STUDIA UBB CHEMIA, LXXI, 2, 2026 (p. 105-123)
(RECOMMENDED CITATION)
DOI:10.24193/subbchem.2026.2.06

COMPARATIVE APPLICATION OF RAW-ANN AND PCA-ANN FOR THE SPECTROPHOTOMETRIC DETERMINATION OF CAFFEINE, PROPYPHENAZONE, AND PARACETAMOL IN TABLETS

Asiye ÜÇER[a,b], Nazangül ÜNAL[c], Erdal DİNÇ[a,*]

ABSTRACT. Comparative chemometric approaches based on raw artificial neural networks (RAW-ANN) and principal component analysis-artificial neural networks (PCA-ANN) were developed and applied to the simultaneous spectrophotometric determination of caffeine (CFN), propyphenazone (PRPN), and paracetamol (PRC) in a commercial ternary pharmaceutical formulation. The spectral overlap of the three components within the 220-300 nm region renders conventional spectrophotometric methods inadequate without prior separation. The proposed ANN-based models enabled direct analysis of raw UV spectral data without requiring any separation procedure over concentration ranges of 2.5-12.0 µg/mL (CFN), 3.0-12.0 µg/mL (PRPN), and 3.0-16.0 µg/mL (PRC).

Both training approaches demonstrated satisfactory analytical performance, with mean recoveries ranging between 97% and 105%. In particular, PCA-ANN yielded relative standard deviations below 2.5%, indicating enhanced precision and predictive stability compared to RAW-ANN. Statistical evaluation confirmed the robustness and reliability of the developed models. The methods were successfully applied to the quantitative analysis of pharmaceutical tablets, demonstrating their suitability for routine quality control of complex multicomponent formulations.

Keywords: *Spectrophotometric determination, artificial neural networks, principal component analysis, caffeine, propyphenazone, paracetamol*

[a] *Department of Analytical Chemistry, Faculty of Pharmacy, Ankara University, 06560 Yenimahalle, Ankara, Türkiye.*
[b] *Department of Analytical Chemistry, Faculty of Pharmacy, Ankara Yıldırım Beyazıt University, 06010 Etlik, Keçiören, Ankara, Türkiye.*
[c] *Department of Pharmacy Services, Eşme Vocational School, Uşak University, 64600, Uşak, Turkey.*
[*] *Corresponding author: dinc@ankara.edu.tr*

INTRODUCTION

In modern medical practice, multi-component pharmaceutical formulations are increasingly preferred over single-agent therapies due to their enhanced therapeutic efficacy and improved clinical outcomes. In particular, triple-component analgesic combinations are widely used in over-the-counter preparations to achieve synergistic effects and improved patient compliance. Among these, the combination of paracetamol, propyphenazone, and caffeine is frequently incorporated into commercial pharmaceutical products.

Paracetamol (acetaminophen; IUPAC: N-(4-hydroxyphenyl) acetamide) is a widely used analgesic and antipyretic agent administered alone or in combination with other drugs in various dosage forms [1]. Propyphenazone (IUPAC: 1,5-dimethyl-2-phenyl-4-propan-2-yl-pyrazol-3-one) is a pyrazolone-derived nonsteroidal anti-inflammatory drug (NSAID) possessing analgesic and antipyretic properties and is commonly included in combination analgesics [2-3]. Caffeine (IUPAC: 1,3,7-trimethyl-3,7-dihydro-1H-purine-2,6-dione), a methylxanthine derivative, enhances analgesic efficacy through central nervous system stimulation and adenosine receptor antagonism [4-5]. Owing to its pharmacological effects, caffeine is frequently incorporated into multicomponent formulations [6-7].

The accurate and reliable quantification of these active ingredients in ternary pharmaceutical formulations is essential for quality control, regulatory compliance, and routine industrial analysis. However, the simultaneous determination of paracetamol, propyphenazone, and caffeine represents a considerable analytical challenge due to their strongly overlapping absorption spectra and the complexity of multicomponent systems.

Numerous chromatographic methods, including RP-HPLC [8–9], gradient HPLC [10], TLC and HPTLC [10–14], gas chromatography [15], pressurized planar electrochromatography [13], micellar electrokinetic capillary chromatography [16], and voltammetric techniques [17], have been reported for the analysis of these compounds in pharmaceutical preparations and biological matrices. Although these approaches offer high selectivity and sensitivity, they often require sophisticated instrumentation, extensive sample preparation, long analysis times, and increased operational costs.

UV spectrophotometry offers a rapid and cost-effective alternative for routine pharmaceutical analysis. Nevertheless, the pronounced spectral overlap of paracetamol, propyphenazone, and caffeine within the 220-300 nm region severely limits the applicability of classical univariate spectrophotometric approaches. To address this limitation, several mathematical and multivariate techniques have been proposed, including derivative spectrophotometric methods [18,22], absorption ratio approaches [19], principal component regression [20], net analyte signal-based strategies [21], and flow-through UV

systems [23]. Although these methods improve selectivity, they may suffer from sensitivity to noise amplification, complex data preprocessing steps, or limited predictive robustness when applied to highly collinear spectral datasets.

