Use of TanDEM-X and Sentinel Products to Derive Gully Activity Maps in Kunene Region (Namibia) Based on Automatic Iterative Random Forest Approach

Miguel Vallejo Orti, Lukas Winiwarter, Eva Corral-Pazos-de-Provens, Jack G. Williams, Olaf Bubenzer, and Bernhard Höfle

***Abstract*—Gullies are landforms with specific patterns of shape, topography, hydrology, vegetation, and soil characteristics. Remote sensing products (TanDEM-X, Sentinel-1, and Sentinel-2) serve as inputs into an iterative algorithm, initialized using a micro-mapping simulation as training data, to map gullies in the northwestern of Namibia. A Random Forest Classifier examines pixels with similar characteristics in a pool of unlabeled data, and gully objects are detected where high densities of gully pixels are enclosed by an alpha shape. Gully objects are used in subsequent iterations following a mechanism where the algorithm uses the most reliable pixels as gully training samples. The gully class continuously grows until an optimal scenario in terms of accuracy is achieved. Results are benchmarked with manually tagged gullies (initial gully labeled area $<0.3\%$ of the total study area) in two different watersheds (408 and 302 km^2, respectively) yielding total accuracies of $>98\%$, with 60% in the gully class, Cohen Kappa >0.5, Matthews Correlation Coefficient >0.5, and receiver operating characteristic Area Under the Curve >0.89. Hence, our method outlines gullies keeping low false-positive rates while the classification quality has a good balance for the two classes (gully/no gully). Results show the most significant gully descriptors as the high temporal radar signal coherence (22.4%) and the low temporal variability in Normalized Difference Vegetation Index (21.8%). This research builds on previous studies to face the challenge of identifying and outlining gully-affected areas with a shortage of training data using global datasets, which are then transferable to other large (semi-) arid regions.**

Manuscript received June 20, 2020; revised October 30, 2020; accepted November 17, 2020. Date of publication November 24, 2020; date of current version January 6, 2021. This work was supported by the Deutsches Zentrum für Luft- und Raumfahrt (DLR) as the donor for the used TanDEM-X datasets, by the Integrated Land Management Institute (ILMI) and by the Kurt-Hiehle-Foundation. *(Corresponding author: Miguel Vallejo Orti.)*

Miguel Vallejo Orti is with the Department of Geo-Spatial Sciences and Technology, Namibia University of Science and Technology, Windhoek 13388, Namibia, and also with the 3D Geospatial Data Processing Group, Institute of Geography, Heidelberg University, 69120 Heidelberg, Germany (e-mail: morti@nust.na, morti@uni-heidelberg.de).

Lukas Winiwarter, Jack G. Williams, and Bernhard Höfle are with the 3D Geospatial Data Processing Group, Institute of Geography, Heidelberg University, 69120 Heidelberg, Germany (e-mail: lukas.winiwarter@uni-heidelberg.de; jack.williams@uni-heidelberg.de; hoefle@uni-heidelberg.de).

Eva Corral-Pazos-de-Provens is with the Departamento de Ciencias Agroforestales, Universidad de Huelva, E-21819, Huelva, Spain (e-mail: eva.corral@dcaf.uhu.es).

Olaf Bubenzer is with the Geomorphology and Soil Science, Institute of Geography, and also with the Heidelberg Center for the Environment, Heidelberg University, 69120 Heidelberg, Germany (e-mail: olaf.bubenzer@uni-heidelberg.de).

Digital Object Identifier 10.1109/JSTARS.2020.3040284

***Index Terms*—Arid regions, automatic classification, gully erosion, iterative learning, land degradation, Namibia, random forest (RF), soil erosion mapping.**

I. Introduction

SOIL health is a global concern affecting modern and traditional societies and is a fundamental aspect that sustains life on Earth [1], [2]. Gully erosion represents a key driver of landscape and soil degradation, and it is a dominant geomorphic process present in many regions worldwide [3], [4]. Gullies are responsible for the loss of large masses of sediment from river catchments and agriculture plots [5]. Given that they threaten important infrastructures such as settlements, roads, and cattle dams, these effects are considered more pronounced as compared to other soil erosion types [6].

With growing land pressure within Africa and changes in patterns of rainfall intensity, an increase in gully erosion rates [7] is projected to occur in the coming decades [8]. Arid and semi-arid countries suffer land degradation in an extreme way [9]–[11], and Namibia represents a significant example of this. Gully erosion in Namibia is considered one of the largest overlooked environmental problems, with critical consequences including desertification, loss of habitats [12], [13], lack of food security, social conflicts in rural areas, and urban migrations [14], [15] affecting national economic development and ecological wealth [16], [17]. In recent years, considerable scientific effort has been directed in Namibia towards controlling [18] and understanding general land degradation [19], [20], in particular, that associated with gullies [21]–[23]. There are also attempts to carry out inventories and distribution maps of gullies as well as land degradation by individual researchers [24], farmers associations [25], and the Ministry of Agriculture, Water and Rural development of Namibia [26], [27]. Mapping initiatives are generally carried out at the local (farm) level and form a valuable basis for a better understanding of the gully phenomena. A more global approach, however, is required to tackle the problem as a critical hazard affecting the entire nation. In this sense, as highlighted by Poesen [5] for a global scale, research activities in Namibia should also cover the development of standardized and reliable measuring techniques for gully identification and mapping at

large spatial scales, which remain scarce. Nevertheless, most of the techniques applied for micro- and catchment scales are not applicable at regional, national, or continental scales [28], with datasets suited to large gully-affected areas generally being less explored [29]. This paucity of suitable data is particularly acute in countries such as Namibia, where the very low population density, limited technology, and difficulty in accessing affected areas hinder campaigns for direct measurement in the field [30], [31].

The development of techniques for gully mapping in large areas has been investigated from several perspectives [3], [32]. Following a rough classification, traditional methods focus on the estimation of gully risk, density, or susceptibility using empirical model-based approximations [33], [34]. More recent approaches draw on direct measurements, generally based on combinations of remote sensing methods [29], [35]–[37], Geographic Information Systems [38], and Machine Learning (ML) [39]–[41]. Empirical model-based estimations have provided impressive results using gully density patterns derived in existing studies. One such model [42] has reasonably predicted (correlation coefficient $r = 0.71$ and $r = 0.78$ for the presence of gullies in 5 and 10 km resolution grids, respectively) gully density at the regional or sub-continental scale in Australia using multivariate modeling, but does not include accurate gully mapping and object identification. One of the most recent and successful approaches to gully mapping was conducted by [43], who developed a continent-scale regression model for Africa based on topography, rainfall, and vegetation cover, which simulated 57% of manually mapped gully heads.

The resolution and scale of datasets, both in the temporal and spatial domains, are increasingly commensurate with the scale and rates of gully change [3], [44], [45]. Visual interpretation of panchromatic Pleiades B imagery [29] has enabled accurate quantification of gully area (mean error 11.1%) and perimeter (mean error 8.9%) in the Loess Plateau (China), yet uncertainty remains regarding the level of required expertise and pixel tolerance during digitization. Approaches that are entirely manual become increasingly laborious to implement as the scale of mapping increases, with automatic gully detection using satellite imagery providing an alternative that is under development. For example, Vallejo *et al.*, [46] developed three methods to detect large gullies in central Namibia based on gully morphology. When tested using a TanDEM-X (TX) Digital Elevation Model (12 m spatial resolution), the approaches achieved total accuracy (TA) of $\sim$80% and a $\sim$50% Kappa value. Efficient mapping of large areas, combining vector (drainage network), and Google Earth (GE) raster data using a random forest classifier (RFC), has also yielded high accuracy levels (79%; [47]). A limitation arises when gully areas exhibit large internal morphological variability and imagery hosted in GE cannot provide a sufficiently clear spectral signal to differentiate gullies. This drawback can be mitigated by introducing additional descriptive gully parameters based on topography, hydrology, local climate, vegetation, and soil characteristics. There is consensus that gully formation is principally controlled by slope, land use, rainfall, stream proximity and lithology [33], [48]–[50], and gullies produce changes in terrain roughness [51] and vegetation patterns [52]. Data-driven approaches, which draw on these parameters through ML, have demonstrated that RF, Support Vector Machine (SVM), and Logistic Model Trees are efficient means of deriving gully density and location in order to generate gully susceptibility maps [10], [11], [39], [40]. They nevertheless share some of the following limitations:

i) they normally require reliable, large, and evenly distributed training data;
ii) they usually do not consider the dynamic character of the gully while building the explanatory variables; and
iii) they define gullies as a point or line locations, but more geomorphologically relevant features such as outline and volume are often not derived.

Attempting to address these limitations is one of the aims of this research.

Two studies have faced the challenge of deriving gully outlines, still under the umbrella of ML methods, but including object-based image analysis (OBIA) [41], [37]. Liu *et al*. [37] performed a two-level gully feature mapping (gully-affected areas and bank gullies) using a 1-m DEM and Worldview-3 imagery. Object-based and RF methods were combined using four topographic parameters for a catchment in the Chinese Loess Plateau. It resulted in a 93% overall accuracy in gully affected areas and 78.5% for bank gully mapping. Shahabi *et al*. [41] tested artificial neural networks, SVM, and RF integrated with geographic object-based image analysis (GEOBIA) in Australia using 12 m ALOS-2 DEM and Sentinel-2 products. An 84% accuracy for gully detection and between 30% and 70% overlap with existent gully inventories was achieved in this case. Although [41] and [37] represent a very valuable reference, the constraint of limited and spatially clustered training data, common to arid/semi-arid regions, must be addressed in order to derive practical solutions for Namibia. Gullies are landforms with large internal heterogeneity that adopt diverse typologies. Insufficient training data would create a narrow initial definition of a gully class. Accurate, class balanced, and reliable training is fundamental for a single pixel-based classification algorithm (i.e., RF). Thus, considering that gullies are continuous structures, the proposed procedure ought to conduct successive searches to strategically broaden the definition of gullies, considering the pixel values and their spatial arrangement. As this search proceeds, the classification criteria can be optimized through object-based spatial analysis and the adaptation of configuration settings in the classification algorithm. In this vein, we propose an automatic iterative learning-based approach, which is inspired by active learning [53], [54] and in line with other iterative training techniques [55]. Training data are dynamic and changes according to a continuous learning process that begins with a small gully portion, surveyed in the field, or digitized using high-resolution imagery in a micro-mapping approach [56]. As gully conditioning factors, elements that are statistically favorable to gully appearance (slope and distance to stream) and those measurable descriptors of gully presence impacts (vegetation patterns, soil characteristics, and terrain roughness [57]) are considered together. Given that gullies are dynamic features, behavioral indicators based on temporal patterns [i.e., Normalized Difference Vegetation Index (NDVI),

radar coherence [58], and backscatter intensity time series] are included as gully descriptors. This is important in light of the fact that trends of historical data differ between adjacent areas [57], [59], [60].

The algorithm, based on RFC and GEOBIA, includes different global remote sensing data sources [TanDEM-X DEM (TX), Sentinel-1 SAR (S1), and Sentinel-2 imagery (S2)]. By applying and validating our approach in two watersheds (WS1 and WS2) in Kunene region (Namibia), we aim to address the problem of classifying and outlining gully-affected areas for large regions using little training data.

II. Study Area and Data

A. Study Area

Our study site is located within the Kunene Region, northwest Namibia, where the presence of gullies is widespread [18]–[20]. The town of Opuwo represents the highest population density spot in the area, nevertheless, many small villages and settlements are scattered in the region. This area has special ethnographic value since it hosts most of the Himba and Herero traditional communities [61]. This territory's governance falls under communal land [62].

The dominant ecosystem type is Western Highlands [64], [65] characterized by grasslands and scattered trees. Terrain elevation ranges between 823 and 1883 m a.s.l., with an average elevation of 1392 m. The slopes are irregular with an average value of 19% (12 m pixel size) and exhibit a mixture of hilly and rolling topography dissected by many tributaries. The dominant vegetation biome in this area is tree and shrub Savanna. In the north-eastern area, the vegetation structure is of woodland type, made up of low trees, which transitions to sparse shrubland toward the southwest, where climate and edaphic conditions are more adverse [66]. *Acacia reficiens* is the dominant species, accompanied by *Commiphora* spp. [63], [67].

The study site is categorized as arid, with $<$500 mm of annual rainfall and a highly variable spatial distribution. Rain falls mostly as summer storms between December and March ($\sim$80% of the annual precipitation). The average annual temperature varies between 19 and 21 °C, with minima between 6 and 10 °C and maxima between 28 and 36 °C, registering high annual potential evaporation rates (3000–3200 mm) [68].

The prevailing land use is communal crop (i.e., corn and pearl millet), goat and cattle livestock farming [19]. The ecosystem is very fragile with land degradation derived from overgrazing. The cattle density is highest in the area´s northeast corner, with $\sim$25 animals km^{-2}, and decreases with distance from Opuwo [see Fig. 1(a)] [63].

Soils are weakly developed, with exposed bedrock in some places. The dominant soils are Leptosols with Eutric and Lithic characteristics, accompanied by Regosols (Eutri-Skeletic and Leptic). In the northwest corner of the study area, the Lithic Leptosols are surrounded by a fringe of Petric Calcisols, and, in the east, occasionally Arenosols and Cambisols also appear [19], [69].

The two WSs extracted for analysis (see Fig. 1) have been selected due to the erodible soil types, significant human presence, proximity to populated areas and main roads, and high livestock density relative to the surrounding area (see Table I).

