



HAL
open science

Computing and Plotting Correlograms by Python and R Libraries for Correlation Analysis of the Environmental Data in Marine Geomorphology

Polina Lemenkova

► **To cite this version:**

Polina Lemenkova. Computing and Plotting Correlograms by Python and R Libraries for Correlation Analysis of the Environmental Data in Marine Geomorphology. Jeomorfolojik Arařtırmalar Dergisi / Journal of Geomorphological Researches, 2019, 3, pp.1 - 16. 10.6084/m9.figshare.10012808 . hal-02327797

HAL Id: hal-02327797

<https://hal.science/hal-02327797>

Submitted on 25 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Jeomorfolojik Arařtırmalar Dergisi

Journal of Geomorphological Researches

© Jeomorfoloji Derneęi

www.dergipark.gov.tr/jader

E - ISSN: 2667 - 4238



Arařtırma Makalesi / Research Article

COMPUTING AND PLOTTING CORRELOGRAMS BY PYTHON AND R LIBRARIES FOR CORRELATION ANALYSIS OF THE ENVIRONMENTAL DATA IN MARINE GEOMORPHOLOGY

Polina LEMENKOVA

Ocean University of China, Department of Marine Geo-sciences, Qingdao, Shandong Province, People's Republic of China

pauline.lemenkova@gmail.com  <https://orcid.org/0000-0002-5759-1089>

Makale Tarihçesi

Geliř 19 Nisan 2019

Düzenleme 13 Mayıs 2019

Kabul 15 Mayıs 2019

Article History

Received April 19, 2019

Received in revised form May 13, 2019

Accepted May 15, 2019

Anahtar Kelimeler

Denizaltı Jeomorfolojisi, Mariana Çukuru, Pasifik Okyanusu, Denizel Jeoloji, Python, R Programlama Dili, Mekânsal Analiz

Keywords

Submarine Geomorphology, Mariana Trench, Pacific Ocean, Marine Geology, Python, R Programming Language, Geospatial Analysis

Atıf Bilgisi / Citation Info

Lemenkova, P. (2019) Computing and Plotting Correlograms by Python and R Libraries for Correlation Analysis of the Environmental Data in Marine Geomorphology, *Jeomorfolojik Arařtırmalar Dergisi / Journal of Geomorphological Researches*, 2019 (3): 1-16

ABSTRACT

The geomorphology of the Mariana Trench, the deepest ocean trench on the Earth, has a complex character: its transverse profile is asymmetric, the slopes are higher on the side of the Mariana island arc. The shape of the Mariana Trench is a strongly elongated, arched in plan and lesser rectilinear depression. The slopes of the trench are dissected by deep underwater canyons with various narrow steps on the slopes of various shapes and sizes, caused by active tectonic and sedimentation processes. Understanding of factors that may affect the shape of the geomorphology of such complex structure requires advanced methods of numerical computing. Current research is focused on the analysis of the geomorphology of the Mariana Trench by application of statistical libraries embedded in Python and R programming languages for the data analysis. Workflow algorithms include processing a data set by analysis, computing and visual plotting of the graphs. The research aims is to understand the environmental interactions affecting submarine geomorphology of the Mariana Trench by statistical data analysis. Technically, used algorithms included libraries of Python (Seaborn, Matplotlib, Pandas, SciPy and NumPy) and libraries of R ({hexbin}, {ggally}, {ggplot2}). Technically, following types of the statistical analysis were tested for computing and plotting: correlograms, histograms, strip plots, ridgeline plots and hexagonal diagrams for the bathymetric and geomorphic analysis. Python, being a high-level language, shown more straightforward approach for the statistical data analysis, while R implies more power in the data visualization. The results of the geospatial data modelling show detected correlation between various factors (geology, bathymetry, tectonics) affecting submarine geomorphology that reveal unevenness in its structure. Both programming languages demonstrated significant functionality for the spatial data analysis. The effective and accurate geospatial data visualization demonstrated by Python and R proves high potential of their application in the geomorphological studies.

© 2019 Jeomorfoloji Derneęi. Tüm hakları saklıdır. All rights reserved.

INTRODUCTION

The study area is focused on the Mariana Trench, the deepest ocean hadal trench on the Earth is located in the west Pacific Ocean. Its geometric shape presents a

strongly elongated, arched in plan and lesser rectilinear depression approximating a form of the arc with a modeled center located at 140°E 18°N (Figure 1).

Geometric shape and location of the Mariana Trench. Bathymetry: 1 arc min ETOPO1 Global Relief Model

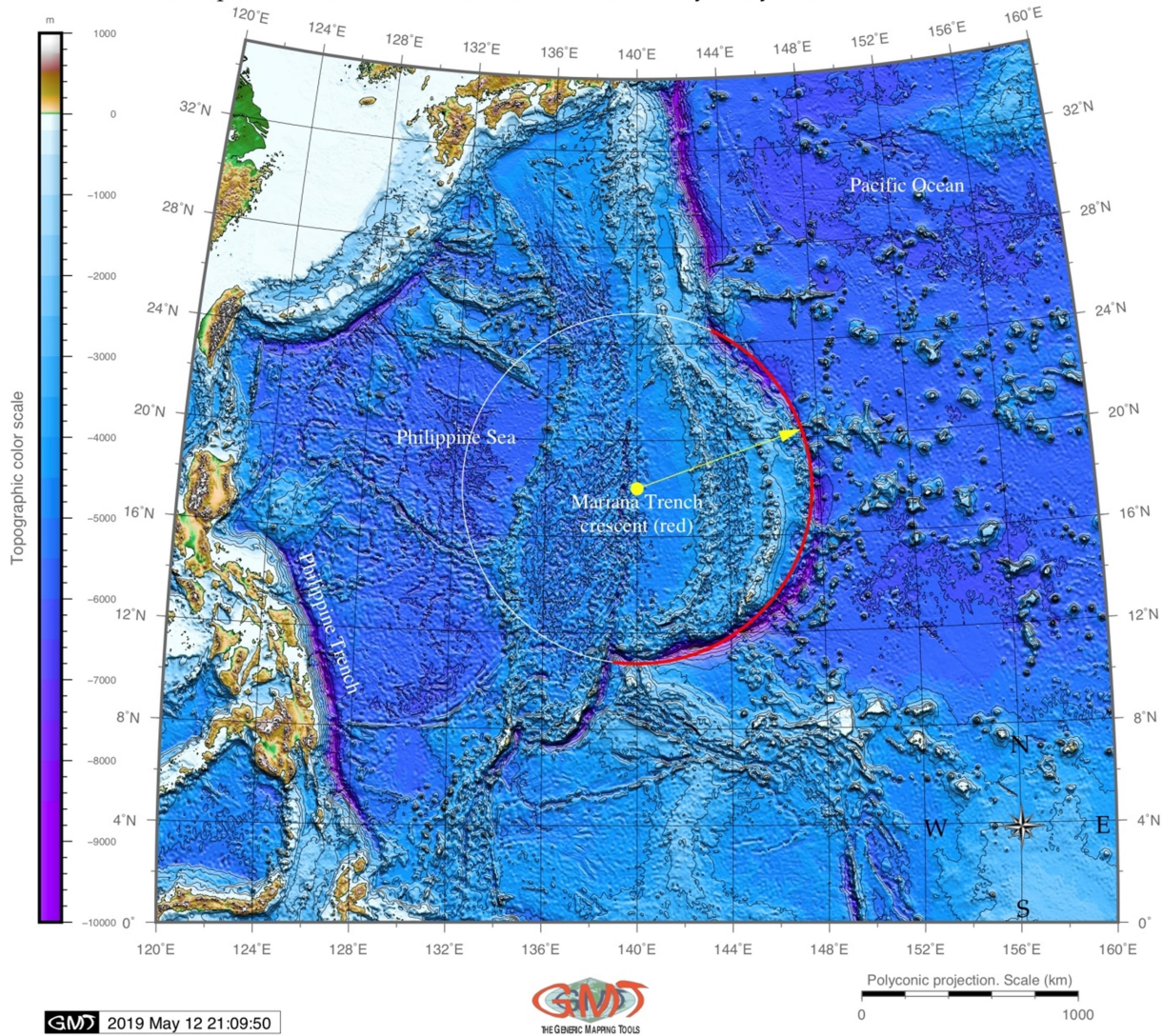


Figure 1. Geographic location and geometric approximated arc form of the Mariana Trench

The transverse profile of the Mariana Trench is asymmetric: the slope is higher on the side of the Mariana island arc. The slopes of the trench are dissected by deep underwater canyons. Various narrow steps are also often found on the slopes of the trench. On the longitudinal profiles of the bottom of the Mariana there are many deep-water detected transverse steps and thresholds connected by the movements of the earth's crust along transverse faults (Zhou et al., 2015). These thresholds can play important role in the evolution of the deep-water trough, for example, to control the intensity of its filling by sedimental precipitation. In this process very important are the streams of the

precipitation moving along the bottom from the part of the trough which is closer to the continent (Taira et al., 2004). At the bottom of the trough there is a transverse threshold, which can present kind of a dam, in front of which there is accumulation of a thick layer of sediment precipitation.

