



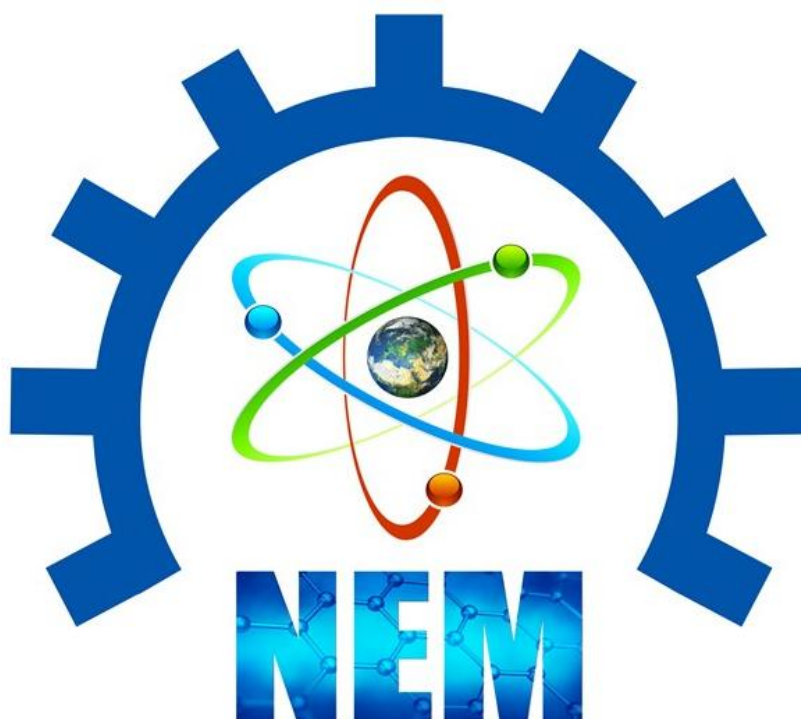
**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

**ISBN: 978-605-68918-2-3**

# **NEM 2026**

## **FULL TEXT BOOK**



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**

**Sep 10-12, 2026, Famagusta, Northern Cyprus**



6<sup>th</sup> International **N**atural Science, **E**ngineering and **M**aterial Technologies Conference  
Sep 10-12, 2026, Famagusta, Northern Cyprus

---

## FOREWORD

It is a pleasure for us to offer you this Book of Abstract for the 6<sup>th</sup> International **N**atural Science, **E**ngineering and **M**aterial Technologies Conference (NEM 2026). Our goal was to create a platform that introduces the newest results on internationally recognized experts to local students and colleagues and simultaneously to display relevant Turkish achievements to the world. The positive feedback of the community encouraged us to proceed and transform a single event into a conference series. Now, NEM 2026 is honored by the presence of over 102 colleagues from various countries. We kept the original NEM 2026 concept and accepted contributions from all fields of materials science and technology to promote multidisciplinary discussions. The main points of the conference emerged spontaneously from the submitted abstracts: energy applications, advanced materials, electronic and optoelectronic devices, organic electronic materials, chemistry, physics, environmental science, medical science, applied and engineering science, computer simulation of organic structures, biomedical applications and advanced characterization techniques of nanostructured materials. Further fields of interest include e.g. new advanced and functional materials, advanced-functional composites, biomaterials, smart materials, dielectric materials, optical materials, magnetic materials, organic semiconductors, inorganic semiconductors, electronic materials, graphene, and more.

Therefore, we hope that getting first-hand access to new results, establishing new connections and enjoying the Famagusta, Northern Cyprus ambience will make you feel that your efforts were spent well in NEM 2026.

Our warmest thanks go to all invited speakers, authors, and contributors of NEM 2026 for accepting our invitation, visiting Northern Cyprus and attending NEM 2026 as a medium for communicating your research results.

We hope that you will enjoy the conference and look forward to meeting you again in the forthcoming **NEM 2027** event.

Best regards,  
Chairmen's of Conference

Assoc. Prof. Burhan COŞKUN

Prof. Dr. Huriye İCİL



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

**Editors:**

**Prof. Dr. Burhan COŞKUN**

**Prof. Dr. Huriye İCİL**

**Published, September-2026**

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned. Nothing from this publication may be translated, reproduced, stored in a computerized system or published in any form or in any manner, including, but not limited to electronic, mechanical, reprographic or photographic, without prior written permission from the publisher.

[nem2026.klu.edu.tr](http://nem2026.klu.edu.tr)

[nem@klu.edu.tr](mailto:nem@klu.edu.tr)

The individual contributions in this publication and any liabilities arising from them remain the responsibility of the authors.

The publisher is not responsible for possible damages, which could be a result of content derived from this publication.



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

**COMMITTEE**

**HONORARY PRESIDENT**

**Prof. Dr. Rengin AK** (Rector of Kırklareli University)

**Prof. Dr. Hasan KILIÇ** (Rector of Eastern Mediterranean University)

**CONFERENCE PRESIDENT**

**Prof. Dr. Huriye İCİL** (Eastern Mediterranean University)

**Prof. Dr. Burhan COŞKUN** (Kırklareli University)

**ORGANIZING COMMITTEE**

Prof. Dr. Fűrüzan ÇELİK / Kırklareli University

Prof. Dr. İlke TAŞÇIOĞLU / İstanbul Topkapı University

Prof. Dr. Mustafa KURBAN / Ankara University

Prof. Dr. Serpil AKÖZCAN PEHLİVANOĞLU / Kırklareli University

Assoc. Prof. Dr. Ayşegöl DERE / Fırat University

Assoc. Prof. Dr. Berna Akgenç HANEDAR / Kırklareli University

Assoc. Prof. Dr. Hale KARAYER / Kırklareli University

Assoc. Prof. Dr. Mehmet Mümin KOÇ / Kırklareli University

Assoc. Prof. Dr. Taner DAŞTAN / Cumhuriyet University

Dr. Instructor Bahadır BEKÂR / Trakya University

Dr. Instructor Ufuk PAKSU / Kırklareli University

Dr. Arwa Abou RAJAB / Eastern Mediterranean University

Dr. Çiğdem Ş. GÜÇLÜ / Gazi University

Dr. Duygu UZUN / Eastern Mediterranean University

Dr. Meltem DİNLEYİCİ / Eastern Mediterranean University

Lect. Sümeyye ÇETİN / Kırklareli University

**CONFERENCE EXECUTIVE ASSISTANT**

Çilem AYDINLIK / Eastern Mediterranean University



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

**SCIENTIFIC COMMITTEE**

- Prof. Dr. Abdel Salam Hamdy MAKHLOUF / University of Texas / UNITED STATES  
Prof. Dr. Abdullah G. AL-SEHEMI / King Khalid University / SAUDI ARABIA  
Prof. Dr. Adem TATAROĞLU / Gazi University / TÜRKİYE  
Prof. Dr. Ali ÖVGÜN / Eastern Mediterranean University / NORTH CYPRUS  
Prof. Dr. Asa H. BARBER / London South Bank University / UNITED KINGDOM  
Assoc. Prof. Dr. Bayram ÇETİN / Kırklareli University / TÜRKİYE  
Dr. Instructor Burhan CEYLAN / Harran University / TÜRKİYE  
Prof. Dr. Burç MISIRLIOĞLU / Sabancı University / TÜRKİYE  
Prof. Dr. Chung GWIY-SANG / Ulsan University / KOREA  
Dr. Instructor Demet TOPAL KOÇ / Kırklareli University / TÜRKİYE  
Prof. Dr. Denis NIKAI / Moldova University / MOLDOVA  
Asst. Prof. Dr. Devrim ÖZDAL / European University of Lefke, NORTH CYPRUS  
Prof. Dr. Dilek NARTOP / Düzce University / TÜRKİYE  
Prof. Dr. Elias STATHATOS / Tecnological-Educational Institute of Patras / GREECE  
Prof. Dr. Elizabeth BLACKBURN / University of Birmingham / UNITED KINGDOM  
Prof. Dr. Hatice ÖĞÜTCÜ / Kırşehir Ahi Evran University / TÜRKİYE  
Assoc. Prof. Elvan HASANOĞLU ÖZKAN / Gazi University / TÜRKİYE  
Asst. Prof. İsmail KILIÇ / Kırklareli University / TÜRKİYE  
Asst. Prof. Dr. Kawa H. SHARIF / Charmo University / IRAQ  
Asst. Prof. Macide CANTÜRK RODOP / Yıldız Technical University / TÜRKİYE  
Asst. Prof. Dr. Mehmet OKCAN / Eastern Mediterranean University / NORTH CYPRUS  
Prof. Dr. Müjdat ÇAĞLAR / Eskişehir Anadolu University / TÜRKİYE  
Prof. Dr. Mustafa ERKOVAN / Beykoz University / TÜRKİYE  
Prof. Dr. Mustafa HALİLSOY / Eastern Mediterranean University / NORTH CYPRUS  
Asst. Prof. Dr. Mustafa RIZA / Eastern Mediterranean University / NORTH CYPRUS  
Assoc. Prof. Naim ASLAN / Munzur University / TÜRKİYE  
Prof. Dr. Niyazi ÖZDEMİR / Fırat University / TÜRKİYE  
Prof. Dr. Numan AKDOĞAN / Gebze Technical University / TÜRKİYE  
Prof. Dr. Nur P. AYDINLIK / Cyprus International University, NORTH CYPRUS  
Prof. Dr. Ömer Suat TAŞKIN / Istanbul University / TÜRKİYE



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

Prof. Dr. Ram K. GUPTA / Pittsburg State University / UNITED STATES

Asst. Prof. Dr. Selin TEMÜRLÜ / Near East University / NORTH CYPRUS

Prof. Dr. Sermet KOYUNCU / Çanakkale Onsekiz Mart University / TÜRKİYE

Prof. Dr. S. MANSOURI / University of Gabès / TUNUSIA

Asst. Prof. Sena Esen BAYER KESKİN / Kırklareli University / TÜRKİYE

Prof. Dr. Serap ŞENTÜRK DALGIÇ / Trakya University / TÜRKİYE

Dr. Simona MANCINI / University of Salerno / ITALY

Assoc. Prof. Dr. Süleyman AŞIR / Near East University, NORTH CYPRUS

Assoc. Prof. Tarık ASAR / Gazi University / TÜRKİYE

Prof. Dr. W. Aslam FAROOQ / King Saud University / SAUDI ARABIA

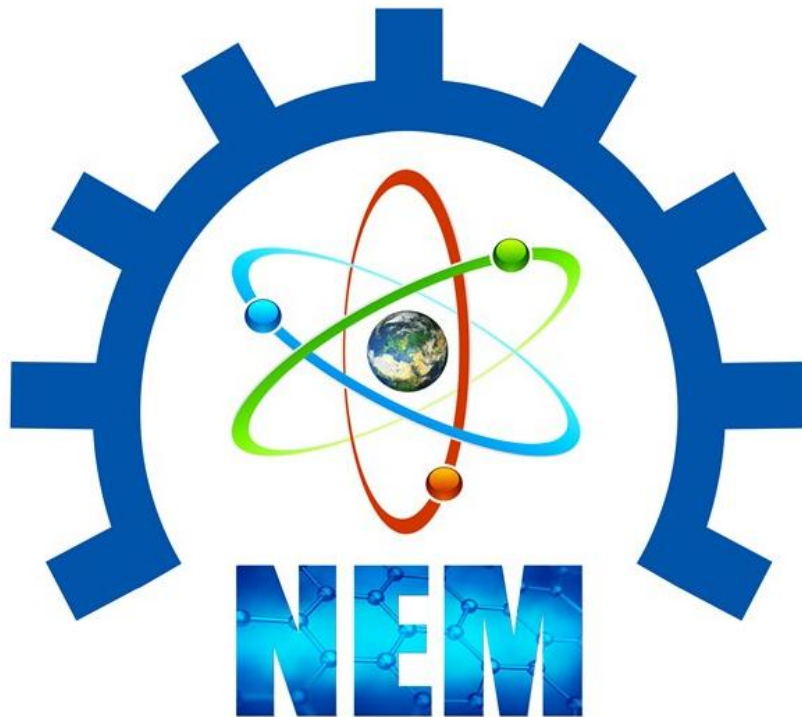
Prof. Dr. Yasin ÜNLÜTÜRK / Kırklareli University / TÜRKİYE

Prof. Dr. Yasemin Şafak ASAR / Gazi University / TÜRKİYE



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---



**FULL TEXT PRESENTATIONS**



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

**FULL TEXT PRESENTATIONS**

| NAME           | TITLE                                                                                                                                    | PAGE NO |
|----------------|------------------------------------------------------------------------------------------------------------------------------------------|---------|
| Ahsen R. Ergün | EFFECT OF LYOPHILIZED RED PEPPER AND SEED ENRICHMENT ON THE ANTIOXIDANT ACTIVITY AND QUALITY PROPERTIES OF FUNCTIONAL GLUTEN-FREE SNACKS | 1-11    |
| Idrissa Adamu  | DECONTAMINATION PARADOX: HOW REMOVING BENCHMARK OVERLAP CAN BIAS CROSS-DATASET PHISHING EVALUATION                                       | 12-20   |
| Murat Özdenefe | AIR INFILTRATION RATE MEASUREMENT OF AN OFFICE MODULE WITH TRACER GAS CONCENTRATION DECAY METHOD                                         | 21-30   |
| Murat Özdenefe | ENERGY PERFORMANCE AND MODELLING OF A CHILLER DURING AN OFF-PEAK SEASON PERIOD                                                           | 31-38   |
| Ülkü Kıymaz    | INFLUENCE of SILVER (Ag) DOPING on the CHARGING-DISCHARGING PROCESSES of MWCNTs ANODES in Li-ion BATTERIES                               | 39-46   |
| Zhala Alızada  | COMPARATIVE EVALUATION OF ARTIFICIAL INTELLIGENCE-BASED MODELS FROM ACTIVATED CARBON PRODUCTION TO ADSORPTION PERFORMANCE                | 47-57   |



6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference  
Sep 10-12, 2026, Famagusta, Northern Cyprus

## EFFECT OF LYOPHILIZED RED PEPPER AND SEED ENRICHMENT ON THE ANTIOXIDANT ACTIVITY AND QUALITY PROPERTIES OF FUNCTIONAL GLUTEN-FREE SNACKS

*A.R. Ergün<sup>1</sup>, A.Ş. Arıkan<sup>1</sup>, M.M. Yıldırım<sup>1</sup>, T. Baysal*

<sup>1</sup>Department of Food Engineering, Faculty of Engineering, Ege University, İzmir, TÜRKİYE

E-mail: [ahsenrayman@gmail.com](mailto:ahsenrayman@gmail.com)

### Abstract

Snack products represent one of the fastest-growing categories in the food industry due to increasing consumption trends, while the evaluation of plant seeds generated as food processing by-products is important for sustainable production; red pepper and its seeds, rich in phenolic compounds and antioxidants, are therefore regarded as innovative and promising ingredients for functional product development. This study aimed to develop functional gluten-free snack formulations containing freeze-dried red pepper and red pepper seed powder obtained through lyophilization and to evaluate the effects of different formulations on the physicochemical, functional, and sensory properties of the final products. Three formulations were prepared: a control sample containing only gluten-free flour; a sample containing gluten-free flour and lyophilized red pepper seed powder; and a sample containing gluten-free flour, lyophilized red pepper seed powder, and lyophilized red pepper. The products were analyzed for chemical composition (moisture, ash, fat, protein), water activity, antioxidant activity, total phenolic content, textural parameters (hardness, brittleness), and sensory properties. The production process completed with baking at 150 °C for 5 min following tray drying at 60 °C and 1.5 m/s air velocity, was more effective in preserving bioactive components and promoting crisp texture than conventional atmospheric cooking. Antioxidant capacity, determined by the DPPH method, was found to be as  $49.126 \pm 6.944$   $\mu\text{mol TE/g}$  in the control group,  $55.381 \pm 0.346$   $\mu\text{mol TE/g}$  in the seed-containing sample, and  $59.504 \pm 0.795$   $\mu\text{mol TE/g}$  in the sample containing both red pepper and seeds; this indicates that the addition of plant-based components increased the antioxidant capacity. The limited literature on lyophilized plant by-products in functional gluten-free snacks highlights the importance of this study, whose findings demonstrate an innovative and applicable approach and provide a scientific basis for industrial-scale production of zero-waste, clean-label, nutritionally enriched gluten-free products.

**Keywords:** Antioxidant activity, Functional snack, Gluten-free product, Red pepper and seed, Lyophilization



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

## **1. INTRODUCTION**

Consumer interest in gluten-free products has increased substantially in recent years [1]. Although a gluten-free diet is essential for individuals with celiac disease and is also adopted for other reasons, the nutritional quality of gluten-free bakery products remains an important concern [1, 2]. Such products often rely on refined starches and flours and may benefit from enrichment with plant-derived ingredients providing proteins, dietary fiber, minerals, and bioactive compounds [3]. Functional food development increasingly focuses on incorporating plant-derived ingredients and food-processing by-products to improve nutritional value while supporting the sustainable utilization of plant materials [4, 5]. Red pepper (*Capsicum annuum* L.) is a valuable source of carotenoids, vitamin C, phenolic compounds, and other antioxidants, although its phytochemical properties are affected by processing and drying conditions [6]. Pepper seeds, often regarded as a by-product, also have considerable nutritional and functional potential. Studies have identified sterols, triterpenes, organic acids, fatty acids, volatile compounds, proteins, lipids, dietary fiber, phenolics, and minerals in *C. annuum* seeds, along with concentration-dependent antioxidant activity [7–10]. Reported compositions include approximately 18.3% crude protein, 11.0% crude oil, and 61.0% dietary fiber, with the lipid fraction being predominantly unsaturated fatty acids [8]. Drying conditions are critical for preserving bioactive compounds in plant-based ingredients. High temperatures can accelerate drying but may adversely affect color, phenolics, carotenoids, and other heat-sensitive compounds [6]. Freeze-drying is therefore frequently preferred for bioactive-rich materials due to its favorable retention of carotenoids and characteristic red color compared with some conventional drying methods [11]. However, freeze-drying the final snack product is economically and technologically demanding. A more practical approach is to freeze-dry the functional ingredients and subsequently produce the crackers using moderate-temperature convective drying, combining bioactive preservation with a simpler and more scalable process. Accordingly, this study aimed to develop gluten-free crackers enriched with freeze-dried red pepper seeds and a combination of freeze-dried red pepper and red pepper seeds. Antioxidant capacity, total phenolic content, moisture, water activity, crude protein, fat, ash, instrumental color, texture, and sensory properties were evaluated against a gluten-free control formulation.



## 2. MATERIALS AND METHODS

### 2.1. Materials

Gluten-free flour and other cracker ingredients were obtained from commercial sources. Fresh red peppers (*Capsicum annuum* L.) and red pepper seeds were used to prepare the functional ingredients. Analytical-grade reagents were used for physicochemical, phenolic, and antioxidant analyses.

### 2.2. Preparation of Red Pepper and Pepper Seed Powders

Red pepper and red pepper seeds were separately frozen and freeze-dried until a stable dry structure was obtained. The dried materials were ground into homogeneous powders and stored in sealed containers until cracker production. Freeze-drying was selected due to its favorable retention of carotenoids and characteristic pepper color compared with intensive thermal drying [11].

