

MATSAS: a small-angle scattering computing tool for porous systems

Amirsaman Rezaeyan,^{a*} Vitaliy Pipich^b and Andreas Busch^a

^aLyell Centre, Institute of GeoEnergy Engineering, Heriot-Watt University, Research Avenue South, Edinburgh EH14 4AS, United Kingdom, and ^bHeinz Maier-Leibnitz Zentrum (MLZ), Forschungszentrum Jülich GmbH, Jülich Centre for Neutron Science (JCNS), Lichtenbergstrasse 1, Garching 85748, Germany. *Correspondence e-mail: ar104@hw.ac.uk, amirsaman.rezaeyan@gmail.com

Received 2 September 2020

Accepted 20 January 2021

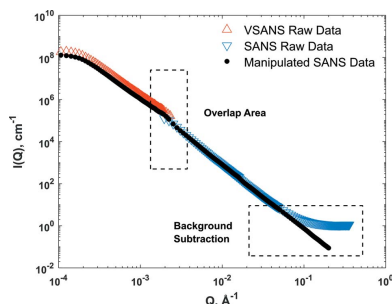
Edited by E. P. Gilbert, ANSTO, Kirrawee DC, Australia

Keywords: MATSAS; small-angle scattering; polydisperse spherical model; porous media; computer programs.

MATSAS is a script-based MATLAB program for analysis of X-ray and neutron small-angle scattering (SAS) data obtained from various facilities. The program has primarily been developed for sedimentary rock samples but is equally applicable to other porous media. *MATSAS* imports raw SAS data from .xls(x) or .csv files, combines small-angle and very small angle scattering data, subtracts the sample background, and displays the processed scattering curves in log–log plots. *MATSAS* uses the polydisperse spherical (PDSP) model to obtain structural information on the scatterers (scattering objects); for a porous system, the results include specific surface area (SSA), porosity (Φ), and differential and logarithmic differential pore area/volume distributions. In addition, pore and surface fractal dimensions (D_p and D_s , respectively) are obtained from the scattering profiles. The program package allows simultaneous and rapid analysis of a batch of samples, and the results are then exported to .xlsx and .csv files with separate spreadsheets for individual samples. *MATSAS* is the first SAS program that delivers a full suite of pore characterizations for sedimentary rocks. *MATSAS* is an open-source package and is freely available at GitHub (<https://github.com/matsas-software/MATSAS>).

1. Introduction

Small-angle scattering (SAS) of neutrons and X-rays (SANS and SAXS, respectively) is widely used for the nondestructive study of the low-resolution structure of natural and engineered systems, including sedimentary rocks, biological macromolecules, composite nanomaterials and polymers on length scales between ångströms and micrometres in a single or combined experiment (Feigin & Svergun, 1987; Binder *et al.*, 2000; Zemb & Lindner, 2002; Radlinski, 2006; Borsali & Pecora, 2008; Anovitz & Cole, 2015; Melnichenko, 2015; Fritzsche *et al.*, 2016). Advances in SAS instrumentation, such as neutron and high-flux X-ray synchrotron beamlines, have significantly increased the use of SANS and SAXS experiments (Melnichenko, 2015; Zemb & Lindner, 2002; Heenan *et al.*, 1997). With the availability of these technologies, modern instruments can provide high-quality data in time- or space-resolved experiments or measurements under various physical and chemical conditions, such as temperature, pressure, humidity *etc.* (Konarev *et al.*, 2006; Schrank *et al.*, 2020). The theoretical and methodological developments obtained over the past few decades have allowed the retrieval of structural information from SAS patterns to address questions revolving around the size, shape, distribution and orientation of scatterers (scattering objects) (Konarev *et al.*, 2006; Petoukhov *et al.*, 2012).



OPEN ACCESS

Neutron and X-ray scattering techniques complement each other, but neutrons and X-rays are different in their charge, energy and interaction with matter, which makes each technique subject to its own experimentation type and/or sample type (Binder *et al.*, 2000; Zemb & Lindner, 2002; Melnichenko, 2015). Fig. 1 illustrates a pinhole SAS experiment. Neutrons or X-rays are collimated and monochromated towards the sample, inside which a neutron or photon is elastically scattered from its wavevector $\mathbf{k}_0$ into a state with wavevector $\mathbf{k}$ under a scattering angle 2θ . The magnitude of a wavevector relates to its wavenumber, which is $|\mathbf{k}| = |\mathbf{k}_0| = k = 2\pi/\lambda$ for elastic scattering, where λ is the neutron or X-ray wavelength. The intensity of the scattered radiation dI is therefore measured in the direction $\mathbf{k}$ as a function of the momentum transfer (the convention $\mathbf{s} = |\mathbf{k} - \mathbf{k}_0|$) or the scattering vector $\mathbf{Q}$. The magnitude of the scattering vector is given by $Q = 4\pi \sin \theta/\lambda$, from which it follows that $\mathbf{Q} = 2\pi\mathbf{s}$, where $s = 2 \sin \theta/\lambda$ (Radlinski, 2006; Melnichenko, 2015).

The incident flux of the scattering objects is denoted by Φ_0 , *i.e.* $\Phi_0 = I_0/A$, where I_0 is the incident intensity (neutrons or X-rays per second) and A is the beam cross-sectional area at the sample position (Radlinski, 2006). The scattered intensity monitored in the solid-angle element $d\Omega$ targeted by $\mathbf{Q}$ can be expressed as

$$dI \propto \Phi_0 \frac{d\Sigma}{d\Omega} d\Omega, \quad (1)$$

where $d\Sigma$ is the elemental scattering cross section. The quantity $d\Sigma/d\Omega$ is called the differential cross section of scattering (Radlinski, 2006). The aim of SAS experiments is to determine volume-averaged information on the spatial distribution of the scattering length density (neutrons) or electron density (X-rays) in the sample from the measured $d\Sigma/d\Omega$ as a function of the scattering vector magnitude Q , thus $(d\Sigma/d\Omega)(Q)$ or $I(Q)$ (Melnichenko, 2015).

For a wide range of substances, SAS data for hard and soft matter can generally be interpreted accurately using a two-phase approximation (Melnichenko, 2015). In this approximation, the scattering volume is viewed as being composed of above-molecular-size phases, each characterized by one of two possible values of the physical property that provides the scattering contrast ($\Delta\rho^*$). For instance, for porous media, these two phases are the solid matrix (phase 1) and the pore space (phase 2) (Radlinski, 2006). The two-phase approximation is a simplification inherent in the SAS method and has

been implicitly or explicitly employed for many years. As such, the general expression of the scattering cross section can be expressed as

$$I(Q) = NV_p^2(\rho_1^* - \rho_2^*)^2 P(Q) S(Q) + B, \quad (2)$$

where N is the number density of scatterers N_p per unit volume, V_p is the volume of the scatterers, and ρ_1^* and ρ_2^* are the scattering length/electron density of phase 1 and phase 2, respectively. B is the sample background, accounting for scattering in the high- Q limit. The high- Q background originates from (i) Q -independent incoherent scattering caused by hydrogen atoms in organic matter and/or water, and (ii) Q -dependent coherent scattering resulting from microscopic inhomogeneities (*e.g.* small pores in the rock matrix; Bahadur *et al.*, 2015; Blach *et al.*, 2020). $P(Q)$ is the form factor which describes the size and shape of the scatterer. There are analytical expressions for the form factor for simple geometrical objects like spheres, cylinders, discs or parallelepipeds (Melnichenko, 2015). $S(Q)$ is the structure factor and contains information about the spatial distribution of the scatterers. The structure factor represents the modification of the intensity due to the spatial correlation of the scatterers (Fritzsche *et al.*, 2016), where the positions of the scatterers are frozen in time and space in solid porous materials (Melnichenko, 2015). In soft-matter systems, the interaction potential between scatterers is also taken into consideration (Melnichenko, 2015). Form and structure factors need to be specified to determine the structural information in the scattering curves.

