

APPLICATION OF MACHINE LEARNING ALGORITHMS IN HYDROCARBON
EXPLORATION AND RESERVOIR CHARACTERIZATION

by

Saba Keynejad

Copyright © Saba Keynejad 2018

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF GEOSCIENCES

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2018



ARIZONA

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that an accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Saba Keynejad

ACKNOWLEDGEMENTS

"It does not do to dwell on dreams and forget to live." -Albus Dumbledore.

I would like to express my deepest gratitude to my advisor Roy A. Johnson for taking a chance on me by admitting me into his research team. The greater chance, beyond the academic and professional opportunities that I have received, is and will always be having and knowing him as my mentor. It is utterly difficult, now, to imagine how I would have accomplished this if it was not for his truly unconditional support and helpful guidance throughout these years.

I am greatly thankful to Marc L. Sbar for his insights and constructive criticisms in the course of this research. I also thank him for all the discussions on political and social issues that helped me keep my sanity in these insane times. My special thanks to Mary M. Poulton, who I was so lucky and delighted to have her on my committee in spite of her various professional roles and responsibilities. I would like to thank Randall M. Richardson, whose resourceful teaching styles have been, and always will be, inspiring me.

I may not be able to thank enough Kiriaki Xiluri for her tireless help and support, which admittedly by everyone who has ever asked for her help, is beyond the scope of her responsibilities and out of the kindness of her heart.

I am very grateful to Lucy who was always there for our once-in-a-while caffeine breaks that seemed to be the only way of reminding us the color of the sky after a long day of staying in our gloomy window-less offices. I am also thankful to my current and former lab mates Daniel, Jared, and Noah, and all other fellow graduate students of reflection seismology group for their friendship and encouragements.

Butterfly effect or not, I believe something very special would have been missed from my life, had it not been for knowing my wonderful friends in Tucson: Goli, the wonderful woman who has always cared for me beyond my appreciation; Alyeh, Sara, Simin, who I had been extremely lucky for having them as my friends, and was always inspired by their wisdom, passion, and perseverance; Nafiseh and Mina who reminded me that the charm of friendship, just as of wine, gets stronger as it gets older. Special thanks to Nader and Faranack, friends and dear family of mine in the US, who encouraged me and supported me for my academic endeavors and the personal challenges I faced.

It is beyond my words to express how much I owe my "being" to my parents who urged us to be brave, strong, and independent women. Special mentions to my sisters, one of whom taught me the alphabet and basic math when I was 3, and the other who taught me English phonetic symbols to properly use her English-Farsi dictionary a bit later.

DEDICATION

To my Mämän,

and to the memory of my Bäbä.

TABLE OF CONTENTS

LIST OF FIGURES.....	10
ABSTRACT.....	11
INTRODUCTION	14
PRESENT STUDY	18
Fields and data	18
Methods.....	20
Simultaneous prestack inversion:	21
Fuzzy inference system.....	21
Bootstrap aggregating tree	22
Artificial neural networks	23
MAJOR RESULTS.....	24
REFERENCES	30

APPENDIX A: COMPARISON OF MODEL-BASED GENERALIZED REGRESSION NEURAL NETWORK AND PRESTACK INVERSION IN PREDICTING POISSON'S RATIO IN HEIDRUN FIELD, NORTH SEA..... 34

ABSTRACT.....	35
INTRODUCTION	37
SIMULTANEOUS PRESTACK INVERSION METHODOLOGY	40
ARTIFICIAL NEURAL NETWORK METHODOLOGY.....	42
RESULTS	44
CONCLUSION.....	46
ACKNOWLEDGMENTS	49

REFERENCES	50
LIST OF FIGURES.....	54
TABLES	57
FIGURES.....	58
APPENDIX B: ASSESSMENT OF MACHINE LEARNING TECHNIQUES IN PREDICTING LITHO-FLUID-FACIES LOGS IN HYDROCARBON WELLS ...	70
ABSTRACT.....	71
INTRODUCTION	72
HEIDRUN FIELD	75
KUPE FIELD	78
APPLICATION OF ML ALGORITHMS.....	79
ANN.....	81
SOM -.....	81
MLFN -	82
Bagged Trees	82
DISCUSSION AND RESULTS	83
Supervised training	84
Heidrun litho-fluid facies logs	86
Kupe litho-fluid facies logs.....	89
CONCLUSIONS.....	90
ACKNOWLEDGEMENTS.....	94
REFERENCES	95
LIST OF FIGURES.....	99

TABLES	102
FIGURES.....	103
APPENDIX C: CREATING PROBABILISTIC 3D MODELS OF LITHO-FLUID FACIES USING MACHINE LEARNING ALGORITHMS.....	114
ABSTRACT.....	115
INTRODUCTION	116
METHODOLOGY	119
FIS.....	121
BT	123
PNN.....	125
DISCUSSION AND RESULTS	126
CONCLUSION.....	131
ACKNOWLEDGEMENTS.....	135
REFERENCES	136
LIST OF FIGURES.....	138
TABLES	141
FIGURES.....	142
APPENDIX D: TERMS AND DEFINITIONS OF THE MACHINE LEARNING TECHNIQUES DISCUSSED IN THIS DISSEERRTATION	153
ARTIFICIAL NEURAL NETWORKS	154
Multilayer feed-forward neural networks	155
Generalized regression and probabilistic neural networks	156
Self-organizing feature maps	158

BOOTSTRAP AGGREGATING TREES	159
GLOSSARY	160
REFERENCES	163

LIST OF FIGURES

Figure 1. Poisson's ratio predicted by GRNN (a) and prestack inversion (b) shown on an inline at well 6507/8-1, Heidrun Field. GRNN result is more consistent with the HC shows in log data from GOC to OWC depth. ANN model is also able to predict an anomalously low-PR zone at the depth of the Melke Formation.	26
Figure 2. BT prediction of litho-fluid facies for well 6507/8-1 of Heidrun Field (left), and well Kupe South-7 of Kupe Field (right). P-impedance, gamma ray, and deep resistivity logs are shown for reference.	27
Figure 3. 3D models of LFF-class probabilities for PNN (left), and BT (right). Classes are: blue: brine sand, green: oil sand, brown: shale, and red: gas sand, with the lighter colors for lower probabilities. The opacity of colors is modified as shown in color-scale insets to better depict the variations, especially for HC classes.....	29

ABSTRACT

This dissertation presents novel approaches to evaluate complex seismic and well-log data using machine learning algorithms with examples from two different hydrocarbon fields. The applicability of these algorithms for predicting and classifying direct or indirect hydrocarbon indicators are assessed and compared to knowledge-driven methods. The efficacy of the various techniques leads to recommendations for utilizing machine learning algorithms in well planning or later cycle hydrocarbon-field development.

In the first study in this dissertation, application of a model-based artificial neural network is compared to the performance of a prestack simultaneous inversion method in predicting hydrocarbon presence in the Heidrun Field, offshore Norway. Low-frequency initial models were used to create 3D Poisson's ratio models to reflect the fluid within this field and the results were compared based on the accuracy and generalization power of the two methods. The results of both methods confirmed Poisson's ratio to be a good direct hydrocarbon indicator within the wells used from this field. The direct dependency of the inversion method on the provided input constraints, however, can raise the risk for well planning decisions beyond the known zones. The generalize regression neural network results better matched the observations at the training wells and provided a lower risk of false discoveries in delineating favorable zones beyond the drilled wells.

The second study was conducted with the aim of classifying different facies from well logs in wells of the Heidrun Field and in the Kupe Field, offshore New Zealand. Different machine learning approaches were utilized in this study and to investigate quantitatively and qualitatively the accuracy and stability of their predictions. Both supervised methods could successfully predict hydrocarbon-bearing units, with the bagged tree algorithm

having a higher overall, and hydrocarbon-related, accuracy rate. Application of the bagged-tree algorithm showed a very low false discovery rate for oil sands and no false discoveries for gas sands in the Heidrun Field. A misclassification of oil sands as brine sands in one Heidrun well is in agreement with relatively high Poisson's ratios as discussed in the first study. Qualitative investigations of Kupe Field results also demonstrated accurate prediction of hydrocarbon-bearing units, including a shaly hydrocarbon sand class defined for low-quality reservoir sands. Hydrocarbon shows reported in one well that were not predicted by the algorithm, in fact, occur in a very low-porosity section of the reservoir that was not identified as reservoir in reports either.

In the last study, the classifications of the litho-fluid facies were extended to three dimensional models using two machine learning methods and were compared with a knowledge-driven approach. The results were examined through a probabilistic approach to reflect the uncertainty of the predicted classes. The probabilistic neural network and the bagged-tree algorithm successfully predicted the variations of litho-fluid facies, especially for hydrocarbon units. Both methods predicted gas sands in certain parts of the field, away from control points, with similar form and lateral dimension. By comparing the results in predicting oil sands and shale, we interpret the bagged-tree method to be more adherent to the known parameters set by the interpreter, such as the OWC and the target classes. Predictions from the probabilistic neural network, however, can deviate from the target facies even close to the wells on which it has been trained.

The efficiency of machine learning techniques in increasing the prediction accuracy and decreasing the procedure time, and their objective approach toward the data, make it highly desirable to incorporate them in seismic data analyses. Along with the emphasis on

the application of machine learning techniques in the study of subsurface properties, this dissertation presents frameworks for utilizing these techniques as new tools for the interpreter, not as a replacement. The knowledge of the data analyst about the field, and the selection and preparation of the attributes and application of the appropriate algorithm are all crucial factors in this procedure.

INTRODUCTION

Advancements in hardware technology and computational power used in seismic data acquisition and processing have resulted in considerable increases in the resolution and volume of the data. While this dramatic rise in data volume and complexity, caused by increases in dimension (3-D, 4-D), resolution, and number of attributes generally is an advantage by itself, there are inevitable challenges stemming from very large, complex data sets that need to be addressed. In seismic data analysis and interpretation, ever-growing computational power can handle part of the data-growth issue when it comes to visualization and calculation tasks. But before getting to this phase, the analyst needs to carefully investigate the data to design a model for the problem at hand and choose the parameters necessary for the calculations.

Much larger and more complex data volumes require deeper analysis to take advantage of advancements and avoid mistakes or oversights in interpretations. It does not seem logical to expect analysts to study the data quantitatively and/or qualitatively, line-by-line and sample-by-sample, to set a framework that grasps as much of the data as is appropriate for the final calculations. Moreover, analyzing the complicated and intertwined relationships among the parameters, and the very often non-linear relations that relate them to the geological targets in a multi-dimensional space, needs something more than an ever-increasing number of analysts in the work force.

Mathematical models and physical assumptions have traditionally related the seismic response to geologically meaningful models and, with the aid of improvements in computational power, these models are more accurate than ever. Incorporating the expanding classes of information into these methods, however, can strain the capabilities

of some of the analytical techniques or can be theoretically problematic. This is especially true of the information content extracted from the seismic response (referred to as seismic attributes, or just attributes). Apart from the arrival time and dip of seismic events, which were used in structural interpretations since the beginning of explorational seismology, introduction of more sophisticated seismic attributes began around the early 1970's (e.g., Balch (1971) and Nigel Anstey (Taner, 2001)). Their work showing color overlays of interval velocity, reflection strength, and mean frequency, was later followed by Hilbert transformation and introduction of complex trace attributes (Taner et al., 1979), which opened the door to hundreds of additional attributes up to now (Tanner, 1992).

However, the practicality, or even the meaning of some of these attributes, is still subject to debate. Some attributes are considered to be duplicates of other attributes, hence not adding any new information to the model, while some are suggested to be discarded as they are “nonsense” or lacking a geophysical or geological meaning (Barnes, 2006). Moreover, many computational models suffer from overfitting, and too many model parameters can be a source of this problem. While the issues about redundancy or theoretical basis of the attributes are mostly correct regardless of the computational platform and must be considered during data analysis, one important point should be noted: even if we cannot “see” how a set of attributes that have theoretical meaningful bases is related to the expected geological solution, it does not necessarily mean that such a relation does not exist. Despite our inability to perceive a higher-dimensional space (greater than 3-D), and our limitation in handling complicated and inter-related parameters in such an environment, evolution has provided us with the right amount of intelligence and exceptional tool-making talents to find a solution to very complex problems!

It did not take long after Turing's (1937) "learning machine" and McCulloch and Pitts' (1943) "threshold logic" as a computational model representing biological nervous systems, that the concept of artificial intelligence developed. Machine learning (ML) is one of the branches of artificial intelligence that has influenced various fields of study and is finding its way to new ones. ML includes any algorithm that can take these three steps for an expected task, such as predicting an output value or identifying different classes:

1. Take in the input data,
2. Analyze the data to learn the governing rules (aka, training),
3. "Generalize" those rules and relations to a new, unseen dataset.

It is the second step that distinguishes ML from other computational methods; the algorithm is not explicitly programmed on the data. In other words, there are no pre-determined equations or rules, set by the human expert, to relate the input data to the outcome. Rather, the algorithm is expected to "learn from experience" during the training phase. Some algorithms are trained to learn the task with respect to a set of accompanied outcomes or targets (supervised ML), while the others are designed to learn the patterns, such as similarity, within the input features only (unsupervised ML). Either way, the training is accomplished when a measure of performance is optimized, for instance, an error value or a distance/similarity parameter.

ML algorithms are not used in geophysical studies only to catch up with new trends in technology; the massive amounts of data and different data types, as explained before, need more powerful analytical capabilities and approaches. And even more importantly, the earth, as a natural phenomenon, imposes more uncertainty and complexity to the problem than what conventional hard-computing mathematical tools usually can handle. ML

algorithms are considered as universal approximators that, if provided with proper data¹, can learn and generalize the patterns and trends of non-linear systems, as is the case in geophysical studies. For these reasons, ML has become a necessary tool in predicting characteristics of a petroleum reservoir based on seismic attributes and petrophysical parameters.

Cases of successful application of ML algorithms, such as artificial neural networks in seismic data analysis, date back to the 1980's (Poulton, 2002; Russell, 2005). As more algorithms are designed, and seismic data quality becomes more efficient, more studies are conducted on the application of ML in geophysics. From predicting reservoir-controlling parameters such as porosity and pressure to mapping structural boundaries, gas-chimneys, and direct hydrocarbon indicators such as lithology and fluid classes, ML algorithms play an important role in understanding subsurface features in one- to four-dimensional seismic data analysis. For a more recent list of publications see the list of references of the following appendices and the scientific journals with special issues designated to this subject (Jayaram et al., 2015; Davidson, 2017; AlRegib et al., 2018).

ML algorithms also have the advantageous ability to make decisions more objectively in comparison with knowledge-driven methods. The quality of the dataset and the selected attributes are important for both knowledge-driven and data-driven methods and can affect the objectivity of both approaches. But in the case of ML, the extent of the analyst's knowledge and their subjective point of view about the reservoir does not play as

¹ In other words, proper selection of number and type of attributes (and target, if applicable) that span the attribute space as much as possible. There are suggestions and tips for some algorithms, but it is eventually up to the expert to select and prepare the data.

significant a role as the data itself. For this reason, the data and framework that MLs are trained on are assessed in this dissertation, to investigate the accuracy in their predictions and to compare their performance to knowledge-driven and/or hard-computing methodologies.

The higher accuracy and shorter calculation times in ML approaches, in comparison with other methodologies, make them powerful tools in reservoir characterization studies. Moreover, their objective approach makes the results more reliable when predictions are made away from the control points and training data. The analysis frameworks proposed in this dissertation emphasize the advantages of implementing robust ML-based methodologies in traditional seismic-data-analysis procedures along with the interpreter's first-hand knowledge and information.

PRESENT STUDY

Different ML approaches are used in this dissertation to predict fluid and lithology presence within test hydrocarbon reservoirs. The results are compared to conventional quantitative methods or other ML approaches. To assess the results, different parameters are considered: 1) quantitative and/or qualitative accuracy of results (i.e. to what degree are the predictions in agreement with the training data), 2) calculation time, and 3) generalization ability of the approaches; depending on data availability, this parameter can be evaluated either by test data, or in an interpretive way with respect to other information about the field such as structural interpretations or drilling reports.

Fields and data

The seismic and well data used in this dissertation are from Heidrun and Kupe Fields, offshore Norway and New Zealand, respectively. The seismic amplitude data sets are in

the form of full-stack 3D volumes, and time-migrated, NMO-corrected CMP gathers, with a 4-ms sample rate from Heidrun Field. Located on the mid-Norwegian continental shelf, Heidrun Field was formed in an extensional tectonic phase during the separation of Eurasia and North America in Late Jurassic to Early Cretaceous. The Middle Jurassic Fangst Group clastic reservoir was deposited in a shallow-marine to fluvial environment. It comprises three mostly clean sandstone formations, with the Not Formation being a more shaly layer that thickens toward the southwest. For more on the stratigraphy of the Fangst reservoir and Heidrun Field geology, see Harris (1989) and Morton (2009). In this dissertation, we have focused on a section that includes the Melke Formation shale (Viking Group, Middle to Upper Jurassic), the Fangst Group, and the underlying Ror muddy sandstone and Tilje sandstone and shale (Båt Group, Early Jurassic).

A subset of seven wells was used in each appendix, of which three are deviated production wells, and the rest are appraisal wells. The drilling and logging information about formation tops and hydrocarbon shows were provided by the Norwegian Petroleum Directorate website (n.d.) for the four appraisal wells (6507/7-3, 6507/7-4, 6507/7-8, 6507/8-1). The producing wells were left as a “blind wells” for assessment and verification purposes (6507/7-A-17, 6507/7-A-35, 6507/7-A-53). Six reflection horizons were picked based on structural and stratigraphic interpretations including: sea bed, two shallow control horizons, Brygge Formation, Fangst Group, and Åre Formation. An rms velocity model was built using available well-tied time-depth logs, and the interpreted structural horizons. The analyses were done on a full-stack seismic volume, from 1000 to 3000 ms, with the focus on the Fangst Group to Åre Formation as the pay zone.

Another hydrocarbon field, used in Appendix B, is Kupe Field in the southeastern Taranaki Basin, offshore New Zealand. The Paleocene Farewell Formation is the primary reservoir for this gas-condensate field, comprised of medium- to coarse-grained sandstones with interbedded shale layers, deposited in fluvial to coastal braided plains (Pang and Collen, 1996). Our primary focus in this study is on sections in eight wells that contain the Farewell sandstone reservoir. The reservoir quality of the Farewell Formation can be highly affected in some parts depending on the type of clay minerals within the sandstone (Martin et al., 1994). We used four wells (Kupe South-6, Kupe South-7ST1, Kupe South-8, and Momoho-1) to train the ML algorithms, using information from the available drilling reports (New Zealand Petroleum & Minerals, 2017).

Methods

Depending on the objectives of the study, different approaches are used in each appendix to investigate a hydrocarbon field from its wells and/or within a seismic volume. The goal in each study is either to predict a value (regression problem) or classify the facies (classification problem) to delineate the potential hydrocarbon-bearing units. These methods can be categorized by two main paradigms: knowledge-driven and ML (data-driven). All these methods are common in the main procedure in which they take in attributes (models, well logs, seismic attributes) and predict the final output based on the rules and relations between them. The key difference is in the way the method has “learned” those rules and relations. In knowledge-driven methods, those rules and relations are observed and dictated by the analyst. In ML, on the other hand, the algorithm extracts the relations and rules during the training phase from the presented attributes (and the target

for supervised cases). More details about machine learning methods used in this dissertation are provided in Appendix D.

Simultaneous prestack inversion:

Simultaneous prestack inversion is a technique for predicting lithology and fluid properties using angle (or offset) gathers (CDP gathers that are arranged by offset or computed angle of reflection incidence). It derives angle-dependent information from CDP offset stacks and estimates P-impedance (Z_P), S-impedance (Z_S), and density (ρ) through an iterative algorithm (Hampson and Russell, 2005). Inverting for the two latter parameters reduces the non-uniqueness risk of the solution, which is inherent within poststack inversion algorithms (Ma, 2002). The method we use, draws on the relationship between $Z_P - Z_S$, and $Z_P - \rho$ couples based on Fatti's re-expression of Aki-Richard's equations, and the generalized Gardner's equation respectively (Hampson and Russell, 2005). These relationships are assumed to be linear in the absence of complicating factors such as hydrocarbon presence. To reduce the instability issue inherent within the inversion problem, an iterative algorithm is chosen, which begins with an initial guess for Z_P and uses the conjugate-gradient method to approach the best possible solution step by step.

Fuzzy inference system

As a knowledge-driven approach, we designed a fuzzy inference system (FIS) of Mamdani type (Mamdani and Assilian, 1975). FISs are decision-making systems that are developed based on Zadeh's fuzzy set theory (Zadeh, 1965). Whether or not the elements belong to a fuzzy set is determined by a range of values called membership degree, contrary to the sharp boundaries of inclusion and exclusion in classical set theory. FIS-based

classification rules are a collection of linguistic statements in an if-then format that sets the path for the FIS to make a decision. For instance:

if (input1 with membership degree1) aggregated (by operator1) with (input2 with membership degree2) then: (output1 with membership degree3) weighted by the confidence/importance level x.

Fuzzy logic is used in ambiguous conditions, by handling the uncertainty based on the analyst's knowledge fed into the system in the form of the membership degrees and the logical framework. We have used this ambiguous state in the FIS results as a tool to reflect the uncertainty of predicting litho-fluid facies (LFF) volumes by a knowledge-driven method. The results of the FIS are a direct reflection of the expert's knowledge, and thus provide a reasonable basis to evaluate ML results. For FIS, to be able to define the aggregation rules, we are limited to those attributes whose relationships to the target and each other can be "observed".

Bootstrap aggregating tree

Decision trees are ML algorithms that relate attributes via "branches" to a "leaf", which is a predicted value (regression) or label (classification). The branches split at several points based on variations in the features (predictors or log values). The bagged-tree (BT) procedure is a method to ensemble the trees to increase the stability and reduce the variance, which are common problems as the trees tend to "grow too deep" and learn the noise within the data. It begins by creating replicates of the learning set through bootstrap resampling, i.e., random sampling with replacement from the training data set (Breiman, 1996). In this way each decision tree is trained on a bootstrap sample set, the outcome of which is an aggregation of the results through "voting" or averaging for

classification and regression problems, respectively. The BTs in this dissertation use a random-forest approach to grow the bagged trees. This means that, in addition to the bootstrap sampling of the training data set for each tree, the features at each split are also randomly sampled without replacement (Breiman, 2001).

Artificial neural networks

ANNs are a particular branch of ML, originally inspired by biological neural networks. ANNs are well known as powerful ML techniques in solving complex and non-linear problems governing a large amount of multi-attribute data sets. Supervised ANNs take in the observations (e.g., logs) and relate them to the associated targets (e.g., known lithofacies) via non-linear activation functions of their innermost hidden layer(s). Unsupervised algorithms on the other hand, have the advantage of not depending on the availability of known targets for making predictions. But for the same reason, their results need to be evaluated more rigorously and validated by some reliable evidence before being authorized in a decision-making procedure.

The number of layers, their neurons (processing elements), the connection weights between layers, and the direction of information flow in an ANN defines its architecture, which in turn depends on the problem. The training algorithms are optimization problems that minimize the difference between prediction and the target by tuning the connection-weight values accordingly.

Four types of ANNs are used in this dissertation: (1) A self-organizing feature map (SOM) neural network, (2) A general regression neural network (GRNN), (4) a probabilistic neural network (PNN), and (4) A multi-layer feed-forward network (MLFN).

SOM is an unsupervised algorithm that clusters data based on the “similarities” of the input data while preserving the topological relationships based on the Kohonen rule (Kohonen, 1987). Both GRNN and PNN are supervised ANNs that use radial basis functions in their hidden layer to assess the closeness of the input data in training (Specht, 1990; Specht, 1991). However, GRNN is used in regression problems and PNN is used for classification and pattern recognition tasks. MLFN is one of the common supervised ANNs that can be used in classification and function-approximation problems, depending on the activation function used in its output layer.

MAJOR RESULTS

The studies presented in this dissertation are aimed at finding direct or indirect hydrocarbon indicators within wells or across the reservoir section of hydrocarbon fields. Proposed methods and frameworks can be incorporated in reservoir-characteristic assessments and well-planning studies to provide more accurate subsurface information in less time. These approaches not only improve the exploration and production operations, but also can be environmentally significant as it reduces the need/risk of un-necessary drilling.

In Appendix A, a prestack inversion and a GRNN approach are used to create a 3D model of Poisson’s ratio for Heidrun Field. Low-frequency models of different seismic parameters such as P-impedance, S-impedance, and Poisson’s ratio are calculated using well logs and structural constraints to serve as initial models for each method. The results show that inversion can predict Poisson’s ratio as a hydrocarbon indicator in very fine detail around four wells that were used to build the model for the inversion. However, GRNN demonstrated better performance beyond the known areas (i.e., for a blind well), as

well as for hydrocarbon shows shallower than logging depth at a training well. The results indicate that GRNN, unlike inversion, did not adhere to the predefined boundaries and limits away from training points (Figure 1). GRNN results also demonstrated a lower risk of showing false-positive responses. This makes GRNN a favorable option for development goals in reservoir evaluation beyond the known zones.

In the second study, presented in Appendix B, the presence of hydrocarbon is investigated by classifying lithofacies and fluid variations within the wells in the form of litho-fluid facies logs. An unsupervised ANN (SOM), a supervised ANN (MLFN) and a BT classifier were trained on the litho-fluid facies logs that were created for wells from two hydrocarbon fields (Heidrun and Kupe Fields). Two supervised classifiers (MLFN and BT) outperformed the unsupervised ML algorithm in classifying litho-fluid facies logs in the training wells. Quantitative and qualitative investigations were performed on training results of the two supervised methods, and results showed that BT had higher accuracy and stability through multiple runs of each algorithm.

A more detailed study of the prediction results for trained BTs showed a very low false discovery rate for oil sands, and no false discovery for gas sands in Heidrun Field. Qualitative investigations of the Kupe Field results also demonstrate no false discoveries for hydrocarbon-bearing units. The results emphasize the importance of ML techniques in well-log interpretation, especially in the absence of crucial logs and/or low core recovery, as it can predict reservoir sections within wells with a very low risk of false discoveries (Figure 2).

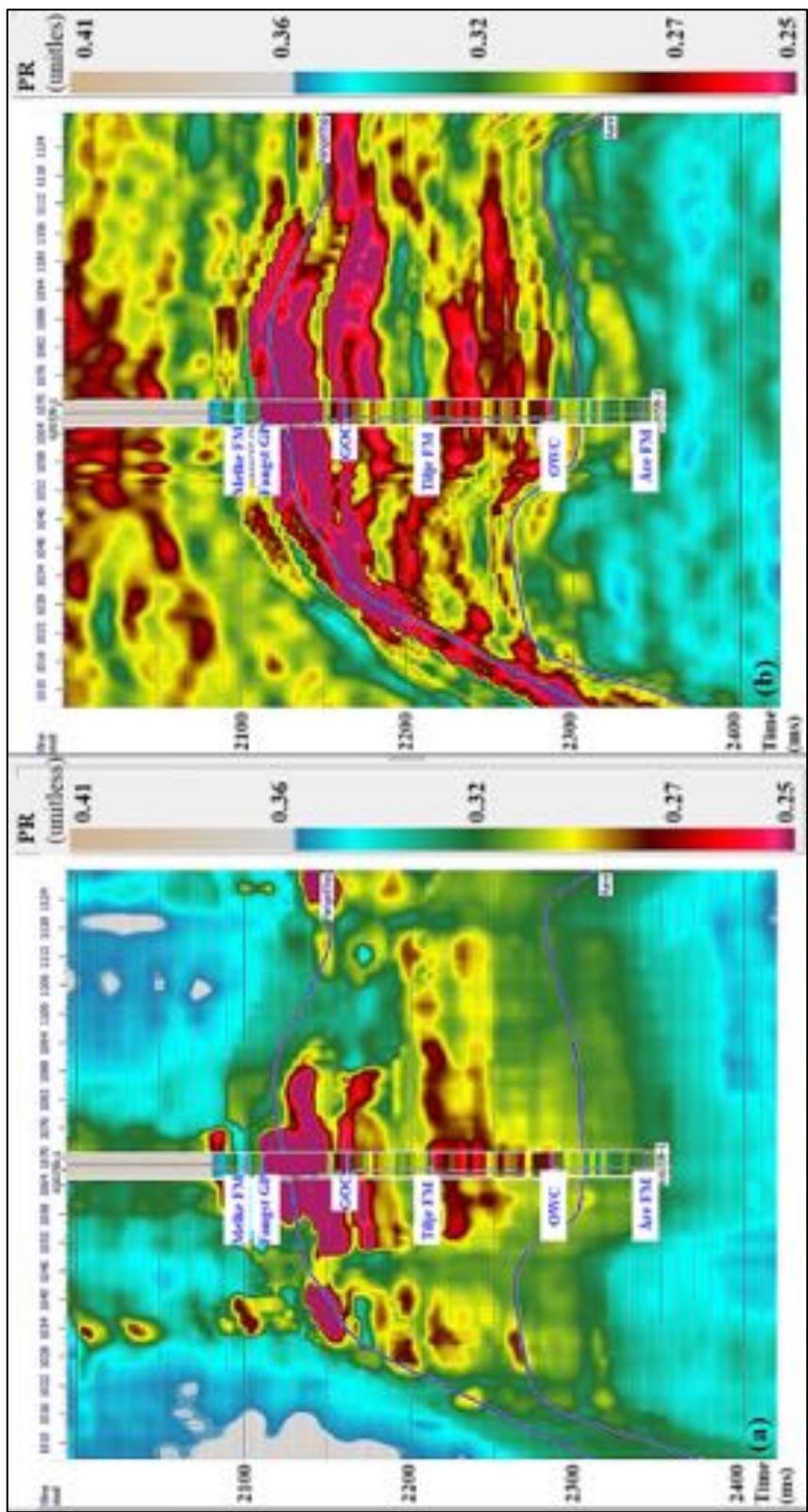


Figure 1. Poisson's ratio predicted by GRNN (a) and prestack inversion (b) shown on an inline at well 6507/8-1, Heidrun Field. GRNN result is more consistent with the HC shows in log data from GOC to OWC depth. ANN model is also able to predict an anomalously low-PR zone at the depth of the Melke Formation.

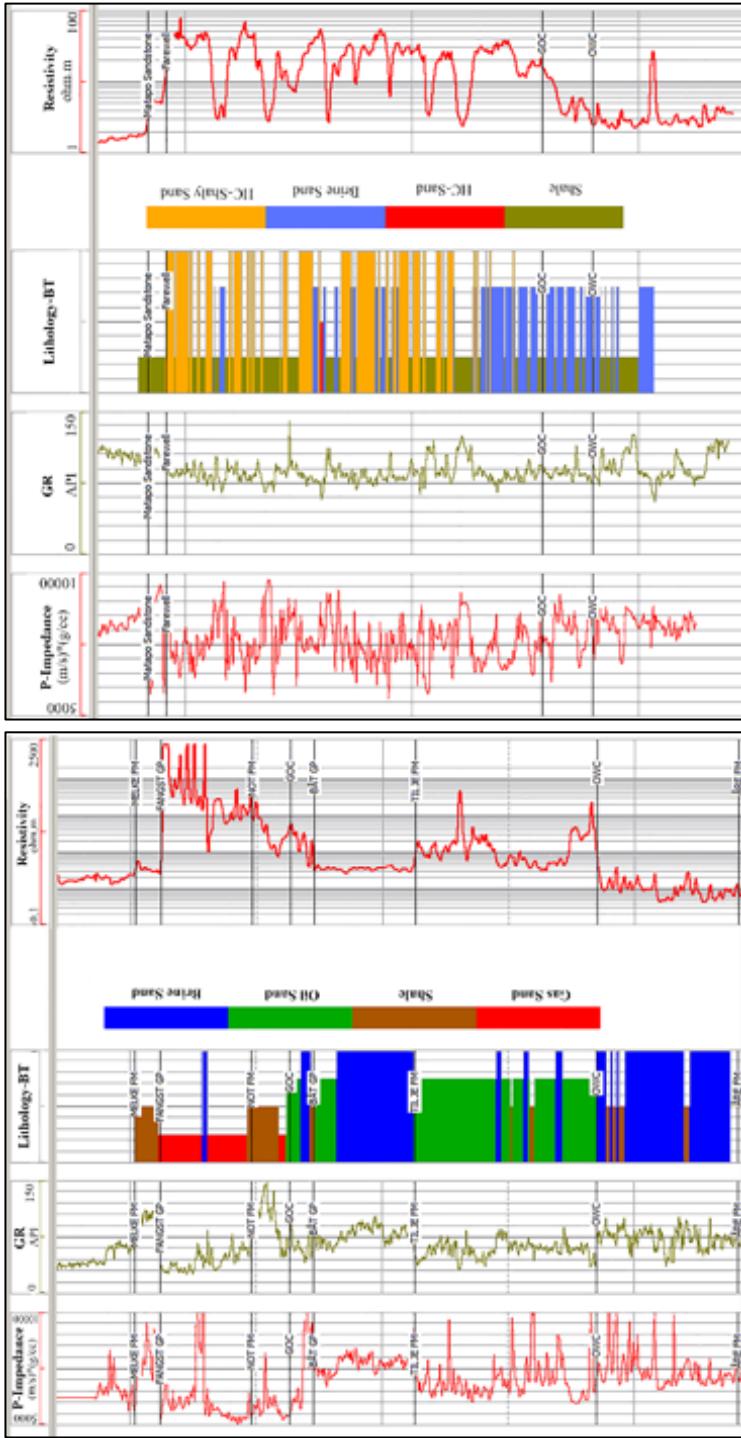


Figure 2. BT prediction of litho-fluid facies for well 6507/8-1 of Heidrun Field (left), and well Kupe South-7 of Kupe Field (right). P-impedance, gamma ray, and deep resistivity logs are shown for reference.

The third study, Appendix C, is built upon the previous one, by creating a 3D litho-fluid facies (LFF) model for Heidrun Field. The more important objective of this study is to provide a way to assess the uncertainty of the predicted classes. A fuzzy inference system (FIS), a BT classifier, and a PNN are used and their performances are investigated using blind wells and qualitative interpretations. FIS is used since it can closely resemble the interpreter's behavior in making decision in classifying LFFs throughout the reservoir, using the attributes with “observable” relationships among themselves and the target LFF. Also, the uncertainty and ambiguousness of classification decisions can be depicted by the “fuzzy” representation of its outcomes. The uncertainty in the results of BT and PNN, however, are derived from the probabilities calculated by these two approaches as part of their classification algorithms. The MLs successfully classified the LFFs within the fluid contact intervals, with BT being in better agreement with the lateral distribution of hydrocarbon-bearing units (Figure 3). Both ML methods drastically outperformed FIS in predicting LFF classes and in the calculation times.

Predicting facies classes and the uncertainty of prediction is crucial in reservoir characterization since various sources of heterogeneity, from the gradual changes of facies to assumptions and simplifications in relating elastic to petrophysical parameters, are usually the biggest problem in mapping facies. Mapping the uncertainties and evaluating LFFs within a confidence interval can be utilized in estimation of reservoir capacity and to reduce the risk of false discoveries in well planning studies.

