Computers & Geosciences 92 (2016) 9–20

Contents lists available at ScienceDirect

Computers & Geosciences

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

Research paper

A new artefacts resistant method for automatic lineament extraction using Multi-Hillshade Hierarchic Clustering (MHHC)

Jakub Šilhavý [a,*], Jozef Minár [b], Pavel Mentlík [c], Ján Sládek [d]

[a] Section of Geomatics, Department of Mathematics, Faculty of Applied Sciences, University of West Bohemia in Pilsen, Univerzitní 22, Pilsen 306 14, Czech Republic
[b] Faculty of Natural Sciences, Comenius University, Bratislava, Slovak Republic
[c] Centre of Biology, Geoscience and Environmental Education, Faculty of Education, University of West Bohemia, Pilsen, Czech Republic
[d] Department of Physical Geography, Geomorphology and Natural Hazards, Institute of Geography, Slovak Academy of Sciences, Štefánikova 49, Bratislava 814 73, Slovak Republic

ARTICLE INFO

Article history:
Received 26 August 2015
Received in revised form
25 March 2016
Accepted 28 March 2016
Available online 31 March 2016

Keywords:
Lineaments
GIS
Automatic extraction
Bohemian forest
Central Western Carpathians

ABSTRACT

This paper presents a new method of automatic lineament extraction which includes the removal of the 'artefacts effect' which is associated with the process of raster based analysis. The core of the proposed Multi-Hillshade Hierarchic Clustering (MHHC) method incorporates a set of variously illuminated and rotated hillshades in combination with hierarchic clustering of derived 'protolineaments'. The algorithm also includes classification into positive and negative lineaments. MHHC was tested in two different territories in Bohemian Forest and Central Western Carpathians. The original vector-based algorithm was developed for comparison of the individual lineaments proximity. Its use confirms the compatibility of manual and automatic extraction and their similar relationships to structural data in the study areas.

1. Introduction

Lineaments are generally considered to be linear features manifesting in the land surface and land cover reflecting discontinuities of geological structures (mainly faults). Although various phenomena can form lineaments (rock boundaries, sedimentary layers, wetness and vegetation changes – see e.g. Gupta, 2003), distinct linear landforms are most frequently used to extract geological structures (Smith and Clark, 2005; Smith and Wise, 2007; Evans, 2012). If lineaments detection is based exclusively on the morphometric properties of the land surface, then the lineaments can be termed 'morpholineaments' (Minár and Sládek, 2009). Although morpholineaments are automatically extracted either directly from a Digital Elevation Model (DEM) (Vaz, 2011; Mallast et al., 2011) or from different derived surfaces, e.g. second derivatives of DEM (Wladis, 1999), shaded relief (hillshade) is the most frequently used derived surface (Abdullah et al., 2010; Masoud and Koike, 2011a; Jordan and Schott, 2005).

Image pre-processing (edge enhancement, noise removal using thresholding) followed by edge linking methods (Hough Transform, Canny edge detector) are mostly used for automated lines extraction (Table 1). In some cases, the pre-processing is part of the extraction (closed-source software modules).

Morpholineaments can be considered not only as a surface expression of particular lithospheric faults, joints and lithological boundaries (e.g. Solomon and Ghebreab, 2006; Štěpančíková et al., 2008; Batayneh et al., 2012), but also as an expression of a morphotectonic field – a manifestation of lithospheric stress fields in the landforms (Urbánek, 2005; Minár and Sládek, 2009; Sládek, 2010). When producing a morphotectonic field model, even a small artificial misrepresentation of the morpholineaments direction (artefacts) can lead to problematic interpretations of results. Artefacts formation during a raster based analysis is pointed out and solved in this paper.

The main objective of this paper is to present a new Multi-Hillshade Hierarchic Clustering (MHHC) artefacts resistant method for automated lineaments extraction. The second goal is to evaluate the correlation between automatically and manually delineated lineaments, test the algorithm's ability to detect linear geological features (such as faults and linear parts of rock boundaries) and extract the main tectonically significant directions for their following evaluation in morphotectonic analysis.

Although subjective visual assessment is the most common approach for validation of extracted lineaments (Kageyama and Nishida, 2004; Jordan and Schott, 2005), more objective approaches have been published. For example, Abdullah et al. (2010) computed simple statistics of count and length of lineaments to

* Corresponding author.
E-mail address: jakub.silhavy@gmail.com (J. Šilhavý).

http://dx.doi.org/10.1016/j.cageo.2016.03.015
0098-3004/

compare different datasets. Vaz et al. (2012) implemented the confusion matrices approach and the distance between lineaments and reference point data was calculated to prove correlation with ground truth datasets as a comparison metric by Mallast et al. (2011). A new method for comparing different lineament datasets has been developed in this study.

2. Data

Two geologically and geomorphologically different study areas were selected (Fig. 1): A) surroundings of Prášilské jezero (lake) in the Bohemian Forest (BF) and B) the boundary area between Žiar (Mts.), Malá Fatra (Mts.) and Turčianska kotlina (basin) in the Central Western Carpathians (CWC).

The input DEMs were generated using 5 m equidistant contour lines from topographic maps – 1:10,000 for CWC and 1:25,000 for the BF area. The vector contour lines were processed using the Create Hydrologically Correct DTM tool (Jedlička et al., 2015).

Existing, manually delineated morpholineaments (from the same source DEMs) were used for comparison (Minár and Sládek, 2009; Mentlík, 2006). This selection allowed the algorithm to be tested in different geological and geomorphological settings and, moreover, it provided the variability of the creation of manually delineated morpholineaments. Faults and lithological boundaries adapted from 1:50,000 geological maps (Pelc and Šebesta, 1994; Káčer et al., 2005) were used for expressing relationships to the geological structure.

Table 1
Overview of several approaches and software products for automatic lineament extraction.

Reference	Approaches and software products
Pradhan et al. (2010)	Manual extraction method based on automatically pre-processed images with enhanced edges
Abdullah et al. (2010)	PCI Geomatica with module LINE
Mallast et al. (2011)	ERDAS Imagine modules and PCI Geomatica
Argialas and Mavrantza (2004)	Optimised Hough Transform method (Fitton and Cox, 1998)
Soto-Pinto et al. (2013)	Hough Transform and software LESSA (Zlatopolsky, 1992)
Vaz (2011)	Wavelet edge analysis and morphological multi-scale gradient
Kageyama and Nishida (2004)	Hierarchical lineament detection method using land cover information in mixed pixels (mixels)
Masoud and Koike (2011a)	Segment Tracing Algorithm (STA) of Koike et al. (1995)

3. MHHC method for automated lineament extraction

MHHC is composed of six steps (Fig. 2): 1) Creation of DEM (P1), 2) Derivation of hillshades from DEM (P2a) plus their rotation (P2b), 3) Line extraction based on edge detection (P3), 4) Noise reduction (P4), 5) Cluster line analysis (P5), 6) Classification of lineaments (P6). The user driven parameters are marked by P in Fig. 2 and also in the text.

The algorithm was written in Python using the functionality of ArcGIS tools via ArcPy library. Line extraction was handled by PCI Geomatica software. The EASI scripting language was used to control PCI Geomatica tools.

3.1. DEM creation

The algorithm works with three types of input data:

1. the vector elevation data (contour lines and elevation spots),
2. the vector LIDAR data (elevation spots),
3. the raster DEM.

The output raster DEM is generated (or resampled in the case of the input raster DEM) within the boundary of the study area and with a specified output cell size. The spatial resolution is chosen depending on the scale and purpose of the analysis (P1 in Fig. 2). The raster pre-processing is part of the line extraction step which uses the LINE module (see Section 3.3 for details).

3.2. Hillshades creation

Masoud and Koike (2011a), Mallast et al. (2011) and Abdullah et al. (2010) mentioned the dependency of results on the

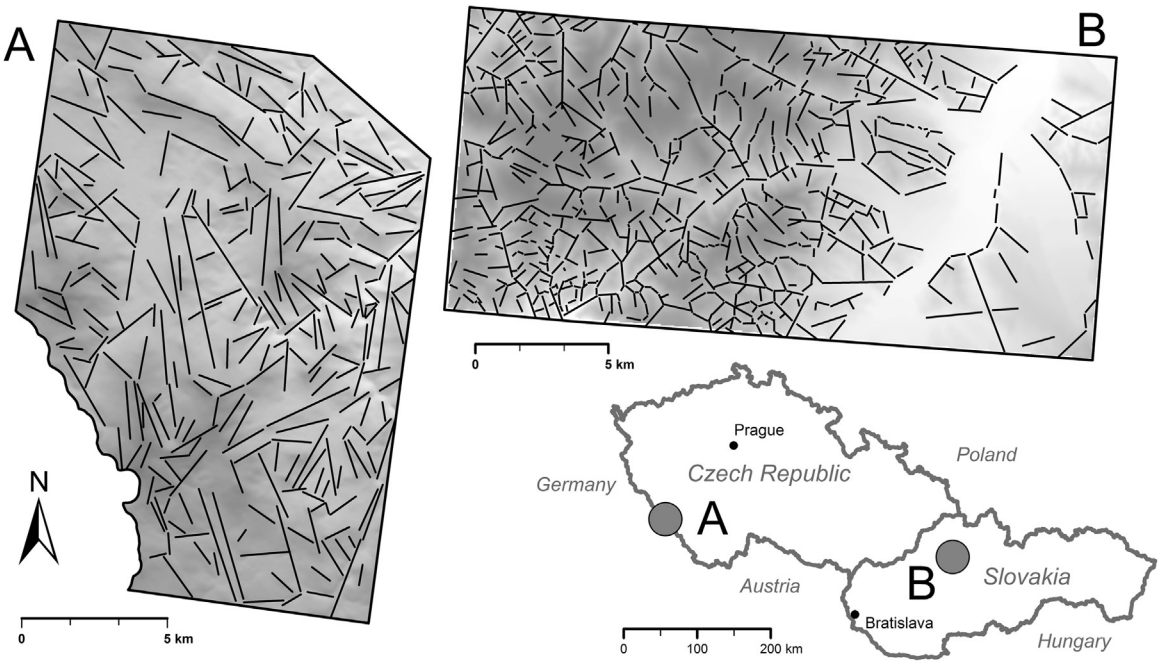

Fig. 1. Manually determined morpholineaments in both study areas. A) Bohemian Forest (surroundings of Prášilské jezero – lake). B) Central Western Carpathians.

