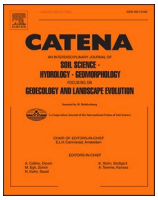
Catena 263 (2026) 109699

Contents lists available at ScienceDirect

Catena

journal homepage: www.elsevier.com/locate/catena



What drives gullies in Spain's olive landscapes? A regional analysis of gully activity

Paula González [a,*], Adolfo Peña [a], Sofie De Geeter [b,c,d], Matthias Vanmaercke [b], Jean Poesen [b,e], Tom Vanwalleghem [f]

[a] University of Cordoba, Department of Rural Engineering, Civil Engineering and Engineering Projects. ETSIAM. Edif. Leonardo Da Vinci, Ctra. Madrid km 396, 14071 Córdoba, Spain
[b] KU Leuven, Division of Geography and Tourism, Department of Earth and Environmental Sciences, Celestijnenlaan 200E, 3001 Heverlee, Belgium
[c] University of Liège, Department of Geography, Clos Mercator 3, 4000 Liège, Belgium
[d] Research Foundation Flanders – FWO, Brussels, Belgium
[e] Institute of Earth and Environmental Sciences, UMCS, Krasnicka Av. 2d, 20-718 Lublin, Poland
[f] University of Cordoba, Dept. of Agronomy, da Vinci Bldg., Cra Madrid km 396, 14071 Córdoba, Spain

ARTICLE INFO

Keywords:
Gully head
Erosion
(critical) Shear stress index
Topographic threshold (TT)
Gully head initiation index (GHI)
Guadalquivir basin
Olive groves

ABSTRACT

Gully erosion poses a significant threat to Mediterranean basins. However, precise spatial data on gullies location at regional scale is scarce. Current models often use the topographic threshold index, TT, which works locally and predicts gully location based on slope and drainage area but does not distinguish between actively retreating and stabilized gullies. We aim to close this gap by (i) constructing a dataset of gully head location and activity in olive groves in southern Spain across different landscape regions and (ii) evaluate a process-oriented model to distinguish gully head presence and activity based on the Gully Head Initiation (GHI) index. This GHI integrates slope, drainage area, precipitation, curve number (CN), soil type, and clay content to predict gully head formation.

Photointerpretation of four 25 km^2 study areas between 2008 and 2019, 3 gully head activity classes were identified, with 261 existing active, 76 new active and 138 stable gully heads identified. The GHI clearly differentiates gully head from non-gully head pixels (AUC = 0.93). Furthermore, it distinguished between existing active and stable gully heads, as well as new active and existing active ones. However, there are no significant differences between stable and new active gully heads. Besides, applying the model separately by landscape type is not as effective as applying it across all landscapes.

These results highlight the potential applicability of GHI at the regional scale and confirm that the factors considered in the GHI index affect both gully head initiation processes and erosion dynamics.

1. Introduction

Soil erosion represents a major threat to food security and soil health, critical for many ecosystem services. At present, 62 % of EU soils are unhealthy (Panagos et al., 2024), with soil erosion being one of the main degradation processes. As Borrelli et al. (2022) have reported, tolerable soil losses are exceeded in half of the arable land in the EU. These issues are a priority on the European political agenda, and an ambitious package of measures to promote soil health was put forward across different policies, such as the 2023–2027 Common Agricultural Policy, the Biodiversity 2030, Farm to Fork and Chemical Strategies (Panagos et al., 2022). In 2024, a proposal for a new Soil health monitoring and resilience Directive was adopted by the Council. The obligatory monitoring of soil health aims to reverse the current situation and ensure that all soils are in a healthy state by 2050. Similarly, the EU Nature Restoration Act, which has been agreed upon by all Member States, aims to regenerate degraded ecosystems within their territories. This initiative is intended to contribute to the achievement of EU climate and

* Corresponding author at: University of Cordoba, Department of Rural Engineering, Civil Engineering and Engineering Projects. ETSIAM. Edif. Leonardo Da Vinci, Ctra. Madrid km 396, 14071 Córdoba, Spain.
E-mail address: z82gogap@uco.es (P. González).

https://doi.org/10.1016/j.catena.2025.109699
Received 9 June 2025; Received in revised form 11 November 2025; Accepted 25 November 2025
Available online 3 December 2025

biodiversity objectives, as well as enhance food security. The Act sets a target of restoring at least 20 % of its land and marine areas by 2030, and all ecosystems that require it, by 2050 (European Parliament, 2024).

In order to design soil health management strategies, it is necessary to identify erosion hotspots and processes. However, one important erosion process, gully erosion, is not yet being adequately addressed in current regional-scale soil erosion assessments (Vanmaercke et al., 2021). Extensive research has been carried out on the modelling of soil erosion rates, mostly using the RUSLE (Revised Universal Soil Loss Equation) model (Tan et al., 2021; Panagos et al., 2015). This model only considers sheet and rill erosion. Other studies using data-based approaches in combination with Deep Learning models (Shojaeezadeh et al., 2024), have also identified this limitation. Current estimates therefore underestimate both soil loss and potentially the location of erosion hotspots. Semi-arid areas of the Mediterranean, and specifically olive groves, are environments highly susceptible to gully erosion (Poesen et al., 2002; Hayas et al., 2017a,b). In order to ensure the effective management of soil health, strategies must be designed at the watershed or broader regional level. These strategies must include prediction of gully initiation and activity.

For gully initiation prediction, one important problem is the regional variability. While many studies have successfully modelled gully initiation locations with the topographic threshold model (TT) (Vandaele et al., 1996; Vandekerckhove et al., 1998, 2000; Gómez-Gutiérrez et al., 2009; Imaizumi et al., 2010; Torri & Poesen, 2014), most of these studies are limited to the local scale (Vanmaercke et al., 2021). Applications at regional, country or continental scale are limited. Hayas et al. (2017b), in their study on gully heads in a Mediterranean olive grove catchment, has shown that the TT varies from year to year, and was directly related to dynamics in precipitation and vegetation cover. Therefore, in larger areas where precipitation, vegetation cover, or soil properties vary, the TT cannot be considered constant. Statistical approaches such as machine learning could offer an alternative (Li et al., 2024) but are still limited to the extent of the training dataset and extrapolation is difficult. Finally, more process-based models, such as LANDPLANER (Rossi, 2014) or the Gully Head Initiation (GHI) index model (De Geeter et al., 2025), show great potential to overcome this limitation. Such models calculate runoff directly using the curve number method, instead of using catchment area as a proxy. The GHI (Gully Head Initiation) index, which is based on the ratio between the shear stress index (SSI) and the critical shear stress index (CSI), is then linked to gully head initiation. De Geeter et al. (2023) used this approach at continental and national scales across the entire African continent, as well as at the local catchment scale in the Ethiopian highlands. A similar approach was proposed in LANDPLANER, using an erosion index calculated based on runoff and slope, which is then linked to gully initiation. (Kariminejad et al., 2023) tested this in a small 5,2 km^2 catchment in Iran.

Such a process-based approach offers several advantages. First, in theory, it offers a more universal model that can be easily applied to a wide range of different study areas, as the spatial variability in soil properties, vegetation and rainfall can be explicitly considered in the runoff calibration. This would eliminate the need for calibrating with new gully head data, as is the case for the TT model. However, the current study of De Geeter et al. (2023) in Africa yielded a fairly weak model efficiency (AUC of 0.67). Of course, this is partly due to the complexity of modelling runoff in ungauged basins (Hrachowitz et al., 2014), which could be improved by using more advanced rainfall-runoff models and rainfall input data. Therefore, further testing is required in different areas, and the performance of both models (GHI vs. TT) must be compared directly. Secondly, a process-based approach such as the GHI index could, in theory, be used to model not only gully head initiation, but also gully head activity. So far, the GHI index has only been applied to gully heads derived from a single mapping time, and not to a time series of gully heads. This implies that gully heads include both active and older, non-actively retreating ones. Only a few studies have modeled gully activity. Fan et al. (2024) modelled gully activity using fuzzy classification in three small study areas in the Sunshui River basin in China and evaluated the controlling factors. The potential of the GHI index to distinguish between active and stable gully heads remains largely unexplored, and this would be a major step forward in modelling the contribution of gully erosion to overall soil erosion and in identifying active erosion hotspots. Finally, such models facilitate scenario analysis, for example considering the effect of changing rainfall patterns (Kariminejad et al., 2023), land use, or the effect of soil conservation measures, including cover crops that mitigate runoff.

The main objective of this study is therefore to improve the prediction of gully head initiation and activity at a regional scale by testing and applying the GHI index to four representatives olive grove landscapes in the Guadalquivir basin (Southwestern Spain). Linked to this, our specific objectives are:

- Mapping of gully heads in the study areas and classification of gully activity from 2008 to 2019.
- Evaluation of the potential of the GHI index to differentiate between newly active, existing active and stable gully heads, as well as points that are not gully heads.
- Evaluation of regional and rainfall differences on GHI index.

2. Materials and methods

2.1. Study areas

The study area is located in the Guadalquivir basin, that covers an area of 57,527 km^2 (Fig. 2.a). The Guadalquivir basin is bordered by the Sierra Morena to the north, the Betic mountain ranges, located to the south with a SW-NE development and the Atlantic Ocean. The mountainous edge that delimits the space, with elevations between 1,000 m and 3,480 m a.s.l., contrasts with the low altitude of the wide valley of the Guadalquivir river. The climate is Mediterranean with dry and hot summers, Csa in the Köppen-Geiger scheme (Peel et al. 2007). The temperature is warm (16.8 °C as annual average) and the rainfall is irregular (annual average of 550 mm). The rain events are frequently torrential with daily maximum depths up to 213.8 mm (Table 1). In addition, soil has a marked susceptibility to erosion with crusting after long periods of drought. In the lowlands, clayey soils, like Vertisols, Luvisols predominate. In the steep, mountainous areas sandy soils with a reduced degree of formation, Cambisols or Regosols, are common. Very often, decades of intensive erosion have generated Cambisols or Leptosols in the agricultural valley areas.

In this study, we wanted to focus on the effect of regional landscape variability on gully head formation, so we limited the land use analyzed to olive groves to eliminate the effects of land use variability. Olive is the main crop in the Guadalquivir basin, covering over 1,7 million ha. The area dedicated to olive groves for oil extraction is increasing at a rapid rate, with a growth in area of more than 100,000 ha since 2008 (Ministerio de Agricultura, Pesca y Alimentación, 2024). Unfortunately, this crop is very affected by gully erosion (Hayas et al., 2017a,b).

