spectroscopy journal

MDPI

Article

Simultaneous Determination of CH_4, C_2H_6 and C_2H_4 Mixtures Using MCPSO-Optimized DKELM

Pengcheng Gu [1], Meixuan Zhao [1], Xinyu Tian [2] and Yuwang Han [1,*]

[1] College of Chemistry and Molecular Engineering, Nanjing Tech University, Nanjing 210009, China; 202361205187@njtech.edu.cn (P.G.); 202361105017@njtech.edu.cn (M.Z.)
[2] Data Science and Information Engineering College, Harbin Huade University, Harbin 150025, China; 20230815286@hhdu.edu.cn
* Correspondence: hanyw@njtech.edu.cn; Tel.: +86-139-1399-6723

Abstract

Photoacoustic spectroscopy (PAS) is a highly sensitive and non-destructive technique widely used for trace gas detection; however, the simultaneous quantification of methane (CH_4), ethane (C_2H_6), and ethylene (C_2H_4) remains challenging due to severe spectral cross-interference and non-linear responses across broad concentration ranges. In this work, we propose a high-precision, end-to-end detection framework based on a Deep Kernel Extreme Learning Machine (DKELM) optimized using a Mutation–Chaotic Particle Swarm Optimization (MCPSO) algorithm. To enhance diagnostic information in the photoacoustic signals, a multi-scale wavelet transform based on a db4 wavelet basis with 5-layer decomposition and a Heursure soft threshold strategy is first employed for denoising and enhancing absorption features. To address the hyperparameter sensitivity and local-optimum trapping inherent in deep models, the MCPSO algorithm integrates hybrid chaotic initialization, adaptive mutation probability control, Cauchy-based perturbation, temperature-controlled mutation amplitude, and elite-guided population updating. The proposed MCPSO-DKELM model is evaluated on an expanded dataset of 470 mixed-gas spectra and benchmarked against other frameworks, including the previously reported SVM-CPSO-KELM architecture. The experimental results demonstrate that MCPSO-DKELM achieves stable, segmentation-free quantification across the full dynamic range, with an average detection error below 3.5% and the maximum relative error constrained to under 15%, which represents a substantial improvement over existing approaches. Thus, the combination of deep kernel feature extraction and mutation–chaotic global optimization provides a robust and reliable solution for simultaneous multi-component hydrocarbon gas analysis in complex industrial environments.

Keywords: photoacoustic spectroscopy; mixed-gas analysis; deep kernel extreme learning machine; mutation–chaotic particle swarm optimization; spectral cross-interference

Check for updates

Academic Editor: Vincent Boudon

Received: 20 April 2026
Revised: 2 June 2026
Accepted: 23 June 2026
Published: 24 June 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The simultaneous monitoring of light hydrocarbons—specifically methane (CH_4), ethane (C_2H_6), and ethylene (C_2H_4)—is a critical operational requirement in the petrochemical [1] and energy sectors [2]. These gases frequently coexist as primary components in natural gas extraction [3], as key intermediates in catalytic cracking pathways [4], and as diagnostic indicators of fault evolution in industrial safety systems [3]. While methane dominates both fuel composition and greenhouse gas considerations [5], variations in the

relative concentrations of ethane and ethylene strongly influence the calorific value of gas mixtures and directly affect the efficiency and selectivity of downstream chemical conversion processes [6,7]. Consequently, the ability to resolve the individual concentrations of these components within a dynamic mixture is essential for both process optimization and environmental compliance [6,8].

Photoacoustic spectroscopy (PAS) has emerged as a leading technique for trace gas analysis as it offers clear advantages over electrochemical and gas chromatography methods [9]. By detecting the acoustic waves generated when molecules undergo non-radiative relaxation following modulated light absorption, PAS operates with an intrinsically low background, which in turn enables exceptionally high sensitivity and a broad linear dynamic range. With the adoption of narrow-band laser sources, this acoustic detection scheme also provides strong molecular selectivity [10]. The versatility of PAS has been demonstrated across diverse multi-component sensing applications, including the precise detection of nitrogen dioxide in environmental and industrial settings, as well as the simultaneous measurement of acetylene, methane, and carbon dioxide [11,12]. Consequently, PAS has been proven to be a robust technology for high-precision, rapid-response gas sensing in demanding operational environments [13].

However, applying PAS to light hydrocarbon mixtures introduces several challenges. The mid-infrared absorption bands of CH_4, C_2H_6, and C_2H_4 overlap heavily in the C–H stretching region near 3.3 μm [14], which produces correlated spectral features and nonlinear cross-interference in the resulting photoacoustic signals [15]. To address this problem, researchers have explored approaches such as selective optical filtering and high-resolution TDLAS-PAS systems [16,17], and more recently have constructed hybrid chemometric models capable of handling the underlying spectral nonlinearity [18,19]. In our previous work [20], we substantially improved prediction accuracy by integrating a Support Vector Machine (SVM) pre-classification stage with a Kernel Extreme Learning Machine (KELM) optimized via Chaotic Particle Swarm Optimization (CPSO). In this SVM–CPSO–KELM framework, the SVM stage partitions samples into high- and low-concentration regimes to enable more specialized local regression. However, this SVM pre-classification step, while effective, creates a disjointed architecture. Any error in the initial classification would propagate through the entire system, and the rigidity of threshold-based segmentation limits the model's ability to adapt smoothly to continuous concentration variations in real-world gas streams. There is a pressing need for a unified end-to-end modeling approach capable of resolving complex cross-interference across the full dynamic range without relying on hard segmentation.

Deep Kernel Extreme Learning Machine (DKELM) offers a promising solution by extending the single-hidden-layer structure of standard KELM into a multi-layer architecture [21]. Because it stacks successive kernel mappings, DKELM can extract increasingly abstract spectral representations that remain robust against noise and concentration-dependent variability, which potentially eliminates the need for explicit pre-classification [22]. However, the performance of deep models such as DKELM depends critically on the careful tuning of hyperparameters, including layer-specific penalty coefficients and kernel parameters [23]. Conventional parameter-search strategies and common metaheuristics, such as Particle Swarm Optimization (PSO) and Gray Wolf Optimization (GWO), often struggle in these high-dimensional spaces where premature convergence and local-optimum trapping can prevent the network from reaching its full generalization capability. To address these limitations, researchers have incorporated chaotic mapping [20] and mutation operators as enhancement mechanisms. Chaotic mappings exploit the ergodicity and non-periodicity of chaotic sequences to generate more uniformly distributed initial populations, which provides a stronger foundation than purely random initialization.

Mutation operators, in turn, introduce controlled perturbations that restore population diversity when candidate solutions begin to cluster around suboptimal regions. These two strategies can also be combined to improve the robustness of the search process: chaos promotes broad global exploration in the early stages, whereas mutation offers a critical escape mechanism during later-stage exploitation.

In this work, we adopted DKELM as the core regression engine to provide a deep, unified learning architecture for the detection of mixed light hydrocarbons. For optimization, we implemented and tested four distinct swarm intelligence algorithms: CPSO and CGWO, and their mutation-enhanced versions, MCPSO and MCGWO. All algorithms share identical parameters: population size of 50, maximum iteration of 100, and convergence criterion of fitness variation less than 1×10^{-6}. This allows us to rigorously evaluate the impact of chaotic initialization and the novelty of mutation-based diversity preservation on model performance. We applied this framework to an expanded dataset of 470 mixed-gas photoacoustic spectra (a significant increase compared to the 130-spectrum dataset in our prior study) and performed a head-to-head comparison against our previous SVM–CPSO–KELM model, to demonstrate the superiority of the deep, non-segmented approach in achieving high-precision, simultaneous detection of CH_4, C_2H_6, and C_2H_4.

2. Algorithms

2.1. Deep Kernel Extreme Learning Machine (DKELM)

The Deep Kernel Extreme Learning Machine (DKELM) extends the Kernel Extreme Learning Machine (KELM) into a multi-layer architecture [24]. KELM is originally derived from Extreme Learning Machine (ELM) [25]; it employs kernel functions to map data into a high-dimensional feature space, which eliminates the dependency on the number of hidden layer nodes and improves stability and generalization [26,27]. However, the performance of KELM depends critically on the regularization parameter C, which controls the trade-off between fitting the training data and maintaining model smoothness, and the kernel parameter γ, which determines the locality and complexity of the induced feature space. Large C or large γ increases model flexibility and can lead to overfitting, whereas small C or small γ produces smoother decision functions and may result in underfitting [28,29]. In contrast, DKELM stacks kernel mappings to achieve layer-by-layer abstraction and deep feature extraction. With this modified architecture, it is particularly suitable for high-dimensional spectral analysis as it retains the advantage of KELM in the analytical solution and captures complex non-linear relationships more effectively [22,30].

2.2. Mutation–Chaotic Particle Swarm Optimization (MCPSO)

MCPSO extends the classical Chaotic Particle Swarm Optimization (CPSO) framework by integrating a multi-level mutation control system designed to counteract premature convergence and preserve population diversity [31,32]. In CPSO, population initialization and parameter adaptation are driven by hybrid chaotic mapping, typically through alternating Logistic and Tent maps, to enhance ergodicity and improve early-stage exploration [33]. In MCPSO, this chaotic architecture is retained, and several important additions are introduced: (1) adaptive mutation probability adjustment based on population fitness distribution, (2) Cauchy-based perturbation for stagnating particles, (3) a temperature-controlled mutation amplitude strategy inspired by simulated annealing, and (4) a population update mechanism that combines elite retention with fission-based directional replacement. Together, these additions substantially extend the exploratory capacity of CPSO and reinforce its ability to maintain population diversity. By combining the global ergodicity of chaotic dynamics with the strong local escape capability of mutation operators, MCPSO achieves a balanced exploration–exploitation process that supports

rapid convergence while remaining resilient to local optima in complex, multi-modal, and high-dimensional optimization settings.

2.3. Mutation–Chaotic Gray Wolf Optimization (MCGWO)

MCGWO is a swarm intelligence method that incorporates chaos theory and mutation strategies into the Chaotic Gray Wolf Optimization (CGWO) algorithm [34,35]. It retains the chaotic initialization of CGWO to ensure uniform search space coverage, but to mitigate the tendency of the standard algorithm to stagnate in later iterations, it introduces an adaptive mutation framework that dynamically triggers operations based on population fitness diversity. Specifically, it applies Lévy flight perturbations to leading individuals (α, β, δ wolves) to enhance exploration and applies polynomial mutations to subordinate individuals (ω wolves) to refine the local search. Furthermore, the mutation amplitude is coupled with the convergence factor a, which enables adaptive adjustment from large-scale exploration in early iterations to fine-grained exploitation in later stages.