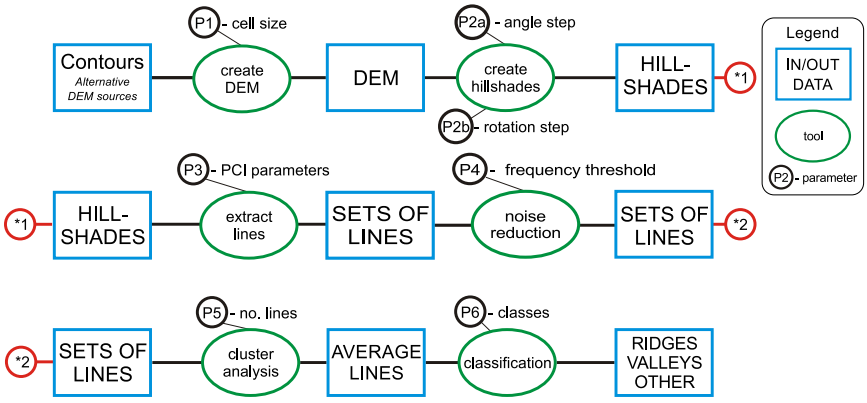

Fig. 2. The workflow of a new method for automated lineaments extraction. See the text for further explanation.

illumination azimuth when the shaded relief is used for lineaments extraction. These authors combine differently illuminated rasters into one raster to avoid this azimuth bias. The variability of the shaded reliefs is utilised as an advantage in our method. The differently illuminated shaded reliefs are used to extract different results which are processed separately.

As the hillshade raster resolution is the same as the input DEM, spatial resolution, illumination altitude (height of light source) and illumination azimuth (angle of light source) are parameters which influence the shaded relief. The illumination altitude influences the image contrast and depends on terrain characteristics. However, a major change to the parameter value causes only a minor change in the results. A constant value of 30° was chosen for our study areas.

The illumination azimuth has a significant impact on the results (Seleem, 2013), and its variations range from 0° to 360° using step P2a (Fig. 1). Decreasing the value of P2a gives statistically more reliable results but it is more time consuming. The value 15° was found to be the optimum between time and reliability. Each of the 24 hillshade rasters is the input for the line extraction module.

3.2.1. *Hillshades rotation – identification of artefacts and proof*

Directional analyses of the extracted lines after step 3 show their relation to the eight directions of a regular raster. Two datasets were used for testing these artefacts. The first dataset was derived from hillshade with illumination angle 0° and the second dataset was derived from the same raster, but rotated by 15° before extraction. The test proved that the extraction algorithm slightly adjusts lines to fit the eight main directions (see Fig. 3). In most cases, the values near the main direction are adjusted to fit this direction.

This effect has a minimal impact on the line length and position but strongly influences the directional analysis – see Fig. 4. Even

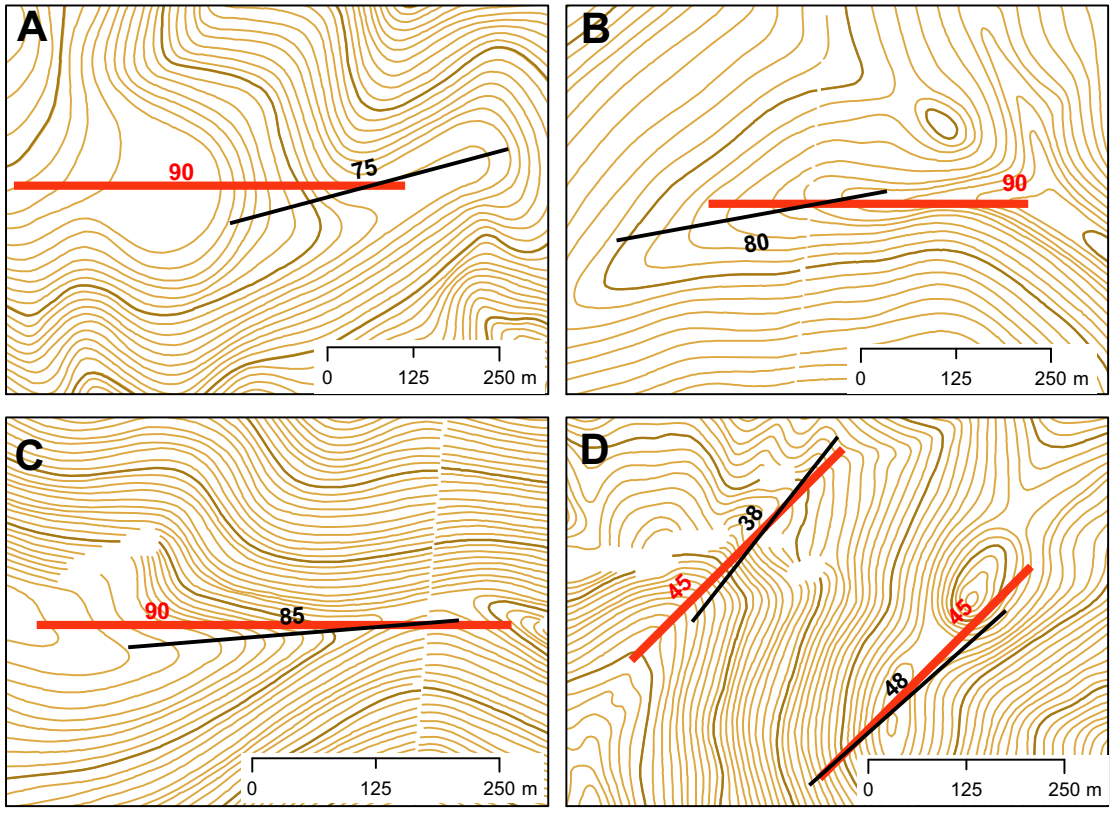

Fig. 3. Artefacts proof – examples of extracted lines from two different datasets. The bold line dataset prefers directions 0°, 45°, 90° and 135°. The narrow line dataset was derived from a raster rotated by 15° (the resulting lines were rotated back by 15° in order to compare their positions). In cases A) and B) the algorithm chooses different parts of the lineament feature to prefer artefact angle 90°. In cases C) and D) the algorithm adjusts the line orientation to fit the main direction.

