MATSAS: A Small Angle Scattering Computer Tool for Porous Systems

Amirsaman Rezaeyan[1*], Vitaliy Pipich[2], and Andreas Busch[1]

[1] *Heriot-Watt University, The Lyell Centre, Research Avenue South, Edinburgh, EH14 4AS, UK;*

[2] *Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ), Forschungszentrum Jülich GmbH, Lichtenbergstrasse 1, 85748 Garching, Germany;*

Abstract: MATSAS is a script-based MATLAB® program for analysis of small angle scattering (SAS) of neutrons and X-rays data obtained from various facilities. The program has primarily been developed for sedimentary rock samples but is equally applicable for other porous media. MATSAS imports raw SAS data from Excel worksheets, combines small angle scattering and very small angle scattering data, subtracts the background signals, and displays the processed scattering curves in interactive log–log plots. MATSAS uses the polydisperse spherical (PDSP) model to obtain structural information of scatterers (scattering objects); for a porous system, 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 (countless); results are then exported to a Microsoft Excel file with separate sheets for individual samples. We included extra I/O programs for CSV import/export, however they permit one sample analysis at a time only. MATSAS is the first SAS program that delivers a full suite of pore characterisations for sedimentary rocks. MATSAS is an open-source package, which is freely available at GitHub (https://github.com/matsas-software/MATSAS).

* correspondence email: ar104@hw.ac.uk ; amirsaman.rezaeyan@gmail.com

1 Introduction

Small-angle scattering (SAS) of neutrons and X-rays (SANS and SAXS) is widely used to non-destructively study the low-resolution structure of natural and engineered systems, including sedimentary rocks, biological macromolecules, composite nanomaterials, or polymers on the length scales between Angstroms and microns in a single or combined experiment (Feigin and Svergun 1987, Binder et al. 2000, Zemb and Lindner 2002, Borsali and Pecora 2008, Melnichenko 2015, Fritzsche et al. 2016). Advances in SAS instrumentation such as neutron radiation and high-flux X-ray synchrotron beamlines have significantly increased the use of SANS and SAXS experiments (Melnichenko 2015, Zemb and Lindner 2002, Heenan et al. 1997). Despite the availability of these technologies, modern instruments 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). Obtained over past decades, theoretical and methodological developments 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). Neutron and X-ray scattering techniques are complementary rather than competitive (Melnichenko 2015). However, differences in the charge and energy, and hence the way they interact with matter, define certain advantages and disadvantages of each type of radiation for conducting a specific type of experiment or dealing with a specific type of sample (Binder et al. 2000, Zemb and Lindner 2002, Melnichenko 2015). Figure 1-1 illustrates a typical SAS experiment. A flux of monochromatic neutrons or X-rays- that travels in the straight trajectory of their wave vector $\boldsymbol{k_0}$- is elastically scattered inside a sample of uniform thickness t and irradiated volume V. The magnitude of $\boldsymbol{k_0}$ is λ^{-1}, where λ is the neutron or X-ray wavelength. The intensity dI scattered in direction $\boldsymbol{k}$ is measured; thereby the convention $|\mathbf{k} - \mathbf{k_0}| = \mathbf{s}$ and the so-called scattering vector $\mathbf{Q} = 2\pi\mathbf{s}$. It follows from Figure 1-1 that s is related to the wavelength λ and the scattering angle 2θ by $s = 2\sin\theta/\lambda$. Q is thus expressed by $Q = 4\pi sin\theta/\lambda$ (Radlinski 2006).

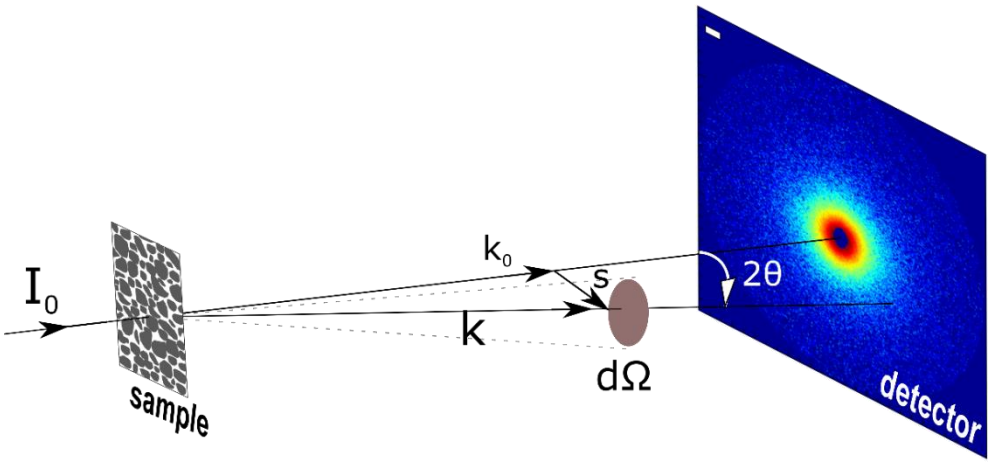

Figure 1-1. The schematic principle of a SAS experiment.

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/X-rays per second) and A is the beam cross sectional area at the sample position (Figure 1-1) (Radlinski 2006). The scattered intensity monitored in the solid angle element $d\Omega$ targeted by the scattering vector 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 scattering length density (neutrons) or the electron density (X-rays) in the sample from the measured $d\Sigma/d\Omega$ as a function of scattering vector Q; $\frac{d\Sigma}{d\Omega}(Q)$ or $I(Q)$ (Melnichenko 2015).

