

# Using TeX Markup Language for 3D and 2D Geological Plotting

Vasiliy Lemenkov, Polina Lemenkova

### ▶ To cite this version:

Vasiliy Lemenkov, Polina Lemenkova. Using TeX Markup Language for 3D and 2D Geological Plotting. Foundations of computing and decision sciences, 2021, 46 (3), pp.43-69. 10.2478/fcds-2021-0004 . hal-03158568

## HAL Id: hal-03158568 https://hal.science/hal-03158568

Submitted on 4 Mar 2021

**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 4.0 International License

sciendo

DOI: 10.2478/fcds-2021-0004

### ISSN 0867-6356 e-ISSN 2300-3405

#### Using TeX Markup Language for 3D and 2D Geological Plotting

Vasiliy Lemenkov\*, Polina Lemenkova\*\*

Abstract. The paper presents technical application of TeX high-level, descriptive markup language for processing geological dataset from soil laboratory. Geotechnical measurements included equivalent soil cohesion, absolute and absolute deformation index, soil compressibility coefficient by time of immersion depth, exposure time to compressive strength to samples and physical and mechanical properties (humidity, density). Dataset was received from laboratory based experimental tests of the phys- ical and mechanical properties of soils. Data were converted to csv table and processed by LaTeX. Methodology is based on LaTeX packages: {tikz}, {tikz-3dplot}, {tikzpicture}, {pgfplot}, {filecontetns}, {spy} for 3D plotting showing correlation in variables and descriptive statistical analysis based on the data array processing. Results demonstrated LaTeX scripts and graphics: 2D and 3D scatterplots, ternaries, bar charts, boxplots, zooming techniques detailing fragment of the plot, flowchart. Research novelty consists in technical approach of TeX language application for geo-logical data processing and graphical visualization. Engineering graphics by TeX was demonstrated with screenshots of the codes used for plotting.

Keywords: LaTeX, TeX, markup language, plotting, data visualization, data analysis, geotechnical modelling, soil compression, 3D modelling, 2D modelling

#### 1. Introduction

The presented research aims at the assessment of the soils properties in engineering and geological works where road construction is planned. In geotechnical engineering, methods of data analysis include studying physical and mechanical properties of soil aimed to estimate its suitability to loads and pressure. The objective of the laboratory tests of the soil compressibility was to estimate bearing capacity of different types of soils and their physical and mechanical properties prior to road constructions by equivalent cohesion of

<sup>\*</sup> Geoprojectsurvey LLC, Moscow, Russian Federation, https://orcid.org/0000-0002-6982-4089 wasiliy.lemenkov@gmail.com

<sup>\*\*</sup>Schmidt Institute of Physics of the Earth, Russian Academy of Sciences. Laboratory of Regional Geophysics and Natural Disasters (No. 303). 10 Bolshaya Gruzinskaya St., Bld. 1, Moscow, 123995, Russian Federation, https://orcid.org/0000-0002-5759-1089 pauline.lemenkova@gmail.com

soils [8], [41], [62]. In principle, three types of strength can be analysed to measure soil structure: tensile strength, compressive strength, and shear strength. The compressive strength of soils reflects their capacity to withstand loads from constructed buildings, roads, railways and other infrastructure, to reduce size of particles and soil mass [19], and is being measured on a testing machine [38]. In this experimental testing, the KrioLab equipment was used. The compressive strength is a key value in design of structures. Soil cohesion is a component of the shear strength in soils, measured prior to the construction works. Soil cohesion is caused by cementing by Fe2O3,CaCO3,NaCl, chemical elements of soil structure [37] and other affecting factors, e.g. capillary pressure.

The question of which software to use for such a research based on data received from the series of the experiments in a soil laboratory is not an unambiguous one and has been discussed previously [13], [14], [34]. There is a traditional software where one can build graphs, tables and carry out primitive data processing. For a quick visualization of a small amount of data, standard programs, e.g. SPSS, MS Excel, ArcGIS, can be applied [26], [48] to visualize correlation between data and plotting their distribution. For instance, it could be line of dot graphs showing variations of category X on Y. However, the graphical view and the design representation of such graphs remains questionable. Therefore, for more advanced cases, the advanced methods of complex data analysis should be applied.

The applications of modelling in geology and civil engineering are diverse. To mention a few examples of such works: technique of soil stabilization testing for construction roads and bridges [55], determining the increase of the stiffness of the subsurface by Spectral Analysis of Surface Waves (SASW) [11], optimizing binder blends for shallow stabilization of fine-grained soils [53], testing compression strength of concrete by the analysis of the volume of cement particles with sizes [70], testing stability and mechanical strength of the trays affected by the heat input [52], studying compaction properties and sensitivity to moisture content of the fine-grained tills by the Moisture Condition Value test [23], determination of potential soil acceptability for earthworks [56], applying spectroscopy methods for detecting early micro damage in concrete and other building materials (stabilized soil) [63], application of digital image analysis and pattern recognition for construction works [22].

Other approaches may include high-level programming languages in civil engineering and geological data analysis. Such cases are especially useful if dealing with a large heterogeneous data arrays and the output requires modeling of different categories, plotting facetted plots, as well as printer-quality graphics. In these cases using advantage of the statistical libraries of the programming languages is evident (e.g. R, Python, Octave, Matlab) [44], [45], [46]. The advantages of the machine learning applied for data analysis in civil engineering consists in combination of both mathematical powerful methods of modelling and a high-quality data visualization. LaTeX, a high-level, descriptive markup language combines a graphical quality of the data visualization similar to the best statistical packages of R or Python, with simplicity and straightforward approach to reading tables, description statistical data visualization and plotting.

Hence, it can be effectively used for quick yet effective data visualization. This paper presents a technical application of TeX [27], [28], [29], [31] for processing geological dataset. The data were received from the experimental sample series of the soil laboratory. The methodology is based on the existing workflow of geological data processing and sampling various soil types by compressive strength [11], [39], [43], [54].

#### 2. Methodology

#### 2.1. LaTeX packages

The methodology of this research is based on TeX language [21], [30], [32], [33]. Hence, the prerequisites included the installation of LaTeX. The main LaTeX package used in this work is a {pgfplot}, which is a free (open source) software, specially designed for 2D and 3D plotting [15], [16]. It works under various OS platforms including Mac OS, Linux, and Windows. The MacTeX release was used in this work.

