), by using the Blaisdell solution to fit the scour equation (Daly et al., 2013). The laboratory Jet Test apparatus, the procedure, and the analysis method used to fit the scour depth equation are described in the Appendix 1.

Furthermore, two topographic indices based on the mean weighted slope by the area (AS_1, equation 3) and by the length of the EG channel (AS_2, equation 4) (Casalí et al., 1999; De Santisteban et al., 2005) were also determined for each EG.

$$AS_1 = A\,\frac{\sum_{i=1}^{n} A_i \cdot S_i}{\sum_{i=1}^{n} A_i} = \sum_{i=1}^{n} A_i \cdot S_i \quad (3)$$

$$AS_2 = A \cdot \frac{\sum_{j=1}^{m} L_j \cdot I_j}{\sum_{j=1}^{m} L_j} \quad (4)$$

where A is the total EG watershed drainage area (m^2), S_i is the slope of each of the *n* sub-watershed units with uniform slope (m^{-1}), A_i is the area of each of the *n*

sub-watershed units (m^2), L_j is the length of each of the *m* segments of the EG channel with uniform slope and length (m), and l_j is the slope of each of the *m* previous segments (m m^{-1}). Table 3 shows the values of the previous morphological attributes in each of the EGs studied.

Several studies (e.g. Capra et al., 2009; Casalí et al., 1999; Castillo et al., 2003; Luffman et al., 2015) have shown that EG formation is conditioned by the antecedent soil moisture, which affects runoff generation during rainfall events. In addition, it is well known that several soil properties (e.g. aggregate stability, shear strength, etc.) are strongly influenced by initial soil moisture conditions (Bryan, 2000). As there were no direct soil moisture measurements available, it was decided to calculate a simple and commonly used surrogate such as the antecedent accumulated rainfall (Capra et al., 2009). In this way, 5 potential surrogate variables of soil moisture at the moment of EG formation were obtained: total rainfall accumulated during 1 hour (aP1, mm), 1 day (aP1d, mm), 5 days (aP5d, mm), 7 days (aP7d, mm) and 21 days (aP21d, mm) before the storm event which triggered the EG formation (Table 2).

Statistical analysis

Soil variables were represented by the mean value of the measurements taken in the field or in the laboratory (see Table 4). On the contrary, a determined specific value was considered for the rest of variables: TSLs, AS_1, AS_2 (Table 3) and the 5 variables of antecedent rainfall to the EG formation used as soil moisture surrogate (Table 2).

The existence of significant relationships between the variable TSLs and the remaining variables was analyzed using 3 different multivariate statistical procedures: Cluster Analysis (CA), Principal Component Analysis (PCA) and

225 Multiple Regression Analysis (MRA). These tools were applied independently
226 and without attributing any preference or prior assumptions of performance to any
227 of them. Thus, the results obtained were interpreted independently. All the
228 statistical analysis statistics techniques in this study were performed using the R
229 statistics software version 3.1.1 (R Core Team, 2015).

230 The CA is a non-supervised data reduction technique designed to classify
231 observations in subgroups denominated clusters, using all the information in the
232 initial dataset and without making previous assumptions (Shrestha and Kazama,
233 2007). In this study, a hierarchical agglomerative CA was applied on the data by
234 using the Ward method. Also, the Euclidean distance was assigned as a similarity
235 measurement among the sample units (i.e. 20 soils studied). This method is
236 characterized by a greater grouping potential, because it uses more information
237 on the contents of the cluster than other methods do (Willet et al., 1987).
238 Therefore, clusters obtained were characterized by the mean values of the
239 variables defining them, which were statistically different from the mean of the
240 total population (Anderberg, 1973). The variables displaying a greater statistical
241 significance during the cluster formation were identified by showing a value of $p < 0.001$ in Student t test. This statistical test was used to assess whether the
243 mean value of those variables differed from the mean value of the total
244 population.

245 The PCA technique provides a reduction in the original dataset dimensionality
246 underlying the most meaningful information with a minimum loss of information
247 (Abdi and Willians, 2010). To achieve this, the PCA calculates new artificial non-
248 correlated variables called Principal Components (PCs), which are obtained
249 through linear combinations of the original variables (Bayat et al., 2013). If

250 necessary, the PCA can be oriented towards those variables of special interest
251 (i.e. supplementary variables), which enables the analysis of the results based
252 on those variables, without interfering in the analytical process itself (Abdi and
253 Willians, 2010). In this study, the variable TSLs was fixed as a supplementary
254 one. To interpret the results obtained more easily, the PCs were subjected to a
255 Varimax type rotation (Westra et al., 2010). Finally, those PCs presenting both
256 an eigenvalue > 1 (Kaiser, 1960) and variables with a correlation factor ≥ 0.50
257 with the supplementary variable were identified (Ollobarren et al., 2017).

258 The MRA aims at obtaining the relationship between two or more explanatory
259 (or independent variables) and one response (or dependent variable); for this
260 purpose, it applies a linear equation to the data observed. In this study, The MRA
261 was used as a weighting tool for the variables of the total population which best
262 fitted an explanatory linear model for the variable TSLs (fixed as a dependent
263 variable). The principle of parsimony was applied to balance the goodness of fit
264 of the model and its complexity, thus preventing its overfitting (Vandekerckhove
265 et al., 2015). Therefore, all possible linear models from 1 to 4 variables were
266 obtained and analyzed to seek the best explanatory model of TSLs. In each of
267 the models obtained, the following evaluation criteria were applied to diagnose
268 the best model: (i) calculation of the variance inflation factor (VIF) to discard the
269 independent variables presenting multicollinearity (i.e. VIF > 2) (Lin, 2008); (ii)
270 Akaike information criteria (AIC) values to select the best model (Akaike, 1974);
271 (iii) verifying the goodness of fit obtained by the model employing the Nash-
272 Sutcliffe efficiency coefficient (NSE) and the mean square error (MSE) (Moriasi
273 et al., 2007); (iv) regression diagnosing based on the significance level ($p < 0.05$,

Student t test) for the regression coefficient of each independent variables conforming the model (Walpole et al., 2011).

Finally, the robustness of the explanatory model candidate obtained after MRA was evaluated by means of the statistics technique FITEVAL (Ritter and Muñoz-Carpena, 2013). This tool develops an objective assessment of the goodness of fit of a proposed model based on the existence of statistical significance. Using a specific bootstrapping technique, followed by the correction of the bias and the calculation of confidence intervals, the approximate distribution probability of 2 statistical indicators of the model's efficiency is obtained: NSE and the root mean square error (RMSE). If the NSE value exceeds a previously fixed threshold value, the validity of the model is statistically accepted or rejected. In this work, the threshold values proposed in Moriasi et al. (2007) were applied to evaluate the model as: unsatisfactory (NSE < 0.50), acceptable (0.50 < NSE < 0.65), good (0.65 < NSE < 0.75), or very good (NSE > 0.75). Furthermore, this technique also evaluates the sensitivity of the above-mentioned indicators to the model's bias, as well as the presence of outliers.

Results

Cluster Analysis (CA)

The CA showed the presence of 2 clusters (Cluster A and Cluster B) in which the soils were grouped as follows: 13 in Cluster A (ABA 1 to 3, AOI 1 to 5, LUM 1, RAD 1, RAD 3, RAD 4 and RAD 6) and 7 in Cluster B (AOI 6, LEO 1, PIT 1, PIT 2, RAD 2, RAD 5 and RAD 7). The Cluster B soils displayed a 2.6-fold higher mean erosion rate (TSLs) than that recorded in Cluster A (median of 0.016 and of 0.006 t MJ^{-1} mm^{-1} h yr^{-1}, respectively). This result suggests that soils'

299 susceptibility to erosion due to EGs could be tentatively related to the range of
300 values of the variables identified in CA.

301 Of the 19 variables identified in the CA (Table 5), only 2 of them showed a
302 value of $p < 0.001$ after the Student t test: CR_4 and CR_5. Although these variables
303 were obtained from a similar granulometric balance, variable CR_5 incorporated a
304 clay fraction, whereas CR_4 did not (see Table 4). Thus, both variables were
305 different and therefore they were selected. In addition, it is worth noting that the
306 cationic exchange capacity (CH_5), although discarded, showed a statistical
307 significance very close to the threshold fixed for the selection of the key variables
308 of the CA ($p = 0.013$, Table 5). The outstanding variables from the CA were CR_4
309 and CR_5.

310 *Principal Component Analysis (PCA)*

311 PCs with a higher eigenvalue than the unit were obtained (Table 6). However,
312 only the first 2 (PC_1 and PC_2) showed a significant correlation (0.14 and 0.37,
313 respectively) (Table 6) with the supplementary variable (TSLs). Therefore, the
314 rest of PCs were discarded, focusing the analysis of the results on PC_1 and PC_2,
315 which, in turn, were capable of explaining 36.4% of the total variance of the
316 original information (20.7% and 15.7%, respectively). It should be noted that only
317 those variables presenting a correlation ≥ 0.50 with the PCs were investigated
318 (Table 6).

319 Thus, 22 variables were selected and grouped into 9 groups in accordance
320 with their typology: (i) organic matter content (CH_1); (ii) soil texture composition
321 and organic matter content (CR_3, CR_4, GF_1, GF_2, GF_3, GF_4, GF_5, SI_1, SI_2, and
322 SI_3); (iii) soil aggregate stability (SI_4, SI_5, and SI_6); (iv) soil sealing susceptibility
323 (CR_2); (v) soil crust hydraulic conductivity measured in field (HY_3); (vi) soil crust

resistance to penetration (PR_1); (vii) resistance to shear strength (SS_3); (viii) cationic exchange capacity (CH_5); and (ix) antecedent moisture in the soil before the formation of the EGs (aP5d, aP7d, and aP21d). For groups with more than one variable (ii, iii, and ix), those with a higher correlation with the PCs and, therefore, with TSL, were selected: CR_4, SI_6, and aP5d. Variables SI_4 and SI_5 were also selected in defining soil aggregate stability against different disaggregation mechanisms from that defined by variable SI_6 (see Table 4). Finally, the variables highlighted by the PCA were 11: CH_1, CR_2, CR_4, SI_4, SI5, SI_6, HY_3, PR_1, SS_3, CH_5, and aP5d.

Multiple Regression Analysis

In MRA, all the possible relationships between the response variable, TSLs, and the independent variables estimated in the study were analyzed. Thus, all the models with 1 variable (56), 2 variables (1540), 3 variables (27720) and 4 variables (367290) were obtained.

After applying regression diagnosis criteria, the best regression models in each situation were obtained (Table 7). Therefore, the best model with 1 variable was procured with E_1 as an independent variable (NSE = 0.78, AIC = -103.73). The best model with 2 variables was the one formed by the variables E_1 and CR_2 (NSE = 0.83, AIC = -107.25). For 3 variables, the best model was obtained with variables E_1, CR_2 and SS_3 (NSE = 0.85, AIC = -107.28). Finally, the best model constructed with 4 variables was defined with variables E_1, CR_2, SS_3 and CH_5 (NSE = 0.87, AIC = -108.44). This last model presented the highest value of NSE and the lowest one of AIC, which indicates a better balance between the goodness of fit and the complexity of the model (Akaike, 1974). Therefore, the model with 4 variables (equation 5) was selected to yield the best relationship

between the variables analyzed (E_1, CR_2, SS_3, and CH_5) and the EG erosion rate (TSLs).

$$TSLn\ x\ 10^3 = -31.8 + 0.4\ E_1 + 0.5\ CR_2 - 1.2\ SS_3 + 1.8\ CH_5 \qquad (5)$$

With the exception of variable E_1 (i.e. erodibility coefficient obtained through the Jet Test methodology) (Table 4), the rest of variables identified by MRA also stood out in the previous statistical analyses (see above). It is worth highlighting the importance of variable E_1 in the MRA, since it has been identified in all the best models obtained (see Table 7). Thus, this variable was able to explain by itself 78% of the TSL values obtained using the linear regression explanatory model (Figure 2). Consequently, and even without having been determined in previous analyses, the relationship between variable E_1 and EG erosion was identified.

