

Article

RadonFAN: Intelligent Real-Time Radon Mitigation Through IoT, Rule-Based Logic, and AI Forecasting

Lidia Abad * , Fernando Ramonet , Margarita González , José Javier Anaya  and Sofía Aparicio 

Institute of Physical and Information Technologies “Leonardo Torres Quevedo” (ITEFI), Consejo Superior de Investigaciones Científicas (CSIC), C/Serrano 144, 28006 Madrid, Spain

* Correspondence: lidia.abad@csic.es

Abstract

Radon (Rn-222) is a major indoor air pollutant with significant health risks. This work presents RadonFAN, a low-cost IoT system deployed in two galleries at the Institute of Physical and Information Technologies (ITEFI-CSIC, Madrid), integrating distributed sensors, microcontrollers, cloud analytics, and automated fan control to maintain radon concentrations below recommended limits. Initially, ventilation relied on a reactive, rule-based mechanism triggered when thresholds were exceeded. To improve preventive control, two end-to-end deep learning models based on regression-to-classification (R2C) and direct classification (DC) are developed. A quantitative analysis of predictive performance and computational efficiency is reported. While the R2C model is hindered by the inherent behavior of the time series, the DC model achieves high classification performance (recall > 0.975) with low computational cost (<4 million parameters, 7 million FLOPs). Modifications to the DC model are studied to identify potential performance bottlenecks and the most relevant components, showing that most limitations arise from feature richness and time series behavior. When evaluated against the existing rule-based ventilation system, the DC model reduces both unsafe radon exposure events and energy consumption, demonstrating its effectiveness for preventive radon mitigation.

Keywords: radon monitoring; indoor air quality; low-cost sensors; IoT; ventilation control; time series; deep learning

1. Introduction

Radon (Rn-222) is a radioactive gas produced by the decay of Radium-226, derived from naturally occurring Uranium-238 in soil, rocks, and some building materials. Odorless, colorless, and tasteless, radon is undetectable without specialized equipment [1]. While outdoor radon disperses rapidly and generally poses minimal risk, poorly ventilated indoor spaces can allow dangerous radon accumulation [2]. Historic buildings are especially vulnerable to such accumulation due to construction characteristics and limited ventilation, highlighting the need for monitoring [3,4]. The resulting exposure represents a significant public health concern, as inhaled radon and its decay products damage DNA in lung epithelial cells, making residential exposure the second leading cause of lung cancer after smoking [5,6]. Even levels below international reference limits can increase risk, especially for smokers. The European Union and the World Health Organization recommend reference levels of 300 Bq/m³ and 100 Bq/m³, respectively [2,7]. Long-term monitoring shows relatively low year-to-year variability in radon levels, indicating that properly designed measurement campaigns can reliably characterize exposure for risk assessment purposes [8].



Academic Editor: Arslan Munir

Received: 19 December 2025

Revised: 27 January 2026

Accepted: 4 February 2026

Published: 11 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

Since exposure to radon indoors is common, yet preventable, long-term monitoring and predictive modeling are essential public health tools. They allow early detection of elevated radon levels and support targeted mitigation measures, such as improved ventilation, to reduce exposure [9].

Low-cost sensors are a compelling alternative to conventional air-quality monitoring stations due to their significantly reduced cost, ease of deployment, and fair performance under diverse conditions. Although they generally exhibit lower accuracy compared to reference-grade instruments and often require periodic calibration and compensation for environmental interferences [10], their affordability enables broader spatial and temporal coverage [11]. The development of Internet of Things (IoT) technologies has further facilitated their integration into distributed systems capable of real-time data acquisition and remote access. Building on these developments, several systems have been designed to monitor indoor radon concentrations [12–14].

Statistical time-series models have traditionally served as a foundation for forecasting environmental variables, relying on assumptions such as stationarity and linear temporal dependencies and providing interpretable baseline predictions through autoregressive and moving-average formulations [15,16]. While effective in simple settings, these approaches often struggle to represent the complex dynamics that characterize indoor radon time series. In operational mitigation systems, control decisions are frequently implemented through rule-based strategies that encode expert knowledge using fixed thresholds and heuristic logic. Although such systems are transparent and easy to deploy, they are inherently reactive, difficult to scale, and unable to adapt to evolving data patterns [17].

Recent advances in Artificial Intelligence (AI), particularly in deep learning (DL), offer data-driven alternatives capable of learning nonlinear temporal dependencies directly from time-series data. DL architectures, including recurrent [18], convolutional [19], and attention-based models [20,21], have demonstrated strong performance in environmental forecasting tasks by capturing both short- and long-range temporal dependencies [22,23]. For pollutant forecasting, Long Short-Term Memory (LSTM)-based architectures are widely employed to predict future concentrations of environmental contaminants [24–26]. Extensions incorporating attention mechanisms, convolutional layers, or hybrid combinations with traditional signal processing techniques have been proposed to improve predictive accuracy, often at the cost of increased computational complexity [22]. Similar methods have also been explored in the context of indoor radon forecasting [27,28].

While most traditional DL methods primarily focus on forecasting continuous values, there is an increasing interest in time series classification, where models directly predict categorical events [29]. In radon monitoring, classification can be used to identify hazardous radon levels, converting continuous forecasts into actionable alerts. However, in this context, DL classification models are still in the early stages of exploration, with most progress coming from anomaly detection [30]. Moreover, most existing studies still treat regression (forecasting pollutant concentrations) and classification (exceedance detection) as separate tasks, and relatively few integrate classification directly from forecasted values within a unified model [31].

Despite progress in both radon monitoring and forecasting, existing systems remain fragmented. Radon-focused works typically emphasize either continuous concentration prediction [27,28] or monitoring and threshold-based mitigation [12–14], but rarely unify these functionalities within a real-time IoT framework. This gap underscores the novelty of this study, which integrates low-cost IoT sensing, DL-based forecasting, and real-time mitigation actions into a unified system designed specifically for radon risk reduction.

This work extends a previous investigation [9] and presents RadonFAN, a low-cost, IoT-enabled system for monitoring and mitigating indoor radon while minimizing energy

use. Deployed in two service underground galleries at the Institute of Physical and Information Technologies (ITEFI-CSIC, Madrid) after elevated radon concentrations were detected, RadonFAN integrates low-cost radon sensors within a distributed architecture that enables real-time data acquisition, remote supervision, and automated ventilation control. In its initial deployment, RadonFAN was governed by a rule-based control strategy, in which a microcontroller activated ventilation fans using fixed parameters. This approach enabled rapid and interpretable mitigation, but was inherently reactive, required repeated manual hyperparameter tuning, and did not adapt to the temporal dynamics of radon fluctuations, limiting its transferability across buildings or galleries. As a result, the rule-based system produced both unnecessary fan activations and delayed responses under certain conditions. These limitations motivated the development of an AI-based control strategy intended to substitute the original rule system. This study investigates time-series classification as a decision-oriented alternative to conventional regression-based forecasting for radon mitigation. Two end-to-end DL approaches are considered. This work builds on a previously deployed IoT-based indoor radon monitoring and rule-based mitigation system, which is here evaluated by comparing it to new AI-driven approaches. It reformulates radon mitigation from regression-based forecasting to a time-series classification problem for preventive control, and compares classification pipelines within the existing system.

2. Materials and Methods

2.1. System Deployment

The RadonFAN system is an IoT-based architecture for monitoring and mitigating radon levels, deployed in two enclosed basement galleries at ITEFI-CISC. Operating from October 2019 to December 2024, the system was developed to maintain radon concentrations below a threshold through automated ventilation control [9]. Figure 1 illustrates the layout of the two service galleries, which are geometrically symmetric. The total internal volume of each gallery is approximately 26 m^3 . Each gallery is equipped with a mechanically forced ventilation system consisting of two air supply fans or inlets (Shenzhen Hongguan Mechatronics Co., Ltd., Shenzhen, China) and one air extraction fan or outlet (Maico Italia S.p.A., Lonato del Garda, Italy). The fans provide a nominal airflow rate of $195 \text{ m}^3/\text{h}$ and have an electrical power rating of less than 100 W each. Basement service galleries, connected to the wastewater infrastructure, exhibited elevated radon levels, which infiltrated upper floors and prompted active ventilation-based mitigation.

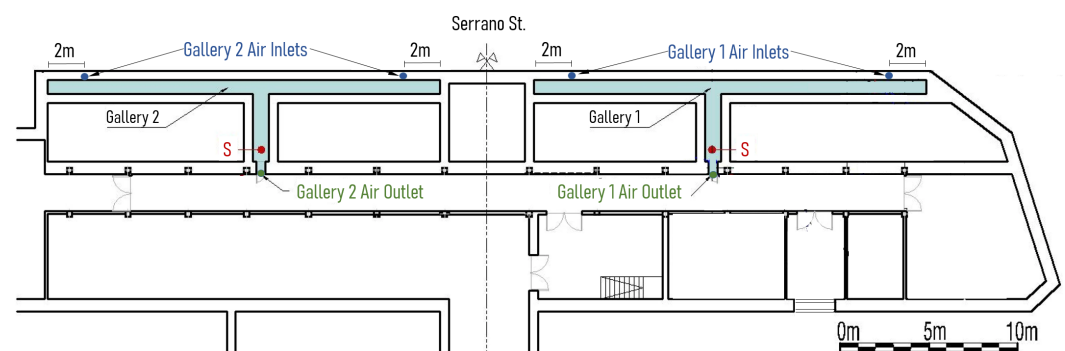


Figure 1. Layout of the two galleries with their corresponding radon sensor (S), two air inlets and air outlet placements.

The architecture of RadonFAN comprises multiple hardware, communication and cloud computing elements. The physical layer incorporates sensors and fans in each gallery, which are controlled by an Arduino (Arduino S.r.l., Monza, Italy) and a Raspberry Pi (Raspberry Pi Ltd., Cambridge, UK). The communication system between devices is based

on the Message Queuing Telemetry Transport (MQTT) protocol and a Wi-Fi network. The ThinkSpeak platform is also employed for cloud communication. Figure 2 shows the configuration of the system.

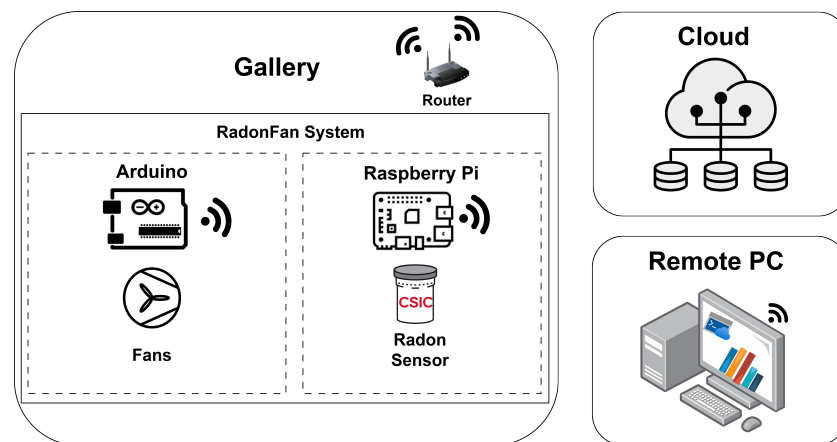


Figure 2. System architecture diagram. The gallery hosts the RadonFan system, composed of fans connected to an Arduino and a radon sensor connected to a Raspberry Pi. The Raspberry and Arduino are connected via a router providing Wi-Fi access.

The sensing layer employs different detectors in each gallery. Gallery 1 uses an FTLab RD200M radon sensor (FTLab Co., Ltd., Seoul, Republic of Korea) with a measurement range of 0–100,000 Bq/m³, accuracy $\pm 10\%$ at 2000 Bq/m³, precision $\pm 10\%$ at 370 Bq/m³. Gallery 2 uses an FTLab RS9A radon sensor (FTLab Co., Ltd., Seoul, Republic of Korea) with a measurement range of 0–50,000 Bq/m³, accuracy $\pm 15\%$ at 1000 Bq/m³, precision $\pm 15\%$ at 370 Bq/m³. Both sensors sample radon every ten minutes and deliver data to the gallery-specific Raspberry Pi via USB. The actuation system includes three fans per gallery, driven by an Arduino MKR1000 through a 3 V opto-isolated relay shield connected to General Purpose Input/Output (GPIO) pin 1 (Jiatong Technology Innovation (Shenzhen) Co., Ltd., Shenzhen, China). Fan operation follows a predefined algorithm. Initially, the system employs a rule-based control logic (Section 2.3), which defines a RadonFan subsystem referred to as RadonFAN-RS. Subsequent efforts aim to enhance this algorithm through the application of AI-based methods, resulting in the development of RadonFan-DC and RadonFAN-R2C (Section 2.4).

The communication infrastructure centers on a Raspberry Pi 3B+, which serves as the system hub. During startup, it configures network connectivity by sending Wi-Fi credentials to the Arduino access points using HTTP POST requests. An MQTT architecture manages device communication, using the Paho MQTT library (v2.1.0) on client devices and a Mosquitto v1.6 broker on the Raspberry Pi to handle message routing. For cloud connectivity, the Raspberry Pi sends radon and fan-state data to ThingSpeak via HTTP POST, enabling remote visualization through customizable dashboards with real-time and historical information. Locally, the system stores timestamped logs for each sensor and gallery, with automated daily rotation for efficient data management.

