SCIENTIFIC HORIZONS

Automated remote sensing system for crops monitoring and irrigation management, based on leaf color change and piecewise linear regression models for soil moisture content predicting

Svetoslav Atanasov*

PhD Student, Magister in Computer System and Technologies
Trakia University
6015, Students campus, Stara Zagora, Bulgaria
https://orcid.org/0000-0002-2658-1611

Abstract. Plants can serve as biological sensors if their "readings" and the feedback they provide us through changes in the colour of their leaves can be correctly interpreted. The study aims to predict soil moisture and, as such, the need for irrigation, using nonlinear mathematical models, describing the relationship between RGB and HSL colour model components and soil moisture and temperature. Nonlinear mathematical models used in the study are based on piecewise linear regression with breakpoint and soil moisture prediction using colour components and soil temperature with a deviation of +-6%. A system for automated irrigation was created and its control program was made, the basic control law of which is based on non-linear piecewise linear models. The automated irrigation management system includes a remote crop monitoring subsystem and an irrigation management subsystem. The program processes the photo received from the camera and activates the actuators when watering is needed. Compared to manual data collection in the first part of the study, the program calculates the average RGB model values from images in the studied row of tomato plantations with an accuracy of over 99% for the R and G components and over 92% for the B component. The program also predicts soil moisture with 98% accuracy. The practical significance of the water-saving efforts of this study lies in the development of a program-controlled automated irrigation system that utilizes plants as biological sensors, employing nonlinear mathematical models based on leaf colour changes to accurately predict soil moisture

Keywords: biosensors; precision irrigation; RGB colourimetry; image processing; digitalization; bioinformatics

*Corresponding author

INTRODUCTION

People knew when their crops required water or treatment by their appearance (e.g., wilting, wrinkling) or their colour (lightening or darkening of their green colour) for centuries. Precision irrigation offers numerous benefits in modern agriculture. This targeted approach also improves crop yields and quality while minimizing environmental impact. Additionally, precision irrigation systems can be remotely controlled and monitored, allowing farmers to make real-time adjustments based on weather conditions and soil moisture levels, ultimately enhancing overall farm productivity and sustainability.

Soil moisture refers to the amount of water present in the soil, typically measured as a percentage of the soil's weight. Monitoring soil moisture is crucial from a precision irrigation perspective as it allows farmers to optimize their water usage. By determining the exact moisture levels in the soil, farmers can apply irrigation precisely, providing just the right amount of water needed. This not only conserves water resources but also ensures that crops receive adequate hydration for optimal growth and yield. Monitoring soil moisture levels helps prevent over-irrigation, which wastes water and causes environmental issues, as well as under-irrigation, which can harm crop health and productivity.

However, irrigation efficiency and water resource preservation are often lower than expected (Zahoor *et al.*, 2019; Laureti *et al.*, 2021). It is estimated that agriculture uses over 70% of the world's freshwater supplies, with about half of this being lost and wasted (United Nations Development Programme, 2021). Furthermore, during the irrigation process, most of the water is absorbed by the soil whereas plants get only a small portion. Tomatoes lose over 90% of absorbed water through transpiration from their leaves, making them more vulnerable to water stress than other parts of the plant, such as fruit (Zhao *et al.*, 2020), therefore, they are more sensitive and a suitable indicator of water stress.

There are two widely recognized traditional approaches for monitoring irrigation: groundwater monitoring (Wood & Cherry, 2021) and plant water monitoring (Simbeye *et al.*, 2023). Various methods for assessing soil moisture are highlighted by several authors, including the gravimetric approach (Reich *et al.*, 2021), TDR dielectric methods (Gnatowski *et al.*, 2018), FDR dielectric methods (Kang *et al.*, 2019), and neutron probe. Measurement techniques for assessing plant water status encompass ZIM probes, dendrometers, infrared thermometry measurements, pressure chambers, infrared gas analysers, pulse methods, and porometers. E. Serrano-Finetti *et al.* (2023) suggest employing non-invasive plant leaf water content monitoring using electrical impedance spectroscopy.

The study describes a different, indirect, intelligent, quick approach based on physical observation of plants with the naked eye, used by farmers for a long time, which also has scientific potential. A method that provides information that the plant is well watered and that eliminates the drawbacks of the available sensors. This method must be sufficiently accurate to ensure optimal irrigation and water saving.

Modern studies (Atanasov, 2021) state that plants require irrigation when their leaves turn dark green and lighten as the soil moisture increases. However, the present study does not include cases of infected or diseased plants with discoloured leaves, but only healthy ones. Atanasov (2021) also proves that the colour change is an early indicator of water stress, which can be detected within half an hour to an hour after irrigation depending on the height of the plant, or it can be monitored continuously with an automated monitoring system proposed late in the present study. The same aforementioned study also demonstrated that young leaves (at the top of the plant) are more informative regarding the need for watering (they provide smaller errors when predicting soil moisture). Therefore, twice as many RGB and HSL colour samples were taken from them in the studies.

The study demonstrates the potential of the natural world. The study author suggests that plants can become reliable living biosensors – essentially high-tech biogadgets. They employ leave colour as an indicator, and a smart and cost-effective solution for directly measuring soil moisture by analysing colour changes in tomato and other plants' leaves grown in a greenhouse environment or outdoors can be created on this basis. As such, the study aims to build an interactive model to predict soil moisture levels and, as a result, provide optimal irrigation timing. This study also seeks to employ the natural traits of plants for further use. All of this can be achieved by investigating the type of functional dependence describing the relationship between the average value of the parameters of the RGB and HSL colour components (R_{avg}, G_{avg}, B_{avg}, H_{avg}, S_{avg}, L_{avg}) and the factors of soil moisture (θ) and soil temperature (T).

MATERIALS AND METHODS

Methodology of the experiment. The experimental approach involves conducting research in heated glass greenhouses situated in Plovdiv, Bulgaria, at coordinates 42.18265, and 24.73989. These greenhouses meticulously control all environmental variables and ensure the well-being of plants by maintaining optimal conditions. To provide winter heating, hot water is sourced from a nearby thermal power plant using steam heating. The greenhouse dimensions are approximately 50 m by 35 m, with 15 columns of 3 m in length and each row spanning about 2 m in width. Tomatoes are cultivated year-round, yielding two harvests – one from January to July and another from August to December.

The investigated indeterminate tomatoes are one variety – Panekra. They were planted at the end of July. The experiments were conducted in September – 24 hours after irrigation and 24 hours before irrigation.

