Improving the positional accuracy of drainage networks extracted from Global Digital Elevation Models using OpenStreetMap data

Elisabete S.V. Monteiro[1*], Cidália C. Fonte[2], João L.M.P. de Lima[3]

[1] UDI-Research Unit for Inland Development - Polytechnic Institute of Guarda / Institute for Systems Engineering and Computers at Coimbra, Av. Dr. Francisco Sá Carneiro, 50 6300-559, Guarda, Portugal. E-mail: emonteiro@ipg.pt

[2] Department of Mathematics University of Coimbra / Institute for Systems Engineering and Computers at Coimbra, Apartado 3008, EC Santa Cruz, 3001-501,Coimbra, Portugal. E-mail: cfonte@mat.uc.pt

[3] Department of Civil Engineering of University of Coimbra / MARE - Marine Environmental Sciences Centre, Faculty of Sciences and Technology, University of Coimbra, Rua Luís Reis Santos, Pólo II University of Coimbra, 3030-788 Coimbra, Portugal. E-mail: plima@dec.uc.pt

[*] Corresponding author. E-mail: emonteiro@ipg.pt

Abstract: Drainage networks allow the extraction of topographic parameters that are useful for basins characterization and necessary for hydrologic modelling. One way to obtain drainage networks is by their extraction from Digital Elevation Models (DEMs). However, it is common that no freely available DEMs at regional or national level exist. One way to overcome this situation is to use the available free Global Digital Elevation Models (GDEMs). However, these datasets have relatively low spatial resolutions, 30 and 90 meters for ASTER and SRTM, respectively, and it has been shown that their accuracy is relatively low in several regions (e.g., Kääb, 2005; Mukul et al., 2017). In this study a methodology is presented to improve the positional accuracy of the drainage networks extracted from the GDEMs using crowdsourced data available in the collaborative project OpenStreetMap (OSM). In this approach only free and global datasets are used, enabling its application to any location of the world. The methodology uses elevation points derived from the GDEMs and the water lines extracted from the collaborative project OSM to generate new DEMs, from which new water lines are obtained. The methodology is applied to two study areas and the positional accuracy of the used data and the obtained results are assessed using reference data.

Keywords: Drainage networks; GDEMs; OpenStreetMap; Positional accuracy.

INTRODUCTION

Drainage networks are necessary for many hydrologic studies, namely the morphometric basin characterization and hydrologic modelling. They are represented accurately in topographic maps, but those may not be available freely in digital form. Another way to obtain them is by extraction from Digital Elevation Models (DEMs). These models can be generated using photogrammetric methods, laser scanning surveys, or radar data and interferometry and may be produced by nationals mapping agencies or companies. These products in most cases are not available for free use. On the other hand, there are Global Digital Elevation Models (GDEMs) freely accessible with no costs, which cover almost the entire surface of the Earth. However, the GDEMs have errors due to several factors, such as limited spatial resolution (Florinsky and Kuryakova, 2000; Lacroix et al., 2002), terrain characteristics (Fisher, 1998; Kyriakidis et al., 1999; Toutin, 2002), or due to the techniques used in data acquisition and raw data processing (e.g., Elkhrachy, 2017). If used to generate drainage networks, the errors in the DEMs are propagated to the drainage networks and also to the parameters derived from them. Therefore, these networks are not as accurate as the ones represented in detailed topographic maps or derived from regional or national DEMs. However, if changes are made to the GDEMs that improve their accuracy, this may also cause an improvement of the topographic parameters and drainage networks extracted from them.

The improvement of DEMs accuracy may be achieved through the combination of DEMs constructed by data acquired with different techniques, namely stereo imaging and interferometry (e.g., Roth et al., 2002) or even, proposed more recently, using data acquired by unmanned aerial systems or terrestrial sensors (e.g., Petrasova et al., 2017). However, these data may be difficult and expensive to obtain. This raises the attention to the currently available Volunteered Geographic Information (Goodchild, 2007), which is obtained by volunteers that collect many types of features, such us, photographs, geotags, geo-names or routes of the Earth surface, acting like sensors. Among the large number of projects that allow citizens to create geographic information, OpenStreetMap (OSM) is one of the most well-known (Goodchild, 2007; Haklay, 2010; Neis and Zielstra, 2014). This initiative allows citizens to create vector data representing different types of geographic features, such as buildings, forests, roads or rivers. Even though large amounts of data related to water bodies and water lines are available, few references are found about the use of these data for the improvement of the GDEMs. Schellekens et al. (2014) developed a study that allows the extraction of information from OSM for hydrological and hydraulic models. Using OSM data, the authors created gridded maps with fraction of paved and unpaved area and of open water in each cell. One of the features used corresponds to the "rivers" represented in OSM, also used in the present study. Another study developed by Monteiro et al. (2015) assessed the positional accuracy of the drainage networks derived from the SRTM and ASTER DEMs and compared them with the positional accuracy of drainage networks extracted from OSM. The results showed that the OSM watercourses had a higher positional accuracy than the drainage networks obtained from the ASTER and the SRTM GDEMs. The fact that the drainage networks available in OSM presented better positional accuracy than the ones derived from the GDEMs, even though they may not be complete, motivated

the current study, which aims to analyse whether the inclusion of OSM watercourses and elevation points extracted from the GDEMs enables the creation of improved DEMs from which drainage networks with higher positional accuracy can be extracted. The developed process is semi-automatic and uses only free geographic data, enabling the identification of drainage networks in regions where high accurate networks are not available or are not freely accessible.

The proposed methodology requires: 1) the extraction of water lines from OSM that represent natural drainage networks; 2) the extraction of elevation points from the GDEMs; 3) generating a new DEM using an interpolation method that creates hydrologically correct DEMs; and 4) extract the drainage networks from the resulting DEMs. To assess if the methodology improved the accuracy of the obtained drainage networks, the study areas used to test the methodology were chosen in regions where authoritative reference data was available. This enabled the assessment of the positional accuracy of the OSM water lines, and of the drainage networks extracted from the original and the transformed DEMs.

This paper is organized in six sections. In the first one an introduction of the theme and the framing is presented. The datasets used in the study are explained in section "DATASETS". Section "METHODOLOGY" is dedicated to the explanation of the proposed methodology. The study areas are presented and characterized in Section "STUDY AREAS". The results are shown and analysed in section "RESULTS" and finally, in section "DISCUSSION AND CONCLUSIONS", some discussion and conclusions are presented, referring the advantages and limitations of this methodology and future work.

DATASETS

Three different datasets were used in this study: 1) Two GDEMs, namely ASTER and SRTM; 2) data extracted from OSM corresponding to waterways; and 3) the rivers of Great Britain available in Ordnance Survey, which were used as reference data.

The GDEMs

The GDEMs ASTER and SRTM are free elevation datasets available in grid format, where the attribute of each cell is the elevation, with spatial resolutions of 30 and 90 metres, respectively. The grid format is appropriate to implement hydrologic tools (e.g., Zhang and Montgomery, 1994), hence their frequent use in hydrologic applications, such as in the morphometric characterization of basins or flood simulations (Arun et al., 2005; Freeman, 1991; Vieux, 1993).