2.2. Dataset

2.2.1. Dataset Description

The dataset used in this study covers the period from October 2019 to December 2024. It exhibits temporal discontinuities due to system-related interruptions, including network connectivity issues, hardware modifications, and multiple rounds of hyperparameter tuning and experimental adjustments during the early deployment phases. The available dataset contains over 74,000 entries for Gallery 1 (G1) and 57,000 entries for Gallery 2 (G2),

with measurements recorded every ten minutes. For each gallery, the data are organized into two files: one containing records from the Raspberry Pi and the other from the Arduino-based sensors.

For evaluating the initial rule-based system, data from 17 April to 21 June 2024 were used, corresponding to a period governed exclusively by that control logic. This subset guided preprocessing and model-design decisions. For developing the AI systems and comparing the rule-based and AI-driven approaches, data from 10 January to 18 March 2024 were selected, a period without fan activation or control interventions. This focus is necessary because other periods were dominated either by interruptions such as administrative network changes or construction works, or by repeated trials aimed at tuning the rule-based system parameters. This underscores the potential value of automated, AI-driven approaches that can adapt to temporal dynamics and improve preventive control.

Table 1 summarizes radon measurement statistics (Bq/m^3) for the whole dataset (Oct 2019–Dec 2024), the subset used for the rule system evaluation (Apr–Jun 2024) and the subset used for the AI model development and evaluation (Jan–Mar 2024). Overall, G2 exhibits higher mean and median radon levels than G1, although its standard deviation and maximum values are generally lower, except during the Jan–Mar 2024 period. Radon concentrations decrease during rule system/fan operation (Apr–Jun 2024), with lower values in G2 than G1; however, maximum values in both galleries still exceed recommended limits [2,7]. During the AI period, G2 shows substantially higher levels and variability, more than double those observed in G1.

Table 1. Summary statistics of radon measurements for Gallery 1 (G1) and Gallery 2 (G2) across different periods.

Gallery	Period	RS System	Mean	Median	Std	Min	Max
G1	Oct 2019–Dec 2024	Mixed	136.94	111.00	136.33	10.73	1895.14
	Apr–Jun 2024	On	114.98	103.60	66.11	12.58	639.73
	Jan–Mar 2024	Off	103.13	87.32	77.79	10.73	551.30
G2	Oct 2019–Dec 2024	Mixed	148.65	125.00	120.99	7.00	1187.00
	Apr–Jun 2024	On	109.97	104.00	58.19	12.00	359.00
	Jan–Mar 2024	Off	218.41	145.00	212.71	7.00	1187.00

These differences may be partly explained by the position of the galleries: G1 is located on the north side of the building, and G2 on the south. Even though both galleries face west and are otherwise symmetrical, local differences in sunlight exposure, temperature gradients, and wind-driven ventilation can lead to distinct radon accumulation patterns. Other contributing factors may include differences in soil radon emission and the degree of sealing of floors and walls. Together, these factors help explain why G2 experiences higher average radon levels and greater variability during certain periods.

Missing values in these periods caused by sensor faults were imputed using a k -nearest neighbors method (kNN, $k = 5$), which estimates each missing value based on the most similar neighboring time points. This method preserves short-term temporal patterns and is particularly effective when data are irregularly spaced or contain intermittent gaps. In contrast, simpler strategies such as mean or median imputation replace missing values with constant estimates, ignoring local dynamics, while linear interpolation assumes uniform spacing and may smooth over sudden fluctuations. These characteristics make kNN a practical choice for maintaining the natural variability of radon time series when handling missing data.

2.2.2. Preliminary Analysis of Time Series

To evaluate the suitability of the radon time series for forecasting, the Augmented Dickey–Fuller (ADF) and the Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests were ap-

plied to both the raw (X) and first-differenced (ΔX) data at multiple seasonal lags (12, 36, 144, 360). ADF and KPSS tests are employed to assess whether a series exhibits constant mean and variance over time [32]. The results are summarized in Table 2. Across all tested lags, the findings are consistent: for the raw series, the ADF test rejects the unit-root hypothesis (p -value < 0.05), suggesting the absence of pure unit-root behavior typical of a random walk, whereas the KPSS test indicates non-stationarity. This implies that the non-stationarity of the raw series is not driven by a pure unit-root process. After first differencing, both tests consistently indicate stationarity, with the ADF test rejecting the unit-root hypothesis and the KPSS test not rejecting stationarity, confirming that the series is integrated of order $d = 1$.

Table 2. ADF and KPSS test statistics and p -values for radon time series in Gallery 1 (G1) and Gallery 2 (G2), evaluated on original (X) and first-differenced (ΔX) data across multiple lags.

Dataset	Series	Lag	ADF Statistic	ADF p -Value	KPSS Statistic	KPSS p -Value
G1	X	12	−8.601	0.000	7.565	0.010
		36	−9.079	0.000	3.002	0.010
		144	−9.079	0.000	1.184	0.010
		360	−9.079	0.000	0.681	0.015
	ΔX	12	−26.165	0.000	0.018	0.100
		36	−17.899	0.000	0.016	0.100
		144	−16.414	0.000	0.056	0.100
		360	−16.414	0.000	0.070	0.100
G2	X	12	−9.588	0.000	12.158	0.010
		36	−9.267	0.000	4.961	0.010
		144	−3.029	0.000	2.107	0.010
		360	−3.308	0.000	1.045	0.010
	ΔX	12	−22.495	0.000	0.008	0.100
		36	−18.943	0.000	0.007	0.100
		144	−13.172	0.000	0.040	0.100
		360	−11.093	0.000	0.043	0.100

Figure 3 shows the AutoCorrelation Function (ACF) and the Partial AutoCorrelation Function (PACF) of raw and differenced series. The ACF and PACF help identifying the strength and lag structure of dependencies, guiding model selection in autoregressive and moving-average frameworks [15]. The raw series exhibits slowly decaying ACF and sharply dropping PACF after lag 1, consistent with a near-unit root process. First differences show rapidly decaying ACF and PACF, indicating short-term AR behavior. In Gallery 2, the slower decay hints at a slight MA component. Longer-lag analysis (Appendix A) reveals minor daily seasonality (lags 144, 288), which does not dominate short-term behavior.

Autoregressive models of orders 1–3 were fitted to both series (Table 3). All models show extreme persistence, with $\sum_i \phi_i \approx 1$ and white-noise residuals, confirming near-random-walk dynamics. The random walk is a nonstationary process for which the optimal one-step-ahead forecast reduces to persistence [33]. The ADF test rejected the null hypothesis, indicating the absence of a pure unit root. However, this analysis points out that the series is so persistent that it causes near-random-walk behavior.

Although the radon series exhibits strong persistence and limited long-range predictability, the authors systematically evaluated the standard forecasting formulations commonly applied in time-series research. Most existing works in environmental monitoring and radon analysis rely on regression-based forecasting. However, in this case, the high persistence and near-random-walk dynamics make reliable multi-step prediction particularly challenging, as the partial autocorrelation drops below 0.5 after lag 1 (Figure 3), indicating very limited correlation beyond the first lag. While some studies mitigate this difficulty by employing a short forecast horizon, e.g., one time step [34], such short-term forecasting is insufficient for this application. A preventive mitigation strategy requires

anticipating radon exceedances several intervals in advance. Differencing is often used to make a time series stationary, but this process also reduces the signal-to-noise ratio. As a result, it becomes more difficult to detect meaningful precursors in the data. For these reasons, the task is reformulated as a time-series classification problem.

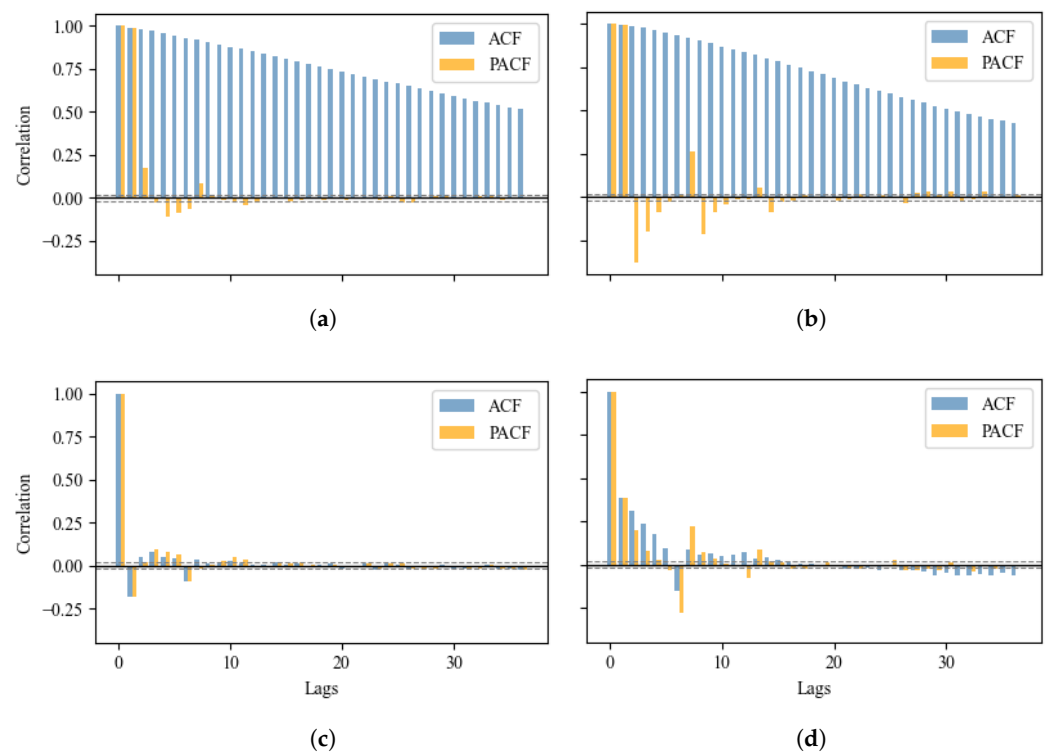


Figure 3. AutoCorrelation function (ACF) and Partial AutoCorrelation function (PACF) plots for (a) absolute values of Gallery 1; (b) absolute values of Gallery 2; (c) first differences in Gallery 1; (d) first differences in Gallery 2. Dashed lines correspond to 95% confidence intervals.

Table 3. Estimated AR(1)–AR(3) coefficients, residual variances, and coefficient sums for Gallery 1 (G1) and Gallery 2 (G2).

Dataset	AR Order	Coefficients (ϕ_i)	Residual Variance (σ^2)	$\sum_i \phi_i$
G1	1	0.9879	151.0	0.9879
	2	0.8132, 0.1770	146.3	0.9902
	3	0.8173, 0.1957, −0.0231	146.2	0.9899
G2	1	0.9953	412.5	0.9953
	2	1.3780, −0.3844	351.5	0.9936
	3	1.3005, −0.1065, −0.2017	337.2	0.9923

2.3. Rule System

2.3.1. System Description

The decision-making mechanism within RadonFAN is initially governed by a rule-based controller (RadonFAN-RS) that combines concentration thresholds and slope information to provide nuanced ventilation control. In this work, the authors propose the following preprocessing of radon measurements. The raw radon measurement at time t , denoted R_t , is first smoothed using a standard exponential weighted average [35]:

$$R_{v_t} = w_t \cdot R_t + (1 - w_t) \cdot R_{v_{t-1}} \quad (1)$$

with $w_t = 0.5$ and measurements sampled every ten minutes ($\Delta t = 10$). The authors define the short-term rate of change (slope) as the standard discrete derivative of the sequence over each interval:

$$s_t = \frac{R_{v_t} - R_{v_{t-1}}}{\Delta t} \quad (2)$$

The rule-based system encodes domain knowledge via if-then logic. The controller evaluates the smoothed radon measurement R_{v_t} and its slope s_t at each time step. Actions are determined as follows:

- If $R_{v_t} \geq T$, then activate the fan.
- If $R_{v_t} < T_w$, then deactivate the fan.
- If $T_w \leq R_{v_t} < T$, then:
 - if $s_t > s$, then activate the fan;
 - if $s_t < -s$, then deactivate the fan;
 - otherwise, maintain current fan state.

