CODEN [USA]: IAJPBB ISSN : 2349-7750

INDO AMERICAN JOURNAL OF
PHARMACEUTICAL SCIENCES
SJIF Impact Factor: 7.187
https://doi.org/10.5281/zenodo.24564641

Available online at: http://www.iajps.com Review Article

CHEMOMETRICS-ASSISTED SPECTROSCOPIC METHODS FOR SIMULTANEOUS ESTIMATION OF DRUGS IN MULTICOMPONENT PHARMACEUTICAL FORMULATIONS

Shivratna Rajesh Sathe[1*], Pavan Namdev Khurd[2], Satish Bhagwan Tarange[3], Prof. Pushpanjali Shinde[4]

[1*, 2, 3] Shri Ganpati Institute of Pharmaceutical Sciences and Research, Tembhurni, Solapur, Maharashtra – 413211

[4] Assistant Professor, Shri Ganpati Institute of Pharmaceutical Sciences and Research, Tembhurni, Solapur, Maharashtra – 413211

Abstract:

Fixed-dose combination medicines place several active pharmaceutical ingredients (APIs) in a single dosage form, and their absorption spectra frequently overlap so strongly that classical single-wavelength UV–visible spectrophotometry cannot quantify the components separately. Chemometrics supplies the mathematical tools needed to extract each analyte's contribution from the entire recorded spectrum, which makes simultaneous estimation possible without prior chromatographic separation. This review surveys chemometrics-assisted spectroscopic methods for the simultaneous estimation of drugs in multicomponent pharmaceutical formulations, with emphasis on UV–visible spectrophotometry. It explains the origin of spectral overlap and the limits of derivative, ratio-spectra and related classical approaches; summarises spectral pre-processing and variable-selection tools; and compares the principal calibration algorithms (classical least squares, principal component regression, partial least squares, multivariate curve resolution–alternating least squares, artificial neural networks and support vector regression) in terms of assumptions, strengths and weaknesses. The design of calibration and validation sets, including multilevel multifactor, Kennard–Stone and Latin hypercube strategies, is discussed together with the figures of merit used to judge models and the validation expectations set out in ICH Q2(R2) and Q14. A survey of ten representative studies published between 2010 and 2024, covering mixtures of two to five analytes, indicates that partial least squares remains the practical default, that MCR-ALS, nonlinear models and variable selection add value for minor components and severely overlapped systems, and that mean recoveries close to 100% are reported for synthetic mixtures and marketed products. Greenness, blueness and whiteness metrics increasingly accompany these procedures and position them as sustainable alternatives to chromatography. Open challenges include calibration transfer, excipient and degradant interference, minor-component quantification and lifecycle management of multivariate procedures. Machine-learning models, explainable chemometrics and harmonised reporting are the most promising future directions.

Keywords: *chemometrics; UV–visible spectrophotometry; partial least squares; multivariate curve resolution; multicomponent formulations; simultaneous estimation; green analytical chemistry; ICH Q2(R2)*

Corresponding author:
Shivratna Rajesh Sathe,
Shri Ganpati Institute of Pharmaceutical Sciences
and Research, Tembhurni, Solapur, Maharashtra – 413211
Email: shivsathe05@gmail.com

QR CODE

*Please cite this article in press **Shivratna Rajesh Sathe** et al., **Chemometrics-Assisted Spectroscopic Methods For Simultaneous Estimation Of Drugs In Multicomponent Pharmaceutical Formulations.,**. Indo Am. J. P. Sci, 2026; 13(11).*

INTRODUCTION:

Medicines that combine two or more active ingredients in one tablet, capsule or solution have become a routine part of therapy. Fixed-dose combinations (FDCs) simplify regimens for hypertension, pain, colds, respiratory disease, infection and many chronic conditions, and they can improve adherence and broaden therapeutic effect. For the quality control laboratory, however, every added ingredient multiplies the analytical burden: each active pharmaceutical ingredient (API) must be identified and assayed against its label claim, usually in the presence of excipients, related substances and sometimes degradation products. Releasing a batch therefore depends on assays that are fast, selective and economical and that can quantify several components in a single measurement.

High-performance liquid chromatography (HPLC) is the established reference technique for such work because it physically separates the analytes before detection. Its strengths are well known, but so are its costs: expensive columns and instruments, organic solvents that must be purchased, handled and disposed of, long equilibration and run times, and trained operators. These costs are difficult to bear in small manufacturing units, hospital pharmacies, teaching laboratories and regulatory laboratories in resource-limited settings. UV–visible spectrophotometry offers an attractive alternative. Absorption in the ultraviolet region follows the Beer–Lambert law, the instrumentation is inexpensive and widely available, a measurement takes seconds, and aqueous or alcoholic solvents are often sufficient [1]. The difficulty is selectivity. Most drug molecules absorb through the same few chromophores (aromatic rings, conjugated systems and heteroatom lone pairs), so the broad, featureless bands of different APIs superimpose, and the absorbance at any single wavelength reports the sum of several contributions.

Chemometrics addresses this difficulty computationally. The term was introduced to describe the use of mathematical and statistical methods to design optimal measurements and to extract the maximum chemical information from chemical data [2,3]. For multicomponent spectrophotometry the idea is straightforward. Instead of choosing one wavelength per analyte, the analyst records the full spectrum of a set of carefully designed calibration mixtures and lets a multivariate model learn how spectral shape relates to concentration. When the model is applied to the spectrum of an unknown sample, it returns the concentration of every modelled component at once. No separation or derivatisation is needed beyond dissolution, extraction and dilution.

Interest in this strategy has grown steadily, and several reviews already cover parts of the field. Nurani and co-workers surveyed chemometrics-assisted UV–Vis spectrophotometry for the assay of APIs, impurities and degradation products in dosage forms [4]. Cerdà and colleagues traced the evolution from mono- to multicomponent UV–Vis and fluorimetric methods [5]. Ríos-Reina and Azcarate reviewed how chemometrics has revived UV–Vis spectroscopy as an analytical sensor, including pre-processing, modelling and non-targeted applications [6]. Mazivila and Olivieri examined chemometrics coupled to vibrational spectroscopy for solid pharmaceutical products [7]. The present review complements these works in three ways. First, it concentrates on the practical problem faced by a pharmaceutical analyst: the simultaneous estimation of several APIs in a marketed multicomponent formulation. Second, it links model building explicitly to experimental design, validation metrics and regulatory expectations, because a model that fits well but is validated poorly is not fit for release testing. Third, it considers the growing role of green, blue and white analytical chemistry metrics, which now accompany most recent method papers.

The regulatory context deserves emphasis. The revised ICH Q2(R2) guideline on the validation of analytical procedures, adopted in November 2023, covers the analytical use of spectroscopic data and contains a section on multivariate analytical procedures, while ICH Q14 on analytical procedure development includes a dedicated chapter on the development of multivariate procedures and on model lifecycle management [8,9]. These documents make it timely to review how chemometric spectroscopic methods are developed and validated. Sustainability pressures point in the same direction. The twelve principles of green analytical chemistry encourage direct analysis, minimal sample treatment, reduced solvent consumption and lower energy demand [10], all of which are characteristics of multivariate UV–Vis procedures.

Spectral overlap and the case for chemometrics

UV absorption arises from electronic transitions in chromophores such as aromatic rings, conjugated carbonyl groups and heterocycles. In solution, vibrational and solvent interactions broaden these transitions into bands that are tens of nanometres wide, so the spectra of different drugs often share the same 200–320 nm window and differ mainly in subtle shape rather than in the position of well-separated peaks [1]. Combination products make matters worse. Component spectra may be similar in shape, and the formulated mass ratios can be extreme. A paracetamol–sodium ascorbate–chlorpheniramine product, for example, contains the three actives at roughly 150:140:1 by mass, so the minor component contributes almost no distinct signal against the other two (Table 2, entry 3).

