Carpathian Journal of Earth and Environmental Sciences, November 2014, Vol. 9, No. 4, p. 237 - 252

3D DIGITAL GEOLOGICAL MODELLING OF PALEO-SURFACES. A CASE STUDY FROM THE NATIONAL NUCLEAR WASTE DEPOSITORY SITE IN HUNGARY

Gergely KORODY[1] & Gyozo JORDAN[2]

[1]Department of Palaeontology, Eötvös University, P.O.Box 120, H-1518 Budapest, Hungary; phone/fax: (36-70) 651 4229 / (36-1) 251 0703, e-mail: korodygergely@gmail.com

[2]Department of Chemistry and Biochemistry, Institute of Environmental Science, Szent Istvan University, 2100 Gödöllő, Pater Karoly utca 1. e-mail: gyozo.jordan@gmail.com

Abstract: The site of the final disposal facility for radioactive waste of the Paks Nuclear Power Plant at Bátaapáti in Hungary is under intensive geological survey, including paleo-environmental and paleo-surface investigations. The objective of this paper is to present the results of a borehole database analysis and modelling for the pre-Quaternary paleo-surface reconstruction in this important area. Data are heights above sea level of this paleo-surface measured in drill cores. The analysis is based on the assumption that a geological process such as hill slope erosion produces a statistically homogeneous distribution of the measured variable, paleo-surface depth in this case. During spatial trend analysis the obtained six statistically significant paleo-surface depth populations are divided into 13 trend surfaces significant at the 95% confidence level. The model gained by merging the 13 local trend surfaces describe the overall slope conditions. The other surface model is obtained by the accurate linear Triangular Irregular Network (TIN) interpolation capturing all the local details of morphological information. The original 10m grid TIN model is generalised by average smoothing filter with window size increasing from 25m to 1,000m in order to reveal morphological 'trends' at various spatial scales. Analysis of morphological lineaments confirm that major valleys run along regional fault lines, while secondary erosion processes forming the smaller side valleys act at scales below 1km. In this way the pre-Quaternary and younger paleo-valleys can be distingushed from the higher spatial scale tectonic forms. The stationary residuals remaining after trend removal enables detailed surface reconstruction by kriging interpolation.

Keywords: granite, paleo-surface, tectonic geo-morphometry, smoothing, TIN, trend

1. INTRODUCTION

The objective of geological survey is the 3D mapping of rock formations and geological features such as terrain surfaces formed in geological times (Walsh, 2009, Piotrowska et al., 2005). These surfaces are often burried by younger sediments and they need to be examined in boreholes, as well as natural and artificial exposures.

In order to build a geological 3D model (Bhuiyan-Anwar, 2009) of a buried (Marache et al., 2009, Eto et al., 2008) and eroded (Susini & De-Donatis, 2009, De-Donatis et al., 2009) paleo-surface depth information is obtained from surfacial surveys (Tonini et al., 2008, 2009), boreholes or geophysical measurements (Huang-Guo-Chin et al., 2009, Boehm et al., 2009). The surface is reconstructed from spatial seismic or electric (Bakkali & Amrani, 2008), as well as discrete point measurements (e.g. boreholes) by means of interpolation methods (Zanchi et al., 2009, Jahn & Riller, 2009). Structural information can be gained by statistical methods of borehole data (Bárdossy, Li, 2008). The assumption is that surface evolution including erosion processes result in systematic morphological pattern that can be analysed and modelled with the appropriate mathematical methods like trend (Merwade, 2009) and periodicity analyses, statistical homogeneity tests and autocorrelation. Reconstruction by interpolation is made difficult in the presence of heterogeneities produced by differential erosion due to lithological diversity (Maxelon et al., 2009) and by faulting causing movement of rock blocks (Bistacchi et al., 2008, Carrera et al., 2009). Thus, various methods have to be combined (Aunon & Gomez-Hernandez, 2000) and prior to surface modelling homogeneous surface areas have to be defined and delineated for

subsequent interpolation. Also, many interpolation methods (Li & Heap, 2008) such as kriging assumes stationarity in the mean (spatially invariant mean), that is systematic changes like trend or periodicity have to be modelled and removed from the dataset before surface reconstruction with geostatistical interpolation (Cressie, 1991, Goovaerts, 2009).

Mining and complex industrial projects are supported by geological exploration and site investiation (Ersoy et al., 2004, Wycisk et al., 2009). Our study is carried out in the Mórágy-Geresd Hills, an area of extensive exploration linked to the programme for the final disposal facility for low and intermediate level radioactive waste of the Paks Nuclear Power Plant in Hungary (Gyalog et al., 2004, Gyalog & Szegő, 2004).The investigation of surface evolution is essential to estimate and model 3D geological characteristics of the granite rock formation (Wan et al., 2002.) designated to keep the dangerous materials safe for centuries.

The objective of this study is to develop and apply methods for the identification of spatial heterogeneities and model them with various methods in order to identify major trends and patterns in burried paleo-surfaces (Peckham & Jordan, 2007). A further objective of detailed statistical and trend analysis is to prepare the borehole data for surface modelling with geostatistical methods like the kriging interpolation that requires stationarity in the data (constant mean with no outliers). A case study is presented for the detailed analysis of the pre-Quaternary paleo-surface in the study area and its evaluation in terms of spatial heterogeneity, spatial scales, exploratory surface modelling and paleo-morphotectonic interpretation.

2. STUDY AREA

The 8x8km study area encompasses the target area of the National Low and Intermediate Level Radioactive Nuclear Waste Repository in south-western Hungary (Fig. 1a). The 350 million years old Palaeozoic granite formation is the basement bedrock. The granite body, the host rock of the underground waste site, outcrops in the south-western corner of the study area at 110m a.s.l. and in incised valleys, and it fast plunges to depths of 60m (40m a.s.l.) toward the north-east (Fig. 1b). The north-western border of the granite body is defined by the South Mecsek Line separating the Paleozoic granite from the Jurassic sedimentary formations (Fig. 1b). To the east, the granite body is neighbouring the alluvial sediments of the River Danube. The other borders of the granite are uncertain but a very large extent under the surface is assumed (Balla, 2004a). The southern border of the target area is the border of the recent catchment boundary of the Lajvér Creek, beyond which the meridional valley system of Baranya Hills prevails (Sebe, 2009), while the Rák Creek is defined as the western border of the area (Fig. 1a).

The granite is increasingly weathered towards the surface and fresh granite is found in 50m depth only, at 60m above the sea level. The granite evolved by hypogenetic magma-mixing during the Variscian Orogenesis. In the last stage of cooling a NE–SW striking regional foliation formed (Király & Koroknai, 2004). The granite body is divided into three blocks. The northern and southern blocks are relatively undisturbed, while the transitional block obtained its position moving along a strike-slip fault which also caused folding in northern and southern blocks (Maros et al., 2004). The granite body uplifted to be exposed to surface denudation in the Carboniferous and Permian, then it slightly subsided in the Triassic and Jurassic. Sediments deposed in this period were eroded during the next uplifting stage in the Cretaceous. There were several sedimentation stages later in the Paleogene but the area remained in an emerged position in the whole period (Császár, 2004). Due to the prevailing subareal erosion in the Neogene, Latter Miocene sedimentary rocks are insignificant but in the northern and eastern areas. These rocks consist of either abrasion off-shore sediments linked to transgressional processes and they are composed of fine-grained sand with frequent well-rounded pebbles and boulders, or they are composed of silty-argillaceous sediments with angular or slightly rounded granite fragments. The granite pluton was completely covered by these sediments which were eroded off latter. A third sedimentary facies is the reddish-brown clay found only very locally. Despite of the intensive tectonic history of the area, the above described Tertiary sedimentary cover is horizontally bedded, irrespectively of the granite basement below (Balla 2004b) (Fig. 1d). Finally, the sequence is closed by Pleistocene loess of 70m thickness at some locations (Fig. 1c) underlain by Clay Formation (Marsi et al., 2004). The tectonic structure of the area is prevailed by fracture zones, magmatic and other plastic structural elements parallel to the NE–SW striking South Mecsek Line (Maros, 2006) (Figs 1a and 1b). The youngest movement was uplifting in the Latter Miocene and Pliocene along the pre-defined NE-SW strike-slip fault zone (Maros et al., 2004). The topographic surface has a slight tilt towards to north-east and the three main stream valleys run parallel in this direction over fracture zones to discharge eventually into the Lajvér Creek in the northern part of the study area (Fig. 1a). The origin and development of present landscape is argued but a combination of

tectonic and eolian processes are the most likely forces (Sebe, 2009).

3. MATERIALS AND METHODS

3.1 Data and data processing

The construction of the National Low and Intermediate Level Radioactive Nuclear Waste Repository was preceded by widespread geological survey resulting in an elaborate borehole database. The database includes 479 boreholes from drillings of different projects developed between 1921 and 1993 in the 16x16 km target area. Archive data were revised and some new drillings were also carried out. As a result, a dense network of boreholes is available offering a unique opportunity to model the pre-Quaternary paleo-surface with reasonable accuracy. The actual modelling area is smaller (8x8 km) in order to avoid edge effects due to interpolation errors in the final paleo-surface models. The borehole data used in this study is composed of the old so-called 'Archive Database' and the new 'Revised Database' which has recently been revised and re-interpreted. In this study, the new database is used, however, there are some boreholes available only in the archive data source. These are exactly the same boreholes used for the manual construction of the existing geological maps (Balla et al., 2009a, b) enabling the comparison of our numerical modelling results with previous studies. In one case our data analysis revealed error in the registered depth value (borehole O-4) that was excluded from further analysis. Some uncertainty in the depth values of the pre-Quaternary surface of the granite emerges from the fact that its surface is strongly weathered and the upper few tens of meters were occasionally interpreted as part of the overlying Quaternary sediments. Moreover, the borehole data is heterogeneous derived from projects carried out with different purposes in different times. Statistical analysis is used in this study to prove the homogeneity of data from drillings of various exploration campaigns.