The tool FITEVAL (Ritter and Muñoz-Carpena, 2013) was applied on the best explanatory model obtained for the variable TSLs (equation 5, see above). Figure 3 shows that the prediction of the variable TSLs is considered very good (NSE > 0.75) in 70.7% of the cases, and only in 9.6% of the situations was the model's fit unsatisfactory (NSE < 0.50). Therefore, the goodness of the fit of the model proposed is statistically valid, since the probability of that model's fit being considered unsatisfactory does not exceed 10% of the cases obtained ($p < 0.10$; Ritter and Muñoz-Carpena, 2013).

Guide values for the variables selected

After applying the 3 statistical techniques, a total of 13 key variables in the control of the vulnerability of the soils studied to EG erosion were identified: aP5d (accumulated rainfall 5 days before the initial event), CH_1 (organic matter percentage), CH_5 (cationic exchange capacity), CR_2 (relative sealing index), CR_4

374 (sealing-crusting index), CR_5 (crusting index), E_1 (soil erodibility coefficient), HY_3
375 (hydraulic conductivity of soil crust), SI_4 (stability of the aggregates from clay
376 swelling due to slow wetting), SI_5 (stability of the aggregates against slaking), SI_6
377 (stability of the aggregates against mechanical breaking-up), PR_1 (soil crust
378 resistance to penetration) and SS_3 (resistance to shear strength). Furthermore, a
379 guide value was determined by taking the mean values reached by the previous
380 variables in the cluster most resistant to erosion (Cluster A, Table 8). Thus, and
381 for the soils analyzed, the transition between resistant soils and those vulnerable
382 to erosion due to EGs (Cluster B) can be roughly defined.

383 Starting from the mean values of the key variables in both clusters (Table 8),
384 Figure 4 defined the existence of 2 areas of susceptibility to EG erosion: an area
385 with high erodibility (TSLs = 0.016 t MJ^{-1} mm^{-1} h yr^{-1}, red area) and another one
386 with lesser erodibility (TSLs = 0.006 t MJ_{-1} mm_{-1} h yr_{-1}, green area).

387 Based on the above, a new soil (with similar soil properties and topography to
388 those analyzed here) could be classified as being less susceptible to EG erosion
389 if it displayed, approximately, values higher than 1.68 mm (CV = 1.68) for aP5d,
390 than 1.29% (CV = 0.29) for CH_1, than 11.15 cmol (+) kg^{-1} (CV = 0.27) for CH_5,
391 than 1.64 mm h^{-1} (CV = 0.53) for HY_3, than 0.60 mm (CV = 0.57) for SI_4, than
392 1.49 mm (CV = 0.46) for SI_5, than 0.89 mm (CV = 0.54) for SI_6, and than 14.32
393 kPa (CV = 0.35) for SS_3; and lower than 35.02 (CV = 0.76) for CR_2, than 4.57
394 (CV = 0.39) for CR_4, than 1.24 (CV = 0.25) for CR_5, than 110.29 (CV = 0.87) cm^3
395 $(N\ s)^{-1}$ for E_1, and than 280.72 kPa for PR_1 (CV = 0.36) (where CV represents the
396 highest coefficient of variation in the variables in the 2 clusters, Table 8).
397 However, these guide values should be interpreted independently of each other,

since, in this study, the possible interactions between the key variables proposed were not evaluated.

Discussion

The individual relationships between a set of 13 key variables and erosion due to EGs (TSLs) were determined statistically. Nevertheless, the interdependence among those variables was not examined. So, vulnerability to erosion from EGs in the soils analyzed was due to the action of several factors.

The first factor is the content of cementing agents of soil particles, namely organic matter and clay. A low percentage of clay reduces the structural stability of the soil and increases its erodibility in the face of concentrated flows (Knapen et al., 2007; Rapp, 1998). Thus, in our study, the soils most susceptible to erosion in Cluster B exhibited a lower clay content than the less susceptible ones in Cluster A: 23.74 and 33.15%, respectively. This result is consistent with the conclusions shown by Sheridan et al. (2000) and Li et al. (2015a), who, on diverse agricultural soils, correlated negatively the clay content with the erodibility from rills and EGs, respectively; although they do not provide any threshold values for those relationships. On the other hand, Cantón et al. (2009) and Li et al. (2015b) observed a significant increase in the level of erodibility in various agricultural soils when the organic matter was less than 2%. In our study, the soils most resistant to erosion (Cluster A), precisely, presented a higher organic matter content than the threshold cited (CH_1 = 2.49%, Table 8), whereas the most erodible ones (Cluster B) were found to be below the cited threshold (CH_1 = 1.29%, Table 8). From all of the above, the importance of variables CR_4 and CR_5 can be deduced in determining both of them by means of a balance between the

soil fine fraction and the organic matter (see Table 4). It was following this reasoning that Pulido-Moncada et al. (2009) established, for five agricultural soils in Venezuela, critical values for CR_4 and CR_5 (3.33 and 0.70, respectively), above which the risk of sealing and soil erosion increased as a result of the dominance of silts and fine sand over the clay and the organic matter present in the soil. In our experiments, values above those thresholds for both variables were precisely found (4.57 and 1.24, respectively) in the soils defined as being more vulnerable (Cluster B, Table 8). However, values below (or very similar to) those thresholds were recorded in the more resistant soils of Cluster A (2.14 and 0.77, respectively, Table 8). Based on this, it can be concluded that, when physico-chemical, non-active particles (silts and sands) dominate over clay and organic matter, erodibility increases. Similar results were obtained by Chaplot et al. (2013) and Lentz et al. (1993) in agricultural soils affected by (ephemeral) gullies.

Secondly, soil erodibility against concentrated flows is related to soil aggregate stability (Govers et al., 1990). Our study has precisely highlighted the 3 variables (SI_4, SI_5, SI_6) proposed by Le Bissonnais (1996) for quantifying the structural stability of the soil against different breaking-up mechanisms (see Table 4). For those variables, Le Bissonnais (1996) found that values of over 0.8 mm would indicate a lower level of crusting, increasing infiltration and reducing erosion. This agrees with our results, in which the soils most resistant to erosion in Cluster A presented values higher than the threshold cited for the 3 previous variables: SI_4 = 1.06 mm, SI_5 = 2.19 mm and SI_6 = 1.89 mm (Table 8). Similarly, Chaplot el al. (2013), in grassland areas in South Africa, related a low value for the rate of erosion by gullies with a high stability for the soil aggregates (SI_6 values

447 comprised between 2.8 and 3.2 mm). However, these higher values for SI_6 could
448 be due to the different use of soil in both research works.

449 The role of variable SI_4 is also noteworthy, as it reflects the stability of the
450 aggregates in the face of breaking-up processes caused when heavy rainfall
451 events fall onto dry soils (Fox et al., 1998). These conditions were recorded on
452 the most erodible soils in Cluster B, in which wetness preceding the formation of
453 the EGs was almost negligible (aP5d = 1.68 mm, Table 8). That is why, on these
454 soils, the lowest values for variable SI_4 (0.60 mm, Table 8) were found. This
455 agrees with the findings demonstrated by Geng et al. (2015), who, on soils of the
456 Chinese loess plateau, related increases in erodibility to concentrated flow when
457 the value of SI_4 was lower than approximately 1 mm. So this result suggests that
458 a low content of antecedent water would be related to a higher instability of the
459 aggregates, followed by greater erosion due to EGs. This hypothesis is supported
460 by the studies of Nachtergaele and Poesen (2002) and Knapen and Poesen
461 (2010), who, on Belgian loess soils affected by rills and EGs, positively correlated
462 the erodibility of the soil with a reduction in the antecedent water to the formation
463 of those erosion phenomena.

464 Thirdly, the effect of cations in the soil on its structural stability stood out
465 through the cationic exchange capacity (CH_5). When the percentage of
466 exchangeable sodium percentage (ESP) is dominant over the bivalent cations
467 (Ca^{2+} and Mg^{2+}), the structural stability diminishes, thus producing erosion
468 (Bronick and Lal, 2005). Rienks et al. (2000) reported a greater vulnerability to
469 gully erosion when ESP value was over 20% and clay content under 25% on
470 South African soils. A similar trend was detected in our study, in which the most
471 erodible soils with a lesser content in clays (23.74%) from Cluster B displayed a

much higher ESP value than that of the more resistant soils and with a higher content in clays (31.15%) from Cluster A: 30.39 and 4.00%, respectively. Van Zijl et al. (2014), on soils in South Africa, fixed 0.67% as the threshold value of PSI, above which the dispersion of the soil and erosion due to gullies proportionally increased.

Fourthly, the susceptibility of the soil to sealing and to surface crusting stood out, which was quantified by means of the relative sealing index (CR_2) and of the permeability of the crust measured in the field (HY_3). Thus, lower values of HY_3 together with higher values of CR_2 would indicate a reduction in the hydraulic permeability of the soil. This situation would generate a higher runoff rate and, therefore, a greater vulnerability to erosion from EGs (Martínez-Casasnovas et al., 2003). Ramos et al. (2003) obtained the highest values for the variable CR_2 (up to 3 times) on agricultural soils which showed poor stability in its aggregates after applying the 3 test proposed by Le Bissonnais (1996). A similar trend was recorded in our study, in which an increase of approximately 50% over the value of CR_2 was detected in soils with a lesser structural stability (Cluster B) with respect to those with greater stability in their aggregates (Cluster A): 35.02 and 23.50 mm h^{-1}, respectively. On the other hand, Lozano et al. (2000) found a minimum value of 1.7 mm h^{-1} for variable HY_3 when the silt and sand contents increased to above 84% on 4 agricultural soils susceptible to crusting in Venezuela. This result is similar to the one obtained in our study, in which the most erodible soils (silt + sand = 76.27%) gave a similar value of HY_3 to that of the cited threshold (1.64 mm h^{-1}, Table 8), whereas the less vulnerable ones (silt + sand = 64.85%) displayed a higher value for that variable (2.61 mm h^{-1}, Table 8). However, this last permeability value is below the threshold of 5 mm h^{-1}

497 proposed by Florentino (1998) to identify those soils prone to undergoing surface
498 sealing. Therefore, the soils in both clusters would be in some way vulnerable to
499 forming a surface seal under the action of erosion agents.

500 In the fifth place, susceptibility of erosion is also influenced by the mechanical
501 resistance of the soil, namely resistance to shear strength (SS_3) and to the
502 penetration of the crust (PR_1). Increases in the value of SS_3 are related to a
503 greater resistance of the soil to concentrated flow (Knapen et al., 2007). Poesen
504 and Govers (1990) and Geng et al. (2015), both on agricultural soils, correlated
505 negatively the erodibility of the soil due to gullies and rills with the highest values
506 of SS_3 –measured with a Torvane device– for a range of values comprised
507 between 2.5 and 15 kPa. Nevertheless, those authors did not report any threshold
508 value of SS_3 to define the vulnerability of the soil before the erosion process.
509 Namely, a similar behavior to the above-mentioned was obtained in our
510 experimentation. Thus, the less erodible soils in Cluster A presented a 50%
511 higher value of SS_3 (21.48 kPa) than the one recorded in the more vulnerable
512 soils in Cluster B (14.32 KPa) (Table 8). Again, high values of PR_1 are related to
513 higher runoff and erosion rates (Gabriels et al., 1997). In our study, average
514 values for variable PR_1 were obtained, lower in Cluster A (263.18 kPa) than in
515 Cluster B (280.72 kPa). Although the difference between PR_1 values in both
516 clusters is not high, it agrees with the results of Bouma and Imeson (2000), who,
517 in semi-arid areas of Alicante (Spain), correlated positively the value of PR_1
518 (measured with a pocket penetrometer) with a greater risk to rill erosion.