The study object is the leaf mass of tomato plants (*Solanum lycopersicum*) and the influence of microclimate parameters in the greenhouse (humidity and soil temperature). Leaf mass was examined for four quality factors – young leaves (C_1) before (A_1) and after irrigation (A_2), old leaves (C_2) before and after irrigation. Young leaves (Fig. 1 – left and Fig. 2 – left) are the new, underdeveloped leaves at the top of the plant. Old leaves (Fig. 1 – right and Fig. 2 – right) are large, fully developed leaves at the bottom of the plant.

Figure 1. *Photos of young (left) and old leaves (right) after irrigation*

Figure 2. *Photos of young (left) and old leaves (right) before irrigation*
Source: *photos by the author*

Throughout the measurements, wireless sensors were employed to monitor soil moisture and temperature, a handheld portable colourimeter (PCE-RGB2) was utilized for colour measurement, and a device (PCE-EM 883) was employed to measure air humidity, temperature, and luminance. The calibration of the soil moisture sensor for the specific greenhouse soil type, completed as outlined by J. Starr and I. Paltineanu (2002) for calibrating capacitive sensors, was conducted, as the greenhouse soil is of an alluvial-meadow nature.

As both factors θ and T are measurable, the study delved into the relationship between them and the average values of each component in the RGB and HSL colour models through multivariate regression analysis.

The influence of qualitative factors, such as "before irrigation", "after irrigation", "young leaves", and "old leaves", were investigated using multivariate analysis of variance. The examination of normal distribution utilized the entire set of obtained experimental data (random variables), while in regression and dispersion analysis, the averaged values of the RGB and HSL colour components for each plant (R_{avg}, G_{avg}, B_{avg}, H_{avg}, S_{avg}, L_{avg}) were employed.

Sample size. In the present study, 168 RGB and HSL tomato leaf colour samples were taken manually using a colourimeter before and 168 after irrigation. Regression and variance analysis were performed and models of the dependence between colour components and θ and T were obtained. The RGB readings for each plant

are rounded to the second digit, although the RGB components have no fractional part. Soil moisture (θ) and soil temperature (T) were measured repeatedly and averaged, close to the plant and the root, in a layer of soil 20-30 cm, which is also the root habitat.

Conducting the experiments. The experiment took place under diverse weather conditions. In conditions of clear and sunny weather, the greenhouse registered a humidity level of 73% and a temperature of 25.7°C. Concurrently, outside the greenhouse, the humidity was 65.8%, and the temperature remained at 25.5°C, with an illumination range of 39-47 kLx within the greenhouse. Additionally, during cloudy and rainy weather, the greenhouse experienced increased humidity at 84%, a lower temperature of 19.8°C, and illumination within the range of 5.2-6.2 kLx. Outside the greenhouse, the humidity increased to 85.8%, and the temperature dropped to 16.8°C.

Based on the experience gained during the experiments, the following experiment plan or the following data collection methodology was created and named "12-9-6-3": twelve randomly selected plants (several from each of the three sections of the row) were studied – nine for mathematical model creation and three to test the model (before irrigation). Measurements used six-fold repeatability for young leaves and three-fold repeatability for old leaves – six colour values were taken randomly from different leaves from the crown (top) of each plant examined and three colour values were randomly taken from the lowest their leaves.

To facilitate the process, the plants are marked with an easily visible marking, e.g. red thread. Measured data is entered into a purpose-built field log. The experiment was conducted at the end of a specially dedicated row. The experiment section is 15.53 m long with controlled irrigation using two valves. Near the first valve, inside the greenhouse, as well as for the entire greenhouse, the irrigation rate is 30 m³/acre. The irrigation rate between the first and second valve is 20 m³/acre and 10 m³/acre near the second valve to the concrete path. Fertilization is carried out by dissolving 20 kg of fertilizer and using the solution for drip irrigation with water – 20 m³/acre (Fig. 3):
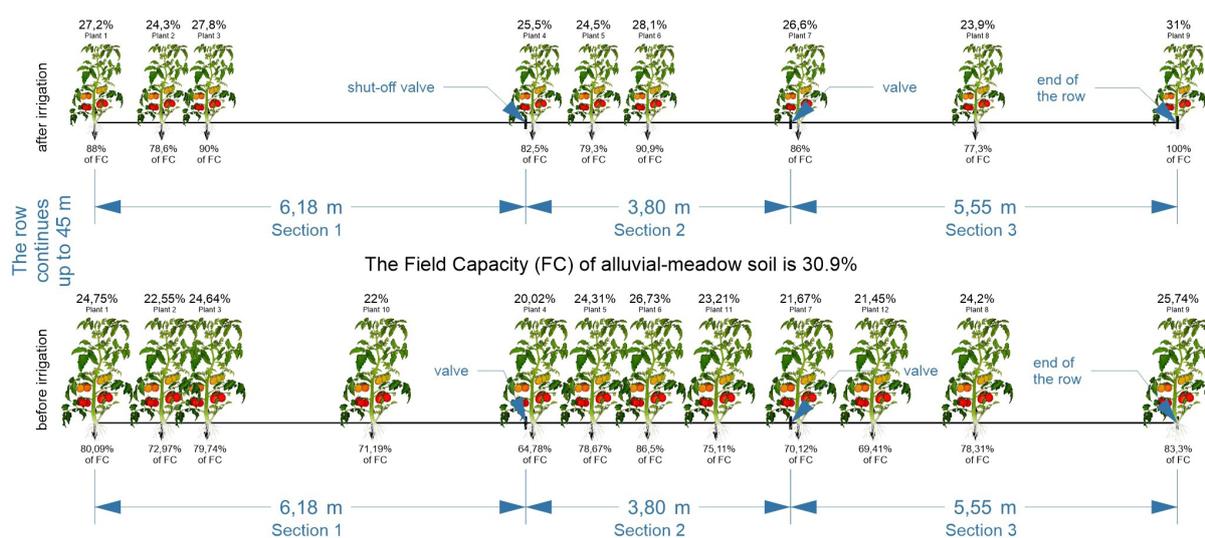
Figure 3. *Visualization of the studied plants and methodology "12-9-6-3" at the row in the greenhouse intended for this purpose*

Source: *compiled by the author*

Multivariate analysis of variance. The variance analysis demonstrated that both "young/old leaves" and "before/after irrigation" factors have a significant influence on the colour component R_{avg}. The "young/old leaves" factor has a stronger influence. Only the "young/old leaves" factor has a proven influence on G_{avg} and B_{avg} colour components. The components R_{avg}, G_{avg} and B_{avg} have the largest values (lighter green colour when combined) in the older leaves after irrigation, which demonstrates that leaves darken when dry.

