Chapter 6

Optimization of banana peel waste based microbial fuel cells by machine learning

Optimization of banana peel waste based microbial fuel cells by machine learning

6.1 Introduction

Fruit waste refers to non-edible/rotten or indigestible parts that are generated during harvesting, handling, transport and food supply chain [9]. Fruit waste has high content of lignocellulose which can be easily used in the generation of renewable energy [10]. Therefore, there is an immediate need to implement technologies which can degrade fruit waste with simultaneous generation of renewable energy [11]. Such technologies are also expected to halt environmental issues like greenhouse gas emissions (methane, carbon dioxide) and soil leaching [12]. Various efforts have been made in past to obtain an efficient renewable energy source from the fruit waste. Treating fruit waste is a challenging task because it has a complex composition, high moisture and organic content [11], [13]. Disposing, burning and composting of fruit waste impact environment negatively [14]. Due to environmental contamination and insufficient resource recovery, the afore mentioned disposal methods are not appropriate for managing fruit waste [11]. MFC system is known for its optimistic potentials such as the compatibleness, compactness and cleanliness [15]. MFC is also a convenient technology for waste resource recovery, wastewater remediation along with power generation. MFC is a bio electrochemical device that uses electrons originating from the anaerobic oxidation of organic compounds to generate electricity [16]–[18]. Many studies have shown the use of convenient organic/inorganic materials found in wastewater as a substrate for power generation [20]. The use of waste as a source of energy offers two benefits: firstly, it helps to manage waste and secondly, the power generated is highly cost effective. Many studies have shown that agricultural wastes are oxidised by various microorganisms, which lead to production of energy [21]. MFC is an innovative way for the valorisation of fruit waste biomass. Moreover, waste management strategies promote the recovery of valuable by-products and energy from waste biomass prior to its disposal. Various studies have explored the potential of fruit waste such as

lemon [22], lime [22], orange [22], [23], tangerine [22] and blueberry waste [24] as substrate in MFC. Banana peel waste (BW) is also one of the such wastes that is available in large amount. However, it has been less explored in MFC operations. Bananas (*Musa paradisiaca*) are abundantly accessible in tropical and subtropical zones across 130 nations. In 2019, 116 million tonnes of bananas were produced worldwide, and the fruits was available all-round the year. The typical fruit weighs around 125 grams, of which 75% is water and 25% is dry substance. China produced 11.6 million tonnes and India 30.4 million tonnes of bananas in 2019. Indonesia was third with 7.2 million tons, followed by Brazil and Ecuador with 6.8 and 6.5 million tonnes, respectively [455]. BW is available throughout the year in India and neighbouring countries. The massive quantity of BW originates from fruit markets and fruit processing industries every day. This shows that managing fruit peel waste is huge project for the environmental safety. BW comprises 30 - 40% weight out of whole weight of banana fruit. It is made up of lignin (5-10 %), hemicellulose (6 - 8 %) and cellulose (60 - 65 %) [12]. Characterization of BW powder revealed that it has a notable quantity of starch (41 % w/w), protein (8.4 % w/w) and cellulose (9.3 % w/w) [11]. Ripened BW holds up to 30% free sugars. It also has calcium, iron, phosphate, magnesium, zinc, copper, potassium and manganese [456]. BW are abundant in phenolic compounds, essential macro and micronutrients. These nutrients help in microbial growth [456]. Thus, BW can be easily utilised as value-added bioresource instead of direct dumping which leads to environmental issues. The concept of utilizing lignocellulosic substrates in the form of hydrolysates or solids to generate bioenergy have been also explored [337], [424]. Since ages, yeast has been applied in various biotechnological processes of the food industry. Such food industry processes generate huge amounts of wastewater rich in yeast and other organic components. The yeast has been used in several production processes, can also be applied as anode biocatalysts in MFC for treating wastewaters emanating from food industries [16]. *Saccharomyces cerevisiae* (baker's yeast)

has been applied as an anode biocatalyst in MFC as it is non-pathogenic, easy to process in extreme environmental conditions [16]. Alcohol is valuable as fuel if extracted as a by-product of yeast and BW based MFC [21], [325], [337]. Previously, *S. cerevisiae* as an anode biocatalyst has been explored with the banana and sweet lemon peel waste [325], [337]. However, optimization of yeast growth parameters by machine learning (ML) has been scarcely reported in the case of fruit peel waste. Power generation in MFC is basically linked to the growth of electrogenic microorganisms and extracellular electron transfer rate [16]. Growth of anodic microbial community is affected by various parameters such as temperature, pH and availability of nutrients [25]. It is necessary to be aware of the physiological response of microorganism towards the operating conditions. By providing the correct combinations of culture conditions, the power generation efficiency can be improved. Temperature is a vital aspect for microbial growth as it influences metabolic actions like enzyme activity, respiration and biomass productivity [26]. pH of the anolyte drastically influences the microbial physiology. It plays a significant role in the bioelectricity generation from MFC [27]. Additionally, temperature, pH and substrate concentration also impact the growth of microbial population and influences MFC performance [28]. Optimum concentration of substrate varies with the microbial diversity. Hence, it is difficult to determine optimal concentration range of substrate. The optimization becomes vital to reduce operational cost. The traditional response surface methodology (RSM) technique to optimize input variable of MFC has been widely used [29]–[34]. A number of optimization methods have been also developed by ML. ML can be endorsed to replace the conventional RSM technique. RSM analysis fits data to a second-order polynomial. RSM optimization works well only if curvature can be fitted to second-order polynomial. Besides this, RSM estimates error using least squares [35]. RSM requires selecting acceptable operational parameter ranges and restricts optimisation to specified scales [36], [37]. RSM cannot work with larger models and predict future outcomes for a system operated outside

the range [36], [37]. It receives poor optimization results when operated with large range of responses [36], [37]. ML models deliver a new way to evaluate the response surface function which is problematic in the conventional RSM method [35]. ML needs the pre-acquisition of a massive volume of experimental data (size of samples should be greater than the extent of coefficients to be evaluated) [35]. ML algorithms promptly evaluate a wide-range of dataset and reveal the most reasonable combinations of predictor variables for increasing current and power density. Patterns spotted using ML algorithms can contribute in building highly productive experimental designs [35]. There is a scarcity of literature where the decision tree regression model has been applied to optimize a MFC with *S. cerevisiae* as electrogenic inoculum and BW as a feed for the microbial growth. The present chapter focus on the application of a decision tree regression algorithm to optimise the operational parameters of an MFC utilising *S. cerevisiae* as anode biocatalyst and BW slurry as substrate. The parameters optimised were temperature (T), slurry concentration (SC), initial pH, pre-treatment (P), and external resistance (R). The construction of the dataset was achieved through the execution of experiments in accordance with the central composite design (CCD) methodology. Subsequently, the dataset was analysed utilising a ML algorithm, specifically the decision tree. The experimental validation of these combinations was conducted to assess the ultimate accuracy of the decision tree. The model generated in the present chapter can be implemented directly by the other researchers.

6.2 Materials and Methods

6.2.1 Substrate Preparation

BW was collected from a regional fruit juice shop situated in the institute's campus. BW used in MFC was processed as BW slurry. The fresh BW was rinsed and ground to fine thick paste in a blender. One portion of thick BW paste was pre-treated with 3% (w/v) NaOH solution in

sealed 2 litre erlenmeyer flasks in an autoclave at 121°C and 15 psig [411]. Remaining portion of BW slurry was kept untreated. Distilled water was added to the pre-treated and untreated BW paste to adjust the concentration of BW slurry as per experimental layout.

6.2.2 Anode Biocatalyst

S. cerevisiae (Kothari's instant yeast make, India) was applied as an external inoculum at anode. *S. cerevisiae* inoculum was prepared with yeast potato dextrose (YPD) medium. 1% (V/V) inoculum was added into BW slurry for MFC operation. Physiochemical characteristics of inoculum source have been shown in Table 6.1. There was possibility that indigenous bacteria were present in both pre-treated and untreated BW slurry [337].

Table 6.1. Physiochemical properties of inoculation source

Physiochemical characteristics	Values
pH	5.21
Reducing sugar (g/L)	24.56 ± 6.0
yeast plate count (CFU/ml)	3.5×10^{1}
Ash (%)	0.37 ± 0.015
Total dissolve solid (g/L)	11.4 ± 0.36

6.2.3 Design of MFC

H-shaped MFC (Figure 6. 1) were constructed using two 500 mL cylindrical plastic vessels. 12 such MFC were constructed. Each of them has analogous cathode and anode features in terms of material and surface area. One notable benefit of utilising a two-chamber MFC as compared to a single-chamber MFC is its ability to optimise cathode performance. This can be done by regulating pH, oxygen purging, flow rate augmentation and the introduction of electron-mediators. These measures enhance the overall performance of the MFC [457].

