

RESEARCH

Open Access



# Application of machine learning assisted multi-variate UV spectrophotometric models augmented by kennard stone clustering algorithm for quantifying recently approved nasal spray combination of mometasone and olopatadine along with two genotoxic impurities: comprehensive sustainability assessment

Ahmed Emad F. Abbas<sup>1\*</sup> , Mohammed Gamal<sup>2</sup>, Ibrahim A. Naguib<sup>3</sup>, Michael K. Halim<sup>1</sup>, Basmat Amal M. Said<sup>4</sup>, Mohammed M. Ghoneim<sup>5</sup>, Mohmeed M. A. Mansour<sup>6</sup> and Yomna A. Salem<sup>7</sup>

## Abstract

The recent approval of the nasal spray combination of mometasone (MOM) and olopatadine (OLO) presents a significant analytical challenge, as only a single reported method exists for its determination, deviating from eco-friendly practices. This study addresses this critical gap by pioneering the application of machine learning techniques to develop robust UV spectrophotometric approach for the simultaneous quantification of MOM and OLO, along with two genotoxic impurities: 4-dimethylamino pyridine (DAP) and methyl para-toluene sulfonate (MTS). By simultaneously determining these highly concerning genotoxic impurities and active pharmaceutical ingredients, this method underscores its paramount significance in upholding rigorous pharmaceutical quality standards and safeguarding patient safety. Applying the multilevel-multifactor experimental design, the calibration set was meticulously chosen at five different concentrations, yielding 25 calibration mixtures with central levels of 4, 46.5, 2.5, and 3  $\mu\text{g}/\text{mL}$  for MOM, OLA, MTS, and DAP, respectively. The key innovation lies in the strategic implementation of the Kennard-Stone Clustering Algorithm to create a robust validation set of thirteen mixtures, resolving the limitations of reported chemometric methods' random data splitting. This approach ensures unbiased evaluation across the full concentration space, improving the method's reliability and sustainability. The robustness of this approach was rigorously tested using five distinct chemometric models: principal component

\*Correspondence:

Ahmed Emad F. Abbas  
dr.ahmedeamad@gmail.com; ahmed.emad.pha@o6u.edu.eg

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

regression, classical least squares, partial least squares, genetic algorithm-partial least squares, and multivariate curve resolution-alternating least squares, demonstrating its broad applicability across diverse modeling techniques. All models successfully determined all components with excellent recovery, low bias-corrected prediction, and adequate limits of detection. The Greenness Index Spider Charts and the Green Solvents Selection Tool were used to choose environmentally conscious solvents. A comprehensive sustainability assessment employed six state-of-the-art tools, including the national environmental method index, complementary green analytical procedure index, analytical greenness metric, blue applicability grade index, carbon footprint analysis, and the red-green-blue 12 metrics. Favorable results across all metrics affirmed the method's eco-friendliness, real-world applicability, and cost-effectiveness, supporting sustainable development goals in pharmaceutical quality control processes.

**Keywords** Mometasone and olopatadine, Machine learning models, Kennard stone clustering algorithm, Greenness whiteness and blueness evaluation, Solvents selection tool and spider charts

## Introduction

The pharmaceutical industry faces an increasing demand for sustainable, cost-effective, and eco-friendly analytical practices [1–4]. The recent approval of the nasal spray combination of mometasone (MOM) and olopatadine (OLO) has presented an important analytical problem. To date, only a single reported method exists for its quantification [5], that departs from the guidelines of sustainability, white analytical chemistry (WAC) and green analytical chemistry (GAC) [5–7]. This method relies on toxic solvents, complex multi-step processes, costly instrumentation, and extensive infrastructure. These unsustainable practices jeopardize the environment, human health, and economic viability, highlighting the need for novel analytical approaches that incorporate emerging sustainability-driven disciplines [8, 9].

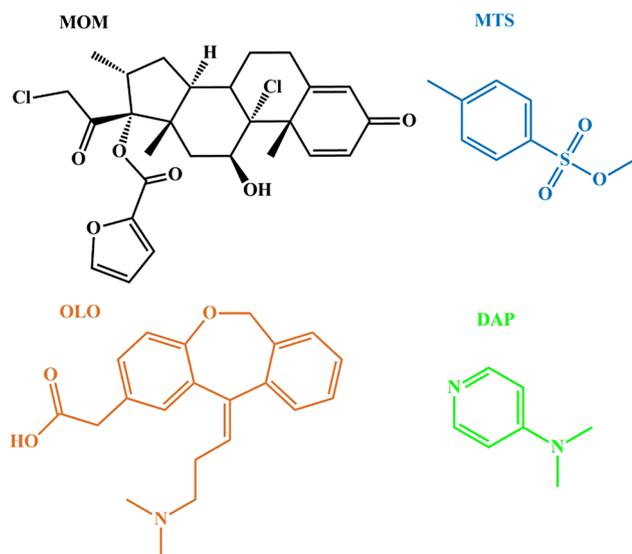
Ultraviolet-visible (UV) spectrophotometry is a viable option that aligns with the sustainability objectives of analytical chemistry. It utilizes affordable reagents and fundamental instrumentation, including light sources and cuvettes, while producing negligible hazardous waste [10]. However, significant spectrum overlaps frequently impede the direct measurement of medications [11]. This challenge can be addressed by strategically applying chemometric techniques, transforming UV spectroscopy into an effective and eco-friendly analytical approach suitable for sustainable and efficient pharmaceutical quality control processes [12, 13].

A prevalent limitation in up-to-date chemometrics studies is the routine use of random splitting to divide data into sets for validation and calibration [14]. While simple to implement, using this method could result in validation sets that don't accurately reflect the whole sample, which could introduce bias or unrealistic model accuracy expectations. To resolve this matter, our study employs the Kennard Stone Clustering (KSC) Algorithm, a powerful statistical technique, to systematically generate well-balanced validation sets [15]. The KSC approach partitions the feature space that was modelled into separate clusters, guaranteeing that validation samples fully capture the range and each variable's distribution. This

approach promotes analytical sustainability by reducing resource usage and waste production while enhancing reliability with less but strategically allocated validation samples.

The assessment and control of genotoxic impurities (GTIs) in pharmaceutical products have acquired significant attention because they have the capacity to cause genetic alterations and raise the risk of cancer [16, 17]. Amongst the impurities linked to MOM and OLO, 4-dimethylamino pyridine (DAP) and methyl para-toluene sulfonate (MTS) (Fig. 1) are especially troubling examples of GTIs [18, 19]. Even at trace amounts, their presence can harm DNA, leading to mutations and chromosomal aberrations. The genotoxicity of MTS and DAP underscores the requirement for sensitive, rapid, and economical analytical methods to accurately quantify and monitor these impurities in MOM and OLO products, ensuring patient safety. Notably, a substantial gap in quality control standards is highlighted by the fact that no analytical approach for the concurrent determination of OLO, MOM, and these crucial GTIs has been described.

This study aims to address all previous critical gaps in pharmaceutical analysis through three interconnected objectives. Firstly, we develop the first chemometric approach for the concurrent quantification of MOM and OLO, and their genotoxic impurities MTS and DAP without separation. Secondly, we implement the KSC for the first time to create robust validation sets, enhancing model reliability across all concentration ranges. Thirdly, we prove the method's sustainability, agreement with GAC and WAC basic concepts, and superiority over the reported method using six cutting-edge tools. These include the National Environmental Method Index (NEMI), complementary Green Analytical Procedure Index (Complex GAPI), Analytical Greenness Metric (AGREE), Blue Applicability Grade Index (BAGI) and Red-Green-Blue 12 (RGB12) metrics, and carbon footprint assessment. Overall, this comprehensive approach aims to pioneer a paradigm shift in pharmaceutical analysis, promoting sustainable, efficient, and economically viable quality control practices.



**Fig. 1** The chemical structure of MOM, OLO, MTS, and DAP

## Experimental

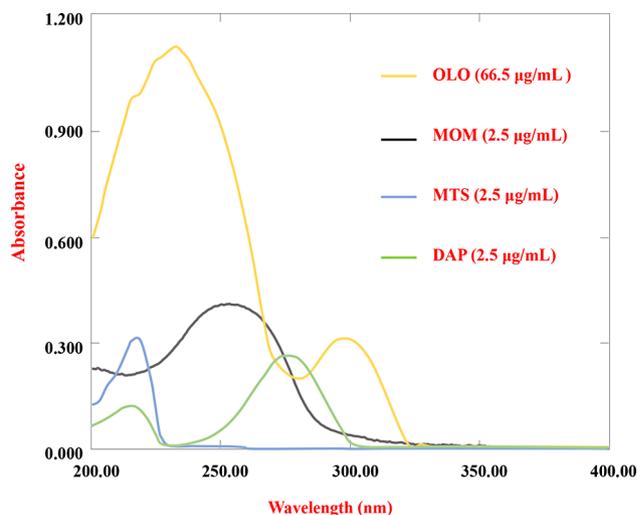
### Analytical instruments and software tools

Dataset were attained utilizing a high-precision dual-beam UV-Vis spectrophotometer UV-1800 (Shimadzu Corporation, Kyoto, Japan) outfitted with 1 cm quartz cells. Measurements were taken utilizing UV-Probe software (version 2.42, Shimadzu Corporation). The instrument settings were set as follows: A 0.1 nm sampling interval, a 1.0 nm slit width, and fast mode with one scan. Shimadzu AGE-220 analytical balance (Shimadzu Corporation, Kyoto, Japan) for precise weighing. Julabo ultrasonic bath (Julabo Labortechnik GmbH, Seelbach, Germany) for sample extraction.

MATLAB R2013a (version 8.2.0.701, MathWorks, Natick, MA, USA) was utilized for data processing and chemometric analysis, in addition to PLS Toolbox version 2.0, designed by Eigenvector Research (Wenatchee, WA, USA) and the freely accessible toolbox of MCR-ALS (<http://www.mcrals.info>). The KSC method was executed utilizing custom MATLAB R2013a scripts, incorporating improved statistical algorithms. SDAGI and One-way analysis of variance (ANOVA) were carried out by using Microsoft Excel (Microsoft Corporation, Redmond, WA, USA). The GSST was employed using the free calculator algorithm available at <http://green-solvent-tool.herokuapp.com/>.

### Reagents and materials

Ultrapure water (18.2 M $\Omega$ .cm at 25 °C) was acquired from a Milli-Q water purification system (Millipore, Bedford, MA, USA). MOM (99.30% purity) and OLO (99.78% purity) were obtained from Glenmark Pharmaceuticals Egypt (Cairo, Egypt). MTS (99.60% purity) and DAP (99.52% purity) were acquired from Sigma-Aldrich



**Fig. 2** The zero-order absorption spectrum MOM, OLO, MTS, and DAP

Co. (St. Louis, MO, USA). Hexane, acetonitrile, ethanol, methanol, ethyl acetate, and chloroform were sourced from Merck KGaA (Darmstadt, Germany). Ryaltris<sup>®</sup> nasal spray (batch no: PL25258-0331, Glenmark Pharmaceuticals, Egypt) containing 665 µg OLO and 25 µg MOM per metered dose was acquired from a local pharmacy.

### Preparation of standard solutions

Individual stock solutions of OLO, MOM, DAP, and MTS were formulated at 100 µg/mL each. For each compound, 10 mg of the reference standard was accurately weighed utilizing the analytical balance and was then transferred thoroughly to a 100 mL volumetric flask. The standards were dissolved in HPLC-grade ethanol with the aid of sonication for 5 min to ensure complete dissolution. Using the same solvent, the volume was brought up to the corresponding mark and mixed thoroughly. Daily preparation of working standard solutions was performed by diluting the stock solutions with ethanol to obtain the planned concentration ranges for analysis. All dilutions were performed using Class A volumetric glassware or calibrated micropipettes to ensure accuracy.

