PANAMGEO CHILE 2024
17th PAN-AMERICAN CONFERENCE ON SOIL MECHANICS AND GEOTECHNICAL ENGINEERING | 2nd LATIN-AMERICAN REGIONAL CONFERENCE OF IAEG
LA SERENA • CHILE / NOVEMBER 12 - 16, 2024

*Proceedings of the 17th Pan-American Conference on Soil Mechanics and Geotechnical Engineering (XVII PCSMGE), and 2nd Latin-American Regional Conference of the International Association for Engineering Geology and the Environment (IAEG), La Serena Chile, 2024.*

# GEOREC: New techniques for processing georeferenced map data

## GEOREC: Nuevas técnicas de procesamiento de datos georreferenciados

Ernesto Cáceres, Camilo Mejías, **Elías Godoy**, Milene Gutiérrez & Roberto Ávila

*Data and research department, Hibring, Chile, egodoy@hibring.com*

ABSTRACT: Chile hosts one of the world's largest mining sectors, encompassing a wide range of geological resources, including significant copper mining operations. Geological mapping generates an extensive volume of data in various formats, requiring considerable time investment for data standardization. We have developed an innovative geotechnical software that presents a distinct approach to underground and open pit mining data analysis and processing, streamlining the transcription of substantial data volumes. Engineered by a multidisciplinary team with a focus on algorithmic efficiency, reliability and robustness, this software utilizes georeferenced data directly obtained from digitized underground mine drawings, being an alternative to AI techniques. Successfully tested in Chile across four mines, GEOREC provides significant time savings and prevents human-induced errors by automating various processes, managing to extract hundreds of joints and faults within seconds, providing visualization tools and reports of such data. This adaptable and user-friendly tool can be used through a platform, allowing users to analyze and process mining data online, digitizing historical drawings and making data available as an input for other software.

KEYWORDS: Geotechnical software, mining data analysis, underground mining, data standardization, georeferenced data.

## 1 INTRODUCTION.

Chile's mining production reaches figures of over 5300 kilotons of fine copper content, positioning the country as a key producer in the global mining scenario. The current challenge in the mining industry resides in competitiveness and sustainability (Bartos, 2007; SERNAGEOMIN, 2022). Small and medium-scale mining is important for the development of the mining industry in Chile, especially in copper exploitation. This part of Chilean's mining accounted for more than 265 kilotons in 2021, translating into figures exceeding 2.8 billion dollars (COCHILCO, 2022). It is in this segment where, thanks to its flexibility, there is room to address problems through innovation, allowing for the piloting, testing, and validation of technologies; that require research and development (R&D).

Mining industry plays a primary role in training experts and developing new technologies; however, its focus is not on R&D, and it is still distant from that of the high-tech industry. Innovation is crucial in facing more complex exploitation scenarios and aligning with process improvement in mining. Process efficiency and cost reduction have been the prevailing trends in recent decades. This is where Knowledge Intensive Business Services (KIBS) companies play a fundamental role in advancing solutions. The main objective of these companies is to meet the needs of the international market in line with the task of improving processes, integrating and adding value to knowledge. (Bartos, 2007; Sanchez & Hertlieb, 2020; Urzúa et al., 2013).

Digital transformation (DT) resonates throughout the mining industry due to the sustained increase in demand for metals. Current trends indicate that Industry 4.0 is moving towards automation and remote operations, the Internet of Things (IoT), and the use of artificial intelligence (AI), digital twin simulation, among other technologies (Barnewold & Lottermoser, 2020; Sanchez & Hertlieb, 2020). In the efforts to keep pace with digital transformation as a present reality, it is necessary to progress in the digitization of geological information from the source. This requires a collective effort to seek relevant improvements from various fronts, integrating interdisciplinary and transdisciplinary tools, from undergraduate and postgraduate to advanced and applied research.

Within the capabilities of small and medium-scale mining, the use of software for both geological and geotechnical resources is notable and widely known in the mining industry. In the last two decades, medium-scale mining has digitized a significant part of its processes, among them the mapping of geological structures. Among several tools, the use of georeferenced data in GIS and CAD formats stands out, along with other geological and mining modeling software to create detailed 3D models of mineral deposits and underground structures, which are dominant in the market. Geological-geotechnical mapping has been adapting to the type of input required by these types of software. Therefore, systematizing data is crucial for developing robust models that enable a better understanding of mining operations despite limitations in fieldwork. This includes information from the past that is not necessarily digitized.

