Scientific Papers. Series B, Horticulture. Vol. LXVIII, No. 2, 2024
Print ISSN 2285-5653, CD-ROM ISSN 2285-5661, Online ISSN 2286-1580, ISSN-L 2285-5653

# A MULTIPLE LINEAR REGRESSION MODEL TO ESTIMATE THE PLANT COVERAGE OF A GREEN WALL SYSTEM

**Ciprian CHIRUŢĂ[1], Iulian STOLERIU[2], Marius CĂLIN[1], Emilian BULGARIU[1], Mirela COJOCARIU[1]**

[1]University of Life Sciences, 3 Mihail Sadoveanu Alley, Iaşi, Romania
[2]Faculty of Mathematics, University "Al. I. Cuza", 11 Carol I Blvd, Iaşi, Romania

Corresponding author email: mirela.cojocariu@iuls.ro

***Abstract***

*Recently, green walls have come to the attention of researchers from several fields, as a result of the concerted effort to find viable solutions to stop/mitigate urban pollution that produces numerous negative effects. The viability of a green wall consists in the use of plants with a high aesthetic appearance, which are resistant to environmental conditions and ensure a quick and compact coverage. The present paper proposes a multiple linear regression model for the plant coverage of a vertical system in which the soil moisture and temperature are explanatory variables. The questions that we address and answer in this paper are related to the dependence of the plant coverage on the orientation of the wall and the influence of the plant coverage on the temperature inside a green wall system.*

***Key words****: green walls, multiple linear regression model, plant coverage.*

## INTRODUCTION

The use of statistical techniques in landscape research helps researchers to establish links between the parameters used, between the chosen variants, to demonstrate the strength of the connections and to characterize their evolution in a unitary and correct way within the experiment. The mere simple description of the execution steps, the observation of the phenomena and their narration within research do not manage to present in a scientific way all the particularities that appear in the development of the living systems.

To highlight the relationship that exists between two or more data sets, mathematical methods of analysis such as correlation analysis or regression analysis are needed (Wackerly et al., 2007). Any researcher wants to use the results obtained in a project to determine whether or not there are connections between the measured data sets, as well as the strength of the determined connections. The dependence or independence of variables is an important goal in any research and it is necessary to resort to mathematical modeling. A series of studies already published, with applicability in the horticultural field, highlight such connections. Thus, for a good characterization of the chemical composition, pH, ascorbic acid and total phenolics for three species of wild berries according to precipitation, air temperature, atmospheric pressure and wind speed, Tripon and colleagues used multiregression analysis and simple correlation. The conclusion of this research was that only precipitation and air temperature influence wild berries dry matter content (Tripon, 2022).

Another study highlights the variation scanning the binden leaf area and cumulative scan leaf area in relation to the position of the leaf on the shoot and cumulative shoots length was evaluated by regression analysis by Rosu and collaborators. Leaves deviating from the theoretically determined model were determined (Rosu, 2022). Also, a study carried out on tomatoes reveals that productivity and yield in tomatoes can be modeled mathematically using the interaction between humidity and water deficit for three levels of fertilizer. Thus, it was found that the optimal irrigation rate and 73% manure represent a good combination (Stoyanova, 2019). Another research in this sense was carried out in the Jidvei viticultural center, Tarnave region, where the Feteasca regala variety was studied. Using linear regression, it was established that the relationships between global solar radiation and parameters: sugars and global acidity are very strong (Ropan, 2023). Also, the relationships between global solar radiation and alcohol concentration, total acidity and reduced dry extract have strong relationships with significant statistical correlations.

Different species of plants develop differently in similar conditions; the same species change their way of evolution if they are placed in new conditions of stress. All these changes that occur are determined correctly and quickly by statistical techniques such as correlations and regressions.

In our project, the plants were removed from the traditional way of planting (in soil) and were included in a planting system on a vertical wall made up of 4 facades, each facing a cardinal point. This way of alignment was chosen in order to observe the changes that take place in the evolution of plants, in the way of development on such a system. The aesthetic aspect of green walls is extremely important. and the degree of plant coverage of the wall in a certain period of time contributes significantly to their visual quality.

In urban areas, it is hard to find available land for new parks and other traditional green areas. Thus, vegetation has been extended to other surfaces, such as roofs and facades of buildings (Ghazalli et al., 2019). The use of roofs and facades of buildings for the expansion of green areas could be a remedy for the greening of the habitat, as well as for a high aesthetic aspect of the cities we live in (Pérez et al., 2020; Francis and Lorimer, 2011). Green wall systems are sustainable solutions (Palermo and Turco, 2020; Sheweka and Nourhan, 2012) and could be a component of modern urban design (Perini et al., 2011) with many benefits for city residents (Fan et al., 2011; Perez-Urrestarazu et al., 2015).

During the course of the experiment, the influence of environmental factors and the substrate used on the behavior of the plants used in the vertical system was observed. The monitoring was done with the help of two different devices, every three days, in the same time interval. The research carried out in the experiment is complex, and many of the results have already been published. The behavior of perennial flowering species such as *Heuchera x hybrida* "Fire Alarm", *Festuca glauca*, *Sedum spurium* "Tricolor", *Carex testacea, Polystichum aculeatum* (Cojocariu et al., 2022a) and *Cineraria maritima* was monitored on the experimental structure, as well as annuals such as *Plectranthus forsteri*, *Coleus blumei* (Cojocariu et al., 2022b), *Begonia sempervirens*, and other.

In another work we compared the percentage of plant coverage (PCP) determined by a multilinear regression (MLR) and by a modeling with artificial neural networks (ANN) using confidence intervals. The model obtained by ANN is more accurate in predicting the phenomenon than the model obtained by multilinear regression (Chiruţă et al., 2023).

In this article, in order to be able to answer a series of questions, we had at our disposal the measurements made on the organic mixture used in the system (humidity and temperature), the internal temperature, the system temperature, the ambient temperature (from the Iasi National Meteorological Center) and the measured percentage of plant coverage. The tests we made with linear regression gave interesting results that we have centralized in this material.