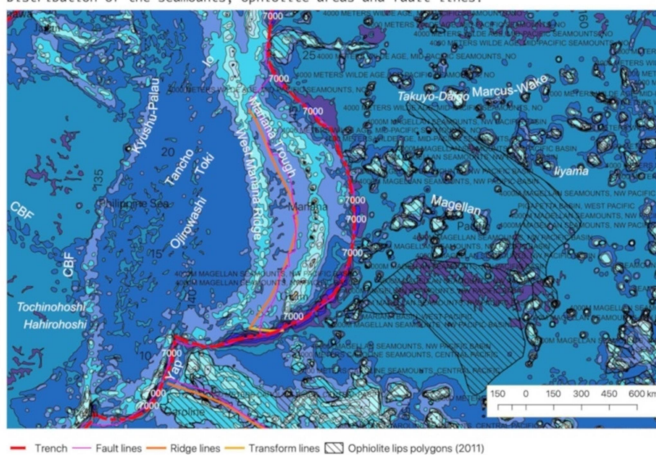
Bathymetry of the Mariana Trench has a complex character and includes a system of cracks. This includes a complicated set of steps of various shapes and sizes, caused by active tectonic and sedimental processes. The steepness of its geomorphic depth averages in 4-5 degrees, but its individual parts can be limited to steeper

slopes as subjects to the gravitational flow system of the submarine canyons and valleys (Kawabe, 1993). Complex distribution of the various geomorphic materials on the adjacent abyssal plains of the ocean contributes to the formation of the geomorphic features of the particular region of the Mariana Trench.

A special focus in this study was laid on the analysis of the geomorphology of the Mariana Trench including the steepness angle of the slopes of the bathymetric cross section profiles. The seafloor bathymetry of the Mariana Trench varies significantly seaward of the trench axis. The Caroline Ridge is located near and off the trench axis at distance of about 0-250km. A prominent trench-parallel belt of seamounts with trench-perpendicular width of about 250 km, is located not far than 1350km (Figure 2, A). Another

prominent group of wide seamounts, intersects the trench at distance of 1600-2000km (Zhang et al 2014). The effects of the volcanism on the biological communities were recently demonstrated by Kitahashi et al (2014), who proved the relationship earthquake events and the assemblage structures. Because rapid sedimentation was induced by the turbidity currents, the organic matter content increased in the topmost layers which indicate that disturbances to deep-sea sediments affects vertical distribution structure of meiofauna in the trenches. The complex character of the geomorphology of the Mariana Trench deserves constant attention of the scientific world (Chadwick et al, 2018; Curtis & Moyer, 2005; Faccenna et al., 2018; Fujioka et al, 2002; Theberge, 2008; Tian et al 2018).

A. Map: Submarine geomorphology of the Mariana Trench, Pacific Ocean. Distribution of the seamounts, ophiolite areas and fault lines.



B. Location of the four tectonic plates (Pacific, Philippine Sea, Mariana and Caroline) around Mariana Trench study area

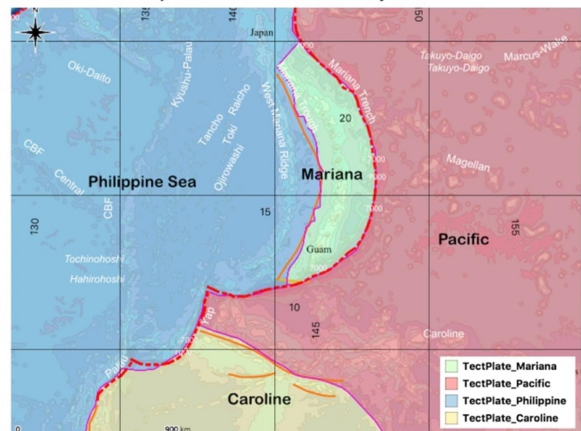


Figure 2. Geologic settings of the Mariana Trench: seamounts, faults, ophiolites (A); tectonic plates (B)

The structure of the submarine geomorphology of the hadal trenches is a complex interplay of various processes. Multiple factors affect the geometry of the hadal trench: plate tectonic movements, sediment thickness, ocean streams determine the local geomorphological conditions that strongly depend on the sediment deposition processes. Besides, the submarine geologic erosion occurring within the main channel of the trench and

the adjusting slope areas, plays an important role in sculpting the geomorphology of the deep hadal trenches. Due to their physical inaccessibility, the ocean hadal trenches are very distinct from other objects on the Earth. Two methods available to study such unreachable areas include first, remote sensing echo-sounding approach, and second, a computer-based modeling. Since the first method is significantly more

resource- and cost-intensive, this paper presents the second one, that is, computer assisted data analysis and modeling.

Strong topographical isolation and complex mix of different water masses distinct hadal trenches from other ocean regions. The difficulties to understand the deep-sea hadal life and habitats hadal zone (6,000–11,000m) that represents the last great frontier in marine science are caused by the remote location of such study object as ocean trenches. Given the above-mentioned difficulties in studies of marine geomorphology, the advanced methods of the statistical analysis were applied for computing correlation matrices of factors affecting the geomorphic structure of the Mariana Trench using methods discussed below.

The novelty of this research consists in the multi-disciplinary approach of this work that combines methods of the statistical analysis and computing with mapping and spatial analysis of the geomorphology and marine geology of the study area. The main aim of the work is to statistically visualize and compute various factors that influence submarine geomorphic shapes causing variability, environmental triggers and interplay between them: geology, bathymetry, tectonics, oceanography.

MATERIAL AND METHOD

The work methodology aims at understanding the geological factors affecting the geomorphology of the Mariana Trench by means of the numerical computing and plotting correlograms. The statistical analysis was based on Python and R libraries using existing algorithms and methods of the machine-based data analysis (Cielen & Meysman, 2016; Halterman, 2011; Harrington, 2015; R Development Core Team, 2014).

The methodology of the research in both hadal and riverine geomorphology is traditionally being supported by the GIS software, mainly in the geological modeling and investigations. Thus, the geomorphic studies of the hydrological effects on the sites location often use GIS modeling and digital terrain analysis for geospatial analysis (Berger, 2011). Various examples of the geomorphological analysis of the landforms by spatial data analysis using GIS include application of satellite imagery, topographic and land cover maps using ESRI ArcGIS, Google EarthView plugin for QGIS and algorithms of the remote sensing data processing (Bertoldi et al, 2015; Arrospide et al, 2018; Bechter et al, 2018; Marchese et al, 2017; Hudson et al, 2008).

The visual schematically mind map summarizing the approaches, tools and methods of the work is given on Figure 3.

The initial map of study area (Figure 1) was done using Generic Mapping Tools using available documentation (Wessel et al., 2017; Wessel & Smith 1995). Digitizing profiles and retrieving information about the geomorphic structure of the Mariana Trench was done using QuantumGIS. Further geostatistical approaches and methods were performed by R and Python programming languages. The initial geodata were extracted from the QGIS project as cross-section bathymetric profiles drawn across the trench with a length of 1000 km (Figure 4). The profiles cross four tectonic plates: Mariana, Caroline, Pacific and Philippine Sea (Figure 2, B).

The initial statistical data analysis in data distribution was visualized by ridge line plots (Figure 5) showing distribution of depths by profiles and observation points by tectonic plates. The attribute data set was derived from the QGIS vector layers:

tectonics, bathymetry, geomorphology and geology. These include cross-section profiles having data from various vector GIS layers: bathymetry, coordinates, bathymetry, geology, tectonics, ophiolites, seamounts, igneous volcanic zones: Figure 2 and Figure 6.

The information was derived in the QGIS project from the attribute tables of the points where profiles crossed the layers. Then the table was stored in .csv format for further statistical processing by R and Python libraries.



Figure 3. Mind map of the project: methods, concepts, approaches and tools. Plotting: LaTeX.