In modelling presented in this research, the 2D and 3D correlations of the geological variables showing soil properties are studied and visualized based on a two- and threedimensional statistical models using LaTeX tools. To work with LaTeX, the latest version of TeXShop Editor was installed and LaTeX compiler was used. It contains necessary packages as well as additional required packages: https://www.latex-project.org. The abbreviation 'pgf' means 'portable graphics format'. Technical tools include following LaTeX packages [17]: {tikz}, {tikz-3dplot}, {tikzpicture}, {pgfplot}, {filecontetns}, {spy}. The 3D-visualization showing correlation between variables and descriptive statistical analysis is based on the processing of a data array to show functional dependence between the characteristics of the soil variables.

The methodology included the assessment of the engineering and geological conditions of soils in the study area where road construction works are planned. The study of the soil deformation was performed by compression strength using existing standards and requirements. The equipment was developed by KrioLab (http://kriolab.ru/). Methods of strengthening and stabilizing compressible soils were applied [10], [40]. The laboratory equipment includes a device designed by KrioLab for compressive strength measuring. The approach was used to maintain objectivity in measuring equivalent cohesion of the soil samples.

The test was based on the compression of a cylindrical soil sample by a vertical load with measurement of soil displacement. The consequent data analysis was performed using existing tested methods [3], [5], [39]. The results were stored in a table and processed further by LaTeX.

#### 2.2. Data

The data were collected during geological fieldwork including survey sampling. The data were processed and analysed after they were collected. The table with dataset was received from the data processed in a laboratory as a series of the experimental tests on physical and mechanical properties of soils. Data processing was carried out by statistical methods and visualization. Data were converted to csv table and processed by LaTeX. Reading data from a .csv file was done using package {filecontents}. By this package the data can be read in to the LaTeX environment and used to plot graphs. In this particular case, we analysed changes in the incompressibility of soils under the deformation conditions. In case of a small table with three categories (conditionally, xyz), these variables were directly added as coordinates. In this case, they are not necessarily geographical, but reflect the dependence of the category values on each other: frozen soil properties, deformation, compressibility,

heave by varied temperature. Other physical and mechanical parameters of the soils include pressure, humidity, immersion time, density, etc. Based on the assessment of the engineering and geological soil properties, it was experimentally determined that soil deformation index increases with the increasing compression strength.

#### 2.3. Geotechnical measurements

The study area is located in Yamal Peninsula with dominating frozen soil due to the permafrost distribution. Located in the extreme climate conditions of Arctic, the structure of the frozen soil consists in the solid particles, including mineral or organic grains, gases, unfrozen water and ice [4], [42], [61]. In contrast with other soil types, frozen soil matrix includes ice and water in its content, which may change significantly with varying temperature conditions and applied stress.

Necessarily, this type of soils in such a harsh climate conditions requires special methods of measurements and careful testing prior to the road constructions. Cohe- sion, weight, and friction angle are ones of the most often tested mechanical charac- teristics of such soils. These characteristics are used for assessment of soil suitability for buildings and roads construction [66].

Cohesion of soils is an important factor of soil consistency which is described according to the definition in soil physics as 'cohesive force that takes place between the adjacent particles' [6]. In soil mechanics it is considered as 'shear strength when compressive stresses are equal to zero' [69]. It is generally assumed that the rela- tionship between the compressive stress (p) and corresponding shear strength (s) in cohesive soils can be represented by an empirical equation, which is known as Coulomb-Terzaghi shear strength equation by the Formula 1,

$$s = c + \sigma \tan \varphi$$
 (1)

where  $tan\varphi$  is the coefficient of plane sliding friction which describes the surface roughness and hardness of the materials constituting the slope; *s* is a total shear strength; *c* is cohesion;  $\sigma$  is a normal stress.

Specifically for frozen soils, their mechanical properties closely reflects the ice behavior, which is filling most of the pore space in soil mass, despite the unfrozen water presented in its structure. The pore ice usually consists of polycrystalline type with a random crystal orientation [57]. In usual environmental conditions, the response of a soil to an external stress is controlled by the motion of the dislocations and can be represented by a power law creep equation of the Norton-Bailey type computed by the algebraic Formula 2, [9]:

$$\varepsilon_{cr} = A \sigma^n t^m \tag{2}$$

where *n* and *m* are temperature dependent coefficients; *t* is time and  $\sigma$  is stress; *A* is a creep strain hardening coefficient of material constants, dependent on temperature, but independent of stress.

The deformation in soil also includes changes in shape (expressed as a shear modulus G) which is calculated by the Formula 3, [58]. The shear modulus of soil shows its rigidity, defined as the ratio of shear stress to the shear strain:

$$G = \frac{\tau_{xy}}{\gamma_{xy}} = \frac{(F/A)}{(\delta x/l)} = \frac{Fl}{(A \,\delta x)}$$
(3)

where  $\tau xy = F/A$  is a shear stress, F is an acting force, A is an area, l is an initial length, yxy is a shear strain,  $\Delta x$  is the transverse displacement.

The elasticity of soils is also measured through the estimation of bulk modulus K by the Formula 4 [7], [18],

$$K = -V(\frac{dP}{dV}) \tag{4}$$

where P is an external pressure, V is a volume, and dP/dV is a derivative of the pressure with respect to volume.

Other important types of soil deformation are expressed by the Young's modulus and the Poisson's ratio, measured in uniaxial compression, that is, tests with constant stress on the vertical surfaces. The Young's modulus enables calculation of the change in a dimension of a sample under the compressive loads, expressed by the Formula 5, [58]:

$$E = \frac{\sigma}{\varepsilon}$$
(5)

where E is the Young's modulus,  $\sigma$  is the uniaxial stress and  $\varepsilon$  is the proportional deformation (strain).

The Poisson's ratio describes the expansion or the contraction of soil in the directions perpendicular to the direction of loading which is tested by the Formula 6, [58]:

$$v = \frac{-(d\varepsilon_{trans})}{(d\varepsilon_{axial})}$$
(6)

where v is a resulting Poisson's ratio,  $\varepsilon_{trans}$  is a transverse strain, i.e. negative for

an axial tension (stretching) and positive for an axial compression;  $\varepsilon_{axial}$  is an axial strain (positive for an axial tension, negative for an axial compression).