Because absorbance is additive when the components do not interact, the recorded spectrum is the sum of the individual spectra weighted by concentration. Figure 1 illustrates the consequence with simulated bands: each analyte has a clear maximum on its own, yet their sum forms one broad envelope in which no analyte can be read directly.

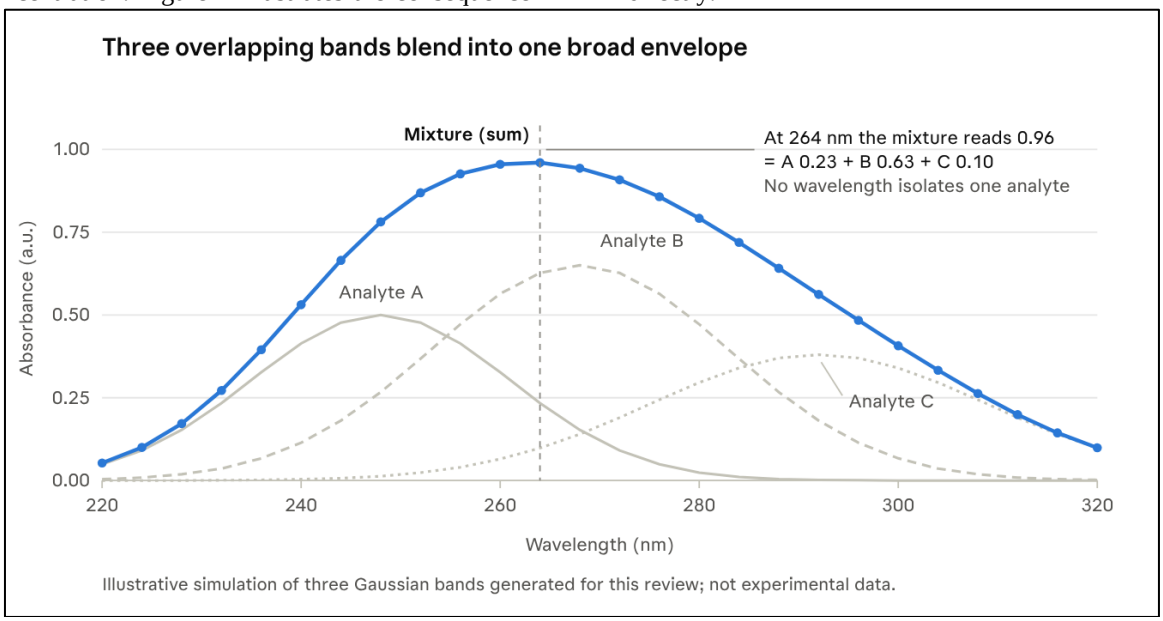

Figure 1. Simulated absorption spectra of three hypothetical analytes (A, B and C) and of their mixture. Individually distinct bands merge into a single broad envelope, so a single-wavelength reading reports the sum of all three contributions.

Classical resolution strategies and their limits

Before multivariate calibration became routine, analysts relied on mathematical manipulations that exploit one feature of each analyte. Simultaneous equations (Vierordt's method) solve for each component from absorbances at as many wavelengths as there are analytes, and they work only if every analyte has at least one wavelength at which its absorptivity differs markedly from that of the others. Derivative spectrophotometry converts broad bands into sharper features and can create zero-crossing points at which one component does not contribute, so that another can be read; the technique has been reviewed extensively [11,12]. Ratio-spectra methods divide the mixture spectrum by the spectrum of one component, so that this component's contribution becomes a constant that can be removed by subtraction or by differentiation of the ratio spectrum [13]. Combinations of these ideas, such as ratio subtraction followed by third-derivative measurement, can resolve ternary mixtures. AlSalem and co-workers used exactly this combination for hydroxyzine, ephedrine and theophylline, optimising the divisor concentration, derivative interval and scaling factor empirically before also applying PCR and PLS to the same system [14].

When chemometrics is justified

A multivariate approach is worth the extra effort when at least two analytes overlap over most of the usable spectral range, when reliable calibration mixtures can be prepared from reference standards, when the formulation matrix is reasonably well characterised, and when the precision required for assay is compatible with spectrophotometric noise. It is less suitable when an analyte is present far below the quantification limit relative to the other components, in which case spiking or standard addition may be needed, or when unknown, strongly absorbing degradants are expected and cannot be included in the calibration design.

Spectral pre-processing and variable selection

The quality of a chemometric model depends as much on the data entering it as on the algorithm applied. For UV–Vis data, which are smooth and low-dimensional compared with near-infrared or Raman spectra, pre-processing is usually light. Nevertheless, three decisions strongly influence the outcome: how the spectra are corrected, which wavelength range is retained, and whether individual wavelengths are selected.

Instrumental and sampling choices

The usable range is set by the solvent cut-off and the detector response. Water and methanol permit work from about 210–220 nm, and many studies therefore begin at 200–230 nm, while the upper limit is placed where the absorbance of all analytes has fallen to the noise level. Spectral data intervals matter as well. De Luca and co-workers found that reducing the interval to 0.5 nm improved the prediction of the minor component in a ternary and a quaternary mixture, because more points describe the weak features that distinguish a low-level analyte [17]. Using a very fine interval does not always help, because neighbouring absorbances are

highly correlated and add little independent information while increasing the risk of overfitting.

Smoothing, derivatives and scatter correction

The Savitzky–Golay filter fits a low-order polynomial within a moving window and can be used both for smoothing and for calculating first or second derivatives [18]. In multivariate calibration, derivatisation can remove constant or sloping baselines and sharpen overlapping bands, but it also amplifies noise, so the window width and polynomial order should be optimised against cross-validation error and not chosen arbitrarily. Standard normal variate (SNV) transformation and multiplicative scatter correction (MSC) compensate for additive and multiplicative scattering effects [19,20]. These are mainly relevant to turbid solutions, suspensions and solid-state or diffuse reflectance measurements, and are seldom needed for clear solutions measured in a cuvette. A systematic comparison of pre-processing options in spectroscopy emphasises that the best treatment is data-dependent and that it should be selected by validation, not by habit [21].

Selecting wavelength regions and variables

Including every recorded wavelength is simple but not always optimal, since regions with only noise or with strongly correlated, uninformative signals degrade the model. Four families of selection strategies appear in the pharmaceutical literature.

- **Manual range selection.** The analyst discards regions of solvent absorption, high noise or saturation and keeps the range where the analytes absorb. This is transparent and reproducible and remains the most common approach.
- **Interval methods.** Interval PLS divides the spectrum into equal sub-intervals, builds a local PLS model in each and retains the intervals giving the lowest cross-validated error [22]. Synergy interval PLS extends the idea by combining several intervals.
- **Filter and significance methods.** Uninformative-variable elimination adds random noise variables to the data, calculates regression-coefficient reliability and removes variables that are no more stable than the noise [23].
- **Stochastic search.** Genetic algorithms treat each wavelength as a gene, evolve populations of variable subsets and retain those giving the best cross-validated fit [24]. They are widely used with PLS and ANN but demand care because the optimised subset can be over-fitted to the calibration set.

A critical review of variable-selection methods in PLS concludes that no method is universally superior, and that performance depends on the structure of the data and on how the selection is validated [25]. In practice the justification for selection is a smaller, more interpretable model with equal or lower prediction error on independent samples. The recent pharmaceutical studies illustrate this. In the analysis of caffeine, codeine, paracetamol and the degradation product p-aminophenol, a genetic algorithm reduced the number of data points to roughly one-third of the original, and the GA–ANN model predicted all validation mixtures [26]. For montelukast and levocetirizine, genetic selection reduced the data matrix by 63% and 51%, respectively, in the two GA-PLS models, giving the lowest prediction errors of the four models compared [27]. Both studies selected variables using the calibration set alone, which is the essential safeguard: the validation set must not influence wavelength choice, otherwise the reported prediction error is optimistic. A manual window remains the usual alternative; the five-component migraine study simply restricted the working range to 220–350 nm in ethanol, giving 131 data points per spectrum [28].