This study determines a model with functional dependence between the studied parameters, where the coefficient of determination R^2 is >= 80%, as suitable, which means that it is capable of predicting soil moisture. That is why the non-linear estimation approach is used to find the relationship between the set of independent variables (θ and T) as predictors and the dependent variables (colour components) which vary over the independent variables.

Modeling with Piecewise linear regression with breakpoint. This method works as follows: the model is divided into two parts and the formula is solved using the Quasi-Newton nonlinear method, which is an iterative, iterative, nonlinear optimization method that greatly minimizes inconsistencies and errors in colour prediction, and which is used to minimizing the sum of squared errors (Araújo-Paredes *et al.*, 2022). Breakpoints are fundamental when using piecewise regression, as

they are points in the range of X where the behaviour of Y changes, thus the name "breakpoint". In some cases, the breakpoint can be known, but normally it is not. In such cases, there is a need for a method to determine it.

In the Quasi-Newton method, the process of constructing a model to predict leaf colour involves calculating the first-order derivative of the function at a given point to ascertain its slope. Subsequently, the second derivative indicates the rate at which the slope changes at that point and its direction. The Quasi-Newton method iteratively evaluates the function at various points during each step, computing both first and second-order derivatives. These derivatives are crucial for determining the minimum of the loss function. The choice of methods to minimize the loss function depends on the objective, with the least squares method being the most prevalent. This method entails squaring the difference between predicted and observed values. The end goal is to minimize the disparity between the observed and predicted values (Araújo-Paredes *et al.*, 2022).

RESULTS AND DISCUSSION

Investigating the distribution of colour components (random variables). During the normal distribution check, almost all RGB component values studied were found to have a Gaussian distribution, except for B in young leaves after irrigation. A normal distribution in HSL for the following colour components was present – S and L in young leaves after irrigation, S and L in young leaves before irrigation and H, S and L in old leaves before irrigation. For those colour components that have a Gaussian distribution, prerequisites are

available for moving on to the next steps: regression and variance analyses.

Multivariate analysis of variance. In the HSL colour model, both "young/old leaves" and "before/after irrigation" factors have a significant influence on the H_{avg} component. The component H_{avg} has the greatest value in young leaves before irrigation. None of the two factors "young/old leaves" and "before/after irrigation" had a proven influence on the S_{avg} component. Only the factor "young/old leaves" has a proven influence on the L_{avg} component. The components H_{avg} and S_{avg} have the greatest values in young leaves before irrigation. The component L_{avg} has the largest values in old leaves after irrigation.

Multivariate (bivariate) regression analysis. Regression analysis concluded that multiple linear regression and quadratic polynomial regression modelling were inappropriate (unsatisfactory coefficient of determination and non-significant models). The studied environmental factors also have an inherent non-linear behaviour concerning leaf colour. In other words, the variation of the θ and T data does not follow any recognizable linear combination concerning leaf colour. Therefore, it is difficult to model a required dynamic relationship using conventional linear methods such as multiple linear regression. A universal non-linear model is believed to best describe the colour components at different soil moistures and temperatures.

Modelling with Piecewise linear regression with breakpoint. The following coefficients of the empirical equation were obtained using this method. The general form of the model, calculated using the least squares method, is as follows:

$$y = (b_{01} + b_{11}x_1 + \ldots + b_{m1}x_m)(y \le b_n)(b_{02} + b_{12}x_1 + \ldots + b_{m2}x_m)(y > b_n), \tag{1}$$

Alternatively, using the parameters R_{avg}, G_{avg}, B_{avg}, H_{avg}, S_{avg}, L_{avg}, θ and T, the model is as follows:

$$\{R_{avg}, G_{avg}, B_{avg}, H_{avg}, S_{avg}, L_{avg}\} = (b_{01} + b_{11}\theta + b_{21}T) \text{ (for } \{R_{avg}, G_{avg}, B_{avg}, H_{avg}, S_{avg}, L_{avg}\} \le \text{breakpoint } b_0)$$
$$\text{Or } (b_{02} + b_{12}\theta + b_{22}T) \text{ (for } \{R_{avg}, G_{avg}, B_{avg}, H_{avg}, S_{avg}, L_{avg}\} > \text{breakpoint } b_0), \tag{2}$$

In this way, two separate linear regression equations are calculated – one for the y values that are less than or equal to the breakpoint (b_0) and one for y values greater than the breakpoint. For the R_{avg} colour component in young leaves after irrigation, according to model (2), the coefficient of determination was 98.67%. It indicates that 98.67% of the variation in R_{avg} is attributed to changes in the θ and T percentages. The remaining 1.33% is attributable to factors not included in the model. The conclusions

are similar for the other investigated colour components possessing a normal distribution – G_{avg}, S_{avg} and L_{avg}.

A very important conclusion was also established – the physical definition of the breaking point coincides with the arithmetic mean value of the given colour component (in this case breakpoint = 119.13). The coefficients of the models concerning the remaining colour components in young leaves after irrigation according to model (2) are summarized in Table 1:

Table 1. *Coefficients and breakpoint for R_{avg}, G_{avg}, S_{avg} and L_{avg} models in young leaves after irrigation*

Model variable	Coefficients	R_{avg}	G_{avg}	S_{avg}	L_{avg}
Constant	b_{01}	868.17	2.35	-0.106	0.781
θ	b_{11}	-3.80	-12.44	-0.007	-0.003
T	b_{21}	-31.49	20.90	0.024	-0.029
Constant	b_{02}	10.35	28.84	0.455	0.052

Table 1. Continued

Model variable	Coefficients	R_{avg}	G_{avg}	S_{avg}	L_{avg}
θ	b_{12}	-1.06	-1.24	-0.002	-0.0005
T	b_{22}	6.66	6.99	-0.007	0.0036
	breakpoint	119.13	140.09	0.235	0.111
R		0.99	0.99	0.88	0.99
Variations covered (%)		98.67%	98.03%	77.54%	99.47%
R^2		0.99	0.98	0.78	0.99

Source: *compiled by the author*

This iterative approach is well-suited for scenarios involving multiple independent variables and a dependent variable (colour component) with values both above and below the breakpoint. The residual values are maintained at acceptably small levels through a nonlinear optimization method, ranging between -1.443 and 1.094. The predicted values closely align with the observed values. All data points from the experiments fall within or very near the narrow confidence interval (at a confidence level of γ=0.95), and the regression line closely tracks the experimental points.