Several SAS programs have been developed in different laboratories that consider various data processing and manipulation methods, fitting models, and form and structure factors to characterize the structure of the scatterers (Table 1). Recognizing the increasing application of SAS data to analyse the pore structure of sedimentary rocks, especially low-permeability rocks such as coal and mudrocks or gas shales (Radlinski, Ioannidis *et al.*, 2004; Radlinski, Mastalerz *et al.*, 2004; Radliński *et al.*, 2009; Mares *et al.*, 2012; Clarkson *et al.*, 2012; Mastalerz *et al.*, 2012; Melnichenko *et al.*, 2012; Bahadur *et al.*, 2014, 2015; Anovitz *et al.*, 2015; Leu *et al.*, 2016; Busch *et al.*, 2017, 2018; Anovitz & Cole, 2018; Sakurovs *et al.*, 2018; Vishal *et al.*, 2019; Blach *et al.*, 2020), we have developed the program package *MATSAS*. It allows the analysis of data obtained from small-angle and very small angle scattering of neutrons and X-rays [very small angle neutron scattering (VSANS), small-angle neutron scattering (SANS), wide-angle X-ray scattering (WAXS), ultra-small-angle X-ray scattering (USAXS) and small-angle X-ray scattering (SAXS)].

MATSAS analyses data from pinhole-geometry, time-of-flight (TOF) and Bonse–Hart machines and was tested using data acquired at FRM-II (Research Reactor Munich II, Garching, Germany) and ORNL (Oak Ridge National Laboratory, Tennessee, USA) (Rezaeyan, Pipich *et al.*, 2019a,b; Rezaeyan, Seemann *et al.*, 2019; Seemann *et al.*, 2019). *MATSAS* does post-processing of data obtained from research facilities. It is assumed that initial corrections for sample thickness, transmission, detector sensitivity, instrument

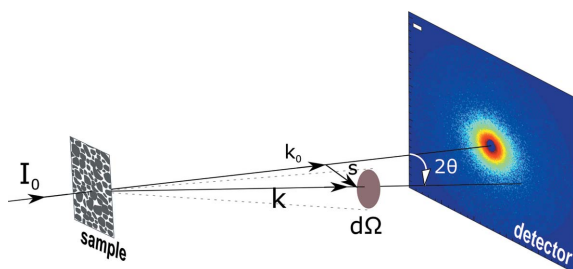


Figure 1
The schematic principle of a SAS experiment.

Table 1

Common SAS programs and their capabilities and applicabilities.

SAS program	Capabilities	Applicability	Reference
<i>FIT2D</i>	2D image data reduction/manipulation and peak fitting		Hammersley (1995, 2016)
<i>BerSANS</i>	Data acquisition/reduction		Keiderling (1997)
<i>DALAI_GA</i>	<i>Ab initio</i> shape determination	Biological systems	Chacón <i>et al.</i> (1998)
<i>FISH</i>	Peak analysis and parametric fitting using various form and structure factors		Heenan (1999)
<i>SAX3D</i>	<i>Ab initio</i> shape determination	Biological systems	Walther <i>et al.</i> (2000)
<i>SAXS/WAXS</i> software system	Data acquisition/reduction		Homan <i>et al.</i> (2001)
<i>GRASP</i>	Data acquisition/reduction		Dewhurst (2002)
<i>SAXSANA</i>	Data reduction, <i>Q</i> determination, data conversion, data correction, analysis of time-resolved data and data extrapolation	Biological systems	Hiragi <i>et al.</i> (2003)
<i>PRINSAS</i> †	Fitting of 1D curves using a spherical form factor for a polydisperse scattering system	Porous systems	Hinde (2004)
<i>ATSAS</i>	Data reduction, data processing and 3D modelling	Biological systems	Konarev <i>et al.</i> (2006), Petoukhov <i>et al.</i> (2012), Franke <i>et al.</i> (2017), Manalastas-Cantos <i>et al.</i> (2021)
<i>DAMMIF</i>	<i>Ab initio</i> shape determination for disordered systems and solutions	Nanostructures	Franke & Svergun (2009)
<i>IRENA</i> †	Plotting SAS data, merging of two overlapping data sets, and fitting form and structural models to data from contrast variation experiments	A wide range of systems	Ilavsky & Jemian (2009)
<i>BioXTAS RAW</i>	Isotropic SAXS data reduction, primary data analysis and calculations of the pair-distance distribution functions, averaging, subtraction and analysis of radius of gyration and molecular weight, calculation of inverse Fourier transforms and envelopes, processing of inline size-exclusion chromatography coupled SAXS data, and data deconvolution	Biological systems	Nielsen <i>et al.</i> (2009), Hopkins <i>et al.</i> (2017)
<i>SCATTER</i>	2D data analysis	Nano- and mesoscale oriented structures	Förster <i>et al.</i> (2010)
<i>SAAF</i>	SANS data analysis using a set of standard models	Polymers	Zhao (2011)
<i>SASTBX</i>	Data reduction, model reconstruction, model refinement and shape retrieval	Biological systems	Liu <i>et al.</i> (2012)
<i>SASET</i>	1D and 2D data analysis and fitting of data using scattering models and anisotropy methods	Anisotropic structures	Muthig <i>et al.</i> (2013)
<i>MolScat</i> and <i>SAFIR</i>	Modelling of 3D macromolecular structures	Biological systems	Hofmann & Whitten (2014)
<i>SASfit</i>	Reduction of oversampled data sets, confidence assessment of optimized model parameters and availability of custom user-provided models	Polymers	Breßler <i>et al.</i> (2015)
<i>QtiSAS/QtiKWS</i> †	Graphical visualization, reduction, analysis and fitting of data using various scattering models	A wide range of systems	https://www.qtisas.com/
<i>SASview</i> †	Data reduction, manipulation and analysis using several form and structure factors with polydispersity and orientational distributions	A wide range of systems	http://www.sasview.org/

† *PRINSAS*, *IRENA*, *QtiSAS* and *SASview* are commonly used for analysis of SAS data obtained from porous systems featuring a wide range of pore sizes.

background, multiple scattering and noise have been made using the instrument-specific settings at the facility itself, providing data in absolute units (Hinde, 2004; Melnichenko, 2015). *MATSAS* is primarily oriented towards the structural analysis of sedimentary rocks using a polydisperse spherical (PDSP) model. The *MATSAS* software is constantly refined to broaden its functionality, making it applicable to isotropic and partially ordered objects such as biological nanoparticle systems, colloidal solutions, and polymers in solution and bulk. It is an open-source computer tool for academic users and is freely available on GitHub (<https://github.com/matsas-software/MATSAS>). Open-source access reflects transparency in the fundamental assumptions and solving approaches employed in the program and allows third parties to interface their in-house programs with the data analysis framework of the program (Liu *et al.*, 2012) and to help in accelerating its development. In this paper, we summarize the main components of *MATSAS* and its development framework.