These rules were derived from domain knowledge and empirical observations of indoor radon dynamics. Specifically, T_w ensure that the fan remains off when radon concentrations are safely below the safety threshold T , and only state changes occur when concentrations approach this threshold. Slope-based conditions were incorporated to anticipate rapid changes. Parameter values are set independently for each gallery to ensure comparable behavior regardless of sensor or environmental differences. Parameters are set to $T_1 = 200$ and $T_2 = 300$ for Galleries 1 and 2, respectively. $T_w = 2/3T$, i.e., $T_{w_1} = 133$ and $T_{w_2} = 200$, and $s_1 = s_2 = 0.3$. These values are selected based on multiple empirical trials, with the aim of keeping the radon levels below the safety limits (T) while avoiding unnecessary ventilation. A notable limitation of this rule-based approach is its sensitivity to hyperparameter tuning, which requires repeated manual adjustment and does not generalize easily across conditions. Table 4 summarizes the impact of varying individual rule-system parameters relative to the base configuration ($T_w = 2/3T$, $s = 0.3$, $w = 0.5$) on fan activation steps and overall fan usage. Fan activation timing is expressed as μ_t , representing the number of intervals (steps) before or after the activation in the base configuration; positive values indicate earlier activation, while negative values indicate later activation. Fan usage is reported as the percentage change relative to the base configuration, with positive values representing increased operation and negative values representing reduced operation. Adjusting the T_w parameter produces the largest impact on both activation timing and fan usage, with lower T_w values improving early response at the expense of increased fan operation. The analysis highlights the sensitivity of the system to parameter choices, making careful tuning necessary, and shows its limited transferability: empirically set parameters must be recalibrated for each building or gallery. In contrast, DL models can be retrained on new data to adapt to different environments.

Table 4. Sensitivity analysis of the rule system. Changes relative to the base configuration ($T_w = 2/3T$, $s = 0.3$, $w = 0.5$) are reported. μ_t is the mean absolute difference in fan activation steps. Fan usage (%) is given as the percentage change relative to base fan usage.

Dataset	T_w/T	s	w	μ_t	Fan Usage (%)
G1	0.50	0.3	0.5	1.500	17.16
	0.95	0.3	0.5	−3.305	−39.46
	2/3	0.1	0.5	−0.926	−0.59
	2/3	0.5	0.5	1.000	−1.44
	2/3	0.3	0.1	−0.780	−4.52
	2/3	0.3	0.9	−1.755	3.21

Table 4. Cont.

Dataset	T_w/T	s	w	μ_t	Fan Usage (%)
G2	0.50	0.3	0.5	1.449	18.38
	0.95	0.3	0.5	−3.113	−21.16
	2/3	0.1	0.5	−0.343	−0.43
	2/3	0.5	0.5	1.215	1.24
	2/3	0.3	0.1	−0.013	3.79
	2/3	0.3	0.9	−0.804	0.52

2.3.2. Preliminary Analysis of System Behavior

To guide the development of the AI-based control strategy, the performance of the existing rule-based ventilation logic was first evaluated. Two types of inefficiencies were identified: (i) insufficient activations, where the fan failed to prevent a radon exceedance, and (ii) unnecessary activations, where the fan operated even though concentrations remained below the threshold.

The sufficiency analysis was conducted during the period when the fan operated according to the rule-based system. An activation was considered insufficient when the radon concentration R_t exceeded the threshold T despite fan operation. Such cases represented 16.6% of activations in Gallery 1 and 9.1 % in Gallery 2. On average, the exceedance occurred $\mu_t = 1.34$ ($\sigma = 0.86$, 95 % CI ± 0.32) time steps after fan activation for Gallery 1 and $\mu_t = 1.33$ ($\sigma = 0.89$, 95 % CI ± 0.56) time steps for Gallery 2. Once the threshold was surpassed, concentrations remained above T for $\mu_t = 3.90$ ($\sigma = 3.03$, 95% CI ± 1.16) time steps in Gallery 1 and $\mu_t = 2.75$ ($\sigma = 1.29$, 95% CI ± 0.62) time steps in Gallery 2. These findings indicate that the rule-based system often reacted only after concentrations had accelerated towards exceedance. To assess unnecessary activations, the same rule-based logic was retrospectively applied to data from the period without fan operation. In this case, the rule would have triggered the fan unnecessarily, i.e., concentrations remained below T throughout, in 58.1% of the cases for Gallery 1 and 17.0% for Gallery 2.

The findings directly informed the design of the AI-based predictive architecture described in the following section, in order to avoid insufficient and unnecessary activations of a reactive system. Moreover, this preliminary analysis indicates that the rule-based system performs better in Gallery 2, exhibiting a lower proportion of both insufficient and unnecessary fan activations.

2.4. Deep Learning System

2.4.1. Model Architecture

Two time series classification models are defined, with their architectures shown in Figure 4. The first, Regression-to-Classification (R2C), performs numerical radon forecasting as a pretraining stage and subsequently maps the predicted values to discrete ventilation states within a unified framework. The second, Direct Classification (DC), bypasses forecasting and directly predicts fan activation states from the radon time series. These approaches highlight two complementary strategies for integrating predictive modeling with control decisions in indoor radon mitigation. The authors hypothesize that the DC model will outperform the R2C model because it directly learns the decision boundary for fan activation, avoiding the error accumulation inherent in regression-based forecasts of near-random-walk radon dynamics (Section 2.2.2).

The two models are built using One-Dimensional Convolutional Neural Networks (1D-CNNs). Radon time series exhibit near-random-walk behavior, with slow drifts and stochastic fluctuations that limit the effectiveness of simpler statistical classifiers based on stationarity or predefined features. CNNs address this by learning local, discriminative temporal patterns directly from raw signals. Compared to recurrent or attention-based

architectures, CNNs have lower computational cost and do not rely on long-memory mechanisms, which are unnecessary in this setting (Section 2.2.2). This makes them well suited for capturing short-term dynamics relevant for timely mitigation decisions.

The input is defined as $X \in \mathbb{R}^{B \times L \times F}$, where B denotes the batch size, L the lookback window length, and F the number of features per time step. The network includes two convolutional blocks, each composed of a 1D causal convolution with kernel size $k = 3$ and hidden dimension $H = 16$. Dilations are set to $d = 1$ and $d = 2$ for the first and second layers, respectively, increasing the receptive field in the second layer. The resulting feature map is flattened and passed through a MultiLayer Perceptron (MLP) with fully connected layers of sizes 96 and 16. All hidden convolutional and linear layers are followed by Parametric Rectified Linear Unit (PReLU) activations. Dropout regularization is employed inside the MLP ($p = 0.3$). The final layer of the DC model produces a single classification logit (z). For clarity and simplicity in Figure 4, the output is depicted as $\hat{y}_{cl}$, representing the predicted probability obtained via the sigmoid function during loss computation ($\hat{y}_{cl} = \sigma(z)$, Section 2.4.3). The head of the R2C model comprises a linear layer that outputs $\hat{y}_{reg}$ of size equal to the predefined forecast length. Then, it employs an extra linear layer to project from forecasted regression values to classification predicted probability $\hat{y}_{cl}$. In Figure 4 the shared backbone of the R2C and DC architectures is shown in purple, while R2C head and DC heads are shown in blue and red, respectively. These heads are mutually exclusive, with only one attached to the backbone at a time.

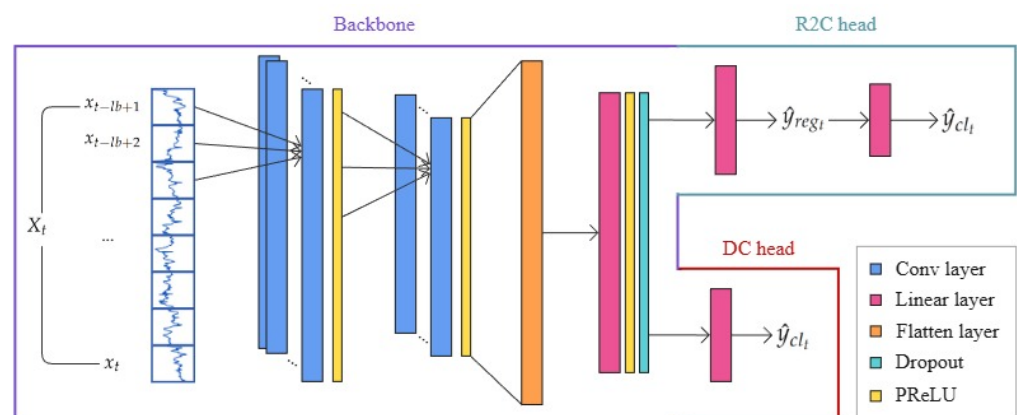


Figure 4. R2C and DC architectures. The shared backbone (purple) connects to one of two alternative heads: R2C (blue) or DC (red). X_t refers to input data from $t - lb + 1$ to t , with lb representing the lookback parameter. The variable $\hat{y}_{cl_t}$ corresponds to the classification output at time t , while $\hat{y}_{reg_t}$ denotes the regression output at time t .

A structured hyperparameter search was conducted to optimize architectural choices. The explored configurations included normalization strategies ({none, batch normalization, layer normalization}), activation functions ({ReLU, PReLU}), dropout probabilities ({0, 0.2, 0.3, 0.5}), kernel values ([2, 4]), strides ([1, 2]), dilations ([1, 4]), pooling strategies ({None, Avg, Max}), hidden dimensions ({16, 32, 64, 128}), number of convolutional layers ([1, 3]), and number of linear layers ([1, 3]). The final configuration was selected based on validation performance and stability. Four modified architectures are further evaluated, as shown in the results section. DCMF employed MaxPooling (kernel $K = 2$) after each convolutional block instead of dilations, reducing the temporal dimension. To allow for a smoother gradient flow, the DCRes modification used residual connections after each convolutional block. DCBN used batch normalization to avoid exploding gradients. Finally, DCnDp eliminated the dropout of the MLP, leading to greater capacity.

2.4.2. Data Preprocessing

To evaluate the model while preserving temporal consistency, the dataset was split into training, validation, and test sets using a 70:15:15 ratio. The split was performed in the order of test-validation-train to maintain approximately similar positive class ratios across sets, although perfect balance is not guaranteed. Stratified sampling was not applied, as it would violate the temporal structure of the time series data. For Gallery 1, the proportion of positive samples ($y_{cl} = 1$) is 9.11% in the training set, 23.15% in the validation set, and 9.91% in the test set, yielding an overall positive rate of 11.32%. Gallery 2 exhibits positive rates of 20.69%, 48.17%, and 37.21% in the training, validation, and test sets, respectively, with an overall rate of 27.24%. Despite these imbalances, the temporal split ensures realistic evaluation conditions. Weighted loss functions and data augmentation were applied during training to mitigate class imbalance (Section 2.4.3). Figure 5 shows the distribution of radon values for each dataset split and gallery, highlighting the higher positive rate in the validation sets and the heavier right tail observed in Gallery 2. The broader distribution of radon values in Gallery 2 indicates greater variability and a less stable behavior, with radon concentrations more widely spread across values. A Robust Scaler was fitted exclusively on the training set and then applied to the validation and test sets to prevent information leakage. This scaler was chosen over MinMax or Standard scaling methods due to the strong right-skewed distribution of the data. The scaler is saved to be applied to online radon monitoring.

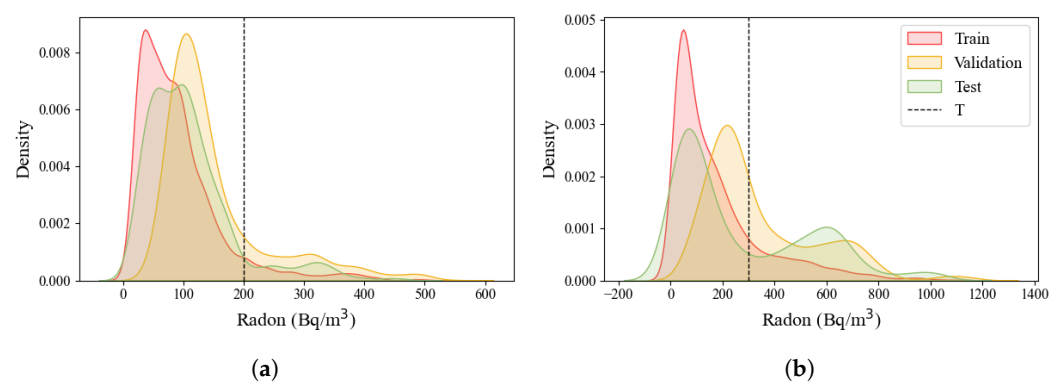


Figure 5. Distribution of train, validation, and test sets for (a) Gallery 1; (b) Gallery 2. T refers to threshold.

Model input and outputs were constructed independently for each set to ensure no overlap and information leakage. At each forecasting step t , the model input was constructed as $X_t = \{x_{t-lb+1}, \dots, x_t\}$, where $lb = 12$ steps (2 h). This value was selected to provide the model with a sufficient time-series window to capture past dynamics and forecast future radon values, while larger lb values provided limited additional information due to the low autocorrelation at longer lags (Figure 3). On top of that, the precise lookback parameter depends on the specific ventilation conditions, including natural ventilation rates, room volume, and radon exhalation coefficient. These factors determine how far back in time past measurements remain informative, delimiting the amount of historical information necessary to extract relevant dynamics for accurate modeling and control. Several configurations were studied ($lb \in [6, 360]$). The input was exclusively built on raw radon measurements, i.e., $X_t = \{R_{t-lb+1}, \dots, R_t\}$, to reduce model size and allow for easier deployment. No external variables, such as temperature, humidity or pressure, were employed in this initial study, but they are considered for future extensions. The regression output, exclusively defined for the R2C model, was $y_{reg_t} = \{R_{t+1}, \dots, R_{t+fc}\}$ with $fc = 6$ as the forecast horizon. The classification output was defined as $y_{cl} = 1$ given that any $R_{t+i} \geq T$ for $i \in [1, fc]$. Extending the forecast horizon increases task complexity, as noted by [34]. Shorter horizons simplify the prediction task, but provide less lead time

to respond to rising radon concentrations. Moreover, preliminary rule-based analysis of unsafe events durations suggested that $fc > 2$ was necessary (Section 2.3.2). The choice of $fc = 6$ represents a compromise between predictive reliability and actionable lead time. Several forecast horizons ($fc \in [1, 12]$) were evaluated.