### 2.3. Experimental Design and Cracker Formulation

Three gluten-free cracker formulations were prepared: control (C), without pepper enrichment; pepper seed (PS), containing freeze-dried red pepper seed powder; and red pepper–pepper seed (RPPS), containing both freeze-dried red pepper and pepper seed powders. The total amount of functional ingredient was kept constant in the enriched formulations, while other formulation and processing conditions remained unchanged.

### 2.4. Cracker Preparation and Drying

Dry ingredients were mixed, followed by incorporation of the remaining ingredients to form the dough. The doughs were sheeted and shaped to uniform thickness and dimensions. Cracker samples were dried in a convective tray dryer at 60 °C and 1.5 m/s air velocity, followed by baking at 150 °C for 5 min. The products were then cooled to room temperature and stored in sealed packaging until analysis.

### 2.5. Moisture Content

Moisture content was determined using AOAC/AACC-based gravimetric procedures. Approximately 5 g of sample was dried at  $103 \pm 2$  °C to constant weight, and moisture content was calculated from the mass loss [12, 13].

### 2.6. Water Activity

Water activity (aw) was measured at room temperature using a calibrated AquaLab-type water activity meter. Measurements were performed in triplicate after equilibrium was reached.

### 2.7. Ash Content

Approximately 3 g of sample was weighed into pre-weighed porcelain crucibles and incinerated in a muffle furnace. Samples were initially heated at approximately 600 °C and maintained at 550 °C until



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

complete mineralization. After cooling in a desiccator, ash content was determined gravimetrically according to AOAC procedures [12].

### **2.8. Crude Protein Content**

Crude protein was determined by the Kjeldahl method. Samples were digested with concentrated H<sub>2</sub>SO<sub>4</sub> using potassium sulfate and copper sulfate catalysts. After alkalization with NaOH, released ammonia was distilled into boric acid and titrated with standardized 0.1 N HCl. Nitrogen content was calculated after blank correction and converted to crude protein using a factor of 6.25 [12].

### **2.9. Crude Fat Content**

Crude fat was determined by solvent extraction. Approximately 3 g of ground sample was extracted with a non-polar solvent, after which the solvent was removed and the lipid fraction was weighed. Fat content was calculated as  $\text{Fat (\%)} = [(W_2 - W_1) / W_s] \times 100$ .

### **2.10. Instrumental Color**

Color was measured using the CIE Lab\* system with a calibrated Hunter colorimeter. L\* indicates lightness, a\* the green-to-red axis, and b\* the blue-to-yellow axis. Five measurements were obtained per formulation and expressed as mean  $\pm$  standard deviation. Total color difference was calculated as  $\Delta E^* = [(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2]^{1/2}$ .

### **2.11. Extraction of Phenolic and Antioxidant Compounds**

Approximately 5 g of ground sample was extracted with 50 mL of 80% methanol. Samples were homogenized, protected from light, and periodically vortexed during extraction. Following centrifugation, the supernatants were used for total phenolic and antioxidant analyses [14].

### **2.12. Total Phenolic Content**

Total phenolic content was determined using the Folin-Ciocalteu method according to Singleton et al. [15]. Extracts were reacted with Folin-Ciocalteu reagent and sodium carbonate, followed by incubation for color development. Absorbance was measured at 760 nm. Gallic acid was used as the calibration standard, and results were expressed as gallic acid equivalents.

### **2.13. Antioxidant Capacity**

Antioxidant capacity was determined using the DPPH radical-scavenging method according to Brand-Williams et al. [16]. Sample extracts were reacted with DPPH solution in methanol under controlled conditions, and absorbance was measured at 517 nm. Results were expressed as  $\mu\text{mol Trolox equivalents per gram } (\mu\text{mol TE/g})$ .



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

#### 2.14. Texture Analysis

Texture characteristics were evaluated using a TA.XT2i Texture Analyzer (Stable Micro Systems, UK) with a three-point bending/fracture configuration. Individual crackers were positioned across two supports and fractured by compression at the midpoint. Maximum fracture force was recorded as hardness/breaking force and expressed in Newtons (N), with fracture behavior also considered as an indicator of brittleness and crispness.

#### 2.15. Sensory Evaluation

Sensory evaluation was performed using a 9-point hedonic scale (1 = dislike extremely; 9 = like extremely) with at least 10 semi-trained panelists [17]. Samples were presented in transparent containers with randomized three-digit codes and randomized serving order. Water was provided for palate cleansing. Appearance, color, odor, taste, texture, and overall acceptability were evaluated.

#### 2.16. Statistical Analysis

Results were expressed as mean  $\pm$  standard deviation. Data were analyzed by one-way ANOVA, followed by Duncan's multiple-range test where applicable. Statistical significance was set at  $p < 0.05$  using SPSS Statistics version 16.0 (SPSS Inc., Chicago, IL, USA).

### 3. RESULTS AND DISCUSSION

#### 3.1. Antioxidant Capacity and Total Phenolic Content

The antioxidant and total phenolic results of the developed crackers are presented in Table 1.

**Table 1. Antioxidant capacity and total phenolic content of gluten-free cracker formulations.**

| Formulation                     | Antioxidant capacity ( $\mu\text{mol TE/g}$ ) | Total phenolic content (mg GAE/g) |
|---------------------------------|-----------------------------------------------|-----------------------------------|
| Control (C)                     | $49.126 \pm 1.944^a$                          | $0.149 \pm 0.040^a$               |
| Pepper seed (PS)                | $55.381 \pm 0.346^b$                          | $0.209 \pm 0.003^b$               |
| Red pepper + pepper seed (RPPS) | $59.504 \pm 0.795^c$                          | $1.512 \pm 0.095^c$               |

\* In the table, different letters within the same row indicate a statistically significant difference ( $p < 0.05$ ) between groups according to the Duncan test results.

Antioxidant capacity increased from  $49.126 \mu\text{mol TE/g}$  in C to  $55.381 \mu\text{mol TE/g}$  in PS and  $59.504 \mu\text{mol TE/g}$  in RPPS, corresponding to increases of 12.9% and 21.1%, respectively. ANOVA confirmed a significant difference among formulations ( $p < 0.05$ ). The increase is consistent with the reported antioxidant potential of pepper seeds and red pepper, which contain phenolic compounds, carotenoids,



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

ascorbic acid, and other bioactive constituents [6, 7]. Total phenolic content increased from  $0.149 \pm 0.040$  mg GAE/g in C to  $0.209 \pm 0.003$  mg GAE/g in PS and  $1.512 \pm 0.095$  mg GAE/g in RPPS ( $p < 0.05$ ). Compared with C, PS showed an approximately 40% increase, while RPPS exhibited approximately tenfold higher phenolic content. This pronounced increase suggests that the freeze-dried red pepper fraction was the major contributor to phenolic enrichment [6, 11]. The substantially greater increase in total phenolic content than DPPH activity is reasonable because these methods reflect different chemical properties. DPPH activity is affected not only by total phenolics but also by individual compounds, synergistic interactions, pigments, vitamins, and other reducing substances.

### 3.2. Protein, Fat and Ash Contents

The proximate composition of the cracker formulations is presented in Table 2.

**Table 2. Protein, fat and ash contents of gluten-free cracker formulations.**

| Formulation | Protein (%)         | Fat (%)             | Ash (%)             |
|-------------|---------------------|---------------------|---------------------|
| C           | $3.664 \pm 0.055^a$ | $4.833 \pm 0.833^a$ | $2.103 \pm 0.427^a$ |
| PS          | $4.758 \pm 0.164^b$ | $7.333 \pm 0.333^b$ | $1.820 \pm 0.048^b$ |
| RPPS        | $4.922 \pm 0.109^b$ | $7.167 \pm 0.167^b$ | $2.269 \pm 0.360^a$ |

\* In the table, different letters within the same row indicate a statistically significant difference ( $p < 0.05$ ) between groups according to the Duncan test results.

Protein content increased from  $3.664 \pm 0.055\%$  in C to  $4.758 \pm 0.164\%$  in PS and  $4.922 \pm 0.109\%$  in RPPS, consistent with the considerable protein content reported for pepper seeds [8]. Fat content increased from  $4.833\%$  in C to  $7.333\%$  in PS and  $7.167\%$  in RPPS, with a significant formulation effect ( $p < 0.05$ ). This increase is attributable to the lipid-rich nature of pepper seeds, whose oils are predominantly composed of unsaturated fatty acids, particularly linoleic acid [8, 9]. Ash contents were  $2.103 \pm 0.427\%$ ,  $1.820 \pm 0.048\%$ , and  $2.269 \pm 0.360\%$  for C, PS, and RPPS, respectively, with no significant difference ( $p > 0.05$ ). Overall, pepper-derived ingredients modified the protein and fat composition of the crackers while having limited effect on ash content.



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

### 3.3. Moisture Content and Water Activity

Moisture and water activity data before and after drying are presented in Table 3.

**Table 3. Moisture content and water activity of dough and dried cracker samples.**

| Parameter            | C              | PS             | RPPS           |
|----------------------|----------------|----------------|----------------|
| Dough moisture (%)   | 40.223 ± 0.033 | 36.893 ± 0.046 | 32.580 ± 0.073 |
| Dough aw             | 0.931 ± 0.003  | 0.922 ± 0.004  | 0.928 ± 0.005  |
| Cracker moisture (%) | 4.220 ± 0.029  | 10.453 ± 0.203 | 5.523 ± 0.256  |
| Cracker aw           | 0.308 ± 0.002  | 0.525 ± 0.005  | 0.407 ± 0.004  |

Moisture content differed significantly among formulations ( $p < 0.05$ ), with values of 4.22%, 10.45%, and 5.52% for C, PS, and RPPS, respectively. Water activity showed a similar trend, with values of 0.308, 0.525, and 0.407. The higher moisture and aw of PS may be related to the protein and fiber content of pepper seeds, which can increase water binding within the cracker matrix [9]. This behavior is technologically important because moisture and aw influence crispness and storage stability. Thus, the PS formulation may require longer or modified drying conditions to achieve a texture comparable to the control. Nevertheless, the tray-drying process successfully produced low-moisture, low-aw crackers in all formulations.

### 3.4. Instrumental Color

The color parameters of the crackers are presented in Table 4.

**Table 4. CIE L\*a\*b\* color parameters and total color difference of gluten-free cracker formulations.**

| Formulation | L*                          | a*                          | b*                          | $\Delta E$ |
|-------------|-----------------------------|-----------------------------|-----------------------------|------------|
| C           | 89.804 ± 0.069 <sup>a</sup> | -0.044 ± 0.016 <sup>a</sup> | 9.026 ± 0.026 <sup>a</sup>  | —          |
| PS          | 88.206 ± 0.106 <sup>b</sup> | 0.508 ± 0.030 <sup>b</sup>  | 10.140 ± 0.069 <sup>b</sup> | 2.02       |
| RPPS        | 71.804 ± 0.040 <sup>c</sup> | 11.380 ± 0.046 <sup>c</sup> | 22.618 ± 0.070 <sup>c</sup> | 25.28      |

\* In the table, different letters within the same row indicate a statistically significant difference ( $p < 0.05$ ) between groups according to the Duncan test results.

Pepper seed enrichment caused only a slight reduction in L\* from 89.804 to 88.206, indicating limited visual change. In contrast, RPPS showed a substantially lower L\* value (71.804) and higher a\* (11.380) and b\* (22.618) values, indicating a darker and more pronounced red-orange color. These changes are associated with the carotenoid pigments of red pepper, whose color properties can be affected by



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

processing and drying conditions [6, 11].  $\Delta E$  values were approximately 2.02 between C and PS and 25.28 between C and RPPS, confirming that red pepper incorporation produced a distinct visual difference.

### 3.5. Texture Properties

Breaking force decreased from  $18.6 \pm 1.4$  N in C to  $11.9 \pm 1.1$  N in PS and  $15.2 \pm 1.3$  N in RPPS ( $p < 0.05$ ). The lower breaking force of PS may be associated with its higher moisture and water activity, which can increase molecular mobility and reduce cracker crispness. Fracturability values were  $2.15 \pm 0.18$ ,  $3.42 \pm 0.25$ , and  $2.76 \pm 0.21$  mm for C, PS, and RPPS, respectively. The greater deformation of PS indicates a less brittle structure, whereas C exhibited the most brittle and crisp characteristics. RPPS showed intermediate texture, suggesting that the combined incorporation of red pepper and seed modified the structure without the pronounced softening observed in PS. Overall, RPPS provided a favorable balance between functional enrichment and texture, maintaining relatively firm and brittle characteristics while substantially increasing phenolic and antioxidant properties.

**Table 5. Instrumental texture properties of gluten-free cracker formulations.**

| Formulation                     | Breaking Force / Hardness (N) | Fracturability (mm)  |
|---------------------------------|-------------------------------|----------------------|
| Control (C)                     | $18.60 \pm 1.40^a$            | $2.15 \pm 0.18^b$    |
| Pepper Seed (PS)                | $11.90 \pm 1.10^c$            | $3.42 \pm 0.25^a$    |
| Red Pepper + Pepper Seed (RPPS) | $15.20 \pm 1.30^b$            | $2.76 \pm 0.21^{ab}$ |

\* In the table, different letters within the same row indicate a statistically significant difference ( $p < 0.05$ ) between groups according to the Duncan test results.

### 3.6. Sensory Characteristics

All formulations were generally acceptable, although pepper-derived ingredients affected individual sensory attributes. The control received high scores for taste ( $7.7 \pm 0.7$ ), texture ( $7.8 \pm 0.6$ ), and overall acceptability ( $7.7 \pm 0.6$ ), likely reflecting its lower moisture and aw and consequently crisp structure. PS showed lower scores for taste ( $6.5 \pm 1.0$ ), texture ( $6.4 \pm 0.9$ ), and overall acceptability ( $6.6 \pm 0.8$ ). These results may be associated with its higher moisture and aw as well as the characteristic flavor of pepper seeds.

RPPS exhibited the most favorable enriched formulation profile, receiving the highest color score ( $8.1 \pm 0.6$ ) and a high appearance score ( $7.8 \pm 0.7$ ). It also received favorable scores for odor ( $7.4 \pm 0.7$ ),



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

taste ( $7.1 \pm 0.8$ ), texture ( $7.2 \pm 0.8$ ), and overall acceptability ( $7.4 \pm 0.7$ ). Its characteristic red-orange color was consistent with the instrumental color results.

**Table 6. Sensory evaluation scores of gluten-free cracker formulations.**

| Sensory Attribute     | Control (C)     | Pepper Seed (PS)   | Red Pepper + Pepper Seed (RPPS) |
|-----------------------|-----------------|--------------------|---------------------------------|
| Appearance            | $7.6 \pm 0.8^a$ | $7.1 \pm 0.9^{ab}$ | $7.8 \pm 0.7^a$                 |
| Color                 | $7.4 \pm 0.7^b$ | $6.9 \pm 0.8^b$    | $8.1 \pm 0.6^a$                 |
| Odor                  | $7.3 \pm 0.8^a$ | $6.8 \pm 0.9^a$    | $7.4 \pm 0.7^a$                 |
| Taste                 | $7.7 \pm 0.7^a$ | $6.5 \pm 1.0^b$    | $7.1 \pm 0.8^{ab}$              |
| Texture               | $7.8 \pm 0.6^a$ | $6.4 \pm 0.9^b$    | $7.2 \pm 0.8^a$                 |
| Overall Acceptability | $7.7 \pm 0.6^a$ | $6.6 \pm 0.8^b$    | $7.4 \pm 0.7^a$                 |

*Note: Values are expressed as mean  $\pm$  standard deviation. Different superscript letters within the same row indicate significant differences among formulations ( $p < 0.05$ ). Sensory evaluation was performed using a 9-point hedonic scale (1 = dislike extremely, 9 = like extremely).*

#### 4. CONCLUSION

This study demonstrated the potential of freeze-dried red pepper and pepper seed powders as functional ingredients for gluten-free cracker production. Pepper enrichment increased antioxidant capacity, with values ranging from  $49.126 \pm 6.944$   $\mu\text{mol TE/g}$  in the control to  $59.504 \pm 0.795$   $\mu\text{mol TE/g}$  in RPPS. Total phenolic content showed a more pronounced increase, reaching  $1.512 \pm 0.095$   $\text{mg GAE/g}$  in RPPS, approximately ten times higher than the control. Protein content also increased to  $4.758 \pm 0.164\%$  in PS and  $4.922 \pm 0.109\%$  in RPPS. Enrichment affected moisture, water activity, and color characteristics. PS exhibited the highest moisture and water activity, suggesting the need for further optimization of drying conditions, whereas RPPS produced a distinct red-orange color with higher  $a^*$  and  $b^*$  values. Overall, RPPS showed the most favorable functional profile, combining increased phenolic content, protein and antioxidant capacity with natural coloration. Pepper seeds also offer potential as value-added ingredients by contributing proteins, lipids, dietary fiber, and bioactive compounds while supporting the utilization of processing by-products. Overall, freeze-dried red pepper and pepper seed powders show promise for developing antioxidant-enriched gluten-free snack products. Further research should focus on optimizing drying conditions and evaluating texture, sensory acceptance, and storage stability.



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

**REFERENCES**

- [1] Melini, V., & Melini, F. (2019). Gluten-free diet: Gaps and needs for a healthier diet. *Nutrients*, 11(1), 170. DOI: 10.3390/nu11010170.
- [2] Vici, G., Belli, L., Biondi, M., & Polzonetti, V. (2016). Gluten free diet and nutrient deficiencies: A review. *Clinical Nutrition*, 35(6), 1236–1241. DOI: 10.1016/j.clnu.2016.05.002.
- [3] Conte, P., Fadda, C., Drabińska, N., & Krupa-Kozak, U. (2019). Technological and nutritional challenges, and novelty in gluten-free breadmaking: A review. *Polish Journal of Food and Nutrition Sciences*, 69(1), 5–21. DOI: 10.31883/pjfn-2019-0005.
- [4] Galanakis, C. M. (2012). Recovery of high added-value components from food wastes: Conventional, emerging technologies and commercialized applications. *Trends in Food Science & Technology*, 26(2), 68–87. DOI: 10.1016/j.tifs.2012.03.003.
- [5] Baiano, A. (2014). Recovery of biomolecules from food wastes—A review. *Molecules*, 19(9), 14821–14842. DOI: 10.3390/molecules190914821.
- [6] Vega-Gálvez, A., Di Scala, K., Rodríguez, K., Lemus-Mondaca, R., Miranda, M., López, J., & Perez-Won, M. (2009). Effect of air-drying temperature on physico-chemical properties, antioxidant capacity, colour and total phenolic content of red pepper (*Capsicum annuum* L. var. Hungarian). *Food Chemistry*, 117(4), 647–653. DOI: 10.1016/j.foodchem.2009.04.066.
- [7] Silva, L. R., Azevedo, J., Pereira, M. J., Valentão, P., & Andrade, P. B. (2013). Chemical assessment and antioxidant capacity of pepper (*Capsicum annuum* L.) seeds. *Food and Chemical Toxicology*, 53, 240–248. DOI: 10.1016/j.fct.2012.11.036.
- [8] Azabou, S., Ben Taheur, F., Jridi, M., Bouaziz, M., & Nasri, M. (2017). Discarded seeds from red pepper (*Capsicum annum*) processing industry as a sustainable source of high added-value compounds and edible oil. *Environmental Science and Pollution Research*, 24(28), 22196–22203. DOI: 10.1007/s11356-017-9857-9.
- [9] Cvetković, T., Ranilović, J., & Jokić, S. (2022). Quality of pepper seed by-products: A review. *Foods*, 11(5), 748. DOI: 10.3390/foods11050748.
- [10] Dobón-Suárez, A., Zapata, P. J., & García-Pastor, M. E. (2025). A comprehensive review on characterization of pepper seeds: Unveiling potential value and sustainable agrifood applications. *Foods*, 14(11), 1969. DOI: 10.3390/foods14111969.