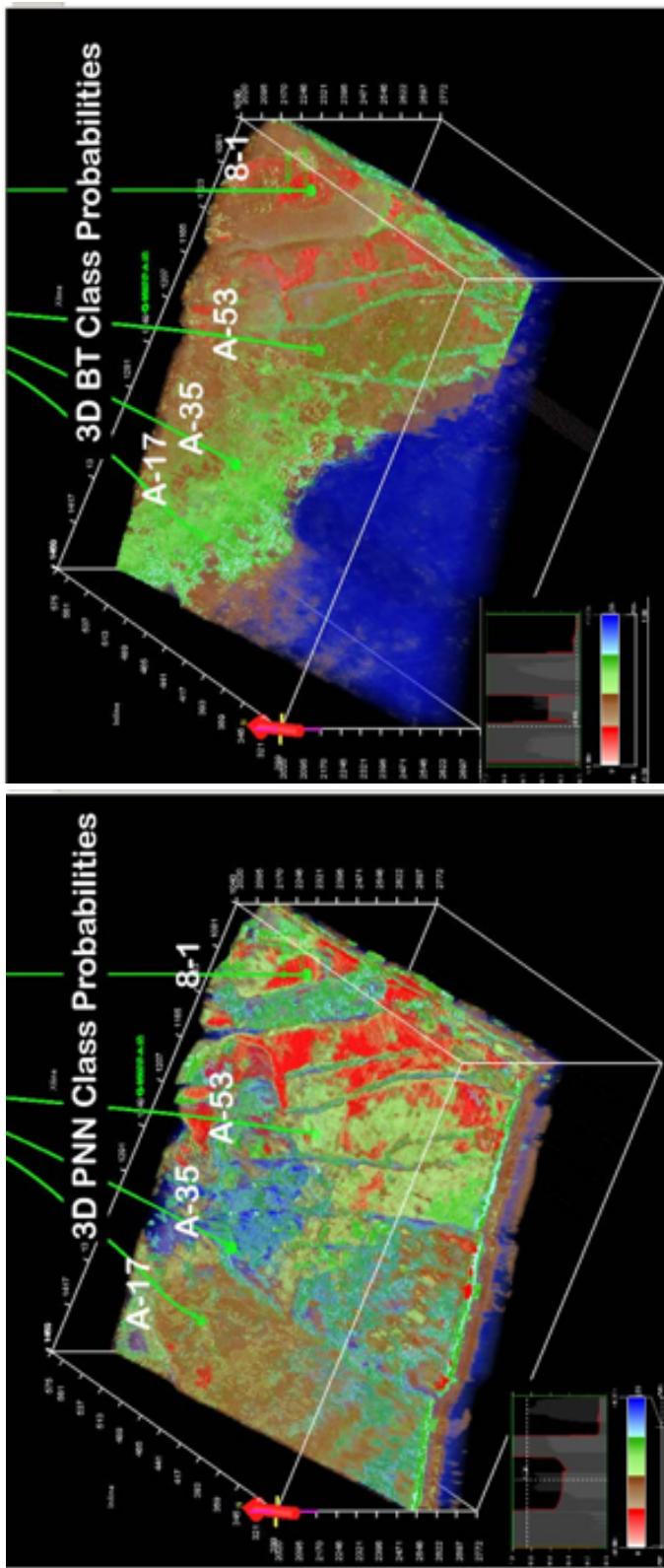


Figure 3. 3D models of LFF-class probabilities for PNN (left), and BT (right). Classes are: blue: brine sand, green: oil sand, brown: shale, and red: gas sand, with the lighter colors for lower probabilities. The opacity of colors is modified as shown in color-scale insets to better depict the variations, especially for HC classes.

REFERENCES

- AlRegib, G., Fomel, S., and Lopes, R. 2018, Subsurface Exploration: Recent Advances in Geo-Signal Processing, Interpretation, and Learning [Special Issue], IEEE Signal Processing Magazine, **35**.
- Balch, A. H. 1971, Color Sonograms: A New Dimension in Seismic Data, Interpretation, Geophysics, **36**. p. 1074-1098.
- Barnes, A. 2006, Too Many Seismic Attributes? CSEG RECORDER Magazine, **31**. p. 41-45.
- Breiman, L. 1996, Bagging Predictors, Machine Learning, Springer, **24**. p. 123-140.
- Breiman, L. 2001, Random Forests, Machine Learning, Kluwer Academic Publishers, **45**. Boston, p. 5-32.
- Davidson, M. 2017, Data Analytics and Machine Learning [Special Section], Society of Exploration Geophysicists, p. 206-266.
- Hampson, D. P., and Russell, B. H. 2005, Simultaneous Inversion of Pre-Stack Seismic Data, SEG Technical Program Expanded Abstracts, **24**. p. 1633.
- Harris, N. B. 1989, Reservoir Geology of Fangst Group (Middle Jurassic), Heidrun Field, Offshore Mid-Norway, AAPG Bulletin, **73**. p. 1415-1435.

Jayaram, V., Avseth, P., Azbel, K., Coléou, T., Devegowda, D., de Groot, P., . . . Kumar, V. 2015, Introduction to Special Section: Pattern Recognition and Machine Learning [Special Section], Interpretation, Society of Exploration Geophysicists, **3**. p. SAE93.

Kohonen, T. 1987, Self-Organization and Associative Memory, Springer-Verlag.

Ma, X. 2002, Simultaneous Inversion of Prestack Seismic Data for Rock Properties using Simulated Annealing, *Geophysics*, **67**. p. 1877.

Mamdani, E. H., and Assilian, S. 1975, An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller, *International Journal of Man-Machine Studies*, **7**. p. 1-13.

Martin, K. R., Baker, J. C., Hamilton, P. J., and Thrasher, G. P. 1994, Diagenesis and Reservoir Quality of Paleocene Sandstones in the Kupe South Field, Taranaki Basin, New Zealand, *AAPG Bull.*, **78**. p. 624-643.

McCulloch, W. S., and Pitts, W. 1943, A Logical Calculus of the Ideas Immanent in Nervous Activity, *Bulletin of Mathematical Biophysics*, **5**. p. 115-133.

Morton, A., Hallsworth, C., Strogen, D., Whitham, A., and Fanning, M. 2009, Evolution of Provenance in the NE Atlantic Rift: The Early–Middle Jurassic Succession in the Heidrun Field, Halten Terrace, Offshore Mid-Norway, *Mar.Pet.Geol.*, **26**. p. 1100-1117.

New Zealand Petroleum & Minerals. 2017, New Zealand Petroleum Exploration Data Pack, New Zealand Petroleum & Minerals,

Norwegian Petroleum Directorate. Heidrun Field. Retrieved from
<http://factpages.npd.no/factpages>

Pang, J., and Collen, J. 1996, Well Log Analysis of the Kupe South Field, Taranaki Basin, New Zealand, SPE Asia Pacific Oil and Gas Conference, Society of Petroleum Engineers, Adelaide, Australia,

Poulton, M. M. 2002, Neural Networks as an Intelligence Amplification Tool: A Review of Applications, *Geophysics*, **67**. p. 979-993.

Russell, B. H. 2005, Neural Networks Applications in Geophysics, p. 339-341.

Specht, D. F. 1991, A General Regression Neural Network, *TNN*, IEEE, **2**. United States, p. 568-576.

Specht, D. F. 1990, Probabilistic Neural Networks, *Neural Networks*, **3**. p. 109-118.

Taner, T. 2001, Seismic Attributes, *CSEG RECORDER Magazine*, **26**. p. 48-56.

Taner, T., Koehler, F., and Sheriff, R. 1979, Complex Seismic Trace Analysis, *Geophysics*, Society of Exploration Geophysicists, **44**. p. 1041-1063.

Tanner, M. T. (1992). Attributes revisited. Retrieved from
<http://www.rocksolidimages.com/attributes-revisited/>

Turing, A. M. 1937, On Computable Numbers, with an Application to the Entscheidungs problem, *Proceedings of the London Mathematical Society*, Oxford University Press (OUP), **s2-42**. p. 230-265.

Zadeh, L. A. 1965, Fuzzy Sets, Information and Control, **8**, p. 338-353.

**APPENDIX A: COMPARISON OF MODEL-BASED GENERALIZED REGRESSION
NEURAL NETWORK AND PRESTACK INVERSION IN PREDICTING POISSON'S
RATIO IN HEIDRUN FIELD, NORTH SEA**

Saba Keynejad¹, Marc L. Sbar¹, Roy A. Johnson¹

¹*University of Arizona, Tucson, AZ, USA*

The text and figures in this appendix were previously published as:

Keynejad, S., Sbar, M., & Johnson, R. (2017). Comparison of model-based generalized regression neural network and prestack inversion in predicting Poisson's ratio in Heidrun Field, North Sea. *The Leading Edge*, 36(11), 938-946.

10.1190/tle36110938.1 Retrieved from <https://doi.org/10.1190/tle36110938.1>

ABSTRACT

Drilling wells in the oil and gas industry is a substantial process, whether it be an appraisal well drilled for reservoir-characteristic assessments at the exploration stage, or a production well drilled following prior assessments. The challenge has always been to reduce drilling-related expenses and natural/environmental hazard by reducing the number of wells drilled, and to evaluate reservoir characteristics with as few calibration wells as possible. Physical and mathematical modeling of seismic data can help us understand the geologic and structural formations with minimal wells, and interpolate reservoir characteristics across large areas between a few drilled wells. In a new comparative approach, simultaneous prestack inversion and Artificial Neural Network (ANN) methods are used to create 3D Poisson's ratio models built upon low-frequency initial models. Training the ANN on initial models similar to those used in the inversion has improved its performance, while creating a valid base of comparison between the two methods. The inversion method was able to model the Poisson's ratio around four wells that had been used in creating the initial models. The Generalized Regression Neural Network (GRNN) that was trained on a Poisson's-ratio initial model, along with other seismic attributes, gave results that were consistent with the existing wells. The results of both methods confirm the existence of a strong relationship between Poisson's ratio and known hydrocarbon presence in these wells. However, examining the results with a blind well showed that ANN was notably more successful than inversion in extrapolating the results beyond the logged sections in the wells, and away from control wells. While this particular conclusion cannot be generalized, and the results obtained from the same methodology may vary from one reservoir to another, such results suggest

that this procedure can become a robust part of a pre-drilling reservoir-evaluation phase in developing hydrocarbon fields.

INTRODUCTION

Seismic imaging is the most powerful tool in hydrocarbon exploration due to its ability to provide detailed subsurface information indirectly. Seismic data analyses can take this a step farther by relating the seismic response to potential physical predictive agents through mathematical models. Those methods, aimed specifically at reservoir characterization, play an essential role not only in defining the limits of a potential oil or gas field, but also in reducing the risk of drilling dry wells.

Inversion methods are among the most widely used tools in this field that are applied to obtain acoustic impedance and related physical properties of reservoirs (Huuse and Feary, 2005; Leite and Vidal, 2011a). Prestack inversion has become an effective method that reduces the inherent risk of non-uniqueness, and provides more reservoir-related information in comparison with other inversion methods (Goodway, 1999; Ma, 2002). It benefits from angle- (or offset-) related variations in seismic response and extracts shear-wave properties such as shear-wave velocity (V_S) and S-Impedance (Z_S), along with conventional acoustic-impedance information.

Artificial Neural Networks (hereinafter referred to as ANN, neural network, or network) are an alternative technology in dealing with seismic problems. ANN is a programmed computational procedure or analysis method that can be trained to learn and generalize non-linear and complex rules and relations governing a set of inputs and associated outputs. As the name implies, an ANN is an imitation of the biological neural system in animals. In fact, its fundamental development was made by psychologists who tried to explain associative memory around the end of the 19th century (Poulton, 2001). Their work then led to the introduction by McCulloch and Pitts (1943) of a mathematical

model for neural excitement. In their model, a neuron works as a threshold logic unit that becomes active if the inputs it gets from the excitatory synapses exceed the threshold. Otherwise, or if it receives inputs from the inhibitory synapses, it stays inactive. This definition of the biological neuron then was used to explain how a neuron in the ANN works.

ANNs have developed throughout the 20th century, getting more attention and introduction into different fields by its revival in 1980's after the proposal of non-linear network algorithms (Poulton, 2001). The application of neural networks in geophysics dates to the late 1980's, when this technique was used to address magnetic and electromagnetic problems, and seismic data analysis, such as waveform recognition for first-break picking (van der Baan and Jutten, 2000; Poulton, 2002; Russell, 2005). The power of ANNs in seismic research is in their inherent ability to "learn" the complicated non-linear relationship between model parameters and data. Consequently, there is a growing trend of successful uses of ANN in predicting reservoir parameters such as porosity, permeability, V_S , and water saturation (Helle et al., 2001; Khoshdel and Riahi, 2011; Leite and Vidal, 2011b; Gholami et al., 2014; Cersósimo et al., 2016; Mohamed et al., 2017). ANN has also been used in lithology classifications, and detecting hydrocarbon migration patterns and traps (Clifford and Aminzadeh, 2011; Silva et al., 2013; Connolly, 2015; Araya-Polo et al., 2017).

V_P/V_S and Poisson's ratio (PR) are directly related, though not linearly, and both are sensitive to rock-fluid properties. P-wave velocity is significantly influenced by the presence of fluid, while S-wave velocity is known to be largely insensitive to it, which makes these ratios effective indicators of hydrocarbon reservoir characteristics. In this

study, we have used two methods, prestack simultaneous inversion and a Generalized Regression Neural Network (GRNN), to predict Poisson's ratio as a hydrocarbon indicator in Heidrun field, North Sea (Figure 1). Poisson's ratio was chosen after examining the results of ANN analysis in predicting both V_P/V_S and PR parameters. The simultaneous inversion result includes a V_P/V_S output, which was later converted to Poisson's ratio for the sake of comparability.

The seismic amplitude data sets were available in the form of a full-stack 3D volume, and time-migrated, NMO-corrected CMP gathers, with a 4-ms sample rate to total times of 4 and 3 seconds respectively. Five wells are used in this study, of which one is a production well, and the rest are appraisal wells. The drilling and logging information about formation tops and hydrocarbon shows were provided by the Norwegian Petroleum Directorate website (n.d.) for the four appraisal wells. This information is utilized in all analysis phases, while the producing well has been left as a "blind well" for assessment and verification purposes. Six horizons were picked based on structural and stratigraphic interpretations including: sea bed, two shallow control horizons, Brygge Formation, Fangst Group, and Åre Formation. In this study, the Fangst Group and Åre Formation tops represent the top and base of the pay zone respectively. An rms velocity model was built using available and well-tied depth-time logs, and the interpreted structural horizons. The analyses are done on full stack seismic volume, from 1000 to 3000 ms, with the focus on Fangst Group to Åre Formation as the pay zone (Figure 2).

One of the appraisal wells (6507/7-3) lacked the shear-wave and Poisson's ratio logs that were needed in inversion and ANN analysis. For the reservoir section (pay

zone), a shear-wave velocity log was calculated by the Greenberg-Castagna method. This method uses rock type, mineral and matrix parameters in a reservoir with a combination of lithologies and minerals to calculate a single V_P - V_S relationship (Dvorkin, 2007). For outside of the reservoir section, the mud-rock line was used to calculate shear-wave velocity. The equation's constants were calibrated to, and consistent with, other wells in the Heidrun field. A Poisson's ratio log was then calculated from the resultant V_P / V_S values. The reservoir was defined as sandstones from Fangst Group down to the oil-water-contact (OWC) depth at the Ror Formation claystones.

For both inversion and ANN, a 4-ms sample rate has been applied in making final volumes. The PR logs in the seismic sections have been resampled using a Backus averaging method (Backus, 1962) to represent the same sample rate for comparison purposes.

SIMULTANEOUS PRESTACK INVERSION METHODOLOGY

Simultaneous prestack inversion is a technique for predicting lithology and fluid properties using angle (or offset) gathers. It derives angle-dependent information from CDP stacks and estimates P-impedance (Z_P), S-impedance (Z_S), and density (ρ) through an iterative algorithm (Hampson and Russell, 2005). Inverting for the two latter parameters in addition to Z_P reduces the non-uniqueness risk of the solution, which is inherent within poststack inversion algorithms (Ma, 2002). The method we use draws on the relationship between the Z_P & Z_S , and Z_P & ρ couples based on Fatti's re-expression of Aki-Richard's equations (Aki and Richards, 1980), and the generalized Gardner's equation respectively (Hampson and Russell, 2005). The relationships

between the logarithm of Z_P and the logarithm of each of Z_S and ρ are assumed to be linear in the absence of complicating factors such as hydrocarbon presence. To reduce the instability issue inherent within the inversion problem, an iterative algorithm is chosen, which begins with an initial guess for Z_P and uses the conjugate-gradient method to approach the best possible solution step by step.

In order to initiate the inversion procedure, a fully processed CDP-gather seismic volume was first converted to a set of angle gathers. To estimate the output angle range, an rms velocity model was built using P-wave logs, and six interpreted structural horizons. The ray-parameter analysis then showed an effective range of angles from 0 to 45 degrees in the zone of interest (Figure 3). The angle gathers were then placed in 15 bins over the range of 0 to 45 degrees. A set of angle-dependent wavelets were extracted for every 15° to calculate a synthetic trace at each well location by convolving the reflectivity with the angle-dependent extracted wavelet. Extracted traces from each of these four wells are then correlated with the synthetics to obtain a better time-depth relationship (Figure 4).

The simultaneous inversion method used in this study is a model-based inversion. It uses an initial model (IM) of the desired parameter as a first guess from which it iterates to a better solution by reducing the misfit. To provide the IMs required for the model-based inversion, P-wave, S-wave, and density logs are used. For this purpose, the logs are interpolated between wells by interpreted structural horizons and an rms velocity model as the constraint, using squared inverse of distance. In the next step, a low-pass frequency filter is applied to the model, passing frequencies up to 10 Hz, and tapering higher frequencies up to 15 Hz as the high-cut frequency. This procedure is

followed to make IMs for Z_P , Z_S , and V_P/V_S values as well. The P-impedance IM obtained by this workflow serves as the initial guess and the low-frequency component that needs to be added to the band-limited seismic data set. A background trend is provided by relating the logarithm of Z_P & Z_S , and Z_P & ρ derived from well logs which is used by the inversion process to detect deviations from the trend.

The acquisition and processing steps impose an arbitrary scaling on the traces that has to be removed. For this purpose, a scaling factor is calculated for each of the three angle ranges by setting each trace equal to the synthetics at the well locations. This scaling factor is calculated as a ratio of rms of real amplitude and the average synthetic amplitude in an analysis window set over the non-pay zone. After getting satisfactory results by modifying the assigned model and scaling parameters, the seismic volume was inverted from 1000 to 3000 ms.

ARTIFICIAL NEURAL NETWORK METHODOLOGY

Different ANNs have been tested for PR and V_P/V_S targets. To train each network, a set of attributes has been extracted from the seismic volume using a stepwise-regression method. This method tries to find the best combination of attributes to predict a property by measuring the prediction error after adding one more attribute to the previously chosen best attribute(s) in each step (Hampson et al., 2001). Since each sample in the log (target) is usually related to a group of samples in neighboring seismic traces, the regression and training steps were done using convolutional operators. This means that each target sample is derived by convolving a set of attribute samples with assigned weights rather than multiplying them on a sample-by-sample basis.

The seismic attributes used for training consist of two types: internal attributes are the result of extracting information from the full-offset seismic amplitude volume (such as phase and frequency, integration or derivative operation on the data traces, etc.), and external attributes that are calculated and supplied as separate seismic files. To improve the training and validation results, we included the low-frequency models built for inversion, as attributes for training ANNs. These low-frequency IMs were imported as external attributes in the multi-attribute analysis. Including the IMs as external attributes significantly improves the network performance, and as expected, were among the most significant attributes in this step depending on the target type (Table 1).

After determining the appropriate set of attributes and operator lengths, different types of networks were trained and validated. Training and validation both use the same part of the data in which the target values are known; in our case, the well logs and attribute traces extracted at well locations. The difference is that the training set is used to modify the weights by which the inputs are related to the target values, while the validation set is used to evaluate the prediction and generalization ability of the trained network. For this purpose, the wells used in training were also used for validation in a cross-validation manner. In this way, we predict the target values for one hidden well at a time, using the other well logs. Based on the training and validation errors of different network types, a Generalized Regression Neural Network (GRNN) was selected and applied on the seismic data set to make the V_P/V_S and PR sections. GRNN is comparable to Probabilistic Neural Networks (PNN) in form, which causes it to be falsely referred to as PNN in many publications. A fundamental difference between these two is in the type of their output: PNN detects patterns and classes by setting decision boundaries,

while GRNN uses an estimation function in its output layer to predict continuous target values (Specht, 1991).

GRNNs apply a Gaussian basis function on the ‘distance’ between input data (attributes) and the unknown, weighted by training data. The so-called ‘distance’ refers to a measure of difference, such as the difference between attribute amplitudes in our case, rather than the Cartesian distance. These distances to the ‘unknown’ (target point) are scaled by a smoothing operator, sigma. By modifying these sigma values, the predicted values are compared to the known data (log values) using a conjugate gradient method. GRNN can be assumed to be an extension of PNN since the weighted distance is normalized by a measure of the probability estimate of the unknown target in the whole training data set (Russell et al., 2003).

For each case of PR and V_P/V_S , the average error and correlation is calculated at the well locations (Figure 5). The inconsistent behavior of wells in predicting each parameter (e.g., the average error of well 7-4) can be explained by noting that different types and number of attributes are chosen for each case. Based on these measures, and quantitative comparisons on seismic sections, PR was selected over V_P/V_S , and the associated network was then applied to the whole seismic volume to predict PR within it.

RESULTS

To compare the performance of the two procedures, different methods were applied in the analyses. We expected a procedure to be successful if its outcome could pass these criteria: 1) showing a meaningful pattern in sections and slices; 2) “matching” the logs quantitatively (misfit, and correlation) and qualitatively (matching low PR zones with

hydrocarbon shows in sections and slices); and, more importantly, 3) successfully predicting beyond the known values (logs, and structural horizons). The last criterion is the most significant factor in evaluating a method's ability in predicting potential hydrocarbon presence away from well control.

Approaching the comparisons quantitatively, the misfit between predicted and original logs are calculated at each well location and are shown in Figure 6. This misfit is presented as an error/difference profile for each well, and also is averaged over the whole calculation window for each well location. In addition to that, we have cross-validation criterion for ANN results shown in Figure 5, by which the effectiveness of the process itself can be measured. The other way of estimating the accuracy of results is to measure the correlation between the original and the predicted logs. This also is presented as an averaged value over the calculation window for each well (Figures 5 and 6), and in the form of cross-plots in Figure 7.

Sections and slices were drawn from *PR* volumes to examine the predictions at well locations and within the pay zone respectively. In each section (Figures 8, 9, and 10), model results are compared to the log data shown in a color column for two appraisal wells and one test well. Another section of two *PR* models is shown in Figure 11, along with the associated seismic amplitude section, to compare the structural features reflected in the model results. This figure is an expansion of Figure 9. Slices are created by averaging *PR* values from Fangst top, over a 50-ms window, to better investigate the lateral variations of *PR* values in the pay zone (Figure 12). The 50-ms window is chosen according to the observed time interval from the top of the pay zone to the OWC depth in most cases.

The time-slice figure indicates that the ANN result is more reliable in predicting the pay zone, and much more clearly shows the faults. This is clear by examining the resultant sections in comparison with the seismic data in Figure 11. The inversion result is primarily faithful to the pre-defined parameters (horizons and log data) when predicting properties in unknown zones. ANN on the other hand, although abiding by those pre-defined parameters (introduced to it via logs and IM-incorporated horizons), is not as profoundly dependent on them. This, in turn, can reduce the risk of showing a false positive result, for instance, the continuous low-*PR* layers found by the inversion. A fault-seal analysis can be conducted in similar studies to evaluate whether or not the fluid flow in the pay zone is consistent with the structural gaps shown in the ANN result. Both procedures show a similar low-*PR* response in the graben areas which should be further investigated since it can be caused by depositional settings and lithology changes different from the reservoir (Avseth et al., 2016).

Another drawback in the inversion results is the risk of presenting false positive responses, as is the case of a fairly continuous layer of low *PR* in well 6507/7-3 (Figure 8). This issue has been discussed by Avseth (2016) in detail, who suggested reasons, such as residual NMOs and underlying assumptions in approximating Zoeppritz equations, introducing more uncertainty in the inversion results.

CONCLUSION

Poisson's ratio values for both methods showed strong correlations with the known hydrocarbon shows at well locations as expected. Prestack inversion is shown to remain a reliable technique in providing subsurface information in very fine detail if coupled by separate geological and structural investigations. The results obtained by this method can

be used in reservoir characteristic assessments, assuming that the uncertainty level in all input parameters in model building and conditioning the results has been lowered as much as possible. Although this direct dependency of inversion results on the provided input makes it possible to monitor the procedure from beginning to end, but it may become an “Achilles heel” when it comes to reservoir evaluation beyond the known zones. One possible solution might be applying a geostatistical investigation and incorporating its results in inversion through the initial model, as a way to introduce a more localized pattern of variation for the target.

ANN outperformed simultaneous inversion in predicting *PR* in the pay zone in all cases, and matched better with the known data comparing the misfit and correlation parameters. Quantitative comparisons also confirmed this improved correlation except beneath the OWC depth in some cases, where inversion matched the higher frequency variations better. This might again be due to inversion’s stronger dependency on well logs and structural horizons provided through IMs.

The ANN technique is not an exception when it comes to the effect of provided input (training data) on its results. Including an IM as an external attribute has improved the results for each well location; however, the results show that ANN is not confined in its lateral behavior by that model. Considering the attributes that GRNN has been trained on, the type of sources provided might seem similar to what inversion has benefitted from (except the angle-dependent information). The key difference is in how this information is being handled within the network through training iterations; each reduces the output misfit by modifying the weights and function parameters associated with input values. The result is that the network has therefore learned the governing

relations between training sets and outputs, while not being bound within the limits of a particular input (e.g., following horizons despite structural disruptions). This makes the ANN results highly desirable in prospecting for new pay zones considering that: 1) ANN was able to show favorable zones beyond the drilled and logged (known) areas for the blind well, as well as for hydrocarbon shows shallower than logging depth, and 2) ANN results held a lower risk of showing false-positive responses. The latter can be related to the fact that ANNs, if not over-trained, can stay relatively insensitive to the noise in the training data. In an inversion procedure on the other hand, the uncertainties involved with the input parameters, noise, and the underlying physical assumptions are reflected in the results particularly away from the well controls.

This ANN methodology, after coupling the results with other controlling factors (geological, structural, etc.) can become an effective procedure in reservoir-characteristic assessments. It can increase the efficiency and accuracy of making decisions regarding new production well locations.

ACKNOWLEDGMENTS

The authors thank CGG Services (U.S.) Inc. for providing licenses for their Geosoftware package, including Hampson-Russel Suite modeling and analysis software. We also thank Halliburton Inc. for providing Landmark Graphics interpretation software through the Landmark University Grant Program. Dr. Mary Poulton provided many insightful comments and suggestions throughout the course of this work for which we are very grateful. We are also grateful to ConocoPhillips and its partners, Statoil, Petoro, and Eni for providing the Heidrun field dataset and permission to publish this work.

REFERENCES

- Aki, K., and Richards, P. G. 1980, Quantitative Seismology: Theory and Methods, W. H. Freeman and Co.
- Araya-Polo, M., Dahlke, T., Frogner, C., Zhang, C., Poggio, T., and Hohl, D. 2017, Automated Fault Detection without Seismic Processing, *The Leading Edge, Society of Exploration Geophysicists*, **36**. p. 208-214.
- Avseth, P., Janke, A., and Horn, F. 2016, AVO Inversion in Exploration — Key Learnings from a Norwegian Sea Prospect, *The Leading Edge*, **35**. p. 405-414.
- Backus, G. E. 1962, Long-Wave Elastic Anisotropy Produced by Horizontal Layering, *Journal of Geophysical Research*, **67**. p. 4427-4440.
- Cersósimo, D., Ravazzoli, C., and Martínez, R. 2016, Prediction of Lateral Variations in Reservoir Properties Throughout an Interpreted Seismic Horizon using an Artificial Neural Network, *The Leading Edge, Society of Exploration Geophysicists*, **35**. p. 265-269.
- Clifford, A., and Aminzadeh, F. 2011, Gas Detection from Absorption Attributes and Amplitude Versus Offset with Artificial Neural Networks in Grand Bay Field, *81st International Annual Meeting San Antonio, SEG Expanded Abstracts*, p. 375-380.
- Connolly, D. 2015, Visualization of Vertical Hydrocarbon Migration in Seismic Data: Case Studies from the Dutch North Sea, *Interpretation, Society of Exploration Geophysicists*, **3**. p. SX27.

Dvorkin, J. 2007, Yet another Vs Equation, SEG Technical Program Expanded Abstracts, **26**. p. 1570.

Gholami, R., Moradzadeh, A., Rasouli, V., and Hanachi, J. 2014, Shear Wave Velocity Prediction using Seismic Attributes and Well Log Data, *Acta Geophysica*, **62**. p. 818-848.

Goodway, B. 1999, Improved AVO Fluid Detection and Lithology Discrimination using Lamé Petrophysical Parameters; “ $\lambda\rho$ ”, “ $\mu\rho$ ”, & “ λ/M Fluid Stack”, from P and S Inversions, SEG Technical Program Expanded Abstracts, **16**. p. 183.

Hampson, D. P., and Russell, B. H. 2005, Simultaneous Inversion of Pre-Stack Seismic Data, SEG Technical Program Expanded Abstracts, **24**. p. 1633.

Hampson, D. P., Schuelke, J. S., and Quirein, J. A. 2001, Use of Multiattribute Transforms to Predict Log Properties from Seismic Data, *Geophysics*, **66**. p. 220-236.

Harris, N. B. 1989, Reservoir Geology of Fangst Group (Middle Jurassic), Heidrun Field, Offshore Mid-Norway, *AAPG Bulletin*, **73**. p. 1415-1435.

Helle, H. B., Bhatt, A., and Ursin, B. 2001, Porosity and Permeability Prediction from Wireline Logs using Artificial Neural Networks: A North Sea Case Study, *Geophysical Prospecting*, Blackwell Science Ltd, **49**. p. 431-444.

Huuse, M., and Feary, D. A. 2005, Seismic Inversion for Acoustic Impedance and Porosity of Cenozoic Cool-Water Carbonates on the Upper Continental Slope of the Great Australian Bight, *Marine Geology*, Elsevier B.V, **215**. p. 123-134.

Khoshdel, H., and Riahi, M. A. 2011, Multi Attribute Transform and Neural Network in Porosity Estimation of an Offshore Oil Field — A Case Study, Journal of Petroleum Science and Engineering, **78**. p. 740-747.

Leite, E. P., and Vidal, A. C. 2011a, 3D Porosity Prediction from Seismic Inversion and Neural Networks, Computers and Geosciences, Elsevier Ltd, **37**. p. 1174-1180.

Leite, E. P., and Vidal, A. C. 2011b, 3D Porosity Prediction from Seismic Inversion and Neural Networks, Computers & Geosciences, **37**. p. 1174-1180.

Ma, X. 2002, Simultaneous Inversion of Prestack Seismic Data for Rock Properties using Simulated Annealing, Geophysics, **67**. p. 1877.

McCulloch, W. S., and Pitts, W. 1943, A Logical Calculus of the Ideas Immanent in Nervous Activity, Bulletin of Mathematical Biophysics, **5**. p. 115-133.

Mohamed, I., Shenkar, O., and Mahmoud, H. 2017, Understanding Reservoir Heterogeneity through Water-Saturation Prediction Via Neural Network — A Case Study from Offshore Nile Delta, The Leading Edge, Society of Exploration Geophysicists, **36**. p. 298-303.

Norwegian Petroleum Directorate. Heidrun Field. Retrieved from:

<http://factpages.npd.no/factpages>

Poulton, M. M. 2001, Computational Neural Networks for Geophysical Data Processing, Pergamon.

Poulton, M. M. 2002, Neural Networks as an Intelligence Amplification Tool: A Review of Applications, *Geophysics*, **67**. p. 979-993.

Russell, B. H. 2005, Neural Networks Applications in Geophysics, CSEG Annual Convention, Expanded Abstracts, p. 339-341.

Russell, B. H., Lines, L. R., and Hampson, D. P. 2003, Application of the Radial Basis Function Neural Network to the Prediction of Log Properties from Seismic Attributes, *Exploration Geophysics*, Australian Society of Exploration Geophysicists, **34**. p. 15-23.

Silva, A., Neto, I. L., Carrasquilla, A., Misságia, R., Ceia, M., and Archilha, N. 2013, Neural Network Computing for Lithology Prediction of Carbonate- Siliciclastic Rocks using Elastic, Mineralogical and Petrographic Properties, p. 1055-1058.

Specht, D. F. 1991, A General Regression Neural Network, *TNN*, IEEE, **2**. United States, p. 568-576.

van der Baan, M., and Jutten, C. 2000, Neural Networks in Geophysical Applications, *Geophysics*, **65**. p. 1032-1047.

LIST OF FIGURES

Figure 1. The workflows of the two methods used in this study. Note that the inputs are similar, but the computational methods are different.

Figure 2. The stratigraphic sequence of the reservoir units with focus on Middle Jurassic formations of the Fangst Group in well 6507/7-4, adapted from Harris (1989). The logs on right are resistivity (ohm.m) and shale content (v/v), in measured depth.

Figure 3. Example of the prestack time-migrated and NMO-corrected CMP gathers. Effective angle range is from 0 to 45° in the pay zone (from Fangst Group to Åre Formation). Gathers are colored based on the incident angles.

Figure 4. Example of a well tie for well 6507/7-3. Synthetic traces (blue) calculated using near angle-dependent wavelets (a) extracted from angle gathers (center, black). Top right inset (b) shows the averaged result of cross-correlating synthetic with composite traces.

Figure 5. Average error (top) vs correlation measured at each well location after applying GRNN to predict PR (a), and VP/VVS (b). Training and cross-validation profiles are in black and red respectively. PR is mostly outperforming VP/VVS in both measurements.

Figure 6. (a): PR-log predicted by GRNN at well locations (in blue) overlain on the original log (in red). The difference logs in black curves show the misfit between the calculated and original logs. The averaged error for each well is noted above the misfit log columns, and the averaged correlation is noted above the prediction/original logs. Logs are in measured depth. (b): PR-log predicted by inversion at well locations (in blue) overlain on the original log (in red). The difference logs in black curves show the misfit between

the calculated and original logs. The averaged error for each well is noted above the misfit log columns, and the averaged correlation is noted above the prediction/original logs. Calculations are limited to the highlighted zone for each well.

Figure 7. (a): original PR-logs (x-axis) vs GRNN-derived PR values (y-axis) for each well, showing 96% average correlation coefficient. (b): original PR-logs (x-axis), vs inversion derived PR values (y-axis) for each well, showing 80% average correlation coefficient.

Figure 8. PR sections for well 6507/7-3: ANN model (a), and prestack inversion (b). Based on drilling reports, there were hydrocarbon shows in Fangst Group. ANN model is in comparable agreement with the log. Inversion model, while showing higher frequency variations, has predicted a false-positive response from Fangst Group, down to the OWC.