### Spectral properties and linearity

The UV absorption spectra of MOM, OLA, MPS, and DAP were separately measured within the 200–400 nm wavelength range, as shown in (Fig. 2). The concentration ranges for MOM, OLA, MPS, and DAP were 1–7, 6.5–86.5, 0.5–4.5, and 1–5 µg/mL, respectively. These ranges were deliberately selected to cover the instrument's linear dynamic range and to reflect the drug formulations' concentrations.

### Experimental design

An organized design of experiment is essential for acquiring optimal results and information-rich spectral

data and in accordance with the plan provided by Brereton et al. [20], we created a calibration set with multiple levels and factors, composed of 25-different mixtures. Concentration ranges stated for each analyte in the calibration set were 1–7, 6.5–86.5, 0.5–4.5, and 1–5  $\mu\text{g/mL}$  for MOM, OLA, MPS, and DAP, respectively. To construct a reliable validation set that consistently assesses the model's performance across the entire concentration space, we used the KSC algorithm. This approach resulted in 13-validation mixtures, each selected from one of 13 strata with equal probability, covering the entire range of concentration. The validation and calibration designs align with GAC and WAC principles, as they are simple, highly sensitive, highly selective, easy to use, economically efficient, use minimal solvent, save time, and are environmentally friendly. Mixtures were prepared using calibrated micropipettes and HPLC-grade ethanol as the solvent in 25 mL Class A volumetric flasks. Absorption spectra were measured utilizing 1 cm quartz cuvettes within the wavelength range of 200–400 nm, with ethanol as the blank. After careful evaluation, we excluded lower and higher spectral ranges due to high levels of noise or insufficient signal. The resulting spectral data matrix extended from 210 to 320 nm with a resolution of 1 nm, resulting in 111 wavelength data points for each spectrum. To ensure data quality and reproducibility, each mixture was prepared in triplicate and analyzed three times. The spectra were averaged for subsequent analysis. We also implemented a randomization strategy for sample preparation and measurement to minimize systematic errors.

### Models development and optimization

For model construction, 5 chemometric techniques were applied, including classical least squares (CLS), principal component regression (PCR), partial least squares (PLS), multivariate curve resolution-alternating least squares (MCR-ALS), and PLS combined with genetic algorithm-based wavelength selection (GA-PLS). Each model was rigorously optimized using the 25-mixture calibration set to prevent overfitting and ensure robustness. In the CLS model, regression calculations were performed independently at each wavelength, employing a moving window method for wavelength selection. Window widths ranging from 5 to 30 nm were evaluated through cross-validation, and the most suitable width was determined based on the balance between noise reduction and information retention. For both PCR and PLS models, the count of Latent Variables (LVs) was incrementally varied from 1 to 10, with the optimal value identified through Venetian blinds cross-validation, utilizing the Root Mean Squared Error of Cross-Validation (RMSECV) as the selection criterion, seeking a balance between complexity and model fit. In the GA-PLS model, the parameters of the Genetic

Algorithm were optimized to identify the most relevant spectral regions, focusing on balancing accuracy, reliability, and generalizability, with selected wavelengths fed into the PLS model for final calibration. The MCR-ALS model employed non-negativity constraints for the spectral and concentration profiles, using a non-negative least squares (nnls) algorithm, with constraints optimized to achieve satisfactory results with minimal iterations.

### Analytical performance parameters

Comprehensively evaluated utilizing a range of analytical metrics to assess their predictive effectiveness, accuracy, precision, robustness, and sensitivity [21]. For the calibration set, we calculated the standard error of calibration (SEC), root mean square error of calibration (RMSEC), and RMSECV to gauge the model's fitting and predictive capabilities. The performance of the validation set was quantified utilizing the relative root mean square error of prediction (RRMSEP), while the model's ability to generalize effectively was assessed via the root mean square error of prediction (RMSEP). Furthermore, we utilized the bias-corrected mean square error of prediction (BCRMSEP) to evaluate the predictability and precision of novel samples.

The subsequent formulas were employed to calculate these metrics [21]:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}}$$

$$Bias = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{n}$$

$$SEC = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i - bias)^2}{n - 1}}$$

$$RRMSEP\% = \frac{\frac{1}{n} \sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2}}{\bar{y}_i} \times 100$$

$$BCRMSEP = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n} - (bias)^2$$

In this equation, the variable  $y_i$  denotes the outcome of the experiment,  $\hat{y}_i$  represents the predicted value,  $n$  represents the samples' number, and  $\bar{y}_i$  represents the experimental values' mean.

For assessment of accuracy, we conducted an analysis of three separate concentration points that fell within the linear range specified for each analyte in triplicate: MOM (2.5, 4, and 5.5  $\mu\text{g/mL}$ ), OLO (26.5, 46.5, and 66.5  $\mu\text{g/mL}$ ), MTS (1.5, 2.5, and 3.5  $\mu\text{g/mL}$ ), and DAP (2, 3, and

4 µg/mL). Recovery percentages (%R) were calculated for each concentration. Precision was evaluated through repeatability (intra-day) and intermediate (inter-day) precision studies. We analyzed three different samples in triplicate on the same day for repeatability, and on three different days for intermediate precision, using the same concentration levels as in the accuracy study. Precision was represented as a percentage relative standard deviation (%RSD). The method's robustness was tested by introducing small, deliberate variations in experimental parameters: spectral bandwidths (1 nm versus 0.8 nm slit width), wavelength intervals (1 nm compared to 0.9 nm), and scan speeds (fast versus medium). The limit of quantification (LOQ) and limit of detection (LOD) were computed depending on the standard error of the regression curve, which was plotted using the experimentally determined concentrations (expressed on the y-axis) against the predicted concentrations by the models (expressed on the x-axis), applying the IUPAC-recommended Eqs. [17, 22]:

$$\text{LOD} = 3.3 / S$$

$$\text{LOQ} = 10 / S$$

Where  $S$  represents the slope, which reflects the sensitivity of the method to variations in concentration, while  $\sigma$  denotes the standard error of the prediction-versus-actual regression, which measures the variability of the predicted values.

#### Analysis of pharmaceutical dosage forms

The proposed method was applied to the analysis of Ryaltris® nasal spray labeled to consist of 665 µg of OLO and 25 µg of MOM per mL. A 50 mL aliquot of the nasal spray was accurately transferred to a 100 mL volumetric flask, to which about 30 mL of HPLC-grade ethanol was placed. The mixture was sonicated for 15 min to ensure complete drug extraction, cooled to room temperature, and diluted to volume with ethanol to acquire concentrations of 12.5 µg/mL MOM and 332.5 µg/mL OLO. This stock solution was then passed through a 0.45 µm PTFE membrane filter, with the initial few milliliters of the filtrate discarded. An appropriate aliquot of the filtered stock solution was subsequently diluted with ethanol. The spectra of the resulting solutions were recorded between 200 and 400 nm, using ethanol as the reference blank. The developed chemometric models were applied to the acquired spectra to ascertain the concentrations of OLO and MOM. The accuracy was assessed by introducing spiked standards in triplicate at four different concentration levels. RSD% and R% were then computed. Ultimately, the outcomes generated by the suggested chemometric techniques were statistically compared to

those from the validated reference method using one-way analysis of variance (ANOVA), ensuring a comprehensive evaluation of the method's performance in quantifying MOM and OLO in the presence of potential matrix effects from the pharmaceutical formulation and demonstrating its applicability for routine quality control analysis.

## Results and discussion

In this research, we have effectively designed and validated novel chemometric approaches for the concurrent quantification of MOM and OLO, along with the genotoxic contaminants MTS and DAP, utilizing UV-visible spectrophotometry. Five different chemometric models were meticulously tested, demonstrating the robustness of the innovative KSC validation method and its broad applicability across these diverse modeling techniques, as well as the superior integrity of the multivariate concentration domain represented by the experimental data. Thorough validation verified the precision, accuracy, and sensitivity of the proposed approach. Adhering to the concepts of GAC and WAC, a thorough evaluation of the method's greenness, whiteness, and blueness was undertaken utilizing six state-of-the-art methods: Complex GAPI, NEMI, AGREE, carbon footprint analysis, RGB12, and BAGI. The advantageous outcomes across all metrics affirmed the method's eco-friendliness, real-world applicability, and cost-effectiveness. Furthermore, the selection of environmentally friendly solvents using the GSST and SDAGI contributed significantly to the method's overall sustainability. This holistic approach to method development not only addresses the analytical challenges posed by the MOM-OLO nasal spray combination but also aligns with broader sustainable development goals in pharmaceutical quality control processes.

### Green solvent selection

#### Green solvent selection tool (GSST)

The concept of a "green solvent" is complex, but it generally refers to solvents with minimal health risks, greater safety profiles, and minimal effect on the environment over their lifecycle. In pursuit of sustainability, major pharmaceutical companies such as Pfizer, Sanofi, and GlaxoSmithKline (GSK) have developed standards for choosing suitable green solvents. The GSK Solvent Sustainability Guidelines, in particular, provide comprehensive solvent sustainability guides that compare the advantages and disadvantages of different solvents based on information from their Safety Data Sheets (SDSs). Building on these guidelines, Christian Larsen's research introduced a novel chemometric tool for green and sustainable solvent selection [23]. This tool calculates a Greenness score ( $G$ ) using the equation:

$G = \sqrt[4]{(H \times S \times E \times W)}$ , where H, S, E, and W represent Health, Safety, Environment, and Waste Disposal factors, respectively. The G score ranges from 1 to 10, with better scores signifying more sustainability and greenness. We applied this tool to evaluate seven polar solvents: ethanol, acetonitrile, ethyl acetate, water, methanol, hexane, and chloroform. The findings indicated that ethanol, methanol, water, acetonitrile, and ethyl acetate attained markedly greater G scores in contrast to the more harmful solvents, such as chloroform and hexane (Fig. S1). Particularly, acetonitrile, water, ethyl acetate, ethanol, and methanol attained G scores of 7.3, 6.7, 6.6, 5.8, and 5.8, respectively, reflecting positive assessments in safety, health impacts, environmental impact, and disposal of waste. Based on these greenness assessments, Acetonitrile, water, ethyl acetate, ethanol, and methanol were chosen for further evaluation due to their superior environmental compatibility. This selection aligns with our objective to develop an analytical technique that follows the guidelines of green chemistry by using safer and more environmentally benign chemicals.

#### **Spider diagram for assessment of the greenness index (SDAGI)**

Despite tools like the GSST facilitating initial solvent selection, a more thorough assessment of reagent greenness depending on comprehensive experimental data is essential. The SDAGI tool offers a valuable qualitative method for this purpose, utilizing SDS data [24, 25]. This technique leverages Safety, Health, and Environmental (SHE) data on reagent effects and properties to compute greenness scores based on important parameters and produce visual spider diagrams. As shown in (Fig. S1), the SDAGI hierarchical spider chart shows points from -5 to +5, which are obtained from 5 evaluation subcategories: general properties, health impact, fire safety, odour, and stability. This visual representation simplifies the process of evaluating and comparing reagents. To examine SDS data for the solvents we had previously chosen, we used SDAGI. The findings showed that ethanol and ethyl acetate had greater overall greenness evaluates (1.33 and 1.44, respectively) and bigger safety areas. At the same time, methanol and acetonitrile showed lower scores of -0.12 and -0.30, as detailed in (Table S1). Based on these findings, we narrowed our selection to water, ethyl acetate, and ethanol as the most eco-friendly solvents for further evaluation. To determine the most appropriate green solvent for our specific analytical needs, we examined the UV spectra of MOM, OLA, MPS, and DAP in these three solvents. Our spectral analysis disclosed that ethanol offered a better UV response and better spectral shape for all four compounds in contrast to water and ethyl acetate. This enhanced analytical performance, combined with ethanol's favorable environmental profile,

led us to conclude that ethanol is the primarily appropriate green solvent for additional studies incorporating MOM, OLA, MPS, and DAP.

