User Interface Explanation and Operational Instruction.

1. Running the software

There are two primary methods to run the software:

- **For machines with Python already installed:** Download all the provided scripts and execute the Python script ***run_PRINSAS.py***. This method is cross-platform compatible and independent of the operating system. The scripts were developed and tested using Spyder on Windows and are available in this GitHub repository.

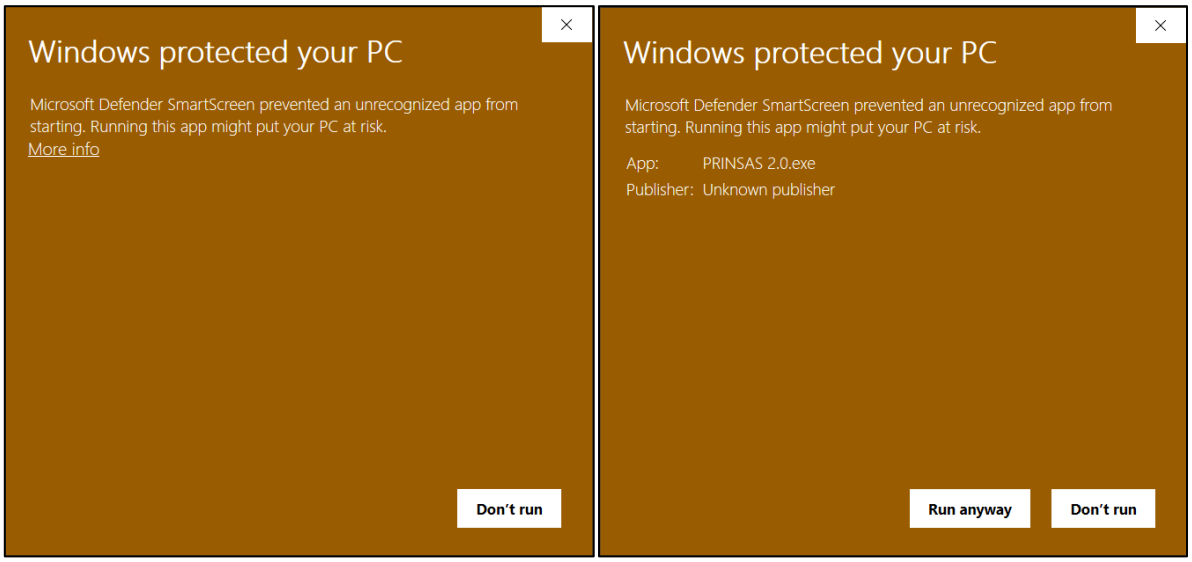

Figure 1. *By passing Windows Defender in Windows 10 during the first launch of the software.*

- For Windows machines: Use the standalone executable .exe package. This version includes all necessary dependencies and runs portably without requiring additional installations. However, as the program is unsigned, it will likely trigger a warning from Windows Defender or other antivirus software. To bypass this warning, follow the following steps during the first run when the antimalware software is triggered (**Figure 1**): Click **More info** > Select **Run anyway**. These steps apply to the default Windows

10 antivirus, Microsoft Defender. The process may vary slightly for other antivirus programs or versions of Windows.