Choosing the number of latent variables

For PCR and PLS the single most important tuning parameter is the number of latent variables. Too few leave systematic variance unmodelled (underfitting), and too many incorporate noise (overfitting) and make predictions on new samples unreliable. The usual procedure is to plot the cross-validated error against the number of factors and select the smallest number at which the error reaches a minimum or stops improving significantly. A model should never be selected solely by its calibration fit, because the calibration error falls monotonically as components are added. Randomisation tests and the one-standard-error rule are useful tools that favour parsimonious models [25].

Chemometric calibration models

Multivariate calibration relates a block of spectra to the known concentrations of the analytes in a set of calibration mixtures, and the fitted model is then used to predict concentrations in unknown samples [29]. A useful way of organising the many algorithms is by the direction in which the spectra–concentration relationship is modelled, as shown in Figure 2. Direct models describe the spectra as a function of concentration, inverse models describe concentration as a function of the spectra, and non-linear models learn a more general mapping when the response is not strictly additive. A comparison of the principal approaches is given in Table 1.

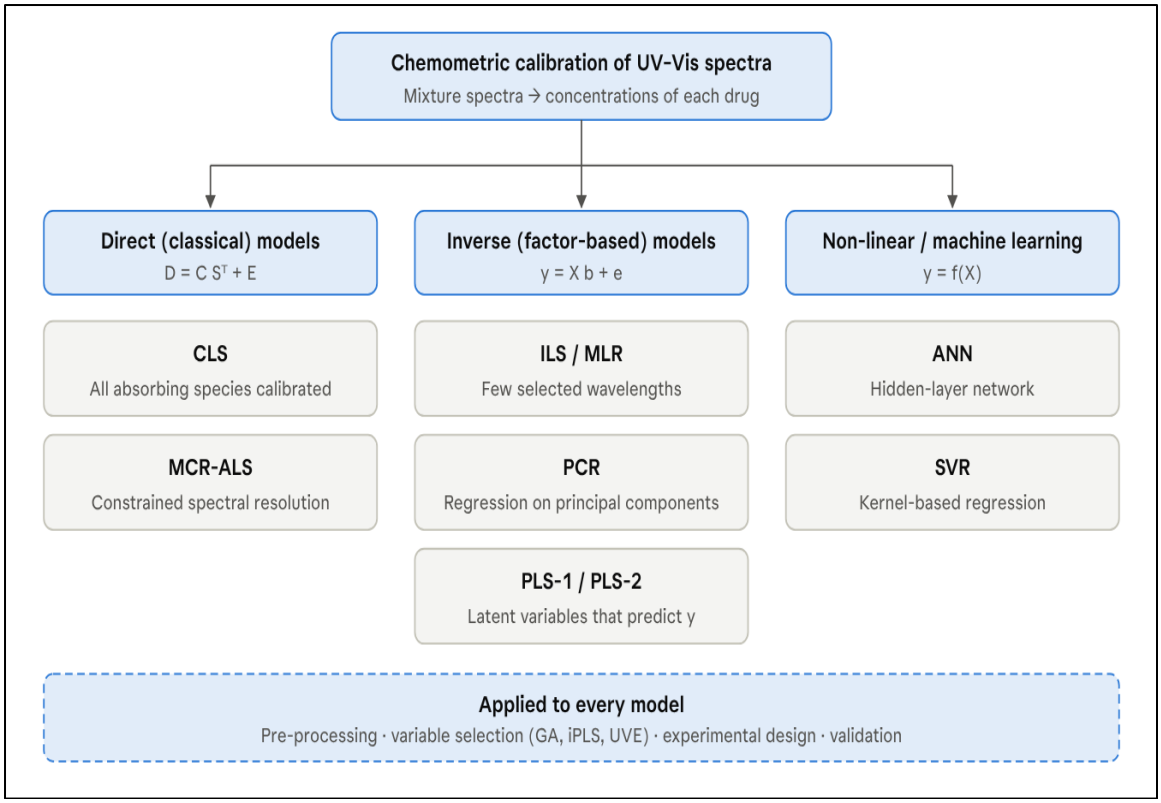

Figure 2. Classification of chemometric calibration methods used for the simultaneous estimation of drugs from UV–Vis spectra. CLS, classical least squares; ILS, inverse least squares; MLR, multiple linear regression; PCR, principal component regression; PLS, partial least squares; MCR-ALS, multivariate curve resolution–alternating least squares; ANN, artificial neural network; SVR, support vector regression.

Classical and inverse least squares

Classical least squares estimates the pure-component spectra matrix S from the calibration data and then predicts the concentrations of an unknown sample by least-squares fitting of its spectrum to those estimated spectra [29]. The method makes full use of the spectrum and is easy to interpret, but it assumes that the Beer–Lambert law holds and, crucially, that every absorbing species in the samples is included in the calibration. An excipient or degradation product that was not modelled therefore biases the result. Inverse least squares (also called multiple linear regression) reverses the regression, so that no knowledge of interferents is needed, but the number of wavelengths used cannot exceed the number of calibration samples and highly correlated absorbances produce unstable coefficients [30]. In practice, ILS requires a variable-selection step, which limits its use for strongly overlapped data.

Principal component regression

PCR addresses collinearity in two steps. Principal component analysis decomposes the mean-centred spectral matrix into orthogonal scores and loadings arranged in decreasing order of explained variance, and the concentrations are then regressed on the first few scores [31]. The retained components describe the dominant sources of spectral variance, which are not necessarily those most relevant for the analyte, and a component that explains little spectral variance but correlates strongly with concentration may be discarded. This limitation is the main reason PLS often gives lower prediction error when an analyte is present at a low level.

Partial least squares regression

PLS also projects the spectra onto latent variables, but the components are chosen to maximise the covariance between spectra and concentrations and not the spectral variance alone, so that the latent variables are relevant to prediction [32]. PLS-1 builds a separate model for each analyte, whereas PLS-2 models all analytes simultaneously with one set of latent variables. PLS-1 usually gives smaller errors, because the number of latent variables can be optimised for each analyte separately, while PLS-2 is simpler to run and may be adequate when the analytes' concentrations are correlated [33]. The algorithm is described step by step in a widely used tutorial [34], and the latent variables are, like principal components, orthogonal linear combinations of the original wavelengths [35]. The technique is the workhorse of pharmaceutical spectroscopic calibration. Multivariate regression of UV spectra has been applied to drug products since the early 1990s, when principal component regression was used to analyse a pharmaceutical preparation without prior separation [36].

Multivariate curve resolution–alternating least squares

MCR-ALS decomposes the mixture data matrix into the product of a concentration profile matrix and a pure-spectra matrix, $D = CS^T + E$, by alternating least-squares optimisation under constraints [37]. Starting from initial estimates, which are obtained by evolving factor analysis [38] or by pure-variable detection with SIMPLISMA [39], the algorithm alternately refines C and S until convergence, while constraints such as non-negativity of concentrations and spectra, or a correlation constraint that links the resolved concentrations to the known ones in the calibration samples, restrict the solutions to chemically meaningful ones [40,41]. Unlike CLS, MCR-ALS can model unknown interferents, since an extra component can be added for them, and it provides estimated spectra that can be compared with those of the reference compounds as a diagnostic. Its main weakness is rotational ambiguity: for overlapped systems, different sets of profiles may fit the data equally well, and constraints only partly reduce this uncertainty [42]. The lack of fit between the experimental and reconstructed data is the standard convergence indicator. In a study of five psychotropic drugs, MCR-ALS with non-negativity and correlation constraints converged in seven iterations with a lack of fit of 0.3541%, giving relative errors of 0.61–1.65% [43]. In a five-component migraine formulation, MCR-ALS gave the best prediction among the models compared [28].

