Mathematical Modelling for Predicting the Yield from Vertical Farming

Mira Doshi
Jamnabai Narsee International School,
Mumbai, India
miradoshimail@gmail.com

Vinay Vishwakarma
On My Own Technology,
Mumbai, India
vinay.vishwakarma@omotec.in

Abstract

Excessive urbanization, global warming and uncontrollable use of herbicides and pesticides have significantly affected the soil's fertility in a negative manner. Above that, with the increase in global population there is a decrease in the amount of land available to each person which impacted the soil productivity. Vertical farming or vertical plantation is a solution to these problems. This paper explores the concept and potential benefits of vertical farming as an innovative and sustainable solution for agricultural production. Moreover, in this paper an Artificial Neural Network (ANN) model is developed to predict the yield from vertical farming. The practicality of a model lies in its ability to be applied in a practical problem. Hence, the ANN model is used in predicting the tomato yield from hydroponics which is a hydroponic technique. The data for the paper are collected from 50 different farmers who are growing tomatoes by hydroponics from 1st April to 30th April, 2023. The collected data contains information about pH, temperature and concentration of fertilizers which are statistically analyzed by Analysis of Variance (ANOVA). Then the data are utilized to develop ANN networks which are analyzed by the root mean square error (RMSE) value. The best network have co-efficient of correlation value (r) for training, testing, validation and overall of 0.98469, 0.97247, 0.94088 and 0.97918 respectively. Also, the network is validated for detecting any machine learning error.

Keywords: Vertical farming, ANN, ANOVA, pH, Superphosphate, Phosphate, Urea, Temperature, Tomato yield.

1. Introduction

A plantation is a large agricultural estate where cash crops are grown, typically in tropical or subtropical regions. Historically, plantations were associated with the production of crops such as sugar, cotton, tobacco, coffee, and tea, often utilizing slave or indentured labor. Plantations played a significant role in the colonial economy and the transatlantic slave trade. Today, the term "plantation" can refer to any large-scale agricultural operation focused on the cultivation of specific crops.

Vertical plantation, also known as vertical farming or vertical agriculture, is a method of cultivating plants in vertically stacked layers or structures. It is an innovative approach to agriculture that maximizes land and space utilization by growing crops in a controlled indoor environment, often using techniques such as hydroponics or aeroponics. Vertical plantations can be implemented in urban areas, abandoned buildings, or specially designed vertical farming facilities. This approach offers advantages such as year-round crop production, efficient resource use, reduced water consumption, and protection from pests and weather conditions. Vertical plantations have the potential to increase food production in urban areas and provide sustainable solutions for agriculture. The advantages of vertical plantation, or vertical farming, includes i) Space efficiency, ii) Year-round crop production, iii) Efficient resource utilization, iv) Reduced transportation and emissions, v) Crop protection, and vi) Sustainability and resource conservation.

Vertical farming allows for the cultivation of crops in vertically stacked layers, maximizing land use and making it possible to grow a large amount of produce in a relatively small footprint. This is particularly beneficial in urban areas where land availability is limited. Vertical plantations can operate in controlled indoor environments, providing optimal conditions for plant growth regardless of external weather conditions. This enables year-round crop production, reducing dependency on seasonal variations and improving overall productivity. Vertical farming utilizes advanced technologies such as hydroponics or aeroponics, which allow for precise control over nutrient delivery and water usage. These methods typically require less water compared to traditional soil-based agriculture and can significantly reduce the need for pesticides and fertilizers.

Vertical plantations can be located closer to urban centers, reducing the distance food needs to travel from farm to table. This can lower transportation costs, decrease food waste, and reduce carbon emissions associated with long-distance transportation. Indoor vertical farming provides a controlled environment that protects crops from pests, diseases, and extreme weather events. This reduces the reliance on chemical pesticides and allows for organic or pesticide-free cultivation. Vertical farming has the potential to minimize land degradation, conserve natural resources, and decrease the overall environmental impact of agriculture. By optimizing resource use, such as water and energy, and implementing sustainable practices, vertical plantations can contribute to a more sustainable and resilient food production system.

1.1. Motivation and Novelties

Although the vertical farming is a modern solution to the farming problems, yet the researchers have limited resources to predict the yield from it. Focusing on this limitation, a deep learning (DL) model is developed in this study for the prediction. In this regard, different factors that influenced the vertical farming are analyzed and those that showed significant relationship are modelled with the yield.

Remainder of the paper are arranged as follows:
Section 2 summarizes the recent literatures on vertical plantation. Then section 3 give a detailed description of tomato cultivation by vertical faming which is followed by methodology used in this paper for analyzing the factors and developing the DL model in section 4. Section 5 outlines results obtained from training and testing the developed DL model. Finally section 6 concludes the paper.

2. Review of Contemporary Literatures

The soil's fertility has been negatively impacted by rapid urbanization, natural disasters, global warming, as well as the indiscriminate use of herbicides and pesticides. Additionally, the amount of land available to each person has decreased, soil fertility has deteriorated, and soil productivity has been severely reduced [1 – 3]. The challenges to the water resources in the watershed include a changing climate, rising temperatures, frequent dry spells, and the unpredictability of the weather. A few threats to the water resources of the watershed include excessive irrigation water use, unregulated water contamination, and a downward trend in groundwater levels [4]. With a projected population of 8.9 billion people by 2050, the globe would need to produce 50% more food, which will require additional arable land that will simply not be accessible [5]. The amount of arable land per person is predicted to be less than 0.20 hectares by 2050, which is less than a third of what it was in 1970 [6]. Figure 1 shows the graph for arable land per person and world population from 1950 to 2050.

Food production is currently a significant concern due to the serious threats these issues pose to traditional soil-based agricultural production systems. Modern agricultural methods that are more productive and environmentally benign must be used in addition to soil-based farming techniques [7]. When putting these new agricultural practises into practise, it is important to take into account problems like climate change, decreased soil productivity, depleted soil nutrient stocks, and limited irrigation water availability. Systems of soilless farming might provide a solution to these contemporary problems. Vertical farming methods may be used in place of soil-based farming systems as a supplementary approach to help address the current scarcity of fertile arable lands and water. One such specialised method for growing plants without soil is liquid culture, often known as hydroponics. Because they are founded on the idea that nature has established as the basis for life, artificial crop production techniques are not unnatural. Other methods of soil-less growing include rooftop farming, aquaponics, and aeroponics [8].

