



Greenness Assessment Metrics (AGREE, GAPI, and Eco-Scale) as Multi-Criteria Tools in Analytical Method Engineering

¹Anselm Oluwaseun Hannah, ^{2*}Jaja Wisdom

¹ Department of Chemical Sciences, College of Science and Information Technology,
Tai Solarin Federal University of Education, Ijagun, Ijebu-Ode, Ogun State, Nigeria
anselmoh@tasued.edu.ng

²Department of Mechanical Engineering, Faculty of Engineering, University of Benin, Benin City, Nigeria
wisdom.jaja@uniben.edu

Corresponding Author's: Jaja Wisdom; wisdom.jaja@uniben.edu

Manuscript History

Received: 25/12/2024

Revised: 28/03/2025

Accepted: 22/04/2025

Published: 28/04/2025

<https://doi.org/10.5281/zenodo.21463491>

[1/zenodo.21463491](https://doi.org/10.5281/zenodo.21463491)

Abstract: Green analytical chemistry (GAC) has turned the vision of a “green” analytical laboratory into a quantitative engineering problem, focusing on the development of analytical methods that reduce the use of hazardous reagents, energy consumption, waste generation and risk to the operator, while ensuring analytical performance. There are multiple greenness assessment metrics, but their functions as decision support tools are not sufficiently researched. The three greenness assessment metrics widely used, namely Analytical Eco-Scale, Green Analytical Procedure Index (GAPI) and Analytical GREENness (AGREE), are critically discussed and analysed in the context of Multi-Criteria Decision Analysis (MCDA). Literature was consulted to compare the conceptual basis, metrics, weighting methods, aggregation methods and applications of relevant greenness assessment tools and the analytical method evaluation literature. This review paper looked into how greenness assessment framework can complement the formal MCDA methods like the Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS), the Analytic Hierarchy Process (AHP) and the Preference Ranking Organisation Method for Enrichment Evaluation (PROMETHEE) for the transparent and reproducible selection of an analytical method. The results indicate that the variations in the measures are not due to flaws in the methodology, but to the different criteria and weights assigned to them and to the different ways they are aggregated. Future research should aim to embed greenness metrics into analytical method engineering and reporting frameworks using artificial intelligence and advanced decision support systems for wider use in analytical laboratories.

Keywords: Green Analytical Chemistry, Greenness Metrics, AGREE, GAPI, Analytical Eco-Scale, Multi-Criteria Decision Analysis, Method Selection, Sustainability Assessment

INTRODUCTION

Analytical chemistry is in a paradoxical place in the sustainability movement. Though measurement is the means by which environmental contamination is detected, regulated, and remediated, the process of measurement consumes a volume of solutes, reagents and energy and disperses contaminants of its own. The concept of this tension led to the development of green analytical chemistry, which encompasses the application of the ten principles of green chemistry advocated by Anastas and Warner (1998) to the functional aspects of sampling, sample preparation, separation and detection. The early formulations focused on solvent reduction and the use of non-toxic chemical components (Anastas, 1999; Armenta et al., 2008; Keith et al., 2007), the field has now developed into a systematic approach with well-defined criteria, specialised journals, a variety of test tools and devices, and a large body of experience. The important next conceptual leap was the codification of twelve principles of green analytical chemistry, and their formation in the SIGNIFICANCE mnemonic, by Gałuszka et al. (2013).

These principles apply to translating the abstract sustainability goals to action, and favour "direct" techniques, miniaturisation, automation, multi-analyte methods, avoiding derivatisation, minimisation of energy and waste, the use of safer or renewable reagents, and improvements to operator safety. The principles are not, however, evaluative. They do not answer the engineering question, which always comes up when developing, adapting, or choosing a method, which is: "How green is this method for which candidate? And is this candidate capable of being green? Is this candidate A green over this candidate B?"

By definition, that question is a multi-criteria problem. A procedure is green or not green from a single perspective, but rather, from multiple aspects at once, including reagent hazard, volume of solvent, energy usage, number of steps, rate of sample throughput, miniaturisation of the process, automation, and waste. The criteria are diverse and expressed in incommensurable units, and often they are contradictory; therefore, an improvement in one needs to be compensated by a compromise in another dimension. The smaller the amount of solvent used in microextraction, for example, the larger the amount of energy required and the more complex the instrument. Multi-criteria decision analysis is the discipline which formally deals with such trade-offs, and the main thesis of this review is that greenness assessment metrics, in fact, represent multi-criteria decision tools that are compact and specifically oriented within a certain field (Tobiszewski & Orlowski, 2015; Nowak *et al.*, 2020). The three are widely adopted and commonly used in the methods sections of analytical papers. The Analytical Eco-Scale of Gałuszka *et al.* (2012) sums up the negative records in deviations from an ideal green analysis as penalty points deducted from the ideal score of one hundred. The entire procedure is described by the Green Analytical Procedure Index Plotka-Wasyłka 2018, which is divided into fifteen fields arranged in five pentagrams, each in a different colour (green, yellow, red). The Analytical GREENness metric of Peña-Pereira *et al.* (2020) quantifies the degree of alignment between the twelve principles in a unit interval and sums them together (with user-chosen weight factors) into a single pictogram, in the form of a clock, with a central score between zero and one, respectively. These differences are responsible for complementary strengths as well as for the occasionally noted differences in performance when the same procedure is executed with one of the tools; these differences represent different ways of normalising and weighting the scores and different ways of aggregating the results, with each of these steps arising from the underlying decision-analytic pipeline depicted in Fig. 1.

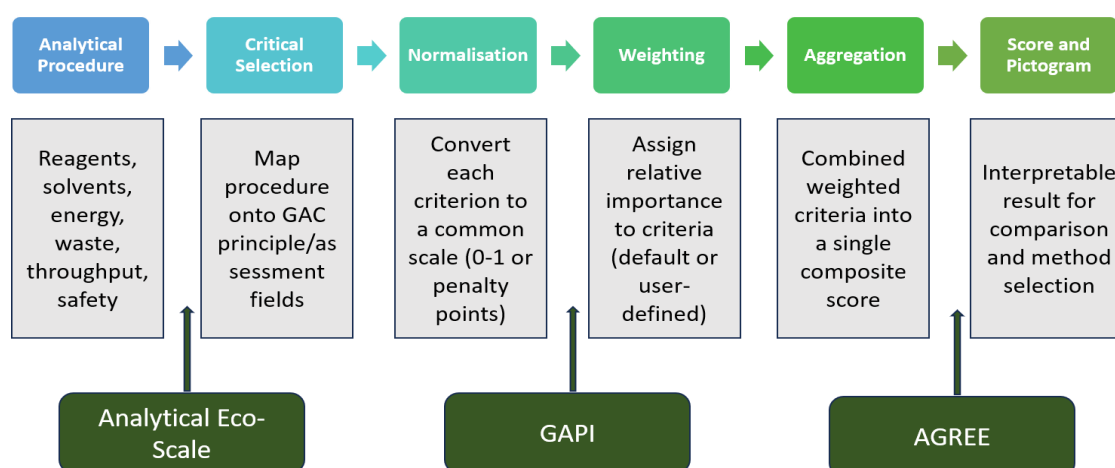


Fig. 1 Greenness assessment metrics as a multi-criteria decision workflow. Each flagship tool implements the same pipeline of criteria selection, normalisation, weighting, and aggregation, but differs in how the steps are resolved and in the form of the final output.

