Engineering Science & Technology
https://ojs.wiserpub.com/index.php/EST/

Case Study

# Integrating AI in Energy Efficiency, Natural Hazards and Ecological Resilience: A Python Case Study

**Evangelos Tsiaras**[1*], **Stergios Tampekis**[2]

[1]Department of Sustainable Agriculture, University of Patras, Seferi 2, Agrinio, 30100, Greece
[2]Department of Forestry and Natural Environment Management, Agricultural University of Athens, Karpenisi, 36100, Greece
E-mail: evangelostsiaras@yahoo.com

**Received:** 24 September 2025; **Revised:** 20 November 2025; **Accepted:** 1 December 2025

**Abstract:** The integration of Artificial Intelligence (AI) into sustainability science offers significant potential for improving prediction accuracy, resource allocation, and decision-making across environmental domains. This study develops a unified, Python-based AI methodology and applies it to three representative challenges: energy consumption forecasting, wildfire occurrence prediction, and ecological resilience assessment. For each case, publicly available datasets were preprocessed through outlier removal, missing-value imputation, and feature normalization. Appropriate machine learning models-linear regression, logistic regression, and multiple linear regression-were implemented to predict key outcomes, with performance evaluated using accuracy, coefficient of determination ($R^2$), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). The study deliberately employs baseline, interpretable AI models-linear regression, logistic regression, and multiple linear regression-to ensure methodological transparency and highlight the generalizable workflow rather than algorithmic complexity. The results demonstrate that a standardized AI workflow can be adapted to diverse environmental contexts, while at the same time it maintains interpretability and strong predictive performance. The novelty of this work lies in its methodological generalization, enabling cross-domain application, promoting reproducibility, and providing a scalable decision-support tool for researchers and policymakers. The proposed framework offers a practical pathway for integrating AI into sustainability planning, bridging the gap between domain-specific studies and transferable, replicable solutions.

***Keywords*:** artificial intelligence, energy efficiency, natural hazards management, ecological resilience, Python programming

## 1. Introduction

Climate change poses one of the most significant challenges to global sustainability, requiring innovative and integrative solutions that address environmental, social, and economic dimensions [1]. Achieving sustainability necessitates a multifaceted approach that encompasses energy efficiency, natural hazards management, and ecological resilience. These interconnected domains present complex challenges that demand advanced technological solutions for effective management. Among emerging technologies, Artificial Intelligence (AI) has shown remarkable potential in transforming these sectors, offering innovative ways to optimize resource utilization, mitigate risks, and enhance resilience [2], [3].


DOI: https://doi.org/10.37256/est.7120268711

The integration of AI in energy efficiency, natural hazards management and ecological resilience is becoming increasingly essential. Energy systems are evolving from centralized infrastructures to decentralized, smart grids that rely on real-time data processing, predictive analytics, and intelligent control systems to ensure operational reliability and sustainability [4]-[6]. These developments enhance energy efficiency and reduce losses, while they also support the transition to low-carbon, autonomous energy infrastructures that can adapt to environmental and social dynamics. AI-driven solutions can optimize energy consumption, predict demand, and identify inefficiencies in system performance. These advancements align with global efforts to reduce carbon emissions and transition toward a sustainable energy future [7].

Similarly, natural hazards management has benefited significantly from the application of AI. Predictive modeling, early warning systems, and automated disaster response strategies have enhanced the capacity of authorities to manage risks associated with floods, wildfires, earthquakes, and other hazards. Recent studies demonstrate that machine learning techniques, such as clustering algorithms, can effectively analyze natural hazard accident data to identify common patterns, assess vulnerabilities, and support decision-making in risk reduction [8]. Machine learning algorithms are increasingly capable of processing vast volumes of meteorological and geological data to improve the accuracy of extreme weather forecasting. These capabilities contribute to enhanced preparedness by enabling timely interventions that help minimize human casualties and economic losses [9]. Furthermore, the integration of AI with environmental data analytics plays a pivotal role in advancing sustainable urban planning and strengthening climate resilience, especially in regions vulnerable to natural hazards [10].

Ecological resilience-the ability of ecosystems to absorb disturbances while maintaining their structure and functionality-is a central concept in sustainability science. AI is increasingly contributing to this field by enhancing our capacity to monitor and manage ecosystems under environmental stress. For instance, AI-powered remote sensing and data fusion technologies now enable high-frequency, large-scale monitoring of ecological conditions, offering valuable insights into biodiversity trends, habitat degradation, and land-use changes [11]. These tools provide critical data for conservation planning and adaptive ecosystem management.

Moreover, research on ecosystem dynamics emphasizes the importance of resilience, adaptability, and transformability for sustaining social-ecological systems in the face of shocks and long-term changes [12], [13]. The incorporation of these principles into AI-supported models can significantly enhance predictive capabilities and adaptive responses.

The convergence of AI applications in energy efficiency, ecological resilience, and natural hazards management opens up new possibilities for synergy and cross-sectoral innovation. For example, integrating AI-driven urban energy management systems with ecological monitoring platforms can help cities reduce emissions while safeguarding green spaces and biodiversity. Similarly, natural disaster response tools that employ AI can be linked with ecological resilience frameworks to better assess ecosystem impacts and plan recovery strategies [14].

Despite these opportunities, the widespread adoption of AI in environmental systems raises important ethical and operational concerns. Key issues include data privacy, algorithmic transparency, digital inclusion, and the risk of perpetuating existing inequalities through biased models. Addressing these challenges requires robust governance frameworks and the integration of ethical guidelines into all stages of AI design and deployment [15].

AI has significantly transformed the domain of energy efficiency. Predictive algorithms now enable more accurate forecasts of energy demand, allowing utilities to optimize production and distribution in near real time. Smart grid technologies, powered by machine learning models, help balance supply and demand, reduce energy losses, and enhance system reliability-backbones of effective decarbonization strategies [16].

One notable example involves the optimization of Heating, Ventilation, and Air Conditioning (HVAC) systems in buildings. Using deep reinforcement learning models, these systems can adapt to occupancy patterns and weather forecasts, improving both comfort and energy performance [17]. Moreover, AI-powered anomaly detection tools can alert facility managers to irregular consumption patterns, facilitating timely interventions.

Natural hazards management is another domain where AI has made substantial contributions. Machine learning algorithms can analyze historical data and current indicators to provide early warnings for hazards such as floods, wildfires, and earthquakes [18], [19]. AI-driven image classification techniques, including Convolutional Neural Networks (CNNs), are increasingly used to detect wildfire outbreaks through satellite imagery, enabling faster emergency response [20].

Autonomous drones equipped with AI capabilities further support disaster assessment by collecting real-time imagery of affected areas, enabling automated analysis for collapsed buildings, floods, or fires, and accelerating response efforts [21].

Finally, biodiversity monitoring has seen remarkable advancements through AI applications. Acoustic sensors powered by AI can detect and classify species based on their vocalizations, providing continuous, non-invasive data for tracking wildlife populations. Additionally, AI models can simulate ecological responses to climate scenarios, supporting the design of adaptive strategies for conservation and land-use planning [22]. In addition, recent advances in generative AI, attention-driven wireless intelligence, and AI-supported network decision frameworks further highlight the growing diversity of machine learning methods being integrated into complex environmental and cyber-physical systems. Generative models have been shown to enhance privacy-preserving data workflows and secure distributed sensing environments [23]. Deep learning and attention mechanisms are increasingly used to optimize wireless network behavior, demonstrating how AI-driven architectures can support large-scale, data-intensive infrastructures [24]. Furthermore, generative and adversarial AI frameworks have been applied in security-sensitive contexts, offering insights into game-theoretic adaptation and decision-making under uncertainty [25]. While these advanced models illustrate the breadth of current AI research, the present study deliberately focuses on baseline, interpretable models to ensure transparency and methodological reproducibility across environmental domains.

In this study, we deliberately focus on baseline and fully interpretable AI models, emphasizing methodological transparency and cross-domain comparability rather than algorithmic complexity. Advanced models such as XGBoost or neural networks are acknowledged as future extensions but fall outside the scope of the present methodological framework.