TABLE I
Main Topographic, Environmental, and Land Use Characteristics of the Two Selected WSs [33], [69]

	Watershed 1 (WS1)	Watershed 2 (WS2)
Total area (km^2)	408.9	302.84
Elevation (m a.s.l)	1166 – 1850	1027 - 1883
Average Slope (%) – 12 m pixel size	14.9	11.1
Average annual temperature (°C)	21 – 22	20 – 22
Cattle density (km^{-2})	Highly variable, 1 - 20	3 – 14
Livestock density (kg/ha)	Highly variable, 0 - 99	0 – 79
Vegetation structure	Woodland	Woodland and sparse shrubland
Vegetation cover (dominant species)	Savanna (*Acacia reficiens*)	Savanna (*Acacia reficiens*)
Soil type	Leptic and Eutri-Skeletic, Regosols with Eutric Leptosols, and Rubi-Ferralic Arenosols	Leptic Regosols and Eutric Leptosols

B. Input Datasets

The gully descriptive parameters used as inputs in our model have various sources. S1 (2017–2019) and S2 (2018–2019) images captured at different times are used to compute time series for NDVI, radar coherence, and sigma backscatter VV intensity. TX imagery is used to generate slope, topographic roughness, stream channel maps, as well as two gully shape factors [Multi-Profile Curvature Analysis (MPCA) and Inverted Morphological Reconstruction (IMR)] developed in [46]. To generate validation and reference data to conduct accuracy assessments, gully affected areas are manually digitized from satellite images (Bing Aerial with ground sampling distance below 0.5 m, corresponding to the period December 2013 to November 2016) by two operators with experience in gully identification. A 30-m grid map of gully distribution with two categories, 1) gully and 2) no-gully, is generated [see Fig. 7(b) and Fig. 10(b)].

The spatial resolution, acquisition interval, and corresponding gully feature are listed in Table II for each product.

Preprocessing was required to generate the final products used as descriptive gully parameters (see Fig. 2).

III. Methods

A. General Framework

The main challenge faced in mapping land cover and producing gully-affected maps in low populated regions is the scarcity of training data and the capacity to create it in order to achieve high classification accuracies [55]. To overcome this limitation, our approach generates a new set of gully training samples with each iteration k and it continuously grows in size. The different explanatory variables [see Fig. 3(2A-F)] are integrated into a pixel-based RF classification algorithm. Small manually tagged

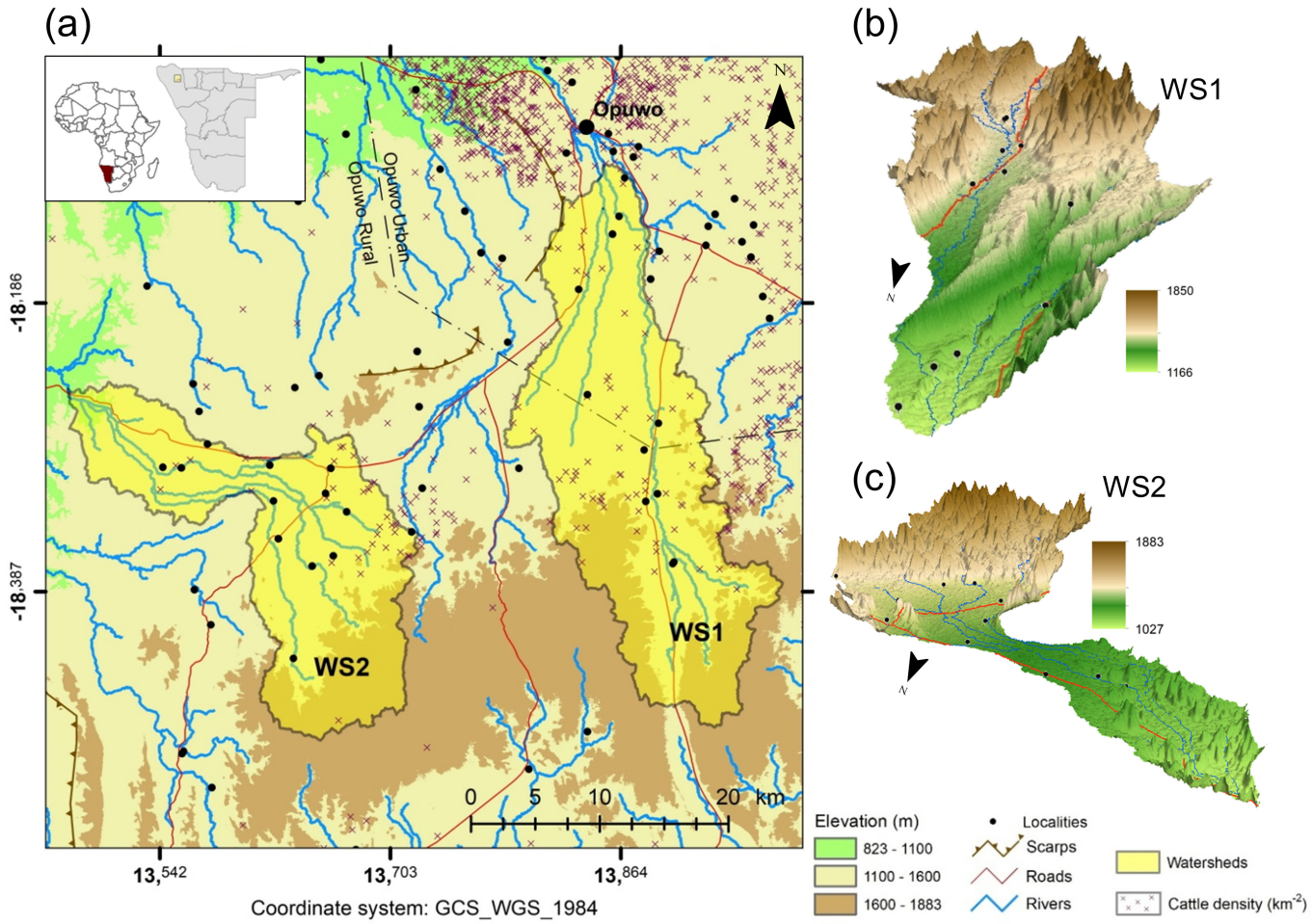

Fig. 1 Gully-affected areas selected in this study. a) Map of the study area and two WSs (WS1 and WS2) selected for analysis including cattle activity as a possible trigger of land degradation; (b) and (c) 3D views of the WSs with main settlements, rivers, and roads [63].

gully portions are used as initial training data [see Fig. 3(1)]. Using an alpha shape algorithm [70], a subsequent geometrical identification of gully-candidate pixels is used to generate gully objects raw [GORs; Fig. 3(5)]. This operation dismisses isolated pixels outside GORs, thereby mitigating salt and pepper effects and reducing RF false positives rates. Pixels inside GOR have an associated probability of being a gully derived from the RF (low, medium, and high). The pixels classified as "very high" are categorized as gully seeds and stored in a seeds pool. A subset of this pool (50% random selection) is added to the training data in subsequent RF iterations [see Fig. 3 (4)]. To reduce the effect of outliers, the gully objects clean (GOC) are calculated as the overlapping area between the GORs of two consecutive iterations *k-1* and *k* [see Fig. 3(5)]. This region (GOC) represents the gully label in the training data for the following iteration *k+1*. The iterations are repeated until an optimal situation in terms of accuracy is achieved. Three separated (alpha shape) gully outlines are generated from the three-pixel groups based on their RF derived gully probability (high, medium, low). Finally, gully outlines are refined using DEM-based morphological operators. The general workflow, including key operations, is presented in Fig. 3.

B. *Methods Description and Settings*

The algorithm comprises eight main steps:

1) initialization;
2) gully feature extraction;
3) gully pixel generation (RF classifier);
4) seed pool generation;
5) object-based classification;
6) analysis/decision; and
7) final classification.

One additional step, 8) for refinement is included outside of the main workflow as an experiment:

1) Initialization: Georeferenced gully data collection in remote locations can be logistically difficult and costly, but a minimum amount of gully labeled data is required as initial training datasets. For initialization of the process, potential data sources such as handheld-GNSS, UAV-borne photogrammetry (e.g., structure from motion), and basemap digitization can be utilized. Assessing the least size of an initial training dataset is of value from a methodological standpoint. In order to do so, we simulated a gully delineation where the area that could be examined within 15–30 minutes using the three abovementioned

TABLE II
DESCRIPTION OF THE INPUT DATA SOURCES USED IN THE STUDY AS CONDITIONING FACTORS OF THE RF CLASSIFIER FOR THE GULLY DETECTION

Source type	Spatial Resolution	Feature name	# Images	Time	Gully feature measured
S1 (GRDH)	10 m	(A)Intensity VV Mean	30	Jan 2017 – Nov 2019	Terrain Roughness/ texture, Soil Moisture,
S1 (SLC)	10 m	(B)Coherence VV Mean	44	Apr 2017 – Oct 2019	Terrain Change detection
S2 (MSI2A)	10 m	(C)NDVI Standard Deviation	18	Feb 2018 – Sept 2019	Vegetation patterns
TanDEM -X (HRTI-3)	12 m	(D)Slope		2015	Topography
		(E)Topographic roughness			Topography
		(F)Distance to Stream			Hydrological
		(G)Shape factor I: IMR			Gully shape factor
		(H)Shape factor II: MPCA			Gully shape factor

Preprocessing was required to generate the final products used as descriptive gully parameters (see Fig. 2).

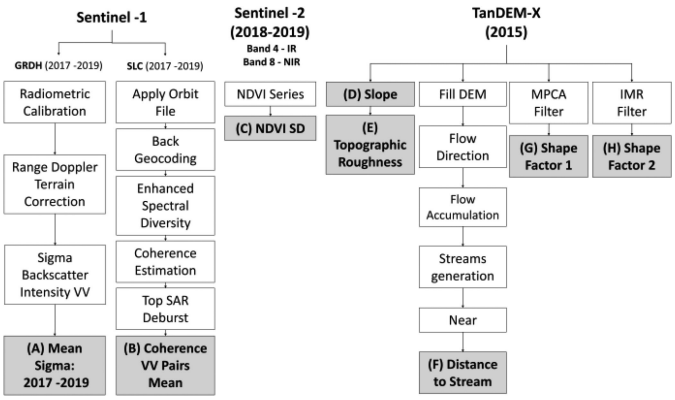

Fig. 2. Primary data sources and pre-processing of explanatory variables datasets. Ground Range Detected High Resolution (GRDH) and Single Look Complex (SLC) radar products from S1 are used to derive Radar Backscatter and Coherence. S2 Multi-Spectral Instrument (MSI) products are used for NDVI calculation. TanDEM- X 12 m HRTI-3 Global DEM is the primary data to derive topographic, hydrologic, and morphological factors.

methods was approximated by manual digitizing. Bing Aerial basemap was used to manually delineate a gully (polygon), while the remaining area of the WS was initially labeled as "no-gully".

2) Gully Features Extraction: This initially digitized small area is labeled as "gully." The remaining area of the WS is temporally labeled as "no-gully." Both labels are extracted on a 30-m regular grid sampling with a total amount of *n* binary samples (c_i), together with a set of explanatory variables values ($v_{i,j}$): radar backscatter intensity VV polarization mean (S1) ($v_{i,1}$), radar VV coherence mean (S1) ($v_{i,2}$), NDVI Standard Deviation - SD (S2) ($v_{i,3}$), slope (TX) ($v_{i,4}$), topographic roughness (TX) ($v_{i,5}$), and 3D distance to stream (TX) ($v_{i,6}$). With this information and the corresponding Cartesian data (x_i, y_i) for each sample point of the grid, the input data D_n is

$$D_n = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_i \\ \vdots \\ X_n \end{pmatrix} \quad \text{with}$$

$$X_i = (x_i\ y_i\ c_i\ v_{i,1}\ v_{i,2}\ v_{i,3}\ \ v_{i,4}\ \ v_{i,5}\ v_{i,6})\,. \tag{1}$$

3) Pixel-Based Classification: This input dataset (D_n) is separated into two subsets, gully (D_g) and no gully (D_{ng}). Training data (D_{tr}) is built after a random selection from each subset in different proportions (80% for D_g and 20% for D_{ng}). As we aim at iterative training, we chose a moderate value of 20% for D_{ng} in order to allow reclassifications in the no-gully (unlabeled) area, while the 80% in D_g is needed to capture the internal variability of the gullies, including areas in the shade or occluded by rough topography. D_{tr} is used to train a RFC (maximum tree depth, $d = 100$, and number of trees, $M = 100$). A reclassification of the whole dataset D_n into gully pixels or no-gully pixels is then conducted. RF assigns a probability ($m_{M,i}$) to each data point (X_i) calculated as

$$m_{M,i}\ (X_i; \theta_1, \ldots, \theta_M, D_n) = \frac{1}{M} \sum_{t=1}^{M} m_i\ (X_i; \theta_t, D_n) \tag{2}$$

where θ_t is one of the random variables and m_i denotes the "gully" ($m_i = 1$) or "no gully"($m_i = 0$) vote of one single tree *t* [71]. This probability $m_{M,i}$ is derived from the fraction of positive votes and the number of trees (100) to define gully ($m_{M,i} \geq \frac{50}{100}$) and no-gully classes ($m_{M,i} < \frac{50}{100}$) as well as to identify the different class probability of gully pixels: low ($\frac{50}{100} \leq m_{M,i} < \frac{67}{100}$), medium ($\frac{67}{100} \leq m_{M,i} < \frac{84}{100}$), and high ($m_{M,i} \geq \frac{84}{100}$).

During training, the degree to which a feature decreases the weighted impurity (Gini Impurity) is calculated to derive each feature's importance. Gini Impurity "*represents the probability (G) of incorrectly classifying a randomly chosen element X in the dataset if it were randomly labeled according to the class distribution in the dataset*" [72]. The feature importance $I(V_j)$ indicates the sum of the decrease in Gini Impurity *(G)* over all the nodes split on that feature *j* [73].

4) Seed Pool Generation: Pixels classified as gullies with $m_{M,i} \geq \frac{95}{100}$ are categorized as seeds and appended to the seeds repository (gully pixel seeds) which is regenerated after each iteration. A new random sample of this repository (50%) is added in each subsequent iteration to the randomly selected training data of the RF, to ensure that a minimum amount of pixels with high gully class probability are included in the training, but avoiding possible false positives used systematically as seeds in the training data throughout the whole process.