The residual values, distributed uniformly and within acceptable limits, contribute to the model's reliability. The high R^2 value of 0.99 for R_{avg} in young leaves after irrigation amplifies a significant dependence of the colour component on the variables incorporated into the model. The points cluster closely around the straight line, suggesting a normal distribution of residuals. Similar reasoning can be extended to all other considered colour components in old leaves after irrigation based on model (2). The coefficients of the models related to colour components in young leaves before irrigation are summarized in Table 2.

Table 2. *Coefficients and breakpoint for R_{avg}, G_{avg}, B_{avg}, S_{avg} and L_{avg} models, young leaves before irrigation*

Model variable	Coefficients	R_{avg}	G_{avg}	B_{avg}	S_{avg}	L_{avg}
Constant	b_{01}	104.72	137.12	57.18	0.357	0.112
θ	b_{11}	-2.12	-2.59	-0.59	-0.0023	-0.0016
T	b_{21}	2.48	2.57	1.63	-0.0032	0.0014
Constant	b_{02}	-746.30	-1727.72	131.70	0.383	0.105
θ	b_{12}	9.26	19.02	0.25	0.006	-0.132
T	b_{22}	33.59	74.11	-2.98	-0.0127	0.189
	breakpoint	108.74	133.89	77.35	0.264	0.113
R		0.98	0.99	0.98	0.90	0.77
Variations explained (%)		96.98%	97.51%	95.14%	81.44%	59.22%
R^2		0.97	0.98	0.95	0.81	0.59

Source: *compiled by the author*

Figure 4 also shows acceptably small residual values, with predicted values very close to the observed values:

	Observed	Predicted	Residuals
1	100.83	99.91	0.92
2	115.50	115.50	-0.00
3	95.50	98.48	-2.98
4	108.50	109.56	-1.06
5	104.33	100.72	3.62
6	129.33	129.33	-0.00
7	107.33	106.53	0.80
8	117.00	117.00	0.00
9	100.33	101.64	-1.30

Model is: (Young leaves before irrigation.sta)
Dep. Var. : R

Figure 4. *Measured, and predicted values and residuals for R_{avg} in young leaves before irrigation*
Source: *compiled by the author*

Figure 5 also demonstrates that all the experimental data points fall within or very close to the narrow confidence interval (at a confidence level of γ=0.95)

and that the regression line passes close to the experimental points:

Predicted = 3.2895+0.9697*x; 0.95 Conf.Int
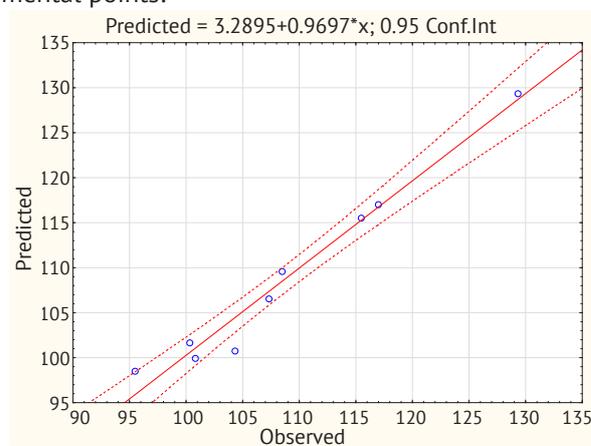
Figure 5. *Scatter plot – predicted versus measured values of colour component Ravg at young leaves before irrigation, model (2)*
Source: *compiled by the author*

Figure 6 demonstrates that the residual values are within acceptable limits and with a uniform distribution. The high value of $R^2=0.99$ for R_{avg} in young leaves after irrigation indicates that the colour component is significantly dependent on the variables included in the model. The points are located close to the straight line, i.e. it is possible to assume that the residuals have a normal distribution. The coefficients of the models regarding the colour components of old leaves before irrigation are summarized in Table 3:
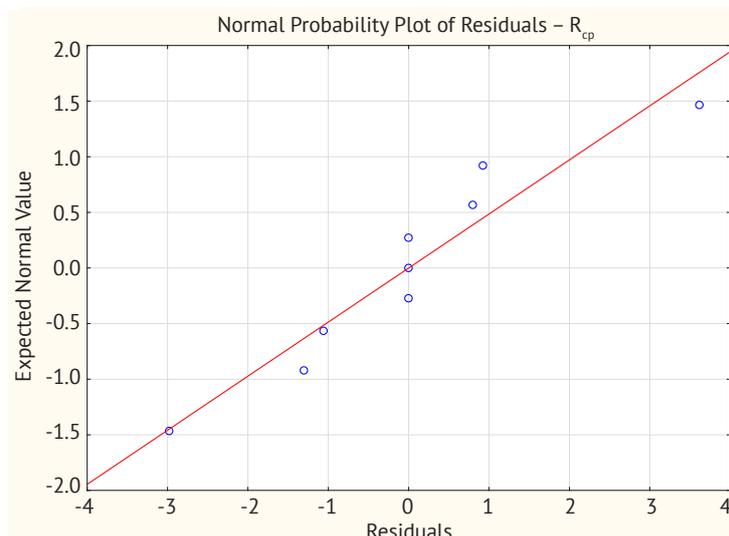
Figure 6. *Normal probability plot of the residuals of model (2) for R_{avg} at young leaves before irrigation*
Source: *compiled by the author*

Table 3. *Coefficients and breakpoint for R_{avg}, G_{avg}, B_{avg}, H_{avg}, S_{avg} and L_{avg} models, old leaves before irrigation*

Model variable	Coefficients	R_{avg}	G_{avg}	B_{avg}	H_{avg}	S_{avg}	L_{avg}
Constant	b_{01}	122.18	169.45	122.02	0.288	0.092	0.146
θ	b_{11}	0.53	1.16	0.42	0.0005	-0.005	0.0005
T	b_{21}	0.02	-2.16	-1.25	-0.0033	0.009	-0.0016
Constant	b_{02}	-785.96	-512.82	-222.52	0.374	0.283	-0.385
θ	b_{12}	2.89	2.96	4.80	-0.0005	-0.0034	0.0064
T	b_{22}	44.99	31.73	11.75	-0.0063	-0.0007	0.0192
	breakpoint	139.15	158.89	112.37	0.239	0.172	0.133
R		0.93	0.95	0.97	0.88	0.96	0.95
Variations explained:		87.42%	91.13%	93.54%	77.10%	92.82%	89.96%
R^2		0.87	0.91	0.94	0.77	0.93	0.90%

