applied sciences

MDPI

*Article*

# A Novel Empirical and Deep Ensemble Super Learning Approach in Predicting Reservoir Wettability via Well Logs

**Daniel Asante Otchere [1,2,*], Mohammed Abdalla Ayoub Mohammed [1,2], Tarek Omar Arbi Ganat [3], Raoof Gholami [4] and Zulkifli Merican Aljunid Merican [5]**

[1] Department of Petroleum Engineering, Universiti Teknologi PETRONAS, Seri Iskandar 32610, Perak Darul Ridzuan, Malaysia; abdalla.ayoub@utp.edu.my
[2] Centre of Research for Enhanced Oil Recovery, Universiti Teknologi PETRONAS, Seri Iskandar 32610, Perak Darul Ridzuan, Malaysia
[3] Department of Petroleum and Chemical Engineering, Sultan Qaboos University, Muscat 123, Oman; t.ganat@squ.edu.om
[4] Department of Energy Resources, University of Stavanger, Kitty Kielland's house Rennebergstien 30, 4021 Stavanger, Norway; raoof.gholami@uis.no
[5] Department of Fundamental & Applied Science, Universiti Teknologi PETRONAS, Seri Iskandar 32610, Perak Darul Ridzuan, Malaysia; zulkifli.aljunid@utp.edu.my
* Correspondence: ascotjnr@yahoo.com

**Abstract:** Accurately measuring wettability is of the utmost importance because it influences several reservoir parameters while also impacting reservoir potential, recovery, development, and management plan. As such, this study proposes a new formulated mathematical model based on the correlation between the Amott-USBM wettability measurement and field NMR $T_2$LM log. The exponential relationship based on the existence of immiscible fluids in the pore space had a correlation coefficient of 0.95. Earlier studies on laboratory core wettability measurements using $T_2$ distribution as a function of increasing water saturation were modified to include $T_2$LM field data. Based on the trends observed, water-wet and oil-wet conditions were qualitatively identified. Using the mean $T_2$LM for the intervals of interest and the formulated mathematical formula, the various wetting conditions in existence were quantitatively measured. Results of this agreed with the various core wettability measurements used to develop the mathematical equation. The results expressed the validity of the mathematical equation to characterise wettability at the field scale. With the cost of running NMR logs not favourable, and hence not always run, a deep ensemble super learner was employed to establish a relationship between NMR $T_2$LM and wireline logs. This model is based on the architecture of a deep learning model and the theoretical background of ensemble models due to their reported superiority. The super learner was developed using nine ensemble models as base learners. The performance of nine ensemble models was compared to the deep ensemble super learner. Based on the RMSE, $R^2$, MAE, MAPD and MPD the deep ensemble super learner greatly outperformed the base learners. This indicates that the deep ensemble super learner can be used to predict NMR $T_2$LM in the field. By applying the methodology and mathematical formula proposed in this study, the wettability of reservoirs can be accurately characterised as illustrated in the field deployment.

**Keywords:** surface wettability; empirical formula; NMR characterisation; artificial intelligence; water saturation; deep neural network



**Citation:** Otchere, D.A.; Abdalla Ayoub Mohammed, M.; Ganat, T.O.A.; Gholami, R.; Aljunid Merican, Z.M. A Novel Empirical and Deep Ensemble Super Learning Approach in Predicting Reservoir Wettability via Well Logs. *Appl. Sci.* **2022**, *12*, 2942. https://doi.org/10.3390/app12062942

Academic Editor: Sohrab Zendehboudi

Received: 21 December 2021
Accepted: 25 January 2022
Published: 14 March 2022



## 1. Introduction

Wettability can be defined as the intermolecular interaction between immiscible fluids and a rock's pore space [1]. This phenomenon leads to a relative affinity of one fluid to smear itself on the grain surface, whereas the non-wetting fluid occupies the central space. The importance of wettability cannot be downplayed as it directly influences capillary

pressure, relative permeability, effective porosity, electrical properties, simulated tertiary oil recovery, and residual oil saturation [2,3]. Wettability is also an essential property upon which reservoir potential, recovery, subsequent plans of development, and reservoir management depend [4,5].

However, many different methods have been developed to quantify the wettability of a system. They include quantitative methods—contact angles, imbibition, and forced displacement (Amott), United States Bureau of Mines (USBM) wettability method, and Nuclear Magnetic Resonance (NMR). The unique criteria in determining the degree of water or oil wetness are shown in Table 1 below. The qualitative methods include imbibition rates, microscope examination, flotation, glass slide method, relative permeability curves, permeability/saturation relationships, permeability curves, permeability/saturation relationships, capillary pressure curves, displacement capillary pressure, reservoir logs, nuclear magnetic resonance, and dye adsorption [5].

Although no single accepted method exists, three industry-standard quantitative methods are generally used: contact-angle measurement, the Amott method (imbibition and forced displacement), and the USBM method. For elementary and accurate measurements, the contact angle method can quantitatively determine the fundamental wetting condition of a specific flat solid surface. However, aside from ageing and preparing samples, uncertainty in measurements arises due to surface roughness and geometrical imperfection. However, this method is not reliable and applicable in complex geometries, heterogeneous reservoirs, and measuring the wettability of multiphase fluid [5].

The Amott and USBM methods are robust industry laboratory methods that measure the average wettability of core plugs. These methods, notwithstanding their accuracy and reliability, are not applicable downhole. In many field applications, wettability is classified as water- or oil-wet, masking the complexity of wettability due to these extreme simplifications [6]. The various degrees of wetting, commonly stated as mixed wettability, were first developed by Salathiel [7]. This wetting condition occurs when the macropores and micropores are wetted by oil and water, respectively. Through their extensive research, Fleury and Deflandre [8] introduced various forms of mixed wettability at the pore level. They defined this phenomenon as the rock surface area having parts of it being wetted by different fluids.

**Table 1.** Summary of the different quantitative wettability measurement techniques identifying key advantages, limitations, and observations.

| Wettability Techniques | Advantages | Disadvantages | Measurement Principle |
|---|---|---|---|
| Contact Angle [4] | Accurate | Surface heterogeneity can affect reading | Involves direct observation and measurement of wetting angles |
| | Simple | Cannot provide information regarding organic coatings<br>Impurities can affect reading.<br>Not applied downhole<br>Liquid drop can form many stable contact angles. | Uses a petrographic microscope or SEM fitted with a stage |
| Amott/Amott–Harvey [9] | Reliable | Not effective to measure neutral wettability | Measure the macroscopic mean wettability |
| | Most reliable and accurate, especially in the neutral wettability region | Insensitive near-neutral wettability<br>Inability to differentiate between important degrees of strong water-wetness<br>Not applied downhole<br>No validity as an absolute measurement | Measures the volume of naturally imbibed and forcefully displaced fluids |
| United States Bureau of Mines [10] | Reliable | No validity as an absolute measurement | Measure the macroscopic mean wettability |
| | It is sensitive to near-neutral wettability. | Cannot determine mixed wettability | Calculates the work required to imbibe and displace fluids |
| | It is sensitive to near-neutral wettability. | It can only be measured on plug-size samples. | Areas under the brine- and oil-drive curves are used for calculations. |
| Nuclear magnetic resonance [11] | Ability to study fractional wettability | Not applied downhole | Manipulates the wetting fluids' hydrogen protons to measure their relaxation times. |
| | Distinguishes between macro-and micro-pores | Requires complex and expensive machinery | Changes in longitudinal relaxation time due to its interaction with the pore surface |
| | Does not intervene in the fluid saturation<br>Wettability can be measured at any given saturation. | | |

However, the measurement and predictability of the various types of wettability are complicated and challenging. Many researchers have constantly been trying new ideas and methods in this field. During the past two decades, many researchers have been inclined to predict reservoir wettability with nuclear magnetic resonance (NMR) data at the forefront. This inclination to NMR is due to its non-invasive nature. It is used to measure fluid and solid surface interaction sensitivity and other rock-fluid properties like porosity, pore size distribution, relative permeability, and fluid saturation [12,13]. The NMR logging tool can quantify the different wettability conditions and is applicable downhole, but it cannot be run in all wells due to economic conditions. This scarceness of NMR logs is unlike conventional wireline logs, which are almost always run in all wells, especially the reservoir section [14]. These conventional wireline logs give essential information about the characteristics of formations in subsurface layers, including physical, mineralogical, and morphological properties. The information provided can be directly or indirectly related to wettability. However, a systematic methodology has not yet been developed to use these logs to quantitatively measure wettability due to complex and non-linear relationships between the wireline logs and wettability.

Machine learning algorithms have proven successful in engineering activities ranging from civil, electrical, and mechanical to mining and petroleum disciplines. The main advantages of these machine learning models are their ability to work with limited data, robustness against noise, and high efficiency in recognising the non-linear relationships between the input and output data [15].

### 1.1. Research Motivation

The application of laboratory wettability measurements downhole remains a challenge and not close to being well-established in the industry due to heterogeneity and complex geological structures [16]. NMR logging and well testing have proven to successfully measure important reservoir properties, including wettability [5]. However, these methods are costly and, in relatively low oil price regimes, do not seem feasible to run in each well, giving conventional wireline logs ubiquitous in all wells.

Conventional wireline logs have a complex and non-linear relationship with surface wettability, but the application of advanced machine learning algorithms may help establish a relationship on a field scale. Many laboratory methods such as spontaneous imbibition, Amott–Harvey, and contact angle measurements have been used to determine the surface wettability of reservoirs. However, these methods are often incapable of providing reliable results once tested in the field due to the heterogeneity of the reservoir and complex geological structures. As a result, few field scale approaches have been proposed based on well testing and NMR log data analysis, but they are also expensive to acquire. Using wireline logs to help predict this essential reservoir parameter is relatively cheaper and time-efficient than any laboratory wettability method and NMR logging. As such, the need for a cheap and accurate field scale approach for quantitative surface wettability determination has been of interest considering current oil prices.

Therefore, this research aims to integrate empirical and machine learning techniques in proposing a field scale, time-efficient, and relatively inexpensive alternative approach to determine the wettability of reservoir rocks using wireline logs.

### 1.2. Background on Wettability Measurements

Quantifying reservoir wettability accurately has gained attention for some time now and has increased research to improve its understanding. The most important reason for determining the surface wettability of reservoir rocks is to increase and improve productivity by employing a suitable Enhanced Oil Recovery (EOR) technique. To decide about the type of EOR technique to employ to achieve maximum reservoir performance, detailed understanding, and prediction of surface wettability and the parameters that affect it is crucial. This statement means that parameters such as mineral composition, morphology, pore structures, surface roughness, salinity, and surface contaminations must be quantified

at the lab and field scales. These parameters are of interest because a wrongly assumed parameter can have severe permanent damage to the reservoir and subsequently affect production, particularly water flooding and EOR techniques [16–18].

For many decades, researchers have come up with several empirical models that best try to postulate a correlation between well log data and water saturation for wettability measurements, with most being a modification or extension of Archie's equation [19]:

$$S_W{}^n = \frac{R_W}{\varphi^m R_t} \quad (1)$$

where Sw is the water saturation, Rw is the water phase resistivity, φ represents porosity, $R_t$ is the sample's resistivity, and m, the cementation factor. Based on Archie's equation which the petroleum industry has widely accepted since 1940, wettability is affected by reservoir properties like fluid flow, pore space geometry, distribution, and connectivity. The fluid phase connectivity is also relied upon by resistivity; hence wettability directly has a partial effect on the resistivity parameter n. The importance of resistivity to wettability cannot be overlooked because of its wide depth of investigation range and direct influence on saturation measurements [20].

Many researchers have come up with several methods over the last couple of decades to quantitatively and qualitatively measure surface wettability accurately in the laboratory and translate such results onto the field level. Wettability can be therefore be measured by several methods such as contact angles [4], Amott–Harvey [9], US Bureau of Mines (USBM) [10], electrical resistivity [21] and Nuclear Magnetic Resonance [11].

### *1.3. Nuclear Magnetic Resonance*

#### 1.3.1. Concept and Application

Nuclear Magnetic Resonance (NMR) is the response of atoms' nuclei, having an odd number of proton and neutron, in the presence of applied magnetic fields at a specific resonance frequency [12]. As a result of nuclei angular momentum, in this case, hydrogen proton, due to its large magnetic moment and abundance in both water and oil, the proton experiences a magnetic moment. The proton then behaves like a spinning bar magnet. It is polarised in the presence of an externally applied magnetic field, resulting in an induced frequency (Larmor frequency) that is a quantum property and can be measured [22]. The schematic nuclei response to the applied magnetic fields is illustrated in Figure 1. Therefore, NMR manipulates the wetting fluids' hydrogen protons to measure its relaxation times (NMR signal decay) due to its interaction with the pore surface [23,24]. Due to this interaction, the wetting fluid experiences a longer relaxation time (smaller diffusion coefficient) due to the surface relaxation acceleration from the surface–liquid interaction (surface relaxivity). As a result, the non-wetting fluid relaxes at a single exponential time (bulk relaxation time). This mechanism has been reported and recognised by several authors [8,22,25–27] that for a rock not being thoroughly water wet, the surface relaxation will accelerate the bulk relaxation time of oil. This acceleration will be dependent on the volume of oil interacting with the pore surface.

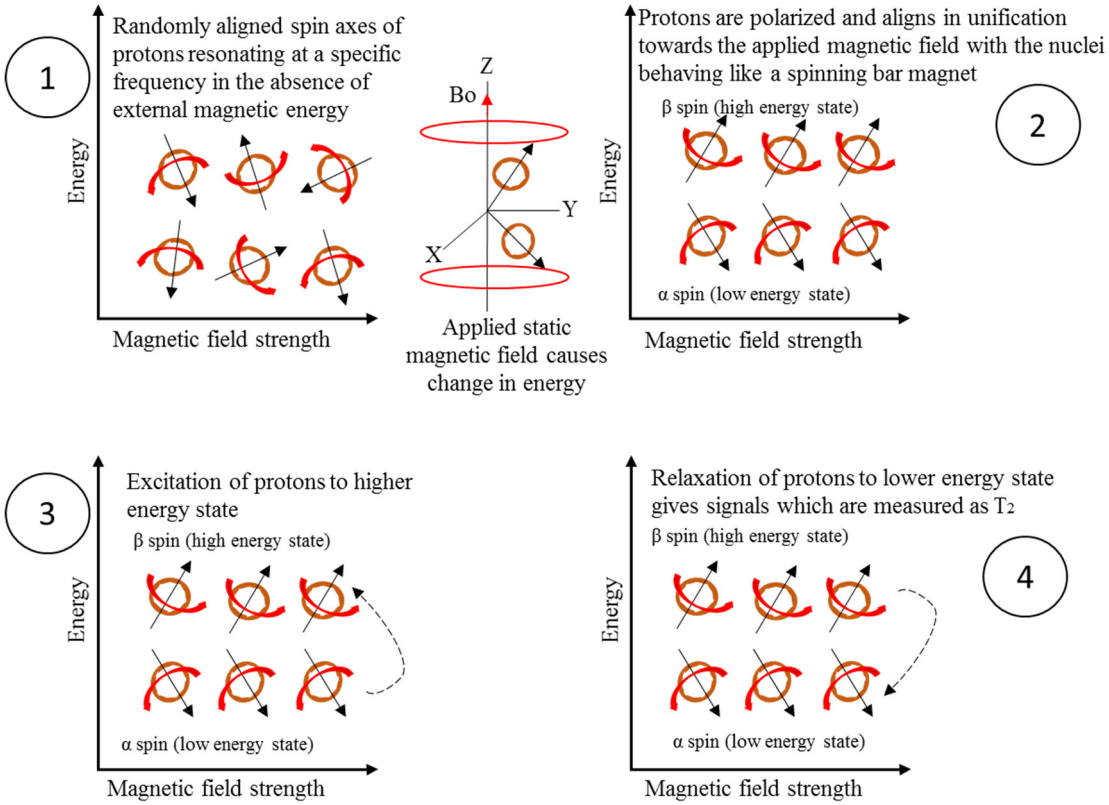


**Figure 1.** Polarisation of hydrogen nuclei illustrating nuclei excitation/relaxation in the presence of applied external magnetic energy.