5) Object-Based Classification: To outline gully objects, the 2D alpha-shape algorithm is applied to the initial pixel-based classification. This technique creates an outline around grouped gully points (pixels) based on the following topology rules:

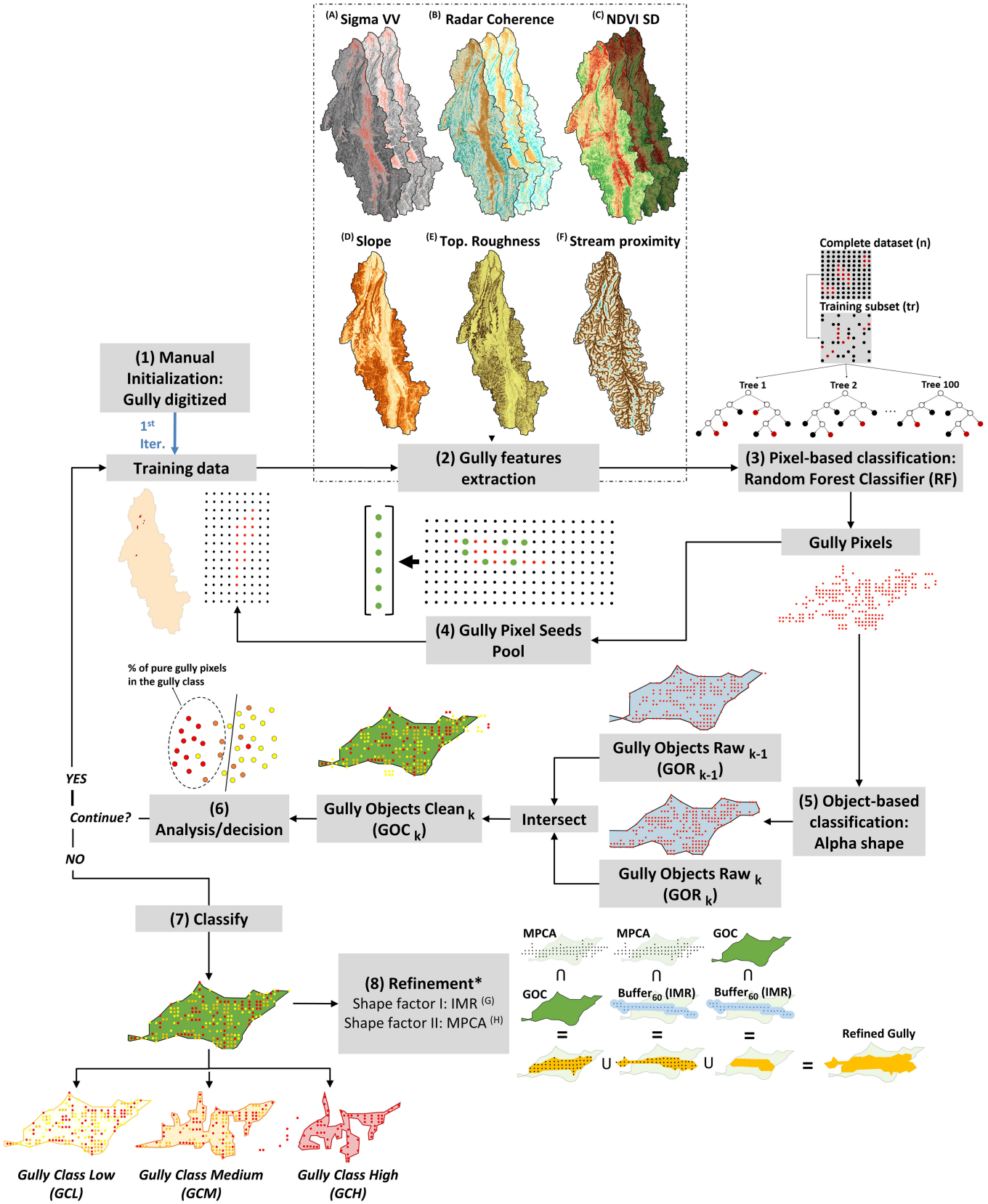

Fig. 3. General classification algorithm diagram with datasets identified with letters (A–H) and important procedure with numbers (1–8). The diagram explains how the algorithm works from the initial manual label of gully sections (1) until the final classification results (7) and refinement (8).

considering each of the two possible circumferences with radius alpha that contain each point pair of the dataset, if at least one circumference does not enclose any other point from the dataset, it is assumed that the straight line connecting the point pair is a boundary of the alpha shape. If another point is contained within either circumference, the connection between the two points is not a boundary of the alpha shape, but the point pair is contained by the alpha shape. Moreover, if one point does not share an alpha-radius circumference with any other point from the dataset, this point is set to fall outside of the alpha shape. Alpha determines the level of detail of the final shape. The larger the alpha, the more points will be enclosed by the shape. If alpha $\rightarrow\infty$, the shape will be approximated to a convex hull of all the points [70]. Here, the value alpha is initialized as 180 m based on the fact that our sampling grid width is 30 m and maximum gully widths are also approximately 180 m in our study area. Therefore, initially, shapes are created out of points separated six times the original grid width. As our gully points density increases, the value of alpha is decreased, and our alpha shape becomes more detailed. This operation generates the polygons above called GOR. For some distributions of gully points, the alpha shape occasionally generates artifacts enclosing empty areas; for example, in the space between individual gully pixels clusters. To remove artifacts in the classification, a GOR generated in the iteration *k* is intersected with the GOR generated in the previous iteration *k-1,* generating a GOC, depicted as follows:

$$GOC_{\mathrm{k}} = \mathrm{GOR}_{\mathrm{k}} \cap \mathrm{GOR}_{\mathrm{k-1}}\,. \tag{3}$$

6) Analysis/Decision: In those cases when sufficient testing data is not available, a critical aspect to achieve optimal results is the definition of criteria to halt subsequent iterations based on indicators derived from the quality of classification. In order to explore algorithm intrinsic quality indicators, the result of the classification is analyzed (4) after each iteration to calculate the percentage of high probability gully pixels inside the GOC, *Q* (5):

$$\begin{aligned}
&\left(\frac{50}{100} \le m_{M,i} < \frac{67}{100}\right) \rightarrow i \in C_{\mathrm{low}}\\
&\left(\frac{67}{100} \le m_{M,i} < \frac{84}{100}\right) \rightarrow i \in C_{\mathrm{med}}\\
&m_{M,i} \ge \frac{84}{100} \rightarrow i \in C_{\mathrm{high}}
\end{aligned} \tag{4}$$

$$Q = \frac{n_{C_{\mathrm{high}}}}{n_{C_{\mathrm{high}}} + n_{C_{\mathrm{med}}} + n_{C_{\mathrm{low}}}} \tag{5}$$

where n_C represents the amount of gully pixel of each class (high, medium, and low). This point is extended in Section III-C.

In order to assess in detail the accuracy of the results and understand the model performance, in addition to Confusion Matrix values, the receiver operating characteristic area under the curve (ROC AUC) [74] and the Matthews Correlation Coefficient (MCC) [75], [76] are derived for each iteration. Extended information about these indexes is included in Appendix III.

7) Final Classification: Once an optimal accuracy scenario is detected, the gully points inside the GOCs are separated into three classes based on the associated probability, from less to more restrictive (low, medium, and high): $C_{\mathrm{high+med+low}}(\, m_{M,i} \ge \frac{50}{100}\,)$, $C_{\mathrm{high+med}}$ $(\, m_{M,i} \ge \frac{67}{100}\,)$, and C_{high} $(\, m_{M,i} \ge \frac{84}{100}\,)$. Subsequently, an alpha shape operation is applied to each point dataset individually, generating gully outlines with three levels of exposure for the gully incidence: Gully High Class, Gully Medium Class, and Gully Low Class.

8) Refinement: This is included as an experiment in the method and is a refinement based on the results of the DEM-based gully classification conducted in [46]. IMR results show very high user accuracy (UA) for gully class and MPCA performs well in outlining gully shapes as its producer accuracy (PA) is generally >50% for the gully class. Using GOCs resultant of the general methodology, an additional methodology is included where we:

1) apply 60 m buffers around IMR pixels;
2) compute the intersection between the MPCA pixels with IMR buffered area and calculate its alpha shape;
3) calculate the intersection between the MPCA pixels and the resultant GOC and calculate its alpha shape;
4) compute the intersection between the GOC and the IMR buffered area; and
5) apply a union between 1), 3), and 4) to generate a final refined gully class (6).

$$\begin{aligned}
\text{Refined Gullies} = {} & \mathrm{alpha}(\mathrm{Buffer}_{60}(\mathrm{IMR}) \cap \mathrm{MPCA})\\
& \cup\ \mathrm{alpha}\,(\mathrm{GOC} \cap \mathrm{MPCA})\\
& \cup\,(\mathrm{GOC} \cap \mathrm{Buffer}_{60}(IMR))\,.
\end{aligned} \tag{6}$$

C. Iterative RF: Control and Stopping Criterion

The main hypothesis of active learning approaches is their ability to select the data they want to learn from according to a specific strategy and gradually add new training samples to enhance the classification power [54]. In this automatic iterative learning approach, the algorithm learns from the previous iterations to configure the settings of the subsequent one. In this line, the main parameter controlling the algorithm is the alpha shape threshold (alpha). Two variables are used to modulate alpha to allow growth in the GOC while avoiding GOR with low densities of gully points: the average GOC area covered by each gully pixel and the growth rate between the GOCs of two consecutive iterations. A key point is to control the amount and quality of newly generated gully pixels to avoid exponential growth of gully labels and ensure that only the most reliable data are selected. In order to create a dynamic operation and automate the learning process, the rules listed below are used to govern and control the algorithm.

1) The ratio ($\mathrm{r} = \frac{\text{Gully Pixels}_{\,k}}{\text{Gully Pixels}_{\,k-1}}$) between the number of pixels classified as gullies by RF in the iteration k and those of the previous iteration $k-1$ remains between 1.0 and 1.5. If $r > 1.5$, points with the lowest gully probability are removed until the ratio lowers to 1.5.

TABLE III
DESCRIPTION OF INPUT TRAINING DATASETS WITH DIFFERENT SIZES FOR THE TWO WSS AND THE EFFECT ON THE CLASSIFICATION ALGORITHM CONSIDERING THEIR BEST ACHIEVED ACCURACIES [PA, UA FOR GULLY (G) AND NO GULLY (NG) CLASSES, PLUS TA, COHEN KAPPA (KA) AND MCC] AND THE ITERATION WHEN THIS ACCURACY WAS ACHIEVED

	Total samples (n)	Simulated Collection Type	Initial Gully sampling area (km^2)	Initial Gully samples	Training area/ Total WS area (%)	Training area/ Real Gully area (%)	Iter. (k)	Highest Accuracy						
								PA (G)	PA (NG)	UA (G)	UA (NG)	TA	KA	MCC
WS1	454226	Digitizing	0.76	845	0.20	10.88	11	0.540	0.994	0.669	0.990	0.984	0.590	0.593
		UAV	0.44	489	0.06	6.30	13	0.554	0.993	0.629	0.990	0.983	0.581	0.581
		GNSS	0.06	71	0.01	0.85	2	0.003	0.999	0.826	0.979	0.979	0.007	0.056
WS2	336410	Digitizing	0.59	659	0.19	25.21	9	0.568	0.994	0.482	0.995	0.990	0.517	0.518
		UAV	0.26	292	0.08	11.11	9	0.518	0.995	0.547	0.995	0.991	0.528	0.528
		GNSS	0.09	107	0.03	3.84	2	0.020	0.999	0.822	0.990	0.990	0.039	0.128

2) The area increment (ΔGOC) between the GOCs of two consecutive iterations must be kept between 40% and 10%.
3) A maximum average area density $\text{a/p} = \frac{\text{Area}(GOC)}{\text{Gully Pixels} \cdot 900}$ should be maintained to ensure a minimum density of gully pixels inside the GOC. The threshold ratio is 3600 m^2/point (four times 900 m^2, the maximum average area density in the ideal situation of 30 m regular grid).
4) The alpha shape begins with a radius alpha of 180 m and progressively modifies its value to meet the requirements set in b) and c). If alpha results in 0%< ΔGOC<40% while the rate (a/p) is <4.0, the algorithm accepts the result and proceeds to the next iteration. In any other case, alpha shape operation is repeated with a new alpha until the condition is met. The criterion to adjust alpha is presented as

$$\begin{aligned}
&\text{if:}\frac{\text{GOC}_k}{\text{GOC}_{k-1}} > 0.4 && \text{alpha}_\text{k} = \text{alpha}_{k-1} - 5\\
&\text{if:}\frac{\text{GOC}_k}{\text{GOC}_{k-1}} < 0.0 && \text{alpha}_\text{k} = \text{alpha}_{k-1} + 3\\
&\text{if: } 0.0 < \frac{\text{GOC}_k}{\text{GOC}_{k-1}} \left\langle 0.4 \text{ and} \frac{\text{a}}{\text{p}} \right\rangle 4 && \text{alpha}_\text{k} = \text{alpha}_{k-1} - 5.
\end{aligned} \tag{7}$$

Reliable criteria for halting iterations underpin the accuracy of gully identification. Due to the previous operational decisions, while a seed repository is created and used as training data, more points are likely to be classified under the category of high probability gully pixels. New gully areas are often formed out of lower probability pixels; however, when an area is stable and no new gullies are classified, the high probability gully pixels are added to the GOCs progressively. Once these points reach a certain classification maturity, generally 35%–45%, iterations can be halted (see Fig. 4).

IV. RESULTS

A. Initialization

In Table III, the summary of the classification result for the simulation of the three cases of gully class initialization (by manual digitizing of satellite imagery, UAV-based photogrammetry, or by GNSS collection) is presented.

It is noticeable that above a certain threshold, smaller initial training datasets do not affect the results obtained significantly.

