Proceedings of the 21st ICSMGE, Vienna, Austria, 14 – 19 June 2026. Pistrol, Adam & Schweiger (eds.)
Published by: ÖGG, Austrian Society for Geomechanics, Salzburg, Austria, ISBN 978-3-9503898-4-5
https://doi.org/10.53243/ICSMGE2026-876

Data-driven subsurface geological mapping along tunnel infrastructure: a case study at CERN

Zahra Hajighasemi, Biao He, Zili Li
[a] *Civil, Structural & Environmental Engineering, University College Cork, Cork, Ireland, 124113275@umail.ucc.ie*

Andrea Visentin [cd]
[c] *School of Computer Science & IT, University College Cork, Cork, Ireland*
[d] *Insight Centre for Data Analytics, University College Cork, Cork, Ireland*

Aohui Ouyang [ab], Vanessa Di Murro [b], John Andrew Osborne [b]
[b] *SCE-SAM-FS, European Organization for Nuclear Research, CERN, Geneva, Switzerland*

ABSTRACT: The maintenance of large underground infrastructure is a challenging task. Hydrogeological profiles such as groundwater level, hydraulic permeability, rock characteristics, and water-rock interaction can lead to cracks, leakage, and other defects, posing risks to tunnel safety and serviceability. A comprehensive geological map and groundwater profiles are two crucial factors for assessing the risk of adverse effects on tunnel functionality. Recently, Machine Learning (ML) methods have emerged for geological mapping. Previous studies used a variety of parameters to train their models and then build a geological map, like borehole coordinates and borehole logging tests. However, the limitation of borehole numbers, paper-based handwritten log records, and imbalanced labels can lead to challenges in implementing the ML models. In this study, the goal is to develop a geological map using ML algorithms and the coordinates of boreholes. As a first step, borehole log reports were collected from specific CERN tunnel sections and then digitized, structured, and integrated. After that, preprocessing steps were applied to the structured dataset, including filtering the rock types, considering additional points, eliminating duplicate points, and one-hot encoding for multiple labels. Subsequently, ML techniques like Random Forest (RF), Support Vector Machine (SVM), Categorical Boosting (CatBoost), and eXtreme Gradient Boosting (XGBoost) were trained using the available data. The Leave-One-Out Cross-Validation (LOOCV) was implemented to split the train and test sets because of the limited number of boreholes. The geological maps for the LHC tunnel from Point 3 to Point 4 were developed. In future studies, potential correlations between the surrounding hydrogeological profiles and structural defects in tunnel infrastructure will be investigated, contributing to the prediction of future defects and risks in underground infrastructure.

KEYWORDS: Geological mapping, Machine learning, Underground tunnel, Borehole log report.

1 INTRODUCTION

Due to the growth of large infrastructure and the impact of the hydrogeological profile on them, many studies have tried to propose innovative methods for tunnel maintenance during their operation period. The hydrogeology of the surrounding environment of the tunnel can affect its performance and can cause deformation, degradation, cracks, and shifts. Several methods for constructing a geological map of the study area have been proposed in the literature, including spatial interpolation, kriging, and inverse distance weighting. However, these traditional approaches have notable limitations: they struggle to manage sparsely distributed data, scale poorly with large datasets, and have limited ability to capture complex and nonlinear relationships. In contrast, machine learning algorithms offer greater flexibility and robustness in overcoming these challenges, as they are capable of uncovering hidden correlations among geological features.

Methods for 3D modelling of subsurface stratigraphy may be divided into two categories: geological-interface-based modelling methods ((Zhu et al., 2012), (Kim & Ji, 2022), (Cracknell & Reading, 2014)) and point-cloud-based modelling methods ((Guo et al., 2019), (Lyu & Wang, 2024), (Liu et al., 2024)). The first group generates interfaces between different strata from available measurements, while the second one represents a 3D domain with distributed points, with a label of the corresponding rock type (Lyu et al., 2024).

The input data for ML algorithms can take various forms, such as images ((Shi & Wang, 2022), (Lyu et al., 2024)) and text to develop a geological map. Geologically speaking, this input data can be extracted from borehole log reports. Some studies focused on using only borehole coordinates as the input features to classify soil/rock types ((Guo et al., 2019), (Liu et al., 2024), (Hu et al., 2024)). Depending on the amount and type of data, previous researchers applied a variety of ML algorithms for their objectives. Where enough data was available, the Neural Networks were able to predict the soil/rock categories ((Guo et al., 2019), (Liu et al., 2024), (Kim & Ji, 2022)). However, with only a limited number of boreholes, other techniques like SVM, RF, and XGBoost were used (Hu et al., 2024).

