BeFo

STIFTELSEN BERGTEKNISK FORSKNING
ROCK ENGINEERING RESEARCH FOUNDATION

AUTOMATED MWD DATA PROCESSING AND UNIFIED DATABASE BUILDING

Chunling Shan
Abbas Abbaszadeh Shahri
Stefan Larsson
Fredrik Johansson

BeFo Report 242

STIFTELSEN BERGTEKNISK FORSKNING
ROCK ENGINEERING RESEARCH FOUNDATION

AUTOMATED MWD DATA PROCESSING AND UNIFIED DATABASE BUILDING

Automatiserad MWD databehandling och uppbyggnad av en enhetlig databas

Chunling Shan

Abbas Abbaszadeh Shahri

Stefan Larsson

Fredrik Johansson

BeFo Report 242
Stockholm 2024
ISSN 1104-1773
ISRN BEFO-R-242-SE

PREFACE

This report presents a generalized and automated tool for big amount of MWD data filtering and normalization, and an idea of applying artificial intelligence (AI) on MWD data for predicting bedrock quality conditions and designing appropriate grouting systems.

Reference group members who provided valuable comments and suggestions was composed of Patrik Vidstrand, Johan Funehag, Mahmoud Yazadani, Thomas Dalmalm, Karl-Johan Loorents, Johan Spross, Almir Draganovic, Catrin Edelbro and Jeroen van Eldert. The project was funded by BeFo and Tyréns.

Stockholm

Patrik Vidstrand

FÖRORD

Denna rapport presenterar ett generaliserat och automatiserat verktyg för datafiltrering och normalisering av stora mängder MWD data, samt presenterar en idé om att tillämpa artificiell intelligens (AI) på MWD-data för att förutsäga berggrundskvalitetsförhållanden och stödja injekteringsdesign.

Referensgruppsmedlemmar som gav värdefulla kommentarer och förslag bestod av Patrik Vidstrand, Johan Funehag, Mahmoud Yazadani, Thomas Dalmalm, Karl-Johan Loorents, Johan Spross, Almir Draganovic, Catrin Edelbro och Jeroen van Eldert. Projektet finansierades av BeFo och Tyréns.

Stockholm

Patrik Vidstrand

SUMMARY

Continuous forecasting of the ground conditions ahead of the tunnel face during construction projects like tunneling is of great interest. Nowadays, in tunneling projects, analyzing the acquired measurement-while-drilling (MWD) data has been demonstrated to be a helpful method for assessing the rock mass conditions. However, the MWD is a monitoring technique that provides significant large records and analysis of a large amount of generated data not only is time-consuming.

The purpose of this work is to provide a generalized and automated tool for large amounts of MWD data filtering and normalization in order to help in more appropriate grouting design systems, at the mean time to investigate the possibilities of applying artificial intelligence (AI) on MWD data for predicting bedrock quality conditions and designing appropriate grouting systems. One test dataset containing MWD, real-time grouting data, and field protocols from Stockholm Bypass project was investigated and analyzed within this project. An automated process for filtering and normalizing MWD data using a combination of Mode and Percentile gate bands was developed for efficient removal of the noisy data caused by rig components, i.e. collaring and coupling effects from rod extensions. The presented approach in the form of single hole and peer group-based methods were developed compared, and automated to evaluate the applicability of the normalizing methods in removing the hole depth dependencies of MWD data.

It is concluded that the hole-based normalization method can more accurately remove the hole depth dependencies and stepwise problems in MWD data. A relational PostgreSQL database for storing the MWD and real-time grouting data was built to provide cost-effective and efficient tools for data extraction. This tool automatically can transfer raw MWD and grouting data to the database. The entire process from reading the data, filtering, normalizing, and then building the data center was automated using the Python programming platform. For future research, an idea for the use of AI by utilizing the MWD and grouting data was proposed. As summary, this idea aims to combine different types of data for predictions of the success of the grouting design. The presented automated MWD data processing and unified database building in combination with the realization of the idea of using AI in MWD data analysis is considered to be a very valuable daily tool for people involved in their decision-making process.

Keywords: Big data, Measurement while drilling (MWD), normalizing index, filtering process

SAMMANFATTNING

I byggprojekt innehållande tunneldrivning är kontinuerlig undersökning och prognos av markförhållandena framför tunnel drivningsfronten är av stort intresse. Numera har analys av MWD data i tunnelprojekt visat sig vara en hjälpsam metod för att bedöma bergmassans förhållanden. MWD är en övervakningsteknik som analyserar stora mängder data, vilket för närvarande är mycket tidskrävande utan för närvarande och som dessutom inte används fullt ut i det dagliga arbetet.

Syftet med detta arbete är att tillhandahålla ett generaliserat och automatiserat verktyg för filtrering och normalisering av stora mängder MWD data, samtidigt som man undersöker möjligheterna att tillämpa artificiell intelligens (AI) på MWD data för att förutsäga bergkvalitetsförhållanden och utforma lämpliga injekteringsdesign.

En testdatauppsättning innehållande MWD data, realtidsdata för injektering och fältprotokoll från projekt Förbifart Stockholm har undersökts och analyseradts inom ramen för detta projekt. En automatiserad process för filtrering och normalisering av MWD-data med hjälp av en kombination av mode och percentil filter utvecklades för effektiv borttagning av icke relevant data orsakad av borrutrustning, dvs. effekter av kragning och koppling från stångförlängningar. De presenterade metoderna för så kallad enskild-hålbaserad och grupp-hålbaserad utvecklades, jämfördes och automatiserades för att utvärdera användbarheten hos normaliseringsmetoderna samt för att ta bort håldjupberoenden hos MWD data.

Slutsatsen är att den enskild hålbaserad metoden kan mer noggrant ta bort håldjupberoenden och stegvisa problem i MWD data. En relationell PostgreSQL-databas för lagring av MWD och realtidsinjektering data byggdes upp för att tillhandahålla kostnadseffektiva och effektiva verktyg för datautvinning. Detta verktyg kan automatiskt överföra rå MWD och injekteringsdata till databasen. Hela processen från att läsa data, filtrera, normalisera och sedan bygga upp databasen automatiserades med hjälp av programmeringsplattformen Python. För framtida forskning föreslogs en idé om att använda AI genom att utnyttja MWD och injekteringsdata. Sammanfattningsvis syftar denna idé till att kombinera olika typer av data för att förutsäga framgången för injekteringsdesignen. Den presenterade automatiserade bearbetningen av MWD data och uppbyggnaden av en enhetlig databas i kombination med förverkligandet av idén att använda AI i MWD-dataanalyser anses vara ett mycket värdefullt dagligt verktyg för personer som är involverade i deras beslutsprocess.

Nyckelord: Big data, Measurement while drilling, normalisering och filtrering

CONTENTS

1. INTRODUCTION

Measurement while drilling (MWD) is a monitoring technology (Gearhart et al., 1986), that can provide valuable technical information on the penetrated rock mass, for rock engineering projects i.e. excavations and stability assessments, characterization of geology (e.g. lithology, porosity and permeability), fluid content and physical properties of the rocks being drilled. (Schunnesson, 1996 and 1998). This implies the ability of MWD records in providing more efficient design and making decisions based on real-time data in direct response to measured changes in the rock mass.

Technologically, the immediacy and relative cheapness of data acquisition using the available sensors on the drilling rig, are the attractiveness of MWD system. In any MWD system the monitored drill parameters are either stored on-board or directly transferred to a storage device for further processing. Depending on the type of drilling, different parameters are monitored. In the case assessed in this report a face drill rig equipped with top hammer rock drills are used, where the monitored parameters are thrust, feed pressure, percussion pressure, rotation speed, penetration rate, rotation pressure, flushing pressure, flushing flow, drilling depth and time are measured (Schunnesson, 1998).

Currently, MWD data monitoring are required and successfully organized in many infrastructure projects in several countries like Sweden (Martinsson and Bengtsson, 2010; van Eldert et al., 2020; Isheyskiy and Sanchidrian, 2020), USA (Rostami et al., 2015), Norway (Nilsen, 2015; Hansen et al., 2022), Spain (Navarro et al., 2018), Canada (Khorzoughi et al., 2018; Khorzougi and Hall, 2016), and Russia (Isheyskiy et al., 2021). More information about the industrial companies with significant contributions in development and interpreting the MWD data can be found in Appendix A.

The MWD data is a typical representation of complex big data in geoengineering applications. Moreover big data analysis in geo-modelling problems has just started in the geoengineering field (Zhang et al.,2021).Therefore, conducting a unified MWD database in this study technically dedicates an essential tool for future research:

- a structured framework for data integrity to reduce redundancy,
- deeper insights and more physical meaningful interpretation on the retrieved information,
- improving data management, i.e. a centralized location with accessible share space during research stages,
- providing a consistent and cost-effective data analysis platform for auditing and optimization across an entire process using data mining and artificial intelligence approaches to get more detailed information on subsurface conditions.

As a result, the unified database facilitates collaborations between engineers with different roles in projects for better communication and promoting more efficient workflows (e.g. Tsatalos and Ioannidis, 1994; Zhussupbekov et al., 2021; Ishaq et al., 2023). Such analysis will then greatly help the geoengineers to identify patterns and trends, data anomalies and thus error elimination to get more informed decision-making and operational improvement (Isheyskiy and Sanchidrian, 2020; van Eldert et al., 2021a; van Eldert, et al., 2021b).

