

TRƯỜNG ĐẠI HỌC QUY NHƠN
QUY NHON UNIVERSITY

TẠP CHÍ KHOA HỌC
TRƯỜNG ĐẠI HỌC QUY NHƠN

QUY NHON UNIVERSITY
JOURNAL OF SCIENCE

KHOA HỌC TỰ NHIÊN VÀ KỸ THUẬT
NATURAL SCIENCES AND ENGINEERING

20 (3)

2026

JUNE 2026

CONTENTS

1.	Performance evaluation of a KNN algorithm implementation using Rust-Webassembly Tran Hoang Viet, Doan Thi Thu Cuc, Vo Thi My, Quang Thi Ngoc Bich	5
2.	Optimization of congo red degradation performance on diatomaceous earth catalyst using Box-Behnken design Pham Dinh Du.....	19
3.	Gridding-based $\mathcal{H}_\infty$ observer design for the Lorenz 63 system using an NLPV reformulation Khanh-Thoai Vo, Thanh-Phong Pham, Tien Dung Le	29
4.	Practical embedded AI on FPGA: A compact CNN achieving high throughput and low power for digit recognition Pham Van Khoa, La Cong Loc, Nguyen Duc Trong, Huynh Minh Quy, Ho Gia Huyen, Vo Quang Huy, Pham Ba Loc, Huynh Thi Thu Hien.....	43
5.	Study on analysis and assessment of exposure levels and accumulation profiles of Polycyclic Aromatic Hydrocarbons (PAHs) in human hair and fingernail samples Nguyen Duc Hieu, Hoang Quoc Anh, Chu Thi Hue, Hoang Thi Huong Hue, Nguyen Thi Duyen, Nguyen Xuan Thao, Pham Quynh Thiem, Do Van Doanh, Tran Duc Chung, Nguyen Thi Anh Huong	55
6.	Enhancing higher education faculty competencies in the age of generative AI: Integrating explainability, contestability, and reflective practice Tran Diem Quynh Nguyen, Thi Cam Mai Truong.....	65
7.	Thermal analysis of low-voltage power cables considering the presence of current harmonics Kim Anh Nguyen, Dai Huynh, Thi Thu Le, Huy Hoang Le, Ba Tu Phung.....	85
8.	Effect of maturation stages and storage conditions on physicochemical properties of three-leave cayratia (Cayratia trifolia) berries from different geographical regions in Vietnam Doan Thi Kieu Tien, Tran Diem Phuong, Huynh Thi Ngoc Mi, Dinh Hoang Lan Chi, Nguyen Ngoc Thanh, Huynh Xuan Phong.....	97
9.	Robust design optimization for overhead power transmission line design Pham Van Phat, Tran Ngoc Thanh, Tuan-Ho Le	109

10.	Research on beamforming optimization in intelligent reflective surface-assisted wireless communication systems	
	Do Dung Nguyen , Van Hao Nguyen, Minh Hung Dao	127
11.	Splitting piecewise cubic Bézier curves and the equivalence constants for some norms	
	Hoang Van Duc	139
12.	Design and development of a computer-vision-based robotic arm system for sorting of vehicle-like objects	
	Dinh Minh Le, Xuan Loc Nguyen, Van Can Do, Xuan Khoa Tran	151
13.	Design and metrological validation of a programmable high-voltage I-V measurement system for dielectric breakdown kinetics	
	Huynh Cong Tu, Nguyen Van Hao	177
14.	Recovery of monoclinic phase of VO ₂ nanoparticles from oxidized samples: Assessment of reusability and optimization of re-annealing conditions	
	Ho Nguyen Hong Tham, Nguyen Thi Thanh Nhan, Nguyen Thi Tuong Vy, Le Phuong Thao, Le Thi My Nhi, Hoang Thi Hang, Le Thi Ngoc Loan	189

Performance evaluation of a KNN algorithm implementation using Rust-Webassembly

Tran Hoang Viet^{1,*}, Doan Thi Thu Cuc¹, Vo Thi My¹, Quang Thi Ngoc Bich²

¹Department of Information Technology, Quy Nhon University, Vietnam

²Le Hong Phong Primary School, Quy Nhon, Vietnam

Received: 29/09/2025; Revised: 22/10/2025;

Accepted: 24/10/2025; Published: 28/06/2026

ABSTRACT

In the context of rapid development of edge computing, shifting processing from server-side to client-side is increasingly important to reduce latency, enhance data security, and support real-time applications such as IoT or artificial intelligence (AI) applications on devices like web browsers in particular and edge devices (e.g., IoT devices, smartphones, or hardware with limited resources...) in general. WebAssembly (WASM) is a technology that supports operations on these devices with near-native speed and portability during deployment. In this paper, we evaluate the performance of the K-Nearest Neighbors (KNN) algorithm implemented in the Rust language and compiled to WebAssembly, and compare it with various other KNN implementations in web environments. The objective is to assess the capabilities of WASM and the KNN algorithm, thereby opening up research directions for implementing other machine learning models that require substantial computational power, utilizing WASM on edge devices to optimize machine learning applications.

Keywords: *K-Nearest Neighbors algorithm, WebAssembly, edge computing, Rust, machine learning.*

1. INTRODUCTION

The rapid proliferation of edge devices, such as smartphones, IoT sensors, and web clients, has created a pressing need for data processing closer to the source, reducing reliance on cloud servers. Edge computing not only minimizes latency and conserves bandwidth but also enhances data security by limiting the transmission of sensitive information to the cloud. In the domain of machine learning, many applications are shifting toward direct execution in web browsers or resource-constrained devices to support tasks like real-time anomaly detection or personalized recommendations. However, classical algorithms like K-Nearest Neighbors (KNN), which demand substantial computational resources

for distance calculations in high-dimensional spaces, face significant efficiency issues on edge devices with limited hardware, particularly in web environments [1] - [3].

WebAssembly (WASM) offers a promising solution, providing a binary instruction format that enables compilation from low-level languages like Rust, C, and C++ to achieve near-native performance in browsers [4] - [5]. Rust, with its memory safety, concurrency support, and robust WASM compilation capabilities, is an ideal choice for deploying applications on edge devices. Recent studies demonstrate that WASM can outperform traditional JavaScript in web environments with reported speedups of 1.5–2x, as evidenced by applications such as Photoshop Web, AutoCAD Web, and Figma [6].

*Corresponding author:

Email: tranhoangviet@qnu.edu.vn

In the field of machine learning, WASM has been adopted for certain algorithms to enable client-side execution in web environments. However, there is a lack of comprehensive studies evaluating the performance of classification algorithms like KNN when implemented with WASM compared to other setups. This study evaluates the performance of the KNN algorithm, a simple yet computationally intensive classifier with a time complexity of $O(n \times d)$ to demonstrate WASM's capabilities in resource-constrained web environments compared to traditional JavaScript implementations, TensorFlow.js-based setups, and server-side models [7].

The primary motivation of this research is to shift processing from large-scale servers to edge devices, reducing cloud dependency and enhancing data privacy by keeping data on-device. Rust provides safety and high efficiency for KNN, while WASM ensures seamless deployment in web environments.

This paper is organized as follows: Section 2 reviews related work; Section 3 describes the experimental methodology; Section 4 presents the experimental results; Section 5 discusses the findings; and Section 6 concludes with future directions.

2. RELATED WORK

2.1. WebAssembly and Rust in edge computing

WebAssembly (WASM) is a critical technology for web and edge applications, particularly when combined with programming languages like C/C++ and Rust to enable machine learning (ML) on client-side devices. In recent years, numerous applications have been compiled to WASM, reflecting a shift from server-side to client-side processing [1], [5].

2.1.1 WASM and RUST

WebAssembly (WASM) is a binary instruction format designed to execute code at near-native speed in sandboxed environments, such as web browsers or edge devices. Developed by the World Wide Web Consortium (W3C) in 2017,

WASM aims to provide a low-level compiled language that operates across platforms without relying on JavaScript, addressing the performance limitations of traditional scripting languages. WASM supports compilation from various source languages, including C/C++, Go, and notably Rust, enabling developers to write high-performance code with strong portability [1], [5].

Rust, a systems programming language developed by Mozilla in 2010, emphasizes memory safety, performance, and concurrency without requiring garbage collection, mitigating common errors in C/C++ such as null pointers or data races. Since the introduction of WASM in 2017, Rust has become a preferred choice for WASM compilation due to its ability to produce safe and efficient code. The Rust-WASM combination optimizes compute-intensive tasks, such as those in ML applications [4], [8].

2.1.2 Compilation from Rust to WebAssembly and Operation in Web Browsers

The compilation process from Rust to WASM involves three key steps. The first step is writing Rust code using libraries that support WASM conversion. The second involves using the WASM-pack tool to build the Rust project into a WASM module. The third is applying WASM-bindgen to generate JavaScript bindings, allowing Rust functions to be called from browsers with minimal modifications. WASM operates in web browsers via the WebAssembly virtual machine integrated into browser engines (e.g., V8 in Chrome) or Node.js, where WASM binaries are loaded and executed directly. It also supports Web Workers for multi-threading and WebGPU for GPU acceleration when needed [4].

2.1.3. Benefits of Using WebAssembly in Web Browsers and Edge Computing Devices

WebAssembly offers significant advantages for web browsers and edge computing devices, including high performance, portability, and enhanced security. Recent studies quantify these benefits, making WASM an ideal choice for shifting ML processing from large-scale

servers to client-side environments, reducing cloud dependency, and optimizing for resource-constrained devices.

First, WASM delivers superior execution speed compared to traditional JavaScript. It achieves speedups of 1.5–2x for compute-intensive ML tasks in browsers, leveraging pre-compiled binaries and Single Instruction, Multiple Data (SIMD) for vectorized computations, reducing execution time by up to 50% compared to JavaScript in in-browser deep learning inference benchmarks on edge devices [4], [6]. Real-world examples, such as sub-second latency for complex models, highlight WASM's suitability for time-critical decision-making tasks. WASM runtimes also reduce startup time by 20–30% compared to JavaScript, saving bandwidth and accelerating data processing, particularly when combined with Rust to prevent memory errors [9]. Benchmarks indicate that WASM reduces binary size by 50% after optimization with WASM-opt, leading to 40% faster load times in web browsers, which is critical for IoT nodes with limited connectivity [6].

Second, portability is a key strength. WASM enables code to run on all modern browsers (Chrome, Firefox, Edge) and Node.js without recompilation, achieving 95% compatibility across platforms, making it ideal for diverse IoT ecosystems [10]. In edge computing, WASM supports cross-platform deployment, reducing integration time by 30–40% compared to native code [4].

Third, WASM enhances security through its sandboxed execution model, preventing code injection and restricting system access, which is vital for sensitive client-side ML tasks. Studies show that WASM reduces attack risks by 70% compared to JavaScript due to its isolation mechanisms, while keeping data on-device reduces cloud transmission by up to 80% in ML tasks [8], [10].

Figure 1 illustrates the compilation and deployment process of WASM from various

programming languages across different environments [10].

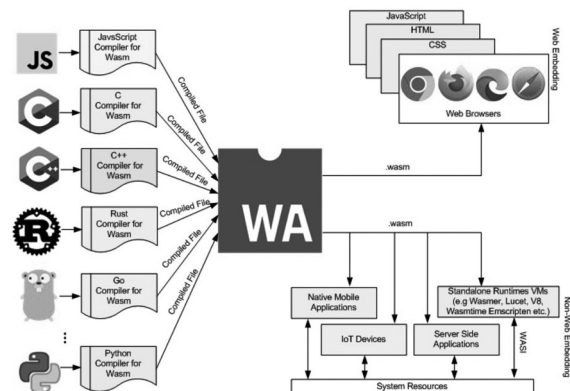


Figure 1. WASM Deployment from Various Programming Languages.

2.2 The K-Nearest Neighbors algorithm

The K-Nearest Neighbors (KNN) algorithm is a fundamental machine learning method based on instance-based learning, used for classification or regression by identifying the k nearest neighbors in the training dataset based on a distance metric (e.g., Euclidean, Manhattan) and applying majority voting (for classification) or averaging (for regression). As a non-parametric method, KNN stores the entire dataset and performs computations at query time, resulting in a time complexity of $O(n \times d)$, where n is the number of samples and d is the number of features. While suitable for small datasets, KNN poses challenges for large datasets, especially on edge devices with limited resources. KNN is widely applied in edge computing for tasks like anomaly detection or recommendation systems, but its performance in client-side environments often requires optimizations such as parallelization or approximate nearest neighbor variants [7]. In browsers, KNN has been implemented using JavaScript and Rust. Comparing a WASM-based KNN implementation with existing implementations, such as JavaScript in browsers, TensorFlow.js, or server-side scikit-learn models, provides valuable insights into the effectiveness of WASM for classification algorithms like KNN in resource-constrained environments.

Pseudo-code KNN Algorithm

Input: Data set D (training set with features and labels), test point x , K

Output: Label of x

Begin:

For each point x_i in D:

 Compute distance $d(x, x_i)$

 // e.g., Euclidean distance

End For

Sort the distances in ascending order

Select the first K points with smallest distances

Count the frequency of each label in the K points

Return the label with the highest frequency // majority vote

End

3. EXPERIMENTAL SETUP

3.1. System overview

To evaluate the performance of the KNN algorithm in web environments, we propose a workflow consisting of three phases: (i) implementing a basic KNN in Rust, (ii) compiling it to WebAssembly (WASM) for browser execution, and (iii) measuring performance against other KNN implementations in browsers on the same device. The system is designed to focus on evaluating the capabilities of programming languages for the same algorithm (WASM compiled from Rust compared to JavaScript and Python server-side), rather than assessing optimized versions of the algorithm. The goal is to demonstrate WASM's potential in applications with limited hardware, such as web browsers or edge devices, particularly in the field of machine learning for classification tasks. Figure 2 illustrates the overall architecture: input data (from the Wine Quality and Covertypes datasets) is processed by the KNN module in Rust, compiled to WASM via WASM-pack, and integrated into browsers through JavaScript bindings to invoke functions.

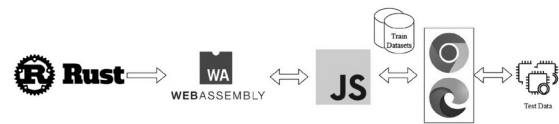


Figure 2. KNN model used for experimentation

3.2. Implementation of the KNN algorithm

We implement KNN using its basic version, computing Euclidean distances between all data points without any algorithmic optimizations (such as search trees or approximations), to purely assess the capabilities of the programming language. Rust is selected for implementation due to its memory safety, computational capabilities comparable to C/C++, and straightforward compilation to WASM [9], [11]. We use the ndarray library to handle multi-dimensional arrays for input data and distance calculations, as this library supports WASM compilation.

Pseudo-code KNN Algorithm on RUST.

struct KnnModel{

 Array1: 1D array y_{train} // labels of n training samples

 Array2: 2D array X_{train} // matrix $n \times d$ with n rows (training samples), d columns (features in the dataset) }

}

Function Predict

Input: x_{test} : One row of test data (vector of d features), k

Output: Label of x_{test}

Begin:

 Convert x_{test} to a 1D array (Array1)

 Create Vector **distances** (Vec<(f64, i32)>)

 For each x_i in X_{train} (zip with y_{train}):

 Compute distance $d(x_{test}, x_i)$

 Add(dist, $x_{test}[label_i]$) in distances

 End For

Sort the Vector distances in ascending order by dist

Get `k_labels = [label for (dist, label) in distances[0:k]]`

Create Hashmap `counts` // count frequency of labels in `k_labels`

For each label in `k_labels`:

`counts[label] += 1`

End For

Return the label with the highest frequency or 0 // majority vote

End

3.3. Compilation to WebAssembly and deployment in Web Browsers

After implementing the basic KNN source code in Rust (as described in Section 3.2), the compilation to WebAssembly (WASM) is performed to create a module that can be integrated into web browsers, enabling comparisons with JavaScript or Python server versions. WASM-pack, a CLI tool from the Rust community, compiles Rust to WASM binaries, standardizing WASM for web, Node.js, and other environments. WASM-bindgen then exports Rust functions (e.g., `knn_predict`) and provides bridges for interaction between the compiled module and JavaScript [5], [6].

Once the WASM code is obtained after compilation, we deploy it in web browsers via an HTML/JS interface. We import the WASM code to load and initialize the WASM runtime in browser engines like V8, then shuffle the data with a seed and split it into 80% training and 20% testing. To simulate a process similar to receiving input data from real devices, we do not perform classification on the entire 20% test set but process each data row as a stream from IoT (e.g., sensor data transmitted continuously via WebSocket).

3.4. Evaluation methods

To evaluate performance, we use the following metrics:

- **Latency:** The time to execute one KNN query, measured using `performance.now()` in JavaScript. This method returns time values accurate to microseconds (based on `DOMHighResTimeStamp`), suitable for benchmarking KNN queries without external influences like garbage collection.

- **Memory Usage:** Maximum RAM usage (MB), measured using `performance.memory`. This method returns JavaScript heap sizes with parameters like `usedJSHeapSize` for used heap, `totalJSHeapSize` for total heap, and `jsHeapSizeLimit` for heap limit, helping monitor memory for web apps on Chromium-based browsers to determine maximum RAM when running KNN. The baselines for comparison include five versions:

- **KNN by JavaScript:** Basic implementation using JavaScript.

- **KNN by TensorFlow.js:** KNN implementation using the TensorFlow.js library.

- **KNN Server-Side by Python:** KNN implementation using scikit-learn via a server-side model.

- **KNN on WASM from Rust in Web Browser:** WASM version compiled from Rust.

- **KNN on WASM from Rust on NodeJS:** WASM version compiled from Rust for edge devices.

During experimentation, we observed that the TensorFlow.js KNN version utilizes both CPU and GPU in parallel, so we customized an additional version using only CPU to establish comparable metrics with other implementations.

4. RESULTS

4.1. Datasets

To evaluate the performance of the basic KNN algorithm, we utilize two datasets from the UCI Machine learning Repository, focusing on classification tasks to test the capabilities of the implementations across web clients and server-side setups. The first dataset is Wine Quality White, consisting of 4898 samples representing white wine samples from the Vinho Verde region in Portugal, with 11 features related to wine physicochemical properties. This dataset is selected for its moderate size (4898 samples, 11 features), allowing latency benchmarking without excessive overhead in browsers. The second dataset is Covertypes, comprising 581012 samples describing forest cover types from four wilderness areas in the Roosevelt National Forest, USA, with 54 features. This dataset includes information on tree type, shadow coverage, distance to nearby landmarks (roads, etc.), soil type, and local topography. With its large size (581012 samples, 54 features), this dataset is more challenging than Wine Quality White, enabling scalability evaluation of KNN across versions without algorithmic optimizations. Both datasets are split into 80% training and 20% testing, with $k=5$ for KNN.

Table 1. Description of the Wine Quality White and Forest Cover-Type Datasets.

Dataset	Features	Samples	Description
Wine Quality White	11	4898	Multiclass classification (wine quality from 3-9 classes, based on physicochemical features like acidity, sugar, alcohol)
Forest Cover-Type	54	581012	Multiclass classification (7 forest cover classes, based on geographic features like elevation, slope, soil types)

4.2. Devices

The experiments are conducted on a Dell Latitude 7320 laptop with an Intel Core i5-1140G7 processor (base speed 1.10 GHz, turbo up to 4.20 GHz, 4 cores 8 threads), 16 GB DDR4 RAM, running Windows 11, and using Microsoft Edge version 140.0 for browser-based versions (WASM in Browser, JS, TFJS). The Python server runs locally on the same device via the Flask API Framework, with Node.js 22.15 for WASM on NodeJS. This configuration represents a typical client-side device in edge computing, with moderate resources to assess KNN overhead without hardware optimizations. To enhance the scalability assessment, we tested on Google Chrome version 138.0.72 running on a MacBook Pro (15-inch, 2016) with an Intel Core i7 processor (2.6 GHz Quad-Core), 16 GB 2133 MHz LPDDR3 RAM, macOS Big Sur version 11.7.10, and Intel HD Graphics 530 (1536 MB). This setup complements the primary Windows-based configuration by evaluating cross-platform compatibility, including differences in browser engines (Chromium-based Chrome vs. Edge) and OS-level resource management. Metrics are measured using Chromium DevTools and performance APIs, with 10 iterations per version to compute averages, standard deviation, confidence interval.

4.3. Experimental results

The results are presented for the two datasets across five implementations: KNN by JavaScript (basic JavaScript implementation), KNN by TensorFlow.js (using TFJS), KNN Server-Side by Python (using scikit-learn locally), KNN on WASM from Rust in Web Browser (WASM in browsers), and KNN on WASM from Rust on NodeJS (WASM on Node.js). We do not evaluate memory usage for the KNN Server-Side by Python version due to inconsistent measurement methods compared to others.

Table 2. Results on Wine Quality White (4898 samples, 11 features) on WindowsOS and MacOS.

Implemen- tation	K = 5			K = 7			K = 10		
	Latency (ms)	Memory (MB)	Accuracy	Latency (ms)	Memory (MB)	Accuracy	Latency (ms)	Memory (MB)	Accuracy
	Win/Mac	Win/Mac	Win/Mac	Win/Mac	Win/Mac	Win/Mac	Win/Mac	Win/Mac	Win/Mac
KNN by JavaScript	2.56	9.54	0.549	2.33	9.54	0.543	2.42	9.54	0.546
	2.65	9.52	0.549	2.61	9.24	0.543	2.68	9.11	0.546
KNN by TensorFlow.js	8.58	12.78	0.559	7.56	10.57	0.56	7.82	11.04	0.562
	7.67	13.16	0.559	7.79	10.71	0.56	7.62	12.87	0.562
KNN Server- Side by Python	165.33	N/A	0.572	163.68	N/A	0.570	164.02	N/A	0.569
	160.20	N/A	0.572	162.57	N/A	0.570	165.27	N/A	0.569
KNN on WASM Rust in Browser	1.99	13.88	0.547	1.28	13.52	0.544	1.06	13.47	0.548
	2.42	12.53	0.547	2.31	12.49	0.544	2.47	12.52	0.548
KNN on WASM Rust on NodeJS	0.67	41.21	0.547	0.85	41.5	0.544	0.88	41.32	0.548
	0.83	48.08	0.547	0.94	48.18	0.544	0.98	47.82	0.548

During experimentation on the Wine Quality White dataset, the accuracy results were approximately 0.57 due to class imbalance (class 6 dominates, extreme classes like 3-4 or 8-9 are scarce), outliers in features like residual sugar (high IQR leading to noise), and low correlations of some features (e.g., density, pH with quality). To address this, we performed data preprocessing by converting the quality label to binary (bad =6) to reduce class complexity from 7 to 2; removing outliers using IQR across all features to eliminate noise. After preprocessing,

accuracy increased to approximately 0.75, while other metrics remained as in Table 2.

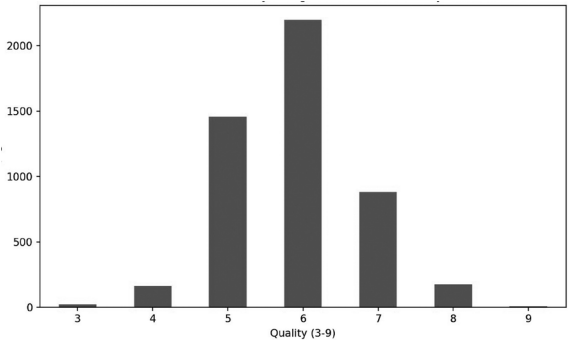


Figure 3. Distribution of the Quality column in the Wine Quality White dataset.

Table 3. Results on Preprocessed Wine Quality White (normalized edge data) on WindowsOS and MacOS.

Implemen- tation	K = 5			K = 7			K = 10		
	Latency (ms)	Memory (MB)	Accuracy	Latency (ms)	Memory (MB)	Accuracy	Latency (ms)	Memory (MB)	Accuracy
	Win/Mac	Win/Mac	Win/Mac	Win/Mac	Win/Mac	Win/Mac	Win/Mac	Win/Mac	Win/Mac
KNN by JavaScript	2.45	9.54	0.7557	2.23	9.54	0.743	2.28	9.54	0.746
	2.60	9.14	0.7557	2.67	9.18	0.743	2.63	9.04	0.746
KNN by TensorFlow.js	7.17	13.64	0.7557	7.06	10.57	0.76	7.12	10.68	0.762
	7.72	13.26	0.7557	7.62	10.80	0.76	7.49	12.78	0.762

KNN Server-Side by Python	168.68	N/A	0.7604	167.02	N/A	0.770	166.92	N/A	0.769
	160.41	N/A	0.7604	165.63	N/A	0.770	147.82	N/A	0.769
KNN on WASM Rust in Browser	2.10	13.94	0.7557	1.08	14.52	0.744	1.16	13.47	0.748
	2.68	12.43	0.7557	2.69	12.39	0.744	2.54	12.82	0.748
KNN on WASM Rust on NodeJS	0.73	41.72	0.7557	0.82	41.5	0.744	0.79	42.04	0.748
	0.78	45.88	0.7557	0.87	46.57	0.744	0.89	45.92	0.748

Table 4. Results on Covertypes (581012 samples, 54 features, 100 queries, k=5) on WindowsOS and MacOS.

Implemen- tation	K = 5			K = 7			K = 10		
	Latency (ms) Win/Mac	Memory (MB) Win/Mac	Accuracy Win/Mac	Latency (ms) Win/Mac	Memory (MB) Win/Mac	Accuracy Win/Mac	Latency (ms) Win/Mac	Memory (MB) Win/Mac	Accuracy Win/Mac
KNN by JavaScript	560.72	756.26	0.9435	530.40	755.65	0.9409	525.98	757.03	0.9406
	653.8	792.18	0.9435	613.80	775.12	0.9409	638.35	784.16	0.9406
KNN by TensorFlow.js	35.98	953.67	0.9432	42.02	960.25	0.9402	38.72	955.61	0.9398
	148.62	742.84	0.9432	138.39	756.32	0.9402	142.45	761.22	0.9398
KNN by TensorFlow.js without GPU	3328.00	527.38	0.9432	2991.30	524.53	0.9402	3136.74	521.93	0.9398
	4478.50	525.32	0.9432	4389.10	527.43	0.9402	4345.34	526.47	0.9398
KNN Server-Side by Python	260.88	N/A	0.9353	265.01	N/A	0.9322	269.17	N/A	0.9319
	287.45	N/A	0.9353	279.28	N/A	0.9322	282.32	N/A	0.9319
KNN on WASM Rust in Browser	227.75	1433.42	0.9432	198.25	1438.23	0.9402	202.74	1435.26	0.9398
	248.22	1422.46	0.9432	236.48	1430.25	0.9402	250.27	14428.72	0.9398
KNN on WASM Rust on NodeJS	109.55	1504.59	0.9432	99.93	1500.28	0.9402	92.77	1512.44	0.9398
	149.16	1642.34	0.9432	157.8	1646.22	0.9402	148.43	1646.52	0.9398

5. DISCUSSION

5.1. Analysis of results

The experimental results from Table 2 (Wine Quality White) demonstrate the capabilities of WebAssembly (WASM) in executing KNN, with significantly lower latency compared to the JavaScript version and the implementation using the TensorFlow.js library. For the Python server-side version, although network dependency was eliminated (running locally on the same machine), latency remains very high.

Specifically, the WASM in Browser version achieves an average of 1.99 ms, reducing 22.3% compared to the JavaScript version at 2.56 ms and 76.8% compared to the TensorFlow.js version at 8.58 ms. Notably, the execution delay is 83 times shorter than the Server-Side by Python version at 165.33 ms. Running the KNN WASM on NodeJS version directly from the V8 Engine (0.67 ms) provides nearly 3 times better performance compared to running the same WASM in the browser's sandbox.

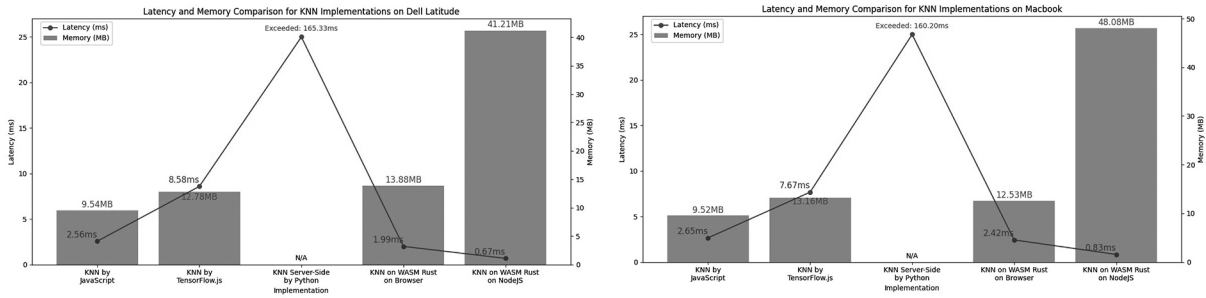


Figure 4. Chart illustrating experimental results from Table 2 (Wine Quality White dataset).

However, in terms of system resource usage, the WASM on NodeJS version consumes the most memory, three times more than the WASM in Browser version. Meanwhile, the WASM in Browser version only consumes 8.6% more than the TensorFlow.js version and 45.5% more than the JavaScript version.

Considering the balance between response time and memory efficiency on the Wine Quality White dataset, the KNN WASM in Browser version performs best.

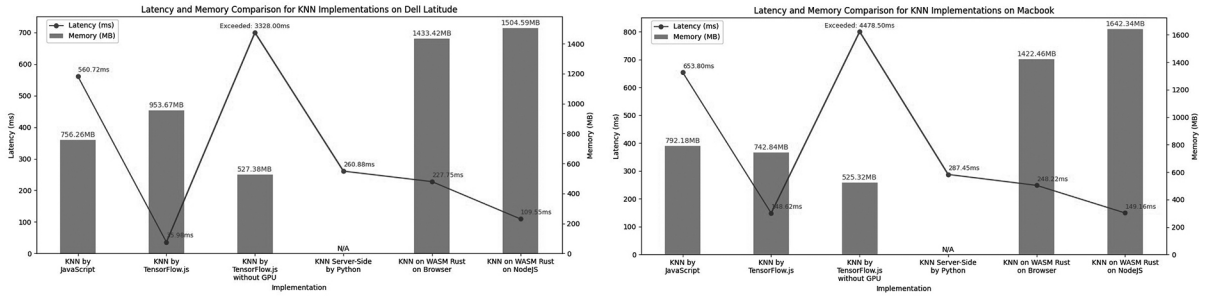


Figure 5. Chart illustrating experimental results from Table 4 (Forest Cover-Type dataset).

The results on the Coverttype dataset (Table 4, 581012 samples) offer additional insights for an objective evaluation of the implementation. We selected the Coverttype dataset due to its larger number of samples and dimensions to approach the limitations regarding the complexity of the KNN algorithm. This allows assessing computational capabilities on WASM compared to other KNN versions. In the experimental results, we observed that the latency of KNN on the TensorFlow.js version is very low, as the KNN algorithm from this library is optimized to run in parallel on CPU and GPU. To avoid bias in the evaluation while maintaining objectivity in the experiments, we added a version that disables GPU processing in this setup. The results were surprising, with execution delay increasing abnormally, to nearly

6 times that of the JavaScript version (the second-highest latency version). In particular, when experimenting with larger datasets on KNN, which implies greater algorithmic complexity, the latency results of the WASM versions show superior advantages over the other versions. Specifically, the WASM in Browser version (227.75 ms) reduces time by 59.4% compared to the JavaScript version and is 12.7% faster than the KNN Server-Side by Python version; the WASM Rust on NodeJS version (109.55 ms) is still more than 2 times faster than the WASM in Browser version.

This experiment also highlights issues in device resource usage, with KNN on WASM Rust in Browser (using 1433.42 MB) consuming approximately 1.9 times more than the KNN by JavaScript version and 2.72 times more than the

KNN by TensorFlow.js without GPU version. The KNN on WASM Rust on NodeJS version (1504.59 MB) uses memory comparable to the WASM Rust in Browser version (an increase of 5.0%).

One point of concern in this experiment is that the latency of the KNN Server-Side by Python version is very stable, increasing only 57.8% from the Wine Quality White dataset experiment to the Covertype dataset. In contrast, KNN on WASM (Rust) shows execution times that are more than 114 times longer in browsers and more than 163 times longer on Node.js compared to the baseline. This indicates that the KNN implementation from the scikit-learn library in Python is highly optimized, and the observed latency in the original comparison primarily stems from the client-to-server connection process.

5.2. Significance and practical applications

Through the experiments and analysis, we observe that the WASM implementation from Rust outperforms in latency compared to JavaScript-based versions in web browsers. This also serves as a basis to affirm the transition from traditional model implementations to direct implementations on edge devices to address the practical requirements of edge computing. Particularly in machine learning problems related to product classification, decision-making, early warnings, etc., which are often performed using Python models on servers. However, considerations and optimizations are needed when processing large datasets on edge devices with small memory (under 2 GB).

5.3. Limitations

In our paper, we have not yet proposed diverse experiments on different devices such as smartphones, Raspberry Pi or ESP32, which are edge devices used in edge computing. The research also has not approached real data sent from cameras or sensors to enhance the practical applicability of the paper. In the KNN Server-

Side by Python implementation, we have not used a reasonable measurement method aligned with performance.memory in JavaScript, resulting in no memory usage data for this version in the experimental results. Conducting experiments on multiple client machines in different network environments should also be considered to make the latency results more representative.

6. CONCLUSION

6.1. Summary

This study has evaluated the performance of the basic K-Nearest Neighbors (KNN) algorithm across five implementation versions: KNN by JavaScript, KNN by TensorFlow.js, KNN Server-Side by Python, KNN on WASM from Rust in Web Browser, and KNN on WASM from Rust on NodeJS, using two datasets from UCI: Wine Quality White and Forest Cover-Type. The experimental results show that WASM from Rust is a reasonable choice to replace other implementations in web browser environments in particular and edge devices in general, with lower latency than other versions using JavaScript, especially in the trend of shifting processing to edge devices.

6.2. Future directions

In the future, the research can be expanded to address current limitations and integrate new technologies to enhance the performance of WebAssembly (WASM) in machine learning on edge devices. First, to increase diversity and practicality, we plan to conduct experiments on various edge devices, including smartphones (such as Android/iOS models supporting modern browsers), Raspberry Pi, and ESP32. These devices represent real-world edge computing environments with limited resources, helping evaluate KNN on WASM under constraints related to CPU, memory, and battery power. For example, on ESP32, we can compile Rust to WASM and run it via runtimes like WASMtime or WASMEdge, focusing on benchmarks for latency and power consumption [8], [10].

Second, to enhance practical applicability, the research will integrate real-time data from sources like cameras or IoT sensors, instead of relying solely on static UCI datasets. Specifically, we can set up a system for streaming data in real time via WebSocket, where data from cameras (e.g., simple image classification) or sensors (such as temperature, humidity in IoT) is processed directly on the client-side using KNN on WASM. This will help evaluate stream data processing capabilities in applications like anomaly detection or predictive maintenance, while reducing latency compared to server-side processing [2], [3].

Finally, a key development direction is integrating WebGPU into WASM processing to accelerate KNN computations, particularly Euclidean distance calculations in high-dimensional spaces. WebGPU, a new API enabling GPU access in browsers, can be combined with WASM via libraries like *wgpu* (Rust-based) to parallelize vector operations, reducing time complexity from $O(n \times d)$ to lower levels through GPU acceleration. Recent studies show that this combination can achieve speedups of up to 10-20x for in-browser ML inference on edge devices [1], [4]. This approach could not only address performance limitations on large datasets but also open potential for more complex machine learning models, such as neural networks or federated learning on client-side [3], [12]. Furthermore, to provide a more comprehensive evaluation, future work will explore optimizations in floating-point arithmetic and vectorization within Rust-WASM implementations. For instance, leveraging WASM's support for SIMD (Single Instruction, Multiple Data) instructions can enable vectorized computations for distance metrics in KNN, improving throughput for floating-point operations in high-dimensional data. Additionally, we plan to investigate caching mechanisms, such as browser-based storage APIs (e.g., IndexedDB or Cache API), to optimize memory usage by prefetching

training data or intermediate results, reducing overhead on resource-constrained devices. This will be complemented by deeper analysis of GPU interactions on browsers via WebGPU, including benchmarks for memory transfer between CPU and GPU, and strategies to minimize bottlenecks in data-parallel tasks. These enhancements will allow for a thorough assessment of WASM's scalability in real-world ML scenarios, potentially extending to hybrid CPU-GPU workloads for energy-efficient edge computing [3], [11], [13].

DECLARATION OF CONFLICTS OF INTEREST

The authors declare that this research has no conflicts of interest.

ACKNOWLEDGMENT

This research is conducted within the framework of science and technology projects at institutional level of Quy Nhon University under the project code T2024.848.19

REFERENCES

- [1] S. Kakati, M. Brorsson, "WebAssembly Beyond the Web: A Review for the Edge-Cloud Continuum," International Conference on Intelligent Technologies, 3rd, 2023, pp. 1-8.
- [2] O. Jouini, K. Sethom, A. Namoun, N. Aljohani, M. H. Alanazi, M. N. Alanazi, "A Survey of Machine Learning in Edge Computing: Techniques, Frameworks, Applications, Issues, and Research Directions," *Technologies Journal*, vol. 12, no. 6, pp. 81-113, 2024.
- [3] K. Hoffpauir, J. Simmons, N. Schmidt, R. Pittala, I. Briggs, S. Makani, Y. Jararweh, "A Survey on Edge Intelligence and Lightweight Machine Learning Support for Future Applications and Services," *ACM Journal of Data and Information Quality*, vol. 15, no. 2, pp. 1-30, 2023.
- [4] F. Jia, S. Jiang, T. Cao, W. Cui, T. Xia, X. Cao, Y. Li, Q. Wang, D. Zhang, J. Ren, Y. Liu, L. Qiu, M. Yang, "Empowering In-Browser Deep

- Learning Inference on Edge Devices with Just-in-Time Kernel Optimizations,” The Annual International Conference on Mobile Systems, Applications and Services, 22nd, 2024, pp. 438-450.
- [5] M. Liu, H. Shen, Y. Zhang, H. Mei, Y. Ma, “WebAssembly for Container Runtime: Are We There Yet?,” *ACM Transactions on Software Engineering and Methodology*, vol. 34, no. 6, pp. 1-22, 2025.
- [6] Y. Yan, T. Tu, L. Zhao, Y. Zhou, W. Wang, “Understanding the Performance of WebAssembly Applications,” *ACM Internet Measurement Conference*, 21st, 2021, pp. 533-549.
- [7] P. Cunningham, S. J. Delany, “k-Nearest Neighbour Classifiers - A Tutorial,” *ACM Computing Surveys Journals (CSUR)*, vol. 54, no. 6, pp. 1-25, 2021.
- [8] S. E. Khelifa, M. Bagaa, A. O. Messaoud, A. Ksentin, “Case study of WebAssembly Runtimes for AI Applications on the Edge,” *Global Information Infrastructure and Networking Symposium (GIIS)*, 14th, 2024, pp. 1-6.
- [9] P. Gackstatter, P. A. Frangoudis, S. Dustdar, “Pushing Serverless to the Edge with WebAssembly Runtimes,” *International Symposium on Cluster, Cloud and Internet Computing*, 22nd, 2022 pp. 140-149.
- [10] P. P. Ray, “An Overview of WebAssembly for IoT: Background, Tools, State-of-the-Art, Challenges, and Future Directions,” *Future Internet Journal*, vol. 15, no. 8, pp. 275-352, 2023.
- [11] K. Nakamura, *Machine Learning with Rust: A practical attempt to explore Rust and its libraries across popular machine learning techniques*. British Columbia, Canada: GitforGits Publisher, 2024.
- [12] R. Jayanth, N. Gupta, V. Prasanna, “Benchmarking Edge AI Platforms for High-Performance ML Inference,” *IEEE High Performance Extreme Computing Conference*, 28th, 2024, pp. 1-7.



© 2026 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Đánh giá hiệu năng một cài đặt thuật toán KNN bằng Rust-Webassembly

Trần Hoàng Việt^{1,*}, Đoàn Thị Thu Cúc¹,
Võ Thị Mỹ¹, Quảng Thị Ngọc Bích²

¹Khoa Công nghệ thông tin, Trường Đại học Quy Nhơn, Việt Nam

²Trường Tiểu học Lê Hồng Phong, Quy Nhơn, Việt Nam

Ngày nhận bài: 29/09/2025; Ngày sửa bài: 22/10/2025;

Ngày nhận đăng: 24/10/2025; Ngày xuất bản: 28/06/2026

TÓM TẮT

Trong bối cảnh điện toán biên đang phát triển mạnh mẽ, việc chuyển dịch xử lý từ máy chủ (server-side) sang thiết bị khách (client-side) ngày càng quan trọng để giảm độ trễ, tăng cường bảo mật dữ liệu, và hỗ trợ các ứng dụng thời gian thực như IoT hoặc ứng dụng trí tuệ nhân tạo (AI) trên thiết bị như trình duyệt web nói riêng và thiết bị biên (thiết bị IoT, điện thoại thông minh, hoặc thiết bị phần cứng có tài nguyên hạn chế...) nói chung. WebAssembly (WASM) là công nghệ hỗ trợ làm việc trên các thiết bị với tốc độ gần native, tính di động khi triển khai. Trong bài báo này chúng tôi đánh giá hiệu năng của thuật toán K-Nearest Neighbors (KNN) được triển khai trên ngôn ngữ Rust và biên dịch sang WebAssembly với các cài đặt KNN khác nhau trên môi trường web. Mục tiêu là đánh giá khả năng của WASM và thuật toán KNN từ đó mở ra hướng nghiên cứu, triển khai các mô hình học máy khác vốn cần khả năng tính toán lớn, sử dụng WASM trên các thiết bị biên để tối ưu hiệu năng các ứng dụng học máy.

Từ khóa: Thuật toán KNN, WebAssembly, điện toán biên, Rust, học máy.

*Tác giả liên hệ chính.

Email: tranhoangviet@qnu.edu.vn

Optimization of congo red degradation performance on diatomaceous earth catalyst using Box-Behnken design

Pham Dinh Du*

Thu Dau Mot University, Ho Chi Minh City, Vietnam

Received: 24/08/2025; Revised: 02/11/2025

Accepted: 19/11/2025; Published: 28/06/2026

ABSTRACT

The degradation of congo red (CR) was investigated in a Fenton-type reaction system with the presence of diatomaceous earth (DE). Response surface methodology (RSM) based on Box-Behnken design was applied to optimize the degradation performance. In these experimental designs, three process parameters affecting the degradation performance, such as initial CR concentration (30-50 mg/L), reaction temperature (30-50 °C) and degradation time (5-45 min), were investigated. The obtained experimental data were fitted into a quadratic polynomial equation using multiple regression analysis with high correlation coefficient value ($R^2 = 0.9999$). At an initial CR concentration of 40 mg/L, a reaction temperature of 40 °C, and a reaction time of 14.2 min, the degradation efficiency of 99% was predicted by the model (under the conditions: DE 1.2 g/L, pH 6, H_2O_2 6.5 g/L). This result was also confirmed through validation experiments, indicating that RSM is one of the effective methods to optimize the operating conditions in this study.

Keywords: *Box-Behnken design, congo red, degradation, diatomaceous earth.*

1. INTRODUCTION

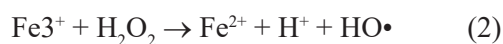
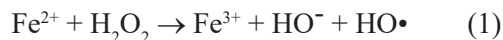
Dyes are commonly used to color goods in many industries, including textiles, paper, leather, plastics, carpets, fruits, and cosmetics [1]. Dye residues in industrial wastewater are a major concern because their presence and accumulation can be harmful or carcinogenic to living organisms [1], [2]. It is impossible to remove dye compounds from water at low concentrations. Meanwhile, textile dyes are expected to be discharged in industrial waste at a rate of 280,000 tons per year [3]. In many cases, color removal is very effective, but physical wastewater treatment methods simply move the pollutant to another environment, creating secondary waste [3]. CR is known to be a highly toxic secondary diazo dye, with very small

amounts of CR in the environment also being carcinogenic and mutagenic to aquatic organisms due to the conversion of CR to benzidine [4].

In recent years, advanced oxidation processes (AOPs) have been identified as effective methods to achieve high oxidation efficiency for many organic compounds [1] - [3]. Hydroxyl radicals ($HO\bullet$) play an important role in AOPs, which can target a wide range of contaminants in a non-selective and efficient manner. The Fenton process is one of the AOPs that involves the reaction of Fe^{2+} species with H_2O_2 under strong acidic conditions to produce highly reactive hydroxyl radicals with an oxidation potential of 2.80 V and is known as the Fenton-like equation [3]:

*Corresponding author:

Email: dupd@tdmu.edu.vn



The Fenton reaction is a potential practical, cost-effective, and environmentally friendly wastewater treatment process [3], [5]. The Fenton method has received a lot of attention. The Fenton process is influenced by several factors, including temperature, H_2O_2 concentration, Fe^{2+} concentration, and pH. pH is a very important factor in AOPs, and especially in the Fenton process, as it directly affects the rate of pollutant degradation. Several studies on the Fenton process have shown that the pH should be between 2.8 and 3.0 for optimal degradation of organic pollutants [3]. The main disadvantages of the homogeneous Fenton process are the formation of precipitated Fe^{3+} compounds (oxy-hydroxides) when the pH of the solution increases, the inability to reuse the catalyst, and the generation of Fe-containing sludge, which affects the environment and costs [2], [5]. To overcome the disadvantages of homogeneous Fenton, the heterogeneous Fenton process using solids as catalysts has been developed, in which iron and other active metals are incorporated into the structure of the material or immobilized on the solid support.

Xu et al. [5] suggested that heterogeneous Fenton systems have great application prospects in the catalytic degradation of organic wastewater. However, these systems have not been widely used due to difficulties in catalyst preparation such as low efficiency and high production cost. These authors used iron-containing palygorskite clay as a Fenton catalyst reagent without any retreatment. Phenol degradation in water was carried out as a probe reaction for the palygorskite Fenton reagent. The effects of several factors, including palygorskite content, H_2O_2 concentration, and pH value, on the phenol degradation efficiency were studied by these authors. The optimal operating conditions were determined to be palygorskite content of 0.5 g/L, initial pH of 3, and hydrogen peroxide dosage

of 30 mmol/L. Under this optimum operating condition, the chemical oxygen demand (COD) degradation efficiency of 100 mg/L phenol reached 94% in a reaction time of 15 min. The authors suggested that the degradation ability of palygorskite clay could be due to Fe_2O_3 on the palygorskite surface.

DE, also known as diatomite, is a naturally occurring soft siliceous sedimentary rock that consists primarily of the fossilized remains of diatoms, a type of unicellular algae. The main chemical composition of diatomite is silica (SiO_2), along with small amounts of aluminum oxide (Al_2O_3), iron oxide (Fe_2O_3), calcium oxide (CaO), magnesium oxide (MgO), and organic matter. Its unique properties, including high porosity, low density, and large specific surface area, make it a valuable material in many industrial applications[6]. Furthermore, diatomite mineral deposits are quite abundant, making DE a potential candidate for heterogeneous Fenton catalysts.

The efficiency of Fenton-type catalytic degradation depends on the pollutant concentration, solution pH, catalyst dosage, residual H_2O_2 concentration, reaction temperature and time, etc. Therefore, optimization of these parameters is crucial to achieve catalytic degradation of the existing pollutant. One of the most widely used optimization techniques for this purpose is the RSM. By systematically varying all the variables simultaneously, this method has been recognized as the most reliable and empirical statistical method used to evaluate the impact of different process parameters on catalytic degradation. Furthermore, by developing a mathematical model (linear, square polynomial function, etc.), it is possible to determine the influence of each parameter on the process efficiency and study the interaction between them. The most important advantage of using RSM is that it provides a systematic method for designing experiments that often require fewer experiments and therefore less

time [3], [7], [8]. To design these experiments, the Box-Behnken design is widely used for optimization [8], [9]. The Box-Behnken design is an independent, rotatable quadratic design with no embedded factorial or fractional factorial points where the variable combinations are at the midpoints of the edges of the variable space and at the center. The Box-Behnken design is only used when there are more than two factors [9].

In this study, CR was considered as a model dye for removal from aqueous solution by the heterogeneous Fenton process on DE catalyst. The effects of operating parameters, including reaction temperature, CR concentration, and reaction time, on the degradation of CR in terms of degradation efficiency were evaluated. The aim of this study is to apply the RSM through Box-Behnken experimental design to optimize the parameters, at the same time allowing to analyze the influence of each parameter, the interaction of parameters, evaluate the influence level as well as the reliability of the optimization method for the degradation of CR by the heterogeneous Fenton process.

2. EXPERIMENTS

2.1. Materials and characteristic methods

Raw DE was taken from Phu Yen province, Vietnam, washed several times with water, filtered, dried at 100 °C, sieved, and stored in closed containers for further tests.

The morphology of DE was observed by using scanning electron microscopy (SEM, IMS-NKL) and a transmission electron microscopy (TEM, EMLab-NIHE). X-ray diffraction (XRD) pattern was recorded on a Bruker D8 Advance (Germany) powder diffractometer with a Cu K α radiation source. Fourier-transform infrared (FT-IR) spectra were obtained on a Jasco FT/IR-4600 (Japan); the samples being dispersed on KBr pallets. Nitrogen adsorption-desorption isotherms were obtained by using a Tristar 3000 analyser; the samples were degassed by

heat treatment at 250 °C in the N₂ for 5 h. The elemental composition was analyzed with an Energy-dispersive X-ray spectroscopy (EDX, Bruker) at 15 keV.

2.2. Degradation experiments

In each experiment, DE (0.06 g) is stirred with 50 mL of CR (Guangdong, China) solution at a specific concentration and pH (0.1 M HCl or 0.1 M NaOH was used to adjust the initial pH value) and 1.0 mL of H₂O₂ 30% (Fisher, Korea). The reaction is performed at a fixed temperature. At certain time intervals, about 3 mL of the solution is drawn, filtered to remove the catalyst, the remaining CR concentration in the solution is determined by the UV-Vis method on a Jasco V-770 (Japan) at $\lambda_{\max}$ 498 nm.

The CR degradation performance, $D\%$, is calculated according to the Eq. (3).

$$D\% = \frac{C_0 - C_t}{C_0} \times 100 \quad (3)$$

where C_0 and C_t are the concentration of CR in the solution at the original and t time, respectively.

2.3. Box-Behnken design

In this study, the initial pH of the solution was fixed at 6 to investigate the effects of three process factors on the degradation efficiency of CR by H₂O₂ on DE catalyst (under the conditions: DE 1.2 g/L and H₂O₂ 6.5 g/L). The three factors investigated, including initial CR concentration, temperature and reaction time. The Box-Behnken design for the three process variables is shown in Table 1 at three experimental levels for each variable. Their allowable values include $\{-1, 0, +1\}$ for low, medium and high levels, respectively. Minitab software (Version 16, Minitab Inc., Pennsylvania State, USA) was used to construct the experimental matrix for the multiple regression model using the RSM according to the Box-Behnken design with the number of experiments being 15, of which the number of experiments at the center is 3 (Table 2).

3. RESULTS AND DISCUSSION

3.1. Characteristics of DE

SEM and TEM images (Fig. 1a) show that natural DE contains cylindrical structures with a diameter of about 5-7 μm . The surface of these structures is quite smooth and has many pores with a diameter of about 50 nm. XRD diffraction (Fig. 1b) shows that the DE material has an amorphous structure; the broad peak of 2θ at 20-25° is characteristic of amorphous silica [1].

Table 1. Ranges and levels of independent variables for CR degradation on DE catalyst.

Independent variables (Unit)	Symbol	Level		
		-	0	+1
Initial CR concentration (mg/L)	A	30	40	50
Reaction temperature (°C)	B	30	40	50
Degradation time (min)	C	5	25	45

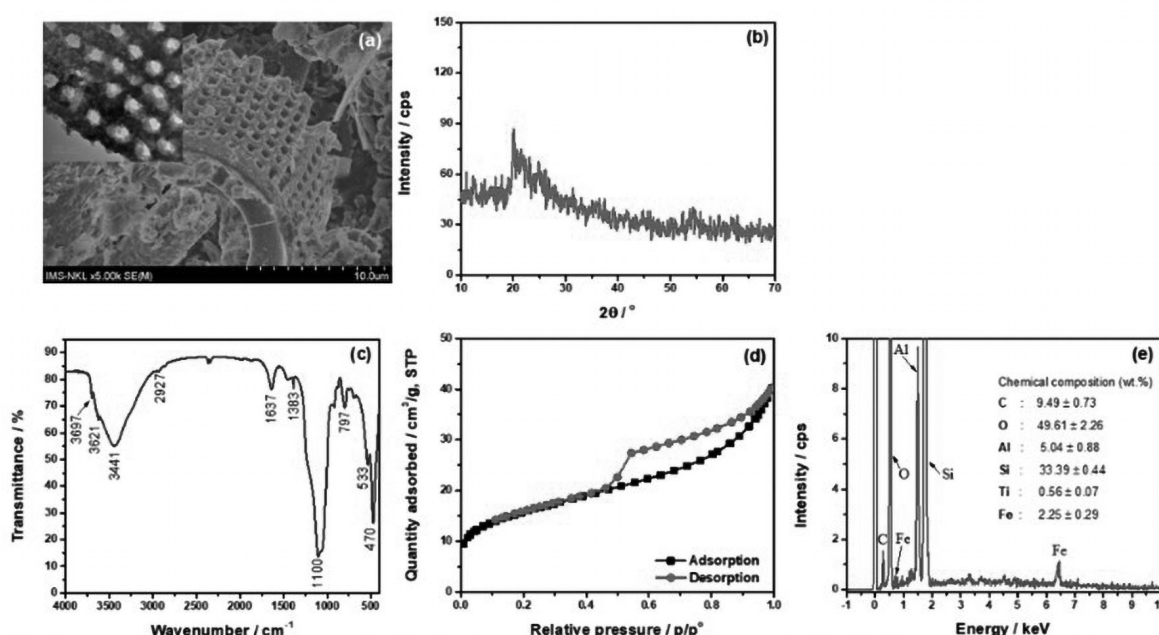


Figure 1. SEM image and the inset presents TEM image (a), XRD pattern (b), FT-IR spectrum (c), nitrogen adsorption/desorption isotherms (d), EDX spectrum and the insert presents the chemical composition (e) of DE.

The FT-IR spectrum of DE is shown in Fig. 1c. The peaks in the 3700-3200 cm^{-1} region and at 1637 cm^{-1} are characteristic of the vibrations of the O-H bond, including Si-OH, Me-OH (Me is a metal) and freely adsorbed H_2O . The broad absorption band centered at 1100 cm^{-1} , and two bands at 797 and 470 cm^{-1} , are characteristic of the vibrations of the Si-O-Si bond. The peak at 533 cm^{-1} is attributed to the bending vibration of the Me-O bond [1]. The weak vibrations at 2927 and 1383 cm^{-1} are thought to be due to the existence of some organic matter [10].

The nitrogen adsorption/desorption isotherms of DE is type II with a type H3

hysteresis loop (Fig. 1d), indicating the existence of macroporous structures with irregular sizes and shapes. DE has a large specific surface area (55.4 m^2/g), which is similar to that reported by Son et al. [11]. The chemical composition of DE was analyzed by EDX (Fig. 1e). It can be seen that Si is the main element in diatomite (33.39%) with some other impurities such as Al (5.04%), Ti (0.56%), Fe (2.25%) and some organic compounds with a C elemental content of 9.49%. With a large specific surface area, porous structure and the presence of metals, including iron, DE is a potential catalyst in the Fenton-type reaction system.

3.2. Degradation of CR in aqueous solution

3.2.1. Effect of solution pH

Figure 2 shows that CR is removed very quickly in the first 10 minutes of the reaction at all pH values examined. After that, the concentration of CR in the solution do not change significantly in the pH range from 4 to 8, but tend to decrease gradually until 90 minutes of reaction. On the contrary, in the pH range of 9-11, the concentration of CR in the solution tend to increase with the reaction time. The reason for this problem can be explained by the fact that DE has a high affinity for adsorbing CR, so the concentration of CR decreases significantly at the first; over time, the desorption occurs making the concentration of CR in the solution increases again. But the oxidation of CR by H_2O_2 on the DE catalyst occurs favorably at pH 4-8, so the concentration of CR still decreases gradually with the reaction time.

3.2.2. Removal of CR under different conditions

The removal of CR in aqueous solution under different conditions is shown in Fig. 3. It can be seen that CR is quite stable under the effect of H_2O_2 (without catalyst), CR is only degraded about 7% after 90 minutes of reaction. In the presence of DE (without H_2O_2), CR is removed about 50% after 90 minutes of contact, proving that DE also has the ability to adsorb CR. Notably, when H_2O_2 and DE are present at the same time, CR is removed very quickly, the degradation efficiency reaches 85% after only 10 minutes of reaction and changes insignificantly during the remaining investigation time. Therefore, the degradation of CR in aqueous solution by H_2O_2 in the presence of DE continues to be studied.

$$D\% = -18.3156 + 3.5972 A + 3.0397 B - 0.9453 C - 0.0665 A^2 - 0.0500 B^2 + 0.0002 C^2 + 0.0250 AB + 0.0201 AC + 0.0004 BC \quad (4)$$

where A, B and C are the coded process variables.

The adequacy and significance of the proposed model were analyzed using analysis of

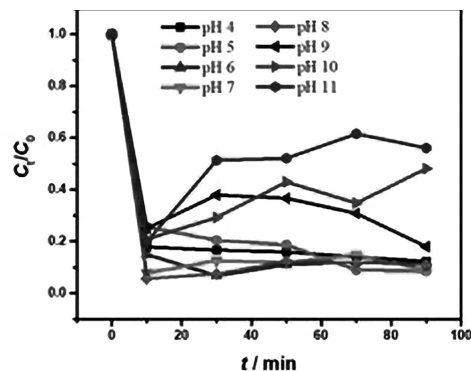


Figure 2. Decomposition of CR by H_2O_2 on DE catalyst at different initial solution pHs (Experimental conditions: DE 1.2 g/L, CR 30 mg/L, H_2O_2 6.5 g/L and ambient temperature).

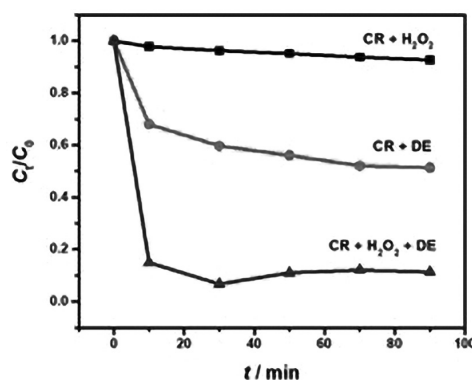


Figure 3. CR removal under different conditions (Experimental conditions: DE 1.2 g/L, CR 30 mg/L, pH 6, H_2O_2 6.5 g/L and ambient temperature).

3.2.3. Applying the Box-Behnken experimental design model

The results of the CR degradation performance determination of 15 experimental studies are given in Table 2. The experimental D% values were found to be in the range of 80.82 – 99.74%. A multiple regression equation was derived using the experimental variable data and the experimental D% values to express the relationship between D% and the process parameters according to the coded variables. Eq. (4) represents this relationship.

variance, the results of which are shown in Table 3. Statistically, a low p -value (< 0.05) indicates the significance of a model or parameter. Therefore, in the interpreting the results of

the variance test, it is known that independent variables with p -values less than 0.05 are effective in the obtained model. Whereas, p -values greater than 0.1 indicate that the terms are non-significant [7], [8]. From Table 3, it can be seen that the interacting factors of C^2 and BC is non-significant, the remaining terms are all significant. The F -value of the regression model is 4747.62 indicating that the model is significant at $p = 0.000$ (< 0.001). The F -value of Lack-of-Fit is 1.36 and the associated p -value is 0.450 which is non-significant due to the pure relative error [9]. Furthermore, there is a good agreement between the correlation coefficient ($R^2 = 0.9999$), the predicted R^2 (0.9987) and the adjusted R^2 (0.9997), between the theoretical and experimental values [7]. Fig. 4 shows the predicted response compared to the actual response as well as the normal plot of residuals. The fact that the residuals lie around a straight line indicates that the errors are distributed reasonably (Fig. 4a) [8]. The plotted points almost lying on the first quadrant bisector in Fig. 4b also indicate that the predicted $D\%$ value is close to the experimental result. These demonstrate the adequacy of the proposed model.

As can be seen from Eq. (4), the parameters that are significant to the model are those with larger coefficients, so with small changes in this value, the reaction will undergo large changes. Furthermore, with Eq. (4), the effect of the variable on the process can be inferred. The combined effects of decomposition time, reaction temperature and initial CR concentration on $D\%$ are depicted in Fig. 5. These plots can be understood as graphical representations of the regression model equation used to solve the optimal conditions of the factors, and are often applied to demonstrate an understanding of the types of interactions between variables that can be used to improve the decomposition performance [8]. The degradation efficiency of CR by H_2O_2 on DE catalyst with a target of about 99.0% was predicted by Minitab software (v. 16) for three process parameters found to be 40 (mg/L), 40 ($^{\circ}C$), and 14.2 (min) corresponding to initial CR concentration, temperature and reaction time, respectively. Under these conditions, the experimental results showed that the $D\%$ value reached 98.6%. The experimental results, under the same conditions, are close to the software prediction results, which confirm the reliability of the mode.

Table 2. Experimental and predicted responses for different experimental runs.

Run	Coded variables			Experimental variables			Response values ($D\%$)	
	A	B	C	Initial CR concentration (mg/L)	Reaction temperature ($^{\circ}C$)	Degradation time (min)	Exp.	Pred.
1	+1	0	+1	50	40	45	90.61	90.65
2	+1	0	-1	50	40	5	87.30	87.24
3	-1	0	+1	30	40	45	86.98	87.04
4	-1	0	-1	30	40	5	99.74	99.70
5	0	+1	+1	40	50	45	91.08	91.07
6	0	+1	-1	40	50	5	95.45	95.54
7	0	-1	+1	40	30	45	90.02	89.93
8	0	-1	-1	40	30	5	94.69	94.71
9	+1	+1	0	50	50	25	86.88	86.85
10	+1	-1	0	50	30	25	80.82	80.86
11	-1	+1	0	30	50	25	86.32	86.27
12	-1	-1	0	30	30	25	90.26	90.29
13	0	0	0	40	40	25	97.81	97.72
14	0	0	0	40	40	25	97.72	97.72
15	0	0	0	40	40	25	97.63	97.72

Table 3. Analysis of variance for *D%* of the CR degradation on DE catalyst.

Source	df	Seq SS	Adj SS	Adj MS	<i>F</i> -value	<i>p</i> -value	Remarks
Regression	9	414.745	414.745	46.083	4747.62	0.000	Significant
Linear	3	83.839	136.875	45.625	4700.45	0.000	Significant
A	1	39.167	59.225	59.225	6101.57	0.000	Significant
B	1	1.933	42.290	42.290	4356.90	0.000	Significant
C	1	42.739	36.406	36.406	3750.66	0.000	Significant
Square	3	241.342	241.342	80.447	8288.00	0.000	Significant
A ²	1	148.183	163.332	163.332	16827.08	0.000	Significant
B ²	1	93.130	92.326	92.326	9511.81	0.000	Significant
C ²	1	0.029	0.029	0.029	3.00	0.144	Non-significant
Interaction	3	89.563	89.563	29.854	3075.71	0.000	Significant
AB	1	25.013	25.013	25.013	2576.89	0.000	Significant
AC	1	64.527	64.527	64.527	6647.83	0.000	Significant
BC	1	0.023	0.023	0.023	2.42	0.181	Non-significant
Residual Error	5	0.049	0.049	0.010			
Lack-of-Fit	3	0.033	0.033	0.011	1.36	0.450	Non-significant
Pure Error	2	0.016	0.016	0.008			
Total	14	414.793					

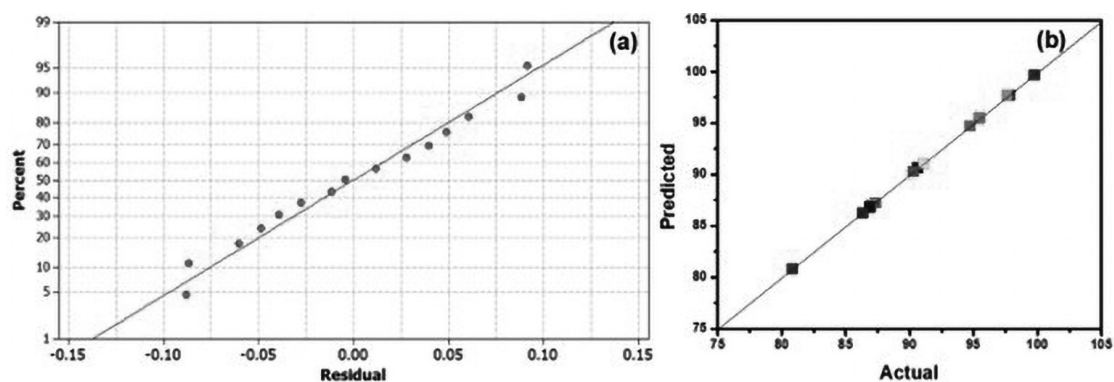


Figure 4. Normal plot of residuals (a) and predicted vs. actual values (b) for CR degradation.

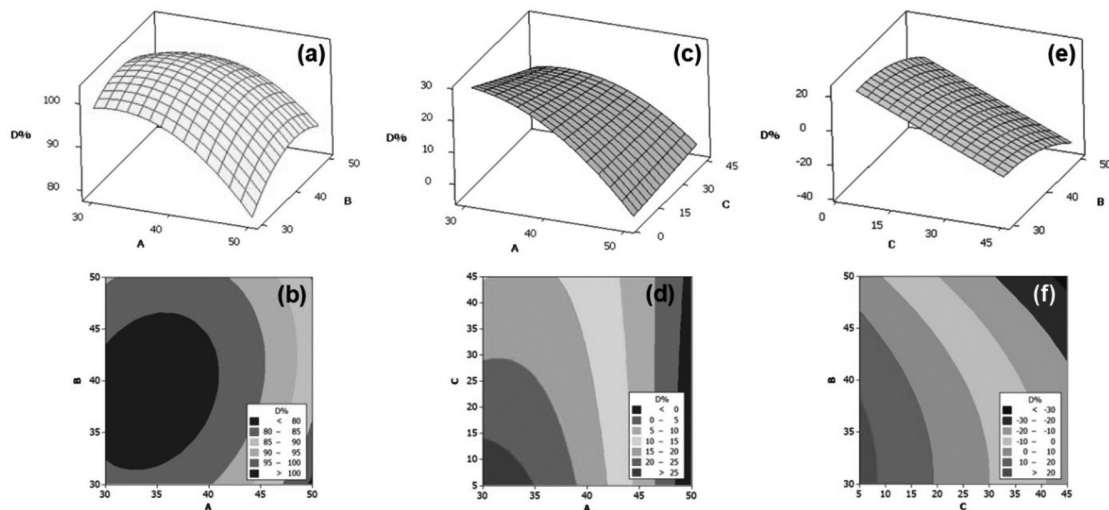


Figure 5. Combined effects of independent variables: (a, b) Initial CR concentration and reaction temperature; (c, d) Initial CR concentration and degradation time; (e, f) Degradation time and reaction temperature.

Table 4 compares the degradation efficiency of CR by H₂O₂ oxidizing agent in the presence of different catalysts. Although the reaction conditions are different, it can be seen that DE is also an effective catalyst for CR oxidation reaction in aqueous solution.

3.3. Recyclability of DE

Used DE catalyst is recovered, washed with distilled water and reused. The catalytic activity of DE was evaluated over four cycles of use, and the results of the CR degradation efficiency are shown in Fig. 6. It can be seen that the D% value decreased sharply after the first use, from 99% to 76% in the second use. After that, the D% only decreased slightly until the fourth use. This shows that DE still maintained its catalytic activity after four cycles of use, so this catalyst

can be used to degrade organic dyes in aqueous solution with hydrogen peroxide as an agent.

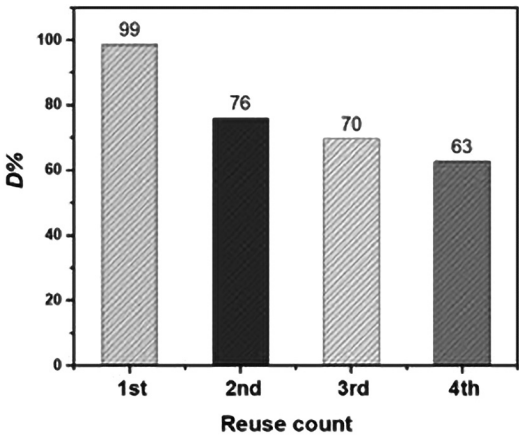


Figure 6. CR degradation efficiency after four times of reusing DE (Experimental conditions: DE 1.2 g/L, CR 40 mg/L, pH 6, H₂O₂ 6.5 g/L, 40 °C, and reaction time of 14.2 min).

Table 4. Degradation of CR using H₂O₂ as oxidizing agent on different catalysts.

Catalyst	Degradation efficiency (%)	Experimental conditions	Ref.
DE	98.6	pH 6, catalyst dosage = 1.2 g/L, [H ₂ O ₂] = 6.5 g/L, initial [CR] = 40 mg/L, temperature = 40 °C, time = 14.2 min	This work
Ag/ZnCo-ZIFs	94	pH 7.1, catalyst dosage = 150 mg/L, [H ₂ O ₂] = 0.05 mol/L, initial [CR] = 30 mg/L, temperature = 30 °C, time = 45 min	[12]
FeSO ₄ (*)	83	pH 3.5, catalyst dosage = 25 mg/L, [H ₂ O ₂] = 5 mM, initial [CR] = 25 mg/L, temperature = 30 °C, time = 60 min	[13]
Fe ²⁺	74.75	pH 3, catalyst dosage = 1.03 mM, [H ₂ O ₂] = 11.77 mM, initial [CR] = 50 mg/L, temperature = 25 °C, time = 120 min	[3]
Fe ₂ O ₃ @CeO ₂ -ZrO ₂ /Palygorskite	95	pH 3.4, catalyst dosage = 4 g/L, [H ₂ O ₂] = 40 mM, initial [CR] = 500 mg/L, temperature = 30 °C, time = 105 min	[14]

*The sono-Fenton process.

4. CONCLUSION

DE material was used as a catalyst in a heterogeneous Fenton oxidation system to decompose CR in aqueous solution. DE has a large specific surface area (55.4 m²/g) with the main component being amorphous silica, and some impurities such as Al (5.04%), Ti (0.56%), Fe (2.25%) (by weight). In the initial solution pH range from 4 to 8, the decomposition occurred favorable and the decomposition efficiency did not change significantly. At the initial solution pH of 6, the experimental results were effectively modeled using Box-Behnken factorial design

for three process variables affecting the CR decomposition efficiency, including initial CR concentration, temperature and reaction time. The model summary statistics showed that the developed model was adequate and accurate with the experimental data. The optimal operating conditions for the three process parameters were found to be 40 (mg/L), 40 (°C), and 14.2 (min) for the initial CR concentration, reaction temperature and decomposition time, respectively. Under these conditions, the CR decomposition efficiency was approximately 99% (at DE 1.2 g/L, pH 6, H₂O₂ 6.5 g/L).

DECLARATION OF CONFLICTS OF INTEREST

The author(s) declare that this research has no conflicts of interest.

REFERENCES

- [1] D. Dai, H. Liang, D. He, H. Potgieter, and M. Li, "Mn-doped Fe_2O_3 /diatomite granular composite as an efficient Fenton catalyst for rapid degradation of an organic dye in solution," *Journal of Sol-Gel Science and Technology*, vol. 97, no. 2, pp. 329-339, 2021.
- [2] W. Chen, M. Zeng, and J. Yang, "Preparation of Fenton Catalysts for Water Treatment," *Catalysts*, vol. 13, no. 11, 2023.
- [3] G. Afef and B. Mohamed, "Removal of Congo Red and Rhodamine B Dyes from Aqueous Solution by Fenton Process: Optimization of Operational Parameters," *Iranian Journal of Chemistry and Chemical Engineering*, vol. 42, no. 3, pp. 801-809, 2023.
- [4] P. O. Oladoye, M. O. Bamigboye, O. D. Ogunbiyi, and M. T. Akano, "Toxicity and decontamination strategies of Congo red dye," *Groundwater for Sustainable Development*, vol. 19, p. 100844, 2022.
- [5] J. Xu, X. Yun, M. Li, Y. Tian, X. Lei, and F. Zhang, "Iron-containing palygorskite clay as Fenton reagent for the catalytic degradation of phenol in water," *RSC Advances*, vol. 11, no. 47, pp. 29537-29542, 2021.
- [6] A. Przybek, M. Łach, M. H. Krupa, K. Miernik, and J. Mikuła, "Research of the physical and chemical properties of diatomite as a carrier of phase change materials for use in advanced building materials," *Frontiers in Materials*, vol. 11, 2024.
- [7] A. F. Sanromán, E. Rosales, M. Pazos, and A. Sanromán, "One-pot synthesis of bimetallic Fe-Cu metal-organic frameworks composite for the elimination of organic pollutants via peroxymonosulphate activation," *Environmental Science and Pollution Research*, vol. 32, no. 16, pp. 10592-10607, 2023.
- [8] M. Seloglu, R. Orhan, V. Selen, and G. Dursun, "Analysis of Photocatalytic Degradation of Phenol by Zinc Oxide Using Response Surface Methodology," *ChemistryOpen*, vol. 13, no. 6, 2024.
- [9] J. P. Maran, S. Manikandan, K. Thirugnanasambandham, C. V. Nivetha, and R. Dinesh, "Box-Behnken design based statistical modeling for ultrasound-assisted extraction of corn silk polysaccharide," *Carbohydrate Polymers*, vol. 92, no. 1, pp. 604-611, 2013.
- [10] Y. Liu, J. Zhang, X. Sheng, N. Li, and Q. Ping, "Adsorption and Release Kinetics, Equilibrium, and Thermodynamic Studies of Hymexazol onto Diatomite," *ACS Omega*, vol. 5, no. 45, pp. 29504-29512, 2020.
- [11] B. H. D. Son, V. Q. Mai, D. X. Du, N. H. Phong, N. D. Cuong, and D. Q. Khieu, "Catalytic wet peroxide oxidation of phenol solution over Fe-Mn binary oxides diatomite composite," *Journal of Porous Materials*, vol. 24, no. 3, pp. 601-611, 2017.
- [12] G. H. Dang, T. T. A. Le, A. K. Ta, T. N. T. Ho, T. V. Pham, T. V. H. Doan, T. H. V. Luong, "Removal of Congo red and malachite green from aqueous solution using heterogeneous Ag/ZnCo-ZIF catalyst in the presence of hydrogen peroxide," *Green Processing and Synthesis*, vol. 9, no. 1, pp. 567-577, 2020.
- [13] S. Nawaz, M. Siddique, R. Khan, N. Riaz, U. Waheed, I. Shahzadi, A. Ali, "Ultrasound-Assisted Hydrogen Peroxide and Iron Sulfate Mediated Fenton Process as an Efficient Advanced Oxidation Process for the Removal of Congo Red Dye," *Polish Journal of Environmental Studies*, vol. 31, no. 3, pp. 2749-2761, 2022.
- [14] J. Ouyang, Z. Zhao, S. L. Suib, and H. Yang, "Degradation of Congo Red dye by a Fe_2O_3 @ CeO_2 - ZrO_2 /Palygorskite composite catalyst: Synergetic effects of Fe_2O_3 ," *Journal of Colloid and Interface Science*, vol. 539, pp. 135-145, 2019.



© 2026 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Tối ưu hóa hiệu suất phân hủy congo red trên xúc tác đất diatomite sử dụng thiết kế Box-Behnken

Phạm Đình Dũ*

Trường Đại học Thủ Dầu Một, Thành phố Hồ Chí Minh, Việt Nam

Ngày nhận bài: 24/08/2025; Ngày sửa bài: 02/11/2025

Ngày nhận đăng: 19/11/2025; Ngày xuất bản: 28/06/2026

TÓM TẮT

Quá trình phân hủy congo red (CR) đã được khảo sát trong hệ phản ứng kiểu Fenton với sự hiện diện của đất diatomite (DE). Phương pháp bề mặt đáp ứng (RSM) dựa trên thiết kế Box-Behnken đã được ứng dụng để tối ưu hóa hiệu suất phân hủy. Trong những thiết kế thực nghiệm này, ba thông số quá trình ảnh hưởng đến hiệu suất phân hủy như nồng độ CR ban đầu (30-50 mg/L), nhiệt độ phản ứng (30-50 °C) và thời gian phân hủy (5-45 phút) đã được khảo sát. Dữ liệu thực nghiệm thu được được đưa vào phương trình đa thức bậc hai sử dụng phân tích hồi quy đa biến với giá trị hệ số tương quan cao ($R^2 = 0,9999$). Ở nồng độ CR ban đầu 40 mg/L, nhiệt độ phản ứng 40 °C, và thời gian phản ứng 14,2 phút, hiệu suất phân hủy đạt được 99% được dự đoán bởi mô hình (ở điều kiện: DE 1,2 g/L; pH 6; H_2O_2 6,5 g/L). Kết quả này cũng được xác nhận thông qua các thí nghiệm xác thực, chỉ ra rằng RSM là một trong những phương pháp hiệu quả để tối ưu hóa các điều kiện vận hành trong nghiên cứu này.

Từ khóa: Congo red, đất diatomite, phân hủy, thiết kế Box-Behnken.

*Tác giả liên hệ chính:

Email: dupd@tdmu.edu.vn

Gridding-based $\mathcal{H}_\infty$ observer design for the Lorenz 63 system using an NLPV reformulation

Khanh-Thoai Vo^{1,*}, Thanh-Phong Pham¹, Tien Dung Le²

¹University of Technology and Education (UTE), The University of Danang (UD), Viet Nam

²University of Science and Technology (DUT), The University of Danang (UD), Viet Nam

Received: 08/08/2025; Revised: 27/02/2026;

Accepted: 06/03/2026; Published: 28/06/2026

ABSTRACT

This paper presents a robust $\mathcal{H}_\infty$ observer design for the Lorenz 63 system based on a Nonlinear Parameter-Varying (NLPV) reformulation. The nonlinear dynamics are approximated by gridding the state space and constructing local linear models. At each grid point, an observer gain is synthesized by solving a linear matrix inequality (LMI), with a common Lyapunov function ensuring stability across the operating range. The observer gain is updated online via barycentric interpolation using the current estimated state. The approach enables real-time state estimation with guaranteed stability and disturbance attenuation. Simulation results under both noisy and noise-free conditions, and a comparison with an Extended Kalman Filter (EKF), confirm the effectiveness of the proposed design. Quantitative evaluations using Root Mean Square Error (RMSE), Normalized RMSE (NRMSE), and the coefficient of determination R^2 demonstrate high estimation accuracy and robustness of the observer across a range of dynamic behaviors in the Lorenz 63 system.

Keywords: Lorenz system, $\mathcal{H}_\infty$ observer, nonlinear parameter-varying (NLPV), LMI approach, gridding method.

1. INTRODUCTION

The Lorenz 63 system is a classical benchmark in nonlinear dynamics and chaos theory [1] - [3]. Due to its strong nonlinearities and sensitivity to initial conditions, it provides an ideal platform for testing observer design techniques. Designing observers for such systems is challenging, especially when dealing with unknown inputs, external disturbances, and nonlinearities [4] - [8].

Traditional methods like the EKF often rely on linearization and statistical assumptions, which may not be robust in chaotic regimes [9], [10].

An alternative is the $\mathcal{H}_\infty$ observer framework, which focuses on worst-case

disturbance attenuation [11] - [13]. However, the direct application of $\mathcal{H}_\infty$ methods to nonlinear systems is difficult due to their non-convexity.

To overcome this issue, the nonlinear system can be reformulated into an NLPV structure. The state space is discretized through gridding, and LMIs are employed to design observer gains at multiple linearization points. During online execution, the observer gain is interpolated in real-time based on the current estimated state using barycentric weights [14]. Although some existing works formulate the Lorenz system within a stochastic framework using Stochastic Differential Equation (SDE) [9], [15] and Itô calculus [16], this paper considers a purely deterministic Ordinary

*Corresponding author.

Email: vkthoai@ute.udn.vn

Differential Equation (ODE) model. It focuses on robust observer design without relying on statistical assumptions. Estimation performance is evaluated under both noisy and noise-free scenarios using standard quantitative metrics, including RMSE, NRMSE, and R^2 . Results confirm the robustness and effectiveness of the observer design across a wide range of operating conditions.

The main contributions of this paper are summarized as follows:

- Provide a rigorous deterministic NLPV reformulation for the Lorenz-63 dynamics (starting from ODE) and clarify noise modeling assumptions (process and measurement).
- A grid-based $\mathcal{H}_\infty$ observer synthesis procedure using LMIs and a common Lyapunov function.
- A real-time gain scheduling strategy using barycentric interpolation based on the observer state.
- Simulation-based performance evaluation under noisy and noise-free conditions using RMSE, NRMSE, and R^2 metrics.

The rest of this paper is organized as follows. Section II presents the NLPV modeling of the Lorenz system. Section III introduces the observer design approach via gridding and LMIs. Section IV provides numerical simulations and evaluation. Section V concludes the paper.