To indicate the importance of the hydrogeological profile on the tunnel performance, (Li et al., 2024) reviewed 155 cases of significant damage or even collapse of some parts of the tunnel. The results demonstrated that geological factors have a higher impact than accidental factors, lining factors, design, construction, and operational hazards. Among these geological factors, half of them were related to groundwater and poor geological bodies. In addition, (Wang et al., 2024) studied hydraulic factors, including groundwater profile, hydraulic permeability, and hydraulic deterioration, influence long-term tunnel performance.

The current research aims to develop a geological map near the CERN tunnels and explore potential correlations with hydrogeological profiles for future investigations.

2 GEOLOGICAL MAPPING

2.1 *Study area description*

CERN, the European Laboratory for Particle Physics, is located on the Franco-Swiss border in Geneva, Switzerland. Two underground tunnels—the Large Hadron Collider (LHC) and the Super Proton Synchrotron (SPS)—host detectors and experimental facilities used to explore fundamental particles.

The rock mass surrounding most CERN tunnels consists largely of marls and sandstones and it is called the red molasse. However, a smooth transition between layers makes predicting unknown layers difficult (Fern et al., 2018).

Figure 1 shows the positioning of the LHC, which is at an altitude of 350 m to 500 m with a slight slope, and for the SPS tunnel the altitude is almost 400 m (Figure 1).

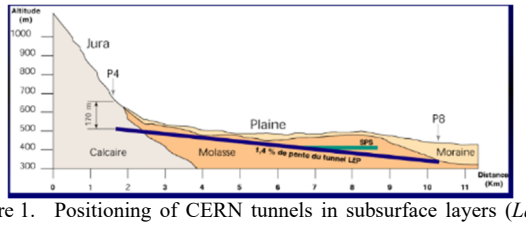

Figure 1. Positioning of CERN tunnels in subsurface layers (*Large Colliders:Civil Engineering and Siting*, n.d.)

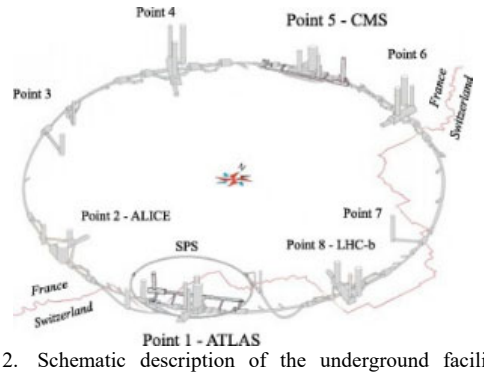

Figure 2. Schematic description of the underground facilities at CERN (Fern et al., 2018)

2.2 *Data collection and structuring*

In the presented study, two data sources were used. Some borehole reports were provided by colleagues at CERN, while others were extracted from a French study (Charollais et al., 2007) conducted on the geology of the area. These documents were PDF files and unstructured. The collected data were in four data sets: "SPS", "SLHC", "LHC point 3 to point 4 "(also may be named as "LHC_3_4"), "LHC French", and the integrated dataset named "all". From each report, information such as coordinates, altitude, depth, stratum geological layer, and type of soil/rock in each specific depth of boreholes was extracted and stored in Microsoft Excel. In these structured files, each row had the following attributes: X coordinate, Y coordinate, altitude, borehole name, and type of rock, which started from the altitude of the upper layer and ended in the same row altitude. After the steps mentioned earlier, a new column was added to store the order of the layers in boreholes (Figure 3).

Due to the lack of information about the coordinates of some boreholes, the available maps in the reports were used to estimate their location.

