UNIVERSIDADE DO VALE DO RIO DOS SINOS - UNISINOS

UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO

DOUTORADO EM GEOLOGIA

THIAGO SANTI BRESSAN

APLICAÇÃO DE INTELIGÊNCIA ARTIFICIAL E MACHINE LEARNING EM DADOS LITOESTRATIGRÁFICOS E GEOFÍSICOS DAS EXPEDIÇÕES DO PROGRAMA INTERNACIONAL DE DESCOBERTAS OCEÂNICAS (IODP)

ORIENTADOR: PROF. DR. FARID CHEMALE Jr.

São Leopoldo

2021

THIAGO SANTI BRESSAN

Aplicação de Inteligência Artificial e Machine Learning em dados litoestratigráficos e geofísicos das Expedições do Programa Internacional de Descobertas Oceânicas (IODP)

Tese de Doutorado apresentada como requisito para obtenção do título de doutor em Geologia, pelo Programa de Pós-graduação em geologia da Universidade do Vale do Rio dos Sinos – UNISINOS

Orientador (es): Profs. Drs. Farid Chemale Jr. e Sandro Jose Rigo

Banca Examinadora:

Prof. Drª. Joice Cagliari
PPGEO – Unisinos

Prof. Drª. Adalene Moreira Silva
Instituto de Geociências – Universidade de Brasília

Dr. Luiz Gustavo Rodrigues Pinto
Serviço Geológico do Brasil - CPRM

São Leopoldo

2021

B843a Bressan, Thiago Santi.

Aplicação de inteligência artificial e machine learning em dados litoestratigráficos e geofísicos das expedições do Programa Internacional de Descobertas Oceânicas (IODP) / Thiago Santi Bressan. – 2021.

257 f. : il. ; 30 cm.

Tese (doutorado) – Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Geologia 2021.

"Orientador (es): Profs. Drs. Farid Chemale Jr. e Sandro Jose Rigo."

1. Inteligência Artificial. 2. Geofísicos. 3. Programa Internacional de Descobertas Oceânicas (IODP). I. Título.

CDU 55

Dados Internacionais de Catalogação na Publicação (CIP)

(Bibliotecária: Amanda Schuster – CRB 10/2517)

AGRADECIMENTOS

Em especial ao meu orientador prof. Dr. Farid Chemale Jr, que com seu conhecimento e grande experiência na área acadêmica e profissional, ajudou no desenvolvimento da tese e na publicação dos artigos. Muito obrigado

Ao prof. Dr. Osmar Gustavo Wohl Coelho, que tive o primeiro contato e ajudou muito na elaboração e definição da proposta inicial do projeto do doutorado. Muito obrigado

Aos demais professores do PPGEO-Unisinos pelos conhecimentos transmitidos durante todo o período das aulas e dos campos práticos.

Aos familiares, em especial a minha esposa Simone e minhas filhas Eduarda e Helena, pela compreensão e suporte familiar neste período dedicado a conclusão do doutorado. A família é essencial e capaz de criar todo o ambiente favorável a construção do conhecimento integrando e interagindo com a realidade, principalmente neste momento difícil da pandemia pela COVID-19.

Ao International Ocean Discovery Program (IODP) e aos cientistas que participaram da expedição 362 pelo ótimo trabalho desenvolvido na elaboração e disponibilização dos dados.

A UNISINOS e todos os setores que ajudaram de alguma forma na finalização deste grande sonho construído durante 3 anos e 8 meses.

Ao Instituto Federal Farroupilha pelo apoio durante o período de realização do doutorado junto com as aulas presenciais e, através da participação de edital específico, o afastamento integral para conclusão do doutorado.

E por fim, a CAPES pelo financiamento da minha pesquisa através da bolsa auxílio taxa.

RESUMO

Ferramentas computacionais específicas auxiliam o geólogo a identificar litologias e o empilhamento estratigráfico em perfuração de poços, reduzindo custos operacionais e gerenciando tempo de trabalho prático dos profissionais, direcionando-os para interpretações eficientes de dados ou mesmo no aprimoramento das pesquisas científicas na região o trabalho ou em regiões geologicamente distintas. Neste estudo, a aplicação de algoritmos de aprendizado de máquina para a classificação supervisionada de litologias foi avaliada. Foram utilizados dados de registros de parâmetros multivariados em poços offshore, relacionados ao Programa Internacional de Descobertas Oceânicas (IODP) com dados supervisionados e não supervisionados (imagens), com a criação de um contexto de aplicação híbrida de algoritmos, dividida em dois manuscritos. No manuscrito I, por meio da análise das litologias propostas em 7 Expedições IODP e do uso da utilização dos algoritmos, foi possível agrupar e dividir os conjuntos litológicos em quatro grupos de litologias e templates. As propriedades geofísicas utilizadas no presente estudos incluíram GRA, PWL, MS, RSC e SRM. Os templates foram submetidos a treinamento e testes pelos métodos Multi-Layer Perceptron (MLP), DecisionTree, RandomForest e Support Vector Machine (SVM) utilizando como avaliação dos resultados as métricas de classificação. Como resultado, observou-se que Template1 obteve melhores resultados no algoritmo MLP, Template2 e Template3 obtiveram melhores resultados no algoritmo RandomForest acima de 80,00% de acurácia. Para Cross-validation, o algoritmo RandomForest obteve ótimo desempenho em todos os cenários. No Practical Template, o grupo de litologia G2 obteve melhor resultado com o algoritmo MLP com acurácia média superior a 85,00%. Para o manuscrito II, a divisão dos dados incluiu a formatação de três conjuntos de dados: dataset0, dataset1 e dataset2, especificamente com os dados obtidos durante a Expedição IODP 362. As propriedades petrofísicas utilizadas incluíram PWL, GRA, RSC, NGR, MAD, MS, RGB e imagens de alta definição. Dataset0 incluiu dados de formação temporária para validar o melhor interpolador. O dataset1 possui os dados interpolados das propriedades petrofísicas, perfazendo um total de 295.945 registros para U1480 e U1481 com 17 features. O dataset2 abrange os dados de textura e cor extraídos da segmentação das imagens, perfazendo um total de 85.058 registros para o U1480 e U1481, com 90 features. Cada conjunto de dados é replicado em dois grupos de litologia: Grupo 1 e Grupo 2. Para o dataset2, novas combinações são adicionadas entre as features formando 102 arranjos práticos com resultados específicos em cada combinação. Os valores foram interpolados por Linear, Spline, Slinear, Quadratic, Cubic, Akima, Pchip e Piecewise. O método de aprendizado de máquina empregado para todos os datasets é o RamdomForest. Os resultados mostram que o melhor interpolador avaliado no dataset0 é o Akima com acurácia igual a 98,22%. Para dataset1, U1480, o valor de acurácia é 96,96% na combinação de 70% de treinamento e 30% de teste no Grupo 1 e 97,71% na combinação de 70% de treinamento e 30% de teste no Grupo 2. Para dataset1, U1481, o valor de acurácia é 99,68% na combinação de 80% de treinamento e 20% de teste no Grupo 1 e 99,74% na

combinação de 80% de treinamento e 20% de teste no Grupo 2. Para o dataset2, os arranjos práticos melhores avaliados são o 51 (Grupo 1) e o 102 (Grupo 2) para U1480 e os arranjos práticos 32, 33, 44, 45 e 51 (Grupo 1) e 102 (Grupo 2) para U1481. Em relação ao novo método Area Superpixel (Apx), os melhores resultados estão no Grupo 1 com maior combinação de treinamento e menor combinação de testes. Os conjuntos de dados avaliados foram agrupados numa organização entre os locais da expedição, em que se gerou um contexto de aplicação prática real na atividade diária do geólogo com excelentes resultados de classificação litológica. A interpolação de propriedades petrofísicas é válida e necessária quando há poucos dados para treinamento respeitando as características de cada propriedade e interpolador. Propriedades extraídas das imagens são relevantes e agrupadas em conjunto com as propriedades petrofísicas criam um contexto de extrema importância na descoberta e apresentação da informação para o profissional geólogo.

ABSTRACT

Specific computational tools help the geologist to identify lithologies and stratigraphic stacking in well drilling, reducing operational costs and managing the professionals practical work time, directing them to efficient data interpretations or even in the improvement of scientific research in the region. in geologically distinct regions. In this study, the application of machine learning algorithms for the supervised classification of lithologies was evaluated. Data from records of multivariate parameters in offshore wells were used, related to the International Ocean Discovery Program (IODP) with supervised and unsupervised data (images), with the creation of a context of hybrid application of algorithms, divided into two manuscripts. In manuscript I, through the analysis of the lithologies proposed in 7 IODP-Expeditions and the use of the use of the algorithms, it was possible to group and divide the lithological sets into four groups of lithologies and templates. The geophysical properties used in the present study included GRA, PWL, MS, RSC and SRM. The templates were submitted to training and testing by the Multi-Layer Perceptron (MLP), DecisionTree, RandomForest and Support Vector Machine (SVM) methods, using the classification metrics as the result evaluation. As a result, it was observed that Template1 obtained better results in the MLP algorithm, Template2 and Template3 obtained better results in the RandomForest algorithm above 80.00% accuracy. For Cross-validation, the RandomForest algorithm achieved excellent performance in all scenarios. In the Practical Template, the G2 lithology group obtained the best result with the MLP algorithm with an average accuracy greater than 85.00%. For manuscript II, the division of the data included the formatting of three datasets: dataset0, dataset1 and dataset2, specifically with the data obtained during the IODP-Expedition 362. The petrophysical data used included PWL, GRA, RSC, NGR, MAD, MS, RGB and high-definition images. Dataset0 included temporary training data to validate the best interpolator. The dataset1 has the interpolated data of the petrophysical properties, making a total of 295,945 records for U1480 and U1481 with 17 features. The dataset2 covers the texture and color data extracted from the segmentation of the images, making a total of 85,058 records for the U1480 and U1481, with 90 features. Each data set is replicated in two groups of lithology: Group 1 and Group 2. For dataset2, new combinations are added between the features forming 102 practical arrangements with specific results in each combination. The values were interpolated by Linear, Spline, Slinear, Quadratic, Cubic, Akima, Pchip and Piecewise. The machine learning method used for all datasets is RamdomForest. The results show that the best interpolator evaluated in dataset0 is Akima with an accuracy of 98.22%. For dataset1, U1480, the accuracy value is 96.96% in the combination of 70% training and 30% testing in Group 1 and 97.71% in the combination of 70% training and 30% testing in Group 2. For dataset1, U1481, the accuracy value is 99.68% in the combination of 80% training and 20% testing in Group 1 and 99.74% in the combination of 80% training and 20% testing in Group 2. For dataset2, the best evaluated practical arrangements are 51 (Group 1) and 102 (Group 2) for U1480 and practical arrangements 32, 33, 44, 45 and 51 (Group 1) and 102 (Group 2) for U1481. Regarding the new Area Superpixel (Apx) method, the best

results are in Group 1 with the greatest combination of training and the least combination of tests. The evaluated datasets were grouped in an organization between the expedition sites, in which a context of real practical application in the daily activity of the geologist was generated with excellent results of lithological classification. The interpolation of petrophysical data is valid and necessary when there is little data for training respecting the characteristics of each property and interpolator. Properties extracted from the images are relevant and grouped together with the petrophysical properties create a context of extreme importance in the discovery and presentation of information to the geologist.

SUMMARY

INTRODUCTION

The geophysical behavior of sedimentary rocks is predominantly influenced by their mineralogical composition, granulometry, organic matter content, iron content, moisture and the presence of salts. Each type of sedimentary rock is individually characterized by carrying specific features in its structure for its identification and classification.

Numerous methods are capable of measuring the geophysical behavior of these sedimentary rocks creating a group of characteristics (properties or features) (AMINZADEH and DASGUPTA, 2013). Each feature has one or more variables depending on the type of reading value appropriate to the selected method. The main methods are Gamma Ray Attenuation bulk density (GRA), Moisture and Density (MAD), Magnetic Susceptibility (MS), Natural Gamma Radiation (NGR), P-Wave Velocity Logger System (PWL), Red Green Blue channels (RGB) and Reflectance Spectrophotometry and Colorimetry (RSC) adjusting its application to the rock at specific intervals or according to the choice and initial analysis of the use of a given method. These methods are embedded in high-resolution equipment capable of acquiring continuous data in solid bodies appropriate to the weather of the extraction site. It is added to the methods the generation of high-resolution images (high definition) in the rock range analyzed in the color channels R, G and B.

The features, formed through the values extracted from the selected methods, form a dataset suitable for advanced classification/regression processing.

The values of the variables stored in the dataset (s) need to go through adjustments (standardization) checking for the presence of outliers, blank data, noise or values in an inappropriate format.

This standardization of values in the dataset induces the analysis of the combination of the variable by variable and removing, changing or adding values to the dataset, thereby creating new data (called metadata) on the existing dataset. From the analysis and processing of these already standardized datasets, it is possible to generate new knowledge, develop notes on theories of processing and classification, and cover a reliable and portable database for countless computational experiments.

The study site, through the IODP-Expedition 362, covers a huge range of sedimentary rock types where the methods can characterize and cover all drilling contexts with or without physical contact with the rock, extracting data with high precision value and knowledge generation.

OBJECTIVES

The objectives of this study:

- Develop a methodology for classifying lithologies in IODP-Expedition using geophysical data and images;

- Evaluate the data increase (interpolation methods) in dataset of geophysical data of IODP-Expedition;

- Present a new model for dataset formation in IODP-Expedition images;

- Provide datasets on the web for wide access and use by availability on the GitHub platform.

USE OF ARTIFICIAL INTELLIGENCE IN GEOSCIENCES

Artificial Intelligence (AI) emerged as a computing challenge to explain, prototype and transfer human knowledge and learning to the computer. Its application in human activities encompasses several contexts and daily actions of people carrying out information feedback, thereby increasing their knowledge base (DENG et al., 2020).

In the Geosciences field, AI is considered an innovative tool for the scientist (researcher), for the entrepreneur or for the liberal professional. AI does not eliminate the support of the human specialist but helps in the processing, organizing and presentation of results in a quick and precise way over the analyzed data. In geoscience, AI is a new field of study with significant advances in recent years, being a differential due to the amount of data with high referential power (for science) or financial (exploratory). Data diversity in geoscience is the key point inducing AI to seek and relate information in numerous formats, creating its own learning process, adapted to the context of each study site. The study site or material is part of an important context of data selection because AI requires that the reading values are reliable,

organized and consistent with the proposed. In this case, data extracted from IODP-Expedition with a high value of added knowledge stands out, covering an environment of continuous and punctual study in the generation of new knowledge. Still, the IODP numerous multidisciplinary scientists in specific studies for the generation of a huge amount of data aimed at exploration, characterization and appropriate processing. It should be noted that the processing with the IODP-Expedition must be fast and simultaneous with feedback to the dataset.

Artificial Intelligence has a broad concept and covers all techniques and methods capable of transferring to computer learning and performing tasks similar to human beings. KOK et al., 2009 describe that AI is imitating of intelligent human behavior using algorithms or even complex computer programs, which can be divided into two categories: Systems that think and act like humans and systems that think and act rationally. GREWAL, 2014 presents a broader concept of AI. Describes AI as being intelligence a computational mental ability to reason, solve problems, learn and cover the context of smart devices (through the Internet of Things - IoT), therefore defining as a new paradigm capable of increasing and revolutionizing the use of AI in everyday life. CIOFFI et al., 2020 define that AI is used when a machine mimics the human brain's functioning, in learning and solving problems. It highlights that problem solving is the basic principle of AI, solving or looking for alternatives to common activities that require a lot of physical or even repetitive effort.

In improving the AI application, two subfields are used for organization, groupings of functions and data processing logic: Machine learning (ML) and Deep Learning (DL). Machine Learning deals with the extraction of patterns from the analysis of a dataset. It is mentioned for this subfield the statistical methods Support Vector Machine (SVM), Decision Tree, Bayers Learning, Regression, among others. Deep Learning is also called a deep neural network or complex network. DL uses neural networks to analyze several factors of its own containing a structure similar to the human neural system with the complexity of processing and sampling the data instantly and accurately (BRESSAN et al., 2020).

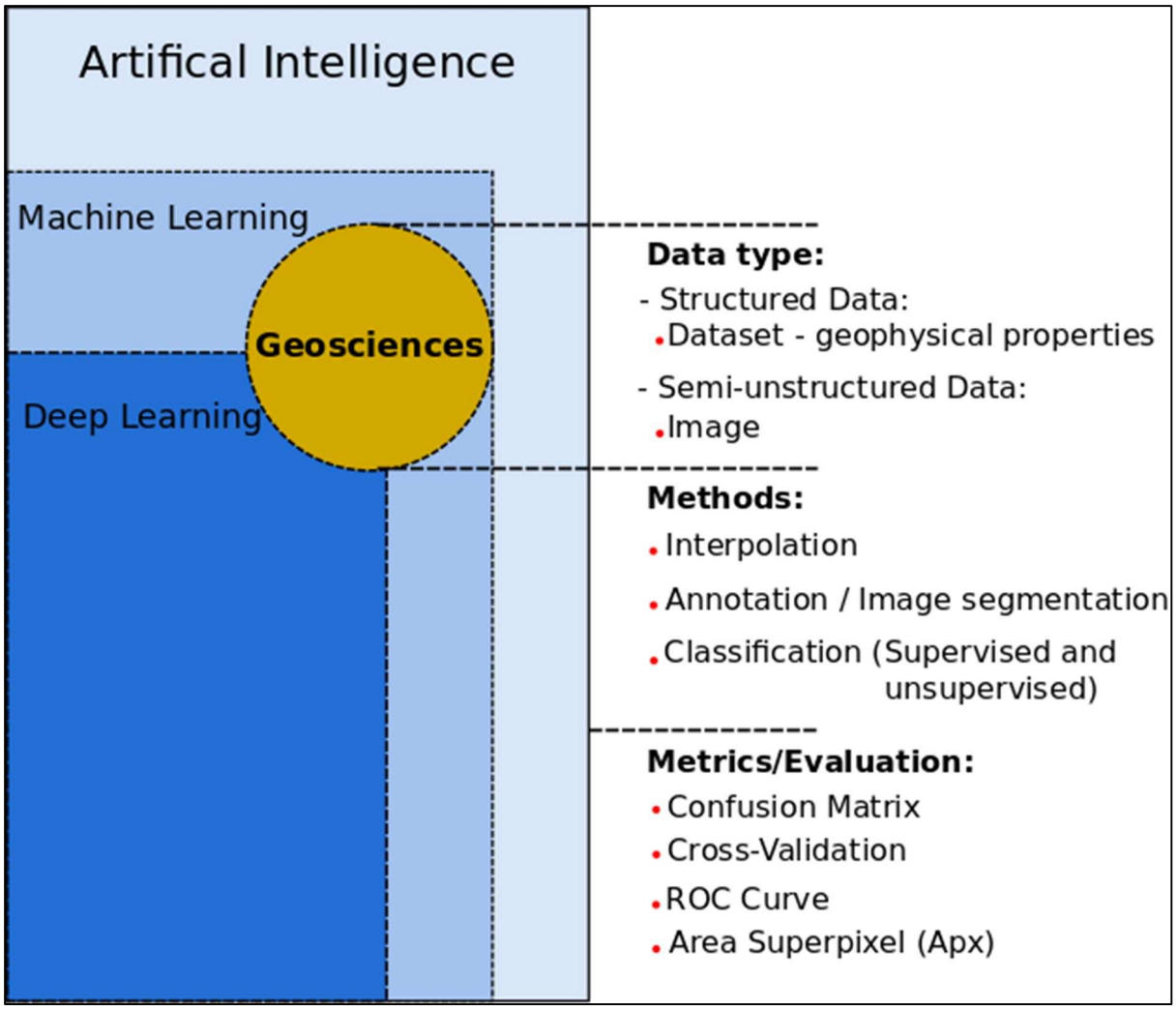


Figure 1 Location of Geosciences in relation to AI and subfields Machine Learning and Deep Learning. The data types are structured and semi-unstructured, the main methods are interpolation, annotation/image segmentation and classification. The evaluation metrics are Confusion Matrix, Cross-Validation, ROC Curve and Area Superpixel (Apx). Source: author.

Geoscience, as shown in Fig. 1, is located between ML and DL. The position between the two subfields is guided by the profile of the data extracted in the Geosciences. It is positioned in ML processing data in structured and unstructured numeric format, from different sources using numerous known statistical methods or even creating new statistical methods customized to the type of data processed. In DL with performance in data processing in structured and semi-unstructured image format, using advanced statistical methods in the context of deep neural networks. For any AI subfield, Geoscience data fit together precisely, creating a vast field of research and application of ML and DL in the discovery of new information. It is added to the ML the formation

and use of large data datasets in line with the use of Data Science and Big Data, strengthening and improving the groups of expected results.

Geosciences and AI form a context of great expansion in the discovery and extraction of information taking into consideration the following factors:

- Large amount of structured and unstructured data to numeric format and structured and semi-unstructured data from images.

- Source/origin of data from different locations.

- Formation of datasets appropriate to processing context.

- Application of various statistical methods with support for programming languages free and easy to use.

- Presentation and visualization of information in a simple and objective way.

- Storage and processing in a secure location (cloud).

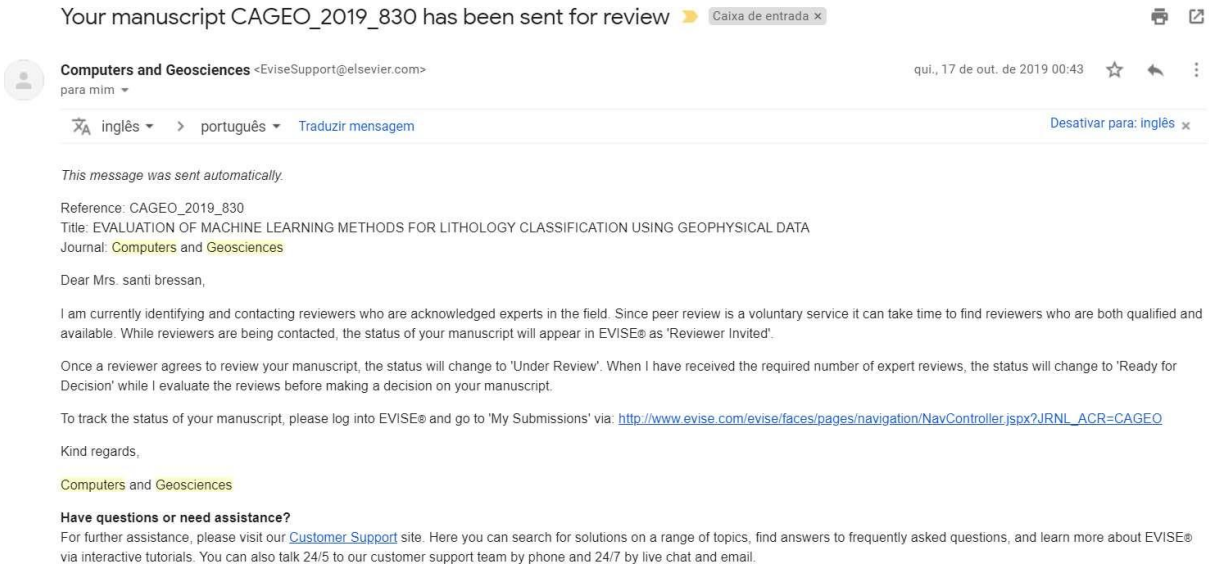
In the context of AI analysis and processing, access to new hardware and cloud computing has improved the expansion of the ML and DP application, from data collection with training and combination of embedded equipment to the large processing of files in the format of the image (video cards with GPU-Graphics Processing Unit support) with total speed and added value in the results. In the scenario of new hardware, IoT covers the formation of equipment aimed at automation and data acquisition in the various application areas, increasing the creation of new reliable data and necessary for the safe expansion of ML and DL.

The evaluation of a model or method in ML and DL requires applying appropriate pre-existing metrics or creating new custom metrics. The results of the metrics need to be in precise, adequate and reliable formation to the decision making by the professional. In this work, as described in Fig.1 we used the types of structured data (geophysical data) and semi-unstructured data (Image) from IODP-Expeditions. The appropriate methods, interpolation (for structured data), Annotation/Image segmentation (for semi-unstructured data) and supervised and unsupervised classification for structured and semi-unstructured data were used. Finally, for metrics, we used the confusion matrix, cross-validation and ROC Curve. A new format called Area Superpixel (Apx) is added

to the metrics specific for the classification of segmented areas (semi-unstructured data).

Manuscript I

Confirming submission to Computers and Geosciences



Your manuscript CAGEO_2019_830 has been sent for review Caixa de entrada ×

Computers and Geosciences <EviseSupport@elsevier.com> qui., 17 de out. de 2019 00:43
para mim

inglês › português Traduzir mensagem Desativar para: inglês ×

This message was sent automatically.

Reference: CAGEO_2019_830
Title: EVALUATION OF MACHINE LEARNING METHODS FOR LITHOLOGY CLASSIFICATION USING GEOPHYSICAL DATA
Journal: Computers and Geosciences

Dear Mrs. santi bressan,

I am currently identifying and contacting reviewers who are acknowledged experts in the field. Since peer review is a voluntary service it can take time to find reviewers who are both qualified and available. While reviewers are being contacted, the status of your manuscript will appear in EVISE® as 'Reviewer Invited'.

Once a reviewer agrees to review your manuscript, the status will change to 'Under Review'. When I have received the required number of expert reviews, the status will change to 'Ready for Decision' while I evaluate the reviews before making a decision on your manuscript.

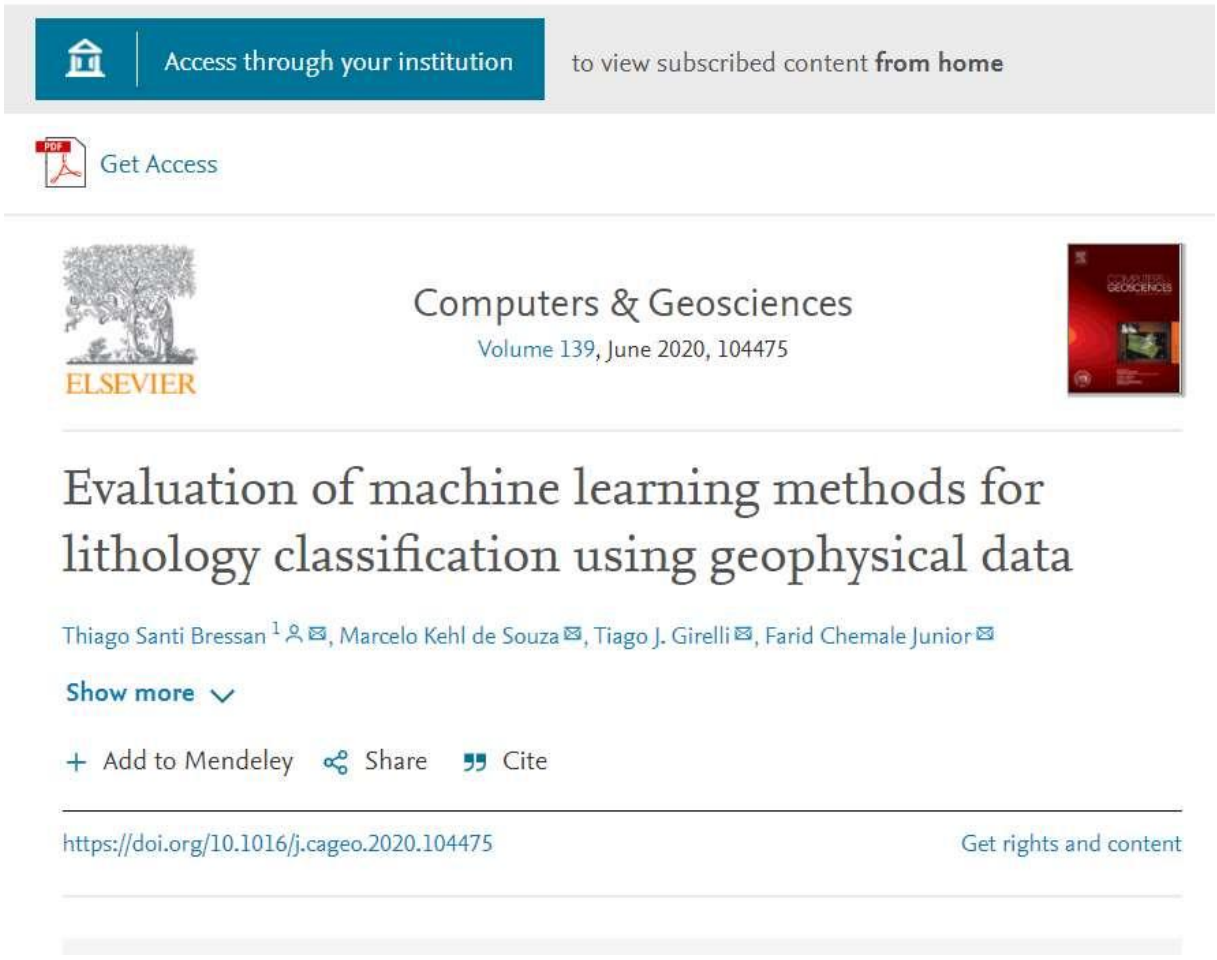
To track the status of your manuscript, please log into EVISE® and go to 'My Submissions' via: http://www.evise.com/evise/faces/pages/navigation/NavController.jspx?JRNL_ACR=CAGEO

Kind regards,

Computers and Geosciences

Have questions or need assistance?
For further assistance, please visit our Customer Support site. Here you can search for solutions on a range of topics, find answers to frequently asked questions, and learn more about EVISE® via interactive tutorials. You can also talk 24/5 to our customer support team by phone and 24/7 by live chat and email.

Confirmation of publication - Computers and Geosciences



EVALUATION OF MACHINE LEARNING METHODS FOR LITHOLOGY CLASSIFICATION USING GEOPHYSICAL DATA

Thiago Santi Bressan[1], Marcelo Kehl de Souza, Tiago J. Girelli, Farid Chemale Junior

Universidade do Vale do Rio do Sinos – Unisinos, São Leopoldo, RS, Brasil

ABSTRACT:

Specific computational tools assist geologists in identifying and sorting lithologies in well surveys, reducing operational costs and practical working time. This allows for the management of professional output, efficient interpretation of data, and completion of scientific research on data collected in geologically distinct regions. Machine learning methods and applications integrate large sets of information with the goal of efficient pattern recognition and the capability of leveraging accurate decision making. The objective of this study is the application of machine learning methods to the supervised classification of lithologies using multivariate log parameter data from offshore wells from the International Ocean Discovery Program (IODP). According to the analysis of the lithologies proposed in the IODP-Expeditions, and for the application of our methods, the lithologies were divided into four groups. IODP-Expeditions were organized into four templates for better results in the analysis of the set of expeditions and practical application of the methods. The templates were submitted to training, validation, and testing by Multi-Layer Perceptron (MLP), DecisionTree, RandomForest, and Support Vector Machine (SVM) methods. Evaluation was randomly divided into training (70%), validation (10%), and testing (20%), using the classification methods as an evaluation of the results. In the results, it was observed that Template1 (IODP-Expedition 362) obtained better results with the MLP method, Template2 (IODP-Expeditions 354, 355, and 359) and Template3 (IODP-Expeditions 354, 355, 359, and 362) obtained better results with the RandomForest method with greater than 80.00% accuracy. For cross-validation, the RandomForest method performed well in all scenarios. In the Practical Template, the G2 group obtained a better result with the MLP method with an average accuracy above 85.00%. It is expected that machine learning methods can help improve the study of geology, with accurate and fast answers related to the interpretation of collected data in different study regions.

KEYWORDS:

lithological group, pattern recognition, multivariate data, sedimentary rocks

[1] Corresponding author
E-mail address: tsbressan@gmail.com

Authorship statement: Thiago Santi Bressan designed and developed the algorithms, worked on the main writing of the manuscript. Marcelo Kehl de Souza contributed to the statistical analysis, algorithm tests, review and writing of the manuscript. Tiago J Girelli contributed to the development of figures and tables and writing of the manuscript. Farid Chemale Junior contributed to the writing and revision of the manuscript.

HIGHLIGHTS:

INSERT HIGHLIGHTS

1. Introduction

The accurate classification of information intensifies work, improves scientific results, and automates the execution of activities with reduced time and work hours. Appropriate methods and computational tools assist geologists with identifying lithologies in onshore or offshore well surveys by allowing for efficient data interpretation. The efficiency of lithological classification can be improved using applications capable of producing more objective decisions, rather than more conventional methods of interpretation (HSIEH, LEWIS and LIN, 2005; JAHDHAMI and ANBOORI, 2017).

Integration between neural networks, machine learning, and geology with numerous scientific and technological studies, search forms, and processes aids in the identification and recognition of patterns in sedimentary rocks. The collection of physical samples of sedimentary rocks is essential, as the professional needs material available for visual interpretations and laboratory analysis in order to generate information with a high degree of validity and accuracy. This process is slow and costly, requiring highly trained and dedicated professionals. In addition, geophysical properties of sedimentary rocks, acquired directly in the field with specific equipment, or in the laboratory, can be utilized, reducing time and costs in the identification of lithologies. These methods include applications in the estimation of permeability and porosity using well profiles (RAFIK and KAMEL, 2017), determination of lithofacies with seismic sections (BRCKOVIC et al., 2017), and geological mapping of surfaces and minerals (LATIFOVIC, POULIOT and CAMPBELL, 2018; BRUNO et al., 2018).

This study differs by its integration of geology and neural networks in intelligent learning in order to assist the professional geologist with practical work in the laboratory or the field. This paper applies the methods of machine learning in the supervised classification of lithologies using multivariate data of log parameters of offshore wells from the International Ocean Discovery Program (IODP).

2. Materials

2.1 Machine Learning

Machine learning can be defined as the most current method of using the computer to discover new knowledge. Since the last decade, the computer was used for simple learning or information storage only. With the advancement of computer science and the creation of new, ultra-fast hardware, one can enhance algorithms and exponentially increase processing by glimpsing computational thinking from the analysis of large amounts of data in a short time. Machine learning is used in areas such as natural language processing, speech recognition, stock prediction, identification of diseases in healthcare, calculation of porosity in sedimentary rocks, identification of specific minerals in sedimentary rocks, and more. Machine learning creates responses from the intensive processing of information, predicting highly reliable outputs for decision making (RASCHKA, 2015). The supervised learning method processes reference data input to create a model for the prediction of new data. For this training, the algorithm requires data in a standard format and type, as well as data with reliable and accurate values, extracted from relevant sources with the ability to improve feedback. In this study, we used a dataset organized from the division of lithologies into GP (Group GP), G1 (Group 1), G2 (Group 2), and G3 (Group 3) according to Table 1. These lithologies are linked to the IODP -Expeditions 349, 354, 355, 356, 359, 361, and 362, and

organized into four templates according to Table 4. They were processed by the application of machine learning methods for supervised classification, including: MLP, DecisionTree, RandomForest, and SVM.

2.1.1 DecisionTree

DecisionTree is a practical, fast, and robust learning method for supervised inductive learning (MAIMON and ROKACH, 2010). It is a useful method in the process of previously unknown information extraction from the analysis of a large volume of data. Examples of applications that use DecisionTree as a learning algorithm include cancer diagnosis (YUE et al., 2018) and image classification (LOUSSAIEF and ABDELKRIM, 2018).

A decision tree is essentially a series of *if-else* statements, aligned through a structure of nodes and sheets. When applied to database records, it results in the classification of records and is a robust method for data with lots of noise, as well as non-standard data (SÁEZ et al., 2013). Configurations such as maximum tree depth, number of features for the best split, maximum number of nodes, maximum number of sheets, and the function for division and choice of nodes can be defined and improved with training.

In relation to the criteria for division and choice of nodes, there are two methods available: Gini (1) and Entropy (2).

$$\text{Gini(E)} = 1 - \sum_{j=1}^{c} p_j^2 \qquad (1)$$

$$\text{H(E)} = - \sum_{j=1}^{c} p_j log\, p_j \qquad (2)$$

The use of each criterion is defined according to the quantity and format of data available for training. For large amounts of data, the Entropy criterion may be altered

using a logarithm, requiring more computational processing and making it relatively slower in relation to the Gini criterion (JING et al., 2017). In relation to data format type, the Gini criterion works well with binary values.

2.1.2 RandomForest

RandomForest is an ensemble method for pattern classification, developed based on decision trees. RandomForest creates a set of multiple decision trees and calculates the average of all processed trees. In this context, it includes groups of calculations, averaging calculations with the aggregation of several estimators and returning the mean of their predictions and reduction of their variances, boosting calculations that integrate several small, specific estimators, seeking an appropriately large return with the sum of the results (JAMES et al., 2013; RASCHKA, 2015).

RandomForest is integral to the configuration criteria for division and choice of nodes (Gini and Entropy), depth of decision trees, and quantity of decision trees.

The averaging method includes a bootstrap aggregation procedure as a statistical approach used to quantify the degree of uncertainty associated with a certain learning estimator.

The RandomForest method is applied in several areas such as meteorology (GRANGE et al., 2018), dengue risk mapping (XU LIU et al., 2018), genetic association studies in Ecology (BRIEUC et al., 2018), and specifically in Geology and Geoscience in group mapping of vegetation (DUBEAU et al., 2017), predicting the spatial distribution of seabed hardness (LI, TRAN and SIWABESSY, 2016), and fault detection (PUGGINI, DOYLE and MCLOONE, 2015).

2.1.3. Support Vector Machine (SVM)

SVM is a supervised learning method that constructs a border (hyperplane) on a representation of the data, thus improving its presentation, grouping, and separation between different instances of data classes (VAPNIK, 1998).

Its application depends on the format and the relationship between the data, with possible application on linearly separable and linearly non-separable data. The method calculates the vector that best represents the hyperplane. Support Vectors (Fig. 2) are points closest to the hyperplane line and the distance between these vectors are the margins. The method chooses the hyperplane that has the greatest margin, being considered the more robust model, less tolerant to errors.

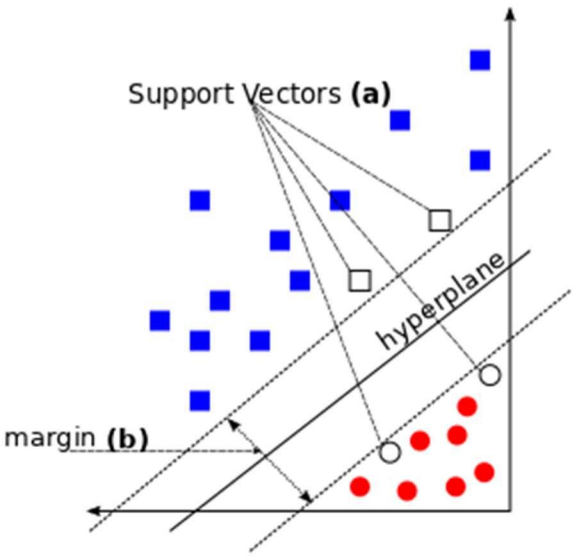


Figure 2 Representation of SVM in a linearly separable model. The groups are converted into data vectors and separated by a hyperplane calculated by the method that best represents the separation of the data. Margins represent the distance between the data vectors. Figure modified from VAPNIK, 1998.

For a model with linearly separable data (as seen in Fig. 2) the representation performed by SVM is fast and accurate, resulting in an excellent practical return for classification output. A model with linearly non-separable data requires a method to

interpret the data group, which applies a kernel method to find patterns and relationships after applying the representation of the hyperplane to the group data.

SVM is applied in several areas such as cancer genomics (HUANG et al., 2018), and in Geology and Geoscience for 3D geological modeling (DE BOISSIEU et al., 2017), regolith-geology (DE BOISSIEU et al., 2017), hyperspectral classification of images (PENG, ZHOU and CHEN, 2015), and automated lithological classification using ASTER remote sensing data (YU et al., 2012).

2.1.4 Multi-Layer Perceptron (MLP)

An Artificial Neural Network (ANN) can be defined as a linear model based on brain architecture, developed in an attempt to transfer learning capacity to a computerized system (CASTRO, 2017; SOUZA et al., 2019). The best-known architecture of an ANN is the MLP, which is made up of a series of layers, neurons, and connections.

It is important to note that due to its definition, the MLP architecture depends on the adjustment of the main points: architecture configuration and training. The configuration of the architecture includes the definition of the connections and quantities of neurons necessary for good learning and adequate resolution of the proposed problem. Consequently, training of MLP networks can be very time-consuming for large datasets. The determination of the ideal architecture configuration is a constant research objective, in which the quadratic error, prediction, precision and its metrics are adjusted for in the expected results. In these points and using these metrics, results are related to the type and quality of the data used in the process of forming the network.

Because it is a traditional network in the context of the generation and application of knowledge, areas such as agriculture (HALAGUNDEGOWDA and ABHISHEK, 2018), biomedicine (CAO et al., 2018; SAVALIA and EMAMIAN, 2018), and road infrastructure (MANSOURKHAKI, BERANGI and HAGHIRI, 2017) utilize the practical organization of this important method.

2.1.5 Validation and Evaluation

A confusion matrix is the representation of real values and predicted values that allows for the visualization of the performance of a machine learning classifier method on the proposed templates (NAVIN and PANKAJA, 2016). The confusion matrix configures itself as a table generated for classification of a binary dataset and is used to describe the performance of a classification or method. The main diagonal indicates the accuracy of the evaluated records, combining the true results in an organized structure in the matrix. Its organization is presented according to Fig. 3.

Accuracy (Eq. 3) is formed by the division between true positive and true negative values and total positive and negative values, according to Equation 3:

$$\text{Accuracy} = \frac{TP+TN}{P+N} \quad (3)$$

It is important to highlight that this classification metric can result in false method performance because this metric calculates the average between the return of the classes of a dataset (HOSSIN and SULAIMAN, 2015).

Precision (Eq. 4) and recall (Eq. 5) belong to the F_1-score metric. Precision is calculated by the division of true positive values by the sum between true positive values and false positive values. Recall is calculated by the division of the true positive

values by the sum between true positive and false negative values. Their equations are presented below:

$$\text{Precision} = \frac{TP}{TP+FP} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (5)$$

F_1-score or F-measure (Eq. 6) is the harmonic mean between precision and recall. Its use stands out in dataset processing with diversified classes that are highly disproportionate. This equation is highlighted below:

$$F_1 = \frac{2*Precision*recall}{Precision+recall} \quad (6)$$

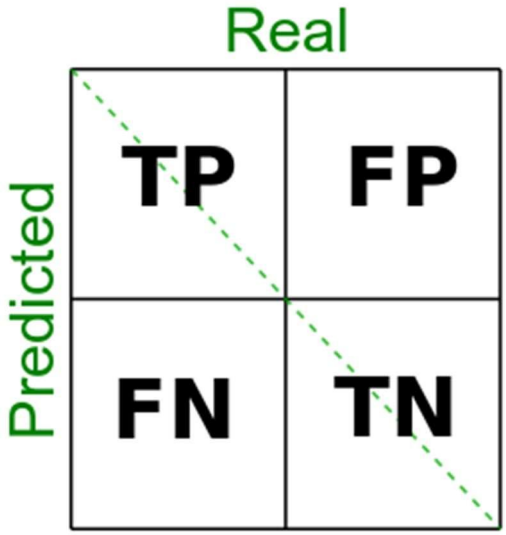


Figure 3 Definition of the confusion matrix. The matrix is divided into real data and predicted data (rows and columns), combining TP (True Positive) data, FP (False Positive) data, FN (False Negative) data and TN (True Negative) data. Figure modified from NAVIN and PANKAJA, 2016.

The cross-validation method is used to evaluate the performance of the data. This method randomly partitions the total untrained dataset into *k* smaller groups of equal size (HAYKIN, 2009). Processing of the data is repeated *k* times until all groups are trained and tested. Processing returns are described through rating metrics such as accuracy, precision, recall, and F_1-score according to Fig. 4. In this way, the entire set of available data is evaluated, returning precise classification of the data and integrating the various characteristics of data formation and grouping.

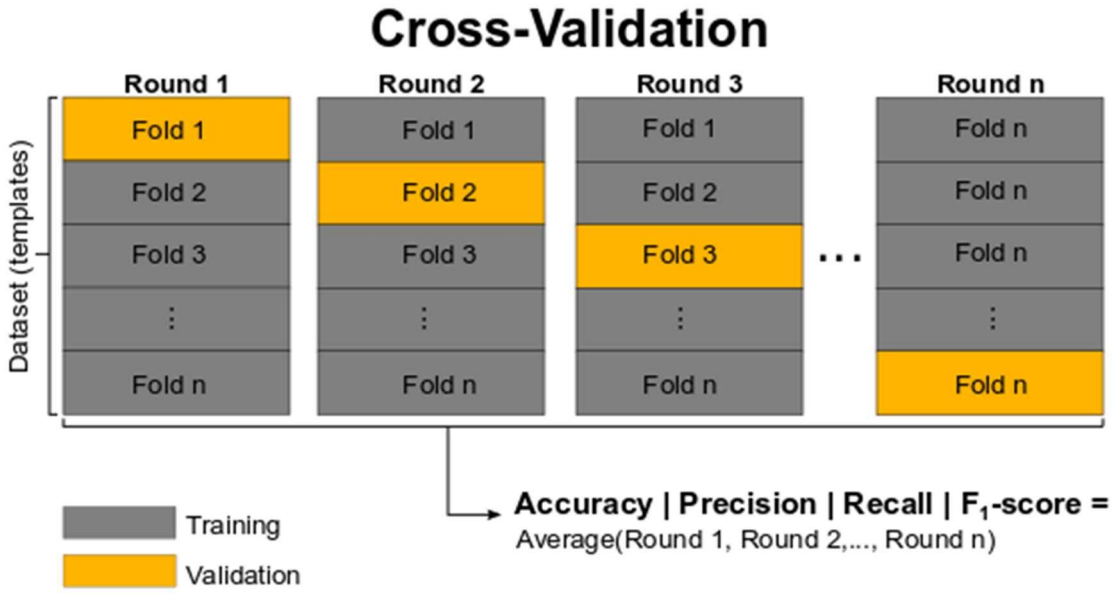


Figure 4 Graphical representation of cross-validation with the selected data in the datasets of the templates. In each round the dataset is divided into n groups of equal size, with n groups for training and n groups for validations. The cross-validation result will produce the measurements of accuracy, precision, recall, and F1-score. Figure modified from HAYKIN, 2009.

Receiver Operating Characteristic (ROC) is a graphical method for evaluation, presentation, and selection of prediction systems (THARWAT, 2018). It uses the confusion matrix to construct the results and represents two parameters of the

223 probability of accuracy (True Positive Rate (TPR) and False Positive Rate (FPR)). TPR
224 (Eq. 7) and FPR (Eq. 8) are defined as follows:

$$TPR = \frac{TP}{TP+FN} \quad (7)$$

$$FPR = \frac{FP}{FP+T} \quad (8)$$

225

226 The ROC curve plots TPR vs. FPR at different classification thresholds. Its
227 multidimensional capability allows for better visualization of the result variables
228 throughout the spectrum of the graph. The descending diagonal (0,1) represents the
229 classification model that plays equally in both classes. Points belonging to the upper
230 left triangle of this diagonal represent the best results and points belonging to the lower
231 right triangle represent the worst results. Its origin is related to the detection of signals
232 and evaluation of the transmission quality of a noise signal (EGAN, 1975). ROC
233 graphics are used in medicine (TILAKI-HAJIAN, 2013), economics (GAJOWNICZEK,
234 ZABKOWSKI and SZUPILUK, 2014), weather forecast (ZHAO, 2011), and geology
235 (VAKHSHOORI and ZARE, 2018).

236

237 2.2 Geological Setting

238 This study is based on the IODP-Expeditions 349, 354, 355, 356, 359, 361, and 362.
239 The holes were drilled in different regions of the Indian Ocean. As a result, there is a
240 large amount of information resulting in a good data group for this study in machine
241 learning. Further description of the study areas can be found in the supplementary
242 material.

All of the expeditions described in the present study contain similar successions of sediments/sedimentary rocks from the Bengal-Nicobar Fan. Therefore, it is possible to perform a grouping of lithologies and to map a pattern between depth and lithology, as well as the distribution of sedimentary rocks in all of the sites and holes surveyed for the seven expeditions.

The grouping of lithologies seeks to organize the sets of lithologies present in the IODP-expeditions by separating the records, creating a wide combination of data and direction for the heterogeneous sedimentary rocks identified and described in the Visual Core Description (VCD). FAN et al. (2019), KOROLEV et al. (2018), and RAHIM et al. (2009) describe the division into groups as integral to the heterogeneous reality of lithological analysis in the field as it is related to the multivariate physical characteristics of the site or core sampled.

For each lithological group, models and lithological code were assigned for the classification of the lithology by machine learning methods. Table 1 presents the model of the division of lithologies into four groups, denominated GP (Group GP), G1 (Group 1), G2 (Group 2), and G3 (Group 3).

Table 1 Division of lithology into groups. Each group contains its composition models, lithological composition, and lithological code.

Groups	Models	Lithological Composition	Lithology Code
GP	**Litho1**	Very fine sand/sandstone, Fine sand/sandstone, Medium sand/sandstone, Coarse sand/sandstone, Sand, Sand/Sandstone	10

	Litho2	Alternating sand/sandstone and mud/mudstone layers, Sandy clay/claystone, Sandy silt/siltstone	11
	Litho3	Alternating silt/siltstone and clay/claystone layers, Clayey silt/siltstone	12
	Litho4	Sand/sandstone-silt/siltstone-clay/claystone, Clayey sand/sandstone, Silty sand/sandstone	13
	Litho5	Silt/Siltstone, Silty clay/claystone, Clay/claystone, Clay, Silt, Mudstone	14
	Litho6	Calcareous silty clay/claystone, Calcareous silt/siltstone, Calcareous ooze, Chalk, Marlstone, Rudstone, Floatstone, Grainstone, Packstone, Wackestone, Boundstone, Limestone, Calcareous claystone	15
G1	**Litho1**	Very fine sand/sandstone, Fine sand/sandstone, Medium sand/sandstone, Coarse sand/sandstone, Sand, Sand/Sandstone	10
	Litho2	Silt/siltstone, Silty clay/claystone, Clay/claystone, Clay, Silt, Alternating silt/siltstone and clay/claystone layers, Clayey silt/siltstone	14
	Litho3	Calcareous silty clay/claystone, Calcareous silt/siltstone, Calcareous ooze, Chalk, Marlstone, Rudstone, Floatstone, Grainstone, Packstone, Wackestone, Boundstone, Limestone, Calcareous claystone	15
G2	**Litho1**	Very fine sand/sandstone, Fine sand/sandstone, Medium sand/sandstone, Coarse sand/sandstone, Sand, Sand/sandstone	10
	Litho2	Sand/sandstone-silt/siltstone-clay/Claystone, Clayey sand/sandstone, Silty sand/sandstone, Alternating sand/sandstone and mud/mudstone layers, Sandy clay/claystone, Sandy silt/siltstone	13
	Litho3	Silt/siltstone, Silty clay/claystone, Clay/claystone, Clay, Silt, Alternating	14

		silt/siltstone and clay/claystone layers, Clayey silt/siltstone	
	Litho4	Calcareous silty clay/claystone, Calcareous silt/siltstone, Calcareous ooze, Chalk, Marlstone, Rudstone, Floatstone, Grainstone, Packstone, Wackstone, Boundstone, Limestone, Calcareous claystone	15
G3	**Litho1**	Very fine sand/sandstone, Fine sand/sandstone, Medium sand/sandstone, Coarse sand/sandstone, Sand, Sand/sandstone	10
	Litho2	Alternating sand/sandstone and mud/mudstone layers, Sandy clay/claystone, Sandy silt/siltstone, Sand/sandstone-silt/siltstone-clay/ claystone, Clayey sand/sandstone, Silty sand/sandstone	11
	Litho3	Alternating silt/siltstone and clay/claystone layers, Clayey silt/siltstone, Silty clay/claystone	12
	Litho4	Silt/Siltstone, Silt	13
	Litho5	Clay/claystone, Clay	14
	Litho6	Calcareous silty clay/claystone, Calcareous silt/siltstone, Calcareous ooze, Chalk, Marlstone, Rudstone, Floatstone, Grainstone, Packstone, Wackestone, Boundstone, Limestone, Calcareous claystone	15

262

263

264 During IODP-Expeditions, the data collected on-board provided the first steps for
265 working with the sampled rocks. The collection of these data occurred in core or
266 collected samples. Among the measurements of the log parameters of rocks, we can
267 highlight Gamma Ray Attenuation bulk density (GRA), P-Wave Velocity Logger

System (PWL or P-wave), Magnetic Susceptibility (MS), Reflectance Spectrophotometry and Colorimetry (RSC), and Magnetic Remanence (SRM).

GRA is a measure of the density expressed in g/cm³, with a high degree of penetration, emitted spontaneously from an atomic nucleus (137 Ce) during radioactive decay. It has a gamma ray peak of 662 KeV and is attenuated as it passes through the core. This attenuation is related to Compton spreading, where a known sample thickness is proportional to the bulk density. Bulk density can also be affected by vertical compaction during the collect of cores. This measure is used in the identification or classification of rocks, mineral composition, grain size, and porosity calculation (ODP, 2007).

PWL values are measurements of the sound wave velocity through a sample. The PWL velocity varies according to the physical composition, porosity, density, and degree of fracture. In marine environments, PWL values are influenced by the degree of consolidation and lithification, fractures, and hydrocarbon occurrence (BRCKOVIC et al., 2017; DOVETON, 1994). Together with the GRA, the PWL measurements are used to calculate the acoustic impedance and reflection coefficients to construct synthetic seismic profiles and to estimate the depths of seismic horizons.

MS is the intensity with which the material can be magnetized in an external magnetic field (BLUM, 1997; BRCKOVIC et al., 2017; MCNEILL et al., 2017). The ratio of magnetization is expressed in units of volume, defined as:

$$k = M/H$$

where *M* is the volume of magnetization applied to a magnetic susceptibility *k* by an applied external field (*H*). Susceptibility is measured with the main recording devices, for which calibration factors must be met for geometry and effects of transport and core

292 coatings. They can be classified by the magnetization volume value (*M*) in three
293 groups: diamagnetic materials (-1 < *M* <0), paramagnetic materials (0 < *M* <= 1), and
294 ferromagnetic materials (*M* >= 1). MS varies according to the type and concentration
295 of magnetic grains and corresponds to variation in sediment composition, mainly the
296 granulometry and mineralogical composition. Sediments with the presence of clay
297 have relatively lower magnetic susceptibility and materials with the presence of water
298 tend to have slightly negative values.

299 RSC is a unit of measurement related to two widely used techniques in the visual
300 identification of rock characteristics: Colorimetry and Reflectance Spectrophotometry.
301 Colorimetry is used to measure the color value of a surface. Many numerical systems
302 have been developed to express the visual values of colors. The International
303 Commission on Illumination (CIE) proposed a standard method for the numerical
304 measurement of colors, the L*, a*, b* system, considering the non-linear perception of
305 the human eye and the combination of illumination and basic colors (HUGHES and
306 LANGLOIS, 2010; CIE, 2004). L* defines the brightness value of the respective
307 sample, with values between L* = 0 (totally dark) and L* = 100 (totally bright). The a*
308 defines the value of the red-green coordinate (a* positive defines values with more red
309 and a* negative defines values with more green) and b* defines the value of blue-
310 yellow coordinate (b* positive defines values with more yellow and b* negative defines
311 values with more blue) (BLUM, 1997; CIE, 2004). Colorimetry is widely used to
312 calculate relative brightness and colors on a surface or sample, combining with other
313 properties in order to create a more accurate evaluation of the analyzed object
314 (HUGHES and LANGLOIS, 2010).

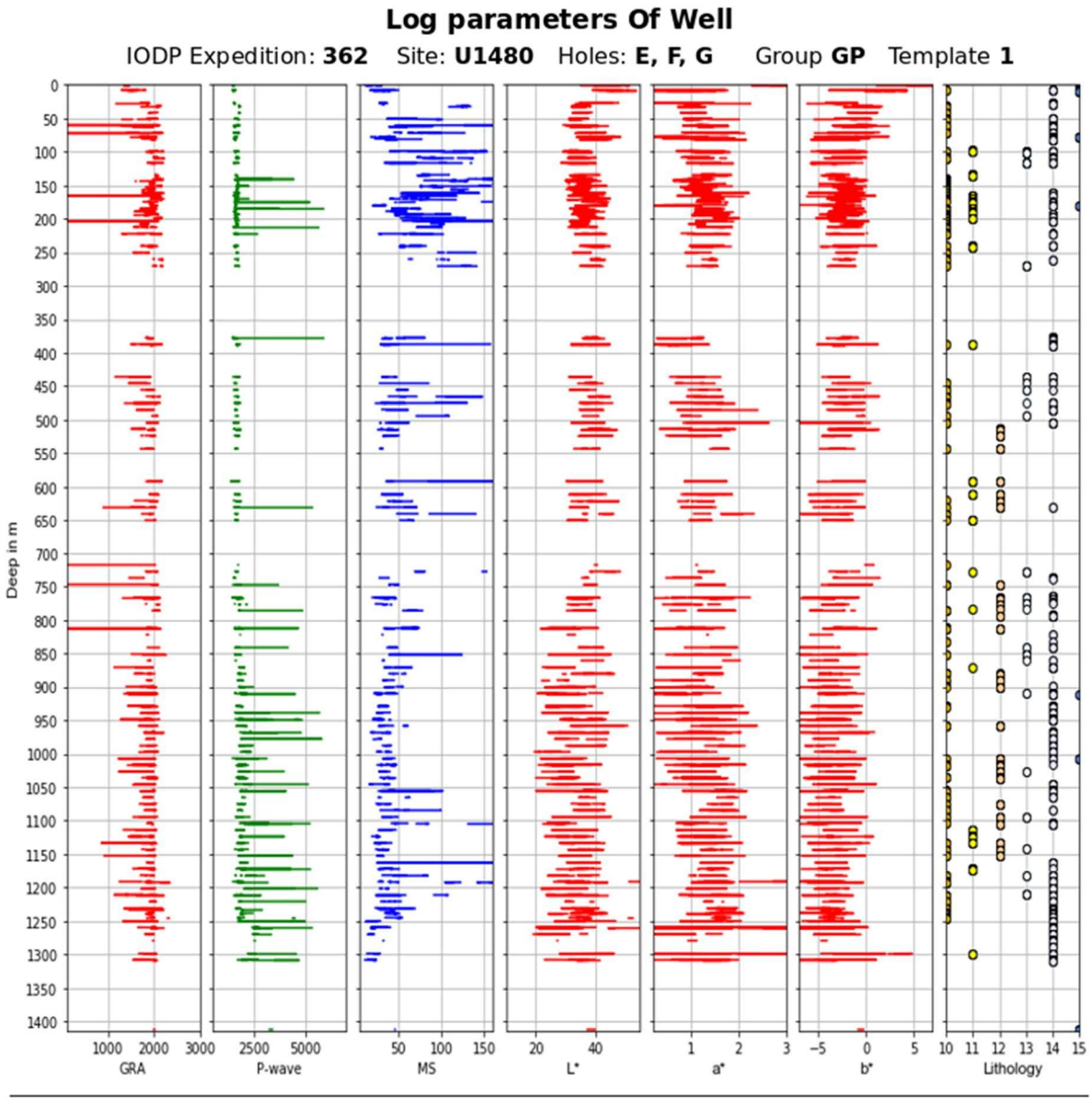
315 Magnetization is correlated with the proportional response of magnetic susceptibility to
316 a magnetic field. In some cases, for pure magnetization, this relationship may undergo

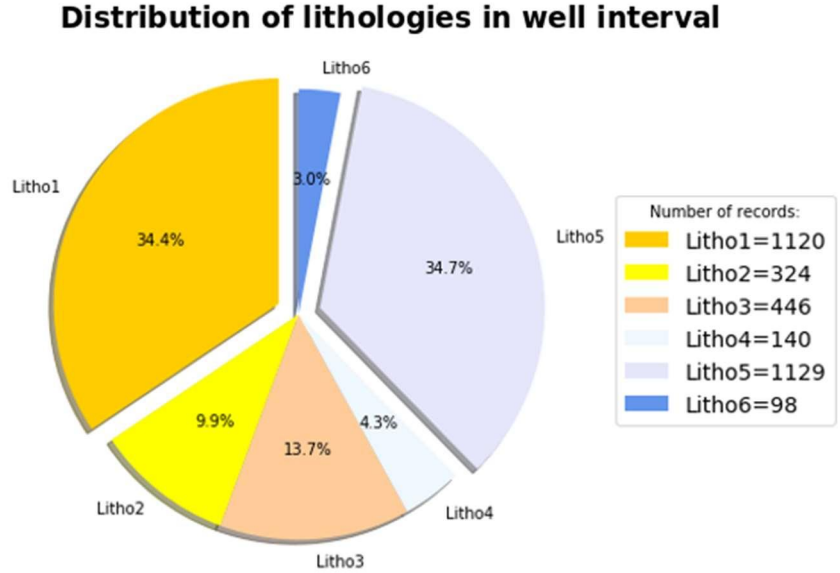
317 changes, where a medium exhibits a magnetic field, even with the absence of a
318 magnetic field applied to it. This process is called Magnetic Remanence (SRM)
319 (BUSCHOW and BOER, 2004; GOODRICH, 2007; JOVANE et al., 2013). The vector
320 of magnetization of an object is the sum of the values of the induced magnetic field
321 and the magnetic field remanence. Magnetic Remanence can be classified into five
322 types, of which the main type, Natural Remanence Magnetization (NRM), was used in
323 the IODP. NRM is a more reliable method, since remanence is transmitted from an
324 object naturally, by its favorable chemical compositions and without the interference of
325 equipment or sensors.

326 The definition of use for these seven log parameters is according to the fact that data
327 acquisition resolution is closest to the depth, and the joint capability of these logs
328 facilitates in the recognition and classification of lithologies. Through analysis of the
329 GRA, PWL, MS, L*, a*, b*, and SRM in a joint processing using machine learning, it is
330 possible to identify which member of the lithological group belongs to the value sought.

331 Fig. 5 shows a model of the log parameters for IODP-Expedition 362, Sites U1480,
332 Holes E, F, and G, which form Template1, Lithology Group GP, with a plot of the
333 classification of the lithologies in relation to the depth of the wells. The other
334 expeditions and sites follow the model of Fig. 5, adjusting the depth with the lithological
335 classification.

336


Log parameters Of Well
IODP Expedition: 362 Site: U1480 Holes: E, F, G Group GP Template 1
Deep in m
GRA
P-wave
MS
L*
a*
b*
Lithology


Distribution of lithologies in well interval
Litho6
3.0%
Litho1
34.4%
Litho5
34.7%
9.9%
13.7%
4.3%
Litho2
Litho4
Litho3
Number of records:
Litho1=1120
Litho2=324
Litho3=446
Litho4=140
Litho5=1129
Litho6=98

337

Figure 5 Log parameters for the IODP-Expedition 362, Site U1480, Holes E, F, and G which form Template1, Lithology Group GP. Lithology Group GP is divided into Litho1, Litho2, Litho3, Litho4, Litho5, and Litho6. Litho1 consists of sand with variable granulation; Litho2 is composed of interlayered sand and mud with layers of clay and silt; Litho3 is composed of layers of clay and silt; Litho4 consists of mixed layers of sand, silt, and clay; Litho5 is composed of lithologies divided between silt, clay, and mudstone; and finally, Litho6 is composed of lithologies based on carbonate composition and its derivatives. The color legend in the figure refers to the lithologies present in the GP group. The unit of measurement for GRA is g/cm³, P-wave is m/s, MS is m³/kg, L* is %, a* is %, and b* is %. The SRM log parameter is not included.

2.3 Methods and data configuration

2.3.1 Data preparation

The total number of records for the IODP-Expeditions, divided in groups, is shown in Table 2. Each record includes values for the seven log parameters GRA, PWL, MS, RSC (L*, a*, b*), and SRM. The data set formed by the log parameters is transformed into the matrix. A table of the complete data set with IODP-Expeditions, groups, and lithologies is found in Appendix 1E of the supplementary material.

Table 2 Total number of records by lithological groups

Number of Records – IODP-Expeditions	**Group**
24504	GP
23967	G1
24504	G2
24504	G3
Total number: 97479	GP, G1, G2 and G3

2.3.2 Programming language and library for machine learning

Python is a high-level programming language that is interpretable, easy to learn and use, and supports numerous add-ons, which makes it a powerful language for

362 calculating and analyzing large amounts of information (SPRONCK, 2017). In this
363 study of machine learning, the Scikit-learn Library was used, which integrates all
364 methods for learning processing, with the support of supervised and unsupervised
365 training on the data, creating extremely elaborate and understandable outputs. The
366 modules used in Scikit-learn are the MLPClassifier module for the MLP algorithm,
367 DecisionTreeClassifier module for the DecisionTree algorithm,
368 RandomForestClassifier module for the RandomForest algorithm, and SVC module for
369 the SVM algorithm.

370

371 2.3.3 Training

372 To carry out the trainings, the data were divided according to the Table 4. Three
373 templates were created with data combinations between IODP-Expeditions 354, 355,
374 359, and 362, using 70.00% for training data, 10.00% for validation data and 20.00%
375 for test data, randomly separated. A Practical Template was created, simulating a real
376 exercise of application of the methods, predicting the lithologies of a hole through the
377 training of the entire dataset. The composition of the expeditions in the templates was
378 given by the proximity of the expeditions in the same geological region, directing the
379 context as a practical analysis in the field or directly on the ship.

380

381 2.3.4 Prediction

382 For each of the templates, the training described in the Train column of Table 4 was
383 performed using the MLP, DecisionTree, RandomForest, and SVM methods. In order
384 to carry out the prediction, data was used according to the Test column of Table 4.

385 The configurations of the MLP, DecisionTree, RandomForest, and SVM methods are
386 shown in Table 3. The configurations use features of the methods to create algorithms

in the Python language, and the main points were adjusted for use with possible characteristics such as number of nodes, trees, depth, and activation function. The SVM method was configured with the default configuration provided by the programming language itself.

Table 3 Configuration of the methods for training, validation and testing

Methods	**Configuration**
MLP	solver = lbfgs activation = relu random_state = 8
DecisionTree	random_state = 8 max_depth = 20 criterion=entropy
RandomForest	max_depth=20 n_estimators = 1000 random_state = 8 n_jobs = -1
SVM	Default setting

The distribution between training, validation, and testing in Templates 1, 2, and 3 was related to the performance of practical tests with the methods and in accordance with STORKEY (2013) and KORJUS, HEBART and VICENTE (2016) with a percentage of 10% for validation and 10-20% for testing. For the Practical Template, training, validation, and testing were divided according to the segmentation of the expeditions.

Table 4 Division for training according to IODP-Expeditions

Groups	**Template**	**Composition IODP**	**Train**	**Validation**	**Test**	**Methods**

		Expeditions – Sites				
GP, **G1**, **G2** and **G3**	Template1	**362 (**U1480)	70.00%	10.00%	20.00%	MLP DecisionTree RandomForest SVM
	Template2	**354 (**U1449, U1450, U1451) **355 (**U1456) **359 (**U1465, U1466, U1467, U1468, U1470, U1471, U1472)	70.00%	10.00%	20.00%	
	Template3	**354 (**U1449, U1450, U1451) **355 (**U1456) **359 (**U1465, U1466, U1467, U1468, U1470, U1471, U1472) **362 (**U1480)	70.00%	10.00%	20.00%	
	Practical Template	**349 (**U1431, U1432, U1433) **354 (**U1449, U1450, U1451) **355 (**U1456) **356 (**U1462) **359 (**U1465, U1466, U1467, U1468, U1470, U1471, U1472) **361 (**U1477, U1478) **362 (**U1480, U1481)	Expeditions **354**,**355**,**359** and **362** (U1480)	Expedition **349** (U1433)	Expeditions **349, 356**, **361** and **362** (U1481)	

400

401

After trainings and predictions, a table was compiled comparing the output result for each template and each method of the analyzed dataset. For each processing, classification metrics and confusion matrix values were tabulated for better organization and presentation. A detailed table with the results is found in Appendix 1E of the supplementary material.

