

Article

Feature Selection to Predict LED Light Energy Consumption with Specific Light Recipes in Closed Plant Production Systems

Martín Montes Rivera ^{1,*}, Nivia Escalante-Garcia ^{2,*}, José Alonso Dena-Aguilar ³, Ernesto Olvera-Gonzalez ² and Paulino Vacas-Jacques ³

¹ Dirección de Posgrados e Investigación, Universidad Politécnica de Aguascalientes, Calle Paseo San Gerardo #201, Fracc. San Gerardo, Aguascalientes 20342, Mexico

² Laboratorio de Iluminación Artificial, Tecnológico Nacional de México/IT de Pabellón de Arteaga, Carretera a la Estación de Rincón Km. 1, Aguascalientes 20670, Mexico; jose.og@pabellon.tecnm.mx

³ Departamento de Ingenierías, Tecnológico Nacional de México/IT de Pabellón de Arteaga, Carretera a la Estación de Rincón Km. 1, Aguascalientes 20670, Mexico; jose.da@pabellon.tecnm.mx (J.A.D.-A.); paulino.vj@pabellon.tecnm.mx (P.V.-J.)

* Correspondence: martin.montes@upa.edu.mx (M.M.R.); nivia.eg@pabellon.tecnm.mx (N.E.-G.)

Abstract: The use of closed growth environments, such as greenhouses, plant factories, and vertical farms, represents a sustainable alternative for fresh food production. Closed plant production systems (CPPSs) allow growing of any plant variety, no matter the year's season. Artificial lighting plays an essential role in CPPSs as it promotes growth by providing optimal conditions for plant development. Nevertheless, it is a model with a high demand for electricity, which is required for artificial radiation systems to enhance the developing plants. A high percentage (40% to 50%) of the costs in CPPSs point to artificial lighting systems. Due to this, lighting strategies are essential to improve sustainability and profitability in closed plant production systems. However, no tools have been applied in the literature to contribute to energy savings in LED-type artificial radiation systems through the configuration of light recipes (wavelengths combination). For CPPS to be cost-effective and sustainable, a pre-evaluation of energy consumption for plant cultivation must consider. Artificial intelligence (AI) methods integrated into the prediction crucial variables such as each input-variable light color or specific wavelengths like red, green, blue, and white along with light intensity (quantity), frequency (pulsed light), and duty cycle. This paper focuses on the feature-selection stage, in which a regression model is trained to predict energy consumption in LED lights with specific light recipes in CPPSs. This stage is critical because it identifies the most representative features for training the model, and the other stages depend on it. These tools can enable further in-depth analysis of the energy savings that can be obtained with light recipes and pulsed and continuous operation light modes in artificial LED lighting systems.

Keywords: light wavelength; energy efficiency; features selection; machine learning



Citation: Montes Rivera, M.; Escalante-Garcia, N.; Dena-Aguilar, J.A.; Olvera-Gonzalez, E.; Vacas-Jacques, P. Feature Selection to Predict LED Light Energy Consumption with Specific Light Recipes in Closed Plant Production Systems. *Appl. Sci.* **2022**, *12*, 5901. <https://doi.org/10.3390/app12125901>

Academic Editors: Luis Hernández-Callejo, Sergio Nesmachnow and Sara Gallardo Saavedra

Received: 3 May 2022

Accepted: 26 May 2022

Published: 9 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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/>).

1. Introduction

1.1. LED Lights in Closed Plant Production Systems

Agriculture in 2050 will have to produce almost 50% more output to meet the demand for food supplies, presenting it with a crucial challenge in meeting the increase in demand [1]. Technological development and innovation can offer alternatives to ensure food security sustainably. The use of closed growth environments, such as greenhouses, plant factories, and vertical farms [2–6], represents a sustainable alternative for fresh food production. In closed plant production systems (CPPSs), several variables can be controlled and optimized, such as water, fertilizers, CO₂ injection, and temperature, as well as the quantity and quality of light thus ensuring minimum greenhouse gas emissions [3]. CPPSs allow growing of any plant variety, no matter the season of the year. Artificial lighting plays an essential role in CPPSs, as it promotes growth by providing optimal conditions for

plant development. LEDs are energy-efficient replacements that contribute to plant growth in agriculture. An outstanding advantage of LED lamps is their ability to operate with specific wavelengths (λ) that considerably reduce energy consumption. LEDs regularly generate continuous light. Likewise, they can radiate pulsed light (on/off in microseconds (μs)) with high power and low energy consumption at a specific frequency and duty cycle without upsetting the vegetative development of plants [4,7,8]. LED technology can produce different colors of light—that is, different qualities—called light recipes (different wavelength combinations). The wavelength combinations (red, blue, green, ultraviolet, and infrared) and the photosynthetic photon flux density (PPFD, given in $\mu\text{mol m}^{-2} \text{s}^{-1}$) are the components that constitute the light recipes. Light recipes impact crop growth from branching to flowering; optimize the biomass; and increase the antioxidant capacity levels of calcium, potassium, magnesium, chlorophyll, iron, vitamins A, B, and E, and other substances [7–9]. Crop quality and productivity rely upon the time and the light quantity supplied to the plants.

CPPSs can offer several advantages (improved management control of all variables involved—temperature, CO_2 , radiation—and increased productivity, growth, and yield) and generate an impact on humanity. Nevertheless, it is a model with a high demand for electricity for the artificial radiation systems needed to enhance the developing plants. Environmental control (refrigeration), the air required to remove the heat produced, and artificial lighting account for approximately 32%, 11%, and 57% of the total energy demand, respectively [10]. Furthermore, according to Avgoustaki and Xydis [11], the artificial lighting system accounts for 80% of the electrical demand, since the overall operability of the CPPS accounts for 40% of the total energy consumption.

Innovative approaches, such as fluid dynamics, evolutionary algorithms [12,13], the derivative integral model, and derivative model [14–16], control the resources in CPPSs. Artificial neural networks predict weather conditions and energy consumption [13–15,17]. Other techniques predict energy consumption performance for plant production [18,19]. Finally, other techniques focus on in the optimization of resources and reducing energy demand in CPPSs [20,21].

1.2. Machine-Learning Modeling

Physical modeling approaches are the most common approaches for predicting system behaviors, but they rely on descriptions of physics concepts. Thus, they tend to be complex, as the detail of the model increased. Therefore, as the principle of Occam's razor states, physical modeling must balance complexity with assumptions in order to produce simplified and representative models [22,23].

On the other hand, artificial intelligence (AI) researchers have proposed several techniques that allow automatic generation of the models and equations based on measurements arranged in datasets. Furthermore, machine learning (ML), a field of AI, applies deterministic and heuristic methods to produce models with less complexity established in the raw measurements [22].

During the last two decades, ML models have exhibited high effectivity, accuracy, and performance in several fields, including energy applications. Furthermore, ML results for modeling have motivated researchers to apply its models to accurately predict the behavior of physical phenomena [22–30].

The ML modeling process can involve several stages, depending on its application, but a general description would include collecting data, preprocessing data, building a model, training, and testing. Furthermore, all the stages must be continually tuned to improve the results; i.e., the stages can repeatedly change across the entire process if the model requires efficiency improvements, as represented in Figure 1 [22].

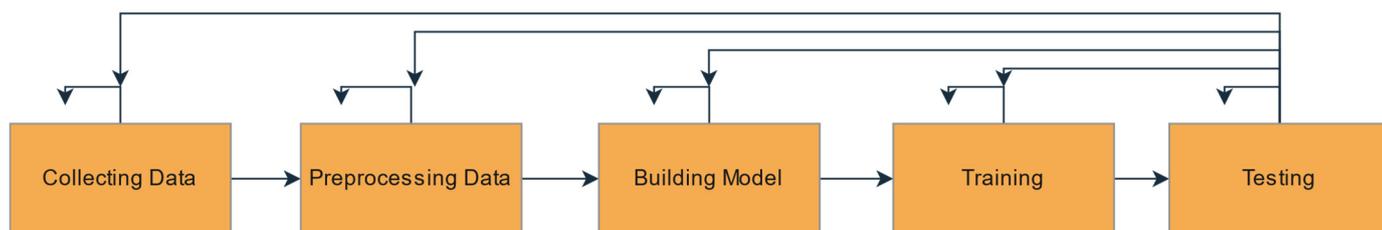


Figure 1. Modeling process with ML.

1.2.1. Collecting Data

ML modeling uses algorithms, statistics, and measurements structured in a dataset to identify the process behaviors and mimic them in a model [31]. The data generation stage depends on the processes contained in the chosen model. They may include electrical, mechanical, optical, thermal, psychic, or chemical variables [22,24,25,32,33]; derive from statistical analysis [26,27]; or be constructed with text, multimedia, or even real-time reports [32–35]. Nevertheless, the datasets can be associated with a specific time and/or frequency domain [36,37].

1.2.2. Preprocessing Data

After collecting and structuring the dataset, its variables need to be cleaned, processed, and filtered for the ML model. The processing stage includes several techniques, which can be human- or AI-designed, and they depend on the nature of the training data. For example, in natural language processing with text, preprocessing removes capitals [38]; in signal processing, wavelet transforms separate signals into their main components [39]; in image processing, convolution with the image filters extracts features [40]; in big data and data mining, dimensionality reduction is achieved [41].