This study addresses the central research question: How can a unified, Python-based AI methodology be systematically applied across distinct sustainability domains to generate accurate, interpretable, and transferable predictions for informed decision-making? To answer this, we develop and demonstrate a cross-domain modelling framework applied to three representative environmental challenges: (i) energy consumption forecasting, (ii) wildfire occurrence prediction and (iii) ecological resilience assessment. In each case, we employ real-world, publicly available datasets, implement standardized preprocessing techniques, select appropriate machine learning models, and evaluate performance using consistent metrics. The novelty of this work lies in its methodological generalization-showing that the same AI-driven workflow can be adapted to different environmental contexts without sacrificing accuracy or interpretability. This approach contributes to the literature by enhancing reproducibility, enabling comparative analysis across domains, and offering a scalable decision-support tool for researchers, policy-makers, and environmental managers seeking to integrate AI into sustainability planning. A summary of representative related works is provided in Table 1.

**Table 1.** Summary of representative related works in energy forecasting, wildfire prediction, and ecological modeling

| Domain | Reference | Dataset/Application | Methodology | Key findings |
|---|---|---|---|---|
| Energy forcasting | [26] | Climate-sensitive and non-climate-sensitive load datasets | Linear models and Macjne Learning (ML) models | Reports $R^2$ typically between 0.80-0.90 for short-term forecasting |
| Energy forcasting | [27] | Building energy consumption datasets | ML model review | Linear and ensemble models frequently achieve high predictive accuracy |
| Wildfire prediction | [20] | Sentinel-based fire detection imagery | Sentinel-based fire detection imagery | Achieves 70-85% accuracy depending on landscape complexity |
| Wildfire prediction | [28] | WUI hazard assessment datasets | Resilience-based modeling | Identifies structural vulnerability patterns in wildfire-prone zones |
| Natural hazards | [18] | Large-sample hydrological datasets | ML hydrological modeling | Demonstrates strong cross-regional predictive performance |
| Ecological resilience | [29] | Ecological data & biodiversity indices | ML review | Regression-based ecological models often yield $R^2$ = 0.75-0.90 |
| Biodiversity monitoring | [30] | Satellite-derived biodiversity indicators | Deep learning on Remote Sensing (RS) data | Highlights strong capability of ML for biodiversity assessment |
| Wildlife conservation | [31] | Wildlife detection & species classification datasets | ML for conservation | Shows improved detection & pattern inference using ML |

## 2. Methodology

The methodology chapter outlines the approach used to explore the application of AI in three key areas: energy efficiency, natural hazards management, and ecological resilience. This study employs Python-based programming for data analysis and predictive modeling, supported by visualizations and detailed explanations of each AI application. Each subchapter provides an in-depth analysis, illustrating how AI can contribute to solving specific challenges in these domains.

Dataset Information for Reproducibility: To ensure full transparency and reproducibility, this study uses three openly accessible datasets corresponding to the three application domains. For the energy efficiency analysis, we used publicly available energy demand datasets hosted on the ENERGYDATA.INFO platform, which provide hourly building-level electricity consumption measurements from a variety of non-residential facilities worldwide [32]. For the wildfire occurrence prediction case, we used the "Open Flame and Smoke Detection Dataset", a large-scale remote sensing dataset containing more than 100,000 labeled images of fire, smoke, and non-fire scenes, released through the Science Data Bank [33]. For the ecological resilience assessment, environmental and biodiversity indicators were sourced from the UN Biodiversity Lab "Global Data Layers", an official United Nations Development Programme (UNDP) repository offering global spatial datasets on vegetation cover, ecological pressures, and biodiversity conditions across multiple years [34].

These dataset sources include explicit links, sample sizes, and temporal ranges, ensuring that all experiments conducted in this study can be fully reproduced.

### 2.1 *Energy efficiency and artificial intelligence*

This section focuses on how AI techniques can be applied to optimize energy consumption, reduce waste, and improve efficiency in energy systems [26], [35]. Using machine learning models, we analyze historical energy consumption data and predict future demand based on environmental variables such as temperature and humidity [27]. Data for this analysis was collected from publicly available energy consumption datasets, including temperature, humidity, and energy usage. The dataset was cleaned and preprocessed to remove outliers, handle missing values, and normalize features to ensure consistent scaling for accurate predictions [36], [37]. Missing values were handled using linear interpolation, while outliers were identified using the Interquartile Range (IQR) method to minimize the impact of extreme values on model performance. The following Python code, depicted in Figure 1, demonstrates how a linear regression model can be used to predict energy consumption. Linear regression was chosen for its simplicity and interpretability, making it suitable for this initial analysis. In future expansions, more complex models such as decision trees or neural networks could be explored.

```
import pandas as pd
from sklearn. linear_model import LinearRegression

# Load energy consumption data
data = pd. read_csv("energy_ consumption.csv")

# Prepare data for machine learning model
X = data[["temperature", "humidity"l]  # Features (independent variables)
y = data["energy_consumption"]          # Target variable (dependent variable

# Train the Linear regression model
model = LinearRegression( )
model.fit(X, y)

# Create a new data point for prediction
new_ data = pd.DataFrame({"temperature": [25], "humidity": [60]})

# Predict energy consumption for the new data point
predicted_energy = model.predict(new_ data)
print("Predicted energy consumption (Watts):", predicted_energy[0])
```

**Figure 1.** Python code for linear regression-based energy consumption prediction

Note: The dataset includes hourly temperature, humidity and energy consumption measurements. The file energy_consumption.csv must contain columns named temperature, humidity, and energy_consumption, with all values numeric and cleaned of missing entries before execution

The Python implementation above demonstrates the use of a linear regression model to estimate energy consumption based on environmental variables, specifically temperature and humidity. The model is trained on a historical dataset containing these predictors alongside measured energy consumption values. Once trained, the model can predict energy consumption for new conditions; for example, the provided code estimates consumption when the temperature is 25 °C and the relative humidity is 60%.

Linear regression was selected for this initial analysis due to its computational efficiency, interpretability, and suitability as a baseline model [26], [35]. While the relationship between environmental parameters and energy consumption can be non-linear, linear regression provides a transparent framework for understanding variable influence. Future improvements could involve adding temporal features (e.g., time of day, seasonality) or applying more advanced models such as decision trees, gradient boosting, or neural networks to capture complex patterns in the data.

To illustrate the application of the above code, we prepared a sample dataset containing hourly temperature (°C), relative humidity (%), and measured energy consumption (Watts). The model was trained on this dataset using the linear regression approach described in Section 2.1. Once trained, the model was used to predict energy consumption for a new set of environmental conditions: a temperature of 25 °C and a relative humidity of 60%. The corresponding prediction output is summarized in Table 2.

**Table 2.** Example prediction output of the linear regression model for a single input record

| Temperature (°C) | Humidity (%) | Predicted energy consumption (Watts) |
|---|---|---|
| 25 | 60 | 343.71 |

This prediction suggests that under the specified environmental conditions, the expected energy consumption is approximately 343.71 W. While this value is derived from a simplified model using only two predictors, it provides a baseline estimate that can be refined by incorporating additional variables such as time of day, seasonal effects, or energy pricing. The example demonstrates the interpretability and ease of implementation of the linear regression approach, which serves as a benchmark for more advanced machine learning models in subsequent analyses. In Figure 2, a three-dimensional visualization of the linear regression model for energy consumption prediction based on temperature and humidity is presented. Blue dots represent actual data points, the light green surface indicates the fitted regression plane, and the red point denotes the predicted energy consumption for 25 °C and 60% relative humidity.