2. Program overview

MATSAS features a script-based package in MATLAB (The MathWorks Inc., Natick, MA, USA), which integrates

computation and visualization in an easy-to-use environment. The *MATSAS* program is a versatile computer tool allowing both users and developers to add additional tools and develop specific novel applications. The flexible user-friendly framework of *MATSAS* for basic routines, such as intensity calculation or model alignment, allows anyone with basic programming skills to improve or adapt *MATSAS* to better reflect user-specific needs. Furthermore, the current version of the package includes the PDSP model to analyse SAS data in terms of theoretical intensity computation, the $f(r)$ probability function of pore size distribution and model refinement. The PDSP model is the method commonly used for SAS analysis of a polydisperse system of randomly oriented independently scattering particles and is ubiquitous for fractal microstructures (e.g. sedimentary rocks) as well as other porous systems (Radlinski, Ioannidis *et al.*, 2004), provided that the particle-shape distribution is independent of the distribution of particle dimensions in the polydisperse system (Schmidt, 1982). The script-based *MATSAS* code allows parameters to be tuned for more features of each routine.

Use of the *MATSAS* program is divided into three steps: (i) pre-processing of raw or facility post-corrected SAS and

very small angle scattering (VSAS) data as well as physical information, (ii) processing of the imported information to produce $I(Q)$ versus Q curves, combine the SAS and VSAS curves, and fit the PDSP model, and (iii) post-processing to display and export structural information obtained from the samples being analysed. Fig. 2 illustrates the main components of the present version of *MATSAS*.

Detailed instructions on how to use the package are available on GitHub. Supporting information and command descriptions are embedded in each module. Errors and bugs can be invoked when no parameters or incorrect data are given to the command. We developed the package in Windows and recommend running it in Windows, Mac or Linux, with any Intel or AMD x86-64 processor with four logical cores and AVX2 instruction set support, as a minimum. Although the program runs satisfactorily without a specific graphics card, a hardware-accelerated graphics card supporting OpenGL 3.3 with 1 GB graphical processing unit (GPU) memory is recommended, as displaying figures and generating Microsoft *Excel* worksheets require more background processing.

3. Data pre-processing

The data pre-processing module is composed of two components: (i) data are prepared in a Microsoft *Excel* spreadsheet [**.xls(x)*] or **.csv* file, including (V)SAS data, neutron scattering length densities or X-ray electron densities of phases 1 and 2 (*e.g.* rock matrix and pore), grain density of the sample, data reduction limits (optional), and sample name, and (ii) the MATLAB *data_input.m* file reads and stores the imported data for the next step. *MATSAS* allows users to run a batch of samples. The units in the input files can be converted between different unit systems (between nm^{-1} and $\text{\AA}^{-1}$ for Q , for instance) by changing appropriate codes. The range(s) of data points can be adjusted for each data set individually or simultaneously for selected groups of files.

4. Data processing

The data processing module is used to manipulate and analyse the information imported. The primary data processing script is developed to manipulate scattering curves. The *data_manipulation.m* program carries out multiple tasks, including $I(Q)$ data sorting, curve fitting, background subtraction, curve merging, curve smoothing and raw data reduction. The secondary data processing script file in *data_analysis.m* is designed to analyse $I(Q)$ – Q curves and produce structural information. An arbitrary size distribution is created first and the PDSP model is then fitted to the processed scattering curve. Pore characteristics are predicted and fractal dimensions (including the pore fractal dimension, D_p , surface fractal

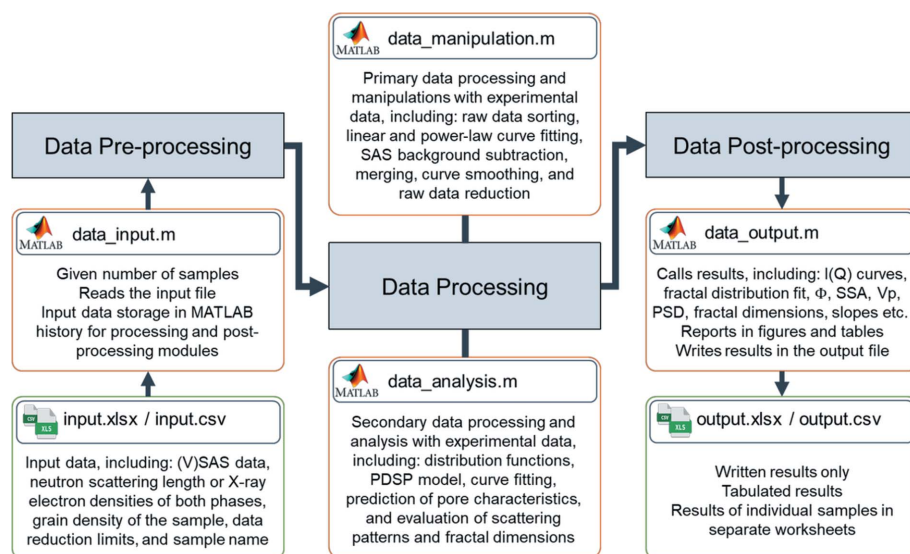


Figure 2
A schematic flow chart of *MATSAS* programs and their functionalities.

dimension, D_s , and general fractal dimension, D_f) are evaluated from the fit in this module.

4.1. Data manipulation

The program *data_manipulation.m* is a data processing module encompassing major SAS data processing steps for isotropic systems, from merging of scattering curves to background reduction. This program performs manipulations with one-dimensional data sets and calls other analysis and fitting programs via user-defined or built-in function files. The SAS data might have been collected at different sample-to-detector distances. Once data from several experimental curves have been combined for one specific instrument (*e.g.* SANS), they may not be sorted, which leads to numerical problems in further analysis. Data sorting is therefore carried out in the data manipulation package using a built-in function. If the SAS data consist of two scattering profiles obtained from two different instruments (*e.g.* VSANS and SANS), *MATSAS* allows users to merge the two curves using a least-squares fit in the overlapping range, as illustrated for example in Fig. 3. The SAS curve is the basis onto which the VSAS curve is rebinned. The high- Q background is subtracted using equation (3),

$$\frac{d\Sigma}{d\Omega}(Q) = A Q^{-a} + \left[\frac{d\Sigma}{d\Omega}(Q) \right]_{\text{inc}}, \quad (3)$$

where the scattering varies with Q^{-a} in the high- Q limit before plateauing (Melnichenko, 2015). The value of the background $[(d\Sigma/d\Omega)(Q)]_{\text{inc}}$ is determined from a linear plot of equation (4),

$$Q^a \frac{d\Sigma}{d\Omega}(Q) = A + Q^a \left[\frac{d\Sigma}{d\Omega}(Q) \right]_{\text{inc}}, \quad (4)$$

where $[(d\Sigma/d\Omega)(Q)]_{\text{inc}}$ is the slope and A is the intercept (Melnichenko, 2015). Fig. 3 shows the background subtraction

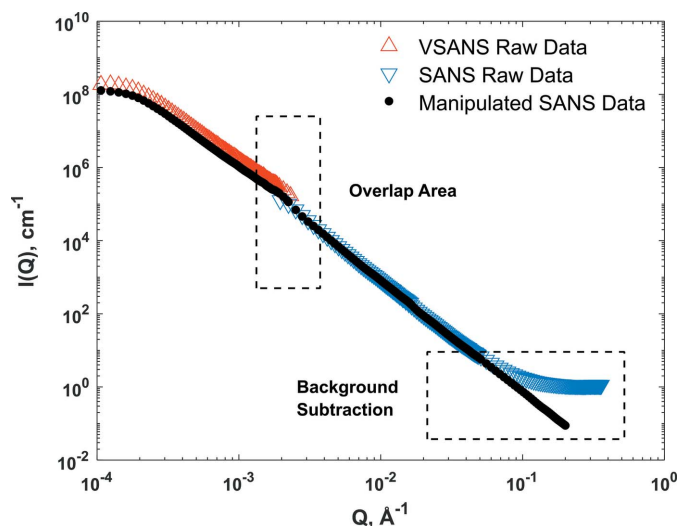


Figure 3

SANS data manipulated and processed on an arbitrary mudrock sample. Red, blue and black curves are the scattering profiles from the VSANS and SANS instruments and the net scattering after manipulation (merging, background subtraction and smoothing), respectively.

in the high- Q limit for a range that users can change manually in the program.