Figure 9. PR sections for well 6507/8-1: ANN model (a), and prestack inversion (b). Based on drilling reports, the only oil shows reported above the target reservoir were two samples from the Melke Formation. The Middle Jurassic Fangst Group and the Early Jurassic Tilje Formation reservoir sandstones were found to be hydrocarbon bearing. Both models are able to show Fangst Group and Tilje Formation low-PR values, ANN result is more consistent with the HC shows in log data from GOC to OWC depth. ANN model is also able to predict an anomalously low-PR zone at the depth of the Melke Formation.

Figure 10. PR sections for well 6507/7-A-53: ANN model (a), and prestack inversion (b). The insets show the predicted (blue) versus the original (red) PR logs. Depth to the tops, hydrocarbon shows, and OWC information are not mentioned in the drilling reports. This producing well is used as the blind well in both cases to evaluate the “generalization” ability of the two methods. The ANN model is in better agreement with the log above the

Fangst Group down to the inferred OWC, predicting a lower PR zone for a thickness more consistent with the log data. The Inversion model, unlike for other wells in the study, is not showing high-frequency variations at this well location. However, below the inferred OWC depth, the inversion result is more in agreement with the log data.

Figure 11. Seismic section (a), and the corresponding sections of the ANN (b), and inversion (c). Five interpreted faults are shown in seismic section. The equivalent location of the faults and a stratigraphic pinch-out are shown by white and black arrows in the two result sections respectively. The anomalously low-PR zone (purple and red) is disconnected at most cases in the ANN section, better reflecting actual structure. The inversion section shows a relatively smooth and continuous low-PR layer throughout the pay zone.

Figure 12. PR slices through the ANN (top), and inversion (bottom) models at the pay zone. Values are averaged over a window from Fangst Group to 50 ms below it.

TABLES

Target : PR			Target : $\log(Vp/Vs)$		
Final Attribute	Training Error	Validation Error	Final Attribute	Training Error	Validation Error
Sqrt(PR "IM")	0.021254	0.024536	(PR "IM") ²	0.183138	0.186143
Integrated Absolute Amplitude	0.018958	0.02232	1 / (Zs "IM")	0.171525	0.179414
Integrated Amplitude	0.017926	0.021597	Log(Zp "IM")	0.158163	0.165357
Average Frequency	0.017071	0.034934	Derivative	0.154537	0.164798
Cosine Instantaneous Phase	0.016426	0.033497	Integrated Absolute Amplitude	0.153038	0.163011
Zp "IM"	0.016064	0.033444	Dominant Frequency	0.151988	0.162846
Sqrt(Zs "IM")	0.014925	0.033956	Average Frequency	0.150947	0.164168
Porosity "IM"	0.014477	0.032524	Filter 25/30-35/40	0.150374	0.164741
Amplitude Weighted Frequency	0.014124	0.031601	Cosine Instantaneous Phase	0.149977	0.165633
Log(Vs "IM")	0.013894	0.031582	Amplitude Envelope	0.149528	0.166218

Table 1. Multiple attribute-regression-analysis results for PR (left) and Vp/Vs (right) as the target. First attribute has the lowest error, the second attribute is chosen in a way that the couple has the lowest error, etc. The top attributes in bold, are selected based on their validation error. The validation error, calculated using cross-validating method, decrease by adding each attribute until it reaches an optimum level, after which the error increases. Red boxes show the low-frequency IMs as external attributes.

FIGURES

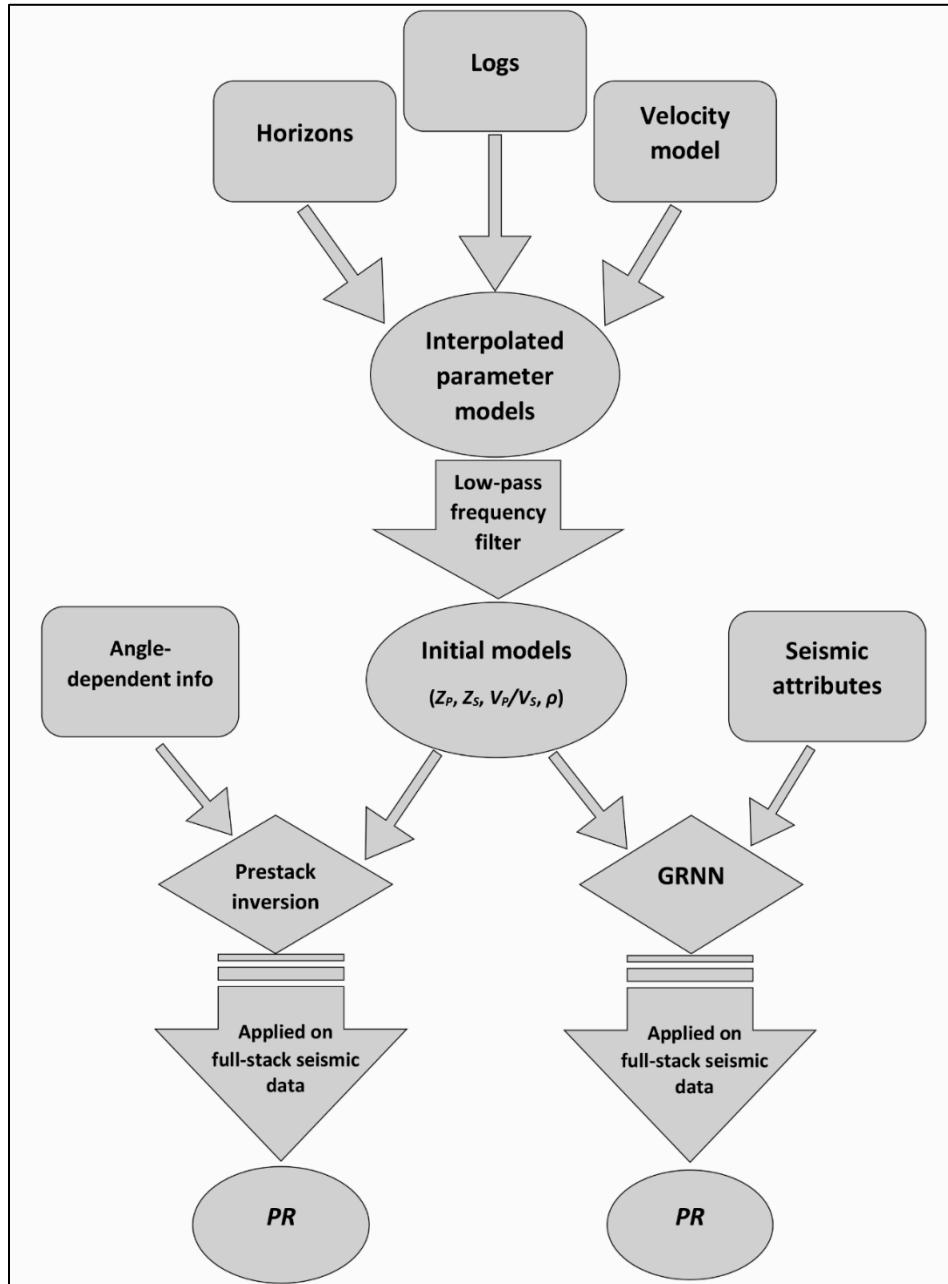


Figure 1. The workflows of the two methods used in this study. Note that the inputs are similar, but the computational methods are different.

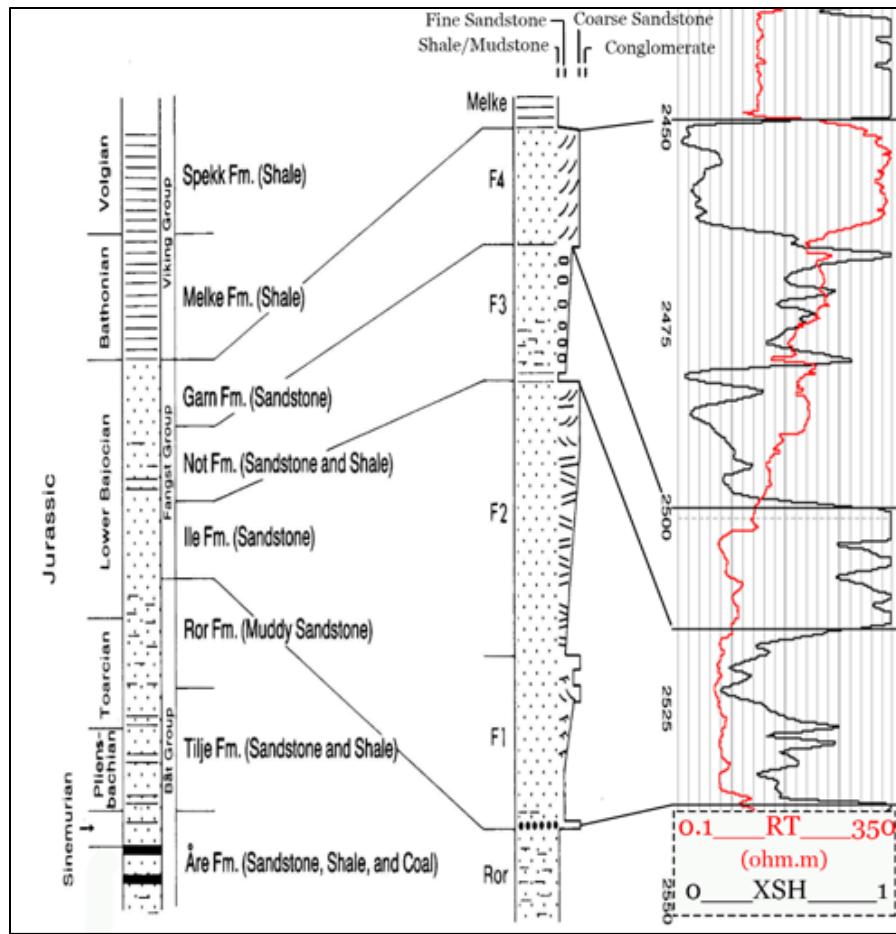


Figure 2. The stratigraphic sequence of the reservoir units with focus on Middle Jurassic formations of the Fangst Group in well 6507/7-4, adapted from Harris (1989). The logs on right are resistivity (ohm.m) and shale content (v/v), in measured depth.

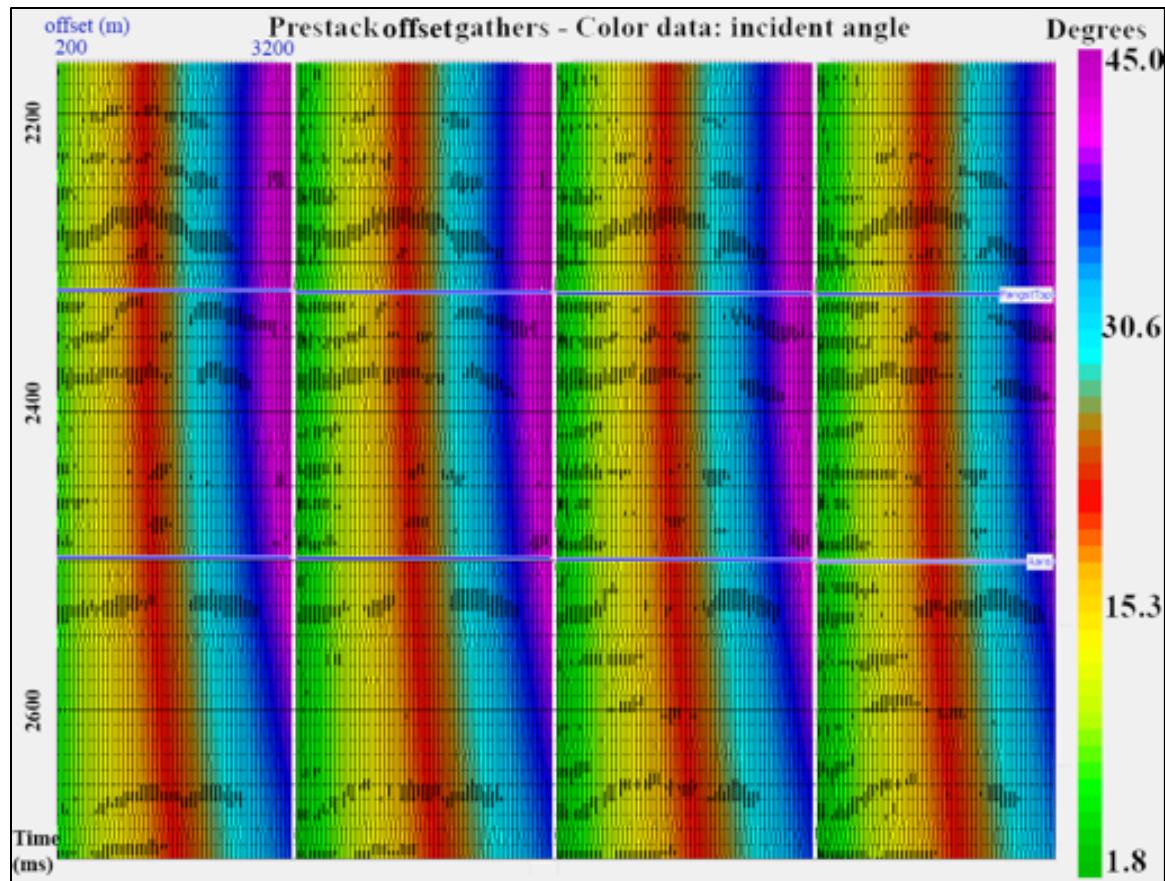


Figure 3. Example of the prestack time-migrated and NMO-corrected CMP gathers. Effective angle range is from 0 to 45° in the pay zone (from Fangst Group to Åre Formation). Gathers are colored based on the incident angles.

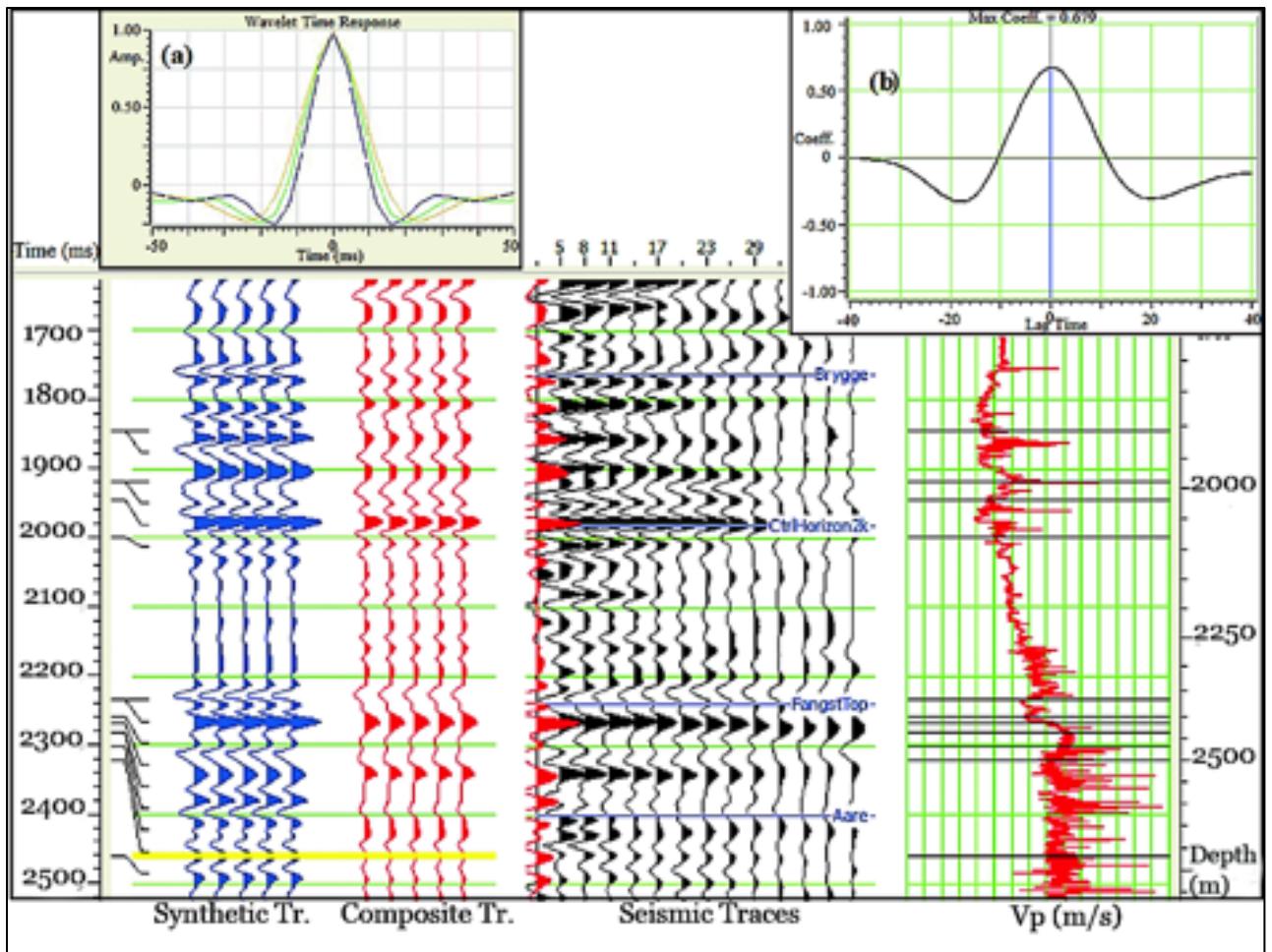


Figure 4. Example of a well tie for well 6507/7-3. Synthetic traces (blue) calculated using near angle-dependent wavelets (a) extracted from angle gathers (center, black). Top right inset (b) shows the averaged result of cross-correlating synthetic with composite traces.

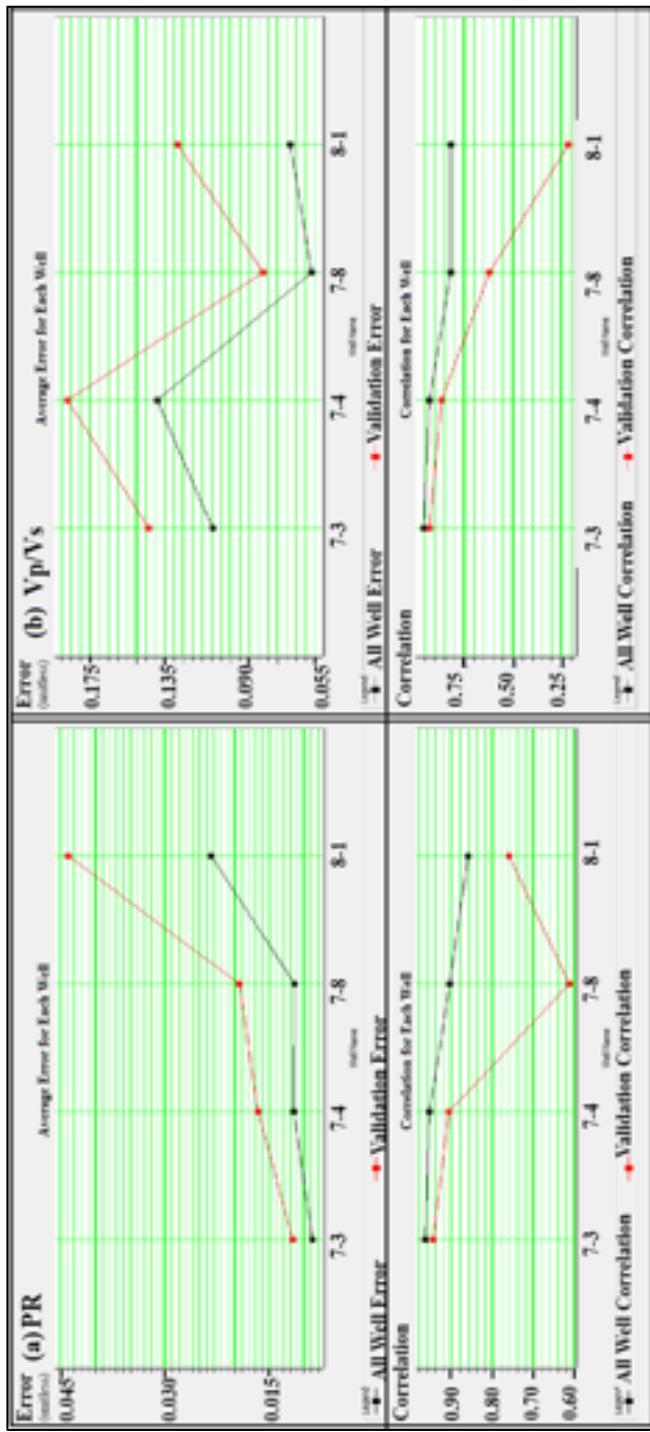


Figure 5. Average error (top) vs correlation measured at each well location after applying GRNN to predict PR (a), and V_p/V_s (b). Training and cross-validation profiles are in black and red respectively. PR is mostly outperforming V_p/V_s in both measurements.

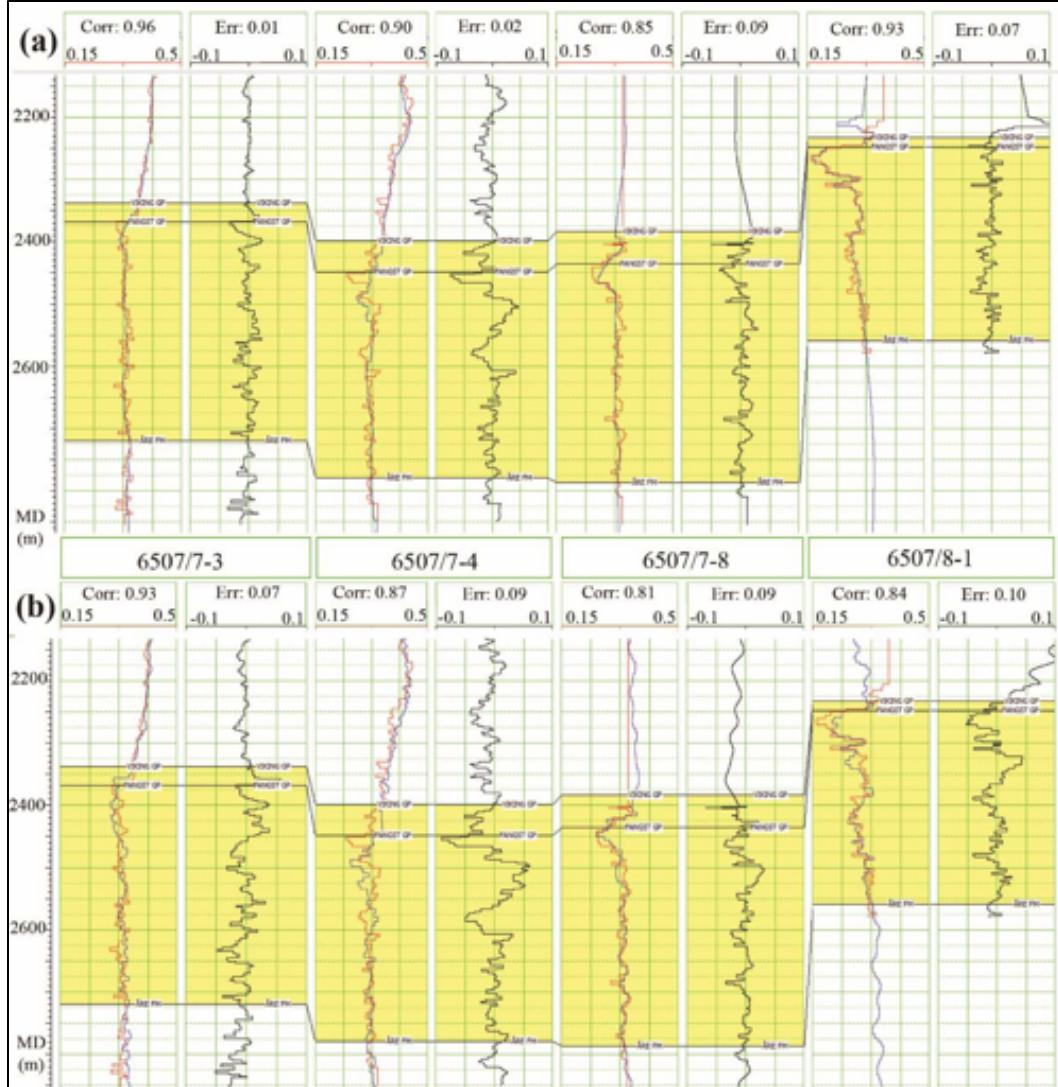


Figure 6. (a): *PR-log* predicted by GRNN at well locations (in blue) overlain on the original log (in red). The difference logs in black curves show the misfit between the calculated and original logs. The averaged error for each well is noted above the misfit log columns, and the averaged correlation is noted above the prediction/original logs. Logs are in measured depth. (b): *PR-log* predicted by inversion at well locations (in blue) overlain on the original log (in red). The difference logs in black curves show the misfit between the calculated and original logs. The averaged error for each well is noted above the misfit log columns, and the averaged correlation is noted above the prediction/original logs. Calculations are limited to the highlighted zone for each well.

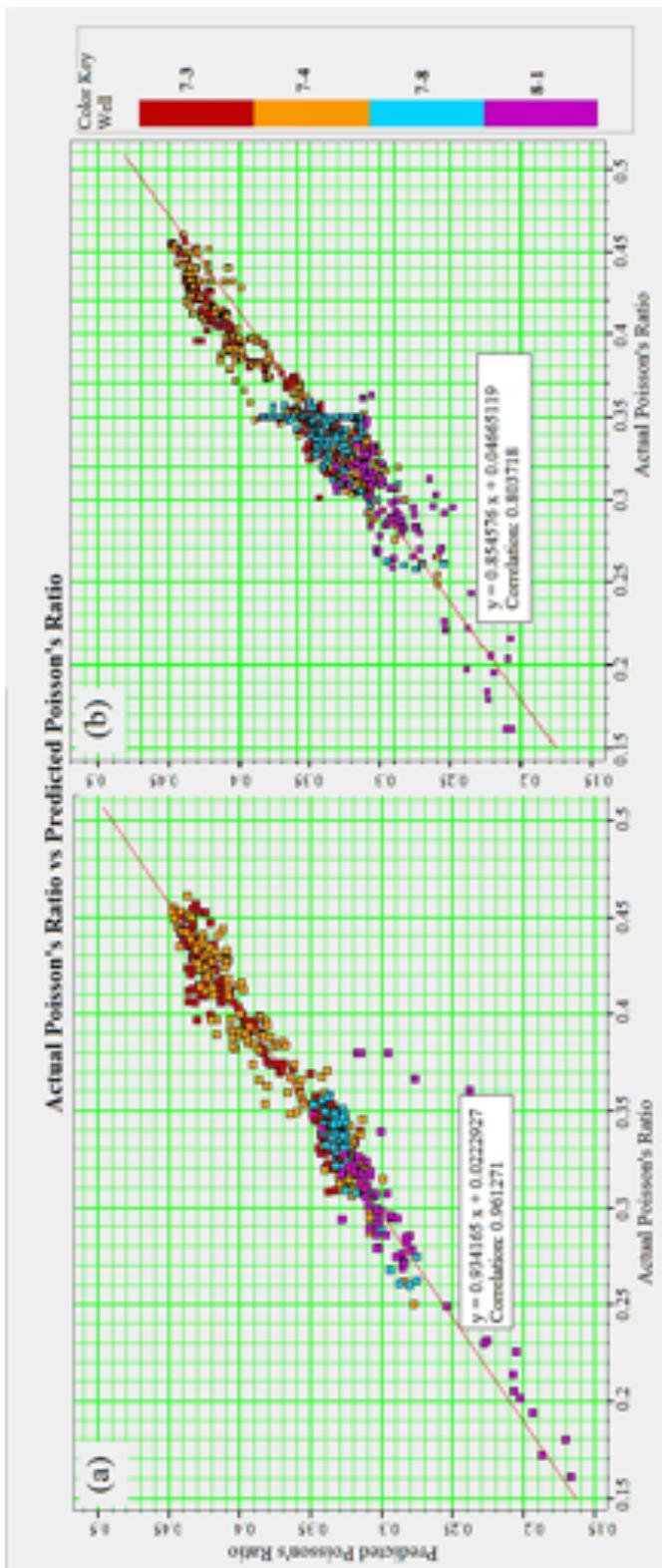


Figure 7. (a): original PR-logs (x-axis) vs GRNN-derived PR values (y-axis) for each well, showing 96% average correlation coefficient. (b): original PR -logs (x-axis), vs inversion derived PR values (y-axis) for each well, showing 80% average correlation coefficient.

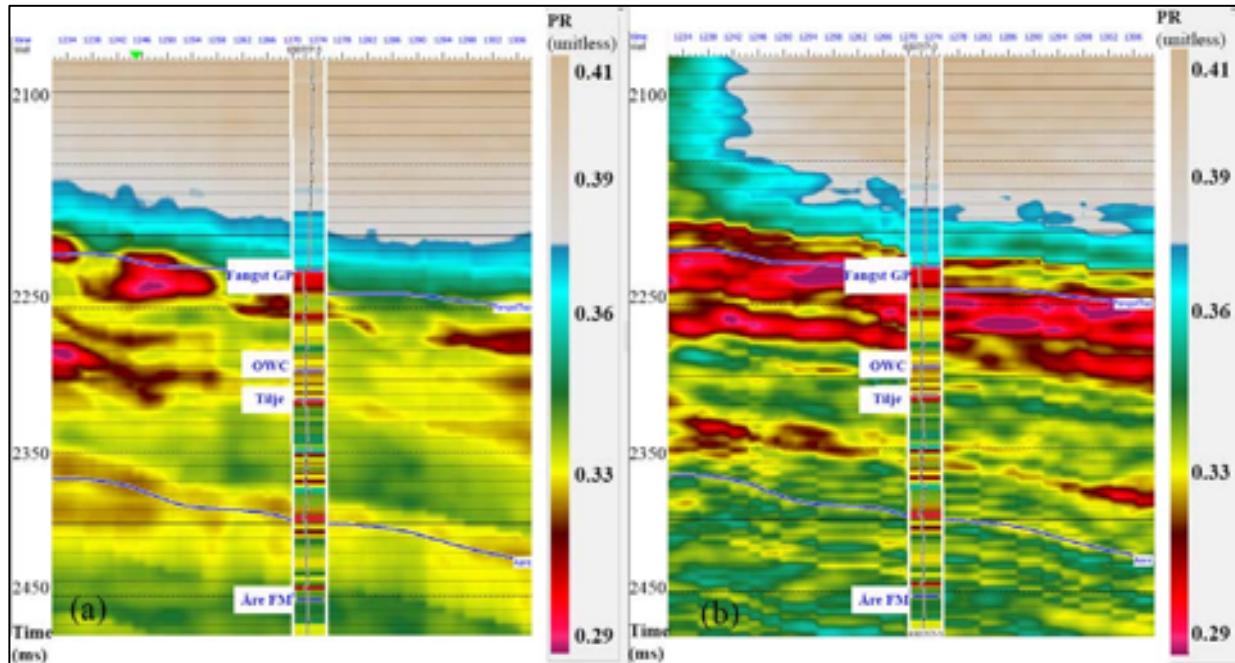


Figure 8. PR sections for well 6507/7-3: ANN model (a), and prestack inversion (b). Based on drilling reports, there were hydrocarbon shows in Fangst Group. ANN model is in comparable agreement with the log. Inversion model, while showing higher frequency variations, has predicted a false-positive response from Fangst Group, down to the OWC.

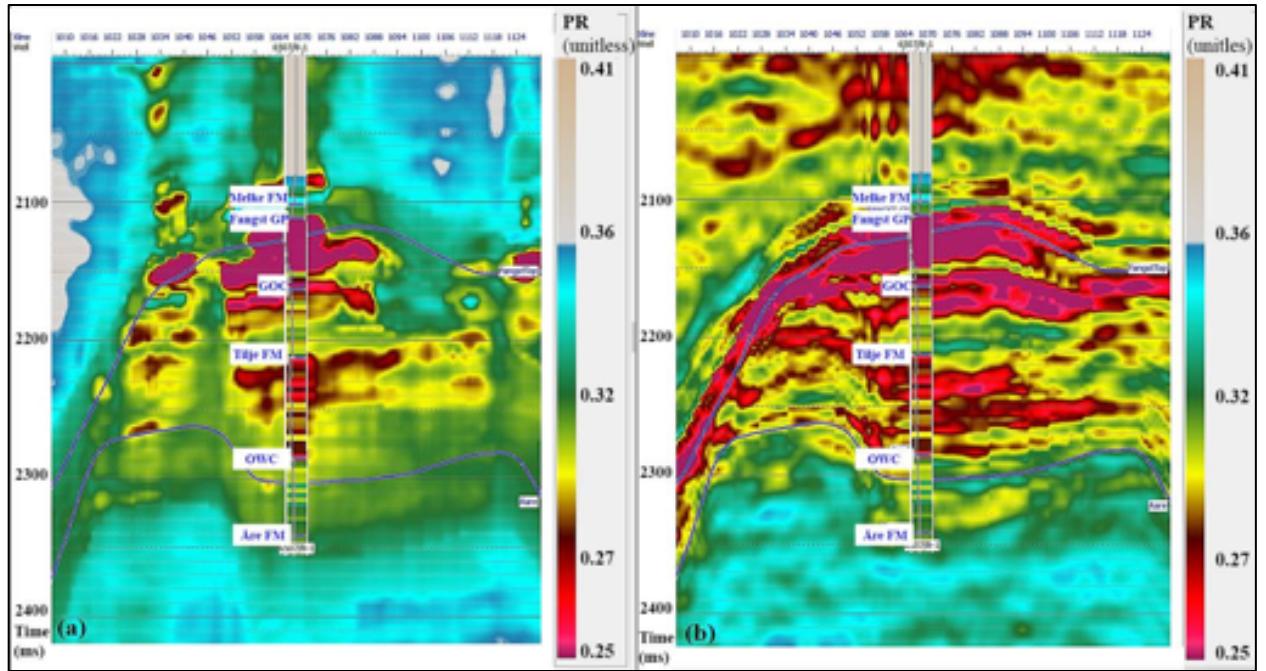


Figure 9. PR sections for well 6507/8-1: ANN model (a), and prestack inversion (b). Based on drilling reports, the only oil shows reported above the target reservoir were two samples from the Melke Formation. The Middle Jurassic Fangst Group and the Early Jurassic Tilje Formation reservoir sandstones were found to be hydrocarbon bearing. Both models are able to show Fangst Group and Tilje Formation low-PR values, ANN result is more consistent with the HC shows in log data from GOC to OWC depth. ANN model is also able to predict an anomalously low-PR zone at the depth of the Melke Formation.