The regional landscape variability is based on the classification by the Andalusian Environmental Information Network (Junta de Andalucía, 2018). This classification marks transitions between categories or

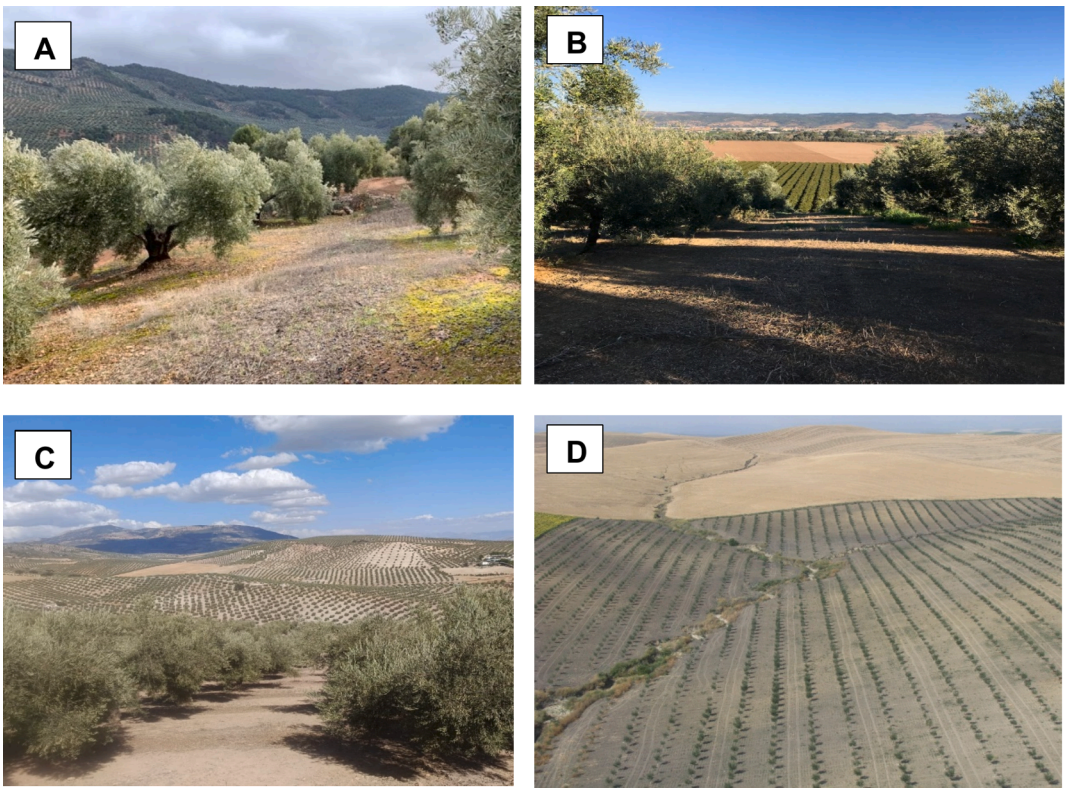


Fig. 1. Photographs of the main landscape units (a) Mid mountains (b) Valley plains (c) Countryside foothills (d) Countryside hills.

geographical situations based on morphological, vegetation cover, or land-use characteristics. The olive growing areas of the Guadalquivir basin cover four different landscapes units: Mid-Mountains, Valley Plains, Countryside Foothills and Countryside hills, which, as we can see in Fig. 2.b. We then randomly selected four 25 km^2 study areas, one for each of the main landscape units (Junta de Andalucía, 2018) (Fig. 2.c), to map gully head features in detail.

Below are some characteristics of each landscape unit and the corresponding study area (Junta de Andalucía, 2018). To better illustrate the differences among these areas, Fig. 1 shows representative photographs of each landscape unit.

2.1.1. *Countryside hills*

Up to 31 % of the territory of the Autonomous Community of Andalusia is occupied by the Campiñas landscape, composed mainly of the hills (more than 17 %) and the surrounding lowlands. These are the traditional agricultural areas par excellence of Andalusia.

Specifically, the study area is located in Campiñas Altas landscape unit, which includes the highest areas of the Seville, Cordoba, and Jaén provinces. Most of this landscape is used for agricultural exploitation, offering a very homogeneous, robust and specialized landscape, considered both within and outside the region to be one of the most representative of Andalusia. Its distinctive character is essentially due to the vast expanses of olive groves, which comprise 73 % of the area, while other dry and irrigated crops have been relegated to smaller portions.

2.1.2. *Countryside foothills*

Distributed to the north and south of the Baetic Depression and comprised of hills, ridges, and sporadic mountain ranges, they occupy the spaces between the depression and the Sierra Morena and the Subbaetic Mountain Range.

Specifically, the study area corresponding to this landscape unit is located in the Subbaetic foothills. Its territory is dominated by low-lying relief forms, hills, and ridges, punctuated by more prominent features, with a strong variation in altitude. Its conditions favor dryland crops, among which olive groves stand out, along with areas of scrubland and grassland.

2.1.3. *Mid-Mountains*

Andalusia contains a significant mountainous area, a 44 % of the total surface.

Specifically, the study area of this landscape unit is located in the Eastern mountains, in the Northern Granada province. The zone has a gentle relief, with elevations from 560 m. Most of the area is dedicated to dry-framing agriculture, with olive groves and herbaceous crops.

2.1.4. *Valley plains*

This landscape unit covers just over 10 % of the territory.

The study area is within the Antequera Depression. The medium–high fertility of the soils and the abundant water resources have promoted an appreciable agricultural production, supported by olive groves, along with irrigated and dryland herbaceous crops.

Table 1 shows the most detailed characteristics of each study area analyzed in the study for each landscape unit. These data were obtained from the REDIAM (Andalusian Environmental Information Network) and from weather stations near the study area.

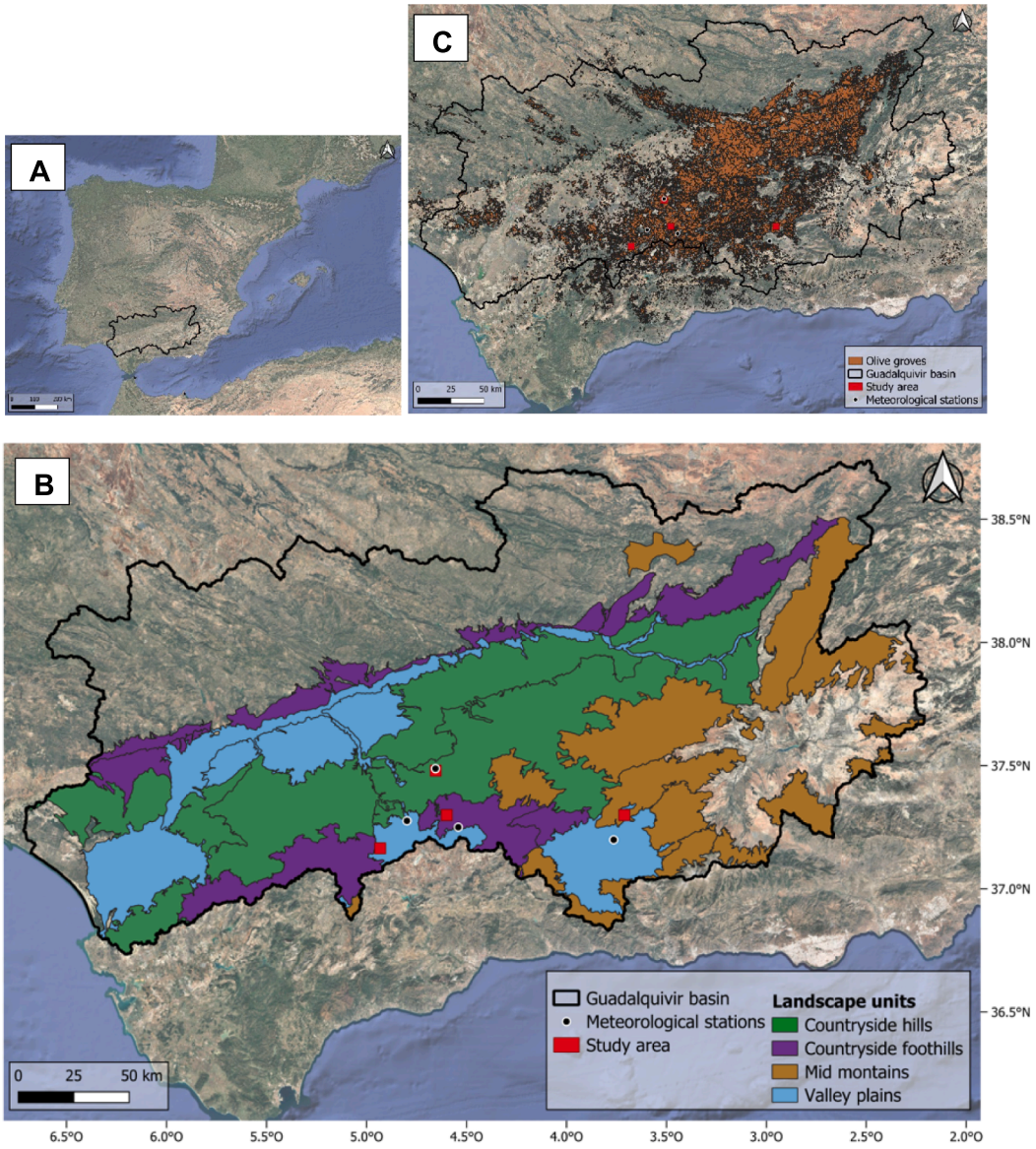


Fig. 2. (a) Guadalquivir basin in Andalusia (Spain) (b) study area, one within each of the principal landscapes units occupied by olive groves and the meteorological stations (c) area occupied by olive groves in Andalusia (SIOSE).

Table 1
Summary characteristics of each study area.

Landscape unit	Countryside hills	Countryside foothills	Mid-Mountains	Valley plains
Slope (m/m)	0.03–0.12	0.05–0.37	0.05–0.31	0.04–0.28
Organic matter (%)	1.05	1.23	0.98	1.29
CN	83	85	83	83
Average annual rainfall (mm)	564.9	468.3	434.5	591.95
Clay Content (% ± SD)	34.10 ± 2.97	36.46 ± 2.26	29.92 ± 2.01	33.71 ± 2.81
Sand Content (% ± SD)	29.94 ± 6.31	39.13 ± 3.65	42.04 ± 3.38	41.88 ± 4.12
Silt Content (% ± SD)	35.96 ± 6.85	24.35 ± 1.79	28.16 ± 1.99	24.52 ± 1.92
Textural class	Clay loam	Clay loam	Clay loam	Clay loam
Lithology	Calcarenites, sands, marls, and limestones	Gypsiferous marls, sandstones, and limestones	Conglomerates, sands, shales, and limestones	Sandstones and limestones
Elevation (m a.s.l.)	302.3–344.2	389.8–423.7	692.8–752.9	370.5–429.2
Coordinates	37.49°N, 4.71°W	37.25°N, 3.58°W	37.19°N, 2.32°W	37.29°N, 5.02°W
Max Precip. 2008–2010 (mm/day)	83.4	37.2	41.8	65.8
Max Precip. 2010–2013 (mm/day)	213.8	76.8	52.6	96.6
Max Precip. 2013–2019 (mm/day)	56.2	38	52.1	79.4

2.2. Generation of DEM, A & S values

The 2x2 m Digital Elevation Model (DEM) of each study area has been generated from the 3D point cloud data from LiDAR of the PNOA (National Plan for Aerial Orthophotography). The coverage for Andalusia has a density equal to 0.5 points/m^2 corresponding to the year 2014, using a Leica ALS60 sensor for this purpose, with the planimetric and altimetric mean square error being equal to 0.3 and 0.2 m respectively. From the classification of the LiDAR point cloud, those corresponding to class 2 (points belonging to bare soil) were filtered for the generation of the DEM. Then, using the QGIS software and the IDW (inverse distance weighting) interpolation method, the generated 2x2 m DEM has been processed and, for each pixel, the drainage or flow accumulation area, A (m^2) and the local slope, S (m/m) have been calculated. The calculation of A has been carried out using the D8 algorithm. The D8 algorithm, as described in the work of (O'Callaghan and Mark, 1984), assigns the water flow from each pixel to one of its eight adjacent neighbors in the direction of the greatest downward slope. This approach has been widely validated in various studies and has been shown to provide accurate estimates of local accumulation areas and slopes, becoming a fundamental tool in geomorphological analysis and hydrological modelling within the QGIS environment (Tarboton, 1997; Quinn et al., 1991; Jones, 2002).

