Response to the Comments of Reviewer 2

We greatly appreciate the reviewer's detailed comments and additional references that have helped us to improve our manuscript significantly. In the section below we repeat the comment and this is followed directly by our responses. The changed text is highlighted in red. For ease of reading we omitted the deleted text.

We apologize for the delay in submitting the response. The calculation of the uncertainty took much longer than anticipated.

REVIEWER 2

Comment 1, line 62:

But also, a change on the rate sometime is related with the rainfall erosivity doi:10.1016/j.still.2009.05.009

Response 1.

We put the above "doi" in Google and Science Citation Index, but we could not find the references. Assuming that "still" stand for "soil and tillage research" we found a tropical soil conference in Mekelle in Ethiopia but none of the articles seemed relevant. Partly, based on the comment we changed the beginning of the paragraph as follows:

> "The drainage area at the gully head is one of the parameters explaining linear, areal and volumetric gully headcut retreat (Vandekerckhove et al., 2003; Frankl et al., 2012, 2013, Vanmaercke et al., 2016). Runoff-contributing drainage area can be used as a surrogate for runoff, especially if it is assumed that the rainfall amount is equal for all drainage areas and that surface conditions and land use are also very similar (Oostwoud Wijdenes and Bryan, 2001). Frankl et al. (2012) reported that among all environmental characteristics in the catchment, only the drainage area had a strong positive association with gully headcut retreat, Hereafter, headcut retreat refers to the longitudinal growth and the bank failure refers to cross-sectional growth."

Comment 2, line 86:

average?

Response 2:

We updated the paragraph as follows

> "The study area, the Debre Mawi watershed, is located in the sub-humid highlands of northwest Ethiopia, 30 km south of Bahir Dar along the road to Adet, and lies between 11º20' and 11º22' N and 37º24' and 37º26' E. The watershed drains an area of 608 ha. The altitude ranges from 2186 to 2366 m (Fig.2).; the elevations of the gullies considered in this study range from 2212 to 2272 m. Rainfall is unimodal with an average value of 1240 mm yr^{-1}. Most rainfall falls between June and the beginning of September and

amounts 900 mm yr^{-1}). The rainfall gauge station in the Debre Mawi watershed has been established since 2008 by Adet Agricultural Research Center to record rainfall in rainy phase only. The dry season lasts between 8 to 9 months. The mean daily temperature is 20 °C"

Comment 3, line 129:

This measuring method has some concerns, please keep in mind for the discussion these relevant papers: doi:10.1016/j.catena.2006.03.005, and from this journal:http://www.soil-journal.net/1/509/2015/soil-1-509-2015.pdf

Response 3:

Apologies for the confusion. In the original manuscript we did not describe clearly in the Methods section the size of the gullies in the Debre Mawi watershed. Our gullies were 1 to 3 orders of magnitude larger that the gullies in the references cited above by Casili et al (2006, 2015). In our case with the larger gullies a cross section spacing of 3.6 m is reasonable to determine short-term linear, areal, and volumetric expansion over 10 to 30 meters. The paragraph is now reads as:

:

"During the 2013 and 2014 monsoon rain phases, we measured for 13 gullies: (1) the headcut retreat (longitudinal growth) and bank widening (or lateral retreat) for the first 10-30 m downslope from the headcut, and (2) the gully expansion rates and associated amount of soil loss along the total gully length.

To measure the headcut retreat and widening of the 13 gullies, the first 10 to 30 m of each gully was divided into 3 to 8 uniform segments. The average distance between two consecutive cross sections was 3.6 m and varied from 1 m to 10 m with a standard deviation of 2.7 m. This method is relatively precise, simple and low-cost compared with other methods (Casali et al., 2006, 2015). Gully cross-sectional geometry was surveyed by dividing the cross section into trapezoidal segments at abrupt changes in the bank, and measuring the width and depth of the gully at each segment (Fig. 3). Cross-sectional area (A) and surface area (S) were then calculated as…."

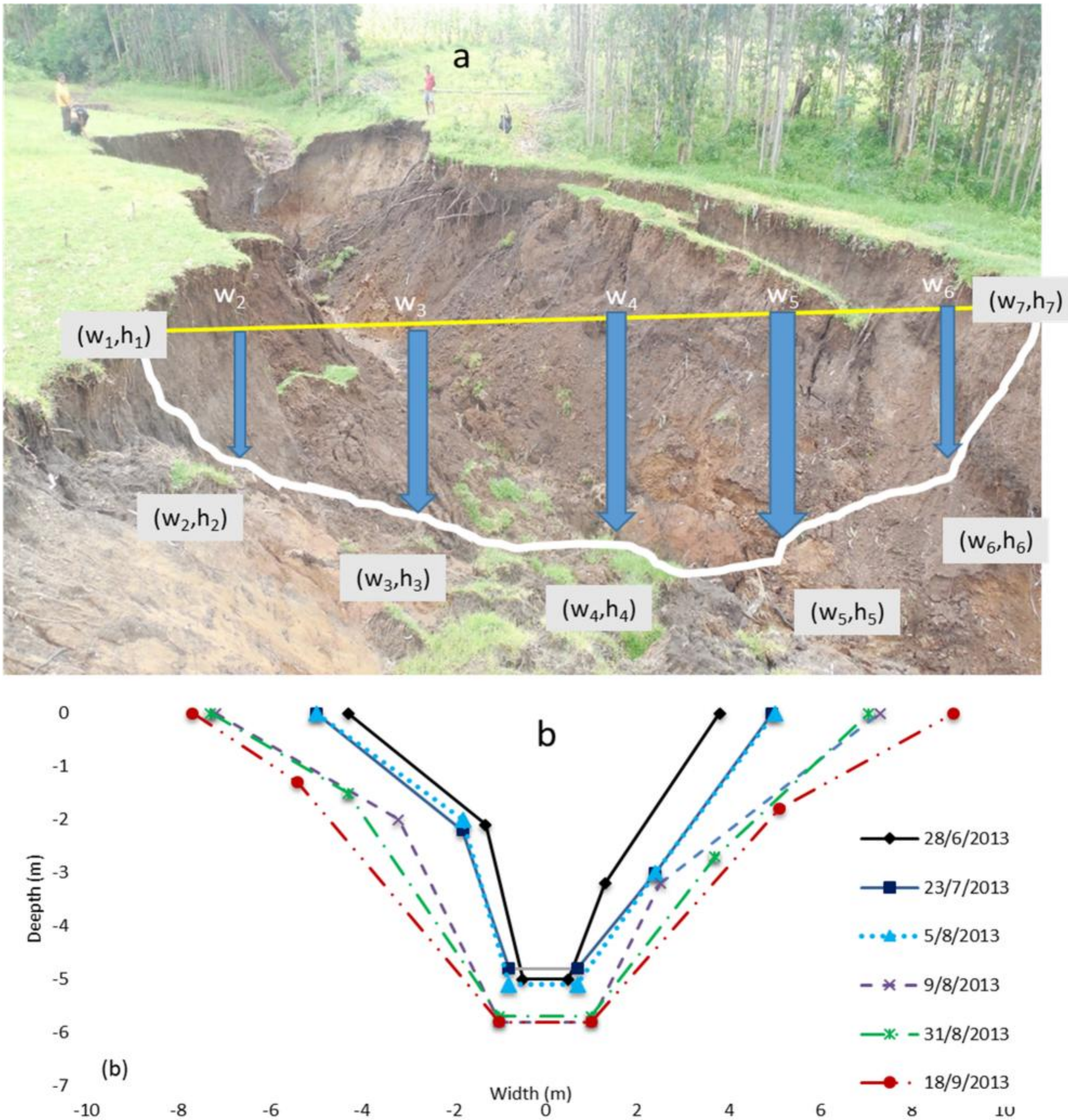


Fig.3. (a) Cross section segmentation methodology to determine the cross-sectional area of the gullies. (b) Measured profiles of a cross-section located on gully G6 during the 2013 rainy season, showing the lateral and downward expansion of the gully.