For a wide range of substances, the 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 comprised of supra-molecular-size phases, each characterised 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), respectively (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 scatterers, B is the background signal, and ρ_1^* and ρ_2^* are the scattering length/electron density of phase 1 and phase 2, respectively. $F(Q)$ is the so-called form factor that describes the size and shape of the scatterer. $S(Q)$ is called the structure factor that contains information about spatial distribution of the scatterers (Melnichenko 2015). There are analytical expressions for the form factor for simple geometrical objects like spheres, cylinders, discs, parallelepipeds. The structure factor is defined by its position component in solid porous materials where the positions of scatterers are frozen in time and space. In soft matter systems, the structure factor is also dependent on the interaction potential between scatterers (Melnichenko 2015). Form and structure factors need to be specified to determine the structural information of 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 characterise the structure of scatterers. The program FIT2D is one of the first packages, which is designed for two-dimensional image data reduction/manipulation and peak fitting (Hammersley 1995). Modelling and fitting of one-dimensional curves can be done using the program FISH (Heenan 1999) for peak analysis and parametric fitting using various form and structure factors, or using the ab initio shape determination programs DALAI_GA (Chacón et al. 1998) or SAX3D (Walther et al. 2000). The packages BerSANS (Keiderling 1997), SAXS/WAXS software system (Homan et al. 2001), and GRASP (Dewhurst 2002) are also processing toolboxes for SAS data acquisition/reduction. Hiragi et al. (2003) developed an interactive processing program, which is called SAXSANA. The program PRINSAS (Hinde 2004) applies spherical form factors for a polydisperse scattering system in order to analyse porous materials such as sedimentary rocks. ATSAS (Konarev et al. 2006, Petoukhov et al. 2012) covers the major processing and interpretation steps from primary SAS data reduction to 3D modelling. DAMMIF (Franke and Svergun 2009) is a program for rapid ab-initio shape determination, which recovers the 3D structure from the 1D SAS pattern. The IRENA package (Ilavsky and Jemian 2009) is data manipulations and analysis toolbox for SAS data that include plotting SAS data, merging of two overlapping data sets, and fitting form and structural models to data from contrast variation

experiments. BioXTAS RAW program (Nielsen et al. 2009) reduces the isotropic SAXS data automatically or manually, conducts the primary analysis, and calculates the pair-distance distribution functions. SAAF (Zhao 2011) program analyses SANS data using analytical model functions with a set of standard models. SASTBX (Liu et al. 2012) contains a range of functions to analyse biological SAS data, from data reduction to model reconstruction, refinement and Shape retrieval. SASview (http://www.sasview.org/) is designed for SAS data reduction, manipulation, and analysis, which includes several form and structure factors with polydispersity and orientational distributions such as rectangular, Gaussian, Schulz, log normal, and array. SASview contains instrument smearing options, fit weighting options, complex constrained fitting, as well as a plug-in model editor. Other widely used software packages for the analysis of SAS data include the SASFIT suite (Kohlbrecher and Bressler 2009), SCATTER (Forster et al. 2010), SASET (Muthig et al. 2013), MolScat and SAFIR (Hofmann and Whitten 2014), and BioXTAS RAW (Hopkins et al. 2017).

Recognising 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 (Clarkson et al. 2012, Mastalerz et al. 2012, Melnichenko et al. 2012, Bahadur et al. 2014, Anovitz et al. 2015, Bahadur et al. 2015, Leu et al. 2016, Busch et al. 2017b, Anovitz and Cole 2018, Busch et al. 2018, Sakurovs et al. 2018, Vishal et al. 2019, Blach et al. 2020), we developed the program package MATSAS. It allows analysing data obtained from small angle and very small angle scattering of neutrons and X-rays (VSANS, SANS, VSAXS, and SAXS). MATSAS analyses data from pinhole geometry, time of flight (TOF) and Bonse–Hart machines and was tested using data acquired from FRM-II (Research Reactor Munich II, Garching, Germany) and ORNL (Oak Ridge National Laboratory, Tennessee, USA) (Rezaeyan et al. 2019a, 2019b, 2019c, Seemann et al. 2019). MATSAS does post-processing of data obtained from research facilities, assuming that initial corrections for sample thickness, transmission, detector sensitivity, instrument backgrounds, and noise have been done 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 the functionality, making it applicable for 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, which is freely available on GitHub (https://github.com/matsas-software/MATSAS). Besides, 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 help accelerating its development. In this paper, we summarise the main components of MATSAS and its development framework.

2 Program overview

MATSAS features a script-based package in MATLAB® (The MathWorks Inc., Cambridge, UK), which integrates computation and visualisation 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 to 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 commonly used PDSP method in SAS analysis, such as theoretical intensity computation, $f(r)$ probability function of pore size distribution, and model refinement. The script-based MATSAS allows tuning parameters for more features of each routine. Nevertheless, use of the MATSAS program is divided into three steps: (1) the pre-processing of raw or facility post-corrected SAS and very-small-angle scattering (VSAS) data as well as physical information, (2) the processing of the imported information to produce $I(Q)$ versus Q curves, combine the SAS and VSAS curves, and fit the PDSP model, and (3) the post-processing to display and export structural information obtained from the samples being analysed. Figure 2-1 illustrates main components of the present version of MATSAS.

The detailed instructions to use the package is available on GitHub. Supporting information and command descriptions are embedded in each module. Errors and bugs can be invoked when no parameter or incorrect data is given to the command. We developed the package in Windows and recommend running it in Windows, Mac, or Linux, with any Intel or AMD x 86-64 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 1GB GPU memory is recommended as displaying figures and generating Microsoft Excel worksheets require more background processing.