The original database contains the borehole ID, the X,Y coordinates of the borehole in the EOV national metric projection system, the field identified topographic heights of the borehole a.s.l. (borehole Z value), and the top and bottom depth of rock formations measured from the topographic surface in metres. The studied surface is the top of the youngest pre-Quaternary rock formation observed in the borehole. For the modelling of the pre-Quaternary surface elevation above sea level, each depth value was substracted from the borehole Z value. Database Z values were checked against the Hungarian Defence Mapping Agency 1:50 000 scale DTM-50 digital elevation model and difference larger than 10m was not found, verifying the high accuracy of the new revised borehole database.

3.2 Statistical data analysis

First, statistical data analysis methods are used to identify statistically homogeneous sub-populations or groups in the pre-Quaternary surface depth data corresponding to paleo-surfaces or blocks at different depths. Sub-population identification followed the 'natural break' method, i.e. the data series was separated where the cumulative distribution function (CDF) had an inflection point (natural break) identified visually on the CDF plot (see Fig. 2) (Hoaglin et al., 1983; Reimann et al., 2008). This point corresponds to a local minimum in the frequency histogram (multi-modal histogram, see Fig. 2). One homogeneous distribution corresponds to a single stochastic process, thus identification of sub-populations reveals separate processes, i.e. uniform pre-Quaternary surfaces formed under uniform tectonic and/or erosion conditions. Separation of sub-populations was confirmed at the 95% confidence level by the Mann-Whitney (Wilcoxon) homogeneity test (Mann, & Whitney, 1947) based on the comparison of medians. This homogeneity test was also applied to verify the homogeneity of the various data sources. Based on the result of the homogeneity tests, the present discussion uses data sources that are homogeneous in the median at the 95% confidence level. Outlying values represent sudden and unusual values, essential for identifying singular heights or depressions in the modeled surface.

Figure 1. Location and geology of the study area. A. Digital elevation model of the study area. Inset shows location of study area in southeast Hungary. B. Geological map of pre-Quaternary rocks. The heights above sea level of the Paleozoic granite body top is shown by gray scale. Contour lines represent the thickness of the Pannonian sediment cover. Note that zero contour line marks the boundary of the sediments. Location of the borehole Üh-37 log in figure 1c and the geological cross-section in figure 1d are also shown. C. Borehole Üh-37 geological section (after Gyalog et al., 2004). For the location see figure 1b. D. Geologial section across the Bátaapáti target area (after Balla, 2004a). For the location see figure 1b. Key: 1. Quaternary sediments (Paks Loess and Tengelic Red Clay Formation); 2: Palaeozoic Mórágy Granite Formation, monzogranite group; 3: Palaeozoic, Mórágy Granite Formation, monzonite group; 4: Cretaceous, Rozsdásserpenyő Formation, trachy-andesite dykes; 5: geological contour; 6: fracture zone; 7: borehole within the section and its ID; 8: foliation of NW dip; 9: foliation of SE dip; 10: foliation of mixed dip.

↓

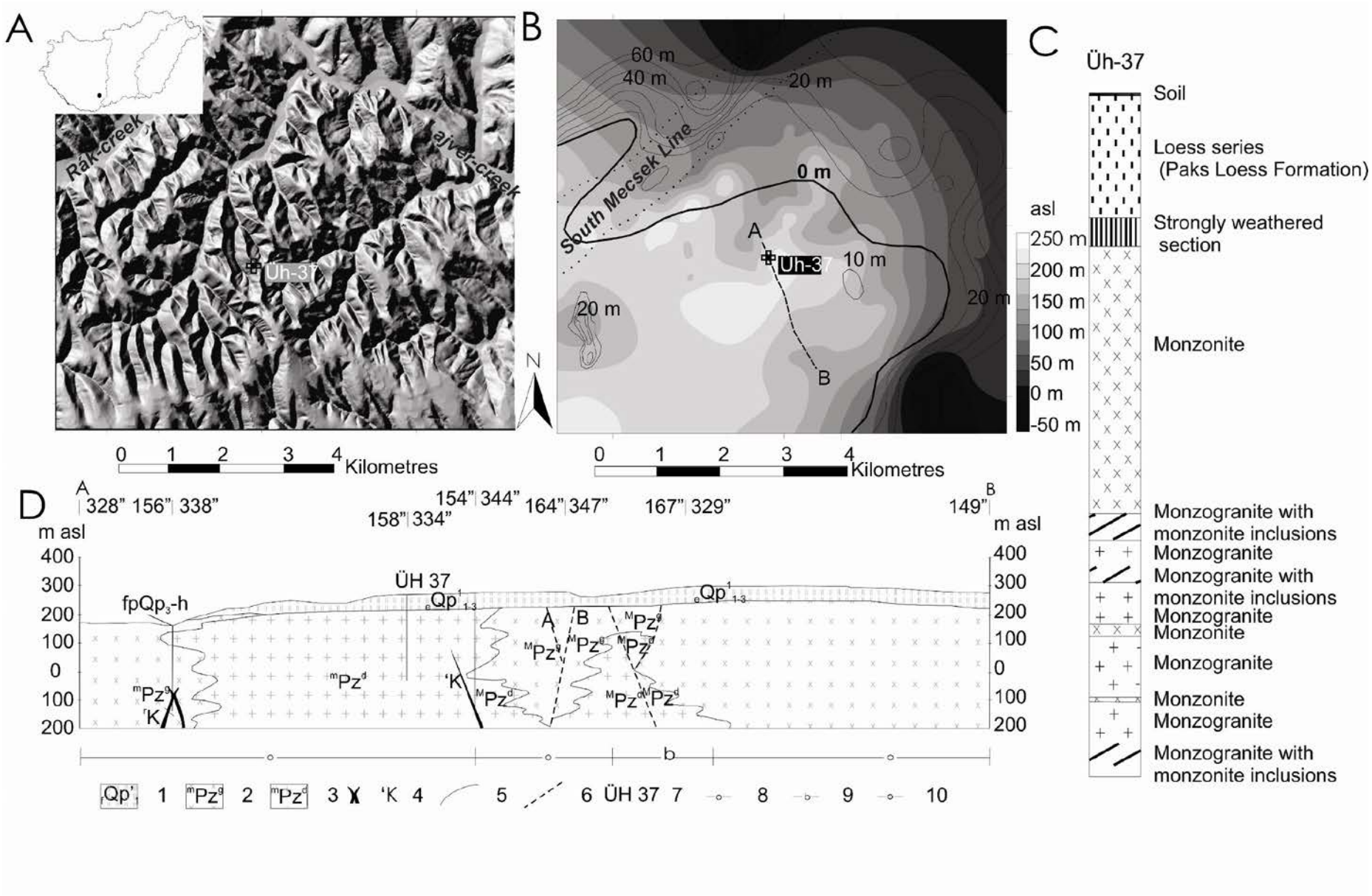
A
Rák-creek
Ajvér-creek
Üh-37
0 1 2 3 4 Kilometres
B
60 m
40 m
20 m
South Mecsek Line
0 m
A
Üh-37
10 m
20 m
20 m
B
N
0 1 2 3 4 Kilometres
C
Üh-37
asl
250 m
200 m
150 m
100 m
50 m
0 m
-50 m
Soil
Loess series
(Paks Loess Formation)
Strongly weathered
section
Monzonite
Monzogranite with
monzonite inclusions
Monzogranite
Monzogranite with
monzonite inclusions
Monzogranite
Monzonite
Monzogranite
Monzonite
Monzogranite
Monzogranite with
monzonite inclusions
D
A
328" 156" 338"
158" 334"
154" 344"
164" 347"
167" 329"
149"
B
m asl
400
300
200
100
0
100
200
ÜH 37
fpQp3-h
Qp1 1-3
Qp1 1-3
mPzg
mPzd
'K
MPzg
MPzd
b
Qp' 1
mPzg 2
mPzd 3
'K 4
5
6 ÜH 37 7
8
9
10

Table 1. Characteristics of homogeneous data populations distinguished by one variable analysises i.e. natural break method and confirmed by Mann -Whitney homogeneity test.

groups	N	Average	Standard deviation	Minimum	Lower quartile	Median	Upper quartile	Maximum	Interquartile range
all data	479	157	46	45	118	150	198	273	80
A	15	72	5	63	67	74	75	79	8
B	33	94	6	82	91	94	98	104	7
C	222	130	16	107	116	127	141	164	25
D	39	174	5	165	169	174	178	180	9
E	82	197	7	183	191	197	203	209	12
F	82	224	1	210	217	224	229	246	12

Tukey's (1977) inner-fence criteria (see Fig. 2 c) was used for outlier definition. The spatial location of the boundaries between areas of sub-populations was delineated by automatically generated contour lines from a simple linear TIN (Triangular Irregular Network) interpolated surface (see below). For example, the major breakline, according to the histogram analysis, in the studied pre-Quaternary surface is at 183m a.s.l., so it was identified and displayed as the 183m contour line (see Fig. 3b).