3 Results and Discussions

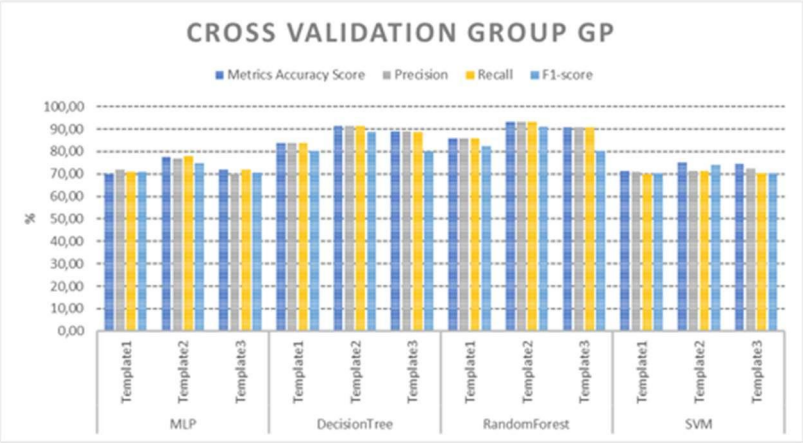
In this section, the results of the application of the four methods are discussed, according to the templates and data set presented.

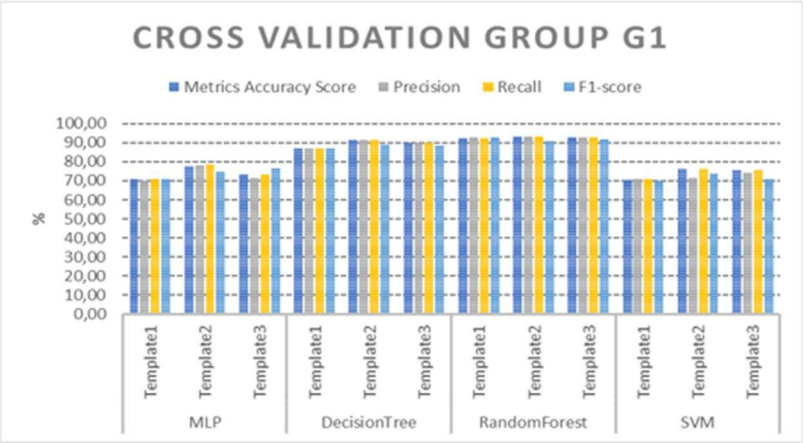
The sum of the records of the data set used by the templates in this study was 97474 records. According to STRAUß (2018) and VENKATARAMAN (2017), in order to obtain results above 70.00% accuracy using machine learning, it is necessary to have thousands or millions of records for training. Such results depend on the organization of a precise and reliable historical database, that adjusts itself with trainings and testing. In this case, it was determined that for the GP group, Template1 obtained an accuracy of 76.00% in the SVM method, Template 2 obtained an accuracy of 82.61% in the RandomForest method, and Template3 obtained an accuracy greater than 82.74% in the RandomForest method. For the G1 group, Template1 obtained an accuracy of 89.51% in the RandomForest method, Template2 obtained an accuracy of over 75.00% in the four methods analyzed, and Template3 obtained an accuracy of over 73.00% in the four methods analyzed. For the G2 group, Template1 obtained an 82.84% accuracy in the RandomForest method, Template2 obtained an 82.61% accuracy in the RandomForest method, and Template3 obtained an 84.20% accuracy in the RandomForest method. Finally, for the G3 group, Template1 obtained a 79.00%

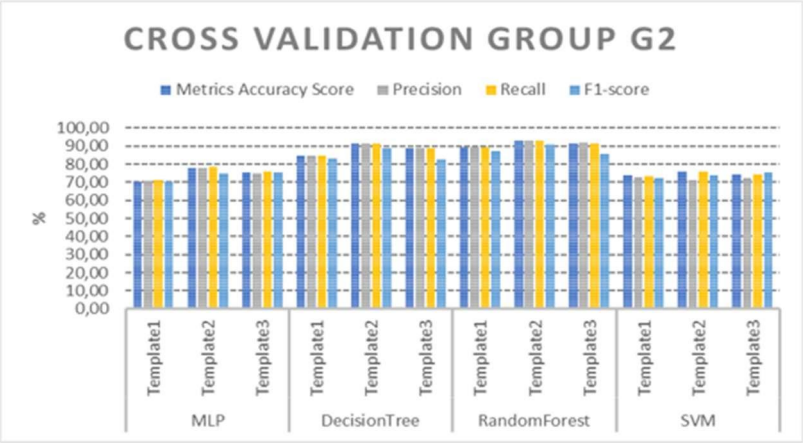
426 accuracy in the SVM method, Template2 obtained an accuracy of 77.89% in the
427 RandomForest method, and Template3 obtained a 79.42% accuracy in the
428 RandomForest method. A detailed table with these results is found in Appendix 1E of
429 the supplementary material.

430 The results of the application of cross-validation is shown in Fig. 6. They were divided
431 into five folds, performing training and testing for all templates and datasets. Accuracy,
432 precision, recall, and F_1-score were used to measure the performance of templates for
433 five-fold cross-validation. The DecisionTree method in the entropy criterion obtained a
434 better result in relation to the Gini criterion, with an increase in accuracy of 5.00%, as
435 well as a proportional increase in cross-validation with 5 *k* folds.

436


CROSS VALIDATION GROUP GP
Metrics Accuracy Score
Precision
Recall
F1-score
%
100,00
90,00
80,00
70,00
60,00
50,00
40,00
30,00
20,00
10,00
0,00
Template1
Template2
Template3
MLP
DecisionTree
RandomForest
SVM


CROSS VALIDATION GROUP G1
Metrics Accuracy Score
Precision
Recall
F1-score
%
100,00
90,00
80,00
70,00
60,00
50,00
40,00
30,00
20,00
10,00
0,00
Template1
Template2
Template3
MLP
DecisionTree
RandomForest
SVM


CROSS VALIDATION GROUP G2
Metrics Accuracy Score
Precision
Recall
F1-score
%
100,00
90,00
80,00
70,00
60,00
50,00
40,00
30,00
20,00
10,00
0,00
Template1
Template2
Template3
MLP
DecisionTree
RandomForest
SVM

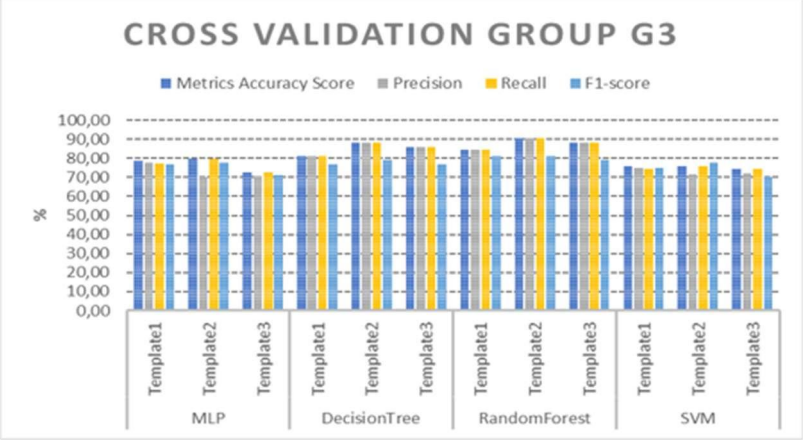

CROSS VALIDATION GROUP G3
Metrics Accuracy Score
Precision
Recall
F1-score
%
100,00
90,00
80,00
70,00
60,00
50,00
40,00
30,00
20,00
10,00
0,00
Template1
Template2
Template3
MLP
DecisionTree
RandomForest
SVM

Figure 6 Results of the application of cross-validation in the four defined groups: GP, G1, G2, and G3, divided into Template1, Template2, and Template3. Template1 covers IODP-Expedition 362, Site U1480; Template2 covers IODP-Expeditions 354, 355, and 359; and Template3 covers IODP-Expeditions 354, 355, 359, and 362. The methods applied to the data were MLP, DecisionTree, RandomForest, and SVM.

In Fig. 7, log parameters are presented for the Practical Template in the defined groups and methods (data of G3 group, IODP-Expedition 362, Site U1481). For the Practical Template, the execution of the methods obtained excellent results, which can be observed in Fig. 7 and Appendix 1E of the supplementary material.

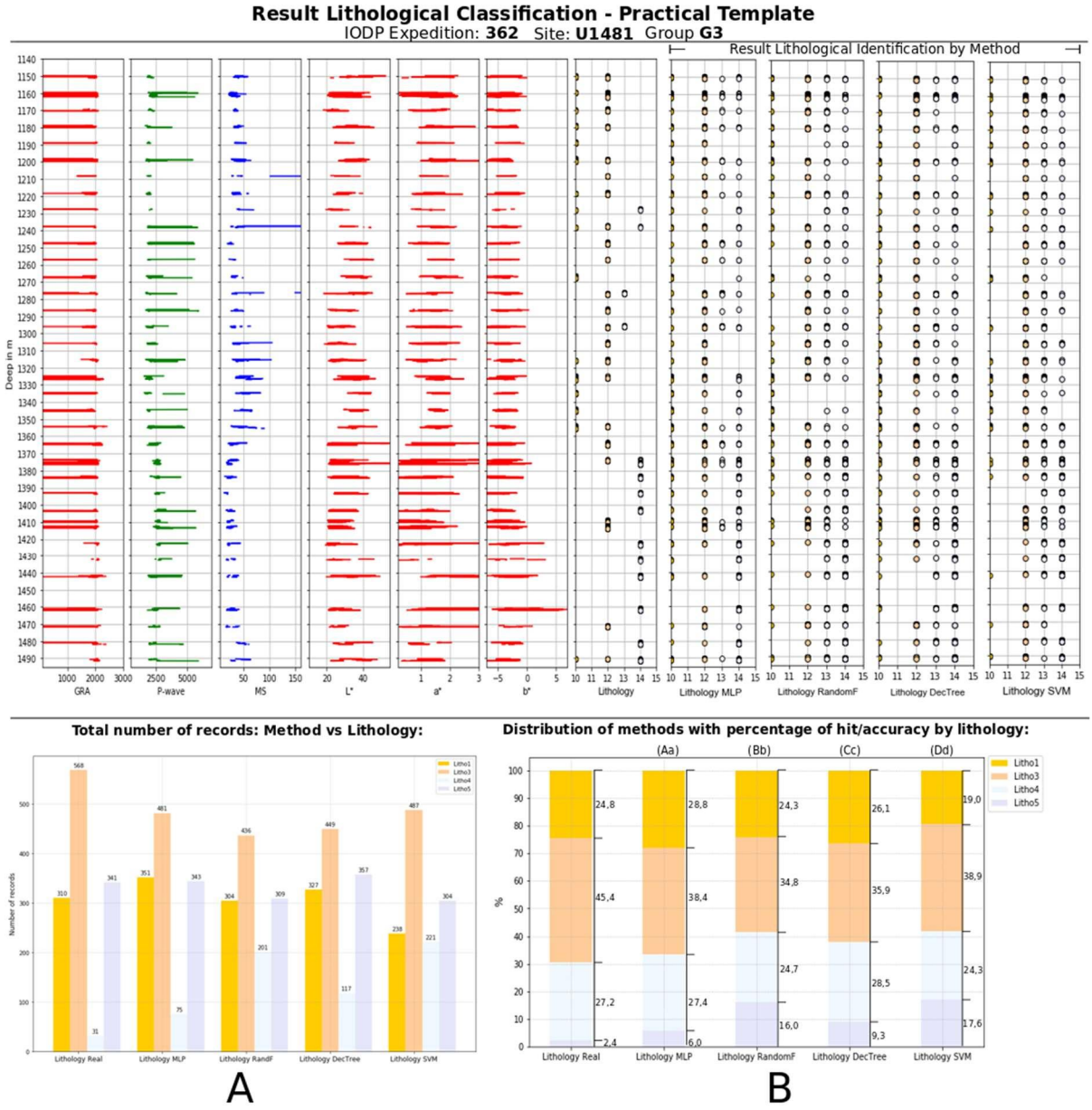


Figure 7 Resulting Lithological Classification – Practical Template, G3 group, IODP-Expedition 362, Site U1481. The log parameters of the well GRA, P-wave, MS, L*, a*, and b* are displayed with respect to the real lithology related to the depth presented. Graph A displays the total number of records per method vs. lithology. Graph B displays the resulting lithological classification by method relative to the real lithology. The methods are MLP, RandomForest, DecisionTree, and SVM, presented in the sequence of identification of the lithologies. (Aa) to (Dd) correspond to the method distribution with percentage of hit by lithology. Litho1 refers to Lithology Code 10, Litho3 refers to Lithology Code 12, Litho4 refers to Lithology Code 13 and Litho5 refers to Lithology Code 14. Litho2 and Litho6 are not recorded in this hole range. The SRM log parameter is not included.

It should be noted that the G2 group obtained better mean results in accuracy for the four methods. The RandomForest method obtained an excellent result for the Practical

462 Template, with more than 79.00% accuracy for classification of lithologies in GP and
463 G1 groups. Emphasis is placed on the MLP method with 85.08% accuracy in the G2
464 group.

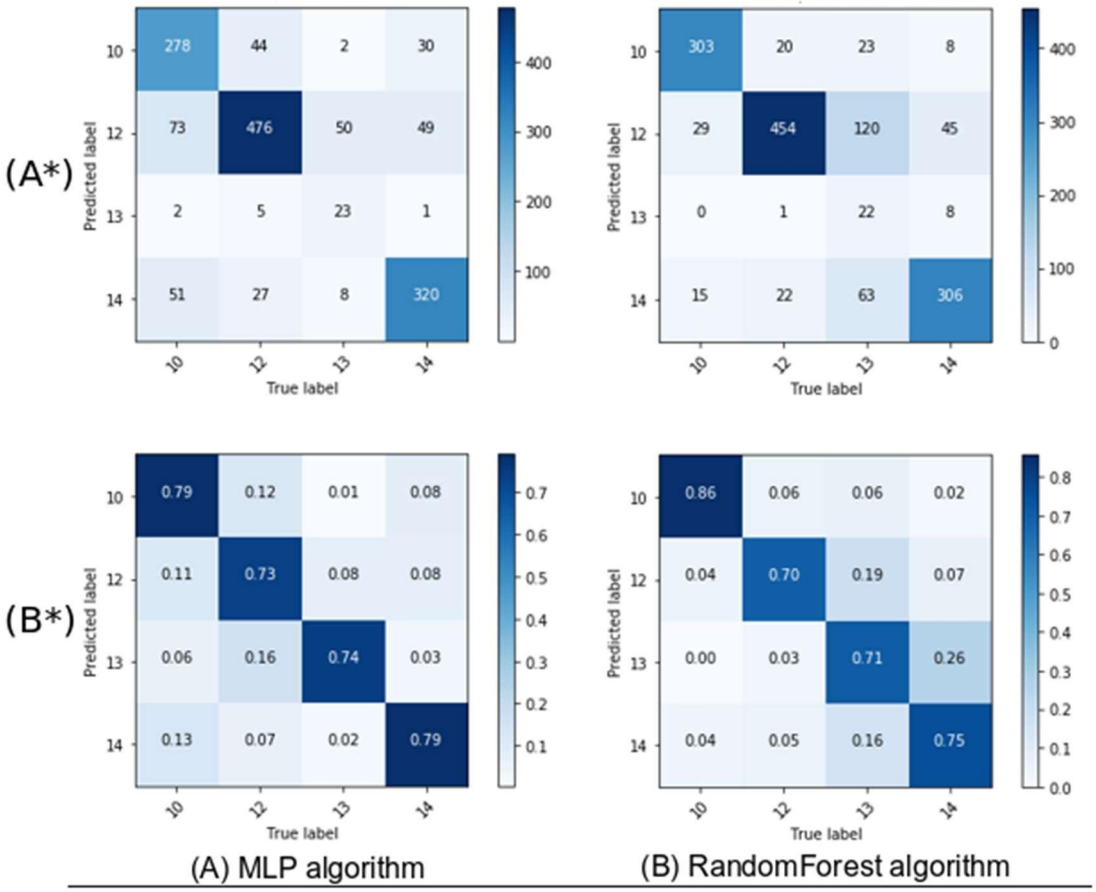
465 IODP-Expedition 362, Site U1481 obtained the best result for the lithological
466 classification in the G3 group, MLP method. IODP-Expedition 361, Site U1478
467 obtained the best result in the GP group, SVM method and site U1477 obtained the
468 best result in the G2 group, MLP method. IODP-Expedition 356, site U1462 obtained
469 the best result in the G2 group, MLP method. IODP-Expedition 349, site U1431
470 obtained the best result in the G2 group, MLP method, site U1432 obtained the best
471 results in the G2 group, MLP and DecisionTree methods, site U1433 obtained
472 excellent results in the G2 group, MLP method. Therefore, it was concluded that the
473 best method for the Practical Template is the MLP, with best results in 85.70% of the
474 total Practical Template.

475 The confusion matrix highlighted in Fig. 8 (related to Fig. 7) represents the method
476 result of the respective lithological classification for IODP-Expedition 362, site U1481
477 (G3 group). The MLP method (A) has the best average result for the hole, with
478 accuracy of 76.30%. For Litho1 (Lithology Code 10) the best accuracy was obtained
479 by the RandomForest method (B). For Litho3 (Lithology Code 12) the best accuracy
480 was obtained by the MLP method (A). For Litho4 (Lithology Code 13) the best accuracy
481 was obtained by the MLP (A) and SVM (D) methods. Finally, for Litho5 (Lithology Code
482 14) the best accuracy was obtained by the MLP (A) and DecisionTree (C) methods.
483 The complete table with classification metrics is found in Appendix 1E of the
484 supplementary material.

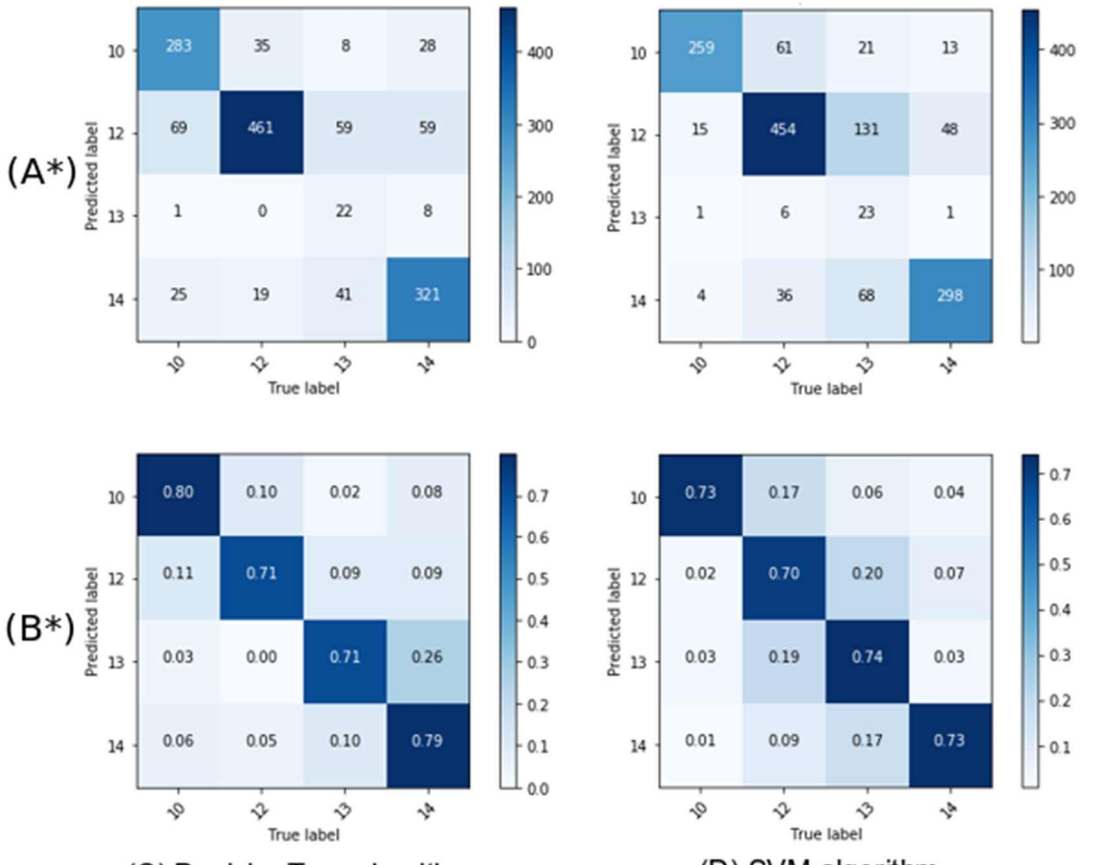
Figure 9 shows the ROC curve representing four methods for the G3 group, Practical Template in IODP-Expedition 362, Site U1481, divided by models (Litho1, Litho3, Litho4, Litho5) and lithological code (10,12,13, and 14). The results are good in all plottings, with a positive ROC curve above the main diagonal (0,1) and corroborate with precise identification (from 70.00% to 86.00%) of the lithologies as seen in Fig. 8. The MLP method has a better and growing stability of the plotted results relative to the other methods.

The best result for the lithological classification linked to the G2 group is due to the fact that the grouping was organized into four lithologies: Litho1, Litho2, Litho3, and Litho4, which formed a larger organization, agglomerating more lithologies and facilitating the recognition by the methods. The GP group did not obtain good results in the general context, due to a greater refinement of the division and grouping of the lithologies, making it difficult to recognize the analyzed methods.

Table 5 summarizes the results of machine learning processing methods on the groups and data templates proposed in this article. The best method was RandomForest. The best data group was the G2 group. The best template was Template2. Referring to Group vs. Method, in the GP group the best method was RandomForest, in the G1 group the best method was RandomForest, in the G2 group the best method was MLP and in the G3 group the best method was RandomForest. In the combination between Template vs. Method, in Template1, Template2, and Template3 the best method was RandomForest, and in the Practical Template the best method was MLP. The best result in applying cross-validation was the RandomForest method.

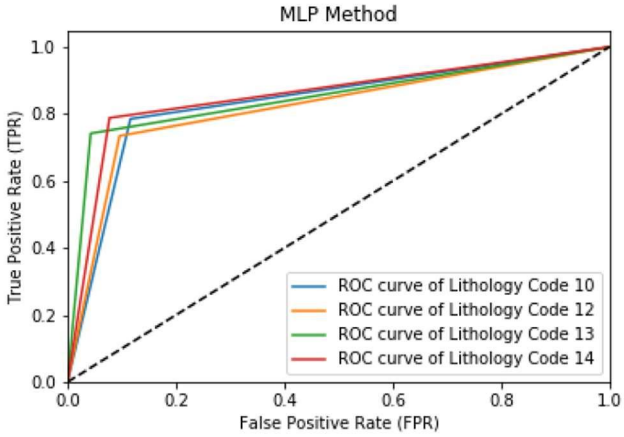


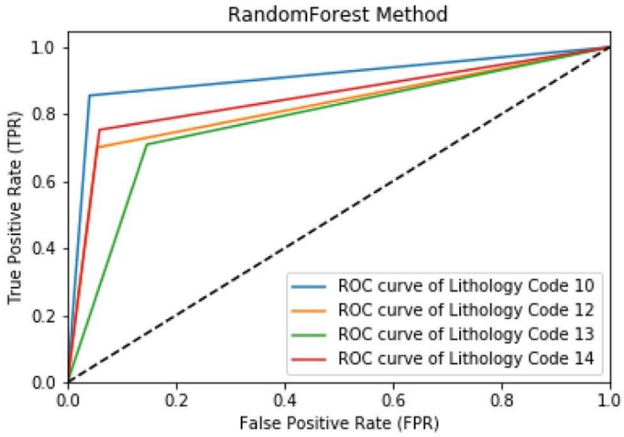
(A) MLP algorithm (B) RandomForest algorithm

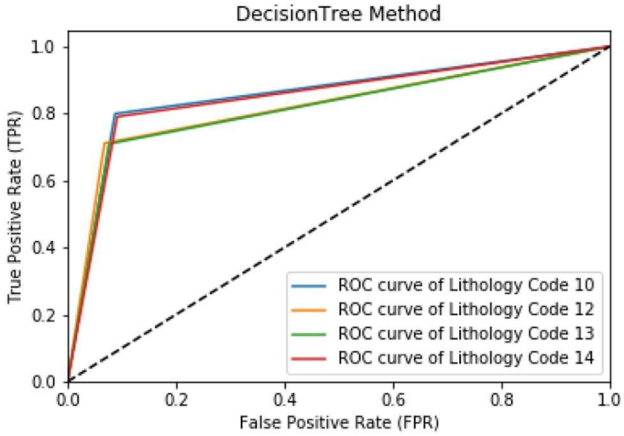


(C) DecisionTree algorithm (D) SVM algorithm

Figure 8 Confusion matrix for the G3 group, Practical Template (IODP-Expedition 362, Site U1481). (A) MLP method, (B) RandomForest method, (C) DecisionTree method, (D) SVM method. A to D represents the diagonal matrix for the result of each method, where the MLP method had the best result. (A*) represents the number of true and predicted G3 group (Lithology Code 10, 12, 13, 14) lithology records without normalization. (B*) highlights the percentage of normalized data between true and predicted. Colors represent the frequency of normalized and non-normalized records. Lithology Code 11 and Lithology Code 15 are not recorded in this hole range.


MLP Method
True Positive Rate (TPR)
False Positive Rate (FPR)
ROC curve of Lithology Code 10
ROC curve of Lithology Code 12
ROC curve of Lithology Code 13
ROC curve of Lithology Code 14


RandomForest Method
True Positive Rate (TPR)
False Positive Rate (FPR)
ROC curve of Lithology Code 10
ROC curve of Lithology Code 12
ROC curve of Lithology Code 13
ROC curve of Lithology Code 14


DecisionTree Method
True Positive Rate (TPR)
False Positive Rate (FPR)
ROC curve of Lithology Code 10
ROC curve of Lithology Code 12
ROC curve of Lithology Code 13
ROC curve of Lithology Code 14

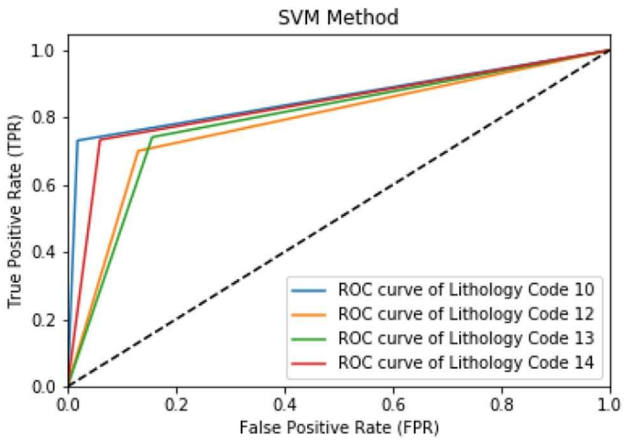

SVM Method
True Positive Rate (TPR)
False Positive Rate (FPR)
ROC curve of Lithology Code 10
ROC curve of Lithology Code 12
ROC curve of Lithology Code 13
ROC curve of Lithology Code 14

Figure 9 ROC Curve of the MLP, RandomForest, DecisionTree, and SVM methods applied to the G3 group, Practical Template (IODP-Expedition 362, Site U1481). Lithology Code 10 refers to the G3 Group, Models Litho1. Lithology Code 12 refers to the G3 Group, Models Litho3. Lithology Code 13 refers to the G3 Group, Models Litho4. Lithology Code 14 refers to the G3 Group, Models Litho5. Lithology Code 11 and Lithology Code 15 are not recorded in this hole range.

Table 5 Summary table with the best accuracy per method, group, and template.

Table summary best accuracy	
Method:	
RandomForest	
Group:	
G2	
Template:	
Template2	
Group vs. Method:	
GP	RandomForest
G1	RandomForest
G2	MLP
G3	RandomForest
Template vs. Method:	
Template1	RandomForest
Template2	RandomForest
Template3	RandomForest
Practical Template	MLP
Cross-Validation:	
RandomForest	

The ash layer was not part of the division of the lithological groups, since its presence in the lithological profile of the wells is restricted to mm- and cm thick layers, with few log records and inadequate resolution, making it impossible for machine learning applications in all of the configurations used at this stage. Further details referring to the ash layer records are found in Appendix 1F of the supplementary material with ash layer records for the IODP-Expedition 362, site U1480.

4 Conclusions

In this study, four different machine learning methods were applied to three standard data templates and a practical data template in a lithological classification problem for wells from IODP-Expeditions.

The results indicate that the MLP method had better results in the lithological classification for the Practical Template, taking into account the lithology of the G2 group proposed in this research and the characteristics of the data sets used.

In data sets Template1, Template2, and Template3, the best results were in the G1 group with the RandomForest method, with an accuracy of more than 85.00%. The G1 group had better results for lithological classification for the organized templates, due to the advantageous grouping of the three lithologies (Litho1, Litho2, and Litho3). The Litho2 grouping composite for joining lithologies between silt and clay increased the accuracy according to the quantity of the dataset, as it facilitated the execution of the classifications by the methods proposed.

The SVM method obtained poor results in cross-validation analysis and accuracy. The geological data analyzed do not have a highly defined pattern, and this method depends on many external and natural factors to obtain excellent results in its application. To successfully run machine learning with this method, it is necessary to have a data set with a large amount of data, and the data must be balanced regarding each training for a specific group of data. Indeed, we must perform an adjustment of the configuration parameters, making it difficult to use this for classification of lithologies as proposed in this study.

The best result for cross-validation was obtained by the RandomForest method in all groups and templates analyzed. The characteristics of the RandomForest method

allowed for better results of lithological classification, as it has a simple configuration and its design includes a set of several decision trees, leading to accurate results with smaller statistical variance.

The main contributions of this work are a quick approach to lithological classification in offshore wells, the proposal of a lithological classification using supervised training methods, lithological classification using multivariate data and the support of a large number of variables, and strengthening and improving learning using neural network methods and machine learning.

In all processes, there are limitations that create restrictions on achieving results with more than 80.00% accuracy (PENG and BAI, 2017; PAPERNOT et al., 2016), especially due to the characteristics of the data sets, the amount of data in the data sets, the number of features (GRA, PWL, MS, L*, a*, b*), and the inability to store training and testing for future use, which strengthens datasets and improves results.

Acknowledgement

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. This research used data provided by the International Ocean Discovery Program (IODP) (www.iodp.org/access-data-and-samples).

Computer Code Availability

Name of code: LithoPy. Developer: Thiago Santi Bressan. Contact address: Programa de Pós-Graduação em Geologia, Universidade do Vale do Rio dos Sinos (Unisinos), Av. Unisinos, 950, Cristo Rei, São Leopoldo, RS, Brasil. Telephone number: +55-55-99601-6723. e-mail. tsbressan@gmail.com. Year first release: 2019. Hardware required: I3 CPU or better with 16GB memory RAM. Software required: Anaconda

(>=5.3), Jupyter Notebook (>=5.2), Sklearn (>=0.20), NumPy (>=1.8.2), SciPy (>= 0.13.3), Matplotlib (>=3.0.0) and Windows (10) or Linux (Ubuntu or like system). Program language: Python. Program size: 10 MB. How to access the source code: Available at: https://github.com/tsbressan/LithoPy.

Abbreviations used in this manuscript

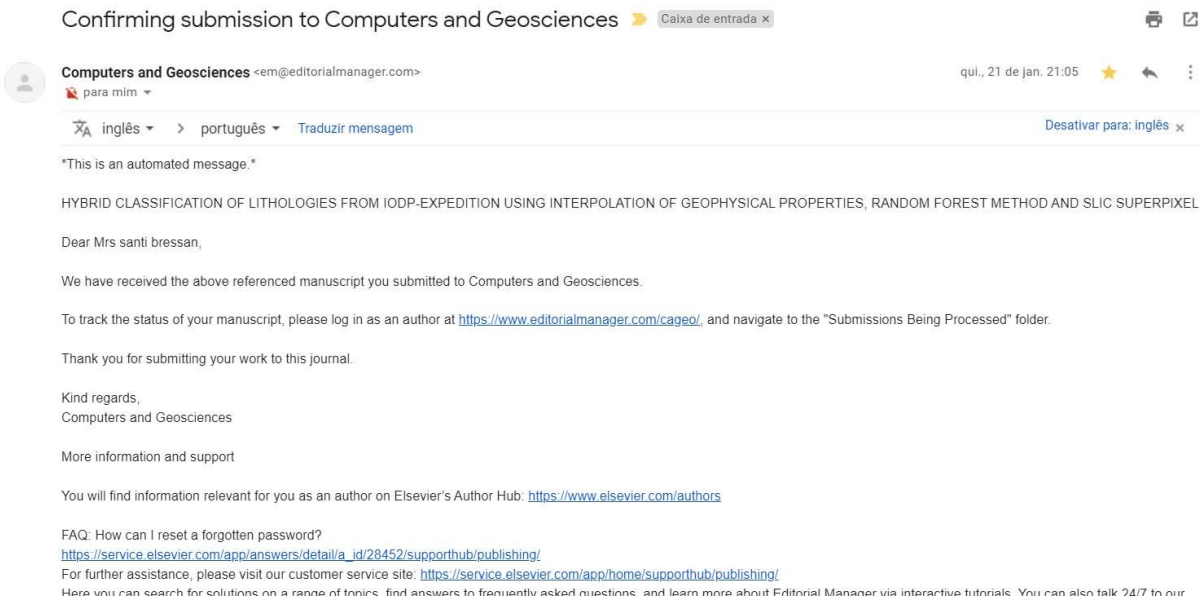
ANN Artificial Neural Network
CIE International Commission on Illumination
GRA Gamma Ray Attenuation
IODP International Ocean Discovery Program
MS Magnetic Susceptibility
MLP Multi-Layer Perceptron
NRM Natural Remanence Magnetization
PWL P-Wave Velocity Logger System
RSC Reflectance Spectrophotometry
ROC Receiver Operating Characteristic
SRM Magnetic Remanence
SVM Support Vector Machine
VCD Visual Core Description

Appendix A – Supplementary data

Supplementary data related to this article can be found at https://github.com/tsbressan/LithoPy.

Manuscript II

Confirming submission to Computers and Geosciences



Confirming submission to Computers and Geosciences Caixa de entrada ×

Computers and Geosciences <em@editorialmanager.com> qui., 21 de jan. 21:05

para mim

inglês › português Traduzir mensagem Desativar para: inglês ×

This is an automated message.

HYBRID CLASSIFICATION OF LITHOLOGIES FROM IODP-EXPEDITION USING INTERPOLATION OF GEOPHYSICAL PROPERTIES, RANDOM FOREST METHOD AND SLIC SUPERPIXEL

Dear Mrs santi bressan,

We have received the above referenced manuscript you submitted to Computers and Geosciences.

To track the status of your manuscript, please log in as an author at https://www.editorialmanager.com/cageo/, and navigate to the "Submissions Being Processed" folder.

Thank you for submitting your work to this journal.

Kind regards,
Computers and Geosciences

More information and support

You will find information relevant for you as an author on Elsevier's Author Hub: https://www.elsevier.com/authors

FAQ: How can I reset a forgotten password?
https://service.elsevier.com/app/answers/detail/a_id/28452/supporthub/publishing/
For further assistance, please visit our customer service site: https://service.elsevier.com/app/home/supporthub/publishing/
Here you can search for solutions on a range of topics, find answers to frequently asked questions, and learn more about Editorial Manager via interactive tutorials. You can also talk 24/7 to our

HYBRID CLASSIFICATION OF LITHOLOGIES FROM IODP-EXPEDITION USING INTERPOLATION OF PETROPHYSICAL PROPERTIES, RANDOM FOREST METHOD AND SLIC SUPERPIXEL

Thiago Santi Bressan[2], Farid Chemale Junior, Sandro Jose Rigo

Universidade do Vale do Rio do Sinos – Unisinos, São Leopoldo, RS, Brasil

ABSTRACT:

Recent advances in the construction of computers and software with high processing and storage combined with large data production allow the geological professional to make decisions based on discovering new information. This study applies the hybrid method to classify lithologies in IODP-Expedition using interpolation of petrophysical data, Random Forest method, and SLIC Superpixel. The division of data includes the creation of three datasets. Each dataset is replicated in two lithology groups. For dataset2 new combinations are added between the features forming 102 practical arrangements. The petrophysical data used include PWL, NGR, GRA, RSC, MAD, MS, and RGB, whose values are interpolators using Linear, Quadratic, Cubic, and Spline Slinear, Akima, Pchip, and Piecewise. SLIC Superpixel segmentation groups image pixels in regions extracting color and texture characteristics. The division of datasets for training and testing comprises combinations between training and testing with specific combinations of the number of trees and the maximum depth of trees. The results show that the best interpolator evaluated is the Akima, with accuracy equal to 98.22%. For dataset1, U1480, the accuracy value is 96.96% in combination of 70% training and 30% testing in Group 1 and 97.71% combined with 70% training and 30% testing in Group 2. For dataset1, U1481, the accuracy value is 99.68% combined with 80% training and 20% testing in Group 1 and 99.74% combined with 80% training and 20% testing in Group 2. In dataset2, U1480 and U1481, the best accuracy (99.75%) is in the combination of 90% training and 10% testing and practical arrangement equal to 51. We demonstrated that petrophysical data's interpolation is valid and necessary when there is little data for training. Dataset2 resulted in excellent accuracy results creating a new format for generating data about images in IODP-Expedition.

KEYWORDS:

SLIC Superpixel, Random Forest, classification lithology, IODP

HIGHLIGTHS:

[2] Corresponding author
E-mail address: tsbressan@gmail.com

Authorship statement: Thiago Santi Bressan designed and developed the algorithms, development and adjusted the figures, worked on the main writing of the manuscript. Sandro Jose Rigo contributed to the statistical analysis, algorithm tests, review and writing of the manuscript. Farid Chemale Junior contributed to the writing and revision of the manuscript.

Integration of physical properties and core images provides a robust Automatized Lithological Hybrid Classification

Use of structured and semi-unstructured data in the lithological classification in IODP-Expedition with average accuracy greater than 90.0%

The best interpolator for petrophysical data is Akima interpolator, with accuracy equal to 98.22%

Lithological classification on a segmented image using SLIC Superpixel with a hit rate of up to 99.75%

1. Introduction

Expeditions carried out by the International Ocean Discovery Program (IODP) cover a vast field of research on ocean platforms, collaboratively collecting and disseminating data recorded in sediments and sedimentary rocks present on the seafloor.

Artificial intelligence methods require in their structure that the data be organized and standardized in a format suitable for the desired processing, whether in text, numeric, or image format. For lithological classification, the data needs to be stored in appropriate datasets and able to integrate the algorithm's training and testing database for Machine Learning (ML). Strengthening and creating new data requires care and support for accurate and validated datasets. ML processing can help professional geologists use datasets with a large amount of data and strengthen the expected lithological classification. In this case, the bulky dataset is increased through interpolation techniques, guaranteeing new and pertinent data for joint computational processing in the lithological characteristics.

The purpose of this article is to develop a hybrid method for classifying lithologies of the IODP-Expedition using the Random Forest Method with data of interpolated petrophysical data and image processing SLIC Superpixel segmentation.

1.1 Related works of Lithological Classification in Sedimentary Rocks

The hybrid format in the lithological classification environment integrates a broad data group multivariates and integrated with characteristics of the analyzed material. This study adds innovation in the context of applying hybrid lithological classification using petrophysical data and high-resolution images and in the context of that multiple information provided by IODP-Expeditions compared to other equivalent studies, as summarized in Table 6. Such study makes it a differential due to the accuracy of the results and the great range of data combinations with precision and added value to reduce time, fewer working hours, and financial resources in the lithological classification process. The Random Forest (RF) method's choice is because it has already been tested on data from the same IODP-Expedition with excellent results (BRESSAN et al., 2020). The ML, being a robust and floating system, can adapt to the processing of new data, adjust the characteristics of the problem features and act in a scaled way to manage large amounts of data (GÉRON, A., 2017). SLIC Superpixel is a modern and adaptable method to the type and format of image data extracted from the expedition, integrating the manipulation and grouping of pixels quickly and precisely in geological data (HARALICK, SHANMUGAM and DINSTEIN, 1973; ACHANTA et al., 2010).

Table 6 Summary of related works of Lithological Classification in Sedimentary Rocks.

Authors	**IODP data**				**Interpolation - data increase**	**Supervised and unsupervised Methods**	**Best accuracy (%)**
	Petrophysical data		**Images**				
	IODP	**Others**	**IODP**	**Others**			
BRESSAN et al., 2020	Yes	No	No		No	MLP, DT, RF, SVM	89.51
BRESSAN, CHEMALE and RIGO, 2021	Yes	No	Yes	No	Yes	RF+SLIC Superpixel	99.75
AO et al., 2020	No	Yes	No		No	LDA, QDA, GMM, KDE,	92.95

					NBAYES, TAN, AODE, PNN and PrRF	
TSE et al., 2018	Yes	No	No	No	K-means, SOM, HC, RF	86.90
SILVA et al., 2020	No	Yes	No	No	DT, RF, GB, KNN, NB	92.06
XIE et al., 2017	No	Yes	No	No	SVM, ANN, RF, GTB, NB	81.10
OJHA and MAITI, 2016	Yes	No	No	No	CA, PCA, ANN, BNN	87.55
ALZUBAIDI et al., 2020	No		No	Yes	No	CNN	93.12
FAN et al., 2020	No		No	Yes	No	CNN	94.55
KUMAR et al., 2020	No		No	Yes	No	LDA, RF, SVM	85.48
BACHRI et al., 2019	No		No	Yes	No	SVM	85.00
CRACKNELL and READING, 2014	No		No	Yes	No	NB, RF, KNN, SVM, ANN	92.40

Studies in the area of lithologies classification using SLIC Superpixel in images and jointly integrating petrophysical data from IODP-Expeditions were not carried out yet.

1.2 Geological Background

The studied area is in the Indian Ocean between the Ninety-East Ridge (NER) and Sumatra Subduction Zone (Fig. 9). In this region, the Nicobar depositional system, separated by the Ninety-East Ridge from the Bengal Fan (CURRAY et al., 2003, MCNEIL et al., 2017, PICKERING et al., 2019). The Nicobar Fan is genetically connected to the Bengal Fan forming the largest submarine fan in the world (CURRAY AND MOORE, 1974; CLIFT et al. 2008). The fan feeder system of Bengal and Nicobar fans record mainly the denudated material of the Himalayas Mountains that they are result of the India-Asia collision onset ~50 Ma (HINSBERGEN et al., 2012). The sediments transport occurred through the Ganges and Brahmaputra rivers, from

95 northeastern India and the Central-Eastern Tibetan-Himalayan Orogen. The large
96 contribution occurs since the hard continent-continent collision at begin of Miocene
97 (HINSBERGERN et al., 2012) and supply the system with an enormous amount of
98 clastic material ($1*10^9$ t/yr.; MILLIMAN AND SYVITSKI, 1992).
99 In this study, we used material provided during the IODP-Expedition 362 - Sumatra
100 Seismogenic Zone: Two sites drilled by the IODP-Expeditions record the pre-Nicobar
101 Fan and the Nicobar Fan rocks. The pre-Nicobar rocks consist, from base to the top,
102 of Late Cretaceous MORB basalts lava (67 Ma), pelagic sedimentation and Plume-
103 related volcanic and intrusive alkaline rock. They record the drifting process of the
104 Indian Ocean and plume-related Kergeulen hot spot during the Cretaceous through
105 Paleocene (MCNEIL et al., 2017, GIRELLI et al., 2021b). Overlying units, represented
106 by Nicobar Fan unconsolidated sediments and sedimentary rocks with intercalation of
107 ash/tuff layers, is directed connected to Himalayas-Tibet orogenic system. The Nicobar
108 Submarine Fan is divided in three units (Unit III, II and I) and eight facies associations
109 (PICKERING et al. 2019), corresponding to siliciclastic sandstone/siltstone, tuff layers,
110 chalk, tuffaceous mudstones and hemipalgic sediments/sedimentary rocks deposited
111 between early Oligocene to Quaternary (MCNEIL et al., 2017, PICKERING et al.,
112 2019). Further description of the studied area with the description of Units I, II, and IIII and
113 Facies Association (FA) from Nicobar Fan and drill core recovery can be found in the
114 SM1_FA_Geology.

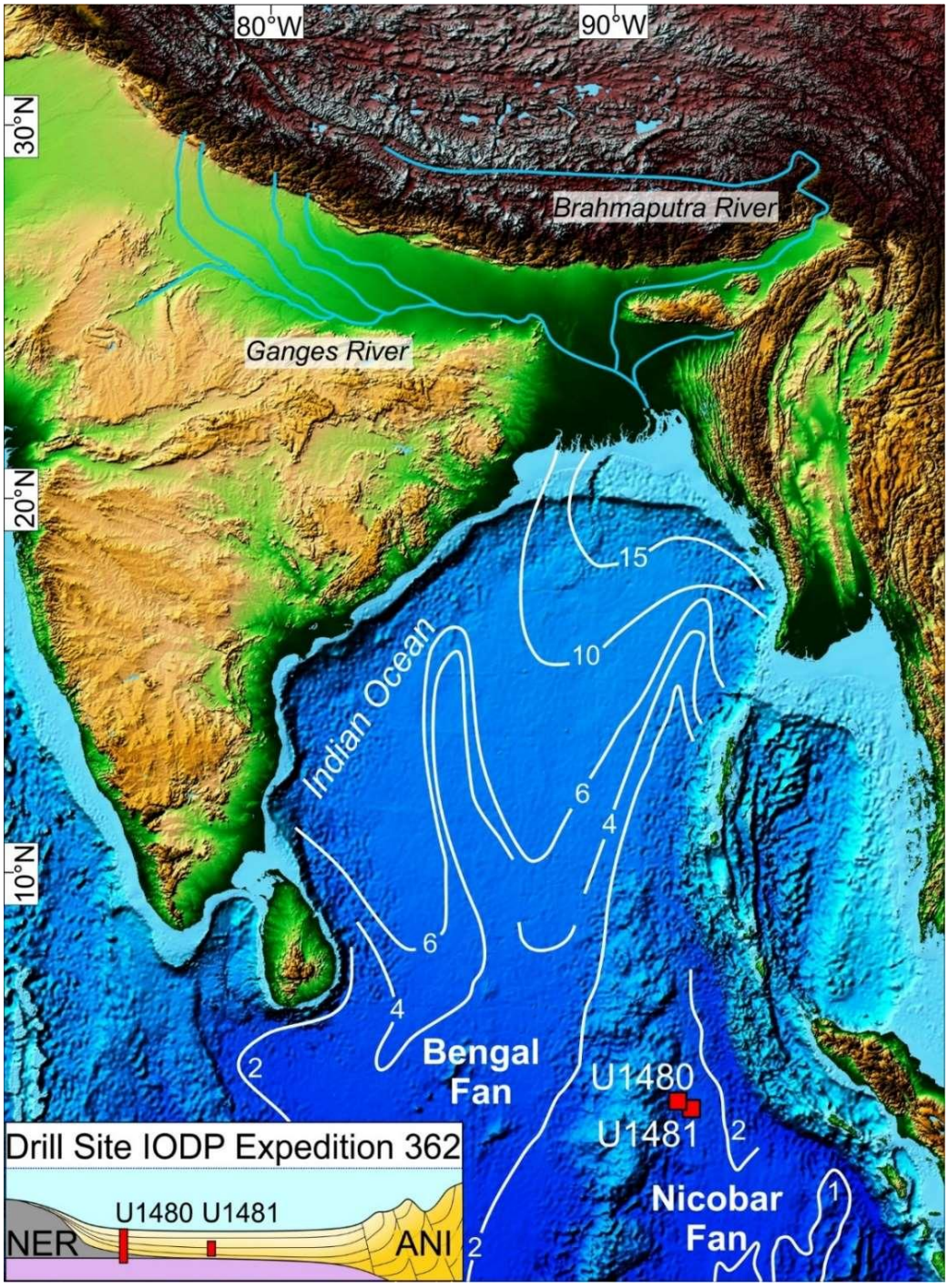


Figure 9 Map of South and East Asia showing the Bengal and Nicobar fans' main feeder system, the Ganges and Brahmaputra rivers, India, Himalayas, and Sumatra Subduction Zone. The Nicobar Fan is located between the Ninetyeast Ridge and the Sunda subduction zone. IODP-Expedition 362 sites U1480 and 1481 used for the present study are shown as red squares. Isopach (white lines) are in km (after CURRAY et al., 2003). Inset shows a schematic model for the transection Ninety-East Ridge (NER), Nicobar Fan Andaman Nicobar Islands (ANI) based on the seismic data of MCNEIL et al., 2017 (modified from GIRELLI. et al., 2021a).

2. Material and Methods

2.1 Material

In the drilling process of sites U1480 and U1481, the collected rock samples were subjected to techniques to capture the petrophysical data Gamma Ray Attenuation bulk density (GRA), Moisture and Density (MAD), Magnetic Susceptibility (MS), Natural Gamma Radiation (NGR), P-Wave Velocity Logger System (PWL), Red Green Blue channels (RGB) and Reflectance Spectrophotometry and Colorimetry (RSC). Add up the list of petrophysical data, high resolution images generated by imaging the core recovery directly at the IODP-Expedition.

The description of the material used in this article is found in the SM2_material.

2.2 Methods

2.2.1 Data Configuration

The flowchart of data configuration and formation of datasets between petrophysical data and images is observed according to Fig. 10.

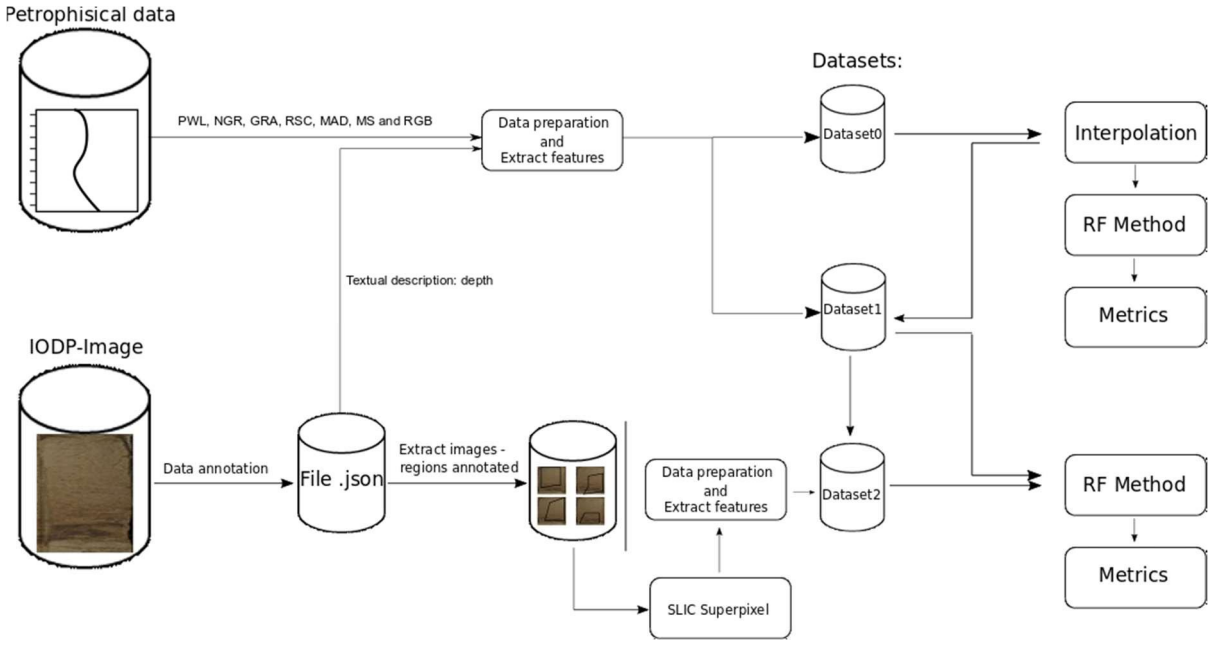


Figure 10 Model of configuration flowchart and formation of datasets

The values of the petrophysical data PWL, NGR, RSC, MAD, MS and RGB are read directly from the original expedition files in the depth range according annotation in the image, with data preparation and extraction of features are performed by the algorithm in Python programming language.

The images generated in the IODP-Expedition 362 are in the 96 DPI resolution setting, 24 bits of color intensity in the RGB channels (Red, Green and Blue), divided by sites U1480 and U1481, holes U1480: E, F, G, H and U1481: A, cores and sections (configured according to the sample size extracted from the hole and numbered sequentially), totaling 685 images from the site U1480 and 128 images from the site U1481. (MCNEILL et al., 2017). The image list and description to download images is found in the supplementary material, file: Computer Code Availability/instruction guide for use. On each image is used online software VIA (DUTTA and ZISSERMAN, 2019) to perform data annotation (lithology and depth) and save file .json with the annotated regions. The SLIC Superpixel method as well as the application of data preparation and feature extraction are performed by the algorithm in Python programming language.

The designation of the types of lithology used follows the standard defined by the methods of IODP-Expedition 362 (MCNEILL et al., 2017) integrate sets of 14 lithologies in consolidated and unconsolidated sedimentary rocks organized into two groups (Group 1 and Group 2). Group 1 covers a broad context and directed to the moment of drilling with a complete division of all the present lithologies. Group 2 provides an organization of the lithologies by similarity and organization between the layers. Table with a summary of original images, annotated images and division of lithologies are found in the SM3_method, table 15 and 16. The file output .json is found in the supplementary material, directory Computer Code Availability/json.

The total number of records collected is divided into three datasets separated by site, hole, lithology group and extracted features, according table 7. For Cod_Feature: Feature1, the name of the features follows the pattern: value_ <Seq> _ <Petro>, where <Seq> is the sequence (quantity) of the features by petrophysical data and <Petro> is the petrophysical data. The quantity of feature by petrophysical data: GRA one feature, MAD six features, MS one features, NGR one feature, PWL one feature, RGB three features and RSC three features. For Cod_Feature: Feature2, the name of the features follows the pattern: <codText> _D_ <dist> _an_ <ang>, where <codText> refers to the texture property used as being C: contrast, Di: Dissimilarity, H: Homogeneity, A: ASM, E: Energy and Cor: Correlation. <dist> refers to the pixel distance related to the GLCM matrix that is defined in 1, 2 and 3. <ang> is the definition of the angle of analysis of the texture properties defined in 0º, 45º, 90º, 135º.

Table 7 Structure of the formation of datasets between site, hole, lithology groups and features.

Cod_features	Features
Feature1	Offset, value_1_gra, value_1_mad, value_2_mad, value_3_mad, value_4_mad, value_5_mad, value_6_mad, value_1_ms, value_1_ngr, value_1_pwl, value_1_rgb, value_2_rgb, value_3_rgb, value_1_rsc, value_2_rsc, value_3_rsc
Feature2	C_D_1_an_0, C_D_1_an_45, C_D_1_an_90, C_D_1_an_135, C_D_2_an_0, C_D_2_an_45, C_D_2_an_90, C_D_2_an_135, C_D_3_an_0, C_D_3_an_45, C_D_3_an_90, C_D_3_an_135, Di_D_1_an_0, Di_D_1_an_45, Di_D_1_an_90, Di_D_1_an_135, Di_D_2_an_0, Di_D_2_an_45, Di_D_2_an_90, Di_D_2_an_135, Di_D_3_an_0, Di_D_3_an_45, Di_D_3_an_90, Di_D_3_an_135, H_D_1_an_0 H_D_1_an_45, H_D_1_an_90, H_D_1_an_135, H_D_2_an_0, H_D_2_an_45, H_D_2_an_90, H_D_2_an_135, H_D_3_an_0, H_D_3_an_45, H_D_3_an_90, H_D_3_an_135, A_D_1_an_0, A_D_1_an_45, A_D_1_an_90, A_D_1_an_135, A_D_2_an_0, A_D_2_an_45, A_D_2_an_90, A_D_2_an_135, A_D_3_an_0, A_D_3_an_45, A_D_3_an_90, A_D_3_an_135, E_D_1_an_0, E_D_1_an_45, E_D_1_an_90, E_D_1_an_135, E_D_2_an_0, E_D_2_an_45, E_D_2_an_90, E_D_2_an_135, E_D_3_an_0, E_D_3_an_45, E_D_3_an_90, E_D_3_an_135, Cor_D_1_an_0, Cor_D_1_an_45, Cor_D_1_an_90, Cor_D_1_an_135, Cor_D_2_an_0, Cor_D_2_an_45, Cor_D_2_an_90, Cor_D_2_an_135, Cor_D_3_an_0, Cor_D_3_an_45, Cor_D_3_an_90, Cor_D_3_an_135, mean_intensity, mean_offset, value_1_gra, value_1_mad,

			value_1_ms, value_1_ngr, value_1_pwl, value_1_rgb, value_1_rsc, value_2_mad, value_2_rgb, value_2_rsc, value_3_mad, value_3_rgb, value_3_rsc, value_4_mad, value_5_mad, value_6_mad
Dataset	**Site-Hole**	**Group**	**Cod_features**
Dataset0	U1480-E	1	**Feature1**
Dataset1	U1480-E,F,G,H	1	**Feature1**
Dataset1	U1480-E,F,G,H	2	**Feature1**
Dataset1	U1481-A	1	**Feature1**
Dataset1	U1481-A	2	**Feature1**
Dataset2	U1480-E,F,G,H	1	**Feature2**
Dataset2	U1480-E,F,G,H	2	**Feature2**
Dataset2	U1481-A	1	**Feature2**
Dataset2	U1481-A	2	**Feature2**

The application of the eight selected interpolation methods on the datasets organized with the petrophysical data standardized and increased the data intervals between property value and reading depth. The interpolated data ranges are specific to each petrophysical data, as noted in the SM3_method, Fig. 41.

Dataset0 covers the formation of a temporary validation dataset seeking to identify the best interpolator on the types of data analyzed. For this, dataset0 has values of the petrophysical data of the site U1480, Hole E, Group 1, being tested with values of all analyzed interpolators. The configuration of the interpolation and depth combinations follows the pattern of the dataset0 is found in the SM3_method, table 17.

Dataset1 contains all the interpolated records of the 7 petrophysical data for site U1480, Hole E, F, G and H, and U1481, Hole A, divided by lithology group (Group 1 and Group 2). The number of original records of petrophysical data and the number of records generated by the interpolation at each site and hole is found in the SM3_method, table 17 and 18.

Dataset2 includes records extracted from the processing of annotation images by lithology, through the SLIC Superpixel with analysis of texture properties and color combinations in each segmented area. In dataset2, features petrophysical data of dataset1 are added according to the depth of the segmentation of the image, by averaging the values of the records. Fig. 11 shows the data selection scheme for a segmented area. The SM3_method, table 19, summarizes the formation of datasets after interpolation, number of lithology records by group and quantity of features.

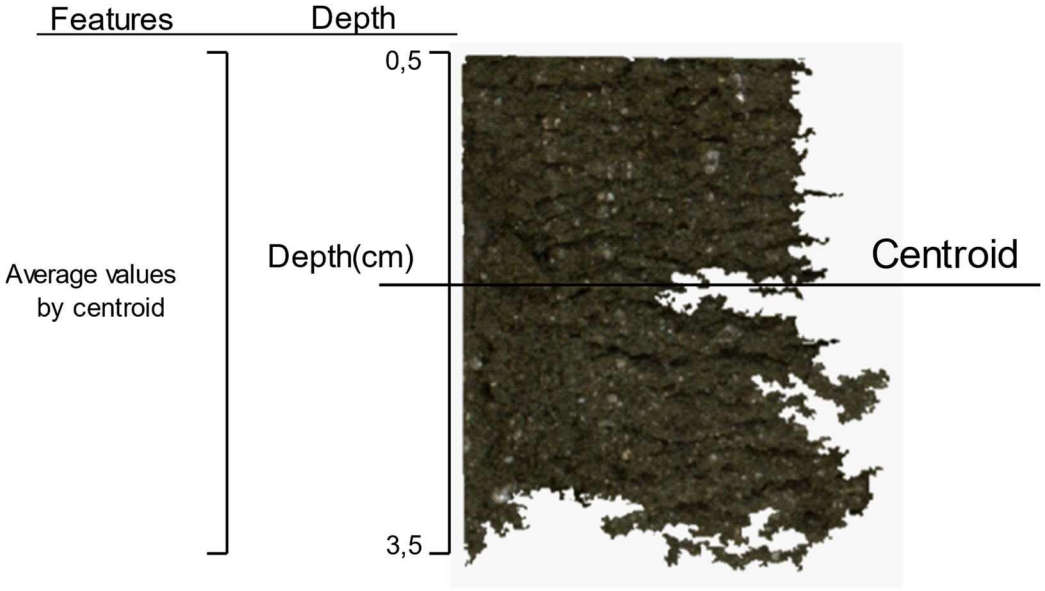


Figure 11 Segmentation region extracted from the image: 362-u1480e-2h-2-a_shlf7853651_20160813145856.jpg for the formation of dataset2. Lithology: Medium-sand_sandstone cod_region: 22, Depth: 0,5 cm a 3,5 cm. Features include all values extracted from dataset2. Centroid is the average of the depth values and texture properties.

2.2.2 Training and Testing

To carry out the training and testing with the proposed datasets, the data were divided into different combinations randomly. For dataset0, dataset1 and dataset2 the data combinations cover four-division configurations for training and testing: 90%-10%, 80%-20%, 70%-30% and 60%-40%. In each division, the RF algorithm is configured

between maximum depth of trees, number of trees in the forest, input features and output feature as shown in the table below.

Table 8 Configuration of the RF method between the maximum depth of trees, the number of trees in the forest, input features and output feature. Max_depth refers to the maximum depth of trees. cod_lit refers to the lithology code according to the division of the groups.

Dataset	Site	Max_depth	Number of Trees	Input Features	Output Features
Dataset0	U1480	20	1000	All	cod_lit
Dataset1	U1480	10, 50, 100, 1000 and 2000	5-100, 1000 and 2000	All	cod_lit
Dataset1	U1481	10, 50, 100, 1000 and 2000	5-100, 1000 and 2000	All	cod_lit
Dataset2	U1480	10, 50, 100, 1000 and 2000	5-100	According combination of practical arrangement	cod_lit
Dataset2	U1481	10, 50, 100, 1000 and 2000	5-100	According combination of practical arrangement	cod_lit

The dataset0 the configuration follows the pattern used in BRESSAN et al., 2020, with a maximum depth of trees of 20 and a number of trees of 1000.

The dataset2 has extra combinations of training and testing among the features that compose it, resulting in 102 Practical Arrangements (Practical_Arrang) with specific results in each combination as is found in the supplementary material, file: Computer Code Availability/Python code/practical_arrangement_to-dataset2. This combination of practical arrangements envisages the possibility of expanding the combinations in the set of features, seeking to identify separate results in each practical arrangement.

2.2.3 Library and Programming Language

In this machine learning study, Python programming language is with the Scikit-learn, Scikit-image and wxPython programming libraries. In Scikit-learn the module used is an ensemble.RandomForestClassifier. For Scikit-image, it covers the use of several specific modules for image processing and data selection. For segmentation, the module used is skimage.segmentation.slic with a configuration of up to 400 segmentation regions. Instructions for using the code and the graphic tool can be found in the supplementary material, file: Computer Code Availability/instruction guide for use.

3. Results

The annotation of the images covers the analysis of the image of the rocks collected at the sites in a total of 813 original image files. On the site U1480, 11.82% are Hole E images, 37.52% are images of Hole F, 36.06% are images of Hole G and 14.60% are images of Hole H. For the site U1481, 100% of the images are from Hole A. For the images annotation, the total number of images extracted for the two sites is 2601 with Apx of 421,514.34 cm² to site U1480 and 87,429.30 cm² to Site U1481. The summary of the composition of the images and annotations are observed in the SM4_results, Fig. 42. The division of lithologies by annotation and a group of lithologies is shown in Fig. 12.

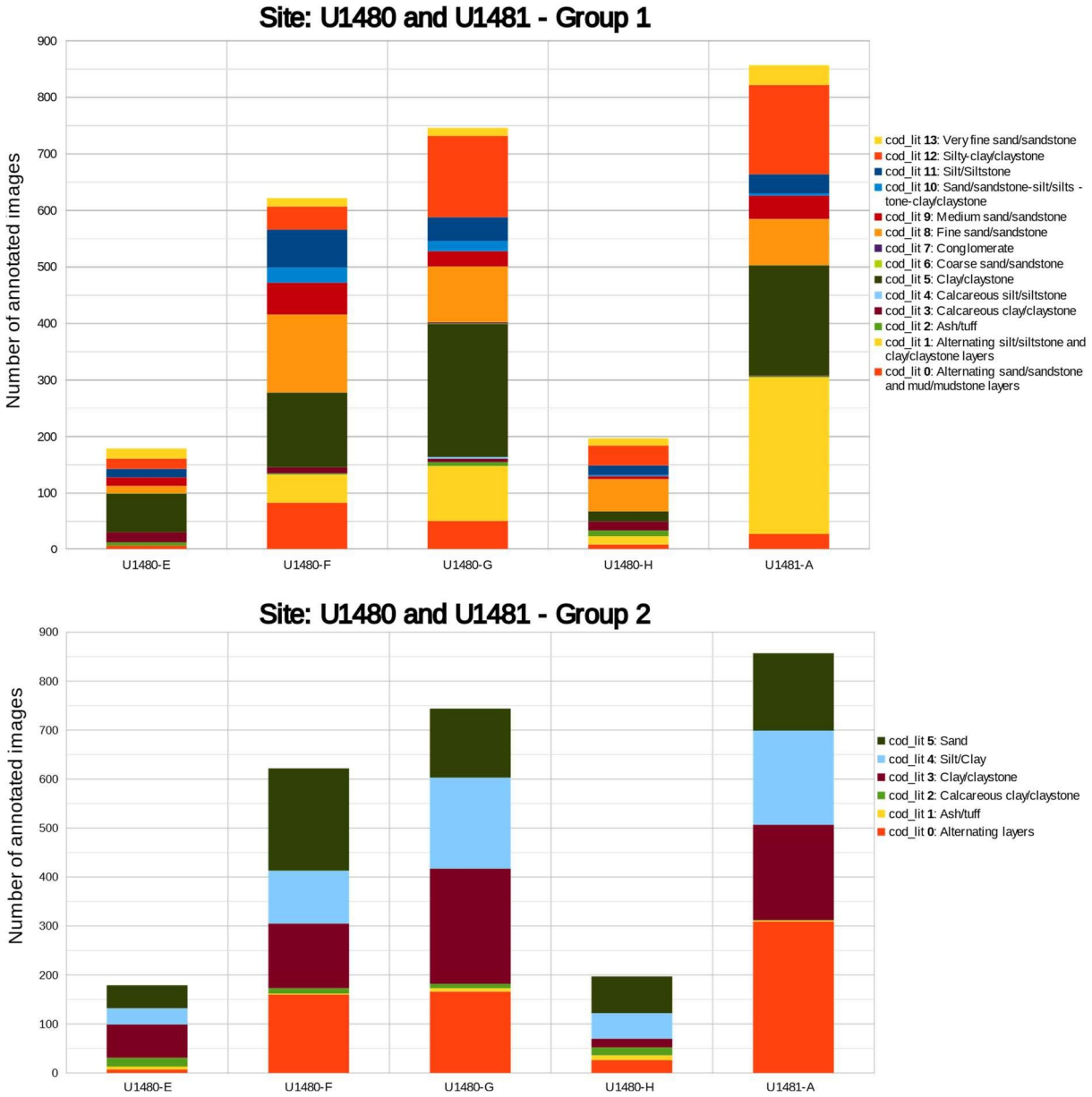
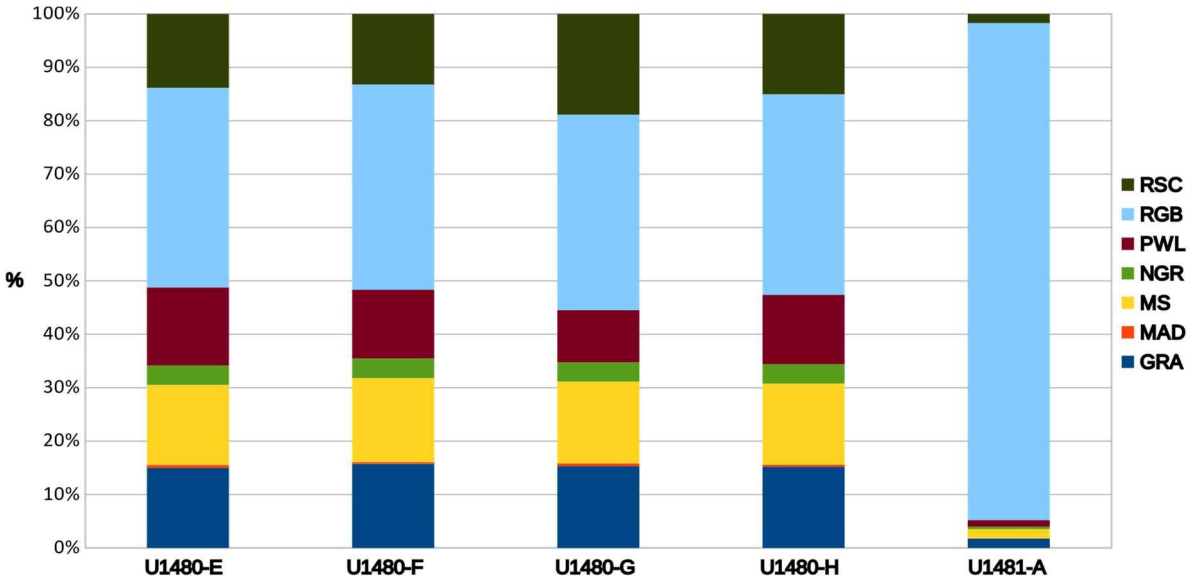


Figure 12 Division of the annotated images by site and hole according to the grouping performed in groups 1 and 2. cod_lit refers to the code that indicates the lithology in the group.