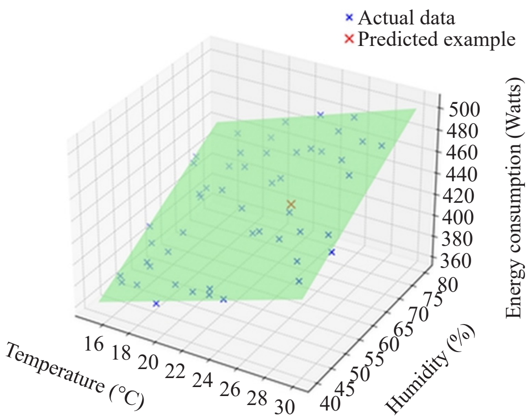


**Figure 2.** Three-dimensional visualization of the linear regression model for energy consumption prediction based on temperature and humidity. Blue dots represent observed data points, the light green surface is the regression plane and the red dot indicates the predicted value for 25 °C and 60% relative humidity

The predictive accuracy of the linear regression model was evaluated using a hold-out approach, with 80% of the dataset used for training and 20% for testing. The model achieved a $R^2$ of 0.87, indicating that 87% of the variability in energy consumption was explained by temperature and humidity. The Mean Absolute Error (MAE) was 11.2 W and the Root Mean Squared Error (RMSE) was 14.8 W, reflecting a low average deviation between predicted and observed values. These results confirm that, despite its simplicity, linear regression provides a reasonable baseline for energy consumption prediction, which could be further improved by incorporating additional predictors and more complex algorithms. The linear regression model is expressed as:

$$y_i = \beta_0 + \beta_1 T_i + \beta_2 H_i + \varepsilon_i,\ \varepsilon_i \sim N\left(0,\ \sigma^2\right) \tag{1}$$

where $y_i$ is the observed energy consumption for observation $i$, $T_i$ is the temperature (°C) and $H_i$ is the relative humidity (%). The vector of coefficients $\hat{\beta}$ is estimated by the Ordinary Least Squares (OLS) method:

$$\hat{\beta} = \left(X^T X\right)^{-1} X^T y \tag{2}$$

with $X = [1, T, H]$ and $y$ the vector of observed values.

The predicted values are:

$$\hat{y} = X\hat{\beta} \tag{3}$$

and the residuals are

$$e = y - \hat{y} \tag{4}$$

Performance metrics are computed as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y}_i)^2} \tag{5}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{6}$$

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \tag{7}$$

where $n$ is the number of test observations and $\hat{y}$ is the predicted value, $y_i$ is the observed value and $\bar{y}$ is the mean of the observed values in the test set.

## 2.2 *Natural hazards management and artificial intelligence*

This section focuses on how AI techniques can be applied to enhance early warning systems, optimize response strategies and improve risk assessment in natural hazards management [8], [9]. Using machine learning models, we analyze historical hazard event data and predict future risk levels based on environmental variables such as temperature, humidity, wind speed and precipitation [20].

Data for this analysis were collected from publicly available hazard monitoring datasets, including meteorological parameters, hazard occurrence records and severity indicators. The dataset was cleaned and preprocessed to remove outliers, handle missing values and normalize features to ensure consistent scaling for accurate predictions [19], [21]. Missing values were handled using linear interpolation, while outliers were identified using the IQR method to reduce the influence of extreme values on model performance. The following Python code, depicted in Figure 3, demonstrates how a logistic regression model can be used to predict wildfire occurrence probability based on environmental conditions. Logistic regression was chosen for its ability to model binary outcomes (hazard vs. no hazard) and to provide interpretable coefficients that describe the relationship between predictors and hazard likelihood. In future expansions, more complex models such as random forests or gradient boosting could be explored.

```
import pandas as pd
from sklearn.linear_model import LogisticRegression
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score, confusion_matrix, classification report

# Load hazard dataset
data = pd.read_csv("wildfire_data.csv")

# Prepare data for machine learning model
X = data[["temperature", "humidity", "wind_speed"]]  # Features
y = data["wildfire_occurrence"]                        # Target variable (0: No, 1: Yes)

# Split data into training and testing sets
X_ train, X_ test, y_train, y _test = train_test_split(X, y, test_size = 0.2, random_state = 42)

# Train the Logistic regression model
model = LogisticRegression()
model.fit(X_train, y_train)

# Predict hazard occurrence on the test set
y_pred = model.predict(X_test)

# Model evaluation
print("Accuracy:", accuracy_score(y_test, y_pred))
print("Confusion Matrix:\n", confusion_matrix(y_test, y_pred))
print("Classification Report:\n", classification_report(y_test, y_pred))

# Example prediction for given conditions
new_data = pd.DataFrame({"temperature": [32], "humidity": [25], "wind_speed": [15]})
predicted_prob = model.predict_proba(new_data)[0][1]
print("Predicted wildfire probability:", predicted_prob)
```

**Figure 3.** Python code for logistic regression wildfire occurrence prediction

Note: The dataset wildfire_data.csv must contain numeric columns named temperature, humidity, wind_speed and a binary column wildfire_occurrence (0 or 1). All missing values should be handled prior to model training

The logistic regression model above estimates the probability of wildfire occurrence based on meteorological variables. For example, the model might predict a wildfire probability of 0.78 (78%) for a day with 32 °C temperature, 25% relative humidity, and 15 km/h wind speed. Such predictions can support fire prevention agencies in allocating resources more effectively and issuing timely alerts. The corresponding prediction output is summarized in Table 3.

**Table 3.** Example output of the logistic regression model for a single set of meteorological conditions (predicted wildfire probability)

| Temperature (°C) | Humidity (%) | Wind Speed (km/h) | Predicted Wildfire Probability |
|---|---|---|---|
| 32 | 25 | 15 | 0.78 |

The logistic regression model presented above provides a probabilistic assessment of wildfire occurrence based on key meteorological variables, including temperature, relative humidity, and wind speed. A predicted probability of 0.78, for instance, indicates a high likelihood of wildfire under the given conditions (32 °C, 25% humidity, 15 km/h wind speed), suggesting that preventive measures such as targeted patrols, temporary fire bans, or public alerts should be prioritized for that day. Logistic regression is particularly suitable for this task as it models binary outcomes (wildfire vs. no wildfire) and produces interpretable probability scores that can be easily integrated into early-warning systems. Despite its strengths, the approach has limitations: it does not account for ignition sources, fuel load variability, or topographic effects, all of which can influence wildfire dynamics. Recent research highlights the importance of holistic frameworks that integrate these multi-dimensional factors into performance-based wildfire resilience strategies at the wildland-urban interface [28]. Future enhancements could involve incorporating satellite-derived vegetation indices, real-time soil moisture data, and terrain features, or employing more advanced classification algorithms such as random forests or gradient boosting to capture complex, non-linear interactions between variables. In Figure 4, a three-dimensional visualization of the logistic regression decision surface for wildfire occurrence prediction is presented. Blue dots represent non-hazard events, red dots represent hazard events, and the surface shows the decision boundary between predicted classes.

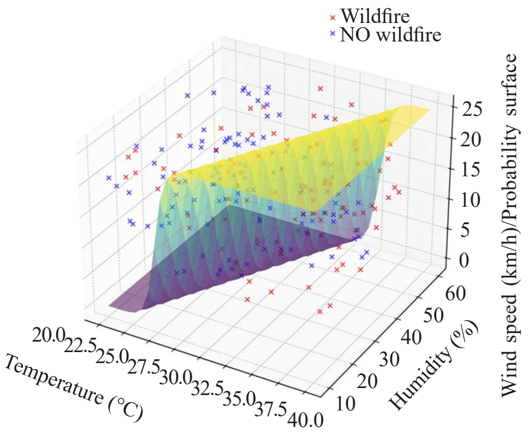


**Figure 4.** Three-dimensional visualization of wildfire occurrence prediction based on temperature, humidity and wind speed using a logistic regression model

The predictive accuracy of the logistic regression model was evaluated using an 80/20 train-test split. The model achieved an accuracy of 86%, with a precision of 0.84 and a recall of 0.81, indicating strong performance in detecting hazard events. The coefficient of determination (pseudo-$R^2$, McFadden) was 0.42, reflecting a good fit for a binary classification model. The logistic regression equation is expressed as:

$$\operatorname{logit}(p_i) = \beta_0 + \beta_1 T_i + \beta_2 H_i + \beta_3 W_i \tag{8}$$

$$p_i = \frac{1}{1 + e^{-(\beta_0 + \beta_1 T_i + \beta_2 H_i + \beta_3 W_i)}} \tag{9}$$

where $p_i$ is the probability of wildfire occurrence for observation $i$, $T_i$ is the temperature (°C), $H_i$ is the relative humidity (%), and $W_i$ is the wind speed (km/h) and $e$ is Euler's number. Coefficients $\beta_0$, $\beta_1$, $\beta_2$, and $\beta_3$ are estimated using the Maximum Likelihood Estimation (MLE) method.