Comment 4, line 140:

Please include the date

Response 4:

"on August 2013" is inserted (see response 5 below)

Comment 5, line 145:

This equation has great implications on it. The geometry is not the same throughout the time. Is sensitive to gully size. Please include on the discussion the uncertainty on it.

Response 5:

The following changes in the paragraphs are included in the METHOD and DISCUSSION sections as shown below.

> *"METHOD*
>
> Gullies were digitized by determining the location of each gully in the watershed using a hand-held GPS with a horizontal accuracy of about 3 m on August 2013, after which its coordinates were imported into Google Earth to situate all gullies on the aerial imagery. The gully edges were then digitized using Google Earth's polygon mapping tool. Finally, the digitized polygons were converted to shape-file format using ESRI's ArcGIS software, which was also used to calculate the surface area and the length of each gully. Since gully volume could not be obtained from aerial measurements, it was derived from the digitized gully surface area through a regression of the surface area and volume of the measurements of the 13 gullies with surface area in 2013 ranged from 260 to 14,050 m^2 (Table 2). The following regression equation was obtained
>
> .
>
> $$V = 0.54\, S^{1.226} \qquad R^2 = 0.98 \qquad (4)$$
>
> where *S* is the gully surface area (m^2) obtained from Google Earth and V is the predicted volume (m^3) of the gully. The total gully volume for the entire watershed is then simply the sum of all individual gully volumes. Obviously Eq. 4 is only valid in the sub-humid Debre Mawi watershed where the valley soils are deep and the depth is not restricted by bedrock. The area to volume relationship developed by Frankl et al., (2013) for gullies in the semi-arid Ethiopian highlands has a different form because of the bedrock at shallow depths that control the vertical growth."
>
> The absolute relative error (E) in predicting gully volume using Eq. (4) was calculated as:
>
> $$E = \frac{\left|V_{T,p} - V_{T,m}\right|}{V_{T,m}} \qquad (10)$$
>
> where, $V_{T,p}$ is predicted volume (Eq. 4) and $V_{T,m}$ is the measured volume for each gully.

Comment 6, line 155:

How many and where were they located?

Response 6:

We clarified the paragraph as:

> "Ground water elevation is believed to be one of the most important factors for gully formation and bank instability (Tebebu et al., 2010). Therefore, ground water depths were measured using a piezometer installed 5-10 m above each gully head (13 piezometers)".

Comment 7, line 161:

Recording frequency?

Response 7:

The following text is included at the beginning of the paragraph

.

> "Daily precipitation was measured with 5-minute intervals using an automatic tipping bucket, self-emptying rain gauge installed in the northern portion of the watershed".

Comment 8, line 165:

How many samples?

Response 8:

The following text is inserted

> "A total of 55 soil samples for bulk density (BD) and for textural analysis were collected from different soil layers along the bank profile of the sidewalls near the gully head (the number of layers varies from 3 to 5 depending on the gully depth). Samples for BD were collected with a 98-cm^3 (5 cm high) cylindrical core sampler. Soil samples were dried for 24 h at 105 °C, and bulk density was calculated by dividing the mass of the oven-dried soil by the volume of the core. The textural analysis was carried out using the hydrometer method after sieving (Day, 1965)".

Comment 9, line 178:

There is more behind this index:doi:10.1016/j.jhydrol.2012.12.004

Response 9:

Thanks for the reference, the text is adjusted as:

> "In general, based on Ritter and Muñoz-Carpena (2013), NSE > 0.65 is considered acceptable and NSE = 1 indicates a perfect fit, while an NSE < 0 suggests that the mean of the observed values is a better predictor than the evaluated model itself."

Comment 10, line . The quantitative value is included on the Table, so please consider to remove it. The scale is really coarse, and makes the information inaccessible to the reader.

Response 10: We improved the figure as shown below and we removed it in the manuscript as per the comment and included in supplementary material

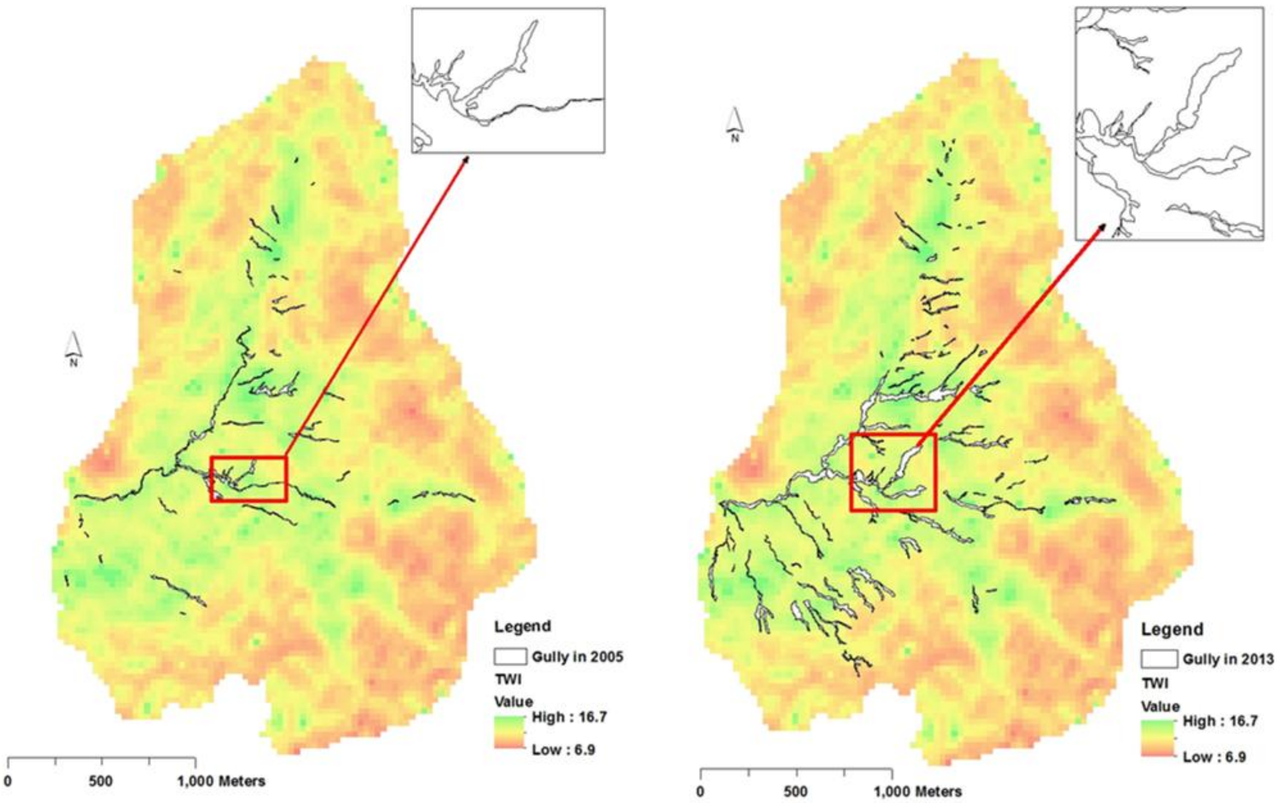


Figure Sx: The relationship between gully formation locations and topographic wetness index (TWI), and gully expansion rate between (a) 2005 and (b) 2013 in the Debre Mawi watershed, Ethiopia. Lines represent gully edges digitized from aerial imagery.

Comment 11, line 184:

All the time, when we report a certain value should be included the error estimation. Otherwise we are assuming that the value is the 100% the real value with is untrue. On following paper is reported a way to estimate the error approximation based on previous literature: DOI: 10.1002/jgrf.20147 Please consider it.

Response 11:

We have added the uncertainty of the values throughout in the revised manuscript. This comment caused the delay in our response. It was difficult to agree how best to do this and what values to use. At the end we presented the uncertainty of most measurements in the revised manuscript and we added the following sections to the manuscript on how the uncertainties were calculated.