Starting from our dataset, our aim was to build the best possible multiple linear regression model, for each cardinally oriented face of the green wall system. Based on these models, we want to explain the variation of the plant cover percentage (PCP) determined by the soil moisture, the soil temperature for each facade of the system and the time in the year these variables were measured.

Because the pH and the moisture content of the organic mixture were highly correlated ($r = -0.9233, p < 0.001$), they were not considered in any of the models. Nutrient substrate moisture and temperature had the greatest influence in predicting percent plant coverage for each system facade. Other important variables for the evolution of the plant coverage percentage were the time of year in which each observation was made and the cardinal direction of each facade of the system were added in the mathematical modelling.

In this article, we try to answer the following questions: Q1: Does the plant coverage percentage depend on orientation of the facade of the GWS [North (N), East (E), South (S), West (W)]? Q2: Does the plant coverage percentage influence the inner temperature of the GWS? Q3: Can one build good regression models of PCP solely on soil Humidity and soil Temperature as explanatory variables? Q4: Will

there be an improvement in the regression models if we also take into account the time of the year when these variables were measured?
In order to determine the influence of environmental factors and the substrate used on the behaviour of plants, the temperatures recorded in the area of the city of Iasi (Temp Iasi) were taken into account - data received from the National Meteorology Center Iasi.

Table 1. Temperatures in the experiment area between 2020-2021

| | Temp Iasi 2020 | Temp Iasi 2021 |
|---|---|---|
| **Mean** | 12.16 | 10.33 |
| **Minimum** | -4.2 | -11.1 |
| **Maximum** | 27.6 | 27.3 |
| **CI mean** | 12.16 ± 0.87 | 10.33 ± 0.91 |

## MATERIALS AND METHODS

### Experimental setup

Our experiment took place in open air, in the didactic field of the Floriculture department of the Faculty of Horticulture within IULS - Iasi University of Life Sciences (decimal GPS lat. N 47.1941, long. E 27.5555). It observed the behavior of some flower species planted vertically, in local climatic conditions, with the aim of including them in an assortment of ornamental species that can be successfully used in the decoration of green facades. Resistant flower species with minimal maintenance needs and high decorative potential were used such as *H. x hybrida* "Fire Alarm", *S. spurium* "Tricolor", *C. maritima, P. aculeatum, P. forsteri, C. blumei*. The flower species can also decorate by their habitus, or by the shape and color of their leaves (Draghia and Chelariu, 2011). This category also includes ornamental grasses. In Romania, their culture is slowly starting to make its presence felt (Chelariu, 2018), especially for their low maintenance and the high degree of resistance to environmental factors they exhibit. For these reasons, species from the category of ornamental grasses such as: *F. glauca* and *C. testacea* were tested on the vertical structure.
The first set of plants, consisting of the species *H. x hybrida* "Fire Alarm", *F. glauca, S. spurium* "Tricolor", *C. Testacea* and *P. aculeatum* were placed on the vertical structure at the end of autumn 2019. Due to the low rate of survival in the vertical system, *C. testacea* was replaced, between June and November 2020, with the annual species *Begonia semperflorens* ("Big"). Later, the assortment was enriched with new plants, so that in 2021, the following species were on the facades of the experimental module: *B. semperflorens, H. x hybrida* "Fire Alarm", *C. maritima, P. fosteri, C. blumei* and *F. glauca* (Chiruță et al., 2023).
The experimental structure was built out of heat-insulating panels, especially for this study. The facades of the structure are of equal size; each face being composed of four overlapping landings. The layout of the structure on the ground was made so that each facade is oriented towards a cardinal point.
The flower species were planted on each landing in equal and uniform numbers. Plots had identical organic matter within the experimental scheme. During the study, catch/attachment percentage, degree of cover, biometric aspects and visual quality of the mentioned species were monitored and comparisons were made between the cardinal orientations.
Data on the evolution of plant characteristics were collected every 3 days at mid-day. The characteristics of organic matter (moisture, temperature, pH) were also measured with the RZ89 4 in 1 3.5-9 ~ 9 pH Meter Digital Magnetic Soil Health Analyzer Machine Soil Moisture Ammonitor Hygrometer Gardening Plant Tester, as well as the internal and external temperatures of the green wall system, for which the Somogyi Elektronic Home HC 12 device with a resolution of 0.1°C was used.

### Multiple linear regression (MLR)

Regression analysis aims to determine the relationship between two (or more) variables of interest, in order to obtain information about one of them from the values of the other(s). The regression is called to be linear when the response variable depends linearly on the parameters. The general equation for a multiple linear regression model with a dependent variable $Y^*$ and $m$ independent variables (or stimuli), denoted by $X_k$, $k = 1 \dots m$, is

$$Y^* = b_0 + b_1X_1 + b_2X_2 + \dots + b_kX_k + b_mX_m, \quad (1)$$

where $b_k$, $k = 1 \ldots m$, are called the regression parameters. Each parameter $b_k$ can be interpreted as the expected change in response $Y^*$ associated with a 1-unit increase in $X_k$, while the other stimuli are held constant. For a given model, the difference between the observed value of $Y$ and the model-predicted value, $Y^*$, at the same given point, is called the residual.

In a good MLR model, each independent variable explains part of the variation in the dependent variable. If the change in the mean value of $Y^*$ associated with a 1-unit increase in one independent variable (say, $X_1$) depends on the value of a second independent variable (say, $X_2$), then there is interaction between these two variables. One can incorporate this interaction into the MLR model by including the product of the two independent variables, $X_1 X_2$.

To quantify how well a multiple linear regression model fits a data set, we calculate the coefficient of determination $R^2$ and test the utility of the model. For a good model, we want $R^2$ to be close to 1 and the *F-statistic* for the utility test to be high enough. For more details on the multiple linear regression models, see (Devore, 2012).