Artificial neural networks (ANNs) have emerged as powerful chemometric tools capable of modeling complex and nonlinear relationships in multivariate analytical data. Raw artificial neural networks (RAW-ANNs) directly utilize original spectral data, enabling the resolution of severely overlapping signals without prior mathematical manipulation. Furthermore, the integration of principal component analysis with ANN (PCA-ANN) reduces data dimensionality, alleviates collinearity, and enhances model stability and convergence performance [24-27]. In comparison with conventional chromatographic and classical spectrophotometric approaches, ANN-based training methods provide several practical and computational advantages for multicomponent pharmaceutical analysis. These methods enable direct quantitative analysis of highly overlapping spectral data without requiring prior separation procedures, thereby reducing analysis time, instrumental complexity, and sample preparation steps. Furthermore, ANN architectures are capable of modeling complex and nonlinear relationships within highly collinear datasets, while PCA-assisted dimensional reduction improves numerical stability, accelerates convergence, and enhances predictive robustness. Despite their proven effectiveness in various multicomponent pharmaceutical systems, a systematic comparative evaluation of RAW-ANN and PCA-ANN for the direct spectrophotometric determination of the paracetamol-propyphenazone-caffeine ternary system has not yet been reported.

In the present study, RAW-ANN and PCA-ANN models were comparatively developed and applied for the first time to the simultaneous spectrophotometric determination of caffeine (CFN), propyphenazone (PRPN), and paracetamol (PRC) in a commercial ternary pharmaceutical formulation without any prior separation step. The predictive performance, precision, and convergence characteristics of both chemometric strategies were systematically evaluated, and their applicability to routine pharmaceutical quality control was demonstrated.

THEORETICAL BACKGROUND

Artificial Neural Networks (ANN)

Artificial neural networks (ANNs) are data-driven modeling tools inspired by biological neural systems and are widely used for nonlinear multivariate training in analytical chemistry. In feedforward neural networks, the relationship between input variables and output responses is established through interconnected processing units (neurons) organized in layers.

For a typical two-layer feedforward ANN, the output Y_k can be expressed as:

$$Y_k = f_o\left(\sum_{j=1}^{m} \omega_{jk}^{(2)} f_h\left(\sum_{i=1}^{n} \omega_{ij}^{(1)} x_i + b_j^{(1)}\right) + b_k^{(2)}\right) \quad (1)$$

where x_i represents the input spectral variables, $\omega_{ij}^{(1)}$ and $\omega_{jk}^{(2)}$ are weight coefficients, $b_j^{(1)}$ and $b_k^{(2)}$ are bias terms, f_h and f_o denote hidden and output layer activation functions, respectively.

In chemometric training, ANN models are particularly advantageous for handling collinear spectral data and nonlinear relationships between absorbance and concentration values. RAW-ANN utilizes the original spectral absorbance matrix directly as the input dataset.

Principal Component Analysis (PCA)

Principal component analysis (PCA) is a multivariate projection method used to reduce the dimensionality of highly correlated datasets while preserving the maximum variance. The original spectral matrix X can be decomposed as:

$$X = TP^T + E \quad (2)$$

where T is the score matrix, P is the loading matrix, and E is the residual matrix. Principal components are orthogonal linear combinations of the original variables and therefore eliminate multicollinearity inherent in spectral data. The removal of multicollinearity improves numerical stability by reducing the condition number of the data matrix, minimizes variance inflation of model parameters, and enhances predictive robustness. Furthermore, by compressing redundant spectral information and filtering low-variance noise components, PCA preprocessing reduces the risk of overfitting and improves convergence behavior when applied before ANN modeling.

In the PCA-ANN approach, the reduced score matrix obtained from PCA is used as the input dataset for ANN modeling instead of the original absorbance matrix. This preprocessing step reduces dimensionality, improves computational efficiency, enhances convergence speed, and minimizes the risk of overfitting.

Therefore, while RAW-ANN directly models the original spectral data, PCA-ANN operates on decorrelated and compressed input variables, potentially leading to improved prediction performance and model robustness.

RESULTS AND DISCUSSION

RAW-ANN and PCA-ANN Modeling Strategy and Advanced Comparative Evaluation

To achieve reliable simultaneous quantification of PRC (Figure 1a), CFN (Figure 1b), and PRPN (Figure 1c), in the presence of severe spectral overlap in the 220-300 nm region, two nonlinear multivariate training strategies, RAW-ANN and PCA-ANN, were systematically developed and compared.