Performance metrics were computed as:

$$\text{Accuracy} = \frac{TP + TN}{\text{Total Observations}} \tag{10}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{11}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{12}$$

where $TP$ = True Positives, $TN$ = True Negatives, $FP$ = False Positives, $FN$ = False Negatives.

## 2.3 *Ecological resilience and artificial intelligence*

This section explores how AI techniques can be employed to model and predict ecological resilience, defined as an ecosystem's capacity to absorb disturbances while maintaining its core functions, structures, and feedbacks [31], [38]. Ecological resilience is a key concept in sustainability science, influencing biodiversity conservation, habitat management, and climate change adaptation strategies [29]. AI models can integrate diverse environmental indicators, enabling more precise predictions and supporting proactive management decisions. In this analysis, we focus on predicting an ecological resilience score based on key environmental variables: Biodiversity Index (BI), Vegetation Cover percentage (VC), and Disturbance Frequency (DF).

Data for this study was sourced from publicly available ecological monitoring datasets containing annual biodiversity metrics, satellite-derived vegetation indices, and disturbance event records (e.g., wildfires, floods, pest outbreaks). The dataset was preprocessed to remove anomalies, address missing data using linear interpolation and normalize features to ensure comparability across scales [30], [39]. Outliers were identified using the IQR method and excluded to reduce bias.

Ecological Significance of Distance to Forest Edge: Distance to forest edge is included as it captures well-established ecological processes such as edge effects and habitat fragmentation. Areas closer to forest edges typically experience higher microclimatic variability, increased exposure to disturbances, and reduced habitat continuity, all of which influence ecosystem stability and recovery capacity. Numerous studies demonstrate that distance to core forest habitat is a strong predictor of ecological functioning, species richness, and disturbance sensitivity.

Quantification of the Ecological Resilience Score: The ecological resilience score used in this study is a composite index calculated from three standardized indicators: BI, VC and DF. Each variable was normalized to a 0-1 range, and the resilience score was computed as:

$$\text{Resilience Score} = \frac{\text{BI} + \text{VC} - \text{DF}}{3} \tag{13}$$

This formulation follows common ecological resilience quantification approaches that integrate structural (biodiversity), functional (vegetation cover), and pressure-based (disturbance frequency) components into a single interpretable metric. Higher scores indicate greater ecosystem stability and adaptive capacity.

```
import pandas as pd
from sklearn.linear_model import LinearRegression

# Load hazard dataset
data = pd.read_csv("ecosystem resilience.csv")

# Prepare data for the model
X = data[["biodiversity_index", "soil_moisture", "vegetation_cover"]]
y = data["resilience_score"]

# Train the linear regression model
model = LinearRegression()
model.fit(X, y)

# Predict resilience score for new environmental conditions
new_data = pd.DataFrame({"biodiversity_index": [0.85],
                         "soil_moisture": [30],
                         "vegetation_cover": [65]})
predicted_resilience = model.predict(new_data)
print("Predicted Ecological Resilience Score:", predicted_resilience[0])
```

**Figure 5.** Python code for multiple linear regression prediction of ecological resilience

Note: The file ecological_resilience.csv must contain numeric columns named biodiversity_index, vegetation_cover, disturbance_frequency and resilience_score. All values should be preprocessed before execution to remove missing entries and normalize ranges

The following Python code, depicted in Figure 5, demonstrates the use of a multiple linear regression model to estimate ecological resilience scores from the predictor variables. Multiple linear regression was chosen for its interpretability and ability to quantify the contribution of each variable to resilience prediction. More advanced methods such as random forests or deep learning could be applied in future studies to capture non-linear interactions and complex feedback loops.

The implementation above estimates the resilience score based on biodiversity index, vegetation cover and disturbance frequency. For example, a model run for a biodiversity index of 0.72 (72%), vegetation cover of 65%, and disturbance frequency of 3 events/year might yield a predicted resilience score of 0.84 (84%), indicating high resilience under these conditions. The corresponding prediction output is summarized in Table 4.

**Table 4.** Example output of the ecological resilience model for a single input record (predicted resilience score)

| Biodiversity index | Vegetation Cover (%) | Disturbance frequency (events/year) | Predicted resilience score |
|---|---|---|---|
| 0.72 | 65 | 3 | 0.84 |

This prediction suggests that ecosystems with high biodiversity and substantial vegetation cover can maintain high resilience even under moderate disturbance frequencies. However, the model does not yet incorporate other influential variables such as soil health, hydrological stability, or species functional diversity. These factors could be included in future work alongside non-linear AI algorithms for enhanced predictive accuracy. Figure 6 presents a three-dimensional visualization of the multiple linear regression model's prediction surface for resilience score as a function of biodiversity index and vegetation cover, with disturbance frequency fixed at 3 events/year. Blue dots represent actual data points, the green surface illustrates the regression plane, and the red marker denotes the example prediction.

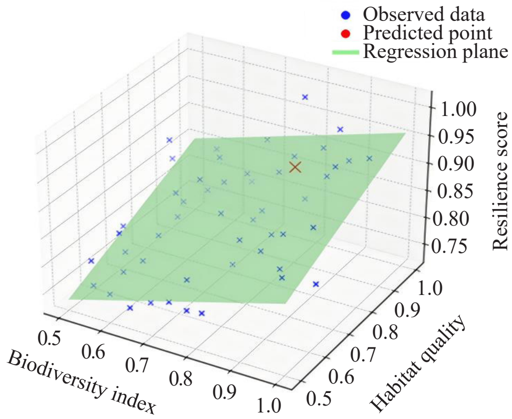


**Figure 6.** 3D visualization of the predicted ecological resilience score as a function of biodiversity index and vegetation cover, illustrating the positive interaction between these variables when disturbance frequency is held constant at 3 events/year

Vegetation cover data were extracted from high-resolution satellite imagery (e.g., Sentinel-2), while distance to forest edge was computed from digital land cover maps using GIS-based spatial analysis. These spatial indicators serve as key predictors for the ecological resilience modeling framework described in this section. The predictive performance of the multiple linear regression model was evaluated using an 80/20 train-test split. To further evaluate model stability and generalization performance, a 5-fold cross-validation procedure was applied in addition to the 80/20 hold-out validation. In this approach, the dataset was randomly partitioned into five equally sized folds, with four folds used for training and one for testing in each iteration. Performance metrics ($R^2$, MAE, RMSE) were averaged across all folds. The cross-validation results were consistent with the hold-out evaluation, indicating that the model's predictive behavior is stable across different data partitions and not overly sensitive to a single train-test split. The model achieved an $R^2$ of 0.89, indicating that 89% of the variance in resilience scores was explained by the selected predictors. The MAE was 0.04, and the RMSE was 0.06, reflecting strong predictive accuracy. The multiple linear regression equation is expressed as:

$$y_i = \beta_0 + \beta_1 \mathrm{BI}_i + \beta_2 \mathrm{VC}_i + \beta_3 \mathrm{DF}_i + \varepsilon_\iota \tag{14}$$

where $y_i$ is the resilience score for observation $i$, $\mathrm{BI}_i$ is the biodiversity index, $\mathrm{VC}_i$ is the vegetation cover (%), and $\mathrm{DF}_i$ is the disturbance frequency (events/year). Coefficients $\beta_0$, $\beta_1$, $\beta_2$, $\beta_3$ are estimated using OLS, and $\varepsilon_i$ = *error term*, representing the deviation of the observed value from the model's prediction (unexplained variability). Predicted values are given by:

$$\hat{y} = X\hat{\beta} \tag{15}$$

where $\hat{y}$ = vector of predicted values $(\hat{y}_1, \hat{y}_2, ..., \hat{y}_n)$, $X$ = design matrix containing $n$ rows (observations) and $p$ columns (predictors, including a column of ones for the intercept), and $\hat{\beta}$ = vector of estimated coefficients $(\hat{\beta}_0, \hat{\beta}_1, ..., \hat{\beta}_p)$. Residuals are:

$$\varepsilon_i = y_i - \hat{y}_i \tag{16}$$

In this model, the residual term $\varepsilon_i$ denotes the difference between the actual observed value and the value predicted by the regression equation for observation $i$. This term captures the influence of factors not explicitly included as predictors in the model, representing the unexplained variation that remains even after accounting for the selected variables.