The general relaxation time of the fluids in the pores of a sample with a uniform magnetic field can be defined as [28]:

$$\frac{1}{T_2} = \frac{1}{T_{2B}} + \rho\frac{A}{V}\frac{f_i}{S_i} \tag{2}$$

where $T_2$ is the observed relaxation time, $T_{2B}$ is the relaxation rate of bulk fluid, $\rho$ is the surface relaxivity, A/V is the pore ratio of the surface area to volume and is used to represent the pore size, $f_i$ is the fraction of pore wetted by fluid and $S_i$ is the saturation of pore by fluid. This equation was then used to measure the Nuclear Wetting Index, NWI, as [11]:

$$\text{NWI} = \frac{\text{Surface wetted by water} - \text{Surface wetted by oil}}{\text{Total Surface}} \tag{3}$$

To correlate to the Amott–Harvey index, results vary from −1 for strongly oil-wet rocks to +1 from strongly water-wet rocks, with 0 representing neutrally wet rocks. This, however, is not fundamentally equal to the Amott–Harvey and the USBM indices. The relationship between the wetting fluid decay time and its interaction with the pore spaces can also be expressed as [3]:

$$\frac{1}{T_i} \approx \rho\frac{A}{V} \tag{4}$$

where the surface relaxivity $\rho$ is expressed as [3]:

$$\rho = \frac{\delta}{T_{iS}} \tag{5}$$

$\delta$ represents the surface area thickness, and $T_{iS}$ represents the surface relaxation time. Measurement of wettability using measurements of $T_2$ is complicated in that it also relies on the measurements of fluid viscosity, rock relaxivity, and pore size distribution. The

most pragmatic way to circumvent this is to quantitatively apply numerical modelling to measure wettability and saturation [5]. The use of the NMR tool, a non-invasive and non-destructive tool, as part of the wireline logging suite makes it possible to locate pay zones in a thinly laminated sandstone reservoir that would have otherwise been missed by the resistivity, density, and gamma-ray logs. This helps for volumes to be quantified and identified, improving completion strategy and maximising the reservoir production.

#### 1.3.2. Wettability Determination

Before the advent of pore size distribution measurement and micro-structural properties being the primary NMR application method, Brown and Fatt [25] published research on fractional wettability measurement by nuclear magnetic relaxation method. Their work focused on using uncoated sand packs and dri-film lubricant treated sand packs as water- and oil-wet porous media, respectively measuring the $T_1$ relaxation time of water in both media. The dri-film lubricant provides a low friction coefficient, which gives excellent lubrication. There was an inverse relationship between the fraction of oil-wet sand and the longitudinal relaxation time. There was extensive analysis and discussion on this done by Zhang et al. [29], where different heterogenous sandstones were used. Applying a low field NMR spectrometer on different sandstone reservoir types using two fluids, Soltrol 130 as refined oil and crude oil from the Gulf of Mexico, alters wettability in aged core analysis. $T_1$ was found to define wettability alteration quantitatively, culminating in the proposal of a systematic approach in interpreting wettability alteration in sandstone reservoirs with varying amounts of shales. This led to further research by Guan et al. [30], where the arithmetic mean of the transverse relaxation time distribution, initial water saturation, residual oil saturation, and irreducible water saturation was defined. This result was then correlated to the fractional wettability index of the samples used. This pioneering work inspired several other notable pieces of research on NMR.

NMR has proven to be very sensitive to the strength of the rock-fluid boundary. The fundamental physics behind it is that the wetting fluid has a relatively shorter relaxation time than the non-wetting fluid. Thus, the wetting fluid has a surface relaxation time similar to the bulk phase value with the reduction in relaxation time of the oil (oil-wet) from the bulk phase value being used as a qualitative [26,31] and quantitative wettability index [8,32,33]. The $T_2$ distribution of the different fluids in a reservoir is illustrated in Figure 2.

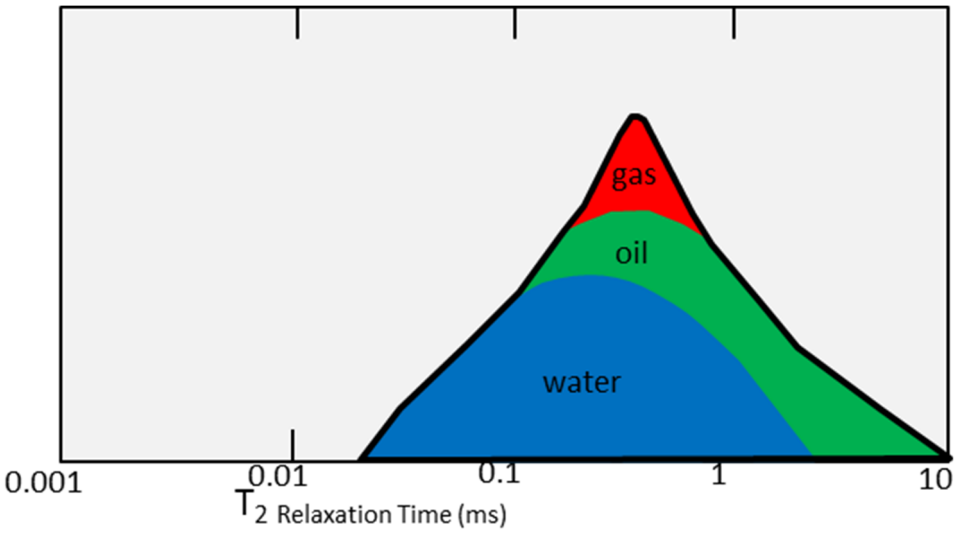


**Figure 2.** Graphical representation of NMR $T_2$ distribution of different fluid phases.

### 1.4. Machine Learning Application in Reservoir Characterisation Related Works

Implementing machine learning methods and Artificial Intelligence (AI) to solve complicated problems has gained momentum in many industries, including the petroleum industry [15]. Most of these sophisticated complications impeded critical decision-making and enhanced advancement in the industry. Hence, researchers progressively moved from using empirical correlations and linear regression models to applying AI techniques that have been welcomed due to their added value in the industry [34]. Researchers [14,35–37] developed AI to predict subsurface geological and petrophysical properties, identify seismic patterns, analyse drill bit problems. The main aim of their research was to improve production, drive down cost and add value to the industry. However, applying a suitable AI modelling technique is also key in achieving the best results; hence many papers have utilised the different AI models suitable to the industry. As a result, AI application has achieved some excellent results compared to actual data [38,39].

Being able to quantify water saturation accurately forms the basis of wettability measurement. However, none of these methods are time-efficient, inexpensive, or gives failproof results due to the many limitations hence the advancement of machine learning models. Therefore, in determining the best algorithm for this research, the accuracy and range of acceptable parameters will be analysed systematically based on performance, complexity, accuracy, reliability, and sensitivity to establish a simplified relationship amongst core data, wireline logs, and NMR wettability measurements. This could invariably overcome complex relationships between surface wettability and other petrophysical data recorded by wireline logs and serves as an immense advantage to accurately predict surface wettability in the field.

### 1.5. Contribution to Study

The prediction of wettability has been made by several researchers using numerous methods. However, some gaps such as testing with only water- or oil-wet samples, using classical mathematical approaches, and its inapplicability on a field scale have been identified, leading to this research. Wireline logs offer the best approach to measure wettability in the field due to them being ubiquitous in every well. The reported success of applying machine learning techniques to predict several reservoir parameters makes its use a viable candidate in predicting NMR logs from wireline logs. Hence, the main contributions of this research are:

1 The integration of a new technique based on laboratory reported results on Amott-USBM and NMR $T_2$ logarithmic mean ($T_2$LM) based wettability to develop a new mathematical model for establishing a relationship between field and laboratory wettability measurement.
2 The development of a systematic field methodology to predict wettability using NMR $T_2$LM and water saturation.
3 Engineering new features from existing log data capable of improving the model prediction of NMR $T_2$LM.
4 The application of deep ensemble super learner to establish a relationship between wireline logs and NMR $T_2$LM and predict NMR $T_2$LM. This predicted log could subsequently be used with the developed methodology and mathematical model to predict reservoir wettability.

## 2. Methodology

### 2.1. Data Collection and Description

The primary samples used in this research were experimental and field sandstone data from Western Australia. The core plugs of the sandstone samples were prepared and utilised for the surface wettability characterisation through Amott-USBM and NMR-based measurements. This data will be used to establish a mathematical correlation between the Amott-USBM wettability index and NMR $T_2$LM field data. Table 2 lists the basic wettability measurements and properties of the core plug samples. Figure 3 indicates the incremental

porosity and multiple pore size distribution with relatively good pore-to-pore connectivity identified using multi-exponential decay time analysis of NMR $T_2$ distribution at 100% brine saturated samples.

**Table 2.** Amott-USBM wettability measurements of core samples used in correlating core to log wettability.

| Sample ID | Depth m | $K_{air}$ md | Helium Porosity, pv | Amott Wettability | | Oil Drive Area | Water Drive Area | USBM Wettability Number | Wettability Indicated |
|---|---|---|---|---|---|---|---|---|---|
| | | | | Index to Water | Index to Oil | | | | |
| 8 | 1271.60 | 4490 | 0.308 | 0.313 | 0.00 | 0.77 | 0.25 | 0.479 | Water wet |
| 9 | 1271.90 | 3130 | 0.318 | 0.158 | 0.00 | 0.75 | 0.27 | 0.443 | Water wet |
| 10 | 1272.21 | 2677 | 0.297 | 0.143 | 0.00 | 1.79 | 0.32 | 0.750 | Water wet |
| 15 | 1273.70 | 2611 | 0.277 | 0.116 | 0.00 | 2.18 | 0.42 | 0.719 | Water wet |
| 62 | 1290.75 | 2540 | 0.249 | 0.513 | 0.00 | 1.33 | 0.35 | 0.582 | Water wet |

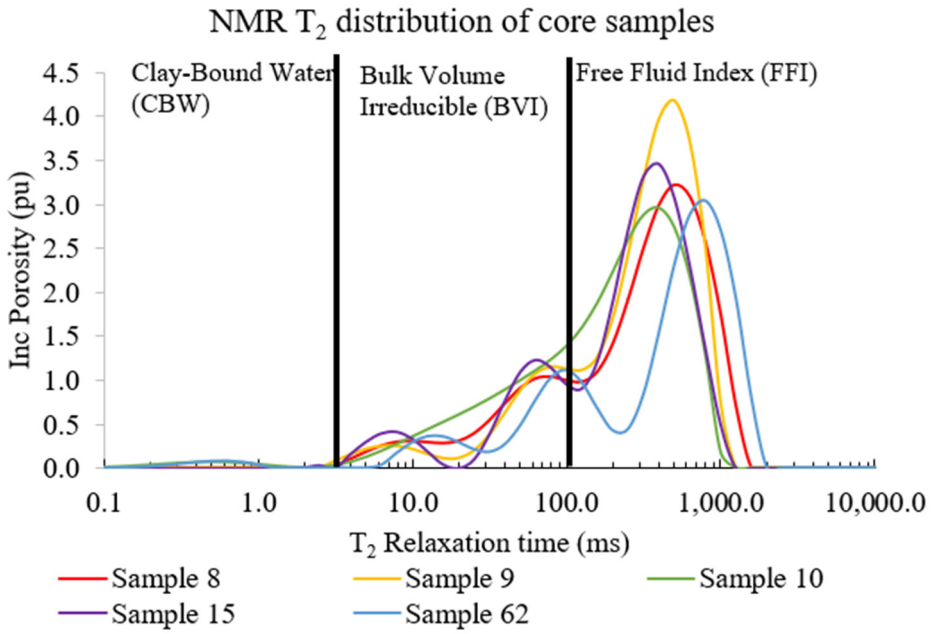


**Figure 3.** NMR $T_2$ multimodal pore size distribution of core samples at 100% water saturation.

Conventional wireline logs from a well in the Coniston field will be used to train and test the deep super learner model in predicting NMR $T_2$LM. This is crucial for the next stage, where attempts are made to establish relationships between NMR related measurements and wireline logs data. The accuracy of this model in predicting NMR $T_2$LM for wells will be essential in the accurate measurement of porosity, pore size distribution, and permeability. In addition, this model will be used to determine wettability via the new empirical formula. This will help test the accuracy of the empirical equation in a different field.

### *2.2. Developing an Empirical Determination of Wettability from Core and NMR $T_2$LM Data*

A correlation between Amott-USBM test results and NMR $T_2$LM from the same well depth was used to develop the empirical correlation between laboratory wettability data and field wettability data. This technique is sought to work for both clastic and carbonate reservoirs and oil and gas fields to determine in situ wettability conditions. A combined Amott-USBM wettability data from the five (5) core samples were used to establish a relationship with NMR $T_2$LM. The depths of these cores were matched with the NMR $T_2$LM log from the same reservoir. The depth matching enabled a cross-plot of the Amott USBM wettability results and NMR $T_2$LM. An exponential function was used for this research because, regardless of how low the $T_2$LM will be, in a hydrocarbon-filled reservoir interval, there will never be the existence of 100% water; hence Amott-USBM wettability

index will never reach 0. This mathematical model will form the basis for determining the wettability from NMR $T_2$LM. In classifying the wettability condition, an integration of Amott-USBM proposed for this study is illustrated in Figure 4. The proposed sub-categories of oil-wet and water-wet will help capture the different strengths of wettability existing in the reservoir. Wireline and core data from a sandstone oil and gas reservoir in Australia were used to test and validate the hypothesis. These data included Gamma ray (GR), Density (DEN), Neutron (NEU), Total porosity (POR), Resistivity (RES_DEP, RES_MED, RES_SLW), water saturation (Sw), and permeability (PERM) logs. The second part of this research involves applying machine learning techniques in predicting NMR $T_2$LM in the field using wireline logs. The developed methodology and empirical equation were then applied to the predicted $T_2$LM to predict the wettability state.

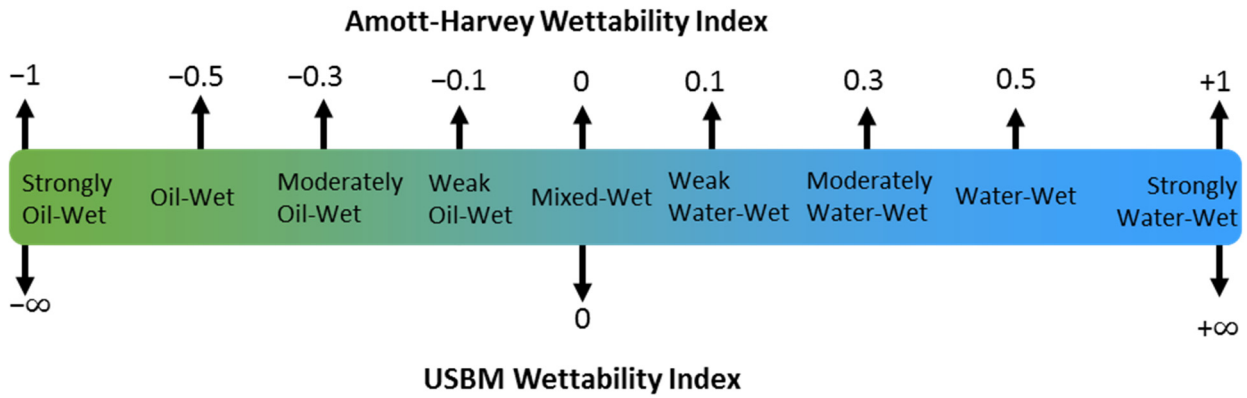


**Figure 4.** Graphical illustration of the integration of Amott–Harvey and USBM wettability index showing the different types of wetting conditions.

### *2.3. Validating the Developed Approach in Wettability Prediction*

To test the reliability of this mathematical model on the field, water saturation and field NMR $T_2$LM log data were used. The integration of water saturation helps to analyse the depth where hydrocarbon and water coexist, leading to the wettability phenomenon. The field cut-off for water saturation was increased from 70% to 80%. The increment was to accommodate enhanced and chemical oil recovery mechanisms and varying field cut-offs for net pay zones based on the field development plans and production forecasts. Based on several chemical enhanced oil recovery at field scale, residual oil saturation at 20% is optimal [40]. The overall reservoir water saturation and NMR $T_2$LM relationship were established using cross plots. The reservoir interval was subsequently divided into either increasing or decreasing water saturation zones, based on water saturation cut-off. That is, where saturation increases up to the point it starts to decrease is delineated and vice versa. This subdivision created several zones where a cross plot of saturation and $T_2$LM was employed to help analyse the effect of hydrocarbon on $T_2$LM. The average $T_2$LM for the observation interval is computed to establish wettability.

### *2.4. Data Analysis and Feature Engineering*

The correlation between the input and output variables was identified using covariance matrix and pair plots via Jupyter notebook python programming language. The descriptive statistics of the data are summarised in Table 3. This analysis helps to determine the data distribution and their expected behaviour.

