Pseudo proximate analysis: Method using wireline logs to estimate components of coal bearing rock matrix without control data

Christopher Robert McLean

A minithesis submitted in partial fulfilment of the requirements for the degree of Magister Scientiae in the Department of Earth Science, University of the Western Cape.

Supervisor: Prof Daniel Mikeš (NMMU)
Co-supervisor: Prof Jan van Bever Donker (UWC)

February 2015
Pseudo proximate analysis: Method using wireline logs to estimate components of coal bearing rock matrix without control data

Christopher Robert McLean

KEYWORDS
Botswana
Kalahari Karoo
Coal Bed Methane
Proximate Analysis
Statistics
Rock Matrix
Wireline Logs
Neutron
Modelling
Relationship
ABSTRACT

Pseudo proximate analysis: Method using wireline logs to estimate components of coal bearing rock matrix without control data

C. R. McLean

MSc minithesis, Department of Earth Science, University of the Western Cape.

Lab conducted proximate analysis of coal bearing rock units calculates the weight percentage of ash, moisture, fixed carbon and volatile matter through a series of combustion steps. The data obtained is quintessential in establishing the coal rank and in the case of coal bed methane the gas-in-place estimates. In this study 105 proximate analysis samples, from 7 drilled wells, are taken from the south-eastern Kalahari Basin in Botswana.

The pseudo proximate analysis, the method proposed in this thesis, calculates the lab proximate analysis results using the neutron, density and gamma ray wireline logs. The uniqueness of the method lies in the fact that no cut off values are needed for the wireline logs, nor are the results of the lab proximate analysis required for calibration. An in depth study of the relationship between the wireline logs and proximate analysis is conducted using a principle component analysis and the results tested using a combination of statistical techniques to determine the significance of the relationship. It is shown that the density and neutron logs model the proportion of ash and volatile matter in the rock matrix, respectively, with a high degree of accuracy. The multiple regression analysis shows that percentages fixed carbon and moisture components of the rock matrix correlate poorly to the proposed well logs, thus most error lies in the determination of these two components.

It is statistically proven that the pseudo proximate analysis results are significantly different to the lab measured proximate analysis. This implies that the proposed pseudo proximate analysis method is unable to accurately determine the components of a coal bearing rock matrix using the density, neutron and gamma ray wireline logs. The application of the proposed method is a model to identity the coal bearing rock matrix and provide a predictive estimation of the coal quality, a priori lab measured data.

February 2015
DECLARATION

I declare that *Pseudo proximate analysis: Method using wireline logs to estimate components of coal bearing rock matrix without control data* is my own work, that it has not been submitted before for any degree or examination in any other university, and that all the sources I have used or quoted have been indicated and acknowledged as complete references.

Christopher Robert McLean
February 2015

Signed: __________________________
ACKNOWLEDGEMENTS

The success of this mini-thesis is due to help from many different parties, and here I would like to take the opportunity to thank them. I thank Synergy for providing the software application, Interactive Petrophysics, on which the majority of the work for this thesis was done. I also thank Sasol for providing the dataset and the opportunity to complete my master’s degree.

I thank my supervisor, Prof. D. Mikeš, for guiding the direction of the thesis and making certain that it follows a strict scientific methodology. I would also like to thank my co-supervisor, Prof. J. van Bever Donker, for his support during those final weeks and for is valuable and excellent knowledge of the English language.

Lastly, without the support of my wife and family I would still be struggling to complete this thesis. Thank you for all the financial support and for pushing me to achieve a good result.
CONTENTS

Title Page ........................................................................................................................................ i
Keywords ......................................................................................................................................... ii
Abstract .......................................................................................................................................... iii
Declaration .................................................................................................................................... iv
Acknowledgements .................................................................................................................. iv

1. Introduction ................................................................................................................................ 1
2. Regional overview ....................................................................................................................... 2
3. Methodology ............................................................................................................................... 4
4. Lab Proximate Analysis ............................................................................................................... 5
5. Previous Work ............................................................................................................................. 6
  5.1. Literature - %TOC ......................................................................................................................... 6
  5.2. Literature - Coal Analysis ............................................................................................................. 7
6. Pseudo Proximate Analysis ......................................................................................................... 9
  6.1. Coal Analysis - Cluster Analysis ................................................................................................... 9
  6.2. Coal Analysis - Normalization ...................................................................................................... 12
    6.2.1. Gamma Ray Normalization ........................................................................................................ 14
    6.2.2. Density Normalization ................................................................................................................ 16
    6.2.3. Neutron Normalization ............................................................................................................. 18
    6.2.4. Normalized Well Log Examples ............................................................................................... 19
  6.3. Identifying coal bearing rock .................................................................................................. 21
  6.4. Coal Analysis - PCA .................................................................................................................. 22
    6.4.1. The Method .............................................................................................................................. 22
    6.4.2. The Results ............................................................................................................................ 23
    6.4.3. Conclusions ........................................................................................................................... 24
  6.5. Coal Analysis – Pseudo Proximate ........................................................................................... 25
    6.5.1. Volume Clay ........................................................................................................................... 25
    6.5.2. Volume Fixed Carbon ............................................................................................................ 26
    6.5.3. Quartz and Pore Volumes ....................................................................................................... 27
  6.6. Coal Analysis – Results ............................................................................................................ 29
1. Introduction

The main focus of this project is to measure the relationship between logging tools and the lab conducted proximate analysis for coal bearing zones. The region chosen to conduct this study is the Central Kalahari Basin, Botswana where 9 wells have been drilled in an area covering 2800 km² (Figure 1).

This paper is divided into two components, namely pseudo proximate analysis (the method proposed in this thesis) and a multiple regression analysis. The first component, the pseudo proximate analysis calculates the coal bearing rock matrix proportions (%fixed carbon, %volatile matter, %ash and %moisture) without the use of the lab measured proximate data to calibrate the data. The multiple regression analysis does the exact opposite, whereby it uses the lab measured proximate data to establish a relationship with the logging tools and then applies the empirical relationship to quantify the coal bearing rock matrix proportions. The results of both methods are tested to determine if the proposed pseudo proximate analysis method is statistically accurate.

For the pseudo proximate analysis, the first step is to determine whether there is a relationship between the logging tools and the coal bearing rock. This is done using a cluster analysis for rock typing. This identifies the logging tools that will be used in the pseudo proximate analysis calculation. The second step is a principle component analysis to better understand the multidimensional relationship between the logging tools and the proximate data in a two dimensional space. The results of this guide the equations of the pseudo proximate analysis to calculate the different proportions of the coal bearing rock matrix.

Using this information a method is proposed, the pseudo proximate analysis (PPA), to quantify the coal bearing rock as a function of multiple logging tools. This method will attempt to match the data described in the lab proximate analysis, by modelling the volume percentages of fixed carbon, volatile matter, ash and moisture. This method is different to other methods due to one main principle, the pseudo proximate analysis does not estimate a density value for coal or for clay. Additionally, the method proposed uses the unprocessed neutron (counts per second) wireline log as direct measurement of the hydrogen concentration in the coal bearing rock matrix. The equations established by the pseudo proximate analysis method can then be applied to the well log, creating multiple continuous logs showing the proportion of fixed carbon, volatile matter, ash and moisture in the rock matrix.

The method looks at all the available and useable well logs as one dataset, rather than modelling each well log separately. Although this has the potential to increase the error, it means that this method (PPA) can then be applied to all wells, even those without proximate sample data. This is the ultimate goal of the pseudo proximate analysis, the accurate calculation of the coal bearing rock matrix properties without lab measured calibration data.

The results of the pseudo proximate analysis are statistically compared to the results of the multiple regression analysis. A comparison between the results of the proximate analysis and multiple regression analysis removes the components of internal variance and measurement error for both the lab measured proximate results, as well as the wireline log measurements. The multiple regression
analysis acts as the best case scenario, to which the pseudo proximate analysis is tested.

The statistical tests prove that there is a significant difference between results of the pseudo proximate analysis and the lab measured proximate analysis. Therefore, proving that the proposed pseudo proximate method does not correctly model the surface coal bearing rock matrix properties. The recommendations for improving the method are discussed. The most crucial component being the established relationships between the wireline logs and the proximate analysis components (%ash, %moisture, %fixed carbon and %volatile matter).

This is the first step to modelling coal bearing intervals using this particular combination of logging tools. The pseudo proximate analysis requires further investigation using more advanced tools (such as the gamma ray spectrometer and micro-resistivity) and better calibration (neutron response calibration to different coal bearing rock types). While there are more direct measurement methods available, such as Schlumberger’s Elemental Capture Spectroscopy log, these cost additional money to run and are not often used.

2. Regional overview
The 9 wells drilled are the first in this area, although the south eastern margin of the basin has been extensively studied for the coal mining activities (Figure 2). (Advanced Resources International, 2003; Bordy et al., 2010; Cairncross, 2001)

The area of interest, the location of the 9 drilled wells, is located to the north west of the coal mining activity. The area is overlain by the Stormberg lava

![Geological map showing location of outcropping formations. Area circled in blue represents the area of coal mining and the area most studied. The outlined red area indicated the location of the drilled wells used in this study. Modified from Smith (1984)](image-url)
group, which restricts the use of gravimetric data. The area is flat, less than 1 degree dip, and has undergone little to no deformation since deposition (Smith, 1984).

There are two coal bearing formations, Serowe and Morupule, these are the target formations for interpretation of the coal bearing rock matrix properties, the pseudo proximate analysis. The reference well, as determined by Smith (1984), describes the formations of the Ecca Group (Figure 3).

Kamatoka Formation lies unconformably over the pre-Karoo basement in the study area; it shows a fining upward sequence of reworked Archaean granite and Waterberg sediment. Lithology of the formation includes feldspathic sandstone, fining upwards into mudstone and in some areas carbonaceous mudstone. The deposition occurred during a period of uplift, the depositional environment has been interpreted as an erosive, highly channelized deltaic plain with little of the overbank deposits remaining (Smith, 1984).

The first target, coal bearing, interval is the Morupule Formation. The formation start is indicated by the first coal seam, Member F1. This indicates a hiatus in the wide spread deposition, allowing for a thick blanket of peat to form in the swampy plains, in a tundra-like climate. The coal thickness is controlled by the differential compaction of the underlying sandstone, and the quality, or rank, is dependent on the water depth. An increase in subsidence and clastic input during Member F2 would account for the development of channels and muddy overbank deposits. This Member has lower rank coal seams, with the main lithology indicated as carbonaceous mudstone. The final Member of the Morupule Formation, F3, shows an increased clastic input with an overall increase in grain size. The coal seams found in this Member are thinner.

The second target interval, the Serowe Formation, has a base, Member G1, of fine grained siltstone typically finely laminated or rippled. There is evidence of bioturbation and root trace fossils towards the top of the member. Generally, this member does not contain coal seams, but is also not laterally extensive. Member G2, may contain thin bright coal seams alternating with carbonaceous mudstone. The final member, G3, marks the end of the Serowe Formation. It is a 1 to 2 meter thick bright coal seam with carbonaceous mudstone laminations. The coal is brighter towards the southern margin of the basin, and dulls the further
away from the coal mining activity. The interpretation of the Serowe Formation, based on Smith (1984), is a low energy, stable environment. The preservation of fine lamina, ripples and bioturbation indicates low clastic input.

The Tlhabala Formation overlies the Serowe Formation, and is a non-coal bearing interval. The majority of the unit is a silty mudstone, with occasional evidence of limestone concretions. The deposition environment, interpreted by Smith (1984), is a widespread, flat and shallow lacustrine environment.

The analysis of the deposition environment for each well is outside the scope of this study. However, the pseudo proximate analysis can be used as a tool for definition of coal layers and correlation between coal layers.

Lastly, the area is affected by a large number of southeast-northwest trending mafic dyke swarms (Le Gall et al., 2005). These cut through the pre-Karoo basement and the formations of interest. This is seen in 6 of the wells drilled, namely C2, C3, C4, C5, C8, C9. However, the effect of this on the coal intervals is not measured.

3. Methodology

This project is divided into three main sections each containing subsections. A detailed overview of each section is discussed in this chapter.

Identify
The first section is to identify useful well logs using a combination of statistical methods. The first approach is to use cluster analysis for rock typing. This method identifies different rock types by clustering the sample points based on a selection of means and standard deviations for multiple wireline logs. The wireline logs that provide the best discrimination for different rock types are used as inputs for the pseudo proximate analysis. The selected tools are the gamma ray, neutron and density.

Normalize
The tools selected all measure the rock property in different units. These are transformed into indices that measure rock matrix properties. The gamma ray is normalized from °API to a measure of the volume of clay. The density log is normalized from g/cm$^3$ into a measure of the ash volume. And lastly the neutron log is transformed from its raw form, counts per second (cps), into a measure of hydrogen.

Relationship
The relationship between the selected logging tools and the results of the proximate analysis must be quantified. This is done using the normalized (or indexed) well logs. The average values of each index over the coal sample intervals are extracted for the gamma ray clay index, the density ash index and the neutron hydrogen index. This results in a table showing the proximate analysis and the average clay, ash and hydrogen index associated with that sample.

A principle component analysis is completed to determine the relationship, in a multidimensional space, between the logging tools and the proximate analysis. This method quantifies the correlation between logging tools and the proximate analysis samples.

Figure 4: Overview of the pseudo proximate workflow from the identification of well logs to use in calculating the pseudo proximate analysis through to the application of the results to the well logs. Each of these steps in the workflow are detailed as sections in this article. A more detailed description of the workflow is shown in figure 10, outlining the inputs for each step.
Calculate
The equations and procedure for calculating a pseudo proximate analysis are discussed. The rock matrix is divided into four sections; these are %clay, %coal, %quartz and %moisture. A division of %coal into its components of %volatile matter and %fixed carbon is also discussed.

The end of this section shows a sensitivity analysis based on the selected minimum and maximum values used to create the index. This is a measurement of the potential error in the pseudo proximate analysis calculation.

Correlate
The calculated pseudo proximate analysis results are then correlated to the lab experiment proximate analysis. A multiple regression analysis is conducted to provide a base line correlation, to which the pseudo proximate analysis results can be tested.

Compare
To determine whether the correlation between the lab measured and pseudo proximate analyses is statistically significant, multiple tests on the effect size are conducted. These tests will prove or disprove the proposed method for calculating a pseudo proximate analysis.

4. Lab Proximate Analysis
A proximate analysis is a common laboratory process for analysing coal bearing rock. It measures, using cored samples, the total weight percent of fixed carbon, volatile matter, ash (clay and quartz) and moisture. The results are generally used to calculate coal rank, gas-in-place, specific heat values and other parameters necessary to quantify the quality of the coal, both for coal mining and coal bed methane extraction (Speight, 2005).

In order to compare the results of the proximate analysis to the subsurface in-place rock matrix a clear understanding of the methods used and measurements taken during the proximate analysis is needed.

There are four steps to the proximate analysis, after each step the weight loss of the powdered coal sample is measured. The proximate analysis method below is briefly outlined without an in-depth explanation. (Speight, 2005)

1) Moisture (wt %)
   • Heated up to 108 ± 2 °C to evaporate the water component of the sample

2) Volatile Matter (wt %)
   • Heated up to 900 ± 15 °C in a closed environment to remove the volatile matter component without oxidation of the carbon

3) Fixed Carbon (wt %)
   • Heated under a Bunsen burner to burn, or oxidize, the carbon from the sample

4) Ash (wt %)
   • This is the remaining weight of the powered sample

The significance of the proximate analysis can be described by the components that they measure.

1) Fixed Carbon
   • Consists mostly of pure carbon with some sulphur, nitrogen, hydrogen and oxygen not lost during the %volatile matter measurement process

2) Volatile Matter
   • This is a measure of the gaseous fuels present in the sample. These include methane, hydrocarbons, hydrogen and carbon monoxide. This may also measure some incombustible gases such as nitrogen and carbon dioxide

3) Ash
   • This is a measure of the impurity of the coal, the proportion of the sample that will not burn

4) Moisture
   • Usually a low proportion of the sample and a measure of the water found either in pore spaces or bound to clay and coal.
As a summary the %fixed carbon is a measure of carbon, the %volatile matter is a measure of hydrogen, the %ash is a measure of the non-combustible components such as quartz, oxides and clays, and lastly %moisture is a rough estimate of the water in the rock matrix.

