


2014

# Classical gully spatial identification and slope stability modeling using high-resolution elevation and data mining technique

Laurimar Gonçalves Vendrusculo  
*Iowa State University*

Follow this and additional works at: <http://lib.dr.iastate.edu/etd>

 Part of the [Agriculture Commons](#), [Bioresource and Agricultural Engineering Commons](#), [Computer Sciences Commons](#), [Environmental Sciences Commons](#), and the [Soil Science Commons](#)

---

## Recommended Citation

Vendrusculo, Laurimar Gonçalves, "Classical gully spatial identification and slope stability modeling using high-resolution elevation and data mining technique" (2014). *Graduate Theses and Dissertations*. 14082.  
<http://lib.dr.iastate.edu/etd/14082>

This Dissertation is brought to you for free and open access by the Graduate College at Iowa State University Digital Repository. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact [digirep@iastate.edu](mailto:digirep@iastate.edu).

**Classical gully spatial identification and slope stability modeling  
using high-resolution elevation and data mining technique**

by

**Laurimar Gonçalves Vendrusculo**

A dissertation submitted to the graduate faculty  
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Co-majors: Agricultural and Biosystems Engineering, and  
Environmental Science

Program of Study Committee:  
Amy Kaleita, Major Professor  
Lie Tang  
Matthew Helmers  
Monica Haddad  
Richard Cruse

Iowa State University

Ames, Iowa

2014

Copyright © Laurimar Gonçalves Vendrusculo, 2014. All rights reserved.

## **DEDICATION**

I dedicate this dissertation to my mother, Maria Auxiliadora Teixeira Pereira, who taught me staples—values of life such as hard work, persistence, ethics and faith. For her no professional, academic and personal project is too late or difficult to be achieved.

I also dedicate this work to my husband, Dr. Edson Adriano Vendrusculo, who has supported me throughout this new challenge, and to my brothers and sisters: Marymar Gonçalves Butruille, Anderson Gonçalves, Everton Gonçalves, Geicimar Gonçalves, Gleicimar Gonçalves, Peterson Gonçalves, and Debora Gonçalves for whom I have deep gratitude.

## TABLE OF CONTENTS

	Page
LIST OF FIGURES .....	v
LIST OF TABLES .....	viii
ACKNOWLEDGMENTS .....	ix
ABSTRACT .....	x
CHAPTER 1. INTRODUCTION .....	1
1.1 Soil Erosion .....	1
1.2 Gully Erosion Process .....	2
1.3 Precision Conservation .....	4
1.4 Soil Erosion Models .....	6
1.5 Data Mining Applied to Soil Erosion .....	6
1.6 Objectives .....	8
References .....	8
CHAPTER 2. IDENTIFICATION OF CLASSIC GULLY USING AIRBORNE LiDAR DATA, TERRAIN ANALYSIS, AND SPATIAL CLUSTERING TECHNIQUES UNDER DENSE VEGETATION AREA .....	17
Abstract .....	17
2.1 Introduction .....	18
2.2 Materials and Methods .....	24
2.2.1 Area description .....	24
2.3 Data Processing .....	28
2.3.1 Data collection and preparation .....	28
2.3.2 Classification methods .....	32
2.3.3 Threshold model to identify concentrated flow path within a gully channel bed .....	33
2.3.4 Validation methods .....	34
2.4 Results and Discussion .....	36
2.4.1 Comparing field survey and gulley concentrated flow paths .....	36
2.4.2 Statistical analysis of terrain attributes .....	40
2.4.3 Classical accuracy .....	42
2.4.4 Gully classification .....	47
2.4.5 Precision conservation application .....	58
2.5 Conclusions .....	51
Acknowledgments .....	54
References .....	54

CHAPTER 3. EVALUATING GULLY SLOPE STABILITY IN MULTI- TEMPORAL SURVEY BOUNDARIES FOR PRECISION CONSERVATION PURPOSE.....	59
Abstract.....	59
3.1 Introduction.....	60
3.2 Material and Methods.....	66
3.2.1 Study site.....	67
3.2.2 Gully perimeters coordinate adjustments in past surveys.....	68
3.2.3 Factor of safety computation and implications to precision conservation.....	74
3.3 Results and Discussion.....	75
3.3.1 Sensitivity analysis.....	80
3.3.2 Critical instable slopes for precision agriculture purpose.....	85
3.4 Conclusions.....	86
Acknowledgments.....	88
References.....	89
CHAPTER 4. CONCLUSIONS.....	92
APPENDIX A. RANKING METHOD.....	94
APPENDIX B. GULLY LOCATION DATASET.....	99
APPENDIX C. PLAIN FAILURE FOR GULLY SIDEWALLS.....	104

## LIST OF FIGURES

Figure 1.1	(a) Bowl-shaped headcut of the gully viewed looking upstream from the foot of the western gully bank; (b) Wooded lower reaches of the mapped section of the gully looking downstream. (Source: Thomas et al., 2004) .....	3
Figure 1.2	Schematic representation of main gullies mass-wasting processes gullies in Western Iowa.....	4
Figure 2.1	Site map showing the location of KBSSR watershed among Iowa counties, USA and detail of study area with gully networks identified by SSURGO lines (dotted lines) and field-surveyed patches (bold-lines). Note that the largest gully in the top left of inset figure it was surveyed in its top edge and floor flow patch because it was in a trapezoidal shape.....	25
Figure 2.2	(a) Example of V-shaped gully and (b) concentrated flow path between two gully branches. Images taken at leaf-off condition in April, 2014. As shown in the images, the area is fully covered by leaves and hardwoods and the gully floors are still muddy and exhibit ephemeral flow.....	27
Figure 2.3	(a) Elevation hillshade map of KBSSR portion area at Story County, Iowa, derived from the LiDAR One Meter Digital Elevation Model highlighting gullies classified at SSURGO (dotted lines); (b) 2013 National Agriculture Imagery Program (NAIP) Four-Band Aerial Photography with field-surveyed pathways (bold lines).....	28
Figure 2.4	General framework including three stages of the gully morphology with include the identification of side walls cluster from CLARA identification and gully concentrated flow (GCF) polylines finalizing with the spatial intersection of these gully elements.....	35
Figure 2.5	Map with original flow accumulation (a) and gully concentrated flow polylines computed (black lines) by the threshold procedure compared with the fieldwork paths (red lines).....	37
Figure 2.6	Distance variation in the largest gully (#1) at KSSBR study area between bed-path survey and flow-accumulation streamlines exceeding 5000 m <sup>2</sup> (e.g. high flow). The first point corresponds to the headcut and further points head toward the South Skunk River.....	39
Figure 2.7	Linear regression models among six measured classic gully length and predicted gully concentrated flow (GCF) lines.....	39
Figure 2.8	Descriptive statistics of four variables used in the clustering considering the entire study area.....	41

Figure 2.9	Maps of (a) Stream power index, (b) profile curvature (c) aspect (d) mean slope deviation at the study area from 1-m LiDAR data. The black lines in the maps represent the survey paths from field work.....	43
Figure 2.10	Sum of squares error curve between representative centroid point and observation labeled in each cluster. According elbow technique the lowest adequate number of clusters is equal to 6.....	44
Figure 2.11	Box plots of attributes used in CLARA unsupervised clustering distributed by cluster. The boxplot shows the data extremes (whiskers and the points for extreme cases), lower and upper quartiles, and the median. ....	46
Figure 2.12	Bivariate plot of principal components according the eight groups classified by CLARA method. ....	47
Figure 2.13	(a) Final study area classified with six groups; (b) refined areas formed by points in cluster 6 and classified as gully side walls (blue areas). The bold lines represent the GCF polylines obtained by the general threshold. ....	49
Figure 2.14	Clustered area surrounding Gully #3 with a cross section of its open channel. The gully edge has points classified as 3 and can be chosen to soil control structures placement. ....	51
Figure 3.1	(a) Geometry of the slope with tension crack placed in upper slope surface and (b) diagram forces showing the relationship of resisting and driving forces to a generic trapezoidal failure surface. ....	63
Figure 3.2	(a) Hillshade of Treynor site Pottawattamie County, Iowa derived from the LiDAR One Meter Digital Elevation Model. (b) 2011 National Agriculture Imagery Program (NAIP) Four-Band Aerial Photography highlighting boundaries of watershed #1 boundary (light line).....	67
Figure 3.3	Examples of (a) $\theta_1$ (b) $\theta_2$ (c) $\theta_2$ rotation in three-dimensional conformal coordinate transformation extracted from Ghilani and Wolf (2006). ....	69
Figure 3.4	Individual lines regarding to November, 1964 and October, 1965 survey consolidated by Thomas et al. (2004) transformed in this study in equal-spaced points sequence.....	72
Figure 3.5	Set of gully perimeters showing the partial surveys executed, for instance, survey lines produced at April 1979 and 1980 campaigns.....	73
Figure 3.6	Boxplots showing the dataset distribution of elevation and slopes in gully edges both from survey periods on 1999 and 2014.....	77

Figure 3.7	Slope map and gully perimeter at the gullied area in Treynor site on (a) February, 1999 originated from Thomas et al. (2009) study, and (b) May, 2014 from LiDAR data.....	78
Figure 3.8	Spatial distribution of factor of safety regarding to partially saturated failure ( $\gamma_w = 15.5 \text{ KN/m}^3$ ) on 1999 (a,b), and 2014 survey (c,d) under dry conditions. Negative values were computed and are placed mainly in the gully bed. ....	79
Figure 3.9	Histograms of slab height (m) and slope (degrees) at Treynor site on May, 2014 at dry and partially saturated conditions coded in a) negative FS (n=961) b) instable slopes (n=1761) c) stable slopes (n=370). ....	81
Figure 3.10	Impacts of variables uncertainties on slope instability: (a) internal friction angle, ( $\theta$ ) b) soil cohesion, c) failure plane angle, ( $\psi_p$ ) and, d) slab thickness, (x) at 1999 survey. ....	83
Figure 3.11	Empirical cumulative distribution (CDF) of factor of safety on 1999 and 2014 surveys at Treynor gullied area. ....	84
Figure 3.12	Factor of safety maps with selected pixels considered instable slopes that are (a) one and (b) two standard deviation of elevation from maximum slope ( $S_{max} = 57.6\text{o}$ ) at 2014 survey. ....	86



**LIST OF TABLES**

Table 2.1	Statistics of surveyed polylines and GCF in the study area.....	36
Table 2.2	General statistics for study area considering terrain attributes of elevation above sea level ( $Z$ ); stream power index (SPI); mean slope deviation, aspect and profile curvature ( $Curv_{pr}$ ).....	42
Table 2.3	Summary statistics of topographic and hydrologic information by classes in the study area .....	45
Table 3.1	Survey timeline with gully length obtained from Thomas et al. (2004) set of perimeters and transformed with two-dimensional conformal transformation.....	76

## ACKNOWLEDGMENTS

I would like to thank my committee chair, Dr. Amy Kaleita, for her guidance, support and encouragement during this research. I am also appreciative of from my committee members: Drs. Lie Tang, Richard Cruse, Mattew Helmers and, especially, Monica Haddad for their guidance and support throughout the course of this research.

I would like to acknowledge the opportunity and financial support of Embrapa Brazilian Corporation of Agricultural Research.

I am also grateful to my friends Giorgi Chighladze, Derek Groenendyk, Ligia Serrano, Ambika Karkee, James Newman, and my colleagues from the Spatial Analysis Lab, and the Agricultural and Biosystem Engineering Department faculty and staff for making my time at Iowa State University a lifetime experience. I am also appreciative of the help from those who assisted me in my field work and provided data for this study.

Special thanks to my family, especially my husband, Edson Adriano Vendrusculo, my son, Henrique, and daughter, Ana, for their encouragement, patience, respect and love.

Last, and foremost, I am thankful to God for keeping me healthy and enriching my faith in his plan for my life.

## ABSTRACT

It is widely known that soil erosion is an issue of concern in soil and water quality, affecting agriculture and natural resources. Thus, scientific efforts must take into consideration the high-resolution elevation dataset in order to implement a precision conservation approach effectively. New advances such as LiDAR products have provided a basic source of information to enable researchers to identify small erosional landscape features. To fill this gap, this study developed a methodology based on data mining of hydrologic and topographic attributes associated with concentrated flow path identification to distinguish classic gully side walls and bed areas. At 0.91 Km<sup>2</sup> region of the Keigley Branch-South Skunk River watershed, an area with gullies, we computed profile curvature, mean slope deviation, stream power index, and aspect gridded in 1-m pixel from Iowa LiDAR project. CLARA (CLustering LARge Applications) algorithm. An unsupervised clustering approach was employed on 913,495 points splitting the dataset in six groups, the number in agreement with within-group sum of squared error (WSS) statistical technique. In addition, a new threshold criteria termed gully concentrated flow (GCF) based upon data distribution of flow accumulation and mean overall slope were introduced to produce polylines that identified the main hydrographic flow paths, corresponding to the gully beds. Cluster #6 was classified as gully side walls. After distinguishing gullies and cliffs areas among points belonging to cluster 6, all six gullies were satisfactorily identified. The proposed methodology improves on existent techniques because identifies distinct parts of gullies which include side walls and bed zone.

Another important concept is assessing gully slope stability in order to generate useful information for precision conservation planning. Although limit-equilibrium concept

has been used widely in rock mechanics its application in precision conservation structures is relatively new. This study evaluated two multi-temporal surveys in a Western Iowa gullied area under the approach of soil stability regarding precision conservation practice. The study computed factor of safety (FS) at the gully area, including headcut and gully side walls using digital elevation models originated from surveys conducted in 1999 and 2014.

Outcomes of this assessment have revealed significantly less instability of the actual slopes compared to 1999 survey slopes. The internal friction angle ( $\theta$ ) had the largest effect on slope stability factor ( $S.D._{1999} = 0.18$ ,  $S.D._{2014} = 0.24$ ), according to the sensitivity analysis, compared to variations of soil cohesion, failure plane angle and slab thickness. In addition, critically unstable slopes within gully, based on units of the slope standard deviation, as a threshold, have produced an area of  $61 \text{ m}^2$  and  $396 \text{ m}^2$  considering the threshold of one and two slope standard deviation, respectively. The majority of these critical areas were located near the headcut and in the border of side walls. Based on current literature, association of processed material (geotextile) and crop cover with high root density might be an alternative to improve slope instability, but empirical tests are necessary to validate this approach. Nevertheless, the slope instability must include other factors that capture the dynamics of failure.

## CHAPTER 1. INTRODUCTION

### 1.1 Soil Erosion

It is widely known that soil erosion is an issue in soil and water quality that affects agriculture and natural resources. Due to erosion, valuable topsoil and its intrinsic organic matter are removed and the newly exposed soil may present imbalanced nutrients which compromise yield productivity. The eroded soil sediments also add toxic substances to natural water bodies and increase turbidity and toxicity (Lal & Lowery, 1999). Even though from 1982 to 1997 there was a 40% soil erosion reduction in the United States as announced by the EPA (2009), there is still room for improvement in mitigating environmental degradation. Furthermore, high soil losses by erosion are still reported in Asia, Africa and South America (30-40 ton ha<sup>-1</sup> year<sup>-1</sup>) as estimated by Pimentel et al. (1995), largely due to inappropriate land use.

Soil erosion is related to factors such as soil properties, topography, drainage, precipitation and land use (Toy, Foster, & Renard, 2002). However, the interaction among those properties has been found complex and remains an object of research. Likewise, the knowledge and judgments from experts and local people are crucial for employing effective methodologies to understand and mitigate this process over time. For the most part, classical models of soil loss analyze sheet and rill erosion. However, gully erosion contributes a disproportional yield of sediment at watershed scales (Pimentel et al. 1995) For this feature, additional factors contribute to the geometry of gully instability, such as groundwater sapping and tension crack development (Istanbulluoglu et al., 2005).

The empirical measurements for modeling soil erosion at large scales, such as watershed, are expensive and time consuming. Furthermore, land use and soil coverage are

frequently field-scale estimated and changing seasonally. Thus, other less costly technologies are desirable, especially in places where fine-resolution topographic variables, soil and climate attributes not are available. One example of this advance has been reported by Watts et al. (2011) who increased the accuracy of recognition of conservation tillage incorporation high temporal MODIS series in his analysis. Furthermore, Nachtergaele and Poesen (1999) reported the potential assessment of soil losses by ephemeral gully erosion using aerial imagery. Recently, cutting-edge computer system techniques have been applied in agricultural and environmental applications, such as fuzzy logic. This technique certainly can be used in soil erosion that needs to overcome sharp boundary problems.

Specifically in gully research, Poesen et al. (2003) reported that those features generated between 10 and 95% of total eroded sediment mass in a watershed scale but, conversely, in that study the gullies channels occupied less than 5% of the total watershed area. Not only do these features contribute for the decreasing of the catchment area, but they are also considered effective links of runoff and sediment among watersheds and river systems.

## **1.2 Gully Erosion Process**

Gullies are an extreme form of soil erosion that degrades diverse environments and compromise crop productivity. Predominantly, gully erosion happens in concentrated overland flow. The permanent or classic type of gully is described as a small, steep-sided channel that cannot be crossed by ordinary farm machinery ( FAO, 1965; Soil Science Society of America, 2012) as depicted in Fig. 1.1. Additionally, the weather action, soil formation and land use are strongly correlated with phenomena that carry large amounts of eroded sediments at the valley bottoms.



**Fig 1. 1** (a) Bowl-shaped headcut of the gully viewed looking upstream from the foot of the western gully bank; (b) Wooded lower reaches of the mapped section of the gully looking downstream. (Source: Thomas et al., 2004)

Complementarily, Poesen et al. (2003) pointed out that gullies represent a clear issue of exceeding threshold. In another words, this geomorphic process occurs when rainfall, flow accumulation, and slope are in extreme magnitude. They are also affected by wetting-drying and freezing-thawing cycles (Poel et al., 1986; Thomas et al., 2009)

As a unique geomorphic feature in the landscape, very often the footprints of gullies are not identified accurately by traditional remote sensing techniques. To help answer this question, recently light detection and ranging (LiDAR) has been used to map and estimate gully erosion ( Galzki et al., 2011; James et al., 2007; Perroy et al., 2010). Remote sensing through high resolution ASTER imagery also has been applied to produce qualitative erosion risk maps (Bouaziz et al., 2009) through a combination of the factors that most influence soil erosion (i.e., topography, soil type, land use, climate and vegetation index).

In addition to the gully location issue at large scales, gully widening may occur in different manners. Gully widening is associated with mass-wasting processes of their walls. Specifically for loess and loess alluvium in Western Iowa, three types were described by Bradford and Piest (1980): deep-seated slumps, trapezoidal slab failures and pop-out failures

(Fig. 1.2). The movement in deep-seated slumps is made by slides that collapse along a concave-upward, circular plane, turning in way that the wasting mass moves downward and the toe moves outward. The trapezoidal slab failures are small blocks located on the bank top that are detached by deep tension cracks. Ultimately, the pop-out failures consist of small blocks of soil that detach from the base of the gully wall forming a bowl-shaped alcove. The same authors advocate that the gully-wall collapses are produced by increments of water content through the soil due to rainfall or snowmelt infiltration.



(a) Deep-seated Slump failures

(b) Trapezoidal slab failure

(c) Pop-up failures

**Fig 1. 2.** Schematic representation of main gullies mass-wasting processes gullies in Western Iowa

### 1.3 Precision Conservation

Precision Conservation (PC) deals with the implementation of conservation management practices upon the landscape or agricultural areas. This broader concept embraces the technologies applied in precision farming and focuses not only on maximizing yield but also on interconnecting cycles and flows of energy in order to reduce environment impact (Berry et al., 2003).

The precision conservation idea was conceived under troublesome facts that project increasing demands for food during the 21st century and also a 60% reduction of per capita arable land by the year 2050 (Lal et al.,1999). Because precision conservation uses the same set of technologies of precision farming in an enlarged spatial resolution, it has the potential



to integrate site-specific fields with off-site conservation practices to support watershed sustainability. Therefore, it is crucial to incorporate spatial variability of soil, hydrologic and topographic properties when designing conservation practices such as buffers, filter strips, grassed waterways and terraces. Furthermore, precision conservation must be able to identify features with small footprints, such as gullies. The identification of these critical areas might be a challenge due to the hydrologic complexity and variability across landscapes. However, fine-resolution elevation measurements such as those obtained with LiDAR have been shown to be useful for topographic indices determination under precision conservation context (James et al., 2007).

Related to the yield issue, Quine and Zhang (2002) found a complex relationship between spatial erosion on crop yield. Eroded areas were assessed with lower yields due to nutrient depletion but also lower yields were related to high soil aggregation. Zhou et al. (2011) investigated the effectiveness and cost-benefit of conservation management practices considering sediment reduction. Using the WEPP model, they revealed that additional conservation practices are crucial for reducing sediment when using chisel plow management with a corn-soybean rotation. Although the authors highlighted that the implementation of conservation practice may be expensive, the benefits of soil loss and off-site impacts are valuable. As a regional example, Arbuckle et al. (2011) strongly suggested the use of conservation practices at the farm level to prevent soil fertility loss and sustained resilience. Consequently, these studies indicate clearly that precision conservation practices must respond to spatial erosion under intensive cropping in a practical and tangible way.

## **1.4 Soil Erosion Models**

The mathematical description of the soil erosion has been represented by several models that are useful to design erosion control structures, impoundments, and environmental planning and assessment. The most widely known and used is the Universal Soil Loss Equation (USLE) which estimates a long-term average annual soil loss per unit area by a product of five factors. The factors used in USLE are: rainfall and runoff factor (R), soil erodibility (K), topography (LS), crop cover and management (C) and effect of cultural practices (P). Primarily based on field-scale experiments, USLE and its modified forms do not account for deposition, one of the important processes in a modern soil erosion theory. The lack of sediment deposition estimates inspired new development of process-based simulation models for soil erosion (Nearing et al., 1990). In addition, the USLE is for estimation of sheet and rill erosion only.

Poesen et al. (2003) reported several attempts employed to model gully erosion. One of them was proposed by Sidorchuk (2001) to calculate the amount of sediment eroded. The method, called the stable gully model, computes the final development state of gully flow line networks. In this model, crucial variables for the understanding of the gully are also calculated, such as the bed critical velocity of erosion initiation and the stable slopes of a gully profile. Nevertheless, Poesen et al. (2003) pointed out that, even though those initiatives have been made to develop empirical and process-based models to estimate gully erosion, there is still a need for more research effort to create reliable models.

## **1.5 Data Mining Applied to Soil Erosion**

Agricultural and environmental processes if monitored in high temporal frequency or spatial scale will produce massive data volume. Thus, advanced techniques of data analysis

and modeling are required to visualize and analyze this information producing reliable predictions about those phenomena.

This is the main purpose of an interdisciplinary field called data mining which enables the discovery of new patterns from large datasets (Han & Kamber, 2006). Data mining intersects methods such as artificial intelligence, machine learning, statistics and database systems. Spatial data mining is the branch of data mining that aims to analyze large geographical databases and extract implicit information from spatial data. It is important to point out that the complexity of spatial data and intrinsic spatial relationships limit the usefulness of conventional data mining techniques for extracting spatial patterns, thus, the need for advanced study is strong.

The application of data mining methods for soil erosion issues is recent (Rouet et al., 2009). Processed SPOT satellite images are used to produce erosion hazard maps using attribute ranking and association rules mining. Licznar and Nearing (2003) obtained similar outcomes of soil erosion compared to WEPP model using artificial neural network (ANN) model; however, a high quality and quantity of available data were required. Another application of ANN estimating soil erosion, dissolved phosphorus and nitrate in runoff was provided by Kim and Gilley (2008). Outcomes from this neural network model revealed a positive correlation of eroded sediments with rainfall and runoff. However, a major drawback using ANN studies is a lack of physical relations produced by the model's results which can misguide conclusions about the erosional process.

## 1.6 Objectives

The overall goal of this research was to create an innovative model to assess, with confidence, classic gullies identification and slope stability modelling. The objectives to achieve this goal were to:

1. Develop a methodology to identify areas (pixels) in the landscape that are prone to classical gully erosion, using influential variables of terrain and hydrologic attributes derived from LiDAR digital elevation model. An unsupervised clustering approach was employed as way to group points with most spatial similarity and a new threshold to map gully side walls were proposed (gully concentrated flow, or GCF). Ground truthing of actual gullies was used to validate the model.
2. Evaluate slope stability analysis in two multi-temporal surveys (1999 and 2014) in a gullied area at Western Iowa regarding to precision conservation management.

## References

- Arbuckle Jr., J. G., M. Helmers, M. Liebman, and L. Schulte. 2011. "From Vulnerability to Resiliency: Iowa Agriculture in the Age of Biorenewables." <http://www.extension.iastate.edu/bioeconomy/biocon2/WP6-Arbuckle.pdf>.
- Baruch, A., and S. Filin. 2011. "Detection of Gullies in Roughly Textured Terrain Using Airborne Laser Scanning Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 66 (September). International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS): 564–78. doi:10.1016/j.isprsjprs.2011.03.001.
- Berry, J. K, J. A. Delgado, R. Khosla, and Fran Pierce. 2003. "Precision Conservation for Environmental Sustainability." 1–26.
- Besler, H. 1987. "Slope Properties, Slope Processes and Soil Erosion Risk in the Tropical Rainfall Forest of Kalimantan Timur (Indonesian Borneo)." *Earth Surface Processes and Landforms* 12: 195–204.
- Bouaziz, M., M. Leidig, and R. Gloaguen. 2009. "Optimal Parameter Selection for Qualitative Regional Erosion Risk Monitoring: A Remote Sensing Study of SE Ethiopia." *Geoscience Frontiers* 2(2): 237–45.

- Bradford, J. M., and R. F. Piest. 1977. "Gully Wall Stability in Loess-Derived Alluvium1." *Soil Science Society of America Journal* 41(1): 115–22.
- Bradford, J. M., D. A. Farrel, and W. E Larson. 1973. "Mathematical Evaluation of Factors Affecting Gully Stability." *Soil Science Society of America* 37.
- Bradford, J. M., and R. F. Piest. 1980. "Erosional Development of Valley-Bottom Gullies in the Upper Midwestern United States." In *Geomorphic Thresholds*, edited by D. R. Coates & J. D. Vitek, 75–101. Stroudsburg: Dowden&Culver.
- Campo-Bescós, M. A., J. H. Flores-Cervantes, R. L. Bras, J. Casali, and J. V. Giráldez. 2013. "Evaluation of a Gully Headcut Retreat Model Using Multitemporal Aerial Photographs and Digital Elevation Models." *Journal of Geophysical Research: Earth Surface* 118(4): 2159–73. doi:10.1002/jgrf.20147.
- Chaplot, V. 2013. "Impact of Terrain Attributes, Parent Material and Soil Types on Gully Erosion." *Geomorphology* 186 (March). Elsevier B.V. 1–11. doi:10.1016/j.geomorph.2012.10.031.
- Chen, Su-Chin, Kuang-Tsung Chang, Shi-Hao Wang, and Jun-Yue Lin. 2010. "The Efficiency of Artificial Materials Used for Erosion Control on Steep Slopes." *Environmental Earth Sciences* 62(1): 197–206. doi:10.1007/s12665-010-0514-6.
- Churchill, R. R. 1982. "Aspect-Induced Differences in Hillslope Processes." *Earth Surface Processes and Landforms* 7 (August 1980): 171–82.
- Daggupati, P., A. Y. Sheshukov, and K. R. Douglas-Mankin. 2014. "Evaluating Ephemeral Gullies with a Process-Based Topographic Index Model." *Catena* 113 (February). Elsevier B.V. 177–86. doi:10.1016/j.catena.2013.10.005.
- De Baets, S., J. Poesen, J. Meersmans, and L. Serlet. 2011. "Cover Crops and Their Erosion-Reducing Effects during Concentrated Flow Erosion." *Catena* 85(3). Elsevier B.V. 237–44. doi:10.1016/j.catena.2011.01.009.
- Di Stefano, C., V. Ferro, and P. Porto. 2000. "Length Slope Factors for Applying the Revised Universal Soil Loss Equation at Basin Scale in Southern Italy." *Journal of Agricultural Engineering Research* 75(4): 349–64. doi:10.1006/jaer.1999.0514.
- EPA. 2009. *Soil Protection*. <<http://www.epa.gov/agriculture/ag101/cropsoil.html#map>> .
- Eustace, A., M. Pringle, and C. Witte. 2009. "Give Me the Dirt: Detection of Gully Extent and Volume Using High-Resolution LiDAR." In *Innovations in Remote Sensing and Photogrammetry*, edited by K. Jones, S., Reinke, 255–69. Berlin Heidelberg: Springer-Verlag.

- Evans, J. S., and A. T. Hudak. 2007. "A Multiscale Curvature Algorithm for Classifying Discrete Return LiDAR in Forested Environments." *IEEE Transactions on Geoscience and Remote Sensing* 45(4): 1029–38. doi:10.1109/TGRS.2006.890412.
- Food and Agriculture Organization (FAO). 1965. *Soil Erosion by Water. Some Measures for Its Control on Cultivated Lands. FAO Agriculture Paper 81*. Rome.
- Galzki, J. C., a. S. Birr, and D. J. Mulla. 2011. "Identifying Critical Agricultural Areas with Three-Meter LiDAR Elevation Data for Precision Conservation." *Journal of Soil and Water Conservation* 66(6): 423–30. doi:10.2489/jswc.66.6.423.
- Ghilani, C. D., and Paul R. Wolf. 2006. *Adjustment Computation - Spatial Data Analysis*. Edited by John Willey & Sons, Inc.. 4th Edition. Hoboken, New Jersey.
- Glenn, N. F., David R. S., D. J. Chadwick, G. D. Thackray, and S. J. Dorsch. 2006. "Analysis of LiDAR-Derived Topographic Information for Characterizing and Differentiating Landslide Morphology and Activity." *Geomorphology* 73(1-2): 131–48. doi:10.1016/j.geomorph.2005.07.006.
- Goddard, T. W. 2005. "An Overview of Precision Conservation in Canada." *Journal of Soil and Water Conservation* 60(6): 456–61.
- Gullo, F., and A. Tagarelli. 2012. "Uncertain Centroid Based Partitional Clustering of Uncertain Data." *Proceedings of the VLDB Endowment* 5(7): 610–21. doi:10.14778/2180912.2180914.
- Han, J, and M. Kamber. 2006. *Data Mining: Concepts and Techniques*. 2nd ed. Amsterdam ; Boston : Elsevier ; San Francisco: Morgan Kaufmann.
- Hengl, T., and H. I. Reuter. 2009. *Geomorphometry: Concepts, Software, Applications*. Edited by T Hengl and H I Reuter. *Developments in Soil Science*. Vol. 33. Developments in Soil Science. Elsevier. doi:10.1016/S0166-2481(08)00036-6.
- Hoek, E., and J. W. Bray. 1981. *Rock Slope Engineering*. Revised Th. The institution of mining and metallurgy.
- Istanbulluoglu, E. 2005. "Implications of Bank Failures and Fluvial Erosion for Gully Development: Field Observations and Modeling." *Journal of Geophysical Research* 110(F1): F01014. doi:10.1029/2004JF000145.
- Istanbulluoglu, E., R. L. Bras, and H. Flores-Cervantes. 2005. "Implications of Bank Failures and Fluvial Erosion for Gully Development: Field Observations and Modeling." *Journal of Geophysical Research* 110(F1): F01014. doi:10.1029/2004JF000145.