These datasets are georeferenced with the World Geodetic Coordinate System 1984 (WGS-84) and are generated from satellite images. They have some advantages over the use of local or regional DEMs as they cover large regions of the Earth's surface and are accessible at no cost for the users. The ASTER GDEM is acquired by photogrammetric methods (Greve, 1996), while the SRTM GDEM is collected using Interferometric Synthetic Aperture Radar (InSAR) technology (Hanssen, 2001). These datasets have some limitations, such as their resolution, which restricts their use in applications where high accuracy is required, as their errors will propagate to the GDEM-derived attributes/products (e.g., Kenward et al., 2000; Rodriguez et al., 2006). Thus, the accuracy assessment of the GDEMs is an issue well covered in literature. Several studies have been made to assess the accuracy of the GDEMs in different regions of the world (e.g., Cook et al., 2012; Eckert et al.,

2005; Mukherjee et al., 2013; Varga and Bašić, 2013; Weydahl et al., 2007; Yadav and Indu, 2016) and also the accuracy of the topographic parameters extracted from them (Ebenezer, 2015; Lin et al., 2005; Liu et al., 2015; Monteiro et al., 2015; Mukherjee et al., 2013; O'Callaghan and Mark, 1984; Tarboton et al., 1991). These studies enable to conclude that the vertical accuracy of DEMs is affected by morphologic terrain characteristics (Holmes et al., 2000), by the land use and land cover, but also by the techniques used for data acquisition and processing. The radar technology, used in SRTM, has some advantages over optical techniques (Massonet and Feigl, 1998), including the fact that it uses an active system that includes the self-transmitting and receiving electromagnetic waves. Therefore, the images acquisition is independent of natural illumination conditions (Forkuor and Maathuis, 2012). Another advantage is that the SRTM images are not affected by clouds. However, the SRTM, as the ASTER, is also not able to penetrate the vegetation (Das et al., 2016; Sefercik, 2012).

OSM data

OSM is a collaborative project that creates and distributes free geographic data of and for the world. This project is designed to create and provide free spatial datasets based on volunteered efforts (e.g., Al-Bakri and Fairbairn, 2010). The OSM database is a collection of vector objects that consist of point, line and polygon features (Fonte et al., 2016). The OSM community is composed of over 3,000,000 registered users (http://www.openstreetmap.org/stats/data_stats.html), from all around the world, collecting and updating geographic information, which is immediately visible to all other users. OSM data are structured by tags (formed by a "key" and a "value") that represent objects on Earth, including roads, rivers, buildings, etc. A list of OSM tags is available in the OSM Wiki webpage at http://wiki.openstreetmap.org/wiki/Map_Features. However, volunteers may choose to use other tags.

The coverage and consistency of OSM data varies widely among different geographic areas (Bossard et al., 2000; Heymann, 1994), with urban zones having usually the highest coverage (Neis and Zipf, 2012). The quality assessment can be performed by comparing OSM data with authoritative geographic databases with higher accuracy, used as reference data. Several researchers have analysed the quality of OSM data using this methodology (e.g., Haklay, 2010; Zielstra and Zipf, 2010). Other approaches have however also been used, such as shape similarity with existing data as an indicator of quality (e.g., Mooney et al., 2010). In general terms, what is evidenced in the literature on the quality of OSM data is that the quality is quite heterogeneous across the globe and in some areas it even exceeds the quality of proprietary data (Brovelli et al., 2016). In the present study, the quality of OSM data is assessed by comparing the water lines position with the corresponding position in the Ordnance Survey data.

Reference data

The reference data used was downloaded from the Ordnance Survey web page (www.ordnancesurvey.co.uk), which is the British National Mapping Agency. This institution has some cartographic products available for free download in their web page, such as roads, names of places and rivers. The data used in this work was the drainage network of Great Britain, which is formed by 144,000 km of watercourses, including freshwater rivers, tidal estuaries and canals. These data are available in GML 3.2 and ESRI Shapefile formats, at scales ranging from

1/15,000 to 1/30,000, and are in the British National Grid Projected Reference System. The attributes of these data are: watercourse name; watercourse type and direction of water flow. This drainage network was used in this study as reference to assess the positional accuracy of OSM watercourses and the positional accuracy of the drainage networks extracted from the original and rebuilt DEMs.

METHODOLOGY

The methodology used in the work presented in this article was applied considering the following steps: 1) The extraction of drainage networks from the original GDEMs and the analysis of their positional accuracy; 2) The extraction from OSM of the waterways that correspond to natural drainage networks and the assessment of their positional accuracy; 3) The creation of new DEMs applying an interpolation method that uses data extracted from the original GDEMs and OSM watercourses; and 4) Extraction of new drainage networks from the rebuilt DEMs and the assessment of their positional accuracy. The methodology flowchart is shown in Figure 1.

Drainage networks extraction from the DEMs

For the extraction of the drainage networks the popular D8 flow algorithm was applied (e.g., Jenson and Domingue, 1988; Lin et al., 2008; Mark, 1984; Martz and Garbrecht, 1998; O'Callaghan and Mark, 1984; Tarboton, 1997). This method enables the extraction of drainage networks from grid elevation data. To correct the DEMs of possible artificial depressions (sinks) and peaks the "Fill" tool available in ArcGIS software was used (Planchon and Darboux, 2002; Tarboton et al., 1991). The matrix of flow direction was generated, where each cell gives the direction of water flow. Then, the flow accumulation matrix was computed, where the value of each pixel is the total number of pixels that drains to itself. To generate the watercourses a critical level (*CL*) needs to be defined. This threshold limit enables to establish if a pixel belongs or not to the drainage network. The pixels that have values of flow accumulation lower than the *CL* value do not belong to the network while the ones with higher values form the stream network. Therefore, lower threshold values generate more branched networks, as more pixels will be included in the network.

To choose the *CL* appropriate values for the drainage networks extracted from the ASTER and SRTM DEMs two factors were considered. One was the spatial resolution of the DEMs and the other was to obtain drainage networks with a

total length similar to the total length of the reference drainage network. To this aim, firstly the appropriate value for the SRTM DEM was selected having into consideration the total length of the reference network. Then, taking into account that one pixel of SRTM DEM corresponds to nine pixels of the ASTER DEM, the threshold value chosen for ASTER was nine times larger than that defined for SRTM. For the rebuilt DEMs derived from ASTER the threshold values were the same as the ones used for the original ASTER, and the same was done for the rebuilt DEMs derived from SRTM.

Extraction of waterways from OSM

To extract the waterways from OSM, an analysis of the tags available in OSM was performed. The tag with key "waterway" is used to describe different types of waterways, such as rivers, streams, drains and canals. As for this study only waterways corresponding to natural watercourses should be used, only the lines corresponding to the values "river" and "stream" were considered.

Creation of the rebuilt DEMs

The reconstruction of the DEMs was performed applying the ANNUDEM interpolation method that uses altimetry (elevation points), and lines (streams). This interpolation method, developed by Hutchinson et al. (2011), generates a surface hydrologically correct and is implemented in the ArcGIS tool "Topo to raster". The method was applied using the elevation points extracted from the GDEMs, which were obtained from the original GDEMs by converting them to the vector data model, extracting an elevation point corresponding to the centre of each raster cell. The linear features used corresponding to the streams were the ones extracted from OSM. The boundary of the study areas was also used to limit the obtained DEMs.

Assessment of the positional accuracy of the drainage networks

The horizontal position of the drainage networks extracted from the original GDEMs, the ones extracted from the rebuilt DEMs, and the ones extracted from OSM was compared to the horizontal position of the reference drainage network.