Particularly in underground operations of small and medium-scale mining, there is a shortage of digital tools that streamline geological mapping; it is common for the initial approach to excavation geology to be carried out with paper and pencil in the field. The collection of structural data is done in written form on plans with underground cartography, or an approximation of this, which is later digitized using software that best suits the working team. This method of gathering information from each tunnel and mine access leads to having extensive databases that do not have an identical and precise drawing format given their transfer from paper to computer. However, they are created following international standard drawing patterns. Extracting important information from these databases becomes complex when done manually into a tabulated format such as .csv or .xls, as it can take several minutes or even

hours before conducting more detailed analyses through specialized software for the mining industry.

This manuscript focuses on the development of new techniques that enable the recognition and post-processing of geological-geotechnical information and its extraction from drawing formats and vectorial data such as CAD, allowing its availability in a user-manageable database compatible with other software's formats.

## 2 MATERIAL AND METHODS

### 2.1 *Input data*

#### 2.1.1 *Underground cartography*

For each tunnel section to be mapped, the topography used is previously known and is brought on paper to the mapping site. Tunnels and accesses are represented by lines, polylines, and polygons throughout each level of the underground operation. The quantity of topographic data depends on the method of topographic information gathering.

This georeferenced data is arranged in 3D lines and polylines that construct both the walls and the floor and ceiling of the tunnel. The amount of available data varies based on the field-mapped and CAD-digitized information, ranging from a complete lack of georeferenced data to having thousands of three-dimensional data points for certain meters of the tunnel. In Figure 1, an example of a tunnel section with mapped geological structures is observed. The lines defining the tunnel were assigned arbitrarily and solely as an example.

#### 2.1.2 *Geological data*

The drawing format adheres to the standard established by the U.S. Geological Survey (FGDC, 2006); however, these drawings are executed based on the individual criteria of the person digitizing the field-mapped data and therefore are not exactly the same. The size of each symbol varies depending on how each geological structure is arranged in the tunnel and the scale of the drawing; nevertheless, their azimuth and dip are easily identifiable. The graphical representation of geological-geotechnical mapping lacks exact dimensions or replicated locations. In Figure 1, a simple example of geological mapping in a tunnel section is observed, where fracture frequency is represented within circles, while thickness is depicted within rectangles.

All these data do not account for their elevation-depth positioning or may contain imprecision in the data due to manual and generalized assignment for large tunnel segments due to the complexity of positioning each mapped structure three-dimensionally in the context of small and medium-scale mining. Additionally, it is possible to find extra information on joints and structures of greater interest, such as their fracture frequency and thickness.

### 2.2 *Data Extraction*

Joint sets and faults are represented in drawings by three abstract objects: line segments, circles and texts. These are differentiated through line colors, thickness and length. Basic information about joint sets and faults, such as polyline coordinates, width and color, texts insertion points, color and content, circles radii, color and center, and hatches coordinates and filling color, are extracted from drawings. In general terms, Figure 2 shows a referential diagram of the entire process in question. In the case of CAD drawings, data is extracted through AutoLISP scripts that return objects in a plain text file.

### 2.3 *Data Recognition*

Once a plain text file containing extracted structure information is generated, objects are reassembled. For instance, following the drawing standard outlined in 2.1.1 (Figure 1), a joint is depicted using two parallel line segments, each with varying width, which can be utilized as a parameter for recognition.

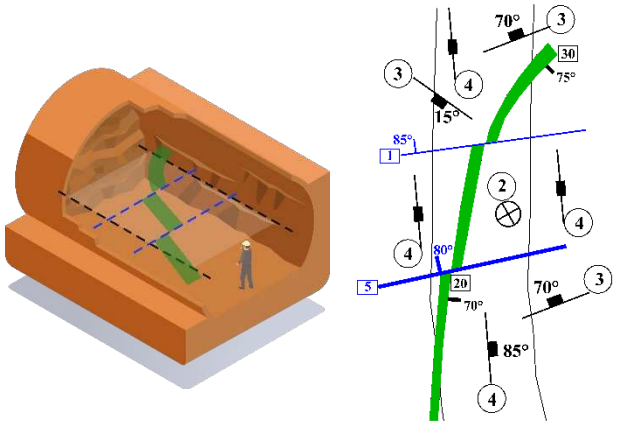


Figure 1. An example of a tunnel mapping and its drawing as a mid-level, which includes joints, a bedding plane (both in black), faults (blue), and a dike (green). Tunnel section was adapted from the design by macrovector - Freepik.