Apart from architecture modifications, the result section reports two feature-related modifications and four time series parameter modifications. The feature-related modifications are DC16F and DC4F. DC16F is trained on a total of sixteen variables derived from the raw radon measurements. DC4F is trained only on the subset of four features that seemed most significant in terms of correlation to the binary label and feature importance analysis (Appendix B). In order to see the effect of the fc and lb parameters, the modifications of DCfc3, DCfc12, DClb6 and DClb24 are reported using the corresponding fc or lb parameter indicated in the name.

Data augmentation was applied solely to the training set to improve generalization, with synthetic samples generated on-the-fly at each epoch. Table 5 indicates the augmentation strategies employed, their probabilities (p), magnitude factors (f) and formulas. Smaller augmentation probabilities and magnitude factors are used for the negative class due to its larger sample size, lower variability, and the observation that data augmentation affects classes differently [36]. Data augmentation parameters were tuned empirically to match the training data distribution, increasing variability within a feasible margin.

Table 5. Data augmentation techniques applied to the training set with their corresponding description, probability (p), magnitude factors (f) and formulas. Probabilities and factor magnitudes differ for positive (+) and negative (−) classes.

Technique	Description	p (+/−)	f (+/−)	Formula
Jittering	Add Gaussian noise $\mathcal{N}(0, \sigma^2)$	0.75/0.25	2.5/1	$x \leftarrow x + \mathcal{N}(0, \sigma^2), \quad \sigma = 0.05 \cdot f \cdot \text{std}(X_{\text{train}})$
Scaling	Multiply signal by random factor	0.75/0.25	2.5/1	$x \leftarrow x \cdot s, \quad s \sim \mathcal{U}[1 - 0.1f, 1 + 0.1f]$
Time warping	Remove points (window of 3 before/after random point) and interpolate using smooth nonlinear function	0.75/0.25	2.5/1	$x[i - 3 : i + 3] \leftarrow \text{interp}(x[i - 3 : i + 3]), \sigma = f$
Magnitude warping	Locally scale segments of 4 points	0.75/0.25	2.5/1	$x[i : i + 4] \leftarrow x[i : i + 4] \cdot s, \quad s \sim \mathcal{U}[1 - 0.1f, 1 + 0.1f]$
Masking	Remove random points from last quarter of the sequence to prevent reliance on recent points; add variability	0.25/0.10	−	$x[-n:] \leftarrow 0, \quad n \sim \mathcal{U}(1, \text{len}(x)/4)$

2.4.3. Training and Evaluation

The different approaches were trained and tested on a workstation that has a NVIDIA GeForce RTX 4060 Ti GPU, a 13th Gen Intel Core i7-13700K CPU and 32 GB of RAM. All training and experiments were performed using Python 3.12.3 with PyTorch 2.5.1. The initial learning rate was set to 1×10^{-4} and the number of epochs to 500, with patience of 50 epochs. AdamW [37] was selected as the optimizer and a scheduler that halved the learning rate every 10 epochs without improvement was employed. A weight decay of 1×10^{-2} was chosen. The convolutional and linear layers were initialized using Kaiming initializers [38]. For the training and validation sets, a batch size of 256 was selected and shuffling was enabled to mitigate imbalanced batches due to seasonality. Although WeightedRandomSampler is a common method to handle class imbalance, it was intentionally avoided due to its potential to introduce instability during training. Instead, to address the challenge of rare events in the data distribution, the authors adopt a weighted loss formulation (see Equations (3)–(6)). A hyperparameter search was conducted over the training settings. Several initial learning rates ($[1 \times 10^{-4}, 1 \times 10^{-3}]$) and weight decay values ($[1 \times 10^{-5}, 1 \times 10^{-1}]$) were evaluated. Different stopping patience epochs ($\{10, 30, 50, 100\}$), learning rate scheduler patience epochs ($\{3, 5, 10\}$), and batch sizes ($\{64, 128, 256, 512\}$) were also explored. The final configuration provided stable convergence and robust generalization for both regression and classification tasks across the two galleries.

For the R2C model, a joint multitask loss can promote shared temporal representations and improve generalization [39]. Weighted combinations have been proposed [40], yet they are highly sensitive to hyperparameter tuning. Initial experiments with multi-term losses and curriculum learning confirmed that classification weighting often overwhelmed regression, leading to early stopping and reduced model quality. To mitigate this, a staged training strategy is adopted: first training the encoder–decoder with the regression head while freezing the classifier (last layer), then freezing the regression to train the classifier. This decoupled optimization prevents gradient interference, preserves forecast accuracy, and enables the classifier to specialize on calibrated regression outputs, yielding more stable results. Three forecasting approaches can be considered: (i) direct multi-step prediction, where the model predicts R_{t+fc} directly from past lags as $\hat{R}_{t+fc} = \phi(R_t, \dots, R_{t-lb+1})$, which often yields persistent forecasts when additional information is weak; (ii) iterative 1-step prediction, where the model is trained for $\hat{R}_{t+1} = \phi(R_t, \dots, R_{t-lb+1})$ and recursively applied as $\hat{R}_{t+fc} = \phi^{fc} R$, optimal for AR(1) series but prone to error accumulation; and (iii) iterative multi-step prediction $\hat{R}_{t+fc} = \phi(\hat{R}_{t+fc-1}, \dots, R_{t-lb+1})$, which balances learning short-term dynamics while reducing error accumulation at longer horizons. The third forecasting strategy is selected in order to balance persistence and error accumulation, and ensure an end-to-end pipeline for deployment. For the forecasting task, the authors employ the conventional weighted Mean Squared Error (MSE) loss [41]:

$$\mathcal{L}_{\text{reg}} = \frac{1}{N} \sum_{i=1}^N w_i \frac{1}{fc} \sum_{t=1}^{fc+1} (\hat{y}_{\text{reg},i,t} - y_{\text{reg},i,t})^2, \quad (3)$$

where N is the number of total time steps in the dataset and w_i is the sample weight, emphasizing the onset of anomalous sequences and encouraging early detection. The authors define w_i as:

$$w_i = (fc - d)^2 \quad (4)$$

with $d \in [0, fc]$ being the minimum number of data points given $y_{cl_t} = 1$ and $y_{cl_{t-d}} = 0$. The training criterion for the classification task was the Binary Cross Entropy With Logits Loss [41], where $\hat{y}_{cl_i} = \sigma(z_i)$ and

$$\mathcal{L}_{\text{cl}} = \frac{1}{N} \sum_{i=1}^N w_i w_p y_{cl_i} \log(\hat{y}_{cl_i}) + w_i (1 - y_{cl_i}) \log(1 - \hat{y}_{cl_i}) \quad (5)$$

where z_i is the model logit, σ is the sigmoid function. The authors define the positive-class weight w_p as:

$$w_p = \sqrt{\frac{\sum_i (1 - y_{cl_i})}{\sum_i y_{cl_i}}} \quad (6)$$

to assign higher importance to errors on positive instances compared to negative ones. Particularly, $w_{p_1} = 3.16$ for Gallery 1 and $w_{p_2} = 1.96$ for Gallery 2. Weights (w_i and w_p) were only applied during training to ensure unbiased validation and correct early stopping.

For evaluation, the DC and R2C models were assessed on classification performance. Metrics such as accuracy and the Area Under the Receiver Operating Characteristic Curve (ROC AUC) could be misleading in this imbalanced setting. Hence, they were not reported. To capture performance more reliably, the reported metrics include precision, recall, F1-score and Brier score. Because the dataset is imbalanced and false negatives increase the risk of unsafe radon exposure, the recall metric was prioritized to ensure hazardous events are detected, while other metrics such as precision and F1-score were also considered. Ninety-five percent confidence intervals (95% CI) for all evaluation metrics were estimated

using non-parametric bootstrap resampling (1000 iterations), where each bootstrap sample was generated by sampling the test set with replacement and recomputing the metric.

Because time series forecasts often exhibit slight phase shifts, where predicted events occur a few steps earlier or later than observed, some studies adopt soft evaluation methods [42], using triangular membership functions to assign partial credit to temporally shifted detections. While this approach provides a fairer assessment of temporal alignment, timely fan activation and deactivation are critical for safety and energy consumption. Therefore, the primary evaluation relies on hard metrics, with soft scores reported only in the Appendix C for completeness.

Large models can lead to slow inference and high power consumption, which undermines real-time performance and portability. As the models are to be deployed in a Raspberry Pi, which has limited memory and processing power, the number of trainable parameters, the model size and the Floating Point Operations (FLOPs) are also examined.

As a baseline model for DL evaluation, InceptionTime [43] is used as a state-of-the-art model. The model was run via sktime for 500 epochs using the default configuration. Attempts to modify parameters such as kernel size and model depth to match those used in the R2C and DC models led to compatibility issues, preventing a direct one-to-one alignment between the InceptionTime configuration and the proposed models.

2.5. Performance Metrics for Control Systems

Evaluating the performance of the control systems (rule-based, R2C and DC) requires a dual focus: ensuring compliance with safety standards and reducing operational costs. To this end, two performance indicators are defined: an unsafety counter (U) and an energy consumption ratio (C). The unsafety counter defined for this work is:

$$U = \sum_{t=1}^N \delta_t, \quad (7)$$

where

$$\delta_t = \begin{cases} 1, & \text{if } F_t = 0 \text{ and } R_{t+i} \geq T \forall i \in [0, tm], \\ & \text{and no } t' \in [t - dr + 1, t - 1] \text{ satisfies the same condition,} \\ 0, & \text{otherwise.} \end{cases} \quad (8)$$

with F_t being the fan state (0:OFF, 1:ON) at time t , tm the time margin for the unsafe event and dr the unsafe event duration. The condition enforces uniqueness: once an episode is triggered, subsequent time points inside the same duration interval are ignored. Based on a preliminary analysis of insufficient activations (Section 2.3.2), the parameters dr and tm were set to the smallest integers above the observed means of unsafe events radon exceedance lag after fan activation and duration under the preliminary rule system. The mean lags to the first radon exceedance were 1.34 (Gallery 1) and 1.33 (Gallery 2), leading to $tm_1 = tm_2 = 2$ for both galleries. The mean durations were 3.9 (Gallery 1) and 2.75 (Gallery 2), yielding $dr_1 = 4$ and $dr_2 = 3$, respectively.

The authors define the energy consumption ratio as the proportion of time the fan is active:

$$C = \frac{1}{N} \sum_{t=1}^N F_t \quad (9)$$

Apart from the rule system and the two DL models, the unsafety counter and energy consumption ratio for (i) the best theoretical model and (ii) the persistence model are also

reported. The authors define the best theoretical model as the ideal case with no unsafe events ($U = 0$) and an energy consumption ratio that accounts for all data points above the threshold plus tm time steps before each threshold exceedance:

$$C = \frac{1}{N} \sum_{t=1}^N \delta_t^{(1)} + tm \cdot \sum_{t=2}^N \delta_t^{(2)}, \quad (10)$$

The persistence model energy consumption ratio is even lower as it only activates the fan when radon values are over the threshold:

$$C = \frac{1}{N} \sum_{t=1}^N \delta_t^{(1)}, \quad (11)$$

which leads to a high number of unsafety events:

$$U = \sum_{t=2}^N \delta_t^{(2)}, \quad (12)$$

with the deltas defined as:

$$\delta_t^{(1)} = \begin{cases} 1, & R_t \geq T, \\ 0, & \text{otherwise,} \end{cases} \quad \delta_t^{(2)} = \begin{cases} 1, & R_{t-1} < T \text{ and } R_t \geq T, \\ 0, & \text{otherwise.} \end{cases} \quad (13)$$

Equations (7)–(13) are defined for this case study and not derived from previous works or the literature.

For the purpose of model evaluation in the DL study, predicted probabilities from the AI models were binarized using a simple threshold of 0.5, i.e., a prediction is considered positive ($y = 1$) if the model output exceeds 0.5 and negative ($y = 0$) otherwise. However, for deployment in the radon mitigation system, a smoother fan control is required to prevent rapid on/off switching caused by small fluctuations around this threshold. To achieve this, a hysteresis-based control that ensures operational stability while maintaining a safety margin is implemented. Specifically, the fan is activated whenever the predicted value exceeds 0.5 and deactivated whenever it falls below 0.35. Within the dead zone [0.35, 0.5], the controller preserves the previous state. This classic hysteresis approach was chosen because it did not rely on signal smoothing or consecutive-step analysis, which could introduce delays critical in the radon mitigation scenario. The dead zone was set below 0.5 to prioritize safety over the energy consumption ratio. The hyperparameter 0.35 was chosen based on grid search and empirical trials.