519 Finally, the susceptibility to erosion from EGs was reflected by the soil
520 erodibility coefficient (variable E_1) obtained through the Jet Test assay. This
521 technique reproduces, under controlled conditions, the physical process of the

formation of (ephemeral) gully headcuts (Hanson and Cook, 2004; Stein and Julien, 1993). Thus, high values for E_1 would be related to a greater susceptibility of the soil to the appearance of a gully headcut in the face of a rainfall event. Since this technique has been oriented towards evaluating the stability and migration of headcuts in dams and streambanks (e.g., Daly et al., 2015; Hanson et al., 2003), the authors are aware of no works that have evaluated the relationship between parameter E_1 and the erosion rate due to (ephemeral) gullies. As an exception, Potter et al. (2002) used the Jet Test to estimate the erodibility of 6 agricultural soils in Mexico in situ, finding a positive correlation between variable E_1 and the silt + very fine sand content in the soil textural fraction. In our study, the highest value of E_1 was precisely determined in the more vulnerable soils (and with a high content of silt and sands, see above) in Cluster B (110.29 cm^3 $(N\ s)^{-1}$), whereas, on the more resistant soils in Cluster A (richer in clays, see above), the value of this variable was of 102.53 cm^3 $(N\ s)^{-1}$. Similarly, Daly et al. (2016), over 3 streambanks with different soil textures, reported that an increase in clay content was related to lower values of E_1. The latter occurred as a result of the increase in bulk density, which diminished the distance between soil particles and reduced susceptibility to swelling in the clays and to erosion (Wynn and Mostaghimi, 2006). Unfortunately, Daly et al. (2016) did not obtain any significant correlation with the previous property or with any of the soil parameters evaluated (for example, texture, bulk density, etc.), and they concluded that parameter E_1 should be measured directly, as it could not be estimated from empirical relationships with soil properties.

The small difference between the E_1 values in both clusters could occur because this variable was not identified by the CA but by the MRA, in which it

gained great importance in the best explanatory model of TSLs (Table 7). Furthermore, the high positive correlation shown between E_1 and TSLs (R^2 = 0.78, Figure 2) reveals E_1 as a promising parameter for identifying the degree of erodibility due to (ephemeral) gullies in agricultural soils in which this erosion process is not completely controlled by their topography, and the soil factor conditions to a great extent the appearance and development of EGs (Taguas et al., 2010). This is maintained in this study, since no statistical relationship was identified between the topography (variables AS_1 and AS_2, Table 3) and the variable TSLs.

Finally, it is important to point out that the 13 variables identified in this study showed a statistical relationship with the rates of erosion from EGs measured in field experimentation. So, the number of rainfall events conditioning the volume of soil removed by the EGs was not the same in all situations, hence the normalization applied (see above). Therefore, any one of the variables proposed here (and their measured values) could differ from the results obtained in studies in which the experiment conditions are homogeneous and are completely controlled, such as the case of interrill erosion reproduced by rainfall simulation (Ollobarren et al. 2017). Further research would be necessary to evaluate the 13 key variables (and their interactions) in other soil types. This would allow the definition of erodibility indices *per se*, which could be incorporated into the current erosion models with the aim of improving the prediction of the location and of the volumes of soil eliminated by this type of water erosion.

Conclusions

In agricultural lands with a smooth surface relief, the soil's nature and conditions play a key role in the erosion process, giving rise to the appearance of ephemeral gullies. Under these circumstances, soil susceptibility to ephemeral gully development has been reflected in a set of 13 soil variables, representing a wide range of soil physico-chemical properties. Among these, a coefficient of erodibility (E_1), determined by means of the Jet Test technique, stands out; to a certain extent, this emulates precisely the genesis of an (ephemeral) gully headcut. It is worth noting that these variables could be of use for the evaluation of large-scale areas (e.g. watersheds), since the techniques used to obtain them are economical and easy to apply.

These variables and their respective guide values –which approximately indicate the transition between soils resistant or vulnerable to erosion due to EGs– were defined for soils in Mediterranean environments. They would also be applicable to soils of a different nature, but it is likely that the guide values, determined empirically, would be different.

Acknowledgements

This study was partly financed by the Spanish Government's 'National Plan for Research, Development and Technological Innovation' (projects CGL2011-24336 (MICIN/FEDER) and CGL2015-64284-C2-1-R (MINECO/FEDER)). The authors thank the Public University of Navarre (UPNA) for funding this work with a grant awarded to the principal author. In addition, the authors are grateful for the support and assistance given by Dr. Antonina Capra and Dr. Carmelo La Spada from the Università degli Studi Mediterranea di Reggio Calabria (Italia).

596 **References**

597 Abdi H, Williams LJ. 2010. Principal component analysis. Wiley Interdisciplinary
598 Reviews: Computational Statistics2: 433-459. DOI: 10.1002/wics.101.

599 Al-Madhhachi AT, Hanson GJ, Fox GA, Tyagi A., Bulut R. 2013. Deriving
600 parameters of a fundamental detachment model for cohesive soils from flume
601 and jet erosion tests. Trans. ASABE 56 (2): 489-504. DOI:
602 http://dx.doi.org/10.13031/2013.42669.

603 Anderberg MR. 1973. Cluster analysis for applications. Academic Press, New
604 York: 359 pp.

605 Akaike H. 1974. NEW LOOK AT THE STATISTICAL MODEL IDENTIFICATION.
606 IEEE Transactions on Automatic Control AC-19: 716-723.

607 Archibold O, Levesque L, De Boer D, Aitken A, Delanoy L. 2003. Gully retreat in
608 a semi-urban catchment in Saskatoon, Saskatchewan. Applied Geography 23:
609 261–279. DOI: 10.1016/j.apgeog.2003.08.005.

610 Bayat H, Neyshaburi MR, Mohammadi K, Nariman-Zadeh N, Irannejad M. 2013
611 Improving water content estimations using penetration resistance and principal
612 component analysis. Soil and Tillage Research 129: 83-92. DOI:
613 10.1016/j.still.2013.01.009.

614 Bennett SJ, Casalí J, Robinson KM, Kadavy KC. 2000. Characteristics of actively
615 eroding ephemeral gullies in an experimental channel. Transactions of the
616 American Society of Agricultural Engineers 43: 641-649.

617 Bingner RL, Theurer FD, Yuan YP. 2015. AnnAGNPS technical processes.
618 Washington, D.C. USDA-ARS.

619 Blaisdell FW, Clayton LA, Hebaus CG. 1981. Ultimate dimension of local scour.
620 J. Hydraul. Div. ASCE 107 (HY3), 327–337.

621 Boiffin J, Monnier G. 1985. Infiltration rate as affected by soil surface crusting
622 caused by rainfall. International symposium on the assessment of soil surface
623 sealing and crusting. En: Assessment of soil surface sealing and crusting,
624 proceedings of the Symposium held in Gent, Belgium: 210-217.

625 Bouma NA, Imeson AC. 2000. Investigation of relationships between measured
626 field indicators and erosion processes on badland surfaces at Petrer, Spain.
627 Catena 40: 147-171. DOI: 10.1016/S0341-8162(99)00046-6.

628 Bouyoucos GJ. 1935. The clay ratio as a criterion of susceptibility of soils to
629 erosion. Journal of American Society of Agronomy 27: 738-741.

630 Bradford JM, Truman CC, Huang C. 1992. Comparison of three measures of
631 resistance of soil surface seals to raindrop splash. Soil Technology 5: 47-56. DOI:
632 10.1016/0933-3630(92)90006-M.

633 Bronick CJ, Lal R. 2005. Soil structure and management: A review. Geoderma
634 124: 3-22. DOI: 10.1016/j.geoderma.2004.03.005.

635 Bryan RB. 2000. Soil erodibility and processes of water erosion on hillslope.
636 Geomorphology 32: 385-415. DOI: 10.1016/S0169-555X(99)00105-1.

637 Bryan RB. 2004. Gully-scale implications of rill network and confluence
638 processes. In: Li, Y., Poesen, J., Valentin, C. (Eds.), Gully Erosion Under Global
639 Change. Sichuan Science and Technology Press, Chengdu, China, pp. 73–95.

640 Cantón Y, Solé-Benet A, Asensio C, Chamizo S, Puigdefábregas J. 2009.
641 Aggregate stability in range sandy loam soils Relationships with runoff and
642 erosion. Catena 77: 192-199. DOI: 10.1016/j.catena.2008.12.011.

643 Capra A, Porto P, Scicolone B. 2009. Relationships between rainfall
644 characteristics and ephemeral gully erosion in a cultivated catchment in Sicily
645 (Italy). Soil and Tillage Research 105: 77-87. DOI: 10.1016/j.still.2009.05.009.

646 Capra A, La Spada C. 2015. Medium-term evolution of some ephemeral gullies
647 in Sicily (Italy). Soil and Tillage Research 154: 34-43. DOI:
648 10.1016/j.still.2015.07.001

649 Casalí J, López JJ, Giráldez JV. 1999. Ephemeral gully erosion in southern
650 Navarra (Spain). Catena 36: 65-84. DOI: 10.1016/S0341-8162(99)00013-2.

651 Casalí J, Loizu J, Campo MA, De Santisteban LM, Álvarez-Mozos J. 2006.
652 Accuracy of methods for field assessment of rill and ephemeral gully erosion.
653 Catena 67: 128-138. DOI: 10.1016/j.catena.2006.03.005.

654 Casalí J, Gastesi R, Álvarez-Mozos J, De Santisteban LM, Lersundi, J D V d,
655 Giménez R, Larrañaga A, Goñi M, Agirre U, Campo MA, López JJ, Donézar M.
656 2008. Runoff, erosion, and water quality of agricultural watersheds in central
657 Navarre (Spain). Agricultural Water Management 95: 1111-1128. DOI:
658 10.1016/j.agwat.2008.06.013.

659 Castillo VM, Gómez-Plaza A, Martínez-Mena M. 2003. The role of antecedent
660 soil water content in the runoff response of semiarid catchments: A simulation
661 approach. Journal of Hydrology 284: 114-130. DOI: 10.1016/S0022-
662 1694(03)00264-6.

663 Castillo C, Pérez R, James MR, Quinton JN, Taguas TV. 2012. Comparing the
664 Accuracy of Several Field Methods for Measuring Gully Erosion. Soil Science
665 Society of America Journal 76: 1319-1322. DOI: 10.2136/sssaj2011.0390.

666 Chaplot V. 2013. Impact of terrain attributes, parent material and soil types on
667 gully erosion. Geomorphology 186: 1-11. DOI: 10.1016/j.geomorph.2012.10.031.

668 Comerma J, Torres S, Lobo D, Fernández N, Delgado R, Madero L. 1992.
669 Aplicación del sistema de evaluación de tierras de la F.A.O. 1985 en la zona de
670 Turén, Venezuela. Cuadernos de Agronomía, año 1 1: 24.

671 Daggupati P, Sheshukov AY, Douglas-Mankin KR. 2014. Evaluating ephemeral
672 gullies with a process-based topographic index model. Catena 113: 177-186.
673 DOI: 10.1016/j.catena.2013.10.005.

674 Daly ER, Fox GA, Al-Madhhachi AT, Miller RB. 2013. A scour depth approach for
675 deriving erodibility parameters from Jet Erosion Tests. Trans. ASABE 56 (6):
676 1343-1351. DOI: http://dx.doi.org/10.13031/trans.56.10350.Daly ER, Fox GA, Al-

677 Madhhachi AST, Storm DE. 2015. Variability of fluvial erodibility parameters for
678 streambanks on a watershed scale. Geomorphology 231: 281-291. DOI:
679 10.1016/j.geomorph.2014.12.016.

680 Daly ER, Fox GA, Fox AK. 2016. Correlating site-scale erodibility parameters
681 from jet erosion tests to soil physical properties. Transactions of the ASABE 59
682 (1): 115-128. DOI: 10.13031/trans.59.11309.

683 De Ploey J, Mücher HJ. 1981. Consistency index and rainwash mechanisms on
684 belgian loamy soils. Earth Surface Processes and Landforms 6: 319-330. DOI:
685 10.1002/esp.3290060311.

686 De Santisteban LM, Casalí J, López JJ, Giráldez JV, Poesen J, Nachtergaele J.
687 2005. Exploring the role of topography in small channel erosion. Earth Surface
688 Processes and Landforms 30: 591-599. DOI: 10.1002/esp.1160.