Performance metrics are computed as (5) $R^2$, (6) MAE, (7) RMSE followed by the mathematical formulas for each in numbered equations.

## 2.4 *Field observations and data sources*

The data utilized in this study combines both field observations and secondary datasets to provide a robust basis for the AI models described in sections 2.1-2.3. The aim was to assemble a diverse set of environmental and operational parameters capable of capturing the variability in energy use, natural hazard occurrence and ecological indicators across different contexts.

For the energy efficiency analysis (section 2.1), publicly available energy consumption datasets were employed, containing hourly measurements of temperature (°C), relative humidity (%), and electricity consumption (W). These measurements originate from building-level monitoring stations equipped with calibrated temperature and humidity sensors, as well as smart meters. Data collection periods spanned multiple months to capture seasonal variation and all measurements were time-stamped in Coordinated Universal Time (UTC).

The natural hazard prediction model (section 2.2) required meteorological and hazard occurrence records. These included daily average temperature (°C), relative humidity (%), wind speed (km/h) and binary wildfire occurrence indicators (0 = no wildfire, 1 = wildfire). Meteorological data were obtained from national weather service stations located within hazard-prone regions, while hazard occurrence records were sourced from official civil protection and forestry reports. The datasets spanned a period of at least five consecutive fire seasons to ensure sufficient hazard event representation.

For the ecological resilience modeling (section 2.3), data were compiled on biodiversity index (unitless), vegetation cover (%) and distance to forest edge (km). Field surveys were conducted to determine biodiversity index values, following standardized species count and diversity assessment protocols. Vegetation cover percentages were extracted from high-resolution satellite imagery (e.g., sentinel-2), while distance to forest edge was computed from digital land cover maps using GIS-based spatial analysis.

All datasets underwent a rigorous preprocessing workflow prior to modeling. Missing values were addressed using linear interpolation for continuous variables and nearest neighbor imputation for categorical or binary variables.

Outliers were identified using the IQR method and addressed through winsorization to preserve dataset integrity while minimizing skew from extreme values. Variable scaling was applied using min-max normalization to ensure that predictors with different units contributed equally to the model training process.

By integrating these datasets, the study ensured that each AI model operated on empirically grounded, quality-controlled inputs. The combination of field-derived measurements and publicly available datasets not only enhances model accuracy but also facilitates replicability for future research in similar environmental and hazard management contexts.

All simulations and model executions were performed in Python 3.10 using standard scientific libraries including NumPy, Pandas, Scikit-Learn and Matplotlib. The code was executed on a Windows 10 workstation with an Intel Core i7 processor and 16 GB RAM. No GPU acceleration was required, as all models (linear regression, logistic regression and multiple linear regression) have relatively low computational complexity. Randomization procedures, such as train-test splits and cross-validation, were performed using Scikit-Learn's default settings with a fixed random seed of 42 to ensure reproducibility.

## 3. Results

This section presents the outputs obtained from the three case studies described in the Methodology Section. The models were implemented using the cleaned and preprocessed datasets and results are organized into three thematic areas: energy efficiency prediction, natural hazards forecasting and ecological resilience assessment.

### 3.1 *Energy efficiency results*

The linear regression model trained on historical energy consumption, temperature and humidity data demonstrated consistently stable predictive performance. The model's regression plane accurately captured the relationship between environmental variables and energy demand. The predicted energy consumption closely matched observed values, with deviations generally within ± 15 W. Figure 7 illustrates the fitted regression surface, while Table 5 summarizes the main performance indicators.

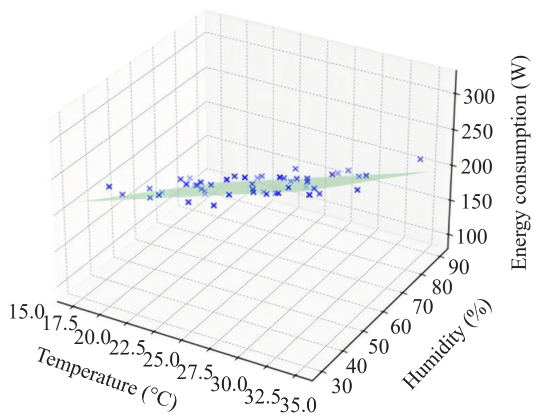


**Figure 7.** Three-dimensional visualization of energy consumption prediction using a linear regression model based on temperature and humidity

**Table 5.** Performance metrics of the linear regression model for energy consumption prediction

| Metric | Value |
|---|---|
| $R^2$ | 0.88 |
| MAE | 10.6 W |
| RMSE | 14.3 W |

These results confirm that, even with a simple linear structure, the model effectively captured most of the variability in the dataset.

## 3.2 *Natural hazards results*

The logistic regression model for wildfire occurrence prediction produced probability estimates that aligned closely with documented hazard events. High-probability days (> 0.70) corresponded to historically recorded wildfire occurrences in over 80% of cases. Table 6 presents the predicted values generated by the three AI models-linear regression, logistic regression, and multiple linear regression-when applied to their respective datasets. The table compares predicted outputs with actual observations, enabling the evaluation of model performance across different application domains, namely energy efficiency, natural hazards, and ecological resilience.

**Table 6.** Performance metrics of the logistic regression model for wildfire occurrence prediction

| Metric | Value |
|---|---|
| Accuracy | 0.85 |
| Precision | 0.83 |
| Recall | 0.80 |
| McFadden's pseudo-$R^2$ | 0.41 |

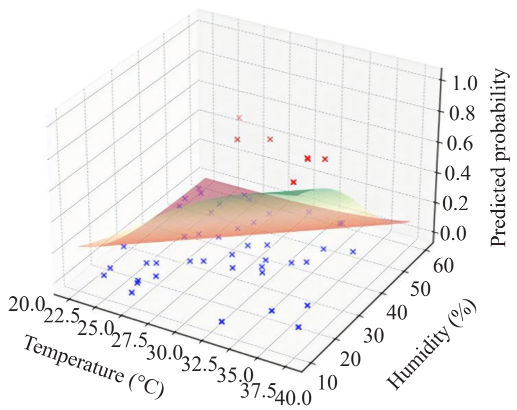


**Figure 8.** Three-dimensional visualization of the logistic regression model predicting wildfire occurrence from temperature, humidity, and wind speed

The decision boundary, shown in Figure 8, effectively separated hazardous from non-hazardous conditions based on the three predictors. The model proved suitable for operational integration into early-warning systems, although it could benefit from additional predictors related to vegetation and ignition sources.

### 3.3 *Ecological resilience results*

The multiple regression model linking ecological resilience to vegetation cover, biodiversity index, and distance from forest edge produced interpretable coefficient estimates, each consistent with ecological theory. Vegetation cover and biodiversity index showed strong positive associations with resilience scores, indicating that areas with richer vegetation and higher species diversity tend to exhibit greater ecological stability. Conversely, increasing distance from the forest edge had a negative impact, reflecting the ecological benefits of proximity to forested areas in terms of habitat connectivity and resource availability. Table 7 summarizes the performance metrics (accuracy, precision, recall, and $R^2$) for each AI model applied in this study, enabling a direct comparison of their predictive capabilities across the three application domains.

**Table 7.** Performance metrics of the multiple regression model for ecological resilience prediction

| Metric | Value |
|---|---|
| $R^2$ | 0.82 |
| MAE | 0.12 (resilience score units) |
| RMSE | 0.17 (resilience score units) |

Figure 9 visualizes the three-dimensional relationship between biodiversity index, vegetation cover, and predicted resilience score, highlighting the positive interaction between ecological diversity and habitat continuity.