In the Material and Methods section we added the following:

Errors of the following measurements were considered; (1) error generated from using the average bulk density to calculate the amount of soil loss, (2) measurement errors of length width and cross-sectional area of the gully and, (3) the accuracy of the drainage area estimated from the DEM

We obtained the measurements errors as follows: The bulk density measurement error was equated with the standard error that was calculated as the standard deviation of three to five samples taken for each layer (there were up to 5 layers in a bank of a gully) of each gully divided by the square root of the number of observations. The measurement error of the length and width was assumed to be related to tape measurement and was estimated at 0.1 m. The measurement error of the cross-sectional area was1 m^2 based on our previous experience.

The drainage area measurement error was mainly attributed to the accuracy of the DEM that was used to delineate the drainage area. For this we used the relationship of the relative errors in 14 sub-catchments studied by Oksanen and Sarjakoski (2005). The digitized surface area error for all gullies' in the watershed was also estimated based on the errors calculated from the 13 gullies.

In order to calculate the uncertainty of the surface area (S), volume (V) and soil loss (SL) of the gullies we used the method presented by Ku (1966) on the propagation error (e) as:

$$e\,(x \pm y) = \left[(e\,(x))^2 + (e\,(y))^2\right]^{1/2} \qquad (8)$$

$$e\,(xy) = |xy| * \left[(e\,(x)/x)^2 + (e\,(y)/y)^2\right]^{1/2} \qquad (9)$$

Where e(x) and e(y) are the measurement errors of x and y, and x or y are variables stands for the errors generated from length, width, area, volume or bulk density.

In the results section we will incorporate the following in the text

The uncertainty derived from the propagation error calculation using Eqs. (8, 9) and prediction errors (Eq. 10) for the 13 gullies are presented in Tables 1-3. The error calculations were made on: (1) the gully expansion caused by all gullies in the watershed (Table 1). (2) the gully expansion caused by the entire length of each thirteen gullies (Table 2); and (3) the gully expansion caused by head cut retreat of the 13 gullies (Table 3).

The uncertainty in calculating the combined volumetric retreat (2013-2014) from the 13 gullies was 66.7 m^3, and the error generated from the combined soil loss (2013-2014) was 122 t (Table 3). The uncertainty in calculating the total soil loss from the 13 gullies between 2005 and 2014 was 7625 t (Table 2) and the estimated error in calculating soil loss in the entire gully network in the Debre Mawi watershed in 2013 was 28,281 ton (Table 1). The uncertainty derived from digitizing the drainage areas of 13 gullies was estimated at 1.6 ha, and for individual gully ranged from 0.12 to 0.9 ha (Table 3).

Table 1. The combined length, area and volume of the total gully network in the 608 ha Debra Mawi watershed obtained from satellite imagery in 2005 and 2013. The "soil loss" in the last column represents the total soil loss from the gully network preceding the date of measurements and is calculated as the volume in column 4 times the bulk density. Errors were estimated using Eqs. (8-10).

	Gully length km	Gully area ha	Gully volume 10^3 m^3	Soil loss 10^3 t
2005	8.7	4.5	140	168
Estimate error in 2005	-	0.03	5	6
2013	26.0	20.4	654	784
Estimated error in 2013	-	0.1.	23	28
Increase from 2005-2013	17.3	15.9	514	616
Relative change, % 2005-2013	197	350	366	366

Table 2. Increase in surface area and corresponding soil loss of the 13 gullies in the Debre Mawi watershed in the period between 2005 and 2014. Surface area up to March 2013 was obtained by digitizing the gully edges on aerial imagery and the next two rainy phases by manual measurement.

Gully name	Gully surface area (m^2)						Bulk density (g cm^{-3})		2005 - 2014						2013-2014			
	From aerial image				Manual measurement													
	6/3/05	4/5/11	3/4/12	23/3/13	18/9/13	18/10/14	measured	error	Change in area (m^2)	Error (m^2)	Chang in volume (m^3)	error (m^3)	Soil loss (t)	Error (t)	Change in volume (m^3)	Error (m^3)	Soil loss (t)	Error (t)
G1	140	265	390	420	440	440	1.26	0.02	300	20.2	709	55.2	894	71.6	52.1	4.1	66	5.3
G2	785	2700	3330	3560	3573	3575	1.19	0.02	2790	113.4	10354	1164.9	12321	1409.3	63.1	7.1	75	8.6
G3	210	530	600	1430	1460	1460	1.14	0.05	1250	26.0	3712	546.3	4232	647.4	102.8	15.1	117	17.9
G4	230	400	450	750	780	785	1.17	0.11	555	16.2	1488	168.7	1740	254.4	104.0	11.8	122	17.8
G5	2820	10700	11500	13700	13960	14050	1.16	0.04	11230	128.8	56511	794.9	65553	2601.9	2000.3	28.1	2320	92.1
G6	1720	6770	8100	9110	9580	9960	1.22	0.18	8240	16.7	38076	142.4	46453	6863.1	4462.8	16.7	5445	804
G7	110	365	365	385	390	390	1.15	0.05	280	19.6	639	108.0	735	127.8	12.7	2.2	15	2.6
G8	365	2140	2860	3740	3850	3890	1.19	0.09	3525	40.2	12856	420.6	15299	1235.7	640.3	21.0	762	61.5
G9	40	730	1050	1120	1150	1180	1.19	0.09	1140	35.6	3102	223.6	3691	391.6	195.3	14.1	232	24.6
G10	50	190	400	455	460	460	1.25	0.11	410	17.1	928	98.6	1160	162.9	13.2	1.4	17	2.4
G11	152	600	750	890	1020	1070	1.23	0.07	918	14.3	2540	42.7	3124	173.3	565.0	9.5	695	38.6
G12	50	170	240	255	255	260	1.22	0.06	210	15.9	428	132.2	522	163.1	11.6	3.6	14	4.4
G13	199	240	345	365	370	370	1.14	0.05	171	12.2	405	90.5	462	105.2	12.6	2.8	14	3.2
Total	6800	25800	30380	36180	37288	37890	1.19	0.30	31019	187.7	131748	1617.0	156186	7624.8	131748	47.54	9894	814

Table 3. List of soil and gully topographic factors for the 13 gully heads in the Debre Mawi watershed and observed gully head erosion during the 2013 and 2014 rain phase (between July and September). BD is bulk density and DA is drainage area.

Gully name	Min. water table depth (m)	Clay content (%)	Mean bulk density (g cm^{-3})		head cut depth (m)	Drainage area (ha)		Linear headcut retreat(m)		Area retreat (m^2)		Volumetric retreat (m^3)		Soil loss (t)	
										2013-2014		2013-2014		2013-2014	
			measured	error		measured	error	2013	2014	measured	error	measured	error	measured	error
G1	1.50	58	1.26	0.02	3.9	12.8	0.43	0.4	0	3.3	0.82	9	2.3	11	2.8
G2	1.22	53	1.19	0.02	2.2	13	0.43	2.2	0.5	10.9	0.62	32	5.7	38.5	6.8
G3	0.02	55	1.14	0.05	1.4	41.6	0.72	36	0	22.5	2.12	146	35.0	167	40.6
G4	0.59	59	1.17	0.11	2	1.7	0.18	7	5	15.7	0.61	61	14.1	72	17.8
G5	0.05	60	1.16	0.04	4.8	68	0.89	10	3	101.5	2.67	1087	50.0	1260	74.4
G6	0.08	67	1.22	0.18	4.6	13.3	0.44	12	0	182.0	2.27	413	16.6	504	77.1
G7	1.36	59	1.15	0.05	1.4	0.7	0.12	0.2	0	0.7	0.33	1	0.5	0.9	0.6
G8	0.07	56	1.19	0.09	3.3	17.4	0.49	24.4	7	108.9	1.94	237	11.0	281	24.5
G9	1.20	59	1.19	0.09	3.4	6.8	0.33	3.8	1.65	21.2	0.76	73	6.2	87	10.0
G10	1.44	55	1.25	0.11	2.5	6.5	0.32	0.7	0	2.7	0.39	6	1.9	7.5	2.4
G11	0.45	66	1.23	0.07	4.2	9.2	0.37	6.2	1.4	123.4	2.72	356	7.0	437	24.7
G12	1.38	66	1.22	0.06	1.9	4.1	0.26	0.07	0	5.0	0.38	3	1.0	4	1.2
G13	1.25	60	1.14	0.05	1.3	4.8	0.28	0.04	0.8	10.1	0.63	3	0.6	2.8	0.5
Total /Ave	**0.82**	**59.5**	**1.19**	**0.30**	**2.84**	**200**	**1.63**	**103**	**19**	**608**	15.13	**2427**	**66.7**	**2873**	**121.8**