- Istanbulluoglu, E., O. Yetemen, E. R. Vivoni, Hugo a. Gutiérrez-Jurado, and R. L. Bras. 2008. "Eco-Geomorphic Implications of Hillslope Aspect: Inferences from Analysis of Landscape Morphology in Central New Mexico." *Geophysical Research Letters* 35(14): L14403. doi:10.1029/2008GL034477.
- James, L., D. Watson, and W. Hansen. 2007. "Using LiDAR Data to Map Gullies and Headwater Streams under Forest Canopy: South Carolina, USA." *Catena* 71(1). Elsevier: 132–44. doi:10.1016/j.catena.2006.10.010.
- James, L. A., D. G. Watson, and W. F. Hansen. 2007. "Using LiDAR Data to Map Gullies and Headwater Streams under Forest Canopy: South Carolina, USA." *Catena* 71(1): 132–44. doi:10.1016/j.catena.2006.10.010.
- Johansen, K., S. Taihei, D. Tindall, S. Phinn, Joint Remote, Biophysical Remote, Sensing Group, et al. 2012. "Object-based Monitoring of Gully Extent and Volume in North Australia using LIDAR data." In *GEOBIA*, I:168–73. Rio de Janeiro.
- Kakembo, V., W. W. Xanga, and K. Rowntree. 2009. "Topographic Thresholds in Gully Development on the Hillslopes of Communal Areas in Ngqushwa Local Municipality, Eastern Cape, South Africa." *Geomorphology* 110 (3-4). Elsevier B.V. 188–94. doi:10.1016/j.geomorph.2009.04.006.
- Kaufman, L., and P. J. Rousseeuw. 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York, NY: John Wiley & Sons.
- Kdrilen, D. L., L. A. Kramer, D. E. James, D. D. Buhler, and M. R. Burkart. 1964. "Field-Scale Watershed Evaluations on Deep-Loess Soils:"
- Kim, M., and J. E. Gilley. 2008. "Artificial Neural Network Estimation of Soil Erosion and Nutrient Concentrations in Runoff from Land Application Areas." *Computers and Electronics in Agriculture* 64(2): 268–75.
- Kirkby, M. J., and L. J. Bracken. 2009. "Gully Processes and Gully Dynamics." *Earth Surface Processes and Landforms* 34: 1841–51. doi:10.1002/esp.
- Knapen, A., and J. Poesen. 2009. "Soil Erosion Resistance Effects on Rill and Gully Initiation Points and Dimensions." *Earth Surface Processes and Landforms* 228 (December 2009): n/a–n/a. doi:10.1002/esp.1911.
- Lal, R., D. Mokma, and B. Lowery. 1999. "Relation between Soil Quality and Soil Erosion." In *Soil Quality and Soil Erosion*, CRC Press, 237–58. Boca Raton, Florida.
- Licznar, P., and M. A. Nearing. 2003. "Artificial Neural Networks of Soil Erosion and Runoff Prediction at the Plot Scale." *Catena* 51(2): 89–114. doi:10.1016/S0341-8162(02)00147-9.

- Lohnes, R. A., and R. L. Handy. 1968. "Slope Angles in Friable Loess." *The Journal of Geology* 76(3): 247–58.
- Lucà, F., M. Conforti, and G. Robustelli. 2011. "Comparison of GIS-Based Gullying Susceptibility Mapping Using Bivariate and Multivariate Statistics: Northern Calabria, South Italy." *Geomorphology* 134 (November). Elsevier B.V. 297–308. doi:10.1016/j.geomorph.2011.07.006.
- Mararakanye, N., and N. S. Nethengwe. 2012. "Title: Gully Erosion Mapping Using Remote Sensing Techniques in the Capricorn District, Limpopo." *South African Journal of Geomatics* 1(2): 109–18.
- Maroco, J., D. Silva, A. Rodrigues, M. Guerreiro, I. Santana, and A. de Mendonça. 2011. "Data Mining Methods in the Prediction of Dementia: A Real-Data Comparison of the Accuracy, Sensitivity and Specificity of Linear Discriminant Analysis, Logistic Regression, Neural Networks, Support Vector Machines, Classification Trees and Random Forests." *BMC Res Notes* 4: 299.
- Molina, A., G. Govers, F. Cisneros, and V. Vanacker. 2009. "Vegetation and Topographic Controls on Sediment Deposition and Storage on Gully Beds in a Degraded Mountain Area." *Earth Surface Processes and Landforms* 767: 755–67. doi:10.1002/esp.
- Momm, H. G., R. L. Bingner, R. R. Wells, J. R. Rigby, and S. M. Dabney. 2013. "Effect of Topographic Characteristics on Compound Topographic Index for Identification of Gully Channel Initiation Locations." *Transactions of the ASABE* 56(2): 523–37.
- Montgomery, D. R., and W. E. Dietrich. 1992. "Channel Initiation and the Problem of Landscape Scale." *Science* (New York, N.Y.) 255 (5046): 826–30. doi:10.1126/science.255.5046.826.
- Montgomery, D. R., and William E. Dietrich. 1988. "© 198 8 Nature Publishing Group." *Nature* 336(17): 232–34.
- Mortazavi, R., and S. Jalili. 2014. "Fast Data-Oriented Microaggregation Algorithm for Large Numerical Datasets." *Knowledge-Based Systems*, May. Elsevier B.V. doi:10.1016/j.knosys.2014.05.011.
- Nachtergaele, J., and J. Poesen. 1999. "Assessment of Soil Losses by Ephemeral Gully Erosion." *Earth Surface Processes and Landforms* 24: 693–706.
- Nearing, M. A., L. Deer-Ascough, and J. M. Laflen. 1990. "Sensitivity Analysis of the WEPP Hillslope Profile Erosion Model." *Transactions of the ASAE* 33(3): 839–49.



- Oostwoud W., D. J., and R. Bryan. 2001. "Gully-Head Erosion Processes on a Semi-Arid Valley Floor in Kenya: A Case Study into Temporal Variation and Sediment Budgeting." *Earth Surface Processes and Landforms* 26(9): 911–33. doi:10.1002/esp.225.
- Patton, P. C., and S. A. Schumm. 1975. "Gully Erosion, Northwestern Colorado : A Threshold Phenomen." *Geology* 3(2): 88–90.
- Perroy, R. L., B. Bookhagen, G. P. Asner, and O. A. Chadwick. 2010. "Comparison of Gully Erosion Estimates Using Airborne and Ground-Based LiDAR on Santa Cruz Island, California." *Geomorphology* 118(3-4). Elsevier B.V. 288–300. doi:10.1016/j.geomorph.2010.01.009.
- Pike, A., T. Mueller, E. Rienzi, S. Neelakantan, B. Mijatovic, T. Karathanasis, and M. Rodrigues. 2012. "Research on Soil Erosion." In , edited by Danilo Godone and Silvia Stanchi., 45–63. In Tech Open Access on-line book. (<http://dx.doi.org/10.5772/51526>).
- Pimentel, D., C. Harvey, P. Resosudarmo, K. Sinclair, D. Kurz, M. Mcnair, S. Crist, et al. 1995. "Environmental and Economic Costs of Soil Erosion and Conservation Benefits." *Science*, 267: 1117–23.
- Poel, P. Van Der, R G Spomer, and R F Piest. 1986. "Slope Indicator Measurements of Subsurface Movement in Gully Walls." *Transactions of the ASAE* 29: 982–87.
- Poesen, J., J. Nachtergaele, G. Verstraeten, and C. Valentin. 2003. "Gully Erosion and Environmental Change: Importance and Research Needs." *Catena* 50(2-4): 91–133. doi:10.1016/S0341-8162(02)00143-1.
- Quine, T. A., and Y. Zhang. 2002. "An Investigation of Spatial Variation in Soil Erosion, Soil Properties, and Crop Production within an Agricultural Field in Devon, United Kindom." *Journal of Soil Water Conservation* 57(55-64).
- Rengers, Francis K, and G. E. Tucker. 2014. "Analysis and Modeling of Gully Headcut Dynamics , North American High Plains." *Journal of Geophysical Research: Earth Surface*, 1–21. doi:10.1002/2013JF002962.Received.
- Rouet, I., D. M. Allenbach, Gay N. Selmaoui, A. G. Ausseil, M. Mangeas, Maura J. P. Dumas, and G. Lille. 2009. "Tools for Regional Scale Soil Erosion Mapping and Hazard Assessment 2p Application to New Caledonia." In *In Proceedings of 18th World IMACS/MODSIM Congress*. Cairns, Australia.
- Santisteban, L. M., De J. Casalí, J. J. López, J. V. Giráldez, J. Poesen, and J. Nachtergaele. 2005. "Exploring the Role of Topography in Small Channel Erosion." *Earth Surface Processes and Landforms* 30(5): 591–99. doi:10.1002/esp.1160.

- Shit, P. K., and R. K. Maiti. 2012. "Mechanism of Gully-Head Retreat - A Study at Ganganir Danga, Paschim Medinipur, West Bengal." *Ethiopian Journal of Environmental Studies and Management* 5(4): 332–42.
- Shruthi, R. B.V., N. Kerle, and V. Jetten. 2011. "Object-Based Gully Feature Extraction Using High Spatial Resolution Imagery." *Geomorphology* 134 (3-4). Elsevier B.V. 260–68. doi:10.1016/j.geomorph.2011.07.003.
- Shruthi, R. B. V., N. Kerle, V. Jetten, and A. Stein. 2014. "Object-Based Gully System Prediction from Medium Resolution Imagery Using Random Forests." *Geomorphology* 216 (July). Elsevier B.V. 283–94. doi:10.1016/j.geomorph.2014.04.006.
- Sidorchuk, A. 2001. "GULTEM - The Model to Predict Gully Thermoerosion and Erosion (Theoretical Framework )" 1: 966–72.
- Simon, A., and P. O. Box. 2006. "Study of the Effects of Lateral Seepage Forces on Tension-Crack Development, Bank-Failure Dimensions and Migration of Edge of Field Gullies." In *PROCEEDINGS of the Eighth Federal Interagency Sedimentation Conference*, 660–66. Reno.
- Skempton, A. W., and J. N. Hutchinson. 1969. "Stability of Natural Slopes and Embankment Foundations." In *Proc. 7th Int. Conf. on Soil Mechanics.State of the Art Report.*, 291–340.
- Soil Science Society of America. 2012. "Glossary of Soil Science Terms." *Oil Science Society of America, Madison, WI*. <https://www.soils.org/publications/soils-glossary#>.
- Staff, Soil Survey. 2014. "Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey." <http://websoilsurvey.nrcs.usda.gov/>.
- Svoray, T., E. Michailov, A. Cohen, L. Rokah, and A. Sturm. 2012. "Predicting Gully Initiation: Comparing Data Mining Techniques, Analytical Hierarchy Processes and the Topographic Threshold." *Earth Surface Processes and Landforms* 37(6): 607–19. doi:10.1002/esp.2273.
- Swift, L.W., K. J. Elliott, R. D. Ottmar, and R. E. Vihnanek. 1993. "Site Preparation Burning to Improve Southern Appalachian Pine-Hardwood Stands: Fire Characteristics and Soil Erosion Moisture, and Temperature." *Canadian Journal of Forest Research* 23(10): 2242–54,.
- Team, R Development Core. 2011. "R: A Language and Environment for Statistical Computing". Vienna, Austria: R Foundation for Statistical Computing. <http://www.r-project.org/>.

- Thiebes, B., R. Bell, T. Glade, S. Jäger, M. Anderson, and L. Holcombe. 2013. "A WebGIS Decision-Support System for Slope Stability Based on Limit-Equilibrium Modelling." *Engineering Geology* 158 (May). Elsevier B.V. 109–18. doi:10.1016/j.enggeo.2013.03.004.
- Thomas, J. T., N. R. Iverson, and M. R. Burkart. 2009. "Bank-Collapse Processes in a Valley-Bottom Gully," *Earth Surface Processes and Landforms*, 34: 109–22. doi:10.1002/esp.
- Thomas, J. T., N. R. Iverson, M. R. Burkart, and L. A. Kramer. 2004. "Long-Term Growth of a Valley-Bottom Gully, Western Iowa." *Earth Surface Processes and Landforms* 29(8): 995–1009. doi:10.1002/esp.1084.
- Tomer, M. D., D. W. Meek, and L. A. Kramer. 2005. "Agricultural Practices Influence Flow Regimes of Headwater Streams in Western Iowa." *Journal of Environmental Quality* 34(5): 1547–58. doi:10.2134/jeq2004.0199.
- Tomer, M. D., T. B. Moorman, J. L. Kovar, D. E. James, and M. R. Burkart. 2007. "Spatial Patterns of Sediment and Phosphorus in a Riparian Buffer in Western Iowa." *Journal of Soil and Water Conservation* 62(5): 329–38.
- Tomer, M. D., C. A. Cambardella, D. E. James, and T. B. Moorman. 2006. "Surface-Soil Properties and Water Contents across Two Watersheds with Contrasting Tillage Histories." *Soil Science Society of America Journal* 70(2): 620. doi:10.2136/sssaj2004.0355.
- Toy, T. J., G. Foster, and K. G. Renard. 2002. *Soil Erosion. Processes, Prediction, Measurement and Control*. Edited by John Willey and Sons. New York.
- Valentin, C., J. Poesen, and Yong Li. 2005. "Gully Erosion: Impacts, Factors and Control." *Catena* 63 (2-3): 132–53. doi:10.1016/j.catena.2005.06.001.
- Vandaele, K., J. Poesen, G. Govers, and B. V. Wesemael. 1996. "Geomorphic Threshold Conditions for Ephemeral Gully Incision." *Geomorphology* 16: 161–73.
- Vandekerckhove, L., J. Poesen, D. Oostwoud Wijdenes, and T. de Figueiredo. 1998. "Topographical Thresholds for Ephemeral Gully Initiation in Intensively Cultivated Areas of the Mediterranean." *Catena* 33(3-4): 271–92. doi:10.1016/S0341-8162(98)00068-X.
- Vandekerckhove, L., J. Poesen, D. Oostwoud Wijdenes, J. Nachtergaele, C. Kosmas, M. J. Roxo, and T. De Figueiredo. 2000. "Thresholds for Gully Initiation and Sedimentation in Mediterranean Europe." *Earth Surface Processes and Landforms* 25: 1201–20.

- Vendrusculo, L. G., and A. L. Kaleita. 2013. "Terrain Analysis and Data Mining Techniques Applied to Location of Classic Gully in a Watershed." *2013 Kansas City, Missouri, July 21 - July 24, 2013*. St. Joseph, MI: American Society of Agricultural and Biological Engineers. doi:10.13031/aim.20131619828.
- Vrieling, A., S. C. Rodrigues, H. Bartholomeus, and G. Sterk. 2007. "Automatic Identification of Erosion Gullies with ASTER Imagery in the Brazilian Cerrados." *International Journal of Remote Sensing* 28(12): 2723–38. doi:10.1080/01431160600857469.
- Watts, J. D., S. L. Powell, R. L. Lawrence, and T. Hilker. 2011. "Improved Classification of Conservation Tillage Adoption Using High Temporal and Synthetic Satellite Imagery." *Remote Sensing of Environment* 115(1): 66–75.
- Weaver, A. V. B. 1991. "The Distribution of Soil Erosion as a Function The Distribution of Slope Aspect and Parent Material in Ciskei , Africa Southern." *GeoJournal* 23 (1): 29–34.
- Wilson, J. P., and J. C. Gallant. 2000. "Digital Terrain Analysis." In *Terrain Analysis and Applications*, 479. John Willey & Sons.
- Wischmeier, W.H., and D D Smith. 1978. *Predicting Rainfall Erosion Losses — a Guide to Conservation Planning*. Agricultur. Washington DC: US Department of Agriculture.
- Wysocki, D. A., P. J. Schoeneberger, and H.E. LaGarry. 2000. "Geomorphology of Soil Landscapes." In *Handbook of Soil Science*, 1–39. New York, NY: CRC Pres.
- Zhou, X., M. J. Helmers, M. Al-Kaisi, and H. M. Hanna. 2011. "Cost-Effectiveness and Cost-Benefit Analysis of Conservation Management Practices for Sediment Reduction in an Iowa Agricultural Watershed." *Journal of Soil and Water Conservation* 64(5): 314–23.

## CHAPTER 2. IDENTIFICATION OF CLASSIC GULLY USING AIRBORNE LiDAR DATA, TERRAIN ANALYSIS, AND SPATIAL CLUSTERING TECHNIQUES UNDER DENSE VEGETATION AREA

A paper to be submitted to the Transactions of the *ASABE Journal*  
Laurimar Gonçalves Vendrusculo, Amy Kaleita, Monica Haddad

### Abstract

LiDAR high-resolution elevation products have been providing an important source of information to enable researchers to identify small erosional landscape features. Such features, especially classical gullies, are responsible for delivering a large percentage of sediment to streams, depending on the scale of the watershed. Identification of gullies is very likely to optimize placement of soil and water conservation structures. It was hypothesized that gullies throughout a given terrain have a similar topography, hydrography, and paths of concentrated flow erosion. Thus, locating gullies in natural landscapes can be performed by using spatial combinations of those parameters computed from high-resolution LiDAR elevation along with zones of concentrated flow (e.g., gully bed). Using a 0.91 Km<sup>2</sup> region of the Keigley Branch-South Skunk River (KBSSR) watershed, an area populated with classical gullies, we computed profile curvature, mean slope deviation, stream power index, and aspect at a 1-m pixel level. CLARA (CLustering LARge Applications) algorithm, an unsupervised clustering approach, was employed on 913,495 points splitting the dataset in six groups. The partition number is in agreement with within-group sum of squared error (WSS) statistical technique. Outcomes from MANOVA analysis (Wilks F= 0.33,  $p < 0.0001$ ) provided strong evidence of non-similarity among six groups. In addition, a new threshold criteria termed gully concentrated flow (GCF) based upon data distribution of flow accumulation and mean overall slope identified the main hydrographic flow paths,

correspond to the gullies beds. Points belonging to cluster #6 were classified as gully side walls because they surrounded GCF lines and also it had maximum first quartile of mean slope deviance among other clusters. The whole gully feature was classified intercepting GCF lines with cluster #6 area.

## 2.1 Introduction

Accuracy in locating classical gullies increases the efficiency of targeting conservation practices for reducing soil erosion from agricultural fields and drainage areas. Generally, classical gully erosion (e.g., permanent gullies) is very likely to degrade precious water and soil resources through transport and deposition of topsoil with high organic matter (Kakembo, Xanga, and Rowntree 2009; Poesen et al. 2003). Previous studies have indicated that sediments from small gully features account for 10 to 94% of total sediment yield in some watersheds (Poesen et al., 2003; Thomas et al., 2004)

Gully channels are characterized by U-, V-, and trapezoidal-shaped cross-sections and steep wall sides. Development of these gullies is predominantly controlled by surface runoff that produces concentrated overland flows that cut channels (Kirkby and Bracken, 2009). However, several studies have shown that other factors, including topographic and hydrological attributes, soil parent material and land use can also control gully development (Chaplot 2013; Poesen et al., 2003; Valentin et al. Li 2005; Vandaele et al., 1996; Vandekerckhove et al., 1998, 2000).

Despite the significance of these features, only a few erosion models account for such channel-sediment losses. These include: AnnAgNPS (Annual Agricultural Non-Point Source pollution), WEPP (Water Erosion Prediction Project), EGEM (ephemeral gully erosion model), and CREAMS (Chemicals, Runoff, and Erosion from Agricultural Management

Systems). Most of these models compute sediment yield initiated by small open channels or ephemeral gullies. They also require user determination of gully locations. Therefore, efforts using the compound topographic index (CTI) threshold have been initiated to implement automatic location of ephemeral gully channel starting points (Momm et al., 2013; Svoray et al., 2012). An approach for identifying the location and length of ephemeral gullies incorporating a process-based simulation of overland flow and combining factors such as drainage area, surface roughness, slope, and soil critical shear stress has been proposed by Daggupati et al. (2014).

One source for predetermined classical gully locations and other erosional linear or point features (e.g., escarpment, gravel pit, boulder spot, sinkhole, etc.) in the United States could be based on the electronic database entitled Natural Resources Conservation Service Soil Survey Geographic – SSURGO (Soil Survey Staff 2014). A classical gully, for example, is defined by a line in a shapefile format. Although there is a rigid certification process underlying this database, intricate gully morphology obtained from SSURGO data is not clearly revealed. Furthermore, this database is static and does not account for gully changes or development of new features.

For more detailed mapping of intricate gully morphology, recently-developed technologies such as airborne and ground LiDAR (Light Detection And Ranging) cloud points may be useful in modeling and mapping soil erosion patterns. The advantages of LiDAR compared to GPS, satellite imagery and field surveying are mainly: high vertical and horizontal accuracy, fast acquisition, able to collect elevation data even in dense canopy, no-weather limited, and higher data density among others. Although this is a potentially useful data source, for long-term monitoring applications in large areas this technology is quite

expensive and, thus, feasible primarily for governmental institutions (Baruch and Filin 2011). Recent studies related to water and soil quality have applied LiDAR, with the majority utilizing first and second derivatives of elevation to produce slope, curvature, or stream-power index. For example, James et al. (2007) tested the capability of LiDAR technology for mapping headwater channels and gully systems under dense vegetation in South Carolina, and further validated the approach with field surveys using differential GPS and cross sections. Galzki, Birr, and Mulla (2011), processed LiDAR-based terrain attributes to identify ephemeral gullies in hydrologically-connected areas located in Minnesota watersheds. Using binary classification (e.g., gully, non-gully), a study conducted by Baruch and Filin (2011) presented an unsupervised model for gully extraction within alluvial landscapes where LiDAR-derived terrain roughness was used to detect gully presence. That study also included concentrated flow path tracking and flow path selection to ascertain flow connectivity. Generating 0.5-m resolution digital elevation models from LiDAR transects, Eustace, Pringle, and Witte (2009) classified gully presence and volume for each pixel at the Fitzroy catchment in Queensland. Furthermore, a random-forest algorithm was used in part of transects sections to discover correlation between gullies and ancillary variables. In a few selected sites of the same catchment, Johansen et al. (2012) compared gully extent over a 3-year period. They used an image-processing approach in which a pixel-based object-resizing algorithm identified gully edges from DTM raster images.

The amount of data generated by LiDAR systems can reach a few millions points per square kilometer, and standard desktop computer systems are not prepared to extract meaningful knowledge from these massive datasets. To handle such amounts of data, advances in data-mining are being developed to enable the discovery of new patterns from



large datasets (Han and Kamber, 2006). Data mining integrates methods such as artificial intelligence, machine learning, statistics, and database systems. Spatial data mining is a branch of data mining that focuses on analyzing large geographical databases and extracting implicit information from spatial data. The complexity of spatial data and intrinsic spatial relationships limit the usefulness of conventional data-mining techniques for extracting spatial patterns, thus specific data mining tools for spatial data handling are needed. Shruthi et al. (2011) applied a rule-set of slope, drainage area, NDVI and texture contrast to identify gully edges from IKONOS and GEOEYE-1 images in a sub-humid region in Morocco. Although their algorithm is able to successfully detect gully main lines, it has a few limitations. It is topographically region-dependent, requires user-defined image thresholds, and an empirical process to remove false positives. Subsequently, to improve subjectivity caused by threshold determination, Shruthi et al. (2014) used Random Forest object-oriented analysis, a semi-automated classification method to identify gully boundaries from medium-resolution imagery. Svoray et al. (2012) also obtained better prediction capability with a decision-tree algorithm rather than a traditional topographic threshold technique. This database included environmental, climatic and human-induced variables. Mararakanye and Nethengwe (2012) tested image objective, an image-based feature-extraction technique, at gullied sites obtained from SPOT 5 satellite imagery. Vendrusculo and Kaleita (2013) compared three unsupervised classification approaches: centroid (K-means and Fuzzy K-means) and medoid (CLARA) to identify gully location in a field scale watershed at a Treynor, Iowa site. A 1-m digital elevation model was used to calculate topographic and hydrologic attributes. At a much larger pixel resolution (4 x 4 km), Vrieling et al. (2007) used ASTER imagery to identify gully sites in a Brazilian cerrados, a vast tropical savanna,

using a maximum-likelihood supervised classification approach. Based in traditional statistics, Lucà et al.(2011) employed bivariate and logistic regression analyses to test occurrence of gullies incorporating in their model diverse attributes such as lithology, land use, slope, and aspect-splitting point and polygon observation in training and validation sets.

In applications where part of the objects population is previously labeled, supervised classification is employed. This involves grouping of observations that are similar to one another. A supervised approach then uses a training set of observations to produce an inferring function for classifying unlabeled observations. Conversely, few studies have employed unsupervised approaches to classify gully erosion features (Shruthi et al., 2011, 2014). Unsupervised classification searches for hidden structure in unlabeled data. Because they are unique geomorphic features in the landscape, the footprints of gullies are very often not accurately identified at large scale. Determining their structure without prior knowledge fits into an unsupervised classification approach that was applied in the current study.

We selected the clustering large application (CLARA) algorithm, proposed by Kaufman and Rousseeuw (1990), to discern gully-clustering among diverse landscape features. To deal with computer memory limitations, CLARA draws multiple samples from the dataset. From each sample, it arbitrarily chooses  $k$  objects, in the initial parameter definition, to form a number of clusters from the original dataset as representative objects and then assigns each remaining object to the group at the nearest distance. Then, for each representative object, the algorithm randomly selects a non-representative object and computes the total cost of swapping a representative object with the random object. The group memberships are updated accordingly and the process continues until a specific criterion is reached. The CLARA's drawback lies in its dependence on sample size and

sample quality (e.g., unbiased sample) as well as the requirement that the user set up number of clusters.

The CLARA algorithm is a k-medoid method. These methods minimize the average dissimilarities between cluster points to the most representative point in the cluster. A medoid point is one of the original points corresponding to the one least dissimilar from all data points. A different approach is given to centroid methods (k-centers). To identify the centroid points, the arithmetic average of point position individually for each dimension (variable) is computed by each cluster. The computed centroid is not necessarily one of the original data points. (Gullo and Tagarelli, 2012). However, according Kaufman and Rousseeuw (1990), k-medoid based algorithms are superior to k-centroid methods because they have a more robust sum of squared error A compared to a centroid approach. In addition, the clusters are not too elongated which allow for better characterization of all points within clusters. One consequence of a more compact cluster shape is the outliers are more easily isolated which can form clusters with more internal similarity.

We hypothesized that gullies throughout a given terrain have similar topography, hydrography, and size and shape of concentrated flow paths in its bed. Thus, identification of gullies in the landscape might be accurately performed by using spatially- combinations of those parameters computed from high-resolution LiDAR elevation associated with zones concentrated flow (e.g., gully bed). Most studies have also focused on DEM-based image processing rather than on assigning attribute values to the LiDAR cloud points mostly because raster image processing is faster. However, we developed a novel technique that initially locates high slope gradient areas based on a data-mining approach using topographic and hydrologic parameters at pixel points computed from airborne LiDAR data, and later

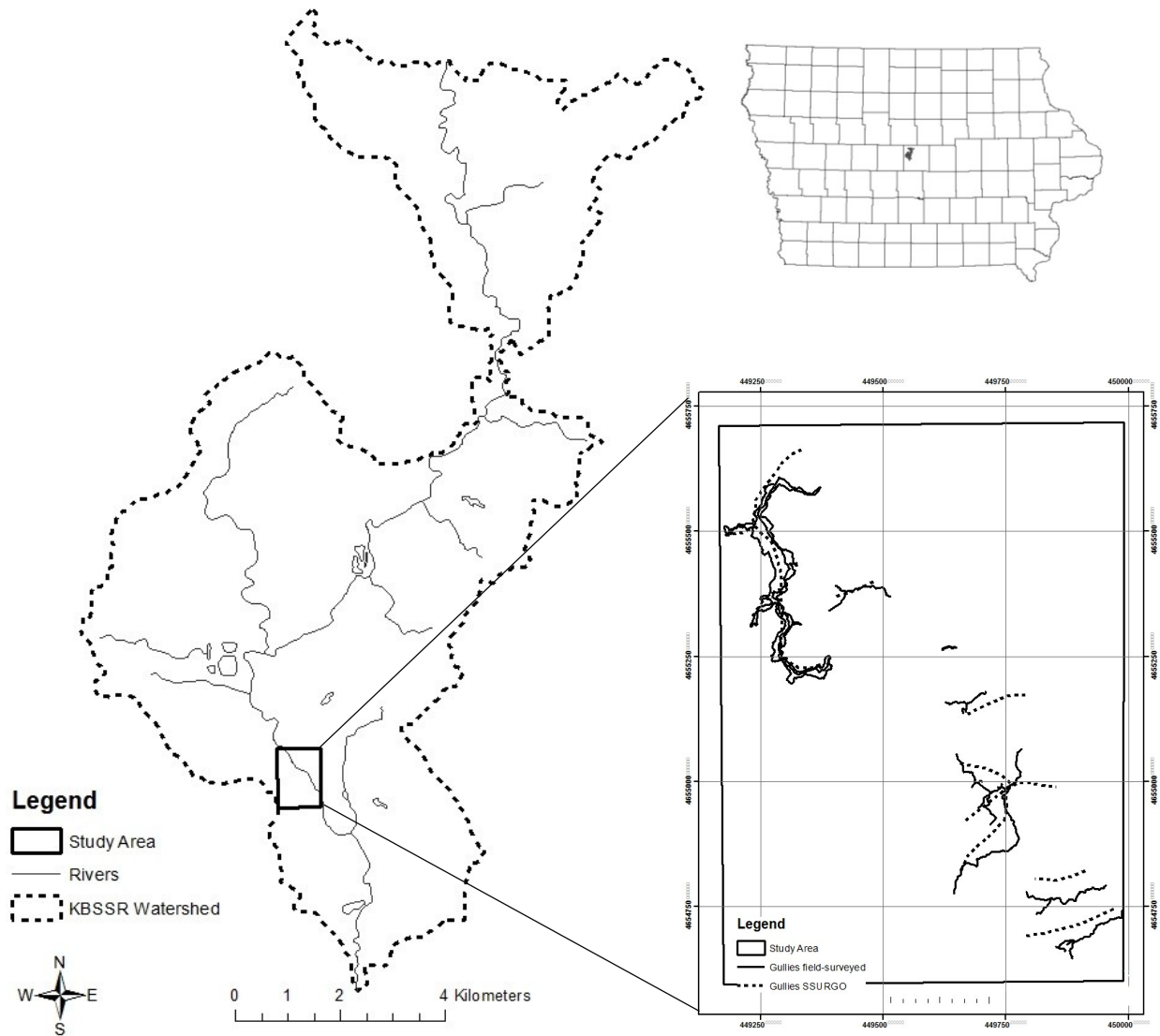
refines the gully location by its concentrated flow path within gully bed. The methodology was validated through ground truth field survey. The main contributions of this study are to: (a) to present a methodology to process gullies dataset, and (b) to evaluate the accuracy of the proposed technique. The next section describes the study area, LiDAR dataset processing, and data-mining techniques.

## **2.2 Material and Methods**

### **2.2.1 Area description**

A study area was selected with a high density of classical incised gullies with steep cross-sections. Most of the gullies in the study area have a well-defined thalweg which represents the line of lowest elevation usually in the gully bed. This area (449,166.511 m E, 4,655,742.51 m N) corresponds to a 0.91 Km<sup>2</sup> portion of the Keigley Branch-South Skunk River (KBSSR) watershed located east of the city of Ames, Iowa as shown in Fig. 2.1. This area is currently used only for recreational purposes.

The study area is located in the Iowa and Minnesota Till prairies common resource area (CRA). This region as defined by United States Department of Agriculture – Natural Resources Conservation Service (USDA – NRCS) is characterized by loamy soils formed in glacial till, with unique features such as potholes and lacustrine areas. The mean annual rainfall ranges from 790 to 840 mm. Rainfall distribution is irregular, with high-intensity rainstorms during late spring and summer seasons, as reported in the soil survey of Story County, Iowa. The annual daily average temperature is 8.3<sup>o</sup>C. Hayden-Storden complex, Coland-Terril complex, Lester, and Hayden are the most extensive soil types in the study area landscape. Hayden soils vary from gently sloping to very steep, are typically well-drained on the upland side, and typically comprise a considerable number of gullies and deep



**Fig. 2.1** Site map showing the location of KBSSR watershed among Iowa counties, USA and detail of study area with gully networks identified by SSURGO lines (dotted lines) and field-surveyed patches (bold-lines). Note that the largest gully in the top left of inset figure it was surveyed in its top edge and floor flow patch because it was in a trapezoidal shape.

drainage ways. Lester soils follow the same slope pattern as Hayden soils; they furthermore present convex side slopes along streams and upland drainage ways. Conversely, Storden soils primarily exhibit convex side slopes bordering similar landscapes.

Analysis of SSURGO gully lines in this area, as a starting point, revealed that all the pre-identified gully lines lie only in Hayden-Storden loams (25 to 50% slopes). Therefore, there is no variation in soil properties under these gullies in the SSURGO data. Without supplemental soil data, we were unable to incorporate any variables related to soil properties.