The gathered MWD data falls within the big data category. Accordingly, the noise in such metadata need to be appropriate removed (filtering), the data need to be scaled for consistent interpretation (normalizing) and then centralized in storing location (unified database). This means that the data can then be retrieved quickly for various analyses. Therefore, technically development of an automated procedure covering the filtering, normalizing and the creation of an integrated unified meta database is highly motivated in geoengineering applications. Such processing paradigm is greatly beneficial because:

- ensures for consistency in MWD data analysis via a standardized and understandable format by the various professions,
- improving the accuracy of the MWD data by eliminating errors or biases that may be introduced by the rig monitoring systems or the drilling environment,
- gaining new insights into the drilling process.

Based on the motivations, a further developed automated filtering and normalization approach based on van Eldert et al. (2020) for MWD data is presented using the mode and percentile filters and linear regression normalizing techniques. As an illustration for practical use, the capacity of the suggested procedure was examined by data processing on acquired data from two rigs from different drilling environment from infrastructure project Stockholm Bypass in Stockholm, Sweden.

2. BRIEF REVIEW OF MWD DATA PROCESSING TECHNIQUES

As referenced by Smith (2002), the use of MWD as a drill monitoring technique in different geoengineering applications has been well recognized since the 1970s. Standardization of data formats (Saunders et al., 1996), data integration (Ziegler and Dittrich, 2007), data cleansing (Wu, 2013), metadata management (Chapman et al., 2009), cloud-based solutions (Alreshidi et al., 2018), and application programming interfaces (*APIs*) (Imieliński et al., 1999) are the most common used techniques for processing and managing a centralized MWD database in geoengineering. The International Real-time Data Exchange Standard (IREDES) can facilitate the real-time exchange of drilling data between different systems ensuring interoperability and seamless communication (Cayeux et al., 2019). However, relying solely on a single standard, such as IREDES may limit flexibility in choosing the best technologies and data exchange methods for specific drilling projects (Geekiyanage et al., 2021). These methods, overall aim to define a consistent form of MWD data processing that can be integrated and shared across different systems and platforms. Figure 1 shows the increasing trend of the geoengineering application of MWD data in recent years.

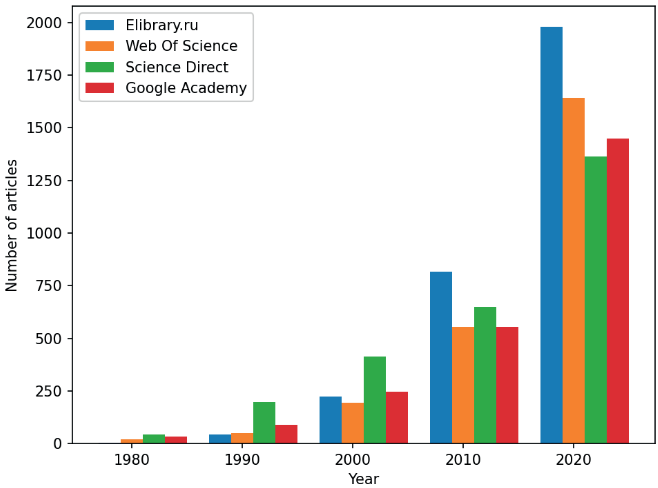

Figure 1. Increasing trend of using MWD data in geoengineering applications in the last five decades (after Isheyskiy and Sanchidrian, 2020)

Despite the fact that the awareness has increased significantly, development of accurate processing of these types of recorded data is still needed The data process contains different types of recorded data that can be categorized as independent (e.g. feed force, rotation speed and air pressure) which are directly controlled and influenced by the rig control system and the operator and then the dependent (e.g. penetration rate, rotation torque, vibration) that represent the mechanical responses altered by the geological and mechanical properties of the rock (e.g. Brown and Barr, 1978; Peck, 1989; Schunnesson and Kristoffersson, 2011).

2.1 Filtration techniques

Since the MWD data is typically acquired by embedded sensors installed directly on the drill rig and can provide information about the conditions at the drill bit, the presence of signals recorded due to various factors, like the drilling environment/ condition, tool wear, and signal interference is significant and cannot be neglected (Navarro 2018; Charoenprasit et al., 2018; Geekiyanage et al., 2021) These factors can not only make it challenging to extract accurate and reliable information from the MWD data but can also introduce operator caused influence on the data. Therefore, filtering of MWD data is a critical task because it fundamentally helps to 1) remove data that is not rock mass dependent, 2) identify potential issues and abnormalities in the drilling process (e.g. tool wear, wellbore instability, drilling-induced vibrations), 3) improve the signal-to-noise ratio to provide higher resolution to detect and interpret trends and patterns in the data. Subsequently, filtration assists in extracting more useful information from data influenced by varying rock mass conditions, control systems and operators to improve the accuracy of the analyses of the results.

Mathematically, the MWD data can be filtered using different techniques such as bandpass (Zhao et al., 2009), moving average (Geekiyanage et al., 2021), Kalman (Yang et al., 2020), and wavelet (Arabjamaloei et al., 2011). However, the choice of filtering technique has a close dependency on the geological conditions and the specific drilling operation for the MWD data (Zhao et al., 2023). Therefore, in the first step, a careful evaluation of the filtering methods and their impact on the data should be executed.

2.2 Normalizing techniques

Normalizing is used to adjust or scale the datasets to a standard or reference condition to eliminate the effects of variations in drilling circumstances, measurement equipment, and other factors that can affect the data. Since the MWD parameters have different units of measurement, the normalization aims to obtain comparable scales of criteria values. The MWD data can be normalized using different methods via various parameters like depth normalizing (e.g. Navarro et al., 2018; van Eldert et al., 2020), time normalizing (e.g. Eren and Ozbayoglu, 2010; Abdelaal et al., 2022.; Leung and Scheding, 2015), lithology normalizing (Deng et al., 2022), mud weight normalizing (Aljubran et al., 2021), tool normalizing (e.g. Ertunc et al., 2001; Rodgers et al., 2020), environmental normalizing (e.g. Purkayastha and Nair, 2017; van Eldert et al., 2020; Isheyskiy and Sanchidrian, 2020), and statistical normalizing (e.g. Basarir et al., 2017; Ghosh et al., 2014). There is thus available normalization techniques for the parameters that are registered. In this study, an improved normalization method was developed based on the linear regression method used in von Eldert et al. (2020).

3. DATASET

In this project, MWD data from one part of Stockholm Bypass project close to Lunda entitled FSE410 were analyzed. The acquired datasets from the MWD at FSE410 also include the real time grouting supplemented by implemented protocols i.e. drilling plans and water flow measurements. These supplementary datasets can further be used to investigate the potential development of modern artificial intelligent (AI)-based modeling approaches for detailed analyses of the MWD parameters and grouting design. A proposed development is presented and discussed in more detail in Section 8.

The employed MWD datasets and their units follow a matrix from. The columns show the measured parameters including hole depth (*HD*, mm), penetration rate (*PR*, dm/min), percussive pressure (*HP*, bar), feed pressure (*FP*, bar), damping pressure (*DP*, bar), rotation speed (*RS*, r/min), rotation pressure (*RP*, bar), water flow (*WF*, l/min) and water pressure (*WP*, bar) and the time of operation (hh:mm:ss) and the rows present the corresponding measured values of each recorded interval.

4. APPLIED METHODOLOGY

4.1 Filtering procedure

In this work the presented flow diagram in Figure 2 was followed. The procedure includes several inner nested loops and mimic an automated process for MWD data processing, where the input data are automatically filtered and normalized and then transferred to the centralized space to store and create an unified database. Block A shows the process to adopt the multi-filtration procedure while Block B expresses the implemented framework for the normalization step. The process was designed in such a way that it covers both hole and peer group-based analyses. In the hole-based procedure, the single MWD data (single hole) is fed, while the peer group is referred to a set of MWD records that are grouped based on analytically relevant criterion (criteria). In the case of MWD data, these criteria are the diameter and hole depth that are related to the rod length. Furthermore, the created database was developed so it can be updated by several possibilities like adding the raw data and utilities for the user in retrieve and extract appropriate information. This process was entirely coded in Python.

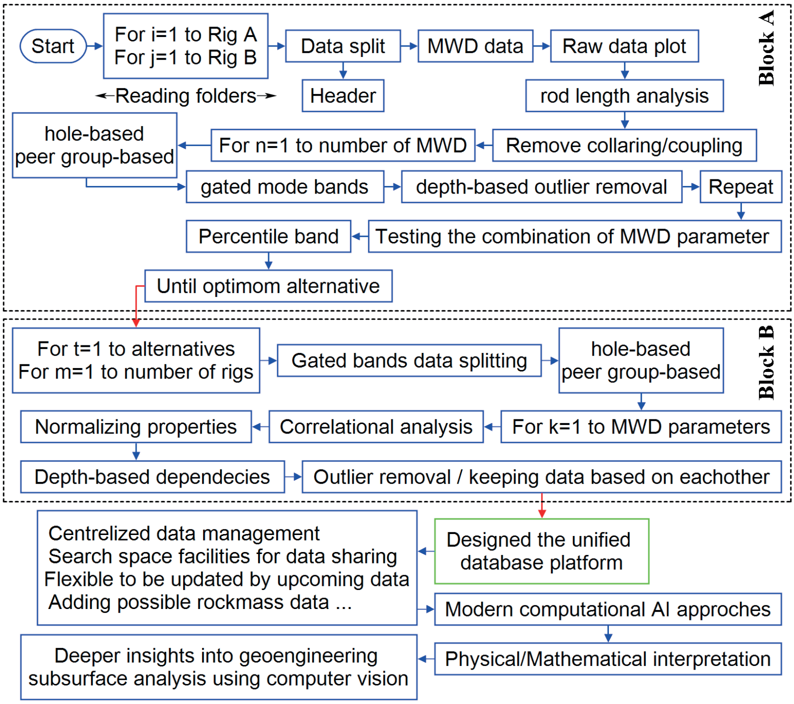

Figure 2. Simplified diagram of applied automated MWD processing procedure and generating unified database

Based on block A (Figure 2), at the first step, the rod length using the raw data was analyzed to ensure the dissimilarities (Figure 3). The results showed different lengths in drilling sequences (e.g. ≈5 m at the first sequence and then ≈ 3 m long in further steps). This issue means that the drill rod length should be dealt with as a variable instead of a constant during the filtration process (peer group criterion).