A noise-removal operation is embedded to remove the sparse data around the beam stop or detector edge. Raw data reduction, whose cut-off limits are determined in the data input files, is carried out as well. Two data smoothing operations are included in the package, which can be employed to obtain a smooth scattering profile for further structural analysis. Fractal dimensions and slope are determined here. For all operations, the propagation of uncertainty is performed using standard equations (Bevington & Robinson, 2003). A MATLAB plotting operation displays the currently active scattering profiles on a log-log scale. An advanced plotting option included in the plot permits users to change the plotting range, enlargement factor *etc.* The data manipulation file contains an output section, where the result of each operation can be used in subsequent data analysis. Information about the operation (type of operation, section names, functions, weights, ranges of points used *etc.*) is written in the package in green and allows modification or change of lines if needed.

4.2. Data analysis

The data analysis program calculates the intensity of SAS from a polydisperse system of scatterers (Porod, 1951, 1952; Guinier & Fournet, 1955). The intensity is expressed in terms of a fractal distribution of scatterers, also called the probability density of the pore size distribution $f(r)$, for greater numerical stability (Ilavsky & Jemian, 2009). SAS curves from sedimentary rocks are usually linear on a log-log scale, particularly in the high- Q region, which reflects fractal behaviour (Melnichenko, 2015). Scattering from a fractal surface is equivalent to scattering from a system of polydisperse spherical scatterers (Schmidt, 1982), with a number-size distribution (the number of spheres with radii between R and $R + dR$) given by

$$f(r)dr \simeq R^{-(1+D_f)} dR, \quad (5)$$

where D_f is the fractal dimension determined from the slope of the power-law scattering (Melnichenko, 2015). In practice, the distribution described in equation (5) and in the range $R_{\min} \leq R \leq R_{\max}$ shows fractal behaviour between the upper and lower cut-off parameters. $f(r)$ is expressed as

$$f(r) = \frac{D_f}{R_{\min}^{-D_f} - R_{\max}^{-D_f}} R^{-(1+D_f)}. \quad (6)$$

This is valid for $R_{\max} > R_{\min} > 0$ and $D_f \in (-1, \infty)$, where $D_f = 6 + \text{slope}$. Scattering from a PDSP featured sample has a linear region with a similar slope $-(1 + D_f)$ and is described by (Radlinski, Ioannidis *et al.*, 2004)

$$I(Q) = \int_{R_{\min}}^{R_{\max}} (\rho_1^* - \rho_2^*)^2 N f(r) V^2 P(Q, r) dr, \quad (7)$$

where $V \equiv V(r) = \frac{4}{3}\pi r^3$ is the volume of a sphere of radius r (volume of a scatterer). In addition, $P(Q, r)$ is the form factor of a sphere of radius r (Guinier & Fournet, 1955):

$$P(Q, r) = \left[3 \frac{\sin(Q, r) - Qr \cos(Q, r)}{Q^3 r^3} \right]^2. \quad (8)$$

N is the total number of scatterers, which is related to the number size distribution as $N(r) = Nf(r)$. $N(r)$ is expressed as

$$N(r) = \frac{\Phi}{\bar{V}(r)} f(r) = \frac{IQ_0}{(\rho_1^* - \rho_2^*)^2} \frac{1}{\bar{V}^2(r)}, \quad (9)$$

where

$$IQ_0 = f(r) (R_{\max} - R_{\min}) \Phi \frac{(\rho_1^* - \rho_2^*)^2}{\bar{V}(r)} \quad (10)$$

is the scattering intensity at $Q = 0$ and $\bar{V}(r) = \int_{R_{\min}}^{R_{\max}} V(r) f(r) dr$ is the average volume of the scatterers (Radlinski *et al.*, 2002). Similarly to the approach of Ilavsky & Jemian (2009), MATSAS calculates equation (7) throughout the integration over a continuous size distribution with a summation over a discrete size histogram:

$$I(Q) = \sum_i (\rho_1^* - \rho_2^*)^2 \sum_{i,j} N_i f_i(r_{i,j}) V_i(r_{i,j})^2 P_i(Q, r_{i,j}) \Delta r_{i,j}, \quad (11)$$

where the subscript i represents different scattering sizes and the subscript j describes bins in the size distribution. $\Delta r_{i,j}$ is the width of bin j and each scattering size has its own binning index i, j . r is the dimension of the scatterer (radius for spheres) and has limits $r_{\max,i}$ and $r_{\min,i}$. The radius r is calculated using $R = 2\pi/Q$, which is $R = 2.5/Q$ in the fractal distribution (Radlinski *et al.*, 2000).

MATSAS uses an arbitrary size distribution to model the scattering volume distribution $V^2(r)P(Q, r)$ and determine $f(r)$. The user can change the theoretical ranges of the various size distributions in the data analysis program. Numerical calculations call limits on the range of dimensions ($r_{\min}$ and $r_{\max}$), the cut-off limits ($R_{\min}$ and $R_{\max}$) and the number of bins (N_{bin}). This method results in a natural logarithmic step in

dimension and uses three parameters, $R_{\min}$, $R_{\max}$ and N_{bin} . The centres of the first ($r_{i,1}$) and last ($r_{i,N_{\text{bin}}}$) bins are $R_{\max}$ and $R_{\min}$, respectively, and extra fractional volumes are discarded for both bins: the volumes associated with $r_{\min,i,1} - r_{i,1}$ and $r_{i,N_{\text{bin}}} - r_{\max,i,N_{\text{bin}}}$ for the first and last bins, respectively. The widths of the bins are equal by selecting associated dimensions at regular increments of the cumulative distribution (Ilavsky & Jemian, 2009), leading to

$$\log(\Delta) = \frac{\log(R_{\max}) - \log(R_{\min})}{N_{\text{bin}}}. \quad (12)$$

However, the numerical operation of the *data_analysis.m* file requires $r_{\min,i,j}$, $r_{i,j}$, $r_{\max,i,j}$, $f_i(r_{i,j})$ and IQ_0 to fit the PDSP model in equation (7) to the measured $I(Q)$ curve. The fitting procedure employs $f(r)$ and IQ_0 as fitting parameters for each iteration to attain a match where the summation of square errors (SSQ) tends to a minimum (Hinde, 2004).

To reduce the computation time taken by numerical integration, we found an analytical solution for the scattering volume distribution

$$\int_{R_{\min}}^{R_{\max}} V^2 P(Q, r) dr = \sum_{i,j} \frac{1}{Q_i^7} \left(16\pi^2 \left\{ -\frac{5}{4} \sin(Q_i r_{i,j}) \cos(Q_i r_{i,j}) + \frac{3}{2} Q_i r_{i,j} \cos(Q_i r_{i,j})^2 + Q_i^2 r_{i,j}^2 \left[\frac{1}{2} \sin(Q_i r_{i,j}) \cos(Q_i r_{i,j}) + \frac{1}{2} Q_i r_{i,j} \right] - \frac{1}{4} Q_i r_{i,j} - \frac{1}{3} Q_i^3 r_{i,j}^3 \right\} \right) \Delta r_{i,j} \quad (13)$$

that transforms equation (11) into

$$I(Q) = \sum_i (\rho_{1i}^* - \rho_{2i}^*)^2 \sum_{i,j} N_i f_i(r_{i,j}) \times \frac{1}{Q_i^7} \left(16\pi^2 \left\{ -\frac{5}{4} \sin(Q_i r_{i,j}) \cos(Q_i r_{i,j}) + \frac{3}{2} Q_i r_{i,j} \cos(Q_i r_{i,j})^2 + Q_i^2 r_{i,j}^2 \left[\frac{1}{2} \sin(Q_i r_{i,j}) \cos(Q_i r_{i,j}) + \frac{1}{2} Q_i r_{i,j} \right] - \frac{1}{4} Q_i r_{i,j} - \frac{1}{3} Q_i^3 r_{i,j}^3 \right\} \right) \Delta r_{i,j}. \quad (14)$$

