3 Biotech (2021) 11:28
https://doi.org/10.1007/s13205-020-02553-2

ORIGINAL ARTICLE

Optimization of process parameters for bio-enzymatic and enzymatic saccharification of waste broken rice for ethanol production using response surface methodology and artificial neural network–genetic algorithm

Payel Mondal[1] · Anup Kumar Sadhukhan[1] · Amit Ganguly[2] · Parthapratim Gupta[1]

Received: 18 December 2019 / Accepted: 12 November 2020 / Published online: 3 January 2021

Abstract
Reducible sugar solution has been produced from waste broken rice by a novel saccharification process using a combination of bio-enzyme (bakhar) and commercial enzyme (α-amylase). The reducible sugar solution thus produced is a promising raw material for the production of bioethanol using the fermentation process. Response surface methodology (RSM) and Artificial neural network-genetic algorithm (ANN-GA) have been used separately to optimize the multivariable process parameters for maximum yield of the total reducing sugar (TRS) in saccharification process. The maximum yield (0.704 g/g) of TRS is predicted by the ANN-GA model at a temperature of 93 °C, saccharification time of 250 min, 6.5 pH and 1.25 mL/kg of enzyme dosages, while the RSM predicts the maximum yield of 0.7025 g/g at a little different process conditions. The fresh experimental validation of the said model predictions by ANN-GA and RSM is found to be satisfactory with the relative mean error of 2.4% and 3.8% and coefficients of determination of 0.997 and 0.996.

Keywords Bio-enzyme (Bakhar) · Artificial neural network · Genetic algorithm · Broken rice · Bio-ethanol

Introduction

Global energy demand has been increasing steeply due to rapid industrialization and alarming growth in the world population (Conesa et al. 2016). Most of the energy demand is met by utilizing conventional fossil fuels such as coal (33%), petroleum (24%) and natural gases (19%), while the rest is sourced from the renewables like wind, solar, biomass, geothermal, hydropower, etc. (Stöcker 2008). Combustion of fossil fuels produces greenhouse gases, like CH_4, CO_2, and other harmful gases like NOx, SOx, PAH, thereby causing environmental pollution, global warming, and ecological imbalance. The use of fossil fuel in the transportation sector contributes to about 19% of CO_2 and 70% of CO emission to the environment (Achinas et al. 2016), emitting about 8 kg of CO_2 per gallon of gasoline (Balat et al. 2009). The CO and SO_2 emission showed a rise of 5.1% and 6.5% from 2011 to 2012, while a 4.5% increase was recorded on both NO_2 and particulate matter emissions. The detrimental effect on the environment due to the use of fossil fuels may be avoided using biofuels such as biodiesel in place of commercial diesel and bioethanol in place of gasoline. Biodiesel production and its performance analysis are well-proven, but large-scale production of bioethanol is still under research (Aditiya et al. 2016). Moreover, the reserve of non-renewable energy sources is depleting very fast (Devarapalli et al. 2015), while the energy demand is expected to increase by 50% in 2030 (U.S. EIA 2019). The United Nations (UN) has targeted a 50–80% reduction in emissions of greenhouse gas by 2050, and the biofuels

✉ Anup Kumar Sadhukhan
t_sadhu@yahoo.com

Payel Mondal
payelmondal2128@gmail.com

Amit Ganguly
amitganguly022@gmail.com

Parthapratim Gupta
parthagupta2000@yahoo.com

[1] Chemical Engineering Department, National Institute of Technology, Durgapur 713209, India

[2] CSIR-Central Mechanical Engineering Research Institute, Durgapur 713209, India

directive of the European Union (EU) Commission has suggested using 20% of biofuels by 2020 (Stöcker 2008).

Bio-ethanol is one of the reliable renewable energy sources and is considered to be a promising energy generation route as it has an oxygen content itself, resulting in less emission of carbon monoxide and greenhouse gases when used in 5–10% as an additive in fuel for automobiles (Betiku et al. 2015; Ahmad et al. 2011). Ethanol can be a very good replacement for methyl tertiary butyl ether (MTBE), which is an additive of gasoline fuel as MTBE is hazardous for human health (Betiku et al. 2015). It is used as gasohol (blended with gasoline) or pure fuel in motor vehicles (Wang et al. 2011). Bioethanol has various advantages over gasoline due to its higher octane number of 108, larger heat of vaporization and reduced emission of harmful gases (Azhar et al. 2017). Waste biomass may be utilized to produce ethanol as it is abundantly available, carbon–neutral and free from 'fuel vs. food' competition (Rosillo-Calle et al. 1987). Moreover, the residual biomass from the biofuel production can be converted to value-added products in a well-integrated biorefinery (Das et al. 2015). An abundant quantity of low-cost starchy raw materials like waste/broken rice, waste potatoes, sugarcane bagasse, corncob, agricultural waste, vegetable waste, etc. is gaining its popularity for the production of bioethanol.