Vertical farming is a successful method for growing a variety of field crops, including rice, wheat, tubers, fodder maize, and numerous vegetable crops like spinach, okra, cucumber, onion, carrot, and tomato. This method exhibits greater production and better nutritional status [9 – 11].

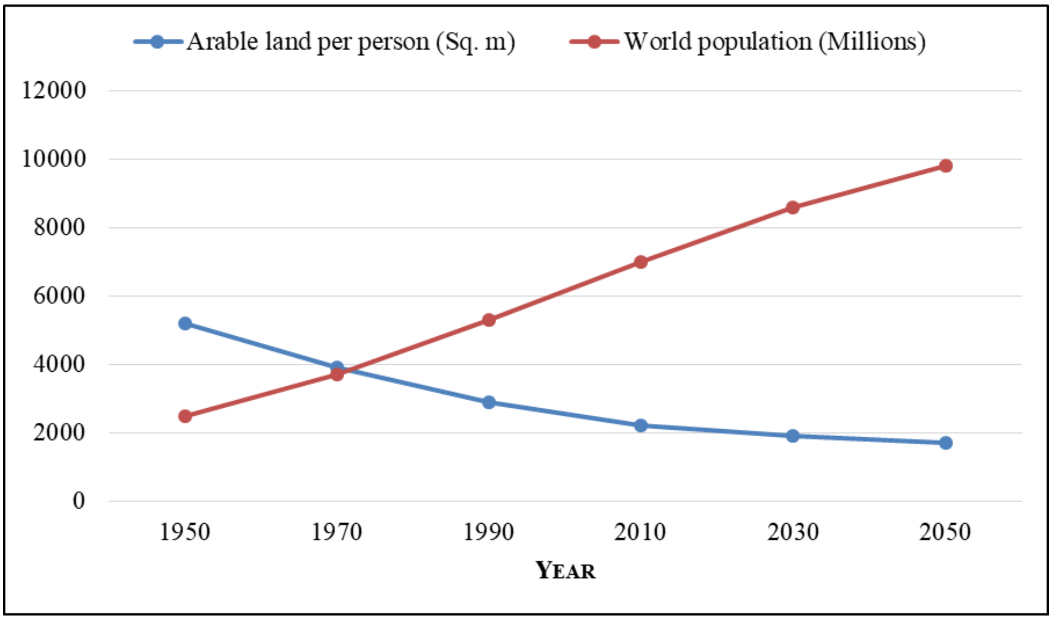

Figure 1: Arable land per person vs World population from 1950 to 2050 [6]

For highly cultivated maize fodder production, the hydroponic vertical farming technology is optimum. Animals enjoy the entire crop of feed, which is ready in 7-8 days. Additionally, compared to conventional field crops, vertical farm crops yield around 10 times more while using about 70–95% less water [12]. Using an artificial growing media and a nutrient solution created to deliver the precise amounts of nutrients required for their development and growth, plants are grown hydroponically [13]. This might be viewed as the art of controlling water, feeding it with nutrients essential to plants, and timely delivering it to their parched roots to produce the best harvests with much less water and labor than would otherwise be needed [14]. Hydroponically grown plants are healthier because their diets are more balanced than those of soil-grown plants [15]. When compared to conventional farming, hydroponics has the advantage of allowing complete control over the crop's nutrition, which enhances both nutrient regulation and water management [16]. For the reasons outlined above, hydroponics is largely recognized as a superior kind of farming in almost every nation on the planet [17]. Hydroponic research is being conducted by an increasing number of affluent countries, especially the United States and China, in an effort to increase food yields and get around existing problems. In affluent nations, a variety of crops, including lettuce, cucumbers, and tomatoes, have been investigated for hydroponic production [18]. Although it is still a new method, additional research is required before it can be applied to agricultural production in developing nations. India has an increasing demand for hydroponically grown food due to the nation's quick population expansion. It has not become a reality due to a preponderance of the evidence and research on the technology's various components.

3. Vertical farming details

Tomato is one of the most commonly cultivated horticultural crop by hydroponics which is a technique of vertical farming. There are several factors that impacts the performance in vertical plantation viz. a) Air Quality, temperature, and Relative Humidity, b) Water Quality, c) Light Quality, d) Substrate, e) Nutrient, f) pH, g) Watering, h) Selection of Ideal Crop and i) Vulnerability [8]. To guarantee that tomatoes are cultivated in an ideal environment, all the factors are regularly monitored. Tomato cultivators maintains an optimal environment by conserving the water quality, light quality, and substrate. However, the components which are varied for tomato cultivations from farmer to farmer includes temperature, pH and the amount of nutrients.

The optimal temperature required for the optimal tomato yield is when the daytime temperature is between 70 and 80^{O} F and the nighttime temperature is between 60 and 70^{O} F. pH plays a very important role on tomato yield. The desirable pH range for optimum yield varies among crops. For maximum possible yield of tomato the optimal pH range is 6 – 7. When the pH value decreases from the optimal range, a huge deviation from the maximum yield is observed. pH value not only determine the acidity but also determines the percentage of effective utilization of fertilizers. The fertilizers used for tomato cultivation are 1) Super phosphate (SP), 2) Potash (Po) and 3) Urea (Ur) in the ratio of 2:1:3. For obtaining the maximum possible yield of tomato the optimal amount of fertilizers required per hectare are 50kg of super phosphate, 25kg of potash and 75kg of urea. But when used in acidic soil the effect of fertilizer is decreased and extra amount of fertilizers have to be used to compensate the waste effect of fertilizers. The effect of fertilizer wasted against pH range is shown in the Table 1.

Table 1: Percentage of Fertilizer Wasted against pH Range.

Sl. no.	pH range	Super phosphate (%)	Potash (%)	Urea (%)	Total effect wasted (%)
1	0<pH<4.5 (extremely acidic)	70.00	77.00	67.00	71.33
2	4.5<pH<5.0 (very strong acid)	47.00	66.00	48.00	53.67
3	5<pH<5.5 (strong acid)	23.00	52.00	23.00	32.67
4	5.5<pH<6 (medium acid)	11.00	48.00	0.00	19.67
5	6<pH<7 (neutral)	0.00	0.00	0.00	0.00

3.1. Data acquisition

In this paper, the data of tomato yield per day, temperature, pH value and the amount of SP, Po and Ur used are considered. The data are collected from 1st April to 30th April, 2023 from 50 different tomato cultivators who uses hydroponics as a way of cultivation. Table 2 shows the data collected for the study.