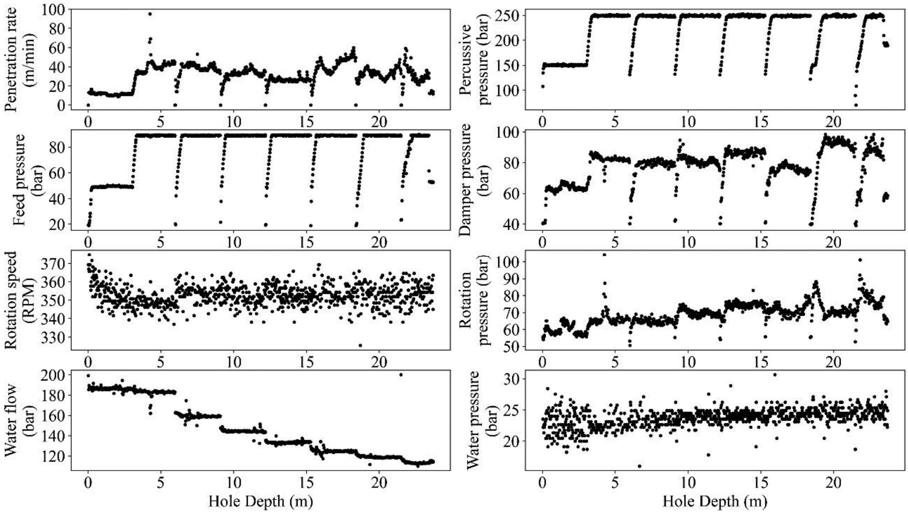

Figure 3. A sample plot of the raw MWD data to check the rod length in different drilling sequences

As presented in Figure 2, a dynamic multi gated band filtering procedure was applied. Through this process, the most appropriate combination of MWD parameters i.e. *PR*, *HP, DP, HP-FP, DP-HP*, *HP-DP-FP*, and etc. are identified to split the entire of data into different gated bands based on the mode and long term average statistics. The designed bands then are supplemented by percentile filter. In the current project the combination of *HP-PR* showed the most optimum results and thus selected as the MWD parameters to define the gated band and percentile filters. The gated bands were applied on *HP* and the percentile was applied on *PR* to remove the outliers. It should be noted that the filtering procedure simultaneously is applied on the other MWD parameters, i.e. removing one data from *HP* means eliminating the entire row of data for all parameters. Referring to the definition of mode filter, the gated bands can be used to split the data into three modes, including high/change/low modes (Figures 4 and 5). The filter band for high mode values i.e. the 'High pressure mode' data (purple dots in Figures 4 and 5) was designed as a combination of mode and average to cover the max of *HP*. Therefore, in this condition the gated band was designed as an interval around max *HP* value and programmed in terms of [max *HP*-15, max *HP*]. This band was obtained from merged *HP* data (peer group analysis) of drilled holes.

To capture the low mode i.e. ‘Low pressure’ data (red dots in Figures 4 and 5), the gated band in terms of [mode value-5, mode value+5] was designed and implemented. Accordingly, the rest of data falls outside of the upper/lower gated bands were then classified as ‘Change mode’ (black dots in Figures 4 and 5) and were considered as noisy data from machine operations e.g. coupling and collaring and the dependency on the rod length for each sequence of drilling is excluded in further analysis (van Eldert et al., 2020). In the current project the given recommendation by van Eldert et al., 2020 (removing a value of 0.5 m from both sides of each rod) was not applicable due to variation of the rod length in drilling sequences. Accordingly, in the peer group analysis after merging the split data i.e. High/Low-pressure values, a Percentile filter was applied to remove noise or outliers. In this work, only lower noisy PR data were removed. The high values were sustained for further analysis to avoid removing important data points collected from poor quality rock mass.

Like any measurement system, MWD tools are not perfect and may have inherent measurement errors. Therefore, identify and handling outliers in MWD data is crucial for maintaining the accuracy of drilling operations and making informed decisions. In the current project some of the outliers i.e. data points that deviate significantly from the overall pattern or trend of the dataset, in MWD data after filtering and normalizing are still remained. Possible causes for outliers in MWD data:

- Tool wear: During the drilling sequences, the tools undergo wear and thus degrading. Consequently the sensors may provide less accurate records, leading to outliers in the data. This implies to consider the tool reliability through combination of new designs, maintenance procedures and operating practices (Martin et al., 1994).
- Formation heterogeneity: The subsurface variability, i.e. changes in rock formations, presence of fractures, and etc., can result unexpected records of MWD parameters, causing outliers in the data (Fernández et al., 2023) and should be kept during filtration process.
- Tool interactions and drilling dynamics: since the drilling rig consists of different tools, they can interact with each other in complex ways(drill string, bit, and the subsurface) and thus introduce noise or anomalies in the MWD parameters leading to outliers (Reckmann et al., 2010).
- Drilling Depth: Understanding the potential impact of drilling depth on MWD data is essential for interpreting the measurements accurately. The increase in hydrostatic pressure with drilling depth can affect the performance of downhole sensors and therefore may impact the accuracy of measurements, potentially resulting in outliers (Ertunc et al., 2001; Rodgers et al., 2020). On the other hand, the problem of vibration and shock also should be considered. Deeper drilling often involves more challenging subsurface conditions, including harder rocks. Increased vibration and shock loads on the drilling tools can influence the reliability of sensors and leading to monitoring the outliers (Song et al., 2022). These signals should be adjusted and normalized before further analysis.

- Data transmission challenges: Since the MWD data is often transmitted in real-time from downhole sensors to the surface, errors or signal interference can lead to corrupted data, resulting in outliers. Data transition from greater depths also poses additional challenges, i.e. the longer drill strings cause more signal attenuation and data transmission delays, or potential signal loss, that contributes to outliers in the received data. Furthermore, the influence of operational-worker errors in recording also should be considered (van Eldert et al, 2020).

As a result of peer group analysis, a visualized filtering result from one fan in terms of rod length is presented in Figure 6, i.e. the split data from 'rod 1' into High/Low pressure modes for the depth interval of 0-6 m.

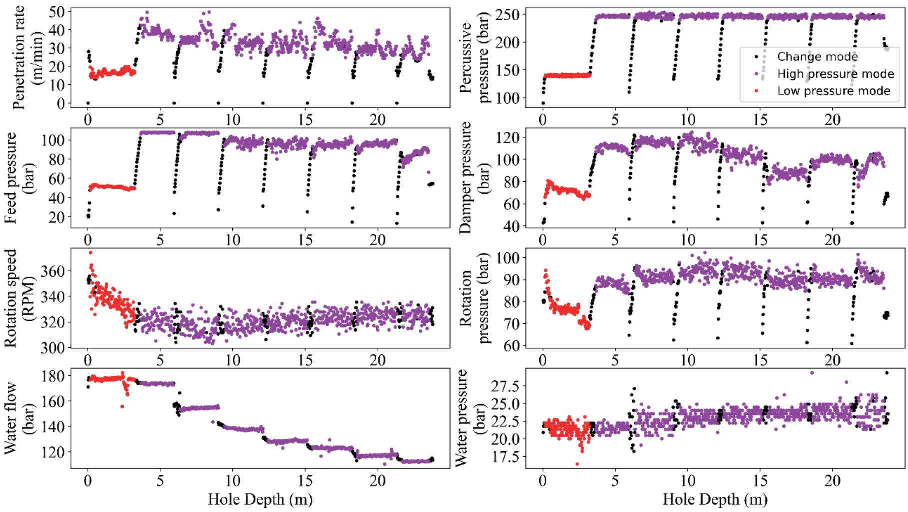

Figure 4. A visualized sample of carried out efforts for hole-based data filtering. Black dots represent Change mode, purple dots represent High pressure mode and red dots represent Low pressure mode.

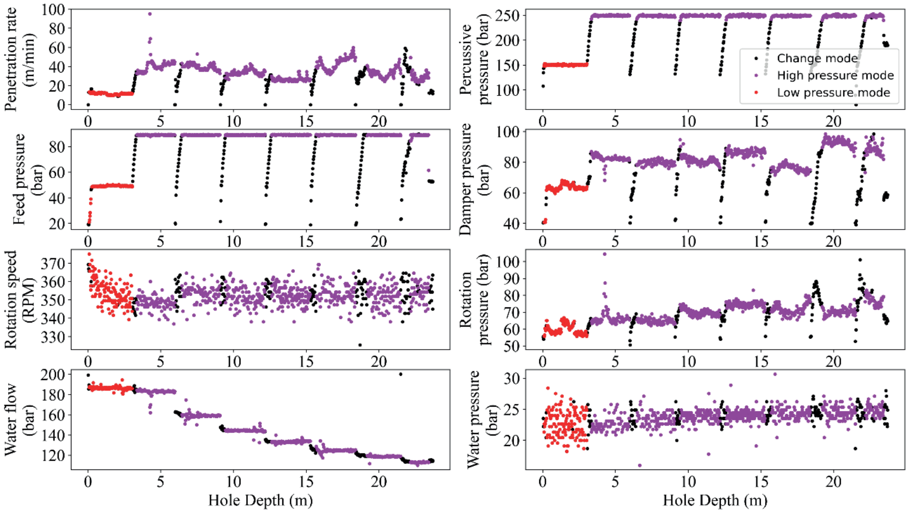

Figure 5. Visualized results of filtering procedure based on gated bands and modes of the MWD data in accordance to *HP*