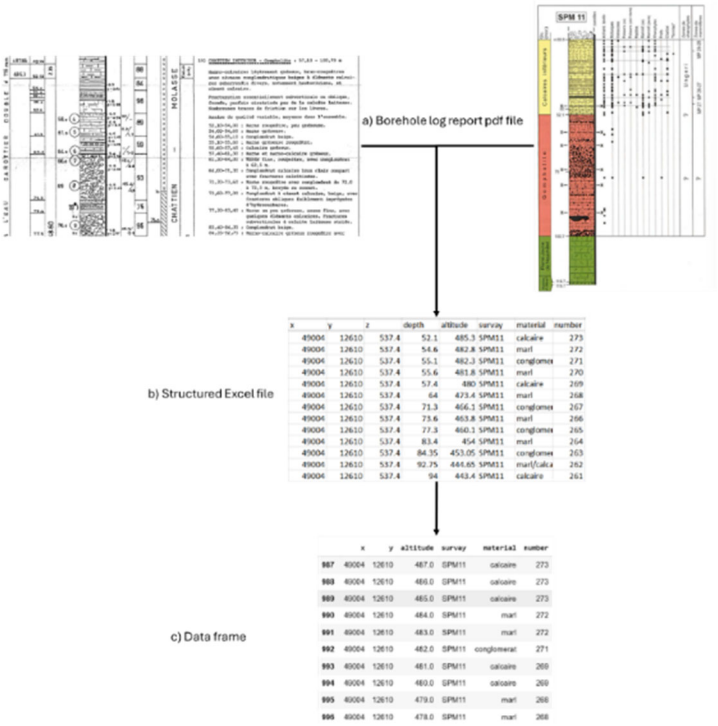

Figure 3. Structuring unstructured data and integration

2.3 *Data integration*

As the available data were taken from various sources, their coordinate system did not have the same origin and scale. Therefore, some common boreholes and maps were used to integrate a unified origin for all the boreholes. The origin of the boreholes from the French report is assumed as a reference, and others are redefined according to that.

For the SPS tunnel, F6 and F15 boreholes were found in both reports, so by using the Get Data software (version 2.24) the new axes were defined, and the location of the other boreholes was retrieved.

In the LHC tunnel, the available data from CERN reports were for the section of the tunnel from point 3 to point 4, and the boreholes SPM11, SPM15, and L135 were used to update the borehole locations. Figure 5 illustrates how the common boreholes were chosen, and the new coordinates are shown on the right section of the image.

Figure 4 shows the boreholes of each data set in the unified coordinates.

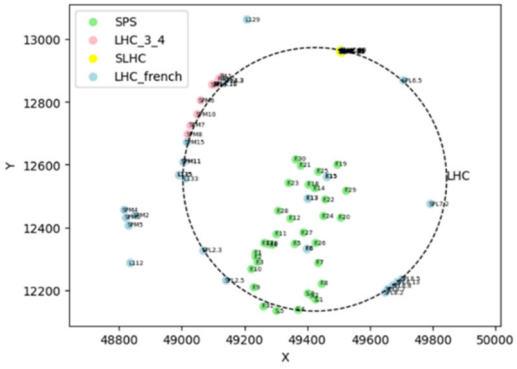

Figure 4. Borehole location visualization for each dataset

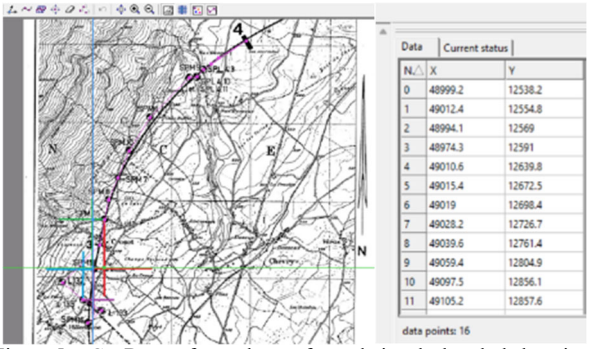

Figure 5. Get Data software layout for updating the borehole locations

Finally, the excel files were converted to data frames using the Pandas library in Python, and the next steps were taken in a Jupyter notebook environment.

The "all" data set represents the integrated data set in which the common boreholes were considered only once, and their coordinates are updated in the same referencing system.

2.4 *Preprocessing*

One of the crucial steps in the ML procedure is data preprocessing. In the present study, the following steps were taken:

1. Adjusting altitude

The accuracy of measurements in different datasets varied. To address this problem, all the altitudes were rounded up if their decimal number was greater than 0.5 and otherwise rounded down. Thus, the accuracy of the Z axis (altitude) became 1 meter.

