Turkish Journal of Engineering
https://dergipark.org.tr/en/pub/tuje
e-ISSN 2587-1366

Analysis of the effect of environmental data collected through the WSN on plant development by the WEKA program

Muhammed Furkan Koşum[*1], **Ali Osman Özkan** [2]

[1]Başkent University, Electronics and Automation, 06980, Ankara, Turkey, mfurkankosum@baskent.edu.tr
[2]Necmettin Erbakan University, Electrical and Electronics Engineering, 42005, Konya, Turkey, alozkan@erbakan.edu.tr

Cite this study: Koşum, M. F., & Özkan, A. O. (2025). Analysis of the Effect of Environmental Data Collected Through the WSN on Plant Development by the WEKA Program. Turkish Journal of Engineering, 9(2), page 202-210.

https://doi.org/10.31127/tuje.1532313

Keywords
Digital Agriculture
Regression Analysis
Water Consumption
WEKA Program
Wireless Sensor Network

Research Article

Received: 12.08.2024
Revised: 28.08.2024
Accepted: 16.09.2024
Published:01.04.2024

check for updates

Abstract

The increasing world population, along with rising water usage for personal, industrial, and agricultural purposes, and climate change issues, have caused water problems in many regions. Unconscious irrigation and the inability to predict plant water needs lead to water waste and yield losses. Innovative technologies are needed to improve efficiency and better manage water resources. One solution is digitalization in agriculture. This study created a greenhouse environment where sensors collected environmental data (soil moisture, temperature, humidity, and light) and transferred it to a farm management system. A gateway was established to enable data exchange between the sensor and actuator nodes and the management system over the network. Instant and hourly data were recorded in the database through the gateway. If the instant value transferred to the database was below the threshold value, automatic irrigation was performed by the actuator. The amount of water used was recorded in the database. The collected data were analyzed with classification and prediction algorithms using the WEKA program. The impact of environmental data on water consumption and plant height was transformed into a mathematical equation.

1. Introduction

With the increasing global population, the demand for food is also rising day by day. Simultaneously, the increase in the amount of water used for personal use, industrial, and agricultural production, coupled with climate change issues in many parts of the world, has created water problems in some regions. Numerous studies have been conducted and continue to be conducted to prevent the wastage of water resources and to use water efficiently [1]. Unconscious irrigation in agriculture and the inability to predict the water needs of plants lead to both water wastage and yield losses.

Innovative ideas and technologies are needed to increase efficiency and better utilize water resources. One of the methods that can meet these needs is digitalization in agriculture. Digital agriculture methods will be effective in increasing productivity, reducing labor, environmental pollution, water wastage, and the chemicals used in agricultural spraying, among other things [2]. Digital agriculture is the application of communication technologies to the modern agricultural industry to meet these needs and sustainably provide food requirements.

In modern agricultural systems, environmental data is collected by sensors and transferred to the farm management system, where necessary actions are taken by actuators. Therefore, a network of sensors and actuators that accurately directs data packets and provides multiple communications is needed.

Wireless sensor and actuator networks consist of many sensor/actuator nodes containing various sensors and actuators, and a central exit node to transmit the collected data to the server. There is a gateway to facilitate network exchange between the sensor and actuator nodes and the management system. This gateway, with its practical structure compared to wired systems, can respond to the needs of farm management at a lower cost and more quickly. In this respect, the use of wireless sensor and actuator networks in agricultural applications is suitable [1-3].

The collected data can be analyzed with classification and prediction algorithms, and the impact of certain environmental data – soil moisture (Mois.), ambient temperature (Temp.), ambient humidity (Humi.), and ambient light (Light) – on water consumption (Water) and plant height (Height) can be transformed into a mathematical equation. Thus, it can be predicted which parameter has the most significant effect on water consumption and plant height.

1.1. Wireless sensor networks

Wireless sensor networks are networks composed of many intelligent devices, called nodes, deployed in a specific study area to perform a task. These networks are designed to monitor, record, and organize data collected on physical and environmental magnitudes such as temperature, sound, pollution levels, humidity, wind, etc., at a central location [4-6]. Thanks to advancements in micro-electro-mechanical systems technology, wireless communications, and digital electronics, it has become possible to develop small-sized, low-cost, low-power, multifunctional sensor nodes capable of wireless communication over short distances [7]. Sensor nodes themselves can act as independent sources, collecting data from their surroundings. Additionally, they can receive data from other nearby sensor nodes and transmit it to the sink node, as shown in Figure 1. The sink node acts as a gateway and transmits the collected data to the user at a remote-control center via wired (internet, etc.) or wireless (GSM, satellite, etc.) communication [8].

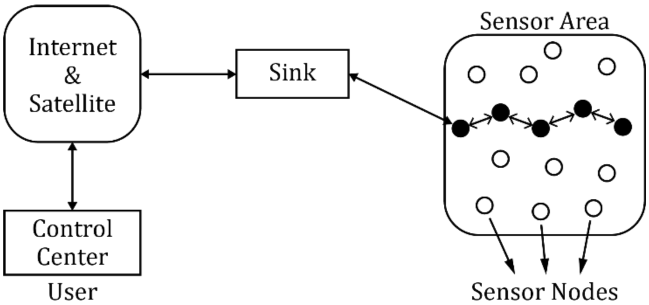

Figure 1. Example sensor node scenario distributed across the work area

Wireless sensor networks enable the execution of automatic sensing, embedded computing, and wireless networking tasks in small, inexpensive, and low-power devices. When connected via a wireless network, sensor nodes, which possess a strong and flexible structure, have quite limited individual capabilities. However, by collaborating with each other, they can easily perform distributed sensing tasks over a wide working area. They can be deployed either in a planned or random manner. The number of required sensor nodes varies depending on the size of the area being studied and the sensitivity required by the application [9].