**The non-parametric one-way ANOVA**

The data is grouped according to the cardinal orientation of the facade of the green wall system. Our first task is to check whether there are significant differences between the average plant cover percentage for the four facades. To this aim, we shall perform the Kruskal-Wallis test. This test provides a nonparametric alternative to the one-way ANOVA. We use this test as not all the requirements for the application of the ANOVA test are met. More specifically, the data is not normally distributed. Within the vertical system we have 4 facades from which we randomly sampled the data, so there are 4 categories of data that we have to compare.

The requirements for the application of the Kruskal-Wallis test are: (1) all samples were randomly selected, (2) the observed values in the samples are independent (which is true, as they were collected independently), and (3) all groups should have similar shape distributions. The last condition could be observed from Figure 1, where we have drawn the boxplots for the four groups of data. From Figure 1 we also observe that there are no outliers in the data.

Since it is a nonparametric test, the Kruskal-Wallis test does not assume a normal distribution of the residuals, unlike the parametric version of this test, the one-way ANOVA test.

The hypotheses for the Kruskal-Wallis test are as follows:

$H_0$: the four samples originate from the same distribution (*the null hypothesis*)

$H_1$: at least one sample stochastically dominates one other sample (*the alternative hypothesis*).

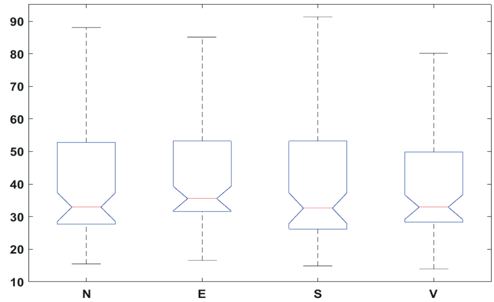


Figure 1. Box-plot for the PCP for each side of the GWS

For the Kruskal-Wallis test and the MLR models derived in this paper we have used the MATLAB software v9.8.0 (R2020a) with the Statistics Toolbox.

## RESULTS AND DISCUSSIONS

### Dependence of the plant coverage percentage on the orientation of the facade

Our first question is whether the plant coverage depend on orientation (N, E, S, W) of the facade of GWS. To answer this, we have performed the Kruskal-Wallis test, a non-parametric variant of the 1-way ANOVA test. This analysis was carried out for the data collected during the period 2020-2021. The table below shows the Kruskal-Wallis ANOVA test results (Figure 2). The p-*value* for the equality of mean coverage percentages on each of the GWS sides is p = 0.2755. The highest observed difference in the mean coverage percentage is 2.74% (between E and W), but it is not significant at the 5% level. In conclusion, the degree of plant cover of the green system does not depend on the cardinal orientation at a significance level of 0.05.

**Kruskal-Wallis ANOVA Table**

| Source | SS | df | MS | Chi-sq | Prob>Chi-sq |
|---|---|---|---|---|---|
| Columns | 33987.1 | 3 | 11329.04 | 3.87 | 0.2755 |
| Error | 2800140.4 | 320 | 8750.44 | | |
| Total | 2834127.5 | 323 | | | |

Figure 2. The Kruskal-Wallis test table

## Correlation between the plant coverage percentage and the inner temperature of the GWS

We are now interested to see whether the plant coverage influence the inner temperature of the GWS. To answer this question, we have calculated the Pearson's correlation coefficient between the inner temperature inside the GWS and the plant coverage percentage for each of the GWS faces. We have carried out this calculation for three types of data: firstly, for the data collected during a 2-year period 2020-2021 (Table 1), secondly, only for the data collected in 2020 (Table 2), and thirdly, only for the data collected in 2021 (Table 3). We observe that, with only one exception, the inner temperature and the plant coverage of the faces are linearly correlated, though not strongly correlated.

Table 1. Pearson correlation coefficients for temperatures and plant coverage percentage for two years (2020-2021)

| **2020-2021 (2 years)** | | **Pearson's coefficient** | **Significance** (p-value) |
|---|---|---|---|
| **Plant coverage percentage for GWS** | *North* | 0.0853 | 0.4488 |
| | *East* | 0.1265 | 0.2605 |
| | *South* | 0.0988 | 0.3804 |
| | *West* | 0.0465 | 0.6800 |
| | *total* **GWS** *coverage* | 0.0914 | 0.4170 |

Table 2. Pearson correlation coefficients for temperatures and plant coverage percentage for 2020 only

| **2020** | | **Pearson's coefficient** | **Significance** (p-*value*) |
|---|---|---|---|
| **Plant coverage percentage for GWS** | *North* | -0.2223 | 0.1681 |
| | *East* | -0.2707 | 0.0912 |
| | *South* | -0.1632 | 0.3143 |
| | *West* | -0.2972 | 0.0625 |
| | *total* **GWS** *coverage* | -0.2368 | 0.1413 |

Table 3. Pearson correlation coefficients for temperatures and plant coverage percentage for 2021 only

| **2021** | | **Pearson's coefficient** | **Significance** (p-*value*) |
|---|---|---|---|
| **Plant coverage percentage for GWS** | *North* | 0.3047 | 0.0527 |
| | *East* | 0.3220 | 0.0401* |
| | *South* | 0.2773 | 0.0793 |
| | *West* | 0.2493 | 0.1160 |
| | *total* **GWS** *coverage* | 0.2911 | 0.0649 |

*Level of significance 0.05

## Regression models of PCP on soil Humidity and soil Temperature

For each face of the GWS, we perform linear regression of the plant coverage percentage ($PCP$) on soil humidity ($H$) and the soil temperature ($T$). We shall consider two models, with or without interaction between the independent variables. For the reliability of the models, the data has been transformed to look fairly normal, using the Box-Cox transformation. All the models below are in terms of the transformed variables.