There are four objectives in this review. First, to explain the conceptual principles that link the 12 principles of green analytical chemistry with the formal framework of multi-criteria decision analysis. Second, an engineering detailed description of the Eco-Scale, GAPI and AGREE, starting from its mathematics and its constituent equations, its scores, its mathematics-governed logic, its characteristic outputs and its principal derivatives. Third, they are compared over a range of attributes that are relevant when a metric is being used to inform a method's development or selection, and they are placed in an overall context of a wider, far more diverse, growing suite of assessments. Fourth, to explore the relationship of these screening metrics with complete multi-criteria decision algorithms, considering analytical performance and make suggestions for more transparent and reproducible reporting of greenness. The review will be useful for analytical chemists who have become involved in method engineering, whose greenness is no longer a luxury; it's part of the design specification to be reported.

2.1 Conceptual Foundations of Greenness Assessment

The concept of green chemistry was developed in 1998 as twelve principles of green chemistry dealing with synthesis, the concept of atom economy, waste prevention and the design of safer chemicals, use of renewable chemicals as feedstocks (Anastas & Warner, 1998). They are adapted to preparative chemistry such that the product is a target molecule, and measures like atom economy, the environmental factor, and reaction mass efficiency evaluate the conversion of mass from the reagents to product (Tobiszewski, Marc, *et al.*, 2015). Structurally, analytical chemistry is different. Unlike a chemical reaction, it yields the product: information, not material; its quality is measured according to its figures of merit: sensitivity, selectivity, accuracy and precision; the reagents are not involved in a product, but rather in measurement. Initial comparisons of mass-based synthetic metrics to analysis, therefore, fall short of a translation of the approach into analysis as was noted early in the development of the field (Gałuszka *et al.*, 2012).

The twelve principles of green analytical chemistry, summarised in Table-1, were developed to fill this gap. They carry the principle of green chemistry, but apply it to the analytical process from sampling to sample preparation to analytical determination. The principles are accompanied by the SIGNIFICANCE mnemonic, which in memory is laid out as a sequence and serves as an indication that it is meant to be used and considered during method design, as a checklist (Gałuszka *et al.*, 2013). An important point for the current discussion, the principles are a criterion set. The dimension described by each principle is one in which a method can become greener or less green; to make a complete assessment, both a method is characterised on every dimension, and the dimensions of its assessment are combined in an overarching judgment. This is the type of input needed by a multi-criteria decision procedure.

Table-1 The twelve principles of green analytical chemistry, their operational meaning for method engineering, and the assessment tool that most directly encodes each principle

Principle	Operational meaning for method engineering	Primary tool coverage
Direct analytical techniques	Avoid or minimise sample treatment by measuring the analyte in situ or with minimal preparation.	AGREE; GAPI
Minimal sample size and number	Reduce the mass or volume of the sample and the number of samples required per determination.	AGREE; Eco-Scale
In situ measurement	Measure the sampling location to avoid transport, storage, and preservation reagents.	GAPI; AGREE
Integration of operations	Combine analytical steps to reduce reagent use, manipulation, and energy demand.	GAPI; AGREE
Automation and miniaturisation	Prefer automated and miniaturised formats that lower reagent and waste volumes.	AGREE; BAGI
Avoidance of derivatisation	Eliminate derivatisation steps that introduce additional reagents and hazards.	AGREE; GAPI

Minimisation of waste	Reduce the volume of analytical waste and provide for its treatment	Eco-Scale; GAPI
Multi-analyte or multi-parameter methods	Determine several analytes or parameters in a single run to amortise reagents and energy.	AGREE; GAPI
Minimisation of energy	Select low-energy techniques and instrumentation per sample	Eco-Scale; AGREE
Renewable reagents	Prefer reagents obtained from renewable or bio-based sources	AGREE; ComplexGAPI
Elimination of toxic reagents	Replace or remove toxic and hazardous reagents and solvents	Eco-Scale; GAPI; AGREE
Operator safety	Increase the safety of the analyst by reducing exposure to hazards	AGREE; Eco-Scale

A multi-criteria decision problem consists of a set of alternatives, a set of criteria along which the alternatives are evaluated, a procedure for measuring the performances (results) of the alternatives on each criterion, a normalisation that makes it possible to compare heterogeneous criteria, a vector of weights indicating relative importance of the criteria, and an aggregation rule that aggregates the performance (or rank) of the alternatives along the weighted normalised criteria. Alternatives in the assessment of analytical methods are candidate procedures, and the criteria are greenness dimensions based on the 12 principles, the overall output being an overall greenness score or profile (Nowak *et al.*, 2020; Tobiszewski & Orlowski, 2015). The three metrics are similar primarily in how they address normalisation, weighting and aggregation, which are designed decisions that are appropriate for an engineering analysis. The general weighted aggregation is the most probable compact representation of general logic shared between multiple indicators, from which composite indicators are built. When a procedure is defined through performance values for a set of criteria, they are each put on a common scale and then combined as a weighted mean as expressed in Equation 1.

$$S = (\sum w_j \cdot n_j) / \sum w_j \quad (1)$$

The S is the composite score of the greenness, n_j is the normalised result of the procedure to criterion j , and w_j is the weight of criterion j . Each of the three metrics is a particular case of this expression. The Eco-Scale overcomes this notion of a 'sweet middle ground to human foibles' by permanently establishing the ideal (which is always 'good') at 100 and acknowledging each of the weighted deficiencies as a decrease in score. GAPI does not purposefully aggregate the numbers, but rather displays the Weighted Per-Criterion performances directly as a coloured profile, which is to be judged solely by the reader. In AGREE, the weighted mean is done explicitly on a unit interval, and is returned both as an aggregate and as per criterion components. Identifying these tools just as another family helps to understand how they are linked but not equal, and why there is no one 'right answer' when it comes to greenness. There are two implications that spring from this. The criterion sets vary from one tool to the next, so a procedure that has a high score for one is low for another because it asks a different type of question of a dimension that is being asked in the other context. Also, since the rules for aggregating the results vary, the same performance on each criterion could generate different sets of rankings of the alternatives. Improperly dealt with, these are not illnesses to cure, rather attributes to know about and encourage, frequently explained in the literary works in this way, with more than one metric being reported in addition to screening scores, and, when ranking is necessary, a formal decision algorithm over the screening scores (Fabjanowicz *et al.*, 2021; Sammut Bartolo *et al.*, 2024).