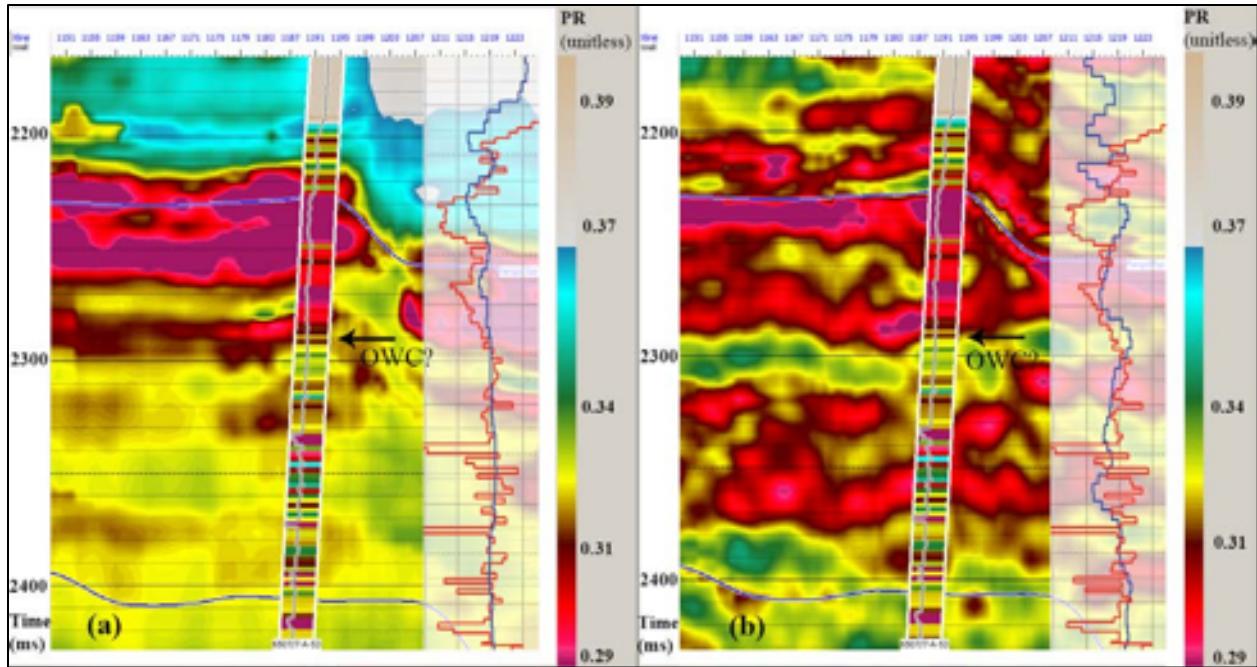


Figure 10. PR sections for well 6507/7-A-53: ANN model (a), and prestack inversion (b).

The insets show the predicted (blue) versus the original (red) PR logs. Depth to the tops, hydrocarbon shows, and OWC information are not mentioned in the drilling reports. This producing well is used as the blind well in both cases to evaluate the “generalization” ability of the two methods. The ANN model is in better agreement with the log above the Fangst Group down to the inferred OWC, predicting a lower PR zone for a thickness more consistent with the log data. The Inversion model, unlike for other wells in the study, is not showing high-frequency variations at this well location. However, below the inferred OWC depth, the inversion result is more in agreement with the log data.

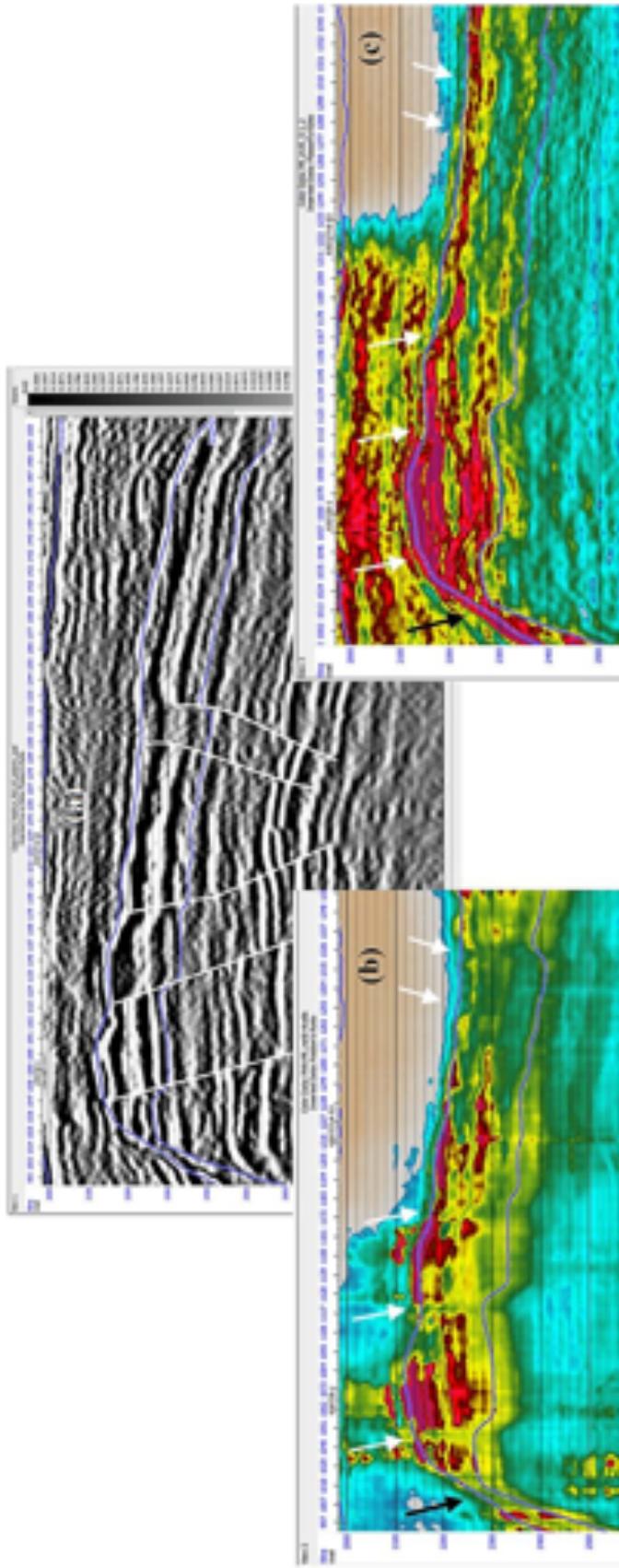


Figure 11. Seismic section (a), and the corresponding sections of the ANN (b), and inversion (c). Five interpreted faults are shown in seismic section. The equivalent location of the faults and a stratigraphic pinch-out are shown by white and black arrows in the two result sections respectively. The anomalously low-*PR* zone (purple and red) is disconnected at most cases in the ANN section, better reflecting actual structure. The inversion section shows a relatively smooth and continuous low-*PR* layer throughout the pay zone.

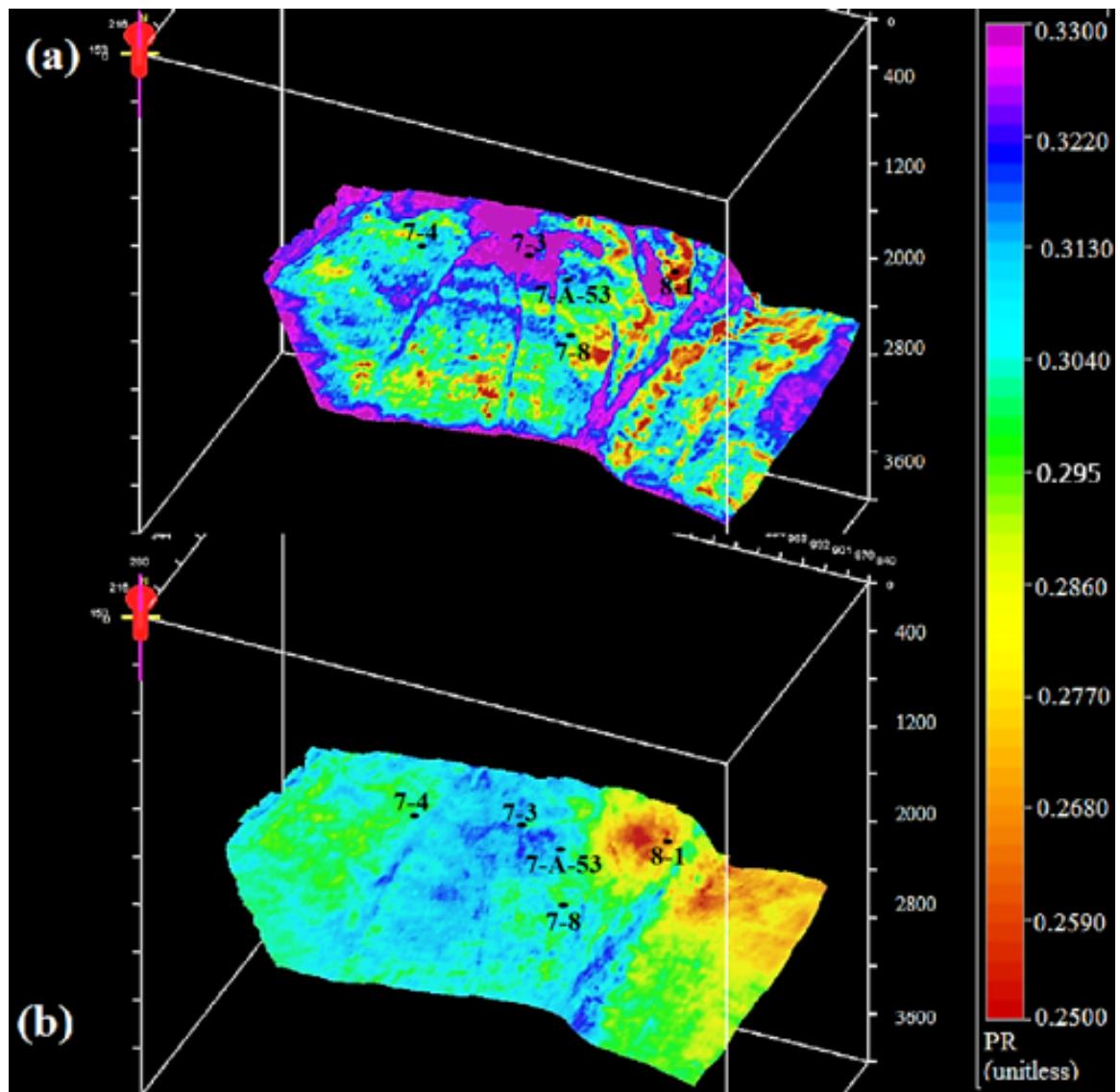


Figure 12. PR slices through the ANN (top), and inversion (bottom) models at the pay zone. Values are averaged over a window from Fangst Group to 50 ms below it.

**APPENDIX B: ASSESSMENT OF MACHINE LEARNING TECHNIQUES IN
PREDICTING LITHO-FLUID-FACIES LOGS IN HYDROCARBON WELLS**

Saba Keynejad¹, Marc L. Sbar¹, Roy A. Johnson¹

¹*University of Arizona, Tucson, AZ, USA*

Submitted to the Journal of Interpretation, 2018

ABSTRACT

Wireline log interpretation is a well-exercised procedure in the oil and gas industry with all its added value from exploration to production stages. It becomes even more important when it is one of only a few available alternatives to compensate for the lack of core samples in a study of lithological and fluid variations in a well. Yet, as with other purely-expert-oriented interpretational techniques, there is always a considerable risk of subjective or technical errors. We propose a hybrid approach that links a machine learning (ML) algorithm to the log interpretation procedure to solve these problems. We have applied this approach to two different hydrocarbon fields with the aim of predicting the hydrocarbon- (HC-) bearing units in form of litho-fluid facies logs at different well locations. The values of these logs are labels of classes that are separated based on their lithological and fluid content characteristics. After training different MLs on designed litho-fluid facies logs, we chose a bagged tree (BT) algorithm to predict these logs for the target wells due to its superior performance. This algorithm predicted HC-units in an accurate interval (above the HC-fluid contact depth) and showed a very low false discovery rate. The high accuracy rate, speed of analysis, and its generalization ability, even in data deficient cases, accentuate why including ML algorithms can improve the understanding of the subsurface at every phase of the exploration and production process. The proposed approach of utilizing ML algorithms, trained and tuned based on expert's knowledge of the reservoir, can be modified and applied to future wells in a hydrocarbon field to significantly minimize the risk of false HC discoveries.

INTRODUCTION

Boreholes and calibration wells are “direct” sources of information in the study of lithology and fluid content in a hydrocarbon reservoir. The cores and drilling reports, where available and well documented, provide us with this first-hand observation of subsurface characteristics. In those cases where this source is not reliable (e.g. low core-recovery ratio or lack of samples at a certain depth), or not available at all, wireline logging can fill the gap to help with interpretation of the lithofacies and potential fluid column at a well location. This interpretation process is not always easy, straightforward, or accurate. In fact, it is usually time consuming, subject to human error, and may be affected by other complicating factors such as poor log calibrations, conflicting log results, lack of data, etc. A better alternative is to use a hybrid approach that implements a computational method to retrieve as much objective information as possible from the logs, while inserting the first-hand information by the analyst to model the vertical lithologic and fluid variations in each well.

One of the quantitative techniques for the first part of this hybrid approach is Machine Learning (ML). ML algorithms are data-driven techniques that can learn the intended properties within a data set such as classes and trends, and then extract those features from un-seen data as well. Different techniques usually fall into one of the categories of supervised or unsupervised algorithms.

Supervised ML algorithms build predictive models with known outputs for different observations. These outputs can be “continuous” values or class labels, for which regression models (function approximators) and classifiers are used, respectively. Unsupervised algorithms, on the other hand, discover some measure of proximity or shared

features within the learning set to subdivide the data into clusters, without knowing the desired outputs. Either way, there is always some measure of training performance, the improvement of which facilitates algorithm performance on the unseen data.

Despite the fast-growing number of successful cases published in recent years, application of ML in hydrocarbon exploration is not yet established as a standard procedure within the Exploration & Production world. Among different fields of application, ML algorithms have enjoyed increasing attention in facies-recognition studies in recent years. Artificial Neural Networks (ANN) have been used in many cases, for instance in complex carbonate reservoirs, where lithofacies recognition based on wireline logs can be very challenging (Qi and Carr, 2006; Al Moqbel and Wang, 2011). Torres and Reverón (2014), and Zhao et al. (2014) have used different Support Vector Machine (SVM) approaches for lithofacies classification in reservoir modeling. Other ML algorithms such as generalized boosted regression modeling and quadratic Discriminant Analysis (DA) have also been successfully tested and applied by Aleardi and Ciabarri (2017) and Al-Mudhafar (2017).

In this study, we are proposing a hybrid approach that merges data-driven (ML) and knowledge-driven methods. The former yields superior computational power, accuracy, and resolution, while the interpreter can get around the potential data deficiency and/or acquisition problems through the latter. We have applied this technique to two different data sets to create litho-fluid facies logs with the main goal of predicting HC-bearing units. This approach can be modified for different fields depending on their data availabilities and reservoir characteristics.

The first step of this framework includes data preparation, qualitatively and statistically investigating the logs, and then designing target litho-fluid facies logs

accordingly. In this step, the expert defines the number and type of target classes and chooses the logs (referred to as features, attributes, or predictors) that demonstrate a desired degree of relevancy to the targets. These logs and classes will be used to train the algorithm on the wells with known litho-fluid facies logs (training wells). After obtaining satisfying results in training, the algorithm can be applied on the un-seen wells (target wells) to predict the litho-fluid facies classes.

In this study, different ML algorithms are tested on two fields: a self-organizing feature map (SOM) as an unsupervised ANN, and a multilayer feed-forward neural network (MLFN) and a bagged tree (BT) classifier as supervised algorithms. We have implemented the same workflow on both fields to better assess the methodologies. All algorithms are trained on each data set, and after evaluating the training results (qualitatively, and/or quantitatively) the best algorithm (BT) is applied to the target wells. Depending on the available information from each data set, we then have assessed our predictions using a previously interpreted reference in each case.

The available data sets for this analysis include four boreholes from Heidrun Field, off-shore Norway, and eight boreholes from Kupe Field in Taranaki Basin, off-shore New Zealand. Heidrun is a producing oil field with associated gas and Kupe is a gas-condensate field. Both fields are clastic reservoirs, but with distinct properties such as the amount of alternating shale layers and the quality of reservoir sand. Even though the framework is the same, the parameters of the algorithms for each field were designed separately based on available wireline logs and the facies types. Each data set includes different wireline logs accompanied by drilling information available for some wells. These are the primary sources of information needed by the expert to design and train the algorithm, and then

interpret the results of its application. According to the objectives of this study, we chose to keep the class labels to a minimum, regardless of the stratigraphic and depositional factors, to highlight the HC-bearing units. To evaluate the methodology and results, we have excluded parts of the available data, either in the form of one whole well, and/or a certain depth interval. The excluded parts form our target wells and sections, while the rest are used as training wells.

This study shows promising results in applying ML algorithms in hydrocarbon exploration. SOM can be used to define target clusters in pilot stages of exploration or underexplored areas where no cored boreholes are available for training. Supervised algorithms, however, could successfully predict HC units in both fields, consistent with fluid contact depths and without false discoveries. This can help in lowering the risk of overestimating a reservoir's capacity in field development stages.

HEIDRUN FIELD

Located on the mid-Norwegian continental shelf, Heidrun Field was formed in an extensional tectonic phase during the Late Jurassic to Early Cretaceous. The Middle Jurassic Fangst Group clastic reservoir was deposited in a shallow marine to fluvial environment. It comprises three mostly clean sandstone formations, with the Not Formation being the more shaly layer that thickens toward the southwest (~19 m in well 6507/7-4). For more on the stratigraphy of the Fangst reservoir and Heidrun Field geology, see Harris (1989) and Morton (2009). In this paper, we have focused on a section within four wells that includes the Melke Formation shale (Viking Group, Middle to Upper Jurassic), the Fangst Group, and the underlying Ror muddy sandstone and Tilje sandstone and shale (Båt Group, Early Jurassic).

To create the training database, we assumed that only one well (6508/8-1) has complete drilling information and another one (6508/7-4) has partial records, as if the core samples of the second well were lost for a certain depth interval. This excluded section (2480-2600m) is, in fact, part of the oil-bearing sandstone in the reservoir. The other two wells (6507/7-3 and 6507/7-8) are assumed to lack any core samples. In this way we can “mask” the excluded parts of our available information only to use them later for testing the approach and validating the results.

The compiled database includes depths and the associated attribute values extracted from in-situ and computed logs. To create the litho-fluid facies logs, crossplots of different logs were analyzed. According to drilling information provided by Norwegian Petroleum Directorate (n.d.), four classes are defined as the dominant facies shaping the reservoir: shale (including mudstone and siltstone), brine-sand, oil-sand, and gas-sand. Since well 6507/8-1 is the only well with gas-sand samples, we chose to use it in the training phase, and evaluate any potential false discovery of gas-sands in other wells. Also note that most ML algorithms cannot predict new classes beyond the limited set of classes on which they are trained. For reservoirs where encountering new facies in un-seen wells are expected, a series of neural networks (ART2) with capability of expanding predicted clusters can be used (Chang et al., 2000).

The best separation was seen in LMR (Lambda-mu-rho) analysis on a crossplot of $\lambda\rho$ (LR) vs $\mu\rho$ (MR) logs, noting that incompressibility (λ) and rigidity (μ) are, by definition, pore-fluid and rock-matrix indicators, respectively. These logs are computed using in-situ logs (V_p , V_s , density) as follows:

$\mu\rho = Z_S^2$, $\lambda\rho = Z_P^2 - 2Z_S^2$, where Z_P and Z_S are P-wave and S-wave acoustic impedances, respectively.

Goodway (1997) compares the Lamé constants embedded in P-wave velocity or in a ratio (λ/μ or V_P/V_S) with LR and MR coefficients, and demonstrated a significant increase in sensitivity toward petrophysical variations with LR and MR coefficients. LMR analysis shows promising results in petrophysical discrimination in hydrocarbon reservoirs. Young and Tatham (2007) identify gas sands by applying LMR inversion on young, unconsolidated sediments. LMR analysis can also become a crucial tool in reservoirs where a decrease in V_P/V_S occurs without a pore-fluid-related increase in V_S (Close et al., 2016).

The results of LMR analysis were used to create the initial litho-fluid facies logs (classes). Each interval was then approved or modified by the reference information to form the class labels for our database. This is the “confirmed” part of the data set on which the algorithms will be trained, and based on that, the missing information will be recovered/predicted.

The other attributes (logs) were selected by plotting the probabilistic density function (PDF) histogram of each log for each target class. Among all wireline logs that were available and well-recorded within the reservoir interval at all well locations, six of them were interpreted to be good target indicators. These six logs, including LR, MR, Porosity (Phi), Poisson’s ratio (PR), Shale volume (V_{Sh}), and P-impedance (Z_P), along with depth, form the seven predicting attributes for Heidrun Field. We found that including depth as an attribute can improve the HC-prediction rate since it can act as a constraint in our 1-D data sets.

KUPE FIELD

The Paleocene Farewell Formation is the primary reservoir for the Kupe gas-condensate Field, in the southeastern Taranaki Basin, off-shore New Zealand. The Farewell Formation is comprised of medium- to coarse-grained sandstones with interbedded shale layers, deposited in fluvial to coastal braided plains (Pang and Collen, 1996). Our primary focus in this study is on sections in eight wells that contain the Farewell sandstone reservoir. The reservoir quality of the Farewell Formation can be highly affected in some parts depending on the type of clay minerals within the sandstone (Martin et al., 1994).

We used four wells (Kupe South-6, Kupe South-7ST1, Kupe South-8, and Momoho-1) to train the ML algorithms for predicting the litho-fluid facies in other wells. The training wells, unlike four target wells, have Vs logs which were used to create LMR and PR logs and crossplots. We also used crossplots of Gamma-ray (GR) vs the difference between deep and shallow resistivity logs (RESD and RESS) to design the litho-fluid facies target logs. Based on these analyses and the available drilling reports (New Zealand Petroleum & Minerals, 2017), four classes are defined: shale, brine-sand, HC-sand, and HC-shaly sand. The latter two classes are both HC-bearing sandstones with different degrees of shale content as an indicator of reservoir quality. The defined classes can be different in any field depending on the reservoir, data availability, or the objective of the study. For instance, minor coal seams occurring throughout Farewell sandstones in some wells did not improve the prediction when included, nor did they relate to the objective of this paper.

The litho-fluid facies logs were then compared to the existing composite logs so that the defined classes and possible fluid contacts are in agreement with the drilling information. This corrects for any misclassifications due to inaccurate wireline log

measurements and/or potentially oversimplified labeling. Based on our investigations of available wireline logs, GR, resistivity, and Z_P logs were selected along with depth values, for predicting litho-fluid facies by ML algorithms. Note that to train and apply the algorithms, the choice of predictor logs is limited to the ones that are available in target wells. So, the LR and PR logs, which were used in creating the litho-fluid facies classes, could not be included in training since the target wells did not have Vs (and consequently, LR and PR) logs. The parallel coordinates plot in Figure 1 shows how each of the selected predicting features can separate the target classes. It also shows the importance of ML algorithms by noticing the challenge of separating overlapping classes in a multi-dimensional attribute space.

APPLICATION OF ML ALGORITHMS

“Machine learning” can be thought of as a self-explanatory term, noting that the word “machine” implies a wider spectrum of definitions such as computational modeling. In other words, any data analysis method, with the aim of discovering trends and features within a data set, and from the data set itself, without relying on a predetermined equation is considered as a machine learning algorithm. The algorithm learns the desired information from the data set, which is, in essence, similar to the procedure of learning from experience in human beings.

ANNs are a particular branch of ML, originally inspired by biological neural networks. ANNs are well known as powerful ML techniques in solving complex and non-linear problems governing a large amount of multi-attribute data sets. Supervised ANNs take in the observations (e.g., logs) and relate them to the associated targets (e.g., known lithofacies) via non-linear activation functions of their innermost hidden layer(s).

Unsupervised algorithms on the other hand, have the advantage of not depending on the availability of known targets for making predictions. But for the same reason, their results need to be evaluated more rigorously and validated by some reliable evidence before being authorized in a decision-making procedure.

The number of layers, their neurons, the connection weights between layers, and the direction of information flow in an ANN defines its architecture, which in turn depends on the problem. The training algorithms are optimization problems that minimize the difference between prediction and the target by tuning the connection weight values accordingly. For a history of the application of ANNs in geophysical studies until 2002, see Poulton (2002).

The bagging or **bootstrap aggregating** method is generally used to increase stability and reduce the variance of an ML algorithm. This procedure is typically applied to the classification and regression trees (CART) which tend to overfit the training data. CARTs are decision-making models that relate features via “branches” to a “leaf” which is a predicted value (regression) or label (classification). The branches split at several points based on variations in the features.

The bagged-tree (BT) procedure begins by creating replicates of the learning set through bootstrap resampling, i.e., random sampling with replacement from the training data set (Breiman, 1996). This way each decision tree is trained on a bootstrap sample set, the outcome of which is an aggregation of the results through “voting” or averaging for classification and regression problems, respectively.

ANN

SOM - SOM is an unsupervised ANN that clusters data while preserving their topological relationships. SOM uses the Kohonen rule in a competitive layer, by which the winning neurons are determined based on their ‘closeness’ to the prototype or initial vectors (Kohonen, 1987). In each iteration, the connection weights to the winning neurons will be updated, putting the neurons with similar features in one cluster. SOM entered the seismic processing field mainly as a tool in horizon tracking and waveform recognition, and is still being used in facies mapping studies in more innovative workflows (Liu et al., 2017; Zhao et al., 2017).

We used SOM as an initial and alternative data-driven method to estimate the number of clusters, independently from the LMR-based clustering results. For this purpose, we measured the silhouette parameter (as defined below) to evaluate the optimum number of SOM-driven clusters. This parameter is calculated as a ratio for each point, using a measure of dissimilarity such as Euclidean distance, correlation, cosine of two vectors, etc., as follows:

$$S_i = \frac{B_i - A_i}{\max(A_i, B_i)},$$
 where A_i is the average distance between point i and the rest of the points in the same cluster, and B_i is the distance between that point and the points in the closest cluster (Rousseeuw, 1987).

After running SOM for feature maps with 3, 4, 5, and 6 neurons (each neuron representing a cluster) and calculating the averages of S_i 's in each case, the optimum number of clusters determined by the silhouette parameter was 4 for both fields (Figure 2). We then applied the SOM network trained for a four-class scenario on each data set.

Because of the unsupervised nature of this network, we must label its predicted clusters after it has been applied.

MLFN - An MLFN is also trained and applied to predict the litho-fluid facies logs. MLFN is one of the common supervised ANNs that can be used in classification and function-approximation problems. It has at least one hidden layer and updates the connection weights and biases in a backward direction (backpropagation) using one of the many gradient-based training algorithms available. The output layer uses either a sigmoid or a linear transfer function for classification and regression problems, respectively.

We used a single-hidden layered MLFN with 10 neurons as a classifier to predict litho-fluid facies logs in the two fields separately. The network uses 15% of the training data, randomly selected, for validating the results. The validation value is a generalization measure that estimates the algorithm's capability in reproducing what is learned in the new "un-seen" data set.

Bagged Trees

In this study, we have used a random-forest approach to grow the bagged trees for the BT algorithm. This means that, in addition to the bootstrap sampling of the training data set for each tree, the features (predictors or log values) at each split are also randomly sampled without replacement (Breiman, 2001). The training procedure can be summarized as follows:

1. Take a bootstrapped sample of the training set to form a tree.
2. Form a splitting node by randomly sampling features (e.g., splitting a branch into $PR > 0.3$ and $PR < 0.3$).
3. Repeat step 2 at each split in the tree until grown as large as desired.

4. Repeat steps 1-3 for a large number of trees.
5. Over all the trees, count the number of predicted classes for a specific observation¹.
6. The majority vote of an observation determines the class assigned to it.

Brieman (2001) showed that the combination of voting over a large number of trees and random sampling of features results in reduction of both variance and bias. Thus, there is no need for pruning the trees since the overfitting problem is already addressed by these two qualities. Figure 3 shows a small part of one of the trained trees as an example.

We tested different ML algorithms (including SVM and DA) trained on the same training data set with a 5-fold cross-validation factor. In this procedure, the data are subdivided into five sections (folds) during the training phase and at each turn one section is predicted by the rest. The accuracy of the algorithm is calculated based on the average of these cross-validation errors. In all trial sessions of training for this study, BT constantly had the highest accuracy among all available ML algorithms.

DISCUSSION AND RESULTS

SOM is an unsupervised method that clusters the data based on similarities without using labels (targets) and, thus, the results of the clustering need interpretive validations after being applied. Our interpretation of the implied litho-fluid facies labels is based on investigating the results in cross-plots and PDF histograms.

The best separation among SOM clusters occurred when the algorithm looked for four classes within each data set, according to the silhouette value. This was in agreement with

¹ In BT, a set of all the splitting nodes on a branch that leads to a certain leaf is called an observation.

our assumption about the target classes. However, a preliminary investigation of the SOM-driven litho-fluid facies logs shows a disagreement between the predicted and actual fluid content (Figure 4). In most cases the predicted HC column extends deeper than the lowest HC observed in the wells (OWC or LKG). Also, in Heidrun Field, gas-sand only exists in well 6507/8-1, while thin layers of this class are predicted in other wells by SOM.

According to the goals of our study, the SOM results were not reliable and thus it was not considered for simulating on the target wells. Ross and Cole (2017) have compared the results of MLFN with unsupervised ANNs in a seismic facies mapping case and concluded that the latter is not well suited for solving non-linear problems as in facies classification.

Supervised training

Two supervised ML algorithms, BT and MLFN, were trained on two data sets. We trained each algorithm seven times independently (not retraining) and then selected the median values to compare the accuracy of the two techniques. The reason behind multiple training sessions is to have a measure of the “stability” of each algorithm, since ML algorithms usually (if not intentionally designed otherwise) make use of randomness at some point in training. For example, the initiation of weights in ANN, the observation order in BT, and sampling data to create validation subgroup(s) happen in a random state. This causes the non-repeatability of the exact training results in different runs of the algorithm. Accuracy values and ranges, and the success rate in predicting HC and non-HC units are summarized in Table 1. The validation errors for MLFN and BT are reflected in their accuracy values.

An important thing to note in this table is the range or spread of accuracy results depicted by the heights of the box plots. BT appears to be a more stable algorithm due to

its lower variation in accuracies (smaller range) in both fields. It means that despite the inherent non-repeatability of results each time the algorithm is initiated, BT can be expected to have an almost predictable accuracy ($\sim\pm0.5\%$). It also shows a higher overall and average accuracy in both fields. This capability of BTs in generalizing results with minimum overfitting lowers the risk of false discoveries.

The application of BTs can be viewed as going through a series of massive networks of if-then rules that were initially extracted from the training data set. Thus, it is a natural choice to use BTs in wireline log interpretation, which is, in fact, a decision-making task based on how each log is behaving.

The success-rate values show the positive predictions within each category. For instance, in Kupe Field, 89.8% of HC units (two HC classes combined) predicted by BT belong to the HC units in the target log, which means 10.2% of what is predicted as HC units, in fact belong to non-HC units. Based on this parameter, both algorithms are more successful in predicting HC units than non-HC units. However, BT has higher success rates in predicting either HC and non-HC units in comparison with MLFN.

Based on these results we chose the BT algorithm to predict the litho-fluid facies logs in each field, the results of which are discussed in the following sections. Figure 5 shows the BT prediction of litho-fluid facies logs in one well at each field, along with some of the wireline logs as reference. As seen in this figure and discussed before, in comparison with the Heidrun Field, there is a larger amount of shale content in the Kupe Field. The shale content, either in the form of alternating layers or clay content within the sandstones, intensively affects the wireline logs, which makes the qualitative log interpretations more difficult and less accurate.

Heidrun litho-fluid facies logs

The data were analyzed qualitatively and using LMR and statistical analysis to find the well logs that can best predict litho-fluid facies. In all cases, PR, LR and Z_P were the best features in distinguishing between all classes, while the other logs had varying performances in separating different classes. In all cases, shale and brine-sand were found to be difficult to distinguish from each other. One reason could be that the shale class is underrepresented in the Heidrun data. The Fangst Group is mainly shallow-marine and fluvial sandstone deposited during the Late Jurassic regression in off-shore mid-Norway and includes one thin marine-shale layer in the Not Formation (Harris, 1989). Two shaly classes are included in data that are outside the reservoir: the overlying Melke marine-shale of the Viking Group, and the underlying Ror and Tilje muddy sandstone and sandstone and shale from the Båt Group. However, these layers are not clean shales or thick enough to completely overcome the “bias” in the number of samples. A sandy-shale target sample for example is easy to be misclassified with brine-sand due to similar Z_P and LR values. An underrepresented class, however, does not necessarily have a lesser chance of receiving votes since we have accounted for the empirical probability of each class in BT training.

The BT algorithm was applied to the logs to predict the target at two well locations (6507/7-3 and 6507/7-8) and to recover the missing part of 6507/7-4. The results are shown in Figure 6. The results demonstrate the capability of BT in predicting HC units: predicted oil-sands agree with the OWC depths, and there are no gas-sands predicted falsely in the target wells. Also, shale layers belonging to the Viking Group, Ror and Tilje Formations (Båt Group), and Not Formation, where they exist, are predicted correctly.

In the following confusion matrix (Figure 7-a), the performance of the trained BT algorithm can be investigated in more detail. The success rate values shown in this matrix are averaged over cross-validations, while one-fifth of the data were randomly selected and predicted by the rest of the data during a training session. The percentages are based on the predicted classes; a false discovery shows the percentage of a predicted class that does not belong to that class in the target log. For example, false discovery of gas sand shows that 4% of the predicted gas sands were, in fact, shale (1%), and brine sand (3%). Oil sands also were correctly classified in 98% of the predictions, with only 1% and less than 1% of the predictions belonging to brine sands and shale, respectively. The second and fourth columns show that most of the false discoveries occurred between shale and brine sands.

Since in this field we have access to the target litho-facies logs, we expanded these calculations to our predictions of the target wells. The confusion matrix in Figure 7-b shows both positive predictions vs false discoveries and the true classification vs misclassification rates in the last column and last row, respectively. A misclassification rate is the percentage of a true class that is predicted as other classes. The positive prediction rates confirm that predicted HC units by BT belong to the right classes with high positive prediction rates: 92.1% for oil-sands, and 0% for gas-sand. Note that the target wells do not have any gas-sand layers, and BT has not falsely discovered any other classes as gas sands in these wells.

Similar to the training results, the shale and brine-sand classes account for most of the false discoveries between each other (see the 2nd and 4th row on the brine-sand and shale target columns, respectively). This can be explained by the previously discussed issue of the similarities between the shale and brine sand in this data set.

However, the overall low accuracy rate in this model (63.2%) is not only due to the shale/brine-sand false discoveries. The last row of this matrix shows the misclassification rate in red (bottom) numbers. The highest misclassification rate belongs to the oil-sands indicating that 60% of actual oil-sands are misclassified as brine sand, and 9% as shale. This is mostly caused by well 6507/7-3, since taking it out of the calculations drops the misclassification of oil-sands from 60% to 25.2% (Figure 7-c). The chart in Figure 8 shows the role of each target well in the overall accuracy rate of BT. Since the errors are normalized by the number of samples at each well, the stacked bars confirm that the majority of the error in the oil-sand class is caused by well 6507/7-3.

The reason behind the distinct behavior of this well is that, in creating the litho-fluid facies logs, we relied on drilling reports to validate the fluid content, and thus a thick layer of oil-sands (~100 m) is designated for an interval with HC-shows as described in the reports. The wireline logs, however, do not show such a consistent and significant anomaly over the reservoir interval. We have determined this by using other wireline logs that were not used by our BT algorithm. For example, formation resistivity logs (RTF) available only in wells 6507/7-4, 6507/7-A-53, and 6507/7-A-17 show this discrepancy among these wells and well 6507/7-3 (Figure 9). The RTF anomaly of the Fangst reservoir in this well is not as high or as blocky as observed in other wells. We can argue that the BT prediction, in contrast to the reference target, is correctly indicating a low-quality reservoir section encountered by this well. This is also in agreement with the results of an ANN-driven PR model of this field (Keynejad et al., 2017).