689 De Santisteban LM, Casalí J, López JJ. 2006. Assessing soil erosion rates in
690 cultivated areas of Navarre (Spain). Earth Surface Processes and Landforms 31
691 (4): 487-506. DOI: 10.1002/esp.1281.

692 Desmet PJJ, Poesen J, Govers G, Vandaele K. 1999. Importance of slope
693 gradient and contributing area for optimal prediction of the initiation and trajectory
694 of ephemeral gullies. Catena 37: 377-392. DOI: 10.1016/S0341-8162(99)00027-
695 2.

696 Dong Y, Li F, Zhang Q, Lei T. 2015. Determining ephemeral gully erosion process
697 with the volume replacement method. Catena 131: 119-124. DOI:
698 10.1016/j.catena.2015.03.021.

699 E.S.R.I. 2000. ArcView GIS. Environmental Systems Research Institute, Inc.,
700 Redlands (U.S.A.).

701 FAO. 1980. Metodología provisional para la evaluación de la degradación de los
702 suelos. F.A.O. 1980, Roma-Italia: 86.

703 Florentino A. 1998. Guía para la evaluación de la degradación del suelo y de la
704 sostenibilidad del uso de la tierra: selección de indicadores físicos. Valores
705 críticos. En: Manejo Sostenible de los Suelos, Manual de Prácticas. Facultad de
706 Agronomía UCV. Maracay- Venezuela: 68-77.

707 Foster GR. 1986. Understanding ephemeral gully erosion. Soil Conservation, vol.
708 2. National Academy of Science Press, Washington, DC, pp. 90–125.

709 Fox DM, Le Bissonnais Y, Bruand A. 1998. The effect of ponding depth on
710 infiltration in a crusted surface depression. Catena 32: 87-100. DOI:
711 10.1016/S0341-8162(98)00042-3.

712 Gabriels D, Horn R, Villagra MM, Hartman R. 1997. Assessment, prevention, and
713 rehabilitation of soil structure caused by soil surface sealing, crusting and
714 compaction. In: Lal R. (ed). Methods for assessment of soil degradation. CRC
715 Press. Boca Raton. New York. 129-167.

716 Geng R, Zhang G, Li Z, Wang H. 2015. Spatial variation in soil resistance to
717 flowing water erosion along a regional transect in the Loess Plateau. Earth
718 Surface Processes and Landforms 40: 2049-2058. DOI: 10.1002/esp.3779.

719 Govers G, Everaert W, Poesen J, Rauws G, De Ploey J, Lautridou JP. 1990. A
720 long flume study of the dynamic factors affecting the resistance of a loamy soil to
721 concentrated flow erosion. Earth Surface Processes and Landforms 15: 313-328.
722 DOI: 10.1002/esp.3290150403.

723 Han Y, Zheng FL, Xu XM. 2017. Effects of rainfall regime and its character indices
724 on soil loss at loessial hillslope with ephemeral gully. Journal of Mountain Science
725 14(3): 527-538. DOI: 10.1007/s11629-016-3934-2.

726 Hanson GJ, Cook KR, Hahn W, Britton SL. 2003. Evaluating erosion widening
727 and headcut migration rates for embankment overtopping tests. ASAE Paper No.
728 032067. St.Joseph, Mich.: ASAE.

729 Hanson GJ, Cook KR. 2004. Apparatus, test procedures, and analytical methods
730 to measure soil erodibility in situ. Applied Engineering in Agriculture 20: 455-462.
731 DOI: 10.13031/2013.16492.

732 Hanson GJ, Hunt SL. 2007. Lessons learned using laboratory jet method to
733 measure soil erodibility of compaction soils. T. ASABE 23 (3): 305-312.

734 Hoober D, Svoray T, Cohen S. 2017. Using a landform evolution model to study
735 ephemeral gullying in agricultural fields: the effects of rainfall patterns on
736 ephemeral gully dynamics. Earth Surf. Process. Landforms 42: 1213–1226. DOI:
737 10.1002/esp.4090.

738 Kaiser HF. 1960. The application of electronic computers to factor analysis.
739 Educational and Psychological Measurement 20: 141–151.

740 Knapen A, Poesen J, Govers G, Gyssels G, Nachtergaele J. 2007. Resistance of
741 soils to concentrated flow erosion: A review. Earth-Science Reviews 80: 75-109.
742 DOI: 10.1016/j.earscirev.2006.08.001.

743 Knapen A, Poesen J. 2010. Soil erosion resistance effects on rill and gully
744 initiation points and dimensions. Earth Surface Processes and Landforms 35:
745 217-228. DOI: 10.1002/esp.1911.

746 Le Bissonnais Y. 1996. Aggregate stability and assessment of soil crustability
747 and erodibility: I. Theory and methodology. European Journal of Soil Science 47:
748 425-437.

749 Le Bissonnais Y, Cerdan O, Lecomte V, Benkhadra H, Souchère V, Martin P.
750 2005. Variability of soil surface characteristics influencing runoff and interrill
751 erosion. Catena 62 (2-3): 111-124. DOI: 10.1016/j.catena.2005.05.001.

752 Lentz RD, Dowdy RH, Rust RH. 1993. Soil property patterns and topographic
753 parameters associated with ephemeral gully erosion. Journal of Soil & Water
754 Conservation 48: 354-360.

755 Li Y, Poesen J, Valentin C. 2004. Gully Erosion Under Global Change. Sichuan
756 S&T Press.

757 Li ZW, Zhang GH, Geng R, Wang H. 2015a. Spatial heterogeneity of soil
758 detachment capacity by overland flow at a hillslope with ephemeral gullies on the
759 Loess Plateau. Geomorphology 248: 264-272. DOI:
760 10.1016/j.geomorph.2015.07.036.

761 Li ZW, Zhang GH, Geng R, Wang H. 2015b. Rill erodibility as influenced by soil
762 and land use in a small watershed of the Loess Plateau, China. Biosystems
763 Engineering 129: 248-257. DOI: 10.1016/j.biosystemseng.2014.11.002.

764 Lin F. 2008. Solving multicollinearity in the process of fitting regression model
765 using the nested estimate procedure. Quality and Quantity 42: 417-426. DOI:
766 10.1007/s11135-006-9055-1.

767 Leica Geosystems. 2006. Leica Geo Office v. 4.0.0.0. Program documentation.
768 Leica Geosystems AG, Heerbrugg (Switzerland).

769 Léonard J, Richard G. 2004. Estimation of runoff critical shear stress for soil
770 erosion from soil shear strength. Catena57: 233-249. DOI:
771 10.1016/j.catena.2003.11.007

772 Lozano Z, Lobo D, Pla I. 2000. Diagnóstico de limitaciones físicas en inceptisoles
773 de los llanos occidentales venezolanos. Bioagro 12 (1): 15-24.

774 Luffman IE, Nandi A, Spiegel T. 2015. Gully morphology, hillslope erosion, and
775 precipitation characteristics in the Appalachian Valley and Ridge province,
776 southeastern USA. Catena 133: 221-232. DOI: 10.1016/j.catena.2015.05.015.

777 Martínez-Casasnovas JA, Antón-Fernández C, Ramos MC. 2003. Sediment
778 production in large gullies of the Mediterranean area (NE Spain) from high-
779 resolution digital elevation models and geographical information systems
780 analysis. Earth Surface Processes and Landforms 28: 443-456. DOI:
781 10.1002/esp.451.

782 Moriasi DN, Arnold JG, Van Liew MW, Bingner RL, Harmel RD, Veith TL. 2007.
783 Model evaluation guidelines for systematic quantification of accuracy in
784 watershed simulations. Transactions of the ASABE 50: 885-900.

785 Nachtergaele J, Poesen J, Steegen A, Takken I, Beuselinck L, Vandekerchove
786 L, Govers G. 2001. The value of a physically based model versus an empirical
787 approach in the prediction of ephemeral gully erosion for loess-derived soils.
788 Geomorphology 40: 237-252. DOI: 10.1016/S0169-555X(01)00046-0.

789 Nachtergaele J, Poesen J. 2002. Spatial and temporal variations in resistance of
790 loess-derived soils to ephemeral gully erosion. European Journal of Soil Science
791 53: 449-463. DOI: 10.1046/j.1365-2389.2002.00443.x.

792 Nash JE, Sutcliffe JV. 1970. River flow forecasting through conceptual models
793 part I - A discussion of principles. Journal of Hydrology 10: 282-290. DOI:
794 10.1016/0022-1694(70)90255-6.

795 Ollobarren P, Giménez R, Campo-Bescós MA. 2016. Assessing soil susceptibility
796 to interrill erosion: an extensive empirical approach. Land Degradation
797 Development. In press.

798 Papadakis J. 1978. Climatic Classification and Terminology. Monthly Climates,
799 Buenos Aires.

800 Pla I. 1982. Sealing index to predict problems of soil and water conservation in
801 tropical rainfed agricultural lands. ASAESSASSSA, Annual Meetings. Anaheim,
802 California. USA. 1982.

803 Poesen J, Govers G. 1990. Gully erosion in the loam belt of Belgium: typology
804 and control measures. Soil erosion on agricultural land: 513-530.

805 Poesen J, De Luna E, Franca A, Nachtergaele J, Govers G. 1999. Concentrated
806 flow erosion rates as affected by rock fragment cover and initial soil moisture
807 content. Catena 36: 315-329. DOI: 10.1016/S0341-8162(99)00044-2.

808 Poesen J, Nachtergaele J, Verstraeten G, Valentin C. 2003. Gully erosion and
809 environmental change: Importance and research needs. Catena 50: 91-133. DOI:
810 10.1016/S0341-8162(02)00143-1.

811 Potter KN, Velázquez-Garcia JDJ, Torbert HA. 2002. Use of a submerged jet
812 device to determine channel erodibility coefficients of selected soils of Mexico.
813 Journal of Soil and Water Conservation 57: 272-277.

814 R Core Team. 2015. R: A language and environment for statistical computing. R
815 Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-
816 project.org/. 3.1.1

817 Rapp I. 1998. Effects of soil properties and experimental conditions on the rill
818 erodibilities of selected soils. Ph. D. Thesis, Faculty of Biological and Agricultural
819 Sciences, University of Pretoria, South Africa.

820 Ramos MC, Nacci S, Pla I. 2003. Effect of raindrop impact and its relationship
821 with aggregate stability to different disaggregation forces. Catena 53: 365-376.
822 DOI: 10.1016/S0341-8162(03)00086-9.

823 Renard KG, Foster GR, Weesies GA, McCool DK, Yoder DC. 1997. Predicting
824 soil erosion by water: A guide to conservation planning with the revised universal
825 soil loss equation (RUSLE). USDA Agricultural Handbook 703: 404 pp.

826 Rieke-Zapp D, Poesen J, Nearing MA. 2007. Effects of rock fragments
827 incorporated in the soil matrix on concentrated flow hydraulics and erosion. Earth
828 Surface Processes and Landforms 32: 1063-1076. DOI: 10.1002/esp.1469.

829 Rienks SM. Botha GA, Hughes JC. 2000. Some physical and chemical properties
830 of sediments exposed in a gully (donga) in northern KwaZulu-Natal, South Africa
831 and their relationship to the erodibility of the colluvial layers. Catena 39 (1): 11-
832 31. DOI: 10.1016/S0341-8162(99)00082-X.

833 Ritter A, Muñoz-Carpena R. 2013. Performance evaluation of hydrological
834 models: Statistical significance for reducing subjectivity in goodness-of-fit
835 assessments. Journal of Hydrology 480: 33-45. DOI:
836 10.1016/j.jhydrol.2012.12.004.

837 Schäuble H. 2003. Hydrotools 1.0 for ArcView. Technical University of Darmstadt.

838 Sheridan GJ, So HB, Loch RJ, Walker CM. 2000. Estimation of erosion model
839 erodibility parameters from media properties. Australian Journal of Soil Research
840 38: 265-284. DOI: 10.1071/SR99041.

841 Shrestha S, Kazama F. 2007. Assessment of surface water quality using
842 multivariate statistical techniques: A case study of the Fuji river basin, Japan.
843 Environmental Modelling and Software 22: 464-475. DOI:
844 10.1016/j.envsoft.2006.02.001.