#### **Chemometric models**

The complex spectral overlapping of MOM, OLA, MPS, and DAP in their UV absorption spectra, as shown in (Fig. 2), required the use of powerful chemometric solutions in order to achieve precise quantification. The acquisition of concentration data hidden within the intricate spectrum patterns was made possible by chemometric modelling [26–28]. Using ternary mixtures' whole spectrum data obtained between 200 and 400 nm, we created multivariate calibration models. A critical aspect of accurate modeling was reducing noise and artefacts while identifying useful spectral regions. The ideal spectral range, recorded at 1 nm intervals, was found to be 210–320 nm through an iterative tuning procedure that was guided by noise levels, undesired signals, and model performance indicators. This resulted in 111 data points that formed the input matrix for our chemometric analyses. This optimized data set was then used to construct and refine five distinct chemometric models using MATLAB including PCR, CLS, PLS, GA-PLS, and MCR-ALS models. Each of these models offers unique advantages in handling spectral data. CLS provides a straightforward approach based on Beer's Law but can be sensitive to spectral interferences. PCR reduces data dimensionality and can handle collinearity effectively. PLS balances the information in both spectral and concentration data, often leading to robust predictions. GA-PLS incorporates evolutionary algorithms to optimize wavelength selection, potentially improving model performance. MCR-ALS offers the ability to resolve pure component spectra and concentration profiles, providing additional insights into the mixture composition. By employing this diverse array of chemometric techniques, we aimed to comprehensively evaluate the spectral data and ensure robust quantification of all four compounds despite their spectral similarities. This approach not only addresses the immediate analytical challenge but also provides a comparative framework for assessing the strengths of different chemometric methodologies in complex pharmaceutical analyses.

#### **Calibration set design**

The calibration set was constructed employing a multifactorial, multilevel experimental strategy, following the methodology outlined by Brereton et al. [20]. This method was applied in order to produce a set of 25 calibration mixtures, ensuring comprehensive coverage of the concentration range for all analytes. The concentration levels for each analyte were strategically selected at five points: -2, -1, 0 (center point), +1, and +2, as detailed

in (Table 1). This design allows for the exploration of both main effects and potential interactions between analytes while minimizing collinearity in the calibration matrix. This design offers several advantages including efficient coverage of the experimental domain with a rather small number of mixtures, the ability to model non-linear responses, reduced risk of overfitting due to the structured nature of the design, and the decrease of

the reagent consumption and waste generation which aligns well with the principles of GAC.

#### Validation set design

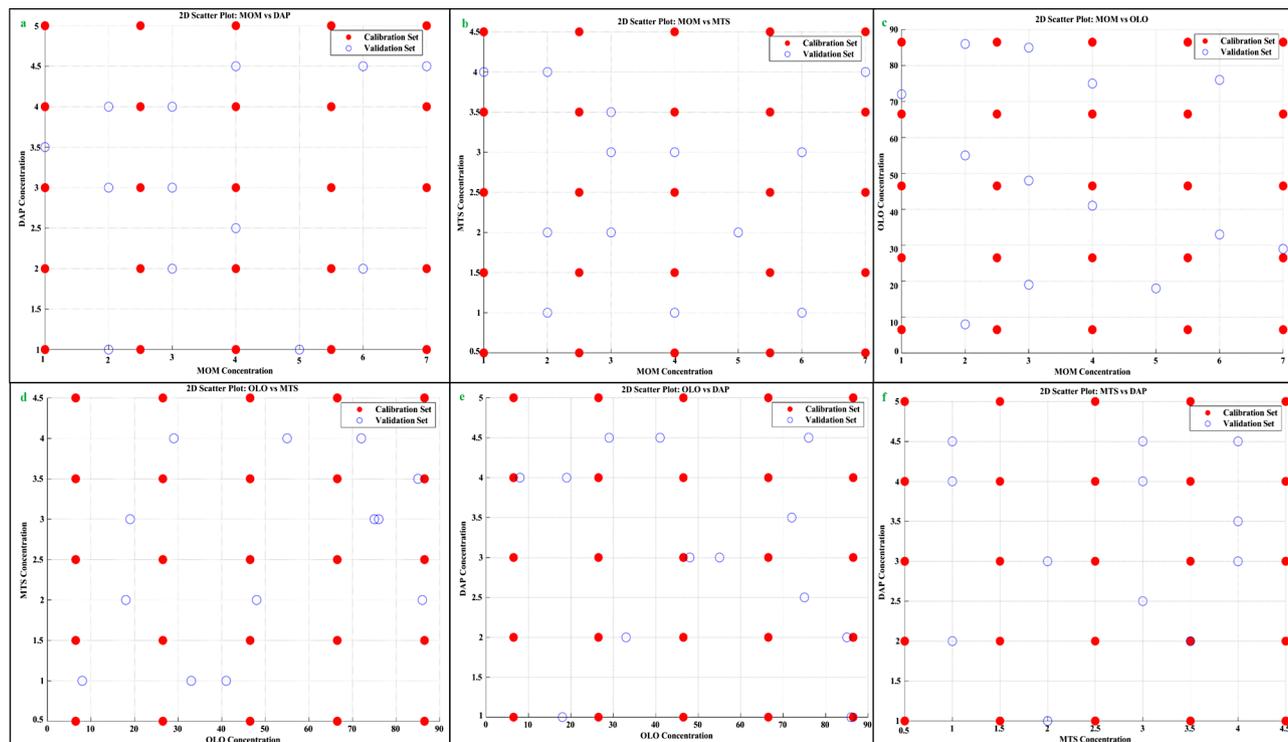
To rigorously assess the predictive performance of our chemometric models, a validation set encompassing the complete range of concentration of the calibration process was essential. To overcome the deficiencies of basic random sampling, which may result in inadequate coverage and skewed accuracy assessments, we employed the KSC algorithm for validation set selection. The KSC algorithm was applied utilizing MATLAB with custom scripts. This approach strategically divides the multivariate concentration space into unique clusters, guaranteeing thorough coverage of all analyte concentration ranges and their combinatorial distributions. Key aspects of the KSC-based validation set design include implementing a structured classification approach to segment multi-component concentration data into well-defined clusters spanning multiple dimensions, selection of a single validation sample from each stratum, equivalent to a concentration regime that is equally likely to occur, and generation of a very informative yet concise thirteen-mixture validation set, as shown in (Table 1). The effectiveness of the KSC approach in achieving uniform coverage of the concentration space is illustrated in (Fig. 3) which shows scatter plots of the validation samples through the range of analyte concentrations. Unlike ordinary random sampling, the KSC approach gave thorough representation of the concentration range with fewer samples compared to random sampling, mitigation of potential biases arising from uneven sample distribution, enhanced efficiency in material usage, agreement with GAC guiding principles, and robust evaluation of model performance across diverse sample compositions. Furthermore, the use of a statistically optimized validation set contributes to the overall sustainability of the method by reducing reagent consumption and waste generation, while maintaining high standards of analytical rigor. The successful application of KSC in this study demonstrates its potential as a valuable tool for enhancing the reliability and efficiency of chemometric method validation in complicated pharmaceutical analyses.

#### CLS model

It is alternatively called K-matrix calibration [29] and was employed as an initial step in our analytical process. The Beer-Lambert law, which asserts a direct proportionality between analyte concentration and absorbance over several wavelengths, and multivariate linear regression serve as the foundation for this technique. In our study, we constructed a calibration matrix using spectral data as predictors and known concentrations of the analytes (MOM, OLA, MPS, and DAP) as response variables.

**Table 1** Experimental design of 25 calibration mixtures formulated using a five-level, five-factor approach, and an additional 13 validation mixtures generated by KSC for chemometric models' validation

Mix no	Calibration set ( $\mu\text{g/mL}$ )			
	MOM	OLO	MTS	DAP
1.	4	46.5	2.5	3
2.	4	6.5	0.5	5
3.	1	6.5	4.5	2
4.	1	86.5	1.5	5
5.	7	26.5	4.5	3
6.	2.5	86.5	2.5	2
7.	7	46.5	1.5	2
8.	4	26.5	1.5	4
9.	2.5	26.5	3.5	5
10.	2.5	66.5	4.5	4
11.	5.5	86.5	3.5	3
12.	7	66.5	2.5	5
13.	5.5	46.5	4.5	5
14.	4	86.5	4.5	1
15.	7	86.5	0.5	4
16.	7	6.5	3.5	1
17.	1	66.5	0.5	3
18.	5.5	6.5	2.5	4
19.	1	46.5	3.5	4
20.	4	66.5	3.5	2
21.	5.5	66.5	1.5	1
22.	5.5	26.5	0.5	2
23.	2.5	6.5	1.5	3
24.	1	26.5	2.5	1
25.	2.5	46.5	0.5	1
Mix no	Validation set ( $\mu\text{g/mL}$ )			
	MOM	OLO	MTS	DAP
1.	7	29	4	4.5
2.	3	19	3	4
3.	3	48	2	3
4.	6	76	3	4.5
5.	3	85	3.5	2
6.	4	75	3	2.5
7.	5	18	2	1
8.	2	86	2	1
9.	4	41	1	4.5
10.	2	55	4	3
11.	2	8	1	4
12.	1	72	4	3.5
13.	6	33	1	2



**Fig. 3** Two-dimensional scatter plot illustrating the validation set relationships: (a) MOM versus DAP, (b) MOM versus MTS, (c) MOM versus OLO, OLO versus MTS, OLO versus DAP, and MTS versus DAP designed by KSC design as optimal-space filling design

This design ensured that the number of variables (wavelengths) was appropriately managed to allow for a fully multivariate analysis.

Mathematically, this is expressed as:

$$A = KC + E$$

where  $A$  is the matrix of absorbance spectra,  $K$  is the matrix of pure component spectra ( $K$ -matrix),  $C$  is the concentration matrix, and  $E$  is the residual matrix accounting for errors or deviations from linearity.