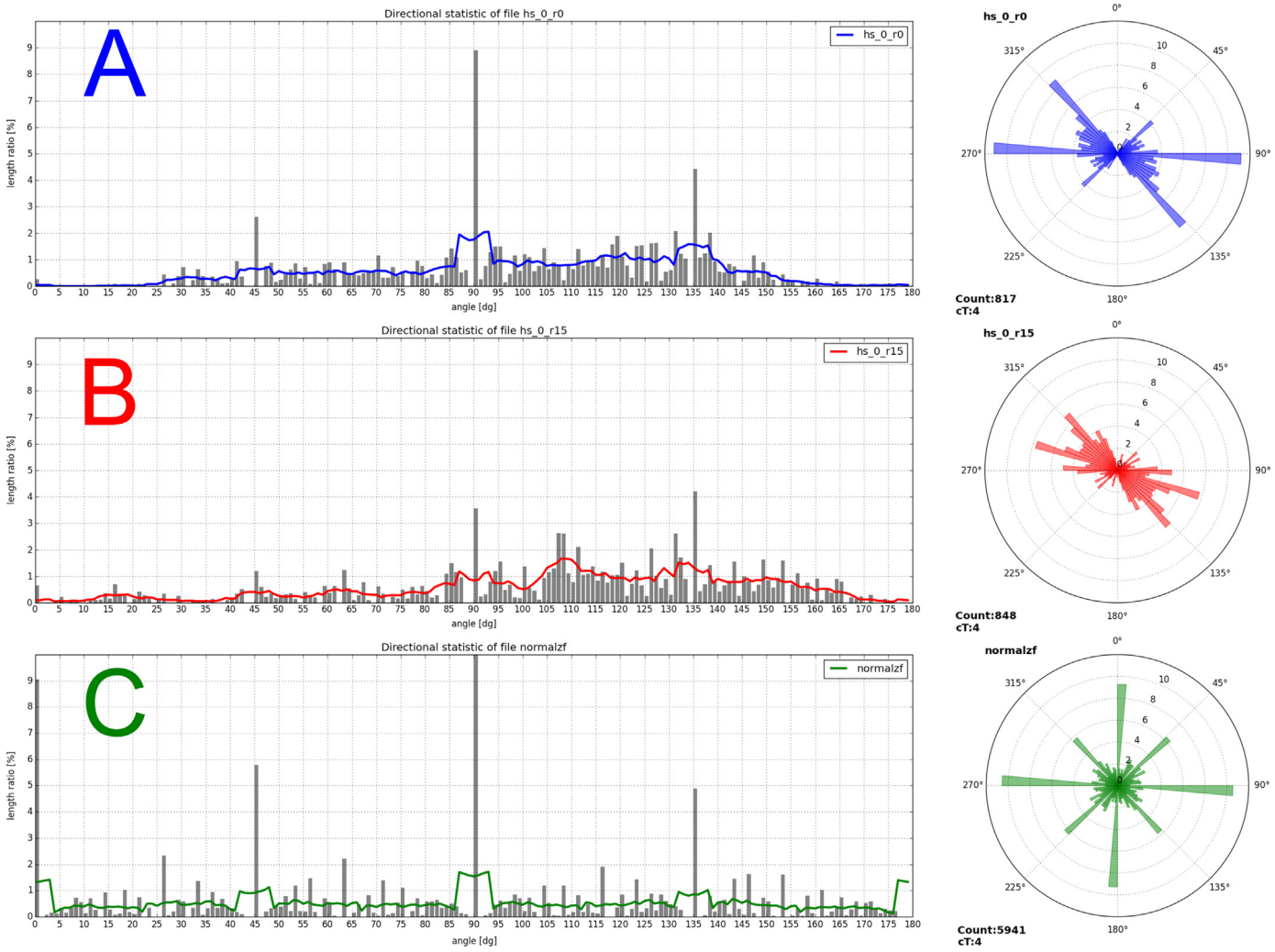

Fig. 4. Artefacts proof – Directional statistic of extracted lines from different rasters. A) Hillshade with illumination angle 0° (peaks 45°, 90°, 135°)*. B) Raster A rotated by 15° (peaks 45°, 90°, 135°)*. C) Raster with randomly generated values (peaks 0°, 45°, 90°, 135°). * In A) and B) the peak 0° is eliminated due to angle of illumination of hillshade.

when the input raster is rotated by 15° (B), the directional histogram has the same peaks as the original raster (A). The raster with randomly generated values clearly shows how the dominant directions are affected by the artefacts (C).

To avoid artefacts formation and misrepresentation of the results the hillshade is rotated. The combination of lines derived from differently rotated rasters eliminates the artefacts in each dataset. The lines are combined using the method described in steps 4 and 5. The rotation angle ranges from 0° to 45° using step P2b (Fig. 2) as the user parameter. The range is given by the interval between two adjacent artefact angles. The tests proved that artefacts are eliminated using a minimum of 5 rotated hillshades (equal to the P2b value of 9°).

3.3. Line extraction

As well as PCI Geomatica software with the LINE module, the Hough Transform (HT) variant used by Fitton and Cox (1998) was tested for the line extraction step.

The HT has a tendency to elongate lineaments beyond clear morphological borders and (in some cases) to create near parallel twin lineaments. Thus the PCI software was chosen to carry out the line extraction step.

The workflow of the LINE algorithm comprises three steps. The edge detection operator (Canny edge detector) is followed by the thresholding producing the binary edge raster. The edge image is processed by many substeps in order to extract the vector lines. A detailed description of the workflow and parameter setting of the LINE module is covered in the internal PCI Help (PCI Geomatics, 2011), papers by Abdullah et al. (2009) and Mallast et al. (2011). These authors used the linking and generalising abilities of the LINE module. In our tests, these options gave noisy and meaningless results (thick light lines in Fig. 5) which led us to eliminate these parameters (see the values of ATHR, DTHR and FTHR in Table 2). The other parameters were set up after several trials (Table 2, parameters P3 in Fig. 1).

3.4. Noise reduction

While Argialas and Mavrantza (2004) and Wladis (1999) reduced noise using different techniques such as thresholding, Mallast et al. (2011) applied morphological operations to the raster data. Koike et al. (1995) eliminated every pixel in a raster which did not satisfy the dynamic distance threshold.

In MHHC, the data to be reduced is represented by a huge number of vector lines – 'protolineaments'. To remove all non-relevant lines, the relevance is defined by the presence of a sufficient number of lines with similar lengths and orientations at the same place. The sufficient number of lines represents the 'frequency threshold' (parameter P4 in Fig. 2) and depends on the total number of line sets

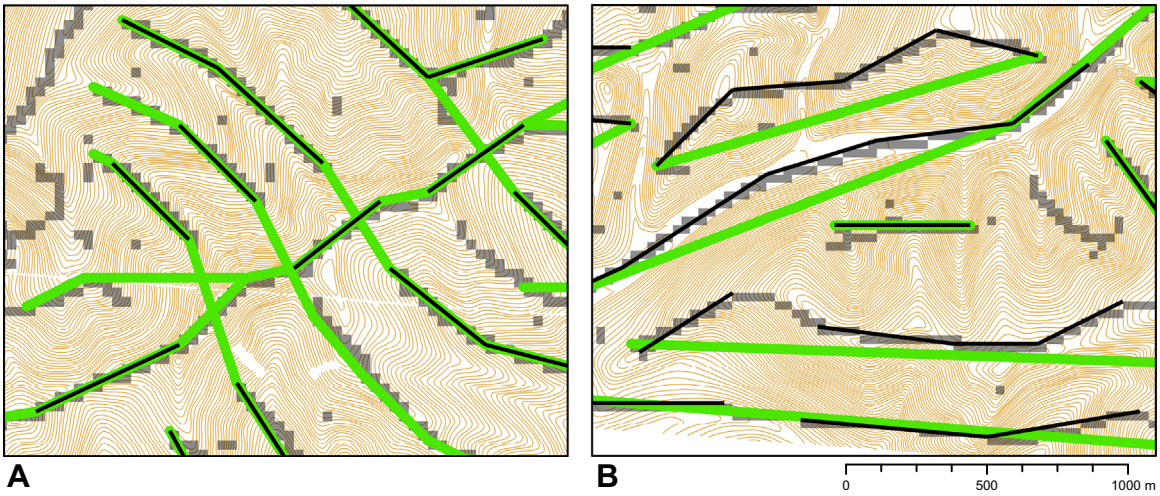

Fig. 5. The influence of parameters on converting edge pixels to lines. A) Thin dark lines – linking switched off, thick light lines – linking switched on. B) thin dark lines – generalization switched off, thick light lines – generalization switched on.

Table 2
Table of parameters of line extraction step.

Parameter	Used value	Result impact
Binary threshold (GTHR)	10 px	Count and length of lines
Length threshold (LTHR)	10 px	Minimum line length
Radius of Gaussian filter (RADI)	10 px	No significant impact
Generalisation (FTHR)	1 px	Spatial accuracy
Angle threshold (ATHR)	0°	Line linking
Distance threshold (DTHR)	0 px	Line linking

(according to parameter P2). It is recommended to choose this parameter interactively based on a visual interpretation of the results after applying noise reduction.

Subsequently, the raster approach is applied to eliminate solitary (or isolated) lines. To achieve this, all vector line sets are converted to binary rasters. The conversion is applied to buffers created around each line in order to set spatial tolerance. The buffer size is chosen with respect to raster spatial resolution and the characteristics of the study area. The 'raster of relevance' is calculated from the raster values in the corresponding cells where high values represent areas with a high occurrence of lines (see Fig. 6 A). The information from the 'raster of relevance' is transferred to each vector line in all the sets and the frequency threshold is applied to every line. Only lines with a higher value for the 'raster of relevance' are preserved. The logical core of this method is also used to classify the final lineaments (see Section 3.6).

Although the raster approach is faster, not every line is well evaluated in comparison with the presented line by line (vector) solution (bold line in Fig. 6B). Noise removal is necessary in order to decrease the total number of lines before the cluster line analysis.