**Table 3.** Summary statistics of the data used in the machine learning modelling.

| | DEN | GR | PERM | NEU | POR | RES_DEP | RES_MED | RES_SLW | Sw | $T_2$LM |
|---|---|---|---|---|---|---|---|---|---|---|
| count | 1529 | 1529 | 1529 | 1529 | 1529 | 1529 | 1529 | 1529 | 1529 | 1529 |
| mean | 2.51 | 100.35 | 272.23 | 0.18 | 0.1 | 8.77 | 8.27 | 4.11 | 0.84 | 55.77 |
| std | 0.16 | 39.41 | 1201.43 | 0.10 | 0.08 | 38.13 | 23.12 | 3.46 | 0.25 | 104.92 |
| min | 2.02 | 40.23 | 0.00 | 0.01 | 0.01 | 0.13 | 0.51 | 0.51 | 0.10 | 0.49 |
| 25% | 2.46 | 66.61 | 0.00 | 0.11 | 0.04 | 2.08 | 2.34 | 2.02 | 0.72 | 2.20 |
| 50% | 2.57 | 97.14 | 0.01 | 0.17 | 0.07 | 3.40 | 3.65 | 3.07 | 1.00 | 6.06 |
| 75% | 2.62 | 128.29 | 1.35 | 0.23 | 0.12 | 6.49 | 6.97 | 4.83 | 1.00 | 64.77 |
| max | 2.75 | 384.84 | 10974.59 | 0.54 | 0.31 | 1051.84 | 457.65 | 27.22 | 1.00 | 700.96 |

Feature engineering was employed by applying domain knowledge to generate a useful new feature that was a better predictor of NMR $T_2$LM. A fundamental feature generation approach summing the RHOB, NPHI, K and Sw log was employed based on the feature importance of most feature selection techniques used in this study. These logs were chosen because wettability occurrence results from two immiscible fluid and their interaction with pore space. In comparison, DEN and water saturation are highly correlated to pore spaces and the coexistence of two fluids, respectively.

A deep ensemble super learner was used in the next stage to predict the NMR $T_2$LM using wireline logs. Feature selection was performed to select the best wireline logs to predict wettability. However, that did not improve the model performance hence was omitted in this study. The feature selection is a crucial step to ensure that representative logs will be chosen for the prediction [41], but the primary purpose of applying machine learning models is to improve performances hence its omission.

### *2.5. Machine Learning Model Application*

The superiority of deep learning over traditional machine learning algorithms has been well documented. The ability of deep learning models to handle big data has made it very popular for tasks involving pattern recognition and predictive modelling [15]. However, deep learning still has some shortcomings with all of its reported improvements in artificial intelligence applications. These drawbacks, such as hyperparameter tuning, slow convergence in small datasets, infinite architectures, and complexities, are mainly experienced. Although conventional machine learning models tend to solve these drawbacks on small datasets, they cannot achieve relatively superior model performance. The vague definition of small datasets also makes it challenging to know which model to use. Based on these assertions, a new ensemble method based on deep learning architecture, result interpretability, and relatively faster convergence on both large and small datasets is used in this study. This deep ensemble super learning algorithm stacks several diverse base learners in its architecture to improve model performances. However, these models on their own do not achieve relatively high performances as compared to the deep ensemble super learner.

#### 2.5.1. Building a Deep Learning Regression with Keras

The deep neural network created takes all the features as input, feeds them into nine connected hidden layers, each of 50 nodes creating a dense network. The features are then connected into an output layer. The data used to evaluate these models were split into train and test data using a 20% holdout cross-validation. The train data was subsequently split into 30% for validation. The sequential constructor model is used since the network consists of a linear stack of layers. The dense layers were used for each hidden layer where the number of neurons was specified. The Rectified Linear Unit (ReLU) activation function was used for all the hidden layers in the network. The ReLU activation function is non-linear, computationally efficient. The activation function takes a derivative function related to all the input features allowing for backpropagation. The ReLU activation function is a good approximator. Another advantage of using ReLU is its ability to work with sparse neurons,

making it much more efficient than the sigmoid and tanh activation functions. The ReLU function gives an output value if the input is positive but returns zero or no output if otherwise. The mean absolute error was used as the error metric between the predicted and actual values. The minimisation algorithm used to optimise the deep learning regression is the Adaptive Movement Estimation algorithm (Adam). The main advantage of Adam is that the learning rate for each input is automatically adapted. Adam saves time from further optimising the learning rate, which is the case when using gradient descent as an optimiser. Adam also uses an exponential decreasing moving average of the gradient to update variables.

#### 2.5.2. Implementation of Deep Ensemble Super Learner Model

The deep ensemble super learner model, also referred to as a stacking ensemble or deep ensemble, is a supervised ensemble algorithm of diverse machine learning models or one model with diverse hyperparameters, otherwise referred to as base learners. Nine base learners were used to develop the deep ensemble super learner, and they are listed in Table 4. The super learner employs the k-fold cross-validation technique in mapping the training dataset (X, y) into a prediction dataset (Z, y) using the base learners. This strategy improved model performances.

**Table 4.** Summary of algorithms used as base learners and corresponding authors.

| Base Learner | Authors |
|---|---|
| Category Gradient Boosting Machine (CatBoost) | Dorogush [42] |
| Light Gradient Boosted Machine (LGBoost) | Ke et al. [43] |
| Decision Tree | Gordon et al. [44] |
| Deep learning | Bengio [45] |
| Adaptive Boosting (AdaBoost) | Freund and Schapire [46] |
| Bootstrap aggregation (Bagging) | Breiman [47] |
| Random Forest | Breiman [48] |
| Extremely Randomised Trees (Extra tree) | Geurts et al. [49] |
| Extreme Gradient Boosting (XGBoost) | Chen and Guestrin [50] |

The development of the super leaner supposes an observation of the learning data stated as $Z_i = (X_i, Y_i) \sim W_0$, $i = 1, \ldots, n$ with the sole aim of estimating the regression function $\psi_0(Y) = E(X|Y)$. This regression function may be expressed as the minimiser of the squared error loss function:

$$\psi_0 = \arg\min E\left[(X - \psi(Y))^2\right] \quad (6)$$

where $(X - \psi(Y))^2$ represents the loss function, X represents the interest outcome, and Y is the n-dimensional covariate set. This loss function can be applied to map any input to output over a parameter space using predictive algorithms for any specified problem.

Although there is a provision of an efficient, super learner algorithm in the python library [51], the algorithm used in this study was manually developed using a straightforward implementation with sci-kit learn [52]. The deep ensemble super learner essentially uses an ensemble model for each layer. In the interest of brevity, this section is simplified. More detailed statistical theory and implementation of the super learner can be found in [53–55]. The overall training process is described in the workflow (Figure 5).

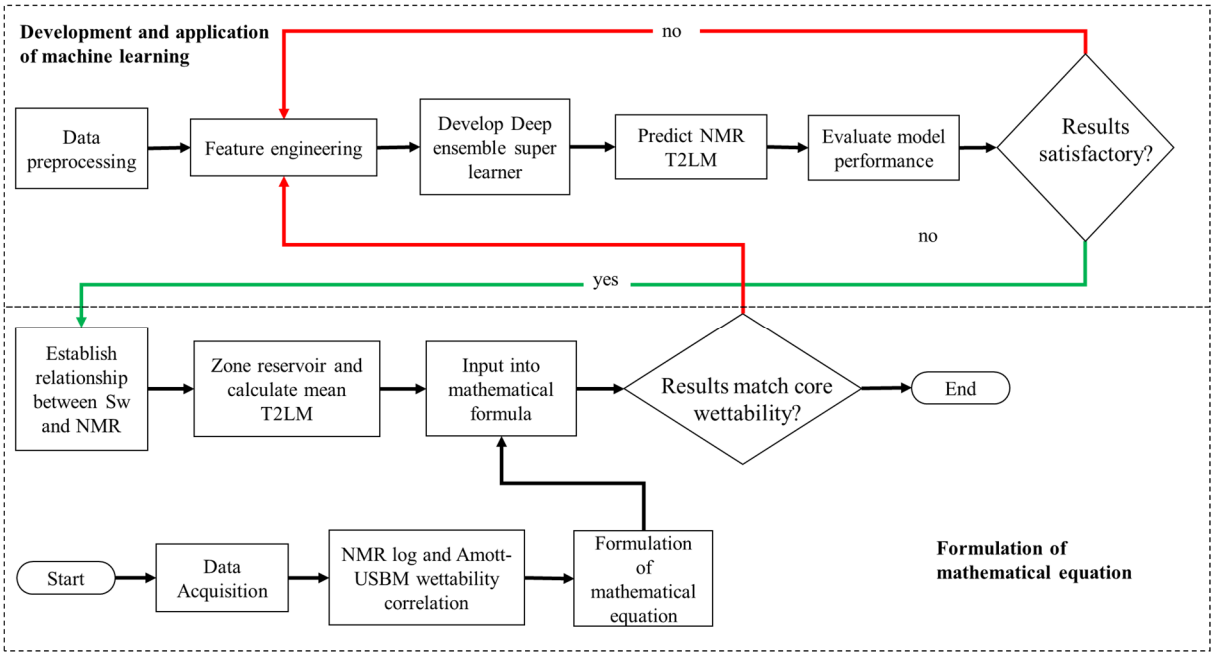


**Figure 5.** Simplified workflow of the methodology used in developing the mathematical formula and machine learning model in this study.

The below code (Algorithm 1) was used to define and stack the diverse regression machine learning models optimised using Bayesian Optimisation. The code in Algorithm 2 was also used to collect out of fold predictions from the k-fold cross-validation. The following summarises the procedure:

1 All base learners were identified and hyperparameter tuned. Where the collection of B candidate learners representing the empirical probability distribution is represented as $X_b$, $b = 1, \ldots, B$A 10 K-fold cross-validation technique creates out-of-fold predictions to train the meta-model (super learner). Where $K\epsilon\ \{1, \ldots, V\}$ index a data split into $K(v) \subset \{1, \ldots, n\}$, which refers to the validation data and T$(v) \subset \{1, \ldots, n\}$ which is the training data, a complement of the validation data set.

2 All the out-of-fold model predictions were used as an input column for the meta-model, where each model's prediction for the columns was stacked horizontally into rows and stored. All the stacked rows were vertically stacked and stored for all the out-of-fold predictions creating a new meta-model matrix dataset of prediction columns. The construction of the new metadata for n is formulated as a vector consisting of the predicted base learner values according to the base learners fitted on the training data. This is represented as $(Y_i, Z_i)$, where $Z_i \equiv \left(\psi_{nbv(i)}(X_i) : \ b = 1, \ldots, B\right)$ with Z being the possible outcome set.

3 All Bayesian optimised base learners were trained on the whole training dataset using the k-fold cross-validation. The performance of these base learners is subsequently stored.

4 The meta-model dataset is used to train the super learner using linear regression for the sequential combination of all base learners.

5 All base learners and super learners' performances on the test data are evaluated.

6 The super learner predictions function using the meta-model will map new data sets into a prediction set to make predictions for new data.

Due to the stochastic nature of the models, numerical variations, and the evaluation criteria used, the results may vary; hence it is encouraged to run the model multiple times to get an averaged outcome.

Algorithm 1 is about model architecture creating a stack of ensemble models as base learners for the Super Learner. Modified after [56].

**Algorithm 1:** Stacking an ensemble of base learners

```
def get_models():
models = list()
models.append(DeepLearning_regression_model ( ))
models.append(AdaBoostRegressor( ))
models.append(DecisionTreeRegressor())
models.append(LGBMRegressor( ))
models.append(CatBoostRegressor( ))
models.append(XGBRegressor( ))
models.append(RandomForestRegressor())
models.append(BaggingRegressor( ))
models.append(ExtraTreesRegressor( ))
return models
```

Algorithm 2 is about collecting out of fold predictions from K-fold cross-validation to create meta model dataset to train and test the Deep Ensemble Super Learner [56].

**Algorithm 2:** Collecting out of fold predictions from k-fold cross-validation

```
def get_out_of_fold_predictions(X, y, models):
meta_X, meta_y = list(), list()
# Data splitting
kfold = KFold(n_splits = 10, shuffle = True)
# Data split computation
for train_ix, test_ix in kfold.split(X):
fold_yhats = list()
# get data
train_X, test_X = X[train_ix], X[test_ix]
train_y, test_y = y[train_ix], y[test_ix]
meta_y.extend(test_y)
# Train and predict with each sub-model
for model in models:
model.fit(train_X, train_y)
ypred = model.predict(test_X)
# store columns
fold_ ypred.append(ypred.reshape(len(ypred),1))
# store fold ypred as columns
meta_X.append(hstack(fold_ypred))
return vstack(meta_X), asarray(meta_y)
```

### *2.6. Criteria for Model Evaluation*

In assessing the performance of machine learning models, the difference between the predicted and the actual data represents the prediction error. In evaluating the appropriateness of the models used in this research, the errors of the models were assessed using the following metric criteria:

Mean Absolute/Percentage Error (MAE/MAPE): This evaluation metric is very sensitive to relative errors and robust against global scaling of the predicted output. This is statistically defined as:

$$\begin{aligned} \text{MAE} &= \frac{\sum_{i=1}^{n} |y_i - \hat{y}_i|}{n} \\ \text{MAPE} &= \frac{1}{n_o} \sum_{i=0}^{n_o - 1} \frac{|y_i - \hat{y}_i|}{\max(\epsilon, |y_i|)} \end{aligned} \tag{7}$$

Root Mean Squared Error (RMSE): This performance evaluation metric is popular because it is interpretable as the standard deviation of models' prediction errors and specifies the closeness of predicted data to actual data. This is written as:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{8}$$

Mean Poisson Deviance (MPD): This metric produces the expected, predicted regression output values by computing the mean Tweedie deviance error with a power parameter of 1.

$$D_{(y,\hat{y})} = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}} \{2(\log(\hat{y}_i/y_i) + y_i/\hat{y}_i) - 1\} \tag{9}$$

Coefficient of determination ($R^2$): Another widely used criterion is the coefficient of determination. It represents how close the dependent value is to the best-fit regression line. This is written as:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \overline{y}_i)^2} \tag{10}$$

Besides $R^2$, low errors are indicative of good model performance.

### 2.7. Model Deployment and Validation

Data from a different field will be used to validate the mathematical formula to assess its determination of water wetness. Since this data does not have core Amott wettability measurements, we will mainly assess the formula on relative permeability. The data will also be segregated into various sections based on water saturation. The Deep Ensemble Super Learner will be used to predict the T2LM for this reservoir. The same methodology applied in this research will be employed to analyse the wettability qualitatively. The mathematical formula will then be used to quantitatively predict the type of wettability in existence based on the various wettability scale used in this research.

## 3. Results and Discussion

### 3.1. Development of a Mathematical Model to Establish a Relationship between Amott-USBM Wettability and NMR T2LM

From the Amott-USBM test conducted, samples primarily showed water-wet wettability states. However, using the Amott-USBM scale used in this study, samples 9, 10, and 15 showed moderately water-wet conditions. Samples 8 and 62 indicated water-wet and strongly water-wet conditions, respectively (Table 5). This scale was used to analyse the various heterogeneous wettability states in existence. From the cross-plot of the NMR $T_2$LM log data and the Amott-USBM experimental data of the core samples, it was observed that there is some form of correlation between them. However, with the understanding of wettability in a reservoir, there can never be only hydrocarbon in the pore space as there will always be irreducible water. The relationship should not be less than or equal to 0 for the Amott-USBM index with this assumption. This is because an oil-wet system is classified in the negative (−1 to 0), and water-wet is positive (0 to 1). Therefore, an exponential function is used, making it impossible for the equation to return negative wettability index values (see Figure 6). This yielded a correlation coefficient of 0.95, and consequently, a new formula to quantitatively predict wettability from NMR $T_2$LM is achieved as:

$$\hat{w} = \pm 0.0528 e^{0.0081(\log T) \times T)} \tag{11}$$

where $\hat{w}$ refers to the Amott-USBM wettability index and T represents the average NMR $T_2$LM against water saturation distribution within the reservoir interval. Since NMR response in the field could be affected by other complicating conditions such as mud filtrates and their degree of saturation, the applicability of the proposed formula will have

some limitations. Hence, the NMR log affected by field conditions needs to be adjusted to cater for their effects. The unavailability of sufficient core wettability and NMR log data provides a major uncertainty. However, from the cross-plot, high $T_2$LM representing the gas phase correlates to high Amott-USBM value, representing water-wet to strong water-wet conditions. This analysis validates the plot, and more data will help improve the accuracy of the proposed mathematical formula.