2.3. Mapping of gully heads and gully activity

For each of the 25 km^2 study areas, all the gully heads have been photo-interpreted from PNOA orthophotos. We have mapped the position of gully heads in imagery for the years 2008, 2010, 2013 and 2019.

In this way, not only the current gully network but also their activity over these periods could be evaluated. Gully erosion activity was subdivided into three classes, and data was obtained for three periods (2008–2010, 2010–2013, 2013–2019), as illustrated in Fig. 3. The definition of gully activity only is based exclusively on the movement of the gully head, which is defined as the upslope termination point of the gully. To summarise, we define an active gully as a gully where movement of the gully head has been detected between two orthophotos. Next, in each of the 4 selected areas, 130 non-gully head points were randomly selected in order to match the number of gully heads for further statistical analysis. So, we have 4 types of points:

(1) Stable gully head (SG): gully head that did not advance between the first and last year of the analysed period.
(2) New active gully head (NAG): gully head that appears in the last year of the analysed period
(3) Existing active gully head (EAG): pre-existing gully head that advances between the first and last year of the analysed period.
(4) Points that are not gully head (NON): there is no gully head present.

The reason for selecting 2008–2019 is because 2010–2011 is the period of most intensive gully erosion in the last 60 years. Hayas et al. (2017a,b) found that due to extreme rainfall in this period, gully length almost doubled in this area. Thus, 2008 represents the most recent PNOA image available prior to this event, providing an appropriate baseline for assessing subsequent changes in gully dynamics. Besides, before 2008 the quality of the images was lower for locating gully heads. Orthophotos from the PNOA have a resolution of 0.5 m per pixel for 2008, and 0.25 m per pixel for the other years, which allows for high precision and detail in the visualization and analysis of the orthophotos.

Following the initial identification of gully heads from the orthophotos, and in order to assure correct alignment with the DEM, the location of the gully heads has been corrected by moving the gully head to the point of greatest accumulation of flow closest to it within a buffer

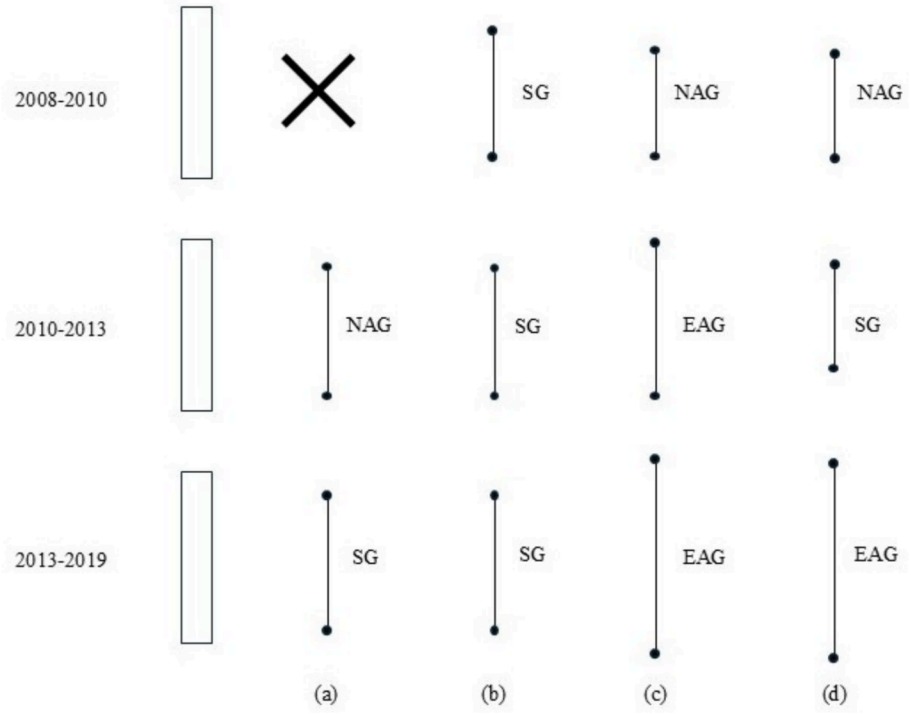


Fig. 3. Graphic illustration of gully head activity over the three study periods, showing four scenarios (NAG is new active gully head, SG is stable gully head, EAG is existing active gully head, X indicates no gully head presence, NON). Scenario a) In 2008–2010 there is no gully head, in 2010–2013 a new gully head appears (NAG) and the gully head remains in the same position in 2013–2019 (SG); Scenario b) A gully head is mapped in the same position in all photos from 2008 to 2019, so the gully remains stable (SG) over all three study periods, Scenario c) A gully head emerges (NAG) in 2008–2010 and advances in 2010–2013 and in 2013–2019. This gully head is considered existing active gully head in the last two periods (EAG) d) A gully head emerges (NAG) in 2008–2010 and the gully head remains stable in 2010–2013 and the gully head advances in 2013–2019.

of 5 m. This procedure was repeated for points that were not gully heads, with a 5 m buffer also applied and the point with the greatest accumulation of flow selected.

Futhermore, field verifications were conducted. In this regard, in 2019, an extensive field campaign was carried out to measure the characteristics of 82 gullies, 16 of which were included in this study. Furthermore, our 2019 field verifications provided us with confidence that the photointerpretations from the previous years of this study, 2013, 2010, and 2008, were correct with respect to the gully head position.

2.4. GHI index

In this study, we followed the methodology developed by index (De Geeter et al., 2025). They proposed to use a gully head initiation index GHI that assumes that a gully forms when the runoff shear stress exceeds the soil resistance to it (Montgomery and Dietrich, 1994):

$$\tau > \tau_{cr} \tag{1}$$

where: τ is the runoff shear stress [Pa] and τ_{cr} is the critical shear stress of the soil [Pa]. That is, it is assumed that gully heads should appear in landscape points where this condition is met, and otherwise, these should be areas of concentrated flow but without gully formation.

To obtain the GHI index, the first step is to obtain the runoff shear stress τ, using the following expression (Nearing et al., 1997):

$$\tau = \rho g R S \tag{2}$$

where: ρ represents the water density [kg/m^3], g the gravity constant [N/kg], R the hydraulic radius [m] and S the local slope [m/m].

R is difficult to determine in field conditions, as potential channel width and flow depths are generally unknown. However, it is reasonable to assume that both are proportional to the maximum runoff flow discharge that occurs (e.g. Nachtergaele et al., 2002). To obtain a relative approximation of these maximum flows discharges, the Curve Number (CN) method is used to obtain runoff depth as a function of the land use and the soil characteristics of the upstream area (Ponce and Hawkins, 1996; Mishra & Singh, 2003; Hawkins et al., 2009; Vanmaercke et al., 2021). This method, initially developed by the US Soil Conservation Service, has been revised and improved over the years, providing a solid basis for its application under a variety of hydrologic conditions. The basic equation for the runoff Curve Number method is (Hawkins et al., 2009):

$$R_d = \frac{(P - \lambda S_\lambda)^2}{P + (1 - \lambda) S_\lambda} \tag{3}$$

where: R_d is the daily runoff depth (mm), P is the daily rainfall depth (mm), S_λ the maximum potential runoff losses, where λ is the initial extraction fraction.

We used the maximum daily precipitation (P) values for each period (2008–2010), (2010–2013), and (2013–2019) and for each of the four study areas (Table 1), obtained from four meteorological stations near each area.

Estimates of CN are generally derived from empirical tables, which are based on land use, management conditions and soil type. These environmental factors are combined into the so-called Curve Number (CN) to estimate S_λ (Hawkins et al., 2009):

$$S_{0,05} = 33.96\left(\frac{100}{CN} - 10\right)^{1.15} \tag{4}$$

The CN was obtained from a layer available in REDIAM (Andalusian Environmental Information Network) using the method proposed by the SCS (Soil Conservation System).

Next, the daily discharge volumes Q [m^3] are calculated:

$$Q = A^* R_d \tag{5}$$

where: A is the drainage area (m^2).

This calculation enables the determination of shear stress index (SSI) that provides a simplified relationship with shear stress τ. This is achieved by using Q as a proxy for the hydraulic radius, while disregardig the constant variables ρ and g:

$$SSI = Q^* S \tag{6}$$

where: S is the local slope (m/m).

Once SSI has been calculated, the next step is to calculate the critical shear stress index.

(CSI), for which the expression of Nachtergaele et al. (2001) is used, based on the percentage of clay in the soil:

$$CSI = 0.311^* 10^{0.0182 Pc} \tag{7}$$

where Pc is the average percentage of clay corresponding to each of the study areas, and CSI is a simplified proxy based on the percentage of clay that can be expected to correlate with the actual critical shear stress.

The average percentage of clay in each sampling area was obtained from a layer available in REDIAM (Table 1). This layer was developed using information provided by the Ministry of Agriculture and Fisheries, calculated as part of the study *"Spatial Inference System for the Physico-chemical and Hydraulic Properties of Soils in Andalusia"*.

There are other factors, such as vegetation and stoniness, that can influence CSI, but they have not been considered due to a lack of data in the study areas.

Finally, by combining equations (6) and (7), the gully initiation index is finally obtained:

$$GHI = \frac{SSI}{CSI} \tag{8}$$

The GHI index obtained is an empirical and dimensionless indicator, so even if GHI > 1, a gully head does not necessarily form. However, the higher *GHI* values mean that there is a greater chance of being in areas with erosion and gully head formation, and vice versa.

The dataset was resampled to standardise the pixel size of the different spatial variables used. This process enabled the establishment of a common resolution of 2 m per pixel, thus facilitating their integration into subsequent analyses. The original datasets had different native resolutions, with the CN layer at 20 m and the clay content percentage layer at 100 m. This resampling process was carried out using the Python programming language.

2.5. Statistical analysis

Once the GHI index was calculated, it was checked if there were significant differences between the three types of gully head activity and points that are not gully heads.