It stands to reason that not all of the components measured in the proximate analysis can be directly measured by the logging tools. However, a combination of logging tools might be able to differentiate the different rock matrix components.

The last point to mention is that in a non-coal bearing rock both %volatile matter and %fixed carbon, by definition, must be equal to zero. Therefore, a method must be determined to define non-coal bearing versus coal bearing rock units. This then also implies that the percentage of coal in the system is the sum of %fixed carbon and %volatile matter.

5. Previous Work
There are two fields in determining organic carbon concentration in a rock matrix. The first is the measuring of total organic carbon (%TOC), this relates to shale units with low concentrations of organic matter and is associated with the production of hydrocarbon. The organic carbon within the rock matrix is usually less than 10%. The second field is the measurement of coal rock matrix properties, coal by definition has a greater than 50% organic matter. This field of study then measures higher concentrations of organic matter.

5.1. Literature - %TOC
The estimation of total organic carbon (TOC) in shale units, although not directly applicable, is similar to the estimation of %fixed carbon in coals. The reason why we look at TOC estimation methods is to try and apply the basic principles to a coal estimation method.

Many methods have been developed to measure TOC with the most popular and widely used being DLogR method developed by Exxon and Esso (Passey et al., 1990). This is one of the few methods that make use of multiple logs to calculate TOC. Using a combination density, sonic and neutron porosity and cross plotting against resistivity the TOC can be estimated, however it requires the interpreter to define a fixed superposition coefficient and to determine the LOM (level of organic metamorphism). These parameters vary locally and for accurate results a complete basin model is needed (Sun et al., 2013).

Another method of calculating TOC, the CARBOLOG method, developed by the French Petroleum Institute in 1988 uses sonic and resistivity (assuming that the resistivity of the rock frame and organic matter is infinite) to determine the proportion of rock frame, clay, water and organic matter. Three of the parameters are needed in order to calculate the fourth. Calculating TOC is achieved by plugging the other three parameters into the chart and reading the result. There are two problems with the method. The first is the complexity of calculating TOC (Sun et al., 2013). The second problem is that the calculation of TOC is dependent on the accuracy of the methods used to estimate clay, rock frame and water. The assumption that the resistivity of the rock frame and organic matter is infinite is a reasonable assumption to make, as long as the rock frame is pure quartz and the organic matter is pure carbon.
It has been proven that no single logging curve can be used to accurately estimate the TOC shown in Figure 5 (Sun et al., 2013).

Sun, et al. (2013) continues with testing three TOC calculating methods against measured TOC in order to determine which method most accurately estimates TOC. The methods that he tests are the DLogR, optimal superposition coefficient DLogR and the CARBOLOG method. The conclusion is that the CARBOLOG method is the most accurate in estimating TOC (Figure 6).

In summary, there have been many attempts to calculate TOC with the most accurate having a correlation coefficient of 83%. Although application of the TOC does not directly apply to the estimation of coal proximate parameters (%fixed carbon, %ash, %moisture and %volatile matter), it does provide an understanding of the following basic principles:

1. No single log curve can be used to estimate organic content. There must be a combination of logging tools
2. One needs to define units containing organic matter versus units not containing organic matter
3. A high correlation coefficient is unlikely

5.2. Literature - Coal Analysis

There are currently multiple methods to quantify coal parameters (fixed carbon, volatile matter, moisture and ash volume percentages) these include density-neutron crossplots, density-sonic crossplots, density matrix vs. Uma and M\text{lihth}-N\text{lihth} crossplots and three mineral models.

All of the aforementioned methods rely on the assumption that matrix density of a coal bearing rock is fixed, usually using a range of values from 1.19 (lignite) to 1.47 (anthracite) g/cm\(^3\) (van Krevelen, 1954). In order to calculate the true matrix density of coal the relative percentages of anthracite, bituminous and lignite coal must be known under subsurface conditions. When using a sonic model to estimate the rock matrix properties a value is selected depending on the coal rank, usually 345 μsec/meter for anthracite and 525 μsec/meter is used for lignite (Rieke, et al., 1979). However, a single deposition environment may have multiple ranks of coal, therefore selecting a single value for sonic may prove erroneous.

Furthermore, it is not possible to select a single correct sonic or density value as coal zones are not homogenous. The grade of coal is dependent on the proportion of starting organic matter, clay and quartz. Varying these rock matrix proportions will have a large effect on the bulk matrix density and the sonic velocity. To correct for this the methods depend heavily on calibration with lab measured proximate analysis results. Without a large amount of lab measured proximate samples from cores these methods are highly speculative.

The majority of coal log analysis is descriptive, using multiple well log cut offs to define coal and non-coal litho-units. There have been few attempts to quantify coal properties based on wireline log responses. Srinaiah, et al. (2014b) propose the use of density, resistivity and gamma ray to identify different lithologies using a cluster analysis.

A descriptive overview on log responses to coal is provided by (Rieke, et al., 1979), they discuss the
theory behind the specific log responses to coal. From their observations the following is determined (Rieke, et al., 1979):

- Gamma Ray
  - Coal seams are identified with very low natural radioactivity. However, under reducing environments there may be secondary uranium enrichment of the coals increasing the natural radioactivity of coals.

- Density
  - Determined to be excellent for coal identification and evaluation. Used, in most cases, to estimate the ash content. Density is a function of rank, water content, type of mineral matter (clay or quartz), maceral composition and gas saturation.

- Spontaneous Potential
  - Coal seams may have some permeability and will therefore respond similar to clastic sediments.

- Resistivity
  - Coal is very resistive and its resistivity is a function of multiple unrelated physical and petrophysical properties. These include mineral composition and degree of metamorphism.

- Sonic Travel Time
  - High travel times are observed in coals. The measurements must be corrected for compaction.

- Neutron Response
  - Based on their understanding the high carbon content moderates the neutron response and it results in low counts per second.

The density tool is commonly used to measure the ash volume within coal bearing rock matrix (Rieke, et al., 1979). Ryan (1990) uses densities measured at different stages of the proximate analysis to determine an equation to measure ash volume in coal given the bulk density of the rock.

There have been more recent attempts to quantify coal parameters using well logs, namely Rai, et al. (2004) and Srinaiah, et al. (2014a). However, careful scrutiny of the work by these authors has demonstrated that they rely on a single logging tool and depend heavily on calibration to the measured proximate analysis, hence they will not be discussed.

The method proposed in this article fundamentally differs from past methods based on one principle. In other methods, such as using sonic-density, density-neutron, $M_{inh-N_{inh}}$ and three mineral model cross plots, the density of both clay and coal must be known or estimated in order to calculate coal bearing rock matrix properties. This is a fundamental error as the density of coal will change depending on multiple factors, these include burial depth, maceral type and the rank of coal. The method proposed in this thesis does not require the use of unknown, or poorly constrained, density cut off values for coal.

A further difference is the absence of the sonic and resistivity logging tools in the pseudo proximate analysis method. The sonic and resistivity log tools measure the rock matrix properties over a larger depth interval (lower resolution) compared to the gamma ray, neutron and density tools. The maximum resolution obtainable is dependent on the tool with the lowest resolution, therefore as this method does not use the resistivity and sonic logs it is able to resolve thinner coal units.

Lastly, there is no method that uses the raw neutron count as a measure of hydrogen in the rock matrix. However, the neutron tool is known to indirectly measure %volatile matter (Thomas, 2002). This new approach is used based on the principle that coal's chemical composition is hydrogen rich compared to clay and water. Therefore, the proposed method uses the neutron log to distinguish between non-coal and coal bearing units, as well as to calculate the proportion of coal within the rock matrix.
6. Pseudo Proximate Analysis

The definition of a “pseudo proximate analysis”, as defined by this thesis, is the estimation of fixed carbon (%fixed carbon), volatile matter (%volatile matter), moisture (%moisture) and ash (%ash) volume percentage of total rock matrix calculated from wireline logging tools. The goal is to calculate the pseudo proximate analysis for the entire depth of the well log using the wireline logs, this includes the zones that have not been cored or sampled. The results should correlate to the measured proximate analysis from the core samples.

The dataset consists of 9 well logs, 7 of the wells contain gamma ray, density, resistivity, neutron and sonic logs. The remaining 2 wells do not contain neutron logs. The wells are drilled to depths between 350 and 550 meters and are cored over 90% of the target intervals, being Serowe and Moropule.

6.1. Coal Analysis - Cluster Analysis

Based on general logging tool theory and the observations discussed in the previous section there are a few tools which measure multiple factors and a few tools that measure a single phenomenon. In order to determine the tools needed to quantify a coal bearing rock matrix multiple cross plots were

![Figure 7: Cluster analysis for rock typing. Above table shows the mean value and standard deviation for each cluster. The means are displayed as coloured points on the cross plots. Six clusters, or rock types, were identified based on varying ratios of coal, clay and silt (and dolerite). The clusters are named based on the predicted bulk rock matrix, shown in the bottom legend. The most interesting curves are analysed in more detail in the following section.](image-url)
used. These cross plots are then run through a multi-parameter cluster analysis to identify logging tool combinations that signal certain rock matrix properties (Figure 7).

Gamma ray is measuring the total natural radioactivity, this is usually directly proportionate to the volume of clay in the rock matrix (Marett, 1978). Factors that might influence this are the amount of background radiation and post-deposition environmental changes (reducing environments, enrichment varying clay composition). If these factors are accounted for the assumption can then be made that the gamma ray reading is only a direct result of the amount of clay in the system. This assumption is used in the petroleum industry to quantify the percentage volume of clay in the rock matrix (Figure 8).

The density log is a measure of the bulk rock matrix. This does not measure a specific property of the rock, as both fluid and the whole rock frame are measured. It is used as a measure of porosity, when the fluid and rock frame densities are known. However, it can also be used as a measure of percentage volume of quartz, only once the majority of the other components (clay and coal) influencing the bulk density are quantified (Figure 8).

The neutron response, in its raw form (counts per second), is a direct measure of the amount of hydrogen in the rock matrix. In a coal, clay and silt mixed deposition system the forms of hydrogen expected are H₂O (moisture), -OH⁻ (clay minerals), CH₄ (methane) and maceral. In order to better understand the relationship between the neutron response to coal, the components of coal must be understood. Coal is described as Type III kerogen (van Krevelen, 1961), and is a combustible sedimentary rock consisting of lithified organic plant matter. Macerals are the microscopic insoluble organic components in coal (Thomas, 2002), these are complex aromatic, poly-aromatic and hydroaromatic compounds formed from terrestrial, marine and lacustrine plant remains. There are many different maceral types depending on the starting organic matter type, initial decomposition and post-deposition diagenetic and maturation processes (Suárez-Ruiz, 2012). Defining a chemical structure and composition for coal is a broad and complex topic that falls outside the scope of this thesis. However, the aromatic compounds of coal minerals can be defined in their simplest form as illustrated in Figure 9 (Heredy & Wender, 1980).

The hydrogen concentration in coal bearing zones is larger than in the pure clastic zones, this is due to the increased concentration of hydrogen bearing gases associated with coal bearing rock (Thomas, 2002). Based on the standard tool theory, the neutron tool
is an inverse measurement of hydrogen concentration. If the hydrogen concentration is proportionate to the amount of coal, then it follows that the neutron response is inversely proportionate to the percentage volume of coal in the rock matrix (Figure 10).

The relationship between the three selected logging tools, shown in Figure 11, is used to define the rock matrix components. A density-neutron relationship is used to determine coal bearing versus non-coal bearing rock units. This relationship is based on the results of the normalization and is discussed in more detail in chapter 6.3.

The density-neutron relationship also acts to measure the amount of coal in the matrix. The greater the separation between the neutron and density index the greater the amount of coal. An example of this is a very low density and a low neutron count (higher hydrogen concentration), indicated as a red circle on figure 12.

The three logging tools used in the pseudo proximate method are the gamma ray, density and neutron tools. The remaining tools that are not used for calculating pseudo proximate analysis are:

- Resistivity
  - The measurement is affected by multiple variables that cannot be quantified and therefore the necessary information needed to calculate the pseudo proximate analysis cannot be extracted (Rieke, et al., 1979)
  - The inherent measurement method (averaging over a specific height) results in a lower resolution compared to other logging tools.

- Sonic (Acoustic Travel Time)
  - Although this provides a good measurement of the bulk rock properties it too has a lower resolution compared to other logging tools (such as density).
  - There is an inverse relationship between density and sonic, essentially showing the same information.

- Spontaneous Potential (SP)
  - The goal is to separate coal from clastic sediments in order to quantify the relative percentages.
While both the acoustic properties and resistivity of the rock do indicate coal bearing zones, using tools with higher resolution allows for the rock matrix to be processed at higher detail (lamination rather than bed scale).

### 6.2. Coal Analysis - Normalization

Determining the equation to calculate a pseudo proximate analysis from well logs is outlined below (Figure 13). The first three steps outline the method used to calculate the equation for the pseudo proximate analysis, and the last three steps describe these results using studies to measure the error range and accuracy of the calculations. They have been discussed above and the next step is to discuss the procedure used to normalize the input well logs into the parameters that they represent.

In this method the rock matrix is divided into four components these are clay, coal, quartz (silt) and pore volume. The goal in this section is to normalize the three selected curves (gamma ray, density and neutron) into curves that represent one or more of the properties of the rock matrix. However, in order to achieve this certain assumptions will be made:

- Gamma ray is an absolute measurement of the total volume of clay in the system
- Density is a measure between pore volume and volume of quartz
- Neutron is directly proportionate to the amount of organic matter (% coal, being the sum of %fixed carbon and %volatile matter)

The assumption that gamma ray represents the absolute volume of clay is valid under certain

---

Figure 12: 3D representation of the relationship between density, gamma ray and neutron using wells C1 to C7. This method does not use cut off values to determine rock matrix components, instead a relationship between density and neutron hydrogen count is used to distinguish coal bearing versus non-coal bearing rock units. This indicates that a rock unit with a lower density may have a higher neutron count (or lower hydrogen concentration) and still be classified as a coal bearing rock. The red circle indicates the points which have the highest coal concentration in the rock matrix. The green arrow represents increasing clay content, and is based on the gamma ray log. Points falling outside the plot area, indicated by the pink circle, represent dolerite.
conditions. This assumption is valid when the clay mineralogy remains relatively constant, there is no or very little change in clay mineralogy. This is interpreted as true due to the fact that the intervals are buried to a maximum of 550 meters restricting the alteration of clay and the formation of secondary clay minerals. Smith (1984) states that the primary source rock material of the Ecca Group is sourced from basement Archean granites, indicating that the source rock mineralogy is uniform. This observation is further proof that the clay mineralogy is relatively uniform throughout the coal bearing intervals. Coal bearing rock, or rock containing organic matter, may contain radioactive elements such as uranium. This may affect the gamma ray readings, however it is interpreted that in this area there is no or very minor uranium in the coal intervals. This assumption is made based on two facts, the first is the age of the Stormberg lava group dates back to 180.5 ± 2.2 Ma in northern Botswana (Jourdan et al., 2007), indicating that the coal are older than this and that the uranium has at one point leached out of the coal during a reducing period. Secondly, the gamma ray response in coal bearing rock units is very low. There is a proportionate decrease in gamma ray with increase in coal rank. The last factor that may influence the gamma ray reading, not linked to clay, is the presence of potassium bearing feldspar. The sedimentary system for the coal units is very fine grained indicating a very distal source rock (Nichols, 2011). There is very limited quartz in the system and no evidence of feldspar.

The assumption that density index is a measure between pore volume and the volume of quartz is only valid under two circumstances. The first is if the rock matrix consists only of quartz and pores (porosity) and that the pores are saturated with mud filtrate. In coal bed methane bearing rock this very unlikely. The second circumstance is that this might prove to be true if all the components of the rock matrix are calculated. If the volume of clay and coal are known in the rock matrix what remains is quartz and pore volume then the relationship may be applied.

The neutron tool is a measure of hydrogen, and it has been discussed in chapter 6.1 that coal contains a much higher proportion of hydrogen when compared to clay, water and mud filtrate. This assumption is therefore valid under most circumstances, the question however remains how to quantify the relationship between the neutron reading and the volume of coal.

