AWK and GNU Octave Programming Languages Integrated with Generic Mapping Tools for Geomorphological Analysis

Polina Lemenkova

AWK AND GNU OCTAVE PROGRAMMING LANGUAGES INTEGRATED WITH GENERIC MAPPING TOOLS FOR GEOMORPHOLOGICAL ANALYSIS

Polina Lemenkova

Ocean University of China, College of Marine Geo-sciences, 238 Songling Rd, Laoshan, 266100, Qingdao, China (P. R. C.)
ORCID: 0000-0002-5759-1089
Email: pauline.lemenkova@gmail.com

ABSTRACT

The paper presents a sequential use of the AWK and GNU Octave programming languages integrated with Generic Mapping Tools (GMT) for geospatial data analysis. The geographic scope of the research is focused on the Kuril-Kamchatka Trench, north Pacific Ocean. Practical research aim is to analyse and compare bathymetry in the southern and northern part of the trench using digitized cross-section profiles. The initial mapping and geospatial analysis was performed in GMT scripting toolset. The GMT was used for cartographic mapping based on the raster ETOPO1 grid and automatic digitizing of the profiles crossing the trench perpendicularly. Besides visualized map, the processed geodata were received in a numerical form as a complex multi-field table for each segment. These tables were generated by the GMT in its native format and could not be directly processed by the MATLAB/Octave. Therefore, the tables were exported to AWK, a data-driven programming language and a powerful tool for data extraction. The table was then restructured, sorted and reshaped by the AWK script. Because the total amount of profiles overstepped 100 (62 and 52 for the northern and southern trench segments), only selected profiles were visualized. For this purpose, at the next step the modified tables were converted to GNU Octave language for visualizing and plotting selected profiles. Finally, the geomorphology was analysed and two segments compared. The results show that the southern part has deeper bathymetric values, vary in geomorphic structure and has steeper gradient slopes comparing to the north, which is caused by the seismicity, volcanism, geologic and tectonic settings. Three full scripts of GMT, AWK and GNU Octave programming languages are presented for replicability in the Appendix.

Keywords: AWK, Octave, GMT, programming language, data analysis, geospatial analysis, trench

1   INTRODUCTION

The conceptual idea of this research consists in a chain of data processes by GMT cartographic scripting toolset and programming languages AWK and Octave. The experiment aims at the analysis of the geomorphic differences of the Kuril-Kamchatka Trench located in the north-west Pacific Ocean (Figure 1) through the analysis of its cross sectioning profiles converted to the AWK and Octave for further processing and plotting. The geographic aim is to visualize the difference in the northern and southern parts of the trench separated by the Bussol Strait (Figure 1). The experimental batch sequence of the data processing through three different tools is advantageous, because they were used sequentially for each step of the data analysis where they suit best: the GMT for mapping, geospatial data analysis and generating series of profiles, AWK for reshaping complex tables, and Octave for plotting graphs of the selected profiles.

The experimental novelty of this research is integration of the three tools, – GMT scripting toolset, AWK and Octave programming languages, – combined together for the geospatial data analysis and modelling aimed to study submarine geomorphology of the hadal trench. While traditional geographic data analysis is mostly based on the GIS software, the technical aim of this research is to demonstrate successful application of the advanced tools of data analysis applied specifically for the geospatial research: AWK and Octave are traditionally used in the IT domains, while geological mapping is mostly based on GIS.

Therefore, the main goal is to develop a technical methodology for the batch processing of the geospatial data by the smooth stepwise GMT-AWK-Octave workflow chain. This technical workflow also supported geomorphic studies helping to visualize and better understand structure of the trench. Usually, similar research is maintained in the GIS software, but applying advanced data processing scripting tools facilitates research through automation. Automatically digitized transect profiles are precise, finely plotted and statistically correct comparing to the hand-made vector drawing using GIS (e.g. ArcGIS or QGIS).

Kuril–Kamchatka Trench area and cross–section profiles

Track segments: northern (red) and southern (green) separated by the Bussol Strait.
Cross–sectional profiles: 400 km long, spaced 10 km, sampled every 2 km.
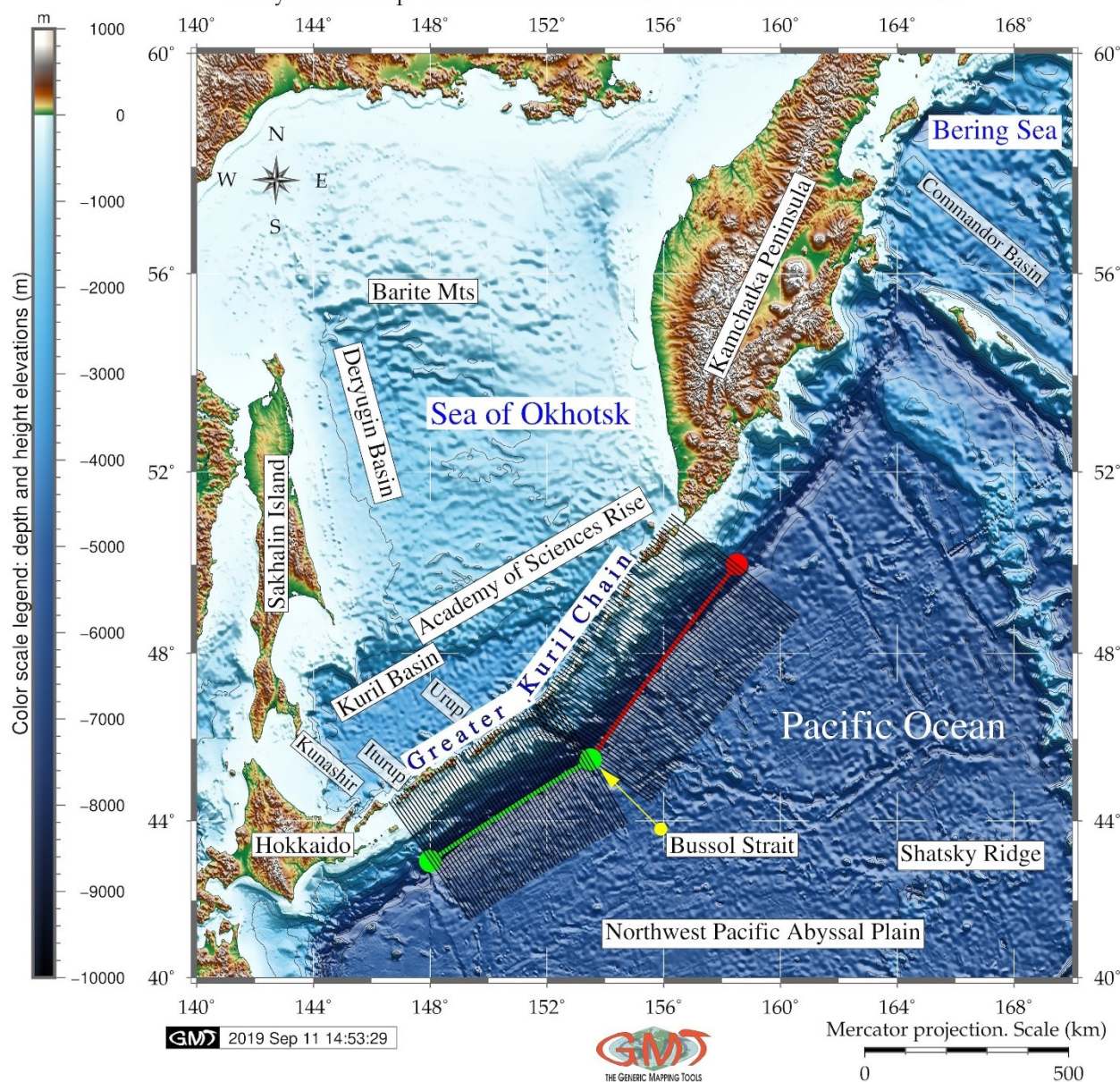Bathymetric map: ETOPO1 Global Relief Model 1 arc min resolution

Figure 1. Bathymetric chart of the study area: the Sea of Okhotsk and Kuril-Kamchatka Trench. (140°E/170°E/40°N/60°N), contour interval 1,000 m. Seafloor topography: ETOPO1. Cross-section profiles shown for the southern (green) and northern (red) study segments

For this purpose, the profiles were automatically digitized by the GMT and then the data were AWK, a domain-specific programming language designed for the text processing, was used in this work for table reshaping. The AWK programming language was created at Bell Labs in the 1970s [1] and its name is derived from the surnames of its authors: Alfred Aho, Peter Weinberger, and Brian Kernighan. The GNU implementation of AWK is called GAWK. The available documentation and information on AWK was used extensively in this research [2], [3], [4]. AWK language is highly suitable for sorting multi-field records in a table due to its specially designed syntax and embedded algorithms. Other possibilities include managing small databases, generate reports based on tables, validating data, producing indexes, performing document-preparation tasks, extracting pieces of data for further processing, sorting data [5]. AWK is a data-driven scripting language integrated with commands of Unix operating system and can be run from the console. It consists of a set of codes-actions that can be taken against to operate with streams of table data [6]. AWK enables to extract or transform text, produce formatted records [7]. It uses string datatype, associative arrays indexed by key strings, and regular expressions [8]. Based on this, AWK was selected as a powerful advanced tool for the processing of tables.

The GNU Octave (https://www.gnu.org/software/octave/) is an open source, high-level interactive language, designed for numerical computations by J. B. Rawlings and J. G. Ekerdt. It is compatible with Matlab and is suitable for a variety of scientific tasks: linear and nonlinear equations, numerical linear algebra, statistical analysis, numerical experiments [9]. A significant well-known advantage of the Octave consists in its compatibility with Matlab: although there are some minor differences in their syntax, GNU Octave is known as a free open-source analogue of the commercial licensed Matlab which is widely used in both academia and industry by thousands of researchers and engineers worldwide. Therefore, Octave scripts may be converted to the Matlab environment using the syntax and power of Matlab. Technically, as Matlab, GNU Octave has various built-in functions, operates with arrays, vectors, strings, booleans, scalars and matrices, enables plotting of high-quality graphics, performs mathematical computations, operates with variables and different data types, as demonstrated further in this paper (see Appendix, Script#3 for GNU Octave). The Octave was used in this work for plotting several geomorphological profiles.