To calculate the horizontal distance between each set of the lines and the reference lines, a set of steps was performed. First, it was necessary to determine which lines of the drainage net-
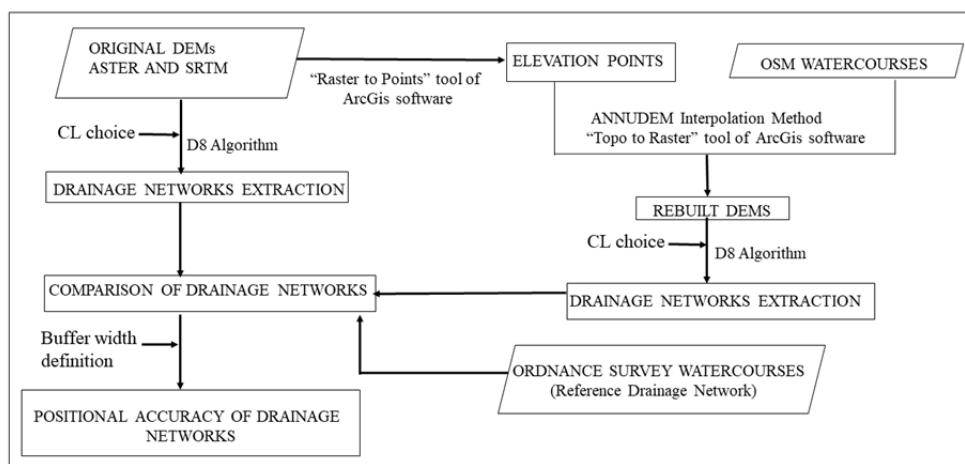

Fig. 1. Methodology flowchart.

work under analysis correspond to the reference lines. For this, a buffer was generated around each reference line, in order to identify the set of lines under study that are inside this buffer and that may be considered as corresponding waterways. The selection of a suitable buffer width is dependent on the dimension of the area and the relative positioning of the corresponding lines. Once this step is completed, to compute the distance between the identified corresponding lines two steps were used: 1) the tool "Feature Vertices to Points" of ArcGIS software was used to transform the selected lines into points (for example, the middle points or end-points of line segments forming the drainage network may be used); 2) the ArcGIS tool "Near" was used to compute the horizontal distance between the points of the waterway under analysis and the reference line. This tool computes the shortest distance between a point and a line.

This computation was only done for the points that are inside the buffers created around the reference drainage network. Then, for each drainage network, the mean M and the standard deviation σ of the obtained distances D_i (Figure 2) were calculated using equations (1) and (2), where n is the number of points considered for the waterway under analysis. The total length of the drainage networks L was also computed and compared.

$$M = \frac{\sum_{i=1}^{n} D_i}{n} \tag{1}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (D_i - M)^2}{n}} \tag{2}$$

STUDY AREAS

The methodology was applied to two regions located in distinct parts of the United Kingdom. Study Area 1 (SA1) corresponds to a drainage basin (in Scotland) with an area of 42 km^2 generated using the tool "Basin" of ArcGIS software applied to both ASTER and SRTM GDEMs, which produced slightly
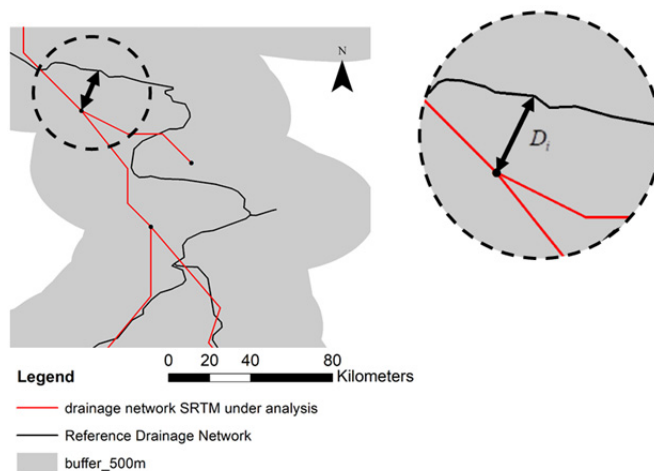

Fig. 2. Distance from a point of the drainage network derived from the DEM to the reference drainage network as computed with the tool "Near" of ArcGIS software.

different basins for the same region due to the differences in the DEMs. Study Area 2 (SA2) corresponds to a 13,899 km^2 rectangular area located in England (Figure 3). The SA1 region has a maximum altitude of 683 m and a minimum of 16 m, while the maximum and minimum altitude of SA2 is 212 m and 2 m, respectively. Both areas have different characteristics in terms of relief and dimension.

The location of the study areas and the ASTER and SRTM DEMs for each are shown in Figure 3.

The reference data (the drainage network of Ordnance Survey) obtained for each study area, are presented in Figure 4.

The OSM watercourses were downloaded in the shapefile format from Geofabrik in September 8th, 2015. These datasets are presented in Figure 5a) and 5b), respectively, for study areas SA1 and SA2. In terms of connectivity, it can be seen that the network obtained for SA2 shows several disconnected watercourses, while all watercourses are totally connected in SA1.

For the extraction of the drainage networks from the original and the rebuilt DEMs the appropriate critical levels had to be
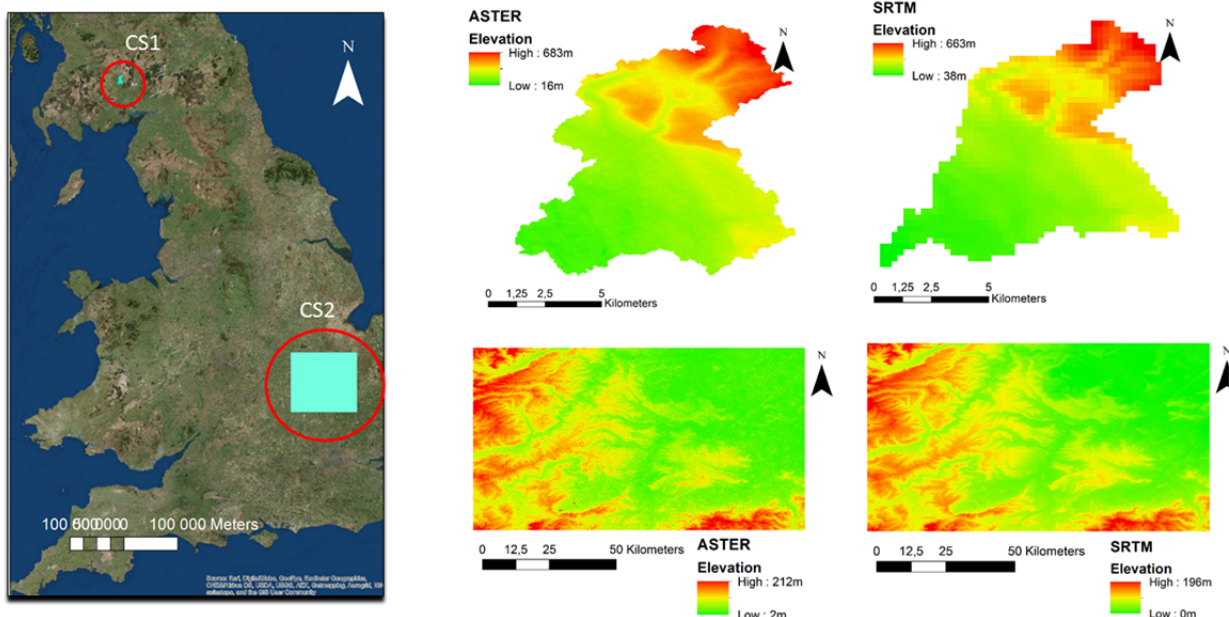

