Workflow model for the digitization of mudrocks

JIM BUCKMAN1*, CAROL MAHONEY2, SHEREEF BANKOLE1, GARY COUPLES1, HELEN LEWIS1, THOMAS WAGNER3, CHRISTIAN MÄRZ4, VLADIMIR BLANCO5 & DORRIK STOW1

1Institute of Petroleum Engineering, Heriot-Watt University, Riccarton, Edinburgh, EH14 4AS, UK
2School of Civil Engineering and Geosciences, Cassie Building, Newcastle University, Newcastle upon Tyne, NE1 7RU, UK
3Lyell Centre, Heriot-Watt University, Riccarton, Edinburgh, EH14 4AS, UK.
4School of Earth and Environment, University of Leeds, Leeds, LS2 9JT, UK
5Instituto Colombiano del Petróleo, Ecopetrol, Bucaramanga, Colombia.

*Correspondence: j.buckman@hw.ac.uk

Abstract: Mudrocks are highly heterogeneous in a range of physical and chemical properties including: porosity/permeability, fissility, colour, particle composition, size, orientation, carbon loading, degree of compaction and diagenetic overprint. It is therefore important that the maximum information be extracted as efficiently and completely as possible. This can be accomplished through high-resolution analysis of polished thin-sections by scanning electron microscopy (SEM), with the collection of large-area images and x-ray elemental map montages, and the application of targeted particle analysis. A workflow model, based on these techniques, for the digitization of mudrocks is presented herein. A range of the data that can be collected, and the variety of analyses that can be achieved are also illustrated. Data collection is discussed in terms of inherent problems with acquisition, storage, transfer and manipulation, which can be time consuming and non-trivial. Similar information and resolutions can be achieved through other techniques, such as QEMSCAN and IR /
Raman spectroscopic mapping. These can be seen as complementary to the workflow described herein.

**Supplementary material:** supplementary materials on typical workflows, based on the specific software packages used herein are available at https://doi.org/xxxx.

The appropriate naming of fine-grained / muddy sedimentary rocks has attracted much attention recently (Milliken 2014; Lazar *et al.* 2015a; Camp *et al.* 2016). For the purpose of this paper, the term mudrock is used in the broadest sense to include all siliciclastic fine-grained (less than 63 µm) muddy semi-consolidated to consolidated materials such as shale, mudstone and siltstone (Stow 2015). For a more in-depth discussion on nomenclature of this rock group please see Lazar *et al.* (2015a). Mudrocks comprise one of the most significant groups of sedimentary rocks (Ibbeken & Schleyer 1991), being increasingly recognized for their significance as oil and gas cap rocks (Al-Bazali *et al.* 2005; Olabode *et al.* 2012), and unconventional reservoirs (Zou *et al.* 2015). They also form seals for CO₂ storage (Olabode *et al.* 2012), potential repositories for radioactive materials (Komameni and Roy 1979; Delage *et al.* 2010), and archives of massive and at times rapid perturbations of Earth’s climate in the past. Despite this, they have historically been considered as superficially bland and monotonous rocks, and as such have not received as much attention from scientific investigations as do other sedimentary materials (e.g., sandstones and carbonates). Over recent years, this focus has gradually changed, and with the increase in studies focusing on high spatial resolution utilizing geochemistry or scanning electron microscopy (SEM), the current consensus is that most mudrocks are highly heterogeneous at the nanometer to decimeter scale (Camp *et al.* 2013; Milliken 2014; Milliken *et al.* 2016; Ma *et al.* 2017). Such heterogeneous properties include: porosity, permeability, fissility, colour, particle type (composition), particle size and particle orientation (see Lazar *et al.* 2015b). They manifest the interplay of a wide range of mechanisms, from orbital forcing, to local provenance and sedimentation, to post-depositional alteration through compaction and process of diagenesis. Many of these properties and processes impact on the physical and chemical attributes of the material, and thus the overall hydrocarbon generation potential (Wagner *et al.* 2013).
It is therefore desirable to have an effective study regime for the examination of mudrocks, to develop a workflow that can access the heterogeneity of such materials at a range of scales, that can provide both qualitative and quantitative data on aspects of microstructure and composition. Automated SEM techniques offer such an opportunity to collect large data sets on the textural and chemical parameters of mudrocks, in a relatively cost effective fashion. The aim here is to outline a generic workflow that is a combination of scanning electron microscopy / energy dispersive x-ray techniques (commonly available from a range of manufacturers). This type of generic workflow is ideally suited to the on-going examination of mudrocks, into the heterogeneity of mudstones, and their mineralogical and textural variability at the nanometer to centimeter scale. Here we address a number of examples of the type and scale of data that can be collected using this approach, with the aim of clearly offering insights into the diversity of data and analyses that are possible. The range of problem areas associated with such data collection and its analysis are also addressed, as an understanding of such issues is critical in terms of data quality and interpretation. Finally, the workflow is compared to other potential SEM and optical techniques, in terms of resolution, and range of data collection achievable.

**Material and methods**

**Materials**

Samples were selected from a variety of previous projects, to illustrate methodology applied in our examinations of mudrocks, and the scope of data that can be collected. Mudrocks illustrated herein include: Quaternary deep-water mudstones from IODP Expeditions 317, 335 and 339 (Canterbury Basin, New Zealand, Laxmi Basin, Arabian Sea, and off the Iberian Peninsula / Gulf of Cadiz), with hemipelagites, turbidites and contourites; a range of calcareous, siliceous and organic-rich Cretaceous shales (from Colombia); the Jurassic Blackstone Shale (from Yorkshire, UK); and a Carboniferous shale (from Fife, Scotland). These sample materials range from semi-consolidated to consolidated / lithified.

