AGREE – Analytical GREEnness metric approach and software

Supporting Information: Software documentation for the Analytical Greenness Calculator

Francisco Pena-Pereira[1*], Wojciech Wojnowski[2*], and Marek Tobiszewski[2*]

[1]Analytical and Food Chemistry Department; Faculty of Chemistry; University of Vigo, Campus As Lagoas-Marcosende s/n, 36310 Vigo, Spain
[2]Department of Analytical Chemistry, Faculty of Chemistry, Gdańsk University of Technology, 11/12 G. Narutowicza Street, 80-233 Gdańsk, Poland
[*]*fjpena@uvigo.es; wojciech.wojnowski@pg.edu.pl; marek.tobiszewski@pg.edu.pl*

Contents

1 Introduction

The Analytical Greenness calculator is a tool for evaluating the environmental and occupational hazards associated with a particular analytical procedure based on the 12 principles of Green Analytical Chemistry [1]. The result of the assessment is translated into a graph that is easy to interpret and contains an overall score, as well as an indication to what degree the evaluated procedure conforms to each of the 12 principles. It can be used to compare different methods in order to select the one with the lowest environmental impact, or to identify possible trouble spots in terms of "greenness" during conceptualization and development of novel analytical procedures.

The calculator was developed in Python 3.7 using the Tkinter module [2]. It utilizes the following modules:

```
from tkinter import *
import matplotlib.pyplot as plt
from matplotlib.backends.backend_tkagg import FigureCanvasTkAgg
from matplotlib.colors import LinearSegmentedColormap
from tkinter import ttk
from tkinter import filedialog
import tkinter.messagebox
import webbrowser
from math import log
from fpdf import FPDF
import os
from datetime import datetime
```

2 Disclaimer

This software is provided “as is,” and any express or implied warranties, including but not limited to the implied warranties of merchant-ability and fitness for a particular purpose, are disclaimed. In no event shall the authors be liable for any direct, indirect, incidental, special, exemplary, or consequential damages (including, but not limited to, procurement of substitute goods or services; loss of use, data, or profits; or business interruption) however caused and in any way theory of liability, whether in contract, strict liability, or tort (including negligence or otherwise) arising in any way out of the use of this software, even if advised of the possibility of such damage.

This software is made available under the MIT license.

3 Installation

The most recent version of the Python code can be obtained from the repository maintained at git.pg.edu.pl/p174235/AGREE.

A compiled version is currently available for the Windows operating system and can be obtained from mostwiedzy.pl/AGREE. After launching the self-extracting archive *ARGEE_sfx.exe* the user is asked to indicate the destination folder (see Figure S1). This will, by default, be the current folder, but any folder can be indicated by clicking the <Browse> button and navigating through the system-default dialog. After extracting the archive, the Analytical Greenness Calculator can be launched by running the *AGREE.exe* executable located in the newly created folder.

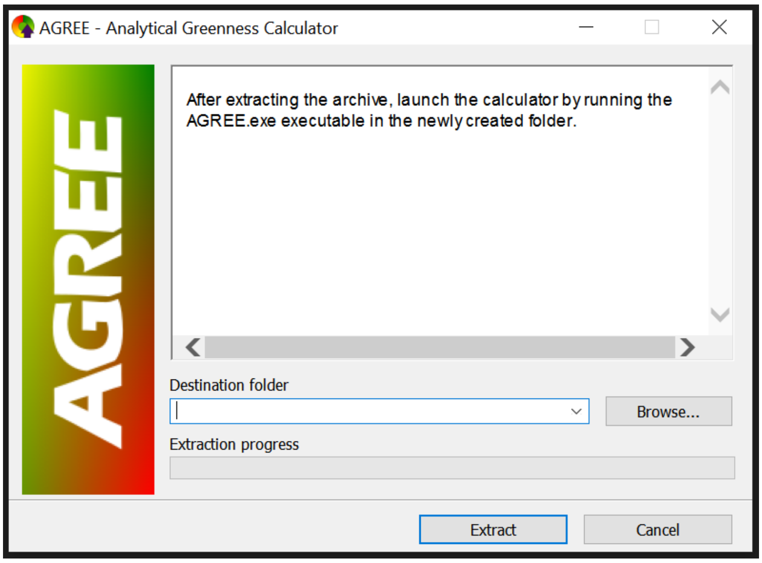

Figure S1: Self-extracting archive dialog.