Based on the preliminary analysis of the rule-based system (Section 2.3.2), substantial reductions in both unsafe events and fan runtime are anticipated when moving to predictive control strategies, mainly in Gallery 1, which exhibits a higher proportion of insufficient activations (reflected in the unsafety counter) and unnecessary activations (reflected in the energy consumption ratio). As both R2C and DC adopt predictive mechanisms, they are expected to markedly reduce the unsafety counter relative to the reactive rule-based system, by anticipating threshold exceedances rather than responding after they occur. However, due to the near-random-walk nature of radon dynamics, regression-based forecasting in R2C is affected by noise accumulation, which can propagate into the classification decision and disproportionately increase false positives. As a result, while R2C is expected to substantially reduce unsafe events, its impact on energy consumption is expected to be more limited. In contrast, by directly optimizing the classification objective, DC is expected to further reduce both unsafe events and unnecessary fan activations, leading to the most favorable trade-off between safety and energy efficiency.

3. Results and Discussion

3.1. AI Model Performance and Efficiency

Table 6 reports the hard classification performance with 95% confidence intervals and efficiency parameters of the InceptionTime [43] model as baseline, the R2C model and the DC model for the two galleries (G1 and G2). Hard classification metrics include precision (Pr), recall (Rec), F1-score (F1) and Brier score (Brier). The reported efficiency metrics are the number of trainable parameters (Param.), the model size (Size) and the number of Floating Point Operations (FLOPs). The best results are indicated in bold and the second best are underlined. Soft classification metric results can be found in Appendix C.

Table 6. Hard classification performance metrics with 95% confidence intervals and efficiency parameters for InceptionTime (IT) [43], R2C and DC. Best results are indicated in bold numbers, second-best are underlined. Up (↑) and down (↓) arrows denote whether higher or lower values are preferable for each variable.

	G1				G2				Param. (M) ↓	Size (MB) ↓	FLOPs (M) ↓
	Pr ↑	Rec ↑	F1 ↑	Brier ↓	Pr ↑	Rec ↑	F1 ↑	Brier ↓			
IT [43]	0.924 [0.883–0.963]	0.825 [0.765–0.880]	0.872 [0.801–0.884]	0.035 [0.026–0.044]	0.975 [0.961–0.987]	0.968 [0.932–0.954]	0.972 [0.961–0.981]	0.021 [0.015–0.030]	0.420	1.604	10.094
R2C	0.470 [0.415–0.526]	1.000 [1.000–1.000]	0.640 [0.588–0.687]	0.111 [0.105–0.117]	0.630 [0.598–0.657]	0.993 [0.984–0.998]	0.770 [0.744–0.793]	0.149 [0.141–0.157]	3.963	<u>0.015</u>	40.800
DC	<u>0.645</u> [0.583–0.709]	<u>0.979</u> [0.951–1.000]	<u>0.778</u> [0.778–0.778]	<u>0.039</u> [0.039–0.039]	<u>0.929</u> [0.907–0.950]	<u>0.980</u> [0.967–0.991]	<u>0.954</u> [0.954–0.954]	<u>0.024</u> [0.024–0.024]	<u>3.956</u>	0.015	6.800

InceptionTime achieves the highest precision, with F1- and Brier scores across both galleries, indicating strong discrimination and well-calibrated probabilities. However, its comparatively lower recall reveals a key limitation in safety-critical contexts, as some hazard radon events remain undetected. Although not all false negatives immediately lead to unsafe exposure, higher recall generally reduces the associated risk. In contrast, R2C achieves the highest recall largely due to its regression-then-classification pipeline. Multi-step regression predictions are noisy (MSE 0.112–0.203, R^2 0.895–0.905), causing the linear classifier to interpret fluctuations as positive instances. This behavior inflates recall but leads to lower precision, degraded probability calibration with higher Brier scores, and unnecessary fan activation, reducing practical usability.

DC bypasses regression and directly learns the decision boundary, avoiding error accumulation from intermediate regression. Its recall exceeds 0.975 in both galleries, improving over InceptionTime, while maintaining remaining metrics close to those obtained by the baseline model. Compared to R2C, DC substantially improves precision, F1 and Brier scores mitigating over-prediction and providing more reliable probability estimates. DC is also highly efficient, with the lowest FLOPs and model size, enabling faster inference and smaller memory footprint.

The performance differences between galleries can be attributed to underlying data characteristics. Gallery 1 exhibits fewer positive events, increasing metric sensitivity and exposing a limitation of all models under class imbalance. Overall, DC combines high recall, near-best classification metrics, and low computational cost, making it the most practical choice for radon time-series classification in safety-critical applications.

3.2. Model Component Analysis

Table 7 shows the differences in hard classification and efficiency metrics for several DC variants. Exact values with the 95% CI can be found on Appendix C. Best values are bold and second-best underlined. As it is already a quite small model an ablation study becomes difficult. Hence the effects of model modifications rather than component ablation

are analyzed. Specifically, the authors evaluate whether model performance is limited by architectural choices, by feature richness, or by the underlying time-series structure, focusing primarily in recall. Four architectural modifications (DCMP, DCRes, DCBN, DCnDp), two data-based modifications (DC16F, DC4F), and four time series parameters modifications (DCfc3, DCfc12, DClb6, DClb24) are reported.

Table 7. Differences relative to DC model for hard classification and efficiency metrics of several DC variants. Best values are bold, second-best are underlined. Up (↑) and down (↓) arrows denote whether higher or lower values are preferable for each variable.

	G1				G2				Param. (M) ↓	Size (MB) ↓	FLOPs (M) ↓
	Pr ↑	Rec ↑	F1 ↑	Brier ↓	Pr ↑	Rec ↑	F1 ↑	Brier ↓			
DC	0.645	0.979	0.778	0.039	0.929	0.980	0.954	0.024	3.956	0.015	6.800
DCMP	+0.017	+0.007	+0.014	+0.000	−0.005	−0.002	−0.004	+0.002	−2.048	−0.008	<u>−2.655</u>
DCRes	−0.011	+0.014	−0.004	+0.003	−0.006	+0.001	−0.003	−0.001	+0.032	+0.000	+0.098
DCBN	<u>+0.067</u>	−0.028	<u>+0.036</u>	−0.009	+0.038	−0.012	<u>+0.014</u>	<u>−0.005</u>	+0.064	+0.000	+0.786
DCnDp	+0.008	−0.007	+0.003	−0.002	+0.010	−0.008	+0.001	−0.001	+0.000	+0.000	−0.001
DCF16	+0.023	+0.021	+0.023	−0.004	+0.004	<u>+0.001</u>	+0.003	−0.001	+0.720	+0.003	+4.423
DCF4	−0.015	<u>+0.021</u>	−0.005	+0.006	−0.006	−0.006	−0.006	+0.004	+0.144	+0.001	+0.884
DCfc12	+0.046	−0.020	−0.047	+0.027	+0.026	−0.032	−0.003	+0.008	+0.000	+0.000	−0.001
DCfc3	+0.107	−0.069	+0.046	−0.006	<u>+0.036</u>	+0.001	+0.019	−0.009	+0.000	+0.000	−0.001
DClb24	−0.047	+0.021	−0.029	+0.011	−0.010	−0.012	−0.011	+0.005	+3.072	+0.012	+6.782
DClb6	+0.025	−0.014	+0.013	−0.005	+0.006	−0.002	+0.002	−0.002	<u>−1.536</u>	<u>−0.006</u>	−3.392

Architectural modifications yield limited and inconsistent gains. Among them, max-pooling (DCMP) benefits from stable temporal patterns, improving Gallery 1 metrics but reducing performance in time series with higher variability, such as Gallery 2. Its lower computational cost arises from halving the temporal dimension. Residual connections (DCRes) consistently improve recall, particularly in Gallery 1, making it the most balanced choice for safety-critical settings, while batch normalization or dropout removal improve precision, F1, and Brier scores at the expense of recall.

Feature richness exerts the strongest influence on performance. Expanding to sixteen features (DCF16) improves all classification metrics, highlighting the importance of additional signals for early transitions, although these gains come with increased computational cost. Reducing the feature set to four (DCF4) decreases performance, showing that seemingly low-importance variables contribute meaningfully to robust decision-making. This exposes a key limitation of the base DC model: performance is constrained by the available feature set.

Time-series parameters also play a key role. Shorter forecast horizons (DCfc3) and lookback windows (DClb6) improve precision, F1-, and Brier scores, but can slightly reduce recall, on top of leaving less margin for early warning in the case of a shorter forecast horizon (DCfc3). Longer horizons (DCfc12) or extended lookbacks (DClb24) tend to degrade recall or overall performance. Efficiency metrics are highly sensitive to the lookback parameter, as longer lookback windows consistently incur higher cost. These effects reflect the near-random-walk dynamics of radon time series, which limit the utility of long-range temporal context.

Overall, the analyses indicate that performance is primarily constrained by feature informativeness and intrinsic time-series characteristics rather than architectural design alone. With the exception of feature expansion, no single modification consistently improves all metrics across both galleries. Architectural refinements cannot overcome these data-driven constraints. The base DC model therefore represents the most balanced configuration, combining strong overall performance with low computational cost. Incorporating additional external variables with weaker near-unit-root behavior emerges as the most promising direction for improving safety-critical performance in future work.

3.3. System Performance

Figure 6 shows the results of the systems performance in terms of unsafety counter and energy consumption ratio. Results are given for the rule-based model (RS), the R2C model, the DC model, the persistence model and the best theoretical model.

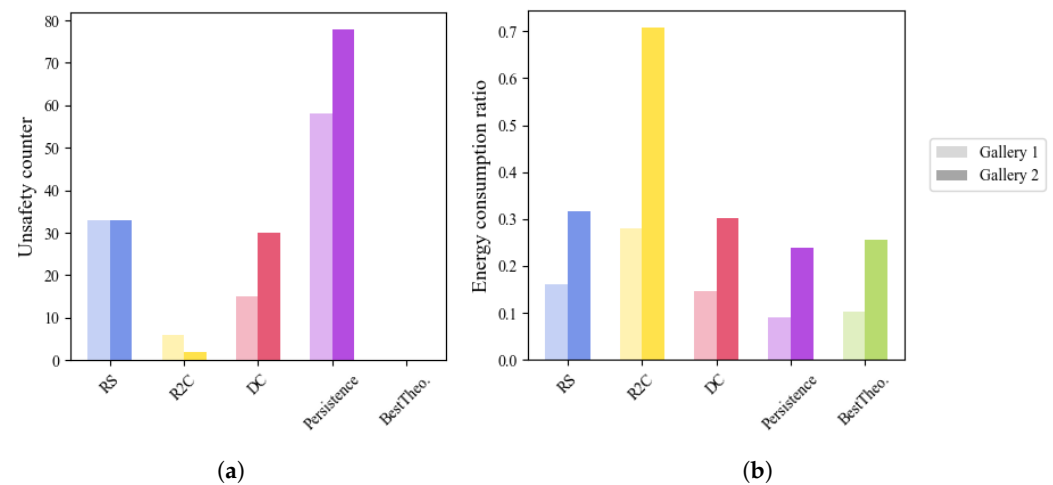


Figure 6. Results of rule-based system (RS), R2C, DC, persistence model, and best theoretical model (BestTheo) for Gallery 1 (light) and Gallery 2 (dark) regarding (a) unsafety counter; (b) energy consumption ratio.

The rule-based system obtains an unsafety event count of 33 and 32 for Galleries 1 and 2, respectively. This comes with an energy consumption ratio of 0.1598 and 0.3163. R2C substantially reduces unsafe events (6 and 2), approaching the theoretical safety bound, but does so at the cost of excessive fan activation and high energy consumption (0.2812 and 0.7091). This behavior confirms that regression-driven control leads to systematic over-prediction and poor robustness. The DC model provides the most balanced compromise. It consistently reduces, relative to the rule-based system, both the unsafety counter (15 and 30) and the energy consumption ratio (0.1468 and 0.3012) in both galleries. Specifically, it achieves reductions of 54.55% and 6.25% in the unsafety counter and 8.14% and 4.77% in the energy consumption ratio for Galleries 1 and 2, respectively.

In this study, with six low-power fans (130 W per gallery), operating all the time (without a control system), the ventilation system would consume approximately

$$E_{\text{always}} = 0.13 \text{ kW} \times 24 \text{ h/day} \times 365 \text{ days/year} \approx 1139 \text{ kWh/year-gallery},$$

corresponding to a total electricity cost of approximately EUR 228/year per gallery (given 0.2 EUR/kWh) and CO₂ emissions of about 322 kg/year per gallery (assuming 0.283 kg CO₂/kWh). The RS reduces fan usage to 15.98–31.63% of total operation time, corresponding to 182–360 kWh/year, a cost of EUR 37–EUR 72/year, and CO₂ emissions of 51–102 kg/year for Galleries 1 and 2, respectively. The DC system further reduces usage to 14.68–30.12%, corresponding to 167–343 kWh/year, a cost of EUR 33–EUR 68/year, and CO₂ emissions of 47–97 kg/year for Galleries 1 and 2, respectively. Across both galleries, these reductions imply annual savings of approximately EUR 355 and 500 kg CO₂ compared to always-on operations, and smaller but meaningful savings of EUR 8 and 9 kg CO₂ compared to the RS. Although the economic and CO₂ reductions relative to the RS are modest for these small galleries, in larger buildings with more fans or higher operational demands, similar strategies could translate into substantial energy and cost savings, highlighting the broader applicability of this approach.