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

- [11] Krzykowski, A., Rudy, S., Polak, R., Biernacka, B., Krajewska, A., Janiszewska-Turak, E., Kowalska, I., Żuchowski, J., Skalski, B., & Dzik, D. (2024). Drying of red chili pepper (*Capsicum annuum* L.): Process kinetics, color changes, carotenoid content and phenolic profile. *Molecules*, 29(21), 5164. DOI: 10.3390/molecules29215164.
- [12] AOAC International. (2019). *Official Methods of Analysis of AOAC International* (21st ed.). AOAC International, Rockville, MD, USA.
- [13] AACC International. (2010). *Approved Methods of Analysis* (11th ed.). AACC International, St. Paul, MN, USA.
- [14] Do, Q. D., Angkawijaya, A. E., Tran-Nguyen, P. L., Huynh, L. H., Soetaredjo, F. E., Ismadji, S., & Ju, Y. H. (2014). Effect of extraction solvent on total phenol content, total flavonoid content, and antioxidant activity of *Limnophila aromatica*. *Journal of Food and Drug Analysis*, 22(3), 296–302. DOI: 10.1016/j.jfda.2013.11.001.
- [15] Singleton, V. L., Orthofer, R., & Lamuela-Raventós, R. M. (1999). Analysis of total phenols and other oxidation substrates and antioxidants by means of Folin–Ciocalteu reagent. *Methods in Enzymology*, 299, 152–178. DOI: 10.1016/S0076-6879(99)99017-1.
- [16] Brand-Williams, W., Cuvelier, M. E., & Berset, C. (1995). Use of a free radical method to evaluate antioxidant activity. *LWT – Food Science and Technology*, 28(1), 25–30. DOI: 10.1016/S0023-6438(95)80008-5.
- [17] Altuğ, T., & Elmacı, Y. (2005). *Gıdalarda Duyusal Değerlendirme*. Meta Basım Matbaacılık Hizmetleri, İzmir, Türkiye.



## DECONTAMINATION PARADOX: HOW REMOVING BENCHMARK OVERLAP CAN BIAS CROSS-DATASET PHISHING EVALUATION

I. Adamu<sup>1</sup>, O. B. Adedoyin<sup>1</sup>, M. Tanriverdi<sup>2</sup>, N. Karaca<sup>3</sup>

*\*Department of Computer Engineering, Faculty of Engineering, University of Kyrenia, Mersin 10, TÜRKİYE*

[Idrissa.adamu@kyrenia.edu.tr](mailto:Idrissa.adamu@kyrenia.edu.tr), [olasilebabatunde.adedoyin@kyrenia.edu.tr](mailto:olasilebabatunde.adedoyin@kyrenia.edu.tr),  
[20240362@std.kyrenia.edu.tr](mailto:20240362@std.kyrenia.edu.tr), [k20230312@std.kyrenia.edu.tr](mailto:k20230312@std.kyrenia.edu.tr)

### Abstract

Phishing URL detectors often report strong results on benchmark datasets, but these results can change when the models are tested on data on another source. This study questioned whether cross-dataset evaluation gives a reliable picture when the two datasets overlap and have different class distributions. We compare two public datasets, PhiUSIIL and LegitPhish, at the URL and registrable-domain levels and find 36,945 shared URLs and 36,531 shared registrable domains. All shared URLs are labelled legitimate in both datasets, with no label disagreements. Cleaning the overlapping domains removes 99.27% of the legitimate URLs while only 2.93% of the phishing URLs in LegitPhish, so the external test set is 99.56% phishing. We evaluate Logistic Regression and XGBoost on lexical URL features. On the external set, both models have very high false-positive rates (90.18% and 94.55%), yet the ROC-AUC of XGBoost stays at 0.967 on both the internal and external sets. The rise in false positives therefore appears to be from applying a fixed decision threshold to a very different class distribution, and not from a full loss of discrimination. SHAP analysis shows that high agreement in feature ranks does not mean stable explanations, because some feature contributions change direction across datasets. Dataset overlap, class composition, decision threshold and explanation stability should all be checked when phishing detectors are evaluated across datasets.

**Keywords:** Phishing URL detection, Machine learning, Cross-dataset evaluation, Dataset decontamination, Explainable artificial intelligence



## **1. INTRODUCTION**

Machine learning classifiers for phishing URL detection often report accuracy above 97% on public benchmarks [1], [2], [3]. These results have not carried over to deployment, where detection systems still produce more false positives than security teams can handle [4]. The gap is increasingly blamed on how models are evaluated more than on model capacity [5], [6]. Most detectors use lexical features of the URL string, which need no network resolution [7], [8], and many report accuracy without explaining which features drive the predictions [9], [10]. Explanations such as SHAP are usually computed inside a single dataset [11].

Cross-dataset evaluation, which trains on one corpus and tests on another, is the standard way to probe the benchmark. Mia et al. [12] combined it with SHAP and found that accuracy fell to 51% and 59% across corpora and that the contribution direction of individual features reversed between datasets. Later studies report similar degradation [13], [14]. This approach assumes that the two corpora are independent. If they share records, the external test set is partly an internal one and the estimate is optimistic, an instance of the data snooping pitfall described by Arp et al. [5]. Duplication, inconsistent labelling and URL overlap have been documented in phishing corpora [14] [15], but independence is rarely measured in cross-dataset studies, and to our understanding no study report an overlap audit.

This study, [12] documented cross-dataset SHAP instability but do not audit whether their corpora are independent. Afonso et al. [5] also, studied overlap and duplication across public phishing corpora but do not examine whether the overlap is asymmetric across the legitimate and phishing classes, or whether it affects explanation stability. Arp et al. [5], also warn about class distributions that differ from deployment, which makes the class balance of an external set part of the same question. We connect these by asking whether the composition of the overlap, and of the decontaminated test set, can influence threshold-dependent performance and SHAP stability.

We measured this independence for PhiUSIIL [16] (235,795 URLs) and LegitPhish [17] (101,219 URLs). They share 36,945 normalized URLs and 36,531 registrable domains. The obvious remedy is decontamination: remove from the external corpus every record that overlaps the training corpus. Our central finding is that this remedy is not class-neutral. Legitimate URLs in both corpora come from popularity-derived or high-authority pools that largely coincide [18], [19] while phishing URLs come from short-lived threat feeds that barely overlap. Domain-level exclusion therefore removed 99.3% of the 37,540 legitimate LegitPhish URLs but only 2.9% of its 63,678 phishing URLs, and the resulting set of 62,087 records is 99.56% phishing. A classifier that always predicts phishing



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

would reach 99.56% accuracy on it, and our trained models beat that by only 17 and 13 records. The false-positive rates above 90% that we report are computed on 275 surviving legitimate URLs, which survived because they did not resemble anything in the training data. We call this the decontamination paradox: a procedure meant to make cross-dataset evaluation trustworthy can shape the result it is supposed to measure. This study, do not argue against decontamination. But argue that its class-differential effect should be measured and reported.

The following four Research Question (RQ) shaped the study.

1. RQ1: how much do the two corpora overlap at the exact URL and registrable domain levels, and are the labels of shared URLs consistent?
2. RQ2: how does decontamination change the size and class composition of the external set?
3. RQ3: how do estimates at a frozen decision threshold compare with threshold-free discrimination and a majority-class baseline?
4. RQ4: do SHAP rank-stability measures capture the stability of explanations, including the direction of feature contributions?

To our knowledge, this is the first overlap and label-consistency audit of this benchmark pair.

## **2. MATERIAL AND METHODS**

**Corpora.** PhiUSIIL contains 235,795 URLs (134,850 legitimate, 100,945 phishing) [16] and LegitPhish contains 101,219 URLs (63,678 phishing, 37,540 legitimate) [17]. Both use 1 = legitimate and 0 = phishing, and we mapped both to 0 = legitimate and 1 = phishing. PhiUSIIL takes its legitimate URLs from a domain popularity source and its phishing URLs from public threat feeds. LegitPhish takes its legitimate URLs from outbound links on high-authority reference sites and from search-engine results, and its phishing URLs from threat-intelligence feeds. Before looking at the overlap, we predicted that it would be concentrated in the legitimate class and that overlap based exclusion would remove legitimate records at a much higher rate than phishing records.

**Decontamination.** Each URL was lower cased, International Domain Names in Applications (IDNA) decoded and stripped of its fragment, keeping scheme, path and query. Hostname and registrable domain (eTLD+1) were derived with an offline Public Suffix List. Within each corpus, records with missing or invalid URLs or labels were removed, URLs that occur under both labels, and same label duplicates (first occurrence kept). This left 234,687 PhiUSIIL and 100,196 LegitPhish



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

unique URLs. Then removed from LegitPhish every record whose registrable domain occurs in PhiUSIIL (38,109 records). The remaining 62,087 records (275 legitimate, 61,812 phishing) form the external set. URLs were processed as text only and were never visited or resolved.

**Features and models.** All 29 features are computed from the URL string alone. PhiUSIIL was split by registrable domain into training (164,273 rows), validation (35,207) and internal test (35,207) sets, with no domain in more than one set. The external set was not used for model selection. Logistic Regression is the linear baseline (standardized features, L2 regularization with  $C = 1.0$ , lbfgs solver, class-balanced weights). XGBoost uses the binary:logistic objective,  $\text{scale\_pos\_weight} = 1.378$ , 300 trees, maximum depth 6, learning rate 0.05, subsample 0.80 and  $\text{colsample\_bytree} = 0.80$ . Hyperparameters were fixed defaults, not tuned, and the random seed was 42. Decision thresholds were chosen on the validation set by maximizing the Matthews Correlation Coefficient (MCC) and then frozen for all later evaluation: 0.18 for Logistic Regression and 0.115 for XGBoost.

**Metrics and explanations.** We report accuracy, precision, recall, F1-score, false positive rate (FPR), MCC, ROC-AUC and PR-AUC. ROC-AUC and PR-AUC are computed over all thresholds and are reported next to threshold dependent metrics, because the external set has a very different class prior from the training corpus. Every evaluation set is also scored against a majority class baseline, since accuracy can mislead under severe imbalance [20]. Proportions come with Wilson 95% confidence intervals and ROC-AUC with stratified bootstrap 95% percentile intervals (400 resamples). For SHAP [21] exact linear contributions for Logistic Regression and TreeSHAP for XGBoost was used. Because the external set is 99.56% phishing, global attributions were computed on two cohorts per evaluation set: a balanced cohort (275 legitimate and 275 phishing records) and a phishing only cohort (5,000 records). Global importance is the mean absolute Shapley value. Then we compare internal and external rankings with Spearman's  $\rho$ , top 10 Jaccard overlap, signed direction consistency and mean absolute rank difference.

### **3. RESULTS AND DISCUSSION**

**Overlap and the cost of decontamination.** Table 1 summarises the audit. All shared URLs are labelled legitimate in both corpora, no phishing URL is shared, and no shared URL has conflicting labels. The corpora are therefore not independent, and the makeup of the overlap matters more than its size. Because the labels always agree, the redundancy reflects a common sampling frame and not a labelling defect, as predicted in Section 2. Decontamination acts almost only on one class, with an attrition ratio of about 34:1.



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

**Table 1.** Overlap between PhiUSIIL and LegitPhish, and class attrition in LegitPhish after decontamination.

| Quantity                          | Count            | Share                     |
|-----------------------------------|------------------|---------------------------|
| Shared unique normalized URLs     | 36,945           | 36.87% of LegitPhish      |
| Shared unique registrable domains | 36,531           | 58.17% of LegitPhish      |
| Legitimate URLs removed           | 37,265 of 37,540 | 99.27%                    |
| Phishing URLs removed             | 1,866 of 63,678  | 2.93%                     |
| Phishing URLs in external set     | 61,812 of 62,087 | 99.56% (62.9% originally) |

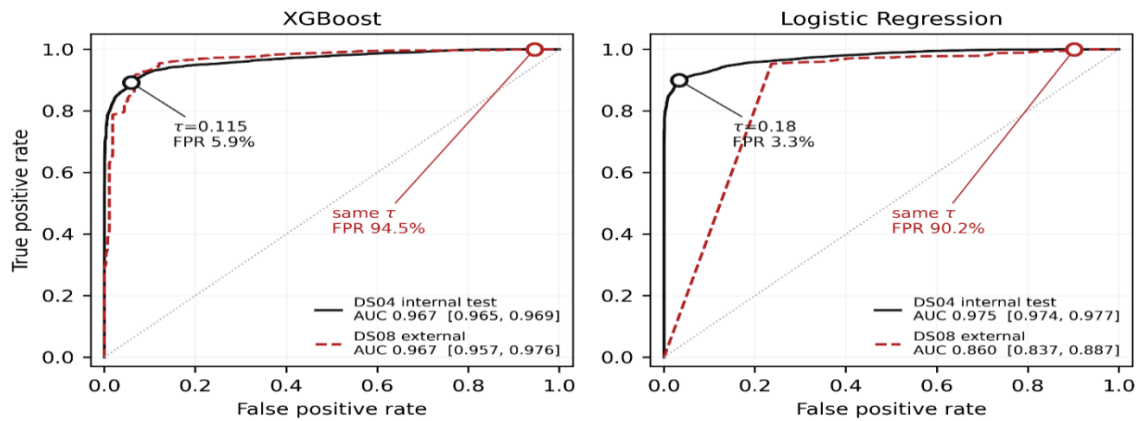
The exclusion granularity does not explain this. Out of all the 37,540 legitimate URLs in LegitPhish, 36,945, nearly (98.4%) are exact duplicates of PhiUSIIL records, and would be removed even under the least restrictive criterion, so no granularity leaves an independent external set with a usable legitimate class. The 275 surviving legitimate URLs are, by construction, those that did not overlap with the training corpus. The external set is therefore not a harder version of the same task but a population selected for its dissimilarity from the training data. Results on it show how the models behave on unusual legitimate URLs, not whether they generalise to independent legitimate traffic. Skipping decontamination is not defensible either, since 58.17% of the registrable domains are shared. Its effect on class composition decides what the evaluation can measure, so it should be reported.

**Model performance.** Table 2 shows strong internal performance, with Logistic Regression slightly ahead of XGBoost. On the external set, performance at the operating point seems to collapse. Logistic Regression labels 248 of 275 legitimate URLs as phishing (FPR 90.18%, 95% Wilson interval 86.1–93.2%) and XGBoost does so for 260 of 275 (94.55%, 91.2–96.7%). Accuracy, precision, recall and F1 all exceed 0.99, however, this mainly reflects the 99.56% phishing prevalence. Mostly, predicting phishing gives 99.557% accuracy and an F1 of about 0.998, and the trained models beat this baseline by only 17 and 13 correctly classified records out of 62,087. MCC, which uses all four confusion-matrix cells and is more informative under imbalance [20], falls from 0.874 and 0.837 internally to 0.267 and 0.219 externally. PR-AUC depends on class prevalence [22], so we do not use it to compare the sets.

**Table 2.** Internal and external performance with the majority-class baseline (LR = Logistic Regression, XGB = XGBoost; accuracy and FPR in %).

| Model, set         | Acc.  | Prec. | Rec.  | F1    | FPR   | MCC   | ROC-AUC |
|--------------------|-------|-------|-------|-------|-------|-------|---------|
| LR, internal       | 93.76 | 0.955 | 0.899 | 0.926 | 3.29  | 0.874 | 0.9755  |
| XGB, internal      | 91.98 | 0.921 | 0.892 | 0.907 | 5.89  | 0.837 | 0.9669  |
| Majority, external | 99.56 | 0.996 | 1.000 | 0.998 | 100   | 0.000 | -       |
| LR, external       | 99.58 | 0.996 | 1.000 | 0.998 | 90.18 | 0.267 | 0.8605  |
| XGB, external      | 99.58 | 0.996 | 1.000 | 0.998 | 94.55 | 0.219 | 0.9669  |

Two results qualify this collapse (Figure 1). XGBoost reaches a ROC-AUC of 0.967 on the internal test set (95% CI 0.965–0.969) and 0.967 on the external set (0.957–0.976), so we find no clear evidence of a change in discrimination. Logistic Regression falls from 0.9755 to 0.8605, with non-overlapping intervals. ROC-AUC summarises ranking across all thresholds and is not affected by the class prior in the way threshold-dependent metrics are. For XGBoost, what changed is therefore the relationship between the frozen threshold and the evaluation set: the threshold was chosen where phishing prevalence was 43.6% and applied where it is 99.56%, and the surviving legitimate URLs were selected by the procedure, so their score distribution need not match that of the internal legitimate URLs. The same threshold gives an FPR of 5.9% internally and 94.5% externally although the two ROC curves are almost identical. We therefore read the collapse at the operating point as threshold transfer under a new class prior, not as a failure of the learned representation. For Logistic Regression both a loss of discrimination and threshold transfer are present. A cross-dataset evaluation that reports only the FPR at a transferred threshold cannot tell which of these explains poor results. When discrimination is preserved, re-selecting the threshold on data representative of deployment is an operational fix. When it degrades, recalibration alone cannot solve the problem.



**Figure 1.** ROC curves on the internal test set (solid) and the decontaminated external set (dashed), with the frozen decision threshold marked on each curve.

**Attribution stability.** Table 3 shows that XGBoost preserves rank order more closely on three of the four measures, while Logistic Regression has the higher top10 Jaccard index, so which model looks more stable depends on the measure. The order is unchanged on the phishing only cohort ( $p = 0.556$  for Logistic Regression and 0.923 for XGBoost).



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

**Table 3.** Cross-dataset stability of global feature attributions (balanced cohort).

| Measure                       | Logistic Regression | XGBoost |
|-------------------------------|---------------------|---------|
| Spearman $\rho$               | 0.651               | 0.918   |
| Top-10 Jaccard index          | 0.667               | 0.538   |
| Signed-direction consistency  | 0.621               | 0.793   |
| Mean absolute rank difference | 4.34                | 2.00    |

Two properties qualify the Spearman coefficient. First, nine of the 29 features receive exactly zero importance from XGBoost on both sets. They share tied ranks and add agreement without reflecting preserved discriminative behaviour. Second, preserved rank does not mean preserved meaning. Table 4 lists three XGBoost features whose mean signed contribution reverses. The most important internal feature, `has_https_scheme`, pushes predictions toward legitimate internally but keeps rank 2 externally while pushing toward phishing. A rank-based measure treats these features as stable although their effect flips, in line with earlier reports of sign changes under dataset shift [12]. For Logistic Regression, the instability is concentrated in `has_ip_address` (rank 26 to 4) and `has_explicit_port` (rank 29 to 5). These structural features are more common among the surviving external records, and XGBoost gives them zero importance on both sets. This plausibly explains part of the difference between the models, but it does not show that tree ensembles are generally more robust. We treat these comparisons as descriptive and did not run a statistical test of attribution invariance.