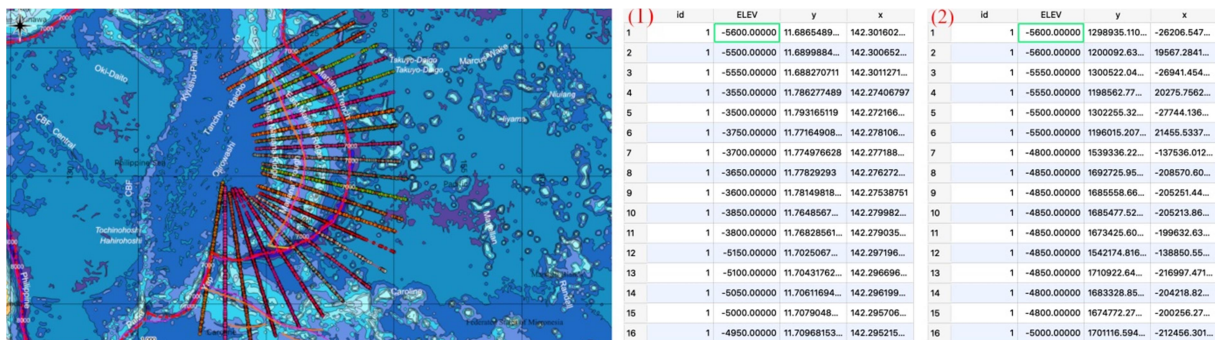


Figure 4. Location of the cross-sectioning track profiles.

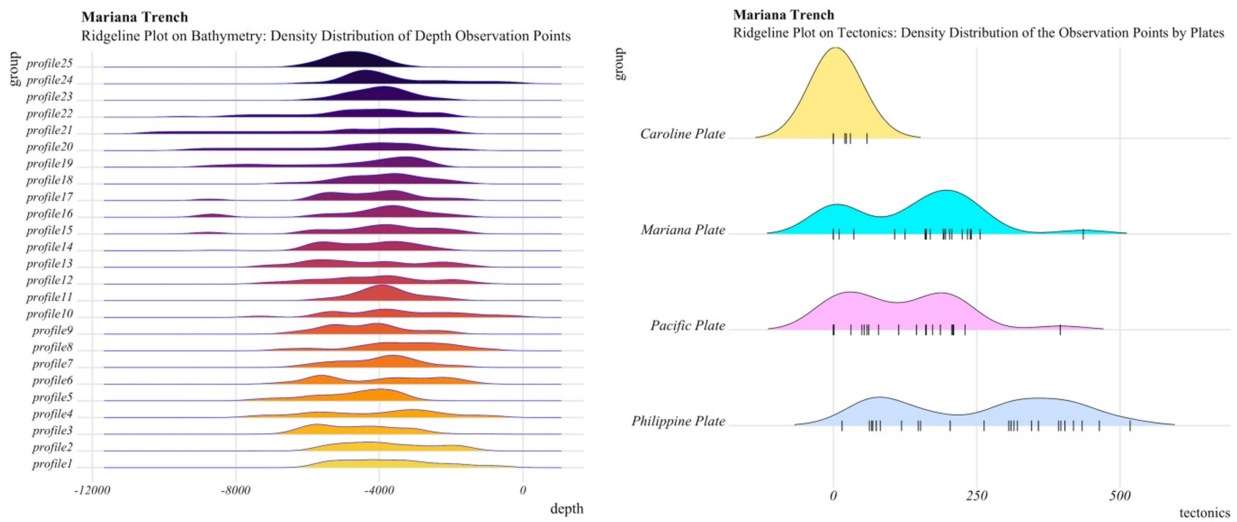


Figure 5. Visualization of the data distribution by bathymetric profiles (left); data distribution by geology and tectonic plates (right) by {ggridges} library.

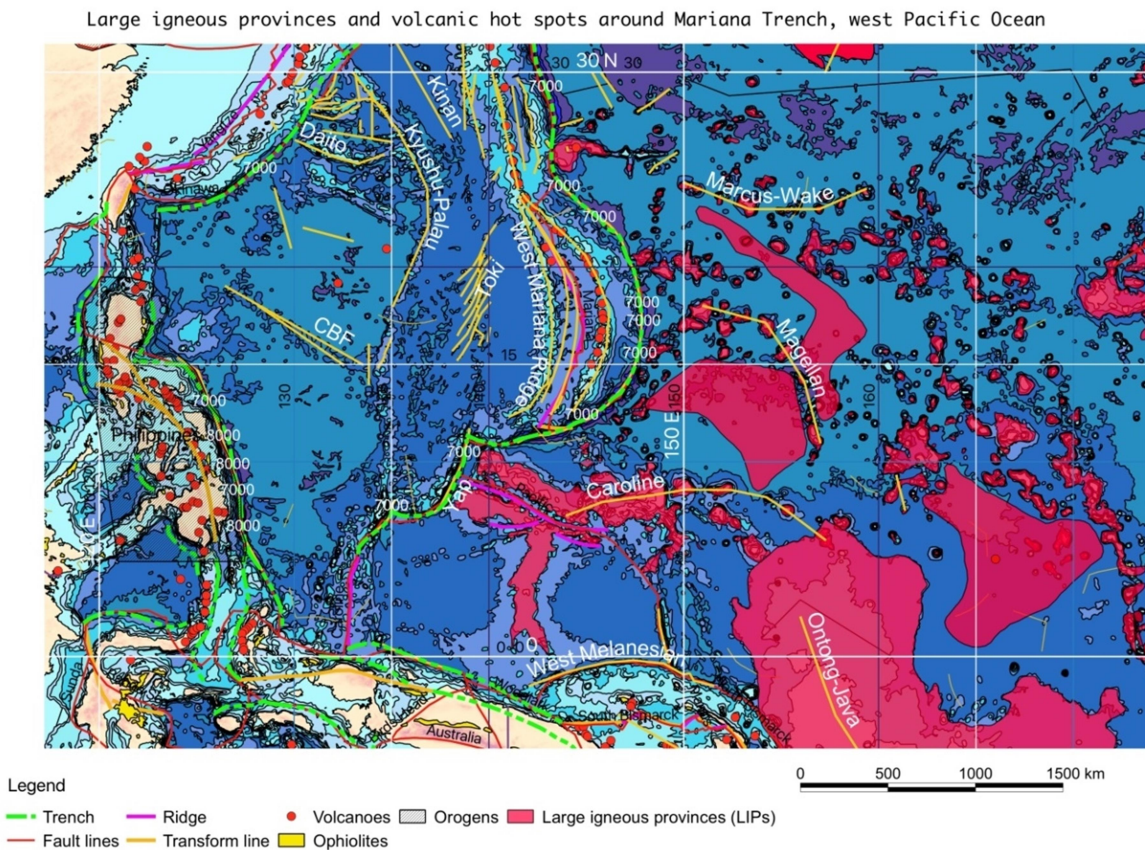


Figure 6. Map of location of the large igneous provinces and volcanoes in the study area.

2.1. Hexagonal plots for modelling bathymetric data range

Plotting hexagonal plot was performed using R language using specially designed library {hexbin}. This plot aims to show how the bathymetric data are distributed by the profiles in terms of frequency. Specifically, it shows how often the values

are being repeated by profiles. Hexagonal, or 'hexbin', plots (Figure 7) serve as a useful alternative to scatter plots, because the bathymetric data in the current data set are too dense (518 observation points by 25 profiles make 12,950 points all together). Therefore, plotting each point individually in this case is not effective: the

data would be indistinguishable on the graph. The R {hexbin} library demonstrates generated hexagonal binning plot of the range of depth's values in Mariana Trench by 25 profiles.

The explanation of the Figure 7 is as follows. The X axis shows profiles from 1 to 25. The Y axis shows recorded depth values (in meters) in observation points by profiles. The colors show the frequency of the depth distribution by observation points in profiles. Thus, there are 3 gradations of the data distribution by repeatability: "hundreds" (very often recorded values, i.e. there are hundreds of points with such depths), "tens" (moderate frequency), and "ones" (very rare values). The frequency is colored in magenta, cyan and blue, respectively. The frequency of the data repetition across the trench, can be read as such: for example, "Ones" (colored blue) are the rarest, i.e. the least frequent recorded depth values. One can see that the profiles 21, 22 and 23 have the deepest values are the rarest in the trench. The cyan-colored "Tens" values

(that is, moderately often recorded depths) correspond to the major depth range (-3,000 to -5,000 meters). Finally, "Hundreds" (colored magenta) are the most frequent depth values that are the most often repeated depths. The thickness of the line of each hexagonal bin (1, 3, 5, 7 and 9) by each of the 3 categories shows the frequency within the group (Figure 7).

