AUTOMATIC INTERPRETATION OF WELL LOGS WITH LITHOLOGY-SPECIFIC DEEP-LEARNING METHODS

Aymeric-Pierre Peyret, Joaquín Ambía, and Carlos Torres-Verdín, The University of Texas at Austin, and Joachim Strobel, Wintershall GmbH

ABSTRACT

Accurate and reliable interpretation of well logs often requires a high level of expertise from a petrophysicist along with enough relevant borehole and core measurements. As an alternative and complementary approach, deep learning has been proposed to capture the experiential knowledge gained from petrophysical interpretations, as well as the physical and heuristic models often used for that purpose. We test the latter idea using a set of wells previously interpreted and explore the critical aspects that can yield a successful automatic well log interpretation.

Some of the questions we attempt to answer here, as a guide for future applications of artificial neural networks (ANNs), are: how much data does the petrophysicist need to explicitly interpret before relying on ANNs? What is the best suited deep-learning network architecture? How widely can the ANNs be generalized? Can we automatically classify different wells for improved ANN usage? In this study, all the wells used come from the same hydrocarbon reservoir and intersect multiple formations. The focus is on estimating clay fraction (VCL), effective porosity (PHIE), water saturation (SW), and permeability (K).

We compare the performance of various architectures of deep artificial neural networks (ANNs) using different numbers of layers and neurons in each layer. Once an ideal ANN architecture was found for a specific formation, it was tested against different formations, but results were relatively poor, corroborating the specificity of ANNs to the lithology where they were trained. Furthermore, we propose a self-organizing map (SOM) as a way of partitioning wells into classes, which should be treated separately (independent ANNs). This strategy yielded the best results.

The amount of data required to train our ANNs was relatively small (5-9 wells), considering the amount of

data typically required for training more general ANNs. This is only possible if different lithologies and/or rock classes are treated independently.

INTRODUCTION

The use of artificial intelligence (AI) in well-log interpretation has been proposed multiple times, and for different purposes. Classification of facies within well trajectories has been done by relying on well logs, and using ANNs (Saumen, et al., 2007; Rogers et al., 1992), modular neural networks (Bhatt, and Helle, 2002(1)), deep convolutional neural networks (CNN) (Imamverdiyev, and Sukhostat, 2019), and SOMs (Fung, et al., 1995).

The use of machine learning techniques has also been proposed to generate synthetic well logs to replace missing measurements, or entire runs using generalized regression ANNs (Rolon, et al., 2009; Khandelwal, and Singh, 2010).

Furthermore, the use of modular ANNs to predict petrophysical properties, including permeability, has been proposed by Rogers et al. (1995). The strategy has been improved by adding a lithology classification step before using an ANN (Fung, et al., 1997). As AI techniques have been evolving in recent years, they have been applied to tackle the same problem. Some of these strategies include using committee machines (Bhatt, and Helle, 2002(2)), including a fuzzy logic classification step (Lim, 2005), using general regression neural networks (GRNN), and support vector machines (SVM) (Gholami et al., 2012).

In this study we propose revisiting legacy logs and applying machine learning methods to understand the nuances of ANN design such as the number of hidden layers and neurons each layer should have. We also assess the amount of data required for a successful ANN. Once an ANN is established for a specific facies, we evaluate its generalizability by testing the network across facies, i.e. ANNs trained in one lithology and applied in a different one. Finally, given that ANNs perform poorly across facies, we consider a further classification of wells

SSSS

themselves (within the same facies). We use a SOM that relies on well logs, and the correlations and auto-correlations of their derivatives. With these studies, we propose a comprehensive and robust workflow, applicable to legacy and new well logs.

FIELD DESCRIPTION

The Wintershall Landau field is a sandstone oil reservoir located in a peri-urban area of Southwestern Germany, in the Rheingraben.
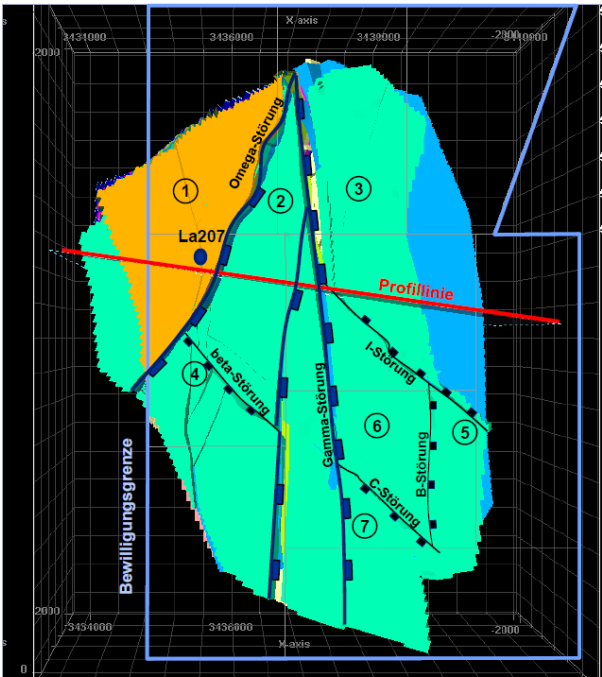

Figure 1: Map of the fault blocks in the Landau field. Blocks are: 1-Nußdorfer Horst, 2-Walsheimer, 3-Knöringer, 4-Nußdorfer, 5-Dammheimer, 6-Bauhorst, and 7-Queichheimer.

The field size is about 4 km x 4 km. Faults (Störungen on the map) separate the reservoir into 7 main blocks ("segments"), named for small villages around the city of Landau: the Nußdorfer Horst (NH), Walsheimer segment (WH), Knöringer segment (KN), Nußdorfer segment (NSE), Dammheimer segment (DHN / DHS), Bauhorst segment (BH) and Queichheimer segment (QH). The relative location of these blocks is shown on **Figure 1**. Most of the wells available for this study were logged in the mid-1950s to mid-1960s, and their location with respect to the fault blocks is shown **Table 1**. The reservoir consists mainly of tertiary clastic sediments originating from the alpine orogenetic processes and fluvially deposited into the extensional opening of the

Upper Rhine valley´s shallow marine environment. Quartz and carbonate exist in equal properties within the numerous reservoir layers of 3-10 m thickness which are separated by 10-100 m of non-reservoir yielding a typical net-to-ross ratio of around 20%. Porosity mean values reach from 20 to 25 percent while permeability ranges from 2 to 500 mD. Further details of the reservoir and its architecture can be found in Böcker (2015) and Strobel (2019).

Table 1: Well list and block location. Because a well may cross multiple blocks, only the most productive one is listed here.

Well	Block
A	DHN
B	DHN
C	DHN
D	DHN
E	BH
F	DHN
G	DHN
H	DHN
I	DHN
J	NSE
K	NH
L	KN
M	KN
N	NSE
O	NH
P	NH
Q	NH
R	NH
S	NH
T	KN
U	KN
V	KN
W	WH
X	WH
Y	WH
Z	WH
AA	WH
BB	WH
CC	WH
DD	NH
EE	NSE
FF	NH
GG	NH

Our main zone of interest is the "Bunte Niederröderner Schichten" (BNS), which may be translated as

"multicolored layers of Niederrödern". The Niederrödern formation is described as "Multicolored marls, sand marls and sandstones of limnal origin. The succession is laterally strongly varying and shows the beginning of limestone banks and channel fills, with sometimes cobble- to pebble-carrying sandstones. Good reservoir sands are found particularly in the upper half. South of Kaiserstuhl, the Niederrödern formation merges into the Tüllingen formation. This is a succession of gray and green marls with partly meter-sized, blocky, in places cavernous limestone." (Jodocy and Strober 2011).

DATA INVENTORY AND QA/QC

Two versions of the dataset have been analyzed for this project: the "raw" dataset where no log was touched before processing, and a "clean" dataset where logs have been cleaned up with the following process: 1- Erroneous extremely large values of gamma ray have been removed (replaced by "no-value" points). 2- Flat sections of logs have been replaced by no-value points. These sections may have been intentionally placed (for instance, porosity logs set to a very small value of porosity in non-net areas) or not (such as the bottom- and top-ends of logs). Whether properties (porosity, permeability) are set to no value or to very small values in non-net sections depends on company practice and requires slightly different machine learning workflows. 3- The gamma ray was normalized in Techlog® with the quantile method. 4- Some negative values found in the resistivity logs were removed. No significant depth shift need was found in this dataset.

Similarly, to many brownfields that are now considered for enhanced oil recovery (EOR), the original data was mostly recorded on paper and had been digitized. However, all header data is fully available. About 10 wells were logged with an emerging digital system, while about 20 wells saw cased hole logs like Sigma and Gr. While digitization of logs has been thoroughly quality checked, it cannot reproduce the original shape of the logging curve to allow curve-shaped based algorithms to be applied which otherwise could have been helpful for the plentiful Microlog data. The typical resistivity log did consist of the short and long Normal log (SN/LN). While the formation water is moderately saline (30-50 ppk Nacl equivalent), most early wells were drilled with fresh mud yielding a very good SP response. The early logging tools were mostly run separately. This creates a non-recoverable slight depth shift between them which adds uncertainty to a digital log analysis. The variety of log types available in the main zone of interest for each well differs from well to

well; a log inventory is shown in **Appendix 1**.