If $v$ is the Box-Cox transformed variable of $u$ (of parameter $\lambda$), then:

$$v = \begin{cases} \ln u, & if\ \lambda = 0 \\ \dfrac{u^{\lambda} - 1}{\lambda}, & if\ \lambda \neq 0 \end{cases} \qquad (2)$$

The inverse transform is

$$u = \begin{cases} e^{v}, & if\ \lambda = 0 \\ exp\left(\dfrac{ln(1 + \lambda v)}{\lambda}\right), & if\ \lambda \neq 0 \end{cases} \qquad (3)$$

In all models below, we denote by $x_1, x_2$ and y the Box-Cox transformed variables of $H$, $T$ and $PCP$ (with parameters $\lambda_H$, $\lambda_T$, $\lambda_{PCP}$), respectively. So, from now on we shall work with the transformed variables $x_1, x_2$ and y.

We are interested in finding regression models (with or without interactions) of the form

$$y = a \cdot x_1 + b \cdot x_2 + c \cdot x_1 \cdot x_2. \qquad (4)$$

In the original variables, this regression model takes the form:

$$\frac{PCP^{\lambda_{PCP}}-1}{\lambda_{PCP}} = a\cdot\frac{H^{\lambda_H}-1}{\lambda_H} + b\cdot\frac{T^{\lambda_T}-1}{\lambda_T} + c\cdot\frac{H^{\lambda_H}-1}{\lambda_H}\cdot\frac{T^{\lambda_T}-1}{\lambda_T} \quad (5)$$

i.e.

$$PCP^{\lambda_{PCP}} = A\cdot H^{\lambda_H} + B\cdot T^{\lambda_T} + C\cdot H^{\lambda_H}\cdot T^{\lambda_T} + D \quad (6)$$

where $A, B, C, D$ are depending on the parameters $a$, $b$, $c$, $\lambda_H$, $\lambda_T$, $\lambda_{PCP}$. Then, for all the models below, the estimated $PCP$ by the model will be:

$$PCP = (A\cdot H^{\lambda_H} + B\cdot T^{\lambda_T} + C\cdot H^{\lambda_H}\cdot T^{\lambda_T} + D)^{1/\lambda_{PCP}} \quad (7)$$

**Model 1** (north side plant coverage percentage model for GWS without interactions among independent variables and no intercept)
The regression model without interactions among the variables and no intercept has the form:

$$y = a\cdot x_1 + b\cdot x_2 \quad (8)$$

The estimated model parameters are given in Table 4.

Table 4. Parameters for linear regression without interactions

| | Estimate | SE | tStat | pValue |
|---|---|---|---|---|
| $a$ | 0.11454 | 0.021727 | 5.2718 | $1.1454\cdot10^{-6}$* |
| $b$ | 0.14164 | 0.028467 | 4.9756 | $3.7226\cdot10^{-6}$* |

*Level of significance 0.05

Other relevant statistics: number of observations = 81, RMSE = 0.86, $R^2$ = 0.5185, adj$R^2$ = 0.5125.
Therefore, the model is:

$$y = 0.11454\cdot x_1 + 0.14164\cdot x_2 \quad (9)$$

For example, if H = 4 g/kg and T = 12.5°C, then the estimated model value for PCP = 14.72%.

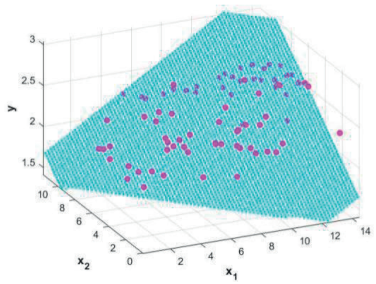


Figure 3. The regression (without interactions) GWS, north side

**Model 2** (north side plant coverage percentage model for GWS with interactions among independent variables, no intercept)

The regression model with interactions among the variables and no intercept has the form:

$$y = a\cdot x_1 + b\cdot x_2 + c\cdot x_1\cdot x_2. \quad (10)$$

The estimated model parameters are given in Table 5.

Table 5. Parameters for linear regression with interactions

| | Estimate | SE | tStat | pValue |
|---|---|---|---|---|
| $a$ | 0.24612 | 0.017285 | 14.239 | $2.1406\cdot10^{-23}$* |
| $b$ | 0.30603 | 0.02221 | 13.779 | $1.3503\cdot10^{-22}$* |
| $c$ | -0.031526 | 0.0026874 | -11.731 | $6.7192\cdot10^{-19}$* |

*Level of significance 0.05

Other relevant statistics: number of observations = 81, RMSE = 0.521, $R^2$ = 0.5034, adj$R^2$ = 0.4906.
Therefore, the model is:

$$y = 0.24612\cdot x_1 + 0.30603\cdot x_2 - 0.031526\cdot x_1\cdot x_2 \quad (11)$$

For example, if H = 4 g/kg and T = 12.5°C, then the estimated model value for PCP = 46.25%.

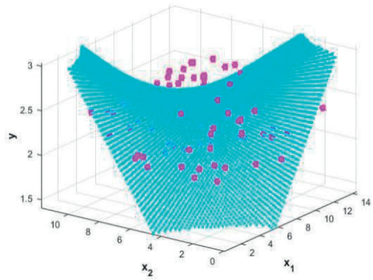


Figure 4. The regression (with interactions) GWS, north side

**Model 3** (east side plant coverage percentage model for GWS without interactions among independent variables, no intercept).
The regression model without interactions among the variables and no intercept has the form:

$$y = a\cdot x_1 + b\cdot x_2. \quad (12)$$

The estimated model parameters are given in Table 6.

Table 6. Parameters for linear regression without interactions

| | Estimate | SE | tStat | pValue |
|---|---|---|---|---|
| $a$ | 0.22674 | 0.046315 | 4.8957 | $5.0893\cdot10^{-6}$* |
| $b$ | 0.18859 | 0.027168 | 6.9419 | $9.6498\cdot10^{-10}$* |

*Level of significance 0.05

Other relevant statistics: number of observations = 81, RMSE = 1.19, $R^2$ = 0.4905, adj$R^2$ = 0.4840.
Therefore, the model is:

$$y = 0.22674 \cdot x_1 + 0.18859 \cdot x_2 \quad (13)$$

If H = 4 g/kg and T = 12.5°C, then the estimated model value for PCP = 18.51%.