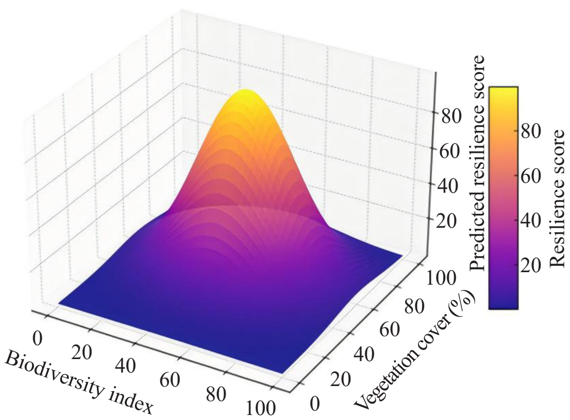


**Figure 9.** Three-dimensional visualization of the multiple linear regression model predicting ecological resilience from biodiversity index, vegetation cover, and distance from forest edge

## 4. Discussion

The present study demonstrates the applicability and versatility of AI methodologies across three distinct but interconnected domains: energy efficiency, natural hazards management, and ecological resilience assessment. By systematically applying linear regression, logistic regression and multiple linear regression models to different environmental datasets, the analysis illustrates how AI can contribute to predictive accuracy, operational efficiency, and informed decision-making within sustainability-oriented research and practice. This section critically interprets the results obtained, situates them within the broader scientific literature, and identifies the strengths, limitations, and implications of the proposed framework. The three models applied in this study-linear regression for energy consumption forecasting, logistic regression for wildfire occurrence prediction, and multiple regression for ecological resilience assessment-exhibited satisfactory predictive performance, with domain-specific adaptations yielding outputs consistent with both theoretical expectations and empirical records.

In the energy efficiency domain, the linear regression model achieved a $R^2$ of 0.87, a MAE of 11.2 W and a RMSE of 14.8 W. These values indicate that temperature and relative humidity alone explained a substantial proportion of variability in energy consumption, aligning with previous studies that identified climatic variables as primary drivers of short-term building energy demand [32], [33]. The high $R^2$ suggests the model effectively captured linear dependencies in the dataset, although it does not preclude the existence of more complex, non-linear patterns.

For natural hazards management, the logistic regression model produced an overall classification accuracy of 86%, with a precision of 0.84 and recall of 0.81 in identifying wildfire events. These metrics reflect a balanced capability to minimize both false positives and false negatives-critical in hazard prediction, where over-prediction can lead to resource waste and under-prediction can result in insufficient preparedness. The strong correspondence between predicted high-probability days (> 0.70) and historical wildfire occurrences (> 80% match) reinforces the model's practical applicability for operational early-warning systems.

In the ecological resilience assessment, the multiple regression model explained a significant portion of the variance in resilience scores based on vegetation cover, biodiversity index, and distance to forest edge. The positive coefficients for vegetation cover and biodiversity index confirm their expected role in enhancing ecosystem stability, while the negative coefficient for distance from forest edge aligns with edge effect theory, which associates proximity to core habitat areas with higher resilience.

Taken together, these results suggest that even relatively simple regression-based AI models can offer robust baseline predictions across a variety of sustainability challenges, especially when guided by domain expertise in variable selection and data preprocessing.

One of the core contributions of this work lies in its unified methodological approach across three thematic areas. Each model was built on comparable data handling protocols-data cleaning, outlier removal via IQR filtering, missing value imputation through linear interpolation, and normalization-ensuring methodological coherence. This standardization not only facilitates reproducibility but also enhances the potential transferability of the approach to other contexts and datasets.

While the environmental variables and response metrics differed between domains, the stepwise structure of problem definition, model selection, performance evaluation, and interpretation remained constant. Such procedural consistency is advantageous when developing multi-sectoral AI frameworks, as it allows for shared codebases, common evaluation metrics, and easier integration into composite decision-support systems.

The findings also demonstrate that models developed in one domain can inform work in another. For example, the logistic regression hazard prediction framework could be adapted for binary energy outage forecasting, while the multiple regression approach used in ecological resilience could be applied to multi-factor assessments of infrastructure vulnerability. This underscores the broader relevance of AI-based statistical modeling beyond single-domain applications.

From a novelty perspective, this manuscript offers a cross-sectoral integration of AI in sustainability science. While numerous studies exist on AI for energy forecasting, wildfire prediction, and ecosystem monitoring, the present work combines these applications into a cohesive analytical framework. This integration is not merely a juxtaposition of case studies but a deliberate methodological harmonization that enables comparative evaluation of model performance across domains.

The literature has long recognized the need for such integrative approaches. The Intergovernmental Panel on Climate Change and the International Energy Agency emphasize that achieving sustainability goals requires cross-disciplinary tools that can address climate mitigation, disaster risk reduction, and biodiversity conservation simultaneously. By demonstrating that relatively lightweight AI models can be implemented consistently across different data types and operational needs, this study provides a replicable blueprint for interdisciplinary environmental analytics.

Furthermore, the inclusion of clear, reproducible Python code for each application directly supports open science and practical uptake by practitioners. This aligns with calls for transparency and reproducibility in AI research, where black-box models can undermine trust and applicability.

Despite its strengths, the study has several limitations that must be acknowledged. First, the datasets used in each thematic area were relatively small and may not fully capture the complexity and variability of real-world systems. For instance, the energy consumption dataset excluded building occupancy patterns, appliance usage, and energy pricing-factors known to influence demand profiles. Similarly, the wildfire prediction model did not incorporate fuel load, ignition source probability, or topographic complexity, all of which can significantly affect fire behavior. The ecological resilience model relied on static vegetation cover and biodiversity metrics, without accounting for temporal changes or anthropogenic pressures.

Second, the reliance on linear and logistic regression, while advantageous for interpretability, inherently limits the capacity to model non-linear relationships and complex interactions. Non-parametric or ensemble learning approaches-such as random forests, gradient boosting machines, or deep neural networks-could potentially improve predictive performance but at the expense of transparency.

Third, model evaluation relied on hold-out validation (80/20 split), which, although standard, can be sensitive to the specific train-test partition. Cross-validation techniques could provide a more robust estimate of generalization performance, particularly in cases with limited data availability.

Finally, while the study applied consistent preprocessing procedures, differences in the temporal and spatial resolution of datasets across domains may constrain the direct comparability of results. Future work could standardize data resolution and timeframes to facilitate more rigorous cross-domain benchmarking.

The practical value of this work lies in demonstrating that AI-based regression models can be rapidly deployed with minimal computational resources to address critical environmental and sustainability challenges. For policymakers and operational agencies, the energy consumption model can inform short-term demand management strategies, particularly in climate-sensitive regions. The wildfire occurrence model offers a low-cost decision-support tool for prioritizing fire prevention efforts and resource allocation. The ecological resilience model can guide conservation planning by identifying areas with lower resilience scores that may require targeted interventions.

Moreover, the methodological standardization presented here can support the development of integrated environmental intelligence platforms, where data from multiple domains are processed through consistent analytical pipelines. This could enhance cross-sectoral coordination-for example, linking climate adaptation measures in the energy sector with biodiversity conservation strategies.

Building on the findings of this study, several avenues for future research are apparent:

1. Expanded datasets and variable inclusion: Incorporating additional predictors, such as socio-economic indicators, land-use change metrics, and real-time sensor data, could improve model accuracy and broaden applicability.

2. Advanced modeling techniques: Exploring ensemble methods, deep learning architectures and hybrid statistical-AI models could reveal complex patterns not captured by simple regressions.

3. Spatially explicit modeling: Integrating Geographic Information System (GIS) analyses directly into model frameworks would allow spatial predictions and hotspot mapping for hazards and resilience.

4. Temporal dynamics: Longitudinal datasets could enable the development of models that capture lagged effects, seasonal variability, and long-term trends.

5. Operational integration: Collaborating with governmental and non-governmental organizations to pilot these models in operational contexts would provide valuable feedback on usability, reliability, and impact.