2.2 Formal Multi-Criteria Decision Analysis in Method Engineering

Before examining the individual metrics, it is useful to make explicit the formal decision algorithms that increasingly accompany them. Screening metrics such as AGREE and the Eco-Scale compress greenness into a single index, but method selection often requires that greenness be balanced against analytical performance and cost. When several criteria of genuinely different character must be reconciled, analysts have turned to established multi-criteria decision analysis methods, the most widely used of which are the Technique for Order of Preference by Similarity to Ideal Solution, the Analytic Hierarchy Process, the Preference Ranking Organisation Method for Enrichment Evaluations, Elimination and Choice Expressing the Reality, and Multi-Attribute Utility Theory (Bigus *et al.*, 2016; Tsakovski *et al.*, 2015). These methods provide a defensible ranking of alternatives and are summarised in Table 2.

The TOPSIS (Technique for Order of Preference by Similarity to Ideal Solution) has proven to be the algorithm most often used together with greenness metrics, due to its ease of implementation and simplicity of calculation. It assumes that the best alternative to a positive ideal solution, where all the values of the criteria are at their best, is the alternative that has the smallest geometric distance, that is, the least distance. On the other hand, the best alternative to a negative ideal solution is the one with the largest geometric distance, that is, the largest distance between the solutions, and it achieves the worst values for all the criteria (Fabjanowicz *et al.*, 2021). Then, it starts with the decision matrix of alternatives that is described with criteria and normalised through Equation 2, followed by weighting in Equation 3.

$$r_{ij} = a_{ij} / \sqrt{(\sum_i a_{ij}^2)} \quad (2)$$

$$v_{ij} = w_j \cdot r_{ij} \quad (3)$$

In these expressions, a_{ij} is the raw performance of alternative i on criterion j , r_{ij} is the vector-normalised value, and v_{ij} is the weighted normalised value. The positive and negative ideal solutions are then identified, and the separation of each alternative from each ideal is computed as a Euclidean distance, as in Equation 4.

$$S_i^+ = \sqrt{(\sum_j (v_{ij} - v_j^+)^2)} \quad ; \quad S_i^- = \sqrt{(\sum_j (v_{ij} - v_j^-)^2)} \quad (4)$$

Finally, the relative closeness of each alternative to the ideal solution, which serves as the ranking index, is obtained from Equation 5. Values approach unity for alternatives close to the positive ideal and approach zero for alternatives close to the negative ideal.

$$C_i = S_i^- / (S_i^+ + S_i^-) \quad (5)$$

The advantage of this scheme for method engineering, because the criteria don't have to be limited to greenness. Some criteria for a decision matrix might also include limits of detection, recovery, precision, analysis time, and cost, to make the matrix more representative of fitness for purpose as well as environmental impact (Fabjanowicz *et al.*, 2021; Tobiszewski & Orlowski, 2015). If weighting is done, it is then an explicitly and audibly expressed set of priorities, which can be assigned based on expert judgment or, in the case of the Analytic Hierarchy Process, by pairwise comparisons that have been structured. This integration is the link between screening for greenness and serious method consideration and is readdressed in Section 7 following the explanation of the individual metrics.

Table-2 Formal multi-criteria decision analysis methods are used to rank analytical procedures by combined greenness and performance

Method	Full name	Core principle	Typical use in method engineering	Reference
TOPSIS	Technique for Order of Preference by Similarity to Ideal Solution	Shortest distance to a positive ideal, farthest from a negative ideal	Ranking procedures by combined greenness and figures of merit	Fabjanowicz et al. (2021)
AHP	Analytic Hierarchy Process	Criterion weights derived from structured pairwise comparisons	Eliciting defensible weights for greenness and performance	Tsakovski, et al. (2015)
PROMETHEE	Preference Ranking Organisation Method for Enrichment Evaluations	Pairwise outranking through preference functions	Outranking solvents, reagents, and candidate procedures	Tsakovski, et al. (2015)
ELECTRE	Elimination and Choice Expressing the Reality	Outranking via concordance and discordance indices	Screening out dominated or unacceptable alternatives	Bigus et al. (2016)
MAUT	Multi-Attribute Utility Theory	Aggregation of single-attribute utility functions	Combining greenness and performance utilities	Bigus et al. (2016)
CA with TOPSIS	Cluster analysis coupled with TOPSIS	Group similar alternatives, then rank within each group	Solvent and procedure grouping and selection	Tobiszewski, Tsakovski, et al. (2015)

2.3 The Analytical Eco-Scale

The Analytical Eco-Scale, presented by [Gałuszka et al. \(2012\)](#), is the earliest of the three flagship tools, and it is the one which is most explicitly arithmetic. Its conceptual model is an ideal green analysis, which would carry no environmental or safety burden and hence would deserve a perfect score of one hundred. Any real procedure that deviates from this ideal is calculated to be a penalty point (points remaining after some are deducted from the maximum, according to Equation 6). The outcome is just one number on a well-known scale from zero to a hundred, with increased numbers indicating greener procedures.

$$\text{Eco-Scale score} = 100 - \sum PP_i \quad (6)$$

In Equation 6, PP_i denotes the penalty points accrued by the procedure for the i -th source of impact. The penalty points are grouped in four categories: the level of reagents, the danger of reagents, the level of energy used by the instruments, the danger of producing an occupational hazard due to vapours or gases and the quantity and treatment of waste. Penalties are based on amounts and level of hazard, with a penalty of an amount factor for a reagent times the hazard factor; a heavy weight of heavily hazardous reagent gets the highest penalty; a small amount of an unlabelled reagent gets a zero penalty. In Table 3, the full penalty reference is provided. Scores above 75 represent an excellent green analysis; scores from 50 to 75 represent an acceptable green analysis; and scores under 50 represent an inadequate green analysis ([Gałuszka et al., 2012](#)).