The total percentage of original records extracted from the annotated images are organized by site, hole, and petrophysical data, as shown in Fig. 13 and attached table, divided into a matrix [line x column] where the lines are the records and the columns are the features.



Site	GRA	MAD	MS	NGR	PWL	RGB	RSC
U1480-E	3333	125	3333	810	3253	8314	3081
U1480-F	8108	166	8108	1885	6636	19815	6806
U1480-G	8753	305	8753	2066	5578	20940	10787
U1480-H	4135	107	4135	994	3530	10230	4099
U1481-A	4679	172	4679	1123	3371	250162	4632

Figure 13 Visualization of the original records of the petrophysical data extracted from the annotated images divided by site and hole. The records are divided into the sites U1480 and U1481, holes E, F, G, H and A. The number of records is directly related to the reading resolution of each property according to the depth. Properties with few records indicate that the reading resolution is lower (10 cm), and many records indicate that the reading resolution is higher (1 cm).

The dataset0 includes original records of the petrophysical data and interpolated records of the U1480 site, Hole E, Group 1, as can be seen in Fig. 14. This dataset was used to validate the best-interpolated data using RF method among the eight defined interpolators. The result of the validation processing can be found in Fig. 15.

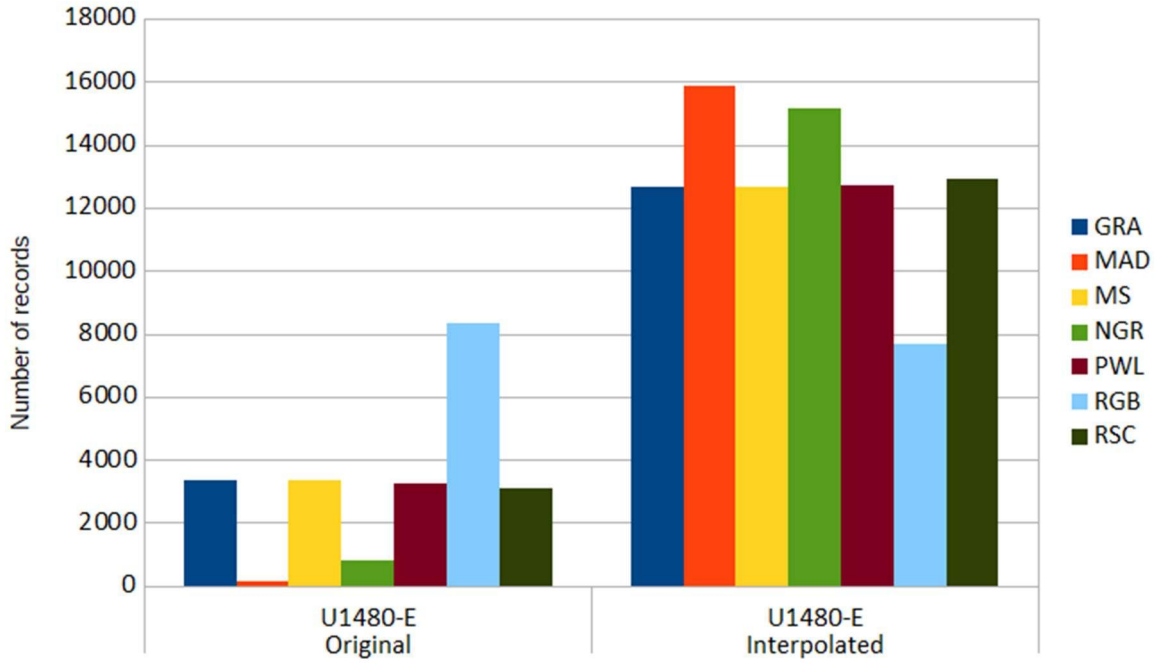
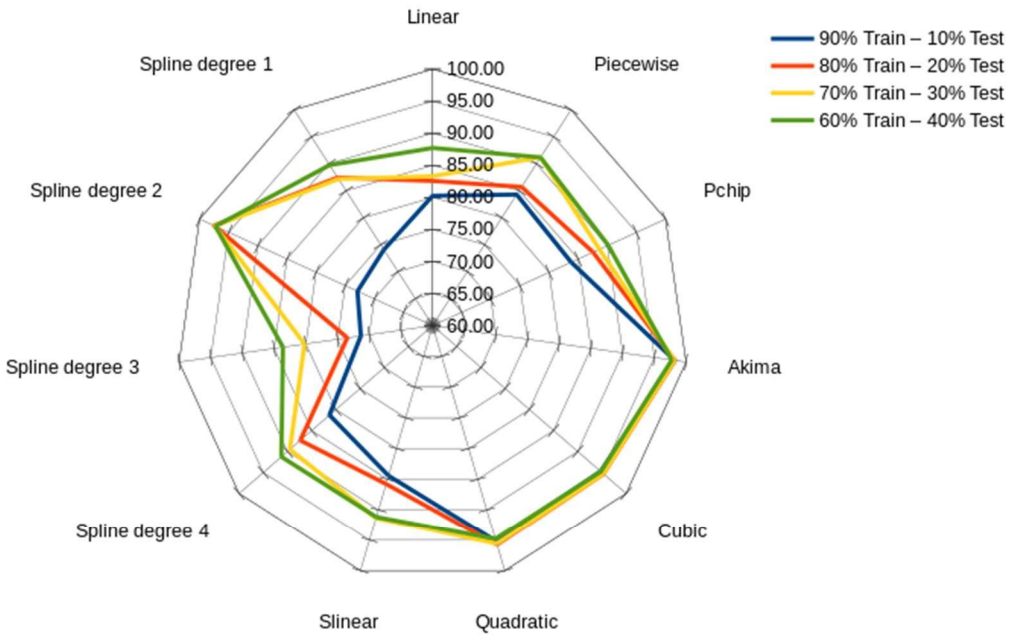


Figure 14 Division between original records and interpolated records for dataset0. The petrophysical data used for the dataset's formation are: GRA, MAD, MS, NGR, PWL, RGB, and RSC. The petrophysical data features with the most interpolated values are MAD and NGR since they have fewer original reading values. The RGB petrophysical data has a smaller number of interpolated values following the original data's reading characteristic with a shorter reading interval than the other properties. The petrophysical data GRA, MS, PWL, and RSC maintained a similar percentage in the number of interpolated values.

The best classification result is the Akima interpolator with accuracy equal to 98.22% in combination 80% train - 20% test. Cross-validation is a 5 k folds configuration with results for the best interpolator using the Accuracy, F1, Precision and Recall metrics. Cross-validation shows good results in the same combination with an accuracy greater than 98%. Observing the confusion matrix, in the four combinations, one can see a few records misclassified, such as in the cases of lithology 9 (Medium sand/sandstone) classified as lithology 8 (Fine sand/sandstone), lithology 11 (Silt/siltstone) classified as lithology 5 (Clay/claystone) and lithology 13 (Very fine sand/sandstone) classified as lithology 8 (Fine sand/sandstone). These inconsistencies could occur due to the great similarity between these lithologies, mainly in the very fine, fine, and medium subdivisions of sand/sandstone.

(A)



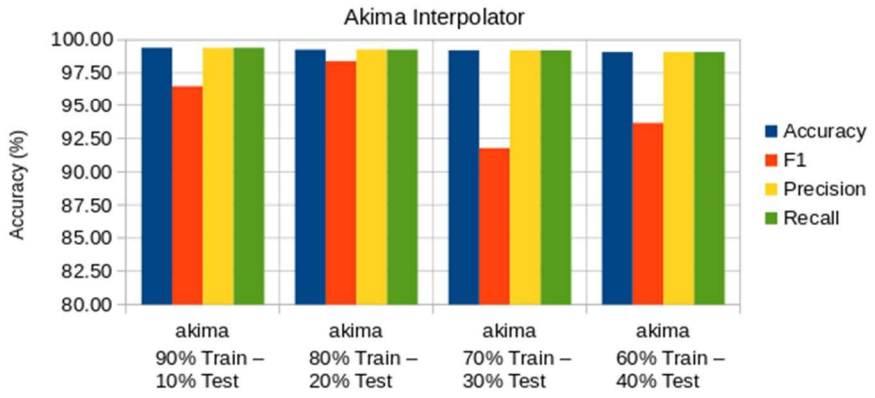
(B)

Accuracy between training and testing

Interpolator	90% Train – 10% Test	80% Train – 20% Test	70% Train – 30% Test	60% Train – 40% Test
Linear	80.24	82.55	83.29	87.70
Spline degree 1	74.08	87.44	87.18	89.75
Spline degree 2	72.90	97.54	97.42	97.30
Spline degree 3	71.32	73.46	80.20	83.61
Spline degree 4	81.31	87.35	89.50	91.26
Slinear	84.42	85.76	91.34	91.15
Quadratic	95.18	95.74	95.60	94.81
Cubic	94.74	95.31	95.29	94.79
Akima	98.19	98.22	98.16	97.75
Pchip	83.79	87.55	88.63	90.16
Piecewise	84.29	85.70	91.22	91.18

Cross-validation 5 k folds

(C)



289

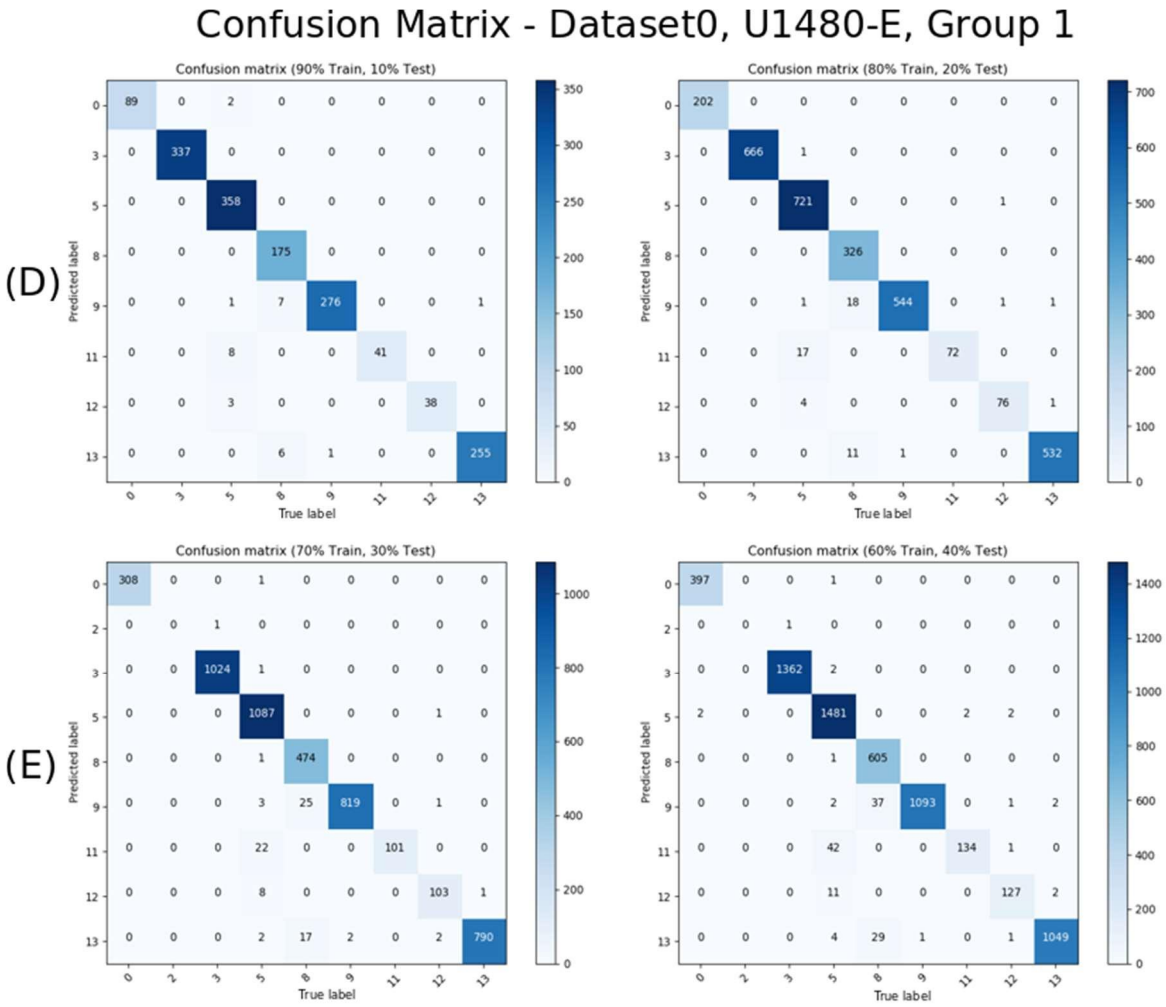
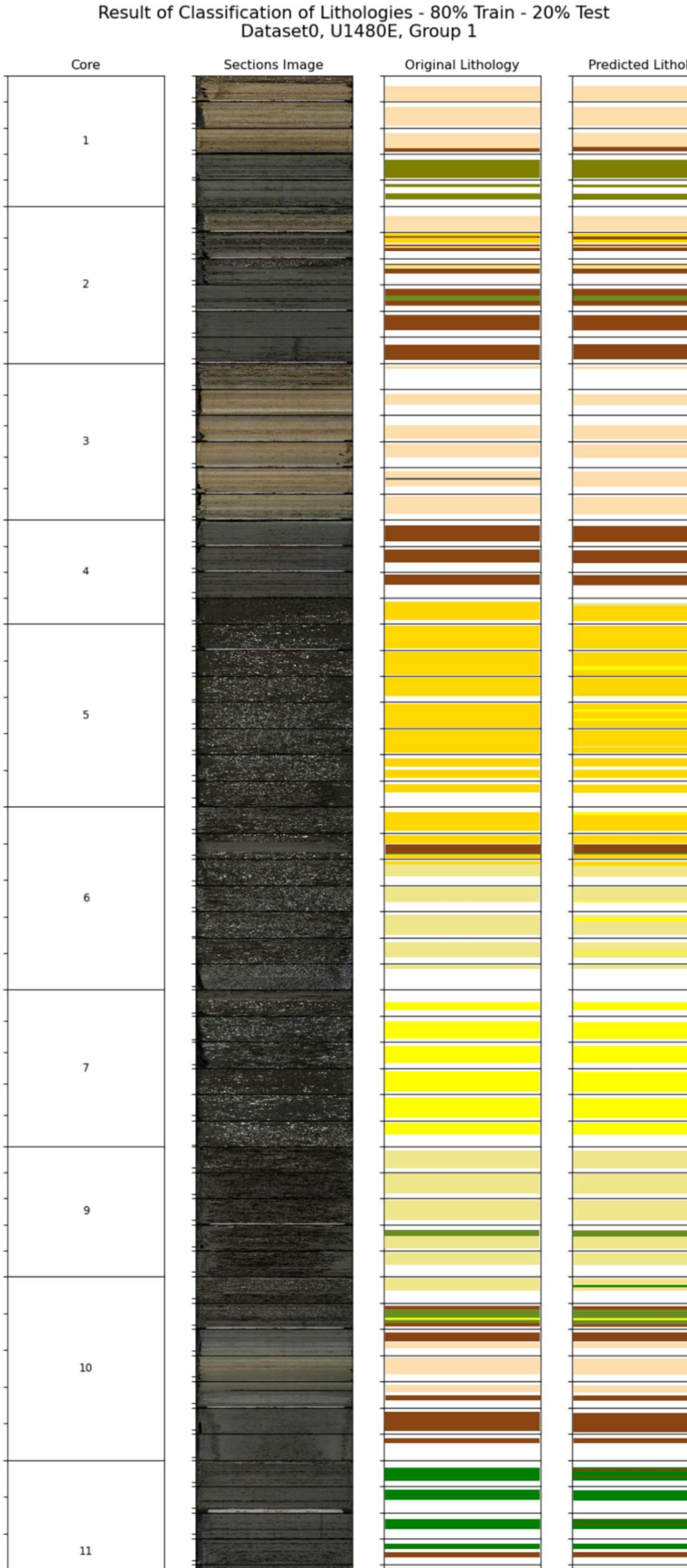


Figure 15 Result of the validation process of interpolators for dataset0. (A) accuracy curve according to the interpolator and the combination of training and testing. The Spline interpolator is subdivided according to grade: Spline degree 1, Spline degree 2, Spline degree 3, and Spline degree 4. (B) refers to the accuracy values in the four combinations of training and testing. (C) the result of the application of Cross-validation 5 k folds is observed for the Akima interpolator. (D) and (E) is displayed result of the confusion matrix for the four combinations of training and testing. The confusion matrix is divided according to the combination of training and testing between Predicted label and True label for each lithology in Group 1. cod_lit refers to the code that indicates the lithology in the group. For Group 1, cod_lit 0: Alternating sand/sandstone and mud/mudstone layers, cod_lit 2: Ash/tuff, cod_lit 3: Calcareous clay/claystone, cod_lit 5: Clay/claystone, cod_lit 8: Fine sand/sandstone, cod_lit 9: Medium sand/sandstone, cod_lit 11: Silt/Siltstone, cod_lit 12: Silty clay/claystone, cod_lit 13: Very fine sand/sandstone.

The lithological organization through the presentation in a log structure with columns, as shown in Fig. 16, describes the best result in the combination of training and testing according to the interpolation interval defined in dataset0, Group 1, in combination 80% train - 20% test, Akima interpolator.


Result of Classification of Lithologies - 80% Train - 20% Test
Dataset0, U1480E, Group 1
Core
Sections Image
Original Lithology
Predicted Litho
1
2
3
4
5
6
7
9
10
11

in the image format described in the expedition (trim, eliminated the depth- ruler over the image). The Original Li
(original) extracted from IODP-Expedition 362, site U1480-E. The Predicted Lithology column describes the litholog
execution of the classification with its specific settings. Records in white (empty) refers to intervals without original reco
records, according to the characteristic of the interpolator used.

317 The formation of dataset1 interpolated records using the best-evaluated interpolator,
318 in this case, Akima, from site U1480-E, F, G, H, and site U1481-A. The dataset1, site
319 U1480, Groups 1 and 2, has 84,009 records divided according to lithologies. For
320 dataset1, site U1481, Groups 1 and 2, has a total of 211,936 records divided according
321 to the lithologies. The SM4_results, Fig. 43, summarizes the formation of dataset with
322 original records and interpolated records.

323 The dataset2 includes records extracted from the processing of the annotated images
324 forming a dataset with 67,944 lines for site U1480 and 17,064 lines for site U1481 with
325 90 features of characteristics stored by processing, divided into Groups 1 and 2. The
326 formation of dataset2, U1480, Groups 1 and 2, includes records of 945 images
327 annotated used with 67,994 segmented regions in the collection of characteristics
328 according to the definitions of texture and color via SLIC Superpixel. For dataset2,
329 U1481, Groups 1 and 2 includes records of 647 images annotated with 17064
330 segmented regions in the collection of characteristics. The SM4_results, Fig. 44,
331 summarizes the dataset configuration.

332 The processing and classification results for dataset1 can be seen in Fig. 17 and in the
333 supplementary material, altogether with tables and images from the other analyses of
334 the lithological classification.

335

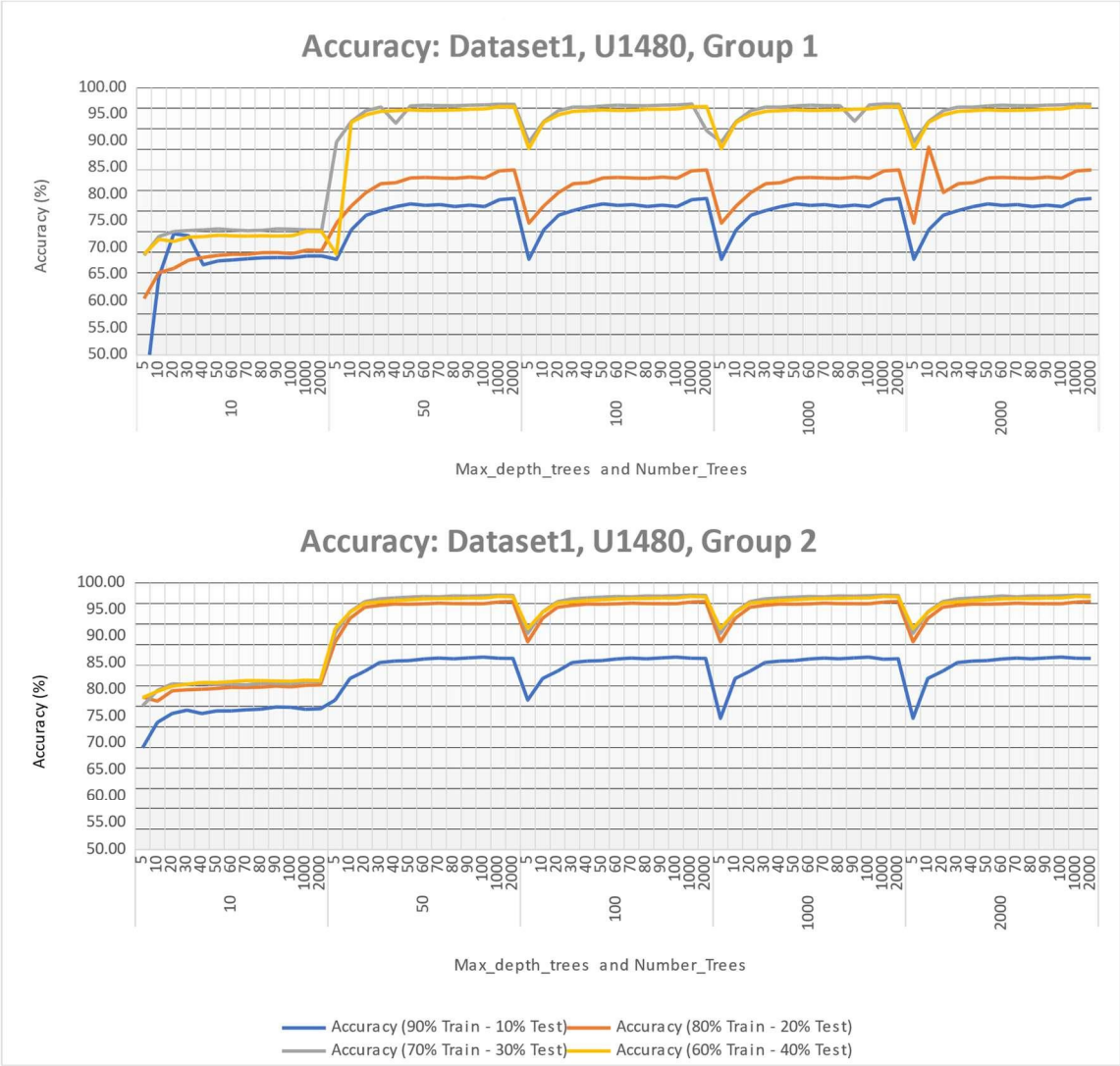


Figure 17 Accuracy result for dataset1, U1480, Groups 1 and 2, divided between a combination of training and testing. The accuracy curve comprises the intervals between the maximum depth of trees and the number of trees in training and test combinations: 90% -10%, 80% -20%, 70% -30%, and 60% -40%. Peaks are observed in the accuracy records in a tree number equal to 5.

The best result of accuracy in Group 1 is the combination of training and testing: 70% train - 30% test with the number of trees equal to 1000 and maximum tree depth equal to 100, 1000 and 2000 with the value of 96.96%, and the best result of accuracy in Group 2 is in the combination of training and testing: 70% train - 30% test with the number of trees equal to 1000 and the maximum depth of trees equal to 50, 100, 1000 and 2000 with a value of 97.71%, as seen in the summary table below.

Table 9 Result of accuracy by group and combination of training and testing for dataset1, U1480, Groups 1 and 2.

Group	Accuracy (90% Train - 10% Test)	Accuracy (80% Train - 20% Test)	Accuracy (70% Train - 30% Test)	Accuracy (60% Train - 40% Test)
1	Max_depth_trees: 50, Number_Trees: 2000 Max_depth_trees: 1000, Number_Trees: 2000	Max_depth_trees: 2000, Number_Trees: 10	Max_depth_trees: 100, Number_Trees: 1000 Max_depth_trees: 1000, Number_Trees: 1000 Max_depth_trees: 2000, Number_Trees: 1000	Max_depth_trees: 1000, Number_Trees: 2000
2	Max_depth_trees: 50, Number_Trees: 100 Max_depth_trees: 100, Number_Trees: 100 Max_depth_trees: 1000, Number_Trees: 100	Max_depth_trees: 50, Number_Trees: 2000 Max_depth_trees: 100, Number_Trees: 2000 Max_depth_trees: 1000, Number_Trees: 2000 Max_depth_trees: 2000, Number_Trees: 2000	Max_depth_trees: 50, Number_Trees: 1000 Max_depth_trees: 100, Number_Trees: 1000 Max_depth_trees: 1000, Number_Trees: 1000 Max_depth_trees: 2000, Number_Trees: 1000	Max_depth_trees: 50, Number_Trees: 1000 Max_depth_trees: 100, Number_Trees: 1000 Max_depth_trees: 1000, Number_Trees: 1000 Max_depth_trees: 2000, Number_Trees: 1000

Group	Accuracy (90% Train - 10% Test)	Accuracy (80% Train - 20% Test)	Accuracy (70% Train - 30% Test)	Accuracy (60% Train - 40% Test)
1	79,28	88,87	96,96	96,49
2	86,13	96,53	97,71	97,46

The results of applying the confusion matrix and ROC curve (Fig. 18) for dataset1, U1480, Groups 1 and 2 obtained excellent results following the values calculated by the accuracy.

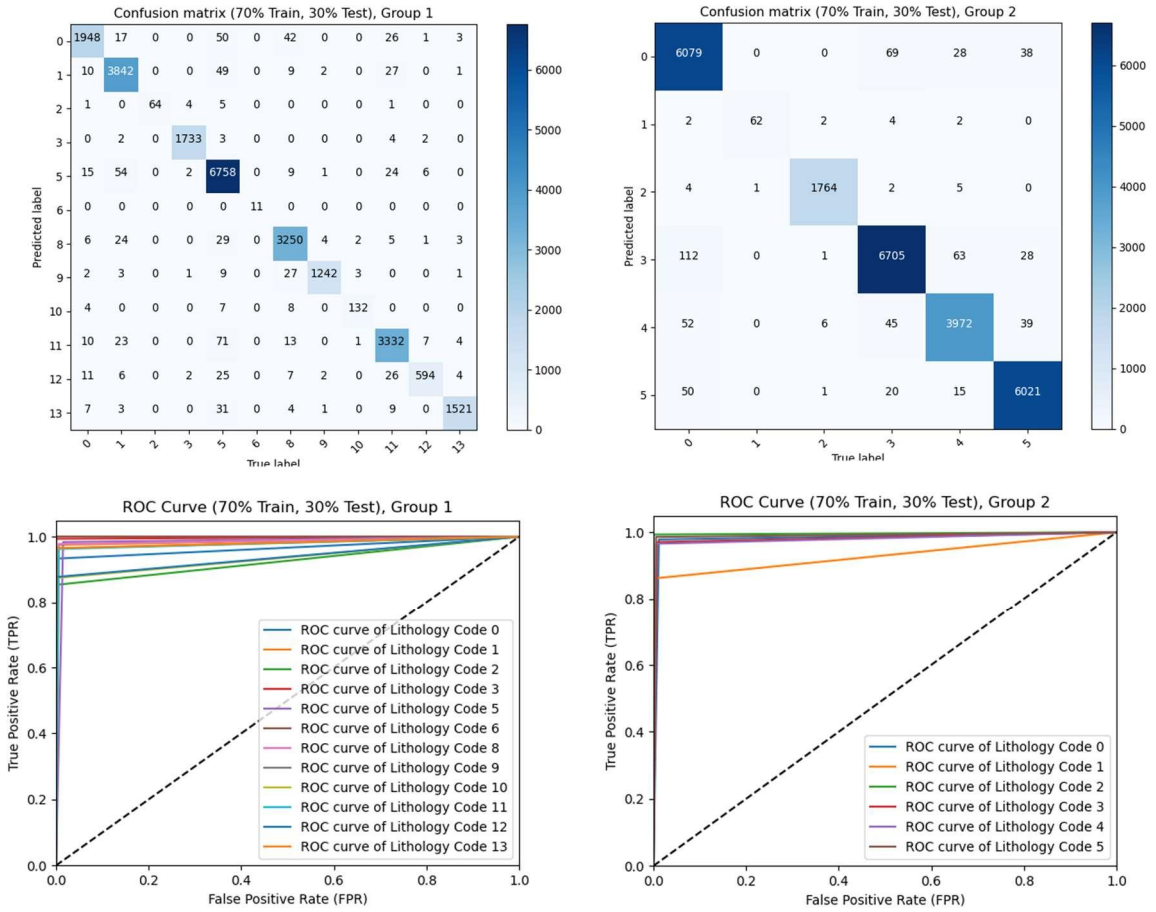


Figure 18 Result of the confusion matrix and ROC curve for dataset1, U1480, Groups 1 and 2. The confusion matrix is divided according to the combination of training and testing between Predicted label and True label for each lithology in Groups 1 and 2. The ROC Curve covers the area calculated between TPR and FPR normalized to the range 0 to 1. cod_lit refers to the code that indicates the lithology in the group. For Group 1, cod_lit 0: Alternating sand/sandstone and mud/mudstone layers, cod_lit 1: Alternating silt/siltstone and clay/claystone layers. cod_lit 2: Ash/tuff, cod_lit 3: Calcareous clay/claystone, cod_lit 5: Clay/claystone, cod_lit 6: Coarse sand/sandstone, cod_lit 8: Fine sand/sandstone, cod_lit 9: Medium sand/sandstone, cod_lit 10: Sand/sandstone-silt/siltstone-clay/claystone, cod_lit 11: Silt/Siltstone, cod_lit 12: Silty clay/claystone, cod_lit 13: Very fine sand/sandstone. For Group 2, cod_lit 0: Alternating layers, cod_lit 1: Ash/tuff, cod_lit 2: Calcareous clay/claystone, cod_lit 3: Clay/claystone, cod_lit 4: Silt/Clay, cod_lit 5: Sand.

For dataset1, U1481, Group 1, the best accuracy values are in the combination of training and testing: 80% train - 20% test and 70% train - 30% test, with a maximum number of trees in 15 and number of trees with values of 90, 100, 1000 and 2000, accuracy greater than 99.60% (Fig. 19). For dataset1, U1481, Group 2, the best accuracy is in the combination of training and testing: 80% train - 20% test and 70% train - 30% test, with maximum tree depth distribution at 15 and number of trees with values of 20, 50, 60, 80 and 100, accuracy greater than 99.70%.

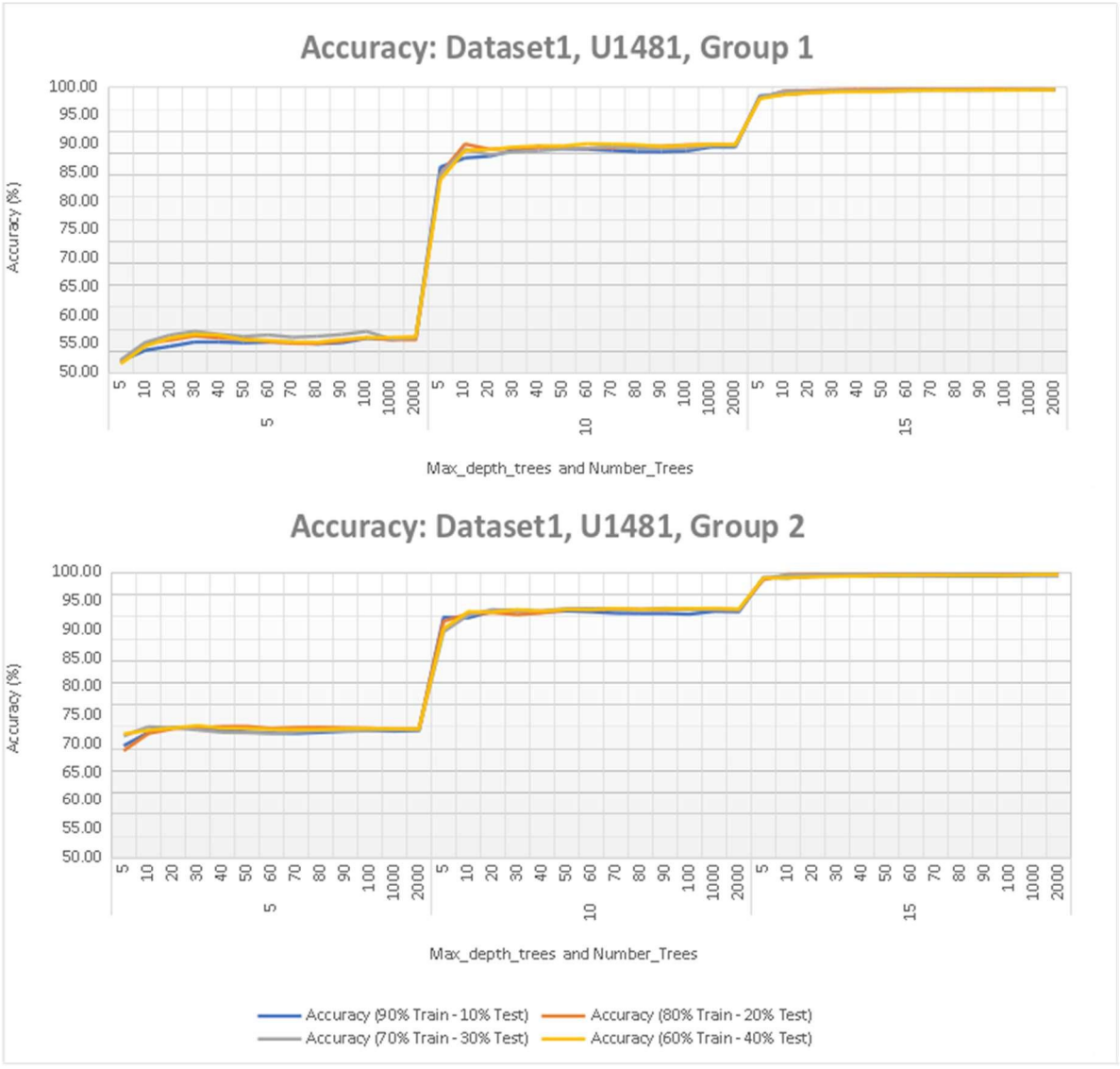


Figure 19 Accuracy result for dataset1, U1481, Groups 1 and 2, divided between combination of training and testing. The accuracy curve comprises the intervals between maximum depth of trees and number of trees in training and test combinations: 90% -10%, 80% -20%, 70% -30% and 60% -40%.

The best result of accuracy in Group 1 is the combination of training and testing: 80% train - 20% test with the number of trees equal to 90, 100, and 1000 and the maximum depth of trees equal to 15 with a value of 99.68%, and the best result of accuracy in Group 2 is in the combination of training and testing: 80% train - 20% test with the number of trees equal to 20, 50 and 100 and maximum tree depth equal to 15 with a value of 99.74%, as is displayed in the summary tables below.

Table 10 Result of accuracy by group and combination of training and testing for dataset1, U1481, Groups 1 and 2.

Group	Accuracy (90% Train - 10% Test)	Accuracy (80% Train - 20% Test)	Accuracy (70% Train - 30% Test)	Accuracy (60% Train - 40% Test)
1	Max_depth_trees: 15, Number_Trees: 1000	Max_depth_trees: 15, Number_Trees: 90 Max_depth_trees: 15, Number_Trees: 100 Max_depth_trees: 15, Number_Trees: 1000	Max_depth_trees: 15, Number_Trees: 2000	Max_depth_trees: 15, Number_Trees: 2000
2	Max_depth_trees: 15, Number_Trees: 1000 Max_depth_trees: 15, Number_Trees: 2000	Max_depth_trees: 15, Number_Trees: 20 Max_depth_trees: 15, Number_Trees: 50 Max_depth_trees: 15, Number_Trees: 100	Max_depth_trees: 15, Number_Trees: 60 Max_depth_trees: 15, Number_Trees: 80	Max_depth_trees: 15, Number_Trees: 1000 Max_depth_trees: 15, Number_Trees: 2000

Group	Accuracy (90% Train - 10% Test)	Accuracy (80% Train - 20% Test)	Accuracy (70% Train - 30% Test)	Accuracy (60% Train - 40% Test)
1	99,52	99,68	99,62	99,55
2	99,53	99,74	99,73	99,66

The results of applying the confusion matrix and ROC curve (Fig. 20) for dataset1, U1481, Groups 1 and 2 obtained excellent results in the calculation of accuracy with very few records classified incorrectly mainly in lithology 0 and 1 (Group 1) and lithology 0 (Group 2) referring to lithologies with intercalation of layers between silt and clay.

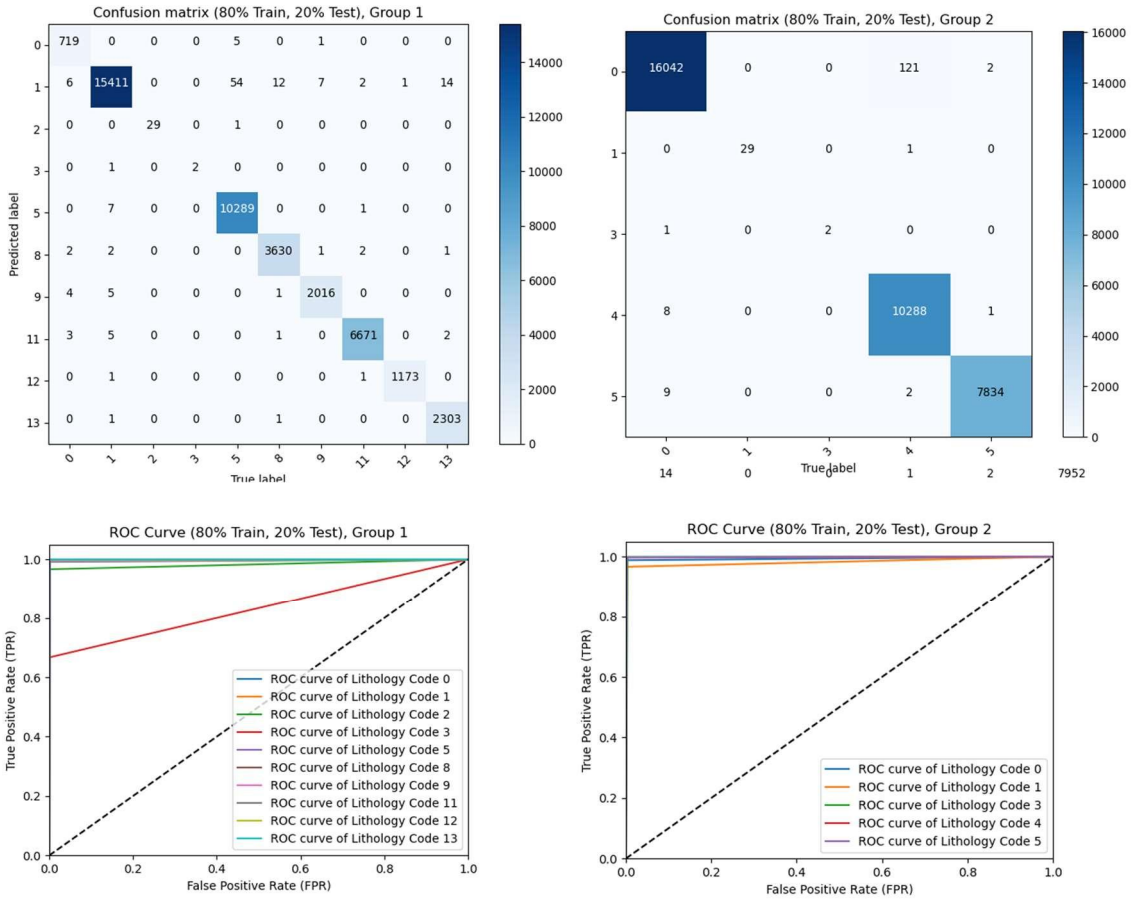


Figure 20 Result of the confusion matrix and ROC curve for dataset1, U1481, Groups 1 and 2, in the combination of training and testing: 80% train-20% test. The confusion matrix is divided according to training and testing between the Predicted label and True label for each lithology in Group 1 and Group 2. The ROC Curve covers the area calculated between TPR and FPR normalized to the range 0 to 1. cod_lit refers to the code that indicates the lithology in the group. For Group 1, cod_lit 0: Alternating sand/sandstone and mud/mudstone layers, cod_lit 1: Alternating silt/siltstone and clay/claystone layers. cod_lit 2: Ash/tuff, cod_lit 3: Calcareous clay/claystone, cod_lit 5: Clay/claystone, cod_lit 8: Fine sand/sandstone, cod_lit 9: Medium sand/sandstone, cod_lit 11: Silt/Siltstone, cod_lit 12: Silty clay/claystone, cod_lit 13: Very fine sand/sandstone. For Group 2, cod_lit 0: Alternating layers, cod_lit 1: Ash/tuff, cod_lit 2: Calcareous clay/claystone, cod_lit 3: Clay/claystone, cod_lit 4: Silt/Clay, cod_lit 5: Sand. Lithologies for cod_lit 2 and 3 (Group 1) and cod_lit 1 and 2 (Group 2) have few records according to dataset and combination of training and testing.

The processing result for dataset2 can be found in Fig. 21, 22, and in the supplementary material. The Apx and the total number of segmented regions for each lithology group and the dataset is shown in Fig. 21.

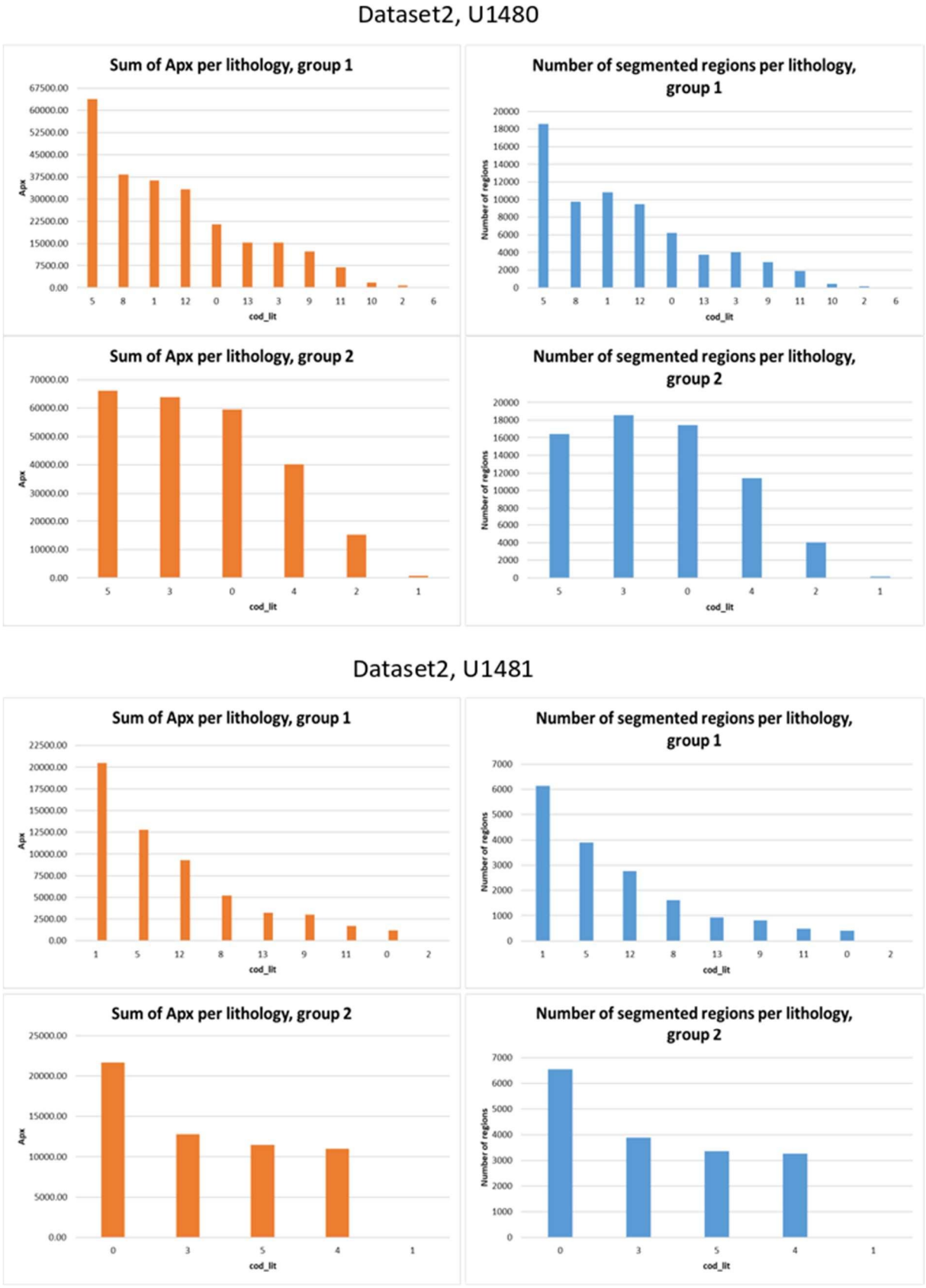
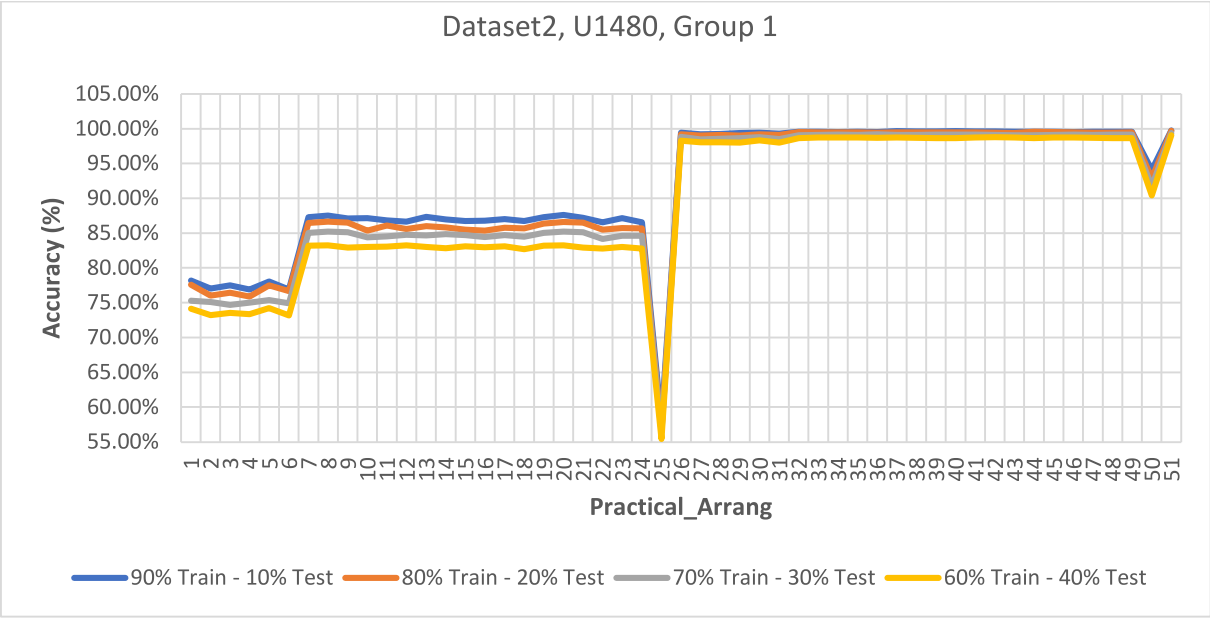
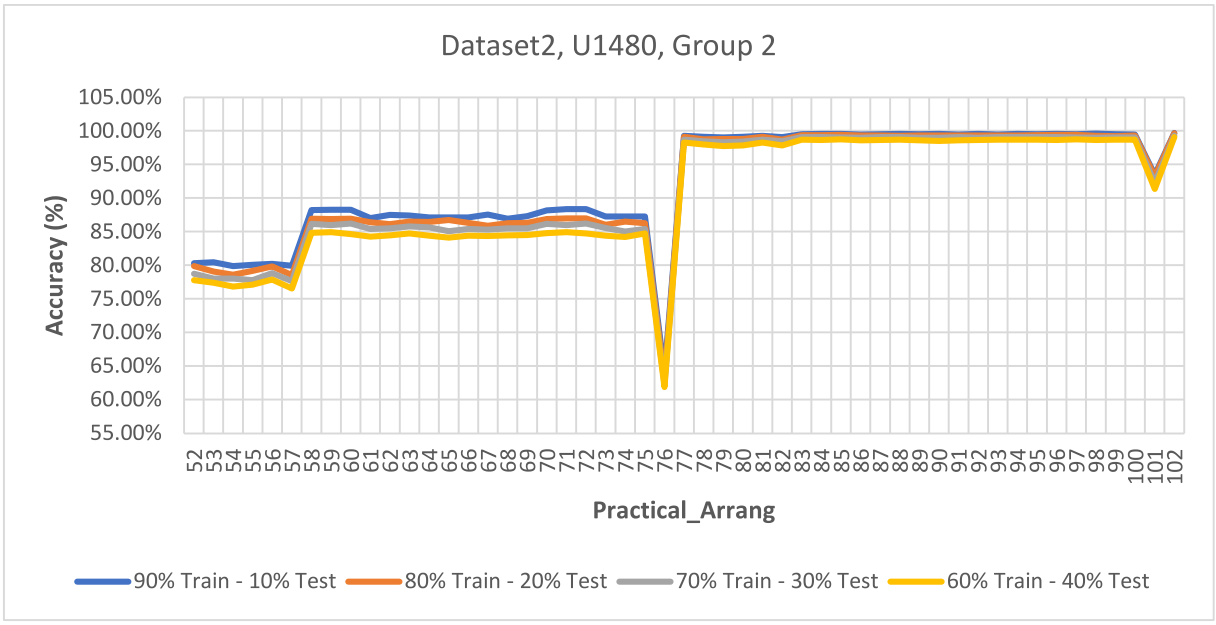


Figure 21 Sum of Apx and number of regions per lithology for dataset2, U1480, Groups 1 and 2, and for dataset2, U1481, Groups 1 and 2. Apx is measured in cm². cod_lit refers to the code that indicates the lithology in the group. For Group 1, cod_lit 0: Alternating sand/sandstone and mud/mudstone layers, cod_lit 1: Alternating silt/siltstone and clay/claystone layers. cod_lit 2: Ash/tuff, cod_lit 3: Calcareous

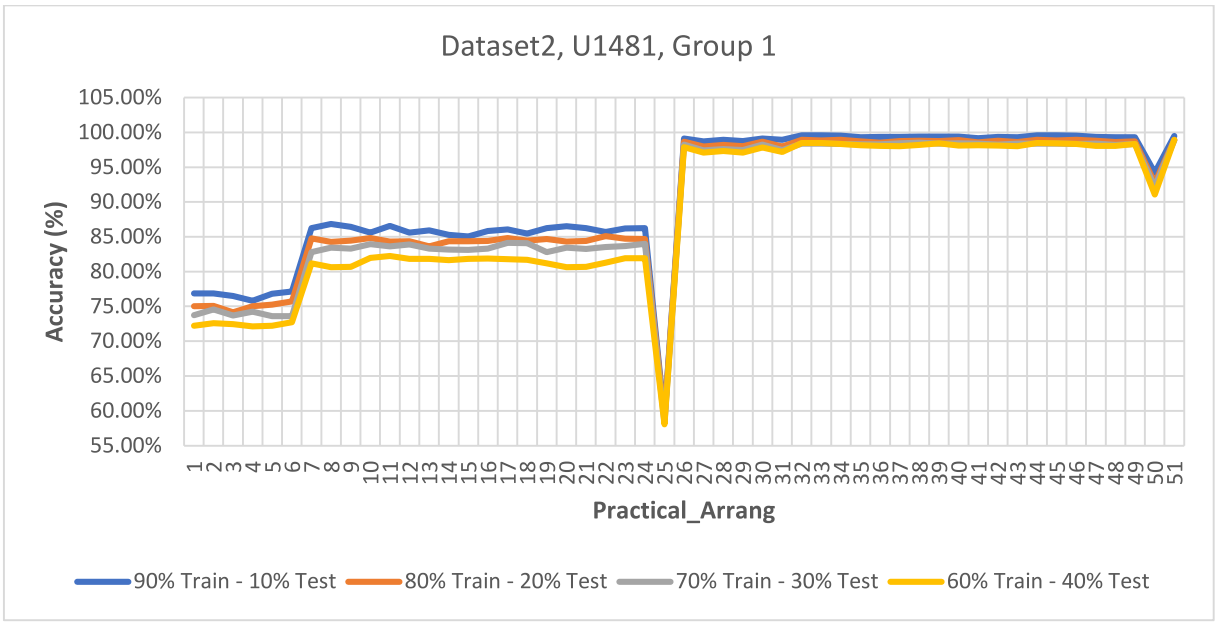
clay/claystone, cod_lit 5: Clay/claystone, cod_lit 6: Coarse sand/sandstone, cod_lit 8: Fine sand/sandstone, cod_lit 9: Medium sand/sandstone, cod_lit 10: Sand/sandstone-silt/siltstone-clay/claystone, cod_lit 11: Silt/Siltstone, cod_lit 12: Silty clay/claystone, cod_lit 13: Very fine sand/sandstone. For Group 2, cod_lit 0: Alternating layers, cod_lit 1: Ash/tuff, cod_lit 2: Calcareous clay/claystone, cod_lit 3: Clay/claystone, cod_lit 4: Silt/Clay, cod_lit 5: Sand. For dataset2, U1480, Group 1, the lithologies cod_lit 5 (Clay/claystone) and cod_lit 8 (Fine sand/sandstone) have the largest area extracted and processed during drilling, representing 41.60% of the total area. For dataset2, U1480, Group 2, by organizing the division of lithologies into new groups, the distribution of the most homogeneous area among the four main lithologies cod_lit 5 (Sand), cod_lit 0 (Alternating layers), cod_lit 3 (Clay/Claystone) and cod_lit 4 (Silt/Clay). For dataset2, U1481, Group 1, the lithologies cod_lit 1 (Alternating silt/siltstone and clay/claystone layers) and cod_lit 5 (Clay/claystone) representing 58.35% of the total area.

In Fig. 14 is shown the analysis of the best accuracy between the training and test combinations with the Practical_Arrang defined for dataset2, U1480 and dataset2, U1481, Groups 1 and 2.




Dataset2, U1480, Group 2
105.00%
100.00%
95.00%
90.00%
85.00%
80.00%
75.00%
70.00%
65.00%
60.00%
55.00%
Accuracy (%)
52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102
Practical_Arrang
90% Train - 10% Test
80% Train - 20% Test
70% Train - 30% Test
60% Train - 40% Test

434


Dataset2, U1481, Group 1
105.00%
100.00%
95.00%
90.00%
85.00%
80.00%
75.00%
70.00%
65.00%
60.00%
55.00%
Accuracy (%)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51
Practical_Arrang
90% Train - 10% Test
80% Train - 20% Test
70% Train - 30% Test
60% Train - 40% Test

435

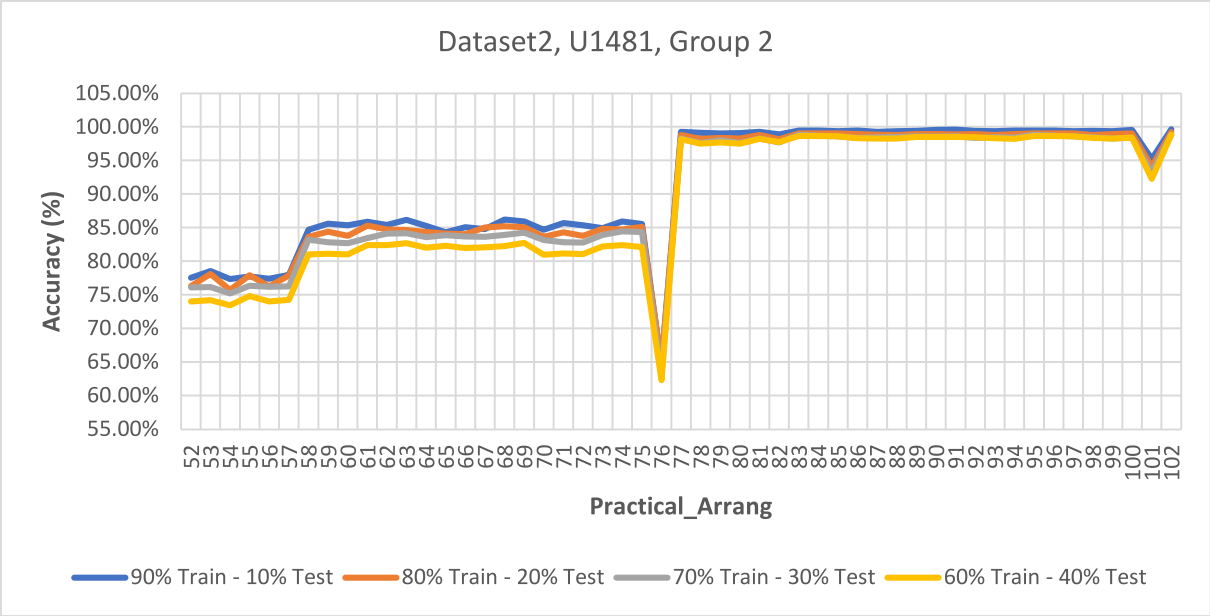


Figure 22 Accuracy curve for dataset2, U1480 and dataset2, U1481 with Practical_Arrang defined by a combination of training and testing for Groups 1 and 2. The accuracy curve comprises the ranges of the Practical_Arrang in the training and test combinations: 90% -10%, 80% -20%, 70% -30% and 60% -40%. For Group 1, the range of Practical_Arrang is 1 to 51. For Group 2, the range of Practical_Arrang is 52 to 102. Each Practical_Arrang covers arrangements between the features in dataset2. The combinations of the Practical_Arrang can be found in the supplementary material, file: Computer Code Availability/Python code/practical_arrangement_to-dataset2.

The final classification of the segmented regions of the images for dataset2, U1480 and U1481, in Groups 1 and 2, combination training and testing, Practical_Arrang, number of trees, and the maximum depth of trees are shown in the table 11. The confusion matrix and ROC curve results for the settings of number of trees, maximum depth of trees and practical arrangements are shown in the supplementary material, directory results.

Table 11 Result of the best accuracy by group, combination of training and testing, Practical_Arrang, number of trees and maximum depth of trees for dataset2, U1480 and U1481, Groups 1 and 2. Max_depth refers to the maximum depth of trees.

Site	Group	Combination	Practical_Arrang	Max_depth	Number of Trees	Accuracy (%)
U1480	1	90% Train - 10% Test	51	50	100	99.75
		80% Train - 20% Test	51	50	100	99.71
		70% Train - 30% Test	51	50	70	99.40
		60% Train - 40% Test	51	50	90	99.08
	2	90% Train - 10% Test	102	50	70	99.72

		80% Train - 20% Test	102	50	90	99.66
		70% Train - 30% Test	102	50	100	99.32
		60% Train - 40% Test	102	50	90	99.07
U1481	1	90% Train - 10% Test	32 33 44 45	50	60 50 60 50	99.59
		80% Train - 20% Test	51	50	100	98.95
		70% Train - 30% Test	51	50	80	98.89
		60% Train - 40% Test	51	50	90	98.95
	2	90% Train - 10% Test	102	50	50	99.65
		80% Train - 20% Test	102	50	50	99.24
		70% Train - 30% Test	102	50	80	98.93
		60% Train - 40% Test	102	50	100	98.87

In table 12 is observed the Apx used in the classification for each combination of training and testing related to sites U1480 and U1481 and to Groups 1 and 2. The SM4_results, Fig. 45, summarizes the segmented regions identification with an example for Group 1.

Table 12 Result of the classification of segmented regions for dataset2, U1480 and dataset2, U1481, Groups 1 and 2, with the division between training and testing combination, Practical_Arrang, Apx to training, Apx to test correct, and Apx to test incorrectly. A

Site	Group	Combination	Practical_Arrang	Apx to training	Apx to test correct	Apx to test incorrect
U1480	1	90% Train - 10% Test	51	242963.87	2553.87	4.45
		80% Train - 20% Test	51	240524.48	4984.55	13.16
		70% Train - 30% Test	51	237947.75	7506.12	68.32
		60% Train - 40% Test	51	235566.15	9889.83	102.75
	2	90% Train - 10% Test	102	242963.87	2551.45	6.87
		80% Train - 20% Test	102	240524.48	4984.42	13.29
		70% Train - 30% Test	102	237947.75	7517.49	56.95
		60% Train - 40% Test	102	235566.15	9883.05	109.53
U1481	1	90% Train - 10% Test	32	51475.40	5443.95	17.95
			33	51462.76	5455.17	19.37
			44	51654.45	5264.90	17.95
			45	51618.05	5299.88	19.37
		80% Train - 20% Test	51	45728.05	11095.56	113.69
		70% Train - 30% Test	51	40166.30	16622.16	175.92
		60% Train - 40% Test	51	34509.74	22226.44	228.20
	2	90% Train - 10% Test	102	51326.95	5587.56	22.79
		80% Train - 20% Test	102	45728.05	11132.97	76.28
		70% Train - 30% Test	102	40166.30	16626.29	171.79
		60% Train - 40% Test	102	34509.74	22204.86	249.78

4. Discussions

The number of images annotated by a group is better distributed in Group 2, with a more uniform amount among the categories used. In Group 1, the number of images per lithology is less uniform, concentrating more in the lithology coded as 5, 8, 12 and 1. In Group 1, lithologies' variation is more significant, with many possible divisions between sand/sandstone, clay/claystone and layer lithologies. For Group 2, lithologies' grouping in larger combinations resulted in a uniform distribution, improving training, and improving the classification results. Group 1 includes a practical model of distribution of lithologies aimed at application in the field (during the lithological data acquisition by the Geologist) at the time of lithological classification. The dataset0, Group 1, has fewer records because it contains values only from the site U1480, hole E, used to validate the best interpolator. It encompasses a range of records in all lithologies, compared to other datasets, creating an excellent collection of appropriate values for choosing the appropriate interpolator. Dataset1, Groups 1 and 2, have more records interpolated by the Akima interpolator, resulting in expanding lithology combinations. Dataset2 has smaller number of lithologies, since it analyzes the segmented areas in the images, creating a lithology record for each segmentation. Dataset1, U1481, Groups 1 and 2, integrates a comprehensive lithological dataset with many interpolated values respecting the reading context and depth range with the original records.

The Akima interpolator was used with the petrophysical data, generating a broader dataset. These data were used in a combination of 80% train and 20% test to create classification models with the RF method. The result obtained presented an accuracy of 98.22% for lithological classification. Analyzing Fig. 8, with the combination of the best result of the interpolation, original characteristics of the lithologies in the images

are all preserved in Predicted Lithology with emphasis on the lithology ranges, Group 1, Calcareous clay/claystone, Fine Sand/Sandstone and Alternating sand/sandstone and mud/mudstone layers. Other interpolators obtain good results like Spline grade 2, Quadratic, and Cubic since they use the polynomial function to calculate the interpolated point's value. The highlight of Akima is related to the fact that it uses six reference points to create the new interpolated point, not acting in smoothing the curve, reducing the oscillation that Spline and Cubic usually produce.

For dataset1, U1480, Group 1, the accuracy metric presented more variations in smaller amounts of several trees and smaller amounts of the maximum depth of trees. The best accuracy is in the combination of training and testing: 70% train - 30% test, with value equal to 96.96%. Stabilization or few learning variations are observed in the number of trees above 50 and the maximum number of trees greater than 1000. In the application of the ROC curve for this dataset, it can be seen that the lithologies cod_lit 3 (Calcareous clay/claystone), cod_lit 5 (Clay/claystone) and cod_lit 8 (Fine sand/sandstone) are the best classified. They are lithologies with well-defined aspects between the interpolated petrophysical data values concerning the other lithologies.

For dataset1, U1480, Group 2, the accuracy values stabilize with few variations from the maximum number of trees greater than 50. Accuracy variations are observed in smaller values of the number of trees in all analyzed combinations. The best accuracy results are in the combinations of training and testing: 70% train - 30% test, which presented the same value equal to 97.71%. The confusion matrix and ROC curve results are excellent, better distributed among the 6 lithologies with few records classified incorrectly. The lithology cod_lit 2 (Calcareous clay/claystone) (Group 1, cod_lit: 3), cod_lit 3 (Clay/claystone) (Group 1, cod_lit: 5) and cod_lit 5 (sand) (Group1, cod_lit: 6, 8, 9, 13) are the best lithologies classified in Group 2.