One of the promising process routes for bio-ethanol production is converting perennial crops like breadfruit (*Artocarpus altilis*), which is abundantly grown in Southern Nigeria (Betiku et al. 2015; Bankole et al. 2005). Sugarcane is another ethanol source, rich in sucrose, in tropical and subtropical countries like Brazil, India, and Colombia. However, pretreatment with alkali or acid for the removal of lignin from sugarcane juice prior to saccharification burdens the process route with additional cost.

Broken rice and other waste grains contain starch, which is a polysaccharide of glucose units linked by α (1–4) and α (1–6) glycosidic bonds (Pandey et al. 2010). Nearly 110 million tons of rice was produced in India in 2017–2018, out of which about 7.5% is broken during polishing to get the fine rice. The huge quantity of waste broken rice is very cheap and is mainly used as cattle feed. This may be utilized to recover reducible sugar at 75–100 °C temperature as it contains 80% starch, which can subsequently act as a substrate for ethanol production (Chu-Ky et al. 2016; Gronchi et al. 2019).

Li et al. 2013, extracted the reducible sugar from broken rice after treatment with enzyme only α-amylase with the final concentration of total reducing sugar (TRS) of 82.7 g/L. Schneider et al. (2018) reported that about 70% starch in the raw broken rice was extracted as reducible sugar on hydrolysis after treatment with two enzymes alpha-amylase and gluco-amylase but the final concentration of reducible sugar solution was not reported. Myburgh et al. (2019) utilized the industrial amylolytic *Saccharomyces cerevisiae* derivatives of Ethanol Red™ Version 1 (ER T12) during saccharification. Although no external enzyme was utilized, the enzyme novel amylase combination secreted by the strains was mainly responsible in hydrolysing the starch in broken rice, with the maximum degree of saccharification of 30% and the concentration of total reducing sugar of 10.5 g/L.

The starch content of broken rice cannot be fermented directly by yeast as it gives a low yield of ethanol. Hence, the starch in broken rice is first converted to glucose through saccharification by treating it with locally available low-cost bio-enzyme called Bakhar and subsequently by the commercially available enzyme α-amylase. Bakhar is the combination of fresh, clean, and dry plant parts (roots, leaves, bark, etc.) of six different species; Akanbindi (*Cissampelo spareira*), Chaulia (*Ruellia turberosa*), Fern (*Lygodium flexuosum*), Asan (*Terminalia alatabark*), Kendu (*Diospyros melanoxylon*) and Red java tea (*Orthosiphon rubicunds*). The starch in broken rice primarily contains polysaccharides, which undergoes de-polymerization under the action of commercial bio-enzyme (bakhar) and enzyme (α-amylase) through hydrolysis and easily gets converted to reducing sugar. Finally, the resulting glucose is treated with an inoculum of yeast (*Saccharomyces cerevisiae*) and allowed to be fermented under controlled conditions of temperature and pH to produce ethanol. The low-cost commercial-grade enzyme and locally made bio-enzyme make the entire process efficient and cost-effective (Hickert et al. 2012). Yeasts are widely used in large-scale industrial production plants due to higher ethanol yield (> 90.0%), and the undiluted fermentation broth has the resistance to inhibitors, thereby reducing the chances of contamination during the growth period (Hickert et al. 2012).

The recovery of reducible sugar from broken rice is a very complex process with several biochemical reaction pathways. The final yield depends on the nature of substrate, process conditions, enzyme loading etc.; such a complex nonlinear process includes reaction kinetics and mass transfer operation. The exact modeling of such individual processes is intricate due to the lack of reliable quantitative estimations of various intermediates. Optimizing one single variable at a time is not efficient and does not identify the exact optimum conditions. Therefore, the process optimization was carried out by Response surface methodology (RSM) and Artificial neural network (ANN)-Genetic algorithm (GA) for studying the relations between different process parameters. RSM tool is employed for evaluating the association between response and independent process parameters through mathematical and statistical equations, while ANN uses non-linear multivariate modeling techniques to capture the interconnected non-linear process response of inputs on the process output at the given experimental conditions (Teslic et al. 2019). RSM also represents a multivariable process using

non-linear quadratic polynomial with all possible combinations of inputs. Hence, the dependency of the yield of reducing sugar on the significant process parameters for the saccharification process is expected to be modeled with reasonable accuracy.In addition ANN has the property of excellent fitting quality, while RSM provides further insight for influence analysis of process parameters (Teslic et al. 2019).