Fig. 3. Location and DEMs of both study areas, SA1 in North and SA2 in South of United Kingdom.
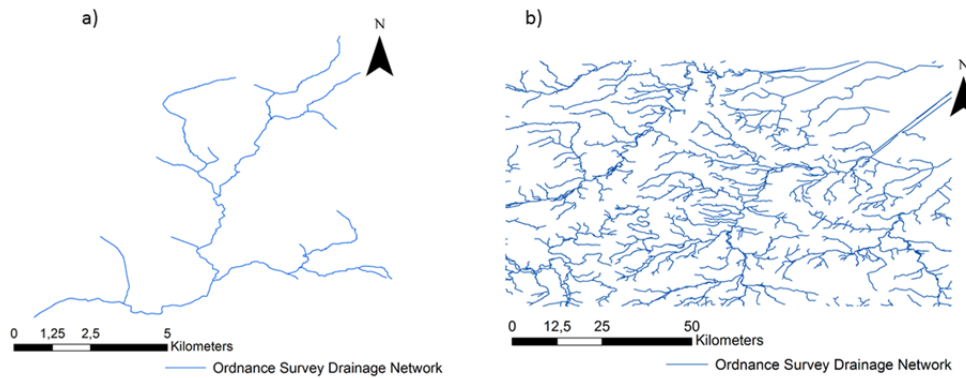
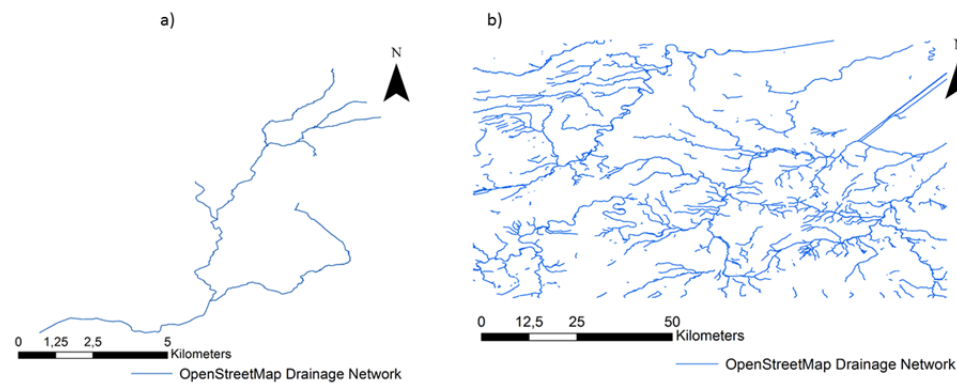**Fig. 4.** Drainage networks of the Ordnance Survey for study area 1 (SA1) (a) and for study area 2 (SA2) (b).

Fig. 5. OSM watercourses available for both study areas - SA1 (a) and SA2 (b).
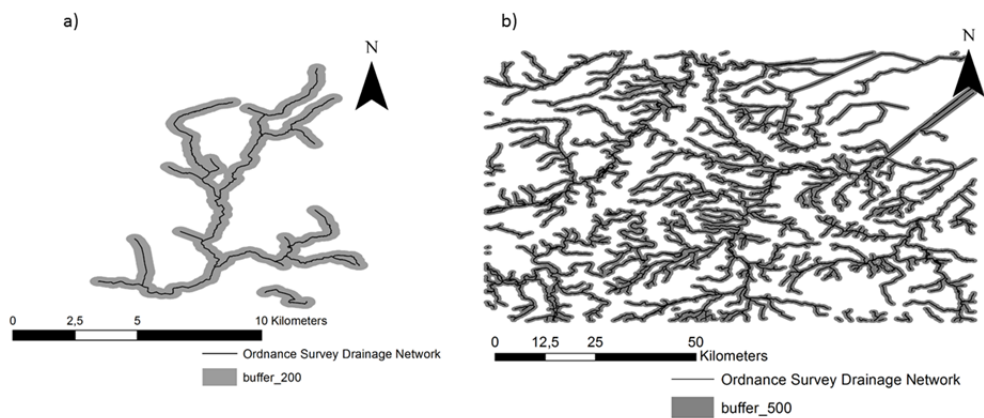

Fig. 6. a) A 200 m buffer around the reference drainage network in SA1; b) A 500 m buffer around the reference drainage network in SA2.

determined. Some tests were made to determine the appropriate *CL* for each SA. For SA1 the critical levels of 14 and 126 pixels were used, respectively, for SRTM and ASTER DEMs, derived as explained in section "Drainage networks extraction from the DEMs". For SA2 the values used were 445 and 4,000 pixels, respectively for the SRTM and ASTER DEMs. The horizontal distances between each drainage network and the reference drainage network was calculated for the parts of the network inside the buffer created as describe in section "Assessment of the positional accuracy of the drainage networks". The appropriate width of the buffers to consider was determined by visually analysis. This resulted in the use of 200 and 500 m for SA1 and SA2, respectively (Figure 6).

To compute the positional accuracy of the different drainage networks all data georeferenced in the World Geodetic System 1984 (WGS-84) was projected to the same projected Reference System (The British National Grid Reference System).

RESULTS

The drainage networks derived from the original and rebuilt DEMs ASTER and SRTM in both study areas are shown in Figure 7.

A visual analysis shows that for SA1 (Figure 7 a) and b)) the drainage network derived from the rebuilt SRTM DEM is more branched than the drainage network derived from the original DEM. The opposite situation occurs for the networks extracted from the ASTER models (original and rebuilt). Some changes in the horizontal position of the drainage network can also be seen. For the study area SA2 (Figure 7 c) and d)), the drainage network derived from the rebuilt ASTER DEM presents a higher similarity with the network extracted from the original ASTER DEM, namely in zones of lower altitude in the Northeast region of this study area.
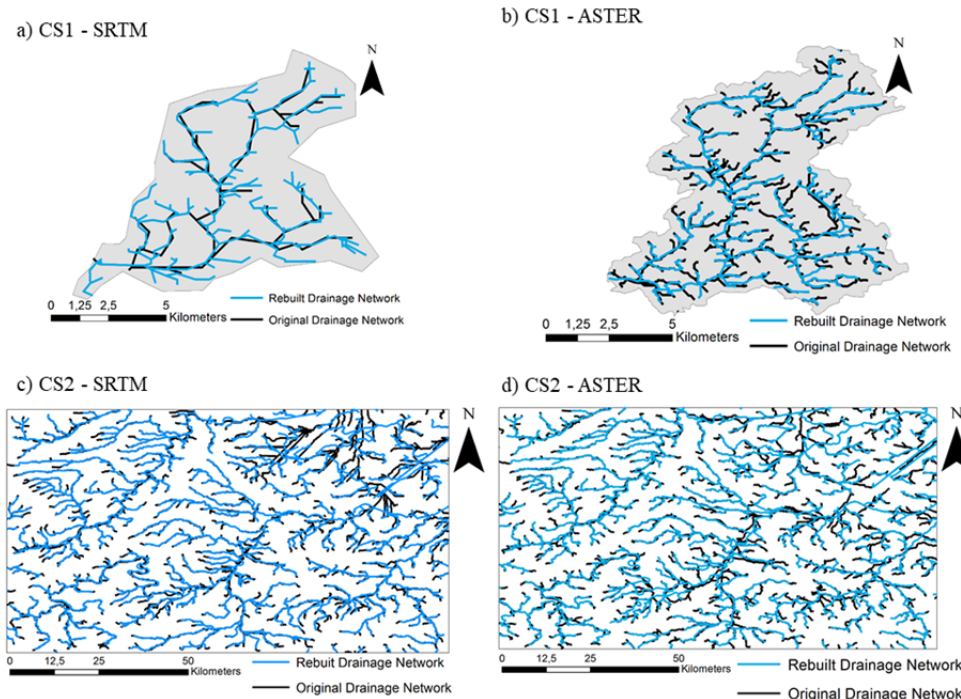
Fig. 7. Drainage networks extracted from the original and rebuilt SRTM DEMs, using the same critical levels a) for SA1 and c) for SA2; and from the original and rebuilt ASTER DEMs b) for SA1 and d) for SA2.