From the caliper, we computed a quality index in the following way for each zone separately: (1) We estimate the mode of the caliper; (2) If $\left|\frac{mode}{bit\ size}\right| > 0.15$, which means that most of the interval shows caliper values 15% over the bit size, i.e. the whole interval is pretty much washed out, we use the bit size as reference size RS. Otherwise, the mode is probably not far away from the bit size (usually just slightly larger) and the mode of the caliper is taken as reference size RS. (3) We compute a first quality factor at every depth as $QF_1 = \frac{1}{1+\left|\frac{caliper}{RS}-1\right|}$, i.e. 1 when caliper equals the reference size, and decreasing when it goes farther away. (4) Based on the derivative of the caliper with respect to depth, a second quality factor QF_2 is computed and scaled to match the range of QF_1 values, i.e. 1 when the derivative is 0, decreasing with the (absolute value of the) inverse of the derivative, down to the same minimum as QF_1 when reaching the maximum value of the derivative. (5) QF_1 and QF_2 are subsequently harmonically averaged into a composite quality factor QF.

To maximize the number of wells that may be used in this study, the type of well logs taken for training the neural networks was reduced to the most common set of logs across all wells; those happen to be caliper, deep and shallow resistivity, spontaneous potential, depth, and composite quality factor. Some initial tests add the Gamma Ray to this list for the purpose of estimating the information added by this log.

METHODS

Given that ANNs and SOMs are well established methods, and the focus of our study is their application, we rely on the Matlab Statistics and Machine Learning Toolbox. This package allows us to modify the number of hidden layers in the ANN, and the number of neurons.

To assess the accuracy of our ANNs predictions, we compare them to the petrophysical properties previously calculated by an expert petrophysicist. The main metrics we use are the adjusted coefficient of determination (R^2) and the root-mean-square deviation (RMSD). The RMSD is simply the square root of the mean squared error but is preferable in our case because it is directly comparable to well logs.

ANN ARCHITECTURE

First, we assume a single hidden layer within our ANN,

SSSS

and test its performance with different numbers of neurons (1 through 12). The test consists of randomly selecting a number of wells for the training set and applying it in the rest of the set (test set). Because we are not certain about how many wells in the training set are optimal for each case, or the effect this might have, we make an educated guess, and run the test for 5 to 8 wells in the training set. All ANNs for this test are predicting VCL, and use shallow resistivity (RS), deep resistivity (RD), spontaneous potential (SP), quality factor (QF), and measured depth as inputs. All tests are run 100 times to get a statistically reliable assessment of their performance.

We are interested in identifying an ANN that yields high R^2, and low RMSD. We also want to avoid ANNs that show R^2 or RMSD with high standard deviation because that would mean that their performance is unreliable.

As shown in **Figure 2**, regardless of the number of wells in the training set, we consistently observe the RMSD of the test set increasing as the number of neurons in the hidden layer increases. This effect is more pronounced after 5 or 6 neurons. We also observe that the RMSD standard deviation grows dramatically after around 7 neurons. Regarding R^2 of the test set, we observe in **Figure 3** that it always shows a maximum at 3 or 4 neurons. With this information we conclude that if a single hidden layer is to be used, it should have between 3 and 5 neurons. The fact that we get acceptable results with a variety of ANN architectures confirms the robustness of using them for this purpose.
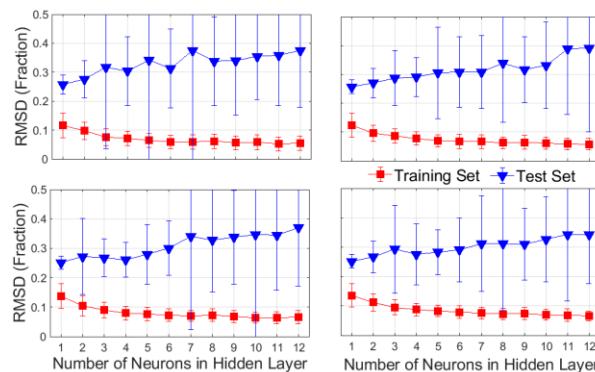

Figure 2: RMSD vs. number of neurons in a single hidden layer. Different plots show results for 5, 6, 7 and 8 wells in training set respectively. Training set in red squares, and test set in blue triangles.

We also evaluated the effect of the ANN depths by performing the same test with different numbers of

hidden layers, and different numbers of neurons in each. For these tests, for simplicity, we always use 6 wells in the training set.
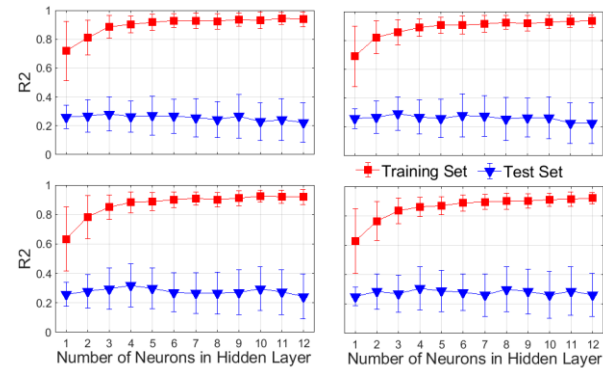

Figure 3: R^2 vs. number of neurons in a single hidden layer. Plots show results for 5, 6, 7 and 8 wells in training set respectively. Training set in red squares, and test set in blue triangles.

In **Figures 4 and 5** we show the RMSD and R^2 for ANNs with 2 hidden layers. We observe that having just one neuron in the first layer (top left plot), consistently yields a smaller RMSD than when more neurons are used in the first layer. Moreover, the standard deviation is significantly smaller, making it a most reliable ANN. Regarding R^2 for the test set, we observe that all ANNs have a consistent R^2 average around 0.3 except for the one neuron in the first layer case, whereas R^2 drops below 0.3 for the cases with 1 through 3 neurons in the second layer. Therefore, we conclude that the one neuron in the first hidden layer, and four neurons in the second hidden layer is the best option for two hidden layers.
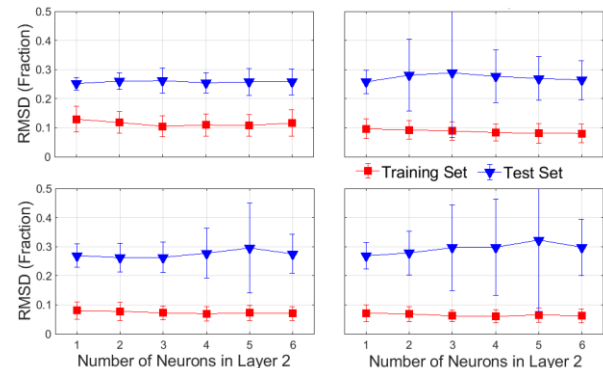

Figure 4: RMSD vs. number of neurons in layer 2. Different plots have different number of neurons in layer one: 1, 2, 3, and 4 respectively. Training set in red

SSSS
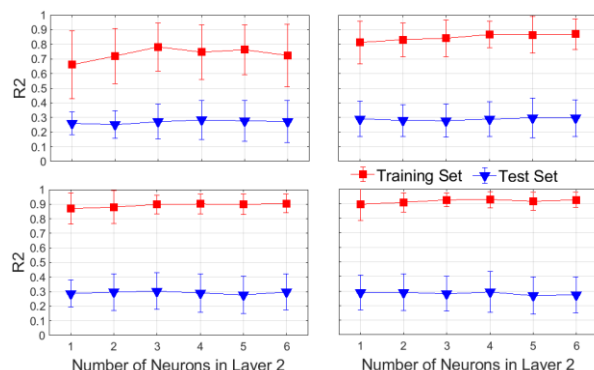
squares, and test set in blue triangles.

Figure 5: R² vs. number of neurons in layer 2. Different plots have different number of neurons in layer one: 1, 2, 3, and 4 respectively. Training set in red squares, and test set in blue triangles.

The performance of the ANN featuring two hidden layers with respectively 1 and 4 neurons vs. a single hidden layer case with 4 neurons, are comparable, with the largest difference being a smaller RMSD standard deviation for the 2-layer case. The extra layer with a single neuron has a stabilizing effect, which is recommended.
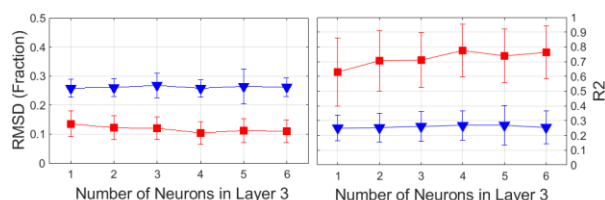
Figure 6: RMSD and R² vs. number of neurons in layer 3 for an ANN with 3 hidden layers. Layers 1 and 2 have only one neuron. Training set in red squares, and test set in blue triangles.

As a follow up of the stabilizing effect of an extra hidden layer with a single neuron, we attempted to add a third layer with another single neuron, but the results (**Figure 6**) show a drop in R² below 0.3. Hence, we do not recommend using a third hidden layer.

NUMBER OF WELLS IN THE TRAINING SET

It is generally assumed that adding consistent training data to a well-designed model improves the accuracy of the model up to a point right before the ANN is being over-trained. In a petrophysical context, the information contained in each well log may not be consistent between wells (difference of tool calibration, difference of cable speed causing a discrepancy in vertical resolution), and the amount of training data is limited. We therefore evaluated the optimal number of wells in a training set to minimize the error in the test set, that is, the number of wells above which adding more training wells only provides a marginal or no improvement in the adjusted correlation coefficient.