The preprocessing data stages include normalization based on algorithms, such as MIN-MAX normalization, decimal scaling, and Z-scores; filtering redundant and inconsistent data; transformations such as linear, quadratic, polynomial, and histogram transformations; removing noisy data with techniques such as ensemble filtering, cross-validated filtering, and interactive partitioning; feature selection with exhaustive, heuristic, filter, and wrapper methods; and discretization to change from analog systems to digital ones [42].

Input features in ML modeling are representative when their information affects the output of the modeled system. Additionally, removing characteristics that are irrelevant or have low correlations from the results produces search spaces with lower complexity, boosting the capabilities of the training algorithm and improving the final model's efficiency [43,44].

One of the most used commonly techniques for removing redundant and inconsistent data in the second stage is feature Selection (FS). FS also makes it possible to reduce size, increase the efficiency and accuracy of predictive learning, and reduce the complexity of the final model [42]. The different FS approaches reported in the literature are constituted theoretically and apply methods such as filtering, wrapping, and embedding through techniques involving search algorithms, statistical criteria, and information, distance, dependency, and consistency measures [42].

1.2.3. Building Model

ML includes several models for predicting behavior that are supported by statistics and artificial intelligence. Different proposals have obtained different results depending on the ML model's application. The most common models are artificial neural networks, evolutionary algorithms, swarm intelligence algorithms, decision trees, naive Bayesian algorithms, logistic regression, fuzzy systems, gradient boosting machines, support vector machines, support vector regression, random forest algorithms, AdaBoost, simulated annealing, and hybrids of these models [22,24,26–28,31,34,44].

1.2.4. Training Model

Each ML model tunes its internal parameters with a training algorithm designed for the learning type. The most common learning types are supervised, unsupervised, reinforced, semi-supervised, transductive, self-trained, ensemble learning, boosting, and generative [31].

1.2.5. Testing Model

The metrics used to evaluate quality in the process of ML modeling depend on the nature of the model, which may be for classification or regression. In regression models, the metrics quantify the reliability of the model and the error between the model output and the real-world system. The most common regression metrics are the root mean square error (RMSE), mean error (ME), mean absolute error (MAE), mean average percentage error (MAPE), and the Nash coefficients E and R^2 [22,44].

1.3. Feature Selection

As mentioned in Section 1.2.2, feature selection is one of the most critical stages of ML modeling since it makes it possible to identify the best relation to the required complexity of the model and its quality at the preprocessing stage. In addition, feature selection makes it possible to find the more representative inputs in the real-world system and to eliminate no representative inputs or those that are redundant. ML models and training algorithms that consider only representative features improve their efficiency and reduce the time required for training [26,27,38,45]. A feature is an observable property in a system. Feature selection aims to select a specific subset of features that maximize the performance of the ML model.

The feature selection (FS) used here applied one of the most common techniques for removing irrelevant data, reducing dimensionality, increasing predictive accuracy and learning efficiency, and reducing the complexity of the final model [42]. Although there are different approaches for FS, all have theoretical support in their use of different methods, such as filtering, wrapping, and embedding, and involve techniques that use search algorithms, statistical criteria, and information, distance, dependence, and consistency measures [42]. The aim was to use linear and nonlinear methods to implement FS with a dataset acquired from an illumination radiation system.

This paper focuses on the feature selection stage in order to train a regression model to predict energy consumption in LED lights with specific light recipes in CPPSs. This stage is critical because it identifies the most representative features for training the model, and the other stages depend on it. These tools can enable further in-depth analysis of the energy savings that can be obtained with light recipes and pulsed and continuous light operation modes in artificial LED lighting systems.

2. Materials and Methods

2.1. Lighting System Features

The Artificial Lighting Laboratory (LIA) at Instituto Tecnológico de Pabellón de Arteaga in Aguascalientes, Mexico, developed the lighting system. An array of eight lamps formed the artificial lighting system. The wattage of each lamp was 25 watts. The ultra-bright LEDs emitted continuous and pulsed irradiation with different qualities (red, blue, green, and white). A programmed controller (a field-programmable gate array (FPGA)) allowed us to configure functions such as pulse frequency, duty cycle, intensity, wavelength, and on–off time.

2.2. Construction of Experiment

This study evaluated 10 light recipes from the literature, as can be seen in Figure 2. After that, the LED artificial radiation system was configured for continuous and pulsed emission to generate the first dataset (see table at the top of figure) to be analyzed. The intensities parameters were 50, 65, 80, 95, 110, 125, 140, 155, 170, and 185 $\mu\text{mol m}^{-2} \text{s}^{-1}$, as

determined by a quantum sensor, and the frequency was set to 100, 500, and 1000 Hz with 40%, 50%, 60%, 70%, 80%, and 90% duty cycles for different treatments.

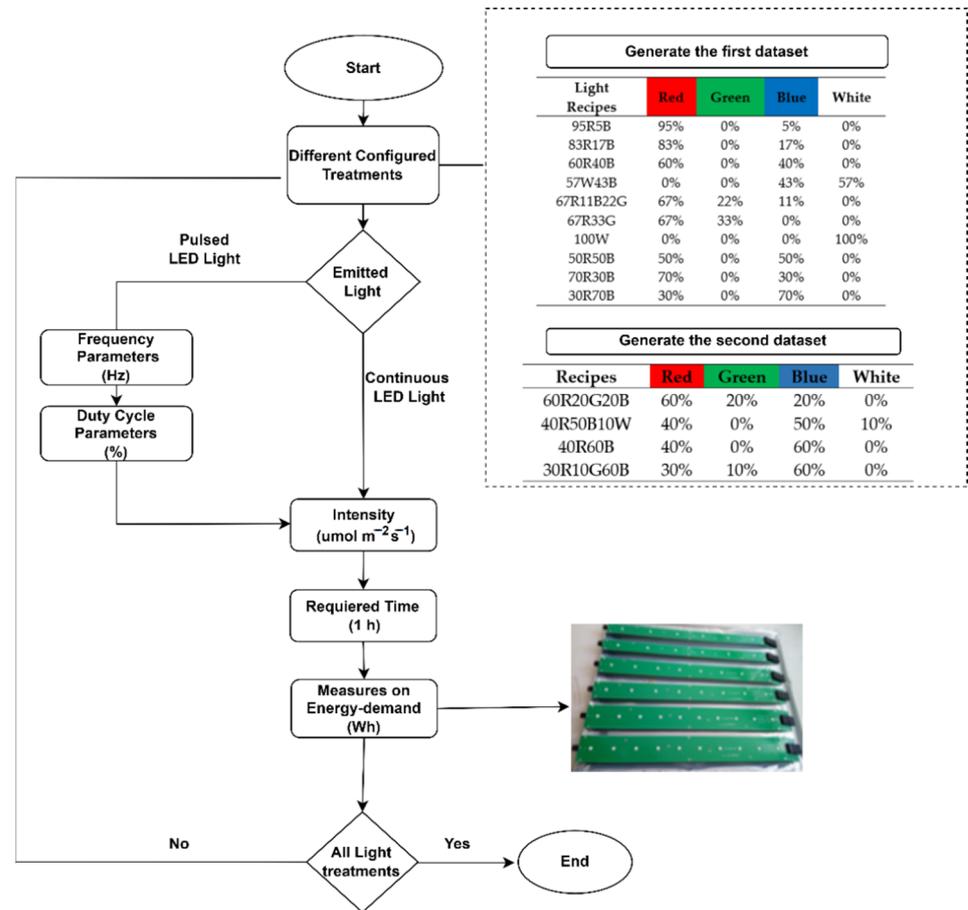


Figure 2. Flow diagram for collection of CPPS measurements for the dataset.

In the generation of the second dataset, four different light recipes were set at intensities of 60, 70, 85, 90, 90, 100, 120, 130, 150, 160, and 180 $\mu\text{mol m}^{-2} \text{s}^{-1}$, the frequency was set at 100, 500, and 1000 Hz, and duty cycles were randomly selected at 60%, 70%, and 80%, depending on the treatment (see table at the bottom of Figure 2).

The artificial illumination system included 14 light recipes (see tables in Figure 2) with all combinations of parameters. After 60 min of radiation, we registered the measurements for the energy demand with a hook-on AC ammeter (Peak Teach, Salerno, Italy) in watts \times hours (Wh). Then, the artificial radiation system was turned off for 15 min to cool down.

2.3. Min-Max Normalization

Normalization linearly transforms variables within specific ranges based on the minimum and maximum median absolute deviations of the variable values, avoiding changes to priorities in the variables because of the scale. Equation (1) represents the standard deviation required in the transformation as X_{std} , and Equation (2) indicates the variable scaling [46,47].

$$X_{std} = \frac{x - X_{min}}{X_{max} - X_{min}} \tag{1}$$

$$X_{scaled} = X_{std} \times (X_{max} - X_{min}) + X_{min} \tag{2}$$

where X_{scaled} is the new value transformed from the original value $x \in X$ and X_{max} and X_{min} are the maximum and minimum values, respectively.

2.4. Pearson Correlation

This association method’s primary goal is to identify two or more correlated variables [45].