The compressive strength of a soil is expressed by the value of the uniaxial compressive stress reached when the material fails completely and is expressed by the Formula 7, [1], [2]:

$$\sigma_e = \frac{F}{A_0} \tag{7}$$

where  $\sigma_e$  is the compressive strength, *F* is the force or load at the point of failure and  $A_{\theta}$  is the initial cross-sectional surface area. Although, as a general rule, soil samples would shorten by a compression test, the degree of the deformation depends on the physical properties of various types of the soils.

The calculation of the compression module of the total strain (Ek) was carried out according to the Formula 8:

$$E_{k} = \frac{(\Delta p)}{(\Delta \varepsilon \beta)}$$
(8)

where  $\Delta p$  is a pressure interval on the soil sample taken as  $\Delta p = p2 - p1 = 0.2 - 0.1 = 0.1M$  Pa.  $\varepsilon$  shows changes in the relative vertical deformation of a soil sample, corresponding to  $\Delta p$ ;  $\beta$  is a coefficient taking into account the absence of the lateral expansion of the soil in a compression device, e.g. for loam,  $\beta = 0.6$ .

The relative vertical deformation of the soil sample was determined by the Formula 9:

$$\Delta \varepsilon = \frac{(\Delta h)}{h} \tag{9}$$

where: h is an initial height of the soil sample in mm;  $\Delta h$  is an absolute vertical deformation of the soil sample.

As briefly demonstrated above, various coefficients are being used to evaluate soils' physical and mechanical properties for construction of buildings and roads, describing elastic deformation in the conditions of high pressure under low temperatures. In this study, following variables were measured during the set of experiments and visualized by LaTeX based plotting:

- 1. equivalent soil cohesion,
- 2. absolute and relative deformation indices
- 3. time of immersion depth
- 4. compressive strength to samples
- 5. physical and mechanical properties (humidity, density)

For the case of a complex table with rows of variables and varying indicators, we used a {filecontetns} package in the preamble of the document. The {filecontetns} package allows LaTeX to read data from the long tables with many variables. In this case, we directly indicated the table in the preamble of the code (see the picture below, Figure 1 which presents the variability of the soil properties during the compression strength testing). The environment between the 'begin {filecontetns}' and 'end {filecontetns}' indicates which file we are working with (.csv), what variables are in, and which of them should be used for plotting. In this example, from the whole long table, we selected only 3 parameters for visualization.

#### 3. Results

#### 3.1. Summary

The summary of the performed geological fieldwork is presented in Figure 10. The experiment on the compression testing of the frozen soils determined the effect of the soil density on a compression coefficient. The experiment shown correlation between the compression coefficient and deformation of the frost heaving soil. It also expanded the experimental base for further studies based on factors affecting the phenomenon of frozen soils heaving. It has been found that with the increase in density and humidity, the average

equivalent cohesion increases for a given soil sample (Figure 5). The decrease in porosity correlates with a gradual increase of the deformation (more steep for the absolute deformation) in an average increase of the compression (Figure 9).

The tested soils are from different locations in northern Russia. All of them were processed in a following workflow: 1) soil samples taken during the fieldwork, 2) soil samples processed using KrioLab equipment; 3) soil samples were analyzed and visualized using LaTeX. The soils were all frozen due to the harsh Arctic conditions. Afterwards the soil was manufactured in the laboratory by comparing different soil types and their properties: porosity coefficient, relative strain index, compressibility coefficient, absolute deformation and relative strain.

Soils of the study area have a high porosity, that is, the ratio of the volume of pores filled with water to the volume of heaving soil. Studied samples included clay, sandy loam and loam. It was found that clay soils better retain moisture, because, when frozen moisture transforms into ice and increases the volume of the soil sample, it also expands, i.e. it causes the process of the frost heave. Frost heave was observed in the soil samples is an upwards swelling of the Arctic soils during freezing conditions caused by a high volume of ice in a sample.

The compression coefficient was the main indicator of the compressibility of the icerich frozen soils in this study. Dry hard soil with high density (1.85g/cm3) and low humidity (29.4%) has equivalent cohesion in range of 0.115–0.132. It was detected that loose moist soil with low density (1.65g/cm3) and a high humidity (49.3%) has an equivalent cohesion fluctuating between the 0.163–0.082 which proves that the cohesion process is high. The accuracy index ( $\rho\alpha$ ) varies within the following values: from 0.06 to 0.40 with an average value of 1.11.

The results are graphically summarized in the presented 10 figures demonstrating a LaTeX script and a graphical output: 2D and 3D scatterplots, ternaries, bar charts and boxplots (i.e. probability of data distribution). Two special techniques were presented additionally: a zooming in a plot detailing fragment of the graph by magnifying glass function, and a flowchart by library {tikz} [65]. The screenshots show codes and graphical results. All the graphical outputs are presented using TeX. The 2D (Figure 1, plot left, Figure 5) and 3D graphs (Figure 1, plot right, Figure 2, Figure 3 and Figure 4) for testing soil samples were plotted against the compression strength with a determination of the absolute strain index and a compressibility coefficient.

The two-dimension plotting of variations in an equivalent soil cohesion agains its physical properties (humidity and density) is performed (Figure 5) by LaTeX library {tikzpicture}. The statistical analysis of the data distribution is shown in Figure 6 by options of {pgfsplotlibrary} and {tikzlibrary} libraries {statistics}. Ternary diagrams demonstrated the granulometric composition of soils using the specialized package of LaTeX {ternary} (Figure 7 and Figure 8). The porosity coefficient, compression and relative strain indices, relative and absolute deformations are shown in Figure 1 and Figure 9 by using the .csv table imported by {filecontents} package of LaTeX using techniques of 2D, 3D and bar chart visualization.



Figure 1. Variability of soil properties during compression strength testing. Visualized by LaTeX. Code for plotting graph and graphical output.

We have tested the six variables and a time category with a maximum of 600 hours for a testing period. In the description of the graph environment, we indicated the parameters of graph in the commands shown in Figure 1 and plotted the graph. The results of the deformation measurements are visualized in Figure 1 showing that the porosity coefficient is decreasing from 5,1 to 4,3 during the course of the experiment (600 h). The absolute deformation increased to 2,3 mm by the end of the experiment along with increased normal pressure (0,7 MPa), while fluctuations in a relative staring remained statistically insignificant.



