

Automatic Rock Recognition from Drilling Performance Data

Hang Zhou, Peter Hatherly, Sildomar T. Monteiro, Fabio Ramos,
Florian Oppolzer, Eric Nettleton and Steve Scheduling

Abstract—Automated rock recognition is a key step for building a fully autonomous mine. When characterizing rock types from drill performance data, the main challenge is that there is not an obvious one-to-one correspondence between the two. In this paper, a hybrid rock recognition approach is proposed which combines Gaussian Process (GP) regression with clustering. Drill performance data is also known as Measurement While Drilling (MWD) data and a rock hardness measure - Adjusted Penetration Rate (APR) is extracted using the raw data in discrete drill holes. GP regression is then applied to create a more dense APR distribution, followed by clustering which produces discrete class labels. No initial labelling is needed. Comparisons are made with alternative measures of rock hardness from MWD data as well as state-of-the-art GP classification. Experimental results from an actual mine site show the effectiveness of our proposed approach.

I. INTRODUCTION

Successful automation of mining activities represents a goal for mining companies. Of the challenges facing mine automation, the construction of representations of the unmined geology so as to determine the quantity and quality of the minerals of interest, represents a significant issue. To help address this challenge, we develop automated rock recognition solutions to extract useful properties such as rock hardness and rock type from measurements of the operation and performance of the drills used to drill blast holes. These holes are drilled on a relatively close grid of 5-6m so that the blasting produces uniform rock breakage. The drills may be operated in manual or autonomous mode. The measurements are known as Measurement While Drilling (MWD) data.

Automated rock recognition provides information that can be used in optimization of mine operations as well as mine planning and design [4]. It can assist the mining industry operate in a more efficient, cost effective and safe manner. For instance, rock type boundary detection is important for blast design as well as general strategic planning.

Conventionally, rock recognition is performed manually by geologists from rock samples collected on site. It is a hard work - time intensive and expensive. This motivates the seeking of an automated solution. One feasible approach is to automate rock recognition from the MWD data that are routinely collected from drilling. MWD data consist of a group of measurements that are primarily used to control and monitor the drilling process. They are collected from the sensors mounted on large drill rigs for blast hole drilling. Figure 1 is an example of the autonomous blast hole drill rig.

The idea of relating drilling measurements to the properties of the rocks has been studied using empirical and statistical approaches [7][5][6][10][9]. Teale's Specific Energy of



Fig. 1. Autonomous blast hole drill rig used for collecting experimental MWD data in this paper.

Drilling (SED) [7] is most widely cited. Machine learning techniques, including unsupervised learning and supervised learning, have also been applied. In [8], an unsupervised learning method is used to classify the geological patterns. Various supervised learning approaches have been developed as well, such as Neural Networks (NN) [11][13][12][4], Conditional Random Field (CRF) model [14][15] and GP classification [16].

Supervised learning methods classify rock types of individual blast holes on the basis of models trained from the labelled MWD data. MWD data are labelled by experienced geologists using other geological data and interpretations. Since it is very difficult for geologists to reliably categorize the rock type changes within one blast hole, the entire hole is assigned a rock type label. A classifier is then trained and used to classify any new coming MWD data. In our context, rocks in each blast hole are classified into one of the three types, i.e. shale, ore and BIF (Banded Iron Formation).

For supervised learning, the major difficulty in accurately predicting rock types from MWD data lies in the inhomogeneity of the rocks. It is possible for one rock type to have different MWD values and for one set of MWD values to represent quite different rock types.