**Sample preparation**

The majority of samples were prepared as polished thin-sections using standard methods (e.g. Miller 1988). A limited number of the IODP samples were additionally
prepared with a Fischione 1060 SEM ion-mill, using argon at 5 kV for 3 hours, at The
University of Oklahoma. All samples were vacuum impregnated with resin.

Techniques
The digitization of mudrocks is approached using scanning electron microscopy
(SEM) combined with energy dispersive x-ray (EDX) analysis, used in our
workflows, which can be applied to whole sample surfaces or selected areas of
interest. Specific software packages used were: ‘Maps (v2.1)’ by FEI (Hillsboro,
USA), ‘AZtec (v3.3) Large Area Maps (LAM) and AZtecFeature’ by Oxford
Instruments (Tubney Wood, UK). These workflows / techniques were used as they
offer a relatively rapid method of obtaining high-resolution large area digital data, that
can be further investigated to extract details on mudstone microstructure at the
nanometer to centimeter scale. The workflow was carried out using an FEI Quanta
FEG 650 SEM. The microscope was operated in low-vacuum mode (0.82 Torr),
between 15 and 20 kV, with a typical spot size of 3.5 - 4.5, a working distance of 10
mm; the samples were examined uncoated.

This workflow can be achieved on a number of alternative systems, with a range of
similar software offered by various manufacturers. Here, a generic workflow is
shown, which illustrates actions undertaken, their order within the workflow, the type
of data output, and an indication of potential areas of use (Fig. 1a-c). The initial
workflow involves the collection of data (Fig. 1a), followed by a secondary stage of
data manipulation (Fig. 1c). Data may either be processed within the software used to
collect it, or with the aid of third party software such as “Fiji / ImageJ” (v1.51n,
National Institute of Health, USA), “Scandium” (v5.2, Olympus Soft Imaging
Solutions GmbH), and “MATLAB” (R2017b, 9.3.0.713579, MathWorks,
Massachusetts, USA). Details on specific workflows utilizing the software packages
within this work are illustrated in the supplementary materials.

The first stage of the workflow involves the collection of a series of low to high-
resolution montages, constructed from a series of tiles collected on a regular grid
across the surface of the sample (Fig. 2). After initially collecting an optical image of
the sample (Fig. 2a), tiles are typically first collected at low-resolution (2 mm
horizontal field of view) across the whole of each sample (Fig. 2b, c). These allow
for accurate location of samples and reduces the amount of empty space scanned
during higher resolution / higher magnification scans. This is followed by higher
resolution montages, taken from selected areas across the thin-section (Fig. 2b-f).
Image montages can be produced from any available detector, which in low vacuum
mode, include backscattered electron (BSE), gaseous secondary electron (GSE) or
cathodoluminescence (CL) detectors. High-resolution image montages are commonly
formed from a series of 100 to 25,000 or more tiles. Tiled images from areas of
specific interest are typically recorded with a horizontal field of view (HFOV) of 518
µm (with 10% overlap), with higher resolution scans having as small as 10 µm HFOV
(with 25% overlap). The imaging of whole polished thin-sections (25 x 48 mm) can
take between 30 minutes and 12 hours, depending on factors such as tile size, pixel
resolution and scan time (Table 1), while the highest resolution images (10 µm tiles)
of large areas may take in excess of 3 days, for further discussion see below
(Montages, time versus resolution conundrum). Once a montage has been acquired it
is typically automatically stitched and exported. Montages can be saved in a number
of formats (tiff, raw etc); with raw being the most commonly used format. Both
stitched montages and individual tiles can be used for analysis purposes (see Fig. 1).
Fully automated energy dispersive x-ray (EDX) mapping is also performed with the
collection of individual tiles, although typically with tens to hundreds rather than
thousands of tiles per montage. Tiles typically have a horizontal field of view (HFOV)
in the region of 518 µm - 2 mm. Best results in terms of elemental x-ray maps are
generated with between 20 and 75 slow scans per tile, although for practical reasons
this is often limited to only 3 or 4. EDX mapping of whole thin-sections takes
considerably longer than comparable image montages, and depends on additional
factors such as number of scans per tile (see Best practice with backups and typical
file size encountered). Elemental maps can be produced for all detected elements, and
these can be further processed to produce phase maps of associated groups of
elements, which can be manually identified and assigned to known minerals or
mineral groups such as calcite, quartz, clays etc. Both elemental maps and phase
maps can be montaged and exported as single large area high-resolution images, in
multiple formats (tiff, jpeg, bmp, gif, png).

The final stage of the workflow involves the analysis of selected particle types.
Backscattered (BSE) images are collected, segmented and binarized, to select the
features of interest through selective thresholding of the images grayscale histogram. Further, physical (e.g. minimum pixel size, equivalent circular diameter) and elemental criteria (e.g. presence / absence / ratios of specific elements) can be defined to determine which particles are selected, as well as the level of analysis performed on each type of particle (physical only, or physical and elemental analysis). Information that can be automatically gathered includes parameters such as size, shape, composition and distribution of particles. With careful selection of criteria and thresholding, many tens of thousands of particles can be detected and analyzed within a timeframe of 1 to 12 hours, dependent upon factors such as the number of particles to be analyzed, total area selected and size of individual tiles to be scanned. Careful thresholding is key to success and can be supplemented, prior to particle analysis, by basic image analysis techniques such as particle separation and noise reduction. Collected data can be further analyzed for a range of purposes.