Some topographic parameters derived from the drainage networks were analysed and discussed. The parameters the order of the basin (O), the orders included in basin (O_i), the number of segments of i order (Ns_i) and the segments length of i order (L_i) and are shown in Table 1 for SA1 for the original and the rebuilt GDEMs. Table 2 show the results for the same parameters for SA2 for the original and rebuilt GDEMs.

In SA1 study area the order of the basin derived from SRTM DEM increased with the DEMs transformation presenting the values of $O = 3$ and $O = 5$, for the basins derived from original and rebuilt DEM, respectively. Also the number of segments and the segments length of each order show an increasing with the values of 34 and 140 numbers of segments for original and rebuilt SRTM DEM, respectively. For the basin derived from ASTER DEM the order maintains the value ($O = 5$), while the number of segments shows a decreasing from 330 to 140, as well as the segments length. In the SA2 the order of the basin varies with the same tendency of the SA1, for both DEMs with $O = 4$ and $O = 5$, for basins derived from original and rebuilt SRTM DEM, respectively. The basins derived from original and rebuilt DEM present the same order, $O = 5$. The number of segments decreased for basins extracted from both DEMs with the values of 1,150 and 772 for SRTM and the values of 1,281 and 776 for ASTER. The segments length also shows a decreasing for both DEMs. The analysis of OSM drainage networks positional accuracy was performed for each study area to assess whether OSM watercourses showed better positional accuracy than the drainage networks derived from the original GDEMs, as this assumption should be fulfilled to apply this methodology. Table 3 shows the mean (M) and standard deviation (σ) of the horizontal distances between the points of OSM watercourses and the reference drainage network. This table also shows the mean (M) and the standard deviation (σ) of the horizontal distances between the reference drainage network and the points of the drainage network obtained from the original ASTER and SRTM DEMs for both study areas.

Table 1. Basin order (O), number and length of segments for stream order O_i of the hydrographic network extracted from the original and rebuilt GDEMs for study area 1 (SA1).

SA1			
GDEMs		ORIGINAL	
	O_i	Number of segments	Length (km)
	1	18	21.2
SRTM	2	10	11.4
(O=3)	3	6	5.9
Total=		34	38.5
	1	170	62.5
ASTER	2	68	24.4
(O=5)	3	51	17.3
	4	35	11.9
	5	6	1.7
Total =		330	117.8

SA1			
GDEMs		REBUILT	
	O_i	Number of segments	Length (km)
	1	73	30.6
SRTM	2	34	18.9
(O=5)	3	16	10.3
	4	9	4.5
	5	8	5.4
Total =		140	69.7
	1	111	34.9
ASTER	2	44	25.3
(O=5)	3	34	9.2
	4	12	5.0
	5	10	1.7
Total =		211	76.1

It can be observed that the positional accuracy obtained for the OSM watercourses in both study areas is better than the positional accuracy obtained for the drainage networks derived from the original GDEMs, and therefore the above mentioned requirement is fulfilled.

Table 2. Basin order (O), number and length of segments for stream order O_i of the hydrographic network extracted from the original and rebuilt GDEMs for study area 2 (SA2).

SA2			
GDEMs		ORIGINAL	
	O_i	Number of segments	Length (km)
SRTM (O=4)	1	603	1,436.7
	2	299	665.5
	3	173	330.5
	4	75	149.0
Total =		1,150	2,581.7
ASTER (O=5)	1	637	1,484.4
	2	316	721.7
	3	163	312.3
	4	59	204.4
	5	43	75.7
Total =		1,218	2,627.5

SA2			
GDEMs		REBUILT	
	O_i	Number of segments	Length (km)
SRTM (O=5)	1	401	1,117.1
	2	202	523.1
	3	73	213.2
	4	63	215.4
	5	33	69.5
Total =		772	2,138.3
ASTER (O=5)	1	407	1,206.7
	2	213	605.3
	3	62	157.4
	4	64	214.8
	5	30	77.3
Total =		776	2,265.1

Table 4 shows the mean and the standard deviation of the horizontal distances between the reference drainage network and the points of the drainage networks derived from the original (M_O and σ_O) and the rebuilt DEMs (M_R and σ_R), for both study areas. The differences between the obtained these values for the networks obtained from the original are rebuilt DEMs are also shown.

The results show that the mean and the standard deviation of the distance between the DEM extracted drainage networks and the reference lines decreases with the transformation in both study areas and for both DEMs, which corresponds to an im-

Table 3. Mean and standard deviation of the horizontal distances between the reference drainage network (from Ordnance Survey) and the OSM watercourses, as well as the drainage networks obtained from the original DEMs, for study areas SA1 and SA2.

Study area	Data set	CL (pixels)	Buffer width (m)	Original GDEM	
				M (m)	σ (m)
SA1	OSM	---	200	10.5	28.0
	SRTM	14	200	45.4	47.0
	ASTER	126	200	38.7	55.4
SA2	OSM	---	500	20.7	57.3
	SRTM	445	500	92.8	116.6
	ASTER	4,000	500	104.9	127.8

provement of their positional accuracy. The results show a larger improvement in the mean distance for SRTM in SA1 (18.2 m for SRTM and 10.9 m for ASTER) and for ASTER in SA2 (48.9 m for ASTER and 38.4 m for SRTM). However, the opposite is observed for standard deviation, even though the values are very similar, in both cases with differences smaller than 1.5 m.

The topographic parameter total length of the drainage network (L) was analysed. Table 5 shows, for both study areas, the results obtained for the total length of the Ordnance Survey drainage network (reference data), the total length of the drainage networks obtained from the original and rebuilt DEMs and the differences between the obtained values.