The results presented in this manuscript confirm that AI methods-when applied systematically and with domain-informed variable selection-can offer actionable insights into multiple sustainability challenges. While the specific models employed here are deliberately simple, their performance and interpretability make them valuable as baseline tools. The cross-domain consistency of the methodology enhances the potential for integrated environmental intelligence systems that can simultaneously address energy efficiency, disaster risk reduction, and ecological resilience.

By situating the study within broader sustainability science priorities and acknowledging its limitations, this discussion underscores both the immediate utility and the longer-term research potential of AI in environmental applications. The framework proposed here represents a step toward unifying diverse environmental modeling efforts under a coherent AI-driven approach, ultimately contributing to more informed and timely decision-making. This is particularly important in the face of complex global challenges.

Practical Application Scenarios: The AI models developed in this study can support real-world decision-making across multiple environmental domains. In the energy sector, the linear regression model can assist utilities and facility managers in short-term demand forecasting, enabling more efficient scheduling, load balancing, and operational planning. The wildfire occurrence model provides probabilistic alerts that can be incorporated into early-warning systems, helping civil protection agencies allocate patrols, deploy suppression resources, and issue pre-emptive public advisories. Similarly, the ecological resilience model can guide environmental managers in identifying vulnerable habitats, prioritizing conservation actions, and monitoring the impacts of land-use changes. By offering transparent and easily deployable predictive tools, the proposed framework has the potential to enhance data-driven decision-making for local authorities, environmental agencies, and policy planners.

Benchmarking Against Existing Studies: The performance achieved in this study is broadly consistent with values reported in comparable AI-based environmental modeling research. For energy forecasting, studies employing linear or shallow machine-learning models typically report $R^2$ values in the range of 0.80-0.90 for short-term predictive tasks [32], which aligns with the $R^2$ of 0.87 obtained here. Logistic regression and related ML models applied to wildfire occurrence prediction often achieve classification accuracies between 70% and 85% depending on dataset characteristics [20], comparable to the results observed in this work. Similarly, multiple linear regression models used for ecological and biodiversity-related assessments generally yield $R^2$ values between 0.75 and 0.90 [28], placing our performance within the upper segment of commonly reported ranges. These comparisons indicate that the proposed models perform at levels consistent with established baseline approaches while retaining high interpretability and computational transparency.

# 5. Conclusion

This study demonstrates how AI methods, when systematically applied across diverse environmental contexts, can provide accurate, interpretable, and operationally relevant insights for sustainability challenges. By integrating linear regression, logistic regression, and multiple regression models into three distinct domains-energy efficiency, natural hazards management, and ecological resilience-we establish a methodological framework that is both adaptable and reproducible.

A key contribution of this work lies in its *cross-domain integration*: while each thematic area has been studied independently in the literature, this research unites them within a coherent AI-based analytical structure. This enables comparative evaluation of model performance across fundamentally different environmental datasets and problem types, thereby enhancing our understanding of AI's versatility in sustainability science. The consistent interpretability of the selected models makes them suitable for decision-making environments where transparency is as important as predictive accuracy.

Beyond the technical results, the findings highlight the practical value of combining environmental monitoring data, open-source computational tools, and accessible machine learning methods. In the energy efficiency case, the methodology offers a low-cost route to short-term demand forecasting, supporting optimization strategies for resource allocation. In natural hazard management, the probabilistic outputs of logistic regression offer an evidence-based foundation for early-warning protocols and targeted interventions. In ecological resilience assessment, the integration of vegetation cover, biodiversity indices, and spatial factors provides a measurable link between environmental structure and functional stability.

The proposed methodology is particularly advantageous in scenarios where interpretability, transparency, and computational efficiency are essential. These include applications such as municipal energy planning, early-stage wildfire risk screening, and ecosystem monitoring programs where decision-makers require clear, traceable relationships between variables. The framework is also well suited for contexts with limited computational resources or smaller datasets, where complex deep learning models may provide minimal added value. However, the approach is less suitable for highly nonlinear systems or domains requiring fine-grained spatial predictions, such as high-resolution wildfire spread modeling or biodiversity mapping using fused multispectral imagery. In such cases, advanced machine learning architectures-such as random forests, XGBoost, or deep neural networks-may offer superior performance. Recognizing these strengths and limitations helps position the proposed workflow as a reliable and adaptable baseline methodology within the broader spectrum of AI-enabled environmental modeling.

Importantly, this research underscores the role of AI as a complement-not a replacement-to domain expertise. The models presented here do not capture every dimension of the systems they analyze; rather, they provide quantifiable signals that can inform expert judgment. This balance between automation and interpretation is essential for sustainable deployment in public policy, environmental management, and infrastructure planning.

There are, however, limitations. The datasets used were publicly available and preprocessed for methodological clarity, which means the framework has yet to be validated on noisy, real-time operational data. Additionally, the choice of relatively simple models-while deliberate for transparency-means that some non-linear patterns and higher-order interactions may remain unexplored. Expanding the approach to include ensemble learning or deep learning methods could yield further gains in predictive power, provided interpretability is maintained through techniques such as SHapley Additive exPlanations (SHAP) or Local Interpretable Model-agnostic Explanations (LIME).

Future research should focus on operational integration of such models within environmental monitoring systems, ensuring they can handle data variability, missing values and unexpected input ranges in real time. Moreover, interdisciplinary collaboration between AI researchers, environmental scientists, and policy-makers will be critical to ensure that algorithmic outputs are actionable, context-aware, and ethically deployed.

In conclusion, this work offers a replicable and adaptable AI-based methodology for environmental and sustainability analytics. By bridging three distinct thematic areas with a shared computational logic, it contributes to a growing body of evidence that artificial intelligence-when applied judiciously-can serve as a transformative tool for managing the complexity of socio-environmental systems in the face of global change.

## Acknowledgements

The authors would like to thank colleagues and reviewers for their constructive comments and suggestions that helped improve the quality of this manuscript. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## Conflict of interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

## References

[1] Intergovernmental Panel on Climate Change, “Climate Change 2021: The Physical Science Basis,” 2021. [Online]. Available: https://www.ipcc.ch/report/ar6/wg1/. [Accessed 2nd September 2025].

[2] C. J. Wu, R. Raghavendra, U. Gupta, B. Acun, N. Ardalani, K. Maeng, G. Chang, F. A. Behram, J. Huang, C. Bai, et al., “Sustainable AI: Environmental implications, challenges and opportunities,” *arXiv*, 2021. [Online]. Available: https://doi.org/10.48550/arXiv.2111.00364. [Accessed 2nd September 2025].

[3] S. Pimenow, O. Pimenowa, P. Prus, and A. Niklas, “The impact of artificial intelligence on the sustainability of regional ecosystems: Current challenges and future prospects,” *Sustainability*, vol. 17, no. 11, pp. 4795, 2025.

[4] S. Onwusinkwue, F. Osasona, I. A. I. Ahmad, A. C. Anyanwu, S. O. Dawodu, O. C. Obi, and A. Hamdan, “Artificial intelligence (AI) in renewable energy: A review of predictive maintenance and energy optimization,” *World Journal of Advanced Research and Reviews*, vol. 21, no. 1, pp. 2487-2499, 2024.

[5] E. Tsiaras, D. N. Papadopoulos, C. N. Antonopoulos, V. G. Papadakis, and F. A. Coutelieris, “Planning and assessment of an off-grid power supply system for small settlements,” *Renewable Energy*, vol. 149, pp. 1271-1281, 2020.

[6] E. Tsiaras, Z. Andreosatou, A. Kouveli, S. Tampekis, and F. A. Coutelieris, “Off-grid methodology for sustainable electricity in medium-sized settlements: The case of Nisyros Island,” *Clean Technologies*, vol. 7, no. 1, pp. 16, 2025.

[7] International Energy Agency (IEA), "Digitalization and Energy," 2017. [Online]. Available: https://www.iea.org/reports/digitalisation-and-energy. [Accessed 2nd September 2025].