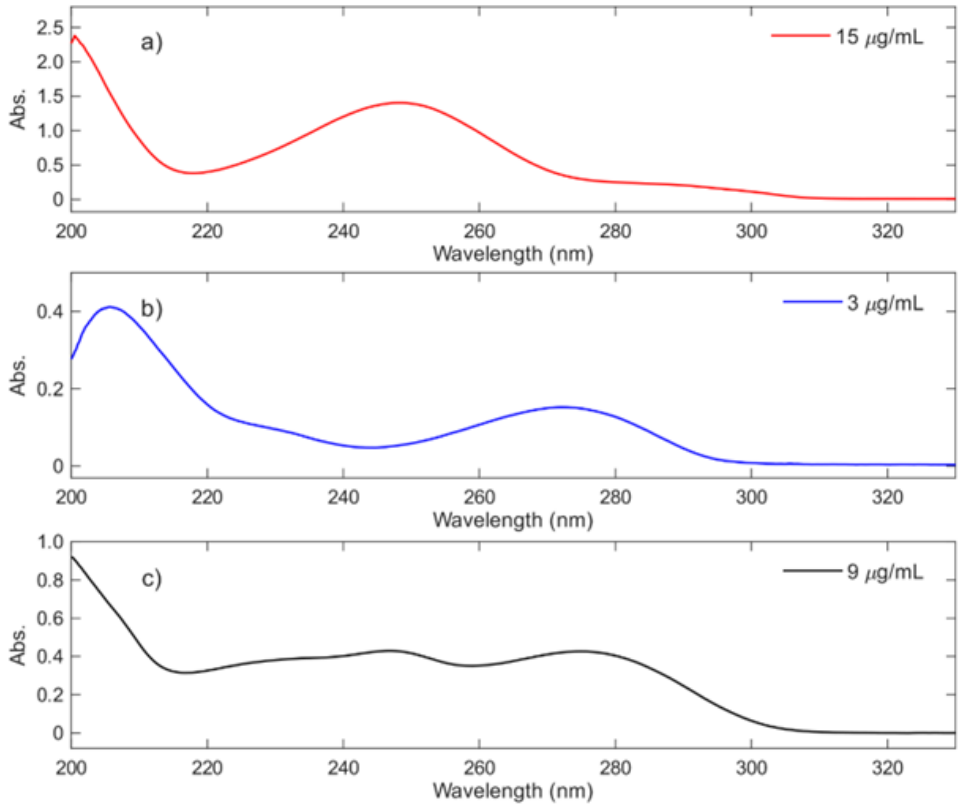

Figure 1. UV absorption spectra of a) 15.0 µg/mL PRC, b) 3.0 µg/mL CFN, and c) 9.0 µg/mL PRPN

The training dataset consisted of 35 synthetic ternary mixtures prepared using a structured concentration design to ensure adequate and balanced coverage of the working ranges of CFN, PRPN, and PRC (Table 1).

For each mixture, UV absorption spectra were recorded over 220-300 nm, generating 700 absorbance variables per sample. This resulted in an initial spectral matrix of 700 × 35 (wavelength × sample), which was subsequently rearranged to 35 × 700 before ANN modeling.

Table 1. A concentration set that includes mixtures of analyzed drugs

Sample No.	Actual CFN (µg/mL)	Actual PRPN (µg/mL)	Actual PRC (µg/mL)	Sample No.	Actual CFN (µg/mL)	Actual PRPN (µg/mL)	Actual PRC (µg/mL)
1	2.5	3	3	19	2.5	6	12
2	2.5	3	8	20	2.5	6	16
3	2.5	3	12	21	4	3	3
4	2.5	3	16	22	4	3	8
5	4	6	3	23	4	3	12
6	4	6	8	24	4	3	16
7	4	6	12	25	8	12	3
8	4	6	16	26	8	12	8
9	8	9	3	27	8	12	12
10	8	9	8	28	8	12	16
11	8	9	12	29	12	9	3
12	8	9	16	30	12	9	8
13	12	12	3	31	12	9	12
14	12	12	8	32	12	9	16
15	12	12	12	33	0	0	15
16	12	12	16	34	0	9	0
17	2.5	6	3	35	3	0	0
18	2.5	6	8				

The same spectral acquisition and preprocessing procedures were applied to independent datasets, including 13 external validation mixtures, 18 intra-day and inter-day precision samples, 12 standard addition samples, and 10 commercial tablet solution samples prepared from the marketed formulation. The latter dataset was used to evaluate the practical applicability of the proposed models to real pharmaceutical samples. The three-dimensional UV absorption spectra of 35 training sets, 43 validation sets (13 external validation mixtures, 18 intra-day and inter-day sensitivity samples, 12 standard additive samples), and 10 commercial triple tablet formulation samples are shown in Figures 2a-c.

Training of both RAW-ANN and PCA-ANN models was performed using the Levenberg–Marquardt backpropagation algorithm (trainlm). The weights and biases between layers were iteratively optimized by minimizing the mean squared prediction error. A maximum of 1000 epochs and a performance goal of 1×10^{-8} were employed, together with validation monitoring to improve generalization and reduce overfitting.

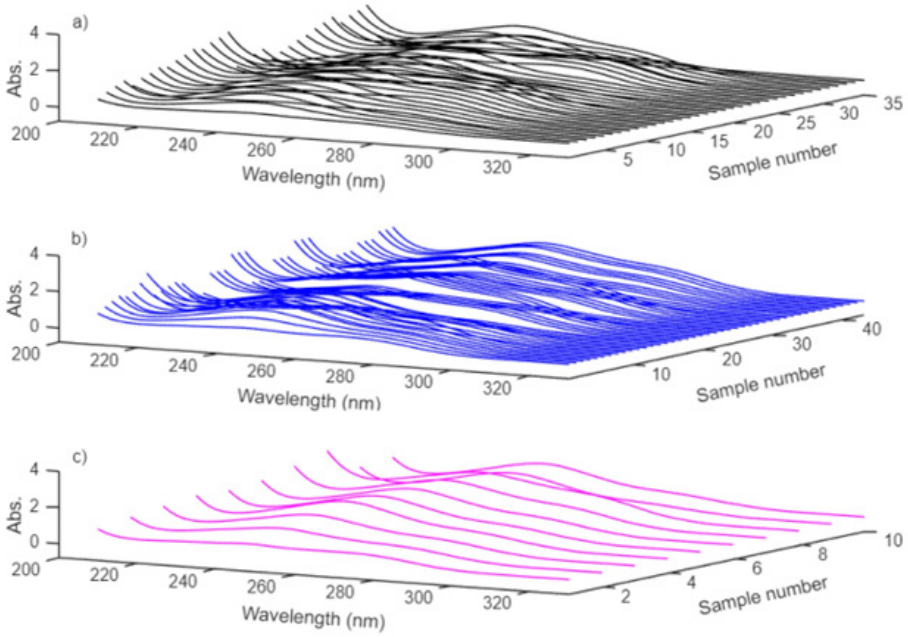

Figure 2. 3D-UV Absorption spectra of a) 35 training sets, b) 43 validation sets (13 external validation mixtures, 18 intra-day and inter-day precision samples, 12 standard addition samples), and c) 10 commercial ternary tablet formulation samples

RAW-ANN Modeling

To reduce spectral redundancy and computational burden, consecutive wavelength variables were averaged, compressing the original 700 absorbance variables into 100 representative inputs. This preprocessing step transformed the original 35 × 700 training matrix into a reduced 35 × 100 dataset used as the ANN input. These 100 correlated inputs were introduced into a multilayer feedforward ANN with architecture:

- Input layer: 100 neurons
- First hidden layer: 50 neurons (logsig transfer function)
- Second hidden layer: 25 neurons (purelin transfer function)
- Output layer: 3 neurons corresponding to CFN, PRPN, and PRC
- Training was performed using the Levenberg–Marquardt backpropagation algorithm (trainlm), with a maximum of 1000 epochs and a performance goal of 1×10^{-8}.