514 For dataset1, U1481, Group 1, were obtained excellent results in the calculation of
515 accuracy (99.68%) with very few records classified incorrectly. These were mainly
516 records belonging to lithology cod_lit 1 (Alternating silt/siltstone and clay/claystone
517 layers). It is worth mentioning that this lithology integrates a context of alternating
518 layers between silt and clay with very close geophysical and image characteristics,
519 making classification and recognition difficult. Another point to highlight, according to
520 MCNEILL et al., 2017, the annotation of the type of lithology is performed by the visual
521 core description manually by the sedimentologist and errors or mistakes may occur in
522 the specific description in the rock at the time of drilling. The application of the ROC
523 curve shows an excellent response in analyzing the relationship between the predicted
524 values. An exception is seen in the lithology cod_lit 3 (Calcareous clay/claystone)
525 because the lithology has few records on the analyzed site (U1481).

526 For dataset1, U1481, Group 2, by the lithological organization of site U1480, lithologies'
527 distribution is more significant in lithologies cod_lit 0, cod_lit 4 and cod_lit 5, accuracy
528 is equal to 99.74%. Few records were classified incorrectly, being located in the
529 lithology cod_lit 0 (Group 1, cod_lit: 0, 1, 10), representing alternating layers between
530 sand, silt and clay. For the ROC curve, the classification results' analysis is excellent,
531 with all lithologies having great returns for the prediction of values.

532 Dataset2 comprises the texture and color characteristics of the images extracted from
533 the lithologies and the petrophysical data linked to dataset1, enabling 102
534 combinations of Practical_Arrang. For the first six Practical_Arrang, the accuracy had
535 a smaller variation between 70.00% and 80.00%, since, in these practical
536 arrangements, we used individually each feature of the texture properties (six
537 properties) with the values of the angles, the distance of pixels, an average of colors
538 of the pixels and the three embracing features originated from dataset1.

For the Practical_Arrang range (8 to 49 and 59 to 10) excellent results were obtained with accuracy equal to 99.63% for dataset2_U1480 and equal to 99.59% for the dataset2_U1481, with few variations between the respective range of analyzed arrangements. The Practical_Arrang (25 and 50 and 76 and 101) did not obtain good results in the accuracy because they use all the features of dataset2 and three features embracing originated from dataset1.

Finally, the Practical_Arrang (51 and 102) with the best result accuracy equal to 99.75% for dataset2, U1480 and the Practical_Arrang (32 and 102) with the best result obtained accuracy greater than 99.65% for dataset2, U1481.

The best values of accuracy between dataset1 and dataset2 are very close, given that the practical arrangement in dataset2 with the best result (51 and 102) has as input all the features of dataset1 plus the average RGB color intensity. As for accuracy only with features of dataset2 (referring to texture properties and average RGB color intensity) the accuracy results are greater than 83.0%. It should also be noted that dataset1 has many records interpolated mainly by the resolution of the petrophysical data RGB with lower reading resolution. Values for texture properties and average RGB color intensity do not have interpolated values. Finally, it is concluded that the features of petrophysical data and texture properties or RGB color intensity must be incorporated in a single dataset (dataset2) to improve the accuracy results. Factors such as the number of records also influence the result of accuracy.

The table below shows a summary of the best accuracy by dataset e group.

Table 13 Summary of the best accuracy by dataset, site, hole, group, input features, output feature and best accuracy. cod_lit refers to the lithology code according to the division of the groups.

Dataset	Site-Hole	Group	Input Features	Output Feature	Best Accuracy (%)
Dataset0	U1480-E	1	All	cod_lit	98.22

Dataset1	U1480-E,F,G,H	1	All	cod_lit	96.96
Dataset1	U1480-E,F,G,H	2	All	cod_lit	97.71
Dataset1	U1481-A	1	All	cod_lit	99.68
Dataset1	U1481-A	2	All	cod_lit	99.74
Dataset2	U1480-E,F,G,H	1	Practical_Arrang: 51	cod_lit	99.75
Dataset2	U1480-E,F,G,H	2	Practical_Arrang: 102	cod_lit	99.72
Dataset2	U1481-A	1	Practical_Arrang: 32	cod_lit	99.59
Dataset2	U1481-A	2	Practical_Arrang: 102	cod_lit	99.65

The best accuracy found in this study, compared with studies of lithological classification using random forest and a group of data equal to IODP (BRESSAN et al., 2020) and like IODP (according to table 6) are superior and capable of assisting the professional geologist in generating new knowledge and creating a dynamic environment to adapt the algorithms to different groups of data linked to IODP.

5. Conclusion

The dataset0, which it is composed of a relatively uniform number of records distributed across all lithologies, was extremely important to validate petrophysical data best interpolator. With accuracy validated at 98.19%, the best Akima interpolator is a simple, fast interpolator in the context of petrophysical data.

Dataset1 has many records distributed in a similar way among all the lithologies defined for classification, resulting in excellent accuracy results in all selected combinations. It is essential to highlight that for the structure of the datasets used in this article, the RF algorithm's necessary configuration has been optimized for its execution in a precise and fast way with excellent results in all selected definitions. The best-evaluated combinations are in 70% train - 30% test for dataset1, U1480, Groups

1 and 2, with accuracy greater than 96.00% and 80% train-20% test for dataset1, U1481, Groups 1 and 2, with an accuracy greater than 99.60%.

The dataset2 because it has records from dataset1 and records extracted from the segmented regions of the images, via texture and color analysis, comprises a robust dataset, precise and adapted to lithological classification in IODP-Expeditions. In this dataset, it was possible to organize various combinations of training and testing, generating reliability and precision in the formation and validation analysis. The practical arrangements with the best accuracy are 51 and 102 for Groups 1 and 2 and the combination of training and testing with the best result is 90% train-10% test for dataset2, U1480 and U1481.

The annotation of data on the images and segmentation of images by the SLIC Superpixel method proved to be robust and reliable for use in extracting characteristics and identifying lithologies in IODP-Expeditions.

With the new techniques developed in this article, future work can be developed in the increment of dataset1 and dataset2 with data from new IODP-Expeditions. It can be applied to the existent data set of earlier expeditions and during the drilling and data acquisition to help construct more robust geological models in the expeditions.

Computer Code Availability

Name of code: LithoPy v.2. Developer: Thiago Santi Bressan. Contact address: Programa de Pós-Graduação em Geologia, Universidade do Vale do Rio dos Sinos (Unisinos), Av. Unisinos, 950, Cristo Rei, São Leopoldo, RS, Brasil. Telephone number: +55-51-3591-1122. e-mail. tsbressan@edu.unisinos.br. Year first available: 2020. Hardware required: I3 CPU or better. Software required (minimum configuration): Anaconda (5.3), Jupyter Notebook (5.2), Sklearn (0.20), NumPy (>=1.8.2), SciPy (>= 0.13.3) and Matplotlib (3.0.0) Program language: Python. Program size: 10 MB. How to access the source code: Available at: https://github.com/tsbressan/LithoPy2.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. FCJ thanks the Joides Resolution Crew and IODP technical team for their support during IODP Expedition 362. This research used data of Expedition IODP 362 provided by the International Ocean Discovery Program (IODP) (www.iodp.org/access-data-and-samples).

Abbreviations used in this manuscript

APX	Area Superpixel
ANI	Andaman Nicobar Islands
ASM	Angular Second Moment
AUC	Area Under the ROC curve
CIELAB	CIE (International Commission on Illumination) L*a*b*
DPI	Dots Per Inch
FA	Facies Association
FPR	False Positive Rate
GLCM	Gray Level Co-occurrence Matrix
GRA	Gamma Ray Attenuation
IODP	International Ocean Discovery Program
JSON	JavaScript Object Notation
MAD	Moisture and Density
ML	Machine Learning
MS	Magnetic Susceptibility
NEAR	Ninety-East Ridge
NRG	Natural Gamma Radiation
PCHIP	Piecewise Cubic Hermite Interpolating Polynomial
PRACTICAL_ARRANG	Practical Arrangements
PWL	P-Wave Velocity Logger System
RF	Random Forest
ROC	Receiver Operating Characteristic
RGB	Red Green Blue

RSC	Reflectance Spectrophotometry
SLIC	Simple Linear Interative Clustering
TPR	True Positive Rate
WRMSL	Whole-Round Multisensor Logger

Appendix A – Supplementary data

Supplementary data related to this article can be found at https://github.com/tsbressan/LithoPy2.

FINAL CONSIDERATIONS

This study integrates the use of high-value geophysical data extracted from IODP-Expeditions with the evaluation and creation of new ML algorithms for specific learning in the range of properties (features) and images available.

The lithological classification using ML in data extracted from IODP-Expeditions has excellent results and is able to assist, streamline and automate all the work of the geologist in a short time during the expeditions. For lithological classification using only five geophysical data (GRA, PWL, MS, RSC and Magnetic Remanence (SRM)), with resolution of reading the data at 2.5 cm, returned accuracy equal to 89.51% in the random forest method. The random forest method integrates a broad, agile analysis, capable of returning excellent results applied to geophysical data IODP-Expeditions. With the addition of three geophysical data (MAD, NGR e RGB), totaling seven geophysical data in the formation of the dataset, with reading resolution of the data at 0.04cm and 1cm (interpolated values), returned accuracy equal to 99.74%. The combination of geophysical data and high-resolution images found excellent results of accuracy, with values equal to 99.75%. The integration between geophysical data and image is necessary due to the characteristics of image extraction and the physical-visual analysis between the divisions of the lithologies present in the study areas of the IODP-Expeditions. The SLIC Superpixel method and its settings and extractions of texture values properly process IODP images maintaining and recognizing patterns between lithologies, being an excellent method suitable the characteristics of description and classification of lithologies.

From these results the main conclusions are based on three main pillars: Education, Exploration and Knowledge Generation. In Exploration, the development of algorithms (a practical tool) covers application in geosciences allowing the geoscientist to develop/interpret models at the time of drilling, orienting new practical studies and improving the drilling strategy in IODP-Expeditions and the like. Exploration requires accurate feedback of information in a short period of time, the algorithms of this study being suitable for this one, adapting to various geological physical contexts both in Brazil and in other locations.

For Education, the presentation of these targeted practical tools allows it to be applied in technological higher education environments (undergraduate and graduate) directed

to the development of scientific projects applied to real data, integrating better training of these students with fast, accurate and capable information to achieve a differential in the discovery of new information. IODP-Expeditions generate a range of data (numerical or images), mainly after the year 2013, stored in an appropriate repository and accessed through the website (https://web.iodp.tamu.edu/LORE/), this material can be used as a practical exercise for students.

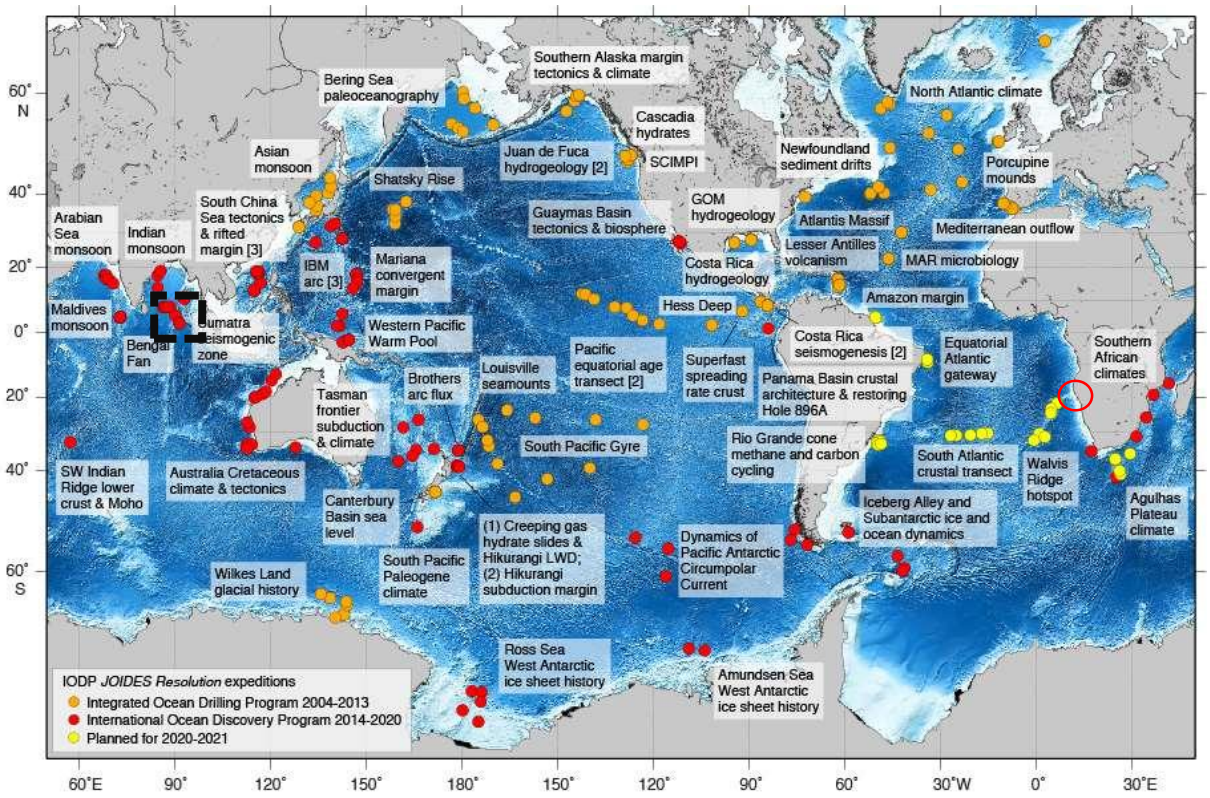


Figure 23 The map showed through the color demarcation the list of deep-water drilling programs, linked to IODP, covering holes in all continents and oceans, with access to high-resolution data (numerical and image) and with great added scientific value (IODP, 2020). Highlighted, the location of IODP-Expedition 362, sites U1480 and U1481.

As an example, according to Fig. 23, we can mention the IODP (and its other programs: DSDP (Deep Sea Drilling Project: from 1965 to 1983), ODP (Ocean Drilling Program: from 1983 to 2003), IODP (Integrated Ocean Drilling Program: from 2003 to 2013) and IODP (International Ocean Discovery Program: from 2013-present)) with data exceeding 1,200 holes, making it a great program currently an international one with participation from more than 25 countries.

In the Knowledge Generation, the use of this tool can improve the availability of the generation of digital libraries/museums of already existing libraries, such as in institutions such as Petrobras, CPRM (Companhia de Pesquisa de Recursos Minerais) and scientific exploration organizations, containing valuable information from the Brazilian soil or even from the Atlantic Ocean itself. The proper use of these algorithms will allow the data to be manipulated quickly and automatically, reducing time and physical and human resources for the institution. Once the data is in digital format, the geoscientist can access the data via the cloud (via the internet) anywhere, define the geological study environment, select the methods and create a favorable environment for decision making and advanced knowledge of the location that is inserted.

In this study, there is an ideal situation that can be applied to exploration with data from different sources and with a variable number of characteristics (features). It should be added that the practical tool is capable of automatically incorporating and increasing the dataset, creating an opportune environment for creating new knowledge with a great learning process on the large database, envisioning a practical relationship with technologies Big Data and Data Science.

REFERENCES

ACHANTA, R. et al., 2010. SLIC Superpixels. EPFL Technical Report 149300, June 2010.

ACHANTA, R. et al., 2012. SLIC Superpixels Compared to State-of-the-Art Superpixel Methods. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34,11. https://doi.org/10.1109/TPAMI.2012.120.

ACHARYA, T., RAY, A. K., 2005. Image Processing. Principles and Applications. Wiley-Interscience.

ALZUBAIDI, F. et al., 2020. Automated lithology classification from drill core images using convolutional neural networks. Journal of Petroleum Science and Engineering, 197, February 2021. https://doi.org/10.1016/j.petrol.2020.107933.

AL-MUDHAFAR, W. J., 2016. Applied Geostatistical Reservoir Characterization in R: Review and Implementation of Rock Facies Classification and Prediction Algorithms-Part I. Offshore Technology Conference. doi:10.4043/26947-MS.

ALAZAWI, S. A., SHATI, N. M., ABBAS, A. H., 2019. Texture features extraction based on GLCM for face retrieval system. Periodicals of Engineering and Natural Sciences, 7, 3. http://dx.doi.org/10.21533/pen.v7i3.787.

AKIMA, H., 1986. A Method of Univariate Interpolation That Has the Accuracy of a Third-Degree Polynomial. Institute for Telecommunication Sciences, National Telecommunications and Information Administration, U.S. Department of Commerce.

AMANTE, C., EAKINS, B. W., 2009. ETOPO1 Global Relief Model converted to PanMap layer format. NOAA-National Geophysical Data Center, PANGAEA, https://doi.org/10.1594/PANGAEA.769615.

AMINZADEH, F., DASGUPTA, S. N., 2013. Chapter 9 - Geophysics for Unconventional Resources. Developments in Petroleum Science, 60, 2013. https://doi.org/10.1016/B978-0-444-50662-7.00009-3.

AO, Y. et al., 2020. Probabilistic logging lithology characterization with random forest probability estimation. Computers & Geosciences, 104556. doi:10.1016/j.cageo.2020.104556.

BACHRI, I. et al., 2019. Machine Learning Algorithms for Automatic Lithological Mapping Using Remote Sensing Data: A Case Study from Souk Arbaa Sahel, Sidi Ifni Inlier, Western Anti-Atlas, Morocco. ISPRS Int. J. Geo-Inf. 2019, 8(6). https://doi.org/10.3390/ijgi8060248.

BETZLER, C. et al., 2017. Expedition 359 summary. Proceedings of the International Ocean Discovery Program (359), https://doi.org/10.14379/iodp.proc.359.101.2017.

BLUM, P., 1997. Physical properties handbook: a guide to the shipboard measurement of physical properties of deep-sea cores, College Station, Texas, USA, http://www-odp.tamu.edu.

BRCKOVIC, A. et al., 2017. Application of artificial neural networks for lithofacies determination based on limited well data. Central European Geology (60)3, 299-315. https://doi.org/10.1556/24.60.2017.012.

BRESSAN, T. S. et al., 2020. Evaluation of machine learning methods for lithology classification using geophysical data. Computers & Geosciences, 139. https://doi.org/10.1016/j.cageo.2020.104475.

BRIEUC, M. S. O. et al., 2018. A practical introduction to Random Forest for genetic association studies in ecology and evolution. Molecular Ecology Resources (18)4, 755-766, https://doi.org/10.1111/1755-0998.12773.

BRUNO, A. E. et al., 2018. Classification of crystallization outcomes using deep convolutional neural networks. PLoS ONE (13)6. https://doi.org/10.1371/journal.pone.0198883.

BUSCHOW, K. H. J., BOER, F. R. de., 2004. Physics of Magnetism and Magnetic Materials, Kluwer Academic Publishers, ISBN 0-306-47421-2, pp.105-115, https://doi.org/10.1007/b100503.

BUZZI-FERRARIS, G., MANENTI, F., 2010. Interpolation and Regression Models for the Chemical Engineer: Solving Numerical Problems. Chapter 1. WILEY-VCH Verlag GmbH & Co. KGaA.

CAO, C. et al., 2018. Deep Learning and Its Applications in Biomedicine. Genomics Proteomics Bioinformatics (16), https://doi.org/10.1016/j.gpb.2017.07.003.

CASTRO, W. et al., 2017. Multilayer perceptron architecture optimization using parallel computing techniques. PLoS ONE (12)12, https://doi.org/10.1371/journal.pone.0189369.

CELANT, G., BRONIATOWSKI, M., 2016. Interpolation and Extrapolation Optimal Designs 1: Polynomial Regression and Approximation Theory. Wiley Online Library. DOI: 10.1002/9781119292272.

CHAI, X. et al., 2020. Deep learning for irregularly and regularly missing data reconstruction. Scientific Reports 10, 3302 (2020). https://doi.org/10.1038/s41598-020-59801-x.

CHEN, Y. et al., 2018. The Interpolation of Sparse Geophysical Data. Springer: Surveys in Geophysics, 40, 73-105 (2019). https://doi.org/10.1007/s10712-018-9501-3.

CHEN, X., ZHANG, F., ZHANG, R., 2017. Medical image segmentation based on SLIC superpixels model. International Conference on Innovative Optical Health Science, 10245 (2017). https://doi.org/10.1117/12.2258384.

CIE, 2004. Technical Report: COLORIMETRY, 3rd Edition. ISBN: 3901906339, https://archive.org/details/gov.law.cie.15.2004, 82pp.

CIOFFI, R. et al., 2020. Artificial Intelligence and Machine Learning Applications in Smart Production: Progress, Trends, and Directions. Sustainability 2020, 12, 492. https://doi.org/10.3390/su12020492.

CLIFT, P.D. et. al., 2008. Correlation of Himalayan exhumation rates and Asian monsoon intensity. Nature Geoscience, 1(12), 875–880. https://doi.org/10.1038/ngeo351.

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, February 2014. https://doi.org/10.1016/j.cageo.2013.10.008.

CURRAY J.R. et al., 2003. The Bengal Fan: morphology, geometry, stratigraphy, history and processes. Marine Petrol Geol 19:1191–1223.

CURRAY J.R., MOORE, D. G., 1974. Sedimentary and Tectonic Processes in the Bengal Deep-Sea Fan and Geosyncline. The Geology of Continental Margins. Springer. https://doi.org/10.1007/978-3-662-01141-6_45.

DE BOISSIEU, F. et al., 2017. Regolith-geology mapping with support vector machine: A case study over weathered Ni-bearing peridotites, New Caledonia. International Journal of Applied Earth Observation and Geoinformation (64), 377-385, http://dx.doi.org/10.1016/j.jag.2017.05.012.

DENG, C. et al., 2020. Integrating Machine Learning with Human Knowledge. iScience, 23 (11), 2020. https://doi.org/10.1016/j.isci.2020.101656.

DOVETON, J. H., 1994. Geological Log Interpretation. SEPM Society for Sedimentary Geology (29), https://doi.org/10.2110/scn.94.29.

DUBEAU, P. et al., 2017. Mapping the Dabus Wetlands, Ethiopia, Using Random Forest Classification of Landsat, PALSAR and Topographic Data. Remote Sensing (9)1056, https://doi.org/10.3390/rs9101056.

DUTTA, A., ZISSERMAN, A., 2019. The VIA Annotation Software for Images, Audio and Video. 27th ACM International Conference on Multimedia. https://doi.org/10.1145/3343031.3350535.

ELAZIZ, M. A. et al., 2020. New machine learning method for image-based diagnosis of COVID-19. PLOS ONE. https://doi.org/10.1371/journal.pone.0235187.

EGAN, J. P., 1975. Signal detection theory and ROC analysis. New York: Academic Press, 1975.

FAN, G. et al., 2020. Recognizing Multiple Types of Rocks Quickly and Accurately Based on Lightweight CNNs Model. IEEE Access (8). https://doi.org/10.1109/ACCESS.2020.2982017.

FAN, B. et al., 2019. Lithologic heterogeneity of lacustrine shale and its geological significance for shale hydrocarbon-a case study of Zhangjiatan Shale. Open Geosci. 2019(11), 101–112, https://doi.org/10.1515/geo-2019-0009.

FARRELL, P., 2018. Numerical Mathematics. Lecture Notes. Hamburg University of Technology, Version 22, August 2018.

FEDOROV, D. V., 2013. Introduction to Numerical Methods. Version 13.05. GNU Licence.

FRANCE-LANORD, C. et al., 2016. Expedition 354 summary. Proceedings of the International Ocean Discovery Program (354). https://doi.org/10.14379/iodp.proc.354.101.2016.

FRITSCH, F. N., 1982. PCHIP FINAL SPECIFICATIONS. Version 8.5. Lawrence Livermore National Laboratory, August 1982.

GALLAGHER, S.J. et al., 2017. Expedition 356 summary. Proceedings of the International Ocean Discovery Program (356), https://doi.org/10.14379/iodp.proc.356.101.2017.

GAJOWNICZEK, K., ZABKOWSKI, T., SZUPILUK, R., 2014. Estimating the Roc Curve and its Significance for Classification Models' Assessment. Quantitative Methods in Economics, 15 (2), 2014,382 – 391pp.

GÉRON, A., 2017. Hands-On Machine Learning with Scikit-Learn and TensorFlow. Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media.

GIRELLI, T. J. et al., 2021a. Source changes in the Nicobar Fan over the Late Cenozoic based on U-Pb-Hf zircon isotopes. Submitted to Geochemistry, Geophysics, Geosystems.

GIRELLI, T. J. et al., 2021b. Alkaline Kerguelen hotspot-derived magmatism in the Indian Ocean – IODP 362 expedition. Submitted to Terra Nova.

GOLLAPUDI, S., 2016. Practical Machine Learning. Packt Publishing, Birmingham B3 2PB, UK.

GONZALES, R. C., WOODS, R. E., 2007. Digital Image Processing, Third Edition. Pearson.

GOODRICH, W. E., 2007. Characterization and Quantification of Magnetic Remanence in Unexploded Ordnance. Master of Science (Geophysics), Faculty and Board of Trustees of the Colorado School, https://mountainscholar.org/bitstream/handle/11124/79225/T06342.pdf?sequence=1.

GRANGE, S. K. et al., 2018. Random forest meteorological normalization models for Swiss PM10 trend analysis. Atmos. Chem. Phys. (18), 6223-6239, https://doi.org/10.5194/acp-18-6223-2018.

GREWAL, D. S., 2014. A Critical Conceptual Analysis of Definitions of Artificial Intelligence as Applicable to Computer Engineering. IOSR Journal of Computer Engineering, 16, 9-13. https://doi.org/10.9790/0661-16210913.

GUO, Z. et al., 2018. Geophysical Field Data Interpolation Using Stochastic Partial Differential Equations for Gold Exploration in Dayaoshan, Guangxi, China. Minerals 2019, 9, 14. https://doi.org/10.3390/min9010014.

HALAGUNDEGOWDA, G. R., ABHISHEK, S., 2018. Multilayer Perceptron Method of Artificial Neural Network for Classification of Farmers Based on Adoption of Drought Coping Mechanisms. Int. J. Pure App. Biosci. (6)2, http://dx.doi.org/10.18782/2320-7051.6405.

HALL, I.R. et al., 2017. Expedition 361 summary. Proceedings of the International Ocean Discovery Program (361), https://doi.org/10.14379/iodp.proc.361.101.2017.

HALL-BEYER, M., 2017. GLCM Texture: A Tutorial v. 3.0. University of Calgary. http://dx.doi.org/10.11575/PRISM/33280.

HAMOUDA, S., ZAINOL, Z., 2019. Semi-Structured Data Model for Big Data (SS-DMBD). 8th International Conference on Data Science, Technology and Applications (DATA 2019). DOI: 10.5220/0007957603480356.

HARALICK, R. M., SHANMUGAM, K., DINSTEIN, I., 1973. Textural Features for Image Classification. IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, 3, 6. https://doi.org/10.1109/TSMC.1973.4309314.

HAYKIN, S., 2009. Neural Networks and Learning Machines. Third Edition. Pearson Prentice Hall, New York City, U. S., 938pp.

HINSBERGEN, D. J. J. van. et al., 2012. Greater India Basin hypothesis and a two-stage Cenozoic collision between India and Asia. PNAS 109 (20) 7659-7664. https://doi.org/10.1073/pnas.1117262109.

HOSSIN, M., SULAIMAN, M. N., 2015. A Review on Evaluation Metrics for Data Classification Evaluations. International Journal of Data Mining & Knowledge Management Process (5)2, 1-11, https://doi.org/10.5121/ijdkp.2015.5201.

HSIEH, B., LEWIS, C., LIN, Z., 2005. Lithology identification of aquifers from geophysical well logs and fuzzy logic analysis: Shui-Lin Area, Taiwan. Computers & Geosciences (31)3, 263-275, https://doi.org/10.1016/j.cageo.2004.07.004.

HUANG, S. et al., 2018. Applications of Support Vector Machine (SVM) Learning in Cancer Genomics. Cancer Genomics Proteomics (15)1, 41-51, https://doi.org/10.21873/cgp.20063.

HUGHES, V. K., LANGLOIS, N. E. I., 2010. Use of reflectance spectrophotometry and colorimetry in a general linear model for the determination of the age of bruises. Forensic Sci Med Pathol. 6:275–28, https://doi.org/10.1007/s12024-010-9171-z.

IODP, 2020. Combined International Ocean Discovery Program and Integrated Ocean Drilling Program Map. https://iodp.tamu.edu/scienceops/index.html.

JAHDHAMI, N. A., ANBOORI, A., 2017. The Application of Specific Drilling Energy to Identify Overburden Lithological Boundaries and Aid Well Operations - Oman Khazzan

Field. Abu Dhabi International Petroleum Exhibition & Conference, Abu Dhabi, UAE. https://doi.org/10.2118/188413-MS.

JAMES, G. et al., 2013. An Introduction to Statistical Learning. Springer, New York City, U.S, doi:10.1007/978-1-4614-7138-7, 187pp.

JIANG, F. et al., 2017. Grain segmentation of multi-angle petrographic thin section microscopic images. IEEE International Conference on Image Processing (ICIP). https://doi.org/10.1109/ICIP.2017.8297009.

JING, S., et al., 2017. An Efficient Algorithm for Parallel Computation of Rough Entropy Using CUDA. International Conference on Computational Intelligence and Security (13), https://doi.org/10.1109/CIS.2017.00009.

JOVANE, L. et al., 2013. Magnetic Methods and the Timing of Geological Processes. Geological Society Special Publication Nº 373, https://doi.org/10.1144/SP373.17.

KAVZOGLU, T., TONBUL, H., 2018. An experimental comparison of multi-resolution segmentation, SLIC and K-means clustering for object-based classification of VHR imagery. International Journal of Remote Sensing, 39, 18. https://doi.org/10.1080/01431161.2018.1506592

KOK, J. N. et al., 2009. Artificial Intelligence: Definition, Trends, Techniques, and Cases. Artificial Intelligence - Volume 1. Encyclopedia of Life Support Systems (EOLSS).

KORJUS, K., HEBART, M. N., VICENTE, R., 2016. An Efficient Data Partitioning to Improve Classification Performance While Keeping Parameters Interpretable. PLoS ONE 11(8), https://doi.org/10.1371/journal.pone.0161788.

KOROLEV, E. A. et al., 2018. Effect of lithological heterogeneity of bitumen sandstones on SAGD reservoir development. IOP Conf. Series: Earth and Environmental Science 155 (2018), doi :10.1088/1755-1315/155/1/012019.

KOTU, V., DESHPANDE, B., 2019. Chapter 4 – Classification. Data Science (Second Edition). Concepts and Practice. https://doi.org/10.1016/B978-0-12-814761-0.00004-6.

KUMAR, C. et al., 2019. Automated lithological mapping by integrating spectral enhancement techniques and machine learning algorithms using AVIRIS-NG hyperspectral data in Gold-bearing granite-greenstone rocks in Hutti, India. International Journal of Applied Earth Observation and Geoinformation, 86, April 2020. https://doi.org/10.1016/j.jag.2019.102006.

LATIFOVIC, R., POULIOT, D. CAMPBELL, J., 2018. Assessment of Convolution Neural Networks for Surficial Geology Mapping in the South Rae Geological Region, Northwest Territories, Canada. Remote Sensing (10)2, 307, https://doi.org/10.3390/rs10020307.

LI, J., TRAN M., SIWABESSY, J., 2016. Selecting Optimal Random Forest Predictive Models: A Case Study on Predicting the Spatial Distribution of Seabed Hardness. Geoscience Australia GPO Box 378, https://doi.org/10.1371/journal.pone.0149089.

LI, C.-F. et al., 2015. Expedition 349 summary. Proceedings of the International Ocean Discovery Program (349), https://doi.org/10.14379/iodp.proc.349.101.2015.

LOUSSAIEF, S., ABDELKRIM, A., 2018. Machine Learning framework for image classification. ASTESJ (3)1, https://doi.org/10.1109/SETIT.2016.7939841.

LYCHE, T., MØRKEN, K., 2018. Spline Methods Draft. Department of Mathematics. University of Oslo.

MAIMON, O., ROKACH, L., 2010. Data Mining and Knowledge Discovery Handbook. 2 editions. Springer, New York City, U.S, Chapter 9, https://doi.org/10.1007/978-0-387-09823-4.

MAITRE, J., BOUCHARD, K., BÉDARD, L. P., 2019. Mineral grains recognition using computer vision and machine learning. Computers & Geosciences, 130, 84-93 (2019). https://doi.org/10.1016/j.cageo.2019.05.009.

MALLADI, S. R. S. P., RAM, S., RODRÍGUEZ, J. J., 2014. Superpixels using morphology for rock image segmentation. Southwest Symposium on Image Analysis and Interpretation. https://doi.org/10.1109/SSIAI.2014.6806050.

MANSOURKHAKI, A., BERANGI, M., HAGHIRI, M., 2017. Comparative Application of Radial Basis Function and Multilayer Perceptron Neural Networks to Predict Traffic Noise Pollution in Tehran Roads. Journal of Ecological Engineering (19)1, https://doi.org/10.12911/22998993/79411.

MARTINEZ-MOSQUERA, D., NAVARRETE, R., LUJAN-MORA, S., 2020. Modeling and Management Big Data in Databases—A Systematic Literature Review. Sustainability, 12, 634. DOI:10.3390/su12020634.

MCNEILL, L. C. et al., 2017. Expedition 362 summary. Proceedings of the International Ocean Discovery Program (362). https://doi.org/10.14379/iodp.proc.362.101.2017.

MILLIMAN, J., SYVITSKI, J., 1992. Geomorphic/Tectonic Control of Sediment Discharge to the Ocean: The Importance of Small Mountainous Rivers. The Journal of Geology 100 (5). https://doi.org/10.1086/629606.

NAVIN, M., PANKAJA, R., 2016. Performance Analysis of Text Classification Algorithms using Confusion Matrix. International Journal of Engineering and Technical Research (6)4, pp. 75-78.

NGUIMBOUS-KOUOH, J. J., MANGUELLE-DICOUM, E., 2019. Evaluating Interpolation Methods by Geostatistical Modeling of the Douala Oil Field Porosity Data (Cameroon). Geoinfor Geostat 7, 1. DOI: 10.4172/2327-4581.1000203.

NISBET, R., MINER, G., YALE, K. D. D. S., 2018. Chapter 11 - Model Evaluation and Enhancement. Handbook of Statistical Analysis and Data Mining Applications (Second Edition) 2018, 215-233. https://doi.org/10.1016/B978-0-12-416632-5.00011-6.

ODP, 2007. ODP Prime Scientific Data: Collection, Archive, and Quality ODP. Technical Note 37. URL: http://www-odp.tamu.edu/publications/tnotes/tn37/TNOTE_37.PDF.

OJHA, M., MAITI, S., 2016. Sediment classification using neural networks: An example from the site-U1344A of IODP Expedition 323 in the Bering Sea. Deep Sea Research Part II: Topical Studies in Oceanography, 125–126, 2016. https://doi.org/10.1016/j.dsr2.2013.03.024.

PANDEY, D. K. et al., 2016. Expedition 355 summary. Proceedings of the International Ocean Discovery Program (355), https://doi.org/10.14379/iodp.proc.355.101.2016.

PAPERNOT, N. et al., 2016. The Limitations of Deep Learning in Adversarial Settings. European Symposium on Security and Privacy, 2016, Germany. https://doi.org/10.1109/EuroSP.2016.36.

PENG, H., BAI, X., 2017. Limits of Machine Learning Approach on Improving Orbit Prediction Accuracy using Support Vector Machine. Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conference, 2017, Hawaii.

PENG, J., ZHOU, Y., CHEN, C. L. P., 2015. Region-Kernel-Based Support Vector Machines for Hyperspectral Image Classification. IEEE Transactions on Geoscience and Remote Sensing (53)9, 4810-4824, https://doi.org/10.1109/TGRS.2015.2410991.

PICKERING, K.T. at al., 2019. Sedimentology, stratigraphy and architecture of the Nicobar Fan (Bengal–Nicobar Fan System), Indian Ocean: Results from International Ocean Discovery Program Expedition 362. Sedimentology. https://doi.org/10.1111/sed.12701.

PROYAN, V., KISELEV, Y., 2010. Statistical Methods of Geophysical Data Processing. World Scientific Publishing Co.

PUGGINI, L., DOYLE, J., MCLOONE, S., 2015. Fault Detection using Random Forest Similarity Distance. IFAC-PapersOnLine (48)21, 583-588, https://doi.org/10.1016/j.ifacol.2015.09.589.

RABBATH, C. A., CORRIVEAU, D., 2019. A comparison of piecewise cubic Hermite interpolating polynomials, cubic splines and piecewise linear functions for the approximation of projectile aerodynamics. Defence Technology 15 (5). https://doi.org/10.1016/j.dt.2019.07.016.

RAFIK, B., KAMEL, B., 2017. Prediction of permeability and porosity from well log data using the nonparametric regression with multivariate analysis and neural network, Hassi R'Mel Field, Algeria. Egyptian Journal of Petroleum 26(3), 763-778, http://dx.doi.org/10.1016/j.ejpe.2016.10.013.

RAHIM, I. A. et al., 2009. Lithological Unit Thickness Approach for Determining Intact Rock Strength (IRS) of Slope Forming Rock Material of Crocker Formation. Borneo Science. 25, pp. 23-32.

RAN, X. et al., 2019. Rock Classification from Field Image Patches Analyzed Using a Deep Convolutional Neural Network. Mathematics, 7, 755. DOI:10.3390/math7080755.

RASCHKA, S., 2015. Python Machine Learning. Packt Publishing Ltd, England, 454pp.

SÁEZ, J.A. et al., 2013. Tackling the problem of classification with noisy data using Multiple Classifier Systems: Analysis of the performance and robustness. Information Sciences (247), http://dx.doi.org/10.1016/j.ins.2013.06.002.

SALOMON, D., 2006. Curves and Surfaces for Computer Graphics. Chapter 2. Springer. https://doi.org/10.1007/0-387-28452-4.

SARKAR, D., BALI, R., SHARMA, T., 2018. Practical Machine Learning with Python. A Problem-Solver's Guide to Building Real-World Intelligent Systems. APRESS. https://doi.org/10.1007/978-1-4842-3207-1.

SAVALIA, S., EMAMIAN, V., 2018. Cardiac Arrhythmia Classification by Multi-Layer Perceptron and Convolution Neural Networks. Bioengineering (5)35, https://doi.org/10.3390/bioengineering5020035.

SHI, J. S. et al., 2019. Effects of annotation granularity in deep learning models for histopathological images. 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). https://doi.org/10.1109/BIBM47256.2019.8983158.

SILVA, A. A. et al. 2020. Petrofacies Classification using Machine Learning Algorithms. GEOPHYSICS,85 (4), 2020. https://doi.org/10.1190/geo2019-0439.1.

SINNOTT, R. O., SUN, H. D., 2016. Chapter 15 - A Case Study in Big Data Analytics: Exploring Twitter Sentiment Analysis and the Weather. Big Data, Principles and Paradigms. https://doi.org/10.1016/B978-0-12-805394-2.00015-5.

SOUZA, J.F.L. et al., 2019. Automatic classification of hydrocarbon ‘‘leads’’ in seismic images through artificial and convolutional neural networks. Computers & Geosciences 132, 23-32, https://doi.org/10.1016/j.cageo.2019.07.002.

SPRONCK, P., 2017. The Coder's Apprentice - Learning Programming with Python 3. Version 1.0.16. Spronck Create Commons Licences, http://www.spronck.net/pythonbook/pythonbook.pdf, 398pp.

STRAUß, S., 2018. From Big Data to Deep Learning: A Leap Towards Strong AI or ‘Intelligentia Obscura’?. Big Data Cognitive Computing 2(3) 16, http://dx.doi.org/10.3390/bdcc2030016.

STREKALOVA, Y.A., BOUAKKAZ, M., 2017. Semi-structured Data. In: Schintler L., McNeely C. (eds) Encyclopedia of Big Data. Springer, Cham. https://doi.org/10.1007/978-3-319-32001-4_183-1.

STORKEY, A. J., 2013. When Training and Test Sets are Different: Characterising Learning Transfer. https://dx.doi.org/10.7551/mitpress/9780262170055.003.0001.

SUN, Z. et al., 2020. A Data-Driven Approach for Lithology Identification Based on Parameter-Optimized Ensemble Learning. Energies, 13, 3903. DOI:10.3390/en13153903.

THARWAT, A., 2018. Classification assessment methods. Applied Computing and Informatics, https://doi.org/10.1016/j.aci.2018.08.003.

TILAKI-HAJIAN, K., 2013. Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. Caspian J Intern Med. 4(2), 627–635pp.

TORRALBA, A., RUSSELL, B. C., YUEN, J., 2010. LabelMe: Online Image Annotation and Applications. Proceedings of the IEEE, 98, 8 (2010). https://doi.org/10.1109/JPROC.2010.2050290.

TSE, K.C. et al., 2018. Unsupervised learning on scientific ocean drilling datasets from the South China Sea. Front. Earth Sci. 13, 180–190 (2019). https://doi.org/10.1007/s11707-018-0704-1.

TURCO, F., AZEVEDO, L., HEROLD, D., 2019. Geostatistical interpolation of non-stationary seismic data. Computacional Geosciences 23, 665-682 (2019). https://doi.org/10.1007/s10596-019-9812-6.

VAPNIK, V. N., 1998. Statistical Learning Theory. A Wiley-Interscience Publication, New York City, U. S., 740pp.

VASUKI, Y. et al., 2017. An interactive image segmentation method for lithological boundary detection: A rapid mapping tool for geologists. Computers & Geosciences, 100, 27-40. https://doi.org/10.1016/j.cageo.2016.12.001.

VAKHSHOORI, V., ZARE, M., 2018. Is the ROC curve a reliable tool to compare the validity of landslide susceptibility maps? Geomatics, Natural Hazards and Risk (9)1, 2018, https://doi.org/10.1080/19475705.2018.1424043.

VENKATARAMAN, S., 2017. System Design for Large Scale Machine Learning. Electrical. University of California at Berkeley, EECS Department, https://www2.eecs.berkeley.edu/Pubs/TechRpts/2017/EECS-2017-219.html.

XIE, Y. et al., 2017. Evaluation of machine learning methods for formation lithology identification: A comparison of tuning processes and model performances. Journal of Petroleum Science and Engineering, 160, 182-193. doi:10.1016/j.petrol.2017.10.028.

XIE, Y. et al., 2020. A Coarse-to-Fine Approach for Intelligent Logging Lithology Identification with Extremely Randomized Trees. Math Geosci (2020). https://doi.org/10.1007/s11004-020-09885-y.

XU LIU, J. O. et al., 2018. Mapping dengue risk in Singapore using Random Forest. PLOS Neglected Tropical Diseases (2), https://doi.org/10.1371/journal.pntd.0006587.

YU, L. et al., 2012. Towards automatic lithological classification from remote sensing data using support vector machines. Computers & Geosciences (45), https://doi.org/10.1016/j.cageo.2011.11.019.

YUE, W. et al., 2018. Machine Learning with Applications in Breast Cancer Diagnosis and Prognosis. Designs (2)13, https://doi.org/10.3390/designs2020013.

ZHAO, L. N. at al., 2011. Verification and comparison of probabilistic precipitation forecasts using the TIGGE data in the upriver of Huaihe Basin. Adv. Geosci. (29), 2011, 95-102pp, https://doi.org/10.5194/adgeo-29-95-2011.

Appendix

Appendix 1A – Description of the Study Site and IODP Expeditions

Fig. 24 shows a map of expeditions and holes related to IODP Expeditions 349, 354, 355, 356, 359, 361 and 362.

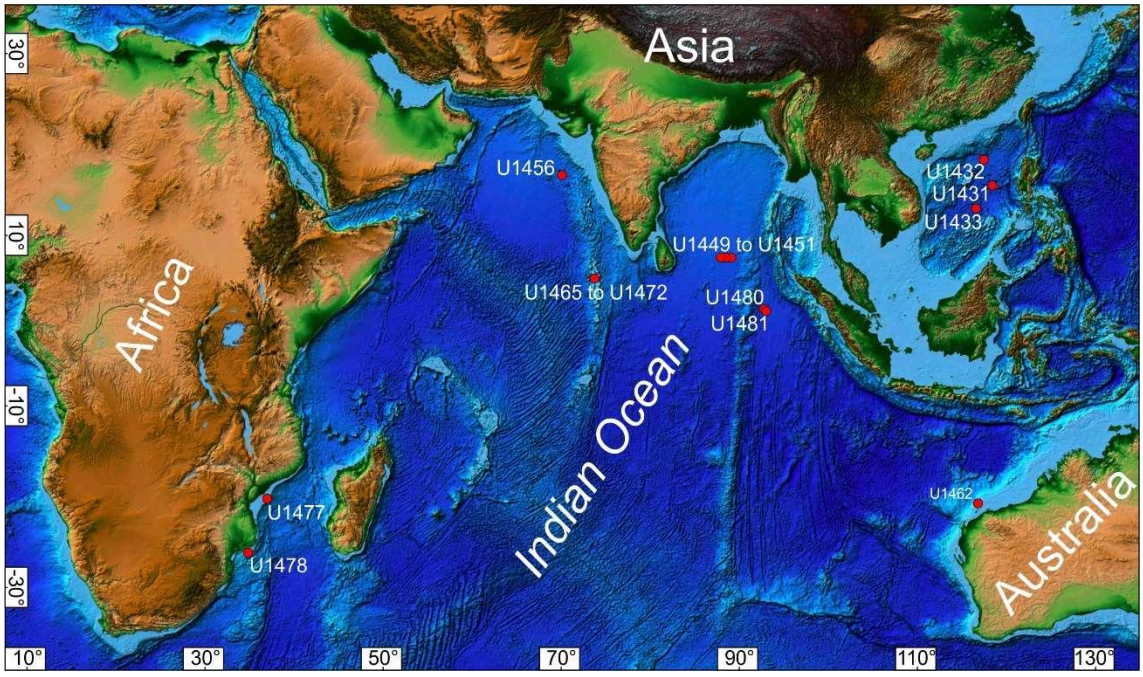


Figure 24 Map of expeditions sites for this study. Sites U1431, U1432 and U1433 – IODP Expedition 349 (LI et al., 2015); Sites U1449, U1450 and U1451 – IODP Expedition 354 (FRANCE-LANORD et al., 2016); Site U1456 – IODP Expedition 355 (PANDEY et al., 2016); Site U1462 – IODP Expedition 356 (GALLAGHER et al., 2017); Sites U1477 and U1478 – IODP Expedition 361 (HALL et al., 2017); Sites U1480 and U1481 – IODP Expedition 362 (MCNEILL et al., 2017); Sites U1465, U1466, U1467, U1468, U1470, U1471 and U1472 – IODP Expedition 359 (BETZLER et al., 2017). Map modified from ETOPO1 Global Relief Model (AMANTE and EAKINS, 2009).

IODP Expedition 362, which occurred from August 6 to October 6, 2016, carried out the drilling of two sites U1480 and U1481, whose main problems to be answered are the establishment of the initial and evolving properties of the Sumatran incoming sedimentary section and the potential effect of these properties on seismogenesis, tsunamigenesis, and forearc development for comparison with global examples (MCNEILL et al., 2017). It is highlighted in this expedition as objectives the determination of the lithology, sedimentation rates, physical, chemical and thermal properties of each entry section.

Sediments and sedimentary rocks were recovered from a depth of 1,415.35 m in nine holes U1480 (A-H) and U1481 (A) (MCNEILL et al., 2017). The unconsolidated

sediments and sedimentary, tuff layers and basic rocks formed between present and Cretaceous Late. The global succession of the holes represents two sections, the pre-Nicobar Fan and Nicobar Fan ones. The pre-Nicobar Fan succession comprises Creataceous Late to Eocene volcanic and siliciclastic-carbonate rocks accumulating between the Sunda subduction zone and the Ninetyeast Ridge. Overlying this pre-Nicobar Fans succession, occur Miocene to early Pleistocene siliciclastic sediments deposited from various gravity flows, derived from Nicobar Fan, with some cm- to mm-thick ash layers. From Pleistocene to recent pelagic sediments with interlayered ash layers covered the drilled sections.

IODP Expedition 354 occurred from January 29 to March 31, 2015, where drilling of seven sites was carried out, 320km long transect across the Bengal Fan. These drillings give a spatial view of the water and sediment deposit of the system comprising the Bengal Fan. The sediments collected in this expedition come from the Himalayan Rivers, generated from terrestrial changes, erosions and weathering of the Himalayas and are transported through a delta and shelf canyon, providing large turbidity currents loaded with a wide spectrum of grain size (FRANCE-LANORD et al., 2016).

IODP Expedition 355 drilled two sites (U1456 and U1457) in Laxmi Basin in the eastern Arabian Sea to document the evolution of mountain construction, weathering, erosion and climate over a range of timescales. The Indian summer monsoon is one of the most intense climatic phenomena on the planet, being linked to the growth of topography in south and central Asia. The expedition collected samples from a depth of 1,109.40m below the seabed. In these drillings, the sediments were dated from 13.50 to 17.70 million years (Site U1456) and ranging from 10.90 to 62.00 million years (site U1457). Site U1457 is more distal in facies, reflecting its more marginal setting. No major active lobe appears to have affected Laxmi Basin since late Pleistocene (PANDEY et al., 2016).

IODP Expedition 359 performed at the Indian Ocean, Maldives Archipelago, from September 30 to November 30, 2015, seeking to study sea level changes and their currents along with the evolution of the Indian Ocean Monsoon. Seven sites (U1465, U1466, U1467, U1468, U1470, U1471 and U1472) were drilled to provide unread records of the evolving ocean Cenozoic icehouse world. The main objective of this

expedition has been to date precisely the onset of the current system (BETZLER et al., 2017).

IODP Expedition 349 drilled five sites in the South China Sea basin (SCS), seeking to deepen basin formation studies and understand east Asian tectonic evolution, 1524.00m of sediments/sedimentary rocks, 78.00m of oceanic basalt were recovered and geophysical information was collected from the wells mainly in places with higher depth (LI et al., 2015).

IODP Expedition 356 traversed the northwest of Australia, searching for precise information on the Indonesian Throughflow (ITF) and all the climatic evolution that has been linked to the Australian monsoons and their variability over time. It should be noted that the expedition sought to identify nature and development of the entire sedimentary structure on the Australian Continent (GALLAGHER et al., 2017).

IODP Expedition 361 drilled six sites on the southeast African margin, totaling 5,175.00m of recovered rocks. The sites are located between the Mozambique channel and the Natal Valley, influenced by the discharge of hydrographic materials from the rivers, resulting in an environment with high potential for temporal studies and with the capacity to identify the complete formation of south-east Africa (HALL et al., 2017).

Appendix 1B – Core imagens of different lithologic groups

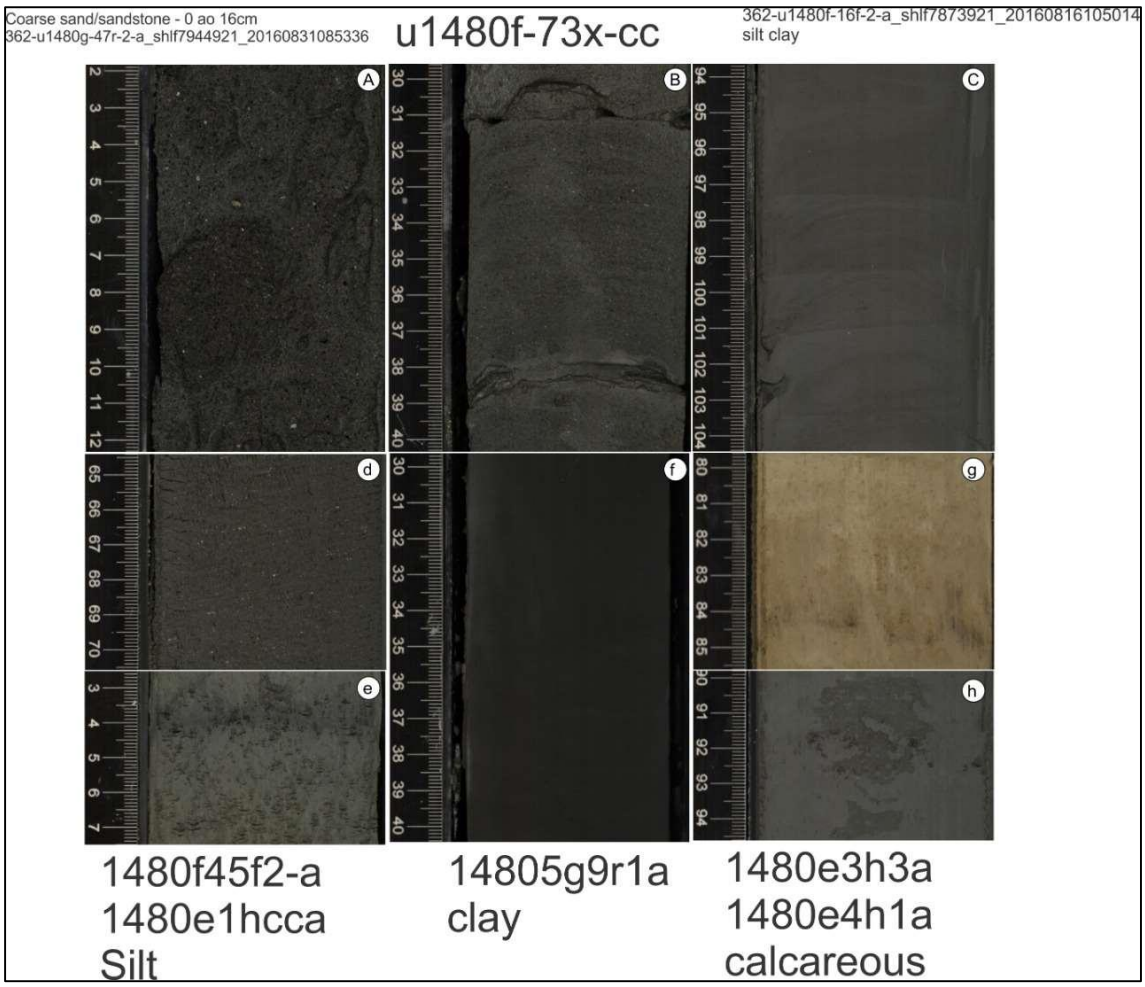


Figure 25 The core imagens of different lithologic groups. **A** represent the image of group GP, G1, G2 and G3, Litho1 - Coarse sand/sandstone. **B** represent the image of group GP, G1, G2 and G3, Litho2 - Alternating sand/sandstone and mud/mudstone layers. **C** represent the image of group GP – Litho5 - Silty clay/claystone, G2 – Litho3 - Silty clay/claystone and G3 – Litho3 - Silty clay/claystone. **d** and **e** represent the image of group GP – Litho5 - Silt/Siltstone, G1 – Litho2 - Silt/Siltstone, G2 – Litho3 - Silt/Siltstone and G3 - Litho4 - Silt/Siltstone. **f** represents the image of group GP – Litho5 - Clay/claystone, G1 – Litho2 - Clay/claystone, G2 – Litho3 - Clay/claystone and G3 – Litho5 - Clay/claystone. **g** and **h** represent the image of group GP – Litho6 - Calcareous clay/claystone, G1 – Litho3 - Calcareous clay/claystone, G2 – Litho4 - Calcareous clay/claystone and G3 – Litho6 - Calcareous clay/claystone.

Appendix 1C – Description of calibration hyperparameters methods

For calibration we used data from IODP-Expeditions 362 Hole U1480, E.

Description:

random-state: range 5 to 9

max-depth: range 1 to 29

n-estimators: range 100 to 1900, every 100

n-jobs: -1 (using all processors)

Accuracy: range 0 to 100

Criterion: entropy

Solver: lbfgs

Activation: Relu

MLP:

Method	**Configuration**			
MLP	**Hiperparameters**			
	Solver	Activation	random-state	Accuracy
	lbfgs	**Relu**	5	61,00
			6	54,00
			7	63,00
			8	**63,10**
			9	60,20

Random Forest:

Method	**Configuration**				
Random Forest	**Hiperparameters**				
	max-depth	n-estimators	random-state	n-jobs	Accuracy
	20	**1000**	**8**	**-1**	**0.9720388349514563**
	7	200	7	-1	0.9611650485436893
	10	500	9	-1	0.9611650485436893
	7	100	8	-1	0.9514563106796117
	8	100	8	-1	0.9514563106796117
	6	200	7	-1	0.9514563106796117
	7	200	8	-1	0.9514563106796117
	8	200	8	-1	0.9514563106796117
	9	200	8	-1	0.9514563106796117
	10	200	7	-1	0.9514563106796117
	10	200	8	-1	0.9514563106796117
	11	200	8	-1	0.9514563106796117
	12	200	8	-1	0.9514563106796117
	13	200	8	-1	0.9514563106796117
	14	200	8	-1	0.9514563106796117
	15	200	8	-1	0.9514563106796117
	16	200	8	-1	0.9514563106796117
	17	200	8	-1	0.9514563106796117
	18	200	8	-1	0.9514563106796117
	19	200	8	-1	0.9514563106796117
	20	200	8	-1	0.9514563106796117
	21	200	8	-1	0.9514563106796117
	22	200	8	-1	0.9514563106796117
	23	200	8	-1	0.9514563106796117
	24	200	8	-1	0.9514563106796117
	25	200	8	-1	0.9514563106796117
	26	200	8	-1	0.9514563106796117
	27	200	8	-1	0.9514563106796117
	28	200	8	-1	0.9514563106796117
	29	200	8	-1	0.9514563106796117
	6	300	8	-1	0.9514563106796117

	7	300	8	-1	0.9514563106796117
	8	300	8	-1	0.9514563106796117
	9	300	7	-1	0.9514563106796117
	9	300	8	-1	0.9514563106796117
	10	300	8	-1	0.9514563106796117
	11	300	8	-1	0.9514563106796117
	12	300	8	-1	0.9514563106796117
	13	300	8	-1	0.9514563106796117
	14	300	8	-1	0.9514563106796117
	15	300	8	-1	0.9514563106796117
	16	300	8	-1	0.9514563106796117
	17	300	8	-1	0.9514563106796117
	18	300	8	-1	0.9514563106796117
	19	300	8	-1	0.9514563106796117
	20	300	8	-1	0.9514563106796117
	21	300	8	-1	0.9514563106796117
	22	300	8	-1	0.9514563106796117
	23	300	8	-1	0.9514563106796117
	24	300	8	-1	0.9514563106796117
	25	300	8	-1	0.9514563106796117
	26	300	8	-1	0.9514563106796117
	27	300	8	-1	0.9514563106796117
	28	300	8	-1	0.9514563106796117
	29	300	8	-1	0.9514563106796117
	6	400	8	-1	0.9514563106796117
	7	400	8	-1	0.9514563106796117
	8	400	8	-1	0.9514563106796117
	9	400	8	-1	0.9514563106796117
	9	400	9	-1	0.9514563106796117
	10	400	8	-1	0.9514563106796117
	11	400	8	-1	0.9514563106796117
	12	400	8	-1	0.9514563106796117
	13	400	8	-1	0.9514563106796117
	14	400	8	-1	0.9514563106796117
	15	400	8	-1	0.9514563106796117
	16	400	8	-1	0.9514563106796117
	17	400	8	-1	0.9514563106796117

	18	400	8	-1	0.9514563106796117
	19	400	8	-1	0.9514563106796117
	20	400	8	-1	0.9514563106796117
	21	400	8	-1	0.9514563106796117
	22	400	8	-1	0.9514563106796117
	23	400	8	-1	0.9514563106796117
	24	400	8	-1	0.9514563106796117
	25	400	8	-1	0.9514563106796117
	26	400	8	-1	0.9514563106796117
	27	400	8	-1	0.9514563106796117
	28	400	8	-1	0.9514563106796117
	29	400	8	-1	0.9514563106796117
	5	500	8	-1	0.9514563106796117
	6	500	8	-1	0.9514563106796117
	7	500	8	-1	0.9514563106796117
	8	500	8	-1	0.9514563106796117
	9	500	9	-1	0.9514563106796117
	10	500	8	-1	0.9514563106796117
	11	500	9	-1	0.9514563106796117
	12	500	9	-1	0.9514563106796117
	13	500	9	-1	0.9514563106796117
	14	500	9	-1	0.9514563106796117
	15	500	9	-1	0.9514563106796117
	16	500	9	-1	0.9514563106796117
	17	500	9	-1	0.9514563106796117
	18	500	9	-1	0.9514563106796117
	19	500	9	-1	0.9514563106796117
	20	500	9	-1	0.9514563106796117
	21	500	9	-1	0.9514563106796117
	22	500	9	-1	0.9514563106796117
	23	500	9	-1	0.9514563106796117
	24	500	9	-1	0.9514563106796117
	25	500	9	-1	0.9514563106796117
	26	500	9	-1	0.9514563106796117
	27	500	9	-1	0.9514563106796117
	28	500	9	-1	0.9514563106796117
	29	500	9	-1	0.9514563106796117

	5	600	8	-1	0.9514563106796117
	6	600	8	-1	0.9514563106796117
	7	600	8	-1	0.9514563106796117
	8	600	8	-1	0.9514563106796117
	9	600	9	-1	0.9514563106796117
	10	600	9	-1	0.9514563106796117
	11	600	9	-1	0.9514563106796117
	12	600	9	-1	0.9514563106796117
	13	600	9	-1	0.9514563106796117
	14	600	9	-1	0.9514563106796117
	15	600	9	-1	0.9514563106796117
	16	600	9	-1	0.9514563106796117
	17	600	9	-1	0.9514563106796117
	18	600	9	-1	0.9514563106796117
	19	600	9	-1	0.9514563106796117
	20	600	9	-1	0.9514563106796117
	21	600	9	-1	0.9514563106796117
	22	600	9	-1	0.9514563106796117
	23	600	9	-1	0.9514563106796117
	24	600	9	-1	0.9514563106796117
	25	600	9	-1	0.9514563106796117
	26	600	9	-1	0.9514563106796117
	27	600	9	-1	0.9514563106796117
	28	600	9	-1	0.9514563106796117
	29	600	9	-1	0.9514563106796117
	5	700	8	-1	0.9514563106796117
	6	700	8	-1	0.9514563106796117
	7	700	8	-1	0.9514563106796117
	8	700	8	-1	0.9514563106796117
	7	800	8	-1	0.9514563106796117
	8	800	8	-1	0.9514563106796117
	10	800	9	-1	0.9514563106796117
	11	800	9	-1	0.9514563106796117
	12	800	9	-1	0.9514563106796117
	13	800	9	-1	0.9514563106796117
	14	800	9	-1	0.9514563106796117
	15	800	9	-1	0.9514563106796117

	16	800	9	-1	0.9514563106796117
	17	800	9	-1	0.9514563106796117
	18	800	9	-1	0.9514563106796117
	19	800	9	-1	0.9514563106796117
	20	800	9	-1	0.9514563106796117
	21	800	9	-1	0.9514563106796117
	22	800	9	-1	0.9514563106796117
	23	800	9	-1	0.9514563106796117
	24	800	9	-1	0.9514563106796117
	25	800	9	-1	0.9514563106796117
	26	800	9	-1	0.9514563106796117
	27	800	9	-1	0.9514563106796117
	28	800	9	-1	0.9514563106796117
	29	800	9	-1	0.9514563106796117
	7	900	8	-1	0.9514563106796117
	8	900	8	-1	0.9514563106796117
	10	900	9	-1	0.9514563106796117
	11	900	9	-1	0.9514563106796117
	7	1000	8	-1	0.9514563106796117
	5	1100	8	-1	0.9514563106796117
	7	1100	8	-1	0.9514563106796117
	8	1100	8	-1	0.9514563106796117
	12	1100	9	-1	0.9514563106796117
	13	1100	9	-1	0.9514563106796117
	14	1100	9	-1	0.9514563106796117
	15	1100	9	-1	0.9514563106796117
	16	1100	9	-1	0.9514563106796117
	17	1100	9	-1	0.9514563106796117
	18	1100	9	-1	0.9514563106796117
	19	1100	9	-1	0.9514563106796117
	20	1100	9	-1	0.9514563106796117
	21	1100	9	-1	0.9514563106796117
	22	1100	9	-1	0.9514563106796117
	23	1100	9	-1	0.9514563106796117
	24	1100	9	-1	0.9514563106796117
	25	1100	9	-1	0.9514563106796117
	26	1100	9	-1	0.9514563106796117

	27	1100	9	-1	0.9514563106796117
	28	1100	9	-1	0.9514563106796117
	29	1100	9	-1	0.9514563106796117
	7	1200	8	-1	0.9514563106796117
	8	1200	8	-1	0.9514563106796117
	12	1200	9	-1	0.9514563106796117
	13	1200	9	-1	0.9514563106796117
	14	1200	9	-1	0.9514563106796117
	15	1200	9	-1	0.9514563106796117
	16	1200	9	-1	0.9514563106796117
	17	1200	9	-1	0.9514563106796117
	18	1200	9	-1	0.9514563106796117
	19	1200	9	-1	0.9514563106796117
	20	1200	9	-1	0.9514563106796117
	21	1200	9	-1	0.9514563106796117
	22	1200	9	-1	0.9514563106796117
	23	1200	9	-1	0.9514563106796117
	24	1200	9	-1	0.9514563106796117
	25	1200	9	-1	0.9514563106796117
	26	1200	9	-1	0.9514563106796117
	27	1200	9	-1	0.9514563106796117
	28	1200	9	-1	0.9514563106796117
	29	1200	9	-1	0.9514563106796117
	6	1300	8	-1	0.9514563106796117
	7	1300	8	-1	0.9514563106796117
	8	1300	8	-1	0.9514563106796117
	5	1400	8	-1	0.9514563106796117
	6	1400	8	-1	0.9514563106796117
	7	1400	8	-1	0.9514563106796117
	8	1400	8	-1	0.9514563106796117
	9	1400	8	-1	0.9514563106796117
	10	1400	8	-1	0.9514563106796117
	11	1400	8	-1	0.9514563106796117
	12	1400	8	-1	0.9514563106796117
	12	1400	9	-1	0.9514563106796117
	13	1400	8	-1	0.9514563106796117
	13	1400	9	-1	0.9514563106796117

	14	1400	8	-1	0.9514563106796117
	14	1400	9	-1	0.9514563106796117
	15	1400	8	-1	0.9514563106796117
	15	1400	9	-1	0.9514563106796117
	16	1400	8	-1	0.9514563106796117
	16	1400	9	-1	0.9514563106796117
	17	1400	8	-1	0.9514563106796117
	17	1400	9	-1	0.9514563106796117
	18	1400	8	-1	0.9514563106796117
	18	1400	9	-1	0.9514563106796117
	19	1400	8	-1	0.9514563106796117
	19	1400	9	-1	0.9514563106796117
	20	1400	8	-1	0.9514563106796117
	20	1400	9	-1	0.9514563106796117
	21	1400	8	-1	0.9514563106796117
	21	1400	9	-1	0.9514563106796117
	22	1400	8	-1	0.9514563106796117
	22	1400	9	-1	0.9514563106796117
	23	1400	8	-1	0.9514563106796117
	23	1400	9	-1	0.9514563106796117
	24	1400	8	-1	0.9514563106796117
	24	1400	9	-1	0.9514563106796117
	25	1400	8	-1	0.9514563106796117
	25	1400	9	-1	0.9514563106796117
	26	1400	8	-1	0.9514563106796117
	26	1400	9	-1	0.9514563106796117
	27	1400	8	-1	0.9514563106796117
	27	1400	9	-1	0.9514563106796117
	28	1400	8	-1	0.9514563106796117
	28	1400	9	-1	0.9514563106796117
	29	1400	8	-1	0.9514563106796117
	29	1400	9	-1	0.9514563106796117
	5	1500	8	-1	0.9514563106796117
	6	1500	8	-1	0.9514563106796117
	8	1500	8	-1	0.9514563106796117
	12	1500	9	-1	0.9514563106796117
	13	1500	9	-1	0.9514563106796117