Table 2: Data collected from different tomato cultivators

Farmer	Cultivated Area (Acre)	pH	Temperature (ºF)	Super Phosphate (Kg)	Potash (Kg)	Urea (Kg)	Total Yield (Quintal)
1	0.7	5.4	69	50	20	50	45
2	1.7	6.2	66	125	50	125	200
3	1.3	6.2	75	100	40	100	190
4	0.5	4.8	68	37.5	15	37.5	26.25
5	1.3	4.6	71	100	40	100	90
6	2.3	6	69	175	70	175	227.5
7	0.7	6.1	72	50	20	50	70
8	1	6	64	75	30	75	90
9	0.7	6.5	64	50	20	50	85
10	1	6.1	70	75	30	75	105
11	1.7	4.6	80	125	50	125	112.5
12	2	5.2	65	150	60	150	165
13	0.3	5	81	25	10	25	27.5
14	0.5	4.9	62	37.5	15	37.5	37.5
15	0.7	4.8	72	50	20	50	40
16	1.2	4.5	64	87.5	35	87.5	78.75
17	1.7	4.2	77	125	50	125	100
18	1	5.1	78	75	30	75	82.5
19	0.7	4	75	50	20	50	35
20	0.7	5	73	50	20	50	25
21	1.3	4.5	74	100	40	100	90

22	3.3	5.9	68	250	100	250	325
23	1.7	6.2	74	125	50	125	162.5
24	3	4	69	225	90	225	157.5
25	1	6	79	75	30	75	127.5
26	1	4.1	68	75	30	75	52.5
27	1	6.1	82	75	30	75	126
28	0.7	4.3	80	50	20	50	44
29	0.7	4.3	72	50	20	50	44
30	1.7	6	72	125	50	125	150
31	1	5.1	76	75	30	75	75
32	0.3	4.5	72	25	10	25	10
33	0.5	4.8	73	37.5	15	37.5	26.25
34	0.7	4.2	80	50	20	50	25
35	0.7	5.9	78	50	20	50	65
36	0.5	4	74	37.5	15	37.5	26.25
37	1.7	6	74	125	50	125	175
38	2.5	6	69	187.5	75	187.5	225
39	2	5.9	65	150	60	150	165
40	0.8	4.2	63	62.5	25	62.5	50
41	1.2	4.1	74	87.5	35	87.5	61.25
42	1	4.8	79	75	30	75	37.5
43	1.3	4.3	78	100	40	100	80
44	0.5	4.5	66	37.5	15	37.5	30
45	2	6.5	63	150	60	150	255
46	0.7	4.5	63	50	20	50	45
47	0.8	5	76	62.5	25	62.5	68.75
48	1.2	5.3	64	87.5	35	87.5	105
49	1.2	4.4	75	87.5	35	87.5	78.75
50	2	5.9	69	150	60	150	225

The tomato yield is the sum of the tomato yield per day for the above mentioned period.

4. Methodology

In this section, the preliminary methods used for developing the DL model is discussed in brief.

4.1. Analyzing the significance of the factors

In this section of the study, the significance of the factors with respect to the tomato yield are tested for the data as shown in table 2. The analysis of variance (ANOVA) test is used to determine the significant factors. The ANOVA test measures the degree by which the data are affected by the external factors or noise. The significance of the factors is tested based F-value and p-value at an upper confidence of 95% [19]. The F – value computed for the data collected measures the mean square for the residual error of a regression model. On the other hand the p – values measures the tendency of the data collected to fit into the model created [20]. The p – value should be less than 0.05 which implies that 5% of permissible tolerance for the data to be affected by the environmental noise. Table 3, 4 and 5 shows the ANOVA table for pH, temperature and fertilizers (SP, Po and Ur) respectively.

Table 3: ANOVA table for pH

	df	*SS*	*MS*	*F*	*p*

Regression	1	102885.2	102885.2	33.22732	0
Residual	48	148627.4	3096.403		
Total	49	251512.5			

Table 4: ANOVA table for temperature

	df	*SS*	*MS*	*F*	*p*
Regression	1	13161.47787	13161.47787	2.650506116	0.11
Residual	48	238351.0584	4965.647049		
Total	49	251512.5363			

Table 5: ANOVA table for fertilizers (SP, Po and Ur)

	df	*SS*	*MS*	*F*	*p*
Regression	1	199728.6	199728.6	185.134	0
Residual	48	51783.95	1078.832		
Total	49	251512.5			

The *df*, *SS*, *MS*, *F* and *p* in table 3 – 5 stands for degree of freedom, sums of square, mean sums of square, F – value and p – value respectively. From the ANOVA table, it is observed that the pH and concentration of fertilizers used for the vertical faming plays a significant role tomato yield. However, the temperature is not a significant factor as evident from the p – value of table 4. The data collected for the study showed the presence of certain noise in the dataset.

4.2. Modelling the factors by DL algorithm

Deep learning algorithms are a subset of machine learning family where multiples layers are used to form a network. Artificial Neural Network (ANN) is a supervised DL algorithm which primarily comprises of three layers namely input layer, output layer and one or more hidden layers []. ANN is inspired by the central nervous system of animals particularly brain. The brain learns from the past experience. Brain is a complex network of neurons which process signals as received by the sensory organs and asked to react according to the situation. Similarly, ANN is a system of interconnected neurons which exchange messages with each other.