2. SYSTEM MODELING AND NLPV REFORMULATION

This section presents a deterministic representation of the Lorenz-63 system and its reformulation into an NLPV structure suitable for LMI-based observer design. All noises in this paper are treated as exogenous additive disturbances, not intrinsic stochastic diffusions. Therefore, the Lorenz model is described by an ODE, consistent with its original formulation in Lorenz in 1963.

2.1. Deterministic ODE form of the Lorenz System

The classical Lorenz-63 system is described by the nonlinear ordinary differential equation:

$$\dot{x}(t) = f(x(t)), x(t) \in \mathbb{R}^3, \quad (1)$$

where the nonlinear vector field is given by

$$f(x) = \begin{bmatrix} \sigma(x_2 - x_1) \\ x_1(\rho - x_3) - x_2 \\ x_1x_2 - \beta x_3 \end{bmatrix}, \quad (2)$$

Here, x denotes the system state, and the classical Lorenz parameters are $\sigma = 10$, $\rho = 28$, and $\beta = \frac{8}{3}$.

To account for modeling uncertainties and external perturbations in a deterministic framework, the system is augmented with an additive process disturbance:

$$\dot{x}(t) = f(x(t)) + B_w w(t), \quad (3)$$

$$y(t) = Cx(t) + D_v v(t),$$

where $w(t)$, $v(t)$ denote process and measurement disturbances, respectively. These signals are treated as bounded-energy exogenous inputs, consistent with the $\mathcal{H}_\infty$ estimation framework. Equation (3) serves as the starting point for building an NLPV representation of the nonlinear dynamics.

2.2. Linearization and Jacobian-Based Approximation

To reformulate the nonlinear ODE into an NLPV structure, we use the Jacobian of $f(x)$:

$$J(x) = \frac{\partial f}{\partial x}(x) = \begin{bmatrix} -\sigma & \sigma & 0 \\ \rho - x_3 & -1 & -x_1 \\ x_2 & x_1 & -\beta \end{bmatrix} \quad (4)$$

Given that the Lorenz system evolves within a bounded region $\mathcal{D} \subset \mathbb{R}^3$, we approximate the nonlinear function locally by:

$$f(x) = J(x^{[i]})x + f_{\text{rem}}^{[i]}(x), \quad (5)$$

with, $x^{[i]}$ is a selected grid point, $f_{\text{rem}}^{[i]}(x) = f(x) - J(x^{[i]})x$ is the remainder term.

2.3. Lipschitz property of the nonlinear remainder

For each grid region $\Omega_i \subset \mathcal{D}$, the remainder satisfies the following Lipschitz inequality:

$$\|f_{\text{rem}}^{[i]}(x_1) - f_{\text{rem}}^{[i]}(x_2)\| \leq L_f \|x_1 - x_2\| \quad (6)$$

$$\forall x_1, x_2 \in \Omega_i$$

A conservative global bound L_f can be computed as: $L_f = \sup_{x \in \mathcal{D}} \|J(x)\|_2$, (7)

where $\|\cdot\|_2$ is the spectral norm. The Lipschitz bound provides a structured way to handle the nonlinear remainder in the Lyapunov-based stability analysis without introducing stochastic assumptions.

2.4. NLPV representation via gridding

Let $\{x[1], \dots, x[N]\}$ be the chosen grid points in $\mathcal{D}$. At each grid point, define:

$$A_i = J(x^{[i]}). \quad (8)$$

Using barycentric interpolation, the state-dependent linearization is:

$$A(\rho(t)) = \sum_{i=1}^N \mu_i(\rho(t)) A_i, \quad (9)$$

with barycentric weights:

$$\mu_i(\rho(t)) \geq 0, \sum_{i=1}^N \mu_i(\rho(t)) = 1. \quad (10)$$

where $\mu_i(\rho(t))$ are interpolation weights depending on the scheduling parameter $\rho(t)$

The scheduling parameter is selected as:

$$\rho(t) = \hat{x}(t), \quad (11)$$

which ensures that all parameter-dependent quantities are available in real time without requiring access to the true state.

Combining the above results, the NLPV representation of the Lorenz–63 system is obtained as:

$$\begin{aligned} \dot{x}(t) &= A(\rho(t))x(t) + B_w w(t) + f_{\text{rem}}(x(t)) \\ y(t) &= Cx(t) + D_v v(t) \end{aligned} \quad (12)$$

where the nonlinear remainder $f_{\text{rem}}(x(t))$, satisfies the Lipschitz bound (6).

This formulation is now suitable for constructing an $\mathcal{H}_\infty$ observer via LMIs in Section 3.

Remark 1

The NLPV representation (11)–(12) is fully consistent with the original deterministic ODE (1)–(2) and does not introduce any stochastic interpretation or approximation of the system dynamics.

The gridding-based NLPV formulation admits a convex interpolation structure, which enables tractable LMI-based observer synthesis.

Retaining the nonlinear remainder explicitly avoids over-linearization and enables rigorous robustness analysis within the $\mathcal{H}_\infty$ framework.

3. OBSERVER DESIGN VIA GRIDDING

This section develops a robust state observer for the NLPV representation of the Lorenz–63 system introduced in Section 2. The proposed method relies on (i) formulating the estimation error dynamics, (ii) constructing an $\mathcal{H}_\infty$ -type dissipation inequality to attenuate disturbances and nonlinear uncertainties, and (iii) enforcing stability across all grid points through convex LMIs using a common Lyapunov matrix.

3.1. NLPV system with nonlinearity and disturbance

The NLPV model obtained in Section 2 is rewritten in compact form as:

$$\begin{aligned} \dot{x}(t) &= A(\rho(t))x(t) + B_w w(t) + f_{\text{rem}}(x(t)) \\ y(t) &= Cx(t) + D_v v(t) \end{aligned} \quad (13)$$

where $x(t) \in \mathbb{R}^n$ is the system state, $w(t) \in \mathbb{R}^{n_w}$ is the process noise, $v(t) \in \mathbb{R}^{n_v}$ is the measurement disturbances. The matrices $A(\rho)$ and B_w are obtained by barycentric interpolation of the Jacobians and disturbance matrices at the selected grid points. The matrix D_v models a scheduling-dependent measurement noise channel. The nonlinear remainder $f_{\text{rem}}(x)$ is globally Lipschitz on the compact domain and satisfies:

$$\begin{aligned} \|f_{\text{rem}}(x_1) - f_{\text{rem}}(x_2)\| &\leq L_f \|x_1 - x_2\|, \\ \forall x_1, x_2 \in \mathcal{D}, \mathcal{D} \subset \mathbb{R}^n \end{aligned} \quad (14)$$

This form isolates the linear parameter-varying component and the nonlinear uncertainty, providing a suitable structure for $\mathcal{H}_\infty$ observer synthesis.

3.2. NLPV observer structure

The structural form of the observer is presented as follows [17]:

$$\begin{aligned}\dot{\hat{x}}(t) &= A(\rho(t))\hat{x}(t) + f_{\text{rem}}(\hat{x}(t)) + L(\rho(t))(y(t) - \hat{y}(t)); \\ \hat{y}(t) &= C\hat{x}(t)\end{aligned}\quad (15)$$

Define the estimation error as:

$$e(t) = x(t) - \hat{x}(t).$$

Substituting (13) – (15) yields the estimation error dynamics:

$$\begin{aligned}\dot{e}(t) &= (A(\rho(t)) - L(\rho(t))C)e(t) + \Delta f(t) \\ &\quad + B_w w(t) - L(\rho(t))D_v v(t)\end{aligned}\quad (16)$$

where $\Delta f(t) = f_{\text{rem}}(x(t)) - f_{\text{rem}}(\hat{x}(t))$ and satisfies $\|\Delta f(t)\| \leq L_f \|e(t)\|$. (17)

Equation (16) explicitly separates the effects of process disturbance $w(t)$ and measurement noise $v(t)$, which facilitates the subsequent $\mathcal{H}_\infty$ analysis and LMI-based observer synthesis.

3.3. LMI condition for $\mathcal{H}_\infty$ observer

The following theorem provides sufficient conditions for exponential stability $\mathcal{H}_\infty$ and performance of the proposed observer.

To ensure robustness against $w(t)$, $v(t)$ and nonlinear uncertainties, a quadratic Lyapunov function $V(e) = e^T P e$ is considered, with $P > 0$. The objective is to ensure: Exponential stability with decay rate $\beta > 0$, and $\mathcal{H}_\infty$ disturbance attenuation with level $\gamma > 0$. These requirements are enforced through the following dissipation inequality:

$$\dot{V}(e) + 2\beta P + e^T e - \gamma^2 w^T w - \gamma^2 v^T v \leq 0 \quad (18)$$

Calculate the derivative:

$$\begin{aligned}\dot{V}(e) &= e^T (A - LC)^T P e + e^T P (A - LC) e \\ &\quad + 2e^T P B_w w - 2e^T P L D_v v + 2e^T P \Delta f\end{aligned}\quad (19)$$

Use Cauchy–Schwarz + Young [18, 19]:

$$2e^T P \Delta f \leq 2 \|Pe\| \|\Delta f\| \leq 2 \|P\| \|e\| \|\Delta f\| \quad (20)$$

For ease, use Young with the parameter: $\epsilon_f > 0$

$$2e^T P \Delta f \leq \epsilon_f \Delta f^T \Delta f + \frac{1}{\epsilon_f} \quad (21)$$

Due $\|\Delta f\|^2 \leq L_f^2 \|e\|^2$, we have:

$$\epsilon_f \Delta f^T \Delta f \leq \epsilon_f L_f^2 e^T e. \quad (22)$$

If we choose ϵ_f and (or) normalize such that

$$\frac{1}{\epsilon_f} P^2 \leq L_f^2 P, \text{ then infer:}$$

$$2e^T P \Delta f \leq L_f^2 e^T P e. \quad (23)$$

Substitute (23) into (19):

$$\begin{aligned}\dot{V}(e) &\leq e^T ((A - LC)^T P + P(A - LC) + L_f^2 P) e \\ &\quad + 2e^T P B_w w - 2e^T P L D_v v\end{aligned}\quad (24)$$

Introduce the standard change of variables:

$$L(\rho) = -P^{-1}Y(\rho) \quad (25)$$

which yields:

$$\begin{aligned}(A - LC)^T P + P(A - LC) &= \\ A^T P + PA + C^T Y^T + YC\end{aligned}\quad (26)$$

At each grid vertex i , define:

$$\Xi_i = A_i P + P A_i^T + C^T Y_i^T + Y_i C \quad (27)$$

Substituting (24), (26)–(27) into (18) gives

$$\begin{aligned}\dot{V} + 2\beta P + e^T e - \gamma^2 w^T w - \gamma^2 v^T v \\ \leq e^T (\Xi_i + 2\beta P + L_f^2 P) e + 2e^T P B_{w,i} w + \\ 2e^T Y_i D_{v,i} v + e^T e - \gamma^2 w^T w - \gamma^2 v^T v < 0\end{aligned}$$

Rearranging terms yields the quadratic form:

$$\begin{bmatrix} e \\ w \\ v \end{bmatrix}^T \begin{bmatrix} \Xi_i + 2\beta P + L_f^2 P & P B_{w,i} & Y_i D_{v,i} \\ B_{w,i}^T P & -\gamma^2 I & 0 \\ D_{v,i}^T Y^T & 0 & -\gamma^2 I \end{bmatrix} \begin{bmatrix} e \\ w \\ v \end{bmatrix} < 0 \quad (28)$$

Inequality (28) is equivalent to the LMI condition

$$M = \begin{bmatrix} \Xi_i + 2\beta P + L_f^2 P & P B_{w,i} & Y_i D_{v,i} \\ B_{w,i}^T P & -\gamma^2 I & 0 \\ D_{v,i}^T Y^T & 0 & -\gamma^2 I \end{bmatrix} < 0 \quad (29)$$

which must hold at each grid point $i = 1, \dots, N$.

Since the NLPV model is a convex interpolation of its vertices and a common

Lyapunov matrix P is used, the feasibility of (19) at all vertices guarantees the dissipation inequality (18) for all admissible scheduling trajectories $\rho(t)$.

+ *Disturbance-free case* ($w = 0, v = 0$)

$$\dot{V} + 2\beta V < 0 \Rightarrow \|e(t)\| \leq ce^{-\beta t} \|e(0)\|,$$

implying exponential stability with decay rate β .

+ *General case*: integrating (18) from 0 to ∞ yields:

$$\int_0^\infty \|e(t)\|^2 dt \leq \gamma^2 \int_0^\infty (\|w(t)\|^2 + \|v(t)\|^2) dt,$$

establishing the desired $\mathcal{H}_\infty$ performance bound.

3.4. Gridding and barycentric interpolation approach

To effectively design an observer for nonlinear parameter-varying systems, the state-dependent matrices $A(\rho(t))$ and $L(\rho(t))$ must be approximated. Direct continuous-time synthesis is generally intractable due to infinite-dimensional dependency on the scheduling parameter $\rho(t)$. Therefore, a gridding approach is adopted to discretize the state space.

The state space $X \subset x(t) \in \mathbb{R}^n$ is partitioned into a finite number of grid points $\{x^{[i]}\}_{i=1}^N$, where a local LMI observer synthesis is performed. At each grid point $x^{[i]}$, an observer gains $L^{[i]}$ is computed by solving the corresponding LMI condition. These gains are stored for online use.

To enable smooth gain variation and avoid chattering between discrete observers, the gains are interpolated during runtime using barycentric weights. Let $\hat{x}(t)$ denote the current observer state. The interpolated gain $L(\hat{x}(t))$ is calculated as:

$$L(\hat{x}(t)) = \sum_{i=1}^N \mu_i(\hat{x}(t)) L^{[i]}, \quad (30)$$

where $\mu_i(\cdot)$ are barycentric interpolation weights satisfying $\sum_{i=1}^N \mu_i(\hat{x}(t)) = 1$, and $\mu_i \geq 0$.

4. NUMERICAL SIMULATION

This section evaluates the performance of the proposed gridding-based $\mathcal{H}_\infty$ NLPV observer on the Lorenz–63 system. The objective is to demonstrate robust state estimation under both process and measurement disturbances and to highlight the advantages of the proposed approach over conventional observers.

4.1. System setup

The Lorenz 63 system is a well-known nonlinear chaotic system governed by the following differential equations:

$$\begin{aligned} \dot{x}(t) &= B_w w(t) + f(x(t)), \\ y(t) &= Cx(t) + D_w w(t) \end{aligned} \quad (31)$$

where $x(t) = [x_1, x_2, x_3]^T \in \mathbb{R}^3$ is the system state, $w(t) \in \mathbb{R}^3$ denotes external disturbances, and the measurement output is $y(t) \in \mathbb{R}$. The matrices are defined as:

$$y(t) = Cx(t), C = [1 \ 0 \ 0]. \quad (32)$$

The nonlinear vector field $f(x)$ is given by [1]:

$$f(x(t)) = \begin{bmatrix} \sigma(x_2 - x_1) \\ x_1(\rho - x_3) - x_2 \\ x_1 x_2 - \beta x_3 \end{bmatrix} \quad (33)$$

with parameters $\sigma = 10$, $\rho = 28$, $\beta = \frac{8}{3}$. To facilitate observer design using convex optimization tools, we reformulate the nonlinear system into an NLPV structure by approximating the dynamics through local linearizations.

The nonlinear vector field $f(x)$ is linearized around multiple grid points $\{x^{(i)}\}_{i=1}^N$ within a bounded region $\mathcal{D} \subset \mathbb{R}^3$. At each point $x^{(i)}$, the Jacobian matrix is computed as:

$$A^{(i)} := J(x^{(i)}) = \frac{\partial f}{\partial x} \Big|_{x^{(i)}} \quad (34)$$

This results in a set of locally linearized models:

$$\dot{x}(t) \approx A^{(i)} x(t) + B_w w(t) + f_{\text{rem}}(x) \quad (35)$$

where $f_{\text{rem}}(x) := f(x) - A^{(i)}x$ is the residual nonlinearity. Assuming that $f(x)$ is Lipschitz continuous over $\mathcal{D}$, the residual satisfies:

$$\|f_{\text{rem}}(x) - f_{\text{rem}}(\hat{x})\| \leq L_f \|x - \hat{x}\|, \forall x, \hat{x} \in \mathcal{D} \quad (36)$$

for some constant $L_f > 0$. To express the system in NLPV form, we introduce a parameter trajectory $\rho(t) = \hat{x}(t)$, leading to:

$$\begin{aligned}\dot{x}(t) &= A(\rho(t))x(t) + B_w w(t) + f_{\text{rem}}(x) \\ y(t) &= Cx(t) + D_v v(t)\end{aligned}\quad (37)$$

The matrix $A(\rho(t))$ is obtained via online interpolation of $\{A^{(i)}\}$ using barycentric weights:

$$A(\rho(t)) = \sum_{i=1}^N \mu_i(\rho(t)) A^{(i)}, \quad (38)$$

where $\sum_i \mu_i = 1, \mu_i \geq 0$

This interpolation ensures a smooth and accurate approximation of the nonlinear dynamics across the grid.

Remark 2

The scheduling parameter is chosen as $\rho(t) = \hat{x}(t)$ to ensure real-time implementability of the observer. This avoids dependence on unmeasurable states while maintaining consistency between the NLPV model and the observer dynamics. As the estimation error converges, the scheduling variable naturally approaches the true state, thereby reducing interpolation mismatch.

4.2. Lipschitz constant estimation

The observer gains are designed at grid points uniformly sampled over $[-20, 20]^3$. At each grid point, the local Jacobian is computed and used to define $A(\rho_i)$.

The nonlinear drift term $f(x)$ of the Lorenz 63 system is defined in equation (27). To facilitate observer design with Lipschitz-type nonlinearities, the Lipschitz constant L_f is required. A vector field $f(x)$ is said to be Lipschitz continuous over the domain $\mathcal{D} \subset \mathbb{R}^n$ if there exists a scalar $L_f > 0$ which satisfies the equation (36).

A sufficient condition to obtain L_f is to evaluate the spectral norm of the Jacobian matrix $J(x)$:

$$L_f = \sup_{x \in \mathcal{D}} \|J(x)\|_2 \quad (39)$$

where $J(x) = \frac{\partial f(x)}{\partial x}$ is the Jacobian matrix of the drift function.

For the Lorenz system, the Jacobian is computed as:

$$J(x) = \nabla f(x) = \begin{bmatrix} -\sigma & \sigma & 0 \\ \rho - x_3 & -1 & -x_1 \\ x_2 & x_1 & -\beta \end{bmatrix} \quad (40)$$

The spectral norm $\|J(x)\|_2$ is the largest singular value of $J(x)$, which can be numerically evaluated over a bounded region $\mathcal{D}$. In this work, the domain is chosen as $\mathcal{D} = [-20, 20]^3$, covering the typical range of Lorenz state trajectories.

A grid-based scan of $J(x)$ across $\mathcal{D}$ yields an upper bound: $L_f \approx 56.6092$.

Figure 1 describes the distribution of $\|J(x)\|_2$ over a bounded domain in (x_1, x_2, x_3) . It is observed that the spectral norm varies significantly, reaching values above $L_f = 56.6092$ in lower regions of the state space (i.e., $x_3 < 0$), while staying below 25 in upper regions (i.e., $x_3 > 10$). This spatial variability reflects the strong local nonlinearity of the Lorenz system, which motivates the use of a gridding-based observer design. By selecting local Lipschitz bounds within each grid cell, the observer gain can be adapted more accurately to the system's local dynamics, avoiding conservatism associated with a global Lipschitz constant.

This bound is used in the LMI formulation to handle the nonlinear remainder term using Lipschitz inequalities, ensuring robust estimation even in the presence of nonlinear uncertainties.

Using the CVX toolbox and semidefinite programming (SDP) (available at <http://cvxr.com/cvx>), a set of LMI conditions is solved to compute the observer gains $\{L_i\}$. During online implementation, the gain $L(\hat{x})$ is obtained via barycentric interpolation, where the weights are determined based on the proximity to the grid centers.

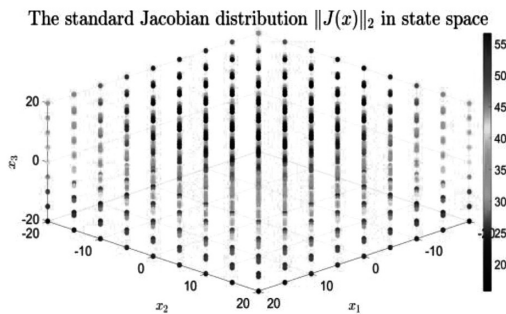


Figure 1. A graph representation of the Jacobian $\|J(x)\|_2$ in the state space.

Specifically, for each grid point $i \in \{1, 2, \dots, 4\}$ of the NLPV system in equation (24), the LMI condition in equation (29) is solved to synthesize a robust $\mathcal{H}_\infty$ observer.

The CVX toolbox in MATLAB is employed to solve the optimization problem and compute the observer gain matrices L_i corresponding to each value of the scheduling parameter ρ , using the system matrices defined in section 3.4.

The optimization yields a disturbance attenuation level of $\gamma = 0.0023$. The resulting observer gain $L(\hat{x})$ of the four grid points of ρ are as follows:

$$L(\hat{x}) = [519.9294; -17.1069; 6.3601]^T$$

4.3. Simulation scenarios and discussion

To evaluate and compare the performance of the proposed $\mathcal{H}_\infty$ observer and the EKF [9], numerical simulations are carried out on the Lorenz 63 system. The initial state of the system is set as $x(0) = [-5, -5, -5]^T$, and both observers are initialized at the origin $[0, 0, 0]^T$. The EKF is implemented using a first-order prediction-correction structure, where the time-varying Jacobian matrix of the Lorenz system is used in the prediction step. The initial covariance matrix is selected as $P_0 = 2 \cdot I_3$ to ensure sufficient initial uncertainty for the EKF. In contrast, the $\mathcal{H}_\infty$ observer uses a gridding structure interpolated via barycentric weights over a predefined grid of the state space.

To provide a comprehensive performance analysis, two sets of simulation conditions are examined:

Noise-Free Scenario: The system evolves without any disturbances to establish a baseline.

Figure 2 illustrates the comparison between the actual states x_1 to x_3 and their corresponding estimated values $\hat{x}_1$ to $\hat{x}_3$. In the plots, the solid green line represents the true system states, the red dashed line indicates the estimates from the proposed $\mathcal{H}_\infty$ observer, while the blue dash-dot line corresponds to the estimates obtained using the EKF.

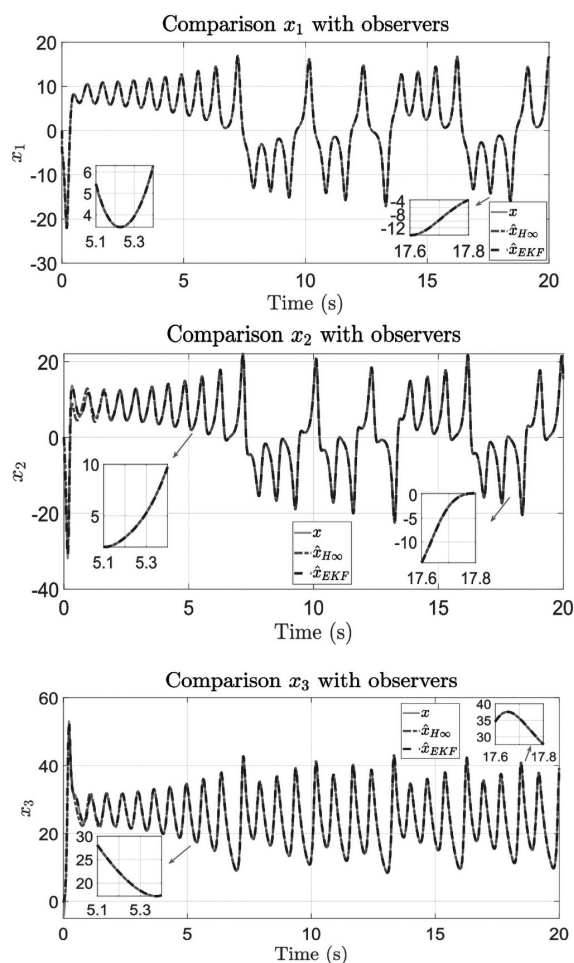


Figure 2. Noise-free state estimation performance of the Lorenz 63 system.

Comparison between the true states (solid green), the proposed gridding-based NLPV $\mathcal{H}_\infty$ observer (red dashed), and the Extended Kalman Filter (blue dash-dot). The insets show zoomed-in views over the intervals $t \in [5.1, 5.3]$ s and $t \in [17.6, 17.8]$ s, highlighting transient convergence behavior and local tracking accuracy under strong nonlinear oscillations.

It is evident from the Figure 2 plots that both $\mathcal{H}_\infty$ observer and the EKF closely track the true system states. The estimation errors are negligible for all state variables, reflecting excellent accuracy in the absence of noise. Notably:

In x_1 , both observers almost overlap with the ground truth across the entire time span, with only very slight divergence in highly dynamic segments. The inset zoomed plots confirm sub-millisecond response agreement.

In x_2 , the estimations remain aligned even during sharp transient oscillations. This highlights the observers' ability to capture rapid nonlinear dynamics.

In x_3 , where chaotic oscillations dominate, both $\hat{x}_{\mathcal{H}_\infty}$ and $\hat{x}_{EKF}$ accurately replicate the system evolution.

In the noise-free scenario, both the proposed gridding-based NLPV $\mathcal{H}_\infty$ observer and the EKF provide accurate state estimation, indicating that both approaches are capable of tracking the Lorenz dynamics under ideal measurement conditions.

Nevertheless, a closer inspection of the zoomed-in sub-figures reveals subtle but consistent differences between the two observers. In particular, the proposed observer exhibits a slightly reduced phase lag during transient convergence, as well as smoother tracking behavior with fewer oscillatory artifacts. These differences become more visible in time intervals characterized by rapid state variations, where local linearization effects may limit the EKF performance.

Although the overall estimation accuracy of both observers is comparable in this noise-free case, the observed improvements indicate that the proposed gridding-based NLPV formulation provides a more consistent representation of the underlying nonlinear dynamics, even in the absence of disturbances.

Noisy Scenario: Both process and measurement noise are activated.

The state dynamics are subjected to an additive zero-mean Gaussian process noise $w(t) \sim \mathcal{N}(0, Q)$ and measurement noise $v(t) \sim \mathcal{N}(0, R)$, where the covariance matrices are chosen as follows:

$$Q = 0.5 \cdot I_3, R = I_1 \quad (41)$$

The process noise $w(t)$ is generated at each time step as:

$$w_k = \sqrt{Q} \cdot \omega_k, \omega_k \sim \mathcal{N}(0, I_3) \quad (42)$$

while the measurement noise is:

$$v_k = \sqrt{R} \cdot v_k, v_k \sim \mathcal{N}(0, I_1) \quad (43)$$

Figure 3 compares the actual states x_1 to x_3 and their corresponding estimated values $\hat{x}_1$ to $\hat{x}_3$, in the presence of noise as described above. The figures demonstrate that both observers track the system well, but the performance diverges during fast transients and in regions of strong nonlinear coupling.

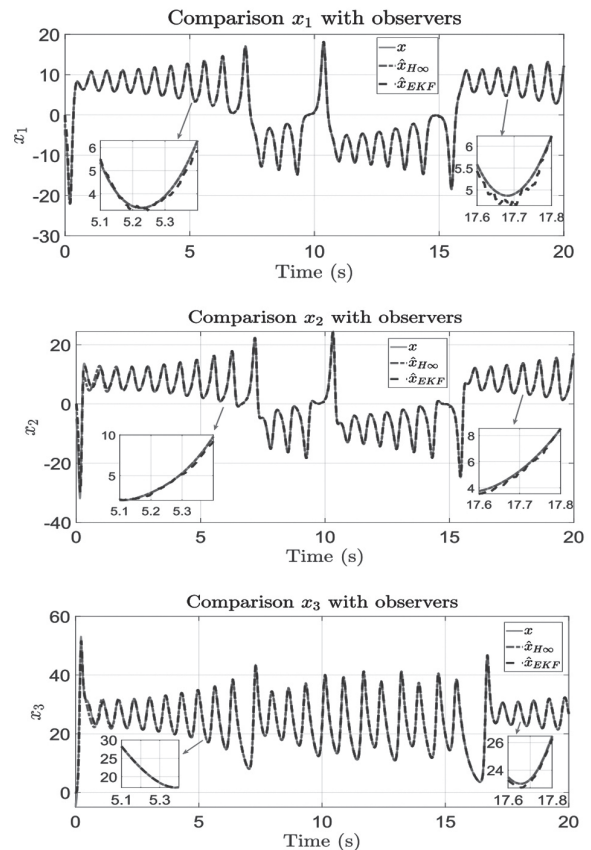


Figure 3. State estimation performance under process and measurement noise.

The system is subjected to Gaussian disturbances as defined in (41)–(43). The plots compare true states (solid green), the proposed $\mathcal{H}_\infty$ observer (red dashed), and the EKF (blue dash-dot). Zoomed-in segments illustrate transient robustness and phase behavior during rapid state transitions in chaotic regions.

Estimation of x_1 : During the early transient phase around $t \in [5.1, 5.3]$ s (left zoom-in), the trajectories of the true state and both observers start to converge after a rapid change in system dynamics. The zoomed view clearly shows that the proposed $\mathcal{H}_\infty$ observer tracks the true state with smaller phase lag and reduced oscillation amplitude compared to the EKF. This behavior indicates a faster convergence rate, which is consistent with the explicit decay-rate constraint incorporated in the LMI-based observer design. In the highly nonlinear region around $t \in [17.6, 17.8]$ s (right zoom-in), the Lorenz system exhibits sharp peaks and rapid variations. In this interval, the EKF estimation shows noticeable deviations from the true state, whereas the proposed $\mathcal{H}_\infty$ observer maintains closer alignment. This demonstrates the improved robustness of the proposed approach against nonlinear amplification effects and disturbances in regions where local linearization becomes less accurate.

For the second state x_2 , the zoomed-in interval $t \in [5.1, 5.3]$ s highlights the convergence behavior during a steep state transition. The proposed observer exhibits a smoother transient response with reduced overshoot compared to the EKF. This suggests that the proposed method effectively balances convergence speed and robustness, avoiding aggressive corrections that can amplify noise.

In the later interval $t \in [17.6, 17.8]$ s, where x_2 undergoes rapid sign changes and strong nonlinear coupling, the EKF trajectory deviates more significantly from the true state. In contrast, the proposed $\mathcal{H}_\infty$ observer remains tightly bounded around the true trajectory. This confirms the advantage of the gridding-based

NLPV formulation, which captures dominant variations of the Jacobian over the attractor domain and provides robustness beyond local linearization.

Estimation of x_3 : The estimation results for x_3 further emphasize the robustness properties of the proposed observer. In the zoomed-in window around $t \in [5.1, 5.3]$ s, both observers converge toward the true state; however, the proposed $\mathcal{H}_\infty$ observer shows smaller steady-state oscillations once convergence is achieved.

More importantly, in the interval $t \in [17.6, 17.8]$ s, where x_3 reaches higher amplitudes and the effect of nonlinear coupling is pronounced, the EKF exhibits increased estimation error. The proposed observer, on the other hand, maintains accurate tracking with reduced sensitivity to noise. This behavior highlights the effectiveness of explicitly accounting for nonlinear remainder bounds and disturbance attenuation in the observer design.

Tables 1–3 present the state estimation performance indices, including RMSE, NRMSE, and the R^2 - for both the EKF and the proposed $\mathcal{H}_\infty$ observer, under noise-free and noisy conditions, respectively.

- Root Mean Square Error (RMSE):

$$\text{RMSE}_i = \sqrt{\frac{1}{N} \sum_{k=1}^N (x_i(k) - \hat{x}_i(k))^2} \quad (38)$$

- Normalized RMSE (NRMSE):

$$\text{NRMSE}_i = \frac{\text{RMSE}_i}{\max(x_i) - \min(x_i)} \quad (39)$$

- Coefficient of Determination (R^2):

$$R_i^2 = 1 - \frac{\sum (x_i - \hat{x}_i)^2}{\sum (x_i - \bar{x}_i)^2} \quad (40)$$

Table 1. RMSE index of states.

RMSE	With noise-free		With Noise	
	EKF	$\mathcal{H}_\infty$	EKF	$\mathcal{H}_\infty$
x_1	0.04042	0.00197	0.25921	0.00313
x_2	0.06177	0.15654	0.33442	0.16664
x_3	0.07490	0.16884	0.27328	0.17989

Table 2. NRMSE index of states.

N RMSE	With noise-free		With Noise	
	EKF	$\mathcal{H}_\infty$	EKF	$\mathcal{H}_\infty$
x_1	0.00114	5.62e-05	0.00757	9.14e-05
x_2	0.00131	0.00336	0.00743	0.00370
x_3	0.00197	0.00455	0.00679	0.00447

Table 3. R^2 index of states.

R^2	With noise-free		With Noise	
	EKF	$\mathcal{H}_\infty$	EKF	$\mathcal{H}_\infty$
x_1	0.99996	1.00000	0.99896	0.99999
x_2	0.99995	0.99968	0.99857	0.99964
x_3	0.99991	0.99955	0.99872	0.99944

The numerical values reported in Tables 1–3 are computed directly from the full simulated state trajectories over the specified time intervals, whereas Figures 2 and 3 provide a qualitative visualization of the corresponding estimation performance.

The simulation results demonstrate a comprehensive comparison between the proposed $\mathcal{H}_\infty$ observer and the EKF under both noise-free and noisy conditions. As presented in Tables 1–3, the $\mathcal{H}_\infty$. The observer exhibits superior robustness, especially in the presence of process and measurement noise.

In the noise-free scenario, EKF achieves slightly better RMSE values for x_2 and x_3 , while the $\mathcal{H}_\infty$ observer delivers the best accuracy for x_1 , achieving an RMSE of only 0.00197 and a coefficient of determination $R^2 = 1.0000$. However, in the noisy case, the performance of EKF significantly degrades across all states. For instance, the RMSE for x_1 increases to 0.25921 under EKF, whereas the proposed observer maintains a remarkably low RMSE of 0.00313.

The NRMSE analysis further supports these findings, with the $\mathcal{H}_\infty$ observer consistently achieves lower normalized errors under noisy conditions. Specifically, for x_1 , the NRMSE of the $\mathcal{H}_\infty$ observer remains as low as 9.14×10^{-5} , compared to 7.57×10^{-3} for EKF.

In terms of the coefficient of determination, the $\mathcal{H}_\infty$ observer consistently attains higher R^2 values in both scenarios, indicating a better match between estimated and actual states. Notably, the observer preserves an R^2 of over 0.9999 for all states, even in the presence of noise, whereas EKF drops to 0.9985 or lower.

In general, these results validate the robustness and estimation accuracy of the proposed $\mathcal{H}_\infty$ observer design. The gridding-based LMI synthesis, combined with barycentric interpolation of the observer gain, enables the observer to maintain high precision under strong nonlinearities and measurement uncertainties. In contrast, EKF performance is more sensitive to noise and model mismatch, highlighting the conservative yet effective design philosophy of the $\mathcal{H}_\infty$ approach.

5. CONCLUSION

This paper has developed a gridding-based NLPV $\mathcal{H}_\infty$ observer for the Lorenz–63 system, addressing the challenges posed by strong nonlinear coupling and chaotic dynamics. By reformulating the original nonlinear system into an NLPV structure with an explicitly bounded nonlinear remainder, the proposed approach enables a convex LMI-based observer synthesis with guaranteed stability properties.

The main results demonstrate that the proposed observer ensures exponential convergence of the estimation error with a prescribed decay rate and achieves effective attenuation of both process disturbances and measurement noise. The use of a small number of representative grid points allows a practical trade-off between robustness and computational efficiency while maintaining the feasibility of a common Lyapunov function.

Simulation results confirm that, in the noise-free case, the proposed observer provides comparable accuracy under ideal conditions and improved robustness in the presence of disturbances compared to the EKF. Under

disturbed conditions, the proposed method exhibits superior robustness, with reduced sensitivity to noise and improved estimation accuracy across all state variables.

Overall, the results validate the effectiveness of the proposed gridding-based NLPV $\mathcal{H}_\infty$ observer as a reliable and robust estimation framework for nonlinear chaotic systems. Future work will extend the proposed framework to higher-dimensional chaotic systems and more complex nonlinear dynamics, including hyperchaotic models. Another promising direction is experimental validation on physical nonlinear platforms that can be reformulated in NLPV form, such as robotic manipulators and electromechanical systems.

In addition, adaptive gridding strategies and data-driven gain scheduling techniques, including machine learning-based interpolation schemes, may be investigated to further reduce conservatism while preserving stability guarantees within the LMI framework.

DECLARATION OF CONFLICTS OF INTEREST

The authors declare that this research has no conflicts of interest.

REFERENCES

- [1] E. N. Lorenz, "Deterministic nonperiodic flow," *Journal of the Atmospheric Sciences* 20, vol. 20, no. 2, pp. 130-141, 1963.
- [2] M. Goodliff, S. Fletcher, A. Kliever, J. Forsythe, and A. Jones, "Detection of non-Gaussian behavior using machine learning techniques: a case study on the Lorenz 63 model," *Journal of Geophysical Research: Atmospheres*, vol. 125, no. 2, pp. 1-15, 2020.
- [3] P. Tandeo et al., "Combining analog method and ensemble data assimilation: application to the Lorenz-63 chaotic system," in *Proceedings of the 4th International Workshop on Climate Informatics*, 2015, pp. 3-12.
- [4] F. Zhu, Y. Fu, T. N. Dinh, "Asymptotic convergence unknown input observer design via interval observer," *Automatica*, vol. 147, p. 110744, 2023.
- [5] C. M. Nguyen, A. T. Nguyen, and S. Delprat, "Cascade Takagi-Sugeno Fuzzy Observer Design for Nonlinear Uncertain Systems with Unknown Inputs: A Sliding Mode Approach," *International Journal of Robust and Nonlinear Control*, vol. 33, no. 15, pp. 9066-9083, 2023.
- [6] T. P. Pham, O. Sename, L. Dugard, "A nonlinear parameter varying observer for real-time damper force estimation of an automotive electro-rheological suspension system," *International Journal of Robust and Nonlinear Control*, vol. 31, no. 17, pp. 8183-8205, 2021.
- [7] S. Meng, F. Meng, F. Zhang, Q. Li, Y. Zhang, and A. Zemouche, "Observer design method for nonlinear generalized systems with nonlinear algebraic constraints with applications," *Automatica*, vol. 162, p. 111512, 2024.
- [8] S. Mohite, M. Alma, A. Zemouche, "Toward enhancing nonlinear observers for Lipschitz system: Exploiting the matrix multipliers-based LMIs," *International Journal of Robust and Nonlinear Control*, vol. 34, no. 15, pp. 10799-10820, 2024.
- [9] C. O. Ongere, D. Angwenyi, R. Oryiema, "Second-order Extended Ensemble Kalman Filter with Stochastically Perturbed Innovation," *Science Mundi*, vol. 4, no. 2, pp. 222-242, 2024.
- [10] M. A. Hazazi, and A. Sihabuddin, "Extended Kalman filter in recurrent neural network: USDIDR forecasting case study," *Indonesian Journal of Computing and Cybernetics Systems*, vol. 13, no. 3, pp. 293-300, 2019.
- [11] S. Mohite, M. Alma, A. Zemouche, and M. Haddad, "LMI-based H_∞ observer design for nonlinear Lipschitz system," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 6745-6750, 2023.
- [12] E. Chnib, P. Bagnerini, and A. Zemouche, "LMI-based H_∞ Observer Design for a Quadcopter Model Operating in an Adaptive Vertical Farm," *IFAC-PapersOnLine*, vol. 56, no. 2, pp. 10837-10842, 2023.

- [13] K. T. Vo, T. D. Le, and T. P. Pham, "Design of an H_∞ observer for nonlinear parameter-varying systems with Lipschitz nonlinearity and model uncertainty," *UD-JST*, vol. 23, no. 9B, pp. 34-39, 2025.
- [14] J. P. Berrut, and L. N. Trefethen, "Barycentric Lagrange Interpolation," *Society for Industrial and Applied Mathematics*, vol. 46, no. 3, pp. 501-517, 2004.
- [15] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York and London: Academic Press, 1970.
- [16] K. Ito, "Stochastic Integral," *Proceedings of the Imperial Academy*, vol. 20, no. 8, pp. 519-524, 1944.
- [17] D. G. Luenberger, "Observing the state of a linear system," *IEEE Transactions on military electronics*, vol. 8, no. 2, pp. 74-80, 1964.
- [18] H. K. Khalil, *Nonlinear Systems*, 3rd Ed. New Jersey: Prentice Hall, 2002.
- [19] S. Boyd and L. El Ghaoui, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA, USA: SIAM, 1997.



© 2026 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Thiết kế bộ quan sát $\mathcal{H}_\infty$ theo phương pháp chia lưới cho hệ Lorenz 63 sử dụng mô hình dạng phi tuyến với tham số thay đổi

Võ Khánh Thoại^{1,*}, Phạm Thanh Phong¹, Lê Tiến Dũng²

¹Trường Đại học Sư phạm Kỹ thuật, Đại học Đà Nẵng, Việt Nam

²Trường Đại học Bách khoa, Đại học Đà Nẵng, Việt Nam

Ngày nhận bài: 08/08/2025; Ngày sửa bài: 27/02/2026;

Ngày nhận đăng: 06/03/2026; Ngày xuất bản: 28/06/2026

TÓM TẮT

Bài báo này đề xuất một phương pháp thiết kế bộ quan sát $\mathcal{H}_\infty$ bền vững cho hệ Lorenz 63 thông qua mô hình hóa lại theo dạng hệ phi tuyến với tham số thay đổi (NLPV). Bằng cách sử dụng phương pháp chia lưới theo không gian trạng thái và xây dựng các mô hình tuyến tính cục bộ, chúng tôi thiết lập một tập hợp các bất đẳng thức ma trận tuyến tính (LMI) để thiết kế bộ quan sát tại mỗi điểm lưới. Các ma trận bộ quan sát thu được sẽ được nội suy theo thời gian thực dựa trên trạng thái của bộ quan sát. Phương pháp này cho phép ước lượng chính xác các trạng thái của hệ trong cả hai trường hợp có nhiễu và không nhiễu. Kết quả mô phỏng và so sánh với bộ lọc Kalman mở rộng (EKF) xác nhận hiệu quả của phương pháp đề xuất thông qua các chỉ số đánh giá Sai số bình phương trung bình (RMSE), Chuẩn hóa RMSE (NRMSE) và hệ số xác định R^2 .

Từ khóa: Hệ thống Lorenz, bộ quan sát $\mathcal{H}_\infty$, hệ phi tuyến với tham số thay đổi (NLPV), tiếp cận LMI, phương pháp chia lưới.

*Tác giả liên hệ chính.

Email: vkthoai@ute.udn.vn

Practical embedded AI on FPGA: A compact CNN achieving high throughput and low power for digit recognition

Pham Van Khoa*, La Cong Loc, Nguyen Duc Trong, Huynh Minh Quy,
Ho Gia Huyen, Vo Quang Huy, Pham Ba Loc, Huynh Thi Thu Hien

Ho Chi Minh City University of Technology and Engineering, Vietnam

Received: 18/10/2025; Revised: 05/01/2026;

Accepted: 23/01/2026; Published: 28/06/2026

ABSTRACT

This work presents the design and deployment of a Convolutional Neural Network on an SoC-FPGA platform for handwritten digit classification using the MNIST dataset. The goal is a compact, efficient FPGA-based CNN accelerator with fewer than 1,000 parameters that integrates seamlessly with the ARM processor on the PYNQ-Z2 board via DMA and AXI interfaces. The accelerator is realized at the register-transfer level and undergoes simulation, synthesis, and resource-focused optimization while preserving inference accuracy. On 10,000 MNIST test images, the system attains 91.28% accuracy - about 5 percentage points below a CPU implementation on dual ARM Cortex-A9 cores (96.26%) - but delivers a 7–8× speedup and a 36% reduction in power consumption. The design highlights effective parallelization and pipelining of convolution operations directly on the FPGA, achieving low resource usage and power draw. These results provide a practical foundation for real-time embedded AI applications - such as character recognition, image monitoring, intelligent IoT systems, and edge computing - on SoC-FPGA platforms.

Keywords: *Convolutional Neural Network Accelerator, FPGA, System on Chip, MNIST, Image Classification.*

1. INTRODUCTION

In recent years, Convolutional Neural Networks (CNNs) have become the dominant approach for image recognition and classification owing to their efficient spatial feature extraction and high accuracy [1]. However, CNN models typically require substantial computation and memory, which makes deployment challenging on embedded systems with limited hardware resources [2]. Field-Programmable Gate Arrays (FPGAs)-with their flexibility, massive parallelism, and low power consumption-have proven to be effective platforms for

accelerating CNNs in embedded applications [1], [3]. Implementing CNNs on FPGAs can reduce inference latency relative to CPU- or GPU-based software while efficiently utilizing hardware resources such as DSP slices, Lookup Tables (LUTs), and block RAM (BRAM). To achieve high performance on FPGAs, many studies quantize weights and activations to replace floating-point operations with integer arithmetic, thereby reducing hardware complexity while maintaining accuracy [2]. In addition, techniques such as parallelization, pipelining, and data reuse are commonly applied

**Corresponding author.*

Email: khoapv@hcmute.edu.vn

to increase throughput and optimize memory bandwidth [3]. In SoC–FPGA architectures, the integration of FPGA Programmable Logic (PL) and the embedded ARM Processing System (PS) provides a balance between performance and flexibility [4]. The PL handles computationally intensive kernels, while the PS manages control and data movement via direct memory access (DMA) over the Advanced eXtensible Interface (AXI) interconnect [5], [6]. On the PYNQ-Z2 platform, the Python Productivity for Zynq (PYNQ) framework enables direct control and testing of the CNN accelerator from Python, facilitating data transfer between Double Data Rate (DDR) memory and the FPGA and thereby simplifying system development and evaluation [6], [7]. This work presents the design and implementation of a lightweight CNN accelerator on an SoC–FPGA platform for handwritten-digit classification using the Modified National Institute of Standards and Technology (MNIST) dataset. The model is optimized to fewer than 1,000 parameters to balance accuracy, memory footprint, and hardware feasibility on the PYNQ-Z2. The system employs DMA/AXI for data exchange between the CPU and FPGA and integrates the PYNQ framework for control and real-time inference. Experimental results demonstrate that the proposed design delivers high performance, efficient resource utilization, and low power consumption, confirming the feasibility of FPGA-based real-time embedded AI systems.

2. RELATED WORK

MNIST inference on FPGA/SoC platforms has been extensively explored along two main directions: (i) extreme-quantization approaches (binary/ternary) that push hardware efficiency and throughput, and (ii) fixed-/integer-point CNN deployments (LeNet-5-based CNN) that target higher accuracy with moderate resource cost. In the first direction, FINN provides a representative toolflow for binarized neural

networks on FPGAs and reports multi-million classifications per second on MNIST with sub-microsecond latency on a Zynq ZC706 platform, together with explicit performance and power reporting [8]. In the fixed-/integer-point CNN category, González et al. implement a LeNet-5 inference accelerator using an SW/HW co-processing scheme on a Zynq-7000 Arty Z7-20, reporting 97.59% MNIST accuracy with 12-bit fixed-point arithmetic and approximately 441 images/s at 100 MHz [9]. More recent full-network deployments using HLS on Zynq devices provide more complete reporting of model scale and system metrics; for example, Liang et al. detail LeNet-5 weight counts and report end-to-end inference time, power, and error rate (e.g., 1.07 ms, 2.193 W, and 0.99% error for their PIPELINE design) [10]. Complementary to binary networks, ternary models have also been proposed to improve efficiency while retaining more representational capacity than binarization; Alemdar et al. present ternary neural networks and discuss FPGA/ASIC realizations that leverage low-precision arithmetic and sparsity effects for energy-efficient inference [11]. Beyond CNNs, compact fixed-point DNN implementations that keep weights entirely in on-chip memory have also been demonstrated to reduce external memory traffic; Park and Sung describe an FPGA-based DNN design using 3-bit weights with on-chip storage (no external DRAM access) and report <5 W full-speed power in their MNIST evaluation [12].

3. DESIGN METHODOLOGY

As shown in Figure 1, the design process comprises four stages: first, we specify the MNIST classification task, target performance and accuracy, hardware resource constraints, and the PS–PL communication scheme within the Zynq SoC to guide subsequent decisions. Next, we design, train, and quantize the CNN in PyTorch, convert the quantized weights to 8-bit integers (INT8), and export them as HEX files

for the hardware stage. We then implement the CNN functional blocks at the Register-Transfer Level (RTL), perform simulation and synthesis to evaluate resource usage (LUTs, DSPs, BRAM), and integrate the accelerator into the Zynq SoC via the AXI4-Stream interface to enable high-throughput PS–PL data movement. Finally, we deploy the generated bitstream on the PYNQ-Z2, control execution through the PYNQ framework on the ARM Cortex-A9, and evaluate the system using 10,000 MNIST test images to measure performance, accuracy, and hardware resource utilization.

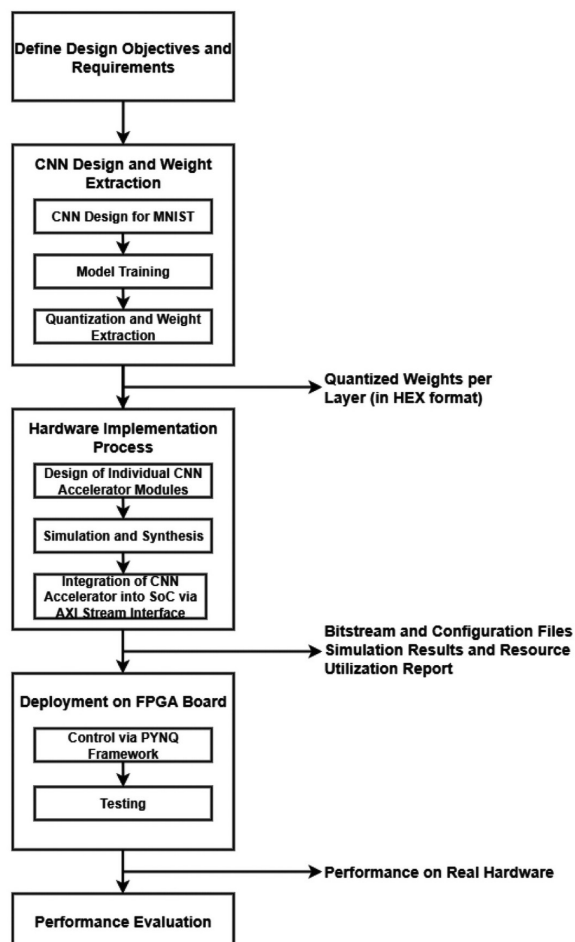


Figure 1. Design implementation flow.

3.1. CNNs for Hand-written digit classification

Selecting an appropriate CNN model is pivotal to the overall system design because it directly influences accuracy, processing latency, and hardware resource utilization on the FPGA.

On an SoC–FPGA platform constrained by DSP slices, LUTs, and BRAM, the model must balance computational complexity with hardware feasibility: an excessive parameter count can exceed on-chip storage, increase DDR access latency, and impede pipelining, whereas an overly simplified network may weaken feature extraction and reduce accuracy. Accordingly, the research team aims to develop a compact, efficient CNN architecture that enables high-throughput inference in an RTL implementation. As shown in Figure 2, the designed CNN model comprises two convolutional layers, two pooling layers with Rectified Linear Unit (ReLU) activation, and a single Fully Connected layer. The architecture follows the LeNet-5 paradigm but is simplified for FPGA deployment. A 5×5 kernel is employed to balance feature extraction quality with streamlined, pipelined Multiply–Accumulate (MAC) operations on DSP units. After the two convolution–pooling stages, the output is flattened into a 48×1 feature vector and passed to a Fully Connected layer that produces class scores over ten outputs corresponding to digits 0–9.

As shown in Figure 2, the proposed CNN contains a small number of parameters across all layers, reflecting its lightweight design for FPGA deployment. The CNN adopts a minimalist architecture with an optimized dataflow tailored to the FPGA’s bandwidth and buffering constraints. The pooling layers progressively reduce the spatial dimensions of the feature maps, which facilitates deep pipelining and lowers the computational load of subsequent layers. With a total of 796 parameters, the model attains approximately 96% accuracy in single-precision (float32), providing a robust foundation for quantization and hardware implementation.

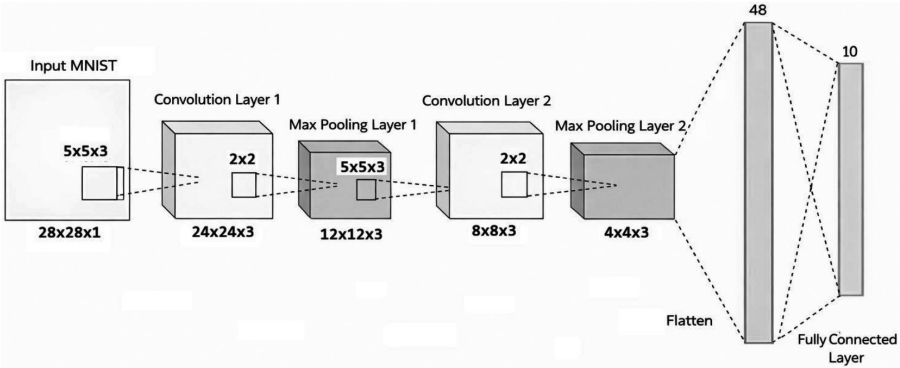


Figure 2. CNN model architecture for MNIST classification.

To reduce hardware cost and accelerate computation, we apply quantization-aware training (QAT) to convert the model to int8, reducing memory usage by approximately 4× while preserving accuracy close to the floating-point baseline. Quantized weights and biases are exported per layer to enable direct inference on the FPGA. The quantization pipeline consists of three steps: (1) normalize weights and biases to the range $[-1, 1]$; (2) scale by 128 to map values to the int8 range $[-128, 127]$; and (3) encode negative values in two’s-complement

form for FPGA storage. The resulting quantized parameters are written as .mem files (one per layer) and loaded directly into BRAM or register files within the RTL design. This workflow yields a CNN optimized for both accuracy and hardware deployability and is ready for accelerator construction and on-FPGA inference. As shown in Figure 3, the quantization process follows a structured three-step pipeline that ensures numerical consistency between software simulation and hardware implementation.

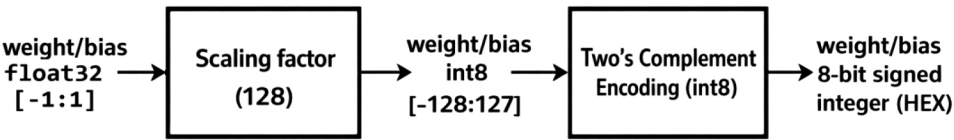


Figure 3. Quantization process of the CNN model.

3.2. Implementation of CNN Accelerator Core on FPGA

After completing the CNN model, the next step is the design of the RTL module. A CNN architecture can be implemented using various approaches, including Naive Convolution, Matrix Multiplication, or Winograd Convolution. In this work, the basic Naive Convolution method

is adopted to construct the CNN hardware architecture.

Figure 4 illustrates the overall system architecture, in which the Buffer, Conv Calc, Max Pooling, ReLU, Fully Connected, and Comparator modules are independently designed and then integrated into a complete CNN block.

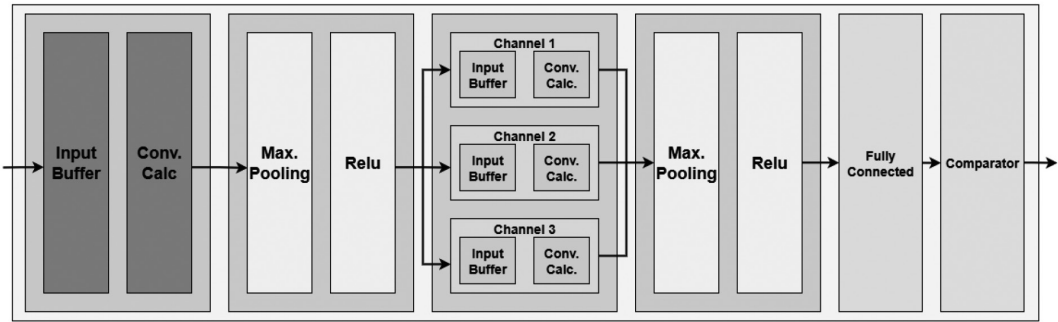


Figure 4. Block diagram of the CNN accelerator.

The Input Buffer Block stores incoming image pixels. The accelerator ingests a 28×28 MNIST bitmap (784 pixels), with each pixel represented in 8 bits. Pixels arrive as a stream—one pixel per clock cycle—in raster-scan order from the top-left corner, proceeding left-to-right and top-to-bottom. This serial loading scheme avoids allocating on-chip memory for the entire image and allows computation to begin immediately, without waiting for full-frame capture. The line buffer holds 140 entries of 8 bits each (corresponding to $5 \text{ rows} \times 28$

columns). After a row is processed, the buffer is overwritten with the next row until the entire image is consumed. For convolution, 5×5 pixel windows are extracted from the buffer and shifted by one pixel horizontally at each cycle, repeating until the end of the row. In every clock cycle, one 5×5 window is emitted and forwarded to the convolution stage. With a 5×5 , stride-1, valid convolution on a 28×28 input, the first layer produces $24 \times 24 = 576$ windows (each containing 25 pixels), matching the output feature-map dimensions of the layer.

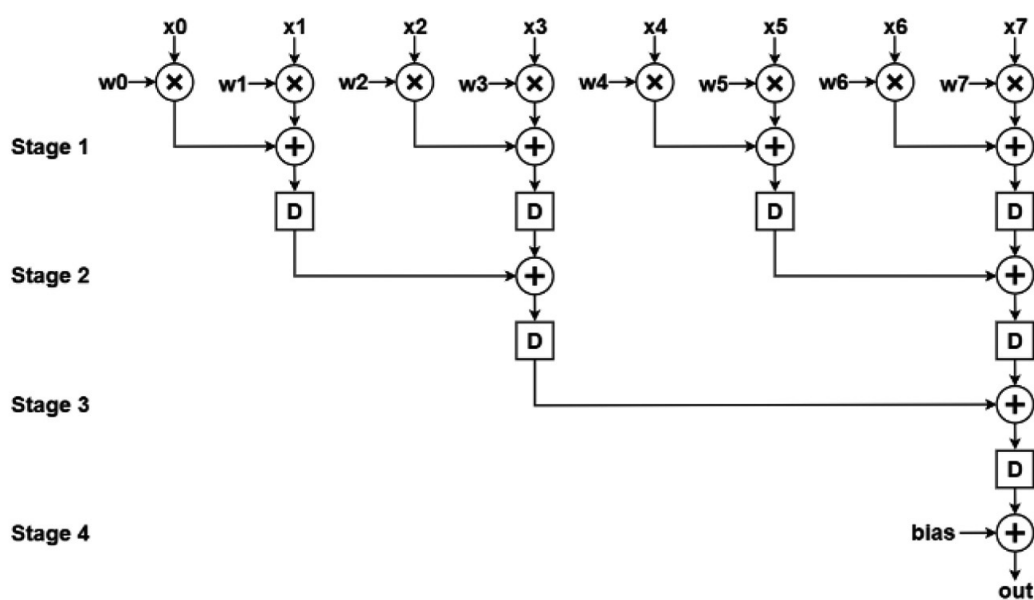


Figure 5. Pipeline stages of the convolution module.

The Convolution Calculation Module performs the convolution between the input data stream from the buffer block and the 5×5 kernel weights. Its input is a stream of 576 windows, each containing 25 parallel pixels, supplied by the buffer. At each clock cycle, one 5×5 window is consumed to compute a dot product with the 5×5 kernel, followed by addition of the bias term. Because arithmetic operations on the FPGA incur propagation delay, the datapath is pipelined to sustain high throughput. Figure 5 illustrates the pipelined structure, realized by inserting registers to partition the computation into multiple stages, thereby shortening the critical path and increasing the achievable

clock frequency. With a four-stage pipeline, the first valid output appears four cycles after the corresponding input window is received; thereafter, the module produces one output per cycle.

For Max-Pooling and ReLU Modules, the first convolution layer yields a 24×24 feature map, emitted as 576 sequential values in a continuous stream, which serves as the input to the Max Pooling and ReLU modules. The 2×2 Max Pooling unit processes pixels in pairs of rows (two from the first row and two from the second row), outputs the maximum among the four, and advances by one pooling stride. The result is then passed to the ReLU activation,

which preserves non-negative values and sets negative values to zero. A line buffer stores 12 elements—one output row of the resulting 12×12 feature map from the Max Pooling–ReLU stage. For each 2×2 cell, the running maximum is updated in the buffer when the current value exceeds the stored value; otherwise, the stored value is retained. The buffered value is then compared with zero to apply ReLU. Pointers and control flags step through the buffer in sync with the 576-pixel input stream, producing a 12×12 feature map with 144 outputs.

The Fully Connected and Comparator Modules receives input from the second convolution and Max Pooling layers. The 4×4 feature maps with three channels are flattened into a 48×1 vector, multiplied by the corresponding weights, and summed with bias terms to produce ten output neurons. Arithmetic operations in the Fully Connected module are pipelined similarly to the convolution module to optimize performance. After computing the ten neuron outputs, the Comparator identifies the neuron with the highest value (argmax). A ten-element line buffer temporarily stores the neuron values from the Fully Connected module, and the index of the maximum value is emitted as the predicted class (digits 0–9).

3.3. Integration of CNN Accelerator Core into Zynq SoC

After completing the CNN accelerator hardware core, we integrated it into the Zynq SoC so that the ARM processing system (PS) can drive the programmable-logic (PL) inference engine through a standard memory-to-stream pipeline. Input images are stored in off-chip DDR, and a Xilinx IP block (AXI DMA) bridges the DDR-based memory subsystem and the streaming accelerator, as depicted in Figure 6. To make the accelerator compatible with the SoC interconnect, we wrapped the RTL core with a lightweight top-level module that presents a compliant AXI4-Stream slave interface for input

pixels and an AXI4-Stream master interface for output results. Since the CNN core natively consumes data as a stream, the AXI-Stream handshake signals (e.g., TVALID/TREADY and frame delimiting) map naturally onto the existing streaming datapath and ensure reliable backpressure handling. The end-to-end dataflow proceeds as follows. (1) The PS configures the AXI DMA through the S_AXI_LITE control port, programming the base DDR addresses and transfer lengths for the input image buffer and the output result buffer. (2) For the MM2S (memory-mapped-to-stream) direction, the DMA fetches pixel data from DDR via M_AXI_MM2S and packetizes it into an AXI4-Stream on M_AXIS_MM2S. (3) The outgoing stream is optionally decoupled using axis_data_fifo_0, which absorbs burstiness from DDR reads and provides elastic buffering so the accelerator can run smoothly even if memory traffic momentarily stalls. (4) The buffered stream is then consumed by the CNN accelerator (axis_cnn_mnist_0), which performs inference in the PL and emits classification outputs as an AXI4-Stream. (5) On the output side, axis_data_fifo_1 buffers the accelerator's result stream and handles any backpressure from the downstream DMA. (6) For the S2MM (stream-to-memory-mapped) direction, the DMA receives the output stream on S_AXIS_S2MM and writes the results back to DDR via M_AXI_S2MM. (7) Finally, the PS reads the result buffer from DDR and uses it for reporting (e.g., predicted digit) or for subsequent software-side processing. Overall, this PS–PL integration turns DDR-resident images into a continuous AXI stream for the accelerator and returns inference outputs back to DDR using the same standardized AXI infrastructure. The two FIFOs isolate the accelerator from memory timing variability, while the AXI-Lite control path lets software orchestrate transfers without modifying the RTL datapath, yielding a reusable and scalable integration pattern for streaming CNN inference on Zynq.

instances required. Finally, where beneficial, fixed inference weights enable replacing some multipliers with LUT-based constant-coefficient arithmetic (e.g., shift-add or distributed arithmetic) and/or applying structured pruning with zero-skipping to reduce the effective MAC count. Collectively, these modifications can significantly lower DSP usage while preserving the AXI-streaming integration and maintaining sufficient throughput for MNIST inference.

Figure 7(a) presents the time-resolved (instantaneous) PYNQ-Z2 board power during single-image inference, clearly capturing the transient transition from the idle baseline to the active computation phase. Figure 7(b)–(d) then compares the corresponding average power across three operating modes: (b) idle (baseline), (c) inference executed on the ARM CPU, and (d) inference executed on the PL-based SoC accelerator.

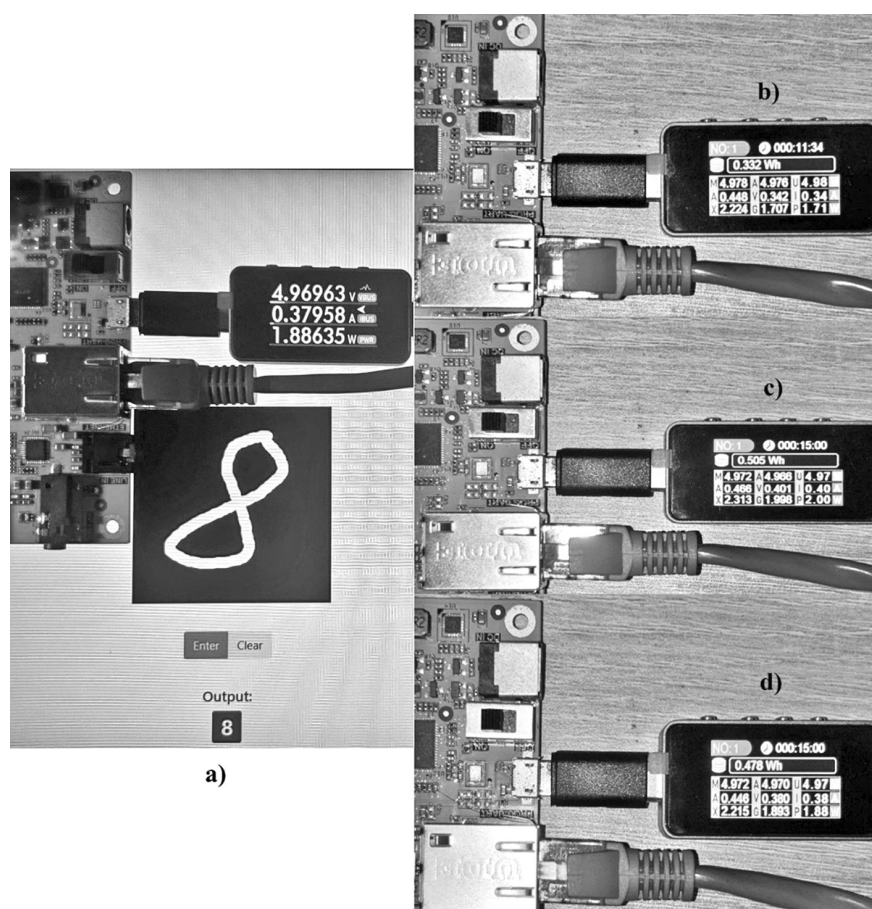


Figure 7. Measurement of instantaneous and average power consumption during CNN inference on the PYNQ-Z2 platform. (a) Instantaneous power profile for a single input image, illustrating the transition from idle to active operation. (b)–(d) Comparison of average power consumption under idle conditions, CPU-based inference, and PL/SoC accelerator-based inference.

Table 2 presents the summarized results—including classification accuracy, frame rate, and average power. The proposed CNN accelerator achieves a classification accuracy of 91.28%, which is approximately 5% lower than the CPU implementation (96.26%). However, processing latency is significantly reduced—from 4.25 ms on the CPU to 0.54 ms on the FPGA—representing

a 7–8 $\times$ speedup, despite the FPGA operating at a much lower frequency (100 MHz vs. 650 MHz). This improvement highlights the benefits of parallel computation and deep pipelining inherent in FPGA-based architectures. In terms of power efficiency, the FPGA implementation consumes 186 mW on average, compared with 291 mW for the CPU, resulting in an overall

36% reduction in power consumption. The corresponding energy per inference decreases from 1.234 mJ per image on the CPU to 0.102 mJ per image on the FPGA, demonstrating a substantial improvement in energy efficiency. Overall, these results indicate that the proposed

CNN accelerator offers a favorable balance between performance and energy consumption, making it well suited for real-time embedded AI applications on resource-constrained edge devices.

Table 2. Experimental results and performance comparison.

Hardware Platform	Latency (ms)	Accuracy (%)	FPS	Power (mW)	Efficiency (mJ/frame)
FPGA (100 MHz)	0.54	91.28	1852	186	0.102
CPU ARM Cortex A9 (650 MHz)	4.25	96.26	235	291	1.234

Table 3 shows that prior FPGA MNIST accelerators typically emphasize either higher accuracy using larger LeNet-5-class models [9], [10] or extremely high throughput via aggressive quantization such as binary or ternary networks [8], [11], [12]. In contrast, our work targets a different operating point by implementing a very small int8 CNN (2 conv + 1 FC) with only 796 parameters directly in RTL and integrating it on the low-cost PYNQ-Z2 (XC7Z020) using an AXI-DMA streaming flow; on 10,000 test images it achieves 91.28% accuracy, 0.54 ms/image latency (1,852 FPS), and 186 mW incremental power (0.102 mJ/image). Compared with earlier studies, the primary strength is model compactness, since the parameter count is far smaller than the larger

LeNet-5 implementations and the 3-layer FC network with substantial parameter storage [8]-[10]. However, the compact model yields lower accuracy than all other works reporting accuracy ($\approx 95.83\%$ – $\approx 99.01\%$) [8]-[11]. In throughput, our design exceeds slower FPGA results such as 441 FPS and 934.6 FPS [9], [10] but remains well below the ultra-high-throughput designs reporting 61,035 – 70,000 FPS or even multi-mega FPS enabled by extreme quantization and heavy parallelism [8], [11], [12]. Finally, while our absolute power is low compared to watt-level implementations [9], [10], the energy per image is not best-in-class because 0.102 mJ/image ($\sim 102 \mu\text{J}/\text{image}$) is higher than the best reported $\mu\text{J}/\text{image}$ figures [11], [12].

Table 3. Comparison of FPGA-based MNIST inference accelerators.

Work	Precision	Platform	Model	Accuracy	Throughput	Power
This work	int8	PYNQ-Z2 (Zynq XC7Z020)	Compact CNN (2 conv. + 1 FC.) 796 params	91.28%	1,852 FPS 0.54 ms/img	186 mW, 0.102 mJ/img
Umuroglu et al. [8]	1-bit	ZC706 (Zynq Z7045)	3-layer FC, 256 neurons/layer; 0.3 Mbits params	95.83%	12.361 M-FPS, 0.31 $\mu\text{s}/\text{img}$	Pchip 7.3 W, Pwall 21.2 W
González et al. [9]	fixed-point (12-bit)	Arty Z7-20 (Zynq-7000)	LeNet-5 (params N/R)	97.59%	441 FPS 2.27 ms/img	1.72 W
Liang et al. [10]	Integer / fixed-point variants	Zynq platform	LeNet-5	$\approx 99.01\%$	934.6 FPS 1.07 ms/img	2.19 W
Alemdar et al. [11]	Ternary NN	Kintex-7 (XC7K160T)	not reported	98.14%	61,035 FPS, 8.09 $\mu\text{s}/\text{img}$	3.63 $\mu\text{J}/\text{img}$
Park et al. [12]	3-bit weights, 8-bit signals	ZC706 (Zynq Z7045)	not reported	not reported	70,000 FPS	71 $\mu\text{J}/\text{img}$

CONCLUSION

This work delivers a deployable, end-to-end SoC-FPGA inference pipeline that goes beyond demonstrating MNIST classification by integrating model design with a practical embedded hardware realization. We (i) design a minimalist int8 CNN (2 conv + pooling/ReLU + 1 FC) with only 796 parameters and export FPGA-ready quantized weights, (ii) implement a fully RTL streaming accelerator that processes input pixels at one pixel per clock using lightweight line buffering and a pipelined datapath, and (iii) integrate the core on the low-cost PYNQ-Z2 (XC7Z020) via AXI4-Stream and AXI-DMA for complete PS-PL deployment. On 10,000 MNIST test images, the PL implementation achieves 91.28% accuracy-about five percentage points below the dual ARM Cortex-A9 CPU baseline-while providing a 7–8× throughput improvement (0.54 ms/image, 1,852 FPS vs. 4.25 ms/image, 235 FPS) and lower incremental power, resulting in markedly better energy per inference. These results highlight the advantages of FPGA-based CNN acceleration for deterministic low-latency and energy-efficient edge inference on resource-constrained platforms, and the presented workflow offers a reproducible template spanning quantization, RTL development, pipelining, resource-aware design, and PYNQ-based system integration. Although accuracy is limited by the intentionally compact model and int8 quantization, and the XC7Z020 imposes tight DSP/BRAM constraints, the proposed system establishes a solid foundation for future improvements such as layer folding and time-multiplexed MACs to reduce DSP usage, refined quantization-aware training to recover accuracy, and scaling to larger FPGAs or modestly larger networks while preserving the same streaming integration methodology. Overall, the paper demonstrates a practical and extensible approach to embedded FPGA inference suitable for real-time edge applications including digit/character recognition and low-power intelligent sensing in IoT systems.

DECLARATION OF CONFLICTS OF INTEREST

The authors declare that this research has no conflicts of interest.

ACKNOWLEDGMENTS

This research was funded by Ho Chi Minh City University of Technology and Engineering, Vietnam, under grant No. T2026-158.

REFERENCES

- [1] A. Shawahna, S. Sait, and A. E. Maleh, "FPGA-based accelerators of deep learning networks for learning and classification: A review," *IEEE Access*, vol. 7, pp. 7823-7859, 2019.
- [2] V. K. Pham, N. Q. Tran, and N. L. Nguyen, "Optimizing the convolutional neural networks for resource-constrained hardware," *Science & Technology Development Journal - Engineering and Technology*, vol. 4, no. 4, p. 906, 2022.
- [3] S. Bouguezzi et al., "An efficient FPGA-based convolutional neural network for classification: Ad-MobileNet," *Electronics*, vol. 10, no. 18, p. 2272, 2021.
- [4] M. Faizan, I. Intzes, I. Cretu, and H. Meng, "Implementation of deep learning models on an SoC-FPGA device for real-time music genre classification," *Technologies*, vol. 11, no. 4, p. 91, 2023.
- [5] Anjali and J. P. Anita, "AXI based DMA memory system testbench architecture using UVM harness technique," in *Proc. 9th International Conference on Advances in Computing and Communication (ICACC)*, 2019, pp. 152-157.
- [6] A. Sharma, "Evaluation of AXI-interfaces for hardware software communication," master's thesis, Technische Universität Chemnitz, Chemnitz, Germany, 2019.
- [7] S. Sun, J. Zou, Z. Zou, and S. Wang, Eds., *Experience of PYNQ: Tutorials for PYNQ-Z2*. Singapore: Springer Nature Singapore, 2023.
- [8] Y. Umuroglu et al., "FINN: A framework for fast, scalable binarized neural network inference," in *Proc. 2017 ACM/SIGDA International*

- Symposium on Field-Programmable Gate Arrays, 2017, pp. 65-74.
- [9] E. González, W. D. V. Luna, and C. A. Fajardo Ariza, "A hardware accelerator for the inference of a convolutional neural network," *Ciencia e Ingeniería Neogranadina*, vol. 30, no. 1, pp. 107-116, 2020.
- [10] Y. Liang et al., "Research on convolutional neural network inference acceleration and performance optimization for edge intelligence," *Sensors*, vol. 24, no. 1, p. 240, 2024.
- [11] H. Alemdar, V. Leroy, A. Prost-Boucle, and F. Pétrot, "Ternary neural networks for resource-efficient AI applications," in Proc. 2017 International Joint Conference on Neural Networks (IJCNN), 2017, pp. 2547-2554.
- [12] J. Park and W. Sung, "FPGA based implementation of deep neural networks using on-chip memory only," in Proc. 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016, pp. 1011-1015.



© 2026 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Mô hình AI trên FPGA: Mô hình CNN gọn nhẹ thông lượng cao và công suất thấp cho bài toán nhận dạng chữ số

Phạm Văn Khoa*, La Công Lộc, Nguyễn Đức Trọng, Huỳnh Minh Quý,
Hồ Gia Huyền, Võ Quang Huy, Phạm Bá Lộc, Huỳnh Thị Thu Hiền

Trường Đại học Công nghệ Kỹ thuật TP. Hồ Chí Minh, Việt Nam

Ngày nhận bài: 18/10/2025; Ngày sửa bài: 05/01/2026;

Ngày nhận đăng: 23/01/2026; Ngày xuất bản: 28/06/2026

TÓM TẮT

Nghiên cứu này trình bày việc thiết kế và triển khai một mạng nơ-ron tích chập trên nền tảng SoC-FPGA để phân loại chữ số viết tay sử dụng bộ dữ liệu MNIST. Mục tiêu là xây dựng một bộ gia tốc CNN gọn nhẹ và hiệu quả, có dưới 1,000 tham số, hoạt động tương thích với bộ xử lý ARM trên bo mạch PYNQ-Z2 thông qua các giao tiếp DMA và AXI. Bộ gia tốc được hiện thực ở mức RTL, với các giai đoạn mô phỏng, tổng hợp và tối ưu hóa tài nguyên, đồng thời vẫn duy trì được độ chính xác của quá trình suy luận. Trên 10,000 ảnh kiểm thử MNIST, hệ thống đạt độ chính xác 91.28% - thấp hơn khoảng 5% so với mô hình chạy trên CPU hai nhân ARM Cortex-A9 (96.26%) - nhưng lại mang lại tốc độ xử lý nhanh hơn 7 lần và giảm 36% mức tiêu thụ điện năng. Thiết kế cho thấy hiệu quả của việc song song hóa và pipeline hóa các phép tích chập trực tiếp trên FPGA, giúp giảm đáng kể mức sử dụng tài nguyên và công suất tiêu thụ. Những kết quả này cung cấp một nền tảng thực tiễn cho các ứng dụng AI nhúng thời gian thực - chẳng hạn như nhận dạng ký tự, giám sát hình ảnh, hệ thống IoT thông minh và tính toán biên - trên các nền tảng SoC-FPGA.

Từ khóa: Bộ tăng tốc mạng nơ-ron tích chập (Convolutional Neural Network Accelerator), FPGA, hệ thống trên Chip (System on Chip), MNIST, phân loại ảnh.

*Tác giả liên hệ chính.

Email: khoapv@hcmute.edu.vn

Study on analysis and assessment of exposure levels and accumulation profiles of Polycyclic Aromatic Hydrocarbons (PAHs) in human hair and fingernail samples

Nguyen Duc Hieu*, Hoang Quoc Anh, Chu Thi Hue, Hoang Thi Huong Hue, Nguyen Thi Duyen, Nguyen Xuan Thao, Pham Quynh Thiem, Do Van Doanh, Tran Duc Chung, Nguyen Thi Anh Huong

University of Science, Vietnam National University, Hanoi, Vietnam

Received: 29/10/2025; Revised: 05/01/2026;

Accepted: 17/01/2026; Published: 28/06/2026

ABSTRACT

Polycyclic aromatic hydrocarbons (PAHs) may accumulate in some non-invasive human samples, which can serve as bioindicators for assessing human exposure to these contaminants. In this study, hair and fingernail samples collected from a group of university students in Vietnam were collected and analyzed to provide biomonitoring information on the accumulation and contamination characteristics of PAHs. This is the first study in Vietnam to simultaneously analyze PAHs in hair and fingernails obtained from the same study subjects. Total concentrations of 16 PAHs in the hair samples ranged from 258 to 811 (mean 493) ng/g, while those in fingernails ranged from 29 to 667 (mean 335) ng/g. Although the total PAH levels in the hair samples were slightly higher than in the fingernails, the latter matrix preferably accumulated high-molecular-weight PAHs with higher toxicity, resulting in higher benzo[a]pyrene equivalents (BaP-EQ) in the fingernails than in hair samples. In contrast, many hair samples accumulated low-molecular-weight PAHs such as naphthalene and phenanthrene, implying the exposure from the air environment and/or endogenous accumulation during hair growing process. Some fingernail samples, especially in women, showed BaP-EQ levels several times higher than the corresponding hair samples, suggesting the role of exogenous adsorption from dust and hand contact with potentially PAH-contaminated surfaces.

Keywords: *PAHs, human fingernail, human hair, non-invasive human biomonitoring, GC-MS.*

1. INTRODUCTION

As a typical class of organic pollutants, polycyclic aromatic hydrocarbons (PAHs) have attracted considerable research attention over the past decades due to their adverse effects on both human health and the environment [1]. Some PAHs, notably benzo[a]pyrene, have been identified as carcinogenic and mutagenic to humans [2]. In terms of origin, PAHs can arise from both natural and anthropogenic

processes [2]. While some natural formation may occur through endogenous synthesis in plants and microorganisms, the majority of PAHs are generated from anthropogenic activities, particularly through incomplete combustion or pyrolysis of organic matter under high temperatures [3]. Humans are exposed to these contaminants mainly through inhalation of polluted air, ingestion of contaminated food and water, dust and soil ingestion, and dermal contact with PAH-containing materials [4].

*Corresponding author:

Email: nguyenduchieu_t65@hus.edu.vn

These accidental exposure events can directly impact public health, especially people affected by extensive industrial, transportation, domestic, and waste processing activities [5]. Exposure assessment and biomonitoring of PAHs in the human body is an essential task.