Kupe litho-fluid facies logs

The cross-validated results of training BT (Figure 10) show a high positive predictive rate for HC-bearing clean sands (92%). The predicted shaly sands with HC-shows have a lower rate with 77% of the predictions being correctly classified. The internal percentages of this class (fourth column) shows that of the 23% false discoveries, 13% belong to the shale class. This was predictable when this class was defined to be an indicator of low-quality reservoir sand, with features between a clean reservoir sand and shale classes in the target log. Note also that most of the clean sand false discoveries are brine sand rather than HC shaly sand. That suggests that the properties of the HC shaly sand are farther from the clean HC sand than for the brine sand.

The predicted litho-fluid facies logs for four target wells are shown in Figure 11. To qualitatively compare the results, the available interpreted composite log and the HC shows are also included. As seen in this figure, there are other interpreted lithofacies that we have not included in our study. These classes are neither related to our goal, nor large enough to affect the algorithm's performance significantly.

The comparison suggests the following points:

- Predicted HC units are in good agreement with HC shows or interpreted fluid contacts; no HC unit is predicted below the expected reservoir zone.
- Predicted shales, especially the sealing shale overlying the Farewell reservoir, are consistent with the interpreted layers of claystone and siltstone and/or proportional to the thickness or amount.
- The HC shaly sands predicted in well Kupe South 7 are consistent with the interpreted argillaceous sandstones.

Kupe South 5 is reported to have hydrocarbons indicated over an interval of ~36 m interval, which is not reflected in BT predictions. The interpreted lithofacies column from drilling reports shows alternating shale layers comprise half of the Farewell Formation in this well, the rest of which are carbonaceous sandstones. Such discrepancies should be investigated through core sample analysis to better understand whether this is only a misclassification. The drilling documents of this well do not mark the Farewell sandstones as reservoir, and report that severe calcite cementations and formation of authigenic clay minerals have affected the reservoir quality (New Zealand Petroleum & Minerals, 2017).

CONCLUSIONS

Discovery and development wells are drilled at various stages of a hydrocarbon field's lifecycle. The main goal is to confirm a prospect or determine the reservoir size by examining lithofacies variations and fluid presence through available cored intervals. The problems can arise when a well or interval of interest is not cored, or the cores are lost due to technical issues. Interpretation of wireline logs, where available, has always been used as an alternative to direct observations to provide a continuous record of subsurface litho-fluid variations. However, it is also possible that some parameters, crucial for the interpretations, are not logged especially at the initial stages of exploration. Also, decision making based on a set of intertwined multi-dimensional attributes is not one of humans' greatest strengths. All of these factors, along with subjective human error, contribute to the uncertainty inherent in the interpretation of the "un-seen target".

Our study suggests the utilization of ML algorithms along with the expert's knowledge to gain objective insight of the subsurface properties while tuning the algorithm and resultant interpretations based on the expectations and information about the reservoir. The

main advantage of ML is its power in resolving complex, multi-criteria problems. Another advantage of this approach is that, once trained, it can be applied to wells with minimum available logs, since the target logs used in training are designed based on reliable information and a complete package of logs in the training wells. We have compared an unsupervised, and two supervised approaches and applied the most successful algorithm on two hydrocarbon fields to interpret its results.

In each field, we used different wireline logs to create litho-fluid facies logs for training wells. The facies defined in this study are not detailed stratigraphic classes; instead, the facies were placed into simplified, inclusive groups with the main aim of distinguishing HC-bearing units from the rest, in accordance with the reservoir's properties. While this approach can be of more importance in appraisal and development phases, ML applications with more class variations based on stratigraphy and/or depositional settings can be similarly designed for explorational stages. Note that misclassification is inevitable in any sort of indirect assessment of the facies, since none of them are completely distinct from the others, physically and or compositionally.

While the un-supervised approach (SOM) was helpful as a pre-processing step to establish the idea about the number of clusters and the predictive attributes without the need to know the targets, the results do not seem to be reliable enough. This approach, coupled with knowledge-driven techniques, can be of especial help in cases where there are no cored wells near the study area.

Both supervised methods could successfully predict HC units in the cross-validation training phase in both fields, with bagged tree (BT) having a higher overall and HC-related accuracy rate. We expect ANNs to be more successful in problems where the relationship

among attributes and targets are more complicated than being mapped with a system of if-then rules. Using seismic attributes for mapping seismic facies is an example in which the rules relating the features to the target, though basically understood, cannot be directly drawn at any given sample.

A more detailed study of the prediction results for trained BTs showed a very low false discovery rate for oil sands (7.9%), and no false discovery for gas sands in Heidrun Field. However, about two-thirds of the existing oil sands were misclassified as brine sands, with well 6507/7-3 as the major contributor to this error. The prediction, though deviating from the previously interpreted fluid column (hence, causing the error), is in agreement with low formation resistivity values and relatively high Poisson's ratio at this well location.

Qualitative investigations of Kupe Field results also demonstrate no false discoveries for HC-units. The HC-shaly sand class was defined for low-quality reservoir sands and was expected to be mostly misclassified either as clean HC-sands or shale. It was successfully predicted at Kupe South 7, which is a deviated well from the training well that contained this class. Thin layers of predicted different classes are in accordance with variations in the associated index logs, which can mean a more reliable resolution than the qualitative interpretations. The HC-shows reported in well Kupe South 5 that occur in a very low-porosity section of the Farewell Formation, which was not marked as reservoir, were not predicted by BT. For such reservoir conditions, predicting an HC-unit would more likely be a falsely discovered reservoir, and cause more harm in field development plans if decisions are based on misleading results.

The apparent misclassification of HC units as non-HC units in this study can be interpreted using core samples and other comprehensive field studies. If the

misclassifications are real, they can cause an underestimation of the reservoir capacity. In general, to lower any misclassification in predicting litho-fluid facies in a reservoir, the observed cases of misclassification should be interpreted and addressed during the training phase.

The advantage of utilizing ML algorithms is their power in scrutinizing large data sets with multi-dimensional, complicated, inter-related attributes. This study showed that ML algorithms can be used in classifying litho-fluid facies to predict HC-bearing units, with minimal to no risk of HC false discoveries. The proposed approach of using ML algorithms, trained and tuned based on an expert's knowledge about the reservoir, can be modified and applied on future wells in a hydrocarbon field to distinguish the pay zone.

ACKNOWLEDGEMENTS

The authors thank MathWorks for providing MATLAB[©], Simulink[©] under the academic license to the University of Arizona. We also thank CGG Services (U.S.) Inc. for providing licenses for their Geosoftware package, including Hampson-Russel Suite modeling and analysis software. We gratefully acknowledge ConocoPhillips and its partners, Statoil, Petoro, and Eni for providing the Heidrun Field data set used in this research, with s special thanks to Graham Hill for his support and assistance in this procedure. We are grateful to Dr. Amanda Hughes for providing us with the Kupe Field database, originally issued by New Zealand Petroleum & Minerals.

REFERENCES

- Al Moqbel, A., and Wang, Y. 2011, Carbonate Reservoir Characterization with Lithofacies Clustering and Porosity Prediction, *Journal of Geophysics and Engineering*, **8**. p. 592-598.
- Aleardi, M., and Ciabarri, F. 2017, Application of Different Classification Methods for Litho-Fluid Facies Prediction: A Case Study from the Offshore Nile Delta, *Journal of Geophysics and Engineering*, **14**. p. 1087-1102.
- Al-Mudhafar, W. 2017, Integrating Well Log Interpretations for Lithofacies Classification and Permeability Modeling through Advanced Machine Learning Algorithms, *J Petrol Explor Prod Technol*, Springer Berlin Heidelberg, **7**. Berlin/Heidelberg, p. 1023-1033.
- Breiman, L. 1996, Bagging Predictors, *Machine Learning*, Springer, **24**. p. 123-140.
- Breiman, L. 2001, Random Forests, *Machine Learning*, Kluwer Academic Publishers, **45**. Boston, p. 5-32.
- Chang, H., Kopaska-Merkel, D. C., Chen, H., and Durrans, S. R. 2000, Lithofacies Identification using Multiple Adaptive Resonance Theory Neural Networks and Group Decision Expert System, *Computers & Geosciences*, **26**. p. 591-601.
- Close, D., Taylor, R., and Nixon, S. 2016, A Case Study in Quantitative Interpretation Ambiguity, Lambda-Mu-Rho, and Rock-Physics Modeling in the Otway Basin, Australia, *The Leading Edge*, **35**. p. 43-50.
- Goodway, B., Chen, T., and Downtown, J. 1997, Improved AVO Fluid Detection and Lithology Discrimination using Lamé Petrophysical Parameters; “ $\lambda\rho$ ”, “ $\mu\rho$ ”, & “ λ/M Fluid Stack”, from P and S Inversions, *SEG Technical Program Expanded Abstracts*, p. 183.

Harris, N. B. 1989, Reservoir Geology of Fangst Group (Middle Jurassic), Heidrun Field, Offshore Mid-Norway, AAPG Bulletin, **73**. p. 1415-1435.

Keynejad, S., Sbar, M., and Johnson, R. 2017, Comparison of Model-Based Generalized Regression Neural Network and Prestack Inversion in Predicting Poisson's Ratio in Heidrun Field, North Sea, The Leading Edge, Society of Exploration Geophysicists, **36**. p. 938-946.

Kohonen, T. 1987, Self-Organization and Associative Memory, Springer-Verlag.

Liu, R., Zhang, B., and Wang, X. 2017, Patterns Classification in Assisting Seismic-Facies Analysis, SEG Technical Program Expanded Abstracts 2017, Society of Exploration Geophysicists, p. 2127-2131.

Martin, K. R., Baker, J. C., Hamilton, P. J., and Thrasher, G. P. 1994, Diagenesis and Reservoir Quality of Paleocene Sandstones in the Kupe South Field, Taranaki Basin, New Zealand, AAPG Bull., **78**. p. 624-643.

Morton, A., Hallsworth, C., Strogen, D., Whitham, A., and Fanning, M. 2009, Evolution of Provenance in the NE Atlantic Rift: The Early–Middle Jurassic Succession in the Heidrun Field, Halten Terrace, Offshore Mid-Norway, Mar.Pet.Geol., **26**. p. 1100-1117.

New Zealand Petroleum & Minerals. 2017, New Zealand Petroleum Exploration Data Pack, New Zealand Petroleum & Minerals.

Norwegian Petroleum Directorate. Heidrun Field. Retrieved from
<http://factpages.npd.no/factpages>

- Pang, J., and Collen, J. 1996, Well Log Analysis of the Kupe South Field, Taranaki Basin, New Zealand, SPE Asia Pacific Oil and Gas Conference, Society of Petroleum Engineers, Adelaide, Australia.
- Poulton, M. M. 2002, Neural Networks as an Intelligence Amplification Tool: A Review of Applications, *Geophysics*, **67**. p. 979-993.
- Qi, L., and Carr, T. R. 2006, Neural Network Prediction of Carbonate Lithofacies from Well Logs, Big Bow and Sand Arroyo Creek Fields, Southwest Kansas, *Computers & Geosciences*, **32**. p. 947-964.
- Ross, C. P., and Cole, D. M. 2017, A Comparison of Popular Neural Network Facies-Classification Schemes, *The Leading Edge*, Society of Exploration Geophysicists, **36**. p. 340-349.
- Rousseeuw, P. J. 1987, Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis, *Journal of Computational and Applied Mathematics*, **20**. p. 53-65.
- Torres, A., and Reverón, J. 2014, Integration of Rock Physics, Seismic Inversion, and Support Vector Machines for Reservoir Characterization in the Orinoco Oil Belt, Venezuela, *The Leading Edge*, Society of Exploration Geophysicists, **33**. p. 774-782.
- Young, K. T., and Tatham, R. H. 2007, Lambda-Mu-Rho Inversion as a Fluid and Lithology Discriminator in the Columbus Basin, Offshore Trinidad, *SEG Technical Program Expanded Abstracts*, Society of Exploration Geophysicists, **26**. p. 214.
- Zhao, T., Li, F., and Marfurt, K. 2017, Constraining Self-Organizing Map Facies Analysis with Stratigraphy: An Approach to Increase the Credibility in Automatic Seismic Facies Classification, *Interpretation*, Society of Exploration Geophysicists, **5**. p. T171.

Zhao, T., Jayaram, V., Marfurt, K. J., and Zhou, H. 2014, Lithofacies Classification in Barnett Shale using Proximal Support Vector Machines, 2014 SEG Annual Meeting, Society of Exploration Geophysicists, Denver, Colorado, USA.

LIST OF FIGURES

Figure 1. The litho-fluid classes plotted on predicting-feature coordinates (parallel coordinates) for New Zealand data: true vertical depth (TVD), gamma-ray (GR), P-impedance (Z_P), and deep resistivity (RESD). Each feature is scaled over its range to better show the separation of classes. Also note how the GR (as an example) separates some of the shales with higher GR from the other lithologies, but other shales with lower GR are similar in value with the brine sands. Most of these low-GR shales, however, can be separated from the brine sands on the TVD or Z_P coordinates.

Figure 2. Silhouette values for Heidrun Field (left) and Kupe Field (right). The values are calculated for different cluster numbers as predicted by SOM networks. The higher value indicates on the best separation between clusters, and hence the optimum number of clusters for the clustering method. Lower values, on the other hand, indicate that more similarities exist between the separated clusters.

Figure 3. A part of one of the trees in a bagged-tree (BT) analysis trained on Heidrun data. Two observations leading to two example leaves are shown in bold. Leaves for this data set are 1: gas sand, 2: shale, 3: oil sand, 4: brine sand. The specifying feature values for those observations are shown at the splitting nodes. TVD: true vertical depth, PR: Poisson's ratio, LR: lambda-rho, Phi: porosity, V_{Sh}: shale volume.

Figure 4. SOM-predicted litho-fluid facies in example wells from Kupe Field (left), and Heidrun Field (right). The dashed lines show the lowest known gas (LKG) and oil-water contact (OWC) for the two wells.

Figure 5. BT prediction of litho-fluid facies for well 6507/8-1 of Heidrun Field (left), and well Kupe South-7 of Kupe Field (right). P-impedance, gamma ray, and deep resistivity logs are shown for reference.

Figure 6. BT-predicted logs are shown above (6507/7-3 and 7-8) and the predicted section of 6507/7-4 (bottom left, zoomed in). The rest of 6507/7-4 and well 6507/8-1 were used to train the BT.

Figure 7. (a) Confusion matrix of trained and cross-validated BT results on Heidrun data with average accuracy 95.7%. On the right, the prediction results are shown for all target wells (b), and all except well 6507/7-3 (c). The cells in the dark columns in (b) and (c) include the positive prediction in green (top) numbers, and the false discovery in red (bottom) numbers. The cells in the dark rows in (b) and (c) show the correct classification in green (top) numbers, and the misclassification in red (bottom) numbers. The blue boxes in (b) and (c) show the overall accuracy and error in green (top) and red (bottom) numbers, respectively.

Figure 8. Error contributed by each target well (6507/7-3, 6507/7-4, and 6507/7-8 in sections from right to left) in BT's overall prediction error (All) at left. MC: misclassification rate, FD: rate of false discoveries, Avg: averaged error. To be comparable with the overall section, the error values of each well are normalized by the number of samples provided by that well. In this way, for example, by adding the lengths of corresponding bar segments of the displayed oil-sand MC's of all three wells, the overall oil-sand MC (the leftmost section) will be obtained.

Figure 9. (a) Crossplot of true formation resistivity (RTF, ohm-m) vs Gamma Ray (API). (b) RTF logs plotted for the same well, colored by RTF amplitude at reservoir depth

interval. The RTF curves are shown in logarithmic scale. The color scale for the curves is cropped at 110 ohm-m to better depict the variations at the reservoir interval (i.e., RTF values of 110-500 ohm-m all are in purple).

Figure 10. Confusion matrix, showing the cross-validated results of training BT algorithm on Kupe Field data, with 88.7% average accuracy.

Figure 11. BT-Predicted litho-fluid facies logs in Kupe Field (right columns), compared to the interpreted lithofacies (left columns). Red arrows for Kupe South 4 point to very thin HC sands predicted by BT.

TABLES

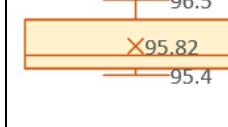
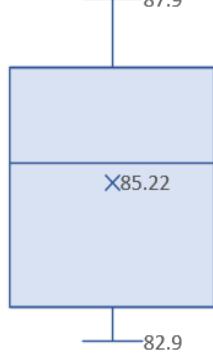
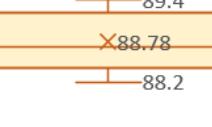
	Heidrun		Kupe	
No. of samples	1910		1986	
ML algorithm	MLFN	BT	MLFN	BT
Accuracy range				
Accuracy (%)	89.9	95.7	85.5	88.7
HC success rate (%)	95.4	97.5	86.2	89.8
Non-HC success rate (%)	87.6	94.7	84.6	88.1

Table 1. Summary of BT and MLFN training results. The boxplots in the accuracy range

show the results of seven independent training sessions. The numbers on these plots are the minimum, average and maximum, from bottom to top, and the line indicates the median. Top and bottom of the boxes are the 3rd and 1st quartiles, respectively. The other rows show the results of the algorithm with the median accuracy. The accuracy is based on the cross-validated training error, and the success rates are positive predictions for each class.

FIGURES

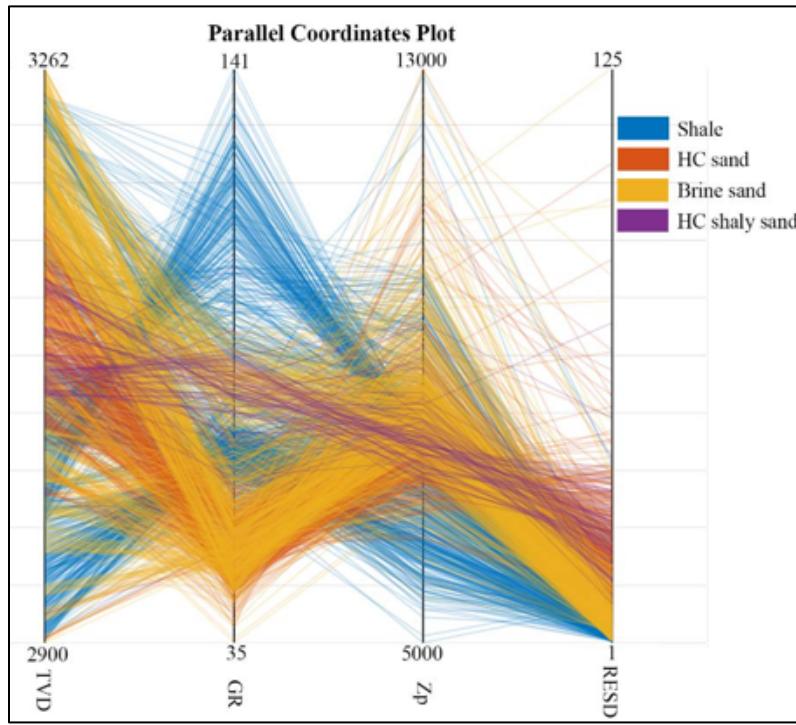


Figure 1. The litho-fluid classes plotted on predicting-feature coordinates (parallel coordinates) for New Zealand data: true vertical depth (TVD), gamma-ray (GR), P-impedance (Z_p), and deep resistivity (RESD). Each feature is scaled over its range to better show the separation of classes. Also note how the GR (as an example) separates some of the shales with higher GR from the other lithologies, but other shales with lower GR are similar in value with the brine sands. Most of these low-GR shales, however, can be separated from the brine sands on the TVD or Z_p coordinates.

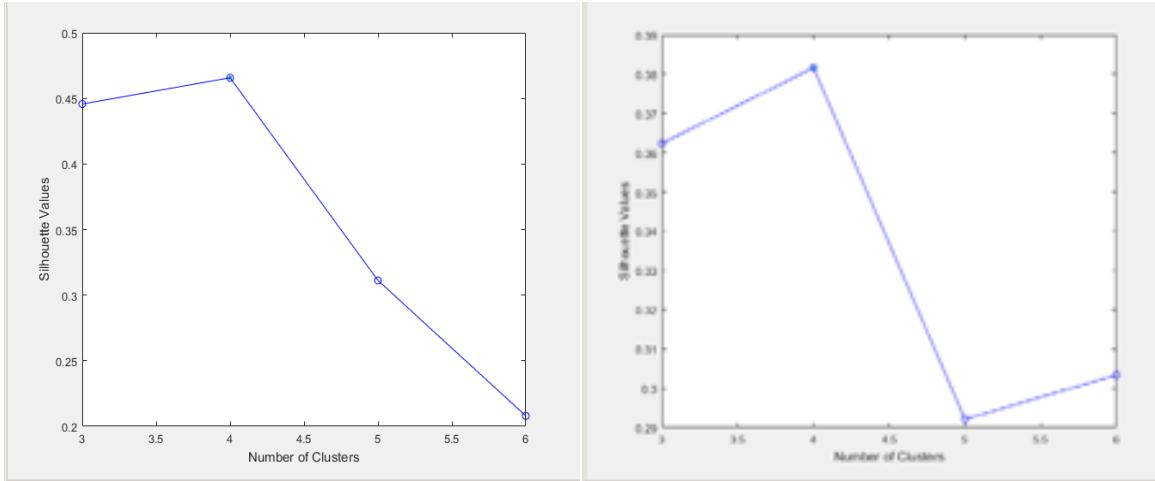


Figure 2. Silhouette values for Heidrun Field (left) and Kupe Field (right). The values are calculated for different cluster numbers as predicted by SOM networks. The higher value indicates on the best separation between clusters, and hence the optimum number of clusters for the clustering method. Lower values, on the other hand, indicate that more similarities exist between the separated clusters.

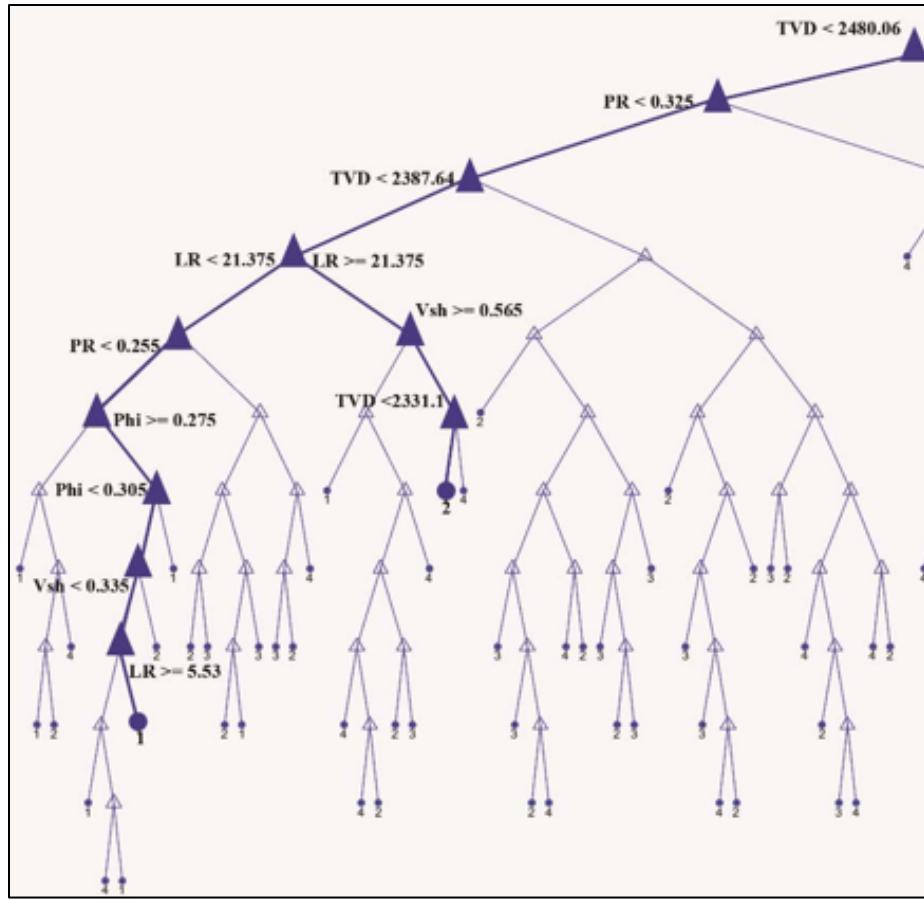


Figure 3. A part of one of the trees in a bagged-tree (BT) analysis trained on Heidrun data. Two observations leading to two example leaves are shown in bold. Leaves for this data set are 1: gas sand, 2: shale, 3: oil sand, 4: brine sand. The specifying feature values for those observations are shown at the splitting nodes. TVD: true vertical depth, PR: Poisson's ratio, LR: lambda-rho, Phi: porosity, V_{Sh}: shale volume.

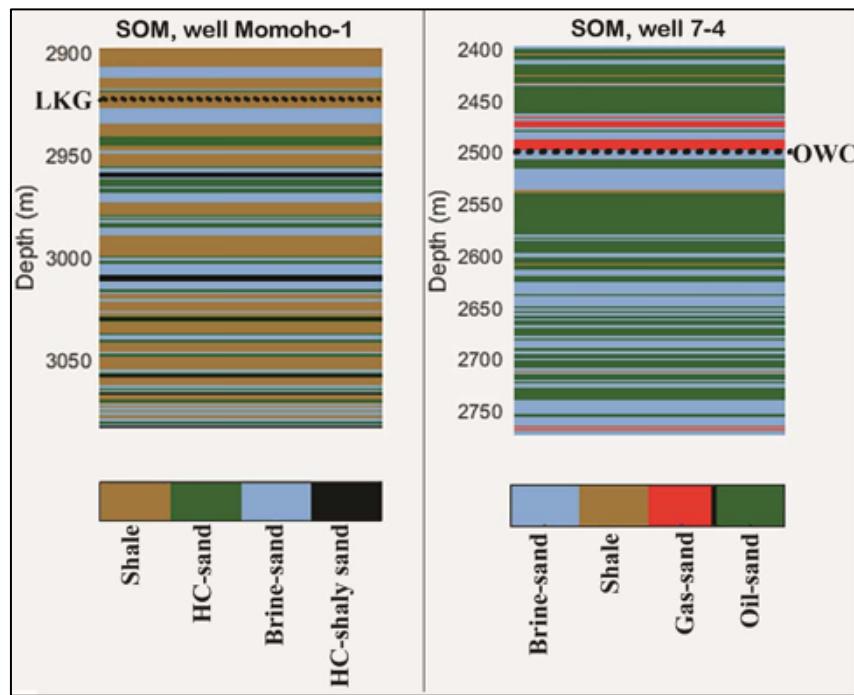


Figure 4. SOM-predicted litho-fluid facies in example wells from Kupe Field (left), and Heidrun Field (right). The dashed lines show the lowest known gas (LKG) and oil-water contact (OWC) for the two wells.

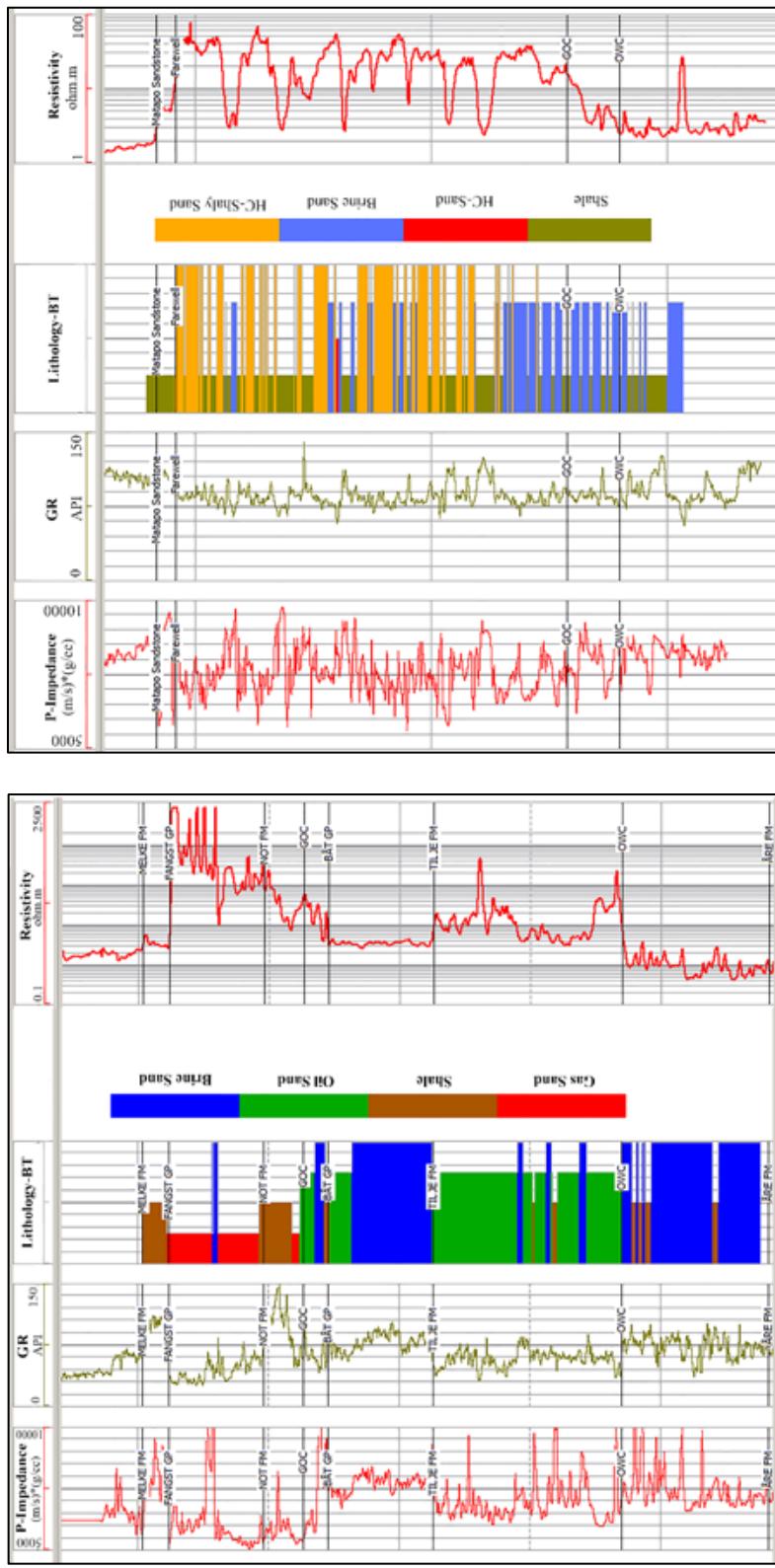


Figure 5. BT prediction of litho-fluid facies for well 6507/8-1 of Heidrun Field (left), and well Kupe South-7 of Kupe Field (right). P-impedance, gamma ray, and deep resistivity logs are shown for reference.

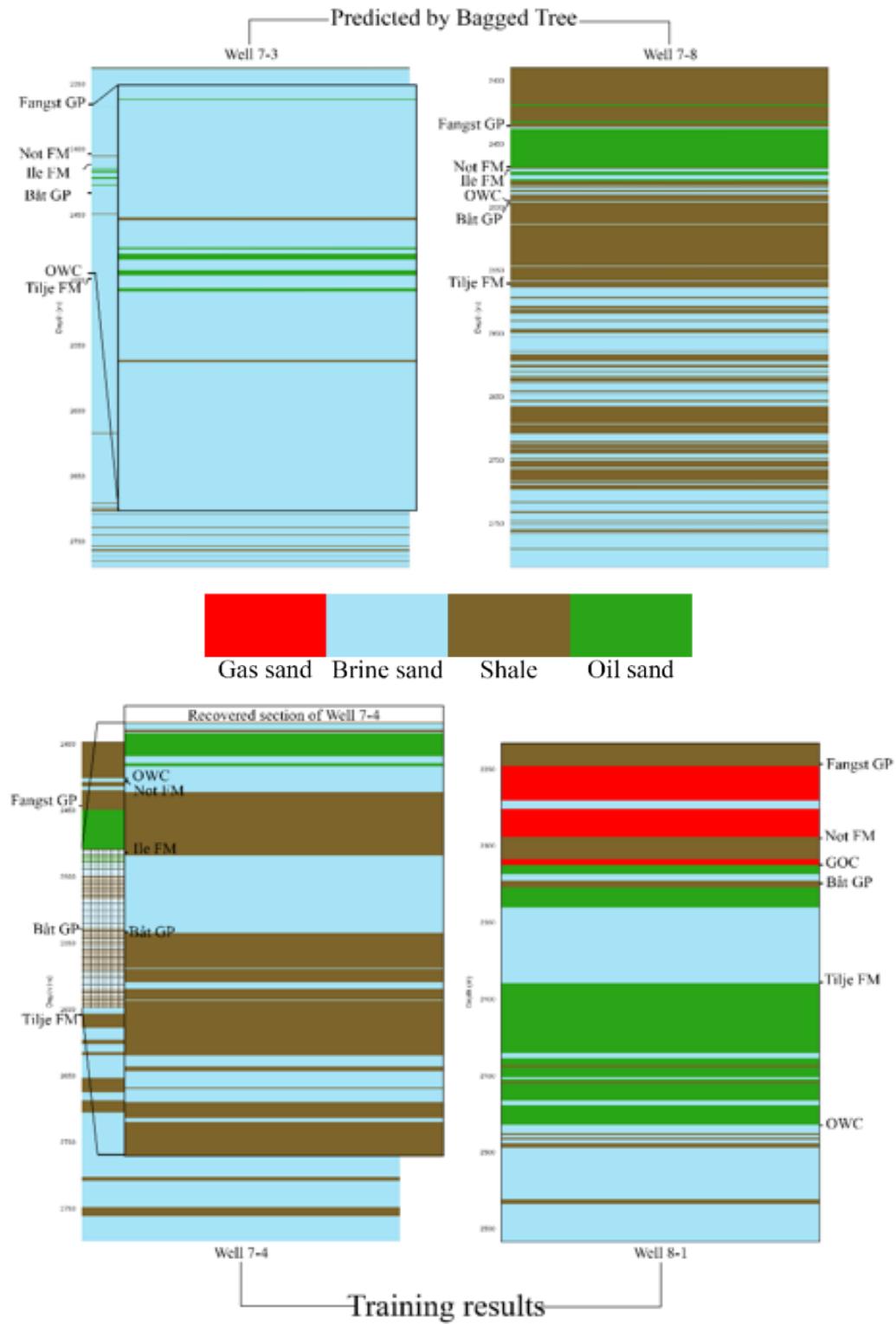


Figure 6. BT-predicted logs are shown above (6507/7-3 and 7-8) and the predicted section of 6507/7-4 (bottom left, zoomed in). The rest of 6507/7-4 and well 6507/8-1 were used to train the BT.