2.2. Histograms of the bathymetric data distribution by tectonic plates

Since Mariana Trench crosses 4 tectonic plates – Mariana, Pacific, Philippine and Caroline (Figure 2, B), the data set contained environmental variables separately for these plates that should be analyzed. This required merging the columns into one category: 'tectonic plates'. Therefore, the columns with the plate names were melted into one class using the Python code. Now the analysis of the bathymetric data divided and classified by four tectonic plates (Pacific, Philippine Sea, Caroline and Mariana) became possible. The data was then read in again to the Python and processed as such.

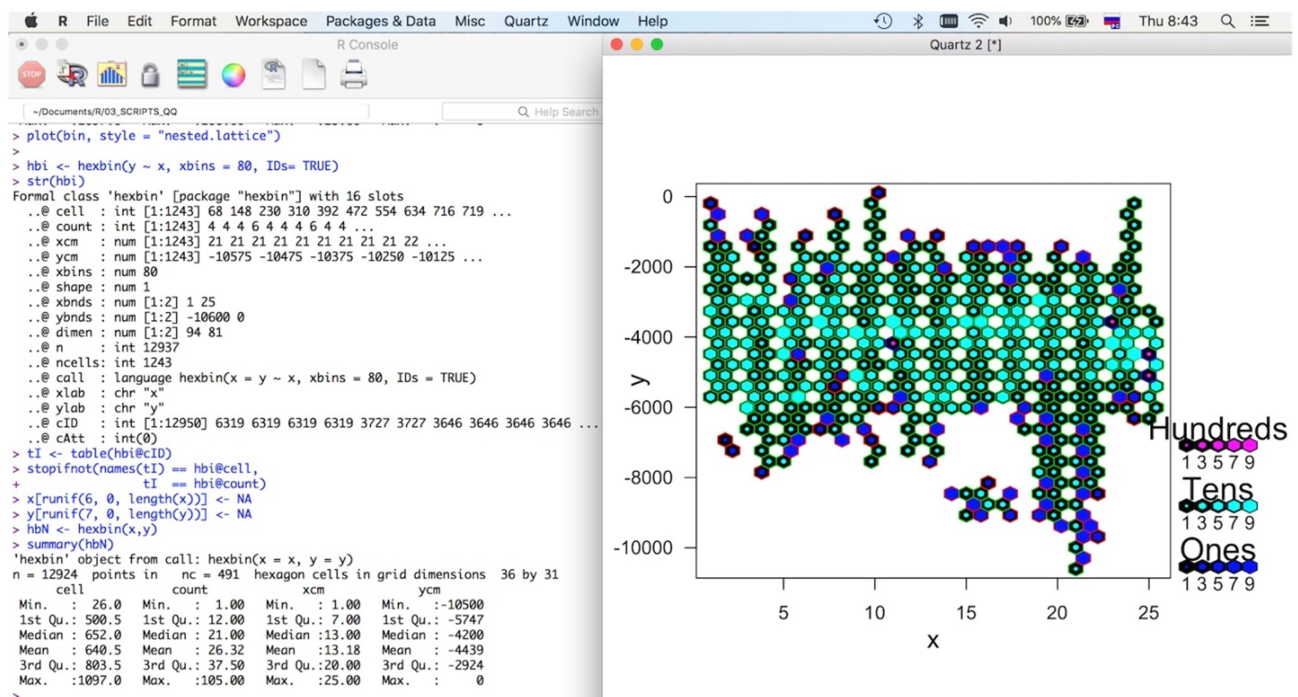


Figure 7. Hexagonal plots showing distribution of depths by bathymetric cross-sectioning profiles. The values density is colored (turquoise, blue, magenta).

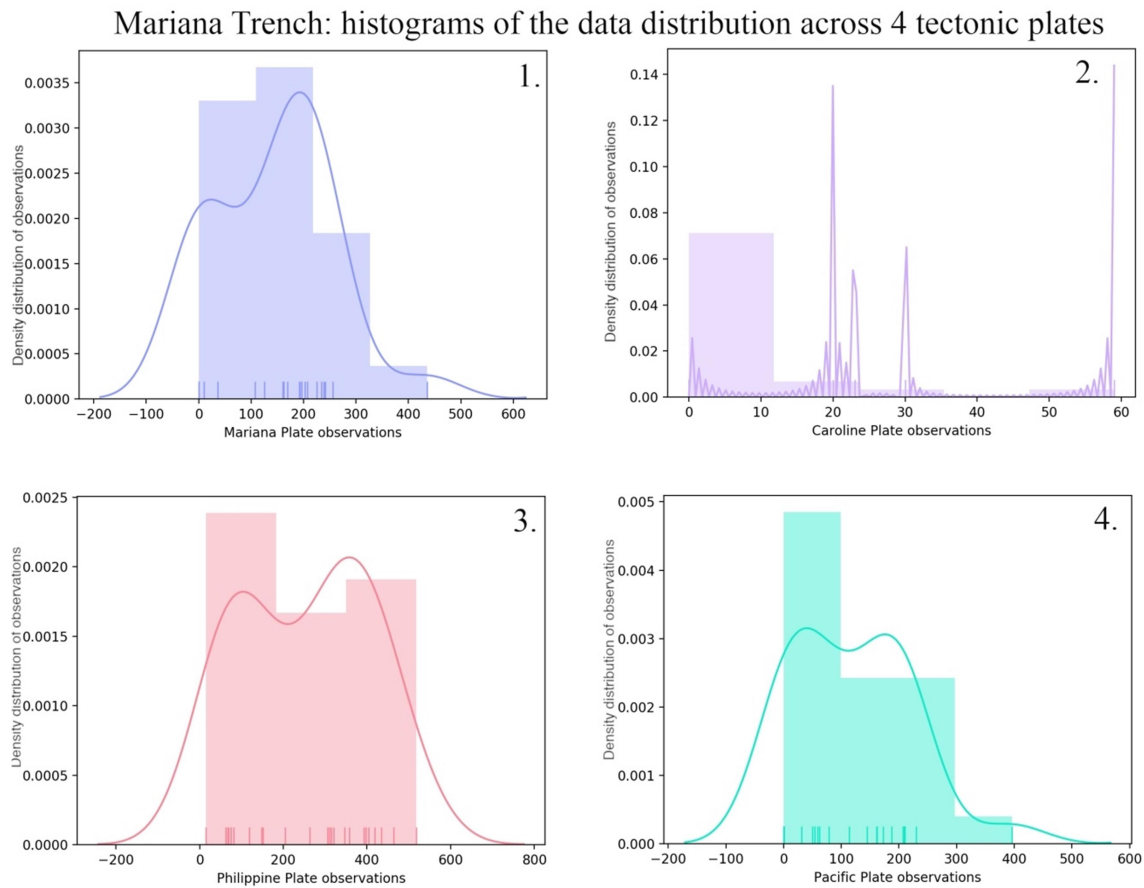


Figure 8. The histograms of the data frequency distribution: bathymetry by 4 tectonic plates.

The histograms show the underlying frequency distribution of a submarine topography of the Mariana Trench (Figure 8). It allows the inspection of the data for the type of distribution that in this case is normal distribution, as well as outliers, skewness. The data are distributed and colored by tectonic plates: Mariana, Caroline, Pacific and Philippine Plates. The histograms of the data distribution across tectonic plates (Figure 8) states that among tested tectonic plates the Caroline Plate shows the least data distribution, which could explain the lower values of the correlation between the data in this particular plate comparing to the Philippine Sea, Mariana and Pacific plates.

2.3. Strip plots showing geological data distribution by four tectonic plates

The next step included visualizing strip plot showing the relationship between the geologic factors (sediment thickness) and

geomorphology (slope angles) as compared by four tectonic plates crossing Mariana Trench. The strip plots visualize categorical data analysis across four tectonics plates. Mariana, Caroline, Pacific and Philippine Sea. The particularity of the scatter strip plots (stripplot()) (Figure 9) lies in its nature: this is a categorical types of plot showing variable dependance (maximal depths values by each profile) upon another variable, i.e. tectonic plates.

It represents the depth values by bathymetric profiles sorted along each of the tectonic plates crossing the trench. On the above plot, one can see the difference on data distribution: the Caroline Plates has just only several points, since it crosses the trench in the south-western small part of the trench. Pacific and Mariana plates are relatively comparable, while Philippine Plate covers the majority of the data.