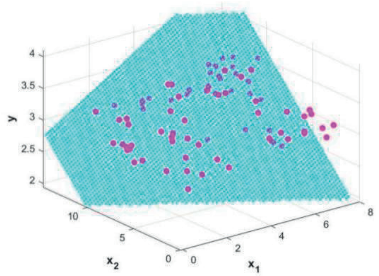

Figure 5. The regression (without interactions) GWS, east side

**Model 4** (east side plant coverage percentage model for GWS with interactions among independent variables, no intercept)

The regression model (*with interactions among independent variables*) has the form:

$$y = a\,x_1 + b\,x_2 + c\,x_1 x_2 \quad (14)$$

The estimated model parameters are given in Table 7.

Table 7. Parameters for linear regression with interactions

| | Estimate | SE | tStat | pValue |
|---|---|---|---|---|
| *a* | 0.49591 | 0.043426 | 11.419 | $2.555 \cdot 10^{-18}$* |
| *b* | 0.34883 | 0.025645 | 13.603 | $2.7549 \cdot 10^{-22}$* |
| *c* | -0.056394 | 0.0061034 | -9.2399 | $3.6987 \cdot 10^{-14}$* |

*Level of significance 0.05

Other relevant statistics: number of observations = 81, RMSE = 0.827, $R^2$ = 0.4612, adj$R^2$ = 0.4473.

Therefore, the model is:

$$y = 0.49591 \cdot x_1 + 0.34883 \cdot x_2 - 0.056394 \cdot x_1 \cdot x_2 \quad (15)$$

If H = 4 g/kg and T = 12.5°C, then the estimated model value for PCP = 34.95%.

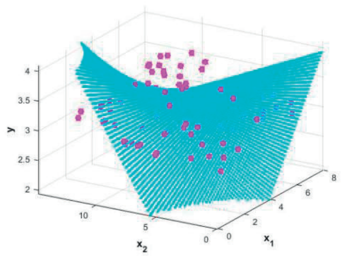

Figure 6. The regression (with interactions) GWS, east side

**Model 5** (south side plant coverage percentage model for GWS without interactions among independent variables, no intercept)

The regression model (without interactions among independent variables) has the form:

$$y = a\,x_1 + b\,x_2 \quad (16)$$

The estimated model parameters are given in Table 8.

Table 8. Parameters for linear regression without interactions

| | Estimate | SE | tStat | pValue |
|---|---|---|---|---|
| *a* | 0.06481 | 0.013107 | 4.9448 | $4.2008 \cdot 10^{-6}$* |
| *b* | 0.11704 | 0.020225 | 5.787 | $1.3813 \cdot 10^{-7}$* |

*Level of significance 0.05

Other relevant statistics: number of observations = 81, RMSE = 0.914, $R^2$ = 0.5082, adj$R^2$ = 0.5020.

Therefore, the model is:

$$y = 0.06481 \cdot x_1 + 0.11704 \cdot x_2 \quad (17)$$

If H = 4 g/kg and T = 12.5°C, then the estimated model value for PCP = 10.16%.

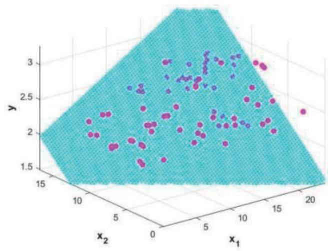

Figure 7. The regression (*without interactions)* GWS, south side

**Model 6** (south side plant coverage percentage model for GWS with interactions among independent variables, no intercept)

The regression model (with interactions among independent variables) has the form:

$$y = a\,x_1 + b\,x_2 + c\,x_1 x_2 \quad (18)$$

The estimated model parameters are given in Table 9.

Table 9. Parameters for linear regression with interactions

| | Estimate | SE | tStat | pValue |
|---|---|---|---|---|
| *a* | 0.14085 | 0.010943 | 12.871 | $5.5242 \cdot 10^{-21}$* |
| *b* | 0.23206 | 0.016748 | 13.856 | $9.8956 \cdot 10^{-23}$* |
| *c* | -0.012987 | 0.0012055 | -10.773 | $4.2101 \cdot 10^{-17}$* |

*Level of significance 0.05

Other relevant statistics: number of observations = 81, RMSE = 0.583, $R^2$ = 0.4623, adj$R^2$ = 0.4485.

Therefore, the model is:

$$y = 0.14085 \cdot x_1 + 0.23206 \cdot x_2 - 0.012987 \cdot x_1 \cdot x_2 \quad (19)$$

If H = 4 g/kg and T = 12.5°C, then the estimated model value for PCP = 30.54%.

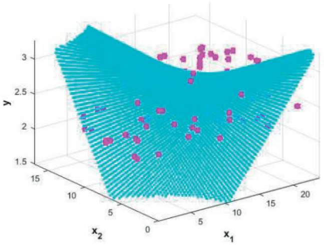


Figure 8. The regression (*with interactions)* GWS, south side

**Model 7** (west side plant coverage percentage model for GWS without interactions among independent variables, no intercept)

The regression model (*without interactions among independent variables*) has the form:

$$y = a\,x_1 + b\,x_2 \quad (20)$$

The estimated model parameters are given in Table 10.

Table 10. Parameters for linear regression without interactions

| | Estimate | SE | tStat | pValue |
|---|---|---|---|---|
| $a$ | 0.10707 | 0.01951 | 5.4881 | $4.7567 \cdot 10^{-7}$* |
| $b$ | 0.20832 | 0.036646 | 5.6846 | $2.1157 \cdot 10^{-7}$* |

*Level of significance 0.05

Other relevant statistics: number of observations = 81, RMSE = 1.28, $R^2 = 0.5069$, adj$R^2 = 0.5007$.