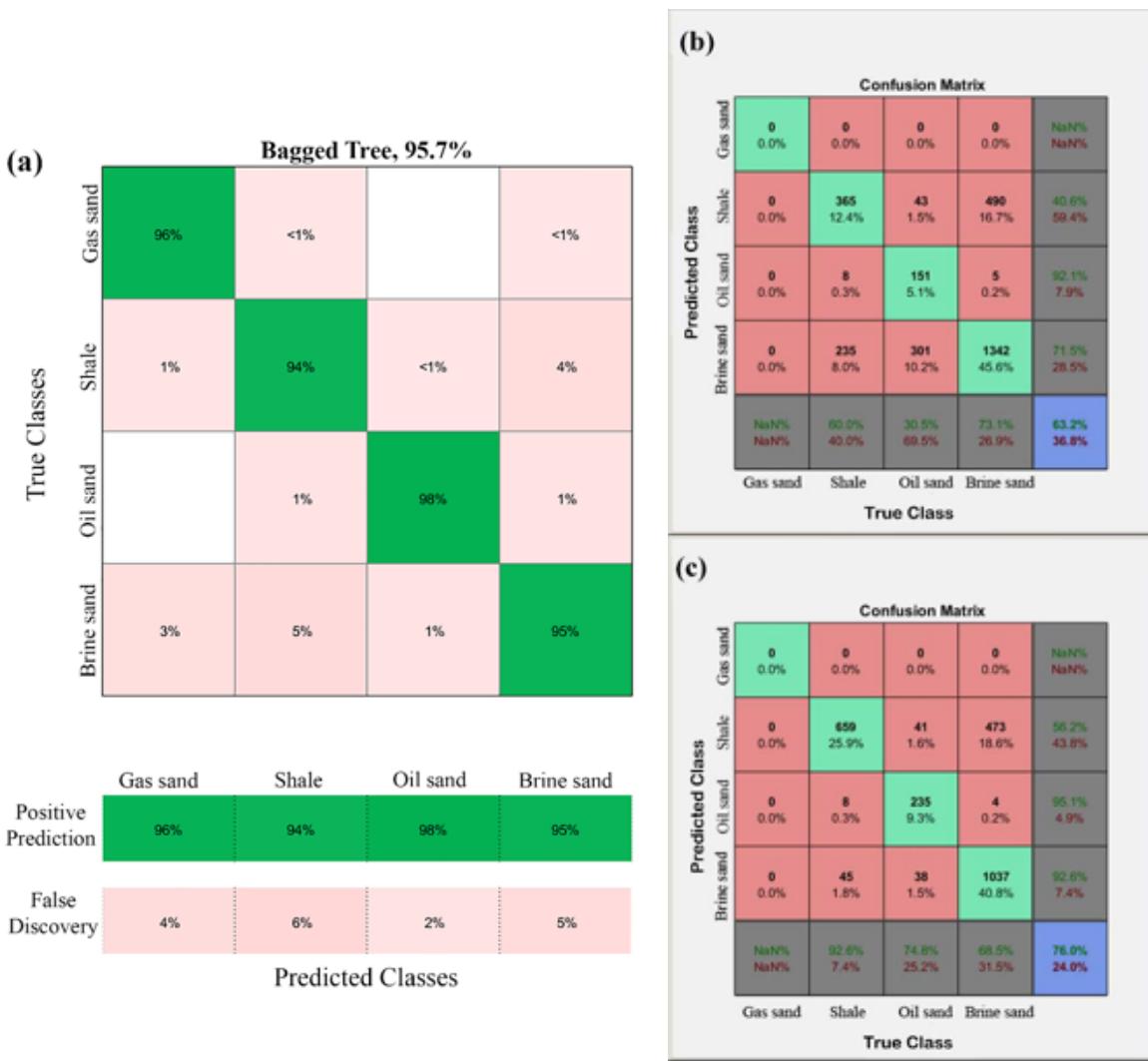


Figure 7. (a) Confusion matrix of trained and cross-validated BT results on Heidrun data with average accuracy 95.7%. On the right, the prediction results are shown for all target wells (b), and all except well 6507/7-3 (c). The cells in the dark columns in (b) and (c) include the positive prediction in green (top) numbers, and the false discovery in red (bottom) numbers. The cells in the dark rows in (b) and (c) show the correct classification in green (top) numbers, and the misclassification in red (bottom) numbers. The blue boxes in (b) and (c) show the overall accuracy and error in green (top) and red (bottom) numbers, respectively.

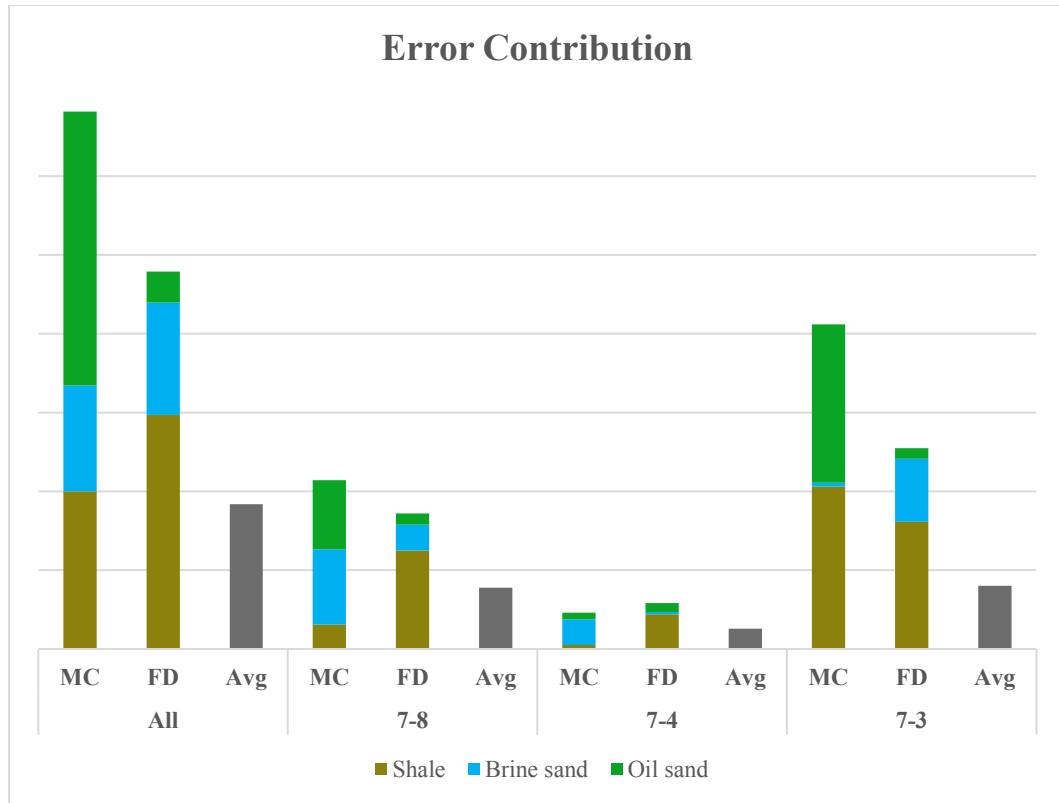


Figure 8. Error contributed by each target well (6507/7-3, 6507/7-4, and 6507/7-8) in sections from right to left) in BT's overall prediction error (All) at left. MC: misclassification rate, FD: rate of false discoveries, Avg: averaged error. To be comparable with the overall section, the error values of each well are normalized by the number of samples provided by that well. In this way, for example, by adding the lengths of corresponding bar segments of the displayed oil-sand MC's of all three wells, the overall oil-sand MC (the leftmost section) will be obtained.

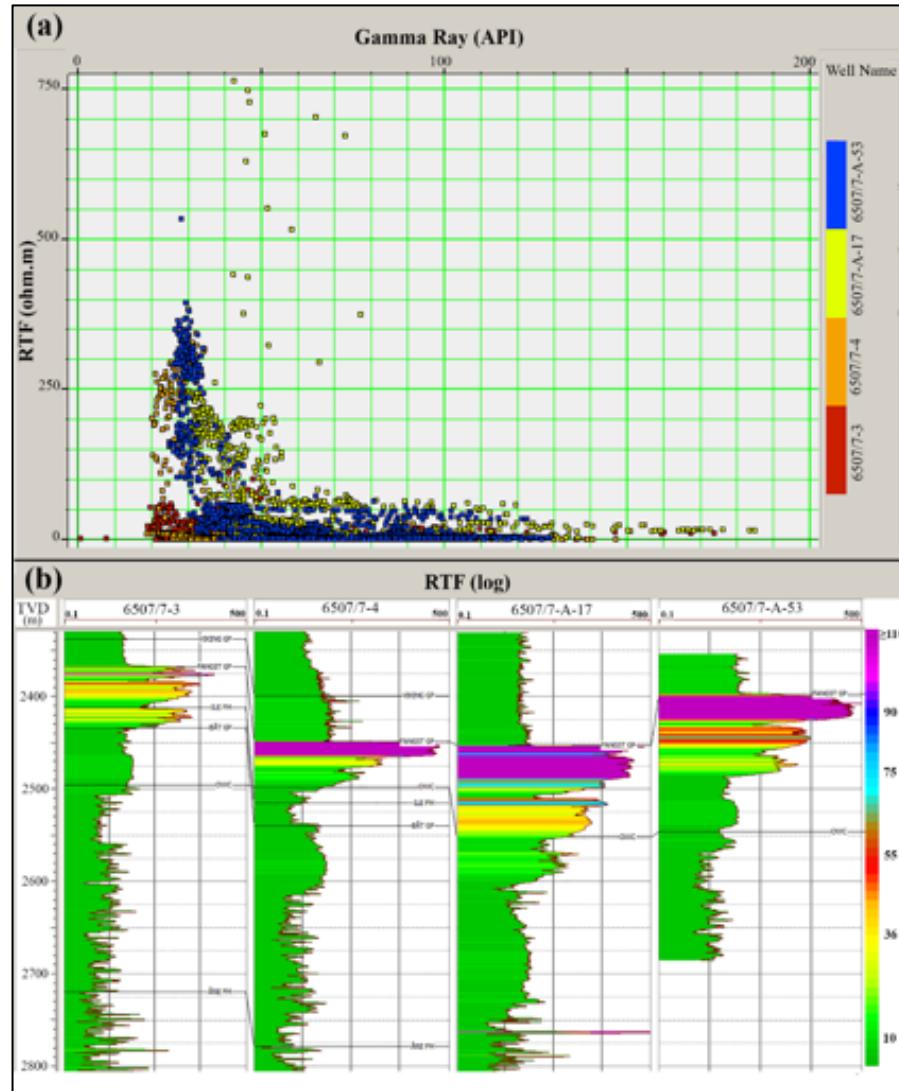


Figure 9. (a) Crossplot of true formation resistivity (RTF, ohm-m) vs Gamma Ray (API). (b) RTF logs plotted for the same well, colored by RTF amplitude at reservoir depth interval. The RTF curves are shown in logarithmic scale. The color scale for the curves is cropped at 110 ohm-m to better depict the variations at the reservoir interval (i.e., RTF values of 110-500 ohm-m all are in purple).

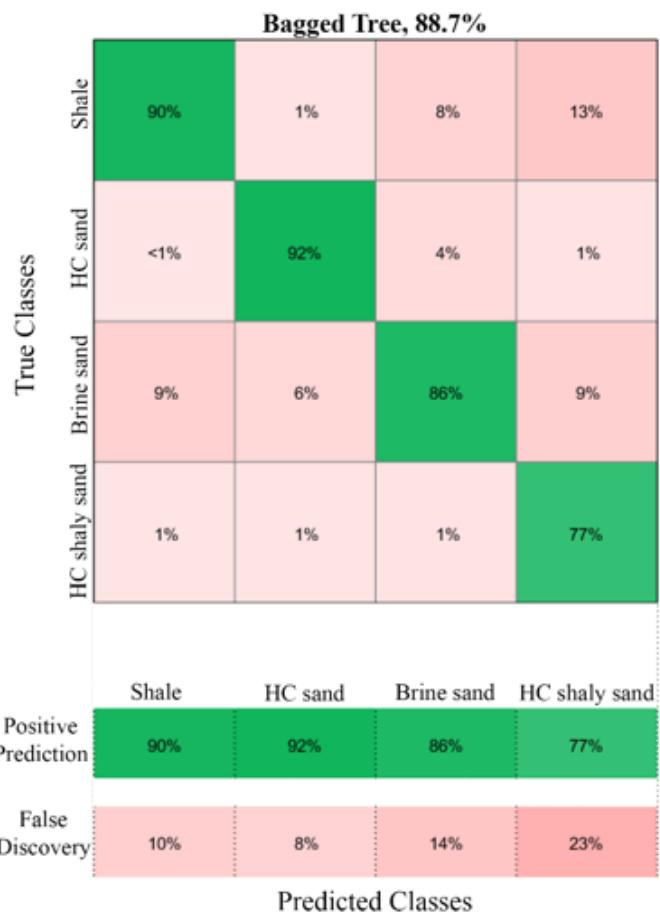


Figure 10. Confusion matrix, showing the cross-validated results of training BT algorithm on Kupe Field data, with 88.7% average accuracy.

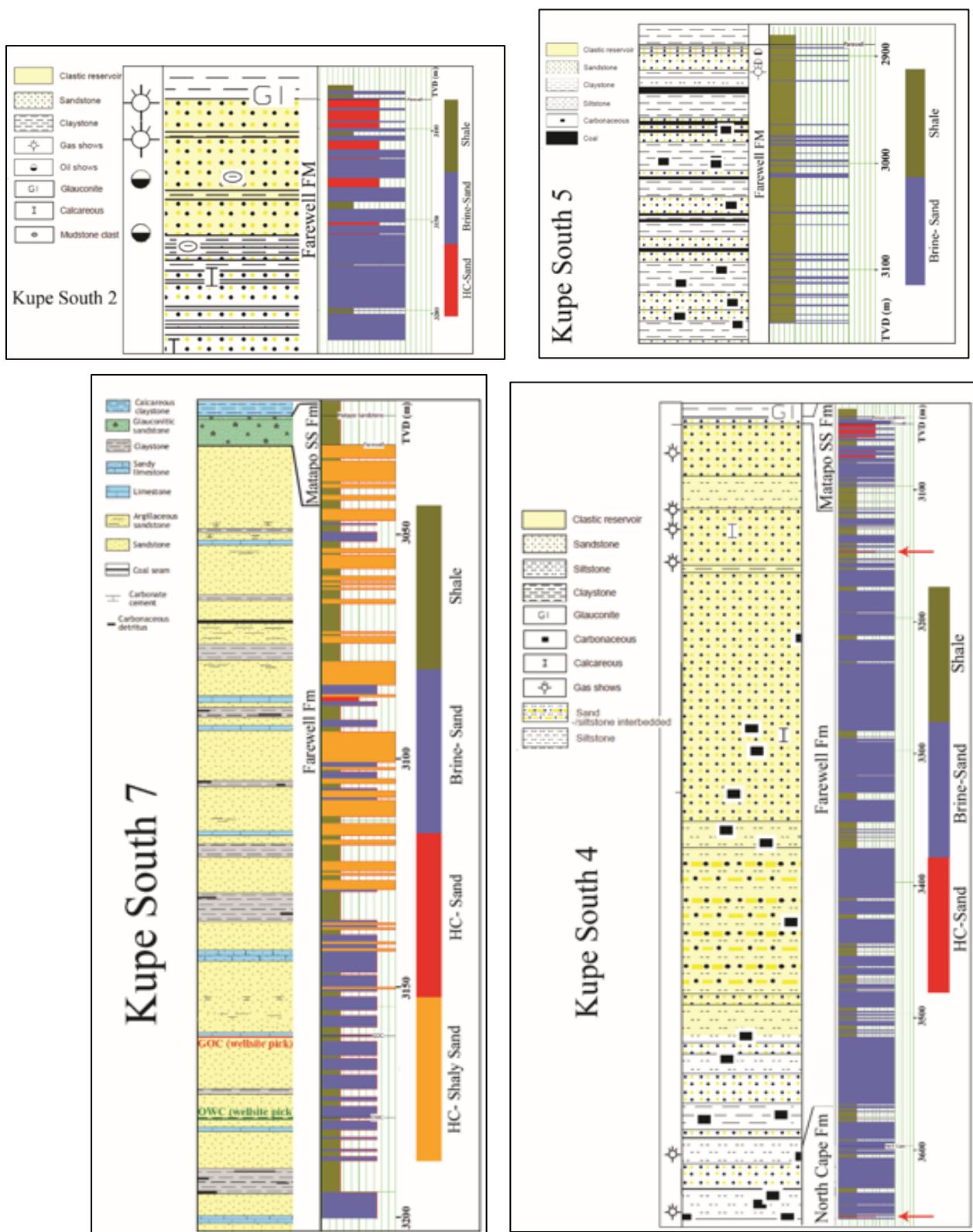


Figure 11. BT-Predicted litho-fluid facies logs in Kupe Field (right columns), compared to the interpreted lithofacies (left columns). Red arrows for Kupe South 4 point to very thin HC sands predicted by BT.

**APPENDIX C: CREATING PROBABILISTIC 3D MODELS OF LITHO-FLUID
FACIES USING MACHINE LEARNING ALGORITHMS**

Saba Keynejad¹, Marc L. Sbar¹, Roy A. Johnson¹

¹*University of Arizona, Tucson, AZ, USA*

Prepared for submission to Journal of Interpretation, 2018

ABSTRACT

Mapping facies variations is a fundamental element in the study of reservoir characteristics. From identifying a pay zone to estimating the reservoir capacity, a hydrocarbon field's development plan depends to a great extent on a reliable model of lithofacies and fluid content variations throughout the reservoir. The starting point is usually creating one-dimensional facies models based on core samples and drilling reports at each well location. Sparse well locations and the inherent heterogeneity in the reservoir properties makes it essential to incorporate the resultant 1D models into a 3D model of facies distribution that includes information about the probability of their occurrence. We propose techniques to build 3D litho-fluid facies (LFF) models that can incorporate the prediction of different lithofacies classes with regard to their potential hydrocarbon content, along with the uncertainties of the prediction. A fuzzy inference system (FIS), as an expert-oriented approach, and two machine learning (ML) algorithms, are applied to different seismic and elastic attributes to model the LFF classes within the Heidrun oil and gas field. The results, compared to the test wells, show that the ML methods could successfully predict the distribution of gas and oil sands within the field, in very good agreement with the known fluid contact intervals. The predictions of shale and brine sands varies depending on the method but are also consistent with our knowledge of this field. The comparison between the results confirms the higher reliability of ML methods. In addition, the ML methods provide a better way of investigating the uncertainty of the predictions. It signifies the advantages of implementing ML algorithms in reservoir characterization by reducing the risk of drilling un-necessary wells due to false discoveries, and by providing a tool to take into account the uncertainty level of predictions.

INTRODUCTION

Interpretation of cored samples and wireline logs is an inevitable step in the study of reservoir characteristics. The results are usually used in creating a model of facies variations either within or between wells. Depending on the approach and parameters, and the objective for which the model is designed, the model may be referred to as lithofacies, seismic facies, electrofacies, etc. The techniques vary not only by the nature, from petrophysical to geochemical and geophysical tools, but also in the way those tools are implemented; in a data-driven or knowledge-driven platform.

Among different data-driven methods, machine learning algorithms are used extensively in facies-classification studies. ML algorithms trained on stratigraphic information and petrophysical and seismic attributes can be used to map facies, for instance, as different classes of lithology, sedimentary cycle and depositional environment (Torres and Reverón, 2014; Ross and Cole, 2017; Roden et al., 2017). MLs are self-sustained predictive systems in the sense that they can extract the desired output from data after passing the training phase, without being explicitly programmed on the existing patterns and relationships within the dataset.

A well-constructed ML trained on properly chosen attributes can have a huge impact on the procedure time and classification accuracy. More importantly, the objective approach of MLs in handling complicated problems in a high-dimensional, attribute space, makes them a desirable option in modelling subsurface properties. This becomes even more important since various sources of heterogeneity are usually the biggest problem in mapping facies by various methods ranging from simple interpolations to more complicated geostatistical techniques. Because these simpler techniques are only

constrained by a limited number of parameters throughout the reservoir, whereas ML considers the selected attributes at each seed point to make a decision based on the previously learned relationships among them.

Of course, the uncertainty in prediction never can be completely removed. Natural changes in facies usually are gradational without distinct boundaries and, thus classification requires simplification in defining target facies for the algorithm. Also, the assumptions and simplifications in relating elastic (seismic) to petrophysical (well-log) parameters, and scale issues while integrating both, are some of the other sources for the remaining uncertainty. To assess the uncertainty in the results and to better reflect the gradual changes in facies, we have investigated results of two data-driven approaches in a probabilistic scheme. We used these techniques to classify litho-fluid facies (LFFs) to model hydrocarbon- (HC-) bearing units in the form of a 3D model of the reservoir. To test the efficiency of ML algorithms, we also included a knowledge-driven method in our study that classifies LFFs based on the analyst's assessment of the attributes and assumptions about the reservoir.

For the knowledge-driven approach, we designed a fuzzy inference system (FIS) of Mamdani type (Mamdani and Assilian, 1975) to map attributes to LFF classes. This is because the results of FIS, as explained in the following methodology section, are a direct reflection of the expert's knowledge, and thus provide a reasonable basis to evaluate ML results. FISs are decision-making systems that are developed based on Zadeh's fuzzy set theory (Zadeh, 1965). Whether or not the elements belong to a fuzzy set is determined by a range of values called membership degree, contrary to the sharp boundaries of inclusion and exclusion in classical set theory. FIS-based classification rules are a collection of

linguistic statements in an if-then format that set the path for the FIS to make a decision. For instance: if (input1 with membership degree1) aggregated (by operator1) with (input2 with membership degree2) then: (output1 with membership degree3) weighted by the confidence/importance level x. Fuzzy logic is used in ambiguous conditions, by handling the uncertainty based on the analyst's knowledge fed into the system in the form of the membership degrees and the logical framework. We have used this ambiguous state in the FIS results as a tool to reflect the uncertainty of predicting LFF volumes by a knowledge-driven method.

Two ML classifiers used in this study are an artificial neural network (ANN) and an ensembled decision tree. A decision tree is similar to the FIS in the way that it connects the input attributes to outputs through a set of conditional rules. Unlike FIS, these rules and conditions are set by the algorithm after learning them during training phase. Single decision trees tend to overfit the data as they usually grow too deep and thus learn very irregular relations within the dataset. To increase stability and reduce the variance in the results, we have used a bagging (**bootstrap aggregating**) method as an ensemble tool, integrating the results from the trees grown by a random forest approach (Breiman, 1996; 2001). We trained the BT both to classify the LFFs, and to estimate the probabilities of various classes as a reflection of the uncertainty in its classification.

Supervised ANNs take in the observations (attributes) and relate them to the associated targets via non-linear activation functions of their hidden layer(s). ANNs are powerful tools in predicting the outcome of complicated systems that have non-linear relationships between the attributes and outcomes. During the learning phase, the network compares its predictions to the actual targets and modifies the connection weights between neurons to

minimize the prediction error. Probabilistic neural networks (PNN) are one type of the ANNs used in pattern recognition problems. These networks use a radial basis function that estimates the probability of an input vector belonging to an output class, and then determine the output based on these probabilities.

Using the probabilistic results, we were able to compare the uncertainty of the resultant LFFs of two MLs and the continuous output of the FIS classification. The MLs successfully classified the LFFs within the fluid contact intervals, with BT being in better agreement with the lateral distribution of HC-bearing units. Both ML methods drastically outperformed FIS in predicting LFF classes.

The key distinction is in who (as in analyst vs algorithm) has set the rules, what is their computational power, and how reliable their predictions would be if we go, for instance, from one BT to another versus from one analyst to another (objectivity). The performance of the ML algorithms primarily depends on the input data. If the data and the algorithm are properly chosen and trained, the results would vary in a highly reliable range of predictions through different runs (Keynejad et al., 2018). For the knowledge-driven approaches on the other hand, although it is very likely that, for example, a skilled petrophysicist can set a better FIS framework and obtain more successful results than this study, it is also very likely that their results would significantly differ from another skilled petrophysicist.

METHODOLOGY

The data sets available for this study included prestack and poststack seismic amplitude volumes, and four appraisal and three producing wells from Heidrun Field, North Sea. The appraisal wells were used for training, while the deviated producing wells were assigned for testing the results. The drilling information about formation tops and

hydrocarbon shows by Norwegian Petroleum Directorate (n.d.) and wireline logs were utilized to define LFF logs for each well. The classification algorithms were applied on the full-stack seismic volume, with the focus on the Fangst Group and the Båt Group (down to the Åre Formation top) as the pay zone. According to the reports and our interpretations, four LFF classes were defined: Gas sand, shale, oil sand, and brine sand (Keynejad et al., 2018).

For training MLs and designing the FIS, a set of in-situ and calculated logs were selected that were, through observation or at least in theory, related to lithology and/or fluid variations. These attributes include: density, lambda-rho ($\lambda\rho$, here referred to as LR), mu-rho ($\mu\rho$, here referred to as MR), porosity, apparent polarity, Poisson's ratio (PR), and P-impedance (Z_P). LR and MR were calculated using P- and S-impedance logs, and are known to be fluid- and matrix-sensitive features respectively (Goodway et al., 1997). The two-way-time (TWT) values of the logs are also included in the attributes as a constraining factor for fluid intervals.

For simulating the techniques on the seismic volume, each attribute needed to be either extracted as an internal attribute (TWT and apparent polarity) or calculated as an external attribute using wireline logs. The porosity volume is a low-frequency initial model created from porosity logs and constrained by the rms velocity model and interpreted structural horizons across the seismic volume. The PR volume is predicted by a generalized regression neural network using internal seismic attributes and initial models (Keynejad et al., 2017). Z_P , Z_S , and density volumes were calculated by a simultaneous prestack inversion technique based on the initial models of these parameters as described by

Keynejad et al. (2017). These impedance volumes were then used to calculate 3D models of the LR and MR attributes.

The abovementioned eight attributes, or a subset of them, was used to classify LFFs, as will be described in the following sections for each method. The general difference here is that, for FIS, we provided the attribute subset based on the observable relations between each attribute and the LFF classes, whereas for MLs the decision was made based on some performance measures for each method. For example, in addition to the lack of a clear separation in the densities of different LFF classes in our case, the interpreter may as well exclude it from the attributes in FIS since, usually, an inversion-derived density is not as well determined as the inversion-derived Z_P . But density is used by MLs since including it improved the prediction accuracy of these algorithms.

FIS

The process of mapping the input attributes to the output in a Mamdani-type FIS consists of these steps:

1. Designing different scenarios for each output based on the input attributes. This is the basis for the following steps.
2. Fuzzifying the input values using membership functions. That is, assigning degrees of membership to the input attributes as needed by each scenario.
3. Fuzzifying the outputs using membership functions.
4. Assigning if-then rules to each scenario.
 - Within these rules, each output is assigned to a set of inputs that are connected to each other by an operator.

- Each rule can be weighted by a value of 0-1, based on the rule's importance and/or the expert's confidence in that rule.

5. Aggregating the weighted rules to get the fuzzified output.

In fuzzy classification, it is common to defuzzify the output into discrete values as class labels. In this study, however, we decided to keep the outputs in their continuous form to reflect the uncertainty in the results.

For FIS, to be able to define the aggregation rules, we are limited to those attributes whose relationships to the target and each other can be “observed”. By plotting the probability density function (PDF) of each attribute for each class in the training wells, five attributes were selected based on the observed and interpreted relations among attributes and targets: travel time, LR, polarity, PR, and Z_p . The attribute-class membership functions were then defined to set a degree by which a range in an attribute belongs to a class. For example, PR_G , PR_{Sh} , RR_O , and PR_B are defined as PR membership functions for gas-sand, shale, oil-sand, and brine-sand classes, respectively.

As can be seen in Figure 1, some attribute-classes are not defined since the class was not separable from the rest by that attribute. Nine rules were defined based on these attribute-classes for different scenarios (Figure 1). In the graphic section of Figure 1, the attribute values for an example point are marked by vertical red lines through the input membership functions in each rule. If the line intersects the function graph of an attribute-class in a rule (i.e., if the example value of an attribute has satisfied the condition defined for it in a rule), the function returns a membership degree for that attribute value (yellow-filled sections). This degree can be interpreted as the likelihood or “probability” of that attribute value belonging to a certain class. This membership degree will then be

aggregated by other (if any) returned degrees in the same rule, multiplied by the weight, to set the output-class membership degree of the rule. At each point, attribute values go through each rule as explained, and the final output is calculated by aggregating all returned output values by different rules. The decision surface in Figure 1 (bottom right) schematically shows how attribute-classes (in LR-PR plane) are integrated into output classes (z-axis) based on the defined rules. Note that in our rules, most decision surfaces use more than two attributes, and as a result cannot be readily depicted in graphical form.

The FIS was designed after testing and refining the rules multiple times to improve the match with training wells. After getting satisfactory results, it was applied to a database of the 3D models of the same five attributes. The simulation time was very slow (more than 5 hours) in comparison to the ML methods (~2 minutes).

BT

A BT algorithm with a random-forest approach was ultimately selected due to its higher prediction accuracy after training different ML classifiers (excluding ANNs). BT, as an ensemble of decision trees, also has the advantage of overcoming the overfitting problem inherent in single decision trees (Breiman, 1996). The rules of a classifying decision tree are the branches of that tree that end at different leaves or class labels. At each splitting node on a branch, one of the attributes is divided into two or more sections, usually as a split in its range if it is a numerical attribute (e.g. $TWT > 2200$ ms, and $TWT \leq 2200$ ms). A set of the splitting nodes on a branch that end at a particular leaf is called an observation. Figure 2 shows a small part of a sample tree with splitting nodes (attributes) on branches ending at different leaves.

The training procedure for BT can be summarized in these steps:

1. Create replicates of the learning set by sampling the training data set randomly with replacement (aka, bootstrap resampling).
2. Grow a tree for each of these replicates.
3. In each tree, form splitting nodes by randomly sampling, but without replacement, from the pool of all features and conditions (e.g., splitting a branch into $TWT > 2200$ ms, and $TWT \leq 2200$ ms).
4. Repeat step 3 in the tree until grown sufficiently large.
5. Repeat steps 2-4 for a number of trees, large enough for the problem in hand.
6. The number of an outcome (e.g., class1) for a specific observation is counted over all the trees. The majority vote of the resultant class by that observation determines the class assigned to it.

There are different ways and “rules of thumb” suggested for determining the number of trees and their sizes (steps 4 and 5), such as cross-validation or “out-of-bag” error, and by trial and error (Liaw and Wiener, 2002). The combination of voting over a large number of trees and random sampling of features results in reduction of variance without increasing the bias (Breiman, 2001). The output can also be in the form of the probability that an observation can lead to a specific class, calculated as the fraction of those observations that have resulted in that class.

After testing with different parameters to determine an optimal operator, a BT with thirty learners (trees) was trained on data with a 5-fold cross-validation parameter to measure its accuracy. This allows the algorithm to cross validate its results by making prediction for one of its randomly divided sections (folds) of the data against the rest, one-fold at a time. Since training BT was fast and reducing the number of attributes did not

increase the cross-validating accuracy (i.e. more attributes did not cause overfitting), we used all eight attributes. The trained BT was then applied to the data base of the 3D model that included all eight attributes.

In addition to predicting the class to which an observation belongs, BT is also able to learn the probability of these predictions. For instance, the probability of an observed point belonging to class 1, equals the fraction of all the similar observations that have ended at class 1, over all trees. This way, we have extracted the probability of each point belonging to one of the classes in four 3D probability volumes.

PNN

PNN compares each input to the training vectors to determine how “close” the input is to each of these vectors. Since training vectors include an assigned output class, the closeness parameters can be summed up to calculate the probability of the input belonging to each class. The competing transfer function at the PNN’s last layer then classifies the input into output classes using this probability information. The output of the PNN is in a discrete form to represent each class, but we have extracted the calculated probability of each class for this study.

Since overfitting can be an issue in ANNs, we first ran a stepwise-regression analysis to choose a subset of available attributes and allowed it to transform the attributes if needed to lower the prediction error. In its first step, this method chooses an attribute that can predict the target with the least prediction error. In following steps, it adds another attribute to the previously chosen one(s) in a way that the selected group has the lowest prediction error (Hampson et al., 2001b). The PNN in this study was trained on the selected

transformed attributes extracted from the training wells, and then was applied to the database including 3D models of these attributes (Table *I*).

Since the basis of the PNN algorithm is to predict the outcome classes according to the previously learned probabilities, a by-product of applying PNN is the probability of each class occurring at each point. These probabilities have been extracted in the form of 3D volumes for each class.

As discussed above, each ML algorithm yields four probability volumes, one for each of the LFF classes. However, to simplify the visualization of the results and better compare the results to the FIS, a combination of the probabilities of the classes is created. To do so, we have compared the probability of the classes at each point to pick the one with the highest probability. The output at each point is the probability of the winner class, transformed to one of the unit-length sections of 0–1 for gas sand, 1–2 for shale, 2–3 for oil sand, and 3–4 for brine sand.

DISCUSSION AND RESULTS

The results of FIS, unlike the ML-driven results, are reflections of the decisions and interpretations made by the analyst. The inference system, including its rules of aggregating attributes and fuzzy membership degrees for each input and output, are designed and dictated by the interpreters, based on his or her observations of, and knowledge about, the field. For this reason, the FIS results cannot be considered as new “predictions”, but rather as the analyst’s point of view of the field. Figure 3 shows the result of FIS on an arbitrary line going through four training wells. Note that the order of the output classes is arbitrary, meaning that the actual litho-fluid facies do not necessarily merge into each other in the same order as the output classes. For instance, shale is mostly

between gas sand and brine sand when considering apparent polarity, whereas for PR, brine sand is between the other two. This becomes more complicated and irregular in a 5-dimensional attribute space, and nearly impossible to depict each facies' gradual change into one another through output fuzzy membership degrees.

For consistency, the color scale for classes are based on the designated ranges of FIS output membership functions. The uncertainty can then be interpretatively assessed as the classes gradually change from one to another. A preliminary investigation of FIS results indicates an overestimation of oil sands, and an underestimation of shales. The former can be caused by the Zs volume (and consequently the LR volume) since the oil sands below the OWC, especially in the middle zone of the figure, is similar to the anomalous zones in those volumes (inset in Figure 3). The misclassification of shale as brine sand on the other hand, was predictable due to the similarities of two classes in their predicting features. These similarities make it difficult to observe and set deterministic FIS-rules for the system to distinguish between the two classes.

Figures 4 and 5 show PNN and BT classification results for the same line. Each of these methods considers the probability of each class first, and then use different approaches to turn these probabilities into class labels. The probability of HC units on the inline through well 8-1 are shown as an example in the insets of each figure. These probability volumes can be more useful than the final classified model. Note that each of these methods assigns the ultimate class label based on the class with the highest probability, which might not be 50% in cases with more than two output classes. But each LFF probability volume can be independently assessed, for example, to estimate the reservoir capacity within a specific confidence interval.

However, to simplify the comparison and visualization of the results, a combination of the probabilities of the classes is provided for each ML algorithm. Figure 6 shows a second arbitrary line going through four training and three test wells for the FIS and ML probabilities. We have used the opacity of colors to reflect the level of uncertainty in the predicted classes, with increasing color opacity corresponding to greater probability. By placing the probability of the winner class at each point in one of the unit-length ranges assigned to that class, each of these ranges represents the probability of that class independent from the adjacent one. This is different than the way the FIS classes are arranged by their output membership functions.