	14	1500	9	-1	0.9514563106796117
	15	1500	9	-1	0.9514563106796117
	16	1500	9	-1	0.9514563106796117
	17	1500	9	-1	0.9514563106796117
	18	1500	9	-1	0.9514563106796117
	19	1500	9	-1	0.9514563106796117
	20	1500	9	-1	0.9514563106796117
	21	1500	9	-1	0.9514563106796117
	22	1500	9	-1	0.9514563106796117
	23	1500	9	-1	0.9514563106796117
	24	1500	9	-1	0.9514563106796117
	25	1500	9	-1	0.9514563106796117
	26	1500	9	-1	0.9514563106796117
	27	1500	9	-1	0.9514563106796117
	28	1500	9	-1	0.9514563106796117
	29	1500	9	-1	0.9514563106796117
	8	1600	8	-1	0.9514563106796117
	12	1700	9	-1	0.9514563106796117
	13	1700	9	-1	0.9514563106796117
	14	1700	9	-1	0.9514563106796117
	15	1700	9	-1	0.9514563106796117
	16	1700	9	-1	0.9514563106796117
	17	1700	9	-1	0.9514563106796117
	18	1700	9	-1	0.9514563106796117
	19	1700	9	-1	0.9514563106796117
	20	1700	9	-1	0.9514563106796117
	21	1700	9	-1	0.9514563106796117
	22	1700	9	-1	0.9514563106796117
	23	1700	9	-1	0.9514563106796117
	24	1700	9	-1	0.9514563106796117
	25	1700	9	-1	0.9514563106796117
	26	1700	9	-1	0.9514563106796117
	27	1700	9	-1	0.9514563106796117
	28	1700	9	-1	0.9514563106796117
	29	1700	9	-1	0.9514563106796117
	6	1800	8	-1	0.9514563106796117
	12	1800	9	-1	0.9514563106796117

	13	1800	9	-1	0.9514563106796117
	14	1800	9	-1	0.9514563106796117
	15	1800	9	-1	0.9514563106796117
	16	1800	9	-1	0.9514563106796117
	17	1800	9	-1	0.9514563106796117
	18	1800	9	-1	0.9514563106796117
	19	1800	9	-1	0.9514563106796117
	20	1800	9	-1	0.9514563106796117
	21	1800	9	-1	0.9514563106796117
	22	1800	9	-1	0.9514563106796117
	23	1800	9	-1	0.9514563106796117
	24	1800	9	-1	0.9514563106796117
	25	1800	9	-1	0.9514563106796117
	26	1800	9	-1	0.9514563106796117
	27	1800	9	-1	0.9514563106796117
	28	1800	9	-1	0.9514563106796117
	29	1800	9	-1	0.9514563106796117
	6	1900	8	-1	0.9514563106796117
	12	1900	9	-1	0.9514563106796117
	13	1900	9	-1	0.9514563106796117
	14	1900	9	-1	0.9514563106796117
	15	1900	9	-1	0.9514563106796117
	16	1900	9	-1	0.9514563106796117
	17	1900	9	-1	0.9514563106796117
	18	1900	9	-1	0.9514563106796117
	19	1900	9	-1	0.9514563106796117
	20	1900	9	-1	0.9514563106796117
	21	1900	9	-1	0.9514563106796117
	22	1900	9	-1	0.9514563106796117
	23	1900	9	-1	0.9514563106796117
	24	1900	9	-1	0.9514563106796117
	25	1900	9	-1	0.9514563106796117
	26	1900	9	-1	0.9514563106796117
	27	1900	9	-1	0.9514563106796117
	28	1900	9	-1	0.9514563106796117
	29	1900	9	-1	0.9514563106796117
	9	100	8	-1	0.941747572815534

	10	100	8	-1	0.941747572815534
	11	100	8	-1	0.941747572815534
	12	100	8	-1	0.941747572815534
	13	100	8	-1	0.941747572815534
	14	100	8	-1	0.941747572815534
	15	100	8	-1	0.941747572815534
	16	100	8	-1	0.941747572815534
	17	100	8	-1	0.941747572815534
	18	100	8	-1	0.941747572815534
	19	100	8	-1	0.941747572815534
	20	100	8	-1	0.941747572815534
	21	100	8	-1	0.941747572815534
	22	100	8	-1	0.941747572815534
	23	100	8	-1	0.941747572815534
	24	100	8	-1	0.941747572815534
	25	100	8	-1	0.941747572815534
	26	100	8	-1	0.941747572815534
	27	100	8	-1	0.941747572815534
	28	100	8	-1	0.941747572815534
	29	100	8	-1	0.941747572815534
	6	200	8	-1	0.941747572815534
	8	200	7	-1	0.941747572815534
	9	200	5	-1	0.941747572815534
	9	200	6	-1	0.941747572815534
	9	200	7	-1	0.941747572815534
	9	200	9	-1	0.941747572815534
	10	200	5	-1	0.941747572815534
	11	200	5	-1	0.941747572815534
	11	200	7	-1	0.941747572815534
	12	200	5	-1	0.941747572815534
	12	200	7	-1	0.941747572815534
	13	200	5	-1	0.941747572815534
	13	200	7	-1	0.941747572815534
	14	200	5	-1	0.941747572815534
	14	200	7	-1	0.941747572815534
	15	200	5	-1	0.941747572815534
	15	200	7	-1	0.941747572815534

	16	200	5	-1	0.941747572815534
	16	200	7	-1	0.941747572815534
	17	200	5	-1	0.941747572815534
	17	200	7	-1	0.941747572815534
	18	200	5	-1	0.941747572815534
	18	200	7	-1	0.941747572815534
	19	200	5	-1	0.941747572815534
	19	200	7	-1	0.941747572815534
	20	200	5	-1	0.941747572815534
	20	200	7	-1	0.941747572815534
	21	200	5	-1	0.941747572815534
	21	200	7	-1	0.941747572815534
	22	200	5	-1	0.941747572815534
	22	200	7	-1	0.941747572815534
	23	200	5	-1	0.941747572815534
	23	200	7	-1	0.941747572815534
	24	200	5	-1	0.941747572815534
	24	200	7	-1	0.941747572815534
	25	200	5	-1	0.941747572815534
	25	200	7	-1	0.941747572815534
	26	200	5	-1	0.941747572815534
	26	200	7	-1	0.941747572815534
	27	200	5	-1	0.941747572815534
	27	200	7	-1	0.941747572815534
	28	200	5	-1	0.941747572815534
	28	200	7	-1	0.941747572815534
	29	200	5	-1	0.941747572815534
	29	200	7	-1	0.941747572815534
	5	300	8	-1	0.941747572815534
	6	300	7	-1	0.941747572815534
	7	300	7	-1	0.941747572815534
	8	300	9	-1	0.941747572815534
	9	300	9	-1	0.941747572815534
	10	300	5	-1	0.941747572815534
	10	300	9	-1	0.941747572815534
	11	300	5	-1	0.941747572815534
	11	300	9	-1	0.941747572815534

	12	300	5	-1	0.941747572815534
	12	300	9	-1	0.941747572815534
	13	300	5	-1	0.941747572815534
	13	300	9	-1	0.941747572815534
	14	300	5	-1	0.941747572815534
	14	300	9	-1	0.941747572815534
	15	300	5	-1	0.941747572815534
	15	300	9	-1	0.941747572815534
	16	300	5	-1	0.941747572815534
	16	300	9	-1	0.941747572815534
	17	300	5	-1	0.941747572815534
	17	300	9	-1	0.941747572815534
	18	300	5	-1	0.941747572815534
	18	300	9	-1	0.941747572815534
	19	300	5	-1	0.941747572815534
	19	300	9	-1	0.941747572815534
	20	300	5	-1	0.941747572815534
	20	300	9	-1	0.941747572815534
	21	300	5	-1	0.941747572815534
	21	300	9	-1	0.941747572815534
	22	300	5	-1	0.941747572815534
	22	300	9	-1	0.941747572815534
	23	300	5	-1	0.941747572815534
	23	300	9	-1	0.941747572815534
	24	300	5	-1	0.941747572815534
	24	300	9	-1	0.941747572815534
	25	300	5	-1	0.941747572815534
	25	300	9	-1	0.941747572815534
	26	300	5	-1	0.941747572815534
	26	300	9	-1	0.941747572815534
	27	300	5	-1	0.941747572815534
	27	300	9	-1	0.941747572815534
	28	300	5	-1	0.941747572815534
	28	300	9	-1	0.941747572815534
	29	300	5	-1	0.941747572815534
	29	300	9	-1	0.941747572815534
	5	400	8	-1	0.941747572815534

	5	400	9	-1	0.941747572815534
	7	400	7	-1	0.941747572815534
	8	400	9	-1	0.941747572815534
	10	400	9	-1	0.941747572815534
	11	400	5	-1	0.941747572815534
	11	400	9	-1	0.941747572815534
	12	400	9	-1	0.941747572815534
	13	400	5	-1	0.941747572815534
	13	400	9	-1	0.941747572815534
	14	400	5	-1	0.941747572815534
	14	400	9	-1	0.941747572815534
	15	400	5	-1	0.941747572815534
	15	400	9	-1	0.941747572815534
	16	400	5	-1	0.941747572815534
	16	400	9	-1	0.941747572815534
	17	400	5	-1	0.941747572815534
	17	400	9	-1	0.941747572815534
	18	400	5	-1	0.941747572815534
	18	400	9	-1	0.941747572815534
	19	400	5	-1	0.941747572815534
	19	400	9	-1	0.941747572815534
	20	400	5	-1	0.941747572815534
	20	400	9	-1	0.941747572815534
	21	400	5	-1	0.941747572815534
	21	400	9	-1	0.941747572815534
	22	400	5	-1	0.941747572815534
	22	400	9	-1	0.941747572815534
	23	400	5	-1	0.941747572815534
	23	400	9	-1	0.941747572815534
	24	400	5	-1	0.941747572815534
	24	400	9	-1	0.941747572815534
	25	400	5	-1	0.941747572815534
	25	400	9	-1	0.941747572815534
	26	400	5	-1	0.941747572815534
	26	400	9	-1	0.941747572815534
	27	400	5	-1	0.941747572815534
	27	400	9	-1	0.941747572815534

	28	400	5	-1	0.941747572815534
	28	400	9	-1	0.941747572815534
	29	400	5	-1	0.941747572815534
	29	400	9	-1	0.941747572815534
	5	500	9	-1	0.941747572815534
	7	500	7	-1	0.941747572815534
	9	500	8	-1	0.941747572815534
	10	500	7	-1	0.941747572815534
	11	500	7	-1	0.941747572815534
	11	500	8	-1	0.941747572815534
	12	500	7	-1	0.941747572815534
	12	500	8	-1	0.941747572815534
	13	500	7	-1	0.941747572815534
	13	500	8	-1	0.941747572815534
	14	500	7	-1	0.941747572815534
	14	500	8	-1	0.941747572815534
	15	500	7	-1	0.941747572815534
	15	500	8	-1	0.941747572815534
	16	500	7	-1	0.941747572815534
	16	500	8	-1	0.941747572815534
	17	500	7	-1	0.941747572815534
	17	500	8	-1	0.941747572815534
	18	500	7	-1	0.941747572815534
	18	500	8	-1	0.941747572815534
	19	500	7	-1	0.941747572815534
	19	500	8	-1	0.941747572815534
	20	500	7	-1	0.941747572815534
	20	500	8	-1	0.941747572815534
	21	500	7	-1	0.941747572815534
	21	500	8	-1	0.941747572815534
	22	500	7	-1	0.941747572815534
	22	500	8	-1	0.941747572815534
	23	500	7	-1	0.941747572815534
	23	500	8	-1	0.941747572815534
	24	500	7	-1	0.941747572815534
	24	500	8	-1	0.941747572815534
	25	500	7	-1	0.941747572815534

	25	500	8	-1	0.941747572815534
	26	500	7	-1	0.941747572815534
	26	500	8	-1	0.941747572815534
	27	500	7	-1	0.941747572815534
	27	500	8	-1	0.941747572815534
	28	500	7	-1	0.941747572815534
	28	500	8	-1	0.941747572815534
	29	500	7	-1	0.941747572815534
	29	500	8	-1	0.941747572815534
	5	600	7	-1	0.941747572815534
	5	600	9	-1	0.941747572815534
	6	600	7	-1	0.941747572815534
	7	600	7	-1	0.941747572815534
	8	600	7	-1	0.941747572815534
	9	600	7	-1	0.941747572815534
	9	600	8	-1	0.941747572815534
	10	600	7	-1	0.941747572815534
	10	600	8	-1	0.941747572815534
	11	600	7	-1	0.941747572815534
	11	600	8	-1	0.941747572815534
	12	600	7	-1	0.941747572815534
	12	600	8	-1	0.941747572815534
	13	600	7	-1	0.941747572815534
	13	600	8	-1	0.941747572815534
	14	600	7	-1	0.941747572815534
	14	600	8	-1	0.941747572815534
	15	600	7	-1	0.941747572815534
	15	600	8	-1	0.941747572815534
	16	600	7	-1	0.941747572815534
	16	600	8	-1	0.941747572815534
	17	600	7	-1	0.941747572815534
	17	600	8	-1	0.941747572815534
	18	600	7	-1	0.941747572815534
	18	600	8	-1	0.941747572815534
	19	600	7	-1	0.941747572815534
	19	600	8	-1	0.941747572815534
	20	600	7	-1	0.941747572815534

	20	600	8	-1	0.941747572815534
	21	600	7	-1	0.941747572815534
	21	600	8	-1	0.941747572815534
	22	600	7	-1	0.941747572815534
	22	600	8	-1	0.941747572815534
	23	600	7	-1	0.941747572815534
	23	600	8	-1	0.941747572815534
	24	600	7	-1	0.941747572815534
	24	600	8	-1	0.941747572815534
	25	600	7	-1	0.941747572815534
	25	600	8	-1	0.941747572815534
	26	600	7	-1	0.941747572815534
	26	600	8	-1	0.941747572815534
	27	600	7	-1	0.941747572815534
	27	600	8	-1	0.941747572815534
	28	600	7	-1	0.941747572815534
	28	600	8	-1	0.941747572815534
	29	600	7	-1	0.941747572815534
	29	600	8	-1	0.941747572815534
	5	700	7	-1	0.941747572815534
	5	700	9	-1	0.941747572815534
	6	700	7	-1	0.941747572815534
	7	700	7	-1	0.941747572815534
	8	700	7	-1	0.941747572815534
	9	700	7	-1	0.941747572815534
	9	700	8	-1	0.941747572815534
	9	700	9	-1	0.941747572815534
	10	700	7	-1	0.941747572815534
	10	700	8	-1	0.941747572815534
	10	700	9	-1	0.941747572815534
	11	700	7	-1	0.941747572815534
	11	700	8	-1	0.941747572815534
	11	700	9	-1	0.941747572815534
	12	700	7	-1	0.941747572815534
	12	700	8	-1	0.941747572815534
	12	700	9	-1	0.941747572815534
	13	700	7	-1	0.941747572815534

	13	700	8	-1	0.941747572815534
	13	700	9	-1	0.941747572815534
	14	700	7	-1	0.941747572815534
	14	700	8	-1	0.941747572815534
	14	700	9	-1	0.941747572815534
	15	700	7	-1	0.941747572815534
	15	700	8	-1	0.941747572815534
	15	700	9	-1	0.941747572815534
	16	700	7	-1	0.941747572815534
	16	700	8	-1	0.941747572815534
	16	700	9	-1	0.941747572815534
	17	700	7	-1	0.941747572815534
	17	700	8	-1	0.941747572815534
	17	700	9	-1	0.941747572815534
	18	700	7	-1	0.941747572815534
	18	700	8	-1	0.941747572815534
	18	700	9	-1	0.941747572815534
	19	700	7	-1	0.941747572815534
	19	700	8	-1	0.941747572815534
	19	700	9	-1	0.941747572815534
	20	700	7	-1	0.941747572815534
	20	700	8	-1	0.941747572815534
	20	700	9	-1	0.941747572815534
	21	700	7	-1	0.941747572815534
	21	700	8	-1	0.941747572815534
	21	700	9	-1	0.941747572815534
	22	700	7	-1	0.941747572815534
	22	700	8	-1	0.941747572815534
	22	700	9	-1	0.941747572815534
	23	700	7	-1	0.941747572815534
	23	700	8	-1	0.941747572815534
	23	700	9	-1	0.941747572815534
	24	700	7	-1	0.941747572815534
	24	700	8	-1	0.941747572815534
	24	700	9	-1	0.941747572815534
	25	700	7	-1	0.941747572815534
	25	700	8	-1	0.941747572815534

	25	700	9	-1	0.941747572815534
	26	700	7	-1	0.941747572815534
	26	700	8	-1	0.941747572815534
	26	700	9	-1	0.941747572815534
	27	700	7	-1	0.941747572815534
	27	700	8	-1	0.941747572815534
	27	700	9	-1	0.941747572815534
	28	700	7	-1	0.941747572815534
	28	700	8	-1	0.941747572815534
	28	700	9	-1	0.941747572815534
	29	700	7	-1	0.941747572815534
	29	700	8	-1	0.941747572815534
	29	700	9	-1	0.941747572815534
	5	800	8	-1	0.941747572815534
	5	800	9	-1	0.941747572815534
	6	800	7	-1	0.941747572815534
	6	800	8	-1	0.941747572815534
	7	800	7	-1	0.941747572815534
	7	800	9	-1	0.941747572815534
	8	800	7	-1	0.941747572815534
	9	800	7	-1	0.941747572815534
	9	800	8	-1	0.941747572815534
	9	800	9	-1	0.941747572815534
	10	800	7	-1	0.941747572815534
	10	800	8	-1	0.941747572815534
	11	800	7	-1	0.941747572815534
	11	800	8	-1	0.941747572815534
	12	800	7	-1	0.941747572815534
	12	800	8	-1	0.941747572815534
	13	800	7	-1	0.941747572815534
	13	800	8	-1	0.941747572815534
	14	800	7	-1	0.941747572815534
	14	800	8	-1	0.941747572815534
	15	800	7	-1	0.941747572815534
	15	800	8	-1	0.941747572815534
	16	800	7	-1	0.941747572815534
	16	800	8	-1	0.941747572815534

	17	800	7	-1	0.941747572815534
	17	800	8	-1	0.941747572815534
	18	800	7	-1	0.941747572815534
	18	800	8	-1	0.941747572815534
	19	800	7	-1	0.941747572815534
	19	800	8	-1	0.941747572815534
	20	800	7	-1	0.941747572815534
	20	800	8	-1	0.941747572815534
	21	800	7	-1	0.941747572815534
	21	800	8	-1	0.941747572815534
	22	800	7	-1	0.941747572815534
	22	800	8	-1	0.941747572815534
	23	800	7	-1	0.941747572815534
	23	800	8	-1	0.941747572815534
	24	800	7	-1	0.941747572815534
	24	800	8	-1	0.941747572815534
	25	800	7	-1	0.941747572815534
	25	800	8	-1	0.941747572815534
	26	800	7	-1	0.941747572815534
	26	800	8	-1	0.941747572815534
	27	800	7	-1	0.941747572815534
	27	800	8	-1	0.941747572815534
	28	800	7	-1	0.941747572815534
	28	800	8	-1	0.941747572815534
	29	800	7	-1	0.941747572815534
	29	800	8	-1	0.941747572815534
	5	900	8	-1	0.941747572815534
	6	900	7	-1	0.941747572815534
	6	900	8	-1	0.941747572815534
	7	900	7	-1	0.941747572815534
	9	900	7	-1	0.941747572815534
	9	900	8	-1	0.941747572815534
	9	900	9	-1	0.941747572815534
	10	900	7	-1	0.941747572815534
	10	900	8	-1	0.941747572815534
	11	900	7	-1	0.941747572815534
	11	900	8	-1	0.941747572815534

	12	900	7	-1	0.941747572815534
	12	900	9	-1	0.941747572815534
	13	900	7	-1	0.941747572815534
	13	900	8	-1	0.941747572815534
	13	900	9	-1	0.941747572815534
	14	900	7	-1	0.941747572815534
	14	900	8	-1	0.941747572815534
	14	900	9	-1	0.941747572815534
	15	900	7	-1	0.941747572815534
	15	900	8	-1	0.941747572815534
	15	900	9	-1	0.941747572815534
	16	900	7	-1	0.941747572815534
	16	900	8	-1	0.941747572815534
	16	900	9	-1	0.941747572815534
	17	900	7	-1	0.941747572815534
	17	900	8	-1	0.941747572815534
	17	900	9	-1	0.941747572815534
	18	900	7	-1	0.941747572815534
	18	900	8	-1	0.941747572815534
	18	900	9	-1	0.941747572815534
	19	900	7	-1	0.941747572815534
	19	900	8	-1	0.941747572815534
	19	900	9	-1	0.941747572815534
	20	900	7	-1	0.941747572815534
	20	900	8	-1	0.941747572815534
	20	900	9	-1	0.941747572815534
	21	900	7	-1	0.941747572815534
	21	900	8	-1	0.941747572815534
	21	900	9	-1	0.941747572815534
	22	900	7	-1	0.941747572815534
	22	900	8	-1	0.941747572815534
	22	900	9	-1	0.941747572815534
	23	900	7	-1	0.941747572815534
	23	900	8	-1	0.941747572815534
	23	900	9	-1	0.941747572815534
	24	900	7	-1	0.941747572815534
	24	900	8	-1	0.941747572815534

	24	900	9	-1	0.941747572815534
	25	900	7	-1	0.941747572815534
	25	900	8	-1	0.941747572815534
	25	900	9	-1	0.941747572815534
	26	900	7	-1	0.941747572815534
	26	900	8	-1	0.941747572815534
	26	900	9	-1	0.941747572815534
	27	900	7	-1	0.941747572815534
	27	900	8	-1	0.941747572815534
	27	900	9	-1	0.941747572815534
	28	900	7	-1	0.941747572815534
	28	900	8	-1	0.941747572815534
	28	900	9	-1	0.941747572815534
	29	900	7	-1	0.941747572815534
	29	900	8	-1	0.941747572815534
	29	900	9	-1	0.941747572815534
	5	1000	8	-1	0.941747572815534
	6	1000	7	-1	0.941747572815534
	6	1000	8	-1	0.941747572815534
	7	1000	7	-1	0.941747572815534
	8	1000	7	-1	0.941747572815534
	8	1000	8	-1	0.941747572815534
	9	1000	7	-1	0.941747572815534
	9	1000	8	-1	0.941747572815534
	9	1000	9	-1	0.941747572815534
	10	1000	7	-1	0.941747572815534
	10	1000	9	-1	0.941747572815534
	11	1000	7	-1	0.941747572815534
	11	1000	9	-1	0.941747572815534
	12	1000	7	-1	0.941747572815534
	12	1000	9	-1	0.941747572815534
	13	1000	7	-1	0.941747572815534
	13	1000	9	-1	0.941747572815534
	14	1000	7	-1	0.941747572815534
	14	1000	9	-1	0.941747572815534
	15	1000	7	-1	0.941747572815534
	15	1000	9	-1	0.941747572815534

	16	1000	7	-1	0.941747572815534
	16	1000	9	-1	0.941747572815534
	17	1000	7	-1	0.941747572815534
	17	1000	9	-1	0.941747572815534
	18	1000	7	-1	0.941747572815534
	18	1000	9	-1	0.941747572815534
	19	1000	7	-1	0.941747572815534
	19	1000	9	-1	0.941747572815534
	20	1000	7	-1	0.941747572815534
	20	1000	9	-1	0.941747572815534
	21	1000	7	-1	0.941747572815534
	21	1000	9	-1	0.941747572815534
	22	1000	7	-1	0.941747572815534
	22	1000	9	-1	0.941747572815534
	23	1000	7	-1	0.941747572815534
	23	1000	9	-1	0.941747572815534
	24	1000	7	-1	0.941747572815534
	24	1000	9	-1	0.941747572815534
	25	1000	7	-1	0.941747572815534
	25	1000	9	-1	0.941747572815534
	26	1000	7	-1	0.941747572815534
	26	1000	9	-1	0.941747572815534
	27	1000	7	-1	0.941747572815534
	27	1000	9	-1	0.941747572815534
	28	1000	7	-1	0.941747572815534
	28	1000	9	-1	0.941747572815534
	29	1000	7	-1	0.941747572815534
	29	1000	9	-1	0.941747572815534
	4	1100	9	-1	0.941747572815534
	6	1100	7	-1	0.941747572815534
	6	1100	8	-1	0.941747572815534
	7	1100	7	-1	0.941747572815534
	9	1100	7	-1	0.941747572815534
	9	1100	8	-1	0.941747572815534
	9	1100	9	-1	0.941747572815534
	10	1100	8	-1	0.941747572815534
	10	1100	9	-1	0.941747572815534

	11	1100	8	-1	0.941747572815534
	11	1100	9	-1	0.941747572815534
	13	1100	8	-1	0.941747572815534
	14	1100	8	-1	0.941747572815534
	15	1100	8	-1	0.941747572815534
	16	1100	8	-1	0.941747572815534
	17	1100	8	-1	0.941747572815534
	18	1100	8	-1	0.941747572815534
	19	1100	8	-1	0.941747572815534
	20	1100	8	-1	0.941747572815534
	21	1100	8	-1	0.941747572815534
	22	1100	8	-1	0.941747572815534
	23	1100	8	-1	0.941747572815534
	24	1100	8	-1	0.941747572815534
	25	1100	8	-1	0.941747572815534
	26	1100	8	-1	0.941747572815534
	27	1100	8	-1	0.941747572815534
	28	1100	8	-1	0.941747572815534
	29	1100	8	-1	0.941747572815534
	5	1200	8	-1	0.941747572815534
	6	1200	7	-1	0.941747572815534
	6	1200	8	-1	0.941747572815534
	7	1200	7	-1	0.941747572815534
	9	1200	9	-1	0.941747572815534
	10	1200	9	-1	0.941747572815534
	11	1200	9	-1	0.941747572815534
	5	1300	8	-1	0.941747572815534
	6	1300	7	-1	0.941747572815534
	7	1300	7	-1	0.941747572815534
	9	1300	8	-1	0.941747572815534
	9	1300	9	-1	0.941747572815534
	10	1300	9	-1	0.941747572815534
	11	1300	9	-1	0.941747572815534
	12	1300	9	-1	0.941747572815534
	13	1300	9	-1	0.941747572815534
	14	1300	9	-1	0.941747572815534
	15	1300	9	-1	0.941747572815534

	16	1300	9	-1	0.941747572815534
	17	1300	9	-1	0.941747572815534
	18	1300	9	-1	0.941747572815534
	19	1300	9	-1	0.941747572815534
	20	1300	9	-1	0.941747572815534
	21	1300	9	-1	0.941747572815534
	22	1300	9	-1	0.941747572815534
	23	1300	9	-1	0.941747572815534
	24	1300	9	-1	0.941747572815534
	25	1300	9	-1	0.941747572815534
	26	1300	9	-1	0.941747572815534
	27	1300	9	-1	0.941747572815534
	28	1300	9	-1	0.941747572815534
	29	1300	9	-1	0.941747572815534
	6	1400	7	-1	0.941747572815534
	7	1400	7	-1	0.941747572815534
	9	1400	9	-1	0.941747572815534
	10	1400	9	-1	0.941747572815534
	11	1400	9	-1	0.941747572815534
	6	1500	7	-1	0.941747572815534
	7	1500	7	-1	0.941747572815534
	7	1500	8	-1	0.941747572815534
	8	1500	7	-1	0.941747572815534
	9	1500	7	-1	0.941747572815534
	9	1500	8	-1	0.941747572815534
	9	1500	9	-1	0.941747572815534
	10	1500	9	-1	0.941747572815534
	11	1500	9	-1	0.941747572815534
	13	1500	8	-1	0.941747572815534
	5	1600	8	-1	0.941747572815534
	6	1600	7	-1	0.941747572815534
	6	1600	8	-1	0.941747572815534
	7	1600	7	-1	0.941747572815534
	7	1600	8	-1	0.941747572815534
	9	1600	8	-1	0.941747572815534
	9	1600	9	-1	0.941747572815534
	10	1600	8	-1	0.941747572815534

	10	1600	9	-1	0.941747572815534
	11	1600	8	-1	0.941747572815534
	11	1600	9	-1	0.941747572815534
	12	1600	8	-1	0.941747572815534
	13	1600	8	-1	0.941747572815534
	13	1600	9	-1	0.941747572815534
	14	1600	8	-1	0.941747572815534
	14	1600	9	-1	0.941747572815534
	15	1600	8	-1	0.941747572815534
	15	1600	9	-1	0.941747572815534
	16	1600	8	-1	0.941747572815534
	16	1600	9	-1	0.941747572815534
	17	1600	8	-1	0.941747572815534
	17	1600	9	-1	0.941747572815534
	18	1600	8	-1	0.941747572815534
	18	1600	9	-1	0.941747572815534
	19	1600	8	-1	0.941747572815534
	19	1600	9	-1	0.941747572815534
	20	1600	8	-1	0.941747572815534
	20	1600	9	-1	0.941747572815534
	21	1600	8	-1	0.941747572815534
	21	1600	9	-1	0.941747572815534
	22	1600	8	-1	0.941747572815534
	22	1600	9	-1	0.941747572815534
	23	1600	8	-1	0.941747572815534
	23	1600	9	-1	0.941747572815534
	24	1600	8	-1	0.941747572815534
	24	1600	9	-1	0.941747572815534
	25	1600	8	-1	0.941747572815534
	25	1600	9	-1	0.941747572815534
	26	1600	8	-1	0.941747572815534
	26	1600	9	-1	0.941747572815534
	27	1600	8	-1	0.941747572815534
	27	1600	9	-1	0.941747572815534
	28	1600	8	-1	0.941747572815534
	28	1600	9	-1	0.941747572815534
	29	1600	8	-1	0.941747572815534

	29	1600	9	-1	0.941747572815534
	5	1700	8	-1	0.941747572815534
	6	1700	7	-1	0.941747572815534
	6	1700	8	-1	0.941747572815534
	7	1700	8	-1	0.941747572815534
	8	1700	8	-1	0.941747572815534
	9	1700	9	-1	0.941747572815534
	10	1700	8	-1	0.941747572815534
	10	1700	9	-1	0.941747572815534
	11	1700	9	-1	0.941747572815534
	12	1700	8	-1	0.941747572815534
	13	1700	8	-1	0.941747572815534
	14	1700	8	-1	0.941747572815534
	15	1700	8	-1	0.941747572815534
	16	1700	8	-1	0.941747572815534
	17	1700	8	-1	0.941747572815534
	18	1700	8	-1	0.941747572815534
	19	1700	8	-1	0.941747572815534
	20	1700	8	-1	0.941747572815534
	21	1700	8	-1	0.941747572815534
	22	1700	8	-1	0.941747572815534
	23	1700	8	-1	0.941747572815534
	24	1700	8	-1	0.941747572815534
	25	1700	8	-1	0.941747572815534
	26	1700	8	-1	0.941747572815534
	27	1700	8	-1	0.941747572815534
	28	1700	8	-1	0.941747572815534
	29	1700	8	-1	0.941747572815534
	5	1800	8	-1	0.941747572815534
	6	1800	7	-1	0.941747572815534
	7	1800	8	-1	0.941747572815534
	8	1800	8	-1	0.941747572815534
	9	1800	9	-1	0.941747572815534
	10	1800	8	-1	0.941747572815534
	10	1800	9	-1	0.941747572815534
	11	1800	9	-1	0.941747572815534
	13	1800	8	-1	0.941747572815534

	5	1900	8	-1	0.941747572815534
	5	1900	9	-1	0.941747572815534
	6	1900	7	-1	0.941747572815534
	7	1900	8	-1	0.941747572815534
	8	1900	8	-1	0.941747572815534
	9	1900	9	-1	0.941747572815534
	10	1900	9	-1	0.941747572815534
	11	1900	9	-1	0.941747572815534
	2	100	9	-1	0.9320388349514563
	5	100	8	-1	0.9320388349514563
	5	100	9	-1	0.9320388349514563
	6	100	7	-1	0.9320388349514563
	6	100	8	-1	0.9320388349514563
	6	100	9	-1	0.9320388349514563
	7	100	7	-1	0.9320388349514563
	7	100	9	-1	0.9320388349514563
	8	100	5	-1	0.9320388349514563
	8	100	9	-1	0.9320388349514563
	9	100	5	-1	0.9320388349514563
	9	100	6	-1	0.9320388349514563
	9	100	9	-1	0.9320388349514563
	10	100	5	-1	0.9320388349514563
	10	100	9	-1	0.9320388349514563
	11	100	9	-1	0.9320388349514563
	12	100	9	-1	0.9320388349514563
	13	100	5	-1	0.9320388349514563
	13	100	9	-1	0.9320388349514563
	14	100	5	-1	0.9320388349514563
	14	100	9	-1	0.9320388349514563
	15	100	5	-1	0.9320388349514563
	15	100	9	-1	0.9320388349514563
	16	100	5	-1	0.9320388349514563
	16	100	9	-1	0.9320388349514563
	17	100	5	-1	0.9320388349514563
	17	100	9	-1	0.9320388349514563
	18	100	5	-1	0.9320388349514563
	18	100	9	-1	0.9320388349514563

	19	100	5	-1	0.9320388349514563
	19	100	9	-1	0.9320388349514563
	20	100	5	-1	0.9320388349514563
	20	100	9	-1	0.9320388349514563
	21	100	5	-1	0.9320388349514563
	21	100	9	-1	0.9320388349514563
	22	100	5	-1	0.9320388349514563
	22	100	9	-1	0.9320388349514563
	23	100	5	-1	0.9320388349514563
	23	100	9	-1	0.9320388349514563
	24	100	5	-1	0.9320388349514563
	24	100	9	-1	0.9320388349514563
	25	100	5	-1	0.9320388349514563
	25	100	9	-1	0.9320388349514563
	26	100	5	-1	0.9320388349514563
	26	100	9	-1	0.9320388349514563
	27	100	5	-1	0.9320388349514563
	27	100	9	-1	0.9320388349514563
	28	100	5	-1	0.9320388349514563
	28	100	9	-1	0.9320388349514563
	29	100	5	-1	0.9320388349514563
	29	100	9	-1	0.9320388349514563
	5	200	6	-1	0.9320388349514563
	5	200	8	-1	0.9320388349514563
	5	200	9	-1	0.9320388349514563
	6	200	9	-1	0.9320388349514563
	7	200	9	-1	0.9320388349514563
	8	200	5	-1	0.9320388349514563
	8	200	9	-1	0.9320388349514563
	10	200	6	-1	0.9320388349514563
	10	200	9	-1	0.9320388349514563
	11	200	6	-1	0.9320388349514563
	11	200	9	-1	0.9320388349514563
	12	200	6	-1	0.9320388349514563
	12	200	9	-1	0.9320388349514563
	13	200	6	-1	0.9320388349514563
	13	200	9	-1	0.9320388349514563

	14	200	6	-1	0.9320388349514563
	14	200	9	-1	0.9320388349514563
	15	200	6	-1	0.9320388349514563
	15	200	9	-1	0.9320388349514563
	16	200	6	-1	0.9320388349514563
	16	200	9	-1	0.9320388349514563
	17	200	6	-1	0.9320388349514563
	17	200	9	-1	0.9320388349514563
	18	200	6	-1	0.9320388349514563
	18	200	9	-1	0.9320388349514563
	19	200	6	-1	0.9320388349514563
	19	200	9	-1	0.9320388349514563
	20	200	6	-1	0.9320388349514563
	20	200	9	-1	0.9320388349514563
	21	200	6	-1	0.9320388349514563
	21	200	9	-1	0.9320388349514563
	22	200	6	-1	0.9320388349514563
	22	200	9	-1	0.9320388349514563
	23	200	6	-1	0.9320388349514563
	23	200	9	-1	0.9320388349514563
	24	200	6	-1	0.9320388349514563
	24	200	9	-1	0.9320388349514563
	25	200	6	-1	0.9320388349514563
	25	200	9	-1	0.9320388349514563
	26	200	6	-1	0.9320388349514563
	26	200	9	-1	0.9320388349514563
	27	200	6	-1	0.9320388349514563
	27	200	9	-1	0.9320388349514563
	28	200	6	-1	0.9320388349514563
	28	200	9	-1	0.9320388349514563
	29	200	6	-1	0.9320388349514563
	29	200	9	-1	0.9320388349514563
	5	300	9	-1	0.9320388349514563
	6	300	9	-1	0.9320388349514563
	7	300	9	-1	0.9320388349514563
	8	300	5	-1	0.9320388349514563
	9	300	5	-1	0.9320388349514563

	9	300	6	-1	0.9320388349514563
	10	300	7	-1	0.9320388349514563
	11	300	7	-1	0.9320388349514563
	12	300	7	-1	0.9320388349514563
	13	300	7	-1	0.9320388349514563
	14	300	7	-1	0.9320388349514563
	15	300	7	-1	0.9320388349514563
	16	300	7	-1	0.9320388349514563
	17	300	7	-1	0.9320388349514563
	18	300	7	-1	0.9320388349514563
	19	300	7	-1	0.9320388349514563
	20	300	7	-1	0.9320388349514563
	21	300	7	-1	0.9320388349514563
	22	300	7	-1	0.9320388349514563
	23	300	7	-1	0.9320388349514563
	24	300	7	-1	0.9320388349514563
	25	300	7	-1	0.9320388349514563
	26	300	7	-1	0.9320388349514563
	27	300	7	-1	0.9320388349514563
	28	300	7	-1	0.9320388349514563
	29	300	7	-1	0.9320388349514563
	4	400	8	-1	0.9320388349514563
	4	400	9	-1	0.9320388349514563
	5	400	7	-1	0.9320388349514563
	6	400	7	-1	0.9320388349514563
	6	400	9	-1	0.9320388349514563
	7	400	9	-1	0.9320388349514563
	9	400	5	-1	0.9320388349514563
	10	400	5	-1	0.9320388349514563
	10	400	7	-1	0.9320388349514563
	11	400	7	-1	0.9320388349514563
	12	400	5	-1	0.9320388349514563
	12	400	7	-1	0.9320388349514563
	13	400	7	-1	0.9320388349514563
	14	400	7	-1	0.9320388349514563
	15	400	7	-1	0.9320388349514563
	16	400	7	-1	0.9320388349514563

	17	400	7	-1	0.9320388349514563
	18	400	7	-1	0.9320388349514563
	19	400	7	-1	0.9320388349514563
	20	400	7	-1	0.9320388349514563
	21	400	7	-1	0.9320388349514563
	22	400	7	-1	0.9320388349514563
	23	400	7	-1	0.9320388349514563
	24	400	7	-1	0.9320388349514563
	25	400	7	-1	0.9320388349514563
	26	400	7	-1	0.9320388349514563
	27	400	7	-1	0.9320388349514563
	28	400	7	-1	0.9320388349514563
	29	400	7	-1	0.9320388349514563
	4	500	9	-1	0.9320388349514563
	5	500	7	-1	0.9320388349514563
	6	500	7	-1	0.9320388349514563
	6	500	9	-1	0.9320388349514563
	7	500	9	-1	0.9320388349514563
	8	500	5	-1	0.9320388349514563
	8	500	9	-1	0.9320388349514563
	9	500	5	-1	0.9320388349514563
	9	500	7	-1	0.9320388349514563
	10	500	5	-1	0.9320388349514563
	11	500	5	-1	0.9320388349514563
	12	500	5	-1	0.9320388349514563
	13	500	5	-1	0.9320388349514563
	14	500	5	-1	0.9320388349514563
	15	500	5	-1	0.9320388349514563
	16	500	5	-1	0.9320388349514563
	17	500	5	-1	0.9320388349514563
	18	500	5	-1	0.9320388349514563
	19	500	5	-1	0.9320388349514563
	20	500	5	-1	0.9320388349514563
	21	500	5	-1	0.9320388349514563
	22	500	5	-1	0.9320388349514563
	23	500	5	-1	0.9320388349514563
	24	500	5	-1	0.9320388349514563

	25	500	5	-1	0.9320388349514563
	26	500	5	-1	0.9320388349514563
	27	500	5	-1	0.9320388349514563
	28	500	5	-1	0.9320388349514563
	29	500	5	-1	0.9320388349514563
	4	600	9	-1	0.9320388349514563
	6	600	9	-1	0.9320388349514563
	7	600	9	-1	0.9320388349514563
	8	600	5	-1	0.9320388349514563
	8	600	9	-1	0.9320388349514563
	9	600	5	-1	0.9320388349514563
	10	600	5	-1	0.9320388349514563
	11	600	5	-1	0.9320388349514563
	12	600	5	-1	0.9320388349514563
	13	600	5	-1	0.9320388349514563
	14	600	5	-1	0.9320388349514563
	15	600	5	-1	0.9320388349514563
	16	600	5	-1	0.9320388349514563
	17	600	5	-1	0.9320388349514563
	18	600	5	-1	0.9320388349514563
	19	600	5	-1	0.9320388349514563
	20	600	5	-1	0.9320388349514563
	21	600	5	-1	0.9320388349514563
	22	600	5	-1	0.9320388349514563
	23	600	5	-1	0.9320388349514563
	24	600	5	-1	0.9320388349514563
	25	600	5	-1	0.9320388349514563
	26	600	5	-1	0.9320388349514563
	27	600	5	-1	0.9320388349514563
	28	600	5	-1	0.9320388349514563
	29	600	5	-1	0.9320388349514563
	4	700	9	-1	0.9320388349514563
	6	700	9	-1	0.9320388349514563
	7	700	9	-1	0.9320388349514563
	8	700	9	-1	0.9320388349514563
	10	700	5	-1	0.9320388349514563
	11	700	5	-1	0.9320388349514563

	12	700	5	-1	0.9320388349514563
	13	700	5	-1	0.9320388349514563
	14	700	5	-1	0.9320388349514563
	15	700	5	-1	0.9320388349514563
	16	700	5	-1	0.9320388349514563
	17	700	5	-1	0.9320388349514563
	18	700	5	-1	0.9320388349514563
	19	700	5	-1	0.9320388349514563
	20	700	5	-1	0.9320388349514563
	21	700	5	-1	0.9320388349514563
	22	700	5	-1	0.9320388349514563
	23	700	5	-1	0.9320388349514563
	24	700	5	-1	0.9320388349514563
	25	700	5	-1	0.9320388349514563
	26	700	5	-1	0.9320388349514563
	27	700	5	-1	0.9320388349514563
	28	700	5	-1	0.9320388349514563
	29	700	5	-1	0.9320388349514563
	4	800	9	-1	0.9320388349514563
	5	800	7	-1	0.9320388349514563
	6	800	9	-1	0.9320388349514563
	8	800	9	-1	0.9320388349514563
	11	800	5	-1	0.9320388349514563
	4	900	9	-1	0.9320388349514563
	5	900	7	-1	0.9320388349514563
	5	900	9	-1	0.9320388349514563
	6	900	5	-1	0.9320388349514563
	6	900	9	-1	0.9320388349514563
	7	900	9	-1	0.9320388349514563
	8	900	7	-1	0.9320388349514563
	8	900	9	-1	0.9320388349514563
	12	900	8	-1	0.9320388349514563
	4	1000	9	-1	0.9320388349514563
	5	1000	7	-1	0.9320388349514563
	5	1000	9	-1	0.9320388349514563
	6	1000	5	-1	0.9320388349514563
	6	1000	9	-1	0.9320388349514563

	7	1000	9	-1	0.9320388349514563
	8	1000	9	-1	0.9320388349514563
	10	1000	8	-1	0.9320388349514563
	11	1000	8	-1	0.9320388349514563
	12	1000	8	-1	0.9320388349514563
	13	1000	8	-1	0.9320388349514563
	14	1000	8	-1	0.9320388349514563
	15	1000	8	-1	0.9320388349514563
	16	1000	8	-1	0.9320388349514563
	17	1000	8	-1	0.9320388349514563
	18	1000	8	-1	0.9320388349514563
	19	1000	8	-1	0.9320388349514563
	21	1000	8	-1	0.9320388349514563
	22	1000	8	-1	0.9320388349514563
	23	1000	8	-1	0.9320388349514563
	24	1000	8	-1	0.9320388349514563
	25	1000	8	-1	0.9320388349514563
	26	1000	8	-1	0.9320388349514563
	27	1000	8	-1	0.9320388349514563
	28	1000	8	-1	0.9320388349514563
	29	1000	8	-1	0.9320388349514563
	5	1100	7	-1	0.9320388349514563
	5	1100	9	-1	0.9320388349514563
	6	1100	5	-1	0.9320388349514563
	6	1100	9	-1	0.9320388349514563
	7	1100	9	-1	0.9320388349514563
	8	1100	7	-1	0.9320388349514563
	8	1100	9	-1	0.9320388349514563
	10	1100	7	-1	0.9320388349514563
	11	1100	5	-1	0.9320388349514563
	11	1100	7	-1	0.9320388349514563
	12	1100	7	-1	0.9320388349514563
	12	1100	8	-1	0.9320388349514563
	13	1100	7	-1	0.9320388349514563
	14	1100	7	-1	0.9320388349514563
	15	1100	7	-1	0.9320388349514563
	16	1100	7	-1	0.9320388349514563

	17	1100	7	-1	0.9320388349514563
	18	1100	7	-1	0.9320388349514563
	19	1100	7	-1	0.9320388349514563
	20	1100	7	-1	0.9320388349514563
	21	1100	7	-1	0.9320388349514563
	22	1100	7	-1	0.9320388349514563
	23	1100	7	-1	0.9320388349514563
	24	1100	7	-1	0.9320388349514563
	25	1100	7	-1	0.9320388349514563
	26	1100	7	-1	0.9320388349514563
	27	1100	7	-1	0.9320388349514563
	28	1100	7	-1	0.9320388349514563
	29	1100	7	-1	0.9320388349514563
	4	1200	9	-1	0.9320388349514563
	5	1200	7	-1	0.9320388349514563
	5	1200	9	-1	0.9320388349514563
	6	1200	9	-1	0.9320388349514563
	7	1200	9	-1	0.9320388349514563
	8	1200	7	-1	0.9320388349514563
	8	1200	9	-1	0.9320388349514563
	9	1200	7	-1	0.9320388349514563
	9	1200	8	-1	0.9320388349514563
	10	1200	7	-1	0.9320388349514563
	10	1200	8	-1	0.9320388349514563
	11	1200	7	-1	0.9320388349514563
	11	1200	8	-1	0.9320388349514563
	12	1200	7	-1	0.9320388349514563
	12	1200	8	-1	0.9320388349514563
	13	1200	7	-1	0.9320388349514563
	13	1200	8	-1	0.9320388349514563
	14	1200	7	-1	0.9320388349514563
	14	1200	8	-1	0.9320388349514563
	15	1200	7	-1	0.9320388349514563
	15	1200	8	-1	0.9320388349514563
	16	1200	7	-1	0.9320388349514563
	16	1200	8	-1	0.9320388349514563
	17	1200	7	-1	0.9320388349514563

	17	1200	8	-1	0.9320388349514563
	18	1200	7	-1	0.9320388349514563
	18	1200	8	-1	0.9320388349514563
	19	1200	7	-1	0.9320388349514563
	19	1200	8	-1	0.9320388349514563
	20	1200	7	-1	0.9320388349514563
	20	1200	8	-1	0.9320388349514563
	21	1200	7	-1	0.9320388349514563
	21	1200	8	-1	0.9320388349514563
	22	1200	7	-1	0.9320388349514563
	22	1200	8	-1	0.9320388349514563
	23	1200	7	-1	0.9320388349514563
	23	1200	8	-1	0.9320388349514563
	24	1200	7	-1	0.9320388349514563
	24	1200	8	-1	0.9320388349514563
	25	1200	7	-1	0.9320388349514563
	25	1200	8	-1	0.9320388349514563
	26	1200	7	-1	0.9320388349514563
	26	1200	8	-1	0.9320388349514563
	27	1200	7	-1	0.9320388349514563
	27	1200	8	-1	0.9320388349514563
	28	1200	7	-1	0.9320388349514563
	28	1200	8	-1	0.9320388349514563
	29	1200	7	-1	0.9320388349514563
	29	1200	8	-1	0.9320388349514563
	5	1300	7	-1	0.9320388349514563
	5	1300	9	-1	0.9320388349514563
	6	1300	9	-1	0.9320388349514563
	7	1300	9	-1	0.9320388349514563
	8	1300	7	-1	0.9320388349514563
	8	1300	9	-1	0.9320388349514563
	9	1300	7	-1	0.9320388349514563
	10	1300	7	-1	0.9320388349514563
	10	1300	8	-1	0.9320388349514563
	11	1300	7	-1	0.9320388349514563
	11	1300	8	-1	0.9320388349514563
	12	1300	7	-1	0.9320388349514563

	12	1300	8	-1	0.9320388349514563
	13	1300	7	-1	0.9320388349514563
	13	1300	8	-1	0.9320388349514563
	14	1300	7	-1	0.9320388349514563
	14	1300	8	-1	0.9320388349514563
	15	1300	7	-1	0.9320388349514563
	15	1300	8	-1	0.9320388349514563
	16	1300	7	-1	0.9320388349514563
	16	1300	8	-1	0.9320388349514563
	17	1300	7	-1	0.9320388349514563
	17	1300	8	-1	0.9320388349514563
	18	1300	7	-1	0.9320388349514563
	18	1300	8	-1	0.9320388349514563
	19	1300	7	-1	0.9320388349514563
	19	1300	8	-1	0.9320388349514563
	20	1300	7	-1	0.9320388349514563
	20	1300	8	-1	0.9320388349514563
	21	1300	7	-1	0.9320388349514563
	21	1300	8	-1	0.9320388349514563
	22	1300	7	-1	0.9320388349514563
	22	1300	8	-1	0.9320388349514563
	23	1300	7	-1	0.9320388349514563
	23	1300	8	-1	0.9320388349514563
	24	1300	7	-1	0.9320388349514563
	24	1300	8	-1	0.9320388349514563
	25	1300	7	-1	0.9320388349514563
	25	1300	8	-1	0.9320388349514563
	26	1300	7	-1	0.9320388349514563
	26	1300	8	-1	0.9320388349514563
	27	1300	7	-1	0.9320388349514563
	27	1300	8	-1	0.9320388349514563
	28	1300	7	-1	0.9320388349514563
	28	1300	8	-1	0.9320388349514563
	29	1300	7	-1	0.9320388349514563
	29	1300	8	-1	0.9320388349514563
	5	1400	7	-1	0.9320388349514563
	5	1400	9	-1	0.9320388349514563

	6	1400	9	-1	0.9320388349514563
	7	1400	9	-1	0.9320388349514563
	8	1400	7	-1	0.9320388349514563
	8	1400	9	-1	0.9320388349514563
	9	1400	7	-1	0.9320388349514563
	10	1400	7	-1	0.9320388349514563
	11	1400	7	-1	0.9320388349514563
	12	1400	7	-1	0.9320388349514563
	13	1400	7	-1	0.9320388349514563
	14	1400	7	-1	0.9320388349514563
	15	1400	7	-1	0.9320388349514563
	16	1400	7	-1	0.9320388349514563
	17	1400	7	-1	0.9320388349514563
	18	1400	7	-1	0.9320388349514563
	19	1400	7	-1	0.9320388349514563
	20	1400	7	-1	0.9320388349514563
	21	1400	7	-1	0.9320388349514563
	22	1400	7	-1	0.9320388349514563
	23	1400	7	-1	0.9320388349514563
	24	1400	7	-1	0.9320388349514563
	25	1400	7	-1	0.9320388349514563
	26	1400	7	-1	0.9320388349514563
	27	1400	7	-1	0.9320388349514563
	28	1400	7	-1	0.9320388349514563
	29	1400	7	-1	0.9320388349514563
	5	1500	7	-1	0.9320388349514563
	5	1500	9	-1	0.9320388349514563
	6	1500	9	-1	0.9320388349514563
	7	1500	9	-1	0.9320388349514563
	8	1500	9	-1	0.9320388349514563
	10	1500	7	-1	0.9320388349514563
	10	1500	8	-1	0.9320388349514563
	11	1500	7	-1	0.9320388349514563
	11	1500	8	-1	0.9320388349514563
	12	1500	7	-1	0.9320388349514563
	12	1500	8	-1	0.9320388349514563
	13	1500	7	-1	0.9320388349514563

	14	1500	7	-1	0.9320388349514563
	14	1500	8	-1	0.9320388349514563
	15	1500	7	-1	0.9320388349514563
	15	1500	8	-1	0.9320388349514563
	16	1500	7	-1	0.9320388349514563
	16	1500	8	-1	0.9320388349514563
	17	1500	7	-1	0.9320388349514563
	17	1500	8	-1	0.9320388349514563
	18	1500	7	-1	0.9320388349514563
	18	1500	8	-1	0.9320388349514563
	19	1500	7	-1	0.9320388349514563
	19	1500	8	-1	0.9320388349514563
	20	1500	7	-1	0.9320388349514563
	20	1500	8	-1	0.9320388349514563
	21	1500	7	-1	0.9320388349514563
	21	1500	8	-1	0.9320388349514563
	22	1500	7	-1	0.9320388349514563
	22	1500	8	-1	0.9320388349514563
	23	1500	7	-1	0.9320388349514563
	23	1500	8	-1	0.9320388349514563
	24	1500	7	-1	0.9320388349514563
	24	1500	8	-1	0.9320388349514563
	25	1500	7	-1	0.9320388349514563
	25	1500	8	-1	0.9320388349514563
	26	1500	7	-1	0.9320388349514563
	26	1500	8	-1	0.9320388349514563
	27	1500	7	-1	0.9320388349514563
	27	1500	8	-1	0.9320388349514563
	28	1500	7	-1	0.9320388349514563
	28	1500	8	-1	0.9320388349514563
	29	1500	7	-1	0.9320388349514563
	29	1500	8	-1	0.9320388349514563
	5	1600	7	-1	0.9320388349514563
	5	1600	9	-1	0.9320388349514563
	6	1600	9	-1	0.9320388349514563
	7	1600	9	-1	0.9320388349514563
	8	1600	9	-1	0.9320388349514563

	9	1600	7	-1	0.9320388349514563
	10	1600	7	-1	0.9320388349514563
	11	1600	7	-1	0.9320388349514563
	12	1600	7	-1	0.9320388349514563
	12	1600	9	-1	0.9320388349514563
	13	1600	7	-1	0.9320388349514563
	14	1600	7	-1	0.9320388349514563
	15	1600	7	-1	0.9320388349514563
	16	1600	7	-1	0.9320388349514563
	17	1600	7	-1	0.9320388349514563
	18	1600	7	-1	0.9320388349514563
	19	1600	7	-1	0.9320388349514563
	20	1600	7	-1	0.9320388349514563
	21	1600	7	-1	0.9320388349514563
	22	1600	7	-1	0.9320388349514563
	23	1600	7	-1	0.9320388349514563
	24	1600	7	-1	0.9320388349514563
	25	1600	7	-1	0.9320388349514563
	26	1600	7	-1	0.9320388349514563
	27	1600	7	-1	0.9320388349514563
	28	1600	7	-1	0.9320388349514563
	29	1600	7	-1	0.9320388349514563
	4	1700	9	-1	0.9320388349514563
	5	1700	7	-1	0.9320388349514563
	5	1700	9	-1	0.9320388349514563
	6	1700	9	-1	0.9320388349514563
	7	1700	7	-1	0.9320388349514563
	7	1700	9	-1	0.9320388349514563
	8	1700	9	-1	0.9320388349514563
	9	1700	7	-1	0.9320388349514563
	9	1700	8	-1	0.9320388349514563
	10	1700	7	-1	0.9320388349514563
	11	1700	7	-1	0.9320388349514563
	11	1700	8	-1	0.9320388349514563
	12	1700	7	-1	0.9320388349514563
	13	1700	7	-1	0.9320388349514563
	14	1700	7	-1	0.9320388349514563

	15	1700	7	-1	0.9320388349514563
	16	1700	7	-1	0.9320388349514563
	17	1700	7	-1	0.9320388349514563
	18	1700	7	-1	0.9320388349514563
	19	1700	7	-1	0.9320388349514563
	20	1700	7	-1	0.9320388349514563
	21	1700	7	-1	0.9320388349514563
	22	1700	7	-1	0.9320388349514563
	23	1700	7	-1	0.9320388349514563
	24	1700	7	-1	0.9320388349514563
	25	1700	7	-1	0.9320388349514563
	26	1700	7	-1	0.9320388349514563
	27	1700	7	-1	0.9320388349514563
	28	1700	7	-1	0.9320388349514563
	29	1700	7	-1	0.9320388349514563
	4	1800	9	-1	0.9320388349514563
	5	1800	9	-1	0.9320388349514563
	6	1800	9	-1	0.9320388349514563
	7	1800	7	-1	0.9320388349514563
	7	1800	9	-1	0.9320388349514563
	8	1800	9	-1	0.9320388349514563
	9	1800	7	-1	0.9320388349514563
	9	1800	8	-1	0.9320388349514563
	10	1800	7	-1	0.9320388349514563
	11	1800	7	-1	0.9320388349514563
	11	1800	8	-1	0.9320388349514563
	12	1800	7	-1	0.9320388349514563
	12	1800	8	-1	0.9320388349514563
	13	1800	7	-1	0.9320388349514563
	14	1800	7	-1	0.9320388349514563
	14	1800	8	-1	0.9320388349514563
	15	1800	7	-1	0.9320388349514563
	15	1800	8	-1	0.9320388349514563
	16	1800	7	-1	0.9320388349514563
	16	1800	8	-1	0.9320388349514563
	17	1800	7	-1	0.9320388349514563
	17	1800	8	-1	0.9320388349514563

	18	1800	7	-1	0.9320388349514563
	18	1800	8	-1	0.9320388349514563
	19	1800	7	-1	0.9320388349514563
	19	1800	8	-1	0.9320388349514563
	20	1800	7	-1	0.9320388349514563
	20	1800	8	-1	0.9320388349514563
	21	1800	7	-1	0.9320388349514563
	21	1800	8	-1	0.9320388349514563
	22	1800	7	-1	0.9320388349514563
	22	1800	8	-1	0.9320388349514563
	23	1800	7	-1	0.9320388349514563
	23	1800	8	-1	0.9320388349514563
	24	1800	7	-1	0.9320388349514563
	24	1800	8	-1	0.9320388349514563
	25	1800	7	-1	0.9320388349514563
	25	1800	8	-1	0.9320388349514563
	26	1800	7	-1	0.9320388349514563
	26	1800	8	-1	0.9320388349514563
	27	1800	7	-1	0.9320388349514563
	27	1800	8	-1	0.9320388349514563
	28	1800	7	-1	0.9320388349514563
	28	1800	8	-1	0.9320388349514563
	29	1800	7	-1	0.9320388349514563
	29	1800	8	-1	0.9320388349514563
	4	1900	9	-1	0.9320388349514563
	6	1900	9	-1	0.9320388349514563
	7	1900	9	-1	0.9320388349514563
	8	1900	9	-1	0.9320388349514563
	9	1900	7	-1	0.9320388349514563
	9	1900	8	-1	0.9320388349514563
	10	1900	7	-1	0.9320388349514563
	10	1900	8	-1	0.9320388349514563
	11	1900	7	-1	0.9320388349514563
	11	1900	8	-1	0.9320388349514563
	12	1900	7	-1	0.9320388349514563
	12	1900	8	-1	0.9320388349514563
	13	1900	7	-1	0.9320388349514563

	13	1900	8	-1	0.9320388349514563
	14	1900	7	-1	0.9320388349514563
	14	1900	8	-1	0.9320388349514563
	15	1900	7	-1	0.9320388349514563
	15	1900	8	-1	0.9320388349514563
	16	1900	7	-1	0.9320388349514563
	16	1900	8	-1	0.9320388349514563
	17	1900	7	-1	0.9320388349514563
	17	1900	8	-1	0.9320388349514563
	18	1900	7	-1	0.9320388349514563
	18	1900	8	-1	0.9320388349514563
	19	1900	7	-1	0.9320388349514563
	19	1900	8	-1	0.9320388349514563
	20	1900	7	-1	0.9320388349514563
	20	1900	8	-1	0.9320388349514563
	21	1900	7	-1	0.9320388349514563
	21	1900	8	-1	0.9320388349514563
	22	1900	7	-1	0.9320388349514563
	22	1900	8	-1	0.9320388349514563
	23	1900	7	-1	0.9320388349514563
	23	1900	8	-1	0.9320388349514563
	24	1900	7	-1	0.9320388349514563
	24	1900	8	-1	0.9320388349514563
	25	1900	7	-1	0.9320388349514563
	25	1900	8	-1	0.9320388349514563
	26	1900	7	-1	0.9320388349514563
	26	1900	8	-1	0.9320388349514563
	27	1900	7	-1	0.9320388349514563
	27	1900	8	-1	0.9320388349514563
	28	1900	7	-1	0.9320388349514563
	28	1900	8	-1	0.9320388349514563
	29	1900	7	-1	0.9320388349514563
	29	1900	8	-1	0.9320388349514563
	3	100	9	-1	0.9223300970873787
	4	100	7	-1	0.9223300970873787
	4	100	8	-1	0.9223300970873787
	4	100	9	-1	0.9223300970873787

	5	100	6	-1	0.9223300970873787
	6	100	5	-1	0.9223300970873787
	7	100	5	-1	0.9223300970873787
	8	100	6	-1	0.9223300970873787
	8	100	7	-1	0.9223300970873787
	9	100	7	-1	0.9223300970873787
	10	100	6	-1	0.9223300970873787
	10	100	7	-1	0.9223300970873787
	11	100	5	-1	0.9223300970873787
	11	100	6	-1	0.9223300970873787
	11	100	7	-1	0.9223300970873787
	12	100	5	-1	0.9223300970873787
	12	100	7	-1	0.9223300970873787
	13	100	7	-1	0.9223300970873787
	14	100	7	-1	0.9223300970873787
	15	100	7	-1	0.9223300970873787
	16	100	7	-1	0.9223300970873787
	17	100	7	-1	0.9223300970873787
	18	100	7	-1	0.9223300970873787
	19	100	7	-1	0.9223300970873787
	20	100	7	-1	0.9223300970873787
	21	100	7	-1	0.9223300970873787
	22	100	7	-1	0.9223300970873787
	23	100	7	-1	0.9223300970873787
	24	100	7	-1	0.9223300970873787
	25	100	7	-1	0.9223300970873787
	26	100	7	-1	0.9223300970873787
	27	100	7	-1	0.9223300970873787
	28	100	7	-1	0.9223300970873787
	29	100	7	-1	0.9223300970873787
	3	200	9	-1	0.9223300970873787
	4	200	7	-1	0.9223300970873787
	4	200	8	-1	0.9223300970873787
	4	200	9	-1	0.9223300970873787
	5	200	7	-1	0.9223300970873787
	6	200	6	-1	0.9223300970873787
	7	200	5	-1	0.9223300970873787

	8	200	6	-1	0.9223300970873787
	3	300	9	-1	0.9223300970873787
	4	300	8	-1	0.9223300970873787
	4	300	9	-1	0.9223300970873787
	5	300	5	-1	0.9223300970873787
	5	300	6	-1	0.9223300970873787
	5	300	7	-1	0.9223300970873787
	6	300	5	-1	0.9223300970873787
	8	300	6	-1	0.9223300970873787
	8	300	7	-1	0.9223300970873787
	11	300	6	-1	0.9223300970873787
	3	400	9	-1	0.9223300970873787
	5	400	5	-1	0.9223300970873787
	5	400	6	-1	0.9223300970873787
	6	400	5	-1	0.9223300970873787
	8	400	5	-1	0.9223300970873787
	8	400	6	-1	0.9223300970873787
	8	400	7	-1	0.9223300970873787
	9	400	6	-1	0.9223300970873787
	9	400	7	-1	0.9223300970873787
	10	400	6	-1	0.9223300970873787
	11	400	6	-1	0.9223300970873787
	12	400	6	-1	0.9223300970873787
	13	400	6	-1	0.9223300970873787
	14	400	6	-1	0.9223300970873787
	15	400	6	-1	0.9223300970873787
	16	400	6	-1	0.9223300970873787
	17	400	6	-1	0.9223300970873787
	18	400	6	-1	0.9223300970873787
	19	400	6	-1	0.9223300970873787
	20	400	6	-1	0.9223300970873787
	21	400	6	-1	0.9223300970873787

	22	400	6	-1	0.9223300970873787
	23	400	6	-1	0.9223300970873787
	24	400	6	-1	0.9223300970873787
	25	400	6	-1	0.9223300970873787
	26	400	6	-1	0.9223300970873787
	27	400	6	-1	0.9223300970873787
	28	400	6	-1	0.9223300970873787
	29	400	6	-1	0.9223300970873787
	4	500	8	-1	0.9223300970873787
	5	500	6	-1	0.9223300970873787
	6	500	5	-1	0.9223300970873787
	7	500	5	-1	0.9223300970873787
	8	500	7	-1	0.9223300970873787
	3	600	9	-1	0.9223300970873787
	4	600	8	-1	0.9223300970873787
	5	600	6	-1	0.9223300970873787
	6	600	5	-1	0.9223300970873787
	7	600	5	-1	0.9223300970873787
	7	600	6	-1	0.9223300970873787
	8	600	6	-1	0.9223300970873787
	9	600	6	-1	0.9223300970873787
	10	600	6	-1	0.9223300970873787
	11	600	6	-1	0.9223300970873787
	12	600	6	-1	0.9223300970873787
	13	600	6	-1	0.9223300970873787
	14	600	6	-1	0.9223300970873787
	15	600	6	-1	0.9223300970873787
	16	600	6	-1	0.9223300970873787
	17	600	6	-1	0.9223300970873787
	18	600	6	-1	0.9223300970873787
	19	600	6	-1	0.9223300970873787
	20	600	6	-1	0.9223300970873787

	21	600	6	-1	0.9223300970873787
	22	600	6	-1	0.9223300970873787
	23	600	6	-1	0.9223300970873787
	24	600	6	-1	0.9223300970873787
	25	600	6	-1	0.9223300970873787
	26	600	6	-1	0.9223300970873787
	27	600	6	-1	0.9223300970873787
	28	600	6	-1	0.9223300970873787
	29	600	6	-1	0.9223300970873787
	4	700	8	-1	0.9223300970873787
	5	700	5	-1	0.9223300970873787
	5	700	6	-1	0.9223300970873787
	6	700	5	-1	0.9223300970873787
	7	700	5	-1	0.9223300970873787
	7	700	6	-1	0.9223300970873787
	8	700	5	-1	0.9223300970873787
	8	700	6	-1	0.9223300970873787
	9	700	5	-1	0.9223300970873787
	9	700	6	-1	0.9223300970873787
	10	700	6	-1	0.9223300970873787
	11	700	6	-1	0.9223300970873787
	12	700	6	-1	0.9223300970873787
	13	700	6	-1	0.9223300970873787
	14	700	6	-1	0.9223300970873787
	15	700	6	-1	0.9223300970873787
	16	700	6	-1	0.9223300970873787
	17	700	6	-1	0.9223300970873787
	18	700	6	-1	0.9223300970873787
	19	700	6	-1	0.9223300970873787
	20	700	6	-1	0.9223300970873787
	21	700	6	-1	0.9223300970873787
	22	700	6	-1	0.9223300970873787