The exposure assessment can be divided into two approaches: external exposure and internal exposure. The external exposure assessment in humans is the process of predicting or estimating the amounts or doses of pollutants enter the human bodies through contacting with environmental media (such as air, soil, or water), based on contamination degree, exposure frequency, and duration [6]. The internal exposure is performed by determining pollutants and their metabolites in human samples such as urine [7], [8], blood [9], [10], or breast milk [9]. Considering the toxicology of PAHs, human biomonitoring plays a crucial role in assessing the accumulation of these substances in the body and the potential health risks [11]. However, collection and storage of human samples require critical protocols under strict conditions [11]. To overcome limitations related to conventional biological matrices, alternative approaches using matrices such as hair or nails have been proposed as non-invasive human biomonitoring. These matrices can reflect average exposure concentrations over longer periods and are less influenced by short-term fluctuations [12]. Additionally, these samples are easily collected and stored, showing insignificant impacts on the health and spirits of the donors [13].

In this study, we developed and applied a rapid, simple, and reliable analytical method to determine 16 priority PAHs in hair and fingernail samples collected from male and female students at several universities in Hanoi. The analytical procedure consisted of four main steps: (1) alkaline digestion of the samples, (2) extraction with an organic solvent, (3) clean-up using solid-phase extraction (SPE) cartridges, and (4) quantitative analysis by gas chromatography-mass spectrometry (GC-MS). This study aimed

to provide a preliminary assessment of PAH contamination and accumulation in these two biological matrices, reflecting human exposure degree and tissue-specific accumulation pattern. To the best of our knowledge, this study is among the first studies in Vietnam to simultaneously elucidate PAH exposure in human hair and fingernail samples.

2. MATERIALS AND METHOD

2.1. Samples collection

The hair and fingernail samples were collected in June 2024 from students at universities in Hanoi. Prior to sample collection, all participants were clearly informed about the purpose of the study and agreed to provide samples. The samples were obtained from 10 volunteers, including 5 males and 5 females. The sample categories were assigned as MH (male hair samples), FH (female hair samples), MN (male nail samples), and FN (female nail samples). The hair samples were cut approximately 2–3 cm from the scalp, while the fingernail samples were clipped using stainless steel clippers. The collected samples (about 0.5 g) were stored in polyethylene zip-lock bags at room temperature until analysis.

2.2. Standards and chemicals

The native PAH standard mixture (H-QME-01 Quebec PAH Mix; AccuStandard, USA) contained naphthalene (Nap), acenaphthylene (Acy), acenaphthene (Ace), fluorene (Flu), phenanthrene (Phe), anthracene (Ant), fluoranthene (Flt), pyrene (Pyr), benz[a]anthracene (BaA), chrysene (Chr), benzo[b]fluoranthene (BbF), benzo[k]fluoranthene (BkF), benzo[a]pyrene (BaP), dibenz[a,h]anthracene (DA), indeno[1,2,3-cd]pyrene (IP), and benzo[ghi]perylene (BP) at a concentration of 500 µg/mL each compound. A deuterated surrogate standard mixture (ES-2044 Surrogate Cocktail; Cambridge Isotope Laboratories, USA) contained Nap-d₈, Acy-d₈, Phe-d₁₀, Pyr-d₁₀, BaP-d₁₂, and BP-d₁₂ at a concentration of 200 µg/mL each compound. The internal standard solution contained Chr d₁₂ (DLM-261-1.2;

Cambridge Isotope Laboratories, USA) at a concentration of 200 µg/mL. Working standard solutions for native, surrogate, and internal standards at a concentration of 100 ng/mL were prepared from the stock standards using hexane as the solvent. The calibration standards were prepared in hexane over a concentration range of 10–500 ng/mL.

Chromatographic-grade solvents used in this study included acetone (99.5%; Macron Fine Chemicals™, Germany), hexane (96%; Daejung, South Korea), dichloromethane (DCM) (99.5%; Daejung, South Korea), and ethyl acetate (EA, 99.98%; Fisher Chemical, Mexico). Sodium hydroxide (> 98.5%) and sodium sulfate (> 99%) were obtained from Sharlau (Spain). A 1 M sodium hydroxide solution for sample digestion was prepared by dissolving 2 g of NaOH in 50 mL of deionized water. Bond Elut SI silica gel SPE cartridges (500 mg, 3 mL, 120 µm) were purchased from Agilent Technologies (USA).

2.3. Chemical analysis

The nail and hair samples (0.2 – 0.4 g) were cut into small pieces (< 0.2 mm), weighed into 10 mL glass tubes, and spiked with 200 µL of the surrogate solution (100 ng/mL). The samples were then digested with 5 mL of 1 M NaOH solution at room temperature, vortexed for 1 min, ultrasonicated for 10 min, and left to stand for 24 h for complete digestion. The sample mixture was then added with 2 mL of hexane, vortexed for 1 min, and centrifuged at 3500 rpm for 5 min. The upper organic phase was collected and transferred to a clean 10 mL glass tube. The extraction process was repeated two more times. The combined organic extract (approximately 6 mL) was dehydrated with anhydrous sodium sulfate, vortexed for 1 min, and centrifuged at 3500 rpm for 5 min. The dried extract was transferred to a concentrator tube and evaporated under a gentle nitrogen stream to 500 µL. The sample cleanup was performed using a silica gel SPE cartridge. The cartridge was preconditioned with 6 mL DCM and 6 mL hexane. The concentrated extract (500 µL) and

rinsate (500 µL) were loaded onto the cartridge, and this loading sample was discarded. PAHs were eluted with 3 mL of DCM/hexane (1:3, v/v) mixture. The eluate was spiked with 200 µL of internal standard solution (100 ng/mL) and concentrated to 200 µL before GC-MS analysis.

2.4. Instrumental analysis

PAHs were separated and quantified using a gas chromatograph (GC 8890, Agilent Technologies) coupled with a triple quadrupole mass spectrometer (7010B GC/TQ, Agilent Technologies) equipped with a DB-5ms capillary column (30 m × 0.25 mm × 0.25 µm; Agilent Technologies). The column oven temperature was programmed as follows: initial temperature 80 °C (held for 1 min), to 170 °C (20 °C/min), to 220 °C (4 °C/min), to 270 °C (3 °C/min), to 310 °C (20 °C/min, held for 20 min). Helium was used as the carrier gas at a flow rate of 1.2 mL/min. Samples (1 µL) were injected in splitless mode. The interface and ion source temperatures were set at 310 °C and 230 °C, respectively. The mass spectrometer operated in electron impact (EI) ionization mode at 70 eV, and data acquisition was performed in selected ion monitoring (SIM) mode.

2.5. QA/QC

Method blanks ($n = 5$) were analyzed in parallel with real samples to control for potential PAH contamination during the analytical procedure. The method detection limit (MDL) was determined using the formula: $MDL = C_B + 3 \times S_B$, where C_B is the mean concentration of PAHs in the blanks and S_B is the corresponding standard deviation. The MDLs for PAHs ranged from 0.04 to 10 ng/g. The recoveries of PAHs in spiked samples (67–102%) and surrogate compounds in real samples (60–121%) satisfied the acceptance criteria of the AOAC International and the US Environmental Protection Agency (US EPA) for analytes at ppb levels. Method repeatability, evaluated using spiked samples at 20 and 50 ng/g ($n = 3$ for each level), showed relative standard deviations (RSD) of less than 25%.

2.6. Statistical analysis

Analytical results were processed using Microsoft Excel® (Microsoft Corporation) and Minitab® 20 Statistical Software (Minitab, LLC). Significance level was set to 0.05.

3. RESULTS AND DISCUSSION

3.1. Concentrations of PAHs in hair samples

Total concentrations of 16 PAHs in the hair samples ranged from 258 to 811 (mean 493) ng/g, as presented in Figure 1. The highest level was observed in sample FH1 (811 ng/g), while the lowest was found in sample FH4 (258 ng/g). The mean PAH concentration in female hair samples (544 ng/g) was higher than that in the male hair samples (441 ng/g). However, this difference was not statistically significant (Mann-Whitney U test, $p > 0.05$), probably due to high variation within each group and the limited sample size. The PAH concentrations detected in our study were in line with those previously reported for hair samples collected from Vietnamese people aged 16 to 71, ranging from 217 to 1173 ng/g [14].

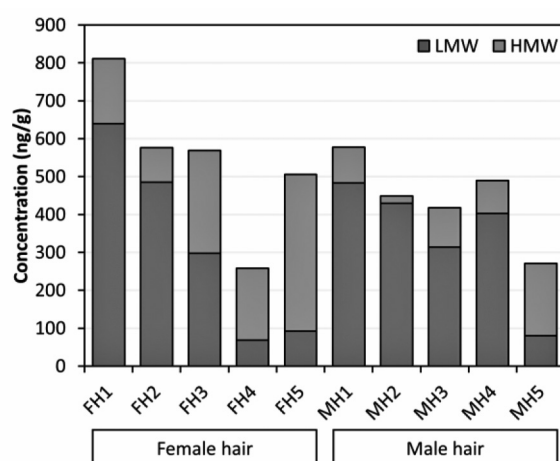


Figure 1. Concentrations of PAHs in hair samples (FH: female hair, MH: male hair).

Lifestyle factors (e.g., personal hygiene, smoking habits, cooking habits, and residential environments) can affect PAH levels in human hair. Frequent exposure to PAH-emitting activities such as smoking, cooking, and traffic may result in higher PAH accumulation in human hair [15], [16]. Tobacco smoke is recognized

as a major source of hair-associated PAHs in human [17], [18]. Previous studies have reported significantly higher PAH concentrations in the hair of smokers (mean 310 ng/g) compared to non-smokers (mean 183 ng/g) [18]. Elevated PAH levels have been observed in the blood and breast milk of smokers [19], [20]. It is estimated that more than 70% of total human exposure to PAHs was resulted from inhalation of tobacco smoke and consumption of contaminated food [17]. All the sample donors of our study are non-smoking students living in dormitories or student accommodations located in urban areas, with similar transportation media as motorcycles. One possible explanation for higher PAH levels in female hair than in male hair is the higher frequency of haircuts in men compared to women, resulting in longer accumulation duration for PAHs in female hair.

In addition, residential locations play an important role in the PAH accumulation in hair. Wang et al. (2020) examined PAH levels in adult hair samples collected from two regions in China (Nanjing and Ningbo) and reported that hair-bound PAH concentrations in Nanjing (317–1580 ng/g, mean 753 ng/g) were several times higher than those in Ningbo (60–440 ng/g, mean 222 ng/g) [21]. This discrepancy was attributed to the higher population density, traffic intensity, and presence of chemical industries in Nanjing, whereas Ningbo is a rural area with mainly agricultural activities and insignificant industrial emissions [21]. Lin et al. (2020) reported the concentrations of 15 PAHs from 6.24 to 692 ng/g in several exposure groups in China, including electronic waste recycling (EW) workers, non-EW workers, adults, and children in China [15]. They indicated that PAH levels in the hair of EW workers were markedly higher than those of the other groups, and that the PAH concentrations in children's hair were approximately 2.5–3 times lower than in adults [15]. These findings highlight that the exposure to direct emission sources of PAHs significantly influences the accumulation of these contaminants in the human body [15], [18], [21].

3.2. Concentrations of PAHs in nail samples

Concentrations of PAHs in the nail samples ranged from 29 to 667 (mean 335) ng/g (Figure 2). Overall, PAH levels in nails were lower than those observed in hair from the same individuals; however, a pronounced inter-individual variability was observed. The female nail sample FN2 exhibited the highest concentration of 667 ng/g, whereas the male nail sample MN2 showed the lowest level of 29 ng/g, suggesting substantial differences in exposure degree and living habits among these participants. A clear sex-dependent pattern was found, with female students showing a considerably higher mean PAH concentration in nails (474 ng/g) than male students (197 ng/g), approximately 2.4 fold higher. This difference can be attributed to both morphological and behavioral factors: (i) longer nails provide a greater surface area for the retention of hydrophobic contaminants and (ii) females generally trim their nails less frequently, prolonging the contact time with contaminated dust and household surfaces [22], [23].

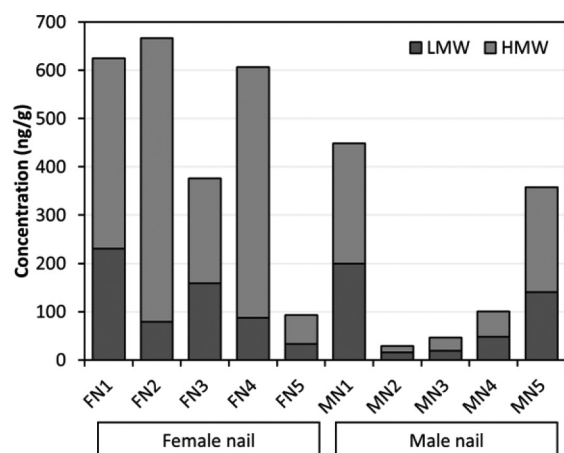


Figure 2. Concentrations of PAHs in nail samples (FN: female nail, MN: male nail).

Unlike hair, nail contamination is strongly influenced by direct external contact, indicating that nails predominantly capture external exposure pathways (hand–object–dust transfer) rather than endogenous accumulation [24], [25]. The primary mechanism behind PAH accumulation in nails is surface adsorption, where hydrophobic aromatic rings interact

with lipids, sebum residues, or particulate matter lodged in the subungual gap through hydrophobic and π - π interactions [26]. Since fingernails consist of avascular hard keratin, endogenous transfer through the bloodstream is negligible, in contrast to hair, which may incorporate contaminants during its growth phase through systemic circulation [26], [27]. Thus, fingernails represent a biomarker for external exposure, whereas hair reflects both internal and external exposure [26], [27].

The PAH levels measured in the fingernail samples of our study were within the ranges reported for similar samples from China. Ma et al. (2021) measured PAH concentrations in nail samples from four different exposure groups, ranging from 7.97–551 (mean 127), 7.05–431 (63.2), 7.93–289 (70.5), and 8.88–1280 (509) ng/g for EW workers, non-EW workers, adults, and children, respectively [24]. Tang (2020) revealed that children had higher nail concentrations than adults, attributed to frequent hand contact with indoor dust [25]. Another study conducted in Guangdong urban areas reported substantially higher PAH levels (98–3690 ng/g; mean 831 ng/g), which may be associated with intense urban–industrial emissions in the region [28]. The sex-related trend of higher PAH concentrations in female nail samples is consistent with international observations, reinforcing the role of nail morphology and hand-contact behavior in determining the level of PAH accumulation [24], [28].

3.3. Implications for sample washing steps

Analytical procedures, especially the washing step of fingernail and hair samples before extraction, can alter PAH concentrations and/or profiles. The removal of particulate matter (e.g., dust and soil) on the sample surface is critical for distinguishing endogenous PAHs from exogenous PAHs. The human hair samples have been decontaminated by washing with sodium dodecyl sulfate (SDS) surfactant followed by methanol [16], acetone [15], or ultrapure water [21]. Meanwhile, Wang et al. (2013) analyzed

PAHs in human hair samples without a washing step [18]. Ma et al. (2020) investigated washing procedures for the fingernail samples with different solvent types, including acetone, dichloromethane, ethyl acetate, hexane, and water [24]. The X-ray fluorescence (XRF) results of the washed fingernail samples indicated that acetone was the most suitable solvent with little damage to the fingernail surface, thus preventing the loss of endogenous PAHs [24]. In our study, we applied a different way to remove exogenous PAHs from the hair and fingernail samples. The sample donors were requested to take their samples immediately after their personal hygiene, thus obtaining clean hair and fingernail samples with minimal loss due to sample damage by organic solvent.

3.4. Profiles of PAHs in nail and hair samples

The compositional characteristics of PAHs in nail and hair samples are presented in Figure 3. A clear difference in PAH accumulation patterns was observed between the two biological matrices, despite being collected from the same individuals. In the nail samples, Phe accounted for the highest proportion (average 23%), followed by Pyr (14%), BbF (13%), Flt (10%), BaP (8%), DA (7%) and Chr (3%). The remaining PAHs, including BkF, IP, Ant, BaA and Flu, contributed approximately 2–3%, while Acy and Ace represented less than 2% of the total PAHs. High-molecular-weight PAHs (HMW-PAHs, 4–6 rings) dominated the profile (70%), substantially exceeding low-molecular-weight PAHs (LMW-PAHs, 2–3 rings; 30.2%), suggesting that the nail largely reflects external accumulation through particulate adhesion. This is consistent with the strong hydrophobicity and environmental persistence of HMW-PAHs, which readily adsorb and remain stable on the highly cross-linked keratin surface of nails. The overall profile is comparable to previous studies conducted in China, where Phe and Pyr were also found to be predominant constituents [24], [25]. However, Zeng et al. reported high levels of Nap, indicating potential differences in emission sources or site-specific exposure [28].

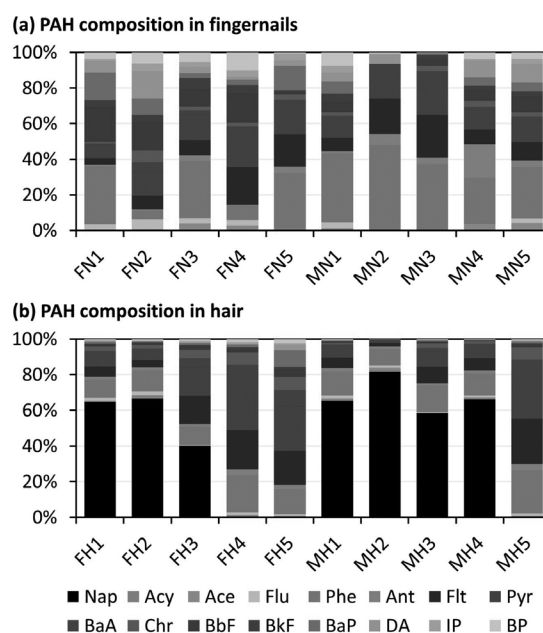


Figure 3. Profile of PAHs in fingernail and hair samples.

In contrast, the hair samples exhibited a markedly different composition, with Nap contributing the largest proportion (50%), followed by Phe (13%), Pyr (12%) and Flt (10%), while HMW-PAHs such as BaP, DA, IP and BP each accounted for only 1–3%. Overall, LMW-PAHs (67%) were dominant over HMW-PAHs (33%). This distribution pattern reflects the dual accumulation pathway of hair, in which LMW-PAHs more readily penetrate the hair shaft during growth or are re-adsorbed to the hair surface via lipid-associated mechanisms, while HMW-PAHs exhibit more limited incorporation unless the exposure source is particularly intense or localized [26]. The large inter-individual variability in Nap levels, with some samples (e.g., FH1) exceeding 60% of total PAHs whereas others showed non-detectable concentrations, also indicates differences in microenvironmental contact such as indoor air, fuel combustion, or personal activity pattern. An exceptional sample, such as FH5, showed higher BaP, DA and BP concentrations than the remaining samples, implying potential exposure from specific sources such as industrial emissions or vehicular exhaust. The compositional patterns observed in our hair samples are consistent with international

literature, where Nap, Flt, Phe and Pyr frequently dominate the PAH profiles in human hair [15], [18], [21].

3.5. Estimation of BaP equivalents

The benzo[a]pyrene equivalents (BaP-EQ) of the PAH mixture were calculated from the measured concentrations with the corresponding toxic equivalency factors (TEFs) [29]. The TEF approach, originally proposed by Nisbet and LaGoy, allows individual PAHs to be expressed relative to the toxicity of BaP, the most critical PAH substance in terms of toxicity and carcinogenicity [29]. The BaP-EQ levels in the samples of this study are shown in Figure 4.

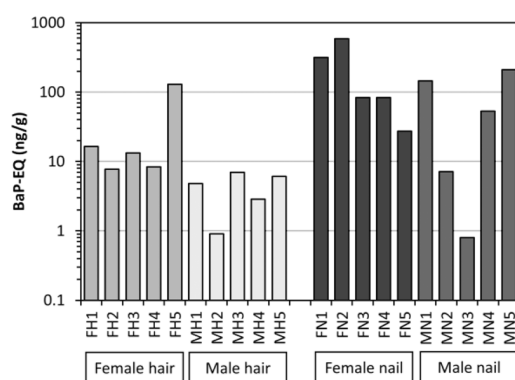


Figure 4. BaP-equivalent toxicity (BaP-EQ) levels in hair and nail samples.

Levels of BaP-EQ ranged from 0.90 to 130 (average 20) ng BaP-EQ/g in the hair samples and from 0.80 to 591 (average 150) ng BaP-EQ/g in the nail samples. In addition to the differences in total PAH concentrations, a clear contrast was also observed in the BaP-EQ levels between nails and hair. Among females, nail samples such as FN1 and FN2 exhibited remarkably high BaP-EQ values (318 and 591 ng/g, respectively), far exceeding those found in the corresponding hair samples, even though the overall PAH levels in hair were not significantly lower. The hair samples showed relatively low BaP-EQs even in cases where the total PAH concentrations are high. Relatively high BaP-EQ levels were found in male nail samples (e.g., 145 and 210 ng BaP-EQ/g in MN1 and MN5), although the levels in corresponding

hair samples were quite low. The BaP-EQ profiles in the hair samples were dominated by BaP (average 36%), DA (28%), BaA (13%), BbF (5%), BkF (2%), and IP (2%). Meanwhile, the BaP-EQs in the nail samples showed the highest contributions of DA (69%), followed by BaP (21%), BbF (5%), and BkF (2%). These results suggest that the accumulation of PAHs in general and highly toxic compounds in particular depends on the physicochemical properties of pollutants, emission source patterns, and tissue-specific distribution. In-depth evaluation of the relationships between PAH exposure levels and specific human health risks in Vietnam should be characterized in future studies.

4. CONCLUSION

This study provides preliminary data on the occurrence of 16 PAHs, including concentrations, accumulation patterns, and BaP-EQ levels in hair and fingernail samples collected from university students in Vietnam. Overall, hair samples exhibited higher total PAH concentrations than nail samples, and female participants generally showed higher levels than males. The ranges observed in this dataset are comparable with those reported in previous studies carried out in residential areas and informal e-waste recycling sites in China. Regarding compositional profiles, nails were dominated by Phe, Pyr, BbF, BaP, and DA, whereas hair was dominated by Nap, Phe, and Pyr. This distribution pattern is consistent with earlier findings in the literature, supporting the view that nails mainly reflect externally deposited PAHs originating from dust and surface contact, while hair captures a mixture of endogenous and external exposure. As a result, BaP-EQ values were markedly higher in nails than in hair, indicating that the nail matrix represents not only the presence of PAHs but also the toxic relevance of exposure, whereas hair better reflects the overall exposure burden. To the best of our knowledge, this is the first study in Vietnam to simultaneously assess PAHs in both hair and nail matrices from the same individuals, thereby offering a new perspective on how these two keratin-based tissues complement each

other in exposure assessment. Further studies should be performed to characterize potential health impacts of PAHs and to clarify how these exposure patterns translate into toxicological or epidemiological outcomes.

DECLARATION OF CONFLICTS OF INTEREST

The authors declare that this research has no conflicts of interest.

ACKNOWLEDGMENTS

This research is funded by Vietnam National University, Hanoi under grant number QG.24.14.

REFERENCES

- [1] C. Wang, S. Zhou, J. Song, S. Wu, "Human health risks of polycyclic aromatic hydrocarbons in the urban soils of Nanjing, China," *Science of The Total Environment*, vol. 162, pp. 750-757, 2018.
- [2] H. I. A. Shafy, M. S. M. Mansour, "A review on polycyclic aromatic hydrocarbons: Source, environmental impact, effect on human health and remediation," *Egyptian Journal of Petroleum*, vol. 25, pp. 107-123, 2016.
- [3] C. Wang, S. Wu, S. Zhou, Y. Shi, J. Song, "Characteristics and Source Identification of Polycyclic Aromatic Hydrocarbons (PAHs) in Urban Soils: A Review," *Pedosphere*, vol. 27, pp. 17-26, 2017.
- [4] K. Sun, Y. Song, F. He, M. Jing, J. Tang, R. Liu, "A review of human and animals exposure to polycyclic aromatic hydrocarbons: Health risk and adverse effects, photo-induced toxicity and regulating effect of microplastics," *Science of The Total Environment*, vol. 773, p. 145403, 2021.
- [5] P. M. Abdulai, O. B. Ojimadu, A. P. Onyena, C. Frazzoli, N. A. Mogborukor, O. C. Ekhatior, G. J. Udom, E. Nwanaforo, O. E. Orisakwe, "Public Health Effects of Polycyclic Aromatic Hydrocarbons Exposure Through Air, Water, Soil, and Food in Ghana: Possible Economic Burden," *Environmental Health Insights*, vol. 19, p. 11786302251343767, 2025.
- [6] D. J. Paustenbach, "The practice of exposure assessment: a state-of-the-art review," *Journal of Toxicology and Environmental Health, Part B*, vol. 3, pp. 179-291, 2000.
- [7] O. Alshaarawy, H. A. Elbaz, M. E. Andrew, "The association of urinary polycyclic aromatic hydrocarbon biomarkers and cardiovascular disease in the US population," *Environment International*, vol. 89-90, pp. 174-178, 2016.
- [8] S. D. Craemer, K. Croes, N. V. Larebeke, I. Sioen, G. Schoeters, I. Loots, T. Nawrot, V. Nelen, L. Campo, S. Fustinoni, W. Baeyens, "Investigating unmetabolized polycyclic aromatic hydrocarbons in adolescents' urine as biomarkers of environmental exposure," *Chemosphere*, vol. 155, pp. 48-56, 2016.
- [9] S. N. Zhou, A. Buchar, S. Siddique, L. Takser, N. Abdelouahab, J. Zhu, "Measurements of selected brominated flame retardants in nursing women: Implications for human exposure," *Environmental Science & Technology*, vol. 48, pp. 8873-8880, 2014.
- [10] X. Zhang, X. Li, Y. Jing, X. Fang, X. Zhang, B. Lei, Y. Yu, "Transplacental transfer of polycyclic aromatic hydrocarbons in paired samples of maternal serum, umbilical cord serum, and placenta in Shanghai, China," *Environmental Pollution*, vol. 222, pp. 267-275, 2017.
- [11] É. Gaudreau, R. Bérubé, J. F. Bienvenu, N. Fleury, "Stability issues in the determination of 19 urinary (free and conjugated) monohydroxy polycyclic aromatic hydrocarbons," *Analytical and Bioanalytical Chemistry*, vol. 408, pp. 4021-4033, 2016.
- [12] A. Alves, A. Covaci, S. Voorspoels, "Method development for assessing the human exposure to organophosphate flame retardants in hair and nails," *Chemosphere*, vol. 168, pp. 692-698, 2017.
- [13] M. Esteban, A. Castaño, "Non-invasive matrices in human biomonitoring: A review," *Environment International*, vol. 35, no. 2, pp. 438-449, 2009.
- [14] P. T. Vi, N. T. Ngoc, T. T. Kim, N. T. Quynh, D. H. Anh, P. H. Viet, "Determination of PAHs content in hair samples of smokers and non-smokers using gas chromatography in hyphenation with mass spectrometry," *VNU Journal of Science: Natural Sciences and Technology*, vol. 35, pp. 118-127, 2019.
- [15] M. Lin, J. Tang, S. Ma, Y. Yu, G. Li, R. Fan, B. Mai, T. An, "Insights into biomonitoring

- of human exposure to polycyclic aromatic hydrocarbons with hair analysis: A case study in e-waste recycling area,” *Environment International*, vol. 136, p. 105432, 2020.
- [16] P. Palazzi, S. Mezzache, N. Bourokba, E. Hardy, A. Schritz, P. Bastien, C. Emond, J. Li, J. Soeur, B. M. R. Appenzeller, “Exposure to polycyclic aromatic hydrocarbons in women living in the Chinese cities of BaoDing and Dalian revealed by hair analysis,” *Environment International*, vol. 121, pp. 1341-1354, 2018.
- [17] T. E. McGrath, J. B. Wooten, W. G. Chan, M. R. Hajaligol, “Formation of polycyclic aromatic hydrocarbons from tobacco: The link between low temperature residual solid (char) and PAH formation,” *Food and Chemical Toxicology*, vol. 45, no. 6, pp. 1039-1050, 2007.
- [18] W. Wang, F. Wu, J. Zheng, M. H. Wong, “Risk assessments of PAHs and Hg exposure via settled house dust and street dust, linking with their correlations in human hair,” *Journal of Hazardous Materials*, vol. 263, pp. 627-637, 2013.
- [19] A. Kamal, M. Qayyum, I. U. Cheema, A. Rashid, “Biological monitoring of blood naphthalene levels as a marker of occupational exposure to PAHs among auto-mechanics and spray painters in Rawalpindi,” *BMC Public Health*, vol. 11, p. 467, 2011.
- [20] L. Zanieri, P. Galvan, L. Checchini, A. Cincinelli, L. Lepri, G. P. Donzelli, M. Del Bubba, “Polycyclic aromatic hydrocarbons (PAHs) in human milk from Italian women: Influence of cigarette smoking and residential area,” *Chemosphere*, vol. 67, no. 7, pp. 1265-1274, 2007.
- [21] C. Wang, S. Zhou, S. Wu, J. Tang, H. Li, J. Du, “Exposure to polycyclic aromatic hydrocarbons (PAHs) in people living in urban and rural areas as revealed by hair analysis,” *Chemosphere*, vol. 246, p. 125764, 2002.
- [22] C. R. Bobbins, *Chemical and Physical Behavior of Human Hair (5th edition)*. USA: Springer, 2012.
- [23] C. H. Liu, S. N. Doan, “Innovations in biological assessments of chronic stress through hair and nail cortisol: Conceptual, developmental, and methodological issues,” *Developmental Psychobiology*, vol. 61, no. 3, pp. 465-476, 2019.
- [24] S. Ma, Z. Zeng, M. Lin, J. Tang, Y. Yang, Y. Yu, G. Li, T. An, “PAHs and their hydroxylated metabolites in the human fingernails from e-waste dismantlers: Implications for human non-invasive biomonitoring and exposure,” *Environmental Pollution*, vol. 283, p. 117059, 2021.
- [25] J. Tang, S. Ma, R. Liu, C. Yue, G. Li, Y. Yu, Y. Yang, T. An, “The pollution profiles and human exposure risks of chlorinated and brominated PAHs in indoor dusts from e-waste dismantling workshops: Comparison of GC–MS, GC–MS/MS and GC×GC–MS/MS determination methods,” *Journal of Hazardous Materials*, vol. 394, p. 122573, 2020.
- [26] M. C. Golpe, A. d. C. Ríos, E. Lendoiro, “Current status of keratinized matrices in Toxicology: Comparison of hair and nails,” *Drug Testing and Analysis*, vol. 17, pp. 436-449, 2025.
- [27] C. R. Daniel, B. M. Piraccini, A. Tosti, “The nail and hair in forensic science,” *Journal of the American Academy of Dermatology*, vol. 50, pp. 258–261, 2004.
- [28] Z. Zeng, Y. Gao, J. Cui, M. Lin, J. Tang, F. Wang, Y. Yang, Y. Yu, S. Ma, “Liquid–liquid extraction combined with online cleanup for the simultaneous determination of PAHs by GC–MS/MS and their hydroxylated metabolites by LC–MS/MS in human fingernails,” *Journal of Chromatography B*, vol. 1188, p. 123057, 2022.
- [29] I. C. T. Nisbet, P. K. LaGoy, “Toxic equivalency factors (TEFs) for polycyclic aromatic hydrocarbons (PAHs),” *Regulatory Toxicology and Pharmacology*, vol. 16, no. 3, pp. 290–300, 1992.



© 2026 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Nghiên cứu phân tích và đánh giá mức độ phơi nhiễm, đặc trưng tích lũy hydrocacbon thơm đa vòng (PAHs) trong mẫu tóc và móng tay người

Nguyễn Đức Hiếu*, Hoàng Quốc Anh, Chu Thị Huệ, Hoàng Thị Hương Huệ, Nguyễn Thị Duyên, Nguyễn Xuân Thảo, Phạm Quỳnh Thiêm, Đỗ Văn Doanh, Trần Đức Chung, Nguyễn Thị Ánh Hoàng

Trường Đại học Khoa học Tự nhiên, Đại học Quốc gia Hà Nội, Việt Nam

Ngày nhận bài: 29/10/2025; Ngày sửa bài: 05/01/2026;

Ngày nhận đăng: 17/01/2026; Ngày xuất bản: 28/06/2026

TÓM TẮT

Hydrocarbon thơm đa vòng (PAHs) có khả năng tích lũy trong một số mẫu sinh phẩm người không xâm lấn và các loại mẫu này có thể được sử dụng làm chỉ thị sinh học để đánh giá mức độ phơi nhiễm các độc chất này ở người. Trong nghiên cứu này, mẫu tóc và móng tay của một nhóm sinh viên đại học tại Việt Nam được thu thập và phân tích nhằm cung cấp các thông tin quan trắc sinh học về mức độ tích lũy và đặc trưng ô nhiễm PAHs. Đây là nghiên cứu đầu tiên tại Việt Nam đánh giá đồng thời PAHs trong tóc và móng tay trên cùng một đối tượng nghiên cứu. Hàm lượng tổng của 16 PAHs trong mẫu tóc dao động từ 258 đến 811 (trung bình 493) ng/g, trong khi hàm lượng trong móng tay dao động từ 29 đến 667 (trung bình 335) ng/g. Mặc dù hàm lượng tổng PAHs trong tóc có xu hướng cao hơn nhưng móng tay lại giàu các PAHs khối lượng phân tử lớn có độc tính cao, dẫn đến giá trị độ độc tương đương quy về benzo[a]pyrene (BaP-EQ) cao hơn so với mẫu tóc. Ngược lại, một số mẫu tóc chủ yếu tích lũy các PAHs khối lượng thấp như naphthalene và phenanthrene, phản ánh sự phơi nhiễm từ môi trường không khí hoặc do sự tích lũy nội sinh trong quá trình hình thành sợi tóc. Một số mẫu móng tay, đặc biệt ở nữ giới, ghi nhận hàm lượng BaP-EQ cao hơn nhiều lần so với mẫu tóc tương ứng, cho thấy vai trò của cơ chế hấp phụ ngoại sinh từ bụi và sự tiếp xúc tay với các bề mặt có nguy cơ ô nhiễm PAHs.

Từ khóa: PAHs, móng tay, tóc, quan trắc sinh học người không xâm lấn, GC-MS.

*Tác giả liên hệ chính.

Email: nguyenduchieu_t65@gmail.com

Enhancing higher education faculty competencies in the age of generative AI: Integrating explainability, contestability, and reflective practice

Tran Diem Quynh Nguyen¹, Thi Cam Mai Truong^{2,*}

¹University of Foreign Language Studies, University of Danang, Vietnam

²Faculty of Natural Sciences, Quy Nhon University, Vietnam

Received: 03/11/2025; Revised: 22/12/2025;

Accepted: 25/12/2025; Published: 28/06/2026

ABSTRACT

The rapid emergence of generative artificial intelligence (AI) and large language models (LLMs) has created both unprecedented opportunities and significant challenges in higher education. While these technologies promise to enhance teaching effectiveness and research productivity, their black-box nature, tendency toward hallucinations, and opacity raise critical concerns about trust, accountability, and pedagogical integrity. This paper addresses the urgent need for a comprehensive framework to enhance faculty competencies in leveraging generative AI responsibly and effectively. We propose TECTRA (Trust through Explainability, Contestability, and Reflective Application), a novel human-centered framework that integrates Explainable AI (XAI) and Contestable AI (CAI) as foundational mechanisms for trustworthy AI adoption in education. The framework is structured around four interdependent pillars: (1) Ethical Grounding, enabled by XAI's transparency; (2) Pedagogical Integration, activated through CAI's dialogic structure; (3) Technical Literacy, developed through XAI's interpretable explanations; and (4) Reflective Practice, sustained through combined feedback loops from both mechanisms. We detail specific, measurable faculty competencies mapped to each pillar and provide concrete development activities and tools. Furthermore, we present a phased implementation strategy roadmap spanning assessment, capacity building, and sustainable scaling, alongside comprehensive policy recommendations that emphasize flexibility, transparency, human oversight, and ethical principles. By positioning XAI and CAI as active, functional elements rather than separate technical considerations, TECTRA transforms generative AI from an opaque tool into a transparent, contestable partner for critical inquiry, ultimately fostering enhanced faculty competencies that are adaptive, evidence-based, and ethically grounded in an increasingly AI-driven educational landscape.

Keywords: *Explainable AI, Contestable AI, Generative AI, Large Language Models, higher education.*

1. INTRODUCTION

The digital era has brought transformative technologies that are reshaping higher education worldwide. Among these, generative artificial intelligence (AI) and large language models (LLMs) such as OpenAI's ChatGPT, Anthropic's Claude, Meta's Llama, and

Google's Gemini have rapidly emerged as powerful tools with the potential to enhance the teaching and research capacities of university and college lecturers. Since the public release of LLMs, educators have been exploring how such AI systems can revolutionize teaching practices and academic research [1]. These AI

*Corresponding author:

Email: truongcammai@qnu.edu.vn

tutors and assistants provide opportunities to innovate pedagogical approaches, personalize learning, automate routine tasks, and support scholarly work. For instance, generative AI can help instructors generate lesson plans, produce quizzes or simulations, provide conversational tutoring to students, and even aid in literature recommending, grading, feedback, or interview preparation [2] - [4]. Likewise, researchers can leverage LLMs to summarize literature, draft manuscripts, or brainstorm research ideas. In theory, these capabilities promise to enhance lecturers' teaching effectiveness and research productivity in the digital age.

However, along with immense potential, generative AI introduces significant challenges and uncertainties in educational contexts. Modern LLMs are often "black boxes", producing answers without revealing clear reasoning or sources [5]. This opacity can erode trust, as educators and students may question how or why an AI arrived at a given answer. Additionally, LLMs are prone to hallucinations (i.e., generating incorrect or fabricated information confidently). In an academic setting, such undetected errors or falsehoods can mislead learners and undermine learning outcomes. Educators also worry about issues of academic integrity (e.g., plagiarism or uncritical use of AI in student work), biases in AI outputs, data privacy, and the broader ethical implications of delegating educational tasks to AI. These challenges highlight that while AI can assist humans, it cannot be blindly trusted, particularly in educational and academic settings, where rigor and accuracy are crucial [6].

To fully realize the benefits of generative AI in education, it is crucial to address these challenges. Two emerging approaches in AI research hold promise in this regard: Explainable AI (XAI) [7] - [10] and Contestable AI (CAI) [4], [11] - [15]. XAI aims to make AI systems more transparent by providing human-understandable explanations for their outputs or decisions. In an educational context, XAI could enable a lecturer or student to see the reasoning behind an

AI-generated answer, thereby improving trust and facilitating error diagnosis [10], [16] - [20].

Recent regulations, including the General Data Protection Regulation (GDPR) and the EU AI Act establish legal requirements for AI interpretability and explainability, fundamentally grounded in the principle of *contestability*, ensuring individuals can meaningfully challenge automated decisions. CAI goes a step further by enabling users to question, dispute, and engage in dialogue with the AI's decisions or reasoning [4], [11] - [14]. A CAI system would not only explain itself, but also allow teachers or students to challenge its responses and have the system revise its answers when justified [15]. Together, these approaches aim to transform AI into a more interactive and accountable assistant, rather than an opaque box.

Hence, this paper presents a comprehensive research study on how explainability and contestability can foster trustworthy use of generative AI in higher education, ultimately enhancing faculty competencies in the digital era. Our contributions are as follows:

1. We propose a novel framework, TECTRA (Trust through Explainability, Contestability, and Reflective Application), which integrates XAI and CAI to create a human-centered, trustworthy generative AI framework in education. We detail how this approach can fill current gaps and empower faculty across disciplines.

2. We design a phased strategy roadmap for implementation and policy recommendations, ensuring that universities can harness the benefits of generative AI while maintaining academic integrity, equity, and human oversight.

By adopting the principles of explainability and contestability for generative AI and LLMs, we aim to provide a practical framework with a phased roadmap, implementation guidelines, and policy recommendations to enhance the trust and fairness of leveraging these technologies in higher education.

2. BACKGROUND

2.1. Generative AI in Education

Generative AI refers to a class of AI models that can produce new content (e.g., text, images, music, or code) by learning patterns from existing data. A prominent example is the LLM, which can engage in human-like dialogue and create texts in response to prompts. OpenAI's ChatGPT, Anthropic's Claude, Meta's Llama, or Google's Gemini, and similar LLMs are generative AI models consisting of billions of parameters trained on massive textual databases and can perform tasks like answering questions, writing essays, summarizing documents, and creating lesson plans. In essence, these models predict text based on learned patterns, stringing together words that are statistically likely to follow a given prompt.

In education, generative AI has swiftly found diverse applications [2], [3], [21] - [27]. Students can utilize AI chatbots to explain complex concepts, generate ideas, or receive feedback on their writing drafts. Faculty members are exploring the use of generative AI to enhance teaching materials and workflows. For instance, an instructor can prompt an AI to generate quiz questions, example problems, or even first drafts of lesson summaries, which can save time on preparation [3], [22], [23]. Generative AI can also assist with research by summarizing scholarly articles or suggesting new research directions [2], [24]. From a learning perspective, these tools offer personalized support, where an AI tutor can converse with a student, quiz them on course content, and adjust the difficulty of questions to the student's level in real-time [25], [26]. This ability to provide instant, tailored feedback and access to information has led many to view generative AI as a powerful aid for both instructors and learners. Early implementations have demonstrated that generative AI can benefit a diverse range of learners. For example, translating or simplifying content for non-native speakers can spark creativity, providing new examples or analogies that enrich the learning experience [28], [29].

However, the use of generative AI in education also raises pedagogical questions. Because LLMs generate text based on patterns rather than genuine understanding, educators must consider how students' use of AI affects learning outcomes. Some studies suggest that while generative AI can increase productivity on certain tasks, it might reduce cognitive effort or lead to more homogeneous student work if over-relied upon [27], [30], [31]. Thus, a consensus is emerging that generative AI should augment teaching and learning, serving as a smart assistant or tutor, rather than replacing the essential human elements of creativity, critical thinking, and mentorship [3], [32]. To realize this vision, faculty members need to guide students in the proper use of AI, and importantly, they themselves must have a clear understanding of how generative AI works and where its outputs can or cannot be trusted. This sets the stage for examining issues of transparency and fairness in generative AI systems.

2.2. Transparency and Fairness: The “Black-Box” and Hallucination Problems

2.2.1. Black-box Problem

Despite their impressive capabilities, most state-of-the-art AI models operate as “black boxes.” In a black-box model, the internal reasoning leading to any given output is hidden or too complex for users and even developers to interpret. Users see both the input and the AI's output, but not the decision-making process that occurs in between [28], [33]. This opacity poses a serious problem in education, where trust and accountability are critical. Educators are understandably uneasy when an AI provides an answer or assessment without any explanation, especially knowing that even the engineers who built the model cannot fully explain how a particular result was generated. The lack of transparency makes it difficult to judge the correctness or bias of AI outputs, and it hinders the ability to contest or audit those outputs. As several studies of AI in education noted [34], [35], the black-box nature of some AI algorithms makes it challenging for stakeholders to understand or challenge

AI-driven decisions, raising ethical concerns about their use in education. In other words, if a generative AI system gives flawed information or an unfair recommendation, its inscrutable logic means educators might not realize the error or have the means to dispute it. This undermines confidence and can erode the educational integrity of AI-assisted processes.

2.2.2. Hallucination of Generative AI

Compounding the transparency issue is the tendency of generative AI models to produce hallucinations. Hallucination refers to the well-documented phenomenon where AI systems, especially generative AI and LLMs, generate incorrect, misleading, or entirely fabricated information or responses that do not accurately reflect the data they were trained on or the input provided to them [36], [37]. For example, generative AI might confidently generate a citation for a non-existent scientific article or incorrectly explain a concept in a way that appears authoritative [38]. These tools do not truly understand the facts or possess a

comprehensive knowledge model of reality. Instead, they base their responses on statistical patterns in the training data. Therefore, an AI might assert false facts because those word combinations seem likely.

Research has highlighted that general-purpose LLMs often draw on poor-quality or incorrect data absorbed during training, which can lead to incorrect outputs, and they typically lack the ability to verify information or cite specific sources for their statements. In academic contexts, such hallucinations can mislead students or propagate misinformation if unchecked. A faculty member using AI to generate lecture notes or a student relying on it for research may be misled by a confidently stated but false piece of information. Instructors also expressed concern about the use of generative AI in education, citing inaccurate or unreliable outputs as a top concern.

2.3. Explainability and Contestability for Generative AI in Higher Education

2.3.1 Explainable AI (XAI)

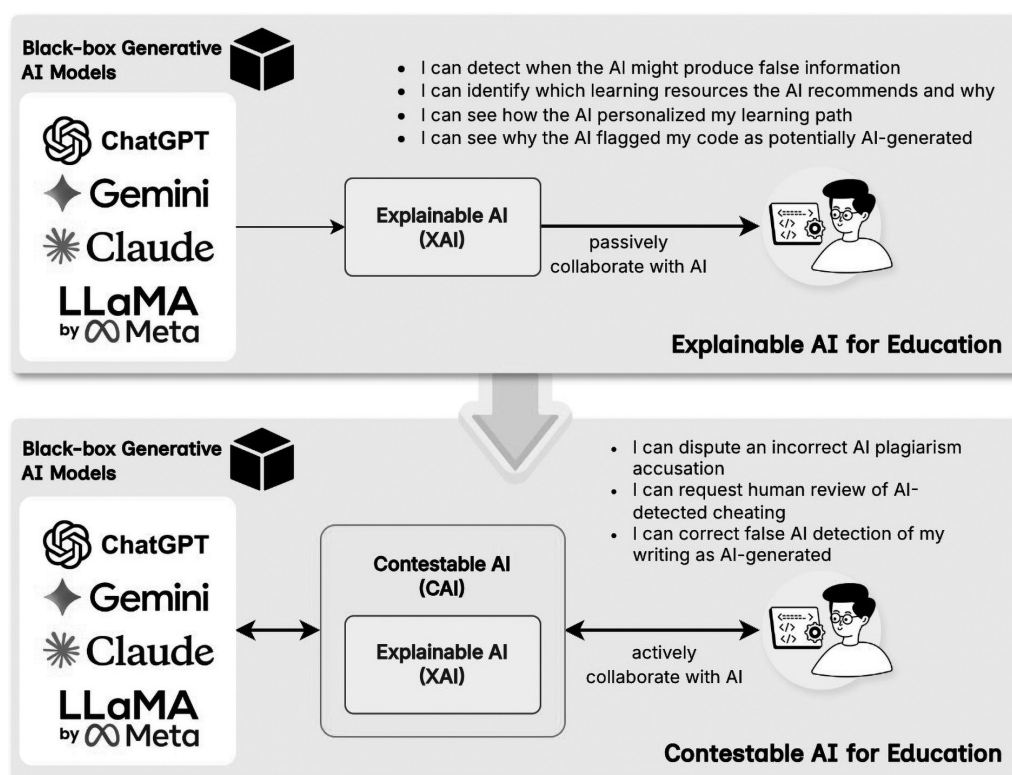


Figure 1. Moving from Explainable AI (XAI) to Contestable AI (CAI) in Enhancing Faculty and Student Competencies in Higher Education.

XAI refers to a set of methods and techniques that make the decisions or outputs of an AI system understandable to human users. Traditional AI models, particularly deep learning (DL) and LLMs, often function as black boxes, producing results without a clear explanation of how they were derived. XAI aims to open this box by providing human-interpretable insights into the model's reasoning, thereby enhancing trust, transparency, and accountability. Common XAI approaches include saliency mapping [9], [39] - [41], feature attribution [42], [43], and counterfactual explanations [44], [45].

For generative AI within the education sector, especially those used in natural language applications with LLMs, XAI techniques increasingly involve prompt engineering for rationale generation, rubric-based evaluations and play a critical role in supporting student learning, assisting teachers, and ensuring the fairness of AI-driven assessments [10], [18]-[20]. XAI technique can provide a justification in natural language, highlight the factors (input features) that most influenced its decision, or produce confidence scores and evidence (e.g, source citations) to support its output. For example, an explainable generative AI writing assistant might underline which parts of its

answer were drawn from which reference texts, or an AI used for student evaluation might show the rubric criteria and how it applied them. Recent works, such as ExASAG [18], CourseEvalAI [19], and QwenScore+ [10], incorporate explainable reasoning strategies into rubric-aligned evaluation to break down automated scoring into human-understandable criteria, enabling clear explanations for grades.

By opening up the AI's reasoning, XAI helps users verify correctness, detect errors or biases, and ultimately decide when to trust the AI and when to be skeptical. This is particularly important in education, where teachers need to understand an AI's suggestion in order to confidently act on it or explain it to students. If an AI can show its work, for instance, by citing the source of a fact or explaining the steps in a solution it provided, the instructor can more easily validate the result. For example, advanced tutoring systems like LPITutor [20] use retrieval-augmented generation (RAG) and user-adaptive prompts to ensure that the content and explanations are grounded in course materials and personalized to the learner's level. Indeed, research confirms that increasing a system's transparency via explainability can strengthen teachers' trust and willingness to accept AI recommendations.

Table 1. An Overview of Prominent Applications of Generative AI Integrated with Explainable AI (XAI) and Contestable AI (CAI) Frameworks in Education.

Method	XAI/CAI	Application	Contributions
ExASAG [18]	XAI	Automatic grading of short, free-text student answers.	Produces human-understandable rationales for each grade using XAI methods; Provides interpretable feedback, enabling teachers to understand and verify AI grading.
CourseEvalAI [19]	XAI	Transparent, rubric-based evaluation of LLMs for grading open-ended student work.	Fine-tunes an LLM using dual-layer rubrics (for answers and explanations); Evaluations and scores stored in a graph database for full traceability; Reduces bias, increases rubric fidelity, and improves the evaluability of model-generated explanations.

Method	XAI/CAI	Application	Contributions
QwenScore+ [10]	XAI	Essay scoring with formative feedback.	Applies LLMs to generate personalized, formative feedback rather than just grades; Emphasizes student trust, ethical design, and transparency in educational AI systems; Evaluates LLMs’ ability to provide scaffolded, rubric-aligned responses that enhance learning outcomes.
LPITutor [20]	XAI	Intelligent tutoring across subjects (adaptive questions and answers, hints).	Increases transparency and accuracy by grounding answers in retrieved documents (students can be shown the source or at least trust the answer is curriculum-aligned); Customizable difficulty: explanations and answers suited to learner’s skill level; Addresses hallucinations by tethering model to real content.
CAELF [15]	CAI	Interactive feedback on essays.	Highly robust to student push-back, maintains logical consistency and only changes grade if student’s counter-argument is valid; Inherently explainable feedback (built from explicit arguments and rubric criteria); Improves LLM’s reasoning and reduces susceptibility to manipulation in an educational dialogue.

These developments illustrate how XAI enhances learning experiences by providing context-aware, pedagogically useful feedback, turning AI systems into interpretable educational partners. Explainability also aids AI literacy, as educators interpret AI outputs, they learn more about the AI’s limits and behavior, becoming better equipped to integrate it appropriately. In short, XAI is a foundation for responsible AI adoption in education, as it addresses the black-box issue by illuminating the AI’s decision logic and thereby fosters user understanding and trust.

2.3.2. Contestable AI (CAI)

However, simply explaining an AI’s output does not fully resolve the power imbalance between human and machine. Even if a teacher knows why an AI produced a certain result, they also need the ability to say “I disagree” or override the AI when it is wrong. This is where the concept of CAI emerges as significant.

Contestability in AI systems means that users have the right and means to challenge or appeal an AI’s decision, and the system can

adapt or be corrected based on that human feedback [13], [14], [46], [47]. In other words, a CAI system is not a one-way authority but rather allows a dialogue or iterative process with human oversight. According to AI ethics frameworks, contestability implies that those affected by an AI outcome are provided with *plain and easy-to-understand* information about how the decision was made, enabling them to challenge the outcome if necessary. Contemporary regulatory frameworks, including the GDPR and the EU AI Act, highlight the critical role of interpretability and contestability. The GDPR’s Articles 13 and 14 establish that individuals subjected to profiling are entitled to receive *meaningful information about the logic involved*. As clarified in official EU guidance, these provisions extend to healthcare settings.

The explainability mandate within the GDPR should be interpreted as establishing a framework for contestability, requiring that AI-driven decisions be sufficiently transparent to enable individuals to challenge them. Article 22

specifies that when automated decision-making, including profiling, is legitimately applied to individuals, data controllers must protect their right “to express his or her point of view and to *contest* the decision”. Contestability is thus increasingly recognized as a key aspect of accountability in AI deployments.

The twin notions of XAI and CAI are complementary and together promise a more trustworthy AI ecosystem in education, as illustrated in Figure 1. Explainability provides transparency (“*Why did the AI say that?*”) and contestability provides agency (“*What can we do if the AI is wrong?*”). By embracing both, we address the earlier challenges: an explainable system reduces the fear of black boxes and helps identify errors or biases, while a contestable system ensures that those errors can be corrected and biases mitigated through human intervention. Importantly, contestability reinforces human authority and accountability. It is seen as a means to facilitate accountability, preventing blind reliance on AI and ensuring that decisions can be audited and overturned if needed. In an educational context, this aligns perfectly with the ethos that teachers (and, where appropriate, students) should have the final say in teaching and assessment decisions where AI can assist, but not autonomously control outcomes without recourse.

To summarize, XAI and CAI are key pillars for integrating generative AI in a way that faculty can trust and leverage effectively. An AI system that can explain its outputs and accept human feedback aligns with educational values of transparency, critical inquiry, and continuous improvement. These principles set the foundation for the framework we propose, which aims to enhance faculty competencies and confidence in working with generative AI by embedding explainability and contestability into both technology and practice.

3. PROPOSED FRAMEWORK

To address the faculty competency gap and guide education institutions toward a more responsible

and pedagogically sound integration of generative AI, this paper proposes the **TECTRA (Trust through Explainability, Contestability, and Reflective Application)** framework, which is a comprehensive, actionable model designed to cultivate the specific skills and dispositions faculty require to build and sustain a trustworthy AI ecosystem.

3.1. Theoretical Grounding

The TECTRA framework was developed through an iterative, literature-informed synthesis that aimed to identify stable and transferable faculty competencies for responsible generative AI adoption in higher education.

The theoretical grounding of the framework draws on trustworthy AI principles, which emphasize transparency, accountability, and human agency, as well as established educational technology models, particularly Technological Pedagogical Content Knowledge (TPACK) [48]. The TECTRA framework is constructed upon four interdependent pillars, which together form a comprehensive approach to faculty competency in education. These pillars are designed to be mutually reinforcing, ensuring that technical skills are always linked to pedagogical purpose and ethical considerations. Each pillar is strengthened by the integration of explainability and contestability mechanisms. As visualized in Figure 2, the TECTRA framework can be conceptualized as an integrated ecosystem focused on creating trustworthy AI in education. Rather than positioning XAI and contestability as external mechanisms, the framework recognizes them as intrinsic to each pillar’s function. (1) Ethical Grounding is made actionable through XAI’s revelatory power; (2) Pedagogical Integration is activated through the dialogic structure of contestability; (3) Technical Literacy is developed through XAI’s explanations; and (4) Reflective Practice is sustained through the combined feedback loops created by both mechanisms.

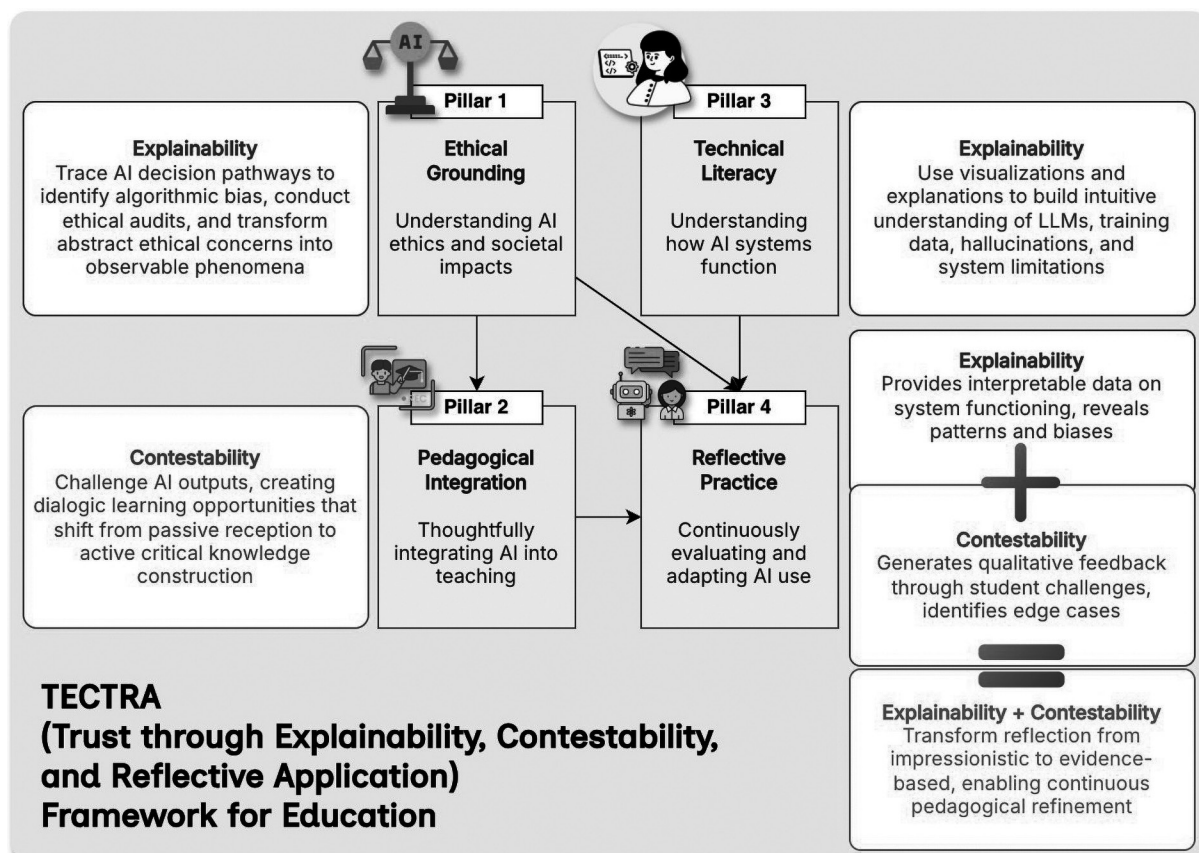


Figure 2. Overview of TECTRA (Trust through Explainability, Contestability, and Reflective Application) Framework for Higher Education.

3.2. Core Principles

3.2.1. Pillar 1: Ethical Grounding (Enabled by Explainability)

This foundational pillar moves the conversation about AI ethics beyond the narrow confines of plagiarism and academic misconduct. It requires faculty to develop a robust understanding of the broader ethical landscape of AI, including the ability to critically analyze issues of algorithmic bias, data privacy, intellectual property, and the potential for AI to perpetuate societal inequities.

The Role of Explainability: XAI serves as the primary mechanism through which faculty can develop and demonstrate ethical competency. When AI systems make their decision-making processes transparent, faculty gain the ability to conduct meaningful ethical audits of the technology they employ. Through XAI dashboards and interpretation tools, faculty can trace how an AI system arrives

at specific outputs, identifying the data points, features, or patterns that influence results. This visibility transforms abstract ethical concerns into concrete, observable phenomena that can be analyzed, discussed, and addressed. Competency in this pillar means faculty can not only articulate ethical risks but can also use XAI tools to actively investigate and demonstrate them. For instance, a faculty member might use XAI techniques to reveal how an automated essay grading system disproportionately penalizes certain writing styles or linguistic patterns associated with multilingual learners. This capability enables faculty to design learning environments and institutional policies that proactively mitigate harm, ensuring that the use of AI aligns with the core values of fairness, accountability, and transparency. Furthermore, XAI empowers faculty to transform ethical instruction into active learning. Rather than lecturing about

algorithmic bias in the abstract, faculty can use XAI as a pedagogical tool, demonstrating in real-time how bias manifests, helping students develop critical evaluation skills, and fostering a more ethically informed user base. This approach creates a culture of ethical vigilance that extends beyond compliance to genuine understanding and advocacy.

3.2.2. Pillar 2: Pedagogical Integration (Enabled by Contestability)

This pillar addresses the practical application of AI in teaching and learning. Drawing inspiration from the TPACK model [48], it emphasizes the crucial intersection of technological knowledge, pedagogical strategy, and subject-matter expertise. Competency here is not merely about using an AI tool but about thoughtfully integrating it into the curriculum to enhance learning outcomes through designs that promote active engagement, critical thinking, and student agency. Pedagogical integration involves redesigning assessments to prioritize process and critical thinking over simple content generation, creating AI-scaffolded activities that support rather than supplant student effort, and leveraging contestable AI systems to create more personalized and adaptive learning experiences. The ability to implement contestable frameworks requires faculty to ask not merely “*What can this tool do?*” but “*How can this tool create productive cognitive friction and dialogic opportunities that help my students achieve our learning objectives more effectively and deeply?*”

The Role of Contestability: Contestability serves as a natural mechanism for formative assessment. When students challenge AI outputs, they reveal their understanding, misconceptions, and the reasoning processes behind their responses. Faculty can use these contestation patterns as diagnostic data, identifying where additional instruction or support is needed and adapting their teaching accordingly. CAI transforms pedagogical integration from passive tool adoption to active, dialogic learning.

When students are empowered to challenge, question, and debate AI-generated outputs (e.g., grades, feedback, content recommendations, or analytical interpretations), the learning dynamic fundamentally shifts from information consumption to critical knowledge construction. Faculty who integrate CAI into their pedagogy create structured opportunities for metacognitive development. Consider an AI-assisted writing tutor that provides feedback on student essays: in a traditional implementation, students might passively accept the AI's suggestions. However, when the system is designed to be contestable, students must engage in reasoned argumentation to challenge feedback they believe is inappropriate. They must gather evidence from their work, present their reasoning, and engage in structured dialogue with both the AI system and their instructor. This process inherently develops higher-order thinking skills (e.g., analysis, evaluation, and creation), positioning AI as a sophisticated learning partner.

3.2.3. Pillar 3: Technical Literacy (Enabled by Explainability)

While deep technical expertise is not required, a foundational understanding of how generative AI systems work is essential for responsible use. This pillar aims to equip faculty with a conceptual understanding of core AI principles, including the nature of LLMs, the role of training data, the statistical foundations of AI outputs, and the inherent limitations of the technology.

The Role of Explainability: XAI serves as both a teaching tool and a competency-building mechanism for technical literacy. Rather than requiring faculty to understand complex machine learning mathematics, XAI systems provide accessible visualizations and explanations that interpret AI behavior. Through interaction with XAI tools, faculty develop an intuitive understanding of how models process information, why they produce certain outputs, and where their limitations lie. A technically literate faculty member, supported by XAI, understands why AI models hallucinate,

recognizes the statistical and probabilistic nature of their outputs, can identify when a model is operating outside its training domain, and can explain these concepts to students using concrete examples drawn from XAI explanations. For instance, when an LLM produces a factual error, XAI tools can help faculty trace the error to limitations in training data, revealing that the model is generating plausible-sounding but unverified content based on statistical patterns rather than verified knowledge. This literacy extends to understanding the source and characteristics of training data. XAI techniques that reveal which aspects of training data most influence specific outputs help faculty recognize potential blind spots, biases, or domain limitations in AI systems. This knowledge prevents both uncritical acceptance and unfounded fear, forming the basis for a more nuanced and effective pedagogical approach. Moreover, XAI supports the development of technical literacy in faculty through guided exploration. Faculty development programs can utilize XAI dashboards as learning environments where educators experiment with different inputs, observe how the AI's explanations evolve, and develop mental models of system behavior. This hands-on, explanation-guided approach to technical literacy is more accessible and pedagogically effective than traditional technical training, making AI competency achievable for faculty across all disciplines.

3.2.4. Pillar 4: Reflective Practice (Enabled by Explainability & Contestability)

This final pillar encourages faculty to adopt a critical, evidence-based, and iterative approach to their use of AI. Competency in this area involves continuously evaluating the impact of AI tools on their own teaching workflows, as well as on student learning and engagement. This pillar also establishes the practice of systematically gathering feedback through both XAI analytics and contestation records, reflecting on successes and failures revealed through transparent system behavior, and adapting pedagogical strategies in response to evidence.

The Role of Explainability and Contestability: Both XAI and CAI function as structured systems for generating the evidence and insights necessary for meaningful reflective practice. XAI provides faculty with interpretable data on how AI systems are functioning in their courses, revealing patterns in automated feedback, highlighting which students are receiving what types of AI support, and making visible any systematic biases or limitations in AI-mediated interactions. This transparency transforms reflection from subjective impression to data-informed inquiry. Meanwhile, CAI creates natural feedback loops that drive reflection. When students challenge AI outputs, they generate rich qualitative data about the AI system's performance, revealing edge cases, misconceptions, and areas where the AI may be falling short of pedagogical goals. Faculty who systematically analyze these contestations, tracking which AI outputs are most frequently challenged, what types of student arguments are most compelling, and how challenges correlate with learning outcomes, gain actionable insights for refining their AI integration strategies.

For example, a faculty member might notice through XAI analysis that an AI tutoring system consistently provides less detailed explanations to students who initially struggle with a concept. This insight, combined with student contestations arguing that the feedback is insufficient, prompts the instructor to reconfigure the system or supplement it with additional human support. The integration of XAI and contestability ensures that reflection is not a solitary, impressionistic activity but a collaborative, evidence-based practice. It positions reflective practice as an ongoing process of inquiry and improvement, keeping pedagogical practice aligned with a rapidly evolving technological landscape. Faculty members become action researchers in their own classrooms, using the transparency and dialogic nature of well-designed AI systems to continuously refine their practice.

3.3. Defining Faculty Competencies

The TECTRA framework is operationalized through a set of specific, measurable faculty competencies, which provide a clear roadmap for professional development. Table 1 maps these core competencies to the four pillars of the framework, now explicitly incorporating XAI and CAI mechanisms, and suggesting development activities and tools that can be used to support them. This structure translates the abstract principles of the framework into a concrete guide for designing effective faculty training programs, drawing upon competency models proposed by leading educational organizations. Our integrated approach for TECTRA ensures that faculty development is not fragmented into disconnected technical skills, but rather unified around the central goal of building trust in AI systems, pedagogical decisions, and in students’

capacity to engage critically with technology. The ultimate outcome of this ecosystem is the cultivation of enhanced faculty competencies that are adaptive, evidence-based, and ethically grounded, which in turn foster greater student agency in an increasingly generative AI educational landscape.

4. IMPLEMENTATION GUIDELINES AND POLICY RECOMMENDATIONS

The transformation from a theoretical model to institutional practice requires a deliberate and structured approach. The TECTRA framework, while providing the conceptual architecture, must be supported by a practical roadmap that guides education institutions through the complex process of implementation and policy development. This section outlines a phased strategy for integrating the TECTRA framework

Table 2. Core Faculty Competencies within the TECTRA Framework.

Pillar	Core Competency	Description	Development Activities & Tools
Ethical Grounding	Critical AI Evaluation	The ability to systematically assess AI-generated outputs for accuracy, veracity, and potential algorithmic bias, and to understand the broader societal and ethical implications of AI use in academia.	Use XAI tools in a sandbox environment to deconstruct biased outputs and understand their origins.
	Ethical AI Pedagogy	The ability to design and articulate clear course policies, assignments, and learning environments that promote the responsible, equitable, and transparent use of AI tools by students.	+ Develop departmental AI usage and citation style guides; engage in case study analysis of complex ethical dilemmas in AI-assisted education. + Craft adaptable syllabus statements.
Pedagogical Integration	AI-Informed Curriculum Design	The ability to strategically redesign learning objectives, activities, and assessments to leverage AI for fostering higher-order thinking skills, rather than allowing AI to circumvent them.	+ Receive training on advanced prompt engineering for educational purposes. + Use generative AI to create diverse and complex case studies or problem sets. + Redesign assessments to focus on process, reflection, and application.
	Adaptive Teaching with AI	The ability to use AI-driven analytics and tools to personalize learning pathways, provide timely and targeted feedback, and offer differentiated support to students based on their individual needs.	Use XAI-enhanced Learning Management System (LMS) dashboards to interpret student performance predictions and identify areas for intervention.

Pillar	Core Competency	Description	Development Activities & Tools
Technical Literacy	Foundational AI Principles	A conceptual understanding of the basics of LLMs, the importance of training data, and the technical reasons for inherent limitations like hallucinations and bias.	Engage in hands-on, guided experimentation with a variety of generative AI tools to understand their capabilities and failure modes.
	XAI Interpretation	The ability to read and interpret the explanations provided by XAI systems to understand a model's behavior, identify key influencing factors, and assess the reliability of its outputs.	Participate in training sessions on interpreting outputs from specific XAI techniques, applying them to educational datasets.
Reflective Practice	AI-Mediated Critical Dialogue	The ability to facilitate and model critical, dialogic interactions with AI systems, encouraging students to question, challenge, and verify AI-generated information.	+ Implement assignments that use CAI frameworks for automated feedback. + Require students to maintain structured reflection journals on their AI usage and findings.
	Continuous Self-Improvement	A commitment to ongoing professional development, staying current with the rapid advancements in AI technology and pedagogy, and actively participating in the institutional conversation around AI.	+ Engage in interdisciplinary faculty learning communities and institutional forums on AI. + Contribute to the work of institutional AI task forces.

into the university’s fabric, providing concrete recommendations for crafting agile, ethical, and enabling AI policies that support rather than hinder innovation, as visualized in Figure 3.

4.1. A Phased Strategy Roadmap

A successful institution-wide integration of trustworthy AI practices requires a strategic, multi-phase approach that builds momentum, secures stakeholder buy-in, and allows for iterative learning and adaptation.

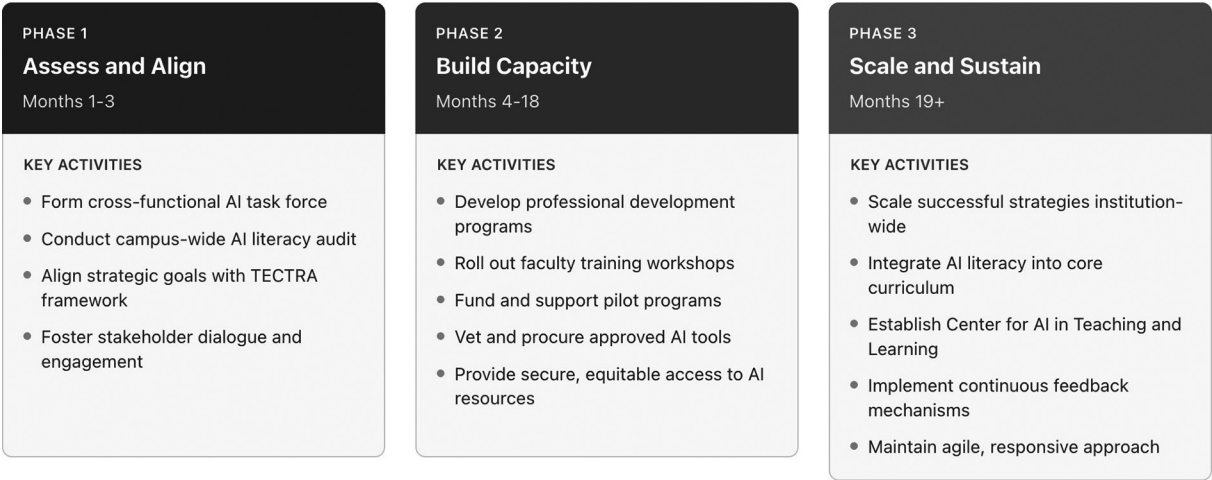
4.1.1. Phase 1: Assess and Align

The foundational phase is dedicated to establishing a shared understanding and a common vision. The first step is to form a cross-functional AI task force comprising administrators, faculty from diverse disciplines, instructional designers, IT professionals, legal experts, and, crucially, students. This group’s initial mandate is to conduct a comprehensive,

campus-wide AI literacy audit to gauge the current knowledge, practices, and attitudes of both faculty and students. Concurrently, institutional leaders must work with this task force to align the university’s strategic goals for AI with the core principles of the TECTRA framework. This phase is not about deploying technology but about fostering dialogue, engaging stakeholders, and collaboratively defining what a successful, human-centric AI integration will look like for the institution.

4.1.2. Phase 2: Build Capacity

The second phase focuses on building the necessary human and technical infrastructure. The institution should develop and roll out a portfolio of targeted professional development programs, workshops, and resources explicitly designed to cultivate the faculty competencies outlined in the TECTRA framework (see Table 2). To move from theory to practice,



**Note: This roadmap emphasizes iterative learning and stakeholder engagement throughout all phases. Institutions should adapt timelines based on their specific context, resources, and existing AI maturity level. The transition from Phase 2 to Phase 3 should be gradual, with ongoing evaluation and refinement of practices.*

Figure 3. Phased Strategy Roadmap for Integrating the TECTRA Framework into the Higher Education Institutions' Fabric.

the university should fund and support pilot programs in various departments, encouraging faculty to experiment with innovative generative AI integrations and assessment redesigns in a low-stakes environment. During this phase, it is also crucial for the institution to vet, procure, and provide secure and equitable access to a suite of approved AI tools, ensuring that faculty and students have a safe and supported environment for exploration.

4.1.3. Phase 3: Scale and Sustain

The final phase transitions from pilot programs to systemic integration. Successful strategies and resources developed during the capacity-building phase should be scaled across the institution. A key element of this phase is the formal integration of AI literacy and ethical AI use into the general education or core curriculum for all students, ensuring that graduates are prepared for an AI-driven world. To sustain this effort, institutions should establish permanent support structures, such as a Center for AI in Teaching and Learning, to provide ongoing consultation, training, and resources. Finally, a continuous feedback loop must be established, allowing the AI task force to regularly gather input from faculty and students

on the effectiveness of policies and support services, ensuring that the institution's approach remains agile and responsive to the rapid pace of technological change.

4.2. Agile and Ethical AI Policies

Effective institutional AI policy and robust faculty competency are not separate initiatives. They are deeply interconnected and mutually reinforcing. A well-crafted policy provides the necessary guidance and security for faculty to innovate, while competent faculty are essential for effectively implementing any policy. An enabling policy cannot be enacted by an unprepared faculty, and faculty cannot develop competency in a policy vacuum. Therefore, policy development must proceed together with capacity building, guided by the following principles:

Define, Not Just Prohibit. Any effective AI policy must begin with a clear and comprehensive definition of generative AI and related terminology, ensuring that all stakeholders are operating from a shared understanding. Rather than resorting to blanket prohibitions, which are often unenforceable and

pedagogically counterproductive, policies should provide nuanced guidelines on permissible and impermissible uses. These guidelines should be flexible and adaptable to the diverse contexts of different academic disciplines.

Mandate Transparency and Citation. To uphold academic integrity in an AI-augmented environment, policies must establish clear and unambiguous standards for transparency and accountability. Students should be required to disclose their use of generative AI tools in all academic work. The policy should provide specific instructions for how to cite and acknowledge this use, drawing from emerging standards in various disciplinary style guides.

Prioritize Human Oversight. Policies must unequivocally affirm that faculty and students are ultimately responsible and accountable for all academic work and educational outcomes. This principle should be operationalized by ensuring that faculty maintain meaningful oversight of any AI-driven assessment or feedback processes. Furthermore, in alignment with the principle of contestability, policies must establish clear and accessible mechanisms for students to appeal or seek review of AI-generated decisions, reinforcing human agency within the system.

Embed Ethical Principles. A robust AI policy must extend beyond academic integrity to explicitly address the core ethical principles of responsible AI. This includes strong provisions for protecting student data privacy and security, particularly when using third-party tools. The policy should also prohibit the use of AI to generate biased, discriminatory, or harmful content and should align with internationally recognized frameworks for trustworthy AI, such as the OECD AI Principles.

Ensure Flexibility and Iteration. Given the breathtaking pace of AI development, any policy written currently will be outdated tomorrow. Therefore, AI policies must be designed as living documents, not as static

regulations set in stone. The institutional AI task force should be charged with conducting regular, periodic reviews of all AI-related policies, ensuring they remain relevant, effective, and aligned with both technological advancements and the evolving pedagogical needs of the university community. This commitment to agility is paramount for navigating the future of AI in education successfully.

5. CONCLUSION

This paper identifies a critical juncture in education, where the asynchronous adoption of generative AI has created a pedagogical crisis rooted in faculty competency gaps and compounded by the risks of AI's opacity, bias, and misinformation. In response, the TECTRA framework provides a proactive, human-centric solution, grounded in Ethical Grounding, Pedagogical Integration, Technical Literacy, and Reflective Practice, positioning XAI and CAI as pivotal mechanisms to transform AI from an opaque tool into a transparent partner for critical inquiry. Beyond mere integration, this technological moment presents an opportunity to fundamentally reimagine education as a more equitable, engaging, and effective ecosystem where AI augments rather than replaces human intellect, creativity, and critical thinking. The TECTRA framework, with its emphasis on human agency and ethical oversight, serves as a foundational step toward this vision, making faculty empowerment the most vital investment institutions can make in securing a human-centric future for education.

This work, however, has the following limitations that should be acknowledged clearly. First, TECTRA is proposed as a conceptual and design framework rather than as a model that has been empirically validated through controlled studies. Its effectiveness will depend on institutional context, disciplinary norms, faculty workload, and the availability of appropriate AI tools and support structures. Second, the framework assumes at least minimal access to

explainability and contestability mechanisms, which may be difficult to implement in institutions with limited financial, technical, or infrastructural resources. In such contexts, faculty may rely on general-purpose AI systems that offer limited transparency or formal avenues for contestation, increasing the burden on human oversight. Third, there is a risk that explanations provided by AI systems may create a false sense of confidence if they are interpreted as guarantees of correctness rather than as aids for critical judgment, while contestation processes may lose educational value if they are treated as procedural formalities without meaningful authority to override AI outputs.

Future work should therefore focus on systematic empirical evaluation and contextual adaptation of the TECTRA framework. Experimental and quasi-experimental studies are necessary to investigate whether TECTRA-aligned faculty development enhances calibrated trust, ethical reasoning, and the quality of AI-integrated pedagogy compared to tool-oriented training approaches. Design-based research across disciplines can explore how the four pillars function in different instructional settings and identify which elements require domain-specific modification. Further research should also develop validated measurement instruments to assess faculty competencies associated with each pillar, combining self-report data with performance-based indicators such as explanation analysis, contestation patterns, and reflective documentation. In addition, equity-focused studies are needed to investigate whether TECTRA can reduce the unintended harms of generative AI, particularly for multilingual learners and students with unequal access to AI resources. Finally, future research should investigate institutional governance and policy models that facilitate the scalable and sustainable implementation of trustworthy AI practices, particularly in resource-constrained environments. By addressing these directions, subsequent work can strengthen the empirical

foundation of TECTRA and support its evolution into an evidence-based model for faculty development in higher education.

DECLARATION OF CONFLICTS OF INTEREST

The authors declare that this research has no conflicts of interest.

REFERENCES

- [1] G. G. Wilkinson, "Enhancing Generic Skills Development in Higher Education in the Era of Large Language Model Artificial Intelligence," *Journal of Higher Education Theory & Practice*, vol. 24, no. 3, pp. 64-76, 2024.
- [2] X. Wang, N. D. Trung, R. R. Bhoyar, and A. M. Jose, "LLM-based literature recommender system in higher education: A case study of supervising students' term papers," *Computers & Education: Artificial Intelligence*, vol. 10, pp. 32-41, 2025.
- [3] H. Nguyen, T. D. Q. Nguyen, H. L. Cao, T. C. T. Tran, T. C. M. Truong, and H. Cao, "SimInterview: Transforming Business Education through Large Language Model-Based Simulated Multilingual Interview Training System," in *Proceedings of International Conference on Economics, Finance, and Management (ICEFM)*, 2025, pp. 1-19.
- [4] H. Nguyen et al., "Heart2Mind: Human-Centered Contestable Psychiatric Disorder Prediction System Using Wearable ECG Monitors," *ACM Transactions on Computing for Healthcare*, vol. 37, no. 4, p. 111, 2026.
- [5] R. Manche and P. K. Myakala, "Explaining black-box behavior in large language models," *International Journal of Computing and Artificial Intelligence*, vol. 3, no. 2, pp. 102-108, 2022.
- [6] H. T. Ho, D. T. Ly, and L. V. Nguyen, "Mitigating Hallucinations in Large Language Models for Educational Application," in *2024 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia 2024)*, 2024, pp. 1-4.