The GMT scripting toolset was selected as a geospatial tool for raster grid processing, mapping and automated digitizing of the geomorphological profiles. GMT was initially developed by the American scientists, P. Wessel and P. Smith [10], further maintained and improved to finally become a powerful open source scripting toolset for the geospatial analysis and cartographic visualization. GMT is completely based on the command-line usage (see Appendix, Script #1 for the example of the GMT code presented in this paper). That being said, it is not a classical GIS. Using the full powerful functionality of GMT requires writing a script with numerous arguments for the fine-tuning and setting up of every cartographic element: projections, major and minor grids, colour palettes, map legends, visualized vector elements and annotations, adding map elements (e.g. coastlines, directional roses, scale bars) etc. However, the high cost of the steep learning curve of GMT as well as its non-visual command line based operation pays off by its powerful functionality and excellent visualization. GMT is compatible with Unix commands (cut, rm, echo) and GDAL raster processing, it operates with both raster and vector data, produces high-quality maps and figures, has embedded functions for the descriptive statistical analysis (e.g. histograms, median, mean). All this makes it one of the most preferred data visualization and mapping toolsets by cartographers and professional geoscientists.

2 GEOGRAPHIC SETTINGS

The geographic focus of this work lies on the Kuril-Kamchatka Trench, an oceanic trench with depths more than 5000 m, located in the Sea of Okhotsk, north-west Pacific Ocean (Figure 1). Geographically, the Kuril-Kamchatka Trench continues south-eastwards from the coast of the Kamchatka Peninsula, in parallel to the Kuril Island chain, and finally meets Hokkaido Island. Geologically, the Kuril-Kamchatka Trench presents a boundary of the large subduction zone of the Pacific tectonic plate under the Okhotsk tectonic plate [11], which causes irregularity in its bathymetric shapes and geomorphic structure. The submarine geomorphology of the seafloor underlying the Kuril-Kamchatka Trench is formed as a result of the complex interplay between the geological substrate, water column and a range of the geophysical processes and phenomena. As a result, the variations in the distribution of the geomorphic landforms in the trench were caused by a continuous effect of various factors through the whole history of the trench formation. The trench is characterized by a geometrically long narrow, elongate, deep and asymmetrical depression of the seafloor with steep slopes [12]. Together with Aleutian Trench, the Kuril–Kamchatka Trench presents contiguous complex covering area of 254,740 km^2 [13].

The Kuril-Kamchatka Trench is distributed along Kuril island arcs and Eurasian continent masses, developed in the association with an island-arc system where Pacific tectonic plate subducts below the North American and Okhotsk plates. The Okhotsk plate represents a region of active grabens lying southwards of a chain of small sedimentary basins in the Cherskii Mountains [14]. The geomorphic and topographic patterns of the land area of Kamchatka Peninsula are influenced by the climate and glaciological effects (glacier cirques) [15]. The Kuril-Kamchatka Trench is divided by the Bussol Strait into southern and northern parts. Because the geologic factors differ in northern and southern parts of the trench, it affects its geomorphology. The southern part of the Greater Kuril Chain includes the large Kunashir, Iturup, and Urup islands constituting the islands-arc-trench system. This part is confined to the northwestern Pacific subduction zone [16]. The topographic and geomorphic features of the trench in its northern part are different from those located further southwards near the group of the seamounts.

3 METHODOLOGY

3.1 Methodological summary

The research aim of this paper is to analyse regional geomorphological variations in the trench in its southern and norther parts through the technical application of GMT, AWK and Octave tools. The methodological workflow aims at stepwise combination of AWK, GMT and GNU Octave, which is reflected in the methodological structure of this paper that consists of the three parts integrating advanced data processing. Brief explanation of the methodology is as follows:

1. The first step consisted in bathymetric mapping, modelling and plotting profiles using GMT by the code provided in Appendix (Script #1). The geospatial analysis was performed using GMT modules for cartographic mapping and automatic digitizing of the orthogonal profiles. As a result of this step the map was visualized (Figure 1) using ETOPO 1 raster grid and the profiles were generated using GMT code (Script #1). The enlarged view of the profiles and their numbering are shown in Figure 2. The stacked profiles with a median as a red line and error bars are shown in Figure 3. The profiles contained information on coordinates and depths in each point across the profile (every 2 km). Sixty-two profiles were plotted for the northern and 52 profiles for the southern parts of the trench. These data were stored as two tables in GMT format. The fragment of the initial table is shown in Table 1. The geological conditions associated with the studied geomorphology were analysed using available literature on the geology of the Kuril-Kamchatka Trench. During GMT mapping (Figure 1), the geospatial data visualization methods included the following ones: selecting cartographic projections, choosing fonts and cartographic layout, ticks, scale bar, legend, annotations, colour palettes, etc. The reason for the use of GMT lies in its high effectiveness and availability of various modules designed for various tasks: plotting lines and polygons, visualizing raster data sets, digitizing cross-section profiles, etc. Because the table produced by GMT automatically during digitizing process (Table 1) was not compatible with GNU Octave and could not be used directly for plotting, the AWK table reshaping was then performed in the second step.

2. The second step consisted of reshaping the tables that were generated automatically from GMT in the previous step. These tables were reshaped using AWK programming language provided in Appendix as Script #2. Only necessary columns and rows were selected from the initial tables using AWK script. As a result of this step, two tables were reshaped into the new ones, as shown in Table 2 (fragment) where each column now contained depth values by the profiles which could be processed by the Octave.

3. The third step consisted in the plotting graphs of the selected profiles by the GNU Octave code provided in Appendix as Script #3. After the table was converted to the necessary format through AWK manipulations (Table 2), it was then processed by the GNU Octave scripting and two graphs showing northern and southern segments of the trench were plotted (Figure 4 and Figure 5). The GNU Octave scripting was used for visualizing and plotting graph and geomorphological analysis for the selected profiles. The trench segments were visualized in graphs (Figure 4 and Figure 5) plotted based on the tables reshaped by AWK in the previous step. The profiles were selected taking into account the qualitative and quantitative assessment of data. Quantitative comparability consists in the repeatability of the sampling intervals: every 7th profile in the northern segment and every 5th profile in the southern segment until profile 40 and profile 52 respectively. Qualitative selection was based on the analysis of the bathymetric shapes: similar profiles were removed to avoid repeatability and only varying profiles were compared.

More detailed explanations of each of the three steps follow below (Sections 3.2, 3.3 and 3.4).

3.2 Mapping and modelling in GMT

The methodological workflow cycle is performed by means of the GMT scripting toolset [17]. The data were collected using automatic digitizing of the series of cross-sectioning orthogonal parallel profiles 400 km long, spaced 10 km, sampled every 2 km. These were stacked using modules of GMT focusing on the bathymetric sampling. Two sample parts of the trench were studied, northern and southern ones, to make a representation of the geomorphological shape patterns of the Kuril-Kamchatka Trench. The available data sets were modelled, visualized and analysed. To achieve high resolution data, the numerical publicly available data sources in .img and .grd format were added to the GMT project from the Scripps University of Oceanography San-Diego California, US ftp://topex.ucsd.edu/pub.
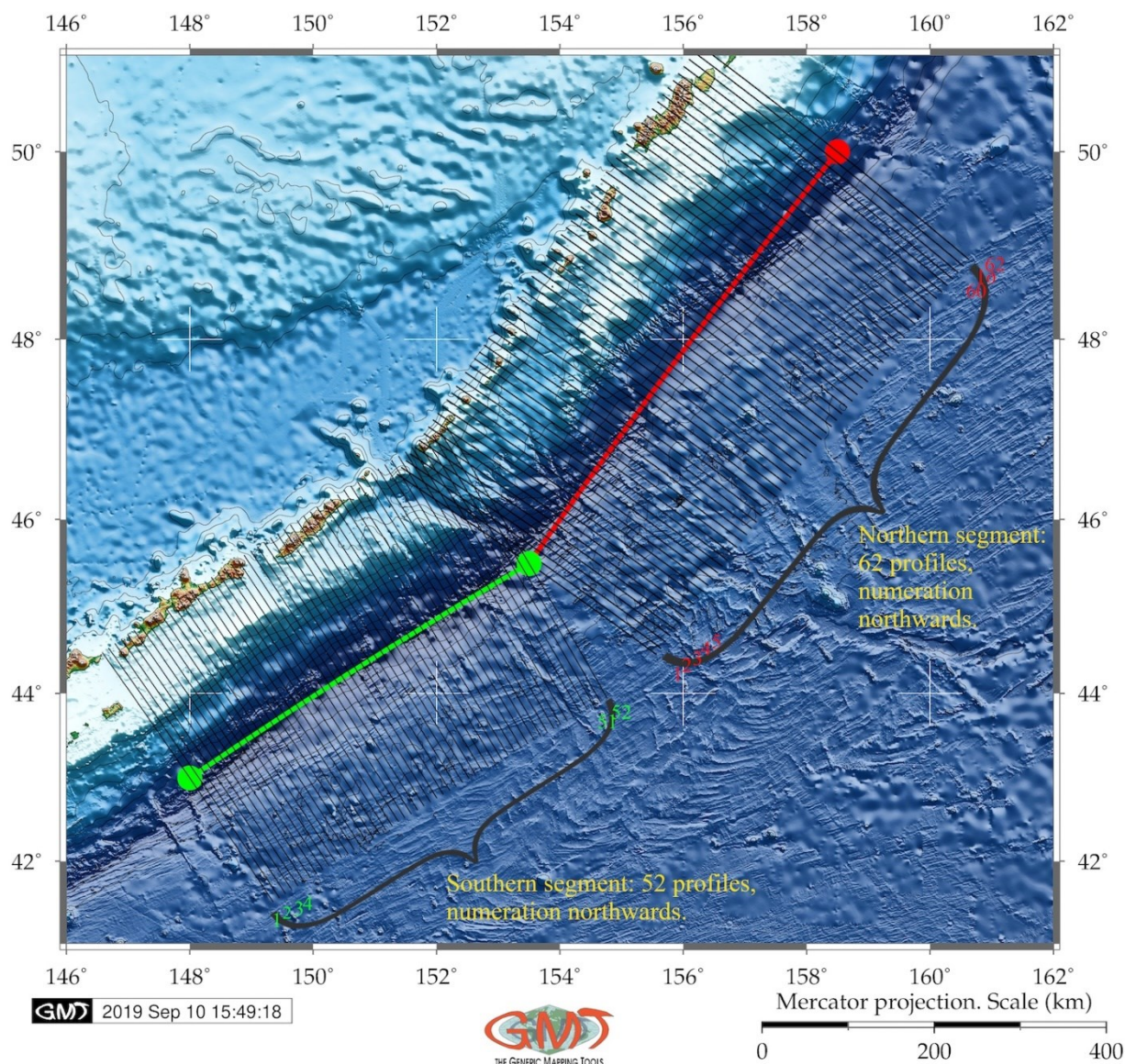
Figure 2. Enlarged view of the profiles: numbering starts from south to north for both segments