845 Stein OR, Julien PY. 1993. Criterion delineating the mode of headcut migration.
846 Journal of Hydraulic Engineering 119: 37-50. DOI: 10.1061/(ASCE)0733-
847 9429(1993)119:1(37).

848 Taguas E, Yuan Y, Pea A, Ayuso JL. 2010. Prediction of ephemeral gullies by
849 using the combined topographical index in a micro-catchment of olive groves in
850 Andalucia, Spain. Agrociencia 44: 409-426.

851 Thorne CR, Zevenbergen LW. 1984. On-site prediction of ephemeral gully
852 erosion. Report to the U.S. Department of Agriculture, Agricultural Research and
853 Soil Conservation Services.

854 Truman CC, Bradford JM. 1990. Effect of antecedent soil moisture on splash
855 detachment under simulated rainfall. Soil Science 150: 787-798.

856 Valentin C, Poesen J, Li Y. 2005. Gully erosion: Impacts, factors and control.
857 Catena 63: 132-153. DOI: 10.1016/j.catena.2005.06.001.

858 Van Zijl GM, Ellis F, Rozanov A. 2014. Understanding the combined effect of soil
859 properties on gully erosion using quantile regression. South African Journal of
860 Plant and Soil 31: 163-172. DOI: 10.1080/02571862.2014.944228.

861 Vandaele K, Poesen J, Govers G, Van Wesemael B. 1996. Geomorphic threshold
862 conditions for ephemeral gully incision. Geomorphology 16: 161-173. DOI:
863 10.1016/0169-555X(95)00141-Q.

864 Vandekerckhove L, Poesen J, Oostwoud Wijdenes D, De Figueiredo T. 1998.
865 Topographical thresholds for ephemeral gully initiation in intensively cultivated
866 areas of the Mediterranean. Catena 33: 271-292. DOI: 10.1016/S0341-
867 8162(98)00068-X.

868 Vandekerckhove J, Matzke D, Wagenmakers EJ. 2015. Model comparison and
869 the principle of parsimony. In Busemeyer JR, Townsend JT, Wang ZJ, Eidels A.
870 (Eds.) Oxford Handbook of Computational and Mathematical Psychology: 300-
871 317. Oxford, UK: Oxford University Press.

872 Verachtert E, Van Den Eeckhaut M, Martínez-Murillo JF, Nadal-Romero E,
873 Poesen J, Devoldere S, Wijnants N, Deckers J. 2013. Impact of soil
874 characteristics and land use on pipe erosion in a temperate humid climate: Field
875 studies in Belgium. Geomorphology 192: 1-14. DOI:
876 10.1016/j.geomorph.2013.02.019.

877 Walpole RE, Myers RH, Myers SL, Ye K. 2011. Probability and Statistics for
878 Engineers and Scientists (9th Edition). Pearson Prentice Hall.

879 Westra S, Brown C, Lall U, Koch I, Sharma A. 2010. Interpreting variability in
880 global SST data using independent component analysis and principal component
881 analysis. International Journal of Climatology 30: 333-346. DOI:
882 10.1002/joc.1888.

883 Willet P. 1987. Similarity and Clustering in Chemical Information Systems.
884 Research Studies Press, Wiley, New York.

885 Wynn T, Mostaghimi S. 2006. The effects of vegetation and soil type on
886 streambank erosion, Southwestern Virginia, USA. Journal of the American Water
887 Resources Association 42 (1): 69-82. DOI: 10.1111/j.1752-1688.2006.tb03824.x.

888 Yoshimura K, Onda Y, Kato H. 2015. Evaluation of radiocaesium wash-off by soil
889 erosion from various land uses using USLE plots. Journal of environmental
890 radioactivity 139: 362-369. DOI: 10.1016/j.jenvrad.2014.07.01.

Appendix 1

The following is a step−by−step listing of the procedure for setting up and conducting a submerged jet test in the laboratory.

1. Field soil sampling:

   For each EG, ten soil samples were extracted in areas close to their erosive path by using a metallic drill. This device allows obtaining undisturbed soil samples in standard metal cylindrical molds, which are representative for the soil erodibility conditions of the topsoil.

2. Preparation of soil samples in the laboratory:

   Soil samples were saturated up to their field capacity by the absorption of water by capillarity. After that, samples were stored for 48 h to give time for the soil porosity to fully hydrate.

3. Laboratory Jet Test apparatus :

   This apparatus consisted of the following parts:

   a) The jet submergence tank was 305 mm in diameter and 305 mm in height. It is made of acrylic tubing.
   b) The jet tube had an 89 mm diameter orifice plate with a 6.4 mm diameter nozzle in the center of the plate. The jet nozzle was set perpendicular and at a variable height above the soil surface.
   c) The point gage can pass through the jet nozzle to the soil sample surface to read the depth of scour. The point gage diameter was

equivalent to the jet nozzle diameter; therefore, the water jet was shut off during scour readings.

d) A deflection plate was attached to the jet tube. This device was used to protect the soil surface, deflecting the impinging jet, during initial filling of the submergence tank.

e) The adjustable head tank was 880 mm in height and was utilized to provide a desired water head upstream of the jet nozzle. The water source was connected to the lower part of the head tank.

f) The air relief valve was used to remove air from the jet tube.

g) The hose was used to connect the head tank with the jet tube.

4. Jet Test procedure:

a) The mold was placed in the jet submergence tank, centering the soil sample directly beneath the jet nozzle.

b) The point gage was used to take three readings at time zero: (1) the height of the jet nozzle, (2) the initial scour depth on the soil surface, and (3) the height of the deflector plate as a reference point.

c) The deflector plate was placed in front of the jet nozzle. Then, the point gage was set against the plate, closing off the nozzle. Therefore, water flow was initiated to the head tank and jet tube.

d) Once the head tank and jet tube were filled with water, the point gage was set upstream of the jet nozzle to eliminate any flow disturbance from the point gage. The water then proceeded to impact the deflector plate and filled the submergence tank, submerging the sample and jet orifice.

e) The deflector plate was moved out of the way of the water jet, starting the test.

f) The point gage reading of the soil scour depth was initially taken every ten seconds. When an increase in scour was not detected, the jet impact time was increased.

g) The maximum depth of scour in the soil sample was monitored for a maximum of one hour or to a depth of scour of 116 mm, whichever occurred first.

5. Analysis of the jet test results:

The soil erodibility (E_1, Table 4) and critical shear stress (SS_5, Table 4) values were determined from the jet test results using the method described by Hanson and Cook (2004). The critical shear stress was estimated by fitting the scour depth versus the time data to a logarithmic hyperbolic function developed by Blaisdell et al. (1981) to determine the final depth of scour. Soil erodibility was determined by a least-squares fit of the test data to a dimensionless form of the excess shear stress. For each EG, the final E_1 and SS_5 values considered were the average of the 10 soil samples tested in the laboratory.

Tables

Table1. Main climate, topography, type and soil management attributes in the analyzed situations.

Soil name	Location	Land use	Climate (Papadakis)	Mean annual rainfall [a] (mm)	Accumulated rainfall [b] (%)	Mean slope (%)	Sand (%)	Silt (%)	Clay (%)	Soil texture (USDA)	Organic matter (%)	Stoniness (%)	Bulk density (g cm^{-3})
PIT 1	Navarre	Soft wheat	Dry temperate Mediterranean	546.50 (24 years)	74.86	9.51	26.21	48.49	25.30	Loam	1.07	0.56	1.56
PIT 2	Navarre	Soft wheat	Dry temperate Mediterranean	546.50 (24 years)	74.86	3.06	10.69	62.39	26.92	Silty loam	1.67	7.04	1.52
AOI 1	Navarre	Soft wheat	Wet temperate Mediterranean	894.40 (18 years)	77.61	22.79	5.49	61.93	32.58	Silty clay loam	3.34	12.10	1.15
AOI 2	Navarre	Soft wheat	Wet temperate Mediterranean	894.40 (18 years)	77.61	17.37	10.40	63.10	26.50	Silty loam	2.28	25.46	1.17
AOI 3	Navarre	Soft wheat	Wet temperate Mediterranean	894.40 (18 years)	77.61	25.47	13.81	48.79	37.40	Silty clay loam	2.34	19.52	1.20
AOI 4	Navarre	Soft wheat	Wet temperate Mediterranean	894.40 (18 years)	77.61	12.39	17.34	49.16	33.50	Silty clay loam	1.88	26.90	1.32
AOI 5	Navarre	Soft wheat	Wet temperate Mediterranean	894.40 (18 years)	77.61	17.53	20.32	44.60	35.08	Clay loam	2.19	17.61	1.38
AOI 6	Navarre	Soft wheat	Wet temperate Mediterranean	772.60 (5 years)	76.11	7.14	66.44	21.80	11.76	Sandy loam	1.07	0.61	1.46
LUM 1	Navarre	Soft wheat	Wet temperate Mediterranean	529.40 (15 years)	75.84	10.96	13.08	55.38	31.54	Silty clay loam	2.34	12.57	1.15
ABA 1	Navarre	Soft wheat	Fresh Maritime	1310.90 (15 years)	82.23	14.12	17.34	55.13	27.53	Silty clay loam	4.29	30.22	1.11
ABA 2	Navarre	Soft wheat	Fresh Maritime	1310.90 (15 years)	82.23	12.90	24.15	47.26	28.59	Clay loam	3.71	5.07	1.20
ABA 3	Navarre	Soft wheat	Fresh Maritime	1310.90 (15 years)	82.23	19.66	21.30	50.00	28.70	Clay loam	3.81	0.00	1.20
LEO 1	León	Rye	Wet temperate Mediterranean	449.49 (5 years)	87.58	16.59	28.80	45.14	26.10	Loam	1.39	0.80	1.47
RAD 1	Sicily	Durum wheat	Dry temperate Mediterranean	ca. 500 (17 years)	ca. 70	22.59	6.91	55.08	38.01	Silty clay loam	1.11	1.13	1.08
RAD 2	Sicily	Durum wheat	Dry temperate Mediterranean	ca. 500 (17 years)	ca. 70	19.68	12.46	60.75	26.79	Silty loam	1.19	25.91	1.15
RAD 3	Sicily	Durum wheat	Dry temperate Mediterranean	ca. 500 (17 years)	ca. 70	9.66	6.56	42.02	51.42	Silty clay	1.58	6.86	1.03
RAD 4	Sicily	Durum wheat	Dry temperate Mediterranean	ca. 500 (17 years)	ca. 70	21.12	8.43	52.63	38.94	Silty clay loam	1.97	17.91	1.13
RAD 5	Sicily	Durum wheat	Dry temperate Mediterranean	ca. 500 (17 years)	ca. 70	17.52	36.53	42.88	20.59	Loam	1.07	41.78	1.11
RAD 6	Sicily	Durum wheat	Dry temperate Mediterranean	ca. 500 (17 years)	ca. 70	17.36	28.20	24.70	47.10	Clay	1.57	20.93	1.07
RAD 7	Sicily	Durum wheat	Dry temperate Mediterranean	ca. 500 (17 years)	ca. 70	20.44	21.27	50.01	28.72	Clay loam	1.54	64.74	1.22

a: in brackets, the years present in each climate database.
b: in the period from October to May.