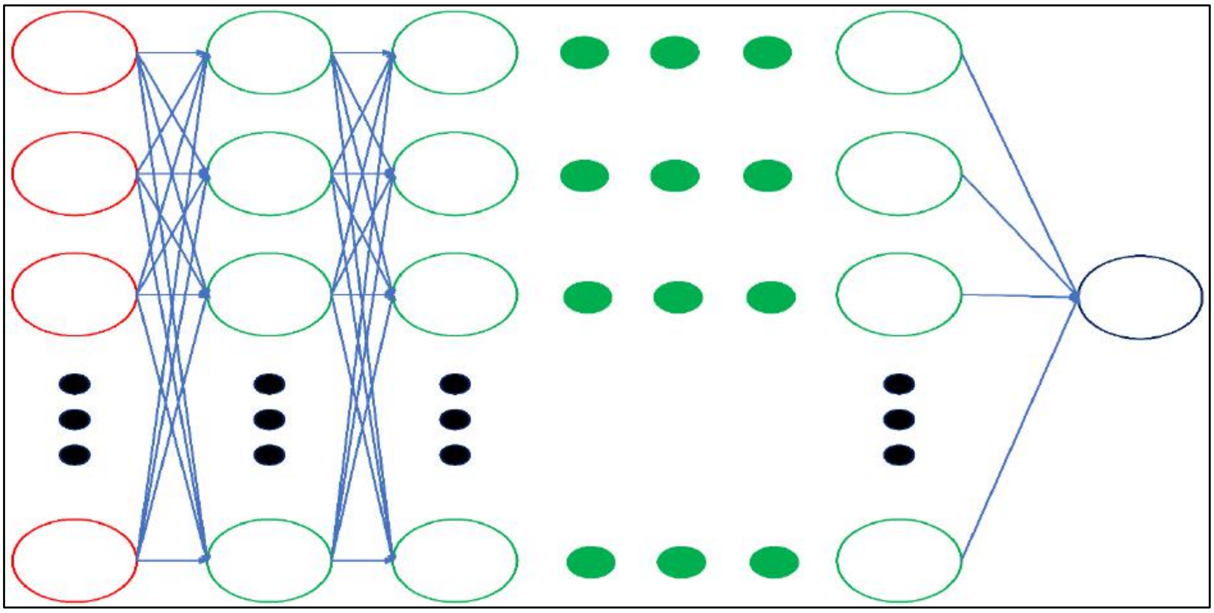

Figure 2: Representation of a typical ANN model

There are some numeric weights at the connections of the neurons. These numeric weights can be adjusted/readjusted through training. Training is the process of learning new task by repeatedly doing it. When an ANN model is trained the predicted output obtained is compared with the actual output and the numeric weights are updated. When a particular ANN model is trained repeatedly the numeric weights shall be updated until the error between predicted output and actual output is the least. A typical representation of ANN model is shown in figure 2.

4.2.1. Basic components of artificial neurons

An artificial neuron comprises of interconnection weights, summing factor and a transfer unit. The outputs from the neurons of previous layer act as input for the present layer which are multiplied by the interconnected weights of each neuron and summed value is transferred to form the output of the present neuron. Figure 3 represents an artificial neuron.

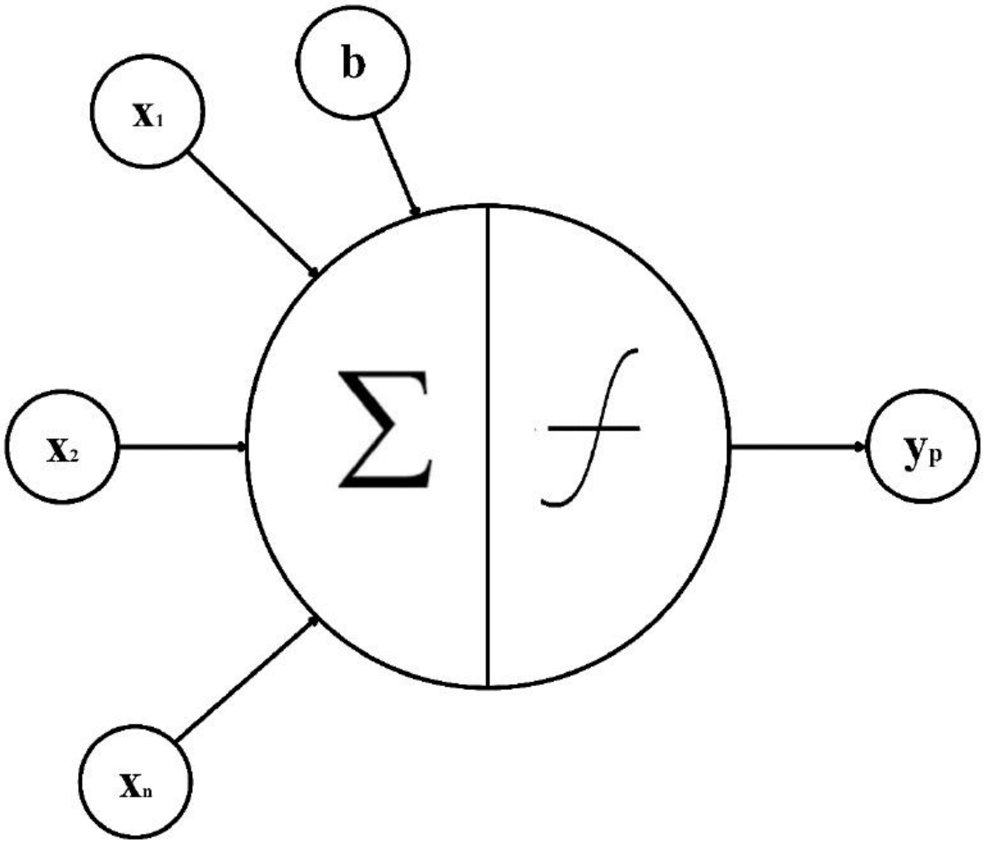

Figure 3: Representation of an artificial neuron

Most commonly used transfer functions include the following:

a) **Tangent hyperbolic sigmoid (tansig function):** This transfer function converts the input data in the range of $-1\ to\ 1$. Mathematically tansig function is represented as:

$$y = \frac{2}{1+exp(-2x)} - 1 \quad (1)$$

b) **Logarithmic sigmoid (logsig) function:** This transfer function converts the input data in the range of $0\ to\ 1$. Mathematically tansig function is represented as:

$$y = \frac{1}{1+exp(-x)} \quad (2)$$

c) Pure linear (purlin) function: Mathematically purlin transfer function is represented as:

$$y = x \tag{3}$$

4.2.2. Different ANN parameters

ANN comprises of different parameters which need to be optimized. These parameters are number of hidden layers and number of hidden neurons. In this study, the number of hidden layers and neurons are optimized by the performance of the developed network which is measured by computing the root mean square error (RMSE) value between the actual and target. The mathematical representation of MSE is shown in Eq. (4).

$$RMSE = \sqrt{\frac{1}{n}\sum_{j=1}^{n}\left(y_j - p_j\right)^2} \tag{4}$$

Where y_i and p_i stands for target and predicted value for the jth observation and n represents the number of observations. The network that has the least MSE value is selected for the tomato yield prediction.