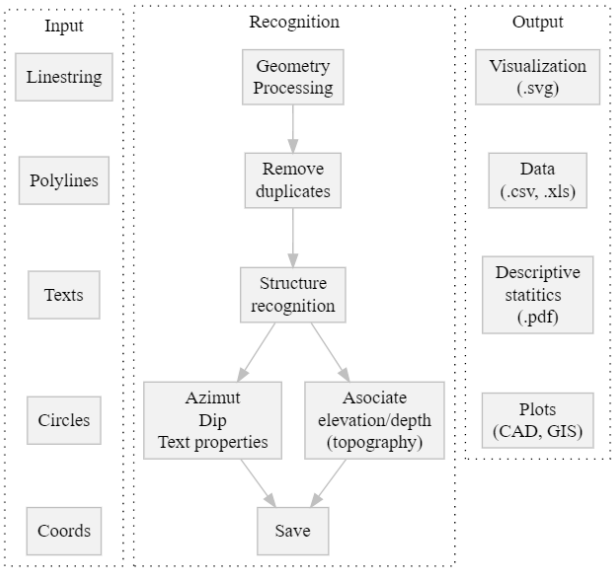


Figure 2. Data is first prepared in a format that makes it easy to read in Python. This input is then processed by the core algorithm so that structures are re-glued together. Finally, tables with the automatized computations are returned.

An orthogonal distance between the thick line and the narrow line, as well as the angle formed between them, uniquely determine joint sets. Their fracture frequencies are usually depicted by circles that contain numbers. Distances between texts' insertion points and circles' centers uniquely determine fracture frequencies, and the minimum distance between circles' centers and joint sets endpoints allow us to associate fracture frequencies to their joint sets. Dips, whose magnitudes are depicted by texts, are associated to their joint sets considering the distance between their insertion points and the joint's thick line segment. Another important data recognition feature is the assignation of structures' altitudes. This is done by projecting structures and tunnels' three-dimensional measurements onto the xy-cartesian plane and by using proximity functions.

### 2.4 *Data Processing*

Recognized and re-glued structural data that was obtained using the procedures described in 2.3 is now ready to process. Some calculations, which are quite tedious to do by hand, are possible to do using computational tools such as Python (Van Rossum & Drake, 2009). For each joint and fracture and considering all possible cases regarding the right-hand rule (Marjoribanks, 2010), azimuth and dip direction (*dipdir*) are computed. Elevation information, which is of great interest, was already assigned in 2.3. Having access to structures' information such as dip and elevation gives rise to two applications that we detail next.

#### 2.4.1 *Trend and plunge*

The orientation of each tunnel is determined by the direction of the vector generated from two points in the drawing, which define the endpoints of each tunnel (start and end of the section). With the association from the underground cartography using a distance norm and since these points are tridimensional, it is possible to compute both trend and plunge of each structure that belongs to the selected tunnel. If the selected line segment that defines a portion of a tunnel has direction vector $v = (\Delta x, \Delta y, \Delta z)$, then its trend $\alpha_t$ is given by the formula:

$$\alpha_t = \text{atan2}(\Delta x, \Delta y) \tag{1}$$

where $atan2(y, x)$ yields the angle measure, in radians, between the x-axis and the ray from the origin to the point $(x, y)$. Since the plunge is the angle between said ray and the y-axis, arguments of the $atan2$ function are swapped. The plunge $\beta_t$ is given by:

$$\beta_t = \text{atan2}(\Delta z, \sqrt{\Delta x^2 + \Delta y^2}) \tag{2}$$

#### 2.4.2 *Terzaghi Weighting*

The Terzaghi Weighting was proposed by Ruth Terzaghi (1965) to correct a directional bias generated by samplings fractures, linear or planar, for data collected from rock surfaces or scanlines. This factor was implemented by following the formula:

$$W = \frac{1}{\sin(\delta)} \tag{3}$$

where W is the weighting factor and δ is the angle between the discontinuity and the uplift direction. This geometric weighting tends to infinity as δ approaches zero, so a limit of 10° is stablished for the angles (Priest, 1993; Wang & Mauldon, 2006), that is:

$$W = \frac{1}{\sin(\delta^*)}, \quad \text{where} \quad \delta^* = max\left\{\delta, \frac{\pi}{18}\right\} \tag{4}$$

and the angle δ is calculated as proposed by:

$$\cos(\delta) = |\cos(\alpha_n - \alpha_s)\cos(\beta_n)\cos(\beta_s) + \sin(\beta_n)\sin(\beta_s)| \tag{5}$$

where $\alpha_s$ is the trend of the selected tunnel and $\beta_s$ its plunge, and for a given discontinuity, $\alpha_n$ and $\beta_n$ are, respectively, the trend and plunge of a line that is normal to the mapped plane, which are obtained as follows:

$$\alpha_n = \alpha_d \pm 180 \tag{6}$$

$$\beta_n = 90 - \beta_d \tag{7}$$

where $\alpha_d$ is the dip direction and $\beta_n$ is the dip of the discontinuity. In formula (5), addition or subtraction of 180 must be chosen so that $\alpha_n$ lies between 0 and 360 degrees.