1 Table 2. Main characteristics of rainfall causing initiation and growth of the EGs in the soils studied.

2

Soil name	Initiating event			Events from the beginning to the experimentation			Antecedent rainfall to the initiator event				
	Star date	P_{INI} (mm)	R_{INI} (MJ mm ha^{-1} h^{-1})	Experimentation date	P_{TOT} (mm)	R_{TOT} (MJ mm ha^{-1} h^{-1})	aP1 (mm)	aP1d (mm)	aP5d (mm)	aP7d (mm)	aP21d (mm)
PIT 1	10/10/2012	30.80	336.76	05/06/2013	703.99	770.63	0.00	0.12	0.12	0.12	23.28
PIT 2	10/10/2012	30.80	336.76	05/06/2013	703.99	770.63	0.00	0.12	0.12	0.12	23.28
AOI 1	19/10/2012	173.33	1106.37	06/06/2013	1388.90	1890.21	0.00	0.14	10.82	11.66	38.79
AOI 2	19/10/2012	173.33	1106.37	06/06/2013	1388.90	1890.21	0.00	0.14	10.82	11.66	38.79
AOI 3	19/10/2012	173.33	1106.37	22/05/2013	1325.80	1795.14	0.00	0.14	10.82	11.66	38.79
AOI 4	19/10/2012	173.33	1106.37	11/06/2013	1483.80	2022.69	0.00	0.14	10.82	11.66	38.79
AOI 5	19/10/2012	173.33	1106.37	11/06/2013	1483.80	2022.69	0.00	0.14	10.82	11.66	38.79
AOI 6	08/10/2014	33.80	252.34	17/10/2014	57.20	282.60	0.00	0.00	5.00	10.80	57.60
LUM 1	19/10/2012	197.83	1440.01	25/06/2013	1104.80	2437.21	0.00	0.00	6.20	6.33	44.65
ABA 1	19/10/2012	203.74	1049.92	14/06/2013	2065.33	2380.19	0.00	0.00	22.44	23.74	70.61
ABA 2	19/10/2012	203.74	1049.92	14/06/2013	2065.33	2380.19	0.00	0.00	22.44	23.74	70.61
ABA 3	08/10/2014	19.50	99.68	24/10/2014	61.00	141.27	0.00	0.00	1.10	1.40	28.90
LEO 1	25/10/2012	21.11	26.86	20/06/2013	451.77	319.08	0.00	0.20	6.50	11.57	59.14
RAD 1	05/11/2013	17.00	108.63	15/02/2014	231.20	248.83	0.00	0.00	0.00	0.00	0.00
RAD 2	05/11/2013	17.00	108.63	15/02/2014	231.20	248.83	0.00	0.00	0.00	0.00	0.00
RAD 3	05/11/2013	17.00	108.63	14/03/2014	274.40	248.83	0.00	0.00	0.00	0.00	0.00
RAD 4	05/11/2013	17.00	108.63	14/03/2014	274.40	248.83	0.00	0.00	0.00	0.00	0.00
RAD 5	05/11/2013	17.00	108.63	14/03/2014	274.40	248.83	0.00	0.00	0.00	0.00	0.00
RAD 6	05/11/2013	17.00	108.63	14/03/2014	274.40	248.83	0.00	0.00	0.00	0.00	0.00
RAD 7	05/11/2013	17.00	108.63	14/04/2014	292.00	248.83	0.00	0.00	0.00	0.00	0.00

3

Table 3. Main morphological, topographic and soil loss attributes recorded in the studied EGs. n = number of measures, WDR: width and depth relation, AS1 and AS2 = topographical indices, V_T = total volume of eroded soil, TSL = EG erosion rate, and TSLn = normalized EG erosion rate.

Soil name	Total length (m)	Cross sections					Topographic attributes			Soil loss		
		n	Mean area (m^2)	Mean width (m)	Mean depth (m)	WDR	Drainage area (m^2)	AS_1 (m^2)	AS_2 (m^2)	V_T (m^3)	TSL (t ha^{-1} yr^{-1})	TSLn (t MJ^{-1} mm^{-1} h yr^{-1})
PIT 1	25.00	20	0.07	0.28	0.25	1.16	4811.54	457.39	523.97	1.80	5.85	0.008
PIT 2	34.26	14	0.14	0.48	0.30	1.62	6755.91	206.78	287.49	4.64	10.45	0.014
AOI 1	44.30	14	0.03	0.20	0.15	1.37	1239.69	281.90	327.17	1.24	11.48	0.006
AOI 2	40.75	15	0.04	0.53	0.07	8.59	4116.18	710.41	379.68	1.50	4.24	0.002
AOI 3	213.10	36	0.09	0.56	0.15	5.11	13888.46	3536.97	2609.03	18.31	15.88	0.009
AOI 4	25.67	11	0.09	0.57	0.16	4.04	6386.31	792.05	1047.94	2.29	4.73	0.002
AOI 5	64.16	20	0.04	0.38	0.11	4.42	4964.09	869.21	872.71	2.76	7.65	0.004
AOI 6	56.30	24	0.05	0.12	0.40	0.33	19950.28	1544.30	1149.99	2.88	2.09	0.007
LUM 1	49.39	23	0.06	0.44	0.11	4.97	4577.81	502.53	535.33	2.34	7.05	0.003
ABA 1	166.79	59	0.05	0.31	0.14	2.35	7756.45	1097.60	1061.17	6.43	9.23	0.004
ABA 2	76.59	28	0.03	0.29	0.10	3.23	5417.50	699.25	447.42	2.37	5.24	0.002
ABA 3	181.40	17	0.08	0.69	0.12	7.34	37222.24	7394.57	6301.03	13.79	4.44	0.031
LEO 1	77.97	21	0.01	0.21	0.07	3.58	2878.81	477.54	388.35	1.08	5.24	0.016
RAD 1	74.50	12	0.08	0.57	0.20	4.02	8649.18	1952.02	2035.37	5.73	7.03	0.028
RAD 2	61.00	9	0.07	0.61	0.17	4.37	8432.14	1658.42	1472.90	4.35	5.88	0.024
RAD 3	180.00	20	0.11	0.47	0.20	3.35	17665.00	1705.14	2018.94	20.13	12.03	0.048
RAD 4	100.00	20	0.05	0.36	0.14	3.81	4241.31	895.35	785.03	3.79	9.95	0.040
RAD 5	140.00	20	0.14	0.73	0.18	5.12	5970.35	856.87	909.34	18.57	35.33	0.142
RAD 6	110.00	20	0.12	0.83	0.15	6.47	7589.37	1317.79	1328.89	12.89	19.29	0.078
RAD 7	81.00	18	0.11	0.59	0.18	3.43	18532.31	3788.38	3022.79	8.27	5.42	0.022

Table 4. Characterization of soil variables. L = laboratory determination, F_{IN} = field determination on the microplot, F_{OUT} = field determination at points surrounding the microplot, SD = standard deviation. Each determination was repeated 3-5 times at each studied soil, and the average values were considered.

Parameter	Description		Observations	Mean & SD	Unit	Reference
CH_1	Percentage of organic matter (L)		Potassium dichromate and potentiometer	1.98 ± 0.87	%	
CH_2	Electrical conductivity (L)		Conductimeter at 25 ºC	2.67 ± 4.88	dS m^{-1}	
CH_3	Exchangeable sodium percentage (L)		AcNH4. Atomic absorption spectrophotometry	12.67 ± 19.67	%	
CH_4	pH of saturated soil paste (L)		pHmeter. Ratio 1:2.5 (p/v)	7.98 ± 0.45	dimensionless	
CH_5	Cation exchange capacity (L)		AcNa. Atomic absorption spectrophotometry	13.78 ± 3.10	cmol (+) kg^{-1}	
CH_6	Calcium carbonate (L)		Bernard calcimeter	21.30 ± 17.66	%	
CR_1	Crusting susceptibility index $C_{5\text{-}10}$ (L) $\frac{1}{W_5 - W_{10}}$		W_5 and W_{10} are the moisture contents in which an incision made on soil paste in a Casagrande's scoop is closed after 5 (W_5) and 10 (W_{10}) blows on the gasket	0.47 ± 0.21	dimensionless	De Ploey & Mücher (1981)
CR_2	Relative sealing index (L) $\frac{KSS}{KCS}$ *KSS*: soil hydraulic conductivity before rainfall, mm h^{-1} *KCS*: soil hydraulic conductivity after rainfall, mm h^{-1}		Variation in the permeability of a layer of aggregates (2-4 mm) before and after being altered by controlled rainfall	24.16 ± 22.51	dimensionless	Pla (1982)
CR_3	Crusting indices (L)	$\frac{\%\text{ silt} + \%\text{ fine sand} + \%\text{ very fine sand}}{\%\text{ clay}}$	Particle separability index	2.38 ± 1.31	dimensionless	Florentino (1998)
CR_4		$\frac{0.550 \cdot (\%\text{ silt} + \%\text{ fine sand} + \%\text{ very fine sand})}{6.743 \cdot (\%\text{ organic matter})}$	Sealing-crusting index	3.12 ± 1.51	dimensionless	
CR_5		$\frac{(1.500 \cdot \%\text{ coarse silt}) + (0.750 \cdot \%\text{ fine silt})}{\%\text{ clay} + 10 \cdot (\%\text{ organic matter})}$	Crusting index	0.97 ± 0.28	dimensionless	FAO (1980)
CR_6		$\frac{(1.125 \cdot \%\text{ fine silt})}{\%\text{ clay} + 10 \cdot (\%\text{ organic matter})}$	Modified crusting index	0.87 ± 0.27	dimensionless	Comerma et al. (1992)
E_1	Soil erodibility coefficient (L)		Jet Test. Based on measuring degree of incision on a soil sample caused by a waterjet	97.64 ± 70.59	cm^3 $(N\ s)^{-1}$	Hanson & Cook (2004)
E_2	K factor of RUSLE (L) $\left[\frac{2.1x10^{-4}\ M^{1.14}(12 - OM) + 3.3(s-2) + 2.5(p-3)}{100}\right] x0.13$ *M*: (% silt + % very fine sand) · (100 - % clay) *OM*: % organic matter *s*: structure class of the soil *p*: permeability class		The permeability classes (p) were defined according to infiltration values measured in the field (HY_1); and the structure classes (p) according to indices defined in the laboratory (SI_4, SI_5, SI_6)	0.036 ± 0.009	t ha h ha^{-1} MJ^{-1} mm^{-1}	Renard *et al*. (1997)
GC_6	Percentage of coarse fragments (> 2mm) (L)			16.91 ± 16.77	%	

GC_{1-5}	Percentiles of the coarse fraction (> 2mm) (L)	Sub-indices 1 to 5 equal to the percentiles 10 [(1)], 38 [(2)], 50 [(3)], 60 [(4)] and 60/10 [(5)], the latter based on Terzaghi and Peck (1969); obtained by dry sieving and discontinuous sedimentation	[(1)] 5.33 ± 5.15 [(2)] 20.22 ± 22.80 [(3)] 26.83 ± 30.47 [(4)] 33.65 ± 37.24 [(5)] 6.06 ± 4.11	mm	
GF_{1-5}	Percentiles of total granulometry (L)		[(1)] 0.0006 ± 0.0007 [(2)] 0.27 ± 1.22 [(3)] 0.58 ± 2.59 [(4)] 1.58 ± 4.66 [(5)] 762.59 ± 2071.86	mm	
GT_{1-5}	Percentiles of fine fraction (< 2mm) (L)		[(1)] 0.0004 ± 0.0002 [(2)] 0.008 ± 0.017 [(3)] 0.014 ± 0.024 [(4)] 0.022 ± 0.031 [(5)] 47.75 ± 29.97	mm	
HY_1	Basic infiltration rate of soil (F_{IN})	Rainfall simulation	47.54 ± 35.50	mm h^{-1}	
HY_2	Hydraulic conductivity of seal (L)	Rainfall simulation	2.30 ± 1.06	mm h^{-1}	Pla (1982)
HY_3	Hydraulic conductivity of crust (F_{IN}) $\frac{4\,V}{\pi\,t\,D^2}$ V: volume of water of the calibration, mm^3 t: total time of water discharge, h D: diameter of wet halo, mm	Based on the measurement of the infiltration rate of a portion of crust previously saturated (=wet halo) (gravitational flow)	2.19 ± 0.87	mm h^{-1}	Boiffin & Monnier (1985)
PH_1	Bulk density (F_{OUT})		1.25 ± 0.16	g cm^{-3}	
PR_1	Crust resistance to penetration (F_{IN})	Pocket penetrometer	276.61 ± 88.24	kPa	Bradford *et al*. (1992)
PR_2	Penetration resistance in the first 3 centimeters of the soil depth (F_{IN})	Digital cone penetrometer	572.72 ± 347.34	kPa	Truman & Bradford (1990)
PR_3	Penetration resistance in the first 6 centimeters of the soil depth (F_{IN})	Digital cone penetrometer	640.56 ± 386.72	kPa	
SI_1	Structure stability indices (L): % silt + % very fine sand		53.93 ± 9.25	dimensionless	Bouyoucos (1935) (original and modifications)
SI_2	Structure stability indices (L): $\frac{\%\text{ silt} + \%\text{ very fine sand}}{\%\text{ clay} + \%\text{ organic matter}}$		1.80 ± 0.52	dimensionless	
SI_3	Structure stability indices (L): $\frac{\%\text{ silt} + \%\text{ very fine sand}}{\%\text{ clay}}$		1.94 ± 0.58	dimensionless	
SI_4	Aggregate stability indices (L) $DMP = \sum_{i=1}^{n} \bar{x}_i w_i$ *DMP*: weighted mean diameter of aggregates, mm $\bar{x}_i$: mean diameter of different sized groups of aggregates, mm [(1)] w_i: % of weight of each group of aggregates with respect to the total weight	The DMP in the soil aggregates (3-5 mm) submitted to 3 disaggregation mechanisms: slaking caused by fast wetting (SI_4), clay swelling caused by slow wetting (SI_5), and mechanical breakdown by shaking (SI_6) [(1)] Soil fractions were obtained by sieving	0.85 ± 0.53	mm	Le Bissonnais (1996)
SI_5			1.97 ± 0.91	mm	
SI_6			1.49 ± 0.79	mm	