In this study, the number of hidden layers are set either as 1 or 2 and the number of hidden neurons are varied from 1 to 15. The network that has the best performance or least MSE value is selected for predicting the tomato yield by the process of hydroponics.

4.2.3. Predictive transfer function developed by ANN

Referring to the figure 2 and 3 of the study, the predictive transfer function obtained can be mathematically represented as follows:

$$h_{iu} = \left(\sum_{k=1}^{m} w_k \times x_k\right) + b \tag{5}$$

$$h_{ou} = \left(h_{iu}\right)^{tr} \tag{6}$$

$$y_v = \left(\sum_{k=1}^{v} w_{ku} \times h_{ou}\right) + b \tag{7}$$

$$y_{vred} = \left(y_v\right)^{tr} \tag{8}$$

Where w_k and x_k are interconnection weight and input for the kth neuron, b is the bias, h_{iu} and h_{ou} is the input and output for the uth hidden layer. The steps for generating the predictive transfer function is shown as flowchart in figure 4.

5. Results and discussions

In this section of the study, the results obtained from the proposed methodology is discussed in brief. The first part of the analysis involves the statistical study of the factor that are significantly affecting the tomato yield. In this regard, ANOVA analysis is conducted. From the analysis, it is observed that the pH and concentration of fertilizers used are significantly effecting the tomato yield whereas, the temperature is not a significant factor for the study. Then data are modelled with the help of ANN. Step by step flow for generating the predictive model is described as such:

a. At first the data are pre-processed and converted into normalized values as per the Eq. (9).

$$n = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{9}$$

Data are pre-processed, in order to bring homogeneity in the data dataset. The data pre-processed by Eq. (9)

b. In the second step, the pre-processed data are randomly split into training and testing data in the ratio of 70:30. 70% of the data are used for training the network whereas the remaining 30% data are used for testing and validation. The testing dataset are also helpful in identifying any overfitting in the network developed. The dataset collected for the study consists of fifty data. Out of which 35 data are used for training and the remaining 15 data are used for testing and validation.

c. In the third step, the performance of different networks build by varying the parameters are noted and tabulated in table 6.