1.2. Modern agriculture studies with wireless sensor networks

A wireless sensor network is a convenient system for sensing and transmitting physical environmental parameters. For efficient agriculture, it is important to monitor and control data such as soil moisture, soil temperature, soil electrical conductivity, and external environmental parameters. Therefore, wireless sensor networks are preferred in applications such as agricultural irrigation, greenhouse cultivation, beekeeping, and plant protection [1-2]. Technological and innovative exemplary studies have been conducted to prevent the wastage of water resources and to ensure efficient production based on knowing the water needs of plants [10-13]. By periodically monitoring soil and crop data through wireless sensors, irrigation systems can be developed to prevent unnecessary water consumption and achieve efficient crop yield [14-17]. The use of wireless sensor networks in greenhouses makes activities such as fertilization, irrigation, monitoring, and spraying more efficient and less costly [3, 18-20].

1.3. WEKA program

WEKA (The Waikato Environment for Knowledge Analysis) is an open-source machine learning software developed under the General Public License at the University of Waikato in New Zealand. The program is named after the WEKA bird, which is native to New Zealand. WEKA includes various algorithms for machine learning tasks (data preprocessing, classification, regression, clustering, association rule mining, visualization, etc.). Additionally, users can create their own algorithms with Java code, process datasets, and analyze results within WEKA [21-22].

The Applications framework includes Explorer, Experimenter, Knowledge Flow, Workbench, and Simple CLI applications. The Explorer is the interface where preprocessing and various data mining applications are performed. The Experimenter is the interface used to test various algorithms on a dataset to determine which one yields better results. The Knowledge Flow is an interface where data mining algorithms are applied by creating data flow diagrams using a drag-and-drop mechanism. The Simple CLI is a console interface where data mining applications can be executed by writing Java code [21-22].

All algorithms in WEKA can take input in the Attribute-Relation File Format (ARFF), which is a simple relational table format used in machine learning. WEKA can also use CSV and C4.5 file formats as input [23].

Data entered in the ARFF format are expressed as character strings at the programming level. The terms @relation, @attribute, and @data are used to define the data structure. The @relation term defines the general purpose or name of the data. The @attribute term defines the attribute names corresponding to the columns in the database. The @data term indicates the row where the raw data begins [24].

1.3.1. Linear regression algorithm

The Linear Regression algorithm is a method for formulating the relationship between a numerical dependent variable and one or more independent variables. The Linear Regression algorithm enables the prediction of the future value of the dependent variable based on the entered data of the dependent and independent variables [25].

1.3.2. Gaussian processes regression algorithm

The Gaussian Processes Regression algorithm is an effective machine learning algorithm for modelling quantitative variables and making classification predictions. Gaussian process models determine the average distance between points with a kernel-based probability function to predict the value of an unknown point from learning data [26].

1.3.3. Multi-layer perceptron algorithm

The Multi-Layer Perceptron algorithm consists of an input layer, an output layer, and one or more hidden middle layers. The input layer receives inputs from the input neurons in the neural network and forwards them to the hidden middle layers. The activity units in the input layer activate the algorithm by connecting to the activity units in the next hidden middle layer. It is preferred for the analysis of nonlinear datasets [27].

1.3.4. Sequential minimal optimization regression algorithm

The Sequential Minimal Optimization Regression algorithm is based on Support Vector Machine (SVM) and is a suitable model for analyzing regression problems. In 1998, Smola and Scholkopf proposed a repetitive algorithm called Sequential Minimal Optimization (SMO) to analyze regression problems using SVM. Shevade and Keerthi successfully developed the SMOreg algorithm, which Smola and Scholkopf defined as SMO for analyzing regression problems. The SMOreg algorithm further enhances the regression value in SMO, consequently providing more efficient and better performance [28].

1.3.5. Multi search algorithm

The Multi Search algorithm searches for a random number of variables without requiring optimization of two variables each time like other algorithms. It then separates the most suitable pair for real selection and learning.

1.3.6. M5Rules algorithm

The M5Rules algorithm prepares decision rules using the separate and conquer method for regression problems. In each iteration, it constructs a model tree using M5. It converts the most suitable leaf into a rule and analyses other examples in the dataset according to this rule [29].

2. Method

2.1. Network design

The proposed network design consists of a sensor node, an advanced node, a database, and a mobile application. The sensor node is responsible for collecting environmental data with various sensors attached to it. The collected data is transmitted to the database instantly via the Wi-Fi unit within the wireless coverage area. Data has been recorded in the database every hour during the operation period. Real-time data tracking has been conducted through a mobile application designed for information exchange with the database. The operational status of the actuator in the advanced node has been adjusted actively, passively, or automatically via the mobile application interface, and the operational status of the actuator has been transferred to the database. The operational status of the actuator has been retrieved from the database through the Wi-Fi unit in the advanced node. For automatic operation mode, the real-time Mois. value obtained from the database has been compared with pre-programmed threshold Mois. values. Based on the comparison result, the actuator has been activated if the instantaneous value is below the lower threshold value and deactivated when it exceeds the upper threshold value. The amount of water consumed has been monitored with the water flow sensor on the advanced node and transferred to the database. Water consumption has been monitored through the mobile application interface. The network architecture used in the study is illustrated in Figure 2.