Figure 13: Pseudo proximate analysis workflow defined in 6 steps. Steps 1 to 3 outline the process of defining the pseudo proximate analysis equations. Steps 4 and 5 determine the accuracy of the method, as well as the conditions under which the method fails to deliver accurate results. The final step, 6, is the application of the method to the available well logs and the analysis thereof.
Histograms are used in order to accurately select the minimum and maximum values to normalize the curves. In order to remove data spikes and erroneous tool readings the maximum will be defined a 98% of the cumulative total (P98). The minimum will be defined as 2% of the cumulative total (P02). A sensitivity analysis, using a tornado diagram, will later show how selecting a different minimum and/or maximum value influences the final pseudo proximate analysis results.

The normalization is used to convert the measurements of the tools into values that represent rock matrix properties. Table 1 shows each individual tool and the results of the conversion process.

<table>
<thead>
<tr>
<th>Tool Measurement</th>
<th>Normalized Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma Ray (°API)</td>
<td>Gamma Ray Index</td>
</tr>
<tr>
<td>Bulk Density (g/cc)</td>
<td>Quartz Index</td>
</tr>
<tr>
<td>Raw Neutron (CPS)</td>
<td>Hydrogen Index</td>
</tr>
</tbody>
</table>

Table 1: Conversion from tool measured values into rock matrix properties.

The relationship between the normalized tools and the rock matrix is shown in table 1. It shows that gamma ray is used as an absolute measurement of the volume of clay in the rock matrix. Once this is defined the next step is to define the volume of coal in the rock matrix, this is done using the separation between the hydrogen index and the density index. In a rock absent of both clay and coal, and only containing quartz and pore fluid the density index is the inverse of the porosity. Therefore, under these conditions the density index shows the percentage of rock volume that is quartz.

For the pseudo proximate analysis the required logging tools are gamma ray, neutron and density. Lab experiment proximate analysis data is required to test the results of the pseudo proximate analysis. Table 2 below describes the well logs that meet the above requirements, and the total thickness of the formations analysed.

<table>
<thead>
<tr>
<th>Well</th>
<th>Serowe (m)</th>
<th>Morupule (m)</th>
<th>Log Tools</th>
<th>Proximate Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>2.48</td>
<td>74.00</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>C2</td>
<td>10.77</td>
<td>136.86</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>C3</td>
<td>11.83</td>
<td>107.07</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>C4</td>
<td>54.3</td>
<td>64.85</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>C5</td>
<td>43.26</td>
<td>31.18</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>C6</td>
<td>13.56</td>
<td>37.68</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>C7</td>
<td>43.03</td>
<td>42.05</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>C8</td>
<td>18.7</td>
<td>93.72</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>C9</td>
<td>18.15</td>
<td>107.1</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 2: Total thickness of the target intervals, Serowe and Morupule. The table also shows the data availability of the proximate analysis and the wireline logs.

Therefore, the pseudo proximate analysis cannot be determined for both well C8 and C9. Well C1 does not have lab proximate analysis data and cannot be used to confirm the results of the pseudo proximate analysis. The data used for comparison with the lab proximate analysis has a total of 53.11 meters of core, analysed over 105 samples. The number of wireline log data points over the same interval of core samples is 5311, at a step interval of 0.01 meters.

6.2.1. Gamma Ray Normalization

The gamma ray histogram (Figure 14) shows a distinct bimodal distribution, the first hump representing the organic matter (coal) and quartz (silt) and the second hump represents clay in varying rock matrix percentages.

The assumption is made that both quartz and coal contribute no natural radioactivity to the total rock matrix. This assumption is only true if there is no secondary uranium enrichment of coal. Therefore, all measured natural radioactivity (using the gamma ray tool) is a direct result of the volume of clay in the rock matrix. This is true if the mineralogical composition of the coal is relatively constant throughout the measured interval. This is not limited to depositional variances and post deposition diagenesis should also be uniform throughout the interval of interest. In this area there is no reason to suspect major compositional changes in the clay, based on observations by Smith (1984).
In order to quantify the volume of clay in the rock matrix, a maximum gamma ray reading must be selected to define the natural radioactivity of pure clay and a minimum gamma ray value is selected to define the background natural radioactivity. These values are selected using the histogram with the maximum (pure clay, indicated by the red line on the gamma ray histogram) being defined as $P_{98}$ equal to 226.67 °API and the minimum (no clay, indicated by the black line) defined as $P_{02}$ equal to 9.875 °API. This results in 2% of the interval (13.46 meters) containing pure clay and 2% of the interval containing no clay. Based on the histogram 50% (336.46 meters) of the interval has a clay volume greater than 50%, already indicating that the deposition system is very clay rich and is thus interpreted as a low energy environment.

The equation to calculate clay volume (Equation 1) is used define the volume of clay in the rock matrix, based on the gamma ray reading (Marett, 1978).

$$%\text{Clay} = \frac{GR_{P_{98}} - GR_{P_{02}}}{GR_{P_{98}} - GR_{P_{02}}}$$

![Figure 14: Gamma ray histogram showing a bimodal distribution. The first distribution, ranging from a gamma ray of 0 to 40 °API, represents both the coal and quartz (silt) rich rock matrix. The increasing gamma ray °API is directly proportionate to the increase in volume clay. The maximum value, 226.67 °API ($P_{98}$), was chosen to represent the clay cut off indicating 100% clay in rock matrix, which results in 2% of the interval of interest being pure clay. Whereas the minimum value, 9.875 °API ($P_{02}$), indicates the cut off at which there is no clay in the system (this value is representative of the background natural radioactivity), which results in 2% of the interval of interest not containing clay. A total of 672.92 meters is under investigation, resulting in the use of 67,292 data point measurements.](image)

<table>
<thead>
<tr>
<th>Well</th>
<th>Top (m)</th>
<th>Bottom (m)</th>
<th>Meters</th>
<th>Min (°API)</th>
<th>Max (°API)</th>
<th>Std. Dev.</th>
<th>Mean (°API)</th>
<th>Mode (°API)</th>
<th>$P_{97}$</th>
<th>$P_{98}$</th>
<th>$P_{99}$</th>
<th>$P_{01}$</th>
<th>$P_{02}$</th>
<th>$P_{03}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>242.4</td>
<td>244.88</td>
<td>2.48</td>
<td>24.75</td>
<td>193.8</td>
<td>39.014</td>
<td>114.06</td>
<td>146.25</td>
<td>173.97</td>
<td>175.62</td>
<td>190.4</td>
<td>36.35</td>
<td>38.05</td>
<td>38.575</td>
</tr>
<tr>
<td>C1</td>
<td>244.88</td>
<td>318.88</td>
<td>74</td>
<td>12.5</td>
<td>330.37</td>
<td>51.249</td>
<td>140.12</td>
<td>164.25</td>
<td>233.17</td>
<td>242.42</td>
<td>257.1</td>
<td>31.375</td>
<td>37.05</td>
<td>41.225</td>
</tr>
<tr>
<td>C2</td>
<td>241.04</td>
<td>251.81</td>
<td>10.77</td>
<td>66.65</td>
<td>197.75</td>
<td>22.028</td>
<td>129.32</td>
<td>132.75</td>
<td>170.07</td>
<td>174.2</td>
<td>183.72</td>
<td>80.675</td>
<td>81.9</td>
<td>87.825</td>
</tr>
<tr>
<td>C2</td>
<td>251.81</td>
<td>388.67</td>
<td>136.86</td>
<td>0.55</td>
<td>299.85</td>
<td>65.768</td>
<td>90.562</td>
<td>20.25</td>
<td>211.27</td>
<td>220.92</td>
<td>234.15</td>
<td>7.35</td>
<td>9.075</td>
<td>10.5</td>
</tr>
<tr>
<td>C3</td>
<td>341.07</td>
<td>352.9</td>
<td>11.83</td>
<td>19.332</td>
<td>209.98</td>
<td>35.184</td>
<td>126.89</td>
<td>146.25</td>
<td>177.55</td>
<td>180.91</td>
<td>190.38</td>
<td>28.904</td>
<td>34.25</td>
<td>36.691</td>
</tr>
<tr>
<td>C3</td>
<td>352.9</td>
<td>459.97</td>
<td>107.07</td>
<td>0.65</td>
<td>320.75</td>
<td>62.043</td>
<td>97.894</td>
<td>20.25</td>
<td>206.86</td>
<td>217</td>
<td>231.91</td>
<td>8.02</td>
<td>9.8799</td>
<td>11.194</td>
</tr>
<tr>
<td>C4</td>
<td>379.72</td>
<td>434.02</td>
<td>54.3</td>
<td>0.425</td>
<td>239.52</td>
<td>51.053</td>
<td>50.606</td>
<td>15.75</td>
<td>154.57</td>
<td>161.9</td>
<td>173.32</td>
<td>4.6</td>
<td>5.275</td>
<td>6.2</td>
</tr>
<tr>
<td>C4</td>
<td>434.02</td>
<td>498.87</td>
<td>64.85</td>
<td>13.025</td>
<td>225.77</td>
<td>36.848</td>
<td>101.56</td>
<td>114.75</td>
<td>168.22</td>
<td>174.77</td>
<td>184.55</td>
<td>23.075</td>
<td>27.2</td>
<td>30.175</td>
</tr>
<tr>
<td>C5</td>
<td>274.09</td>
<td>317.35</td>
<td>43.26</td>
<td>1.1</td>
<td>283.5</td>
<td>61.126</td>
<td>98.754</td>
<td>15.75</td>
<td>201.12</td>
<td>211.02</td>
<td>231.07</td>
<td>6.025</td>
<td>8.35</td>
<td>9.925</td>
</tr>
<tr>
<td>C5</td>
<td>317.35</td>
<td>348.53</td>
<td>31.18</td>
<td>26.7</td>
<td>302.52</td>
<td>46.584</td>
<td>129.06</td>
<td>150.75</td>
<td>216.42</td>
<td>222.1</td>
<td>235.02</td>
<td>36.65</td>
<td>41.325</td>
<td>47.025</td>
</tr>
<tr>
<td>C6</td>
<td>305.44</td>
<td>319</td>
<td>13.56</td>
<td>63.543</td>
<td>424.77</td>
<td>73.606</td>
<td>180.51</td>
<td>132.75</td>
<td>346.77</td>
<td>357.01</td>
<td>366.32</td>
<td>83.852</td>
<td>89.112</td>
<td>91.678</td>
</tr>
<tr>
<td>C6</td>
<td>319</td>
<td>356.68</td>
<td>37.68</td>
<td>2.1238</td>
<td>287.3</td>
<td>59.05</td>
<td>137.73</td>
<td>164.25</td>
<td>223.75</td>
<td>239.72</td>
<td>251.55</td>
<td>15.913</td>
<td>19.713</td>
<td>22.832</td>
</tr>
<tr>
<td>C7</td>
<td>403.32</td>
<td>446.35</td>
<td>43.03</td>
<td>3.0386</td>
<td>271.43</td>
<td>50.347</td>
<td>100.95</td>
<td>128.25</td>
<td>189.82</td>
<td>197.62</td>
<td>209.21</td>
<td>11.446</td>
<td>14.755</td>
<td>17.574</td>
</tr>
<tr>
<td>C7</td>
<td>446.35</td>
<td>488.4</td>
<td>42.05</td>
<td>3.743</td>
<td>318.02</td>
<td>46.443</td>
<td>127.61</td>
<td>132.75</td>
<td>218.64</td>
<td>228.11</td>
<td>247.75</td>
<td>25.425</td>
<td>39.535</td>
<td>45.811</td>
</tr>
</tbody>
</table>

Total of 672.92 meters is under investigation, resulting in the use of 67,292 data point measurements.
Equation 1: Calculating the total volume of clay (%Clay) in the rock matrix. P02 and P98 are defined by the histogram. If the result is \( >1 \) then the %Clay is simply set to 1 (being 100% clay in the rock matrix). If the result is \( <0 \) (negative) then %Clay is set to 0 (0% clay in the rock matrix).

If the gamma ray reading is larger than 226.67 °API then the %Clay calculated is larger than 100%. In order to preserve the data, the %Clay in those instances is set to 100%. The effect of this is analysed and discussed later. The same procedure is used when the results of the calculation are negative (Equation 1).

\[
\therefore \text{%Clay} = \frac{GR_{log} - 9.875}{226.67 - 9.875}
\]

\[
\text{if } \%\text{Clay} > 1 \text{ then set } \%\text{Clay} = 1
\]

\[
\text{if } \%\text{Clay} < 0 \text{ then set } \%\text{Clay} = 0
\]

6.2.2. Density Normalization

The density histogram (Figure 16) shows a tri-modal distribution. On the figure, the first distribution (indicated in red) represents the coal bearing matrix. The second distribution (indicated by the green arrow) represents a clay rich matrix, with a density range between 2.2 and 2.5 g/cc. The final distribution represents the dolerite (indicated in pink) present in the system, however this is outside the scope of this study. The density of quartz is defined as 2.65 g/cc (orange line); from the histogram it is evident that there is very little quartz-rich matrix in the system.

There is an increase in coal volume with a decrease in bulk density. It is interpreted as such due to the fact that pure coal densities range between 1.19 and 1.47 g/cc. Thus increasing coal volume will result in a decrease in bulk density measured by the density tool. However, in order to define a pure coal density the relative proportions of coal maturity (lignite, bituminous or anthracite), maceral type and volume of impurities (mainly the amount of volatile matter and ash in the rock matrix) must be known. Due to the complexity of estimating a pure coal density, the pure coal density is not defined in this method (as it is in other methods) instead the density tool is used to measure the percentage of quartz against the matrix fluid, being the mud filtrate (Equation 2). The results are referred to as the quartz index (QtzInd) (Denoo, 1978).

\[
\text{QtzInd} = \frac{\text{Dens}_{log} - \text{Dens}_{MF}}{\text{Dens}_{Qtz} - \text{Dens}_{MF}}
\]

\[
\therefore \text{QtzInd} = \frac{\text{Dens}_{log} - 1.05}{2.65 - 1.05}
\]

Equation 2: Quartz index (QtzInd) is a measure of total quartz and clay in the rock matrix, and can only be used to calculate quartz volume when both coal and clay volumes have been calculated.

As this is a measure to distinguish between mud filtrate and ash content (Figure 17), then the relative proportion of mud filtrate (water/moisture) and quartz (silt/sandstone) can be calculated, if the proportion of clay and coal is known. If the proportion of clay or coal is unknown, then the quartz index acts as a measurement of the rock frame versus.

This is used as a measure of porosity in non-coal bearing rock units. However, due to the lower density of coal, the results in coal bearing units are inconclusive as coal is falsely measured as porosity due to its lower density range. In order to determine
coal versus non-coal containing rock units a density-neutron relationship is used; this is discussed in chapter 6.3.

Mud filtrate represents the total pore volume, as it is assumed that the drilled mud has replaced the pore volume. Therefore, at a density of 1.05 g/cc the rock matrix consists of pure fluid, this can either be volatile matter or water (mud filtrate) or, more likely, a mixture of both.

Figure 16: The density histogram shows a tri-modal distribution. These are indicated on the histogram as dolerite, clay-rich rock matrix and coal-rich rock matrix. The cut offs are defined not by minimum and maximum as with gamma ray, but are defined using the same parameters when calculating density porosity. The maximum is the expected density of pure quartz, 2.65 g/cc, and the minimum is defined by the fluid density in the pore system, 1.05 g/cc, being the mud filtrate. The dolerite is outside the scope of this study and hence is ignored (for the purposes of calculating the pseudo proximate analysis).
6.2.3. Neutron Normalization

The neutron tool measures the concentration of hydrogen in the rock matrix, and will be referred to as the hydrogen index (HydInd). This proves useful as there is a large hydrogen concentration increase that occurs when a small percentage of coal is present (Thomas, 2002), which results in a lower neutron (cps) reading. Hydrogen is recorded throughout the rock matrix, in non-coal bearing rock units as water and, in some cases, gas (methane). As discussed before, hydrogen is much more abundant in coal as the building blocks of coal, for example the maceral compounds and associated methane, have been recorded to contain a much higher concentration of hydrogen.

In order to calibrate, or normalize (Figure 18), the neutron count, both end members need to be well understood. The maximum neutron count, being the lowest hydrogen concentration, should represent rock units containing only water. This should be proportionate to porosity, but only if there is no clay in the matrix. As the clay volume increases, the neutron count decreases due to the increase in OH...
associated with clay minerals. Once coal is added to the system there is an immediate jump in hydrogen concentration due to the introduction of maceral compounds in the form of aromatic carbon molecules.