This was carried out for each study area representing a type of landscape, and for the entire set of points, to **ascertain** whether the model presents differences between the selected areas or not, or whether it is better to treat all the points at once. The differences in the 3 selected time periods have also been analyzed.

The unpaired two-samples Wilcoxon test was used as a statistical test (with $\alpha = 0.01$), which is a non-parametric test used to compare whether there are significant differences between the four sets of points, obtaining the p-value for each analysis. The Wilcoxon test has been used in recent studies to analyse hydrological data (Saplıoğlu & Güçlü, 2022). The results were displayed using boxplots. All calculations have been performed using Rstudio (R Stats Package) scripting.

In addition, in order to evaluate the model, ROC (receiver operating characteristic curve) curves and AUC (area under the curve) values have been calculated for the 4 sets of points and the 4 representative areas of

each landscape. The ROC curve provides a graphical assessment of the model's accuracy and is constructed by plotting the relationships between the true positive rate (sensitivity, TPR) and the false positive rate (specificity-1, FPR). The Area Under the Curve (AUC) is a performance measure used to evaluate classification models, especially when performing a ROC curve analysis. The AUC allows differentiation between point classes, i.e., pixels that are gully heads and pixels that are not (Kleinbaum, 1994). An AUC of 1 indicates a perfect model, while an AUC of 0.5 suggests performance equivalent to chance (Kleinbaum, 1994; Hosmer & Lemeshow, 2000).

2.6. Determination of TT

Once the GHI was calculated for the three gully head activity types and for the points that are not gully heads, the TT was calculated for these to compare the two methods. Non-gully head points are plotted without the k calculation, as this only applies to gully head points. The following equation was used for the calculation:

$$S \geq kA^{-b} \quad (9)$$

where: S is the local slope [m/m] and A is the drainage area [ha].

As the variables S and A are independent, the S–A relationships were determined by orthogonal regression using R (the R statistical package). The k coefficients were obtained by reducing the previous regression line to the lowest point of an S–A scatter plot, maintaining the original slope or the b coefficient. Thus, the TT data are above this threshold line.

The value of the b coefficient is 0.38 based on field results from several studies (Nachtergaele et al., 2001; Torri & Poesen, 2014).

Keeping the value of the b coefficient constant isolates the effects of land use and precipitation on the k coefficient. Furthermore, the same type of surface flow is assumed throughout the analyzed period (2008–2019) and in all four study areas, which is supported by studies such as Torri & Poesen (2014) and Rossi et al. (2015). A simpler model is therefore obtained where k represents the gully head resistance for the existing active gully heads, the new active gully heads and the stable gully heads.

Finally, for the k coefficient, a p-value was calculated to compare whether there were significant differences between the four sets of points using ANOVA. Furthermore, to evaluate the topographic threshold model and compare it with the GHI, the ROC curves and the area under the curve (AUC) were calculated for the entire model. This was done by considering all points with gully heads versus those without.

3. Results

3.1. Observed gully heads and gully activity

A total of 475 gully heads were mapped. The mapped gully heads are shown in Fig. 4 for each of the four study areas, with different colors for the years 2008, 2010, 2013 and 2019. Mapped gully head varied between 16 and 79, or a respective density of 0,64 to 3,16 gully heads/km^2 depending on the landscape unit (Table 2). The highest densities were observed in the countryside foothills, while mid-mountains were characterised by the lowest densities. Most of the mapped gully heads are existing active gully heads or new active gully heads, although a significant amount of 29 % (138) was still found to be stable. Most of the stable gully heads are located in the countryside foothills, which is also the area with the highest gully head density. This suggests that this area is characterised by heavy erosion processes, but also that many gullies have grown to maturity over the last decades and stopped to grow. In fact, most of the amount of existing active gully heads is similar across the different landscape units, except for the mid-mountains which is a bit lower. Most of the new gully heads are located in countryside hills and foothills (Table 2).

Appendix A.I shows a detailed view of the evolution of a representative active gully over the period 2008–2019.

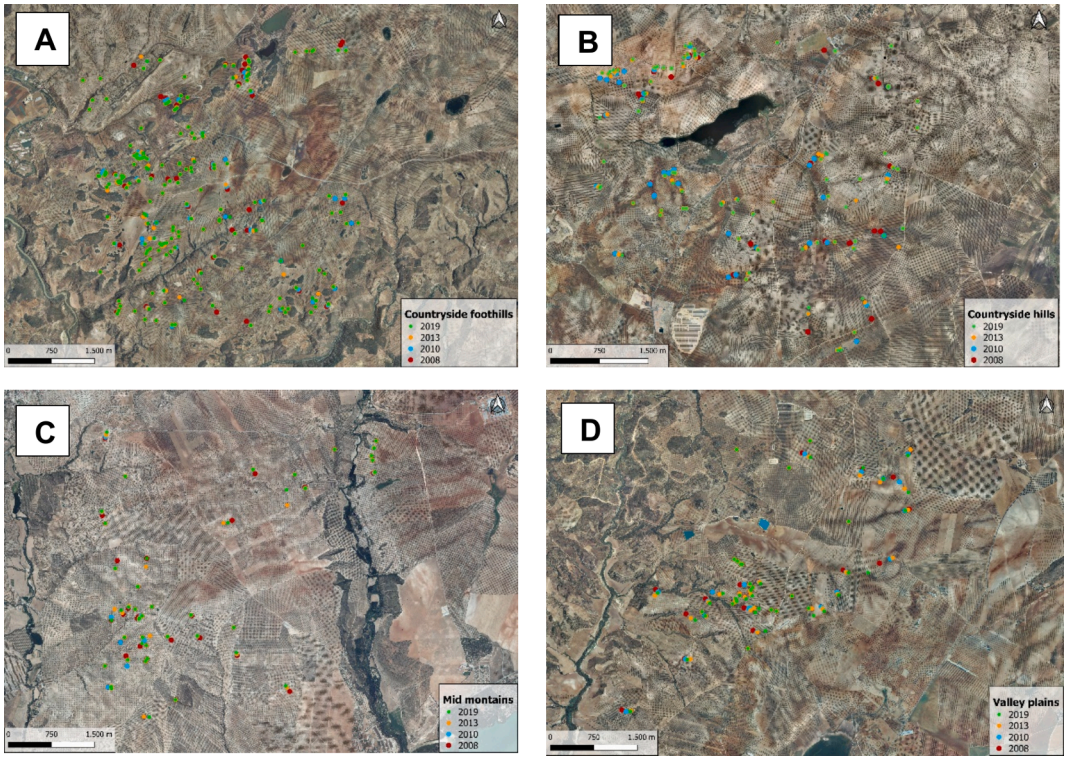


Fig. 4. Observed gully heads in the years 2008, 2010, 2013 and 2019 for each landscape unit: a) countryside foothills, b) countryside hills, c) mid mountains and d) valley plains. Each study area occupies 25 km^2.

Table 2
Number of gully heads by type of activity and time period in the study area.

Landscape unit	Time period	Existing active gully heads	New active gully heads	Stable gully heads	Total gully heads	Total gully head density (gully heads/ km^2)
Countryside hills	2008–2010	24	8	7	39	1,56
	2010–2013	30	11	5	46	1,84
	2013–2019	20	7	1	28	1,12
Countryside foothills	2008–2010	29	19	17	65	2,60
	2010–2013	27	16	30	73	2,92
	2013–2019	30	6	43	79	3,16
Valley plains	2008–2010	23	–	7	30	1,20
	2010–2013	24	–	7	31	1,24
	2013–2019	22	1	4	27	1,08
Mid-Mountains	2008–2010	14	8	2	24	0,96
	2010–2013	11	–	6	17	0,68
	2013–2019	7	–	9	16	0,64
Total	2008–2019	261	76	138	475	4,75

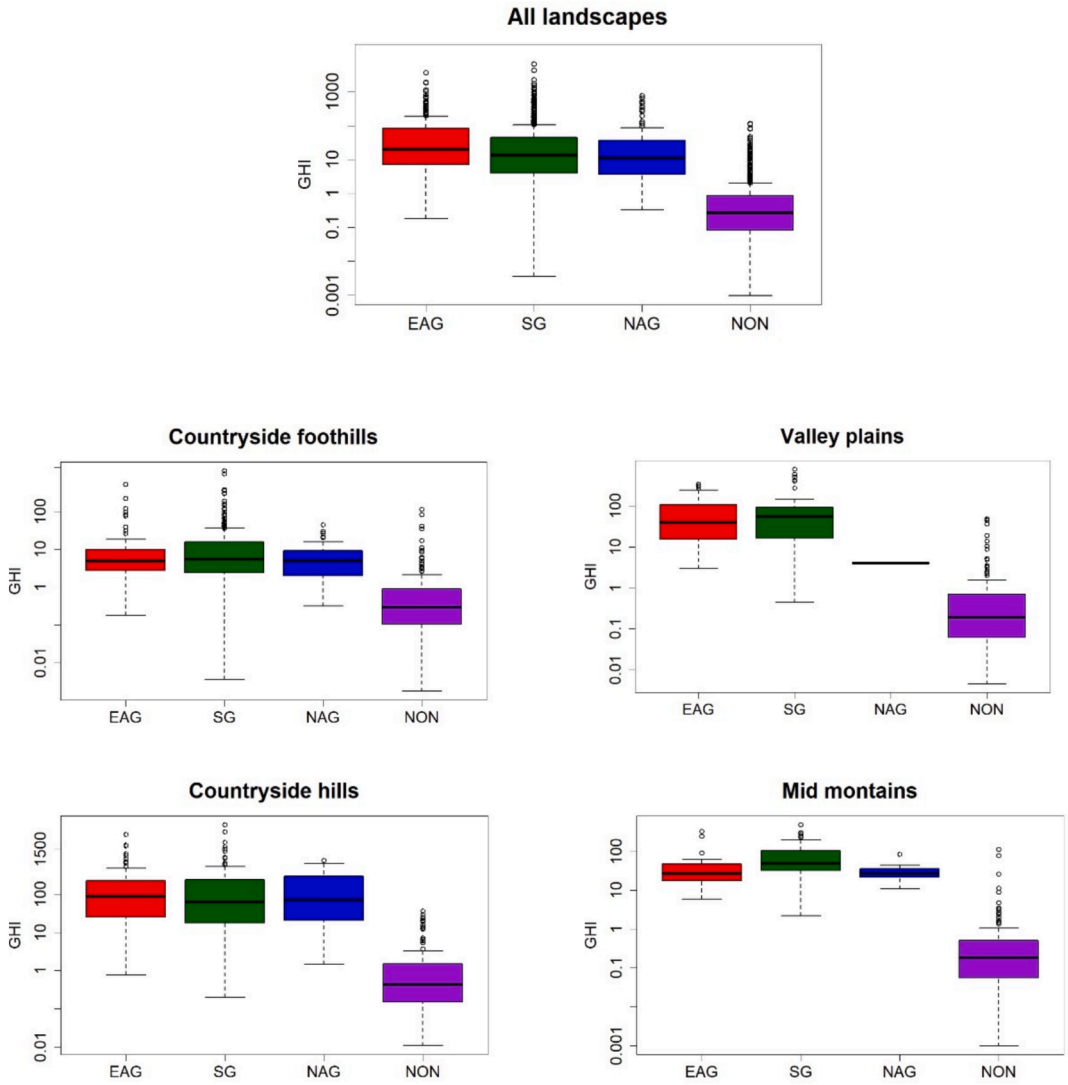


Fig. 5. Gully Head Initiation (GHI) index for the three gully head activity types and points that are not gully heads (EAG – existing active gully heads, SG – stable gully heads, NAG – new active gully heads, and NON– non-gully heads) in the 4 landscape units and for the entire data set (without differentiation between landscapes). The line dividing the box represents the median value.