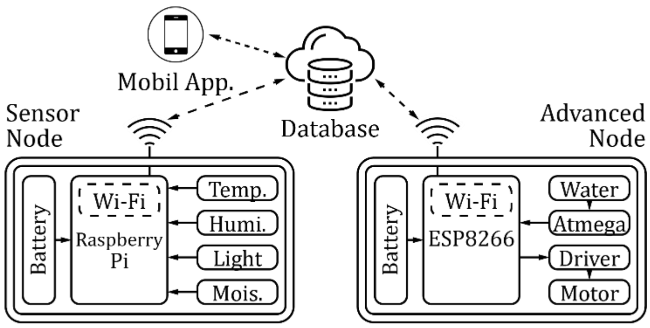

Figure 2. Network architecture used in the study

2.1.1. Sensor node design

The sensor node consists of a microcontroller unit (Raspberry Pi 3 Model B+) with dual-band 2.4 GHz and 5 GHz wireless LAN capabilities compliant with IEEE 802.11 b/g/n/ac standards, an ambient temperature/humidity sensor, a light sensor, a soil moisture sensor, and a power unit.

2.1.2. Advanced node design

The advanced node consists of a wireless transceiver card (ESP8266) operating in the 2.4 GHz band, compliant with IEEE 802.11 b/g/n standards, a water flow sensor (YF-S201), an Atmega328p, an actuator (water pump and driver), and a power unit.

2.1.3. Mobile app design

The mobile application has been developed using the MIT App Inventor mobile application development tool. The application interface consists of five data frames for entering data, three buttons to command the motor, one button to request real-time values, basic visuals, and a background image. When the application screen is opened, real-time Light, Temp., Humi., Mois., and water consumption data are automatically retrieved from the database.

2.1.4. Database

The database has been created using the Firebase database creation platform. Firebase is a web-based platform developed by Google that allows for the creation of real-time databases and data storage for mobile and web applications. The data obtained during the operation is stored in the Firebase database.

2.2. Application design

2.2.1. Research location and features

The research was conducted in Sincan district of Ankara province in 2021. The elevation of the research area is approximately 797 m. The region has a typical continental climate. The annual average total precipitation is 389.1 mm. The highest temperature measured in the region is 40.8 °C in July, while the lowest temperature is -24.9 °C. The annual average temperature is 11.7 °C.

2.2.2. Soil features

The soil structure of the district is brown and grey on the surface, while it is limey and clayey in the lower parts. The alluvial deposits in the upper part are suitable for agriculture. Due to the clayey-loamy structure of the soil, the approximate volumetric percentages are as follows: field capacity is 40%, wilting point is 23%, lime content is 12%, and organic matter content is 0.24%. The effective root depth is 60 cm [30]. The pH value of the soil in the region has been measured to be approximately 7.4 - 7.6 with a TN-200 analogue pH meter and 7.52 with a PH-200 digital pH meter.

2.2.3. Setting up the greenhouse

The research area consists of a 120x100x40 cm plot and a greenhouse roof measuring 180x120x100 cm. The greenhouse is covered with 220-micron UV stabilized greenhouse plastic film. The soil in the greenhouse has been enriched with peat soil, also known as peat, commonly used in greenhouse and seedling cultivation. The pH value of the greenhouse soil has been measured to be approximately 7.2 - 7.5 with a TN-200 analogue pH meter and 7.23 with a PH-200 digital pH meter.

2.2.4. Irrigation water and irrigation system features

In the study, tap water was used. The pH value of the tap water used was measured to be 7.4 with a PH-200 digital pH meter. Watering was done from the water tank by mixing 2.7 ml of nitrogen liquid organic fertilizer per Liter of tap water. The guaranteed content of the liquid organic fertilizer is as follows: 40% organic matter, 18% organic carbon, 2% total nitrogen, 3% soluble potassium oxide, and the pH value is between 4.3 - 6.3. The pH value of the irrigation water given to the plants after mixing was measured to be 6.88 with a PH-200 digital pH meter. A drip irrigation system was used for watering the study area, with 40 cm between drippers.

2.2.5. Soil moisture and ambient data measurement

In the study, soil moisture measurements were conducted using a resistant soil moisture sensor. The sensor probe was placed at a depth of 10 - 15 cm from the soil surface. Instantaneous soil moisture data were obtained and sent to the database via the sensor node. The ambient temperature and ambient humidity information of the greenhouse was obtained by the resistive-type sensor, while the ambient light information was obtained by the light sensor which has a wide spectral range of 200 nm - 400 nm, and both were sent to the database instantly through the sensor node.

2.2.6. Cultivation of plants and cultural processes

Five seedlings of Traditional Pole cucumber (Cucumis sativus), obtained from a seed company, were planted in the working area on July 5, 2021, with a row length of 120 cm and 20 cm spacing between them. The planted seedlings were uniformly watered until July 18, 2021. Two weeks later, the spacing between the seedlings was reduced, leaving 3 sprouted seedlings at intervals of 40 cm in the working area. Starting from July 19, 2021, the seedlings were irrigated using a wireless sensor and actuator network.