3. Experimental Section

3.1. Configuration of Spectral Acquisition System

Figure 1 shows the system for spectral acquisition. To cover the required spectral range, a continuous-wave Interband Cascade Laser (ICL; Nanoplus, Gerbrunn, Germany) centered at 3345 nm served as the excitation source [20]. An LDC-3714C laser driver (ILX Lightwave; Bozeman, MT, USA) maintained the operating temperature at 20 °C and modulated the injection current between 20 and 100 mA, thus providing a wavelength coverage of 3344–3355 nm. The ICL, packaged in a TO66 base with a collimating lens, emitted a Gaussian beam with a 3 mm spot diameter. The acoustic signal was detected by a BO9750CSC electret microphone (sensitivity -26 ± 3 dB). The Dual-Helmholtz Resonance (DHR) photoacoustic cell was custom-built from 304 stainless steel. The main tube was 50 mm in length and 5 mm in inner diameter, and the capillary was 50 mm in length and 2 mm in inner diameter. The total volume was approximately 2.5 mL. The optical path was sealed with 0.5 mm thick calcium fluoride windows [20].

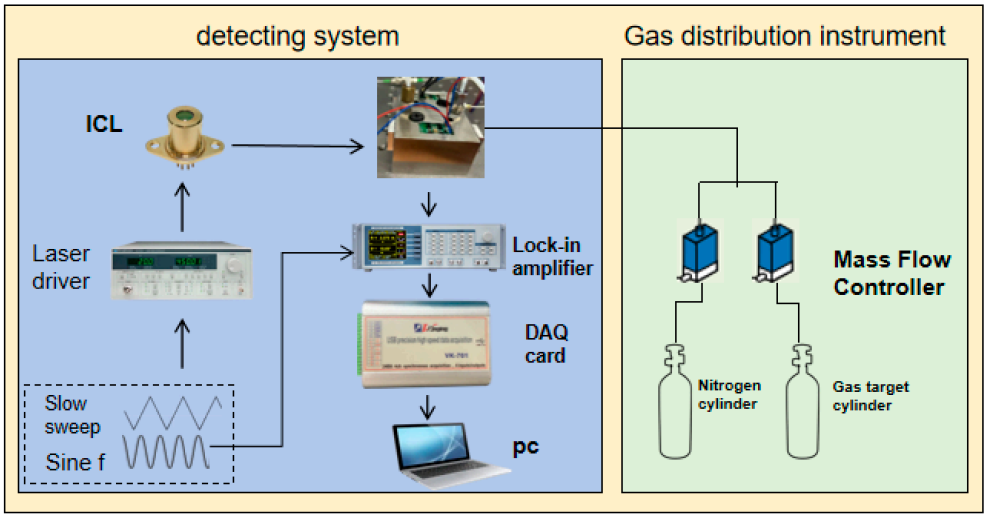


Figure 1. Schematic diagram of the photoacoustic spectroscopy detection system.

3.2. Preparation of Mixed-Gas Samples

A dynamic gas distribution system generated 470 distinct mixtures using high-purity nitrogen as the background gas. Target analytes were diluted with nitrogen via mass flow controllers (MFCs) and mixed at a T-junction to ensure homogeneity. The systematic error

of the MFCs was within 2%. To ensure robust model training, the dataset covered a broad concentration range with randomized profiles (Table S1):

- **Methane (CH_4):** 0–10 ppm (109 samples), 10–100 ppm (202 samples), 100–504 ppm (159 samples).
- **Ethane (C_2H_6):** 0–10 ppm (69 samples), 10–100 ppm (293 samples), 100–266 ppm (108 samples).
- **Ethylene (C_2H_4):** 0–10 ppm (38 samples), 10–100 ppm (145 samples), 100–504 ppm (287 samples).

Sample distribution was generally uniform, except that high-concentration ethane samples were limited to prevent data acquisition saturation due to high sensitivity, and low-concentration ethylene samples were limited due to low sensitivity.

3.3. Data Processing

A workflow based on the MCPSO-DKELM model was adopted to accurately quantify the gas mixtures (Figure 2). Photoacoustic spectral signals were acquired for the 470 samples, with each period containing 10,000 sampling points in the 3343–3347 nm range. To enhance the signal-to-noise ratio (SNR), signals from 16 cycles were averaged. The captured signals were preprocessed using a multi-scale wavelet transform to suppress environmental noise (Section 4.1). The processed spectra were then input into the DKELM model for training and validation.

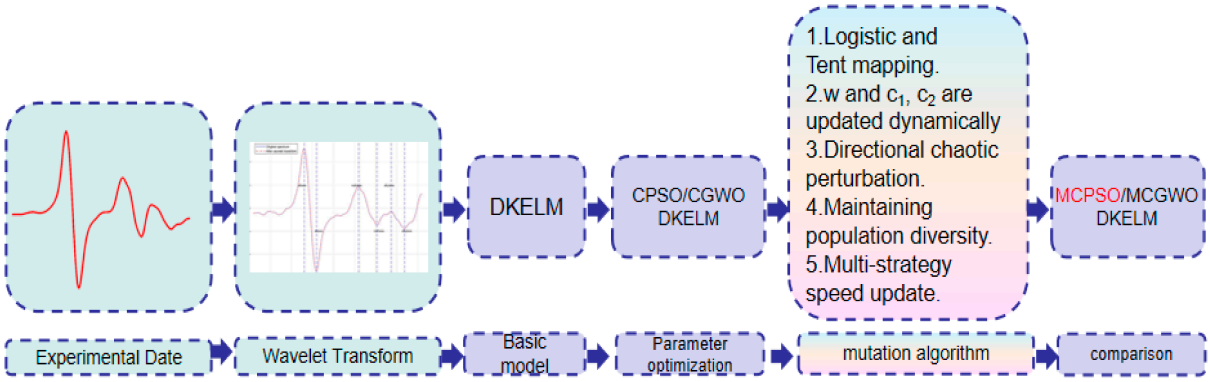


Figure 2. Data processing workflow.

3.4. Evaluation Metrics

The performance of different models is assessed using the following metrics:

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

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (2)$$

$$e_i = \left(\frac{|\hat{y}_i - y_i|}{y_i}\right) \times 100\% \quad (3)$$

$$e_{avg} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{|\hat{y}_i - y_i|}{y_i}\right) \times 100\% \quad (4)$$

where RMSE is the root-mean-square error, R^2 is the coefficient of determination, e_i is the elative error, e_{avg} is the mean percentage error, y_i is the ith true value, $\hat{y}_i$ is the ith predicted value, $\bar{y}$ is the mean of the true values, and n is the sample size. RMSE measures the absolute deviation between predicted and true values, with lower values indicating higher

accuracy [36]. R^2 evaluates the goodness of fit of the model; its value ranges from 0 to 1, and values closer to 1 indicate a better correlation between predicted and actual values [37]. Relative error quantifies the percentage deviation for each individual prediction, providing a direct measure of accuracy per sample. These metrics together provide a comprehensive assessment of model performance.

4. Results

4.1. Signal Preprocessing and Feature Extraction

To mitigate noise arising from environmental factors, temperature fluctuations, and humidity, the raw photoacoustic signals were preprocessed using a multi-scale wavelet transform [38,39]. This technique decomposed the non-stationary spectral signals to separate high-frequency noise components from the fundamental absorption features [40,41]. Analysis of the pure component spectra relative to the HITRAN database allowed for the identification of key absorption features despite the complex overlap in the gas mixture (Figures 3 and 4). Specific characteristic wavelengths were isolated to maximize selectivity:

- **Methane (CH_4):** 3345.6 nm and 3346.03 nm.
- **Ethane (C_2H_6):** 3344.31 nm and 3344.6 nm.
- **Ethylene (C_2H_4):** 3346.17 nm and 3346.69 nm.

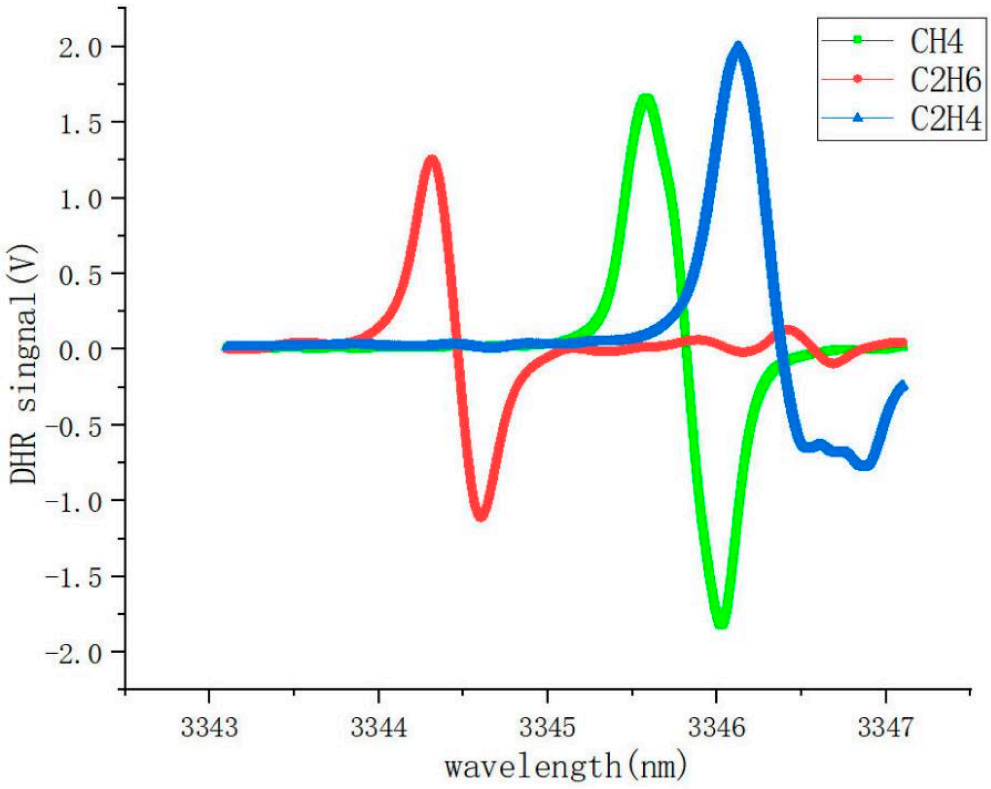


Figure 3. Photoacoustic spectra of pure gases.