### 2.5 *Outputs*

Recognized and re-glued structural data that was obtained using the procedures described in sections 2.2, 2.3 and 2.4 is prepared at this point to be outputted in different ways, such as Excel tables that include the automatized calculations, database tables that are connected through a web access platform, and interactive graphs that allows the user to interpret specific information about discontinuities in an easier way. Additionally, it allows the user to obtain descriptive statistics in PDF format, as well as plots of pole density and rose diagrams for each tunnel, considering trend, plunge, and Terzaghi Weighting computations as described in sections 2.4.1 and 2.4.2. Furthermore, these plots are available in grid format such as CAD or GIS to display georeferenced structural data.

### 2.6 *Limits of applicability*

To achieve high accuracy rates using our software, some considerations regarding the drawing and the input must be taken. The preferred method for categorizing objects drew in a software is to filter them by layer name. Otherwise, alternative filters must be used, such as filtering by line color, width, length, type, etc.

Inputs for algorithms used to re-glue objects together are parametric. This means that to recognize all desired objects we need to give good parameters as input to the program. For example, the average distance between a joint set endpoint and the text that is associated to its fracture frequency is a parameter of interest which can be found by examining the drawing and giving a rough estimate. This allows the program to be optimal. If said parameters are not provided, past uses of the software on dozens of drawing files give default input parameters.

### 2.7 *Testing methods*

To test algorithms, we run experiments on a desktop computer running Windows with an Intel Core i7-12700F processor at 2.10 GHz using 32 GB of RAM.

Data from the CAD drawing used for experiments was extracted using AutoLISP scripts. Drawings and diagrams of this work were made using Python 3.11.2 and Numpy 1.24.2. For visualization of tabulated data, we used Matplotlib 3.7.0 and Microsoft Excel 2023. Data tabulation as figures for this work were created using LaTeX language.

Two experiments were performed, the first one shows the Excel and image output of an CAD drawing of Figure 1, showing how accurate the object recognition algorithm is. The second experiment regards algorithm efficiency and will show how long the software takes in processing objects with respect to the amount of input objects.

To apply the method to a case study, we considered data from mining tunnel drawings sourced from Candelaria Cooper Mining Complex by Lundin Mining. Among these drawings, the number of joints and faults was manually counted, and then the information from each structure was extracted using GEOREC to evaluate the efficiency of the algorithm in recognizing geological-geotechnical information.

## 3 RESULTS

### 3.1 *Method validation*

For testing the algorithm, we represent the scan depicted in Figure 1 in a drawing software. A script that exports information about structures in a plain text file is run. This information includes line coordinates, insertion points for texts and circles, color, among others. A portion of this input file is shown below.

```
Name: Example
Date: yyyymmdd.hhmmss
2D Lines: 32
LINE2D | Dikes | 76.596 | 263.197 | 77.820 | 263.198 | 0 | 255 | 255 | 255
LINE2D | Dikes | 77.820 | 263.198 | 77.820 | 262.183 | 0 | 255 | 255 | 255
LINE2D | Faults | 85.318 | 272.529 | 73.332 | 270.906 | 0.1 | 0 | 0 | 255
LINE2D | Faults | 74.206 | 271.965 | 74.328 | 271.053 | 0.1 | 0 | 0 | 255
LINE2D | Joints | 75.441 | 274.634 | 76.237 | 274.069 | 0.6 | 255 | 255 | 255
LINE2D | Joints | 74.201 | 275.872 | 77.948 | 273.232 | 0 | 255 | 255 | 255
...
Texts: 20
TEXT | Dikes | 76.6727 | 262.3933 | 20 | 255 | 255 | 255
TEXT | Dikes | 83.607 | 277.1383 | 30 | 255 | 255 | 255
TEXT | Joints | 84.4196 | 266.1674 | 4 | 255 | 255 | 255
TEXT | Joints | 73.4215 | 264.1777 | 4 | 255 | 255 | 255
TEXT | Faults | 71.0055 | 262.129 | 5 | 0 | 0 | 255
TEXT | Faults | 72.4292 | 270.5419 | 1 | 0 | 0 | 255
...
Circles: 9
CIRCLE | Joints | 84.8159 | 266.5792 | 0.9 | 255 | 255 | 255
CIRCLE | Joints | 73.8178 | 264.5894 | 0.9 | 255 | 255 | 255
CIRCLE | Joints | 77.7421 | 275.4387 | 0.9 | 255 | 255 | 255
CIRCLE | Joints | 79.2226 | 253.8918 | 0.9 | 255 | 255 | 255
CIRCLE | Joints | 86.0842 | 262.9487 | 0.9 | 255 | 255 | 255
...
Hatches: 3
HATCH | Dikes | 82.861 278.136 82.060 277.283 80.915 275.858 82.163 276.059 83.477
HATCH | Dikes | 82.163 276.059 80.915 275.858 80.052 274.528 80.963 274.598 82.163
HATCH | Dikes | 80.085 273.288 80.963 274.598 80.052 274.528 79.437 273.348 79.088
```

Figure 3. Plain text file containing basic structure information.