	23	700	6	-1	0.9223300970873787
	24	700	6	-1	0.9223300970873787
	25	700	6	-1	0.9223300970873787
	26	700	6	-1	0.9223300970873787
	27	700	6	-1	0.9223300970873787
	28	700	6	-1	0.9223300970873787
	29	700	6	-1	0.9223300970873787
	4	800	8	-1	0.9223300970873787
	5	800	5	-1	0.9223300970873787
	6	800	5	-1	0.9223300970873787
	9	800	5	-1	0.9223300970873787
	9	800	6	-1	0.9223300970873787
	10	800	5	-1	0.9223300970873787
	10	800	6	-1	0.9223300970873787
	11	800	6	-1	0.9223300970873787
	12	800	5	-1	0.9223300970873787
	12	800	6	-1	0.9223300970873787
	13	800	5	-1	0.9223300970873787
	13	800	6	-1	0.9223300970873787
	14	800	5	-1	0.9223300970873787
	14	800	6	-1	0.9223300970873787
	15	800	5	-1	0.9223300970873787
	15	800	6	-1	0.9223300970873787
	16	800	5	-1	0.9223300970873787
	16	800	6	-1	0.9223300970873787
	17	800	5	-1	0.9223300970873787
	17	800	6	-1	0.9223300970873787
	18	800	5	-1	0.9223300970873787
	18	800	6	-1	0.9223300970873787
	19	800	5	-1	0.9223300970873787
	19	800	6	-1	0.9223300970873787
	20	800	5	-1	0.9223300970873787

	20	800	6	-1	0.9223300970873787
	21	800	5	-1	0.9223300970873787
	21	800	6	-1	0.9223300970873787
	22	800	5	-1	0.9223300970873787
	22	800	6	-1	0.9223300970873787
	23	800	5	-1	0.9223300970873787
	23	800	6	-1	0.9223300970873787
	24	800	5	-1	0.9223300970873787
	24	800	6	-1	0.9223300970873787
	25	800	5	-1	0.9223300970873787
	25	800	6	-1	0.9223300970873787
	26	800	5	-1	0.9223300970873787
	26	800	6	-1	0.9223300970873787
	27	800	5	-1	0.9223300970873787
	27	800	6	-1	0.9223300970873787
	28	800	5	-1	0.9223300970873787
	28	800	6	-1	0.9223300970873787
	29	800	5	-1	0.9223300970873787
	29	800	6	-1	0.9223300970873787
	4	900	8	-1	0.9223300970873787
	5	900	5	-1	0.9223300970873787
	7	900	6	-1	0.9223300970873787
	8	900	6	-1	0.9223300970873787
	9	900	6	-1	0.9223300970873787
	10	900	6	-1	0.9223300970873787
	11	900	5	-1	0.9223300970873787
	11	900	6	-1	0.9223300970873787
	12	900	6	-1	0.9223300970873787
	13	900	5	-1	0.9223300970873787
	13	900	6	-1	0.9223300970873787
	14	900	5	-1	0.9223300970873787
	14	900	6	-1	0.9223300970873787

	15	900	5	-1	0.9223300970873787
	15	900	6	-1	0.9223300970873787
	16	900	5	-1	0.9223300970873787
	16	900	6	-1	0.9223300970873787
	17	900	5	-1	0.9223300970873787
	17	900	6	-1	0.9223300970873787
	18	900	5	-1	0.9223300970873787
	18	900	6	-1	0.9223300970873787
	19	900	5	-1	0.9223300970873787
	19	900	6	-1	0.9223300970873787
	20	900	5	-1	0.9223300970873787
	20	900	6	-1	0.9223300970873787
	21	900	5	-1	0.9223300970873787
	21	900	6	-1	0.9223300970873787
	22	900	5	-1	0.9223300970873787
	22	900	6	-1	0.9223300970873787
	23	900	5	-1	0.9223300970873787
	23	900	6	-1	0.9223300970873787
	24	900	5	-1	0.9223300970873787
	24	900	6	-1	0.9223300970873787
	25	900	5	-1	0.9223300970873787
	25	900	6	-1	0.9223300970873787
	26	900	5	-1	0.9223300970873787
	26	900	6	-1	0.9223300970873787
	27	900	5	-1	0.9223300970873787
	27	900	6	-1	0.9223300970873787
	28	900	5	-1	0.9223300970873787
	28	900	6	-1	0.9223300970873787
	29	900	5	-1	0.9223300970873787
	29	900	6	-1	0.9223300970873787
	4	1000	8	-1	0.9223300970873787
	5	1000	5	-1	0.9223300970873787

	7	1000	6	-1	0.9223300970873787
	8	1000	6	-1	0.9223300970873787
	9	1000	6	-1	0.9223300970873787
	10	1000	6	-1	0.9223300970873787
	11	1000	5	-1	0.9223300970873787
	11	1000	6	-1	0.9223300970873787
	12	1000	6	-1	0.9223300970873787
	13	1000	5	-1	0.9223300970873787
	13	1000	6	-1	0.9223300970873787
	14	1000	5	-1	0.9223300970873787
	14	1000	6	-1	0.9223300970873787
	15	1000	5	-1	0.9223300970873787
	15	1000	6	-1	0.9223300970873787
	16	1000	5	-1	0.9223300970873787
	16	1000	6	-1	0.9223300970873787
	17	1000	5	-1	0.9223300970873787
	17	1000	6	-1	0.9223300970873787
	18	1000	5	-1	0.9223300970873787
	18	1000	6	-1	0.9223300970873787
	19	1000	5	-1	0.9223300970873787
	19	1000	6	-1	0.9223300970873787
	20	1000	5	-1	0.9223300970873787
	20	1000	6	-1	0.9223300970873787
	21	1000	5	-1	0.9223300970873787
	21	1000	6	-1	0.9223300970873787
	22	1000	5	-1	0.9223300970873787
	22	1000	6	-1	0.9223300970873787
	23	1000	5	-1	0.9223300970873787
	23	1000	6	-1	0.9223300970873787
	24	1000	5	-1	0.9223300970873787
	24	1000	6	-1	0.9223300970873787
	25	1000	5	-1	0.9223300970873787

	25	1000	6	-1	0.9223300970873787
	26	1000	5	-1	0.9223300970873787
	26	1000	6	-1	0.9223300970873787
	27	1000	5	-1	0.9223300970873787
	27	1000	6	-1	0.9223300970873787
	28	1000	5	-1	0.9223300970873787
	28	1000	6	-1	0.9223300970873787
	29	1000	5	-1	0.9223300970873787
	29	1000	6	-1	0.9223300970873787
	4	1100	8	-1	0.9223300970873787
	7	1100	5	-1	0.9223300970873787
	7	1100	6	-1	0.9223300970873787
	8	1100	6	-1	0.9223300970873787
	9	1100	6	-1	0.9223300970873787
	10	1100	6	-1	0.9223300970873787
	11	1100	6	-1	0.9223300970873787
	12	1100	6	-1	0.9223300970873787
	13	1100	6	-1	0.9223300970873787
	14	1100	6	-1	0.9223300970873787
	15	1100	6	-1	0.9223300970873787
	16	1100	6	-1	0.9223300970873787
	17	1100	6	-1	0.9223300970873787
	18	1100	6	-1	0.9223300970873787
	19	1100	6	-1	0.9223300970873787
	20	1100	6	-1	0.9223300970873787
	21	1100	6	-1	0.9223300970873787
	22	1100	6	-1	0.9223300970873787
	23	1100	6	-1	0.9223300970873787
	24	1100	6	-1	0.9223300970873787
	25	1100	6	-1	0.9223300970873787
	26	1100	6	-1	0.9223300970873787
	27	1100	6	-1	0.9223300970873787

	28	1100	6	-1	0.9223300970873787
	29	1100	6	-1	0.9223300970873787
	4	1200	8	-1	0.9223300970873787
	6	1200	5	-1	0.9223300970873787
	7	1200	6	-1	0.9223300970873787
	8	1200	6	-1	0.9223300970873787
	9	1200	6	-1	0.9223300970873787
	10	1200	6	-1	0.9223300970873787
	11	1200	6	-1	0.9223300970873787
	12	1200	6	-1	0.9223300970873787
	13	1200	6	-1	0.9223300970873787
	14	1200	6	-1	0.9223300970873787
	15	1200	6	-1	0.9223300970873787
	16	1200	6	-1	0.9223300970873787
	17	1200	6	-1	0.9223300970873787
	18	1200	6	-1	0.9223300970873787
	19	1200	6	-1	0.9223300970873787
	20	1200	6	-1	0.9223300970873787
	21	1200	6	-1	0.9223300970873787
	22	1200	6	-1	0.9223300970873787
	23	1200	6	-1	0.9223300970873787
	24	1200	6	-1	0.9223300970873787
	25	1200	6	-1	0.9223300970873787
	26	1200	6	-1	0.9223300970873787
	27	1200	6	-1	0.9223300970873787
	28	1200	6	-1	0.9223300970873787
	29	1200	6	-1	0.9223300970873787
	4	1300	8	-1	0.9223300970873787
	4	1300	9	-1	0.9223300970873787
	7	1300	6	-1	0.9223300970873787
	8	1300	6	-1	0.9223300970873787
	9	1300	6	-1	0.9223300970873787

	10	1300	6	-1	0.9223300970873787
	11	1300	6	-1	0.9223300970873787
	12	1300	6	-1	0.9223300970873787
	13	1300	6	-1	0.9223300970873787
	14	1300	6	-1	0.9223300970873787
	15	1300	6	-1	0.9223300970873787
	16	1300	6	-1	0.9223300970873787
	17	1300	6	-1	0.9223300970873787
	18	1300	6	-1	0.9223300970873787
	19	1300	6	-1	0.9223300970873787
	20	1300	6	-1	0.9223300970873787
	21	1300	6	-1	0.9223300970873787
	22	1300	6	-1	0.9223300970873787
	23	1300	6	-1	0.9223300970873787
	24	1300	6	-1	0.9223300970873787
	25	1300	6	-1	0.9223300970873787
	26	1300	6	-1	0.9223300970873787
	27	1300	6	-1	0.9223300970873787
	28	1300	6	-1	0.9223300970873787
	29	1300	6	-1	0.9223300970873787
	4	1400	8	-1	0.9223300970873787
	4	1400	9	-1	0.9223300970873787
	7	1400	6	-1	0.9223300970873787
	8	1400	6	-1	0.9223300970873787
	9	1400	6	-1	0.9223300970873787
	10	1400	6	-1	0.9223300970873787
	11	1400	6	-1	0.9223300970873787
	12	1400	6	-1	0.9223300970873787
	13	1400	6	-1	0.9223300970873787
	14	1400	6	-1	0.9223300970873787
	15	1400	6	-1	0.9223300970873787
	16	1400	6	-1	0.9223300970873787

	17	1400	6	-1	0.9223300970873787
	18	1400	6	-1	0.9223300970873787
	19	1400	6	-1	0.9223300970873787
	20	1400	6	-1	0.9223300970873787
	21	1400	6	-1	0.9223300970873787
	22	1400	6	-1	0.9223300970873787
	23	1400	6	-1	0.9223300970873787
	24	1400	6	-1	0.9223300970873787
	25	1400	6	-1	0.9223300970873787
	26	1400	6	-1	0.9223300970873787
	27	1400	6	-1	0.9223300970873787
	28	1400	6	-1	0.9223300970873787
	29	1400	6	-1	0.9223300970873787
	4	1500	8	-1	0.9223300970873787
	4	1500	9	-1	0.9223300970873787
	7	1500	6	-1	0.9223300970873787
	8	1500	6	-1	0.9223300970873787
	9	1500	6	-1	0.9223300970873787
	10	1500	6	-1	0.9223300970873787
	11	1500	6	-1	0.9223300970873787
	12	1500	6	-1	0.9223300970873787
	13	1500	6	-1	0.9223300970873787
	14	1500	6	-1	0.9223300970873787
	15	1500	6	-1	0.9223300970873787
	16	1500	6	-1	0.9223300970873787
	17	1500	6	-1	0.9223300970873787
	18	1500	6	-1	0.9223300970873787
	19	1500	6	-1	0.9223300970873787
	20	1500	6	-1	0.9223300970873787
	21	1500	6	-1	0.9223300970873787
	22	1500	6	-1	0.9223300970873787
	23	1500	6	-1	0.9223300970873787

	24	1500	6	-1	0.9223300970873787
	25	1500	6	-1	0.9223300970873787
	26	1500	6	-1	0.9223300970873787
	27	1500	6	-1	0.9223300970873787
	28	1500	6	-1	0.9223300970873787
	29	1500	6	-1	0.9223300970873787
	4	1600	8	-1	0.9223300970873787
	4	1600	9	-1	0.9223300970873787
	6	1600	5	-1	0.9223300970873787
	7	1600	6	-1	0.9223300970873787
	8	1600	6	-1	0.9223300970873787
	8	1600	7	-1	0.9223300970873787
	9	1600	6	-1	0.9223300970873787
	10	1600	6	-1	0.9223300970873787
	11	1600	6	-1	0.9223300970873787
	12	1600	6	-1	0.9223300970873787
	13	1600	6	-1	0.9223300970873787
	14	1600	6	-1	0.9223300970873787
	15	1600	6	-1	0.9223300970873787
	16	1600	6	-1	0.9223300970873787
	17	1600	6	-1	0.9223300970873787
	18	1600	6	-1	0.9223300970873787
	19	1600	6	-1	0.9223300970873787
	20	1600	6	-1	0.9223300970873787
	21	1600	6	-1	0.9223300970873787
	22	1600	6	-1	0.9223300970873787
	23	1600	6	-1	0.9223300970873787
	24	1600	6	-1	0.9223300970873787
	25	1600	6	-1	0.9223300970873787
	26	1600	6	-1	0.9223300970873787
	27	1600	6	-1	0.9223300970873787
	28	1600	6	-1	0.9223300970873787

	29	1600	6	-1	0.9223300970873787
	4	1700	8	-1	0.9223300970873787
	6	1700	5	-1	0.9223300970873787
	7	1700	6	-1	0.9223300970873787
	8	1700	6	-1	0.9223300970873787
	8	1700	7	-1	0.9223300970873787
	9	1700	6	-1	0.9223300970873787
	10	1700	6	-1	0.9223300970873787
	11	1700	6	-1	0.9223300970873787
	12	1700	6	-1	0.9223300970873787
	13	1700	6	-1	0.9223300970873787
	14	1700	6	-1	0.9223300970873787
	15	1700	6	-1	0.9223300970873787
	16	1700	6	-1	0.9223300970873787
	17	1700	6	-1	0.9223300970873787
	18	1700	6	-1	0.9223300970873787
	19	1700	6	-1	0.9223300970873787
	20	1700	6	-1	0.9223300970873787
	21	1700	6	-1	0.9223300970873787
	22	1700	6	-1	0.9223300970873787
	23	1700	6	-1	0.9223300970873787
	24	1700	6	-1	0.9223300970873787
	25	1700	6	-1	0.9223300970873787
	26	1700	6	-1	0.9223300970873787
	27	1700	6	-1	0.9223300970873787
	28	1700	6	-1	0.9223300970873787
	29	1700	6	-1	0.9223300970873787
	4	1800	8	-1	0.9223300970873787
	5	1800	7	-1	0.9223300970873787
	7	1800	6	-1	0.9223300970873787
	8	1800	6	-1	0.9223300970873787
	8	1800	7	-1	0.9223300970873787

	9	1800	6	-1	0.9223300970873787
	10	1800	6	-1	0.9223300970873787
	11	1800	6	-1	0.9223300970873787
	12	1800	6	-1	0.9223300970873787
	13	1800	6	-1	0.9223300970873787
	14	1800	6	-1	0.9223300970873787
	15	1800	6	-1	0.9223300970873787
	16	1800	6	-1	0.9223300970873787
	17	1800	6	-1	0.9223300970873787
	18	1800	6	-1	0.9223300970873787
	19	1800	6	-1	0.9223300970873787
	20	1800	6	-1	0.9223300970873787
	21	1800	6	-1	0.9223300970873787
	22	1800	6	-1	0.9223300970873787
	23	1800	6	-1	0.9223300970873787
	24	1800	6	-1	0.9223300970873787
	25	1800	6	-1	0.9223300970873787
	26	1800	6	-1	0.9223300970873787
	27	1800	6	-1	0.9223300970873787
	28	1800	6	-1	0.9223300970873787
	29	1800	6	-1	0.9223300970873787
	4	1900	8	-1	0.9223300970873787
	5	1900	7	-1	0.9223300970873787
	7	1900	6	-1	0.9223300970873787
	7	1900	7	-1	0.9223300970873787
	8	1900	6	-1	0.9223300970873787
	8	1900	7	-1	0.9223300970873787
	9	1900	6	-1	0.9223300970873787
	10	1900	6	-1	0.9223300970873787
	11	1900	6	-1	0.9223300970873787
	12	1900	6	-1	0.9223300970873787
	13	1900	6	-1	0.9223300970873787

	14	1900	6	-1	0.9223300970873787
	15	1900	6	-1	0.9223300970873787
	16	1900	6	-1	0.9223300970873787
	17	1900	6	-1	0.9223300970873787
	18	1900	6	-1	0.9223300970873787
	19	1900	6	-1	0.9223300970873787
	20	1900	6	-1	0.9223300970873787
	21	1900	6	-1	0.9223300970873787
	22	1900	6	-1	0.9223300970873787
	23	1900	6	-1	0.9223300970873787
	24	1900	6	-1	0.9223300970873787
	25	1900	6	-1	0.9223300970873787
	26	1900	6	-1	0.9223300970873787
	27	1900	6	-1	0.9223300970873787
	28	1900	6	-1	0.9223300970873787
	29	1900	6	-1	0.9223300970873787
	2	100	8	-1	0.912621359223301
	3	100	8	-1	0.912621359223301
	4	100	6	-1	0.912621359223301
	5	100	5	-1	0.912621359223301
	5	100	7	-1	0.912621359223301
	6	100	6	-1	0.912621359223301
	12	100	6	-1	0.912621359223301
	13	100	6	-1	0.912621359223301
	14	100	6	-1	0.912621359223301
	15	100	6	-1	0.912621359223301
	16	100	6	-1	0.912621359223301
	17	100	6	-1	0.912621359223301
	18	100	6	-1	0.912621359223301
	19	100	6	-1	0.912621359223301
	20	100	6	-1	0.912621359223301
	21	100	6	-1	0.912621359223301

	22	100	6	-1	0.912621359223301
	23	100	6	-1	0.912621359223301
	24	100	6	-1	0.912621359223301
	25	100	6	-1	0.912621359223301
	26	100	6	-1	0.912621359223301
	27	100	6	-1	0.912621359223301
	28	100	6	-1	0.912621359223301
	29	100	6	-1	0.912621359223301
	2	200	9	-1	0.912621359223301
	4	200	6	-1	0.912621359223301
	5	200	5	-1	0.912621359223301
	6	200	5	-1	0.912621359223301
	3	300	8	-1	0.912621359223301
	4	300	6	-1	0.912621359223301
	7	300	5	-1	0.912621359223301
	3	400	8	-1	0.912621359223301
	4	400	7	-1	0.912621359223301
	7	400	5	-1	0.912621359223301
	3	500	8	-1	0.912621359223301
	3	500	9	-1	0.912621359223301
	4	500	7	-1	0.912621359223301
	5	500	5	-1	0.912621359223301
	6	500	6	-1	0.912621359223301
	7	500	6	-1	0.912621359223301
	8	500	6	-1	0.912621359223301
	9	500	6	-1	0.912621359223301
	10	500	6	-1	0.912621359223301
	11	500	6	-1	0.912621359223301
	12	500	6	-1	0.912621359223301
	13	500	6	-1	0.912621359223301
	14	500	6	-1	0.912621359223301
	15	500	6	-1	0.912621359223301

	16	500	6	-1	0.912621359223301
	17	500	6	-1	0.912621359223301
	18	500	6	-1	0.912621359223301
	19	500	6	-1	0.912621359223301
	20	500	6	-1	0.912621359223301
	21	500	6	-1	0.912621359223301
	22	500	6	-1	0.912621359223301
	23	500	6	-1	0.912621359223301
	24	500	6	-1	0.912621359223301
	25	500	6	-1	0.912621359223301
	26	500	6	-1	0.912621359223301
	27	500	6	-1	0.912621359223301
	28	500	6	-1	0.912621359223301
	29	500	6	-1	0.912621359223301
	5	600	5	-1	0.912621359223301
	6	600	6	-1	0.912621359223301
	3	700	8	-1	0.912621359223301
	3	700	9	-1	0.912621359223301
	4	700	7	-1	0.912621359223301
	6	700	6	-1	0.912621359223301
	3	800	8	-1	0.912621359223301
	4	800	7	-1	0.912621359223301
	5	800	6	-1	0.912621359223301
	6	800	6	-1	0.912621359223301
	7	800	5	-1	0.912621359223301
	8	800	5	-1	0.912621359223301
	8	800	6	-1	0.912621359223301
	3	900	8	-1	0.912621359223301
	5	900	6	-1	0.912621359223301
	6	900	6	-1	0.912621359223301
	9	900	5	-1	0.912621359223301
	12	900	5	-1	0.912621359223301

	3	1000	8	-1	0.912621359223301
	5	1000	6	-1	0.912621359223301
	6	1000	6	-1	0.912621359223301
	7	1000	5	-1	0.912621359223301
	8	1000	5	-1	0.912621359223301
	9	1000	5	-1	0.912621359223301
	10	1000	5	-1	0.912621359223301
	12	1000	5	-1	0.912621359223301
	3	1100	8	-1	0.912621359223301
	4	1100	7	-1	0.912621359223301
	5	1100	5	-1	0.912621359223301
	5	1100	6	-1	0.912621359223301
	6	1100	6	-1	0.912621359223301
	8	1100	5	-1	0.912621359223301
	9	1100	5	-1	0.912621359223301
	10	1100	5	-1	0.912621359223301
	12	1100	5	-1	0.912621359223301
	13	1100	5	-1	0.912621359223301
	14	1100	5	-1	0.912621359223301
	15	1100	5	-1	0.912621359223301
	16	1100	5	-1	0.912621359223301
	17	1100	5	-1	0.912621359223301
	18	1100	5	-1	0.912621359223301
	19	1100	5	-1	0.912621359223301
	20	1100	5	-1	0.912621359223301
	21	1100	5	-1	0.912621359223301
	22	1100	5	-1	0.912621359223301
	23	1100	5	-1	0.912621359223301
	24	1100	5	-1	0.912621359223301
	25	1100	5	-1	0.912621359223301
	26	1100	5	-1	0.912621359223301
	27	1100	5	-1	0.912621359223301

	28	1100	5	-1	0.912621359223301
	29	1100	5	-1	0.912621359223301
	3	1200	8	-1	0.912621359223301
	4	1200	7	-1	0.912621359223301
	5	1200	5	-1	0.912621359223301
	5	1200	6	-1	0.912621359223301
	6	1200	6	-1	0.912621359223301
	7	1200	5	-1	0.912621359223301
	8	1200	5	-1	0.912621359223301
	9	1200	5	-1	0.912621359223301
	10	1200	5	-1	0.912621359223301
	11	1200	5	-1	0.912621359223301
	12	1200	5	-1	0.912621359223301
	13	1200	5	-1	0.912621359223301
	14	1200	5	-1	0.912621359223301
	15	1200	5	-1	0.912621359223301
	16	1200	5	-1	0.912621359223301
	17	1200	5	-1	0.912621359223301
	18	1200	5	-1	0.912621359223301
	19	1200	5	-1	0.912621359223301
	20	1200	5	-1	0.912621359223301
	21	1200	5	-1	0.912621359223301
	22	1200	5	-1	0.912621359223301
	23	1200	5	-1	0.912621359223301
	24	1200	5	-1	0.912621359223301
	25	1200	5	-1	0.912621359223301
	26	1200	5	-1	0.912621359223301
	27	1200	5	-1	0.912621359223301
	28	1200	5	-1	0.912621359223301
	29	1200	5	-1	0.912621359223301
	3	1300	8	-1	0.912621359223301
	4	1300	7	-1	0.912621359223301

	5	1300	5	-1	0.912621359223301
	5	1300	6	-1	0.912621359223301
	6	1300	5	-1	0.912621359223301
	7	1300	5	-1	0.912621359223301
	8	1300	5	-1	0.912621359223301
	9	1300	5	-1	0.912621359223301
	10	1300	5	-1	0.912621359223301
	11	1300	5	-1	0.912621359223301
	12	1300	5	-1	0.912621359223301
	13	1300	5	-1	0.912621359223301
	14	1300	5	-1	0.912621359223301
	15	1300	5	-1	0.912621359223301
	16	1300	5	-1	0.912621359223301
	17	1300	5	-1	0.912621359223301
	18	1300	5	-1	0.912621359223301
	19	1300	5	-1	0.912621359223301
	20	1300	5	-1	0.912621359223301
	21	1300	5	-1	0.912621359223301
	22	1300	5	-1	0.912621359223301
	23	1300	5	-1	0.912621359223301
	24	1300	5	-1	0.912621359223301
	25	1300	5	-1	0.912621359223301
	26	1300	5	-1	0.912621359223301
	27	1300	5	-1	0.912621359223301
	28	1300	5	-1	0.912621359223301
	29	1300	5	-1	0.912621359223301
	3	1400	8	-1	0.912621359223301
	5	1400	5	-1	0.912621359223301
	5	1400	6	-1	0.912621359223301
	6	1400	5	-1	0.912621359223301
	6	1400	6	-1	0.912621359223301
	7	1400	5	-1	0.912621359223301

	8	1400	5	-1	0.912621359223301
	9	1400	5	-1	0.912621359223301
	10	1400	5	-1	0.912621359223301
	11	1400	5	-1	0.912621359223301
	12	1400	5	-1	0.912621359223301
	13	1400	5	-1	0.912621359223301
	14	1400	5	-1	0.912621359223301
	15	1400	5	-1	0.912621359223301
	16	1400	5	-1	0.912621359223301
	17	1400	5	-1	0.912621359223301
	18	1400	5	-1	0.912621359223301
	19	1400	5	-1	0.912621359223301
	20	1400	5	-1	0.912621359223301
	21	1400	5	-1	0.912621359223301
	22	1400	5	-1	0.912621359223301
	23	1400	5	-1	0.912621359223301
	24	1400	5	-1	0.912621359223301
	25	1400	5	-1	0.912621359223301
	26	1400	5	-1	0.912621359223301
	27	1400	5	-1	0.912621359223301
	28	1400	5	-1	0.912621359223301
	29	1400	5	-1	0.912621359223301
	3	1500	8	-1	0.912621359223301
	5	1500	5	-1	0.912621359223301
	5	1500	6	-1	0.912621359223301
	6	1500	5	-1	0.912621359223301
	6	1500	6	-1	0.912621359223301
	7	1500	5	-1	0.912621359223301
	8	1500	5	-1	0.912621359223301
	9	1500	5	-1	0.912621359223301
	10	1500	5	-1	0.912621359223301
	11	1500	5	-1	0.912621359223301

	12	1500	5	-1	0.912621359223301
	13	1500	5	-1	0.912621359223301
	14	1500	5	-1	0.912621359223301
	15	1500	5	-1	0.912621359223301
	16	1500	5	-1	0.912621359223301
	17	1500	5	-1	0.912621359223301
	18	1500	5	-1	0.912621359223301
	19	1500	5	-1	0.912621359223301
	20	1500	5	-1	0.912621359223301
	21	1500	5	-1	0.912621359223301
	22	1500	5	-1	0.912621359223301
	23	1500	5	-1	0.912621359223301
	24	1500	5	-1	0.912621359223301
	25	1500	5	-1	0.912621359223301
	26	1500	5	-1	0.912621359223301
	27	1500	5	-1	0.912621359223301
	28	1500	5	-1	0.912621359223301
	29	1500	5	-1	0.912621359223301
	3	1600	8	-1	0.912621359223301
	5	1600	5	-1	0.912621359223301
	5	1600	6	-1	0.912621359223301
	6	1600	6	-1	0.912621359223301
	7	1600	5	-1	0.912621359223301
	8	1600	5	-1	0.912621359223301
	9	1600	5	-1	0.912621359223301
	10	1600	5	-1	0.912621359223301
	11	1600	5	-1	0.912621359223301
	12	1600	5	-1	0.912621359223301
	13	1600	5	-1	0.912621359223301
	14	1600	5	-1	0.912621359223301
	15	1600	5	-1	0.912621359223301
	16	1600	5	-1	0.912621359223301

	17	1600	5	-1	0.912621359223301
	18	1600	5	-1	0.912621359223301
	19	1600	5	-1	0.912621359223301
	20	1600	5	-1	0.912621359223301
	21	1600	5	-1	0.912621359223301
	22	1600	5	-1	0.912621359223301
	23	1600	5	-1	0.912621359223301
	24	1600	5	-1	0.912621359223301
	25	1600	5	-1	0.912621359223301
	26	1600	5	-1	0.912621359223301
	27	1600	5	-1	0.912621359223301
	28	1600	5	-1	0.912621359223301
	29	1600	5	-1	0.912621359223301
	3	1700	8	-1	0.912621359223301
	5	1700	5	-1	0.912621359223301
	5	1700	6	-1	0.912621359223301
	6	1700	6	-1	0.912621359223301
	7	1700	5	-1	0.912621359223301
	8	1700	5	-1	0.912621359223301
	9	1700	5	-1	0.912621359223301
	10	1700	5	-1	0.912621359223301
	11	1700	5	-1	0.912621359223301
	12	1700	5	-1	0.912621359223301
	13	1700	5	-1	0.912621359223301
	14	1700	5	-1	0.912621359223301
	15	1700	5	-1	0.912621359223301
	16	1700	5	-1	0.912621359223301
	17	1700	5	-1	0.912621359223301
	18	1700	5	-1	0.912621359223301
	19	1700	5	-1	0.912621359223301
	20	1700	5	-1	0.912621359223301
	21	1700	5	-1	0.912621359223301

	22	1700	5	-1	0.912621359223301
	23	1700	5	-1	0.912621359223301
	24	1700	5	-1	0.912621359223301
	25	1700	5	-1	0.912621359223301
	26	1700	5	-1	0.912621359223301
	27	1700	5	-1	0.912621359223301
	28	1700	5	-1	0.912621359223301
	29	1700	5	-1	0.912621359223301
	3	1800	8	-1	0.912621359223301
	5	1800	5	-1	0.912621359223301
	5	1800	6	-1	0.912621359223301
	6	1800	5	-1	0.912621359223301
	6	1800	6	-1	0.912621359223301
	7	1800	5	-1	0.912621359223301
	8	1800	5	-1	0.912621359223301
	9	1800	5	-1	0.912621359223301
	10	1800	5	-1	0.912621359223301
	11	1800	5	-1	0.912621359223301
	12	1800	5	-1	0.912621359223301
	13	1800	5	-1	0.912621359223301
	14	1800	5	-1	0.912621359223301
	15	1800	5	-1	0.912621359223301
	16	1800	5	-1	0.912621359223301
	17	1800	5	-1	0.912621359223301
	18	1800	5	-1	0.912621359223301
	19	1800	5	-1	0.912621359223301
	20	1800	5	-1	0.912621359223301
	21	1800	5	-1	0.912621359223301
	22	1800	5	-1	0.912621359223301
	23	1800	5	-1	0.912621359223301
	24	1800	5	-1	0.912621359223301
	25	1800	5	-1	0.912621359223301

	26	1800	5	-1	0.912621359223301
	27	1800	5	-1	0.912621359223301
	28	1800	5	-1	0.912621359223301
	29	1800	5	-1	0.912621359223301
	3	1900	8	-1	0.912621359223301
	5	1900	5	-1	0.912621359223301
	5	1900	6	-1	0.912621359223301
	6	1900	5	-1	0.912621359223301
	6	1900	6	-1	0.912621359223301
	7	1900	5	-1	0.912621359223301
	8	1900	5	-1	0.912621359223301
	9	1900	5	-1	0.912621359223301
	10	1900	5	-1	0.912621359223301
	11	1900	5	-1	0.912621359223301
	12	1900	5	-1	0.912621359223301
	13	1900	5	-1	0.912621359223301
	14	1900	5	-1	0.912621359223301
	15	1900	5	-1	0.912621359223301
	16	1900	5	-1	0.912621359223301
	17	1900	5	-1	0.912621359223301
	18	1900	5	-1	0.912621359223301
	19	1900	5	-1	0.912621359223301
	20	1900	5	-1	0.912621359223301
	21	1900	5	-1	0.912621359223301
	22	1900	5	-1	0.912621359223301
	23	1900	5	-1	0.912621359223301
	24	1900	5	-1	0.912621359223301
	25	1900	5	-1	0.912621359223301
	26	1900	5	-1	0.912621359223301
	27	1900	5	-1	0.912621359223301
	28	1900	5	-1	0.912621359223301
	29	1900	5	-1	0.912621359223301

	3	100	6	-1	0.9029126213592233
	4	100	5	-1	0.9029126213592233
	7	100	6	-1	0.9029126213592233
	3	200	6	-1	0.9029126213592233
	7	200	6	-1	0.9029126213592233
	3	300	6	-1	0.9029126213592233
	4	300	5	-1	0.9029126213592233
	4	300	7	-1	0.9029126213592233
	6	300	6	-1	0.9029126213592233
	10	300	6	-1	0.9029126213592233
	3	400	6	-1	0.9029126213592233
	4	400	5	-1	0.9029126213592233
	4	400	6	-1	0.9029126213592233
	6	400	6	-1	0.9029126213592233
	4	500	5	-1	0.9029126213592233
	3	600	8	-1	0.9029126213592233
	4	600	5	-1	0.9029126213592233
	4	600	7	-1	0.9029126213592233
	3	800	9	-1	0.9029126213592233
	7	800	6	-1	0.9029126213592233
	3	900	9	-1	0.9029126213592233
	4	900	5	-1	0.9029126213592233
	4	900	7	-1	0.9029126213592233
	7	900	5	-1	0.9029126213592233
	8	900	5	-1	0.9029126213592233
	10	900	5	-1	0.9029126213592233
	3	1000	9	-1	0.9029126213592233
	4	1000	5	-1	0.9029126213592233
	4	1000	7	-1	0.9029126213592233
	3	1100	9	-1	0.9029126213592233
	4	1100	5	-1	0.9029126213592233
	3	1200	9	-1	0.9029126213592233

	3	1300	9	-1	0.9029126213592233
	6	1300	6	-1	0.9029126213592233
	3	1400	9	-1	0.9029126213592233
	4	1400	7	-1	0.9029126213592233
	3	1500	9	-1	0.9029126213592233
	4	1500	7	-1	0.9029126213592233
	3	1600	9	-1	0.9029126213592233
	4	1600	7	-1	0.9029126213592233
	3	1700	9	-1	0.9029126213592233
	4	1700	7	-1	0.9029126213592233
	3	1800	9	-1	0.9029126213592233
	3	1900	9	-1	0.9029126213592233
	4	1900	7	-1	0.9029126213592233
	2	200	5	-1	0.8932038834951457
	2	200	6	-1	0.8932038834951457
	3	200	8	-1	0.8932038834951457
	4	200	5	-1	0.8932038834951457
	2	300	6	-1	0.8932038834951457
	2	300	9	-1	0.8932038834951457
	7	300	6	-1	0.8932038834951457
	12	300	6	-1	0.8932038834951457
	13	300	6	-1	0.8932038834951457
	14	300	6	-1	0.8932038834951457
	15	300	6	-1	0.8932038834951457
	16	300	6	-1	0.8932038834951457
	17	300	6	-1	0.8932038834951457
	18	300	6	-1	0.8932038834951457
	19	300	6	-1	0.8932038834951457
	20	300	6	-1	0.8932038834951457
	21	300	6	-1	0.8932038834951457
	22	300	6	-1	0.8932038834951457
	23	300	6	-1	0.8932038834951457

	24	300	6	-1	0.8932038834951457
	25	300	6	-1	0.8932038834951457
	26	300	6	-1	0.8932038834951457
	27	300	6	-1	0.8932038834951457
	28	300	6	-1	0.8932038834951457
	29	300	6	-1	0.8932038834951457
	2	400	6	-1	0.8932038834951457
	2	400	8	-1	0.8932038834951457
	2	400	9	-1	0.8932038834951457
	3	400	7	-1	0.8932038834951457
	7	400	6	-1	0.8932038834951457
	2	500	6	-1	0.8932038834951457
	2	500	9	-1	0.8932038834951457
	3	500	6	-1	0.8932038834951457
	4	500	6	-1	0.8932038834951457
	2	600	6	-1	0.8932038834951457
	2	600	9	-1	0.8932038834951457
	3	600	6	-1	0.8932038834951457
	4	600	6	-1	0.8932038834951457
	2	700	9	-1	0.8932038834951457
	3	700	6	-1	0.8932038834951457
	3	700	7	-1	0.8932038834951457
	4	700	5	-1	0.8932038834951457
	4	700	6	-1	0.8932038834951457
	2	800	9	-1	0.8932038834951457
	3	800	6	-1	0.8932038834951457
	3	800	7	-1	0.8932038834951457
	4	800	5	-1	0.8932038834951457
	4	800	6	-1	0.8932038834951457
	2	900	9	-1	0.8932038834951457
	3	900	6	-1	0.8932038834951457
	3	900	7	-1	0.8932038834951457

	4	900	6	-1	0.8932038834951457
	2	1000	9	-1	0.8932038834951457
	3	1000	6	-1	0.8932038834951457
	3	1000	7	-1	0.8932038834951457
	4	1000	6	-1	0.8932038834951457
	2	1100	9	-1	0.8932038834951457
	3	1100	6	-1	0.8932038834951457
	3	1100	7	-1	0.8932038834951457
	4	1100	6	-1	0.8932038834951457
	2	1200	9	-1	0.8932038834951457
	3	1200	6	-1	0.8932038834951457
	3	1200	7	-1	0.8932038834951457
	4	1200	6	-1	0.8932038834951457
	2	1300	9	-1	0.8932038834951457
	3	1300	6	-1	0.8932038834951457
	3	1300	7	-1	0.8932038834951457
	4	1300	6	-1	0.8932038834951457
	2	1400	9	-1	0.8932038834951457
	3	1400	6	-1	0.8932038834951457
	4	1400	6	-1	0.8932038834951457
	2	1500	9	-1	0.8932038834951457
	3	1500	6	-1	0.8932038834951457
	4	1500	6	-1	0.8932038834951457
	2	1600	9	-1	0.8932038834951457
	3	1600	6	-1	0.8932038834951457
	4	1600	6	-1	0.8932038834951457
	2	1700	9	-1	0.8932038834951457
	3	1700	6	-1	0.8932038834951457
	2	1800	9	-1	0.8932038834951457
	3	1800	6	-1	0.8932038834951457
	4	1800	6	-1	0.8932038834951457
	4	1800	7	-1	0.8932038834951457

	2	1900	9	-1	0.8932038834951457
	3	1900	6	-1	0.8932038834951457
	3	100	5	-1	0.883495145631068
	2	200	8	-1	0.883495145631068
	3	200	7	-1	0.883495145631068
	2	300	5	-1	0.883495145631068
	2	300	8	-1	0.883495145631068
	3	400	5	-1	0.883495145631068
	2	500	5	-1	0.883495145631068
	2	500	8	-1	0.883495145631068
	2	600	5	-1	0.883495145631068
	2	600	8	-1	0.883495145631068
	2	700	5	-1	0.883495145631068
	2	700	6	-1	0.883495145631068
	2	700	8	-1	0.883495145631068
	2	800	5	-1	0.883495145631068
	2	800	6	-1	0.883495145631068
	2	800	8	-1	0.883495145631068
	2	900	5	-1	0.883495145631068
	2	900	6	-1	0.883495145631068
	2	900	8	-1	0.883495145631068
	2	1000	6	-1	0.883495145631068
	2	1000	8	-1	0.883495145631068
	3	1000	5	-1	0.883495145631068
	2	1100	6	-1	0.883495145631068
	2	1100	8	-1	0.883495145631068
	3	1100	5	-1	0.883495145631068
	2	1200	6	-1	0.883495145631068
	2	1200	8	-1	0.883495145631068
	3	1200	5	-1	0.883495145631068
	4	1200	5	-1	0.883495145631068
	2	1300	5	-1	0.883495145631068

	2	1300	6	-1	0.883495145631068
	2	1300	8	-1	0.883495145631068
	3	1300	5	-1	0.883495145631068
	4	1300	5	-1	0.883495145631068
	2	1400	5	-1	0.883495145631068
	2	1400	6	-1	0.883495145631068
	2	1400	8	-1	0.883495145631068
	3	1400	5	-1	0.883495145631068
	3	1400	7	-1	0.883495145631068
	4	1400	5	-1	0.883495145631068
	2	1500	5	-1	0.883495145631068
	2	1500	6	-1	0.883495145631068
	2	1500	8	-1	0.883495145631068
	3	1500	5	-1	0.883495145631068
	3	1500	7	-1	0.883495145631068
	4	1500	5	-1	0.883495145631068
	2	1600	5	-1	0.883495145631068
	2	1600	6	-1	0.883495145631068
	2	1600	8	-1	0.883495145631068
	3	1600	5	-1	0.883495145631068
	3	1600	7	-1	0.883495145631068
	4	1600	5	-1	0.883495145631068
	2	1700	5	-1	0.883495145631068
	2	1700	6	-1	0.883495145631068
	2	1700	8	-1	0.883495145631068
	3	1700	5	-1	0.883495145631068
	3	1700	7	-1	0.883495145631068
	4	1700	5	-1	0.883495145631068
	4	1700	6	-1	0.883495145631068
	2	1800	6	-1	0.883495145631068
	2	1800	8	-1	0.883495145631068
	3	1800	5	-1	0.883495145631068

	3	1800	7	-1	0.883495145631068
	4	1800	5	-1	0.883495145631068
	2	1900	6	-1	0.883495145631068
	2	1900	8	-1	0.883495145631068
	3	1900	5	-1	0.883495145631068
	3	1900	7	-1	0.883495145631068
	4	1900	5	-1	0.883495145631068
	4	1900	6	-1	0.883495145631068
	2	100	5	-1	0.8737864077669902
	2	100	6	-1	0.8737864077669902
	3	200	5	-1	0.8737864077669902
	3	300	5	-1	0.8737864077669902
	3	300	7	-1	0.8737864077669902
	2	400	5	-1	0.8737864077669902
	3	500	5	-1	0.8737864077669902
	3	500	7	-1	0.8737864077669902
	3	600	5	-1	0.8737864077669902
	3	600	7	-1	0.8737864077669902
	3	700	5	-1	0.8737864077669902
	3	800	5	-1	0.8737864077669902
	2	900	7	-1	0.8737864077669902
	3	900	5	-1	0.8737864077669902
	2	1000	5	-1	0.8737864077669902
	2	1000	7	-1	0.8737864077669902
	2	1100	5	-1	0.8737864077669902
	2	1100	7	-1	0.8737864077669902
	2	1200	5	-1	0.8737864077669902
	2	1200	7	-1	0.8737864077669902
	2	1300	7	-1	0.8737864077669902
	2	1400	7	-1	0.8737864077669902
	2	1500	7	-1	0.8737864077669902
	2	1600	7	-1	0.8737864077669902

	2	1700	7	-1	0.8737864077669902
	2	1800	5	-1	0.8737864077669902
	2	1900	5	-1	0.8737864077669902
	3	100	7	-1	0.8640776699029126
	2	200	7	-1	0.8640776699029126
	2	400	7	-1	0.8640776699029126
	2	500	7	-1	0.8640776699029126
	2	600	7	-1	0.8640776699029126
	2	700	7	-1	0.8640776699029126
	2	800	7	-1	0.8640776699029126
	2	300	7	-1	0.8543689320388349
	2	1800	7	-1	0.8543689320388349
	2	1900	7	-1	0.8543689320388349
	2	100	7	-1	0.8349514563106796
	1	400	5	-1	0.7475728155339806
	1	200	8	-1	0.7281553398058253
	1	400	8	-1	0.7281553398058253
	1	100	8	-1	0.7184466019417476
	1	300	8	-1	0.7184466019417476
	1	500	5	-1	0.7184466019417476
	1	500	8	-1	0.7184466019417476
	1	600	8	-1	0.7184466019417476
	1	700	8	-1	0.7184466019417476
	1	800	8	-1	0.7184466019417476
	1	1400	8	-1	0.7184466019417476
	1	600	5	-1	0.7087378640776699
	1	700	5	-1	0.7087378640776699
	1	800	5	-1	0.7087378640776699
	1	900	5	-1	0.7087378640776699
	1	900	8	-1	0.7087378640776699
	1	1000	5	-1	0.7087378640776699
	1	1000	8	-1	0.7087378640776699

	1	1100	5	-1	0.7087378640776699
	1	1100	8	-1	0.7087378640776699
	1	1200	5	-1	0.7087378640776699
	1	1200	8	-1	0.7087378640776699
	1	1300	8	-1	0.7087378640776699
	1	1500	8	-1	0.7087378640776699
	1	1600	8	-1	0.7087378640776699
	1	1700	8	-1	0.7087378640776699
	1	1800	5	-1	0.7087378640776699
	1	1800	8	-1	0.7087378640776699
	1	1900	5	-1	0.7087378640776699
	1	1900	8	-1	0.7087378640776699
	1	100	9	-1	0.6990291262135923
	1	200	7	-1	0.6990291262135923
	1	200	9	-1	0.6990291262135923
	1	300	5	-1	0.6990291262135923
	1	900	7	-1	0.6990291262135923
	1	1000	7	-1	0.6990291262135923
	1	1100	7	-1	0.6990291262135923
	1	1300	5	-1	0.6990291262135923
	1	1400	5	-1	0.6990291262135923
	1	1400	7	-1	0.6990291262135923
	1	1500	5	-1	0.6990291262135923
	1	1500	7	-1	0.6990291262135923
	1	1600	5	-1	0.6990291262135923
	1	1700	5	-1	0.6990291262135923
	1	100	5	-1	0.6893203883495146
	1	100	7	-1	0.6893203883495146
	1	200	6	-1	0.6893203883495146
	1	300	7	-1	0.6893203883495146
	1	300	9	-1	0.6893203883495146
	1	400	9	-1	0.6893203883495146

	1	500	9	-1	0.6893203883495146
	1	600	7	-1	0.6893203883495146
	1	600	9	-1	0.6893203883495146
	1	700	7	-1	0.6893203883495146
	1	700	9	-1	0.6893203883495146
	1	800	7	-1	0.6893203883495146
	1	800	9	-1	0.6893203883495146
	1	900	9	-1	0.6893203883495146
	1	1000	9	-1	0.6893203883495146
	1	1100	9	-1	0.6893203883495146
	1	1200	7	-1	0.6893203883495146
	1	1200	9	-1	0.6893203883495146
	1	1300	7	-1	0.6893203883495146
	1	1300	9	-1	0.6893203883495146
	1	1400	9	-1	0.6893203883495146
	1	1500	9	-1	0.6893203883495146
	1	1600	7	-1	0.6893203883495146
	1	1600	9	-1	0.6893203883495146
	1	1700	9	-1	0.6893203883495146
	1	1800	7	-1	0.6893203883495146
	1	1800	9	-1	0.6893203883495146
	1	1900	7	-1	0.6893203883495146
	1	1900	9	-1	0.6893203883495146
	1	200	5	-1	0.6796116504854369
	1	500	7	-1	0.6796116504854369
	1	1700	7	-1	0.6796116504854369
	1	100	6	-1	0.6699029126213593
	1	400	7	-1	0.6699029126213593
	1	300	6	-1	0.6601941747572816
	1	600	6	-1	0.6601941747572816
	1	400	6	-1	0.6504854368932039
	1	500	6	-1	0.6504854368932039

	1	700	6	-1	0.6504854368932039
	1	800	6	-1	0.6504854368932039
	1	900	6	-1	0.6504854368932039
	1	1000	6	-1	0.6504854368932039
	1	1100	6	-1	0.6504854368932039
	1	1200	6	-1	0.6504854368932039
	1	1300	6	-1	0.6504854368932039
	1	1400	6	-1	0.6504854368932039
	1	1500	6	-1	0.6504854368932039
	1	1600	6	-1	0.6504854368932039
	1	1700	6	-1	0.6504854368932039
	1	1800	6	-1	0.6504854368932039
	1	1900	6	-1	0.6504854368932039

DecisionTree:

Method	**Configuration**			
DecisionTree	**Hiperparameters**			
	max-depth	criterion	random-state	Accuracy
	20	**entropy**	**8**	**0.9655388349514563**
	5	entropy	9	0.9611650485436893
	5	entropy	8	0.9320388349514563
	6	entropy	8	0.9320388349514563
	6	entropy	9	0.9320388349514563
	7	entropy	8	0.9320388349514563
	8	entropy	8	0.9320388349514563
	8	entropy	9	0.9320388349514563

	9	entropy	8	0.9320388349514563
	9	entropy	9	0.9320388349514563
	10	entropy	8	0.9320388349514563
	10	entropy	9	0.9320388349514563
	11	entropy	8	0.9320388349514563
	11	entropy	9	0.9320388349514563
	12	entropy	8	0.9320388349514563
	12	entropy	9	0.9320388349514563
	13	entropy	8	0.9320388349514563
	13	entropy	9	0.9320388349514563
	14	entropy	8	0.9320388349514563
	14	entropy	9	0.9320388349514563
	15	entropy	8	0.9320388349514563
	15	entropy	9	0.9320388349514563
	16	entropy	8	0.9320388349514563
	16	entropy	9	0.9320388349514563
	17	entropy	8	0.9320388349514563
	17	entropy	9	0.9320388349514563
	18	entropy	8	0.9320388349514563
	18	entropy	9	0.9320388349514563
	19	entropy	8	0.9320388349514563
	19	entropy	9	0.9320388349514563
	20	entropy	9	0.9320388349514563
	21	entropy	8	0.9320388349514563
	21	entropy	9	0.9320388349514563
	22	entropy	8	0.9320388349514563
	22	entropy	9	0.9320388349514563
	23	entropy	8	0.9320388349514563
	23	entropy	9	0.9320388349514563
	24	entropy	8	0.9320388349514563
	24	entropy	9	0.9320388349514563
	25	entropy	8	0.9320388349514563

	25	entropy	9	0.9320388349514563
	26	entropy	8	0.9320388349514563
	26	entropy	9	0.9320388349514563
	27	entropy	8	0.9320388349514563
	27	entropy	9	0.9320388349514563
	28	entropy	8	0.9320388349514563
	28	entropy	9	0.9320388349514563
	29	entropy	8	0.9320388349514563
	29	entropy	9	0.9320388349514563
	2	entropy	6	0.9223300970873787
	2	entropy	8	0.9223300970873787
	2	entropy	9	0.9223300970873787
	3	entropy	6	0.9223300970873787
	3	entropy	8	0.9223300970873787
	3	entropy	9	0.9223300970873787
	4	entropy	8	0.9223300970873787
	4	entropy	9	0.9223300970873787
	5	entropy	5	0.9223300970873787
	6	entropy	5	0.9223300970873787
	7	entropy	5	0.9223300970873787
	5	entropy	6	0.912621359223301
	8	entropy	6	0.912621359223301
	9	entropy	6	0.912621359223301
	10	entropy	6	0.912621359223301
	11	entropy	6	0.912621359223301
	12	entropy	6	0.912621359223301
	13	entropy	6	0.912621359223301
	14	entropy	6	0.912621359223301
	15	entropy	6	0.912621359223301
	16	entropy	6	0.912621359223301
	17	entropy	6	0.912621359223301
	18	entropy	6	0.912621359223301

	19	entropy	6	0.912621359223301
	20	entropy	6	0.912621359223301
	21	entropy	6	0.912621359223301
	22	entropy	6	0.912621359223301
	23	entropy	6	0.912621359223301
	24	entropy	6	0.912621359223301
	25	entropy	6	0.912621359223301
	26	entropy	6	0.912621359223301
	27	entropy	6	0.912621359223301
	28	entropy	6	0.912621359223301
	29	entropy	6	0.912621359223301
	4	entropy	6	0.9029126213592233
	3	entropy	5	0.8932038834951457
	4	entropy	5	0.8932038834951457
	8	entropy	5	0.8932038834951457
	9	entropy	5	0.8932038834951457
	10	entropy	5	0.8932038834951457
	11	entropy	5	0.8932038834951457
	12	entropy	5	0.8932038834951457
	13	entropy	5	0.8932038834951457
	14	entropy	5	0.8932038834951457
	15	entropy	5	0.8932038834951457
	16	entropy	5	0.8932038834951457
	17	entropy	5	0.8932038834951457
	18	entropy	5	0.8932038834951457
	19	entropy	5	0.8932038834951457
	20	entropy	5	0.8932038834951457
	21	entropy	5	0.8932038834951457
	22	entropy	5	0.8932038834951457
	23	entropy	5	0.8932038834951457
	24	entropy	5	0.8932038834951457
	25	entropy	5	0.8932038834951457

	26	entropy	5	0.8932038834951457
	27	entropy	5	0.8932038834951457
	28	entropy	5	0.8932038834951457
	29	entropy	5	0.8932038834951457
	2	entropy	7	0.883495145631068
	7	entropy	6	0.883495145631068
	3	entropy	7	0.8737864077669902
	2	entropy	5	0.8640776699029126
	5	entropy	7	0.8640776699029126
	6	entropy	7	0.8640776699029126
	7	entropy	7	0.8640776699029126
	7	entropy	9	0.8640776699029126
	8	entropy	7	0.8640776699029126
	9	entropy	7	0.8640776699029126
	10	entropy	7	0.8640776699029126
	11	entropy	7	0.8640776699029126
	12	entropy	7	0.8640776699029126
	13	entropy	7	0.8640776699029126
	14	entropy	7	0.8640776699029126
	15	entropy	7	0.8640776699029126
	16	entropy	7	0.8640776699029126
	17	entropy	7	0.8640776699029126
	18	entropy	7	0.8640776699029126
	19	entropy	7	0.8640776699029126
	20	entropy	7	0.8640776699029126
	21	entropy	7	0.8640776699029126
	22	entropy	7	0.8640776699029126
	23	entropy	7	0.8640776699029126
	24	entropy	7	0.8640776699029126
	25	entropy	7	0.8640776699029126
	26	entropy	7	0.8640776699029126
	27	entropy	7	0.8640776699029126

	28	entropy	7	0.8640776699029126
	29	entropy	7	0.8640776699029126
	6	entropy	6	0.8543689320388349
	4	entropy	7	0.8252427184466019
	1	entropy	5	0.7864077669902912
	1	entropy	9	0.7864077669902912
	1	entropy	6	0.7766990291262136
	1	entropy	7	0.7669902912621359
	1	entropy	8	0.7572815533980582

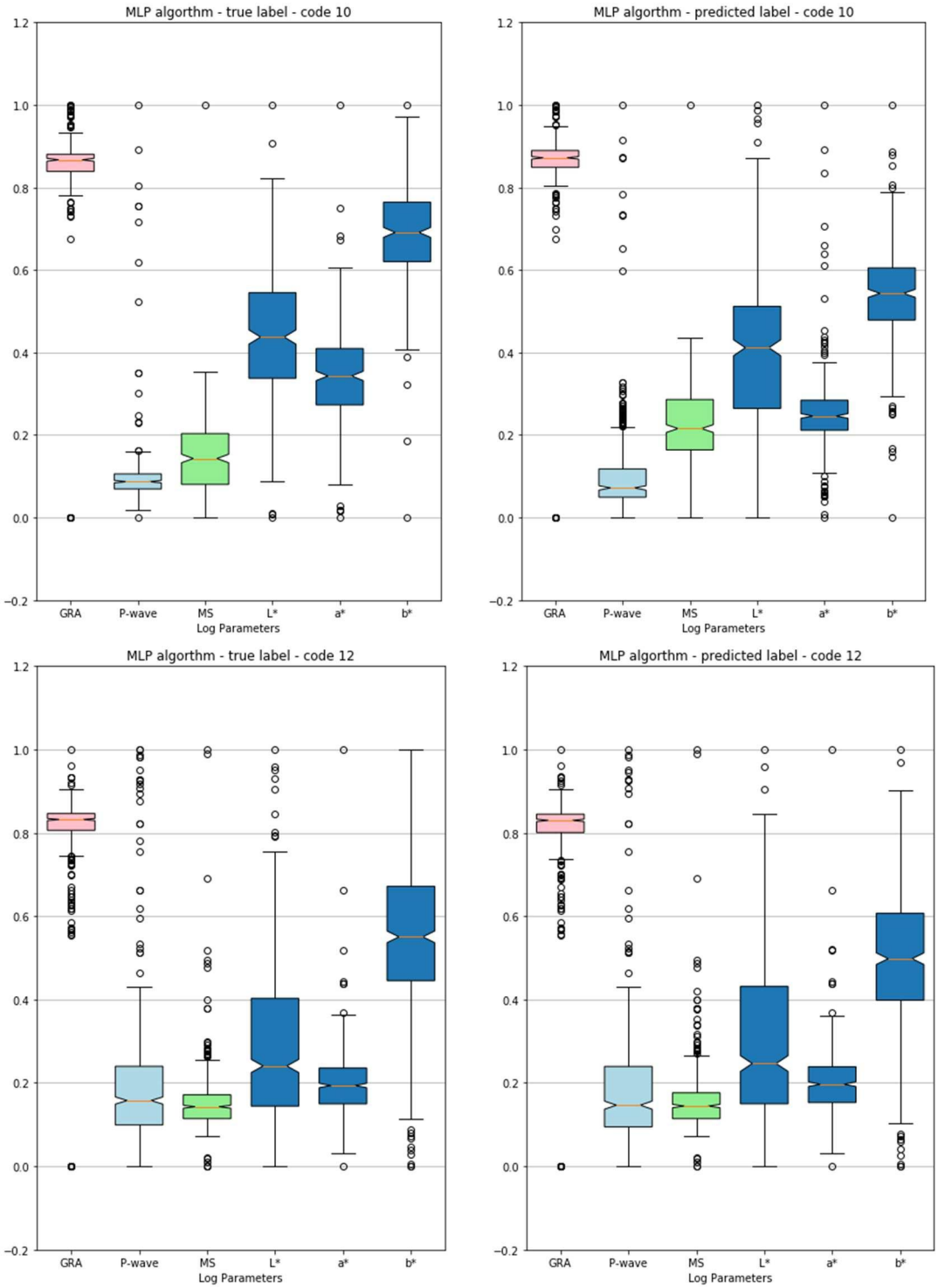
SVM:

Using configuration default (kernel: rbf, random-state: none)

Method	**Configuration**		
SVM	**Hiperparameters**		
	Kernel	random-state	Accuracy
	default (rbf)	**none**	**81,55**
	poly	range 5 to 9	40,77
	linear	range 5 to 9	79,05
	signoid	range 5 to 9	80,15

Appendix 1D – Boxplots of the true label and predicted label to IODP-Expedition 362, site U1481 (G3 group)

Note: log parameters values have been normalized range 0 and 1 using minimum and maximum scalar function. The use of this function is due to a better presentation of the box viewing all within the same range and format.



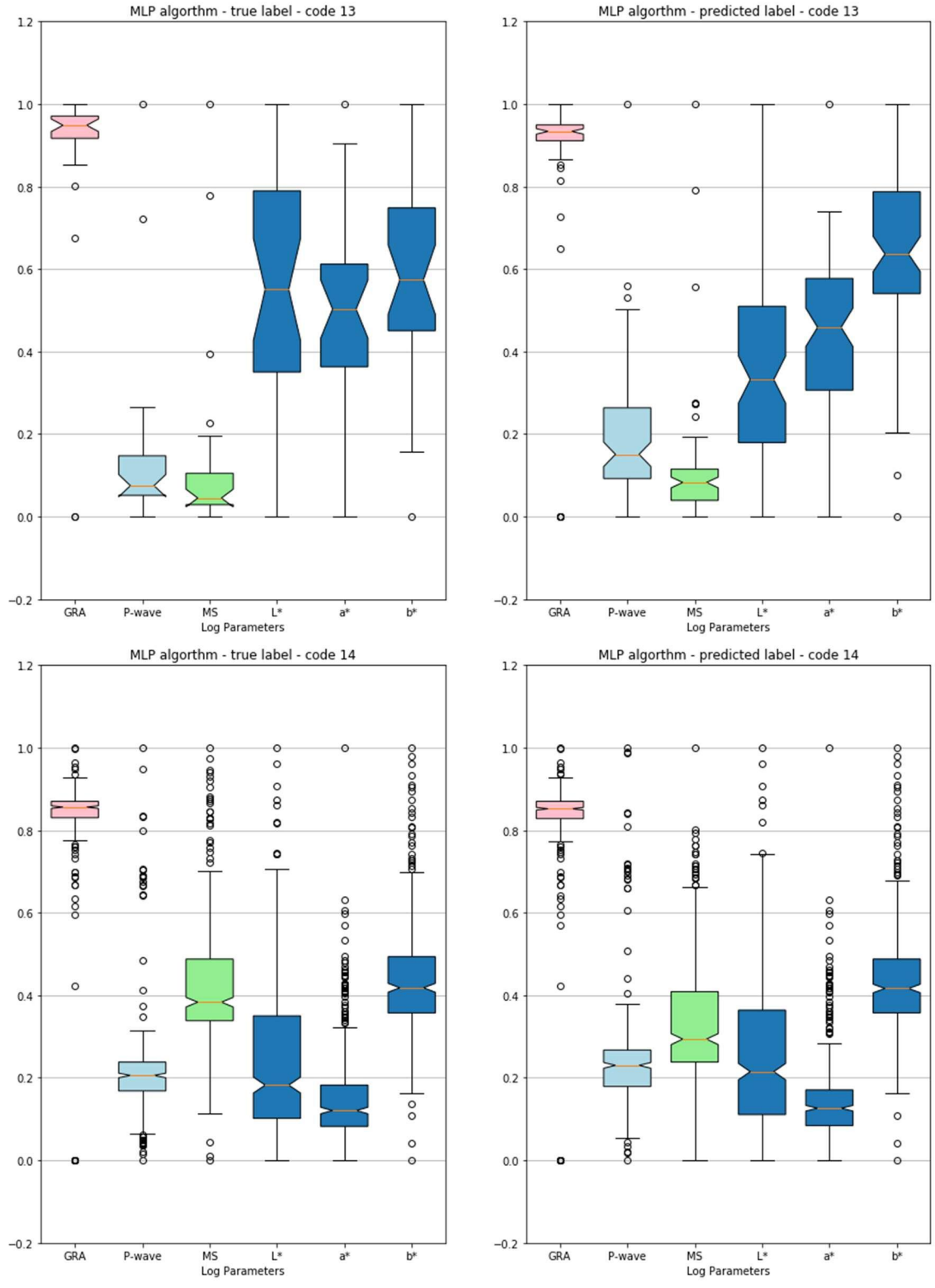
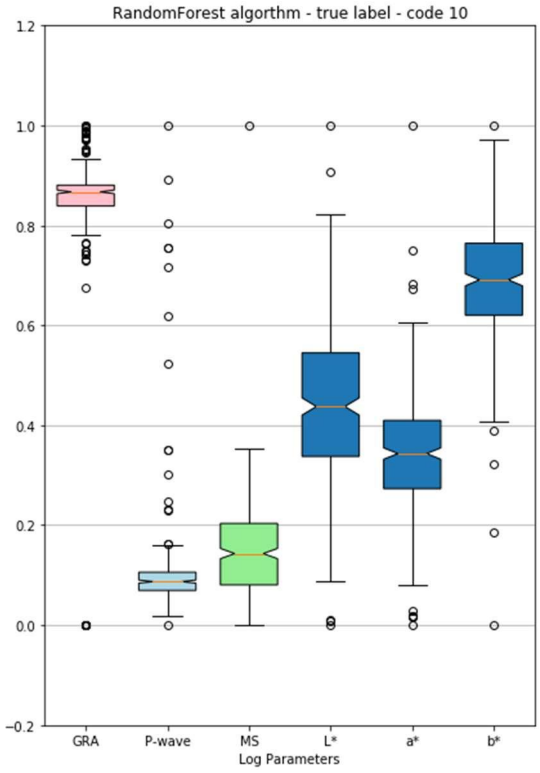
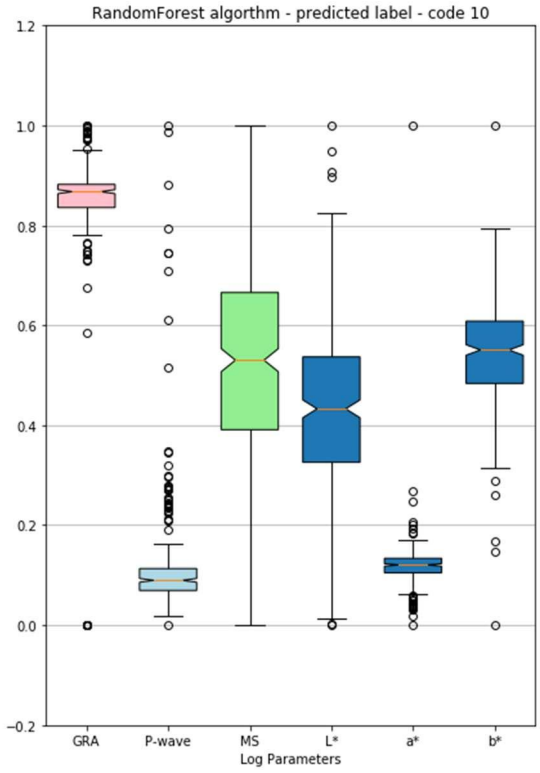
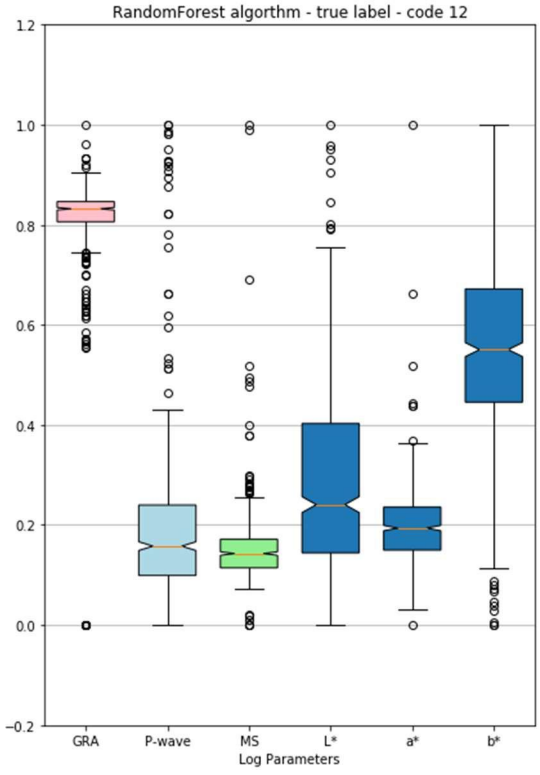
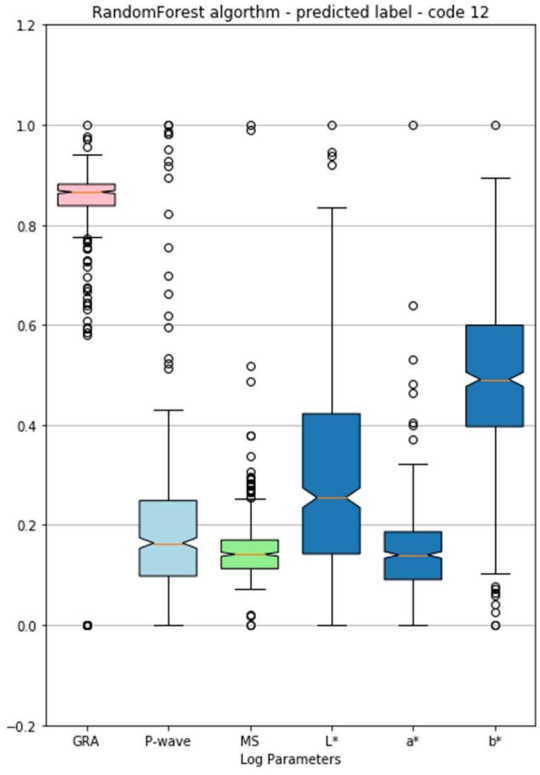


Figure 26 Graphical representation boxplot true label and predicted label, MLP algorithm, code 10 (referring to Litho1), code 12 (referring to Litho3), code 13 (referring to Litho4) and code 14 (referring to Litho5). Lithology Code 11 and Lithology Code 15 are not recorded in this hole range (Al-Mudhafar, 2016).