While ethylene exhibits diffuse spectral bands lacking sharp peaks, the selected wavelengths at 3346.17 nm and 3346.69 nm provided sufficiently distinct absorption profiles for effective discrimination. The application of wavelet denoising significantly enhanced the signal-to-noise ratio at these characteristic points, establishing a high-quality dataset for the subsequent deep learning model.

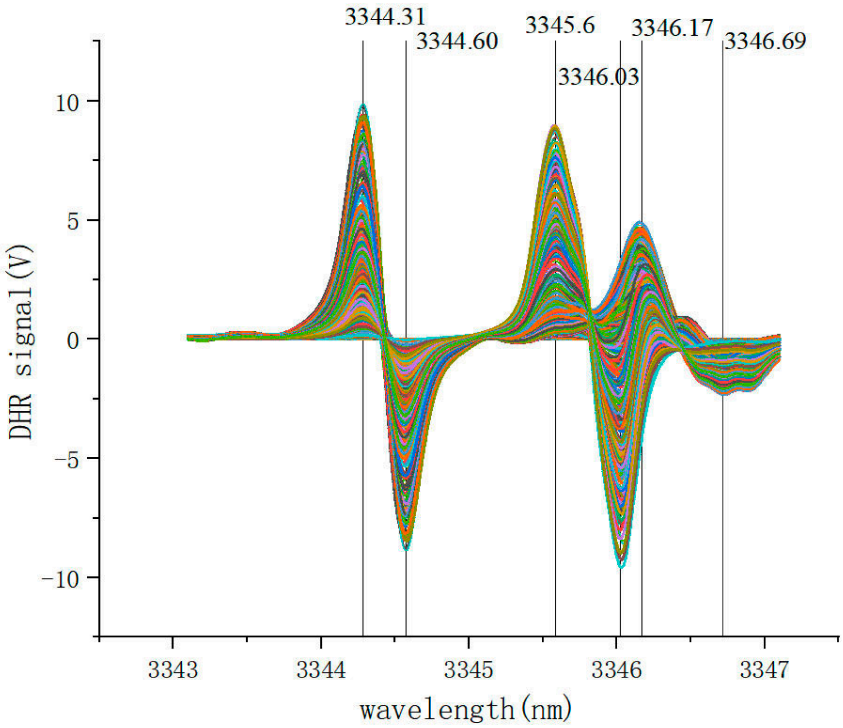


Figure 4. Original spectra of the 470 samples.

4.2. Baseline Performance of Standalone DKELM

To construct a robust and reliable model, the preprocessed spectral data were divided using the Kennard–Stone (KS) algorithm to ensure consistent data distribution between the training and test sets. The samples were split into an 8:2 ratio, with 80% (376 samples) used for training and 20% (94 samples) reserved as an independent test set (Table S1). The precise indices of the split were recorded to ensure experimental repeatability and comparability across different algorithmic iterations.

A DKELM model with a two-layer forward structure was constructed, which comprised hidden layers with 6000 and 5000 nodes respectively. The hyperbolic tangent function (tanh) served as the activation function, and the output layer utilized ridge regression based on a degree-1 polynomial kernel. The regularization penalty factor was set to 100. These node numbers were determined by 5-fold cross-validation with RMSE as the evaluation index within the range of 1000–10,000, and 6000 and 5000 nodes achieved optimal regression performance without obvious overfitting.

The analysis of the training set fitting curves (Figure 5) revealed noticeable dispersion in the predicted concentrations of CH_4, C_2H_6, and C_2H_4, although there were no extreme outliers (Table S2). The magnitude of error was inversely proportional to sample concentration; low-concentration samples exhibited disproportionately large deviations. The significant variance between maximum and minimum errors indicates a deficiency in learning stability and limited adaptability of the base model.

The performance of the base DKELM model on the independent test set (Figure 6) further confirmed these limitations. Prediction deviations became more pronounced in the high-concentration regime and prediction accuracy remained insufficient at lower concentrations (Table S3).

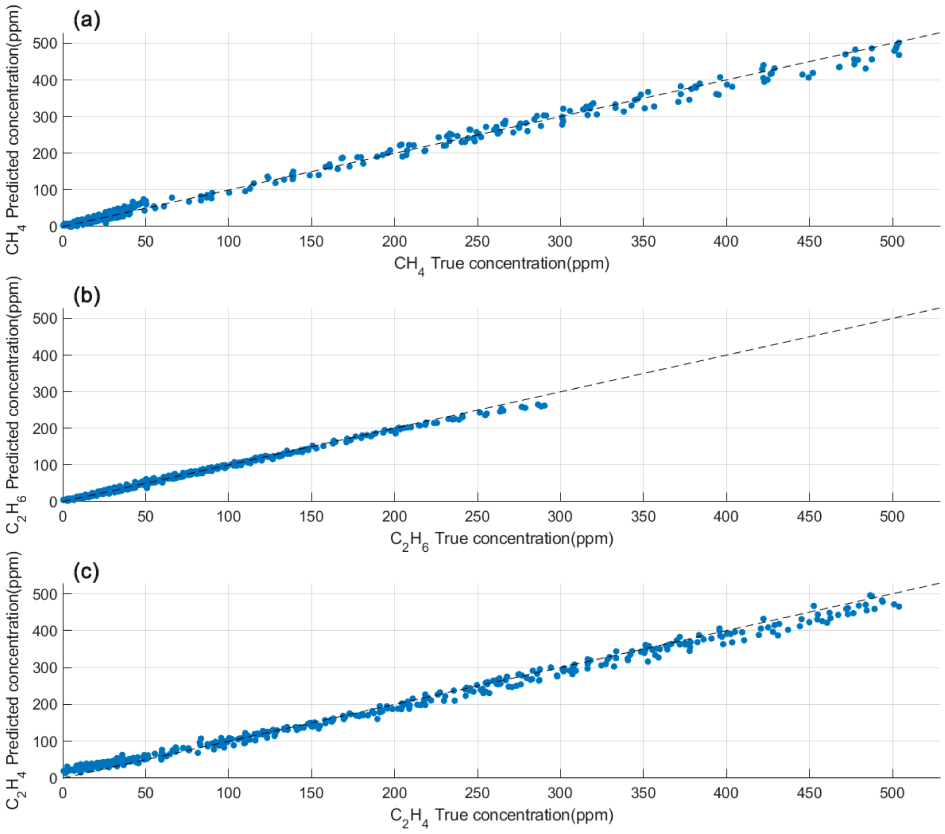


Figure 5. Fitting curves of the base DKELM model on the training set: (**a**) methane, (**b**) ethane, (**c**) ethylene.

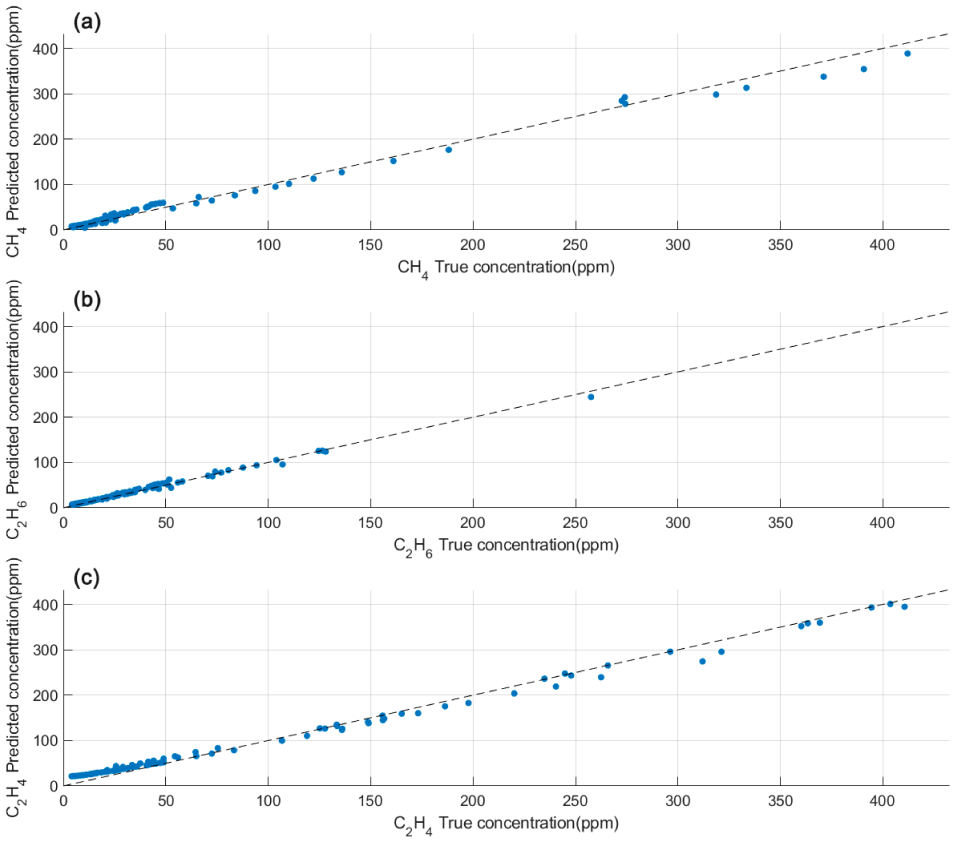


Figure 6. Fitting curves of the base DKELM model on the test set: (**a**) methane, (**b**) ethane, (**c**) ethylene.

As illustrated in the line charts (Figure 7), the base DKELM model struggled to track abrupt concentration fluctuations, and the maximum relative errors reached 202.70%. This indicates poor model stability and suggests that the kernel and regularization parameters were not optimally tuned for complex interference environments, thus causing weak generalization and lack of resistance to noise.