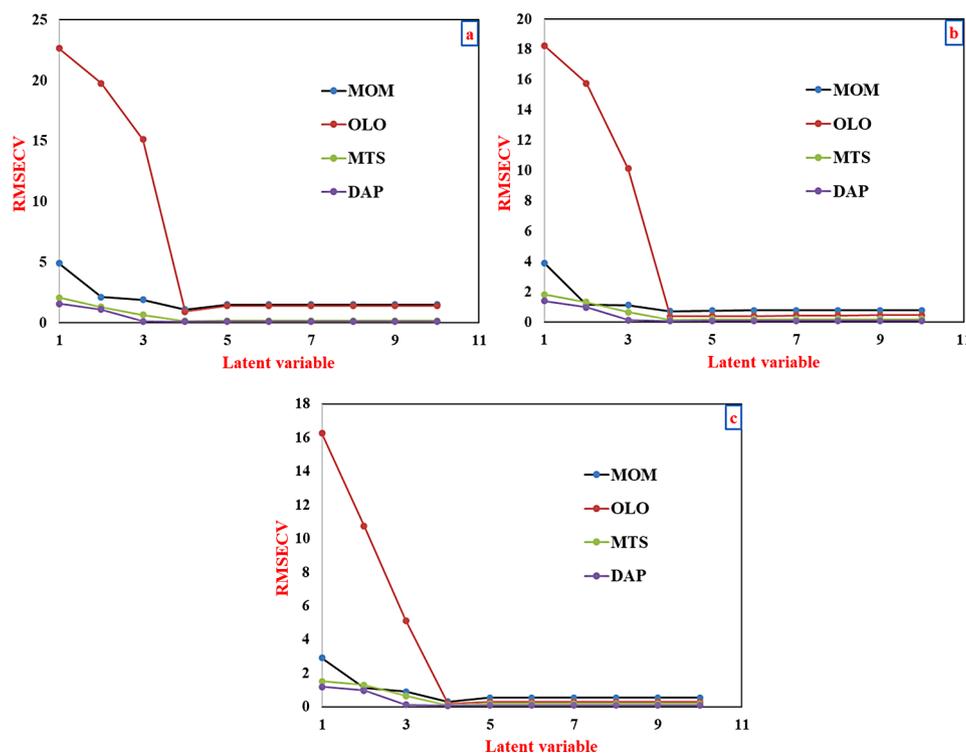
Initially, our predictions using the CLS model were suboptimal, likely due to baseline shifts or other systematic deviations from ideal linearity. To address this, we incorporated an intercept term into the regression model, which significantly improved predictive accuracy, leading to %R of 99.75%, 100.24%, 100.51%, and 99.48% for MOM, OLA, MPS, and DAP, respectively. While these results were promising, we recognize that CLS has inherent obstacles, particularly in managing non-linearity and inter-component interactions, which are common challenges in complex multi-component systems. Despite providing a solid starting point and performing well with our dataset, the CLS model was ultimately outperformed by more advanced chemometric techniques that could better accommodate the complexities of our analysis.

#### PCR model

To address the limitations of CLS, we next implemented the PCR model, which integrates Principal Component Analysis (PCA) with Multiple Linear Regression (MLR) [30–32]. This method effectively reduces data dimensionality while still maintaining much of the spectral fluctuation, making it particularly useful for complex multi-component systems where collinearity among spectral variables may be an issue. Our PCR model development process began with PCA performed on spectral data to obtain Principal Components (PCs). Next, we used cross-validation to determine the ideal number of LVs. The calibration set's mean-centered spectra were used to construct the PCR model, employing leave-one-out cross-validation. RMSECV was calculated after each LV addition to assess model performance. Ultimately, we found that optimal performance was achieved with four LVs for modeling MOM, OLA, MPS, and DAP, as depicted in (Fig. 4). The RMSECV values were 1.1023, 0.9276, 0.16164, and 0.0993 for MOM, OLA, MPS, and DAP, respectively. The PCR model demonstrated improved performance over CLS, effectively handling spectral noise and interference while maintaining good predictive accuracy.

#### PLS model

In our study, we implemented the PLS model, a powerful chemometric model that integrates aspects of regression



**Fig. 4** RMSECV plot of the calibration set as a function of the optimum LVs for the (a) PCR, (b) PLS, and (c) GA-PLS models

modeling and PCA. In contrast to PCR, which maximizes spectral variance, PLS is designed to enhance the relationship between spectral information (independent variables) and concentration levels (dependent variables), emphasizing on retaining variables that are directly associated with the prediction of concentration [14]. We chose to construct separate PLS models for each of the four analytes, MOM, OLA, MPS, and DAP, rather than a single multivariate model. This decision was based on the distinct spectral properties and concentration ranges of each analyte, which necessitated a tailored modeling approach. By developing individualized models, we were able to fine-tune the selection of LVs specific to each analyte through leave-one-out cross-validation, optimizing the predictive accuracy for each component. The optimal performance for each analyte was achieved using four LVs, as demonstrated in (Fig. 4), with RMSECV values of 0.7111, 0.3657, 0.1364, and 0.0823 for MOM, OLA, MPS, and DAP, respectively. These results indicate improved predictive capability compared to the PCR model; a common outcome given PLS's ability to retain concentration-dependent variables more effectively. While constructing individual models required additional computational effort, this approach provided a more nuanced understanding of the spectral-concentration relationship for each analyte. It also minimized the potential issues associated with multicollinearity and noise that could arise in a more generalized, multi-component model. Overall, the

use of separate PLS models for each analyte aligned with our goal of achieving precise and accurate quantification in this complex pharmaceutical analysis.

#### GA-PLS model

To further enhance our analytical approach, we developed a GA-PLS model. This advanced technique uses GA to optimize and refine the PLS models by judicious selection of the primarily relevant wavelengths extracted from the dataset. GA replicates biological evolution through mechanisms such as selection, recombination, and mutation to determine the most suitable set of parameters that enhance the accuracy of predictions [14]. The primary advancement in this method was the removal of unnecessary spectral wavelengths with low analyte signal, achieved through the GA enhancement algorithm. To ensure statistical robustness in variable selection, we fine-tuned GA parameters such as convergence criteria, population size, and mutation rate (Table 2). Several iterative GA operations resulted in the reduction of absorbance matrices for MOM, OLA, MPS, and DAP by 60%, 41%, 53%, and 45%, respectively, keeping only the variables that are most pertinent to measuring each component. The GA-PLS models were then rebuilt using this optimized matrix. Cross-validation found that four LVs were optimally sufficient for the GA-PLS model to represent MOM, OLA, MPS, and DAP, yielding RMSECV

**Table 2** Parameter adjustments optimized for the genetic algorithm used in variable selection

Parameters	Optimum values			
	MOM	OLO	MTS	DAP
Size of the population	44	48	36	44
Highest number of generations	51	73	65	60
Rate of mutation	0.005	0.005	0.005	0.005
Percentage of wavelength utilized at the initiation	15	15	15	15
Variable count within a defined window (window width)	2	2	2	2
Proportion of the population reaching convergence	80	80	80	80
Cross-type	Double	Double	Double	Double
Maximum number of latent variables	4	4	4	4
Cross-validation	Random	Random	Random	Random
Number of subsets to divide data into for cross-validation	5	5	5	5
Iterations per generation in cross-validation	2	2	2	2

values of 0.3111, 0.1987, 0.0996, and 0.0772, respectively, as illustrated in (Fig. 4).

### **MCR-ALS model**

The final chemometric technique employed in our study was the MCR-ALS model. This sophisticated technique seeks to identify the unique spectral and concentration characteristics of each component in an intricate mixture, even in the absence of prior information about the system. MCR-ALS operates by decomposing the data matrix with the help of a bilinear model, initially estimating the substances that are present, and then processing concentration and spectral profiles via the iterative ALS procedure, while adhering to certain constraints [33, 34]. In our implementation, the MCR-ALS analysis was initially performed by conducting an evolving factor analysis, setting a log eigenvalue threshold of -4, which led to the identification of a three-factor model. Non-negativity constraints were applied to both the concentration and spectral profiles, while correlation constraints were also incorporated for the concentration profiles [35]. The iterative process reached a steady state after fifteen iterations, meeting the predetermined convergence criterion of 20%. The model demonstrated excellent performance, achieving a variance percentage ( $r^2$ ) of 100 and a remarkably low fitting error (lack of fit %) of 0.0057. Figure 5 shows the individual spectral records that were resolved using the MCR-ALS model, which display a notable

resemblance to the previously measured absorption spectra. This alignment between resolved and measured spectra not only validates the model's accuracy but also highlights its potential for qualitative component detection alongside quantitative determination. The dual functionality of the MCR-ALS model, precise quantification coupled with qualitative elucidation, sets it apart as a versatile and comprehensive analytical tool.

### **KSC algorithm**

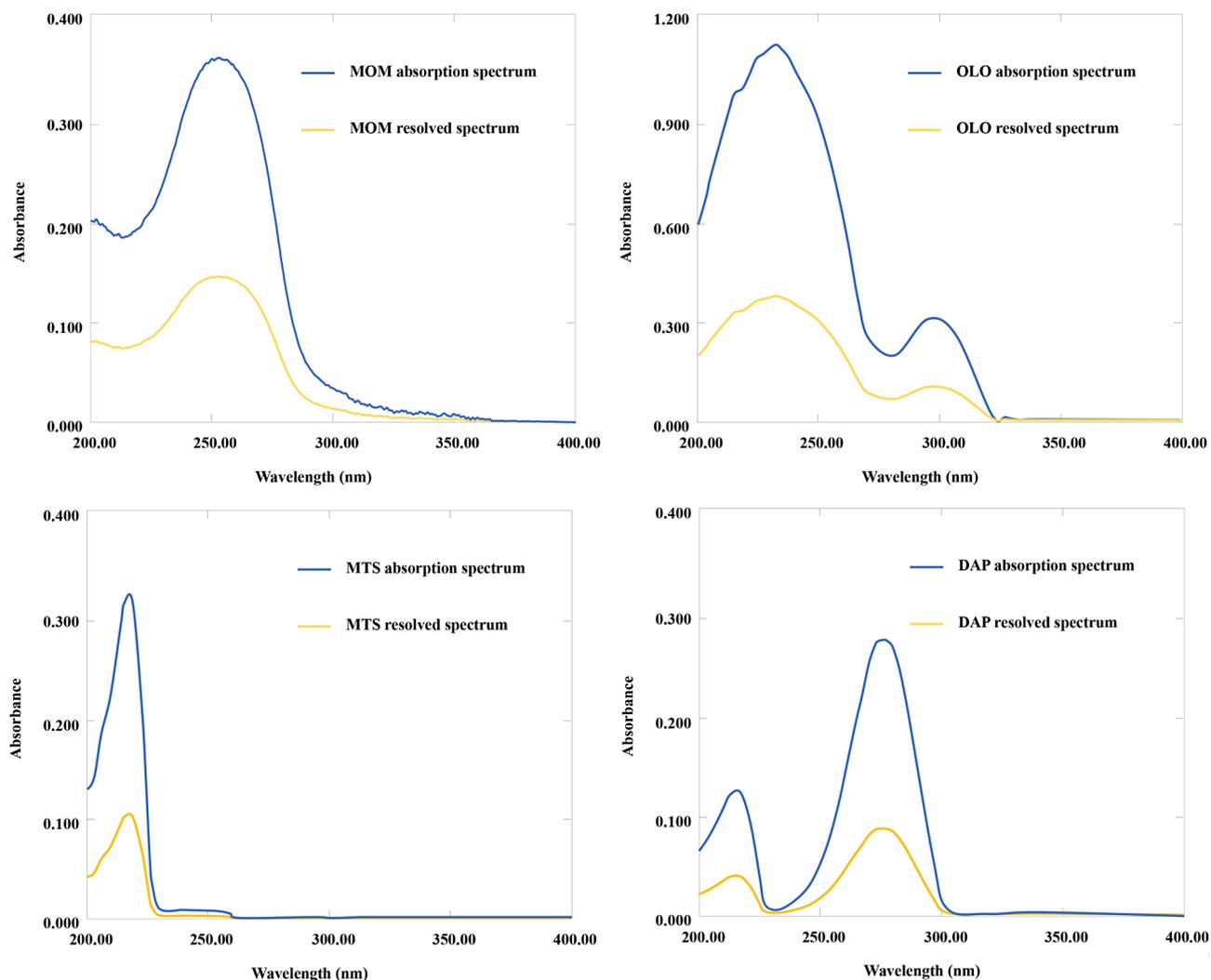
#### **Theoretical background of the Kennard-Stone algorithm**

The Kennard-Stone algorithm, a cornerstone in chemometric analysis, was introduced by R.W. Kennard and L.A. Stone in 1969 [36]. This algorithm is created to choose representative samples from a dataset in an optimal manner, ensuring comprehensive coverage of the feature space. Its primary goal is to enhance the selection process of calibration samples, thereby improving the robustness and accuracy of subsequent analytical models. The algorithm operates by iteratively selecting data points that maximize the minimum distance to the samples already chosen. This selection process guarantees that the chosen samples are diverse and span the entire feature space, leading to a more representative and unbiased calibration set. The Kennard-Stone algorithm's effectiveness in maintaining the integrity of the dataset's variability makes it a valuable tool in various applications within analytical chemistry, particularly in method development and validation. By ensuring that the calibration samples are uniformly distributed across the feature space, the Kennard-Stone algorithm enhances the predictive performance and generalizability of analytical models. This method's ability to systematically address the representativeness of sample selection underscores its importance in achieving reliable and reproducible analytical results.