2. Starting user interface

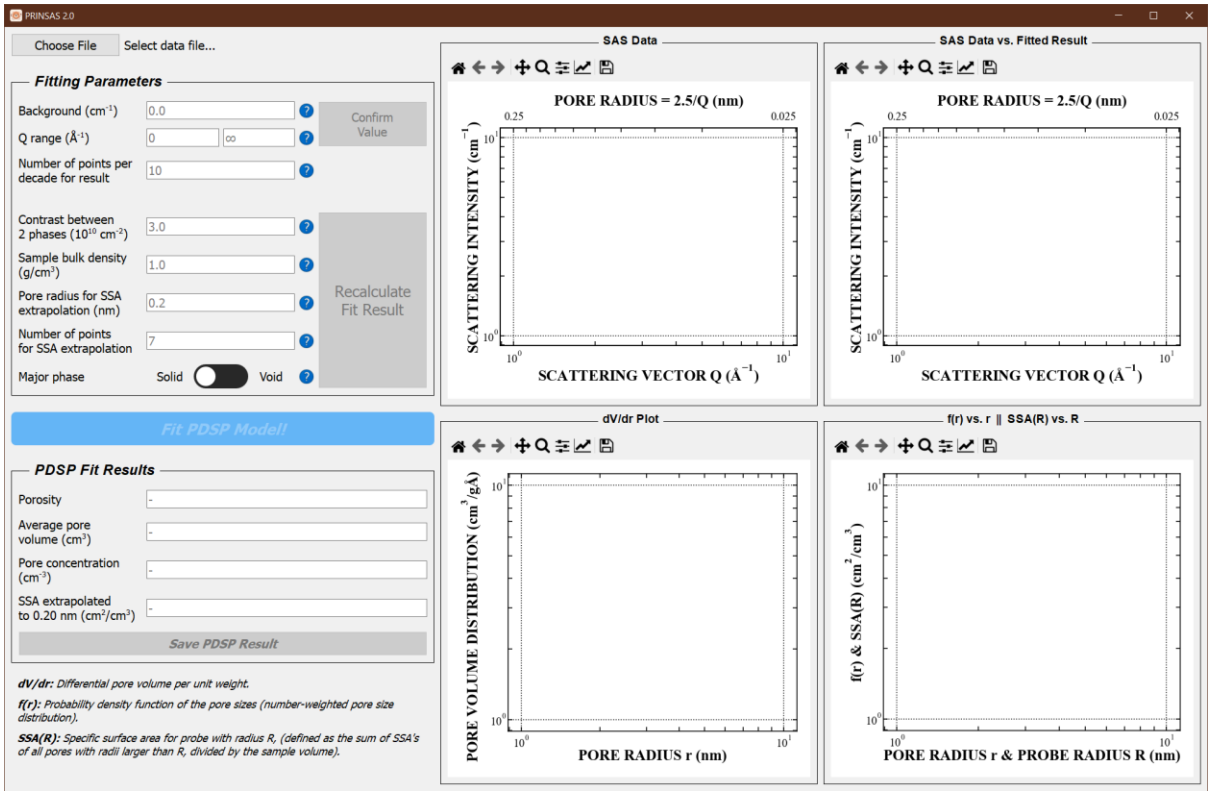

Figure 2.1. *Starting interface layout of the software.*

The software interface (**Figure 2.1**) is divided into two primary regions:

- Left: User input/Fitting result area, including:
  - Data Selection Area: Allows users to load data files.
  - ***Fit Parameters*** panel: Provides input fields for fitting parameters.
  - ***PDSP Fit Results*** panel: Displays the results of the fitting process.
- Right: Plotting area, including:
  - ***SAS Data*** panel: Displays the SAS data and facilitates determination of the flat background and Q range

- ***SAS Data vs. Fitted Result*** panel: Compares the background-subtracted data with the intensity calculated from the PDSP fit result.
- $\boldsymbol{dV/dr}$ ***Plot*** and $\boldsymbol{f(r)}$ ***vs.*** $\boldsymbol{r}$ || $\boldsymbol{SSA(R)}$ ***vs.*** $\boldsymbol{R}$ panels: Display the PDSP fit result, showing the volume-weighted pore size distribution, number-weighted pore size distribution and cumulative specific surface area as a function of pore size.

Initially, all buttons are disabled except the ***Choose File*** button on the top left.

3. Selecting Data Files

To load scattering data for PDSP fitting:

(i) Click the ***Choose File*** button. A pop-up window will appear to allow selection of the desired data file (**Figure 3.1**).

(ii) The program supports ASCII files, even with non-uniform data tables, various delimiters, headers, and footers. The first two columns of data are assigned to Q and $I(Q)$, with the optional third column assigned to $dI(Q)$.

(iii) Once loaded (**Figure 3.2**):

- The file name is displayed next to the ***Choose File*** button.
- The SAS data, including experimental error as error bars, is plotted on the ***SAS Data*** panel. For data files without the optional third column, the error value defaults to 0.
- Background-subtracted SAS data (with the default subtracted background value of 0 cm^{-1}) is also plotted on the ***SAS Data*** and ***SAS Data vs. Fitted Result*** panels.
- The Confirm Background and Fit PDSP Model! buttons are enabled.