To perform this evaluation, we use an ANN with the architecture we previously found to be optimal among the ones we tested: 2 hidden layers with 1 and 4 neurons each. We randomly select a given number of wells to use as training data. Then we apply the ANN to several wells in the test set. We repeat this test for training sets of 1 to 15 wells (for legibility, the plots in this paper only show 1 to 12 wells). The number of wells is kept constant in the test set constant across all tests. Predicted properties are VCL (Volume of Clay), PHIE (Effective porosity), SW (Water saturation) and K (permeability). Inputs are RD (Deep resistivity), RS (Shallow resistivity), SP (Spontaneous Potential), GR_NORM (Normalized Gamma Ray), QF (Quality factor), and measured depth (DEPTH). The statistics for each simulation, R^2_{Adj} and RMSD (root-mean-square deviation) are then averaged over all the runs, and the results ($\overline{R^2_{adj}} \pm \sigma_{R^2_{adj}}$, $\overline{RMSD} \pm \sigma_{RMSD}$) plotted as a function of the number of wells in the training set.

As part of this evaluation, we also varied the type of input and output logs, starting with RS, RD, SP for inputs then adding GR_NORM, and later QF, and DEPTH, for each of the possible outputs VCL, PHIE, SW, K. In addition to these changes in input logs, we changed the way the resistivities were fed into the network, either as the variables or as the logarithm (base 10) of the variables.

Each simulation exhibits the same trends, regardless of the chosen output log and set of input logs: R² increases quickly when adding training wells until about 6 wells, then continues increasing but at a slower rate. Increasing the number of training wells also reduces the uncertainty (uncertainty around R² or RMSD, which also means uncertainty around the computed log values) until about 6 training wells (**Figure 7**). Using this strategy, maximum R² values are about 0.6 for VCL, PHIE, 0.45 for SW and 0.2 for K.
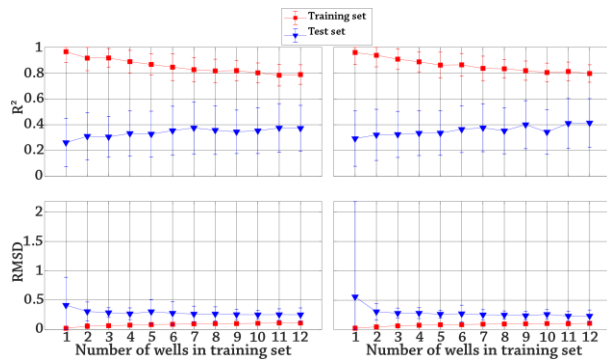
SSSS

Figure 7: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is VCL. Input logs are RS, RD, SP (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP (right column). The Neural Network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.

Adding different types of input logs also increases the accuracy of the neural networks, up to a certain limit. For instance, with 12 training wells, adding the Gamma Ray log moves the adjusted determination coefficient from about 0.4 to about 0.6 (**Figures 7 and 8**) However, adding the quality factor and depth does not further increase R^2 (**Figures 8, 9, and 10**), and may even decrease R^2, especially when the Gamma Ray is not used as an input (**Figures 11 through 32**).
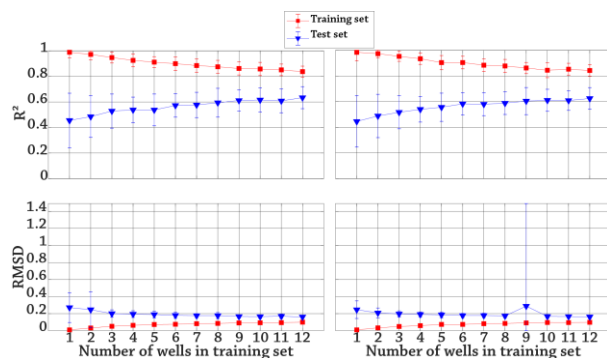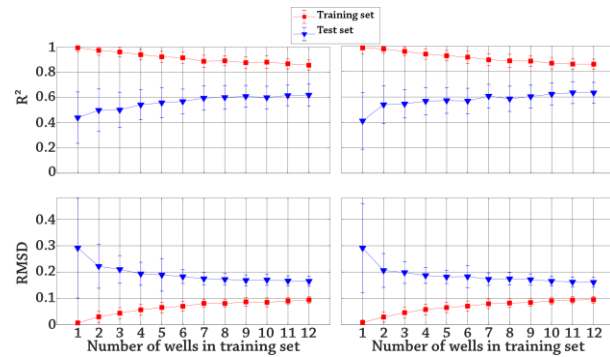

Figure 8: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is VCL. Input logs are RS, RD, SP, GR_NORM (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, GR_NORM (right column). The neural net

Figure 9: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is VCL. Input logs are RS, RD, SP, GR_NORM, QF (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, GR_NORM, QF (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
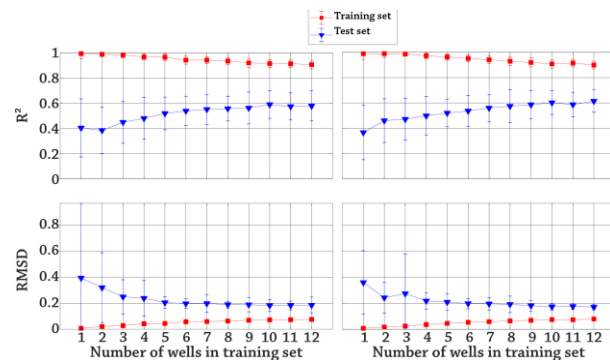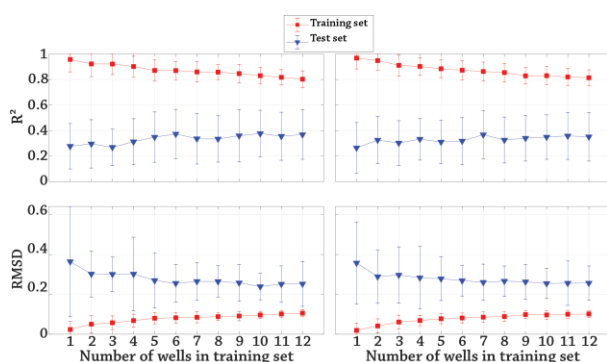

Figure 10: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is VCL. Input logs are RS, RD, SP, GR_NORM, QF, DEPTH (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, GR_NORM, QF, DEPTH (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.

Figure 11: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is VCL. Input logs are RS, RD, SP, QF (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, QF (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
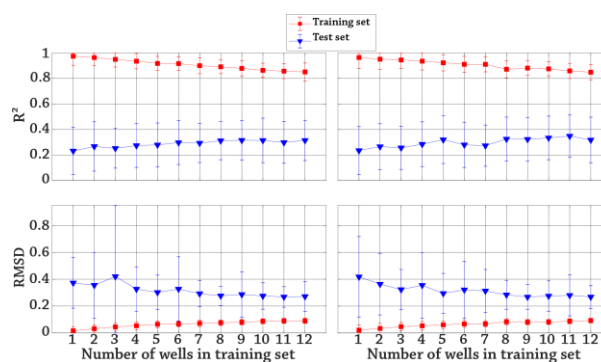

Figure 13: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is VCL. Input logs are RS, RD, SP, QF, DEPTH (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, QF, DEPTH (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
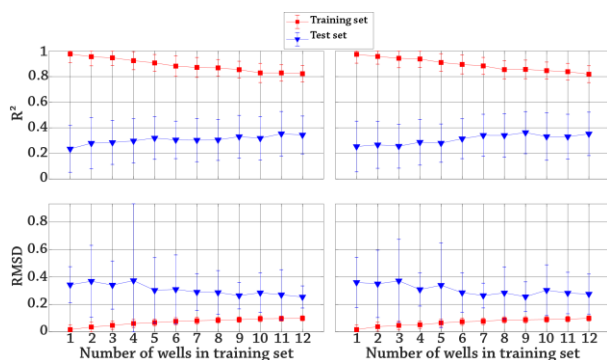

Figure 12: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is VCL. Input logs are RS, RD, SP, DEPTH (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, DEPTH (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
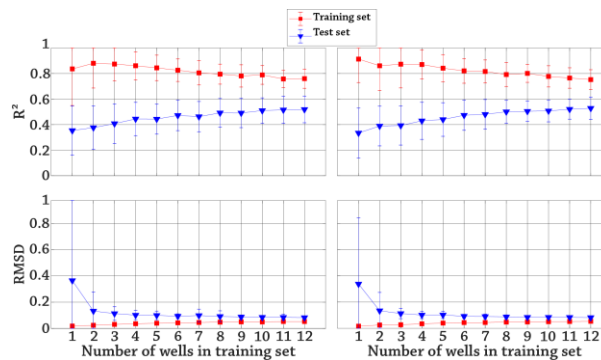

Figure 14: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is PHIE. Input logs are RS, RD, SP (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
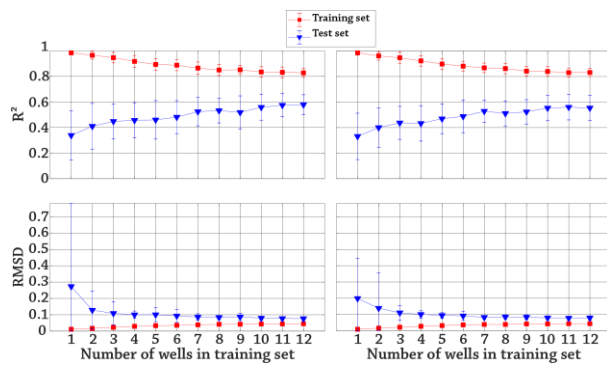
SSSS

7

Figure 15: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is PHIE. Input logs are RS, RD, SP, GR_NORM (left column), $\log_{10}$(RS), $\log_{10}$(RD), SP, GR_NORM (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.

