Formation water geochemistry for carbonate reservoirs in Ordos basin, China: Implications for hydrocarbon preservation by machine learning

Citation for published version:
https://doi.org/10.1016/j.petrol.2019.106673

Digital Object Identifier (DOI):
10.1016/j.petrol.2019.106673

Link:
Link to publication record in Heriot-Watt Research Portal

Document Version:
Peer reviewed version

Published In:
Journal of Petroleum Science and Engineering

Publisher Rights Statement:
© 2019 Elsevier B.V.

General rights
Copyright for the publications made accessible via Heriot-Watt Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
Heriot-Watt University has made every reasonable effort to ensure that the content in Heriot-Watt Research Portal complies with UK legislation. If you believe that the public display of this file breaches copyright please contact open.access@hw.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Formation water geochemistry for carbonate reservoirs in Ordos basin, China: Implications for hydrocarbon preservation by machine learning

Hongyan Yu\textsuperscript{1, 2, 3*}, Zhenliang Wang\textsuperscript{1}, Reza Rezaee\textsuperscript{3}, Yihuai Zhang\textsuperscript{4*}, Lezorgia. N. Nwidee\textsuperscript{3}, Xi Liu\textsuperscript{1}, Michael Verrall\textsuperscript{5}, Stefan Iglauer\textsuperscript{6}

\textsuperscript{1}State Key Laboratory of Continental Dynamics; Department of Geology, Northwest University, Xi’an, 710069 China.
\textsuperscript{2}Research Institute of BGP, CNPC, Zhuozhou, 072750, China.
\textsuperscript{3}WA School of Mines: Minerals, Energy, and Chemical Engineering, Curtin University, 26 Dick Perry Avenue, 6151 Kensington, Australia
\textsuperscript{4}The Lyell Centre, Heriot-Watt University, Edinburgh EH14 4AS, Scotland, United Kingdom
\textsuperscript{5}Earth Sciences and Resource Engineering, CSIRO, 26 Dick Perry Avenue, 6151 Kensington, Australia
\textsuperscript{6}School of Engineering, Edith Cowan University, 270 Joondalup Drive, 6027 Joondalup, Australia

*Corresponding author: Yihuai Zhang (yihuai.zhang@hw.ac.uk); Hongyan Yu (yuhy@nwu.edu.cn)

Abstract

Formation water can in principal be used to identify hydrocarbon reserves. One such potential reserves are the gas reservoirs in the Ordos basin in China. However, there is limited data for this basin; we thus investigated the geochemical properties of a large range of formation water acquired from the Ordovician in the Ordos basin (42 brine samples obtained from different wells at M5 member) and analyzed their chemical characteristics. The results showed that this formation water is associated with a sealed reservoir, which is good for hydrocarbon storage. This is also related to the demonstrated strong diagenetic transformations. We also proposed statistical relationships between these geochemical properties and hydrocarbon storage based on a machine learning method (Decision tree). The results suggest that the salinity, Na\textsuperscript{+} / Cl\textsuperscript{−} ratio, (Cl\textsuperscript{−}-Na\textsuperscript{+}) / Mg\textsuperscript{2+} ratio, (HCO\textsubscript{3}−-CO\textsubscript{3}\textsuperscript{2−}) / Ca\textsuperscript{2+} ratio and Mg\textsuperscript{2+} / Ca\textsuperscript{2+} ratio highly correlate with the gas preservation. The results thus provide
drastically more accurate predictions in terms of where to find gas reservoirs in the Ordos basin, and can thus lead to significantly better exploitation of these resources.

**Keywords:** formation water; ion concentration; hydrocarbon preservation; machine learning; decision tree.

1. Introduction

Formation water as part of the hydrologic cycle in deep sedimentary reservoirs plays a fundamental role in hydrocarbon migration, accumulation, and preservation (Bagheri et al., 2014; Lüders et al., 2010; Phan et al., 2016; Zhang et al., 2013). These hydrocarbon reservoirs abundantly contain formation water (Hitchon and Friedman, 1969; Yoshioka et al., 2015; Kharaka and Hanor, 2003), and the chemical composition of the formation water is significantly affected by their sedimentary environment (Kloppmann et al., 2001; Land, 1995; Salminen et al., 2013). Thus formation water chemistry is modified by water-rock interactions (diagenetic transformations), e.g. dolomitization, anhydrite dissolution, illitization, and/or interaction with meteoric water (Möller et al., 2008). These water-rock interactions mainly occur in sealed reservoirs and they can lead to mineral transformation in the strata (Al-Ramadan et al., 2017; Maleki et al., 2019). Furthermore the ion compositions of the formation fluid (especially in the carbonate rock) can be dramatically changed over geological periods (Castendyk and Webster-Brown, 2007; Khaska et al., 2015; Kloppmann et al., 2001; Leybourne and Goodfellow, 2010; Maleki et al., 2016). This is further complicated by meteoric water from an active aquifer or shallow ground water (with low Total Dissolved Solid (TDS)) which flowed through faults and fractures (Bjørlykke, 1994; Hebig et al., 2012; Kharaka and Hanor, 2003; Knauth and Beeunas, 1986; Tesmer et al., 2007) and can also drastically affect the formation water chemistry (Gislason and Eugster, 1987; Gislason and Hans, 1987; Tang et al., 2008).