The minimum neutron reading, 207.5 cps, is selected to normalize the neutron log to the highest concentration of hydrogen. This is essentially equivalent to the maximum potential coal volume within all 7 wells. All zones with lower concentration result in less potential coal volume. In order to estimate the actual coal/maceral volume in the rock matrix, a combination of neutron and density will be used, this is discussed in the next section.

A neutron count higher than or equal to 840 cps indicates that there is no maceral component (coal) or clay present in the rock matrix. This corresponds to 2% of the interval, and is linked back to the gamma ray normalization where 2% of the interval contains no clay.

Equation 3, below, shows the calculation used to determine the hydrogen index. This equation is founded on the principles of equations 1 and 2, and it is unique to this method.

\[
\text{HydInd} = 1 - \frac{\text{Neutron}_{\log} - \text{Neutron}_{\min}}{\text{Neutron}_{\text{Pp8}} - \text{Neutron}_{\min}}
\]

\[
\therefore \text{HydInd} = 1 - \frac{\text{Neutron}_{\log} - 207.5}{840 - 207.5}
\]

Equation 3: Calculation of the hydrogen index (HydInd) based on neutron response.

In order to determine the most accurate neutron value that represents maximum potential coal, the results of the pseudo proximate analysis must be calibrated to the lab measured proximate analysis. The correct selection of minimum and maximum neutron values can then be made. This is shown as part of the sensitivity analysis in chapter 7.

Figure 19: Hydrogen normalization results indicating the maximum and minimum neutron counts per second. There is an inverse relationship (high neutron results in low hydrogen concentration)

6.2.4. Normalized Well Log Examples

The example well logs below are selected from well C6 and display the entire Moropule formation. The well logs are the gamma ray, density and neutron readings (Figure 20), they are paired with their respective normalized logs, volume clay, quartz index and hydrogen index. This well was not selected for any particular reason, and the results of the remaining wells are shown in Appendix 1.

From the normalized well logs some interpretations can be made; figure 20 highlights four different log responses for the combination of gamma ray, density and neutron readings. Each of the different regions will be discussed.

Table 3 below describes the tool response from the zone marked 1 on figure 20.

<table>
<thead>
<tr>
<th>Tool</th>
<th>Range</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma Ray</td>
<td>98 - 244 °API</td>
<td>167 °API</td>
</tr>
<tr>
<td>Density</td>
<td>2.19 - 2.51 g/cm³</td>
<td>2.37 g/cm³</td>
</tr>
<tr>
<td>Neutron</td>
<td>617 - 776 cps</td>
<td>661 cps</td>
</tr>
</tbody>
</table>

Table 3: The range and average values for the gamma ray, density and neutron logs over the marker 1 depth interval shown in figure 20.

Initial interpretation of the normalized well logs for marker 1 shows that this area has a high volume of clay with some pure clay laminations. The density falls within the expected density of clay and the neutron count indicates that hydrogen is present most likely due to the hydrogen in the clay minerals, however the response is too weak to be associated with coal.

Marker 2 (Figure 20) shows a layered response, alternating between low and high responses. Table 4
below shows the values for each tool associated with the responses.

<table>
<thead>
<tr>
<th>Tool</th>
<th>Range</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gamma Ray</strong></td>
<td>15 - 183 °API</td>
<td>77 °API</td>
</tr>
<tr>
<td><strong>Density</strong></td>
<td>1.25 - 2.02 g/cm³</td>
<td>1.42 g/cm³</td>
</tr>
<tr>
<td><strong>Neutron</strong></td>
<td>325 - 526 cps</td>
<td>405 cps</td>
</tr>
</tbody>
</table>

Table 4: The extracted range and average values for the density, gamma ray and neutron logs over the interval in figure 20 marked as 2.

Interpretation based purely on the results of the normalized well logs indicates that there are alternating bands of high density, high gamma and high neutron readings interbedded with bands of low density, low gamma and low neutron readings. These bands range from 0.8 to 3.3 meters thick. The first set of bands, viz. high values for gamma, density and neutron readings, may be interpreted as a high ash-containing coal. The other band is interpreted to contain lower ash and higher coal percentages. Figure 21 will later explain the definition used to determine coal versus non-coal bearing rock units.

Marker 3 (Figure 20) highlights a strong coal response. The values for this area are shown in table 5 below.

<table>
<thead>
<tr>
<th>Tool</th>
<th>Range</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gamma Ray</strong></td>
<td>10 - 47 °API</td>
<td>19 °API</td>
</tr>
<tr>
<td><strong>Density</strong></td>
<td>1.22 - 1.56 g/cm³</td>
<td>1.35 g/cm³</td>
</tr>
<tr>
<td><strong>Neutron</strong></td>
<td>268 - 415 cps</td>
<td>353 cps</td>
</tr>
</tbody>
</table>

Table 5: The average value and range for density, gamma ray and the neutron well logs over marker 3 in figure 20.

This area has the lowest gamma ray response indicating that there is very little clay in the system.
however the gamma ray logs shows that there is an upward increase in the clay volume recorded over this depth interval. The density response shows that the unit has a low density throughout with an upward increase in clay interpreted with the gamma ray log. The neutron response indicates a high hydrogen response, interpreted to be linked to coal. In summary this area has a high proportion of coal decreasing upwards with increasing percentage of the rock matrix becoming ash-rich.

The final marker, marker 4, represents two mudstone packages with a possible coal stringer between the mudstone units. Table 6 below describes the values seen in marker 4.

<table>
<thead>
<tr>
<th>Tool</th>
<th>Range</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma Ray</td>
<td>155 - 275 °API</td>
<td>206 °API</td>
</tr>
<tr>
<td>Density</td>
<td>1.82 - 2.59 g/cm³</td>
<td>2.51 g/cm³</td>
</tr>
<tr>
<td>Neutron</td>
<td>343 - 778 cps</td>
<td>623 cps</td>
</tr>
</tbody>
</table>

This depth interval shows a relatively constant and high gamma ray, which indicates that throughout the zone there is a high volume percentage of clay. Within the two higher density units a high neutron count is observed. These units are likely to contain 70 to 90% clay, indicated by the gamma ray index, and no coal due to the low hydrogen index. The high density index shows that there might be some quartz in the system, but further analysis is required to verify this. In the centre of the two mudstone units there lies a coal stringer of approximately 60 cm thick. This resolution would not be possible if using tools that average readings over a distance such as resistivity or sonic tools.

### 6.3. Identifying coal bearing rock

In order to calculate the components of the rock matrix a definition between coal bearing and non-coal bearing rocks must be determined. Using an
example from Well C6 (Figure 21) coal bearing units occur when the hydrogen index is greater than the quartz index (density index). This can also be defined as the crossover between the hydrogen and quartz indices (column 4 on figure 21).

The relationship, and its importance, between density and neutron has been mentioned before. Not only is this relationship used to distinguish between coal and non-coal bearing rock units, the separation between the density index and the quartz index (marker 4 on figure 20) is used to calculate the volume of coal in the rock matrix. This is discussed further in chapter 6.5.

6.4. Coal Analysis - PCA

There are multiple dependent and independent variables acting on the calculation of the rock matrix. In order to establish a relationship between the dependent variables a principle component study is conducted. The principle component analysis measures the relationship between the logging tools and the measured lab proximate analysis results. The results of this act to guide in the establishment of multiple equations used to calculate coal bearing rock matrix components.

6.4.1. The Method

There are three steps in the method. The first is to establish a descriptive relationship between all the variables. The second step is to extract the data to have a comparison between the lab proximate analysis and the wireline logs. The final step is the principle component analysis to empirically define the relationship between all variables.

Table 7 is a summary of the descriptive statistics for the test data used for the principle component analysis.

<table>
<thead>
<tr>
<th>Moist</th>
<th>Ash</th>
<th>VM</th>
<th>FC</th>
<th>VM+FC Hyd</th>
<th>GRI</th>
<th>Qtz</th>
<th>Hyd</th>
<th>GRI</th>
<th>Qtz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>3.35</td>
<td>28.67</td>
<td>24.76</td>
<td>43.22</td>
<td>67.98</td>
<td>66.43</td>
<td>26.27</td>
<td>28.15</td>
<td></td>
</tr>
<tr>
<td>Std dev</td>
<td>1.54</td>
<td>15.59</td>
<td>8.20</td>
<td>12.32</td>
<td>15.51</td>
<td>10.47</td>
<td>16.09</td>
<td>11.87</td>
<td></td>
</tr>
<tr>
<td>Skewness</td>
<td>0.46</td>
<td>1.97</td>
<td>-0.90</td>
<td>-0.71</td>
<td>-1.92</td>
<td>-3.04</td>
<td>1.23</td>
<td>1.59</td>
<td></td>
</tr>
<tr>
<td>Kurtosis</td>
<td>-0.68</td>
<td>5.07</td>
<td>0.00</td>
<td>2.12</td>
<td>4.90</td>
<td>14.20</td>
<td>1.45</td>
<td>4.76</td>
<td></td>
</tr>
</tbody>
</table>

Table 7: Descriptive statistics of lab measured proximate analysis and indexed well logs.
Figure 22: The example extraction interval for Well C6, showing the relationship between the density, gamma ray and hydrogen index against the lab proximate analysis results.

Sample 62994-D008, or A, has a lower percentage of ash and a higher proportion of %volatile matter and %fixed carbon compared to Sample 62994-D009, or B.

<table>
<thead>
<tr>
<th>No</th>
<th>Sample No</th>
<th>Moist</th>
<th>Ash</th>
<th>VM</th>
<th>FC</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>62994-D008</td>
<td>5.27</td>
<td>16.07</td>
<td>34.62</td>
<td>44.04</td>
</tr>
<tr>
<td>B</td>
<td>62994-D009</td>
<td>3.89</td>
<td>23.76</td>
<td>33.72</td>
<td>38.63</td>
</tr>
</tbody>
</table>

Table 8: Results of the lab proximate analysis (%) from Well C6 over the depth interval shown in figure 22.

This is then related to the average indices values where the hydrogen index is higher, and the quartz index is lower in A. Based on earlier assumptions the gamma ray index shows that Sample A has a larger proportion of clay in the matrix than Sample B.

Table 9: The extracted average hydrogen, gamma ray and quartz index over the same sample depth interval shown in figure 22.

The most direct correlation is between the sum of %fixed carbon and %volatile matter, referred to as the coal proportion of the rock matrix, and the hydrogen index. Sample A has an average hydrogen index of 76.28% and the proximate analysis coal proportion is 78.66%. Sample B shows a similar correlation with the average hydrogen index of 70.73% and the proximate analysis coal proportion of 72.35%. As this is based on only two samples, the complete principle component analysis is needed for accurate results.

The principle component analysis results in the reduction of the number of dimensions, in this case 7 dimensions, into a number of dimensions that can be interpreted. The relationship between each of the variables is established empirically, these relationships are then used to guide the pseudo proximate calculations.

6.4.2. The Results

The principle component analysis reduces the dimensions from 7 to 2, while still accounting for 80.43% of the variance (Figure 23).

Figure 23: Scree plot showing total variance of 80.43% accounted for by a 2-factor model.

The relationship between each of the variables can be displayed in 2 dimensions, with minimal loss of
data. Figure 24, the normalized PCA biplot, visually shows the relationship between each variable.

Figure 24: Normalized PCA biplot showing the relationship of each variable based on the first and second principle components. The sample data are manipulated to show the distribution in the same plane.

From the biplot two variables stand out, namely %moisture and %fixed carbon, both these variables do not seem to have a correlation to the other variables. %Moisture can be explained by the nature of the data, it has a mean of 3.35% and a standard deviation of 1.54. It is unlikely that measuring the relationship between %moisture and the remaining variables will be achieved with the high variance in the data. The %fixed carbon can be explained by the fact that none of the tools are a direct measurement of carbon. There is a first component relationship between %fixed carbon and %volatile matter, and an inverse relationship between %fixed carbon and %ash. However, the relationships are poor indicating that it is unlikely to achieve a good correlation. This is shown in the reproduced correlation table (Table 10).

Table 10: The reproduced correlation matrix using 2 dimensions. The green text cells indicating the best correlation between wireline tools and the lab proximate analysis. The red text cells show the coefficient of determination for the lab proximate analysis, whereas the blue text shows the same for the wireline tools.

Table 10 empirically shows that in two dimensions there is a poor correlation coefficient when trying to calculate %moisture as well as %fixed carbon. Ash volume percent correlates well to the quartz index, and %volatile matter correlates well with the hydrogen index. The gamma ray index has no good correlation, which further highlights its independence.

6.4.3. Conclusions

The goal of this section is to quantify the relationship between the wireline logs and the proximate analysis measurements. These observations can then be applied to better understand and explain the calculations proposed of the pseudo proximate analysis. The principal component analysis hints towards the degree of error to be expected, based on the nature of the data.

The established relationship, listed in table 11, describes the interpretation of the principle component analysis results.

Table 11: Descriptive correlation between wireline logs and proximate analysis measurements. Strong positive relationship between %volatile matter and hydrogen index, as well as between %ash and the quartz index.

The gamma ray has no strong correlation with any lab measured proximate analysis component. This is due to the fact that the lab measured proximate
analysis does not measure clay in the rock matrix. The method used to conduct the lab proximate analysis measurements will distort the volume of clay.

The hydrogen index measures the hydrogen concentration in the rock matrix, therefore the fact that it has a positive correlation with both the %volatile matter and %moisture is no surprise. The negative correlation with %ash indicates that as the ash proportion of the rock matrix increases, the coal proportion (and hydrogen) decreases.

The quartz index measures the density of the rock matrix. A high ash content is likely to have a higher density. As the coal proportion of the matrix increases the density of the bulk rock decreases, hence the weak negative relationship with %volatile matter and %fixed carbon.

The %coal volume has a better inverse correlation with the quartz index than it has a positive correlation with the hydrogen index (Table 12). This then shows that the example from Well C6, sample A and B, is not necessarily the standard.

### 6.5. Coal Analysis – Pseudo Proximate

During the normalization process, the logging curves are converted into curves that represent components of the rock matrix. The results are three high resolution curves showing hydrogen concentration, density variations and volume of clay. This next section will discuss the method used to calculate a pseudo proximate analysis from the three normalized curves.

First, the components of the rock matrix need to be defined. The rock matrix consists of four components (Figure 25):

1. Volume Clay
2. Volume Fixed Carbon (FC)
3. Volume Quartz
4. Pore Volume (Water + Volatile Matter)

![Figure 25: Schematic representation of the rock matrix.](image)

This definition can be used within this area as there is no limestone and the dolerite that is present is outside the scope of this study and has no direct influence on the calculation of the components of the matrix.

With this a correlation between the lab proximate analysis and the pseudo proximate analysis can be made. From the correlation results the minimum and maximum values selected for the normalization can be calibrated to better match the lab analysis results. This is discussed in more detail in Chapter 7.

#### 6.5.1. Volume Clay

This is only based on the gamma ray response; the gamma ray is normalized to represent the %clay in the rock matrix. The assumptions that must be made
are that the composition of the clay minerals is uniform throughout the interpreted interval, and that there is little to no post-deposition chemical alteration acting to increase the natural radioactivity in non-clay bearing strata.

Figure 26: %Clay is calculated using the gamma ray index. The Remainder rock matrix (RRM) is the calculated as (1 - %Clay).

The remainder rock matrix (RRM) is then defined as:

\[ RRM = 1 - \%Clay \]

Equation 4: The formula used to determine the remainder rock matrix (RRM). This consists of the %Coal, %Quartz and %Pore volume in the rock matrix.

This value corresponds to the remaining rock volume available to be filled with the remaining components. Gamma ray is, therefore, a direct and absolute measure of the volume of clay minerals in the rock matrix and is shown on figure 26.

6.5.2. Volume Fixed Carbon

In order to calculate the volume of coal in the rock matrix a neutron-density relationship is used, in the form of the hydrogen and quartz indices.

This relationship not only measures the amount of coal in the rock matrix, but it is also used to define the coal versus non-coal bearing units (Figure 27).

In this method %coal is defined as the summation of %volatile matter and %fixed carbon. The reason being that the %volatile matter is a result of the organic material, as is the carbon in the rock matrix.