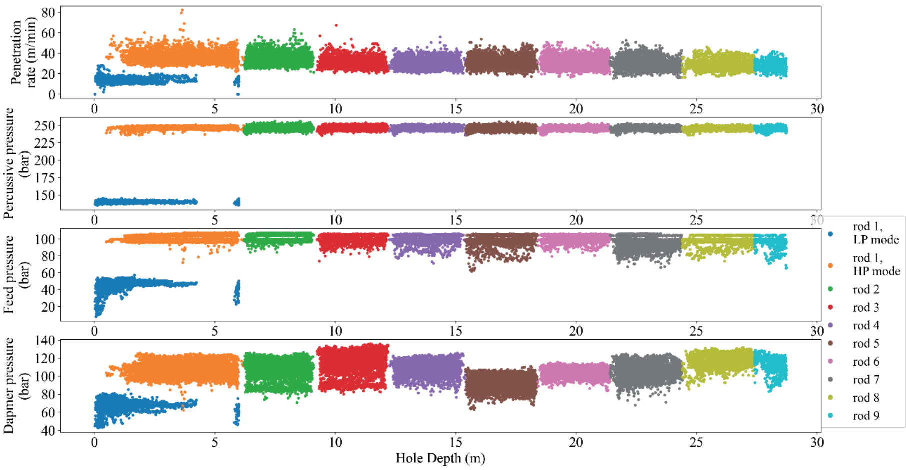

Figure 6. Rod length checking through splitting of merged data for all rods from one umbrella (Checking the mode capability in splitting the high (orange) and low (dark blue) pressure values for rod 1)

4.2 Normalizing process

The normalization process in MWD data aims to adjust and scale the data to a consistent reference or baseline. This process is used to remove variations in the data caused by differences in rig type, drilling setups in the machines and other factors, allowing for more accurate analysis and interpretation of the data. Accuracy improvement, providing comparable conditions, sensitivity analysis and more visualized insights are some of the potential benefits of normalizing MWD data, because:

- Normalization after filtering can assist in removing signals caused by systematic drilling behaviors that may affect the accuracy of the data.
- As it aims to scale the data, then easier comparison among different datasets particularly data from different rigs/drilling conditions is allowed.
- It can highlight smaller changes that may be masked by larger variations in the data.
- Adjusted/scaled data can improve the visual clarity of the data, making it easier to identify trends and patterns.

To execute the depth-normalization, Block B in Figure 2 in the case of both hole and peer group data was followed. The summary of applied procedure in the form of pseudo-code can be found in Appendix B.

The result of hole and peer group-based depth-normalization (single hole and fan holes) in terms of raw data (black dots), normalized data after removing the hole depth dependency (green dots) and adopted regressions for each rod length (red lines) are presented in Figures 7 and 8. The adopted regressions based on peer group data (Figure 8) were conducted for each parameter from all rods with respect to High/Low modes (Figure 6), i.e. two different regression models for 'rod 1'. Referring to Figures 7 and 8, both hole and peer group-based results showed the stepwise problem (energy losses in the couplings for the rod extension) in Feed pressure and Damper pressure at depth ≥ 15 m, where the hole-based normalizing can provide more effective stepwise removal than peer group analysis. However, the low correlation of Rotation pressure (Figure 7) prevented appropriate depth-normalizing, and thereby, the stepwise problem for depth ≥ 15 m is not treated like Feed pressure and Damper pressure. An overview of the compared methods, i.e. hole/peer group-based depth-normalization is shown in Figure 9 that indicates the improper stepwise removal through peer group analysis in Rotation speed around 20 m. Such heterogeneity mechanically can be assigned to the drilled rock mass characteristics which induced uncertainties in the records where the peer group considers all of the holes instead of single data in the hole-based approach. Both the single hole and fan based normalization processes were done in a similar way for all holes from all different sites to make it possible to compare different site conditions.

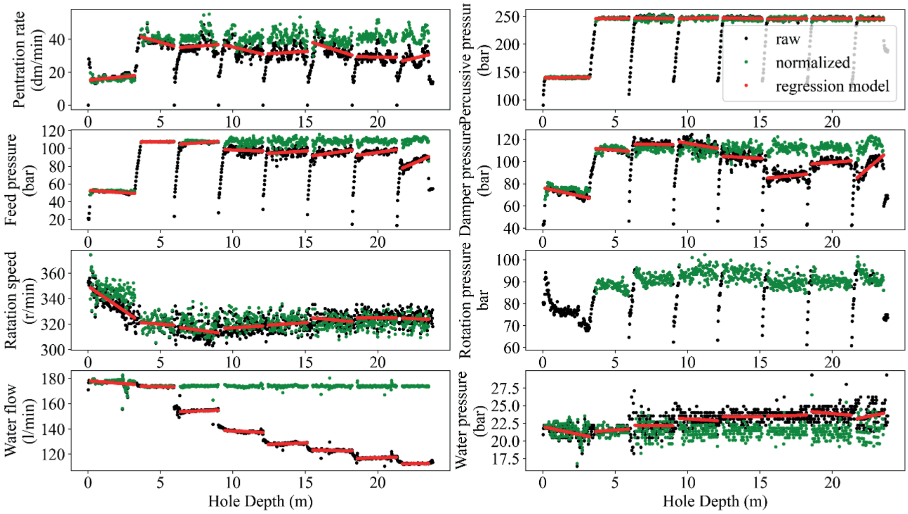

Figure 2. Pattern identification and trend analysis between the normalized and un-normalized MWD data (hole-based). Raw data (black dots), normalized data after removing the hole depth dependency (green dots) and adopted regressions for each rod length (red lines).

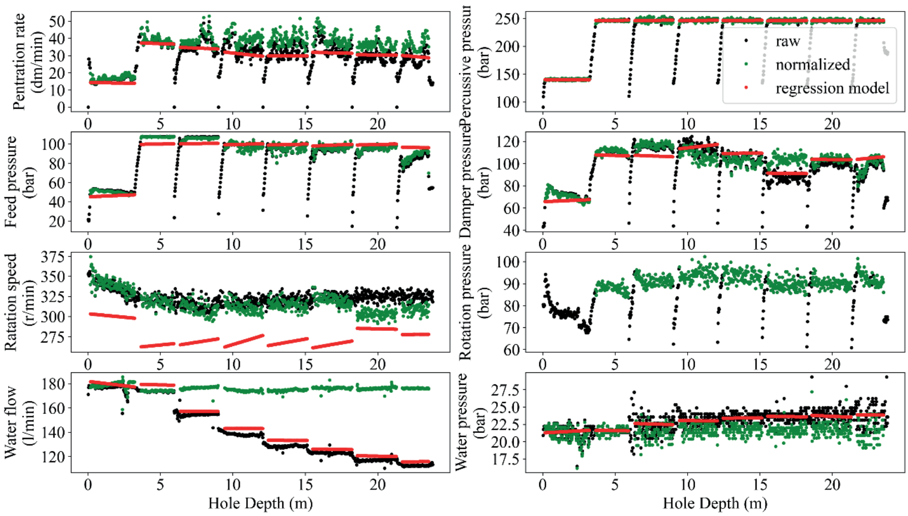

Figure 8. A visualize sample of pattern identification-trend analysis between the normalized and un-normalized MWD data (peer group-based). Raw data (black dots), normalized data after removing the hole depth dependency (green dots) and adopted regressions for each rod length (red lines).

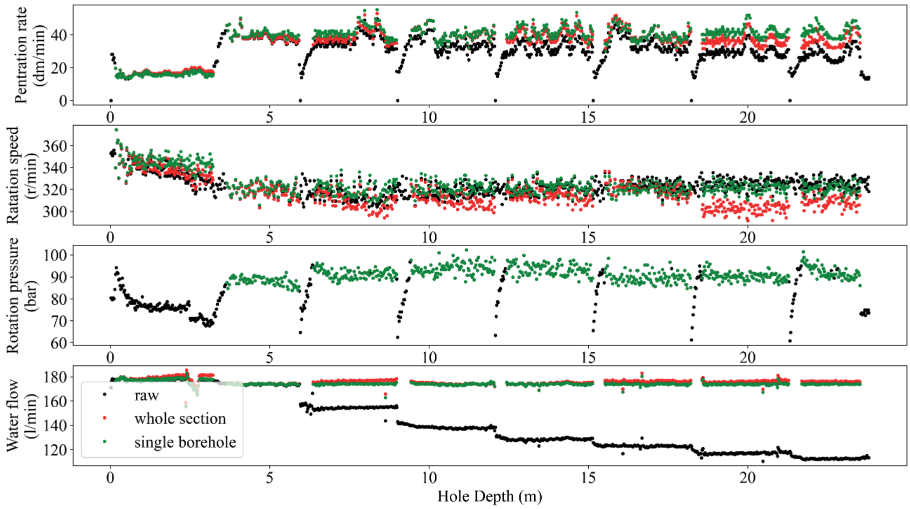

Figure 9. Comparison of two normalization methods for hole depth dependency removing. Raw data (black dots), normalized data after removing the hole depth dependency (green dots) and adopted regressions for each rod length (red lines).