Time series analysis of the variations of the long-term equivalent cohesion of soils  $(C_{eq})$  under increasing depth of the samples. Series-II (M-1-8). 3D:  $\mu$ TeX



#### 3.2. 3D scatterplot of the equivalent cohesion of soils

The Figure 2 shows a workflow for plotting the 3D scatterplot of the correlation between the equivalent cohesion of soils, time of experimental load pressure and the immersion depth of the external loads. A code with explanations (with comments highlighted in red) is describing stepwise which commands were used and which parameters were selected to plot the two correlation graphs. Figure 2 presents the code and graphical plot showing the 3D graph of variability of the equivalent cohesion of soil in LaTeX.

The output shows the dependences of the soil parameters against the external factors (in this case, the depth of the immersion of samples over time indicates the resistivity of soils),

as presented below. In this case, the two 3D graphics from a series of tests were placed as a facetted plot. Following requirements for adjusting plots are demonstrated in Figure 2: rotated axes; rotated axes annotations placed below the axes; small-sized two plotted located sidewise; plotted coordinate ticks; displayed grid; title annotated on top. The source code and the output are visualized in Figure 2.

#### 3.3. Selective zooming of the enlarged fragment

The Figure 3 presents a LaTeX code for plotting zoomed enlarged fragment of a 3D graph and a resulting plot. Creating an enlarged fragment of the graph has been done by {spy} package and magnifier command. Using magnifying glass function was applied to enlarge a fragment of the graph for a more detailed analysis, as shown in Figure 3 which presents the source code above and the output image below.

#### 3.4. Facetted plotting

The principle of the facetted plotting approach (Figure 4) is the same as for the time series graphs. Figure 4 shows a combined plot of six graphs demonstrating a trend of soil behavior for the six different samples. Here we only placed a series of graphs one after another declaring a {tikzpicture} environment from the {tikz} package. The space between the plots is indicated by the {axis} and 'hskip' command that inserts a non-breaking space of the specified width between the graphical objects.

The Figure 4 presents a facetted plot of the several graphics in an experimental set: equivalent soil cohesion, immersion depth and time. Figure 5 presents a scatter 2D plot of the 8-hour equivalent soil cohesion of soil samples against moisture (left) and density (right). The distribution series of the soils parameters are shown on the diagrams of 1D data distribution and a scatter plot in the {pgfplot} LaTeX package. The main tasks of the experiments performed in a soil laboratory included assessment of the soil properties, structure and behavior under varying conditions (pressure).

The measurements of these parameters are performed under various environmental conditions for the soils taken at different depths and wells, as well as factors affect- ing change in the soil characteristics with aim to model and numerically assess soil behavior for the construction purposes. As a study object, soil samples present the heterogeneous multicomponent open dynamic natural systems. Therefore, modeling and visualizing soil compression for the purpose of road construction requires the use of numerical methods of the mathematical and statistical analysis expressed also in a graphical form.



**Figure 3.** Code for plotting zoomed enlarged fragment of a 3D graph of the equivalent soil cohesion with the immersion depth changing in time and graphical result.

The results of the facetted plot (Figure 4) demonstrate a series of data sampling for the six experiments showing variations in the equivalent cohesion of different soil samples. Because the compression of the soil samples is the result of a closer particle packing, a reduction in a pore volume, and an increase in a bulk density, it is possible to visualize the state of the compaction graphically by plotting any of these parameters against the immersion depth and the time of the experiment. The enlarged details of the experiments produced are shown in Figure 3 (zoomed fragment).



Figure 4. Facetted plot of the several graphics on experimental sets: equivalent soil cohesion, immersion depth and time.

#### 3.5. Clustering data

Another important task is the determining of the general nature of data distribution to assess the degree of its homogeneity. The statistical homogeneity of the measured values is characterized by a range of the data variation (scattering) [59]. In this case we tested soil parameters. The mismatch between the values of the specific variables for the statistical units in clusters is assessed as a degree of the data dispersion.

This and other tasks of the basic statistical data processing and visualization can be solved by LaTeX functional. It should be however noted that a complex statistical data analysis including sophisticated methods and mathematical approaches can only be solved using specialized tools, e. g. StatsModel library of Python or powerful functionality of R language. The functionality of LaTeX is more straightforward and can be effectively used for a quick data processing and visualization which enables to avoid learning complex coding syntax of Python or R, yet provides an excellent solution to the graphical data visualization.



Figure 5. Scatter 2D plot of 8-hour equivalent soil cohesion of soil samples against moisture (left) and density (right) from data clusters.

In a series of measurements showing values of the equivalent cohesion of soil samples vary with different loads. We visualized a series of these tests and to evaluate the normality of the data distribution which is commonly described in methods of the statistical data analysis applied to geology [12]. In such a case, visualizing a series of tests and data distribution can be done by LaTeX. Using the 'pgf' functionality, the plot of values was visualized by clusters (Figure 5). In this case, the presence of the object classes that are explicitly defined in the code will be critical: table (meta = class). Following statistical values from the table were entered: median, the minimum and maximum values of the data (the lowest and highest values of the array), lower and upper quartiles (values of the array)

corresponding to 25% for lower quantile and 75% for upper quantile), frequencies on the cumulative graph [24], Figure 5.

#### 3.6. Statistical box plots

The most common case of solving problems in civil engineering, earthworks and geology is a direct task: according to the given characteristics of soils and their reactions to external impact factors we can study behaviour of these soil samples in space and time using methods of the compression strength of samples.

Measured variables include deformation, creep, ductile limit for cohesive soils, heave and cohesion among others. There are also reverse and inverse geological prob- lems, for example, assessing the load pressure on samples by the measured properties of soils [20]. However, the basic essence of the problem in all the cases is the same: interaction between the soil and external pressure, correlations, and reaction of the objects in a dynamic system. The used dataset is discrete since it contains a countable number of the measured values, therefore, the boxplot was used to assess the data distribution. The boxplot (Figure 6) graphically displays the five most important descriptive values for a dataset: the minimum, the maximum, the first quartile, the median and the third quartile. It shows the proportion of the dataset on measured values that belongs to a given interval.