**Table 5.** Initial wettability states and modified wettability states for core samples.

| Sample ID | Depth m | Amott Wettability | USBM Wettability Number | Wettability Indicated | Modified Wettability Indicated Using the Wettability Scale in This Study |
|---|---|---|---|---|---|
| 8 | 1271.60 | 0.313 | 0.479 | Water wet | Water wet |
| 9 | 1271.90 | 0.158 | 0.443 | Water wet | Moderately water wet |
| 10 | 1272.21 | 0.143 | 0.750 | Water wet | Moderately water wet |
| 15 | 1273.70 | 0.116 | 0.719 | Water wet | Moderately water wet |
| 62 | 1290.75 | 0.513 | 0.582 | Water wet | Strongly water wet |

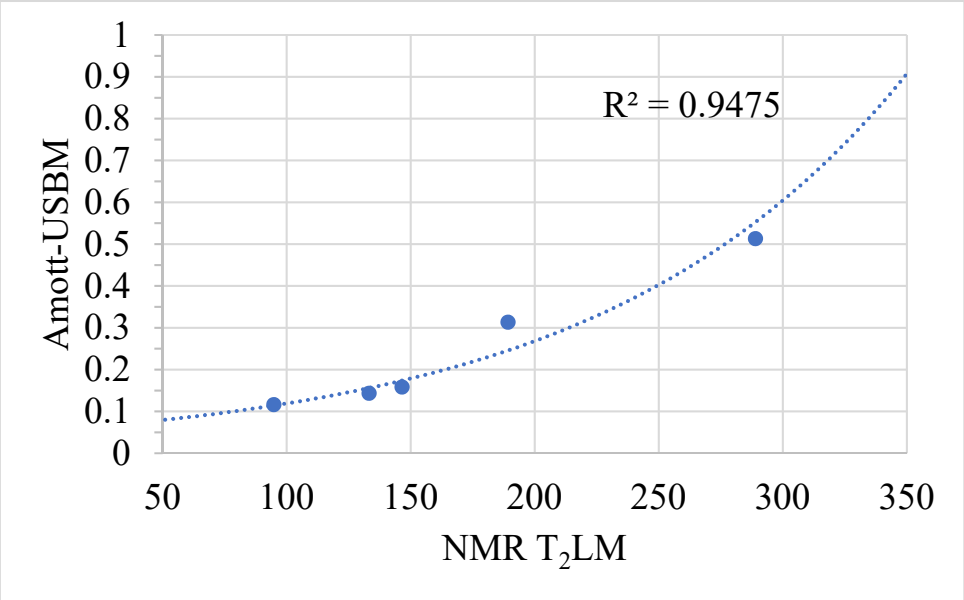


**Figure 6.** Application of Amott-USBM—NMR $T_2$LM cross plot to establish an exponential relationship between core and log wettability data.

### *3.2. Effect of Wettability on $T_2$ Distribution and $T_2$LM*

The ability of crude oils to alter the wettability of a pore surface varies. The $T_2$ relaxation times for fluids existing in rock pores are affected by independent relaxation mechanisms like bulk fluid processes, surface relaxation, and diffusion in the magnetic field gradient. Their relative importance is dependent on the fluids in the pores, pore sizes, surface relaxation strength, and the wetting fluid. The wetting fluid is dominated by surface relaxation. The non-wetting fluid has bulk relaxation as its primary contributing source, while gas is mainly dominated by diffusion in the presence of magnetic field gradients. From NMR data analysis, each fluid exhibits a bulk relaxivity, but when oil molecules wet pore surfaces, they exhibit additional surface relaxivity, thus complicating the $T_2$ relaxation mechanism. This complication is because the effect of the surface relaxivity spectra on oil is quite different from water.

Consequently, when a reservoir exhibits mixed wettability, the situation becomes much more complicated. Based on this notion, this methodology seeks to establish a simplified way of determining reservoir wettability using the relationship between the geometrical mean of $T_2$ relaxation spectra and water saturation. The results of this relationship are

then related to an Amott-USBM wettability measurement from core data through a new empirical equation. Figures 7 and 8 show a schematic representation of wettability's effect on the $T_2$LM log based on water saturation distribution.

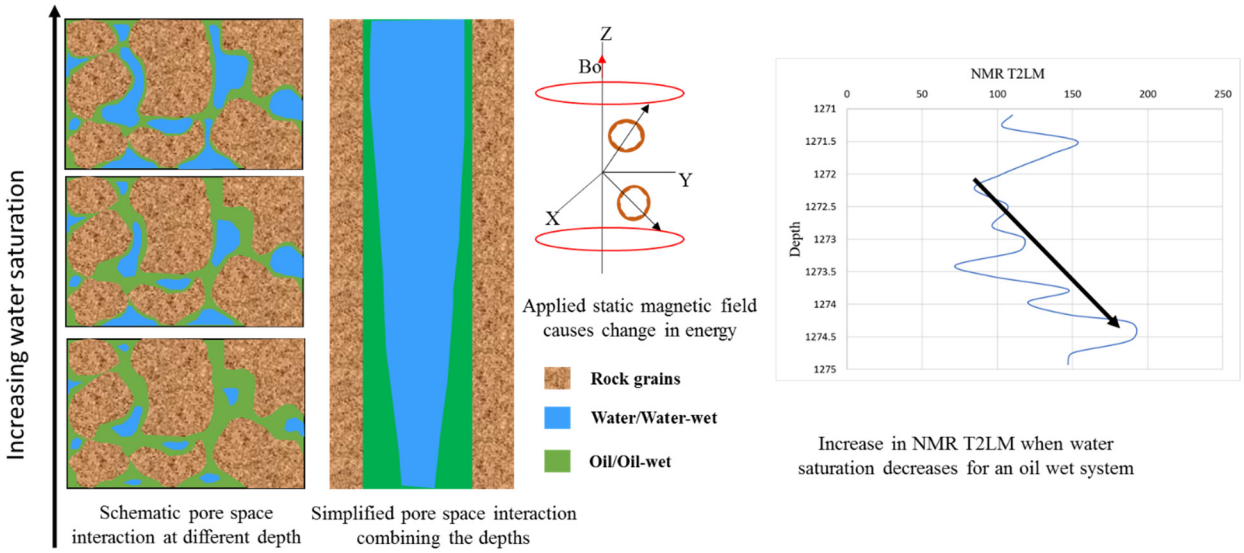


**Figure 7.** Schematic illustration of NMR $T_2$LM distribution as a function of water saturation in a water-wet condition.

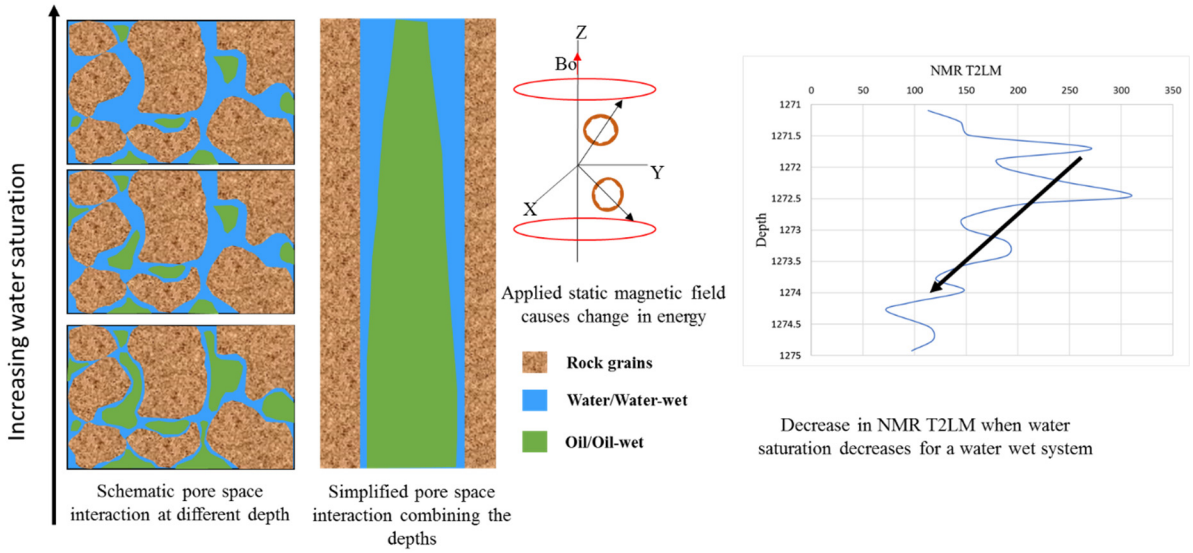


**Figure 8.** Schematic illustration of NMR $T_2$LM distribution as a function of water saturation in an oil-wet condition.

Analysis shows that a variation in water saturation distribution within the pore spaces affects the NMR relaxation times. The wettability effect of the surface relaxivity phenomenon occurs for a water-wet system where under increasing oil saturation, there is an observed but slight decrease in $T_2$LM. When water is wetting the pore surface, the water experiences additional surface relaxation. However, the non-wetting fluid, which is oil, will only experience bulk relaxivity.

Therefore, there is an increase in the water phase $T_2$ relaxation time at decreasing water saturation. The bulk relaxation time for water is between 10 and 100 ms. In contrast, the peak relaxation of oil is around 1000 ms, although this is mainly for oil with high gas-oil-ratio. At 100% water saturation, $T_2$ is usually around 10 ms [57]. When oil is encountered in the interval, $T_2$ tends to peak, signifying the additional oil relaxivity. The

contribution of bulk oil relaxation time will be subtle, whereas the water phase dominates. This behaviour causes the $T_2$ relaxation time to shift to a slower phase. The intensity of the $T_2$ signal for the oil phase increases as oil saturation increases, causing the bulk relaxivity of oil to be at its peak. The increase in the oil phase causes a consistent non-linear increase in the $T_2$LM. However, it is not as pronounced due to the additional contribution of the surface relaxivity influence of water to the wettability phenomenon. Hence, there is a slight and consistent increase in $T_2$LM readings close to the single exponential time for water. In light of this, it is instructive to state that at high water saturation, residual oil in existence within the reservoir pores is non-wetting. The residual oil is subsequently only exhibiting bulk relaxation, resulting in a positive trend as a function of increasing water saturation at the top part of the trend (0.85 Sw and above). When ascribed to weakly water-wet conditions, this analysis results in a negative trend as a function of increasing water saturation. This trend is because an increased volume of oil wetting the pore surfaces leads to the influence of surface relaxivity on oil and water. Therefore, weakly water-wet to mixed wet will have a flat or negative gradient because no spontaneous imbibition occurs at these wettability conditions.

This study did not investigate oil-wet conditions due to the unavailability of cores that have been experimentally analysed for wettability. However, the contributing effect of surface relaxivity on oil will be pronounced for oil-wet conditions, typically the case in carbonate reservoirs. Hence, an increase in oil saturation will invariably cause a decreasing trend in $T_2$LM as a function of decreasing water saturation. This postulation agrees with the reported research by Johannesen et al. [58], where several core analysis investigations into the effect of wettability on $T_2$ distribution were conducted. With the information and methodology detailed in this research, wettability conditions can be established at the field using water saturation and NMR $T_2$LM log data.

Instances, where high water saturation depicts high NMR T2LM measurements, indicate a non-linear correlation between NMR $T_2$LM and water saturation, which can help identify the different wettability conditions in a reservoir. For a mixed wettability condition, surface relaxivity will depend on fluid volume wetting the pore spaces. Although this analysis is based on water-wet conditions, the subsequent discussions will illustrate how heterogeneous wettability states can be classified.

### *3.3. Development of Methodology to Establish a Relationship between Water Saturation and NMR $T_2$LM*

Using wireline log data from the Coniston well, the applicability of the postulated empirical equation was verified. Based on the in-depth literature review, it is known that bulk oil has a slow $T_2$ relaxation time. The opposite occurs for water which has a relatively faster $T_2$ relaxation time. Further analysis has established that the wetting fluid experiences surface relaxivity for strongly water-wet conditions. Whereas the non-wetting fluid decays at its exponential bulk relaxation; hence, there is a shift in $T_2$ relaxation time towards the high $T_2$ readings in an increasing oil saturation. This indicates an increase in $T_2$LM as a function of increasing oil saturation. This agrees with the extensive core analysis carried out by Johannesen et al. [58].

The hydrocarbon coexisting with the water will influence the $T_2$ relaxation times. Hence, the rate of increase or decrease will be affected based on which fluid wets the pore spaces. Based on Figure 9, the trend line for the shift in $T_2$LM as a function of decreasing water saturation primarily indicates a water-wet condition for both the gas and oil zones agreeing with core wettability measurements. For the delineated reservoir zones, the different gradients in plots indicate the existence of heterogeneous wetting conditions (see Figure 10). However, for a weakly water-wet system, it is suggestive that some of the pore surfaces are wetted by both water and oil, creating a complication. This is further understood qualitatively through the gradient of the trend line from the subtle shift towards faster relaxation. The effect of surface relaxation affects water differently from oil; hence a relative shift to help identify the degree of water wetness is impossible. The proposed

empirical formula will help to measure wettability using the average $T_2$LM quantitatively. Based on the average $T_2$LM data for all the intervals under consideration, the new empirical formula is used to determine the varying wettability conditions quantitatively, and this has been summarised in Table 6.

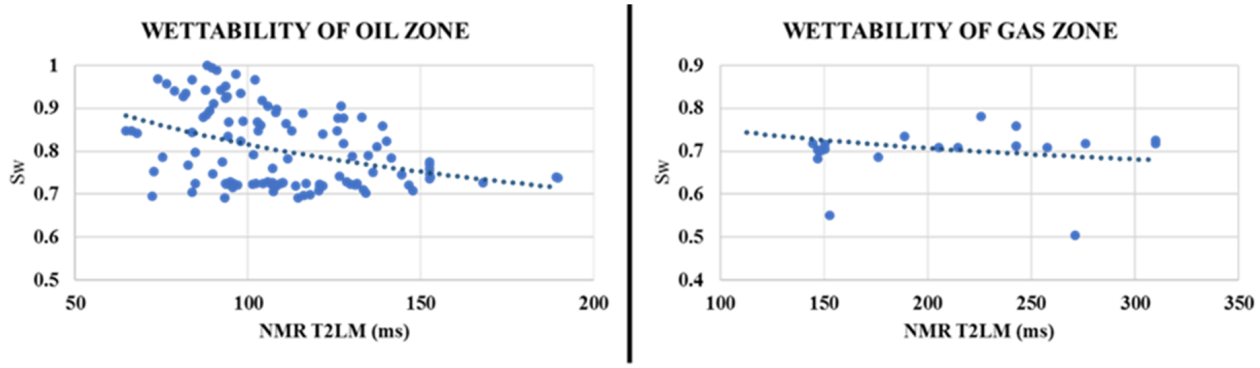


**Figure 9.** Cross-plot correlation between water saturation and NMR $T_2$LM showing the trends for oil and gas zones.

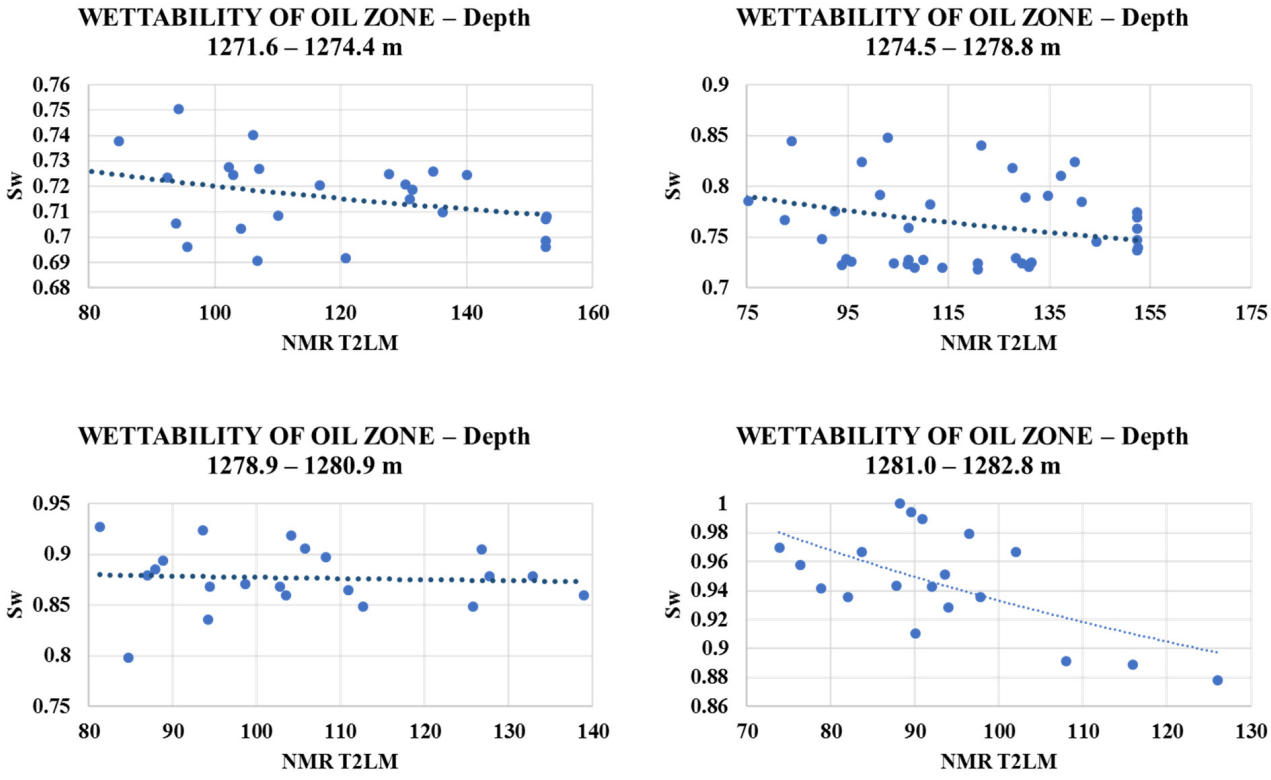


**Figure 10.** Cross-plot correlation between water saturation and NMR $T_2$LM showing the trends in different oil intervals.