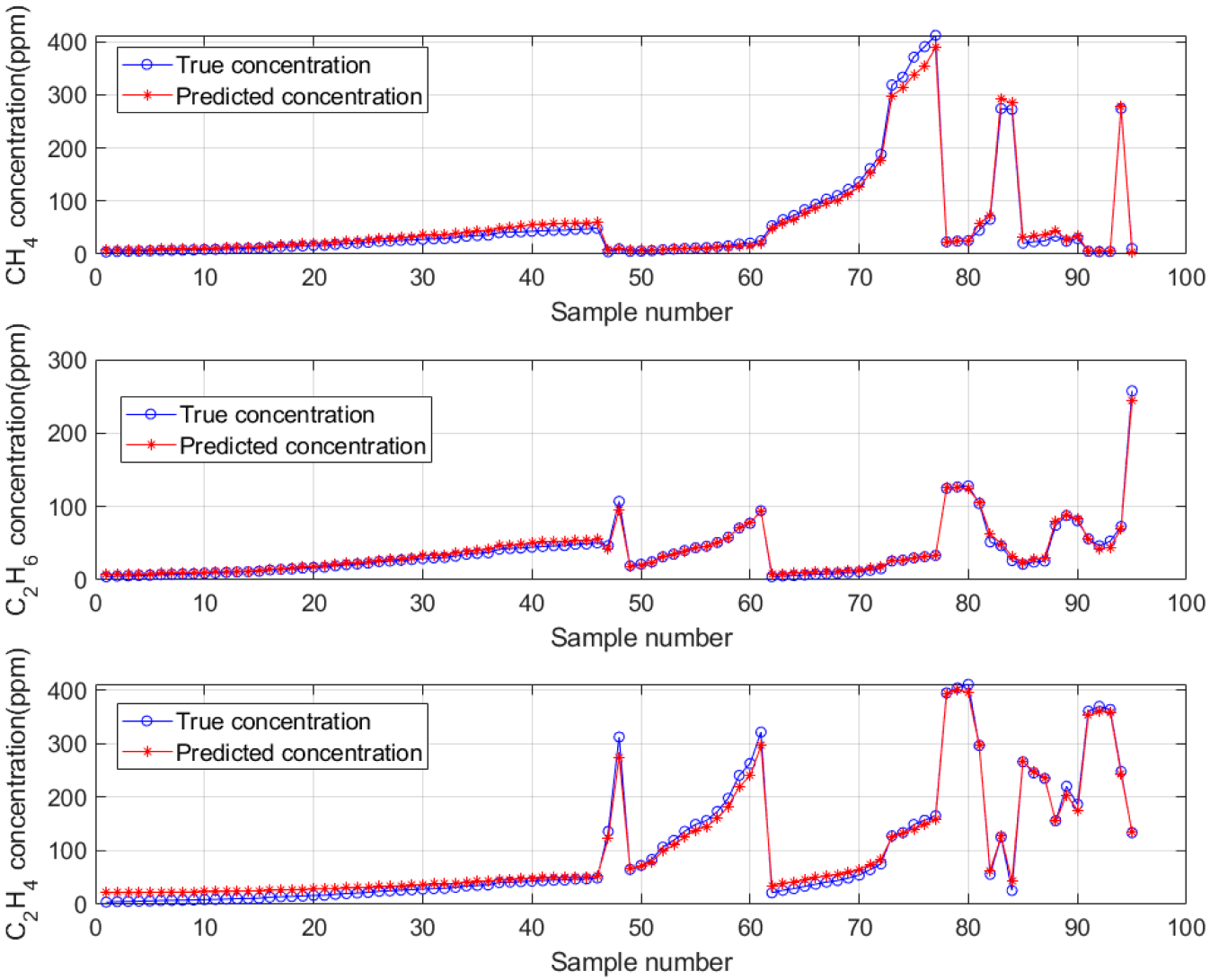


Figure 7. Error distribution of the base DKELM model under concentration fluctuations.

4.3. Performance Analysis of Optimization Algorithm

Performance of Standard Metaheuristic Frameworks

The instability and significant relative errors observed in the standalone DKELM model necessitate global optimization of its hyperparameters (C and γ). Therefore, we evaluated the performance of two standard metaheuristic frameworks: CPSO and CGWO.

These standard algorithms exhibited fundamental limitations in search efficiency when applied to the high-dimensional spectral dataset. The CGWO-DKELM model demonstrated rapid initial convergence (Figure S1), with the fitness error approaching its minimum within the first 10 iterations. However, it plateaued prematurely, indicating that while the gray wolf hierarchy is efficient at identifying promising regions, it lacks the stochastic diversity required to escape local optima once the leading "wolves" converge. Conversely, the CPSO-DKELM model (Figure S2) showed a slower, more erratic descent. The lack of a directional perturbation mechanism meant the particles were prone to "oscillating"

around sub-optimal coordinates, resulting in inferior initial stability and a protracted search process that ultimately failed to identify the global optimum. These results suggest that standard chaotic initialization alone is insufficient to prevent stagnation in complex parameter landscapes.

To overcome the stagnation observed in the Performance of Standard Metaheuristic Frameworks Section, mutation mechanisms were introduced to inject diversity and provide "global escape" capabilities. Therefore, we compared the enhanced variants, MCGWO and MCPSO, against their baseline counterparts.

The introduction of mutation significantly improved the optimization trajectory. In MCGWO (Figure S3), the application of Lévy flight and polynomial mutations to the alpha and omega wolves provided a better exploration–exploitation balance and successfully lowered the final fitness error from approximately 3.5 to 2.7. However, the most profound improvement was observed in the MCPSO algorithm (Figure S4). By integrating a directional chaotic perturbation strategy and a temperature-controlled mutation amplitude, MCPSO maintained population diversity even as the global best converged.

Unlike the other models, MCPSO exhibited a sustained, stepwise decrease in fitness throughout the entire iterative cycle. This behavior confirms that the adaptive mutation probability effectively counteracts the "velocity death" characteristic of standard CPSO. By periodically injecting chaotic perturbations based on the fitness distribution, MCPSO successfully escaped the local minima that trapped the CGWO and CPSO models. Consequently, it identified a more robust parameter set for the DKELM architecture to enable superior detection accuracy.

4.4. Quantitative Analysis of Mixed Gases

4.4.1. Results of Standard Chaotic Optimization (CPSO-DKELM and CGWO-DKELM)

The standard chaotic optimization models provided an initial improvement over the baseline DKELM but failed to achieve the necessary precision for reliable gas quantification. For the CGWO-DKELM model, the training set results (Figure 8) showed a clustering of data points that, although tighter than the base model, still exhibited significant dispersion in the 50–100 ppm range for CH_4. On the test set (Figure 9), the CGWO-DKELM model demonstrated a notable decrease in error compared to the standalone version; however, the relative error still exceeded 30% for several samples in the low-concentration ethylene range (Tables S4 and S5).

Similarly, the CPSO-DKELM model displayed characteristic instability. While the training set fitting (Figure 10) was reasonably consistent for methane, the model struggled with the overlapping absorption features of ethane and ethylene. On the test set, the RMSE values were 3.98 (CH_4), 2.45 (C_2H_6), and 6.12 (C_2H_4). Despite the chaotic initialization, the "velocity death" of particles in the late search stages led to a set of parameters that were insufficiently robust to handle the cross-interference of three-component gas mixtures, and the maximum relative error reached approximately 45% (Tables S6 and S7).

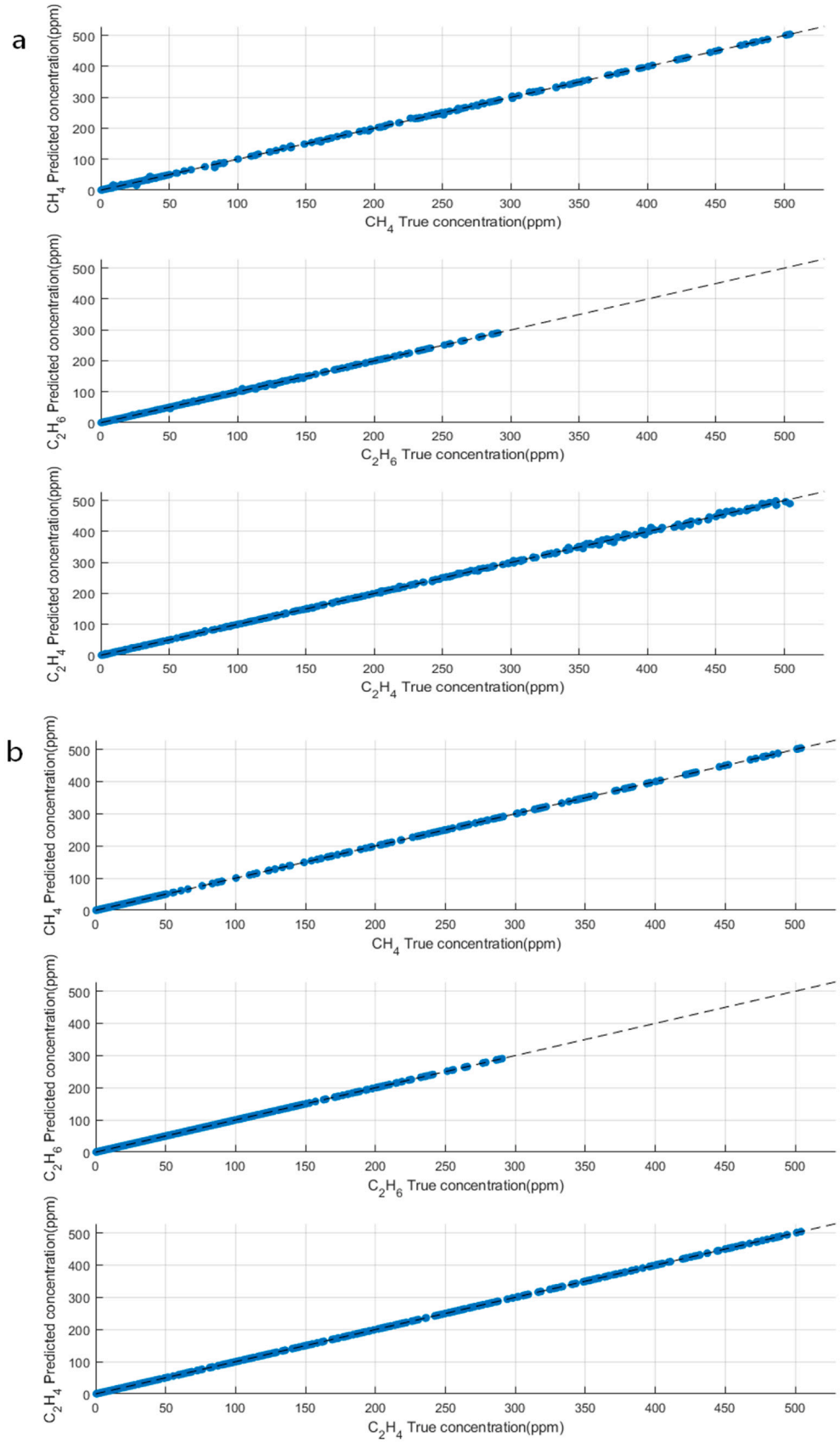


Figure 8. Fitting curves on the training set: (**a**) CGWO-DKELM, (**b**) CPSO-DKELM.

