Script to Estimate Surface Areas of Materials from their Isotherms (SESAMI) 1.0

**User Manual**

Priya Sinha†#, Archit Datar †#, Chungsik Jeong ‡, Xuepeng Deng†, Yongchul G. Chung \*‡, and Li-Chiang Lin\*†

† William G. Lowrie Department of Chemical and Biomolecular Engineering, The Ohio State University, Columbus, Ohio 43210, United States

‡ School of Chemical and Biomolecular Engineering, Pusan National University, Busan 46241, Korea (South)

# These authors contributed equally to this work

\* Corresponding authors: Y.G. Chung ([drygchung@gmail.com)](mailto:greg.chung@pusan.ac.kr)) and L.-C. Lin ([lin.2645@osu.edu](mailto:lin.2645@osu.edu))

# 1. Introduction

This software is a Python script to compute BET, ESW and BET + ESW areas given an adsorption isotherm. A detailed discussion about the working and relative merits of these three methods can be found in the main text and the Supporting Information (SI).

# 2. Requirements and Setup Instructions

This software requires Python >= 3.0, which uses the dependencies mentioned in *Requirements.txt*.

To set up the software, download the directory *SESAMI\_package* and copy it the location where it is to be executed. The files *betan.py* and *mypaper.mplstyle* need to be in the same directory. All other paths can be specified, as discussed in the next section or the user can choose to use the default settings which are also described there.

In subsequent sections, the inputs to and outputs of this software are illustrated with an example for a MOF structure named CIGXIA.

# 3. Inputs

The software reads in two files: the input file, *input.txt*, which specifies the instructions and paths and the isotherm data file, *CIGXIA.txt*, which contains the isotherm data.

## 3.1. Input file (*input.txt)*

The following is an example of an input file *input.txt*.

#Name

CIGXIA

#path to isotherm file

Default

#path to output directory

Default

#ESW data

Default

#Consistency 1 limit data

Default

#R2\_min R2\_max minimum line length

Default Default Default

#path to the executable file

Default

This file is read in as follows:

1. The 1st line is skipped.
2. The 2nd line is the name of the structure for which the analysis is being performed. This name will appear in the output analysis. It is a necessary field.
3. The 3rd line is skipped.
4. The 4th line is the path to the isotherm file. It, like all paths in this software, can be relative to *input.txt* or absolute. ‘Default’ means that the system will search the isotherm data file in the same directory as the input file and will expect it to be named *CIGXIA.txt* (because the name argument in the 2nd line is CIGXIA).

**The user is advised to be extremely careful about specifying the paths in the input file as an error in this argument will cause the program to fail. The user is specifically advised to be mindful if they are running it on a Windows system as the notation of the paths can be complicated. Unless they are confident of the paths, they are advised to use the defaults.**

1. The 5th line is skipped.
2. The 6th line is the path to the output directory. This is the path where all the output files will be stored. ‘Default’ means that the output directory is the same as the input directory.
3. The 7th line is skipped.
4. The 8th line determines if the location of the first ESW minimum is to specified manually. ‘Default’ means ‘No’, i.e. that the location of the first ESW minimum will be computed using the algorithm employed in this software. If this argument is ‘Yes’, the software will read this value from the first line of the isotherm data file as mentioned in the next sub-section.
5. The 9th line is skipped.
6. The 10th line determines if the limit of the consistency 1 criteria—the relative pressure where the quantity has its first maximum—is to be computed manually. Again, ‘Default’ means ‘No’. If this argument is ‘Yes’, the software will read this value from the second line of the isotherm data file.

**For the determination of the ESW minima and consistency 1 maximum, in most cases, ‘Default’ should work. However, the user is advised to check the output to ensure that these values make sense. This is because while finding local maxima or minima in discrete data, noise can be sometimes construed as a local maximum or minimum. The algorithm we have employed tries to tune out the noise, but there is still a chance that maxima or minima chosen might actually be noise.**

1. The 11th line is skipped.
2. The 12th line reads in the values of some additional parameters that might be useful in the analysis. These are:

* *R2min*, the minimum R2 value a chosen region must have to be termed “linear” (see SI Section 3), with ‘Default’ value 0.998.
* *R2cutoff*, the value of R2 beyond which we deem R2 ceases to have a bearing on the goodness of the linear region (see SI Section 3), with ‘Default’ value 0.9995. For example, in this case, if we need to compare two regions: one having an R2 value of 0.9995 and another one having an R2 value of 0.9997, one cannot say that that the latter region is better than the former because at extremely high values of R2, such an inference is not valid. This has been discussed in SI Sec 3.
* *minimum line length*, which is the minimum number of points that a region must have for it to be considered in the analysis. The ‘Default’ value is 4.