Artificial neural networks and support vector regression

When the response is non-linear because of deviations from Beer–Lambert law, strongly interacting species, or instrumental effects, artificial neural networks provide a flexible alternative. A feed-forward network contains an input layer (the spectrum or its principal-component scores), one or more hidden layers of nodes with non-linear transfer functions, and an output layer giving the concentrations; the weights are adjusted by back-propagation [44]. Networks are powerful but require larger training sets than PLS, are sensitive to the choice of architecture and training parameters, and are prone to overfitting, which is why early stopping or an independent monitoring set is used. The interpretability of the weights is also limited. Support vector regression, derived from the support-vector method [45], fits a function that deviates from the training responses by no more than a tolerance ε while keeping the model as flat as possible, and kernels allow non-linear relationships to be handled. SVR has been applied far less than PLS or ANN to routine UV pharmaceutical assays, but it merits wider evaluation. In the analysis of a paracetamol, chlorpheniramine, caffeine and ascorbic acid product, the ANN gave the lowest prediction error among PCR, PLS, MCR-ALS and ANN [46].

Choosing among the models

No algorithm is best in every setting, and the studies are consistent on that point. For a paracetamol, phenylephrine and chlorpheniramine tablet, PLS-1 gave the lowest total relative standard error (3.10%) compared with PLS-2 (3.81%) and PCR (4.41%) [47]. Further comparisons of PLS-1, PCR and hybrid linear analysis for theophylline, montelukast and loratadine in tablets and spiked plasma have been reported [48]. As a rule of thumb, linear systems with a modest number of components are well served by PLS-1; MCR-ALS is preferable when interferents are suspected or resolved spectra are wanted for verification; and ANN or SVR should be reserved for demonstrably non-linear data with sufficient calibration samples. In all cases, models should be compared on an independent validation set, and a simpler model should be preferred when performance is equivalent.

Table 1. Comparison of the principal chemometric algorithms used for multicomponent spectroscopic analysis of drugs.

Method	Principle	Main requirements	Strengths	Limitations
CLS	Fits sample spectrum to estimated pure-component spectra	All absorbing species included in calibration	Simple, uses full spectrum, interpretable	Biased by unmodelled interferents
ILS / MLR	Regresses concentration on a subset of absorbances	Number of wavelengths ≤ number of samples	No need to model interferents explicitly	Collinearity; needs wavelength selection
PCR	PCA of spectra, then regression on scores	Choice of number of principal components	Handles collinearity; noise filtering	Components not chosen for relevance to analyte
PLS-1 / PLS-2	Latent variables maximising spectra–concentration covariance	Number of latent variables chosen by cross-validation	Robust, widely validated, predictive	Less direct chemical interpretation
MCR-ALS	Bilinear decomposition under constraints	Initial estimates; suitable constraints	Resolves spectra; tolerates interferents	Rotational ambiguity; lack of fit must be checked

ANN	Non-linear network trained by back-propagation	Large training set; architecture tuning	Handles non-linearity	Overfitting risk; limited interpretability
SVR	Kernel-based margin regression	Kernel and parameter optimisation	Good generalisation on small sets	Few UV pharmaceutical applications; tuning needed

Design of calibration and validation sets

Why the design matters

A multivariate model is only as reliable as the mixtures used to build and test it. Calibration samples must span the range of concentrations expected in the product, must vary the analytes independently of one another so that the model can distinguish their contributions, and must include enough replicates of the variation encountered in practice. If concentrations are correlated across the calibration set, for instance because a fixed formulation ratio is mimicked in every mixture, the model learns the ratio and not the individual spectra and will fail when the ratio changes. Design of experiments is therefore a core step of method development and not an optional refinement. Figure 3 places this step in the full workflow.

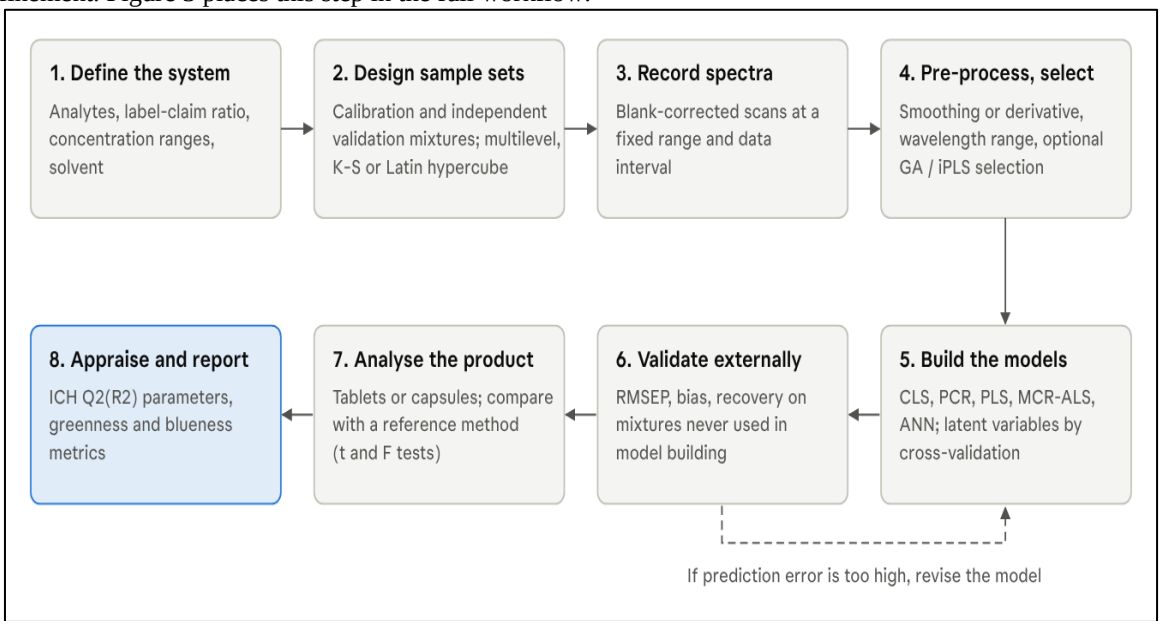

Figure 3. Stepwise workflow for developing, validating and reporting a chemometric UV–Vis method for a multicomponent dosage form. The dashed arrow indicates the iterative return to model building when external prediction error is unacceptable.

Calibration designs

The most widely used design in the recent pharmaceutical literature is the multilevel multifactor design, in which each analyte is set at several concentration levels and the combinations are chosen so that the levels are orthogonal or nearly so [49]. A five-level, four-factor design with 25 mixtures has been employed for both the paracetamol–chlorpheniramine–caffeine–ascorbic acid mixture [46] and the caffeine–codeine–paracetamol–p-aminophenol mixture [26]. The same 25-sample size was used in a five-component migraine product [28], while the psychotropic-drug study combined 25 calibration and 15 validation mixtures [43]. Smaller designs of 20 calibration mixtures have also been reported for a three-drug tablet [47]. Calibration ranges are normally centred on the working concentration obtained from the label claim and extended to roughly ±20–50% so that predictions for real tablets fall within the interpolation region and not at the extremes of the model.

Selecting validation samples

An independent validation set is essential because errors in the calibration set are optimistic and do not indicate how a model behaves on new samples [50]. Validation samples can be generated in three ways. The first is random selection from a larger pool, which is simple but may leave gaps in the design space. The second is deterministic selection with the Kennard–Stone algorithm, which picks samples sequentially so as to be maximally distant from those already chosen and thus covers the space uniformly [51]. The third is Latin hypercube sampling, a stratified random design in which each analyte's range is divided into equal-probability intervals and exactly one sample is drawn from each interval [52]. A recent anti-COVID study used 25 calibration mixtures and 13 Latin-hypercube validation mixtures [27], and the five-component migraine study used 13 validation mixtures chosen

by clustering with Kennard–Stone selection and stratification [28]. All of these approaches share one principle that tutorial accounts of validation stress: the validation set must be independent, representative and untouched by any decision made during model building, including wavelength selection and choice of the number of components [50,53].