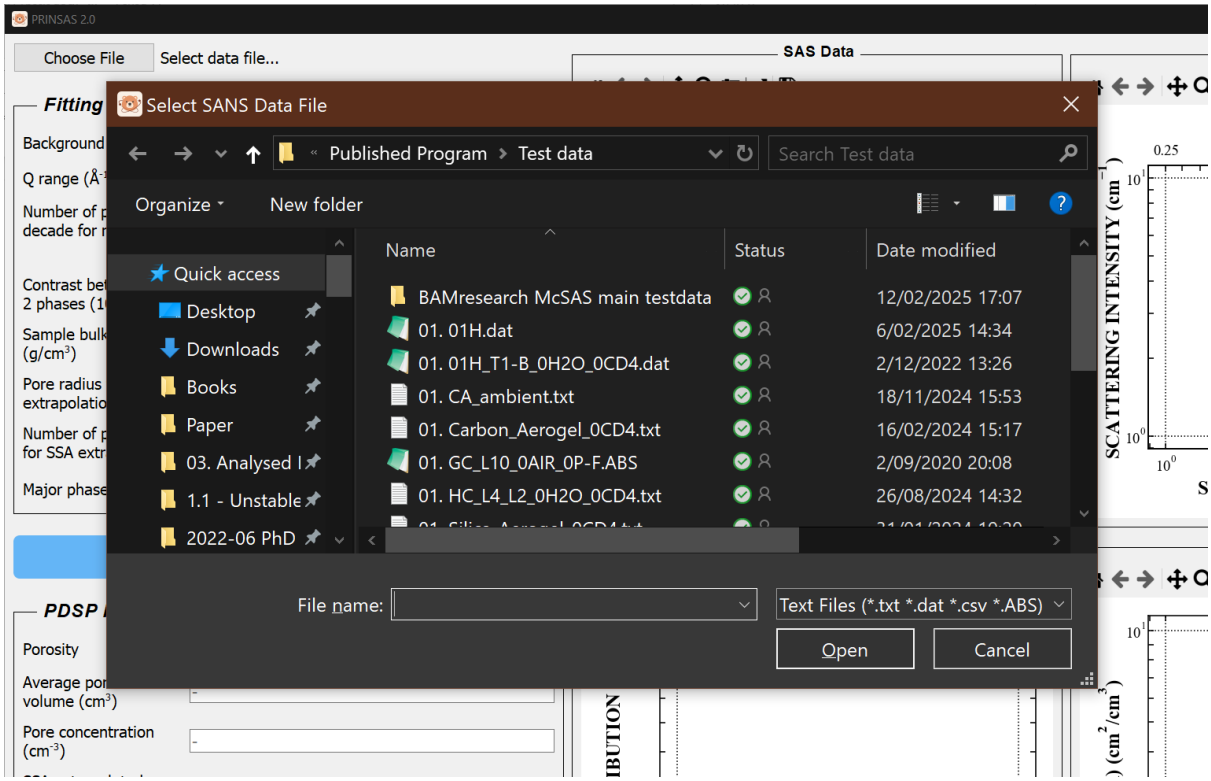

Figure 3.1. *Pop-up window for selecting a data file.*

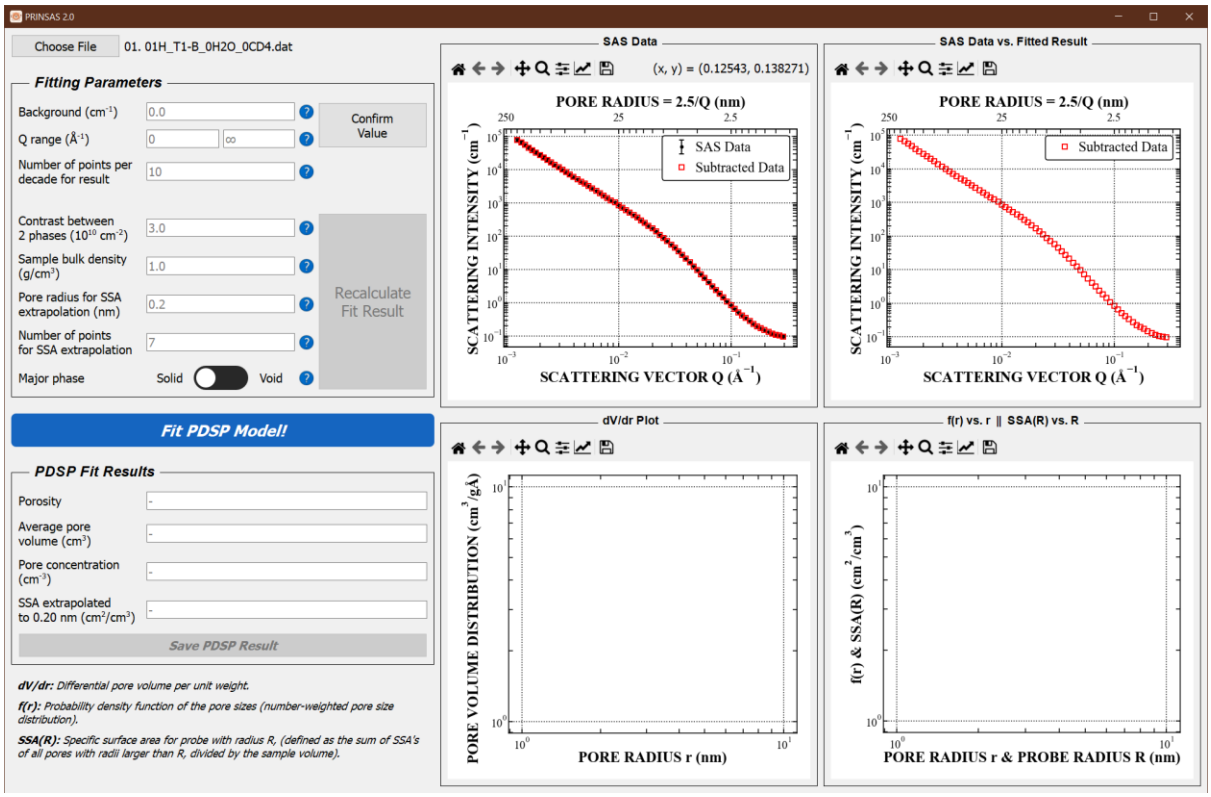

Figure 3.2. *Interface after loading a scattering data file.*

ASCII data files are read using the function ***read_SANS_data()***, supporting a range of delimiters. However, only one delimiter is used per file, and decimal numbers in 'comma' format are not supported.