Source: *compiled by the author*

Verification of the operability of the developed model (2). Next, the performance of model (2) was tested, examining θ as the dependent variable and the colour component and soil temperature as the independent variable. The goal is for the farmer to determine the plant colour and know whether to irrigate (i.e. how wet it is – an indirect smart method). To this end, RGB and HSL values were taken from additional control plants before irrigation (Plants 10-12) during the experiment to verify the need for irrigation. With their help, the adequacy of the model (2) was tested. Analytically, model (2) transforms concerning θ as follows, yielding model (3):

$$\theta = (\{R_{avg}, G_{avg}, B_{avg}, S_{avg}, L_{avg}\} - b_{01} - b_{21}T)/b_{11}) \text{ (for } \{R_{avg}, G_{avg}, B_{avg}, S_{avg}, L_{avg}\} \leq \text{breakpoint } b_0)$$
$$\text{Or } (\{R_{avg}, G_{avg}, B_{avg}, S_{avg}, L_{avg}\} - b_{02} - b_{22}T)/ b_{12}) \text{ (for } \{R_{avg}, G_{avg}, B_{avg}, S_{avg}, L_{avg}\} > \text{breakpoint } b_0), \tag{3}$$

The results of model (3) testing are presented in Table 4:

Table 4. *Approbating the model (3) in young leaves before irrigation*

	R_{avg}	G_{avg}	B_{avg}	S_{avg}	L_{avg}	$\theta_{measured}$	T
Plant 10	108.83	131.67	81.50	0.236	0.104	22.00	19.03

Table 4. Continued

	R_{avg}	G_{avg}	B_{avg}	S_{avg}	L_{avg}	$\theta_{measured}$	T
$\theta_{calculated}$	23.32	20.93	25.30	26.13	21.65		
Predict. error	+5.99%	-4.85%	+14.98%	+18.88%	+13.77%		
Plant 11	118.67	146.33	81.67	0.284	0.111	23.21	18.96
$\theta_{calculated}$	24.63	24.66	25.13	23.63	17.22		
Predict. error	+6.14%	+6.23%	+8.28%	+1.82%	-9.20%		
Plant 12	110.50	133.50	80.67	0.246	0.105	21.45	18.89
$\theta_{calculated}$	24.01	20.09	20.37	21.98	20.90		
Predict. error	+11.92%	-6.36%	-5.05%	+2.42%	+10.66%		

Source: compiled by the author

During the approbating of the model (3) on older leaves before irrigation of large soil, moisture prediction errors (between -66% and 45%) were obtained, therefore they were not demonstrated in this study. Therefore, the older leaves are not suitable for determining the need for irrigation. Young leaves before irrigation are more informative because the error is acceptably small, confirming the result's accuracy.

Description of the components of the automated irrigation system. The model of the remote sensing system for monitoring the tomato crop in the greenhouse and automatic control of the drip irrigation system (irrigation start and stop) consists of the following elements (Fig. 7):

1) Solar panel for autonomous power supply of the camera (40 W); 2) Autonomous 4G IP camera (dust and waterproof); 3) micro-Sim and microSD cards; 4) Protection of the camera lens against condensation (from the high humidity in the greenhouse); 5) Rodent proof cables; 6) Accumulator (20 Ah); 7) Drip irrigation system; 8) Crops greenhouse-grown indeterminate tomato plants; 9) A remote computer executing the control function embedded in a program written in Python; 10) A remote server where a snapshot is stored (e.g. every half hour).
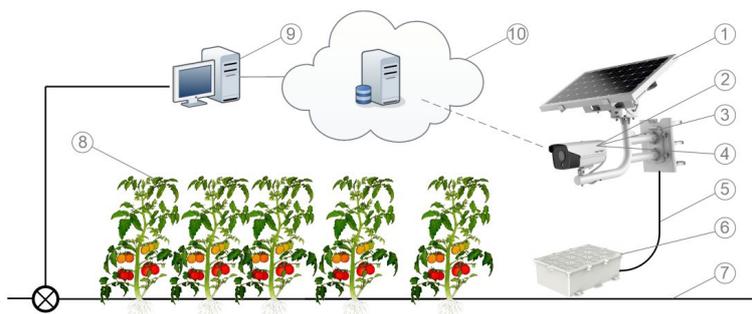
Figure 7. *A model of a remote sensing system for crop monitoring and automated irrigation*
Source: compiled by the author

The camera can be set to automatically take a picture of the crop at a certain time interval (for example half an hour during the daylight hours) through the web interface of its software. This means that to save energy, the camera does not need to be on at all times, but only during the day every half an hour. The cropped photo can be stored in two ways. The first is locally, on the camera's microSD card. The second: is by installing a micro-SIM card in the designated slot, the camera can be connected to the GSM cellular network and send the photos at a certain interval to a cloud storage service or a remote server. Connecting to the GSM cellular network is done by configuring the correct network settings: e.g. through the web interface of the camera, the name of the access point - APN (Access point name) must be set, among other appropriate network settings:

username, password, IP address, subnet mask, gateway, and DNS (if the cellular network requires them). Setting the camera to send the pictures to a cloud service or a remote server is also done through the web interface of the camera. IP address or domain name of the server where the photos will be uploaded, upload interval, file format and image resolution are set. Finally, the photo upload schedule is enabled in the settings.

The image sent to a remote server is automatically retrieved remotely via a script on another computer, processed by a Python-based program, and if the soil moisture calculated from the colour of the leaves is below a certain threshold level, an actuator in the greenhouse is remotely activated to start and subsequently to stop the irrigation system. In the following Figure 8 is presented an example photo of the crop in the greenhouse, taken

during the conduct of the experiments and the acquisition of the experimental data. Field experiments were conducted on row 2 of the image (the right one). If the camera has a wider field of view, more rows can be covered and processed simultaneously.

***Figure 8**. Photo obtained from the camera monitoring the crops in the greenhouse*
***Source:** compiled by the author*

***Description of the image extraction and processing procedure and the program**.* Python is a powerful and versatile programming language that has become increasingly popular among *scientists and researchers* in recent years. Python can be used to create compact and simple code whereas other programming languages require more resources and coding time.