4 User Interface

4.1 Main elements of the GUI

The main window is shown in Figure S2 (the window style might differ depending on the operating system and the system settings). It consists of:

1. File menu containing links to an info page, graph save dialog, GUI re-set dialog, and report generation dialog;
2. Tabs linked to panels corresponding to each greenness criterion which are displayed in the left frame of the main window;
3. Short caption outlining the rationale behind each criterion;
4. Selection/input interface specific to each criterion and panel;
5. Dialog for modifying the default weights of each criterion (default value = 2);
6. Button for re-generating the graph;
7. Status bar with the current value of scores corresponding to the criteria;
8. Graph preview - a visual representation of the Analytical Greenness score.

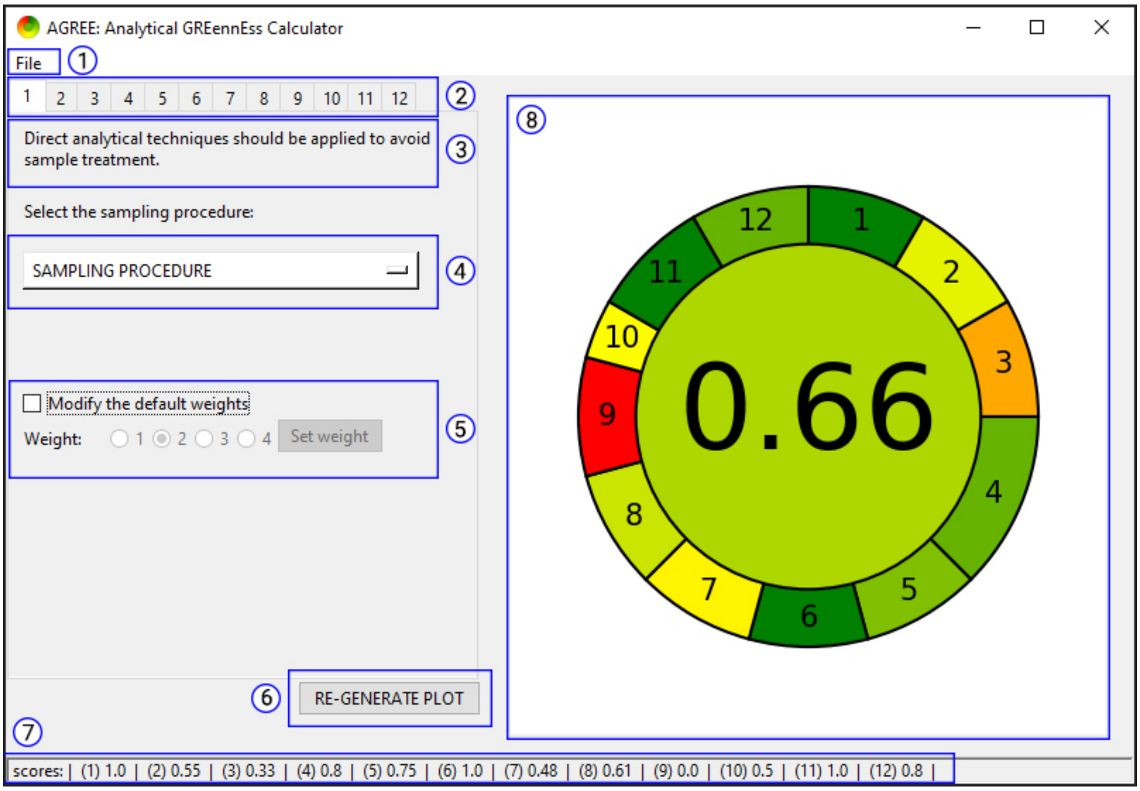

Figure S2: Main GUI elements.

4.2 Graph and the Analytical Greenness score

The Analytical Greenness score is the weighted average of the criteria scores (see Equation 1). It is displayed in the middle of the graph rounded to two decimals, and its value ranges from 0.0 (the lowest score) to 1.0 (perfect score). The graph is a visual representation of the score itself, of the criteria scores, and of the criteria weights.

$$Analytical\ Greenness\ score = \frac{\sum_{i=1}^{12} w_i s_i}{\sum_{i=1}^{12} w_i} \tag{1}$$

The central field corresponds to the Analytical Greenness score, and the outer, numbered segments correspond to respective criteria scores. The relative width of the outer segments indicates the weight of the particular criterion. The default weights of each criterion can be modified by selecting the *"Modify the default weights"* checkbox in the corresponding tab, which unlocks the radio button interface and enables the modification of the default weight by selecting the new weight and pressing the <Set weight> button. De-selecting the checkbox will revert the weight to the default value. It is recommended to retain the default weights unless the User determines that the given criterion has a particularly high/low impact on the method's greenness.