Summary statistics were calculated for the original data series and for the identified sub-populations separately (Fig. 4). Since geological data series, such as borehole measurements, are often characterized by non-normality, heterogeneity and outliers, robust statistics such as the median for location (central tendency), and median averaged deviation (MAD) and the inter-quartile range (IQR) for measure of scale (variability) were used in this study (Velleman & Hoaglin, 1981). Tukey's (1977) robust five-letter display formed the basis for distribution characterization including the minimum, lower quartile, median, upper quartile and maximum values (see Table 1) displayed in Box-and-Whisker plots (see Figs 2b and 2c) (McGill, 1978). The comparison of the average (median) depths of rock units corresponding to depth sub-groups enables the quantitative description of their relative positions. The comparison of variability (IQR) provides an insight into the paleo-surface relief and dissection. Since a trend can be present in the given surface, random variability is calculated after trend removal from the rock unit depth data.

3.3 Surface modelling and interpretation

Following the exploratory modelling approach, two linear interpolations, Triangular Irregular Network (TIN) and least-squares linear trend surface (plane) regression are used to model the pre-Quaternary surface. A preliminary surface grid is created by TIN using all the data points. Although the obtained model is blocky, its advantage is the accurate representation of the original data because TIN is a linear interpolation and it is an exact interpolator, too (Guibas and Stolfi, 1985). The TIN is rasterised to a 10m grid in the 16x16 km target area in order to be able to resolve the shortest distance (12.8m) between the two closest boreholes. The hence obtained 10m TIN grid surface is then further processed by average smoothing filters with window sizes increasing from 25m to 100m, 250m, 410m, 500m and 1,000m in size (see Fig. 5). Grid decimation was not attempted to avoid aliasing expected due to the parallel equidistant paleo-valley system. Besides improving the visual appeal of the TIN surface, smoothing of various levels of generalization can reveal the spatial scale of paleo-morphological features (Peckham & Jordan, 2007) such as major valleys. The smallest size of 25m was determined by the nearest borehole locations (12.8-23m) and the largest kernel size of 1,000m was identified by the widest paleo-valley width estimated visually from the TIN model (see Fig. 5).The smoothed surfaces, together with the original TIN model, were visualized as shaded relief models with Lambertian illumination from the eight cardinal directions (45° from North) using a 45° illumination azimuth and 10x vertical exaggeration.

These models were used to digitize paleo-morphological lineaments along significant linear light-dark features corresponding to valleys, ridges and slope-breaks, according to Jordan (2007a) and Jordan et al., (2005). Unlike the statistically-based local trend surface model (see below), the TIN model does not capture the pre-Quaternary surface sub-populations at various depths. The other surface model used a simple linear least-squares trend surface (plane) fit to all data within the statistically identified homogeneous depth sub-populations representing uniform paleo-surface areas or blocks. Bivariate outliers were identified and removed from the trend models and the presence of linear trend (uniform surface tilt) was accepted at the 95% confidence level. In the lack of significant trend the borehole depth data sub-population represents a horizontal flat surface. For each fitted plane, the slope direction and magnitude was calculated from

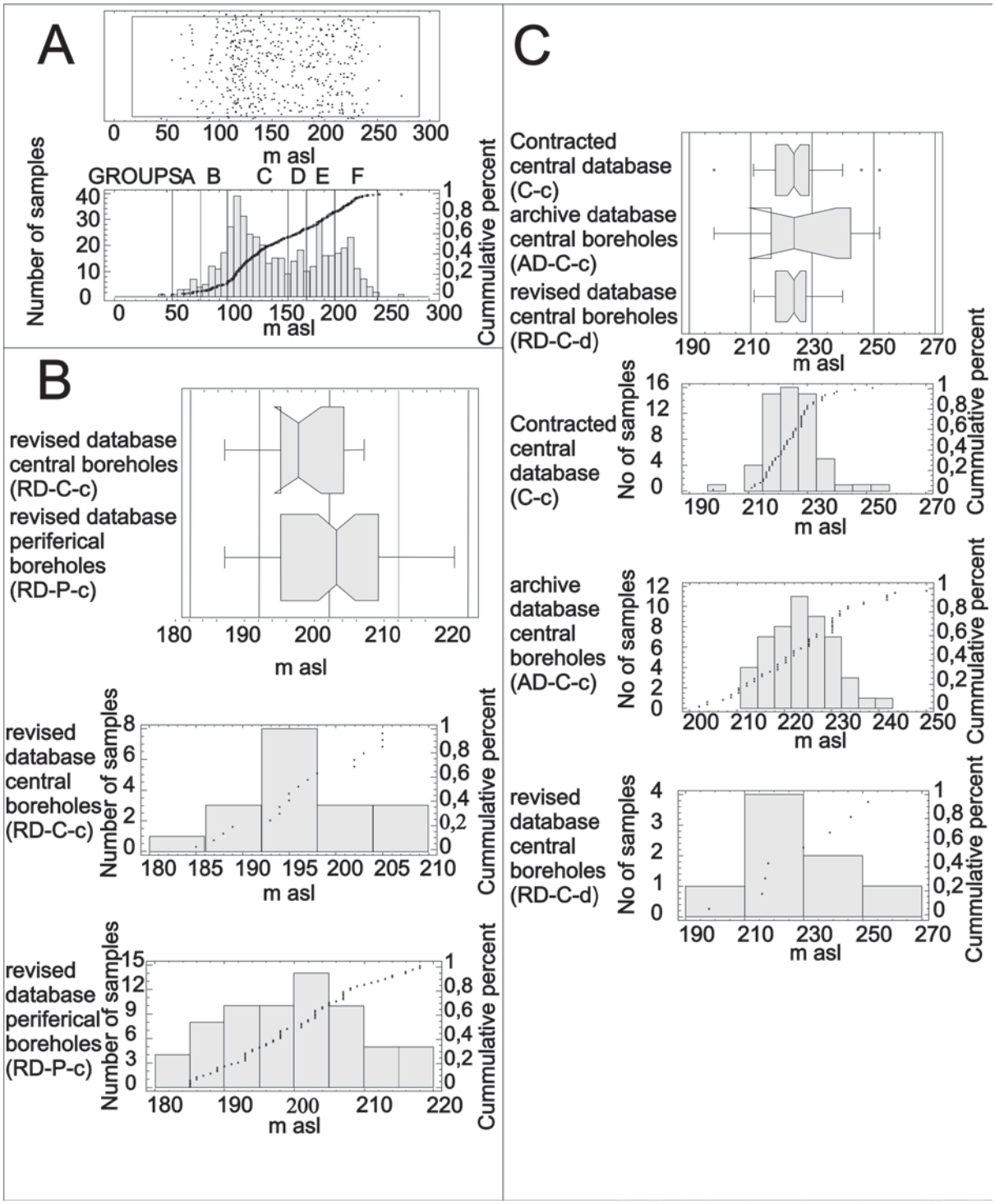

Figure 2. Statistical distribution analysis. A. Scatterplot, frequency histogram (below) and cummulative histogram for all observations for the pre-Quaternary surface elevation (metres a.s.l.) measured in the studied 478 boreholes. The identified six homogeneous depth sub-populations (A-F) are also shown. B-C: plots of homogenity analysis on databases that cannot be contracted (B) and databases that are homogeneous with each other (C). Box-whiskers plots display the five number summary statistics of the borehole data (minimum, lower-quartile, median, upper–quartile and maximum). Outliers are also shown

the regression equation according to Jordan (2007b) (see Table 2 and Fig. 3a). The hence obtained surface models were delineated along the population (block) boundaries and fit together to obtain the pre-Quaternary surface model for the target area (see Fig. 7). Note that this model does capture the pre-Quaternary surface sub-populations at various depths. This model considers the statistically identified surface sub-areas and it is considered the final model at this stage of investigation.

Finally, the residuals remaining after trend removal (Erdogan, 2010) were compared in terms of variability measured by IQR in order to characterise these rock units for surface relief and roughness. The hence obtained residuals are homogeneous and stationary in the mean and thus ready for interpolation by kriging to reconstruct paleo-surface details such as paleo-valleys, which is a subject of a separate study.

In this investigation a combination of various GIS and statistical software were used. TIN surface was generated and smoothed with Golden Software Surfer 10. Contour lines between sub-population areas were also drawn with Surfer. Statistical analysis, including trend fitting was carried out in the Statgraphics Centurion XVI environment. All spatial data processing such as boundary elimination and overlays used the raster GIS software ILWIS 3.7.

4. RESULTS AND DISCUSSION

The pre-Quaternary surface elevation ranges from 45m to 273m a.s.l. with a median central value of 150m (Table 1). Six homogeneous sub-populations were identified by detailed statistical analysis in the pre-Quaternary borehole depth data. The population medians are significantly different at the 95% confidence level. When the areas covering the six sub-populations are displayed a consistent pattern of well-defined concentric zones becomes apparent descending from a median 224m elevation in the southwest to 74m elevation in the northeast by about 150m within 10km in a staircase fashion (Fig. 6). These areas are spatially very homogeneous apart from a few small isolated regions of local heights or depressions being located in the area of other elevation groups. These islands are all represented by one or two boreholes only thus they were eliminated from further spatial analysis. Although the objective is paleo-surfce modelling and not the genetic interpretation of geomorphological processes, it is obvious that the contour lines bordering the six zones reveal three major NE-SW oriented paleo-valleys sub-parallel to the current valley network (Fig. 3b). The median elevation (a.s.l.) values of the 6 populations or zones systematically decrease from the southwest to the northeast in the order F(224m) >E(197m) >D(174m) >C(127m) >B(94m) >A(74m) (Fig. 3b). Populations are vertically separated by an average of 30m. This difference seems small in the absolute sense but represents 10% of the total height difference in the target area. It is interesting that the largest difference (48m) is between zones D and C, while the other elevation differences remain all about 20-30m. Population C has the biggest area (88 km^2) and contains 222 boreholes, representing 46% of the whole dataset. Populations B, D and E are narrow concentric zones only with the smallest areas of 48 km^2, 18 km^2 and 44 km^2 represented by 33, 39 and 82 borehole depth data, respectively. Zones A (31 km^2) and F (37 km^2) represented by 15 and 82 data points are located at the lowermost sedimentary basin area in the north and the opposite uppermost flat plateau area on the top of the granitic pluton in the south, respectively (Fig. 6).