Table-3 Penalty-point reference for the Analytical Eco-Scale, with scoring bands for interpretation

Category	Condition or sub-criterion	Penalty points	Implication
Reagents, amount	Less than 10 mL or 10 g per sample	1	Amount factor
Reagents, amount	Between 10 and 100 mL or g per sample	2	Amount factor
Reagents, amount	More than 100 mL or g per sample	3	Amount factor
Reagents, hazard	Each less severe hazard, signal word Warning	1	Multiply by the amount
Reagents, hazard	Each more severe hazard, signal word Danger	2	Multiply by the amount
Reagents, hazard	Reagent without a hazard pictogram	0	No penalty
Instruments, energy	Up to 0.1 kWh per sample	0	Low energy
Instruments, energy	Up to 1.5 kWh per sample	1	Moderate energy
Instruments, energy	More than 1.5 kWh per sample	2	High energy
Occupational hazard	Hermetic process, no emission of vapours or gases	0	Contained
Occupational hazard	Emission of vapours or gases to the atmosphere	3	Exposure risk

Waste, volume	Less than 1 mL or g	1	Small waste
Waste, volume	Between 1 and 10 mL or g	3	Moderate waste
Waste, volume	More than 10 mL or g	5	Large waste
Waste, treatment	Recycling provided	Reduces	Subtract from waste
Waste, treatment	Degradation or passivation is provided	Reduces	Subtract from waste
Waste, treatment	No treatment provided	Full	Retain waste penalty
Scoring	Starting score for an ideal green analysis	100	Maximum
Interpretation	Excellent green analysis	> 75	Top band
Interpretation	Acceptable green analysis	50 to 75	Middle band
Interpretation	Inadequate green analysis	< 50	Lower band

The Eco-Scale is fast, doesn't require any unique software, and generates a single ratio-scale number that can be easily compared between different procedures as well as be easily integrated into a larger decision-making matrix. The "good and bad" features make it easy to understand and can be easily monitored; it is simple to correlate the four categories with the impacts that are most relevant for a bench analyst: the reagents, the energy, the safety and the waste (Gałuszka *et al.*, 2012; Marc *et al.*, 2015). However, all impacts are condensed into one score, leading to the loss of information about the strengths and weaknesses of a procedure; two methods may have the same score but very different profiles. The penalties are grossly graded and somewhat arbitrary, especially when it comes to hazard severity, and although it rewards miniaturisation, automation and in situ measurement, it isn't specifically making a point of it. The Eco-Scale does not include analytical performance, meaning that a more eco-friendly, but less sensitive method would not face a penalty for poor analytical performance. These restrictions inspired the pictographic and unit-interval tools that did come after them, as well as the subsequent total-scoring versions of GAPI that combined the object-oriented and pictographic views of profiling (Mansour *et al.*, 2024).

2.4 The Green Analytical Procedure Index and Its Extensions

The new Green Analytical Procedure Index introduced by Plotka-Wasyłka (2018) bridges this main drawback, as it keeps and presents the information on each criterion. GAPI is based on the colour logic of the National Environmental Methods Index (NEMINDEX), which defines methods with a four-quadrant symbol, but enriches it to a wider profile (Keith *et al.*, 2007). Each pentagram in the GAPI pictogram represents one point in the analytical workflow process: sample collection (preservation), sample preparation, reagents and solvents, instrumentation, and the characteristics of the end of the method. Inside these stages, 15 fields are assessed, and each field is given a colour according to their environment impact: Green, Yellow, and Red as shown in Fig. 2.

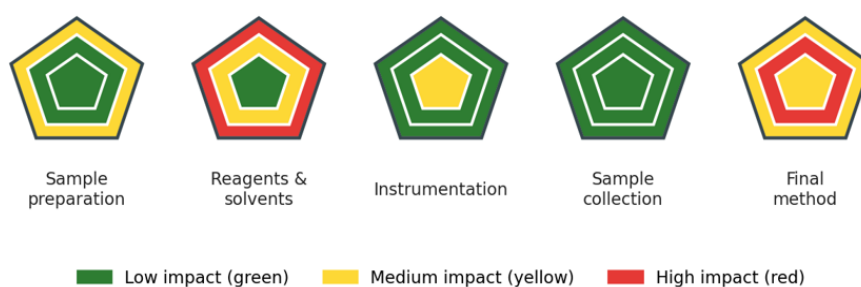


Fig. 2 The Green Analytical Procedure Index pictogram. Five pentagrams represent successive stages of the analytical workflow, and fifteen fields are coloured green, yellow, or red according to the environmental impact of each aspect of the procedure.

The main design choice of GAPI is that it does not require numerical aggregation. The tool not only ranks the performances in a score, but it also always shows the performances on a per-criterion basis, which can then be synthesised by the analyst. When dealing with Section 2 in its own "language," GAPI

asks you to perform the criteria-selection and normalisation steps of the decision pipeline without the weighting or aggregation step. This is its greatest strength concurrently and its main drawback. Though it can provide insights into the problematic steps of a procedure, identified at a glance, which ones to improve the profile cannot be easily compared directly and must undergo an implicit and unstated aggregation by the reader of the profile (Plotka-Wasyłka, 2018; Sajid *et al.*, 2022). There are three extensions of the original tool. The particular approach of the Plotka-Wasyłka and Wojnowski's (2021) complex GAPI introduces a hexagonal field into the pictogram to include the processes preceding the analytical procedure itself, for instance, the synthesis or purification of reagents, materials and instruments, thus broadening the system boundary to a fuller life-cycle perspective. Mansour *et al.*, (2024) developed a new tool, the Modified GAPI (MoGAPI), which adds the total score that the original did not include, and is composed of a pictographic profile and an Eco-Scale-style numerical aggregation, along with software providing the analyst with a profile and a single comparable value.

ComplexMoGAPI then combines the complementary fields of pre-analysis (ComplexGAPI) with the overall scoring of MoGAPI (Mansour, Omer, & Plotka-Wasyłka, 2024). The following are examples of the general direction of development in this area, namely, pictographic and arithmetic logics converging so that value and profile are available for tools and can be aggregated across scores. GAPI is finally interpretable to anyone. It shows the environmental characteristics of the procedure in steps in colours and conveys important information quickly, and is suitable to guide the method improvement as it identifies the disturbing step. It describes the entire process from collection to determination, and ComplexGAPI adds upstream processes to its scope of coverage (Plotka-Wasyłka & Wojnowski, 2021). Meanwhile, the original GAPI does not provide a score, making it difficult to rank and/or use in a simplified decision matrix that would ensure one criterion over another. The 3-level scale is blocky, and the assignments of fields to colours involve the assignment of thresholds that differ depending on the assessor making the colour assignment, thus being less reproducible. This tool, as well as the Eco-Scale, is not used to assess analytical performance. The MoGAPI and ComplexMoGAPI derivatives reduce the subjectivity in scoring, but add the subjectivity inherent in any penalty assignment (Sajid & Plotka-Wasyłka, 2022; Mansour *et al.*, 2024).