The pre-programmed lower threshold Mois. level was set at 50% (equivalent to 23% wilting point). When the instantaneous Mois. level dropped below the wilting point; the water pump was activated to initiate irrigation. When the instantaneous Mois. level exceeded the upper threshold Mois. level of 87% (equivalent to 40% field capacity), the water pump was deactivated to terminate irrigation. Single-stem cultivation was conducted to increase yield [31]. The initial height measurement of the plants was taken on July 20, 2021, and the final height measurement was taken on September 11, 2021. Height measurements were taken seven times during the cultivation period.

In order to see the effect of the wireless sensor and actuator network on the efficient use of water and plant growth, three different saplings were grown between the same dates and irrigated with the traditional agricultural method applying the same cultural processes.

The heights of the seedlings measured according to the dates are shown in Table 1. The study was concluded on September 9, 2021. The last two height measurements of the 2nd sapling in WSN could not be made because the plant sick and died.

Table 1. The Height of the seedlings measured according to the dates

	WSN			Traditional		
Date	1st	2nd	3rd	1st	2nd	3rd
20 July 2021	13.5	12	13.5	11	11	13.5
3 August 2021	52	16	20.5	21	18.5	26
10 August 2021	101	17	25.5	23.5	21	45
16 August 2021	118	22	28	24.5	24.5	58.5
28 August 2021	135	17	29	25	24.5	81
6 June 2021	138	-	29	26	24	85
11 June 2021	136.5	-	27	30	24	80.5

The measurements are given in centimeters.

2.3. Data processing method

The Explorer application interface was used in the study. The Explorer interface consists of preprocess, classify, cluster, associate, select attribute, visualize, and forecast screens.

For the data mining prediction algorithms in WEKA selected for use in the study, Linear Regression (LR), Gaussian Processes Regression (GPR), Multi-Layer Perceptron (MLP), Sequential Minimal Optimization Regression (SMOreg), Multi Search (MS), and M5Rules algorithms were chosen. These six algorithms were deemed suitable for the dataset in the study due to their frequent usage and similarity in previous works.

3. Results

The data obtained from the field studies of the wireless sensor and actuator network were analyzed using the WEKA program, including Mois., Temp., Humi., Light, water consumption, and plant height. Analyses were conducted by calculating the average water consumption and average plant height for a single cucumber plant. The impact of environmental variables (Mois., Temp., Humi., and Light) on water consumption was mathematically expressed. Additionally, the mathematical expression of the impact of environmental variables (Mois., Temp., Humi., and Light) and water consumption on plant height was formulated.

During the study period from 19 July 2021 to 09 June 2021, a total of 53 days of data were collected using the wireless sensor and actuator network, as presented in Table 2.

According to the data collected in Table 2, the average daily Mois. during the study period was approximately 72%, the average daily Temp. was around 32°C, the average daily Humi. was approximately 38%, and the average daily Light was about 57%. The total amount of water consumed for a single cucumber plant was 7416.67 ml, with an average plant height of 54.50 cm observed during the study period. It was observed that the Mois. level was maintained above 50% throughout the study.

In irrigation carried out using traditional agricultural methods in the same cultural processes and on the same dates, the total amount of water consumed for a single cucumber plant was 9999.99 ml, with an average plant height of 44.83 cm observed.

Table 2. Data collected in the study

Day	Mois. (%)	Temp. (°C)	Humi. (%)	Light (%)	Water (ml)	Height (cm)
1	96.67	36.44	27.89	66.00	0.00	13.00
2	97.21	35.29	42.42	65.96	0.00	14.10
3	96.50	34.33	40.88	61.92	0.00	15.20
4	93.46	31.17	45.33	61.63	0.00	16.30
5	89.79	29.96	45.25	61.50	0.00	17.40
6	90.92	29.58	40.46	61.54	0.00	18.50
7	82.83	30.42	37.50	61.50	0.00	19.60
8	74.83	30.13	40.25	61.50	0.00	20.70
9	82.25	31.04	43.13	61.54	0.00	21.80
10	78.83	31.17	34.92	61.67	0.00	22.90
11	75.63	31.33	29.38	61.71	0.00	24.00
12	79.71	32.67	29.83	61.00	500.00	25.10
13	85.21	33.75	33.67	61.29	500.00	26.20
14	78.10	33.75	32.08	61.42	500.00	27.30
15	68.56	32.88	33.50	61.33	500.00	28.40
16	60.00	33.25	32.08	60.25	500.00	29.50
17	50.33	33.67	27.75	59.92	500.00	32.12
18	52.33	33.46	30.38	59.63	1250.00	34.74
19	88.29	34.29	29.42	58.46	1250.00	37.36
20	84.67	31.38	37.46	55.29	1250.00	39.98
21	81.13	29.79	50.67	55.63	1250.00	42.60
22	71.38	30.50	51.92	58.04	1250.00	45.22
23	53.63	30.75	46.00	56.88	1750.00	47.83
24	74.08	28.79	48.25	54.63	1750.00	49.19
25	76.50	29.29	47.63	55.96	1750.00	50.55
26	75.92	29.63	46.33	56.79	1750.00	51.91
27	70.25	29.67	42.71	56.42	1750.00	53.27
28	60.58	30.50	39.21	56.25	2416.67	54.63
29	81.46	30.13	41.29	55.83	2416.67	56.00
30	77.00	32.17	38.04	55.38	2416.67	56.36
31	64.94	32.42	37.58	55.33	3416.67	56.72
32	72.58	31.63	38.13	55.08	3416.67	57.08
33	69.54	31.13	44.04	54.75	3416.67	57.44
34	59.29	30.29	43.88	54.58	4416.67	57.80
35	86.50	32.38	39.46	54.58	4416.67	58.16
36	82.29	31.42	38.96	54.25	4416.67	58.52
37	72.25	31.33	39.12	53.92	4416.67	58.88
38	64.13	32.83	34.25	53.96	5083.33	59.24
39	91.21	34.17	31.71	54.04	5083.33	59.60
40	73.17	33.46	32.96	53.83	5083.33	59.96
41	54.77	35.33	31.33	54.13	5083.33	60.33
42	81.57	36.96	26.54	54.25	5916.67	59.81
43	82.28	36.71	28.67	54.21	5916.67	59.29
44	60.69	35.96	37.42	54.04	5916.67	58.78
45	50.27	34.29	44.67	53.92	6750.00	58.26
46	74.96	31.63	45.75	53.71	6750.00	57.74
47	60.33	29.42	39.13	53.58	6750.00	57.22
48	75.58	28.50	38.13	53.46	7416.67	56.71
49	79.13	29.83	38.13	53.42	7416.67	56.19
50	52.83	28.08	40.75	53.25	7416.67	55.67
51	38.83	26.88	39.42	53.21	7416.67	55.28
52	32.54	29.00	28.46	53.42	7416.67	54.89
53	19.13	31.47	27.53	52.93	7416.67	54.50