Therefore, the model is:

$$y = 0.10707 \cdot x_1 + 0.20832 \cdot x_2 \quad (21)$$

If H = 4 g/kg and T = 12.5°C, then the estimated model value for PCP = 16.55%.

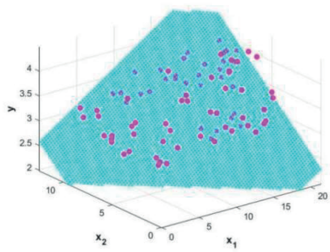


Figure 9. The regression (without interactions) GWS, west side

**Model 8** (west side plant coverage percentage model for GWS with interactions among independent variables, no intercept)

The regression model (with interactions among independent variables) has the form:

$$y = a\,x_1 + b\,x_2 + c\,x_1 x_2 \quad (22)$$

The estimated model parameters are given in Table 11.

Table 11. Parameters for linear regression with interactions

| | Estimate | SE | tStat | pValue |
|---|---|---|---|---|
| $a$ | 0.23439 | 0.017735 | 13.216 | $1.3306 \cdot 10^{-21}$* |
| $b$ | 0.40336 | 0.030502 | 13.224 | $1.2889 \cdot 10^{-21}$* |
| $c$ | -0.027392 | 0.0026507 | -10.334 | $2.8906 \cdot 10^{-16}$* |

*Level of significance 0.05

Other relevant statistics: number of observations = 81, RMSE = 0.839, $R^2 = 0.4847$, adj$R^2 = 0.4715$.

Therefore, the model is:

$$y = 0.23439 \cdot x_1 + 0.40336 \cdot x_2 - 0.027392 \cdot x_1 \cdot x_2 \quad (23)$$

If H = 4 g/kg and T = 12.5°C, then the estimated model value for PCP = 35.27%.

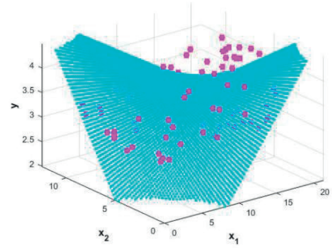


Figure 10. The regression (with interactions) GWS, west side

**Regression models of *PCP* on *soil Humidity, soil Temperature* and *WeekNo* (time of the year).**

We have available 324 observations that were collected during the years 2020-2021. More precisely, we have 81 values recorded for each face of a GWS.

To be able to give better estimates for the plant coverage percentage for each face of the system, we have introduced a new explanatory variable, called *WkNo*. This variable keeps track of the time of the year when the data was recorded. *WkNo* takes values from 1 (first week of the year) to 52 (last week of the year).

In the following models, we perform linear regression of the plant coverage percentage

($PCP$) on soil humidity ($Hum$), the soil temperature ($Temp$) and the week number (WkNo) for each face of a GWS. We shall consider below a linear regression model of $PCP$ on $Hum$, $Temp$ and $WkNo$ with an interaction term between $Temp$ and $WkNo$. For each face, the regression models are based on 81 data sets.

In all models below, we shall consider:

*Explanatory variables:* soil Humidity (*Hum*), soil Temperature (*Temp*), Week number (*WkNo*).

*Explained variable:* Plant coverage percentage (*PCP*)

The regression model has the general form:

$$\boldsymbol{PCP} = \boldsymbol{a} \cdot \boldsymbol{Hum} + \boldsymbol{b} \cdot \boldsymbol{Temp} + \boldsymbol{c} \cdot \boldsymbol{WkNo} + \boldsymbol{d} \cdot \boldsymbol{Temp} \cdot \boldsymbol{WkNo}. \quad (24)$$

with *a, b, c, d* the regression coefficients to be determined.

**Model 1** (plant coverage percentage model for **GWS**, north side):

The estimated model parameters are given in Table 12.

Table 12. Parameters for linear regression with interactions

| | Estimate | SE | tStat | pValue |
|---|---|---|---|---|
| A | 8.3459 | 1.2407 | 6.7266 | $2.7254 \cdot 10^{-9}$* |
| B | -2.0276 | 0.52532 | -3.8598 | 0.00023448* |
| C | 0.41093 | 0.14171 | 2.8997 | 0.004865* |
| D | 0.061527 | 0.016631 | 3.6996 | 0.00040312* |

*Level of significance 0.05

Other relevant statistics: number of observations = 81, RMSE = 14.1, $R^2$ = 0.6194, adj$R^2$ = 0.6046. Therefore, the model is:

$$\boldsymbol{PCP} = \boldsymbol{8.3459} \cdot \boldsymbol{Hum} - \boldsymbol{2.0276} \cdot \boldsymbol{Temp} + \boldsymbol{0.41093} \cdot \boldsymbol{WkNo} + \boldsymbol{0.061527} \cdot \boldsymbol{Temp} \cdot \boldsymbol{WkNo}$$

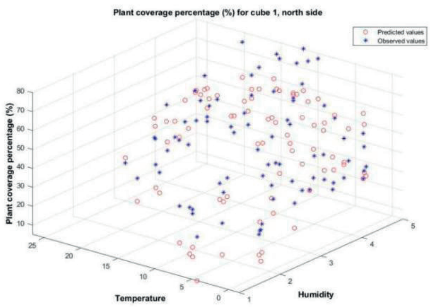


Figure 11. Predicted values vs observed values GWS, noth side

**Model 2** (plant coverage percentage model for **GWS**, east side):

The estimated model parameters are given in Table 13.