2.6 The Analytical GREENness Metric and AGREEprep

The Analytical GREENness metric (Peña-Pereira *et al.*, 2020) is the most straightforward realisation of the multi-criteria decision pipeline and the most connected to the twelve principles. The principles of green analytical chemistry are the criteria used by AGREE, and each principle is transformed into a so-called "sub-score" on the unit interval between 0 and 1, and all 12 "subscores" are combined to a "big score," also on the unit interval between 0 and 1, through a weighted mean. The outcome is presented as a clock with 12 'segments', one for each of the principles, these segments' colour and extent representing how much the sub-score aligned to the segment is from zero, and the central segment's colour and extent reflecting the main result (as deep red being close to zero, and deep green being close to one). The aggregation is expressed in Equation 7. Each principle k receives a sub-score x_k on the unit interval, and a weight w_k , and the overall AGREE score is the weighted arithmetic mean of the twelve sub-scores.

$$AGREE = (\sum_{k=1}^{12} w_k \cdot x_k) / (\sum_{k=1}^{12} w_k) \quad (7)$$

There are two features distinguishing AGREE in the family. The first one is that the weights are adjustable by the user. All twelve principles are equally weighted by default, but if the analyst agrees that some are more applicable to a certain context, they are not hidden but rather given greater weight in the aggregation step, making it explicit and thus auditable. This weighting of freedom that formal decision algorithms offer is incorporated into the metric itself (Peña-Pereira *et al.*, 2020). Secondly, the

index in AGREE consists of both an aggregate score and individual scores for each of the criteria; thus, it makes use of the comparative attractiveness of the one-number reporting provided by the Eco-Scale together with the diagnostic value of each respective criterion provided by the GAPI. Only among the three tools is AGREE able to do all four pipeline steps in a decision-analytic way and present the weights to the user.

A special derivative is concerned with the least environmentally friendly step in the procedure, the sample preparation. The AGREEprep is defined by [Wojnowski et al. \(2022\)](#) as a method that evaluates sample preparation performance based on the criteria mentioned by [Lopez-Lorente et al. \(2022\)](#), among the ten principles of green sample preparation, including criteria from: choice and amount of solvents, materials, and reagents; minimisation of waste; minimisation of energy consumption; sample throughput; and operator safety. It generates a unit-interval score and a pictogram; is subject to weighting; and allows open software to be used for calculations, like AGREE. Sample preparation often accounts for most solvent and energy usage, and is therefore treated in detail separately from the rest of the procedure with AGREEprep, which is complementary, not competing with, any whole procedure tool. AGREE is seen as comprehensive because it looks at all 12 principles and is the only flagship tool to make the aggregation visible to the user, so they can see that weighting has taken place. Both an overall score and per principle components are reported for comparison and diagnosis. The free software reduces the cost of a consistent application, and the unit-interval output can be easily incorporated into a decision matrix ([Peña-Pereira et al., 2020](#)).

However, the difficulty in unambiguously scoring several of the principles in the information commonly reported may result in some level of judgment to be exercised when converting a procedure into the twelve sub-scores, which can diminish the repeatability of scoring between scorers. While user-defined weighting is a good thing, it can make two scores based on the same data not exactly comparable if they are computed with different weights; a warning metric could be the AGREE score. Like the other tools, in the standard form, AGREE takes an assessment of greenness rather than analytical performance; this is where the practicality and performance indices come in handy, recently ([Manousi et al., 2023](#); [Nowak et al., 2025](#)). Eco-Scale, GAPI, and AGREE cannot be mutually substituted. They have separate but similar sets of criteria, handle normalisation and aggregation differently and produce various kinds of output. Table 4 outlines a set of structured comparisons about attributes which are important when one decides to cluster a metric to become the basis for method engineering, including the form of the output, the granularity of feedback, flexibility of weights, software support, and flexibility for reproduction and sensitivity to analytical performance. A difference can be seen in the complementary nature of the tools in Figure 4, which shows the same comparison as a profile along the seven dimensions of decision.

Table-4 Comparative feature matrix of the Analytical Eco-Scale, GAPI, and AGREE across attributes relevant to analytical method engineering

Attribute	Analytical Eco-Scale	GAPI	AGREE
Year introduced	2012	2018	2020
Primary reference	Galuszka et al. (2012)	Plotka-Wasyłka (2018)	Pena-Pereira et al. (2020)
Conceptual basis	Penalty deviation from ideal	Stage and field profiling	Twelve principles on the unit interval
Output form	Single numeric score	Coloured pictogram, no total	Pictogram with central score
Score range	0 to 100	Not a numeric score	0 to 1
Greener direction	Higher score	More green fields	Score closer to 1
Criteria assessed	Reagents, energy, hazard, waste	15 fields in 5 stages	12 GAC principles
Normalisation	Penalty points	Three-level colour scale	Unit-interval sub-scores
Aggregation rule	Subtractive sum	None, left to the reader	Weighted arithmetic mean
Weighting capability	Implicit and fixed	Not applicable	Explicit and user-defined
Whole-procedure coverage	Partial, impact-focused	Full collection of the method	Full, principle-based
Pre-analysis coverage	No	Via ComplexGAPI	Indirect

Sample preparation detail	Aggregated	One stage	Via AGREEprep
Per-criterion feedback	Low	High	High
Visual interpretability	Low	Very high	High
Quantitative comparability	High	Low without total	High
Software availability	Not required	Free pictogram tool	Free calculator
Reproducibility risk	Threshold subjectivity	Colour assignment variance	Sub-score judgement
Sensitivity to figures of merit	None	None	None in standard form
Ease and speed of use	Fast	Moderate	Moderate
Suitability as a decision criterion	High	Low without total	High
Principal derivatives	MoGAPI fused logic	ComplexGAPI; MoGAPI; ComplexMoGAPI	AGREEprep
Typical application domains	Environmental, pharmaceutical	Environmental, food, and pharmaceutical	All analytical domains
Main limitation	Loss of profile information	No native total score	Weighting and sub-score subjectivity

The comparison will help to better understand the engineering reading. If one number can be used, such as for an Eco-Scale rating used as one measure in a ranking matrix, or as one measure in an AGREE rating, then either a single comparable number is suitable, or an AGREE is used (with an AGREE preferred in cases where the value of transparency of the weighting of the rating and per-principle diagnosis is valued). When the intent is to learn about and enhance a procedure, the most comprehensive profile is supplied by GAPI and supplemental extensions. When sample preparation accounts for much of the environmental burden, AGREEprep provides this detail. The practicality and performance indices, already discussed in Section 9, are able to address the gap, namely the analytical performance, which none of the tools, in its standard form, is able to evaluate and which ends up being closed by combining greenness scores with figures of merit implemented by formal decision algorithms.