To calculate the percent fixed carbon in the coal bearing rock matrix the separation between the hydrogen index and quartz index is used (Figure 20, Marker 4). This results in a value that represents the proportion of fixed carbon in the remainder rock matrix. Equation 5 shows the formula to calculate the true percentage of coal in the matrix.

\[ \%FC = (HydInd - QtzInd) \times RRM \]

Equation 5: %Fixed carbon equation. This is a combination of three curves, gamma ray, neutron and density, to calculate the %Coal in the rock matrix.

Ultimately, the %fixed carbon is calculated using the three normalized curves. Where the gamma ray index defines the remainder rock matrix, and the neutron-density (hydrogen-quartz) relationship defines the proportion of coal.

To expand on Equation 5, the hydrogen index is a measure of coal, clay and water. Whereas, the quartz index is a measure of clay and water.
Equation 6: This equation shows that the subtraction of the quartz index from the hydrogen index is equivalent to the coal proportion minus the quartz proportion. If there is no quartz in the system this equation then measures the proportion of coal. This relationship is used to define coal versus non-coal bearing rock units.

It can be reasoned that the percentage of quartz in the coal bearing rock intervals is limited, and for the purpose of calculating the fixed carbon in the rock matrix it can be assumed that there is no quartz in the system for this deposition environment.

The hydrogen index is not a direct measure of the %fixed carbon and is influenced to a degree by the amount of clay and volatile matter in the matrix. The hydrogen-quartz index relationship is used as a measure of the proportion of coal in the matrix and therefore is multiplied with the remainder rock matrix (1 - %Clay). This acts to remove the effect of clay on the hydrogen index.

The calculation of %fixed carbon is only valid when there is organic matter, or coal, in the rock matrix. Therefore, this equation is only applied when the hydrogen index is greater than the quartz index (HydInd > QtzInd). This is the definition of the coal bearing versus non-coal bearing rock units. Where the hydrogen index is less than the quartz index (HydInd < QtzInd), the %fixed carbon is set to zero.

6.5.3. Quartz and Pore Volumes

To calculate the amount of quartz in the rock matrix, the quartz index is used. The remaining rock matrix still unaccounted for is the pore volume and the volume of quartz. This is exactly what the quartz index measures, and can therefore be directly applied to the remaining rock matrix (RRM).

\[
\text{RRM} = 1 - \%\text{Clay} - \%\text{FC}
\]

\[
\%\text{Quartz} = \text{RRM} \times \text{QtzInd}
\]

\[
\therefore \%\text{Pore} = 1 - \%\text{Clay} - \%\text{FC} - \%\text{Quartz}
\]

Equation 7: Formulae to calculate both the %Quartz and %Pore volume using the remainder rock matrix and the quartz index.
The %fixed carbon (FC), %Clay and %Quartz is calculated. From this the remaining rock matrix is equal to the pore volume.

Shown in figure 31, is the result from equation 6 where the %quartz and %pore volume in the matrix are calculated.

The methane component of %volatile matter is stored in the rock matrix as 1) adsorbed gas directly onto the micropores of the internal structure of the coal (Kim, 1977) or 2) is in solution with the water molecules (Duan, et al., 1992). The methane component is per mole the compound with the highest hydrogen concentration. Therefore, it is expected that the hydrogen index is affected by the addition of methane. The remaining rock matrix component, the pore volume, is interpreted as a combination of methane and water. In order to calculate the relative percentages of each component the hydrogen index is used, as it measurement range starting at pure water and will increase with the addition of methane.

\[ \% \text{Volatile Matter} = \% \text{Pore} \times \text{HydInd} \]

\[ \% H_2O = \% \text{Pore} \times (1 - \text{HydInd}) \]

Equation 8: Calculation of the components of the pore space using the hydrogen index. Resulting in %VM and %Water.

Methane and other hydrogen bearing gases are a large proportion of the volatile matter, when compared to non-combustible gases. The %volatile matter calculated is an estimation of the volatile matter in coal bearing rock matrix, the amount of volatile matter will be underestimated as there are non-hydrogen gaseous compounds present, these include but are not limited to carbon dioxide.

The methane component of %volatile matter is stored in the rock matrix as 1) adsorbed gas directly onto the micropores of the internal structure of the coal (Kim, 1977) or 2) is in solution with the water molecules (Duan, et al., 1992). The methane component is per mole the compound with the highest hydrogen concentration. Therefore, it is expected that the hydrogen index is affected by the addition of methane. The remaining rock matrix component, the pore volume, is interpreted as a combination of methane and water. In order to calculate the relative percentages of each component the hydrogen index is used, as it measurement range starting at pure water and will increase with the addition of methane.

\[ \% \text{Volatile Matter} = \% \text{Pore} \times \text{HydInd} \]

\[ \% H_2O = \% \text{Pore} \times (1 - \text{HydInd}) \]

Equation 8: Calculation of the components of the pore space using the hydrogen index. Resulting in %VM and %Water.

Methane and other hydrogen bearing gases are a large proportion of the volatile matter, when compared to non-combustible gases. The %volatile matter calculated is an estimation of the volatile matter in coal bearing rock matrix, the amount of volatile matter will be underestimated as there are non-hydrogen gaseous compounds present, these include but are not limited to carbon dioxide.

Figure 31: The final formulae to calculate the rock matrix

With all the components of the rock matrix calculated, and the results converted into measurements comparable to that of the proximate analysis, a correlation between the pseudo proximate analysis and the lab proximate analysis can be made.
6.6. Coal Analysis – Results

The average results, over the sample intervals, of the pseudo proximate analysis are shown for all samples in Appendix 1.

Well C6 (Figure 34) is used as an example well log of the pseudo proximate analysis calculation to describe the results. Column 2, the results column, shows the %Quartz separate to the %Clay in the rock matrix. This is an added benefit of the method, where it is possible to separate the amount of quartz and clay. However, the lab proximate analysis measures the proportion of %ash in the rock matrix. The ash proportion in the pseudo proximate analysis method is equivalent to the sum of the %Clay and %Quartz. When comparing the results to the lab proximate analysis the %Ash must be used, shown in column 3, the pseudo proximate column.

Column 4, the lithology column, represents an interpretation, completed by members of the Sasol exploration and drilling team, of the extracted core.
There is an accurate match of the depth location of the coal seams between the interpreted lithology and the results of the pseudo proximate analysis. This indicates that the pseudo proximate is able to successfully identify between non-coal and coal bearing lithologies, without the use of density cut off values.

It is important to note that there may be depth inconsistencies between the interpreted lithology, using core depth measurements, and the pseudo proximate analysis, using wireline depth measurements. An example of this is the two coal seams below 350 meters, the core seam location does not match the wireline seam location. The data is depth shifted to remove these inconsistencies for the comparison between the lab measured proximate analysis and the results of the pseudo proximate analysis.

Column 4 only shows the thickness of the coal seams not the quality or rank of the seam. Therefore, based on this dataset, it is not possible to determine the accuracy of the pseudo proximate analysis with regards to the matrix components (%fixed carbon, %volatile matter, %ash and %moisture).

Figure 35 shows the results of the pseudo proximate analysis against the results of the lab measured proximate analysis. The same two samples used for the example in the principle component analysis are selected to example the accuracy of the pseudo proximate analysis rock matrix calculations.

Table 13 below compared the averages of the pseudo proximate analysis (PPA) to the measured lab proximate analysis (Lab) results.

<table>
<thead>
<tr>
<th>Depth</th>
<th>Pseudo Proximate Analysis Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEPTH (M)</td>
<td>Proximate:FC</td>
</tr>
<tr>
<td></td>
<td>Proximate:VM</td>
</tr>
<tr>
<td></td>
<td>Proximate:Ash</td>
</tr>
<tr>
<td></td>
<td>Proximate:Moisture</td>
</tr>
<tr>
<td></td>
<td>Proximate:Coal - VM + FC</td>
</tr>
<tr>
<td></td>
<td>%FC</td>
</tr>
<tr>
<td></td>
<td>%VM</td>
</tr>
<tr>
<td></td>
<td>%Ash</td>
</tr>
<tr>
<td></td>
<td>%Moisture</td>
</tr>
</tbody>
</table>

Table 35: The example Well C6, Sample A and B, interval showing the results of the pseudo proximate analysis. There is a good match between the %Ash measurements and the %Coal (FC + VM) measurements. The %volatile matter is underestimated.

Figure 35: The example Well C6, Sample A and B, interval showing the results of the pseudo proximate analysis. There is a good match between the %Ash measurements and the %Coal (FC + VM) measurements. The %volatile matter is underestimated.

Based on the extracted averages, the %volatile matter is grossly underestimated in both samples, whereas the %moisture content is over estimated. This is likely due to a poor selection of minimum and maximum neutron values used to create the hydrogen index, this is discussed in more detail in the next chapter (Chapter 7), the sensitivity study. It has been mentioned that %volatile matter is likely to be underestimated due to the fact that only hydrogen bearing gaseous components are accounted for, the pseudo proximate analysis does not measure carbon monoxide or carbon dioxide.

In these two example the %fixed carbon is slightly overestimated, this relates again to the selection of
the neutron values used for the hydrogen index. The pseudo proximate analysis uses the hydrogen index as an indirect measurement of %fixed carbon. This is based on the theory that methane is adsorbed to the micropores in the coal structure (Kim, 1977), and that the chemical makeup of coal is hydrogen rich (Heredy & Wender, 1980).

The ash volume percent of the rock matrix in the pseudo proximate analysis is divided into two components, these being the clay and quartz volume percentages. The clay volume is calculated solely using the gamma ray log, whereas the quartz volume is a function of gamma ray and the density logs, as explained in figure 30 (Chapter 6.5.3). Based on the non-coal bearing sensitivity study (Chapter 7) the selection of the maximum density has the largest effect on the quartz volume. Table 14 below shows the separate clay and quartz extracted averages for sample A and B (Figure 35).

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lab</td>
<td>PPA</td>
<td>Lab</td>
</tr>
<tr>
<td>%Quartz</td>
<td>6.74</td>
<td>11.54</td>
</tr>
<tr>
<td>%Clay</td>
<td>16.71</td>
<td>14.36</td>
</tr>
<tr>
<td>%Ash</td>
<td>16.07</td>
<td>23.45</td>
</tr>
<tr>
<td></td>
<td>23.76</td>
<td>25.90</td>
</tr>
</tbody>
</table>

Table 14: The table shows the lab measured and pseudo proximate ash proportion over sample A and B. The pseudo proximate ash proportion is separated into the quartz and clay proportion.

Table 14 shows that sample A lab measured ash content, 16.07%, better matches the gamma ray index, or pseudo proximate analysis clay volume, 16.71%. Whereas, the sum of PPA quartz and clay, 25.90%, better matches the ash proportion, 23.76%, in sample B. This may indicate that sample A consists of more clay than quartz, and thus the lab measured %ash has a better relationship with the gamma ray reading. Sample B has a better relationship with the combination of clay and quartz volume, indicating that the density reading provides more accurate results. However, this is again only two samples from one well, a more complete statistical study is conducted after the sensitivity study (Chapter 7) to measure the error.
7. Sensitivity Study

A sensitivity study using tornado plots on the minimum and maximum (P02 and P98, respectively) values selected to normalize each of the three well logs highlights the tools that have the largest effect on the pseudo proximate results.

Two sets of data were analysed using tornado plots, the first being coal bearing rock (Figure 36) and the second being non-coal bearing rock (Figure 37). Preliminary results show that the value selected for the maximum neutron count (840 cps) is the most sensitive to change the results of the pseudo proximate analysis for coal bearing rock. While in non-coal bearing rock both the maximum value selected for quartz (2.65 g/cm$^3$) and the maximum gamma ray value (267.67 °API) have an equally large effect on the rock matrix component calculation of non-coal bearing rock units.

The tornado plots where ran using ranges based on the histogram to determine the possible error range for each of the rock matrix components. For the maximum value the range is based on the P97 and P99 values. The minimum range is based on the P01 and P03 values. These values are shown on table 15 below.

<table>
<thead>
<tr>
<th>Log Curve</th>
<th>Test Input</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gamma Ray</td>
<td>50</td>
</tr>
<tr>
<td>Density</td>
<td>1.27</td>
</tr>
<tr>
<td>Neutron</td>
<td>400</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Result</th>
<th>Value</th>
<th>Max</th>
<th>Min</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>%Coal</td>
<td>0.455</td>
<td>+0.0451</td>
<td>-0.0392</td>
<td>0.0844</td>
</tr>
<tr>
<td>%Ash</td>
<td>0.235</td>
<td>+0.0257</td>
<td>-0.0285</td>
<td>0.0542</td>
</tr>
<tr>
<td>%Quartz</td>
<td>0.050</td>
<td>+0.0102</td>
<td>-0.0098</td>
<td>0.0200</td>
</tr>
<tr>
<td>%Clay</td>
<td>0.185</td>
<td>+0.0183</td>
<td>-0.0207</td>
<td>0.0390</td>
</tr>
<tr>
<td>%Moisture</td>
<td>0.095</td>
<td>+0.0161</td>
<td>-0.0181</td>
<td>0.0342</td>
</tr>
<tr>
<td>%VM</td>
<td>0.216</td>
<td>+0.0123</td>
<td>-0.0134</td>
<td>0.0257</td>
</tr>
</tbody>
</table>

7.1. Coal Bearing Rock Matrix

The first is a coal bearing rock unit with the tool readings (test input) specified in table 16, below.

Table 16: Coal bearing sensitivity inputs have a low gamma ray of 50 °API, a density of 1.27 g/cm$^3$ and a neutron reading of 400 cps. After running the pseudo proximate analysis calculations the results are 45.5% Coal, 23.5% Ash, 9.5% Moisture and 21.6% Volatile Matter.

The test input data are the inputs used to calculate the rock matrix. A simple example of this would be to look at the %Clay calculated at 18.5% based on the gamma ray input of 50 °API. The equation to calculate %Clay follows:

\[
%\text{Clay} = \frac{GR_{\text{log}} - GR_{P02}}{GR_{P98} - GR_{P02}}
\]

\[
%\text{Clay} = \frac{50 - 9.875}{226.67 - 9.875} = 0.185
\]

All the test inputs are first normalized and then put through the pseudo proximate analysis to determine the rock matrix relative percentages. These results are then processed multiple times using different inputs for normalization based on the histograms. Therefore, to determine the change in %Clay with a change in the selection of the minimum gamma ray an equation is derived:

\[
%\text{Clay} = \frac{GR_{\text{log}} - GR_{P01}}{GR_{P98} - GR_{P01}}
\]

\[
%\text{Clay} = \frac{50 - 7.65}{226.67 - 7.65} = 0.1934
\]

| Table 15: Input parameters used to run the sensitivity analysis and output the tornado diagrams. The input values are based on the histogram values of the percentiles P02 and P98. The minimum value is selected from the histogram percentile at P01 and P03 for the gamma ray and neutron curves. Lastly, the maximum value is selected based on the P03 and P99 percentiles read from the histogram. The density minimum and maximum is determined based on the measurement error of the density tool, +/- 0.02 g/cm$^3$. |
In order to quantify the total effect of the minimum selected value this must be repeated using the P97 percentile:

\[
\%Clay = \frac{GR_{log} - GR_{P02}}{GR_{P97} - GR_{P02}}
\]

\[
.\%Clay = \frac{50 - 9.875}{215.57 - 9.875} = 0.1951
\]

Decreasing the minimum or maximum gamma ray setting results in an overall increase in %Clay. The difference from the standard input, using P98 and P02, is added to determine what the maximum change to %Clay might be if both P01 and P97 were to be selected.

<table>
<thead>
<tr>
<th>Test Input (P02 and P98) %Clay</th>
<th>Results with P97 and P01</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1851</td>
<td>0.1951</td>
<td>0.0100</td>
</tr>
<tr>
<td>0.1851</td>
<td>0.1934</td>
<td>0.0083</td>
</tr>
</tbody>
</table>

Total positive change possible to %Clay: + 0.0183

Table 17: The above table outlines the maximum change, +0.0183, to the %Clay based on changing the minimum gamma ray index inputs, P97 and P01.
This outlines the workflow used to determine the minimum, maximum and range for each pseudo proximate analysis component shown in Table 15. This is repeated for each stage of the pseudo proximate calculation. The final sensitivity results are shown visually as a tornado diagram (Figure 36).

From the tornado diagram it is seen that the greatest error is determined to be caused by two factors; 1) the selection of the maximum neutron count (affecting the maximum hydrogen index) and 2) the selection of the maximum gamma ray value (affecting the maximum volume clay).