In the present study, broken rice was chosen as the main substrate to produce reducible sugar by the saccharification process and subsequently fermentation using *Saccharomyces cerevisiae*. The bio-enzyme (bakhar) and commercial enzyme (α-amylase) are used to enhance the yield of reducible sugar. The pre-treatment of waste broken rice, saccharification, fermentation and suitable separation of ethanol are the four significant steps for the production of bioethanol from waste broken rice. The crucial parameters that affect the effectiveness of the enzymatic and bio-enzymatic saccharification process of broken rice are identified as (a) temperature, (b) time, (c) pH and (d) enzyme dosages. The objectives of the present study are (i) to optimize the yield of reducing sugar by the combination of bio-enzymatic and enzymatic saccharification using RSM (ii) to optimize the same with the help of an Artificial neural network coupled with Genetic algorithm, and (iii) to assess the optimized process parameters with experimental results. The enhanced yield of reducing sugar will make the process the most cost-effective by improving the final yield of bioethanol.

Materials and methods

Broken rice as feedstock

Broken rice (*Oryza sativa*) was procured from Burdwan (West Bengal, India) rice mill. Broken rice composition was examined by the analytical protocol of the National Renewable Energy Laboratory (NREL) (Sluiter et al. 2012). It was found to contain (dry wt. basis) 79.89% of starch, 10.38% protein, 2.76% fat (lipid), 2.51% crude fibers and 4.46% others (non-starch-polysaccharide, minerals, ash content, etc.) by the experimental technique proposed by Sluiter et al. (2012).

The enzymes and microorganism

The suitable operating condition of the commercial-grade α-amylase enzyme was found to be 85–90 °C temperature and 6–6.5 pH. Bakhar was added to enhance the hydrolysis of starch to glucose with a higher yield for the saccharification process. Eventually, the *Saccharomyces cerevisiae* was added; it acts as a fermenting microorganism to convert the reducible sugar into alcohol. Bakhar was collected from a local market in the district of Bankura (West Bengal, India). The dry yeast (*Saccharomyces cerevisiae*) was procured from Kothari fermentation and Biochem Ltd. No. 16, Community Centre, Saket, New Delhi-110017, India.

Pre-treatment of broken rice

One kg of broken rice was washed properly to remove the dirt and other impurities and boiled in a water bath at 100 °C for 2.5 h using the 3 L of water. The boiling process mainly breaks down the long polysaccharide chains of starch into short-chains of monosaccharides.

Saccharification

The boiled mass was kept at a constant temperature water bath (80 °C to 100 °C), and the pH of the solution was varied from 5 to 8. The bio-enzyme bakhar (10 g/kg of the substrate) was added to it to enhance the hydrolysis of starch to glucose. Subsequently, the commercial enzyme α-amylase (1.25 ml/kg of substrate) was added to the mixture, and the resulting solution was allowed to hydrolyze for 8.5 h. During this period, the combined effect of bakhar and α-amylase broke down the starch leading to a higher yield of reducible sugar.

Fermentation

The supernatant from the resulting saccharified solution was stained out and maintained at 30 °C and pH was set at 6.5. The commercial yeast (*Saccharomyces cerevisiae*) was added into the supernatant for converting the glucose to ethanol by the process of fermentation. The media was supplemented with the nutrients- 0.5 g/L NH_4Cl, 2.1 g/L KH_2PO_4, 0.45 g/L $MgSO_4{\cdot}7H_2O$, 0.15 g/L $CaCl_2{\cdot}2H_2O$, 2 mg/L $ZnSO_4{\cdot}7H_2O$, 2.5 mg/L $FeCl_3{\cdot}2H_2O$ and yeast extract 2.0 g/L. The mixture was blended well with the supernatant for a few minutes until the mixture becomes homogeneous (Ahmad et al. 2011). The resulting mixture pH was set to 6.5, put into a glass beaker and was sealed properly to ensure anaerobic conditions suitable for yeast fermentation. Finally, the beaker was kept in an incubator maintained at 30 °C for 2–4 days (Ramaraj et al. 2019). Eight identical samples for fermentation, as discussed above, were made and analyzed for alcohol concentration after 6–12 h interval. The maximum yield of alcohol was detected for the sample kept for 60 h. The concentration of alcohol in the alcohol-water mixture was found to be 70 mg/mL. The common technique for concentrating it includes azeotropic distillation, molecular sieve technology, vacuum distillation, etc. (Taheri et al. 2017). In the present work, the two stages vacuum distillation process was employed at 10

millibar pressure to achieve the final ethanol concentration of 70% (v/v).