Examples of use

High-resolution image montages

Montages made using this technique are usually constructed from a series of BSE image tiles, either from across the whole sample area (Fig. 3a), or from selected areas at a range of resolutions (Fig. 3b, c, d). Resolution of individual tiles, for montages recorded at higher magnifications over specific areas of interest, can be high enough to illustrate fine details within mudrocks such as coccolith plates, clay, pyrite crystals and intergranular micro-porosity (Fig. 3e), with the ability to resolve structures between 10 – 100 nm (Table 2). Images of polished thin-sections (Fig. 4a, b) and ion-beam milled samples (Fig. 4c-e) provide the ability to visualize variability (heterogeneity) at the sub-micron to centimeter scale, while being able to examine the spatial relationships between such areas. This has the advantage of providing an overview of the whole sample, illustrating areas of interest that can be targeted for imaging at higher resolution, while also providing information on heterogeneity, showing the relationship and connectivity between areas comprising different fabrics. Such examples include authigenic siliceous partings interlaminated with more clay-rich lamina (Fig. 4a; see Buckman et al., 2017b), silt-rich burrow fills within a more clay-rich matrix (Fig. 4b) and the distribution of pyrite-rich burrow fills (Trichichnus) within an otherwise more homogeneous hemipelagite (Fig. 4c). Another large area high-resolution montage is
illustrated in figure 4d (42,294 x 40,385 pixels) of a polished block of mudrock that has undergone experimental shearing, exhibiting the pervasive occurrence of authigenic siderite, which along the shear zone has a cataclastic texture (Fig. 4e); illustrating the potential for analysis of mudrocks in geomechanics studies.

High-resolution GSE charge contrast image (CCI) maps can also be used to differentiate carbonate produced through diagenesis from that of original calcareous tests of foraminifera, as well as the occurrence of carbonate rich coccolith-bearing fecal pellets (Fig. 5). The latter case, and the example of siliceous partings, both illustrate how this workflow can prevent the overestimation of primary depositional versus authigenic components.

Image analysis can be used to extract data on parameters such as particle size, perimeter, length, width, equivalent circular diameter (ECD) and orientation (Bankole et al. 2016). A pilot study of six randomly chosen high-resolution montages (each comprising 10x10 tiles) from an example of hemipelagite, Canterbury Basin, New Zealand (Fig. 6a), cut perpendicular to bedding, were analyzed in terms of orientation using the measure function in ImageJ. The results plotted as a series of rose diagrams (Fig. 6b), illustrates the heterogeneous nature of microstructure within mudstones. The three areas towards the bottom of the section (areas 4 to 6) display a consistent particle orientation nearly perpendicular to bedding, while those from the upper part are either obliquely oriented (but parallel to bedding, areas 1 and 2), or have developed a random fabric orientation (area 3). The example illustrated (Fig. 6) is encouraging, as it suggests that it may be possible to analyze particle orientation, from individual tiles, within collected image montages, across areas of several centimeters squared.

Individual BSE tiles used to produce large area high-resolution montages (Fig. 7a, c), can be batch processed using image analysis software such as Fiji, to calculate percentage porosity per tile. These values can then be plotted using software such as MATLAB, to illustrate spatial variation in percentage porosity across the whole montaged area (Fig. 7b, d), as recently demonstrated for carbonates (Buckman et al. 2017a). Although such plots do not indicate degree of porosity connectivity, they can show relatively subtle spatial variations in degree of porosity. Although some care
needs to be taken over the degree of image processing during analysis, and the
presence of surface contaminants (fluff, hairs, grease), dehydration cracks and salt
crystals can increase or decrease apparent porosity within each tile, the technique can
be applied as a visual tool, to rapidly assess spatial trends in porosity that are not
obvious within the BSE montages. Segmentation for porosity can be subjective and
vary from user to user, although a degree of automation can be introduced to reduce
user bias by utilizing plugins such as “Trainable Weka Segmentation” (TWS) within
Fiji.

Representative selected individual BSE tiles can also be used as the basis for three-
dimensional flow simulations (Ma et al. 2014; Song et al. 2017), with the construction
of digital rock physics (DRP) models. In such cases, BSE tiles are binarized to select
porosity (Fig. 8a-c), processed stochastically to produce 3D models that preserve
porosity and fabric relationships (Fig. 8d-f), to calculate representative values for
permeability (Fig. 8g-i). Due to the highly heterogeneous nature of mudrocks, such
models are particularly important, as they allow the possibility of quantitatively
predicting localized variability in permeability, that would otherwise be hard /
impossible to directly measure through more traditional methods.

Individual tiles and whole montages can be further processed through image analysis
(e.g. using ImageJ), to obtain percentage occurrence of specific components (Fig. 3f,
Table 3) through segmentation of the grayscale histogram, which can separate pyrite,
silt particles (quartz and feldspar), clays, biogenic materials (foraminifera, coccoliths
etc) and porosity. Although it is not always possible to clearly separate all phases
using grayscale, due to overlaps, the technique does offer a rapid method of assessing
compositional variability within mudstones. Sample preparation is important, as
variations in topography, such as beveling around grain boundaries, will affect
segmentation and therefore potentially compromise phase separation.

EDX large area maps
EDX large area mapping (LAM) is ideally suited to produce individual qualitative
elemental maps of whole thin-sections. However, due to time constraints it is
common to image narrow strips perpendicular to bedding (Fig. 9, 10). Such elemental
maps can be utilized to illustrate millimeter-scale changes in mineralogy and
elemental composition, due to either original depositional fabric variation (Fig. 9) or
differential diagenetic overprint (Fig. 10).

An EDX montage through a calcareous mudrock from the Cretaceous of Colombia
Illustrates areas rich in calcareous components (Fig. 9c), representing foraminiferal
tests, authigenic calcite grown within the tests, and lens-shaped pods rich in
coccoliths. In addition, consideration of areas containing iron and sulphur illustrates
the occurrence of pyrite and limited amounts of iron oxide particles (Fig. 9d-g).