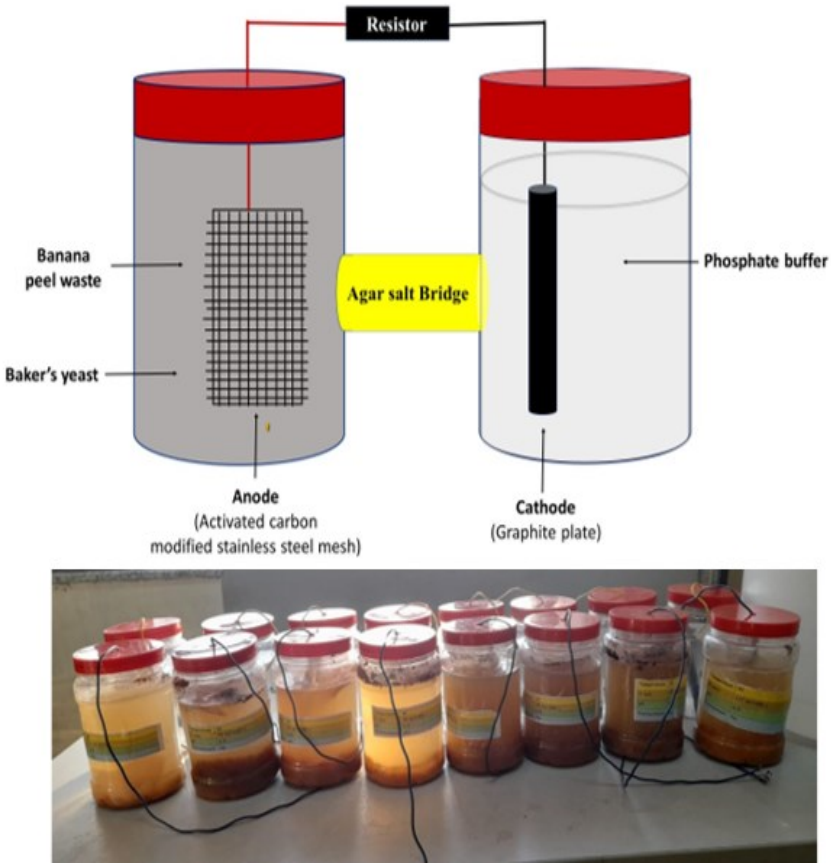

Figure 6. 1. Graphical illustration of H shaped dual-chamber MFC setup

A binder free customization was done over the surface of rectangular stainless steel mesh anode using activated charcoal (Himedia, make India) [337]. Anode has surface area of 60 cm^2 (13 mesh size). Cylindrical graphite rods of radius 0.5 cm, and length 10 cm were used as cathode. Agar salt bridge was used as a separator between anode and cathode chamber. In MFCs, the salt-bridge provides a more cost-effective alternative over proton-exchange membrane, which is notoriously expensive. In an agar salt bridge, the rate of oxygen transport is extremely low that it is almost inaccessible. As compared to MFC that uses cation exchange membrane as a separator, salt bridge has a much lower oxygen permeability. This results in a much larger coulombic efficiency [418].

6.2.4 Experimental Layout and Data Generation

The data for ML analysis was generated through experiments. Voltage, maximum power density and current density were considered as output variables. The input parameters were

initial pH, P/NP, SC, T and R were input variables. All input variables were optimized using decision tree analysis (MATLAB R2017b). Table 6.2 shows all the ranges of predictor variable used during the experiments.

Table 6.2. Predictor variables

Predictor Variable	Range
BW slurry concentration (mL/L)	32.95 to 117.04
Temperature (ºC)	19.88 to 45.11
pH	3.97 to 9.02
Pretreatment (P/NP)	Absent (NP) /Present (P)
Resistance (Ω)	15000 to 10

The experiments were designed using CCD for optimization (MATLAB 2017b) as shown in Table 6.3. All experiments were extended by applying 12 different R (15000 to 10 Ω) that generated dataset containing 480 data entries (40 ×12).

Table 6.3. Design of experiment

Run Order	SC (mL/L)	T(ºC)	pH	P/NP
1	100	25	8	NP
2	75	32.5	3.97	NP
3	75	32.5	6.5	NP
4	75	19.88	6.5	P
5	50	25	5	NP
6	75	32.5	6.5	P
7	75	32.5	9.02	NP

8	75	45.11	6.50	P
9	50	40	8	P
10	50	25	8	NP
11	100	25	5	NP
12	75	32.5	6.5	NP
13	100	40	5	NP
14	75	32.5	6.5	P
15	75	32.5	6.5	NP
16	75	32.5	6.5	P
17	75	32.5	6.5	P
18	75	32.5	6.5	P
19	75	19.88	6.5	NP
20	100	25	8	P
21	75	32.5	3.97	P
22	50	40	5	NP
23	100	40	5	P
24	75	32.5	6.50	P
25	117.04	32.5	6.50	P
26	75	32.5	9.02	P
27	50	40	8	NP
28	75	45.11	6.50	NP
29	75	32.5	6.50	NP
30	50	25	5	P
31	117.04	32.5	6.50	NP
32	75	32.5	6.50	NP

33	100	25	5	P
34	32.95	32.5	6.50	NP
35	50	25	8	P
36	50	40	5	P
37	100	40	8	NP
38	32.95	32.50	6.50	P
39	100	40	8	P
40	75	32.50	6.50	NP

The temperature was varied in a bench top shaker-cum-incubator (Model CIS-18 Plus, Remi make, India). pH was varied by adding 1N HCl / NaOH and determined by using digital pH meter (Eutech pH Tutor make, US). Dual-chambered H-shaped MFC (300 mL working volume) were constructed and run as illustrated in Figure 6. 1. Different concentrations of BW slurry (as per experimental layout) were applied as substrate and inoculated with *S. cerevisiae* inoculum at anode chamber. Copper wires were used as connectors between electrodes. 0.1 M phosphate buffer ($Na_2HPO_4.7H_2O$ and $NaH_2PO_4.H_2O$) of pH 7.4 was used as catholyte. 12 sets of above-mentioned H-type MFC were prepared to conduct experiment according to experimental designs. 12 sets of MFC (according to 12 different R) were operated for a period 15 days (including 5 days of acclimation) for each run order (Table 6.3). MFCs were connected to R varying from 15000 to 10 Ω in set of 12 MFC. In order to validate the accurate selection of the external resistances 15000 to 10 Ω were observed. It has been observed that when MFC was loaded by an external resistance, the closed-circuit voltage was lower than the open circuit voltage, but MFC recovered gradually. If it is not recovered, a larger resistance should be used. In case of no substantial drop in voltage, a lower resistor should be used in the open circuit voltage [393]. A multimeter (Mextech Mas 830l make, India) was used to record voltage.

Stable voltage generation was noticed after 5 days of acclimatization. When the voltage was constant for few hours, voltage values were observed from each MFC continuously for 10 days. Only maximum voltage values obtained (during stable voltage generation) from each MFC were considered in this chapter. Voltage was recorded across each resistance. Current (I) was determined using Eq. 6.1:

$$I = \frac{V}{R} \tag{6.1}$$

where, V represents the voltage (V), R represents resistance (Ω) and I denotes the current (A). Power (P) is evaluated by using Eq. 6.2

$$P = I.V \tag{6.2}$$

To calculate power density (mW/m²), P was divided by surface area (SA) of anode as per Eq. 6.3 to obtain power density [325].

$$\text{Power Density} = \frac{P}{SA} \tag{6.3}$$

The current density (mA/m^2) was determined by Eq. 6.4 [325].

$$\text{Current density} = \frac{I}{SA} \tag{6.4}$$

6.2.5. Decision Tree Regression, Computational and Experimental Validation

Various combinations of predictor variables (T, pH, SC, P/NP and R) ruling to low and high values of output variables were identified by the decision tree regression algorithms. The MATLAB (version 9.2, R2017b) was exploited to execute decision tree regression. Decision trees were built by using "fitrtree" function with the CART algorithm. Optimal tree structure was developed as preliminary step by optimizing the hyperparameters like split criterion, NumVariablesToSample, MinLeafSize and MaxNumSplits. These hyperparameters were optimized by applying the "OptimizeHyperparameters" and choosing value ‘all’.

"OptimizeHyperparameters" option operates by shrinking the cross-validation error. The complete dataset of 480 entries was split into three subsets randomly for training, testing and final validation. Dataset splitting was done via 'cvpartition' and 'HoldOut' function. First subset was called training subset containing 360 entries. Second subset had 90 entries and considered as testing data. For computational validation, 31 entries were chosen randomly for computational validation. In the beginning, different tree structures were constructed using the training subset. Later, the tree having minimal cross-validation error was chosen to test the data from the testing subset to foresee the accuracy. None of the data were taken from the testing subset for the training step of decision tree. The "predictorImportance" function was applied to ascertain the most prominent predictor variables resulting in various values of output variables. The final regression value of the models was determined by 'Evaluation.accuracy' function. The ultimate computational validation of the decision tree models was accomplished by testing them with validation dataset (31 entries).

6.2.6 Software and Statistical Analysis

The performance of developed models was evaluated on the basis of regression coefficient (R^2) for the test data. The regression coefficient (R^2) is shown in Eq. (6.5):

$$R^2 = 1 - \frac{\sum_{i=1}^{N}\left(y_i^{exp} - y_i^{pred}\right)^2}{\sum_{i=1}^{N}\left(y_i^{exp} - y_{av}^{exp}\right)^2} \tag{6.5}$$

where, y_i^{exp} and y_i^{pred} are the experimental and predicted values. y_{av}^{exp} is the average of the experimental values. Validation error was calculated by Eq. 6.6:

$$\text{Validation Error (\%)} = \frac{\text{(Experimental Value – Predicted Value)}}{\text{Predicted Value}} \times 100 \tag{6.6}$$

6.3 Results and Discussion

6.3.1 Decision Tree Evaluation

In this chapter, three distinct decision trees were developed, each for voltage, power and current density. The rules directing to low and high value of output variables were revealed via these

decision trees. The training dataset was applied to build the tree and the testing dataset was employed to verify the accuracy of tree. 'OptimizeHyperparameters' function was applied to optimize the hyperparameters and to expand the trees.