#### **Kennard-stone algorithm for clustering**

The Kennard-Stone algorithm offers a robust theoretical basis for clustering by selecting representative samples that maximize the minimum distance between them, ensuring a diverse and well-distributed sample set as shown in (Fig. 6). This selection process can be adapted for clustering by viewing the representative samples as cluster centroids. Each new sample added to the set further refines these centroids, enhancing the diversity and representativeness of the clusters. This property makes it suitable for creating clusters that are well-distributed across the variable space, ensuring the representation of all significant variations in the dataset.

The new KSC algorithm proceeds as follows:



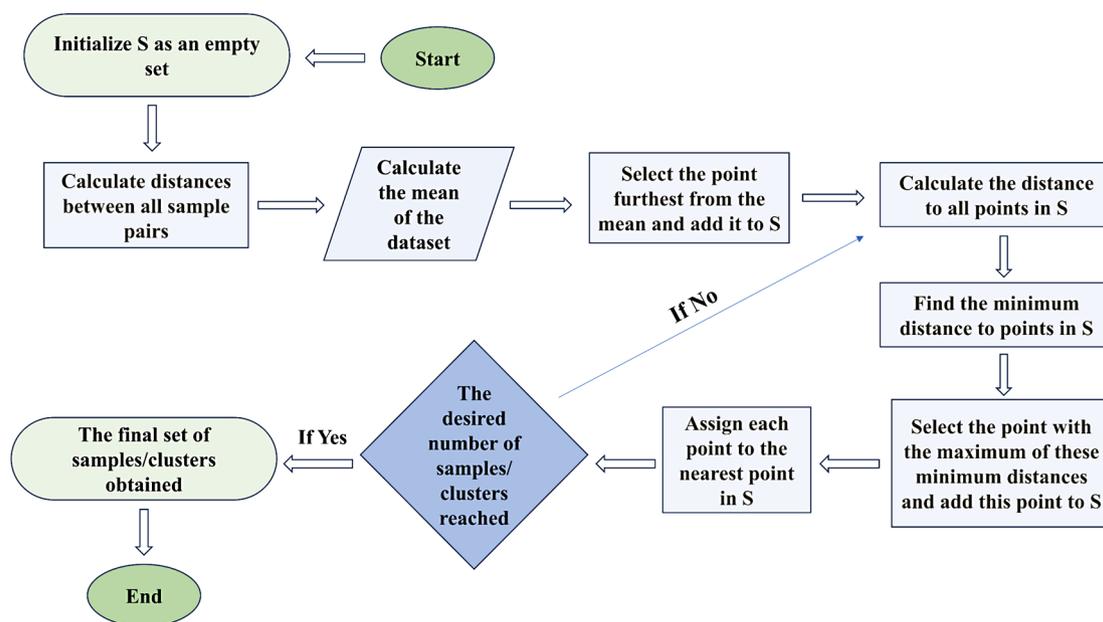
**Fig. 5** Absorption spectra and resolved spectra estimated by MCR-ALS algorithm

- **Initialization:** Begin with an empty set of selected points,  $S = \{\}$ . Calculate the pairwise distances between all points in the dataset  $X$ .
- **Selection of the First Point:** Choose the first point that is furthest from the mean of the dataset and add this point to  $S$ .
- **Iterative Selection:** For each subsequent point, calculate its distance to all points already in  $S$ . Select the point that maximizes the minimum distance to the points in  $S$  and add this point to  $S$ . Repeat this process until the desired number of clusters is reached.
- **Cluster Formation:** Finally, assign each point in the dataset to the nearest centroid based on the Euclidean distance.
- Repeat steps 3 and 4 until the desired number of samples (or clusters) is reached.

We applied the KSC to our dataset of 25 calibration mixtures. The algorithm was used to select 13 mixtures for the validation set, ensuring these samples were well-distributed across the concentration ranges of all four analytes (MOM, OLA, MTS, and DAP). This approach resulted in validation samples that effectively represent the entire calibration space, enhancing the robustness of our chemometric models.

#### **Comparison with traditional clustering methods**

Traditional methods, such as K-means or hierarchical clustering, depend on the initial random selection of centroids followed by iterative refinement. In contrast, the KSC algorithm systematically selects centroids based on distance criteria, ensuring that the clusters are diverse and representative from the outset ensuring a more uniform distribution of samples across the variable space. This is particularly advantageous in analytical chemistry, where capturing the full range of variability is crucial for



**Fig. 6** Flowchart of KSC algorithm steps

developing robust models. While *k*-means and hierarchical clustering aim to minimize within-cluster variance, they may not always select boundary points that are critical for defining the limits of the calibration space. The Kennard-Stone algorithm, by selecting maximally distant points, inherently includes these boundary cases, leading to a more comprehensive representation of the dataset. In our study, applying KSC ensured comprehensive coverage of the analytical feature space, leading to more reliable and reproducible quantification methods.

#### Validation of the chemometric methods

Our proposed chemometric models' validation was carried out through a rigorous evaluation process with the help of an external validation set consisting of thirteen combinations produced by KSC. These validation mixtures, while independent of the model development process, had concentration levels within the calibration range (Table 1), ensuring a comprehensive and unbiased assessment of model performance. We evaluated the models depending on several key metrics, offering a multifaceted view of their predictive capabilities. All models show excellent accuracy, with recovery rates ranging from 98 to 102% for all analytes (Table 3), indicating high reliability in quantitative determinations. The RMSEP, a crucial indicator of model accuracy, showed minimal values across all models, with the MCR-ALS model exhibiting better performance. It achieved RMSEP of 0.00944, 0.42262, 0.01933, and 0.00397 for MOM, OLA, MPS, and DAP, respectively (Table 3). Additionally, we calculated the RRMSEP, which represents prediction accuracy as a percentage of mean analyte concentration. The

MCR-ALS model again showed particularly favorable results, with RRMSEP values of 0.25568%, 0.85179%, 0.75014%, and 0.1306% for MOM, OLA, MPS, and DAP respectively, as shown in (Table S2). To further assess the models' precision, we examined the BCMSEP. All models demonstrated low BCMSEP values, showing a high level of precision in predictions across all components (Table S2). In addition to these metrics, we assessed the overall fit of the models by plotting true concentrations versus estimated ones for all analytes. These plots yielded high coefficients of determination ( $r^2$ ) across the board, demonstrating excellent linearity and predictive performance. The strong linear relationships are visually confirmed in (Fig. S2 to Fig. S5), where all plots exhibit good  $r^2$  values, further validating the accuracy of our models. We also evaluated the sensitivity of our models by calculating LOD and LOQ from net analyte signals, confirming their suitability for quality control analysis. The models exhibited outstanding reproducibility, with intra-day and inter-day assessments showing relative standard deviation (%RSD) values less than 2% (Table S2), demonstrating their precision and robustness under repeated testing conditions. These comprehensive validation results collectively demonstrate the high accuracy, precision, and robustness of our models, particularly the MCR-ALS approach, for the concurrently quantification of MOM, OLA, MPS, and DAP in complex pharmaceutical formulations. The consistent performance across various validation metrics suggests that these models are well-suited for routine quality control applications in pharmaceutical analysis, offering reliable and sensitive quantification

**Table 3** Quantification of analytes in calibration and validation sets using the proposed analytical approaches

	CLS										PCR										PLS										GA-PLS										MCR-ALS									
	MOM		OLO		MTS		DAP		MOM		OLO		MTS		DAP		MOM		OLO		MTS		DAP		MOM		OLO		MTS		DAP		MOM		OLO		MTS		DAP											
	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD	MEAN	SD												
Calibration set	98.84	1.508	99.33	1.480	99.6	1.523	98.57	1.375	99	1.443	99.24	1.427	99.45	1.473	98.64	1.291	98.88	1.258	99.03	0.867	99.18	1.304	99.31	1.347	98.68	1.237	99.03	0.867	99.18	1.304	99.09	0.994	99.15	0.660	98.72	0.994	99.05	1.053	99.13	0.785										
Validation set	99.78	1.175	99.71	1.729	99.13	1.363	100.08	1.173	99.71	1.079	99.79	1.243	99.29	0.873	99.80	1.045	99.88	0.696	99.91	0.384	99.85	0.786	99.51	0.532	99.86	0.656	99.91	0.384	99.72	0.335	99.92	0.188	99.92	0.335	99.93	0.240	99.80	0.178	99.80	0.182										
	0.052	0.052	0.046	0.049	0.035	0.046	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049	0.049									
	0.045	0.045	0.038	0.044	0.039	0.038	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044	0.044						

<sup>a</sup>Root Mean Square Error of calibration<sup>b</sup>Root Mean Square Error of prediction

of both active substances as well as possible genotoxic impurities.

### Statistical analysis

To assess the efficiency of the proposed models, the ANOVA was conducted on the validation data. From the findings, it was determined that there were no statistically significant changes in accuracy amongst the different models ( $p > 0.05$ ). The F-values that were calculated were considerably lower than the critical one, further confirming the comparable accuracy of these models (Table S3). Additionally, a comparative analysis between the suggested chemometric methods and the previously reported method [5] for the quantification of MOM and OLO demonstrated equivalent performance, as demonstrated by the data that was introduced in (Table S3).

### Pharmaceuticals assay

The proposed method was effectively used to the analysis of Ryaltris® nasal spray, demonstrating no significant interference from excipients. To validate the accuracy of the proposed approaches in real pharmaceutical formulations, the standard addition method was used. The findings, summarized in (Table 4), confirm the method's reliability and applicability for routine quality control analysis of the combination product.