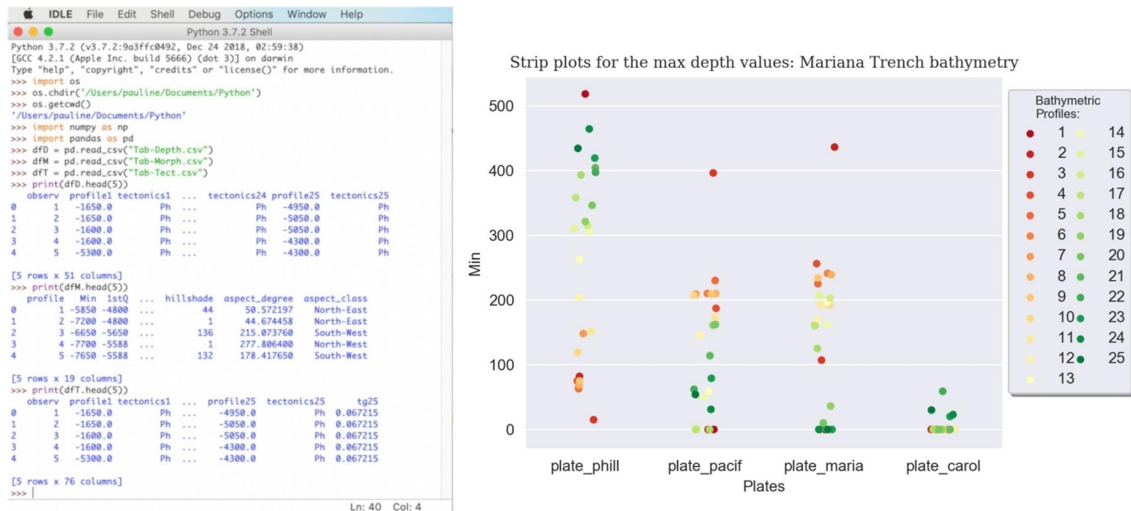


Figure 9. Processing data table by Python Pandas library (left); categorical strip plot showing variations of the bathymetry by four tectonic plates (right).

2.4. Correlograms

The associated names of the environmental variables (e.g. slope angles, sediment thickness, maximal, minimal and mean depths) shown on the correlograms (Figure 10 and Figure 11) were annotated on the axes, because the matrices were

done using NumPy and Pandas objects read in as tables and imported by the Seaborn library. The correlograms show the correlation between the factors affecting the Mariana Trench based on the relationships of the values.

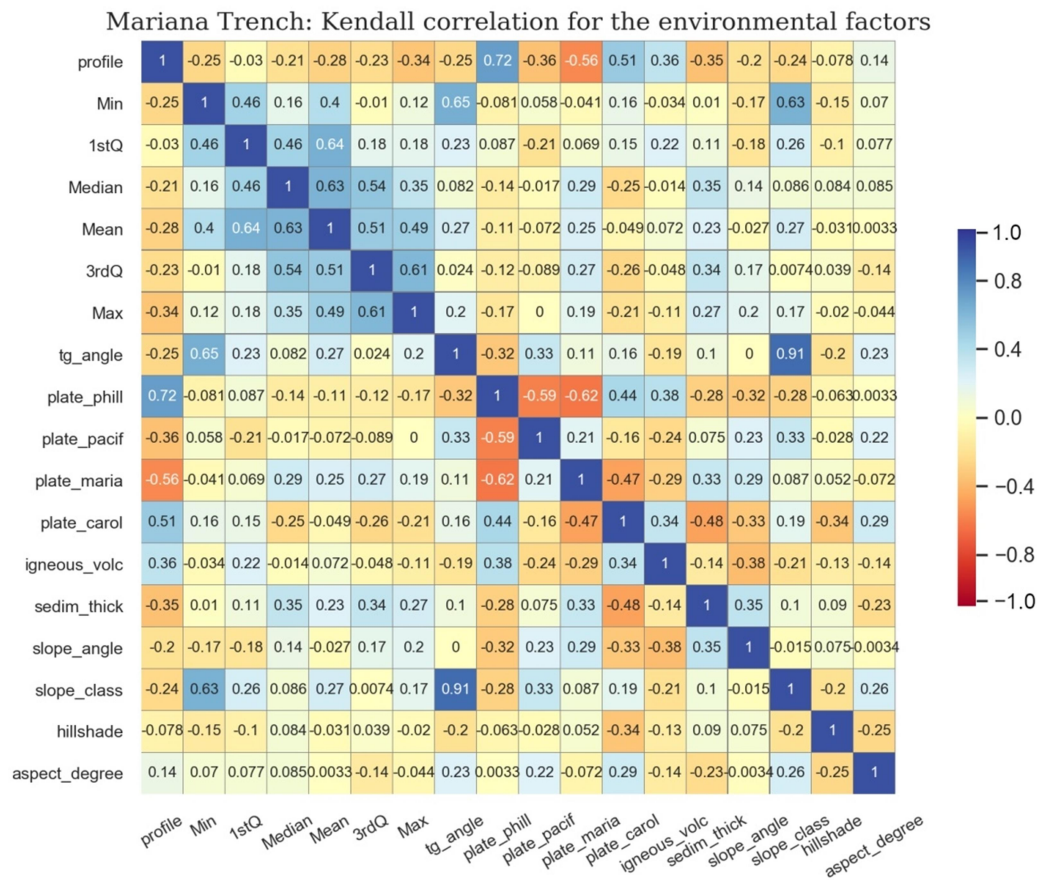


Figure 10. Correlogram matrix of the environmental data affecting Mariana Trench geomorphology, computed by Kendall method.

The external factors potentially affect specific geomorphic features include the following: tectonics, geology and bathymetry of the study area. The effectiveness of the correlograms lies in its properties: the diagonal shows the distribution of each of the environmental variables by a representative square plot which helps to analyze the relationship between the pairs of numerical variables within a matrix. The data were processed using NumPy, Pandas, Matplotlib and Seaborn libraries by Python code presented in the following listing in 6 steps:

```
# Step 1: loading Python libraries
from string import ascii_letters

import numpy as np
import pandas as pd
import seaborn as sb
from matplotlib import pyplot as plt

# Step 2: Importing data
import os
os.chdir('/Users/pauline/Documents/Python
')
sb.set(style="white")
df = pd.read_csv("Tab-Morph.csv")

# Step 3: Computing correlograms
corr = df.corr(method='kendall') # for
Kendall correlogram
corr = df.corr(method='spearman') # for
Spearman correlogram
f, ax = plt.subplots(figsize=(11, 9))
cmap="RdYlBu" # for Kendall correlogram

# Step 4: Selecting palettes
pal = sb.color_palette("RdGy_r") # for
Spearman correlogram reversed color
palette
cmap=pal # for Spearman correlogram
```

```
# Step 5: Drawing the correlogram
heatmap
sb.heatmap(corr, cmap=cmap, vmax=0.3,
center=0, annot=True,
annot_kws={"size":8},
square=True, linewidths=.05,
linecolor='grey',
cbar=True, mask=False, cbar_kws={"shrink":
.5},)
```

```
# Step 6: Adding titles, adjusting subplots
and labelling x/y ticks
plt.title('Mariana Trench: Spearman
correlation for the environmental factors',
fontsize=13, fontfamily='serif')
plt.subplots_adjust(bottom=0.20,top=0.90,
right=0.90, left=0.10)
plt.xticks(rotation=30, size=9)
plt.yticks(rotation=0, size=9)
plt.show()
```

The correlation between each pair of variable is visualize through a numerical symbols and highlighted by colors. The use of both Spearman and Kendall correlation techniques is widely used for exploratory analysis to observe a matrix. In this sense, it may be observed (Figure 10 and Figure 11) that to some degree the interplay between various tested environmental factors has a state of intercorrelation which is characteristic of the interdependencies between the factors: geology, oceanography, tectonics and bathymetry that affects the environmental dynamics and geomorphic structure of the trench as a whole. Thereby can be concluded that the environmental linkages exist between these particular factors presented on the correlograms (Figure 10 and Figure 11) highlight a significant level of dependence between the individuals variables, such as bathymetry (maximal and minimal depths), geography (location of the observation points), tectonics and geology (crossing four tectonic plates:

Mariana, Caroline, Pacific and Philippine Sea). The correlogram ellipses and numerical correlograms (Figure 12 and Figure 13, respectively) were plotted to visualize correlation between the factors

affecting the morphology of the trench. The methodological approach includes calling and executing algorithm of ellipse correlogram {ellipse} library by R programming language.