The data were then recompiled and reprocessed technically using combination of various GMT modules: gmtset, gmtdefaults, grdcut, makecpt, grdimage, psscale, grdcontour, psbasemap, psxy, grdtrack, convert, pstext, logo, psconvert. Following UNIX utilities were used additionally: echo (aim: placing the text), rm (removing tempfiles), cat (visualizing the file content in the shell console). Shorelines and continent vector data are taken from the available embedded GMT layers [18].
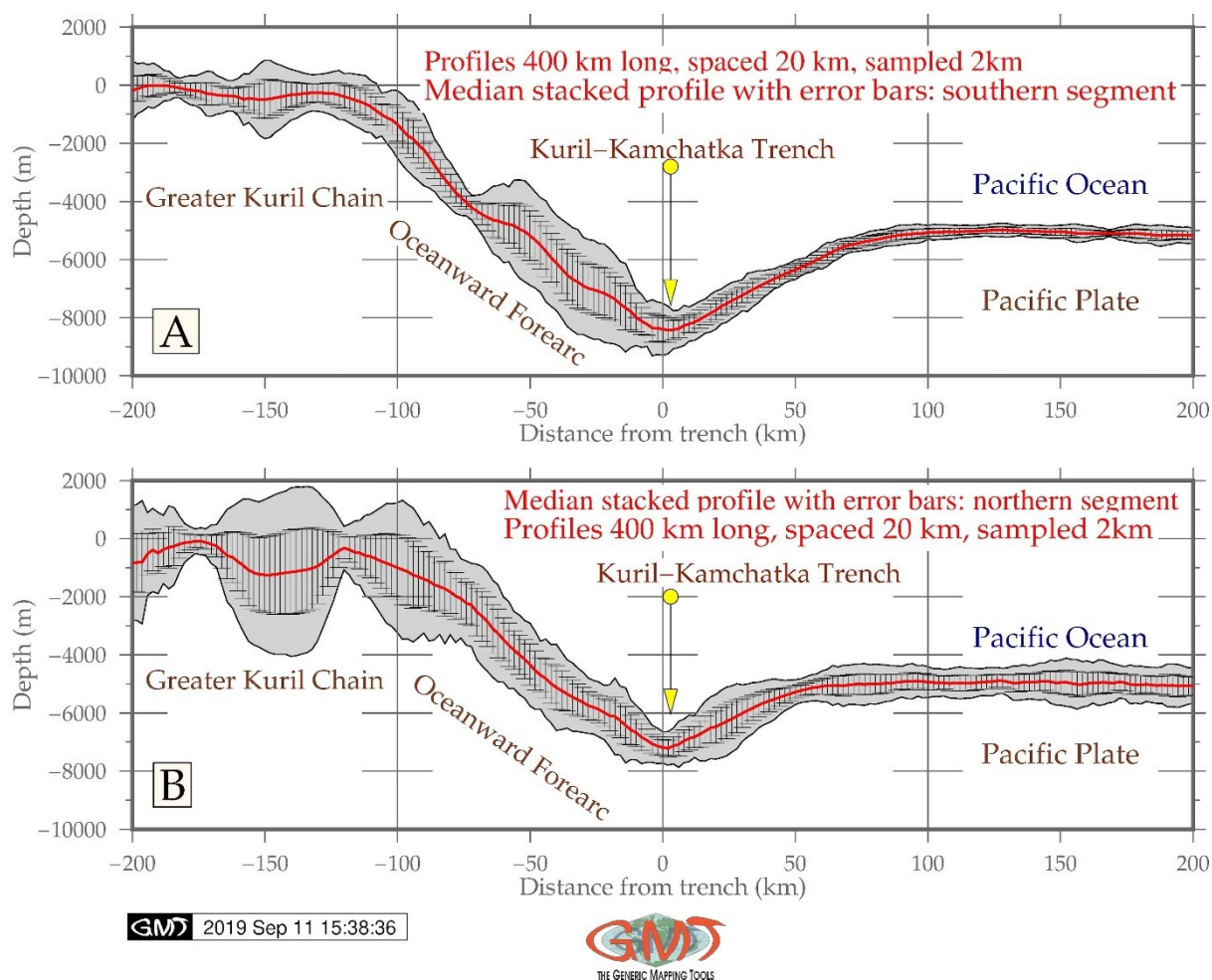
Figure 3. Median stacked profiles with error bars for the southern (A) and northern (B) segments

The essence of this processing is as follows (the explanations below refer to the presented Script#1, see Appendix for the full GMT code).

Step-1, the postscript (ps) file was generated in the Step-1 and saved in the working folder.

Step-2 includes the set-up of the GMT format (map aesthetics: font size, title offset, frames, etc.).

Step-3 includes overwriting of the default GMT values for the aesthetics by the new ones.

Step-4 includes cutting off necessary piece of map from the raster grid using 'grdcut' module. The coordinates of the square are given by the -R command (-R140/170/40/60).

Step-5 includes making colour palette using 'makecpt' module where the range is set up by -T.

Step-6 shows visualization of the raster using gmt 'grdimage' module where the -I+a15+ne0.75 shows illumination (shadows for the relief).

Step-7 consists in adding vertical colour legend. Because some of the annotations were made smaller than on the main map, the font set up was overwritten for this step.

Step-8 adds contour shorelines on the map, where -R -J means the same projection and region as on the main raster grid. '-C1000' means plotting contours every 1000 meters. '-W0.1,black' means the thickness and colour of the contour line.

Step-9 is adding the grid using 'psbasemap' module. -Bxg4f2a4 -Byg4f2a4 means adding major and minor grids every 4 and 2 degrees, and annotations every 4 degrees respectively.

Step-10 consists in dotting two points by coordinates (here for northern segment).

Step 11 means plotting the lines between these two points using module 'psxy'.

Steps 12, 13 and 14 consists in automatic digitizing profiles 400 km long, spaced 10 km, sampled every 2km using module 'grdtrack' and converting data into table by module 'convert'.

Steps 15-18 repeat the same procedure for the southern segment.

Step-19 consists in adding texts on the map by coordinates. For example, '-F+f16p,Times-Roman,blue+jLB' means text of 16 pt size, Times New Roman, blue and Left Bottom justified.

Step-20 is adding vector arrow pointing at the Bussol Strait.

Step-21 is adding a subtitle by the module 'pstext'.

Step-22 is adding the GMT logotype.

The last step is converting the map through the 'psconvert' module into the jpg file with 720 resolutions, portrait orientation and margins 0.5 c.

More detailed explanations can be found in the cartographic documentation of GMT [10].

3.3 AWK programming language for table reshaping

The AWK part of the research followed GMT aimed at table processing. The .csv table was run directly as a file following by a set of the temporary files (tempfiles) created to store the selected columns from the initial tables as part of a script pipeline. The AWK script (see Appendix Script#2) was used for reshaping initial table (Table 1) that was received directly from GMT as a result of the automatically digitized profiles. This table was then reshaped by AWK into the second table (Table 2) that now had a structured view with columns containing only necessary data: depths (m) and profile length (from -200 to 200 m, according to the stretch). The main task of the reshaping was to select useful information from the initial huge table that contained 12726 rows and 5 columns. Therefore, during this procedure the unnecessary rows were removed, and only necessary columns were selected. The AWK research steps were done consequentially as explained below stepwise.

Initially, the working directory was set up to store all the files (both temporary and main ones) in this folder (cd /Users/pauline/Documents/Octave) which was then controlled by 'pwd' check. Then the input file (table2.csv) was read in into AWK environment in units (records) processed by the embedded syntax rules of the AWK one record at a time line by line [19]. Each record was automatically split into pieces, that is fields of the table making it more convenient for AWK to work on the parts of a table [20].

Step-1. At the first step the series of the temporary files was created where all the preliminary results were stored (TMPFILE0… TMPFILE10) using Unix: mktemp utility. Several kinds of tasks occurred repeatedly when working with table file received from GMT: extracting columns one by one from the initial file and then storing them as separate files. Therefore, certain lines were extracted from the table and the rest was discarded at this step.

Step-2. The initial GMT tables for two segments of the profiles contained 62 and 52 profiles for northern and southern parts respectively. Therefore, at this step only two necessary columns were selected using AWK predefined variable 'FS' (field separator) as space. That means, in both input and output data (output field separator as 'OFS') the items were separated by single spaces: 'BEGIN{FS=" "; OFS=" "}'.

Step-3. Now only profile 1 was selected and saved as a separate table as 'table_10pr' with X and Y (location and depths).

Step-4. Only Y column (field 5) was taken from the initial table. This column contained depths for all profiles, for example for the northern segment these are 12726 rows.

Step-5. The profile I was selected consequentially from Y column. This is done using command printing section of file between two regular expressions (regex). Patterns in awk control the execution of rules when its pattern matches the current input record. The following is an example of the AWK snippet code in this case: awk '/-L0-07/,/-L0-08/' $TMPFILE1 > table_p07.csv. This means that in the case all the data between lines L0-07 and L0-08 were taken and a new mini-table was created. This table now contained information for the depths in all points along the profile nr. 7. Then the same procedure was repeated for all other profiles. The regular expression that matches the text of the table ('/-L0-07/,/-L0-08/') is put between the slash lines, like this: /regular expression/.

Consequently, this pattern was repeated for the northern segment for the profiles of 'i' number: 1, 7, 14, 21, 28, 35, 42, 49, 56, 62. The output data are stored as tables (e.g. table_p07.csv).