### Comparative study

A comprehensive comparison of the five chemometric models was conducted to evaluate their performance in quantifying the target analytes. The results, summarized in (Table S2), demonstrate successful model construction for all methods, with MCR-ALS exhibiting better overall performance. The CLS model, while effective, is constrained by its requirement for thorough understanding of every component of the calibration sample, potentially limiting its accuracy and applicability in complicated matrices. In contrast, PCR and PLS models offer greater flexibility, as they operate on the assumption of a linear relationship between response variables and spectral data. This approach eliminates the need for exhaustive component information, allowing for the resolve of key analytes even in the presence of unknown matrix elements. PLS demonstrated various benefits over PCR, encompassing improved handling of collinear variables, additional efficient utilization of spectral data, enhanced management of non-linear relationships, increased robustness to noise, and superior predictive power. Notably, the incorporation of GA data further enhanced PLS model performance, with GA-PLS models outperforming those constructed utilizing raw data. This improvement is attributed to the optimized predictive and resolution capabilities achieved through the selection of fewer, more relevant data points and variables. MCR-ALS emerged as

**Table 4** Quantification of MOM and OLO in medicinal formulations through the utilization of proposed chemometric analytical approaches, coupled with the implementation of the standard addition methodology

Preparation		%Recovery $\pm$ %RSD <sup>(a)</sup>				
		CLS	PCR	PLS	GA-PLS	MCR-ALS
MOM	Application	100.03 $\pm$ 0.862	100.02 $\pm$ 0.632	100.02 $\pm$ 0.797	100.25 $\pm$ 0.527	100.01 $\pm$ 0.621
	Standard addition	100.23 $\pm$ 0.475	100.38 $\pm$ 0.673	99.04 $\pm$ 0.655	100.85 $\pm$ 0.415	99.14 $\pm$ 0.622
OLO	Application	100.21 $\pm$ 0.824	100.52 $\pm$ 0.824	100.11 $\pm$ 0.792	100.25 $\pm$ 0.651	100.54 $\pm$ 0.522
	Standard addition	99.98 $\pm$ 0.705	99.99 $\pm$ 0.646	99.11 $\pm$ 0.622	99.74 $\pm$ 0.591	100.45 $\pm$ 0.692

<sup>a</sup>Average of three determinations

the top-performing model, showcasing superior results across all evaluation metrics. Its exceptional performance can be attributed to its capacity to use iterative learning to understand system behavior and dynamic data analysis under varied conditions. This approach enables a thorough knowledge of the system's complexity in both mathematics and chemistry. Furthermore, MCR-ALS successfully resolved the authentic spectra of each component, offering the additional benefit of qualitative impurity identification. This capability extends to the potential retrieval of qualitative data regarding unknown interferences, if present. In summary, while all models demonstrated satisfactory performance, the MCR-ALS approach offers the most robust and versatile solution for the concurrently quantification of MOM, OLO, and their impurities in the nasal spray formulation.

#### Comprehensive sustainability assessment

The evaluation of analytical methods' environmental and economic impacts is crucial for understanding their overall sustainability. Sustainability in analytical chemistry encompasses multiple facets, including environmental friendliness, waste reduction, safety, efficiency, and cost-efficacy [37]. Given the complexity of these interrelated factors, a single assessment tool is insufficient to provide a thorough evaluation of an analytical method's sustainability profile [38–40]. Therefore, this research utilized a multi-tool methodology to conduct a comprehensive sustainability assessment, considering various complementary perspectives.

#### NEMI tool

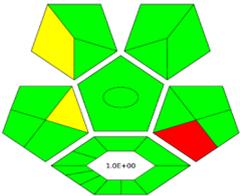
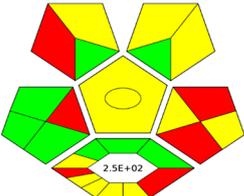
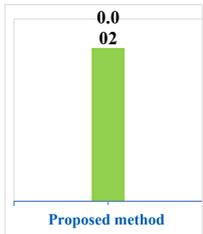
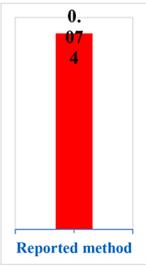
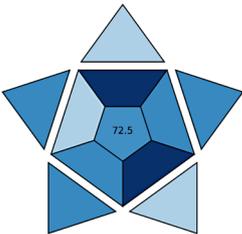
Despite the fact that tools such as GSST and SDAGI aid in selecting greener reagents, evaluating the whole method's environmental impact is crucial. NEMI offers a straightforward, efficient visual evaluation of a method's greenness depending on key parameters such as corrosiveness, toxicity, and waste generation [41]. The NEMI tool utilizes a four-quadrant pictogram (Fig. S6) to represent critical environmental criteria. A method is considered green if it meets the following requirements: (1) None of the chemicals used are categorized by the EPA's Toxic Release Inventory (TRI) as persistent, bio-accumulative, and toxic (PBT). (2) No chemicals are

listed as hazardous waste under the Resource Conservation and Recovery Act (RCRA) U, P, F, D, or TRI lists. (3) The method's pH falls between 2 and 12, which is a non-corrosive range. (4) The total waste generated is less than 50 g. NEMI pictograms were generated for both the reported and proposed methods, as shown in (Table 5). A comparison of these pictograms reveals that the proposed method is significantly greener, with all four quadrants colored green, indicating compliance with all NEMI requirements. The chemicals used in the suggested method were not classified as hazardous or PBT, the pH was within the non-corrosive range, and waste generation was below 50 g. Conversely, the reported method showed blank spaces in the toxic and hazardous classifications due to the use of methanol, which is categorized as both toxic and hazardous under TRI and RCRA criteria. Although NEMI offers a helpful preliminary screening of a method's environmental impact, latest studies have shown limitations in its reliability for comparative evaluations [37]. The tool's simplified pass/fail approach depends on a limited set of requirements and may not fully capture the nuances of method greenness. Therefore, in this study, we employ NEMI as a preliminary screen but supplement it with additional robust, quantitative greenness metrics to allow an additional comprehensive and trustworthy comparison of the suggested approach with current methods. This approach allows for an additional nuanced evaluation of the method's environmental influence, providing a solid foundation for further sustainability assessments using complementary tools.

#### Complex GAPI tool

This tool offers a more comprehensive semi-quantitative assessment of a method's greenness, addressing the limitations of simpler tools like NEMI [42]. Building upon the original GAPI, Complex GAPI incorporates an extra hexagonal field for CHEM21 parameters, facilitating a more nuanced assessment of sustainability across all stages of the analytical process, encompassing the collecting of samples, their preservation, preparation, transportation, analysis, and storage (Fig. S7). This tool employs a color-coded scale (green to yellow to red) to successfully assess ecological impacts at each stage

**Table 5** Comprehensive assessment of environmental sustainability utilizing six advanced assessment tools for both the proposed and existing methods

	The proposed method	The reported method																																																																								
<b>NEMI tool</b>																																																																										
<b>ComplexGAPI tool</b>																																																																										
<b>AGREE tool</b>																																																																										
<b>Carbon footprint (kg CO<sub>2</sub> eq /sample)</b>																																																																										
<b>BAGI tool</b>																																																																										
<b>RGB12 tool</b>	<p><b>Method: Proposed method</b></p> <table border="1"> <tr> <td>R1: Scope of application</td> <td>100.0</td> <td>G1: Toxicity of reagents</td> <td>100.0</td> <td>B1: Cost efficiency</td> <td>90.0</td> </tr> <tr> <td>R2: LOD and LOQ</td> <td>80.0</td> <td>G2: Amount of reagents and waste</td> <td>75.0</td> <td>B2: Time efficiency</td> <td>90.0</td> </tr> <tr> <td>R3: Precision</td> <td>90.0</td> <td>G3: Energy and other media</td> <td>100.0</td> <td>B3: Requirements</td> <td>75.0</td> </tr> <tr> <td>R4: Accuracy</td> <td>100.0</td> <td>G4: Direct impacts</td> <td>100.0</td> <td>B4: Operational simplicity</td> <td>90.0</td> </tr> <tr> <td colspan="2"><b>92.5</b></td> <td colspan="2"><b>93.8</b></td> <td colspan="2"><b>86.3</b></td> </tr> <tr> <td colspan="6" style="text-align: center;"><b>90.8</b></td> </tr> </table>	R1: Scope of application	100.0	G1: Toxicity of reagents	100.0	B1: Cost efficiency	90.0	R2: LOD and LOQ	80.0	G2: Amount of reagents and waste	75.0	B2: Time efficiency	90.0	R3: Precision	90.0	G3: Energy and other media	100.0	B3: Requirements	75.0	R4: Accuracy	100.0	G4: Direct impacts	100.0	B4: Operational simplicity	90.0	<b>92.5</b>		<b>93.8</b>		<b>86.3</b>		<b>90.8</b>						<p><b>Method: TLC-Densitometry</b></p> <table border="1"> <tr> <td>R1: Scope of application</td> <td>75.0</td> <td>G1: Toxicity of reagents</td> <td>75.0</td> <td>B1: Cost efficiency</td> <td>65.0</td> </tr> <tr> <td>R2: LOD and LOQ</td> <td>75.0</td> <td>G2: Amount of reagents and waste</td> <td>60.0</td> <td>B2: Time efficiency</td> <td>65.0</td> </tr> <tr> <td>R3: Precision</td> <td>80.0</td> <td>G3: Energy and other media</td> <td>30.0</td> <td>B3: Requirements</td> <td>72.5</td> </tr> <tr> <td>R4: Accuracy</td> <td>85.0</td> <td>G4: Direct impacts</td> <td>33.3</td> <td>B4: Operational simplicity</td> <td>68.3</td> </tr> <tr> <td colspan="2"><b>78.8</b></td> <td colspan="2"><b>74.6</b></td> <td colspan="2"><b>67.7</b></td> </tr> <tr> <td colspan="6" style="text-align: center;"><b>73.7</b></td> </tr> </table>	R1: Scope of application	75.0	G1: Toxicity of reagents	75.0	B1: Cost efficiency	65.0	R2: LOD and LOQ	75.0	G2: Amount of reagents and waste	60.0	B2: Time efficiency	65.0	R3: Precision	80.0	G3: Energy and other media	30.0	B3: Requirements	72.5	R4: Accuracy	85.0	G4: Direct impacts	33.3	B4: Operational simplicity	68.3	<b>78.8</b>		<b>74.6</b>		<b>67.7</b>		<b>73.7</b>					
R1: Scope of application	100.0	G1: Toxicity of reagents	100.0	B1: Cost efficiency	90.0																																																																					
R2: LOD and LOQ	80.0	G2: Amount of reagents and waste	75.0	B2: Time efficiency	90.0																																																																					
R3: Precision	90.0	G3: Energy and other media	100.0	B3: Requirements	75.0																																																																					
R4: Accuracy	100.0	G4: Direct impacts	100.0	B4: Operational simplicity	90.0																																																																					
<b>92.5</b>		<b>93.8</b>		<b>86.3</b>																																																																						
<b>90.8</b>																																																																										
R1: Scope of application	75.0	G1: Toxicity of reagents	75.0	B1: Cost efficiency	65.0																																																																					
R2: LOD and LOQ	75.0	G2: Amount of reagents and waste	60.0	B2: Time efficiency	65.0																																																																					
R3: Precision	80.0	G3: Energy and other media	30.0	B3: Requirements	72.5																																																																					
R4: Accuracy	85.0	G4: Direct impacts	33.3	B4: Operational simplicity	68.3																																																																					
<b>78.8</b>		<b>74.6</b>		<b>67.7</b>																																																																						
<b>73.7</b>																																																																										

of the method. A key quantitative metric used in Complex GAPI is the Environmental factor (E-factor), which measures the ratio of waste produced to the amount of product obtained. Lower E-factor values indicate larger sustainability because of reduced production of waste. The user-friendly software provided with Complex GAPI facilitates the easy generation of pictograms for the visual representation of results. In our study, Complex GAPI analysis revealed the extraordinary greenness of the developed method, as demonstrated by the predominance of green icons across various stages of the analytical process. The interestingly low E-factor further confirmed the approach's low waste production and favorable environmental profile, as outlined in (Table 5). These results underscore the method's alignment with green chemistry principles and its potential for reducing the environmental footprint of pharmaceutical analysis. However, it is important to note that Complex GAPI mainly emphasizes environmental factors but does not comprehensively cover other key aspects of sustainability, including waste reduction, energy efficiency, and the incorporation of renewable materials. Consequently, while Complex GAPI provides valuable insights into the method's greenness, combining it with extra quantitative tools is suggested for an additional thorough evaluation of overall sustainability.