Our software reads the input file and performs a series of things to it, being the most relevant the following:

- It reads the file and organizes their objects between lines, circles, texts and hatches.
- Duplicate objects are removed and a single copy of each is kept.
- Scripts for recognizing objects are run and objects are re-glued together. Computations for finding objects' elevation are performed.
- Recognized objects are prepared to be exported in different formats such as SVG for drawings and Excel for information summaries.

Figure 4 shows the SVG output generated by GEOREC. Comparing this Figure with Figure 1, we see that all objects and their attributes coincide. Furthermore, green texts in Figure 4 show a correspondence between objects and attributes. Sometimes, for joint sets, the intersection between the line that characterizes each of them and the tunnel gallery (if any) is of interest, as a joint lying outside a tunnel has no physical sense. This intersection is represented by a red point.

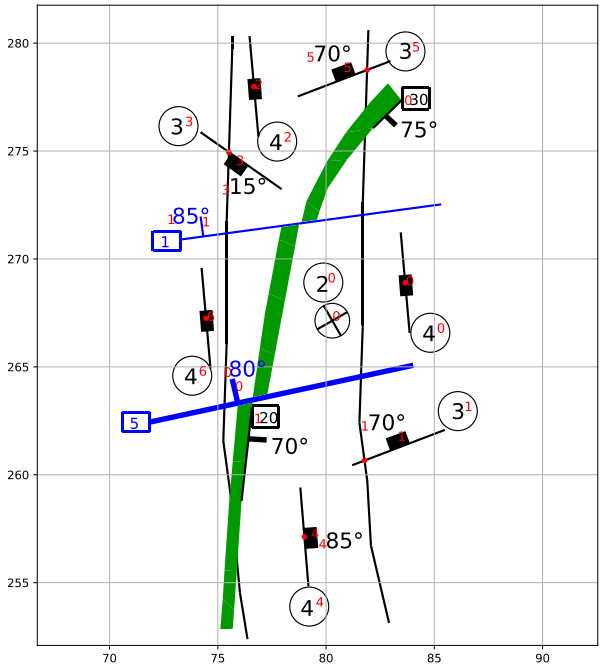


Figure 4. SVG output of GEOREC. Red texts correspond to object indices and red points to line segments intersections (if any).

Excel outputs summarizing the information obtained are provided by GEOREC. In the case of faults, indices, endpoints, dip, azimuth, dip direction, thickness and elevation are included in the spreadsheet. This is shown in Figure 5.

| | $x_1$ | $y_1$ | $x_2$ | $y_2$ | dip | azimuth | dipdir | width | elevation |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 84.02 | 265.07 | 71.89 | 262.44 | 80 | 257.79 | 347.79 | 5 | 136.35 |
| 2 | 85.31 | 272.52 | 73.33 | 270.90 | 85 | 262.28 | 352.28 | 1 | 135.93 |

Figure 5. Excel output with information about faults obtained from the drawing shown in Figure 1, including computed azimuth, dip direction, and elevation.

| | $x_1$ | $y_1$ | $x_2$ | $y_2$ | $x_c$ | $y_c$ | dip | azimuth | dipdir | frequency | elevation |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 83.86 | 266.56 | 83.44 | 271.23 | 83.65 | 268.89 | 90 | 354.95 | 84.95 | 4 | 135.90 |
| 2 | 85.48 | 262.07 | 81.20 | 260.44 | 81.77 | 260.65 | 70 | 249.07 | 339.07 | 3 | 136.63 |
| 3 | 76.87 | 275.66 | 76.46 | 280.32 | 76.67 | 277.99 | 90 | 354.95 | 84.95 | 4 | 135.55 |
| 4 | 74.20 | 275.87 | 77.94 | 273.23 | 75.51 | 274.94 | 15 | 125.16 | 215.16 | 3 | 135.72 |
| 5 | 79.18 | 254.84 | 78.81 | 259.41 | 78.99 | 257.12 | 85 | 355.31 | 85.31 | 4 | 136.80 |
| 6 | 82.97 | 279.17 | 78.69 | 277.53 | 81.89 | 278.75 | 70 | 249.07 | 339.07 | 3 | 135.64 |
| 7 | 74.66 | 264.92 | 74.25 | 269.58 | 74.45 | 267.25 | 90 | 354.95 | 84.95 | 4 | 135.90 |

Figure 6. Excel output with information about joint sets obtained from the drawing showed in Figure 1, and computed azimuth, dip direction, and elevation, as well as the point $(x_c, y_c)$ showed in red in Figure 4 at which, if true, the line representing the joint set intersects the gallery.