To process data, the program scans the entire file to identify the most common delimiter. It then splits each row using this delimiter and attempts to convert the elements into numbers, removing any entries that are negative or not a number. Next, it counts the resulting elements of each row and determines the most frequent column count to structure a table. Finally, data rows with insufficient values are discarded.

4. Configuring Fitting Parameters and Running the Fitting Procedure

The panel ***Fitting Parameters*** provides the necessary input fields for configuring the PDSP fitting process and calculating the resulting structural properties derived from the fit.:

- ***Background (cm^{-1})***: is the flat background value to be subtracted from the original SAS profile prior to the analysis. The background value is determined such that the scattering profile forms a straight, continuous line in the large-Q region after subtraction *(Radlinski et al., 2004)*. To find the appropriate background value, hover the cursor over the desired level on the ***SAS Data*** plot; live cursor coordinates will be displayed in the top right of the plot (as shown in **Figure 3.2**). The background scattering intensity (y-coordinate) can then be entered into the text box next to the ***Background (cm^{-1})*** label, followed by pressing the ***Confirm Background*** button. This updates the ***SAS Data*** and ***SAS Data vs. Fitted Result*** panels with the newly background-subtracted SAS data (**Figure 4.1**).
- ***Q range ($Å^{-1}$)***: is the minimum and maximum Q values to be used for the PDSP analysis. The Q range is chosen to exclude SAS data affected by multiple scattering (at low-Q) and artifacts caused by background subtraction (at high-Q). The desired values can be determined using the x-coordinates of the cursor when hovering over the desired

position on the ***SAS Data*** plot. Once entered, the selected Q range can be visualided in the ***SAS Data*** and ***SAS Data vs. Fitted Result*** panels by pressing the ***Confirm Background*** button.