With the results of the sensitivity study, the tornado plots, it is possible to measure the total error attached to the calculation of each rock matrix component. This error range is shown for coal bearing rocks in Table 16, and for non-coal bearing rock in Table 18. The error range only applies to the selected dataset and will change slightly with different input data. This analysis is used as an example to compare the sensitivity of the inputs using two end member datasets. Effectively the selected datasets describe the maximum error range possible for each rock matrix component.

7.2. Non-Coal Bearing Rock Matrix

The next set of tornado plots (Figure 37) is an example of a non-coal bearing rock matrix. The input data for this analysis (Table 18) is chosen due to the fact the gamma ray value is the same at 50 °API. Comparing the results shown in Table 16 and the Volume Clay tornado plot it is seen that the clay content is the same for coal bearing and non-coal bearing rock units. The maximum clay volume shown in Table 11 of 0.1951 and 0.1934 corresponds to the maximum increase in clay in the clay volume tornado plot (Figure 37).

![Figure 37: Tornado plots for non-coal bearing rock units based on the input test data from table 18. As there is no coal in the rock matrix the hydrogen index does not affect the results. Gamma ray is an absolute measurement of volume clay (by assumption), and therefore it is only the selection of the gamma ray minimum and maximum that will affect the volume of clay. The two most sensitive indices are the selection of the maximum density value for the quartz index and the maximum gamma ray for the clay index.](image-url)
Table 18: Non-coal bearing rock unit sensitivity study inputs with the same gamma ray of 50 °API, a higher density of 2.55 g/cm³ and a slightly higher neutron of 500 cps. The results show a non-coal bearing rock unit due to the high density with 18.5% Clay, 76.4% Quartz and 5.1% Moisture.

As this is a non-coal bearing rock unit, the hydrogen index has no effect on the outcome of the rock matrix components. It is seen that the selection of the maximum density value (2.63 or 2.67 g/cm³) and the selection of the maximum gamma ray value (215.57 or 246.8 °API) has the most influence on the calculation of the rock matrix components. It is also seen that there is a relationship between the %ash and %moisture, where the sum of the two components must equal 100% in a non-coal bearing rock matrix; this is always the result of the pseudo proximate analysis equations.

7.3. Summary

The end result of the sensitivity study proves that the minimum and maximum selected values to create the indices for gamma ray, neutron and density logs do considerably influence the final pseudo proximate analysis result. This is therefore one component that contributes to the error range of the method.

8. Multiple Regression Analysis

The pseudo proximate analysis is a logical approach to calculating the coal bearing matrix components, it is predictive in the sense that lab proximate data are not required as a calibration. With lab proximate analysis data another approach of calculating a continuous pseudo proximate analysis throughout the well log is a multiple regression analysis. This method uses the lab proximate data and the wireline logs to calculate the best fit linear equation. Therefore, the multiple regression analysis provides a method to test the accuracy of the pseudo proximate analysis, while taking into account the inherent error in the lab measured proximate analysis and the readings of the wireline logging tools.

8.1. Data Inputs

The data inputs mimic those used for the principle component analysis, except for the fact that the raw wireline readings are used and not the calculated indices. The inputs to calculate the best fit linear regression equation are the average extracted gamma ray, density and neutron values over the lab proximate sample depth interval.

The example from Well C6 is selected again (Figure 38), however with the multiple regression analysis the raw inputs are used and not the indices. The neutron count is inversely proportionate to the hydrogen concentration. Therefore, an increase in the hydrogen concentration will result in a decrease in the neutron count. The hydrogen index is calculated in order to transform the neutron count into a proportionate measure of the hydrogen concentration (Chapter 6.2.3).

The %moisture, %ash, %volatile matter and %fixed carbon displayed below (Table 19) are the results of the lab proximate analysis.

<table>
<thead>
<tr>
<th>Sample No</th>
<th>Moist</th>
<th>Ash</th>
<th>VM</th>
<th>FC</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 62994-D008</td>
<td>5.27</td>
<td>16.07</td>
<td>34.62</td>
<td>44.04</td>
</tr>
<tr>
<td>B 62994-D009</td>
<td>3.89</td>
<td>23.76</td>
<td>33.72</td>
<td>38.63</td>
</tr>
</tbody>
</table>

Table 19: Lab proximate analysis results for %moisture, %ash, %fixed carbon and %volatile matter over interval A and B for Well C6.

The extracted averages for gamma ray (°API), density (g/cm³) and neutron (cps) readings over the same interval are used as the inputs to calculate the multiple regression analysis (Table 20).

<table>
<thead>
<tr>
<th>Sample No</th>
<th>Neutron</th>
<th>Gamma</th>
<th>Density</th>
</tr>
</thead>
<tbody>
<tr>
<td>A 62994-D008</td>
<td>357.15</td>
<td>46.09</td>
<td>1.34</td>
</tr>
<tr>
<td>B 62994-D009</td>
<td>392.28</td>
<td>41.01</td>
<td>1.45</td>
</tr>
</tbody>
</table>
Table 20: The average for neutron (cps), gamma ray (°API) and density (g/cm³) over the sample depth A and B.

Figure 38 shows the results of the lab measured proximate analysis against the raw density, neutron and gamma ray wireline logs.

8.2. Results

The results for all 105 sample inputs are found in Appendix 1. The proximate analysis components are determined as a function of the wireline logs, however not all wireline logs have a significant effect on the results and can, therefore, be ignored in the equation. This is determined by the p-value of less than 0.05 (the selected alpha value). While this is true, the upper and lower 95% must be taken into account, in that if 0 falls between the upper and lower 95% range, the wireline log may be ignored.

8.2.1. Influencing Logs

Table 21 describes the results for the %ash multiple regression, it shows that both gamma ray readings and neutron count have negligible effect on the calculation of the ash volume percent in the rock matrix.

Table 21: Results of the multiple regression analysis for the %ash component using the density, neutron and gamma ray well logs as inputs.

%Moisture is more complicated, in the sense that all input wireline logs have some effect on calculating the volume of moisture in the rock matrix. However, because of the small data spread with %moisture the selection of the intercept will have the greatest effect of the calculation of moisture. Ultimately, the regression model (Table 22) does not provide an accurate estimation of moisture.

Table 22: The %moisture regression analysis results using the density, gamma ray and neutron logs. The intercept selection and density log has the greatest influence on the calculation of moisture.

%Volatile matter is best represented by neutron and gamma ray, with neutron having the greatest effect. Density can be ignored in the equation of %volatile matter, as both the p-value and upper and lower 95%
requirements are met. Table 23 shows the results of the regression analysis for %volatile matter.

<table>
<thead>
<tr>
<th>VM</th>
<th>Coefficient</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>56.70</td>
<td>9.52E-23</td>
<td>47.87</td>
<td>65.52</td>
</tr>
<tr>
<td>Density</td>
<td>5.20</td>
<td>0.2201</td>
<td>-3.16</td>
<td>13.55</td>
</tr>
<tr>
<td>Neutron</td>
<td>-0.08</td>
<td>4.7E-14</td>
<td>-0.10</td>
<td>-0.08</td>
</tr>
<tr>
<td>Gamma</td>
<td>-0.09</td>
<td>1.5E-05</td>
<td>-0.12</td>
<td>-0.05</td>
</tr>
</tbody>
</table>

Table 23: The %volatile matter regression analysis results based on all three wells logs. Based on the p-value the neutron log has the greatest effect on calculating %volatile matter.

%Fixed carbon has shown to be problematic when attempting correlations with the well logs. This remains true when using the multiple regression method. All wireline logs play a role in the calculation of %fixed carbon, with the greatest effect contributed by the density log. Table 24 shows the results of the %fixed carbon regression analysis.

<table>
<thead>
<tr>
<th>FC</th>
<th>Coefficient</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>106.93</td>
<td>1.1E-21</td>
<td>89.61</td>
<td>124.24</td>
</tr>
<tr>
<td>Density</td>
<td>-69.25</td>
<td>3.2E-13</td>
<td>-85.65</td>
<td>-52.86</td>
</tr>
<tr>
<td>Neutron</td>
<td>0.08</td>
<td>4E-05</td>
<td>0.04</td>
<td>0.11</td>
</tr>
<tr>
<td>Gamma</td>
<td>0.11</td>
<td>0.0301</td>
<td>0.04</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Table 24: The density, gamma ray and neutron logs all play a role in calculating the volume percent of fixed carbon. The neutron and gamma ray play a minor role, while the density has the greatest effect on the calculation of %fixed carbon.

Table 25 shows that the density log has the greatest impact when estimating the percentage coal in the matrix, the sum of %volatile matter and %fixed carbon.

<table>
<thead>
<tr>
<th>VM+FC</th>
<th>Coefficient</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>163.62</td>
<td>3.8E-31</td>
<td>144.38</td>
<td>182.86</td>
</tr>
<tr>
<td>Gamma</td>
<td>0.03</td>
<td>0.5189</td>
<td>-0.05</td>
<td>0.11</td>
</tr>
<tr>
<td>Density</td>
<td>-64.06</td>
<td>3.2E-10</td>
<td>-82.27</td>
<td>-45.84</td>
</tr>
<tr>
<td>Neutron</td>
<td>0.00</td>
<td>0.8767</td>
<td>-0.04</td>
<td>0.04</td>
</tr>
</tbody>
</table>

Table 25: The regression analysis for the coal proportion shows that the neutron and gamma ray logs may be ignored as they have a minor effect on the results. These results are similar, albeit inverse, to the regression analysis determined of the %ash volume, indicating that an inverse relationship between the volume %ash and %coal exists.

The correct logs to calculate a pseudo proximate analysis based on the results of the regression analysis are selected, the well logs selected to calculate each rock matrix component are shown in table 26.

<table>
<thead>
<tr>
<th>%Ash</th>
<th>Gamma</th>
<th>Density</th>
<th>Neutron</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>%Moisture</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>%VM</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>%FC</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>%VM + FC</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 26: The selected wells logs used to calculate the final linear regression equation for each of the rock matrix components.

To calculate the %moisture and %fixed carbon volumes using the multiple regression analysis method all the wells logs are selected. The %coal (VM+FC) and %ash proportion both use only the density well log in the regression equation. Lastly, to calculate the %volatile matter volume using the regression method requires the use of both the neutron log and the gamma ray log.

8.2.2. Multiple Regression Equations

In this section the necessary regressions are recalculated omitting the unnecessary wireline logs. The %moisture and %fixed carbon both result in an r-squared value of less than 50%, indicating that the multiple regression analysis provides a poor correlation.

The correct logs to calculate each proximate rock matrix component are shown from tables 27 to 31.

<table>
<thead>
<tr>
<th>Moist.</th>
<th>Coefficient</th>
<th>P-value</th>
<th>Lower 95%</th>
<th>Upper 95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interceptron</td>
<td>3.47</td>
<td>0.00469</td>
<td>1.09</td>
<td>5.86</td>
</tr>
<tr>
<td>Density</td>
<td>4.79</td>
<td>5.4E-05</td>
<td>2.54</td>
<td>7.05</td>
</tr>
<tr>
<td>Neutron</td>
<td>-0.01</td>
<td>5.9E-08</td>
<td>-0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>Gamma</td>
<td>-0.02</td>
<td>0.00076</td>
<td>-0.03</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

Table 27: The results of the %moisture regression calculation. Only 32% of the data points can be explained by variance alone.

%Fixed carbon has a strong negative relationship with density, and a weak positive relationships with
both the neutron count and gamma ray reading. The overall correlation is poor, with only 44% of data explained by variance.

\[
\text{Fixed Carbon} = (\text{Dens} \times -69.25) + (\text{Neu} \times 0.08) + (GR \times 0.11) + 106.93
\]

Table 28: The regression equation shown above for %fixed carbon, indicate that density had the greatest effect. However, modelling the %fixed carbon based on wireline logs results in a relatively poor correlation.

The calculation of %ash using wireline logs shows that density is the only contributing factor. This is expected as an increase in density is due to an increase in ash volume percent in the rock matrix. The proximate analysis does not distinguish between quartz and clay minerals and therefore gamma ray does not play a role; the reason being that both quartz and clay minerals have a much larger density than the components of coal.

\[
\text{Ash} = (\text{Dens} \times 62.50) - 65.10
\]

Table 29: The %ash regression analysis only uses density as an input. There is a positive strong correlation between density and %ash.

%Volatile matter has a strong correlation to the neutron readings, as proven by the principle component analysis. The correlation is shown to be negative here as a lower raw neutron count indicates a higher hydrogen percentage. Whereas, the hydrogen index, using the indexed neutron readings, measures an increase in hydrogen. From the principle component analysis there is a strong inverse correlation with gamma ray values, this is seen in the regression equation for %volatile matter (Table 30). The gamma ray log has an effect on the calculation of total %volatile matter. The multiple regression analysis works well with the calculation of %volatile matter, and results in an r-squared value of 82%.

\[
\text{Volatile Matter} = (\text{Neu} \times -0.07) + (GR \times -0.07) + 60.85
\]

Table 30: %Volatile matter shows the strongest correlation to the wireline logs, using neutron and gamma ray. Neutron plays the larger role in the regression analysis, this is expected as %volatile matter has the largest hydrogen concentration.

Using the sum of %fixed carbon and %volatile matter the modelling regression equation uses density as the only influencing input. This closely resembles the inverse of the modelled %ash regression equation. The results show a better correlation than %fixed carbon alone.

\[
\text{FC+VM} = (\text{Dens} \times -61.36) + 160.04
\]

Table 31: The sum of %fixed carbon and %volatile matter shows a better correlation compared to that of modelling %fixed carbon separately.

Using the example from Well C6, samples A and B (Figure 39), it is possible to calculate the proximate analysis components using the regression analysis. The table below (Table 32) shows the average calculated proximate analysis from the regression analysis (Regr) versus the lab measured proximate analysis (Lab) of the two sample intervals.

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lab</td>
<td>Regr</td>
</tr>
<tr>
<td>%Fixed Carbon</td>
<td>44.04</td>
<td>49.47</td>
</tr>
<tr>
<td>%Volatile Matter</td>
<td>34.62</td>
<td>30.93</td>
</tr>
<tr>
<td>%Ash</td>
<td>16.07</td>
<td>18.94</td>
</tr>
<tr>
<td>%Moisture</td>
<td>5.27</td>
<td>3.89</td>
</tr>
</tbody>
</table>

Table 32: Shows the comparison between the lab measured proximate analysis and the results of the regression analysis calculated over the average interval of the samples.
There is a similar, albeit less prominent pattern with results (Table 32) of the multiple regression analysis compared to the equations of the pseudo proximate analysis. The table shows that the %fixed carbon is overestimated and the %volatile matter is underestimated. The calculated %ash volume from the regression equation has a strong correlation with the lab proximate analysis. However, as with the pseudo proximate analysis the results are also overestimated.

The same trends are seen here compared to the trends seen in the pseudo proximate analysis, indicating that these trends are independent of the method used and related to the data itself. These results indicate that the relationship between the wireline logging tools and the lab measurements makes it unlikely, if not impossible, to get near perfect correlation result.

8.4. Conclusions

The multiple regression analysis proves to be an accurate method for calculating a pseudo proximate analysis provided that there is a large set of lab measured proximate analysis results on which to model the regression equations.

There are two main concerns when applying the regression equations to the wireline logs:

1. The regression equations are only valid for coal bearing units, and therefore a distinction between non-coal and coal bearing rock units must be applied before the application of the regression equations.

2. The sum of %moisture, %fixed carbon, %volatile matter and %ash volumes calculated using the regression analysis calculations do in most cases not equal 100% rock matrix. This occurs because the multiple regression equation calculates each rock matrix component separately.

The solutions to these problems are outside the scope of this thesis, as they do not affect the results required for comparison.

9. Correlations

The goal of the pseudo proximate analysis method is to calculate an accurate value match for each component of coal bearing rock. A perfect value match between the lab measured proximate analysis and the pseudo proximate analysis rock matrix components would have a regression line equation of \( y = x \), and a correlation coefficient of 1.
All four rock matrix components (%moisture, %ash, %volatile matter and %fixed carbon) of the pseudo proximate and the lab proximate analysis are plotted. Displayed on each plot is the linear regression line equation and the coefficient of determination, r-squared.

The volume %moisture cross plot (Figure 40) shows that there is a very poor negative correlation. The lab measured %moisture has a mean of 3.35 and a standard deviation of 1.54, whereas the pseudo calculated %moisture has a mean of 10.75 with a standard deviation of 5.72. This shows that overall the pseudo proximate analysis is overestimated.

There are large data spikes showing greater than 15% moisture and values up to 48% are observed. These are unrealistic in the clay to silty-mud lithologies, and are associated with very low %fixed carbon extracted averages of less than 10%. The data spike anomalies are caused by the resolution of the measuring tools, linked to the coal versus non-coal bearing detection method discussed in chapter 6.3.

The next cross plot for the proportion of ash in the matrix (Figure 41), shows that there is a positive correlation as expected. The correlation is good with most of poorly correlated data points towards the higher ash concentrations. As discussed above, an increased %ash implies a decreased %fixed carbon proportion, this again relates back to the method used to determine coal versus non-coal bearing rock matrix.

The %volatile matter cross plot (Figure 42) shows a clustering of data between approximately 20% and 25% of the volatile matter calculated with the pseudo proximate method. Within this zone the lab measured %volatile matter is slightly higher than the calculated pseudo proximate %volatile matter, which is expected as the pseudo proximate method measures only the hydrogen containing gasses and the remainder forms part of the moisture content.

There is a very good regression of almost 1 between the two datasets. The coefficient of determination for the pseudo proximate modelled %volatile matter is around 51%, this implies that 51% of the modelled pseudo proximate analysis data points can be explained by the lab measured proximate analysis
based on the variance of the data. The remaining 49% of data points cannot be explained using the variance and must be attributed to either unknown variables or the internal variability either or both of the datasets. This is the best result of the pseudo proximate model. Based on these data, the modelled pseudo proximate %volatile matter is slightly underestimated. This interpretation is based on the intercept of +3.86 and the fact that the gradient is marginally greater than 1.

The %fixed carbon cross plot (Figure 43) shows that there is almost no correlation between the pseudo and lab proximate analysis data. The lower the pseudo proximate percentage, the greater the lack of correlation. This indicates that there is a flaw in the method for calculating %fixed carbon, or the method is not calculating %fixed carbon and is estimating something else.

Figure 43: Pseudo versus lab %fixed carbon data points, showing the familiar poor correlation at lower %fixed carbon proportions. This rock matrix component has the poorest correlation and regression compared to all other components.

The following cross plot (Figure 44) shows that there is a positive correlation between the lab measured %volatile matter and modelled pseudo proximate %fixed carbon. This indicates that the calculations used to model the %fixed carbon correlate more efficiently with the lab measured %volatile matter. The reason for this result is that the equation for %fixed carbon incorporates both the density and the neutron logs, from the principle component analysis it can be deduced that both of these wireline logs are influenced by the %volatile matter. The conclusion is hinting in the direction that %fixed carbon is not directly measureable as a rock matrix component by using wireline logs.

Figure 44: Lab measured %volatile matter versus the pseudo proximate analysis calculated %fixed carbon.

Figure 45 is a cross plot showing the lab measured versus pseudo proximate calculated sum of %volatile matter and %fixed carbon, also referred to as the proportion of coal in the rock matrix. As this is a combination of both %fixed carbon and %volatile matter, the lower proportions show a similar data spread as seen in the %fixed carbon cross plot. Surprisingly, there is a relatively good correlation indicating that the pseudo proximate analysis is able to determine coal proportions in the rock matrix.

Figure 45: Proportion of coal (VM+FC) in the rock matrix, a cross plot between the lab measured coal proportion and the calculated coal proportion from the pseudo proximate analysis method. The lower proportions show a higher variance, associated again with the tool measurement resolution.

The assumption is made in this study that the volume of volatile matter and fixed carbon in the rock matrix are proportionate. This assumption can be made due
to the chemical composition of coal, where an increase in carbon results in the increase of combustible and non-combustible gases, which in turn results in an overall increase in the coal proportion. This then implies that the sum of the %volatile matter and %fixed carbon rock matrix components will result in a better estimation of the coal proportion in the rock matrix.

However, figure 46 shows the lab measured %volatile matter concentration versus the lab measured %fixed carbon concentration; there is no correlation between the two measurements.

The reasons for this lack in correlation are complex, and most likely due to the measurement procedure conducted by the lab. %Volatile matter and %fixed carbon must be associated with both the organic matter and coal in the rock matrix, but figure 46 suggests otherwise.

The pseudo proximate analysis cross plot (Figure 47) however of the same two components as in figure 46 i.e. the %volatile matter versus the %fixed carbon, shows that there is a measureable correlation between the two components.

This indicates that pseudo proximate calculated %volatile matter volume and %fixed carbon volume have a very good logarithmic relationship, this logarithmic relationship forms when a small proportion of carbon introduced to the rock matrix causes a much larger increase in the concentration of volatile matter in the rock matrix. The results shown here (Figure 47) are very different to the results of the lab measured %volatile matter versus the %fixed carbon. The strong correlation between the %volatile matter and %fixed carbon volumes is what one would expect to find, as both these components are associated with coal seams and a proportionate increase in both seems logical. However, based on the lab measured results there is no relationship between the rank of the coal and the amount of volatile matter in the rock matrix. The relationship between %fixed carbon and %volatile matter is very complex and depends on the burial history, biogenic alteration of the gases, and coal rank. Due to this only general statements concerning the relationship between the concentrations of volatile matter and fixed carbon in the rock matrix should be made (Rice, 1993).

**Conclusions**
The cross plots highlight a few important problems and successes of the pseudo proximate analysis method. The first is that the %moisture volume in the rock matrix is not easily calculated and perhaps should be ignored when trying to determine an
estimation of the %moisture volume. If not ignored then it should not be made a priority.

The method used to determine the coal versus non-coal bearing rock matrix is the hydrogen-quartz index cross over, and is only applicable to thicker coal intervals. Due to the tool resolution this method will cause the %fixed carbon to be greatly underestimated in thinner coal units.

The %ash volume of the rock matrix is calculated well, with the results almost mimicking those of the multiple regression analysis. Only in thinner coal units with lower measured %fixed carbon averages, does the %ash calculation perform poorly, this is due to the tool resolution as explained earlier. Using this pseudo proximate analysis to calculate ash content in the matrix has the added benefit that it provides a measure of quartz and clay as well. The accuracy of this must still be tested which is not possible with the available data in this project.

Lastly, and most importantly, the relationship between %volatile matter and %fixed carbon is nonexistent according to the lab measured results. The principal component analysis showed that %fixed carbon has no positive relationship to any other component or logging tool, and only a negative relationship with the percentage ash. Logical reasoning dictates that there will be a relationship between the %fixed carbon and %volatile matter at a chemical level. The apparent lack of this relationship may be explained by three scenarios:

1. It is a product of the lab measurement procedure, where the process leads to %fixed carbon or %volatile matter being incorrectly measured.
2. There is no relationship between the %fixed carbon and %volatile matter. This may be caused by chemical differences in the coal minerals, and can imply that some coal units will produce more volatile matter than others regardless of the carbon content.
3. The majority of the %volatile matter is not located within the coal structures and is found in the surrounding non-coal bearing rock matrix. This could be a result of the drilling process, causing the gaseous material, volatile matter, to become mobile and disperse evenly throughout the drilled zone.

However, more thorough studies as outlined by Rice (1993), dictates that the relationship between the %fixed carbon volume and the gaseous component is very difficult understand as it is dependent on multiple external influencing factors.

There is a weak negative correlation between the proportion of ash and volatile matter, and a weak positive correlation between the proportion of moisture and volatile matter in the rock matrix. Further analysis on this is required to fully understand the reasons behind the lack of a relationship, however this topic falls within the field of reservoir engineering and is outside the scope of this thesis.

10. Lab Proximate Analysis Error

Thus far we have assumed that the lab proximate analysis data is true and without its own error, which is unlikely to be the case. One measure of the error is the internal variance which can be described using the internal variance of the lab proximate analysis data. This variance is determined with the principle component analysis, where the reduction into a 2-dimensional space shows the percentage of data points that can be described by variance alone in a 2-dimensional space (or linear relationship), this is also referred to as the coefficient of determination (Table 33).

<table>
<thead>
<tr>
<th>Moisture</th>
<th>Ash</th>
<th>VM</th>
<th>FC</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.77</td>
<td>-0.02</td>
<td>0.60</td>
<td>-0.47</td>
</tr>
<tr>
<td>-0.02</td>
<td>0.91</td>
<td>-0.61</td>
<td>-0.75</td>
</tr>
<tr>
<td>0.60</td>
<td>-0.61</td>
<td>0.85</td>
<td>0.13</td>
</tr>
<tr>
<td>-0.47</td>
<td>-0.75</td>
<td>0.13</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Table 33: The best possible coefficient of determination described in 2 dimensions as determined by the principle component analysis.
What this implies is that due to the nature of the measured lab data it is impossible based on variance alone to describe the data better than the principle component analysis does in a two dimensional space. The principle component analysis of the lab measured proximate analysis provides a benchmark that describes the maximum achievable coefficient of determination in a 2-dimensional space, or linear relationship. Essentially, this benchmark describes the internal variance of the lab measured data and can be compared to the pseudo proximate analysis results without the influence of the internal variance. This comparison is discussed in Chapter 13.

There are no controls available to reproduce an estimate of the lab measured error, thus the measurements are assumed accurate.

11. Comparison

The inherent inaccuracy or variability with the lab proximate analysis measurements as well as the logging tools provides problems when trying to establish the validity of the pseudo proximate analysis calculations.

The multiple regression analysis calculates the best possible empirical linear relationship between the wireline logs and the lab proximate analysis data, while including the inherent error in the lab measurements and the logging tools. Therefore, the multiple regression analysis provides a dataset to which the pseudo proximate analysis can be compared.

However, the multiple regression analysis is a posteriori knowledge that incorporates the lab measured results to achieve an empirical relationship. Whereas, the pseudo proximate method is a priori knowledge, requiring only the log tool readings and not the proximate analysis results. As the multiple regression analysis is modelled using the results of the lab proximate analysis one would expect the multiple regression analysis to outperform the predictive pseudo proximate analysis in all regards.

There are multiple methods to statistically test the validity of the pseudo proximate analysis. Each method will test the results of the multiple regression analysis and the pseudo proximate analysis against the lab measured proximate results. The ultimate goal is to determine whether the pseudo proximate analysis is statistically significant compared to the measured lab proximate analysis. This is done using the multiple regression analysis to measure the best case scenario, taking into account the inherent variance of the lab measured results.

First, the t-test ('Student', 1908) will determine if there is a statistical significant relationship between datasets, but does not provide a quantitative measure of the differences between the data sets. In order to quantify the differences, two effect sizes are calculated: the first measures the correlation effect size and the second measures the difference from the mean effect size. The results of the three tests are used to determine the statistical validity of the pseudo proximate analysis method.

11.1. T-test (significance)

The t-test tests both the pseudo proximate and the multiple regression analysis against the lab measured proximate analysis. The t-test assesses whether the means of two datasets are statistically different from each other. Essentially, this method tests whether the two datasets are statistically different.

Table 34 displays the results of the t-test. The critical t-value calculated at an alpha of 0.05 is 1.97, all t-stat values below this indicate that 95% of the time there is statistically no difference. The t-test shows that the multiple regression analysis is statistically similar to the lab measured proximate analysis. The pseudo proximate analysis is however not statistically similar to the lab proximate analysis.
This proves that, 95% of the time (α = 0.05), the difference between the pseudo proximate analysis and the lab measured proximate analysis is statistically significant.

However, the t-test does not provide a quantitative measure of the difference between the two datasets. Therefore, we look at two more tests the first being Pearson’s r, or correlation coefficient (Pearson, 1901), to test the effect size based on explained variance. The second test is Cohen’s d (Cohen, 1988) to test the effect size based on difference from the mean.

### 11.2. Correlation Coefficient (r)

The correlation coefficient, Pearson’s r, describes the linear relationship between two samples: between the multiple regression and pseudo proximate analysis against the lab proximate analysis.

Using the coefficient of determination, r-squared, a quantitative measure of the effect size based on variance is measured. These values are shown in the table below (Table 35).

<table>
<thead>
<tr>
<th></th>
<th>MRA</th>
<th>PPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moisture</td>
<td>2.6E-15</td>
<td>12.81</td>
</tr>
<tr>
<td>Ash</td>
<td>1.9E-15</td>
<td>4.19</td>
</tr>
<tr>
<td>VM</td>
<td>6.9E-15</td>
<td>5.71</td>
</tr>
<tr>
<td>FC</td>
<td>3.0E-14</td>
<td>6.13</td>
</tr>
<tr>
<td>VM+FC</td>
<td>1.5E-14</td>
<td>6.97</td>
</tr>
</tbody>
</table>

Table 34: The results of the t-test, showing that the components of the pseudo proximate analysis (PPA) are all statistically different to the lab proximate analysis. The multiple regression analysis (MRA) components are not statistically different. The t-crit value is 1.97.

The results show that there is a linear correlation between the two datasets where the %ash, %volatile matter and %coal (VM+FC) show the highest r-squared. The anomaly lies with the %fixed carbon, where the pseudo proximate analysis fails to correlate to the lab proximate analysis. Volume %moisture proves to poorly correlate regardless of the method used.

With the Fisher transformed Pearson’s correlation coefficient (Fisher, 1915), it is possible to empirically measure the effect size of the correlation difference using Cohen’s q value. This allows for a statistical measure between the correlation coefficients, to determine whether there is a measureable difference based on the sample size used (Cohen, 1988).

### 11.2.1. Correlation Coefficient (r)

The correlation coefficient, Pearson’s r, describes the linear relationship between two samples: between the multiple regression and pseudo proximate analysis against the lab proximate analysis.

Using the coefficient of determination, r-squared, a quantitative measure of the effect size based on variance is measured. These values are shown in the table below (Table 35).

<table>
<thead>
<tr>
<th></th>
<th>MRA</th>
<th>PPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>%Moisture</td>
<td>0.32</td>
<td>0.11</td>
</tr>
<tr>
<td>%Ash</td>
<td>0.58</td>
<td>0.45</td>
</tr>
<tr>
<td>%VM</td>
<td>0.67</td>
<td>0.51</td>
</tr>
<tr>
<td>%FC</td>
<td>0.44</td>
<td>0.06</td>
</tr>
<tr>
<td>%VM+FC</td>
<td>0.56</td>
<td>0.42</td>
</tr>
</tbody>
</table>

Table 35: The r-squared values of both the multiple regression analysis (MRA) and the pseudo proximate analysis (PPA)

Based on Cohen’s description of the correlation difference effect size, Cohen’s q, the %ash and %coal (VM+FC) show a small to medium difference. %Moisture and %volatile matter show a medium difference in correlation, whereas the %fixed carbon shows a large difference. This implies that calculating %fixed carbon using the pseudo proximate analysis method does not correlate to lab measured proximate analysis, even when the variance is accounted for when using the multiple regression analysis.
To conclude this section, the observations made using variance based effect sizes indicate that

- %moisture does not correlate well regardless of the method used. This is due to the nature of the data measurement: limited range and a high standard deviation.
- The %fixed carbon calculation using the pseudo proximate analysis method is flawed and requires review. However,
- the pseudo proximate coal calculation (VM+FC) results in a small difference compared to the multiple regression analysis, the same is true for the %ash calculation procedure. Therefore,
- the pseudo proximate analysis method accurately calculates the proportion of ash versus coal in the rock matrix.

11.3. Cohen’s d

The next step is to test the effect size based on difference between the means of the data sets. This effect size is unlike the variance effect size, in the sense that it focuses on comparing the actual values rather than comparing the correlation. The multiple regression model will not be tested here, as the principles of the multiple regression (Pearson, 1896) result in a sum of residuals equal to zero and a mean equal to the mean of the control group. Therefore, Cohen’s d value will be equal to zero for all multiple regression calculations.

In essence, Cohen’s d will determine whether a sample is over- or under-estimated based on the mean and variance. From this it is possible to determine the proportion of data (from the pseudo proximate analysis) above the mean of the control group (lab measured proximate analysis), the percentage overlap between the data, and the probability of superiority which is the probability that a random sample picked will be larger than the control group. This test is done for both the multiple regression and pseudo proximate analysis, the results are then compared.