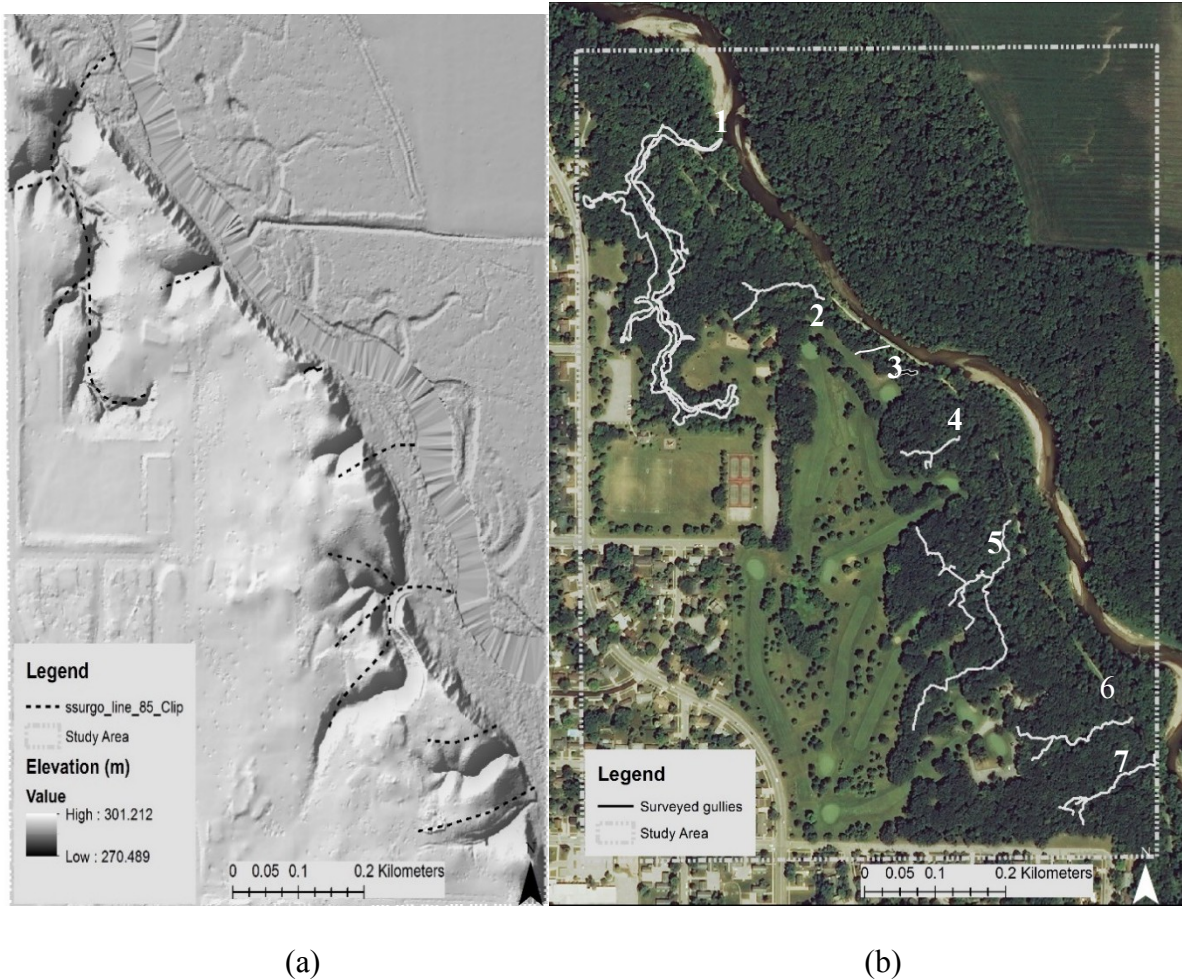
The study area is comprised of a mixed collection of isolated and networked classical gullies. Field surveys were carried out at the study site to delineate the perimeters of the bank and bottom line of the gullies as illustrated in Fig.2.1. Surveys were mainly performed in March and April of 2014, in early spring before leaf-out. Therefore, access to the top edges and bottom area and observation of morphological aspects were facilitated due to lack of leaves on shrubs and trees. A few rainfall events during the survey campaign revealed active erosion process due to recent landslides, and sediment trace deposition and erosion in the drainage floor along with well-defined sharp walls/ sediment transport occurring mainly in the gully floors. V and trapezoidal-shaped gullies were found in the study site. Our strategy during this gully survey was to first delineate only the bottom of V-shaped gullies (Fig 2.2a) because their walls were steep. For trapezoidal-shaped gullies, both the edge and bottom points were collected. The survey was done using a Garmin eTrex legend handheld GPS receiver with 7 meter precision. We initially collected elevation and coordinate points, and in a post-processing procedure those points were transformed into polylines through ArcGis 10. There were sites with visible high-flow concentration; showing recent soil loss in wall sides (Fig. 2.2b).



**Fig. 2.2** (a) Example of V-shaped gully and (b) concentrated flow path between two gully branches. Images taken at leaf-off condition in April, 2014. As shown in the images, the area is fully covered by leaves and hardwoods and the gully floors are still muddy and exhibit ephemeral flow.

The six non-connected gullies in the study area can be visually distinguished in the hillshade image depicted in Fig 2.3a. However, an additional gully (number 3 in Fig. 2.3b) was noticed in the study area. In addition, it can be noted that this area is under dense canopy vegetation (Fig. 2.3b).

During the survey activity, we observed gully faces from recent landslides showing exposed tree roots as well as drainage bottoms with unconsolidated soil material mostly under dense layers of dry leaves. Furthermore, head cuts were observed at an early developmental stage, confirming that gullies were active. A small amount of flow was noted at the widest gully bottom while only saturated surfaces or muddy-covered areas were found in the remainder. However, further investigation was performed to attempt to determine controlling processes of flow within gullies.



**Fig. 2.3** (a) Elevation hillshade map of KBSSR portion area at Story County, Iowa, derived from the LiDAR One Meter Digital Elevation Model highlighting gullies classified at SSURGO (dotted lines); (b) 2013 National Agriculture Imagery Program (NAIP) Four-Band Aerial Photography with field-surveyed pathways (bold lines).

## 2.3 Data Processing

### 2.3.1 Data collection and preparation

An Airborne LiDAR dataset, originally at 1.4 m average bare-earth data spacing and later gridded to 1 m, was obtained for the study area from the Iowa LiDAR Consortium (available at archive <http://geotree2.geog.uni.edu/lidar/>). The point clouds were in a LASer file format (LAS) containing X and Y coordinates (UTM Zone 15N nad83), orthometric elevation Z (NADV88), return level (1, 2, or 12), and intensity (0-255). More than 913,000



data points of LiDAR elevation in that area formed the basis for computing the digital elevation model.

We applied a procedure tailored to smoothing topography in forested environments. The smoothing procedure was necessary to remove points that represented tops of high vegetation and consequently hid the topographic surface below. We applied the approach to extract vegetation implemented by Evans and Hudak (2007) using the multiscale curvature classification (MCC) algorithm. This command-line tool evaluates local-point deviations from a mean surface and interactively erases points above this surface from the ground class. The two parameters required by MCC are nominal-point spacing obtained from the LAS file and curvature tolerance, the maximum height allowed above the mean surface for a point to be classified as a ground point. Valid values of curvature tolerance range from 0 to 1. MCC algorithm used raw elevation points from LiDAR. The resulting bare-earth surface data points were all then interpolated to a 1-m resolution digital elevation model (DEM) in ArcGIS 10 Environment.

The methodology proposed in this study starts with identifying areas with high slope gradient which are based on a combination of hydrologic and terrain variables. Due to accentuated slopes, flow water velocity, might be high and these areas are likely to place linear erosional features such as gullies, ravines and cliffs. Next compute 1-m bare-earth DEM and then calculate the following four features to apply in the data clustering process:

- **Stream power index (SPI):** Predicts the erosive power of overland flow leading to potential erosion through soil scouring (Wilson and Gallant, 2000; Wischmeier and Smith, 1978). SPI is defined as  $\ln(A_s/\beta)$  where  $A_s$  is the local upslope area draining through a certain point per unit length and  $\beta$  is the local slope. Kakembo et al. (2009)

experiment found high values of SPI in the bottomland of gully systems with low slopes. Conversely, the stream power index exhibits low values in upland convex hillslope areas.

- **Mean slope deviation:** Defined as the difference between individual slopes from mean global slope computed at all points within study area. Primarily, slope has great significance for hydrology and geomorphology because it affects velocity of surface and subsurface water flow Kakembo et al. (2009) reported positive correlation between slope gradient and gully depth in an African catchment. Additionally, Santisteban et al. (2005) have found positive relationship in a linear regression analysis between topographic indices and eroded sediments in ephemeral gullies from four small to medium-size watersheds in Spain. The goal of the mean deviation, a measure of dispersion, is indicate how different the local slopes are from their regional mean. In this way we captured locations with abrupt changes in slope.
- **Profile curvature ( $m^{-1}$ ):** Corresponds to the second derivative of slope and it is parallel to the direction of the maximum slope. Negative values characterize surfaces upwardly convex; therefore, decrease flow velocity in these areas. Profile curvature has been reported as significant in controlling flow acceleration and sediment transport ( Istanbuluoglu et al., 2008; Wilson and Gallant, 2000).
- **Hillslope aspect:** A  $\psi$  or slope aspect that corresponds to the direction of steepest descent measured in degrees clockwise from north ranging from  $0^{\circ}$  to  $360^{\circ}$  (Hengl and Reuter, 2009). Additionally, slope aspect represents different degrees of insolation resulting from sunny versus shaded slopes. Numerous empirical studies have examined the influence of slope aspect on soil erosion (Besler, 1987;

Istanbulluoglu et al., 2008; Weaver, 1991). They concluded that north-facing slopes tend to be more eroded than south-facing slopes, primarily due to differences in soil moisture fluctuations. Likewise, Churchill (1982) reported differences in landslide or failure patterns on south-facing versus north-facing slopes. This study in the white river badlands of South Dakota reported that north-facing slopes experienced highly intensive failures or landslides compared to south-facing slopes, where high-frequency failures occurred but carried smaller amounts of sediment. Causes for this aspect-induced difference are related to variations in antecedent soil moisture that leads to soil disaggregation. However, several researchers found opposite patterns. Istanbulluoglu et al. (2008) found south-facing slopes in asymmetric valleys in central New Mexico were more prone to erosion compared to north-facing slopes, even though the former were steeper. Patton and Schümm (1975) described greater stability of steeper slopes facing north because of denser vegetation coverage compared to those facing south. In this study, the aspect values for each point were coded generalizing the main cardinal directions as follows:

- $315^{\circ}$  -  $360^{\circ}$  and  $0^{\circ}$  to  $45^{\circ}$  is equal to 1 (north-faces)
- $46^{\circ}$  to  $135^{\circ}$  is equal to 2 (east-faces)
- $136^{\circ}$  to  $225^{\circ}$  is equal to 3 (south-faces)
- $226^{\circ}$  to  $314^{\circ}$  is equal to 4 (west-faces)

All four variables were created from a raster-smoothed DEM; the value of pixel points in the remaining four variables map were then transformed to point values to be processed in the next step. As the outcome of this process, we produced a dataset textual file

of 913,494 observations with SPI, mean slope deviation, profile curvature, and hillslope aspect point values along with geographical coordinates.

### **2.3.2 Classification methods**

We decided to test unsupervised classification methods to extract gully features because gullies are distinct features in geometric and hydrologic nature compared to the surround landscape. Unsupervised classification can be useful when a training area is not available which means that there is no previous field survey at this area. Generally, in data mining using unsupervised methods no target object is previously introduced; instead there is a search for a pattern or intrinsic structure among parameters or objects forming groups or clusters.

Following a previous preliminary study (Vendrusculo and Kaleita, 2013) that compared diverse types of unsupervised classifiers, the CLARA algorithm, described above, was selected as the classification approach. In our prior study, this algorithm was superior compared to centroid approaches like K-means and Fuzzy K-means in our previous study. However, because of clustering algorithm complexity, enormous computer memory and processing times are required. CLARA algorithm uses the sampling approach to alleviate these problems. The classification of the gullied areas was performed using “Cluster” package available at R Statistic package (R Development Team 2011). A script written in the R programming language was used for this purpose. Since CLARA, like many other clustering algorithms, begins with assigning random initial values in the first run to its centroid points, a fixed seed ( $s=110$ ) was used to permit reproducible results of the CLARA clustering.

Since the CLARA algorithm requires the user to provide the number of clusters, we used the minimization of within-group sum of squared error (WSS) to determine how many

groups were statistically distinct from one another (Mortazavi and Jalili, 2014). WSS is one similarity measure that evaluates how well a clustering algorithm may perform. The algorithm calculates the distance from each classified observation to its representative cluster point. Mainly, as the number of cluster increases, the WSS decrease, because clusters become smaller and better defined. Thus, an adequate cluster number can be selected from the point where error decreases drastically. Finally, a multivariate analysis of variance (MANOVA) using the R package HSAUR was conducted to determine if the cluster were statistically different, thus the null hypothesis was rejected. The null hypothesis in this context asserts that the means of clusters were similar

After the CLARA classification process, the steps outlined in the following sections were used to identify data groups under the gully threshold criteria.

### **2.3.3 Threshold model to identify concentrated flow path within a gully channel bed**

Since the computed flow accumulation given by ArcGis produces the weight of all cells flowing into each downslope cell for the study area, we proposed a general threshold model to identify gully concentrated flow paths (GCF) present only in the channel gully bed. Thus, we identified the flow accumulation that is likely to occur in the gully channel. To do this, the proposed threshold criteria selecting higher values of flow accumulation in places where the slope is also greater than the mean area slope. The mean slope computed to the study area is  $6.14^{\circ}$ . This approach combines statistical data distribution with topographic and hydrologic gully characteristics.

Therefore, in order to select the range of high flow accumulation values, we sorted and divided the values of flow-concentrated paths into 10 parts according to a Natural Jenks classification. Thus, each break in the flow accumulation values represented 10% of the

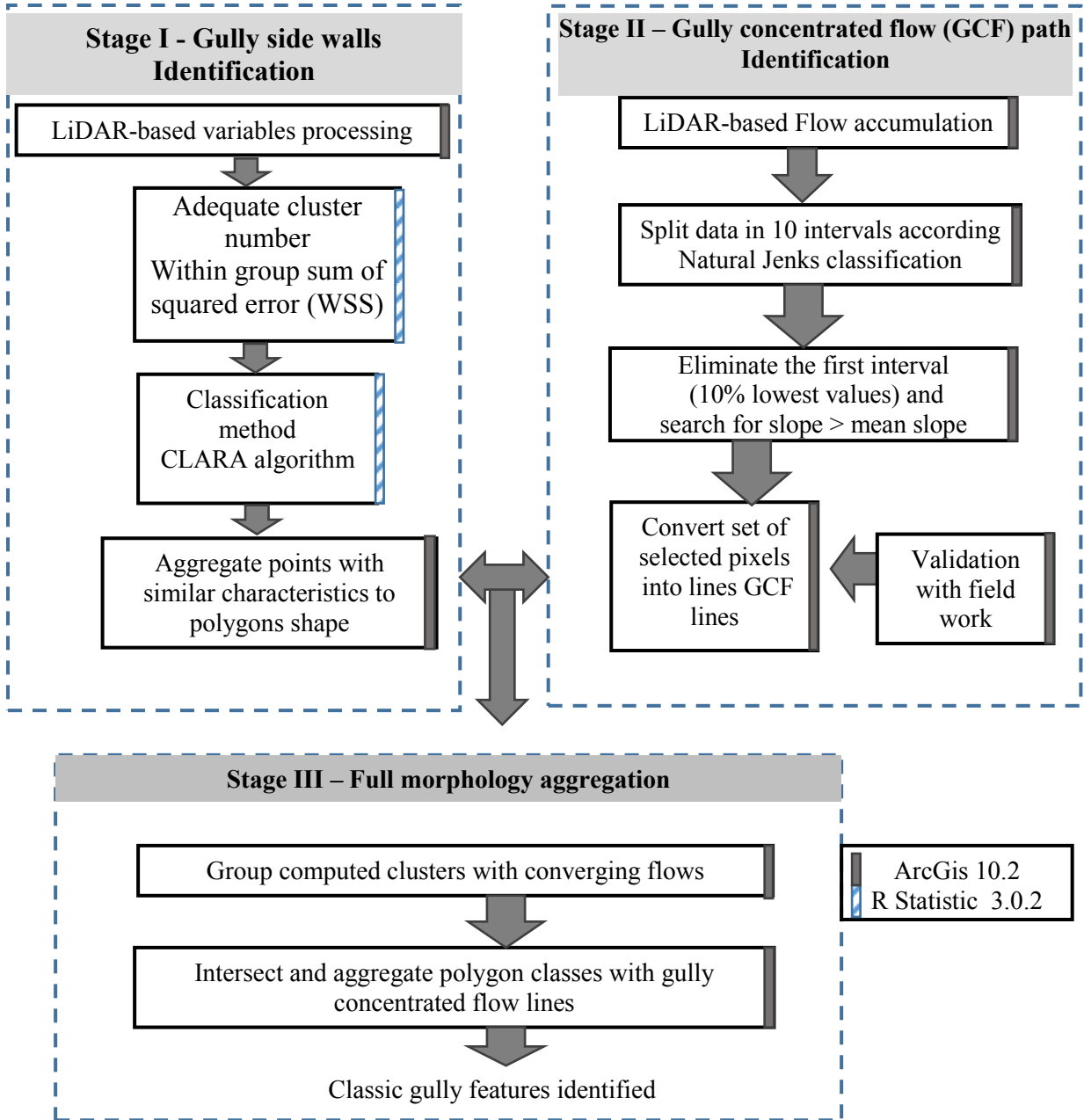
population. Then, using trial-error approach, we removed the 10% lowest values of flow accumulation, after that 20% of the lowest values and so on. For each tentative, we plotted in ArcGis the remaining values to observe if the flow paths, formed by flow accumulation cells, had spatial proximity with the surveyed gully beds. The trial-error approach revealed that the lowest 10% (first group) of flow accumulation values removal (10 % lowest values of flow accumulation  $< 5500 \text{ m}^2$ ) associated to the local slope greater than mean slope was the nearest and obtained approximately similar length comparing with the surveyed gullies beds. A few spurious pixels were manually eliminated. Although 10 was selected as a partition number, a conservative approach is still an educated guess. Scenarios using partitions  $> 10$  and consequent removal of the lowest 10% of values produced an extended concentrated flow patch.

The proposed methodology described above to gully identification is summarized in the framework illustrated in Fig. 2.4.

#### **2.3.4 Validation methods**

Initially, we tested whether the flow path defined by LiDAR DEM was consistent with the survey procedure, e.g., GPS gullies pathways, since the SSURGO polylines seemed somewhat broad to be used as a basis of concentrated flow. Through visual analysis we can infer that SSURGO polylines follow the spatial distribution of soil units and consequently might be shifted from the real gully bottom if some soil classification error occurred. For this reason, SSURGO were used only for reference purpose.

For this reason, the validation purpose only compared features surveyed during field work.



**Fig. 2.4** General framework including three stages of the gully morphology with include the identification of side walls cluster from CLARA identification and gully concentrated flow (GCF) polylines finalizing with the spatial intersection of these gully elements.

## 2.4 Results and Discussion

### 2.4.1 Comparing field survey and gully concentrated flow paths

Fieldwork has identified seven classic gullies and threshold GCF procedure computed six main paths with concentrated flow in the same surveyed area. The longest edge and channel length were 1981 m and 608.3 m (#1), respectively. The shortest length was 32 meters (#3) which was not identified by the GCF threshold procedure. The mean length of seven gullies was mean = 327.8 m with standard deviation of 269.75mm, as shown in Table 2.1. An *in situ* investigation supported by a 3D visualization using LiDAR DEM confirmed the composite of a trapezoidal shape for the largest gully and a V-shape for the remaining gullies.

**Table 2.1** Statistics of surveyed polylines and GCF in the study area

Gully Identification	GPS Survey length (m)	GCF polylines length (m)	Difference <sup>¥</sup> (m)	Shape
1	1,981 *, 608.3 **	962.67***	-	Trapezoidal
2	203.78 *	92.34	111.44	V
3	32 *	0	32	V
4	145.09 *	45.9	99.2	V
5	785.26 *	527.67	257.59	V
6	262.02 *	160	102	V
7	258.57 *	171.12	87.45	V
Total	4,275.92	1,959.7	2,315.3	

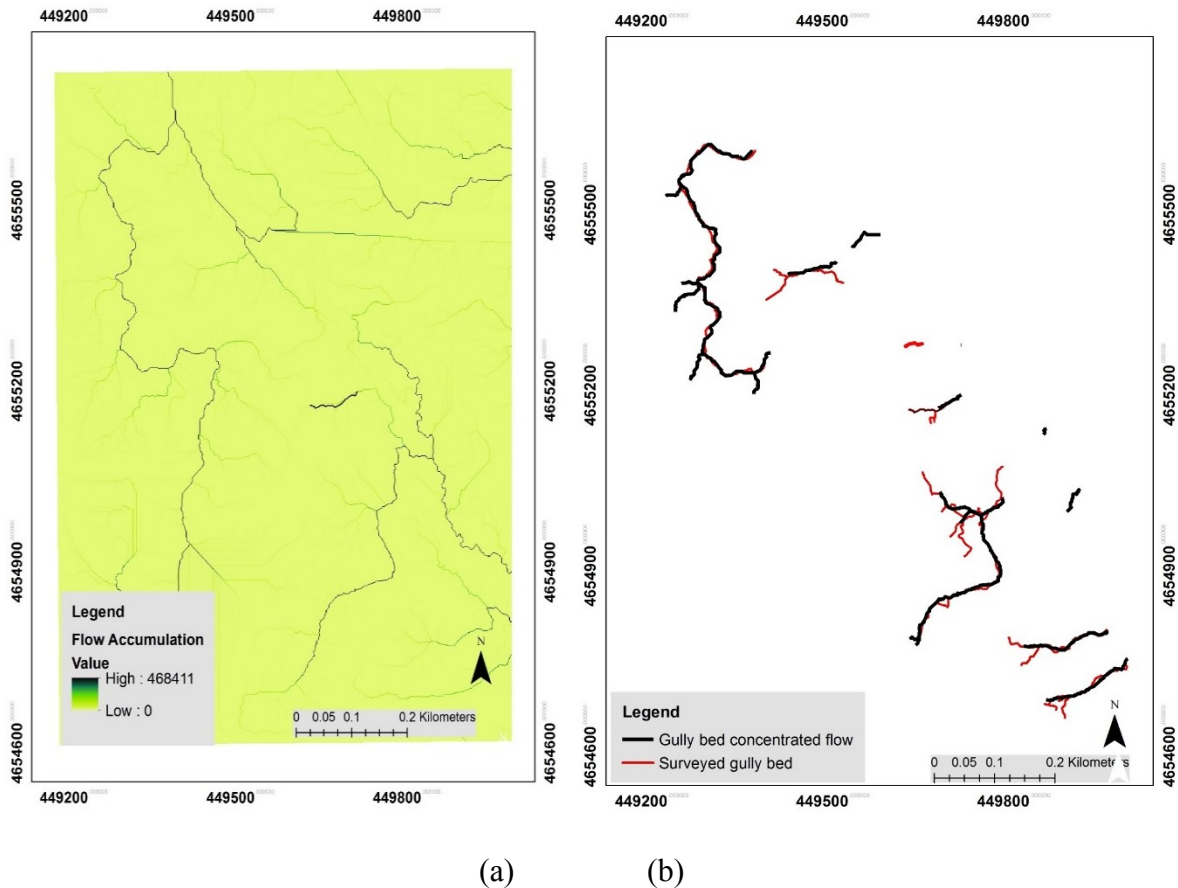
\*Top edge measured separately \*\* Bottom measured separately \*\*\* edge and bottom computed

¥ Positive values represents increase regarding values obtained from survey.



From visual analysis of Fig. 2.5, it is apparent that pathways of gully bottoms exhibit a meandering channel form. These landscape units have a more intricate pattern than computed water flow lines, suggesting that the terrain had changed between the time of the LiDAR flight and the 2014 survey.

Nevertheless, the computed polylines were generally in agreement with the surveyed locations of the bottoms of the V-shaped gullies based in visual analysis. Thus, concentrate flow path (GCF) can be used as a bottom in V-shaped gullies for this area.

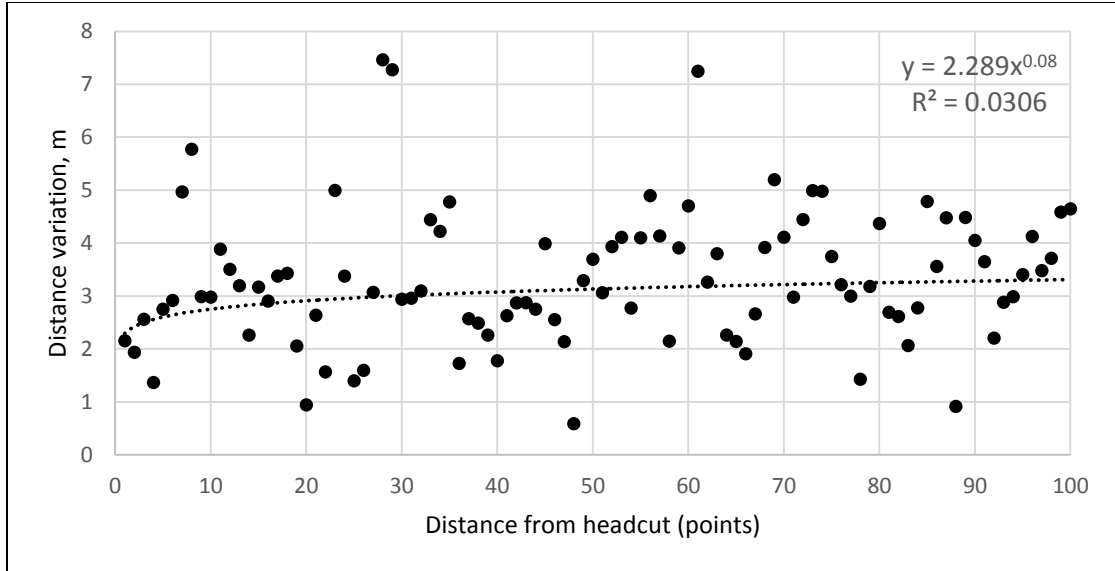


**Fig. 2.5** Map with original flow accumulation (a) and gully concentrated flow polylines computed (black lines) by the threshold procedure compared with the fieldwork paths (red lines).

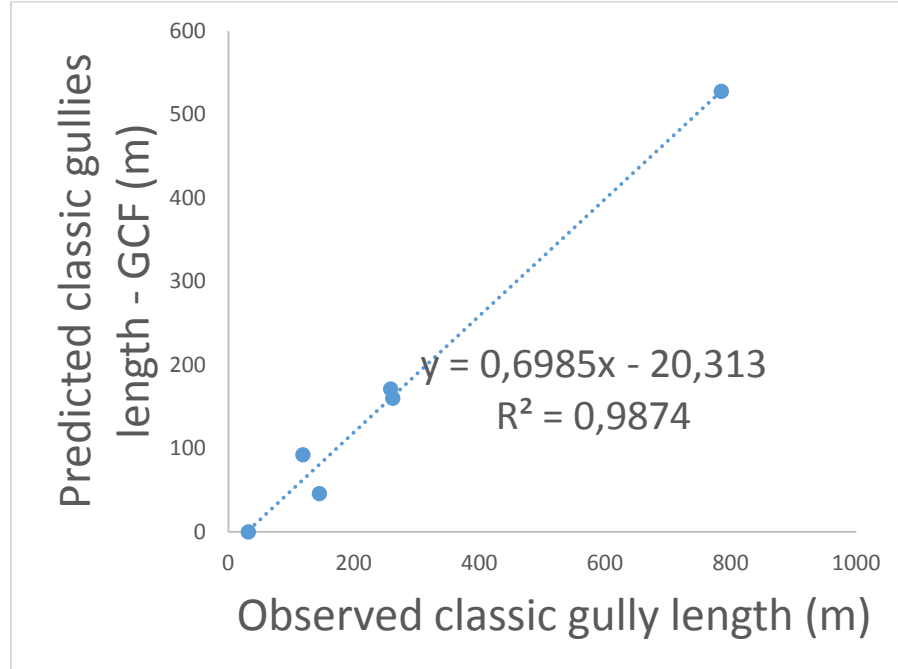
Figure 2.5 (a) shows the original flow accumulation map produced by ArcGis from flow direction and elevation maps. After execution of GCF procedure at the GIS, the resulting polylines are depicted in Fig. 2.5 (b) which reveal, when compared with field work, a much simpler bottom line of gullies existent in that area.

The analysis of distance between the survey channel bed and concentrated flow path lines in the largest (#1) gully is depicted in Fig. 2.6. Using the GIS environment, the largest distances were identified at points that seemed to rapidly change in flow direction, identified during field work. Among seven gullies, only gully line #1 had a headcut completely crossed by the computed polylines. For the remainder of the gully headcuts, the distances are in the range of 23 to 122 m downhill from the flow accumulation starting point. Furthermore, a much larger flow-contributing path than the gully length itself, computed from the concentrated flow network, might explain the greater lengths observed in gullies 1 and 5. Although there was a computed weak power function,  $R^2 = 0.03$  (Fig. 2.7), it might suggest dynamic change in the local landscape. The difference might also be explained due to data acquisition errors, including those of both GPS devices and LiDAR.

An additional comparison among the total length of each gully from the field work survey and the computed GCF paths are shown in Figure 2.6. Despite of the great total length difference between GCF (2,315 m) and fieldwork, R squared is high for the linear regression line that predict length GCF with slope around 0.7.



**Fig. 2.6** Distance variation in the largest gully (#1) at KSSBR study area between bed-path survey and flow-accumulation streamlines exceeding  $5000 \text{ m}^2$  (e.g. high flow). The first point corresponds to the headcut and further points head toward the South Skunk River.



**Fig. 2.7** Linear regression models among six measured classic gully length and predicted gully concentrated flow (GCF) lines.

### 2.4.2 Statistical analysis of terrain attributes

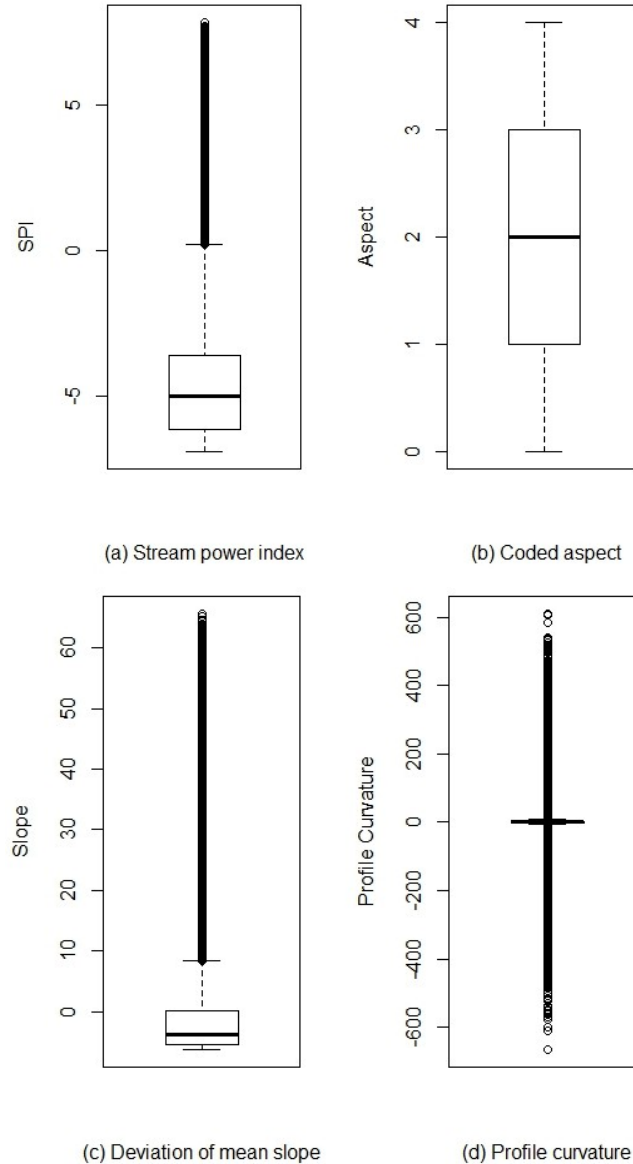
General statistics of terrain and hydrologic attributes are displayed in Table 2.2. Over the 913,494 points, the elevation ranged between 270 m in the Northeast portion and 300 m in the Southwest portion of the study area. SPI ranged between -7 and 7.8 with negative or lower values near gully banks (Fig. 2.8 a). Data distribution, through histogram analysis, exhibits a left-skewed histogram in mean slope deviation characterizing median slope gradient less than mean global slope ( $6.14^\circ$ ). The mean slope deviation also has high data variability from the average (S.D:9.6). Aspect after coding process presented a mean equal to 2 with small standard deviation (S.D.: 1.22). Profile curvature showed a highly concentrated range of points.