MATSAS simplifies the intensity calculation by substituting equation (9) into equation (14), leading to

$$I(Q) = \sum_{i,j} IQ_0 f_i(r_{i,j}) \bar{V}_i^{-2}(r_{i,j}) \times \frac{1}{Q_i^7} \left(16\pi^2 \left\{ -\frac{5}{4} \sin(Q_i r_{i,j}) \cos(Q_i r_{i,j}) + \frac{3}{2} Q_i r_{i,j} \cos(Q_i r_{i,j})^2 + Q_i^2 r_{i,j}^2 \left[\frac{1}{2} \sin(Q_i r_{i,j}) \cos(Q_i r_{i,j}) + \frac{1}{2} Q_i r_{i,j} \right] - \frac{1}{4} Q_i r_{i,j} - \frac{1}{3} Q_i^3 r_{i,j}^3 \right\} \right) \Delta r_{i,j}. \quad (15)$$

Once the match is reached, the data analysis program yields the structural characteristics of the scatterers using the fitted $f(r)$ and IQ_0 values. The specific surface area (SSA) of the scatterers is obtained following Hinde (2004):

$$\text{SSA} = \frac{1}{\rho_g (\rho_1^* - \rho_2^*)^2} \sum_k \frac{4}{3} \pi IQ_0 f(r_k) \bar{V}^{-2}(r_k) \Delta r_k, \quad (16)$$

where the subscript k represents bins in the size distribution. The volume fraction of scatterers per unit volume (Φ) is calculated from equation (9), which results in

$$\Phi = \frac{1}{(\rho_1^* - \rho_2^*)^2} \sum_k IQ_0 \bar{V}(r_k) \bar{V}^{-2}(r_k), \quad (17)$$

and the total volume of scatterers (V_p) is obtained by

$$V_p = \frac{1}{\rho_g} \sum_k \frac{\Phi(r_k)}{1 - \Phi(r_k)}, \quad (18)$$

where the subscript k represents bins in the size distribution. Differential (dV/dr or dA/dr) and logarithmic differential ($dV/d\log r$ or $dA/d\log r$) scatterer size distributions are calculated cumulatively (Meyer & Klobes, 1999).

The scattering intensity decays as Q^{-m} with different power-law exponents m ; this indicates that m is related to the dimensionality of the pore as understood in terms of the concept of fractality (Mandelbrot, 1983). For a fractal pore scatterer, therefore, $D_p = m$ with values $1 < D_p < 3$, and for a surface fractal $D_s = 6 - m$ with values $2 \leq D_s \leq 3$ (Bale & Schmidt, 1984).

The scattering at different length scales indicates the Guinier, mass/pore fractal, surface fractal and Porod regions, suggesting that each fractal region is limited to a specific range of scattering vectors (Fritzsche *et al.*, 2016). Therefore, for sedimentary rocks D_p and D_s are geared to the ranges of $0.0003\text{--}0.003\text{ cm}^{-1}$ and $0.003\text{--}0.03\text{ cm}^{-1}$, respectively. In addition, D_f is included to reflect the fractality of the full pore system over the entire scattering vector range, *e.g.* $0.0003\text{--}0.03\text{ cm}^{-1}$ in sedimentary rocks (Rezaeyan, Pipich *et al.*, 2019a,b; Rezaeyan, Seemann *et al.*, 2019). These ranges can be changed by the user.

For demonstration purposes, we tested the analysis operations on SANS and VSANS data obtained from three rock samples (Opalinus Clay) using batch mode. Opalinus Clay is a Jurassic mudrock that was obtained from the Mont Terri Underground Laboratory in Switzerland and has been described in detail previously (Busch *et al.*, 2017). Fig. 4(a) shows the PDSP modelled $I(Q)$ curves and the measured $I(Q)$ curves after two iterations of the fitting operation. The first iteration starts with initial guesses for $f(r)$ and IQ_0 , which are obtained from the slope of the scattering curves and the Guinier & Fournet (1955) approximation, respectively. SSQ tends to a minimum after the second iteration; two iterations are recommended for most rock samples (Hinde, 2004). Fig. 4(b) shows $f(r)$ after two iterations on a log–log scale. $f(r)$ levels off at scatterer sizes $\gtrsim 2\text{ }\mu\text{m}$ because the scattering intensities of large scatterers are smeared, possibly due to instrument artefacts at the edge of the detector. The error

sensitivity, expressed as $dSSQ/d\log(IQ_0)$, relates SSQ to the number of iterations [Fig. 4(c)]. The value of $dSSQ/d\log(IQ_0)$ varies around zero for all scatterer sizes. However, as illustrated in Fig. 4(c), this can deviate where the fit is rather poor for large scatterer sizes ($3 < \log D < 3.5$) due to different

instrument resolutions or noise within overlap areas. SSQ is magnified when the number of iterations exceeds two, resulting in an attenuation of $f(r)$. Nevertheless, we recommend attaining a smooth $f(r)$ if the optimum fit requires a larger number of iterations for a specific sample. χ^2 tolerance