Figure 17: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of number of training wells. Output log is PHIE. Input logs are RS, RD, SP, GR_NORM, QF, DEPTH (left column), $\log_{10}$(RS), $\log_{10}$(RD), SP, GR_NORM, QF, DEPTH (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
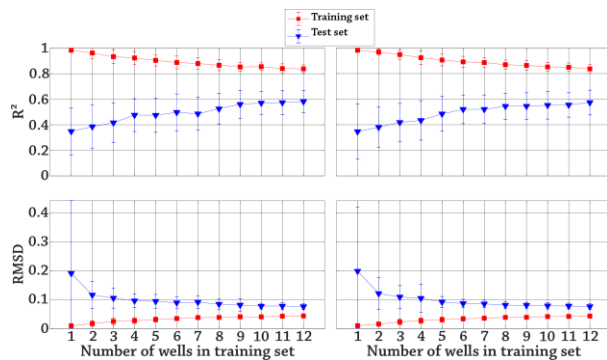

Figure 16: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is PHIE. Input logs are RS, RD, SP, GR_NORM, QF (left column), $\log_{10}$(RS), $\log_{10}$(RD), SP, GR_NORM, QF (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
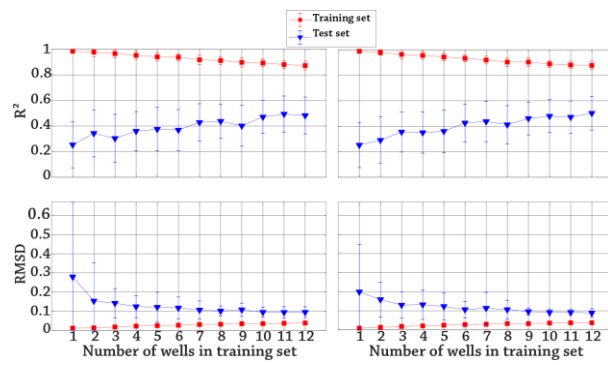

Figure 18: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is PHIE. Input logs are RS, RD, SP, QF (left column), $\log_{10}$(RS), $\log_{10}$(RD), SP, QF (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
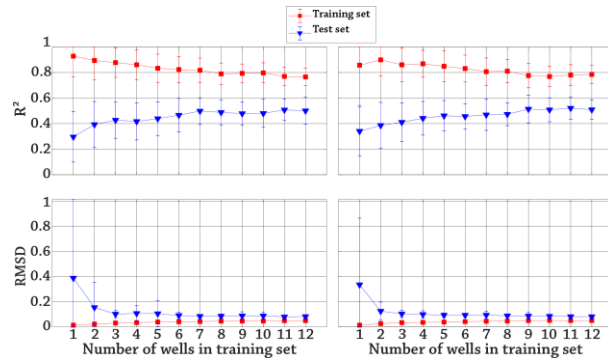
8

Figure 19: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is PHIE. Input logs are RS, RD, SP, DEPTH (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, DEPTH (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.

Figure 21: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is S_W. Input logs are RS, RD, SP (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
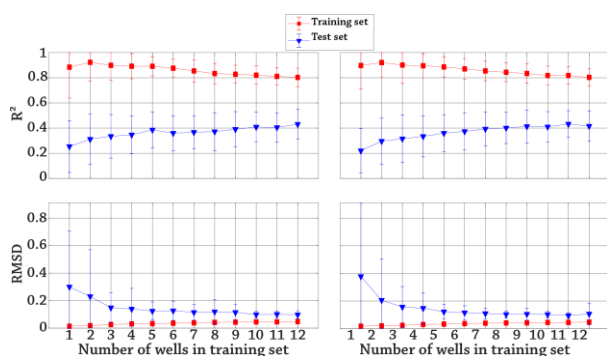

Figure 20: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is PHIE. Input logs are RS, RD, SP, QF, DEPTH (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, QF, DEPTH (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
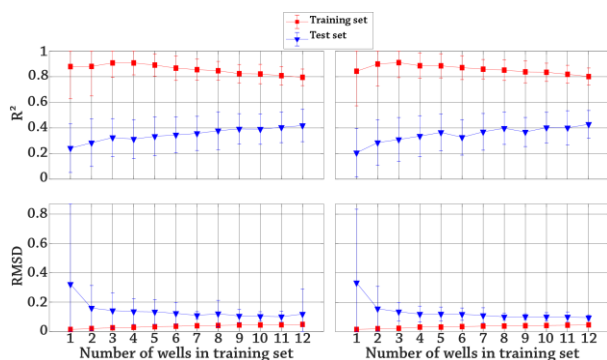

Figure 22: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is S_W. Input logs are RS, RD, SP, GR_NORM (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, GR_NORM (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
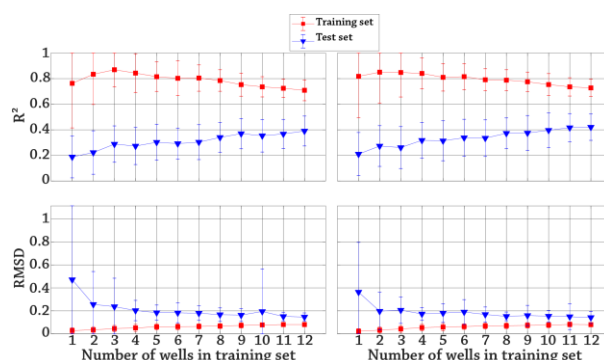
Figure 23: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is S_W. Input logs are RS, RD, SP, GR_NORM, QF (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, GR_NORM, QF (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
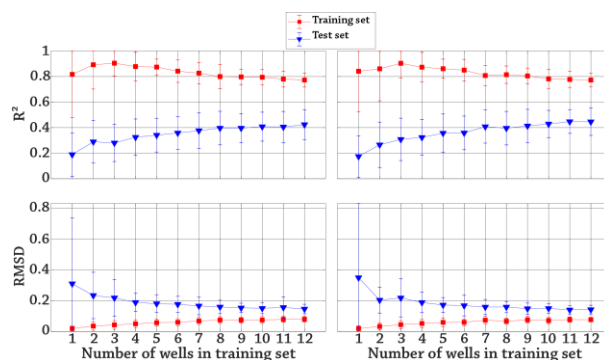

Figure 25: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is S_W. Input logs are RS, RD, SP, QF (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, QF (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.

Figure 24: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is S_W. Input logs are RS, RD, SP, GR_NORM, QF, DEPTH (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, GR_NORM, QF, DEPTH (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
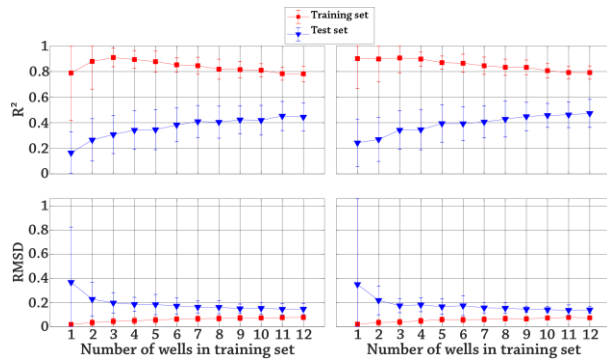

Figure 26: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is S_W. Input logs are RS, RD, SP, DEPTH (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, DEPTH (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
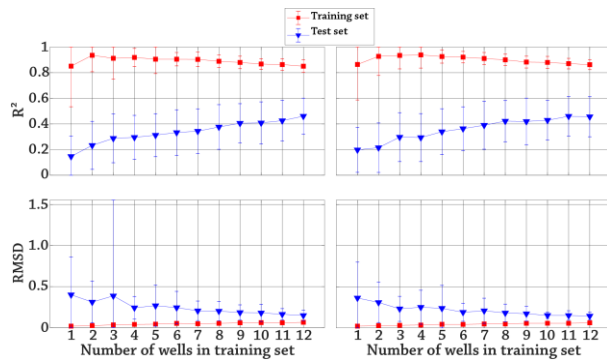
SSSS

10

Figure 27: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is S_W. Input logs are RS, RD, SP, QF, DEPTH (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, QF, DEPTH (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
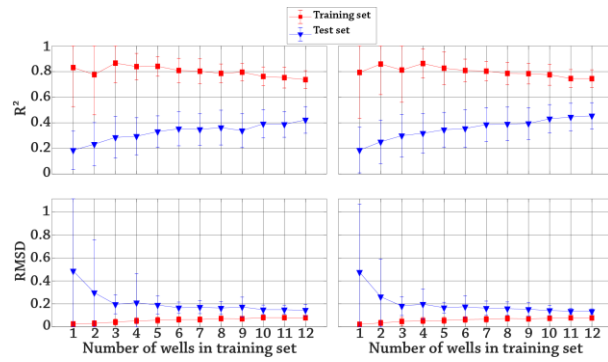

Figure 29: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is K. Input logs are RS, RD, SP, GR_NORM (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, GR_NORM (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.