6.3.1.1 Voltage

The optimal decision tree for voltage is shown in Figure 6. 2. The first partition of the tree was obtained according to the T (root of the tree). Set of rules with high voltage generation (>300 mV) were represented by green nodes. The red nodes denote the set of rules with lower voltage generation.

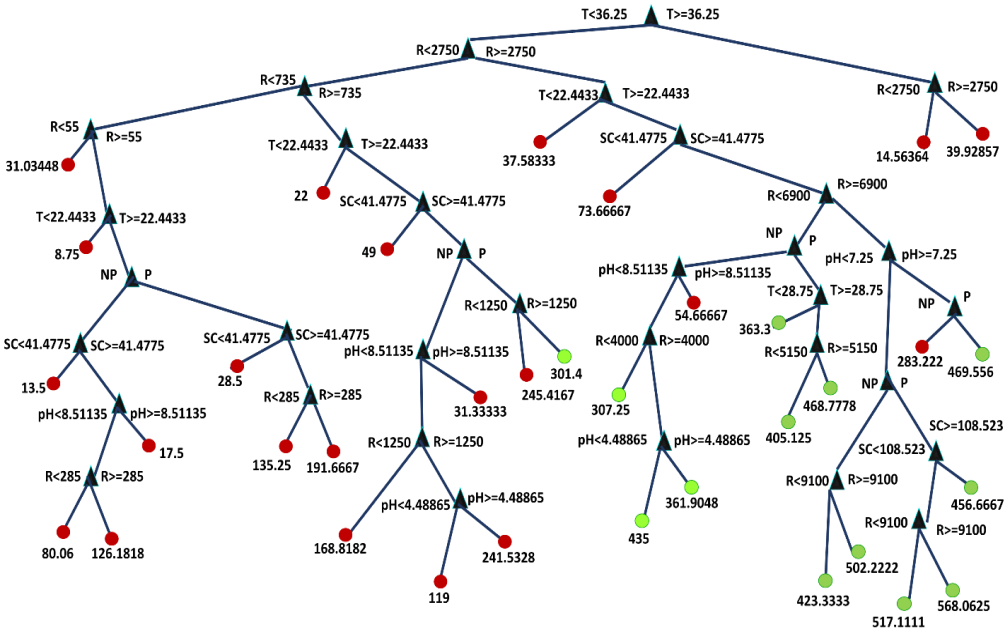

Figure 6. 2 Decision tree regression for voltage generation (in mV)

When T < 36.25 °C, the second level of branching was provided by R (2750 Ω) (LHS of figure 2). There were total 25 paths out of 84 set of rules that yielded high voltage generation (> 300 mV).

Table 6.4. Different combinations of input variables and their predicted value for voltage extracted form decision tree model.

Variable Range	Predicted Value (Voltage mV)
T < 22.44; NP; R<55	1 ± 0
T < 22.44; 735<=R<2750	22 ± 4.51
T < 22.44; 55<=R<735	8.75 ± 5.73
SC<41.47; 22.44<=T<36.25	49 ± 15.11
T<36.25; pH>=7.25; P; R<55	19.33 ± 10.96
SC<41.47; 22.44<=T<36.25; NP; R<55	9 ± 0
SC<41.47; T<36.25; pH<7.25; P; R<55	6 ± 0
SC<41.47; 22.44<=T<36.25; NP; 55<=R<735	13.5 ± 4.94
SC<41.47; 22.44<=T<36.25; P; 55<=R<735	28.5 ± 14.84
41.47<=SC; 22.44<=T<36.25; pH>=8.51; NP; 735<=R<2750	31.33 ± 5.13
SC>41.47; 22.44<=T<36.25; pH>=8.51; NP; R<55	10 ± 0
41.47<=SC<108.5; T<36.25; pH<7.25; P; R<55	46.44 ± 10.26
SC>=108.5; T<36.25; pH<7.25; P; R<55	58 ± 0
SC>=41.47; 22.44<=T<36.25; pH>=8.51; NP; 55<=R<735	17.5 ± 2.12
SC>=41.47; 22.44<=T<28.75; P; 735<=R<1250	213.66 ± 23.18
SC>=41.47; 28.75<=T<36.25; P; 735<=R<1250	256 ± 49.71
41.47<=SC; 22.44<=T<28.75; P; 1250<=R<2750	271.62 ±30.69
SC>=41.47; 22.44<=T<36.25; pH<5.75; NP; R<55	33.33 ± 6.11
SC>=41.47; 22.44<=T<36.25; 5.75<=pH<8.51; NP; R<55	26.66 ± 8.35
SC>=41.47; 22.44<=T<36.25; pH<7.25; P; 55<=R<285	149 ± 45.04

SC>=41.47; 22.44<=T<36.25; pH>=7.25; P; 55<=R<285	94 ± 5.19
SC>=41.47; 22.44<=T<36.25; pH<7.25; P; 285<=R<735	199 ± 40.77
SC>=41.47; 22.44<=T<36.25; pH>=7.25; P;285<=R<735	169.66 ± 25.81
SC>=108.5; 22.44<=T<36.25; pH<8.51; NP; 735<=R<1250	124 ± 0
SC>=41.47; 22.44<=T<36.25; pH<4.48; NP; 1250<=R<2750	119 ± 0
SC>=41.47; 28.75<=T<36.25; P; 1250<=R<1850	297.88 ± 43.17
SC>=41.47; 28.75<=T<36.25; P; 1850<=R<2750	335.12 ± 55.59
41.47<=SC<87.5; 22.44<=T<36.25; pH<8.51; NP; 55<=R<285	75 ± 44.45
SC>=108.5; 22.44<=T<36.25; pH<8.51; NP; 285<=R<735	41 ± 0
41.47<=SC<87.5; 22.44<=T<36.25; pH<8.51; NP; 55<=R<285	179.12 ±72.21
41.47<=SC<108.5; 22.44<=T<36.25; 7.25<=pH<8.51; NP; 735<=R<1250	150 ± 11.31
41.47<=SC; 22.44<=T<36.25; 7.25<=pH<8.51; NP; 1250<=R<2750	197.75 ± 11.87
41.47<=SC<108.5; 22.44<=T<36.25; pH<7.25; NP; 285<=R<735	141.62 ± 71.70
41.47<=SC<108.5; 22.44<=T<36.25; 7.25<=pH<8.51; NP; 285<=R<735	107 ± 7.07
41.47<=SC; 22.44<=T<36.25; 4.48<=pH<7.25; NP; 1250<=R<1850	237.75 ± 73.15
41.47<=SC; 22.44<=T<36.25; 4.48<=pH<7.25; NP; 1850<=R<2750	264.33 ± 82.23
T<22.44; 2750<=R<6900	34 ± 2.60
T<22.44; 6900<=R	41.16 ± 2.22
SC<41.47; 22.44<=T<36.25; NP; 2750<=R	54.83 ±8.63
SC<41.47; 22.44<=T<36.25; P; 2750<=R	92.5 ± 16.56
41.47<=SC; 22.44<=T<36.25; 8.51<=pH; NP; 2750<=R<6900	54.66 ± 8.02
41.47<=SC; 22.44<=T<36.25; 7.25<=pH; NP; 6900<=R	283.22 ± 161.07
41.47<=SC; 22.44<=T<36.25; 7.25<=pH; P; 6900<=R	469.55 ± 54.49

41.47<=SC; 22.44<=T<28.75; P; 2750<=R<4000	319.66 ± 20.13
41.47<=SC; 22.44<=T<28.75; P; 4000<=R<6900	382 ± 23.90
41.47<=SC; 28.75<=T<36.25; P; 5150<=R<6900	468.77 ± 67.22
41.47<=SC; 22.44<=T<36.25; pH<7.25; NP; 6900<=R<9100	423.33 ± 61.80
108.5<=SC; 22.44<=T<36.25; pH<7.25; P; 6900<=R	456.66 ± 43.06
41.47<=SC; 22.44<=T<36.25; 7.25<=pH<8.51; NP; 2750<=R<4000	279.5 ± 30.40
41.47<=SC; 22.44<=T<36.25; pH<4.48; NP; 4000<=R<6900	435 ± 2.82
96.0225<=SC; 28.75<=T<36.25; P; 2750<=R<5150	339 ± 0
41.47<=SC; 22.44<=T<36.25; pH<5.75; NP; 9100<=R	532.66 ± 25.97
41.47<=SC; 22.44<=T<36.25; pH<7.25; P; 6900<=R<9100	517.11 ±63.98
41.47<=SC; 22.44<=T<36.25; pH<5.75; NP; 2750<=R<4000	303.33 ± 5.50
41.47<=SC; 22.44<=T<36.25; 5.75<=pH<7.25; NP; 2750<=R<4000	316.85 ± 80.47
41.47<=SC<96.0225; 28.75<=T<36.25; P; 2750<=R<4000	390.5 ± 57.00
41.47<=SC<96.0225; 28.75<=T<36.25; P; 4000<=R<5150	431.28 ± 66.39
96.0225<=SC; 22.44<=T<36.25; 5.75<=pH<7.25; NP; 9100<=R	521 ± 2.82
41.47<=SC<62.5; 22.44<=T<36.25; pH<7.25; P; 9100<=R	518.5 ± 9.19
41.47<=SC<108.5; 22.44<=T<36.25; 4.48<=pH<8.51; NP; 4000<=R<5150	335.44 ± 65.87
108.5<=SC; 22.44<=T<36.25; 4.48<=pH<8.51; NP; 4000<=R<5150	403 ± 0
41.47<=SC; 22.44<=T<36.25; 4.48<=pH<7.25; NP; 5150<=R<6900	386 ± 60.54
41.47<=SC; 22.44<=T<36.25; 7.25<=pH<8.51; NP; 5150<=R<6900	352 ± 31.11
41.47<=SC<96.0225; 22.44<=T<36.25; 5.75<=pH<7.25; NP; 9100<=R<12500	476.5 ± 39.84
41.47<=SC<96.0225; 22.44<=T<36.25; 5.75<=pH<7.25; NP; 12500<=R	482.66 ± 34.40