Comment 12, line 195:

By definition is not only by headcut retreat, on this value is included the bank failure process. Please clarify on the introduction that when you are refereeing to headcut retreat you are including both process: headcut retreat (longitudinal growth) and bank failure (cross-sectional growth)

Response 12:

We added the definition in the introduction section as:

> "Headcut retreat is a linear longitudinal growth of headcut and bank failure or gully widening is the cross-sectional growth due to the headcut retreat".

Comment 13, line 202:

what is the estimated error on this measurement?

Response 13:

We reported the measurement errors in the tables (Tables 1-3) and reported some of them when appropriate (especially when in cases the quantity included multiple measurement errors) in the text (see comment 11). An example is given for one of the paragraphs that reads as:

> "The recorded precipitation during the 2013 (44 days of rainfall) and 2014 (31 days of rainfall) rainy phases was 917 and 1107 mm, respectively (Fig. 5c). The gully headcut retreat in 2013 ranged from 0.04 to 36 m, with a combined total of 103 m increase in gully length (Fig. 5a, Table 3); whereas the total retreat in 2014 ranged from 0 to 7 m, with a combined total of 19 m (Table 3). Over these two monsoon seasons (2013-2014), about 608 ± 15 m^2 of cultivated land was consumed by only the longitudinal headcut retreat of the 13 gullies. This is equivalent to 44% of the increase in total surface area (both longitudinal and lateral retreat of the entire gully) of the 13 gullies during 2013-2014, and about 4.5% of the total surface area of the 13 gullies since their formation up to 2014. During 2013-2014, the soil loss solely due to headcut migration equaled 2873 ± 122 ton (Tables 3) which represented 30% of the total soil loss from the 13 gullies in the same period, and about 2% starting from their formation up to 2014 (Tables 2 and 3) ……."

Comment 14, line 214:

This is not discussion section

Response 14:

Thanks, we adjusted as:

> "The relationships between the lateral and longitudinal retreat and the associated volumetric soil loss are discussed in Sect.4".

Comment 15, line 223:

Did you explore the use of an saturation index, as the antecedent precipitation d1, d5, d10? If one of the main factors is soil saturation should deliver some results.

Response 15:

Previous studies in the study area showed that saturation is the main driving factor for runoff generation and gully formation. So we included some text and now the paragraph reads as:

> "The correlation between the observed change in linear gully headcut retreat (R_L) and the precipitation recorded during the day of the gully head retreat occurrence varied between -0.23 and 0.88. Some of the big gullies such as G5, G6 and G11 showed strong correlation ($R_{L, G5}$ = 0.88, P = 0.009 and $R_{L, G6}$ = 0.84, P = 0.017), whereas gullies with the greatest linear retreat (L_{G3} = 36 m and L_{G8} = 24 m; Fig. 5a) showed weak relationships with the daily precipitation during the retreat event ($R_{L,G3}$ = 0.27, P = 0.55 and $R_{L,G8}$ = 0.34, P = 0.37). The probable reason for these fairly low correlation coefficients is that there is a time delay between daily rainfall and saturation of the soil surrounding the gully (Tebebu et al., 2010, Tilahun et al., 2013). The probable reason for these fairly low correlation coefficients is that daily rainfall only slowly saturates the surrounding soil, which is partly responsible for destabilizing the gully head. Due to such slow saturation processes, the daily precipitation and gully head retreat may not correlate well. However, after the maximum recorded daily rainfall (94 mm) on 7 Aug 2013, the largest retreat rates were observed on 13 Aug 2013 measurements (Fig. 5a, measurements were carried out 6 days later than the 94 mm rainfall with little or no rainfall within these days). The low correlations indicate that headcut retreat may not always occur during or immediately following precipitation except for very large storm events such as on 7 Aug 2013.

Comment 16, line 235:

error range

Response 16: The error range is a function of DEM accuracy and analysis procedure. The details are given in response 11. Particularly for the measurement error of the drainage area, we included the following in the Material and Methods section

> "The drainage area measurement error was mainly attributed to the accuracy of the DEM that was used to delineate the drainage area. For this we used the relationship of the relative errors in 14 sub-catchments studied by Oksanen and Sarjakoski (2005). The digitized surface area error for all gullies' in the watershed was also estimated based on the errors calculated from the 13 gullies".

In the Results section the uncertainties are specified in the revised manuscript

> "The drainage area for the studied gullies varied from 0.7 (± 0.12) to 68 (± 0.9) ha) with an average value of 15.4 ha and standard deviation of 18.9 ha (Table 3). In order to understand whether drainage area is related to retreat of the gully in 2013, we fitted simple linear regression models (Eqs. 11-12) and power law relationships (Eqs.13-14) between cumulative headcut retreat length (L_T, in m) in 2013 and drainage area (DA, in ha), and between increase in gully volume ($V_{T,}$ in m^3) and DA were developed. Since rainfall in 2014 was not erosive and small gullies did not retreat, we only used the data for 2013 for the regression".

."

Comment 17, line 243:
This value is different than the reported in the equation, please double check or clear it.
Response 17:
Our apologies for the confusion. R^2 = 0.27 is the value when we fitted the predicted V with measured V (not L-DA). The paragraph is reworded as:

> "The predicted L and V using Eqs. 11 and 12 were compared with the measured L and V and tested statistically (Eqs.5-7). The goodness to fit were explained as: R^2 = 0.27 (p = 0.06), NSE = 0.11, and PBIAS = 52% for L; and R^2 = 0.69 ($p << 0.01$), NSE = 0.47, and PBIAS = 49% for V."

Comment 18, line 244:
In all the cases the relationship appears to be statistically insignificant, is it ok?
Response 18:
We agree and omitted the text in which we reported on insignificant differences. The discussion section 4.3 is reworded as:

> "Both the linear (Eqs. 11-12) and power (Eqs. 13-14) type regression relationships indicated that drainage area predicted the volumetric gully erosion (V_T) better than the linear migration of the gully headcut (L_T). This suggests that the larger the drainage area, the greater the lateral gully expansion is by collapsing banks, and hence the greater the sediment production is. Other studies in the semiarid highlands with relatively shallow soils over bedrock have indicated that drainage area (which was not significantly related in the Debre Mawi catchment with deep soils) was a major controlling factor of gully head retreat (Poesen et al., 2003; Frankl et al., 2012)."