The Pearson correlation coefficient measures the degree of correlation between two variables in a linear approach. Let X and Y be those variables, with measurements given by $\{x_1, x_2, x_3, \dots, x_n\}$ and $\{y_1, y_2, y_3, \dots, y_n\}$ and means \bar{x} and \bar{y} . Then, the Pearson coefficient is given by Equation (3) [42].

$$\rho(X, Y) = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\left[\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2 \right]^{\frac{1}{2}}} \tag{3}$$

A Pearson coefficient with the range $\rho = [-1, 1]$ represents the level of correlation when ρ is positive and correlation is direct, and the negative is the inverse [42].

When two variables are highly correlated, one can be redundant. The Pearson correlation works only for linear relations and results in incorrectly measured correlations for nonlinear systems. When classifying with binary outputs, it is possible to identify using Pearson coefficients how an attribute correlates with the target class [42].

Additionally, one can perform a correlation statistical significance test using the p_{value} coefficient, such as a test of the probability that the correlation coefficient ρ is a wrong hypothesis; for example, as a convention from the literature, if $p_{value} > 0.05$, it is unreliable. The alternatives for such a determination include statistical tests, such as the t_{value} , variance analysis (ANOVA), and 1_{tailed} or 2_{tailed} tests [48].

2.5. Variance Threshold

This method is used to identify features with variance. The features eliminated based on variance are those with zero value, near to zero value, or below a specific threshold [49–51].

The variability in a group given with $\{x_1, x_2, x_3, \dots, x_n\}$ is the standard error; in other words, it is the difference between the samples and the average value of the group \bar{x} , as in Equation (4) [52].

$$\sigma^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1} \tag{4}$$

2.6. Mutual Information Gain

Feature selection with mutual information gain enables the discrimination of features based on their interaction measurement, both for linear and nonlinear models [53]. Mutual information measures the uncertainty based on the entropy H of one variable while observing the other one. Let X be a random variable with values $\{x_1, x_2, x_3, \dots, x_n\}$; its entropy is given by Equation (5) [54].

$$H(X) = -\sum_{i=1}^n P(x_i) \log_2 [P(x_i)] \tag{5}$$

Let Y be an output variable with values $\{y_1, y_2, y_3, \dots, y_n\}$ and let X be a features array with values $\{x_1, x_2, x_3, \dots, x_n\}$; $H(X|Y)$ is then given by Equation (6) [54].

$$H(X | Y) = -\sum_{j=1}^n [P(x_i)] \sum_{i=1}^n P(x_i | y_j) \log_2 [P(x_i | y_j)] \tag{6}$$

The mutual information in Equation (7) measures the reduction in the uncertainty of X given Y [54,55].

$$MI(X|Y) = H(X) - H(X|Y) \tag{7}$$

2.7. Univariate Linear F-Regression Selection

This method uses a linear model to measure the degree of linear dependence between two random variables; in other words, it measures the significance of a feature in a linear model [56].

The F-regression equations use the null hypothesis H_0 , indicating that the data only intercept the model, and the alternative hypothesis H_1 , indicating the compatibility of the data with the model. The selection of the true hypothesis relies on the F_{score} given in Equation (8), the explained variance from Equation (9), and the unexplained variance from Equation (10) [56].

$$F = \frac{\text{explained variance}}{\text{unexplained variance}} \tag{8}$$

$$\text{explained variance} = \sum_{i=1}^K n_i \frac{(\bar{Y}_i - \bar{Y})^2}{(K - 1)} \tag{9}$$

$$\text{unexplained variance} = \sum_{i=1}^K \sum_{j=1}^{n_i} \frac{(Y_{ij} - \bar{Y}_{i.})^2}{(N - K)} \tag{10}$$

where Y_{ij} is the j th observation in the i out group in K , which is the number of out groups. N is the overall sample size and n_i is the number of observations.

Additionally, following Section 2.4, one can determine a p_{value} for the hypothesis conclusion, and, like with the Pearson correlation, if $p_{value} > 0.05$, the conclusion is unreliable [56].

2.8. Sequential Feature Selection

Sequential feature selection algorithms are a subset of wrapper algorithms that use greedy search algorithms. They evaluate a solution with certain features in a specific model and decide which feature to remove based on its quality. This technique can use a feedforward or backward approach; i.e., adding or removing features in the model. Figure 3 displays the searching schema for feedforward and backward sequential selection with three features [57,58].

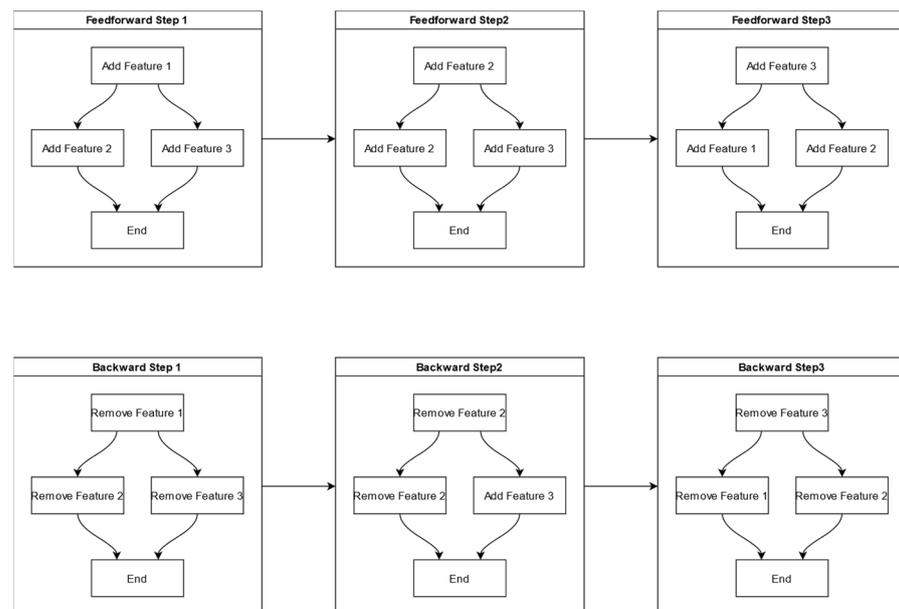


Figure 3. Flow diagram for feedforward and backward sequential feature selection.

For this study, backward sequential feature selection served to remove the worst variables in the energy consumption dataset for the LED lamps of a CPPS.

2.8.1. Linear Regression Model

The linear model structure includes $Y \in \mathbb{R}^{n \times 1}$, where $Y = (y_1, y_2, y_3, \dots, y_n)^T$ is the response variable; $X \in \mathbb{R}^{n \times p}$, where $X = (x_1, x_2, x_3, \dots, x_n)$ represents the design matrix; and $x_i = (x_{i,1}, x_{i,2}, x_{i,3}, \dots, x_{i,p})$ and $\beta \in \mathbb{R}^{p \times 1}$, where n is the number of observations and p is similar to the number of features. Then, the linear regression model is given by Equation (11) [59].

$$Y = \mu + \epsilon \quad (11)$$

where $\mu = \beta \times X$ and ϵ is the regression error.

Then, with a given predictor Y and the design matrix X , Equation (12) solves the β model parameters that reduce ϵ [59].

$$\beta = (X^T X)^{-1} X^T Y \quad (12)$$

2.8.2. Decision Tree Regression Model

Decision trees are hierarchical structures with nodes representing tests of the data with specific attributes and branches representing the test results. Decision tree models include IDS, C4.5, CART, and regression models. For example, the regression decision trees predict continuous random variables by finding the attributes that reduce the mean square error (MSE), obtained with Equation (13) [60].

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y}_i)^2 \quad (13)$$

where $Y = (y_1, y_2, y_3, \dots, y_n)$ is the raw data output variable and $\bar{Y} = (\bar{y}_1, \bar{y}_2, \bar{y}_3, \dots, y_n)$ represents the decision tree model output [60].

For this application, the regression variable used a decision tree with the energy consumption and the node attributes as the features for the energy consumption dataset.

3. Results

3.1. Energy Consumption Dataset

We registered the power consumption emitted by the artificial lighting system as a function of the light recipe, including parameters such as intensity; R, G, B, and W quality; frequency; and duty cycle. Table 1 represents the first dataset obtained through the process described in Section 2 (Figure 2). The evaluated energy consumption contained different ranges depending on the directly configured parameters. However, applying specific value ranges to the inputs affects the priority assigned to each one.

Endeavoring not to affect the input priorities, a new scaled dataset with min-max normalization was generated according to the equations defined in Section 2.3. Table 2 shows the data obtained after applying the equations corresponding to each input and output variable. The data represent the ranges from 0 to 1 after normalization. A value of 0 corresponds to the minimum value identified for that variable, while 1 is the maximum.

Table 1. First 15 measurements of the dataset generated.