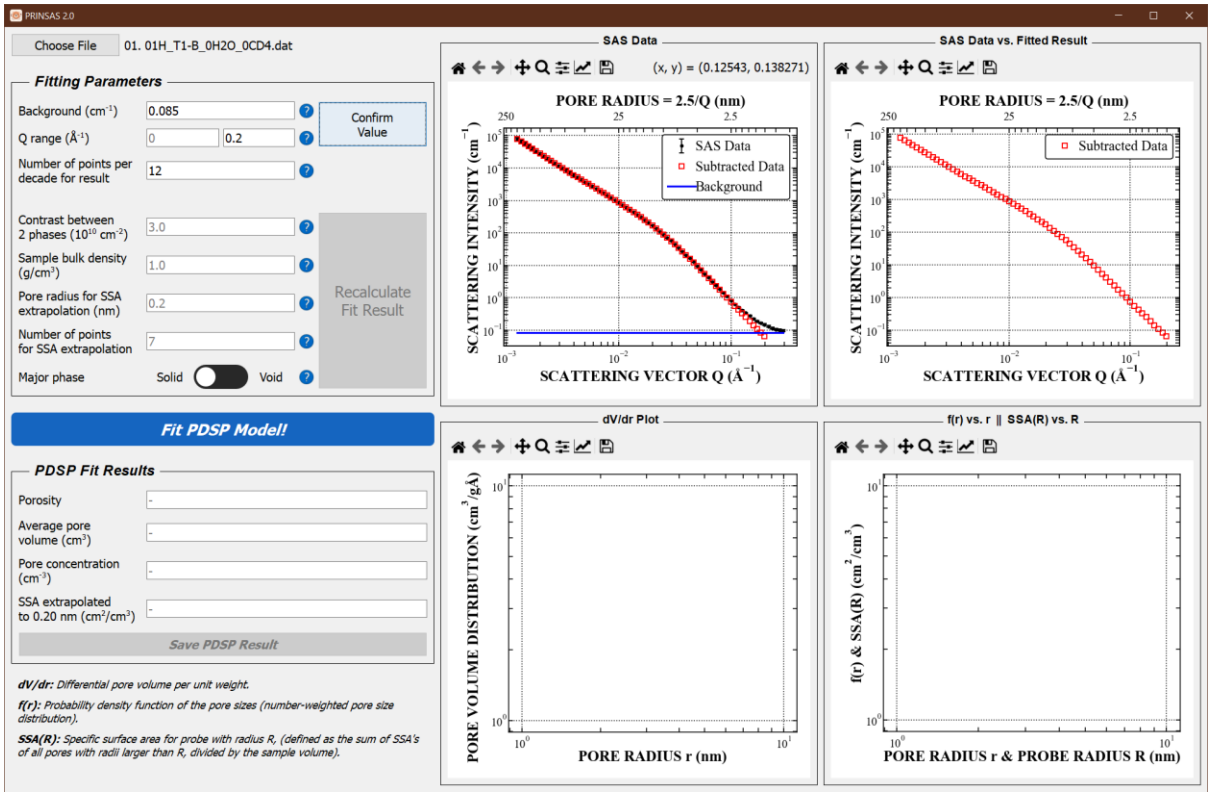

Figure 4.1. *Program interface after the background and Q max value has been specified.*

- ***Number of points per decade for result***: refers to the number of r_i values between each decade of r, where $r = 2.5/Q$, in the fit results. A higher number of r_i values improves the fit to the original intensity profile. However, using too many r_i values may lead to overfitting and increase the time required to achieve a complete fit. The number of r_i values is dynamically updated every 50 ms while the input field is being edited.
- ***Contrast between 2 phases (10^{10} cm^{-2})***: refers to the contrast $(\rho_1^* - \rho_2^*)$ in Equation **Error! Reference source not found.**. For a porous system without any filling medium (*i.e.*, the filling medium is vacuum or air), the contrast is equal to the scattering length density (SLD) of the solid.
- ***Sample bulk density (g/cm^3)***: required for the differential pore volume distribution calculations, as outlined in Equation **Error! Reference source not found.**.

- ***Pore radius for SSA extrapolation (nm)***: is the pore radius to which the SSA value is extrapolated in the $SSA(R)$ vs. R graph, represented by the vertical dashed line on the $SSA(R)$ vs. R plot in **Figure 5.1**. The default R value is 0.2 nm, based on the study of Radlinski *et al.* *(Radlinski et al., 2021)*.
- ***Number of points for SSA extrapolation***: refers to the number of data points used for the SSA extrapolation to the previously specified pore radius in the $SSA(R)$ vs. R graph, indicated by the solid blue marker on the $SSA(R)$ vs. R plot in **Figure 5.1**.
- ***Major phase*** (Void or Solid): since equation **Error! Reference source not found.** is symmetric for both phases of ϕ and $(1-\phi)$, the phase that accounts for the majority of the system's volume must be selected to ensure that the correct porosity value is returned.

Once all the inputs have been entered, the ***Fit PDSP Model!*** button will save the provided input parameters and initiate the PDSP fit procedure. Any values left blank will be assigned the default value, which is displayed in light grey when no input is provided. The fitting and calculation process may take up to a few minutes, depending on the Q range and the number of data points in the selected SAS data, and the input for ***Number of points per decade for result***. During this time, the program becomes unresponsive.

5. Fit Result and Recalculating Structural Properties Post-Fitting

Once the fitting procedure is completed (**Figure 5.1**), the calculated results are presented in the ***PDSP Fit Result*** panel including:

- Porosity (ϕ)
- Average pore volume ($\bar{V}$)
- Pore concentration/number density ($\bar{V}$)
- SSA value extrapolated to the previously specified ***Pore radius for SSA extrapolation***.

The software also plots the fitted $I(Q)$ value to the ***SAS Data vs. Fitted Result*** panel for comparison with the background-subtracted SAS data, and displays dV/dr, $f(r)$, and $SSA(R)$ results in their respective plot panels.

Note that these values are highly dependent on the absolute scattering intensity, which is not always reliable, especially for powdered samples.