Irrespective of the scientific domain, Python proves valuable for researchers in data analysis, visualization, and model development to advance their studies. The developed program focuses on image processing and soil moisture computation for two rows of tomatoes, subsequently regulating irrigation in one of the rows (specifically, row 2 on the right, where the field experiments were conducted) based on the calculated values. The program operates through the following sequence: Initially, it imports essential modules or libraries, such as OpenCV (cv2) for image processing, urllib.request for fetching an image from a URL, numpy for array manipulation, and serial for establishing serial communication with Arduino.

The process begins by retrieving an image from a URL and converting it into a numpy array using

OpenCV's `cv2.imdecode()` function. Subsequently, the program generates two binary masks for distinct regions of interest in the image, specifically the areas corresponding to young leaf regions. The polygonal points defining these regions are manually set during pre-calibration. The `cv2.fillConvexPoly()` function is called to delineate irregular quadrilaterals around each region (refer to Fig. 9). Finally, `cv2.bitwise_and()` is used to apply the masks to the image, resulting in a masked image that exclusively displays the two regions of interest (Fig. 9).

***Figure 9**. Software-masked image showing only the two young leaf regions*
***Source:** compiled by the author using the author-developed program*

Subsequently, a scaling factor is computed to adjust the size of the masked image, ensuring it is 1080 pixels tall – the standard resolution height of a modern monitor. The resizing operation is executed through the `cv2.resize()` function. The program then showcases the resized masked image using `cv2.imshow()`, awaits a key press with `cv2.waitKey()`, and saves the masked image locally via `cv2.imwrite()`. Next, the program counts non-zero pixels in each mask (provided for informational purposes) using `cv2.countNonZero()` and determines the mean RGB colour values for each region using `cv2.mean()`.

For the two rows, the program computes the soil moisture value based on their respective mean green colour values using the non-linear model (3), which forms the core of the control law. This model serves as a

regulator, influencing the actuator to initiate and cease irrigation. According to the assessments in Table 4, the G_{avg} colour component is chosen for soil moisture calculation and irrigation management. This colour component yields the smallest errors in soil moisture prediction (ranging from -6% to 6%, contingent on whether G_{avg} surpasses or falls below the breakpoint). Any prediction errors are rectified in the soil moisture calculation, along with an adjustment factor for the specific soil type (in this instance, the resulting soil moisture readings must be multiplied by a coefficient of 1.1). Ultimately, the program outputs the results of the soil moisture calculation and the irrigation decision for user inspection. Communication with the Arduino board is facilitated using the serial module, transmitting a message to initiate or halt irrigation in the two monitored rows of plantations based on the computed soil moisture value. The program's source code is available on GitHub in the "living-biosensors" repository at this link (n.d.).

Results after the program implementation (Program output)

```
Row 1:
 Mean color values (RGB): [118.79, 142.88, 93.75]
 Number of pixels: 2321756
 Average Soil Moisture: 25.03
 Soil moisture is sufficient in Row 1. Irrigation
not started.
 Row 2:
 Mean color values (RGB): [109.1, 133.79, 83.38]
 Number of pixels: 1926333
 Average Soil Moisture: 23.45
 Soil moisture is sufficient in Row 2. Irrigation
not started.
```

The precision and accuracy of operation of the described manual measurements with a calibrated colourimeter meter were unambiguously confirmed by the calculations with an image processing program, which also confirms the correctness of the calculation methods used, work methods and the sufficiency of the sample used, as shown in the next comparative Table 5:

Table 5. *Comparison of the obtained average RGB values and the soil moisture in the manual measurements and with the help of a program*

Measurement method	Number of measurements	RGB averages	Soil moisture
With hand-held measuring devices	54	[109, 133, 77]	23.85
From photo and program	1 926 333 pixels	[109, 134, 83]	23.45

Source: *compiled by the author*

The average value of the colour components of young leaves before irrigation obtained from several dozen manual measurements with a calibrated colourimeter and obtained from nearly 2 million pixels in row 2 is almost the same. Only the B component diverges, but the accuracy is still above 92%, while R and G it is above 99%. Compared with the handheld measurements, the program predicts the soil moisture with 98% accuracy and activates the actuators if watering is needed. Python allows an automatic program execution, e.g. every half hour using task scheduling tools such as cron (on Unix-based systems) or Task Scheduler (on Windows) to schedule the script to run automatically at set intervals.

Remote access to the greenhouse irrigation system and its start and stop can be done using a GSM card for sending commands via SMS messages and a device that can receive and execute SMS commands (Arduino or other controllers with a GSM receiver). Arduino is a popular microcontroller platform that can be used to control various devices, including irrigation systems. There are various other controllers on the market such as Raspberry Pi, BeagleBone and other microcontroller boards that can be used to control greenhouse irrigation systems. An in-depth state-of-the-art review (Atanasov, 2023) confirmed what was found in previous research (Atanasov, 2021), that there is no information that leaf colour has been used for predictive models of soil moisture, including in the model, as well as soil temperature usage as a dependent factor. Similar studies conducted demonstrate the following results:

Q. Li *et al.* (2022) investigated the interconnection between soil water content, different stages of wheat drought stress, canopy temperature, and spectral response characteristics. C. Ru *et al.* (2020) explored the effectiveness of the Crop Water Stress Index, which relies on leaf temperature, as a tool for assessing the water status of grapevines. N. Chandel *et al.* (2022) introduced a non-invasive approach utilizing computer vision and thermal-RGB imagery to detect water stress in winter wheat crops. Their method incorporated deep learning and function-approximation models to classify crops based on stress levels, utilizing thermal-RGB images and various input parameters. P. López-García *et al.* (2022) conducted experiments to evaluate the impact of different water qualities and initiation times for irrigation on crop growth. They employed Artificial Neural Network models trained on multispectral and RGB data to simulate and analyse water stress in crops. J. Fernández-Novales *et al.* (2021) focused on evaluating the water status of grapevines using measurements of leaf water potential and canopy temperature. They created spatial-temporal maps to examine changes in water status over time and across different locations.