In contrast, in a sample of siliceous mudrock also from the Cretaceous of Colombia
(Buckman et al., 2017b), the distribution of aluminum clearly displays a strong
variation, with a strong negative relationship to the distribution of silica (Fig. 10a-c).
For aluminum, brighter areas (1) are dominated by clay minerals, reflecting the
original lithological / mineralogical character of the mudrock. In contrast, areas with
less aluminum (2) but higher amounts of silica (Fig. 10b, c), represent areas where
high levels of authigenic micro-quartz have been precipitated within the original
mudrock (see Buckman et al. 2017b). From the same sample, consideration of EDX
maps for iron and sulphur (Fig. 10d-f), as with the previous example, can also be used
to qualify the distribution of iron-sulphides and -oxides. In this case, pyrite is
sparingly distributed within the lower part of the section, while thin zones of iron
oxides occur throughout the section. The distribution and abundance of pyrite and
iron oxides may reflect differences in depositional environment, diagenesis, and / or
weathering.

As well as qualitative elemental maps, phase maps can be constructed by comparison
of the distribution of individual elements at each pixel. Quantitative data can then be
extracted for the percentage occurrence of each identified phase (Fig. 11). In the
illustrated example (from the calcareous shale shown in figure 9), mineralogical
phases are easily separated and quantitative data is automatically generated (Table 4).
Such quantitative data is similar to that which can be obtained from BSE montages
(from segmentation of grayscale), but benefits from the input of elemental data,
helping to separate mineral phases with similar grayscale values, which would
otherwise be difficult to quantify.

**Particle analysis**

**Heavy mineral analysis.** Heavy mineral analysis is typically associated with palaeo-
weathering and provenance studies of sandstones and sands (Garzanti & Andò 2007;
Ejeh et al. 2015). Nevertheless, other studies have shown that heavy mineral assemblages are also common within many finer grained mudrock deposits (Totten & Hanan 2007). Such studies usually require the sample to be crushed, releasing the heavy fraction, which is then concentrated and either analyzed optically or through SEM. A preliminary study of a Pleistocene hemipelagite from the IODP Expedition 339, Gulf of Cadiz, clearly illustrates that SEM particle analysis from polished thin-section samples is a viable alternative for such studies (Fig. 12). In this example, an area slightly under 2 square millimetres, identified zircon (114), monazite (102), ilmenite (388) and rutile/anatase (94) particles. Further development of the technique, currently underway, specifically constructing a ‘heavy mineral standard’ block for segmentation/thresholding and criteria calibration, will enhance the practicality of the technique and potentially its use for palaeo-weathering and provenance studies within mudrocks. It also has the advantage of preserving and making accessible information on the relationship between heavy mineral assemblages and mudstone fabrics, mineral association and micro stratigraphy; areas that have not been previously investigated.

Pyrite versus iron oxides. Both pyrite and iron oxides are common subsidiary components within many mudrocks, with less than 1% for pyrite and 3% for iron oxides in average shale (Yaalon 1962) but tens of percent for pyrite in some black shales (März et al. 2011). The occurrence and distribution of iron sulphide and oxide species can provide important information on environment of deposition, changes in pH/Eh values and as indicators of palaeoweathering (Raiswell et al. 1988; Poulton & Canfield 2005; Ding et al. 2014). Current work on artificially weathered Blackstone Shale from the Jurassic of Yorkshire, using data from particle analysis of mudrocks thresholded for pyrite (Fig. 13a, b) and displayed as a scatter plot of sulphur versus iron, can be used to differentiate ‘pristine’ pyrite from sulphur-depleted pyrite (Fig. 13c, d). In this case, pristine pyrite is taken to have a weight % ratio of Fe:S of 0.9 or less, while sulphur-depleted pyrite is plotted having a ratio of greater than 0.9. Backscattered images of pyrite rims, in both pyrite framboids and euhedral crystals, reveal that cores of pyrite are surrounded by variably sulphur-depleted and oxidized pyrite (Fig. 13e, f). This illustrates the ability to examine changes in redox potential within mudrocks at the micron to millimeter scale, through the distribution of pyrite and the occurrence of sulphur-depleted oxidized pyrite, which is crucial for reliable
palaeo-environmental reconstructions and to understand the weathering behavior of pyrite at the Earth’s surface. Particle analysis of experimentally-weathered chips of mudstone clearly illustrate 300 – 500 micron weathered rims of sulphur-depleted pyrite (Fig. 13b).

Organic matter. Organic matter (OM) is of common occurrence within many mudrocks, particularly black shales. Total organic carbon (TOC) contents of 5 to 23 wt% variously recorded by Erdman and Drenzek (2013) from the Marcellus, Haynesville, Woodford and Barnett Shales of the United States, while TOC in the Carboniferous Blackstone Band in the UK reaches around 50 wt% (Huc et al. 1992).

Particle analysis can be used to investigate the character and composition of such organic matter within mudrocks, through thresholding BSE grayscale images, selection of the darkest particles (Fig. 14a, b). Preliminary observations of scatter plots for carbon versus oxygen from such particles shows a negative relationship (Fig. 14c, d). This relationship may reflect differences in the C:O ratio between macerals (vitrinite, inertinite, liptinite). Such relationships have been previously noted for macerals within coals (Mastalerz et al. 2013; Holuszko & Mastalerz 2015) but are less well constrained for TOC-rich marine shale, where amorphous organic matter (AOM) of different origin commonly dominate. Further work in this area is required to determine the best working parameters and the degree of significance of particle analysis in the study of OM in mudrocks. A number of potential problems exist with the technique in relationship to OM, namely that high excitation voltage is not ideally suited for the examination of OM in terms of beam penetration and potential damage through beam heating. In addition, the skirt effect (present in low vacuum use) inevitably leads to a large degree of surrounding inorganic material being captured by EDX when examining small (micron sized) particles of OM. The latter phenomenon, in particular may explain the variation in oxygen and carbon ratio, if a carbon contribution is picked up from calcite, or oxygen from quartz and clays. Neither of these possibilities have been tested during the present work. As water vapour (H₂O) was used as the imaging gas, some oxygen will have derived from the chamber atmosphere, although this would not have produced the observed negative relationship. These problems can be mitigated through low kV high-vacuum analysis (Fig. 13b). It is worth noting that low kV that will not induce surface charging (1 – 2 kV) would be required, as coating with carbon would be counter-productive, and
other conductive coatings such as gold would impact too much on carbon and oxygen x-ray detection. The current work suggests that an SEM capable of low kV high-resolution imaging would be beneficial for such analysis. Another potential problem area is the differentiation of OM from resin used to impregnate the mudrock samples. Careful EDX analysis of well-defined impregnation resin is likely to help clarify this question. The situation could also be improved through preparation techniques that do not require resin impregnation, such as large area ion-beam milling. Despite these issues, we feel that particle analysis of OM within black shales and other organic rich mudrocks has high potential and is worthy of further exploration.