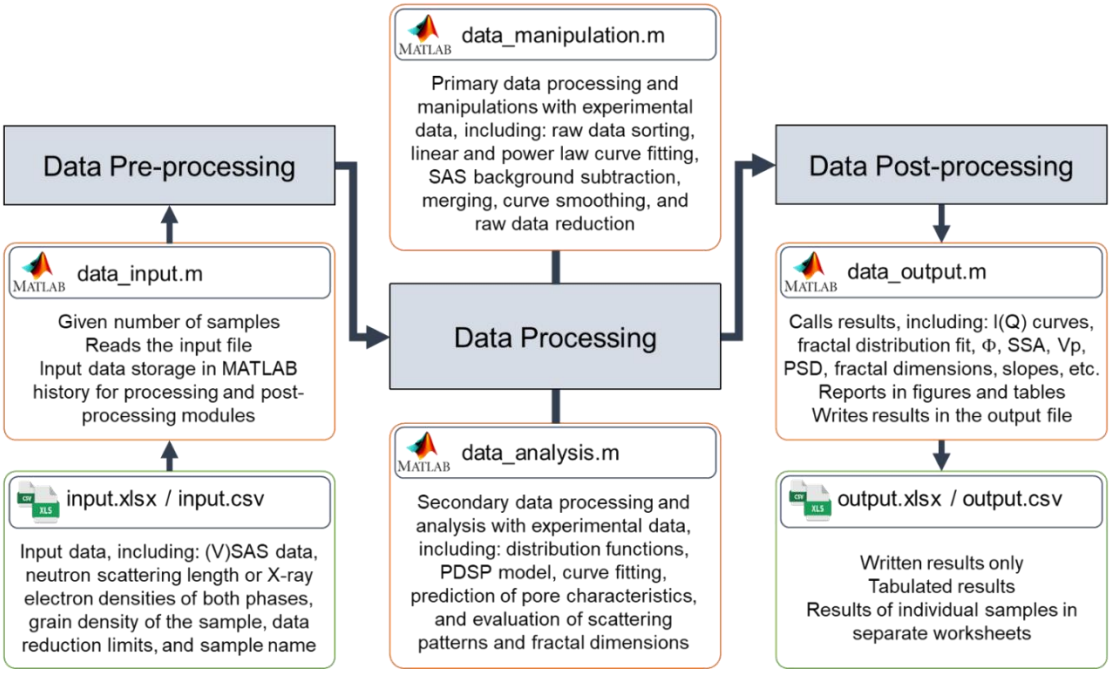

Figure 2-1. A schematic flow chart of MATSAS programs and their functionalities.

3 Data Pre-processing

The data pre-processing module is composed of two compartments: (A) data are prepared in Microsoft Excel worksheets- input.xlsx 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, and sample name, and (B) MATLAB data_input.m file reads and stores the imported data for the next step. Users are required to inform the program on the number of samples to be analysed. 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 $Å^{-1}$ 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. In addition, the csv_input.m file is included to import data in CSV format; in this case one sample can be analysed at a time only.

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 file carries on multiple tasks, including: *I(Q)* data sorting, curve fittings, background subtraction, curves merging, curve smoothing, and raw data reduction. The secondary data processing script file is designed in data_analysis.m to analyse *I(Q)*-*Q* curves and produce structural information. The power law distribution function is employed and PDSP model is fitted to the processed scattering curves. Pore characteristics are predicted and fractal dimensions (including pore fractal dimension, D_p, surface fractal dimension, D_s, and general fractal dimension, D_f) are evaluated from scattering patterns 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 scattering curves merging 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 are possibly collected from more than one detector. Once data of several detectors are combined for one specific instrument, they may not be sorted, which leads to numerical problems in further analysis. Data sorting is the first step and carried out in the data manipulation package using a built-in function. Provided that SAS data consists of two scattering profiles obtained from VSAS and SAS instruments, MATSAS allows users to merge the two curves using a least-squares fitting in the overlapping range as illustrated for sample in Figure 4-1. The SAS curve is the basis onto which the VSAS curve is rebinned. The incoherent background signal is subtracted using Equation (3):

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

in case the scattering varies with Q^{-a} in the high *Q*-limit before plateau (Melnichenko 2015). The value of the background $\left[\frac{d\Sigma}{d\Omega}(Q)\right]_{inc}$ is determined from a linear plot of Equation (4):

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

where $\left[\frac{d\Sigma}{d\Omega}(Q)\right]_{inc}$ is the slope and A is the intercept (Melnichenko 2015). Figure 4-1 shows the background subtraction in the high Q-limit for a range that users can manually change in the program. A noise removal operation is embedded to remove the sparse data around the beam stop or detector edge. The raw data reduction whose cut-off limits are determined in the input.xlsx file is carried out as well. Two data smoothing operations are included in the package that can be employed to make 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 and Robinson 2003). A MATLAB plotting operation displays currently active scattering profiles in logarithmic-logarithmic scale. An advanced plotting option included in the plot permits users to change the plotting range, zoom factor, etc. The data manipulation file contains an output section, where the result of each operation can be further 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 that allows modifying or changing lines if needed.

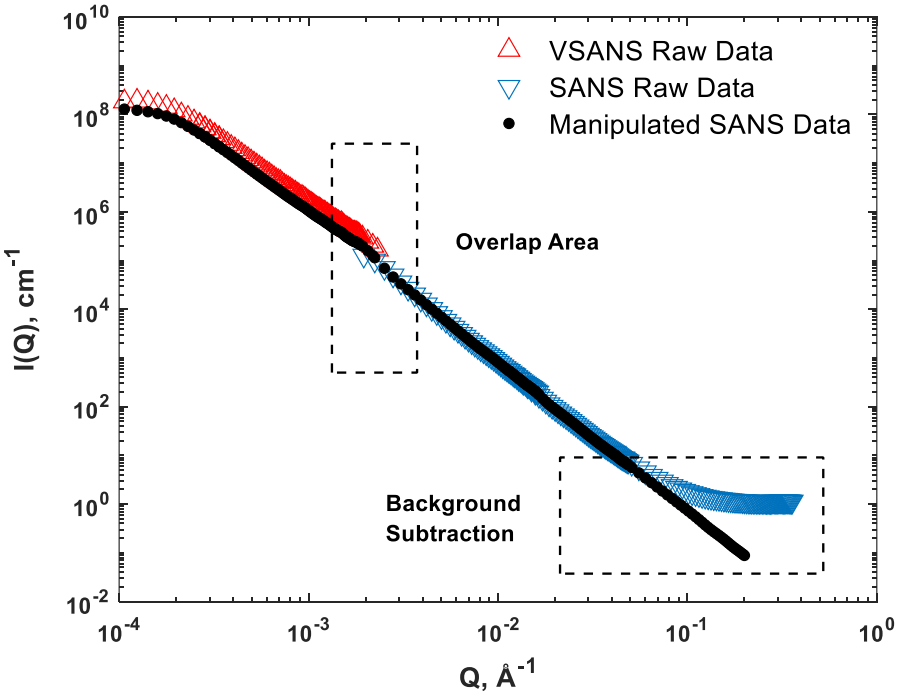

Figure 4-1. SANS data manipulated and processed on an arbitrary mudrock sample. Red, blue, and black curves are the scattering profiles from VSANS instrument, from the SANS instrument, and the net scattering after manipulation (merging, background subtraction, and smoothing), respectively.