For joint sets, apart from segment endpoints, the point $(x_c, y_c)$ at which the line intersects the tunnel gallery, and that is represented by red points in Figure 4 is provided. If no intersection is found, then the base line barycenter is given instead. Other information such as dip, azimuth, dip direction, fracture frequency and elevation are also given. Figure 6 shows an Excel output sheet for joint set. Bedding planes or joint sets that have nearly zero dip are classified in a different category. In this case, and following Figure 4, there is only one such joint set as bedding-plane. The barycenter, fracture frequency and elevation are of interest, as shown in Figure 7.

| | $x$ | $y$ | frequency | elevation |
|---|---|---|---|---|
| 1 | 80.28 | 267.13 | 2 | 136.25 |

Figure 7. Excel output with information about horizontal joint sets obtained from the drawing shown in Figure 1 and plotted in Figure 4, and computed elevation.

In the case of dikes, we see from Figure 4 that there are three green hatches separated by two faults. These represent dikes. However, it is common to find in drawings that each hatch is a union of smaller hatches. Sometimes, faults are associated to dikes as well. Figure 8 shows information about the dikes in Figure 4. In this case the point $(x_b, y_b)$ corresponds to the barycenter of each sub-dike, and the dip and thickness (whenever there is a fault associated to a hatch) and elevation are given.

| | $x_b$ | $y_b$ | dip | width | elevation |
|---|---|---|---|---|---|
| 1 | 82.29 | 276.93 | 75 | 30 | 135.68 |
| 2 | 81.02 | 275.26 | | | 135.71 |
| 3 | 79.88 | 273.43 | | | 135.74 |
| 4 | 79.30 | 272.09 | | | 135.84 |
| 5 | 78.36 | 271.22 | | | 135.80 |
| 6 | 77.95 | 269.48 | | | 135.89 |
| 7 | 77.52 | 267.13 | | | 135.93 |
| 8 | 77.11 | 264.69 | | | 136.14 |
| 9 | 75.80 | 258.33 | | | 136.73 |
| 10 | 75.54 | 254.77 | | | 136.88 |
| 11 | 75.44 | 253.41 | | | 136.92 |
| 12 | 76.16 | 262.32 | 70 | 20 | 136.45 |

Figure 8. Excel output with information about dikes obtained from the drawing showed in Figure 1 and plotted in Figure 4, and computed elevation.

In an operational environment, GEOREC can be integrated into a web platform, and a database can be established to store data originating from a variety of drawings. This stored data can be accessed at any time to generate reports and quantify variables. For instance, Trend and Plunge, as detailed in Section 2.4.1, can be utilized to produce structural information reports. Similarly, Terzaghi weighting, outlined in Section 2.4.2, can be employed to generate a report containing rose diagram and pole diagrams plots.

Saved data can be seen in a plot and can also be filtered out for an easier reading of information. Additionally, a history of data entries at different elevations and mines can be obtained, as well as users who have previously entered data.

### 3.2 *Algorithmic complexity*

In this section, we test running times of GEOREC. To do so, we replicate the drawing in Figure 1 many times. We start with the original drawing containing 22 objects (8 joint sets, 12 polygons constructing a segmented dike in 3 parts, and 4 faults), along with a set of elevation point measures, and duplicate it in the following way: The second drawing contains twice the number of elements from the original drawing, the third contains 4 times the number of elements from the original drawing, and the n-th contains $2^{n-1}$ times the number of elements from the original drawing. We measure the total time that the software takes in reading the file, removing duplicates, re-gluing objects back and generating the excel file and the output figure, and summarize all of these in a table.

Table 1. Summary of number of objects vs execution time.

| Number of objects | Time (s) |
|---|---|
| 352 | 0.06113 |
| 704 | 0.09219 |
| 1408 | 0.1489 |
| 2816 | 0.2805 |
| 5632 | 0.5674 |
| 11264 | 1.202 |
| 22528 | 2.834 |
| 45056 | 7.454 |
| 90112 | 22.36 |

From Table 1, we see that the algorithm's running time increases quasi-linearly with respect to the number of objects. This goes in correspondence with the fact that the algorithm for measuring distances between different objects is not quadratic but quasi-linear.