3.2. Prediction of gully head initiation and activity across landscape units and time periods

The GHI index for all gully heads together (EAG, SG, and NAG), without subdividing gully head activity, is significantly higher compared to non-gully head areas (NON) (Fig. 5, Table 3). This is true across all landscape units, and when pooling all landscape units together, gully head pixels thus have a noticeably higher GHI index

Table 3
P-value per landscape and per gully head activity type and points that are not gully heads (EAG: existing active gully head, SG: stable gully heads, NAG: new active gully heads, and NON: non-gully heads), as well as for all points together.

Landscape unit	P- Value			
	EAG vs SG	SG vs NAG	EAG vs NAG	EAG, NAG, SG vs NON
Countryside hills	0.25	0.86	0.63	< 2.2e-16
Countryside foothills	0.42	0.31	0.59	< 2.2e-16
Mid-mountains	0.00099	0.05	0.87	< 2.2e-16
Valley plains	0.72	0.26	0.1	< 2.2e-16
All landscapes	4.3 e-05	0.77	0.01	< 2.2e-16

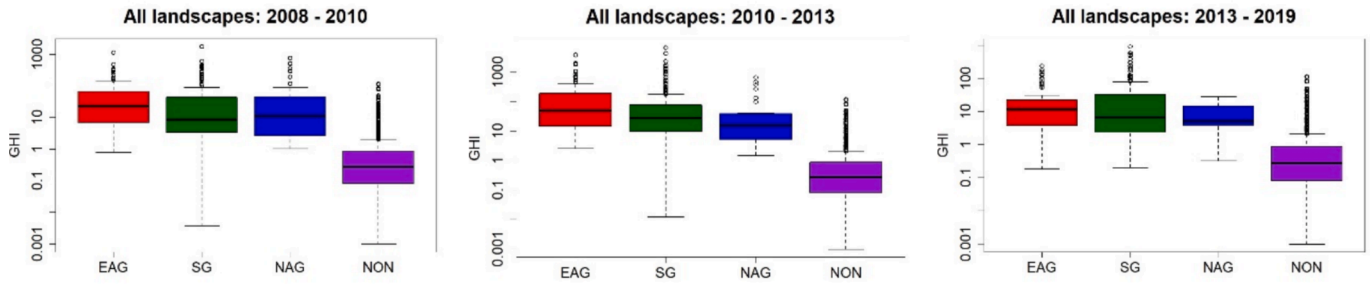


Fig. 6. Comparing GHI across three study periods (EAG: existing active gully heads, SG: stable gully heads, NAG: new active gully heads, and NON: non-gully heads).

compared to non-gully head pixels, reaching values of 1000 in the case of countryside hills and valley plains. In contrast, points that are not gully heads have values typically less than 1. However, statistical analysis of the landscapes reveals minimal variation in gully head types when considering them individually. It should be noted that discrepancies were only observed in a few landscape units. However, when all landscape units are grouped together, the differences between the three gully head activity types and points that are not gully heads become clearer. GHI works well to distinguish existing active gully heads from stable ones and new active gully heads, respectively. The difference between stable and new active gully heads is not statistically significant.

Next, we analyzed whether GHI performs differently according to the time period considered (Fig. 6, Table 4). The period 2010–2013 is marked by the highest maximum daily rainfall values (Table 1), and we know from other studies (Hayas et al., 2017a,b) that this period was marked by a period of considerable gully growth. Indeed, the greatest significance among gully head activity types is observed in the 2010–2013 period (Table 4). All gully head activity types could be well distinguished in this period. For the other two periods, with lower rainfall, there are still clear significant differences when comparing gully head activity types with non-gully head areas. However, apart from EAG vs SG in 2008–2010, the other differences are non-significant.

Table 4
P-values for gully head activity types and non-gully head points across time periods.

	2008–2010	2010–2013	2013–2019
EAG vs NAG	0.186	0.007	0.35
EAG vs SG	0.0024	0.0063	0.22
NAG vs SG	0.56	0.12	0.66
EAG, SG vs NON	$2.2e^{-16}$	$2.2e^{-16}$	$2.2e^{-16}$
NAG vs NON	$2.2e^{-16}$	$6.1e^{-15}$	$8.1e^{-07}$

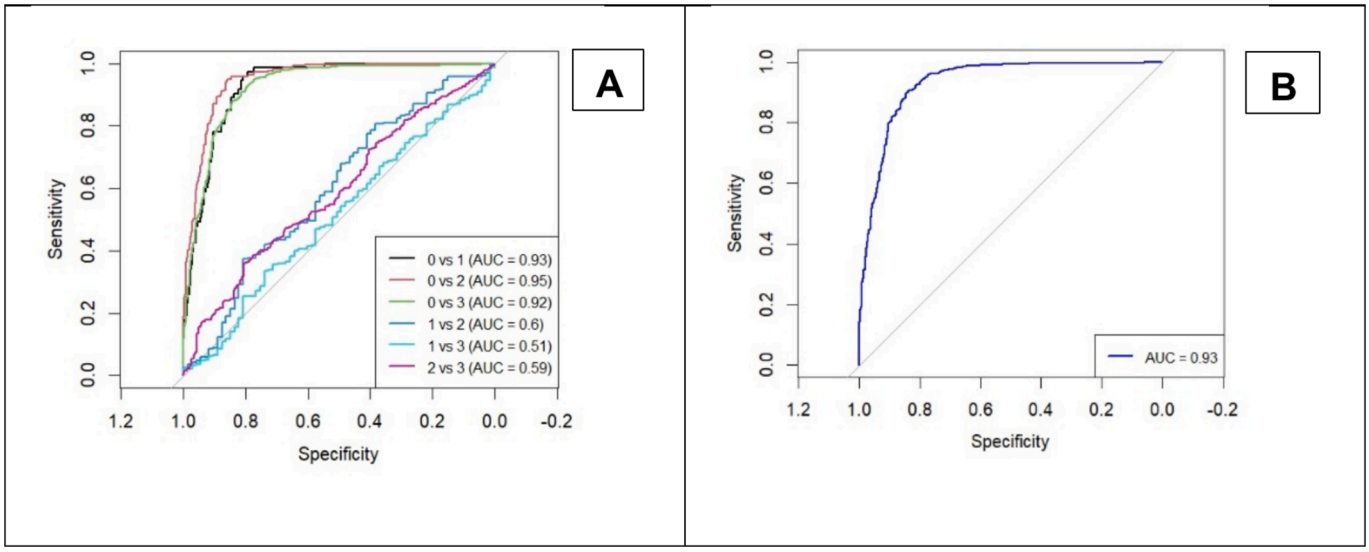


Fig. 7. (a) ROC curves and corresponding AUC values for each pairwise combination of the three gully head activity types, considering all landscapes together. (b) AUC for the entire model together, considering all the points that are gully heads compared to those that are not. 0: Non gully heads, NON; 1: New active gully heads NAG; 2: Existing active gully heads EAG; 3: Stable gully heads, SG.

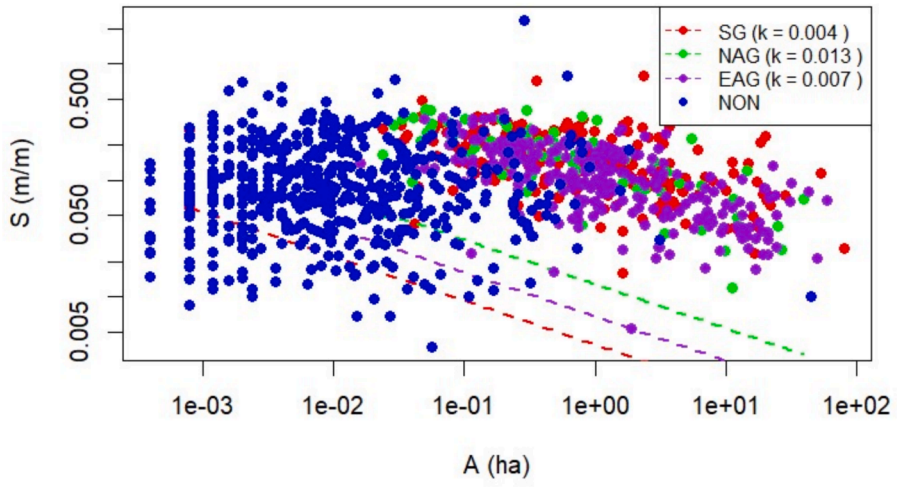


Fig. 8. Topographic threshold for gully head initiation applied to the three classes of gully head erosion activity and points that are not gully heads (SG: Stable gully heads, NAG: New active gully heads, EAG: Existing active gully heads, NON: Non-gully head).

Overall, the model using GHI to distinguish different types of gully head points (new active gully head, existing active gully head, stable gully head) from non-gully head points works very well, with an AUC between 0.92, 0.93 and 0.95 (see Fig. 7 a and b). Distinguishing between different types of gully heads according to their activity is a less effective approach. In the context of a model designed to differentiate between stable and newly active gully heads, the area under the curve (AUC) value reduces to 0.51. This indicates that the model's performance is only marginally superior to that of a completely random model. The model used to distinguish between new active and existing active gully heads, and stable and existing active gully heads, is similar, with an AUC of 0.6 and 0.59, respectively. In this instance, we have combined all the landscape units.

In comparison to the traditional TT approach, the GHI demonstrates significantly superior performance. As demonstrated in the S-A graph (Fig. 8), there is significant scatter. The TT values of the coefficient k obtained showed minimal variation among the gully heads points, with a value of 0.004 for stable gully heads, 0.013 for new active gully heads, and 0.007 for existing active gully heads. Furthermore, the distribution of the points representing the of gully heads (EAG, NAG and SG) is highly comparable, given the similarity in slope and contributing area values, with no substantial differences observed between them. When comparing AUC values for the TT approach (Table 5), they are only 0.64 when comparing all gully heads points with non-gully heads points, compared to 0.93 for the GHI approach.

This can be understood from the differences that exist across the different landscape units that are not taken into account in the TT approach. While the GHI approach is still a simplification of real process, it presents some considerable advantages: through the calculation of the runoff discharge, the differences that might exist in rainfall or soil properties that influence on the runoff characteristics of the curve number approach are accounted for. Also, to some extent, differences in erodibility are taken into account, even though in this study we only considered the effect of clay content on critical shear stress index.