**Table 4.** XGBoost features whose contribution direction reverses between the internal and external sets (mean signed SHAP value).

| Feature                       | Internal rank | Internal mean | External rank | External mean |
|-------------------------------|---------------|---------------|---------------|---------------|
| <code>has_https_scheme</code> | 1             | -0.656        | 2             | +2.247        |
| <code>path_length</code>      | 3             | -0.134        | 1             | +3.506        |
| <code>path_depth</code>       | 10            | -0.022        | 5             | +0.759        |

**Implications and limitations.** Researchers running cross-dataset evaluations should report the effect of decontamination on class composition, describe the class balance and provenance of the external set, and pair threshold dependent metrics with a threshold free measure. A single rank correlation is not enough to describe attribution stability, so directional consistency should be reported and zero importance features identified. Practitioners should re-select thresholds on data representative of the intended deployment instead of transferring one chosen under a different class composition, although this cannot offset a real loss of discrimination. Our findings concern one benchmark pair. The mechanism, class asymmetric overlap from shared legitimate URL provenance, is plausible for other corpora that combine popularity based legitimate sources with independent threat feeds [18], [19], but we did not test a second pair and treat it as a hypothesis. The external legitimate class has only 275



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

records, so the intervals are wide and the balanced SHAP cohort inherits this limit. We did not evaluate the non-decontaminated pair, so this is not a before and after comparison. All features are lexical, so the findings should not be extended to detectors that use host-based, content-based or visual features. Future work should evaluate a second, independently sourced legitimate set, such as a dated Tranco snapshot, under several class priors, and apply the class-disaggregated overlap audit to other corpus pairs.

#### **4. CONCLUSION**

PhiUSIIL and LegitPhish should not be treated as independent for cross-dataset evaluation. Because, they share 36,945 URLs and 36,531 registrable domains, and every shared URL is labelled legitimate in both corpora, which is consistent with shared popularity or authority derived sources. Removing the overlap deletes 99.27% of LegitPhish's legitimate class but only 2.93% of its phishing class, leaving an external set that is 99.56% phishing instead of 62.9%. At the transferred threshold both models have very high FPRs (90.18% and 94.55%), yet the ROC-AUC of XGBoost stays at 0.967, while that of Logistic Regression falls from 0.9755 to 0.8605. For XGBoost, the collapse at the operating point therefore mostly reflects threshold transfer under a new class prior and not a loss of ranking ability. SHAP results show that rank preservation can coexist with reversed contribution direction. Neither decontamination nor explanation stability can be judged by a single headline metric: class composition, operating threshold and attribution direction should all be examined when cross-dataset results are interpreted.

#### **REFERENCES**

- [1] I. Kara, M. Ok, and A. Ozaday, "Characteristics of Understanding URLs and Domain Names Features: The Detection of Phishing Websites with Machine Learning Methods," *IEEE Access*, vol. 10, pp. 124420–124428, 2022, doi: 10.1109/ACCESS.2022.3223111.
- [2] S. S. Shafin, "An explainable feature selection framework for web phishing detection with machine learning," *Data Science and Management*, vol. 8, no. 2, pp. 127–136, Jun. 2025, doi: 10.1016/j.dsm.2024.08.004.
- [3] K. M. M. Uddin, N. Biswas, S. T. Rikta, Md. Nur -A-Alam, and R. Mostafiz, "Explainable Machine Learning for Phishing Site Detection: A High-Efficiency Approach Using Boosting Models and SHAP," *The Journal of Engineering*, vol. 2025, no. 1, Jan. 2025, doi: 10.1049/tje2.70110.
- [4] S. Henry, B. U. A. Stephen, and P. Asuquo, "Ensemble transfer learning for cross-dataset phishing detection," *Franklin Open*, vol. 15, Jun. 2026, doi: 10.1016/j.fraope.2026.100579.
- [5] D. Arp *et al.*, "Dos and Don'ts of Machine Learning in Computer Security," in *In 31st USENIX Security Symposium (USENIX Security 22)*, pp. 3971-3988. 2022., In 31st USENIX Security Symposium (USENIX Security 22), pp. 3971-3988. 2022., 2022. [Online]. Available: <https://www.usenix.org/conference/usenixsecurity22/presentation/arp>
- [6] R. Sommer and V. Paxson, "Outside the closed world: On using machine learning for network intrusion detection," in *Proceedings - IEEE Symposium on Security and Privacy*, 2010, pp. 305–316. doi: 10.1109/SP.2010.25.



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

- [7] Doyen Sahoo, Chenghao Liu, and Steven C.H. Hoi, “Malicious URL Detection using Machine Learning: A Survey,” in *Preprint*, arXiv:1701.07179., 2017, pp. 1–9. doi: 10.1145/nnnnnnn.nnnnnnn.
- [8] O. K. Sahingoz, E. Buber, O. Demir, and B. Diri, “Machine learning based phishing detection from URLs,” *Expert Syst. Appl.*, vol. 117, pp. 345–357, Mar. 2019, doi: 10.1016/j.eswa.2018.09.029.
- [9] T. Kehkashan *et al.*, “Explainable phishing website detection for secure and sustainable cyber infrastructure,” *Sci. Rep.*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-27984-w.
- [10] W. Li, S. Manickam, and Y. W. Chong, “FedPhishLLM: A privacy-preserving and explainable phishing detection mechanism using federated learning and LLMs,” *Journal of King Saud University - Computer and Information Sciences*, vol. 37, no. 8, Oct. 2025, doi: 10.1007/s44443-025-00267-0.
- [11] C. Opara, Y. Chen, and B. Wei, “Look before you leap: Detecting phishing web pages by exploiting raw URL and HTML characteristics,” *Expert Syst. Appl.*, vol. 236, Feb. 2024, doi: 10.1016/j.eswa.2023.121183.
- [12] M. Mia, D. Derakhshan, and M. M. A. Pritom, “Can Features for Phishing URL Detection Be Trusted Across Diverse Datasets? A Case Study with Explainable AI,” in *Proceedings of the 2024 11th International Conference on Networking, Systems and Security, NSysS 2024*, Association for Computing Machinery, Inc, Jan. 2025, pp. 137–145. doi: 10.1145/3704522.3704532.
- [13] A. I. Champa, M. F. Rabbi, and M. Zibran, “Battling Phish: An Empirical Comparison of ML, DL, and LLM in Phishing URL Detection Across Datasets,” in *Proceedings of the ACM Symposium on Applied Computing*, Association for Computing Machinery, Jun. 2026, pp. 1038–1047. doi: 10.1145/3748522.3779846.
- [14] M. Misiek and T. Hyla, “XGBoost-Based URL Phishing Detection Method With Cross-Dataset Validation,” *IEEE Access*, vol. 14, pp. 43200–43212, 2026, doi: 10.1109/ACCESS.2026.3672690.
- [15] P. Afonso, E. Maia, I. Amorim, and I. Praça, “Rethinking Phishing Detection: How Dataset Quality Affects Model Generalization,” in *Proceedings - International Conference on Advanced Computer Information Technologies, ACIT*, Institute of Electrical and Electronics Engineers, 2025, pp. 542–547. doi: 10.1109/ACIT65614.2025.11185656.
- [16] Rachana S. Potpelwar., U.V. Kulkarni, and J.M. Waghmare, “LegitPhish: A large-scale annotated dataset for URL-based phishing detection,” *Data Brief*, 2025, doi: 10.17632/hx4m73v2sf.2.
- [17] A. Das, S. Baki, A. El Aassal, R. Verma, and A. Dunbar, “SoK: A Comprehensive Reexamination of Phishing Research from the Security Perspective,” *IEEE Communications Surveys and Tutorials*, vol. 22, no. 1, pp. 671–708, Jan. 2020, doi: 10.1109/COMST.2019.2957750.
- [18] V. Le Pochat, T. Van Goethem, S. Tajalizadehkhoob, M. Korczyński, and W. Joosen, “TRANCO: A Research-Oriented Top Sites Ranking Hardened Against Manipulation,” in *26th Annual Network and Distributed System Security Symposium, NDSS 2019*, The Internet Society, 2019. doi: 10.14722/ndss.2019.23386.
- [19] D. Chicco and G. Jurman, “The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation,” *BMC Genomics*, vol. 21, no. 1, Jan. 2020, doi: 10.1186/s12864-019-6413-7.
- [20] S. M. Lundberg, P. G. Allen, and S.-I. Lee, “A Unified Approach to Interpreting Model Predictions,” 2017. [Online]. Available: <https://github.com/slundberg/shap>
- [21] T. Saito and M. Rehmsmeier, “The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets,” *PLoS One*, vol. 10, no. 3, Mar. 2015, doi: 10.1371/journal.pone.0118432.



6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference  
Sep 10-12, 2026, Famagusta, Northern Cyprus

---

## AIR INFILTRATION RATE MEASUREMENT OF AN OFFICE MODULE WITH TRACER GAS CONCENTRATION DECAY METHOD

Murat Özdenefe

*Department of Mechanical Engineering, Faculty of Engineering, Eastern Mediterranean University,  
Famagusta, North Cyprus, Mersin 10 TÜRKİYE*

E-mail: [murat.ozdenefe@emu.edu.tr](mailto:murat.ozdenefe@emu.edu.tr)

### Abstract

Air infiltration rate is volumetric rate of air leakage into a space through adventitious cracks and gaps in the building envelope. Typical paths for infiltration are window joints, door cracks, chimneys, ceiling light fixtures and general imperfections in the building envelope. It is an important parameter as it affects the indoor air quality, hence comfort and energy consumption of the buildings. This work is an effort to measure and evaluate the air infiltration rate of an office module using tracer gas concentration decay method.

The office module is on top floor (third floor) of a building which is located at Famagusta North Cyprus. CO<sub>2</sub> which is naturally exhaled by the occupant of the office during the working hours was used as the tracer gas. Testo 400 universal indoor air quality instrument together with CO<sub>2</sub> measuring probe and differential pressure sensor were employed for measuring the decay of the CO<sub>2</sub> concentration in ppm and outdoor-indoor pressure difference in Pa during the unoccupied hours. The simplified tracer mass balance equation i.e. tracer decay equation, was used to evaluate the infiltration rate of the office module. The measurements were done during different days having different outdoor-indoor pressure differences in order to reveal relation between the air infiltration rate and outdoor-indoor pressure difference.

It is found that the minimum infiltration rate occurred as 4.5 m<sup>3</sup>/h when the outdoor-indoor pressure difference is minimum at 0.5 Pa and maximum infiltration rate occurred as 47.7 m<sup>3</sup>/h when the outdoor-indoor pressure difference is maximum at 6.3 Pa.

**Keywords:** Infiltration, Tracer gas, Temporal evolution of CO<sub>2</sub>, Outdoor-indoor pressure difference



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

## **1. INTRODUCTION**

Air infiltration rate is volumetric rate of air leakage into a space through adventitious cracks and gaps in the building envelope under natural conditions [1]. Typical paths for infiltration are window joints, door cracks, chimneys, ceiling light fixtures and general imperfections in the building envelope [2]. It is an important parameter as it affects the indoor air quality, hence comfort and energy consumption of the buildings. This work is an effort to measure and evaluate the air infiltration rate of an office module using tracer gas concentration decay method.

Ventilation is the process of fresh air delivery to the spaces for the comfort and wellbeing of the occupants. It is also the process of thermal energy delivery by distributing the conditioned air to the spaces [1]. There are different methods for measuring the ventilation rates such as tracer gas method, pressurization method, thermography etc. The measurement technique to be selected varies depending on what will be measured e.g. infiltration, airtightness or air permeability as these are different concepts. Infiltration is the uncontrolled volumetric air flow rate into the space which occurs naturally, whereas airtightness is the volumetric air flow rate through the envelope of a building or room when subject to pressure differential of usually 50 Pa. Air permeability on the other hand is the volumetric air leakage rate per square metre of building envelope area [3,4]. Since airtightness and air permeability measurements require significant pressure differential pressurization method should be employed whereas for infiltration measurement there is not necessary to apply mechanical pressurization, hence tracer gas method is used in this study.

In tracer gas method an inert gas is released in the room and the sensors placed in the room are used to measure the temporal evolution of the concentration of the respected gas. The rate of decay of the gas is directly related to the infiltration characteristics the room [3]. In the literature there are numerous studies using this method to find the infiltration rate for spaces [5-9].

## **2. MATERIAL AND METHODS**

The office module is on top floor (third floor) of the Civil Engineering Department building which is in the campus of the Eastern Mediterranean University at Famagusta, North Cyprus. It is a typical office module out of many that exists not only in Civil Engineering Department building but also in other engineering department buildings namely, Industrial, Computer and Electrical Engineering buildings. The office module has a single exterior-facing façade, which incorporates a 1.2 m × 1.8 m

window oriented toward the northwest. It is laterally enclosed by two identical modules and connected to an internal corridor through a door. The module's floor area and volume are 16.75 m<sup>2</sup> and 50.25 m<sup>3</sup> respectively and its height from the floor to the ceiling is 3 m. Site plan of the Civil Engineering building and office module's location within the building is shown in Figure 1, whereas its plan and elevation view is illustrated in Figure 2.



Figure 1. Site plan of the Civil Engineering building and office module's location within the building.

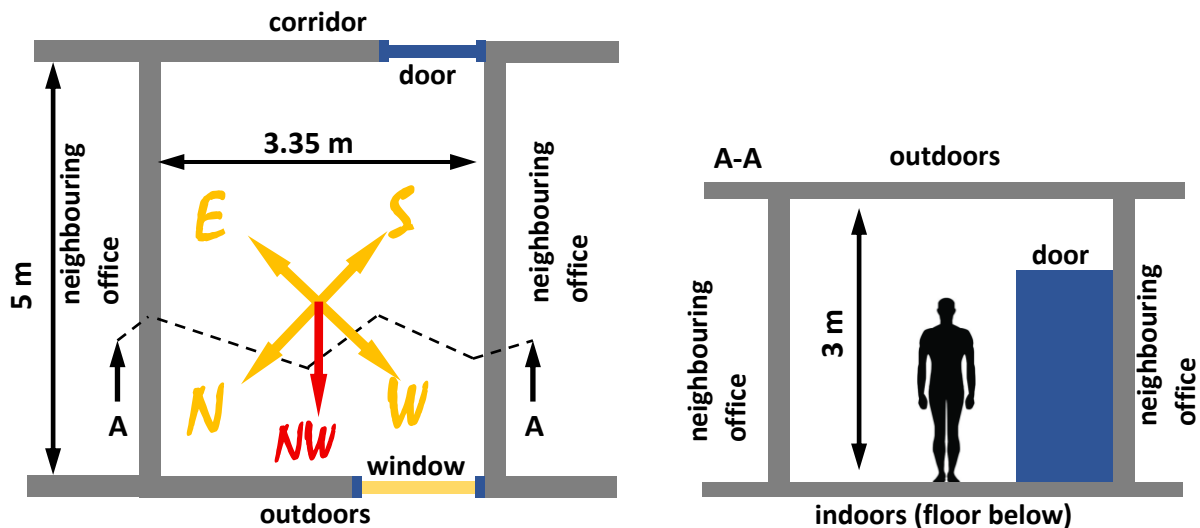


Figure 2. Plan and elevation view of the office module.

CO<sub>2</sub> which is naturally exhaled by the single occupant of the office during working hours was used as tracer gas. The occupant comes to office at about 8 AM and leaves at about 5 PM. The window and the internal door of the office module was kept shut during this period in order to enable the

accumulation of the CO<sub>2</sub>. When the occupant leaves the office the measurement and logging of the CO<sub>2</sub> concentration was initiated. Testo 400 universal indoor air quality instrument together with CO<sub>2</sub> measuring probe and differential pressure sensor were employed for measuring the decay of the CO<sub>2</sub> concentration in ppm and outdoor-indoor pressure difference in Pa during the unoccupied hours. The next day when the occupant comes to the office the measurement and the logging was stopped. Testo 400 universal indoor air quality instrument, CO<sub>2</sub> measuring probe and the tripod equipped with the Testo 400 and the probes which positioned in the office module are presented in Figure 3, whereas the probes' measuring range, accuracy and resolution are given in Table 1.



**Figure 3.** Testo 400 universal indoor air quality instrument and CO<sub>2</sub> measuring probe.

**Table 1.** Testo 400 probes' measuring range, accuracy and resolution.

| Probe                 | Measuring Range                | Accuracy                 | Resolution |
|-----------------------|--------------------------------|--------------------------|------------|
| CO <sub>2</sub> probe | 0 to 10000 ppm CO <sub>2</sub> | 50 ppm+ %3 of m.v.       | 1 ppm      |
| Differential Pressure | -100 to +200 hPa*              | ± (0.3 Pa + 1 % of m.v.) | 0.001 hPa  |
| * 1 hPa=100 Pa        |                                |                          |            |

The temporal evolution of concentration of a pollutant (CO<sub>2</sub> for this work) in a room equipped with an air purifier for CO<sub>2</sub> removal is expressed by the following differential equation [8, 9].



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

$$\frac{dC(t)}{dt} = \frac{G}{V} + N \cdot C_{ext} - N \cdot C(t) - A - S \quad (1)$$

where,  $C(t)$  is the instantaneous concentration of the  $CO_2$  ( $mg/m^3$ ),  $G$  is the  $CO_2$  generation rate ( $mg/h$ ),  $V$  is the room volume ( $m^3$ ),  $N$  is the air change rate ( $h^{-1}$ ) which is given in equation (2),  $C_{ext}$  is the outdoor air  $CO_2$  concentration ( $mg/m^3$ ),  $A$  is the rate of  $CO_2$  removal by the air purifier ( $mg/h.m^3$ ) and  $S$  is the  $CO_2$  sorption rate of the room surfaces ( $mg/h.m^3$ ).

$$N = \frac{Q}{V} \quad (2)$$

where  $Q$  is the infiltration rate in  $m^3/h$ .