4.2 Data Analysis

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

$$f(r)dR \sim R^{-(1+D_f)}dR \tag{5}$$

where D_f is the fractal dimension determined from the slope of the power-law scattering (Melnichenko 2015). Practically, the distribution described by Equation (5) ranging from $R_{min} \leq R \leq R_{max}$ shows fractal behaviour confined between the upper and lower cut-off parameters. Scattering from a PDSP featured sample has a linear region with a similar slope $-(1 + D_f)$ and is described by (Radlinski et al. 2002):

$$I(Q) = \int_{R_{min}}^{R_{max}} (\rho_1^* - \rho_2^*)^2 \, N \, f(r) \, V^2 \, P(Q,r) \, dr \tag{6}$$

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

$$P(Q,r) = \left[3\frac{\sin(Q,r) - Qr \, \cos(Q,r)}{Q^3 r^3}\right]^2 \tag{7}$$

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

$$N(r) = \frac{\Phi}{\bar{V}(r)} f(r) = \frac{IQ_0}{(\rho_1^* - \rho_2^*)^2} \frac{1}{\bar{V}^2(r)} \tag{8}$$

where $IQ_0 = f(r)(R_{max} - R_{min})\Phi \frac{(\rho_1^* - \rho_2^*)^2}{\bar{V}(r)}$ 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 scatterers (Radlinski et al. 2002). Similar to Ilavsky and Jemian (2009), MATSAS calculates Equation (6) throughout the integration over a continuous size distribution with a summation over a discrete size histogram:

$$I(Q) = \sum_i (\rho_{1_i}^* - \rho_{2_i}^*)^2 \sum_{i,j} N_i\, f_i(r_{i,j})\, V_i(r_{i,j})^2\, P_i(Q_i, r_{i,j})\, \Delta r_{i,j} \tag{9}$$

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; each scattering size has its own binning index *i,j*. *r* is the dimension of the scatterer (radius for spheres) that has limits r_{max_i} and r_{min_i}.

MATSAS uses a rectangular size distribution that is used to model the scattering volume distribution $V^2(r)\, P(Q, r)$. Users can change the theoretical ranges of the various size distributions in the data analysis program. Numerical calculations call limits on the range of dimension (r_{min} and r_{max}), cut-off limits (R_{min} and R_{max}), and number of bins (N_{bin}). This method results in a natural logarithmic stepping in dimension and uses three user parameters – R_{min}, R_{max}, and N_{bin}. The centre of the first ($r_{i,1}$) and the last ($r_{i,N_{bin}}$) bins are R_{max} and R_{min}, respectively and extra fractional volumes are discarded for both bins: the volume associated with $r_{min_{i,1}} - r_{i,1}$ and $r_{i,N_{bin}} - r_{max_{i,N_{bin}}}$ for the first and last bins, respectively. The width of bins are equal by selecting associated dimensions at regular increments of the cumulative distribution (Ilavsky and Jemian 2009), leading to $\log(\Delta) = \frac{\log(R_{max}) - \log(R_{min})}{N_{bin}}$. 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_i} to fit the PDSP model in Equation (6) to the measured *I(Q)* curve. The fitting procedure employs $f(r)$ and IQ_0 as fitting parameters for each iteration to attain the match where the summation of square errors (SSQ) tends to a minimum (Hinde 2004). Furthermore, to decrease the computation time due to numerical integration, we found an analytical solution for the scattering volume distribution:

$$\int_{R_{min}}^{R_{max}} V^2\, P(Q,r)\, dr = \sum_{i,j} \frac{1}{{Q_i}^7}\Big(16\pi^2\Big(-\frac{5}{4}\sin\big(Q_i r_{i,j}\big)\cos\big(Q_i r_{i,j}\big) + \frac{3}{2} Q_i r_{i,j} \cos\big(Q_i r_{i,j}\big)^2 + {Q_i}^2 {r_{i,j}}^2\Big(\frac{1}{2}\sin\big(Q_i r_{i,j}\big)\cos\big(Q_i r_{i,j}\big) + \frac{1}{2} Q_i r_{i,j}\Big) - \frac{1}{4} Q_i r_{i,j} - \frac{1}{3} {Q_i}^3 {r_{i,j}}^3\Big)\Big)\Delta r_{i,j} \quad (10)$$

that transforms Equation (9) into:

$$I(Q) = \sum_i (\rho_{1_i}^* - \rho_{2_i}^*)^2 \sum_{i,j} N_i\, f_i\big(r_{i,j}\big)\, \frac{1}{{Q_i}^7}\Big(16\pi^2\Big(-\frac{5}{4}\sin\big(Q_i r_{i,j}\big)\cos\big(Q_i r_{i,j}\big) + \frac{3}{2} Q_i r_{i,j} \cos\big(Q_i r_{i,j}\big)^2 + {Q_i}^2 {r_{i,j}}^2\Big(\frac{1}{2}\sin\big(Q_i r_{i,j}\big)\cos\big(Q_i r_{i,j}\big) + \frac{1}{2} Q_i r_{i,j}\Big) - \frac{1}{4} Q_i r_{i,j} - \frac{1}{3} {Q_i}^3 {r_{i,j}}^3\Big)\Big)\Delta r_{i,j} \quad (11)$$

MATSAS simplifies the intensity calculation by substituting Equation (8) into Equation (11), leading to:

$$I(Q) = \sum_{i,j}\, IQ_{0_i}\, f_i\big(r_{i,j}\big)\, \bar{V}_i^{-2}\big(r_{i,j}\big)\, \frac{1}{{Q_i}^7}\Big(16\pi^2\Big(-\frac{5}{4}\sin\big(Q_i r_{i,j}\big)\cos\big(Q_i r_{i,j}\big) + \frac{3}{2} Q_i r_{i,j} \cos\big(Q_i r_{i,j}\big)^2 + {Q_i}^2 {r_{i,j}}^2\Big(\frac{1}{2}\sin\big(Q_i r_{i,j}\big)\cos\big(Q_i r_{i,j}\big) + \frac{1}{2} Q_i r_{i,j}\Big) - \frac{1}{4} Q_i r_{i,j} - \frac{1}{3} {Q_i}^3 {r_{i,j}}^3\Big)\Big)\Delta r_{i,j} \quad (12)$$

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

$$SSA = \frac{1}{\rho_g(\rho_1^* - \rho_2^*)^2} \sum_k \frac{4}{3}\pi\, IQ_{0_k}\, f(r_k)\, \bar{V}^{-2}(r_k)\, \Delta r_k \quad (13)$$

where the subscript *k* represents bins in the size distribution. The volume fraction of scatterers per unit volume (Φ) is calculated from Equation (8), which results in:

$$\Phi = \frac{1}{(\rho_1^* - \rho_2^*)^2} \sum_k IQ_{0_k}\, \bar{V}(r_k)\, \bar{V}^{-2}(r_k) \quad (14)$$

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

$$V_p = \frac{1}{\rho_g} \sum_k \frac{\Phi(r_k)}{1 - \Phi(r_k)} \quad (15)$$

where the subscript *k* represents bins in the size distribution. Differential (dV/dr or dA/dr) and logarithmic differential scatterer size distributions ($dV/dlogr$ or $dA/dlogr$) are calculated cumulatively (Meyer and Klobes 1999).