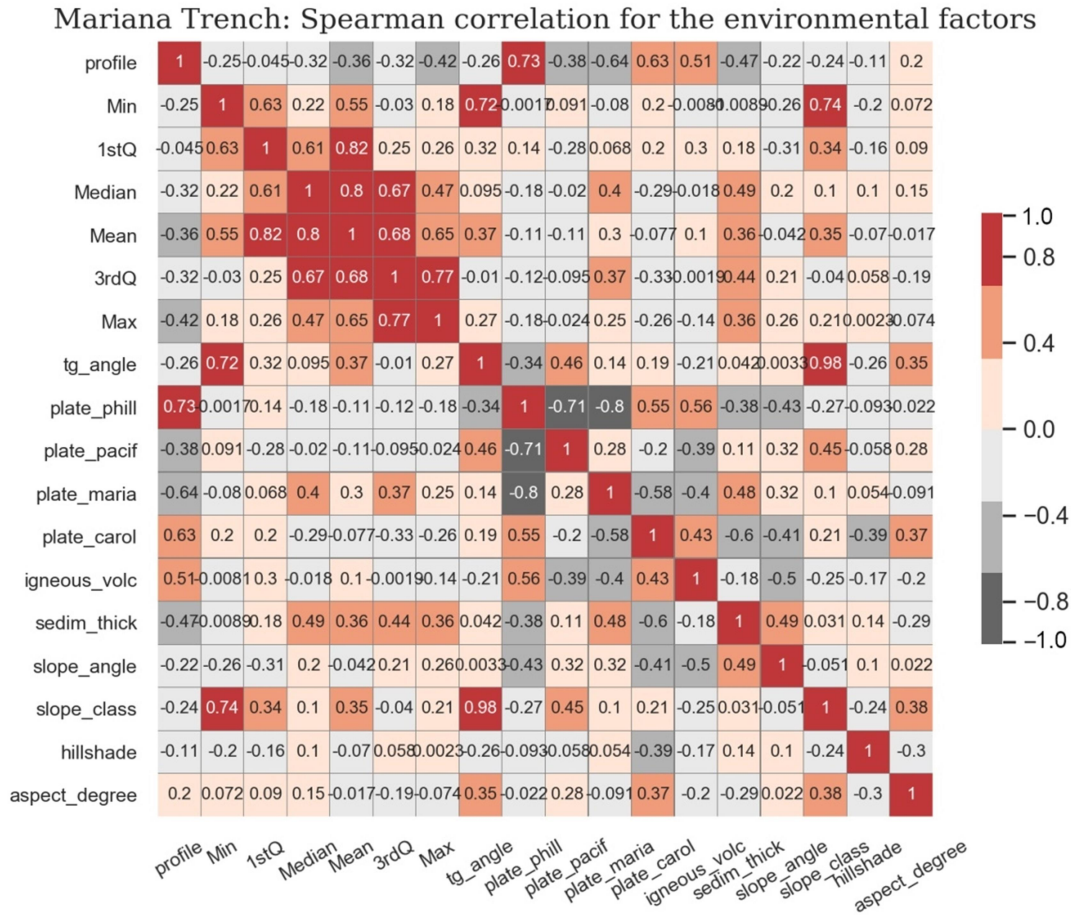


Figure 11. Correlogram of the environmental data affecting Mariana Trench geomorphology, computed by Spearman method.

The R code used to plot correlation ellipses is as follows:

```
data=cor(MDF)
```

```
#step-1 Build a Panel of colors by Rcolor Brewer
```

```
my_colors <- brewer.pal(5, Spectral)
my_colors= colorRampPalette(my_colors)(100)
```

```
#step-2. Order the correlation matrix, to distribute colours by the factors of the values
```

```
ord <- order(data[, ])
data_ord = data[ord, ord]
```

```
#step-3. variant-1, correlation ellipses:
plotcorr(data_ord ,
col=my_colors[data_ord*50+50],
mar=c(1,1,1,1),
outline = TRUE, numbers = FALSE,
main = 'Mariana Trench: Correlation Ellipses',
xlab = 'Geomorphological, bathymetric and geological factors',
ylab = 'Geomorphological, bathymetric and geological factors', cex.lab = 0.7)
```

```
#step-4. variant-3, numerical correlation:
plotcorr(data_ord ,
col=my_colors[data_ord*50+50],
mar=c(1,1,1,1), outline = TRUE, numbers =
TRUE,
main = 'Mariana Trench: Numerical
Correlation',
xlab = 'Geomorphological, bathymetric and
geological factors',
ylab = 'Geomorphological, bathymetric and
geological factors', cex.lab = 0.7)
```

Another way of the correlation data analysis is presented by scatterplot combined matrix that enables to visualize several approaches on one matrix (Figure 14): scatter plots, box plots with residuals, margin plots and surface plots. The scatterplot matrix was done using R language. The scatterplot matrix is plotted to visualize correlation and connections between the environmental variables constituting Marine Trench system.

Mariana Trench: Correlation Ellipses

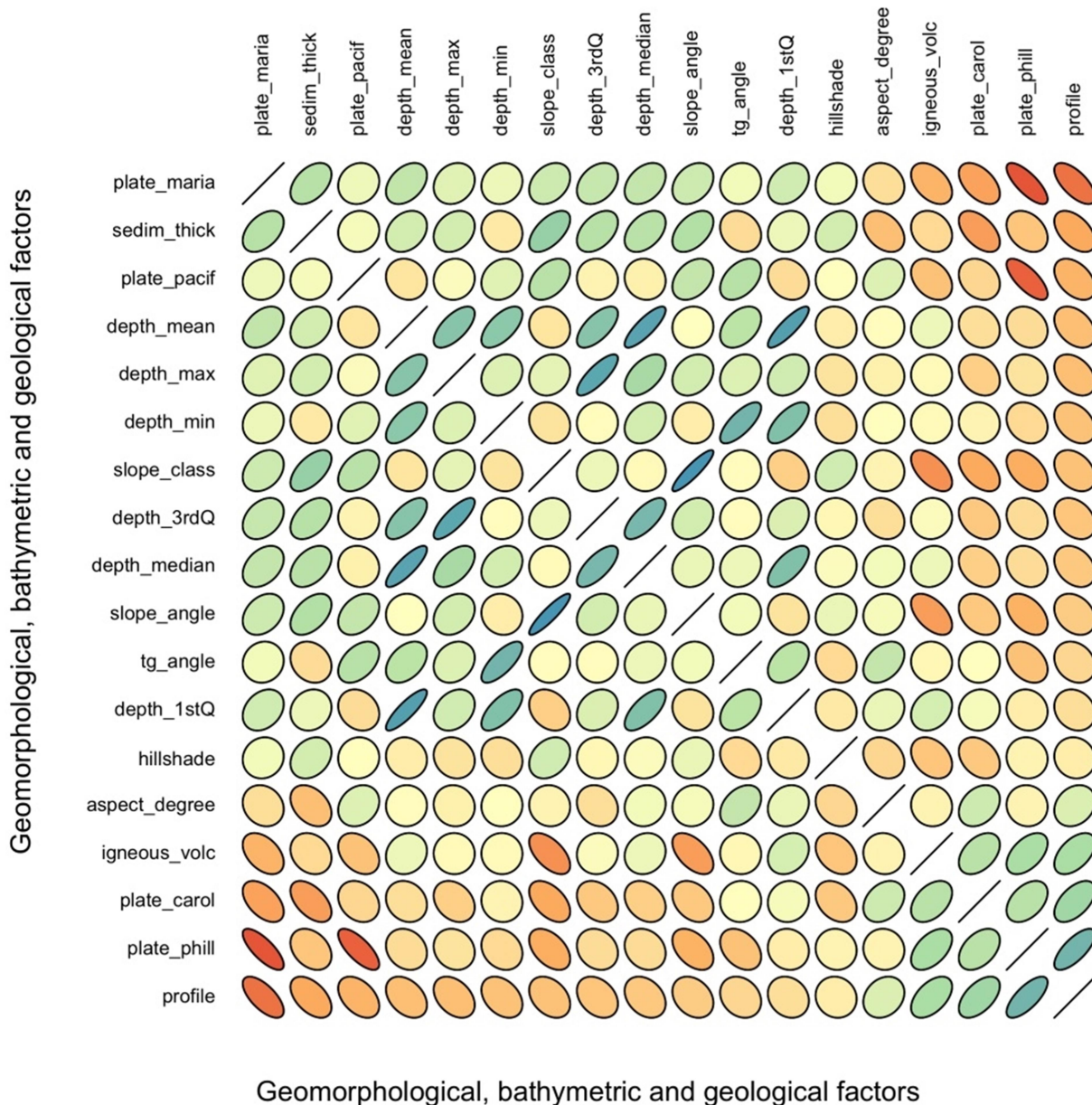


Figure 12. Correlation ellipses.