3. 1 Integration of Green Metrics into Analytical Method Engineering

The utility of greenness metrics can only be realised when they become a part of a method development and selection process instead of being added as a rather decorative pictogram to a completed manuscript. Its decision analytic interpretation of the metrics points towards a definite sequence. While the development progresses, the per-criterion feedback of GAPI and AGREE can serve as a diagnostic tool to pinpoint the location and the principle of where a change would improve the procedure, and help the analyst decide on the specific changes in the procedure to try, such as the replacement of a hazardous solvent, miniaturisation of an extraction, or elimination of a derivatisation step. The latest modification can be re-scored, and those limits may be considered an iterative design goal - similarly, limits of detection and recovery are optimised. This is designed for greenness and relies on measurements that reveal the strengths and weaknesses of a method, a diagnostic function distinction between AGREE and GAPI from the single-number environmental score of the Eco-Scale (Plotka Wasylka, 2018; Peña-Pereira *et al.*, 2020). On selection among finished candidates (often the case when the task is not the improvement of an existing candidate), a single greenness index is generally less useful as the analyst has to account for other parameters, such as sensitivity, selectivity, accuracy, precision, throughput and cost. Here, the screening metrics are inserted in a formal algorithm used to make a decision.

A decision matrix is built for each of the candidate procedures (rows), and the columns allude to a greenness score from AGREE or the eco scale, and the figure of merit and cost data. The matrix is then normalised and weighted according to the environmental impact and fitness for purpose, and then an algorithm, specifically the Technique for Order of Preference by Similarity to Ideal Solution, returns a ranking (Tobiszewski & Orłowski, 2015; Fabjanowicz *et al.*, 2021). The weighting vector clarifies and

puts on display the analyst's priorities; it is better to put the priorities on the screen than hide them in one figure. The wide-ranging use of this has been confirmed in the literature. In Table-5, some representative case studies in which greenness metrics have been used (or together with decision algorithms) in method engineering are summarised, both in the environmental and pharmaceutical domains and in the field of food analysis. The examples span from the use of the least harmful procedure for the phthalate determination in baby diapers to comparing four metrics for the determination of cannabinoids in oils and demonstrate both diagnostic and selection uses.

Table-5 Representative applications of greenness metrics and multi-criteria decision analysis in analytical method engineering

Application area	Analytical context	Metric or method applied	Decision insight	Reference
Phthalate determination	Disposable baby diapers, nine candidate procedures	GAC metrics combined with TOPSIS ranking	Selected the lowest-impact procedure that retained adequate figures of merit	Fabjanowicz et al. (2021)
Aldrin determination	Water, competing analytical procedures	Multi-criteria decision analysis ranking	Ranked procedures by combined environmental and performance criteria	Tobiszewski and Orłowski (2015)
Chlorophenol determination	Water, dispersive liquid-liquid microextraction	MCDA for solvent type optimisation	Identified a greener solvent without loss of extraction efficiency	Bigus et al. (2016)
Cannabinoid determination	Oils	NEMI, Eco-Scale, AGREE, and GAPI compared	Cross-tool comparison exposed discordance and informed reporting	Sammut Bartolo et al. (2024)
Solvent selection	151 solvents, including bio-based options	Chemometrics coupled with TOPSIS	Grouped and ranked solvents by physicochemical and hazard criteria	Tobiszewski, Tsakovski et al. (2015)
Derivatisation agent selection	Liquid, gas, and chiral chromatography	Multi-criteria decision analysis	Produced confidence-graded rankings of safer derivatisation agents	Tobiszewski et al. (2017)
Sample preparation assessment	Microextraction and digestion methods	AGREEprep ten-criterion score	Quantified the environmental burden of the preparation stage	Wojnowski et al. (2022)
Method greenness scoring	Pharmaceutical separations	MoGAPI total score with pictogram	Delivered a comparable score alongside a stage profile	Mansour, Plotka-Wasyłka, and Locatelli (2024)

The three flagship tools are part of an expanding suite of assessment instruments, as summarised in Table-6 (The inventory is ordered by year of publication. Greenness metrics increasingly appear alongside practicality and performance indices, reflecting the field's movement from greenness alone toward balanced, multi-axis assessment of analytical methods. There are two trends throughout this history. The first is methodological convergence because later tools combine the pictographic profile with the arithmetic aggregation of the Eco-Scale and AGREE, like in ComplexMoGAPI and MoGAPI, with the result that analysts would increasingly get both a pictographic profile and a score from the same instrument ([Mansour et al., 2024](#)). The second trend involves removing concepts from their status as just green. All three "flagship metrics" have been criticised for incentivising environmental 'virtue' without consideration for analytical performance; that is, they can have a very high score on greenness without being sensitive or sufficiently strong for their intended purposes. Responding to the community, complementary axes of assessment have been developed. Economic, practical, and analytical attributes are combined with greenness into a composite profile of red, green, and blue additive colour model, developed by the authors [Nowak and Koscielniak \(2019\)](#), with the White Analytical Chemistry framework that elaborates it ([Nowak et al., 2021](#)). The ease and applicability of a method to real-world systems can be measured by using the Blue Applicability Grade Index proposed by [Manousi et al., \(2023\)](#), which takes into consideration some of the key criteria: cost, sample

throughput, and specialised personnel needed for the method. The remaining gap is filled by the Red Analytical Performance Index of Nowak *et al.* (2025) that provides a score for analytical performance itself. These tools, combined with the greenness indicators, can facilitate a balanced assessment in which environmental issues and impacts are not assessed alone, but together with the validity and the practicability.