Intensity (A) ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	Light Color Percentage (%)				Frequency (Hz)	Duty Cycle (%)	Energy Consumption (Wh)
	R	G	B	W			
50	45	0	5	0	0	0	23.5
50	41.5	0	8.5	0	0	0	23.4
50	30	0	20	0	0	0	23.9
50	0	0	21.5	28.5	0	0	25.1
50	33.5	11	5.5	0	0	0	24.4
50	33.5	16.5	0	0	0	0	23.4
50	0	0	0	50	0	0	24.5
50	25	0	25	0	0	0	23.9
50	35	0	15	0	0	0	33.5
50	15	0	35	0	0	0	24.1
50	45	0	5	0	100	40	20.7
50	41.5	0	8.5	0	100	40	20.6
50	30	0	20	0	100	40	20.9
50	0	0	21.5	28.5	100	40	22.2
50	33.5	11	5.5	0	100	40	21.1

Table 2. First 15 scaled dataset measurements.

Intensity (A) ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	R	G	B	W	Frequency (Hz)	Duty (%)	Energy Consumption (Wh)
0.000	0.256	0.000	0.039	0.000	0.000	0.000	0.085
0.000	0.236	0.000	0.066	0.000	0.000	0.000	0.082
0.000	0.171	0.000	0.154	0.000	0.000	0.000	0.097
0.000	0.000	0.000	0.166	0.154	0.000	0.000	0.132
0.000	0.191	0.180	0.042	0.000	0.000	0.000	0.111
0.000	0.191	0.270	0.000	0.000	0.000	0.000	0.082
0.000	0.000	0.000	0.000	0.270	0.000	0.000	0.114
0.000	0.142	0.000	0.193	0.000	0.000	0.000	0.097
0.000	0.199	0.000	0.116	0.000	0.000	0.000	0.378
0.000	0.085	0.000	0.270	0.000	0.000	0.000	0.103
0.000	0.256	0.000	0.039	0.000	0.100	0.444	0.003
0.000	0.236	0.000	0.066	0.000	0.100	0.444	0.000
0.000	0.171	0.000	0.154	0.000	0.100	0.444	0.009
0.000	0.000	0.000	0.166	0.154	0.100	0.444	0.047
0.000	0.191	0.180	0.042	0.000	0.100	0.444	0.015

3.2. Person Correlation Results

The next step was determining the Pearson correlation level with the coefficient ρ supported by the statistic test p_{value} using the equations in Section 2.4. The evidence for Pearson correlation with intensity, R, G, B, W, and frequency was sufficient as $p_{value} \leq 0.05$, but the duty correlation was unreliable because $p_{value} > 0.05$. Thus, duty cycle was the first variable eliminated (Table 3).

Table 3. ρ and p_{value} correlation with energy consumption per input variable.

Elimination Order	Input	ρ	p_{value}
7th	Intensity	0.865312	0
3rd	R	0.091069	5.64×10^{-12}
2nd	G	0.043198	0.001106
5th	B	0.372963	1.3×10^{-187}
6th	W	0.522086	0
4th	Frequency	0.110005	8.18×10^{-17}
1st	Duty cycle	0.014195	0.283926

Figure 4 shows a correlation heat map of the input variables and the energy consumption output to identify the strongest correlations graphically. There is a lower correlation where the graph color is darker.

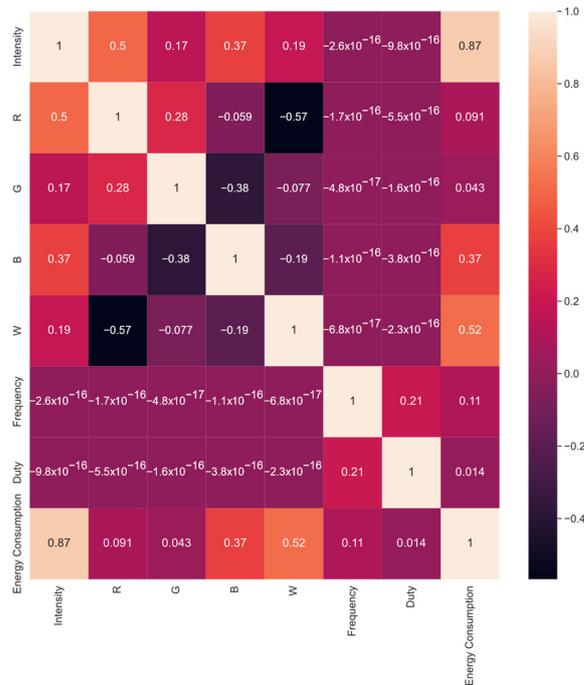


Figure 4. Pearson correlation heat map results.

3.3. Variance Threshold Results

The dataset energy consumption variables were dismissed against the variance threshold value as it gradually increased. The method for eliminating the variables with lower variance was described in Section 2.5. Figure 5 indicates the color associated with each variable in the variance threshold selection.

Table 4 displays the feature variance, the threshold value, and a bar plot showing the eliminated variable. Each threshold value was increased by 0.01 steps until a feature was eliminated from the energy consumption dataset.

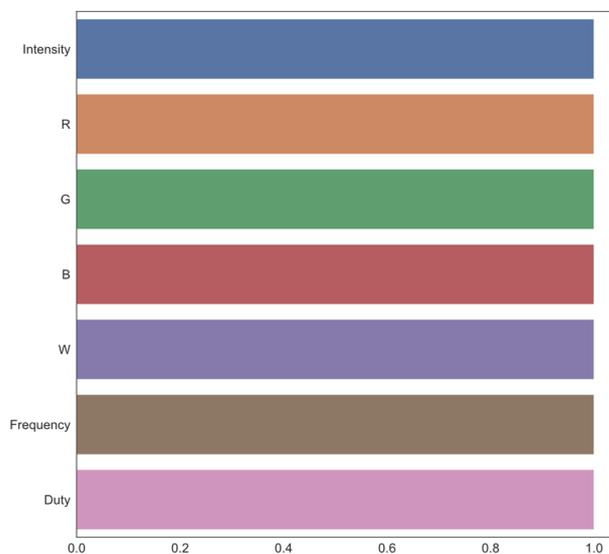


Figure 5. Colors per variable for variance threshold feature selection.

Table 4. Elimination order for features using variance threshold selection.

Elimination Order	Variable	Variance	Threshold	Image
1st	W	0.05079	0.051	
2nd	G	0.05490	0.055	
3rd	B	0.05546	0.056	
4th	Duty cycle	0.06012	0.061	

Table 4. Cont.

Elimination Order	Variable	Variance	Threshold	Image
5th	R	0.06479	0.065	
6th	Intensity	0.10185	0.110	
7th	Frequency	0.14260	N/A	N/A

3.4. Mutual Information Gain Results

Mutual information gain feature selection for the energy consumption dataset, set as the output variable, and all the other variables, assigned to the features array, was applied according to the description in Section 2.6. The results of this analysis are represented in Figure 6 and Table 5.

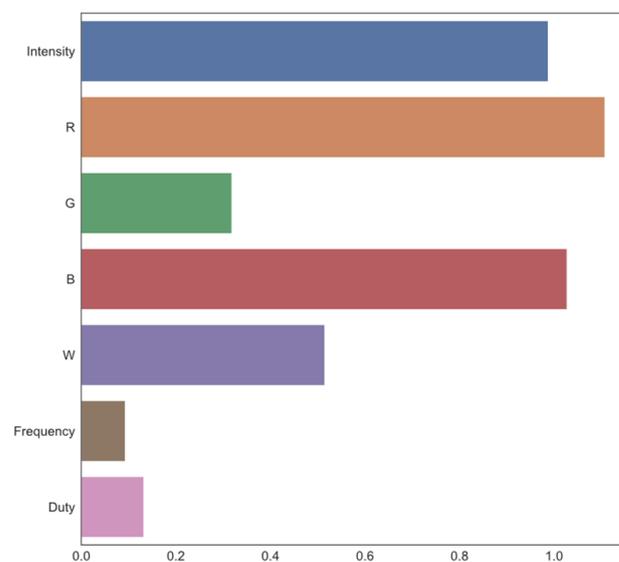


Figure 6. Mutual information evaluated versus the features in the dataset.

Table 5. Mutual information gain values for the energy consumption dataset.

Elimination Order	Input	MI(X Y)
5th	Intensity	0.987600
7th	R	1.107432
3rd	G	0.318185
6th	B	1.027326
4th	W	0.514607
1st	Frequency	0.092839
2nd	Duty	0.131858

3.5. Univariate Linear F-Regression Results

Once again, feature selection by F-regression in the generated dataset employed the energy consumption, such as the outcome variable in the features array (Section 2.7). Figure 7 displays the F_{score} value for hypothesis H_0 for each variable in the features array, and Table 6 shows the F_{score} and p_{value} calculated by H_0 for all input parameters.

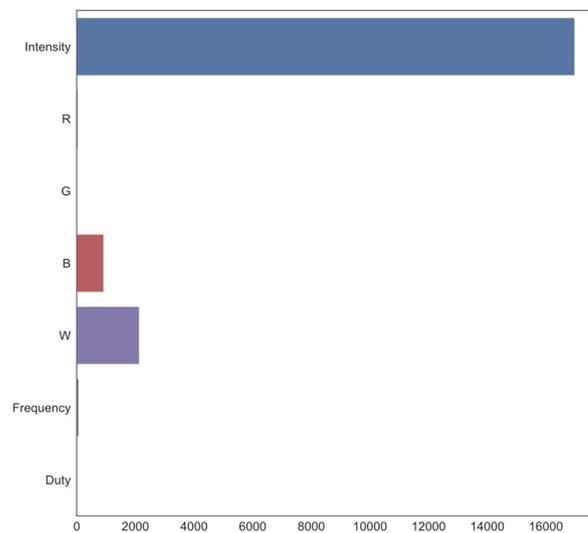


Figure 7. F_{score} comparison and the energy consumption dataset features.