Equation (1) reduces to equation (3) in the absence of an air purifier and neglecting the  $CO_2$  sorption by the room surfaces which are the conditions for the current work [8, 9].

$$\frac{dC(t)}{dt} = \frac{G}{V} + N \cdot C_{ext} - N \cdot C(t) \quad (3)$$

The room volume,  $V$ , external air's  $CO_2$  concentration rate,  $C_{ext}$  and air change rate of the room,  $N$ , for the given conditions are constant and  $G=0$  as there is not any  $CO_2$  emission during the unoccupied hours. Rearranging equation (3) and integrating it from initial time  $t=0$  where  $C=C_0$  (initial  $CO_2$  concentration) to time  $t$  where  $C=C(t)$ , equation (4), which gives the  $CO_2$  concentration at a time, is obtained [8, 9].

$$C(t) = C_{ext} + (C_0 - C_{ext}) \cdot e^{-Nt} \quad (4)$$

Equation (4) can be rewritten as below.

$$C(t) - C_{equi} = (C_0 - C_{equi}) \cdot e^{-Nt} \quad (5)$$

where,  $C_{equi}$  and  $C_0$  are equilibrium and initial concentrations of the  $CO_2$  respectively, whereas  $t$  is the time elapsed. The equilibrium concentration of the  $CO_2$  is the concentration at steady state, which is equal to the external  $CO_2$  concentration,  $C_{ext}$ , since the decay of the  $CO_2$  concentration will lead to  $C_{equi}=C_{ext}$  as  $t \rightarrow \infty$ . Note that the units of all  $CO_2$  concentration terms in equation (4) ( $C(t)$ ,  $C_{equi}$ ,  $C_0$ ) are considered as ppm whereas  $N$  is  $1/h$  and  $t$  is  $h$ .



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

Equation (5) then can be put in the form of exponential decay equation as below.

$$Y = K \cdot e^{-Nt} \quad (6)$$

where,  $Y = C(t) - C_{\text{equi}}$  and  $K = C_0 - C_{\text{equi}}$ .

It is also possible to linearize equation (6) in order to obtain linear form of the decay equation. If natural logarithm of both sides of equation (6) is taken, linearized decay equation is achieved as equation (7) which is in the form of equation of line given in equation (8).

$$\ln[C(t) - C_{\text{equi}}] = -Nt + \ln(C_0 - C_{\text{equi}}) \quad (7)$$

$$Y = mt + b \quad (8)$$

where  $Y = \ln[C(t) - C_{\text{equi}}]$  and  $b = \ln(C_0 - C_{\text{equi}})$ .

During the tests  $C(t)$  values were measured and logged in ppm at every 5 minutes. Tests were done in three different days having different average outdoor-indoor pressure differences ( $\Delta P_{\text{avg}}$ ) over the measurement period.  $C_{\text{equi}}$  which is the outdoor air's  $\text{CO}_2$  concentration rate was considered as 400 ppm which is the typical value for a non-polluted area.  $C(t) - C_{\text{equi}}$  values for each time step are evaluated and entered to Excel spreadsheet program. Subsequently the  $C(t) - C_{\text{equi}}$  vs time scatter plots were generated, exponential curves were fitted and equations in the form of equation (6) were generated for each day. Subsequently the air change rates,  $N$ , were obtained.

### 3. RESULTS AND DISCUSSION

The measurement and logging of the  $\text{CO}_2$  concentration and outdoor-indoor pressure difference was performed for three different days. The days were selected based on the  $\Delta P_{\text{avg}}$  over the measurement period. It is essential to have meaningful  $\Delta P_{\text{avg}}$  variation among different days to be able to observe the variation of  $N$  against  $\Delta P_{\text{avg}}$ . The  $\text{CO}_2$  concentration and  $\Delta P$  scatter plots, exponential curves fitted to  $\text{CO}_2$  concentration data points,  $\Delta P_{\text{avg}}$ , and equations of the fitted curves together with the  $R^2$  values are given in Figure 4, Figure 5 and Figure 6 for day 1, day 2 and day 3 respectively.

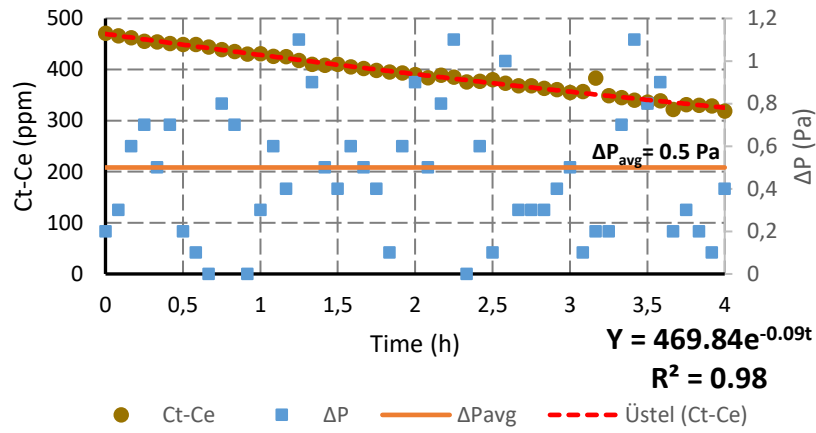


Figure 4. Results for Day 1 ( $\Delta P_{avg}=0.5$  Pa).

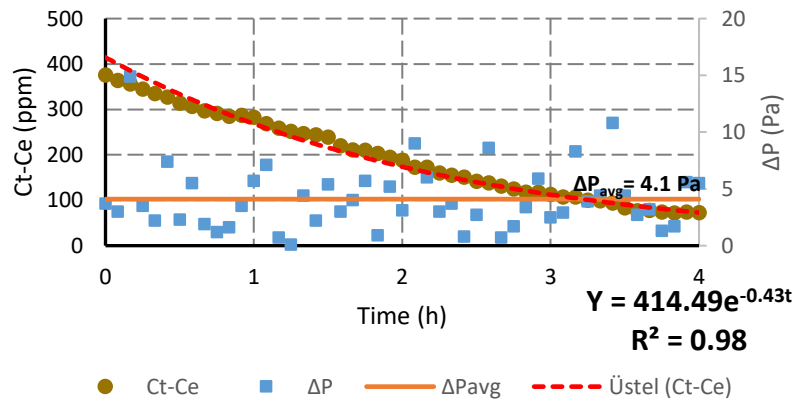


Figure 5. Results for Day 2 ( $\Delta P_{avg}=4.1$  Pa).

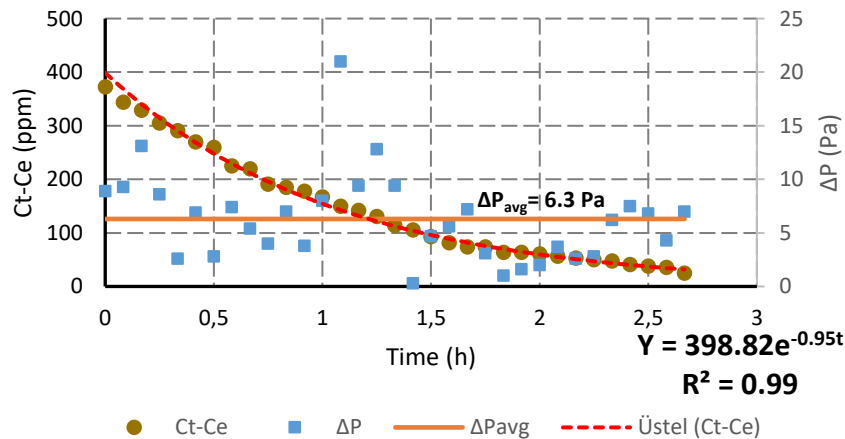
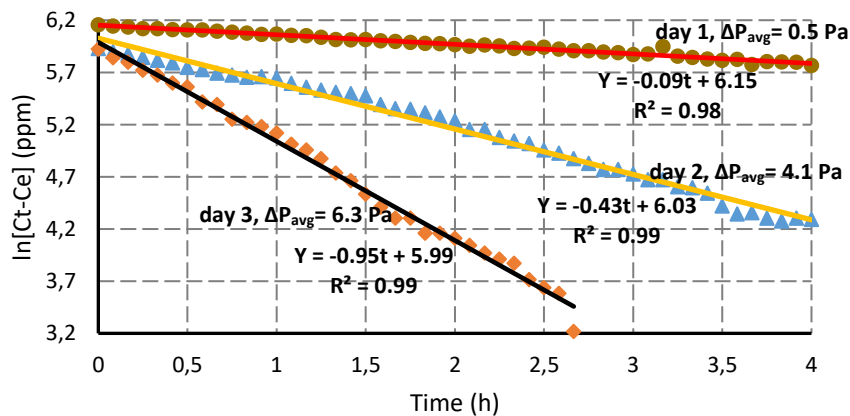


Figure 6. Results for Day 3 ( $\Delta P_{avg}=6.3$  Pa).

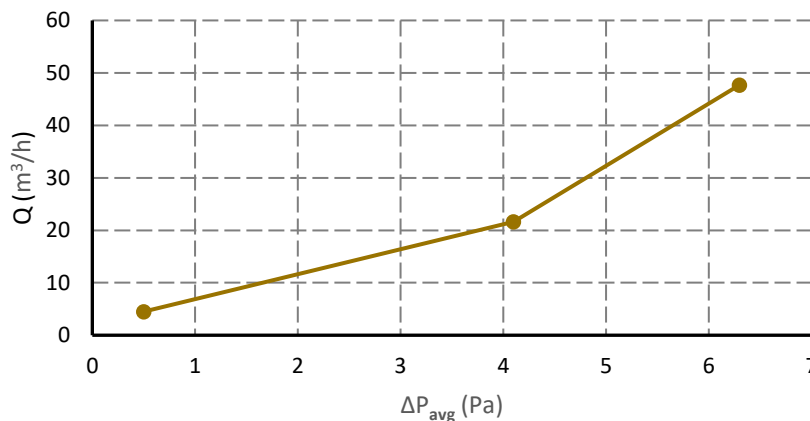


**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

It is seen in above figures that day 3 has the highest  $\Delta P_{avg}$  as 6.3 Pa, whereas day 1 has the lowest  $\Delta P_{avg}$  as 0.5 Pa. Day 2's  $\Delta P_{avg}$  lies between day 1 and day 2 with a value of 4.1 Pa. Fitted exponential curves for each day have high  $R^2$  values, very close to 1, indicating a strong match between the data points and the curve. As anticipated, highest N is obtained in day 3 when the  $\Delta P_{avg}$  is the highest whereas, the lowest N is obtained in day 1 when the  $\Delta P_{avg}$  is the lowest. It is also evident in the figures that the gradient of the CO<sub>2</sub> concentration decay curve for day 3 is maximum suggesting faster decline of CO<sub>2</sub> concentration when the  $\Delta P_{avg}$  is maximum. In contrast gradient of the CO<sub>2</sub> concentration decay curve in day 1 which has the lowest  $\Delta P_{avg}$ , is minimum demonstrating slower drop of CO<sub>2</sub> concentration when the  $\Delta P_{avg}$  is minimum. This phenomenon can also be observed in Figure 7 which presents the linearized decay curves. Figure 8 on the other hand shows the variation of Q with respect to  $\Delta P_{avg}$  and Table 2 presents the summary of the results.



**Figure 7.** Linearized decay curves.



**Figure 8.** variation of Q with respect to  $\Delta P_{avg}$ .



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

**Table 2.** Summary of the results.

| Day # | $\Delta P_{avg}$ (Pa) | Exponential decay eqn. | $R^2$ | Linear decay eqn.   | N ( $h^{-1}$ ) | Q ( $m^3/h$ ) |
|-------|-----------------------|------------------------|-------|---------------------|----------------|---------------|
| 1     | 0.5                   | $Y = 469.84e^{-0.09t}$ | 0.98  | $Y = -0.09t + 6.15$ | 0.09           | 4.5           |
| 2     | 4.1                   | $Y = 414.49e^{-0.43t}$ | 0.98  | $Y = -0.43t + 6.03$ | 0.43           | 21.6          |
| 3     | 6.3                   | $Y = 398.82e^{-0.95t}$ | 0.99  | $Y = -0.95t + 5.99$ | 0.95           | 47.7          |

#### 4. CONCLUSION

In this work infiltration rate for an office module located in a building at Famagusta, North Cyprus is measured through tracer gas concentration decay method. CO<sub>2</sub> exhaled by the occupant of the office was used as the tracer gas. Tests were done in three different days with meaningful different outdoor-indoor pressure differences in order to reveal the effect of  $\Delta P$  on infiltration. It is found that infiltration rates of the office are 4.5, 21.6 and 47.7 m<sup>3</sup>/h for  $\Delta P_{avg}$  values of 0.5, 4.1, 6.3 Pa respectively. Higher outdoor-indoor pressure corresponded to faster CO<sub>2</sub> decay leading to higher infiltration rate.

Air infiltration rate and airtightness of buildings of Cyprus are understudied hence this work will be vanguard for the larger campaign of generating the air infiltration characteristics of the buildings of Cyprus. It is intended to carry out the proposed approach for a sample of office spaces and residential buildings in order to reveal the air infiltration rate characteristics and behaviour of typical Cyprus buildings. The outcomes obviously will be used for comfort and energy studies in buildings.

#### REFERENCES

- [1] CIBSE, Ventilation and ductwork CIBSE Guide B2: 2016, The Chartered Institution of Building Services Engineers London, ISBN 978-1-906846-75-6, May 2016.
- [2] Peter S. Charlesworth, Air Exchange Rate and Airtightness Measurement Techniques-An Application Guide, Air Infiltration and Ventilation Centre, ISBN 0 946075 38 7, 1988.
- [3] CIBSE, Testing buildings for air leakage TM23: 2022, The Chartered Institution of Building Services Engineers London, ISBN 978-1-91454317-3, 2021.
- [4] Martin W Liddament, A guide to Energy Efficient Ventilation, Air Infiltration and Ventilation Centre, ISBN 0 946075 85 9, 1996.



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

- [5] Zijun Xiong, Justin Berquist, H. Burak Gunay, Cynthia A. Cruickshank, An Inquiry into the Use of Indoor CO<sub>2</sub> and Humidity Ratio Trend Data with Inverse Modelling to Estimate Air Infiltration, *Building and Environment*, 206, 2021.
- [6] Weihui Liang , Xiangnan Qiu, Comparative Analysis of CO<sub>2</sub> Tracer Gas Methods for the Measurement of Air Change Rates in Student Dormitory Rooms with Open/Closed Window Statuses, *Energy & Buildings*, 296, 2023.
- [7] C.J. Ghazi, J.S. Marshall, A CO<sub>2</sub> Tracer-Gas Method for Local Air Leakage Detection and Characterization, *Flow Measurement and Instrumentation*, 38, 72–81, 2014.
- [8] Chanhung Shim and Goopyo Hong, Airtightness Assessment under Several Low-Pressure Differences in Non-Residential Buildings, *Energies*, 16(19), 2023.
- [9] Y. Liu, P. K. Misztal, J. Xiong, Y. Tian, C. Arata, W. W. Nazaroff, A. H. Goldstein, Detailed Investigation of Ventilation Rates and Airflow Patterns in A Northern California Residence, *Indoor Air*, 28:572–584, 2018.
- [8] REHVA, Indoor Climate Quality Assessment, Federation of European Heating, Ventilation and Air-conditioning Associations, ISBN 978-2-930521-05-3, 2011.
- [9] Gameiro da Silva, M., Virtual Laboratories for A Course About Indoor Environmental Quality, 3rd International Conference on Integrity, Reliability and Failure, Porto/Portugal, 20-24 July 2009.



6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference  
Sep 10-12, 2026, Famagusta, Northern Cyprus

---

## ENERGY PERFORMANCE AND MODELLING OF A CHILLER DURING AN OFF-PEAK SEASON PERIOD

M. Nouredin<sup>1</sup>, M. Özdenefe<sup>1</sup>

<sup>1</sup>*Department of Mechanical Engineering, Faculty of Engineering, Eastern Mediterranean University, Famagusta, North Cyprus, Mersin 10 TÜRKİYE*

E-mail: MAHMOUD.N.A.MAHMOUD@emu.edu.tr  
[murat.ozdenefe@emu.edu.tr](mailto:murat.ozdenefe@emu.edu.tr)

### Abstract

This study examines how an air-cooled chiller performed during off-peak season in the library building at Eastern Mediterranean University in North Cyprus, using field measurements. This work is also an attempt to develop an early-stage calibrated energy model of the library building. Energy drawn by the building, its chiller, along with outdoor air temperature and global solar radiation, was monitored from 24 October to 23 November. Over this period, the building used 33.2 MWh of active energy, with 7.0 MWh consumed by the chiller, which operated at an average power factor of 0.65. The findings revealed a clear daily cycling pattern, frequent shutdowns, and lower peak power demand as outdoor temperatures dropped. Chiller demand also rose around midday as solar radiation increased, while analysis of selected days showed a 0.5 to 1 hour delay between peak weather conditions and peak chiller demand, reflecting the effects of building thermal inertia and control system response. The DesignBuilder model was able to capture the general operating behaviour reasonably well, including on/off cycling and shutdown periods, but it overestimated peak power, suggesting that additional calibration is still needed.

**Keywords:** Chiller, Energy monitoring, Energy modelling, Building energy efficiency



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

## **1. INTRODUCTION**

Buildings' share in the energy consumption is one of the largest in the world among the other sectors. Breakdown of the energy end-use within buildings shows that heating, cooling and air conditioning systems contributes the most to the energy consumption in buildings [1]. For hot climates usually air conditioning systems and chillers constitutes the maximum energy consumption share in the buildings.

This study focuses on the investigation of the energy consumption of the chiller, which is for the library building located at Famagusta, North Cyprus. Most of the previous research on chillers' performance and energy consumption were done during the hottest months of the year, in which energy consumption is the highest [2-4]. But in the spring and autumn, when demand is low, chillers often run at less than full capacity for long periods of time. In these situations, the energy consumption and efficiency of the system can be very different from what was planned, which can lead to energy use that is often underestimated in energy assessments. Hence this study is an effort to investigate real life performance of a chiller during the off-peak season (October-November). It is also aimed to generate the energy model of the library building and the chiller in order to initiate early-stage calibration process.

## **2. MATERIAL AND METHODS**

This study examines the off-peak performance of the air-cooled chiller in the library building at Eastern Mediterranean University. The followed approach has two tiers as: (i) building and chiller electricity consumption as well as ambient conditions measurements and recording in the field and (ii) chiller energy simulation with early-stage energy model calibration using the measurements.

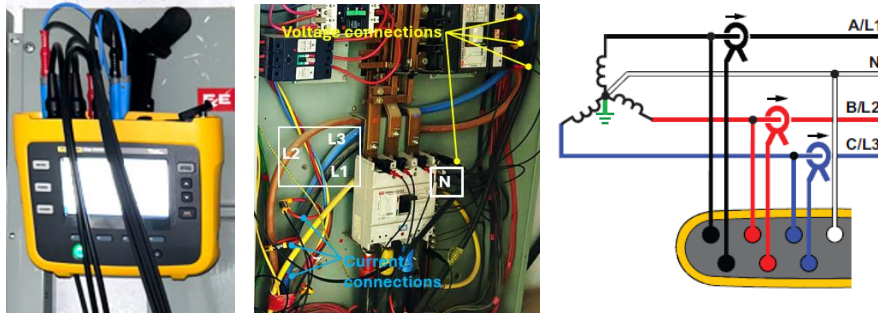
The actual energy consumption of the library building and its air-cooled chiller is measured during an off-peak period i.e. 24<sup>th</sup> October to 23<sup>rd</sup> of November. The chiller energy consumption is then analyzed considering the active energy as well as the power factor. Outdoor air temperature and global solar radiation were also recorded and their effect on energy consumption was investigated.