It is thus evident that formation water chemistry can vary dramatically, while the precise chemistry has a major impact on the optimum location for hydrocarbon preservation (Al-Hajeri and Bowden, 2017). Previous studies suggest that highly saline formation water indicates a good reservoir sealing efficiency and thus high potential for hydrocarbon
accumulations (Hassani and Garjan, 2016; Pattan et al., 2005; Worden et al., 1996; Mindong et al., 2014; Taheri et al., 2019). For instance, Al-Hajeri and Bowden (2017) conducted a field study on the Gotnia Formation (in Kuwait) and successfully identified the subsurface water masses in a seal integrity evaluation. They found several variations in the chemical composition of the formation water, which represent the interaction of migrating fluids with evaporating formations. Hence, we can see that it is of key importance to fully understand and characterize the formation water in the potential reservoirs which would greatly benefit the interpretation of the basins’ geological, hydrological and thermal properties, and thus aid oil and gas exploration (Al-Hajeri and Bowden, 2017; Worden et al., 1996). However, it is still challenging to investigate the effect of salinity, pH and ion concentration in gas reservoirs in the Ordos basin (Martini et al., 1998; Means, 1995; Orem et al., 2014; Shaffer et al., 2013; Zhenping and Lianfu, 2006), which is the second largest sedimentary basin of China, and also has lot of vital gas reservoirs in China (Hongyan et al., 2016; Liu et al., 2009). Thus in this study, we analyzed the chemical composition of the formation water from the M5 Member of Ordovician in Ordos basin, China (which is a typically carbonated reservoir deposited at Lower-Paleozoic), and evaluated the geochemical properties and distribution characteristics of the formation water.

there are a lot of studies where machine learning was used for reservoir evaluation; e.g., (Wong et al., 2005) and (Hall, 2016) used support vector machines to classify reservoir characteristics and lithofacies, respectively; or (Bruce et al., 2000) estimated permeability from well logs with Bayesian neural networks; and (Akande et al., 2015) predicted the permeability of carbonate reservoirs through support vector machines. We also identified the relationship between gas preservation and formation water chemistry based on geostatistical modelling.

2. Methodology

2.1. Materials and reservoir characterization

42 brine samples were collected from exploratory wells of the M5 Member in the northwest Ordos’s basin. Note that the Paleozoic strata of the Ordos basin covers an area of more than 250000 km². There are four upper Paleozoic giant gas fields which contain over 100 billion cubic meters (bcm) of gas reserves, respectively. For the lower Paleozoic strata, natural gas is only preserved in the 5th member of Majiagou formation (Hongyan et al., 2016; Liu et al.,
The Majiagou formation is a typical self-sourced carbonate natural gas reservoir which has significant resource potential. The effective source rocks of the Majiagou formation are mainly distributed in the upper section of the 5th member of Majiagou formation which has a 0.30%-8.45% TOC; the composition of natural gas is mainly dry gas that is almost source oil-type gas because most of these source rocks are at an over-mature stage (Tu et al., 2016; Wei et al., 2017). Hence, the M5 Member is one of the most interesting layers with large hydrocarbon accumulations at the Ordovician, Lower-Paleozoic layer (3000m to 3500m depth) in China (Dai et al., 2005; Yang et al., 2005; Yu et al., 2016). The research area thus covers two cities (H and W, see Figure. 1) and is part of the Jingbian gas field – one of the largest carbonate gas fields in China (Dai et al., 2006; Yang et al., 2005) with more than $1 \times 10^{11}$ m$^3$ gas production (Cai et al., 2005; Wang and Al-Aasm, 2002). Despite hundreds of geologic and petroleum resource investigations completed over the last decades, studies of formation brine in this basin are still limited (Hua et al., 2014; Xinquan et al., 2012; Xinshan et al., 2017; Yao et al., 2008). Furthermore, there are four gas trap types which include paleo-geomorphology traps, lithology traps, paleo-geomorphology-lithology combination traps and structure-lithology combination traps (Li et al., 2008). The M5 Member is mainly marine carbonate rock formed in a supratidal sedimentary environment (Feng and Bao, 1999; Wang and Al-Aasm, 2002; Zhang et al., 2014). The hydrocarbon migration and storage was induced by the pervasive cracks generated after several stages of weathering, denudation, eluviation and mixed dissolution at the top of the M5 Member formation (Lianbo and Xiang-Yang, 2009; Wang and Al-Aasm, 2002; Yang et al., 2008). From the late Ordovician to the middle Carboniferous, the top of the M5 member was uplifted and exposed to air (Jiang et al., 2013; Wang and Al-Aasm, 2002). It consequently experienced weathering, resulting in rock breakdown, and denudation which stripped the weathered products of the surface rocks through wind and water. Also, the fine material and the soluble constituent of the rock were carried away by seeping meteoric water into the underground (eluviation). Such complicated sedimentary and diagenesis processes resulted in rock decomposition and fracturing, and provided seepage channels for eluviation, thus changing porosity, permeability, and also the composition of groundwater (Mindong et al., 2014; Wang and Al-Aasm, 2002).
2.2. Methods

2.2.1. Experimental Procedure

The composition of the 42 brine samples were analyzed in the laboratory; thus acid titration methods were used to determine the bicarbonate concentration; major anions (Cl⁻, CO₃²⁻, SO₄²⁻) were measured via ion chromatography (Dionex-120); cations (Ca²⁺, Mg²⁺, Na⁺, K⁺)
were identified using anion and cation column (Varsányi and Kovács, 2009); pH values were measured in the field using pH meter; and the Total Dissolved Solid (TDS) content was obtained by summing the main ions measured.