Discussion

As has been shown, the workflow used herein can be used to characterize mudstones at the sub-micron to centimeter scale, with potential information collected on grain type, morphology, orientation, composition, pores and porosity, as well as environmental factors such as Eh / pH. Nevertheless, a number of factors need to be taken into serious consideration, some of which are discussed below, as are a number of possible alternative techniques for large area high-resolution data collection.

Montages, time versus resolution conundrum

For image montages, consideration of parameters affecting ultimate image resolution (pixel resolution, scan rate and tile size) it is evident that to scan a typical whole polished thin-section (25 x 48 mm) at 'high-resolution’ is only practical under a limited range of settings (see Table 1). In most cases it is not possible to scan a whole slide at a tile resolution higher than 259 µm, as this would take more than 2 days. For mudrocks, it will generally be necessary to select smaller areas of interest to scan at the highest resolutions. In practice, it has been found that a 10 x 10 tile set, scanned at a pixel resolution of 1536 x 1024 pixels, scan rate of 10 µs and tile horizontal width of 10 µm will take around 30 minutes, which is adequate for most detailed fabric studies of mudrocks. The time versus resolution conundrum is further compounded when considering the incorporation of elemental montage maps, as time also has to be factored in to accommodate multiple EDX scans.

Best practice with backups and typical file size encountered
Given the typical size of files and the time taken to transfer collected data, a centralized storage facility, with high transfer rates, is highly recommended. A number of other issues related to storage and manipulation of large data sets are discussed below.

One of the main issues associated with the collection of large area high-resolution montages is file size. The workflow commonly generates multi gigabyte sized files, which rapidly become difficult to store, manipulate and transfer. Direct observations have indicated that for practical purposes Tiff files are limited to around 2 GB maximum size, while larger files require saving in raw format. Raw format files come in a wide range of standards and not all software packages open raw files. For ease of use, Fiji / ImageJ is recommended. Total project sizes can be 60 GB or larger, requiring careful consideration of storage and archiving protocols, with project transfer times in excess of an hour per project. Good practice involves direct storage to a central archive (with full system redundancy), accessible to both SEM administrator and client, minimizing the need to repeatedly transfer data.

Elemental mapping suffers from the same general issues as found with image maps, being capable of generating large image files that can be hard to manipulate, save and move. Larger EDX maps (within AZtec LAM) are automatically reduced in pixel resolution, which helps with storage and manipulation problems, but limits the resolution of larger EDX maps. Where resolution requires preserved, larger maps can be saved as a series of smaller sub-sets.

The maximum number of particles that can be practically analyzed with particle analysis (using AZtecFeature) is restricted to around 160,000 particles (equating to 1 - 2 GB). Larger data sets become cumbersome and difficult to manipulate. Cutting, pasting and organizing data within larger data sets becomes particularly problematic, and can quickly lead to the introduction of errors. Best practice therefore involves analyzing thin transects across the sample, or smaller targeted areas, with runs of 2 - 4 hours commonly generating data sets with 4,000 or so particles.

Obtaining suitable flat sections for scanning
For the purpose of automated image / data collection by SEM, the prepared surface ideally should be flat, easy to orientate perpendicular to the beam, and have undergone minimal surface damage during sample preparation (cutting, lapping, polishing). With unconsolidated samples, there is the additional potential issue of introducing damage during sampling and vacuum impregnation with resin (essentially cracking and smearing). Polished thin-sections benefit from a flat horizontal surface, parallel to the base of the thin-section, which once focused within the SEM remain in focus, without the requirement of further focusing protocols. However, damage during cutting, lapping and polishing processes can be high, including plucking of silt and biogenic particles, as well as the smearing and swelling of clays due to interaction with water. The use of oil based lubricants and mechano-chemical polishing (using colloidal silica) with non-rotary polishers can be used to help minimize problems; although oil based lubricants may interact with organic matter. Large area (1 centimeter plus) ion-beam milling combined with sample rotation during milling helps with the removal of surface damage introduced during sample preparation. However, focusing problems may occur due to the angled surface produced by the milling process, not necessarily parallel to the sample base, which requires the use of time costly focusing protocols. In addition, ion-beam milling can produce additional surface artifacts (Milliken & Olson 2016). With this in mind, polished thin-sections generally work well where large areas (mm$^2$ to cm$^2$) are being investigated at moderate to high-resolution (3 to 100 nm pixel resolution), while ion-beam milled sections typically offer improved image quality (in terms of lesser damage), but owing to increased effort required to keep such samples focused over larger areas, are best suited to targeted analysis over more defined smaller areas (mm$^2$) at high-resolution (nanometer pixel resolution).