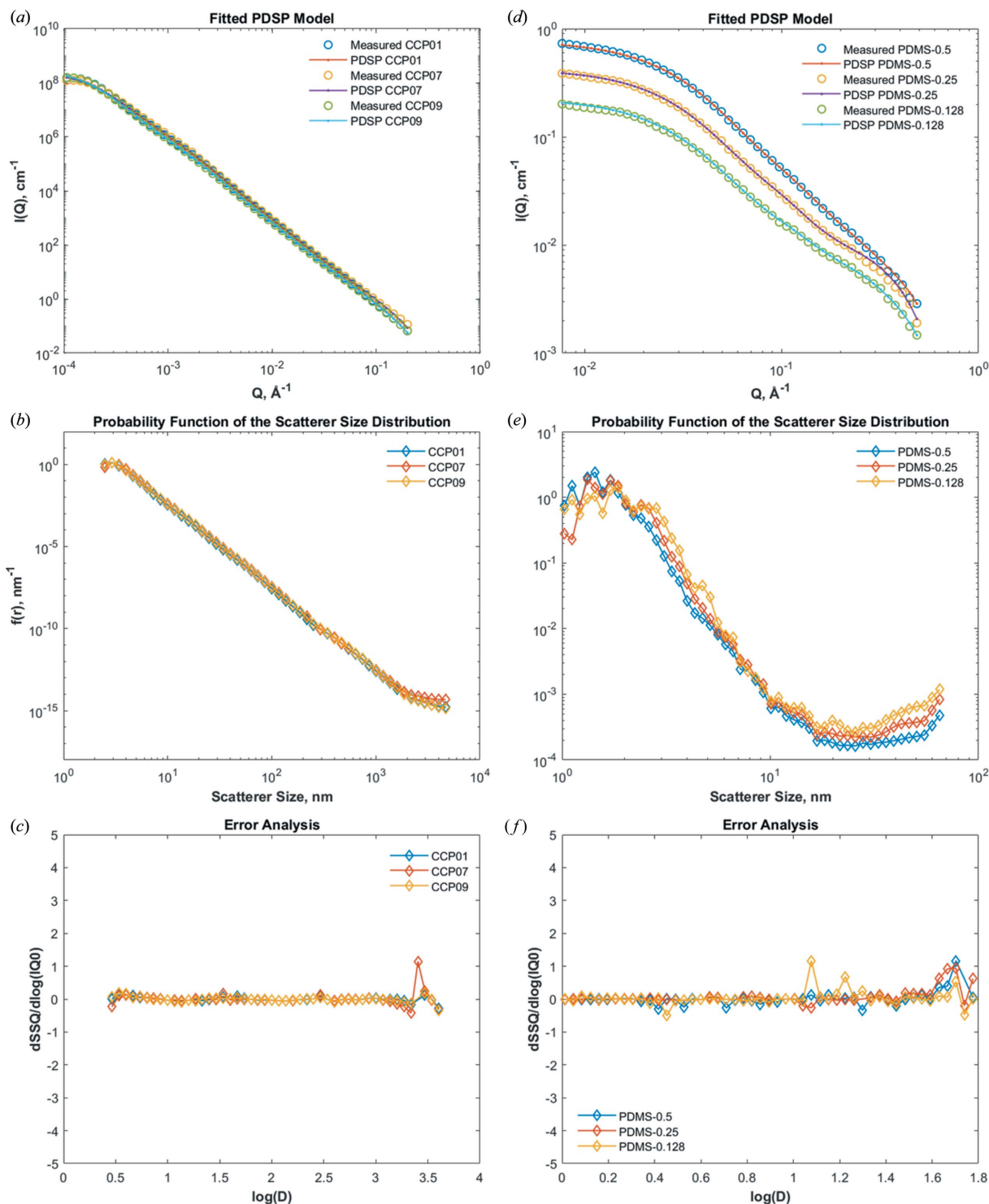


Figure 4

The PDSP model applied to SANS data obtained from three rock samples (Opalinus Clay) and three polydimethylsiloxane (PDMS) polymers of volume fractions 0.128, 0.25 and 0.5 in toluene. Rock samples: (a) measured $I(Q)$ curves after manipulation and $I(Q)$ curves obtained from the PDSP model, (b) probability functions of the pore size distribution $f(r)$, and (c) the error sensitivity $dSSQ/d\log(IQ_0)$ obtained after two iterations. PDMS samples: (d) the fitted PDSP model, (e) probability functions of the scatterer size distribution $f(r)$ and (f) the error sensitivity obtained after 20 iterations.

can be used for the fit when the user has no preference for the number of iterations.

We also tested the PDSP model on three polydimethylsiloxane (PDMS) polymers with volume fractions of 0.128, 0.25 and 0.5 in toluene to demonstrate the applicability of the fitting operation for a non-power-law nanostructure in solution [Figs. 4(d)–4(f)]. Fig. 4(f) displays the numerical flexibility of the fitting procedure after 20 iterations.

5. Data post-processing

The data post-processing module is made of two components, including *data_output.m* in MATLAB and the results reported in figures and tabulated files. The *data_output.m* file calls the results of individual samples, reports the results in figures and tables in the MATLAB command window, and writes the results in *output.xlsx*. The results include measured, processed and predicted scattering curves, fractal distribution fit (f_r), specific surface area (SSA), porosity (Φ), pore volume (V_p), pore size distribution (PSD) by pore volume or pore area, fractal dimensions, slopes of scattering curves, pore characteristics divided into macro-, meso- and micropores, and

Table 2

Slope (m), fractal dimensions (D), incoherent background (I_{BG}) and pore characteristics evaluated by *MATSAS* from the SANS data for three rock samples.

The subscripts meso and macro represent properties in meso- and macropore sizes, respectively.

Sample ID	m	D_t	D_s	D_p	I_{BG}^{VSAS} (cm ⁻¹)	I_{BG}^{SAS} (cm ⁻¹)	SSA (m ² g ⁻¹)	SSA _{macro} (m ² g ⁻¹)	SSA _{meso} (m ² g ⁻¹)
CCP01	-3.06	2.94	2.88	2.84	15317	1.15	31.6	1.4	30.2
CCP07	-3.05	2.95	2.88	2.76	7032	1.23	44.8	1.6	43.1
CCP09	-3.07	2.93	2.88	2.86	14818	0.92	29.6	1.2	28.5

Sample ID	V_p (cm ³ g ⁻¹)	V_{macro} (cm ³ g ⁻¹)	V_{meso} (cm ³ g ⁻¹)	Φ (%)	Φ_{macro} (%)	Φ_{meso} (%)	SSQ	χ^2
CCP01	0.0880	0.0541	0.0339	23.7	14.6	9.1	0.01	0.003
CCP07	0.1036	0.0601	0.0435	28.0	16.3	11.8	0.09	0.006
CCP09	0.0773	0.0472	0.0301	21.1	12.9	8.2	0.01	0.005

background subtraction values. Some results for the Opalinus Clay samples are shown in Fig. 5 and Table 2. The results for individual samples are produced and saved in figure formats (*.tif and *.emf), *Excel* spreadsheets and *.csv* files for