R. Dhillon *et al.* (2019) developed a methodology for forecasting Plant Water Stress (PWS) in almond and walnut trees using a continuous leaf monitoring system. By

gathering data on leaf temperature and environmental factors with the Leaf Monitor system, they achieved a significant average reduction of 39% in water usage compared to the fixed 100% evapotranspiration replacement irrigation approach across all trees. Analysing the colour of strawberry leaves proves to be an effective method for assessing soil conditions and protecting strawberry crops from excessive environmental nutrients, which can lead to financial losses (Madhavi *et al.,* 2022). Their primary goal was to develop machine learning models, specifically multiple linear regression (MLR) and gradient boost regression (GBR), capable of simulating variations in the colour of strawberry leaves. To achieve this, they gathered precise measurements of soil physicochemical properties in the rooting zones from the largest and most diverse coloured leaves of the strawberry plants, using a multifunctional soil sensor. Additionally, they collected 400 strawberry leaflets at various stages of vegetative and reproductive growth, capturing individual leaves using a digital imaging system. Ultimately, their findings indicated that the GBR model outperformed the MLR model, achieving superior performance metrics. F. Hahn *et al.* (2021) focused on mango trees, utilizing dendrometers and capacitors equipped with Teflon clamps to measure Leaf Water Content and investigate the impact of water availability on mango production.

H. Skoneczny *et al.* (2020) explored the potential of non-invasive proximal hyperspectral remote sensing to differentiate between the leaves of apple trees in three states: healthy, infected, and dry. This distinction was achieved by utilizing spectral bands and indices. J. Rodriguez-Perez *et al.* (2018) compared two regression models, ordinary least squares regression and functional linear regression, to predict Leaf Water Content in grapevines using reflectance data and specific wavelengths. T. Zhao *et al.* (2020) employed hyperspectral imaging to assess the water status in tomato leaves, calculating vegetation indices to evaluate Leaf Water Content in various parts of the tomato plants. These studies underscore the promise of advanced techniques in monitoring Plant Water Stress and overall plant health, offering potential benefits for enhanced crop management.

CONCLUSIONS

A significant discovery emerged, highlighting the significance of the breakpoint's physical meaning – it aligns with the arithmetic mean value of the specified colour component. The regression analysis produced two noteworthy non-linear models, namely model (2) and the piecewise linear regression with breakpoint technique. These models demonstrate remarkable accuracy in predicting colour component values based on soil moisture and temperature, as evidenced by the high coefficient of determination (R^2) achieved. The confirmation of a non-linear relationship between the studied dependent variable (colour component) and the two independent variables (θ and T) has been convincingly established.

Model (3) and the piecewise linear regression with breakpoint method prove to be effective in accurately predicting soil moisture values, particularly in young leaves, based on leaf colour and soil temperature. The low error rates in forecasting (+-6%) underscore the suitability of young leaves in determining the need for watering. This validates the methodology of sampling twice as many RGB colour samples from young leaves during the research compared to older leaves, which are found to be unsuitable for assessing watering needs due to less informative results.

Models (2) and (3) form the foundation for the development of an automated crop monitoring system designed to assess watering requirements and control the irrigation system automatically. The precision and accuracy of manual measurements using a calibrated colourimeter are unequivocally confirmed through calculations with an image-processing program. This further affirms the accuracy of the employed calculation methods, work procedures, and the sufficiency of the sample size.

The average colour component values of young leaves before irrigation, derived from both manual measurements with a calibrated colourimeter and analysis of nearly 2 million pixels in the row where field experiments were conducted, align closely. While there is a slight divergence in the B component, the accuracy remains above 92%, with R and G components exceeding 99%. The Python-based program, with a 98% accuracy rate, predicts soil moisture and triggers actuators if there is a need for watering, highlighting the effectiveness of the developed system.

The study results may not solve all subsequent engineering and other problems, as the abovementioned methods do not work in the following situations: when due to diseases, the colour of the leaves of the plants changes; when the plants grow, the camera should change its tilt and/or move up automatically; evening time, when there is no natural light, logically, measurements and calculations cannot be carried out. While the primary application of the study is in terrestrial agriculture, the principles and technologies developed and the ability to predict and manage soil moisture using plant-based sensors can be adopted in outer space exploration and farming. Automated irrigation systems, as described in the paper, could be adapted for use in controlled environments on spacecraft or extraterrestrial bases, ensuring optimal conditions for plant growth.

CONFLICT OF INTEREST

None.

REFERENCES

[1] Araújo-Paredes, C., Portela, F., Mendes, S., & Valín, M.I. (2022). Using aerial thermal imagery to evaluate water status in *Vitis vinifera* cv. Loureiro. *Sensors*, 22(20), article number 8056. doi: 10.3390/s22208056.

[2] Atanasov, S. (2021). Methodology for irrigation water uptake time estimation based on RGB colorimetric measurements of leaves (A visual-graphical observation). *IOP Conference Series: Materials Science and Engineering*, 1031(1), article number 012016. doi: 10.1088/issn.1757-899X.

[3] Atanasov, S. (2023). State-of-the-art technologies for remote sensing of crops water status and nutrients in agriculture: A review. *Scientific Horizons*, 26(9), 167-177. doi: 10.48077/scihor9.2023.167.

[4] Chandel, N.S., Rajwade, Y.A., Dubey, K., Chandel, A.K., Subeesh, A., & Tiwari, M.K. (2022). Water stress identification of winter wheat crop with state-of-the-art AI techniques and high-resolution thermal-rgb imagery. *Plants*, 11(23), article number 3344. doi: 10.3390/plants11233344.

[5] Dhillon, R., Rojo, F., Upadhyaya, S.K., Roach, J., Coates, R., & Delwiche, M. (2019). Prediction of plant water status in almond and walnut trees using a continuous leaf monitoring system. *Precision Agriculture*, 20, 723-745. doi: 10.1007/s11119-018-9607-0.

[6] Fernández-Novales, J., Saiz-Rubio, Tardaguila, J., Valente, J., Barrio, I., Rovira-Más, F., Cuenca-Cuenca, A., Santos Alves, F., & Diago, M.P. (2021). Monitoring and mapping vineyard water status using non-invasive technologies by a ground robot. *Remote Sensing*, 13(14), article number 2830. doi: 10.3390/rs13142830.

[7] GitHub. (n.d.). Retrieved from https://github.com/SSAtanasov/living-biosensors.

[8] Gnatowski, T., Szatyłowicz, J., Pawluśkiewicz, B., Oleszczuk, R., Janicka, M., Papierowska, E., & Szejba, D. (2018). Field calibration of TDR to assess the soil moisture of drained peatland surface layers. *Water*, 10(12), article number 1842. doi: 10.3390/w10121842.

[9] Hahn, F., Espinoza, J., & Zacarías, U. (2021). Mango leaf monitoring with inductive and capacitive sensors and its comparison with trunk dendrometer measurements. *Engineering Proceedings*, 9(1), article number 28. doi: 10.3390/engproc2021009028.