For demonstration purposes, we tested the analysis operations on SANS and VSANS data obtained from 9 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. 2017a). Figure 4-2-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 an initial guess for $f(r)$ and IQ_0, which is obtained from the slope of the scattering curves and the Guinier and Fournet (1955) approximation, respectively. SSQ tends to a minimum after the second iteration; two iterations are recommended for most rock samples. Figure 4-2-B shows $f(r)$ after two iterations on a log-log scale. $f(r)$ levels off at scatterer sizes > ~2 µm because the scattering intensity of large scatterers are smeared, possibly due to instrument artefacts at the edge of the detector. The volume-weighed scatterer size distribution $D^3f(r)$ shows a monomodal scatterer size distribution (Figure 4-2-C). The error sensitivity, expressed as $d\text{SSQ}/d\log(IQ_0)$, relates SSQ to the number of iterations (Figure 4-2-D). The $d\text{SSQ}/d\log(IQ_0)$ varies around zero for all scatterer sizes. However, as illustrated in Figure 4-2-D 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. Furthermore, SSQ magnifies when the number of iterations exceeds 2, 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.

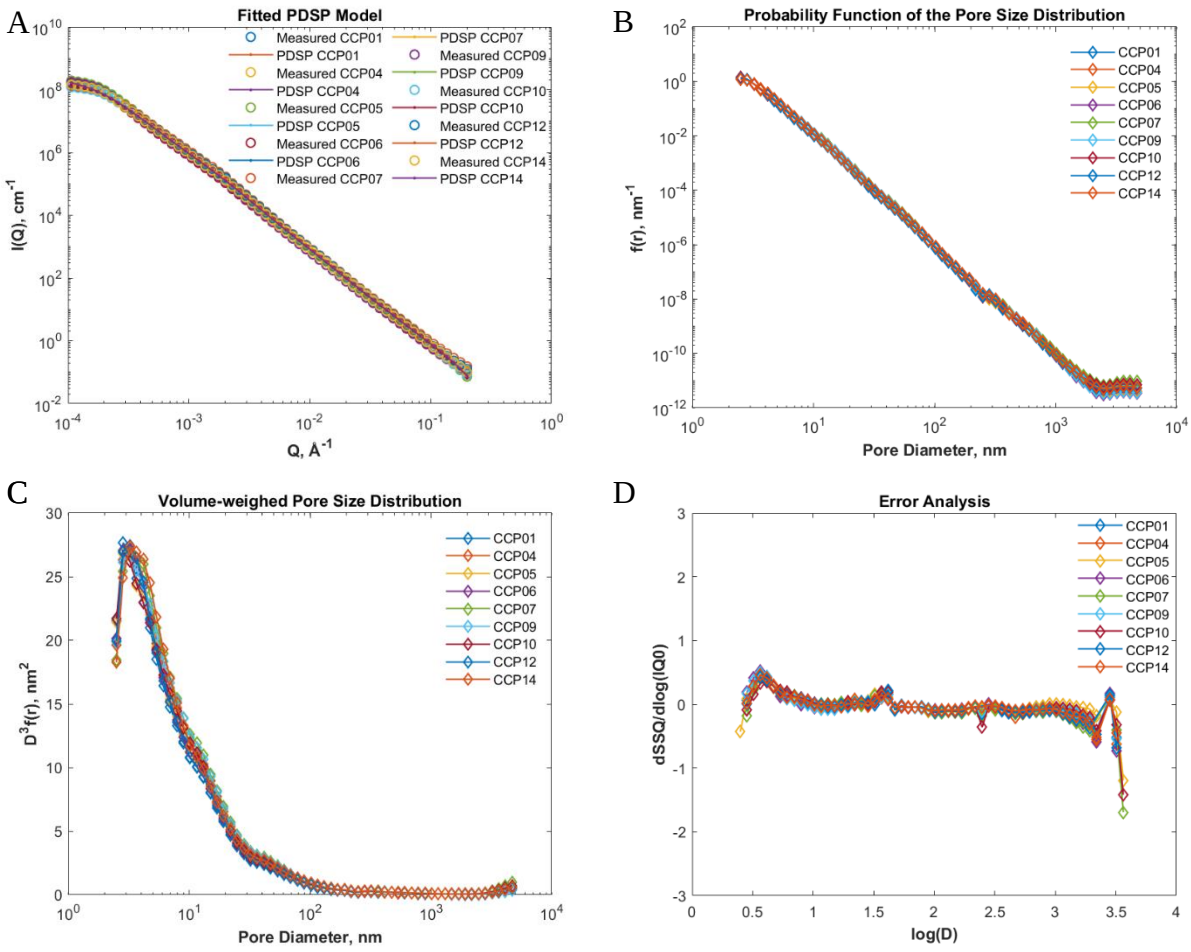

Figure 4-2. PDSP model applied to SANS data obtained from 9 rock samples (Opalinus Clay). (A) Measured *I(Q)* curves after manipulation and *I(Q)* curves obtained from PDSP model; (B) probability functions of the pore size distribution *f(r)*; (C) volume-weighed pore size distribution $D^3f(r)$; (D) error sensitivity $d\text{SSQ}/d\log(IQ_0)$ obtained after two iterations.

5 Data Post-processing

The data post-processing module is made of two compartments, including data_output.m in MATLAB® and output.xlsx in Microsoft Excel. The data_output.m file calls the results of individual samples, reports 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, the slope of scattering curves, pore characteristics divided in macro-, meso-, and micropores, and background subtraction

values. Some of results are shown in Figure 5-1. PDSP model applied to SANS data obtained from 9 rock samples (Opalinus Clay). (A) Cumulative pore area distribution; (B) logarithmic differential pore area distribution; (C) cumulative pore volume distribution; (D) logarithmic differential pore volume distribution.Figure 5-1 and Table 5-1Table 5-1. Slope, fractal dimensions, incoherent background, and pore characteristics evaluated by MATSAS from the SANS data of 9 rock samples.. The results of individual samples are produced in Excel worksheets, separately. No further analysis is needed in Excel. In addition, the csv_output.m file is included to export results in CSV format.