Enzyme activity

The commercial-grade α-amylase enzyme (derived from *Bacillus licheniformis*) was purchased from Om Biosciences, batch no. 962, Ahmedabad, India. Its activity was 30U/L (as per specification mentioned by the supplier).

Activity test of reducing sugar

3, 5 di-nitrosalicylic acid (DNSA) was used as the main reagent for measuring reducing sugar activity using a spectrophotometer (Spectroquant® Prove 300, Merck KGaA, Frankfurter Straße 250, 64293 Darmstadt, Germany) with 540 nm wavelength using di-nitrosalicylic acid solution. For the preparation of 100 mL of DNSA solution, the relative proportion of various reagents were as-100 mL distilled water, 0.77 mL 3, 5 di-nitro salicylic acid, 1.3 g NaOH, 21.61 g potassium sodium tartrate, 0.54 mL phenol, 0.59 g sodium metasulphate (Alam et al. 2019). The 0.2 mL of the sugar solution was added to 0.3 mL of DNSA reagent in a test tube. The test tube was placed in a boiling water bath for 5 min. The standard solution was prepared by mixing 0.2 mL of the distilled water with 0.3 mL of DNSA solution to make the zero reference in the spectrophotometer. The reagents used for the preparation of DNSA solutions were procured from MERCK Specialities Private Limited, Mumbai-400018, India.

Estimation of ethanol concentration

Agilent (Agilent Technologies 1200 series), ZORBEX Phenyl high-performance liquid chromatography (HPLC) was used for measuring the ethanol concentration at the end of saccharification and at various time instances during the process of fermentation. The Refractive index detector (RID) was used for the measurement of ethanol concentration. Operating conditions were: mobile phase 100% pure milli-Q water, flow rate 1 mL/min, injection volume 5 µL, temperature 250 °C, retention time 2.7 min (Kumar et al. 2017).

Broken rice characterization

Scanning electron microscope (SEM)

The surface morphological study of untreated broken rice and that treated with bio-enzyme and commercial enzymes were carried out using the scanning electron microscope (Sigma, Carl Zeiss, UK) images. The dried sample was placed on the aluminum stub and coated with platinum metal for 90 s using an ion-sputter coating system. The sample images were obtained by scanning at 0.02–10 kV acceleration voltage with a beam current of 89–100 pA. The resolution of the present model is 1 nm at the acceleration voltage of 10 kV.

X-ray diffraction analysis (XRD)

XRD analysis of both the untreated broken rice (dry powder) and that treated by bio-enzyme and commercial enzyme were analyzed by a Rigaku MiniFlex II diffractometer (Rigaku, Tokyo, Japan). Cu K-alpha radiation was used with 35 kV beam voltage, 15 mA beam current with the scanning angle (2θ) in the range of 5-40° and the scanning speed of 1°/min. The sample crystallinity index (CI) was calculated after (Das et al. 2014):

$$\text{Crystallinity index, CI}\,(\%) = \left[\mathrm{I}_{002} - \mathrm{I}_{14.7^\circ}\right]/\mathrm{I}_{002} \quad (1)$$

where I_{002} denotes the highest intensity at *002* lattice diffraction and $I_{14.7}$° is the background scatter at 2θ = 14.7◦.

The broken rice retains its high crystallinity until it is treated with enzymes/reagents. On treatment with bio-enzyme (bakhar) and commercial enzyme (α-amylase), it experiences a series of biochemical pathways leading to lowering of its crystallinity with the progress of the saccharification process. Large void space is formed, which enhances the internal surface area providing a better environment for easy hydrolysis of starch towards the higher yield of total reducing sugar (Yoon et al. 2012).