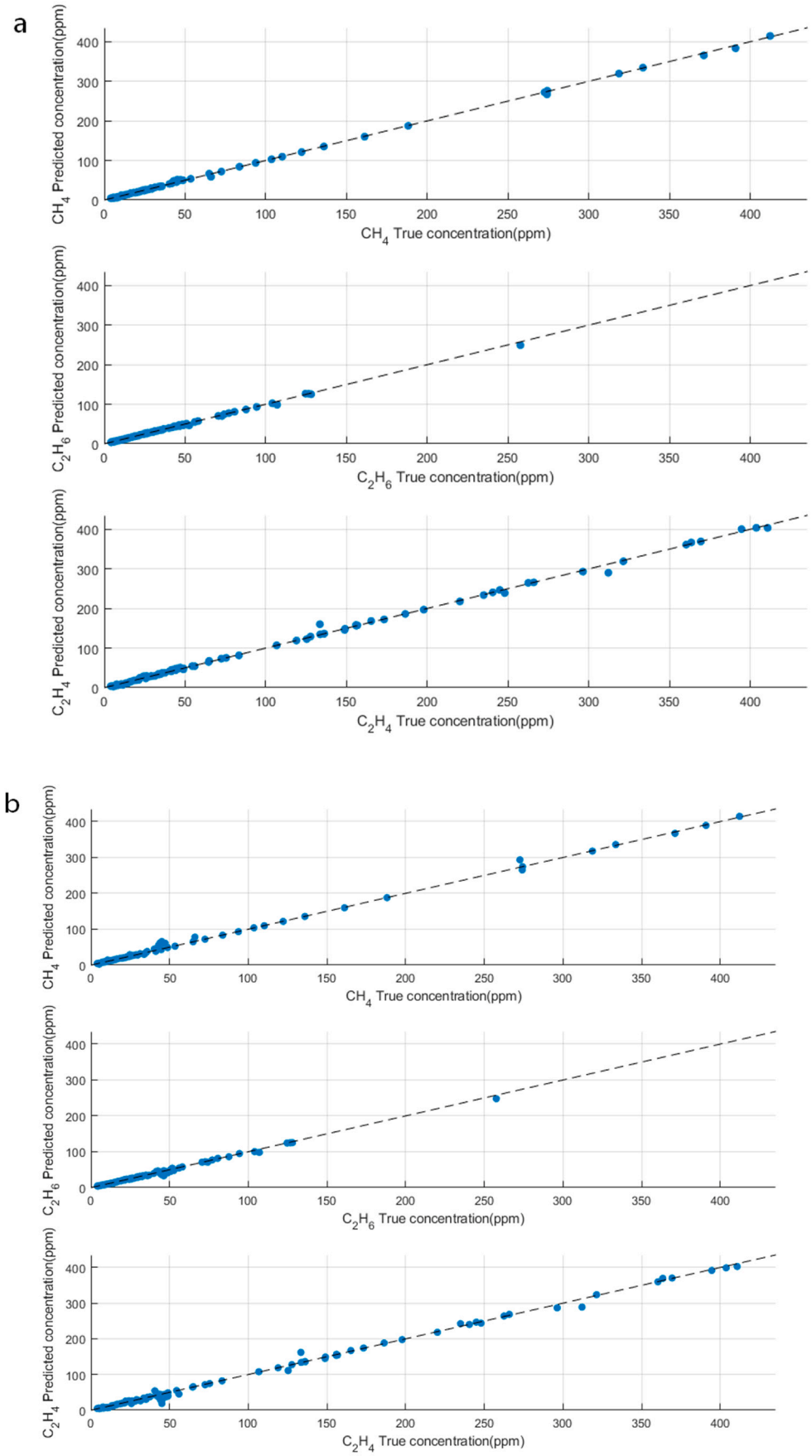


Figure 9. Fitting curves on the test set: (**a**) CGWO-DKELM, (**b**) CPSO-DKELM.

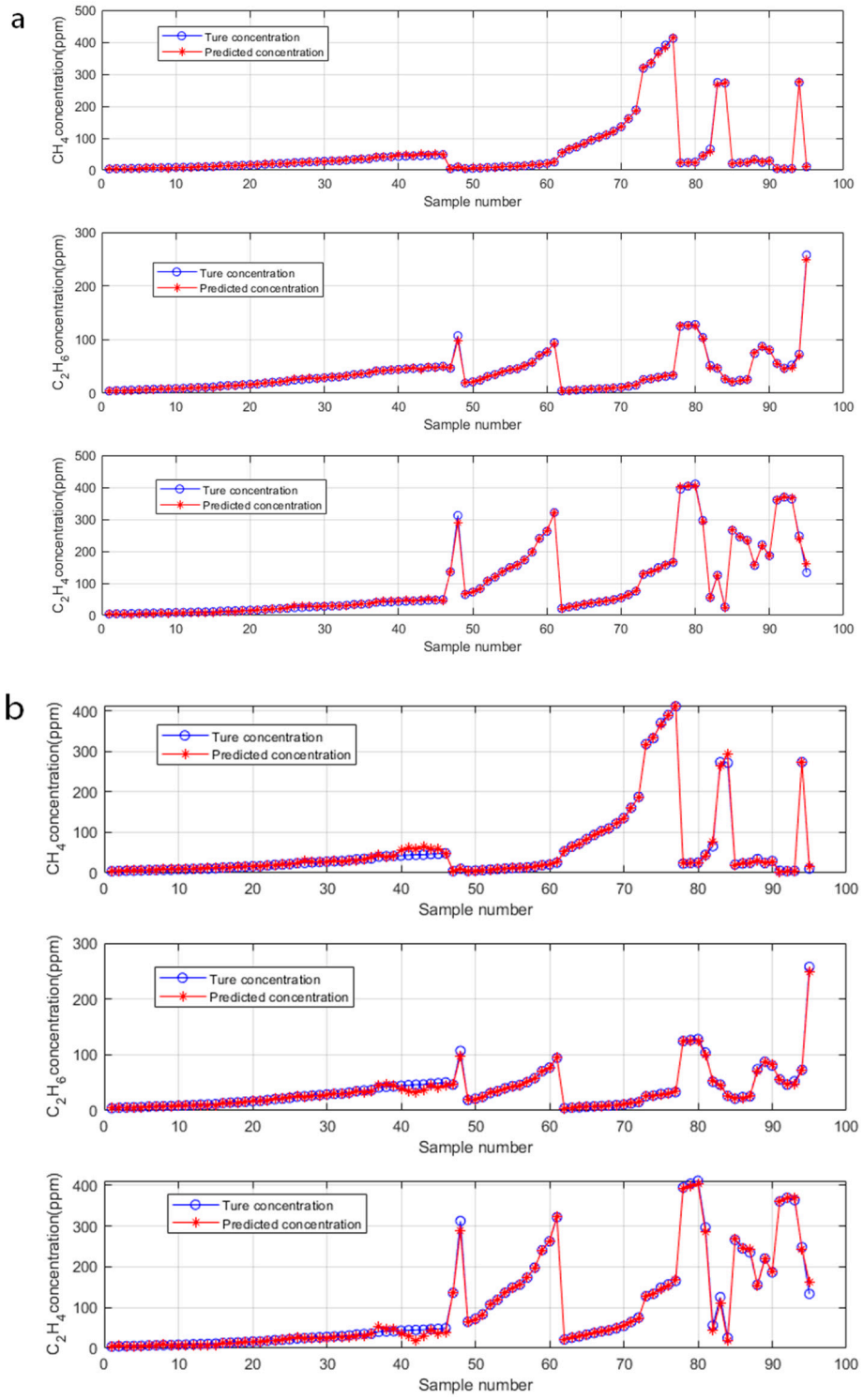


Figure 10. Error distribution of (**a**) CGWO-DKELM and (**b**) CPSO-DKELM under concentration fluctuations.

4.4.2. Results of Mutation-Enhanced Optimization (MCPSO-DKELM and MCGWO-DKELM)

The introduction of mutation strategies led to a marked enhancement in both training stability and test set generalization. The MCGWO-DKELM model showed a significant reduction in fitting dispersion on the training set (Figure 11). By utilizing Lévy flights to perturb the lead wolves, it captured non-linear spectral trends more effectively, achieving an R^2 of 0.995 for the test set. However, a slight tendency toward underfitting was still

observed in samples where ethylene concentrations were below 10 ppm (Tables S8 and S9). The MCPSO-DKELM model yielded the most superior performance across all metrics. On the training set (Figure 11), the fitting curves for CH_4, C_2H_6, and C_2H_4 showed almost perfect alignment with true values, with no discernible dispersion even at concentration extremes. This indicates that the directional chaotic perturbation effectively tuned the DKELM parameters to the global optimum. Evaluation on the independent test set confirmed this exceptional accuracy (Tables S10 and S11); regression plots were tightly clustered around the 1:1 line with high linearity across the full dynamic range (Figure 12).