Figure 28: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is K. Input logs are RS, RD, SP (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
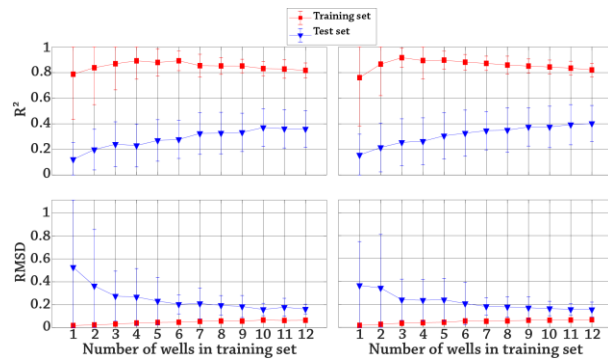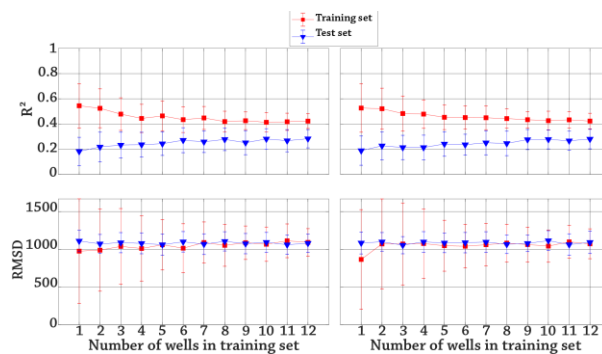

Figure 30: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is K. Input logs are RS, RD, SP, GR_NORM, QF (left column), $\log_{10}(RS)$, $\log_{10}(RD)$, SP, GR_NORM, QF (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
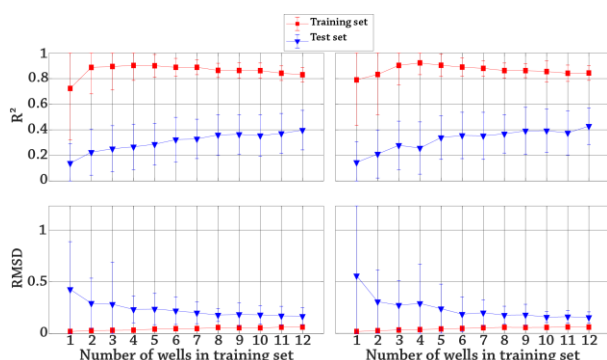
SSSS
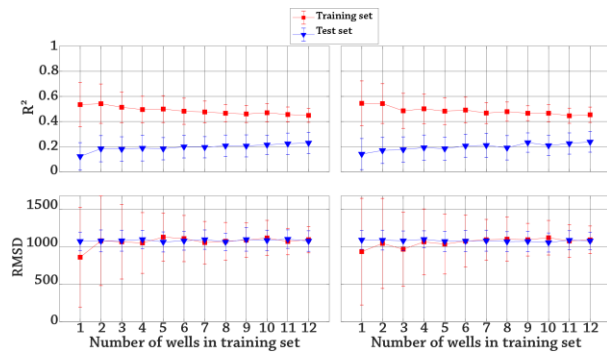
Figure 31: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is K. Input logs are RS, RD, SP, GR_NORM, QF, DEPTH (left column), $\log_{10}$(RS), $\log_{10}$(RD), SP, GR_NORM, QF, DEPTH (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
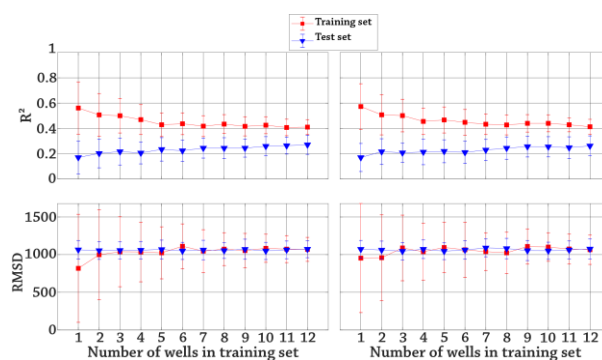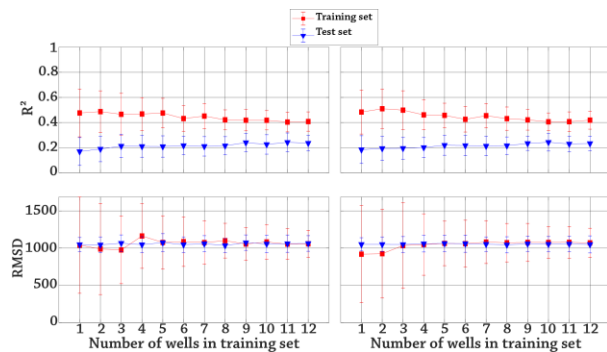

Figure 32: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is K. Input logs are RS, RD, SP, QF (left column), $\log_{10}$(RS), $\log_{10}$(RD), SP, QF (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
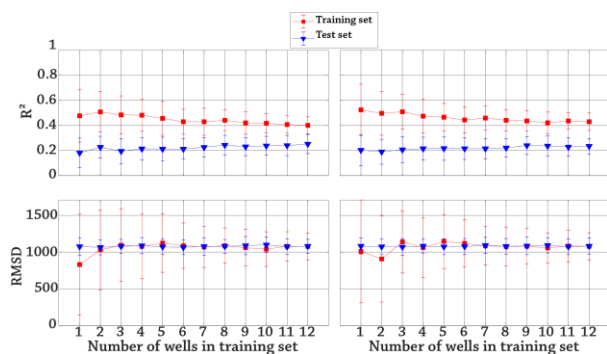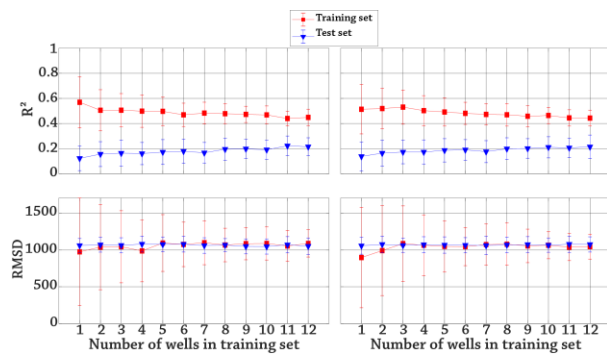

Figure 33: Adjusted coefficient of determination R^2

(upper row) and RMSD (lower row) as function of the number of training wells. Output log is K. Input logs are RS, RD, SP, DEPTH (left column), $\log_{10}$(RS), $\log_{10}$(RD), SP, DEPTH (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.
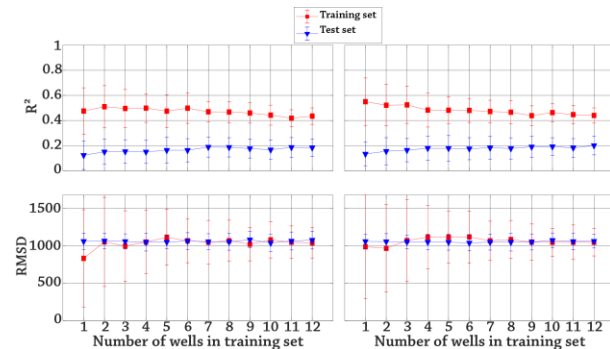
Figure 34: Adjusted coefficient of determination R^2 (upper row) and RMSD (lower row) as function of the number of training wells. Output log is K. Input logs are RS, RD, SP, QF, DEPTH (left column), $\log_{10}$(RS), $\log_{10}$(RD), SP, QF, DEPTH (right column). The neural network used 5 neurons and averages were obtained with 100 runs. Red squares represent the training set and blue triangles the test set.

INFLUENCE OF DISTANCE BETWEEN WELLS

Geological formations have a limited vertical and lateral extent, which influences the spatial variability of petrophysical properties, and, therefore the character and values of well logs. Assuming that our field is large enough to show the effects of that variability, the distance between training and application (test) wells may have an effect on the accuracy of ANNs predictions. To test this hypothesis, we proceed the following way: (1) Training logs are RS, RD, SP, QF, Measured Depth. The output logs are VCL, PHIE, SP, K. (2) The UTM32U Easting and Northing (X and Y) locations of each well head are extracted from the project well-log headers. (3) X and Y are converted to latitude and longitude. (4) The elevation Z of each point is obtained with Google Earth using latitude and longitude. The maximum elevation difference between 2 points is 62m. A relative elevation map of all wells used in the project is shown in **Figure 35**. (5) Four clusters of wells are computed from X, Y, Z using K-means clustering based on Euclidian distance (i.e. neglecting the curvature of the earth for such short distances). The color of the well names on **Figure 35** indicates to which cluster a well