Figure 7 illustrates the time series for Galleries 1 and 2 between the 16th and 17th of March 2024, as an example of how the RS, R2C and DC models work. The fan activation is marked in blue for the RS, in yellow for the R2C and in red for the DC model. In this example, the RS activates earlier during smooth monotonic increases, reflecting its reliance on instantaneous threshold crossings and slope conditions. This behavior is advantageous after long safe intervals, as it enables early reaction to gradual increases. However, under near-threshold conditions, this rigidity becomes detrimental. When radon concentrations remain close to the threshold with negative slopes, the rule-based logic may deactivate ventilation despite sustained exposure risk. Moreover, RS produces multiple short activation periods when radon is still far from the safety limit, resulting in unnecessary fan usage. By contrast, the DC model incorporates temporal context and maintains activation during persistent near-threshold conditions, thereby reducing the likelihood of unsafe events. This behavior produces longer but smoother activation periods, slightly increasing operational cost near the threshold compared to RS. However, it reduces overall cost by avoiding frequent and unnecessary activations when radon levels are far from the threshold. The R2C model, in both galleries, exhibits a strong bias towards fan operability, activating ventilation even when radon levels are clearly below the threshold and failing to provide targeted control.

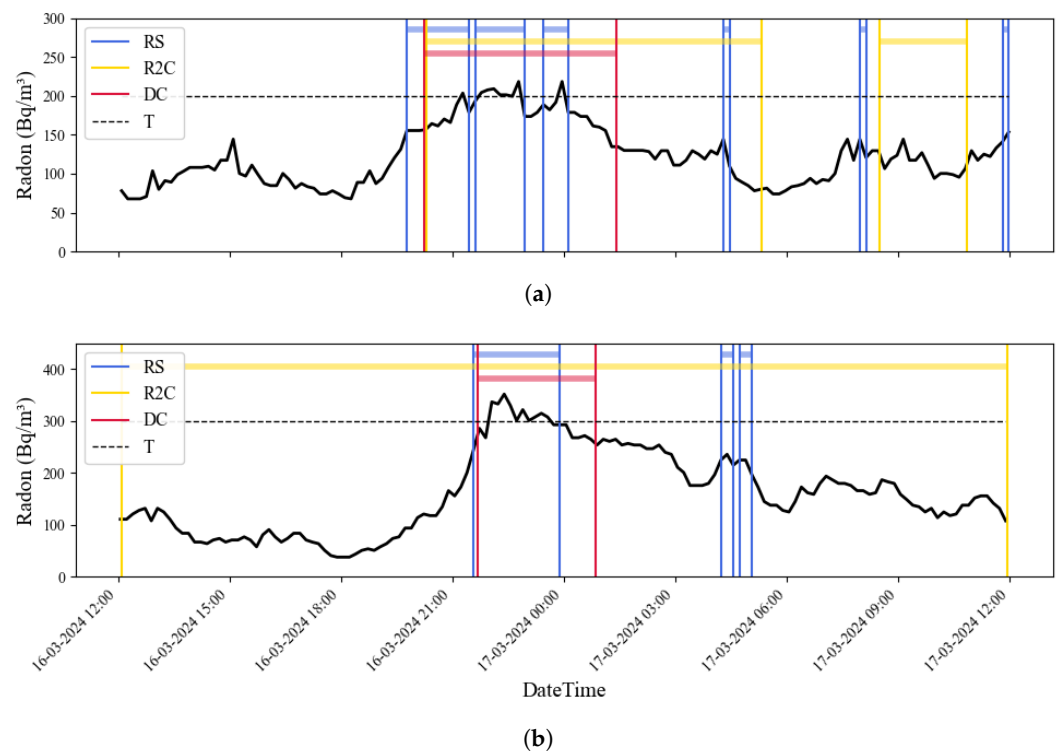


Figure 7. Example of radon time series between the 16th and 17th of March, 2024. Horizontal bands indicate fan activation under the rule system (blue), the R2C model (yellow) and the DC model (red). T indicates the threshold value. Values are shown for (a) Gallery 1; (b) Gallery 2.

Even though the general results (Figure 6) and the time series example (Figure 7) show DC improvements over RS, the magnitude differs across galleries. Improvements are most pronounced in Gallery 1, which exhibits smoother radon dynamics, whereas gains in Gallery 2 are more limited due to higher variability and rapid radon increases. A key limitation of DC is that it may underperform in environments characterized by abrupt radon spikes and weak temporal precursors where near-unit-root dynamics limit the predictive capacity of models trained solely on radon measurements. In addition, all AI-based controllers (DC and R2C) were evaluated offline using historical data; no online closed-

loop deployment has yet been conducted. Although DC exhibits a small computational footprint suitable for real-time operation, live evaluation may reveal additional challenges. Overall, despite these limitations, DC offers improved robustness, scalability, and reduced dependence on manual tuning compared to rule-based control, making it a strong candidate for real-world deployment. Future work will focus on online validation, hybrid control strategies, and the integration of external variables to mitigate the identified failure modes, particularly in highly volatile environments.

3.4. Comparative Analysis with Existing Systems

To contextualize the proposed RadonFAN system within the existing literature, Table 8 provides a comparative overview of representative IoT- and AI-based radon monitoring, mitigation and prediction systems. In this table, the authors refer to RadonFAN-DC as the system that integrates the AI system in the RadonFAN microcontroller. Algorithms employed by the systems include rule-based algorithms, statistical methods (Autoregressive Integrated Moving Average (ARIMA)) and AI models (LSTM, Bidirectional LSTM (BiLSTM) and 1D-CNN). RT refers to Real-Time.

Table 8. Comparison of radon monitoring, mitigation and prediction systems, including description, algorithm, capability to act in real-time (RT), and their advantages and disadvantages.

Study	Description	Monitoring	Mitigation	Prediction	Algorithm	RT	Advantages	Disadvantages
[12]	Alert system	✓	✗	✗	Rule-based	✓	Low-cost, IoT	No mitigation, only alerts
[13]	Continuation of [12]	✓	✓	✗	Rule-based	✓	Low-cost, IoT	No forecasting, rule-based
[14]	Alert and mitigation system	✓	✓	✗	Rule-based	✓	Low-cost, IoT	No forecasting, rule-based
[27]	Short-term prediction	✗	✗	✓ ^{Reg}	LSTM, BiLSTM	✗	Forecast horizon analysis	No mitigation
[28]	Long-term prediction	✗	✗	✓ ^{Reg}	ARIMA, LSTM	✗	Long-term radon analysis	No mitigation, no IoT integration
RadonFAN-DC	This work	✓	✓	✓ ^{Cl}	1D-CNN	✓	Unified IoT + AI + mitigation, low-cost	Not deployed

^{Reg} Regression task, ^{Cl} Classification task.

As shown in Table 8, only a limited number of studies address indoor radon within a complete monitoring–mitigation or prediction framework, and those that do generally do not support concurrent prediction alongside real-time monitoring and mitigation. Rule-based systems primarily focus on monitoring and reactive mitigation, offering low-cost, IoT-enabled, real-time operation; however, they rely on fixed thresholds and lack predictive capabilities. In contrast, data-driven prediction models are typically trained offline and evaluated over either short-term (hours) or long-term (years) forecasting horizons. While short-term predictors could, in principle, support real-time decision-making, long-term forecasting models are primarily informative, serving to characterize radon dynamics and inform strategic or policy-level mitigation planning rather than immediate control actions. RadonFAN-DC addresses this gap by integrating real-time monitoring, predictive inference, and mitigation within a single IoT-based framework. On top of that, the authors explored switching from regression-based prediction to classification-based state forecasting. A current limitation of the proposed approach is that RadonFAN-DC has not yet been deployed in an operational environment; currently, RadonFAN-RS remains the deployed system, with RadonFAN-DC proposed as its data-driven successor for future deployment.

4. Conclusions

This study introduced RadonFAN, a low-cost Internet of Things system for continuous radon monitoring and mitigation, integrating multi-sensor acquisition, embedded control hardware, cloud communication, and data-driven decision-making. In its initial deployment, the RadonFAN decision-making algorithm was a rule-based system (RadonFAN-RS), governed by fixed and manually tuned parameters. While this approach enabled rapid deployment and interpretability, it proved to be highly sensitive to parameter definition,

requiring repeated manual adjustment and exhibiting limited adaptability across different operating conditions. Most critically, the rule-based system did not learn from the underlying temporal dynamics of radon time series, resulting in both unnecessary fan activations and delayed responses under certain scenarios. These limitations motivated the development of an AI-based control strategy capable of replacing the rule system within the embedded architecture.

In contrast to most existing radon-related studies, which predominantly address passive monitoring, reactive mitigation or offline analysis, this work advances toward closed-loop radon mitigation by integrating automated fan actuation within an IoT-enabled system. The proposed RadonFAN-DC framework combines real-time monitoring with data-driven inference to support preventive control decisions that account for both safety and energy efficiency. While deep learning techniques for radon prediction have been explored primarily in offline settings, this study systematically evaluates their integration into a real-time mitigation pipeline. As such, RadonFAN-DC represents a step towards the practical deployment of data-driven radon control systems, bridging the gap between monitoring-focused solutions and predictive, action-oriented mitigation. On top of that, general predictive radon frameworks are regression-oriented. Here, the authors explore switching to classification for a more direct integration of the AI models into the RadonFAN system, addressing a research direction that remains relatively underexplored.

The experimental results obtained from two underground galleries highlight several key findings. First, radon concentration dynamics were characterized by strong persistence, leading to unstable multi-step forecasts. Under these conditions, the Regression-to-Classification (R2C) pipeline suffered from error accumulation during the regression pretraining stage, which propagated into the downstream classification task and degraded control performance. By explicitly characterizing indoor radon time series as exhibiting near random-walk behavior, this study demonstrates that conventional regression-based forecasting is limited for preventive control. Instead, the proposed Direct Classification (DC) approach reframes radon mitigation as a decision-oriented time-series classification problem, yielding superior robustness. Further analyses showed that feature richness and temporal context, rather than network architecture complexity, constitute the primary performance bottleneck. When compared with the original rule-based fan activation strategy, the DC model achieved a more favorable balance between safety maintenance and energy consumption, particularly in relatively stable environments.

Despite its contributions, this study has several important limitations that should be explicitly acknowledged. First, the results are site-specific, as the evaluation was conducted in two underground galleries within the same facility. Indoor radon dynamics are strongly influenced by building architecture, ventilation pathways, geological conditions, and mitigation configurations, which may limit the direct generalizability of the reported performance to other environments. Second, the deep learning models rely exclusively on radon concentration time series as input features. While this design choice favors simplicity, low deployment cost, and robustness to sensor failures, it restricts predictive capability under near random-walk conditions, where exogenous drivers play a dominant role. Third, although the DC approach outperformed both the rule-based and R2C strategies, its effectiveness may decrease in scenarios characterized by abrupt radon increases. In such cases, purely data-driven decision-making may require complementary safety mechanisms. Finally, the deep learning models were primarily evaluated through offline analysis, and long-term online adaptation effects, such as concept drift, seasonal changes, and hardware aging, remain to be assessed in real operational conditions.

Future work should address these limitations through several concrete research directions. Multi-site deployments across buildings with diverse structural and geological

characteristics are essential to assess robustness and scalability. Incorporating additional environmental variables, such as pressure, temperature and humidity, may help mitigate near unit-root behavior. Hybrid control strategies combining rule-based safety constraints with AI-driven preventive control represent a promising avenue for improving robustness in rapidly changing environments. Moreover, long-term online deployment will enable the evaluation of closed-loop adaptation under real-world conditions. Together, these extensions would strengthen both the scientific rigor and practical relevance of RadonFAN as a deployable radon mitigation solution.

Author Contributions: Conceptualization, L.A., F.R., M.G., J.J.A., and S.A.; methodology, L.A., J.J.A., and S.A.; software, L.A.; validation, L.A., F.R., M.G., J.J.A., and S.A.; formal analysis, L.A.; investigation, L.A., F.R., M.G., J.J.A., and S.A.; resources, M.G., J.J.A., and S.A.; data curation, L.A. and S.A.; writing—original draft preparation, L.A.; writing—review and editing, F.R., M.G., J.J.A., and S.A.; visualization, L.A.; supervision, F.R., M.G., J.J.A., and S.A.; project administration, S.A.; funding acquisition, S.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by MICIU/AEI/10.13039/501100011033 and by ERDF/EU, Grant PID2024-159276OB-C41.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original data and the main code used in this study are openly available at the following repositories: data (<https://digital.csic.es/handle/10261/410979>) and code (<https://digital.csic.es/handle/10261/410989>), accessed on 3 February 2026.

Acknowledgments: During the preparation of this manuscript, the authors used ChatGPT (OpenAI, GPT-5.1) solely for language rephrasing and improving clarity in English as non-native speakers. The authors have reviewed and edited all generated text and take full responsibility for the content of this publication.