A whole-building energy simulation model is also generated and a model calibration attempt is done. Building energy modelling and simulation tools are usually used to check building energy consumption without the need to measure real data. This can be helpful in many cases, for example energy audits and making initial decisions about retrofitting. But the accuracy of the results depends on how well the building model is calibrated. To fully calibrate a building a large dataset and long periods

of monitoring are required, which is not possible at the beginning of the analysis. But early-stage calibration based on a small number of measurements can still provide useful insights

Measurements and recording:

The whole building and chiller electricity consumption was measured using a three-phase energy analyzer (Fluke1732/1734) with current probes (one for each phase) and voltage test leads (one for each phase and one for the neutral). The measurement was done in a three-phase star connection configuration, which is common in commercial buildings electrical systems. The energy analyzer and the way it is connected in the electrical panel are shown in Figure 1.



**Figure 1.** Energy analyzer connections.

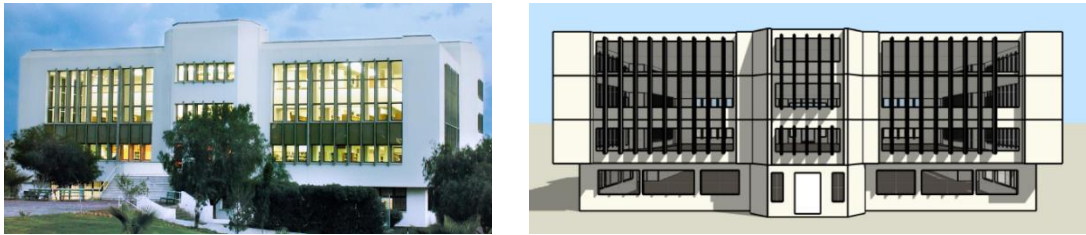
The connections were made for the building and chiller supply, with the current probe direction pointing towards the load. A 30-day continuous measurement was done during an off-peak period which was from October 24<sup>th</sup> to November 23<sup>rd</sup>. The measurement time interval was 1 minute. The dataset measured includes active, apparent, and reactive energy, which enables the assessment of energy consumption and the electrical performance (power factor). Simultaneously, the ambient temperature and global solar radiation were measured. These weather parameters were employed to assess the impact of off-peak ambient conditions on chiller energy performance.

Data Preparation and Analysis:

1-minute time series data were examined for missing data and time-stamped to synchronize the data. Energy performance was aggregated to hourly values to facilitate comparison with weather variability. The correlation between chiller energy performance and (i) outdoor air temperature and (ii) global solar radiation were assessed through scatter plots.

#### Energy Modelling and Simulation:

A comprehensive building energy simulation model was created with the DesignBuilder tool to calculate the electricity demand of the chiller. The building shape and zoning were defined from the available architectural data, while the internal loads and schedules of occupancy, lighting, and equipment were set based on the actual operating conditions of the library. The building envelope properties were introduced into the simulation model through overall heat transfer coefficients (U-values), which included 1.832 W/m<sup>2</sup>K for the external walls, 2.575 W/m<sup>2</sup>K for the roof, 1.639 W/m<sup>2</sup>K for the internal partitions, 0.705 W/m<sup>2</sup>K for the ground floor, and 0.705 W/m<sup>2</sup>K for the internal floors. The outcome of the chiller energy use in the building energy model was then compared against the actual chiller energy use on the site to verify the accuracy of the building energy model and make any necessary adjustments. The real and model view of the library building is shown in Figure 2.



**Figure 2.** Library and its model.

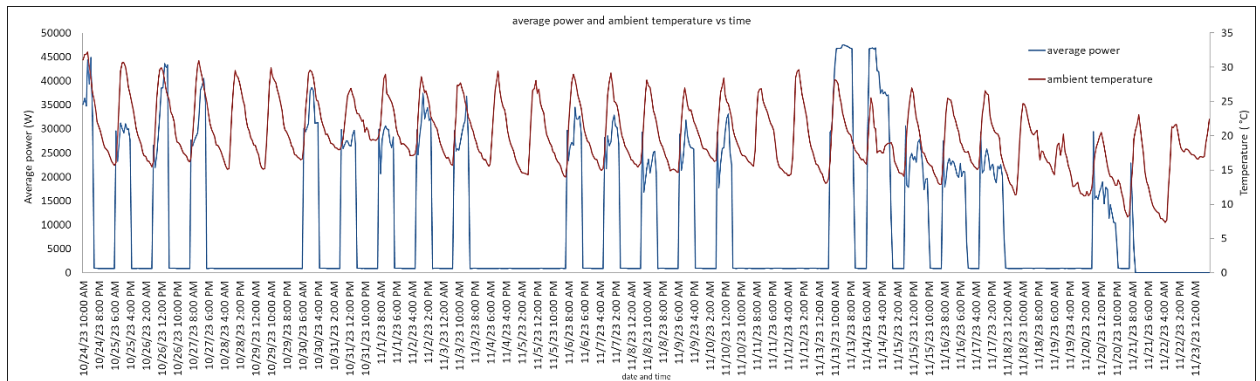
### 3. RESULTS AND DISCUSSION

During the measurement period (24<sup>th</sup> October-23<sup>rd</sup> of November), the cumulative active energy of the building was 33.2 MWh, with 43.7 MVAh apparent energy and 28.3 MVarh reactive energy, indicating an average power factor of 0.76. The cumulative active energy of the chiller was 7.0 MWh, with 10.7 MVAh apparent energy and 7.9 MVarh reactive energy, indicating an average power factor of 0.65.

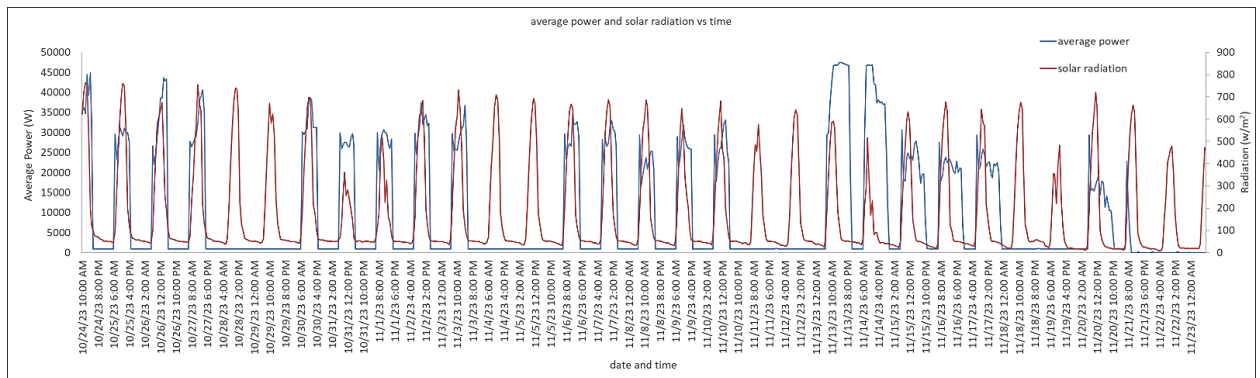
Figure 3 and Figure 4 present the chiller power consumption over the period of October 24<sup>th</sup> to November 23<sup>rd</sup> together with the ambient temperature and solar radiation. The power consumption pattern clearly indicates a daily cycle with frequent shutdown periods (working hours-nonworking hours and weekdays-weekends), which is schedule-driven and cyclic in nature. As the ambient temperature drops towards the end of November, the peak chiller power consumption also reduces, which is consistent with the weather-driven operation during the off-peak period. furthermore, the chiller power

**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

consumption also tends to peak during midday hours when the solar radiation is at its highest, which indicates that solar gains also make a significant contribution to the cooling load, even when the ambient temperature is moderate. the power consumption patterns also show that the chiller power consumption frequently cycles between on and off modes when the demand is low, which is consistent with cycling operation at part-load conditions



**Figure 3.** Average power and ambient temperature for the measurement period.

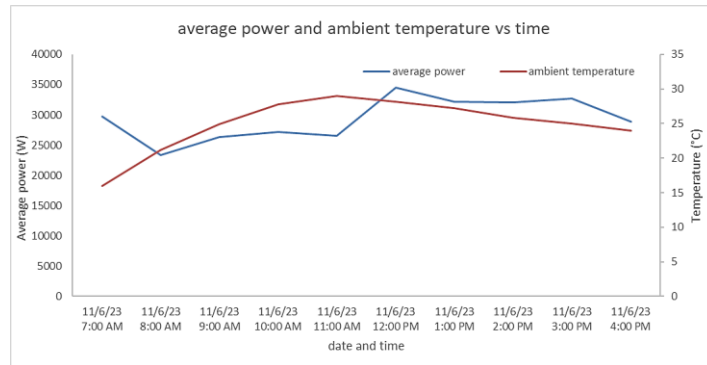


**Figure 4.** Average power and solar radiation for the measurement period.

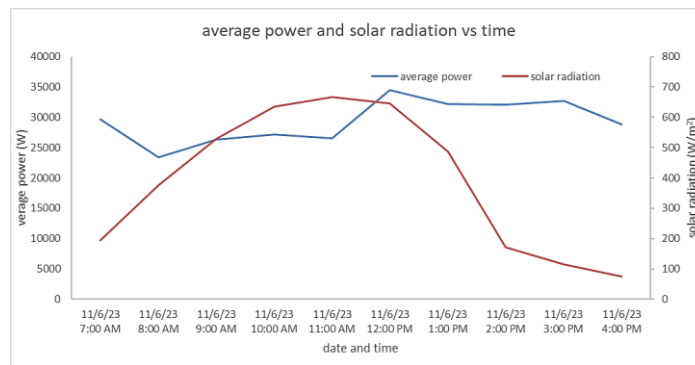
Figure 5 and Figure 6 presents the power consumption of the chiller together with the ambient temperature and solar radiation during a sample day. It is seen that the chiller peak electrical demand occurs after the peak outdoor temperature and solar radiation. Solar radiation and outdoor temperature increase throughout the morning and peak in late morning (~11:00-12:00), while chiller power peaks around noon (~12:00) and remains relatively high afterwards. This 0.5-1-hour delay is as expected because of building thermal inertia and system control response, where solar and outdoor heat gains are first absorbed by the building and then converted into cooling load with a time lag.



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

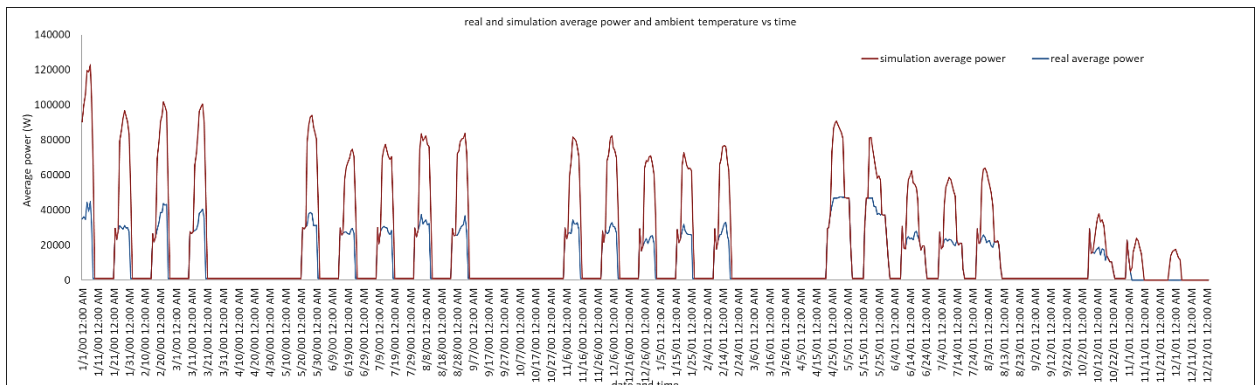


**Figure 5.** Average power and ambient temperature during a sample day.



**Figure 6.** Average power and solar radiation during a sample day.

Figure 7 illustrates the measured and simulated chiller average power. The simulation reproduces the main operational behaviour observed in the real data, including distinct on/off periods, cycling patterns, and extended shutdown intervals, indicating that the model captures the general scheduling and seasonal demand trends. However, simulated power is systematically higher than measured, with peak values often substantially overpredicted.



**Figure 7.** Real vs simulated power consumption.



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

#### **4. CONCLUSION**

In this work the off-peak season energy performance of the library building chiller observed over a 30-day period (from October 24<sup>th</sup> to November 23<sup>rd</sup>) and explored its relationship with outdoor air temperature and global solar radiation. The results of the measured data indicate a strong daily cycle with frequent shutdown periods and strong on/off cycling behaviour under low-demand conditions, which is typical of part-load operation. during the period of monitoring, the building used 33.2 MWh of active energy, with the chiller using 7.0 MWh, and the chiller operated at a relatively low average power factor of 0.65, which indicates an opportunity for electrical performance improvement during off-peak operation.

Weather impacts were observed in both the full-period and sample-day analyses. As the ambient temperature dropped towards the end of November, the peak chiller power correspondingly decreased, which indicates a strong weather-driven component even during the off-peak period. Additionally, chiller power increased around midday when solar radiation was at its highest, which indicates that solar gains have a strong impact on cooling demand even at moderate temperatures. the sample-day analysis also indicated a consistent ~0.5-1 hour lag between the peak of outdoor temperature and solar radiation and the peak of chiller electrical demand, which is consistent with the expected building thermal inertia and HVAC control system response.

a whole-building energy simulation model was created in Design Builder software and compared against actual chiller power consumption in an early-stage calibration exercise. the simulation was able to replicate the overall operating pattern (on/off cycles, tendency to cycle, and shutdown durations) satisfactorily, signifying that the model is able to follow the overall operating patterns and seasonal demand patterns. at the same time, there was a tendency for the simulation to overestimate power consumption, especially during peak periods, signifying that further calibration of loads, plant operating characteristics, and control strategies is required. Future efforts will aim to enhance the calibration process using additional plant and HVAC data (such as chilled water temperatures/flows, staging signals, and more detailed schedules).

The results of this research can help in understanding chiller performance during off-peak periods and also help in validating the early-stage calibration of energy simulation models. the proposed approach can be used by energy auditors to provide energy efficiency recommendations and estimate the potential energy savings with a higher degree of confidence.



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

**REFERENCES**

- [1] N. I. Zulkafli, M. F. Sukri, M. Mohd, A. Muhajir, and D. P. Hanak, “Performance analysis and optimisation of the chiller-air handling units system with a wide range of ambient temperature,” *Clean. Eng. Technol.*, vol. 14, no. May, p. 100643, 2023, doi: 10.1016/j.clet.2023.100643.
- [2] A. Kasman, W. Ruslan, and W. Basuki, “Performance Analysis of Chiller System,” vol. 14, no. 13, pp. 3017–3021, 2019.
- [3] B. S. Rizi, M. Heidarinejad, G. Pavlak, V. Cushing, and W. Hederman, “Evaluating the Performance of Chiller Plant Efficiency Using Random Forest Model : A High-Rise Building Case Study Illinois Institute of Technology , Chicago , IL , 60616 , USA The Pennsylvania State University , University Park , PA , USA QCoefficient Inc ., Chicago , IL , USA,” pp. 234–242, 2022.
- [4] Y. Zhang, H. Li, and S. Wang, “Energy performance analysis of multi-chiller cooling systems for data centers concerning progressive loading throughout the lifecycle under typical climates,” pp. 1693–1708, 2024.



6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference  
Sep 10-12, 2026, Famagusta, Northern Cyprus

---

## INFLUENCE of SILVER (Ag) DOPING on the CHARGING-DISCHARGING PROCESSES of MWCNTs ANODES in Li-ion BATTERIES

Ülkü Kıymaz <sup>1</sup>, Fatma Sarf <sup>2</sup>

<sup>1</sup>Energy Resources and Management, Graduate School, Çanakkale Onsekiz Mart University,  
Çanakkale, TÜRKİYE

<sup>2</sup>Çanakkale Onsekiz Mart University, Çan Vocational School, Çanakkale, TÜRKİYE

E-mail: [ulkukiyamaz@gmail.com](mailto:ulkukiyamaz@gmail.com)

### Abstract

This study investigates the electrochemical performance of silver (Ag)-doped multi-walled carbon nanotubes (MWCNTs) as anode materials for lithium-ion batteries. Pristine MWCNTs and MWCNTs doped with 12 wt% and 32 wt% Ag were comparatively evaluated. The electrodes were prepared using 85% active material, 5% conductive carbon black (Super P), and 10% polyvinylidene fluoride (PVDF), with N-methyl-2-pyrrolidone (NMP) used as the solvent. The electrochemical performance of the prepared electrodes was evaluated through charge–discharge tests. Preliminary electrochemical results showed that the pristine MWCNT sample exhibited a specific capacity of approximately 80 mAh/g, while the Ag-doped samples showed lower specific capacities. The observed decrease in performance may be associated with the Ag distribution and agglomeration on MWCNTs walls, electrode design and electrochemical interfacial processes (such as SEI forming). Further evaluation of the structural, morphological, and electrochemical results will provide a more comprehensive understanding of the effect of Ag doping on the performance of MWCNTs anodes.

**Keywords:** Li-ion battery, MWCNTs, silver doping, anode material, electrochemical performance



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

## **1. INTRODUCTION**

A lithium-ion battery or Li-ion battery is a type of rechargeable battery that uses the reversible intercalation of lithium ions ( $\text{Li}^+$ ) into electronically conducting solids to store energy. Li-ion batteries (LIBs) are recognized as one of the most promising rechargeable batteries due to their long-life span, abundant resources, low toxicity, high thermal stability and high energy density [1]. Traditional LIBs consist of lithium transition metal oxide cathode and graphite anode ( $342 \text{ mAhg}^{-1}$ ); both are typical rigid materials. High-performance from LIBs requires arrange of all components, including electrode (anode and cathode), electrolyte, and separator. Among these components, advanced anode design is so important to improve charging-discharging process [2].

Working principle based on that a lithium-ion battery is fully discharged, all lithium ( $\text{Li}^+$ ) ions are in the cathode while the battery has in charge situation which is connected to an external power source. Lithium ( $\text{Li}^+$ ) ions (which are positively charged) leave the cathode and move through the electrolyte to the anode. The lithiation/delithiation ( $\text{Li}^+/\text{Li}$ ) in the active materials (anode or cathode) is theoretically reversible [3].

Main parameters of Li-ion batteries are defined as energy density, specific capacity and cycle life. They typically exhibit a gravimetric energy density of  $150\text{--}250 \text{ Wh/kg}$  and a volumetric energy density of  $300\text{--}700 \text{ Wh/L}$ . Their specific capacity generally ranges from 250 to  $7000 \text{ mAh/g}$ , depending on the application and electrode materials used. In terms of cycle stability, the cycle life of a Li-ion battery is defined as the total number of charge–discharge cycles that can be completed before its capacity decreases to a specified percentage, typically  $70\text{--}80\%$  of its initial value. Under conventional operating conditions, Li-ion batteries generally provide approximately 300–500 charge cycles, corresponding to an average service life of around 2–3 years depending on usage and maintenance conditions [4-5].