Comment 19, line 249:
The fluctuation is not included in the figure, please consider it.
Response 19:
The average groundwater fluctuation between morning and night is included in the revised draft in the figure 5 and text

> "…..The groundwater table fluctuated between these readings (Fig.5), but the variation was not significant (p = 0.98). The water table decreased between morning and evening readings on average by 0.7 cm with a standard deviation of 4.0 cm. The greatest fluctuations were observed at G2 (Fig.6). The power type regression model between the minimum water table depth during the rainy season (ranging from 0.02 m at G3 to 1.5 m at G1) and the linear retreat and volumetric expansion of the 13 gullies had fairly high coefficients of determination…"

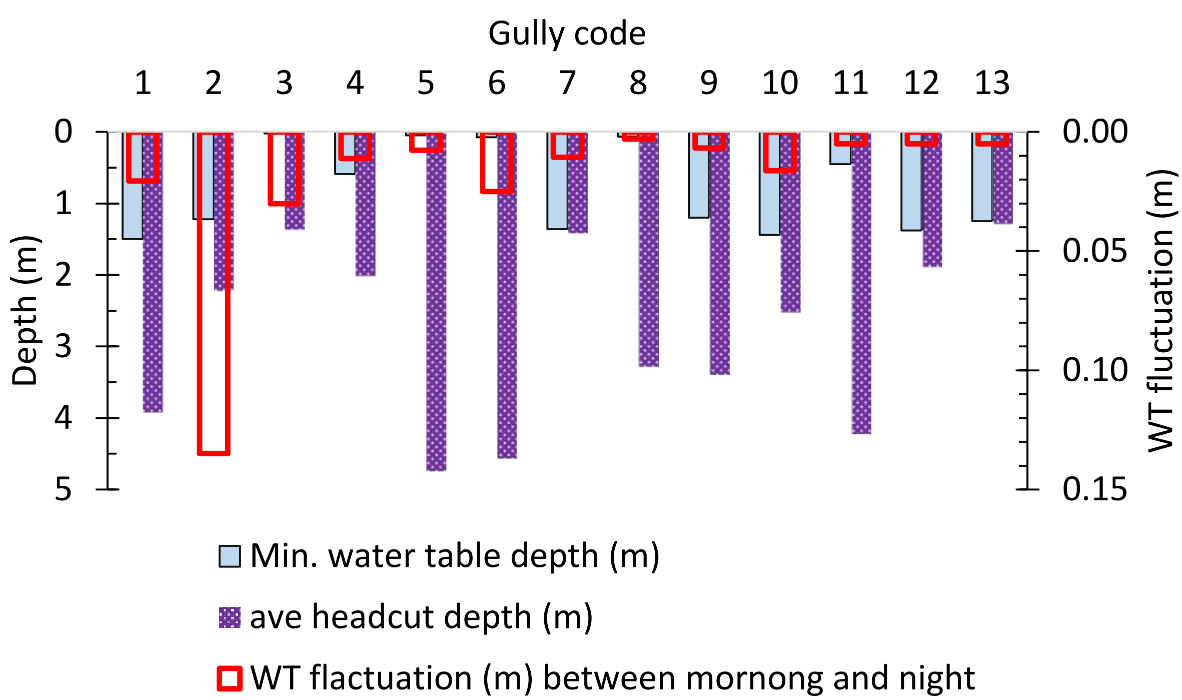


Fig. 5. Comparison of minimum groundwater table depth, gully headcut depth and the average groundwater fluctuation between morning and night for the 13 study gullies in the Debre Mawi watershed, Ethiopia for the 2013 rainy season. WT is water table, Min. is minimum.

Comment 20, line 254:
The height is used to estimate the volume, However, using the r^2 as goodness of fitting it would be classified as unaccepted.

Response 20:

Thanks! We checked the data and some results are changed. The following text is included in Sect 3.3

> "By fitting a simple linear regression, the volumetric gully expansion was significantly related to the height of the gully headcut ($R^2_V = 0.49$, $p = 0.007$). However, the linear retreat of the gully was not well explained by the headcut height ($R^2_L = 0.0004$, $p = 0.9$). The reason is likely the fact that of gully G3 had large linear retreat but small headcut height and therefore influenced the analysis. When this gully is excluded from the analysis, the R^2_L for the linear and power relationship between the gully linear retreat and gully head height increased from 0.0004 to 0.26 ($p = 0.09$) and from 0.21 to 0.52, respectively. In this case, the gully height fairly explained the linear retreat. The mass of potential gully head failure blocks is smaller for lower gully head heights, which corresponds to increased stability of the gully head. An equivalent increase in gully head stability can be obtained by regrading the gully head to a lower slope for a given height."

Comment 21, line 276:

Saturation does not explain the 100%, so I will consider to remove the verb "prohibit"

Response 21:

The word "prohibit" is removed and the word "reduces" is replaced

Comment 22, line 285:

It is an important factor, but it is not the only one. In fact, in your paper there are several things to clarify. First, where were the located the piezometers. Second, Do you assume that the elevated ground water is affecting at the same level all the gully; headcut and walls?

Response 22:

In the revised manuscript, the locations for the piezometers are included in the method section (Sect.2.2.3) as per the comment.

> "Ground water elevation is believed to be one of the most important factors for gully formation and bank instability (Tebebu et al., 2010). Therefore, ground water depths were measured using a piezometer installed 5-10 m above each gully head. Intrusion of silt and sand to the piezometer was prevented by wrapping filter fabric around the 40 cm-long screened bottom end. All piezometers were capped to prevent rainwater entry and were set in concrete to prevent any physical damage. Groundwater table elevations were read using a measuring tape twice a day: in the morning and in the evening"

The elevated ground waters do not equally affect all gullies. But it is a major driving factor for gully formation and expansion in the study watershed. But there are other factors listed in Table 4 that affect the gully retreat. The word "but not only" is added to explain that it is not the only factor. The paragraph in Sect 4.1 reads as

> "Most gullies investigated in the watershed were not stable and have impaired more than 16 hectares of agricultural land from 2005 to 2013. In fact, gully expansion in the Debre Mawi watershed is not distributed evenly over the watershed as the geology of the upper slopes of the watershed (about 50% of the watershed area) reduces gully formation because (but not only) it does not saturate (Tilahun et al., 2013b, Steenhuis et al., 2014; Tebebu et al., 2015). Gully expansion therefore affects mostly the bottomlands where soils do saturate (Fig. 4). A loss of 2 ha of productive farmland per year is considerable for any farmer, but even more significant in a region with smallholder farmers. As farmers' land holding in the Ethiopian highlands is about one hectare of land per household (Sonneveld and Keyzer, 2003), the land loss observed between 2005 and 2013 could have provided farmland for 16 farming households in the watershed".

Comment 23, line 288:
Could be interesting to normalize the gully erosion rate by the rainfall and analyze if the driver is the rainfall or others factors.
Response 23:
Rainfall brings up the groundwater that cause the bank to slip that causes high sediment concentration. However, sediment measurements in the inlet and outlet of gully G6 (Zegeye et al., in preparation) indicates that in most cases peak sediment load and concentration occur several minutes before the peak storm runoff. So it is slightly more complicated that normalizing for rainfall. We will look into this in a follow up paper if times permit.

Comment 24, line 319:
Since the same driver is found in this study, among other, please remark the main difference of this study.
Response 24:
The main difference is that the drainage area didn't explain the linear headcut retreat significantly. So, after the paragraph is rearranged, we included as follows:

> "Both the linear (Eqs. 11-12) and power (Eqs. 13-14) type regression relationships indicated that drainage area predicted the volumetric gully erosion (V_T) better than the linear migration of the gully headcut (L_T). This suggests that the larger the drainage area, the greater the lateral gully expansion is by collapsing banks, and

hence the greater the sediment production is. Other studies in the semi-arid highlands with relatively shallow soils over bedrock have indicated that drainage area (which was not significantly related in the Debre Mawi catchment with deep soils) was a major controlling factor of gully head retreat (Poesen et al., 2003; Frankl et al., 2012)."

Comment 25, line 354:
I would suggest to include this figure at the beginning of the manuscript, in the study area section. A deeper description of each analyzed gully should be provided.
Response 25:
Agreed and we moved it to introduction and cited accordingly in the paragraph 2 and 3 as shown below.