Practical considerations

Several practical points recur in the literature. First, mixtures should be prepared gravimetrically from certified reference standards, in the same solvent and with the same dilution scheme as the sample, to prevent systematic differences between calibration and sample spectra. Second, the number of samples should exceed the number of latent variables by a comfortable margin; designs with fewer than about ten samples per analyte provide limited protection against overfitting. Third, when one component is present at a very low ratio compared with the others, as in the migraine formulation, the analyte may fall below the practical limit of quantification, and standard addition may be necessary [28]. Finally, in multivariate calibration there is no single parameter equivalent to a calibration slope, so documentation of the design, the pre-processing and the model parameters is part of the method description. Cross-validation, whether leave-one-out or venetian blinds, is useful for tuning but cannot replace the external test [53].

Applications to multicomponent pharmaceutical formulations

Scope and selection of the studies

Chemometric UV–Vis methods for drug mixtures have been published for more than three decades, beginning with PLS and PCR applications to binary and ternary products and growing toward five-component formulations and hybrid algorithms. Early reports from several groups applied PLS and PCR to bronchodilator and other combination products, often in parallel with chromatographic methods for comparison [54], and the approach was subsequently extended to many other therapeutic combinations in dosage forms [55,56]. This review does not claim to be a systematic or exhaustive survey. Instead, ten studies published between 2010 and 2024 were selected to illustrate the range of drug classes, numbers of components, algorithms, design strategies and sustainability assessments in current practice.

Table 2. Selected applications of chemometrics-assisted UV–Vis spectrophotometry to multicomponent pharmaceutical formulations (2010–2024).

No.	Analytes (formulation)	Design and spectral range	Models applied	Principal outcome	Ref.
1	Paracetamol, phenylephrine, chlorpheniramine (tablets)	20 calibration + 10 validation mixtures; 200–400 nm	PCR, PLS-1, PLS-2	Total RSE 4.41% (PCR), 3.81% (PLS-2), 3.10% (PLS-1); mean tablet recoveries about 99.5–100.5%	[47]
2	Telmisartan, hydrochlorothiazide (tablets)	25 mixtures; 200–350 nm, 151 wavelengths; 9 latent variables	PCR, PLS-1	RMSEC 0.0539 and 0.0417 (PLS-1) versus 0.0572 and 0.1068 (PCR) for the two drugs; derivative and ratio methods hindered by overlap	[15]
3	Paracetamol, sodium ascorbate, chlorpheniramine (ternary, about 150:140:1); paracetamol, caffeine, phenylephrine, chlorpheniramine (quaternary)	Multilevel mixtures; 0.5 nm data interval evaluated	PCR, PLS-1, PLS-2, MCR	PLS-1 best; finer interval improved the minor component; accuracy 94.24–107.76%	[17]
4	Theophylline, montelukast, loratadine (tablets and spiked plasma)	Laboratory mixtures; tablets and spiked human plasma	PLS-1, PCR, HLA	Three multivariate approaches compared in dosage form and biological matrix	[48]
5	Imipramine, carbamazepine, chlorpromazine, haloperidol, phenytoin (tablets, capsules, ampoules)	25 calibration + 15 validation mixtures; 220–270 nm	MCR-ALS (non-negativity and correlation constraints)	Seven iterations, lack of fit 0.3541%; relative errors 0.61–1.65%; LOD 0.06–0.72 µg/mL; no significant difference from HPLC	[43]
6	Paracetamol, chlorpheniramine,	25 mixtures, five-level four-factor	PCR, PLS, MCR-ALS,	ANN gave lowest RMSEP; MCR-ALS lack of fit	[46]

	caffeine, ascorbic acid (tablets)	design; 220–300 nm (81 points)	ANN	1.5625%; Eco-Scale score 85, AGREE 0.77	
7	Caffeine, codeine, paracetamol, p-aminophenol impurity (tablets)	25 mixtures, four-factor five-level design + 6 validation mixtures	PLS-1, GA-PLS, GA-ANN	GA retained about 31–37% of data points; GA–ANN predicted all validation mixtures	[26]
8	Montelukast, levocetirizine (tablets)	25 calibration + 13 Latin-hypercube validation mixtures; 210–400 nm (191 points)	CLS, PCR, PLS, GA-PLS	GA-PLS best: RMSEP 0.1872 and 0.1779, recoveries 98–102%; AGREE 0.90, BAGI 90, RGB12 90.8; 0.021 kg CO_2 per sample	[27]

Analgesic, antipyretic and cold-remedy combinations

Products combining paracetamol with a decongestant, an antihistamine, caffeine or vitamin C are a favourite test bed because their spectra overlap strongly and their dose ratios are extreme. For the three-drug tablet of paracetamol, phenylephrine and chlorpheniramine, PCR, PLS-1 and PLS-2 were all successful, with PLS-1 giving the smallest overall error [47]. The paracetamol–sodium ascorbate–chlorpheniramine mixture, with a 150:140:1 mass ratio, shows how a minor component can be quantified when the data interval is reduced to 0.5 nm [17]. With four components, the mixture of paracetamol, chlorpheniramine, caffeine and ascorbic acid was resolved with PCR, PLS, MCR-ALS and ANN, and the non-linear network produced the lowest prediction error [46]. In the caffeine–codeine–paracetamol product, the p-aminophenol degradation impurity was included as a fourth analyte, which brings the method closer to stability-indicating use, and a GA–ANN hybrid predicted all of the validation mixtures [26].

Cardiovascular, respiratory and anti-allergic combinations

The telmisartan–hydrochlorothiazide tablet was a clear case in which classical, derivative and derivative-ratio approaches were impeded by overlap throughout the range, whereas full-spectrum PLS-1 and PCR were applicable without further manipulation; PLS-1 gave lower calibration errors than PCR for both drugs [15]. For respiratory products, hydroxyzine, ephedrine and theophylline were resolved both by ratio subtraction with third-derivative measurement and by PCR and PLS-1, and the study is notable for comparing the greenness and blueness of the classical and multivariate routes, which scored very similarly [14]. In the montelukast–levocetirizine tablet, four models were compared, and variable selection with a genetic algorithm gave the lowest prediction error, with high greenness, blueness and whiteness scores and a reported carbon footprint of 0.021 kg CO_2 per sample [27]. A separate study of theophylline, montelukast and loratadine extended the approach to spiked human plasma by using PLS-1, PCR and hybrid linear analysis [48].

Central nervous system and migraine products

The five psychotropic drugs imipramine, carbamazepine, chlorpromazine, haloperidol and phenytoin are chemically diverse and have strongly overlapped spectra in the 220–270 nm window. MCR-ALS with non-negativity and correlation constraints resolved all five in tablets, capsules and ampoules, and a statistical comparison with HPLC showed no significant difference [43]. The five-component anti-migraine product containing ergotamine, propyphenazone, caffeine, camylofin and mecloxamine illustrates the limit of the approach: ergotamine, present at a mass ratio of 0.075 to 20 for propyphenazone, lay below the quantification limit of direct measurement and required standard addition, whereas the other four were determined with recoveries of 98–102% by MCR-ALS [28].

Emerging trends

Three trends are evident. First, formulations of increasing complexity are being addressed: the number of resolved analytes has risen from three to five, and sometimes includes an impurity or degradation product. Second, sample design has become more rigorous, with multilevel designs, Kennard–Stone and Latin hypercube validation sets replacing ad hoc mixtures. Third, hybrid and non-linear algorithms (GA-PLS, GA–ANN) are increasingly combined with variable selection to cut the data matrix by about two-thirds without loss of accuracy [26,27].

Critical appraisal

Several weaknesses limit what can be concluded from the published studies. Validation in this literature often relies on laboratory-prepared mixtures and a single marketed product, so matrix variability between manufacturers and batches is rarely tested. Formal robustness testing, inter-laboratory transfer and periodic re-validation of models appear to be described only rarely in this literature and deserve more attention. Finally, accuracy criteria vary widely between studies, and different authors define RMSEP, relative errors and recoveries on different bases, so numerical values

in Table 2 are not directly comparable and the table should be read as a catalogue of approaches rather than a ranking of methods.