The optimized network topology was presented in Table 2.

Table 2. ANN topologies for the applied training methods

	RAW-ANN	**PCA-ANN**
Input sizes	100	10
Hidden transfer functions	logsig, purelin	logsig, purelin
Output transfer functions	purelin	purelin
epochs	61	9
train time (min.)	5.06	1.06

This 100-50-25-3 configuration resulted in more than 6,600 adjustable parameters (weights and biases), reflecting substantial model complexity relative to the training sample size (n = 35). Although such parameterization enhances nonlinear approximation capability, it increases sensitivity to multicollinearity inherent in highly correlated spectral datasets.

The trained RAW-ANN model was subsequently applied to independent datasets processed using identical preprocessing steps:

- External validation mixtures (13 × 100)
- Intra-day and inter-day precision samples (18 × 100)
- Standard addition samples (12 × 100)
- Commercial tablet solutions (10 × 100)

The corresponding quantitative results are summarized in Tables 4, 5, 6, and 7. Training required 303.90 s (≈5.06 min), with a total processing time of 307.09 s. The training performance curve is presented in Figure 3a, demonstrating slower convergence relative to PCA-ANN (Figure 3b), primarily due to the higher-dimensional and correlated input structure.

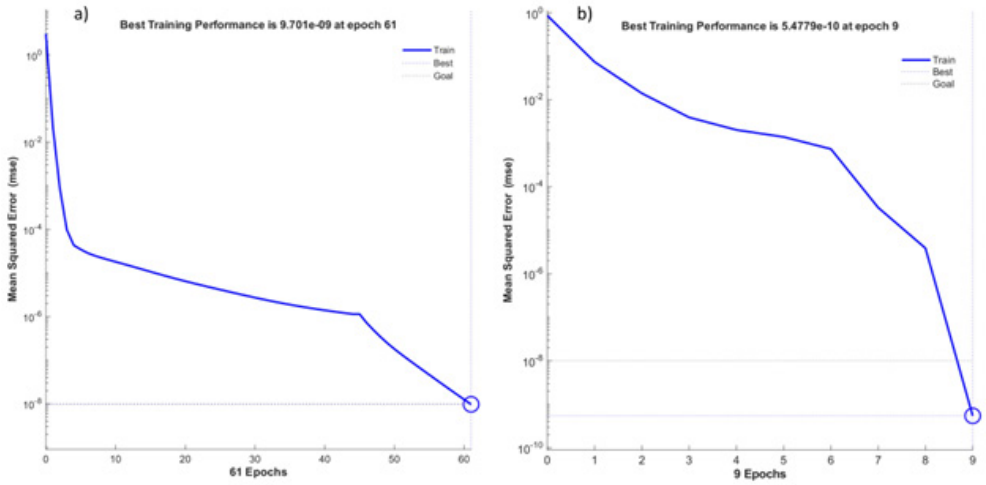

Figure 3. a) Mean Squared Error performance for training of ANN with RAW Data Inputs, b) Mean Squared Error performance for training of ANN with PCA Applied Inputs

Statistical parameters summarized in Table 3 indicate satisfactory training performance; however, higher RMSEP and SEP values compared to PCA-ANN suggest limited predictive stability under high-dimensional correlated input conditions.

PCA-ANN Modeling

To overcome multicollinearity and reduce spectral dimensionality, principal component analysis (PCA) was applied to the original 35 × 700 training matrix prior to ANN modeling. The PCA decomposition generated loading vectors, score matrices, and eigenvalues of the covariance matrix.

Although the first five principal components explained 99.6% of the total spectral variance, ten principal components were retained to ensure preservation of minor spectral contributions and to enhance numerical stability. Consequently, the training dataset was reduced to a 35 × 10 score matrix.

Prior to ANN modeling, the input datasets were normalized to improve numerical stability and network convergence. In the PCA-ANN approach, PCA scores derived from the spectral matrix were used as orthogonalized inputs for ANN training.

The optimized network structure of the PCA-ANN architecture consisted of:

- Input layer: 10 neurons (principal component scores)
- First hidden layer: 10 neurons (logsig transfer function)
- Second hidden layer: 5 neurons (purelin transfer function)
- Output layer: 3 neurons corresponding to CFN, PRPN, and PRC concentrations.
- Training was performed using the Levenberg–Marquardt backpropagation algorithm (trainlm), with a maximum of 1000 epochs and a performance goal of 1×10^{-8}.

Compared to the RAW-ANN (100 input neurons), the PCA-ANN structure significantly reduced the number of adjustable parameters, thereby improving the training sample-to-parameter ratio and enhancing model generalization capability.

Validation datasets (13 × 10), intra/inter-day samples (18 × 10), standard addition samples (12 × 10), and commercial tablet solutions (10 × 10) were projected onto the PCA loading space prior to ANN prediction to ensure methodological consistency.