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

Abbreviations

The following abbreviations are used in this manuscript:

AI	Artificial Intelligence
DL	Deep Learning
LSTM	Long Short Time Memory
MQTT	Message Queuing Telemetry Transport
GPIO	General Purpose Input/Output
ACF	AutoCorrelation Function
PACF	Partial AutoCorrelation Function
ADF	Augmented Dickey–Fuller
KPSS	Kwiatkowski–Phillips–Schmidt–Shin
ARIMA	Autoregressive Integrated Moving Average
kNN	k-Nearest Neighbor
RS	Rule System
DC	Direct Classification
R2C	Regression-to-Classification
1D-CNN	One-Dimensional Convolutional Neural Network
MLP	MultiLayer Perceptron
MSE	Mean Squared Error
BCE	Binary Cross Entropy
BiLSTM	Bidirectional Long Short Time Memory

Appendix A

Figure A1 shows the ACF plots for a total of 360 lags (2.5 days) for radon raw time series of Galleries 1 and 2, with the goal of showing minor daily seasonality with autocorrelation values increasing around lags 144 and 288.

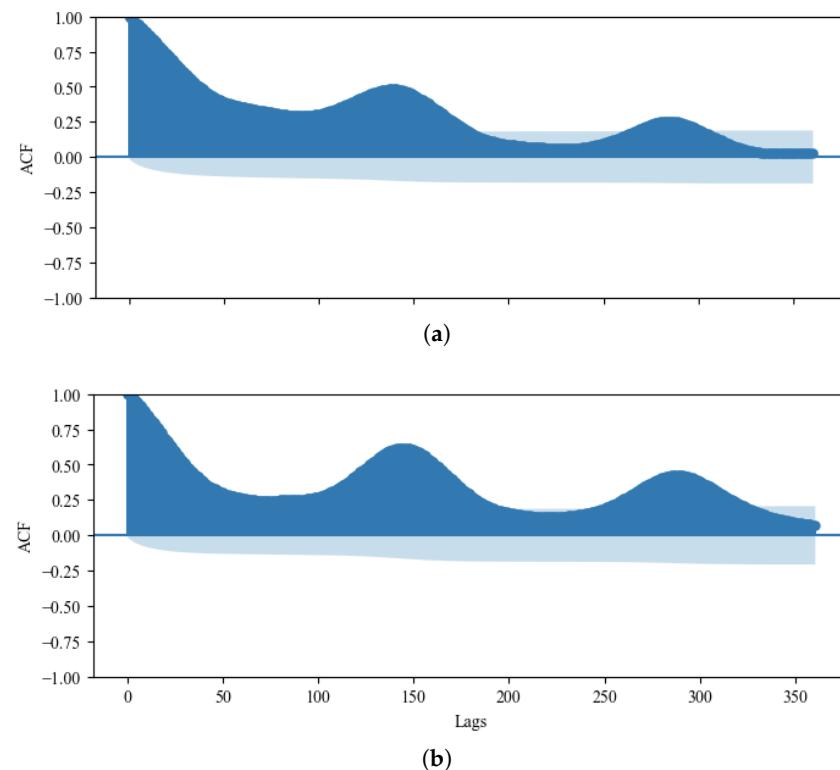


Figure A1. ACF of raw time series with lags = 360 for (a) Gallery 1; (b) Gallery 2. Darker shaded area represents the range of observed autocorrelation values at different time lags. Lighter shaded area represents 95% confidence intervals.

Appendix B

A feature importance and correlation analysis was performed after deriving several features from the raw radon time series. Features studied include Radon (R), Cumulative Sum of radon values in the lookback window (CumSum), smooth Radon Values (Rv), maximum Rt in lookback window (Rt_max), Variance of last three points (V3), Variance of last six points (V6), Slope of last point (S), Smooth Slope Values (Sv), Slope of last three points (S3), Slope of last six points (S6), Difference between S and S3 (Acc), Ratio of mean radon values of last three points with respect to mean of last lookback window (MARatio), Difference between Rv and Rt (Dt), Number of instances above $0.75 \times T$ in lookback window (NumAb075T), difference between Last Rt value and mean Rt values in lookback window (LastRelRt) and difference between last slope and mean slope values in lookback window (LastRelS). Final features selected for the DCF4 model correspond to those with correlation above 0.5 with the binary label (y, last row of Figure A2), with the exception of smooth radon values (Rv), as it shows a correlation of 1 with R and would be redundant. The feature selection based on ablation analysis (Figure A3) identified features whose removal would increase false negatives, highlighting their relevance for maintaining system security. Consequently, the last set of features incorporated in the DCF4 model consists of R, CumSum, Rt_max and NumAb075T.

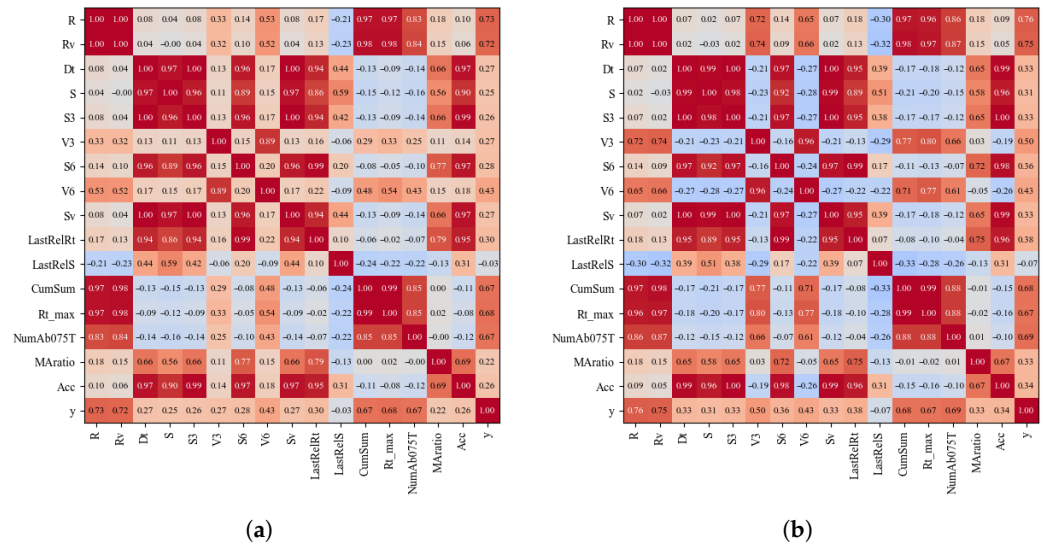


Figure A2. Feature correlation matrices including input features and binary label (y) for (a) Gallery 1; (b) Gallery 2. Colors indicate correlation values, ranging from -1 (blue) to 1 (red).

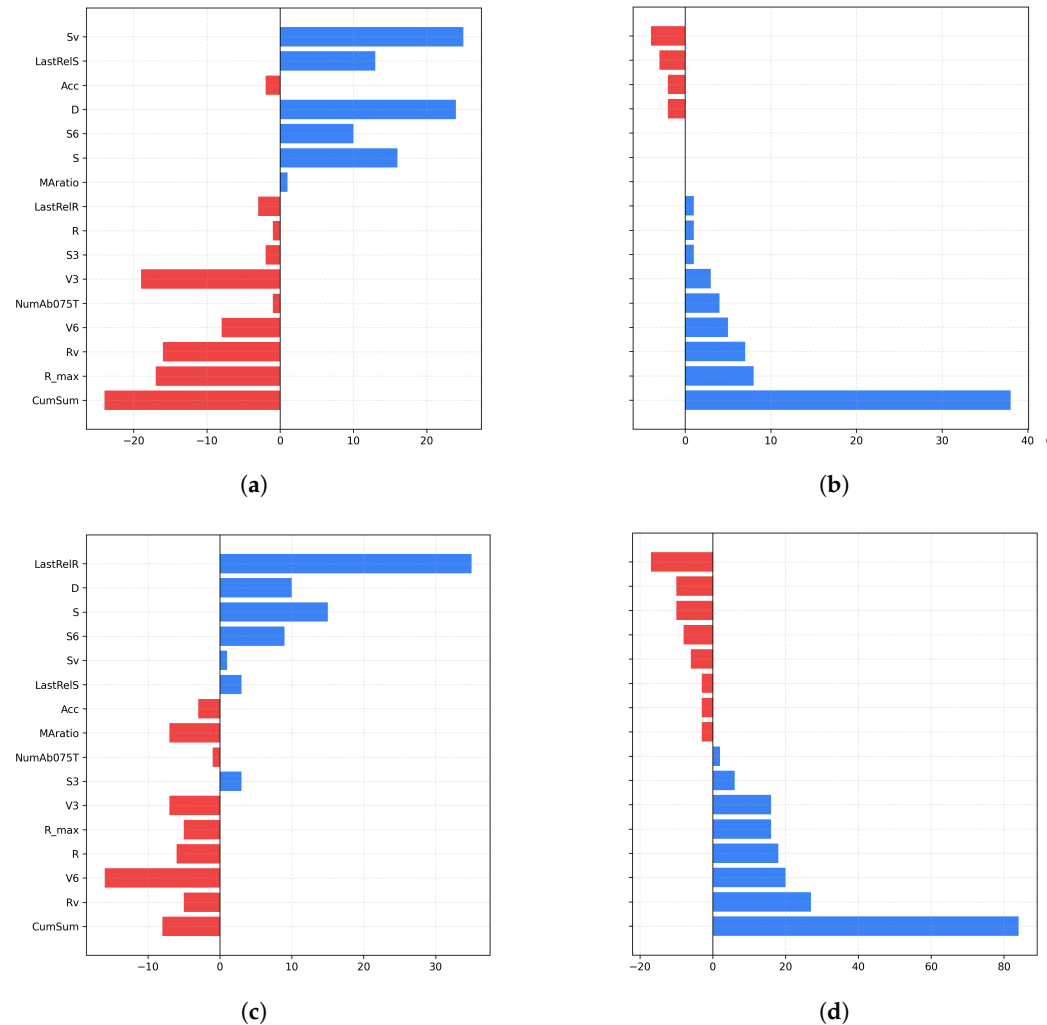


Figure A3. Feature importance analysis based on feature ablation, where y-axis shows the removed feature and x-axis the variation in false predictions. Red indicates a decrease in false predictions, while blue an increase. False predictions correspond to variations in: (a) False Positives for Gallery 1; (b) False Negatives in Gallery 1; (c) False Positives for Gallery 2; (d) False Negatives in Gallery 2.

Appendix C

This appendix reports hard classification metrics and efficiency metrics for the DC model and its modifications (Table A1) and soft classification metrics for the R2C, DC and DC modifications models (Table A2) for Gallery 1 (G1) and 2 (G2).

Table A1. Hard classification metrics (Pr, Rec, F1, Brier) with 95% CI for selected DC modifications.

Model	G1				G2			
	Pr	Rec	F1	Brier	Pr	Rec	F1	Brier
DC	0.645 [0.583–0.709]	0.979 [0.951–1.000]	0.778 [0.778–0.778]	0.039 [0.039–0.039]	0.929 [0.907–0.950]	0.980 [0.967–0.991]	0.954 [0.954–0.954]	0.024 [0.024–0.024]
DCMP	0.662 [0.598–0.725]	0.986 [0.963–1.000]	0.792 [0.792–0.792]	0.039 [0.039–0.039]	0.924 [0.903–0.947]	0.978 [0.964–0.989]	0.950 [0.950–0.950]	0.026 [0.026–0.026]
DCRes	0.634 [0.569–0.695]	0.993 [0.976–1.000]	0.774 [0.774–0.774]	0.042 [0.042–0.042]	0.923 [0.902–0.944]	0.981 [0.969–0.992]	0.951 [0.951–0.951]	0.023 [0.023–0.023]
DCBN	0.712 [0.647–0.775]	0.951 [0.910–0.982]	0.814 [0.814–0.814]	0.030 [0.030–0.030]	0.967 [0.950–0.982]	0.968 [0.953–0.982]	0.968 [0.968–0.968]	0.019 [0.019–0.019]
DCnDp	0.653 [0.588–0.718]	0.972 [0.944–0.994]	0.781 [0.781–0.781]	0.037 [0.037–0.037]	0.939 [0.918–0.958]	0.972 [0.957–0.985]	0.955 [0.955–0.955]	0.023 [0.023–0.023]
DCF16	0.668 [0.605–0.734]	1.000 [1.000–1.000]	0.801 [0.801–0.801]	0.035 [0.035–0.035]	0.933 [0.912–0.953]	0.981 [0.970–0.992]	0.957 [0.957–0.957]	0.023 [0.023–0.023]
DCF4	0.630 [0.566–0.692]	1.000 [1.000–1.000]	0.773 [0.773–0.773]	0.045 [0.045–0.045]	0.923 [0.901–0.944]	0.974 [0.961–0.987]	0.948 [0.948–0.948]	0.028 [0.028–0.028]
DCfc12	0.691 [0.638–0.746]	0.777 [0.723–0.833]	0.731 [0.731–0.731]	0.066 [0.066–0.066]	0.955 [0.937–0.972]	0.948 [0.931–0.965]	0.951 [0.951–0.951]	0.032 [0.032–0.032]
DCfc3	0.752 [0.692–0.810]	0.910 [0.869–0.951]	0.824 [0.824–0.824]	0.033 [0.033–0.033]	0.965 [0.949–0.979]	0.981 [0.968–0.991]	0.973 [0.973–0.973]	0.015 [0.015–0.015]
DCIb24	0.598 [0.537–0.657]	1.000 [1.000–1.000]	0.749 [0.749–0.749]	0.050 [0.050–0.050]	0.919 [0.895–0.942]	0.968 [0.954–0.981]	0.943 [0.943–0.943]	0.029 [0.029–0.029]
DCIb6	0.670 [0.604–0.730]	0.965 [0.932–0.993]	0.791 [0.791–0.791]	0.034 [0.034–0.034]	0.935 [0.914–0.953]	0.978 [0.965–0.989]	0.956 [0.956–0.956]	0.022 [0.022–0.022]

Table A2. Soft classification metrics for R2C, DC and DC variations for Galleries 1 (G1) and 2 (G2). Up (↑) arrows indicate the desired direction of each variable, where higher values are preferred.