"Gullying is a threshold-dependent process controlled by a wide range of factors (Valentin et al., 2005), including rainfall and flowing water, soil properties, and drainage area. Capra et al. (2009) and Campo et al. (2013) found that most of the gully erosion took place during heavy rainfall events, i.e., storm events were one of the main drivers for gully erosion. The mechanic actions of the flowing water can result in a rapid mass movement in the gullies by undercutting of the banks (See Fig. 1, Lanckriet et al., 2015). When these mechanic actions at the gully head exceed the cohesive strength of soil, erosion proceeds upslope through a head ward cutting gully (Munoz-Robels et al., 2010). Interactions between such processes are important as hydraulic erosion promotes bank collapse, which then modifies subsequent hydraulic erosion (Thorne, 1990; Avni, 2005). Similarly, gully formation is initiated with the occurrence of convergent shallow subsurface flow that leads to seepage-induced erosion of surface soils, gully heads and sidewalls (Fig.1f; Vanmaercke et al., 2016; Tilahun et al., 2013a) and sliding (Fig.1d). Active gully networks are therefore predominantly found in the saturated valley-bottomlands (Tebebu et al., 2010; Steenhuis et al., 2014), and the deepest and the most spectacular gullies occur in the bottom of the watershed where in sub-humid monsoonal and wetter climates, the soil becomes saturated starting around the middle of the rainy phase and then remain saturated until the end of the rainy season (Tebebu et al., 2014).

Soil properties and soil types also play a role in gully formation and expansion. For example, Vertisols, heavy clay soils with a high proportion of swelling clays (IUSS Working Group WRB, 2015), form deep wide cracks from the surface downward when they dry out (Fig. 1c) and are prone to the development of pipes (Fig. 1e) that can collapse and thereby turn into rills or gullies (Valentin et al., 2005; Frankl et al., 2014). This may be one of the reasons that most severe gully areas are often associated with Vertisols (Valentin et al., 2005; Tebebu et al., 2014; Frankl et al., 2014). Similarly, in pasture bottom lands, piping often leads to development of permanent gullies (Jones, 1987; Zegeye et al., 2014). These pipes are part of gully networks and during the rainy season, the infiltrating rainfall discharges through the pipes, which increases the lower soil horizon's vulnerability to erosion".

Comment 26, line 363:
Figure 7 shown different gully growth processes, that should be described in the introduction section and study area
Response 26:
Based on comment 25, we moved it in to Introduction section

Comment 27, line 379:
Both erosion types have been mixed in the manuscript, please try to use the proper words
Response 27:
We defined both erosion types as: gully widening is the cross-sectional width increment and gully headcut retreat as the linear longitudinal growth of headcut (as explained in response 12 above)

Comment 28, line 381:
These statements has not been introduced previously.
Response 28:
We agree and it is removed

Comment 29, Fig1:
Gullies are hard to identify on this figure. Please try to clear it. Gully label is unclear. And also a DEM or elevation lines will help to understand the study area. What about, including a land use map?
Response 29:
We have improved the figure as shown below, we mapped the stream and contour lines here and the 13 studied gullies.

We improved the figure as shown below

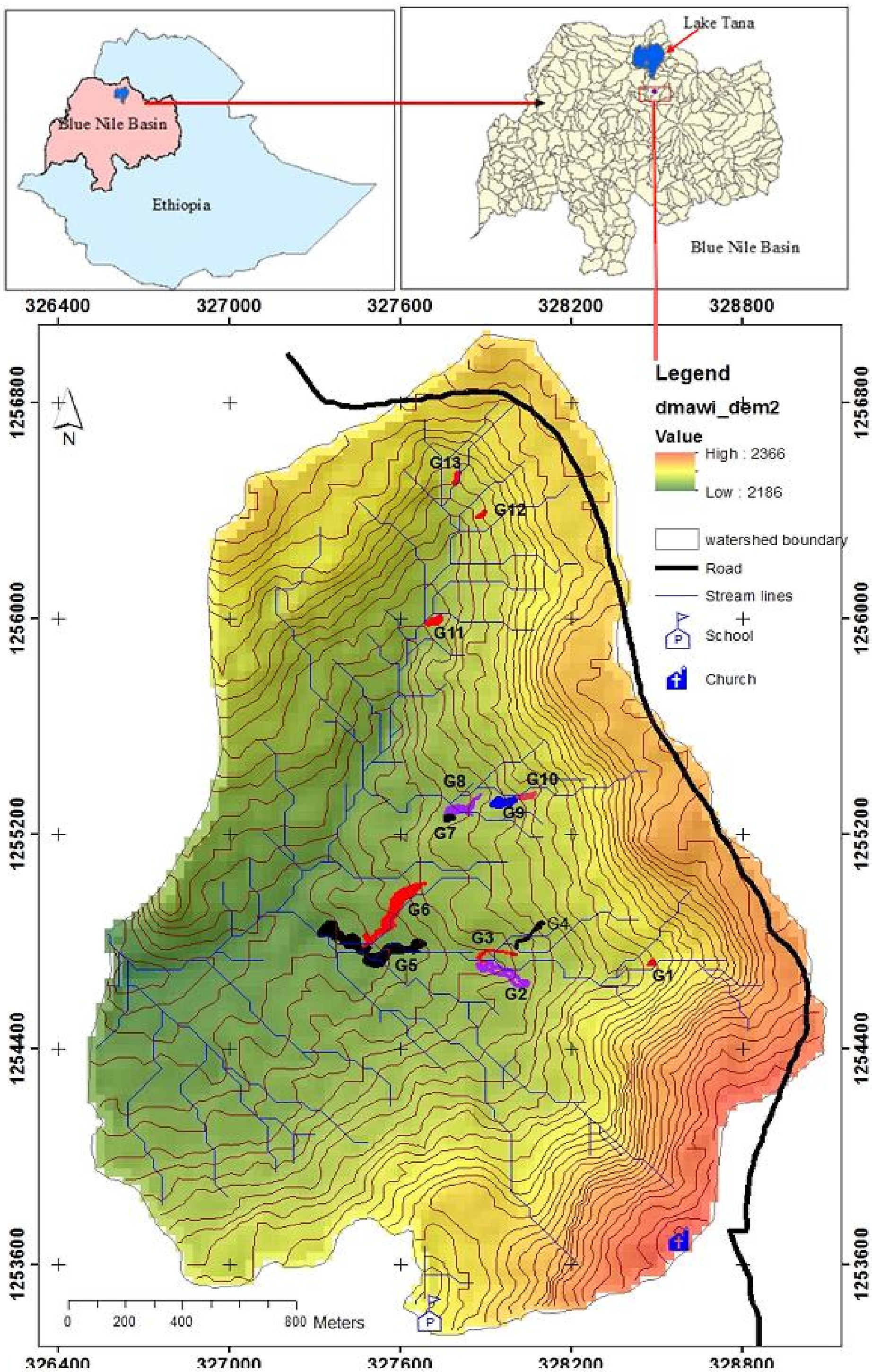
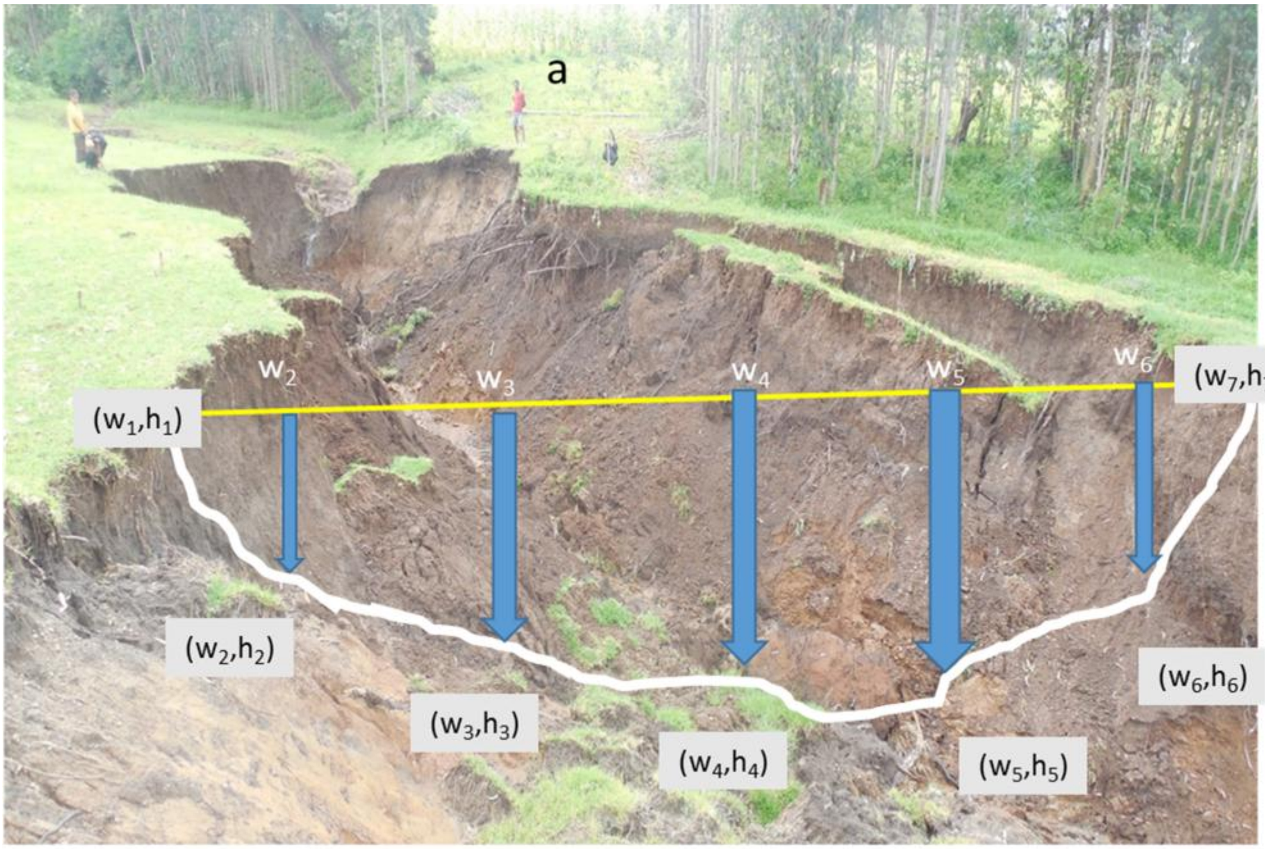