3.5. Cluster line analysis

As can be seen from Fig. 6b, the lines form clusters after lines extraction (P4, Fig. 2). This step is performed to recognise all the clusters obtained after the extraction and to replace them with a representative single line. This approach is similar to Mallast et al. (2011) and Jordan and Schott (2005), who singularised identical results from different sources. However, these authors

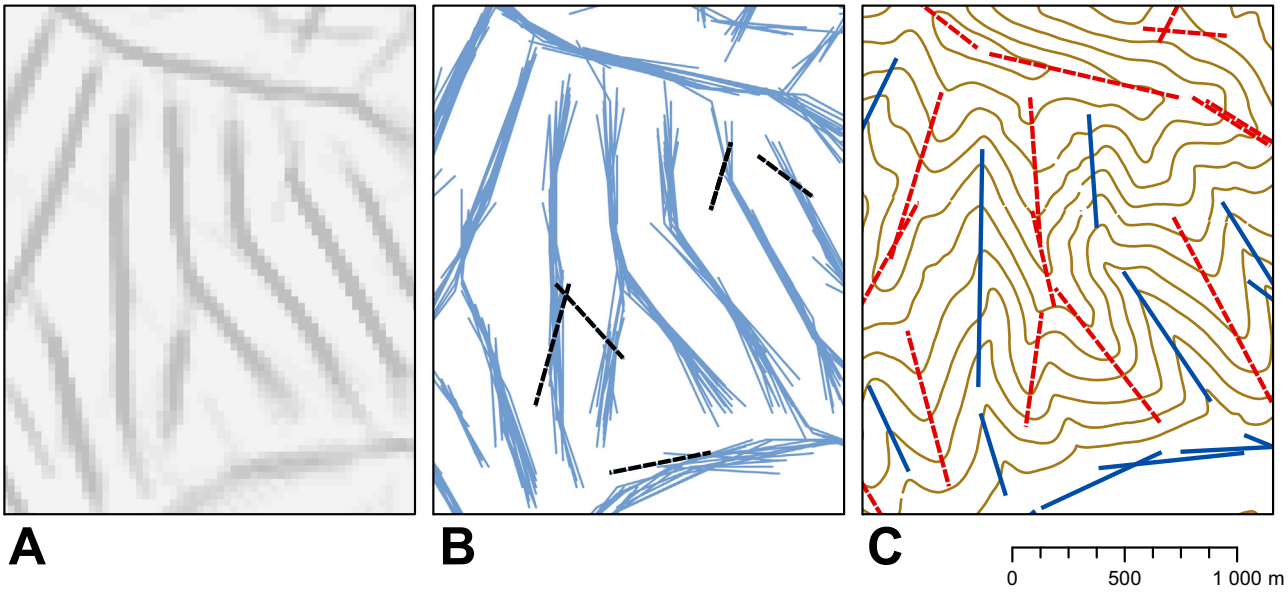

Fig. 6. A) Noise reduction – 'raster of relevance'. B) Sets of lines obviously creating clusters, except several lines marked in bold. C) Classification into positive (dashed line) and negative (solid line) lineaments.

do not use an automated method, only a visual assessment. Koike et al. (1998) and Kim et al. (2004) also grouped the lineaments but for a different purpose. They connected subsequent lineaments to obtain a single long lineament. No cluster algorithms reviewed by Skála and Kolingerová (2011) were suitable for this specific task, so a special algorithm was developed.

To facilitate the analysis, all line sets are merged into one layer and the statistics of azimuth and length are calculated.

The merged set is sorted into descending line length and the following workflow is applied to each set:

(1) Choose the first (longest) line from the set.
(2) Make a buffer around the chosen line.
(3) Select all lines which are completely within the buffer.
(4) Select lines which have azimuth in the range $\pm 20°$ from the chosen line.
(5) If the selection contains more than 4 lines (parameter P5 in Fig. 2), continue to step 6, otherwise continue to step 7.
(6) Create a buffer around selected lines (=cluster) with the following attributes: count of selected lines, average length, average azimuth.
(7) Delete all selected lines from the set.
(8) Repeat from step 1.

The values of the algorithm's variables (buffer size, count threshold, azimuth condition) were chosen based on the characteristics from the study areas (average width, count and azimuth dispersion of line bundles) and with respect to the expert morpholineaments resolution (to achieve a comparable level of detail).

The final clusters are saved as polygonal features with all required attributes. Post-processing makes an average line from each cluster using the centroid of the polygon and the average azimuth and average length. The set of the average lines presents the final lineament layer (Fig. 6C).

3.6. Classification of lineaments

The extracted lines have their origin in the discontinuities of the raster image. The geomorphological interpretation of the extracted lines depends on the geometrical character of the discontinuities. To classify the different geomorphological structures, the positive lineaments (such as ridge lines or convex edges) and negative lineaments (such as valley lines or concave edges) are defined (Fig. 6C). For a different geomorphological interpretation of these types see e.g. Ramli et al. (2010).

Abdullah et al. (2010) distinguished positive and negative lineaments using a specific combination of illumination azimuth within their study area. Because of the dependence of the illumination azimuth on the azimuth of the extracted lines, this approach can only be applied for a location with a specific orientation of positive and negative features.

To solve this problem, we evaluated the proximity of the extracted lineaments to the channel network (high proximity meant negative lineaments). In step 4 (noise reduction), the algorithm which assesses the frequency of lines is presented. A similar approach is adapted here. Instead of the 'raster of relevance', the 'water accumulation raster'[1] is used as an input. For each line, the mean and median value of 'water accumulation' is computed. These two statistics are used to differentiate between lines on ridges and thalwegs respectively. Because there is increased water accumulation on the foothill lines, these important lineaments are classified as negative lineaments. The median is required to avoid uncertainty in some specific cases where the mean could be influenced by a few pixels with a high value of accumulation.

[1] Created by flow accumulation tool (ESRI, 2013).

4. Comparative methods

Datasets of automatically extracted lineaments and control datasets were compared by correlation based on the geometric comparison of individual lines and by evaluation of the local maxima of length weighted directional histograms of all datasets.

4.1. Correlation method

The main principle of vector-based comparison of various line datasets is to imitate visual assessment. In other words, for each test line the method tries to find a corresponding reference line (Fig. 5).

For each line from dataset A, the algorithm finds similar lines from dataset B (Fig. 7 A). The similarity is defined by spatial vicinity and azimuth tolerance. Only the part of the reference line which is within the defined distance is taken into the computation (Fig. 7 B). The correlation index (CI) is computed as the ratio of the length of the clipped part of line B to the length of line A (Fig. 7 D). If the ratio is greater than 1, the value of CI is considered to be 1 ($CI \leq 1$) to ensure that correlation does not exceed 100%.

The method is driven by two parameters, the search radius (size of the line's buffer) and azimuth tolerance. The search radius parameter presents the maximum distance between two lines which are to be considered as similar. The azimuth tolerance parameter expresses the maximum deviance between line azimuths to be considered as similar (Fig. 7 C). A search radius of 200 m and azimuth tolerance 20° was used in the following analyses according to the scale of the analysis and the average size of the lineament clusters.

4.2. Directional analysis

Morphometric statistics of lineaments are widely used (Vaz et al., (2012); Abdullah et al., 2010; Masoud and Koike, 2011b). Statistics which support lineament interpretation are rose diagrams, length and azimuth distribution.

GEORIENT (Holcombe, 1994) and MARD (Moving Average Rose Diagram) (Munro and Blenkinsop, 2012) can be used to plot rose diagrams from this data. However, GEORIENT is not able to plot trend lines and MARD cannot work with length weighted azimuths, which makes the analysis independent of dataset fragmentation (Minár and Sládek, 2009). Therefore we wrote our own Python algorithm to generate a graph with the relative length weighted histogram with a moving average supplemented by a rose diagram from the same data (see Figs. 4, 9 and 11).

5. Applications

5.1. Central western Carpathian study area

Many landforms are scale specific (e.g. Evans, 2012), therefore the automatically delimited morpholineaments were extracted by MHHC in several levels of detail – datasets CWC Auto 20 m, CWC Auto 30 m and CWC Auto 40 m (the number in the dataset name reflects the spatial resolution of the related DEM). Morpholineaments derived from the more detailed datasets can reflect less distinctive younger faults and fractures that are manifested only in smaller line landforms.

The referenced datasets consist of:

1. dataset of morpholineaments derived from a 1:10,000 topographic map by a geomorphology expert (CWC Expert),

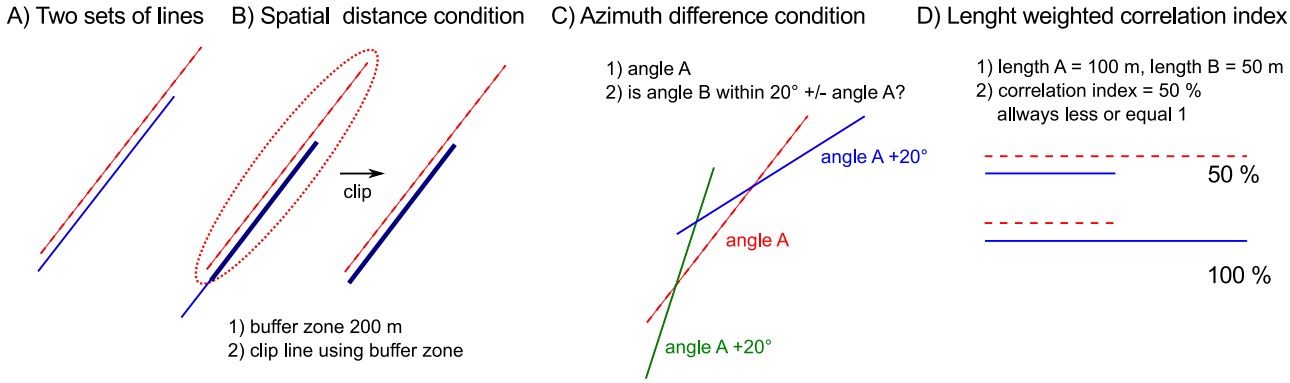

Fig. 7. Workflow of the comparison algorithm.