The evolution of training performance is shown in Figure 3b. PCA-ANN exhibited faster convergence and earlier stabilization compared to RAW-ANN (Figure 3a), requiring approximately 1.06 minutes for training.

Statistical parameters summarized in Table 3 further confirm the superiority of PCA-ANN. RMSEP values decreased notably for CFN (0.2523 → 0.1319) and PRC (0.4353 → 0.2491), while SEP values were consistently lower for all analytes, demonstrating improved predictive robustness.

Comparative Computational and Predictive Performance

The tenfold reduction in input dimensionality (100 RAW variables vs. 10 orthogonal principal components) substantially influenced the bias-variance balance of the ANN models. By eliminating multicollinearity and redundant spectral information, PCA preprocessing improved numerical conditioning, reduced parameter variance, and minimized overfitting risk.

As summarized in Table 2, the PCA-ANN architecture employed a significantly reduced effective input space while maintaining identical training parameters. This dimensional compression translated directly into computational efficiency: training time decreased from 5.06 min (RAW-ANN) to approximately 1.06 min for PCA-ANN, corresponding to nearly a fivefold acceleration.

The training performance curves (Figures 3a and 3b) clearly demonstrate faster convergence and earlier stabilization for PCA-ANN. In contrast, RAW-ANN required longer iterative refinement due to the higher-dimensional and correlated input structure.

Overall, PCA-ANN exhibited improved predictive stability and lower validation errors compared to RAW-ANN, confirming the beneficial effect of dimensional reduction on ANN-based spectrophotometric calibration.

The improved performance of PCA-ANN was primarily associated with the reduced and orthogonalized input space rather than network complexity alone. While RAW-ANN utilized 100 correlated spectral variables, PCA-ANN used 10 principal component scores, resulting in improved numerical stability, faster convergence, and enhanced predictive robustness.

Table 3. Statistical parameters in the training and prediction steps

	RAW-ANN			PCA-ANN		
Analyte	SEC	RMSEP	SEP	SEC	RMSEP	SEP
CFN	0.0017	0.2523	0.2247	0.1071	0.1319	0.1280
PRPN	0.0008	0.3341	0.2990	0.1057	0.3236	0.1797
PRC	0.0012	0.4353	0.3955	0.1054	0.2491	0.2312

The standard error of training (SEC)
The standard error of prediction (SEP)
The root mean squared error of prediction (RMSEP)

Validation of RAW-ANN and PCA-ANN Methods

The predictive reliability of the developed models was evaluated using independent ternary validation mixtures containing CFN, PRPN, and PRC. The mean recovery values and relative standard deviations (RSD%) are summarized in Table 4.

For all analytes, recovery values were close to 100%, and RSD values were generally below 3%, demonstrating satisfactory trueness and precision. PCA-ANN exhibited slightly lower variability, particularly for CFN and PRC, indicating improved stability in prediction.

Further assessment using standard error of calibration (SEC), standard error of prediction (SEP), and root mean square error of prediction (RMSEP) (Table 3) confirmed the enhanced predictive performance of PCA-ANN. The consistently lower SEP values obtained with PCA-ANN indicate improved external prediction capability. Linear regression analysis between actual and predicted concentrations yielded slopes close to unity and high correlation coefficients, confirming the absence of systematic bias.

Table 4. Recovery results obtained by the application of RAW-ANN and PCA-ANN to the validation set consisting of CFN, PRPN, and PRC

Validation set			RAW-ANN			PCA-ANN			RAW-ANN			PCA-ANN		
Added (µg/mL)			Found (µg/mL)			Found (µg/mL)			Recovery (%)			Recovery (%)		
CFN	PRPN	PRC	CFN	PRPN	PRC	CFN	PRPN	PRC	CFN	PRPN	PRC	CFN	PRPN	PRC
3	9	3	2.98	9.01	2.83	3.05	9.03	2.83	99.4	100.1	97.3	101.6	100.3	94.5
3	9	8	2.90	9.44	8.06	2.93	9.25	8.06	96.8	104.9	102.2	97.8	102.8	100.8
3	9	12	2.92	9.17	11.87	2.98	9.24	11.87	97.3	101.9	96.9	99.2	102.7	98.9
3	9	16	2.96	9.12	15.86	2.99	9.43	15.86	98.6	101.4	97.0	99.6	104.7	99.1
2.5	9	15	2.48	8.82	14.99	2.45	9.44	14.99	99.1	98.0	99.3	97.9	104.9	100.0
4	9	15	4.19	8.94	14.95	4.08	9.3	14.95	104.7	99.3	99.4	102.1	103.4	99.7
8	9	15	7.64	9.24	15.49	7.55	9.34	15.49	95.5	102.7	102.3	94.4	103.8	103.2
12	9	15	12.53	8.59	15.58	12.02	9.20	15.58	104.4	95.4	105.6	100.1	102.2	103.8
3	3	15	3.06	2.92	15.31	3.04	2.98	15.31	101.9	97.2	103.3	101.3	99.5	102.1
3	6	15	2.96	5.76	15.2	3.00	5.91	15.20	98.6	95.9	99	99.9	98.4	101.3
3	9	15	3.03	9.17	15.13	3.03	9.40	15.13	100.9	101.9	103.6	100.9	104.5	100.9
3	12	15	2.89	12.50	14.89	2.93	12.24	14.89	96.4	104.2	96.6	97.8	102.0	99.3
3	9	15	3.17	8.72	15.02	3.04	9.21	15.02	105.7	96.8	103.2	101.5	102.3	100.1
								Mean	100	100	100.4	99.5	102.4	100.3
								SD	3.35	3.14	3.07	2.14	2.00	2.32
								RSD	3.35	3.14	3.06	2.15	1.95	2.31

SD: Standard Deviation, RSD: Relative Standard Deviation

As a result, the obtained SEC, SEP, RMSEP, recovery, and precision results confirmed the analytical reliability of both ANN-based training strategies, with PCA-ANN exhibiting superior predictive robustness and lower validation error.