**Table 6.** Re-evaluation of wettability conditions for core samples based on the criteria used in this study.

| Depth from m | Depth to m | Average $T_2$LM | Estimated Amott Index | Wettability Indicated in This Study |
|---|---|---|---|---|
| 1269.5 | 1271.50 | 203.699 | 0.55 | Water-wet |
| 1271.60 | 1274.40 | 112.4131 | 0.27 | Moderately water-wet |
| 1274.50 | 1278.80 | 114.4698 | 0.27 | Moderately water-wet |
| 1278.90 | 1280.90 | 105.2729 | 0.25 | Moderately water-wet |
| 1281.00 | 1282.80 | 93.00 | 0.21 | Moderately to weakly water-wet |

The reservoir used in this research has a gas cap, hence the reservoir was delineated into gas and oil zones. Although there was no core data for the gas zone, the developed mathematical model was applied. When used in the mathematical model, the average $T_2$LM for the gas zone gave an Amott index of 0.55, which shows a strongly water-wet condition. Gas is a non-wetting fluid because oil and water propitiously wet pore spaces in reservoir formations [59]. In general, the oil zone interval from the plot showed a positive relationship with increasing water saturation, indicating that the zone is water-wet. To know the type of water-wet condition, the average of the $T_2$LM was incorporated in the model, and results indicated a moderately water-wet condition with an Amott index of 0.26. This agrees with the average experimental Amott results for the entire oil interval of 0.25. However, the zonation of the oil zones based on decreasing water saturation or vice versa showed several water-wet conditions, shown in Table 6 above. Zones, where core wettability results were available agreed with the results from corresponding depths. Based on the observed trends and average $T_2$LM, the last interval exhibited moderately to weakly water-wet conditions. This occurrence is due to its saturation being above 0.85. As earlier explained, at high water saturation of 0.85 and above, the residual oil exists in the micropores, where it wets the surface of the pores. As a result, the micropores are wet by oil whiles water wets the macropores, exhibiting some weakly water-wet conditions. This interval was added to verify this statement. Alternatively, the occurrence of low water saturation and low $T_2$LM refers to intervals of low permeability as well as a low free fluid index.

### *3.4. Evaluation of New Engineered Feature*

From the results of the feature selection techniques, the GR, DEN, NEU, POR and Sw logs were selected as candidates for feature selection for the NMR $T_2$LM prediction. The variety of logs selected are because the best performing feature selection techniques selected all of these logs as relevant. These logs also relate to properties on which wettability occurs. Simple feature engineering techniques were employed in this study to combine existing features into new variables, and these are shown as:

- New Feature 1: Addition (DEN, Sw)
- New Feature 2: Addition (DEN, NEU)
- New Feature 3: Addition (POR, Sw)
- New Feature 4: Product (POR, Sw)
- New Feature 5: Square root (Product (DEN, GR))
- New Feature 6: Round Sw (2))
- New Feature 7: Scaled (Sw)

The performance of the new feature was analysed using accuracy and Permutation Feature Importance (PFI) on test data. The XGBoost model was used to evaluate the performance of all features in NMR $T_2$LM predictions. In terms of the analysis, the XGBoost feature importance identified Sw and POR as having a low correlation which disagrees with the literature. For the well data used, the most important feature was the DEN log which is a formation dependent log. Figure 11 showed the feature performance of all the 9 variables used in predicting T2LM before feature engineering.

The PFI results of the performance of the new feature for NMR $T_2$LM predictions are illustrated in Figure 12. This study used a similar train/test split in the final XGBoost model prediction to maintain consistency. The initial testing $R^2$ without the new features was 0.8148. The importance score is computed in a way that higher values represent better predictive power. The importance of the values for the most relevant features represents a significant fraction of the accuracy score. From the results, many of the input variables have a considerably low importance score. This outcome suggests that the predictive power of these input variables is condensed into a small number of variables. The permutation importance plot shows that permuting a feature drops the accuracy by at most 0.12. This analysis suggests that some of the features are important. This outcome is consistent with the computed high test accuracy.

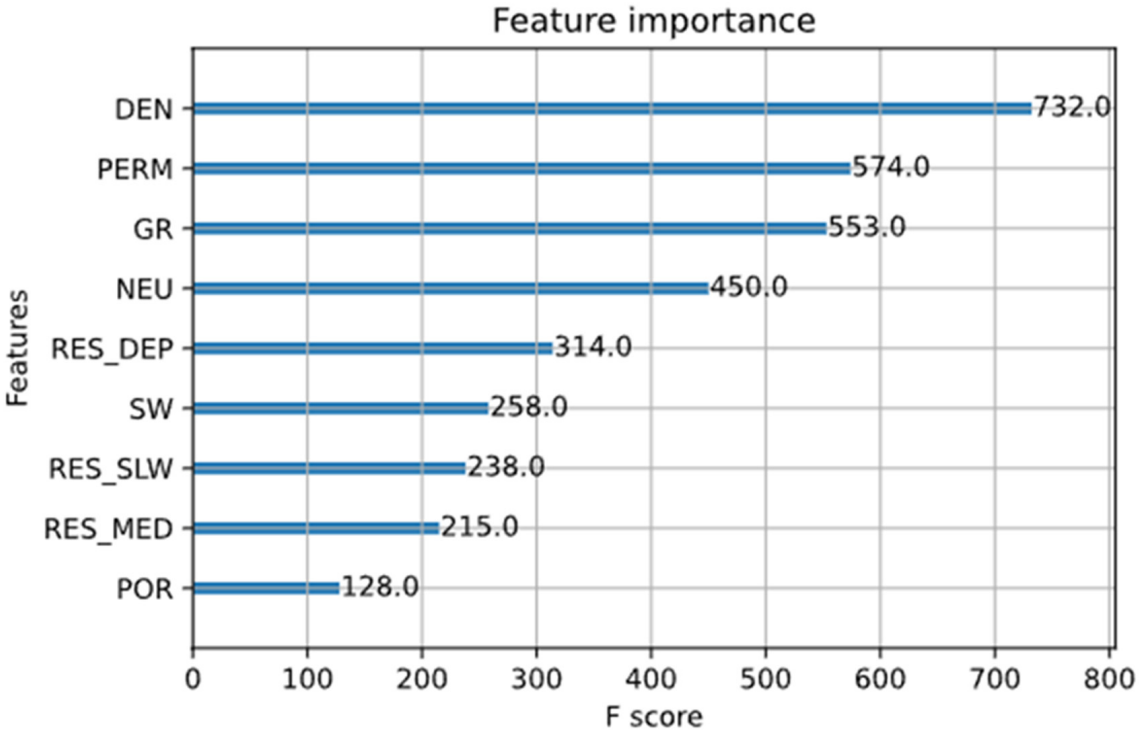


**Figure 11.** XGBoost feature importance in predicting NMR T2LM.

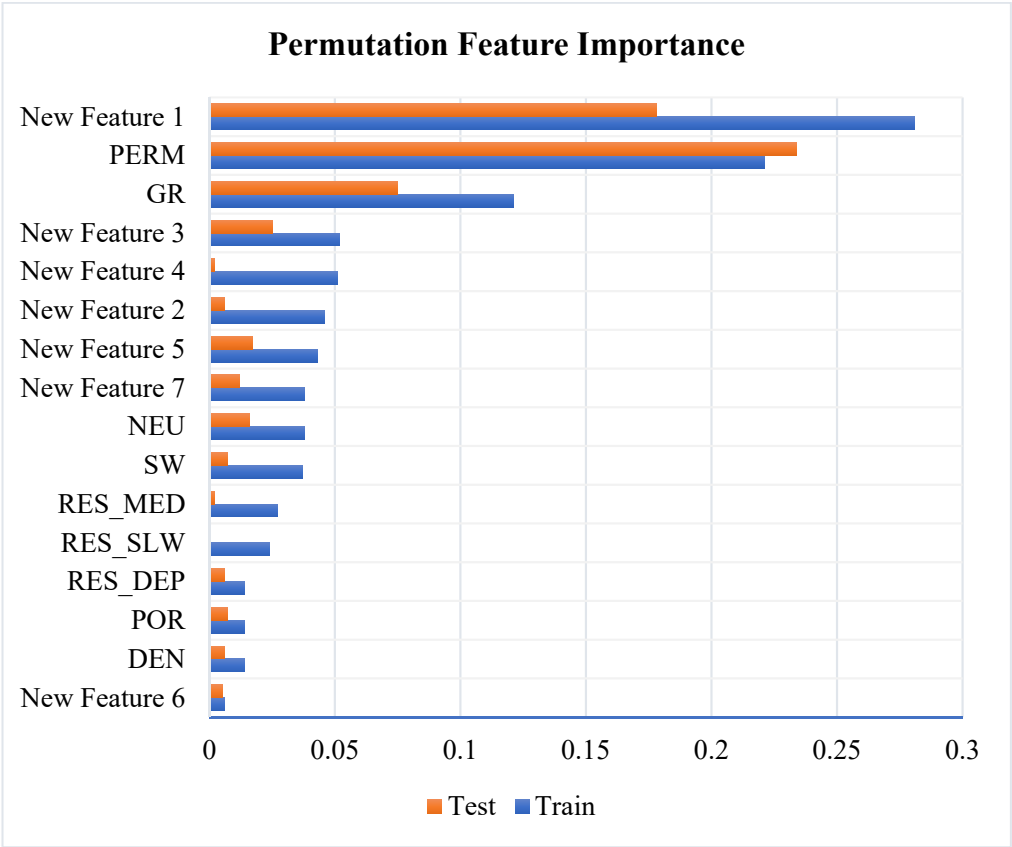


**Figure 12.** Permutation feature importance on all input variables for $T_2$LM prediction.

This outcome goes on to support the engineering of the new feature. As density relates to porosity and the spaces where wettability is formed, these logs should be paramount, whereas saturation relates to fluid interactions. Based on the hypothesis that these logs are important in determining wettability, they had inherent important information for the generalisation of the model. The new feature 1 with DEN and Sw incorporated achieved a much higher relevance score, improving the model's performance and confirming the hypothesis that they have important information that helped improve the model generalisation. By omitting this important feature engineering step and this critical feature would have caused the model to over- or under-fit, which should not be the case. Since there are variabilities from one model to the other on features considered the most important ones, feature engineering needs to be performed based on domain knowledge.

Further analysis was performed using 6 different models for the most important features to ascertain if the new features will be deemed as important when different models are used. This result is shown in Table 7. The models used are the LGBoost, AdaBoost, Bagging, Decision Tree, XGBoost and Extra Tree. All the top performing models assessed some features as not important. The most common irrelevant features are the DEN, Res_Dep and Res_Med logs. However, New feature 1, PERM and GR variables were commonly selected as important features.

**Table 7.** Computed feature importance and accuracy on test data using different models.

| Models | Top 10 Features | Accuracy |
|---|---|---|
| • LGBoost | PERM, GR, New feature 1, Sw, NEU, New feature 2, New feature 3, New feature 5, New feature 6, POR | 0.8056 |
| • AdaBoost | PERM, New feature 2, New feature 5, New feature 3, NEU, DEN, New feature 1, GR, POR, RES_SLW | 0.7119 |
| • Decision Tree | PERM, New feature 1, POR, New feature 7, New feature 2, New feature 5, RES_MED, GR, New feature 3, RES_SLW | 0.6204 |
| • Random Forest | PERM, New feature 1, GR, New feature 5, NEU, New feature 7, POR, New feature 3, New feature 2, DEN | 0.8244 |
| • Bagging | New feature 1, PERM, GR, POR, NEU, New feature 3, New feature 5, Sw, RES_DEP, New feature 7 | 0.7991 |
| • Extra Tree | PERM, New feature 5, GR, New feature 1, NEU, Sw, RES_SLW, POR, New feature 4, New feature 7 | 0.807 |
| • XGBoost | PERM, New feature 1, New feature 7, GR, New feature 2, RES_MED, POR, RES_DEP, New feature 4, New feature 6 | 0.7954 |

### *3.5. Sensitivity Analysis on Features*

Figure 13 shows further evaluation of the top 9 features which overall indicates their superiority over using only the initial features based on all the metrics. The Gradient Boosting model was used to analyse further the performance based on a different model. The results indicate an increase in model accuracy on test data when the top relevant features were selected as input. There was also a reduction in the model's error measurements. The performance of the new features is a testament that wettability, is dependent on the saturation and porosity dependent logs. These logs are quite significant in determining wettability as they resulted in the improvements of the models.

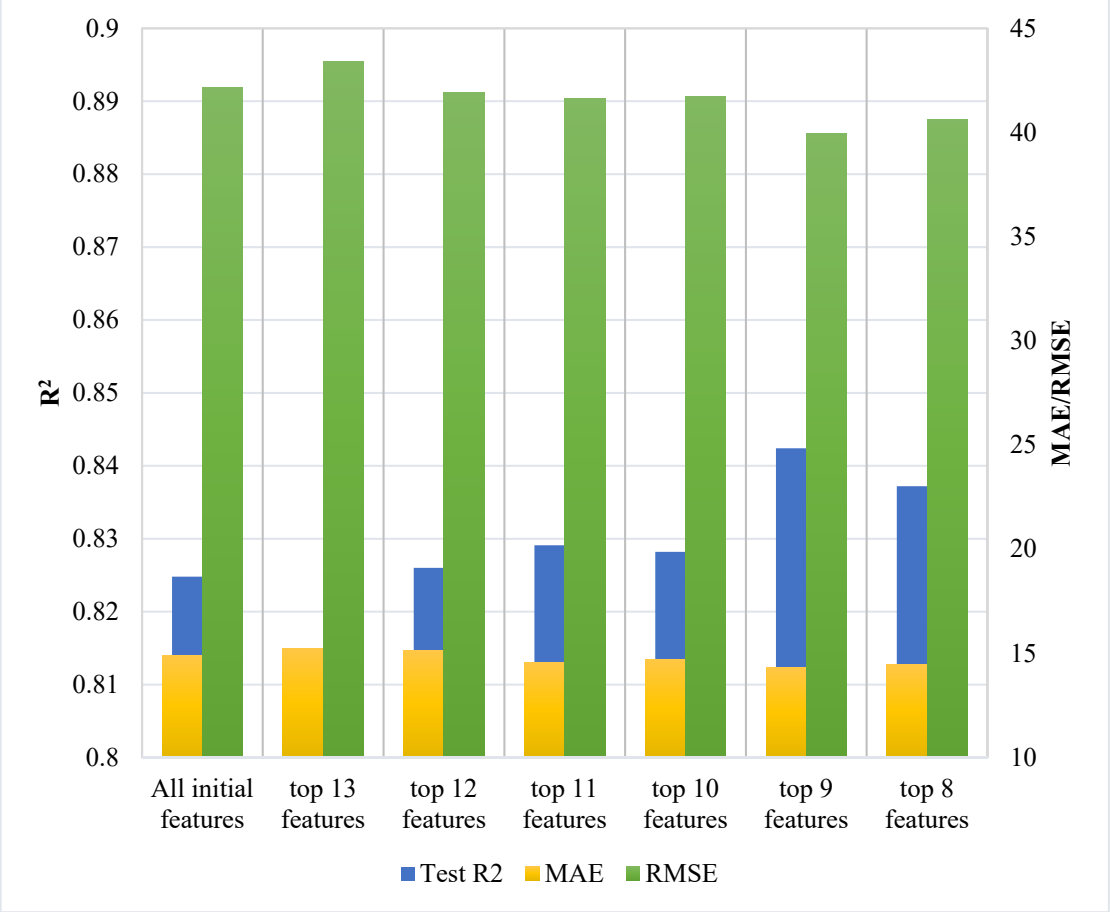


**Figure 13.** Evaluation of the performances of top features using the Gradient Boosting Regressor.

In selecting the optimal model fitting to the data set, all the features discussed in this research can be compared via error metrics. The selected top 9 features are GR, NEU, Sw, PERM, New feature 1, New feature 2, New feature 3, New feature 5 and New feature 7. This result indicates that to characterise an oil and gas reservoir in terms of NMR $T_2$LM, these features should be included as input variables. This inference shows how robust the likelihood of fit the new features is for wettability prediction. As a result, our proposed new features, using Sw, POR and PERM, provide the best likelihood of optimising model fit for NMR $T_2$LM predictions.

### *3.6. Evaluation of Deep Ensemble Super Learner in Predicting NMR $T_2$LM*

Considering the cost of running an NMR log and the success of machine learning models in reservoir characterisation, a deep ensemble super learner was used to characterise the Coniston well in terms of NMR $T_2$LM from wireline logs. The results in Figure 14 specify that in comparison to all selected individual base learners, the deep ensemble super learner exhibited an improved accuracy in NMR $T_2$LM prediction of the holdout data. Based on this alone, it is critical to note that even though none of the models overfitted when tested on the holdout data, the super learner relatively performed more robustly. Therefore, further analysis was carried out to evaluate the prediction errors of all the models.