Both methods have successfully predicted HC units, in very good agreement with the fluid contact depths, according to the test wells. The predictions made at the control points, such as the producing wells, can be quantitatively evaluated where the necessary data is available. The oil-sand class and the highly probable zones of oil sands are limited to the zone above the OWC in the BT and PNN results. Also, the oil sands in Tilje Formation are correctly predicted as a separate layer in well 8-1 by PNN and BT.

Note that gas sand has not been falsely discovered in other wells. However, the PNN, results show a few points of falsely discovered gas sands below the GOC. BT, which has used time as an attribute in predicting LFF, is consistent with the fluid-contact depths, both for gas sands and oil sands. This can be seen more clearly in the more or less straight line of the OWC depicted by BT oil sands, both in classes and in combined probabilities. Both methods, in addition to successfully predicting the gas sands in well 8-1, have predicted more of this LFF class in the vicinity of this well (Figures 4, 5, and 6).

However, those predicted by PNN cover a larger area, with a higher probability. The only gas shows in FIS analysis happens at the gas sands of well 8-1.

There are two main differences in the results of these algorithms:

1) the patchy form of oil sands predicted by PNN vs the continuous shape predicted by BT. Since both methods have correctly predicted the oil sands within the wells, it is not very clear whether the patchy pattern of PNN is a result of the braided fluvial deposition of Fangst sandstones, or that a more continuous oil-sand layer should be predicted away from our control points as in BT.

2) the predicted shales in the Båt Group are significantly different in the two methods. BT has either predicted these shale layers as they are in the training wells (e.g. 7-4) or has predicted them as brine sand (e.g., 7-8), whereas PNN has predicted them as a thick and almost continuous layer of shale.

The prediction made by BT is closer to the target LFF logs it has been trained on, but it does not necessarily indicate that the predictions made by PNN are not correct. Note that the upper formation of the Båt Group, the Ror Formations, is dominantly mudstone with interbedded sequences of sand and silt, whereas the Tilje Formation is mostly sandstone, with high clay content in some intervals, with interbeds of shale and siltstone. Thus, the PNN might have been able to pick the associated variations in attributes to predict shale correctly in spite of thin layers of shale being defined for target LFFs. This characteristic of ANNs, in generalizing the learned complicated relationships among attributes and applying this “knowledge” away from the training points, can be one of their strengths in highly ambiguous conditions such as in subsurface mapping problems (Keynejad et al.,

2017). BT in comparison, uses the ranges of attributes directly, in spite of its detailed and thorough investigations of features to set the decision-making platform.

Figure 7 shows the FIS classes in 3D with well 8-1 and the test wells overlain, along with the time slices picked for the Fangst and Båt Groups (with all wells overlain). The overestimation of oil sands by FIS can be better seen in this figure both down dip from the A-17 and 7-4 wells (3D FIS classes, and Fangst time slice), and in the anomalous zone at the middle of the Båt Group (Båt time slice).

Figure 8 shows the combined probabilities of LFF classes by PNN and BT in 3D with well 8-1 and the test wells overlain. Figures 9 and 10 are time slices showing the LFF classes (a), and probabilities for the Fangst Group (b) and Båt Group (c) with all wells overlain. Notice that an LFF class with a probability lower than 50% may become the final classification result (insets (a) and (b) of Figures 9 and 10). This is an essential factor in reservoir estimations, when a range of probabilities for a class can be more important than a discrete class prediction. In 3D models and time slices, a distinct shale layer can be seen in the Båt time slice predicted by PNN compared to mostly brine sands with scattered shows of shale predicted by BT. On the other hand, BT has predicted a more clearly separated layer of oil sand in the Tilje Formation (Figure 8, lower figures) compared to PNN. Also, note that the lateral variation of oil sands predicted by BT is consistent with the known OWC boundary as implied from seismic amplitudes (Figures 9 and 10).

The distribution and probability of the predicted gas sands can be compared between the two methods in the time slices and 3D models. The predicted gas sands are compartmentalized by faults, indicating the migration of gas along this fault system. Figure 11 shows the distribution of HC classes in relation to the faults on an inline going through

well A-35. Note that the gas sands connected to the observation at well 8-1 are on the far right of this line. Noting that the gas sand has not been falsely identified in other known areas (i.e., the training and test wells), and the fact that their distribution patterns are similar in the two methods and consistent with the structural boundaries, suggests that the predictions made by these methods seem to be far from mere misclassification or false discovery. However, to be confirmed as true or productive gas sands, more data is needed.

CONCLUSION

To create a 3D model of litho-fluid facies (LFF) within the Heidrun Field and to assess the level of uncertainty in classification, we used a fuzzy inference system (FIS) and two ML algorithms. The objective of this study is not to predict the best LFF model by ML techniques, especially since we do not possess the data necessary for a deterministic validation where the results do not match each other or our expectations. Rather, we intend to introduce the ML techniques and their versatility in predicting such a model, when trained and tested by the right collection of data types, from the vast variety of in-house data often available to the oil and gas industry.

FIS, which follows the rules determined and set by the interpreter, showed gas sands only at well 8-1 without false discoveries of this class in other observed cases (test wells). The oil sands shown by FIS, either with high or low degrees of class membership, expanded beyond the known OWC in depth in some parts in the field. In most cases FIS did not correctly show shale layers around or away from the shales at the well locations.

The probabilistic neural network (PNN) and the bagged tree (BT) ML methods successfully predicted the variations of LFF, especially for HC units. Note that, due to different attributes used by each ML algorithm, we cannot objectively determine which

method performed better in this study. Also, given the data available to us, we can only evaluate and compare their performances based on the known information about the field and our interpretations. The following observations and interpretations are made for the ML results:

Gas sands:

- The gas sands were not falsely discovered in other wells.
- Both methods predicted this class in certain parts of the field, away from control points, with similar form and lateral dimension. However, the gas sands predicted by PNN cover a larger area with a higher probability.
- The predicted gas sands can be attributed to the interpreted structural factors and are potentially significant enough to be considered for further evaluations using available data in other studies.

Oil sands:

- Both methods predicted oil sands within the interval down to the known OWC in training and test wells.
- The oil-bearing Tilje Formation observed in well 8-1 was predicted in both cases, but within a more distinct boundary in the BT results.
- The lateral variation of the BT-predicted oil sands closely resembles the OWC boundary interpreted based on seismic amplitudes. This lateral extension and the continuous pattern made by BT (vs the patchy form of PNN oil sands), if not validated by other information, may be due to BT using TWT as an attribute.

Shale:

- PNN predicted a thick shale layer at the top, interpreted to be the dominantly mudstone of the Ror Formation.
- BT on the other hand, has either predicted the top Båt Group as thin shale layers in small patches around the observed shales at the wells, or as brine sands.
- If the existence of the shale layer at this part of Båt Group can be confirmed, the prediction made by PNN shows the power of ANNs in generalizing the rules, away from the control points.
- If the PNN prediction is valid, the reason behind BT's misclassification of this shale formation as brine sand can be explained by noticing the similarities between the predicting features of the two LFF classes.

If the attributes are highly reliable and the FIS framework is “properly” chosen and set by the expert, the FIS can outperform LFF predictions made by some other knowledge-driven methods such as interpolating between wells or simple geostatistical approaches. This is because FIS can use different attributes that were “directly” extracted from each point. For this reason, such an inference system can be a suitable option, though a considerably slow one, for a preliminary study of the reservoir, and to qualitatively assess the relationship among the attributes and target facies.

By comparing the BT and PNN results in this study, we interpret the former to be more adherent towards the known parameters set by the interpreter such as the OWC, and the target LFFs, while predictions from the PNN can be different even around the wells it has been trained on (e.g., the shale layers). Due to the lack of sufficient control data, and because of the complicated and “hidden” nature of ANNs, it is difficult to interpret whether its anomalous predictions away from the control points are misclassifications, or due to it

correctly considering those relationships among attributes that are valid and relevant to the target, yet un-observable for us.

For this reason, BT can be a more suitable method in cases where the interpreter is confident enough in the observations (the attributes, and the created target LFFs). In such a scenario, predictions by BT far from the control points, such as the gas sands in this study, become more reliable since we know its predictions are more abiding by the “seen” examples in comparison with the PNN results. For the same reason, the PNN results can be more desirable in earlier stages of field development planning when the main goal is to delineate potential areas for more detailed explorational studies.

Implementing ML algorithms can significantly shorten the calculation time, making it possible to quickly and easily refine their parameters and even change the inputs if necessary. By incorporating the probability of each class, the uncertainty of the predictions can be assessed qualitatively and quantitatively. These probabilities can be utilized in estimation of reservoir capacity within a certain level of confidence and to reduce the risk of false discoveries in well planning studies.

The results of implementing a knowledge-driven approach, such as FIS, in a multi-attribute environment, can be as good as the interpreter’s level of skills and knowledge, which can be remarkably high. But more importantly, it only will be as good as a human being’s ability to make correct decisions in a highly complex multi-dimensional space of intertwined attributes.

ACKNOWLEDGEMENTS

The authors thank CGG Services (U.S.) Inc. for providing licenses for their Geosoftware package, including Hampson-Russel Suite modeling and analysis software. We also thank MathWorks for providing MATLAB[©], Simulink[©] under the academic license to the University of Arizona. We gratefully acknowledge ConocoPhillips and its partners, Statoil, Petoro, and Eni for providing the Heidrun Field data set used in this research, with a special thanks to Graham Hill for his support and assistance in this procedure. We also thank Halliburton Inc. for providing Landmark Graphics interpretation software through the Landmark University Grant Program.

REFERENCES

- Breiman, L. 1996, Bagging Predictors, Machine Learning, Springer, **24**. p. 123-140.
- Breiman, L. 2001, Random Forests, Machine Learning, Kluwer Academic Publishers, **45**. Boston, p. 5-32.
- Goodway, B., Chen, T., and Downtown, J. 1997, Improved AVO Fluid Detection and Lithology Discrimination using Lamé Petrophysical Parameters; “ $\lambda\rho$ ”, “ $\mu\rho$ ”, & “ λ/M Fluid Stack”, from P and S Inversions, SEG Technical Program Expanded Abstracts, p. 183.
- Hampson, D. P., Schuelke, J. S., and Quirein, J. A. 2001, Use of Multiattribute Transforms to Predict Log Properties from Seismic Data, Geophysics, **66**. p. 220-236.
- Keynejad, S., Sbar, M., and Johnson, R. 2017, Comparison of Model-Based Generalized Regression Neural Network and Prestack Inversion in Predicting Poisson's Ratio in Heidrun Field, North Sea, The Leading Edge, Society of Exploration Geophysicists, **36**. p. 938-946.
- Keynejad, S., Sbar, M. L., and Johnson, R. A. 2018, Assessment of Machine Learning Techniques in Predicting Litho-Fluid-Facies Logs in Hydrocarbon Wells, Manuscript submitted for publication.
- Liaw, A., and Wiener, M. 2002, Classification and Regression by Random Forest&nbs; R News, **2/3**. p. 18-22.

Mamdani, E. H., and Assilian, S. 1975, An Experiment in Linguistic Synthesis with a Fuzzy Logic Controller, International Journal of Man-Machine Studies, **7**. p. 1-13.

Norwegian Petroleum Directorate. Heidrun Field. Retrieved from
<http://factpages.npd.no/factpages>

Roden, R., Smith, T. A., Santogrossi, P., Sacrey, D., and Jones, G. 2017, Seismic Interpretation Below Tuning with Multiattribute Analysis, The Leading Edge, Society of Exploration Geophysicists, **36**. p. 330-339.

Ross, C. P., and Cole, D. M. 2017, A Comparison of Popular Neural Network Facies-Classification Schemes, The Leading Edge, Society of Exploration Geophysicists, **36**. p. 340-349.

Torres, A., and Reverón, J. 2014, Integration of Rock Physics, Seismic Inversion, and Support Vector Machines for Reservoir Characterization in the Orinoco Oil Belt, Venezuela, The Leading Edge, Society of Exploration Geophysicists, **33**. p. 774-782.

Zadeh, L. A. 1965, Fuzzy Sets, Information and Control, **8**. p. 338-353.

LIST OF FIGURES

Figure 1. top: nine FIS rules in symbolic-verbose form (is equal to: $==$, is not equal to: \sim , results in: \Rightarrow), and as membership function graphs of each attribute-class and output (G, O, Sh, and B stand for gas-sand, oil-sand, shale, and brine-sand classes, respectively). An example input's attributes and its resultant output are shown on the membership function graphs by vertical red lines and values. Bottom: an example of a decision surface that relates only two attribute-class sets (LR and PR) to the output fuzzy classes on the z-axis. Bottom left shows membership functions of LR and PR attribute-classes that are used with other attribute-classes in formation of the rules.

Figure 2. A small section of a sample tree. The branches ending at selected leaves are shown in thicker lines. Leaves 1, 2, 3, and 4 are gas-sand, shale, oil-sand, and brine-sand classes respectively. Attributes at splitting nodes are: MR for mu-rho, Pol for apparent polarity, TWT for time, Den for density, PR for Poisson's ratio, and Phi for porosity.

Figure 3. FIS results shown on arbitrary line A (top left inset), going through four training wells. The LFF classes are blue: brine sand, green: oil sand, brown: shale, and red: gas sand. The bottom right inset shows one of the parameters, LR, on a section of the same line between wells 7-3 and 8-1. See text for discussion.

Figure 4. top: PNN results shown on arbitrary line A (top left inset), going through four training wells. Bottom: the probability of oil sand (right) and gas sand (left) occurrences on the inline through well 8-1 (highlighted by red box in top figure), with the actual LFF log overlain at the well location.

Figure 5. top: BT results shown on arbitrary line A (top left inset), going through four training wells. Bottom: the probability of oil sand (right) and gas sand (left) occurrences

on the inline through well 8-1 (highlighted by red box in top figure), with the actual LFF log overlain at the well location.

Figure 6. Arbitrary line B (top left inset) going through all training and test wells. The FIS classes and the combined probabilities of classes in the PNN and BT results are shown in blue: brine sand, green: oil sand, brown: shale, and red: gas sand. The color opacity represents the uncertainty of the predicted classes based on the probability (ML) or membership functions (FIS).

Figure 7. FIS output classes in 3D (left), and in time slices (right). Output membership degrees are: blue: brine sand, green: oil sand, brown: shale, and red: gas sand, with the lighter colors for lower degrees (less probable) of each class. The opacity of colors is modified as shown in color scale insets for 3D cases to better depict the variations, especially for HC classes.

Figure 8. 3D models of LFF-class probabilities for PNN (left), and BT (right). Classes are: blue: brine sand, green: oil sand, brown: shale, and red: gas sand, with the lighter colors for lower probabilities. The opacity of colors is modified as shown in color-scale insets to better depict the variations, especially for HC classes.

Figure 9. Left: LFF prediction in time slices by PNN. Time slices (a) and (b) show the predicted LFF class labels and class probabilities, respectively, for the Fangst Group. Time slice (c) shows the LFF probabilities for the Båt Group. Right: seismic amplitude averaged on an 8-ms window centered at the Fangst top as an indication of the interpreted OWC to compare with the predicted OWC.

Figure 10. Left: LFF prediction in time slices by BT. Time slices (a) and (b) show the predicted LFF class labels and class probabilities, respectively, for the Fangst Group. Time

slice (c) shows the LFF probabilities for the Båt Group. Right: seismic amplitude averaged on an 8-ms window centered at the Fangst top as an indication of the interpreted OWC to compare with the predicted OWC.

Figure 11. An inline of seismic amplitudes, overlain by HC classes predicted by (a) BT, and (b) PNN. The red lines show some of the interpreted faults. Oil sands and gas sands are in green and red, respectively.

TABLES

Table 1. The subset of attributes used by PNN, as selected by a step-wise regression method.

Attribute Name	Attribute Transform
Poisson's ratio (PR)	Log*Sign(PR)
Porosity (Phi)	1/Phi
Density (Rho)	1/Rho
Apparent Polarity	None

FIGURES

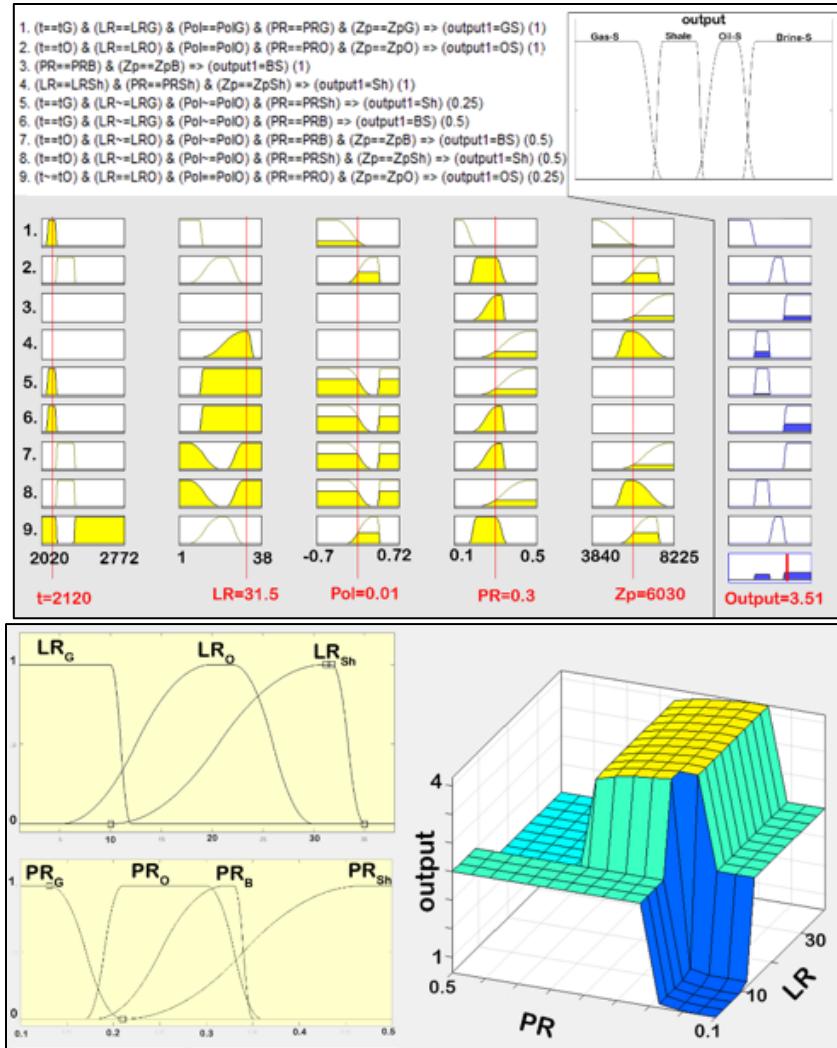


Figure 1. top: nine FIS rules in symbolic-verbose form (is equal to: $==$, is not equal to: $\sim=$, results in: \Rightarrow), and as membership function graphs of each attribute-class and output (G, O, Sh, and B stand for gas-sand, oil-sand, shale, and brine-sand classes, respectively). An example input's attributes and its resultant output are shown on the membership function graphs by vertical red lines and values. Bottom: an example of a decision surface that relates only two attribute-class sets (LR and PR) to the output fuzzy classes on the z-axis. Bottom left shows membership functions of LR and PR attribute-classes that are used with other attribute-classes in formation of the rules.

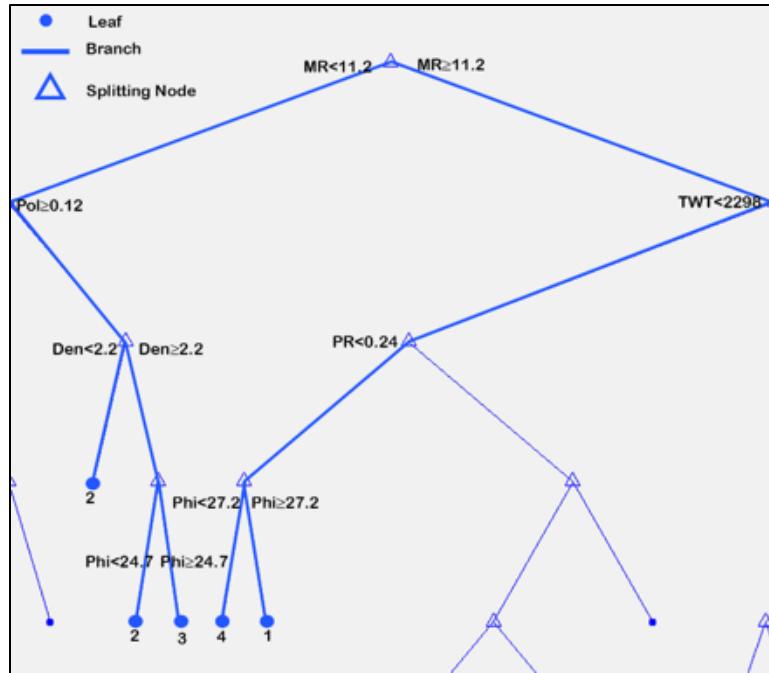


Figure 2. A small section of a sample tree. The branches ending at selected leaves are shown in thicker lines. Leaves 1, 2, 3, and 4 are gas-sand, shale, oil-sand, and brine-sand classes respectively. Attributes at splitting nodes are: MR for mu-rho, Pol for apparent polarity, TWT for time, Den for density, PR for Poisson's ratio, and Phi for porosity.

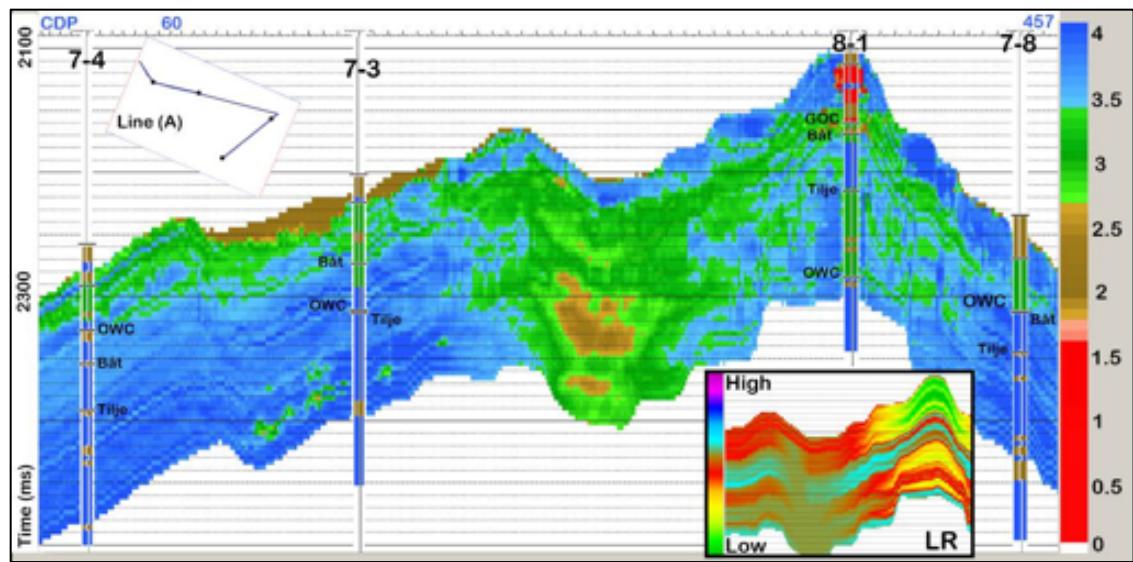


Figure 3. FIS results shown on arbitrary line A (top left inset), going through four training wells.

The LFF classes are blue: brine sand, green: oil sand, brown: shale, and red: gas sand. The bottom right inset shows one of the parameters, LR, on a section of the same line between wells 7-3 and 8-1. See text for discussion.

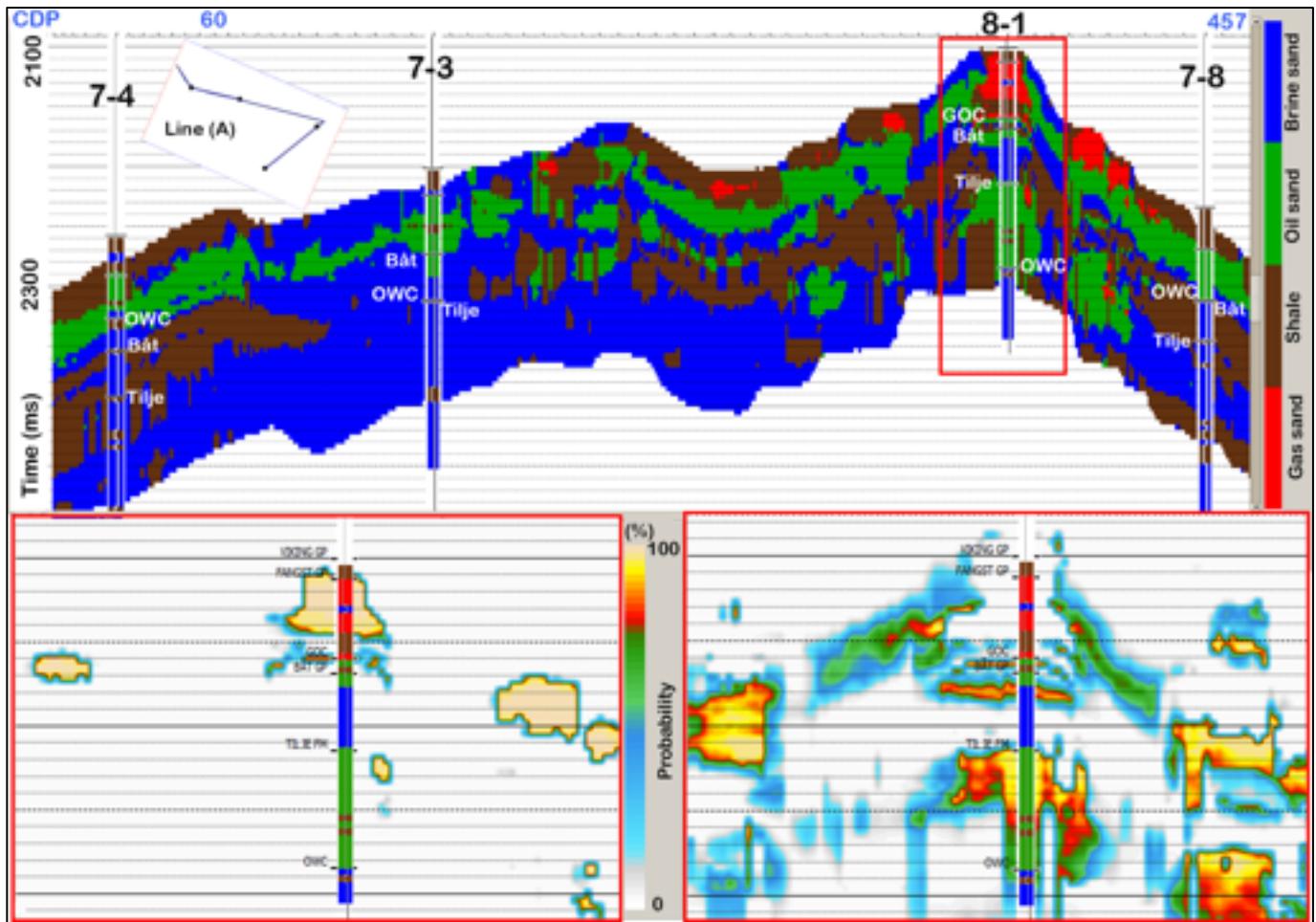


Figure 4. top: PNN results shown on arbitrary line A (top left inset), going through four training wells. Bottom: the probability of oil sand (right) and gas sand (left) occurrences on the inline through well 8-1 (highlighted by red box in top figure), with the actual LFF log overlain at the well location.

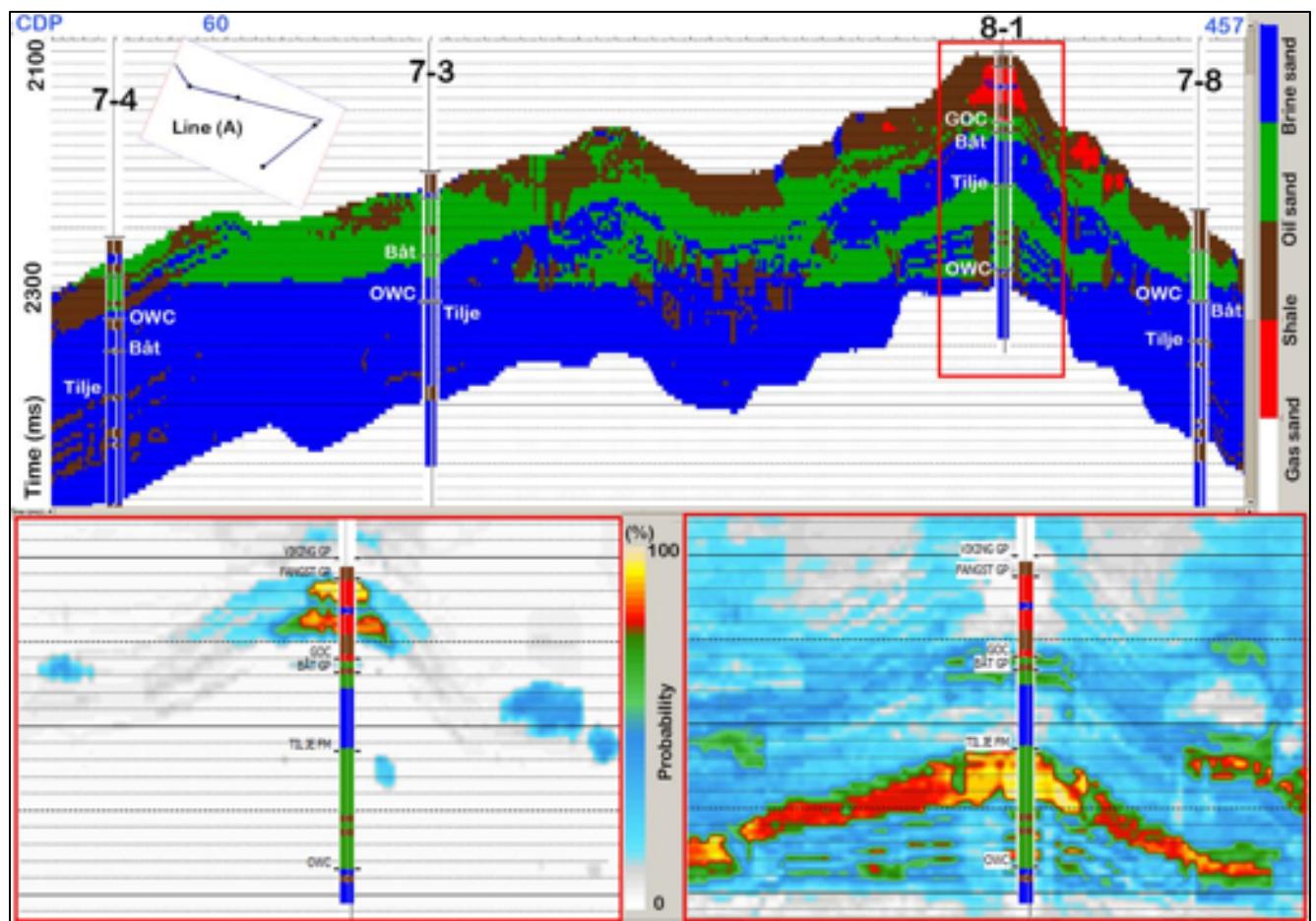


Figure 5. top: BT results shown on arbitrary line A (top left inset), going through four training wells. Bottom: the probability of oil sand (right) and gas sand (left) occurrences on the inline through well 8-1 (highlighted by red box in top figure), with the actual LFF log overlain at the well location.

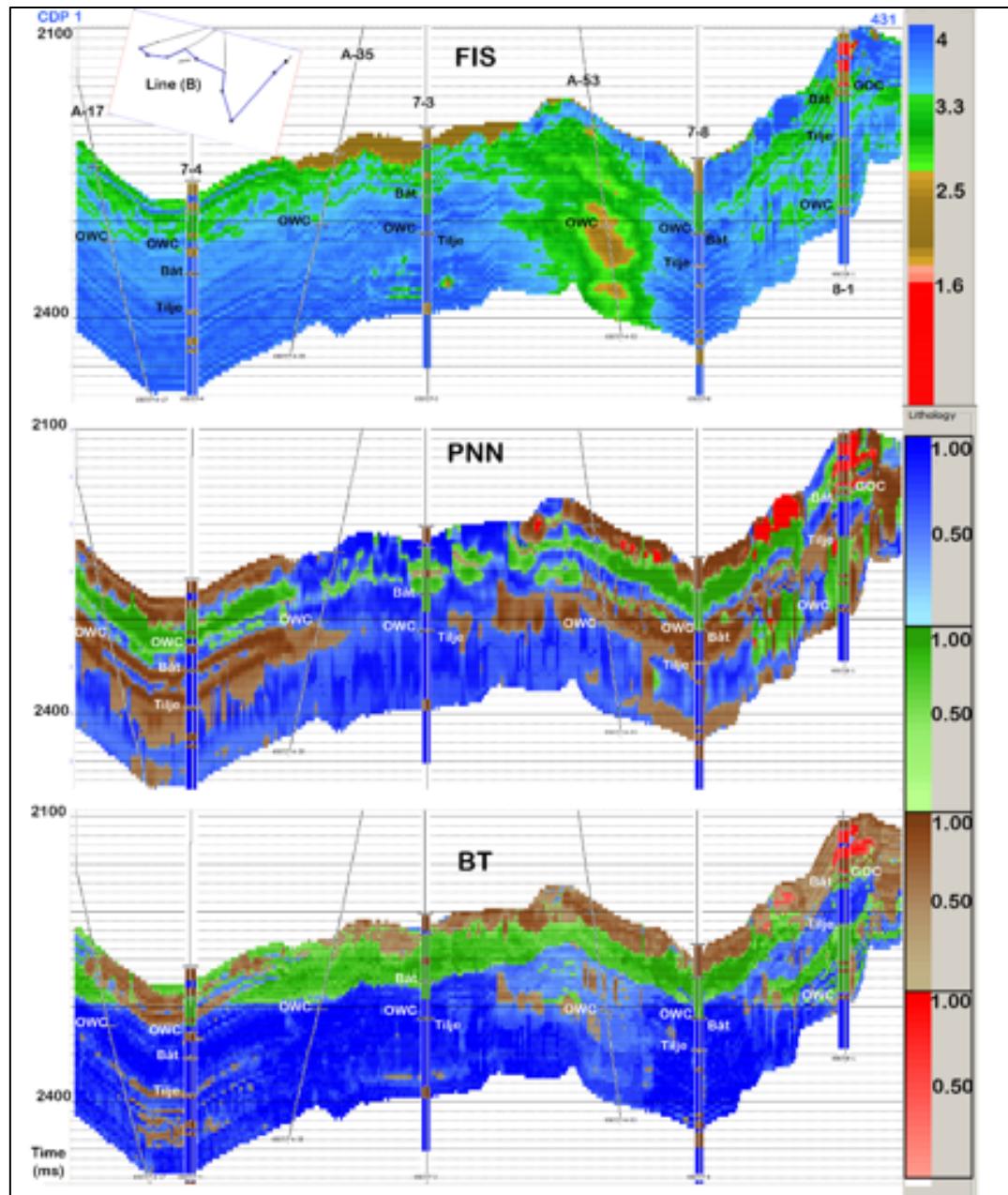


Figure 6. Arbitrary line B (top left inset) going through all training and test wells. The FIS classes and the combined probabilities of classes in the PNN and BT results are shown in blue: brine sand, green: oil sand, brown: shale, and red: gas sand. The color opacity represents the uncertainty of the predicted classes based on the probability (ML) or membership functions (FIS).