To address this difficulty, we propose an adaptive solution with the framework shown in Figure 2. Instead of directly relating MWD data to a specific rock type, we build the connection between MWD data and rock types in two steps. MWD data are firstly mapped to the rock hardness followed

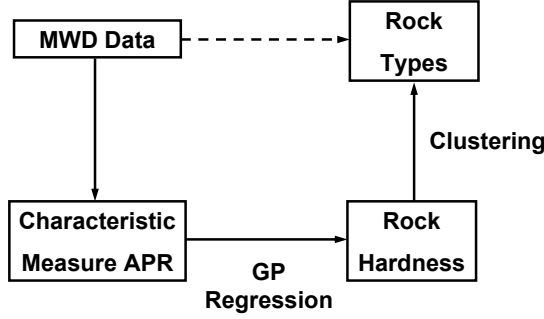


Fig. 2. The framework of the proposed hybrid GP regression and clustering approach for automatic rock recognition.

by another mapping from the rock hardness to rock types. A generalized rock hardness measure - Adjusted Penetration Rate (APR) is extracted from the discrete blast hole MWD data, with which GP regression [1][2] is applied to get a more dense 3D rock hardness distribution. From the derived APR values, general identification of rock types can be attempted. Of the three main rock types present at the site of our work, shale is the softest, BIF is the hardest and ore is in between.

The main contributions of our work are:

- 1) Our proposed approach successfully creates a rock type model from MWD data, where there is not a direct connection between the input (MWD data) and the output (rock types). It reliably classifies rock types without initial labelling.
- 2) A continuous 3D model of rock hardness distribution across a mining site can be obtained by applying GP regression on the rock hardness measure APR from multiple holes. This makes it possible to have a more accurate overall description of the rock hardness.
- 3) In contrast to all previous methods of analyzing rock properties, we introduce and define a new parameter Adjusted Penetration Rate (APR). APR is simply the Penetration Rate divided by the product of Pulldown Pressure and square root of the Rotation Pressure (see Section III for details of MWD data). This makes it possible to reveal the inherent connection between the MWD data and rock hardness, which is not explicitly reflected on the raw MWD data due to the interdependence between them.

For easy reference, major acronyms related to drilling and geology in this paper are listed below,

- 1) APR: Adjusted Penetration Rate
- 2) BIF: Banded Iron Formation
- 3) MWD: Measurement While Drilling
- 4) PR: Penetration Rate
- 5) PP: Pulldown Pressure
- 6) RP: Rotation Pressure
- 7) RS: Rotation Speed
- 8) SED: Specific Energy of Drilling

The rest of the paper is organized as follows. GP regression and GP classification are briefly reviewed in Section II. In Section III, our proposed APR is defined and the APR based rock recognition algorithm is presented. Section IV shows the experimental results. Finally, Section V summarizes the main conclusions.

II. GAUSSIAN PROCESS (GP) REGRESSION AND CLASSIFICATION

In our work, we use Gaussian Process (GP) regression to generate the continuous rock hardness distribution. The results are then compared between GP classification and our proposed method. Both GP regression and GP classification are briefly described as follows.

A GP is a collection of random variables, any finite number of which have a joint Gaussian distribution. It is fully specified by its mean function $\mu(\mathbf{x})$ and kernel function $K(\mathbf{x}, \mathbf{x}')$, i.e., $f \sim \text{GP}(\mu, K)$ [1].

Given a n dimensional dataset with m observations $\{(\mathbf{x}_i, y_i), i = 1, 2, \dots, m\}$, where \mathbf{x} is the n dimensional input vector, comprising the $n \times m$ input matrix X , and y is the scalar as part of the output vector \mathbf{y} , the GP regression mean predictive distribution can be expressed as [1]

$$\bar{\mathbf{f}}_* = K(X_*, X)[K(X, X) + \sigma_n^2 I]^{-1} \mathbf{y}, \quad (1)$$

where X_* is the testing data matrix, $K(X, X)$ is the kernel matrix of the training data, $K(X_*, X)$ is the kernel matrix of the training as well as testing points and σ_n^2 is the Gaussian noise variance of the observations.