5. DATABASE DESCRIPTION

At the final step, a centralized database was designed using PostgreSQL platform because of its robustness and open source object-relational database system. Then, the raw and processed data were transferred into this datacenter. Due to the developed automated coding it can continuously be updated using new upcoming data. Currently, it includes two types of data, the MWD (7252 files, 7252 boreholes, 60110094 data) and real time grouting (1583 files, 39766 boreholes, 6814391 data) for the project FSE410. A significant advantage of this database is that it can facilitate further developing through modern computational approaches like AI. Figure 10 shows the overview of the designed database that contains 6 relational tables and the connections between the tables are based on the settings of primary and foreign keys. The primary and foreign key relationships are used in relational databases to define many-to-one relationships between tables. The ‘ ’ corresponds to the table name. The ‘ID’ is the identifier index linked to the original ‘Raw File’. For example, the ID in ‘Data Type’ shows the type of data, i.e. ‘MWD’ or ‘Grouting’ which can be selected in ‘Column Name’. The table of ‘Raw File’ dedicates the information on name, folder, project and type of the original uploaded files using ‘File ID’, ‘File Name’, ‘Folder Name’, ‘Project Name’, and ‘Data Type ID’. The tables of ‘MWD_header’ and ‘Grouting_header’ store the information of the header of each data type that is linked to the corresponding file in the table of ‘Raw File’ via ‘File ID’. Accordingly, columns T1-T9 are the three-dimensional rotation matrix of the drill wreath for control the spatial direction and columns T10-T12 denote the absolute coordinates of the starting point of the borehole. The (‘ ’) shows the unique identity of each row in that table while (‘ ’) represents a set of attributes in a table that refers to the (‘ ’) of another table. These two keys connect the six tables together and enable users to extract data efficiently from different tables at the same time. Such utilities provide efficient choices to extract both MWD and grouting data through different query conditions and specific field ID values.

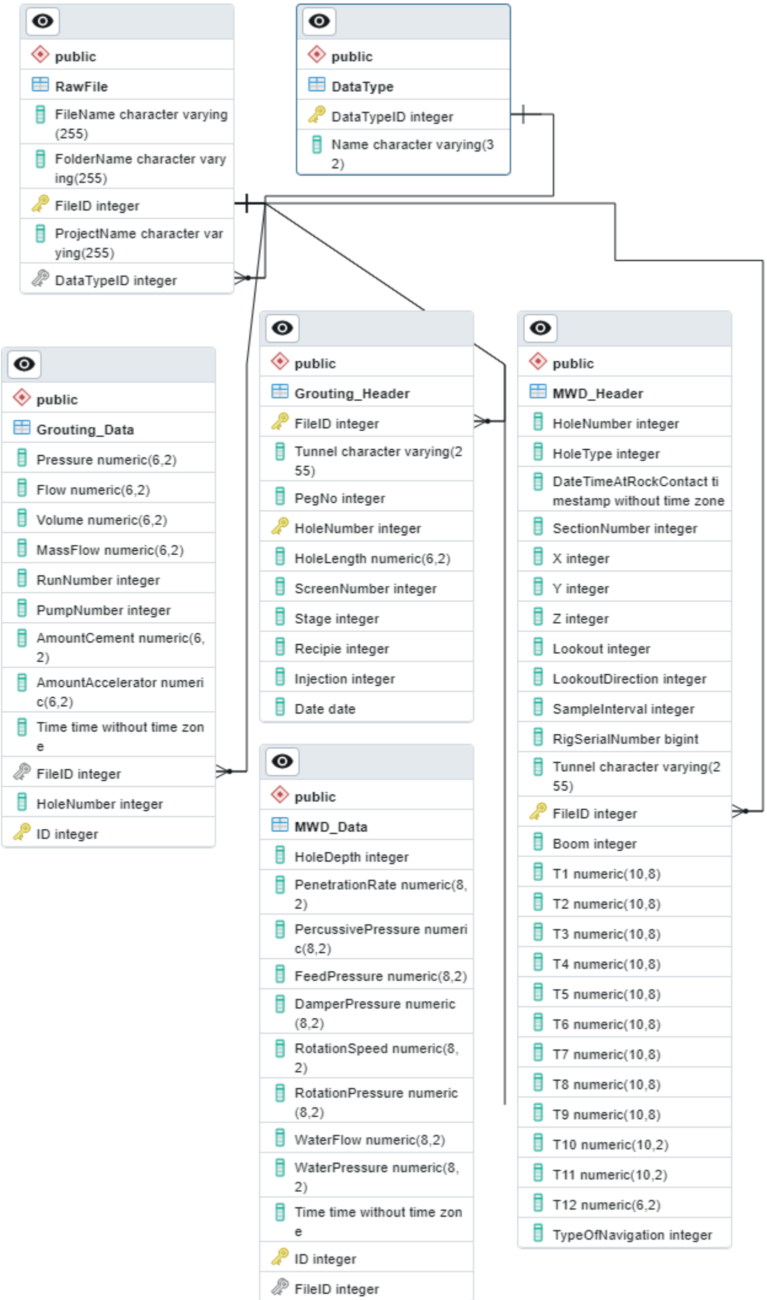

Figure 30. Schema for the PostgreSQL database design

6. DISCUSSION

In the current project, the filtering and normalizing process were only applied on MWD data. The original real time grouting data are stored directly in the database. The executed normalizing after filtering helped to remove some changes caused by differences in rod length, tool geometry, drilling setups in the machines etc. As recommended by Navarro et al., (2018), checking the combination of MWD parameters is preferred to sustain the important data points during the process. This is recommended to be investigated in the future. According to the categorized data state conditions (high/low/change mode) based on combined *PR-HP*, the efficiency of the proposed process in noise removal from the recorded data, i.e. improving the signal-to-noise ratio was approved. However, referring to Navarro et al. (2018), some of the data that fall within the identified Change mode may consist of information of the poor-quality rock that is needed for further investigation using other combined parameters. As an example, the combination of *RS*, *WP* and *WF* may show the variations of the rock mass (Schunnesson et al, 2011 and Navarro et al., 2018). From sensitivity analysis point of view, integrating of the normalized MWD data with other geotechnical information, i.e. rock mass characteristics and geological mapping can assist to recognize the most relevant parameters in MWD datasets and reflect changes in these properties in the MWD data. Therefore, deeper analysis of normalized MWD data can reveal more insights into the anomalies and trends in the formation that may be of interest for drilling (e.g. changes in lithology, porosity, or permeability). This is an important key for geoengineers, where the better understanding of the physical properties and characteristics of the formation being drilled the more accurately interpretation of the normalized MWD data.

The designed database, due to embedded possibilities, dedicates a centralized location with the ability for continuously updating. The provided facilities also show a modern but time/cost-effective tool for big data management and more detailed operational and research analyses.

7. CONCLUSIONS

In the current project, an entire automated process for filtering, normalizing and database creation for big MWD data in both hole and peer group-based was developed. Combination of *PR-HP* parameters was identified as the optimum choice for the filtering procedure. The distinguished states in data (high/low/change mode) using the adopted mode, long term average and percentile gated bands showed efficiency in removal of the noisy data caused by rig components, i.e. collaring and coupling effects from rod extensions. The applicability of normalizing process in removing the hole depth dependencies of MWD data were evaluated using correlational analysis. As a result, the hole-based normalizing method showed better performance in removing the hole depth dependencies and stepwise problem in MWD data. The presented procedure can generally be applied on any retrieved MWD data from each drill rig. The established MWD data center can structure and manage a big amount of MWD and grouting data to facilitate storing and extracting. The generated database mimics the big data characteristic (volume, value, variety, velocity and veracity), which not only can continuously be updated by upcoming data but also the users via the designed queries are able to extract desired data. It is an important tool for further deeper analyses of MWD data through modern approaches i.e. AI modeling. Incorporating the database with other geomechanical data sources can provide more accurate and realistic physical interpretations from MWD and grouting data.

8. AN PROPOSED IDEA OF USING ARTIFICIAL INTELLIGENCE ON MWD AND GROUTING DATA

In the context of geotechnical engineering, AI enables us to process vast amounts of data from geotechnical tests, site investigations, and historical records, and extract meaningful insights. This means that AI can identify patterns and relationships that might be challenging or even impossible for human engineers to spot. Additionally, AI can enhance the speed and accuracy of geotechnical analysis. Tasks that once required hours or even days of manual calculations can be done in a fraction of the time. This allows engineers to focus more on the creative aspects of their work, such as designing innovative solutions to engineering challenges.

The importance and globally fast growth of the AI systems are shaping the future of nearly every industry related to civil/geoengineering and will continue to act as a technological innovator for the foreseeable future. Toward this emerging technology, prediction of grouting design using AI approaches through combination of both MWD and grouting data can be highlighted as a novel idea. The estimated Lugeon values from control holes, as the benchmark criteria, can most probably be used to define grouting success, i.e. evaluating different grouting designs to detect which boreholes will fail or not pass the grouting criteria. A Lugeon is a unit devised to quantify the water permeability of bedrock and the hydraulic conductivity resulting from fractures. Table 1 shows an overview of the data types i.e. MWD and grouting design parameters which will be included in the AI model. The MWD should be acquired from the drilled rock mass. The trained optimum AI model then will predict the Lugeon value, an indicator to evaluate the successfulness of grouting design. The Lugeon values for the training dataset can be read from water flow protocols i.e. the control boreholes after grouting. Other grouting parameters (e.g. maximum pressure (bar), total volume (liter), the fan-based number of holes) except the duration time (second) can directly be acquired from the data files. Grouting duration time should be calculated based on the time-dependent volume data. As an example (Figure 11), the difference between the reaching to maximum grouting volume (21:04:34) and the start time (21:04:17) reflects the duration time of 17s.