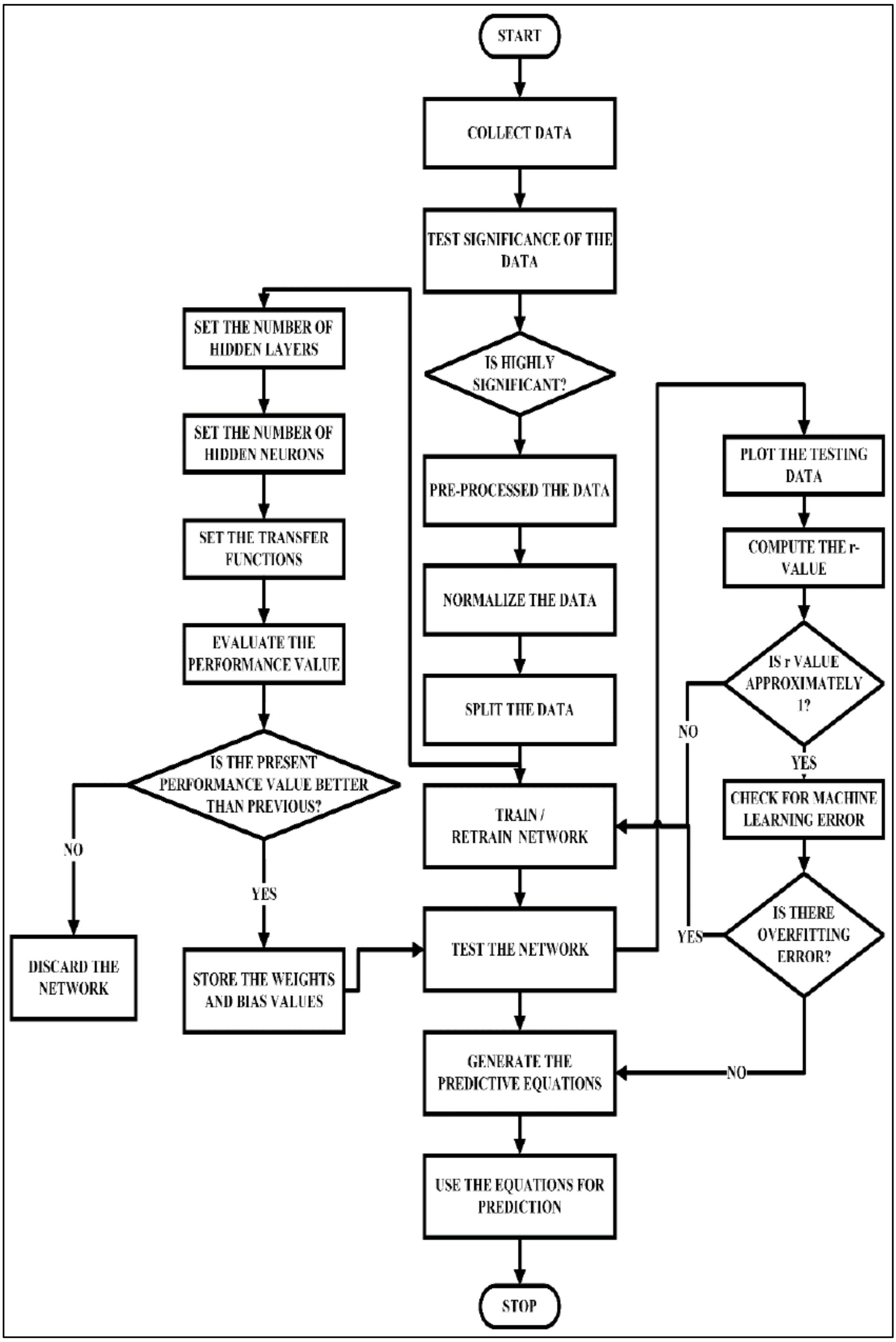

Figure 4: Flowchart for the proposed method

Table 6: RMSE value for different networks developed

Sl. no.	Network	Number of		Transfer function	RMSE value
		Hidden layer	Hidden neurons		
1	Network 1	1	2	tansig – purlin	46.34156
2	Network 2	1	3	tansig – purlin	16.62714
3	Network 3	1	4	purlin – logsig	47.68525
4	Network 4	1	5	purlin – tansig	37.83953
5	Network 5	1	6	purlin – purlin	26.46005
6	**Network 6**	**1**	**7**	**logsig – tansig**	**14.61641**
7	Network 7	1	8	purlin – tansig	16.69627
8	Network 8	1	9	logsig – purlin	24.59627
9	Network 9	1	10	tansig – purlin	48.40644
10	Network 10	1	11	tansig – purlin	25.19582
11	Network 11	1	12	logsig – purlin	15.02284
12	Network 12	1	13	logsig – logsig	28.02103
13	Network 13	1	14	purlin – tansig	17.76559
14	Network 14	1	15	purlin – purlin	32.78696
15	Network 15	2	1 – 5	logsig – logsig – tansig	44.36634
16	Network 16	2	2 – 2	tansig – tansig – purlin	20.40411
17	Network 17	2	3 – 8	logsig – tansig – purlin	24.84235
18	Network 18	2	4 – 7	logsig – purlin – purlin	47.37435
19	Network 19	2	5 – 7	logsig – tansig – tansig	42.05835
20	Network 20	2	6 – 1	tansig – tansig – tansig	40.65514
21	Network 21	2	7 – 4	purlin – logsig – logsig	20.65671
22	Network 22	2	8 – 1	tansig – logsig – logsig	48.50752
23	Network 23	2	9 – 8	logsig – logsig – tansig	20.30917
24	Network 24	2	10 – 7	logsig – tansig – purlin	32.37749
25	Network 25	2	11 – 9	logsig – purlin – logsig	21.94927
26	Network 26	2	12 – 8	logsig – purlin – tansig	43.22033
27	Network 27	2	13 – 5	logsig – logsig – logsig	21.33962
28	Network 28	2	14 – 2	tansig – tansig – logsig	15.12081
29	Network 29	2	15 – 10	purlin - purlin - logsig	15.2836

*Keeping the number of hidden layer and hidden neurons constant the RMSE values for different combination of transfer functions are computed. The combination for which the RMSE value is minimum is shown in the table. For example, in serial number 15, the number of hidden layer is fixed at 2 and the number of hidden neurons in hidden layer 1 and 2 is 1 and 5 respectively. The transfer function between in the different layers are varied. Then the RMSE value for each combination is determined. In serial number 15, the RMSE value for logsig transfer function between input and hidden layer-1, logsig transfer function between hidden layer-1 and hidden layer-2 and tansig transfer function between hidden layer-2 and outer layer have the lowest value. Hence, it is shown in the table 6.

A total of 6201 networks were developed for this study. Out of the networks developed, the network 6 with 1 hidden layer and 7 hidden neurons have the least RMSE value. The training, testing and validation regression graph developed for network 6 is shown in figure 5.

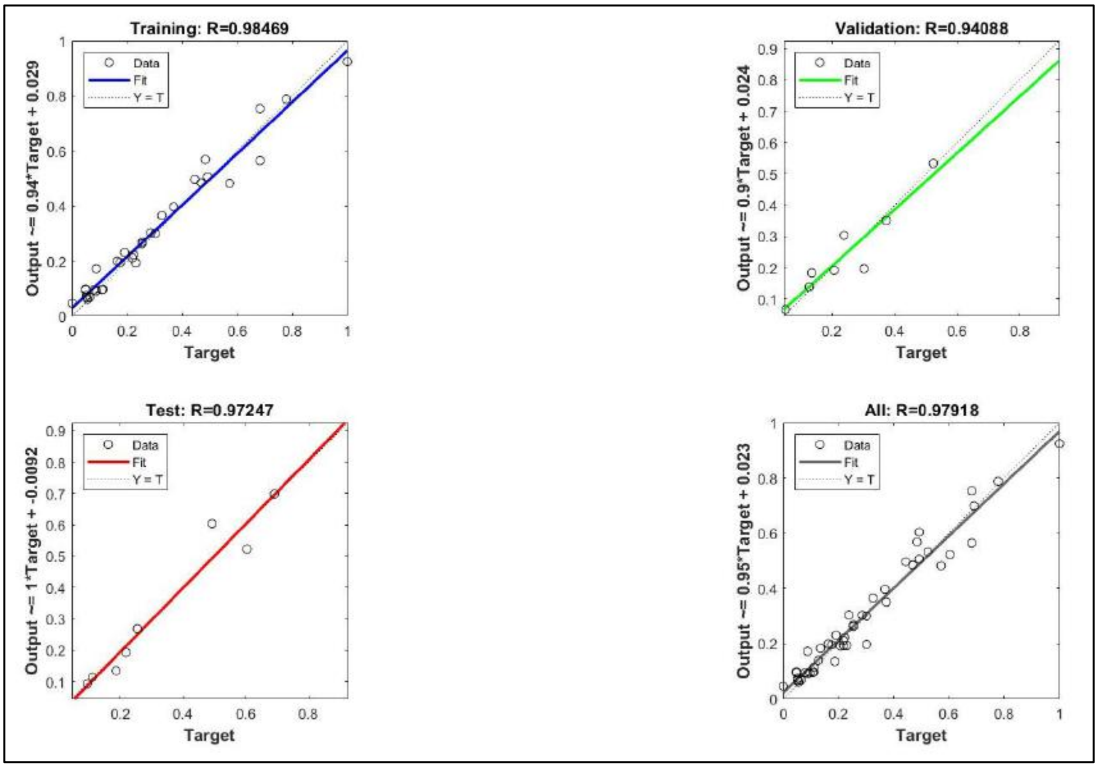

Figure 5: Regression graph for the **network 6**.