If \mathbf{y} has a discrete value of $\{1, -1\}$, it turns out to be a GP binary classification problem. The probabilistic interpretation for the output is [1],

$$p(y_* = +1 | X, \mathbf{y}, \mathbf{x}_*) = \int g(f_*) p(f_* | X, \mathbf{y}, \mathbf{x}_*) df_*, \quad (2)$$

where $p(f_* | X, \mathbf{y}, \mathbf{x}_*)$ is the distribution over the latent function f corresponding to the test case f_* , g can be any sigmoid function that ‘squashes’ the prediction output to guarantee a valid probabilistic value within the range of $[0, 1]$.

III. THE CHARACTERIZING MEASURE: ADJUSTED PENETRATION RATE (APR)

For the MWD data we work with, the data typically includes Penetration Rate (PR), Pulldown Pressure (PP), Torque (RP, reflecting Rotation Pressure), Rotation Speed (RS) and Bit Air Pressure (BAP). Of these, PR, PP and RP are the measurements used in our proposed approach. As indicated in Section I, due to geological variability, there is not a one-to-one correspondence between the MWD data and the rock types. In addition, a full mathematical description of the interaction between the drill bit and the rock is unavailable. Taking these difficulties into consideration, we explore the physical meaning beneath the MWD data and aim to extract a characteristic measure that not only captures the inherent rock hardness but also explicitly relates MWD data to rock types.

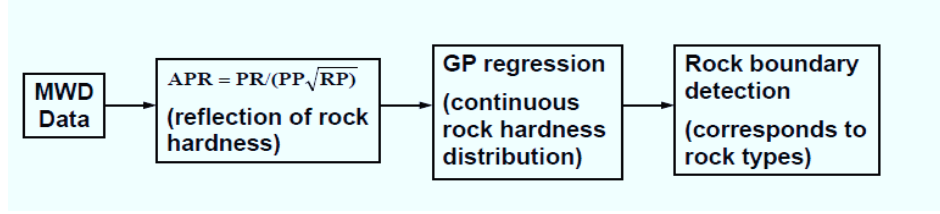


Fig. 3. The major steps of the automatic rock recognition algorithm.

Of the various MWD measurements, PR is a key measurement and responds to the rock hardness as well as the applied forces of PP and RP. Two assumptions are made:

- 1) With constant PP and RP, PR reflects the rock hardness. The lower the PR, the harder the rock and vice versa.
- 2) The rock hardness is generally related to the rock type such that the rock hardness decreases from BIF to ore and then to shale.

However, for both manual drilling and autonomous drilling, PP and RP are variable. We need to define a measure that can capture the inherent properties of the rock regardless of the changes in PP and RP.

We utilize the concept of “capacity factor” from [3]. The capacity factor K_{tc} is defined in eq.(3) as rotation speed (RS) divided by the square root of torque (RP),

$$K_{tc} = \frac{RS}{\sqrt{RP}} \quad (3)$$

As a performance parameter of a torque converter for automatic transmissions, the capacity factor indicates the ability of the converter to generate the observed rotation speed with the applied torque. It is a reflection of the resistance of the fluid in the converter.

In our case, we look for a measure that represents the resistance of the rock being drilled. Similar to the torque converter, this measure can be defined as the ability to generate the observed speed with the applied external forces. What makes it different from the torque converter is that the speed we are concerned with is the vertical PR and the applied forces are from both PP and RP.

As shown in eq.(4), we decompose the definition into three parts. Firstly, similar to eq.(3), we define a capacity factor which indicates the ability of achieving the RS with the applied RP. Since the change of RS and PR are correlated, we then define a ratio of PR over RS, indicating how much PR can be obtained out of RS. Putting the first two parts together, it is equivalent to the capacity factor w.r.t. PR. However, PR is generated by both PP and RP. Therefore, the product of the first two parts is further corrected by PP. In this way, the resistance of the rock is defined. We name this definition as the Adjusted Penetration Rate (APR).

$$APR = \frac{RS}{\sqrt{RP}} \cdot \frac{PR}{RS} \cdot \frac{1}{PP} = \frac{PR}{PP\sqrt{RP}} \quad (4)$$