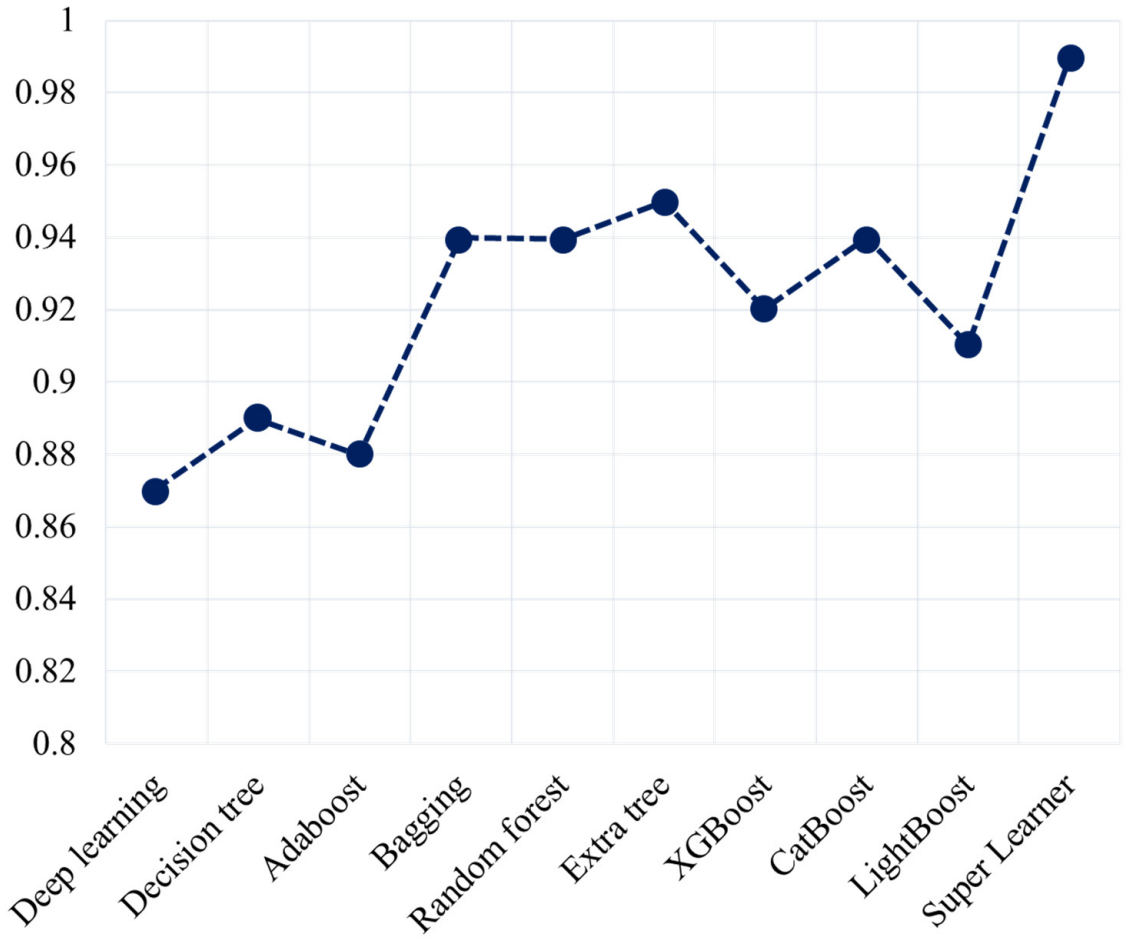


**Figure 14.** Comparison of correlation coefficient results of all machine learning models used in this study.

All the models' precision and consistency were evaluated based on the error measurements, as illustrated in Figure 15. RMSE, MPD, MAE and MAPE results indicate the Deep Ensemble Super Learner's superiority over all the base learners, indicating its consistency and precision in predicting NMR $T_2$LM. However, it is noteworthy to point out that the newly engineered features, based on further sensitivity analysis, improved the prediction performances of all the models. When the top 9 features were used in training the deep ensemble super learner model, it achieved a test $R^2$ of 0.998, MAE of 2.73, RMSE of 4.878, MAPE of 0.192 and MPD of 0.496. The second best model was the optimised Bagging regressor. This model achieved a test $R^2$ of 0.0.846, MAE of 14.7423, RMSE of 39.4848, MAPE of 0.45 and MPD of 9.288. When the Bagging is compared to the Deep Ensemble Super Learner, it shows a gargantuan disparity in terms of all the metrics. The superiority of the Deep Ensemble Super Learner is mainly attributed the stacking nature of the base learners where each error is passed through the base learners iteratively until the last model. The results are then averaged and returned as the final model prediction. Based on all the evaluation criteria, it can be concluded that the new features 1, 2, 3, 5 and 7 were highly relevant and improved model efficiency.

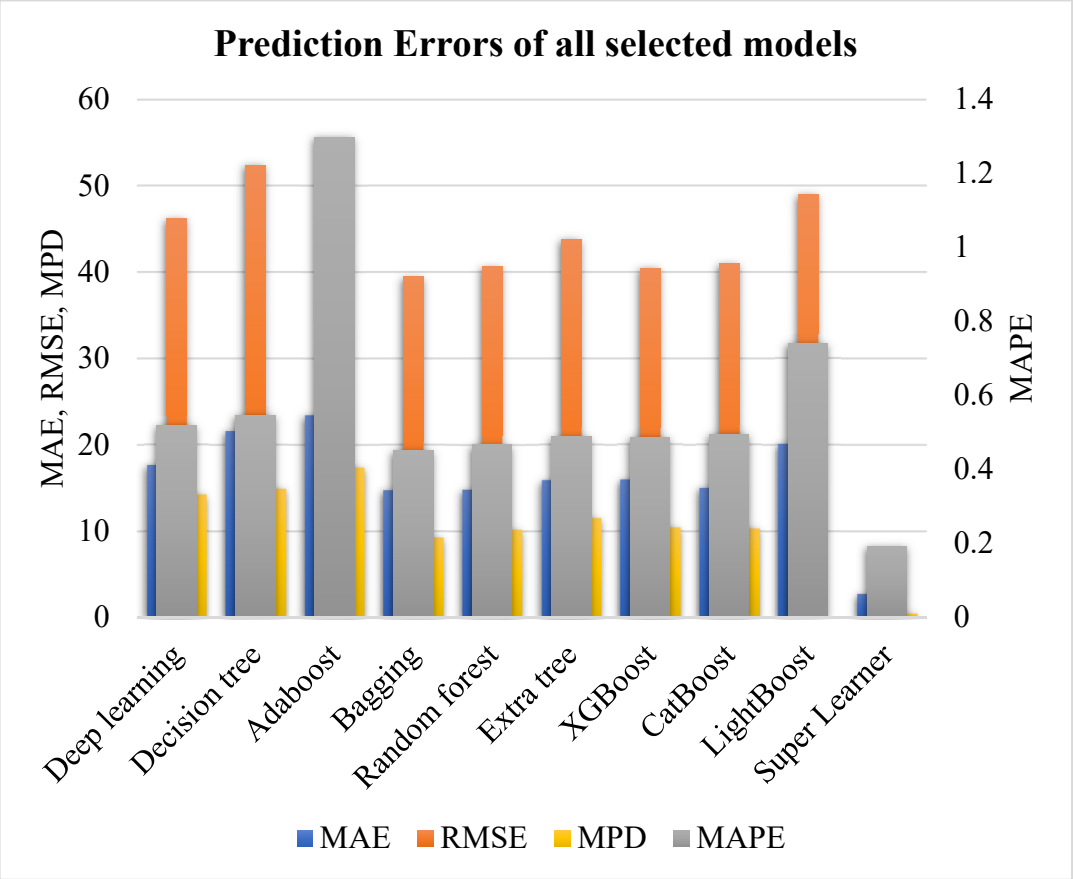


**Figure 15.** Comparison of prediction errors of all machine learning models based on MAE, RMSE, MAPE and MPD on the test data.

Figure 16 below illustrates the kernel density estimate of predicted and actual NMR $T_2$LM data. The predicted values are shown in blue, and the actual values are represented in green. From the observation on test data, the deep ensemble super learner predicted NMR $T_2$LM are much closer to the actual data. This analysis indicates the ability of the super learner to capture the diverse range of values better than the other models hence being appropriate for deployment in predicting NMR $T_2$LM prediction for other wells. The sensitivity analysis, therefore, confirms the evaluation metrics discussed above. The kernel density plot results is further analysed by comparing the actual data points to the predicted data points.

Further analysis shown in Figure 17 compares the predicted and actual data points with the above data points representing the Coniston reservoir. Based on the results, it was observed that the Deep Ensemble Super Learner tends to predict reasonably accurate NMR $T_2$LM data between 0 and 200 ms. This range refers to the water and oil phases. NMR $T_2$LM data points above 250 ms were slightly under-predicted hence contributing to the error evaluations discussed above. Ranges above 250 are typical of the gas phase in a reservoir. Since gas tends to be non-wetting, this under-prediction is relatively inconsequential. The main contributing factor to this under-prediction is the insufficient data within the gas interval of the Coniston well as well as the other reservoir. The gas cap in the Coniston well did not have many data points when compared to the water and oil zones, hence in training the model, a few points in the gas zone will be used. Also, the other well did not have any gas gap but rather light oil or oil with high gas-oil-ratio contributing to the prediction errors. Although the errors are low, it is necessary to understand how they occur and how to mitigate or analyse such results when deployed on different wells that may has a gas interval. With that said, gases tend to be water wet hence the inaccuracies occurring

above $T_2$LM 250 ms which refers to the gas phase will not impact the wettability prediction in any way.

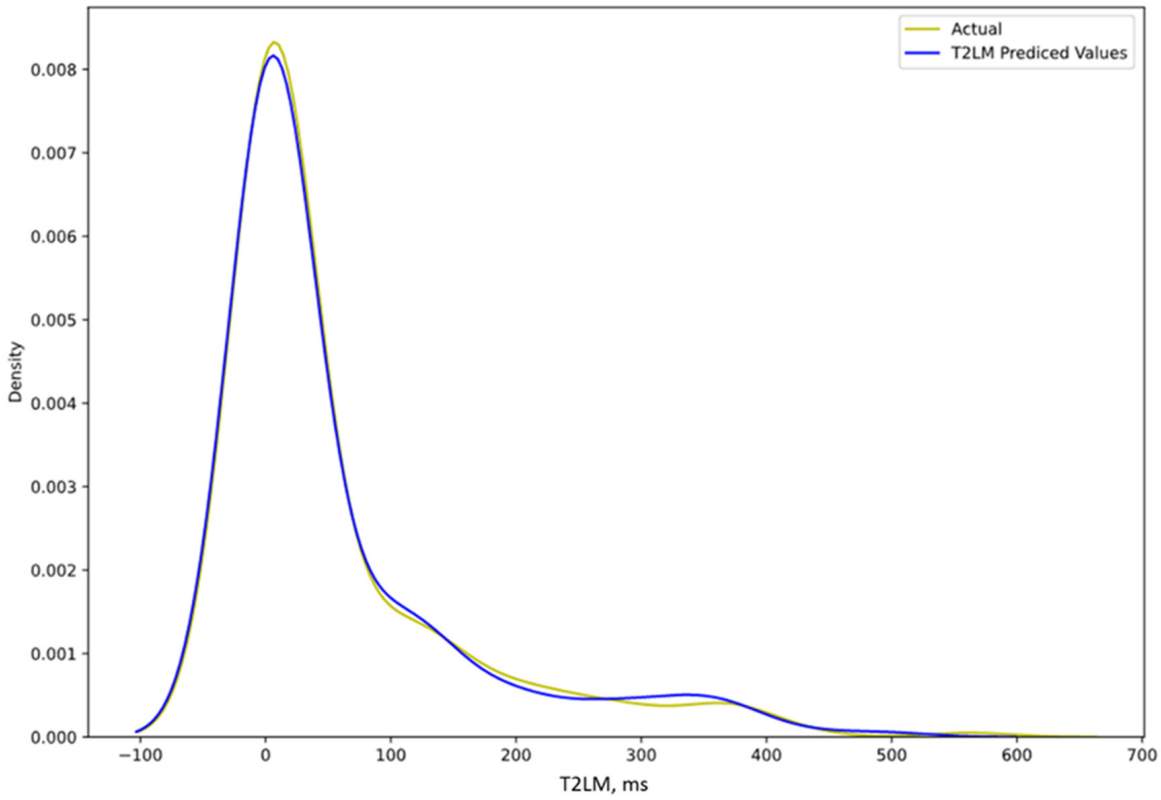


**Figure 16.** Kernel density estimation showing the closeness of predicted to actual NMR $T_2$LM values.

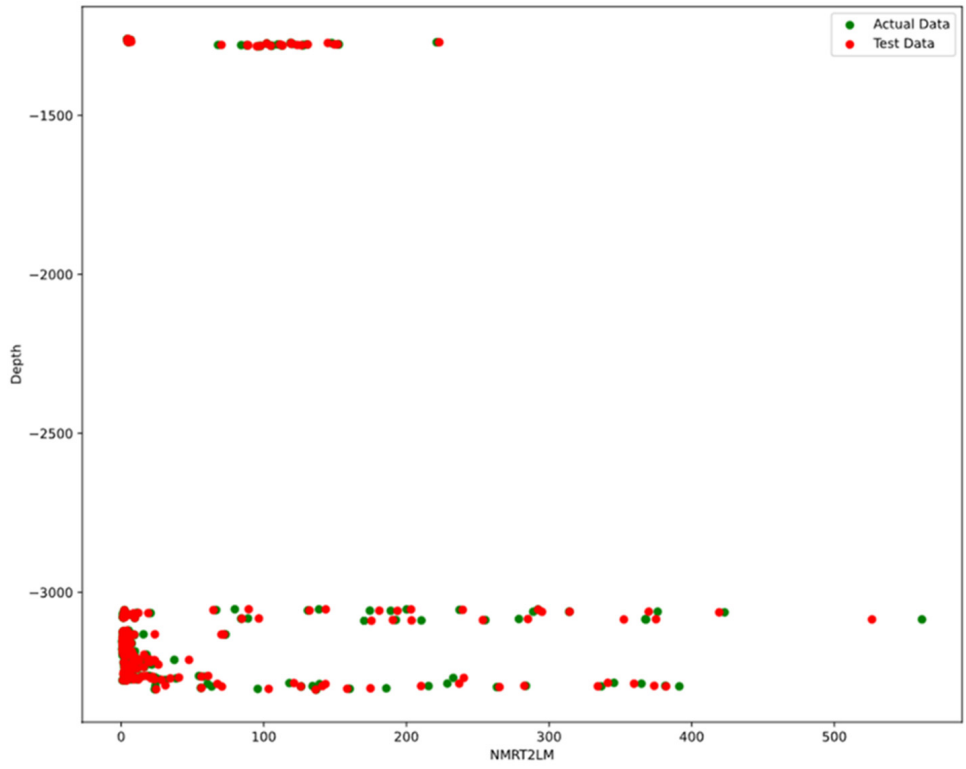


**Figure 17.** Comparison of predicted and test data for the two reservoir intervals.

The excellent performance of the Deep Ensemble Super Learner model is mainly attributed to the stacking of a Deep Neural Network and eight other highly performing ensemble models. Training a Deep Neural Network on small and low dimensional datasets tends to cause memorisation of the data. This typically leads to poor performance of the model on the holdout data. The ineffectiveness of Deep Neural Network regarding small data and its inability to interpret results verifies that it cannot be overly considered the best solution in all cases. However, with ensemble models capable of fitting small datasets, combining a Deep Neural Network in a super learner makes sure both models complement each other. As a downside, the training of the deep ensemble super learner is prohibitively expensive due to the K-fold predictions covering the whole training dataset. However, computational time was not the considering evaluation factor when the model's performance could be significantly improved. Additionally, when the training data has a high variance or is noisy, the Deep Ensemble Super Learner can utilise all the data necessary for training due to the stacking of diverse base learners. This forms the basis of its superior performance over all the individual base learners.

### *3.7. Field Deployment of Deep Ensemble Super Learner and Wettability Determination Methodology*

The proposed methodology will be applied to the entire sandstone reservoir interval of a well from south-eastern coast of the Arabian Peninsula. Field deployment of the proposed methodology is relevant in evaluating the final model. As such, the Deep Ensemble Super Learner is used to predict the $T_2$LM for the well. The well has two different reservoirs which is the focus in this deployment. After the prediction, the developed methodology and mathematical model are employed to predict the wettability of the reservoir. The predicted wettability is then compared to the initial wettability established by relative permeability data from the reservoir. For wettability to be established using relative permeability, the following condition needs to be established:

1 When relative permeability to oil for oil/water displacement (krow) > relative permeability to water (krw) at 50% water saturation, the reservoir is interpreted as water wet.
2 When krow < krw at 50 % water saturation, the reservoir is interpreted as oil wet.
3 When krow = krw at 50 % water saturation, the reservoir is interpreted as neutral wet.