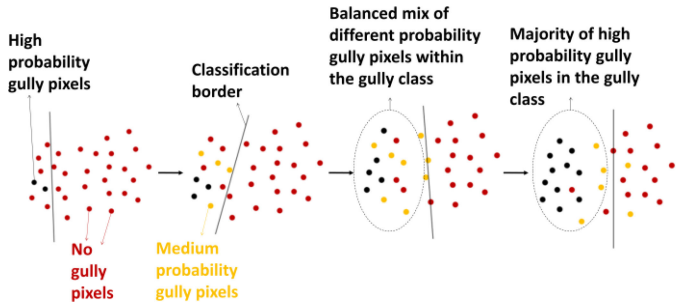

Fig. 4. Simplified graphical explanation of the typical dynamics in the classification results from initial stages until an expected maturity in the iterative classification process.

However, a lower bound occurs when the size of the gully class is insufficient to represent the gully object. In this case, as the algorithm is unable to continue from the second iteration, the gully class decreases until it disappears in subsequent iterations. The results presented below in this section are obtained using the digitizing mode (see Table III) as initial training data.

B. General Classification Performance

We provide an assessment of the algorithm's performance for each WS, with a particular focus on the output metrics outlined in Sections III-B and -C.

1) Watershed 1: Results from all iterations are presented in Fig. 5 and Appendix I, with 22 iterations undertaken before the condition of Q > 60% of high probability gully pixels ($m_{M,i} \geq \frac{84}{100}$) was met. Iteration 11 achieved the highest accuracy and was high performing across all metrics, with >0.5 for the gully class in PA (0.540) and UA (0.669), as well as the highest Kappa value (0.590) and MCC (0.593).

The evolution of the three variables that describe the classification progress (GOC growth, a/p, and gully points increment) is presented in Fig. 5(a). The red line reflects the evolution of the average area covered by each point in proportion to the initial grid of 30 m (a/p). As it is assumed according to the configuration, this indicator is subject to initial adjustments until it achieves a stable value close to 1.2. The black line represents the growth rate (r) of new points generated as gullies by the RF classifier, which remains below 1.5. The blue bars show a change to the GOCs, which remains between 1.0 and 1.4. The

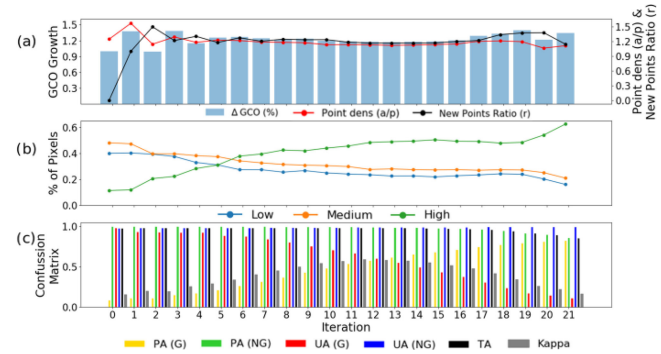

Fig. 5. Algorithm function and accuracy performance for WS1. a) Black line represents the ratio of classified gully pixels by RF between iteration k and $k-1$. Red line is an indicator of gully pixels density inside GOC in units of their proportion to the original grid of 30 m. The bar chart represents the growth ratio for each iteration of the GOC. b) Each line color represents the percentage of pixels created by the RF based on the probability ($m_{M,i}$) of being a gully: blue (C_{low}), orange (C_{med}), and green (C_{high}). c) Accuracy assessment including UA, PA, TA, and Cohen KA for each iteration.

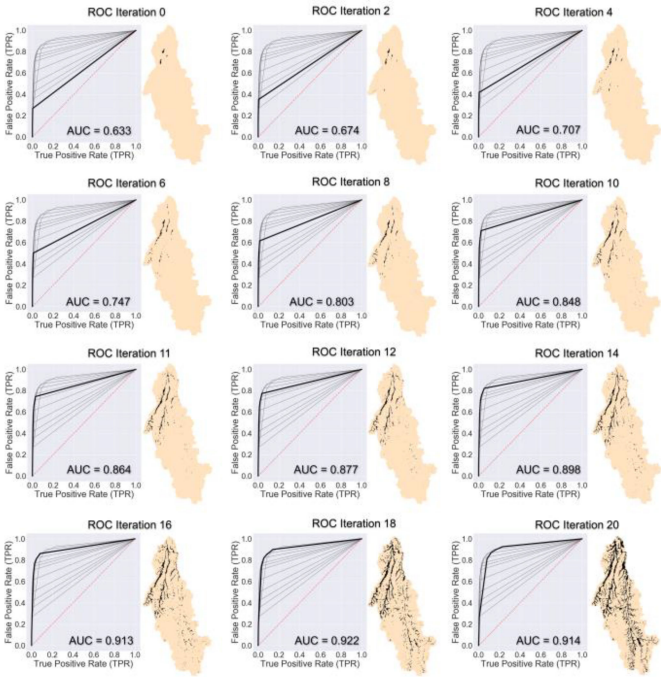

Fig. 6. ROC AUC analysis results for WS1 with ROC curves for the iterations 0, 2, 4, 6, 8, 10, 11, 12, 14, 16, 18, and 20 including the AUC values in the legend and corresponding cartographic outputs of gully class.

proportion of pixels within each quality class (low, medium, and high probability of being gully candidates) is presented in Fig. 5(b). High probability gully pixels (>83% probability) are initially scarce but grow in number until their percentage reaches 35%, 40%, 45%, 50%, and 55% in iterations 6, 8, 11, 15, and 21, respectively. The highest accuracies occur between iterations 9 and 14 when the percentage of high probability gully pixels ranges between 42.7% and 49.1%. Beyond the 14$^{\text{th}}$ iteration, a high false-positive rate of gullies is observed, resulting in a drop-down in Kappa value and UA for Gully class.

Analyzing the outcomes of the RF and the mean significance of each feature from iteration 0 to 11, we can observe that the most important variables when the algorithm achieves its optimal (iteration 11) are radar VV coherence (22.0%), NDVI SD (21.7%), and distance to stream (16.5%; see Appendix I).

Values of ROC AUC shows that the classification model for WS1 starts with values close to 0.6 for the iteration 0, arising to the maximum (AUC $=$ 0.922) in the iteration 18, where a decreasing tendency begins. These results confirm how the model progressively improves its classification capacity from iteration 0 to iteration 11 (highest classification result according to confusion matrix measures), where it reaches AUC $=$ 0.864, meaning that the probability of classifying correctly any point pair, one gully and one nongully, is 86.4%.

Comparing the results at the four predefined milestones for the proportion of high probability gully pixels (35%, 40%, 45%, 50%, and 55%) it is observed that the best results are obtained for iteration 11 when it first exceeds the 45% threshold. Results are presented for the different gully classes as calculated from this iteration in Fig. 7.

The low, medium, and high-quality gully classes cover 9.0, 6.7, and 5.0 km^2, representing 2.2%, 1.6%, and 1.0% of the WS surface, respectively. The main gullies are distributed in the lower regions of the WS to the north. Accuracy results for the three classes are provided in Appendix I, under iteration 11. Values in excess of 0.988 are obtained for the no-gully class (UA and PA), while the gully class reaches 0.598 in PA and 0.747 in UA for the high and low class, respectively. This highlights that generating outcomes into three classes increases the accuracy as compared to a single class, which would otherwise have attained PA $=$ 0.540 and UA $=$ 0.669 for the same iteration. This is consistent with the visual interpretation of the polygons in Fig. 7(c)–(e), where non-gully areas are accurately labeled, thereby avoiding high false-positive rates that are common in many approaches.

The classification is most accurate in areas where the gully is morphologically homogeneous, particularly with respect to the width and longitudinal continuity, but accuracy is reduced with intricate shapes and heavily branched gully areas [see Fig. 7(c) and (d)]. Another common error source is the presence of structurally stable narrow valley bottoms with V-shapes, visible in Fig. 7(e), which may cause the algorithm to classify them as gullies.

2) Watershed 2: For WS2, the algorithm carried out 19 iterations until reaching 60% of high probability gully pixels. The 35%, 40%, 45%, 50%, and 55% proportions were achieved in iterations 7, 10, 13, 14, and 16, respectively. The highest accuracies results are attained between iterations 7 and 10 when the indicator Q ranges between 36.5% and 41.8%. As indicated by the maximum values of Kappa (0.517) and MCC (0.519), the highest accuracy is achieved in iteration 9 with the following values: PA (Gully) $=$ 0.568, PA (No-gully) $=$ 0.994, UA (Gully) $=$ 0.483, UA (No-gully) $=$ 0.996, TA $=$ 0.990 [see Fig. 8(c)]. All operation statistics for WS2 are provided in Appendix II.

As the alpha-shape threshold decreases, the value a/p decreases smoothly (point density increases) until it approaches

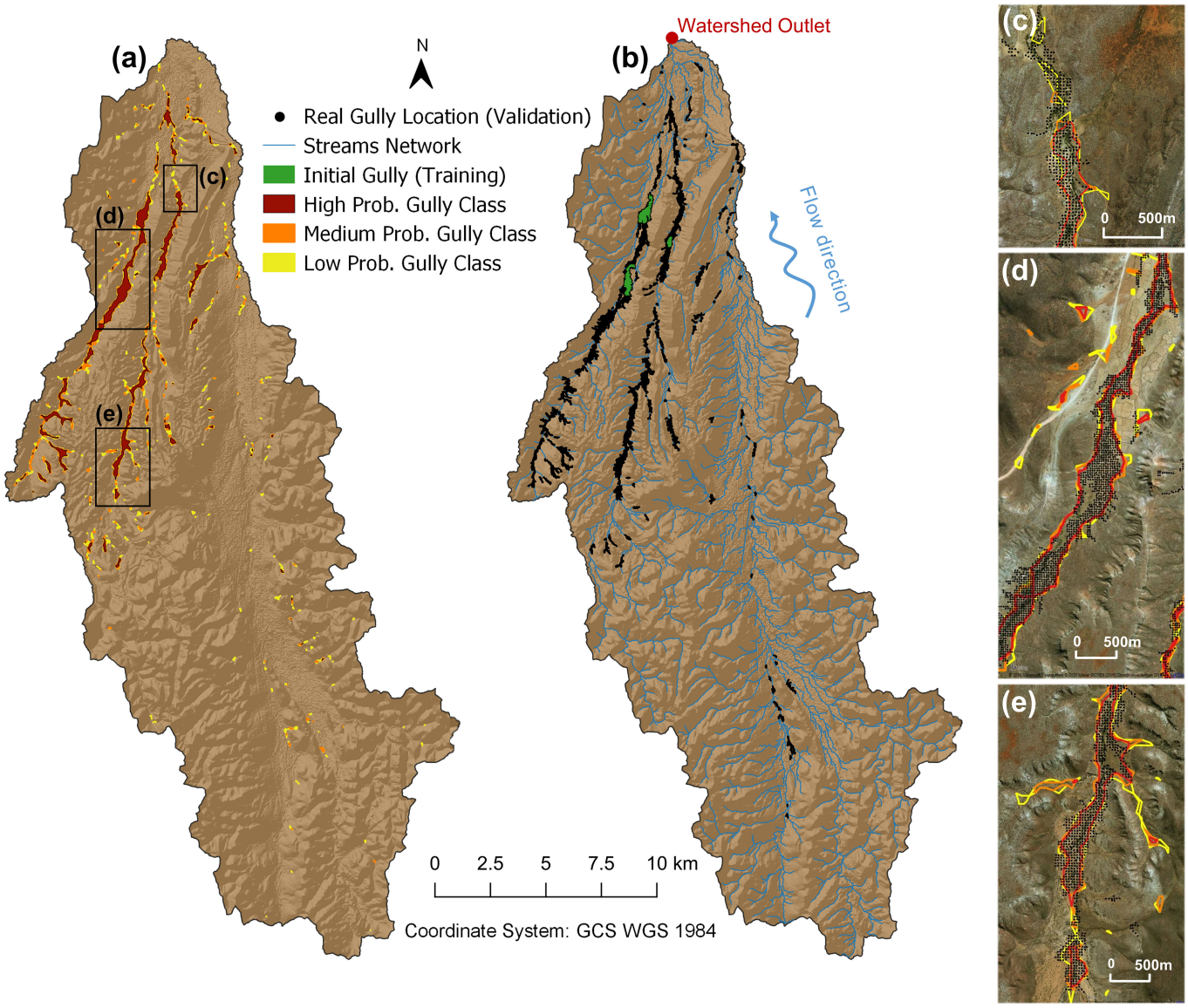

Fig. 7. Classification results for three gully probability classes (high, medium, and low) for WS1. (a) Shows the classification results and (b) the real gully locations used for validation (black dots), initial gully label (green polygons) used for training and streams network for the entire WS. (c)–(e) Detailed view of gully mapping underlain with Microsoft Bing^{TM} Maps Aerial Imagery.

1, at which step the point density is equivalent to the original 30 m grid [see Fig. 8(a)].

In this case, again radar VV coherence (22.8%), NDVI SD (22.0%) register the highest mean feature importance from iterations 0 to 9, with topographic roughness (16.1%) in this case in the third position in terms of significance to generate the classification results. ROC AUC indicates that the classification model achieves a maximum in terms of class separability between iterations 10 and 18, with AUC reaching 0.951 in iteration 14 (see Fig. 9). The evolution of the AUC values in WS2 confirms that our algorithm gradually improves its learning capacity in parallel with the increase in the size of gully training data until an AUC maximum is reached where more training data does not generate an improvement in the classification performance. This is an indicator of the maturity of the classification model.

Once the algorithm detects the best accuracy condition for the milestone of Q $>$ 40% in the iteration 10, three classes are defined as presented in Fig. 10. This allows increasing the accuracy of the results for the gully class, obtaining 0.734 for PA and 0.514 for UA in the high gully class and low gully class, respectively. The low, medium, and high gully classes cover 7.7, 4.7, and 3.6 km^2, accounting for 2.5%, 1.5%, and 1.2% of the WS2 area, respectively. As observed in Fig. 10, gullies are concentrated in the northwestern parts of the WS, dominated by lower areas close to the WS outlet.