The theoretical lithium intercalation capacity of graphite is  $372 \text{ mAh/g}$  and it can reach  $330\text{--}370 \text{ mAh/g}$  and everly decreases in the long-term cycle [6]. Therefore, researchers have resorted to composite materials, which combine two or more different components to take advantage of their complementary qualities, to improve specific capacity. This approach includes integrating conductive carbon matrices with redox-active materials to improve electron transport while reducing structural degradation, or mixing carbon nanotubes with porous materials to form hierarchical porous networks



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

that enhance electrolyte penetration and ion diffusion, thereby boosting both capacitance and rate performance [7].

As an anode material, multi walled carbon nanotubes (MWCNTs) have remarkable properties such as high electrical conductivity, abundant in nature and mechanical stability which arises from electron delocalization throughout the tubular structure. However, pristine MWCNTs anodes suffer from extremely high (1000 mAh/g) irreversible capacity in the first formation cycle, which prevents them from entering the market. In this study, heavily silver doping (12% and 32%) effect on the electrochemical performance of MWCNTs and their usage in Li-ion batteries as anode material were investigated.

## **2. MATERIAL AND METHODS**

In this study, pristine MWCNTs (purity of  $96 \geq$ ) and Ag doped (12 wt% and 32 wt%) MWCNTs (purity of  $99 \geq$ ) were purchased in Nanografi Company-Ankara/ Turkey. In Table 1, structural parameters of pristine and Ag:MWCNTs were shown.

**Table 1.** Structural parameters of pristine and Ag-doped MWCNTs

|                                                       | <b>D (nm)</b> | <b>Average inside parameter (nm)</b> | <b>Average outer parameter (nm)</b> | <b>Specific Surface Area (m<sup>2</sup>/g)</b> |
|-------------------------------------------------------|---------------|--------------------------------------|-------------------------------------|------------------------------------------------|
| <b>Pristine MWCNTs</b>                                | 13            | 5-10                                 | 8-18                                | 220                                            |
| <b>12% Ag:MWCNTs</b>                                  | 20            | 5                                    | >50                                 | 20-25                                          |
| <b>32% Ag: MWCNTs<br/>(to be fixed in the future)</b> | 20            | 5                                    | >50                                 | 20-25                                          |

To obtain slurry and then electrode, similar process was applied in this study as shown in Figure 1. A homogeneous slurry was realized by mixing the active material, conductive additive, binder, and solvent in appropriate proportions. In this study, 85% active material, 5% conductive carbon black (Super P), and 10% polyvinylidene fluoride (PVDF), with N-methyl-2-pyrrolidone (NMP) were used as a solvent to prepare slurry. Then, the prepared slurry was coated onto a metal foil, which serves as the current collector of the electrode. After the coating process, the electrode was dried to remove surface impurities at 80 °C. The drying process affects the porosity of the electrode, the distribution of the binder, and the formation of conductive pathways. The electrode was cut into according to half-cell configuration dimensions and battery components were added and half-cell Li-ion battery was closed.

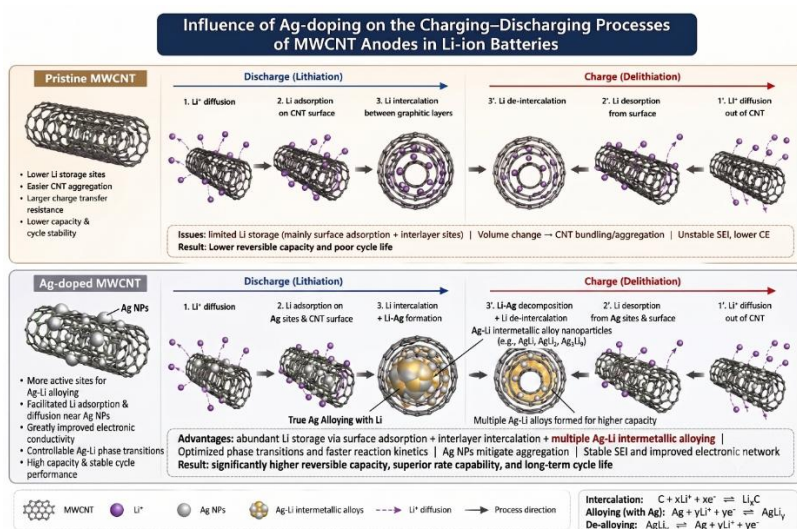
The capacity values were calculated based on the total mass of the active materials. Finally, initial charge-discharge tests were investigated for 8 cycle.



**Figure 1.** Standart slurry and electrode preparation process [8]

### 3. RESULTS and DISCUSSION

As shown in Figure 2, theoretical investigation was presented. With increasing the surface area for lithium charge-discharge process and emerging synergistic interactions causes a superior electrochemical property as compared to pristine MWCNTs in 12% Ag doped MWCNTs electrodes [9]. However, unfortunately, in 32% Ag: MWCNTs, the adverse effect of the agglomerative silver formations on the MWCNTs walls is noteworthy for lithium diffusion and hence SEI forming might be observed.

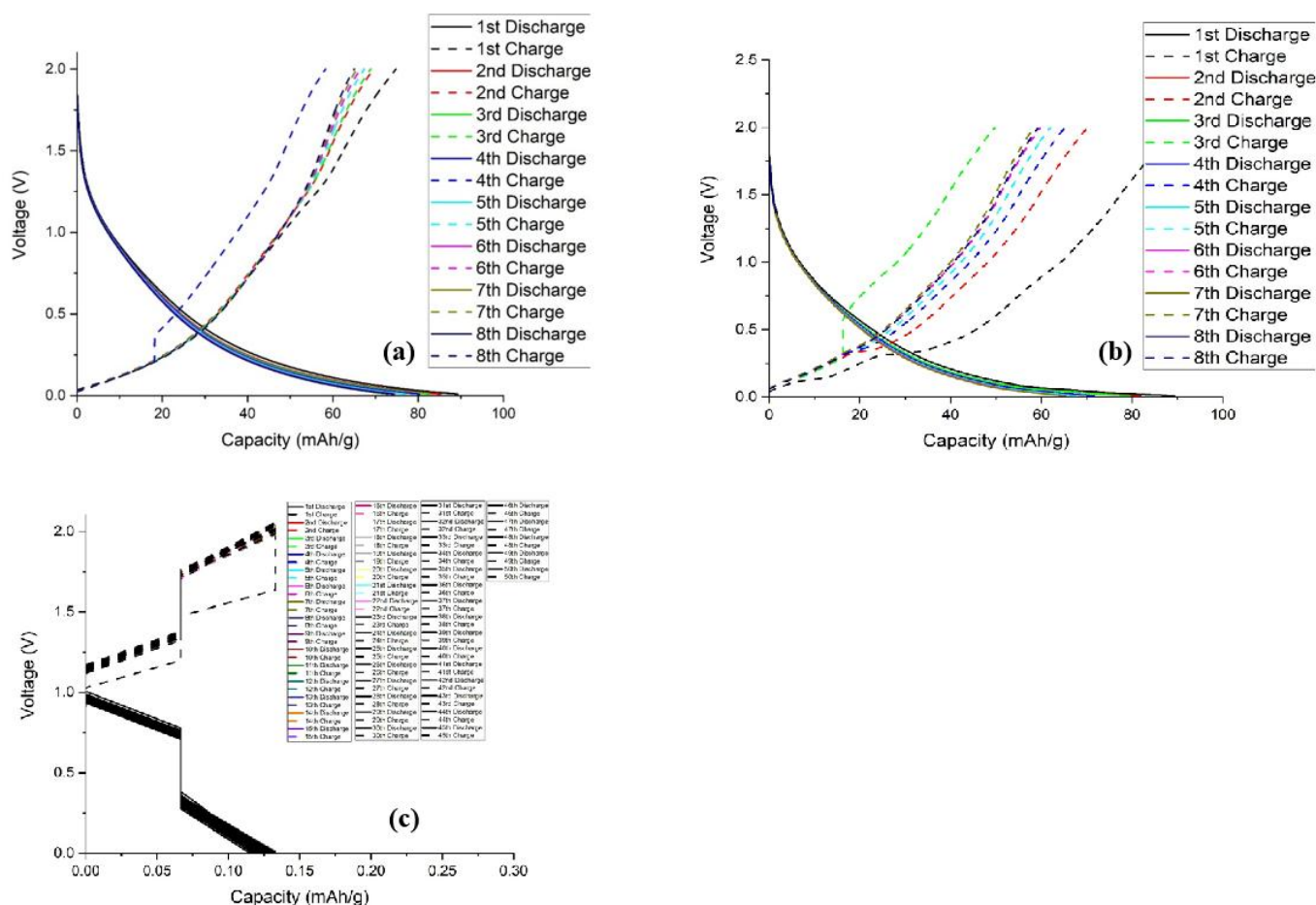


**Figure 2.** Possible charge-discharge process of pristine MWCNTs and Ag: MWCNTs (12% and 32%) anode in Li-ion batteries (*AI was used in this figure*)

In the literature, CNT–Ag hybrid structures have been essential components to obtain high-performance from Li-ion batteries. Among in this approach, the integration of noble metal nanostructures with carbon-based materials has emerged as a promising strategy with significant potential. In particular, silver nanowires (Ag-NWs) have attracted considerable attention owing to their excellent electrical conductivity, mechanical flexibility, and ability to participate in lithium alloying reactions. Furthermore, the optimized Ag-NWs/MWCNTs composite exhibits remarkable electrochemical performance, achieving a high specific capacity of 817 mAh g<sup>-1</sup>, superior rate capability of 423 mAh g<sup>-1</sup> at 5 C, and outstanding cycling stability with a capacity retention of 91.2% after 500 cycles [9].

As shown in Figure 3, the preliminary results obtained from the charge–discharge tests for 8 cycles indicate that the pristine MWCNT sample exhibited a specific capacity of approximately 80 mAh/g. This value is considerably lower compared to graphite-based anode materials. The specific capacity decreased further in the 12% Ag MWCNTs sample, while the 32% Ag MWCNTs sample exhibited a significant decrease in electrochemical performance. The observed decrease in performance may be associated with several factors, including a reduction in the electrochemically active surface area due to the formation of the SEI (Solid Electrolyte Interphase) layer, agglomeration of Ag particles that may limit active sites, or a possible experimental issue during the electrode preparation process

[10]. However, further evaluation of the SEM, XRD, FTIR, and Raman analyses, together with the electrochemical tests for longer cycles, is required to clarify the underlying causes of this performance decrease.



**Figure 3.** Specific capacity/voltage versus of pristine MWCNTs and Ag:MWCNTs for 8 cycles

#### 4. CONCLUSION

In this work, heavily silver doping effect on the electrochemical properties on MWCNTs walls was investigated. Ag concentration was arranged set as 12% and 32%. Structural properties of all samples and SEM image of pristine MWCNTs were explained. Experimental and theoretical results were compared to evaluate the electrochemical properties. The electrochemical performance of the Ag:MWCNTs decreased with increasing Ag concentration along 8 cycles. Then, these electrodes will be evaluated through charge–discharge for more than 100 cycles.



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

For a proper interpretation of the current results for 8 cycles , it is important to determine the chemical forming, particle size, and accumulation of Ag<sup>+</sup> ions on the MWCNTs walls should be investigated using these techniques; SEM-EDX, FTIR and XRD. Future plan of this study is machine learning applying to analyze the electrochemical results and investigate a relationship between the electrode and AI system.

### ACKNOWLEDGMENT

This study was supported by Scientific Project Unit of Çanakkale Onsekiz Mart University under project number: FYL-2026-5675.

### REFERENCES

- [1] N.K. Gautam, A. Alshoaibi et.al. Fabrication and characterization of MWCNT/MXene nanocomposites as an electrode material for supercapacitor applications, *Diamond and Related Materials*, 160,2025,113046.
- [2] Tong Z, Lv C, Bai G-D, Yin Z-W, Zhou Y, Li J-T. *A review on applications and challenges of carbon nanotubes in lithium-ion battery. Carbon Energy.* 2025; 7:e643.
- [3] Namdar, H., Manteghian, M., Jafari, A. *et al.* Novel synthesis method of multi walled carbon nanotube-silica Janus nanostructures. *Sci Rep* **15**, 19525 (2025).
- [4] Alawi, A., Saeed, A., Sharqawy, M. H., & Al Janaideh, M. (2025). A comprehensive review of thermal management challenges and safety considerations in lithium-ion batteries for electric vehicles. *Batteries*, 11(7), 275.
- [5] Pesaran, A. A. (2011). Choices and Requirements of Batteries for EVs, HEVs, PHEVs.
- [6] Ghanooni Ahmadabadi, V., Rahman, M. M., & Chen, Y. (2023). A study on high-rate performance of graphite nanostructures produced by ball milling as anode for lithium-ion batteries. *Micromachines*, 14(1), 191.
- [7] Wu, L., Li, Y., Fu, Z., & Su, B. L. (2020). Hierarchically structured porous materials: synthesis strategies and applications in energy storage. *National science review*, 7(11), 1667-1701.
- [8] Zhang, E. Understanding wet electrode processing in lithium batteries. LinkedIn. [https://www.linkedin.com/pulse/understanding-wet-electrode-processing-lithium-batteries-echo-zhang-zm0mc?trk=public\\_post\\_feed-article-content](https://www.linkedin.com/pulse/understanding-wet-electrode-processing-lithium-batteries-echo-zhang-zm0mc?trk=public_post_feed-article-content)



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

- [9] Yazdanypoor, M., & Khani, A. A. M. (2025). Enhanced lithium-ion battery performance using silver nanowire-decorated multi-walled carbon nanotube composite anodes. *International Journal of Electrochemical Science*, 101193.
- [10] Paravannoor, A., Panoth, D., & Pattathil, P. (2020). Silicon anode design for Li ion batteries: Synergic effects of Ag nanoparticles and ionic liquid electrolytes. *Chemical Engineering Journal Advances*, 4, 100037.



6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference  
Sep 10-12, 2026, Famagusta, Northern Cyprus

## COMPARATIVE EVALUATION OF ARTIFICIAL INTELLIGENCE-BASED MODELS FROM ACTIVATED CARBON PRODUCTION TO ADSORPTION PERFORMANCE

Zhala Alizada

*Department of Environmental Engineer, Sakarya University, Sakarya, TÜRKİYE*

E-mail: [zhala.zada2@ogr.sakarya.edu.tr](mailto:zhala.zada2@ogr.sakarya.edu.tr)

### Abstract

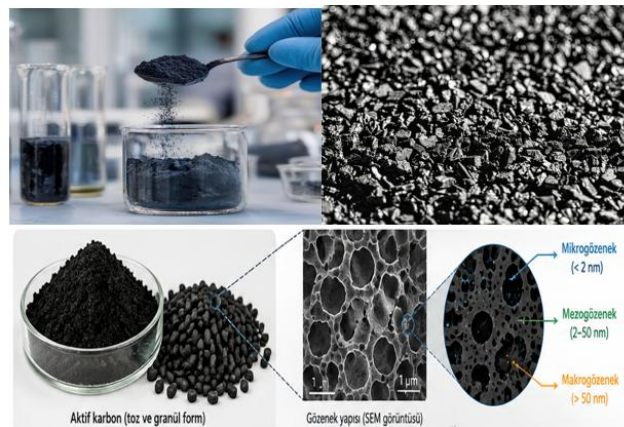
Activated carbon is one of the most widely used carbon-based adsorbents owing to its high specific surface area, well-developed micro- and mesoporous structure, and abundant surface functional groups, making it an indispensable material for water and gas purification applications. The adsorption performance of activated carbon is strongly influenced by production parameters, including the precursor material, activation method, activation temperature, processing time, and chemical activating agents. Because of the complex and nonlinear interactions among these parameters, conventional statistical approaches are often insufficient to accurately describe and predict adsorption behavior. In recent years, artificial intelligence (AI) and machine learning (ML) algorithms have emerged as powerful data-driven tools for modeling activated carbon production processes and predicting adsorption performance. Approaches such as Artificial Neural Networks (ANN), Random Forest (RF), Gradient Boosting (GB), Support Vector Machines (SVM), and Explainable Artificial Intelligence (XAI) have demonstrated remarkable capability in learning complex relationships from experimental data, enabling accurate prediction of adsorption efficiency while identifying the most influential process variables. This review comprehensively examines recent studies employing artificial intelligence-based models for predicting activated carbon production characteristics and adsorption performance. The published studies are comparatively evaluated in terms of datasets, input parameters, model architectures, performance metrics ( $R^2$ , RMSE, and MAE), prediction accuracy, and application fields. Furthermore, the strengths and limitations of different AI models are critically discussed regarding process optimization and the efficient design of laboratory investigations. The primary objective of this review is to provide a comprehensive assessment of the relationship between activated carbon production processes and adsorption performance through the perspective of artificial intelligence-based modeling while highlighting recent advances and emerging research trends in the literature. The novelty of this review lies in the systematic comparison of various AI algorithms within a unified evaluation framework and in addressing both activated carbon production and adsorption processes simultaneously. The comparative analysis indicates that studies involving Explainable Artificial Intelligence (XAI), hybrid machine learning models, and real-time digital twin-based optimization remain relatively limited. Therefore, future research should focus on developing physically interpretable, explainable, and industrially deployable artificial intelligence models capable of supporting intelligent optimization and decision-making in activated carbon adsorption systems.

**Keywords:** Activated carbon; adsorption; artificial intelligence; machine learning; adsorption performance prediction; process modeling; explainable artificial intelligence (XAI)

## 1. INTRODUCTION

Activated carbon is a carbon-based adsorbent characterized by a highly developed porous structure, large specific surface area, and diverse surface functional groups. It is produced from carbon-rich precursor materials through thermochemical processes such as carbonization and physical or chemical activation. These structural and chemical characteristics enable activated carbon to be widely used for the removal of pollutants from aqueous and gaseous systems (Marsh & Rodríguez-Reinoso, 2006). In recent years, biomass and waste-derived materials have attracted increasing attention as activated carbon precursors because they provide opportunities for waste valorization and the development of relatively economical and sustainable adsorbents (Ioannidou & Zabaniotou, 2007; Tan et al., 2015)

The properties of activated carbon are strongly affected by precursor composition, carbonization conditions, activation method, activating-agent type, activation temperature and time, and impregnation conditions. These parameters determine the specific surface area, pore volume, pore-size distribution, surface chemistry, and yield of the resulting carbon material (Ioannidou & Zabaniotou, 2007; Yahya et al., 2015). However, adsorption performance cannot be explained by surface area alone. Pore-size distribution, micropore and mesopore volumes, pore accessibility, surface functional groups, and adsorbate characteristics may also influence adsorption through mechanisms such as electrostatic interactions, hydrogen bonding, surface complexation, and pore filling (Marsh & Rodríguez-Reinoso, 2006; Thommes et al., 2015). Therefore, activated carbon performance represents a multidimensional relationship between production conditions, material properties, and adsorption conditions.