In a real case scenario, and from past applications of the software to mine drawings, the number of objects does not exceed 15000. Since the software takes 1.202 seconds in processing 11264 objects, this translates in very significant time savings and error avoidance with respect to manual processing of the data by a human worker.

### 3.3 *Case of application*

We tested the algorithm with data from drawings compatible with CADs formats provided by Candelaria Copper Mining Complex by Lundin Mining, corresponding to geological data from Candelaria Underground mine. In general, as mentioned before, these drawings have no more than 15,000 objects of interest, as big file sizes may bring complications in its daily manipulation. Considering this data, we tested our software to quantify its efficiency. We manually counted each joint and fault to compare the total data drawn versus data recognized by GEOREC. Global results of this comparation are shown in Table 2.

Table 2. Efficiency of data recognition in a level of Candelaria Mine.

| | *Joints* | *Faults* |
|---|---|---|
| *Manually counted* | 260 | 150 |
| *Recognized by GEOREC* | 243 | 124 |
| *Efficiency (%)* | 93,46 | 82,67 |

In total, 2930 objects were counted with the algorithm. The difference in the recognition of joints and faults is due to the way they were drawn. The general input parameters chosen to recognize each symbol differ from what was found in the drawing, as the lines may have a width outside of the ideal or parameterized range, as well as being more distant or in a different position than ideal or being arranged in layers with a different name than initially proposed, as mentioned in Section 2.6. The recognition of symbols is favorably higher than 80% of the total drawn.

The execution time for this case study was only a few seconds. Consequently, calculating trend and plunge and taking the Terzaghi correction into account involves minimal computational cost, making the running time negligible. Plotting pole density and rose diagrams also does not imply extra time beyond a few seconds.

## 4 CONCLUSIONS

The proposed methodology for extracting, recognizing, and processing data arranged in drawings is applicable to the context of geological-geotechnical mapping in small and medium-scale mining tunnels. By mixing programming languages such as AutoLISP and Python, an efficient recognition tool was developed, while providing an important advantage to optimize calculation times. The integration of these tools facilitated a more comprehensive analysis of the data, enhancing the overall efficiency of the process.

Considering the time consuming and error inducing tasks of entering data and computing quantities such as azimuth, dip direction and elevation, GEOREC is a good alternative for addressing these issues. Since structural information can be digitized very quickly using our software, it is feasible to implement more thorough post-processing data analysis to generate reports almost instantly. In this way, our software has potential avenues to contribute to time savings. Regarding the efficiency in recognizing joints and faults in Auto CAD drawings of tunnels from the Candelaria mine, as mentioned in Section 2.6, GEOREC is limited to the recognition criteria used as input. For the case study, there is some data that falls outside the assigned parameters. This can be due to the distances between segments that compose a symbol, the length and thickness of each line or the dimensions of the symbols, leading to an incomplete recognition of the geological-geotechnical information present on the map.

In terms of software limitations, the authors believe that accurate information recognition is explicitly dependent on the data input format, particularly on the drawer's ability to follow the standard suggested by the U.S. Geological Survey (FGDC, 2006), as the deterministic algorithms used to identify each symbol are adjusted based on this standard. To recognize other symbols, it is necessary to adjust or expand the searching and recognition criteria. Additionally, information post-processing is limited to the methodologies integrated into the software, so it is feasible to perform other types of analysis after some extra development.

Finally, and analogously to the proposed approach, deterministic criteria can be derived for recognizing other symbols indicating relevant geological information. With this, the application of these techniques is feasible in open-pit operations if topographic and geological information is available, as described in Section 2.1. This also makes it feasible to integrate the extraction and recognition of geological-geotechnical information provided by geotechnical classification systems, such as, global stress index or GSI (Hoek & Marinos, 2000; Hoek & Brown, 2018), rock mass rating or RMR (Bieniawski, 1989) and its update by M. Romana (2014), and Q (Grimstad & Barton, 1993; Barton, 2000). Furthermore, criteria can be incorporated to recognize symbols related to ore geology, and detailed three-dimensional topography, to complement and enhance the robustness and data availability.

## 5 LIST OF ABBREVIATIONS

| | |
|---|---|
| AI | Artificial Intelligence |
| CAD | Computer-Aided Design |
| CSV | Comma-Separated Values |
| dipdir | Dip direction |
| FGDC | Federal Geographic Data Committee |
| GIS | Geographic Information System |
| IoT | Internet of Things |
| KIBS | Knowledge Intensive Business Services |
| R&D | Research and Development |
| SVG | Scalable Vector Graphics |

## 6 ACKNOWLEDGEMENTS

As a multidisciplinary team, the authors would like to thank Pablo Escares and Mario Guerrero, geotechnical leaders, for their trust in us to tackle this challenge of geological-geotechnical data and its implementation at Candelaria Copper Mining Complex by Lundin Mining. Additionally, we are deeply grateful for the constant feedback from Lucía Díaz, geologist at Candelaria mine, in the software development process. We appreciate Javiera Torres, graphic designer from Hibring, who kindly provided improvements in the illustrations for this work; also, Víctor Gatica and Guillermo Santana, programmers from Hibring, for their continuous assistance.