The largest group C is the most heterogeneous, too, as shown by the highest IQR value (25m) exceeding 2 times the smaller population variabilities (IQR =7-12). The few (6) high outliers occur in areas covered by Tertiary sediments in the north indicating that erosion did not result in extreme local heights and depth (valleys) in the old granitic block in the south.

Table 2. Characteristics of homogeneous trend surfaces. α (aspect) and γ (slope) gives spatial, R^2 and significance to statistical description of trend surfaces.

ID	α	γ	R^2	N	significance	equation
A	56.99	0.11	96.45	10	0.0013	Z=-0.00153973X+0.00100014Y+1132.99
B1	153.17	0.73	74.04	14	0.0023	Z=-0.0057121X-0.0112932Y+4717.64
B2	-35.18	0.11	74.79	13	0.002	Z=0.00106622X+0.00151277Y-717.408
C1	233.58	1.46	77.93	15	0.0002	Z=0.0204597X-0.0150952Y-10791.4
C2	162.01	0.32	34.36	74	0	Z=-0.00170554X-0.00525197Y+1698.47
C3a	159.78	2.02	46.41	109	0	Z=-0.0121611X-0.0330231Y+10796.5
C3b	3.09	0.53	85.41	9	0.0081	Z=-0.000502741X-0.00930246Y-400.403
D1	252.76	0.42	64.97	18	0.2074	Z=0.00696853X-0.00216228Y-3846.6
D2	265.68	0.05	21.35	18	0.1464	Z=0.000895668X-0.0000676263Y-372.887
E	134.52	0.06	6.15	81	0.0843	Z=-0.00072652X-0.000714492Y+710.402
F1	139.47	0.20	40.31	21	0.0452	Z=-0.00227964X-0.0026662Y+1866.34
F2	150.48	0.36	21.89	32	0.0278	Z=-0.00308095X-0.005441Y+2624.26
F3	155.84	1.10	51.55	24	0.0005	Z=-0.00783828X-0.0174758Y+6713.85

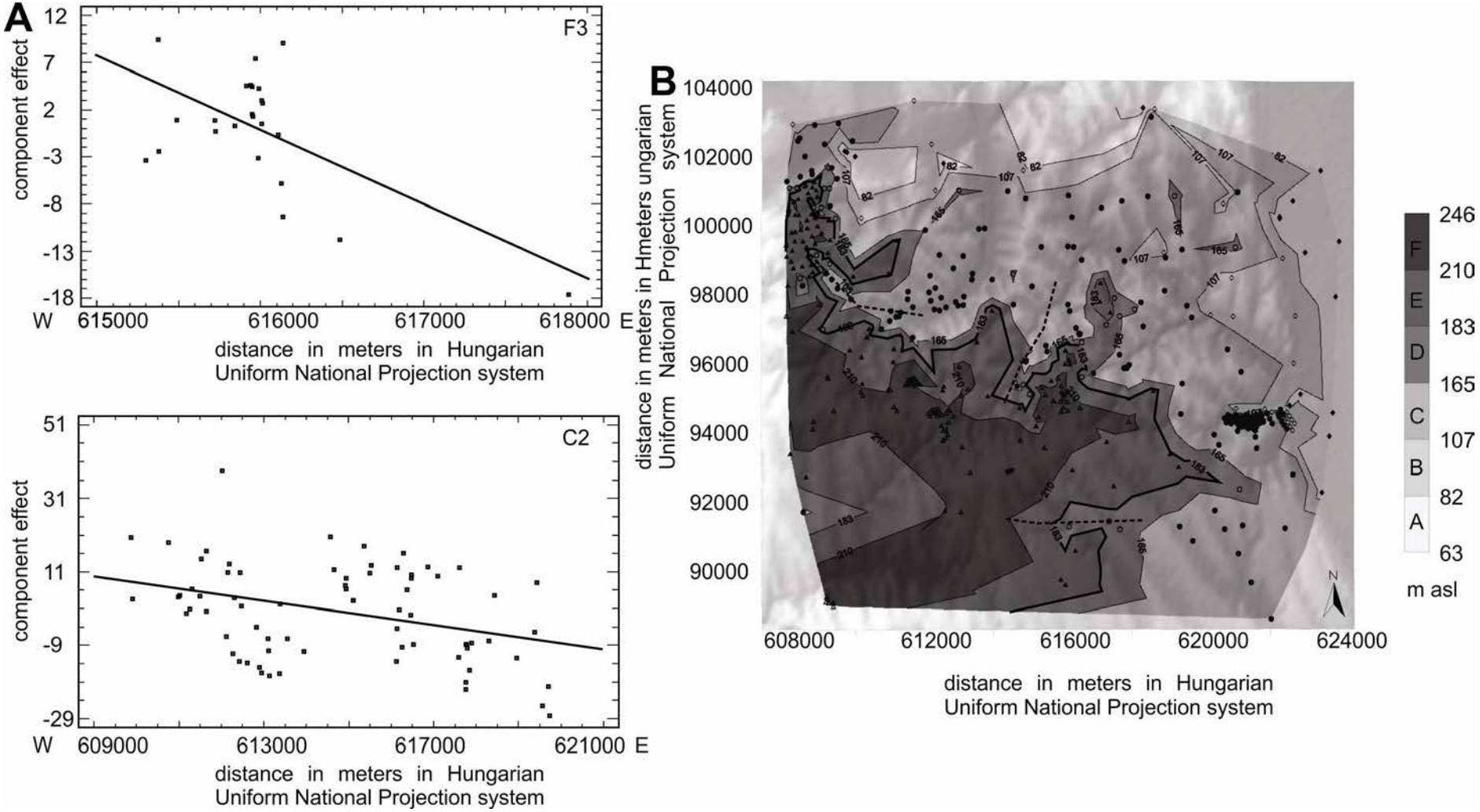

Figure 3.Statistical data analysis and spatial representation. A. Trend surface modelling demonstrated for bivariate sub-populations F3 (top) and C2 (bottom). X-axis represents geographic postion or distance in metres. B. Location of boreholes and areal distribution of the identified six homogeneous sub-populations (A-F) gained from one variable analysis of pre-Quaternary surface depth data. Borders separating the statistically significant depth populations run along height contours of the TIN model generated from the dataset (see Fig. 5a). The main boundary is between zones D and E at 183 m a.s.l. Shaded relief topographic surface is also shown in the background. Dashed lines show the three main SE-NW oriented paleo-valleys revealed by the contour lines along the population boundaries.