**Table 2.2** General statistics for study area considering terrain attributes of elevation above sea level (Z); stream power index (SPI); mean slope deviation, aspect and profile curvature ( $Curv_{pr}$ )

Statistic	Z (m)	SPI m radian	Mean slope deviation Degree	Aspect coded dimensionless	$Curv_{pr}$ $m\ m^2$
Mean	283	-4.6	0.00	2.13	0.9
SD	8.9	1.6	9.64	1.22	21.7
Min	270.5	-7	-6.15	0	-664
Median		-5	-3.83	2	0
Max	300.2	7.8	65.56	4	609.5

The median SPI is close to the average mean considering the whole study area despite of the many upper outliers (values superior to 75<sup>th</sup> quartile). The SPI long upper whisker means that values varied amongst the most positive quartile group. The data are left skewed, meaning that area have SPI values occurring inferior to median (Fig. 2.8a). The coded aspect boxplot demonstrated that there is no preference for a specific slope face in this region (Fig. 2.8b). Zero average of mean slope deviation representing the null difference between global

and local slope was found, despite of the numerous outliers in the 3<sup>rd</sup> percentile of boxplot (Fig. 2.8c). Deviation of slope mean and profile curvature boxplot shows majority of data concentrated around mean and median values. Profile curvature is the one that holds the greatest variability among the variables from the average (Fig. 2.8d).



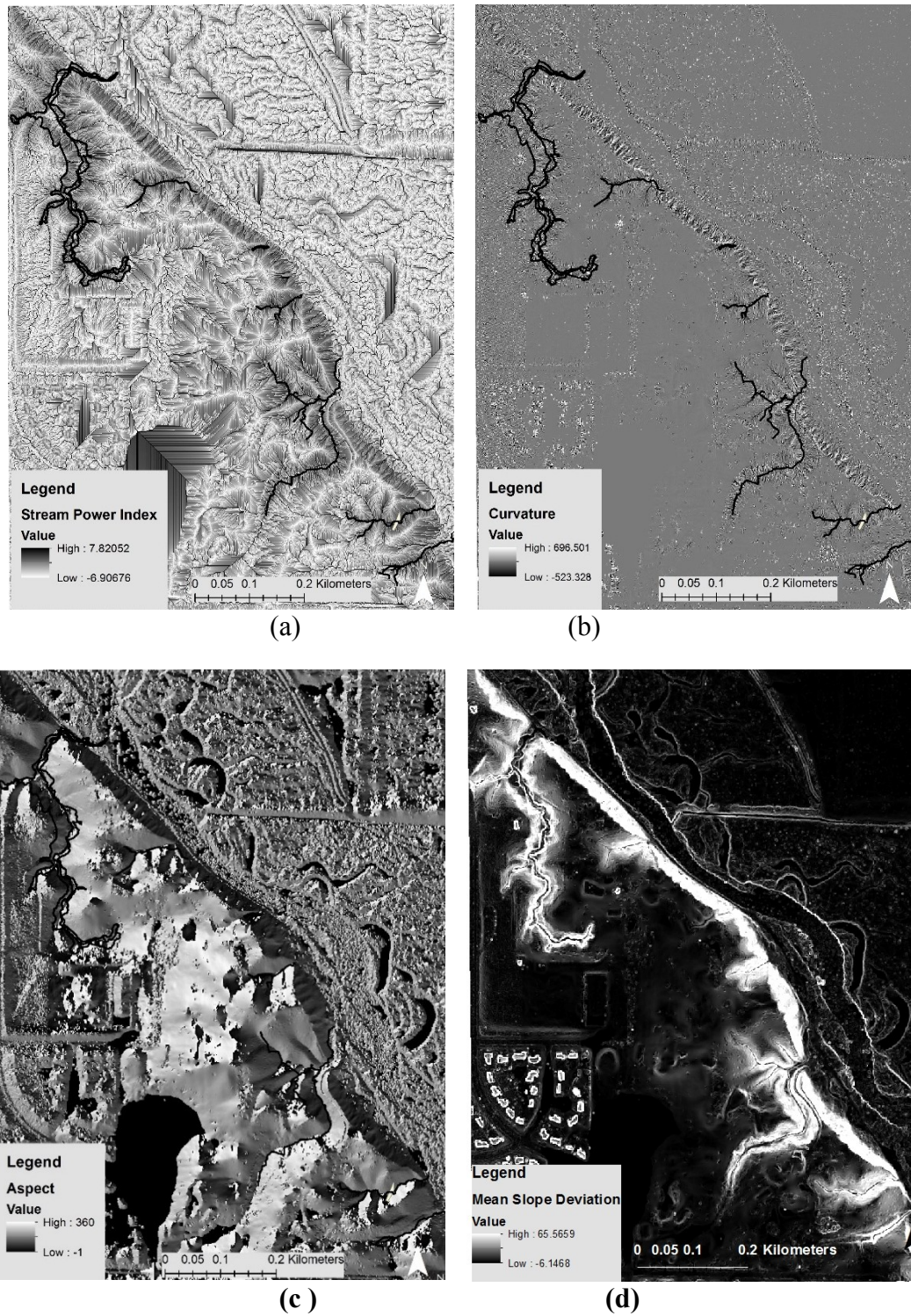
**Fig. 2.8** Descriptive statistics of four variables used in the clustering considering the entire study area.

From a visual interpretation, the stream power index map reveals the high runoff contribution located in the gully bed originating from uphill land (Fig. 2.9 a). Highly-negative values of profile curvature represent concave surfaces and are distinguishable near the floor areas of gully features (Fig. 2.9 b), which are in agreement with the Kakembo et al. (2009) study. Conversely, positive curvature values or convex shapes represent surrounding concave surfaces forming a visible dendritic pattern in the study area (Fig 2.9 c). High variations in mean slope deviation gradient, similar to slope attribute, occurs near the survey flow path, in good agreement with both the fieldwork and digital elevation model (Fig 2.9 d). In general, North and East slope faces in gully banks are dominant.

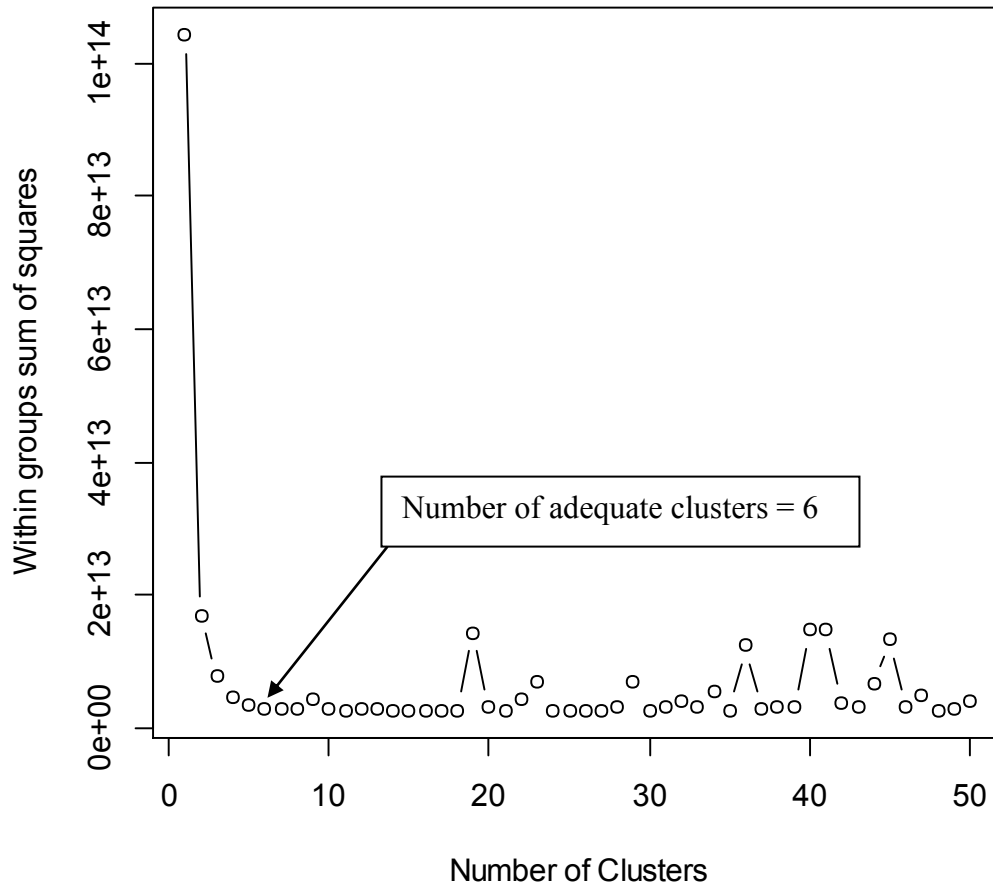
### **2.4.3 Classification accuracy**

An analysis of the WSS curve reveals that a substantially decreased error is reached around six clusters (Fig. 2.10) in this study area. Therefore, we set the cluster number as six during unsupervised CLARA clustering procedure using terrain and hydrologic parameters. The variables were computed at point level from airborne LiDAR data. A few spikes can be noted in the WSS graphic (e.g., 19, 29, 36 groups) which suggest the increasing partitioning is sensitive to the spatial nature of the data.

Large variations in mean slope, such as at gully edges or beds, were shown in cluster boxplots. Pixels clustered at cluster 4 are found on steep slope edges. There is a gradual slope gradient among cluster 4, 5 and 6, as can be observed in Table 2.3, where the points belonging to the later cluster cover the majority of gullied areas. According to the coded aspect boxplot (Fig. 2.11), the data distribution is very similar among clusters with the same median values which represent that all clusters have representatives of slopes facing the four directions (North, South, East, and West. A larger data variation is observed in profile



**Fig. 2.9** Maps of (a) Stream power index, (b) profile curvature (c) aspect (d) mean slope deviation at the study area from 1-m LiDAR data. The black lines in the maps represent the survey paths from field work.



**Fig. 2.10** Sum of squares error curve between representative centroid point and observation labeled in each cluster. According elbow technique the lowest adequate number of clusters is equal to 6.

curvature at clusters 4 and 5, although the data distribution shown in Fig. 2.11 is masked by the presence of high and low values (outliers). Clusters 2 and 5 have only positive values, and the largest standard deviation is exhibited by cluster 4 whereas the lowest is cluster 1. However, clusters 1, 2, and 6 include both convex (profile curv  $< 0$ ) and concave (profile curv  $> 0$ ) shapes that might lead to erosion and soil redistribution (Table 2.3).



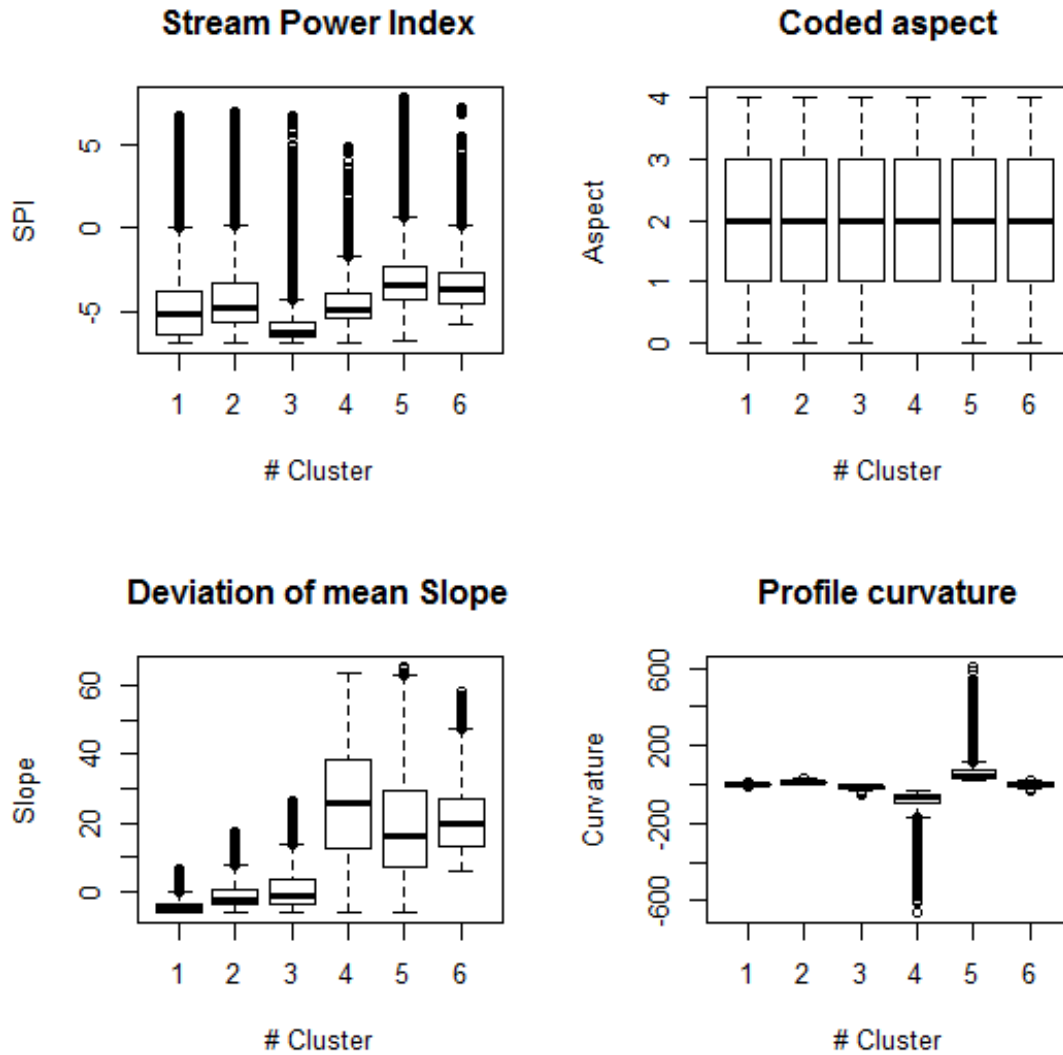
**Table 2.3** Summary statistics of topographic and hydrologic information by classes in the study area.

Statistic	1 <i>n</i> = 57,9651	2 <i>n</i> =127,726	3 <i>n</i> = 87,548	4 <i>n</i> =11,219	5 <i>n</i> =24,070	6 <i>n</i> =83,280
Min	Curv:-7.4 S:-6.14 Asp: 0 SPI: -7	Curv: 3.04 S:-6.12 Asp: 0 SPI:-7	Curv: -49.5 S:-6.13 Asp: 0 SPI: -6.9	Curv: -663.8 S: -6 Asp: 1 SPI: -6.9	Curv: 24.8 S:-6 Asp:0 SPI:-6.7	Curv: -36.6 S:6 Asp:0 SPI: -5.74
Mean	Curv:0.08 S: -4.18 Asp: 2 SPI:-4.8	Curv: 12.2 S: -0.85 Asp: 2.4 SPI: -4.21	Curv: -15 S:0.7 Asp:2.37 SPI:-5.8	Curv:-92.5 S:26 Asp: 2.37 SPI: -4.6	Curv: 67.75 S:19.4 Asp:2.4 SPI: -3	Curv: -1 S: 20.6 Asp: 2.11 SPI: -3.4
Max	Curv: 5.7 S:6.67 Asp: 4 SPI: 6.66	Curv: 35 S: 17.6 Asp:4 SPI: 6.9	Curv: -4.2 S: 26.2 Asp:4 SPI: 6.7	Curv:-35 S: 63 Asp:4 SPI: 4.83	Curv: 609.5 S: 65.6 Asp:4 SPI: 7.8	Curv:26 S:58.3 Asp:4 SPI:7.1
SD	Curv:1.82 S:2.27 Asp:1.3 SPI:1.9	Curv: 6.33 S:4 Asp1: SPI:2	Curv:8.6 S:5.4 Asp:1 SPI:1.22	Curv:75.3 S:16.3 Asp:1 SPI:1.27	Curv:63.5 S:14.8 Asp: 1.1 SPI: 2.1	Curv:11 S:9 Asp1: SPI:1.4

*n* represents the number of observations in each cluster. Curv: Curvature, S: Mean slope deviation, Asp: Aspect, SPI: Stream power index\

A MANOVA statistical analysis was employed on the classified original data to test whether the groups in the study area should be considered to be different. Results from the MANOVA showed that population means for each cluster differed significantly (Wilks  $F=0.33$ ,  $p < 0.0001$ ); thus, hypothesis  $H_0$  was rejected. The six groups found in the clustering process were not similar.

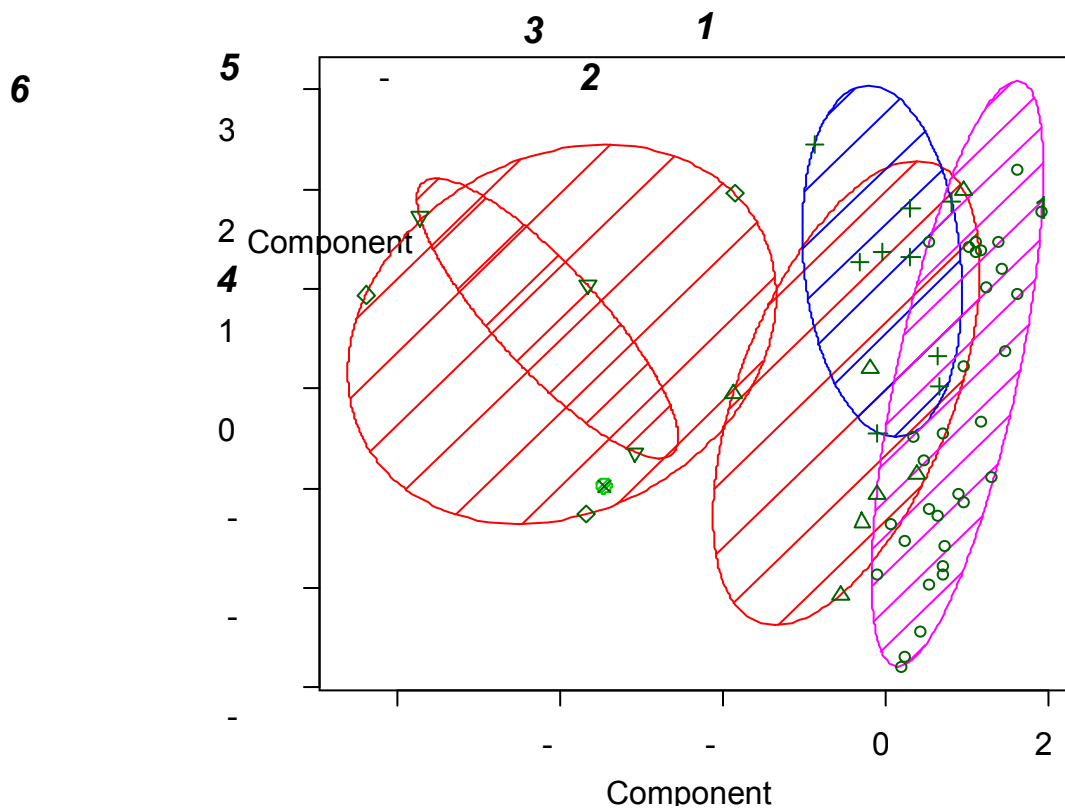
Another approach to visualizing the partitioning results is depicted in Fig. 2.12, where data are shown in a bivariate plot of principal components. The five attributes are reduced to the two principal components which, in this case, would explain 56.2 % of the point variability. The points, as shown in Fig.2.12, represent the best examples of the clustering process. There are a few combinations that are clearly differentiated, such as 1 and 5 or 2 and 3; however, there is an overlap between clusters 6 and 5.



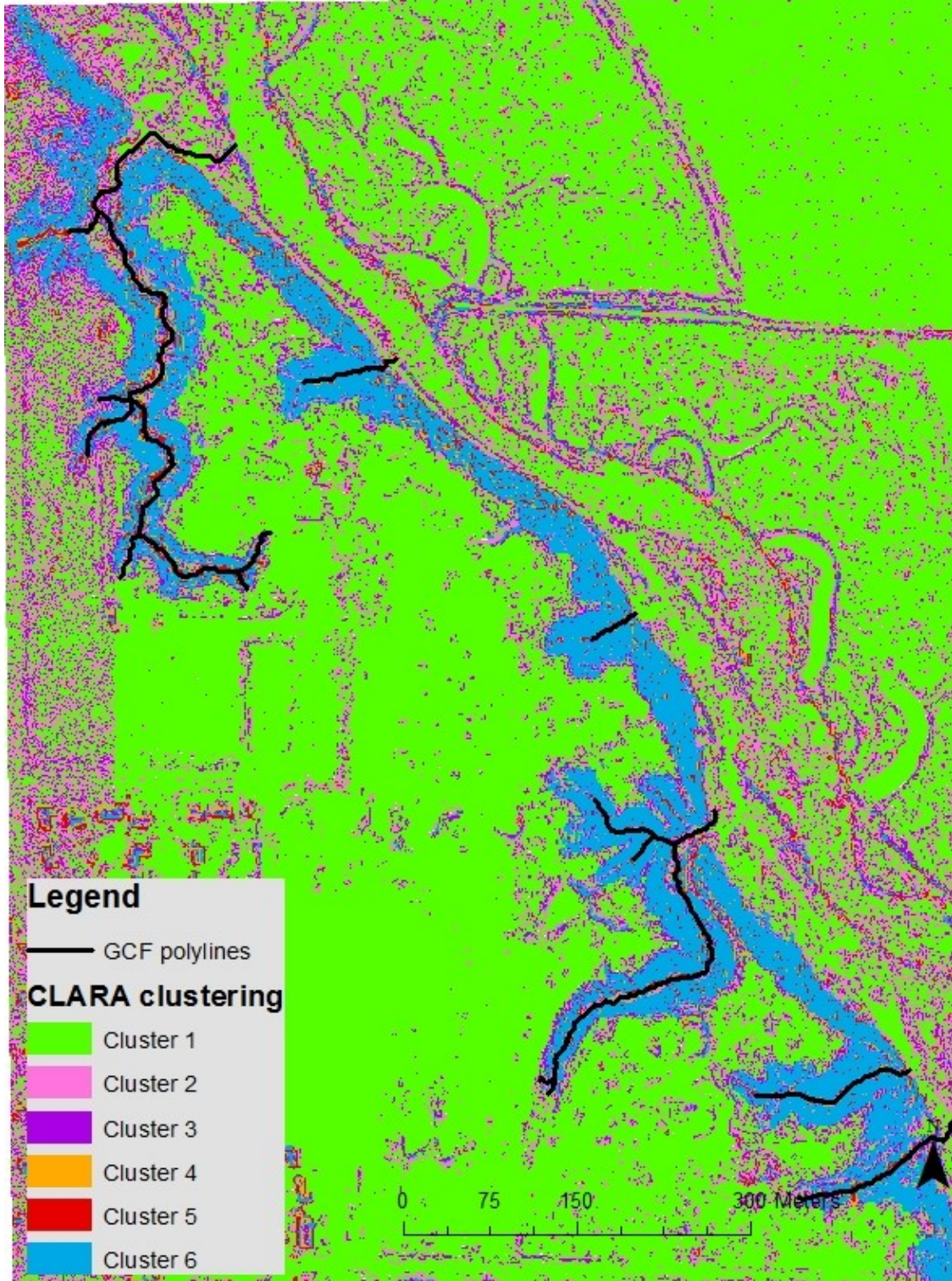
**Fig. 2.11** Box plots of attributes used in CLARA unsupervised clustering distributed by cluster. The boxplot shows the data extremes (whiskers and the points for extreme cases), lower and upper quartiles, and the median.

### 2.4.4 Gully classification

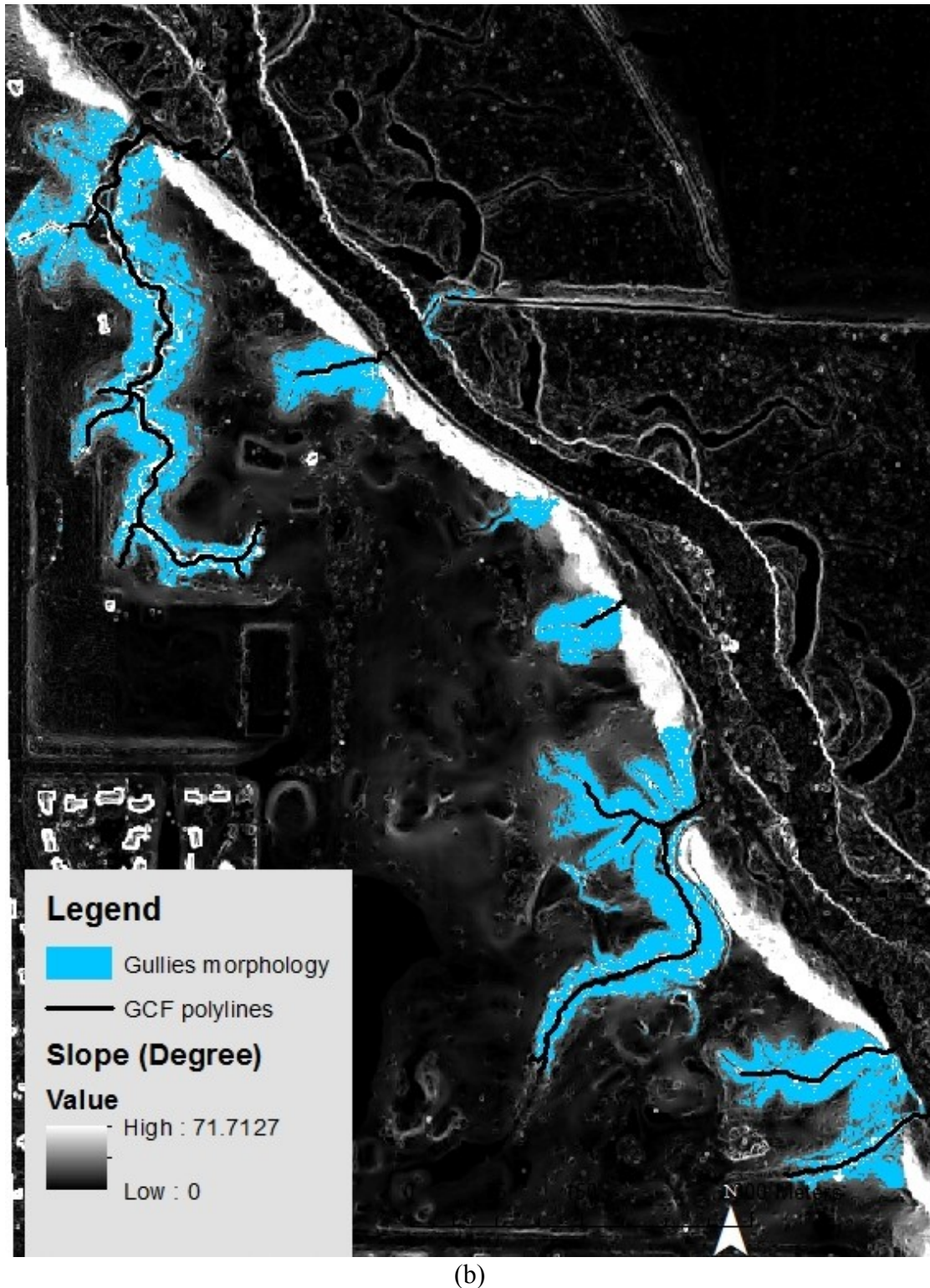
The final classification resulting from the CLARA clustering in the study area under gully erosion is shown in Figure 2.13a. The gully walls cluster was selected as the area surrounding GCF lines and also the maximum first quartile (25%) of mean slope deviance. The cluster that met this requirements was cluster 6. Additionally, points classified as cluster 6 formed the larger area compared with points of cluster 4 and 5 that were placed in the gullies and cliff edges. The approach to compute GCF lines effectively removed areas in urban portion of study area. Initially, due the roofs elevation differences, those points were also classified as cluster 6; however, after executing GCF threshold, those roof zones were excluded as potential gullied areas.



**Fig. 2.12** Bivariate plot of principal components according the eight groups classified by CLARA method.



(a)



(b)

**Fig. 2.13** (a) Final study area classified with six groups; (b) refined areas formed by points in cluster 6 and classified as gully side walls (blue areas). The bold lines represent the GCF polylines obtained by the general threshold.

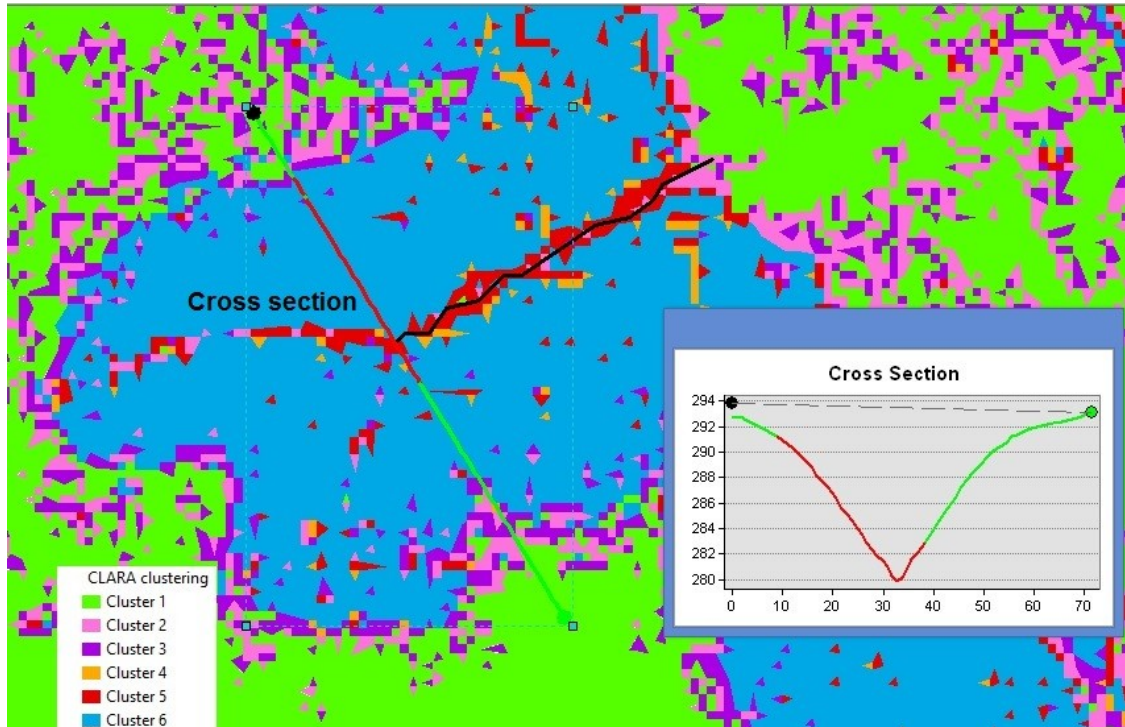
Nevertheless, there were still two types of linear features in the area with points classified as cluster 6. There were gullies and two cliffs classified as special features in SSURGO database. Thus, to separate the two different features, flow direction was used to distinguish flow that contributes to the gullies and also reach the cliffs, as illustrated in Fig. 2.13b. Therefore, the methodology presented is suitable to identifying the whole gully morphology including side walls and bed areas. The majority of the literature referencing classical gully location map this feature as a whole feature. This second contribution is important for design of precision conservation structures that enable engineers to rank the priorities in the bed or side walls areas according effectivity and financial resources available.

#### **2.4.5 Precision conservation application**

The use of precision conservation structures in the area are due to outcomes obtained by the proposed methodology which include location of critical hillslopes that are potential areas of shallow landslides and erosion due particle detachment. With exception of winter, the surrounding area of gullies is located under dense forest. This condition might avoid a solution based in filter or buffer strips around the gullies edges because of issues due to vegetation establishment resulting from a restricted amount of sunlight.