2. Filtering soil/rock

The boreholes contain a variety of geological layers, such as topsoil and clay, but in the current study, only deeper layers were filtered. Therefore, the 5 types of soil/rocks were selected as the final groups, including "marl", "sandstone", "marl-calcaire", "calcaire", and "conglomerate". These were the most frequent layers in the boreholes examined. In this study, other types of soil/rock, including topsoil, clay, etc., were not considered and they were removed from the training and test sets.

3. Adding new points

In the available data frames, the altitude value for each soil/rock layer represents the bottom of that layer, while the top is defined by the altitude of the layer directly above it. Figure 6 illustrates that, to digitize all the boreholes, new points were added for the layers deeper than 1 meter, and the same information was assumed for these new rows, except their altitude, which was changed along that layer.

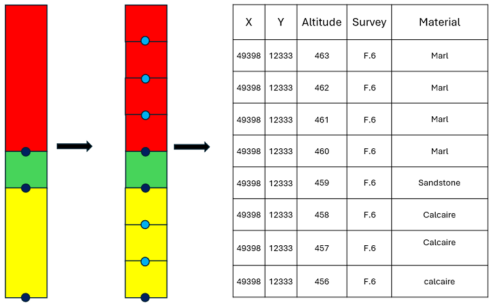

Figure 6. Adding new points for borehole digitalization

4. Removing duplicate points

Rounding the altitude values may result in duplicate records (e.g., altitudes of 450.2 m and 450.4 m both become 450 m). To avoid redundancy and overfitting, only one instance of each duplicate point was retained. Although this process may lead to loss of information, it is acceptable because the model's target resolution is 1 m.

Table 1 shows the number of boreholes in each dataset. Moreover, Table 2 indicates the distribution of the datapoints labeled as each group within a dataset, and Figure 7 illustrates their 3D visualization.

Table 1. Number of boreholes in each dataset

	SPS	SLHC	LHC 3 to 4	LHC French	all
Number of boreholes	36	7	12	24	73

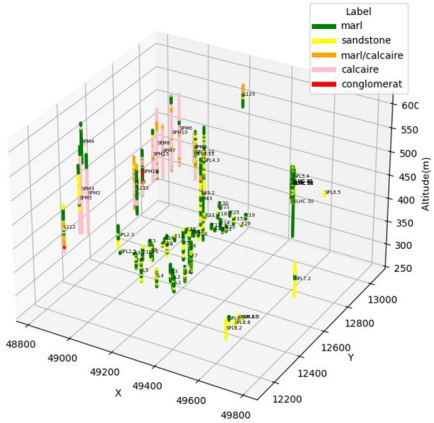

Figure 7. 3D borehole visualization

5. One-hot encoding

Regarding the categorical type of soil/rock, the output of the ML algorithms, one-hot encoding can be helpful. It transforms the label to a binary variable. Instead of a single output value, the output is represented as a one-hot encoded vector, where the length corresponds to the number of unique soil/rock labels. The value for the desired label is set to 1, and all the others are set to 0.

Table 2. The number of data points in the final data sets

Data / soil/rock	SPS	SLHC	LHC_3_4	LHC_French	all
marl	594	479	370	528	1857
sandstone	252	103	226	355	838
calcaire	0	0	713	620	1210
marl-calcaire	0	0	149	93	225
conglomerate	0	0	17	44	23
Total	846	582	1475	1640	4153

2.5 *Machine Learning algorithms*

Different ML algorithms are implemented to build a geological map in the study area. A brief description of them is provided to address the basics and goals of these algorithms.

2.5.1 *Support Vector Machine (SVM)*

This algorithm fits the optimal hyperplane for classifying the input data. For binary classification, it tries to find a line with maximum distance from the nearest point of each class (Mathur & Foody, 2008). This study considered the radial basis function as the kernel for nonlinear relationships.

2.5.2 *Random Forest (RF)*

This ML technique combines decision trees by aggregating their predictions to improve accuracy and prevent overfitting. The base node of each decision tree is chosen randomly, and the final leaf represents the output label of that tree. The most voted label will be the prediction (He et al., 2016).

The summarized steps of the RF are:

1. Bootstrap sampling: sampling training data randomly with replacement to create subsets.
2. Model training: building trees by selecting a random feature and finding the best split based on the criteria
3. Model forecasting: the predicted class from each tree
4. Aggregation: the most voted class as the final prediction.