[10] Kang, S., van Iersel, M.W., & Kim, J. (2019). Plant root growth affects FDR soil moisture sensor calibration. *Scientia Horticulturae*, 252, 208-211. doi: 10.1016/j.scienta.2019.03.050.

[11] Laureti, T., Benedetti, I., & Branca, G. (2021). Water use efficiency and public goods conservation: A spatial stochastic frontier model applied to irrigation in Southern Italy. *Socio-Economic Planning Sciences*, 73, article number 100856. doi: 10.1016/j.seps.2020.100856.

[12] Li, Q., Gao, M., & Li, Z.L. (2022). Ground hyper-spectral remote-sensing monitoring of wheat water stress during different growing stages. *Agronomy*, 12(10), article number 2267. doi: 10.3390/agronomy12102267.

[13] López-García, P., Intrigliolo, D., Moreno, M.A., Martínez-Moreno, A., Ortega, J.F., Pérez-Álvarez, E.P., & Ballesteros, R. (2022). Machine learning-based processing of multispectral and RGB UAV imagery for the multitemporal monitoring of vineyard water status. *Agronomy*, 12(9), article number 2122. doi: 10.3390/agronomy12092122.

[14] Madhavi, B.G.K., Basak, J.K., Paudel, B., Kim, N.E., Choi, G.M., & Kim, H.T. (2022). Prediction of strawberry leaf color using RGB mean values based on soil physicochemical parameters using machine learning models. *Agronomy*, 12(5), article number 981. doi: 10.3390/agronomy12050981.

[15] Reich, M., Mikolaj, M., Blume, T., & Güntner, A. (2021). Field-scale subsurface flow processes inferred from continuous gravity monitoring during a sprinkling experiment. *Water Resources Research*, 57(10), article number e2021WR030044. doi: 10.1029/2021WR030044.

[16] Rodriguez-Perez, J.R., Ordóñez, C., González-Fernández, A.B., Sanz-Ablanedo, E., Valenciano, J.B., & Marcelo, V. (2018). Leaf water content estimation by functional linear regression of field spectroscopy data. *Biosystems Engineering*, 165, 36-46. doi: 10.1016/j.biosystemseng.2017.08.017.

[17] Ru, C., Hu, X., Wang, W., Ran, H., Song, T., & Guo, Y. (2020). Evaluation of the crop water stress index as an indicator for the diagnosis of grapevine water deficiency in greenhouses. *Horticulturae*, 6(4), article number 86. doi: 10.3390/horticulturae6040086.

[18] Serrano-Finetti, E., Castillo, E., Alejos, S., & Hilario, L.L. (2023). Toward noninvasive monitoring of plant leaf water content by electrical impedance spectroscopy. *Computers and Electronics in Agriculture*, 210, article number 107907. doi: 10.1016/j.compag.2023.107907.

[19] Simbeye, D.S., Mkiramweni, M.E., Karaman, B., & Taskin, S. (2023). Plant water stress monitoring and control system. *Smart Agricultural Technology*, 3, article number 100066. doi: 10.1016/j.atech.2022.100066.

[20] Skoneczny, H., Kubiak, K., Spiralski, M., Kotlarz, J., Mikiciński, A., & Puławska, J. (2020). Fire blight disease detection for apple trees: Hyperspectral analysis of healthy, infected and dry leaves. *Remote Sensing*, 12(13), article number 2101. doi: 10.3390/rs12132101.

[21] Starr, J.L., & Paltineanu, I.C. (2002). *Methods for measurement of soil water content*. Retrieved from https://www.researchgate.net/publication/283997019_Methods_for_measurement_of_soil_water_content_Capacitance_devices.

[22] Wood, W.W., & Cherry, J.A. (2021). Food security and inaccurate quantification of groundwater irrigation use. *Groundwater*, 59(6), 782-783. doi: 10.1111/gwat.13122.

[23] Zahoor, S.A., *et al.* (2019). Improving water use efficiency in agronomic crop production. In *Agronomic Crops*. Singapore: Springer. doi: 10.1007/978-981-32-9783-8_2.

[24] Zhao, T., Nakano, A., Iwaski, Y., & Umeda, H. (2020). Application of hyperspectral imaging for assessment of tomato leaf water status in plant factories. *Applied Sciences,* 10(13) article number 4665. doi: 10.3390/app10134665.

Автоматизована система дистанційного зондування для моніторингу посівів та управління зрошенням, що базується на зміні кольору листя та кусково-лінійних регресійних моделях для прогнозування вмісту вологи в ґрунті

Светослав Атанасов

Аспірант, магістр комп'ютерних систем та технологій

Тракійський університет

6015, Студентське містечко, м. Стара Загора, Болгарія

https://orcid.org/0000-0002-2658-1611

Анотація. Рослини можуть слугувати біологічними сенсорами доти, доки ми можемо правильно інтерпретувати їхні «показання» та зворотній зв›язок, який вони надають нам через зміну кольору їхнього листя. Метою даного дослідження було прогнозування вологості ґрунту і, відповідно, потреби в зрошенні на основі створених нелінійних математичних моделей, що описують взаємозв'язок між компонентами колірних моделей RGB і HSL та вологістю і температурою ґрунту. Ці нелінійні математичні моделі базуються на кусково-лінійній регресії з точкою розриву і прогнозують вологість ґрунту за допомогою колірних компонентів і температури ґрунту з похибкою +-6 %. Створено систему автоматизованого поливу та написано програму керування нею, де основним законом керування є створені нелінійні кусково-лінійні моделі. Система автоматизованого управління зрошенням включає підсистему дистанційного моніторингу стану посівів та підсистему управління зрошенням. Програма обробляє фото, отримане з камери, і активує виконавчі механізми, якщо є потреба в поливі. У порівнянні з ручним збором даних у першій частині дослідження, програма розраховує на основі зображення середні значення RGB-моделі в досліджуваному ряду томатних плантацій з точністю понад 99 % для R- і G-компонентів і понад 92 % для B-компонента. Програма також прогнозує вологість ґрунту з точністю 98 %. Практичне значення роботи з точки зору водозбереження полягає в розробці програмно-керованої автоматизованої системи зрошення, яка використовує рослини як біологічні сенсори, застосовуючи нелінійні математичні моделі, засновані на зміні кольору листя, для точного прогнозування вологості ґрунту

Ключові слова: біосенсори; точне зрошення; RGB-колориметрія; обробка зображень; оцифрування; біоінформатика