Step-6. At this step the single Y column of the profile *i* was added to the table (for example, 'table_10p.csv'). Then, the output was saved to 'table_10Kuril-Kamchatka Trench.csv'. Making these modifications in the table was done using patterns (see AWK script), but leaving the rest of the table file alone. The same procedure was repeated for the southern segment using the same regular expression as a template.

Step-7. Now the header and tail were deleted from the output table using the AWK expression that prints only lines which do not match regex.

Step-8. The auxiliary files were removed by the 'rm'.

Step-9. The number of lines was counted for the initial table. In this case, the initial table (table2.csv) had 12726 lines.

Step-10. Finally, the number of lines was counted for the output table. In this case the final table (table_10Kuril-Kamchatka Trench.csv) had only 201 lines (against previous 12726).

As can be noticed, the presented AWK script consists of expressions that are the basic building blocks of AWK patterns and actions. An expression performed in each step of the script evaluates to a value that was printed and consequentially passed to the next expression at the following step in the script. For the case of the temporary files (Step-5), a series of expressions assigned a new value to the table variables using an assignment operator. AWK expressions extensively used in the script served as a pattern for the selected geomorphological profiles along the whole length of the trench, thus facilitating the repeatability of the research steps.

The final output table was reshaped in view, that is columns became rows. Azimuth and absolute coordinates were deleted from the table. As a result, the output table contained 201 rows and 11 columns and initial long table has been reshaped. The new structure of the table is as follows: column 1 contains cross-section line 400 km long (-200:200 km). Columns 2-11 contains depths for the cross-section profiles. Unix programs and utilities used additionally included: 'mktemp' for creating a series of temporary files and 'rm' for removing preliminary temporary files. The final table was then processed in GNU Octave language for the visualization of the cross-section profiles.

3.4 GNU Octave plotting

The final part of the research contains plotting graphs by presented Script #2 (see Appendix) in GNU Octave programming language to visualize selected graphs of the geomorphology. The plotting was done using available technical documentation [21]. The Figures 4 and 5 show selected graphs of the profiles made on the basis of the table created in the previous step by AWK and further visualized in GNU Octave using provided script of 'm-file' with an example of the northern segment (see Script #2). The southern part was repeated using the same script template using corresponding table and necessary minor amendments (numbers of profiles, depths, corresponding for the southern segment).

Table 1. Initial table derived from the GMT for the Profile # 1

> Cross profile number -L0-00				
Coordinates, lon	Coordinates, lat	Location, m*	Azimuth, grad	Depth, m
149.353224965	41.5059926426	-200.065204481	326.61376193	-5388.91250029
149.34000552	41.5210101811	-198.064782729	326.61376193	-5386.2095364
149.326779938	41.5360262051	-196.064356317	326.605000128	-5378.97817425
149.313548214	41.5510407131	-194.063925248	326.596231663	-5363.02735362
149.300310343	41.5660537036	-192.063489519	326.587456531	-5336.36810313
149.287066319	41.5810651752	-190.063049132	326.578674725	-5300.09852962
149.273816136	41.5960751262	-188.062604087	326.569886241	-5249.45260443

149.260559789	41.6110835552	-186.062154384	326.561091074	-5188.08489127
149.247297273	41.6260904606	-184.061700023	326.552289219	-5116.35098797
149.234028582	41.6410958411	-182.061241004	326.54348067	-5051.39489267
149.22075371	41.6560996949	-180.060777327	326.534665422	-5070.64221477
149.207472652	41.6711020207	-178.060308993	326.525843469	-5345.61405245
149.194185402	41.6861028169	-176.059836002	326.517014808	-5504.76112339

[*] The geometric model of each profile is as follows: the trench valley corresponds to the zero value of the segment, that is the origin of the plot. Left and right parts of the segments stretch to the 200 and -200 m, correspondingly as shown on X axis in Figure 2. The direction of each profile is NW-SE, i.e. from the Sea of Okhotsk and Greater Kuril Chain directed to the Pacific Ocean.

Using Octave, a comparative analysis of the geomorphology in selected profiles crossing the Kuril-Kamchatka Trench was performed visualizing the "V" shape of the Kuril-Kamchatka Trench which can be noticed in a series of the cross-section profiles (Figure 4 and Figure 5). The Octave graphs (Figure 4 and Figure 5) highlighted the geomorphic patterns in the northern and southern parts of the trench visualizing variations in depths and geomorphic shapes. Southern part of the trench proved to have deeper bathymetric ranges and steeper trench gradient slope, which is caused by the intensive seismicity and volcanic activities in this particular area of the trench.

The series of the SW-NE-aligned orthogonal bathymetric profiles (Figure 2) show gradual changes in the geomorphology of the Kuril-Kamchatka Trench along the islands from the Greater Kuril Chain northwards from the Bussol Strait: Simushir, Ketoi, Ushishir, Rasshua, Matua, until it finally approaches the north-eastern end of the Sea of Okhotsk and Bering Sea.

Table 2. AWK-reshaped table containing location and depths (m) for the first 15 rows

location	profile1	profile2	profile3	profile4	profile5	profile6	profile7	profile8	profile9	profile10
-200,07	-5388,91	-5066,99	-5096,23	-5007	-5095,57	-5089,61	-5111,86	-5105,17	-5230,93	-5387,38
-198,06	-5386,21	-5056,08	-5087,92	-4990,77	-5111,1	-5113,8	-5126,26	-5116,11	-5242,93	-5395,11
-196,06	-5378,98	-5056,92	-5076,23	-4991,77	-5127,27	-5143,67	-5136,49	-5127,79	-5263,13	-5400,87
-194,06	-5363,03	-5068,51	-5063,31	-4998,02	-5125	-5174,22	-5137,67	-5145,92	-5286,53	-5405,93
-192,06	-5336,37	-5084,85	-5052,82	-5000,48	-5102,27	-5197,73	-5131,9	-5174	-5300,04	-5412,8
-190,06	-5300,1	-5099,18	-5047,83	-5001,55	-5077,16	-5207,11	-5117,67	-5204,43	-5304,92	-5417,57
-188,06	-5249,45	-5109,15	-5047,52	-5005,99	-5057,1	-5203,18	-5092,88	-5222,58	-5303,37	-5410,98
-186,06	-5188,08	-5114,52	-5049,58	-5008,41	-5041,25	-5189,31	-5067,21	-5232,12	-5295,12	-5394,34
-184,06	-5116,35	-5117,15	-5039,51	-5017,05	-5031,61	-5169,68	-5049,59	-5232,31	-5281,9	-5380,4
-182,06	-5051,39	-5123,59	-5024,97	-5029,04	-5029,31	-5146,61	-5043,71	-5220,18	-5265,19	-5367
-180,06	-5070,64	-5133,68	-5020,5	-5038,68	-5034,43	-5125,47	-5044,13	-5190,62	-5245,44	-5364,2
-178,06	-5345,61	-5147,98	-5028,46	-5045,35	-5040,97	-5109,78	-5043,65	-5147,64	-5224,38	-5350,02
-176,06	-5504,76	-5156,31	-5042,72	-5048,1	-5047,01	-5102	-5033,99	-5108,91	-5201,76	-5320,17
-174,06	-5511,42	-5150,97	-5057,82	-5048,89	-5051,76	-5108,89	-5018,05	-5082,59	-5179,36	-5282,81
-172,06	-5475,67	-5137,74	-5073,48	-5057,95	-5052,26	-5132,15	-4991,71	-5070,96	-5155,37	-5246,77

4 RESULTS

GMT/AWK/Octave sequential data processing, automatic digitizing of the orthogonal profiles and graphical plotting based on various algorithms of the data analysis have been tested as an experimental design of this research, aimed to visualize two morphological patterns along the southern and northern segment flanks of the Kuril-Kamchatka Trench. Comparing the resulting two graphs (Figure 4 and Figure 5), two parts of the trench located north- and southwards from the Bussol Strait show that the southern part is deeper reaching -8,200 m depth, while the northern part has -7,800 at maximal records. It can be explained by the fact that the northern part

of the trench affected by numerous earthquakes located on the Kamchatka Peninsula is influenced by a large amount of the accumulated sediments. Necessarily, it forms the shallower depths and gentle flat slopes in the morphology of the northern part of the Kuril-Kamchatka Trench. The results of the dataset analysis were compared and found out that profiles located in the southern and northern parts vary in depths and geomorphic steepness, due to the impact of the geological settings and seismicity.
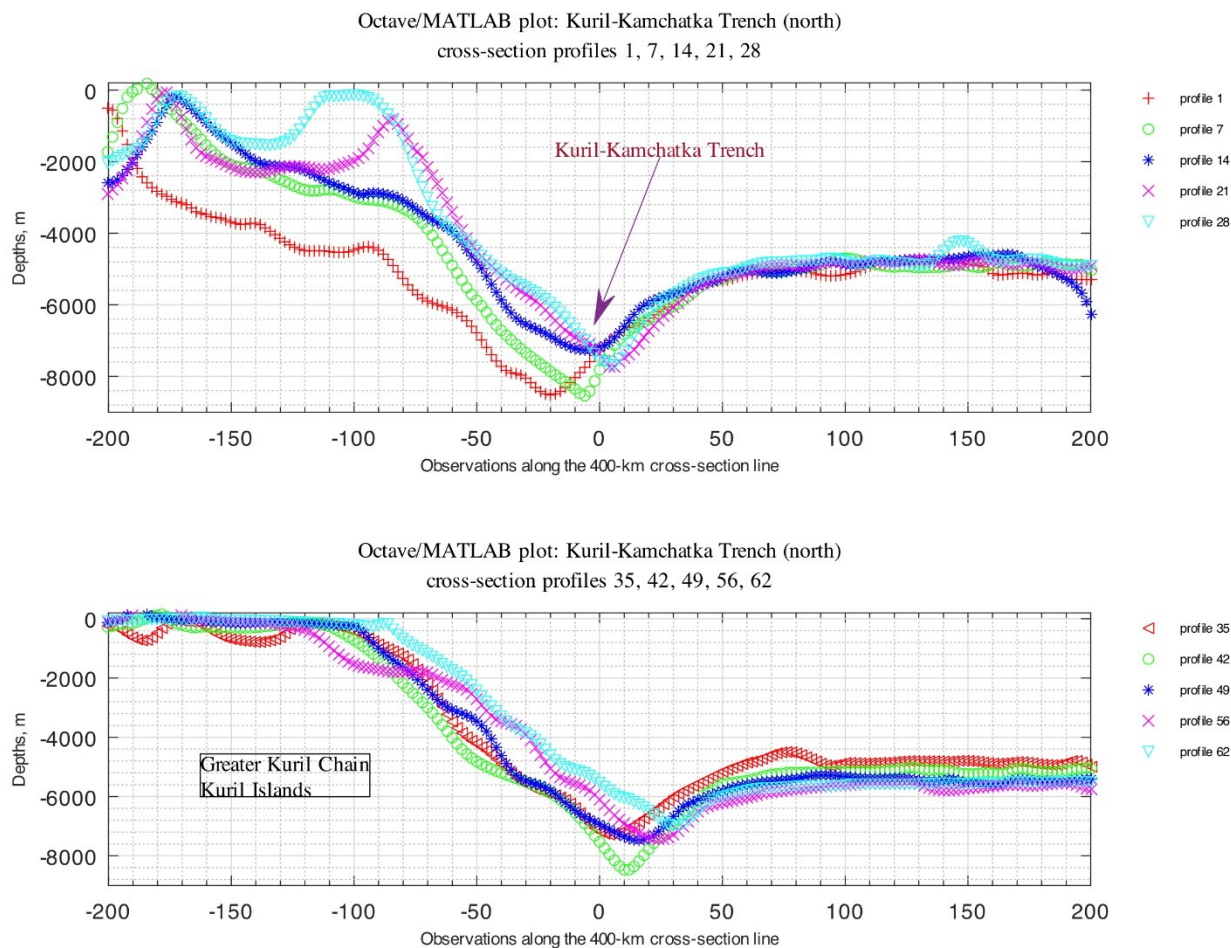