Table 13. Parameters for linear regression with interactions

| | Estimate | SE | tStat | pValue |
|---|---|---|---|---|
| *a* | 4.4503 | 1.1693 | 3.8061 | 0.00028159* |
| *b* | -0.16489 | 0.50551 | -0.32619 | 0.74516 |
| *c* | 0.82419 | 0.13345 | 6.176 | $2.8768 \cdot 10^{-8}$* |
| *d* | 0.010607 | 0.016575 | 0.63993 | 0.52412 |

*Level of significance 0.05

Other relevant statistics: number of observations = 81, RMSE = 16.6, $R^2$ = 0.4827, adj$R^2$ = 0.4625. Therefore, the model is:

$$\boldsymbol{PCP} = \boldsymbol{4.4503} \cdot \boldsymbol{Hum} - \boldsymbol{0.16489} \cdot \boldsymbol{Temp} + \boldsymbol{0.82419} \cdot \boldsymbol{WkNo} + \boldsymbol{0.010607} \cdot \boldsymbol{Temp} \cdot \boldsymbol{WkNo}$$

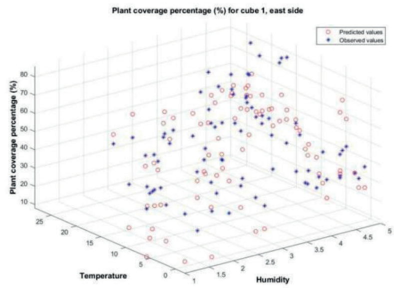


Figure 12. Predicted values vs observed values GWS, east side

**Model 3** (plant coverage percentage model for **GWS**, south side):

The estimated model parameters are given in Table 14.

Table 14. Parameters for linear regression with interactions

| | Estimate | SE | tStat | pValue |
|---|---|---|---|---|
| *a* | 5.7456 | 1.2084 | 4.7548 | $9.0727 \cdot 10^{-6}$* |
| *b* | -1.1205 | 0.47392 | -2.3644 | 0.020581* |
| *c* | 0.7544 | 0.1353 | 5.5758 | $3.5032 \cdot 10^{-7}$* |
| *d* | 0.030779 | 0.014811 | 2.0781 | 0.041031* |

*Level of significance 0.05

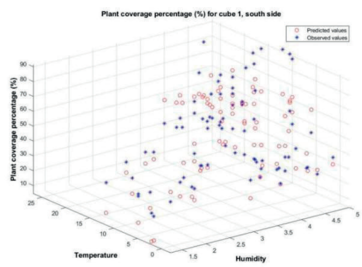


Figure 13. Predicted values vs observed values GWS, south side

Other relevant statistics: number of observations = 81, RMSE = 15.1, $R^2$ = 0.5812, adj$R^2$ = 0.5646. Therefore, the model is:

$$PCP = 5.7456 \cdot Hum - 1.1205 \cdot Temp + 0.7544 \cdot WkNo + 0.030779 \cdot Temp \cdot WkNo$$

**Model 4** (plant coverage percentage model for **GWS**, west side)

The estimated model parameters are given in Table 15.

Table 15. Parameters for linear regression with interactions

| | Estimate | SE | tStat | pValue |
|---|---|---|---|---|
| $a$ | 6.6578 | 1.0354 | 6.43 | $9.7645 \cdot 10^{-9}$* |
| $b$ | -1.2003 | 0.45307 | -2.6492 | 0.0097868* |
| $c$ | 0.54673 | 0.12459 | 4.3884 | $3.5857 \cdot 10^{-5}$* |
| $d$ | 0.034149 | 0.014662 | 2.329 | 0.022481* |

*Level of significance 0.05

Other relevant statistics: number of observations = 81, RMSE = 12.9, $R^2$ = 0.6074, adj$R^2$ = 0.5921. Therefore, the model is:

$$PCP = 6.6578 \cdot Hum - 1.2003 \cdot Temp + 0.54673 \cdot WkNo + 0.034149 \cdot Temp \cdot WkNo$$

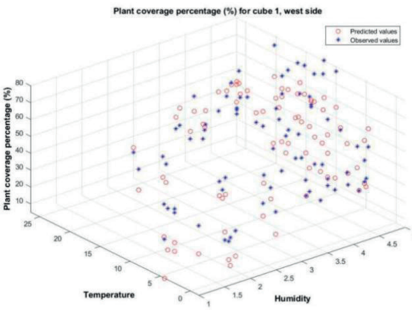


Figure 14. Predicted values vs observed values GWS, west side

## CONCLUSIONS

Based on our linear regression models, we can draw the following conclusions:

1. The plant cover percentage of the green system does not depend on the cardinal orientation at a significance level of 0.05.
2. The plant cover percentage of the green system does not influence the indoor temperature of the system at a significance level of 0.05.
3. We have built regression models in which all regression coefficients are significant at level 0.05. These models may help in predicting the plant coverage percentage of the vertical wall using values of soil moisture and soil temperature.
4. By considering the soil Humidity (*Hum*), soil Temperature (*Temp*), and an extra explanatory variable, *WeekNo* (time of the year), we were searching for linear regression models that can explain the percentage of plant cover of each side of the green system. Among all the models we have tried, the best ones (in terms of lowest MSE, and highest $R^2$ and adj$R^2$) involved the interaction between the variables *Temp* and *WeekNo*. For example, in the case of the regression model for the plant coverage percentage on the north side of the system, we have got the following equation:

$$PCP = 8.3459 \cdot Hum - 2.0276 \cdot Temp + 0.41093 \cdot WkNo + 0.061527 \cdot Temp \cdot WkNo$$

We can interpret this equation as follows:

- an increase by 1 unit in humidity will determine an increase by almost 8.35% in the plant coverage percentage;
- an increase by 1°C in temperature will determine a decrease by almost 2.03% in the plant coverage percentage;
- the influence of the week number on the plant coverage percentage is 0.41093;
- for each week number, an increase by 1°C in temperature will determine an increase by almost 0.06% in plant coverage percentage

5. Apart from the east side of GWS, all models have significant coefficients, with a coefficient of determination $R^2$ about 0.5.
6. We have observed a small improvement in the models when considering the time of the year *WeekNo* as a new explanatory variable.
7. The use of mathematical modeling through regression analysis improved and enriched the results of the study, helping to determine more accurate predictions on the development and evolution of a green wall system.
8. We are aware that other factors (such as watering conditions, exposure to sunlight, human intervention, type of soil etc.) may also influence the plant cover percentage of the GWS, thus our models cannot depict the whole evolution of plants on the green system. Should additional data be available, our models can be further improved.