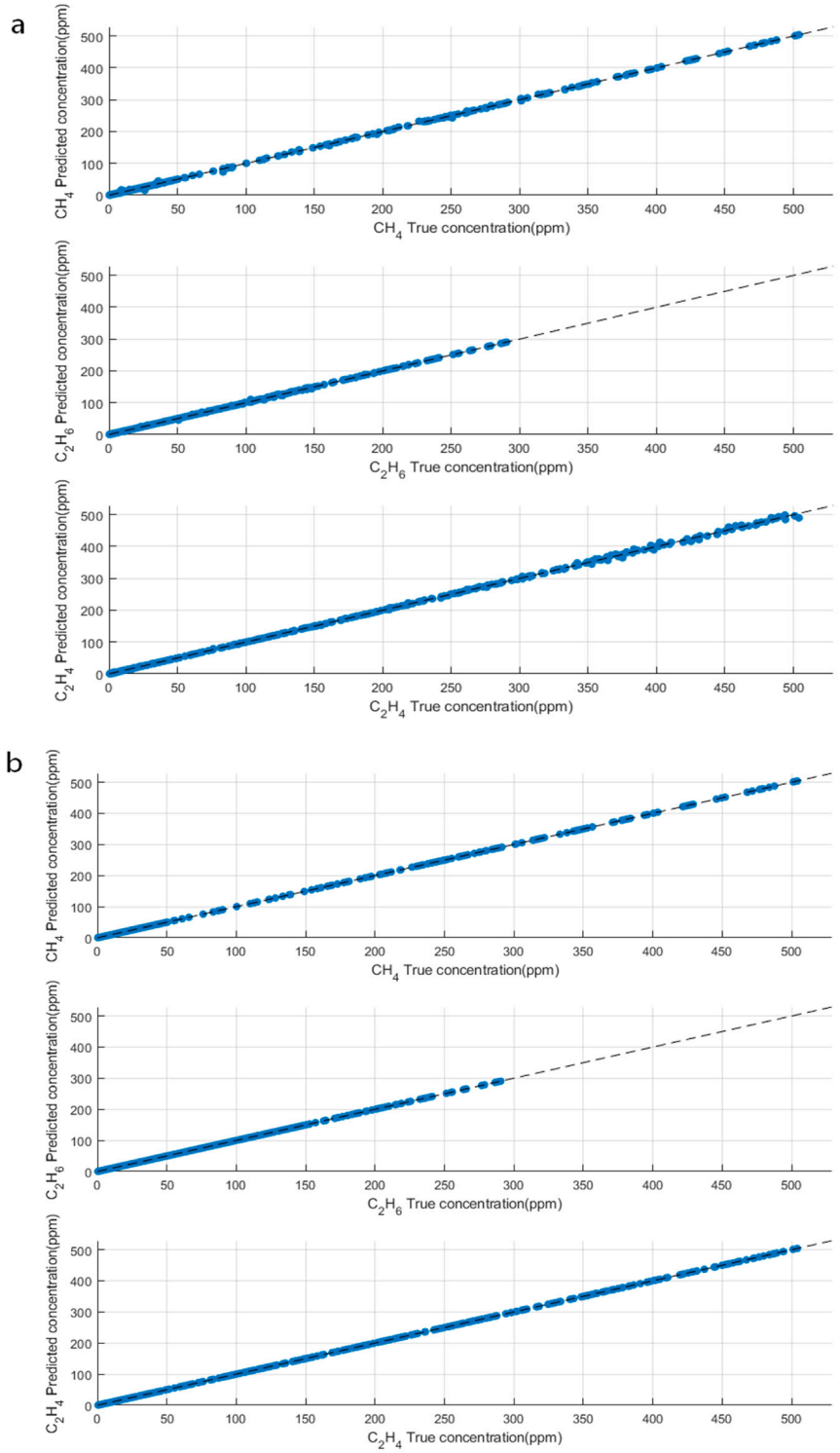


Figure 11. Fitting curves on the training set: (**a**) MCGWO-DKELM, (**b**) MCPSO-DKELM.

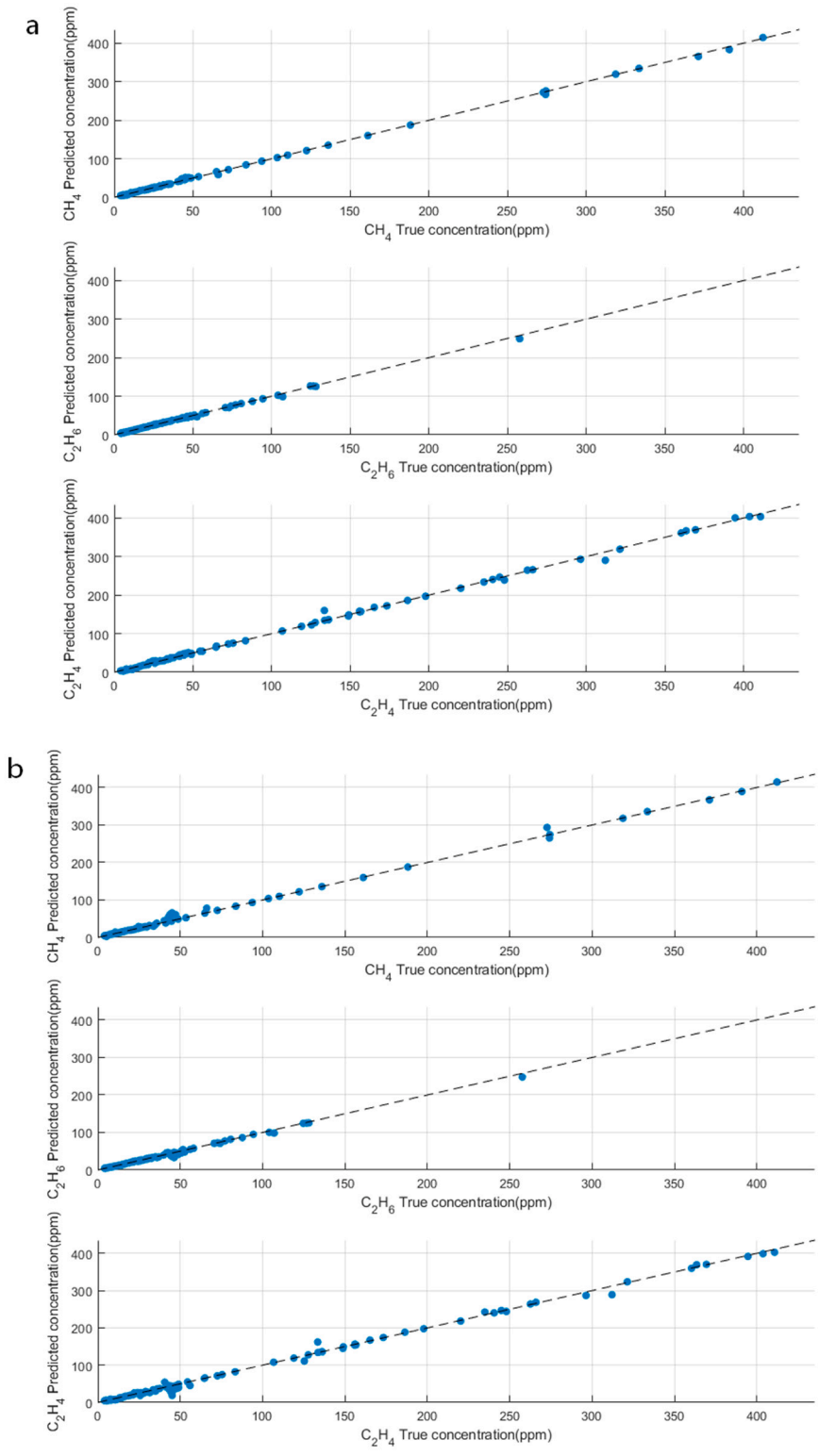


Figure 12. Fitting curves on the test set: (**a**) MCGWO-DKELM, (**b**) MCPSO-DKELM.

The stability and noise resilience of the model across the concentration continuum are further highlighted in the error distribution analysis (Figure 13). The MCPSO-DKELM model successfully tracks concentration fluctuations without the erratic deviations seen in simpler architectures. Table 1 also shows that the MCPSO-DKELM model achieved the lowest errors: the RMSE values were 1.732 (CH_4), 1.545 (C_2H_6), and 4.011 (C_2H_4), the maximum relative error (e_{max}) was constrained to below 15%, and the coefficient of determination (R^2) exceeded 0.998 for all three gases. That is, the MCPSO-DKELM can

capture >99.8% of the variance in the complex spectral data while maintaining excellent resistance to environmental noise and cross-interference.

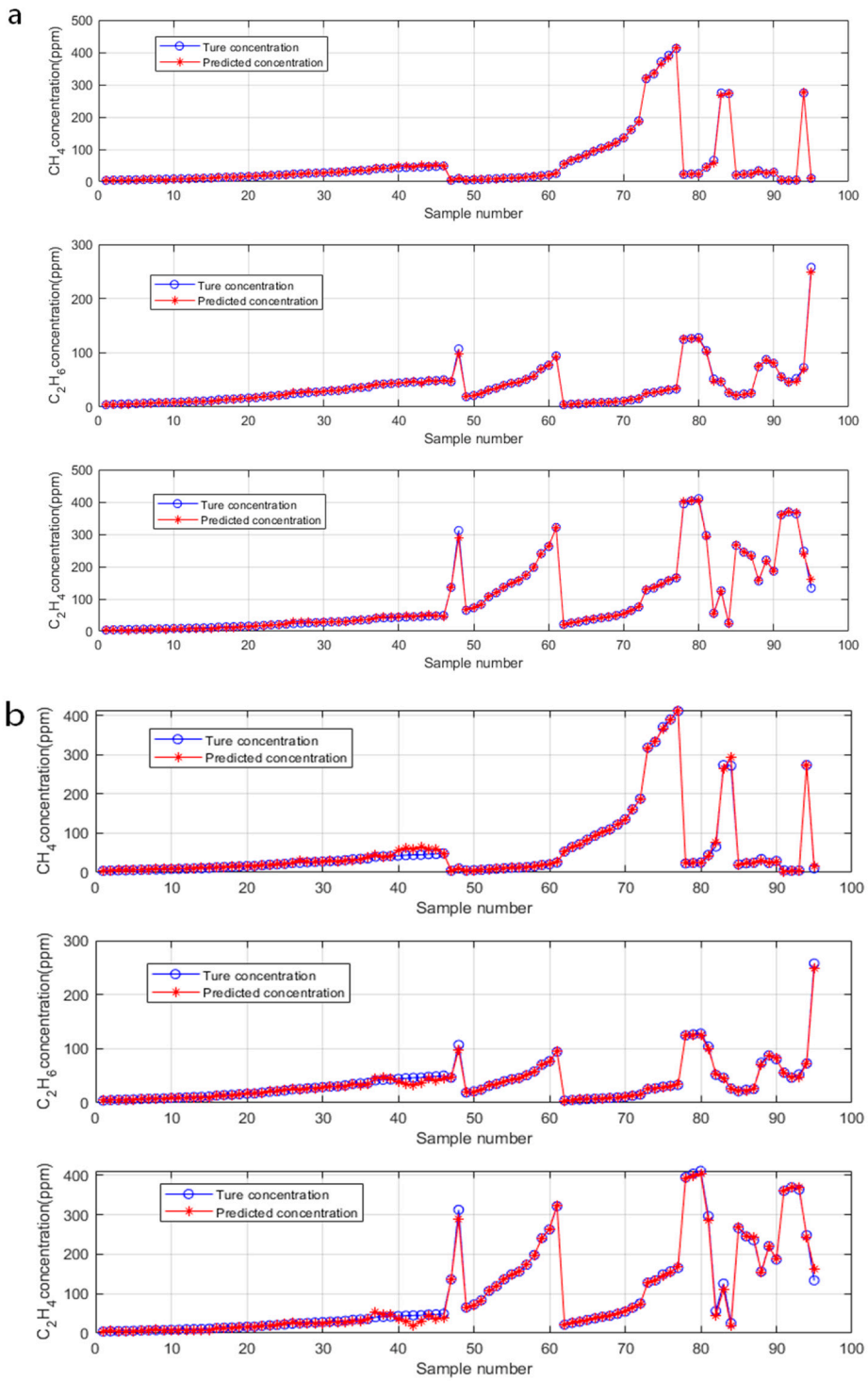


Figure 13. Error distribution of (**a**) MCGWO-DKELM and (**b**) MCPSO-DKELM under concentration fluctuations.