Figure 1 (now Figure 2). Location of the Debre Mawi watershed within the Blue Nile River basin, Ethiopia (top figures). The watershed map (bottom) shows the contour lines, elevation, stream lines, and the 13 studied gullies (indicated by the labels beginning with the letter G). Projected Coordinate System: WGS_1984_UTM_Zone_37N

Comment 30, Fig1:
The coordinate system is incomplete. Please include the reference projection system used on it.
Response 30:
The following text is included in the figure caption (see response 29 above)
Projected Coordinate System: WGS_1984_UTM_Zone_37N

Comment 31, Fig2: in the picture the cross section looks like stable. Please describe if the picture is from the same cross-section included in Fig2.b
Response 31:
Our apologies for the confusion, the picture was used only to show how the gully cross-section measurements were performed but this gully is found in other locations whose cross-section was measured for other studies in 2010. So it is replaced by the actual measured gully in the Debre Mawi watershed as marked by the yellow straight line.



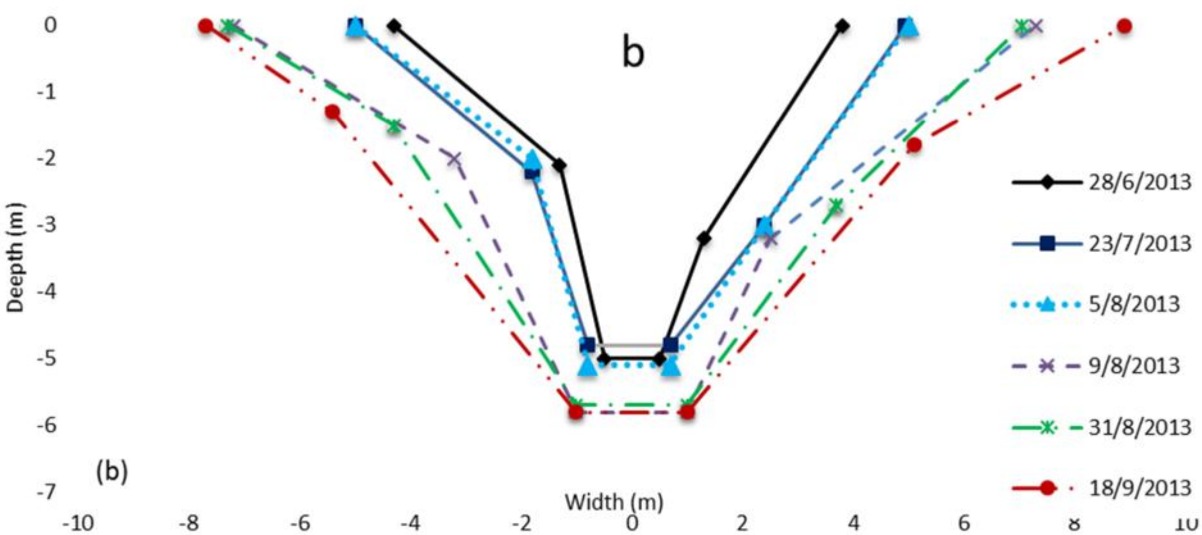


Figure 2 (now Figure 3). (a) Cross section segmentation methodology to determine the cross-sectional area of the gullies. (b) Measured profiles of a cross-section located on gully G6 during the 2013 rainy season, showing the lateral and downward expansion of the gully.

Comment 32, Fig4:
So in this case when you refer to the headcut retreat you are speaking about the longitudinal growth, however many times you use this word for the total gully erosion on the manuscript. Please clear it.
Response 32:
In the revised manuscript, we are consistent and used the term headcut retreat as the longitudinal growth of a gully head. For example: we used these terms in the conclusion section as:

> "Field observations in the Debre Mawi watershed indicate that permanent valley-bottom gully drainage networks and in particular gully widening and headcut retreat are important erosion processes severely impacting the productive farmlands".

Comment 33, Fig4:
Comm RF???
Response 33:
It was intended to say Cumulative RF but was wrongly abbreviated. So adjusted as a full word 'cumulative"