2.2.2. Machine learning

In order to find the relationship between the chemical composition of formation water and gas accumulation, we used the Decision Tree method for correlational analysis. Note that such Machine learning approaches are now popular in petroleum exploration and production (Ahmadi et al., 2014; Al-Anazi and Gates, 2010; Lukoševičius and Jaeger, 2009; Yu et al., 2017a; Yu et al., 2017b). The Decision tree method is a universal and well-known classification algorithm, and it is a white box technique, where testers can check the internal structures or functioning of an application. The decision tree contains an algorithm which has a tree-like nature with branches from the root node base on a Greed Recursive algorithm, and uses the attribute which has the highest weighting of information gain, information gain ratio and Gini coefficient in each step for splitting until the branch node (target attribute) appeared (Figure 2). For this study, we designed a process flow of machine learning analysis specially based on the research area dataset.

Figure 2. Schematic of the Decision tree.

① Data collection: collect formation water parameters
Weighted calculation: compute weights of each formation water parameter based on three weighting algorithms (information gain, information gain ratio, and Gini coefficient).

Selection by weights: select an attribute with the highest weight, respectively, in every step during the decision tree training (Hall, 1999; Langley, 1994).

Machine learning and data training: machine (Decision Tree) learns from the collected data.

Correlation calculation: calculate the parametrical correlations.

Three weighting algorithms (information gain (Lee and Lee, 2006), information gain ratio (Dai and Xu, 2013) and Gini coefficient (Yitzhaki, 1979) were used to simultaneously calculate the weighting of the chemistry parameters. The information gain measured the relevance of attributes to predicted attribute, estimated from the difference between an original requirement and new requirement (Equation 1).

\[
Gain(A) = Info(D) - Info_A(D)
\]

Where \(Info\) is the original expectation information, and \(Info_A\) is the new expectation information. The information gain ratio corrected the shortcoming of information gain, which was always prone to choosing multi-valued attributes by using split information value and reduced the bias caused by the information gain. The split information value represents the potential information generated by splitting the training data set \(D\) into \(v\) partitions, corresponding to \(v\) outcomes on attribute \(A\):

\[
SplitInfo_A(D) = -\sum_{j=1}^{v} \frac{|D_j|}{|D|} \log_2 \left( \frac{|D_j|}{|D|} \right)
\]

And the gain ratio is defined as:

\[
GainRatio(A) = \frac{Gain(A)}{SplitInfo(A)}
\]
Whereas the Gini coefficient measured the impurity of the samples:

\[ Gini(D) = 1 - \sum_{j=1}^{v} p_{j}^{2} \tag{4} \]

The average Gini index is defined as:

\[ Gini(A) = \frac{\sum_{j=1}^{v} |D_j|}{|D|} \times Gini(D) \tag{5} \]

Based on the results provided by the three weighting algorithms, we chose the attribute which has the highest weight among them in every calculation step as the splitting attribute. Finally, Decision tree method computes the correlation between the formation water geochemical parameters and hydrocarbon accumulation.

3. Results and discussion

3.1. Total dissolved solids (TDS)

The formation water samples in this carbonate reservoir showed a broad range of ion compositions. The dominant cations in the formation water were \( \text{K}^+ \), \( \text{Na}^+ \), \( \text{Ca}^{2+} \) and \( \text{Mg}^{2+} \), and the dominant anions \( \text{Cl}^- \), \( \text{SO}_4^{2-} \) and \( \text{HCO}_3^- \) (see Table 1 for details). Two salinity groups were identified in the research area (Figure 3), namely a low salinity group (ranging from 0.25 g/L to 10 g/L), which indicates that the original (NaCl) brine was diluted by mixing with connate and / or even meteoric waters from shallower groundwater (probably due to the absence of a caprock or unconformity) or from active aquifers (Yang et al., 2005). Whereas, the high salinity formation water (53 g/L to 268 g/L) apparently developed in a sealed and reducing environment with high metamorphism, water-rock reaction, and evaporation (Barth and Riis, 1992). Such a hydrological environment with the good sealing is beneficial for hydrocarbon storage (Cramer et al., 1999; Hitchon and Friedman, 1969).

Table 1. The ion concentrations of the water samples analysed.