Table 3
The descriptive statistics of CWC datasets.

Dataset	Number of lines	Total length [km]	Mean length [m]	Min. length [m]	Max. length [m]
CWC Expert	658	402	610	78	3886
CWC Auto 20 m	1044	343	328	217	651
CWC Auto 30 m	531	301	568	360	1313
CWC Auto 40 m	334	211	631	443	1262
CWC Expert G	85	171	2015	434	10,698
CWC Fault lines	3545	301	85	10	929

Table 4
Results of statistical comparison of datasets in study area CWC 1:10,000. The values are in [%].

Reference Compared	CWC expert	CWC Auto 30 m	CWC Expert G	CWC Fault Lines
CWC Expert	–	52	25	18
CWC Auto 30 m	67	–	30	16
CWC Expert G	59	52	–	10
CWC Fault Lines	32	22	6	–

2. dataset made by simplifying and emphasising the main directions of the dataset from CWC Expert (CWC Expert G),
3. dataset of detected and expected fault lines from a 1:50,000 geological map (CWC Fault Lines), where fault lines from the map were split into straight segments.

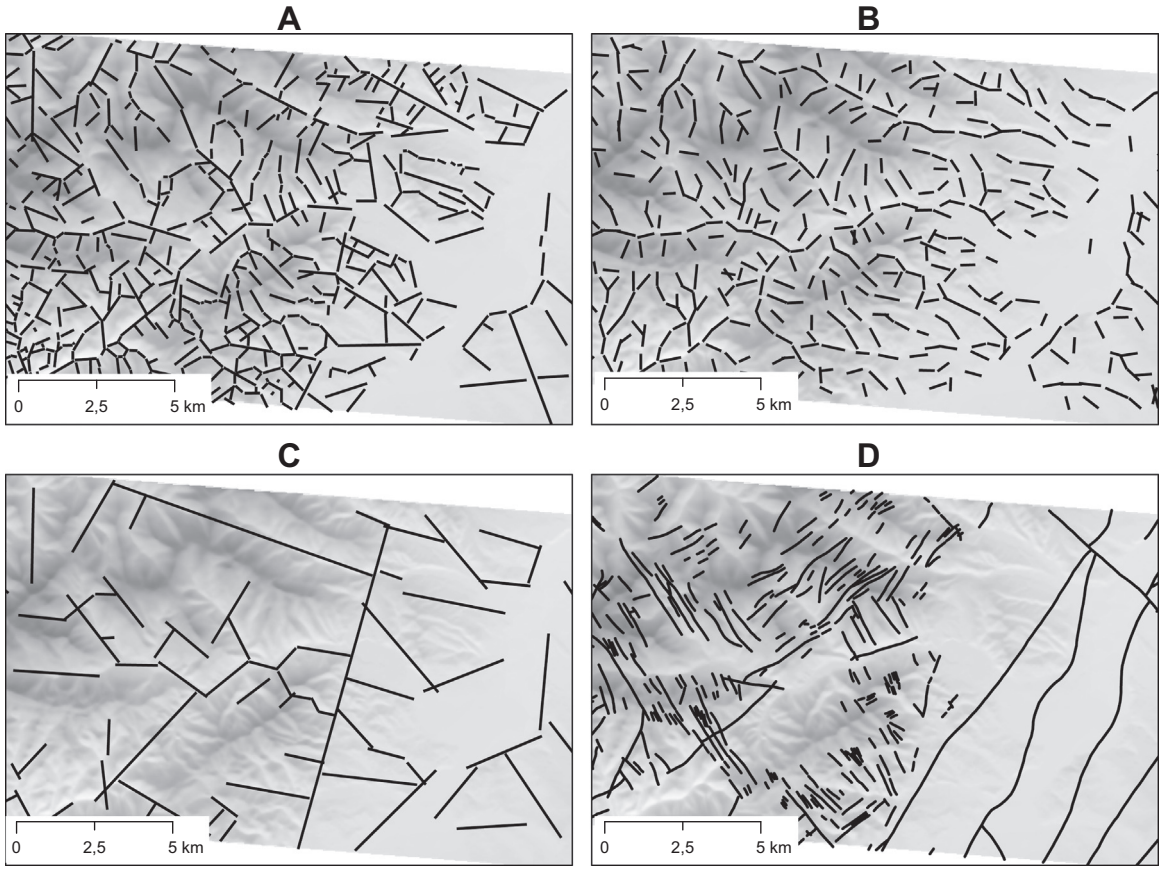

Fig. 8. Compared datasets. A) CWC Expert. B) CWC Auto 30 m. C) CWC Expert G. D) CWC Fault Lines.

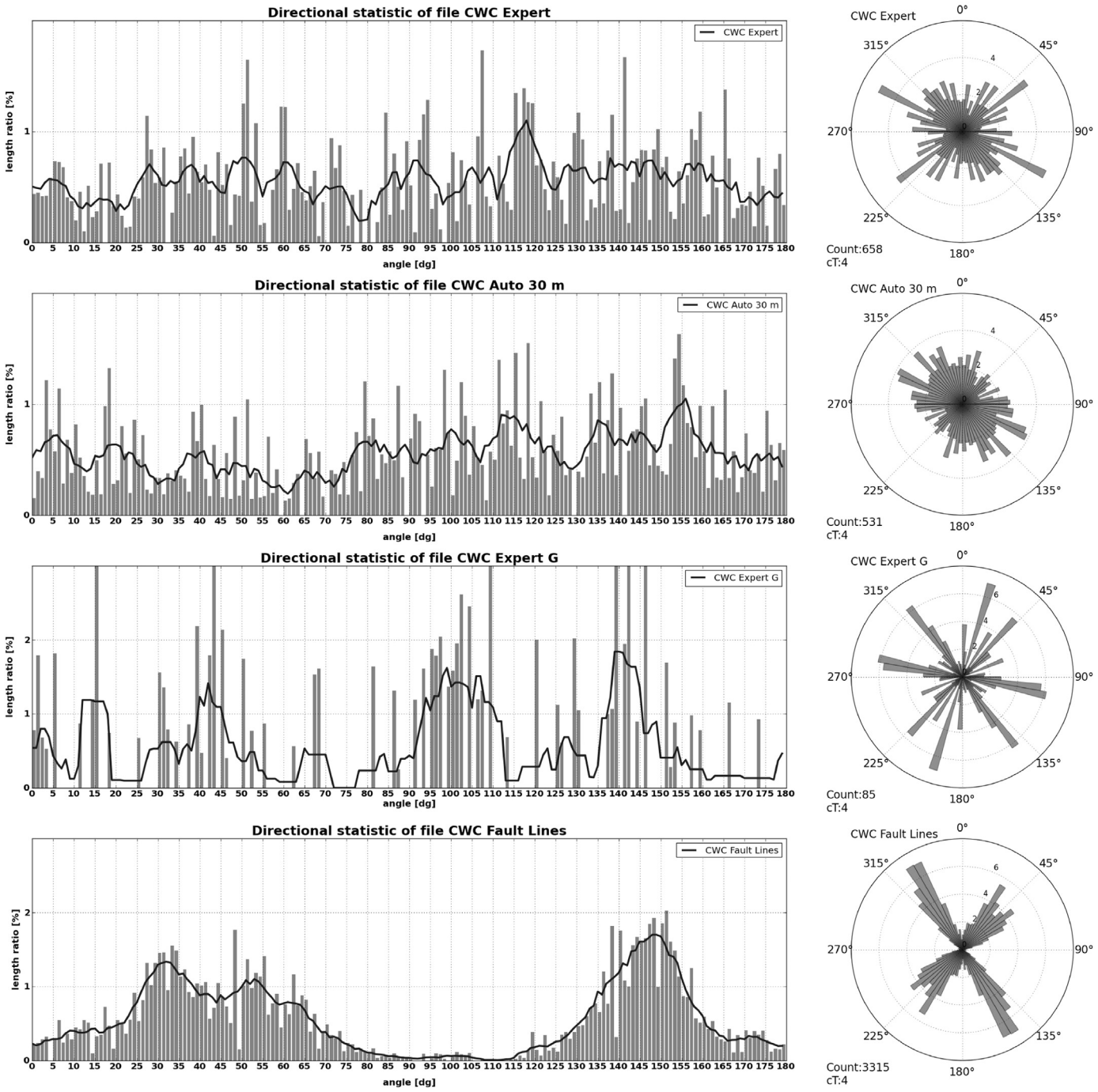

Fig. 9. The results of directional analysis of CWC datasets.

The basic descriptive statistics of all datasets are given in Table 3.

The CWC Auto 30 m dataset was chosen for the following tests because it has a level of detail comparable to the CWC Expert dataset. Research by Minár and Sládek (2009) used negative morpholineaments, thus only the morpholineaments classified by the algorithm as negative were used for comparison. The resultant maps of morpholineaments and fault lines are shown in Fig. 8.