Table 1. Input and output parameters for the AI idea

	Input												Output
	MWD								Grouting				
Borehole Number	Total length of fracture zones	Penetration Rate	Percussive Pressure	Feed Pressure	Damper Pressure	Rotation Speed	Rotation Pressure	Water Flow	Max Grout Pressure	Total Grout Volume	Grout Duration Time	Total number of grout holes	Lugeon Value (Lu)
1… N	val	val	val	val	val	val	val	val	val	val	val	val	val

During this process, AI will learn the relationships between the input and output (Table 1). Subsequently, the AI model will be used to test different grouting designs to give an indication of the design success. The perspective of this idea is to select the optimum grouting design and even an environmental-friendly design by the use of the least amount of grout materials.

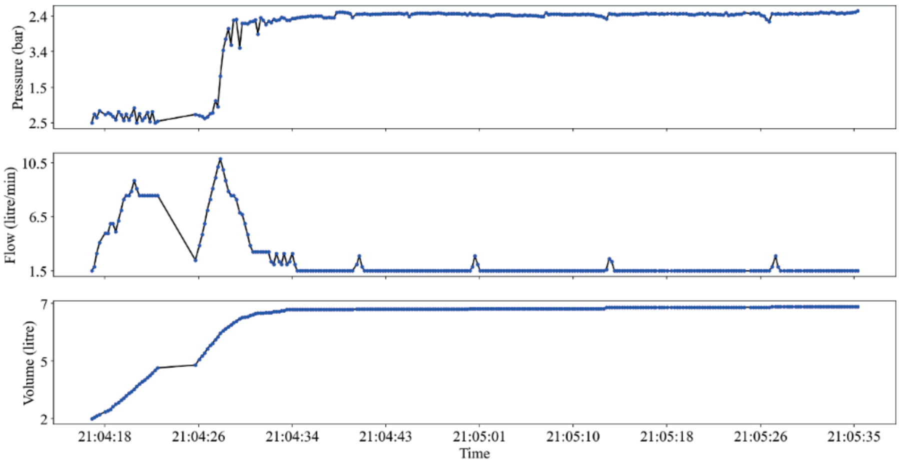

Figure 4. Plot of one grouting hole from real time grouting data

This idea, simply can be described in the following five steps:

Step 1: Correlating the grouting and MWD hole numbers

Data and header split, both MWD and grouting data (e.g. Figure 2, data extraction), should be carried out. The hole number and corresponding hole position can be read out from the MWD data file. However, the borehole number and the position information is not stored in the grouting file (usually .xml file format). The distribution of grouting hole numbers can only be read out from borehole map in .pdf format. Since this idea is hole based, it is important that the MWD data and grouting data are taken and compared from the same hole or spatially very close to each other. The position information from all the boreholes will be compared and used for borehole correlation.

Step 2: Correlate the hole number for MWD and the control hole MWD

Since in a drilling sequence, the number of control holes after grouting are usually less than the total MWD holes, the spatially correlating between the control and MWD holes is important. However, for an appropriate spatial analysis and borehole selection the minimum distance between the control and MWD hole should be defined in advance. Overall, the closer spatially located of MWD to the control holes will be selected for later usage and the rest will be discarded.

Step 3: Fracture index (*FI*) estimation

The fracturing index (*FI*) can be evaluated using the variation of the components of MWD data including penetration rate (*PR*), rotation pressure (*RP*), and percussive pressure (*HP*) as presented by Ghosh et al. (2017), Navarro et al. (2018a and 2019) and van Eldert et al. (2021a and 2021b). Therefore, the following established equations can be used to get the *FI* based on the variation of the MWD parameters.

$$\mathrm{FI} = \sqrt{\frac{HP_{var}}{\overline{HP}}} + \sqrt{\frac{FP_{var}}{\overline{FP}}} + \sqrt{\frac{RP_{var}}{\overline{RP}}} \qquad HP_{var} = \sum_{i}^{i+n} \left| \frac{\sum_{i}^{n+i} HP_i}{N+1} - HP_i \right|$$

$$FP_{var} = \sum_{i}^{i+n} \left| \frac{\sum_{i}^{n+i} FP_i}{N+1} - FP_i \right| \qquad RP_{var} = \sum_{i}^{i+n} \left| \frac{\sum_{i}^{n+i} RP_i}{N+1} - RP_i \right|$$

By estimating the *FI* for each borehole, the total length of the fracture zones can be calculated with a given threshold of *FI*.

Alternatively, the interpolated parameters such as rock hardness, fracture width and water quantities from software ‘GPM Tunnel+’, Rockma can be used in AI modelling. The seven different evaluation parameters categories: strength of intact rock (UCS), small fractures, medium-size fracture, large fracture, small quantities of water, medium quantities of water and large quantities of water can be used as input parameters together with the real time grouting parameters (Zetterlund et al., 2017).

Step 4: Selecting the representative data for the whole of the borehole

Since the data for AI solution is hole-based, the selection of one row of parameters that represents the whole borehole is needed. The selection method can be: 1) average of the data along one borehole, 2) taking the representative values for all the parameters, 3) selection of one row based on one criterion: e.g. the row with the highest *PR*. Different selection methods should be tested and compared to capture the optimum criterion.

This needs to be investigated and it might happen that the original MWD parameters are not necessary for the AI modelling process.

Step 5: Calculation of grouting parameters and database construction

This step will be carried out based on the instruction presented in Table 1 and Figure 11.

9. RECOMMENDATIONS FOR FUTURE WORK

1. Sensitivity analysis of different combinations of MWD parameters is recommended. The supervised Machine Learning methods such as Random Forest and Gradient Boost can be used to evaluate the feature importance of all the parameters and give scores that represent the importance of each parameter.
   Technically, this concern was one of the main critical points why in this project different combinations of MWD parameters in both filtering and normalizing for trend analyses and anomaly identifications were examined. As a result, deeper normalizing analysis can reveal more details about the anomalies and trends in the formation that may be of interest for drilling (e.g. changes in lithology, porosity, or permeability). In this case the sensitivity analysis not only can identify the importance of each recorded parameter in MWD data but also assists to understand how changes in geomechanical characteristics (e.g. characterized rock mass, geological mapping, rock physical properties) might be reflected in the MWD data. AI is one of the main attractive approaches for sensitivity analysis.
2. To extract physical meaningful interpretation from MWD data, the geoengineer should: 1) have good understanding of the formation properties, 2) have an ability in anomaly recognition and trend analysis, 3) be able to compare the data across boreholes at different times, 4) have capability to integrate MWD data with other data sources.
3. To understand the causes of the outliers in the MWD data, the best way is to compare MWD data with the true rock properties. These true rock properties can be obtained from core drilling or borehole filming exactly next to the MWD hole which provides information about rock qualities, fractures, fracture thickness and rock types as a function of drilling depth. By taking these steps (1, 2 and 3), analysts can gain valuable insights into the structure and properties of the formation being drilled to make decisions for further interpretation of outliers. This leads to highlight the importance of pattern/trend analysis to gain more interpretable physical details on the normalized MWD data, specifically when such data are incorporated with other geotechnical information, i.e. rock mass characteristics and geological mapping. Thus, the methodology presented can be a valuable supplementary tool as a basis for decisions.
4. Develop 3D spatial *FI*-based models/maps between boreholes and highlight the identified/predicted fracture zones and then compare with the after-excavation geological mapping.

ACKNOWLEDGEMENT

The authors would like to express their sincere gratitude to the Rock Engineering Research Foundation (BeFo) and Tyréns AB for supporting this research. We also wish to thank our colleagues for valuable contributions to the project: Jennifer Wänseth, Emilia Nordström and Rikard Gothäll, Håkan Schunnesson and the reference group, who provided insight and expertise that greatly assisted us in our research.

10. REFERENCES

Abdelaal A, Elkatatny S, Abdulraheem A (2022) Real-time prediction of formation pressure gradient while drilling. Scientific Reports, 12, 11318, https://doi.org/10.1038/s41598-022-15493-z.

Aljubran M, Ramasamy J, Albassam M, Magana-Mora A (2021) Deep learning and time-series analysis for the early detection of lost circulation incidents during drilling operations. IEEE Access, 9, 76833-76846, doi: 10.1109/ACCESS.2021.3082557.

Alreshidi E, Mourshed M, Rezgui Y (2018) Requirements for cloud-based BIM governance solutions to facilitate team collaboration in construction projects. Requirements Engineering, 23, 1–31, https://doi.org/10.1007/s00766-016-0254-6.

Arabjamaloei R, Edalatkha S, Jamshidi E, Nabaei M, Beidokhti M, Azad M (2011) Exact lithologic boundary detection based on wavelet transform analysis and real-time investigation of facies discontinuities using drilling data. Petroleum Science and technology, 29(6):569-578, https://doi.org/10.1080/10916460903419206.

Arrow KJ, Blackwell D, Girshick MA (1949) Bayes and minimax solutions of sequential decision problems. Econometrica, 17, 213-243.

Basarir H, Wesseloo J, Karrech A, Pasternak E, Dyskin A (2017) The use of soft computing methods for the prediction of rock properties based on measurement while drilling data. In J Wesseloo (ed.), Deep Mining 2017: Proceedings of the Eighth International Conference on Deep and High Stress Mining, Australian Centre for Geomechanics, Perth, 537-551, https://doi.org/10.36487/ACG_rep/1704_36_Basarir.

Bernoulli D (1738) Specimen theoriae novas de mesure sortis. Comentarii Academiae Scientarium Inperialis Petropolitanae, 5, 175-192. (Translated by Sommer L (1954) Exposition of a new theory on the measurement of risk. Econometrica, 22, 23-36.

Brown ET; Barr MV (1978) Instrumented drilling as an aid to site investigations. In Proc. 3rd International Congress of the International Association of Engineering Geology, 21–28, Madrid, Spain.

Börner I (1980) Untersuchungen zur optimierung nach mehreren zielen für aufgaben der bautechnologie. TH Leipzig, Sektion Technologie der Bauproduktion, Diplomarbeit.