2.5.3 *Categorical Boosting (CatBoost)*

The gradient boosted decision trees are the base of CatBoost. A set of decision trees will be updated in a direction to reduce the loss function. With a greedy method, more impactful features will be chosen, and the tree structure will be retrieved. This algorithm can handle categorical features with automatic one-hot encoding (Pandey et al., 2023).

2.5.4 *eXtreme Gradient Boosting (XGBoost)*

This is a supervised algorithm based on ensemble trees. New decision trees will be generated, reducing the error of previous trees. It will be stopped when the expected accuracy is achieved, and the optimal loss function and regularization term are provided.

2.6 *Implementation*

After collecting and preprocessing the data, it must be split into the training and test sets. Regarding the previous studies, the Leave-One-Out Cross-Validation (LOOCV) is the best choice for limited data (Shi & Wang, 2022). LOOCV chooses a random borehole's data points as a test set and uses the rest as the training set in each iteration. This continues, and all the boreholes are used for the test once. Consequently, the average accuracy of the iterations can evaluate the performance of each algorithm.

The ML algorithms were implemented in Python using scikit-learn (SVM, RF), CatBoost, and XGBoost libraries.

After implementing different ML algorithms, the average accuracies and balanced accuracies were used to evaluate the models' performance (Figure 8).

Relying on average accuracy as an evaluation metric alone is insufficient for these datasets because they are imbalanced. For example, in the SPS data set, the labels of 70 % of the datapoints are marl, and 30 % of them are sandstone. Therefore, an average accuracy of 70 % would be achieved by always predicting the majority class, and it is not enough to decide if the model is learning the relationship between features. However, the average balance accuracy will consider the correct predictions of each class and is appropriate for imbalanced datasets. The value of balanced accuracy also depends on the number of classes. For instance, if there are n output labels and the balanced accuracy is about 1/n, the model performs as well as simply predicting the majority class, and higher balanced accuracies are a sign of the learning process by the model.

$$balanced\ accuracy = \frac{1}{2}\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right)$$

After comparing the ML algorithms' performance, the chosen technique was applied to build a geological map near the LHC tunnel from point 3 to point 4. To develop this map, the SVM algorithm was applied to the X coordinates and ground levels of the boreholes. Then, using the circle formula, the corresponding Y coordinates were calculated, allowing the generation of new coordinates for artificial boreholes within that specific tunnel section. Next, the trained ML algorithm was used to predict the soil or rock classification for the data points associated with the artificial boreholes. Based on these predictions, a 3D geological map was subsequently developed for the section of the LHC tunnel between point 3 and point 4.

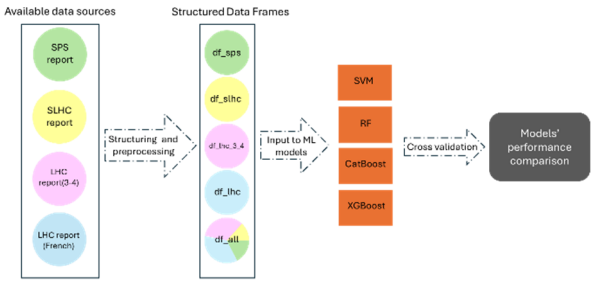

Figure 8. Flow chart of data preparation and ML algorithms' performance comparison

3 RESULTS AND DISCUSSION

After conducting the ML algorithms on the datasets, various metrics were employed to evaluate model performance. In addition to average accuracy and average balanced accuracy, the precision, recall, and F1-score were included to provide more details. The results are presented in Table 3.

The recall values for SVM show that the model predicts the same label for all the datapoints because the recall in most datasets is equal to 1 for a specific class and 0 for the others.

Comparing the average accuracy and average balanced accuracy, in SPS and SLHC, the accuracy is about 0.7, and balanced accuracy is about 0.5, which means none of the ML techniques can learn any trend in these datasets, because of the reasons explained in the implementation section.

In the remaining data sets, there are 5 output classes, and the balanced accuracy should be more than 0.2 to perform better than a majority prediction.

The balanced accuracy is approximately 0.5 for both "LHC_French" and "all" datasets, indicating that the model is learning some relations between the input features. Moreover, the average accuracy exceeds 0.6 when the model is trained on the complete dataset. This suggests that increasing the number of boreholes and distributing them along the tunnel, enhances the model's performance.