The average mutual correlation between CWC Expert and CWC Auto 30 m datasets is about 60%. This means that 60% of the total length of one dataset is located within 200 m of the second dataset and varies less than 20° in azimuth (Table 4). The CWC Expert G dataset has a high (tested) correlation with both CWC Expert and CWC Auto 30 m datasets. The only important correlation is found where CWC Expert G is the tested dataset due to the smaller number and length of morpholineaments. This correlation expresses how many lines from the CWC Expert G dataset are correlated with the reference dataset. In the case of the CWC Fault Lines dataset, the correlation with all morpholineament datasets is too small to confirm a spatial similarity. Fault Lines here poorly correspond with topography and in some cases they are not straight lines. The low density of outcrops and boreholes also complicates comprehensive fault mapping. Therefore, inaccuracy of fault localisation may be higher than in the case of expert morpholineament creation. On the other hand, the morpholineaments also reflect influences other than tectonics (such as rock boundaries, joint systems or sedimentary layers).

The statistical directional analysis (Fig. 9) reveals an interesting regularity in the CWC Auto 30 m dataset. Local maxima of

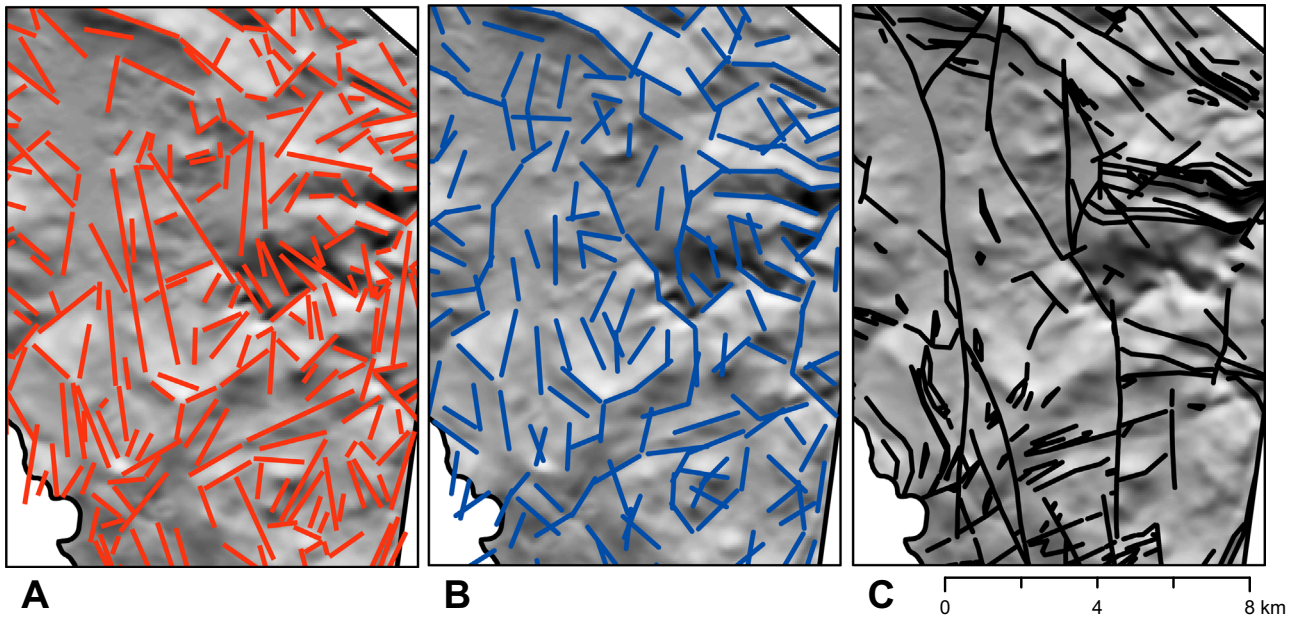

Fig. 10. Compared datasets. A) BF Expert. B) BF Auto 60 m. C) BF Fault Lines.

morpholineament lengths create roughly orthogonal systems (5° to 100°, 20° to 115°, 40° to 135°, 65° to 155°). They are also visible (slightly shifted) in the CWC Expert datasets (3° to 93°, 28° to 118°, 38° to 130°, 62° to 158°). This confirms the compatibility of the CWC Auto 30 m dataset with the Expert datasets. The most important direction of the CWC Fault Lines dataset (150°) is reflected better in CWC Auto 30 m than in the CWC Expert dataset – it can also be seen from the general shape of the rose diagrams (Fig. 9). An interpretation value of automatically derived morpholineaments for the faults detection results from this.

5.2. Bohemian forest study area

In BF the automatically delimited morpholineaments were extracted in the following spatial resolutions of DEM: datasets BF Auto 30 m, BF Auto 40 m, BF Auto 50 m and BF Auto 60 m. In contrast with CWC, negative and positive lines were delineated and analysed.

The referenced datasets consist of:

1. geological faults and rock boundaries which are important structural features in this long-term denuded area (BF Fault Lines). Both features were derived from a 1:50,000 geological map (Pelc and Šebesta, 1994). The obtained lines were split into straight segments,
2. dataset of morpholineaments derived from a 1:25,000 topographic map (BF Expert).

The visual presentation of datasets is shown in Fig. 10 and their descriptive characteristics are presented in Table 5.

Two automatically extracted datasets at a resolution of 30 m and 60 m (BF Auto 30 m and BF Auto 60 m) were chosen for comparison with the expert based morpholineaments and faults. The comparison of descriptive statistics shows similar values between the BF Auto 60 m dataset and the BF Expert dataset (Table 5). In contrast, visual comparison of the datasets reveals more detailed expert based morpholineaments than BF Auto 60 m in some areas. These detailed lines correspond more to the higher resolution dataset of BF Auto 30 m. In this case we argue that the better objectivity of automatic line extraction provided more holistic and balanced results throughout the whole area than the expert method. The expert based segmentation might be more subjective if, during the delineations, an expert unintentionally favours landforms which are supposed to be typical of the region.

The results of correlation are shown in Table 6. The datasets of BF Auto 60 m and expert morpholineaments have an average mutual correlation of over 50%. Both datasets have almost the same correlations to the rest of the datasets.

Correlation of the BF Auto 30 m dataset with the BF Auto 60 m is not taken into consideration because dataset BF Auto 30 m is more detailed than both previous datasets. About 60% of the length of BF Expert and BF Auto 60 m morpholineaments is correlated with dataset BF Auto 30 m.

Two significant directions of tectonic lines with different ages occur in the northwest part of the BF. The older lines run from north to south while the younger lines trend northwest-southeast (Finger et al., 2007). While the first direction is exhibited by north oriented valleys and some glacial cirques (Mentlík et al., 2010), the second direction is more prominent in the relief corresponding with the main direction of the BF ridges and copies the course of the regionally important Pošumavský Fault (Hartvich, 2004; Hartvich and Valenta, 2013). To test the algorithm's ability to detect tectonically significant directions, proportions of line lengths are evaluated in tectonically significant sectors 305° to 325°, 350° to 10° and 125° to 145°, 170° to 190°.

In the investigated datasets between 13.1% and 15.7% of delimited line lengths belong to the evaluated sectors. The lowest proportion in the tectonically significant sectors is evaluated in BF Auto 30 m (13.1% and 13.4%, respectively), while the highest accordance is shown by lines of the BF Expert dataset (15.6% and 15.7%, respectively). The lines of BF Auto 60 m (14.5% and 14.9%, respectively) show the proportion closest to the lines of BF Fault Lines dataset (14.3% and 15%, respectively).

This proves the ability of the algorithm to detect tectonically significant directions. However, the compatibility of the main directions of the automatically extracted morpholineaments and the directions of the principal faults (described in BF) suggest conformity between tectonically influenced features and automatically extracted morpholineaments. In this context an automatic algorithm may provide a tectonically relevant pattern, holistically covering the whole investigated area.