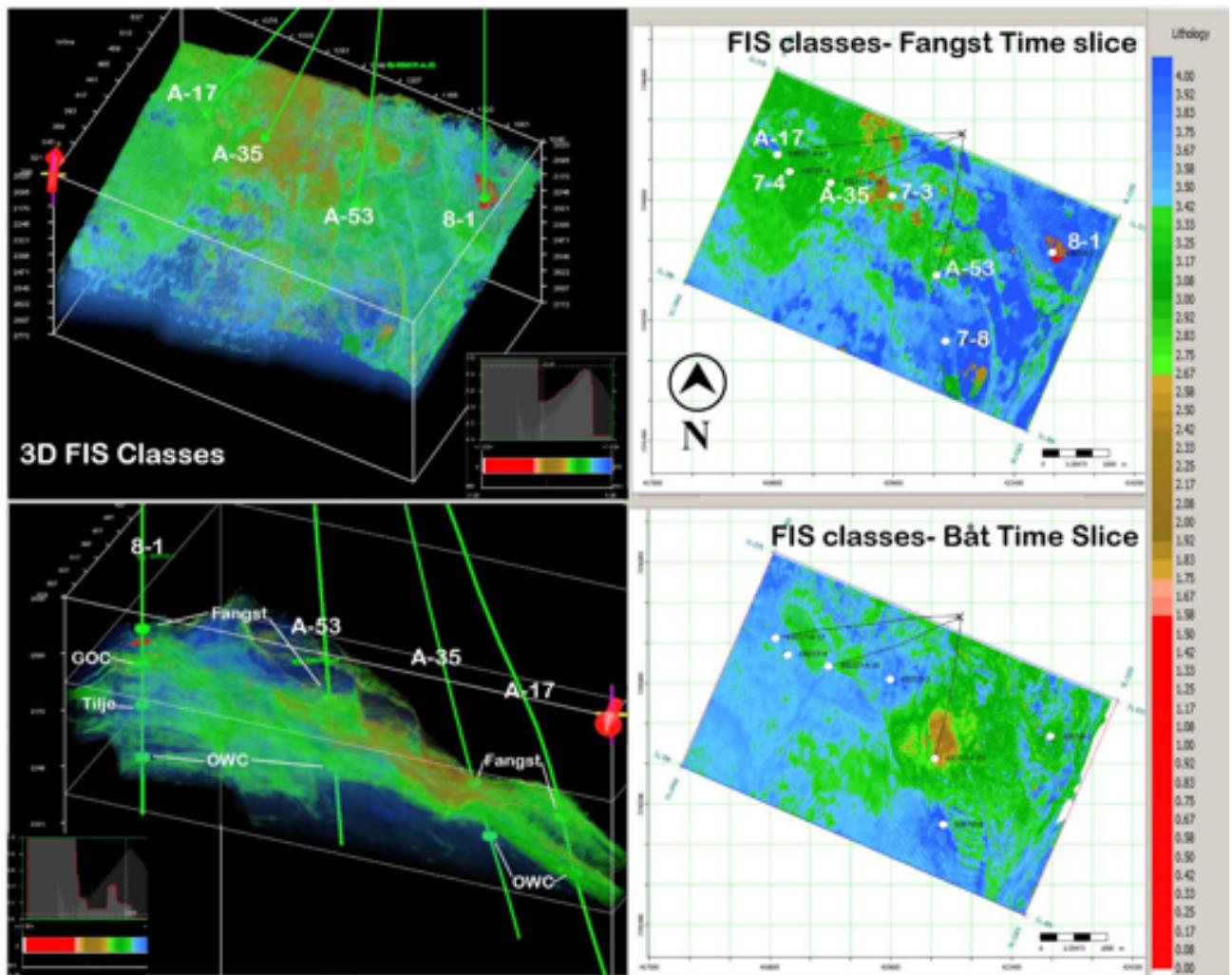


Figure 7. FIS output classes in 3D (left), and in time slices (right). Output membership degrees are: blue: brine sand, green: oil sand, brown: shale, and red: gas sand, with the lighter colors for lower degrees (less probable) of each class. The opacity of colors is modified as shown in color scale insets for 3D cases to better depict the variations, especially for HC classes.

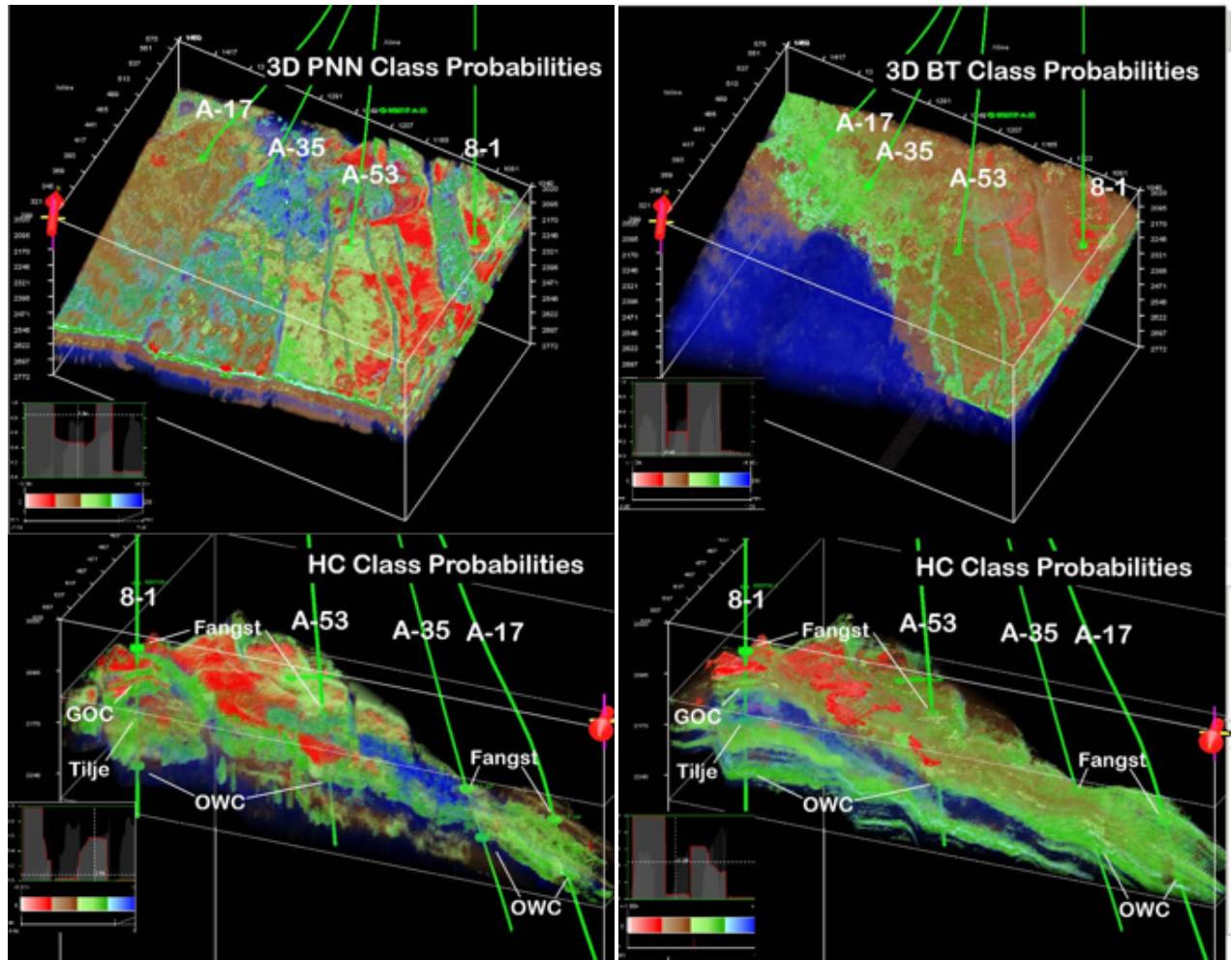


Figure 8. 3D models of LFF-class probabilities for PNN (left), and BT (right). Classes are: blue: brine sand, green: oil sand, brown: shale, and red: gas sand, with the lighter colors for lower probabilities. The opacity of colors is modified as shown in color-scale insets to better depict the variations, especially for HC classes.

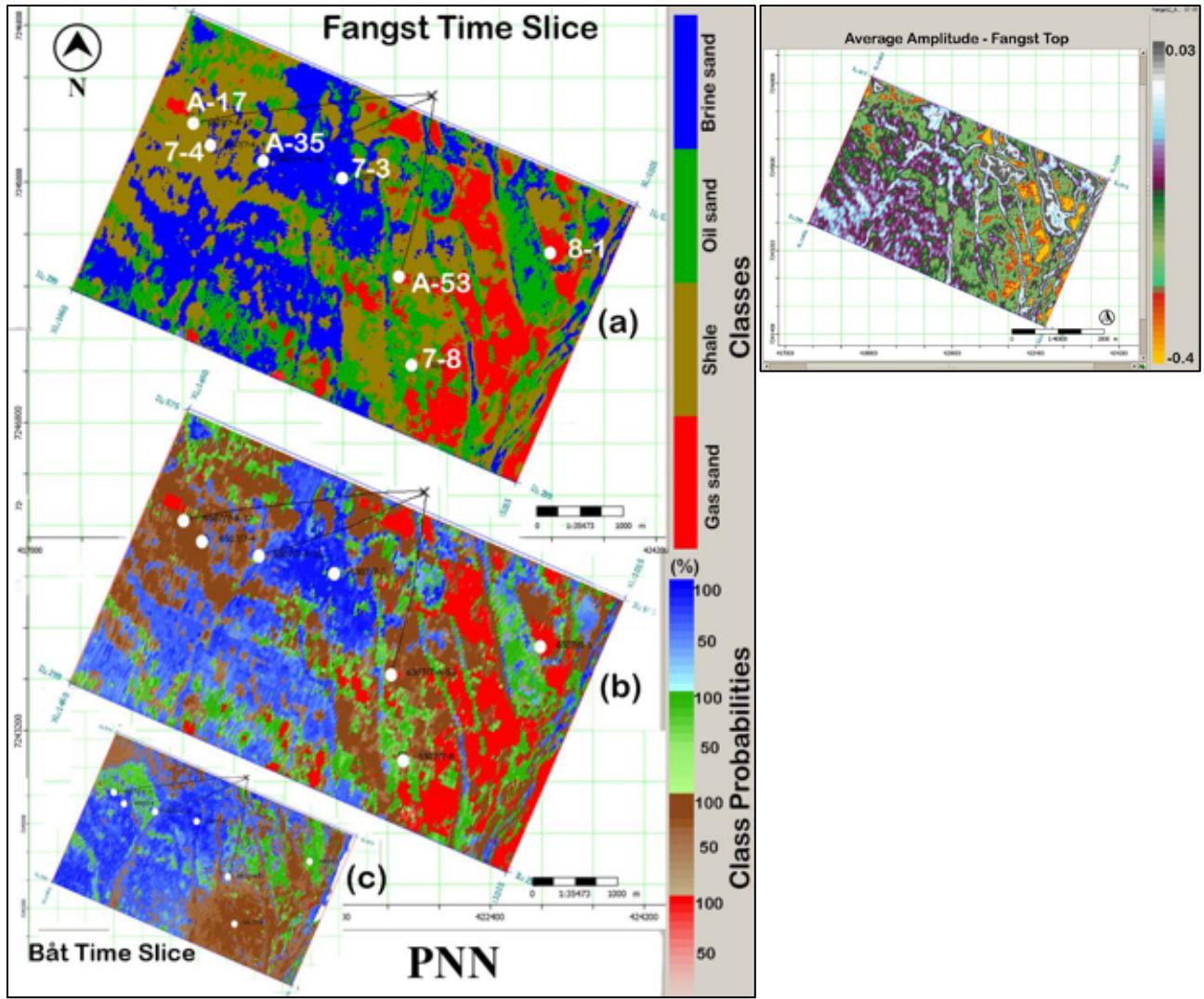


Figure 9. Left: LFF prediction in time slices by PNN. Time slices (a) and (b) show the predicted LFF class labels and class probabilities, respectively, for the Fangst Group. Time slice (c) shows the LFF probabilities for the Båt Group. Right: seismic amplitude averaged on an 8-ms window centered at the Fangst top as an indication of the interpreted OWC to compare with the predicted OWC.

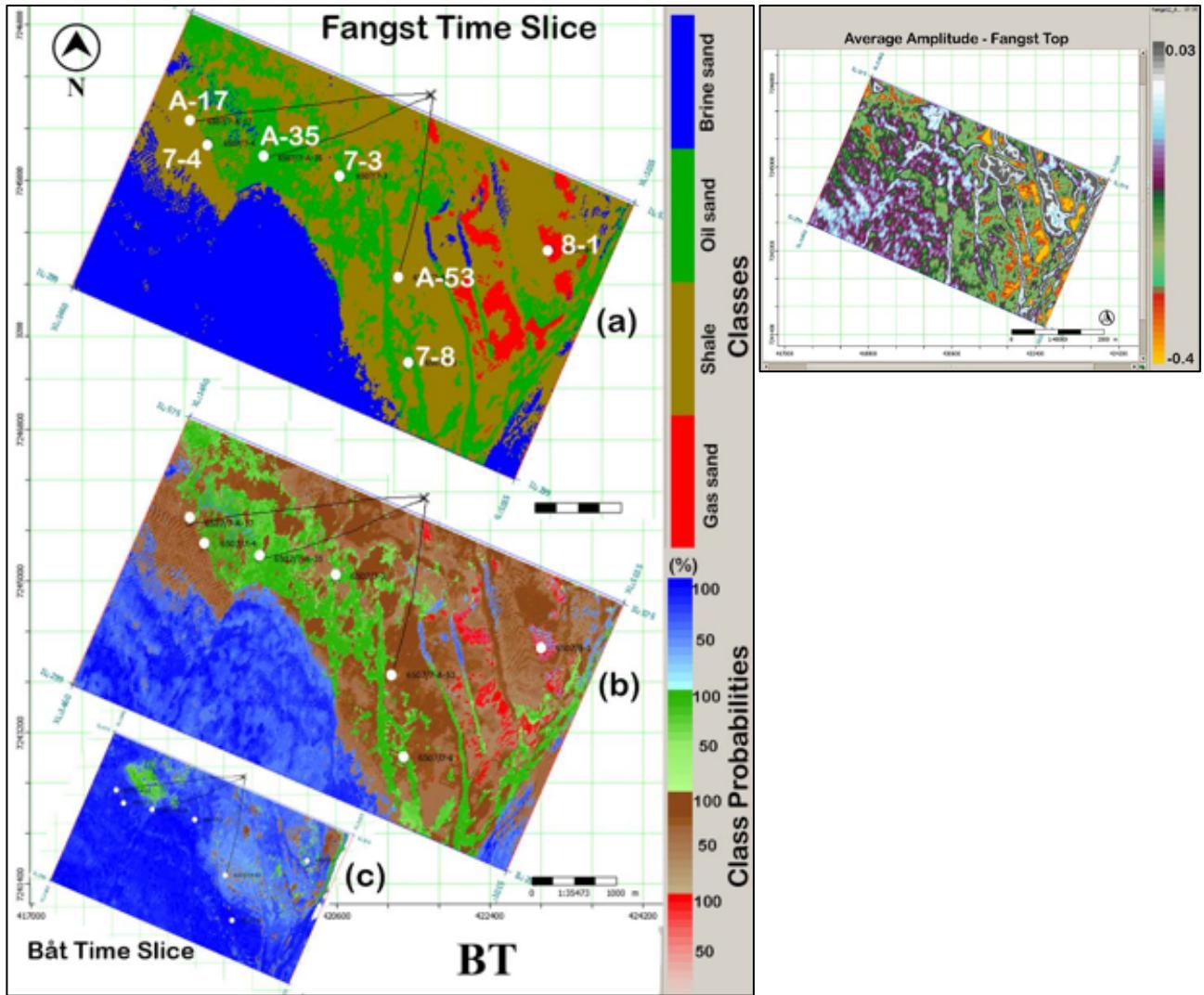


Figure 10. Left: LFF prediction in time slices by BT. Time slices (a) and (b) show the predicted LFF class labels and class probabilities, respectively, for the Fangst Group. Time slice (c) shows the LFF probabilities for the Båt Group. Right: seismic amplitude averaged on an 8-ms window centered at the Fangst top as an indication of the interpreted OWC to compare with the predicted OWC.

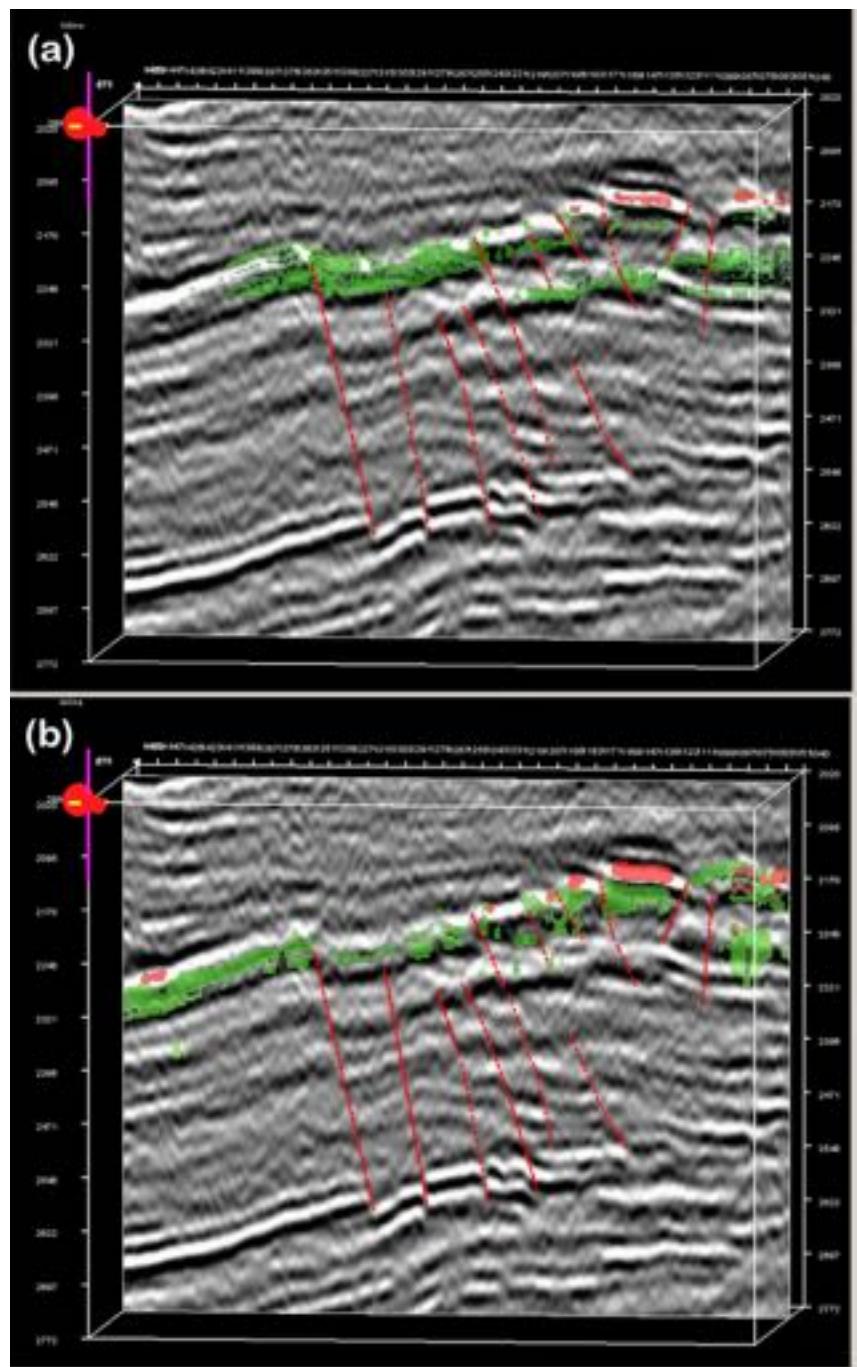


Figure 11. An inline of seismic amplitudes, overlain by HC classes predicted by (a) BT, and (b) PNN. The red lines show some of the interpreted faults. Oil sands and gas sands are in green and red, respectively.

**APPENDIX D: TERMS AND DEFINITIONS OF THE MACHINE LEARNING
TECHNIQUES DISCUSSED IN THIS DISSERRTATION**

In the following sections, description of the machine learning approaches that are used in this dissertation and a list of the associated terms are provided. See the references for more detailed description of the mathematical basis and variations of each algorithm. In the equations, matrices and vectors are denoted as bold upper-case and bold lower-case letters, respectively. Superscript and subscript numbers and letters indicate a layer and a member of a set, respectively. Terms in bold text are defined in the glossary.

ARTIFICIAL NEURAL NETWORKS

The architecture and the learning process of the ANNs used in this dissertation are described in the following sections. A general form of the learning process, in all the supervised ANNs used here, can be summarized in these steps:

1. The network is initiated on a set of (usually) randomly selected **connection weights** and **bias** values.
2. In each training iteration, a set of **input vectors** and the associated **target** values/labels are presented to the network.
3. Based on the calculated error, the connection weights and biases are modified to minimize the **error**.

In the third step, the information is commonly passed backward (from output to input layer), which gives the name “backpropagation” to the algorithm. For the unsupervised case, instead of providing target values in the second step, the input values are clustered based on their similarity to the connection weights.

The output values from the i -th **layer** with S **neurons** to the next hidden or to the output layer are calculated by **activation function** f^i :

$$\mathbf{a}^i = f^i(\mathbf{W}\mathbf{p} + \mathbf{b})$$

In the above matrix-form expression, a^i is the output vector of layer i with S elements, p is the input vector with R elements, W is the $S \times R$ matrix of weights connecting R input elements to S neurons, and b is the vector of S -elements of bias vector connected to each neuron in the i -th layer.

By calculating the derivatives of the error with respect to the weights (and biases) connecting the output neurons to the neurons in its preceding layer and minimizing it, the associated weights are modified by the value:

$$\Delta w^{ji} = \beta \delta^j p^i, \text{ where } \delta^j = \frac{\partial f^j}{\partial a^i} e.$$

Where e is the error as a measure of difference between output d and target t , and p^i is the input for the preceding i -th layer. Parameter β is the “learning rate” that controls the speed of convergence. In the next step, the connection weights between the hidden layer i and its preceding layer, layer $i-1$, are updated by: $\Delta w^{i(i-1)} = \beta \delta^i p^{i-1}$, with $\delta^i = \frac{\partial f^i}{\partial a^{i-1}} \sum \delta^j w^{ji}$, and so on. To read more about the alternative forms of modifications and variations on finding the minimum on the error surface, and on executing the backpropagation algorithm, see sources such as Hagan et al. (1996) and Gurney (1997).

Multilayer feed-forward neural networks

This type of network has at least one hidden layer, propagates data in a feed-forward direction, and uses non-linear, differentiable threshold functions. Figure 1 shows a multilayer feed-forward network with one hidden layer. The output activation function (f) can be a linear or a sigmoid function for regression and classification purposes, respectively. Sigmoid functions usually have the form of either a tangent-sigmoid ($f_{tansig} = \frac{e^n - e^{-n}}{e^n + e^{-n}}$), or a logarithm-sigmoid ($f_{logsig} = \frac{1}{1+e^{-n}}$).

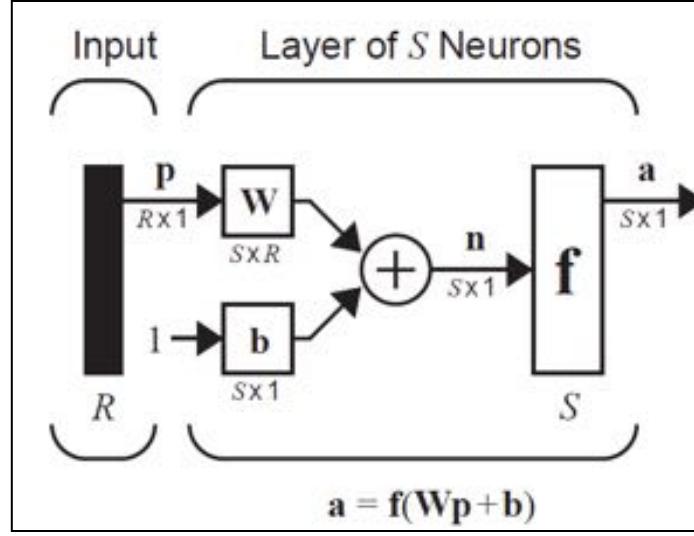


Figure 1. A multilayer ANN with one hidden layer R attributes and S outputs from Hagan et al. (1996).

Generalized regression and probabilistic neural networks

Both of these ANNs are types of radial basis neural networks, a category of ANNs that have radial basis function (RBF) in their first hidden layer (Specht, 1990; Specht, 1991). A Gaussian function is the most commonly used RBF in these networks and is defined as $f = \exp\left(\frac{-a^2}{2\sigma^2}\right)$. Two main differences between multilayer feed-forward and RBF networks happen in this layer; instead of a dot product, the ‘distance’ between the weights and inputs are calculated, and then this measure of closeness is multiplied by bias values ($\frac{1}{\sigma\sqrt{2}}$) that controls the “spread” of the function by σ values (Figure 2). In training, the connection weights, and the sigma values are modified to optimize the network. After passing the distances and biases through RBF, the second layer calculates the ultimate output by output function f .

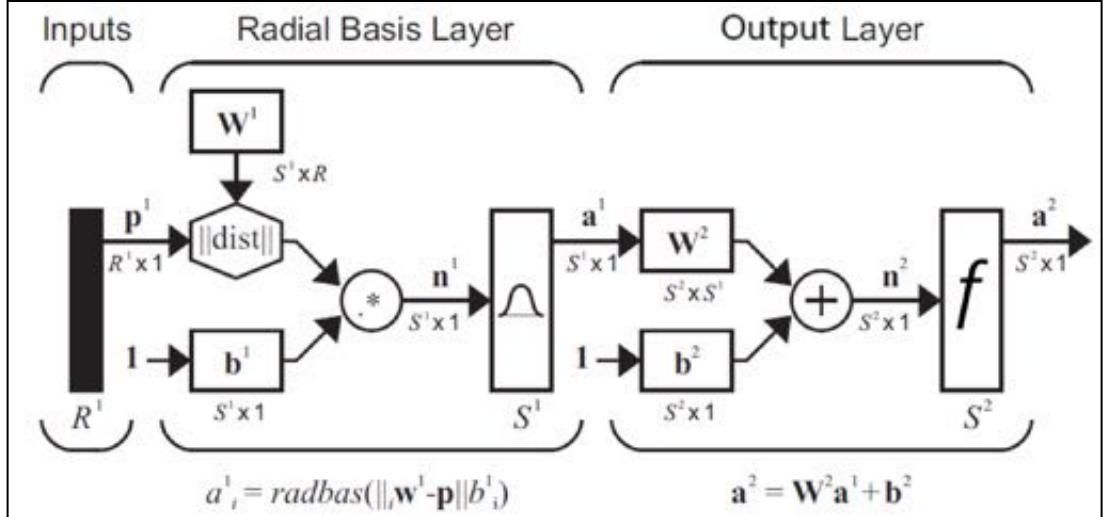


Figure 2. A general representation of a radial basis function (RBF) neural network from Hagan et al. (1996).

With a linear function in the second layer, the outputs will be continuous values. This is the case for the generalized regression neural networks that are used in function-approximation (regression) problems. The weights in the second layer of this network are set to the target values. In this way, when the input vectors are presented, an input vector close to a certain training vector (p_i) will have a distance close to zero, which results in an RBF output close to 1 ($a_i^1 \approx 1$). Multiplying this output in the second layer with the target values (as weights) produces outputs close to the target value associated with the training vector ($a_i^2 \approx t_i$). Optimizing sigma values are the main controlling factors in this procedure; smaller sigma values mean a narrower RBF, which means only a small neighborhood around the “distance zero” is considered close enough to “fire” 1. However, too small neighborhoods can result in poor generalization due to overfitting.

For pattern-recognition problems, the second layer uses a transfer function called the competing function that basically picks the neuron with the highest value. Probabilistic

neural networks are in this group. With the weights in the first layer set to the transpose of the matrix of the training vectors, this layer finds the distance between the input vector and each training vector. The RBF returns a value for each pair of input/training vector, with the higher values for the closer pairs. In other words, the output of the first layer indicates the closeness of the input vector to each training vector. The connection weights of the second layer are set to the vectors of the **target index vector**. The result of multiplying these weights by the previously calculated “measure of closeness” is, in fact, a vector of probabilities of each class for that input vector. The competing function at the output layer fires 1 for the neuron (class) with the highest probability and zeros for the rest.

Self-organizing feature maps

In this unsupervised ANN, there is one hidden layer in which a competitive transfer function generates the ultimate outputs. First, the competitive layer measures the similarity (e.g., distance) between the input data and the training vectors to determine the winner neuron (the most similar) and the neurons in a defined neighborhood around it. Then, the weights for the neurons within that neighborhood are updated according to the Kohonen rule (Kohonen, 1987):

$$w^{new} = w^{old} + \alpha(p - w^{old})$$

This way, each time an input vector p is presented, the weights of the winning neuron(s) will move toward it, and after multiple times the neighboring neurons will learn similar vectors. This approach, which causes neighboring neurons to learn together, makes this network capable of what is known as preserving the topology of the input data. To optimize the learning procedure and increase the stability, the distance and learning rate are adjusted (usually decreased) during the training phase.

BOOTSTRAP AGGREGATING TREES

Classification trees are trained on a dataset of input/target pairs (p_i, t_i), with each p_i being a vector of attribute values, and t_i being the associated class label. The goal is to find the rule or predictor (C) that assigns a class label to each presented vector of attributes. A tree is formed (grown) on the dataset as its root, with branches that split several times until each ending branch reaches a certain **leaf** as the class label. The splitting **nodes** on the branches split the attributes into two or more “conditions” for continuous attributes (e.g., $p_{i1} < x$ and $p_{i1} \geq x$) or for discrete attributes (by feature-specific labels e.g., p_{i1} (color)= *red* and p_{i1} (color)= *blue*). The nodes in a classifier tree are usually selected by means of “information gain” to assess the importance of a feature at a certain node. After a tree is trained, the new data is dropped down the tree to be classified according to the assigned features and labels.

Bagging trees method grows multiple of these trees on subsets of the dataset S , each subset being randomly selected *with replacement* from the original dataset, also known as bootstrap resampling (Breiman, 1996). By using a random forest approach, the splitting nodes at each branch are determined by randomly (but without replacement) selecting m features from M available features. Random forest “de-correlates” different trees, and thus reduces the prediction error (Breiman, 2001). In a classification problem, the assigned class to the input p is determined by:

$$C_{bag}(p) = \bigvee \{C(S^{*b}, p)\}_{b=1}^B$$

Where S^{*b} is a bootstrapped sample of dataset S , and operator \bigvee assigns the majority vote of the classes to the input. It means that the assigned class labels for a certain **observation** is counted over all the bagged trees, and the class with the most votes is

selected as the result. The strength (accuracy) of a tree and the correlation between any two trees, both are directly related to the value m . Using the out-of-bag (**OOB**) error, an optimum m can be found to reduce the prediction error. The bagged trees with a random forest approach do not need to be pruned since overfitting is not an issue as it is in single decision trees.

GLOSSARY

- **Bias (ANN)**: scalar values connected to neurons to add stability and reduce convergence time, and to stay away from a possible saddle point at the origin.
- **Connection weights (ANN)**: The weight w_{ji} between the j -th layer and the neurons in its preceding i -th hidden layer.
- **Cross-validation**: estimating the accuracy of the approach by calculating the error of predicting an excluded part of the training data (one part at a time), by training the algorithm on the rest of the training data. Excluding one part can be done by a random selection (e.g., a percentage of training data), selecting one source of data at a time (e.g., one training well), or the out-of-bag data in bagged trees.
- **Error (e)**: a measure of the difference between a predicted outcome of, and the target assigned to, an input vector estimated after training, usually as the root mean square of all errors.
- **Input/training vectors**: a matrix of $Q \times R$ values. The values of a vector $p_k = (x_{k1}, x_{k2}, \dots, x_{kR})$ are the k -th value of R attributes.
- **Layers (ANN)**: a set of processing elements in an ANN.

- Leaves (**BT**): the ending points at a tree's branches that have a label or a value for classification or regression, respectively.
- Neurons (**ANN**): processing elements of an artificial neural network that receive information (raw or processed data, weights, or bias) and passes processed information to the next layer.
- Nodes or splitting nodes (**BT**): on branches of a tree, each node is a question about, or a separation in, an attribute's features or range.
- Observation (**BT**): the observation X is a direct path from a tree's root (the data set it performs on) to one leaf. It has the elements (x_1, x_2, \dots, x_n) , each of which is created by a node. It differs from input vectors since the length of all observations are not necessarily equal. For instance, from a set of six attributes $(x_1, x_2, x_3, x_4, x_5, x_6)$, two observations on two branches of one tree can be: $X_1 = (x_1 < a, x_2 \geq b, x_5 < c \mid \text{class1})$ and $X_2 = (x_1 < d, x_3 \geq e, x_4 \geq f, x_5 < g \mid \text{class2})$.
- OOB (out-of-bag) data (**BT**): the data that has not been selected in bootstrap sampling for a tree. This data is usually used for cross-validation and performance assessment.
- Target and target index: used in training a supervised algorithm, for each input vector p_i , there is an associated t_i that can be either a value (regression) or a target index (pattern recognition). A target index is a C -length vector for C classes, with 1 for the m -th class and zeros for the rest to denote class m as the response.
- Transfer, threshold or activation function (**ANN**): a function that maps a layer's input data to its output, by aggregating the input, weights and biases connected

to all its neurons. The function type of each layer varies based on the architecture of the ANN, and the task of that layer.

REFERENCES

Breiman, L. 1996, Bagging Predictors, Machine Learning, Springer, **24**. p. 123-140.

Breiman, L. 2001, Random Forests, Machine Learning, Kluwer Academic Publishers, **45**.
Boston, p. 5-32.

Gurney, K. 1997, An Introduction to Neural Networks, UCL Press.p. 234.

Hagan, M. T., Demuth, H. B., and Beale, M. H. 1996, Neural Network Design, PWS Pub.

Kohonen, T. 1987, Self-Organization and Associative Memory, Springer-Verlag.

Specht, D. F. 1991, A General Regression Neural Network, TNN, IEEE, **2**. United States,
p. 568-576.

Specht, D. F. 1990, Probabilistic Neural Networks, Neural Networks, **3**. p. 109-118.