A variety of scenarios can be observed within different gully zones. In Fig. 10(c), there are areas with high gully density surrounded by zones of transition to no-gully, which prove difficult for the algorithm to capture. Although those transition zones exhibit similar spectral and backscatter properties in the remote sensing data, they do not belong to the gully from a morphological and geometrical point of view. This is a common error of the algorithm. It is also observed that although small gully spots are usually difficult to be outlined correctly, the

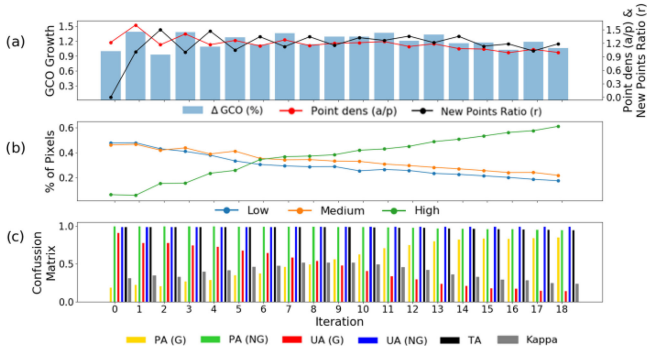

Fig. 8. Algorithm functionality and performance for WS2. (a) Black line represents the ratio of classified gully pixels by RF between iteration k and $k-1$. Red line is an indicator of gully pixels density inside GOC in units of their proportion to the original grid of 30 m. The bar chart represents the growth ratio for each iteration of the GOC. (b) Each line color represents the percentage of pixels created by the RF based on the probability ($m_{M,i}$) of being a gully: blue (C_{low}), orange (C_{med}), and green (C_{high}). c) Accuracy assessment including UA, PA, TA, and Cohen KA for each iteration.

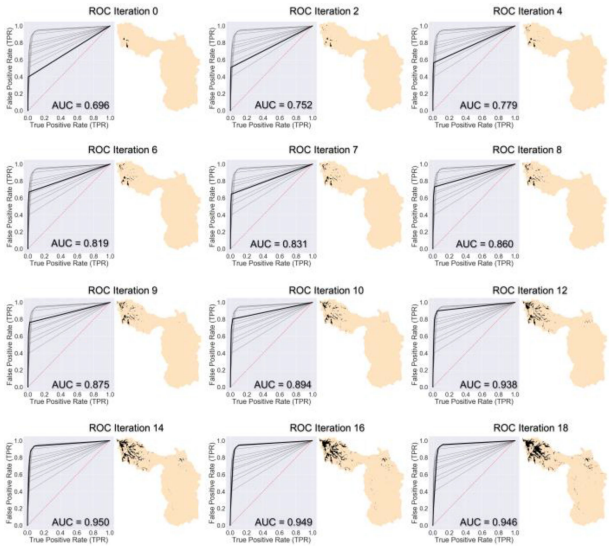

Fig. 9. ROC AUC analysis results for WS2 with ROC curves for the iterations 0, 2, 4, 6, 7, 8, 9, 10, 12, 14, 16, and 18 including the AUC values in the legend and corresponding cartographic outputs of gully class.

algorithm is still able to detect the presence of gully activity generating an overlapping gully object [see Fig. 10(d)]. Conversely, well-defined gullies are outlined accurately when they present linear patterns and clear spectral and morphological differences from their surroundings [see Fig. 10(e)].

C. Role of TanDEM-X IMR, MPCA Factors, and the Iterative Learning

Six factors have been used as conditioning variables within Sections IV-B.1 and IV-B.2 (NDVI SD, slope, topographic roughness, distance to stream, radar backscatter intensity VV, and radar coherence); however, a specific gully form factor was not included. To test the effect of the incorporation of gully morphological factors derived from TanDEM-X DEM resulting from [46], two new datasets were derived from the MPCA and IMR indexes. First, MPCA was included as a seventh explanatory variable in the general methodology and, second, MPCA was combined with IMR as a final geometric correction for the standard iterative solution. The results of both experiments were compared to the outcomes of the iterative learning with six explanatory variables as well as with an optimized (best algorithm settings were selected after several configuration trials) single RF and alpha shape application (not iterative). The experiment is conducted over a specific area where the density of gullies is high for each WS. The results are shown separately again for WS1 and WS2 in Fig. 11.

For both WSs, a different effect is observed by applying a morphological correction to the method. As for the WS1, the PA for gully class is increased and the UA decreased as compared to the Iterative RF solution, the opposite effect is observed in WS2. So, although it seems that this method allows reaching a greater level of detail in areas with branched gullies, the solution is not stable for large areas as the decrease in overall metrics is evident [see Fig. 11(b)–(f)]. On the other hand, the inclusion of MPCA as an explanatory variable has a positive influence in the complex gully network classification of the WS2 by improving the PA(G), Kappa, and MCC values [see Fig. 11(g)]. This result is confirmed with the 14.3% in the MPCA factor importance for WS2 (the third one among all the seven factors), and in contrast to the 8.8% resulted for the WS1, where MPCA is the least influential variable. Additionally, this comparative analysis confirms an evident improvement of the iterative approach toward the single RF approach in both WSs [see Fig. 11(d), (e), (h), and (i)].

V. Discussion

We applied an iterative learning approach framework to generate gully distribution maps in Kunene Region, Namibia, using Optical (S2), SAR (S1), and DEM (TX) imagery. The method combines an RF classifier and alpha shape into an iterative self-adjusting algorithm to first identify gully candidates at pixel level and to subsequently find geometric objects of these landforms. The method proves its ability to detect gullies in large extents with small and geographically aggregated training data, pointing to the potential benefit of integrating micro-mapping and/or crowd-sourcing approaches [56].

As this study represents a continuation of [46], it succeeds to address the research challenges that were articulated in it:

1) achieving an accuracy of at least 50% for gully class (keeping a TA>98%);
2) deriving geomorphometric features from the pixel-based classification, such as gully outline; and
3) integrating TanDEM-X with other products (i.e., S1 and S2) to achieve the mentioned objectives in 1) and 2).

Both studies could be combined in future research though, testing the feasibility of the gully locations generated by [46] as initial training data to supplement those datasets manually collected in the present approach.

In addition, our approach adds value to former studies that used ML and RF to derive gully susceptibility maps [39]. In this sense, one of the achievements of this article is to precisely derive gully outlines assessing in detail the errors types I (false positives) and II (false negatives) for the gully area.

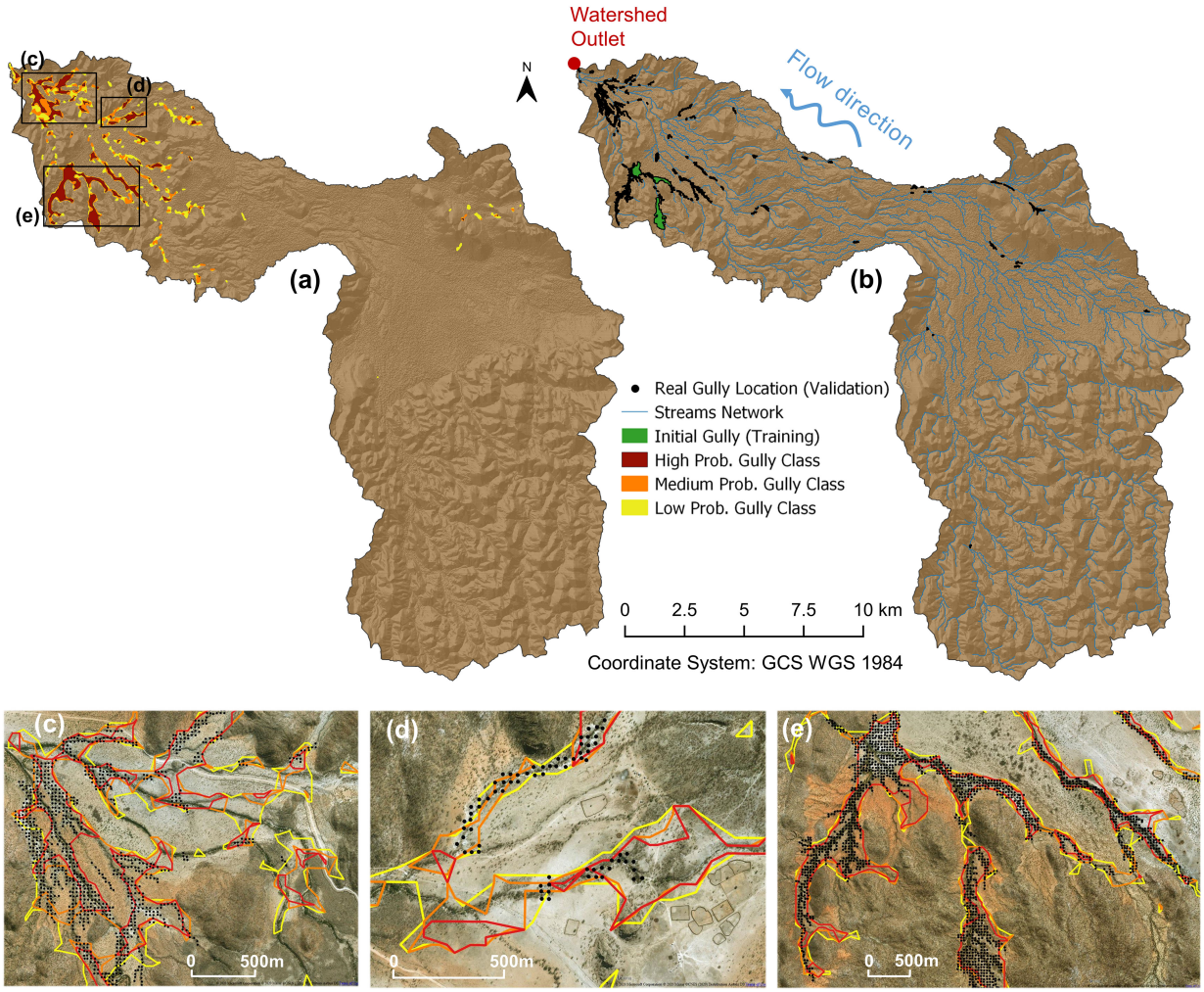

Fig. 10. Classification results for three gully probability classes (high, medium, and low) for WS2. (a) Classification results and (b) the real gully locations used for validation (black dots), initial gully label (green polygons) used for training. (c)–(e) Detailed view of gully mapping underlain with Microsoft Bing Maps Aerial imagery.

In this regard, our outcomes should be also compared with other methods that combine ML and OBIA techniques to derive gully outlines [41], [37]. An additional procedure in this study is the inclusion of time series (i.e., NDVI SD, radar sigma backscatter VV, radar coherence VV) as gully dynamics descriptors reducing seasonal features uncertainties as reported by [37]. A significant achievement is the capability of our method to generate comparable or better accuracy results than [41] and [37], but adjusting the methods to the reality of Namibia of very limited and aggregated initial training datasets. Shahabi *et al.* [41] used ∼70% of the final gully class as training data while [37] applied its methodology into a small study area (1.9 km^2) with approximately 25% of its extension used for training. In our case, gullies were outlined in large testing areas (>300 km^2) but with reduced (<0.3% of the testing area and ∼10% of the final gully class) initial training data.

Since comparisons are complicated for different study areas and input datasets, in this article, a simple RF and segmentation approach was also compared to the iterative solution, noticing that the second approach gives accuracies between 6% and 25% better than the non-iterative baseline.

Keeping in mind that the results of this article represent a solution by its own for large-scale gully classification, several points must be considered for further research. The role of the alpha shape threshold is fundamental to generating a reliable, growing training class after each iteration. To generate more detailed gully outlines, the sample grid width (and in consequence the alpha shape threshold) could be reduced in gully zones by aligning it to the input data maximum spatial resolution to reach higher levels of detail. In this vein, additional experiments to derive more reliable results could involve the modulation of the RF settings (i.e., the proportion of training/testing data) during the iterative process. Moreover, as part of the result to generate an intrinsic stopping criterion, the highest accuracies are achieved when the percentage of high purity gully points (C_{high}) falls between 35% and 45% in relative proportion to the low (C_{low}) and medium (C_{med}) classes. This fact is confirmed by the highest MCC values as a measure of balanced accuracy for different class sizes. However, under the assumption that some gully areas might need more iterations than others to be outlined completely, the possibility of working with individual gully sections and uniform sampling blocks should be considered to define a local

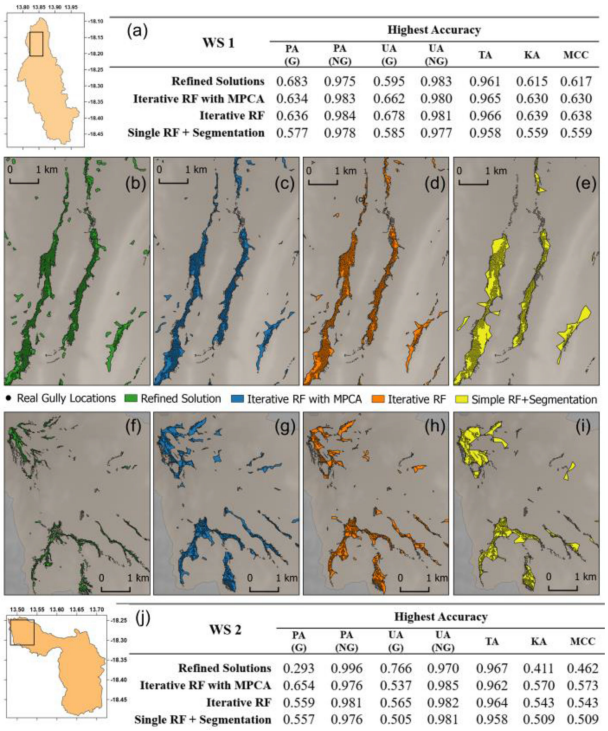

Fig. 11. Comparison of different solutions including single RF application and experimental methods using gully morphological factors. (a) Accuracy results for the WS1 comparing the (b) solution implementing morphological refinement method with IMR and MPCA, (c) the general workflow with MPCA as an explanatory variable, (d) the general workflow analysis without MPCA, and (e) optimal solution applying single RF and alpha shape. (j) Numerical accuracy results for the WS2 comparing the (f) solution implementing morphological refinement method with IMR and MPCA, (g) the general workflow with MPCA as an explanatory variable, (h) the general workflow analysis without MPCA, and (i) optimal solution applying single RF and alpha shape.

ending criterion for the iterations, instead of a global criterion for the whole sampling area. In this line, the addition of other criteria to define gully purity or maturity should be also considered, for example applying multivariate statistical test on gully internal variability [77]. On another note, continuous and unlimited data acquisition can be adopted for the definition of temporal patterns (i.e., NDVI, radar intensity, coherence) as the method can automatically be recomputed, as a new dataset is added (both explanatory and training data). Regarding the algorithm initialization, in addition to the conducted tests concerned to initial training data size, the distribution and spatial autocorrelation of these datasets should be also considered (i.e., aggregated versus disperse) [78] in further studies. Furthermore, to derive gully volume and 3D temporal change, more research is needed using SAR datasets, radar interferometry, coherence estimation techniques, both in the short term (rain event) and over long periods (year scale) [58]. This ideally would allow us to generate gully activity maps for large areas. Finally, extensive fieldwork is required to generate reliable 3D validation data, as well as to test the implications of different data sources as training samples.