The criteria scores and the Analytical Greenness score are linked to a "traffic lights" red-yellow-green sequential colour map, with red assigned to the lowest values, and green to the highest values, as shown in Figure S3.

```python
# connect a float in range 0.0 : 1.0 to a colour in a spectrum from red to yellow to green (256 discrete colour values):
def colorMapper(value):
```

```
    cmap = LinearSegmentedColormap.from_list('rg', ["red", "yellow", "green"], N=256)
    mapped_color = int(value * 255)
    color = cmap(mapped_color)
    return color
```

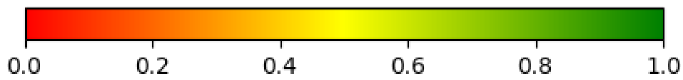

Figure S3: The span of the colour map used in the graph and the corresponding values.

The graph can be exported in the .png format *via* the File menu (*File →Save image*) using the system-default save dialog. Additionally, there is an option to export an annotated image which includes captions clarifying the corresponding rules (*File →Generate annotated result of the assessment*) (Figure S4).

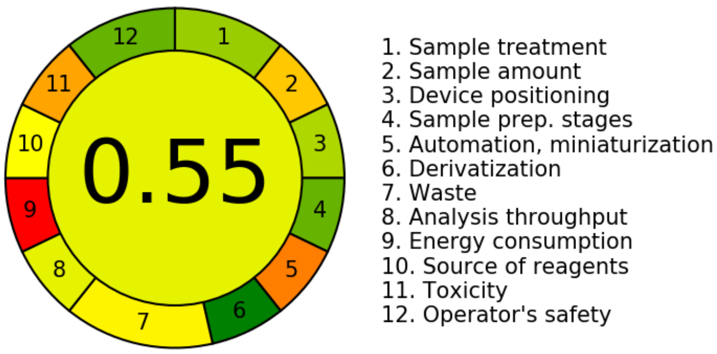

Figure S4: Annotated result of a generic assessment.

4.3 Report

Apart from exporting the graph, the result of the Analytical Greenness assessment can also be exported in the form of a report compiled into a .pdf file using the PyFPDF package [3]. The report can be generated and saved *via* the File menu (*File →Generate report*) using the system-default save dialog. It shall be opened automatically upon saving in the default .pdf browser.

The report contains the current date and time, the Analytical Greenness graph in the upper-right corner, and each of the 12 criteria listed in order, with the corresponding caption, score and weight. The criteria scores are rounded to two decimals. The background colour of the criteria score cells matches the colour assigned to this value using the colour map. The numeric values of the particular weights are displayed next to the respective criteria score cells.

5 Criteria

The panels containing interfaces for assigning particular scores corresponding to each of the 12 criteria are accessed through the tab menu in the upper-left corner of the main window. The scores are, by default, set to 1.0. If a particular criterion is not relevant to the assessed analytical procedure, it should be omitted and left at the default value. The assigned scores are displayed in the status bar at the bottom of the main window and can be modified at any time or globally re-set to the default values *via* the File menu (*File* →*Re-set the scores*). The User is encouraged to go through the tabs one-by-one.

5.1 Tab 1

The sampling procedure utilized in the assessed analytical method is selected through a drop-down option menu. The sampling procedures and the corresponding scores are as follows:

Sampling procedure	Score
Remote sensing without sample damage	1.00
Remote sensing with little physical damage	0.95
Non-invasive analysis	0.90
In-field sampling and direct analysis	0.85
In-field sampling and on-line analysis	0.78
On-line analysis	0.70
At-line analysis	0.60
External sample pre- and treatment and batch analysis (reduced number of steps)	0.30
External sample pre- and treatment and batch analysis (large number of steps)	0.0

5.2 Tab 2

The amount of sample in either g or mL is entered in the entry field. The entry is confirmed either by selecting the <Set> button to the right of the entry field, or by pressing the <Enter> key. The entry can be either a float or an integer; other inputs, such as a text string, will prompt a *Value error* message box. Please note that the '.' sign is used as the decimal separator in floats.