62.5<=SC<108.5; 22.44<=T<36.25; pH<7.25; P; 9100<=R<12500	556.85 ± 46.04
62.5<=SC<108.5; 22.44<=T<36.25; pH<7.25; P; 12500<=R	593.42 ± 16.59
42.55<=T; R<735	15.16 ± 3.97
36.25<=T; 1850<=R<2750	24.77 ± 4.91
36.25<=T; 5150<=R<6900	38.25 ± 4.83
36.25<=T<42.55; 285<=R<735	11.85 ± 1.86
SC<62.5; 36.25<=T; 735<=R<1850	15.85 ± 3.33
36.25<=T; pH<5.75; 2750<=R<5150	35.12 ± 3.83
36.25<=T; pH<5.75; 6900<=R<12500	47.62 ± 4.20
36.25<=T; pH<5.75; 12500<=R	56 ± 6.16
36.25<=T; pH>=5.75; 6900<=R<9100	38.5 ± 2.07
36.25<=T<42.55; R<55	4 ± 2
36.25<=T<42.55; 55<=R<285	7.5±2.44
62.5<=SC; 36.25<=T; NP; 735<=R<1850	18.33 ± 3.26
62.5<=SC; 36.25<=T; P; 735<=R<1850	22.75 ± 2.06
36.25<=T; 5.75<=pH; 2750<=R<4000	28.83 ± 3.06
36.25<=T; 5.75<=pH; 4000<=R<5150	32.66 ± 3.38
36.25<=T; 5.75<=pH; 9100<=R<12500	42.4 ± 1.67
36.25<=T; 5.75<=pH; 12500<=R	46.4 ±2.96

Decision tree regression was pruned up to 46 levels and there were 13 set of rules that resulted in high voltage generation as shown in Figure 6. 2. Pruning was done to reduce the overfitting of the data and to better visualize the tree. It did not affect the accuracy of tree. In all 84 paths,

two paths delivered the most generalised set of rules directing to high voltage generation (Table 6.4). In the first path, SC, T and R ranged from 62.5 to 108.52 mL/L, 22.44 to 36.25 °C and 9100 to 12500 Ω. Other conditional parameters were pH < 7.25 and P. For the second path, SC and T ranged from 62.5 to 108.52 mL/L and 22.44 to 36.25 °C. Other conditional parameters were pH < 7.25, R ≤ 12500 Ω and P. Both paths were supported by 7 data points and showed voltage around 556.85 ± 46 and 593.42 ± 16 mV respectively. When T > 36.25 °C, the second level of branching was provided by R (2750 Ω) (RHS of Figure 6. 2). Right side of decision tree resulted in total 17 set of rules (Table 6.4). Decision tree regression was pruned to 46 levels and ultimately there were 2 set of rules remained after pruning (Figure 6. 2). All 17 paths showed very low voltage generation (< 60 mV) and the maximum voltage 56 ± 6.16 mV was obtained from path condition where T > 36.25 ºC, pH < 5.75 and R ≥ 12500 Ω. Overall, the most decisive path directing to high voltage generation was provided by the left side, where SC and T ranged from 62.5 to 108.52 mL/L and 22.44 to 36.25 ºC, respectively. Other conditional parameters were pH < 7.25, R ≤ 12500 Ω and slurry was P. This path generated voltage of 593.42 ± 16.59 mV and was supported by 7 data points. The final regression value of the model was 0.98. Predictor variables like T, R, SC, initial pH and P highly influenced the voltage generation in MFC. Figure 6. 3 shows relative importance of the predictor variables on the voltage generation.

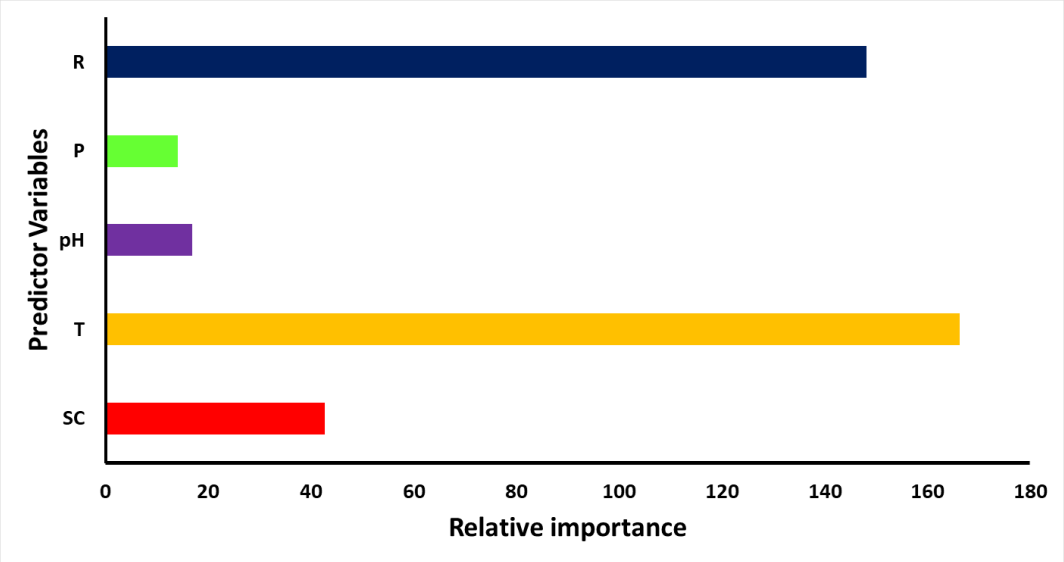

Figure 6. 3. Relative importance of the predictor variables on voltage generation

T was highly influential factor in this study followed by R. SC influenced the voltage generation moderately. However, initial pH and P of slurry had least influence on the voltage generation. Researchers performed many studies on the impact of temperature and attempted to optimize it for attaining high voltage generation [458]–[460]. Very low T values showed longer start-up time due to the following reasons [461]: i. inhibition of yeast cell activity ii. It is diminished growth rate of yeast results in poor microbial attachment at anode surface and iii. increased internal resistance. The effects of operating parameters on the performance of yeast based MFC was investigated [461]. Authors reported that voltage generation increased with the surge in T from 25 to 35 °C. It was also reported that increase in T beyond 40 °C led to decline in the voltage [461]. The impact of the R on the anode biocatalyst and energy production is a critical factor in MFC performance [442]. The reduced level of R value results in generation of sluggish voltage [443]. The R can alter the ratio of cell voltage to the current. Likewise, the higher values of R led to an enhanced voltage. However, it dropped the current generation and vice versa [443]. Present study also showed similar trend for voltage generation against different R values. Optimizing R is also a way to reduce losses during power generation in MFC [462]. Researchers monitored the impact of R values on the voltage generation and

observed that the voltage values escalated with the increase in of R (0 to 4,000 Ω) [463]. Authors achieved the maximum voltage of 358 mV at 4,000 Ω [463].

6.3.1.2 Current density

There were total 5 paths out of 75 set of rules that led to high current density (> 250 mA/m^2).

Table 6.5. Different combinations of input variables and their predicted value for current density extracted form decision tree model

Variable Range	Predicted Value (Current density (mA/m^2))
T<22.44; R<55	41.66 ± 35.35
36.25<=T<42.55; R<55	66.66 ± 33.33
42.55<=T; R<55	175 ± 35.35
22.44<=T<36.25; pH>=7.25; R<55	322.22 ± 152.99
22.44<=T<36.25; pH<7.25; P; R<55	774.07 ± 171.00
SC<41.47; 22.44<=T<36.25; pH<7.25; NP; R<55	150 ± 0
41.47<=SC<87.5; 22.44<=T<36.25; pH<7.25; NP; R<55	460.41 ± 140.55
87.5<=SC; 22.44<=T<36.25; pH<7.25; NP; R<55	591.66 ± 106.06
36.25<=T; 55<=R<285	15 ± 7.54
T<36.25; pH<5.75; NP; 55<=R<285	155.55 ± 37.05
T<22.44; P; 55<=R<285	10 ± 0
T<22.44; pH>=5.75; NP; 55<=R<285	3.33 ± 0
SC<41.47; 22.44<=T<36.25; P; 55<=R<285	30 ± 0