**Figure 1.** Activated carbon and representative porous structures illustrating the relationship between carbon structure and adsorption behavior



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

The simultaneous effects of numerous production and adsorption variables make activated carbon systems challenging to model using conventional approaches. Machine learning (ML) methods provide an alternative by enabling complex and nonlinear relationships among multiple variables to be modeled simultaneously. Production variables such as precursor type, carbonization and activation conditions, activating-agent type, and impregnation ratio can be related to material properties, while adsorption variables such as pH, temperature, initial adsorbate concentration, adsorbent dosage, and contact time can be incorporated to predict adsorption performance (Foo & Hameed, 2010; Wang & Guo, 2020).

Recent studies have demonstrated the potential of ML for both activated carbon production and adsorption modeling. Zou et al. (2024) applied Gradient Boosting Regression (GBR) to literature-derived data on activated carbon production from lignocellulosic biomass and showed that activation temperature and impregnation conditions were important factors affecting material properties. Similarly, Long et al. (2024) compiled 3,674 data points from 151 publications on heavy-metal adsorption and compared Random Forest, Gradient Boosting Decision Tree, and XGBoost models. Their results demonstrated the potential of ML for predicting adsorption capacity, while SHAP analysis was used to identify influential variables (Long et al., 2024).

Beyond prediction, Explainable Artificial Intelligence (XAI) methods such as SHAP provide an opportunity to interpret model outputs from a physicochemical perspective. Such approaches can help identify which production and adsorption variables contribute most strongly to predicted performance and can therefore establish a connection between data-driven modeling and conventional adsorption concepts. Importantly, model performance may vary depending on the dataset, target variable, input features, and validation strategy; therefore, no single ML algorithm can be assumed to be universally optimal for all activated carbon systems. Accordingly, this study aims to comparatively and critically examine artificial intelligence- and machine learning-based approaches applied from activated carbon production to adsorption performance using current literature data. The study focuses on model performance, input variables, validation strategies, influential factors, and the relationships among production conditions, physicochemical properties, and adsorption performance from an integrated AI perspective.



## 2. MATERIAL AND METHODS

### 2.1 Research Design and Methodology

This study was conducted as a literature-based comparative investigation of artificial intelligence (AI) and machine learning (ML) applications in activated carbon production and adsorption. No activated carbon synthesis or new adsorption experiments were performed. Instead, previously published experimental and computational studies were reviewed and compared in terms of input variables, ML algorithms, validation strategies, performance metrics, and model-interpretability approaches.

The literature analysis focused on three main questions: (i) which production, physicochemical, and adsorption variables are commonly used as ML inputs; (ii) which AI/ML algorithms are applied to predict activated carbon properties and adsorption performance; and (iii) which variables are identified as influential through feature-importance and explainable artificial intelligence (XAI) methods. The evaluated variables included precursor type, carbonization and activation conditions, activating agent, and impregnation ratio, together with physicochemical properties such as specific surface area, pore volume, micropore volume, average pore diameter, and yield. For adsorption applications, pH, initial adsorbate concentration, adsorbent dosage, temperature, and contact time were considered. Commonly reported algorithms included artificial neural networks (ANN), random forest (RF), gradient boosting (GB/GBDT), XGBoost, and support vector machine (SVM). Model performance was primarily assessed using  $R^2$ , RMSE, and MAE. Where available, training-validation-test partitioning, cross-validation, and hyperparameter optimization were also considered. Model interpretation was evaluated through feature-importance analyses and XAI methods, particularly SHAP, to identify influential variables and relate model predictions to the physicochemical characteristics of activated carbon and adsorption systems

Overall, the methodological framework followed the sequence production conditions → physicochemical properties → adsorption conditions → ML prediction → model evaluation and interpretation, providing a structured basis for comparing AI-based approaches across activated carbon research.



## 2.2 Literature-Derived Data and Target Variables

The data analyzed in this study were derived from experimental and modeling datasets reported in previously published literature on AI- and ML-based applications in activated carbon production and adsorption. The reviewed studies included variables related to precursor characteristics, carbonization and activation conditions, physicochemical properties, and adsorption operating parameters. Prediction targets included specific surface area, pore volume, activated carbon yield, adsorption capacity, and removal efficiency.

Prediction targets were considered when comparing model performance because production-related properties and adsorption responses reflect different aspects of activated carbon systems. Therefore, reported model performance was interpreted in relation to dataset characteristics, input variables, experimental conditions, and prediction objectives. Long et al. (2024) provided a representative literature-based dataset comprising 3,674 data points from 151 studies on heavy-metal adsorption using carbonaceous adsorbents. RF, GBDT, and XGBoost models were compared, with GBDT achieving an  $R^2$  of 0.9707 and an RMSE of 0.1420. SHAP analysis was also applied to identify influential input variables (Long et al., 2024).

Selected literature-based ML applications are summarized in Table 1, including the carbonaceous material, prediction target, input variables, algorithms, evaluation metrics, and reported performance. The table provides the basis for the subsequent comparison of ML approaches, predictive performance, and influential variables.



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

**Table 1.** Selected examples of machine learning applications on activated carbon and carbon-based adsorbents reported in the literature.

| Study                  | Application                       | Data                       | Models                          | Best Result                            |
|------------------------|-----------------------------------|----------------------------|---------------------------------|----------------------------------------|
| Long et al. (2024)     | Heavy metal adsorption            | 3,674 records / 151 papers | RF, GBDT, XGB                   | <b>GBDT: <math>R^2 = 0.9707</math></b> |
| Sudarsan et al. (2026) | Congo Red adsorption              | 180 experiments            | SVM, ANN, ANFIS, MLR            | Model comparison                       |
| Maleki et al. (2026)   | CO <sub>2</sub> adsorption        | Experimental + modeling    | GB + other ML                   | <b>GB: <math>R^2 = 0.9989</math></b>   |
| Lamsiah et al. (2026)  | Heavy metal adsorption            | 1,528 records              | RF, ET, GBR, XGBR               | <b>GBR: <math>R^2 = 0.9648</math></b>  |
| Shetty et al. (2026)   | Fixed-bed chlorpyrifos adsorption | Experimental               | SVR, XGB, RF, GB, CB, LGBM, ANN | <b>SVR: <math>R^2 = 0.9968</math></b>  |

Overall, the literature-derived framework adopted in this study provides a structured basis for comparing AI-and ML-based approaches across different stages of activated carbon research. The comparison emphasizes the relationships among input variables, prediction targets, algorithm selection, validation procedures, predictive performance, and model interpretability. Importantly, the analysis does not assume that a single algorithm is universally applicable to all activated carbon systems. Instead, reported findings are interpreted in relation to the specific datasets, experimental contexts, prediction objectives, and methodological configurations under which the models were developed.

### **2.3 Data Preprocessing, Model Validation, and Performance Evaluation**

Literature-derived datasets were processed through data cleaning, unit standardization, missing-value and duplicate checks, and categorical-variable encoding. Model validation included training-validation-test partitioning and k-fold cross-validation, while hyperparameter optimization was considered where reported (Breiman, 2001; Cortes & Vapnik, 1995; Haykin, 2009).

Model performance was primarily evaluated using  $R^2$ , RMSE, and MAE. Model interpretability was examined using feature-importance methods and SHAP, which identifies the magnitude and direction of individual feature contributions to predictions (Lundberg & Lee, 2017). These methods support the



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

identification of influential process and physicochemical variables in activated carbon and adsorption systems.

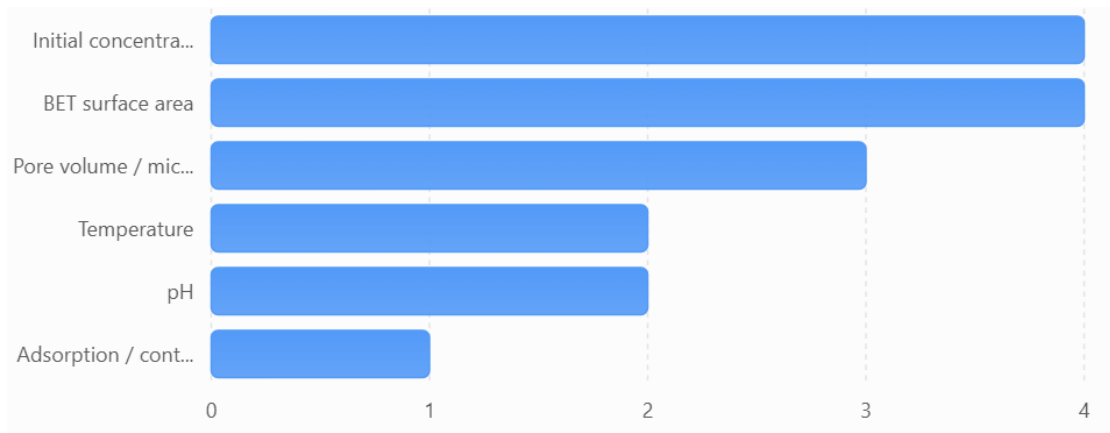
### **3. RESULTS AND DISCUSSION**

#### **3.1 Machine Learning Applications from Activated Carbon Production to Adsorption Performance**

The reviewed studies demonstrate that ML has been applied to both activated carbon production and adsorption performance. At the production stage, models mainly predict physicochemical properties, whereas adsorption studies incorporate both adsorbent characteristics and operating conditions. Zou et al. (2024) used Gradient Boosting Regression (GBR) to predict specific surface area, pore volume, and yield from lignocellulosic biomass. Reported  $R^2$  values ranged from 0.76 to 0.91, with activation temperature and soaking conditions identified as influential variables. At the adsorption stage, Long et al. (2024) compared RF, GBDT, and XGBoost using 3,674 observations from 151 studies. GBDT achieved an  $R^2$  of 0.9707 and RMSE of 0.1420, while SHAP analysis was used to identify influential variables. Recent studies further reported high predictive performance for CO<sub>2</sub> uptake, heavy-metal adsorption, and dynamic breakthrough behavior. Maleki et al. (2026) reported  $R^2 = 0.9989$  using Gradient Boosting for CO<sub>2</sub> adsorption, while Lamsiah et al. (2026) reported  $R^2 = 0.9648$  for heavy-metal adsorption using GBR. Shetty et al. (2026) achieved  $R^2 = 0.9968$  with SVR for fixed-bed breakthrough prediction and identified contact time as an important variable. Overall, the findings indicate that ML performance varies with the prediction target, dataset, input variables, and adsorption system. The principal characteristics of the selected studies are summarized in Table 2.

**Table 2.** Selected literature-based machine learning applications in activated carbon production and adsorption research.

| <b>Study</b>          | <b>Best Model</b> | <b><math>R^2</math></b> |
|-----------------------|-------------------|-------------------------|
| Zou et al. (2024)     | GBR               | 0.91                    |
| Long et al. (2024)    | GBDT              | 0.9707                  |
| Maleki et al. (2026)  | GB                | 0.9989                  |
| Lamsiah et al. (2026) | GBR               | 0.9648                  |
| Shetty et al. (2026)  | SVR               | 0.9968                  |
| Shetty et al. (2026)  | SVR               | 0.9968                  |



**Figure 2.** Frequency of key variables identified as influential in the reviewed machine-learning-based adsorption.

### 3.2. Comparative Evaluation of Model Performance and Influential Variables

The reported ML performance varies considerably among studies because of differences in datasets, prediction targets, adsorbents, experimental conditions, input variables, and validation procedures. Therefore, reported  $R^2$  values should be interpreted within the context of each individual study rather than used to rank algorithms.  $R^2$  values ranged from 0.91 for production-related prediction by Zou et al. (2024) to 0.9989 for CO<sub>2</sub> adsorption prediction by Maleki et al. (2026), while Long et al. (2024), Lamsiah et al. (2026), and Shetty et al. (2026) reported  $R^2$  values of 0.9707, 0.9648, and 0.9968, respectively.

Ensemble-based methods such as Gradient Boosting and GBDT were frequently applied because they can capture nonlinear relationships. However, the strong performance of SVR reported by Shetty et al. (2026) indicates that model suitability depends on the dataset and prediction problem.

For adsorption applications, initial adsorbate concentration and BET surface area were repeatedly identified as influential variables. Pore volume, pH, temperature, and contact time were also important, although their contributions varied according to the adsorption system. SHAP analysis can further reveal the magnitude and direction of feature contributions; however, these results describe model behavior and should not be interpreted as direct causal relationships.



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

### **3.3 From Production Conditions to Adsorption Performance: Integrated Discussion**

The reviewed literature supports a production–structure–performance relationship. Precursor characteristics and activation conditions influence BET surface area, pore structure, and surface chemistry, which subsequently interact with adsorption conditions such as pH, concentration, temperature, dosage, and contact time.

Activation conditions may produce nonlinear effects on pore development, surface area, and yield. ML models are therefore useful for representing complex interactions when sufficient and representative data are available. However, model applicability depends on the adsorption configuration, such as batch, gas, kinetic, or fixed-bed systems. From an engineering perspective, literature-based ML models can support process screening and experimental planning by identifying potentially important variables and parameter combinations. Nevertheless, predictions remain dependent on data quality and coverage and should be experimentally validated before practical process optimization.

Overall, AI and ML provide opportunities to connect production conditions, activated carbon properties, and adsorption performance, while XAI methods such as SHAP improve the interpretation of these relationships.

## **4. CONCLUSION**

This comparative literature study demonstrates that artificial intelligence (AI) and machine learning (ML) have substantial potential as data-driven tools for understanding and predicting activated carbon production and adsorption performance. The reviewed studies collectively show that ML approaches can capture complex and nonlinear relationships among precursor characteristics, production conditions, physicochemical properties, and adsorption parameters. These capabilities have been demonstrated for the prediction of specific surface area, pore volume, activated carbon yield, adsorption capacity, and dynamic breakthrough behavior.

A key finding of the comparative assessment is that ML model performance is strongly dependent on the prediction target, dataset characteristics, input-variable structure, and validation strategy. Although Gradient Boosting-based approaches demonstrated high predictive performance in several of the reviewed studies, strong results were also obtained using SVR and other algorithmic approaches under specific experimental and data conditions. Therefore, the reviewed evidence does not



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

indicate a universally superior algorithm. Instead, algorithm selection should be determined by the characteristics of the dataset, the prediction objective, and the experimental system.

The comparative analysis further highlights the importance of identifying influential variables rather than evaluating models solely according to predictive accuracy. Initial adsorbate concentration and BET surface area were repeatedly identified as influential variables in adsorption-oriented studies, while pore volume, temperature, pH, and contact time showed importance depending on the specific adsorption system. At the production stage, activation temperature and soaking or impregnation conditions were identified as important factors affecting activated carbon properties. These findings support a process–structure–performance perspective, in which production conditions influence material characteristics that subsequently determine adsorption behavior under defined operating conditions.

Another important outcome is the role of explainable artificial intelligence (XAI) in improving the scientific interpretation of ML models. Techniques such as SHAP can provide information about the relative contribution and direction of individual variables, thereby helping to connect data-driven predictions with measurable physicochemical and process characteristics. Consequently, the value of ML in activated carbon research extends beyond obtaining high statistical performance; it also lies in its ability to generate interpretable information that can support scientific analysis and experimental planning.

Overall, this study emphasizes that the future development of AI-assisted activated carbon research should focus on integrating predictive accuracy, model validation, interpretability, physicochemical understanding, and experimental verification. Literature-derived ML models can be used to screen process conditions, identify promising experimental regions, and reduce unnecessary experimental combinations; however, their predictions remain dependent on the quality and representativeness of the available data. Future research should therefore prioritize standardized and larger databases, consistent reporting of experimental variables, independent validation, uncertainty assessment, and physics-informed or hybrid ML approaches.

In conclusion, the principal contribution of AI and ML in activated carbon research is not simply the achievement of a higher  $R^2$  value, but the development of a transparent and validated framework that connects activated carbon production conditions with material properties and, ultimately, adsorption performance. Such an integrated process–structure–performance approach can provide a stronger basis for data-driven process optimization and experimentally testable material design.



**6<sup>th</sup> International Natural Science, Engineering and Material Technologies Conference**  
**Sep 10-12, 2026, Famagusta, Northern Cyprus**

---

phosphor being at 501 K have a general order kinetic model. A whole knowledge of the thermoluminescence processes, determining the kinetic parameters is worth further investigation in order to design the most suitable dosimeter suitable for UVB.

## REFERENCES

- Foo, K. Y., & Hameed, B. H. (2010). Insights into modeling of adsorption isotherm systems. *Chemical Engineering Journal*, 156(1), 2–10. <https://doi.org/10.1016/j.cej.2009.09.013>
- Ioannidou, O., & Zabaniotou, A. (2007). Agricultural residues as precursors for activated carbon production—A review. *Renewable and Sustainable Energy Reviews*, 11(9), 1966–2005. <https://doi.org/10.1016/j.rser.2006.03.013>
- Lamsiah, S., Benhadj, M., Alouiz, I., Bouhadda, M., Amarouch, M. Y., Mazouzi, D., & Sennoune, M. (2026). Prediction of heavy metal adsorption by activated carbon using machine learning. *Artificial Intelligence Chemistry*, 4(1), 100117. <https://doi.org/10.1016/j.aichem.2026.100117>
- Long, X., Huangfu, X., Huang, R., Liang, Y., Wu, S., & Wang, J. (2024). The application of machine learning methods for prediction of heavy metal by activated carbons, biochars, and carbon nanotubes. *Chemosphere*, 354, 141584. <https://doi.org/10.1016/j.chemosphere.2024.141584>
- Maleki, A., Mehdizad, A., Hassani, E., Naveshki, M., Zalakinezhad, A., Rashidi, A., Farhadian, A., Hassanvand Somarin, D., & Aghazadeh, F. (2026). Machine learning-driven design of high-performance activated carbons for enhanced CO<sub>2</sub> capture. *Journal of CO<sub>2</sub> Utilization*, 105, 103336. <https://doi.org/10.1016/j.jcou.2026.103336>
- Marsh, H., & Rodríguez-Reinoso, F. (2006). *Activated carbon*. Elsevier.
- Shetty, S., Srinivasan, N. R., Vinayagam, R., & Selvaraj, R. (2026). Machine learning-assisted fixed-bed column adsorption of chlorpyrifos using sustainable biomass-derived activated carbon. *npj Clean Water*. <https://doi.org/10.1038/s41545-026-00618-y>
- Thommes, M., Kaneko, K., Neimark, A. V., Olivier, J. P., Rodríguez-Reinoso, F., Rouquerol, J., & Sing, K. S. W. (2015). Physisorption of gases, with special reference to the evaluation of surface area and pore size distribution (IUPAC Technical Report). *Pure and Applied Chemistry*, 87(9–10), 1051–1069. <https://doi.org/10.1515/pac-2014-1117>
- Yahya, M. A., Al-Qodah, Z., & Ngah, C. W. Z. (2015). Agricultural bio-waste materials as potential sustainable precursors used for activated carbon production: A review. *Renewable and Sustainable Energy Reviews*, 46, 218–235. <https://doi.org/10.1016/j.rser.2015.02.051>
- Zou, R., Yang, Z., Zhang, J., Lei, R., Zhang, W., Fnu, F., Tsang, D. C. W., Heyne, J., Zhang, X., Ruan, R., & Lei, H. (2024). Machine learning application for predicting key properties of activated carbon produced from lignocellulosic biomass waste with chemical activation. *Bioresource Technology*, 399, 130624. <https://doi.org/10.1016/j.biortech.2024.130624>