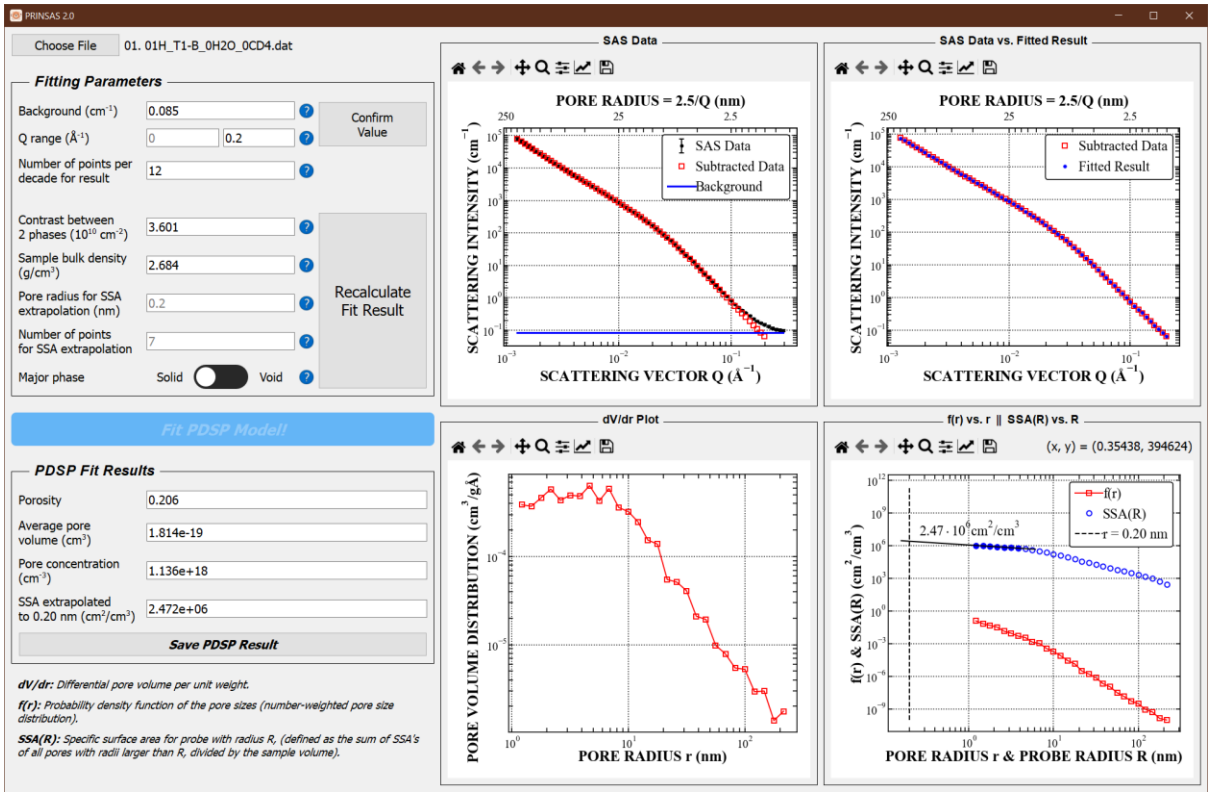

Figure 5.1. *Program interface after the PDSP fitting procedure is completed.*

After the fit, the ***Recalculate Fit Result*** button is enabled, allowing for the recalculation of the structural properties within the ***PDSP fit result*** without repeating the entire fitting process. This recalculation can be applied when the following inputs are changed: ***Contrast between 2 phases***, ***Sample bulk density***, ***Pore radius for SSA extrapolation***, ***Number of points for SSA extrapolation***, and ***Major phase***. After adjusting any of these values, click the ***Recalculate Fit Result*** button, and the results and plots will update accordingly (**Figure 5.2**).

Once the fitting procedure is complete, the ***Fit PDSP Model!*** button will be disabled. It will only become re-enabled if (i) a new background or Q max value is selected, or (ii) a different value for the number of r_i per decade is input.