To account for uncertainty, it is relevant to note the reservoir types used in this research. One of the reservoirs is a vertical sequence of Early Permian clastic deposits. It has been categorized into three sections. The upper two sections are exclusively fluvial deposits, whilst the lowermost part is mostly deltaic. The Coniston field is made up of Upper Barrow Group sandstone from the Berrasian period, with interbedded minor arenaceous sandstone that grades into interbedded siltstone at the bottom. Towards the end of the Jurassic, the reservoir was deposited by the erosion and re-deposition northeast-flowing river systems. This caused parasitic deltaic wedges to be deposited north and west of Coniston.

The reservoir data used to train and test the models were from fluvial and deltaic depositional environments. The reservoir that the methodology is being deployed was also deposited by the same depositional system and is an oil bearing sandstone. The Deep Ensemble Super Learner was deployed on the well to predict the $T_2$LM of the reservoir interval. The predicted $T_2$LM, shown in Figures 18 and 19, is plotted along with some logs to analyse its behaviour in the reservoir. The predicted $T_2$LM data was then plotted against to SW data as shown in Figure 20. The well was further subdivided into their distinct reservoirs with the upper reservoir having 2 zones. The results of this sub-division is shown in Figure 21. From the plots, it can be observed qualitatively that all the net pay intervals are water wet. Table 8 shows the quantitative wettability of the interval when the proposed equation is used. Based on the results, the shallower reservoir exhibited moderately water-wet conditions. The deeper reservoir also exhibited weakly water wet condition. The weakly water-wet condition is results from the micropores being wet by oil whiles water wets the macropores. This situation occurs due to low water saturation and low $T_2$LM refers to intervals of low permeability as well as a low free fluid index. With

$T_2$LM having a direct relationship with these parameters, it may be deduced to be the cause of the low $T_2$LM value in that interval.

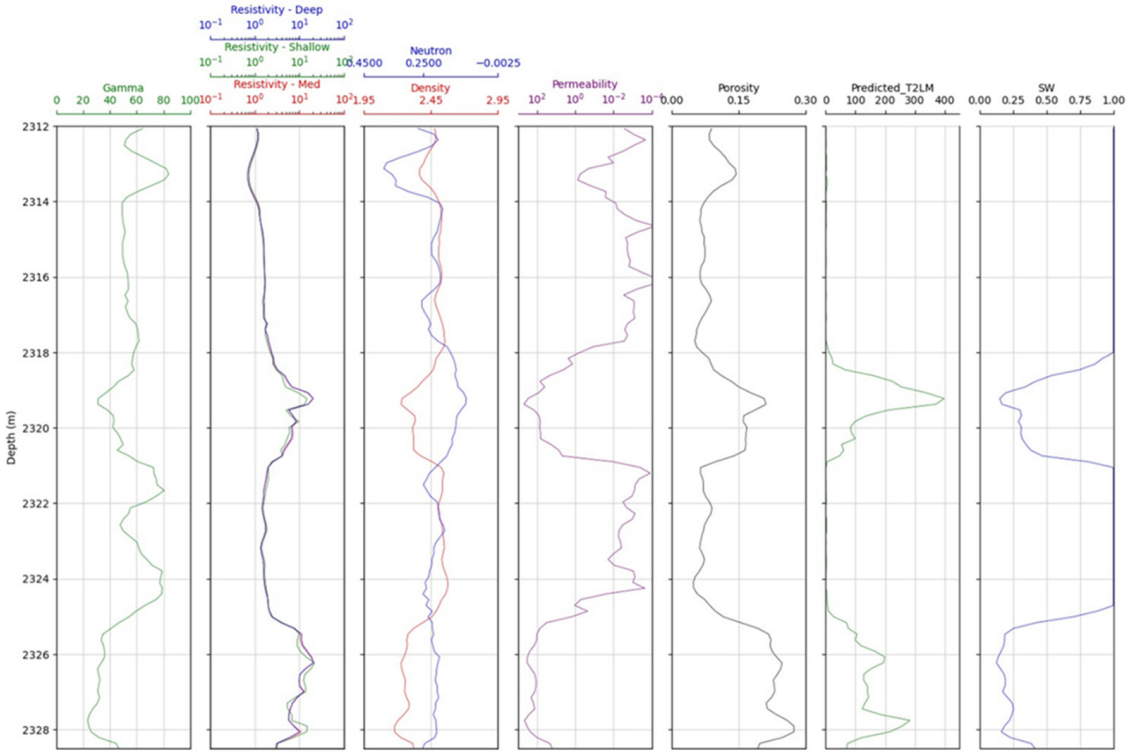


**Figure 18.** Predicted NMR $T_2$LM for Reservoir A in new well.

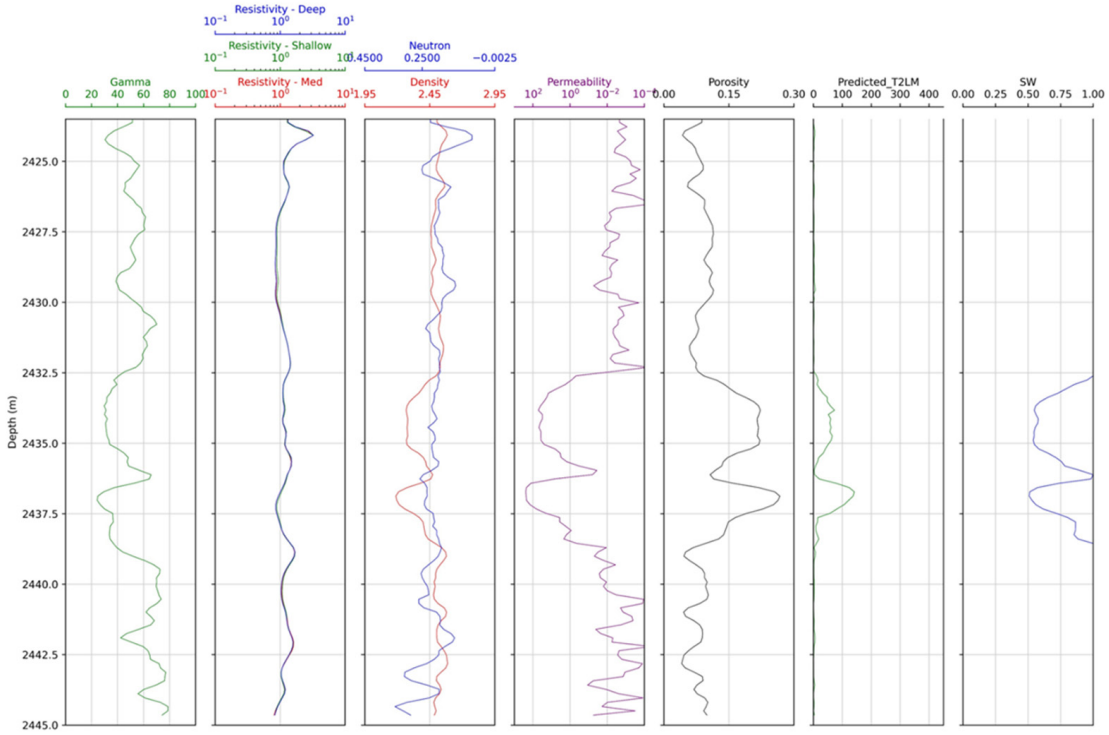


**Figure 19.** Predicted NMR $T_2$LM for Reservoir B in new well.

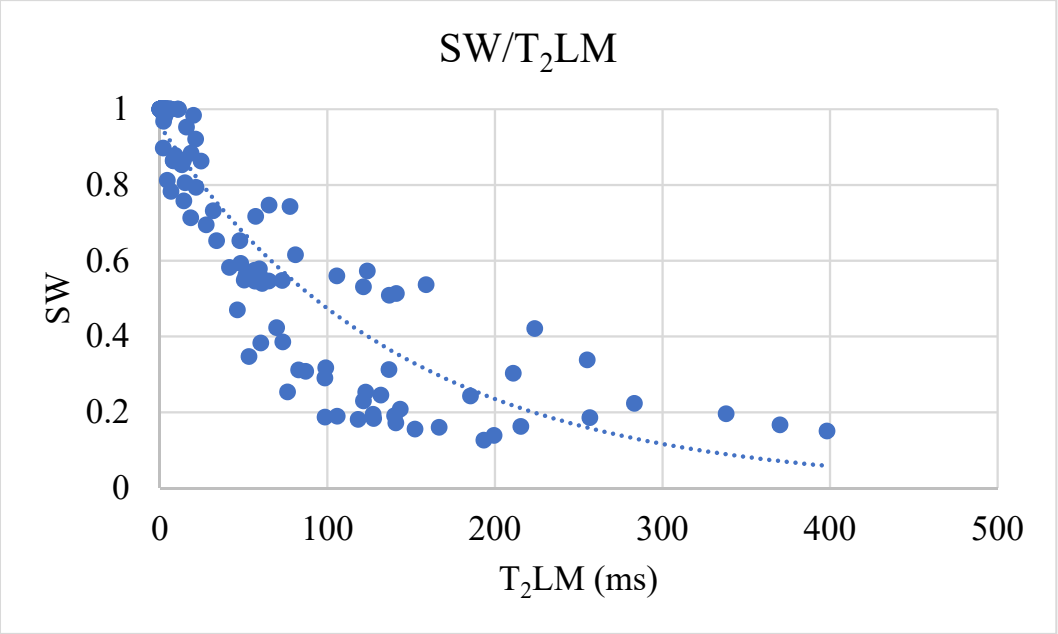


**Figure 20.** Plot showing the SW against $T_2$LM trends in the new well.

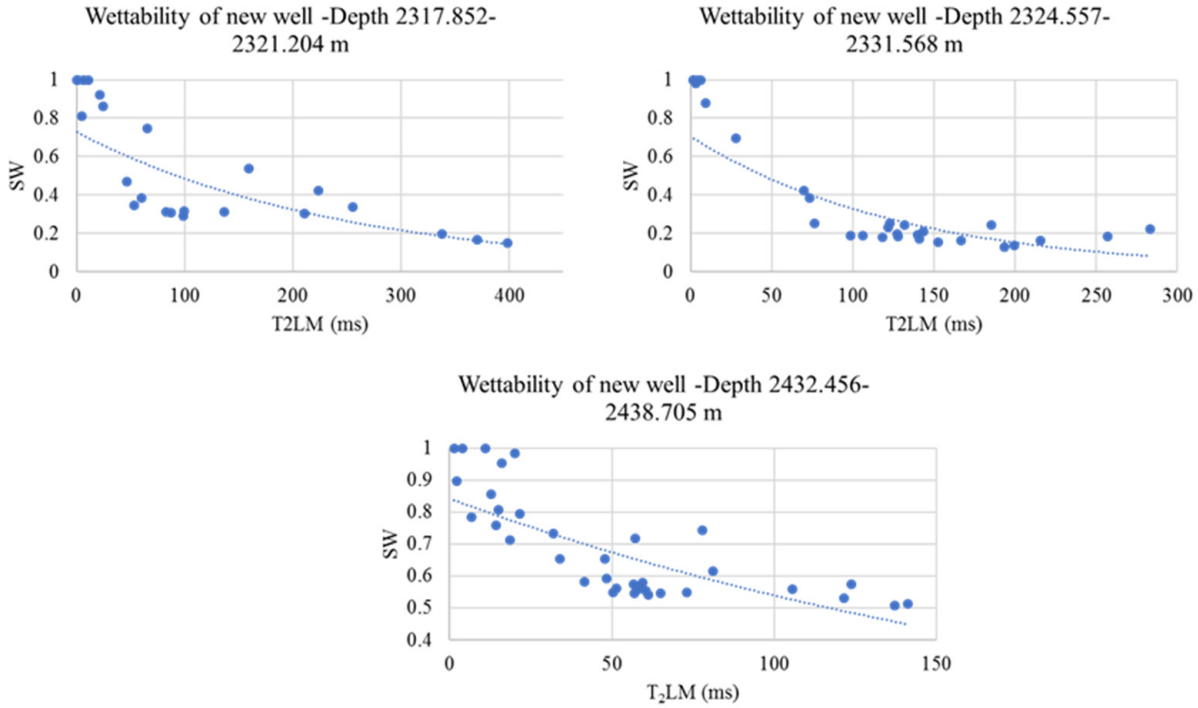


**Figure 21.** Plot of figures showing the SW against $T_2$LM trends in different oil intervals.

**Table 8.** Quantitative evaluation of wettability conditions.

| Depth from m | Depth to m | Average $T_2$LM | Estimated Amott Index | Wettability Indicated in this Study |
|---|---|---|---|---|
| 2317.852 | 2321.204 | 119.742 | 0.3 | Moderately water-wet |
| 2324.557 | 2331.568 | 113.973 | 0.3 | Moderately water-wet |
| 2432.456 | 2438.705 | 61.339 | 0.1 | Weak water-wet |

The relative permeability of the reservoir, shown in Figure 22, confirms the general wettability of the field as water-wet. This confirmation can only state one wetting condition unlike the proposed empirical and machine learning methodology that can quantitatively indicate the different types of wettability in existence.

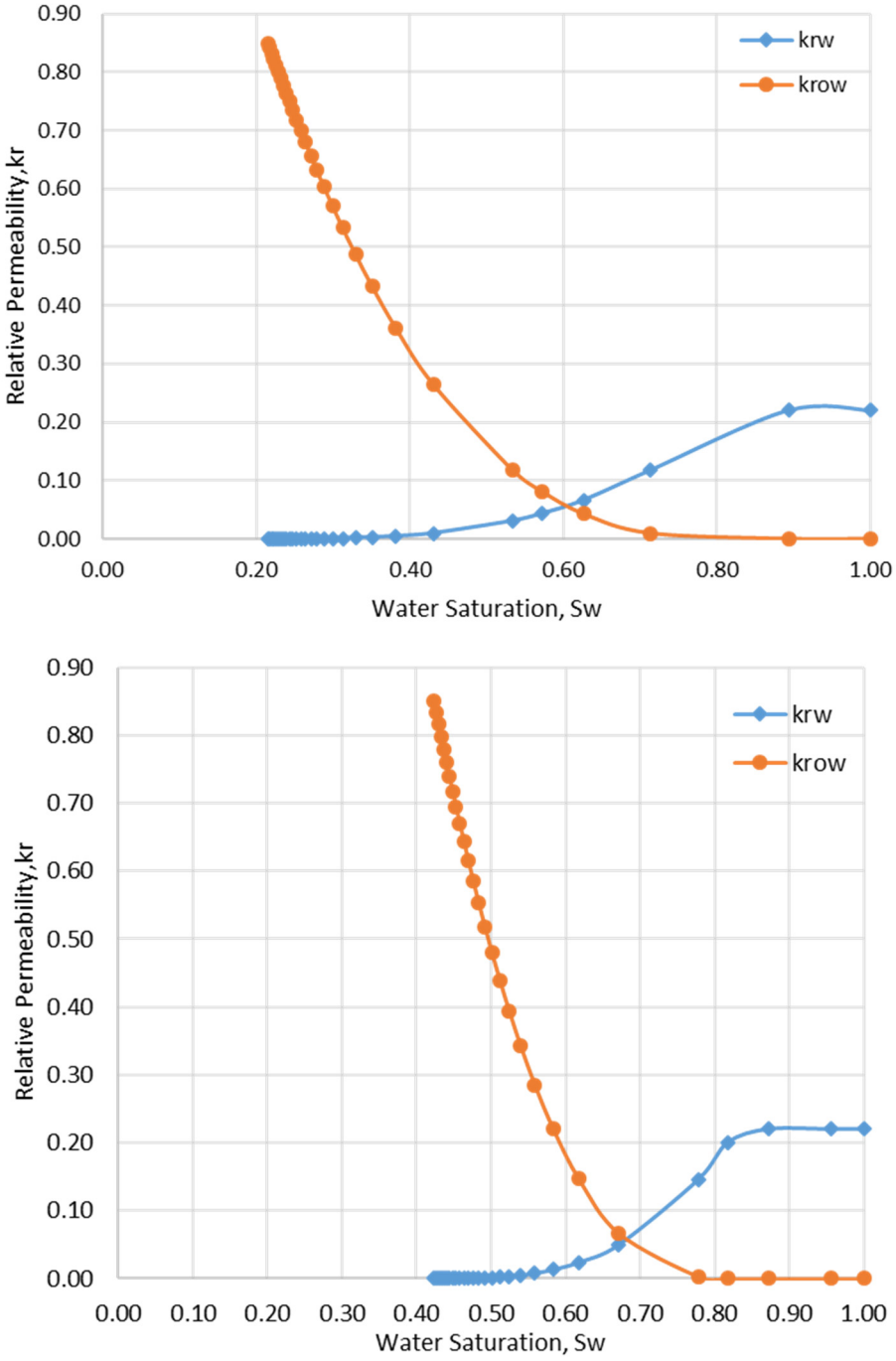


**Figure 22.** Oil-Water relative permeability curves for 2 core samples from the reservoir A (above) and B (below) intervals.