The estimation of the data distribution in space is a common task in the statistical data analysis [20]. In this case, a scatterplot or boxplot are often plotted based on a series of the measured data (Figure 5). In a first approximation, the graph has an ellipse form. In various tasks, the parameters of the ellipse covering points cloud are required for plotting. In other cases, displaying the boxplot of the descriptive statistical data is effective. An ideal boxplot represents the statistical variables (median, quartiles, mean) based on the data distribution [60].

The Figure 6 presents a whisker boxplot showing a diagram of the 1D distribution of statistical probabilities of an 8-hour equivalent cohesion of the soil samples: a series of the 15 measurements. As a general case, the points of cohesion are denser in a center of the box in a boxplot with their density decreased towards the edges and having outliers, which indicates the normal distribution of the data. In other cases, the distribution of the data samples can be graphically displayed as a straight line showing correlation of the measurements against variable. In this case, the boxplot shows a distribution of the equivalent soil cohesion under varied compression strength.



Figure 6. Plotting whisker boxplot: a diagram of 1D distribution of statistical probabilities of an 8-hour equivalent cohesion of the soil samples: a series of 15 measurements.

#### 3.7. Ternary diagrams by LaTeX

The current problems of geological survey are diverse [35], [36] including graphical plotting and visualization, essential in computer data processing. Plotting ternary diagrams is one of the most effective methods of geological data visualization that is less usual comparing to the linear plots. Ternary diagrams are effective for visualization of the relationships providing a basis for studying a wide range of the elements within a complex multi-component system [67], [68]. Ternary plots look like an elegant, logically constructed

triangle showing a triple correlation between the data variables. Since its visualization is not a classic graph showing 2D dependence of x on y or a 3D, plotting ternaries LaTeX requires some explanation.



Sand (0.05-2 mm) **Figure 7.** Ternary diagram: granulometric composition.

80

60

100

0

20

40

-dusty loam

.

100

Omitting the existing solutions possible in R programming which require more advance skills in coding [49], [50], we focus on the more straightforward functionality of LaTeX {pgfplot} and its {ternary} library. The ternary diagram allows to visualize variations in the lithological composition of soils or rock samples in the given territory or to trace their evolution during the geological time in an area through comparing ternaries of the rocks of the same age in a cross-section.

Using ternary technique enables to visualize a three-component system in form of a cloud of points with elements in total composing 100% of the structure. Hence, each of the peaks corresponds to 100% of one of the three components, and the side opposite it corresponds to the zero content of the same component. Figure 7 and Figure 8 present the two ternary diagrams showing the granulometric composition of the soil with Figure 8 using a color-plotted ternary.



Granulometric composition of the soil. Fractions: dust, clay, sand



Figure 8. Ternary diagram: granulometric composition of the soil.

Plotting ternary diagram showing the lithological composition of rocks or soils by LaTeX{pgfplot} can be done using special libraries: pgfplotslibrary {ternary} and tikzlibrary {pgfplots.ternary}. In a LaTeX environment, ternary diagram visualizes a three-component rock system in such a way that the sum of all these components gives 100%. Therefore, it is logical that ternary diagram visually represents a system located on the triangular axes.

Hence, {ternaryaxis} environment of LaTeX works with relative coordinates: each data point consists of the three components x, y, z. Their total amount sums to 100%. So, the while data on one axis stretches in any of these three components, the other two will automatically decrease, but in total the whole will be 100%. The convenience in ternary approach lies in its specifics: if one chooses any absolute values within the data range, the LaTeX will still recalculate them into a system of barycentric (trilinear) coordinates. Therefore, in a standard configuration, we have x, y, z placed within the range (0,1).

Plotting ternary diagram was done by executing commands demonstrated in the code (Figure 7 and Figure 8). Explanations of the specific commands is highlighted in red. The printscreens in Figure 7 and Figure 8 show the techniques of ternary diagram in details, with comment explanations in red color.

At the next step, ternary was plotted in color Figure 8 using the same data with an output shaded by colors selected by the diverse color palettes. The description and the codes are presented here: http://latexcolor.com/ (accessed 16.04.2020) enabling to select color both by name, by Hex RGB Triplet, or by defining individual colors through LaTeX 'definecolor' declaration [25]. For example, used color 'Dark Mandarin' was done using the following example. The RGB triplet was defined explicitly in the document preamble, and the color was then used directly in the work (here: fill = darktangerine).

Since rocks usually consist of many components, visual arrangement of the granulometric ternaries was divided into the three groups according to the dialogical and genetic characteristics. For example, in case of the terrigenous rocks, all fractions of sand were combined in a first group (bottom right). The second group included all fractions of silt. The third group included pelite and clay material (clay and clay rocks). Depending on the goals, data could also be grouped accordingly as independent clusters using other properties, e.g. carbonate rocks, calcite, dolomite and insoluble minerals. The LaTeX {ternary} library can also apply contours and smooth lines function for visualization.

#### 3.8. Bar charts on geological variables

The Figure 9 presents a 2D bar chart showing variables measured against the time and the compressive strength. Another example consisted in a task to build a data distribution diagram (a box plot, or a 'whisker plot'). To do this, the LaTeX function 'boxplot' was used. According to the description of the code, the graph was displayed automatically (see the illustrations below). The source code and the result of the statistical data distribution diagram are shown in Figure 9.

An important technical point: LaTeX can read data from the files in .dat or .csv (comma separated value) formats. In the latter case, when importing data from the Excel tables, one needs to make sure that the distances between the columns are comma separated. Using the available tabular data on measurements of the equivalent soil cohesion (in this case, tests were carried out using the equipment developed by the KrioLab). The data were then filled directly into the body of the code: the upper and lower quartiles, the upper and lower 'whiskers', the extreme values for the entire data sample and the median.



Figure 9. 2D bar charts showing variables measured against time and compressive strength.

The second important moment concerns the data structure: the file should have a period as a decimal separator (not a comma). Hence, because the initial Excel table had a decimal point, the commas were changed in the entire document to points using simple method in the Excel: "Find & Replace". Afterwards, the file was saved with a .csv extension and read into the LaTeX environment according to the described code sample in the screenshots shown above. The third important technical moment concerns the path to the data. Thus, the .csv file should be saved in the same folder as a LaTeX raw .tex file. Otherwise, the path should be declared explicitly. In terms of file management, it is also better to keep all the files of the current project in one directory.