SC<41.47; 22.44<=T<36.25; 5.75<=pH; NP; 55<=R<285	16.66 ± 0
41.47<=SC; 22.44<=T<36.25; 7.25<=pH; P; 55<=R<285	156.6 ± 8.66
41.47<=SC; 22.44<=T<36.25; 5.75<=pH<8.51; NP; 55<=R<285	113.51 ± 73.44
41.47<=SC; 22.44<=T<36.25; 8.51<=pH; NP; 55<=R<285	26.66 ± 0
41.47<=SC; 22.44<=T<28.75; pH<7.25; P; 55<=R<285	175.83 ± 10.60
41.47<=SC; 28.75<=T<36.25; pH<7.25; P; 55<=R<285	256.45 ± 75.91
36.25<=T; 735<=R<1850	2.65 ± 0.60
T<22.44; 285<=R<735	4.78 ± 0.25
T<22.44; 735<=R<1850	2.66 ± 0.50
36.25<=T<42.55; 285<=R<735	4.34 ± 0.72
42.55<=T; 285<=R<735	6.56 ± 0.75
SC<41.47; 22.44<=T<36.25; 735<=R<1850	5.98 ± 1.91
SC<41.47; 22.44<=T<36.25; NP; 285<=R<735	6.02 ± 0
SC<41.47; 22.44<=T<36.25; P; 285<=R<735	13.82 ± 0
41.47<=SC; 22.44<=T<36.25; 8.51<=pH; NP; 285<=R<735	6.73 ± 0
41.47<=SC; 22.44<=T<36.25; pH<7.25; P; 285<=R<735	70.33 ± 14.42
41.47<=SC; 22.44<=T<36.25; 7.25<=pH; P; 285<=R<735	60.16 ± 9.15
41.47<=SC; 22.44<=T<36.25; 8.51<=pH; NP; 735<=R<1850	3.91± 0.82
108.5<=SC; 22.44<=T<36.25; pH<8.51; NP; 285<=R<735	14.53 ± 0
41.47<=SC; 22.44<=T<36.25; pH<4.48; NP; 735<=R<1850	16.11 ± 4.08
41.47<=SC; 22.44<=T<28.75; P; 735<=R<1250	35.37 ± 3.18
41.47<=SC; 28.75<=T<36.25; P; 735<=R<1250	41.81 ± 8.42
41.47<=SC; 22.44<=T<28.75; P; 1250<=R<1850	27.94 ± 2.81
41.47<=SC; 28.75<=T<36.25; P; 1250<=R<1850	32.65 ± 4.92

41.47<=SC<108.5; 22.44<=T<36.25; pH<7.25; NP; 285<=R<735	48.97 ± 24.07
41.47<=SC<108.5; 22.44<=T<36.25; 7.25<=pH<8.51; NP; 285<=R<735	37.94 ± 2.50
108.5<=SC; 22.44<=T<36.25; 4.48<=pH<8.51; NP; 735<=R<1850	17.5 ± 4.47
41.47<=SC<108.5; 22.44<=T<36.25; 4.48<=pH<7.25; NP; 735<=R<1850	28.13 ± 10.07
41.47<=SC<108.5; 22.44<=T<36.25; 7.25<=pH<8.51; NP; 735<=R<1850	22.94 ± 2.61
T<22.44; 1850<=R<4000	1.75 ± 0.32
T<22.44; 4000<=R	0.85 ± 0.28
SC<41.47; 22.44<=T<36.25; 1850<=R<4000	3.46 ± 1.46
SC<41.47; 22.44<=T<36.25; 4000<=R<6900	2.16 ± 0.63
SC<41.47; 22.44<=T<36.25; 6900<=R	1.22 ± 0.44
SC>=41.47; 22.44<=T<36.25; pH>=8.51; NP; 1850<=R<4000	2.58 ± 0.30
41.47<=SC; 22.44<=T<28.75; P; 1850<=R<2750	22.10 ± 1.65
41.47<=SC; 28.75<=T<36.25; P; 1850<=R<2750	25.92 ± 4.25
41.47<SC; 22.44<T<28.75; P; 2750<=R<4000	16.12 ± 1.01
41.47<SC; 28.75<T<36.25; P; 2750<=R<4000	19.17 ±3.10
SC>=41.47; 22.44<=T<36.25; pH>=8.51; NP; 4000<=R<6900	1.89 ± 0.02
41.47<=SC; 22.44<=T<28.75; P; 4000<=R<6900	12.57 ± 1.01
SC>=41.47; 22.44<=T<36.25; pH>=8.51; 12500<=R	0.82 ± 0
SC>=41.47; 22.44<=T<36.25; pH<7.25; NP; 1850<=R<2750	17.61 ± 1.15

SC>=41.47; 22.44<=T<36.25; 7.25<=pH<8.51; NP; 1850<=R<2750	15.71 ±0.48
SC>=41.47; 22.44<=T<36.25; pH<7.25; NP; 2750<=R<4000	15.83 ± 3.44
SC>=41.47; 22.44<=T<36.25; 7.25<=pH<8.51; NP; 2750<=R<4000	14.11 ± 1.53
41.47<=SC; 22.44<=T<36.25; pH<4.48; NP; 4000<R<6900	14.18 ± 1.66
96.0225<=SC; 28.75<=T<36.25; P; 4000<=R<6900	11.47 ± 0.77
41.47<=SC; 22.44<=T<36.25; 7.25<=pH; NP; 6900<=R<12500	5.05 ±2.97
41.47<=SC; 22.44<=T<36.25; 7.25<=pH; P; 6900<=R<12500	8.55 ±1.19
41.47<=SC; 22.44<=T<36.25; pH<7.25; NP; 12500<=R	6.03 ± 0.56
41.47<=SC; 22.44<=T<36.25; 7.25<= pH<8.51; 12500<=R	4.82 ± 0.25
41.47<=SC<108.5; 22.44<=T<36.25; 4.48<pH<8.51; P; 4000<=R<6900	11.53 ± 2.13
108.5<=SC; 22.44<=T<36.25; 4.48<=pH<8.51; NP; 4000<=R<6900	13.35 ± 1.32
41.47<=SC; 28.75<=T<36.25; P; 4000<=R<5150	15.29 ± 2.35
41.47<=SC<96.0225; 28.75<T<36.25; P; 5150<=R<6900	14.33 ± 1.76
41.47<=SC; 22.44<=T<36.25; pH<5.75; NP; 6900<R<12500	9.21 ± 0.61
41.47<=SC; 22.44<=T<36.25; 5.75<=pH<7.25; NP; 6900<R<12500	8.23 ± 1.08
41.47<=SC; 22.44<=T<36.25; pH<7.25; P; 6900<=R<9100	10.16 ± 1.40
41.47<=SC; 22.44<=T<36.25; pH<7.25; P; 9100<=R<12500	8.94 ± 0.86
36.25<=T; 1850<=R<4000	1.75 ± 0.34
36.25<=T; 4000<=R	0.91 ± 0.26

Decision tree regression was pruned up to 35 levels, after pruning, there was 5 set of rules that led to high current density (Figure 6. 4). The optimal decision tree for current density is shown in Figure 6. 4, where R served as root of the tree. Set of rules with high current generation were depicted by green nodes whereas set of rules with lower current densities were represented by red nodes.

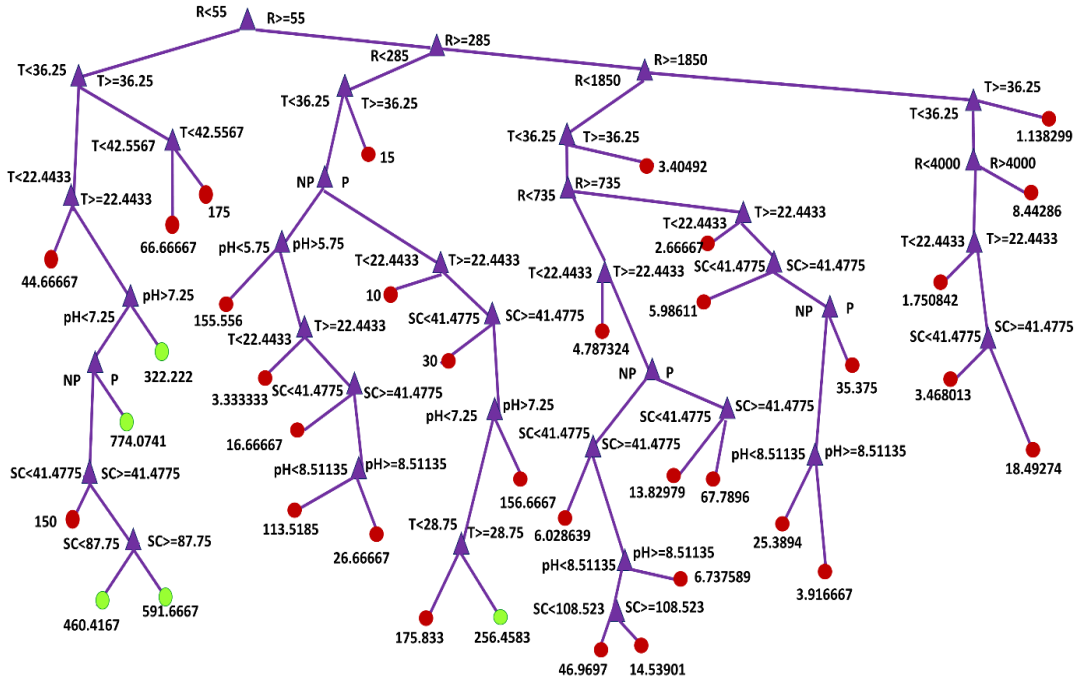

Figure 6. 4. Decision tree regression for current density (mA/m^2)

Out of total 5 set of rules resulting in high current densities, 4 set of rules were from the left side of the tree. Tree first splited the data according to the R (Ω). When R < 55 Ω, the second level of branching was provided by T (36.25 °C) (L.H.S. of Figure 6. 4). Other aspects that influenced the branching were SC, initial pH and slurry was P. Out of all the 5 best paths, the left-hand side of tree generated 4 paths that showed high current density (> 250 mA/m^2) and two of which generated current densities > 500 mA/m^2. First path reflected a current density of 774 ± 171 mA/m^2 when T ranged from 22.44 to 36.25 °C, pH < 7.25, R < 55 Ω and the slurry was P. Second path yielded a current density of 460.41 ± 140.55 mA/m^2 when SC and T ranged from 41.47 to 87.5 mL/L and 22.44 to 36.25°C, pH < 7.25, R < 55 Ω and slurry was