Table 6. F_{score} and p_{value} feature selection with F-regression.

Elimination Order	Input	F_{score}	p_{value}
7th	Intensity	16,981.875086	0
3rd	R	47.651943	5.643556×10^{-12}
2nd	G	10.652646	1.105620×10^{-3}
5th	B	920.664903	$1.349950 \times 10^{-187}$
6th	W	2135.097576	0
4th	Frequency	69.796609	8.176545×10^{-17}
1st	Duty cycle	1.148417	2.839262×10^{-1}

The calculated p_{value} is indicated in Table 6. The Pearson correlations for intensity, R, G, B, W, and frequency showed reliable results since $p_{value} \leq 0.05$, but the duty cycle correlation was unreliable because $p_{value} > 0.05$; that is, the duty cycle was the first eliminated.

3.6. Sequential Feature Selection Results

We used backward sequential feature selection under a linear model and a decision tree regression (no linear model). However, implementation of the sequential feature selection

through the decision tree may have generated overfitting; thus, a 10-fold cross-validation allowed the recognition of the accepted characteristics by modifying the tree depth from 2 to 5.

3.6.1. Sequential Feature Selection with Linear Regression Model

Table 7 presents the results obtained for the linear regression model (Section 2.8.1) determining the feature elimination sequence, which used as attributes the admitted range from 1 to 6. The table follows the logic of Figure 5, showing each color in the elimination ranking with the variance threshold.

Table 7. Sequential feature deletion from the linear regression.

Elimination Order	Variable	Image
1st	Duty cycle	
2nd	W	
3rd	G	

Table 7. Cont.

Elimination Order	Variable	Image
4th	B	
5th	Frequency	
6th	R	
7th	Intensity	N/A

3.6.2. Sequential Feature Selection with Decision Tree Regression Model

The feature recognition through the decision tree regression model used the variance in a specific feature, which ranged from 1 to 6, allowing the elimination order for each input variable (Tables 8–11) to be obtained from the tree depth configuration. A color image of each feature, following the structure for variance threshold selection, is shown in Figure 5.

Table 8. Sequential feature selection by decision tree for depth = 2.

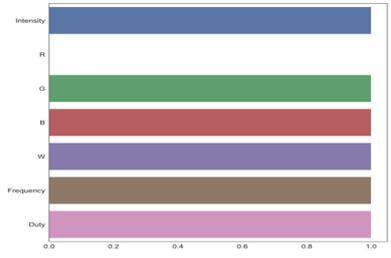
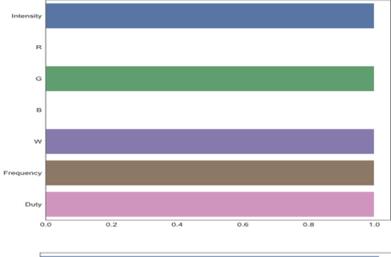
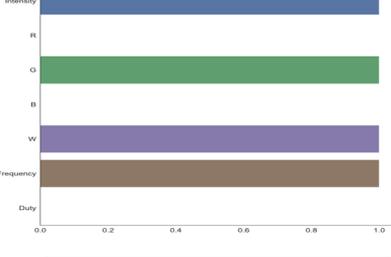
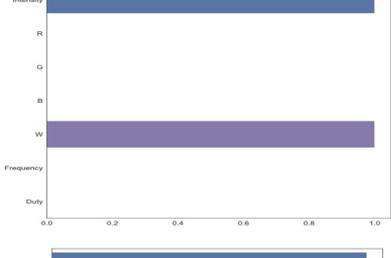
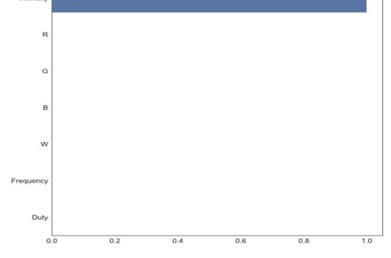
Elimination Order	Variable	Image
1st	R	
2nd	B	
3rd	Duty cycle	
4th	Frequency	
5th	G	
6th	W	
7th	Intensity	N/A

Table 9. Sequential feature selection by decision tree for depth = 3.

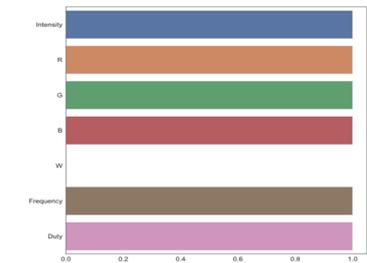
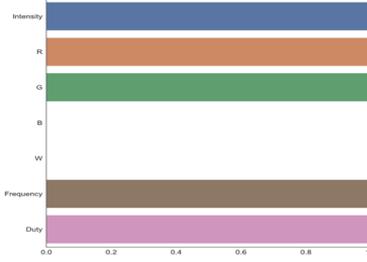
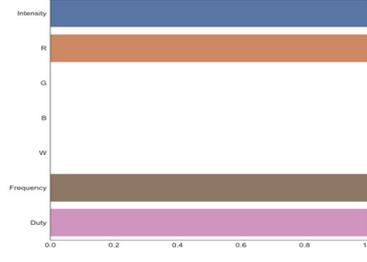
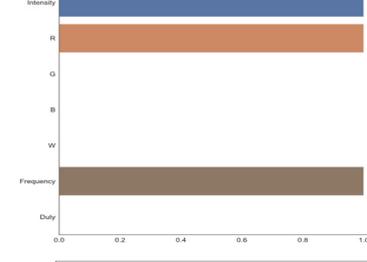
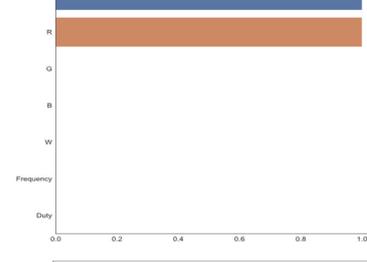
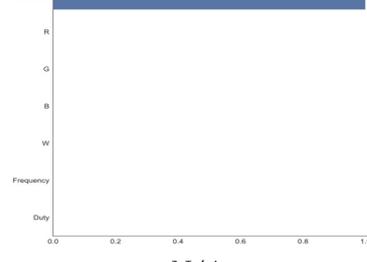
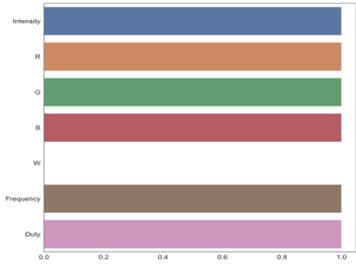
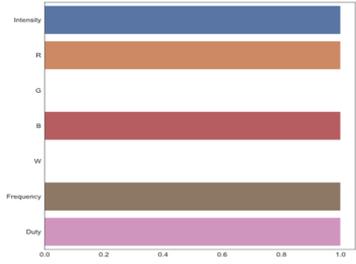
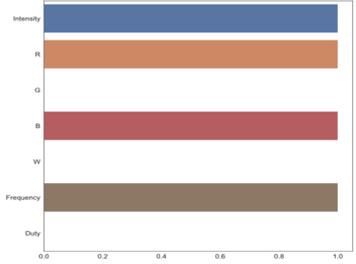
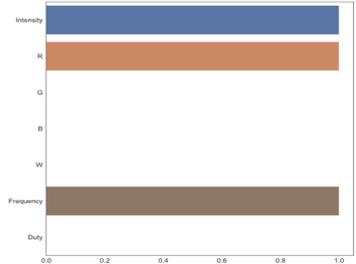
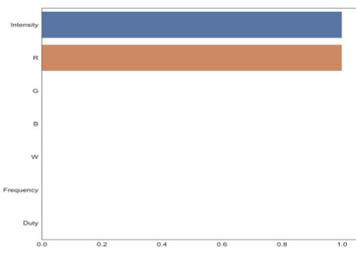
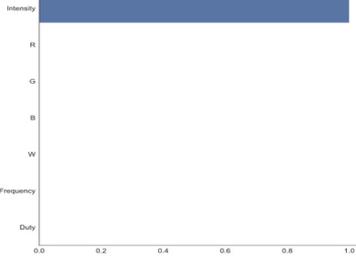
Elimination Order	Variable	Image
1st	W	
2nd	B	
3rd	G	
4th	Duty cycle	
5th	Frequency	
6th	R	
7th	Intensity	N/A

Table 10. Sequential feature selection by decision tree for depth = 4.

Elimination Order	Variable	Image
1st	W	
2nd	B	
3rd	Frequency	
4th	G	
5th	Duty cycle	
6th	R	
7th	Intensity	N/A

Table 11. Sequential feature selection by decision tree for depth = 5.

Elimination Order	Variable	Image
1st	W	
2nd	G	
3rd	Duty cycle	
4th	B	
5th	Frequency	
6th	R	
7th	Intensity	N/A

4. Discussion

The results obtained require division into linear and nonlinear model selection algorithms. The division generated makes it possible to analyze the results according to the model type and to identify the sequence of each feature. Tables 12 and 13 show the algorithms by group, the feature selection order, and the mean.