The water consumption and plant height graph obtained through irrigation using wireless sensor and actuator network between July 19, 2021, and September 09, 2021, are presented in Figure 3.

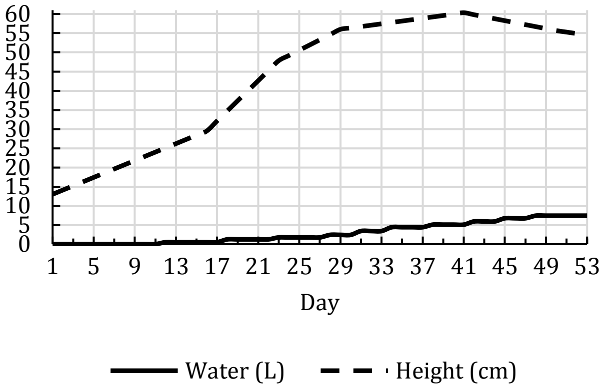

Figure 3. The water consumption and plant height graph (wireless sensor and actuator network)

The water consumption and plant height graph obtained through irrigation using traditional agriculture method between July 19, 2021, and September 09, 2021, are presented in Figure 4.

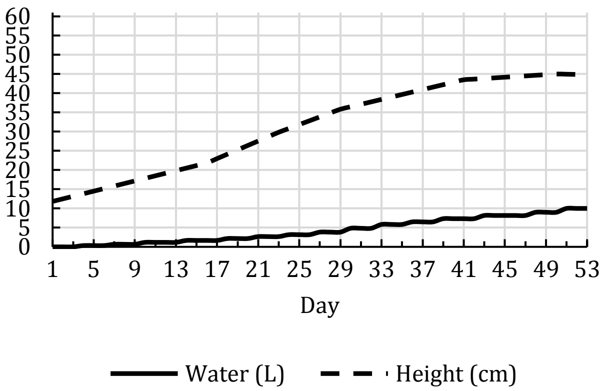

Figure 4. The water consumption and plant height graph (traditional agriculture method)

The regression models were constructed using the dataset obtained from 19 July 2021 to 09 September 2021 in the WEKA program. To compare the regression analysis results with the actual values, the models' Mean Absolute Percentage Error (MAPE) values were examined.

The mathematical expression of the average MAPE value is given in Equation 1 [32-34].

$$MAPE = \frac{100}{N}\sum_{n=1}^{N}\left|\frac{actual\ value - forecast\ value}{actual\ value}\right| \quad (1)$$

3.1. Linear regression algorithm results

The mathematical expressions for the regression equations generated by the Linear Regression algorithm are provided in Equation 2 for Water and in Equation 3 for Height.

$$Water = -21601.4393 + 242.9834 \times Day + 17.0872 \times Mois. + 293.1195 \times Light \quad (2)$$

$$Height = 271.7464 + 0.8326 \times Temp. - 4.453 \times Light \quad (3)$$

The average MAPE values for the estimated Water and Height values based on the equation obtained from the Linear Regression algorithm in the WEKA program are calculated as 19.22% for the Water equation and 10.54% for the Height equation.

3.2. Gaussian processes regression algorithm results

The WEKA program provides estimation results without producing mathematical expressions for the Gaussian Processes Regression algorithm. The average MAPE values for the estimated Water and Height values obtained from the Gaussian Processes Regression algorithm in the WEKA program are calculated as 38.86% for the Water and 13.00% for the Height.