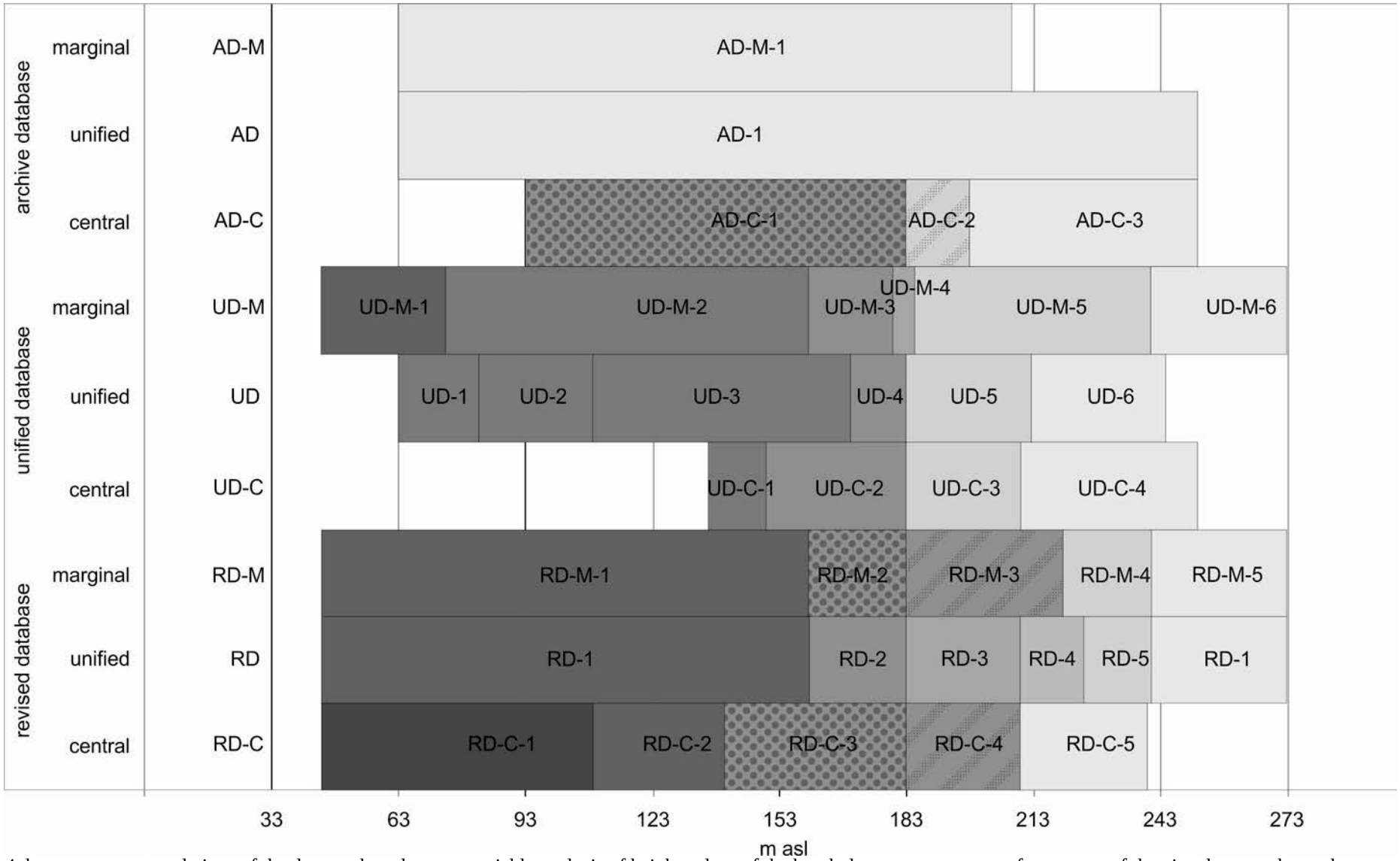

Figure 4. homogeneous populations of the datasets based on one variable analysis of height values of the borehole pre-quaternary surface. seven of the nine datasets shows the border around 183 m a.s.l. The second largest boundary runs at 240m a.s.l

The most significant break is at the 183m a.s.l. population boundary, separating the narrow zones corresponding to populations E and D (Fig. 3b, Fig. 4, Table 1). The principal boundary expresses the strongest relief and the comparison of the statistical results (Figs 2 and 3b) and the geological maps (Fig. 1b; Balla, 2003) reveals that above this height the pre-Quaternary surface consists of the Palaeozoic granite body, while young Pannonian marine sediments are found below.

The second strongest border of 210m delineates the uppermost steep area, zone F, identifying a significant morphological slope break in the granitic surface (see trend analysis below). The level 165m lies close to the main 183m boundary both in height and spatially, separating the largest lower lying zone C from the narow zone D positioned above it (by 48m, see above). Zone C definitely belongs to the sedimentary basin (see Fig. 1). Thus, taking into account the sub-parallel position of pre-Quaternary strata boundaries, the narrow transient zone D most likely represent the paleo-slope scree between the higher lying granitic pluton and the surrounding sedimentary basin. Finally, zone A encompasses the sedimentary basin at the Szekszárd Hills and the Great Plain in N-NE.

Manually constructed geological maps show paleo-valleys based on the present topography, primarily (Balla et al., 2009b). Our detailed statistical analysis identified homogeneous concentric zones in the pre-Quaternary paleo-surface that often contain boreholes aligning along elongated areas intruding into higher lying regions (Fig. 3b). This implies that these areas are valleys that formed together with the lower area in the same geomorphological process such as regional erosion.
Despite of that the used boreholes are very irregularly distributed and they were drilled for various projects and purposes, such as coal exploration, national geological survey and survey for the nuclear waste repository project, still the same homogeneous populations with well-defined spatial distribution are found in the used nine datasets (Fig. 6, Fig. 7). This confirms that the pre-Quaternary surface depth populations identified in this study are both statistically and geologically significant and that the found height differences are significantly larger than the data error within and among the various borehole datasets. Trend surface was fitted to borehole pre-Quaternary depth values within each of the six homogeneous areas. Thirty-five points were excluded during trend analysis because these were identified as bivariate outliers biasing least-squares regression analysis (Table 2). When the trend plane is fitted to all data points with bivariate outliers removed, an overall 1º tilt towards 42º to the northeast is obtained. However, population D located in the centre of the study area with the 183m principal elevation boundary has no trend and it is interpreted as a flat surface (Figs 3b and 6a). This confirms that this narrow zone represents the levelling effect of slope scree deposited at the erosion base of the granite pluton. It is interesting that various trend surfaces became apparent within sub-populations revealing alternating tilt of the pre-Quaternary paleo-surfaces. Populations A (63-79 m) and E (183-209 m) have uniform trends (Fig. 6). Area E has very low dip (slope =0.02º) and it is neighbouring the flat zone D in the middle of the target area. Populations B (82-104m), C (107-164m) and F (210-246m) however have various trends (Fig. 6). Altogether 13 trend surfaces were obtained, most of them tilting towards the northeast. The slope of trend surfaces is generally less than the overall tilt (Fig. 6) and less than the 2º average slope of the TIN surface (Fig. 4a). Trend surfaces show the overall slope conditions while the TIN surface facets correspond to local geomorphology along paleo-valleys with steep hill sides.

Based on the trend analysis implemented on the six elevation zones, the target area can be divided to three parts. The zone below the most important boundary of 183m (zone D, 165-183m) and the one just above (zone E, 183-209m) are the most homogeneous with respect to trends and are essentially horizontal (D1 γ=0º, E γ=0.02º). They constitute the transient zone separating the Pannonian sedimentary strata and the higher lying Palaeozoic plutonic rocks. Since the paleo-valleys cross-cut these horizontal zones, they represent surfaces predating or dating the valley formation. Most likely, these uniform sub-horizontal surfaces (zones D and E) represent the Pannonian age shallow sea shore line above which sub-aerial erosion of the exposed granitic pluton occurred, while Pannonian shallow marine sediments were depositing below, with zone D, together with zone E, interpreted as slope scree.

The three trend surfaces of the granite pluton (F1, F2, F3) above the sub-horizontal transient zone (zones D-E, 165-209m) have N-NE gradient sloping by 0.8º to 1.1º (Fig. 6). Under the transient zone the four trend surfaces of population C (C1, C2, C3a, C3b) and the B2 planes are relatively steep ($\gamma \geq 0.3$º), and their slope directions are diverse (ranging from N, to S, E and W) (Fig. 6). The trends of these areas pointing to S (C3b) and W (C1) are fitted to marginal boreholes, therefore they bear some uncertainty. These four trend surfaces (C1, C2, C3a and C3b) are most likely piedmont based on the divergent tilt of these areas away from the granitic denudation surface. Finally, the lowermost zone A slopes to the north.

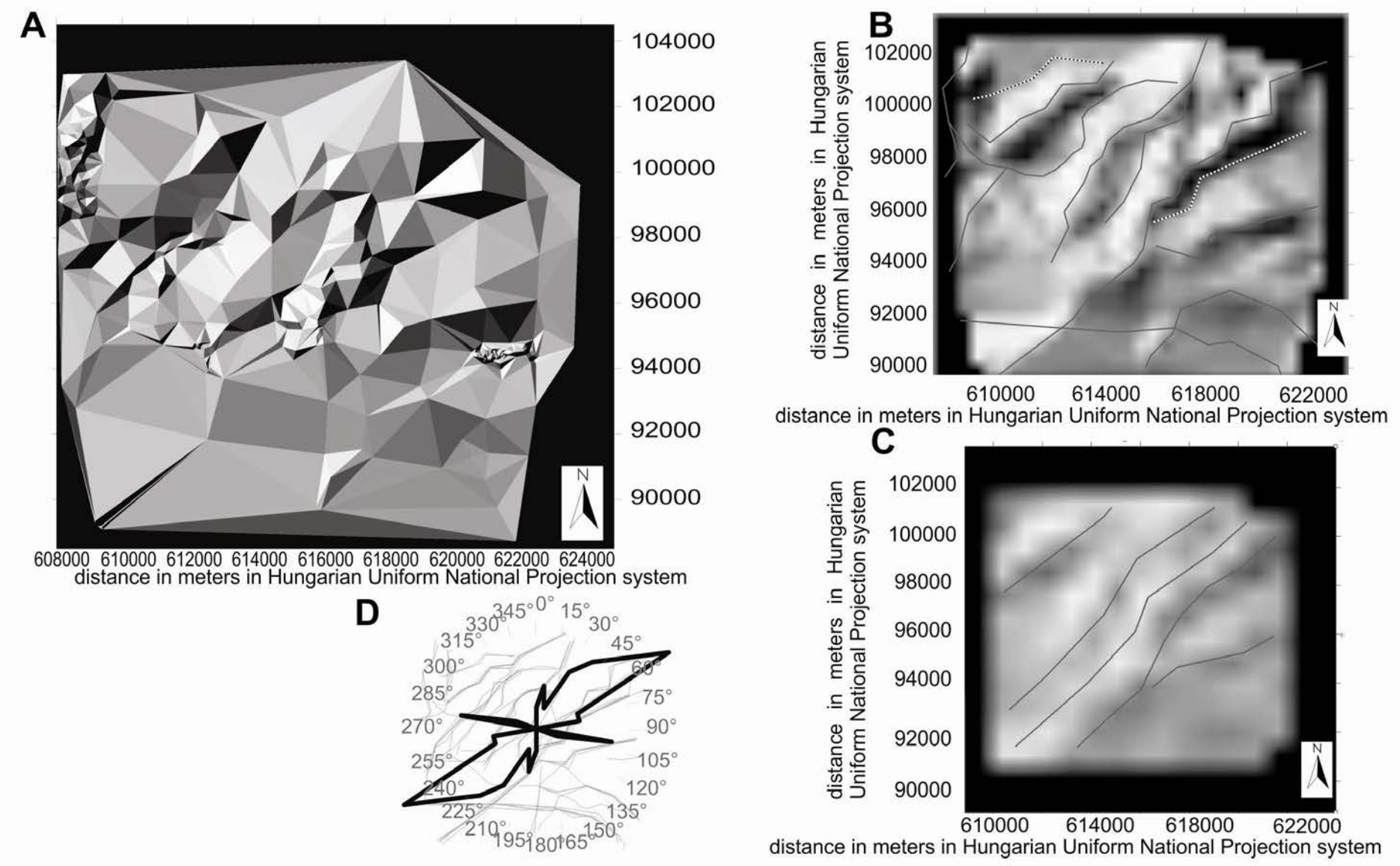

Figure 5. Local trend surface model. A. Contour map for the the local trend surface model. Aspect (α) and slope (γ) values of each trend plain is shown at the gradient vectors (solid arrows). Paleo-valleys cross-cutting the statistically defined homogeneus zones are shown by dashed lines. B. 3D perspective view of the local trend surface model. D and E (horizontal) trend plains are in a transient zone, B and C trend surfaces consist of Pannonian marine sediments surrounding the elevated granite body.