#### 4. Conclusion

#### 4.1. Importance

The importance of data visualization in the engineering geology consists in a precision and correctness in civil engineering works which necessarily set up high requirements for the data processing and visualization. Incorrect assessment of the soil properties may result in a selection of the wrong decisions when constructing roads, railways and buildings. Negative environmental consequences may arise in case of the increased loads on the frozen soils in the harsh climate conditions which can eventually lead to the crumbling roads. The LaTeX libraries for typesetting and document preparation has gained a wide acceptance in both academia and industry, because it provides an excellent performance on the typographic issues. However, using special libraries of LaTeX specifically for the 3D plotting has not been demonstrated widely so far in a scientific graphics and plotting. Instead, specially designed software (such as the SPSS of the Excel MS) is used traditionally for basic plotting, or for the more complex statistical analysis, the programming languages such as R, Python or MATLAB.

Plotting the 2D and 3D graphs accompanying calculations and computing tasks in the engineering geology is an essential part of the earthworks, since the data visualization enables more effective evaluation of the correlations between the parameters and variables of the soil samples. Graphical representation of the correlations between the selected parameters of soils plotted against each other and a compression strength allows to visually display the physical and mechanical properties of the soil samples.

#### 4.2. Importance

In contrast to the existing approaches, this paper demonstrated excellent yet underestimated possibilities of LaTeX for the 2D and 3D data visualization in a scientific plotting, as well as its ability to read in the .csv tables, to perform data processing resulting in high-quality plotting. Given the difficulties of mastering the syntax of the programming languages on the one hand, and insufficient quality of graphs presented by the traditional statistical software on the other, LaTeX might be a good solution for plotting graphics in a geological survey. To illustrated the abilities of LaTeX for workflow diagrams, Figure 10 presented the flowchart of the research project summarizing a cycle of the performed works on soils measurements plotted by LaTeX.

The research process in geology and civil engineering normally includes various steps that can be summarized in a following workflow: collecting data in a fieldwork, sorting and organizing data samples, storing tables, formats conversion, data processing, data modelling and visualization, Figure 10. Despite the equal importance of all these steps in a research organization, the mathematical modeling and visualization of data definitely deserves a special attention. Correct and effective models assist in solving issues related to the calculation of the soil properties, predict possible de- formations during the temperature extremes and loads by infrastructure, determine suitability of road constructions in a given tested region.



Figure 10. Flowchart of the research project: soils measurements.

Construction works (roads, building, infrastructure) usually include collection, sorting and processing the big datasets that include a large amount of stored information. This includes geological documentation of exploration, test series of the soil sampling, as well as the results of the geophysical and geochemical measurements. Such multi-source data providing a large pool of information is then processed to perform a geological analysis of the soil suitability, mapping study areas, plotting cross-sections, drawing 2D and 3D plans, re-projecting the data, modelling geological objects, performing calculations and computations, to mention a few. The mathematical modeling and plotting of correlations in physical and mechanical properties of soils against loads and climate conditions, are essential for planning and managing road constructions and solving other applied geotechnical and civil engineering problems.

#### 4.3. Novelty

The novelty of the presented research consists in a new technical approach of TeX, a highlevel, descriptive markup language applied specifically for the geological data processing and graphical visualization. Technical possibilities of the engineering graphics by TeX were demonstrated with a series of screenshots including both the codes used for plotting and a graphical output. The code snippets can be applied in a similar research. In the present study, and three-dimensional models of the equivalent soils cohesion (Figure 2, Figure 3 and Figure 4) were presented to evaluate the physical-mechanical properties of the soil behaviour under the conditions of load, compression on different frozen soil profiles for different load configurations and the immersion time of pressure.

A correlation between the physical and mechanical properties of soil affects the reliability analysis of the constructed roads, railways and buildings. Therefore, the actuality of the presented study consists in data analysis and effective graphical visualization for the assessment of the equivalent cohesion of soil and a compression strength that are important parameters required for evaluating of the stability and deformation of the geotechnical structures. The authors presented a novel approach of LaTeX packages for geological data visualization with a focus of the study on the equivalent cohesion estimation of soil sampling aimed to quantify the variation of soil reaction under varied loads with a representative number of tested data.

#### 4.4. Actuality

The actuality of the research consists in the importance of high-quality plotting and accurate graphs assisting engineering works in geological analysis. Methods of data analysis and visualization are diverse. therefore, introducing and testing new methods can significantly assist in progress, speed and quality of geological data processing. Incorrect data visualization and analysis may lead to a biased estimation of failure probability. On the contrary, the correct decisions in an engineering geology are largely based on the effective numerical methods that enable modeling of the soils compression. This is especially important in works on constructing buildings and roads. It is therefore crucial to properly quantify and visualize the 2D and 3D graphical plots between the soil characteristics for a geotechnical analysis and design. Thus, plotting the quality graphs and diagrams is a critical step in a geological data analysis.

Effective visualization of such large volumes of information, processed and visualized using the programming and/or mark-up languages enables to avoid significant labor and time costs and to increase the reliability of data processing. Nowadays, the accumulation and processing of the big data containing a multi-source geological information accelerated significantly. This necessarily increased the demand for testing of various methods and tools of data processing. Various approaches can satisfy high criteria of scientific data analysis: precise modelling; effective graphical visualization; embedded mathematical computations by special packages; high speed of data processing; correctness of computations. Given the variety of the existing methods of data processing in geology, current paper contributed by presenting the new approach of TeX language application.

#### 4.5. Recommendations

The functionalities of LaTeX demonstrated in this paper are by no means limited to the above examples and this approach. There are other approaches driven by the recent advances in the machine learning and data availability. Automatization and computer-based assistance in geosciences are rapidly developing and being continuously reported in the existing works using other applications that can be considered besides LaTeX: [47], [51], [64]. As a recommendation for future works it can be also noted that it is possible to visualize solid 3D surfaces by LaTeX, as well as perform a mathematical and statistical analysis of the larger amount of data in geotechnical surveys using special packages of LaTeX that can be recommended for future research.

#### 4.6. Acknowledgements

We cordially thank the two anonymous reviewers for their constructive comments that improved the initial manuscript. This research has been supported and implemented in the framework of the Project No. 0144-2019-0011, Schmidt Institute of Physics of the Earth, Russian Academy of Sciences.