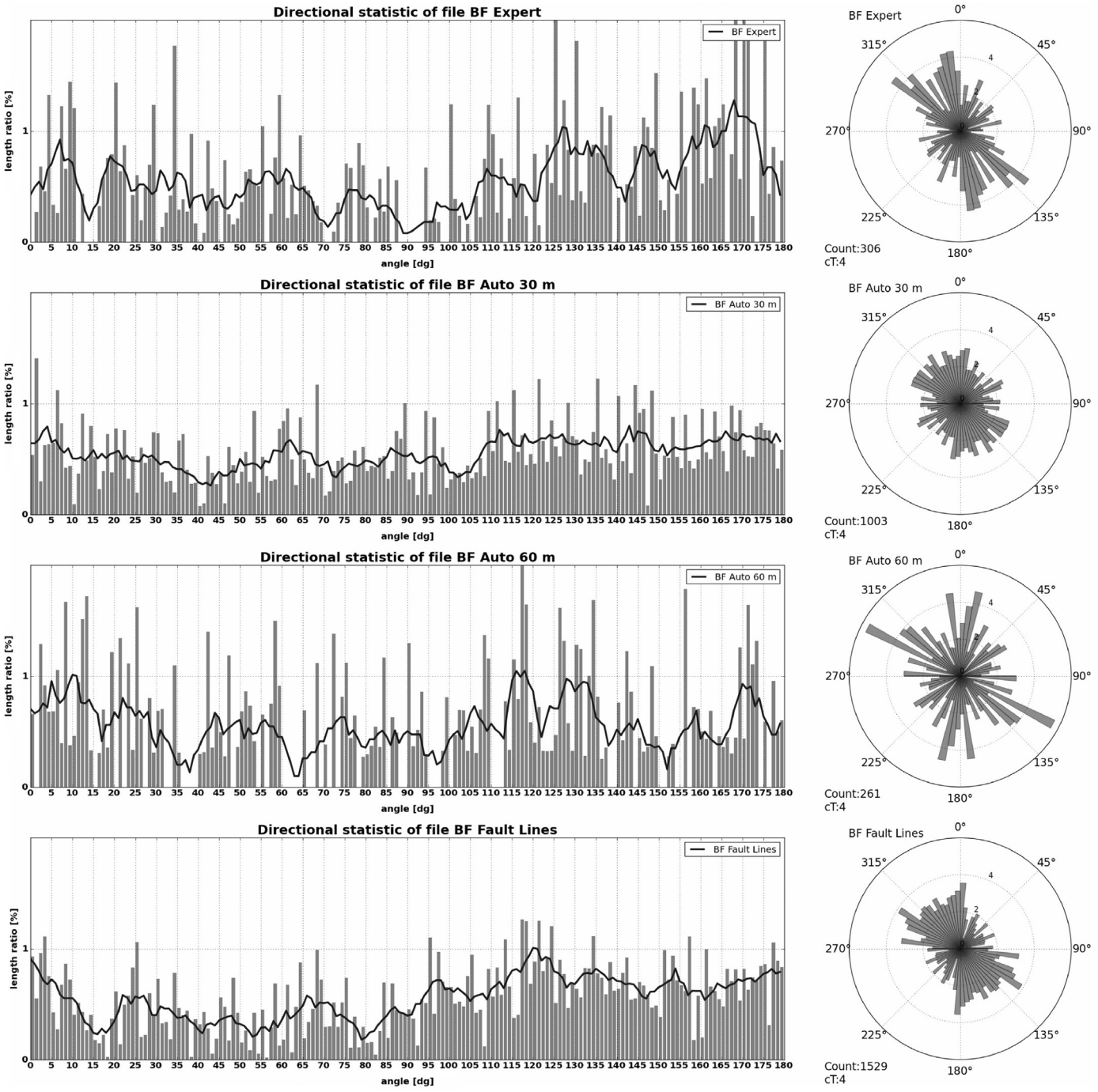

Fig. 11. Results of directional analysis of BF datasets.

The roughly orthogonal system identified in the CWC datasets is also recognisable in the BF, mainly in the BF Auto 60 m dataset. In the BF system it may be related to the joint systems of granitic rocks which also affect the shape of the glacial cirques in the area (Mentlík et al., 2010).

6. Conclusions

The method presented here for automatic extraction of lineaments is based on lines extraction from a raster image.

The originality of our methods in comparison with the cited works is set out below:

A)A robust solution for the illumination of shaded relief – the method works using the results of the lines extraction from all rasters, not only from one raster derived from a combination of all calculated rasters.

B)A more precise solution – the method advantageously combines raster and vector representations. It maintains the speed of raster analysis for removing noise, while retaining the precision of vector representation in terms of the location of the lineaments.

C)The use of spatial clustering of lines to find lineaments – the method clusters lines extracted from variously illuminated rasters. The representation of each lineament as a collection of 'protolineaments' reflects the 'fuzzy' nature of lineaments. In reality these lines (lineaments) approximate to elongated forms, e.g. based on several discrete belts of resistive rocks. Clustering of protolineaments and subsequent generation of lineaments can be

Table 5
The descriptive statistics of BF datasets. See the text for further explanation.

Dataset	Number of lines	Total length [km]	Mean length [m]	Min. length [m]	Max. length [m]
BF Expert	306	363	1186	284	5415
BF Auto 30 m	1003	468	467	328	880
BF Auto 40 m	645	399	619	445	1235
BF Auto 50 m	442	341	771	542	1341
BF Auto 60 m	261	288	1105	706	1732
BF Fault Lines	1529	414	271	4	1938

Table 6
Results of statistical comparison of datasets in study area BF. The values are in [%].

Reference Compared	BF Expert	BF Auto 60 m	BF Auto 30 m	BF Fault Lines
BF Expert	–	48	58	37
BF Auto 60 m	60	–	60	36
BF Auto 30 m	43	35	–	27
BF Fault Lines	40	34	45	–

best understood as establishing the most probable location of a particular lineament (the axis of the area of the lineament). It is a fundamentally new approach for defining lineaments in terms of their 'fuzzy' nature. Clustering used in other algorithms usually serves a different purpose (connecting lineaments into a larger unit, see Koike et al., 1998).

This paper points out the formation of artefacts during a raster based analysis and the Multi-Hillshade Hierarchic Clustering artefacts resistant method is presented. The lines extracted using PCI Geomatica software have to be processed in the specific way described in this paper in order to obtain valuable results of directional analyses.

In both study areas, which have different scales and sizes, the correlation between automatic and expert morpholineaments was around 50% to 60%. Moreover, the relation of the Expert and Auto datasets to the Fault line datasets is similar. This indicates that the presented method is able to delimit morpholineaments with a quality comparable to that of geomorphology experts. Extraction rules for manual delimitation of lineaments are not exact and the result is also influenced by the expert's subjectivity. In contrast, in the automatic extraction the same level of detail is preserved throughout the whole area. On the other hand, the automatic algorithm correctly detects mainly valley and ridge lines, but other types of lineaments (slope edges – e.g. slope foot lines) may be detected less reliably.

The automation mainly eliminates the subjectivity of manual extraction, providing relevant results at various scales. This more objective pattern promises to be a convenient basis for using sophisticated quantitative methods for investigating the expression of a morphotectonic field or the manifestation of rocky stress fields (Urbánek, 2005; Minár and Sládek, 2009; Sládek, 2010). Orthogonal systems of the main directions of morpholineaments have already been detected by previous research in the CWC area and confirmed by the results from the automatic dataset. The algorithm, however, can also be used for delimitation of linear features in the landscape beyond structural geology. For example, for detecting linear features which are broadly connected with human activity.

Application of moving averages is convenient for identifying orthogonal systems in datasets of automatically generated morpholineaments. These orthogonal systems are also reflected in the fault line statistics which can be effectively used in paleostress (morphotectonic) analysis.

Acknowledgements

We wish to thank Professor Ivana Kolingerova and Ondrej Kaas from the Department of Informatics for their valuable help with the optimisation of the cluster line algorithm. The first author also wishes to thank his supervisor Vaclav Cada for his support in this work.

This work was supported by the Slovak Research and Development Agency under the Contract no. APVV-0625-11.

The first author was supported by Project NTIS – New Technologies for Information Society, European Centre of Excellence, CZ.1.05/1.1.00/02.0090. Pavel Mentlík was supported by the Czech Science Foundation project 13-15123S.

The authors would also like to thank Jeremy King of University of West Bohemia for proofreading the manuscript.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.cageo.2016.03.015.

References

Abdullah, A., Akhir, J.M., Abdullah, I., 2010. Automatic mapping of lineaments using shaded relief images derived from digital elevation model (DEMs) in the maran – sungi lembing area, Malaysia. Electron. J. Geotech. Eng. 15 (J), 949–957.

Abdullah, A., Akhir, J.M., Abdullah, I., 2009. A comparison of landsat TM and SPOT data for lineament mapping in hulu lepar area, Pahang, Malaysia. Eur. J. Sci. Res. 34 (3), 406–415.

Argialas, D.P., Mavrantza, O.D., 2004. Comparison of edge detection and hough transform techniques for the extraction of geologic features. In: Proceedings of the XXth ISPRS Congress of the International Society of Photogrammetry and Remote Sensing, Istanbul, Turkey, pp. 1682–1750.

Batayneh, A., Ghrefat, H., Diabat, A., 2012. Lineament characterization and their tectonic significance using gravity data and field studies in the Al-Jufr area, southeastern Jordan plateau. J. Earth Sci. 23 (6), 873–880. http://dx.doi.org/10.1007/s12583-012-0298-6.

ESRI, 2013. ArcGIS Desktop Help 10.1. Redlands (California): ESRI, 2013. [on-line] ⟨http://resources.arcgis.com⟩.

Evans, I., 2012. Geomorphometry and landform mapping: What is a landform? Geomorphology 137 (1), 94–106. http://dx.doi.org/10.1016/j.geomorph.2010.09.029.

Finger, F., Gerdes, A., Janoušek, V., René, M., Riegler, G., 2007. Resolving the variscan evolution of the moldanubian sector of the bohemian massif: the significance of the bavarian and the moravo-moldanubian tectonometamorphic phases. J. Geosci. 52 (1–2), 9–28. http://dx.doi.org/10.3190/jgeosci.005.

Fitton, N.C., Cox, S.J.D., 1998. Optimising the application of the Hough transform for automatic feature extraction from geoscientific images. Comput. Geosci. 24 (10), 933–951.