Table 3. Comparing ML performance on different data sets

		avg_acc	avg_balanced_acc	precision_C	recall_C	F1_score_C	precision_CO	recall_CO	F1_score_CO	precision_M	recall_M	F1_score_M	precision_M/C	recall_M/C	F1_score_M/C	precision_S	recall_S	F1_score_S
SPS	SVM	0.75	0.58	-	-	-	-	-	-	0.7	1	0.82	-	-	-	0	0	0
	RF	0.67	0.53	-	-	-	-	-	-	0.69	0.8	0.74	-	-	-	0.26	0.16	0.2
	CatBoost	0.72	0.57	-	-	-	-	-	-	0.7	0.87	0.77	-	-	-	0.29	0.13	0.18
	XGBoost	0.66	0.53	-	-	-	-	-	-	0.69	0.77	0.73	-	-	-	0.27	0.2	0.23
SLHC	SVM	0.84	0.57	-	-	-	-	-	-	0.82	1	0.9	-	-	-	0	0	0
	RF	0.76	0.56	-	-	-	-	-	-	0.85	0.94	0.89	-	-	-	0.45	0.23	0.3
	CatBoost	0.77	0.53	-	-	-	-	-	-	0.83	0.98	0.9	-	-	-	0.5	0.09	0.15
	XGBoost	0.77	0.57	-	-	-	-	-	-	0.85	0.89	0.87	-	-	-	0.36	0.28	0.32
LHC_3_4	SVM	0.5	0.3	0.48	1	65	0	0	0	0	0	0	0	0	0	0	0	0
	RF	0.41	0.25	0.74	0.67	0.7	0	0	0	0.33	0.43	0.38	0.08	0.07	0.08	0.24	0.22	0.23
	CatBoost	0.45	0.25	0.72	0.73	0.73	0	0	0	0.36	0.41	0.38	0.09	0.07	0.08	0.26	0.23	0.25
	XGBoost	0.46	0.38	0.71	0.73	0.72	0	0	0	0.37	0.41	0.39	0.09	0.09	0.09	0.27	0.19	0.23
LHC_French	SVM	0.12	0.13	0.24	0.5	0.33	0	0	0	0.13	0.09	0.11	0	0	0	0	0	0
	RF	0.53	0.5	0.62	0.45	0.52	0	0	0	0.42	0.62	0.5	0	0	0	0.41	0.42	0.42
	CatBoost	0.53	0.48	0.55	0.65	0.59	0.5	0.11	0.19	0.36	0.44	0.4	0	0	0	0.4	0.21	0.28
	XGBoost	0.46	0.48	0.58	0.52	0.55	0.19	0.11	0.14	0.36	0.55	0.44	0	0	0	0.43	0.26	0.33
all	SVM	0.57	0.47	0	0	0	0	0	0	0.45	1	0.62	0	0	0	0	0	0
	RF	0.61	0.51	0.63	0.62	0.63	0	0	0	0.61	0.69	0.65	0.07	0.04	0.05	0.4	0.34	0.37
	CatBoost	0.65	0.5	0.67	0.8	0.73	0	0	0	0.63	0.71	0.67	0.04	0.01	0.02	0.38	0.25	0.3
	XGBoost	0.6	0.5	0.67	0.63	0.65	0	0	0	0.6	0.76	0.67	0.05	0.03	0.04	0.45	0.26	0.33

abbreviations :

avg_acc	average accuracy	C	Calcaire
avg_balanced_acc	average balanced accuracy	Co	Conglomerate
		M	Marl
		S	Sandstone
		M/C	Marl/calcaire