VI. Conclusion

The ability to quantify and monitor gully erosion is essential for providing a deeper understanding of gully dynamics and planning remediation strategies in (semi-) arid regions. Cost-effective and reliable techniques to derive large-scale gully distribution maps are a critical management tool for preventing and neutralizing erosion and land degradation. In Namibia, the Kunene Region is severely affected by unmapped gully erosion sites, and our analysis contributes to fill this gap using open multisource remote sensing datasets combined with a 12-m global TanDEM-X DEM to delineate gully landforms. The approach demonstrates that with very limited training data and little expert intervention, high accuracies can be achieved automatically. The method is able to outline gullies of very large areas (302–408 km^2) using an initial gully class definition of <1 km^2 to achieve total accuracies $>98\%$, with $>50\%$ for the gully class (both UA and PA), Cohen KA, and MCC values >0.5. Total gully extents are 9.0 and 7.7 km^2 for two WSs, where tested our methodology, respectively. For these two WSs, the highest concentration of gullies is located in the flat area close to the WS outlets. This result points to a classical stream accumulation effect in combination with back wearing linear erosion (retrogressive erosion).

Among the six explanatory variables considered, their importance in decreasing order in the RF classification are radar coherence (22.4%), NDVI standard deviation (21.8%), distance to stream (15.5%), radar backscatter intensity (14.6%), topographic roughness (13.7%), and slope (12.1%). This outcome evidences the importance of factors that involve a temporal component to determine the presence of gullies and the adjusted balance in the importance of the six factors. Additions to the workflow show that a morphological correction using MPCA and IMR acts to increase very locally the level of detail in complex gully shapes, but is generally detrimental to the final accuracy. The inclusion of MPCA as an additional explanatory variable, however, generates moderate improvements in the results as compared to the use of the algorithm without this factor.

The main error source stems from the difficulty in outlining gullies when they present complex and branched forms. This is common in cases of high false-positive rate areas, evident in a higher UA for the gully class. Moreover, the iterative algorithm is very consistent with regard to the amount of initial training data. However, it presents a minimum training area required to be initiated. Therefore it is assumed that the addition of training data based on systematic crowdsourcing methods will benefit the overall accuracy and its applicability to larger areas [78].

This study presents a solution to delineating gullies based on earth surface measurements using global remote sensing products while improving and adapting existent techniques to the real scenario lacking reliable training data. Our methodology is transferable to other (semi-) arid regions, requiring minimal human input, highlighting a key point for its feasibility, and applicability in low populated regions. Finally, although our methodology was initially explored to generate 2D gully results, these outcomes can also be used as the footprint to generate 3D gully activity maps, for example making use of TerraSAR-X datasets [79] and final products such as DEMs of difference [80]. Thus, this research sets a foundation to develop a holistic approach in combination to a crowdsourcing data collection system and SAR techniques to generate gully dynamic maps.

APPENDIX I

A DETAILED REPORT OF ALGORITHM OPERATION FOR THE WS 1 INCLUDING THE FOUR PARAMETERS THAT CONTROL THE ALGORITHM (COLUMNS 2–6: ALPHA, NEW POINTS RATIO, GULLY CLEAN GROWTH, POINT DENSITY, AND QUALITY), THE IMPORTANCE OF THE FEATURES FOR EACH RF CLASSIFICATION (COLUMNS 7–12), AND THE CORRESPONDING ACCURACY VALUES (COLUMNS 13–20: PA, UA FOR GULLY-G AND NO GULLY-NG CLASSES, PLUS TA, COHEN KA, ROC AUC, AND MCC). WHEN THE QUALITY OF THE PARAMETERS ACHIEVES THE VALUES OF 35%, 40%, 45%, 50%, AND 55%, THE RESULT IS SPLIT INTO THREE SEVERITY CLASSES AND THEIR RESPECTIVE ACCURACY IS DISPLAYED

Iterations (k)	Alpha	New points Ratio	Gully Clean Growth	Point Density	Quality	Importance						Accuracy							
						NDVI SD	SIGM	COH	DIST	ROUG	SLOPE	PA (G)	PA (NG)	UA (G)	UA (NG)	TA	Kappa	ROC AUC	MCC
0	180	0.000	1.000	1.249	0.115	**0.207**	0.149	0.199	0.139	0.160	0.147	0.087	1.000	0.983	0.981	0.981	0.157	0.633	0.290
1	90	1.001	1.378	1.573	0.120	0.209	0.150	0.195	0.137	0.166	0.143	0.114	1.000	0.937	0.981	0.981	0.200	0.631	0.323
2	120	1.500	0.991	1.146	0.207	0.213	0.144	0.212	0.148	0.150	0.132	0.113	1.000	0.930	0.981	0.981	0.198	0.674	0.320
3	85	1.221	1.379	1.294	0.224	0.209	0.142	0.221	0.144	0.152	0.133	0.153	1.000	0.927	0.982	0.982	0.258	0.673	0.373
4	85	1.306	1.153	1.184	0.285	0.213	0.133	0.232	0.147	0.157	0.119	0.175	1.000	0.926	0.983	0.982	0.290	0.707	0.399
5	85	1.181	1.255	1.228	0.312	0.222	0.135	0.233	0.152	0.142	0.116	0.213	0.999	0.889	0.983	0.983	0.338	0.719	0.430
6	85	1.276	1.272	1.220	0.380	0.216	0.123	0.243	0.157	0.144	0.117	0.267	0.999	0.877	0.985	0.984	0.404	0.747	0.479
												Gully Class High Prob.	0.172	1.000	0.909	0.983	0.982	0.284		0.391
												Gully Class Medium Prob.	0.257	0.999	0.864	0.984	0.984	0.391		0.467
												Gully Class Low Prob.	0.345	0.998	0.816	0.986	0.985	0.479		0.525
7	85	1.207	1.241	1.188	0.396	0.223	0.116	0.244	0.173	0.139	0.105	0.316	0.999	0.843	0.986	0.984	0.453	0.779	0.510
8	85	1.248	1.219	1.183	0.427	0.225	0.122	0.232	0.180	0.133	0.109	0.370	0.998	0.805	0.987	0.985	0.500	0.803	0.540
												Gully Class High Prob.	0.259	0.999	0.861	0.984	0.984	0.392		0.467
												Gully Class Medium Prob.	0.365	0.998	0.805	0.987	0.985	0.496		0.536
												Gully Class Low Prob.	0.457	0.996	0.733	0.988	0.985	0.556		0.572
9	85	1.236	1.238	1.176	0.421	0.219	0.123	0.226	0.194	0.138	0.099	0.432	0.997	0.758	0.988	0.985	0.544	0.825	0.566
10	85	1.239	1.197	1.135	0.442	0.223	0.125	0.202	0.204	0.142	0.103	0.484	0.996	0.711	0.989	0.985	0.568	0.848	0.579
11	85	1.187	1.187	1.138	0.458	**0.221**	**0.114**	**0.198**	**0.210**	**0.145**	**0.113**	**0.540**	**0.994**	**0.669**	**0.990**	**0.985**	**0.590**	**0.864**	**0.594**
												Gully Class High Prob.	**0.438**	**0.997**	**0.748**	**0.988**	**0.985**	**0.545**		**0.566**
												Gully Class Medium Prob.	**0.525**	**0.994**	**0.670**	**0.990**	**0.985**	**0.581**		**0.586**
												Gully Class Low Prob.	**0.599**	**0.990**	**0.568**	**0.991**	**0.982**	**0.574**		**0.574**
12	85	1.169	1.180	1.132	0.485	0.213	0.126	0.182	0.225	0.148	0.106	0.579	0.992	0.607	0.991	0.983	0.584	0.877	0.584
13	85	1.169	1.172	1.121	0.491	0.206	0.124	0.165	0.238	0.159	0.108	0.620	0.989	0.554	0.992	0.982	0.576	0.891	0.576
14	85	1.169	1.179	1.128	0.496	0.211	0.115	0.155	0.255	0.158	0.107	0.656	0.986	0.497	0.993	0.979	0.555	0.898	0.560
15	85	1.172	1.182	1.138	0.506	0.200	0.118	0.146	0.269	0.160	0.106	0.683	0.981	0.435	0.993	0.975	0.519	0.907	0.548
												Gully Class High Prob.	0.606	0.989	0.551	0.992	0.981	0.568		0.568
												Gully Class Medium Prob.	0.672	0.983	0.455	0.993	0.976	0.531		0.542
												Gully Class Low Prob.	0.727	0.971	0.348	0.994	0.966	0.455		0.488
16	85	1.201	1.205	1.149	0.495	0.193	0.133	0.133	0.289	0.152	0.101	0.713	0.975	0.378	0.994	0.969	0.480	0.913	0.506
17	85	1.222	1.291	1.204	0.493	0.194	0.132	0.127	0.316	0.136	0.096	0.749	0.964	0.307	0.994	0.959	0.419	0.917	0.464
18	85	1.340	1.331	1.213	0.481	0.200	0.132	0.114	0.344	0.120	0.089	0.774	0.947	0.238	0.995	0.943	0.343	0.922	0.410
19	55	1.376	1.396	1.197	0.486	0.195	0.149	0.100	0.370	0.105	0.082	0.797	0.920	0.175	0.995	0.917	0.262	0.921	0.349
20	55	1.381	1.214	1.065	0.543	0.210	0.150	0.084	0.399	0.079	0.078	0.814	0.899	0.148	0.996	0.898	0.222	0.914	0.319
21	45	1.145	1.343	1.115	0.627	0.176	0.162	0.070	0.443	0.072	0.077	0.829	0.859	0.112	0.996	0.859	0.167	0.906	0.272
												Gully Class High Prob.	0.755	0.919	0.166	0.994	0.915	0.247		0.329
												Gully Class Medium Prob.	0.804	0.886	0.131	0.995	0.884	0.197		0.296
												Gully Class Low Prob	0.836	0.838	0.099	0.995	0.838	0.146		0.254

APPENDIX II

A DETAILED REPORT OF ALGORITHM OPERATION FOR THE WS 2 INCLUDING THE FOUR PARAMETERS THAT CONTROL THE ALGORITHM (COLUMNS 2–6: ALPHA, NEW POINTS RATIO, GULLY CLEAN GROWTH, POINT DENSITY, AND QUALITY), THE IMPORTANCE OF THE FEATURES FOR EACH RF CLASSIFICATION (COLUMNS 7–12), AND THE CORRESPONDING ACCURACY VALUES (COLUMNS 13–20: PA, UA FOR GULLY-G AND NO GULLY-NG CLASSES, PLUS TA, COHEN KA, ROC AUC, AND MCC). WHEN THE QUALITY OF THE PARAMETERS ACHIEVES THE VALUES OF 35%, 40%, 45%, 50%, AND 55%, THE RESULT IS SPLIT INTO THREE SEVERITY CLASSES AND THEIR RESPECTIVE ACCURACY IS DISPLAYED