Validation, figures of merit and sustainability assessment

Validation within the ICH framework

Validation of analytical procedures is governed by ICH Q2(R2), adopted in November 2023, which for the first time explicitly considers multivariate procedures, and by ICH Q14, which describes analytical procedure development and the lifecycle approach [8,9]. Q2(R2) retains the familiar performance characteristics (specificity, range, accuracy, precision, detection and quantification limits) but recognises that for multivariate methods the response is a model prediction and not a single signal, so some characteristics are demonstrated differently, typically by comparing predicted with reference values on independent samples. Table 3 sets out how each characteristic is interpreted for a chemometric UV–Vis procedure. The right-hand column lists criteria that authors commonly adopt or report in the reviewed studies; these are not requirements stated in the guideline, and the acceptance limits for a given product should be defined in advance from its intended use.

Table 3. Validation characteristics for chemometrics-assisted spectroscopic procedures: ICH Q2(R2) concepts and their multivariate implementation.

Characteristic	Concept (ICH Q2(R2))	Multivariate implementation	Reported or commonly adopted practice
Specificity / selectivity	Unequivocal assessment of the analyte in the presence of expected components	Net analyte signal and selectivity indices; residual and leverage diagnostics; placebo and excipient-spiked samples	Resolved spectra compared with pure standards (MCR-ALS); no interference from excipients
Range and response	Interval over which the procedure gives suitable accuracy and precision; relationship between response and concentration	Calibration range set by the design; plot of predicted versus nominal concentration within that range	Five-level designs; ranges centred on the label claim; regression of predicted on nominal values
Accuracy	Closeness to the true or accepted value	Percent recovery and bias on independent validation mixtures; RMSEP and RRMSEP	Recoveries of 98–102% in several studies [27,28]; elliptical joint confidence region for slope and intercept
Precision	Repeatability and intermediate precision	Replicate predictions of the same sample on one or several days or instruments	Percent RSD of replicate predictions, usually targeted at 2% or less
Detection and quantification limits	Lowest amount detected or quantified reliably	Net-analyte-signal or error-propagation estimates; 3.3σ/S and 10σ/S where applicable	LOD of 0.06–0.72 µg/mL reported for five psychotropic drugs [43]
Robustness	Capacity to remain unaffected by small deliberate changes	Variation of wavelength window, pre-processing, latent variables, instrument and analyst	Rarely reported; should be part of lifecycle monitoring [9]
Model performance (no single analogue)	Fitness of the predictive model for its purpose	RMSEC, RMSECV, RMSEP, R^2, Q^2; lack of fit for MCR-ALS	External test set; comparison of models on the same samples

Statistical comparison with a reference method

A new procedure is normally compared with an established one, for example HPLC or a pharmacopoeial method, by analysing the same batch of dosage form and applying the Student's t-test for means and the variance-ratio F-test for precision, both at the 95% confidence level. The calculated values must stay below the tabulated ones for the two methods to be considered statistically equivalent. In the psychotropic-drug study, such comparison showed no significant difference between the chemometric and HPLC results [43].

Greenness, blueness and whiteness

A main argument for chemometric spectrophotometry is that it replaces separation with computation, thereby avoiding chromatographic solvents, columns and long run times. Green analytical chemistry formalises this ambition through twelve principles [10], which are evaluated by metric tools of increasing sophistication. The National Environmental Methods Index offers a simple pictogram [60]. The Analytical Eco-Scale subtracts penalty points for reagents, energy and waste from a score of 100, with values above 75 generally regarded as excellent greenness [61]. The Green Analytical Procedure Index and its extension ComplexGAPI use colour-coded pictograms covering the whole analytical life cycle, including sample preparation and synthesis steps [62,63]. The AGREE tool scores all twelve principles and gives a single

number between 0 and 1 together with a clock-shaped diagram [64]. The Blue Applicability Grade Index (BAGI) evaluates practicality and applicability, using ten attributes [65], and white analytical chemistry combines the green, red (analytical performance) and blue (practicality) dimensions in the RGB12 model [66]. A critical review of these metrics warns that they differ in how they weight criteria and that results are tool-dependent [67].

To illustrate current reporting, the paracetamol–chlorpheniramine–caffeine–ascorbic acid method received an Eco-Scale score of 85 and an AGREE score of 0.77 [46]. The montelukast–levocetirizine method scored AGREE 0.90, BAGI 90 and RGB12 90.8 with a carbon footprint of 0.021 kg CO_2 per sample [27], and the five-component migraine method gave AGREE 0.75, BAGI 90 and RGB12 89.8 [28]. The hydroxyzine–ephedrine–theophylline study found AGREE 0.61 for the classical approach and 0.58 for the multivariate approach, with BAGI values of 77.5 and 75, respectively [14]. This last comparison is instructive: the chemometric route is not automatically greener than a well-optimised classical one, because the 25 or more calibration mixtures needed for the model consume reagents and analyst time. Because the scores in different publications were generated with different solvents, sample numbers and software versions, they should be compared within a study rather than across studies, and the number of calibration mixtures, solvent choice and waste handling should all be included in the assessment.

Advantages, limitations and regulatory perspective

Advantages

The attraction of chemometric spectrophotometry is practical. A UV–Vis spectrophotometer is inexpensive, robust and present in nearly every quality-control laboratory, and each spectrum is recorded within seconds, so a sample can be analysed without chromatographic separation, column equilibration, mobile-phase preparation or lengthy sample clean-up. Using the entire spectrum gives redundancy that averages noise and permits diagnostics, since an unusual sample reveals itself through a high spectral residual or leverage, a safeguard that a one- or two-wavelength method lacks. Finally, for analysts in resource-limited settings, the method turns routine instrumentation into a multi-analyte tool at modest cost.

Limitations

The approach also has well-recognised limitations.

- **Dependence on the calibration design.** The model is valid only within the space spanned by the calibration mixtures. Unexpected excipients, degradation products or changes of solvent and pH that were not in the design can invalidate predictions without any obvious warning other than an increased residual.
- **Low-level components.** When one analyte is present at a very small fraction of the total absorbance, its signal may be lost in noise. The ergotamine component of the five-drug migraine formulation could not be quantified directly and required standard addition [28].
- **Chromophore requirement.** Drugs without an adequate UV chromophore, or with spectra that are nearly identical, cannot be resolved however sophisticated the algorithm.
- **Linearity assumptions.** Most models assume additivity of absorbances, so self-association, hydrogen bonding, complexation or high absorbances produce non-linear responses that require ANN or SVR or a restricted range.
- **Instrument and time transfer.** A model built on one spectrophotometer may not give the same predictions on another because of differences in bandwidth, wavelength accuracy and stray light, and calibration-transfer algorithms are needed to avoid rebuilding the model [68,69].
- **Over-fitting and opacity.** Flexible models, particularly GA-optimised and neural-network models, may fit noise, and non-specialist users may accept a model without checking its diagnostics.
- **Stability indication.** Unless degradation products are included in the calibration, as the p-aminophenol impurity was in one study [26], a chemometric assay cannot be claimed to be stability-indicating.

Regulatory perspective

Regulatory acceptance has become clearer in recent years. ICH Q2(R2), which replaced the earlier Q2(R1) text, was adopted in November 2023 and became effective in the European Union in June 2024; it contains a section on multivariate procedures and explains how validation characteristics are applied when the result is a model prediction [8]. ICH Q14 complements it with guidance on analytical procedure development, including the establishment of an analytical target profile, risk-based selection of parameters, and a dedicated chapter on multivariate analytical procedures [9]. Together they imply that a chemometric method should be treated as a lifecycle activity. The development record should describe the calibration design, pre-processing, variable selection and model-building decisions; validation should use independent samples; and routine use should be backed by system-suitability checks, such as monitoring of spectral residuals or leverage, together with a plan for model maintenance and re-validation after changes in

instrument, supplier or formulation. For a student or early-career researcher preparing a method for publication, practical consequences are to document the design and software versions, to report error statistics on a truly independent set, and to include a comparison with a reference method, which also strengthens the case for acceptance by journals.