Gupta, R.P., 2003. Remote Sensing Geology, 2nd edn. Springer, Berlinhttp://dx.doi.org/10.1007/978-3-662-05283-9.

Hartvich, F., 2004. Morphostructural analysis of the surroundings of Pošumavský fault. In: Proceedings of seminar "Stav geomorfologických a kvartérně-geologických výzkumů na Šumavě, Novohradských horách a Českém lese v roce 2004", Miscellanea Geographica 10, pp. 115–127.

Hartvich, F., Valenta, J., 2013. Tracing an intra-montane fault: an interdisciplinary approach. Surv. Geophys. 34 (3), 317–347. http://dx.doi.org/10.1007/s10712-012-9216-9.

Holcombe, R.J., 1994. GEOrient-an integrated structural plotting package for MS-Windows. Geol. Soc. Aust. 36, 73–74.

Jedlička, K., Sládek, J., Šilhavý, J., 2015. Semiautomatic construction of isobase surfaces: a case study from the central Western Carpathians. Comput. Geosci. 78, 73–80. http://dx.doi.org/10.1016/j.cageo.2015.02.012.

Jordan, G., Schott, B., 2005. Application of wavelet analysis to the study of spatial pattern of morphotectonic lineaments in digital terrain models, a case study. Remote Sens. Environ. 94 (1), 31–38.

Káčer, Š., Antalík, M., Lexa, J., Zvara, I., Fritzman, R ., Vlachovič, J., Bystrická, G., Bordianska, M., Madarás, J., Nagy, A ., Maglay, J., Ivanička, J., Gross, P., Rakús, M., Vozárová, A ., Buček, S., Boorová, D., Šimon, L., Mello, J., Polák, M., Bezák, V.,

Hók, J., Teťák, F., Konečný, V., Kučera, M., Žec, B., Elečko, M., Hraško, Ľ., Kováčik, M., Pristaš, J., 2005. Slovak Republic – digital geological map in 1:50,000 and 1:500,000. Manuscript, Archive of Geofond – the informatics branch of State Geological Institute of Dionýz Štúr, pp. 42, Bratislava. [in Slovak].

Kageyama, Y., Nishida, M., 2004. Lineament Detection from land cover information in mixels using landsat-TM data. Electr. Eng. Jpn. 148 (4), 65–73. http://dx.doi.org/10.1002/eej.10342.

Kim, G., Lee, J., Lee, K., 2004. Construction of lineaments maps related to groundwater occurrence with arc view and avenue scripts. Comput. Geosci. 30 (9/10), 1117–1126. http://dx.doi.org/10.1016/j.cageo.2004.09.002.

Koike, K., Nagano, S., Kawaba, K., 1998. Construction and analysis of interpreted fracture lanes through combination of satellite-derived lineaments and digital elevation model data. Comput. Geosci. 24 (6), 573–583.

Koike, K., Nagano, S., Ohmi, M., 1995. Lineament analysis of satellite images using a Segment Tracing Algorithm (STA). Comput. Geosci. 21, 1091–1104.

Mallast, U., Gloaguen, R., Geyer, S., Rodiger, T., Siebert, C., 2011. Derivation of groundwater flow-paths based on semi-automatic extraction of lineaments from remote sensing data. Hydrol. Earth Syst. Sci. 15, 2665–2678. http://dx.doi.org/10.5194/hess-15-2665-2011.

Masoud, A.A., Koike, K., 2011a. Auto-detection and integration of tectonically significant lineaments from SRTM DEM and remotely-sensed geophysical data. ISPRS J. Photogramm. Remote Sens. 66 (6), 818–832. http://dx.doi.org/10.1016/j.isprsjprs.2011.08.003.

Masoud, A.A., Koike, K., 2011b. Morphotectonics inferred from the analysis of topographic lineaments auto-detected from DEMs: application and validation for the Sinai Peninsula, Egypt. Tectonophysics 510, 291–308. http://dx.doi.org/10.1016/j.tecto.2011.07.010.

Mentlík, P., 2006. Geomorphological analysis and development of Geomorphological Information System (GmIS) for the surroundings of Prášilské and Laka Lakes (Šumava Mts., Czech Republic). Comenius University in Bratislava, Bratislava, Slovak Republic, p. 120 [in Czech].

Mentlík, P., Minár, J., Břízová, E., Lisá, L., Táborík, P., Stacke, V., 2010. Glaciation in the surroundings of Prášilské Lake (Bohemian Forest, Czech Republic). Geomorphology 117 (1–2), 181–194. http://dx.doi.org/10.1016/j.geomorph.2009.12.001.

Minár, J., Sládek, J., 2009. Morphological network as an indicator of a morphotectonic field in the central Western Carpathians (Slovakia). Zeitschrift für Geomorphologie, Supplementary Issues 53 (2), 23–29. http://dx.doi.org/10.1127/0372-8854/2009/0053S3-0023.

Munro, M.A., Blenkinsop, T.G., 2012. MARD – A moving average rose diagram application for the geosciences. Comput. Geosci. 21, 112–120. http://dx.doi.org/10.1016/j.cageo.2012.07.012.

Geomatics, P.C.I., 2011. Geomatica Help. [software help]. PCI Geomatics Enterprises. Richmond Hill, Ontario, Canada.

Pelc, Z., Šebesta, J., 1994. Geological Map 1:50,000. Sheet 22–33 Kašperské Hory. ČGÚ, Praha, Czech Republic [in Czech.

Pradhan, B., Pirasteh, S., Varatharajoo, R., 2010. Enhancement of automated lineament extraction from IRS-1B satellite imagery for part of Himalayan region. Int. J. Geoinform. 6 (2), 41–50.

Ramli, M., Yusof, N., Yusoff, M., Juahir, H., Shafri, H., 2010. Lineament mapping and its application in landslide hazard assessment: a review. Bull. Eng. Geol. Environ. 69, 215–233. http://dx.doi.org/10.1007/s10064-009-0255-5.

Seleem, T.A., 2013. Analysis and tectonic implication of DEM-derived structural lineaments, Sinai Peninsula, Egypt. Int. J. Geosci. 4 (1), 183–201. http://dx.doi.org/10.4236/ijg.2013.41016.

Skála, J., Kolingerová, I., 2011. Dynamic hierarchical triangulation of a clustered data stream. Comput. Geosci. 37 (8), 1092–1101. http://dx.doi.org/10.1016/j.cageo.2010.10.005.

Sládek, J., 2010. The Morphotectonic Field Reflection in Geomorphological Network in the Turčianska Kotlina Basin and its Surroundings. Comenius University in Bratislava, Bratislava, Slovak Republic, p. 120 [in Slovak].

Solomon, S., Ghebreab, W., 2006. Lineament characterization and their tectonic significance using Landsat TM data and field studies in the central highlands of Eritrea. J. Afr. Earth Sci. 46, 371–378. http://dx.doi.org/10.1016/j.jafrearsci.2006.06.007.

Smith, M.J., Clark, C.D., 2005. Methods for the visualization of digital elevation models for landform mapping. Earth Surf. Process. Landf. 30 (7), 885–900. http://dx.doi.org/10.1002/esp.1210.

Smith, M.J., Wise, S.M., 2007. Problems of bias in mapping linear landforms from satellite imagery. Int. J. Appl. Earth Obs. Geoinform. 9 (1), 65–78. http://dx.doi.org/10.1016/j.jag.2006.07.002.

Soto-Pinto, C., Arellano-Baeza, A., Sánchez, G., 2013. A new code for automatic detection and analysis of the lineament patterns for geophysical and geological purposes (ADALGEO). Comput. Geosci. 57, 93–103. http://dx.doi.org/10.1016/j.cageo.2013.03.019.

Štěpančíková, P., Stemberk, J., Vilímek, V., Košťák, B., 2008. Neotectonic development of drainage networks in the East Sudeten Mountains and monitoring of recent fault displacements (Czech Republic). Geomorphology 102, 68–80. http://dx.doi.org/10.1016/j.geomorph.2007.06.016.

Urbánek, J., 2005. Geomorphologic analysis of neotectonic forms. Geografický Časopis 57, 57–70 [in Slovak with English summary].

Vaz, D.A., Achille, G.D., Barata, M.T., Alves, E.I., 2012. Tectonic lineament mapping of the Thaumasia Plateau, Mars: comparing results from photointerpretation and a semi-automatic approach. Comput. Geosci. 48, 162–172. http://dx.doi.org/10.1016/j.cageo.2012.05.008.

Vaz, D.A., 2011. Analysis of a Thaumasia Planum rift through automatic mapping and strain characterization of normal faults. Planet. Space Sci. 59 (11–12), 1210–1221. http://dx.doi.org/10.1016/j.pss.2010.07.008.

Wladis, D., 1999. Automatic lineament detection using digital elevation models with second derivative filters. Photogramm. Eng. Remote Sens. 65 (4), 453–458.

Zlatopolsky, A., 1992. Program LESSA (Lineament Extraction and Stripe Statistical Analysis) automated linear image features analysis – experimental results. Comput. Geosci. 18 (9), 1121–1126.