SSSS

belongs. (6) For each cluster: (A) The current cluster is taken as reference. The number of training wells in the reference cluster is set to $N_{Tr_i} = \left\lfloor \frac{3}{4} \cdot N_i \right\rfloor$, where N_i is the number of wells in the cluster, and the number of test wells is $N_{Ts_i} = N_i - N_{Tr_i}$. (B) The number of test wells in the other 3 clusters is set to $N_{Ts_{j \neq i}} = \min\left(N_{Ts_I}, N_{j \neq i}\right)$ (i.e. same number of test wells in each other cluster as in the reference one, except if one cluster does not have enough wells for that). (C) For each of the $\binom{N_i}{N_{Tr_i}}$ combinations of training and test wells in the reference cluster: [i] We choose the $N_{Ts_{j \neq i}}$ test wells randomly in the other clusters. Note that if a test cluster has exactly $N_{Ts_{j \neq i}}$ wells, the test wells in this cluster will be the same for each training well set of the reference cluster. [ii] We compute the following centroids (average of well head coordinates): training wells, test wells in reference cluster, test wells in other clusters [iii] We compute the average root-mean-square-deviation (RMSD) and adjusted correlation coefficient (R_{adj}^2) between the reference curve and the simulated curve for each output variable separately, and also for each centroid (using all the data from the wells used to obtain that centroid). (7) We plot the results (R^2, RMSD for each test well, or $\overline{R^2}, \overline{RMSD}$ for each group of test wells) as function of distances to the reference cluster (to each well when testing individual wells, to the centroid of the test wells in the reference cluster, and to the centroid of the test wells in the other cluster taken all at once).
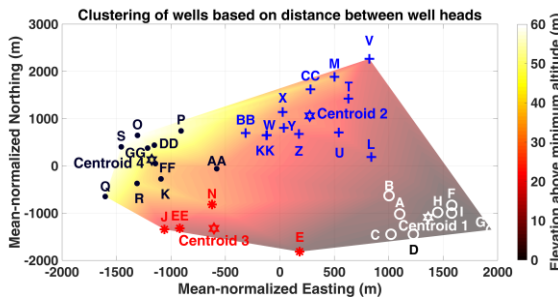
Figure 35: Landau field wells divided into 4 clusters with K-means clustering applied to Euclidian distance. Each cluster is shown in a different color (white, blue, red, black). UTM coordinates X and Y have been mean-normalized and the elevation of the well heads above the lowest point is shown as color shading where lighter colors represent higher elevation.
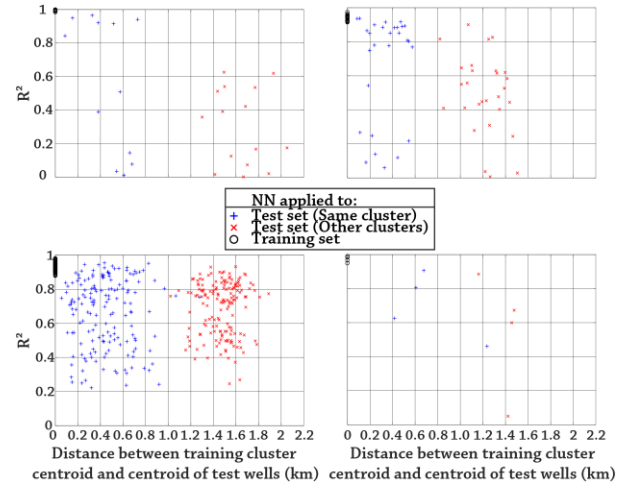
Figure 36: Adjusted coefficient of determination between reference VCL and ANN-computed VCL, as a function of distance between training and testing centroids; each sub-plot represents a different reference cluster. Output log is VCL, input logs are RS, RD, SP, GR_NORM, QF, DEPTH. Black circles are related to training wells (therefore always at distance 0), blue "plus" signs are related to test wells in the reference cluster, and red crosses to test wells in other clusters.
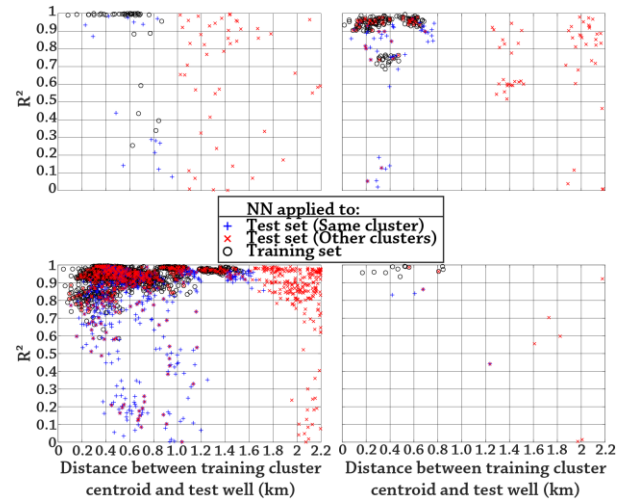
Figure 37: Adjusted coefficient of determination as a function of distance between training centroid and each well; each sub-plot represents a different reference cluster. Output log is VCL, input logs are RS, RD, SP, GR_NORM, QF, DEPTH. Black circles are training wells, blue "plus" signs are test wells in the reference cluster, and red crosses are test wells in other clusters.

For each output variable (VCL, PHIE, SW, K), and whether distances are from training centroid to each test well or from training centroid to training well centroid,

13

results are similar: there does not seem to any correlation between R^2 and distance. Only a slight decrease in maximum average R^2 is noticed between the test wells in the reference cluster and the test wells in other clusters (see **Figure 36**, distances from centroid to centroid).

When considering the distance between training centroid and each individual well, no trend at all may be noticed. The neural network performs very well in some wells far from the training centroid, and very poorly in other wells much closer to the training centroid (**Figure 37**).

USE OF ANN ACROSS FACIES

To test the performance of ANNs across lithologies, we train them in a specific facies, and use them to predict properties in different zones. We assume that the facies previously identified by the operating company are valid, and use the BNS facies to train our ANNs, since that was the zone of interest.

In **Figure 38** we can observe that not surprisingly, the facies that performs the best is BNS itself. However, it is interesting to note that PE performs almost just as well; especially considering that it was classified as a sandstone, while BNS was assumed a shaly carbonate. None the less, PE is neighboring the shaly carbonates, and it might still be transitional.
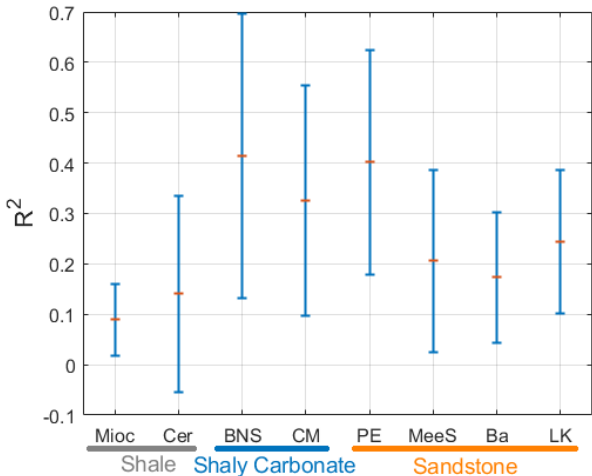
Figure 38: R^2 for the use of an ANN trained in the facies BNS, and applied to all possible facies. The order of the facies follows their depth order.

The facies CM, which is neighboring BNS also outperforms the rest of the facies. This behavior is expected because it is the only one that shares the shaly carbonate classification. All other facies performed

poorly; we conclude that ANNs are better used within specific lithologies.

WELL CLASSIFICATION

A self-organizing map (SOM) was used to classify wells according to their well logs within a single facies (BNS). Well logs used as input for all ANNs are RS, RD, SP, DEPTH, and QF, consistently with previous ANNs. For each well we calculate all possible cross correlations across logs, and the auto-correlation of each log and its derivatives at 1, 2, and 3 sampling steps. The classification, whose output is shown in **Table 2**, is performed with three classes.

Table 2: Output of SOM classification of Landau wells

Class 1 (6 wells)	GG, FF, R, A, D, KK
Class 2 (10 wells)	K, DD, J, EE, BB, AA, B, E, F, I
Class 3 (14 wells)	L, M, O, Q, T, U, V, N, Z, W, X, Y, G, H

We compare ANNs trained exclusively in wells from the same class with ANNs trained using all wells. To this end, we use the ANN with 2 hidden layers featuring respectively 1 and 4 neurons. Given that each class has different number of members, to have a fair comparison, we restrict the test set size for the ANN having all wells as available. For classes 2 and 3 we keep the standard of 6 wells in the training set, but for class 1, we use 4 because the entire class only has 6 wells in total.

Figure 39: RMSD and R^2 for 3 different classes. Each class is compared to a base case using all wells as

SSSS

possible members of training or test set, but restricting the size of the set to match the size possibilities of each class. The training set for class 1 has 4 wells, while that number is 6 for classes 2 and 3.

Figure 39 shows that ANNs trained exclusively in class 1 perform poorly compared to the base ANNs (using all available wells), while ANNs for classes 2 and 3 perform significantly better, reaching R^2 comparable to the training set. 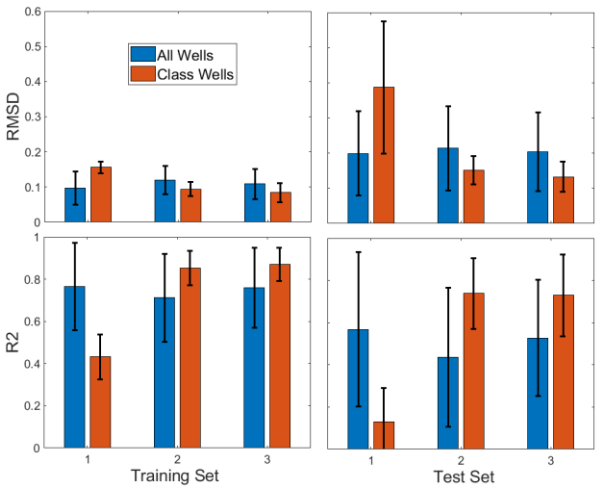Classes 2 and 3 encompass 80% of the total number of wells. Considering that even in the training set, class 1 vastly underperforms the other classes, we can state that it is composed of the wells that have the most difficult relationships between logs, and properties.

It is important to point out that the SOM classification does not show any correlation to geographical location, or any evident way of grouping wells. The features identified by the SOM are buried within the logs and are an important step for successful ANN application.

Next, we take a closer look at the specific results and compare predictions yielded by an ANN using all available wells, which we will call BANN (from Base Case ANN) with predictions using a class specific ANN, which we call CANN.

In **Figure 40** we display some examples of wells from the poorly performing class 1. We observe that neither the CANN nor BANN perform acceptably in any case. Some of the information we are missing might be in some of the well logs we are not using for the training, i.e., Gamma Ray. We can only acknowledge that the SOM was capable of grouping these complicated cases together.

Figure 41 shows examples of class 2 wells. In well I, we observe some improvements from the CANN toward the top, and where VCL varies very rapidly, however, we could state that this is an example where BANN and CANN perform similarly. In well EE, we observe a very similar case where the BANN prediction does not reproduce the extreme values properly. It almost seems artificially constrained to between 0.1 and 0.9, but it does not flatten out to any of these values. The CANN predictions show a significant improvement, although there is a section toward the top, where the data closest to 1 is still being missed.