It can be seen that using the same critical levels for the rebuilt DEMs that were used for the original models, the differences of total length obtained in SA1 between the reference drainage network (L_{OS}) and the drainage networks derived from the rebuilt DEMs (L_R) is –30.9 and –37.3 km for ASTER and SRTM, respectively. This shows that the drainage network obtained with the rebuilt DEMs is larger than the one available in Ordnance Survey and that this difference increased for SRTM, in relation to the difference between the reference drainage network and the drainage networks derived from the original DEMs ($L_{OS} - L_O$ is 0.3 km). However, for ASTER this difference decreased for the rebuilt DEM ($L_{OS} - L_O$ is –79 km, while $L_{OS} - L_R$ is –37.3 km). In SA2 the differences between the reference drainage network (L_{OS}) and the drainage networks derived from the rebuilt DEMs (L_R) are 595.5 and 472.3 km, respectively for the SRTM and ASTER DEMs. The difference between the drainage network obtained from the original and the rebuilt DEMs was also in both cases of the order of 400 km,

Table 4. Mean and standard deviation of the horizontal distances between the reference drainage network and the drainage networks obtained from the original and the rebuilt DEMs for study areas SA1 and SA2, as well as the differences between the values obtained for the original and rebuilt DEMs.

Study area	DEM	CL (pixels)	Buffer width (m)	Original GDEM		Rebuilt GDEM		Differences	
				M_O (m)	σ_O (m)	M_R (m)	σ_R (m)	$M_O - M_R$ (m)	$\sigma_O - \sigma_R$ (m)
SA1	SRTM	14	200	45.4	47.0	27.2	42.5	18.2	4.5
	ASTER	126	200	38.7	55.4	27.8	49.4	10.9	6.0
SA2	SRTM	445	500	92.8	116.6	54.4	89.5	38.4	27.1
	ASTER	4,000	500	104.9	127.8	56.0	102.0	48.9	25.8

Table 5. Total length of the Ordnance Survey drainage network (L_{OS}), of the drainage network extracted from the original DEMs (L_O) and of the one extracted from the rebuilt GDEMs (L_R), as well as the differences between the obtained values.

Study area	L_{OS} (km)	DEM used	CL (pixels)	L_O (km)	L_R (km)	$L_O - L_R$ (km)	$L_{OS} - L_O$ (km)	$L_{OS} - L_R$ (km)
SA1	38.8	SRTM	14	38.5	69.7	–31.2	0.3	–30.9
		ASTER	126	117.8	76.1	41.7	–79.0	–37.3
SA2	2,733.8	SRTM	445	2,581.7	2,138.3	443.4	152.1	595.5
		ASTER	4,000	2,722.8	2,261.5	461.3	11.0	472.3

and in both cases the length of the drainage network decreased when obtained from the rebuilt DEM.

These differences were obtained considering the same *CL* that was used for the original DEMs. However, the choice of the *CL* to use in for each DEM may also be defined for each DEM, and therefore the length of the drainage network will change with this value and can be adjusted according to the needs. To illustrate this, different critical values were used for each DEM, so that drainage networks with similar characteristics could be obtained. This was done first taking into account the same factors and procedures that were used for original DEMs, that is, choose the *CL* that produces drainage networks derived from SRTM with total length next to the total length of reference drainage network. After that, use a *CL* nine times greater than the one used for SRTM DEM to extract the drainage networks from ASTER. Also independent values were used for ASTER, to analyse the differences. The *CL* used and the obtained results are shown in Table 6 and the obtained drainage networks can be seen in Figure 8, for both study areas.

In SA1, a *CL* of 55 pixels was used for SRTM, generating a drainage network with a total length of 39.0 km, presenting a difference of length of only –0.2 km in relation to the reference network. Using a *CL* nine times greater for ASTER model (*CL* = 495 pixels) the drainage network extracted have a total length of 31.3 km, presenting a difference of 7.5 km in relation to the reference network. Choosing an independent value for the DEM rebuilt from ASTER of 250 pixels, a drainage network with a length of 38.9 km is obtained, corresponding to a difference of –0.1 km in relation to the reference one. For SA2 the *CL* of 255 pixels used for the DEM rebuilt from SRTM produced a drainage network with a total length of 2,752.4 km, presenting a difference of –18.6 km in relation to the reference drainage network. Using a *CL* nine times greater for DEM rebuilt from ASTER (2,295 pixels), a drainage network with a total length of 2,925.7 km is obtained, which corresponds to a difference of –191.9 km. If an independent *CL* of 2,650 pixels was used, produce a drainage network with a total length of 2,732.1 km is obtained, which corresponds to a difference of

Table 6. Total length of the Ordnance Survey drainage network (L_{OS}) and of the drainage networks extracted from the rebuilt DEMs (L_R) considering *CL* that provide drainage networks with characteristics similar to the reference drainage network, as well as the differences between the obtained values.

Study area	L_{OS} (km)	DEM used	CL (pixels)	L_R (km)	$L_{OS} - L_R$ (km)
SA1	38.8	SRTM	55	39.0	–0.2
		ASTER	495	31.3	7.5
		ASTER	250	38.9	–0.1
SA2	2,733.8	SRTM	255	2,752.4	–18.6
		ASTER	2,295	2,925.7	–191.9
		ASTER	2,650	2,732.1	1.7

1.7 km. The drainage networks shown in Figure 8 show that when *CL* are chosen that provide networks with similar length the networks obtained from the rebuilt DEMs are in fact very similar to the reference drainage networks.

DISCUSSION AND CONCLUSIONS

The aim of the current work was to determine whether the watercourses available in OSM data could contribute to the creation of DEMs, generated using the ASTER and SRTM GDEMs, which enable the extraction of drainage networks with higher positional accuracy. The methodology uses only free geographic dataset and shows that the combination of OSM watercourses with ASTER and SRTM GDEMs generates DEMs that allow the extraction of more accurate drainage networks.

The study was applied to two areas located in the United Kingdom with different characteristics in terms of relief, dimension, coverage and completeness of OSM watercourses, which are some factors that may influence the accuracy of drainage networks. Authoritative data (from Ordnance Survey) was used as reference to assess the positional accuracy of the data used and results obtained.
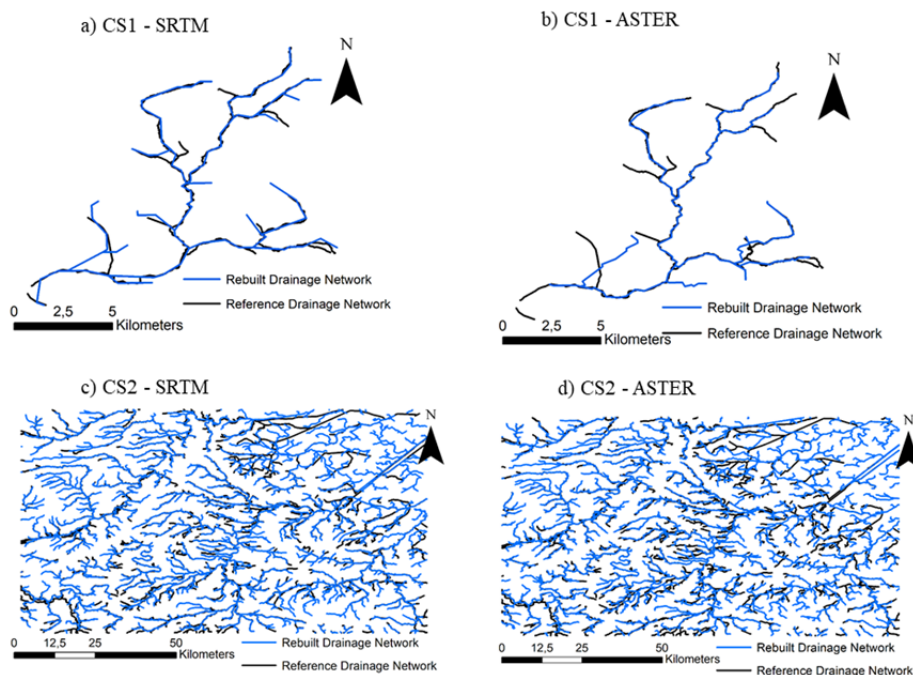

Fig. 8. Reference drainage networks and drainage networks extracted from the rebuilt DEMs, choosing the critical levels that approximate the length of the network under analysis to the reference network, for SRTM DEMs a) for SA1 and c) for SA2; and the rebuilt ASTER DEMs b) for SA1 and d) for SA2.