Looking at eq.(4) from the numerical point of view, it

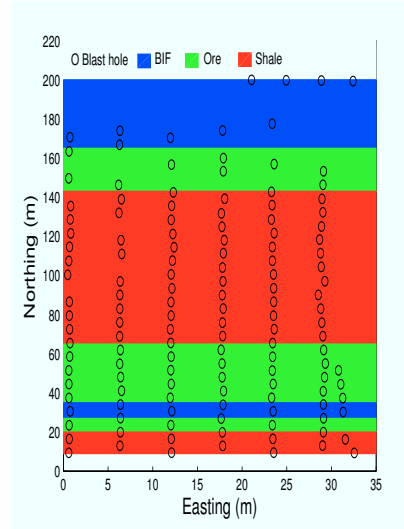


Fig. 4. Blast holes as well as the assumed rock types “ground truth” for our testing.

is different from previous empirical equations [7][6][10][9], because we compensate the PR by dividing it by the product of PP and RP square root, to establish APR¹. In this way, APR explicitly links MWD data with rock hardness, given the two assumptions listed above.

As mentioned in Section I, the blast holes from which the MWD data collected are 5-6m apart. Thus, the calculated APR values (using eq.(4)) are also discretely and sparsely located. At such spacing, it is difficult to identify the exact boundaries between the rock types. With regards to that, we propose to apply GP regression on the APR values of discrete blast hole to obtain a more continuous (dense) distribution of the rock hardness. We take this to be the 3D continuous rock hardness distribution, with which the rock type boundaries can be better identified and extracted by clustering. A summary of our proposed rock recognition algorithm is shown in Figure 3.

IV. EXPERIMENTS AND RESULTS

A. APR Rock Hardness Distribution

We have tested our approach using MWD data collected from an iron ore mine in Australia. Figure 4 shows the

¹Regarding the unit of APR, it is beyond the scope of this paper. In our context, APR is used as a ratio which separates different rock types. Hence, it can be viewed as having similar significance to raw PR.

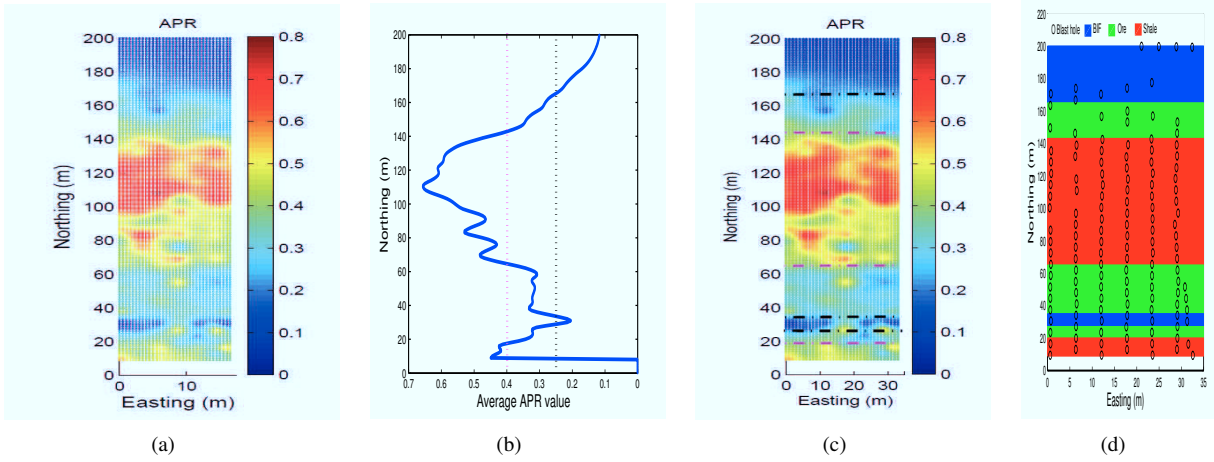


Fig. 5. Boundary detection over hardness distribution. (a) Projection of rock hardness distribution on a 2D plane. (b) Projection of rock hardness distribution into 1D. Black line: BIF-ore boundary threshold, pink line: ore-shale boundary threshold. (c) Detected rock boundary. Black line: BIF-ore boundary, pink line: ore-shale boundary. (d) Rock types 'ground truth'.