Precision and Reproducibility

Intra-day and inter-day precision studies were conducted at three concentration levels for each analyte. The corresponding recovery, RSD, and relative standard error (RSE) values are presented in Table 5.

Table 5. The recovery results obtained in the intra-day and inter-day samples by the proposed RAW-ANN and PCA-ANN methods (n=3 for every concentration level)

				Found (µg/mL)					
	Added (µg/mL)			RAW-ANN			PCA-ANN		
	CFN	PRPN	PRC	CFN	PRPN	PRC	CFN	PRPN	PRC
Inter-day	3	4	5	3.02	3.96	4.98	2.93	4.08	4.98
	6	12	10	5.83	12.08	9.65	5.98	12.31	9.76
	9	16	15	8.77	16.35	14.93	9.02	16.04	14.84
Intra-day	3	4	5	3.00	3.97	4.97	2.95	4.04	4.96
	6	12	10	5.88	12.54	9.71	5.98	12.33	9.72
	9	16	15	8.93	16.31	14.28	9.04	16.12	14.82
				Recovery (%)					
				CFN	PRPN	PRC	CFN	PRPN	PRC
Inter-day				100.7	98.9	99.5	97.8	101.9	99.5
				97.2	100.7	96.5	99.7	102.6	97.6
				97.5	102.2	99.6	100.3	100.3	98.9
Intra-day				100.0	99.2	99.5	98.2	101.0	99.3
				97.9	104.5	97.1	99.6	102.7	97.2
				99.3	101.9	95.2	100.5	100.7	98.8
				RSD (%)					
				CFN	PRPN	PRC	CFN	PRPN	PRC
Inter-day				4.15	3.54	2.35	0.52	1.44	1.18
				0.45	2.82	1.04	0.59	0.49	1.65
				1.6	1.43	1.44	1.14	0.2	1.34
Intra-day				4.18	3.66	0.81	0.78	3.43	1.23
				0.71	2.47	3.42	1.19	0.33	0.98
				2.19	2.34	2.42	0.61	0.19	0.58
				RE (%)					
				CFN	PRPN	PRC	CFN	PRPN	PRC
Inter-day				0.67	-1.08	-0.47	-2.22	1.92	-2.22
				-2.83	0.69	-3.47	-0.28	2.58	-0.28
				-2.52	2.19	-0.44	0.26	0.27	0.26
Intra-day				0	-0.83	-0.53	-1.78	1	-1.78
				-2.06	4.53	-2.93	-0.39	2.72	-0.39
				-0.74	1.94	-4.78	0.48	0.73	0.48

RSD: Relative Standard Deviation, RE: Relative Error

RSD values remained consistently low for both models, demonstrating good repeatability and intermediate precision. Notably, PCA-ANN exhibited reduced inter-day variability, suggesting greater robustness against minor instrumental fluctuations and environmental variations.

Selectivity and Standard Addition Studies

Selectivity and potential matrix effects were evaluated using the standard addition method applied to commercial tablet solutions. Known amounts of CFN, PRPN, and PRC (0-6 µg/mL) were added and analyzed in triplicate. The recovery values summarized in Table 6 confirm that common excipients present in the pharmaceutical formulation did not interfere with the quantitative determination of the analytes. Low RSD values further demonstrate predictive stability in complex matrices. PCA-ANN again showed slightly improved consistency, reinforcing the beneficial impact of dimensional reduction on model conditioning.

Table 6. The recovery results obtained in the standard addition samples by the proposed RAW-ANN and PCA-ANN methods

	Added (µg/mL)			**Found (µg/mL)**					
				RAW-ANN			**PCA-ANN**		
	CFN	**PRPN**	**PRC**	**CFN**	**PRPN**	**PRC**	**CFN**	**PRPN**	**PRC**
Formulation	2	2	2	1.99	2.03	2.00	1.93	2.04	2.00
Formulation	4	4	4	3.94	4.02	3.91	3.96	4.02	3.89
Formulation	6	6	6	5.88	5.96	5.89	6.10	5.98	5.94
				Recovery (%)					
				RAW-ANN			PCA-ANN		
				CFN	PRPN	PRC	CFN	PRPN	PRC
				99.3	101.3	100.2	96.3	102.0	100.2
				98.5	100.6	97.8	98.9	100.5	97.3
				98.0	99.4	98.2	101.7	99.7	99.1
				RSD (%)					
				RAW-ANN			PCA-ANN		
				CFN	PRPN	PRC	CFN	PRPN	PRC
				1.04	3.79	1.04	2.52	1.80	2.02
				2.29	2.32	2.04	1.42	1.39	0.50
				2.85	3.03	2.61	0.35	2.41	1.95

RSD: Relative Standard Deviation

Analysis of Commercial Ternary Tablet Formulation

The developed models were applied to the simultaneous determination of CFN, PRPN, and PRC in commercial tablet formulations. The predicted content values (Table 7) were in close agreement with labeled amounts. Statistical parameters, including standard deviation and percent RSD, confirmed the analytical reliability of both methods. Despite strong spectral overlap and the presence of pharmaceutical excipients, accurate and precise quantification was achieved without any separation step.