The results show that the positional accuracy of the drainage networks improved for both study areas and also for the data extracted from the SRTM and ASTER GDEMs, which proves the usefulness of the methodology. It can also be stressed that the positional accuracy improvement for both study areas is always higher for the data extracted from the SRTM DEM. That is, the data derived from this DEM takes more advantage of the transformation than those extracted from the ASTER DEM. The similarity of the obtained drainage networks with the reference network depends upon the *CL* parameter considered when extracting them from the DEMs, and therefore this value should be chosen independently for each DEM. This parameter also influences the topographic parameters derived from the drainage networks, such as the basin order as well as the number of segments and the segments length obtained for each stream order.

This methodology may be of particular interest for regions where there are no detailed and accurate topographic maps available, as the data sets used are free and available worldwide. It can also be useful when large areas are to be considered, covering more than one country, which would require merging detailed data from different institutions.

A fundamental aspect related to the application of this methodology is the positional accuracy of the OSM waterways. If mistakes are present in OSM data this will have negative impacts in the results. Therefore, the development of methods that enable the identification of erroneous data are particularly useful. Another limitation of the procedure is the coverage and completeness of OSM data, although the trend is that the volume of OSM data increases with an impressed speed (Haklay, 2010), and the waterways are one of the features that start to be inserted by the volunteers. Even if the completeness of OSM data in relation to the waterways is relatively low, the available waterways will likely still be useful, however this needs to be further investigated.

Future work will include the analysis of the impact of this transformation not only on the positional accuracy of the drainage network but also on other types of parameters extracted from the drainage network, such as bifurcation ratio and drainage density (Strahler, 1964), as well as the analysis of potential influence of the region's relief. Further investigations may analyse the variation of positional accuracy of those drainage networks in sub-areas of the global study area. An analysis of the accuracy of the rebuilt DEMs will be performed to ascertain whether they are also more accurate than the original DEMs.

REFERENCES

Al-Bakri, M., Fairbairn, D., 2010. Assessing the accuracy of 'Crowdsourced' data and its integration with official spatial data sets. In: Proc. Symp. Accuracy 2010. Leicester, UK, pp. 317–320.

Arun, P.S., Jana, R., Nathawat, M.S., 2005. A rule base physiographic characterization of drought prone watershed applying remote sensing and GIS. Journal of the Indian Society of Remote Sensing, 32, 189-201.

Bossard, M., Feranec, J. and Otahel, J., 2000. CORINE Land Cover Technical Guide: Addendum 2000. European Environment Agency, Copenhagen.

Brovelli, M.A., Minghini, M., Molinari, E., Zamboni, G., 2016. Positional accuracy assessment of the OpenStreetMap buildings layer through automatic homologous pairs detection: The method and a case study. The International Archives of the Photogrammetry. In: Proc. XXIII ISPRS CongressRemote Sensing and Spatial Information Sciences, vol. XLI-B2, , 12–19 July, Prague, Czech Republic.

Cook, A.J., Murray, T., Luckman, A., Vaughan, D.G., Barrand, N.E., 2012. A new 100-m digital elevation model of the Antarctic Peninsula derived from ASTER Global DEM: Methods and accuracy assessment. Earth System Science Data, 4, 129–142.

Das, S., Patel, P.P., Sengupta, S., 2016. Evaluation of Different Digital Elevation Models for Analysing Drainage Morphometric Parameters in a Mountainous Terrain: A Case Study of the Supin-Upper Tons Basin, Indian Himalayas. Springer Plus, 38 p. DOI: 10.1186/s40064-016-3207-0.

Ebenezer, S.S., 2015. GIS Based Automated Drainage Extraction for the Analysis of Basin Morphometry in Vaniyar Sub-Basin, South India. International Association of Scientific Innovation and Research (IASIR), pp. 31–34.

Eckert, S., Kellenberger, T., Itten, K., 2005. Accuracy assessment of automatically derived Digital Elevation Models from ASTER data in mountainous terrain. International Journal of Remote Sensing, 9, 1943–1957.

Elkhrachy, I., 2017. Vertical accuracy assessment for SRTM and ASTER Digital Elevation Models: A case study of Najran City, Saudi Arabia. Ain Shams Eng J. http://dx.doi.org/10.1016/j.asej.2017.01.007.

Fisher, P., 1998. Improved modelling of elevation error with geostatistics. GeoInformatica, 2, 215–233.

Florinsky, I.V., Kuryakova, G.A., 2000. Determination of grid size for digital terrain modelling in landscape investigations – Exemplified by soil moisture distribution at a micro-scale. International Journal of Geographical Information Science, 8, 815–832.

Fonte, C.C., Minghini, M., Antoniou, V., See, L., Patriarca, J., Brovelli, M.A., Milcinski, G., 2016. An Automated Methodology for Converting OSM Data into a Land Use/Cover Map. In: Proceedings of the 6th International Conference on Cartography & GIS, Albena (Bulgaria), June 13–17, 2016, Bulgarian Cartographic Association, 1, pp. 462–473. ISSN 1314-0604.

Forkuor, G., Maathuis, B., 2012. Comparison of SRTM and ASTER Derived Digital Elevation models over two regions in Ghana – Implications for hydrological and environmental modelling. In: Piacentini T. (Ed.): Studies on environmental and applied geomorphology. InTech, pp. 219–240.

Freeman, T.G., 1991. Calculating catchment area with divergent flow based on a regular grid. Computers and Geosciences, 17, 413–422.

Goodchild, M.F., 2007. Citizens as sensors: the world of volunteered geography. GeoJournal, 4, 211–221.

Greve, C. (Ed.), 1996. Digital Photogrammetry – Addendum to the Manual of Photogrammetry. Publication of the American Society of Photogrammetry and Remote Sensing.

Haklay, M., 2010. How good is the volunteered geographic information? A comparative study of OpenStreetMap and Ordnance Survey datasets. Environment and Planning B: Planning and Design, 37, 682–703.

Hanssen, R.F., 2001. Radar Interferometry: Data Interpretation and Error Analysis. Kluwer Academic Publishers, Dordrecht.

Heymann, Y., 1994. Commission of the European Communities, Directorate-General for Environment, N.S., C.P. CORINE Land Cover: Technical Guide. European Guide. European Commission. Directorate-General. Environment, Nuclear Safety and Civil Protection, Luxemburg.

Holmes, K.W., Chadwick, O.A., Kyriakidis, P.C., 2000. Errors in USGS 30-meter Digital Elevation Model and its impact on terrain modelling. Journal of Hydrology, 233, 154–173.