Comparison with other potential techniques
Other alternative methods also exist that can generate high-resolution results over large areas on mudrocks, in a similar fashion to those presented here. These include QEMSCAN (Omma et al., 2017), and potentially Infra Red (IR) or Raman spectroscopic mapping (Greenberger et al. 2015; Hunt 2017). QEMSCAN offers the potential for rapid mineralogical identification and mapping of mudrocks, and is readily available as a technique from many commercial service providers. It essentially operates in a manner similar to optical petrographic point
counting, with EDX analysis taking place over a fixed grid. EDX analysis, used in conjunction with mineral identification libraries, coupled with BSE imaging, provides a simplified superior method for phase identification. QEMSCAN surveys are typically spatially limited in the area that they scan, particularly so for higher resolution surveys. In addition, the highest resolution maps (typically commercially offered) have a spacing of 4 µm, which is too coarse to pick out the mineralogical variations within mudrocks that typically have a grain / particle size of 1 to 2 µm or less. The technique does, however, present great potential for gathering information on general mineralogical variation across samples at the millimeter to centimeter scale. IR and Raman spectroscopic mapping are optically based techniques using polychromatic and monochromatic light respectively, that identify mineral phases through differences in vibration energy. The techniques have a maximum resolution in the order of 1 µm and the capability to scan a whole polished thin-section. Therefore, IR and Raman spectroscopic mapping, like QEMSCAN, may provide a useful rapid identification aid, and benefits from high-resolution, although the ultimate resolution is still less than that required for many mudrock clay fabrics. For mudrocks, both QEMSCAN and IR / Raman spectroscopic mapping can be seen as additional supportive tools, although not having the resolution required to fully analyze the complex mineralogical and fabric associations typical of many mudrocks.

QEMSCAN is typically more expensive to purchase than the more generic SEM’s equipped with large area and EDX mapping software packages described herein, is specific to a single manufacturer, and tied to the use of one type of EDX detector. Whereas generic large area SEM image montaging and EDX software are available from many SEM/EDX manufacturers, most of which would be adaptable to the workflow illustrated in the present paper. In addition, much of the analysis software required is freeware, or otherwise readily available. All techniques (QEMSCAN, IR / Raman spectroscopy, large area high-resolution mapping) can be considered time consuming and may result in data storage and manipulation problems.

Conclusions

Due to the fine-grained highly heterogeneous nature of mudrocks, a streamlined approach to their efficient digitization is essential. As illustrated, the use of SEM to collect automated high-resolution large area image montages, at a variety of scales over polished-thin sections and ion-beam milled surfaces, in combination with large
area elemental composition maps, and the addition of automated particle analysis provides a flexible path for the digitization of mudrocks, for the observation on heterogeneity at the sub-micron to centimeter scale. Images and elemental analysis can be collected over a range of scales, allowing for visual identification of pores, clay and silt sized particles, the occurrence of mineral phases (clays, quartz, feldspars, calcite, pyrite, iron-oxides, heavy-minerals etc) and general textural features such as laminae. These can be used either directly, or processed via image analysis using freely available software such as Fiji / ImageJ to gather further information on percentage coverage, or extract quantitative data on pores or particles (shape, size, perimeter, orientation). Similar information can also be collected directly through selected particle analysis using software such as AZtecFeature. In addition, BSE images of porosity can be further used to construct 3D models to investigate parameters such as permeability using digital rock physics (DRP) techniques. The typically large size of files generated using the workflow, requires adequate centralized storage, and consideration of file size when selecting areas of interest. In addition, careful consideration of preparation techniques and potential artifacts, are important in obtaining flat representative surfaces suitable for large area high-resolution montaging. Finally, QEMSCAN and IR / Raman spectroscopic mapping, where available may provide supplementary or supporting information that can be used in conjunction with data collected through the illustrated workflow.

The authors would like to acknowledge Mark Curtis of The University of Oklahoma for preparation of ion-beam milled samples, and the Institute of Petroleum Engineering, Heriot-Watt University for providing scanning electron microscopy support. We acknowledge the help of FEI, in the preparation and analysis of the sheared shale sample illustrated in figure 4d, e. The authors acknowledge Ecopetrol and IODP for access to some of the samples examined during the present study. We also thank Zeyun Jiang and Tianshen Huang for help with DRP software developed at the Institute of Petroleum Engineering, Heriot-Watt University.

References
mudrock and the choice of representative sample. *Fifth EAGE Shale Workshop*, 2-4th May 2016, Catania, Italy.


**Figure captions**

**Fig. 1.** Generic workflow used in the analysis of mudrocks. A) Actions undertaken, showing typical order of workflow. Low, moderate and high resolution typically equate to individual tile horizontal field of view of 2mm, 518 µm and 10 – 70 µm respectively. B) Type of data output from various stages of the workflow in (A). C) Potential uses of data generated through use of workflow.

**Fig. 2.** Screenshots, illustrating stages in obtaining high-resolution large area montages using ‘Maps’. A) Nav-cam image of SEM stage. B) Low-resolution montages of whole polished thin-section (1), moderate resolution strip montage (2), higher resolution selected area (5), and highest resolution selected area montages (3, 4). Note that relative positions of optical image and BSE image do not exactly match (shown by double headed arrow), illustrating the importance of lower resolution BSE images to accurately locate samples, and target areas of interest. C) – F) Representative individual tiles from montages taken at different resolutions from area 1, 2, 3 and 5 respectively.

**Fig. 3.** Backscattered (BSE) high-resolution montages of Quaternary hemipelagite from IODP Expedition 339, Iberian Peninsula. A) Montage comprising 15 x 15 tiles, 2.07 mm wide, with 768 x 512 pixel resolution, red box marks position of montage in (B). Darker margins and along cracks due to higher resin impregnation. B) 10 x 10 tiles, 518 µm wide, with 1536 x 1024 pixel resolution, red box marks position of montage in (C). C) 10 x 10 tiles, 70 µm wide, with 1536 x 1024 pixel resolution, red box marks position of montage in (D). D) 10 x 10 tiles, 15 µm wide, with 1536 x 1024 pixel resolution, yellow box marks position of individual tile illustrated in (E). E) Single 15 µm wide tile from montage in (D). C = calcite (biogenic), Q = quartz silt particle, M= clay/ matrix, P= porosity, Py= pyrite. F) Stitched montage (C), thresholded for calcite (yellow), silt (red), clay (brown), pores (green) and pyrite (pale blue), see table 3 for percentage coverage. Scale as in (C).