## 4. Conclusions

Being able to quantify the surface wettability in an early stage is vital for every reservoir. This can give detailed information and advice on the EOR, especially CEOR, to increase production. This research proposes a novel empirical formula for predicting surface wettability in the field. There is also the proposal of a novel deep ensemble super learner that can predict NMR $T_2$LM. The NMR $T_2$LM, when analysed against water saturation data, can indicate the type of wettability condition existing in a reservoir. The average $T_2$LM can then be calculated where the mathematical formula is used to predict the type of wetting condition in existence quantitatively. The integration of these computational techniques thus gives a better indication of surface wettability in the absence of any core wettability data and NMR logs.

Focusing on the findings of this study, the following conclusions can be drawn:

1 An exponential relationship was established between laboratory Amott-USBM wettability index and field NMR $T_2$LM, achieving a high $R^2$ of 0.95.
2 An exponential function was used to formulate a mathematical formula that captures domain knowledge on the existence of 100% water but not oil with the net reservoir zones.
3 The relationship between Sw and NMR $T_2$LM was used to understand surface wettability conditions in the reservoir. The mean NMR T2LM was calculated and used in the mathematical equation to predict wettability based on increasing or decreasing Sw.
4 Since surface wettability occurs in pore spaces and between two immiscible fluids, a new feature was engineered using DEN and Sw. This feature improved the performances of all the machine learning models used in this study.
5 The performance of the deep ensemble super learner was superior to the deep learning and ensemble models in predicting NMR $T_2$LM. Therefore, the deep ensemble super learner could predict NMR $T_2$LM when there are money and time constraints.
6 The deep ensemble super learner is robust and flexible enough to perform better on its own without feature selection.
7 Field deployment of the proposed methodology and deep ensemble super learner was able to successfully estimate the wettability of 3 different reservoir intervals. This result was validated using relative permeability plots.

**Author Contributions:** Conceptualization, D.A.O.; methodology, D.A.O.; software, D.A.O.; validation, D.A.O. and M.A.A.M.; formal analysis, D.A.O.; investigation, D.A.O.; resources, D.A.O.; data curation, R.G.; writing—original draft preparation, D.A.O.; writing—review and editing, D.A.O., T.O.A.G., M.A.A.M., R.G. and Z.M.A.M.; visualization, D.A.O.; supervision, T.O.A.G., M.A.A.M., R.G. and Z.M.A.M.; project administration, D.A.O.; funding acquisition, T.O.A.G. and M.A.A.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Y-UTP grant number [015LCO-105].

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The data and Jupyter Notebooks used in this study is hosted at https://github.com/ascotjnr/A-Deep-Ensemble-Super-Learner-to-Predict-Reservoir-Wettability.

**Acknowledgments:** The authors express their sincere appreciation to University Teknologi Petronas and the Centre of Research in Enhanced Oil recovery for financially supporting this work through Y-UTP grant (015LCO-105).

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Craig, F.F. *The Reservoir Engineering Aspects of Waterflooding*; Henry L. Doherty Memorial Fund of AIME: Dallas, TX, USA, 1971; ISBN 0895202026.
2. Anderson, W.G. Wettability Literature Survey-Part 1: Rock/Oil/Brine Interactions and the Effects of Core Handling on Wettability. *JPT J. Pet. Technol.* **1986**, *38*, 1125–1144. [CrossRef]

3. Johannesen, E.B.; Graue, A.; Baldwin, B.A.; Tobola, D.P. Establishing Mixed Wet Conditions in Chalk-Emphasis on Wettability Alteration and Oil Recovery. In Proceedings of the International Symposium of the Society of Core Analysts, Calgary, AB, Canada, 10–12 September 2007.
4. Anderson, W.G. Wettability Literature Survey—Part 2: Wettability Measurement. *J. Pet. Technol.* **1986**, *38*, 1246–1262. [CrossRef]
5. Valori, A.; Nicot, B. A Review of 60 Years of NMR Wettability. *Soc. Petrophysicists Well-Log. Anal.* **2019**, *60*, 255–263. [CrossRef]
6. Morrow, N.R. Wettability and Its Effect on Oil Recovery. *J. Pet. Technol.* **1990**, *42*, 1476–1484. [CrossRef]
7. Salathiel, R.A. Oil Recovery by Surface Film Drainage in Mixed-Wettability Rocks. *J. Pet. Technol.* **1973**, *25*, 1216–1224. [CrossRef]
8. Fleury, M.; Deflandre, F. Quantitative evaluation of porous media wettability using NMR relaxometry. *Magn. Reson. Imaging* **2003**, *21*, 385–387. [CrossRef]
9. Amott, E. Observations Relating to the Wettability of Porous Rock. *Trans. AIME* **1959**, *216*, 156–162. [CrossRef]
10. Donaldson, E.C.; Thomas, R.D.; Lorenz, P.B. Wettability Determination and Its Effect on Recovery Efficiency. *Soc. Pet. Eng. J.* **1969**, *9*, 13–20. [CrossRef]
11. Looyestijn, W.J.; Hofman, J.P. Wettability-index determination by nuclear magnetic resonance. *SPE Reserv. Eval. Eng.* **2006**, *9*, 146–153. [CrossRef]
12. Coates, G.R.; Xiao, L.; Prammer, M.G. *NMR Logging Principles and Applications*; Halliburton Energy Services Publication: Houston, TX, USA, 1999.
13. Tandon, S.; Rostami, A.; Heidari, Z. A New NMR-Based Method for Wettability Assessment in Mixed-Wet Rocks. In Proceedings of the Proceedings—SPE Annual Technical Conference and Exhibition; Society of Petroleum Engineers (SPE), San Antonio, TX, USA, 9–11 October 2017.
14. Otchere, D.A.; Ganat, T.O.A.; Gholami, R.; Lawal, M. A Novel Custom Ensemble Learning Model for an Improved Reservoir Permeability and Water Saturation Prediction. *J. Nat. Gas. Sci. Eng.* **2021**, *91*, 103962. [CrossRef]
15. Otchere, D.A.; Ganat, T.O.A.; Gholami, R.; Ridha, S. Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models. *J. Pet. Sci. Eng.* **2021**, *200*, 108–182. [CrossRef]
16. Abdallah, W.; Buckley, J.S.; Carnegie, A.; Edwards, J.; Herold, B.; Fordham, E.; Graue, A.; Habashy, T.; Seleznev, N.; Signer, C.; et al. Fundamentals of wettability. *Oilf. Rev.* **2007**, *19*, 44–61.
17. Iglauer, S.; Pentland, C.H.; Busch, A. $CO_2$ wettability of seal and reservoir rocks and the implications for carbon geo-sequestration. *Water Resour. Res.* **2015**, *51*, 729–774. [CrossRef]
18. Feng, C.; Fu, J.; Shi, Y.; Li, G.; Mao, Z. Predicting reservoir wettability via well logs. *J. Geophys. Eng.* **2016**, *13*, 234–241. [CrossRef]
19. Archie, G.E. The Electrical Resistivity Log as an Aid in Determining Some Reservoir Characteristics. *Trans. AIME* **1942**, *146*, 54–62. [CrossRef]
20. Valori, A.; Ali, F.; Abdallah, W. Downhole wettability: The potential of NMR. In Proceedings of the Society of Petroleum Engineers—SPE EOR Conference at Oil and Gas West Asia 2018; Society of Petroleum Engineers, Muscat, Oman, 26 March 2018.
21. Sondenaa, E.; Bratteli, F.; Normann, H.P.; Kollveit, K. Effect of reservoir conditions and wettability on electrical resistivity. In Proceedings of the SPE Asia Pacific Conference, Perth, Australia, 4–7 November 1991; pp. 409–422.
22. Amani, M.; Al-Jubouri, M.; Khadr, S.; Sayed, A. A Comprehensive Review on the Use of NMR Technology in Formation Evaluation. *Comput. Sci.* 2017. Available online: https://www.semanticscholar.org/paper/A-Comprehensive-Review-on-the-Use-of-NMR-Technology-Amani-Al-Jubouri/fe15a754ea33ae2a3c1f2fa689b81f067b398fef (accessed on 19 December 2021).
23. Branco, F.R.; Gil, N.A. NMR study of carbonates wettability. *J. Pet. Sci. Eng.* **2017**, *157*, 288–294. [CrossRef]
24. Looyestijn, W.; Zhang, X.; Hebing, A. How Can NMR Assess the Wettability of a Chalk Reservoir. In Proceedings of the International Symposium of the Society of Core Analysts, Vienna, Austria, 27–30 August 2017.
25. Brown, R.J.S.; Fatt, I. Measurements of Fractional Wettability of Oil Fields' Rocks by The Nuclear Magnetic Relaxation Method. In Proceedings of the Fall Meeting of the Petroleum Branch of AIME, Los Angeles, CA, USA, 14–17 October 1956.
26. Howard, J.J. Quantitative estimates of porous media wettability from proton NMR measurements. In Proceedings of the Magnetic Resonance Imaging, Dallas, TX, USA, 24–26 September 1998; Volume 16, pp. 529–533.
27. Looyestijn, W. Practical Approach to Derive Wettability Index by NMR in Core Analysis Experiments. *Petrophysics* **2019**, *60*, 507–513. [CrossRef]
28. Kenyon, B.; Kleinberg, R.L.; Straley, C.; Gubelin, G.; Morriss, C. Nuclear magnetic resonance imaging—technology for the 21st century. *Oilf. Rev.* **1995**, *7*, 19–33.
29. Zhang, G.Q.; Huang, C.C.; Hirasaki, G.J. Interpretation of wettability in sandstones with NMR analysis. *Petrophysics* **2000**, *41*, 223–233.
30. Guan, H.; Brougham, D.; Sorbie, K.S.; Packer, K.J. Wettability effects in a sandstone reservoir and outcrop cores from NMR relaxation time distributions. *J. Pet. Sci. Eng.* **2002**, *34*, 35–54. [CrossRef]
31. Borgia, G.C.; Fantazzini, P.; Mesini, E. Wettability effects on oil-water-configurations in porous media: A nuclear magnetic resonance relaxation study. *J. Appl. Phys.* **1991**, *70*, 7623–7625. [CrossRef]
32. Chen, J.; Hirasaki, G.J.; Flaum, M. NMR wettability indices: Effect of OBM on wettability and NMR responses. *J. Pet. Sci. Eng.* **2006**, *52*, 161–171. [CrossRef]
33. Chen, P.; Katheeri, A.H.A.; Kalam, M.Z.; Shtepani, E. Integration of Multi-Scale Techniques to Evaluate Reservoir Wettability for Carbonate Reservoirs in Middle East. In Proceedings of the International Symposium of the Society of Core Analysts, Vienna, Austria, 27 August–1 September 2017.

34. Ali, J.K. Neural networks: A new tool for the petroleum industry? In Proceedings of the Society of Petroleum Engineers—European Petroleum Computer Conference 1994, Aberdeen, UK, 15–17 March 1994; pp. 233–242.
35. Al-Bazzaz, W.H.; Al-Mehanna, Y.W.; Gupta, A. Permeability modeling using neural-network approach for complex Mauddud-Burgan carbonate reservoir. In Proceedings of the SPE Middle East Oil and Gas Show and Conference, Manama, Bahrain, 11–14 March 2007; Volume 2, pp. 892–906.
36. Anifowose, F.A.; Abdulraheem, A.; Al-Shuhail, A.A.; Schmitt, D.P. Improved Permeability Prediction from Seismic and Log Data using Artificial Intelligence Techniques. In Proceedings of the SPE Middle East Oil and Gas Show and Conference, Manama, Bahrain, 10–13 March 2013; Volume 3, pp. 2190–2196.
37. Bello, O.; Asafa, T. A functional networks softsensor for flowing bottomhole pressures and temperatures in multiphase flow production wells. In Proceedings of the Society of Petroleum Engineers—SPE Intelligent Energy International 2014, Utrecht, The Netherlands, 1–3 April 2014; pp. 637–651.
38. Helle, H.B.; Bhatt, A. Fluid saturation from well logs using committee neural networks. *Pet. Geosci.* **2002**, *8*, 109–118. [CrossRef]
39. Hamada, G.M.; Elshafei, M.A. Neural network prediction of porosity and permeability of heterogeneous gas sand reservoirs. In Proceedings of the Society of Petroleum Engineers—SPE Saudi Arabia Section Technical Symposium, Al-Khobar, Saudi Arabia, 9–11 May 2009; Volume 10, pp. 1113–1124.
40. Tackie-Otoo, B.N.; Atta, D.Y.; Mohammed, M.A.A.; Otchere, D.A. Investigation into the Oil Recovery Process Using an Organic Alkali–Amino Acid-Based Surfactant System. *Energy Fuels* **2021**, *35*, 11171–11192. [CrossRef]
41. Otchere, D.A.; Ganat, T.O.A.; Ojero, J.O.; Taki, M.Y.; Tackie-Otoo, B.N. Application of gradient boosting regression model for the evaluation of feature selection techniques in improving reservoir characterisation predictions. *J. Pet. Sci. Eng.* **2021**, *208*, 109244. [CrossRef]
42. Dorogush, A.V.; Ershov, V.; Gulin, A. CatBoost: Gradient boosting with categorical features support. *arXiv* **2018**, arXiv:1810.11363.
43. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. *Adv. Neural Inf. Proc. Syst.* **2017**, *30*, 3146–3154.
44. Gordon, A.D.; Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. Classification and Regression Trees. *Biometrics* **1984**, *40*, 874. [CrossRef]
45. Bengio, Y. Learning deep architectures for AI. *Found. Trends Mach. Learn.* **2009**, *2*, 1–27. [CrossRef]
46. Freund, Y.; Schapire, R.E. A Decision-Theoretic Generalisation of On-Line Learning and an Application to Boosting. In *Proceedings of the Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; Springer: Barcelona, Spain, 13–15 March 1995; Volume 904, pp. 23–37.
47. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [CrossRef]
48. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
49. Geurts, P.; Ernst, D.; Wehenkel, L. Extremely randomised trees. *Mach. Learn.* **2006**, *63*, 3–42. [CrossRef]
50. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; ACM: New York, NY, USA, 2016; pp. 785–794.
51. Flennerhag, S.; Moreno, P.G.; Lawrence, N.D.; Damianou, A. Transferring Knowledge across Learning Processes. *arXiv* **2018**, arXiv:1812.01054.
52. Pedregosa, F.; Michel, V.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Vanderplas, J.; Cournapeau, D.; Pedregosa, F.; Varoquaux, G.; et al. Scikit-learn: Machine Learning in Python Gaël Varoquaux Bertrand Thirion Vincent Dubourg Alexandre Passos PEDREGOSA, VAROQUAUX, GRAMFORT ET AL. Matthieu Perrot. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.
53. Van Der Laan, M.J.; Rose, S. Sequential Super Learning. In *Targeted Learning in Data Science*; Springer International Publishing: New York, NY, USA, 2018; pp. 27–34.
54. Van Der Laan, M.J.; Polley, E.C.; Hubbard, A.E. Super learner. *Stat. Appl. Genet. Mol. Biol.* **2007**, 6. [CrossRef]
55. Ledell, E.; Petersen, M.; Van Der Laan, M. Computationally efficient confidence intervals for cross-validated area under the ROC curve estimates. *Electron. J. Stat.* **2015**, *9*, 1583–1607. [CrossRef]
56. How to Develop Super Learner Ensembles in Python. Available online: https://machinelearningmastery.com/super-learner-ensemble-in-python/?unapproved=651132&moderation-hash=7795452e37f54a58f67726270b3881c2#comment-651132 (accessed on 18 January 2022).
57. Wang, X.; Wang, Z.; Feng, C.; Zhu, T.; Zhang, N.; Feng, Z.; Zhong, Y. Predicting oil saturation of tight conglomerate reservoirs via well logs based on reconstructing nuclear magnetic resonance *T2* spectrum under completely watered conditions. *J. Geophys. Eng.* **2020**, *17*, 17–328. [CrossRef]
58. Johannesen, E.B.; Riskedal, H.; Tipura, L.; Howard, J.J.; Graue, A. Wettability Characterization by Nmr *T2* Measurements in Edwards Limestone Rock. In Proceedings of the International Symposium of the Society of Core Analysts, Calgary, AB, Canada, 10–12 September 2007; pp. 1–12.
59. Straley, C. An Experimental Investigation of Methane in Rock Materials. In Proceedings of the SPWLA 38th Annual Logging Symposium, Houston, TX, USA, 15–18 June 1997.