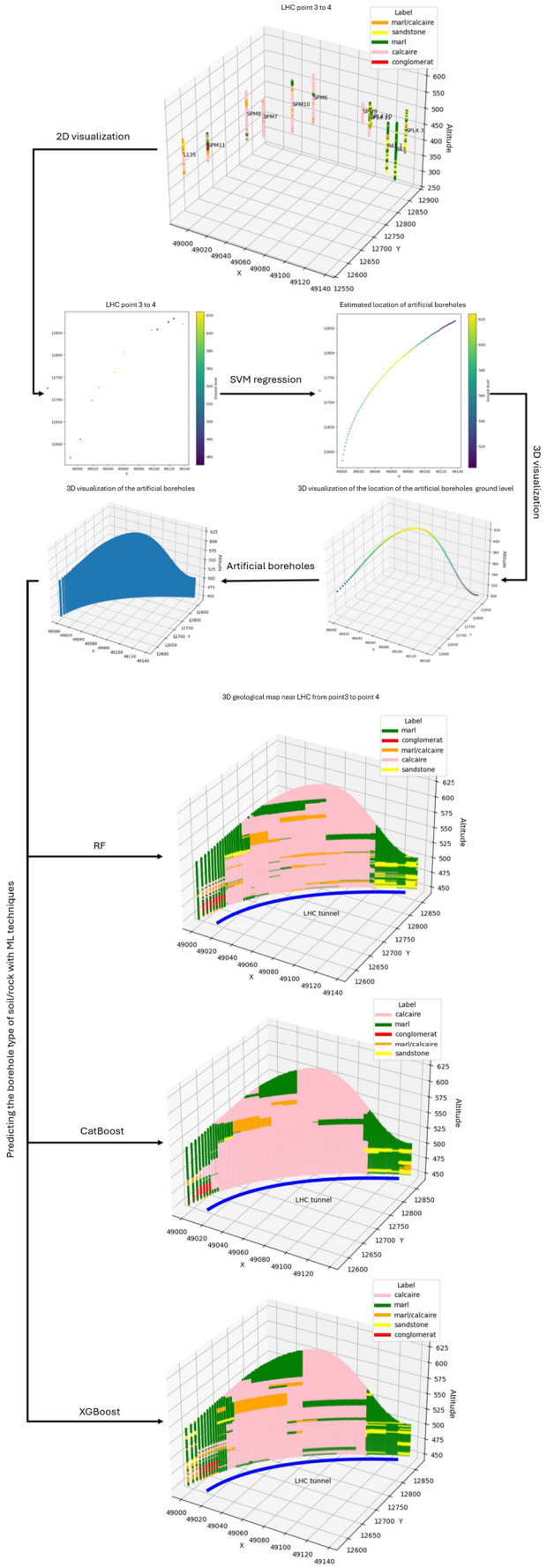

Figure 9. Flowchart of building the geological map using RF, CatBoost, and XGBoost

Figure 9 demonstrates the developed maps using different ML algorithms and the steps taken in this process. RF attempts to predict some thin inner layers as “marl” and “marl/calcaire”, while the CatBoost prefers to predict most of the datapoints in the middle area as “calcaire”. XGBoost was performed with caution and had some “marl” predictions in the middle of the map.

4 CONCLUSIONS

The presented study shows the potential of ML algorithms for building geological maps based on borehole log reports and only using the borehole's spatial information. In this study, the unstructured PDF files were converted to structured Excel files, preprocessed, and integrated to be used as the input of the models.

Among the ML techniques, RF, SVM, CatBoost, and XGBoost were deployed on the available data. The SVM couldn’t perform well and predicted all the data points in one of the most repeated labels. However, the performance of the tree-based models (RF, XGBoost, and CatBoost) was almost the same and can be evaluated more precisely with a larger dataset. Moreover, the results illustrated that the integration of all datasets improved the model's training process and led to a better performance. In conclusion, using the same coordinates for all boreholes and evenly distributed boreholes can result in more accurate geological maps. The more data points and boreholes provided, the better ML algorithms can learn the correlations between features, leading to improved performance.

In future studies, some numerical analysis can be deployed on the coordinate features as additional features to improve predictions. In addition, the developed geological map can be used to assess the potential impact on the tunnel infrastructure and the relationship with the defects.

5 ACKNOWLEDGEMENTS

This publication has emanated from research conducted with financial support of Research Ireland under Grant 13/RC/2092_2 and 12/RC/2289-P2, and CERN. For the purpose of Open Access, the author has applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission.

6 REFERENCES

Charollais, J., Weidmann, M., Berger, J.-P., Engesser, B., Hotellier, J.-F., Gorin, G., Reichenbacher, B., & Schäfer, P. (2007). *La Molasse du bassin franco-genevois et son substratum*.