NP. The current density of 591.66 ± 106 mA/m^2 was observed in third path. The conditional parameters were as follows: SC ≥ 87.5 mL/L, T ranged from 22.44 to 36.25 °C, pH < 7.25, R < 55 Ω and the slurry was NP. The first and second path was supported by 9 and 8 data points respectively, whereas the third path was supported by only 2 data points. On the R. H. S of Figure 6. 4, when R ≥ 55 Ω, the second level of branching was offered by R (285 Ω ≤ R< 285 Ω) which was followed by other factors namely T, SC, initial pH and slurry was P/NP. Right branch of tree generated only one path that reported a current density of 256.4 ± 75.9 mA/m^2 (> 200 mA/m^2) with the condition when SC ≥ 41.47 mL/L and slurry was P. The pH < 7.25, T and R ranged from 28.75 to 36.25 °C and 55 to 285 Ω, respectively. This path was supported by 8 data points. The most conclusive set of rules directing to high current density generation was provided by the left branch of the tree where T ranged from 22.44 to 36.25 °C, pH < 7.25, R < 55 Ω and the slurry was P. This path provided a current density of 774 ± 171 mA/m^2 and was supported by 9 data points. The final regression value of the model was 0.98. Predictor variables like T, R, SC, initial pH and P/NP highly influenced the voltage generation in MFC. Figure 6. 5 shows relative importance of the predictor variables on the current generation. R was highly influential factor in this study followed by T.

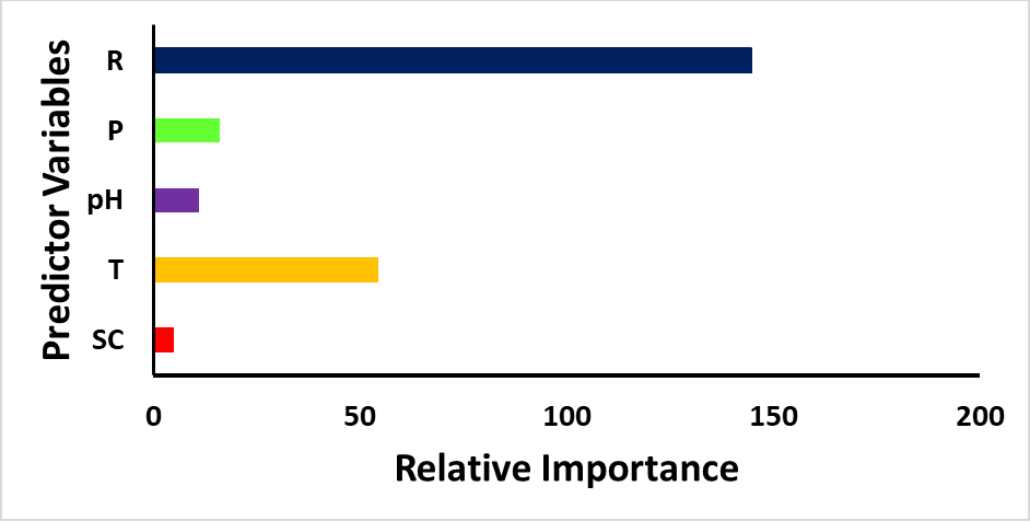

Figure 6. 5. Relative importance of the predictor variables on current density

T affects the current density moderately. The P, SC and initial pH had least influence on the current density. It has been observed that the bit-by-bit drop in the R led to an increase in the current generation with time [395]. The MFC performance with respect to R and applied loading rate was explored [464]. Authors applied R in range of 10.5 to 50 Ω and observed significant increase in the continuous current generation with the decrease in R. The kinetic and mass transfer limitations at lower R values were also noticed. Such limitations resulted in minimum steep descent of the current density [464]. The current generation was minimum and almost stable, when R was very high [463]. With the increase in T beyond 40 °C, a drop in current density was noticed [459]. However, the linear trend between current density and T was noticed at 35 °C which remained up to 45 °C. In the present study, the influence of T was not individually examined. Generally, the performance declines at high T values (about 40 °C and above) [459].

6.3.1.3 Power density

There were total 8 paths out of 32 set of rules that led to high power density (≥10 mW/m^2) (Table 6.6).

Table 6.6 Different combinations of input variables and their predicted value for power density extracted form decision tree model

Variable Range	Predicted Value (Power density (mW/m^2))
42.55<=T; R<285	1.16 ± 0.92
87.5<=SC; T<36.25; NP; R<285	20.07 ± 6.02
T<36.25; 7.25<=pH; P; R<285	11.16 ± 5.65

36.25<=T<42.55; R<55	0.35 ± 0.28
36.25<=T<42.55; 55<=R<285	0.10 ± 0.06
SC<87.5; T<22.44; NP; R<285	0.01 ± 0.00
SC<41.47; T<36.25; pH<7.25; P; R<285	0.57 ± 0.04
SC<41.47; 22.44<=T<36.25; NP; R<285	0.75 ± 0.83
41.47<=SC; T<28.75; pH<7.25; P; R<285	20.78 ± 14.56
41.47<SC<87.5; 22.44<=T<36.25; pH<8.51; NP; R<285	12.68 ± 12.29
41.47<SC<87.5; 22.44<=T<36.25; pH>=8.51; NP; R<285	1.04 ± 0.87
41.47<SC<87.5; 22.44<=T<36.25; pH<5.23; P; R<285	35.62 ± 2.65
41.47<SC<87.5; 22.44<=T<36.25; 5.23<=pH<7.25; P; R<285	41.50 ± 17.97
36.25<=T; 285<=R	0.04 ± 0.01
SC<41.47; T<36.25; NP; 285<=R,2750	0.13 ± 0.02
T<22.44; P; 285<=R<2750	0.05 ± 0.01
SC<41.47; T<36.25; NP; 2750<=R	0.07 ± 0.01
SC<41.47; T<36.25; P; 2750<=R	0.20 ± 0.04
SC>=41.47; T<22.44; 2750<=R	0.03 ± 0.00
SC>=41.47; T<22.44; NP; 285<=R<2750	0.05 ± 0.00
41.47<=SC; 22.44<=T<36.25; 7.25<=pH; NP; 285<=R<2750	2.53 ± 1.85
SC<41.47; 22.44<=T<36.25; P; 285<=R<2750	0.42 ± 0.08
41.47<=SC<62.5; 22.44<=T<36.25; P; 285<=R<2750	6.65 ± 1.33
SC>=62.5; 22.44<=T<36.25; P; 285<=R<735	14.26 ± 4.43
41.47<=SC; 22.44<=T<36.25; pH<8.51; NP; 2750<=R	4.02 ± 1.64
41.47<=SC; 22.44<=T<36.25; 8.51<=pH; NP; 2750<=R	0.10 ± 0.01
41.47<=SC; 22.44<=T<36.25; P; 2750<=R<6900	6.43 ± 1.84
41.47<=SC; 22.44<=T<36.25; P; 6900<=R	4.26 ± 1.26

41.47<=SC<108.52; 22.44<=T<36.25; pH<7.25; NP; 285<=R<2750	6.60±6.56
108.523<=SC; 22.44<=T<36.25; pH<7.25; NP; 285<=R<2750	2.12 ± 1.22
62.5<=SC<87.5; 22.44<=T<36.25; P; 735<=R<2750	10.30 ± 3.31
SC>=87.5; 22.44<=T<36.25; P; 735<=R<2750	7.70 ± 0.94

The optimal decision tree for power density is shown in Figure 6. 6. R served as the root at the top of decision tree. Set of rules leading to high power density generation have been depicted by green nodes whereas set of rules yielding to low power density are shown by red nodes.

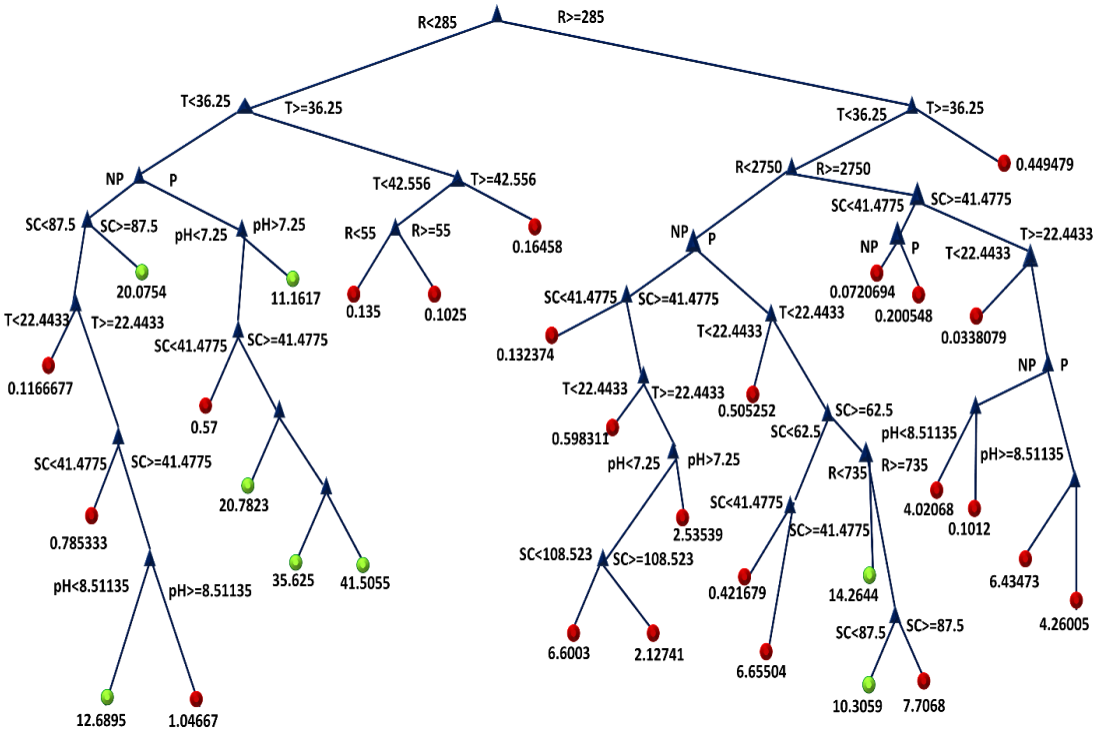

Figure 6. 6. Decision tree regression for power density (mW/m^2)