RandomForest algorthm - true label - code 10
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
GRA
P-wave
MS
L*
a*
b*
Log Parameters


RandomForest algorthm - predicted label - code 10
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
GRA
P-wave
MS
L*
a*
b*
Log Parameters


RandomForest algorthm - true label - code 12
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
GRA
P-wave
MS
L*
a*
b*
Log Parameters


RandomForest algorthm - predicted label - code 12
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
GRA
P-wave
MS
L*
a*
b*
Log Parameters

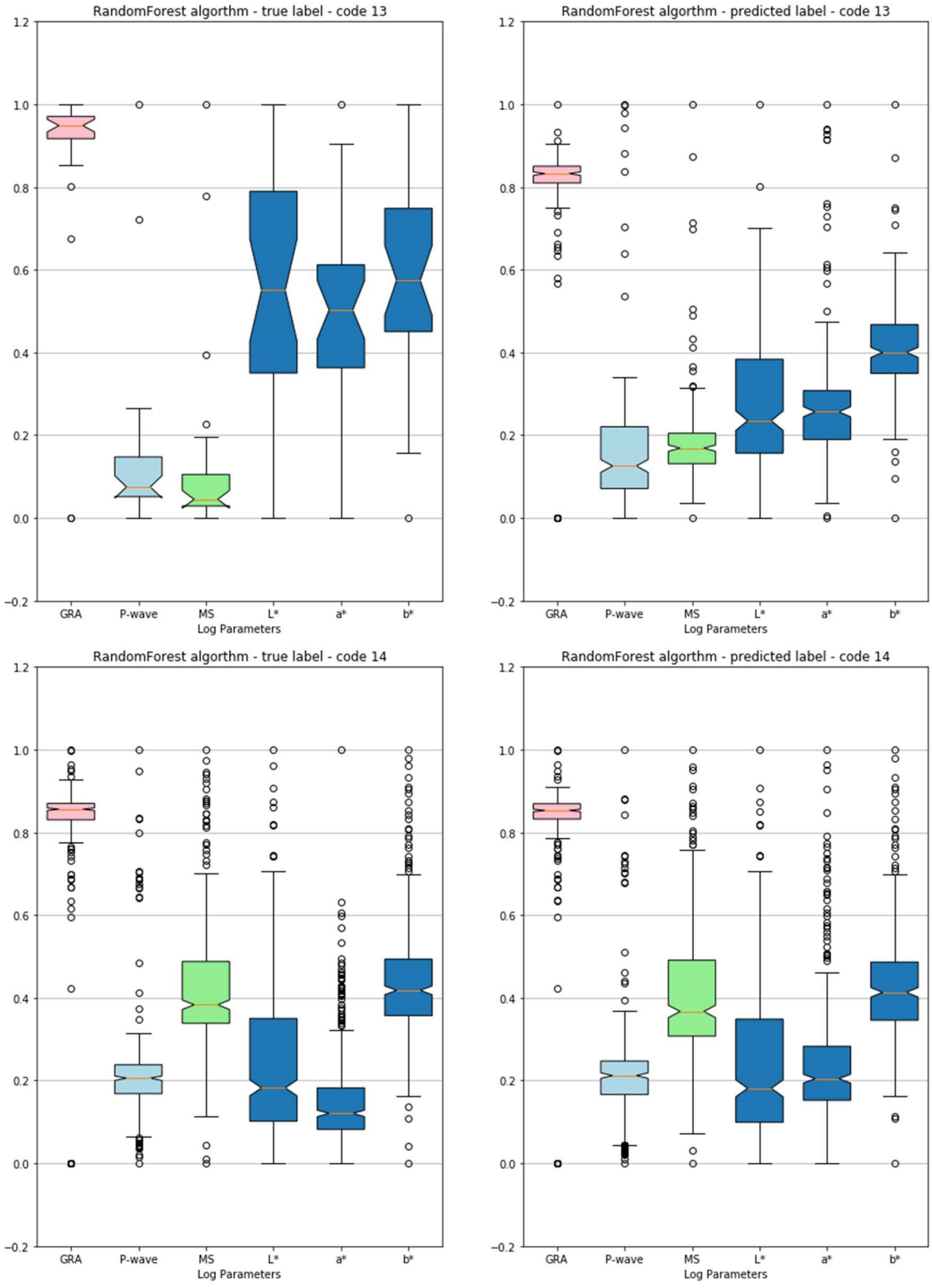
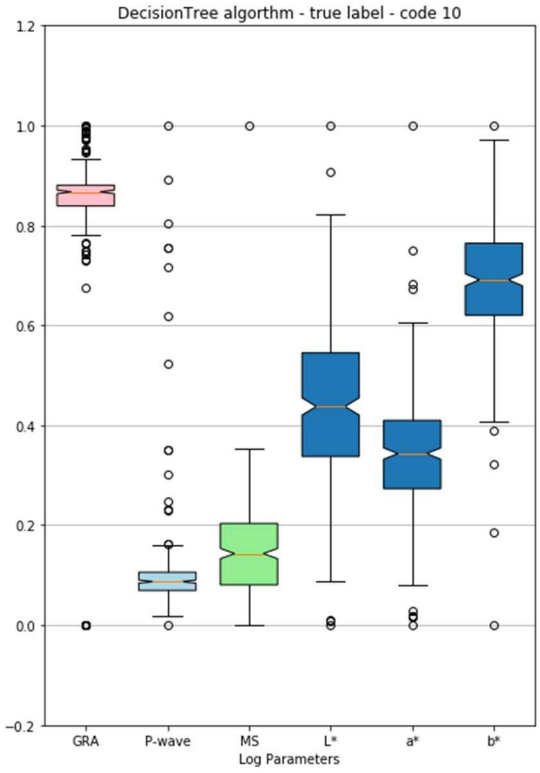
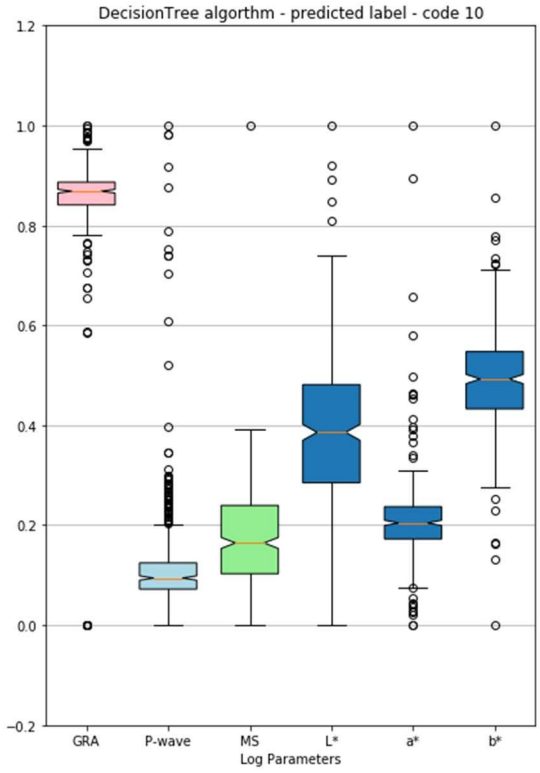
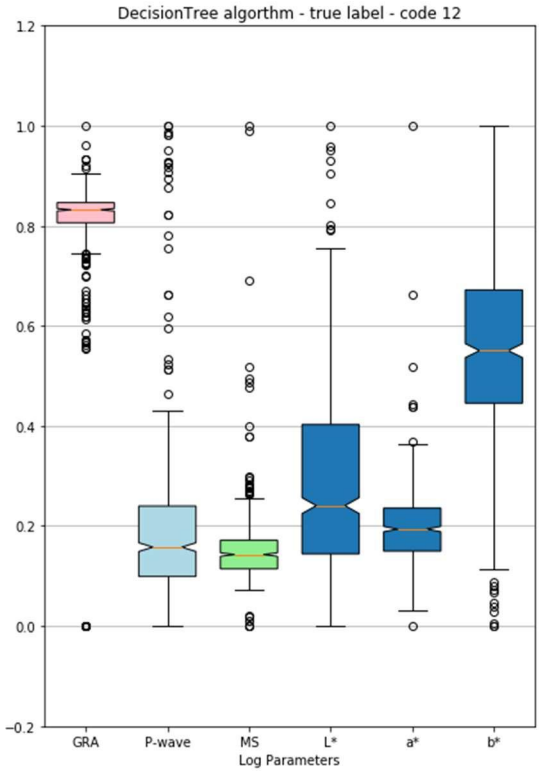
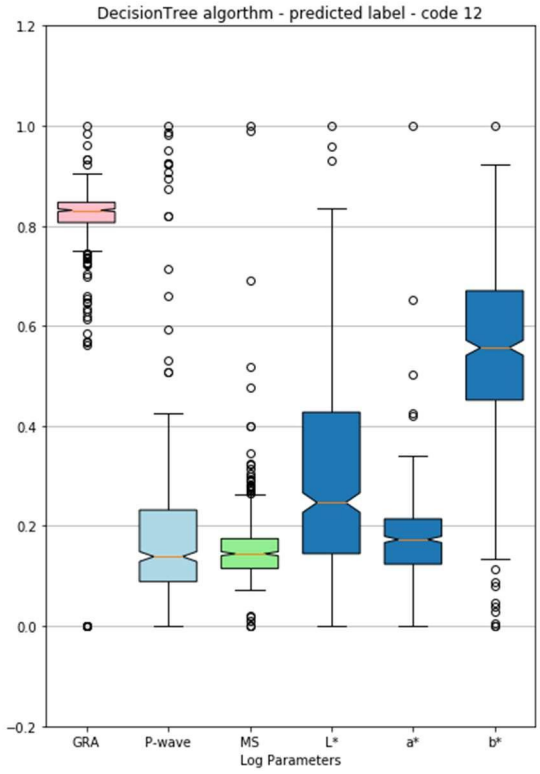


Figure 27 Graphical representation boxplot true label and predicted label, RandomForest algorithm, code 10 (referring to Litho1), code 12 (referring to Litho3), code 13 (referring to Litho4) and code 14 (referring to Litho5). Lithology Code 11 and Lithology Code 15 are not recorded in this hole range (Al-Mudhafar, 2016).


DecisionTree algorthm - true label - code 10
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
GRA
P-wave
MS
L*
a*
b*
Log Parameters


DecisionTree algorthm - predicted label - code 10
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
GRA
P-wave
MS
L*
a*
b*
Log Parameters


DecisionTree algorthm - true label - code 12
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
GRA
P-wave
MS
L*
a*
b*
Log Parameters


DecisionTree algorthm - predicted label - code 12
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
GRA
P-wave
MS
L*
a*
b*
Log Parameters

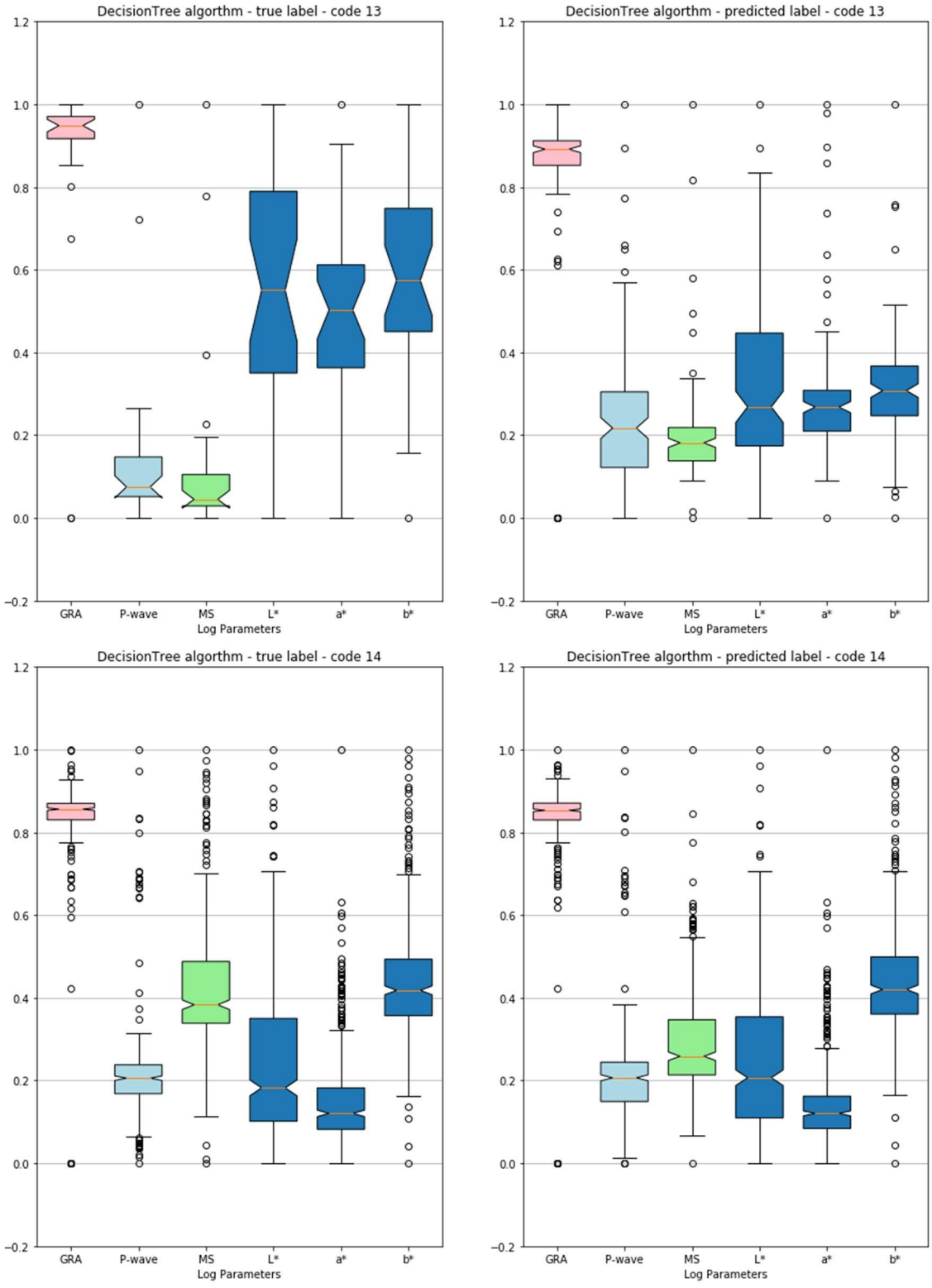
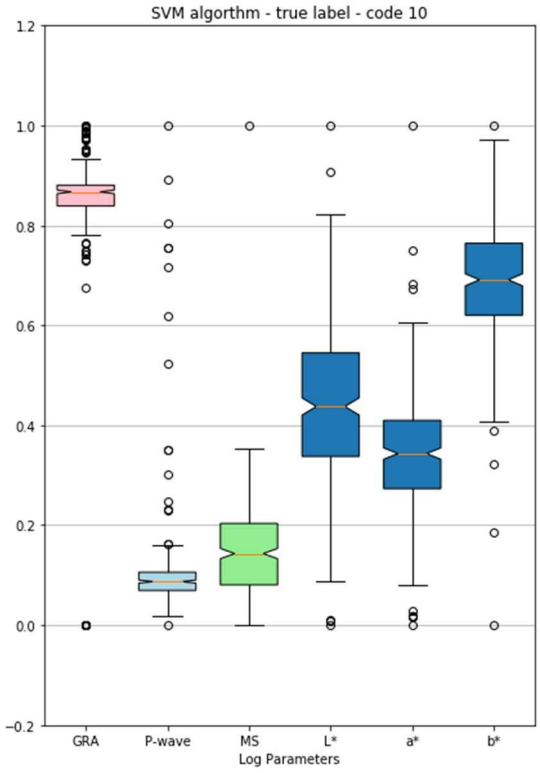
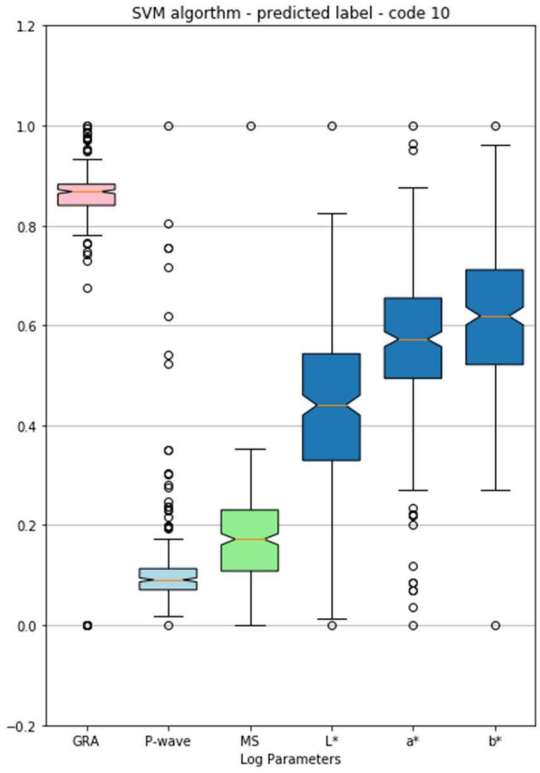
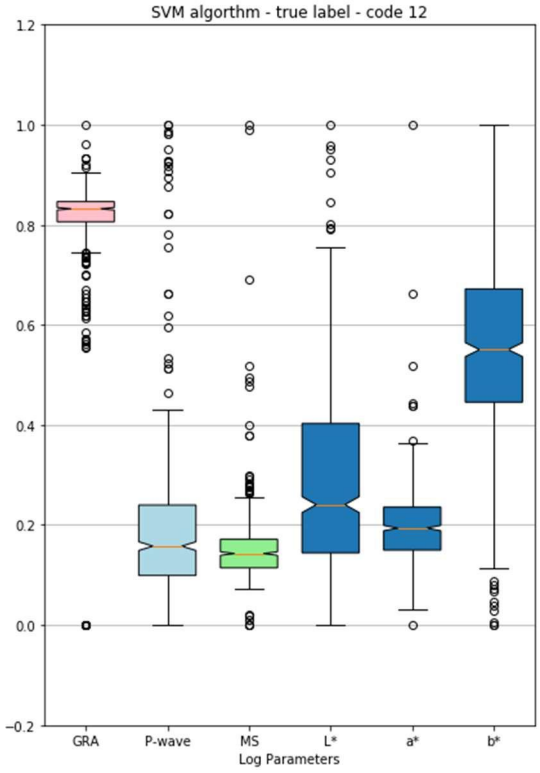
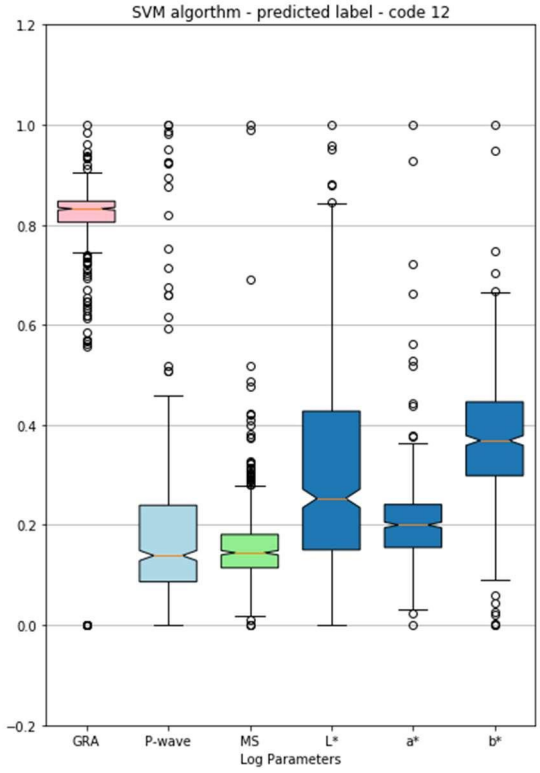


Figure 28 Graphical representation boxplot true label and predicted label, DecisionTree algorithm, code 10 (referring to Litho1), code 12 (referring to Litho3), code 13 (referring to Litho4) and code 14 (referring to Litho5). Lithology Code 11 and Lithology Code 15 are not recorded in this hole range (Al-Mudhafar, 2016).


SVM algorithm - true label - code 10
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
GRA
P-wave
MS
L*
a*
b*
Log Parameters


SVM algorithm - predicted label - code 10
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
GRA
P-wave
MS
L*
a*
b*
Log Parameters


SVM algorithm - true label - code 12
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
GRA
P-wave
MS
L*
a*
b*
Log Parameters


SVM algorithm - predicted label - code 12
1.2
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
GRA
P-wave
MS
L*
a*
b*
Log Parameters

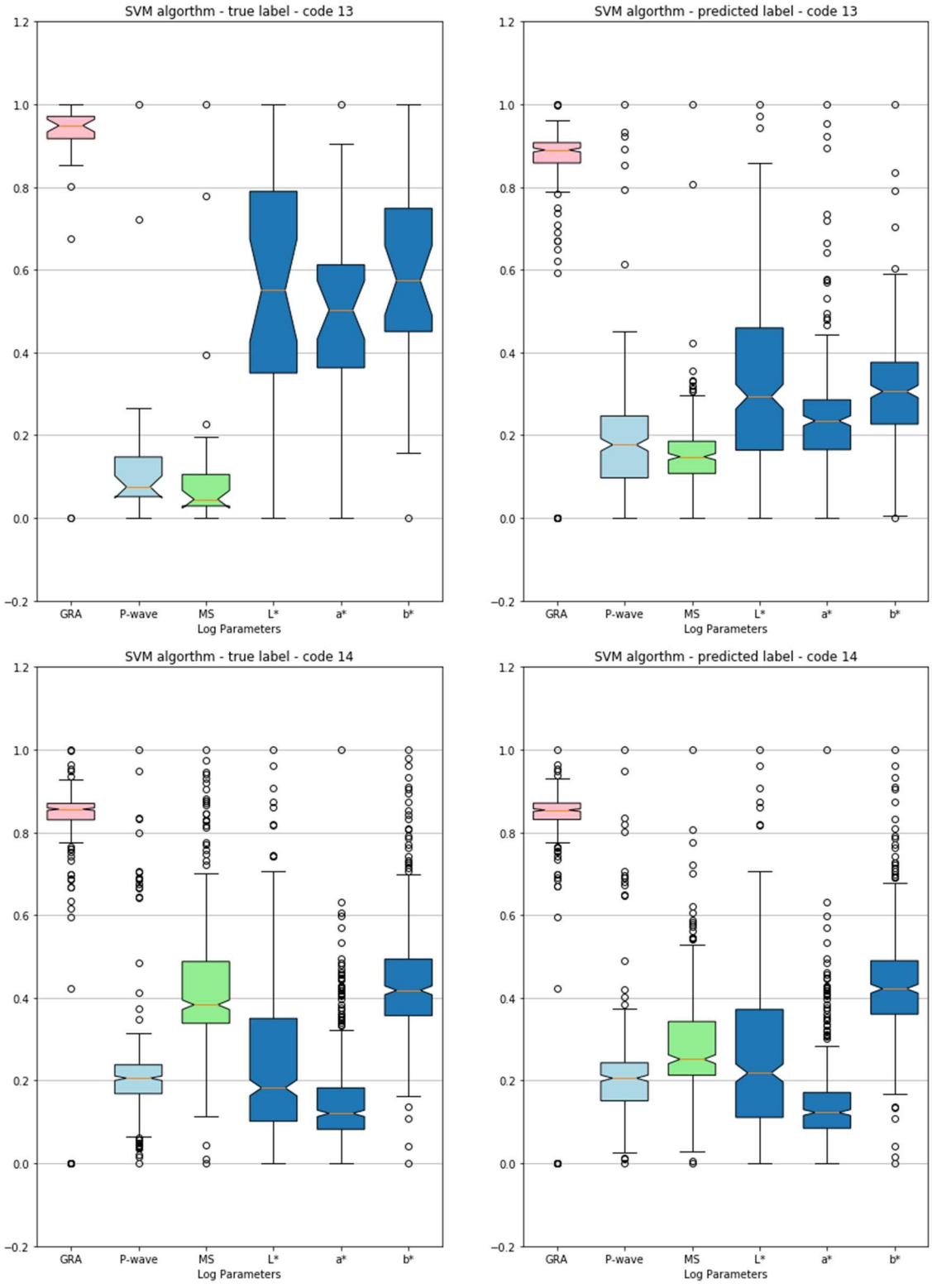


Figure 29 Graphical representation boxplot true label and predicted label, SVM algorithm, code 10 (referring to Litho1), code 12 (referring to Litho3), code 13 (referring to Litho4) and code 14 (referring to Litho5). Lithology Code 11 and Lithology Code 15 are not recorded in this hole range (Al-Mudhafar, 2016).

Appendix 1E – Result methods

GP

Evaluate Performance Machine Learning Algorithm with Datas Analysis

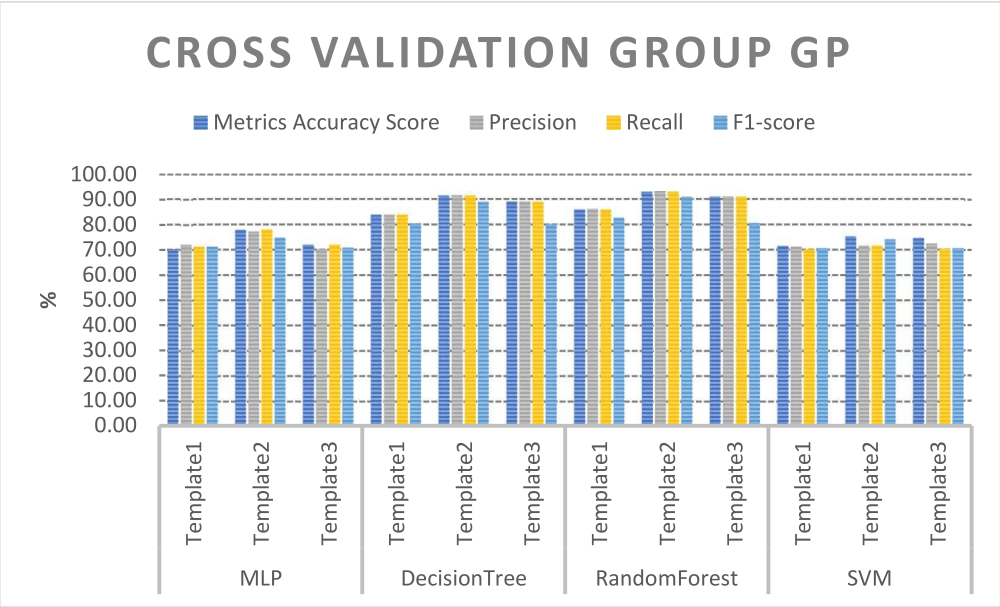
		Accuracy	Metrics Classification Report			
Method	**Template**	metrics accuracy score	Precision	Recall	F1-score	Support
MLP	Template1	70.01	77	71	72	886
	Template2	75.64	85	76	79	3026
	Template3	70.28	76	71	74	3912
	(1)Practical Template - 349 | U1431	79.77	75	80	77	900
	(2)Practical Template - 349 | U1432	84.70	78	85	81	353
	(3)Practical Template - 349 | U1433	84.77	83	85	84	1097
	(4)Practical Template - 356 | U1462	73.17	72	71	73	258
	(5)Practical Template - 361 | U1477	71.42	74	71	77	63
	(6)Practical Template - 361 | U1478	79.08	80	79	75	834
	(7)Practical Template - 362 | U1481	74.73	81	75	75	1439
DecisionTree	Template1	77.00	79	77	79	886
	Template2	77.82	78	78	78	3026
	Template3	74.48	77	74	75	3912
	(1)Practical Template - 349 | U1431	71.33	75	71	72	900
	(2)Practical Template - 349 | U1432	70.14	70.3	70.5	70.36	353
	(3)Practical Template - 349 | U1433	73.83	74	74	78	1097
	(4)Practical Template - 356 | U1462	71.12	71	71	72	258

	(5)Practical Template - 361 | U1477	71.42	72	71	76	63
	(6)Practical Template - 361 | U1478	76.40	74	74	75	834
	(7)Practical Template - 362 | U1481	79.12	73	73	75	1439
RandomForest	Template1	75.84	84	76	78	886
	Template2	82.61	89	83	84	3026
	Template3	82.74	86	83	84	3912
	(1)Practical Template - 349 | U1431	80.44	76	80	78	900
	(2)Practical Template - 349 | U1432	85.83	80	86	83	353
	(3)Practical Template - 349 | U1433	83.77	81	84	82	1097
	(4)Practical Template - 356 | U1462	71.11	71	71	72	258
	(5)Practical Template - 361 | U1477	80.95	81	81	81	63
	(6)Practical Template - 361 | U1478	79.08	81	79	75	834
	(7)Practical Template - 362 | U1481	72.66	83	74	75	1439
SVM	Template1	76.00	77	78	75	886
	Template2	78.48	84	78	79	3026
	Template3	78.99	73	70	70	3912
	(1)Practical Template - 349 | U1431	73.00	70	73	76	900
	(2)Practical Template - 349 | U1432	77.76	76	75	74	353
	(3)Practical Template - 349 | U1433	73.67	88	74	73	1097
	(4)Practical Template - 356 | U1462	72.54	79	79	77	258
	(5)Practical Template - 361 | U1477	71.05	71	70	71	63

	(6)Practical Template - 361 | U1478	79.22	80	79	79	834
	(7)Practical Template - 362 | U1481	75.66	75	77	74	1439

Cross-Validation for 5 k folds

		Metrics Accuracy Score	Metrics Classification Report		
Mehotd	**Template**	**Metrics Accuracy Score**	**Precision**	**Recall**	**F1-score**
MLP	Template1	70.05	72.00	71.00	70.98
	Template2	77.69	77.00	77.79	74.61
	Template3	71.83	70.03	71.83	70.60
DecisionTree	Template1	83.63	83.71	83.71	80.18
	Template2	91.36	91.40	91.36	88.74
	Template3	88.88	88.85	88.80	79.95
RandomForest	Template1	85.72	85.99	85.72	82.50
	Template2	93.02	93.16	93.02	90.92
	Template3	90.75	90.90	90.75	80.39
SVM	Template1	71.23	71.00	70.05	70.42
	Template2	75.09	71.36	71.46	73.96
	Template3	74.47	72.26	70.27	70.38



G1,G2,G3-CrossValidation

group of lithologies: G1

Cross-Validation for 5 k folds

		Metrics Accuracy Score	Metrics Classification Report		
Method	**Template**	**Metrics Accuracy Score**	**Precision**	**Recall**	**F1-score**
MLP	Template1	71.12	70.05	71.00	71.06
	Template2	77.69	77.79	78.27	74.61
	Template3	73.23	71.36	73.23	76.82
DecisionTree	Template1	87.18	87.25	87.18	87.21
	Template2	91.16	91.40	91.36	88.74
	Template3	89.98	90.00	89.98	88.68

RandomForest	Template1	92.42	92.50	92.42	92.58
	Template2	93.02	93.16	93.02	90.92
	Template3	92.69	92.82	92.69	91.75
SVM	Template1	70.56	71.00	70.80	70.04
	Template2	76.09	71.46	76.09	73.96
	Template3	75.72	74.33	75.72	71.04

group of lithologies: G2

Cross-Validation for 5 k folds

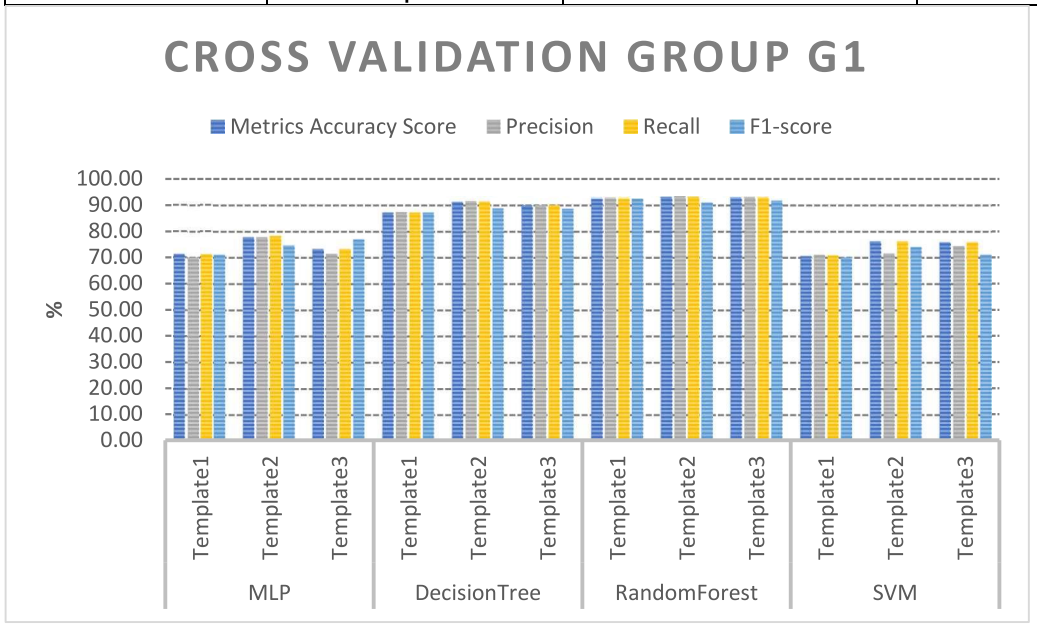
		Metrics Accuracy Score	Metrics Classification Report		
Method	**Template**	**Metrics Accuracy Score**	**Precision**	**Recall**	**F1-score**
MLP	Template1	70.14	70.63	71.42	70.04
	Template2	77.69	77.79	78.27	74.61
	Template3	75.54	75.01	75.99	75.54
DecisionTree	Template1	84.45	84.52	84.45	83.19
	Template2	91.36	91.40	91.36	88.74
	Template3	88.97	89.02	88.97	82.40
RandomForest	Template1	89.11	89.12	89.11	87.28
	Template2	93.02	93.16	93.02	90.92
	Template3	91.57	91.69	91.57	85.68
SVM	Template1	73.62	73.00	73.45	72.46
	Template2	76.09	71.46	76.09	73.96
	Template3	74.51	72.18	74.51	75.47

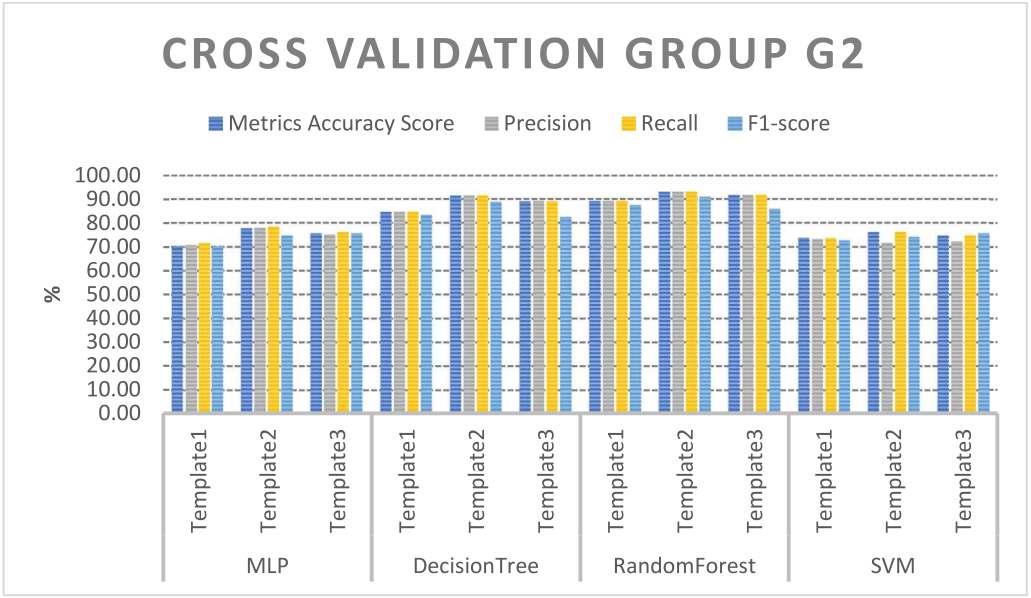
group of lithologies: G3

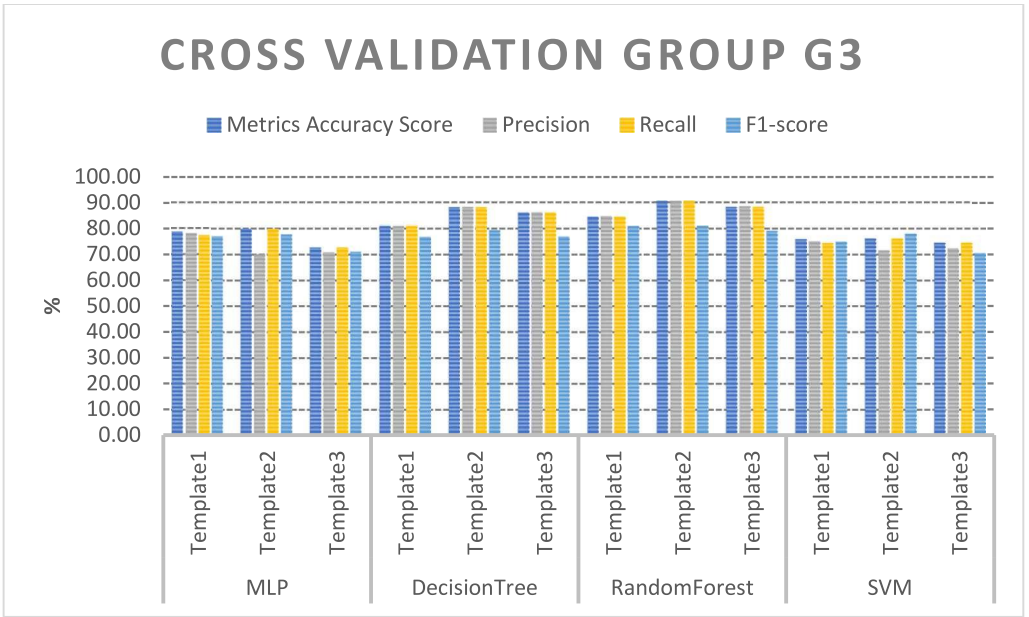
Cross-Validation for 5 k folds

		Metrics Accuracy Score	Metrics Classification Report		

Method	Template	Metrics Accuracy Score	Precision	Recall	F1-score
MLP	Template1	78.58	77.90	77.14	76.89
	Template2	79.75	70.05	79.75	77.71
	Template3	72.66	70.71	72.66	71.03
DecisionTree	Template1	80.87	80.88	80.87	76.66
	Template2	88.00	88.11	88.00	79.32
	Template3	85.87	85.95	85.87	76.78
RandomForest	Template1	84.28	84.54	84.28	80.86
	Template2	90.43	90.60	90.43	81.20
	Template3	88.15	88.30	88.15	79.06
SVM	Template1	75.78	75.00	74.30	74.87
	Template2	76.09	71.46	76.09	77.97
	Template3	74.47	72.25	74.47	70.33




CROSS VALIDATION GROUP G2
Metrics Accuracy Score
Precision
Recall
F1-score
%
100.00
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
Template1
Template2
Template3
MLP
DecisionTree
RandomForest
SVM


CROSS VALIDATION GROUP G3
Metrics Accuracy Score
Precision
Recall
F1-score
%
100.00
90.00
80.00
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
Template1
Template2
Template3
MLP
DecisionTree
RandomForest
SVM

G1,G2,G3- Full

group of lithologies: G1

Evaluate Performance Machine Learning Algorithm with Datas Analysis

		Accuracy	Metrics Classification Report			
Method	**Template**	metrics accuracy score	Precision	Recall	F1-score	Support
MLP	Template1	84.39%	85.00%	84.00%	84.00%	782
	Template2	75.64%	85.00%	76.00%	79.00%	3026
	Template3	73.87%	74.00%	74.00%	73.00%	3808
	(1)Practical Template - 349 | U1431	84.82%	79.00%	85.00%	80.00%	883
	(2)Practical Template - 349 | U1432	83.00%	75.00%	83.00%	79.00%	353
	(3)Practical Template - 349 | U1433	87.14%	86.00%	87.00%	87.00%	1097
	(4)Practical Template - 356 | U1462	74.00%	72.00%	71.00%	73.00%	258
	(5)Practical Template - 361 | U1477	71.74%	71.00%	73.00%	71.00%	63
	(6)Practical Template - 361 | U1478	70.88%	77.00%	74.00%	75.00%	834
	(7)Practical Template - 362 | U1481	76.75%	73.00%	77.00%	73.00%	1439
DecisionTree	Template1	82.48%	83.00%	82.00%	83.00%	782
	Template2	77.82%	78.00%	78.00%	78.00%	3026
	Template3	73.84%	75.00%	74.00%	74.00%	3808
	(1)Practical Template - 349 | U1431	78.28%	71.00%	78.00%	78.00%	833
	(2)Practical Template - 349 | U1432	73.82%	74.00%	74.00%	74.00%	353
	(3)Practical Template - 349 | U1433	73.81%	78.00%	74.00%	74.00%	1097
	(4)Practical Template - 356 | U1462	71.10%	71.00%	72.00%	72.00%	258
	(5)Practical Template - 361 | U1477	78.25%	78.00%	76.00%	76.00%	63
	(6)Practical Template - 361 | U1478	76.28%	74.00%	76.00%	74.00%	834
	(7)Practical Template - 362 | U1481	79.47%	77.00%	74.00%	78.00%	1439
RandomForest	Template1	89.51%	90.00%	90.00%	89.00%	782
	Template2	82.61%	89.00%	83.00%	84.00%	3026

	Template3	86.42%	87.00%	86.00%	86.00%	3808
	(1)Practical Template - 349 | U1431	84.93%	78.00%	85.00%	81.00%	883
	(2)Practical Template - 349 | U1432	85.83%	80.00%	86.00%	83.00%	353
	(3)Practical Template - 349 | U1433	84.50%	81.00%	85.00%	83.00%	1097
	(4)Practical Template - 356 | U1462	71.12%	71.00%	71.00%	72.00%	258
	(5)Practical Template - 361 | U1477	80.95%	81.00%	81.00%	81.00%	63
	(6)Practical Template - 361 | U1478	79.08%	81.00%	79.00%	75.00%	834
	(7)Practical Template - 362 | U1481	70.14%	84.00%	70.00%	71.00%	1439
SVM	Template1	79.41%	81.00%	79.00%	80.00%	782
	Template2	78.48%	84.00%	78.00%	79.00%	3026
	Template3	73.31%	75.00%	73.00%	72.00%	3808
	(1)Practical Template - 349 | U1431	78.39%	78.00%	77.00%	75.00%	883
	(2)Practical Template - 349 | U1432	81.02%	79.00%	79.00%	77.00%	353
	(3)Practical Template - 349 | U1433	72.78%	72.00%	71.00%	73.00%	1097
	(4)Practical Template - 356 | U1462	72.02%	70.00%	71.00%	71.00%	258
	(5)Practical Template - 361 | U1477	76.53%	75.00%	75.00%	74.00%	63
	(6)Practical Template - 361 | U1478	74.88%	74.00%	73.00%	73.00%	834
	(7)Practical Template - 362 | U1481	70.61%	70.00%	71.00%	70.00%	1439

group of lithologies: G2

Evaluate Performance Machine Learning Algorithm with Datas Analysis

		Accuracy	Metrics Classification Report			
Method	**Template**	metrics accuracy score	Precision	Recall	F1-score	Support
MLP	Template1	76.63%	80.00%	77.00%	78.00%	886
	Template2	75.64%	85.00%	76.00%	79.00%	3026
	Template3	79.91%	78.00%	76.00%	77.00%	3912
	(1)Practical Template - 349 | U1431	87.55%	85.00%	88.00%	86.00%	900
	(2)Practical Template - 349 | U1432	88.66%	86.00%	89.00%	87.00%	353
	(3)Practical Template - 349 | U1433	93.71%	100.00%	94.00%	96.00%	1097

	(4)Practical Template - 356 | U1462	93.00%	94.00%	90.00%	92.00%	258
	(5)Practical Template - 361 | U1477	85.71%	90.00%	86.00%	88.00%	63
	(6)Practical Template - 361 | U1478	74.36%	89.00%	74.00%	75.00%	834
	(7)Practical Template - 362 | U1481	72.55%	92.00%	73.00%	81.00%	1439
DecisionTree	Template1	70.89%	73.00%	71.00%	72.00%	886
	Template2	77.82%	78.00%	78.00%	78.00%	3026
	Template3	77.14%	79.00%	77.00%	78.00%	3912
	(1)Practical Template - 349 | U1431	77.77%	71.00%	78.00%	78.00%	900
	(2)Practical Template - 349 | U1432	88.66%	86.00%	89.00%	87.00%	353
	(3)Practical Template - 349 | U1433	78.61%	73.00%	75.00%	74.00%	1097
	(4)Practical Template - 356 | U1462	72.23%	71.00%	73.00%	72.00%	258
	(5)Practical Template - 361 | U1477	73.01%	76.00%	73.00%	70.00%	63
	(6)Practical Template - 361 | U1478	74.41%	74.00%	73.00%	74.00%	834
	(7)Practical Template - 362 | U1481	74.68%	77.00%	75.00%	74.00%	1439
RandomForest	Template1	82.84%	96.00%	83.00%	84.00%	886
	Template2	82.61%	89.00%	83.00%	84.00%	3026
	Template3	84.20%	87.00%	84.00%	85.00%	3912
	(1)Practical Template - 349 | U1431	80.66%	73.00%	81.00%	76.00%	900
	(2)Practical Template - 349 | U1432	85.26%	79.00%	85.00%	82.00%	353
	(3)Practical Template - 349 | U1433	84.32%	81.00%	84.00%	82.00%	1097
	(4)Practical Template - 356 | U1462	71.12%	71.00%	72.00%	71.00%	258
	(5)Practical Template - 361 | U1477	80.95%	81.00%	81.00%	81.00%	63
	(6)Practical Template - 361 | U1478	79.08%	81.00%	79.00%	75.00%	834
	(7)Practical Template - 362 | U1481	70.07%	84.00%	70.00%	77.00%	1439
SVM	Template1	70.88%	81.00%	71.00%	75.00%	886
	Template2	78.48%	84.00%	78.00%	79.00%	3026
	Template3	73.03%	74.00%	73.00%	73.00%	3912
	(1)Practical Template - 349 | U1431	78.09%	78.00%	78.00%	77.00%	900
	(2)Practical Template - 349 | U1432	87.03%	86.00%	87.00%	86.00%	353
	(3)Practical Template - 349 | U1433	75.89%	72.00%	71.00%	71.00%	1097

	(4)Practical Template - 356 | U1462	73.17%	73.00%	73.00%	72.00%	258
	(5)Practical Template - 361 | U1477	79.06%	80.00%	79.00%	79.00%	63
	(6)Practical Template - 361 | U1478	72.38%	77.00%	72.00%	74.00%	834
	(7)Practical Template - 362 | U1481	74.09%	78.00%	73.00%	73.00%	1439

group of lithologies: G3

Evaluate Performance Machine Learning Algorithm with Datas Analysis

		Accuracy	Metrics Classification Report			
Method	**Template**	metrics accuracy score	Precision	Recall	F1-score	Support
MLP	Template1	72.03%	72.00%	71.60%	72.04%	886
	Template2	74.35%	80.00%	74.00%	75.00%	3026
	Template3	70.51%	73.00%	72.00%	72.00%	3912
	(1)Practical Template - 349 | U1431	77.88%	72.00%	78.00%	74.00%	900
	(2)Practical Template - 349 | U1432	78.44%	71.00%	78.00%	74.00%	353
	(3)Practical Template - 349 | U1433	73.12%	72.00%	73.00%	74.00%	1097
	(4)Practical Template - 356 | U1462	70.85%	73.00%	76.00%	74.00%	258
	(5)Practical Template - 361 | U1477	76.98%	78.00%	72.00%	71.00%	63
	(6)Practical Template - 361 | U1478	77.74%	71.00%	75.00%	72.00%	834
	(7)Practical Template - 362 | U1481	76.33%	76.00%	76.00%	75.00%	1439
DecisionTree	Template1	73.00%	74.00%	74.52%	74.10%	886
	Template2	72.40%	72.00%	72.00%	72.00%	3026
	Template3	76.74%	77.00%	77.00%	76.00%	3912
	(1)Practical Template - 349 | U1431	71.33%	74.00%	73.00%	71.00%	900
	(2)Practical Template - 349 | U1432	73.62%	72.00%	74.00%	72.00%	353
	(3)Practical Template - 349 | U1433	74.42%	73.00%	75.00%	75.00%	1097
	(4)Practical Template - 356 | U1462	71.22%	71.00%	74.00%	74.00%	258
	(5)Practical Template - 361 | U1477	76.66%	76.00%	77.00%	75.00%	63
	(6)Practical Template - 361 | U1478	73.05%	72.00%	73.00%	73.00%	834
	(7)Practical Template - 362 | U1481	75.55%	76.00%	76.00%	74.00%	1439

RandomForest	Template1	77.87%	83.00%	78.00%	79.00%	886
	Template2	77.89%	86.00%	78.00%	80.00%	3026
	Template3	79.42%	82.00%	79.00%	80.00%	3912
	(1)Practical Template - 349 | U1431	77.44%	71.00%	77.00%	74.00%	900
	(2)Practical Template - 349 | U1432	84.70%	78.00%	85.00%	81.00%	353
	(3)Practical Template - 349 | U1433	78.66%	72.00%	79.00%	75.00%	1097
	(4)Practical Template - 356 | U1462	71.12%	71.00%	72.00%	71.00%	258
	(5)Practical Template - 361 | U1477	77.77%	75.00%	78.00%	76.00%	63
	(6)Practical Template - 361 | U1478	71.11%	89.00%	71.00%	72.00%	834
	(7)Practical Template - 362 | U1481	75.64%	75.00%	76.00%	71.00%	1439
SVM	Template1	79.00%	79.00%	78.87%	78.00%	886
	Template2	74.52%	80.00%	75.00%	76.00%	3026
	Template3	70.19%	76.00%	71.00%	74.00%	3912
	(1)Practical Template - 349 | U1431	72.00%	73.00%	72.00%	72.00%	900
	(2)Practical Template - 349 | U1432	79.55%	82.00%	79.00%	79.00%	353
	(3)Practical Template - 349 | U1433	75.55%	72.00%	74.00%	74.00%	1097
	(4)Practical Template - 356 | U1462	71.33%	75.00%	72.00%	72.00%	258
	(5)Practical Template - 361 | U1477	76.01%	77.00%	75.00%	76.00%	63
	(6)Practical Template - 361 | U1478	73.81%	75.00%	74.00%	74.00%	834
	(7)Practical Template - 362 | U1481	72.04%	72.00%	72.00%	67.00%	1439

Quantity of sample by lithology

GROUP OF LITHOLOGIES

IODP Expedition	Site	GP						G1			G2				G3					
		10	11	12	13	14	15	10	14	15	10	13	14	15	10	11	12	13	14	15
	U1431			26	17	830	27		856	27		17	856	27		17	26		830	27
349	U1432	15				338		15	338		15		338		15				338	
	U1433			8		1018	71		1026	71			1026	71			8		1018	71
356	U1462						258			258				258						258
361	U1477	6				57		6	57		6		57		6				57	
	U1478	471				363		471	363		471		363		471			353	10	
362	U1481	354		410		675		354	1085		354		1085		354		648	31	406	
362	U1480	1507	361	669	159	1458	276	1507	2127	276	1507	520	2127	276	1507	520	1213	168	746	276
354	U1449	309				465	121	309	465	121	309		465	121	309			6	459	121
	U1450	781				634	439	781	634	439	781		634	439	781			159	475	439
	U1451	801				2016	958	801	2016	958	801		2016	958	825			645	1347	958
355	U1456	534				1106	1270	534	1106	1270	534		1106	1270	534			152	954	1270
359	U1465						94			94				94						94
	U1466						103			103				103						103
	U1467						3037			3037				3037						3037
	U1468						1302			1302				1302						1302
	U1470						67			67				67						67
	U1471						669			669				669						669
	U1472						424			424				424						424

Appendix 1F – Ash layer U1480

Exp	Site	Hole	Core	Type	Sect	A/W	Offset (cm)	Depth CSF-A (m)	Bulk density (GRA)	P-wave	caliper	sonic	MS	Inclinationback	Declinationbac	Intensityback	clinationra	eclinationra	ntensityraw	gneticmom	gneticmom	gneticmom	cod_lit	yback;Inclinationraw;Declinationraw;Intensityraw;MagneticmomentX;MagneticmomentY;MagneticmomentZ;cod_lit
362	U1480	B	1	H	5		103.1	7061	1623	1546	6.54	50.38	41.00											
362	U1480	B	1	H	5		105.6	7086	1610	1528	6.54	50.86	42.00											
362	U1480	C	1	H	6		50.6	8006	1472	1501	6.60	52.05	24.00											
362	U1480	C	1	H	6		53.1	8031	1474	1499	6.58	51.96	23.00											
362	U1480	C	1	H	6		55.6	8056	1472	1500	6.55	51.78	23.00											
362	U1480	C	1	H	6		58.1	8081	1460	1503	6.52	51.48	23.00											
362	U1480	C	1	H	6		60.6	8106	1513	1518	6.51	50.96	24.00											
362	U1480	C	1	H	6		63.1	8131	1526	1530	6.49	50.53	23.00											
362	U1480	C	1	H	6		65.6	8156	1528	1532	6.47	50.27	24.00											
362	U1480	C	1	H	6		68.1	8181	1595	1595	6.42	48.34	22.00											
362	U1480	C	1	H	6		70.6	8206	1686	1617	6.38	47.54	17.00											
362	U1480	C	1	H	6		73.1	8231	1688	1633	6.33	46.88	12.00											
362	U1480	C	1	H	6		75.6	8256	1654	1629	6.33	46.93	12.00											
362	U1480	C	1	H	6		78.1	8281	1600	1621	6.35	47.22	12.00											
362	U1480	C	1	H	6		80.6	8306	1578	1620	6.39	47.50	14.00											
362	U1480	C	1	H	6		83.1	8331	1585	1621	6.41	47.60	14.00											
362	U1480	C	1	H	6		85.6	8356	1578	1617	6.42	47.77	16.00											
362	U1480	C	1	H	6		88.1	8381	1564	1612	6.44	48.02	17.00											
362	U1480	C	1	H	6		90.6	8406	1536	1612	6.44	48.04	18.00											
362	U1480	C	1	H	6		93.1	8431	1562	1614	6.45	48.04	18.00											
362	U1480	C	1	H	6		95.6	8456	1556	1615	6.46	48.07	18.00											
362	U1480	C	1	H	6		98.1	8481	1589	1618	6.46	48.01	19.00											
362	U1480	C	1	H	6		100.6	8506	1565	1616	6.48	48.16	21.00											
362	U1480	C	1	H	6		103.1	8531	1568	1625	6.49	48.02	28.00											
362	U1480	C	1	H	6		105.6	8556	1605	1643	6.50	47.62	41.00											
362	U1480	C	1	H	6		108.1	8581	1621	1663	6.50	47.19	48.00											
362	U1480	C	1	H	6		110.6	8606	1517	1521	6.51	50.90	36.00											
362	U1480	E	1	H	2		5.6	1406	1613	1595	6.54	49.08	35.00											
362	U1480	E	1	H	2		8.1	1431	1598	1621	6.51	48.24	38.00											
362	U1480	E	1	H	2		10.6	1456	1602	1627	6.50	48.05	38.00											
362	U1480	E	1	H	2		13.1	1481	1607	1627	6.50	48.05	38.00											
362	U1480	E	1	H	2		15.6	1506	1488	1639	6.51	47.79	31.00											
362	U1480	E	1	H	2		18.1	1531	1545	1633	6.50	47.91	29.00											
362	U1480	E	1	H	2		20.6	1556	1425	1635	6.50	47.83	29.00											
362	U1480	E	1	H	2		23.1	1581	1409	1633	6.49	47.86	29.00											
362	U1480	E	1	H	2		25.6	1606	1380	1627	6.49	47.98	26.00											
362	U1480	E	1	H	2		28.1	1631	1322	1626	6.48	47.97	26.00											
362	U1480	E	1	H	2		30.6	1656	1348	1610	6.49	48.37	25.00											
362	U1480	E	1	H	2		33.1	1681	1387	1617	6.48	48.17	27.00											
362	U1480	E	1	H	2		35.6	1706	1311	1606	6.48	48.43	30.00											
362	U1480	E	1	H	2		38.1	1731	1415	1577	6.48	49.16	29.00											
362	U1480	E	1	H	2		40.6	1756	1351	1513	6.48	50.89	31.00											
362	U1480	E	1	H	2		43.1	1781	1416	1633	6.48	47.79	52.00											
362	U1480	E	1	H	2		45.6	1806	1558	1645	6.48	47.45	77.00											
362	U1480	E	1	H	2		48.1	1831	1622	1575	6.48	49.26	63.00											
362	U1480	E	3	H	1		32.1	17621	1654	1600	6.52	48.81	20.00											
362	U1480	E	3	H	1		34.6	17646	1715	1654	6.52	47.49	17.00											
362	U1480	E	3	H	1		37.1	17671	1693	1640	6.52	47.81	16.00											
362	U1480	E	3	H	1		39.6	17696	1593	1612	6.51	48.50	20.00											
362	U1480	E	4	H	1		10.0	26900	1570	1603	6.58	49.15	73.00	79.2	355.6	0.07445	79.1	355.9	0.074341	1.4832E-6	-1.0942E-7	1.2679E-5		
362	U1480	G	57	R	5		132.9	1285399	1787				12.67											
362	U1480	G	57	R	5		135.4	1285424	1853	2563	5.55	29.78	16.67											
362	U1480	G	58	R	1		105.0	1289650	1413				6.67	27.5	78.2	0.0014261	27.7	77.6	0.0013037	2.0202E-8	9.3791E-8	8.0530E-8		
362	U1480	G	58	R	1		107.5	1289675	1696				5.67	44	300.3	0.0017761	40.8	297.6	0.0018097	5.1606E-8	-1.0108E-7	1.5710E-7		
362	U1480	G	59	R	1		102.5	1299325	1617				10.00	6.7	280	0.0073301	6.2	279.8	0.007431	1.0185E-7	-6.0566E-7	1.0666E-7		
362	U1480	G	59	R	1		105.0	1299350	1458				11.33	-1	297	0.0079212	-1.4	296.5	0.0080194	2.9129E-7	-5.9670E-7	-2.5601E-8		
362	U1480	G	60	R	4		54.7	1312747	2026				9.00											
362	U1480	G	60	R	4		57.2	1312772	1862	2180	6.43	37.64	10.67											
362	U1480	H	1	H	2		2.4	1524	1298				29.00											
362	U1480	H	1	H	2		4.9	1549	1269	1589	6.43	48.49	31.33											
362	U1480	H	1	H	2		7.4	1574	1316	1588	6.39	48.27	29.00											
362	U1480	H	1	H	2		9.9	1599	1277	1608	6.37	47.63	26.00											
362	U1480	H	1	H	2		12.4	1624	1229	1605	6.36	47.61	24.33											
362	U1480	H	1	H	2		14.9	1649	1214	1604	6.36	47.64	24.00											
362	U1480	H	1	H	2		17.4	1674	1197	1603	6.36	47.68	23.00											
362	U1480	H	1	H	2		19.9	1699	1156	1606	6.36	47.62	22.00											
362	U1480	H	1	H	2		22.4	1724	1159	1608	6.36	47.59	22.33											
362	U1480	H	1	H	2		24.9	1749	1222	1612	6.36	47.51	22.33											
362	U1480	H	1	H	2		27.4	1774	1267	1606	6.36	47.64	23.00											
362	U1480	H	1	H	2		29.9	1799	1307	1606	6.36	47.65	24.00											
362	U1480	H	1	H	2		32.4	1824	1337	1603	6.36	47.71	26.00											
362	U1480	H	1	H	2		34.9	1849	1396	1602	6.37	47.75	27.00											
362	U1480	H	2	H	1		112.5	5625	1806	1506	6.44	50.79	38.67	59.5	44.7	0.018754	59.6	44	0.018615	7.1956E-7	7.1174E-7	2.7878E-6		
362	U1480	H	2	H	1		115.0	5650	1532	1463	6.44	52.05	31.00	79.6	302.7	0.010583	78.9	300.2	0.010511	1.0866E-7	-1.9054E-7	1.7908E-6		

362	U1480	H	2	H	3		72.0	8040	1636	1559	6.47	49.50	46.67																							
362	U1480	H	2	H	3		74.5	8065	1782	1668	6.48	46.84	46.67																							
362	U1480	H	2	H	3		137.0	8690	1924	1560	6.39	48.96	50.00																							
362	U1480	H	3	H	4		0.4	18334					16.00																							
362	U1480	H	3	H	4		2.9	18359	1623				17.67																							
362	U1480	H	3	H	4		5.4	18384	1789				16.00																							
362	U1480	H	3	H	4		7.9	18409	1636				13.00																							
362	U1480	H	3	H	4		10.4	18434	1608	1613	6.38	47.55	13.33																							
362	U1480	H	3	H	4		12.9	18459	1579	1600	6.38	47.89	14.33																							
362	U1480	H	3	H	4		15.4	18484	1545	1595	6.39	48.07	17.00																							
362	U1480	H	3	H	4		17.9	18509	1554	1589	6.40	48.31	26.67																							
362	U1480	H	3	H	4		60.4	18934	2041	2292	6.53	36.52	74.00																							
362	U1480	H	3	H	4		62.9	18959	1828	1526	6.54	50.89	59.00																							
362	U1480	H	4	H	1		52.5	24025	1640	1422	6.42	53.15	41.33	73.2	106.5	0.032689	73.4	106.8	0.032553	-2.8570E-7	9.6639E-7	5.4177E-6														
362	U1480	H	4	H	3		118.6	27686	1645	1563	6.48	49.47	48.67																							
362	U1480	H	4	H	3		121.1	27711	1645				54.33																							
362	U1480	H	4	H	3		123.6	27736	1538				52.67																							
362	U1480	H	4	H	3		126.1	27761	705				37.00																							

Appendix 2A – SM1_FA_Geology – Facies Associations of the Nicobar Fan

SM1_FA_Geology **– Facies Associations of the Nicobar Fan**

Based on drilling of two sites, U1480 and U1481 (MCNEIL et al., 2017) divided the Nicobar, from base to top, in Units IIIB, IIIA, IIC, IIB, IIA, I, which they are summarized in the Fig. 30. The Facies Association (FA) here follows the description of PICKERING et al. 2019 who described eight facies associations for the Nicobar Submarine Fan sediments and the linkage to the described of Nicobar Fan Units.

The basal sedimentary package overlies in unconformity the Late Cretaceous to Paleocene pre-Fan volcano-sedimentary units and corresponds to the Unit IIIB containing chalk and tuffaceous mudstones siliciclastic sediments deposited between early Oligocene to early Miocene (MCNEIL et al., 2017, PICKERING et al., 2019).

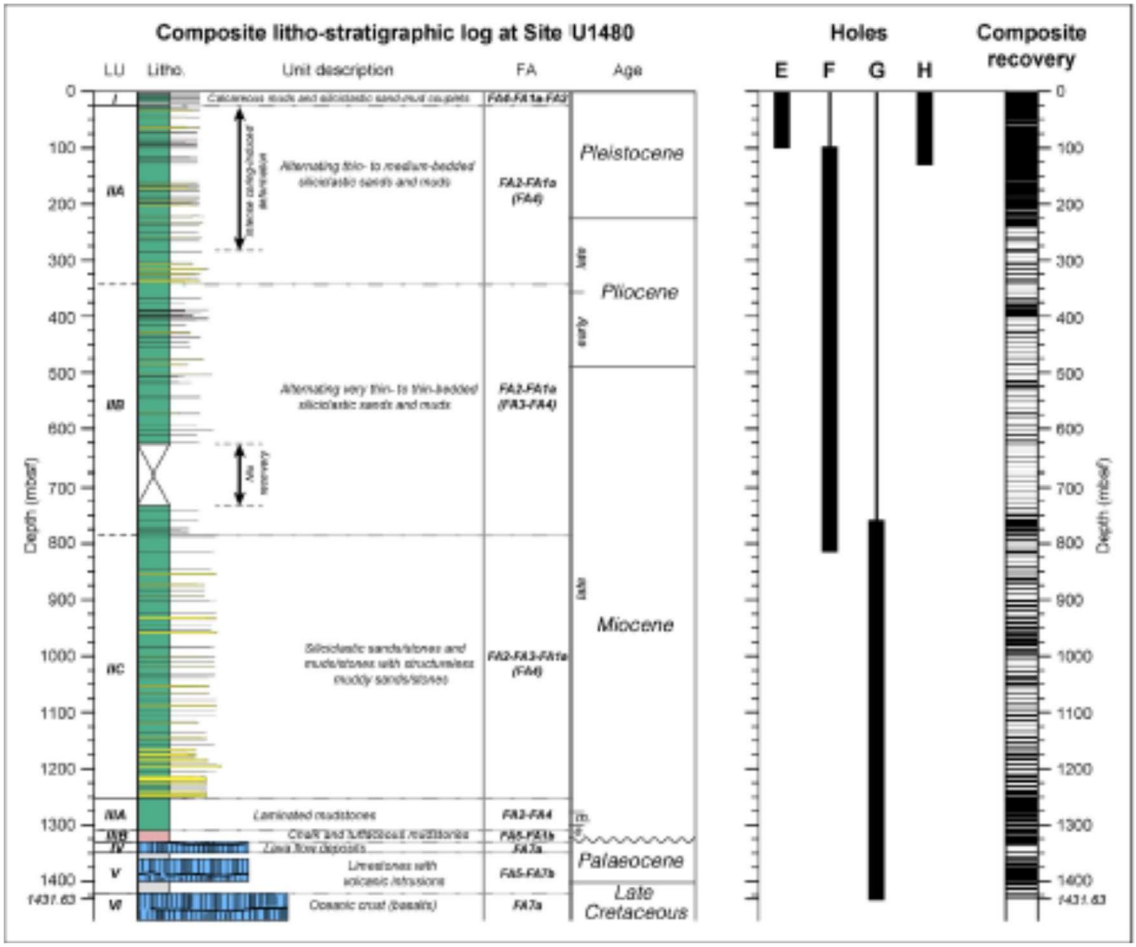
The Unit IIIA (middle Miocene) comprises interlayered thin to medium bedded, gray-green or brown mudstone and siltstones deposited in low-concentration turbidity currents.

The Unit II is the thickest section of the Nicobar Fan, reaching up to 1250 m in the site U1480. The basal unit, the Unit IIC (Upper Miocene), overlies in conformity the Unit IIIA. It consists of siliciclastic sandstone/siltstone and mudstone interlayers with structureless muddy sandstones and subordinate hemipelagic sediments. They are interpreted as structureless mud hemipelagites (FA2) deposited in low-concentration turbidity currents, sediment gravity flows (SGF, FA3) and turbidity sediments (FA1a). Hemipelagites deposited in low turbidite currents occur a common facies association (FA4).