FUTURE PERSPECTIVES:

Several developments are likely to shape the field in the next decade. The first is the migration from classical linear models to deep learning. Convolutional neural networks have been shown to model vibrational spectra without explicit pre-processing and with accuracy competitive with PLS [70], and broader reviews of artificial intelligence in chemometrics describe the opportunities and the risks of data-hungry models [71]. For UV–Vis pharmaceutical assays, where calibration sets of a few dozen mixtures are typical, the immediate prospect is not replacement of PLS but a role for non-linear models when the data genuinely demand it, supported by data augmentation, regularisation and transparent reporting of the training protocol.

Second, explainability will become a requirement. Regulators and reviewers expect evidence that a model is relying on chemically meaningful spectral regions, which can be shown through regression coefficients, variable-importance measures and MCR-ALS-resolved spectra that match reference spectra. Third, miniaturised and portable spectrometers, including smartphone-based devices, together with at-line and in-line probes, open the way to real-time release testing and process analytical technology, provided that calibration transfer between devices is solved [68,69]. Fourth, data fusion of UV–Vis with complementary signals such as fluorescence or infrared spectra could extend the scope to analytes whose UV spectra are too similar to resolve.

Fifth, sustainability assessment will become standard rather than optional, and a harmonised approach to reporting AGREE, BAGI and carbon-footprint scores, including the number of calibration mixtures and the energy used, would make cross-study comparisons meaningful. Sixth, extension of the approach to biological matrices such as plasma, as has been attempted for theophylline, montelukast and loratadine [48], would allow therapeutic drug monitoring and bioequivalence screening, though this requires proof of selectivity against endogenous components. Finally, the field would benefit from open, shared spectral datasets and standardised reporting checklists so that models can be reproduced and benchmarked across laboratories.

CONCLUSION:

Chemometrics converts the main limitation of UV–Vis spectrophotometry, spectral overlap, into an advantage by exploiting the information contained in the full spectrum. For multicomponent formulations, PLS remains a robust default, MCR-ALS adds the ability to resolve and verify the spectra of interfering species, and variable-selection hybrids such as GA-PLS and GA–ANN can reduce data volumes substantially without loss of accuracy. The most reliable studies share common features: a designed calibration set with independent validation mixtures chosen by Kennard–Stone or Latin hypercube sampling, optimisation of the model on the calibration data alone, validation consistent with ICH Q2(R2) and Q14, comparison with a reference method, and transparent greenness assessment. The remaining gaps are robustness, calibration transfer, testing across batches and manufacturers, and the extension to stability-indicating and bioanalytical applications. Addressing these gaps, and reporting results in a harmonised form, will determine how widely chemometric spectrophotometry is accepted in regulated quality control.

CONFLICT OF INTEREST:

The Author declares that there is no conflict of interest.

REFERENCES:

1. Skoog DA, Holler FJ, Crouch SR. Principles of Instrumental Analysis. 7th ed. Boston: Cengage Learning; 2018.
2. Wold S. Chemometrics; what do we mean with it, and what do we want from it? Chemom Intell Lab Syst. 1995;30(1):109–115.
3. Brereton RG. Chemometrics: Data Driven Extraction for Science. 2nd ed. Chichester: Wiley; 2018.
4. Nurani LH, et al. Chemometrics-assisted UV-Vis spectrophotometry for quality control of pharmaceuticals: a review. Indones J Chem. 2023;23(2):542–567. doi:10.22146/ijc.74329
5. Cerdà V, Phansi P, Ferreira SLC. [Title to be verified against the published article]. TrAC Trends Anal Chem. 2022;157:116772.
6. Ríos-Reina R, Azcarate SM. How chemometrics revives the UV-Vis spectroscopy applications as an analytical sensor for spectralprint (nontargeted) analysis. Chemosensors. 2023;11(1):8.
7. Mazivila SJ, Olivieri AC. Chemometrics coupled to vibrational spectroscopy and spectroscopic imaging for the analysis of solid-phase pharmaceutical products: a brief review on non-destructive analytical methods. TrAC Trends Anal Chem. 2018;108:74–87.
8. International Council for Harmonisation. ICH guideline Q2(R2) on validation of analytical

procedures. Adopted November 2023. Geneva: ICH; 2023.
9. International Council for Harmonisation. ICH guideline Q14 on analytical procedure development. Adopted November 2023. Geneva: ICH; 2023.
10. 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.
11. Bosch Ojeda C, Sánchez Rojas F. Recent developments in derivative ultraviolet/visible absorption spectrophotometry. Anal Chim Acta. 2004;518:1–24.
12. Sánchez Rojas F, Bosch Ojeda C. Recent development in derivative ultraviolet/visible absorption spectrophotometry: 2004–2008. A review. Anal Chim Acta. 2009;635:22–44.
13. Şentürk Z, et al. Determination of theophylline and ephedrine HCl in tablets by ratio-spectra derivative spectrophotometry and LC. J Pharm Biomed Anal. 2002;29(1–2):291–298.
14. AlSalem HS, Algethami FK, Magdy MA, Ali NW, Zaazaa HE, Abdelkawy M, et al. Harnessing spectrophotometry resolution power for determining ternary mixture for respiratory disorders treatment in their pharmaceutical formulation. PLoS ONE. 2024;19(10):e0311121. doi:10.1371/journal.pone.0311121
15. Lakshmi KS, Lakshmi S. Design and optimization of a chemometric-assisted spectrophotometric determination of telmisartan and hydrochlorothiazide in pharmaceutical dosage form. J Young Pharm. 2010;2(1):85. doi:10.4103/0975-1483.62224
16. Martens H, Næs T. Multivariate Calibration. Chichester: Wiley; 1989.
17. De Luca M, Ioele G, Spatari C, Ragno G. Optimization of wavelength range and data interval in chemometric analysis of complex pharmaceutical mixtures. J Pharm Anal. 2016;6(1):64–69. doi:10.1016/j.jpha.2015.10.001
18. Savitzky A, Golay MJE. Smoothing and differentiation of data by simplified least squares procedures. Anal Chem. 1964;36(8):1627–1639.
19. Barnes RJ, Dhanoa MS, Lister SJ. Standard normal variate transformation and de-trending of near infrared diffuse reflectance spectra. Appl Spectrosc. 1989;43(5):772–777.
20. Geladi P, MacDougall D, Martens H. Linearization and scatter-correction for near-infrared reflectance spectra of meat. Appl Spectrosc. 1985;39(3):491–500.
21. Rinnán Å, van den Berg F, Engelsen SB. Review of the most common pre-processing techniques for near-infrared spectra. TrAC Trends Anal Chem. 2009;28(10):1201–1222.
22. Nørgaard L, Saudland A, Wagner J, Nielsen JP, Munck L, Engelsen SB. Interval partial least-squares regression (iPLS): a comparative chemometric study with an example from near-infrared spectroscopy. Appl Spectrosc. 2000;54(3):413–419.
23. Centner V, Massart DL, de Noord OE, de Jong S, Vandeginste BM, Sterna C. Elimination of uninformative variables for multivariate calibration. Anal Chem. 1996;68(21):3851–3858.
24. Leardi R. Genetic algorithms in chemometrics and chemistry: a review. J Chemom. 2001;15(7):559–569.
25. Mehmood T, Liland KH, Snipen L, Sæbø S. A review of variable selection methods in partial least squares regression. Chemom Intell Lab Syst. 2012;118:62–69.
26. Kelani KM, Fekry RA, Fayez YM, Hassan SA. Advanced chemometric methods for simultaneous quantitation of caffeine, codeine, paracetamol, and p-aminophenol in their quaternary mixture. Sci Rep. 2024;14:2085. doi:10.1038/s41598-024-52450-4
27. 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. doi:10.1186/s13065-024-01158-7
28. Abbas AEF, Abdelshafi NA, Gamal M, Halim MK, Said BAM, Naguib IA, et al. Simultaneously quantifying a novel five-component anti-migraine formulation containing ergotamine, propyphenazone, caffeine, camylofin, and mecloxamine using UV spectrophotometry and chemometric models. BMC Chem. 2024;18:233. doi:10.1186/s13065-024-01339-4
29. Brereton RG. Introduction to multivariate calibration in analytical chemistry. Analyst. 2000;125:2125–2154.
30. Haaland DM, Thomas EV. Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. Anal Chem. 1988;60(11):1193–1202.
31. Abdi H, Williams LJ. Principal component analysis. WIREs Comput Stat. 2010;2(4):433–459.
32. Wold S, Sjöström M, Eriksson L. PLS-regression: a basic tool of chemometrics. Chemom Intell Lab Syst. 2001;58(2):109–130.
33. Helland IS. Some theoretical aspects of partial least squares regression. Chemom Intell Lab Syst. 2001;58(2):97–107.