The score corresponding to Criterion 2 is calculated with the following equation:

$$s_2 = \begin{cases} 1.0 & \text{if } input < 0.1 \\ -0.142\ln(input) + 0.65 & \text{if } 0.1 \leq input \leq 100 \\ 0.0 & \text{if } input > 100 \end{cases} \tag{2}$$

5.3 Tab 3

The positioning of the analytical device is indicated by choosing an option from a drop-down option menu. The choices and their corresponding choices are as follows:

Positioning	Score
in-line	1.00
on-line	0.66
at-line	0.33
off-line	0.0

5.4 Tab 4

The number of major, distinct steps involved in the sample preparation procedure is indicated through a drop-down option menu. The scores corresponding to the number of steps are as follows:

Steps	Score
≤3	1.0
4	0.8
5	0.6
6	0.4
7	0.2
≥8	0.0

5.5 Tab 5

The degree of automation and miniaturization of the assessed analytical procedure is quantified based on the User's selection in two drop-down option menus:

Sample preparation	Degree of automation		
	automatic	semi-automatic	manual
none or miniaturized	1.0	0.75	0.5
not miniaturized	0.5	0.25	0.0

5.6 Tab 6

The use of derivatization agents is quantified based on a previously developed derivatization agents selection guide [4] which contains more than 260 entries, each with a given score between 0 and 1. If no derivatization agents are used in the evaluated analytical procedure, the default score of 1.0 is retained. Otherwise, a penalty of 0.2 is applied. Derivatization agents are selected from a list (2) (see Figure S5) containing the corresponding CAS numbers. The list can be narrowed down by typing any numeric string contained within the CAS number in the *CAS lookup* field (1). The CAS number chosen from the list is selected using the <Select> button (3), and it is displayed below the list (4). More than one derivatization agent can be selected. The selection can be cleared by selecting the <Clear> button (5). If more than one derivatization agent (DA) is selected, their corresponding scores are multiplied. The score corresponding to Criterion 6 is thus calculated as follows:

$$s_6 = \begin{cases} DA_1 \cdot DA_2 \cdot \ldots \cdot DA_n - 0.2 & \text{if } DA_1 \cdot DA_2 \cdot \ldots \cdot DA_n > 0.2 \\ 0.0 & \text{if } DA_1 \cdot DA_2 \cdot \ldots \cdot DA_n \leq 0.2 \end{cases} \tag{3}$$

5.7 Tab 7

The amount of waste in either g or mL is entered in the entry field. The entry is confirmed either by selecting the <Set> button to the right of the entry field, or by pressing the

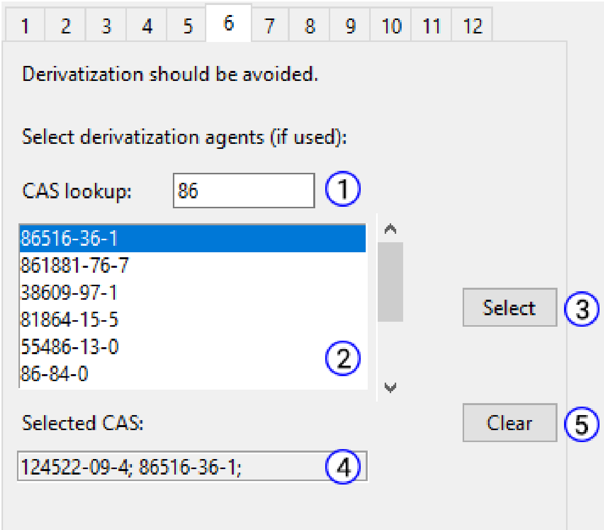

Figure S5: Tab 6 GUI elements.

<Enter> key. The entry can be either a float or an integer; other inputs, such as a text string, will prompt a *Value error* message box.

The score corresponding to Criterion 7 is calculated with the following equation:

$$s_7 = \begin{cases} 1.0 & \text{if } input < 0.1 \\ -0.134\ln(input) + 0.6946 & \text{if } 0.1 \leq input \leq 100 \\ 0.0 & \text{if } input > 150 \end{cases} \tag{4}$$

5.8 Tab 8

The number of analytes analyzed in one hour is determined based on the number of analytes analyzed in a single run and the sample throughput, both entered in the corresponding entry fields. The entries are confirmed by selecting the <Set> button below the entry fields, or by pressing the <Enter> key. The first entry field (the number of analytes determined in a single run, $n_{analytes}$) accepts only integers, and any other entry type will prompt a *Value error* message box. In the second entry field (sample throughput, *throughput*) the valid entry types are floats and integers; other entry types will prompt a *Value error* message box.