- [7] H. Zhao et al., "Explainability for large language models: A survey," *ACM Transactions on Intelligent Systems and Technology*, vol. 15, no. 2, pp. 1–38, 2024.
- [8] H. Nguyen, L. Nguyen, and H. Cao, "XEdgeAI: A human-centered industrial inspection framework with data-centric Explainable Edge AI approach," *Information Fusion*, vol. 116, p. 102782, 2025.
- [9] A. Chattopadhyay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks," in 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), 2018, pp. 839–847.
- [10] W. Xu, M. Shahreeza, W. L. Hoo, and W. Yang, "Explainable AI for Education: Enhancing Essay Scoring via Rubric-Aligned Chain-of-Thought Prompting," *International Journal of Modern Physics C*, vol. 36, p. 2542013, 2025.
- [11] L. P. T. Nguyen, H. T. Do, H. T. T. Nguyen, and H. Cao, "Motion2Meaning: A Clinician-Centered Framework for Contestable LLM in Parkinson's Disease Gait Interpretation," in 9th International Symposium on Chatbots and Human-centred AI, CONVERSATIONS 2025, 2025.
- [12] G. Freedman, A. Dejl, D. Gorur, X. Yin, A. Rago, and F. Toni, "Argumentative Large Language Models for Explainable and Contestable Claim Verification," in Proceedings of the 39th AAAI Conference on Artificial Intelligence, AAAI 2025, 2025, pp. 14930-14939.
- [13] K. Alfrink, I. Keller, G. Kortuem, and N. Doorn, "Contestable AI by design: Towards a framework," *Minds and Machines*, vol. 33, no. 4, pp. 613-639, 2023.
- [14] T. Ploug and S. Holm, "The four dimensions of contestable AI diagnostics-A patient-centric approach to explainable AI," *Artificial Intelligence in Medicine*, vol. 107, p. 101901, 2020.
- [15] S. Hong, C. Cai, S. Du, H. Feng, S. Liu, and X. Fan, "'My Grade is Wrong!': A Contestable AI Framework for Interactive Feedback in Evaluating Student Essays," in International Conference on Artificial Intelligence in Education, 2025, pp. 27-35.
- [16] H. Thai et al., "Educational Technology and Responsible Automated Essay Scoring in the Generative AI Era," in Navigating the Circular Age of a Sustainable Digital Revolution, IGI Global, 2024, pp. 325-356.
- [17] K. S. Geethanjali and N. Umashankar, "Enhancing Educational Outcomes with Explainable AI: Bridging Transparency and Trust in Learning Systems," in 2025 International Conference on Emerging Systems and Intelligent Computing (ESIC), 2025, pp. 325-328.
- [18] M. Tornqvist, M. Mahamud, E. M. Guzman, and A. Farazouli, "ExASAG: Explainable framework for automatic short answer grading," in Proceedings of the 18th workshop on innovative use of NLP for building educational applications (BEA 2023), 2023, pp. 361-371.
- [19] C. Anghel et al., "CourseEvalAI: Rubric-Guided Framework for Transparent and Consistent Evaluation of Large Language Models," *Computers*, vol. 14, no. 10, p. 431, 2025.
- [20] Z. Liu, P. Agrawal, S. Singhal, V. Madaan, M. Kumar, and P. K. Verma, "LPITutor: an LLM based personalized intelligent tutoring system using RAG and prompt engineering," *PeerJ - Computer Science Journal*, vol. 11, p. e2991, 2025.
- [21] G. M. Sekli, A. Godo, and J. C. Véliz, "Generative AI solutions for faculty and students: A review of literature and roadmap for future research," *Journal of Information Technology Education: Research*, vol. 23, p. 14, 2024.
- [22] N. Hashmi and A. S. Bal, "Generative AI in higher education and beyond," *Business Horizons*, vol. 67, no. 5, pp. 607-614, 2024.
- [23] D. S. A. Cufuna, J. M. Duarte, and G. Rangel-de Lazaro, "Augmented reality in higher education: Interactions in LLM-based teaching and learning," in The Learning Ideas Conference (TLIC 2024), 2024, pp. 105-114.

- [24] C. K. Y. Chan and W. Hu, "Students' voices on generative AI: Perceptions, benefits, and challenges in higher education," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, p. 43, 2023.
- [25] E. Dickey and A. Bejarano, "Gaide: A framework for using generative ai to assist in course content development," in 2024 IEEE Frontiers in Education Conference (FIE 2024), 2024, pp. 1-9.
- [26] T. Camarata, L. McCoy, R. Rosenberg, K. R. T. Grellinger, K. Brettschnieder, and J. Berman, "LLM-Generated multiple choice practice quizzes for preclinical medical students," *Advances in Physiology Education*, vol. 49, no. 3, pp. 758-763, 2025.
- [27] S. Haroud and N. Saqri, "Generative AI in higher education: Teachers' and students' perspectives on support, replacement, and digital literacy," *Education Sciences*, vol. 15, no. 4, p. 396, 2025.
- [28] Y. Zhong and K. Zhao, "Application and Research of Large Language Model in Foreign Language Translation," in 2024 International Conference on Information Technology, Communication Ecosystem and Management (ITCEM 2024), 2024, pp. 63-68.
- [29] C. L. Chen et al., "A systematic multimodal assessment of AI machine translation tools for enhancing access to critical care education internationally," *BMC Medical Education*, vol. 25, no. 1, p. 1022, 2025.
- [30] R. Howe, L. Machado, and S. Sneddon, "The Ethical Implications of Generative Artificial Intelligence on Students, Academic Staff, and Researchers in Higher Education," in Artificial Intelligence Applications in Higher Education, 2024, pp. 33-51.
- [31] Q. Zhang, S. B. Siraj, and R. B. Abdul Razak, "Effects of AI chatbots on EFL students' critical thinking skills and intrinsic motivation in argumentative writing," *Innovation in Language Learning and Teaching*, vol. 1, pp. 1-29, 2025.
- [32] A. J. Pereira, L. M. Queiros, A. S. Gomes, and T. T. Primo, "Strategies of Intelligent Tutoring Systems to Engage Students in Online Learning Before LLM Approaches," in Simpósio Brasileiro de Sistemas de Informação (SBSI 2025), 2025, pp. 290-299.
- [33] H. Nguyen et al., "LangXAI: integrating large vision models for generating textual explanations to enhance explainability in visual perception tasks," in Proceedings of the Thirty-Third International Joint Conference on Artificial Intelligence, (IJCAI 2024), 2024.
- [34] O. A. Arise, M. Muzuva, R. Kader, and F. H. Chohan, "Ethical Concerns of Artificial Intelligence in Student Assessments from a Higher Education Perspective," in The Focus Conference (TFC 2024), 2024, pp. 234-254.
- [35] Z. N. Khlaif et al., "University teachers' views on the adoption and integration of generative AI tools for student assessment in higher education," *Education Sciences*, vol. 14, no. 10, p. 1090, 2024.
- [36] J. Waldo and S. Boussard, "GPTs and hallucination: why do large language models hallucinate?," *Queue*, vol. 22, no. 4, pp. 19-33, 2024.
- [37] H. Elsayed, "The impact of hallucinated information in large language models on student learning outcomes: A critical examination of misinformation risks in AI-assisted education," *Northern Reviews on Algorithmic Research, Theoretical Computation, and Complexity*, vol. 9, no. 8, pp. 11-23, 2024.
- [38] W. H. Walters and E. I. Wilder, "Fabrication and errors in the bibliographic citations generated by ChatGPT," *Scientific Reports*, vol. 13, no. 1, p. 14045, 2023.
- [39] K. Nguyen, H. Nguyen, K. Nguyen, B. Truong, T. Phan, and H. Cao, "Efficient and Concise Explanations for Object Detection with Gaussian-Class Activation Mapping Explainer," in Proceedings of the 37th Canadian Conference on Artificial Intelligence, Canadian AI 2024, 2024.
- [40] V. B. Truong, H. Nguyen, V. T. K. Nguyen, Q. K. Nguyen, and Q. H. Cao, "Towards Better Explanations for Object Detection," in Proceedings of the 15th Asian Conference on Machine Learning, ACML 2023, vol. 222, 2024, pp. 1385-1400.

- [41] V. Petsiuk et al., “Black-box explanation of object detectors via saliency maps,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 11443-11452.
- [42] M. T. Ribeiro, S. Singh, and C. Guestrin, ““ Why should I trust you?” Explaining the predictions of any classifier,” in Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, pp. 1135-1144.
- [43] S. M. Lundberg and S. I. Lee, “A unified approach to interpreting model predictions,” in Proceedings of the 31st International Conference on Neural Information Processing Systems, 2017, pp. 4768-4777.
- [44] R. K. Mothilal, A. Sharma, and C. Tan, “Explaining machine learning classifiers through diverse counterfactual explanations,” in Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, FAT 2020, 2020, pp. 607-617.
- [45] R. Poyiadzi, K. Sokol, R. Santos-Rodriguez, T. De Bie, and P. Flach, “FACE: feasible and actionable counterfactual explanations,” in Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, AIES’20, 2020, pp. 344-350.
- [46] H. Nguyen et al., “Human-Centered Explainable Psychiatric Disorder Diagnosis System Using Wearable ECG Monitors,” in Proceedings of the 29th Pacific-Asia Conference on Knowledge Discovery and Data Mining, PAKDD 2025, 2025, pp. 418-429.
- [47] A. A. Tubella, A. Theodorou, V. Dignum, and L. Michael, “Contestable black boxes,” in International Joint Conference on Rules and Reasoning, 2020, pp. 159-167.
- [48] M. Koehler and P. Mishra, “What is technological pedagogical content knowledge (TPACK)?,” *Contemporary Issues in Technology and Teacher Education*, vol. 9, no. 1, pp. 60-70, 2009.



© 2026 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Nâng cao năng lực giảng viên giáo dục đại học trong kỷ nguyên trí tuệ nhân tạo tạo sinh: Tích hợp khả năng giải thích, khả năng tranh luận và thực hành phản tư

Nguyễn Trần Diễm Quỳnh¹, Trương Thị Cẩm Mai^{2,*}

¹Trường Đại học Ngoại ngữ, Đại học Đà Nẵng, Việt Nam

²Khoa Khoa học Tự nhiên, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 03/11/2025; Ngày sửa bài: 22/12/2025;

Ngày nhận đăng: 25/12/2025; Ngày xuất bản 28/06/2026

TÓM TẮT

Sự xuất hiện nhanh chóng của trí tuệ nhân tạo tạo sinh (AI) và các mô hình ngôn ngữ lớn (LLMs) mang đến cả cơ hội lẫn thách thức cho giáo dục đại học. Dù công nghệ này hứa hẹn nâng cao hiệu quả giảng dạy và nghiên cứu, bản chất “hộp đen”, xu hướng “ảo giác” và thiếu minh bạch của chúng đặt ra vấn đề về niềm tin, trách nhiệm và tính toàn vẹn sự phạm. Bài viết đề xuất nhu cầu cấp thiết về một khung năng lực giúp giảng viên ứng dụng AI tạo sinh có trách nhiệm. Chúng tôi giới thiệu TECTRA (Trust through Explainability, Contestability, and Reflective Application), khung tiếp cận lấy con người làm trung tâm, kết hợp AI có khả năng giải thích (XAI) và AI có thể phản biện (CAI) như cơ chế nền tảng để xây dựng niềm tin trong giáo dục. Khung gồm bốn trụ cột: Đạo đức, dựa trên tính minh bạch của XAI; Tích hợp sự phạm, qua cấu trúc đối thoại của CAI; Hiểu biết kỹ thuật, thông qua khả năng giải thích của XAI; và Thực hành phản tư, duy trì bằng phản hồi của hai hệ thống. Bài viết xác định năng lực, hoạt động và công cụ phát triển cho từng trụ cột, cùng lộ trình ba giai đoạn: đánh giá, xây dựng năng lực và mở rộng bền vững. Ngoài ra, khuyến nghị chính sách nhấn mạnh sự linh hoạt, minh bạch, giám sát con người và đạo đức. Bằng cách xem XAI và CAI là yếu tố tương tác thực tiễn, TECTRA đưa AI trở thành đối tác minh bạch, phản biện, thúc đẩy năng lực giảng viên thích ứng, dựa trên bằng chứng và có nền tảng đạo đức.

Từ khóa: AI có thể giải thích, AI có thể tranh luận, AI tạo sinh, mô hình ngôn ngữ lớn, giáo dục đại học.

*Tác giả liên hệ chính.

Email: truongcammai@qnu.edu.vn

Thermal analysis of low-voltage power cables considering the presence of current harmonics

Kim Anh Nguyen*, Dai Huynh, Thi Thu Le, Huy Hoang Le, Ba Tu Phung

*Faculty of Electrical Engineering, The University of Danang -
University of Science and Technology, Vietnam*

Received: 27/11/2025; Revised: 09/02/2026;

Accepted: 12/02/2026; Published: 28/06/2026

ABSTRACT

In practical design and selection of low-voltage power cables, the heating effects caused by current harmonics are often not sufficiently considered. In addition, with the increasing penetration of nonlinear loads and renewable energy sources in low-voltage distribution networks, harmonic current components tend to increase the AC resistance and Joule losses, thereby leading to a rise in cable operating temperature depending on the harmonic magnitude and order. When harmonics are present, the increased conductor losses may cause the actual cable core temperature to significantly exceed that observed under purely fundamental current conditions. In this study, a thermal model originally developed for underground cables is improved and its heat exchange parameters are recalibrated to accurately represent the thermal dissipation characteristics of overhead PVC-insulated low-voltage cables, which are widely used in Vietnamese distribution networks. By employing harmonic current spectra together with the electro-thermal parameters of the cable as input data, the proposed model enables the evaluation of cable heating behavior under various operating scenarios, while simultaneously considering variations in both harmonic magnitude and harmonic order. The study focuses on: (i) improving the thermal balance model for low-voltage cables; (ii) evaluating the impact of current harmonics on cable core temperature; (iii) analyzing insulation lifetime degradation under harmonic conditions; and (iv) assessing the influence of ambient temperature. Based on the numerical results, appropriate operating current derating levels are recommended to maintain cable temperature within permissible limits, thereby ensuring safe operating conditions and reducing the risk of insulation degradation in low-voltage power networks.

Keywords: *Low-voltage power cables, current harmonics, thermal characteristics, operating limits, cable lifetime estimation.*

1. INTRODUCTION

Low-voltage (LV) power cables are a critical component of distribution networks, forming the essential connection between substations and consumers in both residential and industrial sectors. Safe operation of these systems can be maintained only if conductor temperatures are kept within the allowable range. Prolonged thermal stress can, in turn, degrade operational

safety and reduce the expected service life of the cables. When such conditions persist, insulation deteriorates more quickly, the lifetime of materials decreases, and faults including phase-to-phase short circuits or localized arc discharges may arise. These failures not only interrupt electricity supply but also cause considerable economic losses. For this reason, considerable research has been devoted to modeling cable

**Corresponding author:*

Email: nkanh@dut.udn.vn

thermal behavior with greater accuracy, focusing on aspects such as temperature distribution and heat dissipation, which are critical for the design, operation, and maintenance of power networks [1]–[6].

In recent years, numerous studies have been directed towards enhancing the accuracy of thermal modeling for cables by considering real installation conditions that are more complicated than those assumed in standards. For instance, some research has examined the impact of air gaps in the cable structure [7], [8], or in real installations where conduits are not entirely filled [9], showing that such aspects can notably augment the overall thermal resistance and lead to elevated operating temperatures compared to those predicted. In addition, some studies reevaluated the application of IEC standard formulas in unbalanced conditions [10] and in cable trenches [11] with the aim of developing and suggesting more precise analytical models.

A further issue in modern grids concerns the impact of harmonics, which originate from the growing use of nonlinear loads and power electronic devices. Harmonics distort current waveforms, produce additional losses, and can raise conductor temperatures beyond the endurance of insulation. Research on cable thermal modeling has therefore examined harmonic effects on heating and proposed alternative installation practices or derating factors to mitigate them. For example, [12] investigated the effect of harmonic currents produced by electric vehicle chargers on power systems. Other research calculated the temperature of low-voltage [13] and medium-voltage cables [14] by the thermal balance principle applied at every node of an equivalent thermal circuit. In the study by Ratheiser et al. [15], the authors studied the effect of harmonic voltages on electric field distribution and temperature in medium voltage DC (MVDC) cable systems. In addition, [16] presented a dynamic thermal model for LV and MV cables under harmonic loads, revealing a noticeable

decrease in ampacity and faster thermal response. Similarly, Alwan et al. [17] examined the thermal behavior of insulated cables in hot climates and found that harmonics could increase conductor temperature by up to 6°C and accelerate insulation aging, emphasizing the necessity of including harmonic-induced heating in ampacity assessments.

Most previous studies on the impact of harmonics have been limited to idealized current spectra or simplified load conditions, which do not accurately reflect real operating environments. To overcome this limitation, the present study investigates the thermal behavior of a low-voltage single-core Cadivi 0.6/1 kV cable arranged in a trefoil configuration and operating under free-air convection. Three groups of scenarios are considered. Case 1 analyzes the effect of individual harmonic orders while keeping the total RMS current constant, thereby assessing conductor temperature variation with harmonic order. Case 2 focuses on varying the total harmonic distortion (THD) of a fixed harmonic order to evaluate the influence of harmonic amplitude on temperature rise. Case 3 extends the analysis to a more practical scenario, in which the cable supplies current to a DC motor through a fully controlled bridge rectifier with variable firing angles, combining the effects of harmonics, load level, and ambient temperature. The combined results reveal that the conductor temperature depends not only on the RMS current magnitude or total harmonic distortion, but also on the harmonic order, spectral composition, and environmental conditions. These findings highlight the importance of incorporating harmonic spectral characteristics and nonlinear load conditions into thermal assessment and derating factor determination, as well as into accurate cable lifetime prediction, to ensure safe and reliable cable system operation.

The rest of this paper is structured as follows. Section 2 gives the thermal model used in this study. Section 3 outlines the cable parameters, installation conditions, and

harmonic scenarios employed in the calculations. The results are analyzed in Section 4, and the conclusion is made in Section 5.

2. THERMAL MODELLING OF SINGLE-CORE CABLES

In this section, harmonic losses are first calculated using the frequency-dependent AC resistance. These results are then applied in a thermal circuit model to determine the conductor temperature of the single-core cable.

2.1. Harmonic loss calculation

The alternative current flowing through a conductor produces Joule losses. In the presence of harmonics, the total conductor loss in a cable can be determined as:

$$W_{total} = \sum_{h \geq 1} I_h^2 R_{AC}(h) \quad (1)$$

where I_h is the RMS current of the h -th harmonic and $R_{AC}(h)$ is the AC resistance for harmonic order h .

The AC resistance increases with frequency due to skin and proximity effects. According to IEC 60287-1-1, the AC resistance for a given harmonic order can be expressed as:

$$R_{AC}(h) = R_{DC} [1 + Y_s(h) + Y_p(h)] \quad (2)$$

where R_{DC} is the DC resistance of the conductor at the permissible maximum operating temperature per unit length (Ω / m), $Y_s(h)$ and $Y_p(h)$ are the skin effect and proximity effect factors, respectively.

$$Y_s(h) = \frac{x_s^4}{192 + 0.8x_s^4} \quad (3)$$

$$x_s^2(h) = \frac{8\pi f(h)k_s 10^{-7}}{R_{DC}} \quad (4)$$

$$Y_p(h) = \frac{x_p^4}{192 + 0.8x_p^4} \left(\frac{d_c}{S} \right)^2 \left[0.312 \left(\frac{d_c}{S} \right)^2 + \frac{1.18}{\frac{x_p^4}{192 + 0.8x_p^4} + 0.27} \right] \quad (5)$$

$$x_p^2(h) = \frac{8\pi f(h)k_p 10^{-7}}{R_{DC}} \quad (6)$$

where k_s is the coefficient factor for skin effect, k_p is the coefficient factor proximity effect, d_c is the conductor diameter of the cable, S is the distance between core axes.

For low-voltage cables, dielectric losses are typically negligible, and since these cables do not include metallic sheaths or amours, associated losses do not occur. Consequently, harmonic loss calculations in LV cables focus solely on conductor losses arising from skin and proximity effects.

2.2. Cable thermal model

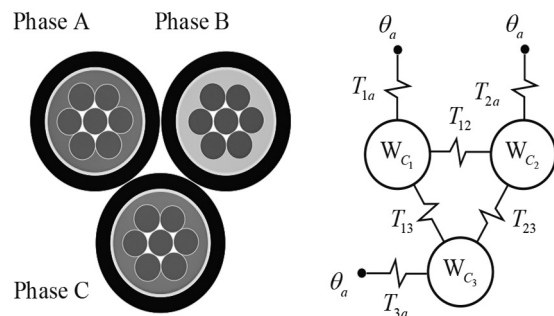


Figure 1. The equivalent thermal circuit of cable.

A generalized two-dimensional finite difference model (FDM) based on the principle of nodal heat balance in an equivalent thermal resistance network was developed in the study by Hiranandani [18]. The governing nodal heat balance equation can be expressed as:

$$W_i + \sum_{j \in \{1,2,3; j \neq i\}} \left(\frac{\theta_j - \theta_i}{T_{ji}} \right) = 0 \quad (7)$$

where W_i is defined as the summation of cable losses at node i , $\theta_j - \theta_i$ is the temperature difference between every two nodes and T_{ji} is the thermal resistance between every two nodes. The equivalent thermal circuit for a three-phase LV cable is illustrated in Figure 1.

According to Figure 1, there are three main heat flow paths for each conductor in the trefoil arrangement. As an example, the heat generated in phase A (W_{c1}) is dissipated: (i) directly to the ambient through the external thermal resistance (T_{1a}), (ii) to phase B through the mutual thermal resistance (T_{12}), and (iii) to

phase C through the mutual thermal resistance (T_{13}). The same principle applies to the other two conductors. By applying the nodal heat balance principle, the following equations can be written for each phase:

$$\theta_1 - \theta_a = W_{C_1} T_{1a} \quad (8)$$

$$\theta_1 - \theta_2 = W_{C_1} T_{12} \quad (9)$$

$$\theta_1 - \theta_3 = W_{C_1} T_{13} \quad (10)$$

where $\theta_1, \theta_2, \theta_3$ are the conductor temperatures of phases A, B, and C, respectively, θ_a is the ambient temperature. T_{1a} is the total thermal resistance from conductor to ambient (including insulation, jacket, and convection), T_{1j} denotes the mutual thermal resistance between phase 1 and phase j. The thermal resistances can be calculated by IEC 60287-2-1.

The thermal resistance from the conductor core to the surface of its insulation:

$$T_i = \frac{\rho_i}{2\pi} \ln \left(1 + \frac{2\delta}{d_c} \right) \quad (11)$$

The thermal resistance from the surface of the insulation to the outer jacket of the cable:

$$T_{ik} = \frac{\rho_i}{2\pi} \ln \left(\frac{d_o}{d_i} \right) \quad (12)$$

The mutual thermal resistance between conductors:

$$T_{ij} = 2(T_i + T_{ik}) \quad (13)$$

In the studies by Gouda [13], [14], the thermal resistance of the surrounding soil is employed to describe heat transfer from the cable surface to the external environment. In this study, the soil thermal resistance is replaced by a modified air thermal resistance, which more appropriately represents the heat dissipation mechanism of PVC-insulated cables installed in free air:

$$T_{air} = \frac{1}{\pi D_e^* h (\Delta\theta_s)^{1/4}} \quad (14)$$

The total thermal resistance from conductor to ambient:

$$T_{1a} = T_i + T_{ik} + T_{air} \quad (15)$$

where ρ_i the thermal resistivity of the insulation material (PVC), δ is the insulation thickness, d_c the conductor diameter, d_o is the jacket outer diameter, d_i is the jacket inner diameter, D_e^* is the jacket outer diameter in meter and $h, (\Delta\theta_s)^{1/4}$ are the heat dissipation coefficient and the excess of cable surface temperature above ambient temperature which can be obtained in IEC 60287-2-1 respectively.

From Equations (8), (9), (10), and using the superposition principle, the following relation is obtained:

$$\frac{\theta_1}{T_{11}} - \frac{\theta_2}{T_{12}} - \frac{\theta_3}{T_{13}} = W_{C_1} + \frac{\theta_a}{T_{1a}} \quad (16)$$

An analogous set of equations can be derived for phases B and C. Collectively, they form the thermal balance system can be represented in matrix form:

$$\begin{bmatrix} \frac{1}{T_{11}} & \frac{-1}{T_{12}} & \frac{-1}{T_{13}} \\ \frac{-1}{T_{21}} & \frac{1}{T_{22}} & \frac{-1}{T_{23}} \\ \frac{-1}{T_{31}} & \frac{-1}{T_{32}} & \frac{1}{T_{33}} \end{bmatrix} \begin{bmatrix} \theta_1(h) \\ \theta_2(h) \\ \theta_3(h) \end{bmatrix} = \begin{bmatrix} W_{C_1}(h) + \frac{\theta_a}{T_{1a}} \\ W_{C_2}(h) + \frac{\theta_a}{T_{2a}} \\ W_{C_3}(h) + \frac{\theta_a}{T_{3a}} \end{bmatrix} \quad (17)$$

where are the copper losses of each phase in the presence of harmonic order h . Hence, an increase in the cable core temperatures in the three phases due to h^{th} harmonics can be obtained:

$$\begin{bmatrix} \Delta\theta_1(h) \\ \Delta\theta_2(h) \\ \Delta\theta_3(h) \end{bmatrix} = \begin{bmatrix} \frac{1}{T_{11}} & \frac{-1}{T_{12}} & \frac{-1}{T_{13}} \\ \frac{-1}{T_{21}} & \frac{1}{T_{22}} & \frac{-1}{T_{23}} \\ \frac{-1}{T_{31}} & \frac{-1}{T_{32}} & \frac{1}{T_{33}} \end{bmatrix}^{-1} \begin{bmatrix} W_{C_1}(h) + \frac{\theta_a}{T_{1a}} \\ W_{C_2}(h) + \frac{\theta_a}{T_{2a}} \\ W_{C_3}(h) + \frac{\theta_a}{T_{3a}} \end{bmatrix} - \theta_a \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad (18)$$

Then, the cable core temperatures of the three phases with harmonics can be obtained:

$$\begin{bmatrix} \theta_{t1} \\ \theta_{t2} \\ \theta_{t3} \end{bmatrix} = \theta_a \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} \sum_{h=1}^{\infty} \Delta\theta_1(h) \\ \sum_{h=1}^{\infty} \Delta\theta_2(h) \\ \sum_{h=1}^{\infty} \Delta\theta_3(h) \end{bmatrix} \quad (19)$$

The developed thermal model is subsequently applied to investigate the impact of harmonic currents on the thermal behavior of the cable.

3. SIMULATION SETUP AND HARMONIC SCENARIOS

Section 3 presents the simulation setup used in the investigation of the thermal response of the single-core cable subjected to harmonic excitations. The first section refers to the adopted cable and cable characteristics, including its electric, thermal, and installing attributes. The second section refers to the harmonic conditions considered.

3.1. Cable parameters

The analysis is carried out on a low-voltage single-core Cadivi cable rated at 0.6/1.0 kV, installed in free air and arranged in a trefoil configuration. Both the insulation and outer sheath are made of PVC, and according to IEC 60502-1, the maximum continuous operating temperature for this material is 70°C. The ambient temperature is assumed to be $\theta_a = 30^\circ\text{C}$. The cable has a nominal cross-sectional area of 70 mm² and a rated current of $I_{rated} = 219\text{ A}$. The principal geometrical dimensions and thermal–electrical properties employed in the simulations are summarized in Table 1.

Table 1. Parameters of the cable.

Parameter	Value	Units
ρ_i	6.5	$\text{k} \times \text{m/W}$
δ	1.4	mm
d_c	9.7	mm
d_i	12.5	mm
d_o	15.3	mm
S	15.3	mm
R_{DC}	0.268	$\text{m}\Omega/\text{m}$

3.2. Harmonic case studies

To investigate the thermal impact of harmonics on the single-core cable, two groups of case

studies are designed with complementary purposes.

3.2.1. Case 1 – effect of individual harmonic orders

In this group, the conductor temperature is evaluated under two conditions: (i) only the fundamental component at 50 Hz is present (without harmonics) and (ii) the fundamental component is superimposed with harmonic components of orders $6n \pm 1$ (e.g., 5th, 7th, 11th, 13th, etc.), which typically occur in three-phase systems. The total RMS current per phase is kept equal to the rated current, $I_{rated} = 219\text{ A}$. To examine the effect of harmonic order, THD is held constant while the harmonic order h is varied. Thus, Case 1 reveals how conductor heating changes with harmonic order for a given harmonic amplitude.

3.2.2. Case 2 – single-harmonic THD variation

In this group, the conductor temperature is evaluated when the fundamental component at 50 Hz is superimposed with only one harmonic order from the sequence $6n \pm 1$. The total RMS current per phase is maintained equal to the rated current, $I_{rated} = 219\text{ A}$. To assess the effect of harmonic amplitude, the harmonic current (expressed as THD) is varied across multiple levels while keeping the harmonic order fixed.

3.2.3. Case 3 – controlled rectifier-motor system simulation

In this case, the cable temperature is studied under practical operating conditions through a simulated system that includes a three-phase low-voltage supply, a fully controlled bridge rectifier, and a DC motor. By adjusting the firing angle of the rectifier, the current waveform on the AC side becomes increasingly distorted, which reflects different motor loading conditions. These current waveforms, containing both the fundamental and harmonic components, are then applied to the thermal model to estimate the conductor temperature. This case serves to connect the previous harmonic-based analyses with a more realistic industrial scenario.

Two operating scenarios are examined. In the first, the cable temperature is evaluated with and without harmonics at load levels ranging from 0% to 100% of the rated current to determine a suitable derating factor. In the second, the ambient temperature is varied from 20°C to 40°C, representing tropical climatic conditions in Da Nang, Vietnam, to assess the effect of environmental temperature on cable thermal limits.

For each condition, the conductor temperature and cable lifetime are estimated, allowing comparison between operation with and without harmonics. These results help to clarify how both harmonic distortion and ambient temperature jointly influence cable loading capability in practical applications.

For Case 1 and Case 2, the thermal calculations are carried out directly in MATLAB by parametrically defining the current harmonic components (magnitudes and orders) and applying them to the thermal model described by Eq. (19). For Case 3, the current waveforms are obtained from a MATLAB/Simulink model of a controlled rectifier–motor system, and the corresponding harmonic spectra are extracted using Fast Fourier Transform (FFT) analysis before being used as inputs to the same thermal model, as illustrated in Figure 2.

4. RESULTS AND DISCUSSION

This section presents the simulation results obtained from the thermal model introduced in Section 2 under the harmonic scenarios defined in Section 3.

4.1. Combined effect of harmonic order and current amplitude on conductor temperature

The values for Case 1 and Case 2 are depicted in Figures 3 and 4, respectively. The plots in the graphics collectively reveal the effects of the harmonic order h and the harmonic current amplitude I_h on the temperature of the conductors, while the overall RMS current per phase remains the same at 219 A.

In Figure 3, the curves show how the conductor temperature changes with harmonic orders of the form $6n \pm 1$ (such as the 5th, 7th, 11th, 13th, and so on) for different THD levels. Since the harmonic current follows $I_h = THD \times I_1$, a higher THD means a larger harmonic current amplitude and a smaller fundamental current I_1 , while the total RMS current remains unchanged. It can be noted that for the same harmonic order, larger harmonic currents generate more heat inside the conductor, which results in a higher temperature rise. Such a relationship follows Joule's law, in which the power dissipation is proportional to the square of the current magnitude.

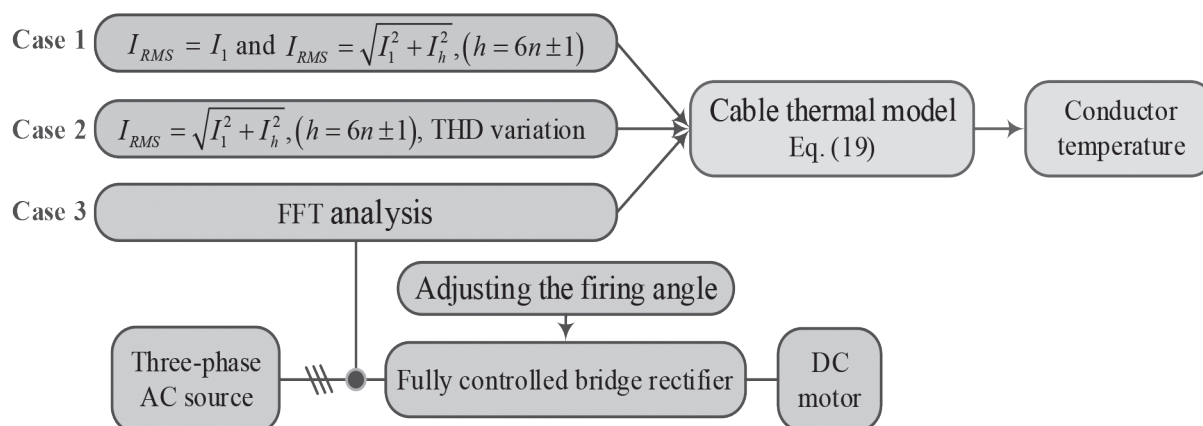


Figure 2. Overview of the three harmonic case studies and the MATLAB/Simulink computational workflow.

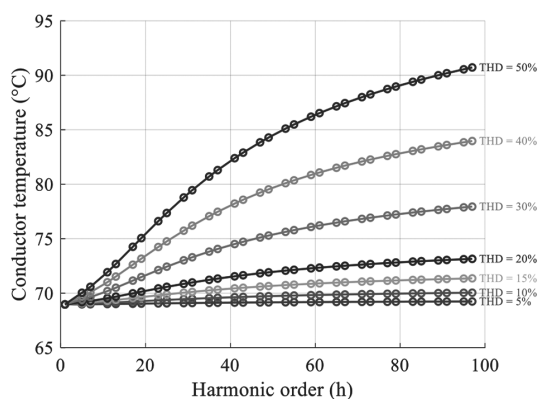


Figure 3. Variation of conductor temperature with harmonic order for different THD levels.

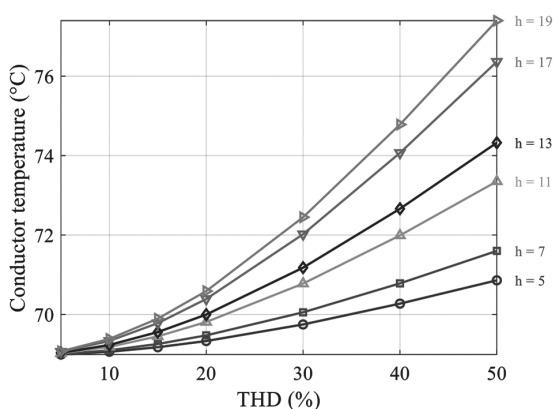


Figure 4. Variation of conductor temperature with THD for different harmonic orders.

Figure 4 illustrates the complementary relationship: each curve is related to a given harmonic order, while the THD varies to mimic changing harmonic amplitudes. On inspecting curves for the same THD levels (i.e., corresponding current amplitudes), it is clear that the temperature increases with harmonic order, so higher-order harmonics have a greater contribution to heating. This result occurs because the AC resistance $R_{AC}(h)$ significantly increases when the frequency increases, by virtue of the skin and proximity effects that become greater for higher harmonic orders and so result in an increase in Joule losses.

Considering the two plots together offers clearer insight into how harmonics influence the cable's heating behavior. The results show that as both the harmonic order and current amplitude rise, each harmonic adds extra loss

to the conductor, which raises its temperature. Hence, ignoring higher-order harmonics in real operation can lead to underestimating the actual temperature of the cable.

4.2. Thermal behavior in controlled rectifier–motor drive system (Case 3)

According to the results of Case 3 shown in Figure 5, changes in the rectifier's firing angle leads to variations in the DC motor load, and this change directly influences the harmonic distortion (THD) of the AC-side current. When the firing angle becomes larger, the current waveform departs further from the sinusoidal shape, its THD rises, while the overall current amplitude drops. Because of this reduction in current, the heat generated in the conductor also decreases, and the cable temperature gradually falls as the firing angle increases.

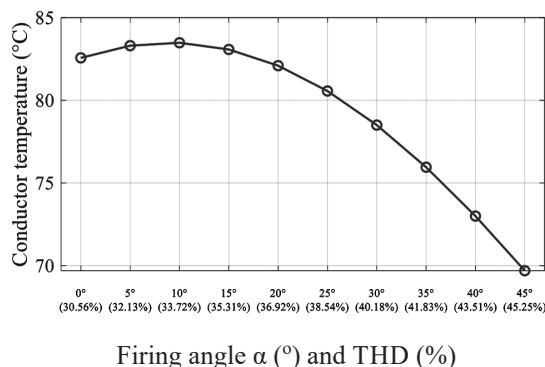


Figure 5. Variation of conductor temperature with firing angle.

Even so, the temperature values remain above the allowable operating limit of 70°C throughout the entire range of operation. This indicates that the presence of harmonic components increases the total electrical losses enough to push the cable into a state of thermal overload, even though the current itself is reduced. To maintain safe operation, both the load level and the harmonic content need to be considered, and suitable control or filtering measures should be applied to keep the cable temperature within acceptable limits.

4.3. Effect of load level with and without harmonics

Figure 6 clearly shows that the presence of harmonic distortion causes the conductor to

operate at a consistently higher temperature compared with the case of a pure sinusoidal current. This occurs even though both cases are evaluated using the same total RMS current. The rise in temperature can be attributed to additional losses generated by higher-frequency current components, which increase the effective AC resistance of the conductor.

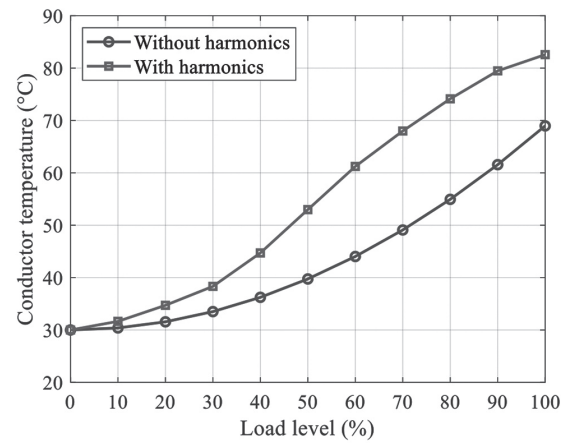


Figure 6. Conductor temperature variation with load level under harmonic and non-harmonic conditions.

To maintain the conductor temperature within the permissible limit, the operating load should be reduced to approximately 70 – 80% of the rated value. Another option is to select a cable type with a higher ampacity. These observations emphasize that harmonic effects must be accounted for during both cable selection and system operation, especially in networks feeding nonlinear loads such as converters or inverters.

To assess how this temperature increases influences insulation aging, the Arrhenius law is applied to estimate the decline in cable lifetime as the operating temperature rises under overload. The relationship is formulated as:

$$L(T_{op}) = L_{ref} \exp \left(\frac{E_a}{k_B} \left(\frac{1}{T_{op}} - \frac{1}{T_{ref}} \right) \right) \quad (20)$$

where $L_{ref} = 20$ years at $T_{ref} = 70^\circ\text{C} = 343\text{K}$, $E_a = 0.7\text{eV}$ is the activation energy of the PVC material, and $k_B = 8.617 \times 10^{-5} \text{ eV} / \text{K}$ is the Boltzmann constant. The estimated lifetimes obtained from the simulated temperatures are summarized in Table 2, comparing harmonic and non-harmonic cases at different load levels.

Table 2. Cable lifetime estimation for various load levels.

Load level	Without harmonics (years)	With harmonics (years)
70%	93.72	22.98
80%	59.07	15.22
90%	35.2	10.91
100%	21.48	8.65

The data reveals a clear degradation trend: the presence of harmonics shortens the insulation lifetime considerably. For an identical loading condition, the expected lifetime under harmonic operation drops to roughly 40 – 50% of that in a purely sinusoidal case. At rated load, the cable lifetime decreases from 20 years to about 8 years, indicating strong thermal aging induced by harmonic-related overheating. Consequently, in practical operation involving loads with nonlinear components, it is necessary to determine an appropriate derating level to ensure the long-term reliability of the cable and the electrical system.

4.4. Influence of ambient temperature

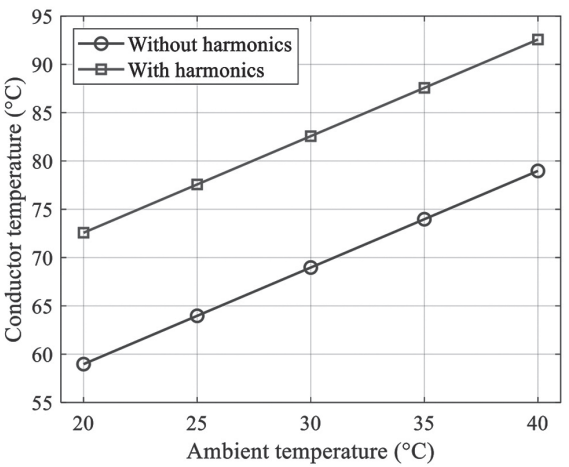


Figure 7. Effect of ambient temperature on conductor temperature.

Figure 7 shows how the cable core temperature varies with ambient temperature once thermal equilibrium is reached. Since the cable is placed in a real thermal environment, its core temperature naturally depends on the surrounding conditions. For that reason,

ambient temperature should be treated as an operating variable rather than a fixed constant. In practice, ambient temperature fluctuates from time to time, but the cable core does not follow those short-term changes immediately because of its thermal inertia. When the surrounding temperature stays roughly constant for a period much longer than the cable’s thermal time constantly-say, over a day or a season, the core temperature gradually settles to a steady value. Therefore, using steady-state core temperatures associated with average ambient conditions is a reasonable way to describe the long-term thermal behavior of the cable.

At steady operation, when the core temperature remains nearly constant, the insulation material ages faster as the temperature increases, following the Arrhenius law. Based on that relation, the expected service life of the cable, both with and without harmonic components, was evaluated at several ambient temperature levels. The calculated results are summarized in Table 3.

From Table 3, it can be observed that higher ambient temperatures correspond to a shorter cable lifetime, reflecting how insulation materials naturally deteriorate under heat. The influence of harmonics varies with temperature: at lower ambient levels, their impact is minor, but as the environment becomes hotter, harmonic losses intensify, leading to a faster reduction in service life. Therefore, in practical operation, in addition to the derating effect caused by harmonics, the cable load should also be adjusted according to ambient temperature conditions to ensure its durability and service lifetime.

Table 3. Conductor lifetime under different ambient conditions.

Ambient temperature (°C)	Without harmonics (years)	With harmonics (years)
20	43.35	16.12
25	30.16	12.33
30	21.21	8.90
35	15.07	6.47
40	10.81	4.47

4.5. Comparative discussion

The three case studies reveal that the thermal response of single-core cables to harmonic distortion depends on multiple interacting factors, including the harmonic order, current level, load, and ambient temperature. As seen in Sections 4.1 and 4.2, higher-order harmonics intensify the skin and proximity effects, thereby increasing the effective AC resistance and raising the conductor temperature even when the RMS current remains unchanged. Extending these observations to the rectifier–motor system in Case 3 leads to the same pattern: when the current waveform becomes more distorted (i.e., higher THD), the conductor operates at a higher temperature despite a smaller overall current.

The comparison between Sections 4.3 and 4.4 shows that both load and ambient temperature amplify the influence of harmonics. At heavier loads, the added harmonic losses drive the operating temperature above safe limits, cutting the expected insulation lifetime to roughly half of that under pure sinusoidal conditions. As the ambient temperature rises, insulation aging accelerates, and the additional heating due to harmonics becomes more evident—an effect that deserves attention in warm or tropical environments where cables often run close to their thermal capacity.

Taken together, these results underline that harmonic effects have a lasting influence on cable thermal performance. Evaluating cable ratings solely from the RMS current can underestimate the true conductor temperature and the rate of insulation aging. In practical operation, both harmonic distortion and ambient temperature must be considered in load limit settings to ensure safe performance and long-term reliability of the cable system.

5. CONCLUSION

This study evaluates the thermal behavior of low-voltage single-core cables under harmonic current loading using an improved steady-state thermal model adapted for overhead

PVC-insulated cables. The analysis shows that harmonic currents significantly increase the AC resistance through enhanced skin and proximity effects, thereby producing additional thermal losses even when the RMS current remains unchanged. As a result, the conductor temperature may exceed its allowable limit under nonlinear loading conditions, accelerating insulation ageing and reducing cable lifetime. This adverse effect becomes more pronounced at elevated ambient temperatures, which is typical in tropical operating environments. The findings indicate that current ratings based solely on RMS values do not adequately represent the actual thermal stress experienced by cables in the presence of harmonics. Accordingly, appropriate current-derating measures should be considered based on both the harmonic spectrum and the ambient temperature to ensure thermal safety and long-term operational reliability.

Future studies need to focus on adding dynamic and transient aspects to the model so that temperature variations over time can be represented more accurately. In addition, finite element (FEM) techniques need to be applied to examine detailed temperature fields within the cable. A combination of numerical modeling and experimental validation is expected to offer a more detailed understanding of the electro-thermal interaction under harmonic-rich conditions.

DECLARATION OF CONFLICTS OF INTEREST

The authors declare that this research has no conflicts of interest.

REFERENCES

- [1] J. Wang, Y. Liang, and C. Ma, "A novel transient thermal circuit model of buried power cables for emergency and dynamic load," *Energy Reports*, vol. 9, pp. 963-971, 2023.
- [2] G. Petrović, M. Cvetković, T. Garma, and T. Kilić, "An approach to thermal modelling of power cables installed in ducts," *Electric Power Systems Research*, vol. 214, p. 108916, 2023.
- [3] W. Wang et al., "Review of calculation methods of power cables temperature based on thermal circuit model," In Proc. 2024 IEEE 7th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), 2024, pp. 1711-1714.
- [4] Y. B. Demireli and O. Kalenderli, "Investigation of effect of laying and bonding parameters of high-voltage underground cables on thermal and electrical performances by multiphysics FEM analysis," *Electric Power Systems Research*, vol. 227, p. 109987, 2024.
- [5] Y. S. Kim, H. N. Cong, B. H. Dinh, and H. K. Kim, "Effect of ambient air and ground temperatures on heat transfer in underground power cable system buried in newly developed cable bedding material," *Geothermics*, vol. 125, p. 103151, 2025.
- [6] E. Mustafa et al., "Thermal degradation and multi-performance aging assessment of low voltage nuclear cable," *Electric Power Systems Research*, vol. 250, p. 112128, 2026.
- [7] P. Wang, "Optimization for thermal rating method of high voltage cable considering the existence of air gap," *Electric Power Systems Research*, vol. 217, p. 109151, 2023.
- [8] M. V. S. D. Silva et al., "Evaluation of low voltage on electrical cables using microCT and COMSOL Multiphysics," *Radiation Physics and Chemistry*, vol. 224, p. 112067, 2024.
- [9] M. Quercio, J. C. D. P. Lopez, S. Grasso, and A. Canova, "Numerical and experimental analysis of thermal behaviour of high voltage power cable in unfilled ducts," *Scientific Reports*, vol. 14, p. 20599, 2024.
- [10] R. Hu et al., "Improved thermal analysis for three-core cable under unbalanced three-phase loads," *Electric Power Systems Research*, vol. 216, p. 108964, 2023.
- [11] M. Xu et al., "Thermal assessment method optimization for three-core cable installation in cable trenches," In Proc. 2025 IEEE 3rd International Conference on Image Processing

- and Computer Applications (ICIPCA), 2025, pp. 450–455.
- [12] H. Wang, W. Zhou, K. Qian, and S. Meng, “Modelling of ampacity and temperature of MV cables in presence of harmonic currents due to EVs charging in electrical distribution networks,” *International Journal of Electrical Power & Energy Systems*, vol. 112, pp. 127-136, 2019.
- [13] O. E. Gouda and A. Z. El. Dein, “Enhancement of the thermal analysis of harmonics impacts on low voltage underground power cables capacity,” *Electric Power Systems Research*, vol. 204, p. 107719, 2022.
- [14] O. E. Gouda, A. Z. El. Dein, E. Tag-Eldin, M. Lehtonen, and M. M. F. Darwish, “Thermal analysis of the influence of harmonics on the current capacity of medium-voltage underground power cables,” *Energy Science & Engineering*, vol. 11, no. 10, pp. 3471-3485, 2023.
- [15] P. Ratheiser, S. Shahtaj, U. Schichler, and B. G. Stewart, “Harmonics in MVDC cable systems - study on thermal and electric stress,” In Proc. 23rd International Symposium on High Voltage Engineering (ISH), 2023, pp. 43-49.
- [16] G. F. A. Osman, T. M. Z. Elsharkawy, and W. A. A. Salem, “Thermal analysis of underground distribution cables under dynamic loading in the presence of harmonic load currents,” In Proc. 22nd International Middle East Power Systems Conference (MEPCON), 2021, pp. 319-325.
- [17] S. H. Alwan, A. K. Jasim, and Y. F. Hassan, “Thermal behavior of insulated cables under the influence of harmonics in hot countries,” *International Journal on Energy Conversion*, vol. 12, no. 1, pp. 24-31, 2024.
- [18] A. Hiranandani, “Calculation of conductor temperatures and ampacities of cable systems using a generalized finite difference model,” *IEEE Transactions on Power Delivery*, vol. 6, no. 1, pp. 15-24, 1991.



© 2026 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Phân tích nhiệt của cáp điện hạ áp có xét đến ảnh hưởng của sóng hài dòng điện

Nguyễn Kim Ánh*, Huỳnh Đại, Lê Thị Thu, Lê Huy Hoàng, Phùng Bá Tú

Khoa Điện, Trường Đại học Bách khoa – Đại học Đà Nẵng, Việt Nam

Ngày nhận bài: 27/11/2025; Ngày sửa bài: 09/02/2026;

Ngày nhận đăng: 12/02/2026; Ngày xuất bản: 28/06/2026

TÓM TẮT

Trong thiết kế và lựa chọn cáp điện hạ áp, ảnh hưởng của gia nhiệt do sóng hài dòng điện thường chưa được xem xét đầy đủ. Sự gia tăng của các phụ tải phi tuyến và nguồn năng lượng tái tạo làm tăng các thành phần dòng điện hài, dẫn đến điện trở xoay chiều và tổn hao Joule của cáp tăng lên, từ đó làm gia tăng nhiệt độ vận hành phụ thuộc vào biên độ và bậc hài. Khi có sóng hài, tổn hao trong ruột dẫn tăng, khiến nhiệt độ lõi cáp có thể vượt đáng kể so với trường hợp chỉ có dòng điện cơ bản. Trong nghiên cứu này, mô hình nhiệt ban đầu áp dụng cho cáp ngầm được cải tiến và hiệu chỉnh để phù hợp với đặc tính tản nhiệt của cáp PVC lắp đặt trên không, loại cáp phổ biến trong lưới điện hạ áp tại Việt Nam. Phổ dòng điện hài cùng các tham số nhiệt–điện của cáp được sử dụng làm dữ liệu đầu vào để đánh giá quá trình gia nhiệt dưới các điều kiện vận hành khác nhau. Nghiên cứu phân tích ảnh hưởng của sóng hài đến nhiệt độ lõi cáp, đánh giá sự suy giảm tuổi thọ cách điện và xem xét tác động của nhiệt độ môi trường. Trên cơ sở kết quả tính toán, mức giảm dòng điện vận hành phù hợp được đề xuất nhằm duy trì nhiệt độ cáp trong giới hạn cho phép, góp phần đảm bảo vận hành an toàn và hạn chế rủi ro suy giảm cách điện trong lưới điện hạ áp.

Từ khóa: *Cáp điện hạ áp, sóng hài dòng điện, đặc tính nhiệt, giới hạn vận hành, ước lượng tuổi thọ của cáp.*

**Tác giả liên hệ chính.*

Email: nkanh@dut.udn.vn

Effect of maturation stages and storage conditions on physicochemical properties of three-leave cayratia (*Cayratia trifolia*) berries from different geographical regions in Vietnam

Doan Thi Kieu Tien^{1,*}, Tran Diem Phuong¹, Huynh Thi Ngoc Mi¹,
Dinh Hoang Lan Chi¹, Nguyen Ngoc Thanh², Huynh Xuan Phong²

¹Can Tho University of Technology, Vietnam

²Can Tho University, Vietnam

Received: 27/10/2025; Revised: 03/02/2026;

Accepted: 06/03/2026; Published: 28/06/2026

ABSTRACT

This study aims at investigating the influence of sampling locations, maturation stages, and storage conditions on total polyphenol content in threeleaf cayratia (*Cayratia trifolia*) berries collected from 09 provinces of the Mekong Delta River in Vietnam, including Kien Giang, An Giang, Dong Thap, Long An, Bac Lieu, Ca Mau, Soc Trang, Tra Vinh, and Ben Tre. The High Performance Liquid Chromatography technique (HPLC) was used to quantify quercetin and resveratrol contents in *Cayratia trifolia* berries samples. The results showed that *Cayratia trifolia* berries had a high polyphenol content from 12.14 to 19.51 mg GAE/g, being highest in samples collected in Dong Thap (DT1) (19.51 mg GAE/g). During the ripening stages from unripened to fully ripened, the total phenolic content of *Cayratia trifolia* berries samples from An Giang (AG3), Tra Vinh (TV4), Dong Thap (DT1, DT2), Ca Mau (CM1, CM2) increased whereas it decreased in those collected in KG1, DT3, DT4, CM3, and CM4. This variation could be due to changes in phenolic compound accumulation during fruit maturation, geographical location, and growing conditions. Besides, total polyphenol content in frozen *Cayratia trifolia* berries frozen showed no significant change (18.62 mg GAE/g) during 5 days of storage, while increased to 19.05 mg GAE/g after 7 days. The results on HPLC further showed that the amount of quercetin and resveratrol in DT samples were 0.14 ± 0.012 and 1.46 ± 0.035 mg/g fresh weight.

Keywords: *Cayratia trifolia*, polyphenol, quercetin, resveratrol.

1. INTRODUCTION

Cayratia trifolia (CT) belongs to the Vitaceae family, found mainly in India, Asia, and Australia. It is a perennial climber, trifoliated leaves with 2-3 cm long petioles and ovate to oblong-ovate leaflets, and small white and brown-greenish flowers[1]. Berries are fleshy fruits with spherical shape, up to 2 cm in diameter [1], dark purple when fully

ripen [2]. CT, is widely distributed in many regions of the Mekong Delta, locally named “Giac” [2]. The whole plant contains tannins, alkaloids, flavonoids, polyphenols, steroids, and yellow waxy oil [1] - [3]. The fruit extract Giac possesses various health benefits such as antioxidation, and anticancer [8].

The berries are rich in polyphenol compounds, exhibiting health and anti-aging

*Corresponding author:

Email: dtktien@ctu.edu.vn

potential [2] - [9]. However, changes in polyphenol content may exist concerning variety, geographic location, environmental conditions, ripening status, and storage condition [10]. This study was interested in the effects of ripening stages and storage conditions on the polyphenol contents of the fruits.

The differences in bioactive compound contents during berry maturation are well documented [11] - [13]. For instance, the total phenolic content (TPC) of plums decreased during development, while anthocyanin content increased in plums as a function of maturity stages [11]. Similarly, the highest TPC of kiwis is identified at the immature stage [13] and young kiwi stage [14]. Also, the effects of maturation on the amounts of bioactive compounds in grapes [15], strawberries [16], redcurrants [17], and raspberries [18], [19] have been demonstrated.

The storage conditions are the other factors that influence the polyphenol content of perishable berries [20] - [22]. Therefore, this study aims to investigate the effect of ripening stages and storage conditions on the polyphenol contents of CT in Vietnam. Further, resveratrol and its glucosides was found in the CT roots culture. Additionally, resveratrol and quercetin are rich stilbenes and flavonols found the fruit berries. Quercetin and other flavonoid components have been extensively found in different types of fruit berries and plant species around the globe. With these evidences, we quantified quercetin and resveratrol contents to evaluate their presence in Giac's fruit flesh.

2. MATERIALS AND METHODS

2.1. Materials

Samples: Berries samples were collected in 9 provinces of the Mekong Delta, Vietnam (Table 1). The provincial names were recorded as they existed at the time of sampling, as they have changed since July 1st 2025.

Table 1. Samples locations of Giac's berries.

Provinces	Locations	Sample names
1. Kien Giang 9.8250° N, 105.1259° E	Giong Rieng	KG1
	Ha Tien	KG2
	Kien Luong	KG3
	Hon Dat	KG4
2. An Giang 10.5216° N, 105.1259° E	Cho Moi	AG1
	Tri Ton	AG2
	Tinh Bien	AG3
	Phu Tan	AG4
3. Dong Thap 10.4938° N, 105.6882° E	Chau Thanh	DT1
	Lai Vung	DT2
	Cao Lanh	DT3
	Dong Thap	DT4
4. Long An 10.6851° N, 106.2051° E	Can Duoc	LA1
	Ben Luc	LA2
	Chau Thanh	LA3
	Thanh Hoa	LA4
5. Bac Lieu 9.2899° N, 105.5005° E	Hoa Binh	BL1
	Hong Van	BL2
	Gia Rai	BL3
	Phuoc Long	BL4
6. Ca Mau 9.1527° N, 105.1961° E	Cai Nuoc	CM1
	Nam Can	CM2
	Tran Van Thoi	CM3
	U Minh	CM4
7. Soc Trang 9.6025° N, 105.9739° E	Ke Sach	ST1
	My Xuyen	ST2
	My Tu	ST3
	Nga Nam	ST4
8. Tra Vinh 9.7627° N, 106.4406° E	Cau Ngang	TV1
	Chau Thanh	TV2
	Duyen Hai	TV3
	Tieu Can	TV4
9. Ben Tre 10.1082° N, 106.4406° E	Ba Tri	BT1
	Mo Cay Bac	BT2
	Thanh Phu	BT3
	Ben Tre	BT4

Chemicals: Sodium carbonate (Na_2CO_3) was purchased from Xilong Scientific (Guangdong, China). Gallic acid was purchased from HiMedia (Maharashtra, India). Folin-Ciocalteu and methanol were purchased from Merck (Darmstadt, Germany).

2.2. Effects of the collected locations

A total of 36 dark purple samples were tested for pH, Brix and TPC contents.

2.3. Effects of the ripening stages

Three stages of ripeness were selected, namely unripe (M1, green and hard pulpy), barely ripe (M2, red-purple and slightly soft pulpy), and fully-ripe (M3, dark purple and soft pulpy) (Figure 1A).

2.4. Effects of the storage conditions

The berries were stored at ambient (28 - 32 °C), refrigerated (4 - 10 °C) and frozen conditions (-18 ± 2 °C), with durations in 3, 5, and 7 days.

2.5. Analysis of physicochemical properties

The pH value was measured using a pH meter (Hanna, USA) and the total soluble solid content was recorded using a manual refractometer (ATAGO, 0-33 °Brix, France).

Total polyphenol content (TPC) was determined following the Folin–Ciocalteu colorimetric method [23]. In detail, the berries were washed thrice under tap water and finally rinsed with distilled water. After draining for 30 mins, 10 g samples were completely crushed and allocated 25 μL into a test tube, and 975 μL methanol was added to the sample. Subsequently, 200 μL of Folin–Ciocalteu reagent 10% and 2500 μL Na_2CO_3 (20% w/v) were added to the mixture. Then, the mixture was homogenized with a vortex for 15 seconds and incubated at room temperature for 45 min in the dark. The absorbance of the sample was measured at 765 nm using a UV-Vis spectrophotometer (ACCURIS SmartReader; Edison, NJ, USA). Gallic acid was used to build a standard curve using a linear regression equation with the relationship

between gallic acid concentrations (20, 40, 60, 80, 100, 120 $\mu\text{g/mL}$) expressed as the X axis and the magnitude of the absorbance results of the reaction of gallic acid with Folin-Ciocalteu reagent stated as the Y axis. TPC estimation was expressed as milligram gallic acid equivalents per gram of sample fresh weight (mg GAE/g).

Additionally, experimental samples were chosen for the quantification of quercetin and resveratrol using high performance liquid chromatography (HPLC). The samples were extracted with a mixture of methanol/water/acetonitrile (30:35:35). Methanol (100%) have been used as mobile phase solution for separation and quantification of the target components. The flow rate was at 0.5 mL/min, the processed temperature in the column was kept stable at 40°C, and the injection volume was 10 μL . UFLC HPLC system (Shimadzu, Kyoto, Japan) was used. Reverse-phase HPLC process were performed using a Shimadzu C18 column (250 x 4.6 mm).

2.6. Data analysis

All the experiments were conducted in triplicate. Data were expressed as means $\pm$ standard deviation (SD). SPSS 20.0 software was used for data analysis including one-way analysis of variance (ANOVA) test and LSD analysis for comparison of means.

3. RESULTS AND DISCUSSION

3.1. Properties of berry samples by regions

The TPC of CT berries varies between 12.14 and 19.51 mg GAE/g while the pH values ranges from 2.94 to 3.64, and the total soluble solids fall within 7.0 to 9.0 °Brix (Table 2). These data indicate the deviation in the flavor of berries.

The berries collected in the different ecological regions in Vietnam exhibit significant differences in TPC ($p \leq 0.05$). The high TPC was identified in DT samples (19.23 - 19.51 mg GAE/g), followed by CM (17.75 - 18.13 mg GAE/g), TV (16.07 - 17.52 mg GAE/g), ST (17.01 - 17.48 mg GAE/g), Kien Giang (14.59 - 17.34 mg GAE/g),

LA (14.89-16.45 mg GAE/g), AG (12.14 - 16.57 mg GAE/g), BT (13.51 - 16.26 mg GAE/g), and BL (13.13 - 13.80 mg GAE/g). Noblely, DT1 had the highest TPC of 19.51 mg GAE/g, and the lowest TPC of 12.14 mg GAE/g was found in AG2, significantly different from other samples at the 95% confidence level. All samples in this study had higher TPC than those extracted by methanol (4.6±0.3 mg GAE/g sample) and water (2.9±0.1 mgGAE/g) in Malaysia [9].

Table 2. pH, Brix, and TPC of CT berries collected in 9 provinces of the Mekong Delta River.

Province	Sample	pH	°Brix	TPC (mgGAE/g)
Kien Giang	KG1	3.07	7.0	17.34 ^{fg}
	KG2	3.05	7.5	14.59 ^{pq}
	KG3	3.45	8.0	15.63 ^m
	KG4	2.98	7.0	16.29 ^{jk}
An Giang	AG1	3.64	8.0	15.01 ⁿ
	AG2	2.94	7.0	12.14 ^u
	AG3	3.59	8.5	16.57 ⁱ
	AG4	3.20	8.0	14.72 ^{op}
Dong Thap	DT1	3.35	7.0	19.51 ^a
	DT2	3.46	7.5	17.79 ^d
	DT3	3.39	9.0	18.05 ^e
	DT4	3.30	8.0	19.23 ^b
Long An	LA1	3.45	8.0	14.89 ^{no}
	LA2	3.30	7.0	16.45 ^{ij}
	LA3	3.48	8.0	15.67 ^m
	LA4	3.38	7.0	16.06 ^l
Bac Lieu	BL1	3.79	8.5	13.80 ^r
	BL2	3.81	10.0	13.13 ⁱ
	BL3	3.86	9.0	14.53 ^q
	BL4	3.65	9.0	13.62 ^s
Ca Mau	CM1	3.09	6.0	17.99 ^e
	CM2	3.74	7.0	18.13 ^e
	CM3	3.15	8.0	18.03 ^e
	CM4	3.59	7.0	17.75 ^d
Soc Trang	ST1	3.50	8.0	17.05 ^h
	ST2	3.37	6.0	17.48 ^{ef}
	ST3	3.31	7.0	17.24 ^g
	ST4	3.60	7.0	17.01 ^h
Tra Vinh	TV1	3.44	9.0	16.28 ^{jk}
	TV2	3.37	9.0	16.35 ^{jk}
	TV3	3.85	8.0	16.07 ^l
	TV4	3.88	9.0	17.52 ^e
Ben Tre	BT1	3.68	6.0	13.51 ^s
	BT2	3.29	7.0	16.26 ^k
	BT3	3.72	7.0	16.39 ^{jk}
	BT4	3.26	6.0	14.47 ^q

The average values in a group with the same letter were not significantly different at the 95% confidence level. KG, AG, DT, LA, BL, CM, ST, TV, and BT are the abbreviations of provinces where the samples were collected; 1-4: number of collected samples of each province.

Likewise, the TPC of berries was higher than that of other tropical berries extracted by H₂O, such as pineapple (34.7 mgGAE/100g), banana pisang mas (27 mgGAE/100g), guava (153 mgGAE/100g)²⁴. Compared to grape species (another genus of the Vitaceae family), berries had higher TPC than both white and red varieties [25]. However, it was lower than that of species of Myrtaceae (49.21 mg GAE/g) [26], Myrciaria cauliflora (31,6 mg GAE/g) and Eugenia aggregate (25,3mg GAE/g) [27]. Thirteen berries with a high TPC were selected for further study on the changes in polyphenol content during fruit maturation in the next experiment.

3.2. TPC of CT berries during maturation

Berries collected in 9 regions showed significant differences in TPC ($p < 0.05$) as affected by stages of ripeness (Table 3). Most fully ripened CT berries acquired the highest TPC, followed by the unripened and barely ripened samples (Figure 1). This trend is partly in accordance with our previous report on the sampling region in Can Tho (Vietnam), where the fully ripened samples were the mutual best option in terms of TPC [28].

Interestingly, TPC gradually increased from unripened to fully ripened berries in AG3, TV4, DT1, DT2, CM1, and CM2. By contrast, this value reduced from unripened to fully ripened in samples collected in KG1, DT3, DT4, CM3, and CM4. This difference might be due to the variation of each compound and its concentration (acid gallic, anthocyanin, quercetin, ellagic acid, etc.) in the total polyphenolic content. According to Guofang, the gallic acid content of all rabbiteye blueberry cultivars increased at the beginning of berry formation and decreased at the berry ripening stage [29]. Meanwhile, the ferulic acid contents of Powderblue and Gardenblue cultivars gradually increased during ripening, and the ellagic acid content of Powderblue blueberries increased at the ripening stage but decreased at this stage for Baldwin blueberries [29]. Tahir Mahmood (2017) documented an increasing trend in TPC

(201 - 2287 mg/100 g GAE) of mulberry fruits [30]. In this study, the TPC was highest in ripe DT1 and lowest in unripe ST2 samples, at 20.66 mg GAE/g and 14.02 mg GAE/g, respectively.

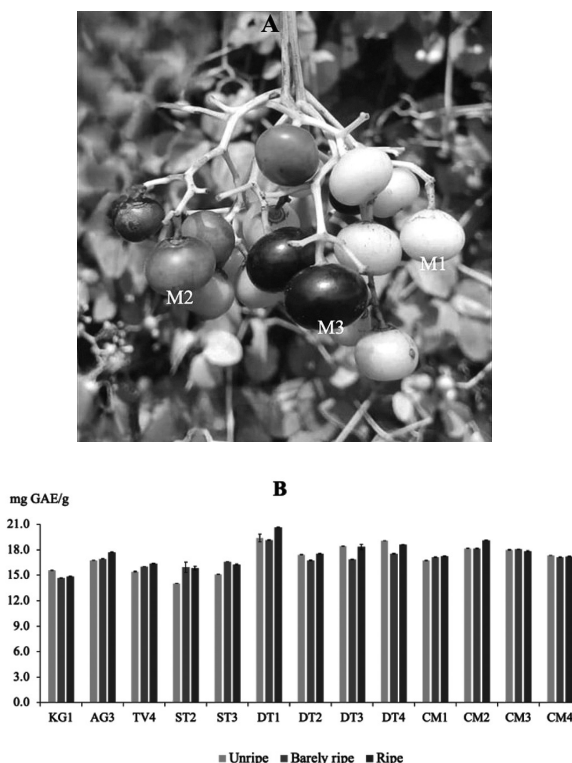


Figure 1. The effect of the ripeness stage on polyphenol content of CT berries. A: The picture of color changes in CT berries during the ripening stage: full-green (M1), purple (M2), and dark purple (M3). B: the variation in the total polyphenol content of *Cayratia trifolia* berries at three different ripeness stages. KG, AG, TV, ST, DT, and CM are the abbreviations of provinces where the samples were collected; 1-4: number of collected samples of each province.

The fluctuation in TPC in this study might depend greatly on the complex alteration of polyphenol composition in different genotypes of the same *Cayratia trifolia* L. variety and the development of individual phenolic compounds, requiring further adequate investigation as maturity progresses. In the food production industry, the ripening stage of berries greatly affects the flavor, color, and nutritional content of the product due to the presence of mineral components and biological compounds [31]. CT wine and cider are two fermented products

from full-ripened CT berries developed in many previous studies [32], [33]. To further diversify CT-based products and improve their quality and yield, a comprehensive assessment of berry quality and origin is essential. Moreover, because Giac is widely distributed across the

Mekong Delta region of Vietnam, the centralized transportation and storage of raw materials for fruit processing, underutilization of this resource represents a significant limitation. The fully ripened berries (DT1-M3) was chosen for the storage stability experiment.

Table 3. pH, Brix, and TPC of Giac’s berries collected at 3-different ripening stages.

Ripening stage	Sample	pH	°Brix	TPC (mgGAE/g)
M1	KG1-M1	2.78	3.0	15.57 ^s
	AG3-M1	2.78	3.0	16.75 ^{op}
	TV4- M1	2.73	3.0	15.44 ^s
	ST2- M1	2.92	3.0	14.02 ^v
	ST3- M1	2.86	4.0	15.11 ^t
	DT1- M1	2.82	3.0	19.41 ^b
	DT2- M1	2.80	3.5	17.42 ^{kl}
	DT3- M1	2.79	3.0	18.43 ^{de}
	DT4- M1	2.77	3.5	19.06 ^c
	CM1- M1	2.81	3.0	16.74 ^{op}
	CM2- M1	2.76	4.0	18.18 ^{fg}
	CM3- M1	2.83	4.0	18.00 ^{gh}
	CM4- M1	2.80	3.5	17.32 ^{lm}
M2	KG1- M2	2.78	4.0	14.69 ^u
	AG3- M2	2.70	4.5	16.93 ^{no}
	TV4- M2	2.74	4.0	16.00 ^r
	ST2- M2	2.88	3.5	15.95 ^r
	ST3- M2	2.87	4.0	16.61 ^p
	DT1- M2	2.78	3.5	19.16 ^c
	DT2- M2	2.79	4.5	16.77 ^{op}
	DT3- M2	2.88	4.0	16.88 ^o
	DT4- M2	2.87	4.0	17.55 ^{jk}
	CM1- M2	2.97	4.0	17.14 ^{mn}
	CM2- M2	2.98	3.0	18.15 ^{fg}
	CM3- M2	2.93	3.0	18.06 ^g
	CM4- M2	2.96	3.5	17.13 ^{mn}
M3	KG1- M3	2.95	5.0	14.85 ^u
	AG3- M3	3.39	7.0	17.72 ^{ij}
	TV4- M3	3.04	5.0	16.38 ^q
	ST2- M3	3.04	5.0	15.84 ^r
	ST3- M3	3.12	6.0	16.27 ^q
	DT1- M3	3.29	9.0	20.66 ^a
	DT2- M3	3.45	7.5	17.54 ^{jk}
	DT3- M3	3.42	7.0	18.35 ^{ef}
	DT4- M3	3.54	9.0	18.61 ^d
	CM1- M3	3.22	7.0	17.25 ^{lm}
	CM2- M3	3.62	7.0	19.13 ^c
	CM3- M3	3.36	9.0	17.83 ^{hi}
	CM4- M3	3.48	8.0	17.15 ^{lm}

The average values in a group with the same letter were not significantly different at the 95% confidence level. KG, AG, DT, LA, BL, CM, ST, TV, and BT are the abbreviations of provinces where the samples were collected; 1-4: number of collected samples of each province; M1: unripened green berry; M2: barely ripened purple berry; M3: full-ripened dark purple berry.

3.3. Effects of the temperature and storage duration on TPC of CT berries

The influence of storage temperature (refrigerating, freezing, and ambient) and storage time (3, 5, and 7 days) on TPC of berries are shown in Table 4. Room temperature caused adverse effects in polyphenol retention in berries, representing the lowest amount of TPC. At this temperature, the berries decayed during 7-day storage since high temperature might promote microbial spoilage, respiration level, and oxidative and nonoxidative reactions²².

Prolong storage duration seemed to degrade polyphenols in samples stored at tropical temperatures while unexpectedly advanced such content in those kept at freezing condition. The total phenolic compound content in frozen samples obtained the highest value of 19.05 mg GAE/g in 7-day storage. In our previous study, CT berries stored at the frozen temperature represented maximal TPC [28]. The best storage temperature to preserve polyphenols in berry was freezing to -20 °C, as evidenced by many authors [22] - [25]. Freezing temperatures crystallize water, reducing water activity, biochemical changes, and microbial growth [34].

Table 4. The effects of temperature and duration storage on TPC of Giac’s berries.

Temperature (°C) - days	Samples	TPC (mgGAE/g)
33°C - 3	DT1-RT3	18.34 ^c
33°C - 5	DT1-RT5	18.42 ^c
33°C - 7	DT1-RT7	-
4°C - 3	DT1-CT3	16.49 ^c
4°C - 5	DT1-CT5	16.18 ^d
4°C - 7	DT1-CT7	16.16 ^d
-5°C - 3	DT1-FT3	18.66 ^b
-5°C - 5	DT1-FT5	18.62 ^b
-5°C - 7	DT1-FT7	19.05 ^a

The average values in a group with the same letter were not significantly different at the 95% confidence level.

3.4. Quantification of resveratrol and quercetin in the berries

Concentration of resveratrol and quercetin was quantified by using RP-HPLC method, at wavelength of 306 nm and 368 nm, respectively. Resveratrol peaks in the CT sample had the retention time at 5.4 - 5.6 minutes. Quercetin peaks in the CT sample had the retention time at 5.5 - 5.8 minutes. The average concentrations of resveratrol in the samples was 1.46 mg/g wet weight while lesser amount counted for quercetin by 0.14 mg/g. A paired sample test showed that the amount of resveratrol was 1.32 ± 0.046 mg higher than that of quercetin, with significant difference ($p < 0.001$). Resveratrol and its glucosides was found in the CT roots culture [35]. Additionally, resveratrol and quercetin are rich stilbenes and flavonols found the fruit berries [36]. Quercetin and other flavonoid components have been extensively found in different types of fruit berries and plant species around the globe [37], [38].

Table 5. Concentrations of resveratrol and quercetin in CT berries.

Compounds	Retention time (minutes)	Concentration (mg/g wet weight)
Resveratrol	5.4 - 5.6	$1.46 \pm 0.35^{***}$
Quercetin	5.5 - 5.8	0.14 ± 0.11

The result values are expressed as means $\pm$ SD, $n = 3$. The stars (***) indicates a significant difference at $p < 0.001$.

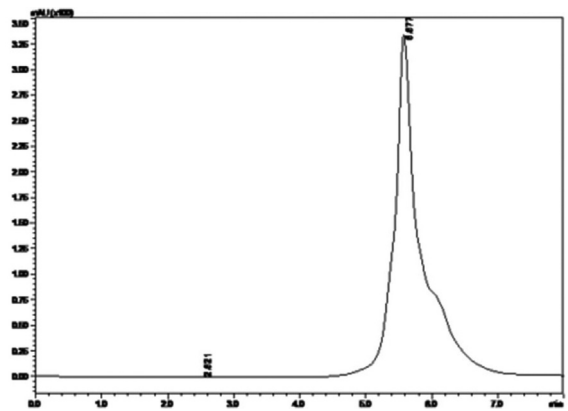


Figure 2. Chromatogram of resveratrol in CT berries.

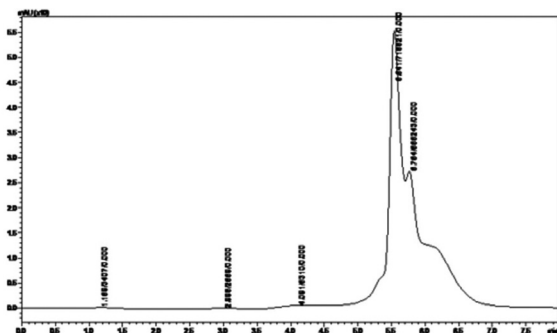


Figure 3. Chromatogram of quercetin in CT berries.

4. CONCLUSION

Wild three-leaf cayratia in Vietnam can be classified into berries with high polyphenol content from 12.14 to 19.51 mg GAE/g. The growing regions, ripening stages, and storage conditions affected the level of bioactive compounds. Dong Thap and Ca Mau provinces highly nourished three-leaf cayratia compared to other sampling locations. The berries exhibited the optimal content of phenolic compounds at the fully ripened stage. The frozen storage favorably accumulated the highest polyphenols in 7 days. The results of this study provided essential information for determining the location, harvesting time and storage conditions for processing products from the three-leaf cayratia with high biological activity due to high polyphenol content. Additionally, the higher content of resveratrol (1.46 mg/g wet weight) than that of quercetin (0.14 mg/g wet weight) has been found in the CT berries. Further researches on other polyphenol targets in the CT products are important to confirm its valuable source for food and health.

DECLARATION OF CONFLICTS OF INTEREST

The authors declare that this research has no conflicts of interest.

REFERENCES

- [1] D. Kumar, S. Kumar, J. Gupta, R. Arya, and A. Gupta, "A review on chemical and biological properties of Cayratia trifolia Linn. (Vitaceae)," *Pharmacognosy Reviews*, vol. 5, no. 10, pp. 184-188, 2011.
- [2] T. K. T. Doan, T. N. M. Huynh, D. D. Nguyen, T. T. Ha, and T. P. D. Ngo, "Total polyphenol content and antioxidant capacity of Cayratia trifolia (L.) Domin berries before and after fermentation using thermotolerant yeast *Saccharomyces cerevisiae* HG1.3," *Vietnam Journal of Science and Technology*, vol. 60, no. 8, pp. 44-60, 2018.
- [3] S. Sowmya, P. C. Perumal, P. Anusooriya, B. Vidya, P. Pratibha, and V. K. Gopalakrishnan, "In vitro antioxidant activity, in vivo skin irritation studies and HPTLC analysis of Cayratia trifolia (L.) Domin," *International Journal of Pharmacology Research*, vol. 7, no. 1, pp. 1-9, 2015.
- [4] M. I. Yusuf, Wahyuni, S. Susanty, Ruslan, and M. Fawwaz, "Antioxidant and antidiabetic potential of galing stem extract (Cayratia trifolia Domin)," *Pharmacognosy Journal*, vol. 10, no. 4, pp. 686-689, 2018.
- [5] M. Yunus, E. Suprihati, and A. Wijaya, "Assessment of relationship between antioxidant activity, toxicity and phenol content of Cayratia trifolia ethanolic extract," *Systematic Reviews in Pharmacy*, vol. 12, no. 1, pp. 1261-1266, 2021.
- [6] B. Meganathan, C. P. Palanisamy, and M. Panagal, "Antioxidant, antimicrobial and cytotoxicity potential of n-hexane extract of Cayratia trifolia L.," *Bioinformation*, vol. 17, no. 3, pp. 452-459, 2021.
- [7] M. Y. Alkandahri, Y. E. Maulana, A. N. Subarnas, A. L. Kwarteng, and A. F. Berbudi, "Antimalarial activity of extract and fractions of Cayratia trifolia (L.) Domin," *International Journal of Pharmaceutics*, vol. 12, no. 1, pp. 1435-1441, 2020.
- [8] S. Hazra, A. S. Ray, S. Das, A. D. Gupta, and C. H. Rahaman, "Phytochemical profiling, biological activities, and in silico molecular docking studies of Causonis trifolia (L.) Mabb. & J. Wen shoot," *Plants*, vol. 12, no. 7, p. 1495, 2023.