<table>
<thead>
<tr>
<th>Well Number</th>
<th>Na(^+) (mg/L)</th>
<th>Ca(^{2+}) (mg/L)</th>
<th>Mg(^{2+}) (mg/L)</th>
<th>Cl(^-) (mg/L)</th>
<th>SO(_4^{2-}) (mg/L)</th>
<th>HCO(_3^-) (mg/L)</th>
<th>TDS (mg/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table Pre-proof
<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13263</td>
<td>29535</td>
<td>3946</td>
<td>83966</td>
<td>0</td>
<td>407</td>
</tr>
<tr>
<td>2</td>
<td>9131</td>
<td>15640</td>
<td>3355</td>
<td>51339</td>
<td>0</td>
<td>316</td>
</tr>
<tr>
<td>3</td>
<td>11651</td>
<td>16272</td>
<td>2972</td>
<td>54959</td>
<td>0</td>
<td>770</td>
</tr>
<tr>
<td>4</td>
<td>17026</td>
<td>24253</td>
<td>3914</td>
<td>79970</td>
<td>0</td>
<td>1009</td>
</tr>
<tr>
<td>5</td>
<td>27613</td>
<td>58827</td>
<td>7828</td>
<td>162902</td>
<td>8657</td>
<td>263</td>
</tr>
<tr>
<td>6</td>
<td>25426</td>
<td>49468</td>
<td>4266</td>
<td>138529</td>
<td>0</td>
<td>1058</td>
</tr>
<tr>
<td>7</td>
<td>234</td>
<td>16</td>
<td>23</td>
<td>47</td>
<td>173</td>
<td>483</td>
</tr>
<tr>
<td>8</td>
<td>224</td>
<td>34</td>
<td>26</td>
<td>47</td>
<td>172</td>
<td>524</td>
</tr>
<tr>
<td>9</td>
<td>98</td>
<td>49</td>
<td>18</td>
<td>135</td>
<td>16</td>
<td>246</td>
</tr>
<tr>
<td>10</td>
<td>23867</td>
<td>13961</td>
<td>2177</td>
<td>57606</td>
<td>5571</td>
<td>243</td>
</tr>
<tr>
<td>11</td>
<td>23882</td>
<td>14742</td>
<td>2613</td>
<td>59024</td>
<td>5692</td>
<td>246</td>
</tr>
<tr>
<td>12</td>
<td>25056</td>
<td>17497</td>
<td>2776</td>
<td>60024</td>
<td>5716</td>
<td>131</td>
</tr>
<tr>
<td>13</td>
<td>26430</td>
<td>19337</td>
<td>3393</td>
<td>65858</td>
<td>6148</td>
<td>352</td>
</tr>
<tr>
<td>14</td>
<td>21418</td>
<td>39458</td>
<td>4246</td>
<td>115058</td>
<td>0</td>
<td>243</td>
</tr>
<tr>
<td>15</td>
<td>21696</td>
<td>40850</td>
<td>3103</td>
<td>114073</td>
<td>646</td>
<td>352</td>
</tr>
<tr>
<td>16</td>
<td>22553</td>
<td>40042</td>
<td>4137</td>
<td>117520</td>
<td>0</td>
<td>246</td>
</tr>
<tr>
<td>17</td>
<td>22297</td>
<td>43094</td>
<td>3974</td>
<td>122115</td>
<td>0</td>
<td>131</td>
</tr>
<tr>
<td>18</td>
<td>275</td>
<td>52</td>
<td>18</td>
<td>236</td>
<td>204</td>
<td>314</td>
</tr>
<tr>
<td>19</td>
<td>20233</td>
<td>25612</td>
<td>3619</td>
<td>84649</td>
<td>3018</td>
<td>283</td>
</tr>
<tr>
<td>20</td>
<td>15438</td>
<td>27740</td>
<td>4355</td>
<td>84785</td>
<td>542</td>
<td>649</td>
</tr>
<tr>
<td>21</td>
<td>5179</td>
<td>94</td>
<td>26</td>
<td>1838</td>
<td>4440</td>
<td>5353</td>
</tr>
<tr>
<td>22</td>
<td>1276</td>
<td>442</td>
<td>92</td>
<td>2514</td>
<td>467</td>
<td>269</td>
</tr>
<tr>
<td>23</td>
<td>11074</td>
<td>14742</td>
<td>1373</td>
<td>46067</td>
<td>1078</td>
<td>499</td>
</tr>
<tr>
<td>24</td>
<td>25</td>
<td>74</td>
<td>20</td>
<td>40</td>
<td>43</td>
<td>264</td>
</tr>
<tr>
<td>25</td>
<td>31</td>
<td>72</td>
<td>23</td>
<td>40</td>
<td>47</td>
<td>287</td>
</tr>
<tr>
<td>26</td>
<td>32</td>
<td>72</td>
<td>19</td>
<td>20</td>
<td>78</td>
<td>264</td>
</tr>
<tr>
<td>27</td>
<td>20</td>
<td>18</td>
<td>19</td>
<td>13</td>
<td>4</td>
<td>173</td>
</tr>
<tr>
<td>28</td>
<td>16598.8</td>
<td>55100</td>
<td>2100</td>
<td>129000</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>29</td>
<td>3396</td>
<td>7727</td>
<td>345</td>
<td>20333</td>
<td>98</td>
<td>150</td>
</tr>
<tr>
<td>30</td>
<td>5888</td>
<td>12469</td>
<td>991</td>
<td>33580</td>
<td>294</td>
<td>394</td>
</tr>
<tr>
<td>31</td>
<td>47</td>
<td>79</td>
<td>4</td>
<td>39</td>
<td>39</td>
<td>271</td>
</tr>
<tr>
<td>32</td>
<td>22417</td>
<td>47519</td>
<td>4423</td>
<td>131281</td>
<td>0</td>
<td>401</td>
</tr>
<tr>
<td>33</td>
<td>23541</td>
<td>45101</td>
<td>6166</td>
<td>133981</td>
<td>0</td>
<td>132</td>
</tr>
<tr>
<td>34</td>
<td>6313</td>
<td>2530</td>
<td>0</td>
<td>12214</td>
<td>1028</td>
<td>2212</td>
</tr>
<tr>
<td>35</td>
<td>16700</td>
<td>28000</td>
<td>6430</td>
<td>93600</td>
<td>446</td>
<td>170</td>
</tr>
<tr>
<td>36</td>
<td>17500</td>
<td>32400</td>
<td>5660</td>
<td>101000</td>
<td>0</td>
<td>177</td>
</tr>
<tr>
<td>37</td>
<td>7724</td>
<td>53319</td>
<td>11810</td>
<td>140744</td>
<td>0</td>
<td>56</td>
</tr>
<tr>
<td>38</td>
<td>4356</td>
<td>24614</td>
<td>3346</td>
<td>59906</td>
<td>0</td>
<td>347</td>
</tr>
<tr>
<td>39</td>
<td>4100</td>
<td>41900</td>
<td>10600</td>
<td>111000</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>40</td>
<td>25889</td>
<td>62105</td>
<td>7750</td>
<td>171196</td>
<td>1148</td>
<td>541</td>
</tr>
<tr>
<td>41</td>
<td>14240</td>
<td>62765</td>
<td>4465</td>
<td>52807</td>
<td>125877</td>
<td>481</td>
</tr>
<tr>
<td>42</td>
<td>17700</td>
<td>12900</td>
<td>1070</td>
<td>53200</td>
<td>0</td>
<td>285</td>
</tr>
</tbody>
</table>
3.2. Ion concentrations