The co-efficient correlation (r) value for training, testing, validation and overall is 0.98469, 0.97247, 0.94088 and 0.97918 respectively. From the validation graph (top right hand graph) shows a linear distribution of the point for the developed dataset. As a result, it can be concluded that there is no overfitting error in the network developed.

d. In the fourth step of the study, the weights and the bias values for the best network is stored in an array. The weights and the bias values are shown in table 7.

Sl. no.	Weights input and hidden layer1					Bias input and hidden layer1
1	-0.50123	2.7576	1.6587	2.1934	0.12679	4.1733
2	-1.3676	-1.9095	1.7655	-1.5925	-1.7729	3.0023
3	-1.6778	-0.25293	-2.3356	-1.2051	2.5086	0.36961
4	2.7183	-1.3271	-0.90811	-2.4306	0.42292	-0.53479
5	-1.6522	1.9084	-1.1902	-1.8822	-0.68109	-2.0408
6	1.2794	-2.3489	1.5796	1.9558	0.020901	3.5406
7	-0.2151	3.4284	-0.88378	1.0762	-1.2751	4.158

	Weights hidden layer1 and output							Bias hidden layer1 and output
8	0.36508	-0.7498	-1.8895	-0.8479	0.24573	0.89013	-1.9705	2.7614

e. In the fifth step, the predictive transfer function is developed as follows:

$$h_{i1-1} = -0.50123 * \text{pH} + 2.7576 * \text{t} + 1.6587 * \text{SP} + 2.1934 * \text{Po} + 0.12679 * \text{Ur} + 4.1733 \quad (10)$$

$$h_{i1-2} = -1.3676 * \text{pH} - 1.9095 * \text{t} + 1.7655 * \text{SP} - 1.5925 * \text{Po} - 1.7729 * \text{Ur} + 3.0023 \quad (11)$$

$$h_{i1-3} = -1.6778 * \mathrm{pH} - 0.25293 * \mathrm{t} + -2.3356 * \mathrm{SP} - 1.2051 * \mathrm{Po} + 2.5086 * \mathrm{Ur} + 0.36961 \quad (12)$$

$$h_{i1-4} = 2.7183 * \mathrm{pH} - 1.3271 * \mathrm{t} - 0.90811 * \mathrm{SP} - 2.4306 * \mathrm{Po} + 0.42292 * \mathrm{Ur} - 0.53479 \quad (13)$$

$$h_{i1-5} = -1.6522 * \mathrm{pH} + 1.9084 * \mathrm{t} - 1.1902 * \mathrm{SP} - 1.8822 * \mathrm{Po} - 0.68109 * \mathrm{Ur} - 2.0408 \quad (14)$$

$$h_{i1-6} = 1.2794 * \mathrm{pH} - 2.3489 * \mathrm{t} + 1.5796 * \mathrm{SP} + 1.9558 * \mathrm{Po} + 0.020901 * \mathrm{Ur} + 3.5406 \quad (15)$$

$$h_{i1-7} = -0.2151 * \mathrm{pH} + 3.4284 * \mathrm{t} - 0.88378 * \mathrm{SP} + 1.0762 * \mathrm{Po} - 1.2751 * \mathrm{Ur} + 4.158 \quad (16)$$

$$h_{o1-1} = \frac{1}{1+exp(-h_{i1-1})} \quad (17)$$

$$h_{o1-2} = \frac{1}{1+exp(-h_{i1-2})} \quad (18)$$

$$h_{o1-3} = \frac{1}{1+exp(-h_{i1-3})} \quad (19)$$

$$h_{o1-4} = \frac{1}{1+exp(-h_{i1-4})} \quad (20)$$

$$h_{o1-5} = \frac{1}{1+exp(-h_{i1-5})} \quad (21)$$

$$h_{o1-6} = \frac{1}{1+exp(-h_{i1-6})} \quad (22)$$

$$h_{o1-7} = \frac{1}{1+exp(-h_{i1-7})} \quad (23)$$

$$y_i = 0.36508 * h_{o1-1} - 0.7498 * h_{o1-2} - 1.8895 * h_{o1-3} - 0.8479 * h_{o1-4} + 0.24573 * h_{o1-5} + 0.89013 * h_{o1-6} - 1.9705 * h_{o1-7} + 2.7614 \quad (24)$$

$$y_{pred} = \frac{2}{1+exp(-2y_i)} - 1 \quad (25)$$

f. In the final step, the output obtained from Eq. (10 – 25) is plotted against the output from the collected data. Figure 6 shows the scatter plot between the predicted and the collected output. The r value for the scatter plot between yield and predict yield is 0.979183.

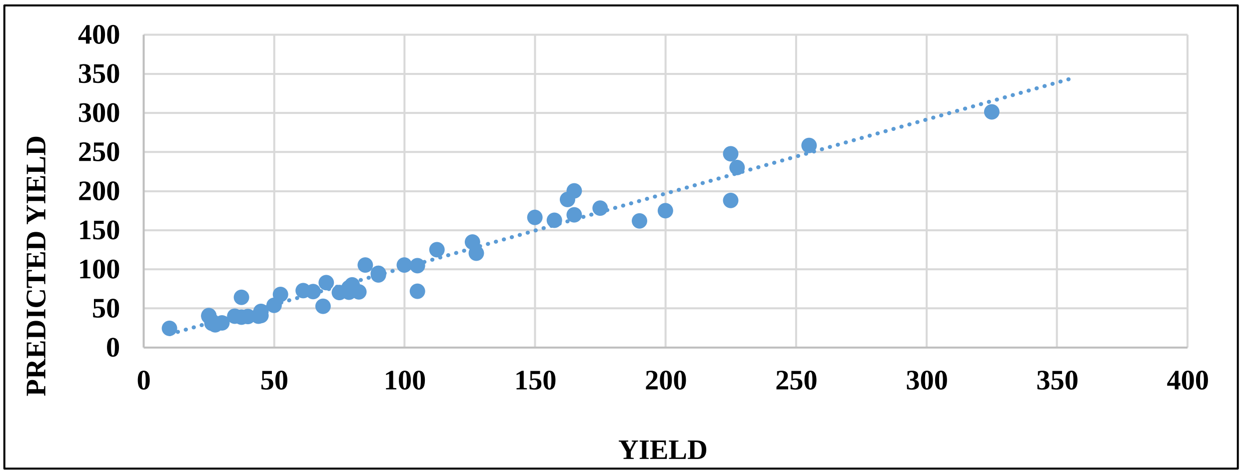

Figure 6: Scatter plot between Yield and Predict Yield

6. Conclusion

This paper has explored the concept and potential benefits of vertical farming as an innovative and sustainable solution for agricultural production. Vertical farming refers to the practice of cultivating plants in vertically stacked layers or structures, using controlled-environment agriculture techniques such as hydroponics, aeroponics, and artificial lighting. The key findings of this paper suggest that vertical farming has the potential to address several pressing challenges faced by traditional agriculture, including land scarcity, water shortage, climate change, and food security. By utilizing vertical space and optimizing resource utilization, vertical farming can significantly increase crop yield per unit area and reduce the environmental footprint associated with conventional farming methods.