Table-6 Chronological inventory of analytical assessment frameworks and metrics relevant to greenness evaluation and method selection

Tool or framework	Output type	Core focus	Reference
EcoScale, organic synthesis	Score out of 100	Penalty-based selection of organic preparations	Van Aken <i>et al.</i> (2006)
National Environmental Methods Index	Four-field pictogram	First widely used profiling symbol for methods	Keith <i>et al.</i> (2007)
Analytical Eco-Scale	Score out of 100	Penalty-based greenness of analytical procedures	Galuszka <i>et al.</i> (2012)
Twelve principles and SIGNIFICANCE	Principle framework	Codified criteria for green analytical chemistry	Galuszka <i>et al.</i> (2013)
Green chemistry metrics review	Review and taxonomy	Mapped metrics across green and analytical chemistry	Tobiszewski <i>et al.</i> (2015)
MCDA ranking of procedures	Procedure ranking	Demonstrated decision analysis for method choice	Tobiszewski and Orlowski (2015)
Solvent selection guide	Clustered ranking	Combined chemometrics and decision analysis for solvents	Tsakovski <i>et al.</i> (2015)
Derivatisation agent selection guide	Confidence-graded ranking	Decision analysis for safer derivatisation agents	Tobiszewski <i>et al.</i> (2017)
Green Analytical Procedure Index	Five-pentagram pictogram	Stage and field profiling of the whole procedure	Plotka-Wasyłka (2018)
RGB additive colour model	Composite colour	Combined greenness with other quality attributes	Nowak and Koscielniak (2019)
HEXAGON tool	Hexagonal pictogram with penalties	Figures of merit, safety, residues, footprint, and cost	Ballester-Caudet <i>et al.</i> (2019)
Analytical Method Greenness Score	Numeric score	Greenness of separation methods in pharmaceuticals	Hicks <i>et al.</i> (2019)
Overview of three multicriteria approaches	Comparative analysis	Compared global multicriteria assessment tools	Nowak <i>et al.</i> (2020)
Analytical GREEnness metric	Clock pictogram with score	Weighted aggregation of twelve GAC principles	Pena-Pereira <i>et al.</i> (2020)
ComplexGAPI	Pictogram with a hexagon	Added pre-analysis processes to GAPI	Plotka-Wasyłka and Wojnowski (2021)
White Analytical Chemistry	Composite RGB profile	Reconciled greenness with functionality and economy	Nowak <i>et al.</i> (2021)
MCDA grouping with TOPSIS	Ranking and grouping	Grouped and ranked procedures for phthalates	Fabjanowicz <i>et al.</i> (2021)
Green analytical chemistry metrics review	Review	Surveyed and compared the metric landscape	Sajid and Plotka-Wasyłka (2022)
AGREEprep	Pictogram with score	Ten-criterion greenness of sample preparation	Wojnowski <i>et al.</i> (2022)
Ten principles of green sample preparation	Principle framework	Criteria for greening the preparation stage	Lopez-Lorente <i>et al.</i> (2022)
Blue Applicability Grade Index	Pictogram with score	Practicality and applicability of methods	Manousi <i>et al.</i> (2023)
Modified GAPI, MoGAPI	Pictogram with total score	Added Eco-Scale-style scoring to GAPI	Mansour <i>et al.</i> (2024)
ComplexMoGAPI	Pictogram with total score	Unified pre-analysis fields and total scoring	Mansour <i>et al.</i> (2024)
Red Analytical Performance Index	Pictogram with score	Direct assessment of analytical performance	Nowak <i>et al.</i> (2025)

CONTRIBUTION TO KNOWLEDGE

This review reframes the Analytical Eco-Scale, GAPI, and AGREE not as standalone scores but as multicriteria decision instruments, decomposing each into criterion selection, normalisation, weighting, and aggregation. It links them to formal algorithms such as TOPSIS, AHP, and PROMETHEE, enabling transparent, reproducible, performance-aware greenness reporting for analytical method engineering.

CONCLUSION

Application of greenness assessment metrics has become commonplace in the development of analytical methods, and in this review, we have suggested that these are best thought of as small domain-specific multi-criteria decision-making tools. The Analytical Eco-Scale is a multi-punctuation score which sums value deviations from the ideal as penalty pursuits and provides a single comparable number. The Green Analytical Procedure Index highlights the whole procedure with red, yellow, green and white columns and is excellent for diagnosis. The Analytical GREENness metric explicitly aggregates on 12 principles and aggregates them on the unit interval in a weighted manner by the user for the purpose of making comparisons, and is also unique in combining the comparison with the diagnosis per principle. Both follow the same pipeline of criterion selection, normalisation of the selected criteria, and aggregation and weighting of the subresult each, and the differences in how they perform these steps give rise to both similarities and disparities in their results. The meaning of this for the analyst who works with patients is clear. Greenness should be designed iteratively and reported as a design specification with the aid of diagnostic feedback (AGREE and GAPI). When methods need to be ranked, greenness should be added, in a clear decision matrix, to the analytical figures of merit and ranked using a formal ranking algorithm (e.g. Technique for Order of Preference by Similarity to Ideal Solution). More than one metric should be reported, weighting schemes and inputs should be made public, and fitness for purpose should not be forgotten when assessing the greenness. With the overall field trend towards balanced assessment systems which incorporate practicality and performance, the decision-analytic perspective presented herein offers a coherent perspective on the various tools and their derivatives, as well as the instruments not yet developed, and how they can be understood, interpreted, and used responsibly in the analytical method engineering process.

CONFLICT OF INTEREST

The author declares no competing financial or personal interests.

REFERENCES

- Anastas, P. T. (1999). Green chemistry and the role of analytical methodology development. *Critical Reviews in Analytical Chemistry*, 29(3), 167-175. <https://doi.org/10.1080/10408349891199356>
- Anastas, P. T., & Warner, J. C. (1998). *Green chemistry: Theory and practice*. Oxford University Press.
- Armenta, S., Garrigues, S., & de la Guardia, M. (2008). Green analytical chemistry. *TrAC Trends in Analytical Chemistry*, 27(6), 497-511. <https://doi.org/10.1016/j.trac.2008.05.003>
- Ballester-Caudet, A., Campins-Falco, P., Perez, B., Sancho, R., Lorente, M., Sastre, G., & Gonzalez, C. (2019). A new tool for evaluating and/or selecting analytical methods: Summarising the information in a hexagon. *TrAC Trends in Analytical Chemistry*, 118, 538-547. <https://doi.org/10.1016/j.trac.2019.06.015>
- Bigus, P., Namieśnik, J., & Tobiszewski, M. (2016). Application of multicriteria decision analysis in solvent type optimisation for chlorophenols determination with a dispersive liquid-liquid microextraction. *Journal of Chromatography A*, 1446, 21-26. <https://doi.org/10.1016/j.chroma.2016.03.065>
- Fabjanowicz, M., Plotka-Wasyłka, J., & Tobiszewski, M. (2021). Multicriteria decision analysis and grouping of analytical procedures for phthalate determination in disposable baby diapers. *Molecules*, 26(22), 7009. <https://doi.org/10.3390/molecules26227009>
- Galuszka, A., Migaszewski, Z. M., Konieczka, P., & Namieśnik, J. (2012). Analytical Eco-Scale for assessing the greenness of analytical procedures. *TrAC Trends in Analytical Chemistry*, 37, 61-72. <https://doi.org/10.1016/j.trac.2012.03.013>