Mariana Trench: Numerical Correlation

Geomorphological, bathymetric and geological factors	plate_maria	sedim_thick	plate_pacif	depth_mean	depth_max	depth_min	slope_class	depth_3rdQ	depth_median	slope_angle	tg_angle	depth_1stQ	hillshade	aspect_degree	igneous_volc	plate_carol	plate_phill	profile
	plate_maria	10	4	1	4	2	1	3	3	4	3	1	3	1	-2	-4	-5	-8
sedim_thick	4	10	1	3	3	-1	6	4	4	5	-2	1	3	-4	-2	-5	-3	-5
plate_pacif	1	1	10	-1	1	2	4	-1	-1	4	4	-2	0	2	-3	-2	-7	-4
depth_mean	4	3	-1	10	7	6	-1	7	8	0	4	9	-1	0	1	-2	-2	-4
depth_max	2	3	1	7	10	2	2	8	5	3	2	3	-1	-1	0	-3	-1	-4
depth_min	1	-1	2	6	2	10	-2	0	3	-1	7	7	-2	0	0	-1	-2	-4
slope_class	3	6	4	-1	2	-2	10	1	0	9	0	-3	3	0	-6	-5	-5	-3
depth_3rdQ	3	4	-1	7	8	0	1	10	7	3	0	2	0	-2	1	-3	-2	-3
depth_median	4	4	-1	8	5	3	0	7	10	1	1	7	1	1	1	-3	-2	-3
slope_angle	3	5	4	0	3	-1	9	3	1	10	1	-1	2	1	-5	-3	-4	-3
tg_angle	1	-2	4	4	2	7	0	0	1	1	10	4	-2	4	0	0	-4	-2
depth_1stQ	3	1	-2	9	3	7	-3	2	7	-1	4	10	-1	2	2	1	-1	-2
hillshade	1	3	0	-1	-1	-2	3	0	1	2	-2	-1	10	-2	-3	-3	-1	-1
aspect_degree	-2	-4	2	0	-1	0	0	-2	1	1	4	2	-2	10	-1	3	-1	2
igneous_volc	-4	-2	-3	1	0	0	-6	1	1	-5	0	2	-3	-1	10	4	5	5
plate_carol	-5	-5	-2	-2	-3	-1	-5	-3	-3	-3	0	1	-3	3	4	10	4	5
plate_phill	-8	-3	-7	-2	-1	-2	-5	-2	-2	-4	-4	-1	-1	-1	5	4	10	7
profile	-7	-5	-4	-4	-4	-4	-3	-3	-3	-3	-2	-2	-1	2	5	5	7	10

Geomorphological, bathymetric and geological factors

Figure 13. Numerical correlograms.

RESULTS AND DISCUSSION

Studies revealed that multiple factors affect Mariana Trench geomorphology. The statistical analysis performed by Python and R languages revealed that the major depth observation points of the Mariana Trench are located in between the -3000 and -5000 meters (Figure 7, hexagonal plot, right). Another significant point is that sediment thickness changes notably both within the trench by profiles and between four tectonic plates that Mariana Trench crosses: Philippine, Pacific, Mariana

and Caroline. There is a notable decrease of depth (Figure 5, ridge line plots) in profiles 20, 21, 22, crossing Caroline tectonic plate may be caused by the changes in the geologic properties of the tectonic plates. Profiles from Nr. 20 to 22 shows the deepest depth data range with Nr 21 the deepest one.

Since the tectonic attribute values differ significantly, the comparative analysis of how the data vary across four distinctive

tectonic plates was undertaken. As can be seen on Figure 7 (hexagonal plot, right) and Figure 9 (categorical strip plots showing the distribution of data by tectonic plates), the majority of the

observation points are located in the Pacific and Philippine plate, following my Mariana plate, while Caroline plate only covers a few points.

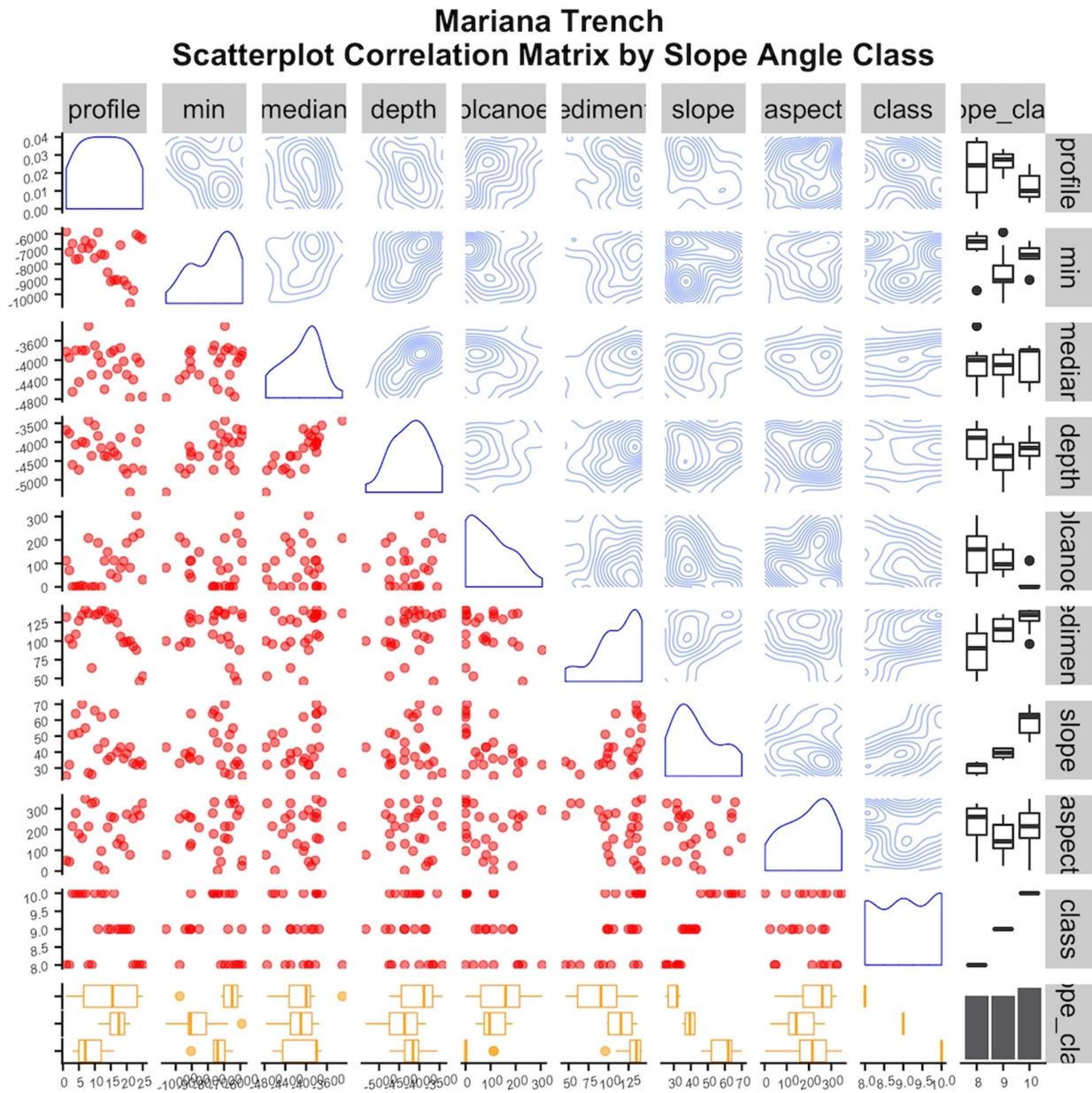


Figure 14. Scatterplot matrix showing geomorphic and geological variables.

Geologic factors, such as sedimentation processes, differ by four tectonic plates (Mariana, Pacific, Philippine and Caroline), due to the overall complexity of the sediment oceanological processes and trigger factors and mechanisms. Furthermore, the results indicate (Figure 14) that sediment sources at the Mariana Trench are connected with trench tectonics. The ridgeline plots (Figure 5)

visualize combine information about statistical data on the 25 bathymetric profiles: arithmetic density distribution, mean and standard deviation along the profiles in a faceted way. Two plots show data distribution as depths by profiles (left) and data density across tectonic plates: Mariana, Philippine, Caroline and the Philippine Sea (right). The lines extending horizontally from the boxes indicate

variability of the trench depths distribution across all the profiles. The ridgeline plots

were generated using {ggridges} and {ggplot2} libraries of R.