After identification of gully morphology, it can be noted visually that points located in cluster 3 are placed in the gullies' edges; in addition, this cluster obtained the lowest SPI, which implies high slopes. Thus, slope stabilization with geotextile or natural fiber blankets is a possible solution in this area. Goddard (2005) also emphasized this landscape position approach to precision conservation purposes. Alternatively, it is possible to use existing trees and shrub debris at locations parallel to contour lines as suggested by Swift et al. (1993). An

example, as depicted in Fig. 2.14, shows where the areas of cluster 3 are placed and their dimensions. This information can reduce time and optimize resources for conservation structures also providing crucial information for possible scenarios of soil erosion controls placement.



**Fig. 2.14** Clustered area surrounding Gully #3 with a cross section of its open channel. The gully edge has points classified as 3 and can be chosen to soil control structures placement.

## 2.5 Conclusions

This study proposed a semi-automated gully identification procedure based on data mining of hydraulic and topographic combined to gully bed concentrated flow path derived from high-resolution elevation data (LiDAR). The contribution this study provides is a novel combination of four variables: coded aspect, mean slope deviation, stream power index and

profile curvature applied at centroid-clustering approach, implemented by CLARA.

However, this non-time consuming technique is computer memory demanding, but can be used at a location of high delivery source of sediments such as gullies. In addition to the relatively complex targets in the study area (e.g., residences roofs, cliffs, roads, etc.) this methodology identified all six surveyed gullies. The only drawback of this technique is that it was not able to identify the bed zone of gully #3, and it also required an additional step to partitioning of cluster 6 in order to distinguish gullies and cliffs. Additional benefits from the proposed methodology are the effective identification of the gully as an open channel comprised by side walls and a bed path. Under a precision conservation approach, there is a need to consider gully location as more than the entire feature as currently cited by several studies ( James, Watson, and Hansen 2007; Mararakanye and Nethengwe, 2012, Momm et al., 2013; Pike et al. 2012). Therefore, as expected, areas with high slope gradient located primarily in the border, can be identified as targets to conservation practices like geotextiles or natural fiber blankets (#cluster 3).

It is, indeed, critical to perform vulnerable areas mapping to broader audiences such as conservationists, engineers, researchers, and especially farmers to help mitigate damage resulting from gully erosion, a fast-changing landscape processes. Identification of degraded areas also helps enhance soil and water conservation structures placement and management practices.

High-resolution elevation measurements, similar to LiDAR, can significantly change the way that landscape features are classified nowadays. Elements not identified in the past because of the abstraction due to lower resolution (e.g., USGS 30 m contour lines), can now be more intensive, related in terms of hydrological and topographical similarities with such



features as in the real world. The proposed methodology may be useful for other applications such as landscape analysis. Because the number of classes used in the data-mining process provides a range of spatial scales, many classes may capture microsites' features.

Conversely, a smaller number of classes might be used to emulate a watershed or even an entire landscape.

Hyperspectral data are also newly available through the LiDAR spectrometer, so evaluation of sensitive models is required to discover various features related to soil erosion. Monitoring gullied areas also requires a time series of LiDAR data. One future solution tailored to small or medium-sized areas might be to obtain LiDAR series using unmanned aerial vehicles (UAV) instrumented with thermal camera or LiDAR devices.

Soil cover is highly correlated with topsoil resistance, especially under highly-concentrated flow conditions that lead to runoff and soil erosion (Knapen and Poesen, 2009). Therefore, an inexpensive, reliable, and indirect approach to quantify ground cover or vegetation and its variation in space and time will be an object of further investigation as this study continues.

Finally, our conclusions only apply to KBSSR study areas with steep-slope topography. Future research activities could test the robustness of our methodology through similar evaluations in other areas with different soil and topographic characteristics. In addition, one remaining research could be carried out to apply our clustering algorithm to large spatial databases to identify clusters with arbitrary shapes and minimal numbers of input parameters.

## Acknowledgments

The authors are grateful to Ligia Serrano and Edson Adriano Vendrusculo for their assistance during the fieldwork. The first author is thankful for research funding provided by EMBRAPA – Brazilian Agricultural Research Corporation which enabled her to complete her PhD program of study at Iowa State University.

## References

- Baruch, A., and S. Filin. 2011. "Detection of Gullies in Roughly Textured Terrain Using Airborne Laser Scanning Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 66 (September). International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS): 564–78. doi:10.1016/j.isprsjprs.2011.03.001.
- Besler, H. 1987. "Slope Properties, Slope Processes and Soil Erosion Risk in the Tropical Rainfall Forest of Kalimantan Timur (Indonesian Borneo)." *Earth Surface Processes and Landforms* 12: 195–204.
- Chaplot, V.. 2013. "Impact of Terrain Attributes, Parent Material and Soil Types on Gully Erosion." *Geomorphology* 186 (March). Elsevier B.V. 1–11. doi:10.1016/j.geomorph.2012.10.031.
- Churchill, R. R. 1982. "Aspect-Induced Differences in Hillslope Processes." *Earth Surface Processes and Landforms* 7 (August 1980): 171–82.
- Daggupati, P., A. Y. Sheshukov, and K. R. Douglas-Mankin. 2014. "Evaluating Ephemeral Gullies with a Process-Based Topographic Index Model." *Catena* 113 (February). Elsevier B.V. 177–86. doi:10.1016/j.catena.2013.10.005.
- Eustace, A., M. Pringle, and C. Witte. 2009. "Give Me the Dirt: Detection of Gully Extent and Volume Using High-Resolution LiDAR." In *Innovations in Remote Sensing and Photogrammetry*, edited by K. Jones, S., Reinke, 255–69. Berlin Heidelberg: Springer-Verlag.
- Evans, J. S., and A. T. Hudak. 2007. "A Multiscale Curvature Algorithm for Classifying Discrete Return LiDAR in Forested Environments." *IEEE Transactions on Geoscience and Remote Sensing* 45 (4): 1029–38. doi:10.1109/TGRS.2006.890412.
- Galzki, J. C., A. S. Birr, and D. J. Mulla. 2011. "Identifying Critical Agricultural Areas with Three-Meter LiDAR Elevation Data for Precision Conservation." *Journal of Soil and Water Conservation* 66 (6): 423–30. doi:10.2489/jswc.66.6.423.

- Goddard, T. W. 2005. "An Overview of Precision Conservation in Canada." *Journal of Soil and Water Conservation* 60 (6): 456–61.
- Gullo, F., and A. Tagarelli. 2012. "Uncertain Centroid Based Partitional Clustering of Uncertain Data." *Proceedings of the VLDB Endowment* 5 (7): 610–21. doi:10.14778/2180912.2180914.
- Han, J., and M. Kamber. 2006. *Data Mining: Concepts and Techniques*. 2nd ed. Amsterdam ; Boston : Elsevier ; San Francisco: Morgan Kaufmann.
- Hengl, T., and H. I. Reuter. 2009. *Geomorphometry: Concepts, Software, Applications*. Edited by T Hengl and H I Reuter. *Developments in Soil Science*. Vol. 33. Developments in Soil Science. Elsevier. doi:10.1016/S0166-2481(08)00036-6.
- Istanbulluoglu, E., O. Yetemen, E. R. Vivoni, H. A. Gutiérrez-Jurado, and R. L. Bras. 2008. "Eco-Geomorphic Implications of Hillslope Aspect: Inferences from Analysis of Landscape Morphology in Central New Mexico." *Geophysical Research Letters* 35 (14): L14403. doi:10.1029/2008GL034477.
- James, L., D. Watson, and W. Hansen. 2007. "Using LiDAR Data to Map Gullies and Headwater Streams under Forest Canopy: South Carolina, USA." *Catena* 71 (1). Elsevier: 132–44. doi:10.1016/j.catena.2006.10.010.
- Johansen, K., S. Taihei, D. Tindall, S. Phinn, Joint Remote, Biophysical Remote, Sensing Group, et al. 2012. "Object-based Monitoring of Gully Extent and Volume in North Australia using LIDAR Data." In *GEOBIA*, I:168–73. Rio de Janeiro.
- Kakembo, V., W. W. Xanga, and K. Rowntree. 2009. "Topographic Thresholds in Gully Development on the Hillslopes of Communal Areas in Ngqushwa Local Municipality, Eastern Cape, South Africa." *Geomorphology* 110 (3-4). Elsevier B.V. 188–94. doi:10.1016/j.geomorph.2009.04.006.
- Kaufman, L., and P. J. Rousseeuw. 1990. *Finding Groups in Data: An Introduction to Cluster Analysis*. New York, NY: John Wiley & Sons.
- Kirkby, M. J., and L. J. Bracken. 2009. "Gully Processes and Gully Dynamics." *Earth Surface Processes and Landforms* 34: 1841–51. doi:10.1002/esp.
- Knapen, A., and J. Poesen. 2009. "Soil Erosion Resistance Effects on Rill and Gully Initiation Points and Dimensions." *Earth Surface Processes and Landforms* 228 (December 2009): n/a–n/a. doi:10.1002/esp.1911.

- Lucà, F., M. Conforti, and G. Robustelli. 2011. "Comparison of GIS-Based Gully Susceptibility Mapping Using Bivariate and Multivariate Statistics: Northern Calabria, South Italy." *Geomorphology* 134 (November). Elsevier B.V. 297–308. doi:10.1016/j.geomorph.2011.07.006.
- Mararakanye, N., and N. S. Nethengwe. 2012. "Title: Gully Erosion Mapping Using Remote Sensing Techniques in the Capricorn District, Limpopo." *South African Journal of Geomatics* 1 (2): 109–18.
- Momm, H. G., R. L. Bingner, R. R. Wells, J. R. Rigby, and S. M. Dabney. 2013. "Effect of Topographic Characteristics on Compound Topographic Index for Identification of Gully Channel Initiation Locations." *Transactions of the ASABE* 56 (2): 523–37.
- Mortazavi, R., and S. Jalili. 2014. "Fast Data-Oriented Microaggregation Algorithm for Large Numerical Datasets." *Knowledge-Based Systems*, May. Elsevier B.V. doi:10.1016/j.knosys.2014.05.011.
- Patton, P. C., and S. A. Schumm. 1975. "Gully Erosion, Northwestern Colorado : A Threshold Phenomen." *Geology* 3 (2): 88–90.
- Pike, A., T. Mueller, E. Rienzi, S. Neelakantan, B. Mijatovic, T. Karathanasis, and M. Rodrigues. 2012. "Research on Soil Erosion." In , edited by Danilo Godone and Silvia Stanchi., 45–63. InTech Open Access on-line book. (<http://dx.doi.org/10.5772/51526>).
- Poesen, J., J. Nachtergaele, G. Verstraeten, and C. Valentin. 2003. "Gully Erosion and Environmental Change: Importance and Research Needs." *Catena* 50 (2-4): 91–133. doi:10.1016/S0341-8162(02)00143-1.
- Santisteban, L. M. De, J. Casali, J. J. López, J. V. Giráldez, J. Poesen, and J. Nachtergaele. 2005. "Exploring the Role of Topography in Small Channel Erosion." *Earth Surface Processes and Landforms* 30 (5): 591–99. doi:10.1002/esp.1160.
- Shruthi, R. B. V., N. Kerle, and V. Jetten. 2011. "Object-Based Gully Feature Extraction Using High Spatial Resolution Imagery." *Geomorphology* 134 (3-4). Elsevier B.V. 260–68. doi:10.1016/j.geomorph.2011.07.003.
- Shruthi, R. B. V., N. Kerle, V. Jetten, and Alfred Stein. 2014. "Object-Based Gully System Prediction from Medium Resolution Imagery Using Random Forests." *Geomorphology* 216 (July). Elsevier B.V. 283–94. doi:10.1016/j.geomorph.2014.04.006.
- Staff, Soil Survey. 2014. "Natural Resources Conservation Service, United States Department of Agriculture. Web Soil Survey." <http://websoilsurvey.nrcs.usda.gov/>.

- Svoray, T., E. Michailov, A. Cohen, L. Rokah, and A. Sturm. 2012. "Predicting Gully Initiation: Comparing Data Mining Techniques, Analytical Hierarchy Processes and the Topographic Threshold." *Earth Surface Processes and Landforms* 37 (6): 607–19. doi:10.1002/esp.2273.
- Swift, L.W., K. J. Elliott, R. D. Ottmar, and R. E. Vihnanek. 1993. "Site Preparation Burning to Improve Southern Appalachian Pine-Hardwood Stands: Fire Characteristics and Soil Erosion Moisture, and Temperature." *Canadian Journal of Forest Research* 23 (10): 2242–54.
- Team, R. Development Core. 2011. "R: A Language and Environment for Statistical Computing". Vienna, Austria: R Foundation for Statistical Computing. <http://www.r-project.org/>.
- Thomas, J. T., N. R. Iverson, M. R. Burkart, and L. A. Kramer. 2004. "Long-Term Growth of a Valley-Bottom Gully, Western Iowa." *Earth Surface Processes and Landforms* 29 (8): 995–1009. doi:10.1002/esp.1084.
- Valentin, C., J. Poesen, and Yong Li. 2005. "Gully Erosion: Impacts, Factors and Control." *Catena* 63 (2-3): 132–53. doi:10.1016/j.catena.2005.06.001.
- Vandaele, K., J. Poesen, G. Govers, and B. V. Wesemael. 1996. "Geomorphic Threshold Conditions for Ephemeral Gully Incision." *Geomorphology* 16: 161–73.
- Vandekerckhove, L., J. Poesen, D. Oostwoud Wijdenes, and T. de Figueiredo. 1998. "Topographical Thresholds for Ephemeral Gully Initiation in Intensively Cultivated Areas of the Mediterranean." *Catena* 33 (3-4): 271–92. doi:10.1016/S0341-8162(98)00068-X.
- Vandekerckhove, L., J. Poesen, D. Oostwoud Wijdenes, J. Nachtergaele, C. Kosmas, M. J. Roxo, and T. D. E. Figueiredo. 2000. "Thresholds for Gully Initiation and Sedimentation in Mediterranean Europe." *Earth Surface Processes and Landforms* 25: 1201–20.
- Vendrusculo, L. G., and A. L. Kaleita. 2013. "Terrain Analysis and Data Mining Techniques Applied to Location of Classic Gully in a Watershed." *2013 Kansas City, Missouri, July 21 - July 24, 2013*. St. Joseph, MI: American Society of Agricultural and Biological Engineers. doi:10.13031/aim.20131619828.
- Vrieling, A., S. C. Rodrigues, H. Bartholomeus, and G. Sterk. 2007. "Automatic Identification of Erosion Gullies with ASTER Imagery in the Brazilian Cerrados." *International Journal of Remote Sensing* 28 (12): 2723–38. doi:10.1080/01431160600857469.
- Weaver, A. V. B. 1991. "The Distribution of Soil Erosion as a Function The Distribution of Slope Aspect and Parent Material in Ciskei , Africa Southern." *GeoJournal* 23 (1): 29–34.

Wilson, J. P., and J. C. Gallant. 2000. "Digital Terrain Analysis." In *Terrain Analysis and Applications*, 479. John Willey & Sons.

Wischmeier, W. H., and D. D. Smith. 1978. *Predicting Rainfall Erosion Losses — a Guide to Conservation Planning*. Agricultur. Washington DC: US Department of Agriculture.

### CHAPTER 3. EVALUATING GULLY SLOPE STABILITY IN MULTI-TEMPORAL SURVEY BOUNDARIES FOR PRECISION CONSERVATION PURPOSE

A paper to be submitted to the *Journal of Earth Surface Processes and Landforms*  
Laurimar Gonçalves Vendrusculo, Amy Kaleita

#### Abstract

A large amount of sediment that originates from gully side walls and headcut causes concerns related to soil and water quality. However, little attention has been paid to soil mechanics associate to the gully slopes. The goal of this study was evaluate the slope stability analysis approach in two multi-temporal surveys carried out in a western Iowa gullied area considering a precision conservation purpose. The study computed spatial failures (Factor of safety indices) at gully head and gully walls of digital elevation models in surveys conducted in 1999 and 2014. Outcomes of this assessment has shown significance less instability in the actual slopes compared to 1999 survey slopes. Internal friction angle ( $\theta$ ) had the largest effect on slope stability factor ( $S.D._{1999} = 0.18$ ,  $S.D._{2014} = 0.24$ ), according the sensitivity analysis, compared to variations of soil cohesion, failure plane angle and slab thickness. In addition, critical instable slopes within gully, based on units of the slope standard deviation as threshold have produced area of  $61 \text{ m}^2$  and  $396 \text{ m}^2$  considering the threshold of one and two slope standard deviation, respectively. This critical areas were located majority near the headcut and in the side walls. Based on current literature, association of processed material and crop cover with high root density might be an alternative to improve slope instability, but empirical tests are necessary to validate this approach.

### 3.1 Introduction

Slope failures are difficult to predict and causes social and economic losses in addition to environmental degradation (Glenn et al. 2006). One important limiting factor in understand slope retreat, regarding to gullying, is its complex interaction with soil properties and local environment which including topography and climate. Conversely, its damages, without distinction, impact population worldwide. In a global scale, USA account for 15% of the total economic annual property damage due landslides which reaches tens of billions of dollars (ITC, 1996).

In a regional scale, mapping slope stability in areas located in the vicinity of agricultural lands play a major role for management practices. Classical gully is a feature in the landscape that generate disproportioned yield of sediments in the watershed (Galzki et al., 2011), thus it require priority in this modeling due its importance in erosion process. Those features are significant drainage channels susceptible to erosion caused by runoff. Their steep side walls and the channel head are locations that experience considerable morphological changes, typically, after intense rainfall events and freeze-thaw cycles (Bradford and Piest, 1977, 1980). Many are the reasons that influence classic gully widening, deepening, and lately stabilization. The leading causes are related to gully-head retreat, upslope sediment movement, hydraulic shear threshold obtained by overland flow on the rim and on the high steep walls, downward mass wasting of walls, and local process at headcut such as rain splash, plunge tools or even biological activity. Also seepage and tunneling contribute for the process of gully erosion (Oostwoud et al., 2001). Channel head is defined by Montgomery and Dietrich (1988) as the most distant upslope location of a channel steep wall banks. The same authors have found that a threshold controls to channel initiation. This threshold is

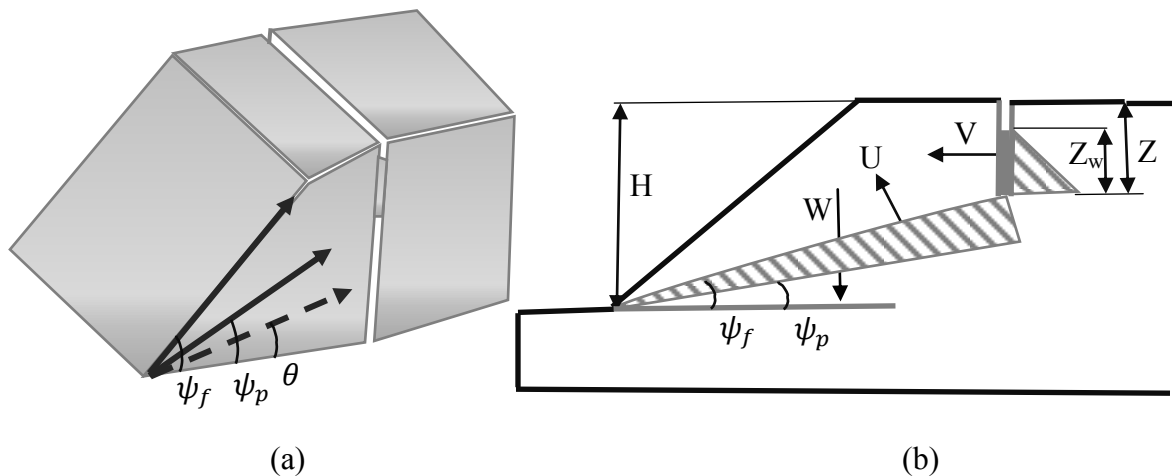


based on the inverse relation between the contributing drainage area and the local valley slope. However, Di Stefano et al. (2000) have pointed out that physical process of sediment transport on hillslopes are not similar to transport within gully network.

Studies have been employed to understand the mechanism of gully-retreat through parameter such as degree slope, slope height, soil density, and strength (Shit and Maiti, 2012; Bradford and Piest, 1980). There are phases associated to this process that might be triggered by high kinetic flow energy, especially, during high intensity rainfall events: (i) failure of gully head and banks (ii) sediment being carry through channels by overland-flow (iii) degradation of the channel which takes to lowering gully bottom surface with debris and sediments deposition. Furthermore, after phase I, has been reported a block of soil overhanging at headcut, solely hold by plant roots, that late will collapse due to desiccation cracks or deterioration of roots strength. Positive pore-water pressure caused by runoff water is considered vital to start mass failure of gully head and stream banks (Simon and Box, 2006). The previous study also has suggested that positive pore-water pressure is associated to macropores infiltration and crack evolution. Other important factor to consider mass fall is due to antecedent moisture conditions, therefore, heavy rainfall precipitation, solely, is not enough to trigger mass movement if you have prior dry conditions (Simon and Box, 2006). Furthermore, climate response and dynamic in land-use also alter the location of the channel head (Montgomery and Dietrich, 1992). As result, study carry out by Molina et al.(2009) found that fluctuations in sediment deposition is explained significantly by vegetation cover in gully bottom at southern Ecuadorian Andes.

Failures in gully slope has been studied under the perspective of plan failure theory (Istanbulluoglu et al., 2005; Bradford and Piest, 1977; Bradford et al., 1973). Chaplot (2013)

also evaluated impact of terrain morphology, parent material and soil types in the gully side wall retreat. Gully slopes are susceptible to resisting forces such as gravitational due to soil mass, weight of water when infiltration and, water table rise and seepage forces of percolating water. Therefore, studies have been early employed to explain instability of slopes in loess regions (Bradford and Piest, 1977; Lohnes and Handy, 1968). This concept is based in soil mechanics which is explained by geometrical conditions to produce rock or gully edge to collapse. The tension crack frequently placed in gully areas are in upper surface as described by Hoek and Bray (1981). This study use the same simplification conditions described by Hoek and Bray (1981) where water is present only along of failure surface and considered the rest of slope mass as impermeable. Thomas et al. (2009) pointed out that the crack depth are difficult to measure, however, their experiment found out the depth close to the water table in a valley-bottom gully at western Iowa. In that way, mathematical equation from slope geometry can simulate conditions where slope are completely drained or when water from heave rain storm event fill in the crack as depicted in Fig. 3.1.



**Fig. 3.1** (a) Geometry of the slope with tension crack placed in upper slope surface and (b) diagram forces showing the relationship of resisting and driving forces to a generic trapezoidal failure surface.

The factor of safety (FS) employ a similar analogy of a block in an inclined lane which to stabilize must to have an equilibrium forces. It is calculated by the coefficient of total force resisting sliding to the total force tending to sliding movement.

$$FS = \frac{c A + (W \cos \psi_p - U - V \sin \psi_p) \tan \theta}{W \sin \psi_p + V \cos \psi_p} \quad (1)$$

Where

$$A = (H - Z) * \operatorname{cosec} \psi_p \quad (2)$$

$$U = \frac{1}{2} \gamma_w Z_w (H - z) \operatorname{cosec} \psi_p \quad (3)$$

$$V = \frac{1}{2} \gamma_w Z_w^2 \quad (4)$$

Where  $z$  is computed in  $H$  and  $x$  terms:

$$Z = H - x * \tan \theta \quad (5)$$

A first approximation of critical head ( $Z_{crit}$ ) need for slab failure, considering  $h$  the height of the topo of the bank above the datum is:

$$Z_{crit} = h - z + Z_w \quad (6)$$

Assuming that the tension crack in located in the upper slope surface

$$W = \frac{1}{2} \gamma H^2 \left( 1 - \left( \frac{z}{H} \right)^2 \right) \cot \psi_p ( \cot \psi_p * \tan \psi_f - 1 ) \quad (7)$$

Where  $c$  is the cohesion of the loess alluvium,  $A$  is the length of failure surface,  $W$  is the total weight of the sliding slab,  $\psi_p$  the angle of the failure surface ( $45 + \frac{\theta}{2}$ ),  $\psi_f$  is the actual slope,  $U$  is the mean water pressure normal to the failure surface and,  $V$  is mean water

pressure in the tension crack (Hoek and Bray, 1981). Considering  $Z$  the depth of the tension crack and  $\gamma$ ,  $\gamma_w$  are the unit weights of the slab and water. These values were also the baseline of Thomas et al. (2009) study.

Under dry conditions, we considered that there are no forces in the tension crack and sliding surface, even though we account for some level of soil moisture, no significant force related to pressure is created. In this case, the forces  $V$  and  $U$  are zero and the equation 1 can be written by:

$$F = \frac{c A}{W \sin \psi_p} + \cot \psi_p * \tan \theta \quad (8)$$

Conversely, when water is present only in the tension crack or slab failure after a long dry period and assuming that the rest of the slab besides the crack area is impermeable, the significant force will be only due to the tension crack. The uplift force, in this case, can be neglected ( $U=0$ ). So, equation 1 is reduced to:

$$F = \frac{c A + (W \cos \psi_p - V \sin \psi_p) \tan \theta}{W \sin \psi_p + V \cos \psi_p} \quad (9)$$

Nowadays, advances in limit-equilibrium modelling are integrated to geographic information systems (GIS) and rather to deal with calculation in a single slope it can process medium and large areas at once. One example is CHASM (Combined hydrology and stability model, implemented in web-based GIS, is a physically-based model with two-dimensional landscape (Thiebes et al. 2013). It is use Bishop's simplified circular method to perform landslide analyses and prediction through Factor of Safety computation.

It is critical to assume that all the measurements must to be in a unique spatial scale. There are many long-term experiments that collect their variables using local and very precise coordinate system (e.g. Tomer et al., 2005). In order to adjust these measurements to

actual source of information (e.g. LiDAR), the process of coordinate transformation from one system to another is indispensable. When the survey has produced only X, Y coordinates, the two-dimensional transformation is employed. If vertical or elevation measurements are included so, three-dimensional coordinate's transformation is recommended. The former transformation imply operations of scaling, rotation and translation. In case of more than 2 control points exists, the least square solution can be applied. Least square adjustment is a mathematically superior method that give the most probable values for unknown points placed in a local coordinate system that need to be transformed in a actual GPS system. The three-dimensional conformal coordinate transformation (e.g. seven-parameter similarity transformation) produce lower RMS errors of rotations, translations and scale factor of x, y and Z axes (Ghilani and Wolf, 2006).

Considerable modeling tools has been developed to understand landscape changes over time. However, there is little investigation on modeling classic gully slopes stability with high resolution elevation data and limit equilibrium methods. Furthermore, a few studies are focused only in the gully headcut height evolution (Rengers and Tucker 2014; Campo-Bescós et al., 2013) rather than processing the entire gullied area. Thus, tools that deal with high spatial resolution and incorporate the entire gullied area will help also to implement precision conservation practices in needed areas.

Early study has pointed out that gully wall stability knowledge advances are achieved according diverse areas: soil characteristics evolved of the mass failure, techniques to compute slope stability, leading factors of gully development, and use of field measurements (Skempton, A. W. and Hutchinson 1969). This study will focus only the second area through the use of factor of safety approach.

The goal of this study is evaluate the slope stability analysis approach in two multi-temporal surveys located in a western Iowa gullied area for precision conservation purpose. The study simulated spatial failures of gully head and gully walls of digital elevation models with 15-year of span. The surveys were obtained with GPS receptor device and from Iowa Light Detection and Ranging (LiDAR) campaign carried out on February of 1999 and May of 2014 respectively. Furthermore, to investigate spatial distribution of critical instable slope in the site area we has employed a cutoff elevation based on multiple of slope standard deviation.

## **3.2 Material and Methods**

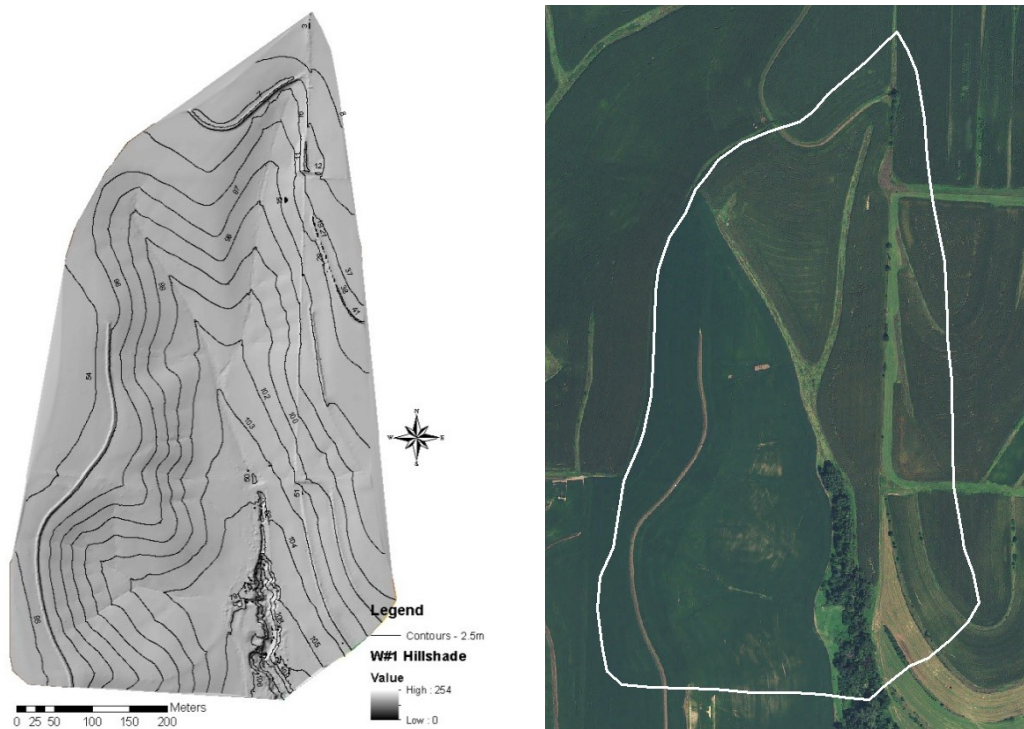
### **3.2.1 Study site**

The study area, a field-scale watershed (0.31 Km<sup>2</sup>), was chosen due the incidence an gully and its location in the Loess Hills landform at Treynor County, IA (41° 9' 44.54"N, 95° 38' 19.94" W). Entitled as Watershed #1 (Fig. 3.2), this field is one of four study areas established by the U.S. Department of Agriculture Research Service (USDA-ARS), in 1964 and nowadays is managed by a private farmer. This site one of most worldwide instrumented and therefore research documented related to gully erosion and runoff under conventional and no-tillage treatments (Poel et al. 1986; Kdrilen et al. 1964; Tomer et al. 2005). The main goal of USDA at the period of experiment was to analyze the impacts of conservation practices in runoff and water-induced soil erosion. Thus, the watershed was instrumented in 1965 to provide measurements of runoff, base flow and sediment concentration<sup>1</sup>. These measurements were quantified using broad-crested V-notch weirs located at the base of each

---

<sup>1</sup> Data available at: <<http://hrsl.ba.ars.usda.gov/wdc/ia.htm#Treynor>>

watershed where the gullies channels are located. Precipitation was measured by rain gauges placed in the watershed perimeter. Since 1963 to 1995, watershed 1 was managed in continuous corn and conventional tillage. From 1996, a no-tillage with rotation of corn and soybean has started (Tomer et al. 2006). The causes and processes of the long-term growth of this valley-bottom gully are described by Thomas et al. (2004). Temporal yield variability among all watersheds is widely explored by Tomer et al. (2005), and recently, Tomer et al. (2007) conducted experiment in riparian buffers placed in this site with the objective to quantify sediment accumulation and respective phosphorus. Related to the gully location in Treynor site, it can be noticed that drainage area outlet is oriented towards south in the bottom of Fig. 3.2.



**Fig. 3.2** (a) Hillshade of Treynor site Pottawattamie County, Iowa derived from the LiDAR One Meter Digital Elevation Model. (b) 2011 National Agriculture Imagery Program (NAIP) Four-Band Aerial Photography highlighting boundaries of watershed #1 boundary (light line).