blast holes as well as the assumed rock types “ground truth” for one of our tests. It covers typical transitions between the three major rock types we are concerned with. The “ground truth” was provided by an experienced geologist using various geological data sources, e.g., chemistry and geological logging analysis. This is assumed to represent a reasonable approximation of the true geology. In Figure 4, the redder colours correspond to higher APR values (lower rock hardness) and blue corresponds to lower APR values (higher rock hardness). Since the rock hardness increases from shale to ore and then to BIF, the lowest hardness red colour zones correspond to shale, the highest hardness blue zones are the BIF and the medium hardness, green, represents the ore.

Our proposed APR rock recognition method was applied to the test site and the results are shown in Figure 5. Figure 5(a) is the APR rock hardness distribution projected onto a horizontal 2D plane. A further 1D projection along the north axis is shown in Figure 5(b). From this 1D projection, the rock types are extracted using preset thresholds. By incorporating the established rock types, the threshold for the BIF-ore boundary is set to 0.25 and the ore-shale boundary is 0.4. The extracted boundaries are shown in Figure 5(c), where the black lines are the BIF-ore boundaries and the pink lines are the ore-shale boundaries. Compared with the “ground truth” in Figure 5(d), it can be seen that the results are satisfactory.

In Figure 6, a comparison is made between the rock hardness distributions based on raw PR and APR. By comparing with the “ground truth” in Figure 6(c), we can see that the results of APR are clearly more consistent with the actual geology.

B. Comparison of APR with an Alternative Measurement - SED

A comparison has also been made between APR and an existing widely cited rock property characterizing measure -

Specific Energy of Drilling (SED) [7], which is defined as,

$$e = \frac{F}{A} + \frac{2\pi NT}{Au}, \quad (5)$$

where F is the thrust (equivalent to PP in our context), N is our RS, T is torque (equivalent to RP in our context), u is our PR and A is the cross-sectional area of the hole which is a constant.

SED is an indication of the energy needed when drilling rocks. Therefore, the SED value has the opposite trend to the APR value, e.g., shale has the lowest hardness and the corresponding SED is low, while APR is high.

In Figure 7, a comparison is made between APR and SED. On the basis of the “ground truth” shown in Figure 7(c), it can be seen that the SED results are less accurate. Figure 7(d) also shows that compared with APR, SED is much less sensitive to the changes in rock type that occur in the northerly direction.

C. Comparison with GP Classification Results

In addition, the rock types and boundaries at the test site were also detected using GP classification [1][2]. The test was performed in two ways: cross validation within the dataset itself as well as using data from three neighbouring test sites as training datasets. To guarantee a more reliable labelling of the training data, for either the cross validation dataset or the training datasets from the three neighbouring test sites, the blast holes near the boundaries were removed.

The classification results are plotted in Figure 8. By comparing with the “ground truth” in Figure 8(c), it can be seen that even with cross validation within the dataset itself, GP classification can only correctly detect most of the shale zones and part of the ore zones. The BIF zone is totally missed. With the GP classification trained by the neighbouring test sites, the results are completely misleading. The shale zone is interpreted as ore and the ore is interpreted as BIF. This result indicates that even with the removal of