Iterations (k)	Alpha	New points Ratio	Gully Clean Growth	Point Density	Quality	Importance						Accuracy							
						NDVI SD	SIGM	COH	DIST	ROUG	SLOPE	PA (G)	PA (NG)	UA (G)	UA (NG)	TA	Kappa	ROC AUC	MCC
0	180	0.000	1.000	1.218	0.061	0.188	0.177	0.209	0.138	0.140	0.148	0.188	1.000	0.915	0.992	0.992	0.310	0.697	0.413
1	105	1.005	1.394	1.603	0.056	0.190	0.181	0.210	0.137	0.138	0.145	0.225	0.999	0.781	0.993	0.992	0.347	0.696	0.417
2	135	1.500	0.930	1.165	0.151	0.196	0.169	0.240	0.133	0.131	0.131	0.209	0.999	0.781	0.992	0.992	0.327	0.753	0.401
3	80	1.000	1.382	1.404	0.154	0.196	0.162	0.236	0.139	0.129	0.139	0.273	0.999	0.748	0.993	0.992	0.397	0.742	0.449
4	80	1.470	1.090	1.166	0.233	0.205	0.157	0.245	0.139	0.130	0.124	0.292	0.999	0.731	0.993	0.992	0.415	0.780	0.459
5	80	1.049	1.277	1.262	0.256	0.219	0.153	0.241	0.141	0.126	0.120	0.354	0.998	0.681	0.994	0.992	0.462	0.792	0.488
6	80	1.343	1.133	1.139	0.343	0.233	0.150	0.240	0.144	0.124	0.110	0.380	0.998	0.647	0.994	0.992	0.475	0.820	0.492
7	80	1.125	1.365	1.273	0.365	0.239	0.153	0.239	0.141	0.118	0.110	0.464	0.997	0.588	0.995	0.992	0.515	0.831	0.518
												Gully Class High Prob.	0.300	0.999	0.682	0.993	0.992	0.413		0.449
												Gully Class Medium Prob.	0.411	0.998	0.629	0.994	0.992	0.494		0.505
												Gully Class Low Prob.	0.522	0.995	0.515	0.995	0.991	0.514		0.514
8	80	1.342	1.143	1.148	0.371	0.272	0.153	0.206	0.153	0.121	0.096	0.496	0.996	0.544	0.995	0.991	0.515	0.860	0.515
9	80	1.153	1.296	1.200	0.382	**0.266**	**0.154**	**0.214**	**0.154**	**0.114**	**0.098**	**0.568**	**0.994**	**0.483**	**0.996**	**0.990**	**0.517**	**0.875**	**0.519**
10	80	1.322	1.296	1.211	0.418	0.284	0.152	0.202	0.166	0.101	0.096	0.629	0.991	0.411	0.996	0.988	0.491	0.894	0.503
												Gully Class High Prob.	**0.513**	**0.995**	**0.515**	**0.995**	**0.991**	**0.509**		**0.509**
												Gully Class Medium Prob.	**0.624**	**0.991**	**0.407**	**0.996**	**0.988**	**0.487**		**0.498**
												Gully Class Low Prob.	**0.735**	**0.985**	**0.314**	**0.997**	**0.982**	**0.432**		**0.473**
11	75	1.262	1.370	1.236	0.429	0.305	0.144	0.190	0.171	0.098	0.092	0.713	0.987	0.342	0.997	0.984	0.455	0.912	0.487
12	75	1.356	1.208	1.126	0.449	0.342	0.147	0.164	0.178	0.088	0.081	0.756	0.983	0.299	0.998	0.981	0.420	0.938	0.468
13	60	1.212	1.332	1.186	0.488	0.358	0.143	0.156	0.186	0.082	0.075	0.803	0.976	0.240	0.998	0.974	0.361	0.941	0.431
												Gully Class High Prob.	0.720	0.986	0.333	0.997	0.984	0.449		0.483
												Gully Class Medium Prob.	0.782	0.979	0.267	0.998	0.977	0.389		0.449
												Gully Class Low Prob.	0.832	0.970	0.208	0.998	0.968	0.323		0.407
14	60	1.351	1.160	1.079	0.507	0.388	0.139	0.134	0.191	0.076	0.072	0.827	0.971	0.214	0.998	0.969	0.329	0.951	0.411
												Gully Class High Prob.	0.784	0.977	0.247	0.998	0.975	0.366		0.432
												Gully Class Medium Prob.	0.828	0.968	0.200	0.998	0.967	0.312		0.397
												Gully Class Low Prob.	0.881	0.939	0.122	0.999	0.938	0.201		0.315
15	45	1.134	1.178	1.069	0.533	0.397	0.131	0.129	0.208	0.069	0.066	0.839	0.964	0.183	0.998	0.963	0.290	0.950	0.382
16	45	1.183	1.025	0.989	0.561	0.436	0.124	0.109	0.207	0.063	0.060	0.837	0.963	0.178	0.998	0.962	0.282	0.949	0.375
												Gully Class High Prob.	0.766	0.972	0.206	0.998	0.970	0.314		0.387
												Gully Class Medium Prob.	0.834	0.960	0.167	0.998	0.959	0.267		0.363
												Gully Class Low Prob.	0.860	0.949	0.140	0.999	0.948	0.229		0.335
17	45	1.027	1.188	1.062	0.576	0.426	0.125	0.102	0.219	0.063	0.064	0.847	0.954	0.152	0.998	0.953	0.245	0.951	0.347
18	45	1.185	1.062	0.994	0.611	0.449	0.139	0.089	0.208	0.057	0.058	0.856	0.951	0.144	0.999	0.950	0.235	0.947	0.340

APPENDIX III

The ROC curve represents the classification performance at all classification thresholds and provides information on how capable the model is of distinguishing between classes. The ROC curve is plotted with true positive rate [TPR (8); a function of true positive, TP, and false negative, FN, cases] against the false positive rate [FPR (9); a function of false positive, FP, and true negative, TN, cases]. Reducing the classification threshold classifies more items as positive, thereby increasing both false positives and true positives [74]

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \tag{8}$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP}+\mathrm{TN}}. \tag{9}$$

The area under the ROC curve or AUC is applied as one of the most widely used classification quality indicators. This takes values between 0.0 (perfectly inaccurate result) to 1.0 (perfectly accurate result), with 0.5 as an indication of random guessing. ROC AUC values can be interpreted as the probability that the model classifies a pair of points correctly, one gully and one non-gully [74].

The MCC is used as a balanced indicator of quality for binary classifications for different class sizes. MCC ranges between -1 and $+1$, where $+1$ is a perfect prediction and -1 represents total disagreement between the estimation and the ground truth [75], [76]. The expression for MCC as a function of TP, TN, FP, and FN is given as

$$MCC = \frac{\mathrm{TP}\cdot\ \mathrm{TN}-\mathrm{FP}\cdot\mathrm{FN}}{\sqrt{(\mathrm{TP}+\mathrm{FN})\cdot(\mathrm{TP}+\mathrm{FP})\cdot(\mathrm{TN}+\mathrm{FP})\cdot(\mathrm{TN}+\mathrm{FN})}}. \tag{10}$$

ACKNOWLEDGMENT

This research is part of the DEM_HYDR2024 project supported by TanDEM-X Science Team, therefore the authors would like to express thanks to the Deutsches Zentrum für Luft- und Raumfahrt (DLR) as the donor for the used TanDEM-X datasets. They acknowledge the financial support provided by the Namibia University of Science and Technology (NUST) within the IRPC research funding programme and to ILMI for the sponsorship of field trips to identify suitable study areas. Finally, they would like to express gratitude toward Heidelberg University and the Kurt-Hiehle-Foundation for facilitating the suitable work conditions during this research.

REFERENCES

[1] T. Gomiero, "Soil degradation, land scarcity and food security: Reviewing a complex challenge," *Sustainability (Switzerland)*, vol. 8, no. 3. pp. 1–41, Mar. 2016.

[2] A. McBratney, D. J. Field, and A. Koch, "The dimensions of soil security," *Geoderma*, vol. 213, pp. 203–213, 2014.

[3] S. J. Bennett and R. R. Wells, "Gully erosion processes, disciplinary fragmentation, and technological innovation," *Earth Surf. Process. Landforms*, vol. 44, no. 1, pp. 46–53, Oct. 2018.

[4] C. Valentin, J. Poesen, and Y. Li, "Gully erosion: Impacts, factors and control," *Catena*, vol. 63, no. 2–3, pp. 132–153, Oct. 2005.

[5] J. Poesen, "Soil erosion in the Anthropocene: Research needs," *Earth Surf. Process. Landforms*, vol. 84, pp. 64–84, Sep. 2017.

[6] V. Chude *et al.*, "Regional assessment of soil changes in Africa south of the Sahara" in *Status of the World's Soil Resources: Main Report, FAO and ITPS*, 2015, ch. 9, pp. 243–274.

[7] G. R. Hancock, T. Wells, C. Martinez, and C. Dever, "Soil erosion and tolerable soil loss: Insights into erosion rates for a well-managed grassland catchment," *Geoderma*, vol. 237, pp. 256–265, 2015.

[8] M. Vanmaercke, S. De Geeter, and J. Poesen, "Modelling Gully erosion rates at regional and continental scales: Challenges and Opportunities," in *Proc. 8th Int. Symp. Gully Erosion Program*, Jul. 2019, p. 27.

[9] *National Policy On Climate Change For Namibia*. Windhoek, Namibia: Ministry of Environment and Tourism, Government of The Republic of Namibia, Oct. 2011.

[10] D. T. Bui *et al.*, "A novel ensemble artificial intelligence approach for gully erosion mapping in a semi-arid watershed (Iran)," *Sensors (Switzerland)*, vol. 19, no. 11, May 2019, Art. no. 2444.

[11] A. Arabameri, K. Rezaei, H. R. Pourghasemi, S. Lee, and M. Yamani, "GIS-based gully erosion susceptibility mapping: A comparison among three data-driven models and AHP knowledge-based technique," *Environ. Earth Sci.*, vol. 77, no. 17, pp. 1–22, Sep. 2018.

[12] H. Pringle, I. Zimmermann, and K. Tinley, "Accelerating landscape incision and the downward spiralling rain use efficiency of Namibian rangelands," *Agricola*, vol. 21, pp. 43–52, 2011.

[13] H. Pringle, I. Zimmermann, K. Shamathe, C. Nott, and K. Tinley, "Landscape incision processes favour bush encroachment over open grasslands in the two extremes of soil moisture balance in arid zones across southern Africa and Australia," *Agricola*, pp. 7–13, 2013.

[14] S. K. Imbamba, "A case study in land degradation and human vulnerability in the drylands of northern Namibia," Univ. of Nairobi, 1996.

[15] W. Pendleton, J. Crush, and N. Nickanor, "Migrant Windhoek: Rural-urban migration and food security in Namibia," *Urban Forum*, vol. 25, no. 2, pp. 191–205, Jan. 2014.

[16] Ministry of Environment and Tourism, *Namibia – Land Degradation Neutrality National Report*. Windhoek, Namibia: Government of the Republic of Namibia, Aug. 2015.

[17] M. Sartori *et al.*, "A linkage between the biophysical and the economic: Assessing the global market impacts of soil erosion," *Land Use Policy*, vol. 86, pp. 299–312, Jul. 2019.

[18] I. Zimmermann, "SASSCAL Task 41: Landscape literacy," *SASSCAL Res. Portfolio*, vol. 1, Aug. 2016.

[19] J. Mendelsohn, S. el Obeid, N. de Klerk, and P. Vigne, "Introducing Namibian farming systems," in *farming systems In Namibia.* Windhoek, Namibia: RAISON (Research & Information Services of Namibia), ABC Press, 2006.

[20] A. Goudie and H. Viles, "Landscapes and landforms of Namibia," in *World Geomorphological Landscapes*. Berlin, Germany: Springer, Jan. 2015.

[21] E. Brunotte and H. Sander, "Loess accumulation and soil formation in Kaokoland (Northern Namibia) as indicators of Quaternary climatic change," *Global Planet. Change*, vol. 26, pp. 67–75, Nov. 2000.

[22] E. Brunotte, B. Maurer, P. Fischer, J. Lomax, and H. Sander, "A sequence of fluvial and Aeolian deposits (desert loess) and palaeosoils covering the last 60 ka in the Opuwo basin (Kaokoland/ Kunene Region, Namibia) based on luminescence dating," *Quat. Int.*, vol. 196, nos. 1/2, pp. 71–85, Mar. 2009.

[23] F. N. Mwazi and G. N. Shamate, "Assessment of the status of soil macro-elements along a gully at Farm Krumhuk, Khomas Region, Namibia (2007)," *Agricola*, vol. 17, pp. 38–41, 2007.

[24] P. Klintenberg and M. Seely, "Land degradation monitoring in Namibia: A first approximation," *Environ. Monit. Assess.*, vol. 99, pp. 5–21, Jan. 2004.

[25] N. Rangeland and M. Policy, "Managing Namibia's rangelands," in *Proc. 13th Namibian Rangeland Forum*, Namibia, Oct. 2009, p. 8.

[26] N. H. Batjes, "SOTER-based soil parameter estimates for Southern Africa," ISRIC World Soil Information, Wageningen, The Netherlands, Rep. 2004/04, Oct. 2004.

[27] M. E. Coetzee, *Namsoter a Soter Database for Namibia*, Distributed by Ministry of Agriculture, Water & Rural Development Republic of Namibia, 2001.

[28] C. Castillo, R. Pérez, M. R. James, J. N. Quinton, E. V. Taguas, and J. A. Gómez, "Comparing the accuracy of several field methods for measuring gully erosion," *Soil Sci. Soc. Amer. J.*, vol. 76, no. 4, pp. 1319–1332, Apr. 2012.

[29] Y. Chen *et al.*, "Accuracy assessment of the planar morphology of valley bank gullies extracted with high resolution remote sensing imagery on the loess Plateau, China," *Int. J. Environ. Res. Public Health*, vol. 16, no. 3, pp. 1–19, Jan. 2019.

[30] B. D. Remmert, "Namibia's media : Facing the digital challenge." Accessed on: Jan 2020. [Online]. Available: https://ieeeauthorcenter.ieee.org/wp-content/uploads/IEEE-Reference-Guide.pdf

[31] FAO. *Aquastat Country Profile Namibia*. Rome, Italy: FAO, 2005.

[32] L. Zhen, "The national census for soil erosion and dynamic analysis in China," *Int. Soil Water Conserv. Res.*, vol. 1, no. 2, pp. 12–18, Sep. 2013.

[33] Y. Zhao and D. Cheng, "Collapsing gullies susceptibility mapping based on entropy information value in Jiangxi province of China," in *Proc. 2019 IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 3471–3474,

[34] A. Hayas, A. Peña, and T. Vanwalleghem, "Predicting gully width and widening rates from upstream contribution area and rainfall: A case study in SW Spain," *Geomorphology*, vol. 341, pp. 130–139, Sep. 2019.

[35] J. Knight, J. R. Spencer, A. P. Brooks, and S. R. Phinn, "Large-area, high-resolution remote sensing based mapping of alluvial gully erosion in Australia's tropical rivers," in *Proc. 5th Aust. Stream Manage. Conf.*, May 2007, pp. 205–210.

[36] A. Vrieling, S. C. Rodrigues, H. Bartholomeus, and G. Sterk, "Automatic identification of erosion gullies with ASTER imagery in the Brazilian Cerrados," *Int. J. Remote Sens.*, vol. 28, no. 12, pp. 2723–2738, Jun. 2007.

[37] K. Liu *et al.*, "An object-based approach for two-level gully feature mapping using high-resolution DEM and imagery: A case study on hilly loess plateau region, China," *Chinese Geogr. Sci.*, vol. 27, no. 3, pp. 415–430, Dec. 2017.