**Fig. 4.** Examples of high-resolution large area BSE montages of various mudrocks. A) and B) Laminated silicified shale and bioturbated shale, from the Cretaceous of Colombia (Z39 and P1+100). C) Hemipelagite from IODP Expedition 339 off the Iberian Peninsula, Pleistocene. D) and E) Siderite rich shale after laboratory shear
inducement, from the Carboniferous, Fife, Scotland. Red arrow in (D) indicates area
of (E). Scales as indicated.

**Fig. 5.** Examples of high-resolution large area montages of calcareous shale (LC224,
Cretaceous, Colombia). A) BSE montage, with solid globular shaped calcitic
foraminifera and lens shaped coccolith fecal pellets. B) GSE montage of same area as
in (A), in which charge contrast imaging (CCI) picks out additional fecal pellets (2)
not seen in the BSE image, as well as ones seen by BSE (1). Also note that the CCI
technique clearly differentiates calcite forming the foraminiferal test from that of
authigenic calcite infill (3).

**Fig. 6.** Illustration of the variability in particle orientation, within a hemipelagite
sample from IODP Expedition 317 (Canterbury Basin, New Zealand), extracted from
high-resolution BSE montages using Fiji. A) BSE montage of polished thin-section
surface, perpendicular to bedding. Dashed line follows bedding junction. Location of
6 individual high-resolution montages indicated with boxes 1-6. B) Rose diagrams of
grain orientation from regions 1-6. Areas below the dashed line all exhibit near
vertical grain orientation in respect to bedding, whereas those above are inclined (but
parallel / sub-parallel to bedding), or show random distribution.

**Fig. 7.** Example of porosity within mudstone contourite deposit from IODP
Expedition 339, Bay of Cadiz, hole U1389e 66R1 18-21. A) and C) BSE montage,
constructed from 10 x 10 tiles, in (A) each tile approximately 70 microns wide, and in
(C) 10 microns. B), D) Porosity maps based on % porosity of individual tiles, plotted
using MATLAB, corresponding to images in (A) and (C) respectively.

**Fig. 8.** Illustrative examples of Digital Rock Physics (DRP) models for mudstone
from turbidite deposits (IODP Expedition 335, Laxmi Basin, Arabian Sea, hole
U1457c 49R6 30-34). A) – C) Each showing three images derived from large area
high-resolution montage tiles. Images binarized, with black representing pores, each
tile has a resolution of 86 nm per pixel. D) – F) DRP generated models, with
dimensions of 400 x 400 x 400 voxels, from binarized images in (A) to (C), showing
pore and grain structure based on stochastic reconstruction. G) – I) Porosity and
permeability values calculated by DRP models.
Fig. 9. EDX montage strip maps using ‘AZtec LAM’, of Cretaceous Colombian calcareous shale (LC224).  A) – E) Individual element maps for aluminium, silica, calcium, iron and sulphur respectively.  F) and G) Composite EDX maps for the elements in (A) to (E), illustrating the distribution of clays (blue to blue-green), calcite (purple), iron-oxide (red) and pyrite (orange).  Scale bar as indicated on individual strips.

Fig. 10. EDX montage maps using ‘AZtec LAM’ through a Cretaceous Colombian siliceous shale (Z39).  A) Map for aluminium, B) silica and C) combined map for aluminium and silica, exhibiting the high level of silicification; 1 = original clay matrix, 2 = silicified layer.  D) Map for iron, E) sulphur and F) sulphur and iron showing distinct layers of iron-oxide (red) and discrete pyrite particles (yellow).  Scale bar as indicated on individual strips.

Fig. 11. Example of phase identification based upon AZtec LAM EDX maps, Cretaceous Colombian calcareous shale (LC224), Colombia.  A) - E) Individual element EDX maps for silica, calcium, aluminium, titanium and iron respectively.  F) BSE image of same area as in (A) – (E).  G) Phase map constructed from (A)-(E), illustrating four identified phases: Matrix (light blue), calcite (pale green), pyrite (red) and Ti-rich particles (dark blue).  See table 4 for percentage coverage figures of each phase.

Fig. 12. Example of heavy mineral analysis by AZtecFeature.  A) BSE image of a Pleistocene hemipelagite from IODP Expedition 339, off the Iberian Peninsula.  B) Particle map generated for the area illustrated in (A), with 698 particles analyzed: zircon (green), monazite (yellow), ilmenite (red) and rutile (purple).  Note pixel quality appears poor due to the need to save image as a screen dump.  C) Exploded view of typical heavy mineral particle.

Fig. 13. Example of iron sulphide analysis using AZtecFeature.  A), B) Distribution plots for pyrite (red) and sulphur-depleted pyrite (yellow), for respectively natural and artificially weathered samples of Blackstone Shale, Jurassic, Yorkshire.  C), D) Plots of thresholded particles (pyrite) with axes for sulphur and iron (weight %), (C) natural
sample, (D) artificially weathered. Colour coding as in (A), (B). E), F) BSE images
of partially weathered pyrite, exhibiting framboidal and euhedral forms respectively.
Brighter areas comprise ‘pristine’ pyrite, whereas darker surrounding areas are
relatively depleted in sulphur (as indicated in (D)).

Fig. 14. BSE images of organic matter from Cretaceous Colombian shale (P145) and
plots of oxygen versus carbon for selected organic particles analyzed by EDX using
AZtecFeature. A) Standard BSE image, B) as in (A) with individual organic particles
marked by AZtecFeature. C) Plot of oxygen versus carbon content for selected
organic particles, taken at 20kV in low vacuum. D) As in (C) but taken in high
vacuum, uncoated at 5kV. Both exhibiting a negative relationship between carbon and
oxygen (weight %) and a number of distinct trends. Red box in (D) indicates
equivalent area from (C).
**Table 1**: Illustrative timing for imaging of whole polished thin-sections (approx. 25 x 48 mm), using Maps software.