34. Geladi P, Kowalski BR. Partial least-squares regression: a tutorial. Anal Chim Acta. 1986;185:1–17.
35. Wold S, Esbensen K, Geladi P. Principal component analysis. Chemom Intell Lab Syst. 1987;2(1–3):37–52.
36. Blanco M, Coello J, González F, Iturriaga H, Maspoch S. Spectrophotometric analysis of a pharmaceutical preparation by principal component regression. J Pharm Sci. 1993;82(8):834–837. doi:10.1002/jps.2600820816
37. Tauler R. Multivariate curve resolution applied to second order data. Chemom Intell Lab Syst. 1995;30(1):133–146.
38. Maeder M, Zuberbuehler AD. The resolution of overlapping chromatographic peaks by evolving factor analysis. Anal Chim Acta. 1986;181:287–291.
39. Windig W, Guilment J. Interactive self-modeling mixture analysis. Anal Chem. 1991;63(14):1425–1432.
40. de Juan A, Jaumot J, Tauler R. Multivariate curve resolution (MCR). Solving the mixture analysis problem. Anal Methods. 2014;6:4964–4976.
41. Jaumot J, de Juan A, Tauler R. MCR-ALS GUI 2.0: new features and applications. Chemom Intell Lab Syst. 2015;140:1–12.
42. Azzouz T, Tauler R. Application of multivariate curve resolution alternating least squares (MCR-ALS) to the quantitative analysis of pharmaceutical and agricultural samples. Talanta. 2008;74(5):1201–1210.
43. Shaaban H, Mostafa A, Al-Zahrani B, Al-Jasser B, Al-Ghamdi R. Simultaneous determination of drugs affecting central nervous system (CNS) in bulk and pharmaceutical formulations using multivariate curve resolution-alternating least squares (MCR-ALS). J Anal Methods Chem. 2020;2020:1684172. doi:10.1155/2020/1684172
44. Zupan J, Gasteiger J. Neural networks: a new method for solving chemical problems or just a passing phase? Anal Chim Acta. 1991;248:1–30.
45. Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20:273–297.
46. Mouhamed AA, Nadim AH, Mostafa NM, Eltanany BM. Application of smart chemometric models for spectra resolution and determination of challenging multi-action quaternary mixture: statistical comparison with greenness assessment. BMC Chem. 2024;18:44. doi:10.1186/s13065-024-01148-9
47. Khoshayand MR, Abdollahi H, Ghaffari A, Shariatpanahi M, Farzanegan H. Simultaneous spectrophotometric determination of paracetamol, phenylephrine and chlorpheniramine in pharmaceuticals using chemometric approaches. Daru. 2010;18(4):292–297.
48. Hassaninejad-Darzi SK, Samadi-Maybodi A, Nikou SM. UV-Vis spectrophotometry and multivariate calibration method for simultaneous determination of theophylline, montelukast and loratadine in tablet preparations and spiked human plasma. Iran J Pharm Res. 2016;15(3):379.
49. Brereton RG. Multilevel multifactor designs for multivariate calibration. Analyst. 1997;122:1521–1529.
50. Esbensen KH, Geladi P. Principles of proper validation: use and abuse of re-sampling for validation. J Chemom. 2010;24(3–4):168–187.
51. Kennard RW, Stone LA. Computer aided design of experiments. Technometrics. 1969;11(1):137–148.
52. McKay MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics. 1979;21(2):239–245.
53. Westad F, Marini F. Validation of chemometric models – a tutorial. Anal Chim Acta. 2015;893:14–24.
54. El-Gindy A, Emara S, Shaaban H. Development and validation of chemometrics-assisted spectrophotometric and liquid chromatographic methods for the simultaneous determination of two multicomponent mixtures containing bronchodilator drugs. J Pharm Biomed Anal. 2007;43(3):973–982.
55. Darwish HW, Hassan SA, Salem MY, El-Zeany BA. Comparative study between derivative spectrophotometry and multivariate calibration as analytical tools applied for the simultaneous quantitation of amlodipine, valsartan and hydrochlorothiazide. Spectrochim Acta A Mol Biomol Spectrosc. 2013;113:215–223.
56. Darwish HW, Hassan SA, Salem MY, El-Zeany BA. Different approaches in partial least squares and artificial neural network models applied for the analysis of a ternary mixture of amlodipine, valsartan and hydrochlorothiazide. Spectrochim Acta A Mol Biomol Spectrosc. 2014;122:744–750.
57. Olivieri AC, Faber NM, Ferré J, Bocqué R, Kalivas JH, Mark H. Uncertainty estimation and figures of merit for multivariate calibration (IUPAC technical report). Pure Appl Chem. 2006;78(3):633–661.
58. Lorber A. Error propagation and figures of merit for quantification by solving matrix equations. Anal Chem. 1986;58(6):1167–1172.
59. Olivieri AC. Practical guidelines for reporting results in single- and multi-component

analytical calibration: a tutorial. Anal Chim Acta. 2015;868:10–22.
60. Keith LH, Gron LU, Young JL. Green analytical methodologies. Chem Rev. 2007;107(6):2695–2708.
61. Gałuszka A, Migaszewski ZM, Konieczka P, Namieśnik J. Analytical Eco-Scale for assessing the greenness of analytical procedures. TrAC Trends Anal Chem. 2012;37:61–72.
62. Płotka-Wasylka J. A new tool for the evaluation of the analytical procedure: Green Analytical Procedure Index. Talanta. 2018;181:204–209.
63. Płotka-Wasylka J, Wojnowski W. Complementary green analytical procedure index (ComplexGAPI) and software. Green Chem. 2021;23:8657–8665.
64. Pena-Pereira F, Wojnowski W, Tobiszewski M. AGREE – Analytical GREEnness metric approach and software. Anal Chem. 2020;92(14):10076–10082.
65. Manousi N, Wojnowski W, Płotka-Wasylka J, Samanidou V. Blue applicability grade index (BAGI) and software: a new tool for the evaluation of method practicality. Green Chem. 2023;25:7598–7604.
66. Nowak PM, Wietecha-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.
67. Sajid M, Płotka-Wasylka J. Green analytical chemistry metrics: a review. Talanta. 2022;238:123046.
68. Workman J Jr. A review of calibration transfer practices and instrument differences in spectroscopy. Appl Spectrosc. 2018;72(3):340–365.
69. Feudale RN, Woody NA, Tan H, Myles AJ, Brown SD, Ferré J. Transfer of multivariate calibration models: a review. Chemom Intell Lab Syst. 2002;64(2):181–192.
70. Acquarelli J, van Laarhoven T, Gerretzen J, Tran TN, Buydens LMC, Marchiori E. Convolutional neural networks for vibrational spectroscopic data analysis. Anal Chim Acta. 2017;954:22–31.
71. Houhou R, Bocklitz T. Trends in artificial intelligence, machine learning, and chemometrics applied to chemical data. Anal Sci Adv. 2021;2(3–4):128–141.