In the carbonate gas reservoirs, which are essentially brine-carbonate-gas systems from a chemical perspective, multiple chemical reactions occur over geological times (Shariatpanahi et al., 2011). Thus the Ca\(^{2+}\) concentration in the brine is usually highest, followed by Mg\(^{2+}\), Na\(^{+}\), K\(^{+}\), Cl\(^{-}\), SO\(_4^{2-}\), HCO\(_3^{-}\), CO\(_3^{2-}\) (Zhu et al., 2007). Note that the SO\(_4^{2-}\) ion content is also higher than in sandstone reservoirs due to the high H\(_2\)S concentration prevalent in most Chinese marine gas fields (Zhu et al., 2007). Variable ion concentrations were observed even in similar salinity samples, where the ion concentrations ranked as Cl\(^{-}\) > Ca\(^{2+}\) > Na\(^{+}\) > Mg\(^{2+}\) > HCO\(_3^{-}\) > SO\(_4^{2-}\) (Table 1). Note that a dominant Ca\(^{2+}\) concentration indicates a carbonate sedimentary environment; and in combination with high K\(^{+}\), Na\(^{+}\), Mg\(^{2+}\) concentrations also indicates intense water-rock reactions (dolomitization, anhydrite dissolution or illitization; Ellis and Mahon, 1967; Giggenbach, 1986; Gunter et al., 2000). Cl\(^{-}\) was dominant in the anion (ahead of HCO\(_3^{-}\), and SO\(_4^{2-}\) (Figure. 4b)), mainly due to the NaCl prevailing in ancient seawater (Hitchon et al., 1971; Land and Macpherson, 1992) and low solubility of sulphate (CaSO\(_4\)) and hydrogen carbonate minerals (calcite and / or dolomite) (Ellis and Mahon, 1964; Ellis and Mahon, 1967). These results also illustrate that the formation water in this region is complicated and highly variable, and includes both, completely isolated formation water (CaCl\(_2\)) and transitional formation water (NaHCO\(_3\) and Na\(_2\)SO\(_4\)).
Figure 4. Dominant ion concentration distributions of the collected samples, (a) cation concentration; (b) anion concentration.
3.3. Sodium / Chloride (Na\textsuperscript{+} / Cl\textsuperscript{-}) ratio

The sodium / chloride (Na\textsuperscript{+} / Cl\textsuperscript{-}) ratio was calculated to elucidate the metamorphism (Warner et al., 2012) and hydro-geochemistry of the reservoirs, (Figure.5). All samples in the plot are below the seawater evaporation line, thus illustrating a removal of Na\textsuperscript{+} from fluids independent of Chloride solubility (Hanson and Mauersberger, 1990; Melaiye et al., 2004). Such results demonstrate that the Na\textsuperscript{+} - Ca\textsuperscript{2+} exchange in feldspar transformation (Lüders et al., 2010). However, although this exchanging is insufficient, the intense water-rock reaction can still be illustrated by the low Na\textsuperscript{+} / Cl\textsuperscript{-} ratios. Thus the east part of the region has an average (Na\textsuperscript{+} / Cl\textsuperscript{-}) ratio of 0.25, and a lowest of 0.1 (refer to Figure. 6). We conclude that the formation water here (the east part) is ancient residual seawater sealed in the reservoir without being disturbed by shallow ground water; it is thus beneficial for hydrocarbon storage (De Choudens-Sanchez and Gonzalez, 2009; Osborn et al., 2011; Rice et al., 2008). The (Na\textsuperscript{+} / Cl\textsuperscript{-}) ratio in the southeast region is higher than 1. This part was exposed to shallow ground water or an active aquifer (Land and Macpherson, 1992), which is an unfavorable condition for hydrocarbon storage (Sturchio et al., 2001; Tesoriero et al., 2004).