Table 5
Performance of the Topographic threshold model for gully head initiation and activity. P-values are reported for the three gully head activity types (Existing Active Gully Heads, New Active Gully Heads, and Stable Gully Heads), as well as for gully head initiation (all gully head points versus non-gully head points). The AUC is provided for gully head initiation (all gully head points versus non-gully head points).

p-value				AUC
EAG vs SG	SG vs NAG	EAG vs NAG	EAG, NAG, SG vs NON	EAG, NAG, SG vs NON
0.07	0.34	0.99	0.0001	0.64

4. Discussion

The GHI applied to olive groves in the Guadalquivir basin has great predictive power at the regional level, for prediction of gully head initiation. It performs better than the TT for gully head initiation, which is the model that has traditionally been used for this purpose (Hayas et al., 2017b; Torri & Poesen, 2014), with an AUC of 0.93 versus an AUC of 0.64 respectively. The difference in significance between the new active gully heads and the existing active gully heads is particularly notable, with a p-value of 0.99 in the case of the TT and a p-value of 0.01 in the case of the GHI, thus improving the results of the traditional method when differentiating between the types of gully head activity described. GHI also offers the advantage that it is conceptually more robust, as it allows variability in land use, soils and rainfall to be directly considered. A key benefit of the GHI approach is that it eliminates the need for calibration and validation when expanding into new study areas. This is in contrast to the TT approach, where thresholds are determined based on observations of S-A values.

The GHI index model applied to the olive grove area of the Guadalquivir basin has given better results compared to the African and Ethiopian study areas of the study by De Geeter et al. (2025) with an AUC of 0.93 and 0.67, respectively. The resolution of the GIS data may be a contributing factor, since the DEM from which the drainage area (A) and slope (S) have been obtained has a resolution of 2x2 m. As a result of this resolution, the difference between the values of these factors for gully heads and non-gully heads is very high, and therefore the significance and AUC are higher. However, in our study, land use was kept constant, olive groves, and more detailed precipitation data was available, which reduces uncertainty.

Regarding gully activity, the GHI in olive groves failed to adequately distinguish all types of activity at the gully head in the analysed landscapes, with a few exceptions. In the case of the mid mountain landscape, we found significance when differentiating between existing gully heads and stable gully heads, and between stable gully heads and new active gully heads. In all other cases there is no significance. However, when considering all landscapes together, the types of gully head activity are significant, apart from the distinction between stable gully heads and new active gully heads (Table 3). In this sense, the fact that the model works best when all points are treated together indicates that it could be standardised across the entire Mediterranean olive-growing

basin.

A surprising and consistent result is that, across all landscapes and periods, existing active gully (EAG) heads exhibit higher GHI values than newly formed active gully (NAG) heads (Figs. 5 and 6). This unexpected pattern underlines the added conceptual and empirical value of the GHI in comparation to traditional slope–area thresholds. This is contrary to what would be expected based on the widely accepted theory that gullies tend to initiate at points in the landscape with the highest slope-area (S-A) values, which usually correspond to the highest GHI. According to this conceptual model (e.g. Vanwalleghem et al., 2005), after the initial formation, gully heads retreat upstream through headcut erosion, and eventually their growth slows as they evolve toward a more stable morphology. Therefore, we expect the S-A values or gully head initiation index (GHI) to vary during the gully's evolution over time. Due to the temporal resolution of the imagery never allows to detect a gully head as soon as it forms, it will almost never be possible to find the S-A or gully head initiation index value of the initiation point, but we expect that the one of the newly formed gully heads to be closer than that of the existing active gully heads. However, our findings suggest the opposite: In terms of topography and rainfall it is easier to form a new gully than to keep it actively expanding. Moreover, the statistical difference between both types of gully heads across all landscapes ($p = 0.01$; Table 3) reinforces the relevance of this differentiation for understanding gully head evolution dynamics. We believe this is a key and innovative finding of our research as it suggests that the conditions leading to gully initiation (which are mainly determined by S & A) are not necessarily sufficient to keep them active. These findings highlight the importance of distinguishing between NAG and EAG in gully head studies.

That the difference between a stable gully head and a new active gully head is insignificant indicates that the GHI may be good at explaining gully head initiation, but not in predicting whether gully heads remain active or not. One potential explanation for this is that, once a gully head forms, slope becomes less important. This is because gully head retreat is mainly caused by plunge pool erosion, rather than incision (shear stress). This plunge pool erosion should be correlated with discharge (Q). Yet, as gully heads retreat: A (and therefore Q) declines, while S typically increases (at least on a typical concave slope). Hence, we could assume that, while the GHI (which includes slope) may be good in predicting initiation, Q (or perhaps Q/CSI) might be better in differentiating between new active gully head and stable gully heads. This interpretation suggests that gully expansion is less controlled by local slope and shear stress, and more directly linked to discharge. In line with this, Vanmaercke et al. (2016) also reported that gully head retreat rates worldwide are mainly explained by rainfall intensity and contributing area, while local topography had very little effect.

In this sense, we need models that describe in more detail the hydrology and erosion processes within the gullies. Flores-Cervantes et al. (2006), propose a model of gully headcut retreat resulting from erosion in plunge pools. This model was implemented in the landscape evolution model CHILD (Channel-Hillslope Integrated Landscape Development), resulting in the retreat of the gully being more significant in areas with gentle slopes or large headcuts. Their model approach was also implemented in the Revised Ephemeral Gully Erosion Model (REGEM) developed by Gordon et al. (2006). They tested their model by running continuous simulations and allowing the gully channel to evolve over multiple runoff events, accounting for seasonal variations in management operations and soil conditions, which could help differentiate between active and permanent gullies. Campo Bescós (2011) evaluated the CHILD model to analyse the dynamics of permanent gullies in the Bardenas Reales in Spain. These types of process-based models could be useful for differentiating gully head activity, as analysed in this study. However, for this, much more information is needed on gully morphology, such as the depth of the plunge pool or height of the headcut, and on rainfall-runoff dynamics, that is often not available at the regional scale.

The absence of statistical significance in certain landscapes may be attributable to a variety of factors. On the one hand, the lack of significant differences between gully heads in mid-mountain and valley plains landscapes could be related to the limited number of points analysed in these areas. Regarding mid-mountain landscape, the number of points analysed is comparatively low, and almost half of these are stable gully heads (Table 2). In the case of countryside foothills, the number of points analysed is much higher, but stable gully heads also predominate (Table 2). Consequently, both areas demonstrate reduced gully head activity. This could explain why there is no significance regarding existing active gully heads and new active gully heads. This difference in the number of gully heads is due to the dynamics of each landscape, defined by its slopes, rainfall, or soil type. This variability between landscapes makes the model more robust. Another reason why the differences between the landscapes are not substantial enough to be detected is that all areas are over the same crop. This means that the differences are limited to climate and soil geology, which is also the same for each study area. Furthermore, since we use maximum precipitation values, the differences in climate and soil-geology that determine runoff characteristics are reduced because we only consider the most extreme and unfavorable conditions. In this sense, when we apply the model to all landscapes together, the greater variability between them can better distinguish the types of gully head activity. These differences in climate and geology between landscapes can be observed in Table 1.

Regarding the analysis of each period (2008–2010, 2010–2013, and 2013–2019), the greatest significance among the gully heads classes analyzed is found in the 2010–2013 period (Table 4). The heightened activity during this period, attributable to substantial rainfall, signifies an elevated risk of the emergence of new active gully heads or the expansion of existing ones, a finding that aligns with the conclusions of numerous studies.

Although the model has been successful in differentiating between areas with and without gully heads, it has certain limitations. Please note that simplifications to the model may lead to significant errors, since it is a dimensionless index intended for application at larger scales or in situations where data are lacking. One of the limitations of this study is the difference in the time intervals used in the analysis. The GHI is expected to perform better over shorter periods, as these allow clearer identification of the precipitation events that triggered gully formation or expansion. However, the available data do not allow such high temporal resolution, as the analysis is constrained by the availability of PNOA orthophotos in the study area. Furthermore, the longest interval, between 2013 and 2019, was selected due to the stability in gully network density observed since 2011 (Hayas et al., 2017a,b). Another limitation concerns the representation of slope in the analysis. Prior to gully head formation, the surface slope can be adequately captured by the 2 m resolution DEM. However, once the gully head stabilises, its dynamics are primarily governed by plunge pool erosion. At this stage, the 2 m DEM may no longer resolve the micro-topographic slopes that exert a pronounced influence, thereby reducing the effectiveness of the GHI in distinguishing between active and stable gully heads. Nevertheless, this is a common problem with regional-scale erosion studies, that cannot be resolved without detailed measurement of individual gullies. With regard to precipitation, the maximum daily precipitation for each period analysed was selected, on the assumption that runoff flows are representative of the peak flows that form the gully heads. It would be interesting to study other, less pronounced precipitation indices to test the influence of rainfall on the GHI index model, such as Rainy Day Normal or any other cumulative rainfall or erosivity index as Hayas et al. (2017b) did. This spatial mismatch may introduce a degree of uncertainty in the precipitation values used to model gully head formation and thus should be considered when interpreting the results. It is important to note that a single precipitation value was assigned to each quadrant and period, assuming a uniform distribution of rainfall within the study area. While we acknowledge the potential spatial variability of precipitation within these quadrants, more detailed spatial

variability data are not available at the resolution of this work. Furthermore, the weather stations used are not located within each study area but are the closest available sources of precipitation data covering the entire study period. This is a source of uncertainty, and stronger results might have been obtained if more accurate rainfall data had been available. With regard to the calculation of the critical shear stress index, only the percentage of clay has been considered (Nachtergaele et al., 2001). It would be beneficial to consider incorporating additional variables, such as vegetation characteristics, into the calculation of this index, given the documented influence of vegetation and root characteristics (Knapen et al., 2007; Vannoppen et al., 2015). Furthermore, critical shear stress is also related to soil properties such as organic matter and rock fragments (Poesen et al., 2003). However, organic matter is similar across all study areas, with a maximum value of 1.29 % in the valley plain and a minimum value of 0.98 % in the mid-mountain, so the results would not be greatly altered in this study. Therefore, it was decided that these characteristics would not be included, since the cultivation is the same in all four areas analysed and detailed information on stoniness is rarely available at the regional level.

In light of our previous results, a key point for considering the broader applicability of our findings is the comparison with the slope–area (S–A) threshold concept, which has traditionally been the most widely used empirical strategy to identify areas prone to gully initiation. However, the S–A threshold approach has several limitations. Firstly, it only reflects the lower limit of gully formation conditions and is often determined by only a few gullies with the lowest slope and/or drainage area values, making it little robust. Secondly, it exclusively incorporates slope and area, while other fundamental factors such as soil properties, vegetation cover, and rainfall are implicitly aggregated in the empirical coefficient and exponent. This complicates the extrapolation of S–A thresholds to other regions and their application in scenario analyses. Thirdly, S–A thresholds merely indicate potential initiation locations but do not provide information on whether or when gullies are actively retreating.