Figure 40: Sample wells from **class 1**. On the left-hand side, VCL predictions for an ANN trained using a pool from all available wells; on the right-hand side ANNs trained exclusively in wells from **class 1**.

Well I

Well L

Well EE

Well H

SSSS

Figure 41: Sample wells from **class 2**. On the left-hand side, VCL predictions for an ANN trained using a pool from all available wells; on the right-hand side, ANNs trained exclusively in wells from class 2.

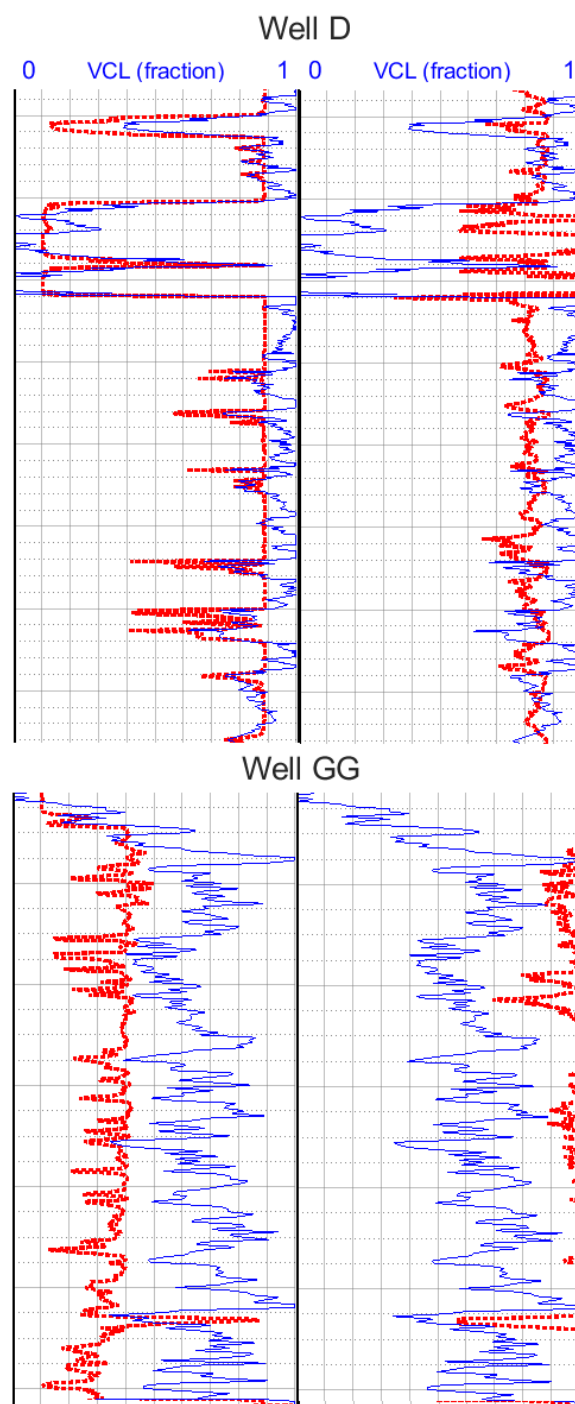**Figure 42**: Sample wells from **class 3**. In the left, VCL predictions for an ANN trained using a pool from all available wells; in the right ANNs trained exclusively in wells from **class 3**.

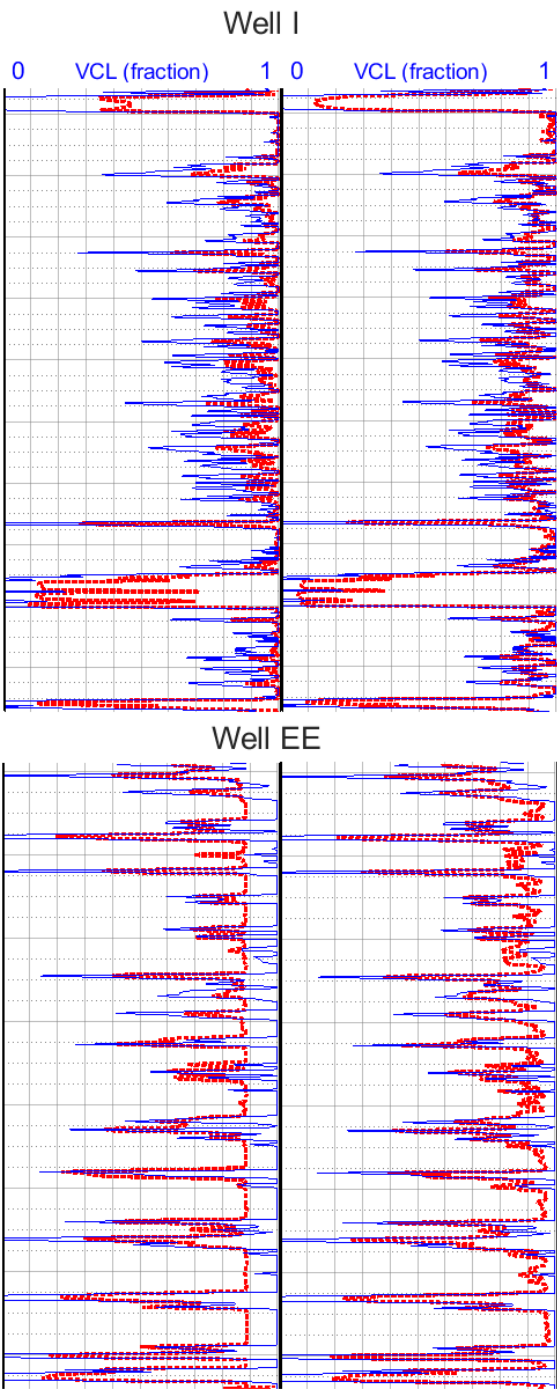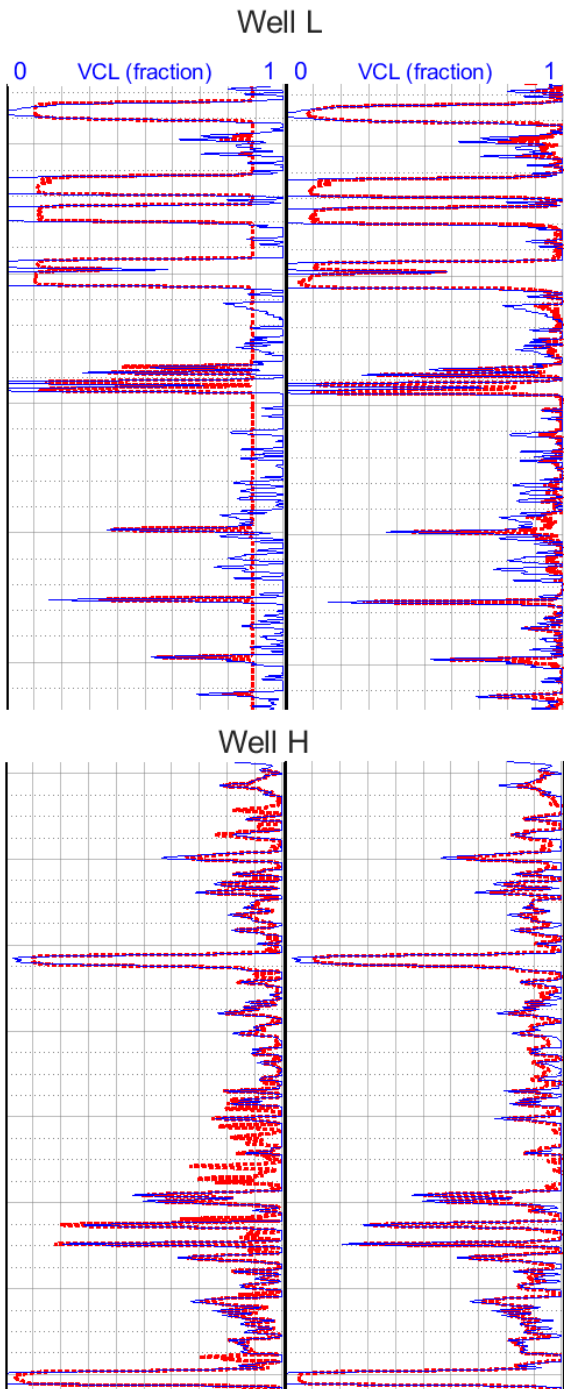In **Figure 42**, we observe that for well L, similarly to well EE from class 2, BANN results do not reach the extreme values of 1, or 0. The CANN predictions do

16

reach the extreme values, although this network still misses some of the subtleties when VCL changes rapidly. In well H, we observe that the BANN can reach the 0 and 1 values correctly, however, this well also exhibits a section toward the bottom with some false positive features. On the other hand, CANN produces a near perfect match.

CONCLUSIONS

The observations we made after running ANNs for which we varied parameters such as number of training wells, number of hidden layers and neurons in each layer, and type of input logs, show that (1) ANNs are a robust enough method to perform acceptably with a variety of architectures. In our field case, we identified the 2 hidden layers with 1 and 4 neurons respectively, as an optimal choice; (2) training and applying ANNs in well groups classified by a self-organizing map (SOM) yields higher accuracy; (3) using more than 6 training wells only marginally improves accuracy and repeatability; (4) the best set of input logs includes spontaneous potential, shallow- and deep-resistivity, and gamma ray. Quality factor and depth do not significantly improve the results but adding quality factor and depth does lead to quality improvement if gamma ray is not present; (5) ANNs work best when applied to the same facies. Cross-facies ANNs show poor accuracy.