Figure 4. Northern part of the geomorphology of the Kuril-Kamchatka Trench: a graph based on the table reshaped by AWK script

In general, the profiles in the northern part of the trench show deeper values and less steep slopes comparing to the southern segment, which refer to the erosional moats and drift of the submarine sediments. On the contrary, the southern part of the trench has deeper bathymetric ranges and steeper trench gradient slope, comparing to the northern part. This can be influenced by the intensive seismicity and volcanic activities in this particular area of the trench. Furthermore, the trench forms a more V- shaped valley in its southern segment with about 3 km wide in profiles and more U-shaped form in the northern part. The shape of the slopes and steeper morphology of the southern part is affected by the specific geophysical conditions of the southern part of the Kuril Islands characterized by the magnetized anomaly zones of the oceanic crust. The V-shaped valley of the trench segment neighbouring Iturup and Urup islands goes in east-northward direction where it faces the Bussol Strait. Afterwards, the trench continues direction in the northern segment along the islands Shiashkotan, Kharimkotan, Onekotan,

Paramushir and Shumshu, and finally reaches Kamchatka Peninsula. Variations in the southern and northern part of the trench reflect the differences in the local geological settings of the respecting segments.
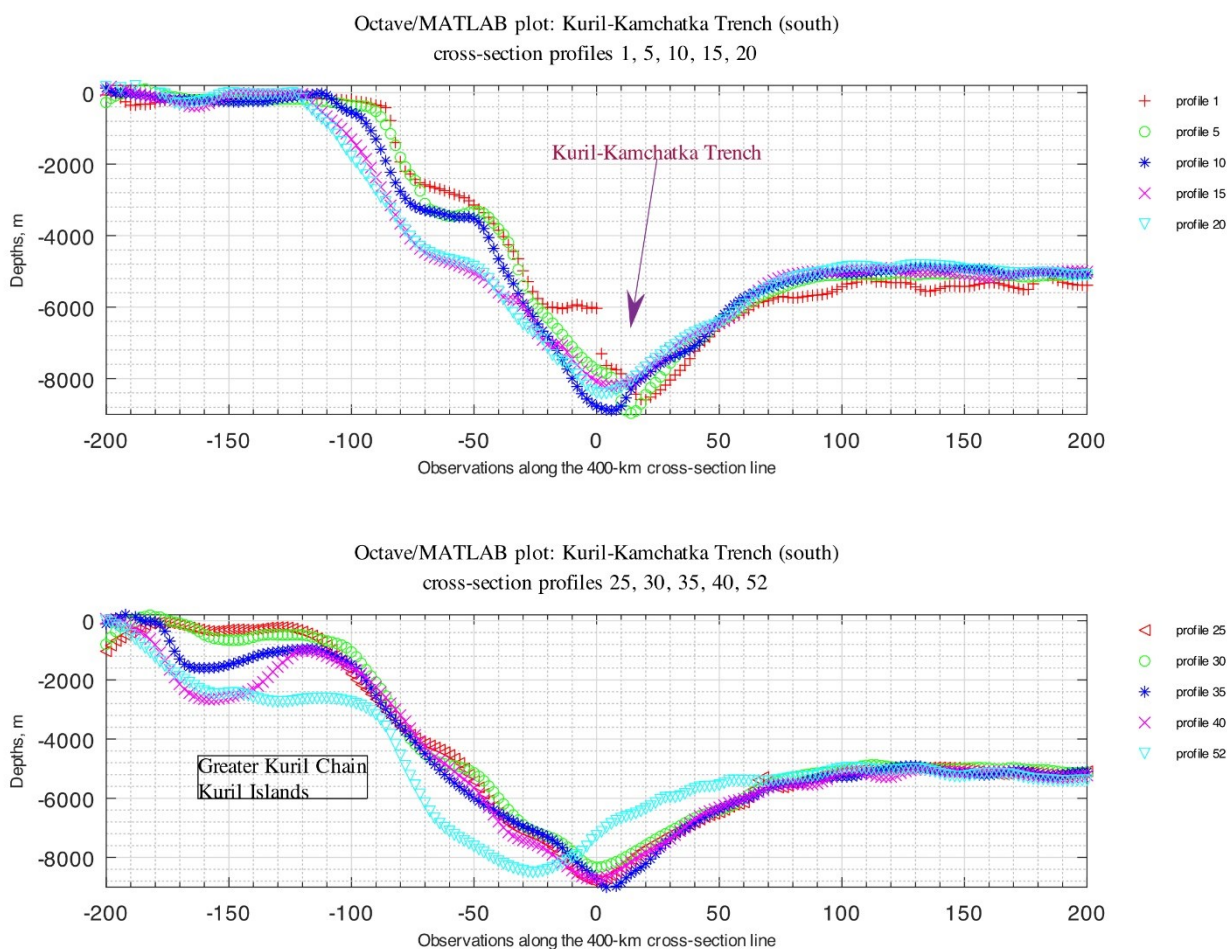

Figure 5. Southern part of the geomorphology of the Kuril-Kamchatka Trench: a graph based on the table reshaped by AWK script

The oceanic trenches are associated with tectonic subduction zones. Therefore, besides the geological factors, the trench geomorphology is necessarily affected by the local tectonic settings. The seafloor spreading creates an axial rift and hills. Spreading ridges are formed by the nearby faults, where subduction of the cooled tectonic plate into the mantle causes the creation of the deep ocean trenches [22]. As a consequence, the major earthquakes occur along the Kuril-Kamchatka trench. Various geophysical factors affect the geomorphological structure of the trench and cause variations in its morphological structure [23], among which tectonics plays the major role in the trench formation. For the case of the Kuril-Kamchatka Trench, two tectonic plate have responses to its tectonic history, which could be caused by the ongoing subduction of the major Pacific Plate under the minor Okhotsk tectonic plate [24]. The Okhotsk plate is a region of the active grabens located southwards of a chain of small sedimentary basins in the Cherskii Mountains, North Russia [25]. In principle, the collision of the tectonic plates could be the one more reason that has affected both the Kuril-Kamchatka trench motion and the morphology of the subducting Pacific plate at depth in the northern and southern parts of the trench that could be sliced in the two distinct morphologies in its northern and southern parts. The distinct changes in the trench geomorphology is illustrated by [26] where they show changes between the slab beneath the horizontal slab and vertical arcs with a case study of Mariana trench area. The variations in physical properties within the slab could be the result of the distortion of the Pacific plate as its shape transforms between near horizontal to near vertical, as well as change in the velocity of the subduction plate, or a combination of all these factors.

The seafloor of the trench presents a background for all geological, oceanological and geochemical processes that are necessarily reflected in its shape. Thus, the main morphostructure elements of the seafloor are the outskirts of the continents, tectonic plates and ocean bedrock. The structure of the trench and the nature of its geomorphic shape are therefore greatly complicated by the multiple secondary tectonic disturbances: system of

faults, displacements of grabens, horsts and lateral geologic shifts [27]. The interconnection of these complex processes is also illustrated by recent findings [28] that demonstrates correlation between the back-arc deformation and upper continental tectonic plate velocity. The trench migration rates are also mainly controlled by the lower continental tectonic plate velocity [29], which in turn depends on the tectonic slab age buoyancy [30]. More detailed study on the tectonic impact goes beyond the scope and size of this manuscript which aims at development of the technical methodology of the batch geodata processing by GMT, AWK and Octave as a chain methodological workflow. As a suggestion for further studies on this phenomena, a comparative analysis of the subduction speed of the tectonic plates in the northern and southern parts of the trench could be performed in a separate research based on the availability of the precise geophysical and seismic data.

5   CONCLUSION

Combination of various data analysis tools, such as GMT scripting toolset, AWK and Octave programming languages for sorting and reshaping data, plotting graphs and mapping enabled to create a geospatial analysis of the selected specific areas of the Pacific Ocean – the Kuril-Kamchatka Trench. Successful application of the programming languages and advanced statistical tools towards geospatial data analysis has been demonstrated in various papers: Python [31], [32], [33], Fortran language [34], [35], [36], R language [37], [38], [39], [40], SPSS IBM Statistics [41], [42], [43], Gretl GNU Regression, Econometrics and Time-series Library for Statistical Analysis [44]. However, only few examples were given so far specifically for AWK language and its application is traditionally limited by the IT domains [45], [46]. In this paper, I contribute to filling the gap in the need for using such effective language as AWK [47] in the geological and Earth sciences and absence of the practical examples and full codes that are always useful besides technical manuals.