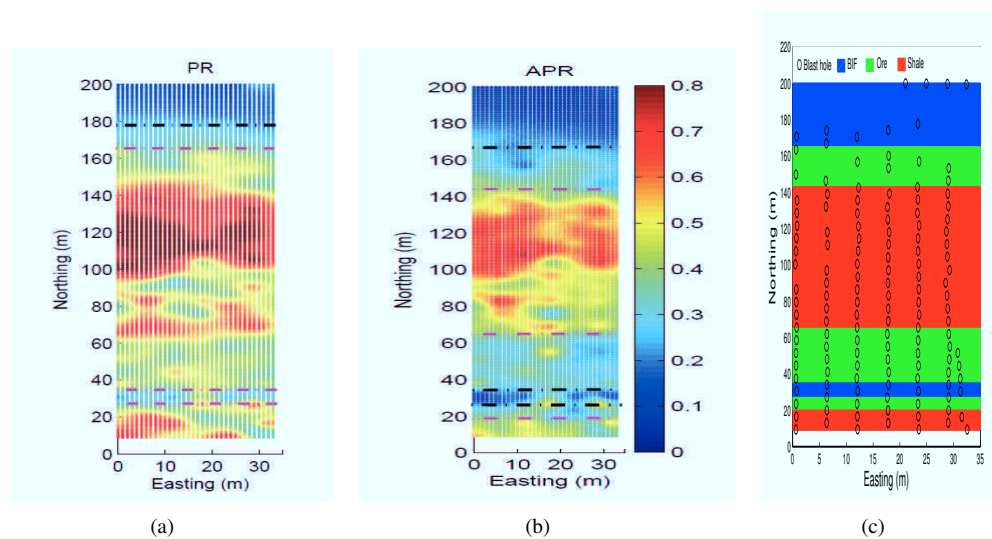


Fig. 6. Comparison of GP regression results based on raw PR and APR. (a) GP regression result on raw PR. Black line: BIF-ore boundary threshold, pink line: ore-shale boundary. (b) GP regression result on APR. Black line: BIF-ore boundary threshold, pink line: ore-shale boundary. (c) Rock types 'ground truth'.

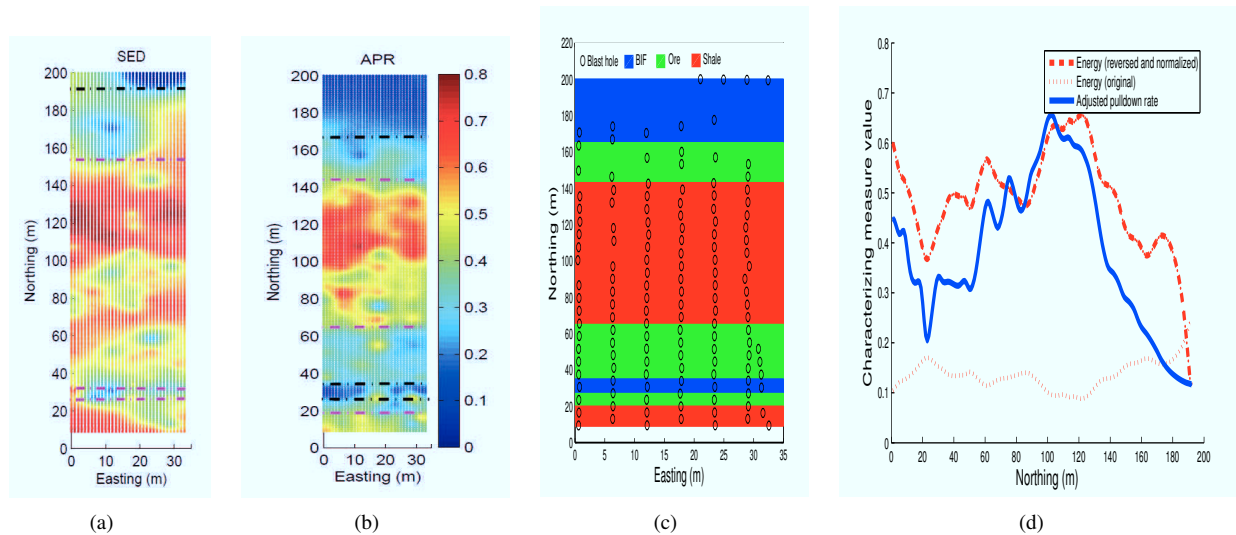


Fig. 7. Comparison between APR and SED. (a) Rock boundary of SED (reversed and normalized). Black line: BIF-ore boundary, pink line: ore-shale boundary. (b) Rock boundary of APR. Black line: BIF-ore boundary, pink line: ore-shale boundary. (c) Rock types 'ground truth'. (d) 1D projection of APR and SED distributions.

the holes of high uncertainty, the classification results are unsatisfactory.