[38] O. Rahmati, A. Haghizadeh, H. R. Pourghasemi, and F. Noormohamadi, "Gully erosion susceptibility mapping: The role of GIS-based bivariate statistical models and their comparison," *Nat. Hazards*, vol. 82, no. 2, pp. 1231–1258, Feb. 2016.

[39] O. Rahmati, N. Tahmasebipour, A. Haghizadeh, H. R. Pourghasemi, and B. Feizizadeh, "Evaluation of different machine learning models for predicting and mapping the susceptibility of gully erosion," *Geomorphology*, vol. 298, pp. 118–137, Dec. 2017.

[40] A. Arabameri *et al.*, "Comparison of machine learning models for gully erosion susceptibility mapping," *Geosci. Front.*, vol. 11, no. 1, pp. 1609–1620, Dec. 2019.

[41] H. Shahabi, B. Jarihani, S. T. Piralilou, D. Chittleborough, M. Avand, and O. Ghorbanzadeh, "A semi-automated object-based gully networks detection using different machine learning models: A case study of bowen catchment, Queensland, Australia," *Sensors (Switzerland)*, vol. 19, no. 22, pp. 1–21, Nov. 2019.

[42] A. Hughes and P. I. Prosser, "Gully erosion prediction across a large region: Murray–Darling Basin, Australia," *Soil Res.*, vol. 50, no. 4, pp. 267–277, Jan. 2012.

[43] S. De Geeter, M. Vanmaercke, C. Isenborghs, and J. Poesen, "Modelling gully erosion rates across Africa: Towards a data-driven and process-oriented model," *EGU Geophys. Res. Abstracts*, vol. 21, 2019, Art. no. 7231.

[44] W. Dai *et al.*, "Effects of DEM resolution on the accuracy of gully maps in loess hilly areas," *Catena*, vol. 177, no. 1, pp. 114–125, Jun. 2019.

[45] Y. Garosi, M. Sheklabadi, H. R. Pourghasemi, A. A. Besalatpour, C. Conoscenti, and K. Van Oost, "Comparison of differences in resolution and sources of controlling factors for gully erosion susceptibility mapping," *Geoderma*, vol. 330, pp. 65–78, Nov. 2018.

[46] M. Vallejo, K. Negussie, E. Corral-Pazos-de-Provens, B. Höfle, and O. Bubenzer, "Comparison of three algorithms for the evaluation of TanDEM-X data for gully detection in Krumhuk farm (Namibia)," *Remote Sens.*, vol. 11, no. 11, Jun. 2019, Art. no. 1327,

[47] K. Liu *et al.*, "Large-scale mapping of gully-affected areas: An approach integrating Google Earth images and terrain skeleton information," *Geomorphology*, vol. 314, pp. 13–26, Aug. 2018.

[48] O. Rahmati, N. Tahmasebipour, A. Haghizadeh, H. R. Pourghasemi, and B. Feizizadeh, "Evaluating the influence of geo-environmental factors on gully erosion in a semi-arid region of Iran: An integrated framework," *Sci. Total Environ.*, vol. 579, pp. 913–927, Feb. 2017.

[49] M. Yibeltal *et al.*, "Morphological characteristics and topographic thresholds of gullies in different agro-ecological environments," *Geomorphology*, vol. 341, pp. 15–27, Sep. 2019.

[50] A. Arabameri, A. Cerda, and J. P. Tiefenbacher, "Spatial pattern analysis and prediction of gully erosion using novel hybrid model of entropy-weight of evidence," *Water (Switzerland)*, vol. 11, no. 6, May 2019, Art. no. 1129.

[51] K. Korzeniowska and O. Korup, "Mapping gullies using terrain surface roughness," in *Proc. 19th AGIL. Int. Conf. Geographic Inf. Sci.*, 2016, pp. 14–17.

[52] G. Lubo, Y. Lei, R. Yi, C. Zhewei, and B. Huaxing, "Spatial and temporal change of landscape pattern in the hilly-gully region of loess plateau," in *Proc. Int. Conf. Environ. Sci. Biotechnol.*, 2011, vol. 8, pp. 103–111.

[53] C. Persello and L. Bruzzone, "Active and semisupervised learning for the classification of remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 11, pp. 6937–6956, Nov. 2014.

[54] Z. Hui *et al.*, "An active learning method for DEM extraction from airborne LiDAR point clouds," *IEEE Access*, vol. 7, pp. 89366–89378, 2019.

[55] Z. Lv, G. Li, Z. Jin, A. Benediktsson, and G. M. Foody, "Iterative training sample expansion to increase and balance the accuracy of land classification from VHR imagery," *IEEE Trans. Geosci. Remote Sens.*, pp. 1–12, 2020

[56] B. Herfort, B. Höfle, and C. Klonner, "3D micro-mapping: Towards assessing the quality of crowdsourcing to support 3D point cloud analysis," *ISPRS J. Photogramm. Remote Sens.*, vol. 137, pp. 73–83, 2018.

[57] C. Nwaogu, O. J. Okeke, S. Assuah Adu, E. Babine, and V. Pechanec, "Land Use—land cover change and soil-gully erosion relationships: A study of Nanka, south-eastern Nigeria using geoinformatics," in *Proc. Dyn. GIsci.*, Mar. 2018, pp. 305–319,

[58] T. Ullmann, J. Sauerbrey, D. Hoffmeister, M. S. May, R. Baumhauer, and O. Bubenzer, "Assessing spatiotemporal variations of Sentinel-1 InSAR coherence on different time scales over the Atacama desert (Chile) between 2015 and 2018," *Remote Sens.*, vol. 11, no. 24, Dec. 2019, Art. no. 2960.

[59] M. Shoshany, "Detection and analysis of soil erodibility patterns using air photographs of the Avisur Highlands, Israel," *Hydrol. Geomorphol. Interface Rainfall, Floods, Sedimentation, Land Use*, vol. 261, pp. 127–139, Jan. 2000.

[60] R. B. V. Shruthi, N. Kerle, V. Jetten, L. Abdellah, and I. Machmach, "Quantifying temporal changes in gully erosion areas with object oriented analysis," *Catena*, vol. 128, pp. 262–277, May 2015.

[61] J. S. Malan, *Peoples of Namibia*. Pretoria, South Africa: Rhino Publishers, 1995.

[62] Namibia Statistical Agency. *Booklet September 2018, Namibia Land and Statistics*, Windhoek, Namibia: Namibia Statist. Agency, Oct. 2018.

[63] Namibia Statistical Agency. "National spatial data infrastructure (NSDI)," Distributed by NSA. Accessed: Sep. 2019. [Online]. Available: https://digitalnamibia.nsa.org.na/

[64] L. Harper-Simmonds *et al.*, *Development of an Inventory of Ecosystem Services in NAMIBIA*. Windhoek, Namibia: Ministry of Environment and Tourism of Namibia, 2011.

[65] G. Menz and J. Richters, "Quantitative classification of landscapes in northern Namibia using an ASTER digital elevation model," *African Landscapes*, vol 4, pp. 105–117, Jan. 2009.

[66] A. Linstädter, "Landscape ecology of Savannas," in *African Landscapes. Studies in Human Ecology and Adaptation*, O. Bubenzer, M. Bollig, Eds., New York, NY: Springer, Oct. 2008, vol. 4, pp. 79–103.

[67] T. Becker and N. Jürgens, "Vegetation along climate gradients in Kaokoland, North-West Namibia," *Phytocoenologia*, vol. 30, no. 3–4, pp. 543–565, 2000.

[68] J. Mendelsohn, A. Jarvis, C. Roberts, and T. Robertson, *The Atlas of Namibia: A Potrait of the Land and its People*. Windhoek, Namibia: Ministry of Environment and Tourism and David Philip, 2002, ch. 3, pp. 68–92.

[69] A. Jones *et al.*, *Soil atlas of Africa*. Luxembourg: Publications Office of the European Union, 2013, pp. 117–124.

[70] H. Edelsbrunner, D. Kirkpatrick, and R. Seidel, "On the shape of a set of points in the plane," *IEEE Trans. Inf. Theory*, vol. 29, no. 4, pp. 551–559, Jul. 1983.

[71] G. Hooker and L. Mentch, "Comments on: A random forest guided tour," *Test*, vol. 25, no. 2, 2016, pp. 254–260,

[72] V. Zhou, "A simple explanation of gini impurity," Accessed: Jan. 2020. [Online]. Available: https://victorzhou.com/blog/gini-impurity/

[73] S. Ronaghan, "The mathematics of decision trees, random forest and feature importance in scikit-learn and spark." Accessed: Jan. 2020. [Online]. Available: https://towardsdatascience.com/the-mathematics-of-decision-trees-random-forest-and-feature-importance-in-scikit-learn-and-spark-f2861df67e3

[74] S. Narkhede, "Understanding AUC – ROC curve." Accessed on: Jan. 2020. [Online]. Available: https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5

[75] B. W. Matthews, "Comparison of the predicted and observed secondary structure of T4 Phage Lysozyme," *Biochim. Biophys. Acta*, vol. 405, pp. 442–451, 1975.

[76] S. Boughorbel, F. Jarray, and M. El-anbari, "Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric," *PLoS ONE*, vol. 12, no. 6, pp. 1–17, 2017.

[77] G. Willems, G. Pison, P. J. Rousseeuw, and S. Van Aelst, "A robust Hotelling test," *Metrika*, vol. 55, no. 1, pp. 125–138, Apr. 2002.
[78] K. Millard and M. Richardson, "On the importance of training data sample selection in Random Forest image classification: A case study in Peatland ecosystem mapping," *Remote Sens.*, vol. 7, no. 7, pp. 8489–8515, Jul. 2015.
[79] German Aerospace Center. "TanDEM-X Ground Segment: Announcement of Opportunity," DLR, Oberpfaffenhofen, Germany, Rep. TD-PD-PL-0032, May 2014.
[80] German Aerospace Center. "EOC geoservice data guide," DLR, Oberpfaffenhofen, Germany. Accessed: Feb. 2020. [Online]. Available: https://geoservice.dlr.de/web/dataguide/tdm90/

Miguel Vallejo Orti received the B.Eng. degree in forest engineering from the University of Huelva, Huelva, Spain, in 2010, and the M.Sc. degree in photogrammetry and geoinformatics from the Stuttgart University of Applied Sciences (HFT), Stuttgart, Germany, in 2015. He is currently working toward the Ph.D. degree at the Institute of Geography, Heidelberg University, Heidelberg, Germany.

Since 2016, he is a Lecturer with the Department of Geo-Spatial Sciences and Technology, Namibia University of Science and Technology, Windhoek, Namibia. His main research interests include the study and monitoring of land and environmental conditions in arid regions and the development of analytical methods based on geoinformation technologies.

Lukas Winiwarter received the B.Sc. degree with distinction and the Dipl.-Ing. (equivalent to a M.Sc. degree) with distinction from Geodesy and Geoinformatics at TU Wien, Vienna, Austria, in 2016 and 2018, respectively.

During his studies, he was working with the research groups Photogrammety and Higher Geodesy at TU Wien in several projects as a student/research assistant. Since 2018, he has been working as a full-time Research Assistant and Ph.D. Student at the Institute of Geography, Heidelberg University, Germany. His main research interests include the automation of 3D point cloud analyses, with a focus on multi-sensor data, error modeling, and time series, as well as the development of machine learning methods.

Eva Corral-Pazos-de-Provens received the B.Eng. degree in forest engineering from University of Huelva (UHU), Huelva, Spain and a Master in environmental law from the International University of Andalusia, Seville, Spain.

She is a Researcher with UHU. She has been conducting soil science research with UHU, focused on soil mapping for ten years. She was the Head of the soil science team with Ecuador in the National Thematic Mapping Project, for 13 months (Ministry of Agriculture of Ecuador). She also worked for seven years in a Project for Modeling the Growth of Forest Plantations based on Environmental Variables (UHU). She is currently working on the European Project POCTEP CILIFO, on planning and management in forest fire prevention.

Jack G. Williams received the Ph.D. degree in physical geography from Durham University, Durham, UK.

He undertook two Post-Doctoral Research Associate positions in Durham's Department of Geography and Institute for Hazard, Risk and Resilience and a subsequent Research Associate position in the 3DGeo research group with the Interdisciplinary Center for Scientific Computing, Heidelberg University, Germany. His research interests include twofold: (1) the use of near-continuous terrestrial LiDAR/point cloud monitoring for investigating the nature and drivers of a range of geomorphic processes, principally landslides and rockfalls, and (2) the evolution of coseismic landslide hazard during and after large earthquake events.

Olaf Bubenzer received the Ph.D. degree from the University of Cologne, Germany, in 1998.

He studied geography, geology, meteorology, and soil sciences in Cologne and Bonn, Germany. Since 2017, he is leading the research group Geomorphology and Soil Geography with the Institute of Geography, Heidelberg University, Heidelberg, Germany, a position he held once before, between 2007 and 2012. In the intervening years, he conducted research with the University of Cologne, where he headed the working group Quaternary Research and Applied Geomorphology and the Section on Africa Research. His main scientific research interests include the investigation of dryland environments, including human-nature interactions.

Dr. Bubenzer is a Fellow of the Marsilius Kolleg, Heidelberg University, and founding director of the Heidelberg Center for the Environment.

Bernhard Höfle received the Ph.D. degree with excellence in the field of physical geography and geoinformatics from the University of Innsbruck, Innsbruck, Austria, in 2007.

He was a Junior Professor of GIScience and 3D spatial data processing with the Institute of Geography, Heidelberg University, Heidelberg, Germany, from 2011 to 2017. In 2017, he became a Full Professor of 3D spatial data processing at the Institute of Geography and head of the 3DGeo Research Group with the Interdisciplinary Center for Scientific Computing, Heidelberg University. Since 2017, he has been the Dean of Studies with the Faculty of Chemistry and Geosciences, Heidelberg University. His research interests include the development of computational methods for the automatic extraction of geoinformation from remote sensing data, and 3D/4D point clouds acquired by different sensor systems and platforms.