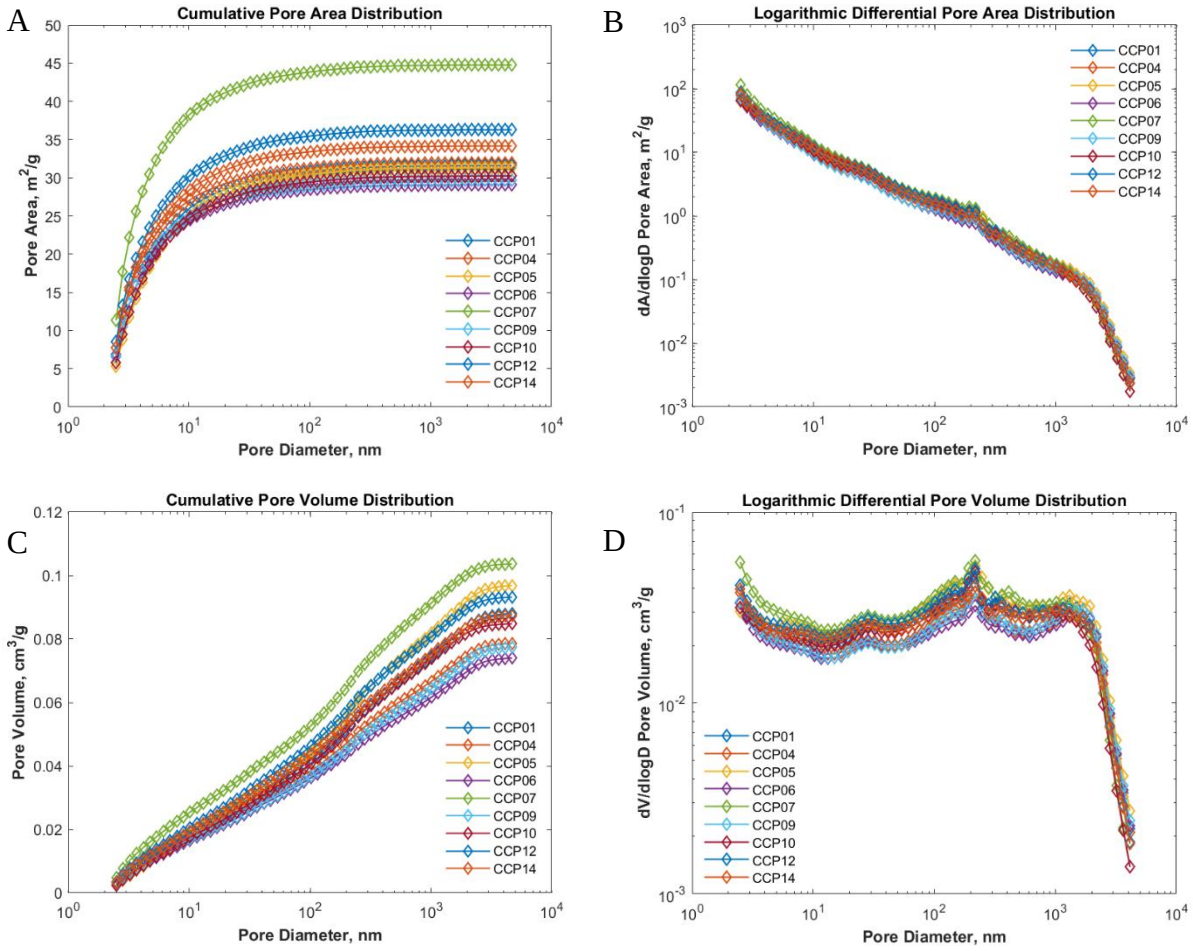

Figure 5-1. PDSP model applied to SANS data obtained from 9 rock samples (Opalinus Clay). (A) Cumulative pore area distribution; (B) logarithmic differential pore area distribution; (C) cumulative pore volume distribution; (D) logarithmic differential pore volume distribution.

Table 5-1. Slope, fractal dimensions, incoherent background, and pore characteristics evaluated by MATSAS from the SANS data of 9 rock samples.

Sample ID	**m**	**D_f**	**D_s**	**D_p**	**$I_{BG\ VSAS}$**	**$I_{BG\ SAS}$**	**SSA**	**SSA_{macro}**	**SSA_{meso}**
	-	-	-	-	cm^{-1}	cm^{-1}	m^2/g	m^2/g	m^2/g
CCP01	-3.06	2.94	2.88	2.84	15317	1.15	31.6	1.4	30.2
CCP04	-3.06	2.94	2.89	2.83	10888	0.91	31.9	1.2	30.7
CCP05	-3.07	2.93	2.87	2.79	15286	1.21	31.3	1.6	29.7
CCP06	-3.05	2.95	2.90	2.87	11135	0.92	29.1	1.1	28.0
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
CCP10	-3.06	2.94	2.86	2.78	23855	1.06	30.2	1.4	28.8
CCP12	-3.04	2.96	2.88	2.78	13734	1.08	36.3	1.5	34.8
CCP14	-3.04	2.96	2.91	2.83	13359	0.56	34.2	1.4	32.8

Table 5-1. (continued).

V_p	**V_{macro}**	**V_{meso}**	**Φ**	**Φ_{macro}**	**Φ_{meso}**	**SSQ**
cm^3/g	cm^3/g	cm^3/g	%	%	%	-
0.0880	0.0541	0.0339	23.7	14.6	9.1	0.017
0.0786	0.0469	0.0318	21.4	12.7	8.6	0.038
0.0968	0.0613	0.0356	26.2	16.6	9.6	0.026
0.0740	0.0443	0.0297	20.1	12.0	8.0	0.004
0.1036	0.0601	0.0435	28.0	16.3	11.8	0.149
0.0773	0.0472	0.0301	21.1	12.9	8.2	0.022
0.0848	0.0519	0.0328	23.0	14.1	8.9	0.032
0.0932	0.0555	0.0377	25.4	15.1	10.3	0.019
0.0874	0.0517	0.0357	23.8	14.1	9.7	0.036