SL_1	Interrill erosion rate (F_{IN})		Rainfall simulation	6.30 ± 5.52	t ha^{-1} h^{-1}	
SS_1	Shear strength (F_{OUT})	d = 330 mm; p = 100 mm	Vane shear apparatus For the variable SS_1 a disk was constructed *ad hoc* with a high depth (p)/diameter (d) ratio. Measurements were taken before wetting the soil up to its saturation	15.10 ± 4.02	kPa	Léonard & Richard (2004)
SS_2		d = 470 mm; p = 40 mm		10.62 ± 2.26	kPa	
SS_3		d = 250 mm; p = 40 mm		20.52 ± 5.73	kPa	
SS_4		d = 190 mm; p = 40 mm		23.32 ± 8.20	kPa	
SS_5	Critical shearing (L)		Jet Test. Based on measuring the degree of incision on a soil sample caused by a waterjet	2.19 ± 1.14	Pa	Hanson & Cook (2004)

1

Table 5. Main variables that define the soils grouped in the 2 clusters identified after the execution of the cluster analysis. Order of the variables according to the lowest probability value (p-value). SD = standard deviation.

Variables	Cluster A			Cluster B			All dataset	
	Mean in cluster	SD cluster	p-value [a]	Mean in cluster	SD cluster	p-value [a]	Overall mean	SD overall
CR_4	**2.14**	**0.80**	**0.0007**	**4.57**	**1.13**	**0.0007**	**2.99**	**1.49**
CR_5	**0.77**	**0.19**	**0.0008**	**1.24**	**0.18**	**0.0008**	**0.94**	**0.29**
CH_5	15.85	1.60	0.0014	11.15	2.76	0.0014	14.21	3.06
CR_1	0.37	0.11	0.0025	0.67	0.20	0.0025	0.48	0.21
SI_2	1.50	0.41	0.0037	2.18	0.25	0.0037	1.74	0.49
E_2	0.03	0.01	0.0065	0.04	0.01	0.0065	0.04	0.01
CH_3	4.00	3.13	0.0071	30.39	26.78	0.0071	13.24	20.39
SI_3	1.63	0.48	0.0082	2.31	0.29	0.0082	1.87	0.53
CH_1	2.49	0.94	0.0092	1.29	0.23	0.0092	2.07	0.96
GT_1	0.00	0.00	0.0101	0.00	0.00	0.0101	0.00	0.00
CR_3	1.73	0.46	0.0109	3.27	1.60	0.0109	2.27	1.25
PH_1	1.17	0.09	0.0129	1.35	0.18	0.0129	1.23	0.16
CH_2	0.92	0.74	0.0140	6.87	6.95	0.0140	3.00	5.03
SI_6	1.89	0.81	0.0148	0.89	0.45	0.0148	1.54	0.85
GF_1	0.00	0.00	0.0215	0.00	0.00	0.0215	0.00	0.00
SS_4	21.48	6.00	0.0223	14.32	4.59	0.0223	18.97	6.52
CR_6	0.75	0.21	0.0274	1.02	0.24	0.0274	0.85	0.26
HY_2	2.61	0.87	0.0422	1.76	0.52	0.0422	2.31	0.86
GF_4	0.01	0.00	0.0471	0.04	0.04	0.0471	0.02	0.03

a: in bold the variables with a p-value less than 0.001.

Table 6. Principal components (PCs) with an eigenvalue greater than 1 and factorial correlation coefficients of the variables with the first two PCs for the 20 soils.

Number of components	Eingenvalue	% Variance	% Accumulated variance
1	11.38	20.69	20.69
2	8.63	15.68	36.38
3	6.86	12.47	48.85
4	5.96	10.83	59.67
5	5.81	10.57	70.24
6	3.63	6.60	76.84
7	2.91	5.28	82.13
8	2.10	3.82	85.94
9	1.62	2.95	88.90
10	1.41	2.57	91.47
11	1.15	2.08	93.55

Active variables – Factorial correlation [a]

Variable	PC_1	PC_2
aP21d	0.31	**0.81**
aP5d	-0.04	**0.90**
aP7d	0.12	**0.89**
CH_1	-0.28	**0.77**
CH_5	**-0.72**	-0.07
CR_2	-0.13	**-0.50**
CR_3	**0.64**	0.03
CR_4	**0.95**	-0.44
GF_1	**0.95**	0.00
GF_2	**0.96**	-0.02
GF_3	**0.98**	-0.01
GF_4	**0.99**	-0.01
GF_5	**0.80**	0.08
HY_3	-0.21	**0.60**
PR_1	0.23	**0.50**
SI_1	**-0.50**	0.20
SI_2	**0.54**	0.12
SI_3	**0.54**	0.20
SI_4	-0.12	**0.76**
SI_5	-0.02	**0.89**
SI_6	-0.04	**0.91**
SS_3	-0.18	**0.57**
Variable supplementary – Factorial correlation		
TSLn	**0.14**	**-0.37**

a: factors in bold show a factor correlation greater than 0.50.

Table 7. Best linear regression models explaining TSLn with one, two, three and four dependent variables. Different letters show different levels of probability significance (p) for regression coefficients.

Variables	Regression coefficients					NSE	AIC	MSE	VIF			
	Intercept	1	2	3	4				1	2	3	4
E_1	-0.03 a	4.80E-04 a	-	-	-	0.78	-103.73	2.43E-04	-	-	-	-
E_1, CR_2	-0.03 a	4.19E-04 a	3.86E-04 b	-	-	0.83	-107.18	1.84E-04	1.23	1.23	-	-
E_1, CR_2, SS_3	-0.02 b	3.87E-04 a	4.44E-04 b	-7.03E-04 b	-	0.85	-107.25	1.67E-04	1.18	1.44	1.32	-
E_1, CR_2, SS_3, CH_5	-0.03 b	3.82E-03 a	4.73E-04 b	-1.17E-04 b	1.82E-03 b	0.87	-108.44	1.42E-04	1.57	1.45	1.34	1.36

a: $p<0.005$.
b: $p<0.05$.

Table 8. Range of values in the 2 clusters for the 13 key variables in the control of erosion by EGs after the accomplishment of the 3 multivariate statistical approaches. Min = minimum value for the variable in the cluster, Max = maximum value for the variable in the cluster, SD = standard deviation, CV = coefficient of variation.

Variable	Units	Cluster A					Cluster B				
		Min	Mean	Max	SD	CV	Min	Mean	Max	SD	CV
aP5d	mm	0.00	8.18	22.44	7.98	0.98	0.00	1.68	6.50	2.82	1.68
CH_1	%	1.11	2.49	4.29	0.98	0.29	1.07	1.29	1.67	0.25	0.19
CH_5	cmol (+) kg^{-1}	13.00	15.85	19.12	1.67	0.11	6.38	11.15	16.03	2.98	0.27
CR_2	dimensionless	3.41	23.50	72.76	20.11	0.86	9.27	35.02	72.58	26.55	0.76
CR_4	dimensionless	1.15	2.14	4.47	0.83	0.39	3.09	4.57	6.34	1.22	0.27
CR_5	dimensionless	0.41	0.77	1.15	0.19	0.25	0.93	1.24	1.52	0.19	0.16
E_1	cm^3 $(N\ s)^{-1}$	50.50	102.53	176.33	38.26	0.30	55.58	110.29	320.28	95.88	0.87
HY_3	mm h^{-1}	1.87	2.61	5.40	0.90	0.42	0.33	1.64	2.76	0.87	0.53
PR_1	kPa	85.81	263.18	353.04	94.16	0.36	160.58	280.72	367.75	75.00	0.27
SI_4	mm	0.45	1.06	2.13	0.60	0.57	0.16	0.60	0.95	0.30	0.51
SI_5	mm	0.51	2.19	3.16	1.00	0.46	0.72	1.49	2.64	0.63	0.43
SI_6	mm	0.56	1.89	3.06	0.85	0.45	0.15	0.89	1.66	0.48	0.54
SS_3	kPa	10.66	21.48	32.97	6.24	0.29	7.48	14.32	19.67	4.96	0.35

Figures

Figure 1. Approximate location and some examples of the 20 ephemeral gullies (EGs) analyzed: (A) general view of the study areas, (B) León, northwest of Spain (1 EG), (C) Navarre, north of Spain (12 EGs), (D) Sicily, south of Italy (7 EGs), (E) ephemeral gully PIT 2, and (F) ephemeral gully AOI 3.

Figure 2. Best linear regression explanatory model of TSLs with one dependent variable (E_1).

Figure 3. Results obtained after applying the FITEVAL technique to the best linear explanatory of 4 variables. The mean values and the range of variation for the statistics, the Nash-Sutcliffe (NSE) coefficient, and the root mean square error (RMSE) are shown at the top on the left. Bottom right, the evaluation, the outliers and the model bias can be seen.

Figure 4. Distribution of the guide values for the 13 key variables identified, obtained from the mean values in the 2 clusters identified. Two areas have been delimited: lesser erodibility by EGs (green area) and greater erodibility due to EGs (red area). The variables CR_2, CR_4, CR_5, E_1 and PR_1 are positively correlated with the rate of erosion by EGs, whereas the variables aP5d, CH_1, CH_5, HY_3, SI_4, SI_5, SI_6 and SS_3 are negatively correlated with the EG erosion rate (e.g a value of approximately below 1.3% for CH_1 would imply a high soil erosibility level).