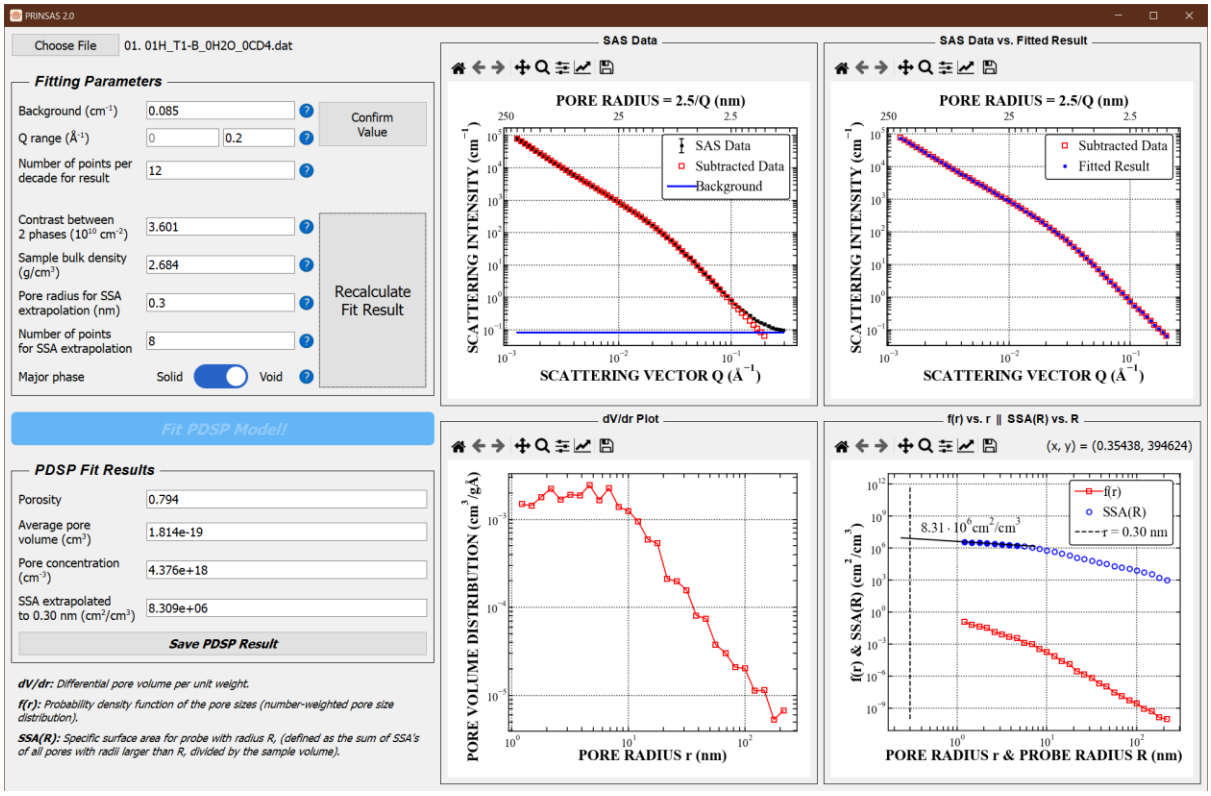

Figure 5.2. *Recalculation of results after entering new inputs (targeted pore radius and number of points used for SSA extrapolation, and major phase) without refitting.*

6. Saving the result

Once the user is satisfied with the fit result, it can be saved to an ASCII file generated by the program. This file includes all the fitting parameters as well as the fit result (**Figure 6.2**). To save the result, press the ***Save PDSP Result*** button (enabled after the fitting procedure is completed). A pop-up will appear, allowing the user to choose where to save the file (**Figure 6.1**).