D. Discussion

To summarize, extensive comparisons of results from actual MWD datasets have shown that our hybrid GP regression and clustering approach for rock recognition using APR has stable and satisfactory performance.

Currently, straight line boundaries are used to separate different rock types. Since the ground truth of an unmined site is never precisely known due to the local and more general variations in the geology, such an approximated separation has already provided valuable information for mine operations. Further work could be done to extract more flexible boundary curves together with the development of a

solid way of validation from the geological point of view.

It should also be noted that our proposed approach represents an attempt to handle situations when there is not a clear causal relation between the input data features and the expected output class labels. Mapping from the raw MWD data to APR tunes the feature(s) more explicitly related to the output (rock type) labels. GP regression on APR encompasses the inherent correlation between neighbouring APR values to properly reconstruct the feature space at a desired resolution. As a result, it leads to a more reliable and accurate classification result. All these allow the subsequent clustering to be implemented in a very simple way. Hence, this approach has the potential of being applied to other applications, where a principal measure can be extracted through a good analysis and understanding of the underlying

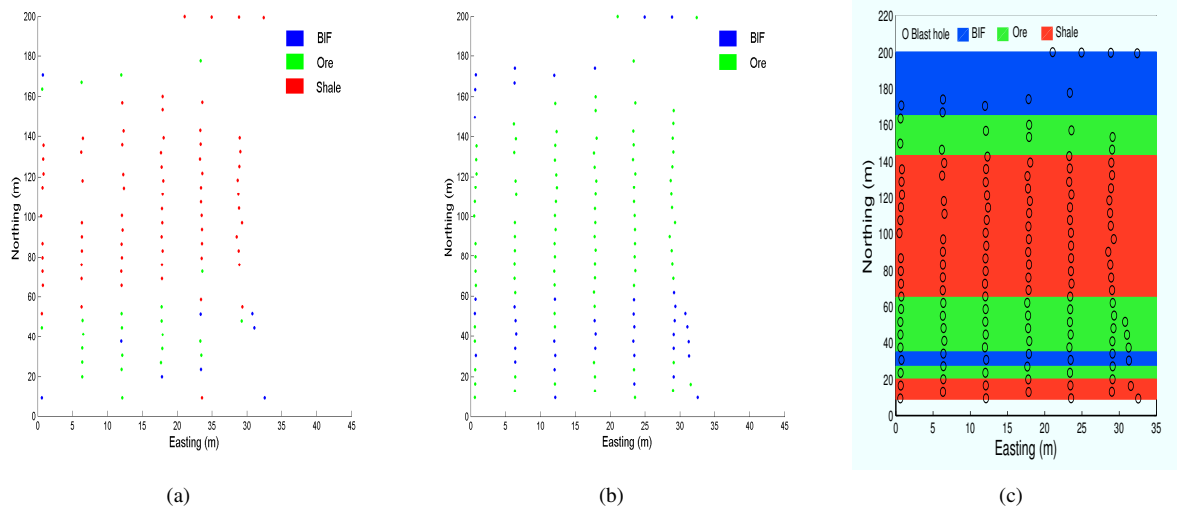


Fig. 8. 2D projection of Gaussian Process classification results. Gaps have been left in the vicinity of the boundaries. (a) Cross validation results within one bench. (b) Classification results using training data from three neighbouring benches. (c) Rock boundary 'ground truth'.

physical meaning, together with a GP regression to create a more separable new feature space.