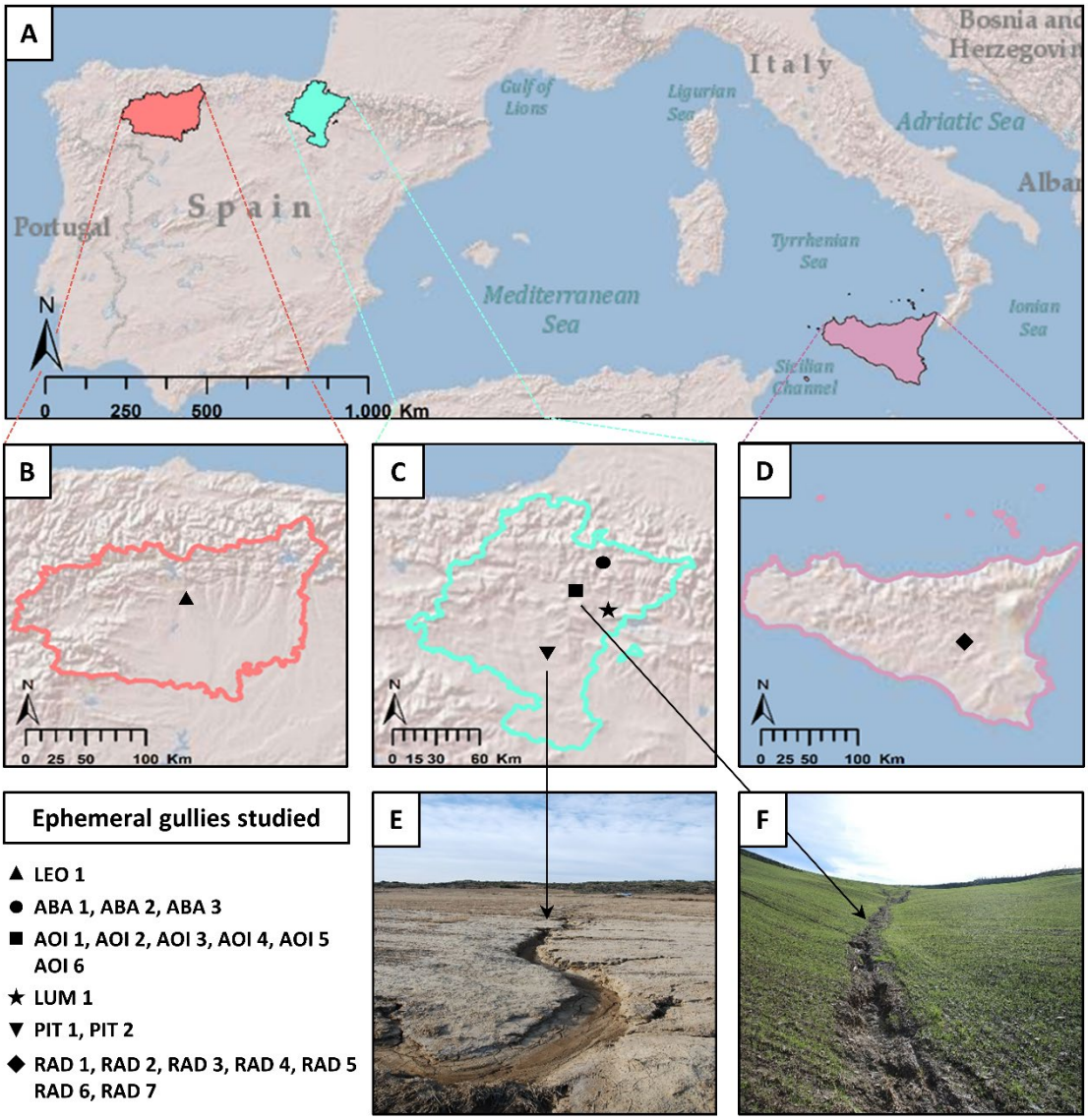

Figure 1. Approximate location and some examples of the 20 ephemeral gullies (EGs) analyzed: (A) general view of the study areas, (B) León, northwest of Spain (1 EG), (C) Navarre, north of Spain (12 EGs), (D) Sicily, south of Italy (7 EGs), (E) ephemeral gully PIT 2, and (F) ephemeral gully AOI 3.

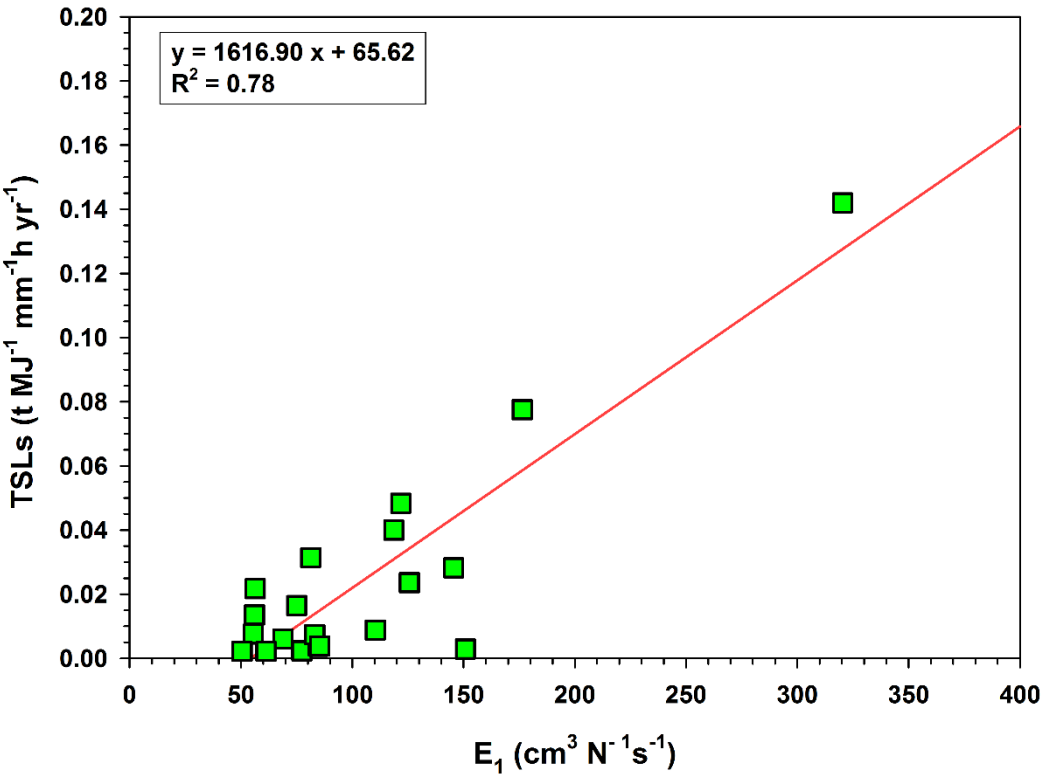

Figure 2. Best linear regression explanatory model of TSLs with one dependent variable (E_1).

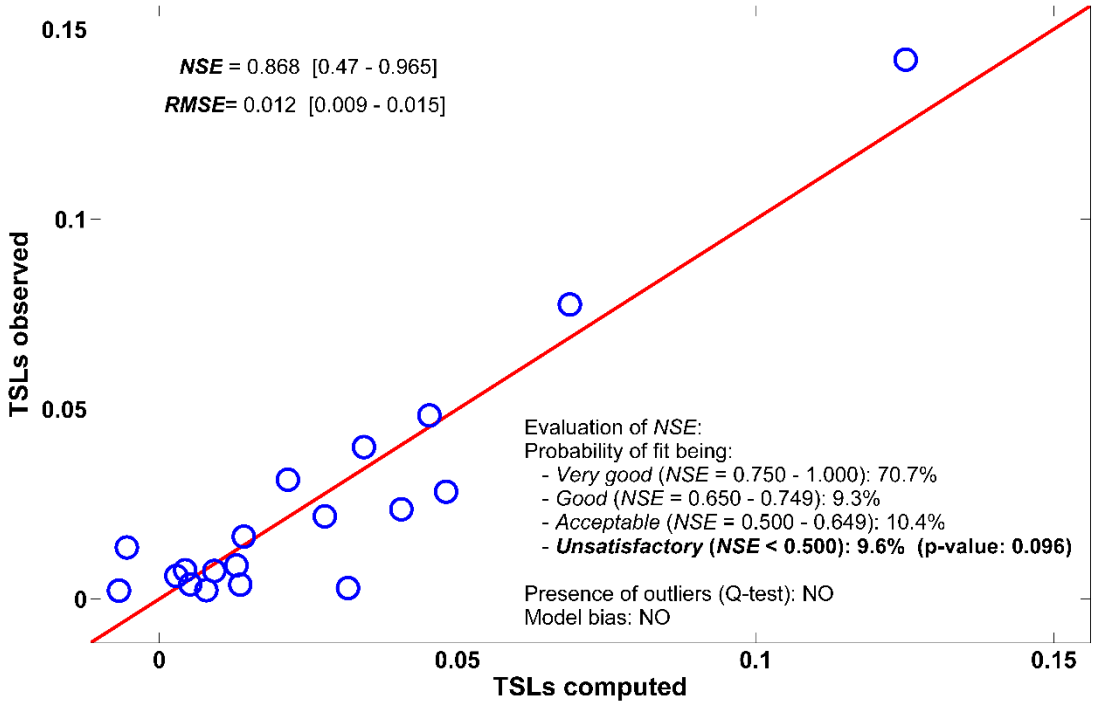

Figure 3. Results obtained after applying the FITEVAL technique to the best linear explanatory of 4 variables. The mean values and the range of variation for the statistics, the Nash-Sutcliffe (NSE) coefficient, and the root mean square error (RMSE) are shown at the top on the left. Bottom right, the evaluation, the outliers and the model bias can be seen.

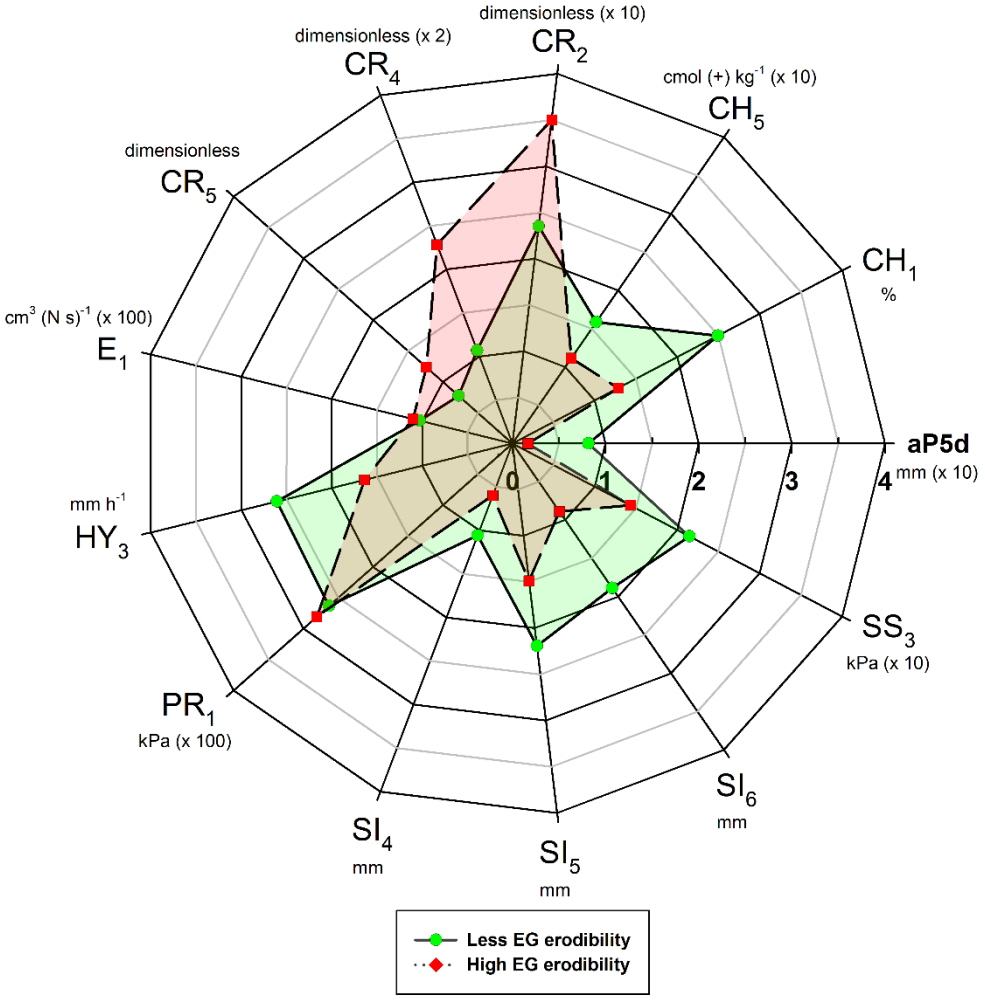

Figure 4. Distribution of the guide values for the 13 key variables identified, obtained from the mean values in the 2 clusters identified. Two areas have been delimited: lesser erodibility by EGs (green area) and greater erodibility due to EGs (red area). The variables CR_2, CR_4, CR_5, E_1 and PR_1 are positively correlated with the rate of erosion by EGs, whereas the variables aP5d, CH_1, CH_5, HY_3, SI_4, SI_5, SI_6 and SS_3 are negatively correlated with the EG erosion rate (e.g a value of approximately below 1.3% for CH_1 would imply a high soil erosibility level).