![Figure. 5. Cross –plot of Na\textsuperscript{+} vs Cl\textsuperscript{-} with theoretical evaporation and dissolution trends (McCaffrey et al., 1987).](image)
Figure. 6. Map of Na\(^+\)/Cl\(^-\) ratios for the study area, contour lines show different ratios.

### 3.4. Magnesium / Calcium (Mg\(^{2+}\)/Ca\(^{2+}\)) ratio

The dolomitization leads to high Ca\(^{2+}\) concentration versus low Mg\(^{2+}\) concentration, Equation (6; Bozau et al., 2015; Stüben et al., 1996) in formation water. We thus plot Ca\(^{2+}\) concentration vs. Cl\(^-\) concentration (Figure. 7), and the results show that most of the samples are between the CaCl\(_2\) solubility line and the seawater evaporation path from McCaffrey 2007; it can be concluded that evaporation is not significant. However, dolomitization (Birkle et al., 2009; Bozau et al., 2015) plays an essential role in controlling the concentration of Ca\(^{2+}\) and Mg\(^{2+}\) (Equation 6); and Figure. 8). Clearly, the concentration of Mg\(^{2+}\) is below the
seawater evaporation path which indicates that Mg\textsuperscript{2+} is displaced by Ca\textsuperscript{2+} besides evaporation.

\[ 2\text{CaCO}_3(\text{Calcite}) + \text{Mg}^2+(\text{Brine}) \rightarrow \text{CaMg(CO}_3)_2(\text{Dolomite}) + \text{Ca}^2+ \]  \hspace{1cm} (6)

Dolomitization occurs in sealed reservoirs (Nader and Swennen, 2004), thus, Mg\textsuperscript{2+} / Ca\textsuperscript{2+} ratios can indicate the sealing condition. Thus well sealed reservoirs have a low Mg\textsuperscript{2+} / Ca\textsuperscript{2+} ratio (De Choudens-Sanchez and Gonzalez, 2009; Folk and Land, 1975; Koleini et al., 2018; Lear et al., 2002; Nielsen et al., 2016; Shariatpanahi et al., 2011; Wilkinson and Algeo, 1989); whereas, high Mg\textsuperscript{2+} / Ca\textsuperscript{2+} ratios indicate reservoirs which are not well sealed, leading to less hydrocarbon storage, as the west and northeast part in Figure 9.

![Figure 7](image1.png)

**Figure. 7.** Cross–plot of Ca\textsuperscript{2+} vs Cl\textsuperscript{−} with theoretical evaporation and dissolution trends (McCaffrey et al., 1987).

![Figure 8](image2.png)

**Figure. 8.** Cross–plot of Mg\textsuperscript{2+} vs Cl\textsuperscript{−} with theoretical evaporation and dissolution trends (McCaffrey et al., 1987).
3.5. Metamorphic coefficient \((\text{Cl}^-\text{Na}^+) / \text{Mg}^{2+}\) ratio

The metamorphic coefficient – (i.e. the Cl\(^-\)-Na\(^+\) / Mg\(^{2+}\)) ratio (Ranasinghe et al., 2005; Zheng, 2012) is assessed to evaluate the degree of water-rock interaction and associated ion replacement. The (Cl\(^-\)-Na\(^+\) / Mg\(^{2+}\)) ratio of M5 Member in the area widely ranged from 2 to 33 with an average value of 18.57, and high metamorphic coefficients are observed in the diagonal line from northwest to southeast. This can be explained by the weak erosion in the gentle hill of the karst slope and remnant hill of the karst highland (Hua and Xinshe, 2014; Jiang et al., 2013; Li et al., 2008; Yu et al., 2016), where less contact with groundwater (or
active aquifers) led to strong water-rock interactions. Thus Na$^+$ and Mg$^{2+}$ were mostly displaced by Ca$^{2+}$, resulting in high metamorphic coefficients. In contrast, the northeast and southwest of the research area (Figure. 10) showed weak metamorphism that is thus unfavorable for hydrocarbon storage.

Figure. 10. Map of (Cl$^-$ - Na$^+$) / Mg$^{2+}$ ratios for the study area, contour lines indicate different ratios.

3.6. Carbonate / bicarbonate and calcium ((HCO$_3^-$-CO$_3^{2-}$) / Ca$^{2+}$) ratio

(HCO$_3^-$-CO$_3^{2-}$) / Ca$^{2+}$ ratio is analysed to test decarbonation and fluid properties. HCO$_3^-$ is a product of carbonate dissolution by oxygenated recharge water and of sulfate reduction methane fermentation (Van Voast, 2003). The ratios of (HCO$_3^-$-CO$_3^{2-}$) / Ca$^{2+}$ in the M5
formation are mostly lower than 0.5 (Figure 11), and the area around H city shows the lowest ratios with an increase to the northwestern and southwestern direction. Previous studies showed that low \( \frac{(\text{HCO}_3^- - \text{CO}_3^{2-})}{\text{Ca}^{2+}} \) ratios are normally associated with high acidic content, which is closely related to the oil field water organics (Deines et al., 1974; Yuan et al., 2007; Zeebe, 1999); these thus indicate suitability for light oil or gas accumulation (Chen et al., 2005; Kurita et al., 2008; Palandri and Reed, 2001).