3.3. Sequential minimal optimization regression algorithm results

The mathematical expressions for the regression equations for Water and Height obtained from the Sequential Minimal Optimization Regression algorithm are provided in Equation 4 for Water and Equation 5 for Height. Normalized values are used in the equations.

$$Water = -0.4835 + 1.4917 \times Day + 0.097 \times Mois. - 0.0288 \times Temp. - 0.0365 \times Humi. + 0.2898 \times Light \quad (4)$$

$$Height = 0.5166 + 0.7638 \times Day + 0.0697 \times Mois. + 0.2513 \times Temp. + 0.1657 \times Humi. - 0.9791 \times Light - 0.4214 \times Water \quad (5)$$

The average MAPE values for the estimated Water and Height values obtained from the Sequential Minimal Optimization Regression algorithm are calculated as 16.43% for the Water equation and 8.28% for the Height equation.

3.4. Multi-Layer perceptron algorithm results

The WEKA program provides estimation results without producing mathematical expressions for the Multi-Layer Perceptron algorithm. The average MAPE values for the estimated Water and Height values obtained from the Multi-Layer Perceptron algorithm are calculated as 12.42% for Water and 1.52% for Height.

3.5. Multi search algorithm results

The mathematical expressions for the regression equations for Water and Height obtained from the Multi Search algorithm are given in Equation 6 for Water and Equation 7 for Height.

$$Water = -23962.3973 + 249.875 \times Day + 16.7371 \times Mois. - 0.6716 \times Temp. + 12.9238 \times Humi. + 323.4106 \times Light \quad (6)$$

$$Height = 67.6021 + 1.6731 \times Day + 0.0994 \times Mois. + 1.0691 \times Temp. + 0.295 \times Humi. - 1.8158 \times Light - 0.006 \times Water \quad (7)$$

The average MAPE values for the estimated Water and Height values obtained from the Multi Search algorithm are calculated as 18.43% for the Water equation and 6.91% for the Height equation.

3.6. M5Rules algorithm results

The WEKA program provides estimation results without producing mathematical expressions for the M5Rules algorithm. The average MAPE values for the estimated Water and Height values obtained from the M5Rules algorithm are calculated as 9.28% for Water and 4.75% for Height.

As a result, the comparison of the average MAPE values for Water and Height obtained using six different algorithms employed between July 19, 2021, and September 09, 2021, is presented in Table 3.

Table 3. Average MAPE values of Water and Height with six different algorithms using data between 19 July 2021 and 09 June 2021

	Average MAPE (%)	
Algorithms	Water	Height
LR	19.22	10.54
GPR	38.86	13.00
SMOreg	16.43	8.28
MLP	12.42	***1.52***
MS	18.43	6.91
M5Rules	***9.28***	4.75

4. Discussion

When examining the average MAPE values in Table 3, it is observed that the M5Rules algorithm provides the best result for the Water data, while the MLP algorithm yields the best result for the Height data.

The impact of input data on output data can be observed by examining the WEKA program algorithm equations. Accordingly:

In Equations. 2 and 3, according to the WEKA program LR algorithm, it is observed that the day, Mois., and Light input data have an increasing effect on the Water output data. The temperature input data has an increasing effect on the Height output data, while the Light input data has a decreasing effect.

In Equations. 4 and 5, according to the WEKA program SMOreg algorithm, it is observed that the day, Mois., and Light input data have an increasing effect on the Water output data, while the Temp. and Humi. input data have a decreasing effect. The day, Mois., Temp., and Humi. input data have an increasing effect on the Height output data, while the Light and Water input data have a decreasing effect.

In Equations. 6 and 7, according to the WEKA program MS algorithm, it is observed that the day, Mois., Humi., and Light input data have an increasing effect on the Water output data, while the Temp. input data has a decreasing effect. The day, Mois., Temp., and Humi. input data have an increasing effect on the Height output data, while the Light and Water input data have a decreasing effect.

5. Conclusion

At the end of the study, the following conclusions have been reached. The impact of Mois., Temp., Humi., and Light data between July 19, 2021, and September 09, 2021, on water consumption, as well as the effect of Mois., Temp., Humi., Light, and water consumption data on plant height, has been expressed.

The use of a wireless sensor and actuator network has been observed to consume 2583.32 ml less water compared to the traditional agricultural method. Accordingly, the use of a wireless sensor and actuator network is 25.83% more efficient in terms of water consumption compared to the traditional agricultural method.

The use of a wireless sensor and actuator network has been observed to result in average plant heights that are 9.67 cm taller compared to the traditional agricultural method. Accordingly, the use of a wireless sensor and actuator network is 21.57% more efficient in terms of plant height compared to the traditional agricultural method.

Table 4 illustrates the effects of input data on output data according to the equations obtained from the WEKA program algorithms. Since the WEKA program does not produce mathematical expressions for Gaussian Processes Regression, Multi-Layer Perceptron and M5Rules algorithms, they are not included in the table.

Table 4. Effect of environmental data on Water and Height

Algorithms	Input	Output Water (ml)	Output Height (cm)
LR	Day	Increasing	Unrelated
	Mois. (%)	Increasing	Unrelated
	Temp.(°C)	Unrelated	Increasing
	Humi. (%)	Unrelated	Unrelated
	Light (%)	Increasing	Decreasing
	Water (ml)	Unrelated	Unrelated
SMOreg	Day	Increasing	Increasing
	Mois. (%)	Increasing	Increasing
	Temp.(°C)	Decreasing	Increasing
	Humi. (%)	Decreasing	Increasing
	Light (%)	Increasing	Decreasing
	Water (ml)	Unrelated	Decreasing
MS	Day	Increasing	Increasing
	Mois. (%)	Increasing	Increasing
	Temp. (°C)	Decreasing	Increasing
	Humi. (%)	Increasing	Increasing
	Light (%)	Increasing	Decreasing
	Water (ml)	Unrelated	Decreasing

Considering the effects of input data on the output data for water consumption in Table 4; in all three equations obtained with the WEKA program, the input data of day, Mois., and Light have an increasing effect on water consumption output. According to these three equations, the input data of Temp. has a 2/3 decreasing effect, and the input data of Humi. has a 1/3 increasing and 1/3 decreasing effect.