V. CONCLUSIONS

A novel hybrid approach has been proposed to reliably categorize rock types from MWD data. It captures the rock hardness measure - Adjusted Penetration Rate (APR). From the APR collected at discrete blast holes, a continuous rock hardness distribution is created by GP regression, which can then be reliably related to rock types by clustering. The proposed solution does not rely on initial labelling. It works well for the geological conditions present at the iron ore mine used for our tests.

VI. ACKNOWLEDGMENTS

This work has been supported by the Australian Centre for Field Robotics and the Rio Tinto Centre for Mine Automation. The authors also acknowledge the support of Danielle Robinson, Annette Pal and Mike Wake of Rio Tinto, as well as Ross Hennessy and Surya Singh of Rio Tinto Centre for Mine Automation.

REFERENCES

- [1] C.E. Rasmussen and C.K.I. Williams, *Gaussian Processes for Machine Learning*, The MIT Press, Cambridge, Massachusetts; 2006.
- [2] N. D. Lawrence, M. Seeger and R. Herbrich, Fast Sparse Gaussian Process Methods, The Informative Vector Machine, *Advances in Neural Information Processing Systems*, vol. 15, 2003, pp. 609-616.
- [3] J. Y. Wong, *Theory of Ground Vehicles*, John Wiley & Sons; 2001.
- [4] J. E. J. M. Gonzalez, Application of Pattern Recognition Techniques to Monitoring-While-Drilling on a Rotary Electric Blasthole Drill at an Open-Pit Coal Mine, *Thesis for Master of Science*, Queen's University, Canada, 2007.
- [5] M.J. Scoble and J. Peck, A Technique for Ground Characterization Using Automated Production Drill Monitoring, *International Journal of Mining, Reclamation and Environment*, vol. 1(1), 1987, pp. 41-54.
- [6] M.J. Scoble, J. Peck and C. Hendricks, Correlation between Rotary Drill Performance Parameters and Borehole Geophysical Logging, *Mining Science and Technology*, vol. 8, 1989, pp. 301-312.
- [7] R. Teale, The Concept of Specific Energy in Rock Drilling, *International Journal of Rock Mechanics and Mining Sciences*, vol. 2, 1965, pp. 57-73.
- [8] R.L. King, M.A. Hicks, S.P. Signer, Using Unsupervised Learning for Feature Detection in a Coal Mine Roof, *Engineering Applications of Artificial Intelligence*, vol. 6, 1993, pp. 565-573.
- [9] H. Schunnesson, Rock Characterisation Using Percussive Drilling, *International Journal of Rock Mechanics and Mining Sciences*, vol. 35(6), 1998, pp. 711-725.
- [10] G.L. Finfinger, G. Wilson, S. Peng, B. Thomas and Q. Gu, "An Approach to Identifying Geological Properties from Roof Bolter Drilling Parameters", *Conference on Ground Control in Mining*, 2000.
- [11] W.K. Utt, "Neural Network Technology for Strata Strength Characterization", *International Joint Conference on Neural Networks*, 1999.
- [12] G.L. Finfinger, A Methodology for Determining the Character of Mine Roof Rocks, *PhD Thesis*, West Virginia University, U.S.A., 2003.
- [13] D. Labelle, J. Bares and I. Nourbakhsh, "Material Classification by Drilling", *17th International Symposium on Automation and Robotics in Construction*, 2000.
- [14] S. Monteiro, F. Ramos and P. Hatherly, "Learning CRF Models from Drill Rig Sensors for Autonomous Mining", *NIPS workshop: Learning from Multiple Sources with Applications to Robotics*, 2009.
- [15] S. Monteiro, F. Ramos and P. Hatherly, "Conditional Random Fields for Rock Characterization Using Drill Measurements", *In Proceedings of International Conference on Machine Learning and Applications*, 2009.
- [16] H. Zhou, S. Monteiro, P. Hatherly, F. Ramos, F. Oppolzer and E. Nettleton, "Spectral Feature Selection for Automated Rock Recognition Using Gaussian Process Classification", *In Proceedings of Australian Conference on Robotics and Automation*, 2009.