- [9] S. P. Rabeta and S. P. Lin, "Effects of different drying methods on the antioxidant activities of leaves and berries of *Cayratia trifolia*," *Sains Malaysiana*, vol. 44, no. 2, pp. 275-280, 2015.
- [10] N. M. Eid, B. A. Awadi, D. Vauzour, M. J. O. Concha, and J. P. Spencer, "Effect of cultivar type and ripening on the polyphenol content of date palm fruit," *Journal of Agricultural and Food Chemistry*, vol. 61, no. 10, pp. 2453–2460, 2013.
- [11] Q. Li, X. X. Chang, H. Wang, C. S. Brennan, and X. B. Guo, "Phytochemicals accumulation in Sanhua plum (*Prunus salicina* L.) during fruit development and their potential use as antioxidants," *Journal of Agricultural and Food Chemistry*, vol. 67, no. 9, pp. 2459-2466, 2019.
- [12] A. Ndou, P. P. Tinyani, R. M. Slabbert, Y. Sultanbawa, and D. Sivakumar, "An integrated approach for harvesting Natal plum (*Carissa macrocarpa*) for quality and functional compounds related to maturity stages," *Food Chemistry*, vol. 293, pp. 499-510, 2019.
- [13] J. Gull, B. Sultana, F. Anwar, R. Naseer, M. Ashraf, and M. Ashrafuzzaman, "Variation in antioxidant attributes at three ripening stages of guava (*Psidium guajava* L.) fruit from different geographical regions of Pakistan," *Molecules*, vol. 17, no. 3, pp. 3165–3180, 2012.
- [14] Y. Jiao, D. Chen, M. Fan, and S. Y. Quek, "UPLC-QqQ-MS/MS-based phenolic quantification and antioxidant activity assessment for thinned young kiwifruits," *Food Chemistry*, vol. 281, pp. 97–105, 2019.
- [15] G. Bombai, F. Pasini, V. Verardo, O. Sevindik, M. Di Foggia, and P. Tessarin, "Monitoring of compositional changes during berry ripening in grape seed extracts of cv. Sangiovese (*Vitis vinifera* L.)," *Journal of the Science of Food and Agriculture*, vol. 97, no. 9, pp. 3058-3064, 2017.
- [16] P. Bertolini, U. Vrhovsek, and E. Baraldi, "Polyphenols variation in fruits of the susceptible strawberry cultivar Alba during ripening and upon fungal pathogen interaction," *Journal of the Science of Food and Agriculture*, vol. 64, no. 9, pp. 1869-1878, 2016.
- [17] Z. Zorenc, R. Veberic, D. Koron, S. Miosic, O. S. Hutabarat, and H. Halbwirth, "Polyphenol metabolism in differently colored cultivars of red currant (*Ribes rubrum* L.) through fruit ripening," *Planta*, vol. 246, no. 2, pp. 217-226, 2017.
- [18] R. Kobori, S. Yakami, T. Kawasaki, and A. Saito, "Changes in the polyphenol content of red raspberry fruits during ripening," *Horticulturae*, vol. 7, no. 12, p. 569, 2021.
- [19] X. Li, J. Sun, Z. Chen, J. Jiang, and A. Jackson, "Characterization of carotenoids and phenolics during fruit ripening of Chinese raspberry (*Rubus chingii* Hu)," *RSC Advances*, vol. 11, no. 18, pp. 10804–10813, 2021.
- [20] M. C. Johnson, A. L. Thomas, and C. M. Greenlief, "Impact of frozen storage on the anthocyanin and polyphenol contents of American elderberry fruit juice," *Journal of Agricultural and Food Chemistry*, vol. 63, no. 23, pp. 5653–5659, 2015.
- [21] O. A. Fawole and U. L. Opara, "Effects of storage temperature and duration on physiological responses of pomegranate fruit," *Industrial Crops and Products*, vol. 47, pp. 300–309, 2013.
- [22] G. L. S. Orbea, R. G. Villalba, M. J. Bernal, A. H. Jiménez, J. A. Egea, and F. A. T. Barberán, "Effect of storage conditions on the stability of polyphenols of apple and strawberry purees produced at industrial scale," *Journal of Agricultural and Food Chemistry*, vol. 71, no. 5, pp. 2541–2553, 2023.
- [23] V. L. Singleton and J. A. Rossi, "Colorimetry of total phenolic compounds with phosphomolybdic-phosphotungstic acid reagents," *American Journal of Enology and Viticulture*, vol. 16, no. 3, pp. 144–158, 1965.
- [24] M. Alothman, R. Bhat, and A. A. Karim, "Antioxidant capacity and phenolic content of selected tropical fruits from Malaysia, extracted with different solvents," *Food Chemistry*, vol. 115, no. 3, pp. 785–788, 2009.
- [25] E. Brighenti et al., "Total polyphenols contents in different grapevine varieties in highlands

- of southern Brazil,” *BIO Web of Conferences*, vol. 9, p. 01024, 2017.
- [26] T. N. Lai et al., “Nutritional composition and antioxidant properties of the sim fruit (*Rhodomyrtus tomentosa*),” *Food Chemistry*, vol. 168, pp. 410-416, 2015.
- [27] K. A. Reynertson, H. Yang, B. Jiang, M. J. Basile, and E. J. Kennelly, “Quantitative analysis of antiradical phenolic constituents from fourteen edible Myrtaceae fruits,” *Food Chemistry*, vol. 109, no. 4, pp. 883-890, 2008.
- [28] D. T. K. Tien et al., “The change of total polyphenol content of three-leaf cayratia (*Cayratia trifolia* L.) by harvest period and juice treatment,” *TNU Journal of Science and Technology*, vol. 228, no. 13, pp. 366-373, 2023.
- [29] X. Guofang, X. Xiaoyan, Z. Xiaoli, L. Yongling, and Z. Zhibing, “Changes in phenolic profiles and antioxidant activity in rabbiteye blueberries during ripening,” *International Journal of Food Properties*, vol. 22, no. 1, pp. 320-329, 2019.
- [30] T. Mahmood et al., “Influence of ripening stages and drying methods on polyphenolic content and antioxidant activities of mulberry fruits,” *Food Measurement*, vol. 11, pp. 2171-2179, 2017.
- [31] L. Kapoor, A. J. Simkin, C. G. P. Doss, and R. Siva, “Fruit ripening: dynamics and integrated analysis of carotenoids and anthocyanins,” *BMC Plant Biology*, vol. 22, no. 1, p. 27, 2022.
- [32] D. T. K. Tien et al., “Fermentation conditions, total polyphenol content, and antioxidant activity of three-leaf cayratia (*Cayratia trifolia* L.) wine prepared using thermotolerant yeast *Saccharomyces cerevisiae* HG1.3,” *Asia-Pacific Journal of Science and Technology*, vol. 27, no. 6, 2022.
- [33] T. T. Doan et al., “The fermentation conditions of low alcoholic three-leaved (*Cayratia trifolia* (L.) Domin) cider using *Saccharomyces cerevisiae* HG1.3,” *Journal of Applied Biology and Biotechnology*, vol. 12, no. 5, pp. 170-176, 2024.
- [34] Y. Grover and P. S. Negi, “Recent developments in freezing of fruits and vegetables: Striving for controlled ice nucleation and crystallization with enhanced freezing rates,” *Journal of Food Science*, vol. 88, no. 12, pp. 4799-4826, 2023.
- [35] J. Arora, C. Roat, S. Goyal, and K. Ramawat, “High stilbenes accumulation in root cultures of *Cayratia trifolia* (L.) Domin grown in shake flasks,” *Acta Physiologiae Plantarum*, vol. 31, pp. 1307-1312, 2009.
- [36] M. Careri, C. Corradini, L. Elviri, I. Nicoletti, and I. Zagnoni, “Direct HPLC analysis of quercetin and trans-resveratrol in red wine, grape, and winemaking byproducts,” *Journal of Agricultural and Food Chemistry*, vol. 51, no. 18, pp. 5226-5231, 2003.
- [37] D. D. Rio, G. Borges, and A. Crozier, “Berry flavonoids and phenolics: bioavailability and evidence of protective effects,” *British Journal of Nutrition*, vol. 104, no. 3, pp. S67-S90, 2010.
- [38] A. Vollmannová, T. Bojnanská, J. Musilová, J. Lidiková, and M. Cifrová, “Quercetin as one of the most abundant represented biological valuable plant components with remarkable chemoprotective effects - A review,” *Heliyon*, vol. 10, no. 12, p. e33342, 2024.



© 2026 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Ảnh hưởng của độ chín và điều kiện bảo quản đến thành phần polyphenol của trái giác (*Cayratia trifolia*) từ các vùng địa lý khác nhau ở Việt Nam

Đoàn Thị Kiều Tiên^{1,*}, Trần Diễm Phượng¹, Huỳnh Thị Ngọc Mi¹,
Đinh Hoàng Lan Chi¹, Nguyễn Ngọc Thanh², Huỳnh Xuân Phong²

¹Trường Đại học Kỹ Thuật - Công nghệ Cần Thơ, Việt Nam

²Đại học Cần Thơ, Việt Nam

Ngày nhận bài: 27/10/2025; Ngày sửa bài: 03/02/2026;

Ngày nhận đăng: 06/03/2026; Ngày xuất bản: 28/06/2026

TÓM TẮT

Nghiên cứu này nhằm khảo sát ảnh hưởng của địa điểm thu mẫu, độ chín và điều kiện bảo quản đến tính chất hóa lý, bao gồm hàm lượng polyphenol tổng số trong trái giác (*Cayratia trifolia*) được thu thập từ 9 tỉnh Đồng bằng sông Cửu Long tại Việt Nam như Kiên Giang, An Giang, Đồng Tháp, Long An, Bạc Liêu, Cà Mau, Sóc Trăng, Trà Vinh và Bến Tre. Kỹ thuật sắc ký lỏng hiệu năng cao được sử dụng để định lượng hàm lượng quercetin và resperatrol trong mẫu trái giác. Kết quả cho thấy trái giác có hàm lượng polyphenol cao từ 12,14 đến 19,51 mg GAE/g, cao nhất ở mẫu thu thập tại Đồng Tháp (DT1) (19,51 mg GAE/g). Trong giai đoạn chín từ lúc chưa chín đến chín hoàn toàn, hàm lượng polyphenol tổng số của mẫu trái giác thu ở An Giang (AG3), Trà Vinh (TV4), Đồng Tháp (DT1, DT2), Cà Mau (CM1, CM2) tăng lên trong khi hàm lượng này lại giảm ở các mẫu thu thập tại Kiên Giang (KG1), DT3, DT4, CM3 và CM4. Sự khác biệt này có thể là do sự thay đổi trong quá trình tích lũy hợp chất phenolic ở quá trình chín của trái, vị trí địa lý và điều kiện trồng trọt. Bên cạnh đó, trái giác đông lạnh có hàm lượng polyphenol tổng số không thay đổi đáng kể (18,62 mg GAE/g) trong thời gian bảo quản 5 ngày, trong khi giá trị này tăng lên 19,05 mg GAE/g sau 7 ngày bảo quản. Kết quả phân tích HPLC cũng cho thấy hàm lượng quercetin và resveratrol trong mẫu DT lần lượt là $0,14 \pm 0,012$ mg và $1,46 \pm 0,035$ mg/g trọng lượng tươi.

Từ khóa: *Cayratia trifolia*, polyphenol, quercetin, resveratrol.

*Tác giả liên hệ chính.

Email: dtktien@ctu.edu.vn

Robust design optimization for overhead power transmission line design

Pham Van Phat¹, Tran Ngoc Thanh¹, Tuan-Ho Le^{2,*}

¹Department of Electrical Engineering, Quy Nhon College of Engineering and Technology, Vietnam

²Faculty of Engineering and Technology, Quy Nhon University, Vietnam

Received: 23/12/2025; Revised: 02/02/2026;

Accepted: 14/02/2026; Published: 28/06/2026

ABSTRACT

Most practical design methods for overhead power transmission lines primarily rely on empirical formulas. This paper proposes a robust design-based two-stage optimization framework for overhead power transmission line design. In the first stage, corresponding to the design of experiments or experimental design, four control factors—namely conductor type, conductor size, span length, and pretension—and two noise factors—ambient temperature and wind speed—are considered, each at three levels. Three output responses—maximum sag, power loss, and life cycle cost—are evaluated. The control factors and noise factors are arranged in the inner and outer arrays, respectively, using a full factorial design. These arrays are then combined to form a full factorial L_{81} crossed-array experimental design. In the second stage, desirability functions are employed to simultaneously minimize the three output responses. Finally, a comparative analysis of different desirability formulations is conducted, and based on the obtained results, the optimal design solutions for the overhead transmission line are identified.

Keywords: Robust design, overhead power transmission line design, full factorial design, desirability functions.

1. INTRODUCTION

Overhead power transmission lines play a critical role in electric power systems by enabling the transmission and distribution of electrical energy from generation sources to consumption areas. Compared with alternative transmission technologies, overhead lines offer significant advantages in terms of lower investment cost, scalability, and ease of operation and maintenance. Consequently, they remain the predominant solution for both transmission and distribution networks across a wide range of voltage levels. The quality of transmission line design has a direct impact on supply reliability, power losses, operational safety, and the overall economic efficiency of the power system. The design of overhead power transmission lines

constitutes a complex engineering problem involving coupled electrical, mechanical, structural, and environmental factors. Key design variables include conductor type and cross-sectional area, span length, pole or tower height, phase configuration, insulation system, as well as conductor sag and tension. These parameters are strongly interdependent rather than independent; variations in a single parameter can simultaneously influence current-carrying capability, electrical clearances, mechanical integrity, and both construction and operational costs. As a result, transmission line design requires a holistic and systematic consideration of multiple interacting constraints. In practical engineering applications, transmission line design is commonly conducted in accordance

*Corresponding author.

Email: tuanhole@qnu.edu.vn

with technical standards and deterministic calculation methods, with conservative assumptions adopted to ensure safety margins and operational reliability. Although such approaches provide robust designs, they may result in suboptimal solutions characterized by excessive material usage, increased capital expenditure, and higher energy losses. These limitations are further exacerbated by growing transmission demand and increasingly variable environmental conditions. Moreover, uncertainties associated with ambient temperature, wind loading, ice accretion, and long-term material aging introduce additional challenges in both the design process and the accurate assessment of transmission line performance.

Extensive research efforts have been devoted to improving the modeling and analysis of overhead transmission line behavior. In sag–tension and mechanical design, detailed analytical procedures accounting for thermal states, conductor creep, and insulator string characteristics have been developed [1]. Modified ruling span methods have been proposed to calculate sag and tension in multi-span line segments while considering the rotational stiffness of suspension insulator strings [2]. Field measurements and statistical analyses have been used to evaluate discrepancies between designed and actual sag, highlighting the effects of installation practices and modeling uncertainties on design accuracy [3]. Reduced-order, least-squares-based state estimators have also been introduced to estimate sag in leveled span configurations, providing useful tools for mechanical design validation and conductor behavior monitoring [4]. Furthermore, sag–tension relationships under maximum wind loading conditions have been investigated to assess mechanical reliability under extreme weather events [5]. In terms of thermal modeling and ampacity assessment, computational fluid dynamics-based thermal models have demonstrated adequate accuracy in predicting conductor temperature, offering a viable approach for dynamic analysis of

distribution networks operating under variable load and weather conditions [6]. The application of high-temperature low-sag conductors and their design implications for capacity enhancement have been evaluated in accordance with international standards such as those issued by CIGRE and IEEE [7]. Transmission line ratings based on the relationship between conductor temperature and sag have been determined using the ruling span principle and tension monitoring systems [8]. For design validation and condition monitoring, predictive sag–load models incorporating temperature effects and mechanical deformation have been proposed [9], while vibration-based tension estimation techniques using accelerometer measurements have been introduced to support mechanical monitoring [10]. In addition, non-contact electric field sensing methods have been applied to transmission line sag measurement and simulation [11]. Advanced conductor technologies, including ACCC, ACSS, and ACCR conductors, have been compared in terms of ampacity and sag performance using reconductoring approaches relevant to design optimization [12]. A comprehensive review of conductor design parameters and associated monitoring technologies is provided [13]. Despite these advances, conventional transmission line design practice continues to rely largely on deterministic procedures and empirically derived design rules. In recent years, optimization-based approaches have attracted increasing attention as effective tools for improving transmission line planning and design. Cost assessment and optimization models employing genetic algorithms have been proposed to minimize the total cost of overhead transmission lines while satisfying design constraints [14]. Integrated optimization frameworks for route selection and tower placement using graph algorithms and dynamic programming techniques have also been developed [15]. Heuristic optimization methods have been applied to optimize conductor bundle geometry with the objective of enhancing transmission capacity while reducing overall cost [16]. Multi-objective optimization

approaches considering trade-offs between electrical and mechanical performance have been investigated using genetic algorithms [17]. More recently, metaheuristic techniques such as the whale optimization algorithm have been employed to estimate optimal parameters of overhead AC transmission lines, enabling improved tuning of design variables under electrical constraints [18].

Although substantial progress has been made in the modeling, analysis, and optimization of overhead power transmission lines, existing studies predominantly emphasize performance improvement under nominal operating conditions. Mechanical and thermal analyses typically focus on accurate prediction of sag, tension, and ampacity, while optimization-based approaches aim to minimize cost or maximize capacity through advanced search algorithms. However, these studies often assume fixed design parameters and operating environments, thereby limiting their applicability in practical scenarios where uncertainties are unavoidable.

Moreover, existing studies have not yet proposed a unified framework for optimal design calculation of overhead power transmission lines that can be generally applied to different line configurations. Factors such as ambient temperature and wind speed, which can significantly influence transmission line design, are often not explicitly considered in the optimal design process. In addressing the technical optimization problem, the robust design methodology proposed by Taguchi can be employed as an effective solution and may be regarded as an alternative to conventional optimization techniques. In quality engineering, products or processes are commonly modeled as a box or system characterized by inputs and outputs. Taguchi classifies the input variables into two categories: control factors (denoted by x) and noise factors (denoted by z). Taguchi proposed a two-step procedure for solving optimal design problems, consisting of the construction of a design of experiments and the use of signal-to-noise ratio (SNR) analysis

to identify optimal solutions. In the design of experiments stage, Taguchi introduced a set of orthogonal arrays instead of traditional experimental design methods. Among more than 60 types of SNRs, three have been most widely adopted: nominal-is-best, larger-is-better, and smaller-is-better. However, the orthogonal arrays and SNR formulations proposed by Taguchi have been criticized by statisticians and researchers in the literature [19] - [21].

Based on the Taguchi philosophy, this paper develops a general multi-objective optimization framework for the design of overhead power transmission lines. In the experimental design stage, a general full factorial design is employed with four control factors-namely conductor type, conductor size, span length, and pretension-each considered at three levels, and two noise factors-ambient temperature and wind speed-also considered at three levels. Three output responses are evaluated: maximum sag, power loss, and life cycle cost (LCC). With the objective of minimizing these three responses, several desirability functions are applied to identify optimal solutions. The results obtained from different desirability functions are compared to identify the best solutions. To the best of our knowledge, this is the first attempt to apply the Taguchi philosophy in this area. An overview of the proposed robust design-based two-stage optimization framework for overhead transmission line design is shown in Figure 1.

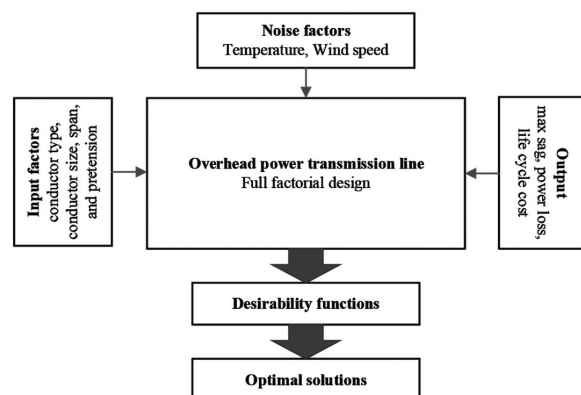


Figure 1. Proposed robust design-based two-stage optimization framework for overhead transmission line design.

2. PROPOSED ROBUST DESIGN OPTIMIZATION

This section presents the proposed robust design-based two-stage optimization framework for overhead power transmission line design. The methodology integrates the Taguchi robust design philosophy with an advanced desirability function-based optimization algorithm to achieve design solutions that are not only optimal under nominal conditions but also robust against environmental and operational uncertainties.

2.1. Design of experiments stage

The overhead transmission line design problem is formulated as a multi-input–multi-output engineering system, in which a set of controllable design variables and uncontrollable noise factors jointly influence the performance of the transmission line. The objective is to determine the optimal combination of design parameters that minimizes key performance indices while maintaining robustness under varying environmental conditions.

Unlike the classical Taguchi approach that relies on predefined orthogonal arrays, this study employs a full factorial experimental design to comprehensively capture interactions among control and noise factors. This choice ensures accurate modeling of complex interdependencies between electrical and mechanical parameters that are common in overhead transmission line systems. For each experimental run, the transmission line responses are evaluated under all combinations of noise factor levels, allowing the influence of environmental variability on system performance to be explicitly quantified.

Based on engineering practice, four control factors reflecting realistic engineering choices are selected: conductor type (x_1), conductor size (x_2), span (x_3), and pretension (x_4). To account for environmental uncertainty, two noise factors, representing typical, adverse, and extreme operating conditions, are incorporated: ambient temperature (z_1) and wind speed (z_2).

Three output responses are selected to evaluate transmission line performance:

- ✓ Maximum conductor sag (y_1), which directly affects electrical clearance and mechanical safety;
- ✓ Power loss (y_2), representing the electrical efficiency of the line;
- ✓ LCC (y_3), which includes material, construction, operation, and maintenance costs.

The four control factors were selected based on standard overhead transmission line design practice, as they are directly controllable by designers and have a dominant influence on both electrical and mechanical performance. The two noise factors represent key environmental uncertainties that cannot be controlled in operation but significantly affect sag, resistance, and power loss. Each factor was defined at three levels to represent low, nominal, and high operating conditions, allowing nonlinear trends and robustness effects to be captured while maintaining a manageable experimental size.

2.1.1. Maximum conductor sag

For typical overhead transmission spans (≤ 500 m), the conductor is modeled as a flexible cable subjected to uniformly distributed loading. Under normal operating conditions, the parabolic approximation is widely accepted and yields sufficiently accurate results for design and optimization studies.

The mid-span sag is expressed as:

$$f = \frac{\omega L^2}{8H} \quad (1)$$

where f is mid-span sag (m), ω is resultant weight per unit length (N/m), L is span length (m), and H is horizontal component of conductor tension (N).

The weight per unit length is given by:

$$\omega = mg \quad (2)$$

where m is the conductor mass per unit length (kg/m) and g is the gravitational acceleration (9.81 m/s²).

The horizontal tension is determined by the pretension ratio:

$$H = \eta H_{rated} \quad (3)$$

where η denotes the pretension level and H_{rated} is the rated horizontal tension of the conductor.

To account for environmental uncertainties such as temperature variation and wind speed, the maximum sag is corrected as

$$f_{max} = \frac{\omega L^2}{8H} [1 + \alpha_s (T - T_0) + \beta_\omega V_\omega] \quad (4)$$

where T is ambient temperature ($^{\circ}\text{C}$), T_0 is reference temperature (typically 20°C), V_ω is wind speed (m/s), and α_s , β_ω are temperature and wind correction coefficients.

The parabolic approximation employed in this study assumes relatively small sag-to-span ratios and uniformly distributed loading, which is generally valid for typical transmission spans under normal operating conditions. For very long spans or extreme weather events involving high wind speeds, ice accretion, or large temperature gradients, catenary effects and nonlinear geometric behavior may become significant. In such cases, the simplified sag model may introduce deviations, and more advanced catenary-based or finite-element models should be used in detailed design stages.

2.1.2. Power loss

The active power loss in an overhead conductor is governed by Joule heating and is expressed as:

$$P_{loss} = I^2 R(T) \quad (5)$$

where I is the line current (A) and $R(T)$ is the conductor resistance at operating temperature T .

The temperature-dependent resistance is modeled as:

$$R(T) = R_{20} [1 + \alpha_{20} (T - 20)] \quad (6)$$

where R_{20} is the resistance at 20°C and α_{20} is the temperature coefficient of resistance. For a given span length L , the total power loss is:

$$P_{loss,span} = I^2 R(T) \frac{L}{1000} \quad (7)$$

2.1.3. Life cycle cost

The LCC of an overhead transmission line is defined as the sum of capital investment, energy loss cost, and maintenance cost over the service lifetime [22].

$$LCC = C_{capital} + C_{loss} + C_{main} \quad (8)$$

The capital cost is primarily determined by the conductor material and size:

$$C_{capital} = C_{cond} (x_1, x_2) \quad (9)$$

The cumulative energy loss cost over the service life is calculated as:

$$C_{loss} = P_{loss} h_{life} C_{energy} \quad (10)$$

where h_{life} denotes the total operating hours over the service life and C_{energy} is the unit cost of electrical energy.

Mechanical stress and increased sag result in higher inspection and maintenance requirements. A linear sag-dependent maintenance cost model is adopted:

$$C_{main} = k_m f_{max} \quad (11)$$

where k_m is maintenance cost coefficient.

A linear sag–maintenance cost relationship is adopted to maintain model simplicity and computational efficiency. Increased sag typically leads to higher inspection and clearance management efforts, for which linear approximations are widely used in preliminary economic analyses. The model is therefore intended for comparative design assessment within normal operating sag ranges.

The LCC formulation is a simplified model for robust, preliminary design optimization and does not explicitly account for stochastic failures, aging, or condition-based maintenance. The framework therefore supports early-stage decision-making under uncertainty rather than detailed reliability or asset management analyses.

2.2. Optimization stage

The design problem is therefore formulated as a multi-objective minimization problem:

$$\min\{y_1(\mathbf{x}, \mathbf{z}), y_2(\mathbf{x}, \mathbf{z}), y_3(\mathbf{x}, \mathbf{z})\} \quad (12)$$

subject to $\mathbf{x}^{\min} \leq \mathbf{x} \leq \mathbf{x}^{\max}$

where $\mathbf{x}$ represents the vector of control factors (conductor type, conductor size, span, pretension).

Three different types of desirability functions for three different cases are as follows:

If the target is the best, the desirability function is

$$d_i(y_i) = \begin{cases} 0 & \text{if } y_i < L_i \\ \left(\frac{y_i - L_i}{T_i - L_i}\right)^s & \text{if } L_i \leq y_i \leq T_i \\ \left(\frac{y_i - U_i}{T_i - U_i}\right)^t & \text{if } T_i \leq y_i \leq U_i \\ 0 & \text{if } y_i > U_i \end{cases} \quad (13)$$

If the maximize is better, the desirability function is

$$d_i(y_i) = \begin{cases} 0 & \text{if } y_i < L_i \\ \left(\frac{y_i - L_i}{T_i - L_i}\right)^s & \text{if } L_i \leq y_i \leq T_i \\ 1 & \text{if } y_i > T_i \end{cases} \quad (14)$$

If the minimize is better, the desirability function is

$$d_i(y_i) = \begin{cases} 1 & \text{if } y_i < T_i \\ \left(\frac{y_i - U_i}{T_i - U_i}\right)^t & \text{if } T_i \leq y_i \leq U_i \\ 0 & \text{if } y_i > U_i \end{cases} \quad (15)$$

where L_i , T_i , and U_i are the lower, target, and upper values of the i^{th} function, respectively. r , s , and t are the exponents of each desirability function. The lower, target, and upper values of the desirability functions were determined based on acceptable engineering limits, relevant design standards, and the observed response ranges obtained from the experimental design. The exponents (r , s , t) were selected to reflect the relative importance of each performance objective and the severity of penalties for

deviations from desired values. To assess the influence of desirability formulation, different desirability functions were compared, providing insight into the sensitivity of the optimal solution to these parameters.

The overall desirability function is shown as

$$D = [d_1(y_1)d_2(y_2)...d_k(y_k)]^{1/k} \quad (16)$$

where k is the number of responses.

3. CASE STUDY

In this study, a 22-kV overhead power transmission line is selected as the case study to demonstrate the proposed optimization framework. For the experimental design, each control factor and each noise factor is defined at three discrete and practically meaningful levels, enabling a systematic evaluation of their effects on the transmission line performance as:

- ✓ Conductor type (x_1): ACSR, AAAC, and ACCC (coded value: 1, 2, 3);
- ✓ Conductor size (x_2): 95, 150, and 240 (mm²) (coded value: 1, 2, 3);
- ✓ Span (x_3): 100, 150, and 200 (m) (coded value: 1, 2, 3);
- ✓ Pretension (x_4): 0.3, 0.5, and 0.7 (of rated) (coded value: 1, 2, 3).

Similarly, each noise factor is defined at three discrete levels to reflect realistic variations in environmental and operating conditions encountered in practice, as summarized as follows:

- ✓ Ambient temperature (z_1): -10, 20, and 40 (°C) (coded value: 1, 2, 3);
- ✓ Wind speed (z_2): 0, 15, and 30 (m/s) (coded value: 1, 2, 3).

The parameters of the ACSR, AAAC, and ACCC conductors with different sizes are summarized in Tables 1, 2, and 3, respectively, including their electrical, mechanical, and physical characteristics relevant to transmission line design and performance evaluation.

Table 1. Parameters of conductor ACSR.

Size (mm ²)	Mass (kg/m)	R20 (Ω/km)	Cost (USD/km)
95	0.38	0.193	8000
150	0.62	0.126	12000
240	0.98	0.078	18000

Table 2. Parameters of conductor AAAC.

Size (mm ²)	Mass (kg/m)	R20 (Ω/km)	Cost (USD/km)
95	0.35	0.195	8500
150	0.58	0.127	12500
240	0.92	0.079	18500

Table 3. Parameters of conductor ACCC.

Size (mm ²)	Mass (kg/m)	R20 (Ω/km)	Cost (USD/km)
95	0.32	0.50	11000
150	0.50	0.125	15000
240	0.82	0.077	23000

Table 4. Full factorial design L₈₁ array for maximum conductor sag.

					Runs	Outer array									
						1	2	3	4	5	6	7	8	9	
					z_1	1	2	3	1	2	3	1	2	3	
					z_2	1	1	1	2	2	2	3	3	3	
Runs	Inner array						Maximum conductor sag (m)								
	x_1	x_2	x_3	x_4	y_{11}		y_{12}	y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}	y_{19}	
1	1	1	1	1	0.3105		0.3107	0.3107	0.3338	0.3339	0.3340	0.3571	0.3572	0.3573	
2	1	1	1	2	0.1863		0.1864	0.1864	0.2003	0.2004	0.2004	0.2143	0.2143	0.2144	
3	1	1	1	3	0.1331		0.1331	0.1332	0.1431	0.1431	0.1432	0.1531	0.1531	0.1531	
4	1	1	2	1	0.6987		0.6990	0.6991	0.7511	0.7514	0.7516	0.8036	0.8038	0.8040	
5	1	1	2	2	0.4192		0.4194	0.4195	0.4507	0.4508	0.4509	0.4821	0.4823	0.4824	
6	1	1	2	3	0.2994		0.2996	0.2996	0.3219	0.3220	0.3221	0.3444	0.3445	0.3446	
7	1	1	3	1	1.2422		1.2426	1.2429	1.3353	1.3358	1.3361	1.4285	1.4290	1.4293	
8	1	1	3	2	0.7453		0.7456	0.7457	0.8012	0.8015	0.8017	0.8571	0.8574	0.8576	
9	1	1	3	3	0.5324		0.5325	0.5327	0.5723	0.5725	0.5726	0.6122	0.6124	0.6126	
10	1	2	1	1	0.5067		0.5069	0.5070	0.5447	0.5449	0.5450	0.5827	0.5829	0.5830	
11	1	2	1	2	0.3040		0.3041	0.3042	0.3268	0.3269	0.3270	0.3496	0.3497	0.3498	
12	1	2	1	3	0.2171		0.2172	0.2173	0.2334	0.2335	0.2336	0.2497	0.2498	0.2499	
13	1	2	2	1	1.1400		1.1404	1.1407	1.2255	1.2259	1.2262	1.3111	1.3115	1.3117	
14	1	2	2	2	0.6840	0.6842	0.6844	0.7353	0.7356	0.7357	0.7866	0.7869	0.7870		
15	1	2	2	3	0.4886	0.4887	0.4889	0.5252	0.5254	0.5255	0.5619	0.5621	0.5622		

Moreover, the additional parameters used in the overhead power transmission line design are as follows: the rated line current is 300 A; the gravitational acceleration is 9.81 m/s²; the rated horizontal tension is 50000 N; the thermal sag coefficient is 1.2×10⁵ °C⁻¹; the wind sag coefficient is 0.005 (m/s)⁻¹; the resistance temperature coefficient is 0.004 °C⁻¹, the service life hours is 30 years, corresponding to 262800 h; the energy price is 0.10 USD/kWh, and the maintenance cost factor is 1000 USD/m.

Based on the selected control and noise factors and their associated coded levels, a full factorial L₈₁ crossed-array experimental design is constructed for the three output responses, as summarized in Tables 4, 5, and 6. In this crossed-array structure, the inner array represents all possible combinations of the control factors, while the outer array accounts for the combinations of the noise factors, enabling a systematic evaluation of the robustness of the design against environmental and operational variations. This design framework facilitates the simultaneous assessment of factor interactions and variability effects on the system performance.

16	1	2	3	1	2.0267	2.0274	2.0279	2.1787	2.1795	2.1799	2.3308	2.3315	2.3320
17	1	2	3	2	1.2160	1.2164	1.2167	1.3072	1.3077	1.3080	1.3985	1.3989	1.3992
18	1	2	3	3	0.8686	0.8689	0.8691	0.9337	0.9341	0.9343	0.9989	0.9992	0.9994
19	1	3	1	1	0.8009	0.8012	0.8013	0.8609	0.8612	0.8614	0.9210	0.9213	0.9215
20	1	3	1	2	0.4805	0.4807	0.4808	0.5166	0.5167	0.5169	0.5526	0.5528	0.5529
21	1	3	1	3	0.3432	0.3434	0.3434	0.3690	0.3691	0.3692	0.3947	0.3949	0.3949
22	1	3	2	1	1.8019	1.8026	1.8030	1.9371	1.9378	1.9382	2.0723	2.0730	2.0734
23	1	3	2	2	1.0812	1.0816	1.0818	1.1623	1.1627	1.1629	1.2434	1.2438	1.2440
24	1	3	2	3	0.7723	0.7725	0.7727	0.8302	0.8305	0.8307	0.8881	0.8884	0.8886
25	1	3	3	1	3.2034	3.2046	3.2054	3.4438	3.4449	3.4457	3.6841	3.6853	3.6861
26	1	3	3	2	1.9221	1.9228	1.9232	2.0663	2.0670	2.0674	2.2105	2.2112	2.2116
27	1	3	3	3	1.3729	1.3734	1.3737	1.4759	1.4764	1.4767	1.5789	1.5794	1.5797
28	2	1	1	1	0.2860	0.2861	0.2862	0.3075	0.3076	0.3077	0.3289	0.3290	0.3291
29	2	1	1	2	0.1716	0.1717	0.1717	0.1845	0.1846	0.1846	0.1974	0.1974	0.1975
30	2	1	1	3	0.1226	0.1226	0.1227	0.1318	0.1318	0.1319	0.1410	0.1410	0.1410
31	2	1	2	1	0.6435	0.6438	0.6439	0.6918	0.6921	0.6922	0.7401	0.7403	0.7405
32	2	1	2	2	0.3861	0.3863	0.3864	0.4151	0.4152	0.4153	0.4441	0.4442	0.4443
33	2	1	2	3	0.2758	0.2759	0.2760	0.2965	0.2966	0.2967	0.3172	0.3173	0.3174
34	2	1	3	1	1.1441	1.1445	1.1448	1.2299	1.2303	1.2306	1.3158	1.3162	1.3164
35	2	1	3	2	0.6865	0.6867	0.6869	0.7380	0.7382	0.7384	0.7895	0.7897	0.7899
36	2	1	3	3	0.4903	0.4905	0.4906	0.5271	0.5273	0.5274	0.5639	0.5641	0.5642
37	2	2	1	1	0.4740	0.4742	0.4743	0.5095	0.5097	0.5098	0.5451	0.5453	0.5454
38	2	2	1	2	0.2844	0.2845	0.2846	0.3057	0.3058	0.3059	0.3271	0.3272	0.3272
39	2	2	1	3	0.2031	0.2032	0.2033	0.2184	0.2184	0.2185	0.2336	0.2337	0.2337
40	2	2	2	1	1.0665	1.0668	1.0671	1.1465	1.1469	1.1471	1.2265	1.2269	1.2271
41	2	2	2	2	0.6399	0.6401	0.6403	0.6879	0.6881	0.6883	0.7359	0.7361	0.7363
42	2	2	2	3	0.4571	0.4572	0.4573	0.4913	0.4915	0.4916	0.5256	0.5258	0.5259
43	2	2	3	1	1.8959	1.8966	1.8971	2.0382	2.0388	2.0393	2.1804	2.1811	2.1815
44	2	2	3	2	1.1376	1.1380	1.1382	1.2229	1.2233	1.2236	1.3082	1.3087	1.3089
45	2	2	3	3	0.8125	0.8128	0.8130	0.8735	0.8738	0.8740	0.9345	0.9348	0.9349
46	2	3	1	1	0.7518	0.7521	0.7523	0.8082	0.8085	0.8087	0.8646	0.8649	0.8651
47	2	3	1	2	0.4511	0.4513	0.4514	0.4849	0.4851	0.4852	0.5188	0.5189	0.5191
48	2	3	1	3	0.3222	0.3223	0.3224	0.3464	0.3465	0.3466	0.3706	0.3707	0.3708
49	2	3	2	1	1.6916	1.6922	1.6926	1.8185	1.8191	1.8195	1.9454	1.9461	1.9465
50	2	3	2	2	1.0150	1.0153	1.0156	1.0911	1.0915	1.0917	1.1673	1.1676	1.1679
51	2	3	2	3	0.7250	0.7252	0.7254	0.7794	0.7796	0.7798	0.8338	0.8340	0.8342
52	2	3	3	1	3.0073	3.0084	3.0091	3.2329	3.2340	3.2348	3.4586	3.4597	3.4604
53	2	3	3	2	1.8044	1.8050	1.8055	1.9398	1.9404	1.9409	2.0751	2.0758	2.0762
54	2	3	3	3	1.2889	1.2893	1.2896	1.3855	1.3860	1.3863	1.4822	1.4827	1.4830
55	3	1	1	1	0.2615	0.2616	0.2617	0.2811	0.2812	0.2813	0.3007	0.3008	0.3009
56	3	1	1	2	0.1569	0.1570	0.1570	0.1687	0.1687	0.1688	0.1804	0.1805	0.1805
57	3	1	1	3	0.1121	0.1121	0.1121	0.1205	0.1205	0.1205	0.1289	0.1289	0.1290
58	3	1	2	1	0.5884	0.5886	0.5887	0.6325	0.6327	0.6329	0.6767	0.6769	0.6770
59	3	1	2	2	0.3530	0.3532	0.3532	0.3795	0.3796	0.3797	0.4060	0.4061	0.4062
60	3	1	2	3	0.2522	0.2523	0.2523	0.2711	0.2712	0.2712	0.2900	0.2901	0.2902
61	3	1	3	1	1.0460	1.0464	1.0467	1.1245	1.1249	1.1251	1.2030	1.2034	1.2036
62	3	1	3	2	0.6276	0.6278	0.6280	0.6747	0.6749	0.6751	0.7218	0.7220	0.7222
63	3	1	3	3	0.4483	0.4485	0.4486	0.4819	0.4821	0.4822	0.5156	0.5157	0.5158
64	3	2	1	1	0.4086	0.4088	0.4088	0.4393	0.4394	0.4395	0.4699	0.4701	0.4702
65	3	2	1	2	0.2452	0.2453	0.2453	0.2636	0.2636	0.2637	0.2819	0.2820	0.2821
66	3	2	1	3	0.1751	0.1752	0.1752	0.1883	0.1883	0.1884	0.2014	0.2015	0.2015

67	3	2	2	1	0.9194	0.9197	0.9199	0.9883	0.9887	0.9889	1.0573	1.0576	1.0579
68	3	2	2	2	0.5516	0.5518	0.5519	0.5930	0.5932	0.5933	0.6344	0.6346	0.6347
69	3	2	2	3	0.3940	0.3942	0.3942	0.4236	0.4237	0.4238	0.4531	0.4533	0.4534
70	3	2	3	1	1.6344	1.6350	1.6354	1.7570	1.7576	1.7580	1.8797	1.8803	1.8806
71	3	2	3	2	0.9806	0.9810	0.9812	1.0542	1.0546	1.0548	1.1278	1.1282	1.1284
72	3	2	3	3	0.7005	0.7007	0.7009	0.7530	0.7533	0.7534	0.8056	0.8058	0.8060
73	3	3	1	1	0.6701	0.6704	0.6705	0.7204	0.7206	0.7208	0.7707	0.7709	0.7711
74	3	3	1	2	0.4021	0.4022	0.4023	0.4322	0.4324	0.4325	0.4624	0.4625	0.4626
75	3	3	1	3	0.2872	0.2873	0.2874	0.3087	0.3088	0.3089	0.3303	0.3304	0.3305
76	3	3	2	1	1.5077	1.5083	1.5086	1.6209	1.6214	1.6218	1.7340	1.7345	1.7349
77	3	3	2	2	0.9046	0.9050	0.9052	0.9725	0.9728	0.9731	1.0404	1.0407	1.0409
78	3	3	2	3	0.6462	0.6464	0.6466	0.6947	0.6949	0.6950	0.7431	0.7434	0.7435
79	3	3	3	1	2.6804	2.6814	2.6820	2.8815	2.8825	2.8831	3.0826	3.0836	3.0843
80	3	3	3	2	1.6083	1.6088	1.6092	1.7289	1.7295	1.7299	1.8496	1.8502	1.8506
81	3	3	3	3	1.1488	1.1492	1.1494	1.2349	1.2354	1.2356	1.3211	1.3215	1.3218

Table 5. Full factorial design L_{81} array for power loss.

Runs	Power loss (kW)								
	y_{21}	y_{22}	y_{23}	y_{24}	y_{25}	y_{26}	y_{27}	y_{28}	y_{29}
1	1.5286	1.7370	1.8760	1.5286	1.7370	1.8760	1.5286	1.7370	1.8760
2	1.5286	1.7370	1.8760	1.5286	1.7370	1.8760	1.5286	1.7370	1.8760
3	1.5286	1.7370	1.8760	1.5286	1.7370	1.8760	1.5286	1.7370	1.8760
4	2.2928	2.6055	2.8139	2.2928	2.6055	2.8139	2.2928	2.6055	2.8139
5	2.2928	2.6055	2.8139	2.2928	2.6055	2.8139	2.2928	2.6055	2.8139
6	2.2928	2.6055	2.8139	2.2928	2.6055	2.8139	2.2928	2.6055	2.8139
7	3.0571	3.4740	3.7519	3.0571	3.4740	3.7519	3.0571	3.4740	3.7519
8	3.0571	3.4740	3.7519	3.0571	3.4740	3.7519	3.0571	3.4740	3.7519
9	3.0571	3.4740	3.7519	3.0571	3.4740	3.7519	3.0571	3.4740	3.7519
10	0.9979	1.1340	1.2247	0.9979	1.1340	1.2247	0.9979	1.1340	1.2247
11	0.9979	1.1340	1.2247	0.9979	1.1340	1.2247	0.9979	1.1340	1.2247
12	0.9979	1.1340	1.2247	0.9979	1.1340	1.2247	0.9979	1.1340	1.2247
13	1.4969	1.7010	1.8371	1.4969	1.7010	1.8371	1.4969	1.7010	1.8371
14	1.4969	1.7010	1.8371	1.4969	1.7010	1.8371	1.4969	1.7010	1.8371
15	1.4969	1.7010	1.8371	1.4969	1.7010	1.8371	1.4969	1.7010	1.8371
16	1.9958	2.2680	2.4494	1.9958	2.2680	2.4494	1.9958	2.2680	2.4494
17	1.9958	2.2680	2.4494	1.9958	2.2680	2.4494	1.9958	2.2680	2.4494
18	1.9958	2.2680	2.4494	1.9958	2.2680	2.4494	1.9958	2.2680	2.4494
19	0.6178	0.7020	0.7582	0.6178	0.7020	0.7582	0.6178	0.7020	0.7582
20	0.6178	0.7020	0.7582	0.6178	0.7020	0.7582	0.6178	0.7020	0.7582
21	0.6178	0.7020	0.7582	0.6178	0.7020	0.7582	0.6178	0.7020	0.7582
22	0.9266	1.0530	1.1372	0.9266	1.0530	1.1372	0.9266	1.0530	1.1372
23	0.9266	1.0530	1.1372	0.9266	1.0530	1.1372	0.9266	1.0530	1.1372
24	0.9266	1.0530	1.1372	0.9266	1.0530	1.1372	0.9266	1.0530	1.1372
25	1.2355	1.4040	1.5163	1.2355	1.4040	1.5163	1.2355	1.4040	1.5163
26	1.2355	1.4040	1.5163	1.2355	1.4040	1.5163	1.2355	1.4040	1.5163
27	1.2355	1.4040	1.5163	1.2355	1.4040	1.5163	1.2355	1.4040	1.5163
28	1.5444	1.7550	1.8954	1.5444	1.7550	1.8954	1.5444	1.7550	1.8954
29	1.5444	1.7550	1.8954	1.5444	1.7550	1.8954	1.5444	1.7550	1.8954
30	1.5444	1.7550	1.8954	1.5444	1.7550	1.8954	1.5444	1.7550	1.8954
31	2.3166	2.6325	2.8431	2.3166	2.6325	2.8431	2.3166	2.6325	2.8431

32	2.3166	2.6325	2.8431	2.3166	2.6325	2.8431	2.3166	2.6325	2.8431
33	2.3166	2.6325	2.8431	2.3166	2.6325	2.8431	2.3166	2.6325	2.8431
34	3.0888	3.5100	3.7908	3.0888	3.5100	3.7908	3.0888	3.5100	3.7908
35	3.0888	3.5100	3.7908	3.0888	3.5100	3.7908	3.0888	3.5100	3.7908
36	3.0888	3.5100	3.7908	3.0888	3.5100	3.7908	3.0888	3.5100	3.7908
37	1.0058	1.1430	1.2344	1.0058	1.1430	1.2344	1.0058	1.1430	1.2344
38	1.0058	1.1430	1.2344	1.0058	1.1430	1.2344	1.0058	1.1430	1.2344
39	1.0058	1.1430	1.2344	1.0058	1.1430	1.2344	1.0058	1.1430	1.2344
40	1.5088	1.7145	1.8517	1.5088	1.7145	1.8517	1.5088	1.7145	1.8517
41	1.5088	1.7145	1.8517	1.5088	1.7145	1.8517	1.5088	1.7145	1.8517
42	1.5088	1.7145	1.8517	1.5088	1.7145	1.8517	1.5088	1.7145	1.8517
43	2.0117	2.2860	2.4689	2.0117	2.2860	2.4689	2.0117	2.2860	2.4689
44	2.0117	2.2860	2.4689	2.0117	2.2860	2.4689	2.0117	2.2860	2.4689
45	2.0117	2.2860	2.4689	2.0117	2.2860	2.4689	2.0117	2.2860	2.4689
46	0.6257	0.7110	0.7679	0.6257	0.7110	0.7679	0.6257	0.7110	0.7679
47	0.6257	0.7110	0.7679	0.6257	0.7110	0.7679	0.6257	0.7110	0.7679
48	0.6257	0.7110	0.7679	0.6257	0.7110	0.7679	0.6257	0.7110	0.7679
49	0.9385	1.0665	1.1518	0.9385	1.0665	1.1518	0.9385	1.0665	1.1518
50	0.9385	1.0665	1.1518	0.9385	1.0665	1.1518	0.9385	1.0665	1.1518
51	0.9385	1.0665	1.1518	0.9385	1.0665	1.1518	0.9385	1.0665	1.1518
52	1.2514	1.4220	1.5358	1.2514	1.4220	1.5358	1.2514	1.4220	1.5358
53	1.2514	1.4220	1.5358	1.2514	1.4220	1.5358	1.2514	1.4220	1.5358
54	1.2514	1.4220	1.5358	1.2514	1.4220	1.5358	1.2514	1.4220	1.5358
55	1.5048	1.7100	1.8468	1.5048	1.7100	1.8468	1.5048	1.7100	1.8468
56	1.5048	1.7100	1.8468	1.5048	1.7100	1.8468	1.5048	1.7100	1.8468
57	1.5048	1.7100	1.8468	1.5048	1.7100	1.8468	1.5048	1.7100	1.8468
58	2.2572	2.5650	2.7702	2.2572	2.5650	2.7702	2.2572	2.5650	2.7702
59	2.2572	2.5650	2.7702	2.2572	2.5650	2.7702	2.2572	2.5650	2.7702
60	2.2572	2.5650	2.7702	2.2572	2.5650	2.7702	2.2572	2.5650	2.7702
61	3.0096	3.4200	3.6936	3.0096	3.4200	3.6936	3.0096	3.4200	3.6936
62	3.0096	3.4200	3.6936	3.0096	3.4200	3.6936	3.0096	3.4200	3.6936
63	3.0096	3.4200	3.6936	3.0096	3.4200	3.6936	3.0096	3.4200	3.6936
64	0.9900	1.1250	1.2150	0.9900	1.1250	1.2150	0.9900	1.1250	1.2150
65	0.9900	1.1250	1.2150	0.9900	1.1250	1.2150	0.9900	1.1250	1.2150
66	0.9900	1.1250	1.2150	0.9900	1.1250	1.2150	0.9900	1.1250	1.2150
67	1.4850	1.6875	1.8225	1.4850	1.6875	1.8225	1.4850	1.6875	1.8225
68	1.4850	1.6875	1.8225	1.4850	1.6875	1.8225	1.4850	1.6875	1.8225
69	1.4850	1.6875	1.8225	1.4850	1.6875	1.8225	1.4850	1.6875	1.8225
70	1.9800	2.2500	2.4300	1.9800	2.2500	2.4300	1.9800	2.2500	2.4300
71	1.9800	2.2500	2.4300	1.9800	2.2500	2.4300	1.9800	2.2500	2.4300
72	1.9800	2.2500	2.4300	1.9800	2.2500	2.4300	1.9800	2.2500	2.4300
73	0.6098	0.6930	0.7484	0.6098	0.6930	0.7484	0.6098	0.6930	0.7484
74	0.6098	0.6930	0.7484	0.6098	0.6930	0.7484	0.6098	0.6930	0.7484
75	0.6098	0.6930	0.7484	0.6098	0.6930	0.7484	0.6098	0.6930	0.7484
76	0.9148	1.0395	1.1227	0.9148	1.0395	1.1227	0.9148	1.0395	1.1227
77	0.9148	1.0395	1.1227	0.9148	1.0395	1.1227	0.9148	1.0395	1.1227
78	0.9148	1.0395	1.1227	0.9148	1.0395	1.1227	0.9148	1.0395	1.1227
79	1.2197	1.3860	1.4969	1.2197	1.3860	1.4969	1.2197	1.3860	1.4969
80	1.2197	1.3860	1.4969	1.2197	1.3860	1.4969	1.2197	1.3860	1.4969
81	1.2197	1.3860	1.4969	1.2197	1.3860	1.4969	1.2197	1.3860	1.4969

Table 6. Full factorial design L_{81} array for LCC.

Runs	LCC (10 ³ USD)								
	y_{31}	y_{32}	y_{33}	y_{34}	y_{35}	y_{36}	y_{37}	y_{38}	y_{39}
1	484.8109	539.5901	576.1095	485.0439	539.8231	576.3425	485.2769	540.0561	576.5755
2	483.5688	538.3475	574.8666	483.7086	538.4873	575.0064	483.8484	538.6271	575.1462
3	483.0364	537.8150	574.3340	483.1363	537.9148	574.4338	483.2361	538.0147	574.5337
4	689.5455	771.7150	826.4947	690.0697	772.2392	827.0190	690.5939	772.7635	827.5432
5	686.7506	768.9192	823.6982	687.0652	769.2337	824.0127	687.3797	769.5482	824.3273
6	685.5528	767.7210	822.4997	685.7775	767.9456	822.7244	686.0022	768.1703	822.9490
7	895.8327	1005.3932	1078.4336	896.7646	1006.3252	1079.3655	897.6966	1007.2571	1080.2975
8	890.8641	1000.4228	1073.4620	891.4232	1000.9820	1074.0211	891.9824	1001.5411	1074.5803
9	888.7346	998.2926	1071.3313	889.1341	998.6920	1071.7307	889.5335	999.0914	1072.1301
10	387.3201	423.0837	446.9261	387.7002	423.4638	447.3063	388.0803	423.8440	447.6864
11	385.2934	421.0563	444.8982	385.5215	421.2844	445.1263	385.7495	421.5125	445.3544
12	384.4248	420.1874	444.0292	384.5877	420.3503	444.1921	384.7506	420.5132	444.3550
13	524.7801	578.4269	614.1915	525.6354	579.2822	615.0468	526.4907	580.1375	615.9021
14	520.2201	573.8653	609.6287	520.7333	574.3785	610.1419	521.2464	574.8916	610.6551
15	518.2658	571.9103	607.6733	518.6323	572.2768	608.0398	518.9989	572.6434	608.4064
16	664.7735	736.3044	783.9917	666.2940	737.8250	785.5122	667.8146	739.3455	787.0328
17	656.6668	728.1948	775.8802	657.5791	729.1071	776.7925	658.4914	730.0195	777.7048
18	653.1925	724.7193	772.4038	653.8441	725.3709	773.0554	654.4958	726.0226	773.7071
19	350.3559	372.4971	387.2579	350.9568	373.0980	387.8587	351.5577	373.6988	388.4596
20	347.1525	369.2925	384.0525	347.5130	369.6530	384.4130	347.8735	370.0135	384.7735
21	345.7796	367.9191	382.6788	346.0371	368.1766	382.9363	346.2946	368.4341	383.1938
22	441.5404	474.7543	496.8969	442.8923	476.1062	498.2488	444.2443	477.4582	499.6008
23	434.3326	467.5439	489.6848	435.1438	468.3551	490.4960	435.9550	469.1663	491.3071
24	431.2436	464.4538	486.5939	431.8230	465.0332	487.1733	432.4024	465.6126	487.7527
25	536.7291	581.0172	610.5426	539.1326	583.4207	612.9460	541.5360	585.8241	615.3495
26	523.9153	568.1988	597.7211	525.3574	569.6409	599.1632	526.7995	571.0829	600.6053
27	518.4237	562.7052	592.2262	519.4538	563.7353	593.2562	520.4838	564.7653	594.2863
28	493.7285	549.0753	585.9731	493.9431	549.2898	586.1877	494.1577	549.5044	586.4022
29	492.5845	547.9308	584.8283	492.7132	548.0595	584.9570	492.8420	548.1883	585.0858
30	492.0941	547.4403	584.3377	492.1861	547.5322	584.4296	492.2781	547.6242	584.5216
31	700.2380	783.2588	838.6060	700.7208	783.7416	839.0889	701.2036	784.2245	839.5717
32	697.6638	780.6837	836.0303	697.9535	780.9734	836.3200	698.2432	781.2631	836.6097
33	696.5605	779.5801	834.9264	696.7675	779.7870	835.1333	696.9744	779.9939	835.3403
34	908.1775	1018.8730	1092.6700	909.0359	1019.7314	1093.5284	909.8943	1020.5898	1094.3867
35	903.6012	1014.2950	1088.0909	904.1162	1014.8100	1088.6059	904.6312	1015.3251	1089.1209
36	901.6399	1012.3330	1086.1284	902.0077	1012.7009	1086.4963	902.3756	1013.0688	1086.8642
37	394.0745	430.1219	454.1535	394.4302	430.4775	454.5091	394.7858	430.8331	454.8647
38	392.1786	428.2253	452.2564	392.3920	428.4387	452.4698	392.6054	428.6520	452.6831
39	391.3661	427.4125	451.4434	391.5185	427.5649	451.5958	391.6709	427.7173	451.7482
40	532.1667	586.2390	622.2872	532.9668	587.0391	623.0873	533.7669	587.8392	623.8874
41	527.9008	581.9716	618.0188	528.3809	582.4517	618.4989	528.8610	582.9318	618.9790
42	526.0726	580.1428	616.1895	526.4156	580.4857	616.5324	526.7585	580.8286	616.8753
43	672.6287	744.7268	792.7922	674.0511	746.1493	794.2147	675.4736	747.5717	795.6371
44	665.0450	737.1404	785.2040	665.8985	737.9939	786.0575	666.7519	738.8473	786.9109
45	661.7949	733.8891	781.9519	662.4045	734.4987	782.5615	663.0141	735.1083	783.1711
46	356.9470	379.3718	394.3217	357.5111	379.9359	394.8857	358.0751	380.5000	395.4498
47	353.9397	376.3634	391.3125	354.2781	376.7018	391.6510	354.6166	377.0403	391.9894
48	352.6508	375.0741	390.0229	352.8926	375.3158	390.2647	353.1343	375.5576	390.5064
49	448.5592	482.1985	504.6246	449.8284	483.4676	505.8938	451.0976	484.7368	507.1629

50	441.7928	475.4296	497.8541	442.5543	476.1911	498.6156	443.3158	476.9526	499.3771
51	438.8928	472.5286	494.9524	439.4368	473.0725	495.4964	439.9807	473.6165	496.0403
52	543.9306	588.7856	618.6889	546.1869	591.0419	620.9452	548.4432	593.2982	623.2015
53	531.9013	576.7520	606.6525	533.2551	578.1058	608.0062	534.6089	579.4596	609.3600
54	526.7459	571.5947	601.4940	527.7129	572.5617	602.4610	528.6799	573.5287	603.4279
55	508.0765	562.0040	597.9557	508.2727	562.2002	598.1519	508.4689	562.3964	598.3481
56	507.0305	560.9576	596.9090	507.1482	561.0753	597.0267	507.2659	561.1930	597.1445
57	506.5822	560.5091	596.4605	506.6663	560.5932	596.5445	506.7504	560.6773	596.6286
58	709.0760	789.9680	843.8960	709.5175	790.4095	844.3374	709.9589	790.8509	844.7789
59	706.7225	787.6136	841.5410	706.9874	787.8785	841.8059	707.2522	788.1433	842.0707
60	705.7138	786.6046	840.5317	705.9030	786.7938	840.7209	706.0922	786.9830	840.9101
61	911.3831	1019.2400	1091.1446	912.1679	1020.0248	1091.9294	912.9527	1020.8096	1092.7142
62	907.1990	1015.0544	1086.9580	907.6699	1015.5253	1087.4289	908.1408	1015.9962	1087.8997
63	905.4058	1013.2606	1085.1637	905.7422	1013.5969	1085.5001	906.0785	1013.9333	1085.8364
64	414.2580	449.7375	473.3905	414.5646	450.0441	473.6970	414.8712	450.3506	474.0036
65	412.6236	448.1025	471.7551	412.8076	448.2864	471.9390	412.9915	448.4704	472.1230
66	411.9232	447.4018	471.0542	412.0545	447.5332	471.1856	412.1859	447.6646	471.3170
67	549.4516	602.6719	638.1521	550.1413	603.3616	638.8418	550.8311	604.0514	639.5316
68	545.7741	598.9931	634.4724	546.1880	599.4070	634.8863	546.6019	599.8208	635.3002
69	544.1981	597.4165	632.8955	544.4937	597.7121	633.1911	544.7893	598.0077	633.4867
70	686.6881	757.6500	804.9579	687.9144	758.8763	806.1842	689.1406	760.1025	807.4104
71	680.1505	751.1100	798.4164	680.8862	751.8458	799.1521	681.6220	752.5815	799.8879
72	677.3486	748.3071	795.6128	677.8742	748.8327	796.1384	678.3997	749.3582	796.6639
73	396.9670	418.8239	433.3951	397.4698	419.3267	433.8979	397.9726	419.8294	434.4007
74	394.2866	416.1425	430.7131	394.5883	416.4442	431.0148	394.8899	416.7458	431.3164
75	393.1378	414.9933	429.5637	393.3533	415.2088	429.7791	393.5688	415.4243	429.9946
76	485.4764	518.2635	540.1215	486.6076	519.3947	541.2528	487.7388	520.5259	542.3840
77	479.4454	512.2303	534.0869	480.1241	512.9091	534.7657	480.8029	513.5878	535.4444
78	476.8607	509.6447	531.5007	477.3455	510.1295	531.9855	477.8303	510.6143	532.4703
79	577.3363	621.0548	650.2005	579.3473	623.0659	652.2115	581.3584	625.0769	654.2226
80	566.6145	610.3292	639.4723	567.8211	611.5358	640.6790	569.0278	612.7425	641.8856
81	562.0195	605.7325	634.8745	562.8814	606.5944	635.7364	563.7432	607.4563	636.5983

To enable an optimization-based solution search, the three output responses obtained from the L_{81} experimental array are transformed into individual desirability functions. Each desirability function maps a response value onto a normalized scale ranging from 0 to 1, where smaller response values correspond to higher desirability. In this study, all three output responses are to be minimized; therefore, equation (15) is employed to construct the desirability function for each response.

To identify optimal solutions, the exponent of each individual desirability function is varied from 0.1 to 2.0 with an increment of 0.1. The overall desirability is then defined as the geometric mean of the individual desirabilities and is used as a scalar objective

function for optimization. All computations and optimization procedures are implemented in MATLAB. By varying the exponents of the individual desirability functions, different overall desirability surfaces are obtained, as illustrated in Figure 2. A higher value of the overall desirability function indicates a more favorable optimal solution.

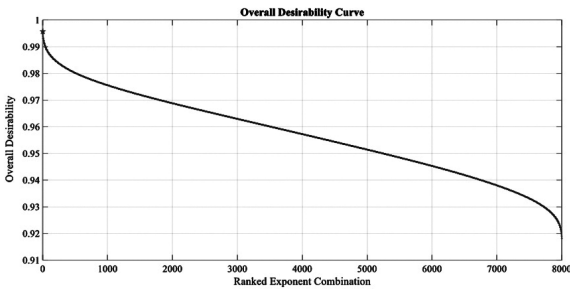


Figure 2. Overall desirability functions.

The optimal solution is obtained when the exponent of each individual desirability function is set to 0.1, resulting in an overall desirability value of 0.9957. Therefore, the optimal design parameters for the 22-kV overhead power transmission line in this case are as follows:

- ✓ Conductor type: ACSR;
- ✓ Size: 240 mm²;
- ✓ Span: 100 m;
- ✓ Pretension: 70%;
- ✓ Maximum sag: 0.3691 ± 0.0223 m;
- ✓ Power loss: 0.6926 ± 0.0612 kW;
- ✓ LCC: 36571.6667 ± 1608.5526 USD.

The main effects plots showing the influence of the four control factors on the three output responses are presented in Figures 3, 4, 5, 6, 7, and 8.

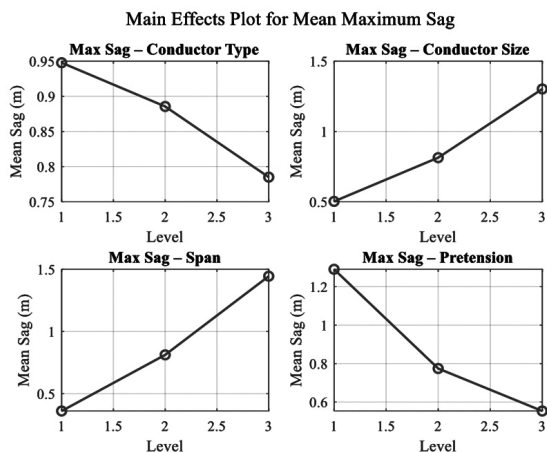


Figure 3. Main effects plot for mean maximum sag.

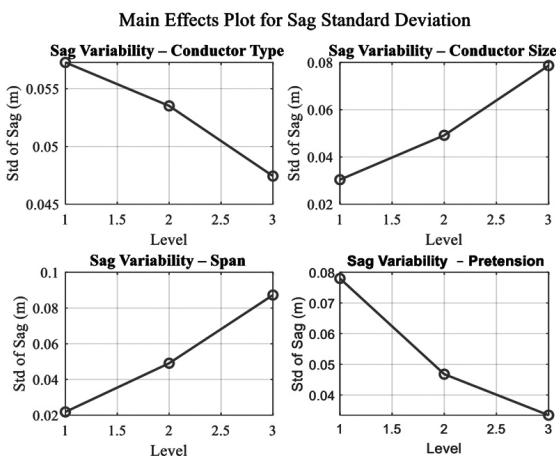


Figure 4. Main effects plot for standard deviation maximum sag.

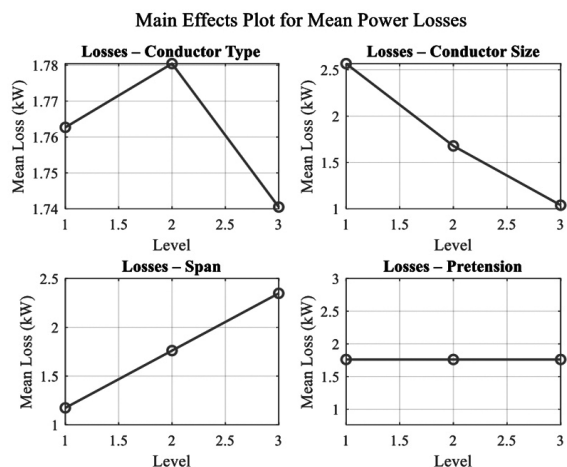


Figure 5. Main effects plot for mean power losses.

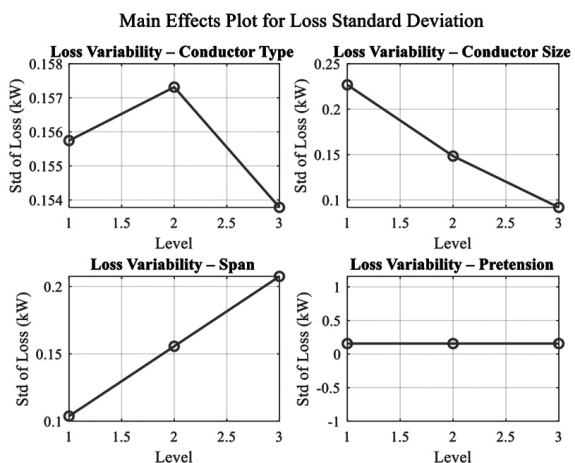


Figure 6. Main effects plot for standard deviation power losses.

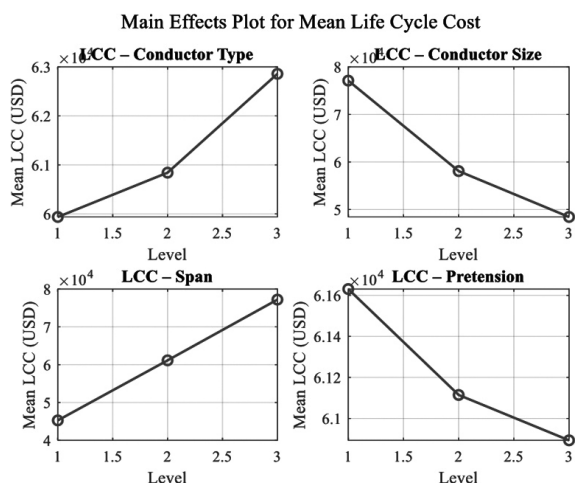


Figure 7. Main effects plot for mean LCC.

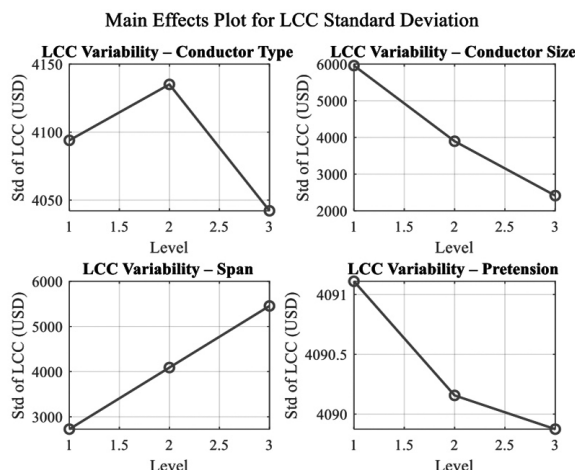


Figure 8. Main effects plot for standard deviation LCC.

Figures 3–8 collectively present the main effects of the control factors on the robust performance of the 22-kV overhead transmission line design. Figures 3, 5, and 7 show the main effects of conductor type, conductor size, span length, and pretension on the mean values of maximum sag, power losses, and LCC, respectively, where each point represents the average response at a given factor level under all noise conditions. In all three responses, span length exhibits the steepest trends, indicating that it is the dominant parameter influencing both mechanical and economic performance, which is physically consistent with the quadratic dependence of sag and resistive losses on span. Pretension also demonstrates a significant influence, particularly in reducing sag and stabilizing LCC, confirming its role in enhancing mechanical robustness. By contrast, conductor type and conductor size show more moderate but non-negligible effects, reflecting their indirect influence through mass, resistance, and material cost.

Figures 4, 6, and 8 show the main effects of conductor type, conductor size, span length, and pretension on the standard deviation values of maximum sag, power losses, and LCC, respectively, where each point represents the standard deviation response at a given factor level under all noise conditions. In all three responses, span length also exhibits the steepest trends,

indicating that it is the dominant parameter influencing both mechanical and economic performance, which is physically consistent with the quadratic dependence of sag and resistive losses on span. Pretension also demonstrates a significant influence, particularly in reducing sag and stabilizing LCC, confirming its role in enhancing mechanical robustness. By contrast, conductor type and conductor size show more moderate but non-negligible effects, reflecting their indirect influence through mass, resistance, and material cost.

4. CONCLUSIONS

This paper proposes a robust design-based two-stage optimization framework for overhead power transmission line design by integrating a full factorial L_{81} crossed-array experimental design with desirability function-based multi-objective optimization. The framework explicitly incorporates both control factors and environmental noise factors, enabling robust design solutions that account for uncertainties in ambient temperature and wind speed. A case study on a 22-kV overhead transmission line demonstrated that the proposed approach can effectively minimize maximum sag, power loss, and LCC simultaneously. The results showed that high overall desirability values can be achieved through appropriate selection of desirability function parameters, leading to balanced and robust design solutions. The MATLAB-based implementation confirms the practical applicability of the proposed methodology for transmission line design optimization.

Future work may extend the proposed framework by considering additional design and environmental factors, such as tower configuration, ice loading, and solar radiation. Moreover, efficient experimental designs and advanced metaheuristic optimization techniques can be explored to reduce computational effort and further improve solution quality. The integration of real-time monitoring data is another promising direction for adaptive and data-driven transmission line design and operation.

DECLARATION OF CONFLICTS OF INTEREST

The authors declare that this research has no conflicts of interest.

REFERENCES

- [1] M. Kampik, P. Kubek, B. Krupanek, and R. Bogacz, "Sag and Tension Calculations for High-Voltage Overhead Line Conductors," *Energies*, vol. 17, no. 12, p. 2967, 2024.
- [2] M. Keshavarzian and C. H. Priebe, "Sag and tension calculations for overhead transmission lines at high temperatures-modified ruling span method," *IEEE Transactions on Power Delivery*, vol. 15, no. 2, pp. 777–783.
- [3] A. Nudler, R. Yozevitch, and E. Holdengreber, "Enhanced method for assessing designed sag accuracy in overhead transmission electric lines using field data and statistical analysis," *International Journal of Electrical Power & Energy Systems*, vol. 165, p. 110457, 2025.
- [4] P. Kumar and A. K. Singh, "Optimal mechanical sag estimator for leveled span overhead transmission line conductor," *Measurement*, vol. 137, pp. 691-699, 2019.
- [5] Z. A. Muhammad, A. A. Muhammad, K. Akhtar, and M. E. Munir, "Overhead Transmission Lines Analysis Considering Sag-Tension under Maximum Wind Effect," *Journal of Mechanics of Continua and Mathematical Sciences*, vol. 13, no. 5, pp. 185-192, 2018.
- [6] I. Makhkamova, K. Mahkamov, and P. Taylor, "CFD thermal modelling of Lynx overhead conductors in distribution networks with integrated Renewable Energy Driven Generators," *Applied Thermal Engineering*, vol. 58, no. 1, pp. 522–535, 2013.
- [7] F. Margita, L. Beňa, W. Malska, and P. Pijarski, "Possibilities of Increasing the Ampacity of Overhead Lines Using High-Temperature Low-Sag Conductors in the Electric Power System of the Slovak Republic," *Applied Sciences*, vol. 14, no. 17, p. 7846, 2024.
- [8] T. O. Seppa, "Accurate ampacity determination: Temperature-Sag Model for operational real time ratings," *IEEE Transactions on Power Delivery*, vol. 10, no. 3, pp. 1460-1470, 1995.
- [9] Y. Luo, C. Gao, D. Wang, Z. Jiang, Y. Lv, and G. Xue, "Predictive model for sag and load on overhead transmission lines based on local deformation of transmission lines," *Electric Power Systems Research*, vol. 214, p. 108811, 2023.
- [10] S. H. Kim and K. H. Chun, "Overhead Power Line Tension Estimation Method Using Accelerometers," *Energies*, vol. 18, no. 1, p. 181, 2025.
- [11] J. Zuo, J. Fan, Y. Ouyang, H. Liu, C. Yang, and C. Hao, "Transmission Line Sag Measurement and Simulation Research Based on Non-Contact Electric Field Sensing," *Sensors*, vol. 22, no. 21, p. 8379, 2022.
- [12] M. Ahsan, M. N. R. Baharom, I. U. Khalil, and Z. Zainal, "Analysis of high-ampacity and low-sag conductors of 275 kV overhead transmission lines using reconductoring technique," *Electric Power Systems Research*, vol. 246, p. 111719, 2025.
- [13] S. Jayathilake, P. Rajeev, and E. Gad, "Review of the design and condition monitoring of overhead power distribution conductors," *Journal of Civil Engineering and Management*, vol. 31, no. 4, pp. 338-361, 2025.
- [14] S. K. Teegala and S. K. Singal, "Optimal costing of overhead power transmission lines using genetic algorithms," *International Journal of Electrical Power & Energy Systems*, vol. 83, pp. 298-308, 2016.
- [15] A. H. M. Santos, R. M. de Lima, C. R. S. Pereira, O. Reinis, O. S. M. Giulia, R. Q. Anderson, K. F. Bárbara, R. P. C. C. Arthur, C. J. Luiz, A. S. Renato, and L. C. J. Eden, "Optimizing routing and tower spotting of electricity transmission lines: An integration of geographical data and engineering aspects into decision-making," *Electric Power Systems Research*, vol. 176, p. 105953, 2019.

- [16] J. S. Acosta and M. C. Tavares, "Methodology for optimizing the capacity and costs of overhead transmission lines by modifying their bundle geometry," *Electric Power Systems Research*, vol. 163, pp. 668-677, 2018.
- [17] J. R. J. Octavio, O. L. Garcia, E. Pilo, and A. Carnicero, "Coupled Electromechanical Optimization of Power Transmission Lines," *CMES*, vol. 25, no. 2, pp. 81-98, 2008.
- [18] M. S. Shaikh, C. Hua, M. Hassan, S. Raj, M. A. Jatoli, and M. M. Ansari, "Optimal parameter estimation of overhead transmission line considering different bundle conductors with the uncertainty of load modeling," *Optimal Control Applications and Methods*, vol. 43, no. 3, pp. 652-666, 2022.
- [19] A. C. Shoemaker, K.-L. Tsui, and C. F. J. Wu, "Economical Experimentation Methods for Robust Design," *Technometrics*, vol. 33, no. 4, pp. 415-427, 1991.
- [20] R. H. Myers, "Response Surface Methodology-Current Status and Future Directions," *Journal of Quality Technology*, vol. 31, no. 1, pp. 30-44, 1999.
- [21] G. G. Vining and R. H. Myers, "Combining Taguchi and Response Surface Philosophies: A Dual Response Approach," *Journal of Quality Technology*, vol. 22, no. 1, pp. 38-45, 1990.
- [22] W. Zeng, J. Fan, W. Zhang, L. Yu, Z. Bin, H. Ruirui, X. Xiao, and L. Junyong, "Whole Life Cycle Cost Analysis of Transmission Lines Using the Economic Life Interval Method," *Energies*, vol. 16, no. 23, p. 7804, 2023.



© 2026 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Thiết kế tối ưu bền vững cho thiết kế đường dây truyền tải điện trên không

Phạm Văn Phát¹, Trần Ngọc Thanh¹, Lê Tuấn Hộ^{2,*}

¹Khoa Điện, Trường Cao đẳng Kỹ thuật và Công nghệ Quy Nhơn, Việt Nam

²Khoa Kỹ thuật và Công nghệ, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 23/12/2025; Ngày sửa bài: 02/02/2026;

Ngày nhận đăng: 14/02/2026; Ngày xuất bản: 28/06/2026

TÓM TẮT

Hầu hết các phương pháp thiết kế đường dây truyền tải điện trên không trong thực tế hiện nay chủ yếu dựa trên các công thức thực nghiệm. Do đó, bài báo này nhằm đề xuất một quy trình thiết kế tối ưu gồm hai giai đoạn dựa trên thiết kế bền vững để thiết kế đường dây truyền tải điện trên không. Trong giai đoạn thứ nhất, tương ứng với giai đoạn thiết kế thí nghiệm, bốn biến đầu vào gồm loại dây dẫn, kích thước dây dẫn, chiều dài nhịp và lực căng ban đầu; cùng với hai yếu tố nhiễu là nhiệt độ môi trường và tốc độ gió được xem xét, mỗi yếu tố ở ba mức. Ba đáp ứng đầu ra được đánh giá bao gồm: độ võng lớn nhất, tổn thất công suất và chi phí vòng đời. Các yếu tố biến đầu vào và yếu tố nhiễu lần lượt được thiết kế trong ma trận trong và ma trận ngoài theo thiết kế giai thừa toàn phần. Sau đó, các ma trận này được kết hợp để hình thành thiết kế thí nghiệm giai thừa toàn phần dạng mảng chéo L_{81} . Trong giai đoạn thứ hai, các hàm mong muốn được sử dụng để tối thiểu hóa đồng thời ba đáp ứng đầu ra. Cuối cùng, một phân tích so sánh giữa các dạng hàm mong muốn khác nhau được thực hiện, và dựa trên các kết quả thu được, các nghiệm thiết kế tối ưu cho đường dây truyền tải điện trên không được xác định.

Từ khóa: Thiết kế bền vững, thiết kế đường dây truyền tải điện trên không, thiết kế giai thừa toàn phần, hàm mong muốn.

*Tác giả liên hệ chính.

Email: tuanhole@qnu.edu.vn

Research on beamforming optimization in intelligent reflective surface-assisted wireless communication systems

Do Dung Nguyen*, Van Hao Nguyen, Minh Hung Dao

Faculty of Engineering and Technology, Quy Nhon University, Vietnam

Received: 06/10/2025; Revised: 17/12/2025;

Accepted: 23/12/2025; Published: 28/06/2026

ABSTRACT

This paper investigates the optimization of beamforming in intelligent reflecting surface (IRS)-assisted heterogeneous wireless communication systems, aiming to meet the stringent requirements of ultra-reliable low-latency communication (URLLC). By designing the transmit beamforming vectors at the small cell (SC), the proposed algorithm transforms a non-convex optimization problem into a solvable convex form using successive convex approximation (SCA) and semi-definite relaxation (SDR) techniques. Simulation results demonstrate that the proposed method significantly enhances system throughput and user data rates compared to benchmark algorithms, especially under varying IRS configurations and transmission power levels. The study confirms the effectiveness and scalability of IRS in improving network performance in URLLC scenarios.

Keywords: *Intelligent Reflecting Surface (IRS), heterogeneous network, beamforming design, UltraReliable Low-Latency Communication (URLLC).*

1. INTRODUCTION

Future wireless networks are anticipated to evolve toward intelligent and software-reconfigurable architectures, enabling ubiquitous communication among humans and mobile devices. These networks will possess the capability to sense, control, and optimize the wireless environment, thereby supporting low-power operation, high throughput, massive connectivity, and ultra-low-latency communication [1]. As communication scenarios in the 6th Generation (6G) mobile communication era become increasingly complex, large-scale MIMO and small-cell deployments are receiving growing attention as key network architectures [2], [3]. This is particularly relevant for practical communication scenarios in densely populated

areas with high user density. Ultra-Reliable and LowLatency Communications (URLLC) represents a critical standard for 6G networks [4], employing short-packet transmission to satisfy stringent reliability and latency requirements. URLLC is especially targeted at mission-critical applications, including industrial automation, remote healthcare, and intelligent transportation, and other scenarios that demand ultra-low latency and ultra-high reliability [5].

Intelligent Reconfigurable Surface (IRS) is a technology capable of reconfiguring the wireless propagation environment, acting like a passive metal mirror or “wave collector.” It can be programmed to modify the impinging electromagnetic field in a customizable manner. IRS is considered a highly promising innovation

*Corresponding author:

Email: nguyendodung@qnu.edu.vn

due to its immense potential to achieve low power consumption, high energy efficiency, high-speed communication, massive connectivity, and low-latency wireless communication [1]. Simultaneously, accurate and low-overhead channel estimation is a critical challenge in IRS-based systems due to the large number of IRS elements and their unique hardware constraints. Reference focuses on an IRS-enabled multi-user Multi-Input Single-Output (MISO) uplink communication system and proposes a channel estimation framework based on parallel factor decomposition to obtain the cascaded channel model. It presents two iterative estimation algorithms for the channel between the base station and the IRS, and the channel between the IRS and the users. It demonstrates that the total rate achieved using the estimated channel always reaches the total rate of the perfect channel under various settings, thereby verifying the effectiveness and robustness of the proposed estimation algorithms [6]. Reference [7] investigates secure beamforming in a multi-IRS assisted millimeterwave (mmWave) system.

These methods have demonstrated superiority over traditional approaches. Currently, much research combines IRS with URLLC or heterogeneous network scenarios. For instance, the paper [8] provides a comprehensive survey of resource allocation (RA) strategies in 5G heterogeneous networks (HetNets), highlighting current research, challenges, and future directions. HetNets integrate various small cells (micro, pico, femto) with macrocells to enhance spectrum efficiency (SE), coverage, and quality of service (QoS). The authors also identify several challenging open issues and propose future research directions in this field. Furthermore, two effective approaches for addressing RA problems in sixth-generation (6G) communications are discussed. In reference [9], the article performed maximizing the ratio of the total data rate to the total cross-layer interference for femtocell users in a two-

tier OFDMA heterogeneous network with layer interference under imperfect channel state information. The optimization problem is studied by optimizing the femtocell base station's transmit power and subcarrier allocation factors. The original problem is transformed into a convex optimization problem using quadratic transformation, variable relaxation, and Lagrange duality theory, and the results indicate an improved interference efficiency compared to traditional schemes.

However, references [6], [7], [9] either focus solely on URLLC without considering heterogeneous networks, or they only address the Shannon capacity scenario. In situations where network conditions become complex, it is important to consider cases where Shannon capacity alone may not meet user demands. To address this, and to satisfy the quality of service requirements for short-packet communication users in heterogeneous networks, this paper proposes an alternating iterative optimization algorithm that is more aligned with practical application needs. The main contributions are as follows:

(1) We propose a novel beamforming optimization algorithm for IRS-assisted heterogeneous wireless networks that designs the transmit beamforming vectors at the small cell with different assumptions on IRS reflection angles of the IRS to meet URLLC requirements.