## 7 REFERENCES

Barnewold, L., & Lottermoser, B. G. (2020). Identification of digital technologies and digitalisation trends in the mining industry. International journal of mining science and technology, 30(6), 747-757. https://doi.org/10.1016/j.ijmst.2020.07.003

Barton, N. R. (2000). TBM tunnelling in jointed and faulted rock. Crc Press.

Bartos, P. J. (2007). Is mining a high-tech industry?: Investigations into innovation and productivity advance. Resources Policy, 32(4), 149-158. https://doi.org/10.1016/j.resourpol.2007.07.001

Bieniawski Z. T. (1989). Engineering Rock Mass Classifications. A Complete Manual for Engineers and Geologists in Mining, Civil and petroleum Engineers. Wiley, New York.

Bravo-Ortega, C., & Muñoz, L. (2021). Mining services suppliers in Chile: A regional approach (or lack of it) for their development. Resources Policy, 70, 101210. https://doi.org/10.1016/j.resourpol.2018.06.001

COCHILCO. 2022. Monitoreo de variables e indicadores relevantes de la mediana y pequeña minería chilena. Dirección de Estudios y Políticas Públicas, Comisión Chilena del Cobre, Ministerio de Mineria. Nº 2022-A-10575.

Federal Geographic Data Committee [FGDC] [prepared for the Federal Geographic Data Committee by the U.S. Geological Survey], 2006, FGDC Digital Cartographic Standard for Geologic Map Symbolization: Reston, Va., Federal Geographic Data Committee Document Number FGDCSTD-013-2006, 290 p., 2 plates, accessed December 27, 2023, at https://ngmdb.usgs.gov/fgdc_gds/.

Grimstad E. & Barton N. (1993). "Updating of the Q-System for NMT", International Symposium on Sprayed Condrete – Modern Use of Wet Mix Sprayed Concrete for Underground Support, Fagernes

Hoek E. & Marinos P. G. (2000). "Predicting tunnel squeezing problems in weak heterogeneous rock masses". Tunnels and Tunnelling International;132(11):45-51.

Hoek, E., & Brown, E. T. (2019). The Hoek–Brown failure criterion and GSI–2018 edition. Journal of Rock Mechanics and Geotechnical Engineering, 11(3), 445-463. https://doi.org/10.1016/j.jrmge.2018.08.001

Marjoribanks, R. (2010). Geological methods in mineral exploration and mining. Springer Science & Business Media. https://doi=10.1007/978-3-540-74375-0

Priest S. D. (1993) Discontinuity analysis for rock engineering, 1st edn. Springer Science & Business Media, Netherlands, Dordrecht. https://doi.org/10.1007/978-94-011-1498-1

Romana, M. (2014). RMR New Guidelines for Tunnels. In Proceedings of the 13th ISRM International Congress of Rock Mechanics, Montreal, QC, Canada, 10–13 May 2015.

Sánchez, F., & Hartlieb, P. (2020). Innovation in the mining industry: Technological trends and a case study of the challenges of disruptive innovation. Mining, Metallurgy & Exploration, 37(5), 1385-1399. https://doi.org/10.1007/s42461-020-00262-1

SERNAGEOMIN, S. (2022). Anuario de la mineria de Chile 2022. Servicio Nacional de Geología y Minería. Available at: https://www.sernageomin.cl/anuario-de-la-mineria-de-chile/

Terzaghi, R. (1965) Sources of Error in Joint Surveys. Géotechnique, 15, 287-304. Available at: https://doi.org/10.1680/geot.1965.15.3.287

Urzúa, O. (2013). The emergence and development of knowledge intensive mining service suppliers in the late 20th century (Doctoral dissertation, University of Sussex). https://doi.org/10.13140/RG.2.1.4802.0328

Wang, X., & Mauldon, M. (2006). Proportional errors of the Terzaghi correction factors. Proceedings of the 41st US. Rock Mechanics Symposium–ARMA's Golden Rocks 2006-50 Years of Rock Mechanics (pp. ARMA-06). https://onepetro.org/ARMAUSRMS/proceedings-abstract/ARMA06/All-ARMA06/ARMA-06-1055/115968