Table 7. Analysis results obtained by applying RAW-ANN and PCA-ANN to tablets

	mg/tablet					
	RAW-ANN			PCA-ANN		
Exp No.	**CFN**	**PRPN**	**PRC**	**CFN**	**PRPN**	**PRC**
1	50.0	148.7	249.3	49.5	149.4	248.3
2	50.8	147.9	256.2	53.3	144.9	251.2
3	50.3	144.2	240.4	50.8	143.2	243.8
4	48.7	140.8	262.6	48.7	142.3	256.2
5	47.7	144.5	259.7	49.7	148.8	255.7
6	50.0	149.3	249.3	49.5	149.5	248.3
7	50.7	144.3	249.8	50.0	151.7	251.0
8	47.0	147.5	255.3	51.0	150.8	256.0
9	51.2	149.0	255.2	50.2	153.2	252.8
10	48.2	154.7	261.6	50.2	154.2	255.3
Mean	49.5	147.1	253.9	50.3	148.5	251.9
SD	1.46	3.84	6.81	1.27	3.06	4.13
RSD	2.95	2.61	2.68	2.52	2.06	1.64

SD: Standard Deviation, RSD: Relative Standard Deviation,
Label claim: 50 mg CFN/150 mg PRPN/250 mg PRC per tablet)

CONCLUSIONS

This study demonstrates that nonlinear ANN-based training strategies can successfully resolve highly collinear and severely overlapping UV spectral data without requiring prior chromatographic separation. By comparatively evaluating RAW-ANN and PCA-ANN architectures, the present work systematically illustrates how input dimensional organization directly influences predictive stability, computational efficiency, and model generalization in multicomponent spectrophotometric analysis.

Although both ANN strategies provided accurate and precise quantification of CFN, PRPN, and PRC in ternary pharmaceutical formulations, the incorporation of principal component compression prior to ANN modeling fundamentally improved the balance between model complexity and predictive robustness. Transforming correlated spectral variables into orthogonal principal components enhanced numerical conditioning, reduced overparameterization risk, accelerated convergence, and improved validation performance, while maintaining analytical accuracy.

Beyond the specific formulation investigated, the findings underscore a broader chemometric principle: in high-dimensional analytical systems, structured dimensional reduction can be as critical as the nonlinear modeling algorithm itself. The PCA-ANN framework presented herein offers a computationally efficient, scalable, and transferable strategy for addressing multicollinearity in spectroscopic datasets and provides a practical alternative for routine quality control analysis of complex multicomponent pharmaceutical preparations.

EXPERIMENTAL SECTION

Instruments and software

A Shimadzu UV-2550 UV-VIS spectrophotometer (Kyoto, Japan), equipped with Shimadzu UVPC software, was used to record absorption spectra in the wavelength range of 200-300 nm ($\Delta\lambda$ = 0.1 nm). Although the spectra were recorded over the 200-300 nm range, only the 220-300 nm region was used for RAW-ANN and PCA-ANN modeling in order to minimize potential interference from the solvent and pharmaceutical matrix components at lower wavelengths. Data acquisition and statistical analyses were performed using MATLAB 7.0 (MathWorks, USA) and Microsoft Excel. Custom-written m-files were developed within the MATLAB environment for the implementation of the RAW-ANN and PCA-ANN training models.

Chemicals and reagents

Reference standards of caffeine (CFN, 99.7% purity), propyphenazone (PRPN, ≥99% purity), and paracetamol (PRC, 99.6% purity) were obtained from Sigma-Aldrich (Steinheim, Germany). Analytical-grade methanol suitable for gradient applications was also purchased from Sigma-Aldrich (USA) and used as the solvent throughout the study. All working solutions were freshly prepared on a daily basis and protected from light during analysis to maintain chemical stability.

The commercial tablet formulation Minoset® Plus 30 (Bayer Türk Chemical Industry Ltd., Ümraniye, Istanbul, Turkey) was selected for pharmaceutical analysis. Each tablet is labeled to contain 50 mg CFN, 150 mg PRPN, and 250 mg PRC, along with common pharmaceutical excipients such as microcrystalline cellulose, hydroxypropyl methylcellulose, formaldehyde casein (Esma spreng), corn starch, talc, magnesium stearate, and colloidal anhydrous silica.

Preparation of stock standard, calibration, and validation solutions

Individual stock standard solutions of caffeine (CFN), propyphenazone (PRPN), and paracetamol (PRC) were prepared by accurately weighing 10.0 mg of each compound and dissolving them separately in methanol using 100 mL volumetric flasks. These primary stock solutions were subsequently diluted with methanol to obtain working solutions required for calibration, validation, and precision studies.

For the construction of the RAW-ANN and PCA-ANN training models, a total of 35 calibration mixtures were prepared within the concentration ranges of 2.5-12.0 µg/mL for CFN, 3.0-12.0 µg/mL for PRPN, and 3.0-16.0 µg/mL for PRC (Table 1).

Additionally, thirteen synthetic ternary mixtures covering different concentration combinations within the working range were prepared to evaluate model performance (Table 2). Precision studies were carried out by preparing intra-day and inter-day samples at three concentration levels for each analyte. The selected concentration levels were 3.0, 6.0, and 9.0 µg/mL for CFN; 4.0, 12.0, and 16.0 µg/mL for PRPN; and 5.0, 10.0, and 15.0 µg/mL for PRC (Table 4).

For accuracy assessment via the standard addition method, known amounts of CFN, PRPN, and PRC standard solutions (0, 2.0, 4.0, and 6.0 µg/mL) were spiked into aliquots of the commercial tablet solution. Each level was prepared in triplicate.