Figure. 11. Map of \( \frac{(\text{HCO}_3^- - \text{CO}_3^{2-})}{\text{Ca}^{2+}} \) ratios for the study area, contour lines indicate different ratios.
3.7. Correlation between hydrocarbon preservation and hydrogeochemical properties

We first chose the attributes including the fractions of Na\(^+\), Ca\(^{2+}\), Mg\(^{2+}\), Cl\(^-\), SO\(_4\)^{2-}, HCO\(_3\)^-, salinity, PH, and ratios of (Na\(^+\) / Cl\(^-\)), ((Cl\(^-\)-Na\(^+\)) / Mg\(^{2+}\)), ((HCO\(_3\)^-CO\(_3\)^{2-}) / Ca\(^{2+}\)) and (Mg\(^{2+}\) / Ca\(^{2+}\)). Moreover, we also created a new parameter, HCO\(_3\)^- / Cl\(^-\), for the accurate evaluation of the formation water properties. This ratio was proposed by considering that most of the HCO\(_3\) ion from meteoric water while most of Cl\(^-\) ion is for oil field water (Engle et al., 2016; Kharaka et al., 1993). The formation water chemistry has been investigated, and the various ion concentrations and ratios have been correlated with gas preservation. Subsequently the information gain, information gain ratio, and Gini coefficient (the criteria of attribute selection) were used to compute the weighting of attributes. Clearly, the HCO\(_3\)^- / Cl\(^-\) ratio had the highest weighting values in the first decision step. However, the weighting of each parameter will be changed in different decision steps. The weighting in each step is further calculated and the highest one (the sum of three weighting from three algorithms) is used to establish the relationship.

Table 2. An example of the weighting for hydrocarbon storage and hydrogeochemical properties.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Information gain</th>
<th>Information gain ratio</th>
<th>Gini coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>HCO(_3)^- / Cl(^-)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>TDS</td>
<td>0.826357992</td>
<td>0.804737774</td>
<td>0.970814278</td>
</tr>
<tr>
<td>Cl(^-)</td>
<td>0.826357992</td>
<td>0.804737774</td>
<td>0.970814278</td>
</tr>
<tr>
<td>Na(^+)</td>
<td>0.800762678</td>
<td>0.63718087</td>
<td>0.803112965</td>
</tr>
<tr>
<td>(HCO(_3)^-CO(_3)^{2-}) / Ca(^{2+})</td>
<td>0.750257359</td>
<td>0.796541491</td>
<td>0.751219512</td>
</tr>
<tr>
<td>Ca(^{2+})</td>
<td>0.6588943</td>
<td>0.635726478</td>
<td>0.786432927</td>
</tr>
<tr>
<td>Mg(^{2+})</td>
<td>0.645357369</td>
<td>0.579609279</td>
<td>0.74138047</td>
</tr>
<tr>
<td>Na(^+)/Cl(^-)</td>
<td>0.586648184</td>
<td>0.579609279</td>
<td>0.74138047</td>
</tr>
<tr>
<td>Mg(^{2+})/Ca(^{2+})</td>
<td>0.510044104</td>
<td>0.318949758</td>
<td>0.451219512</td>
</tr>
<tr>
<td>(Cl(^-)-Na(^+))/Mg(^{2+})</td>
<td>0.240230002</td>
<td>0.326430889</td>
<td>0.512400328</td>
</tr>
</tbody>
</table>

The Decision tree then computed the correlations between the main chemical formation water properties and gas storage (Table 3). Thus the HCO\(_3\)^- / Cl\(^-\) ratio and (Cl\(^-\)-Na\(^{2+}\)) / Mg\(^{2+}\) ratio significantly important, as the chemical environments ① and ② in Table 3 correlated highly with gas preservation, and the production data also show that these reservoirs produce gas. Semi-sealed reservoirs ③ normally contain typeformation water, and thus produce gas and water together. The other reservoirs in the research area are not suitable for gas storage.
The correlation efficiency shows the correlation is not completely accurate, but the accuracy for the rules are higher than 80%. A validation for these correlations was also performed for another four wells which were not included in the training dataset. The well-section (Figure. 12) shows that the tectonic elevation difference did not affect gas migration due to complex sub-reservoirs distributed in this area (Hao et al., 1997; Liu et al., 2009; Mingjian et al., 2011; Yang et al., 2005). This indicates that these four wells are not in the same sub-reservoir, and they were independent of each other with different formation water characteristics. Table 4 shows the geochemical properties of the formation water in these four wells. The geochemical characteristics show Wells B and D are good for gas storage, while Well C is a semi-sealed reservoir, and Well A is not beneficial for gas storage, consistent with the field well production data, see Figure.12.

Table 3. Correlation between hydrocarbon preservation and hydrogeochemical properties.

<table>
<thead>
<tr>
<th>Formation water chemistry</th>
<th>Gas preservation</th>
</tr>
</thead>
<tbody>
<tr>
<td>① (HCO$_3^-$ / Cl) $\leq$ 0.006 &amp; (Cl-Na$^+$) / Mg$^{2+}$ $&gt;$ 17.855 &amp; (HCO$_3^-$-CO$_3^{2-}$) / Ca$^{2+}$ $&lt;$ 0.023</td>
<td>Good (accuracy: 88.12%)</td>
</tr>
<tr>
<td>② (HCO$_3^-$ / Cl) $\leq$ 0.006 &amp; (Cl-Na$^+$) / Mg$^{2+}$ $\leq$ 17.855 &amp; Mg$^{2+}$ / Ca$^{2+}$ $&lt;$ 0.139</td>
<td>Medium (accuracy: 80%)</td>
</tr>
<tr>
<td>③ (HCO$_3^-$ / Cl) $\leq$ 0.006 &amp; (Cl-Na$^{2+}$) / Mg$^{2+}$ $\leq$ 17.855 &amp; Mg$^{2+}$ / Ca$^{2+}$ $&gt;$ 0.005</td>
<td>Unsuitable (accuracy: 81.82%)</td>
</tr>
<tr>
<td>Others</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Correlation between hydrocarbon preservation and hydrogeochemical properties.