(2) The original non-convex optimization problem is effectively transformed into a solvable convex form using successive convex approximation (SCA) and semi-definite relaxation (SDR) techniques.

(3) We conduct extensive simulations to evaluate the performance of the proposed algorithm under various system configurations, demonstrating its superiority over benchmark methods such as random phase shift and no-IRS schemes.

2. SYSTEM MODEL

2.1. Method of signal transmission to users

As shown in Figure 1, the system considered in this paper is a downlink heterogeneous network, where the overall communication system consists of a base station (BS), small cell (SC), users in BS, and users in SC assisted by intelligent reflecting surface (IRS) to meet the communication requirements of high reliability and low latency. Both the base station and small cell are equipped with multiple antennas ($M > 1$). The total number of users in the communication area is $\mathcal{U}_B$, and the number of small cell users assisted by IRS is $\mathcal{U}_S = \{1, 2, \dots, K\}$, with each user being a single-antenna user. The IRS is composed of $\mathcal{N} = \{1, 2, \dots, N\}$ reflecting units, and the corresponding channel matrix is $\Phi = \text{diag}(\mathbf{u}) \in \mathbb{C}^{N \times N}$, where $\mathbf{u} = [u_1, u_2, \dots, u_N]^H = [e^{j\theta_1}, e^{j\theta_2}, \dots, e^{j\theta_N}]^H \in \mathbb{C}^{N \times 1}$, with $\theta_i \in [0, 2\pi)$, $\forall i \in \mathcal{N}$.

The system operation involves the following coefficient channels: the transmission channel from SC station to IRS, $\mathbf{G}_{S,I} \in \mathbb{C}^{N \times M}$; reflective channel from IRS to users in SC k , $\mathbf{h}_{r,k}^H \in \mathbb{C}^{1 \times N}$ and channel directly from SC to user in small cell, $\mathbf{h}_{d,k}^H \in \mathbb{C}^{1 \times M}$ with $k \in \mathcal{U}_S$; the transmission channel from BS to user in base station k , $\mathbf{h}_{b,k}^H \in \mathbb{C}^{1 \times M}$ with $k \in \mathcal{U}_B$. The IRS is equipped with an intelligent controller to coordinate its reception and reflection of signals during the transmission process, and it can obtain channel state information (CSI) from the SC to the IRS via a wireless connection [10]. Simultaneously, a basic baseband unit (BBU) is implemented between the BS and SC through a wireless backhaul, with both the SC and IRS controlled via wireless connection. It is assumed that the CSI of the entire system can be obtained at the BBU [11].

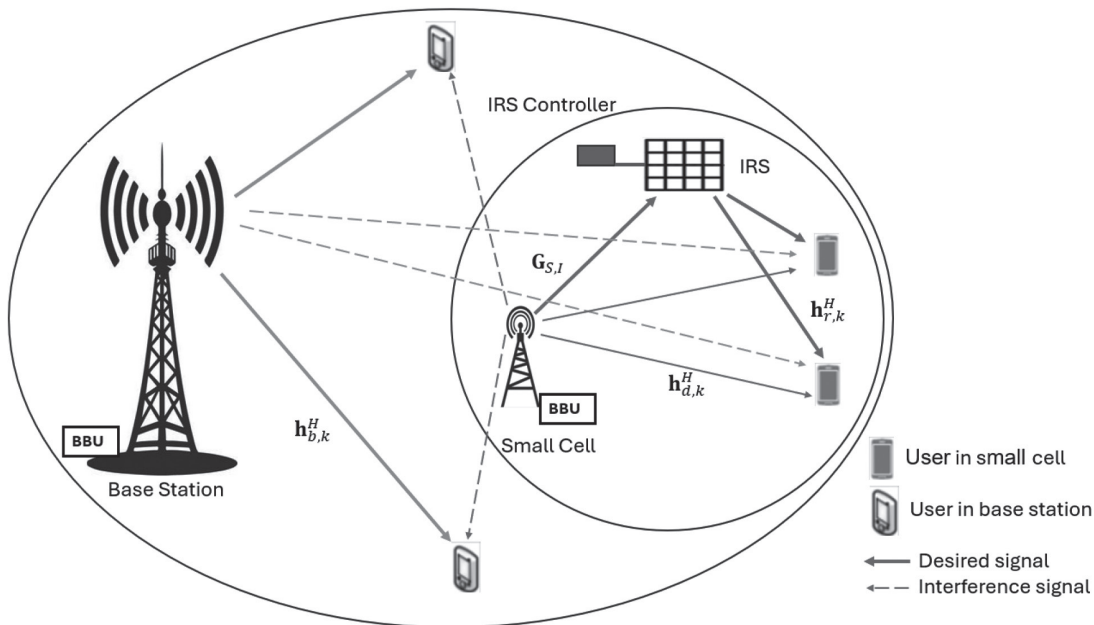


Figure 1. The heterogeneous network model with IRS support.

The signal transmitted from the BS station is given by

$$\mathbf{x}_B = \sum_{i \in \mathcal{U}_B} \mathbf{v}_i s_{b,i}, \quad (1)$$

where $\mathbf{v}_i \in \mathbb{C}^{M \times 1}$ is the beamforming vector for the i -th base station user transmitted from the BS, and $s_{b,i} \sim \mathcal{CN}(0, 1)$, $i \in \mathcal{U}_B$ is the data sent to the i -th base station user with $\mathbb{E}(|s_{b,i}|^2) = 1$. The

transmission signal from the SC is expressed as:

$$\mathbf{x}_S = \sum_{k \in \mathcal{U}_S} \mathbf{w}_k s_{s,k}, \quad (2)$$

where $\mathbf{w}_k \in \mathbb{C}^{M \times 1}$ is the beamforming vector for the k -th small cell user transmitted from the SC, and $s_{s,k} \sim \mathcal{CN}(0, 1)$, $k \in \mathcal{U}_S$ is the data sent to the k -th small cell user with $\mathbb{E}(|s_{s,k}|^2) = 1$. The

maximum transmitted power at the SC, $P_{\max}$, is constrained by the following condition:

$$\sum_{k \in \mathcal{U}_S} \|\mathbf{w}_k\|^2 \leq P_{\max}. \quad (3)$$

2.2. Maximum transmission rate for small cell users

To meet the high-reliability communication requirements and low-latency constraints, the received signal at the small cell users consists of three components: the desired signal from the SC, interference from the BS, and interference from other small cell users. The k -th small cell user is represented as

$$u_{s,k} = (\mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H \Phi \mathbf{G}_{S,I}) \mathbf{w}_k s_{s,k} + \sum_{i \in \mathcal{U}_S, i \neq k} (\mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H \Phi \mathbf{G}_{S,I}) \mathbf{w}_i s_{s,i} + \mathbf{h}_{b,k}^H \mathbf{x}_B + n_k, \quad (4)$$

where $n_k \sim \mathcal{CN}(0, \sigma^2)$ is the additive white Gaussian noise (AWGN) at user k -th. Since this paper focuses on small cell users within the SC coverage area, the BS employs maximum ratio transmission (MRT) for precoding. The MRT beamforming vector [12] is given by $\mathbf{v}^{\text{MRT}} = \frac{\mathbf{h}_{b,k}}{\|\mathbf{h}_{b,k}\|}$. The purpose of dividing the channel vector by its own magnitude normalizes the beamforming vector to have unit norm (i.e., $\|\mathbf{v}^{\text{MRT}}\| = 1$). This ensures the beamformer aligns perfectly with the channel's direction to maximize signal gain at the receiver (the core principle of MRT), while separating the direction of transmission from the magnitude of the transmit power. Thus, the signal to interference plus noise ratio (SINR) for a low-latency user k -th is:

$$\gamma_{s,k} = \frac{|\mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H \Phi \mathbf{G}_{S,I}) \mathbf{w}_k|^2}{\sum_{i \in \mathcal{U}_S, i \neq k} |\mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H \Phi \mathbf{G}_{S,I}) \mathbf{w}_i|^2 + \rho_k + \sigma_k^2}, \quad (5)$$

where $\rho_k = \sum_{i \in \mathcal{U}_B} |\mathbf{h}_{b,k}^H \mathbf{v}_i|^2$. Here $\mathbf{h}_{b,k}$ represents the channel from the BS to the k -th user in BS, and $\mathbf{v}_i$ denotes the beamforming vector associated with the i -th user. In equation (5), the term $\mathbf{w}_k$ represents the precoding vector for the k -th user. In such formulations, the transmit power allocated to user k is typically embedded within this vector. Therefore, the term $\|\mathbf{w}_k\|^2$ (or specifically

$\|\mathbf{w}_k\|^2$) corresponds to the power allocated to that specific user.

The signal received by a base station user l -th also consists of three parts: the desired signal from the BS, the received multi-user interference signals from other users in BS, and the interference signals from the IRS. Thus, the received signal is denoted as:

$$u_{b,l} = \mathbf{h}_{b,l}^H \mathbf{v}_{l s_{m,l}} + \sum_{i \in \mathcal{U}_B, i \neq l} \mathbf{h}_{b,l}^H \mathbf{v}_{i s_{m,i}} + \mathbf{h}_{d,l}^H \mathbf{x}_s + n_l \quad (6)$$

where $n_l \sim \mathcal{CN}(0, \sigma_l^2)$ is the additive white Gaussian noise for the l -th user. The SINR of the base station user l -th is given by

$$\gamma_{b,l} = \frac{|\mathbf{h}_{b,l}^H \mathbf{v}_l|^2}{\sum_{i \in (\mathcal{U}_B - \mathcal{U}_S), i \neq l} |\mathbf{h}_{b,l}^H \mathbf{v}_i|^2 + \sum_{j \in \mathcal{U}_S} |\mathbf{h}_{d,l}^H \mathbf{w}_j|^2 + \sigma_l^2} \quad (7)$$

The transmission rate of a small cell user (in bits per second per Hz) is approximated by equation [13]:

$$R(\gamma_{s,k}) = \log_2(1 + \gamma_{s,k}) - aQ^{-1}(\epsilon) \sqrt{1 - (1 + \gamma_{s,k})^{-2}} \quad (8)$$

This paper aims to maximize the system sum rate for $\mathcal{U}_S$ small cell users within the IRS-assisted region, while ensuring that their QoS [14] requirements are met. For base station users, outside the communication area, it is only necessary to satisfy the minimum SINR requirement. By optimizing the transmit beamforming vectors $\{\mathbf{w}_k\}_{k \in \mathcal{U}_S}$ at the SC and the IRS reflection angles $\theta_i \in [0, 2\pi)$ for all $i \in \mathcal{N}$, the combined optimization problem can be formulated as:

$$\left. \begin{aligned} & \max_{\mathbf{w}_k, \mathbf{u}} \sum_{k \in \mathcal{U}_S} R(\gamma_{s,k}) \\ & \text{s.t.} \\ & \text{C1: } R(\gamma_{s,k}) \geq L_k, \quad \forall k \in \mathcal{U}_S, \\ & \text{C2: } \gamma_{b,l} \geq \text{SINR}_{BS}, \quad \forall l \in (\mathcal{U}_B - \mathcal{U}_S), \\ & \text{C3: } \sum_{k \in \mathcal{U}_S} \|\mathbf{w}_k\|^2 \leq P_{\max}, \\ & \text{C4: } |u_i| = 1, \quad \forall i \in \mathcal{N}. \end{aligned} \right\} \quad (9)$$

Constraint C1 in the optimization problem ensures the minimum QoS requirement for each small cell user; constraint C2 guarantees the minimum SINR requirement for base station users; and constraint C3 ensures that the transmission power at the SC does not exceed the maximum power requirement P_{max} . Constraint C4 is the unit modulus constraint for the IRS elements. Problem (9) is a non-convex problem. The non-convexity arises mainly from the coupling of the SINR and optimization variables in equation (9), and the non-convex nature of the rate expression in equation (8) is reflected in constraint (C1).

For most non-convex problems, there is currently no systematic solution. In the following sections of this paper, a series of approaches are proposed to tackle these problems, and an efficient alternative algorithm based on iterative optimization is presented.

3. DESIGNED ALGORITHM FOR OPTIMIZED BEAMFORMING

3.1. Optimization problem transformation

In this section, we solve optimization problems (9) with coupled variables to obtain optimal solution. This article focuses exclusively on the optimal beamforming at the SC. Therefore, we apply an optimization approach using a fixed-point iteration for the IRS phase shift vector $\mathbf{u}$. Finally, the optimization method is employed to solve the sub-problems in (9).

For the processing $R(x)$ it can be expressed as $R(x) = f_1(x) - aQ^{-1}(\varepsilon)f_2(x)$, where $f_1(x)$ and $f_2(x)$ are two convex functions of x . $Q^{-1}(\varepsilon)$ is a positive constant, so the objective function $R(x)$ is the difference of two convex functions [15]. Using the first-order Taylor expansion of $f_2(x)$, we have:

$$\log_2(1+x) - A_k x - B_k \quad (10)$$

where $A_k = aQ^{-1}(\varepsilon)f_2'(x^t)$, and $B_k = aQ^{-1}(\varepsilon)(f_2(x^t) - f_2'(x^t)x^t)$. Thus, the objective function for the optimal solution of this problem can be expressed in terms of the transmission rate in equation (10).

Next, we aim to optimize the beamforming vector $\mathbf{w}_k$ at the SC, assuming that the IRS reflection vector $\mathbf{u}$ is fixed. The following quantities are

defined: $\mathbf{W}_k \triangleq \mathbf{w}_k \mathbf{w}_k^H$; $\mathbf{M}_k \triangleq \mathbf{m}_k \mathbf{m}_k^H$; $\mathbf{m}_k^H = \mathbf{h}_{d,k}^H + \mathbf{h}_{r,k}^H \Phi \mathbf{G}$; $\mathbf{m}_k = \mathbf{h}_{d,k} + \mathbf{G}^H \Phi \mathbf{h}_{r,k}$. Therefore, the SINR in (5) can be rewritten as:

$$\gamma_{s,k} = \frac{\text{Tr}(\mathbf{W}_k \mathbf{M}_k)}{\sum_{i \in \mathcal{U}_S, i \neq k} \text{Tr}(\mathbf{W}_i \mathbf{M}_k) + \rho_k + \sigma_k^2}. \quad (11)$$

This formula expresses SINR as the ratio of the useful signal power to the sum of interference and noise powers for user k -th at the SC.

3.2. Optimization of the transmit beamforming vector at the SC

From (11), the optimization problem in expression (9) can be rewritten as follows:

$$\left. \begin{array}{l} \max_{\mathbf{w}_k} \sum_{k \in \mathcal{U}_S} \log_2(1 + \gamma_{s,k}) - A_k \gamma_{s,k} - B_k \\ \text{s.t.} \\ \text{C1, C2, C3: } \sum_{k \in \mathcal{U}_S} \text{Tr}(\mathbf{W}_k) \leq P_{\max}, \\ \text{C5: } \mathbf{W}_k \geq 0, \quad \forall k \in \mathcal{U}_S, \\ \text{C6: } \text{rank}(\mathbf{W}_k) \leq 1, \quad \forall k \in \mathcal{U}_S. \end{array} \right\} \quad (12)$$

Here, constraints C5 and C6 are introduced to ensure that $\mathbf{W}_k \triangleq \mathbf{w}_k \mathbf{w}_k^H$ remains valid after optimization. Next, an approximate treatment is made to the problem in expression (12). First, the objective function is expressed as the difference between two convex functions (DC programming) [16], $J_1 - J_2$, where

$$J_1 = \sum_{k \in \mathcal{U}_S} \log_2 \left(\sum_{i \in \mathcal{U}_S} \text{Tr}(\mathbf{W}_i \mathbf{M}_k) + \rho_k + \sigma_k^2 \right) \quad (13)$$

and

$$J_2 = \sum_{k \in \mathcal{U}_S} \log_2 \left(\sum_{i \in \mathcal{U}_S, i \neq k} \text{Tr}(\mathbf{W}_i \mathbf{M}_k) + \rho_k + \sigma_k^2 \right) \quad (14)$$

For any positive semi-definite matrix $\mathbf{W}_k, k \in \mathcal{U}_S$ where t represents the iteration index, we use the first-order Taylor expansion for J_2 at $\mathbf{W}_k^{(t_1)}$:

$$J_2(\mathbf{W}_k) \leq J_2(\mathbf{W}_k^{(t_1)}) + \sum_{k \in \mathcal{U}_S} \text{Tr}(\nabla_{\mathbf{W}_k}^H J_2(\mathbf{W}_k^{(t_1)})(\mathbf{W}_k - \mathbf{W}_k^{(t_1)})) \quad (15)$$

where $\nabla_{\mathbf{W}_k}^H J_2(\mathbf{W}_k^{(t_1)})$ is the gradient of J_2 with respect to $\mathbf{W}_k$. Therefore, the objective function can be rewritten as:

$$\sum_{k \in \mathcal{U}_S} \log_2(1 + \gamma_k) \geq J_1 - J_2(\mathbf{W}_k^{(t_1)}) - \sum_{k \in \mathcal{U}_S} \text{Tr}(\nabla_{\mathbf{W}_k}^H J_2(\mathbf{W}_k^{(t_1)})(\mathbf{W}_k - \mathbf{W}_k^{(t_1)})) \triangleq N_k \quad (16)$$

Since constraint C6 is also a non-convex constraint, to solve this type of problem, the semidefinite relaxation (SDR) technique [17] is employed. We define J_3 and J_4 to represent the following expressions:

$$J_3 = \text{Tr}(\mathbf{W}_k \mathbf{M}_k) \quad (17)$$

and

$$J_4 = \sum_{j \in \mathcal{K}, j \neq k} \text{Tr}(\mathbf{W}_j \mathbf{M}_k) + \rho_k + \sigma_k^2. \quad (18)$$

Both J_3 and J_4 are function of $\mathbf{W}_k$. According to reference [18], for $x > 0$, where $x \in \mathbb{R}$, x can be treated as a specific value after each iteration, so the following inequality (19) holds:

$$\frac{1}{x} \geq \frac{2}{\bar{x}} - \frac{x}{\bar{x}^2} \quad (19)$$

Substituting $\gamma_k = \frac{J_3}{J_4}$, we get:

$$\frac{J_3}{J_4} \geq J_3 \left(\frac{2}{\bar{J}_4^{(t_1)}} - \frac{J_4}{(\bar{J}_4^{(t_1)})^2} \right) = \frac{2J_3}{\bar{J}_4^{(t_1)}} - \frac{J_3 J_4}{(\bar{J}_4^{(t_1)})^2}. \quad (11)$$

In particular, to approximate the product $J_3 J_4$, we write:

$$J_3 J_4 = \frac{(J_3 + J_4)^2}{4} - \frac{(J_3 - J_4)^2}{4}. \quad (20)$$

Similarly, a first-order Taylor expansion is applied to $(J_3 - J_4)^2$, after this treatment, the result is as follows:

$$\frac{J_3}{J_4} = \frac{2J_3}{\bar{J}_4} - \frac{1}{\bar{J}_4^2} \left(\frac{(J_3 + J_4)^2}{4} - \frac{1}{4}(\bar{J}_4 - \bar{J}_3)^2 \right) + 2(\bar{J}_4 - \bar{J}_3)\mathbf{M}_k(\mathbf{W}_k - \mathbf{W}_k^{(t_1)}), \quad (21)$$

where $\bar{J}_4 = \sum_{i \in \mathcal{U}_S, i \neq k} \text{Tr}(\mathbf{W}_i \mathbf{M}_k) + \rho_k + \sigma_k^2$, and $\bar{J}_3 = \text{Tr}(\mathbf{W}_k^{(t_1)} \mathbf{M}_k)$. Furthermore, a Taylor

approximation is applied to the convex function $(J_3 + J_4)^2$, and by rearranging inequality (11), the following result is obtained:

$$\frac{J_3}{J_4} \geq \frac{2J_3}{\bar{J}_4} - \frac{1}{\bar{J}_4^2} \left(\bar{J}_3 \bar{J}_4 + 2\bar{J}_4 \text{Tr}(\mathbf{M}_k(\mathbf{W}_k - \mathbf{W}_k^{(t_1)})) \right) \triangleq J_{3,4} \quad (22)$$

After this processing, the ratios of the two convex functions J_3 / J_4 . Thus, the objective function in expression (12) becomes:

$$\sum_{k \in \mathcal{U}_S} R(\gamma_{s,k}) = \sum_{k \in \mathcal{U}_S} N_k - A_k J_{3,4} - B_k. \quad (23)$$

Next, constraint C2 needs to be addressed. According to the minimum SINR requirement for macro-cell users, MRT precoding is applied for processing. Therefore, constraint C2 can be rewritten as

$$\overline{\text{C2}}: \sum_{k \in \mathcal{U}_S} \text{Tr}(\mathbf{W}_k \mathbf{H}_{r,l}) + \rho_l + \sigma_l^2 \leq \frac{1}{\text{SINR}_{\text{BS}}}, \quad \forall l \in (\mathcal{U}_B - \mathcal{U}_S). \quad (24)$$

Due to $\mathbf{W}_k \triangleq \mathbf{w}_k \mathbf{w}_k^H$, it is included in constraint C6. To solve the problem, the semidefinite relaxation (SDR) technique is applied, and the original optimization problem (12) is reformulated as:

$$\left. \begin{aligned} & \max_{\mathbf{W}_k} \sum_{k \in \mathcal{U}_S} N_k - A_k J_{3,4} - B_k \\ & \text{s.t.} \\ & \text{C1, } \overline{\text{C2}}, \overline{\text{C3}}: \sum_{k \in \mathcal{U}_S} \text{Tr}(\mathbf{W}_k) \leq P_{\max}, \forall k \in \mathcal{U}_S \\ & \text{C5: } \mathbf{W}_k \geq 0, \forall k \in \mathcal{U}_S. \end{aligned} \right\} \quad (25)$$

The paper transforms the original non-convex beamforming optimization of problem (9) into an iterative convex problem (25) by using successive convex approximation and semidefinite relaxation. Specifically, $\mathbf{W}_k$ is optimized directly under constraints $\mathbf{W}_k \geq 0$ (C5) and power limits, while dropping the non-convex rank-1 constraint (C6, $\text{rank}(\mathbf{W}_k) = 1$). After iterative convergence of the SCA algorithm Table 1, the rank-1 approximation $\mathbf{w}_k$ is extracted from the dominant eigenvector of $\mathbf{W}_k$, leveraging SDR's

property that optimal solutions are often rank-1 or near-rank-1 in such beamforming problems. Specifically, expression (25) is now a convex optimization problem that can be solved using standard convex optimization tools. To obtain the optimal solution, an iterative algorithm is employed to solve expression (25) for $\mathbf{W}_k$ across time slots t . The proposed algorithm is summarized in Table 1.

Table 1. Iterative beamforming vector design algorithm based on successive convex approximation.

Initialize the maximum number of iterations, $t^{\max}$, the iteration index t_1 , and the variable set $\{\mathbf{W}_k^{(t_1)}\}$; set the optimization variables $\mathbf{u} = [u_1, u_2, \dots, u_N]^H$ as a constant.
1. For $i = 1, 2, \dots$, do:
2. Under the given values of $\mathbf{W}_k^{(t_1)}$ and $\mathbf{u}$, solve problem (25) to obtain $\mathbf{W}_k^{(t_1+1)}$.
3. Update the iteration index: $t_1 = t_1 + 1$
4. Repeat steps 2–3 until convergence or until $t_1 = t_1^{\max}$
5. End for

This SDR-SCA approach effectively handles the non-convex SINR ratios and DC-programming objective by successive Taylor expansions with respect to expressions (10, 15, 21, 22) and inequality approximations (19), yielding a solvable convex SDP at each iteration. Simulations demonstrate fast convergence within 400 iterations and superior sum-rate performance over random phase-shift and no-IRS benchmarks, especially with larger IRS elements ($N = 64$) and moderate powers ($P_{\max} = 30$ dBm) as illustrated in Figure 2.

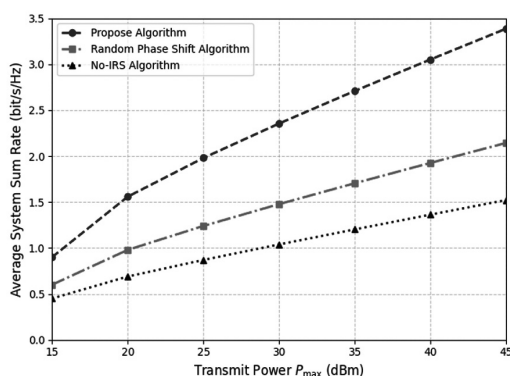


Figure 2. Relationship between $P_{\max}$ of the Small Cell (SC) and the Average Sum System Rate, with $K = 2$ and $N = 16$.

4. NUMERICAL RESULTS

This section presents the simulation results for the proposed system and algorithm. The study considers all communication participants in an IRS-assisted heterogeneous network, which are assumed to be located on the same two-dimensional plane assumed as follows: i) Network Setup: Within the small cell coverage area, there are 2 IRS-assisted small cell users and 2 base station users. The base station region is centered at (0, 0) meters, with the small cell base station located at (20, 0) meters and the base station at (−150, 0) meters. The IRS is positioned at (70, 70) meters; ii) User distribution: Small cell users are randomly distributed within a circle of radius 5 meters centered at (0, 80) meters. Base station users are randomly distributed within a circle of radius 5 meters centered at (0, 25) meters; iii) System parameters: The noise power spectral density is set to −175 dBm/Hz, and the system bandwidth is 240 kHz. The channel is modeled with a Rician factor of 10, and the decoding error probability for lowlatency users is set to $\varepsilon = 10^{-5}$. Each blocklength is $L_k = 11$ bits. The system is equipped with 4 antennas at the base stations, and the IRS is configured with $N = 16, 36$, or 64 reflecting elements. For base station users, the minimum SNR requirement is set to 15 dBm. In this study, the performance evaluation focuses on the average system throughput and the average data rate of small cellusers, which are referred to as the system throughput and average user rate, respectively.

To evaluate and compare the performance of the proposed algorithm, two benchmark algorithms are considered:

(1) Random Phase Shift Algorithm: In this method, each IRS element is assigned a random phase shift. Specifically, for the optimization variable $\mathbf{u}$, a random number is selected uniformly from the interval $[0, 2\pi)$ and fixed at the beginning of the simulation. Subsequently, only the beamforming $\mathbf{W}_k$ is optimized.

(2) No-IRS Algorithm: In this approach, the channels associated with the IRS, namely

$\mathbf{G}_{S,I}$ and $\mathbf{h}_{r,k}$, as well as the phase shift Φ , are removed. Optimization is then performed solely on the beamforming $\mathbf{W}_k$.

Figure 2 illustrates the relationship between the maximum transmitted power (P_{max}) of small cell and the average system sum rate. As P_{max} increases, the average system sum rate shows a monotonically increasing trend. This behavior is influenced by constraint C3 in the optimization problem, where P_{max} directly impacts the feasible range of the beamforming optimization variable $\mathbf{W}_k$. At $P_{max} = 15$ dBm, the three algorithms exhibit different performance characteristics: the proposed optimization algorithm yields the highest average system sum rate, followed by the Random phase shift algorithm, while the “without IRS” case performs the worst. In the range from 15 to 40 dBm, there is a noticeable gap in the rate of increase between the proposed algorithm and the other two. However, the magnitude of this difference remains modest. The performance trend of the random phase shift algorithm closely follows that of the “without IRS” case. For P_{max} in the range of 40 to 45 dBm, the differences among the three algorithms become more pronounced. The proposed algorithm begins to exhibit a steeper growth rate compared to the other two. Similarly, the Random Phase Shift algorithm shows an increase in its growth rate relative to the “without IRS” case, which continues to show the smallest rate of increase. Figure 3 successfully validates the proposed algorithm’s convergence and underscores the pivotal role of the IRS in enhancing network capacity. Specifically, Figure 3 illustrates the relationship between the number of iterations and the average system sum rate (in bits/s/Hz) for a proposed algorithm under varying values of IRS reflecting elements $N = 8, 16, 36, 64$, with a fixed transmit power $P_{max} = 30$ dBm. The results clearly indicate that the proposed algorithm benefits significantly from increased IRS element counts. As N increases, the system sum rate improves across all iteration counts. For instance, at

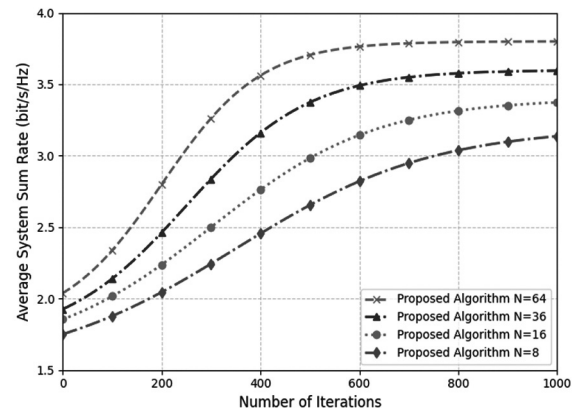


Figure 3. The relationship between number of iterations and the system’s average sum rate with $N = 8, 16, 36, 64$, $P_{max} = 30$ dBm.

1000 iterations, the algorithm with $N = 64$ achieves approximately 3.6 bits/s/Hz, while the configuration with $N = 8$ reaches only around 2.5 bits/s/Hz. In addition, the convergence of the algorithm is evident: all curves exhibit rapid initial growth in the system sum rate, particularly within the first 400 iterations. Beyond this point, the rate of improvement diminishes, indicating convergence toward a performance ceiling. The proposed algorithm demonstrates robust scalability and convergence properties. The consistent upward trend across all configurations confirms its effectiveness in iterative optimization settings.

From Figure 4, it can be observed that as transmitted power at the base station P_{BS} increases, the average system sum rate decreases for all three algorithms. The influence of the BS on the average system sum rate for the SC users can be understood from expressions (9) and (12). In (9), the increase in BS transmission power results in a larger SC component in the denominator, thereby increasing interference for the SC users. This leads to a decrease in the SINR of the SC users, which in turn lowers the average system sum rate. Additionally, for the proposed algorithm, when the transmitted power at the BS is relatively low (e.g., between 20 and 32 dBm), the enhancement effect of the IRS remains significant, and as a result, the average system sum rate decreases more gradually. However,

once the transmitted power at the BS exceeds the 32 dBm threshold, the excessive increase in BS power exacerbates its impact on SINR, causing a sharp acceleration in the rate of decrease of the average system sum rate. Moreover, it can be noted that the proposed algorithm does not always outperform the random phase shift and no-IRS cases. When the decoding error probability is set to $\varepsilon = 10^{-7}$, the average system sum rate is lower compared to all three algorithms when $\varepsilon = 10^{-5}$. This indicates that, as the transmitted power at the BS increases, the change in ε has a more substantial impact on the average system sum rate than in the random phase shift and “No-IRS” scenarios.

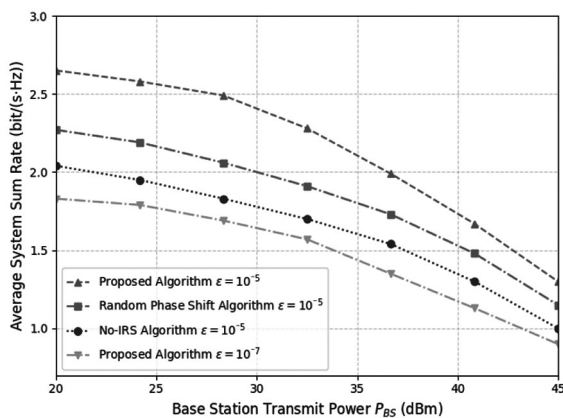


Figure 4. The relationship between the maximum transmitted power of base station (P_{BS}) and the system’s average sum rate with $K = 2$, $P_{max} = 30$ dBm.

5. CONCLUSION

In this study, we proposed an efficient beamforming optimization algorithm for IRS-assisted heterogeneous wireless networks, targeting the enhancement of system throughput under URLLC constraints. By reformulating the original nonconvex problem using convex approximation techniques and iterative optimization, the algorithm achieves notable improvements in average system sum rate and convergence speed. Simulation results validate the superiority of the proposed method over traditional random phase shift and no-IRS schemes, particularly in scenarios with increased IRS elements and moderate transmission power. Furthermore, the analysis reveals the sensitivity

of system performance to base station power and decoding error probability, highlighting the importance of careful parameter tuning. Overall, the proposed approach offers a promising solution for future high-performance wireless communication systems leveraging IRS technology.

DECLARATION OF CONFLICTS OF INTEREST

The authors declare that this research has no conflicts of interest.

ACKNOWLEDGMENTS

This research is conducted within the framework of science and technology projects at institutional level of Quy Nhon University under the project code T2025.896.16.

REFERENCES

- [1] C. Nwufo, O. Simpson, Y. Sun, “Reconfigurable Intelligent Surfaces (RIS) and Their Role in Next-Generation Wireless Networks: An Overview,” *Transactions on Emerging Telecommunications Technologies*, vol. 36, no. 6, p. e70164, 2025.
- [2] S. Rajoria, A. Trivedi, W. Wilfred Godfrey, “A comprehensive survey: Small cell meets massive MIMO,” *Physical Communication*, vol. 26, pp. 40-49, 2018.
- [3] F. Qamar, S. H. A. Kazmi, K. A. Z. Ariffin, M. Tayyab, Q. N. Nguyen, “Multiantenna array-based massive MIMO for b5g/6g: State of the art, challenges, and future research directions,” *Information*, vol. 15, no. 8, p. 442, 2024.
- [4] N. H. Mahmood, I. Atzeni, E. A. Jorswieck, O. L. A. López, “Ultra-reliable lowlatency communications: Foundations, enablers, system design, and evolution towards 6G,” *Foundations and Trends® in Communications and Information Theory*, vol. 20, no. 5-6, pp. 512-747, 2023.
- [5] D. Feng, L. Lai, J. Luo, Y. Zhong, K. Ying, “Ultra-reliable and low-latency communications: applications, opportunities and challenges,”

- Science China Information Sciences*, vol. 64, no. 2, p. 120301, 2021.
- [6] W. Li, H. Chongwen, G. C. Alexandropoulos, C. Yuen, Z. Zhang, M. Debbah "Channel estimation for RIS-empowered multiuser MISO wireless communications," *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 4144-4157, 2021.
- [7] Y. Xiu, J. Zhao, C. Yuen, Z. Zhang, G. Gui, "Secure beamforming for multiple intelligent reflecting surfaces aided mmWave systems," *IEEE Communications Letters*, vol. 25, no. 2, pp. 417-421, 2021.
- [8] Y. Xu, G. Gui, H. Gacanin, F. Adachi, "A survey on resource allocation for 5G heterogeneous networks: Current research, future trends, and challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no.2, pp. 668-695, 2021.
- [9] Y. Xu, G. Gui, T. Ohtsuki, H. Gacanin, B. Adebisi, H. Sari, F. Adachi, "Robust resource allocation for two-tier HetNets: An interference efficiency perspective," *IEEE Transactions on Green Communications and Networking*, vol. 5, no. 3, pp. 1514-1528, 2021.
- [10] B. Zheng, C. You, R. Zhang, "Fast Channel Estimation for IRS-Assisted OFDM," *IEEE Wireless Communications Letters*, vol. 10, no. 3, pp. 580-584, 2021.
- [11] Z. Zhang, L. Dai, "A Joint Precoding Framework for Wideband Reconfigurable Intelligent Surface-Aided Cell-Free Network," *IEEE Transactions on Signal Processing*, vol. 69, pp. 4085-4101, 2021.
- [12] R. Corvaja, A. G. Armada, "Phase Noise Degradation in Massive MIMO Downlink With Zero-Forcing and Maximum Ratio Transmission Precoding," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 10, pp. 8052-8059, 2016.
- [13] Erseghe, Tomaso, "Coding in the FiniteBlocklength Regime: Bounds Based on Laplace Integrals and Their Asymptotic Approximations," *IEEE Transactions on Information Theory*, vol. 62, no. 12, pp. 6854 - 6883, 2016.
- [14] S. He, Z. An, J. Zhu, J. Zhang, Y. Huang, Y. Zhang, "Beamforming Design for Multiuser uRLLC With Finite Blocklength Transmission," *IEEE Transactions on Wireless Communications*, vol. 20, no. 12, pp. 8096-8109, 2021.
- [15] X. Yu, D. Xu, Y. Sun, D. W. K. Ng, R. Schober, "Robust and Secure Wireless Communications via Intelligent Reflecting Surfaces," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2637-2652, 2020.
- [16] S. Hu, Z. Wei, Y. Cai, D. W. K. Ng, J. Yuan. "Sum-Rate Maximization for Multiuser MISO Downlink Systems with Self-sustainable IRS", in *GLOBECOM 2020*, IEEE Global Communications Conference, 2020, pp. 1-7.
- [17] Q. Wu, R. Zhang, "Weighted Sum Power Maximization for Intelligent Reflecting Surface Aided SWIPT," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 586-590, 2020.
- [18] A. A. Nasir, H. D. Tuan, H. H. Nguyen, M. Debbah, H. V. Poor. Resource Allocation and Beamforming Design in the Short Blocklength Regime for URLLC, *IEEE Transactions on Wireless Communications*, vol. 20, no. 2, pp. 1321-1335, 2021.



© 2026 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Nghiên cứu tối ưu chùm tia trong hệ thống truyền không dây có sự trợ giúp của bề mặt phản xạ thông minh

Nguyễn Đỗ Dũng*, Nguyễn Văn Hào, Đào Minh Hưng

Khoa Kỹ thuật và Công nghệ, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 06/10/2025; Ngày sửa bài: 17/12/2025;

Ngày nhận đăng: 23/12/2025; Ngày xuất bản: 28/06/2026

TÓM TẮT

Bài báo này nghiên cứu bài toán tối ưu hóa thiết kế chùm tia trong hệ thống truyền thông không dây có sự hỗ trợ của bề mặt phản xạ thông minh (IRS), nhằm đáp ứng các yêu cầu nghiêm ngặt của giao tiếp độ trễ thấp cực kỳ đáng tin cậy (URLLC). Bằng cách thiết kế các vectơ định dạng chùm tia truyền tại ô nhỏ (SC), thuật toán đề xuất sẽ biến đổi bài toán tối ưu hóa không lồi thành dạng lồi có thể giải được bằng cách sử dụng kỹ thuật xấp xỉ lồi liên tiếp (SCA) và kỹ thuật thư giãn bán xác định (SDR). Kết quả mô phỏng cho thấy phương pháp đề xuất giúp cải thiện đáng kể thông lượng hệ thống và tốc độ dữ liệu của người dùng so với các thuật toán tham chiếu, đặc biệt là trong các cấu hình IRS và mức công suất truyền khác nhau. Nghiên cứu khẳng định hiệu quả và khả năng mở rộng của IRS trong việc nâng cao hiệu suất mạng trong các trường hợp URLLC.

Từ khóa: Bề mặt phản chiếu thông minh (IRS), mạng không đồng nhất, thiết kế chùm tia, giao tiếp độ trễ thấp, cực kỳ đáng tin cậy (URLLC).

*Tác giả liên hệ chính.

Email: nguyendodung@qnu.edu.vn

Splitting piecewise cubic Bézier curves and the equivalence constants for some norms

Hoang Van Duc*

Department of Mathematics and Statistics, Quy Nhon University, Vietnam

Received: 12/11/2025; Revised: 05/03/2026;

Accepted: 06/03/2026; Published: 28/06/2026

ABSTRACT

This article is concerned with splitting piecewise cubic Bézier curves and the equivalence constants equivalence relations for some norms defined through control points. Piecewise cubic Bézier curves are most common to approximate curves. These curves are established by control points. The norms $\|\cdot\|_p^{B_{N,3}}$ and $\|\cdot\|_p^{B_{2N,3}}$ on the space $B_{N,3}$ are determined through control points. With the purpose of keeping the degree of the curves and offering additional flexibility for curve design, we often split an N -piece cubic Bézier curve to become a $2N$ -piece cubic Bézier curve. We will concentrate on the equivalence constants for the norm $\|\cdot\|_p^{B_{N,3}}$ and the norm $\|\cdot\|_p^{B_{2N,3}}$ on the space $B_{N,3}$ of N -piece cubic Bézier curves. So, we can use the norm $\|\cdot\|_p^{B_{N,3}}$ to check the convergence for sequences of piecewise cubic Bézier curves. This result is important for applying piecewise cubic Bézier curves to detect optimal orbits.

Keywords: *Cubic Bézier curves, equivalence constants, norm, split, piecewise.*

1. INTRODUCTION

In mathematics and engineering, there are many curves which have complex shapes or are defined by a set of points. To overcome this difficulty, we create a new curve that closely matches an existing one, often to simplify a complex shape or to fit a set of data points. Some methods to approximate the curve include using simpler curves or interpolating polynomials or Bézier curves.

In the method using simpler curves to approximate the curve, we divide the curve into a series of points and connect them with straight lines or circular arcs, with more points resulting in a closer approximation. The main advantage of approximating curves with straight lines or circular arcs is computational simplicity and efficiency, as line segments are defined by

fewer parameters than higher-order curves. This method is also useful for data compression and noise reduction by simplifying complex curves. However, approximating a curve with simple curves can lead to inaccuracies, especially in areas of high curvature, and may result in poor extrapolation beyond the measured data range [1] - [9].

Besides, a curve can be approximated by an interpolating polynomial such as a Lagrange polynomial or a Newton polynomial, which fits a curve through a set of known points on the curve. This method improve the approximation as the degree of the polynomial increases. Interpolating polynomials passing exactly through specified points. The main advantages of using interpolating polynomials for approximation include their high accuracy

*Corresponding author:

Email: hoangvanduc@qnu.edu.vn

for small datasets, the ability to obtain an explicit function for calculations, and the ease of differentiation and integration. They also provide exact results at the given data points and can be used for data points that are not equally spaced. The main disadvantages of using interpolating polynomials for curve approximation are Runge's phenomenon (oscillations, especially at the endpoints), computational expense for high-degree polynomials, and poor extrapolation properties, where the curve can behave erratically outside the range of the data points (see [10] - [12]).

Approximating a curve with a Bézier curve involves selecting key points on the original curve to serve as control points and endpoints for the Bézier curve(s). For a given curve, this is often achieved by dividing it into segments and approximating each segment with a Bézier curve, using techniques like the de Casteljau's algorithm for subdivision or fitting algorithms like the Adaptive Extension Fitting Scheme to find the optimal control points for a set of segments. Bézier's construction has many benefits. Initially, each Bézier curves is presented by a few control points, then it requires very little memory. Besides, these curves is intuitive, compact and beautiful. It's easy to compute and design, so the designer without mathematical background can use them. More, we can easily change, move, turn Bézier curves just by changing, moving, turning their control points (see [1], [13] - [17]).

In 1959, the mathematician Paul de Casteljau built Bézier curve by using de Casteljau's algorithm while working for the French automaker Citroen. He was the first to apply this method to computer-aided design (CAD). However, his work remained a company secret and was not published for many years. So, his contributions were not widely known at the time. The Bézier curve was publicized by the French engineer Pierre Bézier in 1962. He defined the Bézier curve based on Bernstein polynomials. Pierre Bézier applied Bézier curves

for designing the bodywork of Renault cars. Bézier built definitions, symbols, formulas and specially control points of Bézier curves. These things make it easy and convenient to represent curves in computers software.

The computer program language PostScript uses Bézier curves as the standard curve. Many vector graphics editor and design software such as CorelDRAW, Adobe Illustrator and Inkscape apply Bézier curves. Its importance is due to the fact that, Bézier curves are used in many areas of life, not only mathematics. Bézier curves are applied in computer graphics, computer-aided design system, robotic, industry, walking, communication, path-planning and aerospace (see [15] - [22]). Bézier curves are also used to find plane shape optimization which appears in many fields such as environment design, aerospace, structural mechanics, networks, automotive, hydraulic, oceanology and wind engineering (see [23] - [28]).

Bézier curves are presented in many books and articles for instance [1], [13], [14]. A continuous curve can be approximated by a Bézier curve. However, when the curve is long and complex, the degree of the Bézier curve is high. As a result, the computation is more difficult. Then, the most common use of Bézier curves is as N -piece cubic Bézier curves. We will focus uniform N -piece cubic Bézier curves.

From [29], we have the norm $\|\cdot\|_p^{B_m}$ on the space B_m of Bézier curves of degree m and the norm $\|\cdot\|_p^{B_{N,m}}$ on the space $B_{N,m}$ of uniform N -piece Bézier curves of degree m . These norms are computed through control points.

With the purpose of keeping the degree of the curves and offering additional flexibility for curve design, we often split N -piece uniform cubic Bézier curves to become $2N$ -piece uniform cubic Bézier curves. Splitting piecewise cubic Bézier curves keep huge part of applying these curves. Then, we study the equivalence constants for the norm $\|\cdot\|_p^{B_{2N,3}}$ and the norm $\|\cdot\|_p^{B_{N,3}}$ on the space of piecewise cubic Bézier curves.

Theorem 1.1. Let $p \in [1, \infty]$ and let $\beta \in B_{N,3}$ be an N -piece cubic Bézier curve. Then

$$\min \left\{ \frac{1}{24^{1/p}}, \frac{1}{4} \right\} \|\beta\|_p^{B_{N,3}} \leq \|\beta\|_p^{B_{2N,3}} \leq 3^{\frac{1}{p}} \|\beta\|_p^{B_{N,3}}.$$

2. PRELIMINARIES

We recall some definitions and notations that will be used through the article

Definition 2.1. (in [12]) Given four points W_0, W_1, W_2 and W_3 , the cubic Bézier curve associated with the four control points $W_0, \dots, W_3$ is defined by

$$B([W_0, \dots, W_3], t) := \sum_{i=0}^3 W_i b_{i,3}(t) \quad (1)$$

for $t \in [0,1]$,

where $b_{i,3}(t) = \binom{3}{i} t^i (1-t)^{3-i}$ is the Bernstein polynomial.

The points W_0, W_1, W_2 and W_3 are control points of the cubic Bézier curve. The quadrilateral $W_0 W_1 W_2 W_3$ is the control polygon of the curve. The cubic Bézier curve lies in its control polygon.

For any cubic Bézier curve

$$\begin{aligned} \beta(t) &= B([W_0, W_1, W_2, W_3], t) = \sum_{i=0}^3 W_i b_{i,3}(t) \\ &= (1-t)^3 W_0 + 3(1-t)^2 t W_1 \\ &\quad + 3(1-t) t^2 W_2 + t^3 W_3, \end{aligned}$$

for $t \in [0,1]$,

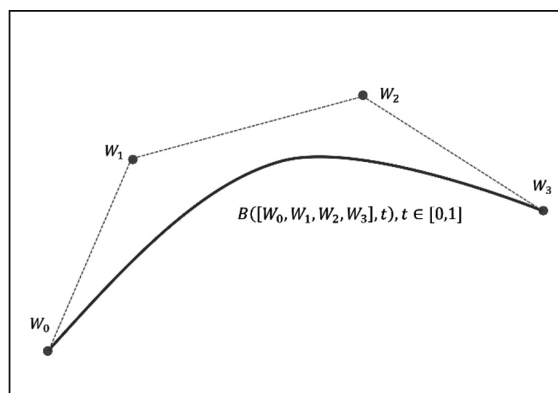


Figure 1. A cubic Bézier curve.

From the recursive property of Bernstein polynomials, a cubic Bézier curve can be

recursively determined as a convex combination of two quadratic Bézier curves as

$$\begin{aligned} \beta(t) &= B([W_0, W_1, W_2, W_3], t) \\ &= (1-t)((1-t)^2 W_0 + 2(1-t)t W_1 + t^2 W_2) \\ &\quad + t((1-t)^2 W_1 + 2(1-t)t W_2 + t^2 W_3) \\ &= (1-t)B([W_0, W_1, W_2], t) \\ &\quad + tB([W_1, W_2, W_3], t). \end{aligned}$$

Since $b'_{i,n}(x) = n(b'_{i-1,n-1}(x) - b'_{i,n-1}(x))$. The derivative with respect to t of the cubic Bézier curve is another Bézier curve which has a lower degree as follows

$$\begin{aligned} \frac{d}{dt} \beta(t) &= \frac{d}{dt} B([W_0, W_1, W_2, W_3], t) \\ &= 3(1-t)^2 (W_1 - W_0) + 6(1-t)t (W_2 - W_1) \\ &\quad + 3t^2 (W_3 - W_2) \\ &= B([W_1 - W_0, W_2 - W_1, W_3 - W_2], t). \end{aligned}$$

A uniform N -piece cubic Bézier curve is formed by the combination of N cubic Bézier curves and the connecting points of the pieces are the points at $t = \frac{j}{N}$ for $j = 0, \dots, N-1$. For convenience, we will omit the word “uniform”.

Definition 2.2. (see [12]) Let N be a positive integer and $W_0, \dots, W_{N+1}$ be $2N+1$ points in $\mathbb{R}^n$. The N -piece cubic Bézier curve associated with control points $W_0, \dots, W_{N+1}$ is defined by

$$\begin{aligned} \beta: [0,1] &\rightarrow \mathbb{R}^n \\ t \mapsto \beta(t) &= B([W_{3j}, \dots, W_{3j+3}], Nt - j) \\ &\quad \text{if } t \in \left[\frac{j}{N}, \frac{j+1}{N} \right]. \end{aligned}$$

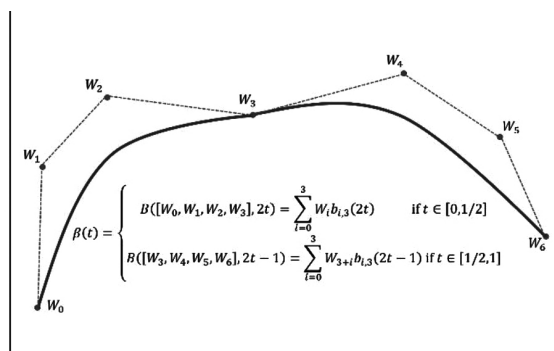


Figure 2. A two-piece cubic Bézier curve.

Notation 2.3.

- The vector space of cubic Bézier curves is denoted by the symbol B_3 .
- The vector space of N -piece cubic Bézier curves is denoted by the symbol $B_{N,3}$.

We will define some norms and distances through control points on the space B_3 of cubic Bézier curves and on the space $B_{N,3}$ of N -piece cubic Bézier curves.

Definition 2.4. Let $p \in [1, \infty]$. The function $\|\cdot\|_p^{B_3}: B_3 \rightarrow \mathbb{R}$ is defined by: For any $\beta(t) = \sum_{i=0}^3 W_i b_{i,3}(t) \in B_3$,

$$\|\beta\|_p^{B_3} := \begin{cases} \left(\sum_{i=0}^3 \|W_i\|_p^p \right)^{1/p} & \text{if } p \in [1, p[\\ \max_{i=0, \dots, 3} \{\|W_i\|_\infty\} & \text{if } p = \infty, \end{cases}$$

where $\|\cdot\|_p$ is the p -norm on $\mathbb{R}^n$.

Using the Minkowski inequality and the properties of the p -norm on $\mathbb{R}^n$, it is easily seen that $\|\cdot\|_p^{B_3}$ is a norm on the vector space B_3 of cubic Bézier curves. Indeed, this is a norm of their control polygons. Naturally, we get an induced distance on B_3 by $d_p^{B_3}(\beta, \gamma) := \|\beta - \gamma\|_p^{B_3}$ for $\beta, \gamma \in B_3$.

Definition 2.5. Let $p \in [1, \infty]$. The function $\|\cdot\|_p^{B_{N,3}}: B_{N,3} \rightarrow \mathbb{R}$ is defined by: For any $\beta(t) = \beta^{(j)}(Nt - j) = \sum_{i=0}^3 W_{j+3i} b_{i,3}(Nt - j)$

if $t \in \left[\frac{j}{N}, \frac{j+1}{N}\right], j = 0, \dots, N-1$,

$$\|\beta\|_p^{B_{N,3}} := \begin{cases} \frac{1}{N^{1/p}} \left(\sum_{j=0}^{N-1} \left(\|\beta^{(j)}\|_p^{B_3} \right)^p \right)^{1/p} & \text{if } p \in [1, \infty[\\ \max_{j=0, \dots, N-1} \left\{ \|\beta^{(j)}\|_\infty^{B_3} \right\} & \text{if } p = \infty. \end{cases}$$

By $\|\cdot\|_p^{B_3}$ is a norm on B_3 and the Minkowski inequality, it is easily seen that $\|\cdot\|_p^{B_{N,3}}$ is a norm on the vector space $B_{N,3}$ of N -piece cubic Bézier curves. Naturally, we get an induced distance on $B_{N,3}$ determined by $d_p^{B_{N,3}}(\beta, \gamma) := \|\beta - \gamma\|_p^{B_{N,3}}$ for any $\beta, \gamma \in B_{N,3}$.

The norm $\|\cdot\|_p^{B_{N,3}}$ is determined through the control points. It is more convenient than the L_p norm. From [29], we have the equivalence relations between the norms $\|\cdot\|_p^{B_{N,3}}$ and L_p as follows:

For $p \in [1, \infty]$. The inequalities

$$\|\beta\|_{L_p} \leq \|\beta\|_p^{B_{N,3}} \leq 2^{10} \|\beta\|_{L_p}$$

hold for any $\beta \in B_{N,3}$.

We can present an N -piece cubic Bézier curve as an N -piece Bézier curve of degree 4 (see more [29]). So, The norm $\|\cdot\|_p^{B_{N,3}}$ and the norm $\|\cdot\|_p^{B_{N,4}}$ are two norms on the space $B_{N,3}$ of N -piece cubic Bézier curves. By [29], we get a corollary about equivalence constants between these norms as follows:

For $p \in [1, \infty]$. The inequalities

$$\frac{1}{8} \|\beta\|_p^{B_{N,3}} \leq \|\beta\|_p^{B_{N,4}} \leq 2 \|\beta\|_p^{B_{N,3}}$$

hold for any $\beta \in B_{N,3}$.

When we need more freedom for designing the curve, we will increase the number of piece in piecewise cubic Bézier curves. To solve this problem, we will split piecewise cubic Bézier curves. By [24], a Bézier curve β can be split at any $t_0 \in (0,1)$ to become two-piece Bézier curves. However, when splitting at $t \neq \frac{1}{2}$, the connecting points of the pieces do not coincide with the point $\beta\left(\frac{1}{2}\right)$. This means that the obtained curve is not uniform. So, we need to split a cubic Bézier curves at $t = \frac{1}{2}$ to get a uniform two-piece cubic Bézier curve as follows:

$$\begin{aligned} \beta(t) &= B([W_0, W_1, W_2, W_3], t) \\ &= \sum_{i=0}^3 W_i b_{i,3}(t), t \in [0,1], \end{aligned}$$

we have

$$\beta(t) = \begin{cases} \beta^{(0)}(2t) = \sum_{i=0}^3 P_i b_{i,3}(2t) & \text{if } t \in \left[0, \frac{1}{2}\right] \\ \beta^{(1)}(2t-1) = \sum_{i=0}^3 P_{3+i} b_{i,3}(2t-1) & \text{if } t \in \left[\frac{1}{2}, 1\right], \end{cases} \quad (2)$$

where

$$\begin{cases} P_i = \sum_{l=0}^i b_{l,i} \left(\frac{1}{2}\right) P_{i-l}, & i = 0, \dots, 3, \\ P_{3+i} = \sum_{l=0}^i b_{l,i} \left(\frac{1}{2}\right) P_{3-i+l}, & i = 0, \dots, 3. \end{cases}$$

Thus, we can consider a cubic Bézier curve as a uniform two-piece cubic Bézier curve.

More generally, let $\beta \in B_{N,3}$ be an N -piece cubic Bézier curve with control points $W_{j+3+i} \in \mathbb{R}^n, i = 0, \dots, 3, j = 0, \dots, N-1$. We have

$$\beta(t) = \beta^{(j)}(Nt - j) = \sum_{i=0}^3 W_{j+3+i} b_{i,3}(Nt - j) \quad \text{if } t \in \left[\frac{j}{N}, \frac{j+1}{N}\right], j = 0, \dots, N-1.$$

In order to get a uniform $2N$ -piece cubic Bézier curve β , we need to split at the middle point of each piece.

$$\beta(t) = \begin{cases} \beta^{(2j)}(2Nt - 2j) = \sum_{i=0}^3 P_{2j+3+i} b_{i,3}(2Nt - 2j) \\ \quad \text{if } t \in \left[\frac{2j}{2N}, \frac{2j+1}{2N}\right], \\ \beta^{(2j+1)}(2Nt - 2j - 1) \\ \quad = \sum_{i=0}^3 P_{(2j+1)3+i} b_{i,3}(2Nt - 2j - 1) \\ \quad \text{if } t \in \left[\frac{2j+1}{2N}, \frac{2j+2}{2N}\right], \\ \quad j = 0, \dots, N-1, \end{cases} \quad (3)$$

where

$$\begin{cases} P_{2j+3+i} = \sum_{l=0}^i b_{l,i} \left(\frac{1}{2}\right) W_{j+3+i-l}, \\ P_{(2j+1)3+i} = \sum_{l=0}^i b_{l,i} \left(\frac{1}{2}\right) W_{j+3-i+l}, \\ i = 0, \dots, 3, j = 0, \dots, N-1. \end{cases}$$

So, β is also a $2N$ -piece cubic Bézier curve and the space $B_{N,3}$ of N -piece cubic Bézier curves is a subspace of the space $B_{2N,3}$ of $2N$ -piece cubic Bézier curves. Therefore, the norms $\|\cdot\|_p^{B_{2N,3}}$ and $\|\cdot\|_p^{B_{N,3}}$ are two norms on the space

$B_{N,3}$. Next, we will consider the equivalence relations between these norms.

3. EQUIVALENCE CONSTANTS FOR THE NORM $\|\cdot\|_p^{B_{2N,3}}$ AND $\|\cdot\|_p^{B_{N,3}}$ ON $B_{N,3}$

We first show a constant K such that $\|\cdot\|_p^{B_{2N,3}} \leq K \|\cdot\|_p^{B_{N,3}}$ on $B_{N,3}$. We will consider two cases $p \in [1, \infty[$ and $p = \infty$.

Lemma 3.1. Let $p \in [1, \infty[$ and let $\beta \in B_{N,3}$,

$$\|\beta\|_p^{B_{2N,3}} \leq 3^{\frac{1}{p}} \|\beta\|_p^{B_{N,3}}.$$

Proof. For any N -piece cubic Bézier curve associated with control points $W_{j+3+i} \in \mathbb{R}^n, i = 0, \dots, 3, j = 0, \dots, N-1$. We have

$$\beta(t) = \beta^{(j)}(Nt - j) = \sum_{i=0}^3 W_{j+3+i} b_{i,3}(Nt - j) \quad \text{if } t \in \left[\frac{j}{N}, \frac{j+1}{N}\right], j = 0, \dots, N-1.$$

By (3), we split β to become a $2N$ -piece cubic Bézier curve as follows

$$\beta(t) = \begin{cases} \beta^{(2j)}(2Nt - 2j) = \sum_{i=0}^3 P_{2j+3+i} b_{i,3}(2Nt - 2j) \\ \quad \text{if } t \in \left[\frac{2j}{2N}, \frac{2j+1}{2N}\right], \\ \beta^{(2j+1)}(2Nt - 2j - 1) \\ \quad = \sum_{i=0}^3 P_{(2j+1)3+i} b_{i,3}(2Nt - 2j - 1) \\ \quad \text{if } t \in \left[\frac{2j+1}{2N}, \frac{2j+2}{2N}\right], \\ \quad j = 0, \dots, N-1, \end{cases}$$

where

$$\begin{cases} P_{2j+3+i} = \sum_{l=0}^i b_{l,i} \left(\frac{1}{2}\right) W_{j+3+i-l}, \\ P_{(2j+1)3+i} = \sum_{l=0}^i b_{l,i} \left(\frac{1}{2}\right) W_{j+3-i+l}, \\ i = 0, \dots, 3, j = 0, \dots, N-1. \end{cases}$$

Case $p \in [1, \infty[$. Since

$$\begin{aligned}
 (\|\Gamma^{(2j)}\|_p^{B_3})^p &= \sum_{i=0}^3 \left\| \sum_{l=0}^i b_{l,i} \left(\frac{1}{2}\right) W_{j3+i-l} \right\|_p^p \\
 &\leq 3 \max_{i=0,\dots,3} \|W_{j3+i}\|_p^p \\
 &\leq 3 \sum_{i=0}^3 \|W_{j3+i}\|_p^p \\
 &= 3 (\|\beta^{(j)}\|_p^{B_3})^p, \\
 \forall j &= 0, \dots, N-1,
 \end{aligned}$$

and similarly

$$\begin{aligned}
 &(\|\Gamma^{(2j+1)}\|_p^{B_3})^p \\
 &= \sum_{i=0}^3 \left\| \sum_{l=0}^i b_{l,i} \left(\frac{1}{2}\right) W_{j3+3-i+l} \right\|_p^p \\
 &\leq 3 (\|\beta^{(j)}\|_p^{B_3})^p, \forall j = 0, \dots, N-1,
 \end{aligned}$$

we obtain

$$\begin{aligned}
 &\|\beta\|_p^{B_{2N,3}} \\
 &= \frac{1}{(2N)^{1/p}} \left(\sum_{j=0}^{N-1} (\|\Gamma^{(2j)}\|_p^{B_3})^p \right. \\
 &\quad \left. + (\|\Gamma^{(2j+1)}\|_p^{B_3})^p \right)^{1/p} \\
 &\leq \frac{1}{(2N)^{\frac{1}{p}}} \left(\sum_{j=0}^{N-1} 6 (\|\beta^{(j)}\|_p^{B_3})^p \right)^{\frac{1}{p}} \leq 3^{\frac{1}{p}} \|\beta\|_p^{B_{N,3}}.
 \end{aligned}$$

Case $p = \infty$. Since

$$\begin{aligned}
 \|\Gamma^{(2j)}\|_\infty^{B_3} &= \max_{i=0,\dots,3} \left\| \sum_{l=0}^i b_{l,i} \left(\frac{1}{2}\right) W_{j3+i-l} \right\|_\infty \\
 &\leq \max_{i=0,\dots,3} \|W_{j3+i}\|_\infty \\
 &= \|\beta^{(j)}\|_\infty^{B_3}, \forall j = 0, \dots, N-1
 \end{aligned}$$

and similarly

$$\begin{aligned}
 &\|\Gamma^{(2j+1)}\|_\infty^{B_3} \\
 &= \max_{i=0,\dots,3} \left\| \sum_{l=0}^i b_{l,i} \left(\frac{1}{2}\right) W_{j3+3-i+l} \right\|_\infty \\
 &\leq \|\beta^{(j)}\|_\infty^{B_3}, \forall j = 0, \dots, N-1,
 \end{aligned}$$

we obtain

$$\begin{aligned}
 &\|\beta\|_\infty^{B_{2N,3}} \\
 &= \max_{j=0,\dots,N-1} \max \{ \|\Gamma^{(2j)}\|_\infty^{B_3}, \|\Gamma^{(2j+1)}\|_\infty^{B_3} \} \\
 &\leq \max_{j=0,\dots,N-1} \|\beta^{(j)}\|_\infty^{B_3} = \|\beta\|_\infty^{B_{N,3}}.
 \end{aligned}$$

From the above two cases, we have the proof of the lemma. ■

In order to show a constant k such that $k \|\cdot\|_p^{B_{2N,3}} \leq \|\cdot\|_p^{B_{N,3}}$ on the space $B_{N,3}$, we also study two case $p \in [1, \infty[$ and $p = \infty$.

Lemma 3.2. Let $p \in [1, \infty]$ and let $\beta \in B_{N,3}$, we have

$$\frac{1}{24^{1/p}} \|\beta\|_p^{B_{N,3}} \leq \|\beta\|_p^{B_{2N,3}}.$$

Proof. For any N -piece cubic Bézier curve associated with control points $W_{j3+i} \in \mathbb{R}^n, i = 0, \dots, 3, j = 0, \dots, N-1$. We have

$$\begin{aligned}
 \beta(t) &= \beta^{(j)}(Nt - j) = \sum_{i=0}^3 W_{j3+i} b_{i,3}(Nt - j) \\
 &\text{if } t \in \left[\frac{j}{N}, \frac{j+1}{N} \right], j = 0, \dots, N-1.
 \end{aligned}$$

By (3), we split β to become a $2N$ -piece cubic Bézier curve as follows

$$\beta(t) = \begin{cases} \beta^{(2j)}(2Nt - 2j) = \sum_{i=0}^3 P_{2j3+i} b_{i,3}(2Nt - 2j) \\ \quad \text{if } t \in \left[\frac{2j}{2N}, \frac{2j+1}{2N} \right], \\ \beta^{(2j+1)}(2Nt - 2j - 1) \\ \quad = \sum_{i=0}^3 P_{(2j+1)3+i} b_{i,3}(2Nt - 2j - 1) \\ \quad \text{if } t \in \left[\frac{2j+1}{2N}, \frac{2j+2}{2N} \right], \\ \quad j = 0, \dots, N-1, \end{cases}$$

where

$$\begin{cases} P_{2j3+i} = \sum_{l=0}^i b_{l,i} \left(\frac{1}{2}\right) W_{j3+i-l}, \\ P_{(2j+1)3+i} = \sum_{l=0}^i b_{l,i} \left(\frac{1}{2}\right) W_{j3+3-i+l}, \\ i = 0, \dots, 3, j = 0, \dots, N-1. \end{cases}$$

We first consider $(\|\Gamma^{(2j)}\|_p^{B_3})^p + (\|\Gamma^{(2j+1)}\|_p^{B_3})^p, j = 0, \dots, N-1$. Set

$A =$

$$\max \left\{ \|W_{j3}\|_p, \frac{1}{2} \|W_{j3+1}\|_p, \frac{1}{2} \|W_{j3+2}\|_p, \|W_{j3+3}\|_p \right\}.$$

- Case 1: $A = \|W_{j3}\|_p$.

We have

$$\begin{aligned} & \left(\|\Gamma^{(2j)}\|_p^{B_3} \right)^p + \left(\|\Gamma^{(2j+1)}\|_p^{B_3} \right)^p \\ & \geq \left(\|\Gamma^{(2j)}\|_p^{B_3} \right)^p \geq \|P_{j3}\|_p^p \\ & = \|W_{j3}\|_p^p \geq \frac{1}{6} \sum_{i=0}^3 \|W_{j3+i}\|_p^p \\ & = \frac{1}{6} \left(\|\beta^{(j)}\|_p^{B_3} \right)^p. \end{aligned}$$

- Case 2: $A = \frac{1}{2} \|W_{j3+1}\|_p$.

We have

$$\begin{aligned} & \left(\|\Gamma^{(2j)}\|_p^{B_3} \right)^p + \left(\|\Gamma^{(2j+1)}\|_p^{B_3} \right)^p \\ & \geq \left(\|\Gamma^{(2j)}\|_p^{B_3} \right)^p \geq \|P_{2j3+1}\|_p^p \\ & = \left\| \frac{1}{2} W_{j3} + \frac{1}{2} W_{j3+1} \right\|_p^p \\ & \geq \frac{1}{4} \|W_{j3+1}\|_p^p \\ & \geq \frac{1}{12} \sum_{i=0}^3 \|W_{j3+i}\|_p^p \\ & = \frac{1}{12} \left(\|\beta^{(j)}\|_p^{B_3} \right)^p. \end{aligned}$$

- Case 3: $A = \frac{1}{2} \|W_{j3+2}\|_p$.

We will estimate $\Gamma^{(2j+1)}$ in this case. This case is similar to Case 2. Then, we get

$$\begin{aligned} & \left(\|\Gamma^{(2j)}\|_p^{B_3} \right)^p + \left(\|\Gamma^{(2j+1)}\|_p^{B_3} \right)^p \\ & \geq \left(\|\Gamma^{(2j+1)}\|_p^{B_3} \right)^p \\ & \geq \frac{1}{12} \left(\|\beta^{(j)}\|_p^{B_3} \right)^p. \end{aligned}$$

- Case 4: $A = \|W_{j3+3}\|_p$.

We will estimate $\Gamma^{(2j+1)}$ in this case. This case is similar to Case 1. Then, we get

$$\begin{aligned} & \left(\|\Gamma^{(2j)}\|_p^{B_3} \right)^p + \left(\|\Gamma^{(2j+1)}\|_p^{B_3} \right)^p \\ & \geq \left(\|\Gamma^{(2j+1)}\|_p^{B_3} \right)^p \\ & \geq \frac{1}{6} \left(\|\beta^{(j)}\|_p^{B_3} \right)^p. \end{aligned}$$

From the results of the above four cases, we obtain

$$\begin{aligned} & \left(\|\Gamma^{(2j)}\|_p^{B_3} \right)^p + \left(\|\Gamma^{(2j+1)}\|_p^{B_3} \right)^p \\ & \geq \frac{1}{12} \left(\|\beta^{(j)}\|_p^{B_3} \right)^p, \forall j = 0, \dots, N-1. \end{aligned}$$

Thus

$$\begin{aligned} & \|\beta\|_p^{B_{2N,3}} \\ & = \frac{1}{(2N)^{1/p}} \left(\sum_{j=0}^{N-1} \left(\|\Gamma^{(2j)}\|_p^{B_3} \right)^p + \left(\|\Gamma^{(2j+1)}\|_p^{B_3} \right)^p \right)^{1/p} \\ & \geq \frac{1}{(2N)^{\frac{1}{p}}} \left(\sum_{j=0}^{N-1} \frac{1}{12} \left(\|\beta^{(j)}\|_p^{B_3} \right)^p \right)^{\frac{1}{p}} \\ & = \frac{1}{24^{\frac{1}{p}}} \|\beta\|_p^{B_{N,3}}. \end{aligned}$$

■

Lemma 3.3. Let $\beta \in B_{N,3}$, we get

$$\frac{1}{4} \|\beta\|_{\infty}^{B_{N,3}} \leq \|\beta\|_{\infty}^{B_{2N,3}}.$$

Proof. For any N -piece cubic Bézier curve associated with control points $W_{j3+i} \in \mathbb{R}^n, i = 0, \dots, 3, j = 0, \dots, N-1$. We have

$$\begin{aligned} \beta(t) &= \beta^{(j)}(Nt - j) = \sum_{i=0}^3 W_{j3+i} b_{i,3}(Nt - j) \\ & \text{if } t \in \left[\frac{j}{N}, \frac{j+1}{N} \right], j = 0, \dots, N-1. \end{aligned}$$

By (3), we split β to become a $2N$ -piece cubic Bézier curve as follows

$$\beta(t) = \begin{cases} \beta^{(2j)}(2Nt - 2j) = \sum_{i=0}^3 P_{2j3+i} b_{i,3}(2Nt - 2j) & \text{if } t \in \left[\frac{2j}{2N}, \frac{2j+1}{2N} \right], \\ \beta^{(2j+1)}(2Nt - 2j - 1) = \sum_{i=0}^3 P_{(2j+1)3+i} b_{i,3}(2Nt - 2j - 1) & \text{if } t \in \left[\frac{2j+1}{2N}, \frac{2j+2}{2N} \right], \\ & j = 0, \dots, N-1, \end{cases}$$

where

$$\begin{cases} P_{2j3+i} = \sum_{l=0}^i b_{l,i} \left(\frac{1}{2}\right) W_{j3+i-l}, \\ P_{(2j+1)3+i} = \sum_{l=0}^i b_{l,i} \left(\frac{1}{2}\right) W_{j3+3-i+l}, \\ i = 0, \dots, 3, j = 0, \dots, N-1. \end{cases}$$

First, we will consider

$$\max \left\{ \|\Gamma^{(2j)}\|_{\infty}^{B_3}, \|\Gamma^{(2j+1)}\|_{\infty}^{B_3} \right\}, j = 0, \dots, N-1.$$

Set

$$A =$$

$$\max \left\{ \|W_{j3}\|_{\infty}, \frac{1}{2} \|W_{j3+1}\|_{\infty}, \frac{1}{2} \|W_{j3+2}\|_{\infty}, \|W_{j3+3}\|_{\infty} \right\}$$

$$\bullet \text{ Case 1: } A = \|W_{j3}\|_{\infty}.$$

We have

$$\begin{aligned} \max \left\{ \|\Gamma^{(2j)}\|_{\infty}^{B_3}, \|\Gamma^{(2j+1)}\|_{\infty}^{B_3} \right\} &\geq \|\Gamma^{(2j)}\|_{\infty}^{B_3} \\ &\geq \|P_{2j3}\|_{\infty} = \|W_{j3}\|_{\infty} \\ &= \|\beta^{(j)}\|_{\infty}^{B_3}. \end{aligned}$$

$$\bullet \text{ Case 2: } A = \frac{1}{2} \|W_{j3+1}\|_{\infty}.$$

We have

$$\begin{aligned} \max \left\{ \|\Gamma^{(2j)}\|_{\infty}^{B_3}, \|\Gamma^{(2j+1)}\|_{\infty}^{B_3} \right\} &\geq \|\Gamma^{(2j)}\|_{\infty}^{B_3} \\ &\geq \|P_{2j3+1}\|_{\infty} \\ &= \left\| \frac{1}{2} W_{j3} + \frac{1}{2} W_{j3+1} \right\|_{\infty} \\ &\geq \frac{1}{4} \|W_{j3+1}\|_{\infty} = \frac{1}{4} \|\beta^{(j)}\|_{\infty}^{B_3}. \end{aligned}$$

$$\bullet \text{ Case 3: } A = \frac{1}{2} \|W_{j3+2}\|_{\infty}.$$

We will estimate $\Gamma^{(2j+1)}$ in this case. This case is similar to Case 2. Thus, we get

$$\begin{aligned} \max \left\{ \|\Gamma^{(2j)}\|_{\infty}^{B_3}, \|\Gamma^{(2j+1)}\|_{\infty}^{B_3} \right\} &\geq \|\Gamma^{(2j+1)}\|_{\infty}^{B_3} \\ &\geq \frac{1}{4} \|\beta^{(j)}\|_{\infty}^{B_3}. \end{aligned}$$

$$\bullet \text{ Case 4: } A = \|W_{j3+3}\|_{\infty}.$$

We will estimate $\Gamma^{(2j+1)}$ in this case. This case is similar to Case 1. Thus, we get

$$\begin{aligned} \max \left\{ \|\Gamma^{(2j)}\|_{\infty}^{B_3}, \|\Gamma^{(2j+1)}\|_{\infty}^{B_3} \right\} &\geq \|\Gamma^{(2j+1)}\|_{\infty}^{B_3} \\ &\geq \frac{1}{4} \|\beta^{(j)}\|_{\infty}^{B_3}. \end{aligned}$$

From the result of the above four cases, we obtain

$$\max \left\{ \|\Gamma^{(2j)}\|_{\infty}^{B_3}, \|\Gamma^{(2j+1)}\|_{\infty}^{B_3} \right\} \geq \frac{1}{4} \|\beta^{(j)}\|_{\infty}^{B_3}.$$

Thus

$$\begin{aligned} &\|\beta\|_{\infty}^{B_{2N,3}} \\ &= \max_{j=0, \dots, N-1} \max \left\{ \|\Gamma^{(2j)}\|_{\infty}^{B_3}, \|\Gamma^{(2j+1)}\|_{\infty}^{B_3} \right\} \\ &\geq \max_{j=0, \dots, N-1} \frac{1}{4} \|\beta^{(j)}\|_{\infty}^{B_3} = \frac{1}{4} \|\beta\|_{\infty}^{B_{N,3}}. \end{aligned}$$

■

By the above result, we obtain the following theorem.

Theorem 1.1. Let $p \in [1, \infty]$ and let $\beta \in B_{N,3}$ be an N -piece cubic Bézier curve. Then

$$\min \left\{ \frac{1}{24^{1/p}}, \frac{1}{4} \right\} \|\beta\|_p^{B_{N,3}} \leq \|\beta\|_p^{B_{2N,3}} \leq 3^{\frac{1}{p}} \|\beta\|_p^{B_{N,3}}.$$

Proof. Using Lemma 3.1, 3.2 and 3.3, we get the proof of Theorem 1. ■

By the above theorem, we obtain the corollary as follows:

$$\begin{aligned} \min \left\{ \frac{1}{24^{1/p}}, \frac{1}{4} \right\} d_p^{B_{N,3}}(\beta - \gamma) &\leq d_p^{B_{2N,3}}(\beta - \gamma) \\ &\leq 3^{\frac{1}{p}} d_p^{B_{N,3}}(\beta - \gamma), \end{aligned}$$

for any $\beta, \gamma \in B_{N,3}$.

4. CONCLUSION

This article presents the norm $\|\cdot\|_p^{B_{N,3}}$ of piecewise cubic Bézier curves which is defined by control points. This norm is more convenient to compute than the L_p norm. An N -piece cubic Bézier curve can be split and reparametrized to become a $2N$ -piece cubic Bézier curve. This way creates extra control points in order to give additional freedom for curve design and avoids increasing the degree of the curve. We also show the equivalence constants for the norm $\|\cdot\|_p^{B_{N,3}}$ and the norm $\|\cdot\|_p^{B_{2N,3}}$. These equivalence constants do not depend on the number of pieces. So, the norm $\|\cdot\|_p^{B_{N,3}}$ can be applied to check the convergence for sequences of piecewise cubic Bézier curves. This result is important for applying piecewise cubic Bézier curves to detect optimal orbits.

DECLARATION OF CONFLICTS OF INTEREST

The author(s) declare that this research has no conflicts of interest.

REFERENCES

- [1] D. Marsh, *Applied geometry for computer graphics and CAD*. London: Springer-Verlag, 2005.
- [2] J. Gallier, *Curves and surfaces in geometric modeling*. San Francisco, California: Morgan Kaufmann, 2000.
- [3] R. Hassin, "Approximation algorithms for hitting objects with straight lines," *Discrete Applied Mathematics*, vol. 30, no. 1, pp. 29-42, 1991.
- [4] C. M. Williams, "Bounded straight-line approximation of digitized planar curves and lines," *Computer Graphics and Image Processing*, vol. 16, no. 4, pp. 370-381, 1981.
- [5] S. Arthur, *Linear Approximation*. Providence, Rhode Island: American Mathematical Society, 1963.
- [6] L. N. Trefethen, *Approximation Theory and Approximation Practice, Extended Edition*. Philadelphia: Society for Industrial and Applied Mathematics, 2019.
- [7] E. Bodansky and A. Gribov, "Approximation of polylines with circular arcs," in *Lecture Notes in Computer Science*, 2004, pp. 193-198.
- [8] S. M. Thomas and Y. T. Chan, "A simple approach for the estimation of circular arc center and its radius," *Computer Vision, Graphics, and Image Processing*, vol. 45, pp. 362-370, 1989.
- [9] C. Ichoku, B. Deffontaines and J. Chorowicz, "Segmentation of Digital Plain Curves: A Dynamic Focusing Approach," *Pattern Recognition Letters, Elsevier BV*, vol. 17, pp. 741-750, 1996.
- [10] R. L. Burden and D. J. Faires. *Numerical Analysis*. Boston: Cengage Learning, 2010.
- [11] S. C. Chapra and R. P. Canale, *Numerical Methods for Engineers*. New York, McGraw-Hill Education, 2015.
- [12] T. Heister, L. G. Rebholz and F. Xue, *Numerical Analysis: An Introduction*. Berlin: De Gruyter, 2019.
- [13] J. Gallier, *Curves and surfaces in geometric modeling*. San Francisco, California: Morgan Kaufmann, 2000.
- [14] T. W. Sederberg, *Applications to computer aided geometric design*. London: American Mathematical Society, 1998.
- [15] J. E. McIntyre, *Guidance, flight mechanics and trajectory optimization*. Washington, D.C.: National Aeronautics and Space Administration, 1968.
- [16] F. Bullo, J. Cortés, S. Martínez, *Distributed control of robotic networks*. Princeton, New Jersey: Princeton University Press, 2009.
- [17] M. Farber, *Invitation to topological robotics*. Zurich: European Mathematical Society (EMS), 2008.
- [18] J. V. Neumann, *Collected works. Vol. VI: Theory of games, astrophysics, hydrodynamics and meteorology*. New York: The Macmillan Co., 1963.
- [19] K. J. Worsley, "Boundary corrections for the expected Euler characteristic of excursion sets of random fields, with an application to astrophysics," *Advances in Applied Probability*, vol. 159, no. 27, pp. 943-959, 1995.
- [20] M. Farber, "Topological complexity of motion planning," *Discrete and Computational Geometry*, vol. 29, no. 2, pp. 211-221, 2003.
- [21] M. Farber, S. Tabachnikov, S. Yuzvinsky, "Topological robotics: motion planning in projective spaces," *International Mathematics Research Notices*, vol. 2023, no. 34, pp. 1853-1870, 2003.
- [22] R. M. Murray, S. S. Sastry, "Nonholonomic motion planning: steering using sinusoids," *IEEE Transactions on Automatic Control*, vol. 38, no. 5, pp. 700-716, 1993.
- [23] J. Sokolowski, J. P. Zolésio, *Introduction to shape optimization*. Springer-Verlag, Berlin, 1992.