#### References

- [1] Agterberg F. P., Statistical techniques for geological data, *Tectonophysics*, 1, 1964, 233–255.
- [2] Agterberg F. P., Geomathematics. Mathematical Background and Geo-Science Applications, Elsevier Science & Technology, 1974.
- [3] Aitchison G. D., The strength of quasi-saturated and unsaturated soils in relation to the pressure deficiency in the pore water, in: *Proceedings of the 4<sup>th</sup> International Conference on Soil Mechanics and Foundation Engineering*, 1957, 135–139.
- [4] Anderstand O. B., Lananyi B., Frozen Ground Engineering, Wiley, 2<sup>nd</sup> ed., 2003.
- [5] Babu G. L. S., Rao R. S., Peter J., Evaluation of shear strength functions based on soil water characteristic curves, *Journal of Testing and Evaluation*, 33, 2005, 461–465.
- [6] Baver L. D., Soil Physics, John Wiley & Sons, Inc., New York, U. S., 3rd ed., 1959.
- [7] Bezukhov N. I., *Fundamentals of the theory of elasticity, plasticity and creep*, Moscow, Vysshaya shkola publishing, 1968.
- [8] Boldyrev G. G., Idrisov I. K., Valeev D. N., Determination of parameters for soil models, *Soil Mechanics and Foundation Engineering*, 43, 2006, 101–108.
- [9] Bratten J. Creep Mechanics, Springer, Berlin, 2008.
- [10] Casagrande L., Loughney R. W., *Method of strengthening and stabilizing compressible soils*, United States Patent Office, **3**, 1966, 386.

- [11] Dahlin T., Svensson M., Lindh P., DC Resistivity and SASW for validation of effeiciency in soil stabilisation prior to road construction, in: *Procs. EEGS'99*, 1999.
- [12] Davis J. C., Statistics and Data Analysis in Geology, Wiley, 1986.
- [13] Davis J. C., Herzfeld U. C., Computers in Geology: 25 Years of Progress, Oxford University Press, 1993.
- [14] Driel J. N. V., Davis J. C., Digital Geologic and Geographic Information Systems, Short Course in Geology, 10. American Geophysical Union, 2<sup>nd</sup> ed., 1995.
- [15] Feuersänger C., Manual for Package pgfplots 2D/3D Plots in LaTeX, Version 1.5.1., 2011.
- [16] Feuersänger C., PgfplotsTable package Loading, rounding and formatting tables in LaTeX, 2011. Available as separate package usepackage{pgfplotstable}, as part of pgfplots.
- [17] Feuersänger C., *Programming in TeX and Library Functions from pgf and pgfplots*, 2011. Available as part of pgfplots.
- [18] Fredlund D. G., Rahardjo H., Soil Mechanics for Unsaturated Soils, John Wiley & Sons, Technology & Engineering, 1993.
- [19] Gan J. K.-M., Fredlund D. G., Rahardjo H., Determination of the shear strength parameters of an unsaturated soil using the direct shear test, *Canadian Geotechnical Journal*, 25, 1988, 500–510.
- [20] Gorokhovsky V. M., Tkachuk E. I., *Modeling in engineering geology*, NPI, Novocherkassk, 1980.
- [21] Jeffrey A., Lists in TeX's Mouth, TUGboat, 11, 1990, 237–245.
- [22] Källén, H., Heyden A., Lindh P. Estimation of grain size in asphalt samples using digital image analysis, in: *Proceedings of SPIE – The International Society for Optical Engineering*, 37, 2014, 921714–921714.
- [23] Källén H., Heyden A., Åström K., Lindh P., Measuring and evaluating bitumen coverage of stones using two different digital image analysis methods, *Measurement*, 84, 2016, 56–67.
- [24] Kaputin Y. E., Yezhov A. I., Henley S., Geostatistics in mining and geological practice, Kolskiy Scientific Center, Mining Institute RAS, Apatity, 1995.
- [25] Kern U, Extending LATEX's color facilities: the xcolor package, 2016.
- [26] Klaučo M., Gregorová B., Stankov U., Marković V., Lemenkova P., Determination of ecological significance based on geostatistical assessment: a case study from the Slovak Natura 2000 protected area, *Open Geosciences*, 5(1), 2013, 28–42.
- [27] Knuth D. E., *The Art of Computer Programming, chapter Fundamental Algorithms*, 3168. Addison-Wesley, U. S., 3<sup>rd</sup> ed., 1968.
- [28] Knuth D. E., Computers & Typesetting, chapter The TeXbook, Addison-Wesley, U. S., Reading, MA, 1984.

- [29] Knuth D. E., Computers & Typesetting, chapter TeX: The Program. Addison-Wesley, Reading, MA, 1986.
- [30] Knuth D. E., The New Versions of TeX and METAFONT, *TUGboat*, **10**, 1989, 325–328,. Reprinted as chapter 29 of Digital Typography.
- [31] Knuth D. E., The future of TeX and METAFONT, *Bijlage GG*, **90**, 1990, 145–145. Bijlage GG Reprint MAPS 5 (90.2).
- [32] Knuth D. E., *Selected Papers on Analysis of Algorithms. Lecture Notes*, Stanford, CA: Center for the Study of Language and Information CSLI, 2000.
- [33] Knuth D. E., *Selected Papers on Computer Languages. Lecture Notes*, Stanford, CA: Center for the Study of Language and Information CSLI, 2003.
- [34] Lemenkov V. A., The Project of Engineering and Geological Survey for the Development of Railway Infrastructure in Zabaykalsky Krai, Russian State Geological Prospecting University, Moscow, Russia, 2015. Licenciate Thesis.
- [35] Lemenkov V. A., Die Profiltypen der Richtbohren und Methoden Ihrer Gestaltung bei der Ingenieurgeologischen Erkundung, in: Actual Directions of Scientific Researches of the XXI Century: Theory and Practice, 4, 2016, 5–9.
- [36] Lemenkov V. A., Die regionale hydrogeologische Verh"altnisse und deren Ein- fluss auf die bautechnischen Eigenschaften, in: Actual Directions of Scientific Researches of the XXI Century: Theory and Practice, 4, 2016, 9–13.
- [37] Lemenkov V. A., Analysis of the Effects of the Mineral Soil Composition on the Cohesion Between its Structural Elements, in: *Modern Solutions to Scientific and Industrial Problems in Chemistry and Petrochemistry*, 2018, 617–625.
- [38] Lemenkov V. A., Computing Deflection and Compressibility of the Clay Loams for the Problem of the Construction Works in the Northern Tyumen (Bovanenkovo-Sabetta Geologic Cross Section), in: *Geography in the Modern World: Progress and New Priorities*, 2018, 506–509.
- [39] Lemenkov V. A., Deformation properties of the clay soil heave with a case study of sandy loam and clay by compression tests, in: *Development Strategy of the Geological Exploration of the Subsoils: Present and Future*, 2, 2018, 258–259.
- [40] Lemenkov V. A., Determination of correlation in deformation, strength and vis- cosity of the frozen soils through external loads by uniaxial compression, in: *Current Trends* and Innovations in Science and Industry, 2018, 64–65.
- [41] Lemenkov V. A., Laboratory tests of the different types of soils for compressibility by the compression and filtration device 'Odometer KFP-2-40 (60)', in: *Science, Education and Innovation in the Modern World*, **2**, 2018, 275–281.
- [42] Lemenkov V. A., Methods of the Determining Physico-Mechanical Parameters of the Frozen Ground Using Uniaxial Compression Strength, in: *Innovation of technical* solutions in mechanical engineering and transport, 2018, 202–206.