Understanding complex interrelation and spatial variation of the geological and geophysical factors affecting the Kuril-Kamchatka Trench is crucial to analysis of its geomorphic patterns. Various factors contribute to the trench geomorphology and shape: motion of the Pacific plate subduction to the adjacent small Okhotsk plate, seismicity and location of the earthquakes. Special geophysical settings, geological parameters and special sedimentation processes induced by active volcanism and seismicity in the region are among the crucial factors affecting trench geomorphology. Geological settings of the Kamchatka region in general are notable for the mountains chains in the adjacent land, recorded active seismicity and magmatism, steep gradients of the mountains Hokkaido and Sakhalin islands, and Kamchatka Peninsula, as well as frequent earthquakes and erodible lithology. Necessarily, these factors contribute towards the active sedimentation processes which in turn affect the geomorphology of the trench. Large amounts of the sediments are being transported to the Okhotsk Sea and then driven in SE direction from the Hokkaido and NW from the Kamchatka Peninsula to the neighbouring Kuril-Kamchatka Trench.

The study illustrated variations in the geomorphic patterns of the Kuril-Kamchatka Trench by means of a sequential usage of the advanced scripting toolset: GMT, AWK and GNU Octave programming languages. The orthogonal cross-sections from the 2D profiles along the two studied parts of the trench were automatically digitized and processed by GMT and initial tables converted to GNU Octave via AWK. Three scripts used in GMT, AWK and GNU Octave are presented in full to enable replicability of this research. The paper has a multi-disciplinary nature combining IT programming instrumentation with geospatial data analysis and cartography in GMT.

Competing interests: The author declares no competing interests.

APPENDIX

Script 1: GMT code for the cartographic mapping and modelling profiles (used for map visualized on the Figure 1)

```sh
#!/bin/sh

# Purpose: Generating and plotting stacked cross-sectioning bathymetric profiles

# Area: along the track of the Kuril-Kamchatka Trench

# GMT modules: grdcut, makecpt, grdimage, psscale, grdcontour, psbasemap, psxy, grdtrack, convert, pstext, logo, psconvert

# Unix prog: cat
```

```
# Step-1. Generate a file
ps=cross2KKT.ps
# Step-2. GMT set up
gmt set FORMAT_GEO_MAP=ddd \
    MAP_FRAME_PEN=dimgray \
    MAP_FRAME_WIDTH=0.1c \
    MAP_TITLE_OFFSET=1.7c \
    MAP_ANNOT_OFFSET=0.2c \
    MAP_TICK_PEN_PRIMARY=thinner,dimgray \
    MAP_GRID_PEN_PRIMARY=thinnest,white \
    MAP_GRID_PEN_SECONDARY=thinnest,white \
    FONT_TITLE=16p,Palatino-Roman,black \
    FONT_ANNOT_PRIMARY=10p,Palatino-Roman,dimgray \
    FONT_LABEL=10p,Palatino-Roman,black \
# Step-3. Overwrite defaults of GMT
gmtdefaults -D > .gmtdefaults
# Step-4. Extract a subset of ETOPO1m for the Kuril-Kamchatka Trench area
grdcut earth_relief_01m.grd -R140/170/40/60 -Gkkt_relief.nc
# Step-5. Make color palette
gmt makecpt -Cgeo -V -T-10000/1000 > myocean.cpt
# Step-6. Make raster image
gmt grdimage kkt_relief.nc -Cmyocean.cpt -R140/170/40/60 -JM6i \
    -P -I+a15+ne0.75 -Xc -K > $ps
# Step-7. Add color legend
gmt psscale -Dg133.8/40+w6.5i/0.15i+v+o0.3/0i+ml \
    -Rkkt_relief.nc -J -Cmyocean.cpt \
        --FONT_LABEL=10p,Helvetica,black \
        --FONT_ANNOT_PRIMARY=8p,Helvetica,black \
    --MAP_ANNOT_OFFSET=0.1c \
        -Baf+l"Color scale legend: depth and height elevations (m)" \
        -I0.2 -By+lm -O -K >> $ps
# Step-8. Add shorelines
gmt grdcontour kkt_relief.nc -R -J -C1000 \
-B+t"Kuril-Kamchatka Trench area and cross-section profiles" \
    -W0.1,black -O -K >> $ps
# Step-9. Add grid
gmt psbasemap -R -J \
```

```
    -Lx13.4c/-1.3c+c50+w500k+l"Mercator projection. Scale (km)"+f \
    -Bxg4f2a4 -Byg4f2a4 \
  --MAP_TITLE_OFFSET=0.3c \
  --MAP_ANNOT_OFFSET=0.2c \
  --MAP_LABEL_OFFSET=0.2c \
  --FONT=10p,Palatino-Roman,black \
  -Tdx1.0c/13.5c+w0.3i+f2+l+o0.15i \
    -UBL/-15p/-35p -O -K >> $ps
```

 # Step-10. Select two points for northern segment along the Kuril-Kamchatka Trench

cat << EOF > trench2.txt

153.5 45.5

158.5 50.0

EOF

 # Step-11. Plot northern trench segment and end points

gmt psxy -Rkkt_relief.nc -J -W2p,red trench2.txt -O -K >> $ps *# my line*

gmt psxy -R -J -Sc0.15i -Gred trench2.txt -O -K >> $ps *# points*

 # Step-12. Generate cross-track profiles 400 km long, spaced 10 km, sampled every 2km

gmt grdtrack trench2.txt -Gkkt_relief.nc -C400k/2k/10k+v -Sm+sstack2.txt > table2.txt

 # Step-13. write stacked profile into table

gmt psxy -R -J -W0.5p table2.txt -O -K >> $ps

 # Step-14. Show upper/lower values encountered as an envelope

gmt convert stack2.txt -o0,5 > env2.txt

gmt convert stack2.txt -o0,6 -I -T >> env2.txt

 # Step-15. Select two points for southern segment along the Kuril-Kamchatka Trench

cat << EOF > trench1.txt

148.0 43.0

153.5 45.5

EOF

 # Step-16. Plot southern trench segment and end points

gmt psxy -Rkkt_relief.nc -J -W2p,green trench1.txt -O -K >> $ps *# my line*

gmt psxy -R -J -Sc0.15i -Ggreen trench1.txt -O -K >> $ps *# points*

 # Step-17. Generate cross-track profiles 400 km long, spaced 10 km, sampled every 2km

gmt grdtrack trench1.txt -Gkkt_relief.nc -C400k/2k/10k -Sm+sstack1.txt > table1.txt

gmt psxy -R -J -W0.5p,black table1.txt -O -K >> $ps

 # Step-18. Show upper/lower values encountered as an envelope

gmt convert stack1.txt -o0,5 > env1.txt

gmt convert stack1.txt -o0,6 -I -T >> env1.txt

 # Step-19. Add annotation texts and labels

gmt pstext -R -J -N -O -K \

-F+f16p,Times-Roman,blue+jLB -Gwhite >> $ps << EOF

148.0 53.0 Sea of Okhotsk

EOF

gmt pstext -R -J -N -O -K \

-F+f12p,Times-Roman,black+jLB+a-75 -Gwhite -Wthinnest >> $ps << EOF

145.0 54.5 Deryugin Basin

EOF

gmt pstext -R -J -N -O -K \

-F+f12p,Times-Roman,black+jLB+a-295 -Gwhite -Wthinnest >> $ps << EOF

158.0 54.0 Kamchatka Peninsula

EOF

gmt pstext -R -J -N -O -K \

-F+f12p,Times-Roman,black+jLB+a-270 -Gwhite -Wthinnest >> $ps << EOF

143.0 49.0 Sakhalin Island

EOF

gmt pstext -R -J -N -O -K \

-F+f12p,Times-Roman,black+jLB -Gwhite -Wthinnest >> $ps << EOF

142.0 43.2 Hokkaido

EOF

gmt pstext -R -J -N -O -K \

-F+f10p,Times-Roman,black+jRB+a-40 -Gwhite@30 -Wthinnest >> $ps << EOF

145.3 44.7 Kunashir

147.0 44.9 Iturup

149.0 46.5 Urup

168.8 55.8 Commandor Basin

EOF

gmt pstext -R -J -N -O -K \

-F+f20p,Times-Roman,white+jLB >> $ps << EOF

160.0 46.0 Pacific Ocean

EOF

gmt pstext -R -J -X0.0c -Y0.0c -N -O -K \

-F+f12p,Times-Roman,blue=0.2p,black+jLB+a-329 -Gwhite >> $ps << EOF

147.4 45.2 G r e a t e r   .

EOF

gmt pstext -R -J -X0.0c -Y0.0c -N -O -K \

-F+f12p,Times-Roman,blue=0.2p,black+jLB+a-307 -Gwhite >> $ps << EOF

151.4 47.0 K u r i l C h a i n

EOF

gmt pstext -R -J -N -O -K \

-F+f14p,Times-Roman,blue+jLB -Gwhite >> $ps << EOF

165.2 58.5 Bering Sea

EOF

gmt pstext -R -J -N -O -K \

-F+f12p,Times-Roman,black+jLB -Gwhite -Wthinnest >> $ps << EOF

154.0 41.0 Northwest Pacific Abyssal Plain

164.1 43.0 Shatsky Ridge

156.2 43.2 Bussol Strait

146.0 55.5 Barite Mts

EOF

gmt pstext -R -J -N -O -K \

-F+jTL+f12p,Times-Roman,black+jLB+a-330 -Gwhite -Wthinnest>> $ps << EOF

145.0 46.5 Kuril Basin

147.7 48.4 Academy of Sciences Rise

EOF

 # Step-20. Plot arrow for Bussol Strait

gmt psxy -R -J -Sv0.15i+bc+ea -Gyellow -W0.5p,yellow -O -K >> $ps << EOF

155.9 43.8 135 1.5c

EOF

 # Step-21. Add subtitle

gmt pstext -R0/10/0/15 -JX10/10 -X0.5c -Y8.1c -N -O -K \

-F+f12p,Palatino-Roman,black+jLB >> $ps << EOF

0.0 14.6 Track segments: northern (red) and southern (green) separated by the Bussol Strait.