Although, there are many benefits of vertical farming, yet, there are very limited work in predicting the yield. In this regard, an ANN model is developed in this study which is further applied to predict the tomato yield by the process of hydroponics. The data are collected from 50 different farmers who are growing tomatoes by hydroponics from 1st April to 30th April, 2023. The data collected are statistically analyzed to determine the significant factors by ANOVA analysis. The pH and concentration of fertilizers which is mixture of superphosphate, phosphate and urea shows significant relation with the tomato yield. However temperature is an insignificant factor. Then the data are utilized to develop ANN networks which are analyzed by the RMSE value. A total of 6201 ANN networks were developed by varying the hidden layer from 1 to 2, hidden neurons from 1 to 15 and the transfer functions. The architecture of the best network have 1 hidden layer with 7 hidden neurons and logsig and tanssig as transfer function between input and hidden layer and hidden and output layer respectively. The $r - value$ for training, testing, validation and overall is 0.98469, 0.97247, 0.94088 and 0.97918 respectively. Also, from the validation it was observed that the data are linearly arranged and hence there is no case of overfitting. Therefore, it can be concluded that the developed model is best suited for predicting yield from vertical farming.

Conflict of Interest:

The authors would like to inform that the study is not sponsored in any form by government, private and semi-government organizations.

References

1. Lambin, E. F. (2012). Global land availability: Malthus versus Ricardo. *Global Food Security*, *1*(2), 83-87.

2. Lal, R. (2015). Restoring soil quality to mitigate soil degradation. *Sustainability*, *7*(5), 5875-5895.

3. Lehman, R. M., Cambardella, C. A., Stott, D. E., Acosta-Martinez, V., Manter, D. K., Buyer, J. S., ... & Karlen, D. L. (2015). Understanding and enhancing soil biological health: the solution for reversing soil degradation. *Sustainability*, *7*(1), 988-1027.

4. Bhanja, S. N., Mukherjee, A., & Rodell, M. (2018). Groundwater storage variations in India. *Groundwater of South Asia*, 49-59.

5. Baudoin, W., Nono-Womdim, R., Lutaladio, N., Hodder, A., Castilla, N., Leonardi, C., ... & Duffy, R. (2013). Good agricultural practices for greenhouse vegetable crops: principles for mediterranean climate areas. *FAO plant production and protection paper (FAO)*.

6. Conforti, P. (2011). Looking ahead in world food and agriculture: perspectives to 2050. ROME. FAO.

7. Lambin, E. F., & Meyfroidt, P. (2011). Global land use change, economic globalization, and the looming land scarcity. *Proceedings of the national academy of sciences*, *108*(9), 3465-3472.

8. Mir, M. S., Naikoo, N. B., Kanth, R. H., Bahar, F. A., Bhat, M. A., Nazir, A., ... & Ahngar, T. A. (2022). Vertical farming: The future of agriculture: A review. *Pharma Innov. J, 11*, 1175-1195.
9. Balashova, I., Sirota, S., & Pinchuk, Y. (2019, November). Vertical vegetable growing: creating tomato varieties for multi-tiered hydroponic installations. In *IOP Conference Series: Earth and Environmental Science* (Vol. 395, No. 1, p. 012079). IOP Publishing.
10. Ranawade, P. S., Tidke, S. D., & Kate, A. K. (2017). Comparative cultivation and biochemical analysis of Spinacia oleraceae grown in aquaponics, hydroponics and field conditions. *International Journal of Current Microbiology and Applied Science*, *6*(4), 1007-1013.
11. Sankhalkar, S., Komarpant, R., Dessai, T. R., Simoes, J., & Sharma, S. (2019). Effects of soil and soil-less culture on morphology, physiology and biochemical studies of vegetable plants. *Current Agriculture Research Journal*, *7*(2), 181.
12. KHAN, F. A. (2018). A review on hydroponic greenhouse cultivation for sustainable agriculture. *International Journal of Agriculture Environment and Food Sciences*, *2*(2), 59-66.
13. Savvas, D. (2003). Hydroponics: A modern technology supporting the application of integrated crop management in greenhouse.
14. Eigenbrod, C., & Gruda, N. (2015). Urban vegetable for food security in cities. A review. *Agronomy for Sustainable Development*, *35*, 483-498.
15. Hayden, A. L. (2006). Aeroponic and hydroponic systems for medicinal herb, rhizome, and root crops. *HortScience*, *41*(3), 536-538.
16. Rouphael, Y., Colla, G., Battistelli, A., Moscatello, S., Proietti, S., & Rea, E. (2004). Yield, water requirement, nutrient uptake and fruit quality of zucchini squash grown in soil and closed soilless culture. *The Journal of Horticultural Science and Biotechnology*, *79*(3), 423-430.
17. Schmilewski, G. (2007, September). Growing medium constituents used in the EU. In *International Symposium on Growing Media 2007 819* (pp. 33-46).
18. Lee, S., & Lee, J. (2015). Beneficial bacteria and fungi in hydroponic systems: Types and characteristics of hydroponic food production methods. *Scientia Horticulturae*, *195*, 206-215.
19. Kim, H. Y. (2014). Statistical notes for clinical researchers: Two-way analysis of variance (ANOVA)-exploring possible interaction between factors. *Restorative dentistry & endodontics*, *39*(2), 143-147.
20. Singh, R., Hussain, S. A. I., Dash, A., & Rai, R. N. (2020). Modelling and optimizing performance parameters in the wire-electro discharge machining of Al5083/B 4 C composite by multi-objective response surface methodology. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, *42*, 1-32.