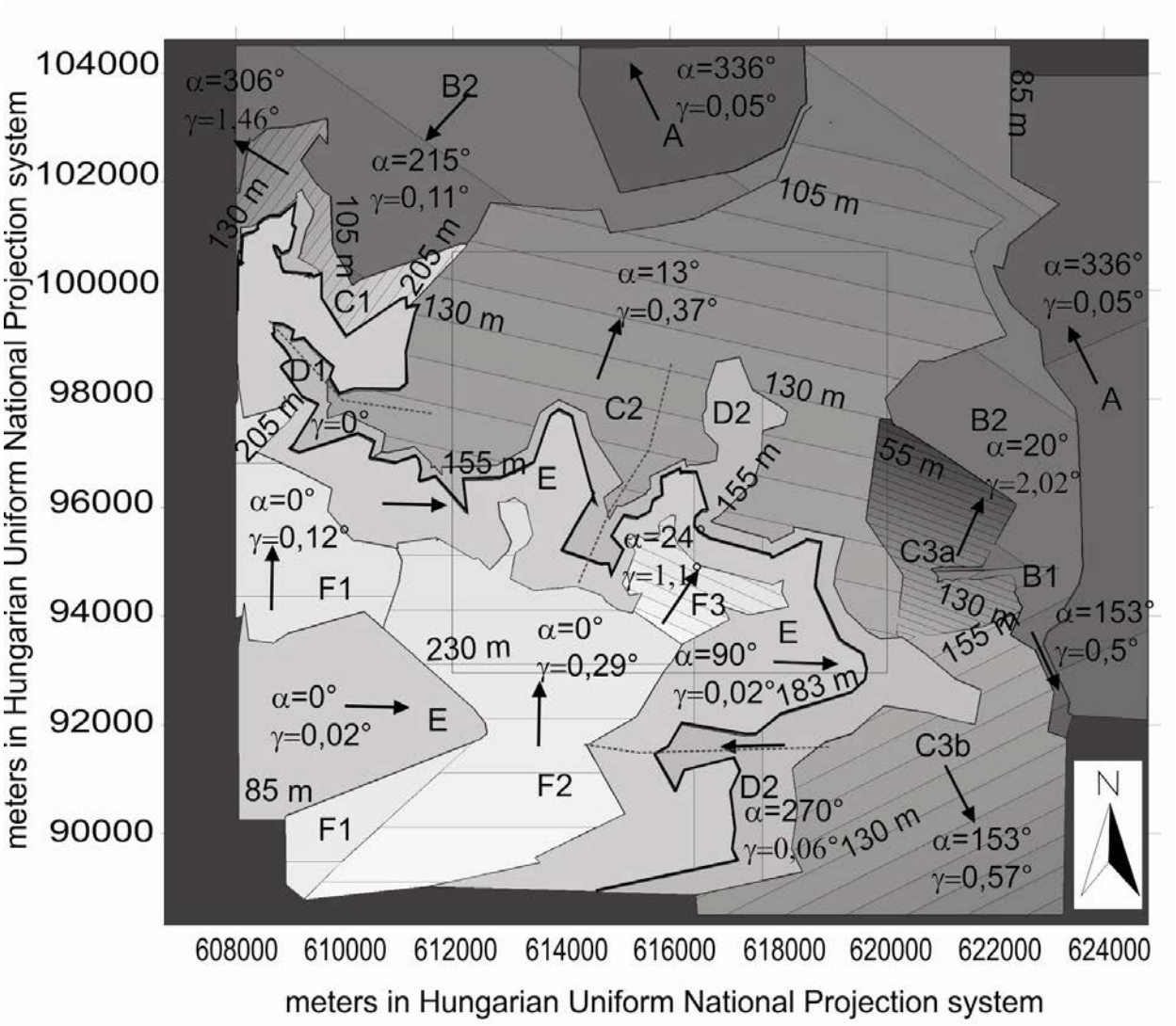

Figure 6. Homogeneous populations of the datasets based on one variable analysis of height values of the borehole pre-Quaternary surface. Seven of the nine datasets shows the border around 183 m a.s.l. The second largest boundary runs at 205m a.s.l.

The other trend sub-population at the margin of the target area, area B1 verges southward and it is interpreted as the southern slope of the Szekszárd Hills (Fig. 6). On the whole, the Palaeozoic granite pluton suffered an asymmetric denudation along the NE-SW transect. The fitted trend surfaces capture and quantitatively characterise the paleo-slope conditions that are also discernible in the manually constructed geological maps. According to trend analysis, zones around the major 183 m level (zones D and E) are horizontal. This is consistent with the contour lines of height levels of pre-Quaternary surface shown in the geological map by Balla (2003). Trends around the elevated pluton diverge in the northerly directions (Fig. 6) suggesting that these areas are slightly tilting eroded surfaces. Finally, the Szekszárd Hills and the area below of the River Danube alluvium (zones B1 and A) in the north slopes to south, similar to the topographic surface (Fig. 1a).

Despite of the overall northerly slope of the target area, the presence of sub-areas with various tilts including horizontal areas show that no overall tectonic tilt of the area can be assumed and the obtained overall tilt and slopes of sub-areas is attributable to erosion processes.

The goodness of fit of the fitted surfaces, local population-based trends and smoothed TIN, was described by the calculation of the differences between the model surfaces and the depth values measured in the boreholes. The median central value of the hence obtained residuals is zero in both cases showing the lack of systematic error, thus confirming the appropriateness of the selected models. Also, the overall variations of differences as measured by the robust MAD statistics are equally a small value (42m) in both cases. This shows that the two models represent similarly well the pre-Quaternary surface, in essence. However, there are numerous high outlying values among the residuals

for the local trend surface, unlike for the smoothed TIN. The smoothed TIN model based on the local moving average window is more sensitive to local deviations and follow closer the original data points, while the local trend fits based on data excluding outlying values (see above) really capture the major overall tendencies in the studied geological unit. The distribution of residuals for the local trend surface is homogeneous and systematically distributed, while the smoothed TIN residuals display sub-populations. Again, the least-squares trend planes properly model the surface tendencies producing an essentially white noise error, while the smoothed TIN surface model does not consider the identified depth sub-populations and is sensitive to sudden changes at population boundaries. Thus, the smoothed surface is appropriate for describing the terrain undulations such as major paleo-valleys but the local trend fit based on thorough statistical analysis is the proper model for describing accurately and quantitatively the characteristic tendencies of the studied pre-Quaternary surface resulting from geological processes.

According to the rose diagram constructed for lineaments identified in the shaded relief models of each successive smoothed TIN surface model, the prevailing lineament orientation is NE-SW in the target area (Figs 5b, c d). This is in good agreement with the findings of detailed tectonic investigations (Maros, 2003). It is interesting that shorter linear morphological features in all directions are discernible at the 500x500m average smoothing window size (not shown in this paper), while at the largest smoothing of 1000x1000m only the NE-SW morphological feature such as major valleys and ridge lines remain in the model (Fig. 5). In this way we captured the spatial scales of the paleo-surface morphology that can be divided into features at scale below and above about 1000m.