## ABBREVIATION

| Notation | Explanation |
|---|---|
| **GWS** | Green wall system |
| **PCP** | Plant coverage percentage |
| **Hum** | Soil Humidity |
| **Temp** | Soil Temperature |
| **WkNo** | Week number |
| **N** | Cardinal point North |
| **E** | Cardinal point East |
| **S** | Cardinal point South |
| **W** | Cardinal point West |

## ACKNOWLEDGMENTS

We thank the National Meteorological Administration - Moldova Regional Meteorological Center for the climate data provided.

## REFERENCES

Chiruţă, C., Stoleriu, I., Cojocariu, M. (2023). Prediction Models for the Plant Coverage Percentage of a Vertical Green Wall System: Regression Models and Artificial Neural Network Models. *Horticulturae, 9*, 419, DOI: 10.3390/horticulturae9040419.

Cojocariu, M., Chelariu, E.L., Chiruţă, C. (2022a). Study on Behavior of Some Perennial Flowering Species Used in Vertical Systems for Green Facades in Eastern European Climate, *Applied Sciences, 12*(1), 474, https://doi.org/10.3390/app12010474.

Cojocariu, M., Chelariu, E.L., Chiruță, C., Pașcu, R., Avarvarei, B.V. (2022b). Comparative Study on ohe Behaviour of *Plectranthus Forsteri* and *Coleus Blumei* Species Growing on the Ground and in Vertical Systems for Green Façades in the Climate of North-East Romania*, Scientific Papers. Series B, Horticulture*, Vol. LXVI, No. 1, pp. 657-666.

Chelariu E L, Draghia L, Brînză M, Cojocariu M, Avarvarei B-V. (2018), Studies regarding the behaviour of ornamental species *Lagurus ovatus* in cropping conditions from N-E area of Romania, *Scientific Papers. Series B, Horticulture*, Vol. LXII, pp 581-584.

Devore, J.L.; Berk, K.N. (2012). *Modern Mathematical Statistics with Applications*; Springer Texts in Statistics; Springer: New York, NY, USA, ISBN 978-1-4614-0390-6.

Draghia, L.; Chelariu, E.L. (2011). *Floriculture (in Romanian)*. "Ion Ionescu de la Brad" Publishing House, Iasi, Romania.

Fan, Y.; Das, K.V.; Chen, Q. (2011). Neighbourhood Green, Social Support, Physical Activity and Stress: Assessing the Cumulative Impact, *Health Place*, 17, 1202–1211.

Francis, A.R.; Lorimer, J. (2011). Urban reconciliation ecology: The potential of living roof sand walls. *Journal of Environmental Management*. 92(6), 1429-1437.

Ghazalli Aini Jasmin, Brack, C., Bai Xuemei, Said, I. (2019). Physical and Non-Physical Benefits of Vertical Greenery Systems: A Review, *Journal of Urban Technology*, doi.org/10.1080/10630732.2019.1637694.

*MATLAB and Statistics Toolbox Release, 2022b*. The MathWorks, Inc.

Pérez G., Chocarro, C., Juárez, A., Coma, J. (2020). Evaluation of the development of five Sedum species on extensive green roofs in a continental Mediterranean climate, *Urban Forestry & Urban Greening*, 48, doi.org/10.1016/j.ufug.2019.126566.

Palermo, S.A.; Turco, M. (2020). Green Wall Systems: Where do we stand? *IOP Conference Series Earth and Environmental Science*. 2020, 410. doi.org/10.1088/1755- 1315/410/1/012.013.

Perini, K.; Ottel´e, M.; Haas, E.M.; Raiteri, R.; Ungers, O.M. (2011). Greening the building envelope, façade greening and living wall systems. *Open Journal of Ecology*, 1, 1-8. doi.org/10.4236/oje.2011.11001.

Perez-Urrestarazu, L.; Fernandez-Canero, R.; Franco-Salas, A.; Egea, G. (2015). Vertical Greening Systems and Sustainable Cities. *J. Urban Technol.*, 22, 65–85.

Ropan, L.M., Trusca, (Popescu) R.I., Giugea, N., Simon, A.M., Horsia, C.A. (2023). Global Radiation and its Importance for White (Young) Wine Quality in the Jidvei Wine-Growing Center. *Scientific Papers. Series B, Horticulture*, Vol. LXVII, Issue 2, Print ISSN 2285-5653, 275-284.

Rosu S., Herbei M.V., Sala F. (2022). Peculiarities of the Linden Leaf Area in Relation to the Leaf Position on the Shoot. *Scientific Papers. Series B, Horticulture*, Vol. LXVI, Issue 1, Print ISSN 2285-5653, 893-900.

Sheweka, S.M.; Nourhan, M.M. (2012). Green Façades as a New Sustainable Approach towards Climate Change. *Energy Procedia*, 18, 507–520.

Stoyanova, A., Kuneva, V., Valchev, N. (2019). Survey of the Influence of Fertilization and Irrigation in the Tomatoes, Greenhouse Production. *Scientific Papers. Series B, Horticulture*, Vol. LXIII, Issue 1, Print ISSN 2285-5653, 315-320.

Tripon, (Burduhos) D., Odagiu, A., Oroian, I., Balint, C., Arnautu, V. (2023). Climatic Inputs Influence on the Yields of Wild Berries Harvested from Transylvanian Testing Spontaneous Flora. A Case Study: *Rubus idaeus l.*, and *Ribes nigrum l.* in the Season of 2022. *Scientific Papers. Series B, Horticulture*, Vol. LXVII, Issue 1, Print ISSN 2285-5653, 860-866

Wackerly, D., Mendenhall, W., Scheaffer, R.L. (2007). *Mathematical Statistics with Applications,* Duxbury Press,7th edition.