The Unit IIB deposited between Upper Miocene and early Pliocene contains alternating very thin- to thin-bedded siliciclastics sands and muds represented by muddy turbidites (FA2), SGF`s (FA3), and sandy turbidites (FA1A). Structureless muds (hemipelagites, FA4)) are also present in the unit IIB.

The upper unit, the Pliocene-Pleistocene Unit IIA, consists of alternating thin to medium bedded siliciclastics sands and muds represented by the turbidites of Facies Associations FA2-FA1a (FA4).

The upper most unit, 26-m thick Quaternary Unit I, comprises biogenic calcareous mud with ash beds and siliciclastic fine-grained sand and muds.



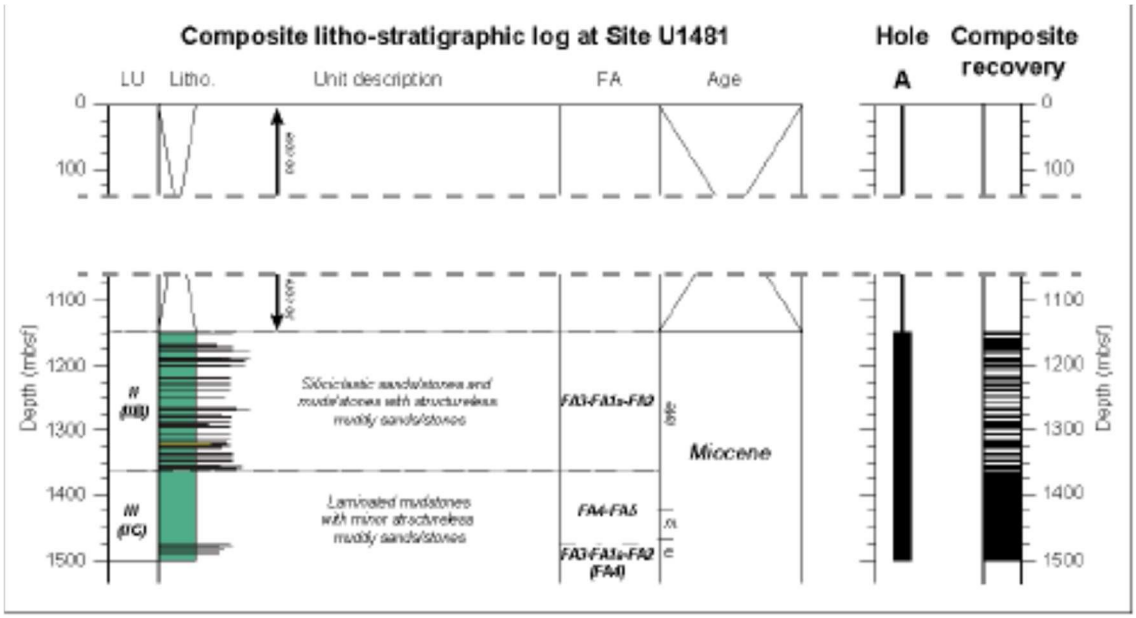


Figure 30 Simplified lithostratigraphic columns of cores at IODP-Expedition 362, sites U1480 and U1481, with the division between units, Facies Association (FA), age and composite recovery. Composite recovery represents the recovery of samples per lithological unit and describes an organization of the structure taken from the seabed. (after MCNEIL et al., 2017 in PICKERING et al., 2019).

Appendix 2B – SM2_material – Description of material

SM2_material – **Description of material**

GRA is a geophysical data acquired by the sensor Whole-Round Multisensor Logger (WRMSL) using a Cs-137 collimated source and NaI scintillation detector, whose values are calibrated using water standard and aluminum standard (BLUM, 1997; MCNEILL et al., 2017). In this article, GRA has a unit of measurement in g/cm^3 in a single feature used.

MAD is the geophysical data that measures the relationship between mass and volume. Core samples of approximately 8 cm² are processed by combining 4 methods (A, B, C, D). The wet mass, by combining methods A, B, C and dry mass (A, B, C, D)

is determined by the use of movement signals by the Mettler-Toledo balance (BLUM, 1997; MCNEILL et al., 2017). In this article, MAD has a measurement unit in $wt\%$, g/cm, and $vol\%$ with 6 features used.

MS is the measure of the degree to which a material can be magnetized at a specific external location. In practice, the susceptibility measurement requires calibration considering factors established by the core geometry, type of material, and core profiling type (BLUM, 1997; MCNEILL et al., 2017). In this article, MS has measurement unit in $instr.units$ in a single feature used.

NGR is acquired through scintillation detectors arranged along with the core interval at every 20 cm distance. Electromagnetic waves (Gamma Rays) act on the frequency between 1019 and 1021 Hz emitted spontaneously by an atomic nucleus during the radioactive decay, divided into packages called Photons. Photons are related to the energy transported (E) and the wavelength (λ) or frequency V. (BLUM, 1997; MCNEILL et al., 2017). In this article, NGR has measure units in cps in a single feature used.

PWL is the reading value of the ultrasonic P-wave velocity, transmitted in a 500 kHz wave pulse, through the core section at an interval and specific repetition. The sonic speed is the relation between the distance covered by the rocky material and the signal's duration in the same material (in seconds) (BLUM, 1997; MCNEILL et al., 2017). In this article, PWL has the unit of measure in m/s in a single feature used.

RGB is the measure of the values of the Red, Green, and Blue channels extracted from the digitization of the core, with reading intervals between 0.04 cm to 1 cm and values ranging from 0 to 255 (0 is completely black and 255 is entirely white) (MCNEILL et al., 2017). In this article, RGB has the unit of measure in R, G, B with three features.

RSC is the measurement of spectral counts in the range of 380 to 925 nm, using the light spectrophotometer, acting on the entire spectrum of the visible (BLUM, 1997; MCNEILL et al., 2017). In this article, RSC measures units in Reflectance L*, Reflectance a*, and Reflectance b* with three features used.

2.1 Machine Learning and Random Forest

Machine Learning (ML) is a set of computational techniques capable of inducing the computer to discover new knowledge (GÉRON, 2017). Its conception integrates finding hidden knowledge or in a more agile and practical way than traditional human techniques. ML can be classified into two main categories regarding training and type of supervision: supervised and unsupervised (GÉRON, 2017; GOLLAPUDI, 2016). In this article, the records of geophysical data extracted from IODP-Expedition 362 are supervised learning, data classification, and the RF method. The records extracted from the images, in their design, are selected in the category of unsupervised learning using the SLIC Superpixel. It is important to note that for the images, the dataset is organized in a training format supervised by the addition of a specific label of the type or group of lithology in each segmented region.

The RF method in the supervised learning category was based on the decision trees method. It is a set learning technique organized in the combination of many trees, based on different data samples and resource combinations (depth and number of nodes) (SINNOTT and SUN, 2016; KOTU and DESHPANDE, 2019). The RF creates multiple combinations of decision trees and averages, processing each combination of tree and nodes. For each iteration in the execution of the algorithm, the RF method selects the best combination (model) of training configuration. It stores its result

(bagging) for later use during its execution with the nodes' choice and configuration using the Gini or Entropy functions.

The RF method is applied in several areas, especially in lithological classification (XIE et al., 2020; AO et al., 2020; BRESSAN et al., 2020; KUMAR et al., 2019).

2.2 Interpolation

The method allows estimating or approximating a value or range of values from a set of specific multivariate data related to a given theme, based on the context of sampling or experiment (CELANT and BRONIATOWSKI, 2016; FARRELL, 2018). In geology, interpolation processes are necessary due to the characteristics of the data collected on the rocks, either by the type of lithology, the range of data collected, or the geophysical data used.

According to CHEN et al., 2018 the geophysical data used in this work are of irregular sparse data. Irregular because they are collected at irregular depths between all the properties analyzed. The reading intervals are to site U1480: RGB: 1 cm, GRA, MS, RSC and PWL: 2.5 cm, NGR: 10 cm and MAD: non-standard and site U1481: RGB: 0.04 cm, GRA, MS, RSC, and PWL: 2.5 cm, NGR: 10 cm and MAD: non-standard. They are sparse because the data reading intervals create large gaps in depths without records in several cases. Based on the type of data sampled, it was defined to use the following interpolation methods: Linear, Quadratic, Cubic, Spline, Slinear, Akima, Pchip and Piecewise.

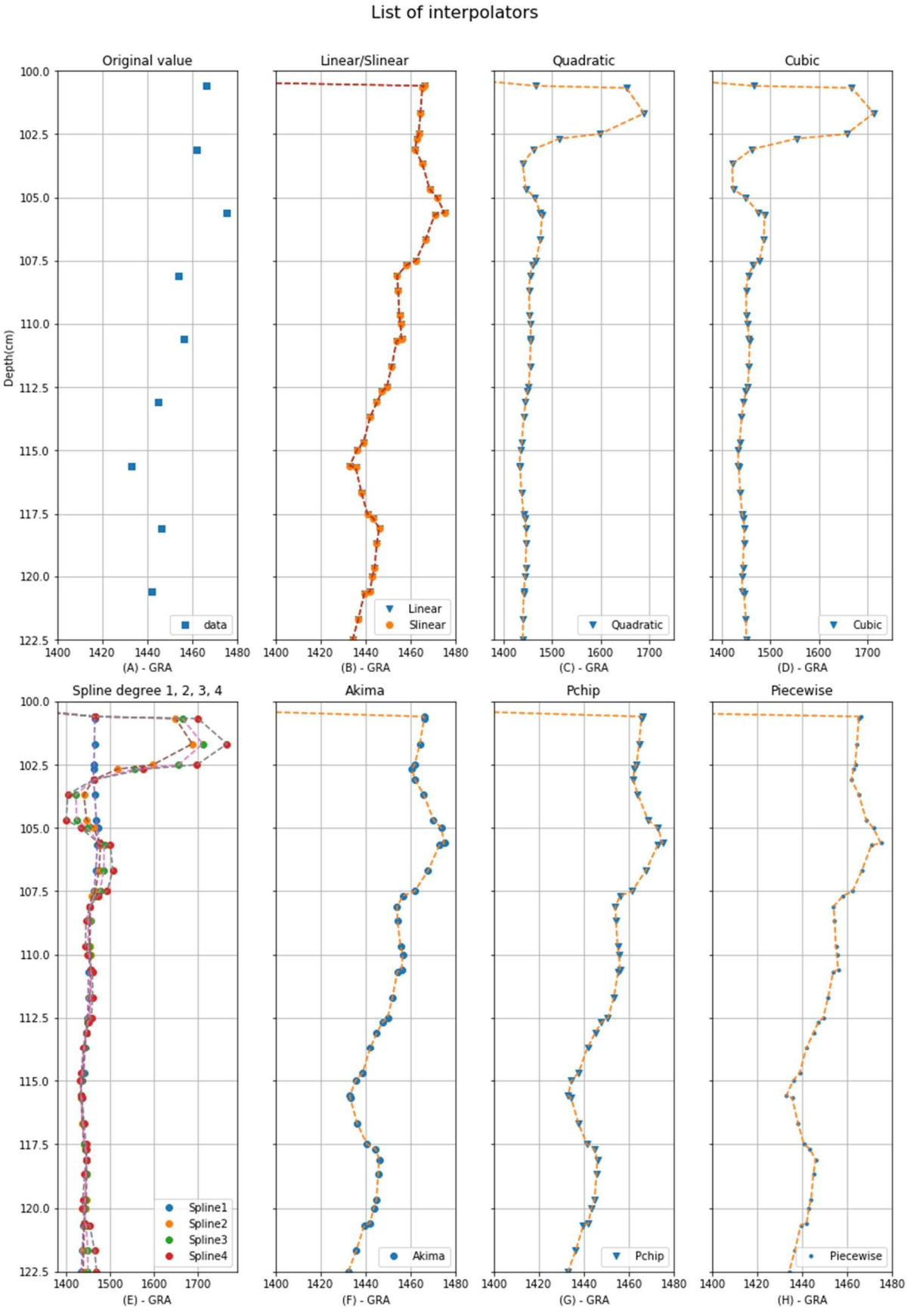


Figure 31 List of interpolators. Data example: GRA geophysical data, site U1480, hole E, Core 1H, Section 2 (range 100 cm to 122.5 cm). The depth is measured in cm according to the shipping

definitions. The geophysical data GRA, MS, RSC, PWL, NGR, MAD and RGB have specific range and reading values (column y of this figure). (A) are the points of reading the original data. (B) are the data interpolated by the Linear and Slinear interpolator. (C) are the data interpolated by the Quadratic interpolator. (D) are the data interpolated by the Cubic interpolator. (E) are the data interpolated by the Spline interpolator degree 1, 2, 3 and 4. (F) are data interpolated by the Akima interpolator. (G) are data interpolated by the Pchip interpolator. (H) are data interpolated by the Piecewise interpolator.

Linear

Linear interpolation or Interpolation 1-D is the process of interpolation between two points in the same dimension (SALOMON, 2006; PROYAN and KISELEV, 2010). A line is used between two neighboring samples and the appropriate point is calculated along that line according to the interval defined in the depth line (cm) as shown in the Fig. 31 - B e Fig. 32.

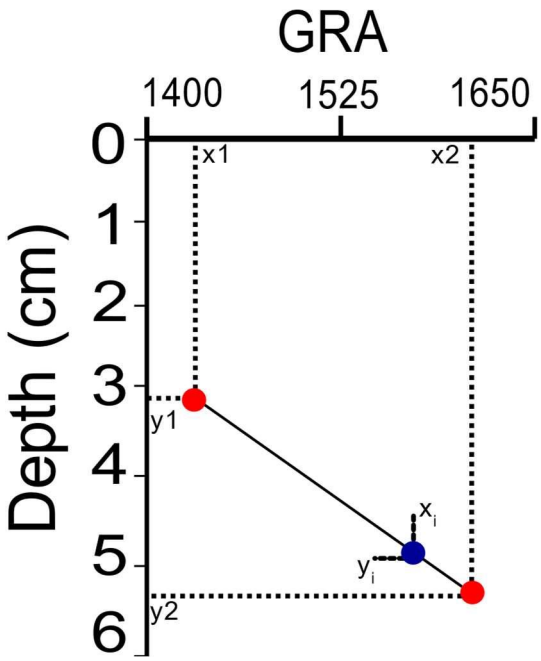


Figure 32 Value interpolation model using linear interpolator. Example for GRA geophysical data, site U1480, hole E, Core 1H, Section 2 (range 0cm to 6cm). The depth is measured in cm according to the shipping definitions. The geophysical data GRA, MS, RSC, PWL, NGR, MAD and RGB have specific range and reading values (column y).

Column x is defined as the value of the geophysical data, in this case it is GRA, and line y is the depth value. y1, y2, x1, x2 are original values obtained by the IODP-Expedition. The equation for calculating the interpolation value x_i at depth x_i is defined as:

$$x_i = y1 + (x_i - x1)(\frac{y2-y1}{x2-x1}),\ y1 < x_i < y2. \quad (1)$$

Quadratic and Cubic

Quadratic interpolation and cubic interpolation are performed using polynomials of degrees 2 and 3 at different points p'$_0$, p'$_1$, ..., p'$_n$ (Fig. 31 - C, D) each with polynomial function $fp_{p'_n}$ defined by (PROYAN and KISELEV, 2010; BUZZI-FERRARIS and MANENTI, 2010):

$$fp_{p'_n} = a_0 + a_1 x + a_2 x^2, \text{ for degrees 2} \quad (2)$$

$$fp_{p'_n} = a_0 + a_1 x + a_2 x^2 + a_3 x^3, \text{ for degrees 3} \quad (3)$$

Where a_0, a_1, a_2 and a_3 are the pairs of coefficients to be calculated, in this case, depth and value of the geophysical data with interpolation flow from the smallest to the largest in relation to the reading value of the geophysical data and the depth, and with quantity reading values of the geophysical data greater than two units per core and section.

The list of original points (geophysical data) are organized in a matrix where each ready (matrix line with three point interpolation (for degrees 2) and four point interpolation (for degrees 3)) represents the calculation of the $fp_{p'_n}$ function resulting in a matrix of non-zero determinant and admitting a single result.

Spline

Spline interpolator is composed of a set of polynomial functions that are connected by certain nodes in certain sections of a curve defining that at each point of interpolation two polynomials connect and their first derivatives must have the same value as all derivatives (K- 1) must be continuous (SALOMON, 2006; LYCHE and MØRKEN,

2018). Spline acts in the processing of curves and smoothing being related to its data set (points) and its functions (formulas).

The interpolator is classified according to the combination of the type of degree (K) of the polynomials present, its complexity in the analysis of the combination of points and smoothing of the nodes or the curve in the interpolated section, with the cubic spline being the most used. In this article, combinations of the spline in the degree of the polynomial were used ($K - 1, 2 \leq K \leq 5$).

Spline degree 1 or Spline Linear consists of several linear polynomials ($Sp_{(x)}$) joined in order to achieve continuity between the original points and the interpolated points. Fig. 31 - E and Fig 33 represents visual analysis in the application of the grade 1 spline and Eq. 4 describes how the operation is performed.

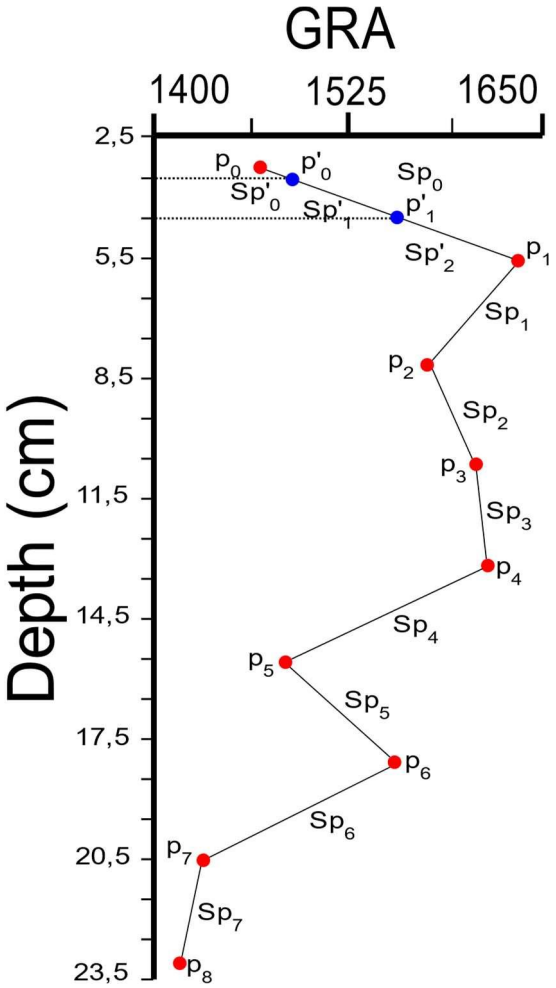


Figure 33 Value interpolation model using Spline degree 1 interpolator. Example for GRA geophysical data, site U1480, hole E, Core 1H, Section 2 (range 2.5 cm to 23.5 cm). The p_i (0 to 8) points are original values of the property reading on the expedition. The p'_i (0 to 1) points are values to be interpolated according to the defined depth interval setting. Sp_i and Sp'_i are segments between the original points and points calculated according to the quantity and degree of polynomials using, resulting in the shape of the connection (line or curve).

$$Sp'_{(x)} = y_i + m_i(x - p_i) = y_i + \frac{y_{i+1} - y_i}{p_{i+1} - p_i}(x - p_i) \quad (4)$$

where $Sp'_{(x)}$ is the segment to be calculated according to the depth value, x is the geophysical data value, m_i is the slope value of the curve calculated by the derivative, y_i is the depth value (in cm) and p_i is the point value (interpolated value found).

Importantly, $Sp_{(x)}$ is an interval between min (x) and max (x), continuous between max (x) and min (x), and min (x) = p_0 < p_1 < p_2 ... = max (x), such that $Sp_{(x)}$ is a linear polynomial at each interval between points p_i.

Splines of higher degree such as degree 2, 3 and 4 are applied at the approach level when it requires greater smoothness in calculating the curve. By the degree of the polynomial (K), it is necessary to determine the value of K-1 control points on the position of $Sp'_{(x)}$ resulting in an accuracy in the angle and shape of the curvature of the new segment. The $Sp'_{(x)}$ function for splines of higher degree follows the linear spline pattern by adjusting the quantity and degree of polynomials in each $Sp_{(x)}$ segment.

Slinear

The Slinear interpolation method acts in a similar way to the linear interpolator and spline interpolator degree 1 (Fig. 31 - B) using the same calculation bases (equations). Its main characteristic is the interpolation of values according to a pattern (within a defined range) of the starting and ending points (SALOMON, 2006). Outliers are discarded, making it impossible to continue interpolation. Slinear depends on the values being in a logical sequence of operation, their interpolation order follows the flow $x \rightarrow x$, where x is the value of the geophysical data with an initial value in relation to the non-zero depth.

Akima

Akima or Sub-Spline interpolator is a method of interpolation in the form of a cubic polynomial in parts acting in a univariate manner similar to a Spline degree 3 (AKIMA, 1986; FEDOROV, 2013), according to the equation:

$$fp_{p'_i} = a_i + b_i\,(x - x_i) + c_i\,(x - x_i)^2 + d_i(x - x_i)^3 \quad (5)$$

Where a_i, b_i, c_i and d_i are determined through the derivative of each point p_i. x is the value of geophysical data and x_i is the value of the geophysical data to be calculated at the point p'_i.

The method is based on a function composed of a set of polynomials of degree 3 and applied to successive data point intervals. It acts in order to estimate the first derivative of the function at each point (respective slope of the curve) based on the analysis of up to six reference points according to Fig. 31 - F and Fig. 34. The resulting curve adapts to various types of univariate data with the presence of several original random points.

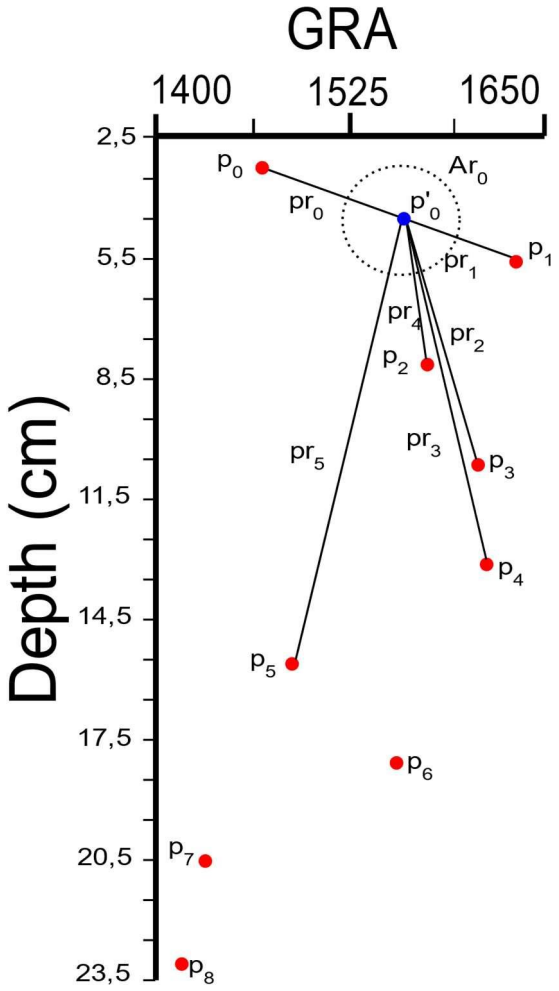


Figure 34 Value interpolation model using Akima interpolator to a point p'$_0$ with six reference points. Example for GRA geophysical data, site U1480, hole E, Core 1H, Section 2 (range 2.5 cm to 23.5 cm).

p_i are points with reading values between depth and geophysical data. p'_i point to be interpolated. Ar_0 is the possible location area for the interpolated point. pr_i are the connections to the reference points p_i.

Contrary to the Spline Degree 3, the Akima method does not require continuity of the second derivative, does not act in smoothing the curve, reduces the oscillation that Spline usually produces and uses the other derivatives as free parameters between the location and the reference points analyzed.

Pchip

Piecewise Cubic Hermite Interpolating Polynomial (Pchip) is an interpolator that uses data in the cubic Hermite format by dividing the interpolation into equal parts or subsets of cubic polynomials. Pchip has a more suitable application in relation to cubic Spline if the data has flat and steep sections preserving the geometry of the data location and the monotonicity between the points (RABBATH and CORRIVEAU, 2019; FRITSCH, 1982). Pchip identifies 4 different points and analyzes their slopes (through derivatives degree 1, 2 and 3) and the average between the connections of the points through a linear interpolator returning the value of the interpolation function as defined in the equation fp_x:

$$fp_x = d_0 + d_1 * x + d_2 * x^2 + d_3 * x^3 \quad (6)$$

Where x is the point to be calculated, d_0, d_1, d_2 and d_3 are derived from the four points analyzed, respecting the value with the respective sign (negative or positive).

The main characteristics of the Pchip are in the sense that it analyzes the slope of the four points analyzed using its derivatives, respects the monotonicity between the data intervals, presents less overshoot / undershoot than the Spline interpolator and

presents itself as a method to perform curve adjustments for all data points (Fig. 31 - G).

Piecewise

Piecewise interpolator interpolates the value between two distinct irregular points using polynomials in parts and respective derivatives of the points (FEDOROV, 2013). It uses the same equation as the Spline interpolator, the definition of the degree of the polynomial is dynamic and adapts to the degree necessary to cover all derivatives at the analyzed points (or interval).

Piecewise seeks to use an equal number of derivatives (polynomial degree) at each point in the interval. If the number of possible derivatives is different, it will use the lowest value or the value of the righter point of the segment (curve). An exception will be reported if it is not possible to calculate the derivatives of the analyzed points or the number of derivatives for interpolation results in high values making an appropriate calculation impossible (Fig. 31 - H).

Operation flow of the interpolators

It is important to highlight that each interpolator analyzes the original data range of the geophysical data and performs the operation flow for interpolation according to the location characteristics between depth and number of features present by dataset in relation to the core and section. The table 14 shows the operation flow of each interpolator based on the geophysical data used in this article.

Table 14 Operating flow of interpolators.

Interpolator	Operation Flow
1. Linear	$x \leftrightarrow x$

2.1 Quadratic	$x \rightarrow x$
2.2 Cubic	$x \rightarrow x$
3. Spline	$x \leftrightarrow x$
4. Slinear	$x \rightarrow x$
5. Akima	$x \rightarrow x$
6. PChip	$x \leftrightarrow x$
7. Piecewise	$x \rightarrow x$

The operation flow described in the interpolators comprises the depth range of the hole, core and section respectively with the sequence of features of the GRA, MAD, MS, NGR, PWL, RGB and RSC geophysical data.

The Linear interpolator being a 1-D dimension interpolator processing data between two variables (two points p_i) covers the creation of new data interpolated in the flow $x \leftrightarrow x$ of greater or lesser value of x, where x is the value of the geophysical data in relation to depth.

Quadratic, Clubic and Piecewise interpolators are interpolators that, in their design, use polynomials in the interpolation processing, requiring initial values for the interpolation flow by performing the operation in the $x \rightarrow x$ direction where x is the geophysical data value in relation to non-zero depth and reading value greater than two units per core and section.

Slinear interpolator, as already described, by the initial and final data range characteristic follows the operation flow $x \rightarrow x$ where x is the value of the geophysical data with initial value in relation to the non-zero depth.

Spline and Pchip interpolator cover interpolation with polynomial functions in calculating the segment between the points, but the continuity between the other points will occur through the processing of a linear polynomial reaching an operation flow of the interpolation equal to the Linear Interpolator.

Akima interpolator uses a set of polynomials in parts grouped according to the distribution of the points and their connections. In addition to the characteristic, it requires reference points when designing the interpolation without making it impossible to perform the operation. The operation flow follows the logical sequence $x \rightarrow x$ where x is the value of the geophysical data with an initial value in relation to the non-zero depth.

The validation of the interpolated data will occur by applying the classification method according to the defined training and testing configuration, seeking to identify the best accuracy result in the interpolated data range.

Examples of the application of numerical interpolation techniques in seismic data (CHAI et al., 2020; TURCO, AZEVEDO and HEROLD, 2019), in mineral exploration (GUO et al., 2018), and oil and gas exploration (NGUIMBOUS-KOUOH and MANGUELLE-DICOUM, 2019).

2.3 Image Data Annotation

Visual marking or human annotation on image, video, or audio indicating an area of interest. Each area of interest is defined as a region being automatically or manually indicated and can have the shapes of rectangle, circle, ellipse, polygons, point and polyline (DUTTA and ZISSERMAN, 2019; TORRALBA, RUSSELL and YUEN, 2010). The textual description of each region of interest is the key point of annotating the image collecting information in depth range and lithology name.

This description includes marking tags indicating this relevant information for further processing. Semi-structured data type is used as output, which allows changes to the

structure and content at run time, forming a dynamic data structure integrated with the analyzed data source. Semi-structured data can be in .json, .xml or .csv file format.


```
" 362-u1480e-2h-2-
a_shlf7853651_20160813145856.jpg6748938": {
   "filename": " 362-u1480e-2h-2-
a_shlf7853651_20160813145856.jpg",
   "size": 6748938,
   "regions": [
    {
     "shape_attributes": {
      "name": "polygon",
      "all_points_x": [
       400,
       1630,
       1660,
       1120,
       710,
       410,
       400
      ],
      "all_points_y": [
       1590,
       1600,
       4930,
       4840,
       4890,
       4960,
       1590
      ]
     },
     "region_attributes": {
      "Lithology": "Medium-sand_sandstone",
      "Depth": "0.5_17.0"
     }
    }
   ],
   "file_attributes": {}
 }
```

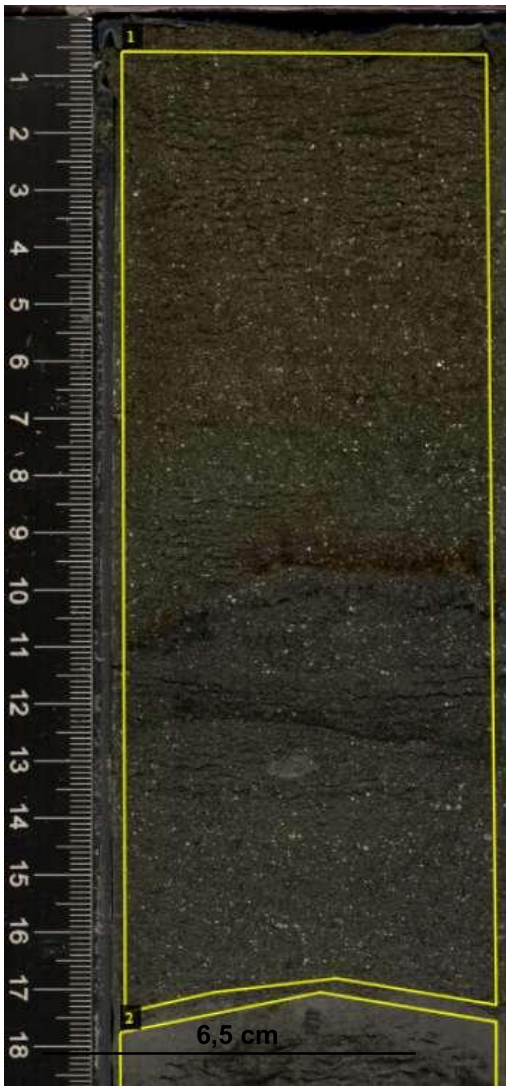



Figure 35 Example of .json file format content of the main tags used in this study. The "name" tag defines the shape of the region of interest. The tags “all_points_x” and “all_points_y” are X and Y pixel coordinates of the analyzed image forming the polygon with 7 coordinates. The tag "region_attributes" includes the attributes defined for "Lithology" and "Depth". Attribute "Depth" is measured in cm, in the format: start_and separated by ‘_’. The example image belongs to site U1480, Hole E, Core 2h, Section 2. (DUTTA and ZISSERMAN, 2019).

The json output file type is widely used in markup structures that require quick, secure and easy handling (Fig. 35). It is a format widely used on the internet and compatible with unstructured databases. (HAMOUDA and ZAINOL, 2019; STREKALOVA and BOUAKKAZ, 2017; MARTINEZ-MOSQUERA, NAVARRETE and LUJAN-MORA,

2020). The file name is defined as a primary initial tag designated as an access key. The tags "filename", "size", "regions" and "file_attributes" define the primary structure with information from the analyzed file (image, video or audio). The subtag “shape_attributes” stores information on areas of interest with pixel X and Y coordinates, in addition to specific user-defined attributes, which in this case are “lithology” and “depth” attributes.

The annotation of images together with ML are appropriate tasks for the visual processing of images, being included in procedures in the health area such as disease identification (ELAZIZ et al., 2020; SHI et al., 2019) and geology such as mineral identification and seismic processing (RAN et al., 2019; MAITRE, BOUCHARD and BÉDARD, 2019)

2.4 Segmentation SLIC Superpixel

Simple Linear Interative Clustering (SLIC) or SLIC Superpixel is a method of grouping or segmenting pixels, based on the k-means method, acting more quickly and reducing the complexity of tasks in image processing. The grouped areas are called Superpixel regions or segmented regions having their characteristic identification, location, and separation as for future processing (ACHANTA et al., 2010; ACHANTA et al., 2012; KAVZOGLU and TONBUL, 2018). In Fig. 36, segmentation of an example image from the lithology Medium-sand_sandstone.

Figure 36 Example of lithology image Medium-sand_sandstone, original file: 362-u1480e-2h-2-a_shlf7853651_20160813145856.jpg, segmented by the SLIC Superpixel resulting in 26 regions. Apx: 115,07 cm². Original image size: 1700 x 32007 pixels Resolution: 96 DPI. Bits intensity: 24. The grouping of pixels takes place in an appropriate way to the process, covering the entire image, separating characteristics and textures relevant to each lithology.

The computational logic for the operation of the SLIC Superpixel method begins with the definition of the maximum number of K groups with approximately the same pixel size. For color images in the CIELAB space, the procedure starts with the step of defining the initial cluster centroid $C_i = [L_i a_i b_i x_i y_i]$ where C_i are arranged in a regular G pixel grid. To generate SLIC Superpixels of approximate sizes, the G= √(N/K) where the centroids are located in a lower gradient pixel region, avoiding the centroid being close to the edge of the Superpixel or having a noisy pixel (ACHANTA et al., 2012; KAVZOGLU and TONBUL, 2018).

When grouping pixels, each pixel i is associated with the cluster centroid C_i nearest, whose search region is related to the distance measurement Da which determines the cluster center closest to each pixel analyzed, as seen in Eq. 7 and Fig. 37.

$$Da = d_{Lab} + \left(\frac{d_{x,y}}{G}\right)m \tag{7}$$

Where Da is the distance between the centroid C_i and the pixel i, m is the compression between the maximum and minimum color distance in the Superpixel, d_{Lab} are the color values in the CIELAB space, G is grid interval and $d_{x,y}$ is the value of the row and column in pixels.

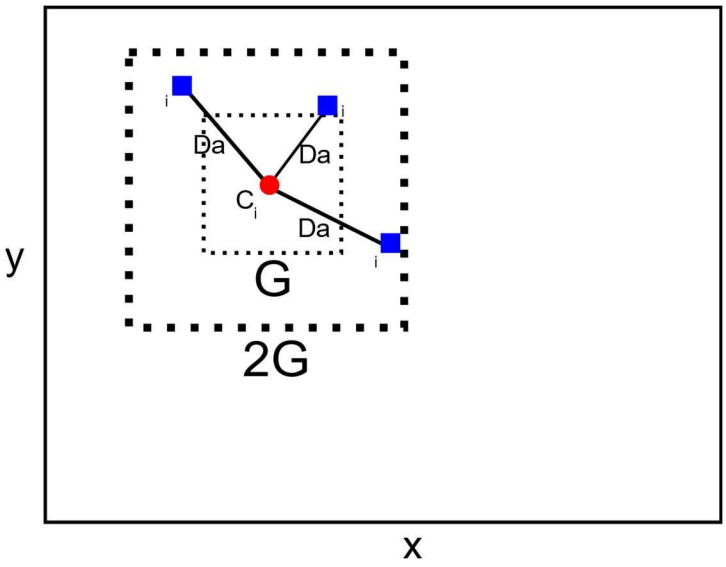


Figure 37 SLIC processing model in the allocation of a pixel i in the centroid C_i according to the distance Da in the region 2G x 2G. x and y are defined as rows and columns of the image, measured in pixels. The size x, y is variable and depends on the image analyzed.

As the size of the expected Superpixel space region is approximated to G x G, the search for similar pixels is done in a 2G x 2G region around the centroid C_i. The main point to speed up the analysis is here because the processing is limited to the size of the research region reducing the calculation of a and increasing the speed of the

grouping in relation to the conventional analysis using K-means in which each pixel i is compared with all centroid C_i. At the end, after assigning all pixels to the centroid, the adjustment step is processed where for each centroid C_i the average value of its location in relation to the total grouped pixel is calculated. The SLIC method acts to minimize the error of grouping and location of pixels in relation to the centroid through testing and updating steps using the Euclidean Norm or L2 Norm.

The distance Da (Eq. 7) represents processing in the CIELAB color space, whose range of values varies in combinations of L_i which is the luminance of black to white, a_i as being the combination of green to red and b_i as being the combination of blue to yellow, with the pixel position in line and column (x, y). The range of CIELAB values varies with the type, size of the image and number of colors present.

SLIC based on the K-means method, being a faster iterative grouping, more efficient in using computational memory, exhibits adherence between the targeted regions and improves the performance of the segmentation algorithm.

2.4.1 Texture and Image color

Images collected in IODP-Expedition 362 are of the RGB type. The image's size is calculated through the resolution of the pixel quantity and the number of colors per pixel. IODP images, the resolution used is 96 Dots Per Inch (DPI) and 24 bits of color in the 3 channels. The size of the image (width and height) is variable and depends on the sample size collected at the time of drilling (MCNEILL et al., 2017).

The segmented regions follow the RGB image pattern, processed individually, maintaining all the initial characteristics (Fig. 38).

Figure 38 Segmentation region extracted from the image: 362-u1480e-2h-2-a_shlf7853651_20160813145856.jpg, lithology: Medium-sand_sandstone cod_region: 22, mean_intensity color: 10.03, Apx: 3.56 cm², Pixel count: 134368

Each region, the $mean\_intensity$ feature is extracted and stored in an appropriate dataset and calculated by the equation:

$$mean\_intensity = \frac{\frac{\sum_n^1 R}{n} + \frac{\sum_n^1 G}{n} + \frac{\sum_n^1 B}{n}}{3} \quad (8)$$

Where:

$mean\_intensity$ is defined as the average RGB color intensity in the segmented region. R, G, B are the color values ranging from 0 to 255, and n is the total number of pixels in the segmented region. Mean values of $mean\_intensity$ are float numbers and range from 0 to 255.

The brightness values or gray levels for the segmented regions are calculated using the Gray Level Co-occurrence Matrix (GLCM). GLCM is defined as the analysis and tabulation of gray level values at a given displacement in the image segment (HALL-

BEYER, 2017). The displacement of the pixel block in the image follows the pattern defined according to Fig. 39.

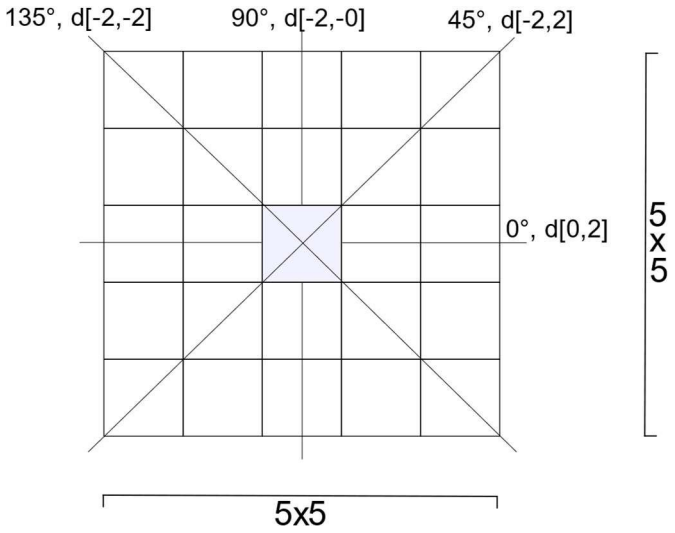


Figure 39 Symmetrical block 5x5 with the pixel of interest indicating the possibilities of angles (0 °, 45 °, 90 °, 135 °) and distance (d) for processing. The block maintains a symmetrical pixel distribution pattern with normalized gray values on the sum scale equal to 1 before calculating the texture properties.

The pixel of interest in GLCM follows the context of binary value (second-order) advancing and processing in groups of two pixels in the original image. The advance of the pixel of interest includes the configuration of two properties: angle and distance. For angle, there are four possible combinations of distribution in the horizontal, vertical and diagonal (0º, 45º, 90º, 135º). Distance comprises the list of the displacement of the pixel pair in the chosen angle defined as an array of initial and final values return a 4-d size matrix with the dimensions GLCM [i, j, d, Ɵ], where i and j are the gray values in the row and column dimensions, d is the distance from the pixel pair and θ is the angle in degrees.

The texture properties calculated using the GLCM are Contrast, Dissimilarity, Homogeneity, Angular Second Moment (ASM), Energy and Correlation.

According to the equations (HALL-BEYER, 2017; ALAZAWI, SHATI and ABBAS, 2019):

$$Contrast = \sum_{i,j=0}^{levels-1} P_{i,j}(i-j)^2 \quad (9)$$

$$Dissimilarity = \sum_{i,j=0}^{levels-1} P_{i,j}|i-j|^2 \quad (10)$$

$$Homogeneity = \sum_{i,j=0}^{levels-1} \frac{P_{i,j}}{1+(i-j)^2} \quad (11)$$

$$ASM = \sum_{i,j=0}^{levels-1} P_{i,j}^2 \quad (12)$$

$$Energy = \sqrt{ASM} \quad (13)$$

$$Correlation = \sum_{i,j=0}^{levels-1} P_{i,j}\left[\frac{(i-\mu i)\,(j-\mu j)}{\sqrt{(\sigma_i^2)(\sigma_j^2)}}\right] \quad (14)$$

Where:

i and j are defined as position between row and column of the matrix, P is the GLCM matrix [i, j, d, Θ]. The result of the properties is an array 2-d Property [d, Θ], with property value at the distance d and angle Θ.

$Contrast$ (Eq. 9) is the sum of the measurement related to the distance of the GLCM diagonal, allocating specific weights. They receive weight 0 when i and j are equal, weight 1 when the difference between i and j is equal to 1, weight 4 when the difference between i and j is 2 and thus increase exponentially as the difference between i and j increases.

$Dissimilarity$ (Eq. 10) follows the Contrast calculation pattern by modifying the weight value in a linear way.

$Homogeneity$ (Eq. 11) acts in order to weight the values by the inverse of the Contrast weights, that is, the weights decrease exponentially in relation to the GLCM diagonal. Its value ranges from 0 to 1.

ASM (Eq. 12) calculate the sum of the weight measurement of each position i and j. High values for ASM indicate high pixel ordering with the same values.

$Energy$ (Eq. 13) is calculated using the square root of ASM.

$Correlation$ or Linear Correlation (Eq. 14) is the sum of the linear dependence measure of the gray level values of the GLCM matrix with the closest neighbors. For a symmetric matrix, the correlation value is equivalent to variance (σ^2).

Research in geology how to use of SLIC Superpixel covers few practical applications of this method with studies in the field of rock morphology (MALLADI, RAM and RODRÍGUEZ, 2014), classification of mineral grains in microscopic images (MAITRE, BOUCHARD and BÉDARD, 2019; JIANG et al., 2017) and detection of lithological limits in images from remote sensing platforms (VASUKI et al., 2017).

2.5 Validation and Evaluation

Confusion matrix represents the visualization of the data classification performance according to the division between real data and predicted data, being the main way to tabulate the results of a classification, being present in its structure the metrics of evaluation accuracy, precision, recall and F_1-score (BRESSAN et al., 2020; NISBET, MINER and YALE, 2018), according to the organization shown in Fig. 40. The metrics for evaluating the confusion matrix are employed in a precise and correct way to classify geophysical data with excellent results in the binary classification of lithologies in IODP-Expeditions (BRESSAN et al., 2020).

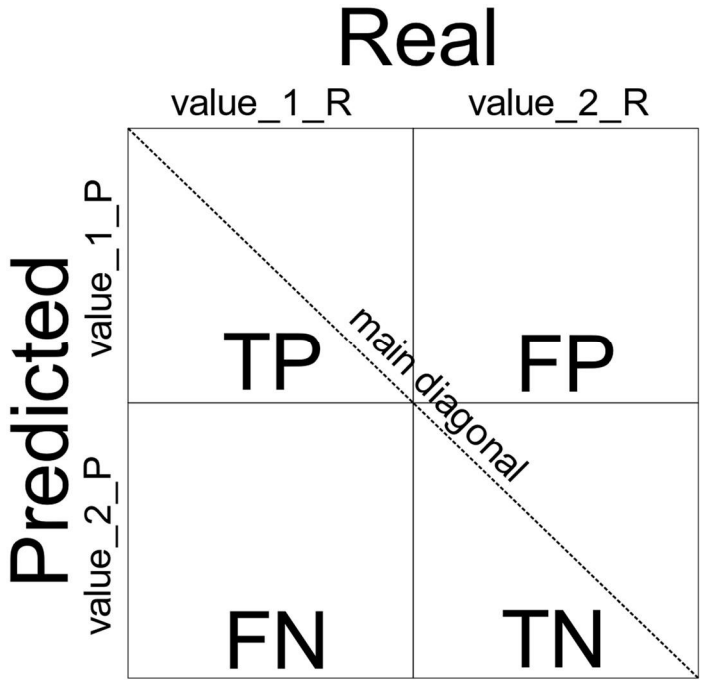


Figure 40 Organization of the confusion Matrix. The rows and columns are divided into predicted data and real data combining TP (True Positive), FP (False Positive), FN (False Negative) and TN (True Negative). The representation in the matrix format is TP [value_1_P: value_1_R], FP [value_1_P: value_2_R], FN [value_2_P: value_1_R] and TN [value_2_P: value_2_R].

The metrics $Accuracy$, $Precision$, $Recall$ and $F_1 - score$ are the results of sampling the classification on the confusion matrix, being defined by the following equations:

$$Accuracy = \frac{TP+TN}{P+N} \quad (15)$$

$$Precision = \frac{TP}{TP+FP} \quad (16)$$

$$Recall = \frac{TP}{TP+FN} \quad (17)$$

$$F_1 - score = 2 * \frac{\text{Precision} * \text{recall}}{\text{Precision} + \text{recall}} \quad (18)$$

$Accuracy$ (Eq. 15) is formed by the division between the sum of true positive and true negative values and the total sum of positive and negative values. The calculated value is presented as a percentage, from 0 to 100%, returning how much the model got right from the possible classification.

$Precision$ (Eq. 16) and $Recall$ (Eq. 17) are formed by the division between TP and the sum of TP and FP values and the sum of TP and FN values, respectively. These two metrics assess the model as to whether the proportion of identifications (class to be classified) is correct.

$F_1 - score$ (Eq. 18) uses values of $Precision$ and $Recall$ calculated by dividing the multiplication of $Precision$ and $Recall$ with the sum of the values of $Precision$ and $Recall$. The calculated value shows a balance between $Precision$ and $Recall$ in the classification of the proposed class and interpreted as the reliability value of the found accuracy (SARKAR, BALI and SHARMA, 2018).

Cross-validation is a method used to evaluate the classification data's performance using metrics from the confusion matrix. Cross-validation processing divides data into equal-sized groups called K groups or K folds. The divided groups are trained and tested in n possible combinations in the same configuration as the machine learning method used on all data (BRESSAN et al., 2020; GÉRON, 2017).

The results of the confusion matrix give rise to new analyzes related to error and success rates. Main values are related to True Positive Rate (TPR) and False Positive Rate (FPR) and visual analysis can be applied using the Receiver Operating Characteristic (ROC) method and its area on the ROC curve called Area Under the ROC curve (AUC) according to the equations (GÉRON, 2017; SUN et al., 2020):

$$TPR = \frac{TP}{TP+FN} \quad (19)$$

$$FPR = \frac{FP}{FP+TN} \quad (20)$$

The new metric applied in this article is called Area Superpixel (Apx). In this analysis, the images extracted from the IODP-Expedition are analyzed after processing,

segmentation and classification of the segmented regions according to the definition of the lithology groups, and the Apx is calculated as defined:

$$Apx = \frac{(A' * a')}{1000} \quad (21)$$

$$ApxTotal = \sum_{n}^{1} Acm \quad (22)$$

Where:

Apx is the resulting total area in cm² of the segmented region. A' is the total number of pixels in the segmented region. a' is the pixel size, in cm, according to the DPI resolution of the analyzed image (ACHARYA and RAY, 2005; GONZALES and WOODS, 2007). In this work, the images are in 96 DPI resolution, the value of a' is 0.02646 cm. $ApxTotal$ of the analyzed image is calculated (Eq. 22) by the sum of the Apx of all the segmented regions.

Appendix 2C – SM3_method – Description of methods – Figure and tables

SM3_method **– Description of methods – Figure and tables**

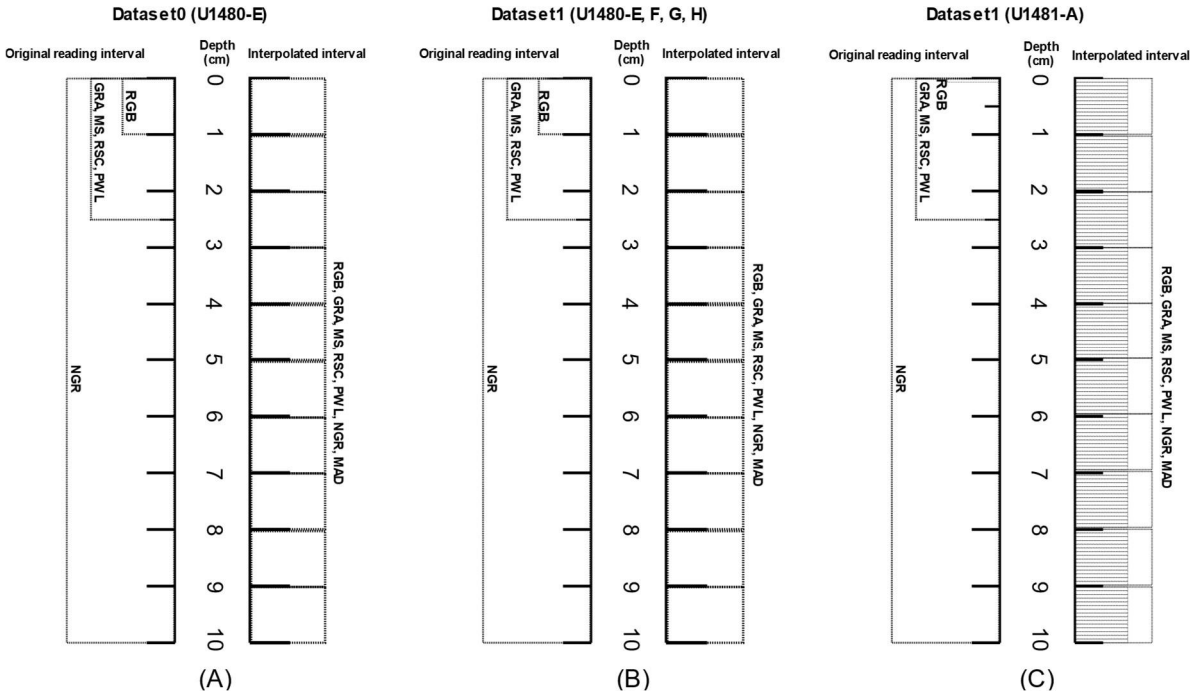


Figure 41 Original reading interval of the geophysical data according to the formation of the dataset. Dataset0 is formed by the original values of the geophysical data GRA, MS, RSC, PWL, NGR and RGB with values interpolated at a depth of 1 cm on site U1480, Hole E. Dataset1 (B) is formed by the original values of the geophysical data GRA, MS , RSC, PWL, NGR and RGB with values interpolated at a depth of 1 cm on site U1480, Holes E, F, G, H. Dataset1 (C) is formed by the original values of the geophysical data GRA, MS, RSC, PWL, NGR and RGB with values interpolated at a depth of 0.04 cm and 1 cm at site U1481, Hole A. Original value for the MAD geophysical data follows a different reading pattern than other properties.

Table 15: Organization, division and grouping of images IODP-Expedition 362 according to site, hole, total of original images, annotated images and lithology group.

Site	Hole	Total Original Images	Annotated Images	Lithologies present by Groups (cod_lit)
U1480	E	81	179	Group 1: 0, 2, 3, 5, 8, 9, 11, 12, 13 Group 2: 0, 1, 2, 3, 4, 5
	F	257	622	Group 1: 0, 1, 2, 3, 5, 8, 9, 10, 11, 12, 13 Group 2:

				0, 1, 2, 3, 4, 5
	G	247	746	Group 1: 0, 1, 2, 3, 5, 6, 8, 9, 10, 11, 12, 13 Group 2: 0, 1, 2, 3, 4, 5
	H	100	197	Group 1: 0, 1, 2, 3, 5, 8, 9, 10, 11, 12, 13 Group 2: 0, 1, 2, 3, 4, 5
U1481	A	128	857	Group1: 0, 1, 2, 3, 5, 8, 9, 10, 11, 12, 13 Group 2: 0, 1, 2, 3, 4, 5

Table 16: Division of lithologies by group and dataset with number of records interpolated by the Akima interpolator

Group	Cod_lit	Lithology or Group Litology	Record Dataset0: U1480-E	Record Dataset1: U1480	Record Dataset1: U1481	Record Dataset2: U1480	Record Dataset2: U1481
1	0	Alternating sand/sandstone and mud/mudstone layers	1046	7130	3626	6193	402
	1	Alternating silt/siltstone and clay/claystone layers	0	12776	78035	10823	6142
	2	Ash/tuff	3	258	163	175	7
	3	Calcareous clay/claystone	3330	5880	21	4020	0
	4	Calcareous silt/siltstone	0	0	0	0	0
	5	Clay/claystone	3833	22963	51347	18563	3893
	6	Coarse sand/sandstone	0	38		42	0
	7	Conglomerate	0	0	0	0	0
	8	Fine sand/sandstone	1508	11120	17567	9770	1619
	9	Medium sand/sandstone	2816	4232	10151	2856	808
	10	Sand/sandstone-silt/siltstone-clay/ claystone	0	472	0	436	0
	11	Silt/Siltstone	461	11524	5801	1895	487
	12	Silty clay/claystone	349	2280	34032	9492	2771

	13	Very fine sand/sandstone	2634	5336	11193	3729	935
2	0	Group 1, cod_lit: 0, 1, 10 (Alternating layers)		20378	81661	17452	6544
	1	Group 1, cod_lit: 2 (Ash/tuff)		258	163	175	7
	2	Group 1, cod_lit: 3 (Calcareous clay/claystone)		5880	21	4020	0
	3	Group 1, cod_lit: 5 (Clay/claystone)		22963	51347	18563	3893
	4	Group 1, cod_lit: 11, 12 (Silt/Clay)		13804	39833	11387	3258
	5	Group 1, cod_lit: 6, 8, 9, 13 (Sand)		20726	38911	16397	3362

Table 17: Interpolation intervals separated at site U1480 and site U1481 using the defined interpolator. The lowest reading resolution on the sites U1480 and U1481 is in the RGB geophysical data.

Site	Interpolador	Depth range for interpolation	Lower reading resolution
U1480	Akima	1,00 cm	1,00 cm
U1481	Akima	1,00 cm	0,04 cm

Table 18: Original records and interpolated records by Akima separated by site and core in each geophysical data used. The total number of records interpolated in U1480-E: 15980, U1480-F: 16538, U1480-G: 33876, U1480-H: 17615, totaling 84009. For the original records, the sum of records does not meet in all depths. In U1480-F, the RGB geophysical data does not require interpolation. The total number of records interpolated in U1481-A: 211936. For the original records, the sum of records does not meet in all depths. In U1481-A, the RGB geophysical data does not require interpolation.

Site	Hole	Original Records	Interpolated Records
U1480	E	GRA – 3333 MAD – 125 MS – 3333 NGR – 810 PWL – 3253 RGB – 8314 RSC – 3081	GRA – 12647 MAD – 15855 MS – 12647 NGR – 15170 PWL – 12727 RGB – 7666 RSC – 12899
	F	GRA – 8108 MAD – 166 MS – 8108 NGR – 1885 PWL – 6636	GRA – 8430 MAD – 16372 MS – 8430 NGR – 14653 PWL – 9902

		RGB – 19815 RSC – 6806	RGB – -3277 RSC – 9732
	G	GRA – 8753 MAD – 305 MS – 8753 NGR – 2066 PWL – 5578 RGB – 20940 RSC –10787	GRA – 25123 MAD – 33571 MS – 25123 NGR – 31810 PWL – 28298 RGB – 12936 RSC – 23089
	H	GRA – 4135 MAD – 107 MS – 4135 NGR – 994 PWL – 3530 RGB – 10230 RSC – 4099	GRA – 13480 MAD – 17508 MS – 13480 NGR – 16621 PWL – 14085 RGB – 7385 RSC – 13516
U1481	A	GRA – 4679 MAD – 172 MS – 4679 NGR – 1123 PWL – 3371 RGB – 250162 RSC – 4632	GRA – 207257 MAD – 211764 MS – 207257 NGR – 210813 PWL – 208565 RGB – -38226 RSC – 207304

Table 19: Formation of datasets with number of records per group and number of features used.

Dataset	Site-Hole		Group	Quantity interpolated records	Quantity features	Record interval
Dataset0	U1480-E		1	15980	17	depending on the interpolation configuration
Dataset1	U1480-E,F,G,H		1	84009	17	depending on the interpolation configuration
Dataset1	U1480-E,F,G,H		2	84009	17	depending on the interpolation configuration
Dataset1	U1481-A		1	211936	17	depending on the interpolation configuration
Dataset1	U1481-A		2	211936	17	depending on the interpolation configuration
Dataset2	U1480-E,F,G,H		1	67994	90	according to the segmentation interval via SLIC
Dataset2	U1480-E,F,G,H		2	67994	90	according to the segmentation interval via SLIC
Dataset2	U1481-A		1	17064	90	according to the segmentation interval via SLIC

Dataset2	U1481-A		2	17064	90	according to the segmentation interval via SLIC

Appendix 2D – SM4_results – Description of results - Figures

SM4_results **– Description of results - Figures**

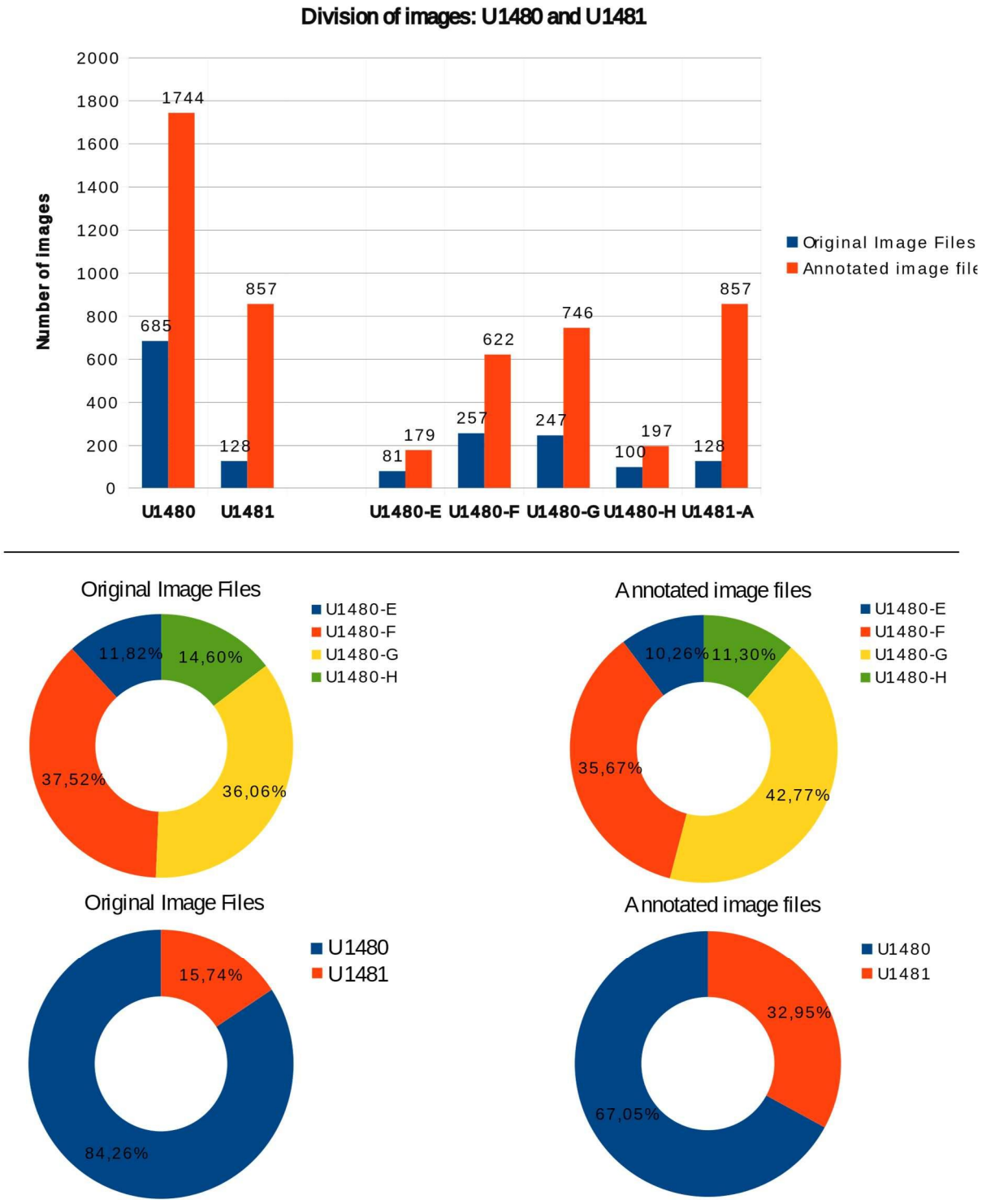
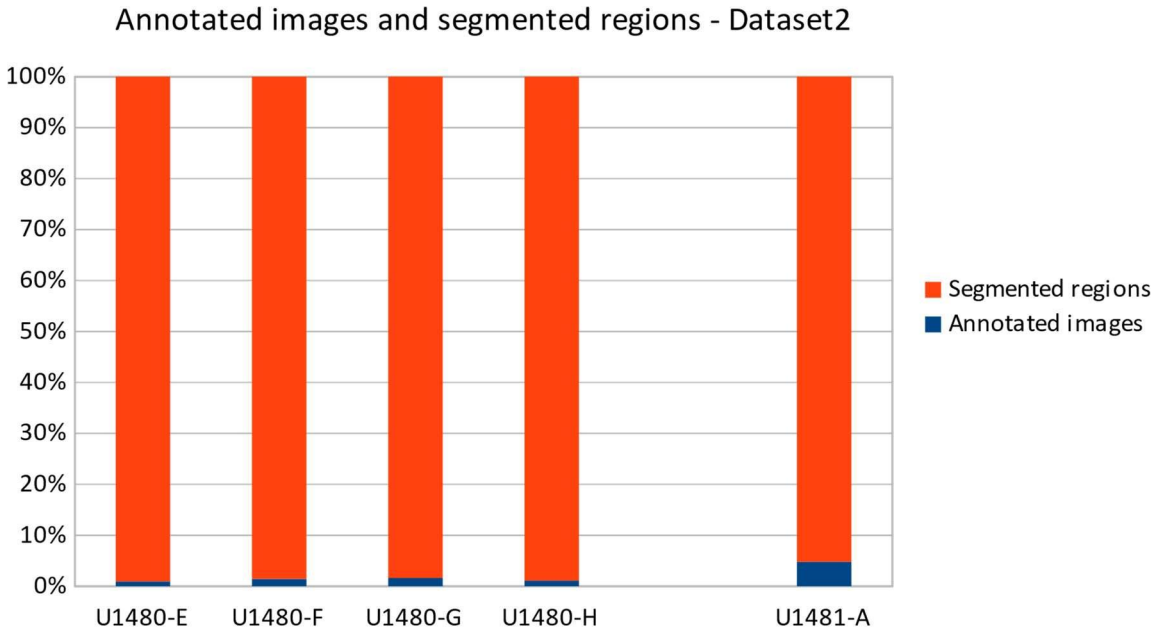


Figure 42 Summary of the composition of the original images and images annotated by site U1480 and site U1481 and holes E, F, G and H to U1480 and A to U1481. Original images are files acquired directly from the expedition, in high-resolution format. Annotation images are files created from the original images through annotation as described in chapter 2.3.

Figure 43 Division between original records and interpolated records for dataset1, U1480 and U1481. The geophysical data are: Gamma Ray Attenuation bulk density (GRA), Moisture and Density (MAD), Magnetic Susceptibility (MS), Natural Gamma Radiation (NGR), P-Wave Velocity Logger System (PWL), Red Green Blue channels (RGB) and Reflectance Spectrophotometry and Colorimetry (RSC).



Site/Hole	Annotated images	Segmented regions
U1480-E	79	8282
U1480-F	230	15869
U1480-G	481	29668
U1480-H	155	14175
U1481-A	647	17067

Figure 44 Summary of the formation of dataset2 between annotated images and segmented regions. Site/Hole covers sites U1480 and U1481 in the defined range.

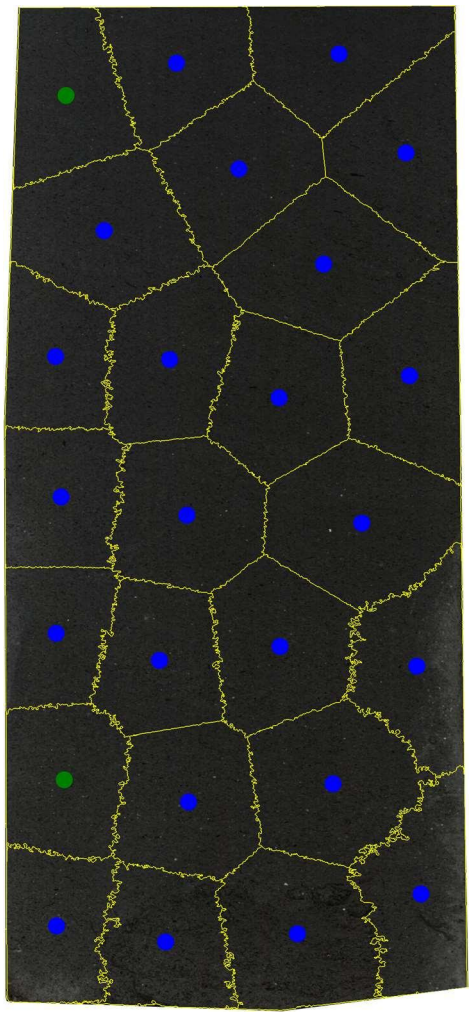


Figure 45 Classification of the segmented regions for the image: 362-u1480g-52r-3-a; region-1; Fine-sand_sandstone.png, lithology: Fine sand/sandstone, total segmented regions: 25, combination training and testing: 90%Train – 10% Test, Apx: 71,48 cm². Caption in blue: Training, Caption in green: Test with correct classification and Caption in red: Test with incorrect classification.