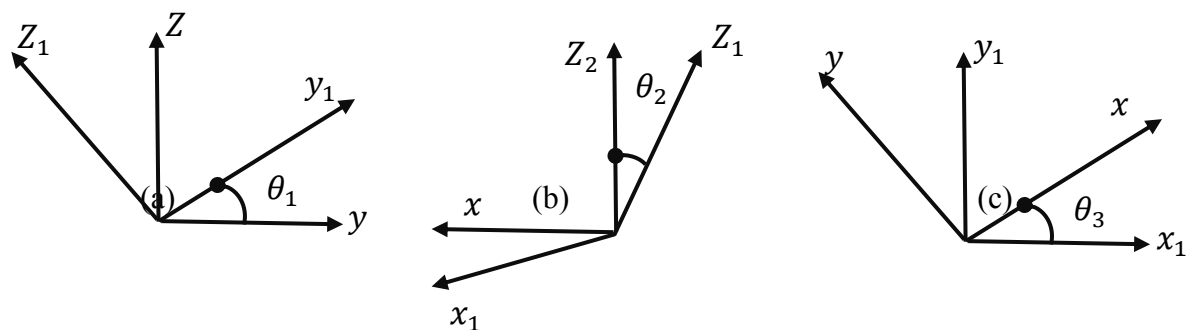
With authorization of the actual land owner, a topographic survey was conducted on May, 7<sup>th</sup>, 2014 in watershed 1, Treynor, IA. A GPS Garmin GPSMAP 62sc receiver device with horizontal accuracy of 5 meters was used to collect elevation points in the gully edges and gully floor. The gully bed flow path was defined using existent surface stream flow in this area. Points were set in universal transverse Mercator (UTM) coordinate system.

### **3.2.2 Gully perimeters coordinate adjustments in past surveys**

Thomas et al. (2004) data were used in our study. The dataset include elevation points and lines of gullies perimeter both in shapefiles format. They consolidated surveys in a single map file, represented as lines, from the boundaries of Treynor site gully top edges. However, all the lines, that registered the widening of this gully over the period of 1964 to 2000, were collected in a local coordinate system. To aggregate recent topographic and hydrologic features based on high resolution data it was crucial employ a technique that transfer local 2D and 3-dimensional coordinate system to GPS coordinate systems. For this purpose a two and three-dimensional conformal coordinate transformation were employed according steps discussed in Ghilani & Wolf (2006).

The three-dimensional conformal coordinate transformation requires seven parameters - three rotations, three translations, and one scale factor. The rotation matrix is obtained using three two-dimension rotations successively around x, y, and z axes. Fig. 3.3 shows examples of rotation  $\theta_1, \theta_2, \theta_3$ , about the x, y, and z axes expressed in matrix format.





**Fig. 3.3** Examples of (a)  $\theta_1$  (b)  $\theta_2$  (c)  $\theta_3$  rotation in three-dimensional conformal coordinate transformation extracted from Ghilani and Wolf (2006).

The rotation  $\theta_1$ , about x axis can be expressed in matrix form (Fig 3.3a)

$$X_1 = R_1 X_0 \quad (10)$$

Where

$$X_1 = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} \quad R_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_1 & \sin \theta_1 \\ 0 & -\sin \theta_1 & \cos \theta_1 \end{bmatrix} \quad X_0 = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Fig. 3.3b shows the rotation  $\theta_2$ , about axis y expressed in matrix format

$$X_2 = R_2 X_1 \quad (11)$$

Where

$$X_2 = \begin{bmatrix} X_2 \\ Y_2 \\ Z_2 \end{bmatrix} \quad R_2 = \begin{bmatrix} \cos \theta_2 & 0 & -\sin \theta_2 \\ 0 & 1 & 0 \\ \sin \theta_2 & 0 & \cos \theta_2 \end{bmatrix}$$

Fig. 3.3c illustrates, the rotation  $\theta_3$ , about axis z expressed in matrix format

$$X_3 = R_3 X_2 \quad (12)$$

$$X = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad \text{and} \quad R_3 = \begin{bmatrix} \cos \theta_3 & -\sin \theta_3 & 0 \\ -\sin \theta_3 & \cos \theta_3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Substituting Equation (a) into (b) and turn into (c) yields

$$X_2 = R_3 R_2 R_1 X_0 = R X_0 \quad (13)$$

Multiplying matrices  $R_3, R_2$ , and  $R_1$  in equation (d) result in a single rotation matrix R for the transformation whose individual elements can be obtained by

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (14)$$

Where

$$r_{11} = \cos\theta_2 \cos\theta_3$$

$$r_{12} = \sin\theta_1 \sin\theta_2 \cos\theta_3 + \cos\theta_1 \sin\theta_3$$

$$r_{13} = -\cos\theta_1 \sin\theta_2 \cos\theta_3 + \sin\theta_1 \sin\theta_3$$

$$r_{21} = -\cos\theta_1 \sin\theta_2$$

$$r_{22} = -\sin\theta_1 \sin\theta_2 \sin\theta_3 + \cos\theta_1 \cos\theta_3$$

$$r_{23} = \cos\theta_1 \sin\theta_2 \sin\theta_3 + \sin\theta_1 \cos\theta_3$$

$$r_{31} = \sin\theta_2$$

$$r_{32} = -\sin\theta_1 \cos\theta_2$$

$$r_{33} = \cos\theta_1 \cos\theta_2$$

Taking advantage of the rotation matrix which is orthogonal, we can use one of its property that is inverse is equal to its transpose. So, using its transpose form and multiplying matrix X by a scale factor, S, also adding translation factor  $T_x$ ,  $T_y$ , and  $T_z$  it was possible to

convert values to a common origin yields. The following mathematical equation consider this transformation

$$\begin{aligned}
 X &= S (r_{11}x + r_{21}x + r_{31}x) + T_x \\
 Y &= S (r_{12}x + r_{22}x + r_{32}x) + T_y \\
 Z &= S (r_{13}x + r_{23}x + r_{33}x) + T_z
 \end{aligned} \tag{15}$$

Equation set (15) contain seven unknowns variables (S,  $\theta_1$ ,  $\theta_2$ ,  $\theta_3$ ,  $T_1$ ,  $T_2$ ,  $T_3$ ). For a unique solution the same number of equations must to be written. Other requirement is a minimum of two control stations with known XY coordinates (e.g. UTM), and also local xy coordinates (e.g. Cartesian). Meeting the requirement of number control points, a least-squares solution can be applied. The equations (15) are nonlinear considering their unknown variables and therefore have to be linearized in order to find theirs a solution. The linearized equations are described following for each point

$$\begin{bmatrix}
 \left(\frac{\partial X}{\partial S}\right)_0 & 0 & \left(\frac{\partial X}{\partial \theta_2}\right)_0 & \left(\frac{\partial X}{\partial \theta_3}\right)_0 & \left(\frac{\partial X}{\partial \theta_3}\right)_0 & 1 & 0 & 0 \\
 \left(\frac{\partial Y}{\partial S}\right)_0 & \left(\frac{\partial Y}{\partial \theta_1}\right)_0 & \left(\frac{\partial Y}{\partial \theta_2}\right)_0 & \left(\frac{\partial Y}{\partial \theta_3}\right)_0 & \left(\frac{\partial Y}{\partial \theta_3}\right)_0 & 0 & 1 & 0 \\
 \left(\frac{\partial Z}{\partial S}\right)_0 & \left(\frac{\partial Z}{\partial \theta_1}\right)_0 & \left(\frac{\partial Z}{\partial \theta_2}\right)_0 & \left(\frac{\partial Z}{\partial \theta_3}\right)_0 & \left(\frac{\partial Z}{\partial \theta_3}\right)_0 & 0 & 0 & 1
 \end{bmatrix}
 \begin{bmatrix}
 dS \\
 d\theta_1 \\
 d\theta_1 \\
 d\theta_1 \\
 dT_x \\
 dT_y \\
 dT_z
 \end{bmatrix}
 =
 \begin{bmatrix}
 X - X_0 \\
 Y - Y_0 \\
 Z - Z_0
 \end{bmatrix} \tag{16}$$

Where

$$\frac{\partial X}{\partial S} = r_{11}x + r_{21}x + r_{31}x \quad \frac{\partial Y}{\partial S} = r_{12}x + r_{22}x + r_{32}x \quad \frac{\partial Z}{\partial S} = r_{13}x + r_{23}x + r_{33}x$$

$$\frac{\partial Y}{\partial \theta_1} = -S[r_{13}x + r_{23}x + r_{33}x] \quad \frac{\partial Z}{\partial \theta_1} = s[r_{12}x + r_{22}x + r_{32}x]$$

$$\frac{\partial X}{\partial \theta_2} = S(-x \sin\theta_2 \cos\theta_3 + y \sin\theta_2 \sin\theta_3 + z \cos\theta_2)$$

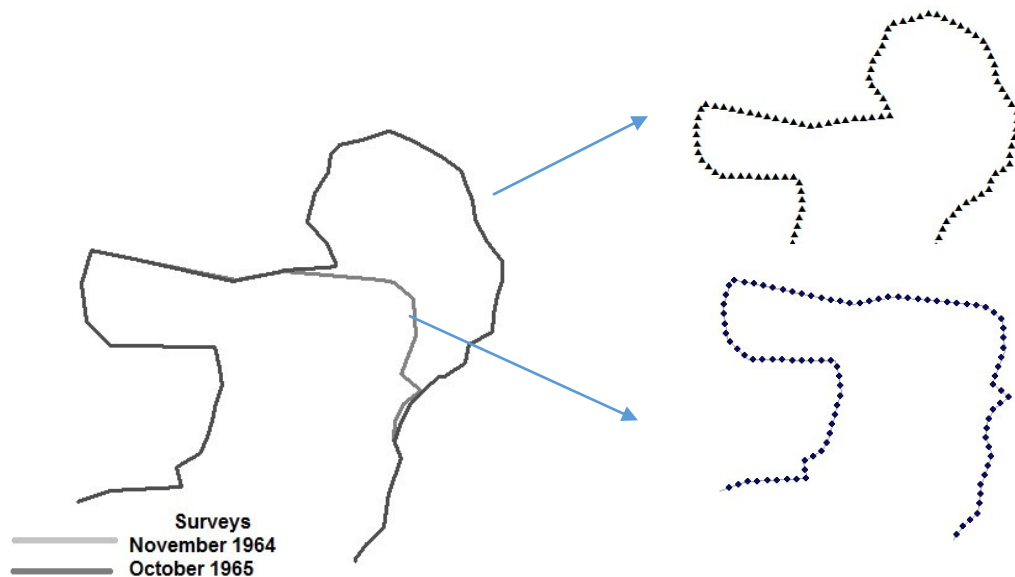
$$\frac{\partial Y}{\partial \theta_2} = S (x \sin \theta_1 \cos \theta_2 \cos \theta_3 - y \sin \theta_1 \cos \theta_2 \sin \theta_3 + z \sin \theta_1 \sin \theta_2 )$$

$$\frac{\partial Z}{\partial \theta_3} = S (-x \cos \theta_1 \cos \theta_2 \cos \theta_3 + y \cos \theta_1 \cos \theta_2 \sin \theta_3 + z \cos \theta_1 \sin \theta_2 )$$

$$\frac{\partial X}{\partial \theta_3} = S(r_{21}x + r_{11}y) \quad \frac{\partial Y}{\partial \theta_3} = S(r_{22}x + r_{12}y) \quad \frac{\partial Z}{\partial \theta_3} = S(r_{23}x + r_{13}y)$$

In this study, initially, individual line correspondent to an individual survey was converted in a sequence of points spaced by 1 meter, to be consistent with further steps as illustrated by Fig. 3.4. Each sequence of points was lately adjusted to a WGS83 datum by 3D or 2D conformal coordinate transformation trough least square approximation technique.

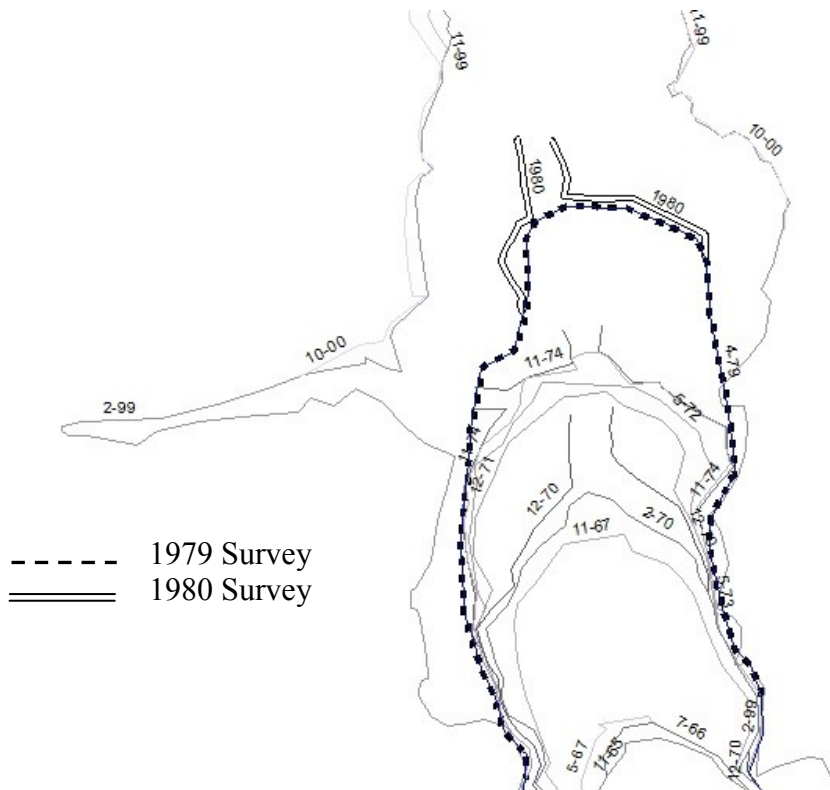
Elevation values and Northing and Easting coordinates of 47 controls stations with known coordinate values in a local and reference coordinate system were used to adjust coordinates in unknown local points. Survey performed in 1999 which have captured



**Fig. 3.4** Individual lines regarding to November, 1964 and October, 1965 survey consolidated by Thomas et al. (2004) transformed in this study in equal-spaced points sequence.

elevation variable was the only one to be transformed through 3D coordinate transformation. The others survey only contained Northing and Easting coordinates information, so two-dimensional transformation was employed.

It is worthy to mention that some surveys, available from Thomas et al.( 2004) were not performed in the entire gully perimeter as depicted in Fig. 3.5. So, at that time, the outcome survey encompassed only a few lines. So, for this study, we supposed that newly eroded areas were composed by the actual line plus the previous surveys. This step was executed using the actual polylines to reconstitute the entire perimeter in each survey. A few times, we have to digitize manually the parts of the perimeter, however, dealing to not introduce new errors in the perimeter position.



**Fig. 3.5** Set of gully perimeters showing the partial surveys executed, for instance, survey lines produced at April 1979 and 1980 campaigns.

This study focused in the surveys carried out in February, 1999 and May, 2014. The later besides the topographic in situ survey also the LiDAR-based digital elevation model ( $n=330,678$ , mean point spacing = 1.4 m), was produced in the gully area. The digital elevation model produced to former survey was originated by GPS points collected by Thomas et al. (2004) within the same gully area ( $n= 1,200$ ). Both surveys used IDW interpolation technique to produce elevation and slope maps, gridded at 1-m grid for individual pixels. In addition, it was performed a procedure tailored to smoothing topography in dense vegetated environments. The smoothing procedure deal to remove points that were lying in high vegetation and consequently hidden bare topography. We applied the approach to extract vegetation implemented by Evans & Hudak (2007) using the multiscale curvature classification (MCC) algorithm.

### **3.2.3 Factor of safety computation and implications to precision conservation**

In this study, topography is represented by a set of independent points with 1-m pixel resolution. We considered that computed factor of safety equal to one represents a limiting equilibrium. In this condition there is a balance between the forces tending to sliding the slope and the one resisting to this failure. FS greater than one, the slope is consider stable because the force to resist to slide is greater than the force trying to slide the block. Slope instability occur when a plan failure is given, as expected, with factor of safety smaller than 1. So, this is the threshold condition to trigger elevation development and gully edges displacement. In this way, gully edge wide by the same scale of slab thickness ( $x$ ). This is considered an instable geometry condition and plan failures happen by gravitational and resistant forces.

Since this research is focused in the same watershed of Thomas et al.(2009) study, we assume the same boundary conditions for the majority of attributes values required by Equations 8 and 9. The attributes values are: friction angle ( $\theta = 22.9^{\circ}$ ), failure plane angle ( $\psi_p = 45^{\circ} + \theta/2$ ), height of the banks top above the datum used for measured head ( $H = 5$  m), unit weight of the water ( $\gamma_w = 9.81$  KN/ m<sup>3</sup>), unit weight of the dry slab ( $\gamma_{dry} = 13$  KN/ m<sup>3</sup>), unit weight of the slab with 50% water content ( $\gamma_{50\% \text{ sat}} = 15.581$  KN/ m<sup>3</sup>), minimum range of slab thickness ( $x = 0.3$  m), and critical head required for bank failure ( $Z_{crit} = 3.54$  m) and, depth of the tension crack ( $Z = 2$  m). The later, height of failure is reported by Poel et al. (1986) as between 3 and 4 meters. Slab (H) height values were calculated through focal analysis function available at ArcGis using a rectangle shape (2x2 pixel cell) which calculate the range of elevation at pixels neighborhood. Slope was computed from digital elevation model obtained at Thomas et al. (2009) field work. Automation of factor of safety estimates was implemented through a R script.

Pixels points with the computed factor of safety values considered instable ( $0 < FS < 1$ ) are target of precision conservation practices. In this study, we simulated two threshold conditions considering points that exceeded one and two multiple of slope standard deviation of slope once the data follow Gaussian distribution.

### 3.3 Results and Discussion

The surveys described at Thomas et al. (2009) work were transformed through two- and three-dimensional conformal technique according data availability. The computed length after the transformation is shown in Table 3.1. The maximum error from original length to transformed is 12.4 % related to 1973-year, May survey and the minimum is 1.1 % in 1980 survey. The total average of length difference is 2.84%. The transformation error for all

surveys were around 0.01 for rotations at x and y axis. Thus, the approach taken in this study was able to transpose the survey at local coordinates system to UTM without significant differences.

Comparing the digital elevation models of 1999 and 2014 survey, this study found out that computed slopes in the gully edge points originated from 2014 survey had higher maximum, and mean values ( $S_{\text{mean}}=29.62^{\circ}$ ,  $S_{\text{max}}=75.43^{\circ}$ ) compared to 1999 survey (Fig. 3.6). The difference between the mean elevation at the study period of 2014 and 1999 survey is 4.8 m. There is strong evidence that paired means of slopes and elevation in both survey are not similar (p value < 0.01). Furthermore, the methods of obtain these slopes were different. In a former survey the slopes were calculated from airborne-LiDAR campaign ( $n= 330,678$ ,

**Table 3.1** Survey timeline with gully length obtained from Thomas et al. (2004) set of perimeters and transformed with two-dimensional conformal transformation.

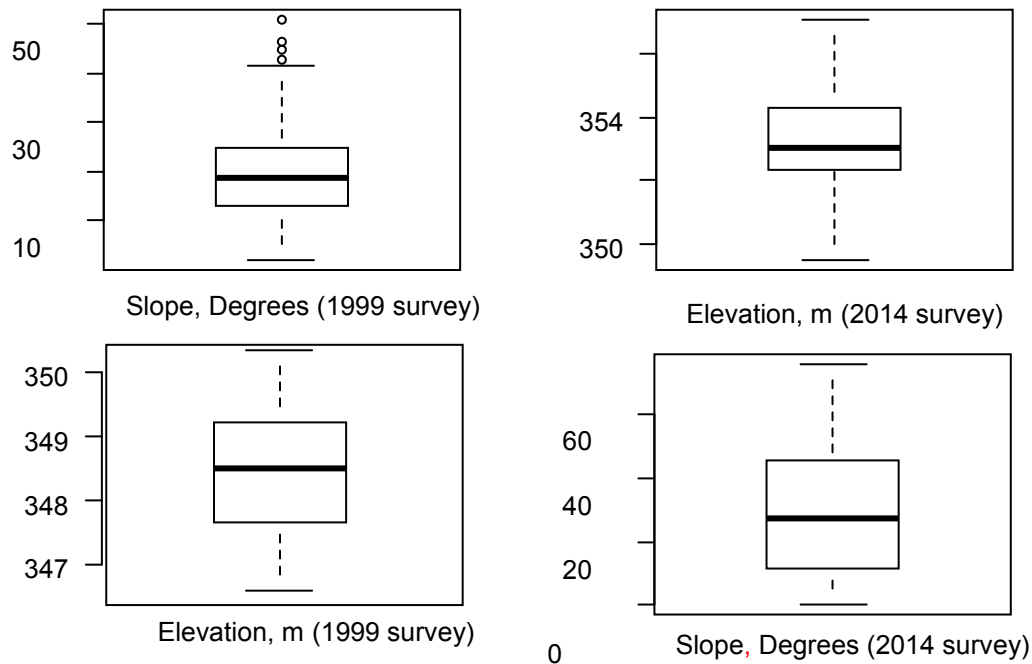
Survey Year	Survey Month	Original surveyed perimeter length (m)	Transformed perimeter length (m)	Difference (%)
1965	April	100.52	97.701	2.8
1965	August	124.95	121.53	2.7
1966	July	137.1	134.46	1.9
1967	May	140	136.33	2.6
1967	November	170.82	167.53	1.9
1970	February	181.15	178.263	1.7
1970	December	191.8	189.52	1.2
1971	December	199.5	195.854	1.8
1972	May	207	202.511	2.1
1973	May	223.5	252.349	12.4
1974	November	208.1*	216.87	4.21
1979	April	236	232.989	1.27
1980	-	247.6	244.864	1.1
1999	February	445.7	436.363	2
1999	July	461	439.934	4.6
1999	November	451.7*	441.977	2.15
2000	October	448.57*	439.669	2

\*The headcut was not delineated from these surveys, only the parallel two lines of wall banks. So, the real length might be greater than those described here.

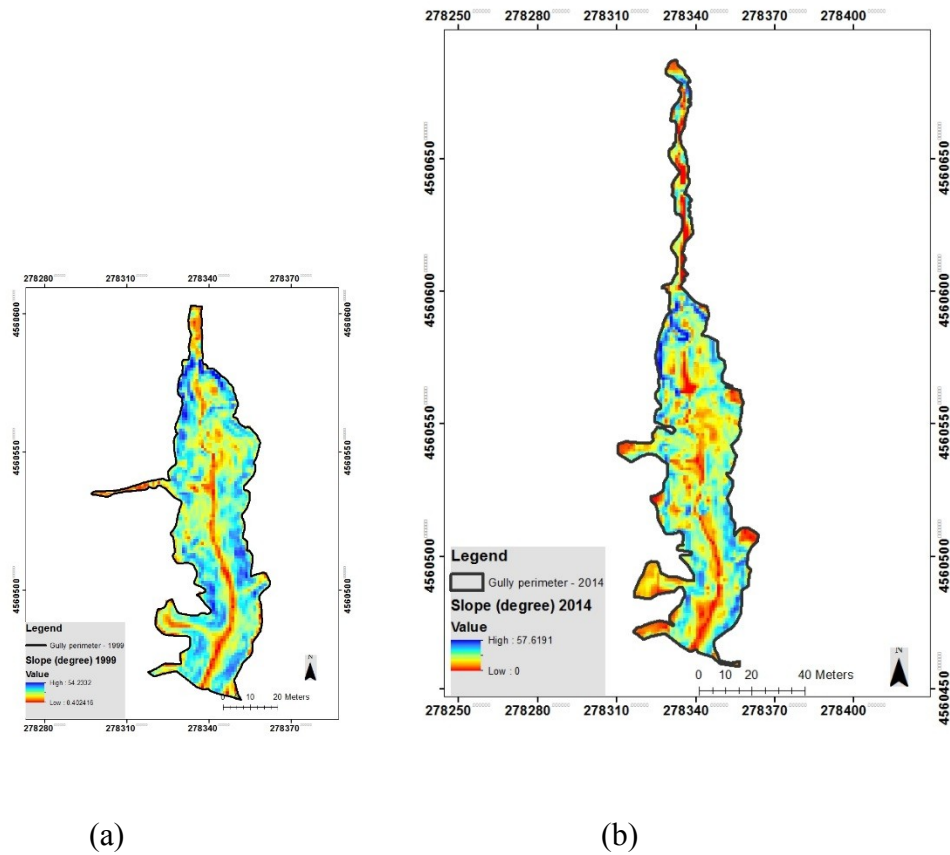


mean point spacing = 1.4 m) for the whole watershed 1, while in the later a dense grid of points was collected manually ( $n=1742$ ) mainly into the gully feature. Considering only the gullied area, each digital elevation model was discretized in 3986 points for 2014 survey and 3128 for 1999-survey with 1-m pixel size.

The slope elevation related to 1999 survey, depicted in Fig. 3.7a, revealed clear pattern of lower and flat elevation in the gully channel bed comparing to the side walls. In addition, steep side walls are located in both sides of headcut. The LiDAR-derived slope map (Fig 3.7b) showed a much detailed terrain. However, because the original slope map was noisy due to presence of high vegetation captured by LiDAR survey it was necessary smooth slope values using the multiscale curvature (Evans and Hudak 2007) approach. A small and narrow channel was mapped as extension of gully headcut in 2014. Since the Figure 3.7b map was obtained overlapping the GPS pathway with LiDAR Digital elevation model, we



**Fig. 3.6** Boxplots showing the dataset distribution of elevation and slopes in gully edges both from survey periods on 1999 and 2014.

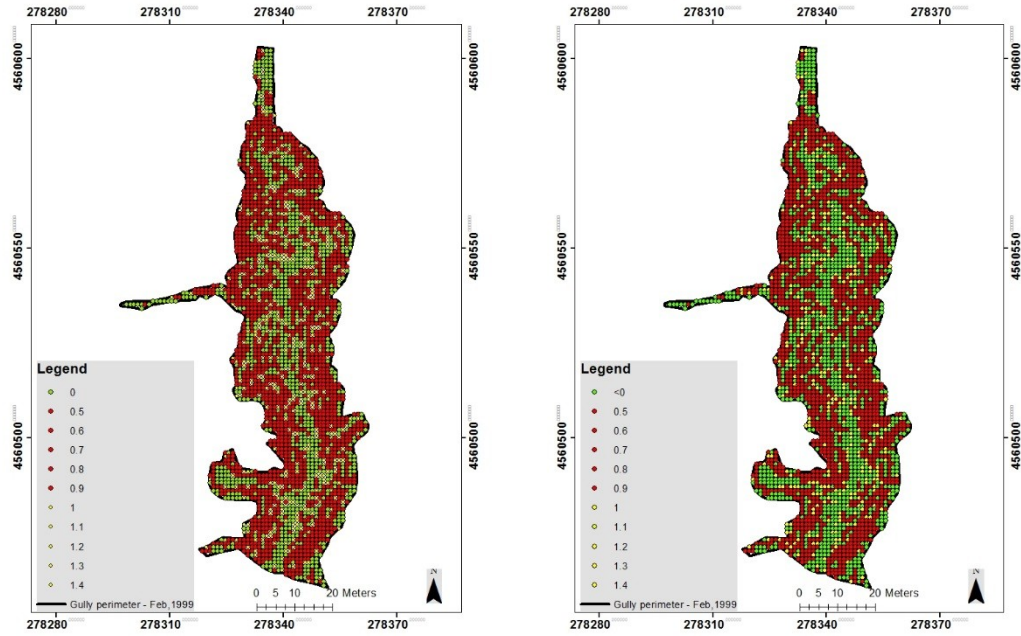


**Fig 3.7** Slope map and gully perimeter at the gullied area in Treynor site on (a) February, 1999 originated from Thomas et al. (2009) study, and (b) May, 2014 from LiDAR data.

have noted some area with small slope, probably errors due GPS resolution and LiDAR inaccuracy.

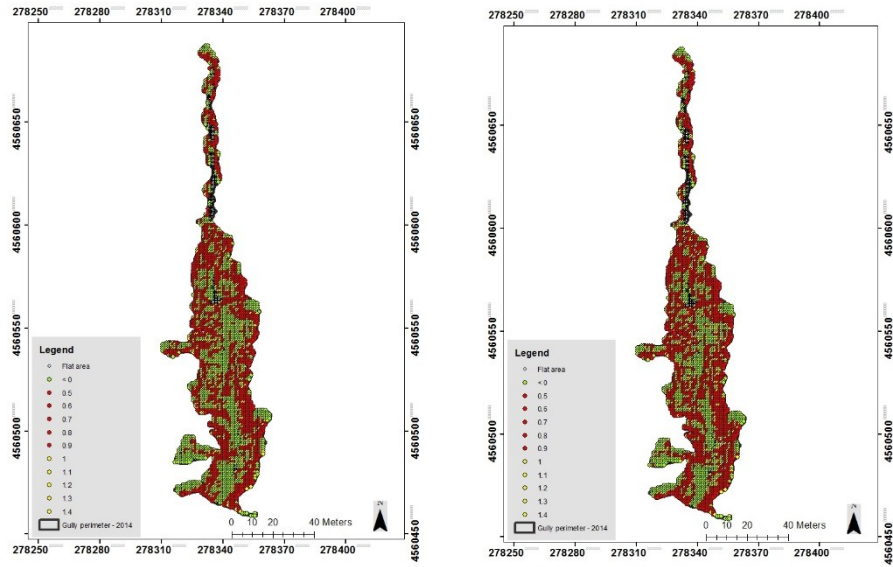
Since we decided to include the extended gully headcut at 2014 where it was observed recent side wall erosion, the area ( $3991.87 \text{ m}^2$ ) and perimeter (728.5 m) were larger and longer compared with 1999 survey area ( $3134 \text{ m}^2$ ) and perimeter (475.8 m).

Spatial distribution of computed factor of safety in the whole gullied area on 1999 and 20014 survey is shown in Fig. 3.8 Excluding the negative computed FS, which represent small slopes that are not prone to sliding, totalized 81.3% of the pixels on 1999 survey that had  $FS < 1$  against 78.2 % on 2014 survey, both regarding to partially saturated conditions.



(a)

(b)



(c)

(d)

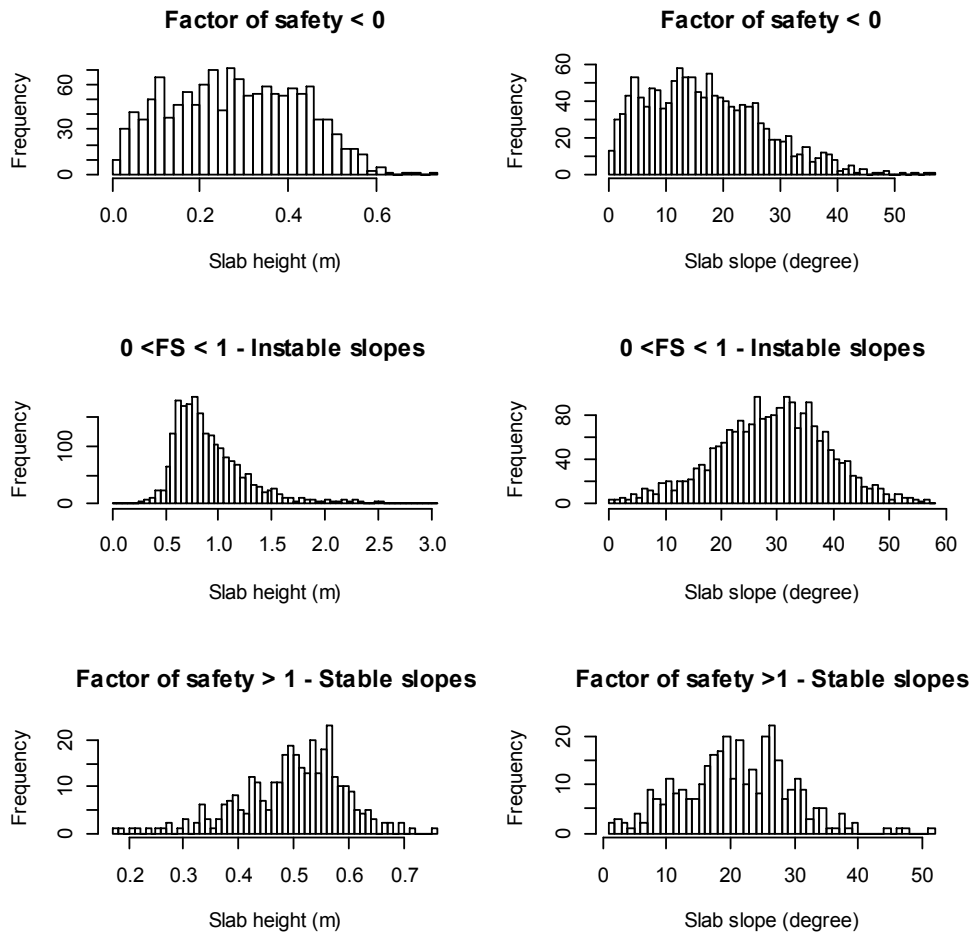
**Fig 3.8** Spatial distribution of factor of safety regarding to partially saturated failure ( $\gamma_w = 15.5 \text{ KN/m}^3$ ) on 1999 (a,b), and 2014 survey (c,d) under dry conditions. Negative values were computed and are placed mainly in the gully bed.