Hutchinson, M.F., Xu, T., Stein, J.A., 2011. Recent progress in the ANNUDEM elevation gridding procedure. In: Hengel, T., Evans, I.S., Wilson, J.P., Gould, M. (Eds.): Geomorphometry 2011, pp. 19–22.

Jenson, S.K., Domingue, J.O., 1988. Extracting topographic structure from digital elevation data for Geographic Information System analysis. Photogrammetry Engineering and Remote Sensing, 54, 1593–1600.

Kääb, A., 2005. Combination of SRTM3 and repeat ASTER data for deriving alpine glacier flow velocities in the Himalaya. Remote Sensing Environment, 4, 463–474.

Kenward, T., Lettenmaier, D.P., Wood, E.F., Fielding, E., 2000. Effects of Digital Elevation Model accuracy on hydrologic predictions. Remote Sensing and Environment, 3, 432–444.

Kyriakidis, P.C., Shortridge, A.M., Goodchild, M.F., 1999. Geostatistics for conflation and accuracy assessment of Digital Elevation Models. International Journal of Geographical Information Science, 7, 677–707.

Lacroix, M.P., Martz, L.W., Kite, G.W., Garbrecht, J., 2002. Using digital terrain analysis modelling techniques for the parametrization of a hydrologic model. Environmental Modelling Software, 17, 127–136.

Lin, W.T., Chou, W.C., Lin, C.Y., Huang, P.H., Tsai, J.S., 2005. Automated suitable drainage network extraction from Digital Elevation Models in Taiwan's upstream watersheds. Hydrological Processes, 20, 2, 289–306. DOI: 10.1002/hyp.5911.

Lin, W.T., Chou, W.C., Lin, C.Y., Huang, P.H., Tsai, J.S., 2008. WinBasin: using improved algorithms and the GIS technic for automated watershed modelling analysis from Digital Elevation Models. International Journal of Geographical Information Science, 22, 47–69.

Liu, T., Yan, H., Zhai, L., 2015. Extract relevant features from DEM groundwater potential mapping. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-7/W4, International Workshop on Image and Data Fusion, USA, pp. 113–119.

Mark, D.M., 1984. Automated detection of drainage networks from Digital Elevation Models. Cartographica, 21, 168-178.

Martz, L.W., Garbrech, J., 1998. The treatment of flat areas and depressions in automated drainage analysis of Raster Digital Elevation Models. Hydrological Processes, 12, 843–855.

Massonet, D., Feigl, K.L., 1998. Radar interferometry and its applications to changes in the Earth's surface. Reviews of Geophysics, 36, 441–500.

Monteiro, E.V., Fonte, C.C., de Lima, J.L.M.P., 2015. Assessing positional accuracy of drainage networks extracted from ASTER, SRTM and OpenStreetMap. In: Proceedings of AGILE 2015, Lisbon, Portugal, June 9–12, 5 p.

Mooney, P., Corcoran, P., Winstanley, A.C., 2010. Towards quality metrics for OpenStreetMap. In: SIGSPATIAL GIS'10, pp. 514–517.

Mukherjee, S., Joshi, P.K., Mukherjee, S., Gosh, A., Garg, R.D., Mukhopadhyay, A., 2013. Evaluation of vertical accuracy of open source Digital Elevation Model (DEM). International Journal of Applied Earth Observation and Geoinformation, 21, 205–217.

Mukul, M., Srivastava, V., Jade, S., Mukul, M., 2017. Uncertainties in the Shuttle Radar Topography Mission (SRTM) heights: Insights from the Indian Himalaya and Peninsula. Scientific Reports, DOI: 10.1038/srep41672, www.nature.com/scientificreports, 10 p.

Neis, P., Zipf, A., 2012. Analysing the contributor activity of a volunteered geographic information project – The case of OpenStreetMap. ISPRS International Journal of Geo-Information, 1, 146–165.

Neis, P., Zielstra, D., 2014. Recent developments and future trends in volunteered geographic information research: The case of OpenStreetMap. Future Internet 2014, 1, 76–106. DOI: 10.3390/fi6010076.

O'Callaghan, J.F., Mark, D.M., 1984. The extraction of drainage networks from Digital Elevation Data. Computer Vision, Graphics, and Image Processing, 28, 323–344.

Petrasova, A., Mitasova, H., Petras, V., Jeziorka, J., 2017. Fusion of high-resolution DEMs for water flow modelling. Open Geospatial Data, Software and Standards. DOI: 10.1186/s40965-017-0019-2.

Planchon, O., Darboux, F., 2002. A fast, simple and versatile algorithm to fill the depressions of Digital Elevation Models. Catena, 2, 159–176.

Rodriguez, E., Morris, C., Belz, J., 2006. A global assessment of SRTM performance. Photogrammetric Engineering and Remote Sensing, 72, 249–260.

Roth, A., Knopfle, W., Strunz, G., Lehner, M., Reinartz, P., 2002. Towards a global elevation product: Combination of multi-source Digital Elevation Models. In: Proceedings of Symposium on Geospatial Theory, Processing and Applications, Ottawa 2002, 5 p.

Schellekens, J., Brolsma, R.J., Dahm, R.J., Donchyts, G.V., Winesemius, H.C., 2014. Rapid setup of hydrological and hydraulic models using OpenStreetMap and the SRTM derived Digital Elevation Model. Environmental Modelling & Software, 61, 98–105.

Sefercik, U.G., 2012. Performance estimation of ASTER Global DEM depending upon the terrain inclination. Journal of Indian Society of Remote Sensing, 4, 565–576.

Strahler, A.N., 1964. Quantitative geomorphology of drainage basins and channel networks. In: Chow, V.T. (Ed.): Handbook of Applied Hydrology, pp. 4–39.

Tarboton, D.G., Bras, R.L., Rodriguez-Iturbe, I., 1991. On the extraction of channel networks from Digital Elevation Data. Hydrologic Processes, 5, 81–100.

Tarboton, D.G., 1997. A new method for the determination of flow directions and upslope areas in Grid Digital Elevation Models. Water Resources, 2, 309–319.

Toutin, T., 2002. Impact of terrain slope and aspect on radargrammetric DEM accuracy. ISPRS Journal of Photogrammetry and Remote Sensing, 57, 228–240.

Varga, M., Bašić, T., 2013. Quality assessment and comparison of Global Digital Elevation Models on the territory of Republic of Croatia. Cartography and Geoinformation, 20, 4–17.

Vieux, B.E., 1993. DEM aggregation and smoothing effects on surface runoff modelling. Journal of Computing in Civil Engineering, 3, 310–338.

Weydahl, D.J., Sagstuen, J., Dick, O.B., Ronning, H., 2007. SRTM DEM accuracy over vegetated areas in Norway. International Journal of Remote Sensing, 16, 3513–3527.

Yadav, S., Indu, J., 2016. Estimation of vertical accuracy of Digital Elevation Models over complex of Indian Subcontinent. IGARSS, 978-1-5090-3332-4/16, IEEE, 6036-6039.

Zhang, W., Montgomery, D.R., 1994. Digital Elevation Model grid size, landscape representation, and hydrologic simulation. Water Resources Research, 30, 1019–1028.

Zielstra, D., Zipf, A., 2010. OpenStreetMap quality research in Germany. In: Sixth International Conference on Geographic Information Science, pp. 15–17.

Note: Colour version of Figures can be found in the web version of this article.