Table 1. Model performance metrics [§].

Model	CH_4				C_2H_6				C_2H_4			
	e_{max}	e_{avg}	RMSE	R^2	e_{max}	e_{avg}	RMSE	R^2	e_{max}	e_{avg}	RMSE	R^2
DKELM	1109.52%/ 121.13%	47.94%/ 17.57%	7.758/ 4.9635	0.9973/ 0.9970	547.47%/ 26.04%	16.37%/ 5.41%	3.706/ 2.277	0.9966/ 0.9964	1595.31%/ 202.70%	38.29%/ 47.94%	9.9145/ 7.361	0.9953/ 0.9959
CGWO-DKELM	284.13%/ 19.81%	6.16%/ 3.09%	1.801/ 1.727	0.9999/ 0.9996	66.60%/ 17.67%	2.14%/ 2.11%	0.822/ 1.479	0.9998/ 0.9985	148.67%/ 28.95%	3.03%/ 3.71%	2.024/ 4.113	0.9998/ 0.9987
CPSO-DKELM	65.10%/ 22.66%	2.10%/ 4.34%	0.772/ 2.162	1/ 2.1619	18.14%/ 10.91%	0.77%/ 2.83%	0.326/ 1.902	1/ 0.9975	62.40%/ 31.35%	1.03%/ 5.28%	0.877/ 4.899	1/ 0.9982
MCGWO-DKELM	87.12%/ 21.25%	5.09%/ 3.34%	1.679/ 1.555	0.9999/ 0.9997	66.72%/ 14.63%	1.78%/ 2.13%	0.815/ 1.508	0.9997/ 0.9983	82.18%/ 21.99%	2.19%/ 4.26%	1.983/ 4.417	0.9997/ 0.9985
MCPSO-DKELM	77.14%/ 13.85%	4.97%/ 3.24%	1.901/ 1.732	0.9998/ 0.9996	35.67%/ 10.67%	1.97%/ 2.21%	1.019/ 1.545	0.9997/ 0.9983	72.99%/ 14.72%	2.05%/ 3.13%	2.480/ 4.011	0.9997/ 0.9986
SVM-CPSO-KELM	196.77%/ 43.80	6.77%/ 6.13%	5.765/ 5.889	0.9985/ 0.9958	174.36%/ 41.20	5.16%/ 5.38%	1.369/ 2.033	0.9979/ 0.9894	189.29%/ 38.28%	6.86%/ 4.63%	5.255/ 4.165	0.9977/ 0.9959

[§] In each cell, the first row denotes the metric for the training set (376 samples) and the second row denotes the metric for the test set (94 samples).

4.5. Benchmark Testing

To validate the advantages of the deep architecture over segmented shallow learning, the MCPSO-DKELM model was compared against our previously established SVM-CPSO-KELM model. This benchmark comparison utilized the full dataset of 470 samples to ensure statistical parity. The results (Table 2) demonstrated that the "deep architecture + mutation–chaotic optimization" strategy of MCPSO-DKELM yields superior accuracy compared to the "classification + regression" strategy of SVM-CPSO-KELM. From SVM-CPSO-KELM to MCPSO-DKELM, the RMSE on the test set was reduced for all gases: from 5.889 to 1.732 (~71%) for CH_4, from 2.033 to 1.545 (~24%) for C_2H_6, and from 4.165 to 4.011 (~4%) for C_2H_4. In contrast, for the SVM-CPSO-KELM model itself, increasing the sample size from 130 to 470 reduced the RMSE by ~35% and ~2% for CH_4 and C_2H_6 respectively, and increased the RMSE by ~1% for C_2H_4. Thus, the MCPSO-DKELM model fully leveraged the larger dataset to minimize errors, and the reported accuracy improvements stem from the robust feature extraction capabilities of the deep kernel architecture rather than mere data augmentation.

Table 2. Model benchmarking [§].

Model	CH_4			C_2H_6			C_2H_4		
	Sample Size	RMSE	R^2	Sample Size	RMSE	R^2	Sample Size	RMSE	R^2
SVM-CPSO-KELM [†]	104/ 26	6.377/ 9.127	0.9984/ 0.9940	104/ 26	1.837/ 2.075	0.9977/ 0.9899	104/ 26	6.146/ 4.122	0.9977/ 0.9943
SVM-CPSO-KELM	376/ 94	5.765/ 5.889	0.9985/ 0.9958	376/ 94	1.369/ 2.033	0.9979/ 0.9894	376/ 94	5.255/ 4.165	0.9977/ 0.9959
MCPSO-DKELM	376/ 94	1.901/ 1.732	0.9998/ 0.9996	376/ 94	1.019/ 1.545	0.9997/ 0.9983	376/ 94	2.480/ 4.011	0.9997/ 0.9986

[§] In each cell, the first row denotes the metric for the training set (376 samples) and the second row denotes the metric for the test set (94 samples). [†] Dataset reported in Ref. [22].

5. Discussion

5.1. Rationale for Mutation–Chaotic Enhancements

The empirical results in Section 4 demonstrate that standard metaheuristics often fail to navigate the complex hyperparameter landscape of a multi-layer DKELM. The stagnation observed in standard CPSO and CGWO can be traced to "search blind spots" and population homogeneity. In standard PSO, particles primarily follow their own historical best and the current global best; as the search progresses, individuals become increasingly similar, leading to "velocity death" where the population crowds into a local optimum.

The MCPSO algorithm resolves these issues through a multi-strategy intervention. The hybrid initialization (alternating Logistic and Tent maps) ensures that the initial population is uniformly distributed rather than clustered in a single corner of the solution space.

Furthermore, the introduction of a "rebound mechanism" for boundary handling and a 10% replacement strategy for low-performing individuals ensures that the search remains active even in later iterations. Crucially, the directional chaotic perturbation using Chebyshev mapping specifically targets "poor" particles (those with fitness 1.5 times worse than the global best), guiding them toward promising regions without sacrificing the stochastic nature of the search. This prevents the "search stagnation" that characterized the baseline results in Section 4.2.

5.2. Architectural Evolution: From SVM-KELM to DKELM

A primary finding of this study is that the MCPSO-DKELM model achieves superior accuracy without requiring the Support Vector Machine (SVM) classification stage used in our previous work (e.g., SVM-CPSO-KELM). The necessity of SVM in the previous work stemmed from the "single-layer" limitation of KELM. Because KELM maps features at a single scale, it cannot simultaneously adapt to the drastically different slopes of the working curves for low-concentration "weak signals" and high-concentration "strong signals." A single-layer regressor typically produces an "averaged" fit, leading to significant relative errors at the concentration extremes. Therefore, SVM was required to "divide and conquer" the data into segments.

In contrast, the DKELM architecture employed here uses multi-layer kernel mapping to achieve hierarchical learning. In this framework,

- Shallow layers extract basic spectral features and absorption peak positions;
- Middle layers disentangle complex cross-interference patterns between the gases;
- Deep layers capture the non-linear concentration–signal response relationships across the entire continuum.

This inherent hierarchy automates the feature adaptation for different concentration intervals, effectively replacing the explicit segmentation of SVM. After DKELM is coupled with the global search stability of MCPSO, the final model maintains high precision across the full range (0–500 ppm), which not only simplifies the processing workflow but also breaks the bottleneck of shallow architectures in detection accuracy.

6. Conclusions

This study successfully developed a high-precision method for detecting the concentrations of CH_4, C_2H_6, and C_2H_4 in gas mixtures using photoacoustic spectroscopy. By integrating a deep kernel architecture with a mutation-enhanced chaotic optimization framework, we addressed the critical challenge of spectral cross-interference and non-linear response variability, and demonstrated the following:

1. **Algorithmic Superiority:** The MCPSO algorithm, featuring adaptive diversity monitoring and directional chaotic perturbations, effectively prevents premature convergence. It outperformed standard metaheuristics in tuning the DKELM hyperparameters and enabled a sustained optimization trajectory.
2. **Detection Accuracy:** The MCPSO-DKELM model reduced the maximum relative error from over 200% (DKELM alone, baseline) to below 15% while keeping the average error under 3.5%. The coefficient of determination (R^2) exceeded 0.998 for all three target gases, which confirms excellent model stability and generalization.
3. **Architectural Efficiency:** The deep hierarchical structure of DKELM eliminates the need for a separate SVM classification stage. This "deep architecture + mutation–chaotic optimization" strategy represents a significant advancement over segmented shallow learning models, particularly for large-scale datasets where feature complexity is high.

Thus, this work provides a robust and reliable solution for the simultaneous detection of multi-component gas mixtures, with significant potential for application in industrial safety, dissolved-gas analysis in transformers, and environmental monitoring. Future research will focus on extending this deep learning framework to more complex gas matrices.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/spectroscj4030012/s1.