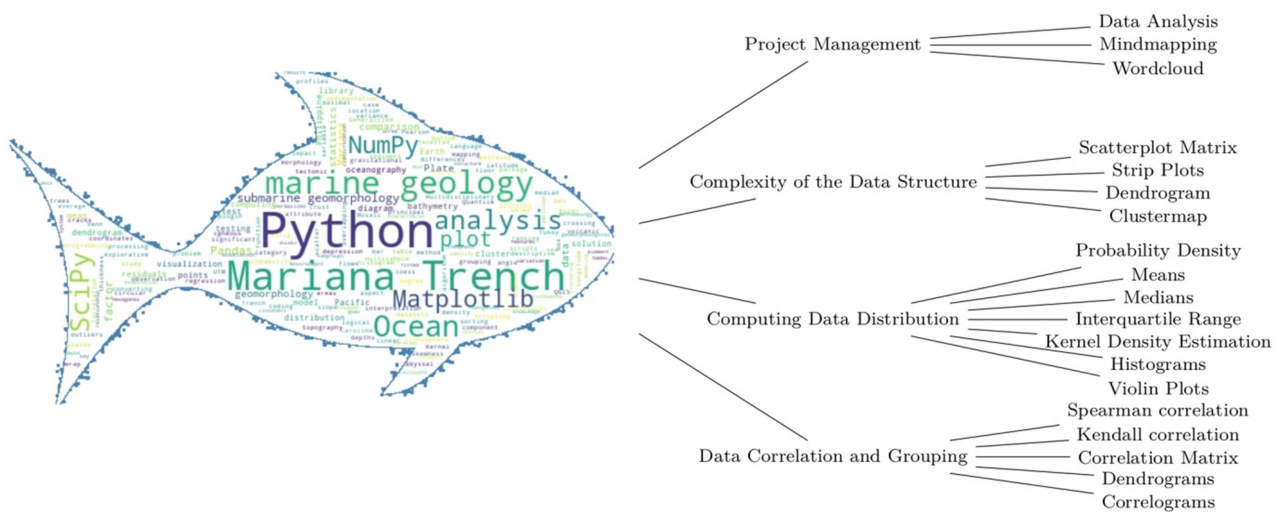


Figure 15. A graphical word cloud and hierarchical structure of the used algorithms representing the conceptual scheme of the research

To schematically demonstrate the use of various approaches, tools and ideas used, merged, combined and applied in this study, the attempt to visualize it as a graphical word cloud was made (Figure 15). This research uses the statistical approaches to analysis a large dataset

(12,950 observations taken from the attribute tables) using Python and R languages. The application of the statistical algorithms embedded into programming languages towards studies in marine geomorphology proved to be an effective tool for data analysis.

REFERENCES

- Arrospide, F., Mao, L. & Escuriaga, C. (2018) Morphological evolution of the Maipo River in central Chile: Influence of instream gravel mining, *Geomorphology*, 306, 182–197. doi: 10.1016/j.geomorph.2018.01.019
- Berger J.F. (2011) Hydrological and post-depositional impacts on the distribution of Holocene archaeological sites: The case of the Holocene middle Rhône River basin, France, *Geomorphology*, 129, 167–182. doi: 10.1016/j.geomorph.2011.02.021
- Bechter, T., Baumann, K., Birk, S., Bolik, F., Graf, W. & Pletterbauer, F. (2018) LaRiMo – A simple and efficient GIS-based approach for large-scale morphological assessment of large European rivers, *Science of the Total Environment*, 628–629, 1191–1199. doi: 10.1016/j.scitotenv.2018.02.084
- Bertoldi, W., Welber, M., Gurnell, A. M., Mao, L., Comiti, F. & Tal, M. (2015) Physical modelling of the combined effect of vegetation and wood on river morphology, *Geomorphology*, 246, 178–187. doi: 10.1016/j.geomorph.2015.05.038
- Chadwick, W.W.Jr., Merle, S.G., Baker, E.T., Walker, S.L., Resing, J.A., Butterfield, D.A., Anderson, M.O., Baumberger, T. & Bobbitt, A.M. (2018) A Recent Volcanic Eruption Discovered on the Central Mariana Back-Arc Spreading Center, *Frontiers in Earth Science*, 6, 1–16.
- Cielen, D., Meysman, A.D.B. & Ali M. (2016) *Introducing Data Science. Big Data, Machine Learning and More, Using Python Tools*. Manning, Shelter Island, U.S., 322.
- Curtis, A.C. & Moyer, C.L. (2005) Mariana forearc serpentine mud volcanoes harbor novel communities of extremophilic Archaeal. *Geomicrobiology Journal*, 30(5), 430–441. doi: 10.1080/01490451.2012.705226
- Faccenna, C., Holt, A.F., Becker, Th., Lallemand, S. & Royden, L.H. (2018) Dynamics of the Ryukyu/Izu-Bonin-Marianas double subduction system.

- Tectonophysics, 1, 229–238, ISSN: 0040-1951. doi: 10.1016/j.tecto.2017.08.011
- Fujioka, K., Okino, K., Kanamatsu, T. & Ohara, Y. (2002) Morphology and origin of the Challenger Deep in the Southern Mariana Trench. *Geophysical Research Letters*, 29 (10), 13-72. doi: 10.1029/2001gl013595
- Halterman, R.L. (2011) *Learning to program with Python*. Python Software Foundation, 283.
- Harrington, A.N. (2015) Hands-on Python Tutorial Release 1.0 for Python Version 3.1+. Loyola University Chicago, US, 191.
- Hudson, P.F., Middelkoop, H. & Stouthamer, E. (2008) Flood management along the Lower Mississippi and Rhine Rivers (The Netherlands) and the continuum of geomorphic adjustment. *Geomorphology*, 101, 209–236. doi: 10.1016/j.geomorph.2008.07.001
- Kawabe, M. (1993) Deep water properties and circulation in the western North Pacific. In: Teramoto, T. (ed.) *Deep Ocean Circulation: Physical and Chemical Aspects*, 17–37.
- Kitahashi, T., Jenkins, R.G., Nomaki, H., Shimanaga, M., Fujikura, K., & Kojima, S. (2014) Effect of the 2011 Tohoku Earthquake on deep-sea meiofaunal assemblages inhabiting the landward slope of the Japan Trench. *Marine Geology*, 358, 128–137. doi: 10.1016/j.margeo.2014.05.004
- Marchese, E., Scorpio, V., Fuller, I., McColl, S. & Comiti, F. (2017) Morphological changes in Alpine rivers following the end of the Little Ice Age *Geomorphology*, 295, 811–826. doi: 10.1016/j.geomorph.2017.07.018
- R Development Core Team, (2014) R: A language and environment for statistical computing. Vienna, Austria. URL: <https://www.r-project.org/>
- Taira, K., Kitagawa, S., Yamashiro, T. & Yanagimoto, D. (2004) Deep and Bottom Currents in the Challenger Deep, Mariana Trench, Measured with Super-Deep Current Meters. *Journal of Oceanography*, 60(6), 919–926. doi: 10.1007/s10872-005-0001-y
- Theberge, A. (2008) Thirty years of discovering the Mariana Trench. *Hydro International*, 12, 38–39. issn: 1385-4569 url: <https://www.hydro-international.com/content/article/thirty-years-of-discovering-the-mariana-trench>
- Tian, J., Fan, L., Liu, H., Liu, J., Li, Y., Qin, Q., Gong, Z., Chen, H., Sun, Z., Zou, L., Wang, X., Xu, H., Bartlett, D., Wang, M., Zhang, Y.-Z., Zhang, X.H., & Zhang, C.L. (2018) A nearly uniform distributional pattern of heterotrophic bacteria in the Mariana Trench interior. *Deep-Sea Research Part I: Oceanographic Research Papers*, 0, 00–00. doi: 10.1016/j.dsr.2018.10.002.
- Wessel, P., Smith, W.H.F., Scharroo, R., Luis, J. & Wobbe, F. (2017) GMT Documentation Release 5.4.2. 314 pp. [Online] <http://gmt.soest.hawaii.edu/projects/gmt/wiki/Documentation>
- Wessel, P., & Smith, W.H.F. (1995) New version of the Generic Mapping Tools released, *EOS Trans. AGU*, 76, 329.
- Zhang, F., Lin, J., & Zhan, W. (2014) Variations in oceanic plate bending along the Mariana trench. *Earth and Planetary Science Letters*, 401, 206–214. doi: 10.1016/j.epsl.2014.05.032
- Zhou, Z., Lin, J. & Behn, M.D. (2015) Mechanism for normal faulting in the subducting plate at the Mariana Trench. *Geophysical Research Letters*, 42. 4309–4317. doi: 10.1002/2015GL063917