- [24] J. Haslinger, R. A. E. Makinen, *Introduction to shape optimization*. Philadelphia, Pennsylvania: Society for Industrial and Applied Mathematics (SIAM), 2003.
- [25] G. Allaire, *Shape optimization by the homogenization method*. New York: Springer-Verlag, 2002.
- [26] B. Mohammadi, O. Pironneau, *Applied shape optimization for fluids*, 2nd edition. Oxford: Oxford University Press, 2010.
- [27] F. D. Gournay, G. Allaire, F. Jouve, “Shape and topology optimization of the robust compliance via the level set method,” *ESAIM. Control, Optimisation and Calculus of Variations*, vol. 14, no. 1, pp. 43-70, 2008.
- [28] D. Bucur, G. Buttazzo, *Variational methods in some shape optimization problems*. Pisa: Scuola Normale Superiore, 2002.
- [29] H. V. Duc, “Equivalence constants for some norms on the space of N -piece cubic Bézier curves,” *Quy Nhon University Journal of Science*, vol. 17, no. 5, pp. 91-101, 2023.



© 2026 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Chia tách đường cong Bézier bậc ba từng mảnh và các hằng số tương đương cho một vài chuẩn

Hoàng Văn Đức*

Khoa Toán và Thống kê, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 12/11/2025; Ngày sửa bài: 05/03/2026;

Ngày nhận đăng: 06/03/2026; Ngày xuất bản: 28/06/2026

TÓM TẮT

Bài báo này liên quan đến việc chia tách các đường cong Bézier từng khúc bậc ba và các hằng số tương đương của một số chuẩn được định nghĩa bởi các điểm điều khiển. Các đường cong Bézier từng khúc bậc ba được sử dụng phổ biến nhất để xấp xỉ các đường cong. Chúng được hình thành từ các điểm điều khiển. Các chuẩn $\|\cdot\|_p^{B_{N,3}}$ và $\|\cdot\|_p^{B_{2N,3}}$ trong không gian $B_{N,3}$ được xác định bởi các điểm điều khiển. Với mục tiêu giữ bậc của đường cong và thêm tính linh hoạt trong việc thiết kế đường cong, chúng ta thường chia tách một đường cong Bézier N khúc bậc ba thành một đường cong Bézier $2N$ khúc bậc ba. Chúng ta sẽ tập trung vào các hằng số tương đương cho chuẩn $\|\cdot\|_p^{B_{N,3}}$ và $\|\cdot\|_p^{B_{2N,3}}$ trong không gian $B_{N,3}$ của các đường cong Bézier N khúc bậc ba. Do đó, chúng ta có thể sử dụng chuẩn $\|\cdot\|_p^{B_{N,3}}$ để kiểm tra sự hội tụ của chuỗi các đường cong Bézier từng khúc bậc ba. Kết quả này là quan trọng trong việc áp dụng đường cong Bézier từng khúc bậc ba để tìm ra quỹ đạo tối ưu.

Từ khóa: Đường cong Bézier bậc ba, hằng số tương đương, chuẩn, chia tách, từng khúc.

*Tác giả liên hệ chính.

Email: hoangvanduc@qnu.edu.vn

Design and development of a computer-vision-based robotic arm system for sorting of vehicle-like objects

Dinh Minh Le, Xuan Loc Nguyen, Van Can Do, Xuan Khoa Tran*

Faculty of Engineering and Technology, Quy Nhon University, Vietnam

Received: 24/12/2025; Revised: 03/03/2026;

Accepted: 24/03/2026; Published: 28/06/2026

ABSTRACT

The application of computer vision in vehicle-like object sorting systems contributes to improving productivity, accuracy, and reducing human-induced errors in industrial production lines. By exploiting visual features and combining image-processing techniques, deep learning enables reliable and automated recognition of complex objects under varying lighting conditions, viewing angles, and working distances. This paper presents the design and implementation of an automated sorting system integrating a robotic arm with a YOLO-based deep learning model for real-time detection and classification of three vehicle categories-cars, trucks, and buses-while extending classification capability through color recognition and QR code identification. The system supports three recognition modes: YOLO combined with color detection, QR code-based recognition, and a hybrid approach. Recognition results are transmitted to a Siemens S7-1200 PLC to control the robotic arm, while monitoring and operation are performed via a SCADA interface. Experimental results demonstrate that the proposed system operates stably in real time and achieves high classification accuracy under different working conditions, confirming the feasibility and effectiveness of integrating deep learning-based computer vision with PLC control for industrial automation systems.

Keywords: *Computer vision, YOLO, QR code, robotic arm, SCADA.*

1. INTRODUCTION

In the context of the rapidly advancing Industry 4.0 revolution, intelligent automation systems have become essential for enhancing productivity, reducing operational errors, and optimizing resource utilization [1]. One of the most prominent technological trends is the integration of industrial robots with computer vision, enabling machines not only to perform mechanical actions but also to “see” and make autonomous decisions based on real-time visual data [2].

In product-sorting applications, especially for vehicle-like objects such as model cars,

trucks, and buses, accurate classification becomes challenging due to varying illumination, background noise, inconsistent object sizes, and dynamic conveyor movement [3].

Among deep-learning-based object detectors, the YOLO (You Only Look Once) family has gained significant attention due to its ability to perform real-time object detection while maintaining competitive accuracy [4]. However, many existing studies rely solely on object shape recognition and do not incorporate additional attributes such as color classification or QR code decoding. Furthermore, only a few works have addressed the integration of deep-learning vision systems with industrial PLCs [5].

*Corresponding author:

Email: tranxuankhoa@qnu.edu.vn

To address these limitations, this paper presents the design and implementation of an automatic sorting system that integrates a robotic arm, a YOLO-based computer vision module, and a Siemens S7-1200 PLC. The proposed system detects and classifies three vehicle categories-cars, trucks, and buses-and further analyzes their colors using HSV color-space processing. In addition, a QR code recognition module is incorporated to enhance flexibility. The overall architecture is designed to ensure real-time operation and high classification accuracy. Moreover, the system's modular design allows easy scalability and adaptation to different industrial environments, supporting future integration of additional sensors or sorting criteria.

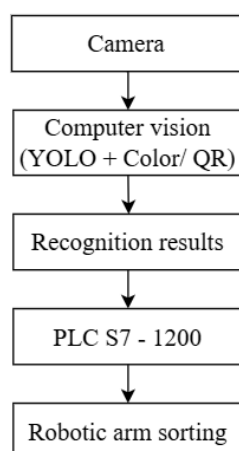


Figure 1. Overview system.

Unlike conventional vision-based recognition approaches that focus primarily on algorithmic improvements, this work emphasizes a system-level contribution. The proposed solution integrates a hybrid YOLO-QR recognition architecture with a priority-based decision logic and real-time PLC-SCADA integration, targeting practical deployment in industrial environments. The system supports three recognition modes-YOLO with color detection, QR-based recognition, and a hybrid mode-allowing adaptive operation under different operational scenarios. Recognition results are transmitted to the PLC for real-time

robotic arm control, while a SCADA interface enables continuous monitoring and supervision of system behavior.

It should be emphasized that this study does not aim to propose new detection algorithms or modify the internal structure of the YOLO model. Instead, the primary technical contributions of this work lie at the system level and can be summarized as follows:

1) A hybrid computer-vision architecture that combines YOLO-based visual detection and QR-code-based identification for real-time industrial sorting applications;

2) A priority-based recognition strategy in which QR decoding is treated as the primary identification method, while YOLO detection is automatically activated as a fallback mechanism when QR labels are missing, damaged, or unreadable, thereby improving robustness without increasing computational complexity;

3) A complete real-time integration framework connecting the vision system, a Siemens S7-1200 PLC, a robotic arm, and a SCADA interface, enabling closed-loop control and synchronized operation between perception, decision-making, and actuation;

4) An experimental system-level evaluation demonstrating consistent and stable operation under varying lighting conditions and continuous conveyor motion in real industrial operating environments.

These contributions position the proposed work as a practical and scalable system-level solution for intelligent industrial sorting rather than an algorithm-level enhancement.

The remainder of this paper is organized as follows: Section 2 reviews related works, Section 3 presents the system design and methodology, Section 4 presents the experimental results, Section 5 provides discussion, and Section 6 provides conclusion.

2. RELATED WORKS

In recent years, the integration of computer vision and deep learning technologies into industrial automation has received significant attention from researchers and engineers. Traditional sensing-based classification methods, such as RGB color sensors, proximity sensors, and photoelectric detectors, are generally suitable only for simple objects and stable operating conditions. However, these methods often fail when facing challenges such as variable illumination, complex backgrounds, or objects with diverse shapes and colors, leading to reduced accuracy and system instability [6]. These limitations have motivated the use of image-processing and deep-learning approaches to enhance recognition performance in real-world environments [7].

In object detection, the YOLO (You Only Look Once) family has emerged as one of the most widely adopted deep-learning models due to its ability to perform fast and accurate real-time detection. Models such as YOLOv3, YOLOv4, YOLOv5, and the more recent YOLOv8 have been applied to various industrial tasks, including defect detection [8], agricultural sorting, traffic monitoring, and vehicle recognition. Existing studies show that YOLO-based systems can achieve more than 30 FPS while maintaining high detection accuracy even when objects are partially occluded or captured under low-light conditions [9]. Some research also integrates YOLO with color-space analysis to extend classification capabilities when objects share similar shapes but differ in color.

Additionally, QR code recognition has been increasingly used in manufacturing due to its high storage capacity, robustness to noise, and ease of integration with modern vision systems. Libraries such as ZBar and Pyzbar enable real-time decoding and are widely used in inventory management, traceability systems, and automatic product routing. The combination of visual recognition and QR decoding enhances

system flexibility by allowing classification based not only on visual appearance but also on encoded product information.

In industrial control, Programmable Logic Controllers (PLCs) play a critical role in bridging computer vision modules with hardware equipment such as robotic arms and conveyor belts. Several studies have explored integrating computer vision with PLCs-such as Siemens or Mitsubishi-through communication protocols like Modbus/TCP, OPC-UA, or Snap7. These works highlight the importance of reliable and low-latency communication in practical automation systems [10]. However, many existing studies focus on isolated components, such as visual detection alone or robotic control alone, without developing a fully integrated system that combines deep learning, multi-modal recognition, and PLC-based control.

From the reviewed literature, it is evident that a gap remains in designing a comprehensive solution that integrates:

- (1) YOLO-based object detection
- (2) Color recognition
- (3) QR code decoding
- (4) Real-time communication with a PLC to perform robotic sorting.

The present paper aims to address this gap by developing a unified, practical product-sorting system suitable for small and medium-sized automation environments.

3. SYSTEM DESIGN AND METHODOLOGY

It should be emphasized that the novelty of the proposed approach does not stem from individual vision algorithms, but from the way multiple recognition modalities are organized and coordinated within a unified real-time system architecture. The design focuses on reliability, priority-based decision logic, and seamless integration between perception and industrial control.

3.1. System architecture

The proposed system is designed as a unified real-time industrial sorting framework that integrates perception, decision-making, and actuation within a closed-loop control architecture. The system combines computer vision, industrial control, and robotic manipulation to enable reliable vehicle classification and sorting under dynamic conveyor conditions. A hybrid recognition strategy is employed to enhance robustness by leveraging complementary identification modalities. Real-time communication between system components ensures synchronized operation and continuous monitoring through a supervisory interface. The overall system architecture is illustrated in Figure 2. This integrated design facilitates stable, efficient, and scalable deployment in practical industrial environments.

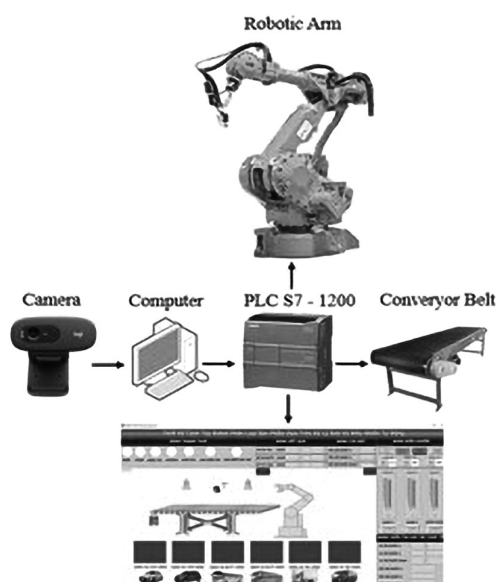


Figure 2. System architecture.

The system, as shown in Figure 2, comprises a camera, a computer running the vision modules, a Siemens PLC S7-1200, a robotic arm, a conveyor belt, and an SCADA interface. The camera captures images of vehicles moving on the conveyor belt and transmits them to the computer for processing.

The vision system adopts a hybrid YOLO–QR recognition architecture with a priority-based

decision strategy. The system first attempts to decode the QR code attached to the vehicle due to its high reliability and direct data encoding capability. When QR decoding is successful, the extracted identification information is immediately forwarded to the PLC for real-time control execution.

In cases where the QR code is missing, damaged, or unreadable, the system automatically activates the YOLOv8-based visual recognition module as a fallback mechanism to identify the vehicle type. After vehicle identification, the color information is determined using the HSV color space. This conditional hybrid strategy ensures continuous system operation without manual intervention while maintaining robustness under real-world industrial conditions.

Once the classification results are obtained, the recognized vehicle type and color information are transmitted to the PLC via the Snap7 communication protocol. Based on the received data, the PLC controls the robotic arm to sort vehicles into designated locations on the conveyor system.

The PLC functions as the central real-time controller, coordinating data exchange between the vision system, robotic arm, and SCADA interface. The SCADA system enables real-time monitoring, data visualization, and operator interaction, forming a closed-loop industrial control architecture. This integrated PLC–SCADA framework ensures synchronized operation across all system components, thereby enhancing reliability and practical deployability of the proposed system.

3.2. Image acquisition and processing

The image acquisition and processing module serves as the foundation of the entire vision pipeline, ensuring stable input data and enabling downstream recognition algorithms to achieve optimal performance. A camera mounted above the conveyor continuously captures frames of objects as they move through the inspection

area. These frames are immediately transferred to the vision-processing program running on the computer, where they are handled sequentially with minimal latency to maintain synchronization between the conveyor motion and the robotic arm's pick-and-place cycle.

To ensure reliable operation, the camera position and capture parameters are carefully configured to minimize motion blur and illumination variations. Continuous frame acquisition allows the system to maintain temporal consistency between consecutive observations, which is essential for stable real-time control. In addition, buffering and frame-handling mechanisms are employed to prevent data loss during high-throughput operation. This design choice contributes to maintaining deterministic system behavior and smooth interaction between perception and actuation components. Overall, the image acquisition stage establishes a reliable data foundation for subsequent recognition and decision-making processes.

Upon receiving a frame, the system first performs a preprocessing step by converting the image from the BGR format into the HSV color space. HSV is selected due to its strong separation between chromatic and luminance components,¹³ providing robust color recognition under varying lighting conditions [11]. After conversion, the frame is routed into one of three processing pipelines depending on the selected operating mode.

In the YOLO-based detection mode, the frame is fed into the YOLOv8 model to identify predefined vehicle categories [12], including cars, trucks, and buses. The detector outputs bounding boxes, class labels, and confidence scores. The bounding box also provides a precise region of interest for subsequent color analysis, effectively reducing the influence of background noise and irrelevant pixels.

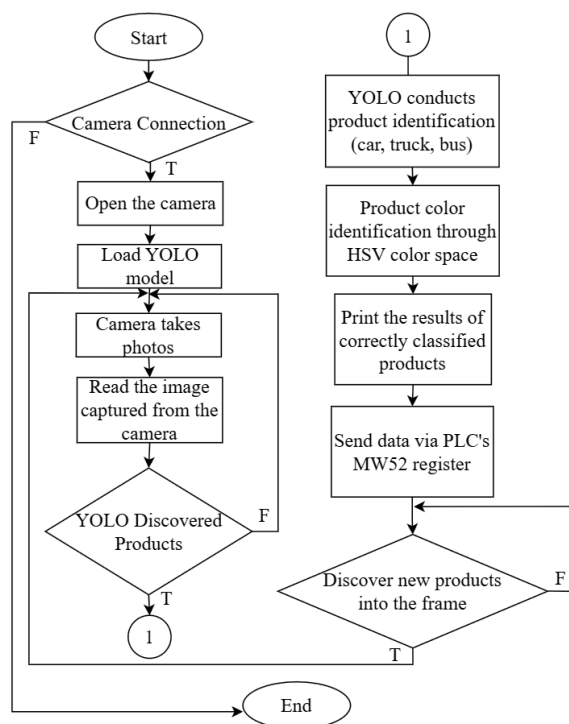


Figure 3. Flowchart of image processing algorithm using YOLO.

In the QR-recognition mode, the system utilizes the Pyzbar library to perform comprehensive scanning and decoding of QR patterns within each processed frame [13]. Pyzbar is selected due to its stable recognition capability even under challenging imaging conditions such as drastic lighting variations, partial occlusions, motion blur, or low-resolution inputs. During operation, the system analyzes the shape and position of regions suspected to contain QR codes, verifies the structural validity of the detected patterns, and then applies Pyzbar's decoding algorithm to extract the embedded data [14]. Once a valid QR code is identified, the information encoded within it is immediately adopted as the final classification output. This mechanism allows the system to bypass the YOLO-based processing steps, thereby reducing latency, improving processing speed, and enhancing overall accuracy. Thanks to these advantages, the QR-recognition mode is particularly suitable for automated production lines, inventory inspection, warehouse management, traceability systems,

and applications that require encoded data with higher reliability compared to conventional visual recognition. Moreover, this mode contributes to consistent real-time performance while maintaining system stability.

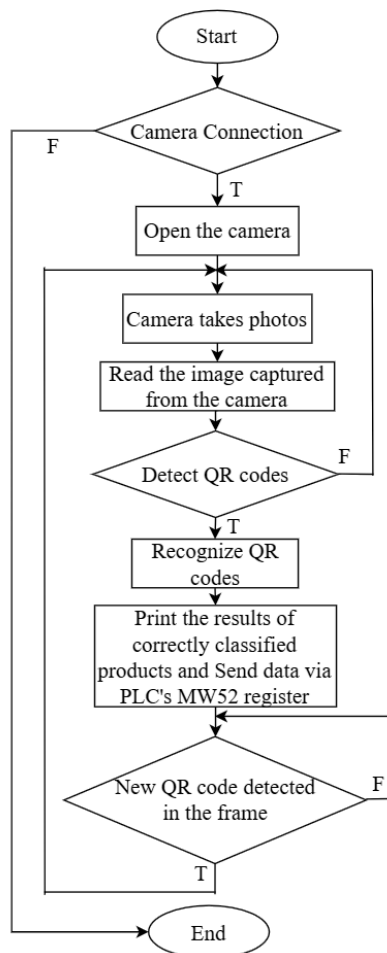


Figure 4. Flowchart of image processing algorithm using QR code.

When YOLO successfully detects an object, the color-recognition module extracts the central region of the bounding box to reduce the effects of shadows, object edges, and background gradients. The selected pixels are converted into HSV values and compared with predefined threshold ranges, while a voting-based algorithm determines the dominant color with greater stability. To prevent misclassification caused by motion blur or temporary lighting variations, a temporal consistency filter ensures that recognition

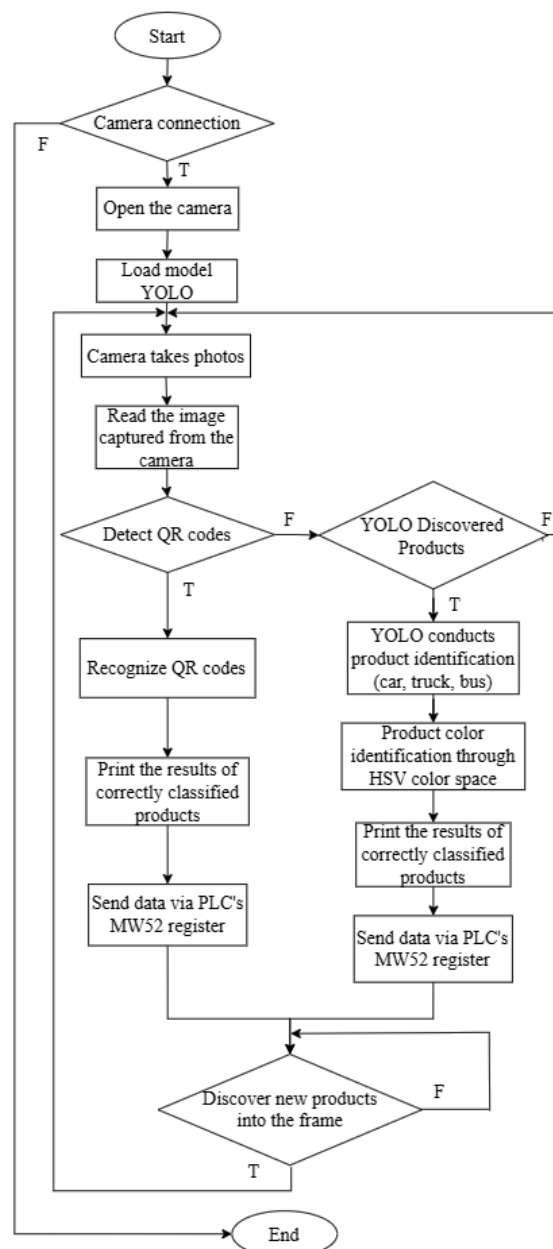


Figure 5. Flowchart of image processing algorithm using YOLO and QR code.

results remain stable across multiple consecutive frames before being accepted. This combination of spatial filtering and temporal validation significantly improves the reliability of the final color output and ensures robust recognition under fluctuating environmental conditions. In addition, the system applies a normalization step to mitigate brightness fluctuations, helping the HSV values remain consistent during operation. The processing pipeline is further optimized to minimize latency, ensuring that color classification

is synchronized with the motion of objects on the conveyor. By integrating these enhancements, the color-recognition subsystem achieves higher resilience against noise, reflections, and rapid illumination changes.

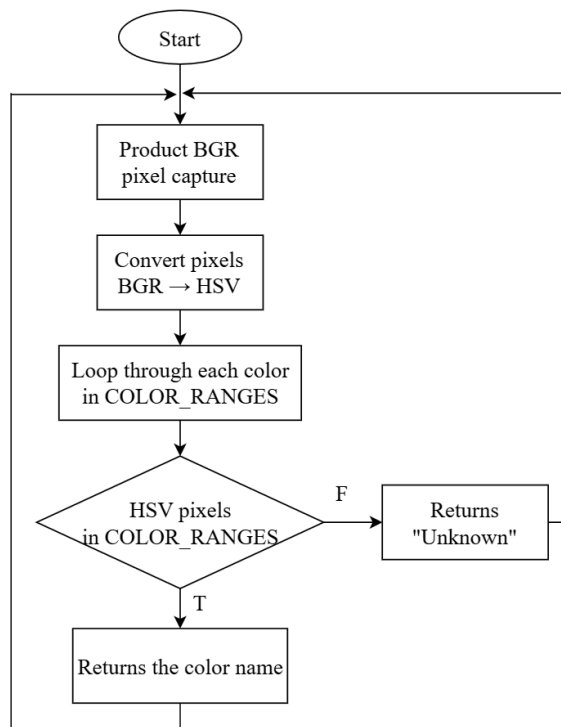


Figure 6. Flowchart of color recognition algorithm using HSV color space.

After the object type and color are confirmed, the final classification result is packaged and transmitted to the Siemens S7-1200 PLC via the Snap7 protocol [15]. The PLC interprets the received data and triggers the corresponding robotic actions according to the predefined workflow. This streamlined communication and control pipeline enhances responsiveness, operational robustness, and fault tolerance. As a result, the system maintains stable performance in industrial automation environments, meeting strict requirements for accuracy, speed, and continuous operation. Furthermore, the PLC continuously monitors the communication status and the operational condition of the robotic arm to prevent unexpected failures. In the event of abnormal data or communication delays, the system automatically activates predefined safety responses, ensuring uninterrupted workflow

and reliable high-throughput operation. This design approach supports long-term autonomous operation while minimizing downtime and maintenance overhead.

3.3. YOLOv8 model training

The YOLOv8 model used in this study was trained using a custom dataset collected directly from the experimental setup. The dataset consists of vehicle-like objects belonging to three categories: car, truck, and bus. All images were captured using a high-resolution camera under adequate lighting conditions, resulting in clear and sharp image quality.

A total of approximately 300 images were collected, including variations in object orientation, position, color, and illumination conditions. The dataset was designed to represent practical industrial scenarios where objects move continuously on the conveyor belt and may appear at different viewing angles and distances. This diversity helps improve the generalization capability of the trained model.



Figure 7. Sample images from the training dataset.

All images were manually annotated using bounding boxes to label the object categories. The annotation process was performed using standard labeling tools, ensuring consistent and accurate ground-truth data for training and evaluation. To improve robustness and reduce overfitting, data augmentation techniques were applied during training. These techniques included image rotation, brightness variation, scaling, and horizontal flipping. Data augmentation increased the effective dataset diversity and improved detection stability under varying environmental conditions.

The YOLOv8 model was selected due to its favorable balance between detection accuracy and computational efficiency for real-time industrial applications. Training was performed for 100 epochs with a batch size of 16 and an input image resolution of 640×640 pixels. The training process was executed on a computer equipped with an NVIDIA RTX 3050 GPU, ensuring efficient convergence and stable model performance.

After training, the best-performing model was selected based on validation accuracy and used in the experimental evaluation presented in Section 4.

3.4. PLC integration and control logic

Once the objects have been fully identified and recognized, the resulting classification data are transmitted to the Siemens S7-1200 PLC through the Snap7 communication protocol. This protocol ensures fast, stable, and reliable data transfer between the vision-processing unit and the control system, even under continuous high-frequency operation [16]. Upon receiving the data, the PLC analyzes the classification information and determines the corresponding control commands that the robotic arm must execute according to the predefined sorting logic. Each classification type is mapped to a specific container, storage cell, or designated physical location on the conveyor system, ensuring that every object is delivered to the correct destination with high accuracy. Moreover, the system can dynamically adjust sorting parameters in real-time based on varying production rates or unexpected changes in object flow, further enhancing operational efficiency.

The control logic of the system is developed in Siemens TIA Portal using the ladder diagram programming language, which provides a clear and intuitive structure for handling sequential control operations [17]. Within this environment, additional safety checks, fault-handling routines, and motion-coordination mechanisms are implemented to guarantee stable and collision-free operation of the robotic arm. The machine-vision module communicates with the PLC through an optimized data-exchange

routine, enabling minimal latency and ensuring that classification results are synchronized with the conveyor's movement and the robotic arm's pick-and-place cycle. In addition, data logging and automated alerts are integrated to notify operators of abnormal conditions immediately, allowing preventive measures to be taken without halting the production line.

To support supervision and operator interaction, a SCADA interface is integrated into the control architecture. This interface provides real-time visualization of the entire sorting workflow, including the detected object type, current operating mode, communication status, and the live position of the robotic arm [17]. Operators can monitor system performance, adjust parameters, and intervene when necessary, thereby improving operational transparency and enabling efficient system troubleshooting. The combination of PLC-based control and SCADA-based supervision ensures a highly reliable, safe, and easy-to-manage automation system. Additionally, the SCADA interface records

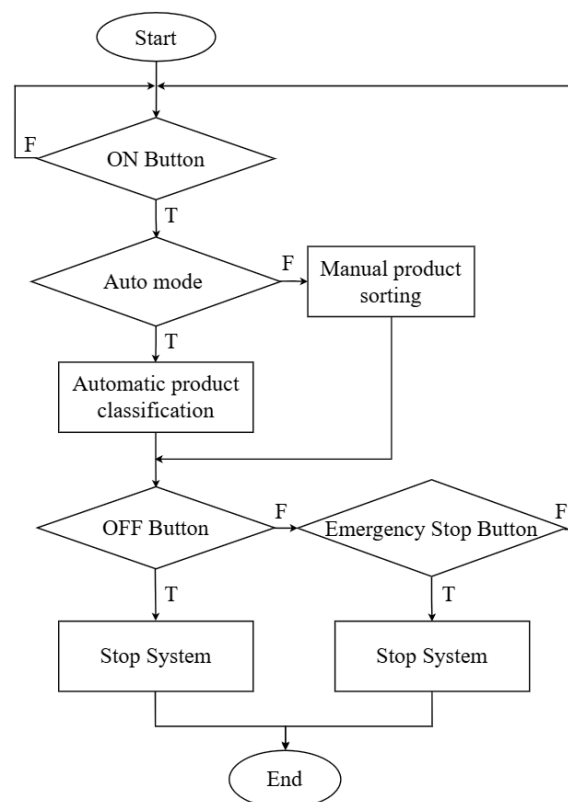


Figure 8. General product classification algorithm flowchart.

historical operational data, allowing engineers to review trends and diagnose recurring issues more effectively. These features collectively enhance the overall maintainability of the system and contribute to long-term operational stability. Additionally, the SCADA interface

records historical operational data, allowing review trends, analyze throughput efficiency, and diagnose recurring issues more effectively. The system also supports predictive maintenance by generating performance reports and forecasting potential equipment failures before they occur.

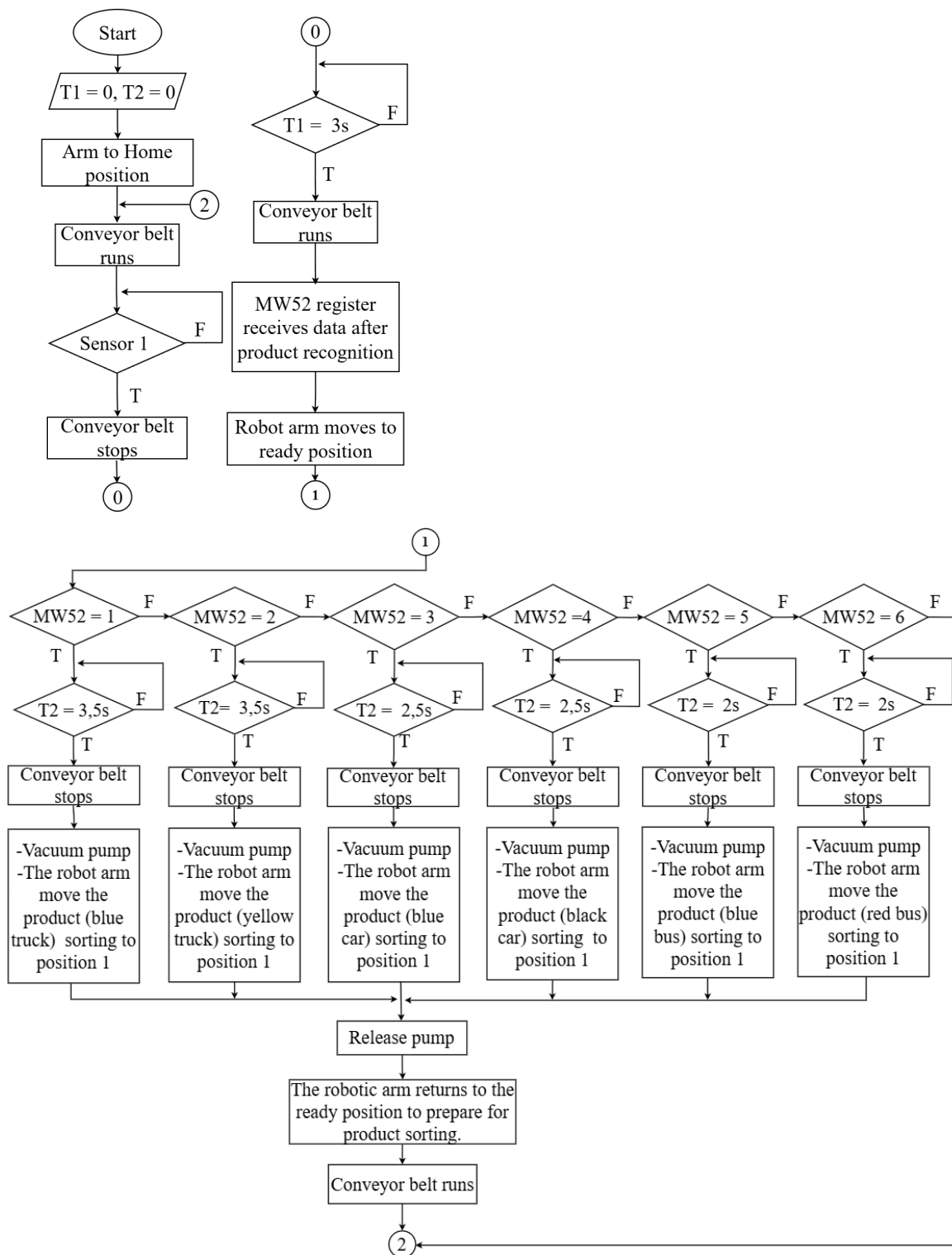


Figure 9. Flowchart of product classification algorithm in Auto mode.

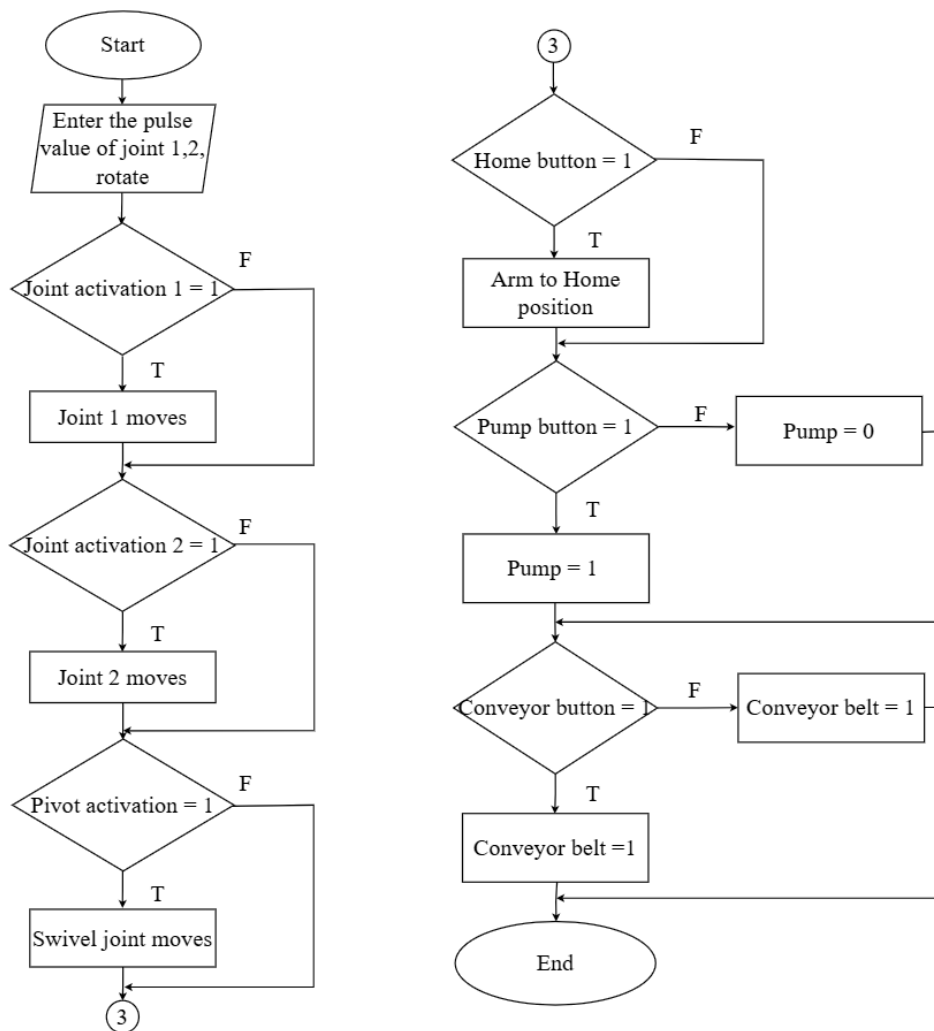


Figure 10. Flowchart of product classification algorithm in Manual mode.

3.5. Formulas

3.4.1. YOLO-Based Image Processing Formulations

In the proposed system, YOLOv8 is employed as the object detection module due to its proven balance between detection accuracy and real-time processing capability [18]. YOLO follows a single-stage detection paradigm, in which object localization and classification are performed simultaneously within a single forward pass of the network. This design makes it particularly suitable for real-time industrial applications where low latency is required [19].

In practice, YOLOv8 processes each input frame captured from the conveyor-mounted camera and outputs bounding boxes, class labels, and confidence scores for detected vehicle-like

objects. These outputs provide reliable spatial information that is subsequently used by higher-level system logic, including color recognition, QR prioritization, and PLC-based decision making.

It is important to emphasize that this work does not aim to modify or improve the internal architecture or loss formulation of YOLOv8. Instead, the model is adopted as a mature and stable detection component within a larger system-level framework focused on robustness and real-time integration.

3.4.2. HSV Color Space–Based Image Processing Formulations

For color recognition, the system adopts the HSV (Hue–Saturation–Value) color space, which is

widely used in industrial vision applications due to its ability to separate chromatic information from illumination intensity. Compared with the RGB representation, HSV provides improved robustness when lighting conditions vary, which is common in conveyor-based environments [20].

After an object is detected by the YOLO module, the corresponding region of interest is extracted and converted into the HSV color space. Color classification is then performed by comparing pixel values within this region against predefined threshold ranges for each target color. A voting-based mechanism is applied to determine the dominant color, reducing the influence of noise, shadows, and minor illumination fluctuations [21].

The HSV-based approach is computationally efficient and well suited for real-time operation [22]. In this study, it serves as a complementary module that extends object classification by incorporating color attributes without increasing system complexity.

3.4.3. Kinematic Equations of the Robotic Arm

The forward kinematic model is introduced to establish a mathematical relationship between joint variables and the end-effector pose, which is essential for defining reachable pick-and-place positions in the sorting workspace. Although real-time trajectory generation is executed by the PLC using predefined motion commands, the kinematic formulation provides the theoretical foundation for workspace analysis, joint limit verification, and collision-free motion planning during system design and validation.

The robotic manipulator under consideration is a three-link revolute configuration (R–R–R), as illustrated in Figure 11. The coordinate frames are assigned following the Denavit–Hartenberg (D–H) convention [23], and the corresponding parameters are summarized in Table 1. This formulation enables a systematic derivation of the forward kinematic model describing the end-effector pose as a function of joint variables [24].

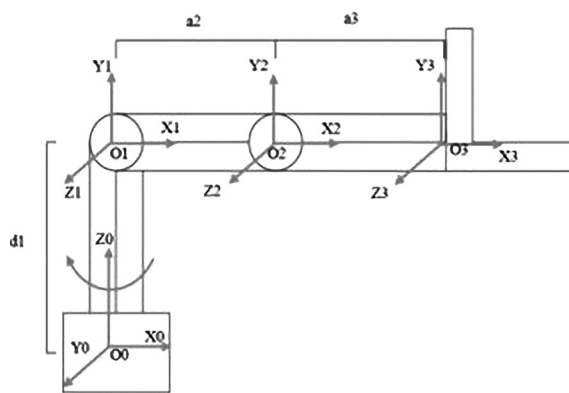


Figure 11. 3-DOF R-R-R robotic arm.

Using the standard Denavit–Hartenberg (D–H) convention, each joint of the robotic manipulator is described by a homogeneous transformation matrix A_i , which relates coordinate frame i to frame $i - 1$. The general form of the D–H transformation matrix is expressed as [25]:

$$A_i = \begin{bmatrix} c_{\theta_i} & -s_{\theta_i} \cdot c_{\alpha_i} & s_{\theta_i} \cdot s_{\alpha_i} & a_i \cdot c_{\theta_i} \\ s_{\theta_i} & c_{\theta_i} \cdot c_{\alpha_i} & -c_{\theta_i} \cdot s_{\alpha_i} & a_i \cdot s_{\theta_i} \\ 0 & s_{\alpha_i} & c_{\alpha_i} & d_i \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (1)$$

where: c is $\cos$ and s is $\sin$, θ_i the joint angle, a_i is the link length, d_i is the link offset, and α_i is the link twist angle.

Based on the mechanical structure of the proposed 3-DOF robotic arm, the corresponding D–H parameters are summarized in Table 1.

Table 1. Denavit–Hartenberg kinematic parameters.

Joint	θ_i	α_i	a_i	d_i
1	θ_1	90°	0	L_1
2	θ_2	0°	L_2	0
3	θ_3	0°	L_3	0

Using these parameters, the individual transformation matrices A_1 , A_2 , and A_3 are obtained. The overall forward kinematics of the manipulator is computed by multiplying the transformations sequentially:

$$T_3 = A_1 \cdot A_2 \cdot A_3 \quad (2)$$

where:

$$A_1 = \begin{bmatrix} c_{\theta_1} & 0 & s_{\theta_1} & 0 \\ s_{\theta_1} & 0 & -c_{\theta_1} & 0 \\ 0 & 1 & 0 & L_1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3a)$$

$$A_2 = \begin{bmatrix} c_{\theta_2} & -s_{\theta_2} & 0 & L_2 c_{\theta_2} \\ s_{\theta_2} & c_{\theta_2} & 0 & L_2 s_{\theta_2} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3b)$$

$$A_3 = \begin{bmatrix} c_{\theta_3} & -s_{\theta_3} & 0 & L_3 c_{\theta_3} \\ s_{\theta_3} & c_{\theta_3} & 0 & L_3 s_{\theta_3} \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (3c)$$

After substituting the D–H parameters and simplifying, the resulting homogeneous transformation matrix T_3 can be expressed in the following form:

$$T_3 = \begin{bmatrix} n_x & O_x & a_x & P_x \\ n_y & O_y & a_y & P_y \\ n_z & O_z & a_z & P_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (4)$$

The position of the end-effector with respect to the base frame is given by:

$$Px = c_{\theta_1} \cdot (L_2 \cdot c_{\theta_1+\theta_2} + L_2 \cdot c_{\theta_2}) \quad (5a)$$

$$Py = s_{\theta_1} \cdot (L_3 \cdot c_{\theta_2+\theta_3} + L_2 \cdot c_{\theta_2}) \quad (5b)$$

$$Pz = L_1 + L_3 \cdot s_{\theta_2+\theta_3} + L_2 \cdot s_{\theta_2} \quad (5c)$$

These equations explicitly define the Cartesian position of the end-effector as a function of the joint angles and link dimensions.

The orientation of the end-effector is represented by the rotation submatrix of T_3 , whose column vectors n , O , and a describe the directions of the local coordinate axes attached to the end-effector. These components are defined as:

$$n_x = c_{\theta_2+\theta_3} \cdot c_{\theta_1} \quad (6a)$$

$$n_y = c_{\theta_2+\theta_3} \cdot s_{\theta_1} \quad (6b)$$

$$n_z = s_{\theta_2+\theta_3} \cdot n_y = c_{\theta_2+\theta_3} \cdot s_{\theta_1} \quad (6c)$$

$$O_x = -s_{\theta_2+\theta_3} \cdot c_{\theta_1} \quad (7a)$$

$$O_y = -s_{\theta_2+\theta_3} \cdot s_{\theta_1} \quad (7b)$$

$$O_z = c_{\theta_2+\theta_3} \quad (7c)$$

$$ax = s_{\theta_1} \quad (8a)$$

$$ay = -c_{\theta_1} \quad (8b)$$

$$az = 0 \quad (8c)$$

The derived forward kinematic equations provide a direct mapping between joint space and Cartesian space. In the proposed system, the computed end-effector position (P_x , P_y , P_z) is used for trajectory planning and object positioning during pick-and-place operations, while the orientation components ensure consistent alignment of the gripper when handling vehicle-like objects.

The above expressions describe the motion of the robotic manipulator through the forward kinematic equations, using the D–H parameters derived from the joints of the system.

To ensure safe and collision-free motion during pick-and-place tasks, each joint is restricted to a predefined angular range. These constraints are implemented in the PLC control logic to prevent the manipulator from exceeding its mechanical limits.

Joint limits:

- Joint 1: $\theta_1 \in [0^\circ, 360^\circ]$;
- Joint 2: $\theta_2 \in [0^\circ, 78^\circ]$;
- Joint 3: $\theta_3 \in [0^\circ, 102^\circ]$.

3.6. Hardware setup

The hardware system consists of five main components: an image acquisition camera, a YOLOv8 processing computer, a Siemens S7-1200 PLC, a robotic arm, and a conveyor belt responsible for vehicle-like objects sorting. Figure 12 illustrates the overall layout of the experimental setup.

A Logitech C270 camera is mounted above the conveyor to capture input images for vehicle recognition, color classification, and QR code detection. The captured frames are then process on the computer to ensure real-time inference performance. The recognition results are transmitted to the Siemens S7-1200 PLC via the Snap7 communication protocol.

The PLC processes the signals and controls the robotic arm to pick and sort the products according to the detected results. This hardware configuration ensures stable operation and meets the requirements of the experimental tests.

In addition, the system is designed to maintain robust synchronization between visual processing and mechanical motion, minimizing latency during sorting. All hardware components are connected through a structured wiring system to enhance reliability and reduce electromagnetic interference. The robotic arm operates within a predefined workspace optimized for the conveyor layout, ensuring smooth and collision-free movements. Furthermore, the modular structure of the hardware setup allows for easy expansion and modifications for future experiments or industrial applications.

To support continuous real-time operation, the processing computer is configured to handle concurrent image acquisition and inference tasks without interrupting PLC communication. The communication latency between hardware components is kept minimal to preserve synchronization during high-speed conveyor movement. This coordinated hardware design enables reliable closed-loop control throughout the sorting process.

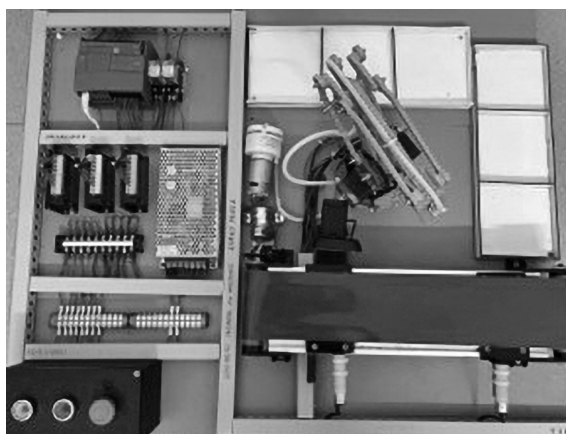


Figure 12. Experimental model.

4. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed system, a series of experiments were conducted

on a product classification model consisting of three vehicle types (car, truck, and bus), multiple color variations, and corresponding QR codes. The experiments were carried out under realistic operating conditions, with the conveyor running continuously, the robotic arm performing pick-and-place actions, and a Logitech C270 camera fixed above the observation area. The vision processing module was executed on a computer equipped with an RTX 3050 GPU, ensuring real-time inference throughout the experiments.

4.1. Accuracy of YOLOv8 and color recognition

To evaluate the performance of the trained YOLOv8s model [8] in vehicle classification, the system was tested on a dataset consisting of three object categories: car, truck, and bus. The YOLOv8s model was trained using the dataset described in Section 3.3 and then validated under realistic operating conditions. Each category included multiple samples with different sizes, shapes, and colors to ensure diversity in the input data. The recognition results were collected under realistic operating conditions with varying ambient lighting and continuously moving objects on the conveyor belt.

Experimental results show that the trained YOLOv8s model achieves high accuracy in identifying the three vehicle types, with an overall accuracy of approximately 94%. The detected vehicles exhibit relatively consistent accuracy across all categories, indicating that the trained model generalizes well to practical operating conditions.

In addition to vehicle recognition, the HSV-based color analysis algorithm was evaluated using a dataset containing multiple color samples. The experiments indicate that the method provides stable performance, particularly when ambient lighting does not change rapidly. Examples illustrating the vehicle detection and color classification processes are presented in Figure 13.

Furthermore, the combined use of the trained YOLOv8s model and HSV color analysis enables the system to classify both the type and color of vehicles simultaneously, providing a comprehensive solution for automated sorting. This integration also allows real-time monitoring and evaluation, which is crucial for practical industrial applications. Additionally, the system can adapt to minor variations in vehicle orientation and position, maintaining high accuracy under typical conveyor conditions. The results demonstrate the robustness of the proposed approach, making it suitable for continuous operation in real-world manufacturing environments.

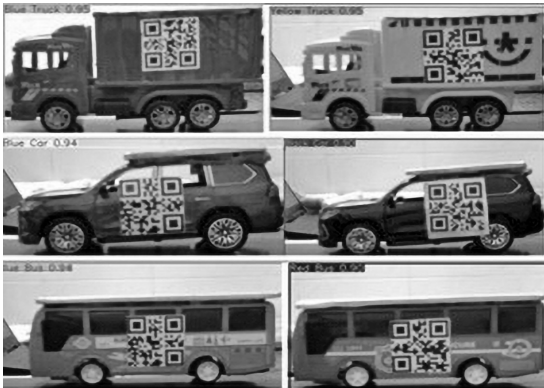


Figure 13. Vehicle and color recognition using YOLO and HSV.

To evaluate the recognition accuracy, the system was tested by performing 10 consecutive recognition trials for each vehicle type with its corresponding colors. Table 2 presents the recorded values for all 10 measurements.

Table 2. Short-term repeatability results over 10 recognition trials for each vehicle type and color.

Vehicle Type	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	Mean (μ) $\pm$ Std
Blue Truck	95.5%	94.2%	94.2%	95.8%	95.5%	95.3%	95.5%	95.0%	95.3%	95.3%	95.16 $\pm$ 0.55 %
Yellow Truck	95.8%	95.1%	94.2%	94.0%	95.6%	94.0%	95.1%	95.7%	95.7%	95.2%	95.04 $\pm$ 0.72 %
Blue Car	96.7%	97.2%	96.1%	97.0%	96.4%	97.1%	96.5%	97.0%	97.2%	96.9%	96.81 $\pm$ 0.37 %
Black Car	97.0%	98.1%	97.2%	97.0%	98.0%	98.2%	97.5%	98.0%	97.2%	97.3%	97.55 $\pm$ 0.48 %
Blue Bus	96.9%	96.3%	95.9%	96.7%	95.7%	96.5%	96.1%	96.2%	96.0%	96.0%	96.23 $\pm$ 0.37 %
Red Bus	96.2%	96.5%	96.1%	95.8%	96.3%	96.1%	96.0%	96.2%	95.5%	96.0%	96.07 $\pm$ 0.28 %

Each experiment was repeated ten times under identical operating conditions to evaluate short-term repeatability and operational consistency of the proposed system. The reported mean and standard deviation values describe variability across repeated trials and are not intended for large-scale statistical inference.

The accuracy of vehicle detection and color classification was evaluated over ten repeated recognition trials for each vehicle type and color. The results are illustrated in a line chart (Figure 14), which presents the

accuracy obtained in each trial. As observed, the recognition accuracy consistently exceeds 92% across all tested scenarios.

In addition to the mean accuracy, the standard deviation remains low, typically below $\pm 1.0\%$, indicating that the observed variations between trials are minimal. Such a small deviation suggests that the system’s performance is highly stable and that the reported accuracy values are experimentally consistent rather than incidental.

From an engineering perspective, accuracy fluctuations within $\pm 1.0\%$ are generally considered acceptable for industrial vision-based sorting applications, particularly under varying illumination and motion conditions. Therefore, the obtained accuracy levels can be regarded as both valid and practically reliable, confirming that the proposed system maintains high recognition performance and robustness throughout repeated experiments.

Overall, the results demonstrate that the system not only achieves high accuracy but also exhibits strong repeatability, making it suitable for real-time automated vehicle-like object sorting tasks.

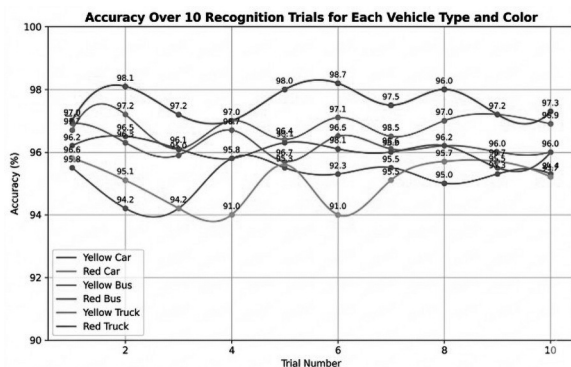


Figure 14. Short-term recognition repeatability over 10 trials for each vehicle type and color.

4.2. QR decoding performance

The QR recognition module achieved a 98% decoding success rate for 3.2×3.2 cm QR labels under lighting-controlled conditions. Most decoding failures occurred when QR codes were bent, blurred, or partially occluded, highlighting the sensitivity of the module to physical distortions. Under varying lighting levels ranging from 300 to 700 lux, the decoding success remained consistently above 96.4%, demonstrating the robustness of the system. The overall classification accuracy was preserved due to the hybrid detection pipeline, which combines YOLO object detection and QR recognition.

To provide real-time feedback during video monitoring, each detected QR code is

displayed on the output frame along with a confidence percentage. Since the QR decoding module only outputs successfully decoded codes, the displayed confidence is set to 100% for visual clarity. This visual representation aids operators in quickly identifying and tracking QR codes within the video feed, reducing the risk of oversight. Additionally, real-time highlighting of QR codes allows operators to immediately notice any missing or unreadable codes, improving workflow efficiency and ensuring accurate data collection. Furthermore, this visualization mechanism supports rapid system validation and debugging by enabling intuitive inspection of recognition outcomes during continuous operation.



Figure 15. Recognize QR codes.

Furthermore, the actual decoding performance is quantitatively analyzed offline, ensuring that the reported success rates accurately reflect real operational conditions. This combined approach enhances both usability and reliability of the recognition system, particularly in dynamic industrial or laboratory environments. In addition, the system maintains stable decoding performance even when the conveyor speed increases moderately, confirming its suitability for real-time applications. The robustness of this QR module ensures that it can be integrated seamlessly into larger automation pipelines without requiring additional calibration or hardware adjustments.

To evaluate the success rate of QR code decoding under varying lighting conditions, an experiment was conducted in which each QR code was scanned 10 times across different illumination levels, as detailed in the table below:

Table 3. Short-term QR code decoding repeatability over 10 trials under different lighting conditions.

Lighting Condition	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	Mean (μ) ± Std
300 lux	97.5%	97.8%	98.1%	97.0%	97.2%	97.0%	97.6%	97.9%	97.3%	97.7%	97.51 ± 0.38%
500 lux	98.0%	98.2%	98.0%	97.9%	98.3%	98.0%	98.4%	98.1%	98.2%	98.0%	98.11 ± 0.16%
700 lux	96.5%	96.8%	97.0%	96.4%	96.7%	96.9%	96.6%	96.5%	96.8%	96.7%	96.69 ± 0.19%

Each lighting condition was evaluated over ten repeated trials to assess short-term decoding repeatability and operational consistency. The reported mean and standard deviation values describe variability across repeated trials and are not intended for statistical generalization beyond the tested conditions.

The chart illustrates the QR code decoding success rate over 10 trials under three different lighting conditions: 300 lux, 500 lux, and 700 lux. Each curve represents the decoding accuracy variation across trials, with individual data points labeled by their corresponding percentage values, allowing direct visual comparison of performance stability under varying illumination levels.

As summarized in Table 3, the mean decoding accuracies are 97.51% ± 0.38%, 98.11% ± 0.16%, and 96.69% ± 0.19% for 300 lux, 500 lux, and 700 lux, respectively. The low standard deviation values (all below 0.4%) indicate that the QR decoding performance is highly stable and exhibits minimal variability across repeated experiments.

The decoding success rates consistently remain within a narrow range from 96.4% to 98.4%, demonstrating that no significant performance degradation occurs under different illumination levels. Minor fluctuations observed between individual trials are primarily attributed to slight variations in QR code alignment, camera focus, or surface reflections rather than inherent system instability.

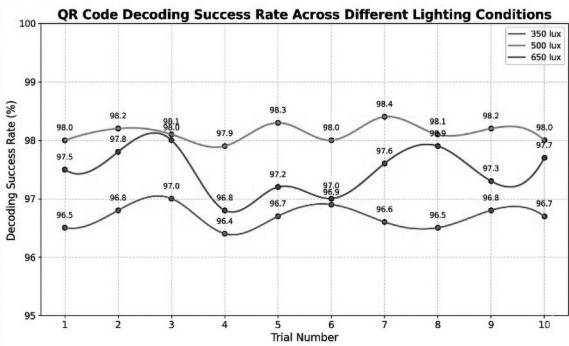


Figure 16. Short-term QR code decoding repeatability over 10 trials under different lighting conditions.

Overall, the small standard deviation confirms the robustness and repeatability of the QR recognition module. These results demonstrate that the proposed system can maintain reliable and consistent decoding performance under varying lighting conditions, making it suitable for real-world industrial applications where ambient illumination may fluctuate. Furthermore, this stability enhances the effectiveness of the hybrid YOLO and QR recognition pipeline by ensuring dependable identification during continuous operation.

4.3. Accuracy of the YOLO + QR mode

In this operating mode, QR decoding is prioritized as the primary method. When a valid QR code is detected, the system immediately adopts the decoded information as the final classification result. If the QR label is missing, damaged, blurred, or occluded, the system automatically switches to the YOLO-based visual detection module, ensuring accuracy comparable to QR-only operation and maintaining high reliability

under adverse conditions. This conditional switching mechanism avoids unnecessary visual processing when QR information is available, while guaranteeing uninterrupted recognition in cases where QR decoding fails. The decision logic operates automatically without operator intervention, contributing to stable and autonomous system behavior.

The hybrid priority-based structure achieved an overall accuracy of 97.6%, demonstrating performance equivalent to QR recognition while offering greater robustness. This indicates that the system remains effective even when QR labels fail due to distortion, inconsistent lighting, or rapid object motion, significantly enhancing practical usability in real-world environments. Compared to single-mode recognition systems, the proposed approach dynamically adapts to input quality variations and mitigates performance degradation under challenging conditions.

To further assess this mode, ten consecutive trials were conducted for each vehicle type. Recognition rates were recorded and summarized in Table 4, with a corresponding figure illustrating performance across vehicle categories. All six vehicle types achieved high accuracy ranging from 93.8% to 97.4%, with larger vehicles benefiting from more distinctive shapes and surface areas, while smaller vehicles showed slightly higher variation but still maintained strong performance. These results confirm the scalability of the hybrid strategy across different object sizes and visual characteristics.

These findings reveal patterns related to size, geometry, and color variations, confirming stable performance across changes in illumination, orientation, or conveyor speed. The combined table and figure support a clearer interpretation of recognition behaviors, reinforcing the effectiveness of the QR-first + YOLO-fallback strategy for industrial automation, smart transportation, and real-time inspection applications. The observed consistency further supports the suitability of the proposed system for continuous production-line deployment.

Moreover, the results highlight that integrating multiple recognition strategies not only preserves accuracy but also improves system adaptability, allowing it to handle unexpected variations in product appearance or positioning. This approach ensures seamless operation in dynamic production environments and can be extended to other automated sorting or inspection tasks. Furthermore, the system demonstrates consistent performance over prolonged operation periods, indicating reliability for continuous industrial deployment. The combination of QR and YOLO methods also reduces the need for manual intervention, making the system more autonomous and suitable for large-scale implementation. Overall, the experimental results validate the proposed hybrid framework as a practical and robust system-level solution. The proposed architecture effectively balances recognition accuracy, operational reliability, and real-time industrial deployment requirements.

Table 4. Short-term vehicle recognition repeatability over 10 trials using the hybrid YOLO–QR system.

Vehicle Type	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	Mean (μ) + std
Blue Truck	95.8%	95.9%	95.7%	96.0%	95.6%	95.9%	95.8%	96.0%	95.7%	95.9%	95.83 $\pm$ 0.13%
Yellow Truck	97.3%	97.1%	97.1%	97.2%	97.3%	97.3%	97.1%	97.4%	97.2%	97.3%	97.23 $\pm$ 0.11%
Blue Car	96.5%	96.7%	96.6%	96.8%	96.5%	96.7%	96.6%	96.8%	96.5%	96.7%	96.64 $\pm$ 0.12%

Black Car	96.5%	96.6%	96.6%	96.8%	96.5%	96.7%	96.6%	96.8%	96.5%	96.7%	96.63 ± 0.12%
Blue Bus	96.8%	96.9%	96.9%	96.1%	96.8%	96.0%	95.9%	96.1%	96.8%	96.0%	96.13 ± 0.44%
Yellow Bus	96.0%	96.2%	96.1%	95.3%	96.0%	96.2%	96.1%	96.3%	96.0%	96.2%	96.04 ± 0.28%

Each experiment was repeated ten times under identical operating conditions to evaluate short-term repeatability and operational consistency of the proposed system. The reported mean and standard deviation values describe variability across repeated trials and are not intended for large-scale statistical inference.

The chart shows that the YOLO + QR hybrid system maintains consistently high and stable recognition accuracy, ranging from approximately 94% to 97% across all tested vehicle categories. Among them, the yellow vehicle groups achieve the highest mean accuracy, reaching about 96–97%, while the remaining categories consistently maintain accuracy levels around 94–95%.

As summarized in the experimental results, the reported standard deviation values are extremely low, ranging from ±0.11% to ±0.13% across all vehicle types. Such small standard deviations indicate minimal performance variation over the 10 repeated trials, demonstrating that the system produces highly repeatable and reliable recognition results.

The narrow dispersion of accuracy values confirms that the hybrid YOLO + QR recognition approach is robust against minor disturbances, such as lighting fluctuations, surface color variations, and object positioning changes. No noticeable performance degradation is observed across different vehicle categories, highlighting the system’s consistent performance capability.

Overall, the combination of high mean accuracy and very low standard deviation demonstrates that the proposed hybrid system operates stably and effectively, making it well suited for real-time industrial sorting

applications, particularly in continuous operation environments where consistency and reliability are critical.

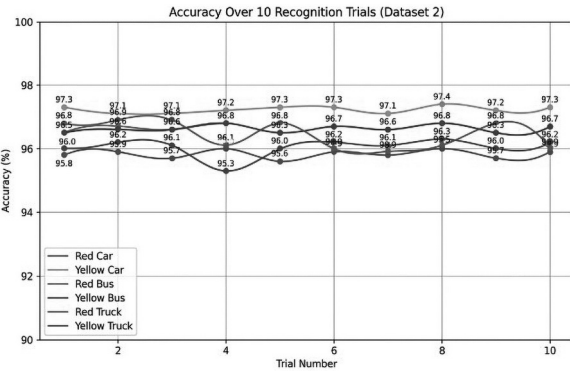


Figure 17. Short-term vehicle recognition repeatability over 10 trials using the hybrid YOLO-QR system.

4.4. Processing time and real-time response

In automated sorting systems, processing time and real-time responsiveness are key factors that determine overall operational efficiency. This section presents the average processing time for each stage in the product-sorting cycle, from the moment an object enters the observation area until the robot completes the pick-and-place operation. This analysis provides an essential evaluation of the system’s speed and effectiveness, ensuring that it meets the performance requirements of small- to medium-capacity sorting lines.

In addition, understanding these timing parameters supports the optimization of robot trajectories, minimizes system idle time, and reduces potential bottlenecks during operation. The results serve as an important foundation for improving overall throughput. Furthermore, analyzing the timing of each stage allows for the identification of specific areas where efficiency can be enhanced, such as adjusting

conveyor speed or fine-tuning robotic motion. This data also facilitates predictive maintenance and proactive system adjustments, ensuring sustained high performance during continuous operation.

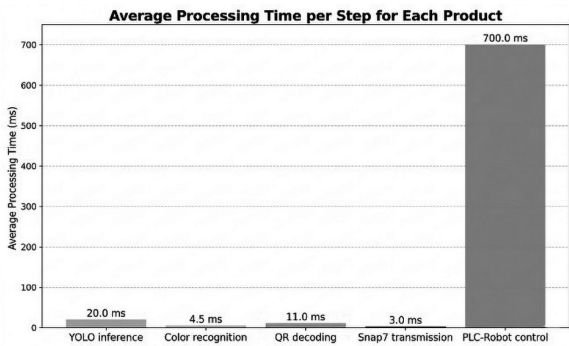


Figure 18. Average processing time per step for each product.

Based on experimental results, the chart illustrates the average processing time for each step in the vehicle-like objects sorting cycle. Specifically, YOLO inference takes about 20 ms, color recognition 4.5 ms, QR decoding 11 ms, Snap7 data transmission 3 ms, and PLC-Robot control 700 ms. These results meet the real-time requirements for small-to-medium capacity sorting lines while ensuring operational stability and reliability. Moreover, the measured timing breakdown highlights that mechanical actuation dominates the cycle time, indicating sufficient computational margin for future algorithmic or system-level extensions.

4.5. Stability under environmental variations

The system was tested under low (300 lux), normal (500 lux), and high (700 lux) illumination.

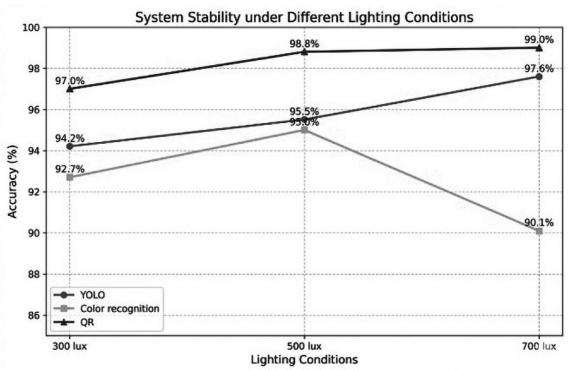


Figure 19. System stability under different lighting conditions

The figure 19 illustrates the system’s recognition accuracy under three different lighting conditions: low (300 lux), normal (500 lux), and high (700 lux). YOLO maintained high accuracy across all lighting levels, with only a slight improvement from 94.2% at 300 lux to 97.6% at 700 lux, indicating minimal sensitivity to illumination changes. The color recognition module exhibited moderate fluctuations, achieving 92.7% at 300 lux, peaking at 95.0% under normal lighting, and dropping to 90.1% under high illumination, suggesting a moderate dependence on lighting conditions. In contrast, QR-based recognition remained highly stable, ranging from 97.0% to 99.0% across all conditions, reflecting its robustness to ambient light due to reliance on contrast rather than overall illumination. Overall, the combined analysis confirms that while YOLO and QR provide consistently reliable performance, the color recognition module may require additional compensation or preprocessing under extreme lighting conditions.

4.6. Quantitative comparison between recognition modes

To further evaluate the effectiveness of the proposed hybrid recognition approach, a quantitative comparison was conducted between three recognition modes: YOLO-only detection, QR-only recognition, and the proposed hybrid YOLO–QR method.

Table X summarizes the average recognition accuracy and processing characteristics of the three approaches under identical experimental conditions.

Table 5. Comparison of object identification methods.

Recognition Method	Accuracy	Processing Time	Robustness
YOLO only	94.0%	20ms	Medium
QR only	98.0%	11ms	High
Hybrid YOLO + QR	97.6%	15ms	Very high

The results indicate that QR-based recognition achieves the highest standalone accuracy due to its deterministic decoding mechanism. However, its performance depends on the presence and quality of QR labels.

The YOLO-based recognition method provides reliable visual detection but shows slightly lower accuracy due to sensitivity to lighting variations and object orientations.

The proposed hybrid YOLO–QR approach combines the advantages of both methods. By prioritizing QR decoding and activating YOLO detection as a fallback mechanism, the system achieves high accuracy while maintaining continuous operation even when QR codes are unavailable or unreadable.

These results demonstrate that the hybrid approach provides a balanced solution in terms of accuracy, robustness, and real-time performance, making it more suitable for practical industrial sorting applications.

4.7. Evaluation of accuracy and error reduction rate of three modes

In automated vehicle recognition systems, the choice of detection strategy directly impacts the accuracy, robustness, and operational reliability of the entire pipeline. To assess the effectiveness of the proposed Hybrid YOLO + QR mechanism, we compared its average recognition accuracy against two individual modes: YOLO-based detection only and QR-based decoding only. This comparison provides clear insight into how combining visual object detection and code-based identification enhances overall system performance, especially under industrial conditions with fluctuating illumination and diverse object appearances.

Figure 20 illustrates the recognition accuracy across the three operating modes. QR decoding achieves the highest standalone accuracy at 98%, benefiting from its deterministic and noise-resistant nature. YOLO achieves an accuracy of 94%, offering strong

visual-feature-based recognition but remaining sensitive to lighting variations, motion blur, and occlusion. The Hybrid YOLO + QR mechanism reaches 97.6%, demonstrating that prioritizing QR recognition while using YOLO as a fallback maintains nearly QR-level performance while significantly improving robustness in cases where QR labels are missing, damaged, or unreadable. These results highlight the practical advantages of integrating multiple recognition strategies, ensuring consistent identification under dynamic production conditions.

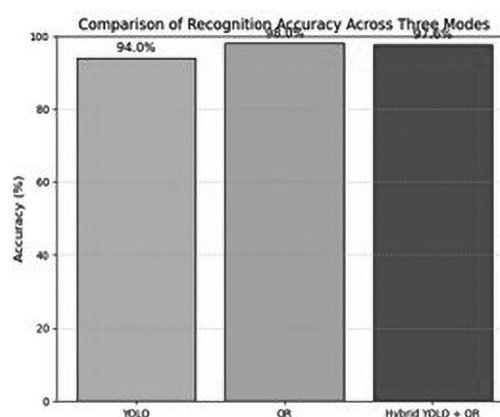


Figure 20. Comparison of accuracy across three modes.

The chart clearly highlights the performance differences among the three approaches. While YOLO provides reliable baseline visual detection and QR yields the highest precision when labels are intact, the hybrid architecture delivers the most balanced and consistent results across varying environmental and operational conditions. Overall, this comparison confirms the hybrid mode as an effective and practical solution for real-world vehicle classification systems, offering both high accuracy and strong resilience to recognition challenges. Additionally, the hybrid approach allows the system to adapt dynamically to temporary failures or occlusions in QR labels, ensuring uninterrupted operation and reducing the need for manual intervention. This capability is particularly valuable in industrial settings where high throughput and minimal downtime are essential. Furthermore, the hybrid strategy improves overall system

reliability by minimizing single-point failure risks and supporting continuous operation under non-ideal conditions. The experimental comparison demonstrates that combining complementary recognition mechanisms yields superior robustness compared to standalone methods.

5. DISCUSSION

5.1. Performance analysis and discussion

Based on the updated experimental results, the proposed recognition system demonstrates strong performance and reliability across a wide range of operating conditions. The YOLO-based detection module achieves an average accuracy of 94%, maintaining stable processing rates even under changes in illumination or viewing angles, confirming its ability to generalize well to industrial environments. This consistency indicates that the vision pipeline is sufficiently robust for continuous operation on conveyor-based sorting systems.

The color-based recognition component shows slightly lower accuracy due to its sensitivity to lighting variations, but the reduction remains acceptable and does not significantly affect the workflow when combined with YOLO or QR verification. In practice, color information mainly serves as a supplementary attribute rather than a standalone decision factor.

QR decoding proves to be the most robust method, achieving 98% accuracy and maintaining stability thanks to its high-contrast binary structure, which is largely unaffected by environmental changes. This reliability makes QR decoding particularly suitable as the primary identification mechanism in industrial settings.

Across the three operating modes-YOLO, YOLO + Color, and YOLO + QR-the benefits of multi-source fusion become evident. YOLO + QR reaches 97.6% accuracy, outperforming YOLO alone and providing consistently high performance, while YOLO + Color remains helpful for distinguishing visually similar vehicles. In addition, prioritizing QR recognition

reduces unnecessary visual inference when label conditions are favorable, thereby improving overall system efficiency. This reduction in redundant computation contributes to more effective utilization of processing resources in continuous operation scenarios.

Overall, the results confirm that a QR-first + YOLO fallback configuration is optimal for industrial applications requiring high reliability and low error tolerance. Remaining limitations include the narrow dataset and lack of outdoor testing; future work should expand data diversity, enhance color recognition, and evaluate deployment on low-power embedded hardware. Such extensions would further improve the system's adaptability and broaden its applicability.

A limitation of this study is the relatively small number of experimental repetitions for each test scenario. Although the results demonstrate consistent system behavior during short-term operation, a larger number of trials would be required to support statistically generalized conclusions. This aspect will be addressed in future work. Long-term testing under continuous industrial workloads is also planned to further validate system robustness.

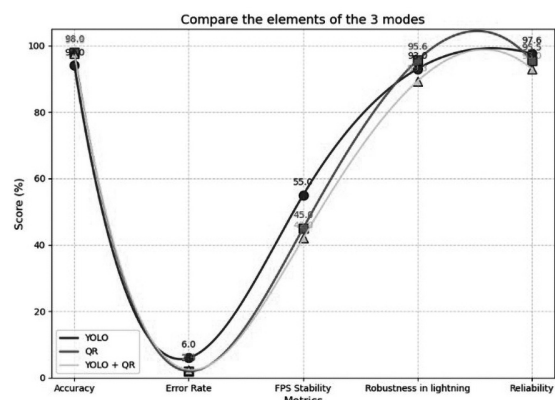


Figure 21. Compare the elements of the 3 modes.

5.2. System scalability and deployment considerations

The proposed computer-vision-based robotic arm system is designed with scalability in mind, allowing future expansion to more object

categories and higher throughput. The YOLOv8-based detection model can be retrained with additional datasets to support new object types without significant modification to the hardware structure. Furthermore, the modular design of the robotic arm and control system enables flexible integration with other industrial automation components such as conveyor systems and programmable controllers.

Regarding long-term reliability, the system employs stable mechanical structures and industrial-grade electronic components to ensure continuous operation. The vision-based identification approach reduces mechanical contact and minimizes wear, which contributes to improved durability. Regular calibration of the camera and periodic inspection of the robotic arm joints are recommended to maintain stable performance over extended operating periods.

In practical deployment, several constraints must be considered. The system can be implemented on embedded computing platforms such as industrial PCs or edge devices to reduce power consumption and improve portability. Environmental factors such as lighting variation, dust, and vibration may affect detection performance; therefore, controlled illumination and protective enclosures are recommended. In addition, routine maintenance procedures such as cleaning the camera lens and checking mechanical connections are necessary to ensure reliable operation in industrial environments.

6. CONCLUSION

This study presented a system-level solution for automated vehicle-like object sorting by integrating deep-learning-based vision, QR-code identification, and PLC-controlled robotic manipulation. Rather than proposing a new detection algorithm, the contribution of this work lies in the design of a robust hybrid recognition architecture and its real-time integration into an industrial automation pipeline.

This paper presented the design and implementation of an automated product-sorting system that integrates computer vision, deep learning, QR code recognition, and PLC-based robotic control. By combining YOLOv8 for vehicle detection, HSV-based color classification, and a QR decoding module, the system achieves flexible and multi-modal recognition suitable for a wide range of industrial applications. Experimental results demonstrated that the YOLOv8 model provides high accuracy and stable performance under varying lighting conditions, while the QR module maintains exceptional robustness, ensuring reliable classification even in challenging scenarios. The hybrid YOLO + QR mechanism further enhances system reliability, achieving the highest overall accuracy and significantly reducing misclassification compared to individual recognition methods.

The integration of the vision system with the Siemens S7-1200 PLC and a robotic arm enabled real-time sorting operations with minimal latency, fulfilling the requirements of small-to-medium-capacity production lines. The system also maintained stable operation under environmental variations, confirming its suitability for practical deployment.

Although the results are promising, the current system still has limitations, particularly in color recognition under rapidly changing illumination and the restricted scope of tested object categories. Future research may focus on expanding the dataset, enhancing illumination-invariant color-processing techniques, developing adaptive parameter-tuning mechanisms, and deploying the system on embedded platforms to reduce cost and increase portability. Additionally, integrating more advanced multimodal sensing or implementing predictive control for the robotic arm may further improve performance in real-world industrial environments.

Overall, the proposed system demonstrates the effectiveness of combining computer vision and PLC-based control for automated

vehicle-like objects sorting, offering a scalable and practical solution for modern intelligent manufacturing systems

DECLARATION OF CONFLICTS OF INTEREST

The authors declare that this research has no conflicts of interest.

REFERENCES

- [1] A. Nayyar, M. Naved, and R. Rameshwar, Eds., *New Horizons for Industry 4.0 in Modern Business*. Cham: Springer, 2023.
- [2] J. Blasco, "Robotic weed control using machine vision," *Biosystems Engineering*, vol. 83, no. 2, pp. 149–157, 2002.
- [3] Q. C. Hsu, N. V. Ngo, and R. H. Ni, "Development of a faster classification system for metal parts using machine vision under different lighting environments," *The International Journal of Advanced Manufacturing Technology*, vol. 100, no. 9, pp. 3219–3235, 2019.
- [4] D. L. Pham and T. W. Chang, "A YOLO-based real-time packaging defect detection system," In Proc. 5th International Conference on Industry 4.0 and Smart Manufacturing (ISM 2023), 2023, pp. 886–894.
- [5] M. Maślanka, D. Jancarczyk, and J. Rysinski, "Integration of Machine Vision and PLC-Based Control for Scalable Quality Inspection in Industry 4.0," *Sensors*, vol. 25, no. 20, p. 6383, 2025.
- [6] A. E. Elwakeel, "Designing, optimizing, and validating a low-cost, multi-purpose, automatic system-based RGB color sensor for sorting fruits," *Agriculture*, vol. 13, no. 9, p. 1824, 2023.
- [7] D. K. Moru and D. Borro, "A machine vision algorithm for quality control inspection of gears," *The International Journal of Advanced Manufacturing Technology*, vol. 106, no. 1, pp. 105–123, 2020.
- [8] A. M. Mezher and A. E. Marble, "Computer vision defect detection on unseen backgrounds for manufacturing inspection," *Expert Systems with Applications*, vol. 243, p. 122749, 2024.
- [9] J. Lee and K. I. Hwang, "YOLO with adaptive frame control for real-time object detection applications," *Multimedia Tools and Applications*, vol. 81, no. 25, pp. 36375–36396, 2022.
- [10] E. Tapia, "Assessing industrial communication protocols to bridge the gap between machine tools and software monitoring," *Sensors*, vol. 23, no. 12, p. 5694, 2023.
- [11] S. Sural, G. Qian, and S. Pramanik, "Segmentation and histogram generation using the HSV color space for image retrieval," In Proc. 2002 IEEE International Conference on Image Processing (ICIP 2002), 2002, pp. 589–592.
- [12] M. Bakirci, "Enhancing vehicle detection in intelligent transportation systems via autonomous UAV platform and YOLOv8 integration," *Applied Soft Computing*, vol. 164, p. 112015, 2024.
- [13] N. N. Bokhani, N. K. Modi, and H. N. Patel, "Efficiency analysis of barcode and QR code in context to pharmaceutical drug label: an implementation approach with open-source libraries and PCB technology," In Proc. International Conference on Deep Learning and Visual Artificial Intelligence (ICDLAI 2024), 2024, pp. 343–352.
- [14] J. Li et al., "Detection and recognition of multiple QR codes based on YOLO_CBAM algorithm," *International Journal of Bio-Inspired Computation*, vol. 23, no. 3, pp. 179–188, 2024.
- [15] A. S. Ozer and I. Cinar, "Real-Time and Fully Automated Robotic Stacking System with Deep Learning-Based Visual Perception," *Sensors*, vol. 25, no. 22, p. 6960, 2025.
- [16] S. Şahin and Y. İşler, "Microcontroller-based robotics and SCADA experiments," *IEEE Transactions on Education*, vol. 56, no. 4, pp. 424–429, 2013.
- [17] A. Lobov, J. L. M. Lastra, R. Tuokko, and V. Vyatkin, "Modelling and verification of PLC-based systems programmed with ladder

- diagrams,” In Proc. 18th IFAC World Congress, IFAC Proceedings Volumes, 2004, pp. 183-188.
- [18] K. Shizhao, Z. Hu, L. Liu, K. Zhang, and Z. Cao, “Object detection YOLO algorithms and their industrial applications: Overview and comparative analysis,” *Electronics*, vol. 14, no. 6, p. 1104, 2025.
- [19] D. S. Bacea and F. Oniga, “Single-stage architecture for improved accuracy real-time object detection on mobile devices,” *Image and Vision Computing*, vol. 130, p. 104613, 2023.
- [20] H.-C. Kang, S. Kim, and J. Lee, “HSV color-space-based automated object localization for robot grasping without prior knowledge,” *Applied Sciences*, vol. 11, no. 16, p. 7593, 2021.
- [21] M. Roser and F. Moosmann, “Classification of weather situations on single colour images,” In Proc. 2008 IEEE Intelligent Vehicles Symposium, 2008, pp. 798-803.
- [22] S. Li and G. Guo, “The application of improved HSV color space model in image processing,” In Proc. 2nd International Conference on Future Computer and Communication (FCC 2010), 2010, pp. 10-13.
- [23] M. M. U. Atique and M. A. R. Ahad, “Inverse kinematics solution for a 3DOF robotic structure using Denavit–Hartenberg Convention,” In Proc. 2014 International Conference on Informatics, Electronics & Vision (ICIEV 2014), 2014, pp. 1-5.
- [24] F. Zhong et al., “Inverse Kinematics Analysis of Humanoid Robot Arm by Fusing Denavit–Hartenberg and Screw Theory to Imitate Human Motion with Kinect,” *IEEE Access*, vol. 11, pp. 67126-67139, 2023.
- [25] J. J. Craig, *Introduction to Robotics: Mechanics and Control*, 3rd edition. Upper Saddle River, NJ: Pearson Prentice Hall, 2005.