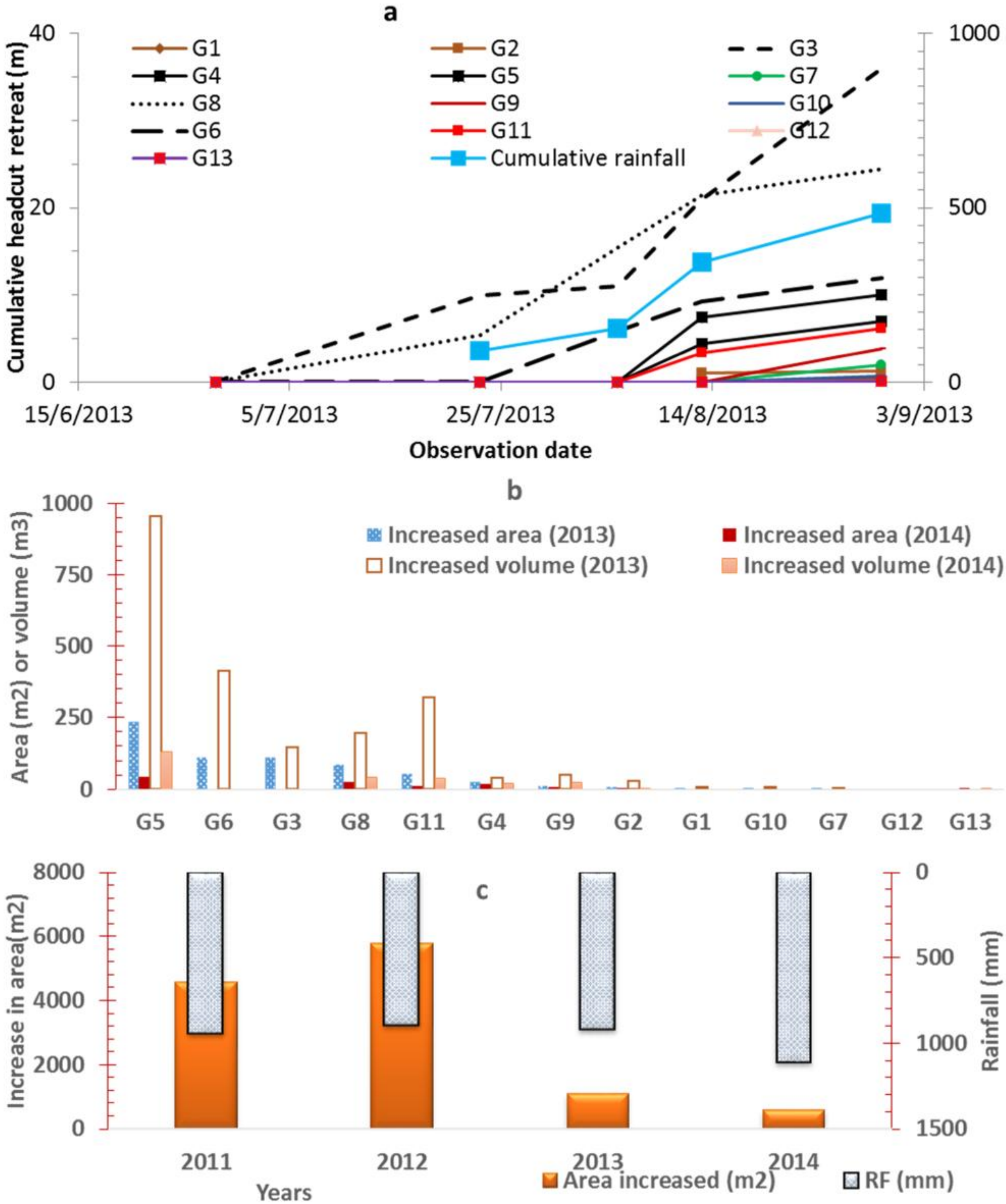


Fig. 4. The observed expansion of the 13 study gullies in the Debre Mawi watershed (see Fig. 2 for gully location): (a) cumulative headcut retreat and rainfall during the 2013 rainy season, (b) increase in gully surface area and volume during the 2013 and 2014 rainy seasons, and (c) increase in the combined gully surface area and the total summer rainfall (RF) between 2011 and 2014.

Comment 34, Fig4:
Please remove the border line.
Response 34:
We removed (as shown in response 33)

Comment 35, Fig4:
When was the rainfall gauge station set up? Form the study area I believe it was established in 2013.
Response 35:
The rainfall gauge station was set up in the watershed since 2008 for previous studies. The paragraph in Sect. 2.1 reads:

> "The study area, the Debre Mawi watershed, is located in the sub-humid highlands of northwest Ethiopia, 30 km south of Bahir Dar along the road to Adet, and lies between 11º20' and 11º22' N and 37º24' and 37º26' E. The watershed drains an area of 608 ha. The altitude ranges from 2186 to 2366 m (Fig.2).; the elevations of the gullies considered in this study range from 2212 to 2272 m. Rainfall is unimodal with an average value of 1240 mm yr^{-1}. Most rainfall falls between June and the beginning of September and amounts 900 mm yr^{-1}). The rainfall gauge station in the Debre Mawi watershed has been established since 2008 by Adet Agricultural Research Center to record rainfall in rainy phase only. The dry season lasts between 8 to 9 months. The mean daily temperature is 20 °C"

Comment 36, Fig 5:
Why do you connect the triangle symbols by lines?
Response 36:
It was to reduce the complexity in viewing the graph. Points are now disconnected and shown by bar graph as shown in the figure in response 19 above.

The following additional references are included in the revised draft
Campo-Bescós, M. A., Flores-Cervantes, J. H., Bras, R. L., Casalí, J. and Giráldez, J. V.: Evaluation of a gully headcut retreat model using multitemporal aerial photographs and digital elevation models, JOURNAL OF GEOPHYSICAL RESEARCH: EARTH SURFACE, 118, 2159–2173, 2013.

Capra, A., Porto, P. and Scicolone, B.: Relationships between rainfall characteristics and ephemeral gully erosion in a cultivated catchment in Sicily (Italy), Soil Tillage Res., 105(1), 77–87, 2009

Casalí, J., Loizu, J., Campo, M. A., De Santisteban, L. M., and Álvarez-Mozos, J.: Accuracy of methods for field assessment of rill and ephemeral gully erosion, Catena, 67, 128–138, 2006.

Casalí, J., Giménez, R. and Campo-Bescós M. A.: Gully geometry: what are we measuring? SOIL, 1, 509–513, 2015, doi:10.5194/soil-1-509-2015

Frankl, A., Nyssen, J., De Dapper,M., Haile, Mitiku, Billi, P., Munro, R.N., Deckers, J., Poesen, J.: Linking long-termgully and river channel dynamics to environmental change using repeat photography (North Ethiopia), Geomorphology, 129, 238–251, 2011

Frankl A, Poesen J, Scholiers N, Jacob M, Haile M, Deckers J, Nyssen J.: Factors controlling the morphology and volume (V)–length (L) relations of permanent gullies in the Northern Ethiopian Highlands, Earth Surface Processes and Landforms, 38(14), 1672–1684, 2013.

Frankl, A. Poesen, J., Haile, M., Deckers J. and Nyssen, J.: Quantifying long-term changes in gully networks and volumes in dryland environments: The case of Northern Ethiopia, Geomorphology, 201, 254-263, 2013.

Frankl, A. Deckers, J. Moulaert, L. Van Damme, A. Haile, M. Poesen J. and Nyssen J.: Integrated solutions for combating gully erosion in areas prone to soil piping: innovations from the drylands of Northern Ethiopia, Land Degradation & Development, 2014.

Ritter, A., and Muñoz-Carpena, R.: Performance evaluation of hydrological models: Statistical significance for reducing subjectivity in goodness-of-fit assessments, Journal of Hydrology 480, 33–45, 2013.

Vanmaercke, M., Poesen, J., Bert Van Mele, B.V., Demuzere, M., et al.: How fast do gully headcuts retreat? Earth-Science Reviews, 154, 336–355, 2016, doi:10.1016/j.earscirev.2016.01.009

Sørensen, R., Zinko, U., and Seibert J.: On the calculation of the topographic wetness index: evaluation of different methods based on field observations, Hydrology and Earth System Sciences, 10, 101–112, 2006

Grabs, T., Seibert, J., Bishop, K., and Laudon, H.: Modeling spatial patterns of saturated areas: A comparison of the topographic wetness index and a dynamic distributed model, Journal of Hydrology, 373, 15–23, 2009.

Ali, G., Birke C., Tetzlaff, D., Soulsby,C., McDonnell, J.J., and Tarolli, P.: A comparison of wetness indices for the prediction of observed connected saturated areas under contrasting conditions, Earth Surf. Process. Landforms, 2013, DOI: 10.1002/esp.3506.

Oostwoud Wijdenes, D.G. and Bryan, R.B.: Gully-head erosion processes on a semi-arid valley floor in Kenya: a case study into temporal variation and sediment budgeting, Earth Surf. Process. Landforms 26, 911–933, doi: 10.1002/esp.225, 2001.

Beven, K.J., Kirkby, M. J.: A physically based, variable contributing area model of basin hydrology, Hydrolological Science Bulletin, 24, 43–69, 1979.