6 Conclusions

MATSAS encompasses a set of modules allowing for a full analysis of (V)SANS and (V)SAXS data from sedimentary rocks and other isotropic systems. MATSAS is written in MATLAB® that 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 worksheets or a CSV file 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 the structural information of porous systems. The post-processing module displays results in forms of tables and figures and exports them in Microsoft Excel worksheets or a CSV file. MATSAS is the first SAS program that provides a full suite of pore characterisations. 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 here were performed at KWS-1 and KWS-3 instruments of the Jülich Centre for Neutron Science (JCNS) at Heinz Maier-Leibnitz Zentrum (MLZ) in Garching, Germany. We are very grateful for the beam time obtained. We also thank Dr Masoud Ghaderi Zefreh of the University of Edinburgh for assisting with MATLAB programs as well as Dr Gernot Rother of the Oak Ridge National Laboratory, Dr Artem Feoktystov of the Forschungszentrum Jülich GmbH, and Dr Lester Barnsley of the Australian Synchrotron for testing and reporting on MATSAS.

References

Anovitz, Lawrence M. and Cole, David R. 2018. Analysis of the Pore Structures of Shale Using Neutron and X-Ray Small Angle Scattering. In *Geological Carbon Storage*, ed. Stéphanie Vialle, Jonathan Ajo-Franklin and J. William Carey, Chap. 4, 71-118. American Geophysical Union and John Wiley & Sons, Inc.

Anovitz, Lawrence M., Cole, David R., Jackson, Andrew J. et al. 2015. Effect of quartz overgrowth precipitation on the multiscale porosity of sandstone: A (U)SANS and imaging analysis. *Geochimica et Cosmochimica Acta* **158**: 199-222.

Bahadur, J., Melnichenko, Y. B., Mastalerz, Maria et al. 2014. Hierarchical Pore Morphology of Cretaceous Shale: A Small-Angle Neutron Scattering and Ultrasmall-Angle Neutron Scattering Study. *Energy & Fuels* **28** (10): 6336–6344.

Bahadur, Jitendra, Radlinski, Andrzej P., Melnichenko, Yuri B. et al. 2015. Small-Angle and Ultrasmall-Angle Neutron Scattering (SANS/USANS) Study of New Albany Shale: A Treatise on Microporosity. *Energy & Fuels* **29** (2): 567-576.

Bevington, P. and Robinson, D.K. 2003. *Data Reduction and Error Analysis for the Physical Sciences*: McGraw-Hill Education.

Binder, K., Erman, B., Mark, J. E. et al. 2000. *Methods of X-Ray and Neutron Scattering in Polymer Science*. Oxford: Oxford University Press.

Blach, Tomas, Radlinski, Andrzej P., Edwards, Dianne S. et al. 2020. Pore anisotropy in unconventional hydrocarbon source rocks: A small-angle neutron scattering (SANS) study on the Arthur Creek Formation, Georgina Basin, Australia. *International Journal of Coal Geology*: 103495.

Borsali, Redouane and Pecora, Robert. 2008. *Soft Matter Characterization*. New York, USA: Springer.

Busch, A., Kampman, N., Bertier, P. et al. 2018. Predicting effective diffusion coefficients in mudrocks using a fractal model and small angle neutron scattering measurements. *Water Resources Research* **0** (ja).

Busch, Andreas, Schweinar, Kevin, Kampman, Niko et al. 2017a. Determining the porosity of mudrocks using methodological pluralism. *Geological Society, London, Special Publications* **454**: 15-38.

Busch, Andreas, Schweinar, Kevin, Kampman, Niko et al. 2017b. Determining the porosity of mudrocks using methodological pluralism. *Geological Society, London, Special Publications* **454**.

Chacón, P., Morán, F., Díaz, J. F. et al. 1998. Low-resolution structures of proteins in solution retrieved from X-ray scattering with a genetic algorithm (in eng). *Biophysical journal* **74** (6): 2760-2775.

Clarkson, C. R., Freeman, M., He, L. et al. 2012. Characterization of tight gas reservoir pore structure using USANS/SANS and gas adsorption analysis. *Fuel* **95** (0): 371-385.

Dewhurst, C. . 2002. *GRASP software package*. Institute Laue–Langevin, Grenoble, France (Reprint).

Feigin, L. A. and Svergun, D. I. 1987. *Structure Analysis by Small-Angle X-ray and Neutron Scattering*. New York: Plenum Press.

Forster, S., Apostol, L., and Bras, W. 2010. Scatter: software for the analysis of nano- and mesoscale small-angle scattering. *Journal of Applied Crystallography* **43** (3): 639-646.

Franke, Daniel and Svergun, Dmitri I. 2009. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *Journal of Applied Crystallography* **42** (2): 342-346.

Fritzsche, Helmut, Huot, Jacques, and Fruchart, Daniel. 2016. *Neutron Scattering and Other Nuclear Techniques for Hydrogen in Materials*. Switzerland: Springer International Publishing

Guinier, A. and Fournet, G. 1955. *Small Angle Scattering of X-rays*. New York: John Wiley.

Hammersley, A. P. . 1995. ESRF Internal Report Exp/AH/95-01, FIT2D V5.18 Ref. Manual, Grenoble, France.

Heenan, R. K. . 1999. FISH, program for peak analysis, Didcot, UK.

Heenan, R. K., Penfold, J., and King, S. M. 1997. SANS at Pulsed Neutron Sources: Present and Future Prospects. *Journal of Applied Crystallography* **30** (6): 1140-1147.

Hinde, Alan. 2004. PRINSAS - a Windows-based computer program for the processing and interpretation of small-angle scattering data tailored to the analysis of sedimentary rocks. *Journal of Applied Crystallography* **37** (6): 1020-1024.

Hiragi, Yuzuru, Sano, Yoh, and Matsumoto, Tomoharu. 2003. SAXSANA: an interactive program for the analysis and monitoring of static and time-resolved small-angle X-ray solution scattering measurements. *Journal of Synchrotron Radiation* **10** (2): 193-196.

Hofmann, Andreas and Whitten, Andrew E. 2014. Two practical Java software tools for small-angle X-ray scattering analysis of biomolecules. *Journal of Applied Crystallography* **47** (2): 810-815.

Homan, E., Konijnenburg, M., Ferrero, C. et al. 2001. The SAXS/WAXS software system of the DUBBLE CRG beamline at the ESRF. *Journal of Applied Crystallography* **34** (4): 519-522.