Preparation of Tablet Samples

Ten tablets of Minoset® Plus (Bayer Türk Chemical Industry Ltd.) were accurately weighed, thoroughly crushed in a mortar until a homogeneous mixture, and the quantity corresponding to one-twentieth of a tablet was transferred to a 100 mL volumetric flask after thorough pulverization of the tablets in a mortar. The flask was filled with methanol, and the powdered sample was mixed with a mechanical stirrer for one hour. Filtration was conducted through a 0.45 µm pore size nylon filter (ISOLAB, Wertheim,

Germany). The resulting solution was diluted with methanol to reach a concentration of 3.0 µg/ mL CFN, 9.0 µg/ mL PRPN, and 15 µg/ mL PRC. The diluted tablet solution was subjected to UV-VIS analysis for the application of the RAW-ANN and PCA-ANN methods.

ACKNOWLEDGMENTS

This work was carried out at the Chemometric Laboratory of Ankara University, Faculty of Pharmacy funded by the scientific research project number 10A3336001.

REFERENCES

1. K. Palur; S. C. Archakam; B. Koganti; *Spectrochim. Acta A Mol. Biomol. Spectrosc.*, **2020**, *243*, 118801.
2. V. Boerlin; B. Maeglin; W. Hägler; M. Kuhn; E. Nüesch; *Eur. J. Clin. Pharmacol.*, **1986**, *31*, 127-131.
3. J. K. Aronson; Meyler's Side Effects of Drugs: The International Encyclopedia of Adverse Drug Reactions and Interactions (Fifteenth Edition), Elsevier Science, Amsterdam, **2006**, pp. 2954-2955.
4. V. S. Reddy; S. Shiva; S. Manikantan; S. Ramakrishna; "Pharmacology of caffeine and its effects on the human body". *European Journal of Medicinal Chemistry Reports*. 2 March **2024**, *10,* 100138.
5. G. Faudone; S. Arifi; D. Merk; *Journal of Medicinal Chemistry*, **2021**, *64 (11),* 7156-7178.
6. A. Rajput; R. Sharma; R. Bharti; *Materials Today: Proceedings,* **2022**, *48,* 1407-1415.
7. L. C. Adam; D. Repantis; N. Boris; B. N. Konrad; M. Dresler; S. Kühn; *Brain and Cognition,* **2021***, 154*, 105802.
8. K. Delvadiya; P. Kabra; R. Kimbahune; N. Patel; L. V. G. Nargund; *Indian J Pharm Educ Res*, **2013***, 47(4)*, 65-72.
9. F. Soponar; D. Staniloae; G. Moise; B. Szaniszlo; V. David; *Rev. Roum. Chim*, **2013**, *58(4-5)*, 433-440.
10. H. Ibrahim; A. M. Hamdy; H. A. Merey; A. S. Saad; *J. Chromatogr. Sci.*, **2021**, *59(2)*, 140-147.
11. H. Tománková; M. Vasatova; *Pharmazie*, **1989**, *44(3)*, 197-198.
12. N. Kosmeas; J. T. Clerc; *Pharm. Acta Helv.*, **1989**, *64(1)*, 2-7.
13. A. Hałka-Grysińska; P. Ślązak; G. Zaręba; W. Markowski; A. Klimek-Turek; T. H. Dzido; *Anal. Methods*, **2012**, *4(4)*, 973-982.
14. K. Bober-Majnusz; A. Pyka-Pająk; *Processes*, **2024**, *12(6)*, 1153.
15. S. Markovic; Z. Kusec; *Pharmazie*, **1990**, *45(12)*, 935-936.
16. R. Injac; B. Srdjenovic; M. Prijatelj; M. Boskovic; K. Karljikovic-Rajic; B. Strukelj; *J. Chromatogr. Sci.*, **2008**, *46(2)*, 137-143.

17. W.P. Silva; L.A.J. Silva; C.H. França; R.M.F. Sousa; R.A.A. Muñoz; E.M. Richter; *Electroanalysis*, **2017**, *29*, 1860-1866.
18. E. Dinç; G. Kökdil; F. Onur; *J. Pharm. Biomed. Anal.*, **2001**, *26(5-6)*, 769-778.
19. A. F. Pulungan; D. L. P. Effendy; M. S. Siti; *AJPRD*, **2018**, *6(5)*, 5-8.
20. H. Rosaldi; A. F. Pulungan; R. Ridwanto; A. S. Daulay; *IJSP*, **2023**, *1(2)*, 74-79.
21. G. Gergov; A. Alin; P. Katsarov; V. Simeonov; D. Yankov; Y. Al-Degs; *Bulg. Chem. Commun*., **2018**, *50*, 265-273.
22. M. Ü. Özgür; G. Alpdoğan; B. Aşçi; *Monatsh. Chem.*, **2002**, *133(2)*, 219-223.
23. A. Domínguez Vidal; P. Ortega Barrales; A. Molina Díaz; *Microchim. Acta*, **2003**, *141(3)*, 157-163.
24. J. M. Xia; X. J. Wu; Y. J. Yuan; *Metabolomics*, **2007**, *3(4)*, 531-537.
25. G. Absalan; M. Soleimani; *Anal. Sci.*, **2004**, *20(5)*, 879-882.
26. G. Ioele; M. De Luca; E. Dinç; F. Oliverio; G. Ragno; *Chem. Pharm. Bull.*, **2011**, *59(1)*, 35-40.
27. E. Dinç; D. Baleanu; N. Ş. Köktaş; *Rev. Anal. Chem.,* **2011**, 30 (1), 11-15.