- [43] Lemenkov V. A., Variations in porosity and deformation in dehydrated loam samples, in: Development Strategy of the Geological Exploration of the Subsoils: Present and Future, 2, 2018, 256–257.
- [44] 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*, **2**, 2018, 35–42.
- [45] Lemenkova P., AWK and GNU Octave Programming Languages Integrated with Generic Mapping Tools for Geomorphological Analysis, *GeoScience Engineering*, **65**, 2019, 1–22.
- [46] Lemenkova P., Geospatial Analysis by Python and R: Geomorphology of the Philippine Trench, Pacific Ocean, *Electronic Letters on Science and Engineering*, 15, 2019, 81–94.
- [47] Lemenkova P., GMT Based Comparative Analysis and Geomorphological Mapping of the Kermadec and Tonga Trenches, Southwest Pacific Ocean, *Geographia Technica*, 14, 2019, 39–48.
- [48] Lemenkova P., Numerical Data Modelling and Classification in Marine Geology by the SPSS Statistics, *International Journal of Engineering Technologies*, 5, 2019, 90– 99.
- [49] Lemenkova P., Plotting Ternary Diagrams by R Library ggtern for Geological Modelling, *Eastern Anatolian Journal of Science*, **5**, 2019, 16–25.
- [50] Lemenkova P., Statistical Analysis of the Mariana Trench Geomorphology Using R Programming Language, *Geodesy and Cartography*, 45, 2019, 57–84.
- [51] Lemenkova P., Topographic surface modelling using raster grid datasets by GMT: example of the Kuril-Kamchatka Trench, Pacific Ocean, *Reports on Geodesy and Geoinformatics*, **108**, 2019, 9–22.
- [52] Leminen V., Tanninen P., Matthews S., and Niini A. The Effect of Heat Input on the Compression Strength and Durability of Press-formed Paperboard Trays. Procedia Manufacturing, 47:6–10, 2020.
- [53] Lindh P. Optimising binder blends for shallow stabilisation of fine-grained soils. In Proceedings of the Institution of Civil Engineers Ground Improvement, vol- ume 5, pages 23–34, 2001.
- [54] Lindh P., Compaction- and strength properties of stabilised and unstabilised finegrained tills, Lund University, Lund, Sweden, 2004. Doctoral Thesis.
- [55] Lindh P., Dahlin T., Svensson M., Comparisons Between Different Test Methods for Soil Stabilisation, in: *GeoEng 2000*, 2000, 1–7.
- [56] Lindh P., Winter M. G., Sample preparation effects on the compaction properties of Swedish fine-grained tills, *Quarterly Journal of Engineering Geology and Hydrogeology*, 36, 2003, 321–330.
- [57] Merriam D. F., Random Processes in Geology, Springer-Verlag Berlin Heidelberg, 1<sup>st</sup> ed., 1976.

- [58] Mills I., Cvitaš T., Homann K., Kallay N., Kuchitsu K., Quantities, Units and Symbols in Physical Chemistry, Blackwell Science, 2<sup>nd</sup> ed., 1996.
- [59] Pincus H. J., Statistical methods applied to the study of rock fractures, *Bulletin of the Geological Society of America*, 62, 1951, 81–130.
- [60] Porotov G. S., *Mathematical modeling methods in geology: Textbook*, SPbGGU (TU) Press, St. Petersburg, 2006.
- [61] Rathmayer H., Frost in Geotechnical Engineering, in: *International Symposium*, **2**. Espoo: VTT, 1989.
- [62] Ridley A. M., Burland J. B., A new instrument for the measurement of soil moisture suction, *Géotechnique*, **43**, 1993, 321–324.
- [63] Ryden N., Dahlen U., Lindh P., Jakobsson A., Impact non-linear reverber- ation spectroscopy applied to non-destructive testing of building materials, *The Journal of the Acoustical Society of America*, 140, 2016, 3327–3327.
- [64] Schenke H. W., Lemenkova P., Zur Frage der Meeresboden-Kartographie: Die Nutzung von AutoTrace Digitizer f
  ür die Vektorisierung der Bathymetrischen Daten in der Petschora-See, *Hydrographische Nachrichten*, 81, 2008, 16–21.
- [65] Tantau T., TikZ and pgf manual, Manual for Version 2.10-cvs, 2012.
- [66] Statens vegvesen, Håndbok 2016 Geoteknikk i vegbygging, Innholdsfortegnelse. Oslo: Trykk, 2010.
- [67] West D. R. F., Ternary Equilibrium Diagrams, Springer, 2<sup>nd</sup> ed., 1982.
- [68] West D. R. F., Ternary Phase Diagrams in Materials Science (Matsci), CRC Press, 3<sup>rd</sup> ed., 2017.
- [69] Yokoi H., Relationship between soil cohesion and shear strength, *Soil Science and Plant Nutrition*, **14**, 1968, 89–93.
- [70] Zhou S., Zhang S., Shen J., Guo W., Effect of cattle manure ash's particle size on compression strength of concrete, *Case Studies in Construction Materials*, 10, 2019, 1– 13.

Received 19.04.2020, Accepted 31.10.2020