Hopkins, Jesse Bennett, Gillilan, Richard E., and Skou, Soren. 2017. BioXTAS RAW: improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. *Journal of Applied Crystallography* **50** (5): 1545-1553.

Ilavsky, Jan and Jemian, Peter R. 2009. Irena: tool suite for modeling and analysis of small-angle scattering. *Journal of Applied Crystallography* **42** (2): 347-353.

Keiderling, U. 1997. A new software package for SANS data processing at the Hahn-Meitner-Institut in Berlin, Germany. *Physica B: Condensed Matter* **234-236**: 1111-1113.

Kohlbrecher, J. and Bressler, I. 2009. *SASfit*. ETHZ PSI, Zurich, Switzerland (Reprint).

Konarev, Petr V., Petoukhov, Maxim V., Volkov, Vladimir V. et al. 2006. ATSAS 2.1, a program package for small-angle scattering data analysis. *Journal of Applied Crystallography* **39**: 277–286.

Leu, L., Georgiadis, A., Blunt, M. J. et al. 2016. Multiscale Description of Shale Pore Systems by Scanning SAXS and WAXS Microscopy. *Energy & Fuels* **30** (12): 10282-10297.

Liu, Haiguang, Hexemer, Alexander, and Zwart, Peter H. 2012. The Small Angle Scattering ToolBox (SASTBX): an open-source software for biomolecular small-angle scattering. *Journal of Applied Crystallography* **45** (3): 587-593.

Mastalerz, M., He, L., Melnichenko, Y. B. et al. 2012. *Conference Paper*. Porosity of coal and shale: Insights from gas adsorption and SANS/USANS techniques. *Energy and Fuels* **26** (8): 5109-5120.

Melnichenko, Y. B. 2015. *Small-Angle Scattering from Confined and Interfacial Fluids: Applications to Energy Storage and Environmental Science*, 329. TN, USA: Springer.

Melnichenko, Yuri B., He, Lilin, Sakurovs, Richard et al. 2012. Accessibility of pores in coal to methane and carbon dioxide. *Fuel* **91**: 200–208.

Meyer, K. and Klobes, P. 1999. *journal article*. Comparison between different presentations of pore size distribution in porous materials. *Fresenius' Journal of Analytical Chemistry* **363** (2): 174-178.

Muthig, Michael, Prevost, Sylvain, Orglmeister, Reinhold et al. 2013. SASET: a program for series analysis of small-angle scattering data. *Journal of Applied Crystallography* **46** (4): 1187-1195.

Nielsen, S. S., Toft, K. Noergaard, Snakenborg, D. et al. 2009. BioXTAS RAW, a software program for high-throughput automated small-angle X-ray scattering data reduction and preliminary analysis. *Journal of Applied Crystallography* **42** (5): 959-964.

Petoukhov, Maxim V., Franke, Daniel, Shkumatov, Alexander V. et al. 2012. New developments in the ATSAS program package for small-angle scattering data analysis. *Journal of Applied Crystallography* **45** (2): 342-350.

Porod, G. 1951. Die Röntgenkleinwinkelstreuung von dichtgepackten kolloiden Systemen - I. Teil. *Kolloid-Zeitschrift* **124** (2): 83-114.

Porod, G. 1952. Die Röntgenkleinwinkelstreuung von dichtgepackten kolloiden Systemen - II. Teil. *Kolloid-Zeitschrift* **125** (1): 51-57.

Radlinski, A.P., Ioannidis, M.A., Hinde, A.L. et al. 2002. Multiscale characterization of reservoir rock microstructure: combining small angle neutron scattering and image analysis. *Proc.*, Proceedings of 2002 International Symposium of the Society of Core Analysts (SCA2002-35), Monterey, California, Sept. 23-27.

Radlinski, Andrzej P. 2006. Small-Angle Neutron Scattering and the Microstructure of Rocks. *Reviews in Mineralogy and Geochemistry* **63** (1): 363-397.

Rezaeyan, Amirsaman, Pipich, Vitaliy, Bertier, Pieter et al. 2019a. Microstructural Investigation of Mudrock Seals Using Nanometer-Scale Resolution Techniques. Presented at the Sixth EAGE Shale Workshop, Bordeaux, France. 29 April 2019.

Rezaeyan, Amirsaman, Pipich, Vitaliy, Bertier, Pieter et al. 2019b. Quantitative Analysis of the Pore Structure of Premature-To-Postmature Organic Rich Mudrocks Using Small Angle Neutron Scattering. Presented at the Sixth EAGE Shale Workshop, Bordeaux, France. 30 April 2019.

Rezaeyan, Amirsaman, Seemann, Timo, Bertier, Pieter et al. 2019c. Understanding Pore Structure of Mudrocks and Pore-Size Dependent Sorption Mechanism Using Small Angle Neutron Scattering. Presented at the SPE/AAPG/SEG Asia Pacific Unconventional Resources Technology Conference, Brisbane, Australia. 2019/11/15/.

Sakurovs, Richard, Koval, Lukas, Grigore, Mihaela et al. 2018. Nanometre-sized pores in coal: Variations between coal basins and coal origin. *International Journal of Coal Geology* **186**: 126-134.

Seemann, T., Bertier, P., Maes, N. et al. 2019. Resolving the Pore Structure and Sorption Properties of Methane in Mudrocks - A Small Angle Neutron Scattering Study. **2019** (1): 1-5.

Vishal, Vikram, Chandra, Debanjan, Bahadur, Jitendra et al. 2019. Interpreting Pore Dimensions in Gas Shales Using a Combination of SEM Imaging, Small-Angle Neutron Scattering, and Low-Pressure Gas Adsorption. *Energy & Fuels*.

Walther, Dirk, Cohen, Fred E., and Doniach, Sebastian. 2000. Reconstruction of low-resolution three-dimensional density maps from one-dimensional small-angle X-ray solution scattering data for biomolecules. *Journal of Applied Crystallography* **33** (2): 350-363.

Zemb, Th. and Lindner, P. 2002. *Neutron, X-rays and Light. Scattering Methods Applied to Soft Condensed Matter* North Holland: North-Holland Delta Series.

Zhao, Jinkui. 2011. SAAF: small-angle neutron scattering data analysis using analytical functions. *Journal of Applied Crystallography* **44** (6): 1277-1280.