Similarly, when considering the effects of input data on the output data for plant height in Table 4; in all three equations obtained with the WEKA program, the input data of day, Mois., and Humi. have a 2/3 increasing effect on plant height output. According to these three equations, the input data of Temp. has a 3/3 increasing effect, the input data of Light has a 3/3 increasing effect, and the input data of Water has a 2/3 decreasing effect.

As a result, it has been observed that irrigation using wireless sensor and actuator networks is not only more practical than traditional agricultural methods, but also more efficient in terms of both water consumption and plant growth. All the same, it has been observed that the most important environmental data to be considered in irrigation control with a wireless sensor and actuator network are soil moisture and ambient light data. It also shows that using the M5Rules algorithm to estimate water consumption based on day and environmental data and using the MLP algorithm to estimate plant height based on day and environmental data can achieve approximately accurate results.

Author contributions

Muhammed Furkan Koşum: Conceptualization, Data curation, Methodology, Software, Validation, Visualization and Writing-Original draft preparation, **Ali Osman Özkan:** Investigation, Writing-Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

1. De Lima, G. H. E. L., Silva, L. C. e, & Neto, P. F. R. (2010). WSN as a tool for supporting agriculture in the precision irrigation. *2010 Sixth International Conference on Networking and Services*, 137-142. https://doi.org/10.1109/ICNS.2010.26
2. Xiong, S., Wang, L., Qu, X., & Zhan, Y. (2009). Application research of WSN in precise agriculture irrigation. *2009 International Conference on Environmental Science and Information Application Technology*, 297-300. https://doi.org/10.1109/ESIAT.2009.231
3. Soy, H., Dilay, Y., & Özkan, A. (2016). Fuzzy control of agricultural irrigation system through wireless sensor/actuator networks. *Journal of Multidisciplinary Engineering Science and Technology (JMEST), 3*(11), 2458-9403.
4. Al Ameen, M., Islam, S. R., & Kwak, K. (2010). Energy saving mechanisms for MAC protocols in wireless sensor networks. *International Journal of Distributed Sensor Networks, 6*(1), 163413. https://doi.org/10.1155/2010/163413
5. Mahalik, N. P. (2007). *Sensor networks and configuration: fundamentals, standards, platforms, and applications*. Springer Science & Business Media. https://doi.org/10.1007/3-540-37366-7
6. Çıbuk, M., Arı, D., Ağgün, F., & Budak, Ümit. (2022). Investigation of failed node method to support healthy communication for linear wireless sensor networks. *Advanced Engineering Science, 2*, 21–26.
7. Akyildiz, I. F. & Vuran, M. C. (2010). *Wireless sensor networks*. John Wiley & Sons Ltd. https://doi.org/10.1002/9780470515181
8. Akyildiz, I. F., Weilian Su, Sankarasubramaniam, Y., & Cayirci, E. (2002). A survey on sensor networks, *IEEE Communications Magazine, 40*(8), 102-114. https://doi.org/10.1109/MCOM.2002.1024422
9. Soy, H. (2013). *Kablosuz algılayıcı ağlar için eşik tabanlı fırsatçı paket gönderim planı tasarımı (Tez No. 339817)*. Konya: Doktora Tezi, Selçuk Üniversitesi, Fen Bilimleri Enstitüsü. Yükseköğretim Kurulu Ulusal Tez Merkezi.
10. Dilay, Y., Soy, H., & Bayrak, M. (2012). Hassas tarımda kablosuz algılayıcı ağların kullanımı ve uygulama alanlarının incelenmesi. *Journal of the Institute of Science and Technology, 2*(A), 21-26.
11. Gondchawar, N. & Kawitkar, R. S. (2016). IoT based smart agriculture. *International Journal of Advanced Research in Computer and Communication Engineering, 5*(6), 2319-5940.
12. Keshtgari, M. & Deljoo, A. (2012). A wireless sensor network solution for precision agriculture based on zigbee technology. *Wireless Sensor Network, 4*(1), 25-30. https://doi.org/10.4236/wsn.2012.41004
13. Li, X., Deng, Y., & Ding, L. (2008). Study on precision agriculture monitoring framework based on WSN. *2008 2nd International Conference on Anti-Counterfeiting, Security and Identification*, 182-185.
14. Pandiyaraju, V., Logambigai, R., Ganapathy, S., & Kannan, A. (2020). An energy efficient routing algorithm for wsns using intelligent fuzzy rules in precision agriculture. *Wireless Personal Communications, 112*(1), 243–259. https://doi.org/10.1007/s11277-020-07024-8
15. Popescu, D., Stoican, F., Stamatescu, G., Ichim, L., & Dragana, C. (2020). Advanced UAV-WSN system for intelligent monitoring in precision agriculture. *Sensors, 20*(3), 817. https://doi.org/10.3390/s20030817
16. Soy, H., Dilay, Y., Özkan, A., Aydın, C., & Bayrak, M. (2013). Intelligent control of agricultural irrigation system based on wireless sensor and actuator networks. *In Proceedings of the 16th International Multiconference Information Society – IS*, 102-105.
17. Soy, H., Dilay, Y., Özkan, A., Aydın, C., & Bayrak, M. (2013). Kablosuz algılayıcı ve eyleyici ağlar yardımıyla konya ovasında tarımsal su tüketiminin azaltılması. *Ulusal Kop Bölgesel Kalkınma Sempozyumu, 14-16 Kasım 2013, Konya.*
18. Kassim, M. R. M. & Harun, A. N. (2016). Applications of WSN in agricultural environment monitoring systems. *2016 International Conference on Information and Communication Technology Convergence (ICTC)*, 344-349. https://doi.org/10.1109/ICTC.2016.7763493
19. Roham, V. S., Pawar, G. P., Patil, A. S. & Rupnar, P. R. (2015). Smart farm using wireless sensor network. *National Conference on Advances in Computing (NCAC2015), 6*, 8-11.