<table>
<thead>
<tr>
<th>Scan Rate (µs)</th>
<th>Pixel Resolution</th>
<th>Horizontal Tile Width (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2000</td>
<td>1000</td>
</tr>
<tr>
<td>5</td>
<td>768x512</td>
<td>0.22H</td>
</tr>
<tr>
<td></td>
<td>1536x1024</td>
<td>1.5H</td>
</tr>
<tr>
<td></td>
<td>3072x2048</td>
<td>6H</td>
</tr>
<tr>
<td></td>
<td>6144x4096</td>
<td>1D</td>
</tr>
<tr>
<td>10</td>
<td>768x512</td>
<td>0.45H</td>
</tr>
<tr>
<td></td>
<td>1536x1024</td>
<td>3H</td>
</tr>
<tr>
<td></td>
<td>3072x2048</td>
<td>12H</td>
</tr>
<tr>
<td></td>
<td>6144x4096</td>
<td>2D</td>
</tr>
<tr>
<td>20</td>
<td>768x512</td>
<td>1.5H</td>
</tr>
<tr>
<td></td>
<td>1536x1024</td>
<td>6H</td>
</tr>
<tr>
<td></td>
<td>3072x2048</td>
<td>1D</td>
</tr>
<tr>
<td></td>
<td>6144x4096</td>
<td>4D</td>
</tr>
<tr>
<td>50</td>
<td>768x512</td>
<td>3.75H</td>
</tr>
<tr>
<td></td>
<td>1536x1024</td>
<td>15H</td>
</tr>
<tr>
<td></td>
<td>3072x2048</td>
<td>2.5D</td>
</tr>
<tr>
<td></td>
<td>6144x4096</td>
<td>10D</td>
</tr>
</tbody>
</table>

H = hours, D = days, Y = years. Areas highlighted in pale green represent practical setting.
**Table 2:** Indication of achievable pixel resolution (nm) for different combinations of tile width (microns) versus tile dimensions (pixels); indicating the possibility of resolving structures between 10 – 100 nm. Colours indicate bands of similar tile resolution.

<table>
<thead>
<tr>
<th>Tile Dimensions (pixels)</th>
<th>Tile Width (µm)</th>
<th>2000</th>
<th>1000</th>
<th>518</th>
<th>259</th>
<th>130</th>
<th>60</th>
<th>30</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>6144x4096</td>
<td>325</td>
<td>163</td>
<td>84</td>
<td>42</td>
<td>21</td>
<td>10</td>
<td>5</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3072x2048</td>
<td>651</td>
<td>325</td>
<td>169</td>
<td>84</td>
<td>42</td>
<td>19</td>
<td>10</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>1536x1024</td>
<td>1302</td>
<td>651</td>
<td>337</td>
<td>169</td>
<td>85</td>
<td>39</td>
<td>19</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>768x512</td>
<td>2604</td>
<td>1302</td>
<td>674</td>
<td>337</td>
<td>169</td>
<td>78</td>
<td>39</td>
<td>13</td>
<td></td>
</tr>
</tbody>
</table>
Table 3. *Percentage phase occurrence calculated from grayscale image using image analysis (Image J).*

<table>
<thead>
<tr>
<th>Phase</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Silt</td>
<td>31.05</td>
</tr>
<tr>
<td>Clay</td>
<td>37.50</td>
</tr>
<tr>
<td>Calcite</td>
<td>18.38</td>
</tr>
<tr>
<td>Sulphides</td>
<td>0.13</td>
</tr>
<tr>
<td>Pores</td>
<td>12.92</td>
</tr>
</tbody>
</table>
Table 4. Percentage phase occurrence calculated using EDX maps collected from AZtec LAM.

<table>
<thead>
<tr>
<th>Phase</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Matrix (clay/silt)</td>
<td>79.16</td>
</tr>
<tr>
<td>Calcite</td>
<td>19.76</td>
</tr>
<tr>
<td>Pyrite</td>
<td>0.35</td>
</tr>
<tr>
<td>Ti rich</td>
<td>0.17</td>
</tr>
<tr>
<td>Unassigned</td>
<td>0.56</td>
</tr>
</tbody>
</table>
**Macro-microfabric analysis & composition**

- QUALITATIVE INTERPRETATION (descriptive analysis)
  - (BSE contrast variation)
  - (charge contrast imaging)

- MICRO FABRIC ANALYSIS (grain-size, orientation, porosity etc)

- 3D POROSITY / PERMEABILITY MODELLING

- IMAGE ANALYSIS OF MINERAL PHASES (through segmentation)

**Compositional analysis**

- QUALITATIVE DISTRIBUTION IN ELEMENTAL COMPOSITION
  - (clays, carbonates,)
  - (silicification ...)

- QUANTITATIVE DISTRIBUTION IN ELEMENTAL COMPOSITION (phase analysis)

**Particle analysis**

- PLOTS
  - (Fe sulphides Eh/pH)
  - (organics, C/O ratio)
  - (in situ heavy mineral analysis)
Figure 4

Click here to download Figure figure 4 final.jpg

Figure 3
Figure 8

A

B

C

D

E

F

G

POROSITY: 11.74%
PERMEABILITY:
X = 0.0124711 mD
Y = 0.00882547 mD
Z = 0.00730663 mD

H

POROSITY: 11.00%
PERMEABILITY:
X = 0.00876178 mD
Y = 0.0065001 mD
Z = 0.00644255 mD

I

POROSITY: 11.98%
PERMEABILITY:
X = 0.0196878 mD
Y = 0.0103445 mD
Z = 0.0102907 mD
Figure 10