Cracknell, M. J., & Reading, A. M. (2014). Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information. *Computers & Geosciences*, *63*, 22–33. https://doi.org/10.1016/J.CAGEO.2013.10.008

Fern, E. J., Di Murro, V., Soga, K., Li, Z., Scibile, L., & Osborne, J. A. (2018). Geotechnical characterisation of a weak sedimentary rock mass at CERN, Geneva. *Tunnelling and Underground Space Technology*, *77*, 249–260. https://doi.org/10.1016/J.TUST.2018.04.003

Guo, J. T., Liu, Y. H., Han, Y. F., & Wang, X. L. (2019). Implicit 3D Geological Modeling Method for Borehole Data Based on Machine Learning. *Dongbei Daxue Xuebao/Journal of Northeastern University*, *40*(9). https://doi.org/10.12068/j.issn.1005-3026.2019.09.021

He, X., Chaney, N. W., Schleiss, M., & Sheffield, J. (2016). Spatial downscaling of precipitation using adaptable random forests. *Water Resources Research*, *52*(10), 8217–8237. https://doi.org/10.1002/2016WR019034

Hu, Y., Wang, Z. Z., Guo, X., Kek, H. Y., Ku, T., Goh, S. H., Leung, C. F., Tan, E., & Zhang, Y. (2024). Three-dimensional reconstruction of subsurface stratigraphy using machine learning with neighborhood aggregation. *Engineering Geology*, *337*, 107588. https://doi.org/10.1016/J.ENGGEO.2024.107588

Kim, H. S., & Ji, Y. (2022). Three-dimensional geotechnical-layer mapping in Seoul using borehole database and deep neural network-based model. *Engineering Geology*, *297*, 106489. https://doi.org/10.1016/J.ENGGEO.2021.106489

Large Colliders:Civil engineering and Siting. (n.d.).

Li, L., Wang, G., Chen, D., Wang, X., Bai, J., Lei, B., Ding, W., Huang, X., & Zhang, Q. (2024). Analysis of tunnel collapse disasters during operation and exploration of disaster damage mechanisms. *IOP Conference Series: Earth and Environmental Science*, *1333*(1), 012049. https://doi.org/10.1088/1755-1315/1333/1/012049

Liu, L., Li, T., & Ma, C. (2024). Research on 3D Geological Modeling Method Based on Deep Neural Networks for Drilling Data. *Applied Sciences*, *14*(1). https://doi.org/10.3390/app14010423

Lyu, B., & Wang, Y. (2024). Immersive visualization of 3D subsurface ground model developed from sparse boreholes using virtual reality (VR). *Underground Space*, *17*, 188–206. https://doi.org/10.1016/J.UNDSP.2023.11.004

Lyu, B., Wang, Y., & Shi, C. (2024). Multi-scale generative adversarial networks (GAN) for generation of three-dimensional subsurface geological models from limited boreholes and prior geological knowledge. *Computers and Geotechnics*, *170*, 106336. https://doi.org/10.1016/J.COMPGEO.2024.106336

Mathur, A., & Foody, G. M. (2008). Multiclass and Binary SVM Classification: Implications for Training and Classification Users. *IEEE Geoscience and Remote Sensing Letters*, *5*(2), 241–245. https://doi.org/10.1109/LGRS.2008.915597

Pandey, M., Karbasi, M., Jamei, M., Malik, A., & Pu, J. H. (2023). A Comprehensive Experimental and Computational Investigation on Estimation of Scour Depth at Bridge Abutment: Emerging Ensemble Intelligent Systems. *Water Resources Management*, *37*(9), 3745–3767. https://doi.org/10.1007/S11269-023-03525-W

Shi, C., & Wang, Y. (2022). Machine learning of three-dimensional subsurface geological model for a reclamation site in Hong Kong. *Bulletin of Engineering Geology and the Environment*, *81*(12), 1–18. https://doi.org/10.1007/S10064-022-03009-Y/FIGURES/12

Wang, C., Friedman, M., Wu, W., Zhang, D., & Li, Z. (2024). Hydraulic influences on the long-term performance of tunnels: A review. *Transportation Geotechnics*, *48*, 101329. https://doi.org/10.1016/J.TRGEO.2024.101329

Zhu, L., Zhang, C., Li, M., Pan, X., & Sun, J. (2012). Building 3D solid models of sedimentary stratigraphic systems from borehole data: An automatic method and case studies. *Engineering Geology*, *127*, 1–13. https://doi.org/10.1016/J.ENGGEO.2011.12.001