Table 12. Elimination order for the linear model group.

Feature	Pearson Correlation	Variance Threshold	Univariate Linear F-Regression	Sequential Backward Linear	Mean
Intensity	7	6	7	7	6.8
R	3	5	3	6	4
G	2	2	2	3	2.2
B	5	3	5	4	4.4
W	6	1	6	2	4.2
Frequency	4	7	4	5	4.8
Duty cycle	1	4	1	1	1.6

Table 13. Elimination order for the nonlinear model group.

Feature	Variance Threshold	Mutual Information Gain	Sequential Backward Deep Tree Values				Mean
			2	3	4	5	
Intensity	6	5	7	7	7	7	6.5
R	5	7	1	6	6	6	5.17
G	2	3	5	3	4	2	3.17
B	3	6	2	2	2	4	3.17
W	1	4	6	1	1	1	2.33
Frequency	7	1	4	5	3	5	4.17
Duty cycle	4	2	3	4	5	3	3.5

The averages calculated and reported in Tables 12 and 13 indicate two different behaviors depending on the model performance (linear or nonlinear). Figure 8 shows an alternative way to visualize the performance between linear and nonlinear models.

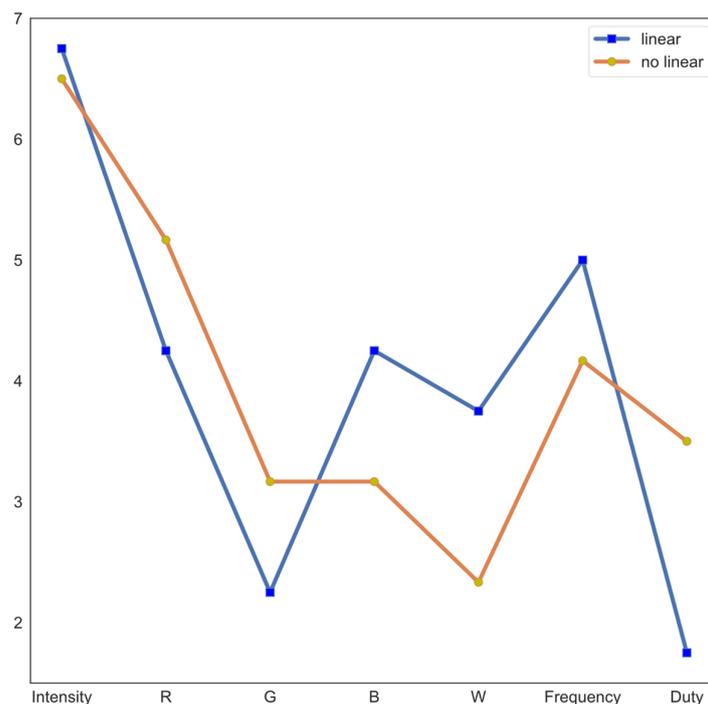


Figure 8. Feature elimination order distributions with algorithms from the two models used.

After dividing the models into linear and nonlinear groups, we validated the ordinal elimination variables on a scale from 1 to 7 and tested the distributions with the Kruskal–Wallis test (Table 14) [61]. The sequence elimination distribution for the proposed models is shown in Figures 9 and 10 (linear and nonlinear models, respectively).

Table 14. Values obtained with the Kruskal–Wallis test.

Group	F_{score}	p_{value}
Linear	16.27232	0.012364
Nonlinear	17.65278	0.007161

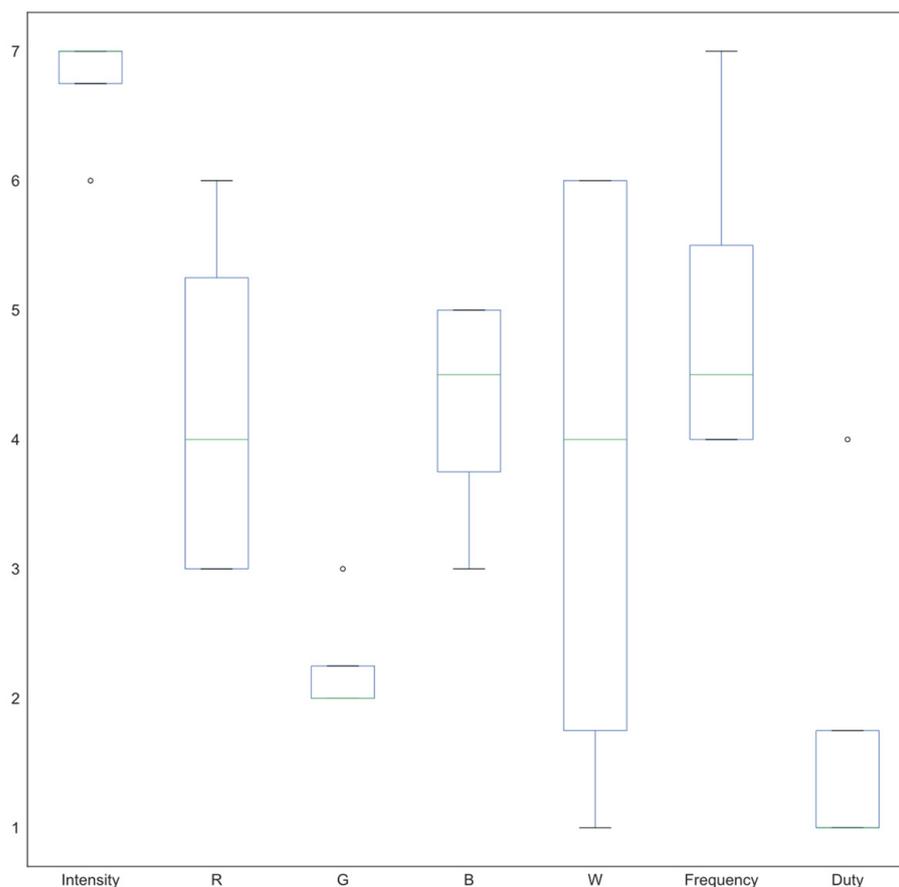


Figure 9. Order of elimination for features in linear models.

The linear model indicated that the essential characteristic was intensity, while the least significant was the duty cycle (Figure 9 and Table 14). If the appropriate sequence for any variable is required, the mean value can be found in Table 12. This means that the elimination order for the linear models was duty cycle, G, R, W, B, frequency, and intensity.

The nonlinear model found that the most crucial characteristic was intensity, while the least important was W (white color), with sufficient significance $p < 0.05$. If the correct sequence of the other variables is required, we can rely on the mean values for the feature distribution (Table 13). Overall, the elimination sequence was W, G, B, duty cycle, frequency, R, and intensity.

The elimination order for the duty cycle and R in the linear and nonlinear models suggests that they are nonlinear features, mainly because several linear algorithms selected them as the first variables to eliminate but nonlinear algorithms selected them as the most important ones.

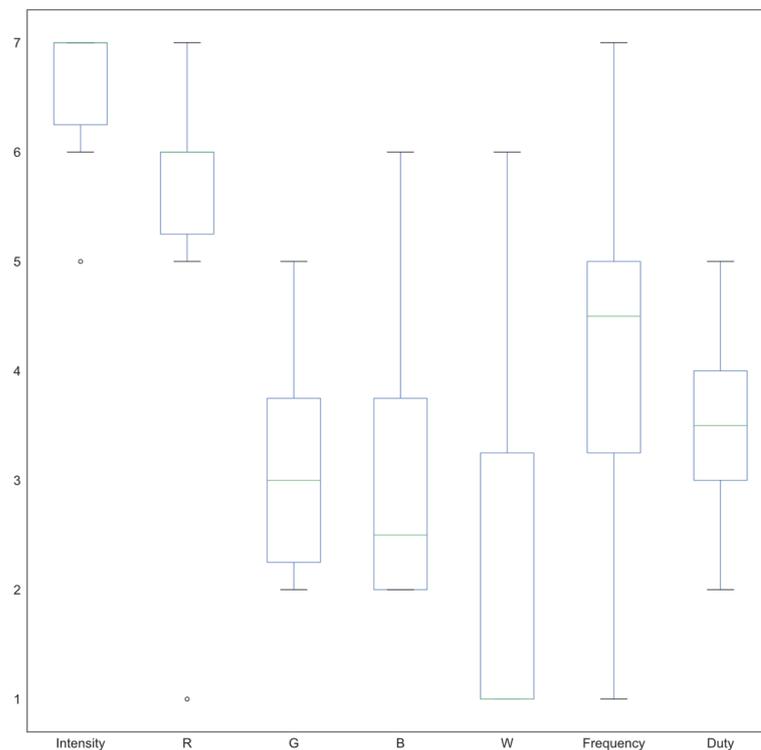


Figure 10. Order of elimination for features in nonlinear models.

5. Conclusions

In this study, we performed feature selection in order to prioritize inputs in the prediction of energy consumption in an artificial illumination system for a CPPS using linear and nonlinear regression models. A dataset was generated with electrical measurements for properties such as intensity, light wavelength (RGB and W), frequency, and duty cycle.