More complex, process-based models have the potential to overcome some of these limitations, but they generally require extensive datasets that are often unavailable, especially at regional to global scales (Vanmaercke et al., 2021). This situation highlights the need for a model that is both conceptually sound and applicable with feasible data requirements. The Gully Head Initiation (GHI) index aims to address this methodological gap. As demonstrated in our boxplot analysis (Fig. 5), the GHI allows thresholds of gully initiation to be defined across the entire distribution of gully heads, rather than being restricted to the most extreme cases. Furthermore, it explicitly integrates environmental drivers such as topography, rainfall, soil, and land use in a logical and transparent manner, offering potential not only for prediction but also for scenario analyses. Importantly, as shown here, the GHI index can also be used to explore patterns of gully expansion, which is an innovative and novel application compared to existing approaches.

Furthermore, as demonstrated by De Geeter et al. (2025), the GHI was first proposed and tested in Ethiopia, demonstrating promising predictive capacity for gully initiation. In this work, we demonstrate that the GHI is equally applicable in a completely contrasting Mediterranean environment. Crucially, the index does not depend on site-specific calibration but instead relies on predefined factors and parameter tables that align with current understanding of gully formation processes. Despite this, the index clearly distinguishes gully from non-gully locations and even separates active from stable gully heads, albeit with some limitations regarding newly formed gullies. It is evident that the findings are in alignment with the results obtained by De Geeter et al. (2025), even though these were tested under wholly distinct conditions. This confirms that the GHI has real potential for wider application. The strength of this approach lies in its balance between sufficient conceptual complexity—by integrating the dominant factors of gully erosion—and data feasibility for application across larger regions. However, further development and testing will be required before full global applicability can be confirmed.

5. Conclusions

In this study, a new Gully Initiation Index (GHI) model developed by index (De Geeter et al., 2025) was applied to four representative types of olive grove landscapes in the Mediterranean Guadalquivir basin (Andalusia, Spain). To this end, a dataset of 475 gully heads analysed from 2008 to 2019 was created. Gully head activity was classified into three classes: existing active gully heads, new active gully heads and stable gully heads. Furthermore, 520 non-gully head points were randomly selected for analysis.

After applying the GHI to all the gully heads points, promising results were obtained. The model's area under the curve (AUC) for differentiating between gully heads and non-gully heads was 0.93, demonstrating the potential of this model for predicting gully head initiation. Furthermore, the GHI index outperforms the topographic threshold for gully head initiation (AUC = 0.64) and for the different gully head types. This confirms that considering factors other than drainage area (A) and local slope (S) influence gully head formation processes. Importantly, this performance was considerably higher than that obtained in previous applications in Ethiopia (AUC = 0.67; De Geeter et al., 2025), which underlines the robustness of the index under Mediterranean conditions.

Regarding gully activity, the model was not as effective in differentiating between new active gully heads and stable gully heads. The lack of significant differences between stable and newly formed gully heads suggests that the GHI is adequate for predicting initiation but not for assessing gully head persistence, which appears to be more closely controlled by plunge pool erosion processes than by slope or drainage area. However, the model performed reasonably well in differentiating between new active gully heads and existing active gully heads (AUC = 0.60), and between existing active gully heads and stable gully heads (AUC = 0.59). This would help differentiate the areas with the greatest erosion activity and inform decisions for better olive grove management.

A key and innovative result of this work is that existing active gully heads consistently exhibit higher GHI values than newly formed gully heads. This pattern challenges the traditional slope-area threshold theory, which assumes that gullies initiate at points with higher threshold values. Our findings suggest that, while initiation can occur at locations with lower GHI, maintaining gully activity requires more restrictive conditions. In light of these findings, and the significance of the relationship between these two types of gully heads (p-value = 0.01), it is important to consider them for future studies.

Regarding the individual analysis by landscape type, the results were inconclusive, except for the mid-mountain landscape, where there was a significant difference between existing active gully heads and stable gully heads. In contrast, when grouping all the points without differentiating between landscapes, a significant difference was identified between existing active gully heads and stable gully heads (p-value = 4.3 e-05), as well as between active and new active gully heads (p-value of 0.01), indicating that we can scale this model to all olive groves in the Guadalquivir basin located in these landscapes. This outcome is indicative of limited number of points and the predominance of stable gully heads in some landscapes, as well as the homogenising effect of land use, which reduces variability when analysed separately. Integrating all landscapes together has been shown to enhance the model's capacity to detect significant differences between gully head activity classes, due to the broader range of topographic and climatic conditions this approach integrates.

One important limitation of the study is that the model is based on several assumptions, such as using daily discharge volumes instead of hydraulic radius and excluding factors like vegetation and stoniness in the calculation of critical shear stress index, given that all four study areas share the same crop type. Nevertheless, despite these simplifications, the process-based nature of the GHI provides valuable insights into gully initiation and dynamics.

In addition to its predictive accuracy, the GHI also proves to be a conceptually robust and widely applicable tool. Unlike traditional slope-area thresholds, which are limited in scope and difficult to extrapolate, the GHI seamlessly integrates topography, rainfall, soil, and land use, making it suitable for scenario analysis and the identification of trigger zones. On the other hand, process-based models are often limited by the need for data that are difficult to access. Furthermore, our results confirm that the index performs consistently in contrasting environments (Ethiopian highlands and Mediterranean olive groves) without the need for site-specific calibration, underscoring its potential for broader application. This balance between conceptual robustness and data viability positions the GHI as a promising framework for advancing gully erosion modelling beyond local scales. However, further development and testing are needed before its full global applicability can be assumed, as there are still uncertainties regarding its performance in areas with different land use, land characteristics, and meteorological conditions.

Data availability

Data will be made available from the corresponding author upon reasonable request.

Author contributions

1. Paula González: Writing – Original draft, Methodology, Article search, screening, analysis, Conceptualization
2. Adolfo Peña: Screening, supervision, final approval.
3. Sofie De Geeter: Methodology.
4. Matthias Vanmaercke: Methodology & review.
5. Jean Poesen: Review & Conceptualization
6. Tom Vanwalleghem: Review, editing & final approval.

Funding sources

This research has been carried out with funding from the "Consejería de Universidad Investigación e Innovación de la Junta de Andalucía", within the scope of the Andalusian Plan for Research, Development and Innovation (PAIDI 2020), which covered the Article Processing Charge (APC).

CRediT authorship contribution statement

Paula González: Writing – original draft, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Adolfo Peña:** Validation, Supervision, Project administration. **Sofie De Geeter:** Methodology. **Matthias Vanmaercke:** Writing – review & editing, Methodology. **Jean Poesen:** Writing – review & editing, Conceptualization. **Tom Vanwalleghem:** Writing – review & editing, Validation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Paula Gonzalez Garrido reports financial support was provided by Consejería de Universidad Investigación e Innovación de la Junta de Andalucía. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This research is part of the research project "CARCAVA. Climatic and Agronomic Influence on the Formation and Evolution of the Gully Network in the Andalusian Countryside," funded by the Ministry of University, Research, and Innovation of Government of Andalucia.

Also, this research was supported by ENIA International Chair in Agriculture, University of Cordoba, who covered the Article Processing Charge (APC).

Finally, we would also like to thank the University of KU Leuven for its collaboration in the predoctoral stay of the first author of this publication.

Appendix A.1

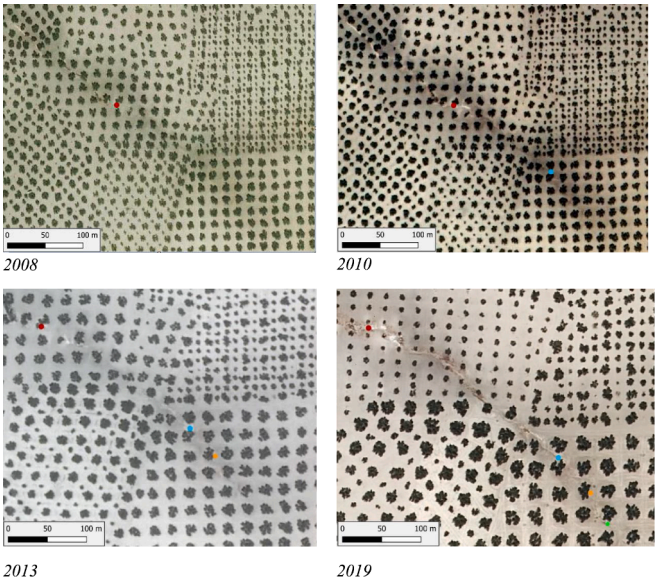


Fig. A.1. Evolution of a representative active gully head over the years of the study. This gully head already exists in the first analyzed image but strongly retreats between 2008–2010. It continues growing over the period 2010–2013 and 2013–2019, although less fast. This example is an existing active gully head, as it is advancing during the 3 periods analyzed.

References

Bescós, M. Á. C. (2011). *Caracterización de la red de cárcavas permanentes en una cuenca semi-árida mediante fotogrametría y modelado* [Doctoral dissertation, Universidad Pública de Navarra]. https://dialnet.unirioja.es/servlet/tesis?codigo=146041.

Borrelli, P., Panagos, P., Alewell, C., Ballabio, C., De Oliveira Fagundes, H., Haregeweyn, N., Lugato, E., Maerker, M., Poesen, J., Vanmaercke, M., Robinson, D. A., 2022. Policy implications of multiple concurrent soil erosion processes in European farmland. Nat. Sustainab. 6 (1), 103–112. https://doi.org/10.1038/s41893-022-00988-4.

De Geeter, S., Verstraeten, G., Poesen, J., Campforts, B., Vanmaercke, M., 2023. A data driven gully head susceptibility map of Africa at 30 m resolution. Environ. Res. 224, 115573. https://doi.org/10.1016/j.envres.2023.115573.

De Geeter, S., Setargie, T.A., Haregeweyn, N., Verstraeten, G., Poesen, J., Tsunekawa, A., et al., 2025. Advancing gully initiation modelling by means of a Curve Number (CN) method: a way forward? Earth Surf. Proc. Landforms 50 (11), e70145. https://doi.org/10.1002/esp.70145.

Parliament, E., 2024. EU Nature Restoration Act. European Union. Retrieved from https://ec.europa.eu.

Fan, F., Gu, X., Luo, J., Zhang, B., Liu, H., Yang, H., Wang, L., 2024. Identification of gully erosion activity and its influencing factors: a case study of the Sunshui River Basin. PLoS One 19 (11), e0309672. https://doi.org/10.1371/journal.pone.0309672.

Flores-Cervantes, J.H., Istanbulluoglu, E., Bras, R.L., 2006. Development of gullies on the landscape: a model of headcut retreat resulting from plunge pool erosion. J. Geophys. Res. Atmos. 111 (F1). https://doi.org/10.1029/2004jf000226.

Gordon, L.M., Bennett, S.J., Bingner, R., Theurer, F., Alonso, C., 2006. REGEM: the revised ephemeral gully erosion model. Proceedings of the Eighth Federal Interagency Sedimentation Conference (8thFISC), Reno, NV, USA.