Gałuszka, A., Migaszewski, Z., & Namieśnik, J. (2013). The 12 principles of green analytical chemistry and the SIGNIFICANCE mnemonic of green analytical practices. *TrAC Trends in Analytical Chemistry*, 50, 78-84. <https://doi.org/10.1016/j.trac.2013.04.010>

Hicks, M. B., Farrell, W., Aurigemma, C., Lehmann, L., Weisel, L., Nadeau, K., Lee, H., Moraff, C., Wong, M., Huang, Y., & Ferguson, P. (2019). Making the move towards modernised, greener separations: Introduction of the analytical method greenness score (AMGS) calculator. *Green Chemistry*, 21(7), 1816-1826. <https://doi.org/10.1039/C8GC03875A>

Keith, L. H., Gron, L. U., & Young, J. L. (2007). Green analytical methodologies. *Chemical Reviews*, 107(6), 2695-2708. <https://doi.org/10.1021/cr068359e>

López-Lorente, A. I., Peña-Pereira, F., Pedersen-Bjergaard, S., Zuin, V. G., Ozkan, S. A., & Psillakis, E. (2022). The ten principles of green sample preparation. *TrAC Trends in Analytical Chemistry*, 148, 116530. <https://doi.org/10.1016/j.trac.2022.116530>

Manousi, N., Wojnowski, W., Plotka-Wasyłka, J., & Samanidou, V. (2023). Blue applicability grade index (BAGI) and software: A new tool for the evaluation of method practicality. *Green Chemistry*, 25(19), 7598-7604. <https://doi.org/10.1039/D3GC02347H>

Mansour, F. R., Omer, K. M., & Plotka-Wasyłka, J. (2024). A total scoring system and software for a complex modified GAPI (ComplexMoGAPI) application in the assessment of method greenness. *Green Analytical Chemistry*, 10, 100126. <https://doi.org/10.1016/j.greeac.2024.100126>

Mansour, F. R., Plotka-Wasyłka, J., & Locatelli, M. (2024). Modified GAPI (MoGAPI) tool and software for the assessment of method greenness: Case studies and applications. *Analytica*, 5(3), 451-457. <https://doi.org/10.3390/analytica5030030>

Nowak, P. M., & Koscielniak, P. (2019). What colour is your method? Adaptation of the RGB additive colour model to analytical method evaluation. *Analytical Chemistry*, 91(16), 10343-10352. <https://doi.org/10.1021/acs.analchem.9b01872>

Nowak, P. M., Koscielniak, P., Tobiszewski, M., Ballester-Caudet, A., & Campins-Falco, P. (2020). Overview of the three multicriteria approaches applied to a global assessment of analytical methods. *TrAC Trends in Analytical Chemistry*, 133, 116065. <https://doi.org/10.1016/j.trac.2020.116065>

Nowak, P. M., Wietecha-Posluszny, R., & Pawliszyn, J. (2021). White Analytical Chemistry: An approach to reconcile the principles of Green Analytical Chemistry and functionality. *TrAC Trends in Analytical Chemistry*, 138, 116223. <https://doi.org/10.1016/j.trac.2021.116223>

Nowak, P. M., Wojnowski, W., Manousi, N., Samanidou, V., & Plotka-Wasyłka, J. (2025). Red Analytical Performance Index (RAPI) and software: The missing tool for assessing methods in terms of analytical performance. *Green Chemistry*, 27, 5546-5553. <https://doi.org/10.1039/D4GC05298F>

Peña-Pereira, F., Wojnowski, W., & Tobiszewski, M. (2020). AGREE - Analytical Greenness metric approach and software. *Analytical Chemistry*, 92(14), 10076-10082. <https://doi.org/10.1021/acs.analchem.0c01887>

Plotka-Wasyłka, J. (2018). A new tool for the evaluation of the analytical procedure: Green Analytical Procedure Index. *Talanta*, 181, 204-209. <https://doi.org/10.1016/j.talanta.2018.01.013>

Plotka-Wasyłka, J., & Wojnowski, W. (2021). Complementary green analytical procedure index (ComplexGAPI) and software. *Green Chemistry*, 23(21), 8657-8665. <https://doi.org/10.1039/D1GC02318G>

Sajid, M., & Plotka-Wasyłka, J. (2022). Green analytical chemistry metrics: A review. *Talanta*, 238, 123046. <https://doi.org/10.1016/j.talanta.2021.123046>

Sammuto Bartolo, N., Gallo, L. L., Szyrmer, K., Buhagiar, P. I., & Vella Szijj, J. (2024). Greenness assessment of analytical methods for the determination of cannabinoids in oils using NEMI, Analytical Eco-Scale, AGREE and GAPI. *Analytical Methods*, 16(35), 5931-5942. <https://doi.org/10.1039/D4AY01083C>

Tobiszewski, M., Marc, M., Gałuszka, A., & Namieśnik, J. (2015). Green chemistry metrics with special reference to green analytical chemistry. *Molecules*, 20(6), 10928-10946. <https://doi.org/10.3390/molecules200610928>

Tobiszewski, M., Namieśnik, J., & Peña-Pereira, F. (2017). A derivatisation agent selection guide. *Green Chemistry*, 19(24), 5911-5922. <https://doi.org/10.1039/C7GC03108D>

Tobiszewski, M., & Orłowski, A. (2015). Multicriteria decision analysis in ranking of analytical procedures for Aldrin determination in water. *Journal of Chromatography A*, 1387, 116-122. <https://doi.org/10.1016/j.chroma.2015.02.009>

Tobiszewski, M., Tsakovski, S., Simeonov, V., Namieśnik, J., & Peña-Pereira, F. (2015). A solvent selection guide based on chemometrics and multicriteria decision analysis. *Green Chemistry*, 17(10), 4773-4785. <https://doi.org/10.1039/C5GC01615K>

Van Aken, K., Strekowski, L., & Patiny, L. (2006). EcoScale is a semi-quantitative tool to select an organic preparation based on economic and ecological parameters. *Beilstein Journal of Organic Chemistry*, 2, 3. <https://doi.org/10.1186/1860-5397-2-3>

Wojnowski, W., Tobiszewski, M., Peña-Pereira, F., & Psillakis, E. (2022). AGREEprep - Analytical greenness metric for sample preparation. *TrAC Trends in Analytical Chemistry*, 149, 116553. <https://doi.org/10.1016/j.trac.2022.116553>