The algorithms used for the linear models to identify the elimination order of the features included the variance threshold, Pearson correlation, univariate linear F-regression, and sequential backward feature selection with linear regression.

On the other hand, for nonlinear models, the algorithms used were the variance threshold, mutual information gain, and sequential backward feature selection with tree decision regression, using a tree depth from 2–5. The Kruskal–Wallis test served to validate the elimination order distributions.

The best order for eliminating features with the linear model was duty cycle, light color, frequency, and intensity, with $p_{value} = 0.012364$. The best order with nonlinear models was white, green, blue, duty cycle, frequency, red, and intensity, with significance at $p_{value} = 0.007161$. The elimination order for the duty cycle and R in the linear and nonlinear models differed enormously because the linear algorithms considered them the most suitable elimination features, while nonlinear algorithms marked them as essential features. This discrepancy was because the duty cycle and R were nonlinear features. Thus, only nonlinear models could map them correctly. Moreover, this supports the hypothesis that the energy consumption in LED lamps for CPPSs has nonlinear behavior and that nonlinear models should be used to predict it.

This technique allows various deductions to be drawn from the analysis of the data obtained, including the estimation of the average energy consumption and its comparison with the quality of the crop, as well as the determination of the circumstances under which energy use is efficient. The selection of characteristics can be used as a reference for the agro-industrial community.

Author Contributions: All authors conceived the experiments; E.O.-G., N.E.-G. and J.A.D.-A. collected the data for the experiments; M.M.R., N.E.-G., E.O.-G. and P.V.-J. conducted the experiments, performed the statistical analysis, and generated the figures. All authors wrote and reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: We acknowledge the support of the Consejo Nacional de Ciencia y Tecnología (CONACYT) in Mexico for supporting this work through funds for the projects INFRA-2016-01, Project No. 270665, and CB-2016-01, Project No. 287818.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. FAO. *FAO Publications Catalogue*; FAO: Quebec City, QC, Canada, 2021.
2. Massa, G.D.; Kim, H.H.; Wheeler, R.M.; Mitchell, C.A. Plant Productivity in Response to LED Lighting. *HortScience* **2008**, *43*, 1951–1956. [[CrossRef](#)]
3. Kozai, T.; Fujiwara, K.; Runkle, E.S. *LED Lighting for Urban Agriculture*; Springer: Singapore, 2016; ISBN 9789811018480.
4. Domurath, N.; Schroeder, F.G.; Glatzel, S. Light Response Curves of Selected Plants under Different Light Conditions. *Acta Hort.* **2012**, *956*, 291–298. [[CrossRef](#)]
5. Eaves, J.; Eaves, S. Comparing the Profitability of a Greenhouse to a Vertical Farm in Quebec. *Can. J. Agric. Econ.* **2018**, *66*, 43–54. [[CrossRef](#)]
6. Benke, K.; Tomkins, B. Future Food-Production Systems: Vertical Farming and Controlled-Environment Agriculture. *Sustain. Sci. Pract. Policy* **2017**, *13*, 13–26. [[CrossRef](#)]
7. Mickens, M.A.; Skoog, E.J.; Reese, L.E.; Barnwell, P.L.; Spencer, L.E.; Massa, G.D.; Wheeler, R.M. A Strategic Approach for Investigating Light Recipes for ‘Outredgeous’ Red Romaine Lettuce Using White and Monochromatic LEDs. *Life Sci. Sp. Res.* **2018**, *19*, 53–62. [[CrossRef](#)] [[PubMed](#)]
8. Ahmed, H.A.; Yu-Xin, T.; Qi-Chang, Y. Optimal Control of Environmental Conditions Affecting Lettuce Plant Growth in a Controlled Environment with Artificial Lighting: A Review. *S. Afr. J. Bot.* **2020**, *130*, 75–89. [[CrossRef](#)]
9. Meng, Q.; Kelly, N.; Runkle, E.S. Substituting Green or Far-Red Radiation for Blue Radiation Induces Shade Avoidance and Promotes Growth in Lettuce and Kale. *Environ. Exp. Bot.* **2019**, *162*, 383–391. [[CrossRef](#)]
10. Graamans, L.; Baeza, E.; van den Dobbelsteen, A.; Tsafaras, I.; Stanghellini, C. Plant Factories versus Greenhouses: Comparison of Resource Use Efficiency. *Agric. Syst.* **2018**, *160*, 31–43. [[CrossRef](#)]
11. Avgoustaki, D.D.; Xydis, G. Energy Cost Reduction by Shifting Electricity Demand in Indoor Vertical Farms with Artificial Lighting. *Biosyst. Eng.* **2021**, *211*, 219–229. [[CrossRef](#)]
12. Hwang, P.W.; Chen, C.H.; Chang, Y.J. A Study on Energy Strategy of a Plant Factory Using Sustainable Energy Combined with Computational Fluid Dynamics Simulation: An Innovative Practice of Green Information Systems. In Proceedings of the Proceedings of Computing Conference, London, UK, 18–20 July 2017; IEEE: Piscataway, NJ, USA, 2018; pp. 517–522.
13. Sørensen, J.C.; Kjaer, K.H.; Ottosen, C.O.; Jørgensen, B.N. DynaGrow—Multi-Objective Optimization for Energy Cost-Efficient Control of Supplemental Light in Greenhouses. In Proceedings of the 8th International Joint Conference on Computational Intelligence (IJCCI 2016), Porto, Portugal, 9–11 November 2016; pp. 41–48.
14. Francik, S.; Kurpaska, S. The Use of Artificial Neural Networks for Forecasting of Air Temperature inside a Heated Foil Tunnel. *Sensors* **2020**, *20*, 652. [[CrossRef](#)]
15. Jung, D.H.; Kim, H.S.; Jhin, C.; Kim, H.J.; Park, S.H. Time-Serial Analysis of Deep Neural Network Models for Prediction of Climatic Conditions inside a Greenhouse. *Comput. Electron. Agric.* **2020**, *173*, 105402. [[CrossRef](#)]
16. Escamilla-García, A.; Soto-Zarazúa, G.M.; Toledano-Ayala, M.; Rivas-Araiza, E.; Gastélum-Barrios, A. Applications of Artificial Neural Networks in Greenhouse Technology and Overview for Smart Agriculture Development. *Appl. Sci.* **2020**, *10*, 3835. [[CrossRef](#)]
17. Singh, V.K.; Tiwari, K.N. Prediction of Greenhouse Micro-Climature Using Artificial Neural Network. *Appl. Ecol. Environ. Res.* **2017**, *15*, 767–778. [[CrossRef](#)]
18. Gros, S.; Zanon, M.; Quirynen, R.; Bemporad, A.; Diehl, M. From Linear to Nonlinear MPC: Bridging the Gap via the Real-Time Iteration. *Int. J. Control* **2020**, *93*, 62–80. [[CrossRef](#)]
19. Ouammi, A.; Achour, Y.; Zejli, D.; Dagdougui, H. Supervisory Model Predictive Control for Optimal Energy Management of Networked Smart Greenhouses Integrated Microgrid. *IEEE Trans. Autom. Sci. Eng.* **2020**, *17*, 117–128. [[CrossRef](#)]
20. Xu, H.; Zhai, Z.; Wang, K.; Ren, S.; Wang, H. Multiobjective Distributed Model Predictive Control Method for Facility Environment Control Based on Cooperative Game Theory. *Turk. J. Electr. Eng. Comput. Sci.* **2017**, *25*, 4160–4171. [[CrossRef](#)]
21. Lin, D.; Zhang, L.; Xia, X. Hierarchical Model Predictive Control of Venlo-Type Greenhouse Climate for Improving Energy Efficiency and Reducing Operating Cost. *J. Clean. Prod.* **2020**, *264*, 121513. [[CrossRef](#)]