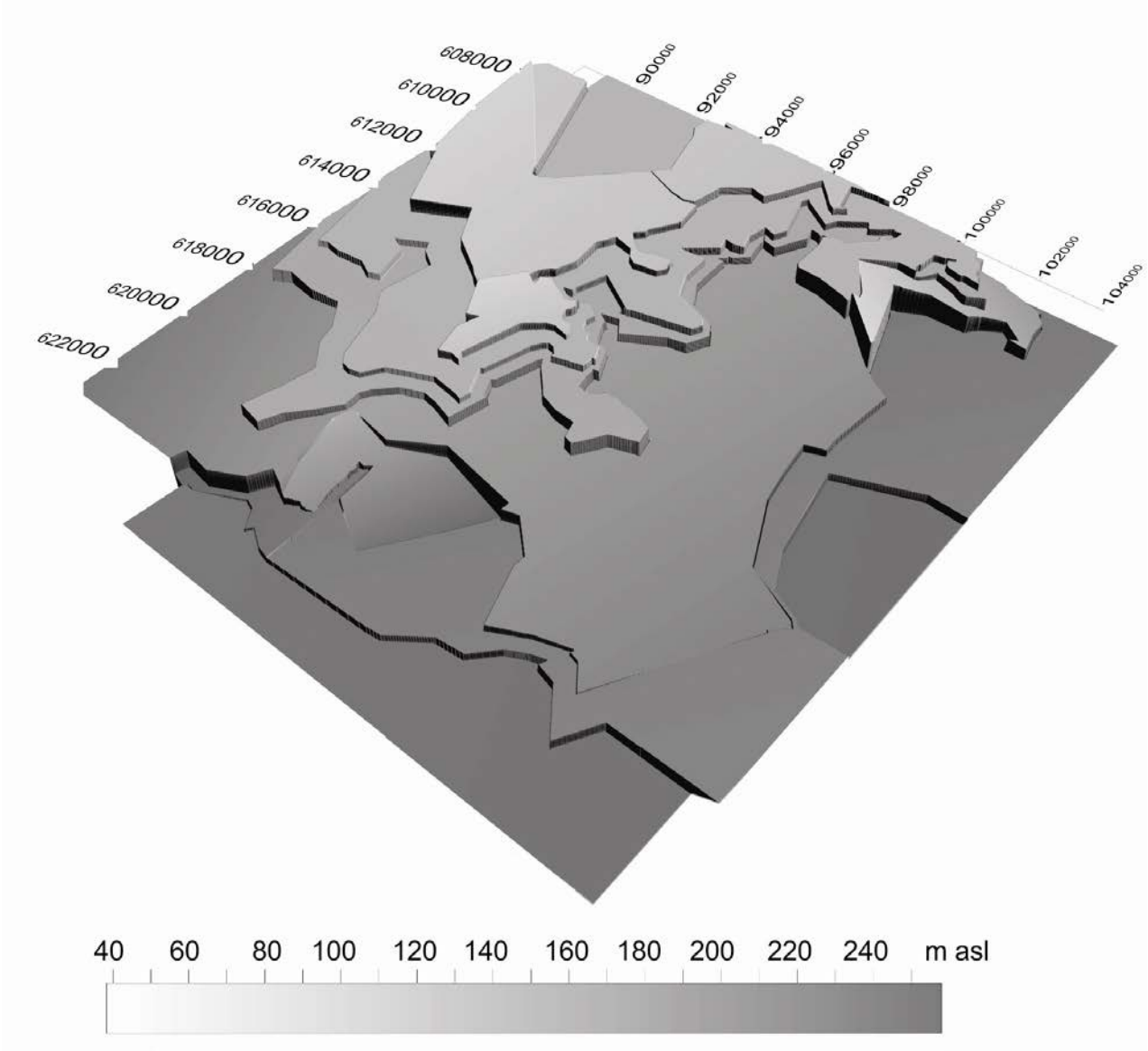

Figure 7. 3D representaion of surface model shown in figure 6.

Since the major lineaments are valleys along regional fault lines (Csontos et al., 2002), it can be concluded that the spatial scale of morphotectonic processes is >1000m, while secondary erosion processes forming the smaller side valleys act at scales below 1km.

The tectonically pre-defined negative geomorphological features of paleo-valleys are present in all the six concentric zones corresponding to the statistically identified populations (Fig. 3b). Thus, these valleys post-date the formation of the paleo-surface horizons. Furthermore, the traces of the pre-Quaternary surface valleys can be seen in the geologic map (Balla et al., 2009b) and they coincide with the recent valley lines in the DEM, too (Fig. 1a). Since these points are coherent belonging to the same statistical population, being statistically and spatially independent from the surrounding higher areas, the relationship between these valleys and the associated lower areas is demonstrated. It should be noted that trend plane boundaries mostly coincide with the lineaments identified in the TIN model suggesting that minor block moving could have occurred along these faults.

Finally, the identification of statistically homogeneous zones and the subsequent removal of trends yielded residuals that are stationary in the mean, a pre-requisite for creating a detailed surface model with kriging (Cressie, 1991), the subject of further research to be published elsewhere.

5. CONCLUSIONS

In this study the pre-Quaternary surface consisting of Palaeozoic granite and Pannonian marine sediments is analysed that is buried and thus not affected by recent erosion. The study is based on boreholes drilled in the last century in the target area for the National Low and Intermediate Level Radioactive Nuclear Waste Repository near Bátaapáti, Hungary. Detailed data preparation, including outlier identification, data errors in source database, homogeneity tests for various data sources, definition of inner and outer modelling kernel to handle interpolation edge effects and careful data analysis yielded consistent geomorphological results. A novel finding was the identification of statistically significant zones of elevation that can be interpreted in terms of paleo-shoreline and associated slope scree surfaces.

Trend surfaces capture and describe various erosional surfaces on and around the Paleozoic granite body. Stationary residuals remaining after trend removal enable detailed surface reconstruction by kriging interpolation.

Smoothing of the TIN surface model with successively increasing moving average window revealed the dominant NE-SW tectonically pre-defined pre-Quaternary paleo-valley network and its spatial scale exceeding erosion features above 1km. The identified paleo-valleys are present in three independent depth populations, they cross cut the spatial population boundaries showing that these valleys are important morphological features.

In this study, two paleo-surface reconstruction models were built, compared and interpreted. The trend model describes the overall paleo-geomorphological tendencies at spatial scales larger than paleo-valley formation (Coleman-Michael et al., 2009) and captures processes like slope mass movement and associated morphological features and regional marine shoreline formation.The smoothed TIN model describes valley formation along fault lines and smaller erosional valley incision. It is noted that both models are low-pass filters in terms of signal processing leaving only low frequency, large morphological features in the output surfaces.

Acknowledgements

This paper reports on the research at the GEM-RG Geochemistry, Modelling and Decisions Research Group

REFERENCES

Aunon, J. & Gomez-Hernandez, J., 2000. *Dual kriging with local neighborhoods; application to the representation of surfaces*, Mathematical Geology. 32; 1 69-85.

Bakkali, S. & Amrani, M., 2008. *About the use of spatial interpolation methods to denoising Moroccan resistivity data phosphate "Disturbances" map*, Acta Montanistica Slovaca. 13; 2 216-222.

Balla Z., Gulácsi Z., Maros Gy. & Síkhegyi F., 2009a. *Relief and Geological Map of the Pre-Cainozoic Basement, in: Geology of the North-eastern Part of the Mórágy Block*, 2. Enclousre, Geological Institute of Hungary, Budapest, pp. 283

Balla Z., Gulácsi Z., Maros Gy. & Síkhegyi F., 2009b. *Relief and Geological Map of the Pre-Quaternary Complexes, in: Geology of the North-eastern Part of the Mórágy Block*, 2. Enclousre, Geological Institute of Hungary, Budapest, pp. 283

Balla Z., 2003. *Az atomerőművi kis és közepes aktivitású radioaktív hulladékok végleges elhelyezésére irányuló program. A felszíni földtani kutatás zárójelentése, Bátaapáti (Üveghuta), 2002-2003; A Bátaapáti (Üveghutai)-telephely környezetének földtani térképsorozata - 6. melléklet: A negyedidőszaki üledékek feküdomborzati térképe*, Magyar Állami Földtani Intézet, Budapest. [in Hungarian: *Programme for final disposal of low- and intermediate-level radioactive waste from the nuclear power plant. Final report of the geological exploration from the ground surface, Bátaapáti (Üveghuta), 2002–2003]. – 6. Enclosure: Map of the Contour Lines at the Base of Quaternary*

Sediments ,Geological Institute of Hungary, Budapest] pp. 392.

Balla Z., 2004a. *General characteristics of the Bátaapáti (Üveghuta) Site (South-western Hungary)*, Annual Report of the Geological Institute of Hungary, 2003, pp. 73-85.

Balla Z., 2004b. *Pannonian sediments of the north-eastern Mórágy Block*, Annual Report of the Geological Institute of Hungary, 2003, pp. 333-337.

Bardossy, A.& Li, J., 2008. *Geostatistical interpolation using copulas*, Water Resources Research. 44; 7.

Bistacchi, A., Massironi, M., Dal-Piaz, Giorgio-V, Dal-Piaz, Giovanni, Monopoli, B., Schiavo, A. & Toffolon, G., 2008. *3D fold and fault reconstruction with an uncertainty model; an example from an Alpine tunnel case study*, Computers and Geosciences. 34; 4, 351-372.

Boehm, G., Ocakoglu, N., Picotti, S. & De-Santis L., 2009. *West Antarctic ice sheet evolution; new insights from a seismic tomographic 3D depth model in the eastern Ross Sea (Antarctica)*, Marine Geology. 266; 1-4, 109-128.

Bhuiyan-Anwar, H., 2009. *Three-dimensional modelling and interpretation of CSEM data from offshore Angola*, Petroleum Geoscience. 15; 2, 175-189.

Carrera, N., Munoz, J. A. & Roca, E., 2009. *3D reconstruction of geological surfaces by the equivalent dip-domain method; an example from field data of the Cerro Bayo Anticline (Cordillera Oriental, NW Argentine Andes)*, Structural Geology. 31; 12 1573-1585.

Coleman, M. L., Niemann-Jeffrey, D. & Jacobs-Elaine, P., 2009. *Reconstruction of hillslope and valley paleotopography by application of a geomorphic model*, Computers and Geosciences. 35; 9, 1776-1784.

Cressie, N. A. C., 1991, *Statistics for Spatial Data*, John Wiley and Sons, Inc., New York, 900.

Császár Géza, 2004. *Alpine burial history of the Mórágy Block and its environs*, Annual Report of the Geological Institute of Hungary, 2003, pp. 395-401.