Cayeux E, Daireaux B, Saadallah N, Alyeav S (2019) Toward seamless interoperability between real time drilling management and control applications. SPE/IADC International Drilling Conference and Exhibition, SPE-194110-MS, The Netherlands, DOI:10.2118/194110-MS.

Chapman JW, Reynolds D, Shreeves SA (2009) Repository metadata: approaches and challenges. Cataloging & Classification Quaterly, 47(3-4):309-325, https://doi.org/10.1080/01639370902735020.

Charoenprasit S, Seemuang N, Slatter T (2018) Monitoring tool wear in drilling process spindle noise features. International Journal of Mechanical Engineering and Robotics Research, 7(5):564-568, DOI: 10.18178/ijmerr.7.5.564-568.

Cheng S, Hu Y (2021) A novel signal processing algorithm for measurement while drilling (MWD) system. IEEE 40th Chinese Control Conference (CCC), Shanghai, China, 3105-3109, doi: 10.23919/CCC52363.2021.9550142.

Dekking MF, Kraaikamp C, Lopuhaa P, Meester LE (2005) A modern introduction to probability and statistics. Springer, London, doi: 10.1007/1-84628-168-7.

Deng LC, Li XZ, Xu W, Xiong Z, Wang J, Qiao L (2022) Measurement while core drilling based on a small-scale drilling platfrom: mechanical and energy analysis. Measurement, 204, 112082, https://doi.org/10.1016/j.measurement.2022.112082.

Eren T, Ozbayoglu ME (2010) Real time optimization of drilling parameters during drilling operations. SPE Oil and Gas India Conference and Exhibition, SPE-1291126-MS, Mumbai, India, https://doi.org/10.2118/129126-MS.

Ertunc HM, LoparoKA, Ocak H (2001) Tool wear condition monitoring in drilling operations using hidden Markov models (HMMs). International Journal of Machine tools and Manufacture. 41(9):1363-1384, https://doi.org/10.1016/S0890-6955(00)00112-7.

Fernández A, Sanchidrián JA, Segarra P, Gómez S, Li E, Navarro R (2023) Rock mass structural recognition from drill monitoring technology in underground mining using discontinuity index and machine learning techniques. International Journal of Mining Science and Technology, 33(5):555-571, https://doi.org/10.1016/j.ijmst.2023.02.004.

Gearhart M, Moseley LM, Foste M (1986) Current state of the art of MWD and its application in exploration and development drilling. In Proc. the International Meeting on Petroleum Engineering, Beijing, China, SPE-14071-MS, https://doi.org/10.2118/14071-MS.

Geekiyanage SCH, Tunkiel A, Sui D (2021) Drilling data quality improvement and information extraction with case studies. Journal of Petroleum Exploration and Production technology,11, 819–837, https://doi.org/10.1007/s13202-020-01024-x.

Ghosh R, Danielsson M, Gustafson A, Falksund H, Schunnesson H (2017) Assessment of rock mass quality using drill monitoring technique of Hydraulic ITH drills. International Journal of Mining and Mineral Engineering, 8(3):169-186, https://doi.org/10.1504/IJMME.2017.085830.

Ghosh R, Schunnesson H, Kumar U (2014) Evaluation of rock mass characteristics using measurement while drilling in Boliden Minerals Aitik Copper Mine, Sweden. In: Drebenstedt, C, Singhal R (eds) Mine Planning and Equipment Selection, 81-91, Springer, Cham. https://doi.org/10.1007/978-3-319-02678-7_9.

Hansen TF, Erharter GH, Marcher T, Liu Z, Tørresen J (2022) Improving face decisions in tunnelling by machine learning-based MWD analysis. Geomechanics and Tunneling, 15(2):222-231, https://doi.org/10.1002/geot.202100070.

Hodges JL, Lehmann EL (1952) The use of previous experience in reaching statistical decision. Annals of Mathematics Studies, 23, 396-407.

Hurwicz L (1951) Optimality criteria for decision making under ignorance. Cowles Commission Paper, Statistics, 370, 45-52.

Imieliński T, Virmani A, Abdulghani A (1999) DMajor- Application programming interface for database mining. Data Mining and Knowledge Discovery, 3, 347-372, https://doi.org/10.1023/A:1009841028985.

Ishaq M, Abid A, Farooq MS, Manzoor MF, Farooq U, Abid K, Helou MA (2023) Advances in database systems education: methods, tools, curricula, and way forward. Education and Information Technologies, 28, 2681–2725, https://doi.org/10.1007/s10639-022-11293-0.

Isheyskiy V, Martinyskin E, Smirnov S, Vasilyev A, Knyazev K, Fatyanov T (2021) Specifics of MWD data collection and verification during formation of training datasets. Minerals, 11, 798. https://doi.org/10.3390/min11080798.

Isheyskiy V, Sanchidrian JA (2020) Prospects of applying MWD technology for quality management of drilling and blasting operations at mining enterprises. Minerals, 10(10), 925, https://doi.org/10.3390/min10100925.

Jiao S, Zhang Q, Zhou Y, Chen W, Liu X, Gopalakrishnan G (2018) Progress and challenges of big data research on petrology and geochemistry. Solid Earth Science, 3(4):105-114, https://doi.org/10.1016/j.sesci.2018.06.002.

Jüttler H (1966) Untersuchungen zu Fragen der Operationsforschung und ihrer Anwendungsmöglichkeiten auf ökonomische Problemstellungen unter besonderer Berücksichtigung der Spieltheorie, Dissertation A, Wirtschftswissenschaftliche Fakultät der Humbold-Universität Berlin.

Kaas R, Buhrman JM (1980) Mean, median and mode in binomial distribution. Statistica neerlandica, 34(1):13-18, https://doi.org/10.1111/j.1467-9574.1980.tb00681.x.

Khorzoughi MB, Hall R, Apel D (2018) Rock fracture density characterization using measurement while drilling (MWD) techniques. Int. J. Min. Sci. Technol., 28, 859–864, https://doi.org/10.1016/j.ijmst.2018.01.001.

Khorzougi MB, Hall R (2016) Processing of measurement while drilling data for rock mass characterization. International Journal of Mining Science and Technology, 26(6):989-994, https://doi.org/10.1016/j.ijmst.2016.09.005.

Leighton J (1982) Development of a correlation between rotary drill performance and controlled blasting powder factors. M.Sc. Thesis, Department of Mining and Mineral Process Engineering, University of British Columbia, Canada.

Leung R, Scheding S (2015) Automated coal seam detection using modulated specific energy measure in a monitor-while-drilling context. International Journal of Rock Mechanics and Mining Sciences, 75, 196-209, https://doi.org/10.1016/j.ijrmms.2014.10.012.

Lopez JC, Lopez JE, Carcedo FJA(1995) Drilling and Blasting of Rocks, CRC Press.

Martin CA, Philo RM, Decker DP, Burgess TM (1994) Innovative advances in MWD. Proceeding of the IADC/SPE Drilling Conference, Dallas, Texas, Paper Number: SPE-27516-MS, https://doi.org/10.2118/27516-MS.

Martinsson L and Bengtsson F (2010) Tillämpning av MWD vid utformning av injektering. Svenska Bergteknikföreningen.

Navarro J (2018) The use of measure while drilling for rock mass characterization and damage assessment in blasting. PhD dissertation. University politecnica de Madrid.

Navarro J, Sanchidrian JA, Segarra P, Castedo R, Paredes C, Lopez LM (2018) On the mutual relations of drill monitoring variables and the drill control system in tunneling operations. Tunn. Undergr. Space Technol., 72, 294–304, https://doi.org/10.1016/j.tust.2017.10.011.

Navarro J, Ghosh R, Schunnesson H, Johansson D, Sanchidrián JA, Segarra P (2019) Application of drill-monitoring for chargeability assessment in sublevel caving. International J Rock Mech Min Sci., 119, 180-192, https://doi.org/10.1016/j.ijrmms.2019.03.026.

Navarro J, Segarra P, Sanchidrián JA, Castedo R, Pérez Fortes AP, Natale M, López LM (2018a) Application of an in-house MWD system for quarry blasting. In Proc. of 12th International Symposium on Rock Fragmentation by Blasting (Fragblast12), Luleå Sweden, 203-207.

Neumann von J, Morgenstern O (1943) Theory of Games and Economic Behavior. Princeton University Press.

Nilsen B (2015) Main challenges for deep subsea tunnels based on Norwegian experience. J of Korean Tunn Undergr Sp Assoc 17(5)563-573, https://doi.org/10.9711/KTAJ.2015.17.5.563.

Park J, Kim K (2020) Use of drilling performance to improve rock-breakage efficiencies: A part of mine-to-mill optimization studies in a hard-rock mine. International Journal of Mining Science and Technology, 30(2):179-188, https://doi.org/10.1016/j.ijmst.2019.12.021.

Peck J (1989) Performance monitoring of rotary blast hole drills. PhD dissertation, McGill University, Montreal, Canada.

Peldschus F, Vaigauskas E, Zavadskas EK (1983) Technologische entscheidungen bei der berücksichtigung mehrerer ziehle. Bauplanung–Bautechnik, 37(4):173-175.

Peldschus F (1986) Zur anwendung der theorie der spiele für aufgaben der bautechnologie. Dissertation B, Technischen Hochschule Leipzig, pp. 119.

Peldschus F (2008) Experience of the game theory application in construction management, Technological and Economic Development of Economy, 14(4):531-545.

Purkayastha AD, Nair PV (2017) Prospect level normalization of offset pore pressure measurements: analysis of approaches and their association with regional geology. SPE Oil and Gas India Conference and Exhibition, SPE-185394-MS, Mumbai, India, https://doi.org/10.2118/185394-MS.