#### **AGREE tool**

This tool provides a quantitative method for evaluating the environmental friendliness of an analytical technique, in alignment with the 12 guiding principles of green analytical chemistry [43]. One of the primary benefits of AGREE is its adaptability, permitting the customization of parameter weights according to their importance for the specific analytical technique. This capability allows for assessments that are specifically tailored to highlight the most relevant greenness criteria for a particular application. The AGREE method produces a final score within a range of 0 to 1, providing a concise summary of overall method greenness. Additionally, the visual clock pictogram generated by the AGREE software identifies specific areas for potential improvement, facilitating targeted optimization of the method's environmental performance. In this study, AGREE analysis was conducted for both the proposed method and the previously reported method. Detailed results of these assessments are provided in the supplementary material. The suggested method shows exceptional greenness, attaining an elevated AGREE score of 0.90 as shown in (Fig. S8). This score indicates better performance in terms of adherence to green chemistry principles compared to existing methods. The graphical representations shown in (Table 5) clearly highlight the outstanding greenness profile of the proposed method, emphasizing its environmentally

sustainable features and validating its conformity with eco-friendly analytical standards. These results underscore the method's potential to significantly reduce the environmental impact of pharmaceutical analysis. Nevertheless, it is significant to note that AGREE mainly concentrates on environmental factors based on the principles of green chemistry. However, other important sustainability aspects, such as safety, analytical performance, and cost-efficiency, are not specifically covered by this tool. Therefore, while AGREE provides valuable insights into a method's greenness, it is suggested to couple this approach with additional tools that evaluate other aspects of sustainability for a more thorough assessment.

#### **Carbon footprint analysis**

This analysis offers a valuable quantitative metric for evaluating how analytical techniques affect the environment in terms of kilos of CO<sub>2</sub> equivalent (kg CO<sub>2</sub> eq), which is a measure of greenhouse gas emissions [44–48]. Unlike qualitative or semi-quantitative assessments, this analysis allows direct comparison of the environmental effects of various analytical techniques on operations and lifecycles on a standardized, quantitative scale. This makes it an excellent complement to other greenness evaluation tools like NEMI, Complex GAPI, and AGREE, as it covers additional aspects of environmental impact not captured by these techniques. In this study, we computed the carbon footprint utilizing the following standard equation:

$$\text{Carbon footprint (kg CO}_2 \text{ eq)} = \sum \text{Instrument Power (kW)} \\ \cdot \text{Analysis time (h)} \\ \cdot \text{Emission factor (kg CO}_2 \text{/kWh)}$$

Our proposed method shows a significantly lower carbon footprint of 0.021 kg CO<sub>2</sub> eq per sample analyzed, as shown in (Table 5). In contrast, the previously reported method exhibited a substantially higher environmental impact, with a carbon footprint of 0.074 kg CO<sub>2</sub> equivalent per examined sample. The decreased carbon footprint of our method can be ascribed to multiple factors including lower electricity usage due to shorter analysis times, elimination of the derivatization step, and replacement of hazardous solvents, leading to reduced emissions associated with transportation. Overall, the proposed method achieves a remarkable 59% reduction in CO<sub>2</sub> emissions compared to the reported method, further confirming its favorable environmental profile.

#### **BAGI tool**

The Blue Applicability Grade Index (BAGI) offers a novel approach to evaluating analytical methods by quantifying their “blueness,” which reflects their practical applicability and real-world effectiveness [49]. Unlike

traditional metrics that primarily focus on environmental aspects (“greenness”), BAGI provides a comprehensive assessment of a method’s suitability based on crucial operational criteria. BAGI evaluates ten key factors that influence a method’s practicality: sample productivity, the type of analysis, the number of analytes used, the equipment used, the requirements for sample preparation, the number of samples analyzed in an hour, the materials and reagents needed, the necessity of pre-concentration steps, the degree of automation, and the amount of sample needed. Each factor is measured from 1 (the least) to 10 (the highest) on a scale. The last BAGI score is calculated as the individual scores’ geometric mean, with higher scores demonstrating greater appropriateness, usefulness, and suitability for the intended use. In our study, the suggested approach attained an impressive BAGI score of 90, demonstrating exceptional real-world applicability. As shown in (Table 5), this high BAGI score underscores the method’s numerous benefits in terms of reduced hazards, cost and time effectiveness, and overall usability in practical settings. However, it’s important to note that while BAGI provides useful information about the method’s practical applicability, it doesn’t provide a comprehensive assessment of overall sustainability. To accomplish a more holistic assessment that considers ecological friendliness, analytical performance, and practicality, we complemented the BAGI assessment with the RGB12 tool.

#### **RGB12 tool**

The RGB12 algorithm, presented by Paweł-Nowak et al. in 2021 [50], offers a comprehensive and easy-to-use quantitative technique for assessing the “whiteness” of methods. This innovative approach evaluates methods based on 12 WAC principles, providing a holistic view of sustainability that encompasses economic, analytical, and environmental factors [51]. The RGB12 algorithm comprises 12 distinct algorithms divided into three subgroups. Green subgroup (G1-G4): emphasizes key GAC considerations, includes waste reduction, toxicity, energy conservation, and reagent minimization, and impacts on animals, humans, and genetic changes. Red subgroup (R1-R4): discusses validation criteria like accuracy, precision, LOD, LOQ, and application scope. Blue subgroup (B1-B4): Evaluates practical and economic factors, including cost-efficiency and time savings. The final “whiteness” score is calculated by summing the method’s performance across all three-color domains, reflecting its adherence to WAC principles. In our study, the proposed method achieved a score of 90.8, as shown in (Table 5). This high score illustrates the method’s multifaceted advantages in terms of, practical applicability, and analytical effectiveness. By integrating the RGB12 assessment with other sustainability measurements, we

have conducted an entire and reliable assessment of the method’s overall sustainability. This systems-oriented approach, utilizing multiple complementary tools, overcomes the limitations of single-metric assessments and provides a more objective and thorough analysis of the method’s sustainability profile.

#### **Conclusion**

In this research, we successfully developed and validated an innovative UV-visible spectrophotometric method combined with advanced chemometric techniques for employing UV-visible spectroscopy to concurrently determine MOM, OLO, and the genotoxic impurities MTS and DAP. The implementation of the KSC for constructing the validation set ensured a comprehensive and unbiased assessment of the method’s performance throughout the whole concentration range, overcoming the limitations associated with traditional random data splitting techniques. Our rigorous assessment of 5 models, including PLS, PCR, CLS, GA-PLS, and MCR-ALS demonstrated the robustness, sensitivity, accuracy, and precision of the proposed approach. Each model achieved excellent recovery rates, low bias-corrected prediction errors, and adequate limits of detection, confirming the method’s reliability and applicability in pharmaceutical quality control with MCR-ALS becoming the best performer by both quantitatively and qualitatively resolving spectra of pure components. In accordance with the concepts of GAC and WAC, the method employed environmentally friendly solvents and underwent a comprehensive sustainability assessment using six state-of-the-art tools: NEMI, Complex GAPI, AGREE, Carbon Footprint Analysis, RGB12 Metrics, and the BAGI. The favorable outcomes across these metrics highlighted the method’s eco-friendliness, affordability and practicality, contributing significantly to sustainable development goals in the pharmaceutical industry.

#### **Supplementary Information**

The online version contains supplementary material available at <https://doi.org/10.1186/s13065-025-01391-8>.

Supplementary Material 1

Supplementary Material 2

#### **Acknowledgements**

The authors extend their appreciation to Taif University, Saudi Arabia, for supporting this work through the project number (TU-DSPP-2024-49).

#### **Author contributions**

A.E.F.A. Methodology, design of the work, investigation, writing-original draft, and writing review and editing. M.G. Supervision, and investigation. I.A.N. Writing review and editing. M.K.H. Writing-original draft, and visualization. B.A.M.S., M.M.G., and M.M.A.M. Interpretation of data and figures preparation. Y.A.S. Supervised analysis procedures and carried out sample preparation. All the authors read, reviewed, and approved the manuscript.

### Funding

This research was funded by Taif University, Saudi Arabia, Project No. (TU-DSPP-2024-49).

### Data availability

Data was collected using a spectrophotometer and software. The corresponding author will provide the datasets created and/or analyzed during the current study upon reasonable request.

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

### Author details

<sup>1</sup>Faculty of Pharmacy, Analytical Chemistry Department, October 6 University, October 6 City, Giza 12585, Egypt

<sup>2</sup>Pharmaceutical Analytical Chemistry Department, Faculty of Pharmacy, Beni-Suef University, Alshaheed Shehata Ahmad Hegazy St., Beni-Suef, Egypt

<sup>3</sup>Department of Pharmaceutical Chemistry, College of Pharmacy, Taif University, Taif, Saudi Arabia

<sup>4</sup>College of Pharmacy, Al-Mustaqbal University, Babylon 51001, Iraq

<sup>5</sup>Department of Pharmacy Practice, College of Pharmacy, AlMaarefa University, Ad Diriyah, Riyadh 13713, Saudi Arabia

<sup>6</sup>Chemistry Department, Faculty of Science, University of Al-Jufra, P.O. Box 61602, Al-Jufra, Libya

<sup>7</sup>Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Sinai University, Kantara Branch, Ismailia 341636, Egypt