Author Contributions: P.G.: software, validation, writing—original draft; M.Z.: data curation, formal analysis, methodology; X.T.: formal analysis; Y.H.: writing—review and editing, supervision. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Materials. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Andrawes, F.F.; Gibson, E.K. Simultaneous determination of trace amounts of hydrogen, oxygen, nitrogen, carbon monoxide, carbon dioxide, methane, ethane, ethylene, and acetylene by two gas chromatographic columns in parallel and one detector. *Anal. Chem.* **1979**, *51*, 462–463. [CrossRef]
2. Dudek, R.B.; Li, F.X. Selective hydrogen combustion as an effective approach for intensified chemical production via the chemical looping strategy. *Fuel Process. Technol.* **2021**, *218*, 106827. [CrossRef]
3. Luo, P.; Harrist, J.; Menduni, G.; Mesdour, R.; StMichel, N.; Sampaolo, A. Simultaneous detection of methane, ethane, and propane by QEPAS sensors for on-site hydrocarbon characterization and production monitoring. *ACS Omega* **2022**, *7*, 3395–3406. [CrossRef] [PubMed]
4. Odjo, A.O.; Garcia, A.N.; Marcilla, A. Refinery nonconventional feedstocks: Influence of the coprocessing of vacuum gas oil and low density polyethylene in fluid catalytic cracking unit on full range gasoline composition. *Energy Fuels* **2014**, *28*, 1579–1593. [CrossRef]
5. Intergovernmental Panel on Climate Change. *Climate Change 2007: Synthesis Report*; World Meteorological Organization: Geneva, Switzerland, 2007; Volume 52, p. 1.
6. Postma, R.S.; Lefferts, L. Effect of ethane and ethylene on catalytic non oxidative coupling of methane. *React. Chem. Eng.* **2021**, *6*, 2425–2433. [CrossRef] [PubMed]
7. Chen, Y.Q.; Kuo, M.J.; Lobo, R.; Ierapetritou, M. Ethylene production: Process design, techno-economic and life-cycle assessments. *Green Chem.* **2024**, *26*, 2903–2911. [CrossRef]
8. Lei, Y.; Pan, W.Y.; Hu, S.B.; Kuang, Y.C.; Liu, X.Y.; Chen, Y.Q. Ionic liquid-assisted extractive distillation for ethylene/ethane separation: Molecular design and process assessment. *Ind. Eng. Chem. Res.* **2024**, *63*, 19187–19199. [CrossRef]
9. Zhou, J.; Er, Z.; Gong, P.; Xie, L.; Jiang, X.K. A compact gas monitoring system for methane based on photoacoustic spectroscopy. *Infrared Phys. Technol.* **2022**, *123*, 104186. [CrossRef]
10. Wijesinghe, D.R.; Zobair, M.A.; Esmaeelpour, M. A review on photoacoustic spectroscopy techniques for gas sensing. *Sensors* **2024**, *24*, 6577. [CrossRef] [PubMed]
11. Wu, L.; Chen, T.; Xiang, B.; Xing, L.K. Design of NO_2 photoacoustic detection system based on finite element simulation. *Appl. Sci.* **2024**, *14*, 11343. [CrossRef]
12. Zhang, C.; Qiao, S.D.; He, Y.; Liu, C.M.; Ma, Y.F. Multi-resonator T-type photoacoustic cell based photoacoustic spectroscopy gas sensor for simultaneous measurement C_2H_2, CH_4 and CO_2. *Sens. Actuators B* **2025**, *427*, 137168. [CrossRef]
13. Zhang, J.L.; Liu, L.X.; Sun, J.L.; Zhang, X.S.; Chen, B.S.; Liang, Y.Z.; Zhao, B.X.; Huan, H.T.; Xu, X.S.; Xu, H.L.; et al. Open type solid photoacoustic trace gas sensor with multi-pass absorption enhancement. *Anal. Chem.* **2025**, *97*, 16663–16671. [CrossRef] [PubMed]
14. Wu, Z.; Shi, Y.X.; Han, Y.W. Methane, ethylene, and ethane detection by differential helmholtz resonance spectroscopy using a 3345 nm mid-infrared tunable diode laser source. *Appl. Sci.* **2023**, *13*, 3169. [CrossRef]

15. Cantatore, A.F.P.; Menduni, G.; Zifarelli, A.; Patimisco, P.; Giglio, M.; Gonzalez, M.; Seren, H.R.; Luo, P.; Spagnolo, V.; Sampaolo, A. Methane, ethane, and propane detection using a quartz-enhanced photoacoustic sensor for natural gas composition analysis. *Energy Fuels* **2025**, *39*, 638–646. [CrossRef] [PubMed]
16. Du, Z.H.; Zhang, S.; Li, J.Y.; Gao, N.; Tong, K.B. Mid-infrared tunable laser-based broadband fingerprint absorption spectroscopy for trace gas sensing: A review. *Appl. Sci.* **2019**, *9*, 338. [CrossRef]
17. Li, Z.; Zhang, Q.R.; Wang, Z.T.; Dai, J.M. A highly sensitive low-pressure TDLAS sensor for detecting dissolved CO and CO_2 in transformer insulating oil. *Opt. Laser Technol.* **2024**, *174*, 110622. [CrossRef]
18. Plehiers, P.P.; Symoens, S.H.; Amghizar, I.; Marin, G.B.; Stevens, C.V.; Van Geem, K.M. Artificial intelligence in steam cracking modeling: A deep learning algorithm for detailed effluent prediction. *Engineering* **2019**, *5*, 1027–1040. [CrossRef]
19. Wang, Y.S.; Kong, Q.M.; Gao, J.Q.; Chen, S.M. Research on on-line detection of CO and CO_2 mixed gas concentration based on improved extreme learning machine and TDLAS. *Infrared Phys. Technol.* **2025**, *148*, 105868. [CrossRef]
20. Zhao, M.X.; Gu, P.C.; Han, Y.W. A hybrid SVM-CPSO-KELM model for the simultaneous detection of methane, ethane, and ethylene via photoacoustic spectroscopy. *Chemom. Intell. Lab. Syst.* **2026**, *269*, 105620. [CrossRef]
21. Li, J.J.; Xi, B.B.; Du, Q.; Song, R.; Li, Y.S.; Ren, G.B. Deep kernel extreme-learning machine for the spectral–spatial classification of hyperspectral imagery. *Remote Sens.* **2018**, *10*, 2036. [CrossRef]
22. Afzal, A.L.; Nair, N.K.; Asharaf, S. Deep kernel learning in extreme learning machines. *Pattern Anal. Appl.* **2020**, *24*, 11–19. [CrossRef]
23. Wu, D.; Li, T.; Wan, Q. A hybrid deep kernel incremental extreme learning machine based on improved coyote and beetle swarm optimization methods. *Complex Intell. Syst.* **2021**, *7*, 3015–3032. [CrossRef]
24. Ding, S.F.; Guo, L.L.; Hou, Y.L. Extreme learning machine with kernel model based on deep learning. *Neural Comput. Appl.* **2016**, *28*, 1975–1984. [CrossRef]
25. Huang, G.B.; Zhou, H.; Ding, X.; Zhang, R. Extreme learning machine for regression and multiclass classification. *IEEE Trans. Syst. Man Cybern. B Cybern.* **2012**, *42*, 513–529. [CrossRef] [PubMed]
26. Li, J.F.; Hai, C.; Feng, Z.; Li, G.X. A transformer fault diagnosis method based on parameters optimization of hybrid kernel extreme learning machine. *IEEE Access* **2021**, *9*, 126891–126902. [CrossRef]
27. Wong, C.M.; Vong, C.M.; Wong, P.K.; Cao, J. Kernel-based multilayer extreme learning machines for representation learning. *IEEE Trans. Neural Netw. Learn. Syst.* **2018**, *29*, 757–762. [CrossRef] [PubMed]
28. Mohanty, F.; Rup, S.; Dash, B. Automated diagnosis of breast cancer using parameter optimized kernel extreme learning machine. *Biomed. Signal Process. Control* **2020**, *62*, 102108. [CrossRef]
29. Lai, Z.; Huang, Z.; Xu, M.; Wang, C.; Xu, J.; Zhang, C.; Zhu, R.; Qiao, Z. High-performance adaptive weak fault diagnosis based on the global parameter optimization model of a cascaded stochastic resonance system. *Sensors* **2023**, *23*, 4429. [CrossRef] [PubMed]
30. Xi, B.B.; Li, J.J.; Li, Y.S.; Song, R.; Sun, W.W.; Du, Q. Multiscale context-aware ensemble deep KELM for efficient hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **2021**, *59*, 5114–5130. [CrossRef]
31. Abdullah, H.N. An improvement in LQR controller design based on modified chaotic particle swarm optimization and model order reduction. *Int. J. Intell. Eng. Syst.* **2021**, *14*, 157–168. [CrossRef]
32. Ardiansyah, A.; Ferdiana, R.; Permanasari, A.E. MUCPSO: A modified chaotic particle swarm optimization with uniform initialization for optimizing software effort estimation. *Appl. Sci.* **2022**, *12*, 1081. [CrossRef]
33. Safaei, N.M.; Tavakkoli-Moghaddam, R.; Kiassat, C. Annealing-based particle swarm optimization to solve the redundant reliability problem with multiple component choices. *Appl. Soft Comput.* **2012**, *12*, 3462–3471. [CrossRef]
34. Tsai, H.-C.; Shi, J.-Y. Potential corrections to grey wolf optimizer. *Appl. Soft Comput.* **2024**, *161*, 111776. [CrossRef]
35. Zarchi, D.A.; Vahidi, B. Multi objective self adaptive optimization method to maximize ampacity and minimize cost of underground cables. *J. Comput. Des. Eng.* **2018**, *5*, 401–408. [CrossRef]
36. Chicco, D.; Warrens, M.J.; Jurman, G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [CrossRef] [PubMed]
37. Alexander, D.L.J.; Tropsha, A.; Winkler, D.A. Beware of R^2: Simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models. *J. Chem. Inf. Model.* **2015**, *55*, 1316–1322. [CrossRef] [PubMed]
38. Floryan, D.; Graham, M.D. Discovering multiscale and self-similar structure with data-driven wavelets. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2021299118. [CrossRef] [PubMed]
39. Ghobber, S.; Mejjaoli, H. Two versions of dunkl linear canonical wavelet transforms and applications. *Mathematics* **2025**, *13*, 3225. [CrossRef]

40. Han, N.N.; Pei, Y.Z.; Song, Z.J. Signal separation operator based on wavelet transform for non-stationary signal decomposition. *Sensors* **2024**, *24*, 6026. [CrossRef] [PubMed]
41. Yulong, D.; Ke, D.; Chunsheng, O.Y.; Yingshe, L.; Yu, T.; Jianyi, F.; Wei, W.; Yaguang, D. Wavelets and curvelets transform for image denoising to damage identification of thin plate. *Results Eng.* **2023**, *17*, 100837. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.