Reckmann H, Jogi P, Kpetehoto F, Chandrasekaran S, Macpherson J (2010) MWD failure rates due to drilling dynamics.Procceding of ADC/SPE Drilling Conference and Exhibition, New Orleans, Louisiana, USA, Paper Number: SPE-127413-MS, https://doi.org/10.2118/127413-MS.

Rodgers M, McVay M, horhota D, Hernando J, Paris J (2020) Measuring while drilling in Florida limestone for geotechnical site investigation. Canadian Geotechnical Journal, 57(11):1733-1744, https://doi.org/10.1139/cgj-2019-0094.

Rostami J, Kahraman S, Naeimipour A, Collins C (2015) Rock characterization while drilling and application of roof bolter drilling data for evaluation of ground conditions. J. Rock Mech. Geo. Eng. 7(3):273–281, https://doi.org/10.1016/j.jrmge.2015.01.006.

Saunders MR, Shields JA, Taylor MR (1996) Improving the value of geological data: a standardized data model for industry. Geological Society, 97, 41-53, https://doi.org/10.1144/GSL.SP.1995.097.01.06.

Savage LJ (1951) The theory of statistical decision. Journal of the American Statistical Association, 46, 55-57.

Schunnesson H (1996) RQD predictions based on drill performance parameters. Tunn Undergr Space Technol 11(3):345–351

Schunnesson H (1998) Rock characterization using percussive drilling. Int. J. Rock Mech. Min. Sc., 35(6):711–725, https://doi.org/10.1016/S0148-9062(97)00332-X.

Schunnesson H, Kristoffersson T (2011) Rock mass characterization using drill and crushability monitoring – a case study. Int J COMADEM., 14(21):44–52.

Schunnesson, H., Elsrud, R. and Rai, P. (2011). Drill Monitoring for Ground Characterization in Tunnelling Operations. Proceedings of the Twentieth International Symposium on Mine Planning and Equipment Selection MPES 2011, 12-14 October 2011, Almaty, Republic of Kazakhstan. ISBN 978-601-7146-15-3

Smith B (2002) Improvements in blast fragmentation using measurement while drilling parameters. International Journal for Blasting and Fragmentation, 6(3–4):301-310.

Song S, Zhang T, Wang Z, Pei R, Yan S, Chen K (2022) Full waveform vibration and shock measurement tool for measurement-while-drilling. AIP Advances, 12, 085114, https://doi.org/10.1063/5.0090505.

Tsatalos OG, Ioannidis YE (1994) A unified framework for indexing in database systems. In: Karagiannis D (eds) Database and Expert Systems Applications, DEXA 1994, Lecture Notes in Computer Science, vol 856, 183-192, Springer, Berlin, Heidelberg, https://doi.org/10.1007/3-540-58435-8_183.

van Eldert J, Schunnesson H, Saiang D, Funehag J (2020) Improved filtering and normalizing of Measurement-While-Drilling (MWD) data in tunnel excavation. 103, 103467, https://doi.org/10.1016/j.tust.2020.103467.

van Eldert J, Funehag J, Schunnesson H, Saiang D (2021a) Drill Monitoring for Rock Mass Grouting: Case Study at the Stockholm Bypass. 54:501-511. https://doi.org/10.1007/s00603-020-02279-w.

van Eldert J, Funehag J, Saiang D, Schunnesson H (2021b) Rock support prediction based on measurement while drilling technology. 80:1449-1465. https://doi.org/10.1007/s10064-020-01957-x.Wald A (1945) Statistical decisions functions which minimise the maximum risk. Annals of Mathematics. 46, 265-280.

Weitendorf D (1976) Beitrag zur optimierung der räumlichen struktur eines gebäudes. Dissertation A, Hochschule für Architektur und Bauwesen Weimar.

Wu S (2013) A review on coarse warranty data and analysis. Reliability Engineering & System Safety. 114, 1-11, https://doi.org/10.1016/j.ress.2012.12.021.

Yang Y, Li F, Gao Y, Mao Y (2020) Multi-sensor combined measurement while drilling based on the improved adaptive fading square root unscented Kalman filter. Sensors, 20(7), 1897, https://doi.org/10.3390/s20071897.

Yao K, Gao J (2016) Law of large numbers for uncertain random variables. IEEE Trans Fuzzy Syst., 24(3):615–621, doi: 10.1109/TFUZZ.2015.2466080.

Zavadskas EK, Ustinovičius L, Turskis Z, Peldschus F, Messing D (2002) LEVI 3.0-multiple criteria evaluation program for construction solutions. Journal of Civil Engineering and Management, 8(3):184-191, https://doi.org/10.3846/13923730.2002.10531275.

Zavadskas EK, Turskis Z (2008) A new normalization method in games theory. Informatica, 19(2):303-314, https://doi.org/10.15388/Informatica.2008.215.

Zetterlund M, Martinsson L and Dalmalm T (2017) Implementation of MWD-data for grouting purposes in a large infrastructure project- the Stockholm Bypass. Grouting 2017 GSP 288, 61-70.

Zhang W, Ching J, Goh A and Leung A (2021) Big data and machine learning in geoscience and geoengineering: Introduction. Geoscience Frontiers, 12(1): 327-329. https://doi.org/10.1016/j.gsf.2020.05.006.

Zhao R, Shi S, Li S, Guo W, Zhang T, Li X, Lu J (2023) Deep learning for intelligent prediction of rock strength by adopting measurement while drilling data. International Journal of Geomechanics, 23(4), https://doi.org/10.1061/IJGNAI.GMENG-8080.

Zhao Q, Zhang B, Hu H (2009) Novel two-step filtering scheme for a logging while-drilling system. Computer Physics Communications, 180(9):1566-1571, https://doi.org/10.1016/j.cpc.2009.04.011.

Zhussupbekov A, Alibekova N, Akhazhanov S, Sarsembayeva A (2021) Development of a unified geotechnical database and data processing on the example of Nur-Sultan City. Applied Sciences, 11(1), 306, https://doi.org/10.3390/app11010306.

Ziegler P, Dittrich KR (2007) Data integration- problems, approaches, and perspectives. In: Krogstie J, Opdahl AL, Brinkkemper S (eds) Conceptual Modelling in Information Systems Engineering, 35-98, Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-540-72677-7_3.

APPENDIX A:

Some of the industrial companies with great efforts in the development of MWD data collection and analysis:

Sweden

Sandvik (https://www.rocktechnology.sandvik)

Epiroc (https://www.epiroc.com)

Rockma (https://www.rockma.com/)

Norway

Bever Control (https:www.bevercontrol.com)

Some of the companies and organizations with great efforts in applying MWD data in infra geoengineering projects:

SIP-STRIM (https://www.sipstrim.se/project/innovative-dth-drill-monitoring-a-pre-study)

Swebrec: Swedish Blasting Research Centre (https://www.ltu.se/centres/swebrec/Vara-projekt)

Russia

Zyfra mining (https:www.zyfra.com)

Kyrgyzstan

Blast maker (http://blastmaker.kg)

Canada

Peck Tech Consulting (http://pecktech.ca)

Spain

MAXAM (https://www.maxamcorp.com)

APPENDIX B:

The pseudo code of the automated procedure in the case of single hole and peer group-based normalization.

1. The implemented pseudocode for hole-based:

1. Read data from one borehole

2. For i=1 to number of holes (Figure 5)

3. Rod length checking through data splitting (e.g. Figure 2)

4. Automating process through inner loops (Figure 8)

5. Repeat

5.1. Calculate the correlation coefficient between each parameter and hole depth

5.2. Condition analysis of correlational dependencies:

5.3. If coefficient >= 0.1, continue with normalization;

5.4. else,

5.5. Terminating process as there is no normalization for the parameter

5.6. Linear regression analysis for each rod and each parameter, $y = b_1x + b_0$ (b_0, b_1: the regression parameters estimated for each rod and each parameter)

5.7. Removing the entire depth dependency based on the coefficient. $y_{norm} = y_{obs} - coef * b_1x - b_0 + b_{0,rod1}$ ($b_{0,rod1}$: intercept of the linear model from rod number 1 for each parameter; y_{obs} andy_{norm}: the measured and normalized data respectively).

6. Until the number of holes (Repeating steps 5-7 for i= number of holes)

7. End For

2. The implemented process for peer group (merged)-based data:

1. For i=1 to number of holes

2. Read all data

3. Peer group analysis (merging procedure)

3.1. All data

3.2. Based on folder project according to holes from one sequence (fan) for each parameter)

3.3. Rod length checking through data splitting (e.g. Figure 7)

4. End For i

5. For J=1 to number of peer groups (Figure 8)

6. Repeat

6-1. Calculate the correlation coefficient between each parameter and hole depth

6.2. Condition analysis of correlational dependencies:

6.3. If coefficient >= 0.1, continue with normalization;

6.4. else,

6.5. Terminating process as there is no normalization for the parameter

6.6. Linear regression models for each rod and each parameter (merged boreholes) in the form of $y = b_1 x + b_0$ (b_0, b_1: the regression parameters estimated for each rod and each parameter from merged data).

6.7. Removing the entire depth dependency for each hole based on the coefficient using $y_{norm} = y_{obs} - coef * b_1 x - b_0 + b_{0,rod1}$ ($b_{0,rod1}$: the intercept of the linear model from rod number 1 for each parameter; y_{obs} and y_{norm}: the measured and normalized data respectively).

7. Until the number of peer groups (fan data) ((Repeating steps 6-8 for J= number of peer groups)

8. End For J

Box 5501
SE-114 85 Stockholm
info@befoonline.org • www.befoonline.org
Visiting address: Storgatan 19, Stockholm

ISSN 1104-1773