Pruning was not applied on power density's decision tree regression. Out of 8 set of rules that led to high power densities, 6 set of rules were obtained from left hand side of the tree. On the L. H. S of Figure 6. 6, the tree first splited the data according to the R (285 Ω). The second level of branching was offered by T (36.25 °C ≤ T <36.25 °C). Other factors that influenced the branching were SC, initial pH and slurry was P. Out of total 6 paths from left branch of the

tree, 2 set of rules resulted in power density ≥ 30 mW/m^2. First path provided set of rules, where SC and T ranged from 41.47 to 87.5 mL/L and 22.44 to 36.25 °C. The other conditional parameters were pH < 5.2, R ≤ 285 and slurry was P. First path generated a power density of 35 ± 2.6 mW/m^2 which was supported by only 2 data points. On the other hand, second path yielded a power density of 41.5 ± 17.9 mW/m^2, where SC and T ranged from 41.47 to 87.5 mL/L, and 22.44 to 36.25°C, respectively. The pH ranged from 5.2 to 7.2, R < 285 and slurry was P. The second path was supported by 13 data points. The most conclusive set of rules originated from the second path was supported by 13 data points and led to a power density of 41.5 ± 17.9 mW/m^2. The final regression value of the model was 0.98. Predictor variables like initial pH highly influenced the voltage generation in MFC. Figure 6. 7 shows relative importance of the predictor variables on the power density.

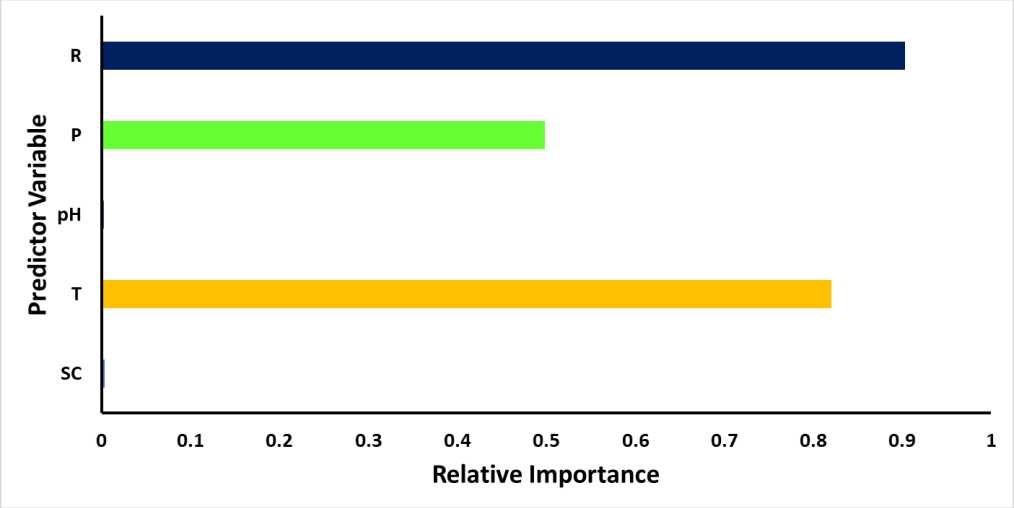

Figure 6. 7. Relative importance of the predictor variables on power density

R was highly influential factor in this study followed by T. P of slurry affects the power density moderately. However, pH and SC had negligible influence on the power density. Approximately, 4.5% reduction in maximum power density at 40 °C was noted as compared to the maximum power density obtained at 34 °C. This behaviour was a deviation from the consistent increase in power output obtained when the T was increased from 20 to 34 °C [459].

This trend was different from the steadily rising power output which was observed when the T was raised from 20 to 34 °C [459]. It has been reported that there was a consistent rise in the power density when T was increased from 4 to 35 °C [465]. The performance of MFCs improved when T was raised from 24 to 35 °C. However, the power density decreased when T was raised above 38 °C [466]. Likewise, the drop in power density was observed when T was raised from 37 to 43 °C, although the same MFC system had a linear trend when the temperature was raised from 10 to 37 °C [467]. When the applied R value was lowered, the voltage dropped. However, there was an increase in current density. Therefore, in order to achieve optimum power density, the appropriate voltage loss required [468].

6.3.2 Model Validation

6.3.2.1 Computational Validation

31 entries from dataset were chosen randomly for the computational validation. It should be noted that the dataset used for validation purpose were not applied in any training or testing step of the decision tree models. Table 6.7, Table 6.8 and Table 6.9 show the actual and predicted values of voltage, current and power density respectively after validation step.

Table 6.7 Actual and predicted voltage obtained after computational validation

Run no.	SC (mL/L)	T(°C)	pH	P/NP	R (Ω)	Actual Voltage value (mV)	Predicted Voltage value (mV)
1	75	45.11	6.5	NP	5600	35	35
2	50	25	5	P	100	101	135.22
3	75	32.5	6.5	P	10000	542	546.85
4	100	40	5	NP	5600	39	35
5	100	40	5	P	1500	28	24.44

6	117.04	32.5	6.5	P	3300	318	375
7	75	32.5	3.9	P	470	201	192.44
8	50	25	5	P	3300	330	304.4
9	75	32.5	6.5	P	4700	399	438.4
10	50	25	5	P	5600	402	411.5
11	75	32.5	6.5	NP	10000	475	488
12	75	19.88	6.5	P	10	4	1
13	100	25	8	P	1000	208	235.55
14	75	32.5	6.5	NP	100	49	84.37
15	75	32.5	3.9	NP	1000	114	179.12
16	75	45.11	6.5	P	1000	21	16.66
17	75	32.5	6.5	NP	10000	455	488
18	75	32.5	6.5	P	15000	463	593.42
19	100	25	5	NP	4700	364	421.8
20	117.04	32.5	6.5	NP	100	38	84.37
21	117.04	32.5	6.5	NP	1500	129	244
22	75	32.5	3.9	NP	2200	223	237
23	75	32.5	6.5	P	2200	399	335.12
24	100	40	5	P	470	15	11.85
25	32.95	32.5	6.5	NP	1500	39	55.75
26	75	45.11	6.5	NP	10000	40	43.25
27	50	40	8	NP	15000	45	43.25
28	117.04	32.5	6.5	NP	8200	511	407.6
29	75	32.5	3.9	NP	470	110	137.66
30	50	40	8	P	1500	19	20

31	100	40	8	P	2200	28	20

Table 6.8. Actual and predicted current density obtained after computational validation

Run no.	SC (mL/L)	T (°C)	pH	P/NP	R (Ω)	Actual Current density (mA/m^2)	Predicted current density (mA/m^2)
1	32.95	32.5	6.5	P	10000	1.78	1.27
2	50	25	5	P	3300	16.67	16.21
3	100	40	5	NP	1500	2.00	2.72
4	50	40	5	P	3300	1.72	0.98
5	50	25	5	NP	2200	19.92	17.39
6	75	32.5	6.5	P	4700	17.62	15.42
7	75	32.5	6.5	P	3300	20.96	19.17
8	75	19.88	6.5	NP	2200	1.97	2.55
9	100	40	8	NP	8200	0.75	0.98
10	75	32.5	9.0	P	15000	6.23	0.82
11	32.95	32.5	6.5	P	3300	3.59	2.30
12	50	25	5	NP	10000	8.35	9.21
13	117.04	32.5	6.5	P	10	966.67	781.25
14	100	40	8	NP	1500	2.00	2.72
15	75	32.5	6.5	NP	5600	12.41	11.78
16	75	32.5	6.5	P	10000	9.62	8.92
17	75	32.5	6.5	P	470	73.40	73.65
18	100	40	8	NP	3300	1.31	0.98

19	50	40	5	P	100	20.00	16.42
20	75	32.5	6.5	P	1500	36.67	31.97
21	100	40	8	P	15000	0.56	0.98
22	75	32.5	6.5	P	15000	5.14	6.13
23	100	25	8	NP	15000	4.60	4.82
24	75	32.5	6.5	P	1000	49.50	39.85
25	100	40	5	NP	3300	1.67	0.98
26	75	32.5	6.5	NP	2200	36.29	17.39
27	50	25	8	NP	4700	12.06	10.31
28	75	32.5	6.5	P	8200	11.52	10.15
29	100	40	5	P	15000	0.61	0.98
30	32.95	32.5	6.5	P	10	100.00	781.25
31	75	32.5	6.5	NP	15000	5.07	6.13

Table 6.9. Actual and predicted power density obtained after computational validation