Therefore, comparing FS maps of both survey, Treynor site had experienced a reduction of points prone to slope failure according slope equilibrium technique over time. In addition, Welch's t-test employed between 1999 and 2014 factor of datasets, for both water content conditions, obtained p-value  $< 0.001$  which agree with the alternative hypothesis that the surveys have statistically different mean values. Commission errors, GPS accuracy, and the intrinsic morphologic-entropic landscape changes may explain the reasons because 1999 survey perimeter is not totally confined in 2014 gully perimeter as illustrated in Fig. 3.8. For example, it was noticed from field work that the distinct branch, located in west portion of gully that appeared in 1999 survey is, at the time of this study, surrounded by a strip grassed area which stabilized gully erosion and smooth slope angles. In addition, concentrated flow shown in Fig. 3.8 reveals that points selected from factor of safety greater than 1 are in the patch or very close to the concentrated flow path.

Related to the data distribution of slope stability, histograms depicted in Fig. 3.9 showed that instable slopes ( $0 < FS < 1$ ) at LiDAR DEM (2014 survey) had average height of 0.92 m with mean slope of  $30^\circ$ ; the maximum slope it was  $57^\circ$  and highest side wall height of 3 m. Conversely, stable slopes ( $FS > 0$ ) in the same survey had shorter mean side walls (0.5 m) and were less steep angles ( $21^\circ$ ). Flat areas, mainly located in the gully bed ( $FS < 0$ ), were evenly distributed with the smaller mean height (0.3 m) and mean slope ( $16.7^\circ$ ). Similar data distribution of slope and slopes height at 1999 survey, not shown here, were found similar compared to 2014.

### **3.3.1 Sensitivity analysis**

Although several studies have been performed in Treynor site (Thomas et al., 2009; Thomas et al. 2004; Bradford et al., 1973; Bradford and Piest, 1977) related to slope stability,



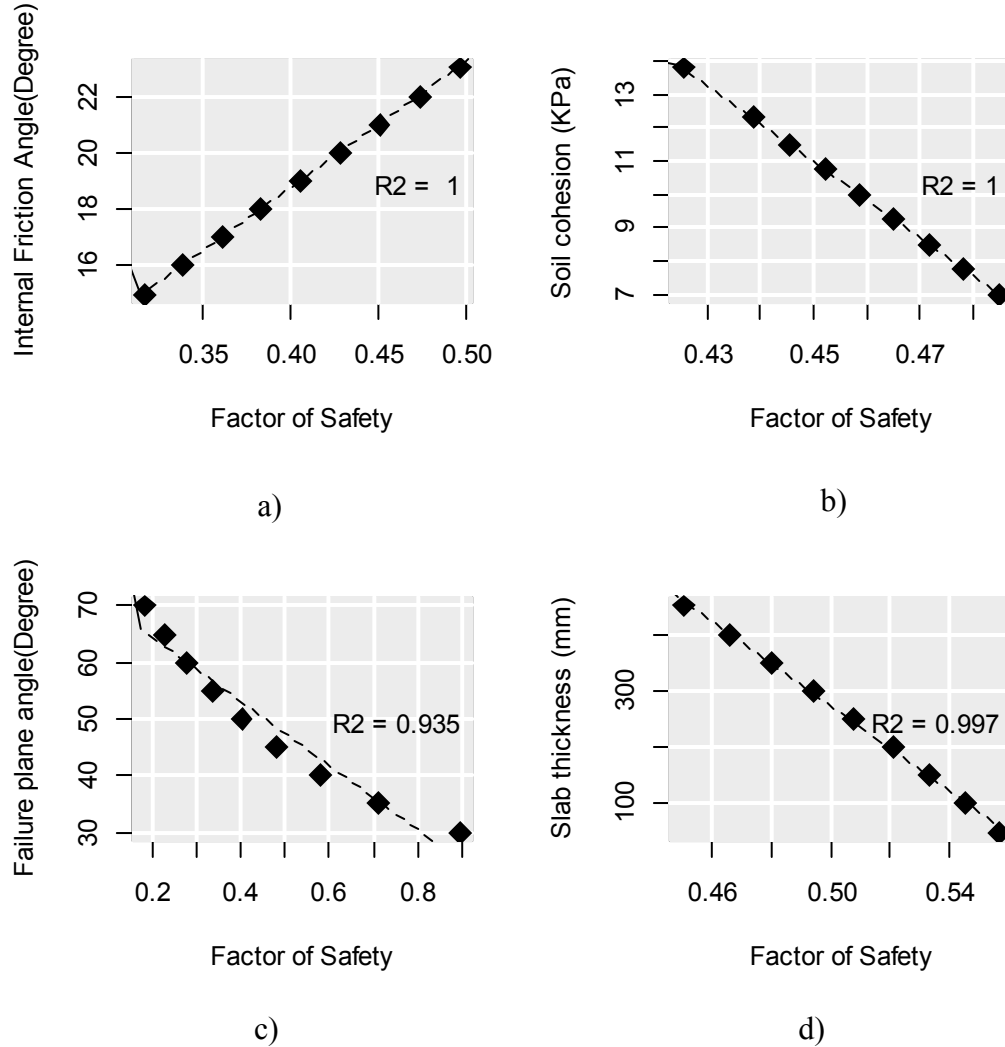
**Fig 3. 9** Histograms of slab height (m) and slope (degrees) at Treynor site on May, 2014 at dry and partially saturated conditions coded in a) negative FS (n=961) b) instable slopes (n=1761) c) stable slopes (n=370).

there are still some limitations into many variables values which cause uncertainty on simulations outcomes. Furthermore, there are different effects in slope equilibrium due to soil and topographic variables variation. Gray and Megahan's (1980) study suggested that models of slope instability are most susceptible to soil cohesion, root cohesion, soil depth, and slope angle. Bradford and Piest (1977) also, stressed out that elevation of water table/phreatic surface decrease the ratio of resisting forces to the driving forces resulting in a slope failure. Many are the variables critical to study to simulation purposes. Therefore, we evaluate here

only the effects of internal friction, angle, soil cohesion, failure plane angle, and crack thickness. In order to proceed the sensitivity analysis, one parameter is changed while others are held constant. We used tests of direct shear strength results performed by Bradford and Piest (1980) on loess alluvium of the Deforest Formation. The range of friction angles is  $15^\circ$  to  $23^\circ$  with cohesion varying from 7 to 13.8 KPa, respectively. Thickness of the slab varies from 50 to 450 mm consistent with was found by (Bradford and Piest, 1980) in the loessial area in Western Iowa and failure plane angle ranged from  $30^\circ$  to  $70^\circ$ . In this study, the sensitivity analysis was performed in one point chosen at each survey. Those were instable points according slope instability analysis. The results showed in Figure 3.10 are associated to the point with highest elevation near to the sidewalls. This point has the following boundary conditions:  $x=278336.7$  m,  $y=4560582$  m; elevation difference: 2.51 m; slope= $53.3^\circ$ ; FS = 0.32.

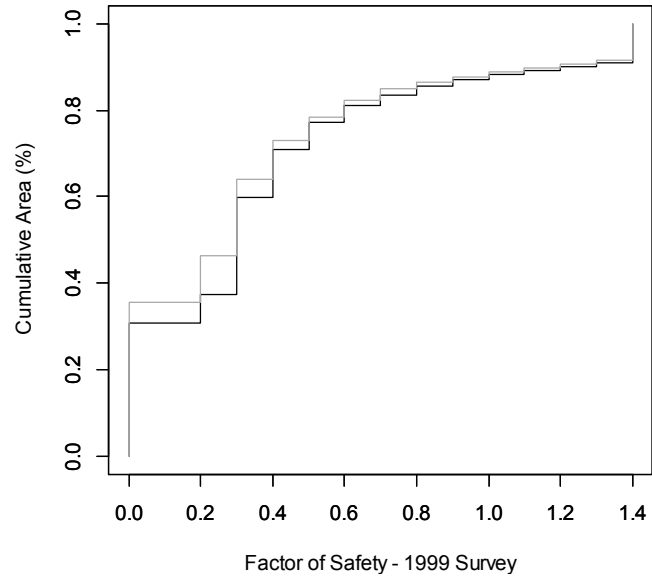
Sensitivity analysis underscores the importance of the four keys factors affecting slope failure in the gully topography. Variance analysis demonstrated that changes in the internal friction angle ( $\theta$ ) have the largest effect on slope stability factor (S.D.<sub>1999</sub> = 0.18, S.D.<sub>2014</sub> = 0.24) comparing to variations of soil cohesion, failure plane angle and slab thickness. The variable  $\theta$  was the only one that increase factor of safety, even though the slopes are still instable. Increase of soil cohesion which imply decrease of water content produced an unexpected decrease of factor of safety of 114-fold.

If the weight of the slab varies from to 13 (dry slopes) to  $15.5 \text{ KN/m}^3$  (partially saturated slopes), failures volumes increases of 7.13% in 1999 survey and 7.53% at 2014 survey. Figure 3.10a and b show the changes of factor of safety distribution caused by 16.13% variation in unit weight in different water content.

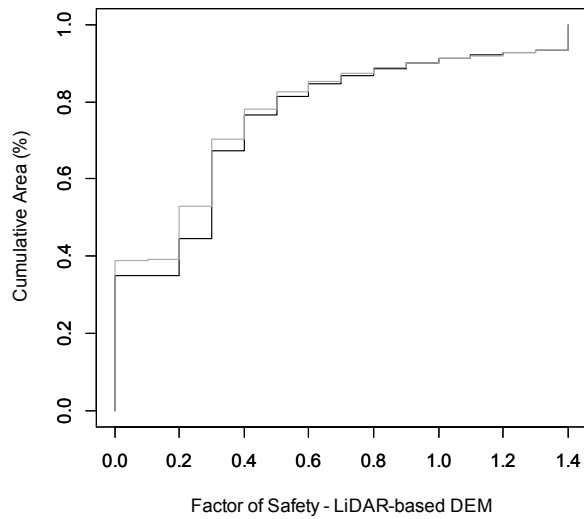


**Fig. 3.10** Impacts of variables uncertainties on slope instability: (a) internal friction angle, ( $\theta$ ) b) soil cohesion, c) failure plane angle, ( $\psi_p$ ) and, d) slab thickness, ( $x$ ) at 1999 survey.

The approach given in this study is different from that implemented in Tomer et al., (2007). In Tomer's work a multispecies riparian buffer, including switch grasses and trees, were installed in the west side of gully at watershed 1 dealing capture sediment and phosphorus from runoff events. Rather than design a crop cover outside of the gullied area; and, since this area is unlikely to be used as a farmland, we presented areas within the gully



(a)



— Partially saturated condition ( $15.5 \text{ KN/m}^3$ )  
 — Dry condition ( $13 \text{ KN/m}^3$ )

(b)

**Fig. 3.11** Empirical cumulative distribution (CDF) of factor of safety on 1999 and 2014 surveys at Treynor gullied area.

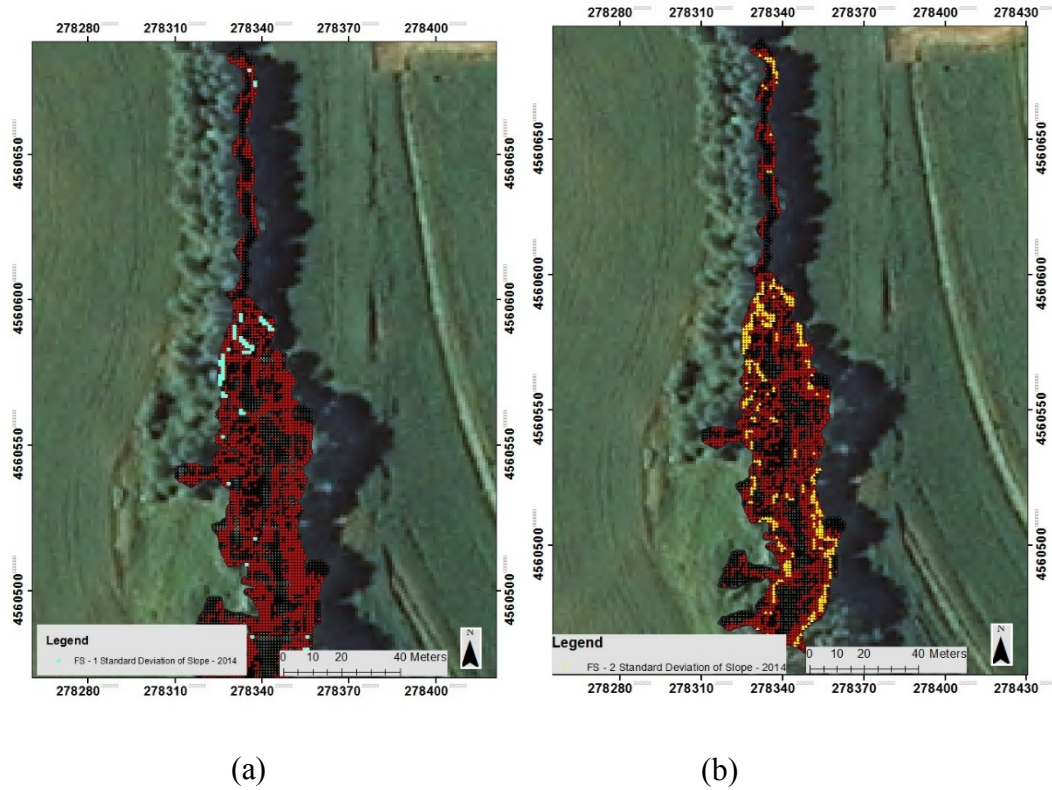


that are instable, so requiring an alternative approach. The areas were simulated computing the difference of one ( $S_{1SD} = 9.95^\circ$ ) and two units of standard deviation of slope dataset from maximum slope ( $S_{max} = 57.6^\circ$ ) for 2014 survey (Fig. 3.11).

### **3.3.2 Critical instable slopes for precision agriculture purpose**

The maps with instable points represented by pixels that exceeded one and two standard deviation of maximum slope considering partially saturated as a conservative scenario are illustrated in Fig. 3.12. With one standard deviation, 61 pixels were found and with condition with two SD, 396 pixels (1 pixel = 1m<sup>2</sup>) were found. The majority of the points in both condition area located in the headcut and side walls, however, not necessary cover the entire gully perimeter. So, this approach seems to fits in the conservation agriculture concept and prioritize areas that are more prone to failure due their geometrical features.

Since the gullied area has greater slope gradient, structures that increase soil stress shear and internal friction angle are required. Chen et al. (2010) empirical outcomes testing diverse material have found that woven mat with fertilizer strips is the material with greatest soil erosion resistant, for slopes of 35°. Additionally, De Baets et al. (2011) study have identified ryegrass, rye, oats and white mustard as the most appropriate plants for control high overland flow due their root density. Therefore, a strategy for control concentrated flow erosion that combine rigid structure, for instance, geotextile and crop cover with high root density seems suitable to sites that does not exhibits flashy stream flow when high intensity rainfall events occur.



**Fig. 3.12** Factor of safety maps with selected pixels considered instable slopes that are (a) one and (b) two standard deviation of elevation from maximum slope ( $S_{max} = 57.6\%$ ) at 2014 survey.

### 3.4 Conclusions

High uncertainty on mass failure on steep slope which impact soil and water quality have been promoting research studies that take into account spatial variability in agricultural and natural environments to implement precision conservation practices.. The main objective of this study was assess slope stability analysis in two multi-temporal surveys at a site located at western Iowa gullied area for precision conservation purpose. The study computed spatial failures (Factor of Safety index) in gully head, side walls, and bed using digital elevation models with 15-year spam. The surveys were obtained with GPS receptor device

and from Iowa Light Detection and Ranging (LiDAR) campaign carried out on February of 1999 and May of 2014 respectively

The main contributions of this study were: 1. The evaluation of two distinct topographic surveys was conducted in Treynor site under the perspective of slope instability in an entire gullied area. Many studies have computed FS for a few points in the study area instead (Bradford et al., 1973; Istanbuluoglu, 2005; Thomas et al., 2009). Outcomes of this assessment has shown statistically significance less instability in the actual slopes compared to 1999 survey slopes These results are based on the computed factor of safety index under dry and partially saturated conditions. 2. The computation of critical instable slopes within gully, based on multiples of the slope standard deviation as threshold. Thus, maps showing spatial distribution of instable slopes points were produced. One unit of slop standard deviation produced 61 m<sup>2</sup> with slopes ranging from 57.6° to 47.7° of more instable area. Two units of slope standard deviation threshold has produced a 396 m<sup>2</sup> area with slopes varying from 57.6° to 37.7°.

Relevant findings, including the good efficiency of natural or processed material to hold high slope gradient associated with high density root system cover plants, corroborate with the approach to prioritize conservation practice in needed areas as presented in this study. However, further empirical studies are need to test this approach. Considering that precision conservation is a relatively new concept, the use of slope instability is a novelty presented for this research considering natural environment affected by agricultural lands. Internal friction angle ( $\theta$ ) had the largest effect on slope stability factor ( $S.D._{1999} = 0.18$ ,  $S.D._{2014} = 0.24$ ), according the sensitivity analysis, compared to variations of soil cohesion, failure plane angle and slab thickness

As an intermediate result we highlight the transformation of past gully perimeters to actual coordinate systems obtained low error rates (mean = 2.14%) of x, y coordinate axis. So, the previous perimeters of gully monitoring over 36 year consolidated by Thomas et al. (2009) were converted in WGS84 coordinate system and available to be fully integrated with other layers of information such as Iowa high resolution elevation LiDAR data and soil survey provided by USDA

In conclusion, this simplified methodology was able to confirm field observation where zones in the vicinity of head walls are the most instable and have shown recent slope slumping due to rainfall events. Other factors must be included in the limit equilibrium model in order to obtain more realistic pattern of gully failure. As future research we will include the dynamics of failure in the headcut and side wall due to infiltration and seepage process. Rengers and Tucker (2014), for example reported the gully floor vegetation as a best model fit to trigger gully headcut dynamics. Include precipitation events with more granularity, considering intensity (mm/hour) rather than daily total precipitation and antecedent moisture will be further research.

### **Acknowledgments**

The authors are thankful to Mr Edward Marshall who allowed the field work at the study site, Edson Adriano Vendrusculo for his assistance during the fieldwork, John T. Thomas for data support, and the research funding provided by EMBRAPA - *Brazilian Agricultural Research Corporation* - supporting the first author in her PhD program at Iowa State University.

## References

- Bradford, J. M., Farrel, D. A., & Larson, W. E. (1973). Mathematical evaluation of factors affecting gully stability. *Soil Science Society of America*, 37.
- Bradford, J. M., and Piest, R. F. (1977). Gully Wall Stability in Loess-derived Alluvium1. *Soil Science Society of America Journal*, 41(1), 115–122.
- Bradford, J. M., & Piest, R. F. (1980). Erodional Development of Valley-Bottom Gullies in the Upper Midwestern United States. In D.R. Coates & J. D. Vitek (Ed.), *Geomorphic Thresholds* (pp. 75–101). Stroudsburg: Dowden&Culver.
- Campo-Bescós, M. a., Flores-Cervantes, J. H., Bras, R. L., Casalí, J., & Giráldez, J. V. (2013). Evaluation of a gully headcut retreat model using multitemporal aerial photographs and digital elevation models. *Journal of Geophysical Research: Earth Surface*, 118(4), 2159–2173. doi:10.1002/jgrf.20147
- Chaplot, V. (2013). Impact of terrain attributes, parent material and soil types on gully erosion. *Geomorphology*, 186, 1–11. doi:10.1016/j.geomorph.2012.10.031
- Chen, S.-C., Chang, K.-T., Wang, S.-H., & Lin, J.-Y. (2010). The efficiency of artificial materials used for erosion control on steep slopes. *Environmental Earth Sciences*, 62(1), 197–206. doi:10.1007/s12665-010-0514-6
- De Baets, S., Poesen, J., Meersmans, J., & Serlet, L. (2011). Cover crops and their erosion-reducing effects during concentrated flow erosion. *Catena*, 85(3), 237–244. doi:10.1016/j.catena.2011.01.009
- Di Stefano, C., Ferro, V., and Porto, P. (2000). Length Slope Factors for applying the Revised Universal Soil Loss Equation at Basin Scale in Southern Italy. *Journal of Agricultural Engineering Research*, 75(4), 349–364. doi:10.1006/jaer.1999.0514
- Evans, J. S., & Hudak, A. T. (2007). A Multiscale Curvature Algorithm for Classifying Discrete Return LiDAR in Forested Environments. *IEEE Transactions on Geoscience and Remote Sensing*, 45(4), 1029–1038. doi:10.1109/TGRS.2006.890412
- Galzki, J. C., Birr, a. S., & Mulla, D. J. (2011). Identifying critical agricultural areas with three-meter LiDAR elevation data for precision conservation. *Journal of Soil and Water Conservation*, 66(6), 423–430. doi:10.2489/jswc.66.6.423
- Ghilani, C. D., & Wolf, P. R. (2006). *Adjustment computation - Spatial data analysis*. (I. John Willey & Sons, Ed.) (4th Editio., p. 611). Hokoben, New Jersey.

- Glenn, N. F., Streutker, D. R., Chadwick, D. J., Thackray, G. D., & Dorsch, S. J. (2006). Analysis of LiDAR-derived topographic information for characterizing and differentiating landslide morphology and activity. *Geomorphology*, 73(1-2), 131–148. doi:10.1016/j.geomorph.2005.07.006
- Hoek, E., and Bray, J. W. (1981). *Rock Slope Engineering* (Revised Th.). The institution of mining and metallurgy.
- Istanbulluoglu, E. (2005). Implications of bank failures and fluvial erosion for gully development: Field observations and modeling. *Journal of Geophysical Research*, 110(F1), F01014. doi:10.1029/2004JF000145
- Kdrilen, D. L., Kramer, L. A., James, D. E., Buhler, D. D., & Burkart, M. R. (1964). Field-scale watershed evaluations on deep-loess soils:
- Lohnes, R. A., & Handy, R. L. (1968). Slope Angles in Friable Loess. *The Journal of Geology*, 76(3), 247–258.
- Molina, A., Govers, G., Cisneros, F., & Vanacker, V. (2009). Vegetation and topographic controls on sediment deposition and storage on gully beds in a degraded mountain area. *Earth Surface Processes and Landforms*, 767, 755–767. doi:10.1002/esp
- Montgomery, D. R., & Dietrich, W. E. (1988). © 198 8 Nature Publishing Group. *Nature*, 336(17), 232–234.
- Montgomery, D. R., & Dietrich, W. E. (1992). Channel initiation and the problem of landscape scale. *Science (New York, N.Y.)*, 255(5046), 826–30. doi:10.1126/science.255.5046.826
- Oostwoud Wijdenes, D. J., and Bryan, R. (2001). Gully-head erosion processes on a semi-arid valley floor in Kenya: a case study into temporal variation and sediment budgeting. *Earth Surface Processes and Landforms*, 26(9), 911–933. doi:10.1002/esp.225
- Poel, P. Van Der, Spomer, R. G., & Piest, R. F. (1986). Slope indicator measurements of subsurface movement in Gully walls. *Transactions of the ASAE*, 29, 982–987.
- Rengers, F. K., and Tucker, G. E. (2014). Analysis and modeling of gully headcut dynamics , North American high plains. *Journal of Geophysical Research: Earth Surface*, 1–21. doi:10.1002/2013JF002962.Received
- Shit, P. K., and Maiti, R. K. (2012). Mechanism of gully-head retreat - A study at Ganganir Danga, Paschim Medinipur, West Bengal. *Ethiopian Journal of Environmental Studies and Management*, 5(4), 332–342.

- Simon, A., and Box, P. O. (2006). Study of the effects of lateral seepage forces on tension-crack development, bank-failure dimensions and migration of edge of field gullies. In *Proceedings of the Eighth Federal Interagency Sedimentation Conference* (pp. 660–666). Reno.
- Skempton, A. W. and Hutchinson, J. N. (1969). Stability of natural slopes and embankment foundations. In *Proc. 7th Int. Conf. on Soil Mechanics. State of the art report.* (pp. 291–340).
- Thiebes, B., Bell, R., Glade, T., Jäger, S., Anderson, M., & Holcombe, L. (2013). A WebGIS decision-support system for slope stability based on limit-equilibrium modelling. *Engineering Geology*, 158, 109–118. doi:10.1016/j.enggeo.2013.03.004
- Thomas, J. T., Iverson, N. R., & Burkart, M. R. (2009). Bank-collapse processes in a valley-bottom gully. *Earth Surface Processes and Landforms*, (34), 109–122. doi:10.1002/esp
- Thomas, J. T., Iverson, N. R., Burkart, M. R., and Kramer, L. a. (2004). Long-term growth of a valley-bottom gully, western Iowa. *Earth Surface Processes and Landforms*, 29(8), 995–1009. doi:10.1002/esp.1084
- Tomer, M. D., Cambardella, C. A., James, D. E., and Moorman, T. B. (2006). Surface-Soil Properties and Water Contents across Two Watersheds with Contrasting Tillage Histories. *Soil Science Society of America Journal*, 70(2), 620. doi:10.2136/sssaj2004.0355
- Tomer, M. D., Meek, D. W., and Kramer, L. a. (2005). Agricultural practices influence flow regimes of headwater streams in western Iowa. *Journal of Environmental Quality*, 34(5), 1547–58. doi:10.2134/jeq2004.0199
- Tomer, M. D., Moorman, T. B., Kovar, J. L., James, D. E., and Burkart, M. R. (2007). Spatial patterns of sediment and phosphorus in a riparian buffer in western Iowa. *Journal of Soil and Water Conservation*, 62(5), 329–338.

## CHAPTER 4. CONCLUSIONS

In Chapter 2 a methodology was proposed to map classical gullies based on data mining of hydraulic and topographic parameters derived from high-resolution elevation data (LiDAR). An unsupervised classification termed CLARA (Large databases applications) was employed to group different topographic areas in the landscape. Since the clustering technique capture mostly the converging flow in gully edges we proposed a simple threshold approach to fuse the flow patch in the gully floor and map the full feature.

The methodology was validated with field work and shows that the approach to delineate concentrated gully flow (GCF) in the gully floor was consistent with observed data. Investigating the predictive power of supervised classifiers a few were performed and among linear discriminant analysis and logistic regression, the regression classification tree showed the larger overall classification accuracy (Acc = 99.99%) with the highest sensitivity and specificity (99.54% and 99.99 %) respectively

Chapter 3 evaluated the limit-equilibrium approach in topographic data of two multi-temporal surveys carried out in a western Iowa gullied area considering precision conservation purpose. The study computed spatial failures (Factor of safety indices) at gully head and gully walls of digital elevation models in surveys conducted in 1999 and 2014. Outcomes of this assessment has shown significance less instability of the actual slopes compared to 1999 survey slopes. Internal friction angle ( $\theta$ ) had the largest effect on slope stability factor (S.D.<sub>1999</sub> = 0.18, S.D.<sub>2014</sub> = 0.24), according the sensitivity analysis, compared to variations of soil cohesion, failure plane angle and slab thickness. In addition, critical instable slopes within gully, based on units of the slope standard deviation as threshold have produced area of 61 m<sup>2</sup> and 396 m<sup>2</sup> considering the threshold of one and two slope



standard deviation, respectively. These critical areas were located majority near the headcut and in the border of side walls. Based on current literature, association of processed material (geotextile) and crop cover with high root density might be an alternative to improve slope instability, but empirical tests are necessary to validate this approach. An intermediate outcome described in this chapter was a good performance of least squares technique employed to transform local coordinate systems in actual WGS83 datum. The mean error was 2.14% between original length survey and transformed ones. However, the slope instability must include other factors that capture the dynamics of failure in the headcut and side wall due to infiltration and seepage process and climate impacts.

Lastly, even though, this study have only performed data mining with functions available in R package is clear that there is a need for algorithm that can better handle big data set. For instance, unsupervised or supervised classifier that deal with proximity matrix (e.g. Partitioning around Medoids – *PAM*, random forest) require a huge amount of RAM memory.