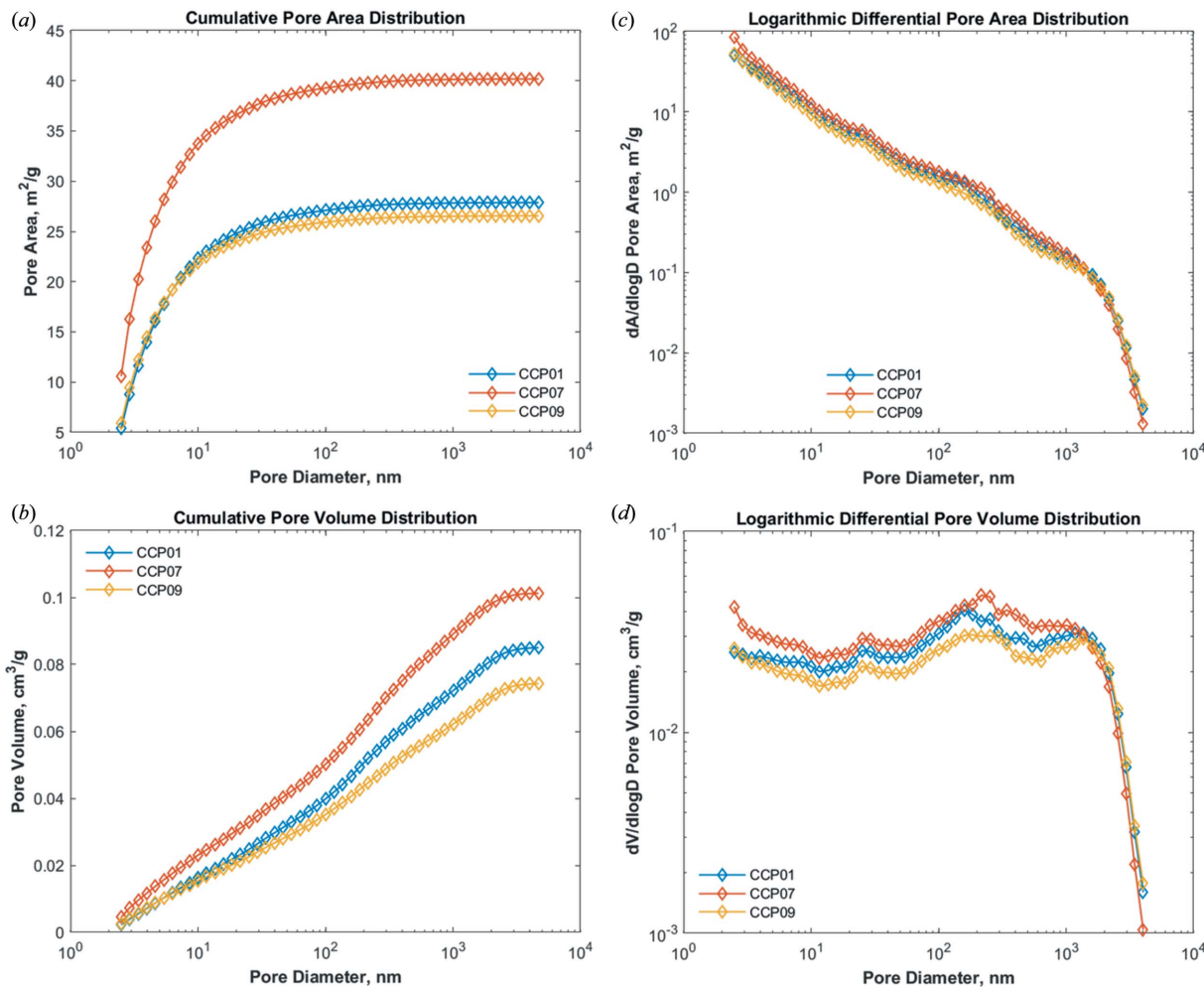


Figure 5 The PDSP model applied to SANS data obtained from three rock samples (Opalinus Clay). (a) Cumulative pore area distribution, (b) logarithmic differential pore area distribution, (c) cumulative pore volume distribution and (d) logarithmic differential pore volume distribution.

use in further specific analyses. The results are usable if the raw SAS data are provided in absolute units; otherwise users must report pore characteristics in arbitrary units.

6. Conclusions

MATSAS encompasses a set of modules allowing for a full analysis of (V)SANS and (V)SAXS data from porous systems, e.g. sedimentary rocks. *MATSAS* is written in MATLAB and combines a desktop environment tuned for data processing and structural analyses with pre- and post-processing modules. The pre-processing module is used to import data from Microsoft *Excel* spreadsheets or .csv files into MATLAB. The main module performs data manipulation and analysis in which $I(Q)$ – Q curves are processed and the PDSP model is fitted to produce structural information for porous systems. The post-processing module displays results in the form of tables and figures and exports them in Microsoft *Excel* spreadsheets or .csv files. *MATSAS* is the first SAS program that provides a full suite of pore characterizations. The programs included in *MATSAS* are publicly available on GitHub (<https://github.com/matsas-software/MATSAS>) for academic users.

Acknowledgements

SANS and VSANS measurements on the rock samples tested in this study were performed on the KWS-1 and KWS-3 instruments at the Jülich Centre for Neutron Science (JCNS) at the Heinz Maier-Leibnitz Zentrum (MLZ) in Garching, Germany. We are very grateful for the beam time obtained. We also thank Masoud Ghaderi Zefreh at the University of Edinburgh for assisting with MATLAB programs, and Gernot Rother at Oak Ridge National Laboratory, Artem Feoktystov at the Forschungszentrum Jülich GmbH and Lester Barnsley at the Australian Synchrotron for testing and reporting on *MATSAS*.

References

- Anovitz, L. M. & Cole, D. R. (2015). *Rev. Mineral. Geochem.* **80**, 61–164.
- Anovitz, L. M. & Cole, D. R. (2018). *Geological Carbon Storage*, edited by S. Vialle, J. Ajo-Franklin & J. W. Carey, pp. 71–118. Washington, DC: American Geophysical Union/John Wiley & Sons, Inc.
- Anovitz, L. M., Cole, D. R., Jackson, A. J., Rother, G., Littrell, K. C., Allard, L. F., Pollington, A. D. & Wesolowski, D. J. (2015). *Geochim. Cosmochim. Acta*, **158**, 199–222.
- Bahadur, J., Melnichenko, Y. B., Mastalerz, M., Furmann, A. & Clarkson, C. R. (2014). *Energy Fuels*, **28**, 6336–6344.
- Bahadur, J., Radlinski, A. P., Melnichenko, Y. B., Mastalerz, M. & Schimmelmann, A. (2015). *Energy Fuels*, **29**, 567–576.
- Bale, H. D. & Schmidt, P. W. (1984). *Phys. Rev. Lett.* **53**, 596–599.
- Bevington, P. & Robinson, D. K. (2003). *Data Reduction and Error Analysis for the Physical Sciences*. 3rd ed. Boston: McGraw-Hill Education.
- Binder, K., Erman, B., Mark, J. E. & Roe, R. J. (2000). *Methods of X-ray and Neutron Scattering in Polymer Science*. Oxford University Press.
- Blach, T., Radlinski, A. P., Edwards, D. S., Boreham, C. J. & Gilbert, E. P. (2020). *Int. J. Coal Geol.* **225**, 103495.
- Borsali, R. & Pecora, R. (2008). *Soft Matter Characterization*. New York: Springer.
- Breßler, I., Kohlbrecher, J. & Thünemann, A. F. (2015). *J. Appl. Cryst.* **48**, 1587–1598.
- Busch, A., Kampman, N., Bertier, P., Pipich, V., Feoktystov, A., Rother, G., Harrington, J., Leu, L., Aertens, M. & Jacops, E. (2018). *Water Resour. Res.* **54**, 7076–7091.
- Busch, A., Schweinar, K., Kampman, N., Coorn, A., Pipich, V., Feoktystov, A., Leu, L., Amann-Hildenbrand, A. & Bertier, P. (2017). *Geol. Soc. London Spec. Publ.* **454**, SP454.1
- Chacón, P., Morán, F., Díaz, J. F., Pantos, E. & Andreu, J. M. (1998). *Biophys. J.* **74**, 2760–2775.
- Clarkson, C. R., Freeman, M., He, L., Agamalian, M., Melnichenko, Y. B., Mastalerz, M., Bustin, R. M., Radlinski, A. P. & Blach, T. P. (2012). *Fuel*, **95**, 371–385.
- Dewhurst, C. (2002). *GRASP*, <https://www.ill.eu/users/support-labs-infrastructure/software-scientific-tools/grasp>
- Feigin, L. A. & Svergun, D. I. (1987). *Structure Analysis by Small-Angle X-ray and Neutron Scattering*. New York: Plenum Press.
- Förster, S., Apostol, L. & Bras, W. (2010). *J. Appl. Cryst.* **43**, 639–646.
- Franke, D. & Svergun, D. I. (2009). *J. Appl. Cryst.* **42**, 342–346.
- Franke, D., Petoukhov, M. V., Konarev, P. V., Panjkovich, A., Tuukkanen, A., Mertens, H. D. T., Kikhney, A. G., Hajizadeh, N. R., Franklin, J. M., Jeffries, C. M. & Svergun, D. I. (2017). *J. Appl. Cryst.* **50**, 1212–1225.
- Fritzsche, H., Huot, J. & Fruchart, D. (2016). *Neutron Scattering and Other Nuclear Techniques for Hydrogen in Materials*. Cham: Springer International Publishing.
- Guinier, A. & Fournet, G. (1955). *Small Angle Scattering of X-rays*. New York: John Wiley.
- Hammersley, A. P. (1995). *FIT2D Version 5.18 Reference Manual*. ESRF Internal Report Exp/AH/95-01. ESRF, Grenoble, France.
- Hammersley, A. P. (2016). *J. Appl. Cryst.* **49**, 646–652.
- Heenan, R. K. (1999). *FISH, Program for Peak Analysis*. Rutherford Appleton Laboratory Internal Publication 89-129, Didcot, United Kingdom.
- Heenan, R. K., Penfold, J. & King, S. M. (1997). *J. Appl. Cryst.* **30**, 1140–1147.
- Hinde, A. L. (2004). *J. Appl. Cryst.* **37**, 1020–1024.
- Hiragi, Y., Sano, Y. & Matsumoto, T. (2003). *J. Synchrotron Rad.* **10**, 193–196.
- Hofmann, A. & Whitten, A. E. (2014). *J. Appl. Cryst.* **47**, 810–815.
- Homan, E., Konijnenburg, M., Ferrero, C., Ghosh, R. E., Dolbnya, I. P. & Bras, W. (2001). *J. Appl. Cryst.* **34**, 519–522.
- Hopkins, J. B., Gillilan, R. E. & Skou, S. (2017). *J. Appl. Cryst.* **50**, 1545–1553.
- Ilavsky, J. & Jemian, P. R. (2009). *J. Appl. Cryst.* **42**, 347–353.
- Keiderling, U. (1997). *Physica B*, **234–236**, 1111–1113.
- Konarev, P. V., Petoukhov, M. V., Volkov, V. V. & Svergun, D. I. (2006). *J. Appl. Cryst.* **39**, 277–286.
- Leu, L., Georgiadis, A., Blunt, M. J., Busch, A., Bertier, P., Schweinar, K., Liebi, M., Menzel, A. & Ott, H. (2016). *Energy Fuels*, **30**, 10282–10297.
- Liu, H., Hexemer, A. & Zwart, P. H. (2012). *J. Appl. Cryst.* **45**, 587–593.
- Manalastas-Cantos, K., Konarev, P. V., Hajizadeh, N. R., Kikhney, A. G., Petoukhov, M. V., Molodenskiy, D. S., Panjkovich, A., Mertens, H. D. T., Gruzinov, A., Borges, C., Jeffries, C. M., Svergun, D. I. & Franke, D. (2021). *J. Appl. Cryst.* **54**, 343–355.
- Mandelbrot, B. B. (1983). *The Fractal Geometry of Nature*. New York: W. H. Freeman and Company.
- Mares, T. E., Radlinski, A. P., Moore, T. A., Cookson, D., Thiagarajan, P., Ilavsky, J. & Klepp, J. (2012). *Int. J. Coal Geol.* **94**, 173–181.
- Mastalerz, M., He, L., Melnichenko, Y. B. & Rupp, J. A. (2012). *Energy Fuels*, **26**, 5109–5120.