20. Yoo, S., Kim, J., Kim, T., & Ahn, S., Sung, J., Kim, D. (2007). A2S: automated agriculture system based on WSN. *2007 IEEE International Symposium on Consumer Electronics*, 1-5. https://doi.org/10.1109/ISCE.2007.4382216
21. Özkan, A. O., Özkan, A., & Soy, H. (2019). Türkiye'nin enerji talebinin regresyon analiz teknikleriyle uzun dönem tahmini. *8. Uluslararası Meslek Yüksekokulları Sempozyumu (UMYOS'19), 3*, 98-105.
22. İşler, Y. & Narin, A. (2012). WEKA yazılımında k-ortalama algoritması kullanılarak konjestif kalp yetmezliği hastalarının teşhisi. *Teknik Bilimler Dergisi, 2*(4), 21-29.
23. Bouckaert, R. R., Frank, E., Hall, M., Kirkby, R., Reutemann, P., Seewald, A., & Scuse, D. (2016). *WEKA manual for version 3-8-1*. Hamilton, New Zealand: University of Waikato.
24. Şaylan, Ç. A. (2013). *Böbrek nakli geçirmiş hastalarda akıllı yöntem tabanlı yeni öznitelik seçme algoritması geliştirilmesi (Tez No. 333151)*. İstanbul: Yüksek Lisans Tezi, Kadir Has Üniversitesi, Fen Bilimleri Enstitüsü. Yükseköğretim Kurulu Ulusal Tez Merkezi.
25. Karaca, C. & Karacan, H. (2016). Çoklu regresyon metoduyla elektrik tüketim talebini etkileyen faktörlerin incelenmesi. *Selçuk Üniversitesi Mühendislik, Bilim ve Teknoloji Dergisi, 4*(3), 182-195. https://doi.org/10.15317/Scitech.2016320514
26. Agunwamba, J. C., Tiza, M. T., & Okafor, F. (2024). An appraisal of statistical and probabilistic models in highway pavements. *Turkish Journal of Engineering, 8*(2), 300-329. https://doi.org/10.31127/tuje.1389994
27. Sebik, N. B. & Bülbül, H. İ. (2018). Veri madenciliği modellerinin akciğer kanseri veri seti üzerinde başarılarının incelenmesi. *TÜBAV Bilim Dergisi, 11*(3), 1-7.
28. Chaoqun, L. & Liangxiao, J. (2006). Using locally weighted learning to improve SMOreg for regression. In: Yang, Q. & Webb, G. (eds), *PRICAI 2006: Trends in Artificial Intelligence (375-384)*. Springer. Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-36668-3_41
29. Ebren Kara, Ş. & Şamlı, R. (2021). Yazılım projelerinin maliyet tahmini için WEKA'da makine öğrenmesi algoritmalarının karşılaştırmalı analizi. *Avrupa Bilim ve Teknoloji Dergisi (23)*, 415-426. https://doi.org/10.31590/ejosat.877296
30. Karaca Bilgen, G., Özbahçe, A., Yeter, T., Görgişen, C., Bahçeci, Alsan P., & Avağ, K. (2018). Farklı sulama seviyeleri ve malç uygulamalarında turşuluk hıyarın verim su ilişkileri. *Ziraat Fakültesi Dergisi*, 328-339.
31. Alan, R. (2011). Farklı budama sistemlerinin serada yetiştirilen hıyarda (cucumis sativus l) meyve özelliklerine ve verime etkisi. *Atatürk Üniversitesi Ziraat Fakültesi Dergisi, 19*(1-4).
32. Demirtop, A., & Sevli, O. (2024). Wind speed prediction using LSTM and ARIMA time series analysis models: A case study of Gelibolu. Turkish Journal of Engineering, 8(3), 524-536. https://doi.org/10.31127/tuje.1431629
33. Ayyıldız, E., & Murat, M. (2024). A lasso regression-based forecasting model for daily gasoline consumption: Türkiye Case. Turkish Journal of Engineering, 8(1), 162-174. https://doi.org/10.31127/tuje.1354501
34. Zela, K., & Saliaj, L. (2023). Forecasting through neural networks: Bitcoin price prediction. *Engineering Applications, 2*(3), 218-224.