0.0 14.0 Cross-sectional profiles: 400 km long, spaced 10 km, sampled every 2 km.

0.0 13.4 Bathymetric map: ETOPO1 Global Relief Model 1 arc min resolution

EOF

 # Step-22. Add GMT logo

gmt logo -Dx6.2/-10.2+o0.1i/0.1i+w2c -O >> $ps

 # Step-23. Convert to image file using GhostScript (portrait orientation, 720 dpi)

gmt psconvert cross2KKT.ps -A0.5c -E720 -Tj -P -Z

Script 2: AWK code for table reshaping (given example is for the northern segment, repeated accordingly for the table of the southern segment)

#!/bin/sh

```
# Step -1.
```

```
TMPFILE0=$(mktemp)
```

```
TMPFILE1=$(mktemp)
```

```
TMPFILE2=$(mktemp)
```

```
TMPFILE3=$(mktemp)
```

```
TMPFILE4=$(mktemp)
```

```
TMPFILE5=$(mktemp)
```

```
TMPFILE6=$(mktemp)
```

```
TMPFILE7=$(mktemp)
```

```
TMPFILE8=$(mktemp)
```

```
TMPFILE9=$(mktemp)
```

```
TMPFILE10=$(mktemp)
```

```
# Step -2.
```

```
awk 'BEGIN{FS=" "; OFS=" "} {print $3, $5}' table2.txt > $TMPFILE0
```

```
# Step-3.
```

```
awk '/profile -L0-00/,/profile -L0-01/' $TMPFILE0 > table_10p.csv
```

```
# Step-4.
```

```
awk '{print $5}' table2.txt > $TMPFILE1
```

```
# Step-5.
```

```
awk '/-L0-07/,/-L0-08/' $TMPFILE1 > table_p07.csv
```

```
awk '/-L0-14/,/-L0-15/' $TMPFILE1 > table_p14.csv
```

```
awk '/-L0-21/,/-L0-22/' $TMPFILE1 > table_p21.csv
```

```
awk '/-L0-28/,/-L0-29/' $TMPFILE1 > table_p28.csv
```

```
awk '/-L0-35/,/-L0-36/' $TMPFILE1 > table_p35.csv
```

```
awk '/-L0-42/,/-L0-43/' $TMPFILE1 > table_p42.csv
```

```
awk '/-L0-49/,/-L0-50/' $TMPFILE1 > table_p49.csv
```

```
awk '/-L0-56/,/-L0-57/' $TMPFILE1 > table_p56.csv
```

```
awk '/-L0-62/,EOF' $TMPFILE1 > table_p62.csv
```

```
# Step-6.
```

```
awk '{print $0}' < table_p07.csv | tee check.csv | paste table_10p.csv - > $TMPFILE2
```

```
awk '{print $0}' < table_p14.csv | paste $TMPFILE2 - > $TMPFILE3
```

```
awk '{print $0}' < table_p21.csv | paste $TMPFILE3 - > $TMPFILE4
```

```
awk '{print $0}' < table_p28.csv | paste $TMPFILE4 - > $TMPFILE5
```

```
awk '{print $0}' < table_p35.csv | paste $TMPFILE5 - > $TMPFILE6
```

```
awk '{print $0}' < table_p42.csv | paste $TMPFILE6 - > $TMPFILE7
```

```
awk '{print $0}' < table_p49.csv | paste $TMPFILE7 - > $TMPFILE8
```

```
awk '{print $0}' < table_p56.csv | paste $TMPFILE8 - > $TMPFILE9
```

awk '{print $0}' < table_p62.csv | paste $TMPFILE9 - > $TMPFILE10

 # Step-7.

awk '!/-L0-/' $TMPFILE10 > table_10Kuril-Kamchatka Trench.csv

 # Step-8.

rm -f table_p07.csv table_p14.csv table_p21.csv table_p28.csv table_p35.csv table_p42.csv table_p49.csv table_p56.csv table_p62.csv check.csv

rm -f $TMPFILE0 $TMPFILE1 $TMPFILE2 $TMPFILE3 $TMPFILE4 $TMPFILE5 $TMPFILE6 $TMPFILE7 $TMPFILE8 $TMPFILE9 $TMPFILE10

 # Step-9.

awk 'END{print NR}' table2.csv

 # Step-10.

awk 'END{print NR}' table_10Kuril-Kamchatka Trench.csv

Script 3: GNU Octave code for plotting profiles

cd /Users/pauline/Documents/Octave

pwd

load table into Octave

P = dlmread ('table_10KKT.csv');

P

assign variables

K = P(:,1); % assign 400-kilometer cross-section line (km) to x axis

D1 = P(:,2); % assign depths to y axis

D2 = P(:,3); % assign depths to y axis

D3 = P(:,4); % assign depths to y axis

D4 = P(:,5); % assign depths to y axis

D5 = P(:,6); % assign depths to y axis

D6 = P(:,7); % assign depths to y axis

D7 = P(:,8); % assign depths to y axis

D8 = P(:,9); % assign depths to y axis

D9 = P(:,10); % assign depths to y axis

D10 = P(:,11); % assign depths to y axis

check up variables in the current session

whos

#

now plot figure

figure;

clf;

```
hold on;

subplot(2,1,1);

plot(K, D1, 'r+', 'markersize', 5 , K, D2, 'go', 'markersize', 5, K, D3, 'b*', 'markersize', 5, K, D4, 'mx', 'markersize',
5, K, D5, 'cv', 'markersize', 5); % plot x against y

    grid on;

    grid minor on;

    axis tight;

    set(gca,'xlim',[-200 200], 'xminortick', 'on');

    title({'Octave/MATLAB plot: Kuril-Kamchatka Trench' ; 'cross-section profiles 1, 7, 14, 21, 28'}, 'FontSize',
10, 'Color', 'black', 'FontWeight', 'normal', 'FontName', 'Times');

    xlabel('Observations along the 400-km cross-section line', 'FontSize', 8);

    ylabel('Depths, m', 'FontSize', 8);

    x = [0.55 0.50];

    y = [0.85 0.68];

    annotation('textarrow', x, y, 'Color', [0.4940 0.1840 0.5560], 'HeadWidth', 5,

            'String', 'Kuril-Kamchatka Trench', 'TextColor', [0.6350 0.0780 0.1840],

            'FontSize', 10, 'FontName', 'Times');

    ylim ([-9000 200]);

    xlim ([-200 200]);

    axis([-200 200 -9000 200]);

    xticks([-200:50:200]);

    legend ({'profile 1', 'profile 7', 'profile 14', 'profile 21', 'profile 28'}, 'FontSize', 6, 'box', 'off', 'location',
'northeastoutside');

subplot(2,1,2);

plot(K, D6, 'r<', 'markersize', 5 , K, D7, 'go', 'markersize', 5, K, D8, 'b*', 'markersize', 5, K, D9, 'mx', 'markersize',
5, K, D10, 'cv', 'markersize', 5); % plot x against y

    grid on;

    grid minor on;

    axis tight;

    set(gca,'xlim',[-200 200], 'xminortick', 'on');

    title({'Octave/MATLAB plot: Kuril-Kamchatka Trench' ; 'cross-section profiles 35, 42, 49, 56, 62'},

        'FontSize', 10, 'Color', 'black', 'FontWeight', 'normal', 'FontName', 'Times');

    xlabel('Observations along the 400-km cross-section line', 'FontSize', 8);

    ylabel('Depths, m', 'FontSize', 8);

    dim = [0.2 0.2 0.3 0.2];

    str = {'Greater Kuril Chain','Kuril Islands'};

    annotation('textbox', dim, 'String', str, 'FitBoxToText', 'on',

            'FontSize', 10, 'FontName', 'Times');
```

```
ylim ([-9000 200]);

xlim ([-200 200]);

axis([-200 200 -9000 200]);

xticks([-200:50:200]);

legend ({'profile 35', 'profile 42', 'profile 49', 'profile 56', 'profile 62'},

     'FontSize', 6, 'box', 'off', 'location', 'northeastoutside');
```

#saveas(gcf,'plot_KKT.jpg');

 print('plot_KKTprofiles','-dpdf', '-r300', '-bestfit')

REFERENCES

[1] AHO, A.V., B.W. KERNIGAN and P.J. WEINBERGER. AWK-A Pattern Scanning and Processing Language. *Software Practice and Experience.* 1979, Vol. 9, No. 4, pp. 267-279. DOI: 10.1002/spe.4380090403

[2] STUTZ, M. *Get started with GAWK: AWK language fundamentals*. IBM Developer Works. [online]. 2006. [cited 2019-09-24]. Available from: https://developer.ibm.com/

[3] AHO, A.V., B.W. KERNIGHAN and P.J. WEINBERGER. *The AWK Programming Language*. New York, NY: Addison-Wesley, 1988. ISBN 0-201-07981-X.

[4] LIEBERMAN, J.E. GRef2End: a GeoRef® to EndNote® bibliography translator written in awk. *Computers & Geosciences.* 1992, Vol. 18, No. 9, pp. 1271-1275. DOI: 10.1016/0098-3004(92)90045-S

[5] ROBBINS, A.D. *GAWK: Effective AWK Programming: A User's Guide for GNU Awk, for the 5.0.0 (or later) version of the GNU implementation of AWK*. Ed. 5.0, Boston, USA, 2019. ISBN 1-882114-28-0

[6] ROBBINS, A.D. *Effective Awk Programming: A User's Guide for Gnu Awk (1.0.3 ed.).* Bloomington, IN: iUniverse, 2000. ISBN 0-595-10034-1.

[7] ROBBINS, A.D. Awk by example, Part 2: Records, loops, and arrays. Common threads. IBM DeveloperWorks. [online]. 2001. [cited 2019-09-24]. Available from: https://developer.ibm.com/tutorials/l-awk2/

[8] ROBBINS, A.D. Awk by example, Part 3: String functions and checkbooks? Common threads. IBM DeveloperWorks [online]. 2001. [cited 2019-09-24]. Available from: https://developer.ibm.com/tutorials/l-awk3/

[9] EATON, J.W., D. BATEMAN, S. HAUBERG and R. WEHBRING. *GNU Octave. A high-level interactive language for numerical computations*. Free Your Numbers. 5th ed. for Octave version 5.1.0, 2019.