	G1			G2		
	Soft Prec ↑	Soft Rec ↑	Soft F1 ↑	Soft Prec ↑	Soft Rec ↑	Soft F1 ↑
R2C	0.470	1.000	0.640	0.634	1.000	0.776
DC	0.659	1.000	0.794	0.939	0.990	0.964
DCMP	0.671	1.000	0.803	0.936	0.990	0.962
DCRes	0.638	1.000	0.779	0.935	0.994	0.964
DCndp	0.671	1.000	0.803	0.946	0.980	0.963
DCF16	0.668	1.000	0.801	0.943	0.993	0.967
DCF4	0.630	1.000	0.773	0.932	0.984	0.957
DCfc12	0.723	0.813	0.766	0.963	0.957	0.960
DCfc3	0.815	0.986	0.893	0.977	0.992	0.985
DCIb24	0.598	1.000	0.749	0.931	0.980	0.955
DCIb6	0.694	1.000	0.819	0.945	0.989	0.966

References

1. Kholopo, M.; Rathebe, P.C. Radon exposure assessment in occupational and environmental settings: An overview of instruments and methods. *Sensors* **2024**, *24*, 2966. [CrossRef] [PubMed]
2. World Health Organization. Radon and Health. Online Fact Sheet. 2023. Available online: <https://www.who.int/news-room/fact-sheets/detail/radon-and-health> (accessed on 3 February 2026).
3. Rizo-Maestre, C.; Echarri-Iribarren, V.; Prado-Govea, R.; Pujol-López, F. Radon Gas as an Indicator for Air Quality Control in Buried Industrial Architecture: Rehabilitation of the Old Británica Warehouses in Alicante for a Tourist Site. *Sustainability* **2019**, *11*, 4692. [CrossRef]
4. Rizo Maestre, C.; Echarri Iribarren, V. The Importance of Checking Indoor Air Quality in Underground Historic Buildings Intended for Tourist Use. *Sustainability* **2019**, *11*, 689. [CrossRef]
5. Hamza, Z.M.; Jasim, A.S.; Heydaryan, K.; Kadhim, S.A. Quantitative Assessment of Uranium and Alpha-Emitting Radionuclides in Lung Cancer Patients Using CR-39 Detectors: Influence of Gender and Smoking Behavior. *Appl. Radiat. Isot.* **2025**, *225*, 112069. [CrossRef]

6. Lorenzo-Gonzalez, M.; Ruano-Ravina, A.; Torres-Duran, M.; Kelsey, K.T.; Provencio, M.; Parente-Lamelas, I.; Piñeiro-Lamas, M.; Varela-Lema, L.; Perez-Rios, M.; Fernandez-Villar, A.; et al. Lung cancer risk and residential radon exposure: A pooling of case-control studies in northwestern Spain. *Environ. Res.* **2020**, *189*, 109968. [\[CrossRef\]](#)
7. Council of the European Union. *Council Directive 2013/59/Euratom of 5 December 2013 Laying Down Basic Safety Standards for Protection Against the Dangers Arising from Exposure to Ionising Radiation*; Council of the European Union: Brussels, Belgium, 2014; Article 74.
8. Antignani, S.; Venoso, G.; Ampollini, M.; Caprio, M.; Carpentieri, C.; Di Carlo, C.; Caccia, B.; Hunter, N.; Bochicchio, F. A 10-year follow-up study of yearly indoor radon measurements in homes, review of other studies and implications on lung cancer risk estimates. *Sci. Total Environ.* **2021**, *762*, 144150. [\[CrossRef\]](#)
9. Sicilia, I.; Aparicio, S.; González, M.; Anaya, J.J.; Frutos, B. Radon Transport, Accumulation Patterns, and Mitigation Techniques Applied to Closed Spaces. *Atmosphere* **2022**, *13*, 1692. [\[CrossRef\]](#)
10. Abad, L.; Ramonet, F.; Ortega, J.; Sanz-Honrado, P.; Gonzalez, M.; Anaya, J.J.; Aparicio, S. Sensor Calibration in Drone-Based Environmental Monitoring: A Machine Learning Approach. In Proceedings of the IEEE International Conference on Cyber Humanities (IEEE CH), Florence, Italy, 8–10 September 2025.
11. Idrees, Z.; Zheng, L. Low cost air pollution monitoring systems: A review of protocols and enabling technologies. *J. Ind. Inf. Integr.* **2020**, *17*, 100123. [\[CrossRef\]](#)
12. Alvarellos, A.; Gestal, M.; Dorado, J.; Rabuñal, J.R. Developing a Secure Low-Cost Radon Monitoring System. *Sensors* **2020**, *20*, 752. [\[CrossRef\]](#)
13. Alvarellos, A.; Chao, A.; Rabuñal, J.; García-Vidaurrázaga, M.; Pazos, A. Development of an Automatic Low-Cost Air Quality Control System: A Radon Application. *Appl. Sci.* **2021**, *11*, 2169. [\[CrossRef\]](#)
14. Blanco-Novoa, O.; Fernández-Caramés, T.M.; Fraga-Lamas, P.; Castedo, L. A Cost-Effective IoT System for Monitoring Indoor Radon Gas Concentration. *Sensors* **2018**, *18*, 2198. [\[CrossRef\]](#)
15. Box, G.E.; Jenkins, G.M.; Reinsel, G.C.; Ljung, G.M. *Time Series Analysis: Forecasting and Control*; John Wiley & Sons: Hoboken, NJ, USA, 2015.
16. Hyndman, R.J.; Athanasopoulos, G. *Forecasting: Principles and Practice*; OTexts: Melbourne, Australia, 2018.
17. Gasperis, G.D.; Facchini, S.D. A Comparative Study of Rule-Based and Data-Driven Approaches in Industrial Monitoring. *arXiv* **2025**, arXiv:2509.15848. [\[CrossRef\]](#)
18. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nature* **1986**, *323*, 533–536. [\[CrossRef\]](#)
19. LeCun, Y.; Boser, B.; Denker, J.S.; Henderson, D.; Howard, R.E.; Hubbard, W.; Jackel, L.D. Backpropagation applied to handwritten zip code recognition. *Neural Comput.* **1989**, *1*, 541–551. [\[CrossRef\]](#)
20. Bahdanau, D.; Cho, K.; Bengio, Y. Neural machine translation by jointly learning to align and translate. *arXiv* **2014**, arXiv:1409.0473.
21. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In *Advances in Neural Information Processing Systems, Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–7 December 2017*; Curran Associates Inc.: Red Hook, NY, USA, 2017; Volume 30.
22. Subramaniam, S.; Raju, N.; Ganesan, A.; Rajavel, N.; Chenniappan, M.; Prakash, C.; Pramanik, A.; Basak, A.K.; Dixit, S. Artificial Intelligence Technologies for Forecasting Air Pollution and Human Health: A Narrative Review. *Sustainability* **2022**, *14*, 9951. [\[CrossRef\]](#)
23. Lim, B.; Zohren, S. Time-series forecasting with deep learning: A survey. *Philos. Trans. R. Soc. A* **2021**, *379*, 20200209. [\[CrossRef\]](#)
24. Kök, İ.; Şimşek, M.U.; Özdemir, S. A deep learning model for air quality prediction in smart cities. In Proceedings of the 2017 IEEE International Conference on Big Data (Big Data), Boston, MA, USA, 11–14 December 2017; pp. 1983–1990. [\[CrossRef\]](#)
25. Reddy, V.N.; Mohanty, S. Deep Air: Forecasting Air Pollution in Beijing, China. *Environ. Sci.* **2017**. Available online: https://www.ischool.berkeley.edu/sites/default/files/sproject_attachments/deep-air-forecasting_final.pdf (accessed on 3 February 2026).
26. Bui, T.C.; Le, V.D.; Cha, S.K. A Deep Learning Approach for Forecasting Air Pollution in South Korea Using LSTM. *arXiv* **2018**, arXiv:1804.07891. [\[CrossRef\]](#)
27. Mpinga, V.; Rosado da Cruz, A.M.; Lopes, S. Forecasting Short-Term Indoor Radon: A Machine Learning Approach Using LSTM Networks. In Proceedings of the 2023 18th Iberian Conference on Information Systems and Technologies (CISTI), Aveiro, Portugal, 20–23 June 2023. [\[CrossRef\]](#)
28. Khan, S. Application of Deep Learning LSTM and ARIMA Models in Time Series Forecasting: A Methods Case Study analyzing Canadian and Swedish Indoor Air Pollution Data. *Austin. J. Med. Oncol.* **2022**, *9*, 1073. [\[CrossRef\]](#)
29. Ismail Fawaz, H.; Forestier, G.; Weber, J.; Idoumghar, L.; Muller, P.A. Deep learning for time series classification: A review. *Data Min. Knowl. Discov.* **2019**, *33*, 917–963. [\[CrossRef\]](#)
30. Hoshen, Y. Time Series Anomaly Detection by Cumulative Radon Features. *arXiv* **2022**, arXiv:2202.04067. [\[CrossRef\]](#)

31. Cerqueira, V.; Torgo, L. Exceedance Probability Forecasting via Regression for Significant Wave Height Prediction. *arXiv* **2024**, arXiv:2206.09821. [[CrossRef](#)]
32. Sjösten, L. A Comparative Study of the KPSS and ADF Tests in Terms of Size and Power. Bachelor's Thesis, Uppsala University, Uppsala, Sweden, 2022.
33. Hamilton, J.D. *Time Series Analysis*; Princeton University Press: Princeton, NJ, USA, 2020.
34. Valcarce, D.; Alvarellos, A.; Rabuñal, J.R.; Dorado, J.; Gestal, M. Machine Learning-Based Radon Monitoring System. *Chemosensors* **2022**, *10*, 239. [[CrossRef](#)]
35. Welsing, A.; Nienkötter, A.; Jiang, X. Exponential Weighted Moving Average of Time Series in Arbitrary Spaces with Application to Strings. In *Structural, Syntactic, and Statistical Pattern Recognition: Joint IAPR International Workshops, S+SSPR 2020, Padua, Italy, 21–22 January 2021*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 45–54. [[CrossRef](#)]
36. Kirichenko, P.; Ibrahim, M.; Balestrierio, R.; Bouchacourt, D.; Vedantam, R.; Firooz, H.; Wilson, A.G. Understanding the Detrimental Class-level Effects of Data Augmentation. *arXiv* **2023**, arXiv:2401.01764. [[CrossRef](#)]
37. Loshchilov, I.; Hutter, F. Decoupled weight decay regularization. *arXiv* **2017**, arXiv:1711.05101.
38. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. *arXiv* **2015**, arXiv:1502.01852. [[CrossRef](#)]
39. Kannan, A. Benefits of Multiobjective Learning in Solar Energy Prediction. *arXiv* **2023**, arXiv:2301.12282. [[CrossRef](#)]
40. Reihanifar, M.; Danandeh Mehr, A.; Tur, R.; Ahmed, A.T.; Abualigah, L.; Dąbrowska, D. A New Multi-Objective Genetic Programming Model for Meteorological Drought Forecasting. *Water* **2023**, *15*, 3602. [[CrossRef](#)]
41. Terven, J.; Cordova-Esparza, D.M.; Romero-González, J.A.; Ramírez-Pedraza, A.; Chávez-Urbiola, E. A comprehensive survey of loss functions and metrics in deep learning. *Artif. Intell. Rev.* **2025**, *58*, 195. [[CrossRef](#)]
42. Salles, R.; Lima, J.; Reis, M.; Coutinho, R.; Pacitti, E.; Masegla, F.; Akbarinia, R.; Chen, C.; Garibaldi, J.; Porto, F.; et al. SoftED: Metrics for soft evaluation of time series event detection. *Comput. Ind. Eng.* **2024**, *198*, 110728. [[CrossRef](#)]
43. Ismail Fawaz, H.; Lucas, B.; Forestier, G.; Pelletier, C.; Schmidt, D.F.; Weber, J.; Webb, G.I.; Idoumghar, L.; Muller, P.A.; Petitjean, F. InceptionTime: Finding AlexNet for time series classification. *Data Min. Knowl. Discov.* **2020**, *34*, 1936–1962. [[CrossRef](#)]

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