In general, the default values should suffice. However, they can be altered if the data points are few or the data are noisy.

1. The 13th line is skipped.
2. The 14th line is the path to the directory containing the file *betan.py*. The ‘Default’ is that it is one directory above the file *input.txt*.

## 3.2. Isotherm file (*CIGXIA.txt*)

This file contains data about the isotherm that is to be analyzed. The following example shows the first 10 lines of this file.

ESWMinima 27

Con1Maxima 44

Pressure Loading

1.0 0.8405165792

5.0 0.8569143274

10.0 0.877571175

20.0 0.9199421724

30.0 0.9637902317

40.0 1.0082090411

50.0 1.054743685

The important features of this file are as follows.

1. This must be a tab-separated file.
2. The 1st line is skipped if line 8 of *input.txt* is ‘Default’. This means that the ESW minima will be determined by the software. Otherwise, this is the index of the ESW minima. **We emphasize that this value is 0-indexed.** For example, if the user wants to say that the 29th point is the ESW minima, they should input 28 here.
3. Similarly, the 2nd line is skipped if line 10 of *input.txt* is ‘Default’. This means that the consistency 1 maxima will be determined by the software. Otherwise, this is the index of the point which is the consistency 1 maxima. Again, this is 0-indexed.
4. Line 3 is skipped.
5. Line 4 onwards are the isotherm data. The left column is the pressure in Pa. Users must ensure that these are float values in order to ensure that the software can provide the correct results.
6. The right column is the loading in mol/kg. As of version 1.0, we cannot allow inputs in other units. However, we will update that in later versions. Again, these must be float values.

For pressure as well as loading, please ensure that there is only float values present, else, the software might not work properly. If it is necessary to create placeholders, please make sure to use ‘NaN’, but nothing else.

# 4. Outputs

The output has three parts: the text out file *Output.txt*, the image summary of the results *CIGXIASummary.png* and the isotherm data that is processed *dataframe.txt*.

## 4.1. *Output.txt*

This provides the data for the output. It is a tab-delimited TXT file, which can be easily read into Excel as well as processed with Bash. An example is shown below.

CIGXIA

BET data

LowIndex HighIndex LowPressure HighPressure Area Nm C Con1 Con2 Con3 Con4 R2

34 40 3500.0000000000 6000.0000000000 2375.155566 27.77557 2.64331e+02 Yes Yes Yes Yes 0.9999

Consistency 1 limit data

Index Pressure

44 8500.0000000000

ESW minima data

Index Loading Pressure Area

27 19.198 900.0000000000 1641.898

BET-ESW data

LowIndex HighIndex LowPressure HighPressure Area Nm C Con1 Con2 Con3 Con4 R2

23 33 700.0000000000 2500.0000000000 2075.415607 24.27035 3.90015e+02 Yes Yes No No 0.9998