[10] WESSEL, P. and W.H.F. SMITH. New version, of the Generic Mapping Tools released, *EOS, Transactions of the American Geophysical Union.* 1995, Vol. 76, pp. 329.

[11] BIRD, P. An updated digital model of plate boundaries. *Geochemistry Geophysics Geodesy*. 2003, Vol. 4, No. 3, pp. 10-27. DOI: 10.1029/2001GC000252

[12] IHO. *Standardization of Undersea Feature Names: Guidelines Proposal Form Terminology*, 4th ed. International Hydrographic Organization and Intergovernmental Oceanographic Commission, Monaco, 2008.

[13] HARRIS, P.T., M. MACMILLAN-LAWLER, J. RUPP and E.K. BAKER Geomorphology of the oceans. *Marine Geology*. 2014, Vol. 352, pp. 4-24. DOI: 10.1016/j.margeo.2014.01.011

[14] SAVOSTIN, L.A., L. ZONENSHAIN and B.V. BARANOV. Geology and plate tectonics of the Sea of Okhotsk, Geodynamics of the Western Pacific-Indonesian Region, *Geodynamic Series.* 1983, Vol. 11, pp. 189-221, AGU, Washington, D.C.

[15] BARR, I.D. and M. SPAGNOLO. Understanding controls on cirque floor altitudes: Insights from Kamchatka. *Geomorphology.* 2015, Vol. 248, pp. 1-13. DOI: 10.1016/j.geomorph.2015.07.004

[16] KHOMICH, V.G., N.G. BORISKINA and S.A. KASATKIN. Geology, magmatism, metallogeny, and geodynamics of the South Kuril Islands. *Ore Geological Revue.* 2019, Vol. 105, pp. 151-162.

[17] WESSEL, P., W. H. F. SMITH, R. SCHARROO, J. F. LUIS and F. WOBBE. Generic mapping tools: Improved version released, *EOS Transactions of the American Geophysical Union.* 2013, Vol. 94 No. 45, pp. 409-410, DOI:10.1002/2013EO450001

[18] WESSEL, P. and W.H.F. SMITH. A Global Self-consistent, Hierarchical, High-resolution Shoreline Database, *Journal of Geophysical Research.* 1996, Vol. 101, pp. 8741-8743.

[19] DOUGHERTY, D. and A.D. ROBBINS. *sed & awk* (2nd ed.). Sebastopol, CA: O'Reilly Media, 1997. ISBN 1-56592-225-5.

[20] ROBBINS, A.D. Awk by example, Part 1: An intro to the great language with the strange name. Common threads. IBM DeveloperWorks. [online]. 2000. [cited 2019-09-24]. Available from: https://developer.ibm.com/tutorials/l-awk1/

[21] LONG, P.J.G. *Introduction to Octave*. Department of Engineering University of Cambridge. Based on: P. Smith *Tutorial Guide to Matlab*, September 2005.

[22] LEMENKOVA, P. Factor Analysis by R Programming to Assess Variability Among Environmental Determinants of the Mariana Trench, *Turkish Journal of Maritime and Marine Sciences.* 2018b, Vol. 4, pp. 146-155, DOI: 10.6084/m9.figshare.7358207

[23] LEMENKOVA, P. R scripting libraries for comparative analysis of the correlation methods to identify factors affecting Mariana Trench formation, *Journal of Marine Technology and Environment.* 2018a, Vol. 2, pp. 35-42, DOI: 10.6084/m9.figshare.7434167

[24] BIRD, P. An updated digital model of plate boundaries. *Geochemistry Geophysics Geosystems*. 2003, Vol. 4 No. 3, pp. 10-27, DOI:10.1029/2001GC000252

[25] SAVOSTIN, L.A., L. ZONENSHAIN and B.V. BARANOV. Geology and plate tectonics of the Sea of Okhotsk, Geodynamics of the Western Pacific-Indonesian Region, *Geodynamics Series.* 1983, Vol. 11, pp. 189-221, AGU, Washington, D. C.

[26] MILLER, M.S., A. GORBATOV and B.L.N. KENNETT. Heterogeneity within the subducting Pacific slab beneath the Izu–Bonin–Mariana arc: Evidence from tomography using 3D ray tracing inversion techniques. *Earth and Planetary Science Letter*s. 2005, Vol. 235, pp. 331-342. DOI:10.1016/j.epsl.2005.04.007

[27] LEMENKOVA, P. Scatterplot Matrices of the Geomorphic Structure of the Mariana Trench at Four Tectonic Plates (Pacific, Philippine, Mariana and Caroline): a Geostatistical Analysis by R. In DEGTYAREV, K. E. *Problems of Tectonics of Continents and Oceans. Proceedings of the 51st Tectonics Meeting. 2019, Vol. 1.* Moscow: GEOS, pp. 347-352. ISBN: 978-5-89118-. DOI:10.6084/m9.figshare.7699787.v1

[28] HEURET, A. and S. LALLEMAND. Plate motions, slab dynamics and back-arc deformation. *Physics of the Earth and Planetary Interiors*. 2005, Vol. 149. pp. 31-51.

[29] LALLEMAND S., A. HEURET, C. FACCENNA and F. FUNICIELLO. Subduction dynamics as revealed by trench migration. *Tectonics*. Vol. 27, TC3014, pp.1-15. DOI: 0.1029/2007TC002212

[30] CHASE, C.G. Extension behind island arcs and motions relative to hot spots. *Journal of Geophysical Research*. 1978. Vol. 83. pp. 5385-5388.

[31] LEMENKOVA, P. Processing oceanographic data by Python libraries NumPy, SciPy and Pandas, *Aquatic Research.* 2019b, Vol. 2, pp. 73-91, DOI: 10.3153/AR19009

[32] YU, Q.-Y., L. BAGAS, P.-H. YANG and D. ZHANG. GeoPyTool: A cross-platform software solution for common geological calculations and plots. *Geoscience Frontiers.* 2019, Vol. 10, No. 4, pp. 1437-1447. DOI: 10.1016/j.gsf.2018.08.001

[33] LEMENKOVA, P. Testing Linear Regressions by StatsModel Library of Python for Oceanological Data Interpretation, *Aquatic Sciences and Engineering.* 2019, Vol. 34, pp. 51-60, DOI: 10.26650/ASE2019547010

[34] AGARWAL, B.N.P. and S. SRIVASTAVA. FORTRAN codes to implement enhanced local wave number technique to determine the depth and location and shape of the causative source using magnetic anomaly. *Computers & Geosciences.* 2008, Vol. 34, No. 12, pp. 1843-1849. DOI: 10.1016/j.cageo.2007.10.012

[35] PRIETO, G.A., R.L. PARKER and F.L. VERNON III. A Fortran 90 library for multitaper spectrum analysis. *Computers & Geosciences.* 2009, Vol. 35, No. 8, pp. 1701-1710. DOI: 10.1016/j.cageo.2008.06.007

[36] CHIANG, G.-T., T.O.H. WHITE, M.T. DOVE, C.I. BOVOLO and J. EWEN. Geo-visualization Fortran library. *Computers & Geosciences.* 2011, Vol. 37, No. 1, pp. 65-74. DOI: 10.1016/j.cageo.2010.04.012

[37] ROBERTS, N.M., B. TIKOFF, J.R. DAVIS and T. STETSON-LEE. The utility of statistical analysis in structural geology. *Journal of Structural Geology.* 2019, Vol. 125, pp. 64-73.

[38] LEMENKOVA, P. Hierarchical Cluster Analysis by R language for Pattern Recognition in the Bathymetric Data Frame: a Case Study of the Mariana Trench, Pacific Ocean. In KRASNYANSKY, M. N. *Virtual Simulation, Prototyping and Industrial Design. Proceedings of the 5th Int'l Sci.-Pract. Conference*, 2018, Vol. 2 (5), Tambov: TSTU Press, pp. 147-152. DOI: 10.6084/m9.figshare.7531550

[39] LEMENKOVA, P. An Empirical Study of R Applications for Data Analysis in Marine Geology. *Marine Science and Technology Bulletin.* 2019, Vol. 8, No. 1, pp. 1-9. ISSN: 2147-9666. DOI: 10.33714/masteb.486678.

[40] LEMENKOVA, P. Statistical Analysis of the Mariana Trench Geomorphology Using R Programming Language. *Geodesy and Cartography.* 2019, Vol. 45 No. 2, pp. 57-84. ISSN: 2029-6991. DOI: 10.3846/gac.2019.3785

[41] PAN, A. Study on mobility-disadvantage group' risk perception and coping behaviors of abrupt geological hazards in coastal rural area of China. *Environmental Research.* 2016, Vol. 148, pp. 574-581. DOI: 10.1016/j.envres.2016.04.016

[42] LEMENKOVA, P. Numerical Data Modelling and Classification in Marine Geology by the SPSS Statistics. *International Journal of Engineering Technologies.* 2019, Vol. 5 No. 2, pp. 90–99. DOI: 10.6084/m9.figshare.8796941

[43] ABDELMALIK, K.W. Role of statistical remote sensing for Inland water quality parameters prediction. The *Egyptian Journal of Remote Sensing and Space Science*. 2018, Vol. 21, (2), pp. 193-200. DOI: 10.1016/j.ejrs.2016.12.002

[44] LEMENKOVA, P. Regression Models by Gretl and R Statistical Packages for Data Analysis in Marine Geology. *International Journal of Environmental Trends*. 2019, Vol. 3(1), pp. 39–59, DOI: 10.6084/m9.figshare.8313362.v1

[45] PRASAD, K.V.K.K. and T.S. LAMBA. Natural language interface based on keyword extraction using AWK. *Microprocessors and Microsystems.* 1987, Vol. 11, No. 3, pp. 157-160. DOI: 10.1016/0141-9331(87)90123-2

[46] VIGNERESSE, J.L. Which language should be used to sort multifield records? *Computers & Geosciences.* 1995, Vol. 21, No 10, pp. 1131-1137. DOI: 10.1016/0098-3004(95)00025-9

[47] LEVY, L.S. AWK – A Prototyping Language. In LEVY, L.S. *Taming the Tiger*. New York: Springer. Springer Books on Professional Computing, 1987. DOI: 10.1007/978-1-4612-4718-0_6