- Melnichenko, Y. B. (2015). *Small-Angle Scattering from Confined and Interfacial Fluids: Applications to Energy Storage and Environmental Science*. Cham: Springer.
- Melnichenko, Y. B., He, L., Sakurovs, R., Kholodenko, A. L., Blach, T., Mastalerz, M., Radliński, A. P., Cheng, G. & Mildner, D. F. R. (2012). *Fuel*, **91**, 200–208.
- Meyer, K. & Klobes, P. (1999). *Fresenius J. Anal. Chem.* **363**, 174–178.
- Muthig, M., Prévost, S., Orglmeister, R. & Gradzielski, M. (2013). *J. Appl. Cryst.* **46**, 1187–1195.
- Nielsen, S. S., Toft, K. N., Snakenborg, D., Jeppesen, M. G., Jacobsen, J. K., Vestergaard, B., Kutter, J. P. & Arleth, L. (2009). *J. Appl. Cryst.* **42**, 959–964.
- Petoukhov, M. V., Franke, D., Shkumatov, A. V., Tria, G., Kikhney, A. G., Gajda, M., Gorba, C., Mertens, H. D. T., Konarev, P. V. & Svergun, D. I. (2012). *J. Appl. Cryst.* **45**, 342–350.
- Porod, G. (1951). *Kolloid-Z.*, **124**, 83–114.
- Porod, G. (1952). *Kolloid-Z.*, **125**, 51–57.
- Radlinski, A. P. (2006). *Rev. Mineral. Geochem.* **63**, 363–397.
- Radliński, A. P., Boreham, C. J., Lindner, P., Randl, O., Wignall, G. D., Hinde, A. & Hope, J. M. (2000). *Org. Geochem.* **31**, 1–14.
- Radliński, A. P., Busbridge, T. L., Gray, E. M. A., Blach, T. P. & Cookson, D. J. (2009). *Int. J. Coal Geol.* **77**, 80–89.
- Radlinski, A. P., Ioannidis, M. A., Hinde, A. L., Hainbuchner, M., Baron, M., Rauch, H. & Kline, R. (2002). *Proceedings of the 2002 International Symposium of the Society of Core Analysts*, 23–27 September 2002, Monterey, California, USA, SCA2002-35.
- Radlinski, A. P., Ioannidis, M. A., Hinde, A. L., Hainbuchner, M., Baron, M., Rauch, H. & Kline, S. R. (2004). *J. Colloid Interface Sci.* **274**, 607–612.
- Radlinski, A. P., Mastalerz, M., Hinde, A. L., Hainbuchner, M., Rauch, H., Baron, M., Lin, J. S., Fan, L. & Thiagarajan, P. (2004). *Int. J. Coal Geol.* **59**, 245–271.
- Rezaeyan, A., Pipich, V., Bertier, P., Seemann, T., Leu, L., Kampman, N., Feoktystov, A., Barnsley, L. C. & Busch, A. (2019a). *Microstructural Investigation of Mudrock Seals Using Nanometre-Scale Resolution Techniques*. Sixth EAGE Shale Workshop, 28 April–1 May 2019, Bordeaux, France.
- Rezaeyan, A., Pipich, V., Bertier, P., Seemann, T., Leu, L., Kampman, N., Feoktystov, A., Barnsley, L. C. & Busch, A. (2019b). *Quantitative Analysis of the Pore Structure of Premature-To-Postmature Organic-Rich Mudrocks Using Small-Angle Neutron Scattering*. Sixth EAGE Shale Workshop, 28 April–1 May 2019, Bordeaux, France.
- Rezaeyan, A., Seemann, T., Bertier, P., Pipich, V., Leu, L., Kampman, N., Feoktystov, A., Barnsley, L. & Busch, A. (2019). SPE/AAPG/SEG Asia Pacific Unconventional Resources Technology Conference, 18–19 November 2019, Brisbane, Australia, p. 13.
- Sakurovs, R., Koval, L., Grigore, M., Sokolova, A., Ruppert, L. F. & Melnichenko, Y. B. (2018). *Int. J. Coal Geol.* **186**, 126–134.
- Schmidt, P. W. (1982). *J. Appl. Cryst.* **15**, 567–569.
- Schrank, C. E., Gioseffi, K., Blach, T., Gaede, O., Hawley, A., Milsch, H., Regenauer-Lieb, K. & Radlinski, A. P. (2020). *J. Petrol.* **61**, ega041.
- Seemann, T., Bertier, P., Maes, N., Rezaeyan, A., Pipich, V., Barnsley, L., Busch, A. & Cnudde, V. (2019). *Resolving the Pore Structure and Sorption Properties of Methane in Mudrocks – A Small Angle Neutron Scattering Study*. Sixth EAGE Shale Workshop, 28 April–1 May 2019, Bordeaux, France.
- Vishal, V., Chandra, D., Bahadur, J., Sen, D., Hazra, B., Mahanta, B. & Mani, D. (2019). *Energy Fuels*, **33**, 4835–4848.
- Walther, D., Cohen, F. E. & Doniach, S. (2000). *J. Appl. Cryst.* **33**, 350–363.
- Zemb, T. & Lindner, P. (2002). *Neutrons, X-rays and Light: Scattering Methods Applied to Soft Condensed Matter*. North-Holland Delta Series. Amsterdam: Elsevier.
- Zhao, J. K. (2011). *J. Appl. Cryst.* **44**, 1277–1280.