![Image of well-section](image_url)
Figure. 12. Gas and water wells (well A-D) for the tested wells.

Table 4. Hydrogeochemical properties of formation water of the four wells tested.

<table>
<thead>
<tr>
<th>Well name</th>
<th>Water type</th>
<th>TDS (g/L)</th>
<th>Na⁺ / Cl⁻ (g/L)</th>
<th>(Cl⁻-Na⁺)⁺ / Mg²⁺ (g/L)</th>
<th>HCO₃⁻ / Cl⁻ (g/L)</th>
<th>(HCO₃⁻-CO₃²⁻)⁻ / Ca²⁺ (g/L)</th>
<th>Mg²⁺ / Ca²⁺ (g/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>CaCl₂</td>
<td>0.747</td>
<td>0.475248</td>
<td>16.99293</td>
<td>1.242574</td>
<td>0.024676</td>
<td>0.164126</td>
</tr>
<tr>
<td>B</td>
<td>CaCl₂</td>
<td>102.858</td>
<td>0.134312</td>
<td>26.99293</td>
<td>0.005901</td>
<td>0.003477</td>
<td>0.04572</td>
</tr>
<tr>
<td>C</td>
<td>CaCl₂</td>
<td>2.346</td>
<td>0.319479</td>
<td>16.123</td>
<td>0.005657</td>
<td>0.005123</td>
<td>0.052299</td>
</tr>
<tr>
<td>D</td>
<td>CaCl₂</td>
<td>245.306</td>
<td>0.201068</td>
<td>17.06781</td>
<td>0.001562</td>
<td>1.446602</td>
<td>0.120485</td>
</tr>
</tbody>
</table>

4. Conclusions

We present an interpretation of the chemical formation water properties in the northwest part of the Ordos basin, China. The study showed that the chemistry of the formation water in the M5 Member of the Ordos basin varied significantly due to two key reasons; namely 1) seawater evaporation and intense water-rock interactions during burial in the hydrodynamically isolated sub-reservoirs, and 2) mixture with and exposure to shallow connate or meteoric fluids or active aquifers of hydrodynamically active sub-reservoirs. The major conclusions of this study are summarized as follows:

(1) Firstly, the ion characteristics showed that water-rock interactions were extremely strong; thus Na⁺ and Cl⁻ concentrations were lower than in original seawater due to ion exchange with the minerals in the rock; dolomitization was also highly significant and it controlled the Ca²⁺ and Mg²⁺ concentrations; metamorphism, however, was the main process in this gas reservoir; while decarbonation reflected high acid content in the reservoirs.

(2) The relation between geochemical characteristics and distribution of the formation water were analyzed, including parameters such as TDS, Na⁺ / Cl⁻ ratio, (Cl⁻-Na⁺)⁺ / Mg²⁺ ratio, (HCO₃⁻-CO₃²⁻)⁻ / Ca²⁺ ratio, Mg²⁺ / Ca²⁺ ratio, which give an indication of sealing conditions and gas storage capabilities. Based on Decision Tree analysis (with three weighting evaluation algorithms, i.e. information gain, information gain ratio and Gini coefficient), we identified the significant correlations between formation water chemistry and gas preservation in this investigated area.

(3) These correlations were also tested and validated in other wells and successfully
delineated the areas for both high water content and good hydrocarbon accumulation. We thus conclude that formation water chemistry clearly correlates with hydrocarbon preservation, and can thus be used a geological indicator to identify hydrocarbon reserves.

**Acknowledgements**

This study was supported by National and Local Joint Engineering Research Center for Carbon Capture Utilization and Sequestration and Special Fund from the State Key Laboratory of Continental Dynamics at Northwest University in China, also supported By Young Talent fund of University Association for Science and Technology in Shaanxi, China. The authors have no real or perceived financial conflicts of interest to declare. Data are available in the tables.

**References**


Giggenbach, W., 1986. Graphical techniques for the evaluation of water/rock equilibration conditions by use of Na, K, Mg and Ca contents of discharge waters, Proc. 8th New Zealand Geothermal Workshop, pp. 37-44.


Hassani, H. and Garjan, Y.S., 2016. Impact of Low Saline Formation Water on the Swept Leg's Reservoir Property at the EOR Target Sand, 22nd Formation Evaluation


Lüders, V. et al., 2010. Chemistry and isotopic composition of Rotliegend and Upper Carboniferous formation waters from the North German Basin. Chemical Geology,


Orem, W. et al., 2014. Organic substances in produced and formation water from


• The geochemistry data of the brine water from the M5 Member, Ordovician in Ordos basin, China is investigated.

• A new model is proposed to predict the gas / water distribution from brine geochemistry properties based on machine learning (Decision tree) method.

• The newly proposed model successfully validated on other wells.
Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: