Multi-scale porosity characterization of rocks via data mining and geomechanical applications

Pebble Arrambi Diaz

Submitted in accordance with the requirements for the degree of

Master of Philosophy

The University of Leeds

School of Earth and Environment

Institute of Applied Geoscience

January 2026

Declaration

The candidate confirms that the work submitted is her own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Professor Quentin Fisher, for his guidance, patience, and continued support throughout this long and challenging journey. I am deeply thankful for his commitment to my development as a researcher and for the considerable time and effort he dedicated to my work, particularly during periods when progress was most difficult.

I would also like to thank Dr Chrysothemis Paraskevopoulou, who was part of my supervisory team during the early stages of my studies and contributed to shaping the initial direction of this research. I am also grateful to Professor Paul Glover, who joined my supervisory team more recently; his independent assessment of my work provided both clarity and essential encouragement at an important stage of my studies.

A significant portion of this journey was navigated through the Progress Support Process (PSP). I wish to thank Dr Paul Brockway, Dr Milena Buchs, and Rachel Furniss for their compassionate involvement. Their willingness to engage with my circumstances and provide structured guidance was instrumental in helping me navigate the complexities of my studies. I am also grateful to the wider School and Graduate School teams for their professional support and understanding.

I would like to acknowledge the extensive support I received from the University of Leeds more broadly, including Disability Services, Student Wellbeing, and mentoring support. Their consistent assistance was vital in helping me manage significant challenges and sustain my academic momentum.

This research journey, which began as a PhD and later transitioned to an MPhil, unfolded against a backdrop of exceptional personal and global adversity, including the COVID-19 pandemic and the loss of my father. These experiences shaped both my academic focus and my personal growth. While this period tested my limits, it also taught me the true meaning of resilience, self-awareness, and perseverance.

On a personal level, I am forever indebted to my family for their love and unwavering belief in me. Their encouragement was an anchor during moments of doubt. To my partner, Tom, thank you for being the heart of my support system and for walking beside me through every high and low of this long journey. I am also deeply grateful to the friends who became my family in Leeds; your kindness and companionship turned a difficult road into one I did not have to walk alone.

Finally, I wish to acknowledge my own commitment to this work. Completing this degree required persistence in the face of immense uncertainty, and I am proud of the knowledge I have gained and the resilience I have proven to myself throughout this journey.

I acknowledge the use of ChatGPT-4 (OpenAI https://chat.openai.com/) for assistance with proofreading my final draft and refining the academic tone of my writing.

Abstract

Accurately characterizing the geomechanical properties of subsurface reservoirs is critical for assessing wellbore stability, designing hydraulic fracture treatments, and predicting surface subsidence. While direct laboratory measurements on core plugs remain the definitive standard, the limited availability of physical core material necessitates the development of alternative predictive methodologies, such as microstructural image analysis. This research presents an optimized workflow for estimating porosity and dynamic elastic properties utilizing high-resolution Backscattered Electron Micrographs (BSEM).

Initial investigations utilizing moderate-magnification BSEM images yielded significant statistical scatter $(R^2 = 0.555)$, with measurement errors ranging from -600% to 100% in heterogeneous datasets. These inaccuracies were systematically attributed to macroscopic sample heterogeneity and resolution-dependent bias, where fine-scale pore details are unresolved during digital acquisition. To mitigate these effects, a standardized imaging protocol was established, identifying a scan speed of level 4 as the optimal balance between signal-to-noise ratio and acquisition efficiency for high-volume automated scanning.

The core methodology employs fractal geometry to address resolution limits. By applying a power-law regression to iterative 50% downscaled resolution levels, porosity values were extrapolated to a theoretical *infinite magnification*. This approach is theoretically grounded in the scale-independent nature of sandstone pore networks established by Thompson, Katz, and Krohn (1987). Furthermore, the study demonstrates that satisfying the Representative Elementary Volume (REV) through large-scale montages (~1 cm²) is essential for mitigating spatial variability and windowing bias.

The synthesis of optimized BSEM data with the PETGAS ultrasonic dataset provides a critical *geomechanical bridge*, allowing for the calculation of dynamic Young's modulus and Poisson's ratio. Ultimately, this thesis establishes a robust framework for integrating micro-scale image data and data mining workflows to enhance the predictive accuracy of subsurface geomechanical models.

Table of Contents

List of Figures

Chapter 1 Overview

1.1. Project Rationale and Background

Knowledge of the mechanical properties of rocks is essential for a wide range of subsurface applications, including assessing wellbore stability, predicting surface subsidence resulting from fluid withdrawal, and designing hydraulic fracture treatments. In the energy sector, particularly within tight gas sandstone reservoirs, these properties dictate the economic viability and safety of extraction processes.

Ideally, geomechanical properties, such as Young’s modulus, Poisson’s ratio, and uniaxial compressive strength (UCS), would be measured directly through carefully controlled laboratory experiments on core plugs. However, a significant challenge in reservoir characterization is that only a small proportion of the subsurface is ever physically sampled. Furthermore, when core material is unavailable, geoscientists must often rely on irregular fragments, such as drill cuttings, which are insufficient for traditional displacement or pressure-based geomechanical testing.

1.2. The Role of Microstructure and Image Analysis

As an alternative to bulk physical testing, it is possible to estimate geomechanical properties based on a rock's microstructure. Porosity is recognized as a primary control on mechanical properties, fundamentally influencing a rock's strength and elastic deformation behaviour. Consequently, accurately quantifying the void space within a rock is a critical first step in predicting its macroscopic geomechanical response.

Scanning Electron Microscopy (SEM) has emerged as a powerful tool for this purpose. By utilizing Backscattered Electron (BSE) imaging, researchers can estimate porosity directly from the rock's microstructure, providing detailed textural and spatial information that traditional bulk measurements cannot resolve. However, the accuracy of these image-based estimates is highly sensitive to acquisition parameters, including magnification, resolution, and scanning speed.

1.3. Data Mining and Petrophysical Integration

Recent advancements in data science offer new opportunities to improve petrophysical predictions. Regressions and digital rock characterisation algorithms can be deployed to identify scaling patterns within microstructural image data. By integrating high-resolution BSEM micrographs with established laboratory datasets, such as the PETGAS ultrasonic velocity and density records, this research seeks to develop more robust empirical models for predicting geomechanical properties from micro-scale observations.

1.4. Aims and Objectives

The primary aim of this research is to develop and optimize a methodology for estimating geomechanical properties from BSEM image data. The specific objectives are:

- Evaluate the Accuracy of BSEM-derived Porosity: Compare image-analysis results with laboratory-measured helium porosity to identify sources of error.
- Optimize Imaging Parameters: Investigate the impact of scanning speed and magnification on signal quality and data reliability.
- Address Resolution Bias: Apply fractal-based power-law extrapolation to estimate porosity at a theoretical "infinite magnification".
- Determine Representative Elementary Volume (REV): Assess the impact of sample heterogeneity and field-of-view on the representativeness of microstructural data.
- Predict Dynamic Elastic Properties: Utilize optimized image data to calculate geomechanical parameters essential for subsurface engineering.

1.5. Thesis Structure

- **Chapter 2** provides a comprehensive review of geomechanical properties, microstructural controls, and machine learning applications in petrophysics.
- **Chapter 3** establishes the baseline methodology for estimating sandstone porosity from BSEM images and discusses the initial causes of measurement scatter.

- **Chapter 4** investigates the impact of scan speed and resolution, introducing fractal extrapolation and the REV concept to improve accuracy.
- **Chapter 5** synthesizes the research findings, discusses the limitations of the current methodology, and provides recommendations for future work regarding the estimation of geomechanical properties from image data

Chapter 2 Estimation of the geomechanical properties of rocks from microstructure: a review

2.1. Introduction

The aim of the research project on which this thesis is based is to estimate the geomechanical properties of rocks based on their microstructure. The following chapter reviews key aspects of this topic. It begins by reviewing key mechanical properties and their controls (**Section 2.2**). The review highlights that porosity is a key control on mechanical properties. The review therefore discusses methods to estimate porosity from image data (**Section 2.3**). To improve estimates of mechanical properties beyond correlations with porosity it may be possible to use machine learning techniques; these are therefore reviewed in **Section 2.4**.

2.2. Mechanical properties

Geoscientists often require a knowledge of a wide range of properties to predict the deformational behaviour of the subsurface. For example: -

- Elastic properties, such as Young's modulus and Poisson's ratio are needed to provide the relationship between stress and strain prior to yield.
- Failure criteria such cohesion, UCS, and friction angle to assess under what stress conditions a rock is likely to experience brittle failure.
- Plastic yield criteria such as the apparent pre-consolidation pressure (p*) to differentiate the elastic and inelastic deformations ranges.

Measurement of seismic and ultrasonic velocities allows the measurement of dynamic elastic properties. Other geomechanics properties (e.g. UCS, p* etc.) can only really be measured in laboratory environment so correlations are often developed between these properties and index properties such as porosity. In the following section a review is provided on elasticity, seismic velocity, and porosity. Later in this research, this review will be extended to include the relationship between these properties and more complex geomechanical properties needed to predict subsurface deformation.

2.2.1. Elastic properties

Stress and strain are two inherent concepts to elasticity of materials. Additionally, certain properties such as porosity, permeability, and pore space and sample volume are considered stress-sensitive (Mavko et al., 2009). Stress is the intensity of loading and is described as the load per unit area (Budhu, 2010).

$$\sigma = \frac{F}{A}$$

where *A* is the cross-sectional area and *F* is the applied force acting on that area. The unit used for stress can be Pa (Pascals) or psi (pounds per square inch) (Fjaer et al., 2008). Fjaer (2008) explains that the sign of the stress is not exclusively defined by the physics involved in the situation, but by convention the compressive stresses are positive in rock mechanics.

A change in external force on a sample caused is to deform. The amount of deformation is termed strain. It is denoted by the ratio of change in a dimension to the original dimension (Budhu, 2011).

$$\varepsilon = \frac{\Delta \mathrm{L}}{\mathrm{L}}$$

As this comes from the ratio of two quantities with the same unit, it is a non-dimensional measure (Hearn, 1997).

2.2.1.1. Young's Modulus

The Young's modulus, *E,* is the slope of the stress strain line for a linear isotropic material undergoing compression in one direction (Budhu, 2011) and is defined as: -

$$E = \frac{\sigma}{\varepsilon} = \frac{\mathrm{F\,L}}{\mathrm{A\,\Delta L}}$$

Young's modulus is a measure of the stiffness of an elastic material (Ma et al., 2015). Mavko (2009) defines it as the ratio of the extensional stress to the extensional strain in a uniaxial stress state. This uniaxial stress state can be achieved by carrying out a standard compression test on a specimen, which is the usual approach for determining its actual value (Hearn, 1997). It is generally assumed that Young's modulus is the same in tension and compression and for most rocks and sediments generally exceeds 2 GPa and can range up to 100 GPa (Hearn, 1997).

2.2.1.2. Poisson’s ratio

Poisson's ratio describes the ratio of deformation of a material relating the change of lateral and longitudinal dimensions when this is subjected to a load (Hearn, 1997). In other words, it is a ratio of strain magnitudes.

$$\upsilon = -\frac{\varepsilon_{lat}}{\varepsilon_{lon}}$$

To understand the negative sign that precedes this ratio, is important to remember that the axial strain induces a diametrical strain of opposite sign, as this would cause a tensile-compressive effect on the longitudinal and lateral strains (Hearn, 1997). Hearn (1997) states that for most engineering materials, the value of υ lies between 0.25 and 0.33, however, Goodman (1991) explains that for linearly elastic and isotropic rocks the value of υ must lie in the range of 0 to 0.5; often it is assumed equal to 0.25.

2.2.1.3. Bulk Modulus

Another useful elastic modulus is the bulk modulus (*K*). It expresses the relationship between mean effective stress (σ_0) and volumetric strain. *K* is the ratio of these two values, and it is broadly used to describe the compliance of a liquid, gas or solid, as it is the reciprocal of the compressibility (Mavko, 2009).

$$\frac{\Delta \mathrm{V}}{\mathrm{V}} = \varepsilon_{\mathrm{lon}} + 2\varepsilon_{lat}$$

$$\sigma_0 = \frac{\sigma_1 + \sigma_2 + \sigma_3}{3}$$

$$K = \frac{\sigma_0}{\frac{\Delta \mathrm{V}}{\mathrm{V}}}$$

2.2.2. Yield/Failure criteria

2.2.2.1. Mohr-Coulomb yield criteria

The concept *criterion of failure* best describes the variation of peak stress, σ_1, with confining pressure, σ_3, (Goodman, 1991). There are several methods available to predict failure in a reservoir model, however, the Mohr-Coulomb (MC) failure criterion is the simplest and best-known failure model for rocks (Goodman, 1991).

The MC failure criterion describes the conditions for which an isotropic material will fail expressed in a set of linear equations in principal stress space (Labuz et al., 2012). It consists of a linear envelope touching all Mohr's circles, which represent combinations of principal stresses that result in failure (Goodman, 1991). Accordingly, Labuz (2012) notes that this failure criterion appears to apply reasonably well to rocks when all principal stresses are compressive.

Aiding this purpose, Mohr's circles help to visualize the stress levels acting on a body and include several features that allow us to work out what is called Mohr's envelope criterion of failure (Coates et al., 1981). This envelope is in essence a definition of the maximum stress level that a rock can sustain (Coates et al., 1981).

Figure 2-1 illustrates the variation in maximums stresses for different confining pressures of a rock specimen, such data can be obtained and assessed after carrying several triaxial tests. Any stress beyond that envelope will cause failure, which indicates that failures in tension or shear can be predicted (Coates et al., 1981).

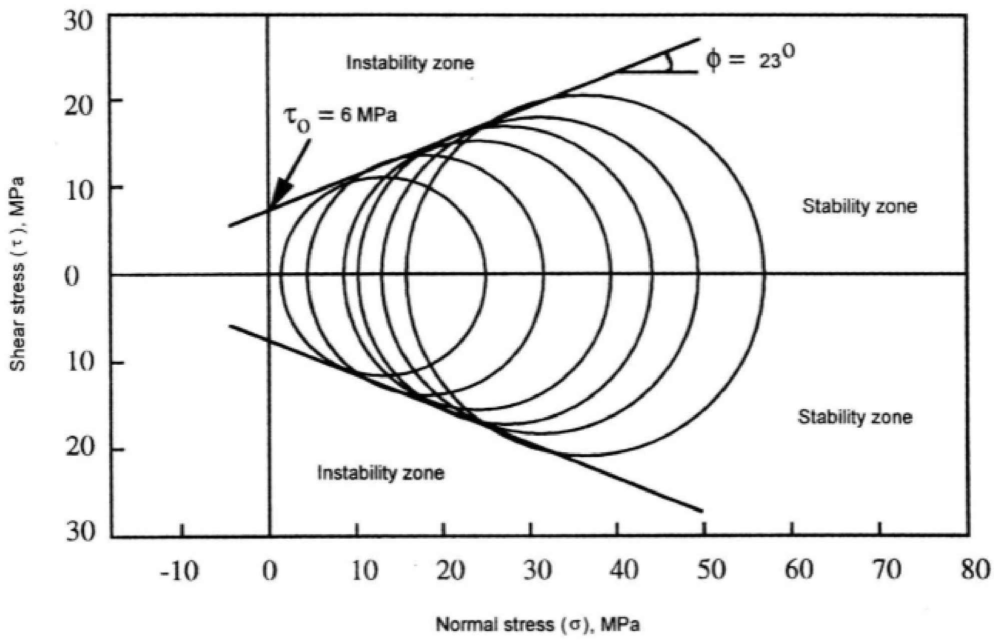


***Figure 2-1** Example of Mohr-Coulomb failure criteria (Al-Awad, 2002)*

Al-Awad (2002) explains that the MC failure criterion is a function of the apparent cohesion (τ_0) and the angle of internal friction (ϕ). This relationship is expressed mathematically as:

$$|\tau| = \tau_0 + \sigma \tan \phi$$

Budhu (2011) proposes that the term cohesion (i.e. the interception of the envelope with the shear stress axis) should be described as the action of intermolecular forces on the shear strength of soils.

2.2.2.2. Cap-plasticity models

Mohr-Coulomb type constitutive models are useful for brittle materials. However, it is well-known that under high confining pressures (relative to their strength) samples can undergo ductile deformation and this is not described by the MC-type models. Around the 1960's and 70's workers at the University of Cambridge, studying the deformation of soils, developed the so-called Cam-Clay (Roscoe, Schofield, and Wroth, 1958; Schofield and Wroth, 1968) and Modified Cam Clay (Roscoe and Burland, 1968) models. The models accounted for observations that soils hardened when exposed to stresses that were high compared to their strength (e.g. unconfined compressive strength.

Typically, these constitutive relationships create an elliptical yield surface on a plot of differential stress, Q, against mean effective stress, P, where: -

$$P = \sigma_1 - \sigma_3$$

$$Q = \frac{\sigma_1' - 2\sigma_3'}{3}$$

where σ'_1 and σ'_3 are the maximum and minimum effective stresses (i.e. stress minus pore pressure) respectively.

Commonly referred to as cap-plasticity models, these constitutive models (e.g. **Figure 2-2**) are particularly useful in that they take into account the behaviour of rocks that are weak with respect to their effective stress conditions. The position at which the subsurface stresses intersect the yield surface is provides a good indication of whether the rock is likely to experience dilation or compaction. The collapse pressure under hydrostatic conditions, *p**, is a very useful geomechanical property that can allow the prediction of how a rock is likely to deform under given stress conditions (Wong et al., 1997; Fisher et al., 2003).

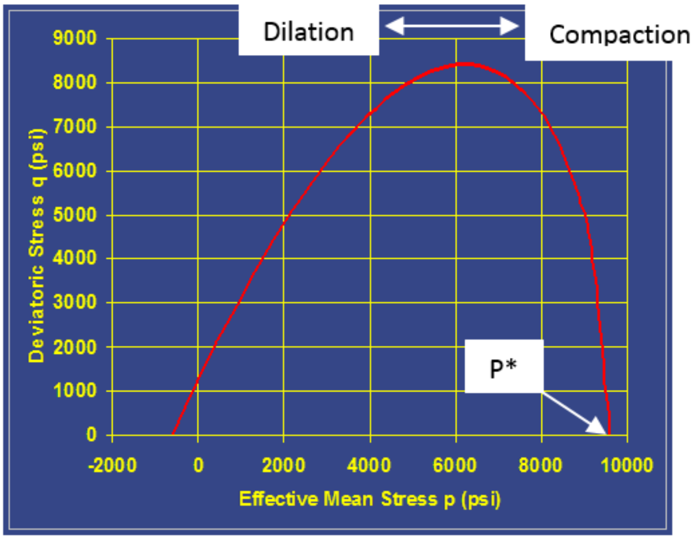


Figure 2-2 *Cap-plasticity model for rock deformation highlighting the collapse pressure under hydrostatic conditions (p*).*

2.2.3. Estimation of geomechanical properties

Ideally, geomechanical properties would be measured from carefully controlled laboratory experiments. However, this is often not possible because only a small proportion of the subsurface is ever sampled (i.e. core taken). It is therefore necessary to estimate subsurface geomechanical properties using other methods. In the following section two of these are described: correlations with other properties that can be measured in the subsurface and deterministic modelling.

2.2.3.1. Estimation of static elastic properties from seismic and ultrasonic velocities

It has been recognised that information about rock properties can be obtained from the results of seismic surveys (Dai et al., 1995). Seismic waves enable us to obtain information about parts of a formation that might be considered inaccessible by any other media as they can travel very long distances through the earth (Fjaer et al. 2008). There are two types of body waves that propagate as an elastic medium (Sjogren, 1984). The first one to be detected on an earthquake record is the P-wave, where P stands for primary, but it is also known as a longitudinal or compressional wave (Sjögren, 1984). In **Figure 2-3** it can be seen that the particle movement in this type of wave occurs in the same direction that the wave is moving (Endsley, 2019). The second wave type is called transverse, shear or S-wave (Sjögren, 1984). The motion of the particles in this type of wave is at right angles of the propagation direction (Sjögren, 1984).

The response of elastic waves is controlled by the elastic properties and density conditions of the material through which it is propagating (Sjögren, 1984). The velocity on which waves propagate is given by elastic stiffness and density of the rock, and these parameters are simultaneously dependent upon other parameters such as porosity. (Fjaer et al., 2008). At the same time, the elastic response of a porous material is affected by the presence of a pore fluid (Fjaer et al., 2008).

Goodman (1991) suggests that if the rock were an ideal, elastic isotropic solid, then the Young's Modulus can be calculated from the following equations:

$$E = V_l^2 \rho$$

$$E = 2(1 + \nu)\rho V_s^2$$

In addition, Mavko (2009) also stressing the isotropic, homogenous, and elastic features of the media involved, illustrates that the P-wave and S-wave velocities can be represented in terms of Poisson's ratio.

$$\frac{{V_P}^2}{{V_S}^2} = \frac{2(1 - \upsilon)}{(1 - 2\nu)}$$

$$\nu = \frac{{V_P}^2 - 2{V_S}^2}{2({V_P}^2 - {V_S}^2)}$$

Adding to this, Mavko (2009) shows that the bulk modulus can be extracted from measurements of density and any two wave velocities.

$$K = \rho({V_P}^2 - \frac{4}{3}{V_S}^2)$$

Velocity and porosity often have an inverse relationship (Barton, 2006). However, this relation has significant scatter as other factors (e.g. mineralogy, fluid saturation, the presence of micro fractures etc.) impact velocity (King, 2005). There are several factors that influence compressional P and S-waves in terms of their velocities and attenuation in sedimentary rocks. For example, it is well known that variations in the clay content of sandstones have a significant impact on velocity (e.g. Wilkens et al., 1984). In particular, Han et al. (1986) found empirical regressions that relate the presence of clay and includes it in the velocity - porosity relation.

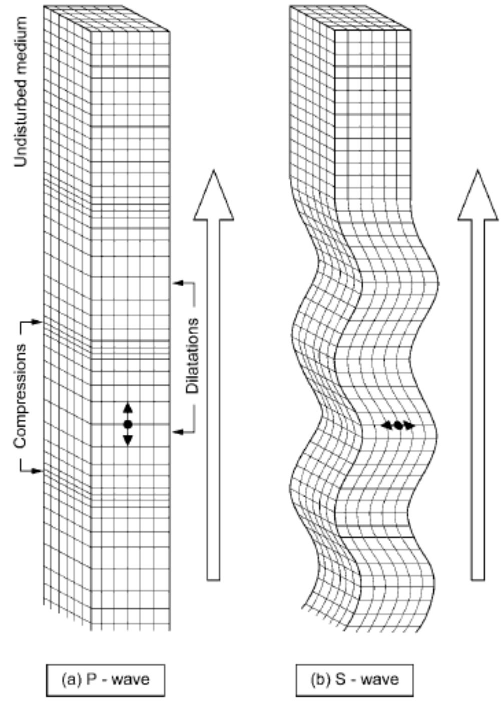


Figure 2-3 *Elastic deformations and particle motions associated with the propagation of body waves (Barton, 2006)*

2.2.3.2. Estimation of mechanical properties from porosity

Porosity is one of the reservoir parameters that is essential for reserve estimation, reservoir simulation and pore pressure prediction (Dvorkin et al., 2001). The porosity and saturation of rocks play a fundamental role in determining the velocities of their compressional and shear waves (Toksöz et al., 1976). Toksöz et al. (1976) analysed velocities of seismic waves in porous rocks with different saturated conditions to compare theoretical calculated data with data measured in the laboratory. A greater effect was shown on compressional velocities relying on the properties of the saturating fluid (Toksöz et al., 1976). This can be explained by the fact that fluids cannot transmit shear waves (Fjaer et al., 2008).

Elastic properties

Nur et al. (1991,1998) discussed the idea that the P and S velocities of porous rocks should have values between the velocities of the mineral grains in the limit of low porosity and the values for a mineral-pore-fluid-suspension in the limit of high porosity. The contact of the grains determines the stiffness of the material in the lower porosity domain, nonetheless, in the higher-porosity domain the grains are suspended in water and no longer in contact (Dvorkin et al., 2001). This threshold that bounds where the rock can exist only as a suspension is named critical porosity (Dvorkin et al., 2001).

Figure 2-4 illustrates the concept of critical porosity, $\emptyset_c$, that separates grain supporting and fluid supporting domains (Mavko et al., 2009). In addition, it is stated that the value of $\emptyset_c$ depends on the sorting and angularity of the grains from which the sediment is composed (Mavko et al., 2009).

Several authors have recognized empirical relations between velocity and porosity (see Mavko et al., 2009 and references within). One of the most discussed approaches to correlate such properties, are the studies carried by Wyllie et al. (1956, 1958, 1962), which unveiled a relatively simple monotonic relation that often can be found between velocity and porosity in sedimentary rocks, providing certain circumstances: (i) they have relatively uniform mineralogy; (ii) they are fluid-saturated, and (iii) they are at high effective pressure(Wyllie et al., 1956, 1958, 1962).

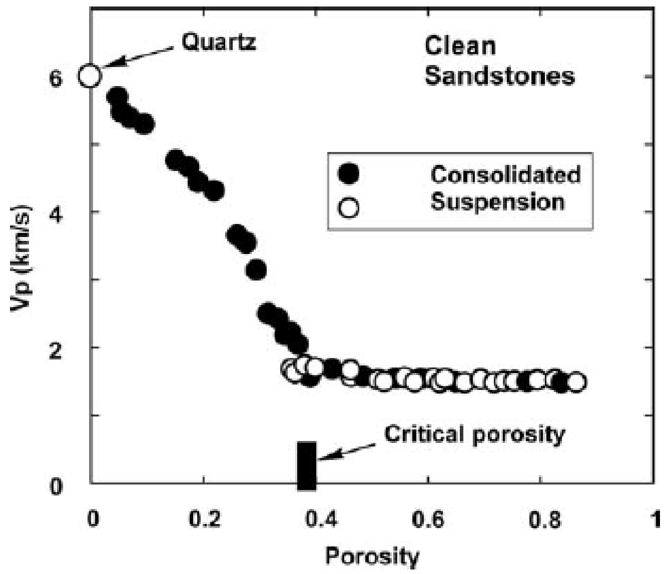


Figure 2-4 *P-wave velocity versus porosity for compact sandstones, and for a suspension of component grains, behaviour that is separated by the critical porosity concept (Barton, 2006)*

Wyllie et al., approximated these relations with the following expression (Mavko et al., 2009).

$$\frac{1}{V_P} = \frac{\emptyset}{V_{P-fl}} + \frac{1-\emptyset}{V_{P-0}}$$

In this equation, a relation is shown between P-wave velocities of the saturated rock (V_P) of the mineral material that constitutes the rocks and the pore fluid, illustrated by $V_{P\text{-}0}$ and $V_{P\text{-}fl}$, respectively and the porosity (Mavko et al., 2009).

Han et al. (1986) also studied the effect of clay content on wave velocities and concluded that clay content is the next most important parameter to porosity in reducing velocities. Likewise, Mavko et al. (2009) observed that when clay is present the correlation with porosity becomes poor but becomes very accurate if clay volume is considered in the regression, concluding that clay content is quite a relevant parameter for quantifying velocity.

Yield/failure criteria

Chang et al. (2006) summarized 31 empirical equations relating unconfined compressive stress (UCS) to physical properties such as velocity, Young's modulus and porosity in sandstone, shale, limestone, and dolomite. It can be argued that for sandstone and shale, the equations presented in his work worked fairly well to relate the porosity-strength relationship.

UCS is a very important mechanical property, necessary to investigate a range of geomechanical problems. Al-Awad’s (2002) study suggests the development of a correlation between the apparent cohesion and the UCS of materials. The laboratory data to which this is based comes from more than 200 rock samples of different types obtained from literature. According to the results and comparisons to other references, the average error of estimation is equal to 10%. The correlation made is based from the understanding that rock apparent cohesion is a measure of the degree of grain -to-grain bonding, and that the uniaxial compressive strength is also a measure of grain-to-grain bonding magnitude. ***Figure 2-5*** shows the trend that was obtained from the plotting of the apparent cohesion and UCS from the literature cited within Al-Awad’s study.

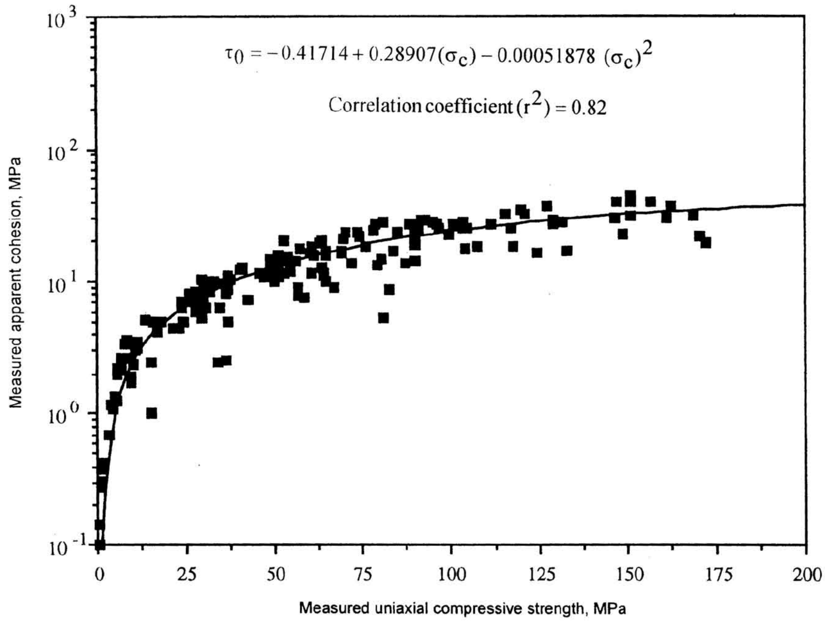


Figure 2-5 *Relationship between measured uniaxial compressive strength and measured apparent cohesion of various types of rocks obtained from literature (Al-awad, 2002)*

Apparent pre-consolidation pressure

The pressure at which pore collapse occurs in sandstones under hydrostatic pressure (*p** described in **Figure 2-2**) has been measured in many sandstone samples (e.g. Wong et al. 1997). A key finding is that p* is correlated with the product of porosity and grain-size (**Figure 2-6**). This observation indicates that it is possible to obtain an important mechanical property from image data.

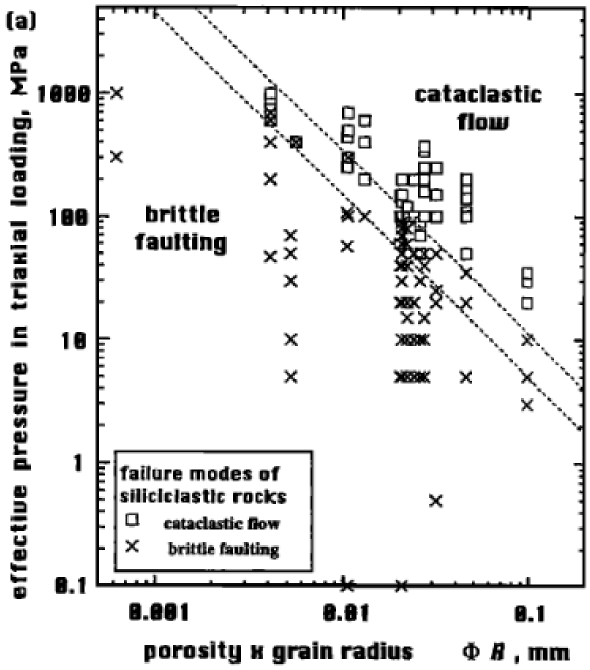


Figure 2-6 *Plot of effective pressure required for pore collapse against the product of porosity and grain-size (from Wong et al. 1997)*

2.2.3.3. Estimation of mechanical properties from microstructure

Knowledge on the microstructure of a rock may aid in the estimation of its geomechanical properties (Zum Gahr, 1987). This is due to the existent dependency between the variation of the properties and the details of the microstructure, which vary extensively as it depends on the properties of the materials (Zum Gahr, 1987).

Studies have been undertaken to predict the mechanical responses of geomaterials using finite element models constructed by digital image processing of microstructural images (e.g. Yue et al., 2003). The study undertaken by Yue et al. (2003) uses images from a digital camera of the fresh cross-sections, which after analysing non-homogeneous materials under the Brazilian split test, showed that the method can be used to carry out the micro-mechanics analysis with high efficiency and accuracy.

2.3. Estimating porosity from images

Estimating porosity from image data has become an essential method in petrophysical analysis, particularly in the study of rock samples such as sandstone, shale, and carbonate rocks. Most studies have concentrated on estimating porosity from SEM images, so the review concentrates on this method. CT scanning and optical microscopy have also been used to estimate porosity, so a brief discussion of this technique is also presented.

2.3.1. Scanning Electron Microscope (SEM)

SEM image analysis is a widely used technique for estimating porosity of rocks. The following section begins by describing how this method involves capturing high-resolution images of rock samples and applying image processing techniques to quantify porosity and pore size distribution. Saraf et al. (2019) used SEM images to evaluate the porosity of shale samples from the Jhuran formation in India. By using thresholding algorithms and the ImageJ software, they successfully estimated porosity values that were validated against helium porosimeter measurements, demonstrating the effectiveness of non-destructive SEM image analysis in petrophysical studies.

Similarly, Solymar and Fabricius (1999) utilized backscattered SEM images to estimate the porosity and permeability of the Amager Greensand. By simplifying images into two phases (pores and particles), they calculated the specific surface area of the solid phase and applied the Kozeny equation to estimate permeability. Their findings emphasized the reliability of SEM image analysis in providing detailed pore structure information crucial for fluid flow interpretation.

2.3.1.1. Imaging using SEM

A SEM scans a focussed electron beam across as sample in a raster pattern and the position of the beam is combined with the intensity from various detectors to produce an image. The electron beam interacts with the sample to produce several different forms of radiation (**Figure 2-7**) that are useful for sample characterization including:

- **X-rays** are generated due to the movement of electrons between different shells. The energy of the electrons is characteristic of the elements present in the phase on which the beam is focussed. The X-ray spectrum produced (e.g. **Figure 2-8)** provides the chemical composition of the phase present and allows is very valuable for identifying the minerals present.

- **Secondary electrons (SE)** are produced from the atoms of the sample due to inelastic interactions with the beam. A secondary electron detector is positioned at above the sample at around 45^o to the electron beam. The number of SE reaching the detector will be controlled by the angle of the surface on which the electron beam is focused and from which the secondary electron is generated to the detector. More secondary electrons will reach the detector from surfaces that are orientated towards the detector than those that face the opposite direction. The SE signal therefore generates information on the morphology of the surface of the sample (**Figure 2-9a**).
- **Back-scattered electrons (BSE)** are reflected from the samples surface as a result of elastic interactions. The BSE electron detector is positioned directly above the sample and the number of electrons that are generated and reach the detector is directly proportional to the mean atomic number of the phase onto which the electron beam is focussed. BSE images are therefore extremely useful for obtaining information on the distribution of different minerals and porosity (**Figure 2-9b**) because the brightness of different phases varies depending upon their mean atomic number; minerals with the same mean atomic number will have the same brightness on BSEM images. BSE images are particularly useful for imaging resin-filled porosity because the resin has a very low mean atomic number and therefore appears darker than the minerals from which rocks are composed. In addition, discrimination of the pore space against the minerals present is facilitated due to the brightness of each phase being proportional to its mean atomic number.
- **Cathodoluminescence (CL)** is light that is generated when the electron beam interacts with a material. The intensity and wavelength distribution of the CL signal is controlled by subtleties in the chemical composition of minerals and is a valuable tool for analysing crystal growth structures (e.g. **Figure 2-9c**).

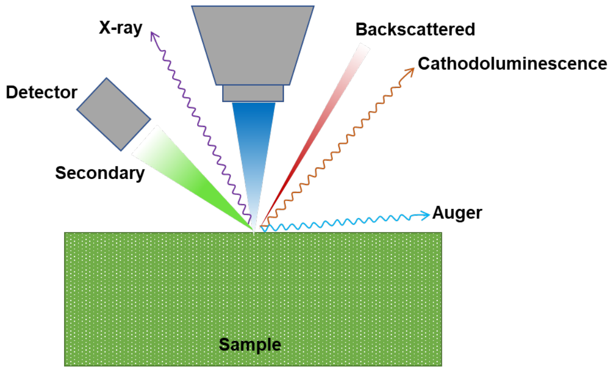


Figure 2-7 *Diagram showing the different types of radiation generated as an electron beam is scanned across a sample.*

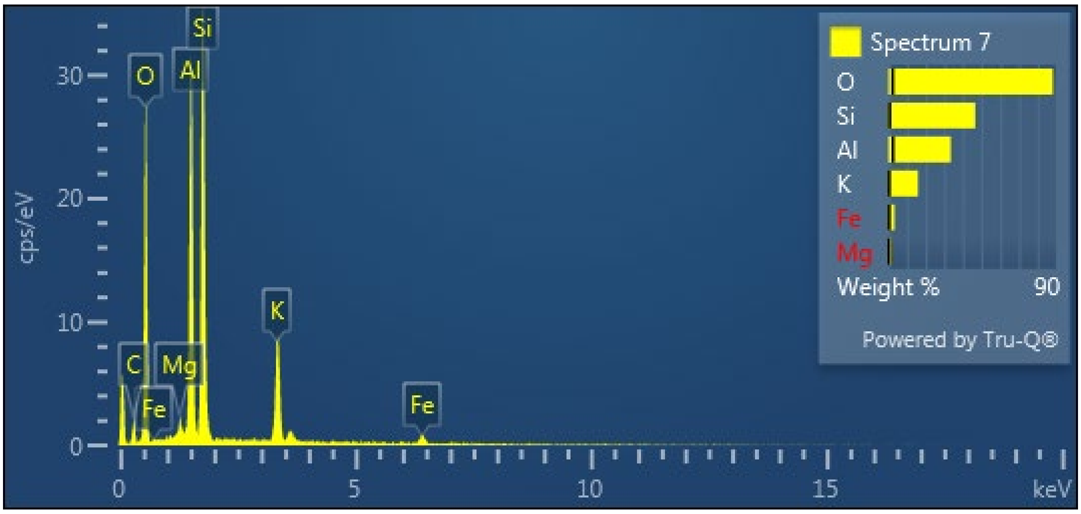


Figure 2-8 *Example of an EDS spectrum obtained from a clay mineral using an Oxford instruments EDAX detector fitted to the Tescan SEM at the University of Leeds.*

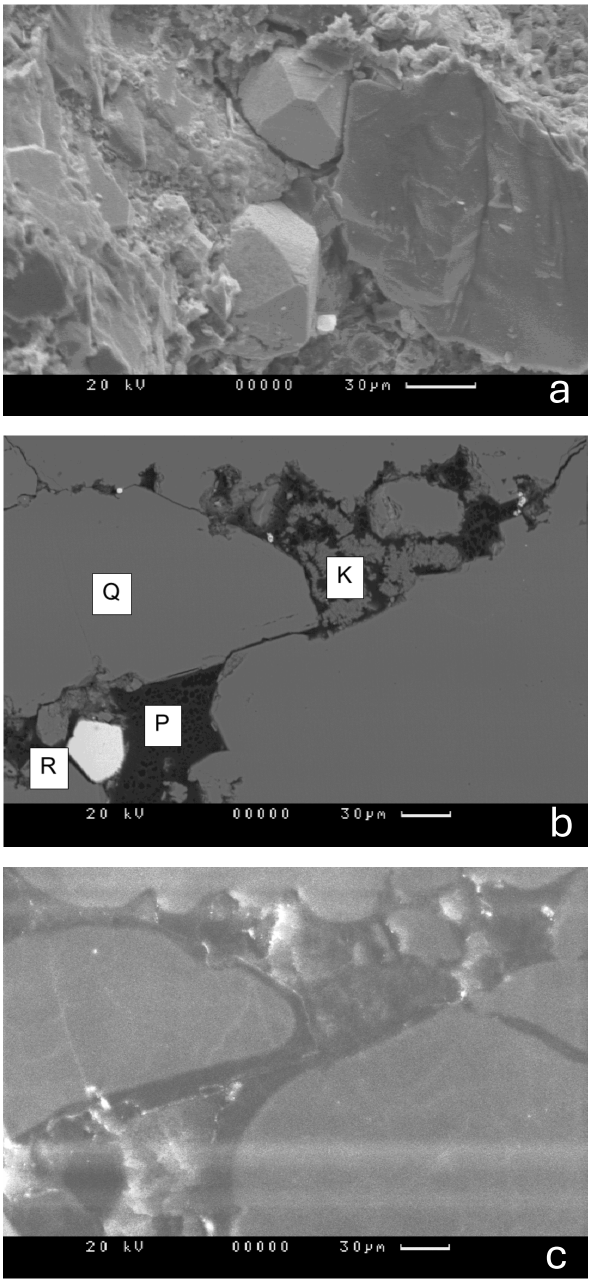


Figure 2-9 *a.) Secondary electron image showing the surface morphology of a sandstone sample, b.) BSE image showing of a sandstone sample; the brightness of each phase is proportional to the mean atomic number of phase on which the beam is focused. Here quartz (Q), rutile (R), kaolin (K) and porosity (P) have different brightness; the porosity is the darkest as it has the lowest mean atomic number, and c.) CL image taken from exactly the same part of the sample as the BSE image shown in image b; the brightness of phases is controlled by subtleties in their chemistry, which makes CL a valuable tool for identifying different growth stages of minerals. In this case, an authigenic quartz overgrowth (OG) is clearly darker than the detrital quartz (DQ).*

These different signals mean that the SEM is a powerful tool imaging the microstructure and morphology of materials. The small wavelength of the electrons generated in an SEM allows imaged to be generated at the sub-μm resolution. Indeed, state-of-the-art field electron SEMs can generate images with a resolution close to 1 nm (**Figure 2-10**). At the opposite end of the scale, SEMs can also produce ultra-low magnification images. For example, in this study a SEM has been used to image an entire thin section in a single frame (**Figure 2-11**).

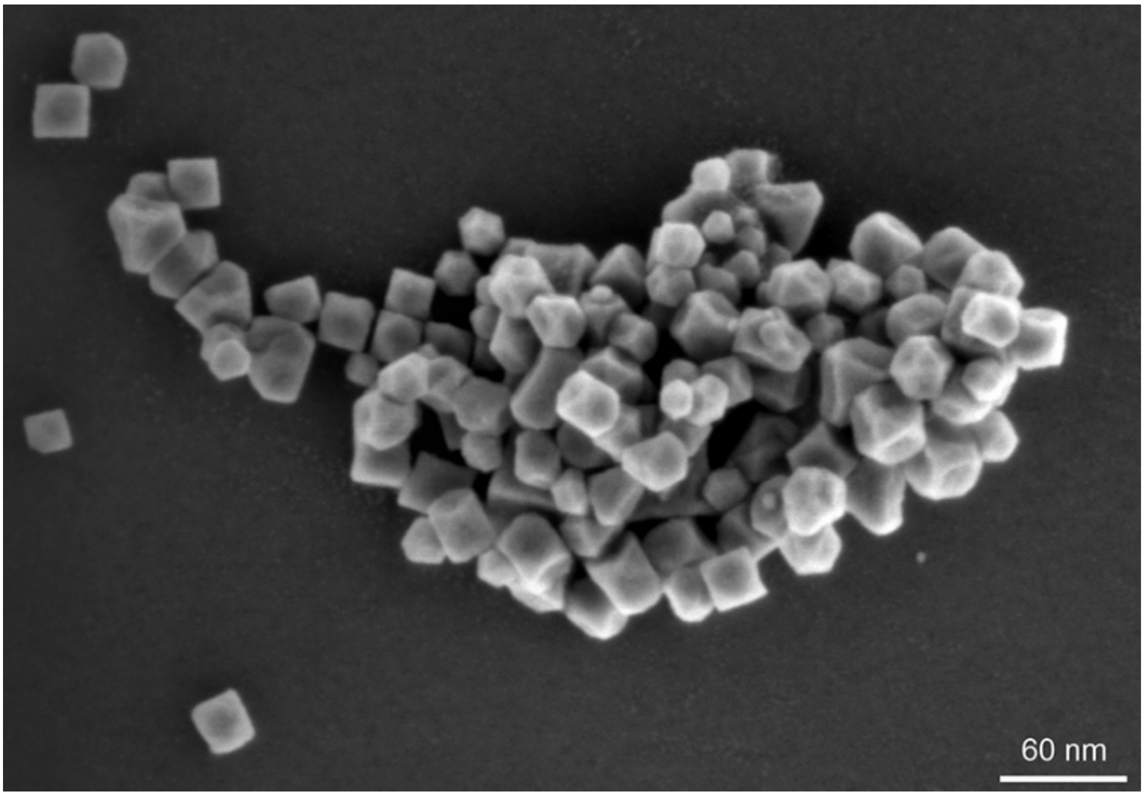


Figure 2-10 *Ultra-high resolution SEM image of magnetic nanoparticles (from https://www.zeiss.com/microscopy/en/products/sem-fib-sem/sem/geminisem-family.html)*

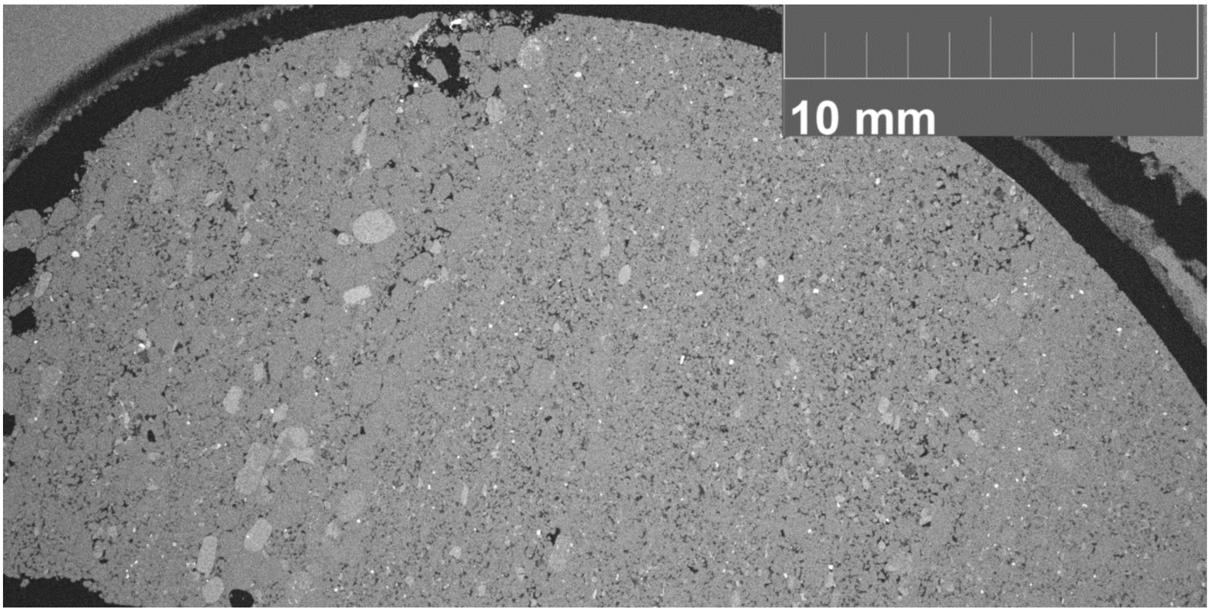


Figure 2-11 *Ultra-low magnification BSEM image of an entire thin section taken using the Tescan SEM at the University of Leeds; the field of view is around 2.5 cm.*

2.3.1.2. Estimation of porosity from BSE images

Estimates of porosity from BSE images are generally made using image analysis software. In particular, a grey level is identified which separates porosity (as shown in **Figure 2-12a**) from the surrounding mineral matrix. The image analysis software, displayed in **Figure 2-12b**, is then used to calculate the area containing pixels that have grey levels below this threshold. There has been variable success at estimating porosity using this technique. For example, (Bøe et al., 1999) showed that the porosity estimated by images analysis was considerably lower than measurements made in the laboratory using helium expansion (**Figure 2-13).** On the other hand, (Ruzyla, 1986) showed a good 1:1 correlation between the porosity estimated from BSE images and that measured in the laboratory (**Figure 2-14)**.

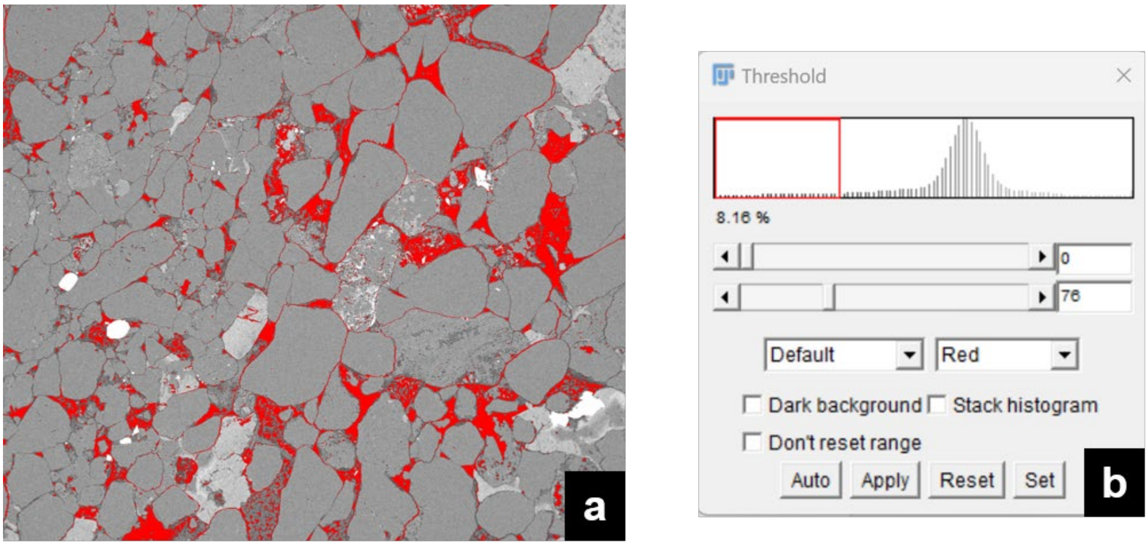


Figure 2-12 *a.) BSEM image in which a grey level threshold has been chosen to distinguish the porosity (red) from the mineral matrix b.) Snapshot of the ImageJ software threshold selection interface used for segmentation process.*

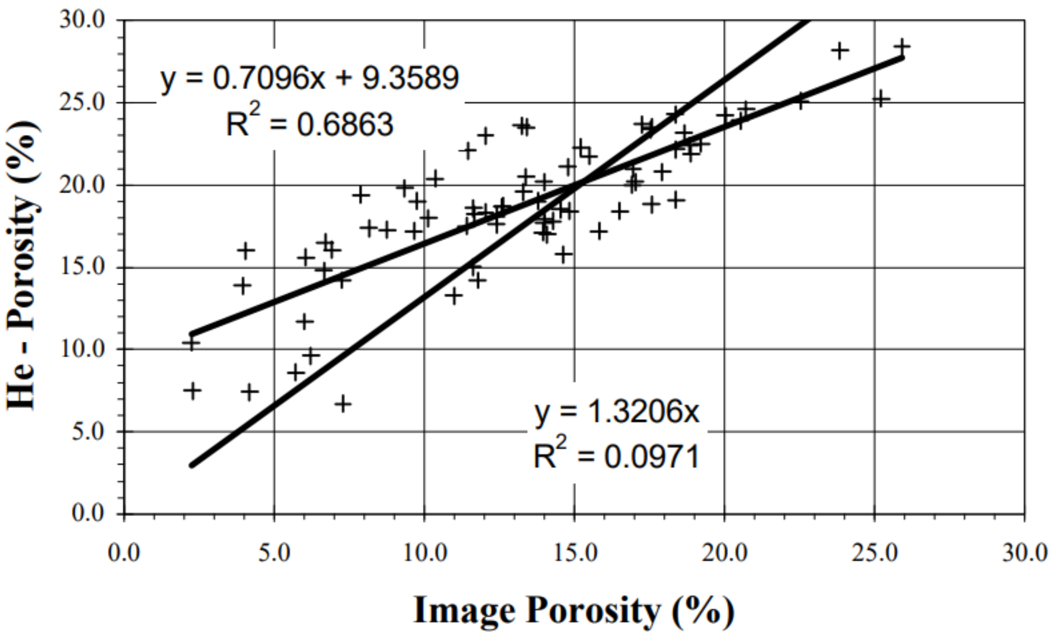


Figure 2-13 *Plot showing that porosity estimated by image analysis is consistently lower than that measured using helium porosimetry on plugs (from Bøe et al. (1999). Two regression lines with equation and correlation coefficients are included in the diagram; one of the lines represents forced fit through the origin.*

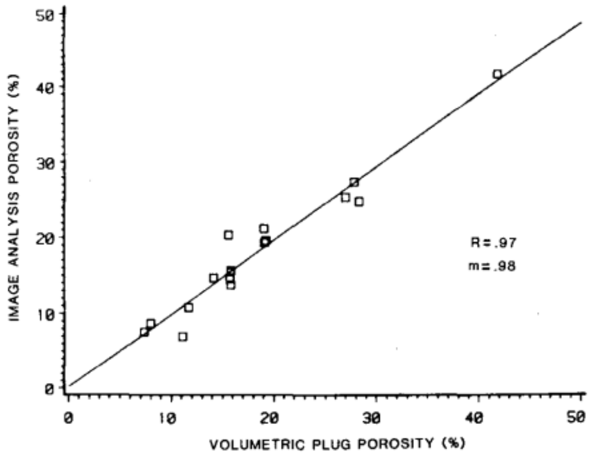
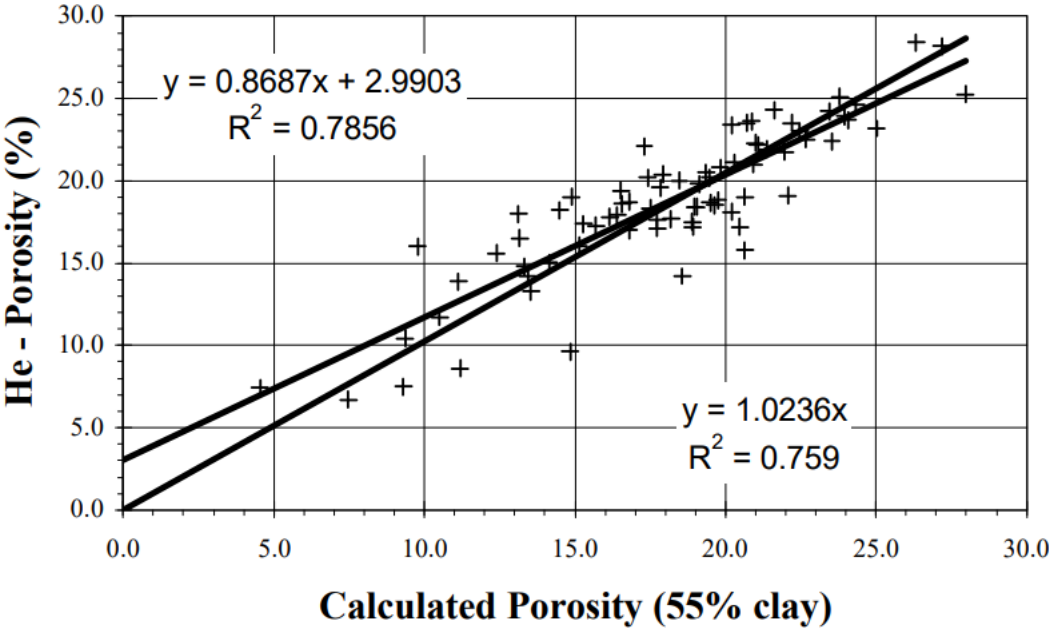


Figure 2-14 *Plot showing a good 1:1 correlation between the porosity estimated by image analysis and that measured in the laboratory.*

To create reliable estimates of porosity from image analysis of BSE images it is important to obtain a better understanding of the key factors controlling the accuracy of the technique. Bøe et al. (1999) showed that a far better estimate of porosity from image analysis was made by assuming that the clay content determined by image analysis included a significant amount (55%) of microporosity (**Figure 2-15**) This example demonstrates a fundamental problem that resin cannot be forced into very small pores and therefore these will not be detected by simply showing a threshold of a BSEM image to highlight resin impregnated pore space.



.

Figure 2-15 *Plot from the same samples showing in Figure 2-13 but in this case it was assumed that 55% of the clay content being assumed to be microporosity (from Bøe et al. (1999). Two regression lines with equation and correlation coefficients are included in each diagram; one of the lines represents forced fit through the origin.*

Different studies agree that the preparation of the samples has significant influence on the quality of data that can be collected and affect the processed outcome. To prevent plucking out of the rock material is conventional to perform the analysis on epoxy-impregnated samples (Ruzyla, 1986; Mowers and Budd, 1996; Solymar and Fabricius, 1999; Cardell et al., 2002). Sample preparation is a key to success in image analysis, and epoxy impregnation serves to enhance discrimination of pore space with the SEM (Ruzyla, 1986). The pioneering work of Ruzyla (1986) focused on using the distribution of pores observed in thin sections of rocks to obtain measurements of porosity, pore size and shape.

The prediction of rock properties by using two-dimensional images, and the microstructural data that can be yielded from them has been a research area of great interest. Different authors have worked to acquire information from two-dimensional images, and from the analysis of the pore space predict different properties, such as porosity and permeability. An example of this, is the work of Lock et al. (2002) where they developed a method for predictions of the hydraulic permeability of consolidated sedimentary rocks, with an absolute value error of approximately 48%, due to the method commonly overpredicting or underpredicting the measured permeability. This could be explained by the fact that the model developed is based in measured attributes of the pore space and how calculations made based on them can introduce a greater error in their predictions.

Pioneers in this research area such as Mowers and Budd (1996), and Solymar and Fabricius (1999) based their estimations of porosity and permeability in empirical relations. Porosity is estimated from the measurements of pore area, and permeability modelled through the derivation of an empirical equation using the Kozeny-Carman expression that relates pore area and specific surface to core permeability (Mowers and Budd, 1996). Closely related, in the work of Solymar and Fabricius (1998) porosity and permeability are determined using Kozeny equation to estimate permeability, and then porosity and permeability obtained from image analysis were compared to results using laboratory methods, as they needed to be supported by information about the pore structure for a better interpretation. Both Mowers and Budd (1996) and Solymar and Fabricius (1999) suggest that the petrographic image analysis cannot resolve microporosity present in the sample, and therefore the imaged porosity is closer to effective porosity.

Studies have been undertaken to predict the mechanical responses of geomaterials using finite element models constructed by digital image processing of microstructural images (e.g. (Yue et al., 2003). The study undertaken by Yue et al. (2003) uses images from a digital camera of the fresh cross-sections, which after analysing non-homogeneous materials under the Brazilian split test, showed that the method can be used to carry out the micro-mechanics analysis with high efficiency and accuracy.

Although using the scanning electron micrographs as a source of imaging is a common practice, Gove and Jerram (2011) developed a method to measure total optical porosity of blue resin-impregnated thin sections, by acquiring the digital images using a conventional 35-mm film scanner. In their work, these images are later analysed using the JPOR macro which is capable of run porosity measurements in a batch of files. Initially JPOR automatically thresholds the image using the default values, but this can be manually adjusted by the user to refine the porosity selection until is consider satisfactory by the user, the next file will be loaded for analysis once the user indicates is ready for the next one (Grove and Jerram, 2011). This indicates the need for an operator to decide the optimal threshold to be used before moving forward. Despite Grove and Jerram's method showed to perform well when evaluated against He injection porosity, the inherent limitations of interoperator variability were also observed, as it introduces a counting error when defining porosity (Grove and Jerram, 2011).

Marcelino et al. (2007) presented a comparison of soil microporosity measurements between three common sample and image acquisition procedures against three 2D image analysis methods with different levels of manual intervention. After comparison of fluorescent and BSE image, they found that BSE images provide much more detail and fluorescent underestimate porosity (Marcelino et al., 2007).

2.3.2. SEM Analysis for estimation of properties from microstructure

Using the Scanning Electron Microscope (SEM) for examining microstructures dates back to the late 40s (Brabazon, 2018) . This technique has been widely accepted by researchers as it allows to analyse samples in terms of visualization, measurements, and composition (Brabazon, 2018; Knaup et al., 2019). Although the Scanning Electron Microscope is considered a key tool to understanding the microstructure of reservoir rocks, is also recognised that the current limitations on image analysis methods can produce significant errors in porosity and pore size distributions (Knaup et al., 2019). Knowledge on the microstructure of a rock aids in the estimation of its geomechanical properties (Zum Gahr, 1987). This is due to the existent dependency between the variation of the properties and the details of the microstructure, which vary extensively as it depends on the properties of the materials (Zum Gahr, 1987).

Using thresholding as part of image analysis is a well-established method for creating binary images (Ruzyla, 1986; Mowers and Budd, 1996; Solymar and Fabricius, 1999; Cerepi et al., 2001; Cardell et al., 2002; Lock et al., 2002; Jurgawczynski et al., 2004; Grove and Jerram, 2011; Suri, 2011; Buckman et al., 2017), and it has been shown that BSE images facilitate the usage of thresholding techniques due to their bimodal grey level histogram nature (Marcelino et al., 2007).

Lock et al. (2002) aimed to develop a method for predicting absolute permeability using only a small number of scanning electron micrograph images, by collecting area and perimeter measurements from BSE images. This approach is based on using the effective-medium theory, where input values of the individual pore conductance are estimated by using the hydraulic radius approximation, which requires knowledge of the cross-sectional areas and perimeters of the pores, in this case estimated by analysing backscattered electron micrograph images (Lock et al., 2002).

Cardell et al. (2002) introduced a method combining image analysis techniques with SEM-EDX and BSE images for identification and quantification of porosity and salts with depth in porous media, which can help to have better understanding of mineral textures and pore distributions in rocks. For the porosity quantification, a cut-off value of 60 for thresholding was selected, considering original images to be 256 grey levels. Quantitative data were obtained relating each area (foreground and background) to the total studied area (Cardell et al., 2002).

SEM produces electronic images which are available for computer manipulation, and such availability enables to obtain quantitative data directly from the image (Webb and Holgate, 2003). However, computers require images to be represented in discrete numerical values in order to process them, therefore these must be converted into digitized grey-scale images, from which pore space is distinguished from the other composing minerals by having a higher grey scale value. Leveraging from this intensity difference, the images are then threshold to yield binary images, in which pores can be distinguished in black, and the mineral grains are white (Suri, 2011).

More recent work has taken advantage of modern SEM's capability for automated image collection of high-resolution images over largescale areas. Buckman et al. (2017) studied BSE images composed from 200 to up to 25 000 tiles per image, which were collected automatically. Thresholding was carried out visually using the default setting and manually adjusting to incorporate the maximum degree of observed porosity, however this method introduces a level of subjectivity into the analysis, as the thresholding process is dependent on the operator's judgment. (Buckman et al., 2017). Despite the potential discrepancies introduced by operator-dependent thresholding, the study was successful in predicting porosity and provided detailed insights into the spatial distribution of porosity in geological samples. The combination of high-resolution imaging, automated acquisition, and batch processing techniques significantly enhanced the reliability and reproducibility of the results.

2.3.3. X-ray Computed Tomography (CT)

X-ray computed tomography (CT) is another powerful imaging technique for porosity estimation, offering the advantage of non-destructive analysis and the ability to capture three-dimensional structures. Taud et al. (2005) developed a grey level method to estimate porosity from CT images, avoiding traditional segmentation techniques that often introduce uncertainties. Their method integrated the image histogram to calculate porosity, providing results comparable to conventional laboratory methods. This approach highlighted the potential of CT imaging in providing accurate and reliable porosity measurements without extensive image processing.

2.3.4. Optical microscopy for porosity estimates

Optical microscopy techniques are essential tools for characterizing porous materials. These methods offer detailed insights into pore structures. This significantly improves our understanding of material properties and performance under various conditions.

2.3.5. Challenges

One of the primary challenges in estimating porosity from image data is the accurate segmentation of pore spaces, especially in heterogeneous rock samples. Traditional segmentation methods, such as thresholding, can be sensitive to noise and variations in image quality. Advances in image processing algorithms, such as machine learning-based segmentation and the grey level method developed by Taud et al. (2005), have addressed some of these challenges by providing more robust and automated analysis techniques.

Another significant advancement is the use of three-dimensional imaging techniques, such as X-ray CT, which allow for the visualization and quantification of complex pore structures in three dimensions. This capability is crucial for understanding the spatial distribution of porosity and its impact on fluid flow properties.

2.4. Data Mining and Statistical Approaches in Petrophysics

2.4.1. Overview of analytical data techniques

Data science is the branch of computer science that involves the extraction of knowledge and insights from data. Artificial Intelligence (AI) can be described as a narrower field of data science, and it refers to the ability of machines to perform intelligent and cognitive tasks (Theobald, 2018).

Machine learning (ML) as well as data mining are subfields of AI. The latter focuses on discovering and revealing patterns in datasets to extract valuable insight from the past, whereas ML focuses on the incremental process of self-learning and data modelling to predict the future (Theobald, 2018); both are closely related fields. Such relationships rely on the fact that self-learning consists of the application of statistical modelling to detect patterns and improve performance based on data as well as on empirical information (Theobald, 2018).

Machine learning comprises three main learning categories: supervised, unsupervised and reinforcement.

Supervised learning:

The most prominent feature of this branch is that it works with labelled datasets. The label or tag of the output is the explanation of its respective input (Mohammed et al., 2017). This means that both the input and output values are known, and they are fed to the machine sample data (Theobald, 2018). The ML algorithm then discovers existent patterns between the input and output values, and after that uses this knowledge to inform further predictions (Mohammed et al., 2017; Theobald, 2018). Following this step, a model based on the rules and trends observed from the data is produced. The model is an algorithm that allows predictions to be made from new data (Theobald, 2018).

Mohammed et al. (2017) explains that the term supervised refers to the fact that labels of the outputs are provided by a supervisor, which can be either a human or a computer. Although, the error rates in data labelled by machines suggest that human judgement is superior, manual labelling is more expensive (Mohammed et al., 2017). Theobald (2018) mentions regression analysis, decision trees, k-nearest neighbours, neural networks, and support vector machines as examples of algorithms used in this learning category.

Unsupervised learning:

The idea of unsupervised learning is to find a hidden pattern or structure in unlabelled data, which might contribute to discover unexpected patterns (Mohammed et al., 2017; Theobald, 2018). Clustering analysis, association analysis, social network analysis and descending dimension algorithms are examples of unsupervised learning algorithms (Theobald, 2018).

Reinforcement learning:

Theobald (2018) describes that the purpose of reinforcement learning is to achieve a specific goal (output) by developing a prediction model through random trial and error. Similarly, Mohammed et al. (2017) points out that the aim of reinforcement learning is to learn from the interaction with the environment which actions would maximize the reward, and which would minimize the risk. This interaction is observed during the iterations of the learning process, where a great number of input combinations is used, and the feedback of their performance is assessed with grades (Theobald, 2018).

In this category, the model is trained through continuous learning. Unlike supervised and unsupervised learning, the outputs obtained are graded instead of labelled (Theobald, 2018). The iterations comprised in reinforcement learning are meant to optimize its future actions, as a result of the gradual improvement of the behaviour and performance of the algorithm based on the acquired experience (Theobald, 2018).

The labelled characteristic of datasets used in supervised learning algorithms shows a resemblance with the type of dataset of properties measured in the laboratory that is available to work with in this research. This manually labelled data which is further described in the document, includes images of the core samples and several values of each of its properties (e.g. porosity). Acknowledging the capability of supervised learning algorithms of analysing the data, retrieving insights and ultimately make predictions, as well as its close relationship with data mining and the advantages that it offers, the implementation of such methods is proposed to support the aim of this study.

Artificial Neural Networks (ANNs)

Artificial Neural Networks (ANN) have structure that can be described in terms of nodes and links, where each node represents a processing unit and the links between them express the causal relationship among them (Kantardzic, 2011). One of the main advantages of using ANNs is ability to learn and generalise, as well as its parallel distributed structure which results in computing power (Kantardzic, 2011). ANNs are flexible because they are highly parameterized enabling them to accurately model irregularities in functions that could be considered as small (Hand et al., 2001). In addition, ANNs offer useful properties and capabilities such as nonlinearity and adaptivity, which allow them to process real-world data that characterises itself for being nonlinear, also, to be designed to adopt parameters in real time according to environmental conditions (Kantardzic, 2011).

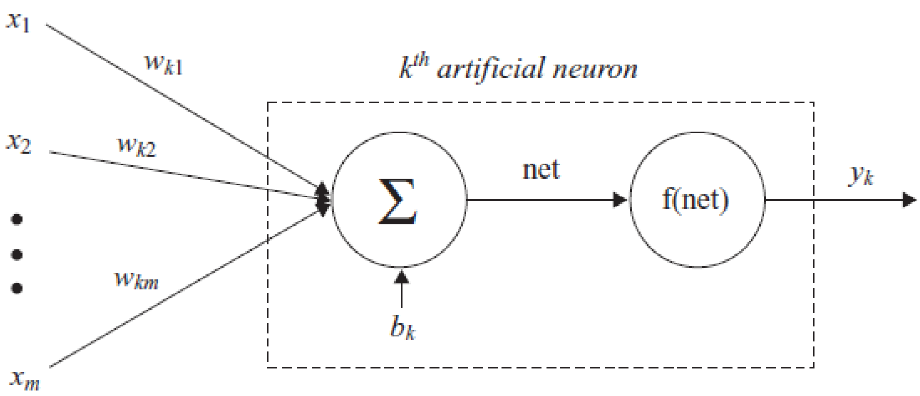


Figure 2-16 *Model of an artificial neuron (Kantardzic, 2011)*

An artificial neuron is the fundamental unit to the operation of an ANN (***Figure 2-16***); it comprises a set of connecting links from different inputs x_i, each of them characterised by a weight or strength w_{ki}. An adder for summing the input signals and an activation function *f* for limiting the amplitude of the output y_k of a neuron (Kantardzic, 2011). ***Figure 2-16*** also illustrates an externally applied bias, b_k, which increases or lowers the net input of the activation function depending on whether it is positive or negative (Kantardzic, 2011).

Some researchers (e.g. Juwaied, 2018; Yousefpour and Fallah, 2018; Li et al., 2019) have studied the performance of ANN models in predicting mechanical properties. Juwaied (2018) concluded that ANNs performed at least as good as conventional AI methods. Similarly, Yousefpour et al. (2018) analysed various instances of applications of machine learning and opted to apply ANNs for the prediction of undrained shear strength. The assessment of its performance showed reasonable accuracy in terms of the prediction of the laboratory test results. Likewise, Li et al. (2019) sought to establish the correlation between effective mechanical property and the mesoscale structure of heterogeneous material, finding a promising accuracy and efficiency of the method developed.

While advanced stochastic architectures like Artificial Neural Networks (ANNs) have been successfully applied to evaluate effective mechanical properties from heterogeneous mesoscale structures (e.g., Li et al., 2019), they introduce significant algorithmic complexity and lack physical transparency. To maintain direct physical accountability, this study utilizes deterministic data mining techniques, specifically power-law and linear regressions grounded in fractal geometry, as a more stable alternative for resolving scale-dependent resolution bias.

Data Mining

Data mining can be understood as the extraction of useful information from data. It is about analysing data present in databases, extracting implicit and unknown information that can be used to solve problems (Witten et al., 2005). It is also defined as the process of discovering patterns, models, summaries, and values from a given collection of data, where this process must be either automatic or semiautomatic (Hand et al., 2001, Witten et al., 2005). An illustrative image of the phases in the data mining process is provided in **Figure 2-17**.It can be seen from this figure that data mining is an iterative process.

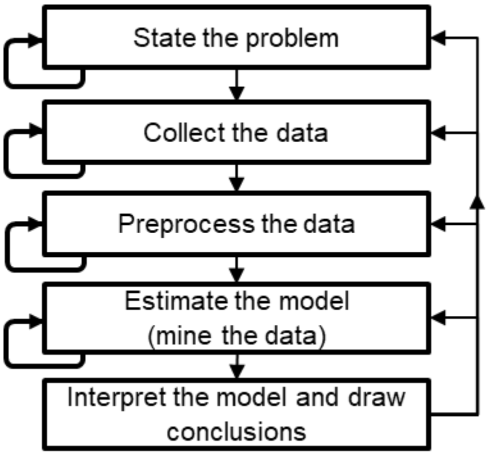


Figure 2-17 *The Data mining process (after Hand et al., 2001, p.9)*

Hand et al. (2001) explains that the process begins with the comprehension of the problem that needs to be solved. The next step is concerned with how the data are collected and how this affects its theoretical distribution (Hand et al., 2001). This understanding is of special interest for the modelling stage and for the interpretation of results (Hand et al., 2001). However, when the data generation process is not under control of the modeller, an observational approach is taken (Hand et al., 2001). In other words, the data is studied and analysed by analytic technique for the purpose of achieving either better or different results; the modeller may go back to the beginning and apply another data analysis tool (Hand et al., 2001).

The two main objectives of data mining tend to be prediction and description. The first attempts to predict unknown values of other variables of interest and ultimately to build a model of the system that will permit the value of one variable to be predicted from variables or information known from the data set (Hand et al., 2001; Kantardzic, 2011). This model is expected to be used to perform tasks such as classification, prediction, and estimation (Kantardzic, 2011). The goal for data description is to gain understanding, discover patterns and relationships, with a view to produce new, nontrivial information (Kantardzic, 2011).

Models produced from data mining aim to enable us to see the main features of the data by obtaining insight about its structure (Hand et al., 2001). Evaluation is the key to making real progress in data mining (Witten et al., 2005). A way to assess the quality of the model is to use it to generate synthetic data, which to be considered good will have the same characteristics displayed in the original data, otherwise if it fails to possess such characteristics it would be considered a poor model (Hand et al., 2001)

2.5. Discussion/conclusions

Determining mechanical properties is key in geomechanics and petrophysics, but obtaining samples for its physical testing is often impractical. This limitation underscores the value of deriving mechanical properties from alternative methods, such as image analysis.

A key factor in understanding the geomechanical properties of rocks is how their mechanical properties correlate to porosity. Porosity, which refers to the amount of empty space within a rock, influences a rock's strength, elasticity, and overall behaviour under stress. Therefore, accurately estimating porosity from images can serve as a critical first step in predicting mechanical properties. Backscattered electron (BSE) images are particularly useful for this purpose. BSE images highlight the differences in brightness between the resin-filled pores and the minerals phases. This contrast in BSE images allows for a more accurate estimation of porosity.

The accuracy of estimating porosity from images has been extensively researched, with studies demonstrating a high degree of precision. Recent studies have shown that BSE image analysis can yield porosity measurements that closely match those obtained from traditional laboratory methods. Which indicates that image analysis can be used as a reliable alternative to traditional laboratory testing methods.

Integrating Machine Learning enhances the potential of image analysis in estimating geomechanical properties. Machine learning algorithms can be trained to recognize patterns and correlations within image data that indicate various mechanical properties.

Chapter 3 Estimation of sandstone porosity from back scattered electron micrographs

3.1. Introduction

Knowledge of the mechanical properties of rocks is important for a range of subsurface applications such as assessing well bore stability, the amount of surface subsidence likely to occur due to fluid withdrawal etc. Often there is not sufficient rock material available to conduct high quality mechanical tests. As an alternative, it may be possible to estimate mechanical properties based on their microstructure. Porosity is known to have a major impact on the mechanical properties of rocks (Wong et al., 1997; Chang et al., 2006). A first step in determining the mechanical properties of rocks from their microstructure is to estimate porosity. The aim of this chapter is to develop a methodology to estimate porosity from images and then to assess the accuracy of the method. Backscattered electron (BSE) images from resin impregnated samples are particularly useful for estimating porosity because the brightness of individual phase is directly proportional to their mean atomic number and the resin impregnated porosity has a much lower mean atomic number than the minerals from which the rock is composed.

The research begins by outlining the methodology applied to estimate porosity from BSE images (**Section 3.2**). The results are then presented (**Section 3.3**) and discussed (**Section 3.4**) before conclusions are made (**Section 3.5**).

3.2. Methodology

A dataset was collected and analysed to assess the accuracy of estimating porosity of samples using BSEM images and the key controls on this accuracy. This was achieved by comparing the results from the image analysis of BSEM images with porosity values obtained from core plugs of the same samples using standard laboratory techniques (i.e. calculated from measurements of grain and bulk volume of the samples). The section begins by describing the samples analysed (**Section 3.2.1**). The methodologies used to measure porosity in the laboratory are described (**Section 3.2.2**). The methods used to obtain BSEM images is provided in **Section 3.2.3**. Finally, the image processing and analysis techniques used to estimate porosity are given in **Section 3.2.4**. This workflow is summarized in **Figure 3-1**.

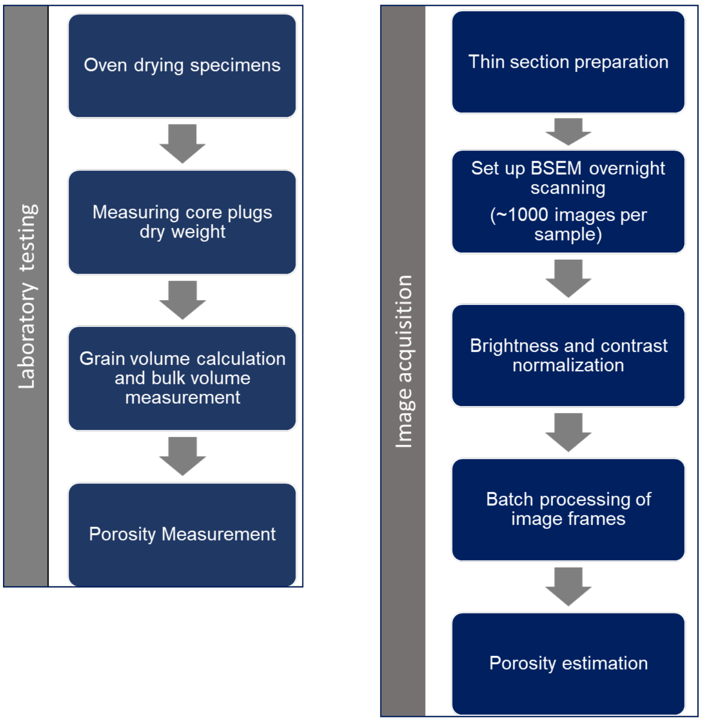


Figure 3-1 *Flow diagram showing methodology for achieving porosity measurement and estimation for further comparison.*

3.2.1. Samples analysed

The samples used in the project were provided by industry sponsors of PETGAS, which is a long-running joint industry project (ran between June 2008 to January 2023) that aimed to improve understanding of the petrophysical properties of tight gas sandstones. The project was led by Prof. Quentin Fisher from the University of Leeds. Around 375 samples from subsurface reservoirs were provided to the project. The samples have a wide geographical distribution, including, Argentina, Australia, Netherlands, Norway, Oman, Republic of Ireland, United Kingdom, and Ukraine. The samples have porosities ranging from <1% to >25% and permeabilities of <0.0001 mD to >500 mD. The samples had already been extensively analysed at the start of the project including the measurement of ambient porosity (**Section 3.2.2**) as well as a small number of moderate resolution BSEM images from each section (**Section 3.2.3**). As described below, the current project dramatically increased the number of SEM images for the database.

3.2.2. Laboratory determination of porosity

Porosity at ambient stress conditions was measured in the Wolfson Multiphase Flow Laboratory at the University of Leeds. It is calculated from measurements of the grain volume, V_g, and bulk volume, V_b, using the formula:

Equation 3-1

$$\emptyset = \frac{V_b - g}{V_b}$$

Grain volume was measured using a helium pycnometer as shown in **Figure 3-2a**. McPhee et al. (2015) explains that this instrument operates on the principal of Boyle’s law, which states that assuming constant temperature for an ideal gas, such as helium, the product of pressure and volume in a closed system remains constant, i.e.

Equation 3-2

$$P_1V_1 = P_2V_2$$

A consequence of this law is that if two containers with volumes of V_1 and V_2 and pressures of P_1 and P_2 are allowed to equilibrate the equilibrium pressure can be calculated using

Equation 3-3

$$P_1V_1 + P_2V_2 = P_3(V_1 + V_2)$$

The schematic for the instrument used is best described in **Figure 3-2b**. The sample is placed in the sample chamber of known volume, V_1, and helium is then allowed to enter the chamber at a pressure of around 20 psi, P_1. After the pressure has equilibrated in the sample chamber a valve is opened to allow gas to expand into an expansion chamber that has a volume of V_2 and is at atmospheric pressure P_2. The system is allowed to equilibrate and the final pressure of P_3 is recorded. The grain volume of the sample can then be determined using: -

$$V_g = V_1 + \frac{V_2(P_2 - P_3)}{P_1 - P_3}$$

The volumes V_1 and V_2 are determined by conducting the experiment using standards.

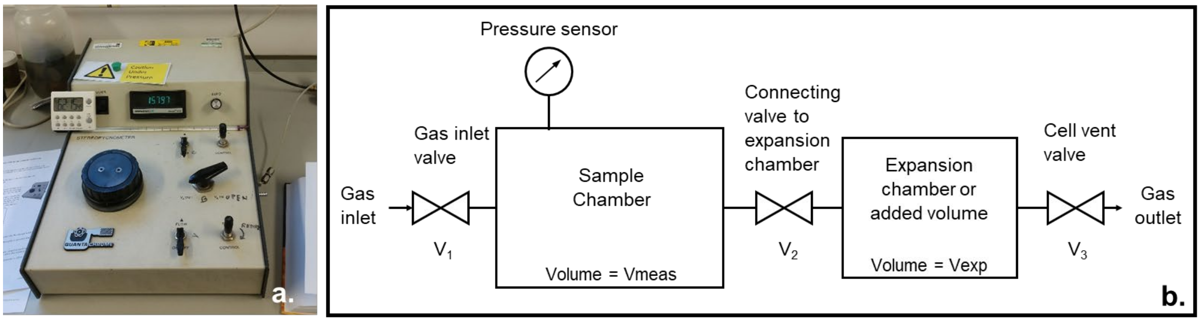


***Figure 3-2** a.) Helium grain volume equipment; and b.) Schematic representation of the instrument.*

Bulk volume of the samples analysed in the first phase of PETGAS (Petgas I) were calculated based on the length, *L*, and diameter, *D*, of core plugs measured by digital calliper using the formula: -

$$V_b = \pi L \frac{D^2}{4}$$

An average of five measurements was taken for *L* and *D*. Unfortunately, this method was found to slightly underestimate V_b due to the imperfect shape of core plugs so subsequent measurements were made using the Archimedes mercury immersion method following the Recommended Practices for Core Analysis.

The mercury immersion method allows for the inclusion of irregular shapes (American Petroleum Institute, 1998); a schematic of the equipment using in this study is shown in **Figure 3-3.**The core plug being measured is placed in a cup filled with liquid mercury, a heavy metal that is non-wetting and does not react with most rocks. A “fork” is then used to submerge the sample 3 to 7 millimetres under the mercury. After removing the fork, the plug is carefully placed on the mercury surface, without touching the sides of the vessel. The balance is tared to consider the weight of the fork and the submerged sample. The fork is inserted again to the same reference mark. The weight displaced by the sample after the immersion is measured. The bulk volume is then calculated using the following equation:

Equation 3-4

$$BV = \frac{Mass\ of\ mercury\ displaced}{Density\ of\ mercury\ at\ measurement\ temperature}$$

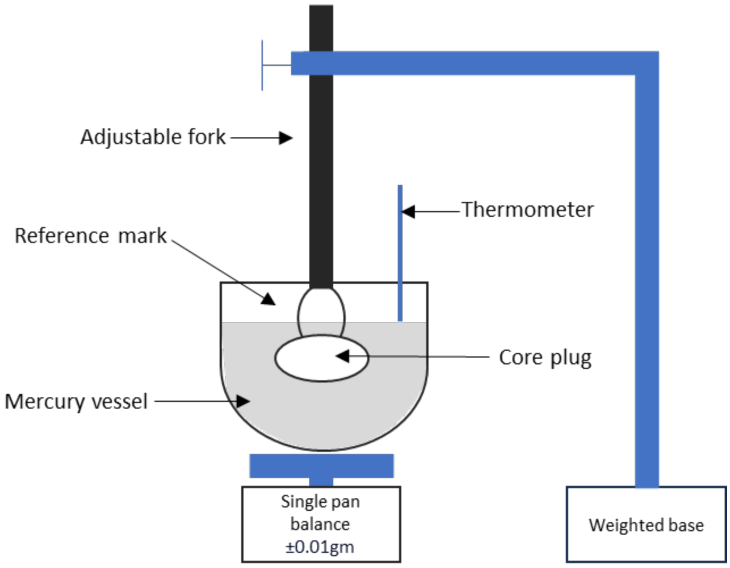


Figure 3-3 *Archimedes mercury immersion apparatus (API, 1998)*

3.2.3. Collection of BSEM images

3.2.3.1. Thin section preparation

Polished thin sections were prepared by Independent Petrographic Services Ltd., which is service company that has 30 years' experience of sample preparation for academia and industry. The sandstone samples were first cut into rectilinear blocks of around 0.5 x 1.5 x 2.5 cm and then cleaned in a Soxhlet extractor using a 50:50 mixture methanol-toluene/dichloromethane. The samples were then dried in a humidity-controlled oven at 60°C before being vacuum impregnated with a low viscosity resin containing blue dye. After the resin had hardened one face of the block was ground down by at least 2 mm and this side was glued to a glass slide. A diamond saw was then used to reduce the amount of rock on the slide to around 0.5 mm thickness before it was ground with successively finer grained diamond until the rock was 30 μm thick at which stage it received a final polish with 0.3 μm aluminium oxide slurry. Finally, the thin sections were coated with a 10 nm thick layer of carbon in preparation for SEM analysis.

3.2.3.2. SEM analysis

The initial sample examination and image collection was undertaken using an FEI Quanta 650 field emission gun environmental SEM (FEGESEM), which has an Oxford Instruments INCA 350 energy dispersive X-ray (EDX) analysed as we as secondary electron (SE), back-scattered electron (BSE) and cathode luminescence (CL) detectors. The sample analysis and image collection were undertaken by Prof. Quentin Fisher for the PETGAS JIP. Between 3-5 representative BSE images were taken of each sample at a magnification of x150 (**Figure 3-4**) as well as high magnification images (**Figure 3-5**) that provide a record of the mineralogy and the diagenetic processes that have impacted the pore structure of the samples. Each image with a magnification of x150 covers an area of around 4 mm^2 of the surface of the sample.

Building upon the theoretical principles of BSEM imaging established in **Section 2.3.1**, the segmentation of pore space in this study was predicated on the contrast produced by variations in the mean atomic number of the constituent phases. It is critical to note for the methodology that BSEM contrast is fundamentally determined by the mean atomic number of the phase being scanned; materials with higher mean atomic numbers appear brighter, while those with lower mean atomic numbers appear darker. This dependence is critical for porosity analysis. Effective segmentation requires resin impregnation of the pore space, as the resin acts as a low-atomic-number medium that provides the necessary contrast against the mineral matrix. In low-permeability rocks, the absence of resin in isolated pore spaces results in these voids lacking the consistent low-atomic-number signal required for thresholding. Consequently, such voids may be misidentified as mineral grains, as the SEM scan captures a flat, polished surface where the pore structure remains effectively 'invisible' to the detector.

To ensure consistency in porosity quantification across the dataset, greyscale levels were standardized using quartz as a reference mineral. Quartz grains were identified manually in each thin section. For each image, the contrast and brightness were adjusted to align the intensity of these identified quartz grains with a predefined reference value. This procedure was essential to distinguish between a mineral that inherently appears dark and darkness arising as an artifact of the SEM's automated image adjustment algorithms.

Later into the PETGAS project, the University of Leeds purchased a Tescan VEGA3 XM SEM. This instrument is capable of taking ultralow magnification images and so was used to obtain single BSEM images of each thin section (**Figure 3-6**) to provides a record of the level of heterogeneity within the sample.

As shall be discussed below, estimates of porosity based on the original images were not particularly accurate so the TESCAN SEM was set up to automatically collect ~1000 high resolution images covering approximately 1 cm^2 of over 350 samples. These images were analysed separately but also stitched together used the TESCAN software to create large high-resolution montages of the sample (**Figure 3-7**).

Images were generally saved in 8-bit tiff format. However, occasionally images were accidently saved as 16-bit bmp format images, so these were converted to 8-bit tiff format at a later stage.

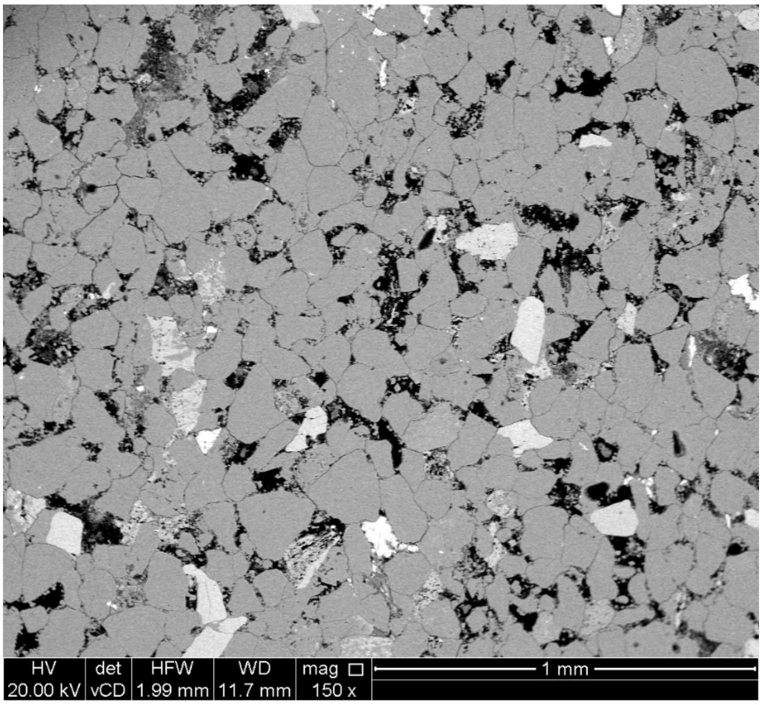


Figure 3-4 *Typical BSEM 150 x magnification image collected for each sample that were initially used to estimate porosity.*

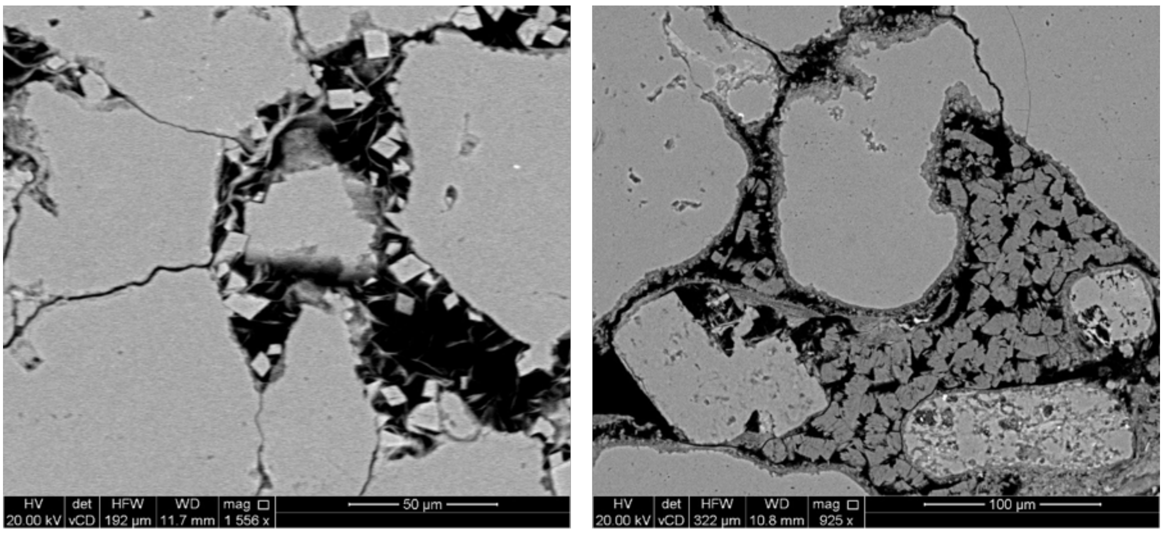


Figure 3-5 *Typical high magnification BSEM images taken to record the key diagenetic processes that have affected each sample.*

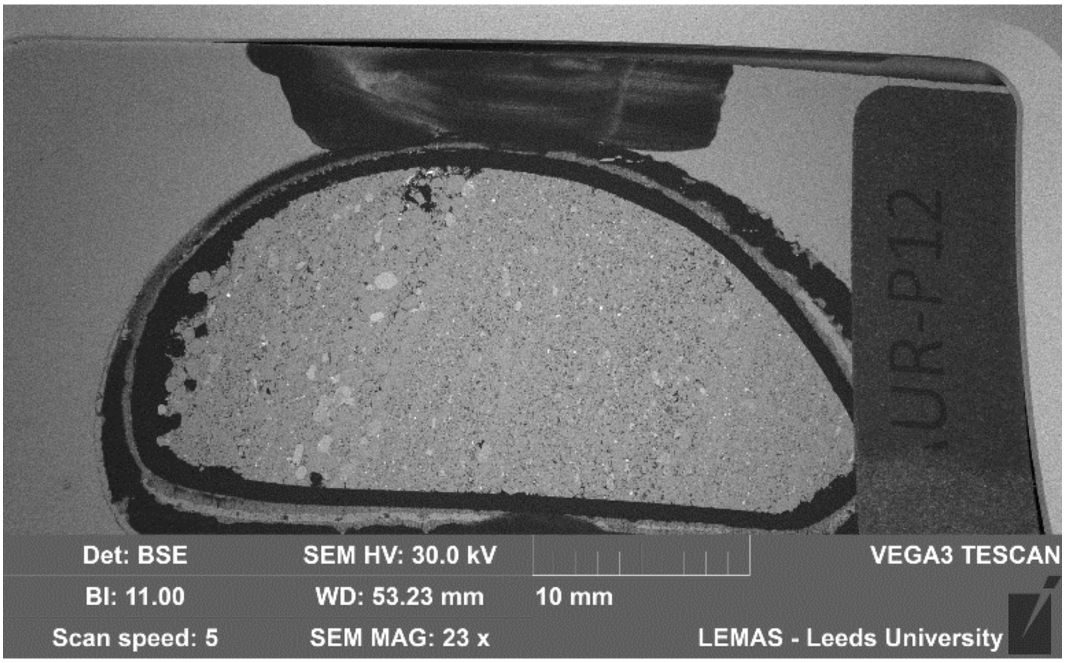


Figure 3-6 *Low magnification BSE image taken to record the level of heterogeneity within each sample.*

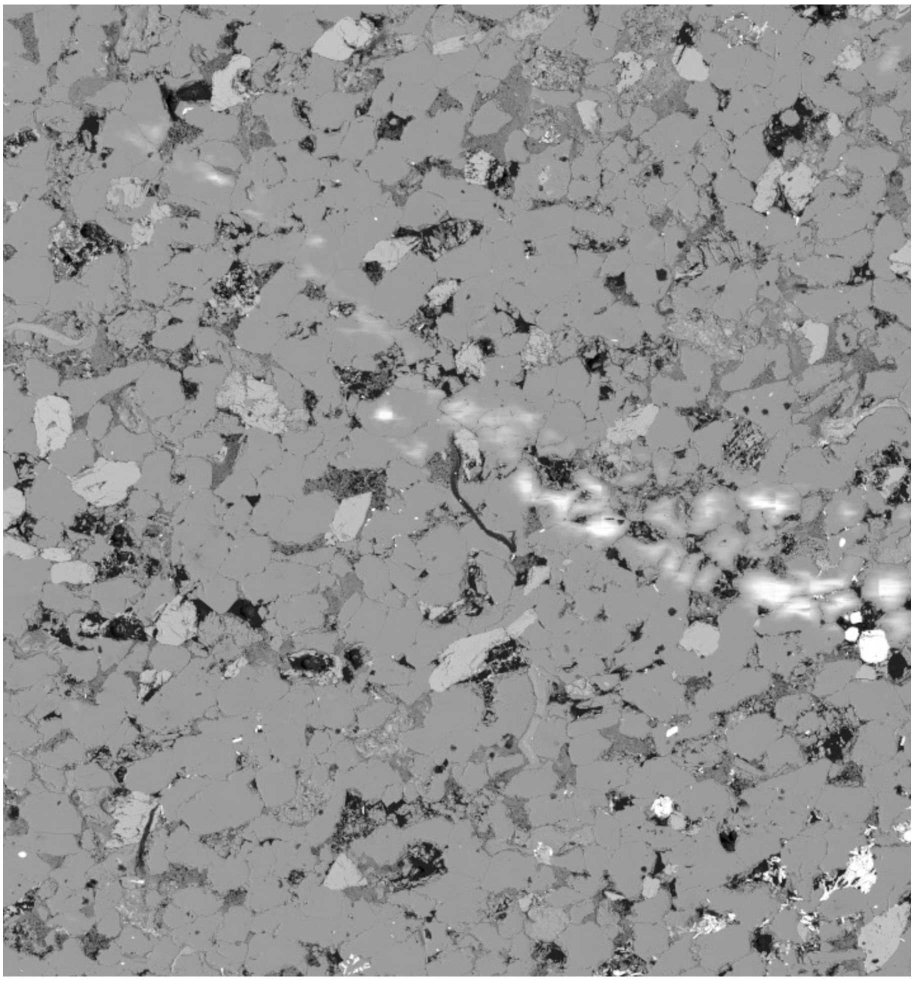


Figure 3-7 *Photomontage created by combining ~1000 high resolution BSEM images.*

3.2.4. Image Analysis

The BSE images were collected with non-uniform brightness and contrast levels. A first step in the research was therefore to normalize the brightness and contrast of each image so that all phases with the same mean atomic number had the same brightness, and the same minerals and porosity had consistent grey levels in each image; this is described in more detail in **Section 3.2.4.1** Once the images were normalized, the porosity was estimated as described in **Section 3.2.4.2**.

3.2.4.1. Normalization of PETGAS Images

Many of the images on the PETGAS database proved to have a 16-bit RGB format rather than 8-bit, so these needed to be converted into 8-bit for further processing. SEM images vary in terms of how details of the images are stored. In some cases, they are stored in a separate file for each image taken, but in others case the metadata are embedded in the image file itself. The latter situation requires the images to be processed in a different way, so the embedded data is not lost when processing the images.

As the original RGB images contained embedded metadata from the SEM (i.e. magnification and user) it was necessary to store such information when the image is being converted from RGB to 8-bit. Existing image processing software, such as Fiji, provides the user with many plugins for image analysis. Early tests showed that this tool did not preserve metadata when used to convert RGB to 8-bit images. Therefore, a script in Python was written that converted RGB to 8-bit images while preserving the metadata in the images. The script takes as input a directory folder containing images of samples, and changes each to 8-bit and writes the original metadata to the resultant file.

A prerequisite for this script is the following inputs:

- File containing sample name and quartz percentage.
- Selected ground truth image.
- Directory containing the images to be processed in tiff format.

The script starts by extracting from the sample-data file the quartz percentage of the ground truth image to analyse. It then obtains an array of the pixel intensities in the image as well as the number of occurrences for each of the intensities as depicted in **Figure 3-8a.**

Considering quartz has a pixel intensity discernible between 40 and 80 % of the pixel intensity values, it selects that subset of pixels (i.e. if the sample goes from 0 to 255 in intensity variation, the subset is 102 to 204 if is from 50 to 225, the subset is 90 to 180). After this is done, it finds the most repeated pixel intensity on that array and that is considered the value for quartz.

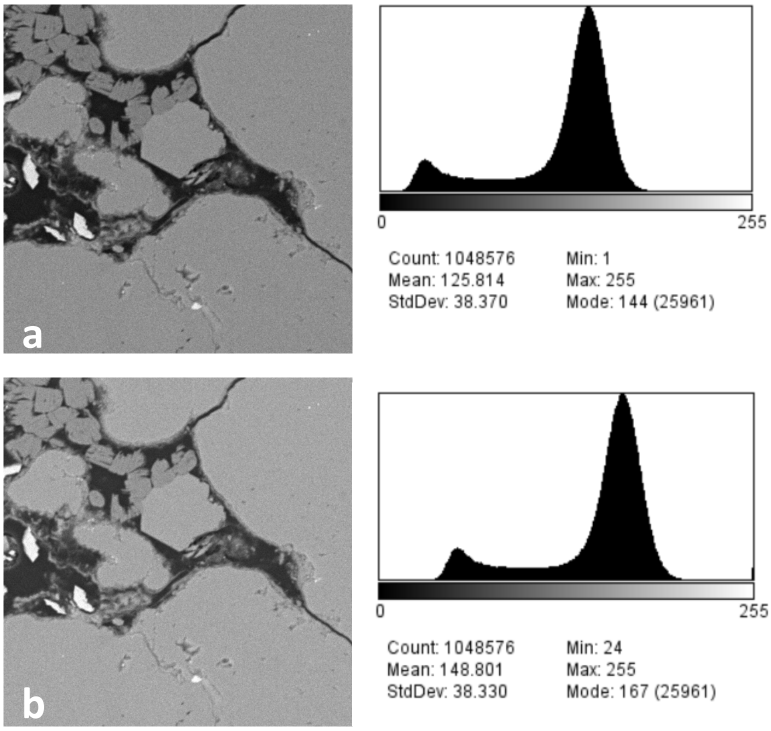


***Figure 3-8** BSEM Image and pixel intensity distribution of a sample a.) before and b.) after normalization*

On the following images (batch) finds the most repeated pixel intensity value (x) and calculates the difference to the ground truth quartz value (Q), then, the image that is being processed (normalised), interpreted as an array of intensities by the computer, gets deducted or added the difference in value to each of the intensities and a new array with these values is created and saved as a new image, as the example in **Figure 3-8b** illustrates.

$$\boldsymbol{new\ image}\,[\quad] = \boldsymbol{image}\,[\quad] - (\boldsymbol{Q} + \boldsymbol{x})$$

3.2.4.2. Estimation of porosity from existing images

Two approaches were followed to determine the porosity present in the samples and its correlation to their available data. The first one, was using the ImageJ software, which offers different algorithms for thresholding allowing image segmentation, which is a widely used technique in pattern recognition applications (Suri, 2011; Sarkar and Das, 2013; Liu et al., 2015; Goh et al., 2018).

An important step in pre-processing in image analysis is separating the foreground out from the background of an image, and this can be done through thresholding (Liu et al., 2015).

The second approach involved the use of a Python script to analyse the 8-bit images. A related point to consider is that 8-bit images can display a maximum of 256 intensities (i.e., grey levels), which illustrated in **Figure 3-9a**. Thus, the Python script allowed the user to input a number between 0 and 255 which will return the estimated porosity using the cumulative value of dark pixels for each of the images, using as boundary the number that was input at the beginning.

In contrast to the approach used with Fiji where the separation between background and foreground of the micrograph is done through a thresholding algorithm provided by the software (**Figure 3-9b**), the Python script functions by calculating the percentage of each pixel level of intensity that is a part of the image under analysis. First, the user sets a value for the intensity level that is between the range of 0 and 255 as the normalized images are in an 8-bit format. Understanding the distribution of grey levels of one of the normalized images by analysing the grey level histogram allows the user to better select a number that is inclusive of the intensity level of interest. Any pixels that remain under this boundary are added cumulatively. In this sense, the boundary established by the user will act as a threshold for separating the background and binarize the image.

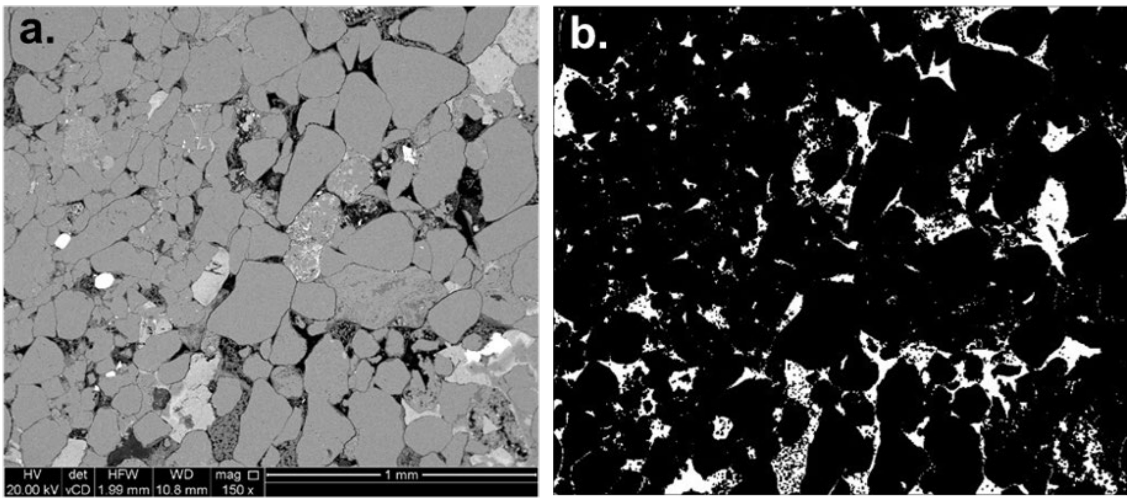


Figure 3-9 *BSEM image displaying a.) Normalized grey levels and pixel intensities; and b.) Binarization after application of threshold algorithm in Fiji.*

To obtain the cumulative area, the first step was to proceed with the discretization of the image to quantify the pixel information. The calculation of the percentage of pixels for each intensity is done by iterating over a list of tuples[1] that relate the brightness and number of pixels contained in each of the images. By accessing these values, it is then possible to calculate the percentage area that corresponds to such pixel intensity levels.

This data is contained in a file that can be read into a spreadsheet. In the far left, column the identifier of the frame that was discretized, in the adjacent columns information such as size in pixels, estimated fraction porosity and estimated percentage porosity. The intensity values are also displayed as a header for each column from 0 to 255, and each of the cells contains the respective quantity of pixels for that given intensity. A column was inserted in this spreadsheet for inspecting the variability of the cumulative area covered by the threshold boundary, this allowed the dynamic calculation of the cumulative number of pixels below a certain boundary, namely the estimated porosity.

The linear regression functions were then used to identify, which grey-level threshold provided the best 1:1 correlation of estimated porosity with the measured porosity value.

3.3. Results

3.3.1. Porosity estimates from the moderate magnification images

In this section, work is presented in which an attempt is made to estimate porosity by conducting images analysis on the moderate magnification BSEM images that were collected prior to the start of this project. The first attempt to estimate the number of darkest pixels in the SEM images was made with Fiji using the histogram-based threshold algorithms offered by this software. The results obtained are shown in **Figure 3-10**. The regression in **Figure 3-10** has three components, a constant, A, a coefficient, B, and a correlation coefficient, R^2. These values are 0.17, -0.47 and 0.042 respectively compared to ideal values of 0, 1 and 1 respectively. The pore results obtained using this procedure meant that it was necessary to use a different approach, namely a Python script, to analyse the 8-bit images.

[1] The term 'tuple' refers to an ordered pair or sequence of elements, often used in programming to represent a composite data structure.

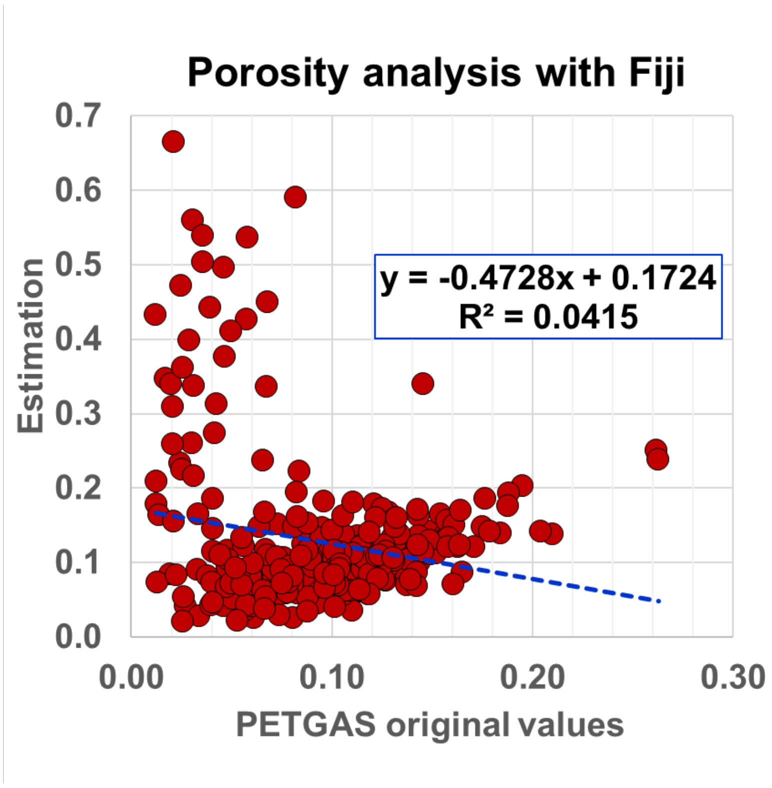


Figure 3-10 *Comparison between porosity estimated from SEM images using Fiji and porosity measured from PETGAS.*

An example of the output for this program is showed in **Figure 3-11**. To improve this approach is necessary to establish the input value that will perform better and produce a best fitted correlation between the estimation and the measured value.

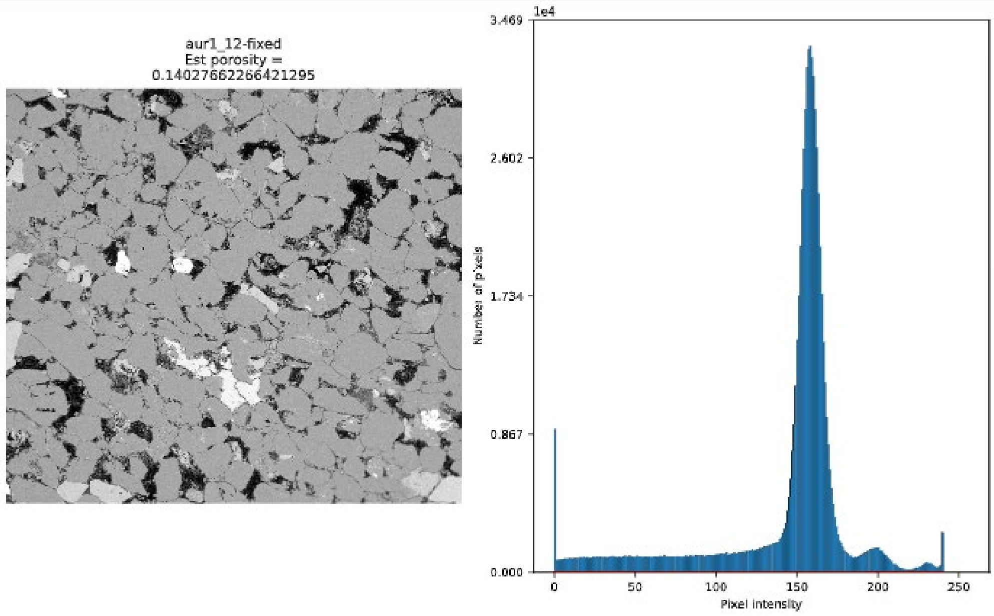


Figure 3-11 *Example of porosity estimation output, showing estimated value and histogram of the image.*

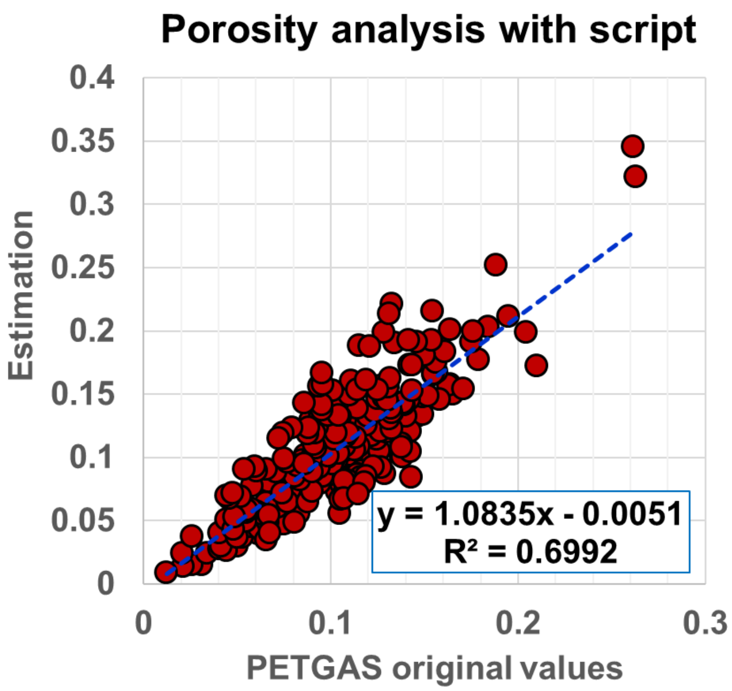


Figure 3-12 *Comparison between porosity estimated from SEM images using script and the values measured in the laboratory.*

After using linear regression functions and trying different boundaries for the collection of images from PETGAS after being normalized, different correlation factors were found for different boundaries parameters that were tested, **Figure 3-12** illustrates the results obtained for setting this boundary at 104 pixels, which is one of the values that showed the best-fit. The porosity estimates from the higher resolution images produced a far stronger correlation with measured porosity values. However, a considerable amount of scatter still existed. For this particular binarization boundary a correlation factor of 0.70 was found, showing what is arguably a strong relationship between the estimation and the measured value of porosity, nonetheless, the percentage error calculated for this data is 21.3%, this lack of precision that can be observed in the outliers of the data suggests inconsistency on the dataset of images, which could be interpreted as lack of homogeneity throughout the sample microstructure. The porosity exhibited in the images does not correctly represent the features of the whole sample. Thus, BSEM images that show a stronger correlation with the measured values are needed. To achieve this, it is proposed to gather high resolution micrographs which enable us to further assess the heterogeneity present in the samples by scanning a larger area of the specimen rather than focusing on arbitrary features.

3.3.2. Analysis of high-resolution scans

A total of 350 samples from PETGAS have been scanned in high resolution, obtaining an approximate of one thousand frames per specimen scanned, which bring us to a total of approximately 300,000 images.

The porosity of each sample was calculated using the Python script described above. Overall, the results (**Figure 3-13**) are far better than the initial results from the moderate magnification images (i.e. **Figure 3-10**) but there is still considerable scatter (R^2 = 0.555), and the constant A and coefficient B are 0.64 and 0.014 respectively. The discussion below assesses the reason for the scatter.

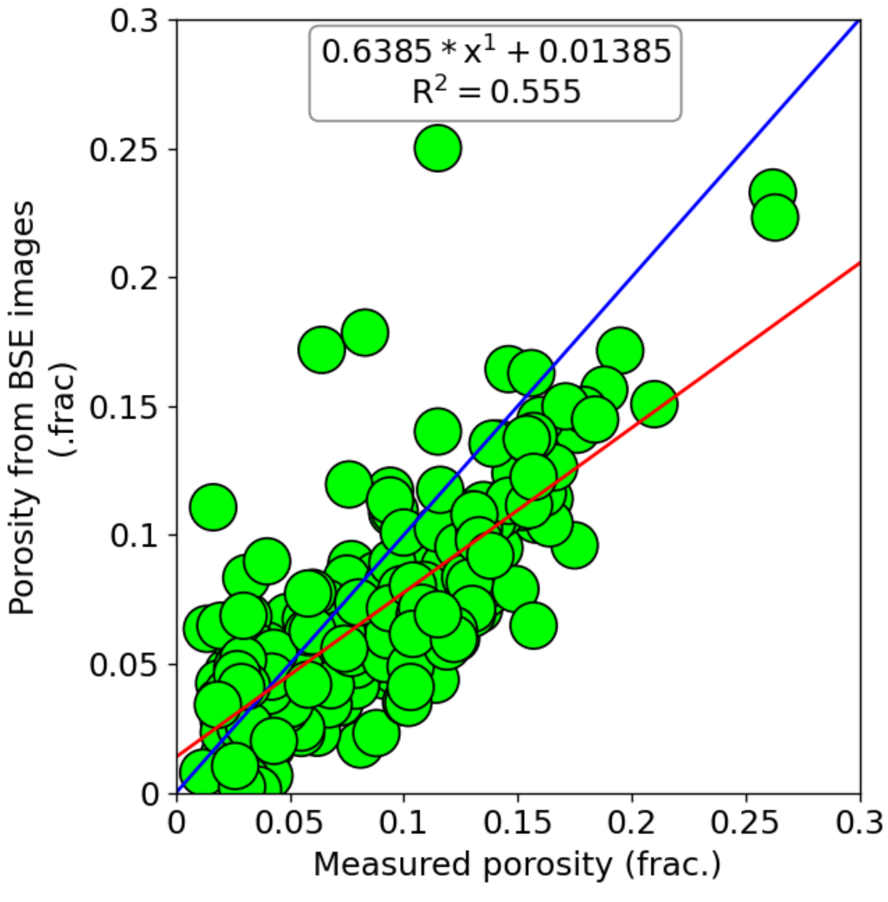


Figure 3-13 *Mean value for porosity per sample compared to PETGAS measured values.*

3.3.3. Limitations of BSEM for Microstructural Images Acquisition

Many of the samples within the PETGAS database show a high degree of heterogeneity in their microstructure, so the small number of images in the PETGAS database was insufficient to illustrate the complexity of the image features across the sample, becoming necessary to collect more, which contributes to the creation of a more robust database.

The expansion of the collection of micrographs has been done using the Tescan Vega 3 SEM. This allowed around 12 samples to be imaged during overnight sessions where the instrument is set to scan continuously throughout a set of samples. Although, this process seems to be straightforward, as the steps taken for setting up the overnight run are always the same, there are a number of factors that influence the quality of the image obtained with this resource including:

- Magnification level and Grain size ratio captured per frame.
- Brightness and contrast settings.
- Human errors such as the inclusion of thin section borders in the scanned area.

An overnight session in this sense, consist in setting up the instrument to work without an expert's supervision throughout several consecutive hours but following a set of parameters that indicate the main characteristics of the scanning that has to be performed. On that note, is worth mentioning that this way of obtaining the images allows to gather a massive number of high-resolution micrographs for each of the samples, segmented in high resolution frames that make it possible to cover a larger area of the surface of the samples than a regular scanning does. Factors such as the number of samples available to be scanned, the area of the sample to be covered, the scan speed, as well as the resolution of the individual frames expected, have to be specified at the beginning of the session in compliance with the number of consecutive hours available for using the instrument.

The Tescan VEGA 3 is a shared resource at the University Electron Microscopy Facility, so being aware of the initial parameters necessary to use it becomes crucial to maximize the number of frames per sample that can be captured within the time limit restrictions.

It is also important to understand that the area selected for scanning will have a great impact on any conclusion that can be draw from this microstructural information. Which is why is relevant to first have an idea on what degree of heterogeneity in the sample we are dealing, so the area selection can be decided based on the inclusion of the variety of features that can be observed on the specimen. Whether this is formed by a range of grain sizes and geometries, a variation of mineral contents, the spatial arrangement that is displayed and how it is distributed, and the amount of space filling that can be observed.

To reduce the irrelevant data that might have been captured with the SEM due to an incorrect set up of the instrument, is important to proceed to perform data cleaning, which is a critical step for the creation of a reliable dataset. The result is an improvement of the quality of the data that will be fed into the regression and extrapolation pipelines.

3.4. Discussion

Numerous outliers are present in the data, which can distort the results. To ensure the accuracy of our analysis, it is key to understand the underlying causes for the deviations from the 1:1 line and eliminate true outliers from the dataset. The PETMiner data visualization software was used to assess the reasons why many of the datapoints did not fall on a 1:1 line. The software allows points to be replaced by images, which often provides clear reasons why the porosity estimated using image analysis differs from that measured in the laboratory. Identification of the reasons for the differences between measured and estimated porosity values provides a robust understanding of the limitations of the method. Removal of genuine outliers from the plot then provides an estimate of the accuracy that may be obtained in close to optimal circumstances.

3.4.1. Causes for inaccurate estimates of porosity from image analysis of BSEM

The following section is divided according to the different processes/reasons responsible for differences between measured porosity and that estimated from BSE images.

3.4.1.1. Sample heterogeneity

Integration of CT images with the position of samples on **Figure 3-15** suggests that sample heterogeneity is a major reason why thresholding of BSE images fails to accurately predict porosity. The reason that this impacts the results is that the polished thin section used for SEM analysis are not representative of the sample as a whole. For example, **Figure 3-14** shows estimated vs measured porosity data from the samples from GDF7 in which the points have been replaced by CT scans. It is clear that these samples are extremely heterogeneous and several of the samples have measured porosities that are far lower than estimated from BSEM images. Removal of the samples from well GDF7 from the dataset instantly improves the R^2 value (**Figure 3-15**). Removing all samples that appear heterogeneous on the CT scans. (**Figure 3-18**) seems to make the results worse, the reason for this is not clear.

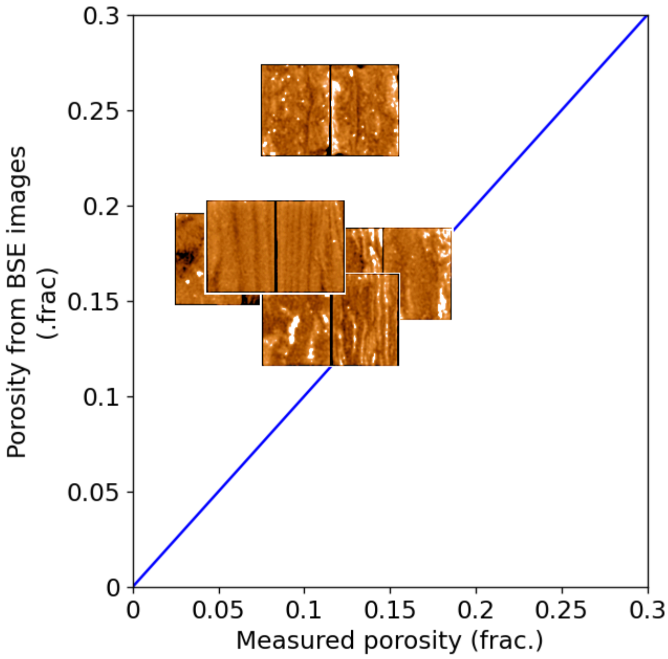


Figure 3-14 *Plot of porosity estimated from BSE images against that measured in the laboratory for the tight gas sandstones from well GDF7 in which points have been replaced by CT images. Note that the samples are very heterogeneous and three deviate significantly from the 1:1 correlation.*

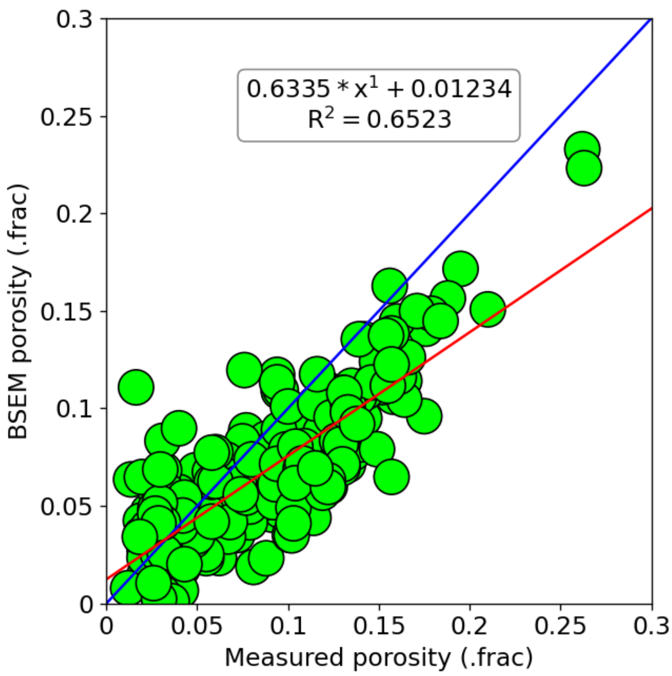


Figure 3-15 *Plot of porosity estimated from BSE images against that measured in the laboratory for the tight gas sandstones analysed during this study with samples from GDF7 excluded.*

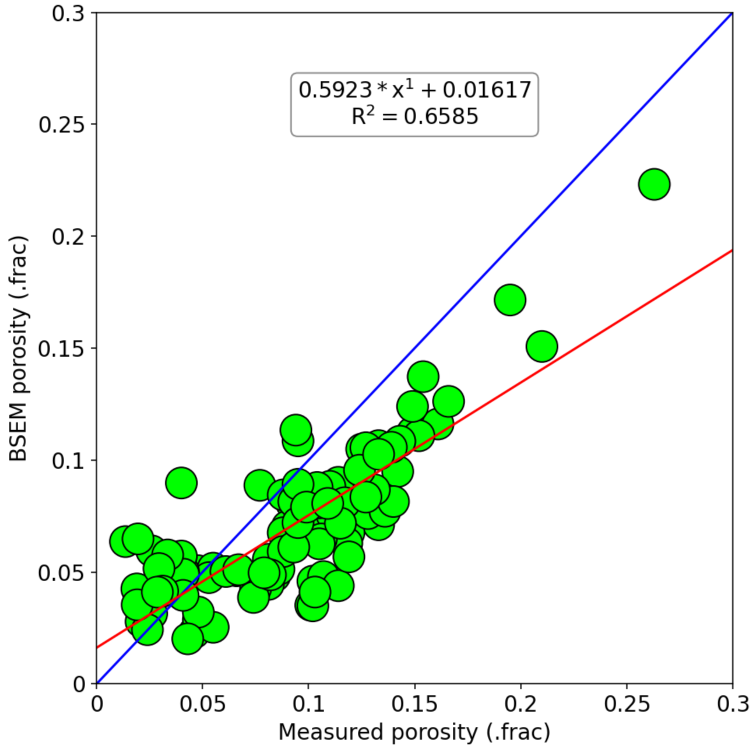


Figure 3-16 *Plot of porosity estimated from BSE images against that measured in the laboratory for the tight gas sandstones analysed during this study with all macroscopically heterogenous samples removed.*

3.4.1.2. Poor labelling of samples

Mistakes in sample labelling can often happen when analysing large datasets. While analysing the causes of scatter in **Figure 3-15** it became clear that thin sections for two of the samples from a particular well were mislabelled. A subsequent review of the laboratory records and imaging logs confirmed that these specimens had been incorrectly identified, leading to inconsistencies between the measured and BSE derived porosity values.

Figure 3-17 Illustrates that removing these mislabelled specimens from the dataset resulted in an improvement in the correlation between estimated and laboratory-measured porosity. This finding underscores the importance of accurate sample identification and precise labelling for maintaining the overall integrity of the dataset and the reliability of any further analysis.

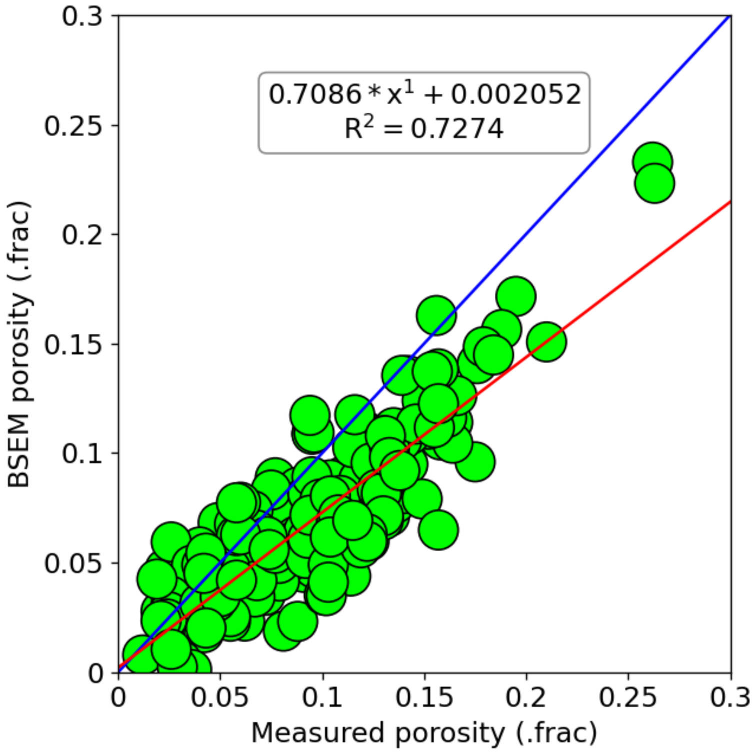


Figure 3-17 *Plot of porosity estimated from BSE images against that measured in the laboratory for the tight gas sandstones analysed during this study two sample removed that had been labelled incorrectly.*

3.4.1.3. Low permeability

The success of estimating the porosity from image analysis relies on impregnating the pore space with resin. This is not possible for samples with extremely low permeability resulting in inaccurate estimates of permeability. For example, **Figure 3-18** only shows data for samples with a permeability of >0.1 mD. The R^2 value has increased to 0.74 for the linear regression, the constant has reduced to -0.007 and the coefficient increased to 0.8.

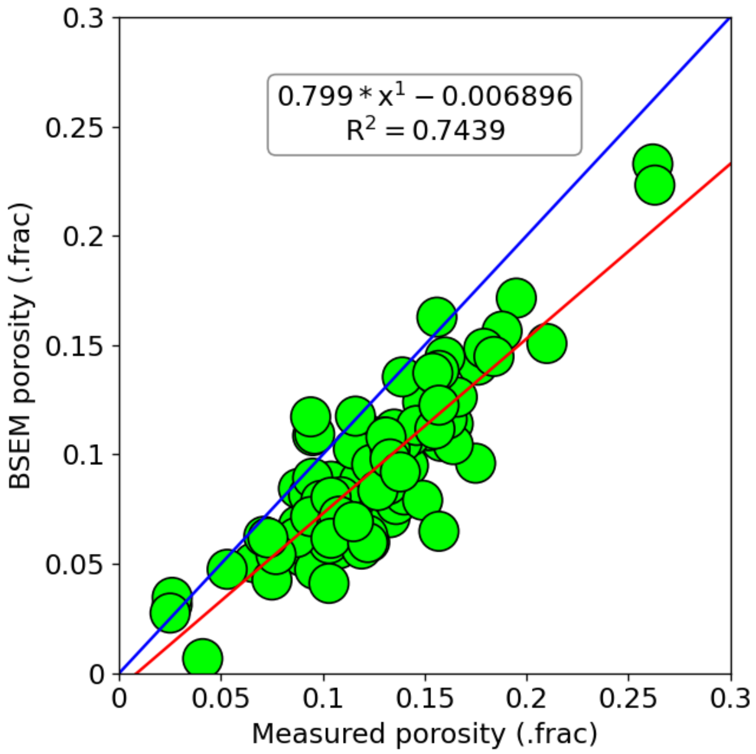


***Figure 3-18** Plot of porosity estimated from BSE images against that measured in the laboratory for the tight gas sandstones analysed that have gas permeabilities at 500 psi confining pressure of >0.1 mD.*

3.4.1.4. Microporosity

Samples can be very permeable but still contain microporosity that cannot be impregnated by resin, which will result in an underestimation of porosity from BSEM images. For example, samples from well EBN5 contain a large amount of glauconite, which contains a significant amount of microporosity that does not get filled by resin during impregnation (***Figure* 3-19**). The resistance of glauconite-rich samples to resin impregnation is attributed to the specific physical and chemical properties of the mineral. Chemically, glauconite is an Fe-rich dioctahedral mica (Bansal et al., 2018). Physically, it typically forms as micro-porous, peloidal aggregates (Triplehorn, 1966; Bansal et al., 2018). Internally, these pellets possess a highly complex, layered architecture composed of homogeneous aggregates of overlapping micaceous flakes (Triplehorn, 1966). High-resolution imaging demonstrates that glauconite pellets feature a pitted texture caused by the random arrangement of micro-platelets, alongside well-developed, slightly sinuous, and sub-parallel lamellae (Bansal et al., 2018). This extremely tight, disordered micro-structure creates highly tortuous, nano-scale intra-pellet pore spaces with a high internal surface area. Consequently, this specific geometry induces a significant capillary resistance that inhibits the entry of the low-viscosity resin, even under sustained vacuum conditions. This results in the resin effectively bypassing the intra-pellet pore space, which remains inaccessible and therefore unrepresented in the BSEM-derived porosity estimates. Removing samples from EBN5 from the dataset increases the R^2 and the coefficient for the linear regression have increased and the constant is closer to 1 (**Figure 3-20**).

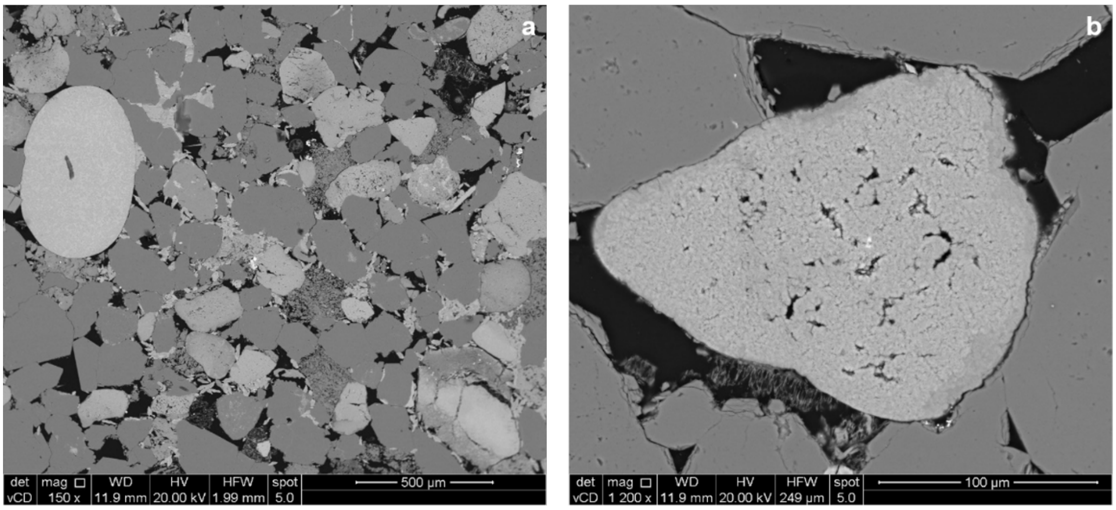


Figure 3-19 *BSEM images from sample EBN5-1. a.) General microstructure with glauconite grains b.) Higher magnification view of the glauconite; note that only larger pores have been impregnated with resin.*

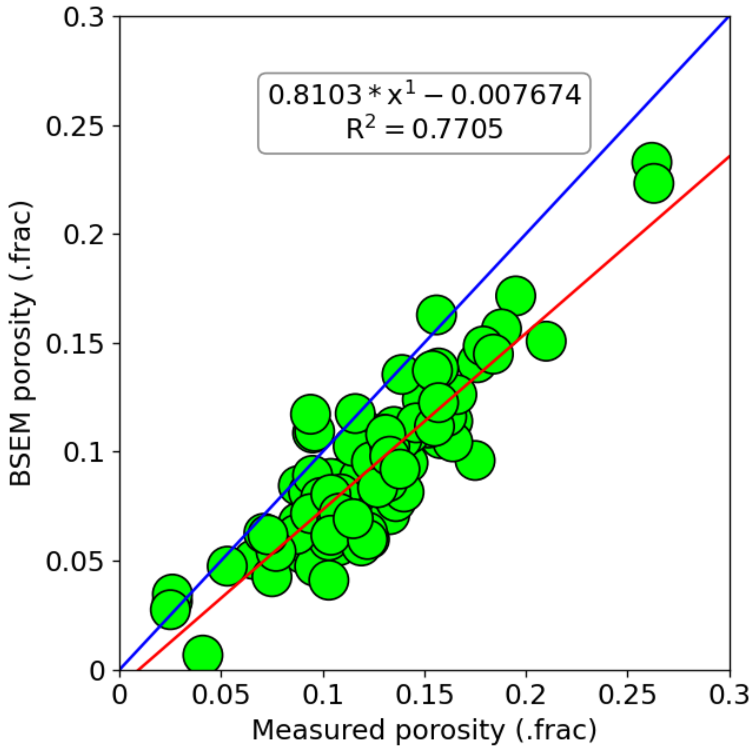


Figure 3-20 *Plot of porosity estimated from BSE images against that measured in the laboratory for the tight gas sandstones analysed that have gas permeabilities at 500 psi confining pressure of >0.1 mD and samples from well EBN5 have been removed.*

3.4.1.5. Other potential causes for discrepancies between measured and estimated porosity

Sections 3.4.1.1 to **3.4.1.4** describe the main causes for differences between measured and estimated porosity identified in the tight gas sandstone database analysed during this study. These potentially mask other sources of discrepancies between results, which are listed below.

Poor sample polishing

Samples that are not polished sufficiently contain “pits” within detrital grains created during cutting or grinding with coarse grid/diamond.

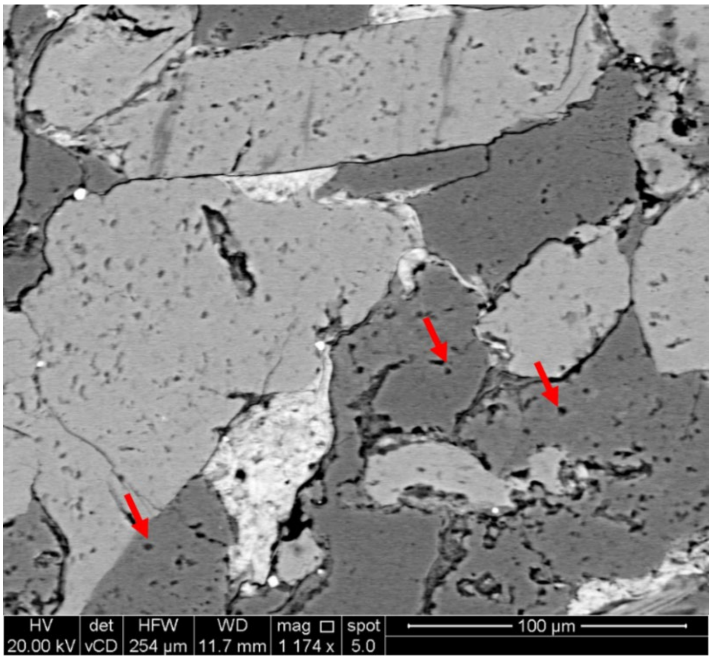


Figure 3-21 *BSEM image of a sample that was not polished properly and contains dark "pits" (arrows) that can be mistaken as porosity.*

Incorrect magnification

It is possible that a single pixel could be contain data from a mixture of minerals and pores, which could be incorrectly assigned to either porosity or mineral depending upon the proportion of each phase present (**Figure 3-22**); this will be discussed in more detail in the following chapter.

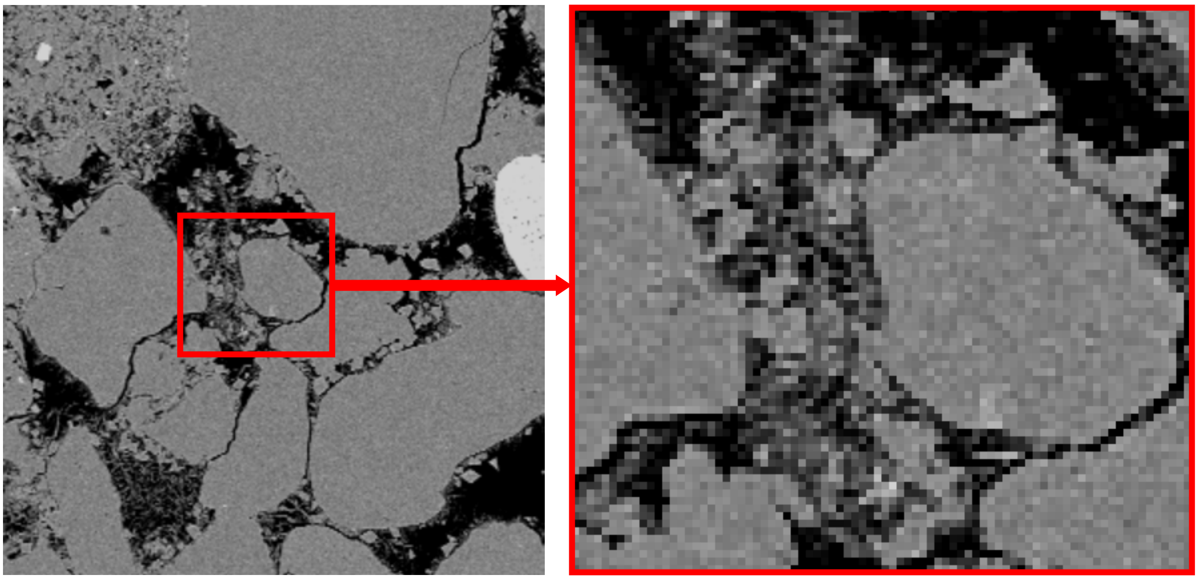


Figure 3-22 *Higher-magnification detail of the red region marked in the left panel. Note that pixels at the grain-pore interface sample both phases simultaneously. This mixed signal introduces significant uncertainty during automated image thresholding.*

Presence of organic matter

Organic matter has a low mean atomic number and is difficult to distinguish from porosity (***Figure* 3-23**). The presence of organic matter could therefore lead to an overestimation of porosity from BSE images.

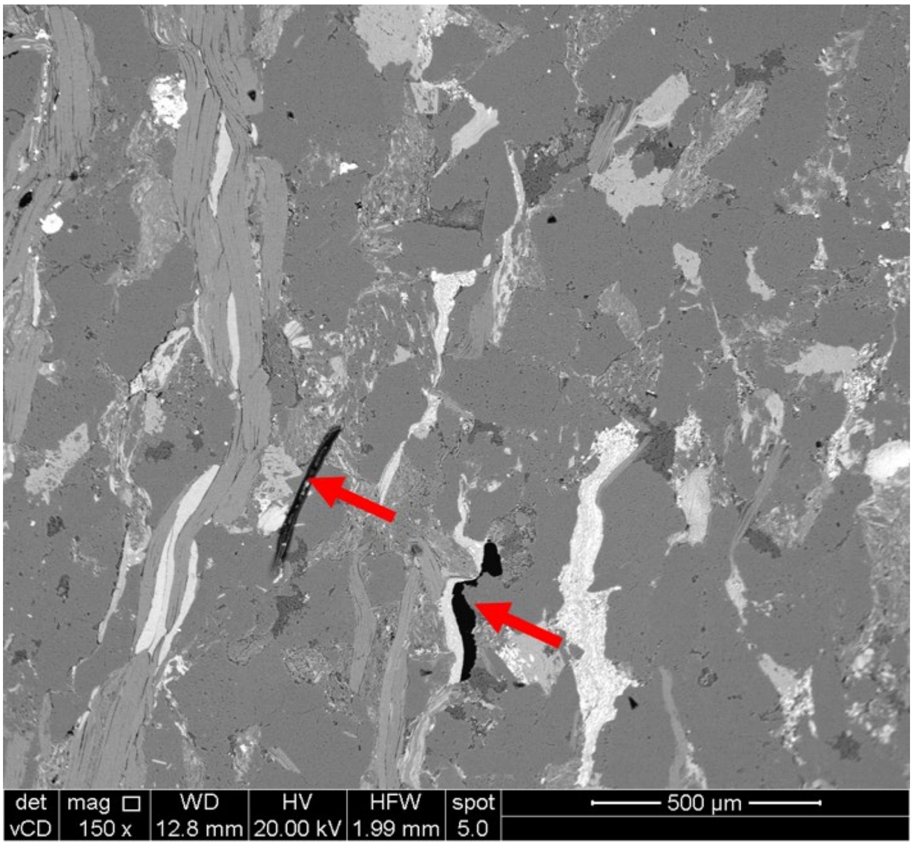


Figure 3-23 *BSEM image of a sample containing organic matter (arrow). Note that it has the same grey level as the resin impregnated porosity.*

A critical distinction must be made regarding the direction of measurement errors observed across the dataset. **Figure 3-20** illustrates a systematic trend of porosity underestimation, which is fundamentally driven by resolution bias. In these tightly compacted fabrics, sub-micron pore networks and clay-hosted microporosity fall entirely below the digital resolution limit, causing void space to be misclassified as mineral matrix.

Conversely, the porosity overestimation highlighted in **Figure 3-25** is a product of sample preparation artifacts and mineralogical anomalies. In these specific samples, pluck-out ‘pits’ generated during polishing, alongside the presence of low-atomic-number organic matter, create artificial dark regions. Because the segmentation script isolates phases based purely on greyscale thresholds, these artifacts are counted as void space, overstating the true core plug porosity.

Therefore, the direction of error is dictated by whether a sample is dominated by sub-resolution micro-features (causing underestimation) or structural preparation damage (causing overestimation).

3.4.2. Comparison with Optical Microscopy

The methodology employed in this study, utilizing high-resolution BSEM-derived porosity, provides a distinct resolution advantage over standard optical microscopy. While optical microscopy of resin-impregnated thin sections is an established petrographic technique for identifying macro-scale grain architecture and porosity, it is fundamentally limited by the resolution of visible light (approximately 1 μm). In contrast, the high-resolution BSEM workflow established here resolves sub-micron pore structures, which are critical for characterizing tight gas sandstones. A formal comparison with optical thin-section data was not conducted in this study due to the unavailability of corresponding optical datasets. While this is acknowledged as a limitation, the methodology serves to isolate the specific impact of resolution bias and scale-dependence in tight gas sandstones. Future work incorporating a dual-approach comparison (where optical microscopy provides a broad view of grain architecture and BSEM provides the necessary high-resolution porosity data) would provide a more comprehensive textural context, potentially strengthening the identification of the Representative Elementary Volume (REV) across varying scales of observation.

3.4.3. Implications for estimating porosity of samples from BSEM images

The porosity estimates for the entire dataset are in error of between around -600 to 100% (**Figure 3-24**). Filtering out data from samples that are heterogeneous and low permeability decreased this error to -35 and 60% (**Figure 3-25**). The implications for these errors in terms of estimating mechanical properties will be discussed later in the thesis. A key issue, however, is whether it is possible to assess whether a reservoir is heterogeneous or low permeability based on the data available. It is possible that heterogeneity could be assessed by analysis of image log data (Alizadeh et al., 2014). It has also been shown that it is possible to estimate porosity-permeability relationships from the microstructure of cuttings (Reyna Flores et al., 2021) and such information could be used to identify whether a sample is likely to be properly impregnated with resin.

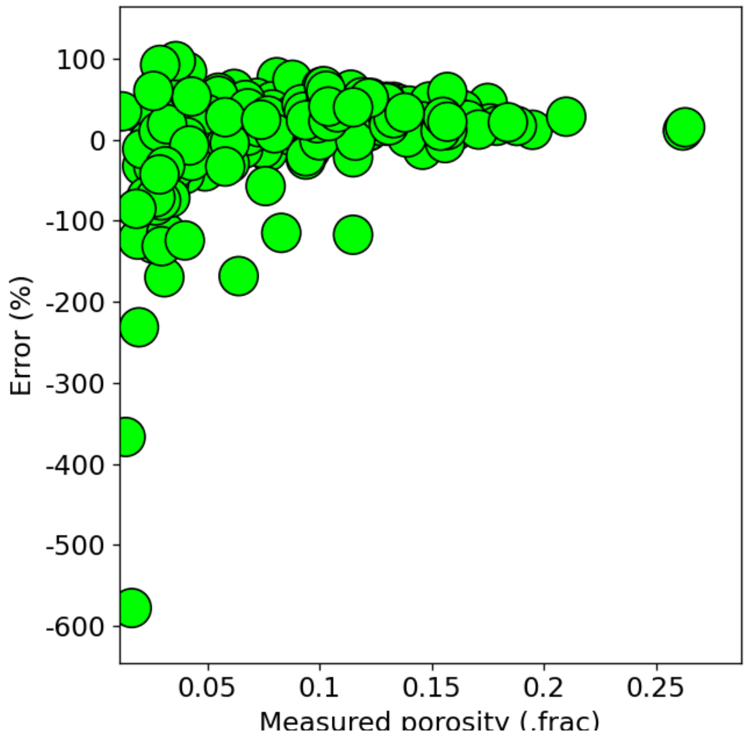


Figure 3-24 *Plot of measured porosity against the error in the porosity estimated from BSEM images for the entire dataset analysed during the current study.*

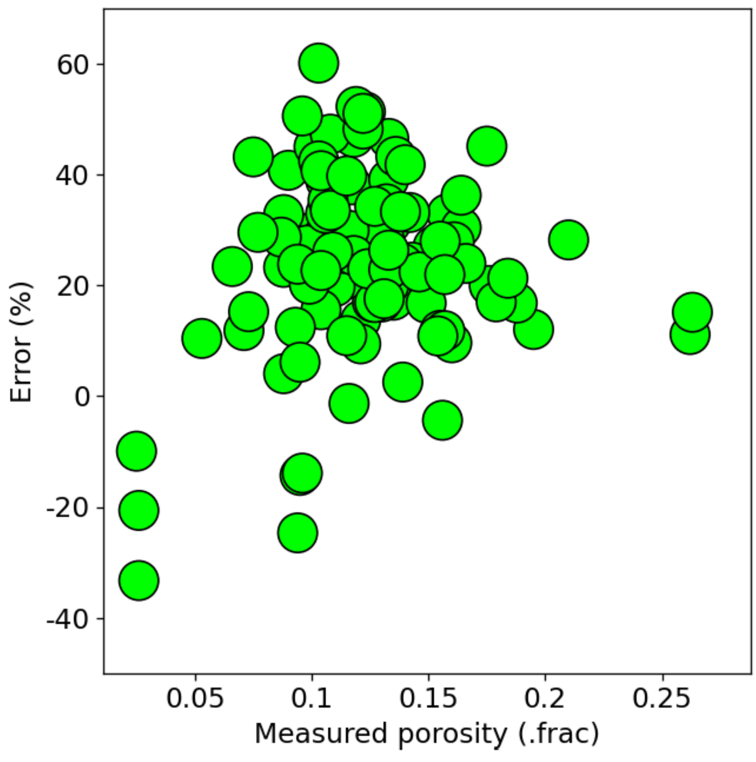


Figure 3-25 *Plot of measured porosity against the error in the porosity estimated from BSEM images with heterogeneous and low permeability samples filtered out.*

3.5. Conclusion

Porosity is the fraction of void space in a rock, determined by the pore space between grains that is not occupied by solid matter. An estimation of porosity from two-dimensional images has been attempted in this document, by evaluating the void space observed in the backscattered electron images, taking advantage of their grey scale bimodal nature.

Geomechanical properties would be ideally measured from controlled laboratory experiments, however the amount of subsurface sampled is often small and it becomes necessary to find alternate methods to obtain this information. Acquiring a high volume of high-resolution backscattered electron micrographs provides access to a massive amount of quantitative data from the microstructure of the sample, which moulds the properties of the rocks analysed.

Opting for using a bespoken script to analyse the SEM images and compare them to the real values from PETGAS yielded a better correlation than using Fiji when analysing the 150x magnification micrographs, when a boundary for thresholding was chosen at 104 pixels. The outliers found suggested inconsistency throughout the dataset of images, indicator of heterogeneity throughout the sample.

It was determined that it was necessary to gather micrographs that covered a larger area of the specimen. After the analysis of the high resolution micrographs for 60 of the samples of PETGAS, 38% of the samples showed a very strong correlation to the measured porosity, however, the number of micrographs that do not accurately represent the known porosity for the samples are likely to be caused by samples scanned with a high magnification that results in a small view field for each frame, where the grain size is larger than each of the frames captured per sample, generating images with focus either on pores or on one particular grain, without the ability to observe both structures on them. Magnification settings as the described above were avoided in further scans to achieve a higher grain ratio per image.

Chapter 4 Impact of image magnification on accuracy of porosity determination from BSEM micrographs

4.1. Introduction

Accurately determining the porosity of rock samples is fundamental in many geoscience and engineering applications because porosity is closely linked to several other crucial properties, such as permeability, mechanical strength, and fluid storage capacity. In fields such as petroleum engineering, mining, and geotechnical engineering, precise porosity estimates are essential for evaluating reservoir quality, assessing structural stability, and planning extraction processes. While laboratory-based methods such as helium pycnometry are efficient and cost-effective for measuring porosity in bulk core samples, they face significant challenges when only small or irregular fragments, such as drill cuttings, are available. In these scenarios, where sample volume is insufficient for traditional displacement or pressure-based techniques, Scanning Electron Microscopy (SEM) provides a critical alternative.

By utilizing high-resolution Backscattered Electron (BSE) imaging, SEM enables researchers to estimate porosity directly from the rock's microstructure, offering detailed visual information that conventional laboratory tests cannot provide. In particular, backscattered electron micrographs (BSEM) provide detailed visual information that can be used to assess pore networks within rock matrices. While traditional core analysis is a non-destructive process, image-based analysis offers a significant advantage when sample volume is too limited for standard laboratory tests. Furthermore, BSEM provides spatial and textural information about pore geometries that traditional bulk measurements cannot resolve. Nevertheless, the accuracy of porosity determination from BSEM micrographs is influenced by several key imaging factors, which need to be carefully considered to ensure reliable results.

As demonstrated in Chapter 3, initial attempts to estimate porosity from BSEM images yielded significant scatter (R^2 = 0.555) and high percentage errors, largely due to sample heterogeneity and the limitations of arbitrary image acquisition. These results highlight the necessity of a more standardized imaging protocol. One of the primary factors identified as affecting these estimates is the magnification level, which directly relates to the resolution of the SEM images.

While high magnification reveals fine details of the pore structure, it can also introduce noise if not properly managed.

Equally important is the scanning speed of the SEM. Results from the previous chapter suggested that inconsistent image quality across large datasets contributed to the observed outliers. Faster scanning speeds can compromise clarity and resolution, while slower speeds, though yielding higher quality images, may not be practical for the high-volume scanning required to overcome the heterogeneity issues identified in Chapter 3. By investigating the relationship between scan speed, resolution, and signal-to-noise ratio, this chapter seeks to establish the optimal parameters required to minimize the inaccuracies found in the preliminary analysis.

This chapter aims to explore how variations in scanning speed and image resolution affect the quality of BSEM micrographs and, consequently, the accuracy of the porosity data extracted from them. By investigating these relationships, the study seeks to identify the optimal balance between scanning speed and magnification settings that yields reliable porosity measurements without compromising efficiency. In doing so, it will also address the compromises between image clarity and processing time, providing valuable insights for researchers who rely on SEM techniques for rock characterization.

Ultimately, the goal of this chapter is to contribute to a more robust understanding of how image acquisition parameters influence porosity determination. By optimizing these factors, we can improve the accuracy of porosity estimates, thereby enhancing our ability to characterise other critical rock properties through data mining correlations. This work not only has implications for laboratory practices but also for field applications where rapid and accurate assessments are essential.

4.2. Methods

4.2.1. Sample Preparation and Imaging Acquisition

The sample preparation protocols for the specimens analysed in this chapter were identical to those detailed in **Section 3.2.3**, involving cutting, resin impregnation, polishing, and carbon coating to ensure microstructural integrity.

The primary dataset consists of the same high-resolution Backscattered Electron (BSE) micrographs acquired using the Tescan VEGA3 XM SEM described in the previous chapter. These original images, often approaching 20,000 pixels in width, provide the necessary greyscale intensity range (where brightness correlates to the mean atomic number) to distinguish between the mineral matrix and the void space.

To specifically investigate the impact of acquisition parameters on data accuracy, additional tests were conducted. These involved re-scanning representative specimens while systematically varying the scanning speed and magnification levels, followed by the normalization and downscaling procedures detailed in the following sections.

4.2.1.1. General SEM Settings and Image Acquisition Parameters

The SEM was operated using the same baseline acquisition parameters employed for the high-resolution scans in the previous chapter. These included a beam intensity of 17 and a working distance of 17-18 mm. Automated scanning sessions were configured to maintain a view field of 200–300 μm per image to ensure a consistent grain-to-frame ratio, preventing the capture of single grains that would bias porosity results. Under these conditions, the acquisition time per frame was approximately 14 seconds, resulting in a total imaging duration of around 35 hours as indicated by the TESCAN VEGA Image Snapper.

Images were acquired in a 1024 × 1024 pixel TIFF format with a 20% overlap between fields, totalling 800–1000 fields per sample. While the core project settings, including Auto signal and Panorama Correlation Matching, were kept identical to those in Chapter 3 to provide a stable baseline, this chapter introduces additional scanning sequences where scan speed and magnification were systematically varied to test their impact on the accuracy of the resulting porosity data.

Parameter	Setting
Microscope model	Tescan Vega3 XM
Beam intensity	17
Working distance	17–18 Mm
Image format	1024 × 1024 pixels, tiff
Field overlap	20%
View field	200–300 µm
Scan speed	4 (Baseline)
Frame accumulation	4 (Live and saved)

Table 4-1 *Summary of standardized SEM acquisition and project parameters used for automated scanning.*

To ensure reproducibility and a consistent baseline for comparison, the SEM acquisition parameters were kept identical to those established in Chapter 3. The specific settings used for all automated scanning sessions are summarized in **Table 4-1**.

While these parameters provided the control group, this chapter focuses on the additional tests where scan speed and magnification were systematically adjusted to isolate their effects on porosity accuracy.

4.2.2. Scanning Speed Methodology

To isolate the influence of scanning speed on image quality and porosity accuracy, additional tests were conducted by systematically varying the speed while holding all other parameters in **Table 4.1** constant. Scans were performed at speed levels 3, 4 (baseline), and 5 to evaluate the impact of acquisition time on the signal-to-noise ratio and the resulting greyscale histograms used for porosity thresholding.

This methodology follows established principles in scanning electron microscopy, where the probe diameter and scan rate determine the volume of the sample from which the signal is averaged (Goldstein et al., 2017). By systematically adjusting the scan speed, this study seeks to quantify the 'noise' introduced into the greyscale histograms, which often obscures the distinction between resin-filled porosity and the mineral matrix.

4.2.3. Image Processing for Resolution and Intensity

Following initial image acquisition and pre-processing, a series of digital image processing techniques were applied to manipulate image resolution and refine pixel intensity levels. These steps were fundamental to investigate the impact of various imaging parameters on the accuracy of porosity determination and to optimize the overall quality of the image data for subsequent analysis.

4.2.3.1. Controlled Image Downscaling Methodology

A central objective of this research was to investigate the influence of image resolution on quantitative pore space analysis, specifically porosity measurements. To achieve this, a controlled and iterative image downscaling methodology was developed and rigorously applied to the cropped BSEM micrographs. The process was undertaken using the ImageMagick tool, which is a powerful open-source command-line software suite renowned for its robust image manipulation capabilities, including high-quality resampling algorithms. ImageMagick was chosen due to its efficiency in batch processing and its flexibility in defining precise scaling parameters, which were essential for managing the large volume of image data.

The resampling utilized a box-averaging algorithm. Each step calculates the arithmetic mean of the greyscale intensities within a 2x2 pixel block to generate a single "super-pixel" in the lower-resolution image.

This method effectively models the physical signal averaging that occurs when the SEM electron beam spot size exceeds the dimensions of individual pore throats (Goldstein et al., 2017).

Unlike conventional downscaling approaches that might reduce each image directly from its original resolution to a predefined set of target percentages (e.g., 75%, 50%, 25%), the methodology employed here used a progressive, stepwise reduction. Specifically, the image dimensions were successively reduced by exactly 50% at each step. This process can be conceptualized as a sequence, as visually depicted in **Figure 4-1**: starting from the original 100% width dimension (designated as 'A'), the image was scaled down to 50% resolution ('B'). Subsequently, the 50% resolution image ('B') served as the source for generating the 25% resolution image ('C'), and so on, continuing to 12.5% ('D'), 6.25% ('E'), 3.13%('F'), 1.56%('G'), 0.78%('H'), 0.39%('I'). The precise impact of these iterative reductions on image dimensions and area is summarized in **Table** 4-2. The downscaling process was concluded at resolution level 'I' because images at this scale consistently yielded dimensions of approximately 100 pixels per side. Further reductions beyond this point were determined to be impractical, as they led to a significant loss of essential image features and overall data integrity, making reliable quantitative analysis unfeasible.

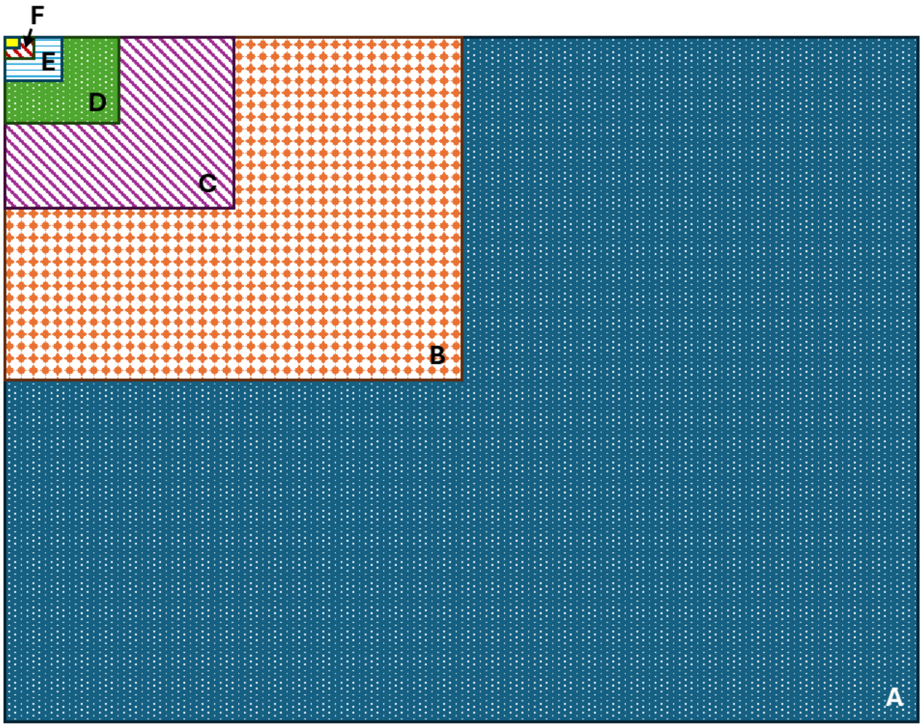


Figure 4-1 *Schematic illustration of the iterative 50% downscaling methodology. (Image A represents the original (100%) resolution micrograph. Each subsequent nested square (B, C, D, E, F) visually demonstrates a progressive 50% reduction in image dimensions derived from the preceding level)*

Label	Percentage of original width (%)	Percentage of original area (%)
A	100.00	100
B	50.00	25
C	25.00	6.25
D	12.50	1.57
E	6.25	0.39
F	3.13	0.098
G	1.56	0.024
H	0.78	0.0061
I	0.39	0.0015

Table 4-2 *Downscaling Resolution Levels and Corresponding Image Dimensions.*

This successive approach was deliberately chosen for its fundamental advantage in preserving the scaling hierarchy. By generating each subsequent resolution level from the immediately preceding one, rather than from the original high-resolution image, the aim was to more accurately simulate the effects of progressively lower optical magnification or reduced digital sampling. The interpolation artifacts are progressively accumulated across successive resolution levels because each lower-resolution image is created by repeatedly shrinking from the immediately preceding version. This compounding build-up of these distortions is intended to more accurately simulate the gradual loss of clarity or detail that occurs with a physical degradation of imaging resolution, such as might be experienced with increasingly less magnification. Furthermore, it helps to isolate the effect of the resolution step itself, as opposed to cumulative resampling from a single high-resolution source, which could potentially introduce different types of interpolation errors at each scale. The specific 50% reduction factor was selected for its clean mathematical progression and its ability to rapidly generate a wide range of distinct resolution states with a relatively small number of steps, thus managing computational burden while providing sufficient data points for trend analysis.

To manage the substantial number of images generated through this downscaling process, a highly organized folder-based structure was implemented. Each specific resolution level (e.g., 100%, 50%, 25%) was assigned its own dedicated subfolder. Within each resolution folder, all corresponding images from different samples were stored. This systematic organization was critical for enabling batch automation. This not only significantly enhanced efficiency by eliminating repetitive manual tasks but also ensured the precise and reproducible application of the downscaling parameters across the entire dataset. The file naming convention for the downscaled versions was also standardized: the highest resolution images retained their original names or were designated with an 'A' suffix, while progressively lower resolutions were appended with 'B', 'C', 'D', and so forth, facilitating easy identification of their relative resolution level throughout the analysis workflow.

4.2.3.2. Gray Level Averaging

Gray level averaging was applied as a noise-reduction technique to mitigate artifacts introduced by pixel intensity variations. This process involves modifying an image’s intensity values by averaging the grey levels of surrounding pixels within a specified region. It is primarily used to smooth images, reduce noise, and improve overall data quality, which enhances the accuracy of pore space segmentation and microstructural analysis. The primary objective was to smooth the image transitions and improve the signal-to-noise ratio, thereby enhancing the accuracy of pore space segmentation. However, care was taken to avoid excessive averaging, which carries the risk of losing fine-scale features critical for porosity analysis in tight gas sandstones.

A comparative analysis of images with and without grey level averaging showed that moderate averaging enhanced pore structure clarity, while excessive smoothing caused porosity underestimation. Optimizing image resolution and grey level averaging is essential for accurate SEM-based porosity assessments.

4.2.4. Greyscale Histogram Extraction and Data Aggregation

Following the meticulous image downscaling, the subsequent crucial step involved the quantitative characterization of the pixel intensity distributions within each image. This was achieved through the extraction of greyscale histograms. A greyscale histogram is a graphical representation of the distribution of pixel intensities in an image, typically ranging from 0 (pure black) to 255 (pure white) for an 8-bit image. In BSEM micrographs, these pixel intensity values are directly correlated with the material's composition; brighter pixels generally correspond to higher average atomic number materials (e.g., mineral phases), while darker pixels typically represent lower atomic number materials or void spaces (pores). Therefore, a greyscale histogram provides a fundamental statistical summary of the constituent phases within the image.

For the purpose of extracting these histograms, a custom macro was written in the ImageJ software. This macro, was specifically designed to iterate through each processed BSEM micrograph at every resolution level. For each image, the macro systematically analysed every pixel, categorizing it into one of the 256 possible greyscale bins (0 to 255) and counting the frequency of pixels within each bin. The output of this process was a separate .csv (Comma Separated Values) file for each individual image, with each row in the file representing a greyscale bin and its corresponding pixel count. This output format was chosen for its universality and ease of parsing by various data analysis software.

The generation of hundreds of individual .csv histogram files necessitated a robust method for data consolidation. To this end, a further macro was developed specifically for data aggregation. The primary function of this macro was to read and parse all the individual greyscale histogram .csv files generated in the previous step and combine their data into a single, comprehensive spreadsheet. This master dataset was meticulously structured to facilitate subsequent analysis within PETMiner. Each row in the aggregated spreadsheet contained key identifiers, allowing for unambiguous data interpretation:

- Sample name: Clearly identifying the geological sample from which the image originated.
- Resolution level: Indicating the specific downscaled resolution (e.g., 100%, 50%, 25%, identified by the 'A', 'B', 'C', (…) and ‘I' labels).

- Greyscale bin: The specific pixel intensity value from 0 to 255.
- Frequency/count: The number of pixels within that particular greyscale bin for that specific image at that specific resolution.

This aggregated dataset was crucial for enabling efficient and accurate ingestion into the PETMiner data visualization and mining software, preventing manual data entry errors and providing a centralized source for all subsequent quantitative porosity analyses. Furthermore, for comprehensive comparison and to provide different perspectives on the data, both raw histogram versions (showing direct pixel counts per bin) and cumulative histogram versions (showing the total proportion of pixels below a given greyscale value) were generated and included in the aggregated data, allowing for flexibility in thresholding and phase differentiation.

4.2.5. Porosity Analysis and Extrapolation to Infinite Magnification using PETMiner

With the aggregated greyscale histogram data in a structured format, the next critical phase involved its ingestion and analysis within PETMiner. PETMiner is a specialized software platform designed for the integration and interpretation of diverse petrophysical data, including core analysis, well logs, and, critically for this study, image-derived properties. Its analytical capabilities are particularly well-suited for handling large datasets and performing complex regressions essential for petrophysical characterization.

The core of the porosity analysis within PETMiner revolved around defining a greyscale cutoff value. In BSEM images, the distinction between pore space (voids) and the solid mineral matrix is often made by applying a threshold to the greyscale intensity. Pixels with intensities below a certain cutoff are typically classified as pore space, while those above are considered solid matrix. The determination of this optimal greyscale cutoff is often an empirical process, and its selection can significantly influence the calculated porosity. To address this, rather than relying on a single, fixed cutoff, a range of plausible greyscale cutoff values (e.g., 80, 81, 83, 85, etc.) were carefully examined. For each image at each resolution, and for each tested greyscale cutoff, the porosity was calculated as the sum of all pixel frequencies falling below that cutoff, divided by the total number of pixels in the image. This approach allowed for a sensitivity analysis of porosity to the chosen threshold.

A primary analytical objective was to extrapolate porosity values to an "infinite magnification" scenario. This concept addresses the inherent limitation of finite image resolution: even at the highest practical magnifications, the image resolution might not be sufficient to fully resolve the smallest pores in the sample. As magnification decreases (and thus resolution worsens), the measured porosity tends to decrease as fine pore details become unresolved or merge with the solid matrix. To quantify this relationship and estimate the "true" porosity free from resolution artifacts, the calculated porosity values were plotted against the inverse of magnification (1/magnification) for each sample, as shown in F**igure 4-2**. Using 1/magnification on the x-axis linearly transforms the relationship, allowing for an easier extrapolation to zero, which conceptually represents infinite magnification.

This methodology is theoretically grounded in the established fractal geometry of porous rocks, as pioneered by Thompson, Katz, and Krohn (1987), who demonstrated that pore networks in sandstones are self-similar across multiple scales. Because the pore network exhibits this scale independence, the relationship between observed porosity and resolution follows a predictable power-law.

Recent research by Luo, Glover, et al. (2020) further supports this approach, demonstrating that when using digital rock models, porosity is frequently underestimated due to the limitations of finite image resolution. Their work confirms that fractal parameters are essential for describing pore structure complexity and that digital image resolution is a primary control on the reliability of petrophysical estimates.

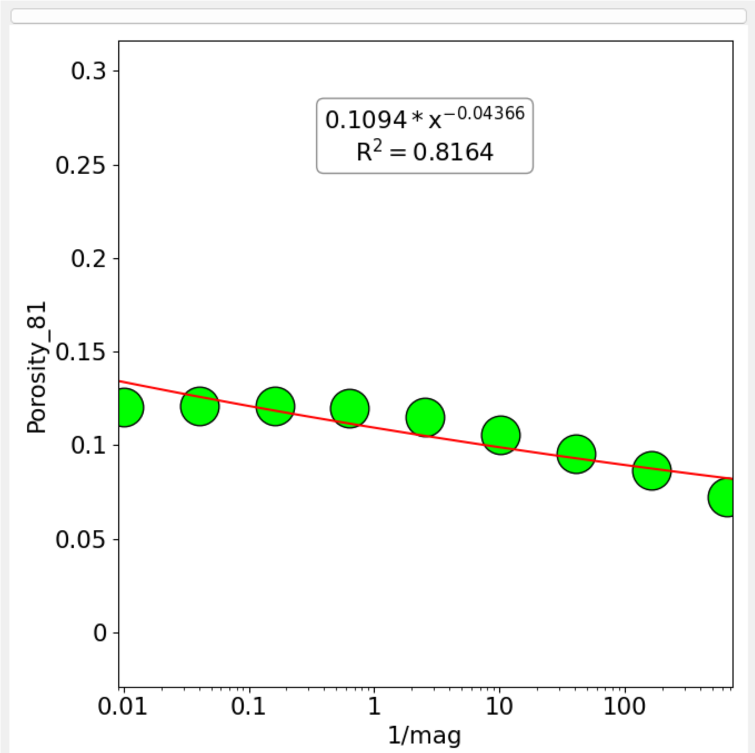


Figure 4-2 *Measured porosity (greyscale cutoff 81) as a function of inverse magnification for aur1_10. This regression provides the basis for extrapolating porosity to infinite magnification (1/mag=0).*

To quantify the trend and perform the extrapolation, a power-law regression was fitted to this porosity vs. 1/magnification plots for each sample. This power-law relationship is dictated by the scale-dependent fractal behaviour of the pore network, taking the general form:

$$\text{Porosity} = A \cdot (1/\text{Magnification})^{B} + C$$

within this formulation, the mathematical fitting coefficients directly map onto the physical fractal scaling properties of the sandstone. Coefficient A acts as a structural scaling constant related to the total volume of the macro-porous network at the upper baseline scale. The exponent, Coefficient B, represents the fractal scaling parameter; it governs the power-law entry rate of sub-resolution features, reflecting the complexity and roughness of the pore-grain interfaces as digital resolution changes. The intercept coefficient, C, represents the asymptotic limit of porosity as 1/magnification approaches zero. This extrapolated baseline provides a resolution-independent estimate of the true pore volume fraction, theoretically eliminating the systematic underestimation caused by pixel-size limitations and capturing sub-resolution features that remained unresolved in the single-frame snapshots from Chapter 3.

4.3. Results

The comparative scans revealed that higher scan speed levels (e.g., speed 5), which correspond to longer acquisition times, resulted in increased image clarity. Conversely, lower scan speed levels (e.g., speed 3) allowed for faster data collection but exhibited increased 'grainy' textures, distortions, and a significant loss of contrast. These artifacts made thresholding-based porosity estimations less reliable, as the noise in the greyscale histogram obscured the clear separation between the mineral matrix and the resin-filled pore space. These results confirm that an optimal scan speed must balance image quality with time efficiency, particularly for high-volume automated scanning.

4.3.1. Impact of Scanning Speed on Image Quality and Porosity Accuracy

The comparative scans across speed levels 3, 4, and 5 demonstrated that acquisition speed is a primary control on the signal-to-noise ratio and the resulting clarity of the pore-mineral boundaries. Higher scan speed levels (e.g., speed 5), which correspond to longer dwell times per pixel, yielded significantly higher image clarity and a smoother greyscale histogram.

Conversely, lower scan speed levels (e.g., speed 3) allowed for rapid data collection but introduced substantial "grainy" textures and periodic distortions in the micrographs. These artifacts led to a loss of contrast that obscured the distinct bimodal nature of the greyscale histogram, which is essential for accurate thresholding.

As noise increases at these faster speeds, the distinction between the dark resin-filled porosity and the grey mineral matrix becomes blurred, often leading to a systematic overestimation or underestimation of porosity depending on the thresholding algorithm used. These results confirm that while baseline speed level 4 provides a viable balance for high-volume scanning, deviations toward

faster speeds (level 3) compromise the data integrity required for the high-resolution analysis conducted in this chapter.

4.3.2. Resolution vs. Porosity Accuracy

To validate the accuracy of the image-derived porosity values extrapolated to infinite magnification, these values were correlated with laboratory measured ambient porosity for the analysed samples. **Figure 4-3**, **Figure 4-4**, **Figure 4-5**, and **Figure** 4-6 illustrate these correlations across different greyscale cutoff values used for porosity segmentation.

This analysis differs significantly from Chapter 3 in two ways:

1. **Removal of Resolution Bias:** Unlike the results in Chapter 3, which utilized fixed-magnification images prone to underestimating fine-scale porosity 5, these results use Coefficient 1 (the intercept from the porosity vs. 1/magnification curves) to estimate porosity at a theoretical infinite magnification.
2. **Improved Correlation:** By using the extrapolated value, the correlation with ambient porosity is more robust than the initial Chapter 3 results. A clear positive linear correlation is observed, indicating that the image-analysis pipeline successfully captures sample variability even when the physical resolution of the SEM is limited.

The blue line in each plot represents the ideal 1:1 correlation, while the red line represents the linear regression fit for each specific greyscale cut-off. In summary, while all tested cut-offs demonstrate a strong positive correlation, lower cut-offs (e.g., 71) provide a slightly better overall linear fit (higher R^2) whereas higher cutoffs (e.g., 80) exhibit slopes that more closely approach the ideal 1:1 relationship. This suggests that higher cutoffs provide better proportionality but carry a potential for increased overestimation at the lowest porosities.

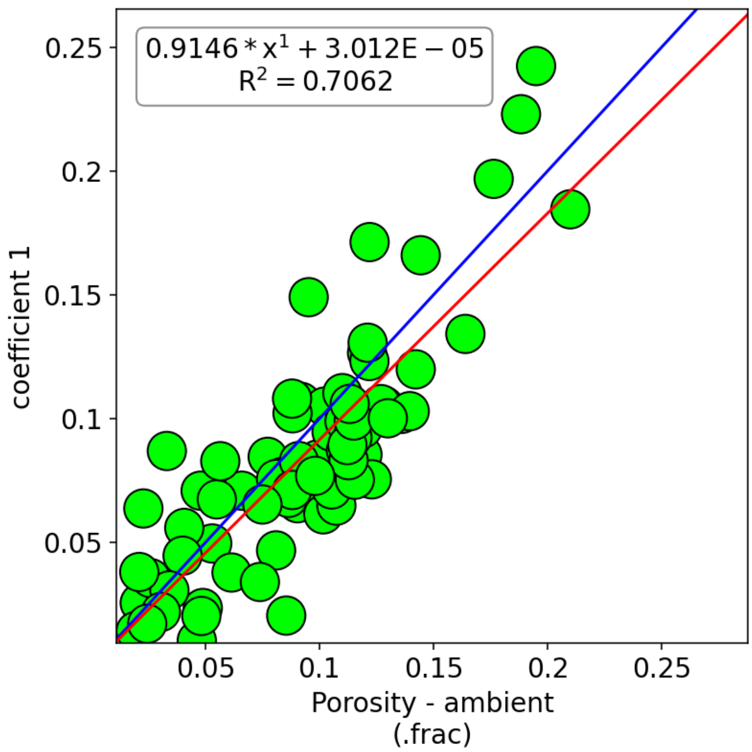


Figure 4-3 *Correlation between Image-Derived extrapolated porosity (Coefficient 1) and ambient porosity for greyscale cut-off value of 71*

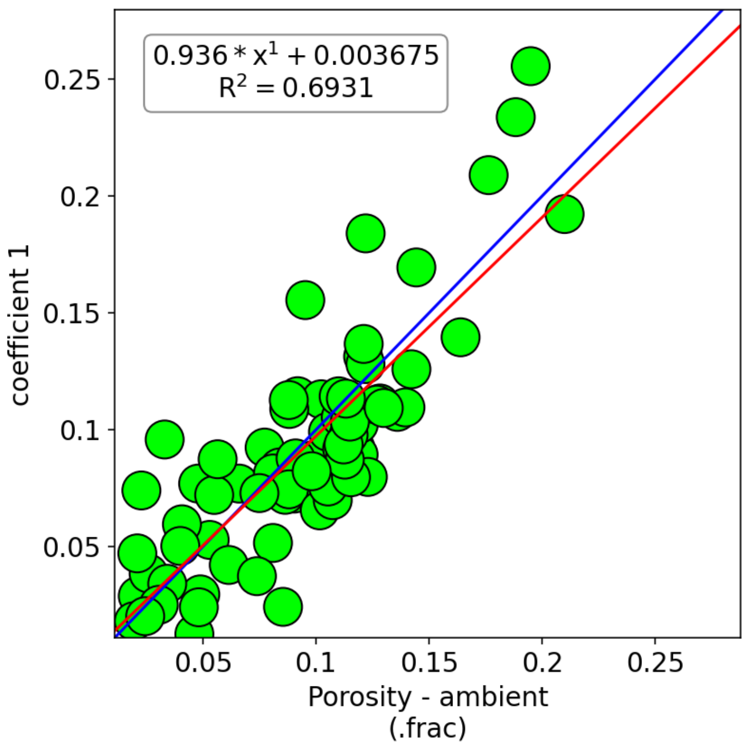


Figure 4-4 *Correlation between Image-Derived extrapolated porosity (Coefficient 1) and ambient porosity for greyscale cut-off value of 75*

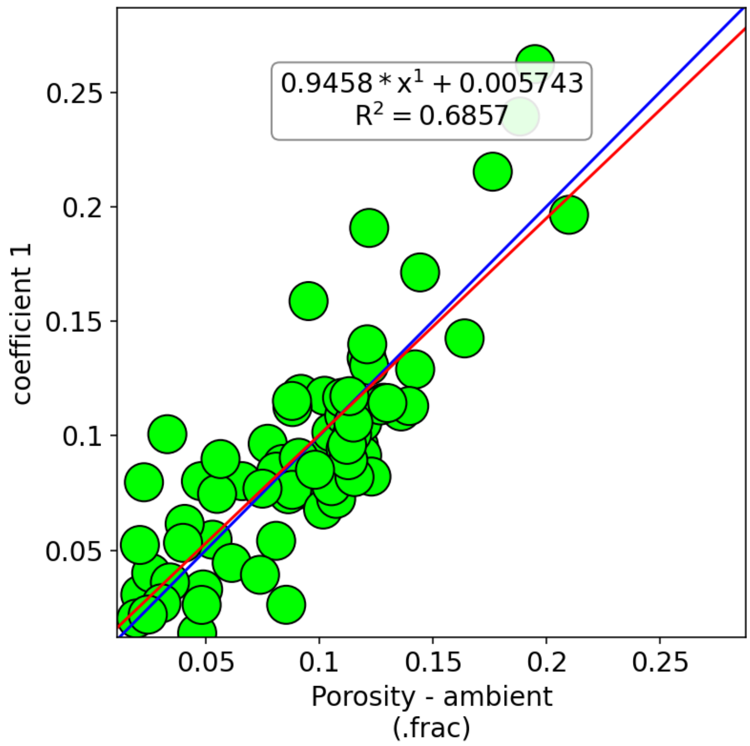


Figure 4-5 *Correlation between Image-Derived extrapolated porosity (Coefficient 1) and ambient porosity for greyscale cut-off value of 77*

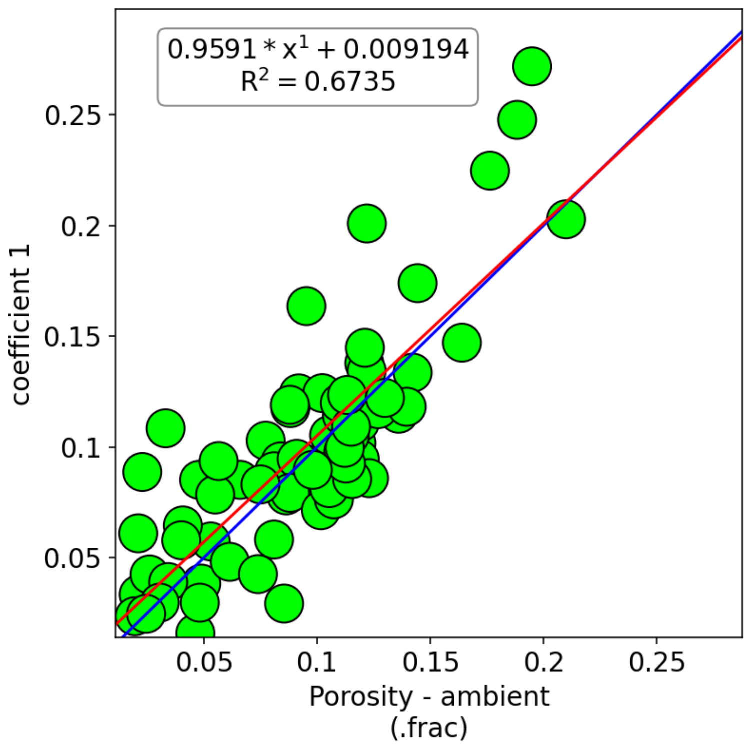


Figure 4-6 *Correlation between Image-Derived extrapolated porosity (Coefficient 1) and ambient porosity for greyscale cut-off value of 80*

4.4. Discussion

4.4.1. Link to Published Work on Representative Element Volume

The successful application of fractal theory to extrapolate porosity highlights a critical trade-off in petrophysical image analysis: the requirement for high resolution to capture fine-scale features versus the necessity of a sufficiently large field of view to satisfy the Representative Elementary Volume (REV). The REV is defined as the volume at which measured properties, such as porosity, become independent of the of the sample size. Failing to reach this scale results in significant fluctuations in measured data, as the volume is too small to capture a representative amount of the rock's internal structural heterogeneity. As noted by Halisch (2013), the REV is only reached when property variations become minimal and converge toward an in-situ value.

Standard deterministic REVs can be difficult to define in heterogeneous media because the required scale varies spatially. Zhang et al. (2000) suggest that for such materials, a 'Statistical REV' (SREV) is a more practical measure, defined as the volume beyond which the mean value becomes approximately constant, and the coefficient of variation (CV) drops below a specific threshold.

In Chapter 3, the observed scatter was largely attributed to sample heterogeneity, while Chapter 4 addresses the resolution-dependent bias through fractal extrapolation, the accuracy of these results remains constrained by the REV.

The methodology in this chapter bridges these concepts by using fractal extrapolation to address resolution-dependent bias while maintaining a large enough initial domain to satisfy the SREV. By calculating Coefficient 1 (infinite magnification) from a base dataset that approaches the rock's representative scale, the resulting porosity estimates provide a more technically accurate representation of the bulk rock fabric than the arbitrary snapshots used in previous attempts.

4.4.2. Limitations and challenges

While the application of fractal-based power-law extrapolation significantly improves the correlation between BSEM-derived data and laboratory measurements, several inherent limitations and challenges persist. These challenges can be broadly categorized into sampling bias and signal processing artifacts.

A critical petrophysical distinction must be maintained between resolution bias and sampling bias within this workflow. The power-law fractal extrapolation framework to a theoretical infinite magnification yields Coefficient 1, which successfully addresses resolution bias by correcting the systematic underestimation of total pore volume that is otherwise dictated by finite pixel boundaries and digital sampling constraints. However, this mathematical correction is strictly resolution-dependent and possesses no capacity to mitigate sampling bias. As established by Halisch (2013), if the initial physical scan area or field of view fails to encompass the Representative Elementary Volume (REV) of the rock fabric, the measured property remains highly sensitive to local spatial variations and becomes fundamentally unrepresentative of the bulk sample. This bounding constraint is explicitly highlighted by the macroscopically heterogeneous sample subsets from well GDF7, where no degree of sub-resolution pixel mathematical manipulation could compensate for an initial scanning matrix that was spatially decoupled from the true macro-scale architecture of the core plug.

In highly heterogeneous tight gas sandstones, reaching a deterministic REV may require scanning areas larger than practically feasible within SEM time constraints. Following the concept of Statistical REV (SREV) proposed by Zhang et al. (2000), there is an upper bound to the range of scales where property statistics remain constant. In samples with significant mineralogical layering or large-scale diagenetic features, the mean porosity may begin to fluctuate again as the "window" of observation moves across different facies, potentially leading to the scatter observed in **Figure 3.13**.

The methodology confirmed that scanning speed is a primary control on the signal-to-noise ratio. A significant challenge remains in balancing the need for high-volume data collection (~1,000 frames per sample) with the requirement for slow, high-quality scans. Faster speeds (Level 3) introduce "grainy" noise that complicates the bimodal histogram nature essential for accurate thresholding. While grey level averaging can mitigate some noise, excessive smoothing risks erasing fine-scale fractal features, creating a new source of inaccuracy.

Finally, a fundamental physical limitation exists: image analysis relies on the successful vacuum impregnation of resin into the pore space. In extremely low permeability samples or those containing isolated microporosity (e.g., glauconite-rich specimens from EBN5), the resin cannot enter the smallest pores. Consequently, even a mathematical extrapolation to "infinite magnification" will only ever reflect the "impregnated" porosity rather than the true total porosity measured by helium pycnometry.

Beyond these impregnation constraints, the 2D BSEM approach is further limited by its inability to distinguish between connected and disconnected porosity. While grey-level thresholding quantifies total unmineralized void areas within a single 2D plane, it offers no volumetric topological context regarding pore throat connectivity in three dimensions. This missing context represents a critical boundary condition for both permeability prediction and geomechanical modelling. Disconnected or isolated intra-granular micro-pores contribute significantly to the bulk rock's elastic compliance and compression behaviour under effective stress fields, yet they provide zero contribution to fluid flow pathways. Consequently, while the workflow captured total porosity area parameters efficiently, the lack of 3D path-connectivity metrics prevents the direct derivation of absolute hydraulic permeability from these 2D pixel statistics alone.

4.5. Conclusions

This chapter has thoroughly evaluated the impact of image acquisition parameters on the accuracy of BSEM-derived porosity measurements. The investigation confirms that both scanning speed and image resolution are primary controls on the reliability of microstructural data extraction. Based on the results of the comparative analysis, the following conclusions are drawn:

- Optimal Scanning Speed: While high-volume automated scanning requires time efficiency, faster scan speeds (e.g., speed level 3) introduce significant "grainy" noise and distortions that obscure pore-mineral boundaries. A baseline scan speed of level 4 or higher is recommended to maintain the signal-to-noise ratio necessary for accurate bimodal histogram segmentation.

- Resolution and Fractal Scaling: The iterative downscaling methodology demonstrated that measurable porosity is scale dependent. By applying a power-law extrapolation based on the fractal nature of sandstone pore networks, it is possible to calculate a theoretical "infinite magnification" porosity (Coefficient 1). This method effectively corrects the systematic underestimation of porosity caused by finite resolution limits.

- Significance of REV: Even with high-resolution imagery and fractal correction, large inaccuracies persist if the initial scan does not encompass a Representative Elementary Volume (REV). The use of large-scale montages (~1 cm^2) is essential to mitigate the effects of microscopic inhomogeneity and spatial variability.

- Technique Improvement: The methodology developed in this chapter represents a significant technical leap over the arbitrary snapshots utilized in Chapter 3. The resulting image-analysis pipeline captures sample variability with higher precision and shows a stronger correlation with laboratory-measured ambient porosity.

Chapter 5 Discussion and Conclusions: Estimation of mechanical properties from image data

5.1. Limitations due to inaccuracy of porosity determinations

As discussed in the following sections, the microstructural inaccuracies identified in Chapter 4 propagate through geomechanical models, impacting everything from reservoir compaction to wellbore stability. These risks are summarized in **Table 5.1**.

The primary finding of Chapter 4 is that image-derived porosity is intrinsically scale-dependent and prone to systematic underestimation without fractal-based power-law extrapolation. While the Coefficient 1 method improves precision, the remaining inaccuracies in porosity fundamentally limit the reliability of subsequent geomechanical predictions.

1. Impact on Elastic Property Estimation: As reviewed in **Section 2.2.3**, elastic wave velocities V_P and V_S are the primary inputs for calculating dynamic Young's modulus and Poisson's ratio. These velocities generally exhibit an inverse relationship with porosity.
   Literature models, such as the Wyllie time-average equation and the critical porosity concept, demonstrate that even a small error in porosity (e.g., 2-3%) can lead to significant miscalculations in rock stiffness. Because porosity determines the "frame" support of the rock, the inaccuracies identified in Chapter 4 (ranging from -35% to 60% in unfiltered datasets) suggest that geomechanical properties derived purely from image-based porosity may suffer from cumulative errors when compared to laboratory triaxial data.
2. Limits on Yield and Failure Prediction: The accuracy of porosity is also critical for predicting failure criteria. In sandstones, the apparent pre-consolidation pressure p^* (the stress required for pore collapse) is highly correlated with the product of porosity and grain size.
   As shown by Wong et al. (1997), if porosity is underestimated due to resolution bias (as seen in **Section 4.3.2**), the predicted collapse pressure will be artificially high, leading to unsafe estimates of reservoir stability. Similarly, the relationship between uniaxial compressive strength (UCS)

and apparent cohesion is mediated by the grain-to-grain bonding magnitude, which is directly influenced by the pore-space distribution captured in BSEM images.

3. Heterogeneity and the REV Constraint: Finally, the geomechanical "Representative Elementary Volume" is often larger than the petrophysical REV for porosity. While Chapter 4 established an "optimal" scan setup for porosity, Zhang et al. (2000) and Vik et al. (2013) highlight that properties like permeability and mechanical strength require even larger sampling domains to account for directional anisotropy and layering. Consequently, the sampling limitations identified in **Section 4.4.2** represent the most significant hurdle in moving from 2D microstructural snapshots to 3D geomechanical field-scale predictions.

Geomechanical Property	Petrophysical Control	Consequence of Porosity Underestimation (Chapter 4 Inaccuracy)
Dynamic Young's Modulus (E)	Inverse relationship with Vp and Vs.	Overestimates rock stiffness, leading to incorrect calculations of reservoir compaction.
Pore Collapse Pressure ($p*$)	Product of porosity (ϕ) and grain radius (R).	Misses sub-resolution porosity, leading to dangerously high estimates of formation strength.
Failure Envelope (UCS/Cohesion)	Grain-to-grain bonding and pore distribution.	Results in an inaccurate Mohr-Coulomb failure criteria, risking wellbore instability.
Dynamic Poisson's Ratio (ν)	Derived from velocity ratios (Vp/Vs).	Porosity errors skew the ratio of lateral to longitudinal strain, impacting hydraulic fracture designs.

Table 5-1 *Impact of Image Acquisition Inaccuracies on Geomechanical Property Prediction*

5.2. Geomechanical Application: Predicting Rock Strength via the Yilmaz (2018) Correlation

To evaluate the engineering utility of the multi-scale, fractal-corrected porosity characterization developed in Chapter 4, the resolution-independent porosity data (Coefficient 1) was integrated into empirical rock strength validation models. While direct laboratory testing remains the definitive standard, alternative predictive frameworks are necessary when bulk core material is restricted or unavailable. Following the geomechanical index testing framework established by Yilmaz (2018), a strong exponential correlation exists between a rock's absolute porosity (n) and its Unconfined Compressive Strength (UCS):

$$UCS = 130.23 \cdot \ e^{-0.079n}$$

Where UCS is the unconfined compressive strength in Megapascals (MPa), e is the base of the natural logarithm, and n represents the absolute porosity percentage (%). This empirical baseline model carries a high correlation coefficient (R^2 = 0.93), confirming that microstructural void space exerts a primary control on macroscopic rock failure mechanics.

Furthermore, Yilmaz (2018) demonstrates that the Core Strangle Index (CSI, measured in kN/cm) can be mapped linearly to predict rock strength via the following predictive model:

$$UCS = 6.7058 \cdot \ CSI + 5.780$$

This predictive framework yields an exceptionally high-performance verification metric (R^2 = 0.95), with a Value Account For (VAF) of 95.6% and a Root Mean Square Error (test RMSE) of 0.04. The application of these values to sample subsets is detailed in **Table 5-2.**

Utilizing the resolution-corrected porosity values derived from the image-analysis pipeline ensures that these geomechanical equations are fed accurate volumetric inputs, eliminating the risk of mathematically inflating rock strength predictions through sub-resolution measurement bias.

By feeding the fractal power-law extrapolated porosity (Coefficient 1) into this empirical relationship, the workflow establishes a functioning geomechanical bridge. It successfully translates 2D pixel statistics into a crucial engineering input for assessing subsurface wellbore failure or reservoir compaction. The practical application of these integrated parameters to our sample subsets is validated in **Table 5-2** below:

Sample ID	Image Porosity (Uncorrected normalised, %)	Corrected Porosity (Coefficient 1, %)	Predicted UCS (MPa)
win 1_11	1.37	2.56	106.40
bp2_4	18.73	13.07	46.37
ebn1_4	13.27	12.66	47.89
bp3_2	8.37	10.60	56.38
ebn2_12	19.90	19.86	27.12

Table 5-2 *Geomechanical Strength Predictions and Sensitivity Analysis Contrasting Baseline Normalised Image and Fractal-Corrected Porosity Metrics*

Geomechanical Error Propagation

This application illustrates the critical necessity of resolving resolution-dependent bias via fractal extrapolation. As established in **Section 4.3.2**, uncorrected, fixed-resolution images systematically underestimate total pore volume because fine-scale pore details fall below the pixel acquisition limit. Due to the exponential form of the Yilmaz (2018) strength relationship, plugging an uncorrected, artificially low porosity value into the model propagates a severe mathematical error, yielding an inflated estimation of unconfined compressive strength. The rock fabric is modelled as structurally stiffer and more competent than its true physical state.

The sensitivity analysis in **Table 5-2** highlights this exact petrophysical trend, proving that the impact of this resolution bias is highly dependent on the rock fabric:

- **In intermediate matrices (e.g., bp3_2):** Failing to account for microstructural fabrics shifts the predicted asset framework by more than 10 MPa (67.22 MPa uncorrected vs. 56.38 MPa corrected), directly demonstrating how easily sub-resolution errors propagate into dangerous overestimations of structural safety.
- **In open, high-porosity fabrics (e.g., ebn2_12):** Conversely, well-sorted networks exhibit a negligible variance between the baseline uncorrected

metrics (19.90%) and Coefficient 1 values (19.86%), as the macro-porous network is already fully resolved at standard operational resolutions.

This variance confirms that the automated data mining pipeline selectively resolves sub-resolution deficiencies only where fine-grained micro-porosity dominates, validating its precision and engineering utility across diverse geological fabrics without artificially distorting well-resolved samples.

5.3. Recommendations for further work

The findings of this study have established a robust baseline for using high-resolution BSEM micrographs to estimate geomechanical properties. However, the identified limitations regarding Representative Elementary Volume (REV) and resolution bias suggest several avenues for future research to enhance predictive accuracy.

While 2D BSEM montages provide high-resolution textural data, they remain inherently limited by "windowing" bias and the inability to capture the true connectivity of the pore network. Future studies should integrate X-ray Computed Tomography (micro-CT) to perform 3D Representative Elementary Volume (REV) analysis. As demonstrated by Halisch (2013) and Vik et al. (2013), 3D imaging allows for the calculation of directional porosity logs and variograms, which are essential for determining the scale at which petrophysical properties become homogeneous in anisotropic sandstones.

The manual selection of greyscale cut-offs remains a primary source of uncertainty in image-based porosity determination. Building upon the review in Section 2.4, future work should implement Convolutional Neural Networks (CNNs) or supervised deep learning algorithms to automate pore-space segmentation. Such models can be trained on datasets where ground truth porosity is established via helium pycnometry, potentially reducing the inter-operator variability associated with traditional thresholding.

To bridge the gap between microscopic images and macroscopic mechanical strength, future research should utilize Digital Rock Physics (DRP) modelling. By performing numerical simulations, such as the Lattice Boltzmann (LB) method, on 3D reconstructed geometries, researchers can assess how fractal pore characteristics influence fluid flow and effective elastic moduli. This approach would allow for a more direct validation of the Coefficient 1 (infinite magnification) values against simulated geomechanical responses.

The current study focused primarily on tight gas sandstones. Given the extreme heterogeneity of carbonate pore classes described by Vik et al. (2013), the fractal extrapolation and REV methodologies developed here should be tested on more complex lithologies. Extending the study to unconventional shales would also provide an opportunity to test the limits of BSEM resolution in resolving organic-matter-hosted porosity, which often falls below the current imaging limits.

List of References

Al-Awad, M.N.J. 2002. Simple Correlation to Evaluate Mohr-Coulomb Failure Criterion Using Uniaxial Compressive Strength. *Journal of King Saud University - Engineering Sciences*. 14(1), pp.137–144.

Alizadeh, M., Movahed, Z., Junin, R., Mohsindd, R. and Alizadehe, M. 2014. How to Determine the Heterogeneity Using Image Logs. *Journal of Advanced Research Design*. 3, pp.23–33.

American Petroleum Institute 1998. *API Recommended Practice for Core-analysis Procedure Second*. American Petroleum Institute, Division of Production.

Bansal, U., Banerjee, S., Ruidas, D.K. and Pande, K., 2018. Origin and geochemical characterization of the glauconites in the Upper Cretaceous Lameta Formation, Narmada Basin, central India. Journal of Palaeogeography, 7(2), pp. 99–116.

Barton, N. 2006. *Rock Quality, Seismic Velocity, Attenuation and Anisotropy*. CRC Press.

Bøe, R., Mjaaland, S., Lieng, E., Holt, T. and Wessel-Berg, D. 1999. Prediction of Petrophysical Parameters in Sandstones By Use of Sem/Image Analysis. *Society of Core Analysts*.

Boone, S.J. 2010. A critical reappraisal of “preconsolidation pressure” interpretations using the oedometer test. *Canadian Geotechnical Journal*. 47(3), pp.281–296.

Brabazon, D. 2018. Chapter 16 - Nanocharacterization techniques for dental implant development In: K. Subramani and W. Ahmed, eds. *Emerging Nanotechnologies in Dentistry (Second Edition)*. William Andrew Publishing, pp.327–354.

Buckman, J., Bankole, S., Zihms, S., Lewis, H., Couples, G. and Corbett, P. 2017. Quantifying Porosity through Automated Image Collection and Batch Image Processing: Case Study of Three Carbonates and an Aragonite Cemented Sandstone. *Geosciences*. 7(3), p.70.

Budhu, M. 2010. *Soil Mechanics and Foundations* [Online]. Wiley.

Cardell, C., Yebra, A. and Van Grieken, R.E. 2002. Applying digital image processing to SEM-EDX and BSE images to determine and quantify porosity and salts with depth in porous media. *Mikrochimica Acta*. 140(1–2), pp.9–14.

Chang, C., Zoback, M.D. and Khaksar, A. 2006. Empirical relations between rock strength and physical properties in sedimentary rocks. *Journal of Petroleum Science and Engineering*. 51(3–4), pp.223–237.

Coates, G.R. and Denoo, S.A. 1981. Mechanical properties program using borehole analysis and Mohr's circle In: *SPWLA 22nd Annual Logging Symposium*. Society of Petrophysicists and Well-Log Analysts.

Dai, N., Vafidis, A. and Kanasewich, E.R. 1995. Wave propagation in heterogeneous, porous media: A velocity-stress, finite-difference method. *GEOPHYSICS*. 60(2), pp.327–340.

Dvorkin, J. and Nur, A. 2001. Critical-Porosity Models In: *Pressure Regimes in Sedimentary Basins and Their Prediction* [Online]. American Association of Petroleum Geologists, pp.33–41.

Fjær, E., Holt, R.M., Horsrud, P. and Raaen, A.M. 2008. *Petroleum Related Rock Mechanics* [Online]. Elsevier Science.

Goh, T.Y., Basah, S.N., Yazid, H., Safar, M.J.A. and Saad, F.S.A. 2018. Performance analysis of image thresholding: Otsu technique. *Measurement*. 114, pp.298–307.

Goldstein, J.I., Newbury, D.E., Echlin, P., Joy, D.C., Romig, A.D., Lyman, C.E., Fiori, C. and Lifshin, E. 2017. *Scanning Electron Microscopy and X-Ray Microanalysis*. 4th edn. New York: Springer.

Goodman, R.E. 1991. *Introduction to Rock Mechanics* [Online]. Wiley.

Grove, C. and Jerram, D.A. 2011. JPOR: An ImageJ macro to quantify total optical porosity from blue-stained thin sections. *Computers and Geosciences*. 37(11), pp.1850–1859.

Halisch, M. 2013. The REV Challenge – Estimating Representative Elementary Volumes and Porous Rock Inhomogeneity from High Resolution Micro-CT Data Sets. SCA2013-069.

Han, D., Nur, A. and Morgan, D. 1986. Effects of porosity and clay content on wave velocities in sandstones. *GEOPHYSICS*. 51(11), pp.2093–2107.

Hand, D.J., Mannila, H. and Smyth, P. 2001. *Principles of Data Mining* [Online]. MIT Press.

Hearn, E.J. 1997. *Mechanics of Materials Volume 1: An Introduction to the Mechanics of Elastic and Plastic Deformation of Solids and Structural Materials* [Online]. Elsevier Science.

Juwaied, N.S. 2018. *APPLICATIONS OF ARTIFICIAL INTELLIGENCE IN GEOTECHNICAL ENGINEERING* [Online]. Available from: https://api.semanticscholar.org/CorpusID:160001401.

Kantardzic, M. 2011. *Data Mining: Concepts, Models, Methods, and Algorithms* [Online]. Wiley.

Knaup, A.S., Jernigen, J.D., Curtis, M.E., Sholeen, J.W., Borer, J.J., Sondergeld, C.H. and Rai, C.S. 2019. Unconventional reservoir microstructural analysis using SEM and machine learning In: *SPE/AAPG/SEG Unconventional Resources Technology Conference 2019*.

Labuz, J.F. and Zang, A. 2012. Mohr–Coulomb Failure Criterion. *Rock Mechanics and Rock Engineering*. 45(6), pp.975–979.

Li, X., Liu, Z., Cui, S., Luo, C., Li, C. and Zhuang, Z. 2019. Predicting the effective mechanical property of heterogeneous materials by image based modelling and deep learning. *Computer Methods in Applied Mechanics and Engineering*. 347, pp.735–753.

Liu, L., Yang, N., Lan, J. and Li, J. 2015. Image segmentation based on gray stretch and threshold algorithm. *Optik*. 126(6), pp.626–629.

Lock, P.A., Jing, X., Zimmerman, R.W. and Schlueter, E.M. 2002. Predicting the permeability of sandstone from image analysis of pore structure. *Journal of Applied Physics*. 92(10), pp.6311–6319.

Luo, M., Glover, P.W.J., Zhao, P. and Li, D. 2020. 3D digital rock modeling of the fractal properties of pore structures. *Marine and Petroleum Geology*. 122, p.104706. https://doi.org/10.1016/j.marpetgeo.2020.104706.

Ma, Y.Z. and Holditch, S. 2015. *Unconventional Oil and Gas Resources Handbook: Evaluation and Development* [Online]. Elsevier Science.

Marcelino, V., Cnudde, V., Vansteelandt, S. and Carò, F. 2007. An evaluation of 2D-image analysis techniques for measuring soil microporosity. *European Journal of Soil Science*. 58(1), pp.133–140.

Mavko, G., Mukerji, T. and Dvorkin, J. 2009. *The Rock Physics Handbook*. Cambridge University Press.

McPhee, C., Reed, J. and Zubizarreta, I. 2015. *Core Analysis: A best practice guide*. Developments in Petroleum Science. 64, pp.1–15.

Mowers, T.T. and Budd, D.A. 1996. Quantification of porosity and permeability reduction due to calcite cementation using computer-assisted petrographic image analysis techniques. *AAPG Bulletin*. 80(3), pp.309–322.

Nur, A., Marion, D. and Yin, H. 1991. Wave Velocities in Sediments In: *Shear Waves in Marine Sediments*. Springer Netherlands, pp.131–140.

Nur, A., Mavko, G., Dvorkin, J. and Galmudi, D. 1998. Critical porosity: A key to relating physical properties to porosity in rocks. *The Leading Edge*. 17(3), pp.357–362.

Reyna Flores, A.G., Fisher, Q. and Lorinczi, P. 2021. Convolutional Neural Networks for the Classification of the Microstructure of Tight Sandstone In: *IPTC*.

Roscoe, K.H. and Burland, J. 1968. *On the Generalized Stress-Strain Behavior of Wet Clays*.

Roscoe, K.H., Schofield, A.N. and Wroth, C.P. 1958. On The Yielding of Soils. *Géotechnique*. 8(1), pp.22–53.

Ruzyla, K. 1986. Characterization of Pore Space By Quantitative Image Analysis. *SPE Formation Evaluation*. 1(4), pp.389–398.

Saraf, S., Singh, A. and Desai, B.G. 2019. Estimation of Porosity and Pore size distribution from Scanning Electron Microscope image data of Shale samples. *Exploration Geophysics*. 2019(1).

Sarkar, S. and Das, S. 2013. Multilevel image thresholding based on 2D histogram and maximum Tsallis entropy—a differential evolution approach. *IEEE Transactions on Image Processing*. 22(12), pp.4788–4797.

Schofield, A. and Wroth, C. 1968. *Critical State Soil Mechanics*.

Sjogren, B. 1984. *Shallow refraction seismics*.

Solymar, M. and Fabricius, I.L. 1999. Image Analysis and Estimation of Porosity and Permeability of Arnager Greensand, Upper Cretaceous, Denmark.

Suri, Y. 2011. Predicting petrophysical properties using SEM Image In: *SPE All Days*, pp.1–12.

Taud, H., Martinez-Angeles, R., Parrot, J.F. and Hernandez-Escobedo, L. 2005. Porosity estimation method by X-ray computed tomography. *Journal of* Petroleum Science and Engineering. 47(3–4), pp.209–217.

Triplehorn, D.M., 1966. Morphology, internal structure, and origin of glauconite pellets. Sedimentology, 6, pp. 247–266

Thompson, A.H., Katz, A.J. and Krohn, C.E. 1987. Fractal Sandstone Pores: Implications for Conductivity and Pore Formation. Physical Review Letters. 58(20), pp.2131–2134.

Toksöz, M.N., Cheng, C.H. and Timur, A. 1976. Velocities of seismic waves in porous rocks. *GEOPHYSICS*.

Vik, B., Bastesen, E. and Skauge, A. 2013. Evaluation of representative elementary volume for a vuggy carbonate rock – Part: Porosity, permeability, and dispersivity. Journal of Petroleum Science and Engineering. 112, pp.36–47.

Wong, T., David, C. and Zhu, W. 1997. The transition from brittle faulting to cataclastic flow in porous sandstones: Mechanical deformation. *Journal of Geophysical Research: Solid Earth*.

Wyllie, M.R.J., Gardner, G.H.F. and Gregory, A.R. 1962. Studies of elastic wave attenuation in porous media. *Geophysics*. 27(5), pp.569–589.

Wyllie, M.R.J., Gregory, A.R. and Gardner, G.H.F. 1958. An experimental investigation of factors affecting elastic wave velocities in porous media. *Geophysics*. 23(3), pp.459–493.

Wyllie, M.R.J., Gregory, A.R. and Gardner, L.W. 1956. Elastic wave velocities in heterogeneous and porous media. *Geophysics*. 21(1), pp.41–70.

Yang, X., Jia, Y., Liu, H. and Shan, H. 2009. Characteristics and causes of the preconsolidation stress of soils in the Yellow River Delta. *Journal of Ocean University of China*. 8(3), pp.215–222.

Yilmaz, I. 2018. Performance of Core Strangle Test (CST) on Porous Rocks. Geotechnical and Geological Engineering. 36(6), pp.3551–3557.

Yousefpour, N. and Fallah, S. 2018. Application of Machine Learning in Geotechnics.

Yue, Z.Q., Chen, S. and Tham, L.G. 2003. Finite element modelling of geomaterials using digital image processing. *Computers and Geotechnics*. 30(5), pp.375–397.

Zhang, D., Zhang, R., Chen, S. and Soll, W.E. 2000. Pore scale study of flow in porous media: Scale dependency, REV, and statistical REV. Geophysical Research Letters. 27(8), pp.1195–1198.

Zum Gahr, K.H. 1987a. *Microstructure and wear of materials*. Elsevier Science Publishers B.V.

Zum Gahr, K.H. 1987b. *Microstructure and wear of materials*. Elsevier Science Publishers B.V.