Csontos, L., Benkovics, L., Bergerat, F., Mansy, J. L. & Wórum, G., 2002. *Tertiary deformation history from seismic section study and fault analysis in a former European Tethyan margin (the Mecsek–Villány area, SW Hungary)*, Tectonophysics 357, 81–102

De-Donatis, M., Borraccini, F. & Susini, S., 2009. *Sheet 280-Fossombrone 3D; a study project for a new geological map of Italy in three dimensions*, Computers & Geosciences 35; 1, 19-32

Erdogan, S., 2010. *Modelling the spatial distribution of DEM error with geographically weighted regression; an experimental study*, Computers and Geosciences 36; 1, 34-43

Ersoy, A., Yunsel, T.Y. & Cetin, M., 2004. *Characterization of land contaminated by past heavy metal mining using geostatistical methods*, Archives of Environmental Contamination and Toxicology 46; 2 162-175

Eto, C., Ishihara, Y., Tanabe, S., Kimura, K. & Nakayama, T., 2008. *Three dimensional models of N-values and lithofacies by using borehole logs; an example of incised valley fills under the northern part of the Tokyo Lowland, central Japan*, Chishitsugaku Zasshi = Journal of the Geological Society of Japan. 114; 4, 187-199

Fodor, L., Bada, G., Csillag, G., Horváth, E., Ruszkiczay-Rüdiger, Zs. & Síkhegyi, F., 2005. *New data on neotectonic structures and morphotectonics of the western and central Pannonian Basin*, Occasional Papers of the Geological Institute of Hungary, volume 204 35-44

Goovaerts, P., 2009. *AUTO-IK: A 2D indicator kriging program for the automated non-parametric modeling of local uncertainty in earth sciences*, Computers &Geosciences 35 1255–1270.

Guibas, L., & J. Stolfi, 1985. *Primitives for the Manipulation of General Subdivisions and the Computation of Voronoi Diagrams*, ACM Transactions on Graphics, v. 4, n. 2, pp. 74-112.

Gyalog L. & Szegő I., 2004. *Boreholes at the Bátaapáti (Üveghuta) Site*, Annual Report of the Geological Institute of Hungary, 2003, pp. 93-107.

Gyalog L., Havas G., Maigut V., Maros Gy. & Szebényi G., 2004. *Geological-tectonic documentation in the Bátaapáti (Üveghuta) Site*, Annual Report of the Geological Institute of Hungary, 2003, pp. 171-187.

Hoaglin D. C., Mosteller F. & Tukey J. W., 1983. *Understanding Robust and Exploratory Data Analysis*. John Wiley and Sons Inc., New York., pp 472.

Huang-Guo-Chin, Wu-Francis, T., Roecker-Steven, W. & Sheehan-Anne, F., 2009. *Lithospheric structure of the central Himalaya from 3-D tomographic imaging*, Tectonophysics. 475; 3-4, 524-543.

Jahn, A. & Riller, U., 2009. *A 3D model of first-order structural elements of the Vredefort Dome, South Africa; importance for understanding central uplift formation of large impact structures*, Tectonophysics 478; 3-4, 221-229.

Jordan G., Meijninger B.M.L., van Hinsberen D.J.J., Meulenkamp J.E. & van Dijk P.M., 2005. *A GIS framework for digital tectonic geomorphology: case studies.* International Journal of Applied Earth Observation & Geoinformation, 7:163-182.

Jordan, G., 2007. *Adaptive smoothing of valleys in DEMs using TIN interpolation from ridgeline elevations: An application to morphotectonic aspect analysis*, Computers & Geosciences 33, 573–585.

Jordan G., 2007. *Digital terrain modelling in a GIS environment. In: Peckham R.J. and Jordan (eds), 2007. Digital Terrain Modelling. Development and Applications in a Policy Support Environment*, Series: Lecture Notes in Geoinformation and Cartography. Springer, Berlin, pp. 1-44.

Király E. & Koroknai B., 2004. *The magmatic and metamorphic evolution of the north-eastern part of the Mórágy Block*, Annual Report of the Geological Institute of Hungary, 2003, pp. 299-310.

Li, J. & Heap, A., 2008. *A review of spatial interpolation methods for environmental scientists*, Geoscience Australia. Canberra, A.C.T., Australia. pp 137.

Mann, H. B. & Whitney, D. R., 1947. *On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other*, Annals of

Mathematical Statistics 18, 50–60., Ohio State University, Columbus, USA.

Marache, A., Dubost, J., Breysse, D. & Denis, A., 2009. *Understanding subsurface geological and geotechnical complexity at various scales in urban soils using a 3D model*, Georisk. 3; 4, 192-205.

Maros Gy., 2003. *Az atomerőművi kis és közepes aktivitású radioaktív hulladékok végleges elhelyezésére irányuló program. A felszíni földtani kutatás zárójelentése, Bátaapáti (Üveghuta), 2002-2003; A Bátaapáti (Üveghutai)-telephely környezetének földtani térképsorozata – 9. melléklet: tektonikai térkép*, Magyar Állami Földtani Intézet, 2003. [in Hungarian: *Programme for final disposal of low- and intermediate-level radioactive waste from the nuclear power plant. Final report of the geological exploration from the ground surface, Bátaapáti (Üveghuta), 2002–2003]. – 9. Enclosure: Tectonics Map*, Geological Institute of Hungary, Budapest] pp. 392.

Maros Gy., Koroknai B., Palotás K., Fodor L., Dudko, A., Florián-Szabó M., Zilahi-Sebess L. & Bán-Győry E., 2004. *Tectonic analysis and structural evolution of the north-eastern Mórágy Block*, Annual Report of the Geological Institute of Hungary, 2003, pp. 371-386.

Maros Gy., 2006. *A Mórágyi Gránit szerkezeti fejlődése az IMAGEO magszkennerrel történt értékelések alapján*, Doktori értekezés, Mikoviny Sámuel Doktori Iskola, Miskolci Egyetem, pp. 73-80. [In Hungarian: *Structural evolution of the Mórágy Granite based on ImaGeo core loggings*, PhD dissertation, University of Miskolc], pp. 143.

Marsi I., Don Gy., Földvári M., Koloszár L., Kovács-Pálffy P., Krolopp E., Lantos M., Nagy-Bodor E. & Zilahi-Sebess L., 2004. *Quaternary sediments of the north-eastern Mórágy Block*, Annual Report of the Geological Institute of Hungary, 2003, pp. 343-359.

Maxelon, M., Renard, P., Courrioux, G., Braendli, M. & Mancktelow, N., 2009. *A workflow to facilitate three-dimensional geometrical modelling of complex poly-deformed geological units*, Computers and Geosciences 35; 3 644-658.

McGill, R., Tukey, J. W. & Larsen, W. A., 1978. *Variations of Box Plots*; The American Statistician 32 (1): 12–16.

Merwade, V., 2009. *Effect of spatial trends on interpolation of river bathymetry*, Journal of Hydrology. 371; 1-4 169-181.

Peckham R.J. & Jordan G. (editors), 2007. *Digital Terrain Modelling. Development and Applications in a Policy Support Environment*. Series: Lecture Notes in Geoinformation and Cartography. Springer, Berlin. pp. 313

Piotrowska, K., Ostaficzuk, S., Malolepszy, Z. & Rossa, M., 2005. *The numerical spatial model (3D) of geological structure of Poland, from 6000 m to 500 m b.s.l.*, Przeglad Geologiczny. 53; 10/2, 961-966.

Reimann C., Filzmoser P., Garrett R. & Dutter R., 2008. *Statistical Data Analysis Explained: Applied Environmental Statistics with R*. John Willey and Sons Ltd., England. pp. 343

Sebe Krisztina, 2009. *A Nyugat-Mecsek és környezete tektonikus geomorfológiai elemzése*, PhD Dissertation, University of Pécs, Hungary, pp. 68 [In Hungarian: *Tectonic geomorphological analysis of the Western Mecsek Mts. and their surroundings*]

Susini, S. & De-Donatis, M., 2009. *3D model of a sector of the south Scotia Ridge (Antarctica)*, Computers and Geosciences 35; 1, 83-91.

Tonini, A., Guastaldi, E., Massa, G. & Conti, P. 2008. *3D geo-mapping based on surface data for preliminary study of underground works: A case study in Val Topina (Central Italy)*, Engineering Geology 99, 61–69.

Tonini A., Guastaldi, E. & Meccheri, M., 2009. *Three-dimensional reconstruction of the Carrara Syncline (Apuane Alps, Italy); an approach to reconstruct and control a geological model using only field survey data*, Computers and Geosciences 35; 1 33-48.

Tukey, J.W., 1977 *Exploratory Data Analysis*, Addison-Wesley. pp. 688

Velleman P. F. & Hoaglin D. C., 1981. *Applications, Basics and Computing of Exploratory Data Analysis*, Duxbury Press, Boston. pp. 354

Wan, Y., Wang, X. & Zhang, J., 2002. *The construction of a GIS granite data base in China*, Yanshi Kuangwuxue Zazhi = Acta Petrologica et Mineralogica. 21; 2 186-191.

Walsh G.J., 2009. *A method for creating a three dimensional model from published geologic maps and cross sections*, Open-File Report - U. S. Geological Survey, US Geological Survey, Reston, VA, United States. pp. 16

Wilcox, R. 2005. *Kolmogorov–Smirnov Test*, In: Encyclopedia of Biostatistics, University of Southern California, Los Angeles, CA, USA.

Wycisk, P., Hubert, T., Gossel, W. & Neumann, C., 2009. *High-resolution 3D spatial modelling of complex geological structures for an environmental risk assessment of abundant mining and industrial megasites*, Computers and Geosciences. 35; 1, 165-182.

Zanchi, A., Francesca, S., Stefano, Z., Simone, S. & Graziano, G., 2009. *3D reconstruction of complex geological bodies: Examples from the Alps*, Computers & Geosciences 35, 49–69.

Received at: 16. 03. 2014
Revised at: 02. 08. 2014
Accepted for publication at: 14. 10. 2014
Published online at: 20. 10. 2014