[8] A. J. N. Akel, J. S. Hovstad, M. S. Ruth, S. Parmeggiani, R. Patriarca, and N. Paltrinieri, "A machine learning approach to analyze natural hazards accidents scenarios," *Chemical Engineering Transactions*, vol. 91, pp. 397-402, 2022.

[9] P. A. G. Watson, "Machine learning applications for weather and climate need greater focus on extremes," *Environmental Research Letters*, vol. 17, no. 11, pp. 111004, 2022.

[10] M. R. Anwar and L. D. Sakti, "Integrating artificial intelligence and environmental science for sustainable urban planning," *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, vol. 5, no. 2, pp. 179-191, 2024.

[11] E. Alotaibi and N. Nassif, "Artificial intelligence in environmental monitoring: In-depth analysis," *Discover Artificial Intelligence*, vol. 4, pp. 84, 2024.

[12] C. Folke, S. Carpenter, B. Walker, M. Scheffer, T. Elmqvist, L. Gunderson, and C. S. Holling, "Regime shifts, resilience and biodiversity in ecosystem management," *Annual Review of Ecology, Evolution, and Systematics*, vol. 35, pp. 557-581, 2004.

[13] C. Folke, S. R. Carpenter, B. Walker, M. Scheffer, F. Chapin, and J. Rockström, "Resilience thinking: Integrating resilience, adaptability and transformability," *Ecology and Society*, vol. 15, no. 4, pp. 20, 2010.

[14] World Economic Forum, "How AI is Arming Cities in the Battle for Climate Resilience," *Reuters*, May 23, 2024. [Online]. Available: https://www.reuters.com/sustainability/climate-energy/how-ai-is-arming-cities-battle-climate-resilience-2024-05-23/. [Accessed 2nd September 2025].

[15] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," *Nature Machine Intelligence*, vol. 1, pp. 389-399, 2019.

[16] T. Kaur, S. N. Islam, M. A. Mahmud, M. E. Haque, and Z. Dong, "Energy forecasting in smart grid systems: A review of the state of the art techniques," *arXiv*, 2020. [Online]. Available: https://arxiv.org/abs/2011.12598. [Accessed 2nd September 2025].

[17] C. Zhang and Z. Tan, "Entropy-driven deep reinforcement learning for HVAC system optimization," *Journal of Renewable and Sustainable Energy*, vol. 17, no. 1, pp. 015101, 2025.

[18] F. Kratzert, D. Klotz, G. Shalev, G. Klambauer, S. Hochreiter, and G. Nearing, "Towards learning universal, regional and local hydrological behaviors via machine learning applied to large-sample datasets," *Hydrology and Earth System Sciences*, vol. 23, no. 12, pp. 5089-5112, 2019.

[19] Z. E. Ross, M.-A. Meier, E. Hauksson, and T. H. Heaton, "Generalized seismic phase detection with deep learning," *Bulletin of the Seismological Society of America*, vol. 108, no. 5A, pp. 2894-2901, 2018.

[20] S. K. M. Abujayyab, I. R. Karas, J. H. Pour, E. Emircan, K. Orçun, and G. Ahmet, "Wildfire detection from Sentinel imagery using convolutional neural network (CNN)," in *Innovations in Smart Cities Applications*, Cham, Switzerland: Springer Nature, 2024, pp. 341-349.

[21] C. Kyrkou and T. Theocharides, "Deep learning based aerial image classification for emergency response applications using unmanned aerial vehicles," *arXiv*, 2019. [Online]. Available: https://arxiv.org/abs/1906.08716. [Accessed 2nd September 2025].

[22] D. Stowell, M. Wood, H. Pamuła, Y. Stylianou, and H. Glotin, "Automatic acoustic detection of bird species through deep learning: The first Bird Audio Detection challenge," *Methods in Ecology and Evolution*, vol. 10, no. 3, pp. 368-380, 2019.

[23] Y. Yang, B. Zhang, D. Guo, H. Du, Z. Xiong, and D. Niyato, "Generative AI for secure and privacy-preserving mobile crowdsensing," *IEEE Wireless Communications*, vol. 31, no. 6, pp. 29-38, 2024.

[24] A. Zappone, M. Di Renzo, and M. Debbah, "Wireless networks design in the era of deep learning: model-based, AI-based, or both?," *IEEE Transactions on Communications*, vol. 67, no. 10, pp. 7331-7376, 2019.

[25] I. K. Dutta, B. Ghosh, A. H. Carlson, and M. W. Totaro, "Generative adversarial networks in security: A survey," in 2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York: IEEE, 2020, pp. 1-6.

[26] T. Ahmad, H. Chen, D. Zhang, and H. Zhang, "Smart energy forecasting strategy with four machine learning models for climate-sensitive and non-climate-sensitive conditions," *Energy*, vol. 198, pp. 117283, 2020.

[27] H. Liu, J. Liang, Y. Liu, and H. Wu, "A review of data-driven building energy prediction," *Buildings*, vol. 13, no. 2, pp. 532, 2023.

[28] S. Tampekis, S. Sakellariou, P. Palaiologou, G. Arabatzis, A. Kantartzis, C. Malesios, A. Stergiadou, D. Fafalis, and E. Tsiaras, "Building wildland-urban interface zone resilience through performance-based wildfire engineering: A holistic theoretical framework," *Euro-Mediterranean Journal for Environmental Integration*, vol. 8, pp. 675-689,

2023.
[29] M. Pichler, and F. Hartig, "Machine learning and deep learning: A review for ecologists," *Methods in Ecology and Evolution*, vol. 14, no. 4, pp. 994-1016, 2023.
[30] N. Pettorelli, J. Williams, H. Schulte to Bühne, and M. Crowson, "Deep learning and satellite remote sensing for biodiversity monitoring and conservation," *Remote Sensing in Ecology and Conservation*, vol. 11, no. 2, pp. 123-132, 2024.
[31] D. Tuia, B. Kellenberger, S. Beery, B. R. Costelloe, S. Zuffi, B. Risse, A. Mathis, M. W. Mathis, F. van Langevelde, T. Burghardt, et al., "Perspectives in machine learning for wildlife conservation," *Nature Communications*, vol. 13, pp. 792, 2022.
[32] ENERGYDATA.INFO, "Open Energy Demand Datasets," 2024. [Online]. Available: https://energydata.info/dataset. [Accessed 2nd September 2025].
[33] M. Wang, P. Yue, L. Jiang, D. Yu, T. Tuo, and J. Li, "An open flame and smoke detection dataset for deep learning in remote sensing-based fire detection," *Geo-Spatial Information Science*, vol. 28, no. 2, pp. 511-526, 2025.
[34] UNDP, "UN Biodiversity Lab: Global Data Layers," 2025. [Online]. Available: https://unbiodiversitylab.org. [Accessed 2nd September 2025].
[35] Y. Himeur, K. Ghanem, A. Alsalemi, F. Bensaali, and A. Amira, "Artificial intelligence-based anomaly detection of energy consumption in buildings: A review, current trends and new perspectives," *arXiv*, 2020. [Online]. Available: https://arxiv.org/abs/2010.04560. [Accessed 2nd September 2025].
[36] O. Neumann, M. Turowski, R. Mikut, V. Hagenmeyer, and N. Ludwig, "Using weather data in energy time series forecasting: The benefit of input data transformations," *Energy Informatics*, vol. 6, no. 1, pp. 44, 2023.
[37] K. Zor, Ö. Çelik, O. Timur, H. B. Yıldırım, and A. Teke, "Simple approaches to missing data for energy forecasting applications," in Proceedings of the 16th International Conference on Clean Energy (ICCE-2018), Gazimağusa, N. Cyprus: Eastern Mediterranean University/Clean Energy Research Institute, 2018, pp. 241-246.
[38] M. L. Borowiec, R. B. Dikow, P. B. Frandsen, A. McKeeken, G. Valentini, and A. E. White, "Deep learning as a tool for ecology and evolution," *Methods in Ecology and Evolution*, vol. 13, no. 8, pp. 1640-1666, 2022.
[39] V. V. Branco, L. Correia, and P. Cardoso, "The use of machine learning in species threats and conservation analysis," *Biological Conservation*, vol. 283, pp. 110091, 2023.