22. Mosavi, A.; Ozturk, P.; Chau, K.W. Flood Prediction Using Machine Learning Models: Literature Review. *Water* **2018**, *10*, 1536. [[CrossRef](#)]
23. Hosseinzadeh, A.; Zhou, J.L.; Altaee, A.; Li, D. Machine Learning Modeling and Analysis of Biohydrogen Production from Wastewater by Dark Fermentation Process. *Bioresour. Technol.* **2022**, *343*, 126111. [[CrossRef](#)]
24. Alizamir, M.; Kisi, O.; Ahmed, A.N.; Mert, C.; Fai, C.M.; Kim, S.; Kim, N.W.; El-Shafie, A. Advanced Machine Learning Model for Better Prediction Accuracy of Soil Temperature at Different Depths. *PLoS ONE* **2020**, *15*, e0231055. [[CrossRef](#)]
25. Nemati, S.; Holder, A.; Razmi, F.; Stanley, M.D.; Clifford, G.D.; Buchman, T.G. An Interpretable Machine Learning Model for Accurate Prediction of Sepsis in the ICU. *Crit. Care Med.* **2018**, *46*, 547. [[CrossRef](#)] [[PubMed](#)]
26. Sneha, N.; Gangil, T. Analysis of Diabetes Mellitus for Early Prediction Using Optimal Features Selection. *J. Big Data* **2019**, *6*, 13. [[CrossRef](#)]
27. Haq, A.U.; Li, J.; Memon, M.H.; Hunain Memon, M.; Khan, J.; Marium, S.M. Heart Disease Prediction System Using Model of Machine Learning and Sequential Backward Selection Algorithm for Features Selection. In Proceedings of the 2019 IEEE 5th International Conference for Convergence in Technology, Bombay, India, 29–31 March 2019. [[CrossRef](#)]
28. Fan, X.; Wang, X.; Zhang, X.; Yu, P.A. Machine Learning Based Water Pipe Failure Prediction: The Effects of Engineering, Geology, Climate and Socio-Economic Factors. *Reliab. Eng. Syst. Saf.* **2022**, *219*, 108185. [[CrossRef](#)]
29. Ahmed, H.W.; Alamire, J.H. A Review of Machine Learning Models in the Air Quality Research. *Int. J. Adv. Res. Comput. Eng. Technol.* **2020**, *9*, 30–36.
30. Zoabi, Y.; Deri-Rozov, S.; Shomron, N. Machine Learning-Based Prediction of COVID-19 Diagnosis Based on Symptoms. *npj Digit. Med.* **2021**, *4*, 3. [[CrossRef](#)] [[PubMed](#)]
31. Mahesh, B. Machine Learning Algorithms—A Review. *Int. J. Sci. Res.* **2020**, *9*, 381–386. [[CrossRef](#)]
32. Ahmad, A.; Khan, M.; Paul, A.; Din, S.; Rathore, M.M.; Jeon, G.; Choi, G.S. Toward Modeling and Optimization of Features Selection in Big Data Based Social Internet of Things. *Future Gener. Comput. Syst.* **2018**, *82*, 715–726. [[CrossRef](#)]
33. Khan, M.A.; Kadry, S.; Alhaisoni, M.; Nam, Y.; Zhang, Y.; Rajinikanth, V.; Sarfraz, M.S. Computer-Aided Gastrointestinal Diseases Analysis from Wireless Capsule Endoscopy: A Framework of Best Features Selection. *IEEE Access* **2020**, *8*, 132850–132859. [[CrossRef](#)]
34. Genova, K.; Cole, F.; Maschinot, A.; Sarna, A.; Vlastic, D.; Freeman, W.T. Unsupervised Training for 3D Morphable Model Regression. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 8377–8386.
35. Gehrig, D.; Gehrig, M.; Hidalgo-Carrio, J.; Scaramuzza, D. Video to Events: Recycling Video Datasets for Event Cameras. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 3586–3595.
36. Simao, M.; Mendes, N.; Gibaru, O.; Neto, P. A Review on Electromyography Decoding and Pattern Recognition for Human-Machine Interaction. *IEEE Access* **2019**, *7*, 39564–39582. [[CrossRef](#)]
37. Combes, P.P.; Gobillon, L.; Zylberberg, Y. Urban Economics in a Historical Perspective: Recovering Data with Machine Learning. *Reg. Sci. Urban Econ.* **2021**, *94*, 103711. [[CrossRef](#)]
38. Uysal, A.K.; Gunal, S. The Impact of Preprocessing on Text Classification. *Inf. Process. Manag.* **2014**, *50*, 104–112. [[CrossRef](#)]
39. Zhang, X.; Zhao, Z.; Wang, Z.; Wang, X. Fault Detection and Identification Method for Quadcopter Based on Airframe Vibration Signals. *Sensors* **2021**, *21*, 581. [[CrossRef](#)] [[PubMed](#)]
40. Choras, R.S. A Survey on Methods of Image Processing and Recognition for Personal Identification. In *Machine Learning and Biometrics*; IntechOpen: Vienna, Austria, 2018. [[CrossRef](#)]
41. Mohammed, B.; Hasan, S.; Mohsin Abdulazeez, A. A Review of Principal Component Analysis Algorithm for Dimensionality Reduction. *J. Soft Comput. Data Min.* **2021**, *2*, 20–30. [[CrossRef](#)]
42. García, S.; Luengo, J.; Herrera, F. *Data Preprocessing in Data Mining*; Springer: Cham, Switzerland, 2015; Volume 72, ISBN 9783319102467.
43. Arslan, S.; Ozturk, C. Feature Selection for Classification with Artificial Bee Colony Programming. In *Swarm Intelligence-Recent Advances, New Perspectives and Applications*; IntechOpen: Vienna, Austria, 2019. [[CrossRef](#)]
44. Olvera-Gonzalez, E.; Rivera, M.M.; Escalante-Garcia, N.; Flores-Gallegos, E. Modeling Energy LED Light Consumption Based on an Artificial Intelligent Method Applied to Closed Plant Production System. *Appl. Sci.* **2021**, *11*, 2735. [[CrossRef](#)]
45. Jia, K.; Yang, L.; Liang, S.; Xiao, Z.; Zhao, X.; Yao, Y.; Zhang, X.; Jiang, B.; Liu, D. Long-Term Global Land Surface Satellite (GLASS) Fractional Vegetation Cover Product Derived From MODIS and AVHRR Data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *12*, 508–518. [[CrossRef](#)]
46. Kappal, S. Data Normalization Using Median Median Absolute Deviation MMAD Based Z-Score for Robust Predictions vs. Min—Max Normalization. *Lond. J. Res. Sci. Nat. Form.* **2019**, *19*, 39–44.
47. Saranya, C.; Manikandan, G. A Study on Normalization Techniques for Privacy Preserving Data Mining. *Int. J. Eng. Technol.* **2013**, *5*, 2701–2704.
48. Curran-Everett, D. Explorations in Statistics: Hypothesis Tests and P Values. *Am. J. Physiol. Adv. Physiol. Educ.* **2009**, *33*, 81–86. [[CrossRef](#)]

49. Kumar, S.C.; Ramasree, R.J. Dimensionality Reduction in Automated Evaluation of Descriptive Answers through Zero Variance, near Zero Variance and Non Frequent Words Techniques—a Comparison. In Proceedings of the 2015 IEEE 9th International Conference on Intelligent Systems and Control (ISCO), Coimbatore, India, 9–10 January 2015. [[CrossRef](#)]
50. Roberts, A.G.K.; Catchpoole, D.R.; Kennedy, P.J. Variance-Based Feature Selection for Classification of Cancer Subtypes Using Gene Expression Data. In Proceedings of the International Joint Conference on Neural Networks, Rio de Janeiro, Brazil, 8–13 July 2018. [[CrossRef](#)]
51. Siti Ambarwati, Y.; Uyun, S. Feature Selection on Magelang Duck Egg Candling Image Using Variance Threshold Method. In Proceedings of the 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems, Yogyakarta, Indonesia, 10–11 December 2020; pp. 694–699. [[CrossRef](#)]
52. Bewick, V.; Cheek, L.; Ball, J. Statistics Review 9: One-Way Analysis of Variance. *Crit. Care* **2004**, *8*, 130–136. [[CrossRef](#)]
53. Chehreh Chelgani, S.; Shahbazi, B.; Hadavandi, E. Support Vector Regression Modeling of Coal Flotation Based on Variable Importance Measurements by Mutual Information Method. *Measurement* **2018**, *114*, 102–108. [[CrossRef](#)]
54. Mamun, M.M.R.K.; Alouani, A.T. Cuffless Blood Pressure Measurement Using Linear and Nonlinear Optimized Feature Selection. *Diagnostics* **2022**, *12*, 408. [[CrossRef](#)] [[PubMed](#)]
55. Xiong, H.; Fan, C.; Chen, H.; Yang, Y.; ANTWI, C.O.; Fan, X. A Novel Approach to Air Passenger Index Prediction: Based on Mutual Information Principle and Support Vector Regression Blended Model. *SAGE Open* **2022**, *12*. [[CrossRef](#)]
56. Toğaçar, M.; Ergen, B.; Cömert, Z. Classification of Flower Species by Using Features Extracted from the Intersection of Feature Selection Methods in Convolutional Neural Network Models. *Measurement* **2020**, *158*, 107703. [[CrossRef](#)]
57. Khaire, U.M.; Dhanalakshmi, R. Stability of Feature Selection Algorithm: A Review. *J. King Saud Univ. Comput. Inf. Sci.* **2019**, *34*, 1060–1073. [[CrossRef](#)]
58. Wang, M.; Lu, Y.; Qin, J. A Dynamic MLP-Based DDoS Attack Detection Method Using Feature Selection and Feedback. *Comput. Secur.* **2020**, *88*, 101645. [[CrossRef](#)]
59. Zhang, D.; Khalili, A.; Asgharian, M. Post-Model-Selection Inference in Linear Regression Models: An Integrated Review. *Stat. Surv.* **2022**, *16*, 86–136. [[CrossRef](#)]
60. Darwin, D.; Christian, D.; Chandra, W.; Nababan, M. Comparison of Decision Tree and Linear Regression Algorithms in the Case of Spread Prediction of COVID-19 in Indonesia. *J. Comput. Netw. Archit. High Perform. Comput.* **2022**, *4*, 1–12. [[CrossRef](#)]
61. Johnson, R.W. Alternate Forms of the One-Way ANOVA F and Kruskal-Wallis Test Statistics. *J. Stat. Data Sci. Educ.* **2022**, *30*, 82–85. [[CrossRef](#)]