Gutiérrez, Á.G., Schnabel, S., Contador, F.L., 2009. Gully erosion, land use and topographical thresholds during the last 60 years in a small rangeland catchment in SW Spain. Land Degrad. Dev. 20 (5), 535–550. https://doi.org/10.1002/ldr.931.

Hawkins, R.H., Ward, T.J., Woodward, D.E., Van Mullem, J.A., 2009. Curve Number Hydrology: State of the Practice. ASCE/EWRI. ISBN 978-0-7844-1004-2, 10.1061/(asce)he.1943-5584.0000119.

Hayas, A., Vanwalleghem, T., Laguna, A., Peña, A., Giráldez, J.V., 2017a. Reconstructing long-term gully dynamics in Mediterranean agricultural areas. Hydrol. Earth Syst. Sci. 21 (1), 235–249. https://doi.org/10.5194/hess-21-235-2017.

Hayas, A., Poesen, J., Vanwalleghem, T., 2017b. Rainfall and vegetation effects on temporal variation of topographic thresholds for gully initiation in Mediterranean cropland and olive groves. Land Degrad. Dev. 28 (8), 2540–2552. https://doi.org/10.1002/ldr.2805.

Hosmer, D. W., & Lemeshow, S. (2000). *Applied logistic regression* (2nd ed.).

Hrachowitz, M., Fovet, O., Ruiz, L., Euser, T., Gharari, S., Nijzink, R., Freer, J., Savenije, H.H.G., Gascuel-Odoux, C., 2014. Process consistency in models: the importance of system signatures, expert knowledge, and process complexity. Water Resour. Res. 50 (9), 7445–7469. https://doi.org/10.1002/2014wr015484.

Imaizumi, F., Hattanji, T., Hayakawa, Y.S., 2010. Channel initiation by surface and subsurface flows in a steep catchment of the Akaishi Mountains, Japan. Geomorphology 115 (1–2), 32–42. https://doi.org/10.1016/j.geomorph.2009.09.026.

Jones, R., 2002. Algorithms for using a DEM for mapping catchment areas of stream sediment samples. Comput. Geosci. 28 (9), 1051–1060. https://doi.org/10.1016/s0098-3004(02)00022-5.

Junta de Andalucía, J., de Agricultura, C., Ganadería, P., y Desarrollo Sostenible, 2018. Paisajes de Andalucía. Secretaría General de Medio Ambiente, Agua y Cambio Climático.

Kariminejad, N., Pourghasemi, H.R., Hosseinalizadeh, M., Rossi, M., Mondini, A., 2023. Evaluating land degradation by gully erosion through soil erosion indices and rainfall thresholds. Nat. Hazards 117 (3), 3353–3369. https://doi.org/10.1007/s11069-023-05990-3.

Kleinbaum, D.G., 1994. Modeling strategy guidelines. Springer eBooks, pp. 161–189. https://doi.org/10.1007/978-1-4757-4108-7_6.

Knapen, A., Poesen, J., Govers, G., Gyssels, G., Nachtergaele, J., 2007. Resistance of soils to concentrated flow erosion: a review. Earth Sci. Rev. 80 (1–2), 75–109. https://doi.org/10.1016/j.earscirev.2006.08.001.

Li, H., Jin, J., Dong, F., Zhang, J., Li, L., Zhang, Y., 2024. Gully erosion susceptibility prediction using high-resolution data: evaluation, comparison, and improvement of multiple machine learning models. Remote Sens. (Basel) 16 (24), 4742. https://doi.org/10.3390/rs16244742.

Ministerio de Agricultura, Pesca y Alimentación. (2024). *Encuesta sobre superficies y rendimientos cultivos (ESYRCE): Encuesta de marco de áreas de España.* https://www.mapa.gob.es/es/estadistica/temas/estadisticas-agrarias/agricultura/esyrce/.

Mishra, S.K., Singh, V.P., 2003. Soil conservation service curve number (SCS-CN) methodology. Water Sci. Technol. Lib. https://doi.org/10.1007/978-94-017-0147-1.

Montgomery, D.R., Dietrich, W.E., 1994. Landscape dissection and drainage area-slope thresholds. Process Models Theor. Geomorphol. 221–246.

Nachtergaele, J., Poesen, J., Steegen, A., Takken, I., Beuselinck, L., Vandekerckhove, L., Govers, G., 2001. The value of a physically based model versus an empirical approach in the prediction of ephemeral gully erosion for loess-derived soils. Geomorphology 40 (3–4), 237–252. https://doi.org/10.1016/s0169-555x(01)00046-0.

Nachtergaele, J., Poesen, J., Sidorchuk, A., Torri, D., 2002. Prediction of concentrated flow width in ephemeral gully channels. Hydrol. Process. 16 (10), 1935–1953. https://doi.org/10.1002/hyp.392.

Nearing, M.A., Norton, L.D., Bulgakov, D.A., Larionov, G.A., West, L.T., Dontsova, K.M., 1997. Hydraulics and erosion in eroding rills. Water Resour. Res. 33 (4), 865–876. https://doi.org/10.1029/97wr00013.

O'Callaghan, J.F., Mark, D.M., 1984. The extraction of drainage networks from digital elevation data. Comput. Vision Graph. Image Proc. 28 (3), 323–344. https://doi.org/10.1016/s0734-189x(84)80011-0.

Panagos, P., Borrelli, P., Poesen, J., Ballabio, C., Lugato, E., Meusburger, K., Montanarella, L., Alewell, C., 2015. The new assessment of soil loss by water erosion in Europe. Environ. Sci. Policy 54, 438–447. https://doi.org/10.1016/j.envsci.2015.08.012.

Panagos, P., Montanarella, L., Barbero, M., Schneegans, A., Aguglia, L., Jones, A., 2022. Soil priorities in the European Union. Geoderma Reg. 29, e00510. https://doi.org/10.1016/j.geodrs.2022.e00510.

Panagos, P., Broothaerts, N., Ballabio, C., Orgiazzi, A., De Rosa, D., Borrelli, P., Jones, A., 2024. How the EU soil observatory is providing solid science for healthy soils. Eur. J. Soil Sci. 75 (3). https://doi.org/10.1111/ejss.13507.

Peel, M.C., Finlayson, B.L., McMahon, T.A., 2007. Updated world map of the Köppen-Geiger climate classification. Hydrol. Earth Syst. Sci. 11 (5), 1633–1644. https://doi.org/10.5194/hess-11-1633-2007.

Poesen, J., Vandekerckhove, L., Nachtergaele, J., Wijdenes, D.O., Verstraeten, G., Van Wesemael, B., 2002. Gully erosion in dryland environments. In: Dryland Rivers: Hydrology and Geomorphology of Semi-arid Channels, p. 229.

Poesen, J., Nachtergaele, J., Verstraeten, G., Valentin, C., 2003. Gully erosion and environmental change: Importance and research needs. Catena 50 (2–4), 91–133. https://doi.org/10.1016/s0341-8162(02)00143-1.

Ponce, V.M., Hawkins, R.H., 1996. Runoff curve number: has it reached maturity? J. Hydrol. Eng. 1 (1), 11–19. https://doi.org/10.1061/(ASCE)1084-0699(1996)1:1(11).

Quinn, P., Beven, K., Chevallier, P., Planchon, O., 1991. The prediction of hillslope flow paths for distributed hydrological modelling using digital terrain models. Hydrol. Process. 5 (1), 59–79. https://doi.org/10.1002/hyp.3360050106.

Rossi, M., 2014. Modeling of landslide phenomena and erosion processes triggered by meteo-climatic factors. Università Degli Studi Di Perugia, p. 3835 https://doi.org/10.13140/2.1.

Rossi, M., Torri, D., Santi, E., 2015. Bias in topographic thresholds for gully heads. Nat. Hazards 79 (S1), 51–69. https://doi.org/10.1007/s11069-015-1701-2.

Saplıoğlu, K., Güçlü, Y.S., 2022. Combination of Wilcoxon test and scatter diagram for trend analysis of hydrological data. J. Hydrol. 612, 128132. https://doi.org/10.1016/j.jhydrol.2022.128132.

Shojaeezadeh, S.A., et al., 2024. Soil erosion in the United States: present and future (2020–2050). Catena 242, 108074. https://doi.org/10.1016/j.catena.2024.108074.

Tan, Z., Leung, L.R., Li, H., Cohen, S., 2021. Representing global soil erosion and sediment flux in Earth system models. J. Adv. Model. Earth Syst. 14 (1). https://doi.org/10.1029/2021ms002756.

Tarboton, D.G., 1997. A new method for the determination of flow directions and upslope areas in grid digital elevation models. Water Resour. Res. 33 (2), 309–319. https://doi.org/10.1029/96wr03137.

Torri, D., Poesen, J., 2014. A review of topographic threshold conditions for gully head development in different environments. Earth Sci. Rev. 130, 73–85. https://doi.org/10.1016/j.earscirev.2013.12.006.

Vandaele, K., Poesen, J., Govers, G., Van Wesemael, B., 1996. Geomorphic threshold conditions for ephemeral gully incision. Geomorphology 16 (2), 161–173. https://doi.org/10.1016/0169-555x(95)00141-q.

Vandekerckhove, L., Poesen, J., Wijdenes, D.O., De Figueiredo, T., 1998. Topographical thresholds for ephemeral gully initiation in intensively cultivated areas of the Mediterranean. Catena 33 (3–4), 271–292. https://doi.org/10.1016/s0341-8162(98)00068-x.

Vandekerckhove, L., Poesen, J., Wijdenes, D.O., Nachtergaele, J., Kosmas, C., Roxo, M., De Figueiredo, T., 2000. Thresholds for gully initiation and sedimentation in Mediterranean Europe. Earth Surf. Proc. Land. 25 (11), 1201–1220. https://doi.org/10.1002/1096-9837(200010)25:11.

Vanmaercke, M., Poesen, J., Van Mele, B., Demuzere, M., Bruynseels, A., Golosov, V., Bezerra, J.F.R., Bolysov, S., Dvinskih, A., Frankl, A., Fuseina, Y., Guerra, A.J.T., Haregeweyn, N., Ionita, I., Imwangana, F.M., Moeyersons, J., Moshe, I., Samani, A. N., Niacsu, L., Yermolaev, O., 2016. How fast do gully headcuts retreat? Earth Sci. Rev. 154, 336–355. https://doi.org/10.1016/j.earscirev.2016.01.009.

Vanmaercke, M., et al., 2021. Measuring, modelling and managing gully erosion at large scales: a state of the art. Earth Sci. Rev. 218, 103637. https://doi.org/10.1016/j.earscirev.2021.103637.

Vannoppen, W., Vanmaercke, M., De Baets, S., Poesen, J., 2015. A review of the mechanical effects of plant roots on concentrated flow erosion rates. Earth Sci. Rev. 150, 666–678. https://doi.org/10.1016/j.earscirev.2015.08.011.