Received: 14 May 2024 / Accepted: 14 January 2025

Published online: 15 April 2025

### References

- Prajapati P, Salunkhe M, Pulusu VS, Shah S. Integrated Approach of White Analytical Chemistry and Analytical Quality by Design to multipurpose RP-HPLC method for synchronous estimation of multiple fixed-dose combinations of paracetamol. *Chem Afr*. 2024;7:1353–71.
- Prajapati P, Rana B, Pulusu VS, Mishra A. Multipurpose RP-HPLC method for simultaneous estimation of fixed-dose combinations of anti-diabetic drugs: integrating Green, Economical, and robust approaches with design of experiments and White Analytical Chemistry. *Chem Afr*. 2024;7:1385–400.
- Prajapati P, Patel K, Patel A, Shakar Pulusu V, Haque A, Ahmad S, et al. Integrated approach of white analytical chemistry and design of experiments to microwave-assisted sensitive and eco-friendly spectrofluorimetric estimation of mirabegron using 4-chloro-7-nitrobenzofuran as biosensing fluorescent probe. *Spectrochim Acta Mol Biomol Spectrosc*. 2024;319:124521.
- Prajapati P, Patel M, Kansara Y, Shah P, Pulusu VS, Shah S. Green LC-MS/MS method for in-vivo pharmacokinetics of mirabegron-encapsulated nanostructured lipid carriers in rat plasma: integrating white analytical chemistry and analytical quality by design approach. *Sustain Chem Pharm*. 2024;39:101523.
- Patel B, Patel S. A specific high-performance thin-layer chromatography method validated for estimation of mometasone furoate and olopatadine hydrochloride. *Sep Sci Plus*. 2023;6:2300032.
- Prajapati P, Pulusu VS, Shah S. Principles of White Analytical Chemistry and Design of experiments to Development of Stability-Indicating Chromatographic Method for the simultaneous estimation of Thiocolchicoside and Lornoxicam. *J AOAC Int*. 2023;106:1654–65.
- Prajapati P, Rana B, Pulusu VS, Shah S. Method operable design region for robust RP-HPLC analysis of pioglitazone hydrochloride and teneligliptin hydrobromide hydrate: incorporating hybrid principles of white analytical chemistry and design of experiments. *Futur J Pharm Sci*. 2023;9:93–102.
- El Hamd MA, Soltan OM, Abdelrahman KS, Fouad A, Saleh SF, Obaydo RH, et al. Roth's switch-on fluorometric probe for green tracking and quantifying of 1,4-dihydropyridine medication: evaluation of greenness, whiteness, and blueness. *Sustain Chem Pharm*. 2023;36:101294.
- Elbordiny HS, Elonsy SM, Daabees HG, Belal TS. Design of trio-colored validated HPLC method for synchronized multianalyte quantitation of four top selling antihyperlipidemic drugs in different fixed-dose combined tablets. *Green Anal Chem*. 2024;8:100100.
- Attimarad M, Chohan MS, Katharigatta Narayanaswamy V, Nair AB, Sreeharsha N, Shafi S, et al. Mathematically processed UV Spectroscopic Method for quantification of Chlorthalidone and Azelnidipine in Bulk and Formulation: evaluation of greenness and whiteness. *J Spectrosc*. 2022;9:1–13.
- Elsonbaty A, Serag A, Abdulwahab S, Hassan WS, Eissa MS. Analysis of quinary therapy targeting multiple cardiovascular diseases using UV spectrophotometry and chemometric tools. *Spectrochim Acta Mol Biomol Spectrosc*. 2020;238:118415.
- Prajapati P, Shahi A, Acharya A, Shah S. Chemometry and Green Chemistry-Based Chromatographic Analysis of Azilsartan Medoxomil, Cilindipine and Chlorthalidone in Human plasma using Analytical Quality by Design Approach. *J Chromatogr Sci*. 2024;62:201–12.
- Katamesh NS, Abbas AEF, Halim MK, Abdel-Lateef MA, Mahmoud SA. Green micellar UPLC and complementary eco-friendly spectroscopic techniques for simultaneous analysis of anti-COVID drugs: a comprehensive evaluation of greenness, blueness, and whiteness. *BMC Chem*. 2024;18:130.
- Attia KAM, El-Olemy A, Eid SM, Abbas AEF. A Green-and-White Integrative Analytical Strategy combining Univariate and Chemometric techniques for quantifying recently approved Multi-drug Eye Solution and potentially Cancer-causing impurities: application to the aqueous humor. *J AOAC Int*. 2024;107:146–57.
- Li T, Fong S, Wu Y, Tallon-Ballesteros AJ. Kennard-stone balance algorithm for time-series big data stream mining. In: *International Conference on Data Mining Workshops*, vol. 24; 2020. pp. 851–858.
- Halim MK, Badran OM, Abbas AEF. Sustainable chemometric methods boosted by latin hypercube technique for quantifying the recently FDA-approved combination of bupivacaine and meloxicam in the presence of bupivacaine carcinogenic impurity: Comprehensive greenness, blueness, and whiteness assessments. *Microchem J*. 2024;200:110276.
- Halim MK, Badran OM, Abbas AEF. Greenness, blueness and whiteness evaluated-chemometric approach enhanced by latin hypercube technique for the analysis of lidocaine, diclofenac and carcinogenic impurity 2,6-dimethylaniline. *Sustain Chem Pharm*. 2024;38:101463.
- Ferreira F, Resina L, Esteves T, Ferreira FC. Comparison and combination of Organic Solvent Nanofiltration and Adsorption processes: a Mathematical Approach for Mitigation of active Pharmaceutical ingredient losses during Genotoxin Removal. *Membr (Basel)*. 2020;10:73.
- Székely Gy, Henriques B, Gil M, Ramos A, Alvarez C. Design of experiments as a tool for LC–MS/MS method development for the trace analysis of the potentially genotoxic 4-dimethylaminopyridine impurity in glucocorticoids. *J Pharm Biomed Anal*. 2012;70:251–8.
- Brereton RG, Jansen J, Lopes J, Marini F, Pomerantsev A, Rodionova O, et al. Chemometrics in analytical chemistry—part II: modeling, validation, and applications. *Anal Bioanal Chem*. 2018;410:6691–704.
- Mostafa A, Shaaban H. Chemometric assisted UV-Spectrophotometric methods using multivariate curve resolution alternating least squares and partial least squares regression for determination of Beta-antagonists in formulated products: evaluation of the ecological impact. *Molecules*. 2022;28:328.
- Attia KAM, El-Olemy A, Abbas AEF, Eid SM. A sustainable data processing approach using ultraviolet-spectroscopy as a powerful spectral resolution tool for simultaneously estimating newly approved eye solution in the presence of extremely carcinogenic impurity aided with various greenness and whiteness assessment perspectives: application to aqueous humor. *J Chem Res*. 2023;47:1–9.
- Larsen C, Lundberg P, Tang S, Råföls-Ribé J, Sandström A, Mattias Lindh E, et al. A tool for identifying green solvents for printed electronics. *Nat Commun*. 2021;12:4510.
- Kayali Z, Obaydo RH, Alhaj Sakur A. Spider diagram and sustainability evaluation of UV-methods strategy for quantification of aspirin and sildenafil citrate in the presence of salicylic acid in their bulk and formulation. *Heliyon*. 2023;9:1–7.
- Shen Y, Lo C, Nagaraj DR, Farinato R, Essenfeld A, Somasundaran P. Development of Greenness Index as an evaluation tool to assess reagents: evaluation based on SDS (Safety Data sheet) information. *Min Eng*. 2016;94:1–9.

26. Ferraro MCF, Castellano PM, Kaufman TS. Chemometric determination of amiloride hydrochloride, atenolol, hydrochlorothiazide and timolol maleate in synthetic mixtures and pharmaceutical formulations. *J Pharm Biomed Anal.* 2004;34:305–14.
27. Saad AS, Elzanfaly ES, Halim MK, Kelani KM. Comparing the predictability of different chemometric models over UV-spectral data of isoxsuprine and its toxic photothermal degradation products. *Spectrochim Acta Mol Biomol Spectrosc.* 2019;219:444–9.
28. Kelani KM, Shalaby AA, Elmaamy MY. Spectrophotometric and chemometric methods for simultaneous determination of two anti-hypertensive drugs in their combined dosage form. *Pharm Anal Acta.* 2015;6:10–9.
29. Abdelazim AH, Shahin M, Abu-khadra AS. Application of different chemometric assisted models for spectrophotometric quantitative analysis of velpatasvir and sofosbuvir. *Spectrochim Acta Mol Biomol Spectrosc.* 2021;252:119492.
30. Abdelazim AH, Shahin M, Abu-khadra AS. Application of different chemometric assisted models for spectrophotometric quantitative analysis of velpatasvir and sofosbuvir. *Spectrochim Acta Mol Biomol Spectrosc.* 2021;252:119540.
31. Prajapati P, Shah S, Prajapati B, Shakar V, Jariwala H, Salunkhe M. Principal component analysis and DoE-driven Green Analytical Chemistry Concept to Liquid Chromatographic Method for Estimation of co-formulated anti-hypertensive drugs. *J AOAC Int.* 2023;3:1–10.
32. Prajapati P, Radadiya K, Shah S. Principal Component Analysis and DoE-Based AQbD Approach to Multipurpose HPTLC Method for Synchronous Estimation of multiple FDCs of Metformin HCl, Repaglinide, Glibenclamide and Pioglitazone HCl. *J Chromatogr Sci.* 2024;62:108–19.
33. Rahman MAA, Elghobashy MR, Zaazaa HE, El-Mosallamy SS. Novel analytical method based on chemometric models applied to UV-Vis spectrophotometric data for simultaneous determination of Etoricoxib and Paracetamol in presence of paracetamol impurities. *BMC Chem.* 2023;17:176.
34. Farid JF, Mostafa NM, Fayed YM, Essam HM, ElTanany BM. Chemometric quality assessment of paracetamol and phenylephrine hydrochloride with paracetamol impurities; comparative UV-spectrophotometric implementation of four predictive models. *Spectrochim Acta Mol Biomol Spectrosc.* 2022;265:1–9.
35. Hassan SA, Nashat NW, Elghobashy MR, Abbas SS, Moustafa AA. Advanced chemometric methods as powerful tools for impurity profiling of drug substances and drug products: application on bisoprolol and perindopril binary mixture. *Spectrochim Acta Mol Biomol Spectrosc.* 2022;267:11–20.
36. Kennard RW, Stone LA. Computer aided design of experiments. *Technometrics.* 1969;11:137–48.
37. Eid SM, Attia KAM, El-Olemy A, Emad F, Abbas A, Abdelshafi NA. An innovative nanoparticle-modified carbon paste sensor for ultrasensitive detection of lignocaine and its extremely carcinogenic metabolite residues in bovine food samples: application of NEMI, ESA, AGREE, ComplexGAP1, and RGB12 algorithms. *Food Chem.* 2023;426:136579.
38. Mahmoud SA, Abbas AEF, Katamesh NS. Greenness, whiteness, and blueness assessment with spider chart solvents evaluation of HPTLC-densitometric method for quantifying a triple combination anti-helicobacter pylori therapy. *Sustain Chem Pharm.* 2024;37:101412.
39. Attia KAM, El-Olemy A, Serag A, Emad F, Abbas A, Eid SM. Environmentally sustainable DRS-FTIR probe assisted by chemometric tools for quality control analysis of cinnarizine and piracetam having diverged concentration ranges: validation, greenness, and whiteness studies. *Spectrochim Acta Mol Biomol Spectrosc.* 2023;302:123161.
40. Attia KAM, Abbas AEF, El-Olemy A, Abdelshafi NA, Eid SM. A recycled-material-based Electrochemical Eco-sensor for Sensitive Detection of Antischistosomal Drug residues in bovine-derived food samples. *Biochip J.* 2024;18:257–74.
41. Keith LH, Gron LU, Young JL. Green analytical methodologies. *Chem Rev.* 2007;107:2695–708.
42. Plotka-Wasyłka J, Wojnowski W. Complementary green analytical procedure index (ComplexGAP1) and software. *Green Chem.* 2021;23:8657–65.
43. Gałuszka A, Migaszewski Z, Namieśnik J. The 12 principles of green analytical chemistry and the significance mnemonic of green analytical practices. *TRAC - Trends Anal Chem.* 2013;50:78–84.
44. Pla-Tolós J, Serra-Mora P, Hakobyan L, Molins-Legua C, Moliner-Martinez Y, Campins-Falcó P. A sustainable on-line CapLC method for quantifying antifouling agents like irgarol-1051 and diuron in water samples: estimation of the carbon footprint. *Sci Total Environ.* 2016;570:611–8.
45. Jiménez J, De la Cruz L, Carballo J, Doménech A. Enfoques metodológicos para el cálculo de la Huella de Carbono. *Observatorio de la sostenibilidad en España* 2011, 2:1–9.
46. Ballester-Caudet A, Campins-Falcó P, Pérez B, Sancho R, Lorente M, Sastre G, et al. A new tool for evaluating and/or selecting analytical methods: summarizing the information in a hexagon. *TRAC Trends Anal Chem.* 2019;118:538–47.
47. Katamesh NS, Abbas AEF, Mahmoud SA. Four chemometric models enhanced by latin hypercube sampling design for quantification of anti-COVID drugs: sustainability profiling through multiple greenness, carbon footprint, blueness, and whiteness metrics. *BMC Chem.* 2024;18:54.
48. El-Masry AA, Abbas AEF, Salem YA. A dual methodology employing ion-pair chromatography and built-in UV spectrophotometry for quantifying recently approved combination of mometasone and indacaterol in a novel combined metered dose inhaler: assessing the greenness, carbon footprint, blueness, and whiteness. *BMC Chem.* 2024;18:54129.
49. Manousi N, Wojnowski W, Plotka-Wasyłka J, Samanidou V. Blue applicability grade index (BAGI) and software: a new tool for the evaluation of method practicality. *Green Chem.* 2023;25:7598–604.
50. Nowak PM, Więtecha-Posłuszny R, Pawliszyn J. White Analytical Chemistry: an approach to reconcile the principles of Green Analytical Chemistry and functionality. *TRAC Trends Anal Chem.* 2021;138:116223.
51. Nowak PM, Kościelniak P. What color is your method? Adaptation of the RGB additive color model to analytical method evaluation. *Anal Chem.* 2019;91:10343–52.

## Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.