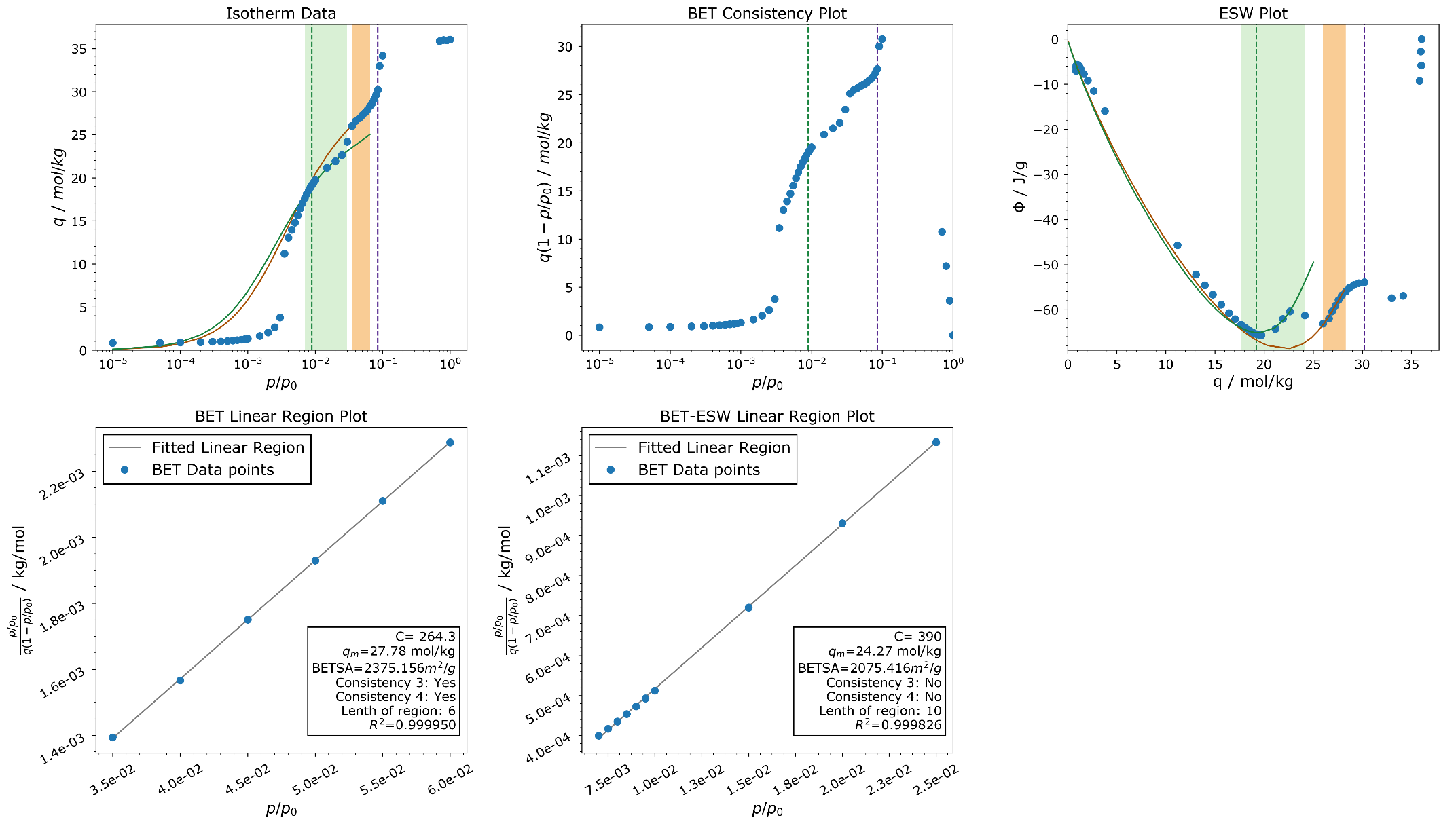
“BET data” describes the properties of the linear region chosen for the BET analysis. “Low index” and “high index” refer to the indices of the start and end of the linear region in terms of the isotherm points, which are listed in the file *dataframe.txt*. “LowPressure” and “HighPressure” refer to the lower and upper limits of the pressure of the chosen region. “Area” is the BET area and “Nm” and “C” are the BET parameters. “Con1”-“Con4” refer to the satisfaction of the consistency criteria and “R2” is the R2 value for the fit.

“Consistency 1 limit data” describes the data related to the consistency criteria 1 upper limit, i.e., the pressure point at which has its first maximum. “Index” and “pressure” refer to the corresponding index and pressure.

“ESW minima data” describes the properties related to the ESW minimum. “Index”, “Loading”, “Pressure” and “Area” refer to the index, loading, pressure and surface area corresponding to the ESW minimum.

“BET-ESW data” conveys the same information as the BET data except for the BET-ESW area.

## 4.2. *CIGXIASummary.png*



This figure summarizes the results of the selection of the linear region from the isotherm. “Isotherm data” shows the isotherm. The BET region is depicted by ochre and the BET-ESW region by green. The purple dashed line corresponds to the consistency 1 maximum while the green dashed line corresponds to the ESW minimum. The orange curve represents the BET fit to the region while the green curve denotes the BET-ESW fit.

“BET Consistency Plot” and “ESW Plot” follow the same color scheme. We haven’t highlighted the linear regions on the “BET Consistency Plot” as they are not useful.

“BET Linear Region Plot” and “BET-ESW Linear Region Plot” show the selected linear region. Users can visually inspect the fit to ensure that it is adequate.

## 4.3. *dataframe.txt*

This outputs the isotherm data so that the users know exactly how the data were used. A part of this output file is shown below:

Index Pressure Loading P\_rel BETy BET\_y2 phi

0 1.0 0.8405165792 1e-05 1.18975642461664e-05 0.8405081740342081 -6.999407042537805

1 5.0 0.8569143274 5e-05 5.835180779007507e-05 0.8568714816836299 -6.138395079161189

2 10.0 0.877571175 0.0001 0.000113962267505083 0.8774834178825001 -5.846383535677127

3 20.0 0.9199421724 0.0002 0.00021744845926535147 0.91975818396552 -5.667431758266294

“Index” denotes the index for the isotherm points. “Pressure”, “Loading” and “P\_rel” denote the pressure in Pa, loading in mol/kg and relative pressure respectively. “BETy” refers to the linearized transformation of the BET equation () which is used to fit to the relative pressure to get the BET area, where is the loading in mol/kg. BET\_y2 refers to the transformation (), which is plotted on the figure “BET Consistency Plot”. “phi” refers to the ESW function in J/g which is plotted on the figure “ESW plot”.