Run no.	**SC (mL/L)**	**T (°C)**	**pH**	**P/NP**	**R (Ω)**	**Actual Power density (mW/m^2)**	**Predicted Power density (mW/m^2)**
1	75	32.5	6.5	P	10000	5.62	5.06
2	75	32.5	9.02	NP	15000	0.06	0.09
3	100	40	5	P	10	0.15	0.33
4	75	32.5	6.5	P	4700	3.15	7.48
5	75	32.5	6.5	NP	8200	2.65	4.51
6	50	25	8	NP	8200	3.00	3.45

7	100	40	8	P	1500	0.07	0.04
8	75	32.5	3.97	NP	4700	6.65	4.51
9	75	32.5	6.5	NP	100	2.41	12.68
10	75	32.5	3.97	P	15000	3.96	3.52
11	100	25	8	NP	470	3.69	2.84
12	100	25	8	NP	1500	3.89	2.84
13	75	19.88	6.5	P	3300	0.05	0.03
14	50	25	5	P	470	12.67	14.43
15	75	32.5	6.5	P	8200	6.06	5.06
16	75	19.88	6.5	P	10	0.27	18.98
17	75	19.88	6.5	NP	1000	0.07	0.05
18	100	40	8	NP	5600	0.03	0.04
19	100	25	8	NP	100	12.62	20.07
20	75	32.5	6.5	P	10000	4.90	5.06
21	75	19.88	6.5	P	8200	0.03	0.03
22	75	32.5	6.5	P	10	56.07	49.72
23	100	40	5	P	8200	0.05	0.04
24	75	32.5	6.5	P	1500	12.17	9.56
25	75	32.5	6.5	NP	1500	4.76	8.85
26	117.04	32.5	6.5	NP	100	2.41	20.07
27	75	32.5	6.5	NP	1000	3.80	8.85
28	50	40	8	P	8200	0.03	0.04
29	75	32.5	9.02	NP	1500	0.10	2.84
30	100	40	8	P	3300	0.05	0.04
31	32.95	32.5	6.5	P	5600	0.23	0.22

Figure 6. 8 (a-c) represents trend of actual and predicted values for each experimental run for voltage, current density and power density respectively. Figure 6. 8 (a-c) clearly depicted the same path for actual and predicted values of voltage, current and power density. Computational validation depicted < 10% error in most of the cases.

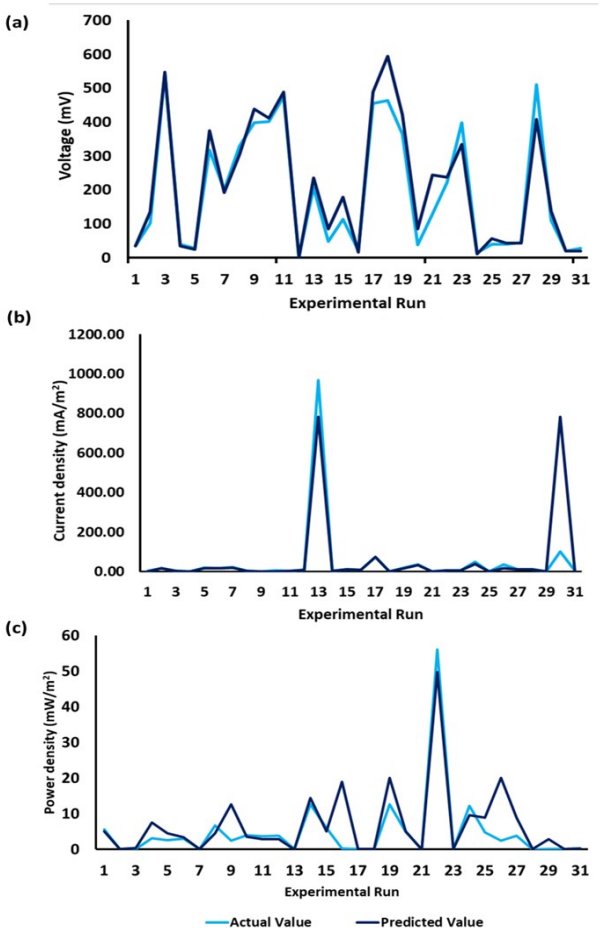

Figure 6. 8. Actual and predicted value for (a) voltage (b) current density and (c) power density

6.3.3.2 Experimental Validation

Final experimental validation was done for the power density by testing the rules extracted from Table 6.6 through experimental verification. Experimental setup was same as that was used for the data generation. All experiments were conducted in triplicate and average values

were recorded. There were total five paths in the Table 6.6 that generated power density ≥ 10 mW/m^2. These all pathways are tabulated in the Table 6.10.

Table 6. 10. Set of rules and their predicted values for model validation

Set of Rules	Predicted Value for Power density (mW/m^2)
41.47 ≤ SC; T < 28.75; pH < 7.25; P; R < 285	20.78
41.47 < SC < 87.5; 22.44 ≤ T < 36.25; pH < 8.5; NP; R < 285	12.6
41.47 < SC < 87.5; 22.44 ≤ T < 36.25; pH < 5.2; P; R < 285	35.6
41.47 < SC < 87.5; 22.44 ≤ T < 36.25; 5.23 ≤ pH < 7.25; P; R < 285	41.5
SC ≥ 62.5; 22.44 ≤ T < 36.25; P; 285 ≤ R < 735	14.2

According to the variable range shown in Table 6. 10, all the values taken for validation of R (< 285Ω) were 100Ω, 120 Ω, 150 Ω, 180 Ω and 200 Ω. Similarly, values taken for 285 ≤ R < 735 Ω, were 470 Ω and 560 Ω. Apart from R values all other input variables were kept same as given in the rules. It should be noted that the data presented in Table 6.10 were not applied in any training and testing step of the decision tree regression models.

Table 6.11. Experimental validation of the decision tree models

S No.	Input Variables	Predicted value	Experimental value	Error (%)

	T (°C)	pH	SC (mL/ L)	P/NP	R (Ω)	PD_P (mW/m^2)	PD_E (mW/m^2)	[(PD_P- PD_E)/ PD_P] ×100
1	27	6	41	P	100	20.8 ± 14.56	21.7 ± 1.05	4.2
2	27	6	41	P	120	20.8 ± 14.56	21.7 ± 1.37	4.4
3	27	6	41	P	150	20.8 ± 14.56	21.2 ± 0.99	1.8
4	27	6	41	P	180	20.8 ± 14.56	22.2 ± 1.08	7.0
5	27	6	41	P	220	20.8 ± 14.56	22.4 ± 0.55	7.8
6	30	7.5	65	NP	100	12.7 ± 12.29	13.5 ± 0.05	6.4
7	30	7.5	65	NP	120	12.7 ±12.29	15.6 ± 0.86	23.0
8	30	7.5	65	NP	150	12.7 ±12.29	15.5 ± 0.92	21.9
9	30	7.5	65	NP	180	12.7 ±12.29	14.7 ± 0.84	15.8
10	30	7.5	65	NP	220	12.7 ±12.29	13.0 ± 1.05	2.5
11	30	4.5	65	P	100	35.6 ±2.65	31.7 ± 0.95	10.9
12	30	4.5	65	P	120	35.6 ±2.65	30.8 ± 2.21	13.4
13	30	4.5	65	P	150	35.6 ±2.65	34.4 ± 0.66	3.4
14	30	4.5	65	P	180	35.6 ±2.65	35.2 ± 1.10	1.2
15	30	4.5	65	P	220	35.6 ±2.65	36.3 ± 1.55	2.0
16	30	6.25	65	P	100	41.5 ± 17.97	42.7 ± 1.70	2.8
17	30	6.25	65	P	120	41.5 ± 17.97	42.1 ± 0.98	1.3
18	30	6.25	65	P	150	41.5 ± 17.97	41.8 ± 0.97	0.8
19	30	6.25	65	P	180	41.5 ± 17.97	41.6 ± 0.89	0.3
20	30	6.25	65	P	220	41.5 ± 17.97	40.4 ± 2.22	2.6
21	30	7.9	65	P	470	14.3 ± 17.97	15.9 ± 0.98	11.5
22	30	7.9	65	P	560	14.3 ± 17.97	15.1 ± 0.97	5.4

The model was able to correctly classify all the 5 sets of rules with an accuracy of 77-99% for the power density (Table 6.11). This signified the reliability of rules obtained from the decision tree models.

6.4 Conclusion

The present chapter has been focused on determining the best combination of input parameters (SC, P/NP, pH, T, R) for achieving high power density in MFC by decision tree regression analysis applying BW slurry as substrate. The best combination was determined by accuracy level, computational and experimental validation. Computational validation delivered < 10% error in most of the cases. The best combination for high power density was: $41.47 < SC < 87.5$; $22.44 \leq T < 36.25$; $5.23 \leq pH < 7.25$; P; $R < 285$. This path showed maximum accuracy level up to 97 - 99%. Results indicated that T, R and P were the most influential parameters to achieve high power density. In order to obtain high current and power densities, R and T values must be lie between 55 to 285Ω and 22.44 to 36.25 ºC, respectively. Slurry pretreatment was also found effective in improving power density. SC between 41.47 to 87.5 mL/L with pH 5.2 was effective in obtaining higher power density. This study can be redesigned with many other parameters such as COD, volume/surface area of electrode, different kinds of electrodes/proton exchange membranes, variation in anode biocatalyst etc., as input variables. More complex algorithms such as random forest can be used for optimization in order to increase the efficiency of MFC setups. Results of study are immediately applicable at pilot and industrial scale without the need for further wet lab trials. This study also eliminates the need of further optimization of operational conditions for similar kind of waste and microbial inoculum.