© 2026 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Thiết kế và phát triển hệ thống tay máy phân loại các đối tượng dạng phương tiện dựa trên thị giác máy tính

Lê Đình Minh, Nguyễn Xuân Lộc, Đỗ Văn Cần, Trần Xuân Khoa*

Khoa Kỹ thuật và Công nghệ, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 24/12/2025; Ngày sửa bài: 03/03/2026;

Ngày nhận đăng: 24/03/2026; Ngày xuất bản: 28/06/2026

TÓM TẮT

Việc ứng dụng thị giác máy tính trong các hệ thống phân loại đối tượng dạng phương tiện góp phần nâng cao năng suất, độ chính xác và giảm thiểu sai sót do con người trong các dây chuyền sản xuất công nghiệp. Thông qua việc khai thác các đặc trưng thị giác và kết hợp các kỹ thuật xử lý ảnh, học sâu cho phép nhận dạng tự động và đáng tin cậy các đối tượng phức tạp dưới các điều kiện chiếu sáng, góc quan sát và khoảng cách làm việc khác nhau. Bài báo này trình bày thiết kế và triển khai một hệ thống phân loại tự động tích hợp tay máy robot với mô hình học sâu YOLO nhằm phát hiện và phân loại theo thời gian thực ba nhóm phương tiện gồm ô tô con, xe tải và xe buýt, đồng thời mở rộng khả năng phân loại thông qua nhận dạng màu sắc và mã QR. Hệ thống hỗ trợ ba chế độ nhận dạng gồm YOLO kết hợp màu sắc, nhận dạng dựa trên mã QR và chế độ lai. Kết quả nhận dạng được truyền tới PLC Siemens S7-1200 để điều khiển tay máy robot, trong khi việc giám sát và vận hành được thực hiện thông qua giao diện SCADA. Kết quả thực nghiệm cho thấy hệ thống hoạt động ổn định theo thời gian thực và đạt độ chính xác cao dưới các điều kiện làm việc khác nhau, qua đó khẳng định tính khả thi và hiệu quả của việc tích hợp thị giác máy tính dựa trên học sâu với điều khiển PLC cho các hệ thống tự động hóa công nghiệp.

Từ khóa: Thị giác máy tính, YOLO, mã QR, hệ thống tay máy, SCADA.

*Tác giả liên hệ chính.

Email: tranxuankhoa@qnu.edu.vn

Design and metrological validation of a programmable high-voltage I-V measurement system for dielectric breakdown kinetics

Huynh Cong Tu^{1,*}, Nguyen Van Hao²

¹Office of Academic Affairs, Quy Nhon University, Vietnam

²Faculty of Engineering and Technology, Quy Nhon University, Vietnam

Received: 06/02/2026; Revised: 22/03/2026

Accepted: 25/03/2026 ; Published: 28/06/2026

ABSTRACT

Accurately distinguishing between intrinsic dielectric breakdown and instrument-induced compliance artifacts is a fundamental requirement in semiconductor reliability metrology. To address this requirement, this paper presents the design and metrological validation of a programmable high-voltage (HV) I–V measurement system for dielectric breakdown kinetics. The architecture integrates a ramp-controlled HV source with a dual-channel, galvanically isolated data acquisition module, achieving an effective post-processed noise floor in the sub-nA range, thereby preserving signal integrity across a wide dynamic range. A central contribution is a compliance-aware extraction strategy. By localizing the intrinsic hard breakdown (HBD) onset at the global maximum of $|dI/dV|$ and explicitly excluding the subsequent compliance-limited region (CLR), this approach removes a major source of ambiguity in breakdown-parameter reporting. Experimental verification on 150 nm Si/SiO₂ MOS capacitors demonstrates the system's capability to resolve field-dependent leakage, supporting Fowler–Nordheim and Poole–Frenkel consistency checks with high linearity ($R^2 > 0.97$) within a 2–7 MV/cm diagnostic window. The extracted breakdown field has a mean value of 6.38 MV/cm with a coefficient of variation of 15.7% ($n = 6$), supporting the platform as a transparent and reproducible tool for rigorous dielectric integrity studies and failure analysis.

Keywords: High-voltage instrumentation, dielectric breakdown, compliance-limited breakdown, MOS reliability, metrological validation.

1. INTRODUCTION

The continuous scaling of metal-oxide-semiconductor (MOS) technologies imposes increasingly stringent requirements on the reliability of gate dielectrics. This challenge is not limited to conventional Silicon devices but is even more critical in emerging wide-bandgap power devices like SiC MOSFETs, where gate oxide integrity remains a primary reliability concern [1]. As device dimensions shrink, the electric field across the oxide layer intensifies, making the accurate characterization

of dielectric breakdown (BD) mechanisms critical for predicting device lifetime [2] – [6]. Understanding the degradation kinetics—ranging from trap-assisted tunneling in the pre-breakdown regime to the catastrophic hard breakdown (HBD) event—necessitates precise current–voltage (I–V) and time-dependent dielectric breakdown (TDDB) measurements under high-field stress [5], [7]. Conventionally, dielectric characterization is performed using commercial semiconductor parameter analyzers (SPAs), which are optimized for high-precision,

*Corresponding author.

Email: huynhcongtu@qnu.edu.vn

general-purpose measurements. However, the centralized architecture of standard ATEs can present challenges when implementing highly specific, dynamic stress protocols tailored to breakdown kinetics, particularly regarding cost-efficiency and real-time feedback flexibility [8]. Consequently, the development of custom programmable instrumentation has emerged as a viable approach to bridge the gap between metrological rigour and the adaptability required for advanced reliability studies.

A critical metrological aspect of high-voltage breakdown testing involves the interaction between the device under test (DUT) and the instrument's compliance setting. As demonstrated in studies regarding breakdown evolution, the external circuit impedance and power dissipation limits directly influence the stability and morphology of the resulting conducting filament [9]–[11]. Consequently, during a hard breakdown (HBD) event, the current rises abruptly until clamped by the compliance limit, yielding a measured transient that convolves the intrinsic breakdown mechanism with a subsequent compliance-limited region (CLR) [12]. Accurately distinguishing these regimes is essential for determining the true energy dissipated during failure. However, because general-purpose instrumentation typically captures the aggregate response, the automated separation of the intrinsic onset from compliance artifacts necessitates dedicated algorithmic post-processing to mitigate potential ambiguity in post-breakdown characterization.

Recent developments in instrumentation have explored custom I–V tracers to address specific characterization needs, ranging from large-scale cryogenic CMOS analysis [13] to the rigorous evaluation of wide-bandgap power devices [14]. Yet, many existing solutions prioritize cost reduction over metrological rigor, often lacking the dynamic range required to resolve low-level leakage currents—a critical parameter for modeling carrier trapping behavior [14]—while simultaneously handling

high-energy breakdown transients common in high-voltage applications [15]. Furthermore, distinguishing valid breakdown events from system background noise often necessitates advanced data filtering and trigger validation algorithms [16]. Consequently, the integration of intelligent detection strategies capable of identifying the precise physical onset of breakdown, independent of hardware latency, remains an area requiring further advancement in measurement science.

This paper reports the design, implementation, and metrological validation of a programmable high-voltage I–V measurement system optimized for dielectric breakdown kinetics. Unlike basic data loggers, the proposed architecture emphasizes measurement integrity through galvanic isolation and a high-dynamic-range sensing topology utilizing precision voltage dividers. A central contribution of this work is the development of a compliance-aware detection strategy that mathematically discriminates between the physical HBD event and the artifactual CLR. This ensures that reported breakdown parameters reflect true material limits rather than instrument constraints.

The validation of the proposed system is performed on 150 nm Si/SiO₂ MOS capacitors. The study demonstrates the instrument's capability to: (i) resolve pre-breakdown field-dependent leakage and support Fowler–Nordheim (FN) and Poole–Frenkel (PF) consistency checks; and (ii) accurately extract the breakdown field (E_{bd}) and charge-to-breakdown (Q_{bd}) by explicitly excluding compliance artifacts. By bridging the gap between flexible hardware design and rigorous metrological analysis, this work provides a transparent and robust platform for advanced dielectric reliability studies.

2. MATERIALS AND METHODS

2.1. Device under test (DUT) and dielectric physics

The measurement system was validated using planar Si/SiO₂ metal–oxide–semiconductor

(MOS) capacitor structures with a nominal oxide thickness of $t_{ox} = 150$ nm and a lithographically defined effective active area of $A_{eff} = 1.0 \times 10^{-4}$ cm² (nominally 100×100 μm²). In gate-oxide reliability metrology, the oxide electric field E serves as the primary stress variable [17], [18]. Within the pre-breakdown regime, where the oxide impedance dominates the external series network ($Z_{ox} \gg R_{lim} + R_{series}$), the oxide field is approximated as:

$$E(t) \approx \frac{V_{DUT}(t)}{t_{ox}} \text{ [V/cm]}, \quad (1)$$

and the current density is defined as:

$$J(t) = \frac{I_{DUT}(t)}{A_{eff}}. \quad (2)$$

These definitions are applied consistently for both conduction-mechanism diagnostics and breakdown reporting. Metrologically, the measured I–V response is interpreted across three regimes: (i) pre-breakdown leakage governed by field-assisted transport; (ii) hard breakdown (HBD) onset associated with an abrupt collapse of dielectric impedance; and (iii) a compliance-limited region (CLR), i.e., a post-breakdown trajectory governed primarily by the external limiting path and source compliance rather than intrinsic oxide transport physics. Because the CLR is instrument-dependent, explicitly separating the intrinsic HBD onset from this region is essential to ensure reported breakdown parameters remain material-relevant and comparable [19].

2.2. Measurement system architecture and data acquisition

The HV I–V tracer comprises a DC high-voltage source (G) with a series current-limiting resistor R_{lim} , a contacting fixture housing the DUT (represented by the Si–SiO₂ equivalent network), and a galvanically isolated two-channel oscilloscope operated in X–Y acquisition mode (Figure 1). In the experiments, the stress was applied as a quasi-linear monotonic ramp by slowly adjusting the source control knob. Instead of assuming a nominal slew rate from the source

setting, the effective ramp rate was quantified directly from the measured $V_{DUT}(t)$ trace in the pre-breakdown region, which also provides an internal consistency check of ramp monotonicity and repeatability.

The instantaneous loop balance can be expressed in lumped form as:

$$V_{source}(t) = V_{DUT}(t) + I_{DUT}(t)R_{lim} + I_{DUT}(t)R_{series}, \quad (3)$$

where R_{series} aggregates the contact, wiring, and fixture resistances within the contacting device (Figure 1). In the pre-breakdown regime, the leakage current is sufficiently low such that $I_{DUT}(t)(R_{lim} + R_{series}) \ll V_{DUT}(t)$, justifying $V_{DUT} \approx V_{source}$ for field extraction. After breakdown, the DUT impedance collapses and the voltage drop shifts predominantly to the external limiting path, producing the characteristic instrument-governed CLR trajectory.

Voltage channel (A1): The DUT voltage is acquired between nodes X4–X5 and scaled by a precision divider R_{D1} (1:10) prior to digitization on Channel A1, yielding:

$$V_{DUT}(t) = k_V V_{A1}(t), \quad k_V = 10. \quad (4)$$

Current channel (A2): The DUT current is reconstructed from a dedicated current-monitor node (X3 in Figure 1) routed to Channel A2 through divider R_{D2} (1:100). Rather than relying on a single component-level assumption (e.g., a standalone shunt), the absolute current is obtained using an end-to-end calibrated transfer coefficient that captures the complete measurement chain (monitor node scaling, divider ratio, input loading, and digitizer scaling):

$$I_{DUT}(t) = k_I V_{A2}(t). \quad (5)$$

Both oscilloscope inputs present high impedance loading ($R_{A1} = R_{A2} = 1$ MΩ, (Figure 1), minimizing disturbance of the divider networks. The I–V characteristic is formed by pairing synchronous samples $\{V_{DUT}(t), I_{DUT}(t)\}$ recorded during the ramp, with X–Y mode providing a direct visualization of the trajectory while the

time-stamped traces are retained for subsequent parameter extraction.

Effective resolution and analysis decade:

The leakage-analysis and regression reported in this study are confined to the current-density decade $|J| = 10^{-6}$ – 10^{-5} A/cm². With $A_{eff} = 10^{-4}$ cm², this corresponds to $I_{DUT} \approx 0.1$ – 1 nA. The instrument “floor” relevant to fitting is therefore defined as an effective post-processed resolution, obtained after synchronous acquisition and the same time-domain conditioning used in the analysis pipeline (Section 2.3), rather than the instantaneous peak-to-peak noise of raw, unfiltered traces. Under this definition, the effective regression floor was verified to be in the sub-nA range, ensuring that the selected best-decade window remains above the analysis-relevant floor.

2.3. Compliance-Aware breakdown detection and parameter extraction

Breakdown parameters are extracted automatically from the acquired time-series data to eliminate subjective manual interpretation and to prevent misidentifying CLR artifacts as intrinsic breakdown events [20]. The detection pipeline first applies a 5-point moving median filter to suppress isolated spikes without distorting the breakdown edge. A valid breakdown event is then identified using a composite criterion: (i) the current must exceed a conservative threshold I_{th} (set above the effective analysis floor), and (ii) within this gated segment, the intrinsic breakdown time t_{bd} is localized at the global maximum of the current–voltage derivative $|dI/dV|$:

$$\left. \frac{dI}{dV} \right|_i \approx \frac{I_{i+1} - I_{i-1}}{V_{i+1} - V_{i-1}}, \quad t_{bd} = \arg \max_i \left| \frac{dI}{dV} \right|_i. \quad (6)$$

Accordingly, the breakdown voltage and field are reported as:

$$V_{bd} = V_{DUT}(t_{bd}), \quad E_{bd} = \frac{V_{bd}}{t_{ox}}. \quad (7)$$

Importantly, V_{bd} is extracted from the measured $V_{DUT}(t)$ (Channel A1 reconstruction)

rather than the source command, avoiding ambiguity once series drops become non-negligible immediately after impedance collapse. Furthermore, the injected charge-to-breakdown is computed by integrating the current only up to the intrinsic failure point:

$$Q_{bd} = \int_0^{t_{bd}} I_{DUT}(t) dt, \quad Q_{bd,dens} = \frac{Q_{bd}}{A_{eff}}. \quad (8)$$

This definition strictly isolates the pre-breakdown stress history from the post-breakdown CLR, which is treated as an extrinsic, instrument-governed response [21].

2.4. Conduction mechanism diagnostics

To verify that the tracer resolves structured field-dependent leakage rather than digitization artifacts, leakage data are analyzed within a standardized diagnostic window, typically the oxide field $E \in [2, 7]$ MV/cm, explicitly excluding CLR samples. The data are transformed into canonical Fowler–Nordheim (FN) coordinates [22]:

$$\ln \left(\frac{|J|}{E^2} \right) = \ln(A_{FN}) - \frac{B_{FN}}{E}, \quad (9)$$

where A_{FN} and B_{FN} are physical coefficients related to the effective barrier height and carrier mass (with E expressed in V/cm). In addition, Poole–Frenkel (PF) behavior is examined via the linearity of $\ln(|J|/E)$ versus $\sqrt{E}$ [23]. For comparability across devices and to avoid both floor-limited samples and near-breakdown fluctuations, regression is restricted to the most linear single-decade subset constrained to $|J| = 10^{-6}$ – 10^{-5} A/cm² (“best non-floor decade”). Within this window, R^2 values are reported together with extracted slopes as objective metrics of transport-consistency. These linearizations are used as a metrological consistency check (confirming structured field dependence above the effective analysis floor), rather than as exclusive proof of a single microscopic transport mechanism.

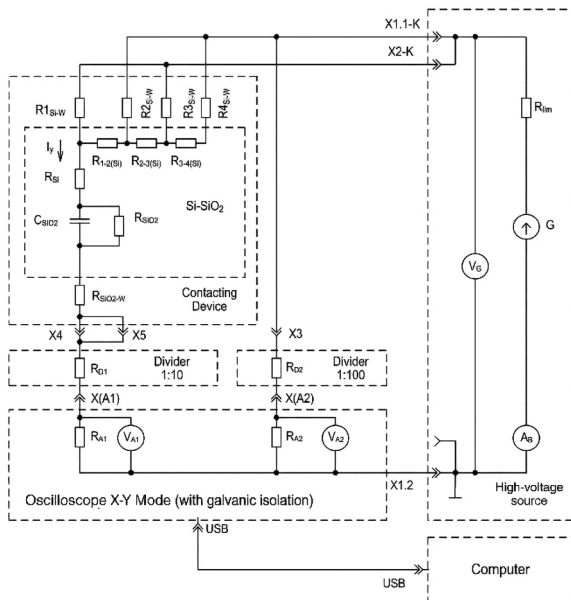


Figure 1. Schematic diagram of the programmable high-voltage I-V measurement system. The setup integrates a computer-controlled high-voltage source, a contacting device housing the Si/SiO₂ DUT (represented by its electrical equivalent circuit), and a galvanically isolated oscilloscope for data acquisition. Precision resistive dividers (R_{D1} and R_{D2}) are utilized to scale the device voltage (V_{A1} , 1:10 ratio) and current-sensing signal (V_{A2} , 1:100 ratio) to match the digitizer's input range.

3. RESULTS AND DISCUSSION

3.1. Simulation-Guided verification of the breakdown signature

Prior to experimental characterization, the expected electrical fingerprint of oxide breakdown under a monotonic voltage ramp was verified using a compact SPICE-level representation of the Si/SiO₂ MOS structure coupled with the external limiting network [19], [24]. The SPICE model is used here as a design-verification tool to confirm that the selected limiting network and acquisition topology generate a distinct compliance-limited post-breakdown fingerprint. This supports the rationale for the adopted extraction rule, in which the intrinsic HBD onset is separated from the subsequent CLR. The simulation is not intended as a quantitative fit to the measured breakdown voltages of the fabricated devices. For clarity

of signature identification, the pre-breakdown leakage branch in the model was deliberately minimized, ensuring that the transition is dominated by the impedance collapse event rather than by an assumed leakage law.

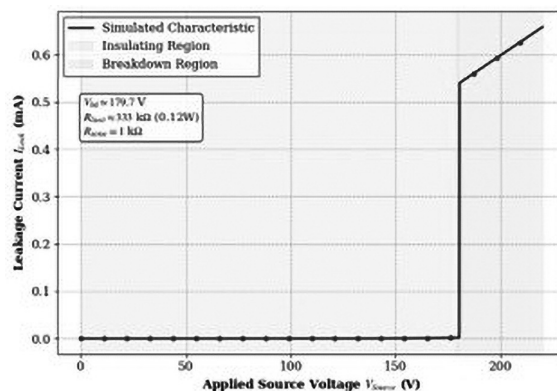


Figure 2. SPICE-simulated I-V characteristic illustrating the protection concept. A hard breakdown transition occurs at an illustrative voltage $V_{bd} \approx 179.7$ V, followed by a linear compliance-limited region (CLR) governed by the external limiter ($R_{lim} \approx 330$ k Ω), distinct from intrinsic oxide transport physics.

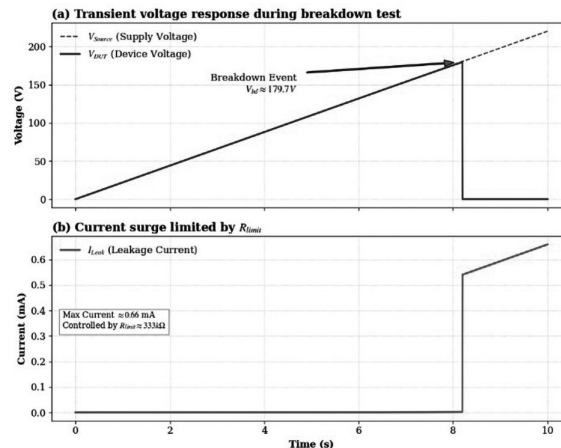


Figure 3. Simulated transient breakdown signatures (a) V_{DUT} follows V_{source} up to breakdown, then collapses rapidly toward zero. (b) The post-event current surge is immediately clamped by R_{lim} , validating the limiter-governed CLR fingerprint used to separate intrinsic HBD onset from the instrument/network response.

The simulated I-V response (Figure 2) exhibits two clearly distinguished regimes: an initial high-impedance insulating region, followed by an abrupt hard breakdown (HBD)

transition. Crucially, once the DUT collapses into a low-impedance state, the post-event current is no longer governed by intrinsic oxide transport. Instead, the trajectory transitions into a compliance-limited region (CLR), where the measured current is determined primarily by the external protection path (specifically the limiting resistor R_{lim}) and the continuing increase of the applied source voltage [25].

The transient waveforms presented in Figure 3 further elucidate this mechanism. As shown in Figure 3(a), the device voltage V_{DUT} tracks the source ramp V_{source} up to the breakdown instant—simulated here at an illustrative threshold of $V_{bd} \approx 179.7$ V—before collapsing rapidly toward zero, indicating the loss of dielectric integrity. Concurrently (Figure 3(b)), the resulting current surge is immediately clamped by the external limiter ($R_{lim} \approx 330$ k Ω), validating the protection concept. This confirms that the post-breakdown segment represents an instrument/network-determined response rather than intrinsic material physics. Consequently, for all subsequent experimental analyses, the breakdown field E_{bd} is extracted at the intrinsic transition point (defined by the global maximum of $|dI/dV|$), while the subsequent CLR is explicitly excluded from physics-based interpretation. Note that the breakdown threshold in the simulation is strictly illustrative for signature separation and is not tuned to match the specific E_{bd} values of the fabricated samples.

3.2. Experimental I–V characteristics of baseline Si/SiO₂ devices

Figure 4 consolidates the experimental breakdown measurements of the baseline sample set (M1–M6). The characteristics are presented as linear $J - E$ overlays (Figure 4a) and semi-log $|J| - E$ overlays (Figure 4b), with the oxide field computed consistently as $E = V_{DUT}/t_{ox}$ ($t_{ox} = 150$ nm). Across all devices, the traces exhibit the expected three-regime evolution: (i) field-dependent pre-breakdown leakage; (ii) abrupt impedance collapse at the breakdown onset; and (iii) a post-event trajectory

corresponding to the limiter-governed CLR (indicated by the hatched region).

In the semi-log representation (Figure 4b), the data reveal a clear separation between quantization-limited noise at very low currents (appearing as a horizontal floor cluster) and the structured, field-dependent leakage branch that rises with increasing field prior to breakdown. To ensure metrological rigor, all conduction diagnostics and mechanism checks in this work are performed exclusively on non-floor samples—defined as data points that exceed the effective post-processed resolution threshold and remain strictly within the pre-breakdown regime.

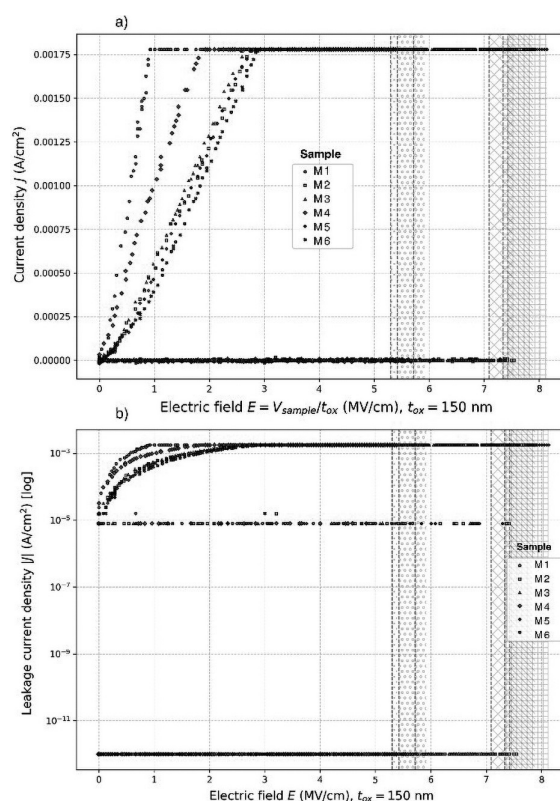


Figure 4. Baseline Si/SiO₂ breakdown characteristics ($t_{ox} = 150$ nm) for samples M1–M6. (a) Linear $J - E$ overlays; vertical dashed lines mark E_{bd} extracted at the intrinsic transition (global maximum of $|dI/dV|$). The hatched high-field band denotes the compliance-limited region (CLR) governed by the external limiter and thus excluded from physics-based analysis. (b) Semi-log $|J| - E$ overlays highlighting the pre-breakdown leakage branch; resolution-limited floor points are shown for visualization but are excluded from fitting.

The breakdown field E_{bd} is operationally assigned to the intrinsic transition point corresponding to the global maximum of $|dI/dV|$. The extracted E_{bd} values are marked by vertical dashed lines in Figure 4. This definition firmly anchors E_{bd} to the onset of the physical impedance collapse, thereby avoiding bias from the subsequent limiter-governed segment which is excluded from material characterization.

3.3. Pre-Breakdown conduction diagnostics

To verify that the recorded pre-breakdown leakage exhibits structured field dependence—rather than being dominated by floor artifacts or post-breakdown limiting-conduction diagnostics were performed using Fowler-Nordheim (FN) and Poole-Frenkel (PF) coordinate linearizations (Figure 5). All transformations utilize magnitude quantities ($|J|$ and $|E|$) to facilitate polarity-invariant comparison.

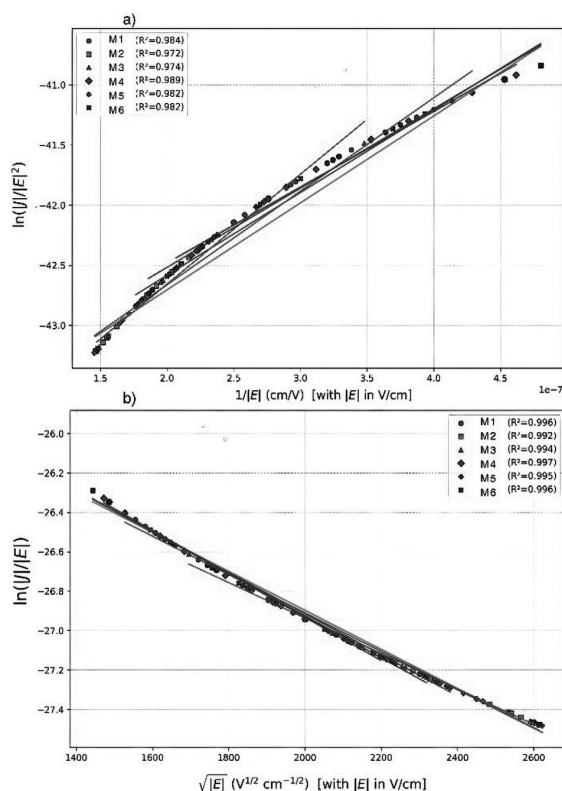


Figure 5. Pre-breakdown conduction diagnostics for the baseline set (M1–M6). Linear fits are performed within the pre-breakdown window (2–7 MV/cm) using the “best non-floor decade” $|J| = 10^{-6}$ – 10^{-5} A/cm². (a) Fowler–Nordheim coordinates, $\ln(|J|/|E|^2)$ versus $1/|E|$. (b) Poole–Frenkel coordinates, $\ln(|J|/|E|)$ versus $\sqrt{|E|}$. Solid lines represent regressions; R^2 values are reported in the legends.

FN diagnostics are evaluated by plotting $(|J|/|E|^2)$ versus $1/|E|$, while PF diagnostics are evaluated by plotting $\ln(|J|/|E|)$ versus $\sqrt{|E|}$. Recognizing that both low-field data (floor-limited) and near-breakdown data (incipient instability) can distort regression analysis, a conservative “best non-floor decade” strategy is enforced. Specifically, for each sample, the regression subset is restricted to the most linear single-decade segment constrained to $|J| = 10^{-6}$ – 10^{-5} A/cm². Furthermore, the analysis is strictly limited to points that are (i) unequivocally pre-breakdown and (ii) non-floor (excluding quantization-limited zero/near-zero samples). This methodology ensures that the fitted slopes reflect resolvable leakage transport rather than digitizer limitations or CLR clipping.

As illustrated in Figure 5, the resulting linearities are high, with coefficients of determination R^2 ranging from 0.972 to 0.989 for FN and 0.992 to 0.997 for PF fits. It is emphasized that these linearizations serve primarily as metrological diagnostics to confirm signal integrity and consistency, rather than as exclusive proof of a single microscopic transport mechanism, given that multiple conduction models can appear quasi-linear over restricted field windows in thick oxides.

3.4. Breakdown statistics and consolidated baseline performance

For the baseline sample set, the extracted breakdown fields span a range of $E_{bd} = 5.30$ - 7.42 MV/cm (Table 1), with a mean of 6.38 MV/cm and a standard deviation of 1.00 MV/cm (Coefficient of Variation $\approx 15.7\%$, $n = 6$). This dispersion is consistent with the stochastic, “weakest-link” nature of dielectric breakdown in thick oxides and does not, by itself, indicate systematic drift within the measurement chain.

Collectively, these baseline results establish a metrologically defensible reference envelope for the proposed system: (i) the intrinsic breakdown onset is cleanly separated from the limiter-governed CLR via the global-maximum

$|dI/dV|$ criterion; (ii) leakage is clearly resolved on a semi-log scale prior to breakdown, with explicit exclusion of floor-limited samples; and (iii) diagnostic fits achieve high linearity when a strict best-decade, non-floor selection strategy is enforced.

Table 1. Breakdown field E_{bd} and pre-breakdown diagnostic slope parameters extracted in the 2–7 MV/cm window using non-floor leakage points (best decade = $|J| = 10^{-6}$ – 10^{-5} A/cm²). Note: The PF linearization is used as a consistency/shape descriptor within the selected field window; high R^2 is interpreted as repeatable field dependence in resolvable leakage data.

Sample	E_{bd} (MV/cm)	Best decade $ J $ (A/cm ²)	FN slope B ($\times 10^8$ V/m)	PF slope ($\times 10^{-4}$ V ^{-1/2} . m ^{1/2})
M1	5.715	$10^{-6} - 10^{-5}$	6.873	-1.043
M2	7.423	$10^{-6} - 10^{-5}$	7.226	-0.990
M3	7.090	$10^{-6} - 10^{-5}$	9.130	-0.900
M4	5.298	$10^{-6} - 10^{-5}$	6.304	-1.100
M5	7.340	$10^{-6} - 10^{-5}$	7.773	-0.975
M6	5.423	$10^{-6} - 10^{-5}$	6.591	-1.067

4. CONCLUSION

This paper has reported the design, implementation, and metrological validation of a programmable high-voltage I–V measurement system tailored for dielectric breakdown kinetics. By integrating a ramp-controlled high-voltage source with a galvanically isolated acquisition architecture, the platform achieves an effective post-processed noise floor in the sub-nA range. This sensitivity enables the simultaneous resolution of low-level pre-breakdown leakage currents and catastrophic breakdown transients within a single experimental setup.

A central contribution of this work is the formalization of a compliance-aware extraction strategy. By mathematically distinguishing the intrinsic hard breakdown (HBD) onset—localized at the global maximum of $|dI/dV|$ —from the subsequent artifactual

compliance-limited region (CLR), the proposed framework removes a major source of parameter ambiguity often associated with general-purpose instrumentation. Experimental verification on 150 nm Si/SiO₂ MOS capacitors demonstrated the system’s robustness, yielding a mean breakdown field of 6.38 MV/cm with a coefficient of variation of 15.7% ($n = 6$), alongside high-linearity conduction diagnostics ($R^2 > 0.97$) supporting Fowler–Nordheim and Poole–Frenkel consistency checks.

Overall, the proposed hardware–software workflow provides a transparent, reproducible, and rigorous alternative to fixed-function commercial testers for reliability-oriented dielectric metrology. Future work will extend this architecture to support Time-Dependent Dielectric Breakdown (TDDB) protocols and to quantify the impact of laser-induced thermal modifications on gate-oxide integrity.

DECLARATION OF CONFLICTS OF INTEREST

The authors declare that this research has no conflicts of interest.

ACKNOWLEDGMENTS

The authors would like to thank colleagues and mentors for their valuable comments and support during this work. This research received no external funding.

REFERENCES

[1] J. Wang and X. Jiang, “Review and analysis of SiC MOSFETs’ ruggedness and reliability,” *IET Power Electronics*, vol. 13, no. 3, pp. 445–455, 2020.

[2] M. Wang, “A Review of Reliability in Gate-All-Around Nanosheet Devices,” *Micromachines*, vol. 15, no. 2, p. 269, 2024.

[3] D. A. Buchanan, “Scaling the gate dielectric: Materials, integration, and reliability,” *IBM Journal of Research and Development*, vol. 43, no. 3, pp. 245–264, 1999.

- [4] J. Shen et al., "The TDDB Characteristics of Ultra-Thin Gate Oxide MOS Capacitors under Constant Voltage Stress and Substrate Hot-Carrier Injection," *Advances in Condensed Matter Physics*, vol. 2018, no. 1, p. 5483756, 2018.
- [5] S. Cha, D. H. Kim, T. Liu, and L. S. Milor, "The die-to-die calibrated combined model of negative bias temperature instability and gate oxide breakdown from device to system," *Microelectronics Reliability*, vol. 55, no. 9, pp. 1404–1411, 2015.
- [6] F. Palumbo et al., "A Review on Dielectric Breakdown in Thin Dielectrics: Silicon Dioxide, High-k, and Layered Dielectrics," *Advanced Functional Materials*, vol. 30, no. 18, p. 1900657, 2020.
- [7] E. Rosenbaum and J. Wu, "Trap generation and breakdown processes in very thin gate oxides," *Microelectronics Reliability*, vol. 41, no. 5, pp. 625–632, 2001.
- [8] B. Steinwender, S. Einspieler, M. Glavanovics, and W. Elmenreich, "Distributed power semiconductor stress test & measurement architecture," in 2013 11th IEEE International Conference on Industrial Informatics (INDIN), 2013, pp. 129–134.
- [9] J. S. Suehle, B. Zhu, Y. Chen, and J. B. Bernstein, "Detailed study and projection of hard breakdown evolution in ultra-thin gate oxides," *Microelectronics Reliability*, vol. 45, no. 3, pp. 419–426, 2005.
- [10] T. S. Lee, N. J. Lee, H. Abbas, H. H. Lee, T.-S. Yoon, and C. J. Kang, "Compliance Current-Controlled Conducting Filament Formation in Tantalum Oxide-Based RRAM Devices with Different Top Electrodes," *ACS Applied Electronic Materials*, vol. 2, no. 4, pp. 1154–1161, 2020.
- [11] J. Yao, L. Zhong, D. Natelson, and J. M. Tour, "In situ imaging of the conducting filament in a silicon oxide resistive switch," *Scientific Reports*, vol. 2, no. 1, p. 242, 2012.
- [12] P. R. Shrestha et al., "Analysis and Control of RRAM Overshoot Current," *IEEE Transactions on Electron Devices*, vol. 65, no. 1, pp. 108–114, 2018.
- [13] J. Eastoe, G. M. Noah, D. Dutta, A. Rossi, J. D. Fletcher, and A. G. Saiz, "Method for Efficient Large-Scale Cryogenic Characterization of CMOS Technologies," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1–10, 2025.
- [14] M. S. Nazir, A. Pampori, Y. H. Zarkob, A. Kar, and Y. S. Chauhan, "Characterization and Modeling of I-V, C-V and Trapping behavior of SiC Power MOSFETs," in 2023 7th IEEE Electron Devices Technology & Manufacturing Conference (EDTM), 2023, pp. 1–3.
- [15] D. Xing et al., "3.3-kV SiC MOSFET Performance and Short-Circuit Capability," in 2020 IEEE Workshop on Wide Bandgap Power Devices and Applications in Asia (WiPDA Asia), 2020, pp. 1–6.
- [16] J. Schmitz, H. P. Tuinhout, H. J. Kretschmann, and P. H. Woerlee, "Comparison of soft-breakdown triggers for large-area capacitors under constant voltage stress," *IEEE Transactions on Device and Materials Reliability*, vol. 1, no. 3, pp. 150–157, 2001.
- [17] A. Padovani, P. L. Torracca, J. Strand, L. Larcher, and A. L. Shluger, "Dielectric breakdown of oxide films in electronic devices," *Nature Reviews Materials*, vol. 9, no. 9, pp. 607–627, 2024.
- [18] H. Satake and A. Toriumi, "SiO₂ dielectric breakdown mechanism studied by the post-breakdown resistance statistics," *Semiconductor Science and Technology*, vol. 15, no. 5, p. 471, 2000.
- [19] T. Pompl, C. Engel, H. Wurzer, and M. Kerber, "Soft breakdown and hard breakdown in ultra-thin oxides," *Microelectronics Reliability*, vol. 41, no. 4, pp. 543–551, 2001.
- [20] K. E. Falidas, M. B. Everding, A. E. Viegas, M. Czernohorsky, and J. Heitmann, "Automated reliability calculation of failure rate, lifetime extrapolation and prediction for embedded Metal-Insulator-Metal capacitors using an optimized Time-Dependent-Dielectric-Breakdown model," *Microelectronics Reliability*, vol. 150, p. 115191, 2023.

- [21] J. Qian et al., "Modeling of Charge-to-Breakdown with an Electron Trapping Model for Analysis of Thermal Gate Oxide Failure Mechanism in SiC Power MOSFETs," *Materials*, vol. 17, no. 7, p. 1455, 2024.
- [22] M. Lenzlinger and E. H. Snow, "Fowler-Nordheim Tunneling into Thermally Grown SiO₂," *Journal of Applied Physics*, vol. 40, pp. 278–283, 1969.
- [23] J. Frenkel, "On Pre-Breakdown Phenomena in Insulators and Electronic Semi-Conductors," *Physical Review*, vol. 54, no. 8, pp. 647–648, 1938.
- [24] A. Padovani, D. Z. Gao, A. L. Shluger, and L. Larcher, "A microscopic mechanism of dielectric breakdown in SiO₂ films: An insight from multi-scale modeling," *Journal of Applied Physics*, vol. 121, no. 15, p. 155101, 2017.
- [25] F. Palumbo, S. Lombardo, and M. Eizenberg, "Physical mechanism of progressive breakdown in gate oxides," *Journal of Applied Physics*, vol. 115, no. 22, p. 224101, 2014.



© 2026 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Thiết kế và đánh giá đặc tính đo lường của hệ thống đo I–V cao áp có khả năng lập trình phục vụ nghiên cứu động học đánh thủng điện môi

Huỳnh Công Tú^{1,*}, Nguyễn Văn Hào²

¹Phòng Đào tạo, Trường Đại học Quy Nhơn, Việt Nam

²Khoa Kỹ thuật và Công nghệ, Trường Đại học Quy Nhơn, Việt Nam

Ngày nhận bài: 06/02/2026; Ngày sửa bài: 22/03/2026

Ngày nhận đăng: 25/03/2026; Ngày xuất bản: 28/06/2026

TÓM TẮT

Phân biệt chính xác đánh thủng điện môi nội tại và các tín hiệu giả do giới hạn dòng của thiết bị là yêu cầu thiết yếu trong đo lường độ tin cậy bán dẫn. Bài báo trình bày thiết kế và đánh giá đo lường một hệ thống đo I–V cao áp (HV) khả trình phục vụ nghiên cứu động học đánh thủng điện môi. Hệ thống tích hợp nguồn cao áp dạng dốc và mô-đun thu thập dữ liệu hai kênh cách ly điện, đạt nhiều nền hiệu dụng sau xử lý dưới nA (sub-nA), đảm bảo tính toàn vẹn tín hiệu trên toàn dải động. Đóng góp trọng tâm là việc kiểm chứng chiến lược phát hiện có nhận biết giới hạn dòng. Bằng cách định vị điểm khởi phát đánh thủng cứng (HBD) nội tại ở cực đại toàn cục của độ dốc vi phân $|dI/dV|$ và loại trừ vùng giới hạn dòng (CLR) theo sau, phương pháp này triệt tiêu sự mơ hồ khi báo cáo tham số. Thực nghiệm trên tụ MOS Si/SiO₂ 150 nm chứng minh khả năng phân giải dòng rò phụ thuộc điện trường của hệ thống, nhận diện cơ chế dẫn điện hỗn hợp Fowler–Nordheim và Poole–Frenkel với độ tuyến tính cao ($R^2 > 0.97$) trong khoảng 2–7 MV/cm. Nghiên cứu ghi nhận trường đánh thủng trung bình 6.38 MV/cm, hệ số biến thiên 15.7% ($n = 6$), khẳng định đây là giải pháp minh bạch và có tính tái lập cao cho nghiên cứu độ bền điện môi và phân tích hỏng hóc.

Từ khóa: Thiết bị đo lường cao áp, đánh thủng điện môi, đánh thủng giới hạn dòng, độ tin cậy MOS, đánh giá đặc tính đo lường.

*Tác giả liên hệ chính.

Email: huynhcongtru@qnu.edu.vn

Recovery of monoclinic phase of VO₂ nanoparticles from oxidized samples: Assessment of reusability and optimization of re-annealing conditions

Ho Nguyen Hong Tham¹, Nguyen Thi Thanh Nhan¹, Nguyen Thi Tuong Vy¹,
Le Phuong Thao¹, Le Thi My Nhi², Hoang Thi Hang^{2,3}, Le Thi Ngoc Loan^{2,*}

¹Faculty of Education, Quy Nhon University, Vietnam

²Faculty of Natural Sciences, Quy Nhon University, Vietnam

³Quantum Solid State Physics, Department of Physics and Astronomy, KU Leuven, Belgium

Received: 27/02/2026; Revised: 24/03/2026;

Accepted: 25/03/2026; Published: 28/06/2026

ABSTRACT

Vanadium dioxide (VO₂) nanoparticles in the monoclinic (M) phase exhibit a reversible metal–insulator transition (MIT) at ~68 °C, underpinning thermochromic applications such as smart windows and thermal sensors. However, prolonged atmospheric exposure causes progressive surface oxidation that converts V⁴⁺ to V⁵⁺, leading to phase degradation and loss of thermochromic function. This study demonstrates that oxidized VO₂ nanopowders can be fully recovered by thermal re-annealing in an inert argon atmosphere. Starting from VO₂(M) nanoparticles synthesized via a hydrothermal route, controlled oxidation under ambient conditions produced a degraded sample rich in V₂O₅. Systematic re-annealing at 450–600 °C for 2–3 h was investigated to identify the optimal recovery conditions. XRD, Raman spectroscopy, and UV-Vis-NIR transmittance measurements confirm complete phase restoration and recovery of thermochromic switching performance equivalent to that of the as-synthesized material, establishing a practical strategy for reusing degraded VO₂ nanopowder stocks.

Keywords: Vanadium dioxide, smart window, transition material, thermochromic recovery.

1. INTRODUCTION

Vanadium dioxide (VO₂) is a polymorphic transition-metal oxide capable of adopting several crystal structures-monoclinic (M), VO₂(A), VO₂(B), and VO₂(D) - despite sharing the same chemical composition [1] - [3]. The differences among these polymorphs arise from distinct VO₆ octahedral arrangements within the crystal lattice, giving each phase unique structural and electronic properties (Figure1).

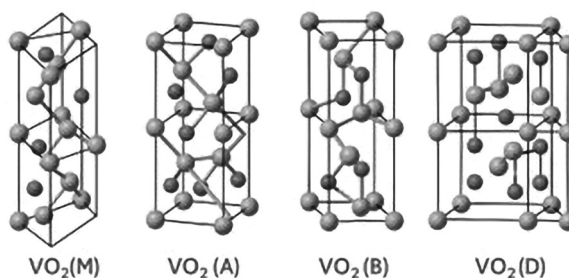


Figure 1. Crystal lattice structures of the four VO₂ polymorphs: VO₂(M), VO₂(A), VO₂(B), and VO₂(D). Orange: V atoms; red: O atoms.

*Corresponding author:

Email: lethingocloan@qnu.edu.vn

Among these, monoclinic $\text{VO}_2(\text{M})$ is thermodynamically stable at room temperature and is of exceptional importance owing to its reversible semiconductor-metal phase transition near 340 K ($\sim 68^\circ\text{C}$), first reported by Morin [4]. In the monoclinic state, vanadium atoms form paired dimers along the c-axis, opening a band gap and yielding semiconducting behavior. Upon heating, dimerization disappears, the density of states near the Fermi level increases sharply, and metallic conductivity is recovered [4] - [5]. This structural reorganization also reverses the optical response: below the transition temperature, VO_2 transmits near-infrared (NIR) radiation; above it, the metallic phase strongly reflects NIR light [6] - [7]. This reversible electro-optical switching underpins thermochromic smart windows, thermal sensors, and optical limiters.

A variety of synthesis strategies have been developed for VO_2 , including physical vapor deposition, magnetron sputtering, sol-gel, and hydrothermal methods [8] - [9]. Hydrothermal synthesis is particularly attractive because it permits direct control of crystal phase, particle size, and morphology through adjustment of temperature, pH, and redox atmosphere [10] - [11]. Typically, a V^{5+} precursor such as V_2O_5 is reduced to V^{4+} under hydrothermal conditions, yielding metastable $\text{VO}_2(\text{B})$ or $\text{VO}_2(\text{A})$ that is subsequently annealed to phase-pure $\text{VO}_2(\text{M})$ in an inert atmosphere [10] - [12].

Despite extensive research, the storage stability of synthesized VO_2 nanopowders remains a practical challenge. Even in nominally sealed environments, prolonged ambient exposure causes gradual surface oxidation of V^{4+} to V^{5+} , forming V_2O_5 -rich surface layers that broaden the MIT, reduce thermochromic contrast, and impair switching performance [1] - [13]. The nanoscale morphology of these particles accelerates this degradation relative to bulk material owing to high surface-to-volume ratios. Recovery of degraded VO_2 nanopowders by thermal re-annealing in inert or reducing

atmospheres has been proposed [12] - [14], yet systematic optimization of re-annealing conditions for nanopowder specimens remains sparse in the literature.

In the present work, we synthesized $\text{VO}_2(\text{M})$ nanoparticles by a hydrothermal route, intentionally oxidized them under ambient conditions, and conducted systematic re-annealing in argon across temperatures of $450\text{--}600^\circ\text{C}$ and durations of 2–3 h. XRD, Raman spectroscopy, and UV-Vis-NIR transmittance measurements are used to track phase evolution and verify restoration of thermochromic function. Our results identify 500°C for 2 h as the optimal condition, demonstrating a straightforward, cost-effective protocol for reusing degraded VO_2 nanoparticle batches.

2. EXPERIMENTAL SECTION

2.1. Materials

Vanadium pentoxide (V_2O_5 , $\geq 98\%$), concentrated sulfuric acid (H_2SO_4 , 98%), hydrazine hydrate ($\text{N}_2\text{H}_4\cdot\text{H}_2\text{O}$, 50–60%), sodium hydroxide (NaOH , $\geq 95\%$), polyvinylpyrrolidone (PVP, MW = 40,000 g/mol), and absolute ethanol (99.6%) were used as received. Deionized (DI) water was used throughout.

2.2. Synthesis of $\text{VO}_2(\text{M})$ nanoparticles

$\text{VO}_2(\text{M})$ nanoparticles were prepared via a four-step hydrothermal procedure illustrated in Figure 2.

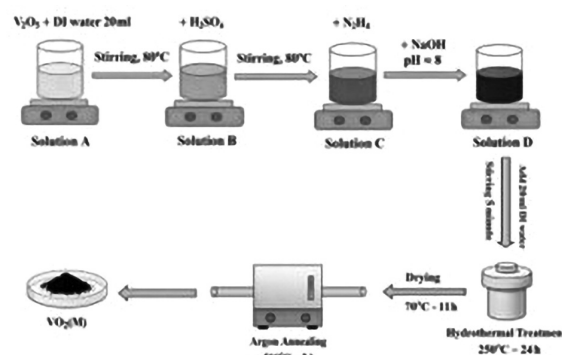


Figure 2. Schematic of the hydrothermal synthesis route for $\text{VO}_2(\text{M})$ nanoparticles, from V_2O_5 precursor dissolution to argon annealing.

Step 1: Precursor preparation: 0.9 g of V_2O_5 was dispersed in 20 mL DI water under stirring at 60 °C (Solution A, yellow suspension). Then 1.5 mL of H_2SO_4 was added dropwise at 80 °C to give a deep-yellow solution (Solution B).

Step 2: Reduction and pH adjustment: 0.5 mL of N_2H_4 was introduced dropwise, raising the temperature to 90–100 °C and yielding a blue solution (Solution C, pH $\approx$ 0.8-1). NaOH (2.3 g) was then added to adjust pH to $\sim$ 8, producing a dark-brown suspension (Solution D). The color evolution is shown in Figure 3.

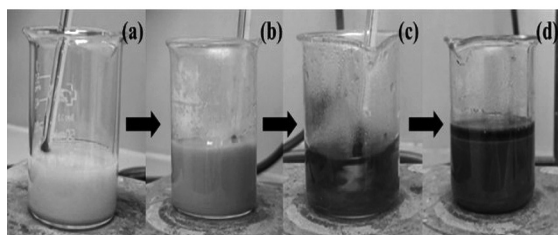


Figure 3. Optical photographs of the precursor at each synthesis step: (a) yellow V_2O_5 suspension; (b) deep-yellow Solution B; (c) blue Solution C after N_2H_4 reduction; (d) dark-brown Solution D after pH adjustment.

Step 3: Hydrothermal treatment: the precursor was transferred to a 25 mL Teflon-lined stainless-steel autoclave and heated at 250 °C for 24 h (5 °C min^{-1}).

Step 4: Post-treatment and annealing: the black precipitate was washed twice with DI water and once with ethanol, dried at 70 °C for 11 h, then annealed at 500 °C for 2 h in flowing Ar (10 °C min^{-1}) to obtain phase-pure $VO_2(M)$.

2.3. Controlled oxidation

Synthesized $VO_2(M)$ powders were stored in an open beaker under ambient laboratory conditions ($\sim$ 25 °C, relative humidity 60–70%) for a controlled period (1 week) to simulate real-world aging. Color changes were recorded photographically.

2.4. Recovery by re-annealing

Oxidized powders were re-annealed in a tube furnace under continuous Ar flow. Stage 1 (temperature optimization): re-annealing at 450,

500, 550, and 600 °C for 2 h. Stage 2 (duration optimization): re-annealing at 500 °C for 2 h and 3 h. All heating ramps were 10 °C min^{-1} .

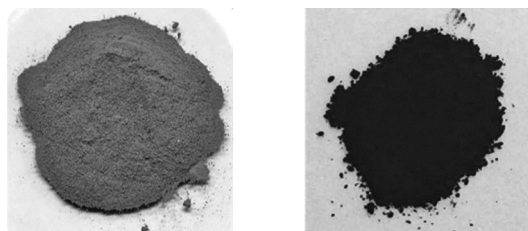
2.5. Characterization

The crystal structure of the samples was determined by X-ray diffraction using a D2 Phaser diffractometer (Bruker, Rheinstetten, Germany) equipped with a Cu $K\alpha$ radiation source ($\lambda = 1.54 \text{ \AA}$). Raman spectra were acquired using a Horiba XPlora Plus spectrometer (Horiba, Villeneuve d'Ascq, France) with a 532 nm laser excitation. The surface morphology was examined using a JSM-IT200 scanning electron microscope (JEOL, Japan). Thermochromic performance was evaluated by UV-Vis-NIR transmittance spectroscopy over a range of 300 – 2000 nm using a JASCO V-770 spectrophotometer (JASCO, Hachioji, Japan) equipped with a temperature-controlled stage during heating and cooling cycles between 30 and 100 °C.

3. RESULTS AND DISCUSSION

3.1. As-synthesized $VO_2(M)$ nanoparticles

The hydrothermal synthesis and Ar annealing steps produced a jet-black powder (Figure 4, right), in stark visual contrast to the yellow V_2O_5 precursor (Figure 4, left). The color change reflects successful $V^{5+} \rightarrow V^{4+}$ reduction and $VO_2(M)$ phase formation.



Photograph of yellow V_2O_5 powder

Photograph of black $VO_2(M)$ powder

Figure 4. Optical photographs of the yellow V_2O_5 precursor powder (left) and the as-synthesized black $VO_2(M)$ nanoparticle powder (right).

The UV-Vis-NIR transmittance spectra of original sample (Figure 5) show a pronounced divergence between the 30 °C (semiconducting) and 100 °C (metallic) curves in the NIR region (800 – 2000 nm), confirming active

thermochromic switching through the MIT at $\sim 68^\circ\text{C}$ [15]. This large NIR transmittance contrast is the key functional signature of high-quality $\text{VO}_2(\text{M})$.

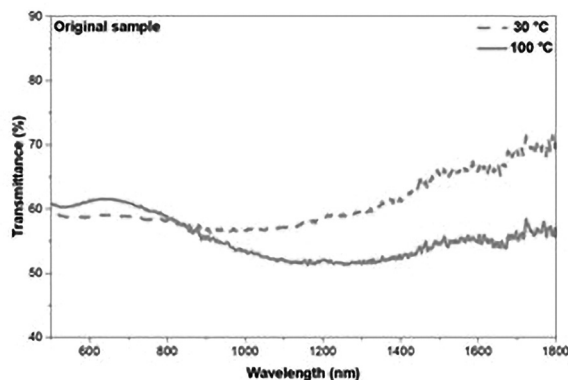


Figure 5. UV-Vis-NIR transmittance spectra of as-synthesized $\text{VO}_2(\text{M})$ film (Original sample) at 30°C (dashed line) and 100°C (solid line), showing clear thermochromic NIR switching.

3.2. Oxidation behavior under ambient conditions

When $\text{VO}_2(\text{M})$ powder was stored in an open beaker under ambient air (Figure 6), a progressive color transformation was observed over the course of days to weeks: black $\rightarrow$ bluish-black $\rightarrow$ moss green $\rightarrow$ dark moss green (Figure 7a). This sequence mirrors the gradual oxidation of V^{4+} to V^{5+} and accumulation of V_2O_5 -like surface species. The SEM image in Figure 7b reveals that re-annealing at 500°C for 3 hours resulted in a mixed morphology, consisting predominantly of nanoparticles with a partial formation of nanorods.

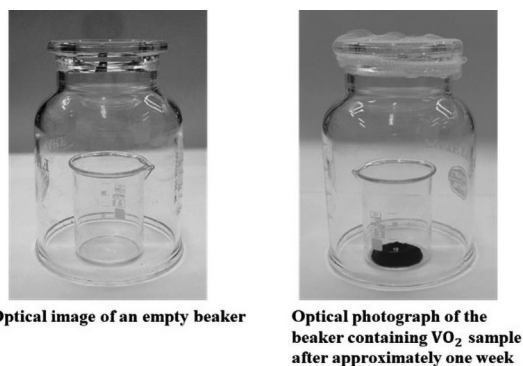


Figure 6. Storage experiment: empty reference beaker (left) and VO_2 powder sample after $\sim$ one week of ambient air exposure (right).

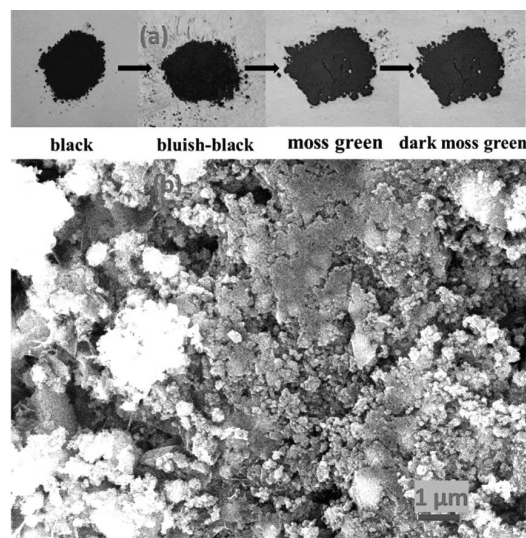


Figure 7. Time-dependent color evolution of VO_2 powder during ambient oxidation: black $\rightarrow$ bluish-black $\rightarrow$ moss green $\rightarrow$ dark moss green (a) and SEM image of the recovery sample annealed at 500°C in 3h.

XRD analysis of the oxidized sample (Figure 8) reveals emergence of V_2O_5 reflections alongside diminished $\text{VO}_2(\text{M})$ peaks, confirming phase transformation rather than mere surface contamination [16]. The original sample (orange) displays only sharp $\text{VO}_2(\text{M})$ reflections, while the oxidized sample (cyan) shows clear additional peaks at positions characteristic of V_2O_5 at 2 theta positions of 26.2° , 31° , 32.4° , 41° , 52° and 62° . Nevertheless, the persistence of lower-intensity peaks at 27.9° , 37° , 55° , 65° and 70° indicates a residual $\text{VO}_2(\text{M})$ component. Furthermore, the diffractogram suggests the secondary presence of the $\text{VO}_2(\text{B})$ phase as an impurity.

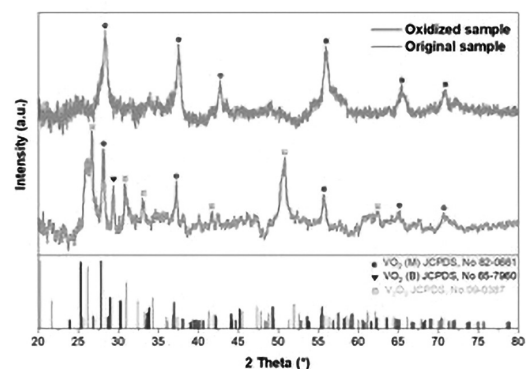


Figure 8. XRD patterns of the original $\text{VO}_2(\text{M})$ (orange) and oxidized sample (cyan) compared with reference stick patterns for $\text{VO}_2(\text{B})$ (blue), V_2O_5 (green) and $\text{VO}_2(\text{M})$ (red). Dashed lines mark key $\text{VO}_2(\text{M})$ peak positions.

The optical consequences are decisive: Figure 9 shows UV-Vis-NIR spectra of the oxidized sample at 30 and 100 °C. The near-identical curves confirm complete abolition of thermochromic switching [17]. This finding is consistent with reports that even partial V^{5+} contamination can substantially suppress MIT sharpness, particularly in nanoparticle systems with high surface-to-volume ratios.

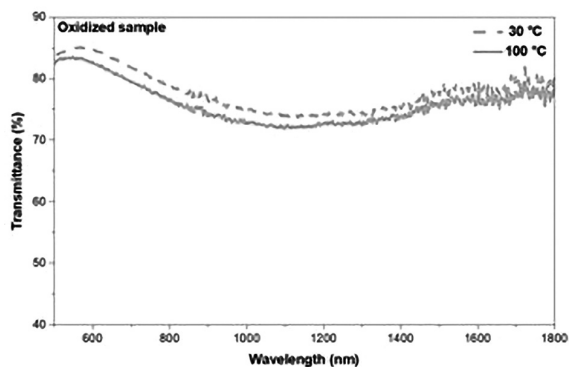


Figure 9. UV-Vis-NIR transmittance spectra of the oxidized sample at 30 °C (dashed line) and 100 °C (solid line). Near-identical curves confirm complete suppression of the thermochromic MIT.

Raman spectroscopy further corroborates the phase loss (Figure 10). The oxidized sample displays a featureless, monotonically rising background with no discernible $VO_2(M)$ phonon modes, consistent with a V_2O_5 -dominated surface that lacks the characteristic V-V dimerization modes of the monoclinic lattice.

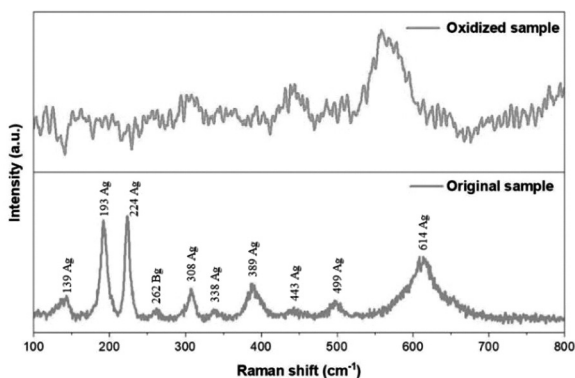


Figure 10. Raman spectrum of the oxidized VO_2 sample versus synthesized one. The featureless background and absence of $VO_2(M)$ modes confirm complete disruption of the monoclinic lattice.

3.3. Recovery of $VO_2(M)$ by thermal re-annealing

Re-annealing in Ar establishes a low oxygen partial pressure environment that suppresses re-oxidation of intermediate phases, enabling the thermally driven, stepwise lattice oxygen release sequence $V_2O_5 \rightarrow V_3O_5 \rightarrow V_4O_9 \rightarrow V_6O_{13} \rightarrow VO_2$ to proceed without introducing unwanted dopants or secondary chemical reactions [18]. The phase evolution was tracked visually (Figure 11) and by XRD (Figure 12).

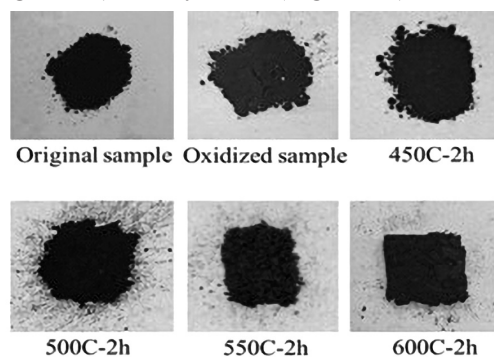


Figure 11. Optical photographs comparing original $VO_2(M)$, oxidized sample, and re-annealed samples at 450, 500, 550, and 600 °C (2 h each). The 500 – 600 °C samples recover the characteristic black color.

In Figure 12, re-annealing at 450 °C (black trace) yields broad, poorly resolved peaks with poor agreement to the $VO_2(M)$ reference - indicating incomplete reduction. At 500 °C (pink trace), a clean pattern matching the $VO_2(M)$ reference (red sticks) is obtained with no V_2O_5 or $VO_2(B)$ impurities. At 550 °C and 600 °C, additional peaks inconsistent with $VO_2(M)$ appear, attributable to minor $VO_2(B)$ or over-reduced secondary phases. [18]

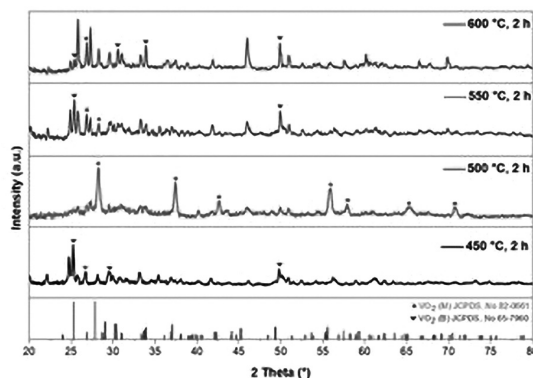


Figure 12. XRD patterns of samples re-annealed at 450 °C (black), 500 °C (pink), 550 °C (dark green), and 600 °C (purple), with $VO_2(B)$ (blue) and $VO_2(M)$ (red) reference patterns.

Raman spectroscopy of the 500 °C for 2 h sample, Figure 13 confirms the reappearance of sharp $\text{VO}_2(\text{M})$ phonon modes at ~ 224 and ~ 614 cm^{-1} , with additional features at ~ 139 , ~ 193 , ~ 262 , ~ 308 , ~ 338 , ~ 389 , ~ 443 , and ~ 499 cm^{-1} - all characteristic of the monoclinic lattice and V-V dimerization. No V_2O_5 Raman bands are visible, providing unambiguous spectroscopic confirmation of complete phase recovery.

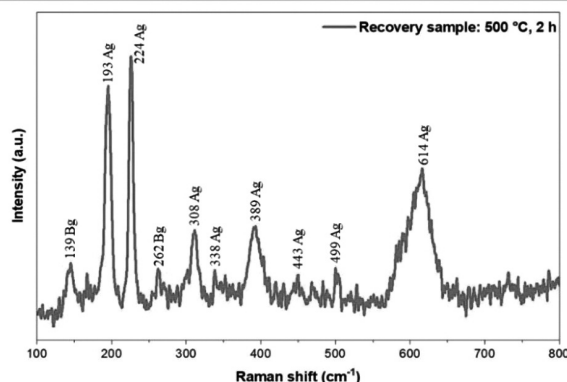


Figure 13. Raman spectrum of recovered $\text{VO}_2(\text{M})$ (500 °C for 2 h). Characteristic $\text{VO}_2(\text{M})$ modes at ~ 224 and ~ 614 cm^{-1} are clearly restored.

3.4. Optimization of re-annealing duration

Comparing re-annealing at 500 °C for 2 h versus 3 h reveals a critical narrow processing window (Figure 14). The 2 h sample (left, pink) yields a phase-pure $\text{VO}_2(\text{M})$ XRD pattern with no extraneous reflections. Extending the duration to 3 h (right, purple) introduces $\text{VO}_2(\text{B})$ peaks

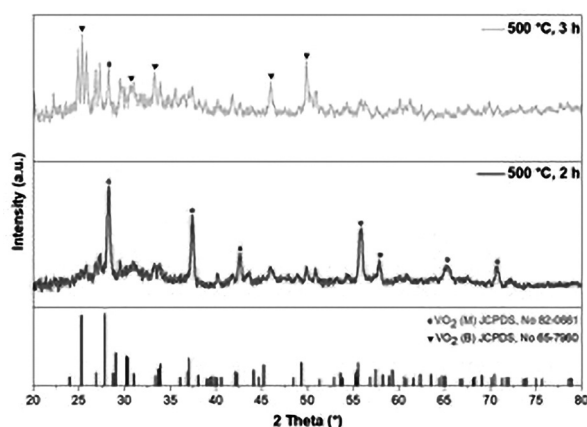


Figure 14. XRD comparison of re-annealing at 500 °C for 2 h (pink, left) and 3 h (lime green, right). The 3 h sample shows $\text{VO}_2(\text{B})$ impurity peaks (dashed lines), confirming 2 h as optimal.

(highlighted by dashed lines), indicating the onset of over-reduction or phase instability under prolonged thermal treatment. Therefore, 2 h is the optimal duration-sufficient for complete reduction yet short enough to avoid unwanted phase evolution.

3.5. Verification of thermochromic performance restoration

Figure 15 presents UV-Vis-NIR data for the optimally recovered sample (500 °C for 2 h). The spectra show a pronounced divergence between the 30 and 100 °C curves in the NIR region, fully restoring the thermochromic switching behavior of the original material. The recovered NIR contrast and switching temperature ($\sim 68^\circ\text{C}$) are quantitatively comparable to those of the original as-synthesized sample, with transmittance differences at 1800 nm of approximately 15% and 12.5% ($\sim 83.33\%$) for the original and re-annealed samples, respectively.

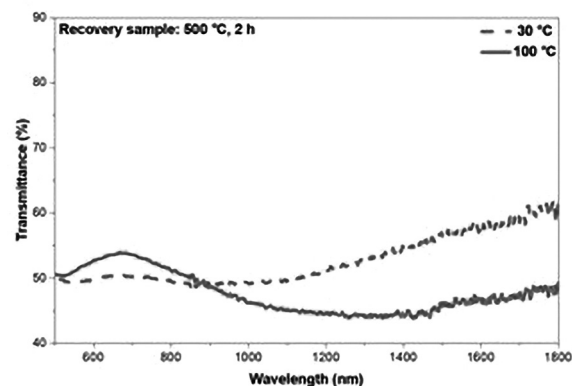


Figure 15. Characterization of recovered $\text{VO}_2(\text{M})$ (500 °C for 2 h): UV-Vis-NIR transmittance spectra at 30 °C (dashed line) and 100 °C (solid line) confirming full thermochromic recovery.

The complete restoration of the MIT - a property exquisitely sensitive to stoichiometry, lattice order, and surface chemistry - confirms that re-annealing at the optimal condition drives thorough phase recovery, extending from the nanoparticle surface into the bulk. This result is consistent with previous reports on thin-film VO_2 systems where post-synthesis reduction treatments restored MIT sharpness and thermochromic contrast [2], [14], [19], [20]. The practical significance is considerable:

VO₂ nanoparticle synthesis requires multi-step procedures, specialized equipment, and expensive precursors. Rejuvenating oxidized stocks via a single 2-hour Ar annealing step - identical to the original synthesis annealing - eliminates the need to discard and resynthesize degraded material, reducing cost, waste, and synthesis time.

4. CONCLUSION

This study has systematically demonstrated the feasibility of recovering phase-pure monoclinic VO₂(M) nanoparticles from ambient-oxidized nanopowders through thermal re-annealing in argon. Progressive atmospheric oxidation converted V⁴⁺ to V⁵⁺, yielding V₂O₅ - rich phases that completely suppressed the MIT, as evidenced by XRD, Raman spectroscopy, and UV-Vis-NIR transmittance. Systematic re-annealing experiments across 450 – 600 °C and 2-3 h identified 500 °C for 2 h in Ar as the optimal recovery condition. Under these conditions, XRD and Raman spectroscopy confirmed complete VO₂(M) phase restoration, and UV-Vis-NIR measurements verified quantitative recovery of thermochromic performance equivalent to the original as-synthesized nanoparticles. These results provide a practical, scalable protocol for extending the operational lifetime of VO₂ nanopowder batches, with direct implications for reducing material waste and cost in thermochromic device fabrication.

DECLARATION OF CONFLICTS OF INTEREST

The authors declare that this research has no conflicts of interest.

ACKNOWLEDGMENTS

The authors gratefully acknowledge VLIR-UOS (Belgium) and Quy Nhon University for supporting this work under project codes VN2022IUC044A101 and S2025.1152.102. The authors also thank Hoang Nhat Hieu and Le Thi Thanh Lieu for SEM and XRD measurements.

REFERENCES

- [1] P. Hu, T. D. Vu, M. Li, S. Wang, Y. Ke, X. Zeng, L. Mai, and Y. Long, "Vanadium Oxide: Phase Diagrams, Structures, Synthesis, and Applications," *Chemical Reviews*, vol. 123, no. 8, pp. 4353-4415, 2023.
- [2] J. Liu, Q. Li, T. Wang, D. Yu, and Y. Li, "Thermochromic VO₂ thin film prepared by post annealing treatment of V₂O₅ thin film," *Advanced Materials Research*, vol. 79-82, pp. 747-750, 2009.
- [3] S. Liang, J. Shi, H. Zheng, B. Liu, Q. Liu, and C. Li, "A facile pathway to prepare VO₂ and V₂O₃ powders via carbothermal reduction," *Journal of Solid State Chemistry*, vol. 265, pp. 299-305, 2018.
- [4] D. K. Manousou, E. Gagaoudakis, G. Michail, V. Binas, G. Kiriakidis, V. Diakonov, and M. Kompitsas, "VO₂ thin films fabricated by reduction of thermally evaporated V₂O₅ under N₂ flow," *Materials Letters*, vol. 299, p. 130086, 2021.
- [5] Z. Yang, C. Ko, and S. Ramanathan, "Oxide electronics utilizing ultrafast metal-insulator transitions," *Annual Review of Materials Research*, vol. 41, pp. 337-367, 2011.
- [6] C. G. Granqvist, "Electrochromics for smart windows: oxide-based thin films and devices," *Thin Solid Films*, vol. 564, pp. 1-38, 2014.
- [7] M. M. Qazilbash et al., "Mott transition in VO₂ revealed by infrared spectroscopy and nano-imaging," *Science*, vol. 318, no. 5857, pp. 1750-1753, 2007.
- [8] A. Kumar, A. Kumar, A. Kandasami, and V. R. Singh, "A comprehensive review on synthesis, phase transition, and applications of VO₂," *Journal of Superconductivity and Novel Magnetism*, vol. 37, no. 3, pp. 475-498, 2024.
- [9] J. Chen, X. Tian, L. Ye, P. Lu, Q. Luo, J. Yan, and D. Zhang, "High-performance VO₂ thin film prepared via room-temperature-deposition and air-annealing treatment," *Journal of Electronic Materials*, vol. 54, no. 9, pp. 7935-7941, 2025.

- [10] W. Burkhardt et al., "Tungsten and fluorine co-doping of VO₂ films," *Thin Solid Films*, vol. 402, no. 1-2, pp. 228-234, 2002.
- [11] Y. Gao, H. Luo, Z. Zhang, L. Kang, Z. Chen, J. Du, M. Kanehira, and C. Cao, "Nanoceramic VO₂ thermochromic smart glass: a review on progress in solution processing," *Nano Energy*, vol. 1, no. 2, pp. 221-246, 2012.
- [12] M. Soltani, M. Chaker, E. Haddad, R. V. Kruzelecky, and J. Margot, "Effects of Ti-W codoping on the optical and electrical switching of vanadium dioxide thin films grown by a reactive pulsed laser deposition," *Applied Physics Letters*, vol. 85, no. 11, pp. 1958-1960, 2004.
- [13] Y. Gao et al., "Enhanced chemical stability of VO₂ nanoparticles by the formation of SiO₂/VO₂ core/shell structures and the application to transparent and flexible VO₂-based composite foils with excellent thermochromic properties for solar heat control," *Energy & Environmental Science*, vol. 5, pp. 6104-6110, 2012.
- [14] Z. Zhang, Y. Gao, Z. Chen, L. Kang, H. Luo, "Thermochromic VO₂ Thin Films: Solution-Based Processing, Improved Optical Properties, and Lowered Phase Transformation Temperature," *Langmuir* vol. 26, pp. 10738-10744, 2010.
- [15] Z. Liang et al., "Long-term stability and efficiency of VO₂ nanostructure-based thermochromic smart windows," *ACS Applied Nano Materials*, vol. 7, no. 18, pp. 21683-21691, 2024.
- [16] Y. Yang et al., "Suppression of photoinduced surface oxidation of vanadium dioxide nanostructures by blocking oxygen adsorption," *ACS Omega*, vol. 4, no. 18, pp. 17735-17740, 2019.
- [17] J. Cao et al., "Strain engineering and one-dimensional organization of metal-insulator domains in single-crystal vanadium dioxide beams," *Nature Nanotechnology*, vol. 4, no. 11, pp. 732-737, 2009.
- [18] Y. Ningyi, L. Jinhua, L. Chenglu, "Valence reduction process from sol-gel V₂O₅ to VO₂ thin films," *Applied Surface Science*, vol. 191, no. 1-4, pp. 176-180, 2002.
- [19] M. Kumar, S. Rani, J. P. Singh, K. H. Chae, Y. Kim, J. Park, and H. H. Lee, "Structural phase control and thermochromic modulation of VO₂ thin films by post thermal annealing," *Applied Surface Science*, vol. 529, pp. 147093, 2020.
- [20] S. Wang, M. Liu, L. Kong, Y. Long, X. Jiang, and A. Yu, "Recent progress in VO₂ smart coatings: strategies to improve the thermochromic properties," *Progress in Materials Science*, vol. 81, pp. 1-54, 2016.



© 2026 by the authors. This Open Access Article is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Phục hồi pha tinh thể đơn tà của hạt nano VO_2 từ mẫu bị oxy hóa: đánh giá khả năng tái sử dụng và tối ưu điều kiện nung phục hồi

Hồ Nguyễn Hồng Thắm¹, Nguyễn Thị Thanh Nhân¹, Nguyễn Thị Tường Vy¹,
Lê Phương Thảo¹, Lê Thị Mỹ Nhi², Hoàng Thị Hằng^{2,3}, Lê Thị Ngọc Loan^{2,*}

¹Khoa Sư phạm, Trường Đại học Quy Nhơn, Việt Nam

²Khoa Khoa học Tự nhiên, Trường Đại học Quy Nhơn, Việt Nam

³Khoa Vật lý và Thiên văn học, Đại học KU Leuven, Vương quốc Bỉ

Ngày nhận bài: 27/02/2026; Ngày sửa bài: 24/03/2026;

Ngày nhận đăng: 25/03/2026; Ngày xuất bản: 28/06/2026

TÓM TẮT

Các hạt nano vanadium dioxide (VO_2) pha đơn tà (M) có khả năng chuyển đổi pha giữa kim loại - bán dẫn (MIT) thuận nghịch tại $\sim 68^\circ\text{C}$ đã mở ra các ứng dụng nhiệt sắc như cửa sổ thông minh và cảm biến nhiệt. Tuy nhiên, việc phơi nhiễm trong môi trường khí quyển gây ra hiện tượng oxy hóa bề mặt, chuyển đổi V^{4+} thành V^{5+} , dẫn đến suy giảm chất lượng pha tinh thể hoặc thay đổi thành phần pha làm mất chức năng nhiệt sắc. Trong nghiên cứu này, các hạt nano VO_2 bị oxy hóa cho thấy khả năng được phục hồi hoàn toàn bằng cách nung lại trong môi trường khí trơ. Từ các hạt nano $\text{VO}_2(\text{M})$ được tổng hợp bằng phương pháp thủy nhiệt, quá trình oxy hóa được thực hiện trong điều kiện môi trường không khí tạo ra mẫu có chất lượng tinh thể bị suy giảm và chứa thành phần V_2O_5 . Để khảo sát khả năng hồi phục pha tinh thể $\text{VO}_2(\text{M})$, quá trình nung lại được khảo sát một cách hệ thống từ $450 - 600^\circ\text{C}$ trong 2-3 giờ, kết quả cho thấy nung ở điều kiện tại nhiệt độ 500°C trong 2 giờ là điều kiện phục hồi tối ưu. Phép phân tích XRD, phổ Raman và đo phổ truyền qua UV-Vis-NIR cho thấy $\text{VO}_2(\text{M})$ khôi phục hoàn toàn pha và hiệu suất nhiệt sắc gần tương đương với vật liệu ban đầu.

Từ khóa: Vanadium dioxide, cửa sổ thông minh, vật liệu chuyển pha, hồi phục tính chất nhiệt sắc.

*Tác giả liên hệ chính.

Email: lethingocloan@qnu.edu.vn

MỤC LỤC

1.	Đánh giá hiệu năng một cài đặt thuật toán KNN bằng Rust-Webassembly Trần Hoàng Việt, Đoàn Thị Thu Cúc, Võ Thị Mỹ, Quảng Thị Ngọc Bích	5
2.	Tối ưu hóa hiệu suất phân hủy congo red trên xúc tác đất diatomite sử dụng thiết kế Box-Behnken Phạm Đình Dũ.....	19
3.	Thiết kế bộ quan sát $\mathcal{H}_\infty$ theo phương pháp chia lưới cho hệ Lorenz 63 sử dụng mô hình dạng phi tuyến với tham số thay đổi Võ Khánh Thoại, Phạm Thanh Phong, Lê Tiến Dũng	29
4.	Mô hình AI trên FPGA: Mô hình CNN gọn nhẹ thông lượng cao và công suất thấp cho bài toán nhận dạng chữ số Phạm Văn Khoa, La Công Lộc, Nguyễn Đức Trọng, Huỳnh Minh Quý, Hồ Gia Huyền, Võ Quang Huy, Phạm Bá Lộc, Huỳnh Thị Thu Hiền	43
5.	Nghiên cứu phân tích và đánh giá mức độ phơi nhiễm, đặc trưng tích lũy hydrocacbon thơm đa vòng (PAHs) trong mẫu tóc và móng tay người Nguyễn Đức Hiếu, Hoàng Quốc Anh, Chu Thị Huệ, Hoàng Thị Hương Huế, Nguyễn Thị Duyên, Nguyễn Xuân Thảo, Phạm Quỳnh Thiêm, Đỗ Văn Doanh, Trần Đức Chung, Nguyễn Thị Ánh Hoàng.....	55
6.	Nâng cao năng lực giảng viên giáo dục đại học trong kỷ nguyên trí tuệ nhân tạo tạo sinh: Tích hợp khả năng giải thích, khả năng tranh luận và thực hành phản tư Nguyễn Trần Diễm Quỳnh, Trương Thị Cẩm Mai.....	65
7.	Phân tích nhiệt của cáp điện hạ áp có xét đến ảnh hưởng của sóng hài dòng điện Nguyễn Kim Ánh, Huỳnh Đại, Lê Thị Thu, Lê Huy Hoàng, Phùng Bá Tú.....	85
8.	Ảnh hưởng của độ chín và điều kiện bảo quản đến thành phần polyphenol của trái giác (<i>Cayratia trifolia</i>) từ các vùng địa lý khác nhau ở Việt Nam Đoàn Thị Kiều Tiên, Trần Diễm Phượng, Huỳnh Thị Ngọc Mi, Đinh Hoàng Lan Chi, Nguyễn Ngọc Thạnh, Huỳnh Xuân Phong	97
9.	Thiết kế tối ưu bền vững cho thiết kế đường dây truyền tải điện trên không Phạm Văn Phát, Trần Ngọc Thạnh, Lê Tuấn Hộ.....	109

10. Nghiên cứu tối ưu chùm tia trong hệ thống truyền không dây có sự trợ giúp của bề mặt phản xạ thông minh
Nguyễn Đỗ Dũng, Nguyễn Văn Hào, Đào Minh Hưng 127
11. Chia tách đường cong Bézier bậc ba từng mảnh và các hằng số tương đương cho một vài chuẩn
Hoàng Văn Đức 139
12. Thiết kế và phát triển hệ thống tay máy phân loại các đối tượng dạng phương tiện dựa trên thị giác máy tính
Lê Đình Minh, Nguyễn Xuân Lộc, Đỗ Văn Cần, Trần Xuân Khoa 151
13. Thiết kế và đánh giá đặc tính đo lường của hệ thống đo I–V cao áp có khả năng lập trình phục vụ nghiên cứu động học đánh thủng điện môi
Huỳnh Công Tú, Nguyễn Văn Hào 177
14. Recovery of monoclinic phase of VO₂ nanoparticles from oxidized samples: Assessment of reusability and optimization of re-annealing conditions
Hồ Nguyễn Hồng Thắm, Nguyễn Thị Thanh Nhân, Nguyễn Thị Tường Vy, Lê Phương Thảo, Lê Thị Mỹ Nhi, Hoàng Thị Hằng, Lê Thị Ngọc Loan 189