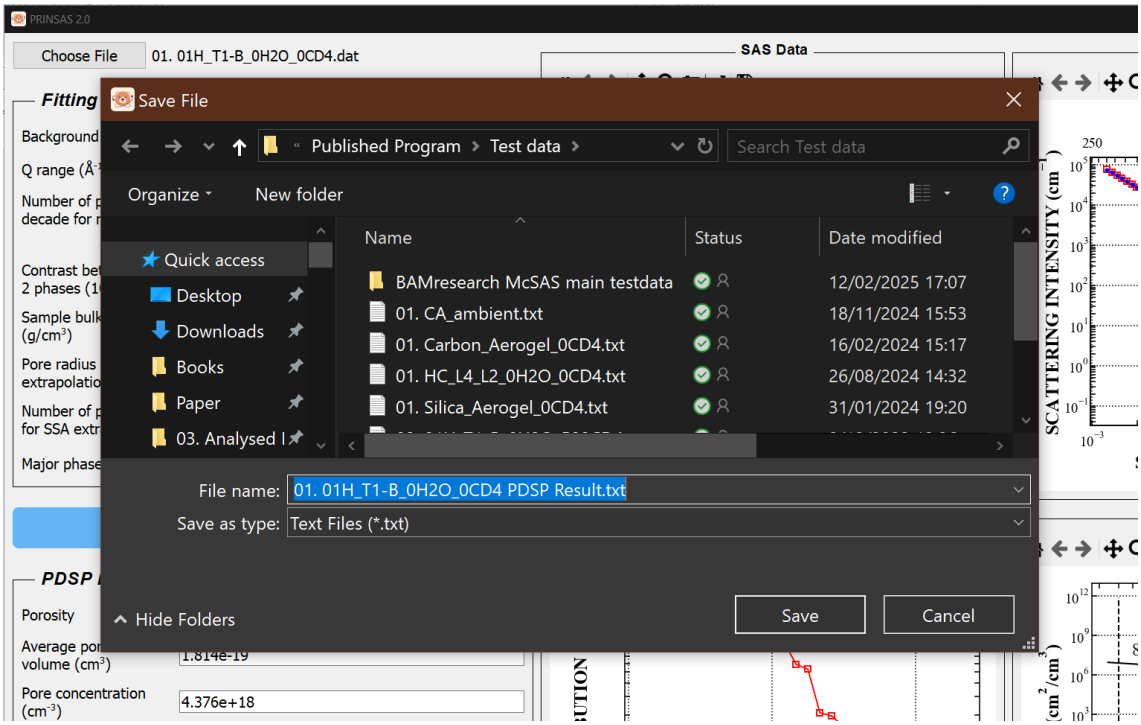

Figure 6.1. *Pop-up window for saving results.*

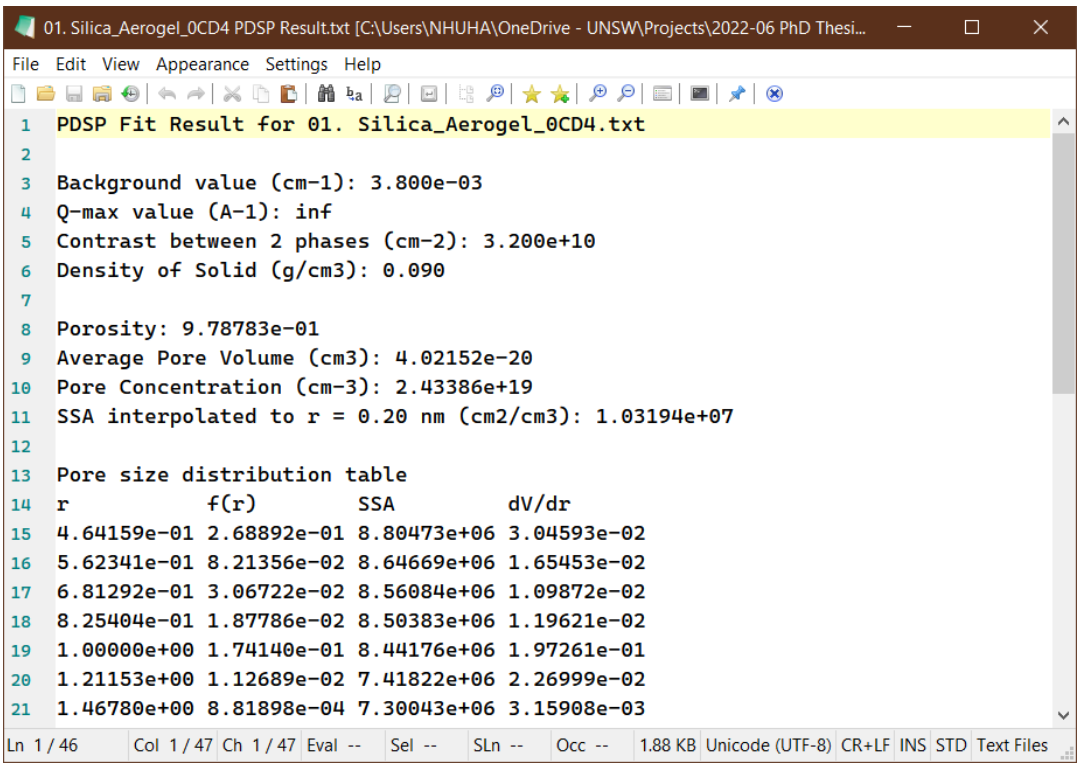

Figure 6.2. *Example of a saved result file.*

7. Plotting window functionality

Each plotting window includes functionality to facilitate the inspection and visualisation of the data and results. In addition to displaying live cursor coordinates, as mentioned previously, the toolbar located to the left of the coordinates provides the following features (arranged from left to right):

- ***Home***: Resets the plot to its original state.
- ***Back/Forward***: Undoes or redoes adjustments made to the plot.
- ***Pan***: Enables panning functionality.
- ***Zoom***: Allows zooming functionality.
- ***Configure subplots***: Adjusts the spacing and border of the plotting area relative to the window frame.
- ***Edit axis/curve***: Provides options for customising plot elements, such as the plot title, axis limits, line and marker shape or colour, and other properties of the data series. However, this function has two significant limitations:
  - The inability to customise error plots, as this feature is embedded in the plotting and rendering engine Matplotlib (referenced in this Github issue).
  - For SAS plots, the values on the top and bottom axes (scattering vector Q and corresponding pore radius $\approx 2.5/Q$) are not automatically linked. They must be adjusted independently when changing the range of horizontal axis. To ensure the r values on the top axis match the Q values on the bottom axis, update the top axis limits using the following formulas: $(x_{min})_{top} = 0.1/(x_{max})_{bottom}$; $(x_{max})_{top} = 0.1/(x_{min})_{bottom}$. This issue occurs only when modifying the axes using the ***Edit axis/curve*** function. It does not affect the zoom or pan functionality.

Save: Exports the plot in various image formats (*.png*, *.jpeg*, *.svg*, *etc.*)

8. Reference

Radlinski, A. P., Blach, T., Vu, P., Ji, Y., de Campo, L., Gilbert, E. P., Regenauer-Lieb, K. & Mastalerz, M. (2021). *International Journal of Coal Geology* **248**, 103850.

Radlinski, A. P., Mastalerz, M., Hinde, A. L., Hainbuchner, M., Rauch, H., Baron, M., Lin, J. S., Fan, L. & Thiyagarajan, P. (2004). *International Journal of Coal Geology* **59**, 245-271.