The score corresponding to Criterion 8 is calculated with the following equation:

$$s_8 = \begin{cases} 1.0 & \text{if } n_{analytes} \cdot throughput > 70.0 \\ 0.2429\ln(n_{analytes} \cdot throughput) - 0.0517 & \text{if } 1.0 \leq n_{analytes} \cdot throughput \leq 70.0 \\ 0.0 & \text{if } in_{analytes} \cdot throughput < 1.0 \end{cases} \tag{5}$$

5.9 Tab 9

There are two options for quantifying the use of energy in the evaluated analytical procedure. The first one is based on selecting the most energy-intensive technique used in the procedure, or its closest equivalent, from a drop-down option menu. The choices and the corresponding scores are as follows:

Selection	Score
None	1.0
FTIR	
Hot plate solvent evaporation (<10 min)	
Rotary evaporation	
Needle evaporation	
Ultrasound-assisted extraction	
SPE and SPME	
Microbiological assays	
Immunoassay	
Spectrofluorometry	
Titration	
UPLC	
UV-Vis Spectrometry	
Energy dispersive X-ray fluorescence	
Potentiometry	
Non-instrumental detection	
Hot plate solvent evaporation (10-150 min)	0.5
Accelerated solvent extraction	
Supercritical fluid extraction	
Microwave assisted extraction	
Flame atomic absorption spectrometry	
Electrothermal atomic absorption spectrometry	
GC	
ICP-MS	
ICP-OES	
LC	
Hot plate solvent evaporation (>150 min)	0.0
Soxhlet extraction	
NMR	
GC-MS	
LC-MS	
X-ray diffractometry	

Alternatively, the User can estimate the total energy use of the evaluated analytical procedure and enter it in the entry field in kWh. The entry is confirmed by selecting the <Set> button next to the entry field, or by pressing the <Enter> key. Valid entry types are floats and integers; other entry types will prompt a *Value error* message box.

If the second option is used, the score corresponding to Criterion 9 is calculated with the following equation:

$$s_9 = \begin{cases} 1.0 & \text{if } input < 0.1 \\ -0.7143 \cdot input + 1.0714 & \text{if } 0.1 \leq input \leq 1.5 \\ 0.0 & \text{if } input > 1.5 \end{cases} \tag{6}$$

5.10 Tab 10

The source of reagents and whether or not they are renewable is indicated by choosing an option from a drop-down option menu. The choices and their corresponding scores are as follows:

Reagents	Score
No reagents	1.0
All reagents are bio-based	1.0
Some reagents are bio-based	0.5
None of the reagents are from bio-based sources	0.0

5.11 Tab 11

The User indicates whether the evaluated analytical procedure involves the use of toxic reagents or solvents through the drop-down option menu. If the 'No' option is selected, the default score of 1.0 is assigned to the Criterion. Selecting the 'Yes' option in the drop-down menu incurs a 0.2 penalty and invokes an entry field in which the User can enter the amount of toxic reagents or solvents in either g or mL. The entry is confirmed by selecting the <Set> button below the entry field, or by pressing the <Enter> key. Valid entry types are floats and integers; other entry types will prompt a *Value error* message box.

If the procedure involves the use of toxic reagents or solvents, the score corresponding to Criterion 11 is calculated with the following equation:

$$s_{11} = \begin{cases} 0.8 & \text{if } input < 0.1 \\ -0.129\ln(input) + 0.5012 & \text{if } 0.1 \leq input \leq 50.0 \\ 0.0 & \text{if } input > 50.0 \end{cases} \tag{7}$$

5.12 Tab 12

Threats to the operator's safety associated with performing the evaluated analytical procedure are indicated by selecting the corresponding boxes. The possible threats are:

- toxic to aquatic life;
- bioacumulative;
- persistent;
- highly flammable;
- highly oxidizable;
- explosive;
- corrosive.

After all the relevant threats are chosen, the entry is confirmed by selecting <Set> button. The score corresponding to Criterion 11 is determined based on the number of selected threats:

$$s_{12} = \begin{cases} 1.0 & \text{if } \sum threats = 0 \\ 0.8 & \text{if } \sum threats = 1 \\ 0.6 & \text{if } \sum threats = 2 \\ 0.4 & \text{if } \sum threats = 3 \\ 0.2 & \text{if } \sum threats = 4 \\ 0.0 & \text{if } \sum threats \geq 5 \end{cases} \tag{8}$$

6 Bibliography

[1] Agnieszka Gałuszka, Zdzisław Migaszewski, and Jacek Namieśnik. The 12 principles of green analytical chemistry and the SIGNIFICANCE mnemonic of green analytical practices. *TrAC Trends in Analytical Chemistry*, 50:78–84, oct 2013.

[2] Python Software Foundation. Tkinter - Python interface to Tcl/Tk - Python 3.7.7 documentation. `https://docs.python.org/3.7/library/tkinter.html`. Accessed: 2020-04-02.

[3] PyPDF - FPDF for Python - documentation. `https://pyfpdf.readthedocs.io/`. Accessed: 2020-04-05.

[4] Marek Tobiszewski, Jacek Namieśnik, and Francisco Pena-Pereira. A derivatisation agent selection guide. *Green Chemistry*, 19(24):5911–5922, 2017.