LIST OF ACRONYMS

AI: Artificial Intelligence
ANN: Artificial Neural Network
BANN: Base Case Artificial Neural Network
CANN: Class Specific Artificial Neural Network
CNN: Convolutional Neural Networks
GRNN: General Regression Neural Network
MSE: Mean Squared Error
R^2: Coefficient of Determination

RMSD: Root Mean Squared Deviation
SOM: Self Organizing Map
SVM: Support Vector Machines

LIST OF SYMBOLS

DEPTH: Measured Depth
GR: Gamma Ray
K: Permeability
PHIE: Effective Porosity
$d\rho$: Bulk density correction
QF: Quality Factor
RD: Deep Resistivity
RS: Shallow Resistivity
SP: Spontaneous Potential
SW: Water Saturation
VCL: Volume of Clay Fraction
$\lfloor X \rfloor$: Integer part of X such that $\lfloor X \rfloor \leq X \, \forall X \in \mathbb{R}$ ("floor" function)
σ_X: Standard deviation of X
$\overline{X}$: Arithmetic average of X

REFERENCES

Bhatt, A., and Helle, H. B., 2002, Determination of facies from well logs using modular neural networks, Petroleum Geoscience, Vol. 8, 217-228.

Bhatt, A., and Helle, H. B., 2002, Committee neural networks for porosity and permeability prediction from well logs, Geophysical Prospecting, Vol. 50, 645-660.

Böcker J (2015): Petroleum system and thermal history of the Upper Rhine Graben : implications from organic geochemical analyses, oil-source rock correlations and numerical modelling. PhD Thesis RWTH Aachen, HBZ: HT018945941.

Fung, C.C., Wong, K.W., Eren, H. and Charlebois, R., 1995, Lithology classification using self-organizing map, IEEE International Conference on Neural Networks, Perth, Western Australia, November 27 to December 1, Vol. 1, 526-531.

Fung, C.C., Wong, K.W., and Eren, H., 1997, Modular artificial neural network for prediction of petrophysical properties from well log data, IEEE transactions on Instrumentation and Measurement, Vol. 46, 1295-1299.

Gholami, R., Shahraki, A. R., and Paghaleh, M. J., 2012, Prediction of hydrocarbon reservoirs permeability using

SSSS

17

support vector machine, Mathematical Problems in Engineering, Vol. 2012, 1-18, DOI: 10.1155/2012/670723.

Imamverdiyev, Y., and Sukhostat, L, 2019, Lithological facies classification using deep convolutional neural network, Journal of Petroleum Science and Engineering, Vol. 174, 216-228, DOI: 10.1016.j.petrol.2018.11.023.

Jodocy, M. and Stober, I., 2011, Kurzer Überblick über die Formationen des Tertiärs im Oberrheingraben im Hinblick auf hydrothermale Nutzungen, Report for the Regierungspräsidium Freiburg.

Khandelwal, M., and Singh, T.N., 2010, Artificial neural networks as a valuable tool for well log interpretation, Petroleum Science and Technology, Vol, 28, 1381-1393. DOI: 10.1080/10916460903030482.

Lim, J., 2005, Reservoir properties determination using fuzzy logic and neural networks from well data in offshore Korea, Journal of Petroleum Science and Engineering, Vol. 49, 182-192. DOI: 10.1016/j.petrol.2005.05.005.

Rogers, S. J., Fang, J. H., Karr, C. L., and Stanley, D. A., 1992, Determination of lithology from well logs using a neural network, The American Association of Petroleum Geologists, Vol. 76, No. 5, 731-739.

Rogers, S.J., Chen, H. C., Kopaska-Merkel, D. C., and Fang, J. H., 1995, Predicting permeability from porosity using artificial neural networks, The American Association of Petroleum Geologists, Vol. 79, No. 12, 1786-1797.

Rolon, L., Mohaghegh, S. D., Ameri, S., Gaskari, R., and McDaniel, B., 2009, Using artificial neural networks to generate synthetic well logs, Journal of Natural Gas Science and Engineering, Vol. 1, 118-133. DOI: 10.1016/j.jngse.2009.08.003.

Strobel J: Application of Artificial Intelligence to the Petrophysical Interpretation of Logs. DGMK annual conference 2019, Celle (Germany).

Sumen, M., Tiwari, R.K., and Kümpel, H.J., 2007, Neural network modeling and classification of lithofacies using well log data: a case study from KBT borehole site, Geophysical Journal International, Vol. 169, 733-746.

ABOUT THE AUTHORS

Aymeric-Pierre Peyret received an MSE in Multi-Disciplinary Engineering from Ecole Centrale de Lille (France), an MSE in Petroleum Engineering from The University of Texas at Austin in 2007, and a PhD in Geological Sciences from the University of Texas at Austin in 2011. From 2011 to 2018 he worked as a studies Petrophysicist at Shell. Since 2019 he has been a Research Fellow in the Formation Evaluation joint Industry Research Consortium, in the Hildebrand Department of Petroleum and Geosystems at University of Texas at Austin, and has been working on complex formation evaluation and applications of machine learning to petrophysics.

Joaquín Ambía received a PhD in Computational Bio-Physics at the University of Houston in 2011. From 2011 to 2014 he worked on the Rosetta software for computational modeling of structural biology systems. Since 2014 he has been a research associate and software developer in the Formation Evaluation joint Industry Research Consortium, in the Hildebrand Department of Petroleum and Geosystems at University of Texas at Austin. He contributes to the 3D UTAPWeLS well logging simulation, and earth model integration software. He is responsible for incorporating newly developed algorithms and developing new algorithms and methods himself.

Joachim Strobel received a Diploma in Geology & Crystallography from the University of Munich and from the University of Swansea. He worked for Schlumberger as field engineer and petrophysicist from 1987 to 1999. He then joined the German operator RWE-DEA and managed the Formation Evaluation group there. In 2012 he joined Wintershall to head a mixed group of Reservoir Engineers and Petrophysicists before becoming the company's lead petrophysicist. While Joachim enjoys all facets of petrophysics, he is specialized in harmonizing saturation models across disciplines and explaining uncommon log responses through numerical simulations.

Carlos Torres-Verdín received a Ph.D. in Engineering Geoscience from the University of California at Berkeley in 1991. During 1991-1997, he held the position of Research Scientist with Schlumberger-Doll Research. From 1997-1999, he was Reservoir Specialist and Technology Champion with YPF (Buenos Aires, Argentina). Since 1999, Dr. Torres-Verdín has been affiliated with the Hildebrand Department of Petroleum and Geosystems of the University of Texas at Austin,

SSSS

where he is currently Full Professor, and holds the Brian James Jennings Memorial Endowed Chair in Petroleum and Geosystems Engineering. Dr. Torres-Verdín is the founder and director of the Research Consortium on Formation Evaluation at the University of Texas at Austin. He is the recipient of the Cockrell School of Engineering's 2016-2017 Lockheed Martin Aeronautics Company Award for Excellence in Engineering Teaching, the 2014 Gold Medal for Technical Achievement from the SPWLA, the 2008 Formation Evaluation Award from the SPE, and the 2006 Distinguished Technical Achievement Award from the SPWLA. He is a Distinguished Member of the Society of Petroleum Engineers (SPE), Honorary Member of the Society of Exploration Geophysicists (SEG), and receiver of the Conrad Schlumberger Award from the European Association of Geoscientists and Engineers (EAGE).

SSSS

APPENDIX 1

Variable well-log inventory.

Well	Measured logs and core data												Interpreted logs				
	BS	Cali	GR	RHOB	NPHI, NEUT	RD, RS, SP	MINV, MNOR	MLL	Sigma, TPHI	PHI_Core	K_Core	RHO_Gcore	Vclay	PHIT	PHIE	SW, SXO	K
A	x	x	x			x							x				
B	x	x	x			x	x	x	x				x	x	x	x	x
C	x		x			x		x	x				x	x	x	x	x
D	x	x	x			x	x	x	x				x	x	x	x	x
E	x	x	x			x	x	x	x				x	x	x	x	x
F	x	x	x			x	x	x	x				x	x	x	x	x
G	x	x	x			x	x	x	x				x	x	x	x	x
H	x	x	x			x	x	x	x				x	x	x	x	x
I	x	x	x			x	x	x	x				x	x	x	x	x
J	x	x	x			x	x	x					x	x	x	x	x
K	x	x				x	x						x	x	x	x	
L		x	x			x	x		x				x	x	x	x	x
M	x	x	x			x	x						x	x	x	x	x
N	x	x	x			x	x						x	x	x	x	x
O	x	x	x		x	x	x						x	x	x	x	x
P	x					x							x				
Q	x	x	x			x	x						x	x	x	x	
R	x	x				x							x				
S	x		x	x		x	x						x	x	x	x	x
T	x	x	x			x			x				x	x	x	x	x
U	x	x	x			x	x						x	x	x	x	x
V	x	x	x			x	x			x	x		x	x	x	x	x
W	x	x	x			x	x						x	x	x	x	x
X	x	x	x			x	x		x				x	x	x	x	x
Y	x	x	x			x	x						x				
Z	x	x	x	x		x	x						x	x	x	x	x
AA	x	x	x			x	x						x	x	x	x	x
BB	x	x	x	x		x	x		x				x	x	x	x	x
CC	x																
DD	x	x				x	x						x	x			
EE	x	x	x	x		x	x						x	x	x	x	x
FF	x	x				x	x						x				
GG	x	x	x	x		x							x	x	x	x	x
KK	x	x	x	x	x	x	x			x	x	x	x	x	x	x	X