## APPENDIX A. RANKING METHOD

```
## original Ranking method.R
## Article 1 - R script
## Classification script applied to gullies in Ames,IA study area – Besides Large application
## clustering (CLARA) there are other clustering techniques such as K-means
## and Fuzzy clustering. Also, WSS statistics to search for adequate cluster number.
## by Laurimar G. Vendrusculo
## ABE, ISU - 01/20/2014 – Last version
```

```
library("cluster")
require(cluster)
```

```
## Reading the original csv files from ArcGis with coordinates and variable values
twi    <- read.csv(file="C:/sk_idf_twi.txt",header = TRUE, sep = ",")
spi    <- read.csv(file="C:/sk_idf_spi.txt",header = TRUE, sep = ",")
curvprf <- read.csv(file="C:/sk_idf_prcur.txt",header = TRUE, sep = ",")
slop2  <- read.csv(file="C:/sk_idf_slope.txt",header = TRUE, sep = ",")
aspect <- read.csv(file="C:/sk_idf_aspec.txt",header = TRUE, sep = ",")
intens <-read.csv(file="C:/sk_intensity_join.txt",header = TRUE, sep = ",")
flowacc <-read.csv(file="C:/sk_idf_facp.txt",header = TRUE, sep = ",")
elevR  <-read.csv(file="C:/sk_idf_elev.txt",header = TRUE, sep = ",")
```

```
## Reading the first five rows and dimensions of dataset (# cols and # rows)
head(inten) ;dim(inten)
head(curv) ;dim(curv)
head(slop2) ;dim(slop2)
head(spi) ;dim(spi)
head(twi) ;dim(twi)
head(aspect) ;dim(aspect)
head(flowacc);dim(flowacc)
head(elevR); dim(elevR)
```

```
sort.elev <- elevR[order(elevR$POINT_X),]
head(sort.elev)
# FID POINTID GRID_CODE POINT_X POINT_Y
#2577 2576 2577 279.311 449165.9 4655710
#3401 3400 3401 279.560 449165.9 4655709
#4225 4224 4225 279.637 449165.9 4655708
#5049 5048 5049 279.833 449165.9 4655707
#5873 5872 5873 279.875 449165.9 4655706
```

```
### timing function in R
tic()
for (i in 1:50) mad(stats::runif(500))
```

```
toc()
```

```
plot(sort.elev$POINT_X[1:4000],sort.elev$POINT_Y[1:4000])
```

```
## Pre-processing dataset selecting only the column needed
## Select only the variable values and creating a new dataset with all of them (all_ar1)
inten1 <- intens[,11]; head(inten1)
curv1 <- curvprf[,3]; head(curv1)
slope1 <- slop2[,3]; head(slope1)
spi1 <- spi[,3]; head(spi1)
twi1 <- twi[,3]; head(twi1)
aspect1 <- aspect[,3]; head(aspect1)
flowacc1 <- flowacc[,3]; head(flowacc1)
elev <- elevR[,3]; head(elev)
```

```
aspect2 <- aspect1
```

```
head(aspect2)
```

```
## recoding aspect to only four values (0 = nothing, 1 = North, 2= East, 3= South and 4=
West)
```

```
## 315 - 360 -- 0 to 45 is equal to 1
```

```
## 46 to 135 is equal to 2
```

```
## 136 to 225 is equal to 3
```

```
## 226 to 314 is equal to 4
```

```
aspect2[aspect1 >= 315 & aspect1 <= 360] <- 1
```

```
aspect2[aspect2 == 315] <- 1
```

```
aspect2[aspect1 >= 0 & aspect1 <= 45.99999] <- 1
```

```
aspect2[aspect1 >= 46 & aspect1 <= 135.99999] <- 2
```

```
aspect2[aspect1 >= 136 & aspect1 <= 225.99999] <- 3
```

```
aspect2[aspect1 >= 226 & aspect1 <= 314.99999] <- 4
```

```
aspect2[aspect1 <= 0 ] <- 0 ## flat areas
```

```
head(aspect2)
```

```
summary(aspect2)
```

```
hist(aspect2)
```

```
## asp4 <- subset(aspect2, aspect1 >= 226 & aspect1 <= 314.999)
```

```
## asp5 <- subset(aspect2, aspect1 >= 314.999 & aspect1 <= 315 )
```

```
## Joining all the previous attributes (curvature, slope, aspect, spi, twi, intensity and flow
accumulation
```

```
all_ar1 <- cbind(curv1,slope1,aspect1,spi1,twi1,inten1,flowacc1,elev);
```

```
head(all_ar1)
```

```
summary(all_ar1)
```

```
dim(all_ar1)
```

```

#### Deviation of mean slope
#### population standard deviation = sqrt(Somatory (X - xmean)^2)/Nobserv)
mean(slope1)
## Slope: Mean : 6.1468
##Deviation_from_mean = X - mean(slope1)
Dev_meanSl = slope1 - mean(slope1)
head(Dev_meanSl)

all_ar1 <- cbind(curv1,slope1,aspect1,spi1,twi1,inten1, flowacc1,Dev_meanSl,aspect2);
head(all_ar1)
summary(all_ar1)

hist(spi1, data=all_ar1)
hist(Dev_meanSl, data=all_ar1)
quartile(spi1, data=all_ar1)

#### Box plot of the whole set
##final version of the figures
par(mfrow=c(1,2))
boxplot(spi1,data=all_ar1, main="", xlab="(a) Stream power index", ylab="SPI")
boxplot(aspect1,data=all_ar1, main="", xlab="(b) Coded aspect", ylab="Aspect")
boxplot(Dev_meanSl,data=all_ar1, main="", xlab="(c) Deviation of mean slope",
ylab="Slope")
boxplot(curv1,data=all_ar1, main="", xlab="(d) Profile curvature", ylab="Profile
Curvature")

#### CLARA clustering using cluster package - seed set up for outcomes reproducibility
purpose
library("cluster")
require(cluster)
## run CLARA
set.seed(110)

## 1      2      3      4      5      6      7      8      9
## curv1, slope1, aspect1, spi1, twi1, inten1, flowacc1,Dev_meanSl,aspect2
cluN = 6 # number of clusters
## Clustering using CLARA at the following variables: curvature, deviation mean
## slope, aspect and SPI
clarax1 <- clara(all_ar1[,c(1,8,9,4)] , cluN , metric = "euclidean",stand = FALSE, samples =
200, sampsize = min(5000, 40 + 2 * 8) )
Inis_clara <- as.data.frame(cbind(curvprf[,4:5],all_ar1[,c(1,8,9,4)],slope1,elev, cluster =
clarax1$cluster))
##%% renaming cluster columns
#colnames(Inis_clara) <- c("PointX","PointY","curvprf","slopeD","aspect2","spi","cluster")
colnames(Inis_clara) <-
c("PointX","PointY","curvprf","Dev_meanSl","aspect2","spi","slop","elev","cluster")

```

```

head(Inis_clara)
summary(Inis_clara)
summary(subset(Inis_clara, clarax1$cluster == 2),na.rm=TRUE)# example of summary by
each cluster

sd(subset(Inis_clara$curvprf, clarax1$cluster == 4),na.rm=TRUE)# example of summary
by eac
sd(subset(Inis_clara$Dev_meanSl, clarax1$cluster == 6),na.rm=TRUE)# example of
summary by eac
sd(subset(Inis_clara$aspect2, clarax1$cluster == 6),na.rm=TRUE)# example of summary
by eac
sd(subset(Inis_clara$spi, clarax1$cluster == 6),na.rm=TRUE)# example of summary by
eac

hist(Inis_clara$Dev_meanSl)
hist(Inis_clara$aspect1)
#write.csv(Inis_clara, file="C:/Users/laurimarvend/users/laurimar1/summer2012/skunk-
river_MCC/clustering/clarax1_skunk_16894_6.csv")

##final version of the figures
par(mfrow=c(2,2))
boxplot(spi~clarax1$cluster,data=Inis_clara, main="Stream Power Index", xlab="# Cluster",
ylab="SPI")
boxplot(aspect2~clarax1$cluster,data=Inis_clara, main="Coded aspect" , xlab="# Cluster",
ylab="Aspect")
boxplot(Dev_meanSl~clarax1$cluster,data=Inis_clara, main="Deviation of mean Slope"
,xlab="# Cluster", ylab="Slope")
boxplot(curvprf~clarax1$cluster,data=Inis_clara, main="Profile curvature" , xlab="#
Cluster", ylab="Curvature")
boxplot(elev~clarax1$cluster,data=Inis_clara, main="Elevation" , xlab="# Cluster",
ylab="Elevation")
box_dm <- boxplot(Dev_meanSl~clarax1$cluster,data=Inis_clara, main="Mean slope
deviation" , xlab="# Cluster", ylab="Elevation")

## multivariable analysis
Y <- cbind(curv1, Dev_meanSl, aspect2, spi1);head(Y)

summary(manova(Y ~ cluster , data = Inis_clara), test = "Hotelling-Lawley")
summary(manova(Y ~ cluster , data = Inis_clara), test = "Wilks")

#      Df Hotelling-Lawley approx F num Df den Df  Pr(>F)
#cluster  1      2.0343  464589   4 913489 < 2.2e-16 ***
#Residuals 913492

# The p--value associated with the Hotelling-Lawley statistic is less than 5%
# and there is strong evidence that #the mean vectors of the two variables

```

```
# are not the same in the clusters.
```

```
tapply(Inis_clara$spi, Inis_clara14$cluster, mean)
tapply(Inis_clara$aspect2, Inis_clara14$cluster, mean)
tapply(Inis_clara$Dev_meanSl, Inis_clara14$cluster, mean)
tapply(Inis_clara$curvprf, Inis_clara14$cluster, mean)
```

```
### Principal components
```

```
## Print cluster characteristics
```

```
clarax1$medoids; print(clarax1)
```

```
Inis_clara14$cluster14
```

```
## Distribution of points in each cluster 255463 + 241961 + 168235 + 247835
```

```
### best for 4 classes
```

```
best_sample <- data.frame(c(Inis_clara[c(
  29119, 29481, 76901, 78643, 91690, 139779, 161802, 188174, 217321, 235943,
  237643, 265577, 269187, 293385, 294263, 317568, 389688, 389827, 398999, 418848,
  437651, 446349, 450656, 455967, 484026, 487901, 489127, 499735, 507791, 541259,
  548130, 572969, 594351, 597125, 629422, 653926, 657759, 658275, 660143, 680256,
  727467, 744863, 751846, 752097, 754745, 781159, 809399, 835646, 856359, 870159,
  895444, 899695, 902915, 906191, 907501, 911794),]))
```

```
head(best_sample);dim(best_sample)
```

```
## Plot the scatter plot of all variables coded by cluster
```

```
plot(best_sample[,c(1,2,3,4,6)],col=best_sample$cluster)
```

```
points(best_sample$cluster, col = 1:4, pch = 10, cex = 2)
```

```
## Plot 2 Principal components of best samples data
```

```
clusplot(best_sample,best_sample$cluster,color=TRUE, shade=TRUE, labels=2, lines=0)
color=TRUE
```

```
## Plot 2 Principal components of original data
```

```
clusplot(all_ar1[,c(1,2,3,4,6)], fit$cluster, adiss = TRUE,
col=(c("yellow", "darkgreen", "red", "blue", "pink", "gray", "darkblue", "orange")),
shade=TRUE, labels=2, lines=0)
```

```
##### EXTRACTING SUMMARY of classical statistic in the clustered dataset
```

```
summary(subset(Inis_clara, clarax1$cluster == 1),na.rm=TRUE)
```

```
summary(subset(Inis_clara, clarax1$cluster == 2),na.rm=TRUE)
```

```
summary(subset(Inis_clara, clarax1$cluster == 3),na.rm=TRUE)
```

```
summary(subset(Inis_clara, clarax1$cluster == 4),na.rm=TRUE)
```

```
summary(subset(Inis_clara, clarax1$cluster == 5),na.rm=TRUE)
```

```
summary(subset(Inis_clara, clarax1$cluster == 6),na.rm=TRUE)
```

## APPENDIX B. GULLY LOCATION DATASET

```

## Exploratory task of gully location using supervised methods
## Article 1 - R script
## Exploratory task of gully location using supervised methods applied to gullies at the
Study are in Ames,IA
## Linear discriminant analysis (MASS package);logistic regression (nnet package), ##
regression tree (rpart Package)
## by Laurimar G. Vendrusculo
## ABE, ISU – 08/17/2014 – Last version

# Routine to reading the gully dataset
#### GPS paths transformed in points 1 meter apart
gul_gps3 <-read.csv(file="C:/gully_slope_aspec.txt ",header = TRUE, sep = ",")
gul_gps4 <-read.csv(file="C:/ gully_spi.txt",header = TRUE, sep = ",")
gul_gps5 <-read.csv(file="C:/ gully_curvat.txt",header = TRUE, sep = ",")

head(gul_gps3);dim(gul_gps3) ## Slope and aspect
head(gul_gps4);dim(gul_gps4) ## SPI
head(gul_gps5);dim(gul_gps5) ## profile curvature

# Worked variables
#aspect - profile curvature - mean slope deviation - SPI
##slope, aspect,spi,curv
gul_gps_red = cbind(gul_gps3[,c(7,8,10,14)],gul_gps4[,c(10)],gul_gps5[,c(10)])
head(gul_gps_red)
colnames(gul_gps_red) = c("PointX", "PointY", "slope", "asp", "spi", "curv")
summary(gul_gps_red)

#### Deviation of MEAN
#### population standard deviation = sqrt(Somatory (X - xmean)^2)/Nobserv)
mean(slope1)
## Slope: Mean : 6.1468
##Deviation_from_mean = X - 6.1468

Dev_meanSIG = gul_gps_red[,3] - 6.1468
head(Dev_meanSIG)

## recoding aspect to only four values (0 = nothing, 1 = North, 2= East, 3= South and 4=
West)
## 315 - 360 -- 0 to 45 is equal to 1
## 46 to 135 is equal to 2
## 136 to 225 is equal to 3
## 226 to 314 is equal to 4

aspect11 <- gul_gps_red[,4]

```

```

aspect22 <- gul_gps_red[,4]
aspect22[aspect11 >= 315 & aspect11 <= 360] <- 1
aspect22[aspect11 == 315] <- 1
aspect22[aspect11 >= 0 & aspect11 <= 45.99999] <- 1
aspect22[aspect11 >= 46 & aspect11 <= 135.99999] <- 2
aspect22[aspect11 >= 136 & aspect11 <= 225.99999] <- 3
aspect22[aspect11 >= 226 & aspect11 <= 314.99999] <- 4
aspect22[aspect11 <= 0 ] <- 0 ## flat areas

head(aspect22);dim(aspect22)
summary(aspect22)
hist(aspect22)

##slope, aspect,spi,curv,cluster1 means gully bottom
gul_gps1 =
cbind(gul_gps5[,c(10)],Dev_meanSIG,aspect22,gul_gps4[,c(10)],rep(1,length(aspect22)))
head(gul_gps1);dim(gul_gps1)
colnames(gul_gps1) = c("curv","slopeD","asp","spi","cl")
summary(gul_gps1)

## all_arR comes from appendix A script
gull_tot = rbind(all_arR, gul_gps1)
head(gull_tot)
dim(gull_tot)

### Sampling processing - Taking 30% of each class at training phase and leaving
### 70% to test

set.seed(503) ## set for outcomes reproducibility

cl1 = gull_tot[which(gull_tot[,5]=="1"),]; dim(cl1)
cl2 = gull_tot[which(gull_tot[,5]=="2"),]; dim(cl2);

indx.cl1 = sample(seq_len(nrow(cl1)),size = floor(0.3*2279))
indx.cl2 = sample(seq_len(nrow(cl2)),size = floor(0.3*913494))

train_cl1 = cl1[indx.cl1,] ;dim(train_cl1)
test_cl1 = cl1[-indx.cl1,] ;dim(test_cl1)

train_cl2 = cl2[indx.cl2,] ;dim(train_cl2)
test_cl2 = cl2[-indx.cl2,] ;dim(test_cl2)

train_tot = data.frame(rbind(train_cl1,train_cl2))
test_tot = data.frame(rbind(test_cl1,test_cl2))
dim(test_tot)[1]
summary(train_tot)

```



```

summary(test_tot)

# Plotting the histogram of the test and train sets
hist(test_tot[,5], border="red")
table(test_tot[,5])
hist(train_tot[,5], border="red")
table(train_tot[,5])

# Linear discriminant analysis (LDA)
library(MASS)
require(MASS)

# LDA applied in train and test sets
gull.ldat<-lda(V5~.,train_tot)
table(train_tot[,5],predict(gull.ldat,train_tot)$class)
#      1      2
# 1    11   672
# 2   393 273655
par(mfrow=c(1,2))
plot(predict(gull.ldat,train_tot,dimen=2)$x, pch=as.character(train_tot[,5]))
title("True class")
plot(predict(gull.ldat,train_tot,dimen=2)$x,pch=as.character(predict(gull.ldat,train_tot,dimen
=2)$class))
title("Predicted class")

gull.ldatest<-lda(V5~.,test_tot) # the variable 6 has equal values (fiber)
table(test_tot[,5],predict(gull.ldatest,test_tot)$class)
#      1      2
# 1     6  1590
# 2   376 639070
(6*100)/(6+376) # 1.570681 %
(639070*100)/(1590+639070) #99.75182%

# Overall error - Error
z<-table(test_tot[,5],predict(gull.ldatest,test_tot)$class)
((1-sum(diag(z)))/dim(test_tot)[1])*-100 # 99.69316

# Appling Test and Train set for logistic regression
library(nnet)
gull_totf = data.frame(gull_tot)
gull.log<-multinom(V5~.,data=gull_totf)
table(gull_tot[,5],predict(gull_tot[,-5]))
plot(gull.log)

gull.logtr<-multinom(V5~.,train_tot)
table(train_tot[,5],predict(gull.logtr,train_tot, type="class"))

```

```

#      1      2
# 1    2    681
# 2   65 273983

gull.logtest<-multinom(V5~.,test_tot)
table(test_tot[,5],predict(gull.logtest,test_tot, type="class"))
#      1      2
# 1    7   1589
# 2   157 639289

Sensitivity= 7/(7+1589)*100 # 0.4385965 %
Specificity= 639289/(1590+639070) *100 # 99.78

(7*100)/(7+157) #4.268293
(639289*100)/(1589+639289) #99.75206

# Overall error
z<-table(test_tot[,5],predict(gull.logtr,test_tot, type="class"))
((1-sum(diag(z)))/dim(test_tot)[1])*-100 #99.72779

# Appling Test and Train set for Classification Tree

# Using train dataset - tree
gull.rptr <- rpart(V5~., data=train_tot, method="class", parms=list(split=' information'))
gull.rptr

# table for the train - tree
table(train_tot[,5],predict(gull.rptr, type="class"))
#      1      2
# 1   592    91
# 2    24 274024
sensitivity (true positive - total number of observations of classe 1

# table for the test - tree
gull.rpts <- rpart(V5~., data=test_tot, method="class", parms=list(split=' information'))
table(test_tot[,5],predict(gull.rpts,test_tot,type="class"))
#      1      2
# 1  1326   270
# 2    6 639440
sensitivity (true positive - total number of observations of classe 1
Sensitivity= 1326 /(1326+6)*100 # 99.54
Specificity= 639289/(270+639070) *100 # 99.99

# Overall error - Error
z<-table(test_tot[,5],predict(gull.rpts ,test_tot, type="class"))

```

```

((1-sum(diag(z)))/dim(test_tot)[1])*-100 # 99.95679%

p-value for the values of sensitivity and specificity
specify = c(99.97,99.78,99.99)
t.test(specify)

sensity = c(1.57, 0.44,99.54)
t.test(sensity)

# Random Forest
library(randomForest)
require(randomForest)
### Require a vector of size 1 GB
gull.rf <- randomForest(factor(V5) ~ ., data=train_tot,importance=TRUE,
proximity=TRUE,mtry=3)
print(gull.rf)

gull.rf$importance
# importance attribute can be used to construct parallel coordinate and 2D tour

head(gull.rf$votes)

# table for the train - tree
gull.rfTr <- randomForest(factor(V5) ~ ., data=train_tot,importance=TRUE,
proximity=FALSE,mtry=2)
table(train_tot[,5],predict(gull.rfTr , type="class"))

# table for the test - tree
gull.rfTs <- randomForest(factor(V5) ~ ., data=train_tot,importance=TRUE,
proximity=TRUE,mtry=3)
table(test_tot[,5],predict(gull.rfTs ,test_tot,type="class"))

# Overall error - Random Forest
z<-table(test_tot[,5],predict(gull.rf,test_tot, type="class")) ## error
1-sum(diag(z))/dim(test_tot)[1] # 9.03

```

## APPENDIX C. PLAIN FAILURE FOR GULLY SIDEWALLS

```

## Article 2 – R script
## Plain failure applied to gully top sidewalls - Equations from Hoek & Bray book
## Reference: Hoek, E., and J. W. Bray. 1981. Rock Slope Engineering. Revised 4 th.
## by Laurimar G. Vendrusculo
#### ABE, ISU - 08/04/2014
##### Treynor - Entire gully area

## Reading input data: Elevation and slope from LiDAR data preprocessed in ArcGis 10.1
survey2014 <-
read.csv(file="C:/Users/laurimarvend/users/laurimar1/summer2012/treynor_site_MCC/survey2014/slab_height2014.txt",header = TRUE, sep = ",")
head(survey2014 ); dim(survey2014 )
#hist(survey2014$GRID_CODE)
sortS2014 <- sort(survey2014$GRID_CODE);sortS2014
head(survey2014)

## Focal analysis obtained the elevation around a pixel of 2x2
slope2014 <-
read.csv(file="C:/Users/laurimarvend/users/laurimar1/summer2012/treynor_site_MCC/survey2014/slope2014.txt",header = TRUE, sep = ",")
head(slope2014); dim(slope2014)
#hist(slope2014$GRID_CODE)
sortE2014 <- sort(slope2014$GRID_CODE);sortE2014
head(slope2014)

survey_2014 <- cbind(survey2014[,c(4,5,3)],slope2014[,3]); head(survey_2014)
colnames(survey_2014) <- c("PointX","PointY","elev","slope");
dim(survey_2014)

##### Treynor boundaries conditions according Thomas et al. (2009) work #####
x      = 300/1000 ;x ## meters Slab thickness
rad    = (3.1415/180); rad
phiF   = survey_2014$slope*rad; phiF
phiP   = (45 + (22.15/2))*rad;phiP
tetha  = 22.15*rad;tetha ## angle of friction
cotphiF = 1/tan(phiF);cotphiF
cotphiP = 1/tan(phiP);cotphiP
sinphiP = sin(phiP);sinphiP
cosphiP = cos(phiP);cosphiP
cosecphiP = 1/sin(phiP);cosecphiP
tanphiP = tan(phiP);tanphiP

H = survey_2014$elev;H

```

```

#### Values in SI
C      = 7.9 # kPa Cohesion
Z      = (H - x*tan(tetha));Z
#Zw    = (Zcrit - h + Z) ,m depth of water filled tension crack - worst scenario - From Eq. 7
from Thomas (2004)
Zw     = (3.54 - 5 + Z);Zw
yw     = 9.81 # KN/m3 unit weight of the water N/m3
y      = 13 # KN/m3 unit weight of the slab dry N/m3
y      = 15.5 # KN/m3 unit weight of the slab with 50% water content

W = (1/2)*y*H^2 * ( (1-(Z/H)^2)*cotphiP - cotphiF ); W
A = (H - Z)*cosecphiP;A
U = (1/2)*yw*Zw *A; U
V = (1/2)*yw*(Zw^2); V

## Estimation of Factor of safety
F = ((C*A) + ((W*cosphiP - U - V*sinphiP)* tan(tetha))) / (W*sinphiP + V*cosphiP) ;F

hist(F)
## Seaching for a FS occurrence greater than 1
GreatFS <- F[which(F >= 1)];GreatFS
SmallFS <- F[which(F < 0)];SmallFS
#### Rearranging the equations in a dimensionless

P = (1-Z/H)*cosecphiP ;P
R = (yw/y)*(Zw/Z)*(Z/H) ; R
S = (Zw/Z)*(Z/H)*sinphiP ; S
Q = ((1- (Z/H)^2)*cotphiP - cotphiF)*sinphiP ;Q

## Factor of safety
FS.dim = (((2*C)/(y*H))*P + (Q*cotphiP - R*(P+S))*tan(tetha) )/(Q + R*S*cotphiP);F

GreatFS.dim <- F[which(F >= 1)];GreatFS.dim
boxplot(F)

#### Recoding FS values between zero to 1.4 value
FS.rec = FS.dim;FS.rec
FS.rec = F
FS.rec[FS.rec<=0] = 0.0
FS.rec[FS.rec>0 & FS.rec<=0.1] = 0.1
FS.rec[FS.rec>0.1 & FS.rec<=0.2] = 0.2
FS.rec[FS.rec>0.2 & FS.rec<=0.3] = 0.3
FS.rec[FS.rec>0.3 & FS.rec<=0.4] = 0.4
FS.rec[FS.rec>0.4 & FS.rec<=0.5] = 0.5
FS.rec[FS.rec>0.5 & FS.rec<=0.6] = 0.6

```

```

FS.rec[FS.rec>0.6 & FS.rec<=0.7] = 0.7
FS.rec[FS.rec>0.7 & FS.rec<=0.8] = 0.8
FS.rec[FS.rec>0.8 & FS.rec<=0.9] = 0.9
FS.rec[FS.rec>0.9 & FS.rec<=1] = 1
FS.rec[FS.rec>1 & FS.rec<=1.1] = 1.1
FS.rec[FS.rec>1.1 & FS.rec<=1.2] = 1.2
FS.rec[FS.rec>1.2 & FS.rec<=1.3] = 1.3
FS.rec[FS.rec>1.3 ] = 1.4;FS.rec

```

```

survey2014FS.dim <- cbind(survey_2014,F,FS.rec); head(survey2014FS.dim)
colnames(survey2014FS.dim ) <- c("PointX","PointY","elev","slope","FS","FSC");
hist(survey2014FS.dim$FSC)

```

```
##### Plotting the associated FS in a Easting and Northing
```

```

h <- qplot( PointX, PointY, data=survey2014FS.dim, geom="point",colour=factor(FS.rec),
xlab="UTM Easting, (meters) ", ylab=" UTM Northing, (meters) ", labels=FS.rec, main = " -
1999 Survey " )
h + scale_colour_manual( values = c("0.0" = "black","0.1" = "red","0.2" = "red","0.3" =
"red","0.4" = "red","0.5" = "red","0.6" = "blue","0.7" = "darkorchid","0.8" = "black","0.9" =
"darkolivegreen4","1" = "yellow","1.1" = "orange","1.2" = "darksalmon","1.3" =
"sienna","1.4" = "lawngreen","1.5" = "lawngreen" ))

```

```
###
```

```
par(mfrow=c(3,2))
```

```

FS.lzero <- survey2014FS.dim[which(survey2014FS.dim[, 'FSC'] == 0), ]; FS.lzero
dim(FS.lzero)
# mean(FS.lzero$elev)
#[1] 0.9223632
# mean(FS.lzero$slope)
#[1] 29.02021
hist(FS.lzero$elev, breaks=50,xlab="Slab height (m)", main="Factor of safety < 0")
hist(FS.lzero$slope,breaks=50,xlab="Slab slope (degree)", main="Factor of safety < 0")

```

```

FS.one =survey2014FS.dim[which(survey2014FS.dim[, 'FSC'] >0
&survey2014FS.dim[, 'FSC'] <=1 ), ]; head(FS.one)
dim(FS.one); max(FS.one$elev);max(FS.one$slope);sd(FS.one$FS);
sd(FS.one$slope)#9.956563
# y= 13
#[1] 1977 6
#[1] 3.041016
#[1] 57.61907
## 57.61907 - 9.956563 = 47.66251 One standard deviation with y - 15.5
### 57.61907 - 2*9.956563 = 37.70594 two standard deviation with y - 15.5

```

```
# y=15.5
```

```

[1] 2126 6
## (2126*100)/1977 - 100 = -7.5

# max(FS.one$slope)
#[1] 57.61907
# max(FS.one$elev)
#[1] 3.041016
# mean(FS.one$elev)
#[1] 0.9223632
# mean(FS.one$slope)
#[1] 29.02021

hist(FS.one$elev, breaks=50,xlab="Slab height (m)", main="0 <FS < 1 - Instable slopes ")
hist(FS.one$slope,breaks=50,xlab="Slab slope (degree)", main="0 <FS < 1 - Instable
slopes")

FS.Gone <- survey2014FS.dim[which(survey2014FS.dim['FSC'] > 1), ]; FS.Gone
dim(FS.Gone); max(FS.Gone$elev);max(FS.Gone$slope)

# max(FS.Gone$slope)
#[1] 51.60528
# max(FS.Gone$elev)
#[1] 3.041016
# mean(FS.Gone$elev)
#[1] 0.9223632
# mean(FS.Gone$slope)
#[1] 29.02021

hist(FS.Gone$elev, breaks=50,xlab="Slab height (m)", main="Factor of safety > 1 - Stable
slopes")
hist(FS.Gone$slope,breaks=50,xlab="Slab slope (degree)", main="Factor of safety >1 -
Stable slopes")

plot(ecdf(survey2014FS.dim['FSC']))
hist(survey2014FS.dim['FSC'])
plot.ecdf(survey2014FS.dim['FSC'],datadensity="density")

## Export the Factor safety computed and information associated
write.csv(survey2014FS.dim, file="C:/survey2014FS15.csv")

## Cumulative histogram

## Plot the density and cumulative density
d <- density(na.omit(survey2014FS.dim['FSC']))
lines(x = d$x, y = d$y * length(survey2014FS.dim) * diff(h$breaks)[1], lwd = 2)
lines(x = d$x, y = cumsum(d$y)/max(cumsum(d$y)) * length(x), lwd = 2)

```

```

FSC.sat2014 = survey2014FS.dim[, 'FSC']; head(FSC.sat2014)
FSC.dry2014 = survey2014FS.dim[, 'FSC']

## Cumulative density line
library(Hmisc)
require(Hmisc)

Ecdf(FSC.sat2014, xlab="Factor of Safety - LiDAR-based DEM", ylab="Cumulative Area (%)", col="black")
par(new=TRUE)
Ecdf(FSC.dry2014, xlab="", ylab="", col="dark grey")

hist(survey2014$GRID_CODE, xlab="Factor of Safety - LiDAR-based DEM", ylab="Cumulative Area (%)", col="grey", title="")
par(new=TRUE)
Ecdf(survey2014$GRID_CODE, xlab="", ylab="", col="black")

hist(FSC.sat2014, xlab="Factor of Safety - LiDAR-based DEM", ylab="Cumulative Area (%)", col="grey", title="")
par(new=TRUE)
Ecdf(FSC.sat2014, xlab="", ylab="", col="black")

## Sensitivity analysis

# The point with highest slope placed in the 2014 perimeter (x=278325.7197
#y=4560571.001, # elevation difference: 2.15, slope=57 o, FS = 0.26) was chosen to
investigate #influence # of shear stress and drive force due to rainfall.

survey2014FS.dim[891,]
#   PointX   PointY   elev   slope   FS     FSC
891  278325.7197 4560571.001 2.15402222 56.99679565 0.2610528502 0.3
## Friction angle 15 to 23 and cohesion varies from 7 to 13.8 KPa.

x      = 300/1000 ;x ## meters Slab thickness
rad    = (3.1415/180); rad
phiF   = 56.99679565*rad; phiF
tetha  = 22.15;tetha ## angle of friction
#phiP  = 70*rad
phiP   = (45 + (22.15/2))*rad;phiP
cotphiF = 1/tan(phiF);cotphiF
cotphiP = 1/tan(phiP);cotphiP
sinphiP = sin(phiP);sinphiP
cosphiP = cos(phiP);cosphiP
cosecphiP = 1/sin(phiP);cosecphiP

```



$\tan\phi_P = \tan(\phi_P); \tan\phi_P$

### Values in SI

$H = 2.15402222; H$

$C = 7.9$  # kPa Cohesion

$Z = (H - x \cdot \tan(\theta)); Z$

$Z_w = (Z_{crit} - h + Z)$ , m depth of water filled tension crack - worst scenario - From Eq. 7 from Thomas (2004)

$Z_w = (3.54 - 5 + Z); Z_w$

$\gamma_w = 9.81$  # KN/m<sup>3</sup> unit weight of the water N/m<sup>3</sup>

$\gamma_y = 13$  # KN/m<sup>3</sup> unit weight of the slab dry N/m<sup>3</sup>

$\gamma = 15.5$  # KN/m<sup>3</sup> unit weight of the slab with 50% water content

$W = (1/2) \cdot \gamma \cdot H^2 \cdot ((1 - (Z/H)^2) \cdot \cot\phi_P - \cot\phi_F)$ ; W

$A = (H - Z) \cdot \operatorname{cosec}\phi_P$ ; A

$U = (1/2) \cdot \gamma_w \cdot Z_w \cdot A$ ; U

$V = (1/2) \cdot \gamma \cdot (Z_w^2)$ ; V

## Factor of safety

$F = ((C \cdot A) + ((W \cdot \cos\phi_P - U - V \cdot \sin\phi_P) \cdot \tan(\theta))) / (W \cdot \sin\phi_P + V \cdot \cos\phi_P)$ ; F

## considering Saturated condition

Frict\_angle = c(15,16,17,18,19,20,21,22,23) # 9 intervals is theta

cohesion = c(7,7.75,8.5,9.26, 10,10.75,11.5, 12.3,13.8) # 9 intervals (C)

$\phi_{PV} = c(30,35,40,45,50,55,60,65,70)$  # failure plane angle  $(45 + \theta/2)$

FS\_VTetha =

c(0.2587968576,0.2762166607,0.2937045078,0.3112685116,0.3289167302,0.3466571569,0.3644977045,0.3824461859,0.4005102875)

FS\_VCohesion =

c(0.4114891518,0.4008302407,0.3901713295,0.3793702996,0.3688535072,0.3581945961,0.3475356849,0.3361661797,0.3148483574)

FS\_VphiP =

c(0.6767636311,0.5662985558,0.4763638426,0.4014429443,0.3374609933,0.2815299558,0.2315797578,0.1860878062,0.1438980464)

var(FS\_VTetha);var(FS\_VCohesion);var(FS\_VphiP)

sd(FS\_VTetha);sd(FS\_VCohesion);sd(FS\_VphiP)

## Student Test

t.test(FS\_VTetha)

t.test(FS\_VCohesion)

t.test(FS\_VphiP)

## Scatter plots with regression line and coefficient of determination (R2)

par(mfrow = c(1,3))

```
scatter2D(FS_VTetha,Frict_angle, bty = "g", col = "black", pch = 18, cex = 2, ylab="Internal  
Friction Angle(Degree)", xlab="Factor of Safety")  
abline(lm(Frict_angle~FS_VTetha),col="black",lty=2)
```

```
scatter2D(FS_VCohesion,cohesion, bty = "g", col = "black", pch = 18, cex = 2, ylab="Soil  
cohesion (KPa) ", xlab="Factor of Safety")  
abline(lm(cohesion~FS_VCohesion),col="black",lty=2)
```

```
scatter2D(FS_VphiP,phiPV, bty = "g", col = "black", pch = 18, cex = 2, ylab="failure plane  
angle(Degree)", xlab="Factor of Safety")  
abline(lm(phiPV~FS_VphiP),col="black",lty=2)
```