The first step is to analyse the control group, or lab measured proximate analysis, to determine the mean and variance for each component. Table 37 displays these results for the lab proximate analysis.

<table>
<thead>
<tr>
<th>Component</th>
<th>Mean</th>
<th>Variance</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moisture</td>
<td>3.35</td>
<td>2.36</td>
<td>1.54</td>
</tr>
<tr>
<td>Ash</td>
<td>28.67</td>
<td>243.17</td>
<td>15.59</td>
</tr>
<tr>
<td>VM</td>
<td>24.76</td>
<td>67.16</td>
<td>8.20</td>
</tr>
<tr>
<td>FC</td>
<td>43.22</td>
<td>151.74</td>
<td>12.32</td>
</tr>
<tr>
<td>VM+FC</td>
<td>67.98</td>
<td>240.66</td>
<td>15.51</td>
</tr>
</tbody>
</table>

Table 37: The results of the lab measured proximate analysis showing mean, variance and standard deviation.

This can be displayed visually as a frequency cross plot, with the category as the percentage rock matrix against the count or frequency (Figure 48).

![Figure 48: Lab proximate analysis frequency cross plot, showing the category percentage versus the count. This is the template dataset using to test both the results of the multiple regression and pseudo proximate analysis.](image)

A successful pseudo proximate analysis will show a similar pattern to figure 48. The degree of success is determined by Cohen’s d and associated interpretations. Table 38 shows results of the comparison to table 37 for the pseudo proximate analysis.

<table>
<thead>
<tr>
<th>Component</th>
<th>Mean</th>
<th>Variance</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Moisture</td>
<td>10.75</td>
<td>32.64</td>
<td>5.71</td>
</tr>
<tr>
<td>Ash</td>
<td>38.78</td>
<td>367.15</td>
<td>19.16</td>
</tr>
<tr>
<td>VM</td>
<td>19.29</td>
<td>29.01</td>
<td>5.39</td>
</tr>
<tr>
<td>FC</td>
<td>31.18</td>
<td>253.44</td>
<td>15.92</td>
</tr>
<tr>
<td>VM+FC</td>
<td>50.47</td>
<td>420.87</td>
<td>20.52</td>
</tr>
</tbody>
</table>

Table 38: The mean, variance and standard deviation for the different components of the pseudo proximate analysis.

As with the lab proximate analysis the data can be displayed as a frequency cross plot (Figure 49).
Initial interpretations show that there are inconsistencies between the %moisture and %volatile matter distributions. The %ash and %fixed carbon distribution for the pseudo proximate analysis is skewed towards the lower percentages compared to that of the lab measured proximate analysis %ash and %fixed carbon distribution.

The Cohen’s $d$ values, shown in table 39, quantify these anomalies. However, it does not provide all the necessary information therefore the Cohen’s $U_3$, percentage overlap (%Overlap) and probability of superiority (P.O.S) are also tabulated below (Cohen, 1988).

<table>
<thead>
<tr>
<th>Source</th>
<th>Cohen $d$</th>
<th>Cohen $U_3$</th>
<th>%Overlap</th>
<th>P.O.S</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example</td>
<td>0</td>
<td>50.0%</td>
<td>100.0%</td>
<td>50.0%</td>
</tr>
<tr>
<td>Moisture</td>
<td>1.77</td>
<td>96.2%</td>
<td>37.6%</td>
<td>89.5%</td>
</tr>
<tr>
<td>Ash</td>
<td>0.58</td>
<td>71.9%</td>
<td>77.2%</td>
<td>65.9%</td>
</tr>
<tr>
<td>VM</td>
<td>0.79</td>
<td>21.5%</td>
<td>69.3%</td>
<td>28.8%</td>
</tr>
<tr>
<td>FC</td>
<td>0.85</td>
<td>19.8%</td>
<td>67.1%</td>
<td>27.4%</td>
</tr>
<tr>
<td>VM+FC</td>
<td>0.96</td>
<td>16.9%</td>
<td>63.1%</td>
<td>24.9%</td>
</tr>
</tbody>
</table>

Table 39: The results of the Cohen analysis. The example row shows an equal data distribution similar to that of all the components of the multiple regression analysis.

From this it is interpreted that %moisture is greatly overestimated with 96.2% of samples being above the mean of the lab measured proximate analysis. There is an overlap of 37.6% indicating that most of the data points do not share a common range. Lastly 89.5% of the time, if chosen randomly, the %moisture of the pseudo proximate analysis will be higher than the lab measured %moisture.

The pseudo proximate analysis results for the %volatile matter and %fixed carbon are both underestimated as rock matrix components when compared to the lab measured proximate analysis. The %volatile matter results show that 78.5% of samples are below the mean of the lab %volatile matter. There is a good overlap of 69.3% between the lab and pseudo proximate analysis groups. There is a 28.8% probability that a randomly selected pseudo proximate data point will be larger than the mean of the lab measured %volatile matter. With 105 data points this indicates that only 30 data points lie above the mean of the lab measured result.

The volume percent of fixed carbon is underestimated to a higher degree compared to that of volatile matter. There is a difference of 12.08% between the lab measured and pseudo proximate method calculated %fixed carbon means. Furthermore there is a larger standard deviation in the pseudo proximate analysis compounding the degree of underestimation.

With both the %fixed carbon and %volatile matter rock matrix components underestimated it stands to reason that the proportion of coal (VM+FC) is also underestimated to a greater degree than either the
%volatile matter or %fixed carbon. This proves true, with only 16.9% of sample being greater than the mean of the lab measured %coal volume.

11.4. Conclusion
Firstly, the t-test proves that there is a statistical difference between the lab measured proximate analysis and the pseudo proximate analysis results, this is true for all components of the rock matrix. The degree and reason for the statistical difference are not measured using the t-test.

The correlation effect size, Pearson’s r, shows that the calculations for %moisture do not correlate with either the pseudo proximate analysis or the multiple regression analysis. This is due to the nature of the data, with a small range and high variance.

Following this, the correlation effect size for %fixed carbon is the second worst performer for the multiple regression analysis. However, with an r-squared value of 44%, a relationship between the multiple regression analysis and lab measured proximate analysis does exist. This is not true for the pseudo proximate calculated %fixed carbon, as the r-squared value is 6%, indicating that there is a fundamental error in the pseudo proximate method used to calculate %fixed carbon.

The standardized normal distribution correlation coefficient, or Fishers’ transformed Pearson’s r, allows for a direct and normalized comparison between correlation coefficients with different data distributions. Cohen’s q provides a measure of how accurately, compared to the multiple regression analysis, the pseudo proximate analysis performs. It shows that the pseudo proximate calculated %ash correlation performs better than the %moisture, %fixed carbon and %volatile matter pseudo proximate calculated results. Lastly, it shows that %moisture, %volatile matter and %fixed carbon have similar results. These results show that the pseudo proximate analysis method can be improved, starting with the method to calculate %fixed carbon.

Lastly, the effect size of the statistical difference based on the mean of the data was analysed to determine the actual difference in values and whether a component is over- or under-estimated. The findings show that the %coal volume is underestimated, while the ash and moisture proportions of the rock matrix are both overestimated. Taking this into account the procedure to modify this would be to modify the minimum and maximum values selected to calculate the quartz, clay and hydrogen indices. To achieve more accurate results, a simultaneous increase the hydrogen index and a decrease in the quartz and clay (ash) index is required. This is further discussed in chapter 12, where the necessary modifications to the pseudo proximate analysis are discussed and reasoned.

12. Recommendations
The pseudo proximate analysis proposed does not statistically match the data obtained from the lab measured proximate analysis. It has been demonstrated that the pseudo proximate analysis fails to correctly calculate the rock matrix components in a coal bearing unit. However, the approach used gives a broader understanding of the relationship between the three wireline logs for neutron, density and gamma ray values, to coal bearing rock matrix components: %ash, %moisture, %volatile matter and %fixed carbon.

12.1. Neutron Tool
The neutron tool, in particular, proves to be a very accurate measure of the %volatile matter based on the principle component analysis. Using the tool in its raw form, or not calibrated to limestone porosity units, allows for an accurate estimate of the proportion of hydrogen in the rock matrix. Where the method can be improved upon is to quantitatively determine the relationship between the neutron tool response and varying degrees of hydrogen concentration in a rock matrix. This will have to be done using samples where the hydrogen concentration is known for varying coal and non-coal
bearing samples, similar to the calibration used for limestone porosity units.

Furthermore, based on the analysis of the data, the relationship between the neutron reading, or the hydrogen index, and %volatile matter is not linear but rather an exponential relationship exists. The neutron tool is sensitive to a change in hydrogen, as a large proportion of %volatile matter consists of combustible hydrogen-bearing gases. It is expected that the addition of these gases would cause a large change in the neutron reading. Figure 50 below shows that a small proportion of %volatile matter has a large effect on the neutron reading, confirming that an exponential relationship exists.

Figure 50: Lab measured %volatile matter against the hydrogen index. The cross plot shows a good exponential correlation, with an r-squared value of 67% at the y-intercept set to 0.

This shows that there is a direct exponential relationship between the neutron tool and the %volatile matter. This implies that with a hydrogen index between 70% and 80% the resulting volatile matter volume percentage can be estimated to be between 25% and 35%. Statistically, there is a standard error of 4.17 based on two standard deviations, resulting in most samples (95%) falling within 4% of the measured %volatile matter. Lastly, the hydrogen index does not correlate to either of the remainder rock matrix components: %fixed carbon, %ash, %moisture, and %coal.

12.2. Density Tool
The bulk density of the rock matrix is controlled by varying proportions of coal minerals, volume of volatile matter, moisture and ash. Therefore, using a density cut off to determine proportions of coal and clay will result in erroneous estimations. The density of coal varies depending on its rank and grade. There is less of a density difference between clay minerals, however a small proportion of quartz will affect the bulk density. It is for this reason that the crossover between the quartz (density) and hydrogen (neutron) index is used to determine coal bearing versus non-coal bearing lithology. This process eliminates the need to estimate an average density cut-off for clay and coal, and at the same time identifies all the coal bearing units interpreted using the core samples.

After the identification of coal versus non-coal bearing rock units, the density tool proves very useful as an estimation of the ash in the rock matrix. The figure below (Figure 51) shows the good linear relationship between the values of the density index and those for the lab measured %ash volume.

Figure 51: Lab measured %ash versus the density index (QtzInd), with the intercept set to 0. This indicates that there is a direct linear relationship between the lab measured %ash volume and the density index (using mud filtrate and quartz densities to calculate the index).

Therefore, it is possible to estimate the proportion of ash in the rock matrix using the density tool only. This will not, however, make a distinction between quartz and clay. The standard error indicates that 95% of the calculated indexed ash samples fall within approximately 16% of the lab measured %ash value. This indicates that while the density tool provides a good measure of the ash proportion in the rock matrix, further investigation is needed to reduce this error range.
12.3. Gamma Ray

The gamma ray tool adds no value to the pseudo proximate analysis other than to define the proportion of clay in the rock matrix. As the definition of the pseudo proximate analysis is to determine components of the rock matrix as measured by the lab proximate analysis, the clay proportion would seem unnecessary. However, this step is necessary as it defines the remaining rock matrix, assuming that the gamma ray tool is an absolute measurement of clay. If the gamma ray is an absolute measure of clay, it then stands to reason that in a clay-rich-only deposition system the gamma ray would correlate to the lab measured %ash volume. However, when using these samples, found in a clay-rich system, the gamma ray does not correlate well to the %ash volume. In fact the gamma ray does not correlate to any of the rock matrix components.

12.4. Conclusion

The neutron tool proves to be very useful in defining the proportion of volatile matter in coal bearing rock matrix. The inaccuracies measured by the neutron tool are caused by the %volatile matter that does not contain hydrogen; two of the more common examples include carbon monoxide and carbon dioxide. An increase in non-hydrogen bearing gases relative to the hydrogen bearing gases, will cause the pseudo proximate calculated %volatile matter proportion to be underestimated. Whereas, a relative decrease in non-hydrogen bearing gases will cause the calculated volatile matter proportion of the rock matrix to be overestimated.

The density tool provides the most accurate measure of the proportion of ash in the rock matrix. These results are based on a minimum bulk density of mud filtrate and a maximum at the quartz percentage. Based on these results, the more accurate measurements are located below an ash volume of 40% when the concentration gets higher than this, the error increases.

It has been demonstrated in this study that based on this dataset and including the principal component analysis, the gamma ray log fails to add value to a pseudo proximate analysis calculation. The reasons for this point to post-deposition alteration of the clay minerals, specifically leaching of uranium in the coal, distorting and homogenizing the readings in coal bearing rock units.

%Fixed carbon does not correlate to either of the three selected logging tools. Based on the dataset used during this study, %fixed carbon cannot be calculated using neutron, density or gamma ray. If the volume proportions of moisture, volatile matter and ash are known for a rock matrix, the remaining rock matrix can be estimated as the volume of fixed carbon. However, the calculation of %moisture presents its own set of problems, as it too does not correlate to any of the logging tools. There is a good negative correlation between the lab measured %ash and the %fixed carbon volume in the rock matrix, this relationship may be used to estimate the %fixed carbon. Lastly, the moisture component of the rock matrix is relatively small and insignificant when determining coal grade, based on this it may be ignored as a rock matrix component.

13. Discussion

Ultimately a pseudo proximate analysis must provide accurate results with limited data availability. This provides the interpreter with the tools needed to make quick estimates of the quality and rank of the coal. The pseudo proximate analysis is applied throughout the entire length of the coal, making initial gas-in-place estimates possible before the results of the lab proximate analysis.

This study set out to determine a method of calculating a pseudo proximate analysis from wireline logs, in order to achieve this following topics were covered:

1. Determine which wireline logging tools will provide the most accurate results at the highest possible resolution.
   - This was achieved using literature and the cluster analysis for rock typing
1. The density, gamma ray and neutron tools provided the best results with the highest resolution.

2. Understand the relationship between the response of the wireline logging tools and coal bearing rock matrix components.
   - This was achieved by using a literature and tool theory to understand the results of each indexed well log.
   - The principal component analysis confirmed interpretation of the indexed well logs and quantified the strength of the relationship between the wireline logs and the rock matrix properties.
   - The neutron log has a strong proportionate relationship with %volatile matter. The density log has an inverse relationship with the %fixed carbon and the proportionate relationship with %ash. %Moisture does not correlate well to any well log.

3. The objective of the thesis was to establish the method and equations used to calculate the pseudo proximate analysis, without the use of lab measured calibration data.
   - This was achieved using the information learned from the principal component analysis, normalization results and the cluster analysis.

4. Calculate the error range for each of the pseudo proximate calculated rock matrix components, as the success of the determined pseudo proximate analysis method relies heavily on the correct selection of the minimum and maximum log values used to normalize the gamma ray, density and neutron wireline log.
   - A sensitivity study was conducted to measure the error range for each rock matrix component, in both a coal and non-coal bearing rock matrix. An example was presented in Chapter 7, however this concept should be applied to the entire well.
   - The results proved that the selection of the neutron minimum and maximum values used to create the hydrogen index had the greatest effect on the calculated pseudo proximate analysis.

5. Test the results of the predictive pseudo proximate analysis method against the results of a calibrated model.
   - A linear multiple regression analysis was selected as the calibration model.
   - A statistical analysis then followed to determine the accuracy of the pseudo proximate analysis using the multiple regression analysis as a comparison.
   - The results proved that the difference between the calculated pseudo proximate analysis and the lab measured proximate analysis are statistically significant.

6. Recommend future changes and enhancements to better define a pseudo proximate analysis.
   - The density and neutron tool have proven to accurately describe the %volatile matter and %ash. The %fixed carbon is more complex and not directly measured by any wireline logging tool. Lastly, %moisture should be ignored when calculating a pseudo proximate analysis.

While this study fails to deliver an accurate method of calculating a pseudo proximate analysis, it succeeds in describing the relationship between the three selected well logs and the results of the proximate analysis. The relationship between density and the proportion of ash in a rock matrix is well documented; however the relationship between the neutron log and %volatile matter is not well understood. This study highlights the importance of the neutron log and its uses.

The project succeeds further by establishing a method using both successful tools, neutron and density, to accurately determine coal versus no coal bearing intervals. The recommendation is that additional tools be tested in the same manner to determine if there is a relationship that can accurately describe an additional rock matrix component.
Bibliography