Fourier-transform infrared spectroscopy (FTIR)

Fourier-transform infrared spectroscopy is a unique technology to identify the various functional groups in any sample. In the present study, the IR spectroscopy (PerkinElmer FT-IR C109292, UK) was used to carry out the experiments. Both the untreated broken rice and that treated by bio-enzyme and commercial enzyme were analyzed. The dry powder sample was mixed with dry KBr in the proportion of 1:10 ratio for pellet preparation and the pellets were placed for the FTIR scan. The FTIR scan was also performed at the beginning with the pure KBr pellet for background correction. The resolution of 4 cm^{-1} and the wavenumber range was 4000 to 400 cm^{-1} for the present study. The range generally covers the fundamental vibrations by almost every organic compound (Ghaffar et al. 2013) in rice.

Experimental designs and optimization strategy

Response surface methodology

The response surface methodology is used in this study for optimizing the process parameters of the saccharification process, as discussed above. The major process parameters are temperature, time, pH and enzyme dosages. In the present study, five levels (one upper limit, one lower limit, and three intermediates levels) have been considered for each parameter. It is difficult to carry out these 5^4 numbers of experiments to assess the exact effect of the parameters on the yield of total reducing sugar. RSM is a tool that helps design the experiments considering both mathematical relationships among the inputs and the outputs and the interaction of various inputs without sacrificing the effect of local minima or maxima (Rao et al. 2000). In the present study, 30 experiments proposed by the central composite design (CCD) were obtained from RSM to optimize the input parameters, like temperature, time, pH and enzyme dosages for maximum yield of TRS. The input temperature has a lower range of 80 °C and the upper range is 100 °C; for the time, it is 0 and 500 min, respectively. The pH lies in the range 5 to 8, while 0.5 (mL/kg) is the minimum range and 2 (mL/kg) is the maximum range for enzyme dosages. The CCD for experimental design and the experimentally obtained yield of reducing sugar and that predicted by RSM, are presented in Table 1.

A second-order polynomial equation was used to describe the effects of input parameters regarding linear and quadratic interactions:

$$Y = \alpha_0 + \sum_{i=1}^{k} \alpha_i X_i + \sum_{1}^{k} \alpha_{ii} X_i X_i + \sum_{i=1}^{k-1} \sum_{j=i+1}^{k} \alpha_{ij} X_i X_j, \quad (2)$$

Table 1 Comparison of experimental total reducing sugar with the prediction by RSM and ANN

Run no	Temperature (°C)	Time (min)	pH	Enzyme dosage (ml/kg)	Experimental yield (g/g)	RSM predicted yield (g/g)	ANN predicted yield (g/g)
1	85	125	5.75	1.625	0.217	0.197	0.217
2	90	250	6.50	1.250	0.689	0.682	0.674
3	85	125	7.25	1.625	0.259	0.262	0.259
4	95	375	5.75	0.875	0.283	0.287	0.283
5	90	250	8.00	1.250	0.250	0.248	0.250
6	100	250	6.50	1.250	0.650	0.648	0.649
7	90	250	6.50	1.250	0.677	0.682	0.674
8	95	375	7.25	0.875	0.412	0.427	0.411
9	90	250	6.50	1.250	0.675	0.682	0.674
10	85	125	7.25	0.875	0.390	0.383	0.390
11	90	250	6.50	1.250	0.689	0.682	0.674
12	95	375	7.25	1.625	0.222	0.227	0.222
13	90	250	6.50	1.250	0.657	0.682	0.674
14	95	125	5.75	1.625	0.298	0.297	0.298
15	95	375	5.75	1.625	0.278	0.271	0.278
16	85	375	7.25	1.625	0.268	0.258	0.268
17	85	375	7.25	0.875	0.380	0.386	0.380
18	90	250	6.50	2.000	0.141	0.152	0.127
19	95	125	5.75	0.875	0.310	0.306	0.310
20	95	125	7.25	0.875	0.453	0.452	0.453
21	85	375	5.75	1.625	0.182	0.189	0.189
22	95	125	7.25	1.625	0.273	0.258	0.273
23	90	250	6.50	1.250	0.687	0.682	0.674
24	80	250	6.50	1.250	0.488	0.498	0.487
25	85	125	5.75	0.875	0.123	0.124	0.123
26	85	375	5.75	0.875	0.131	0.132	0.131
27	90	250	6.50	0.500	0.312	0.300	0.312
28	90	000	6.50	1.250	0.156	0.179	0.156
29	90	500	6.50	1.250	0.170	0.156	0.170
30	90	250	5.00	1.250	0.022	0.027	0.032

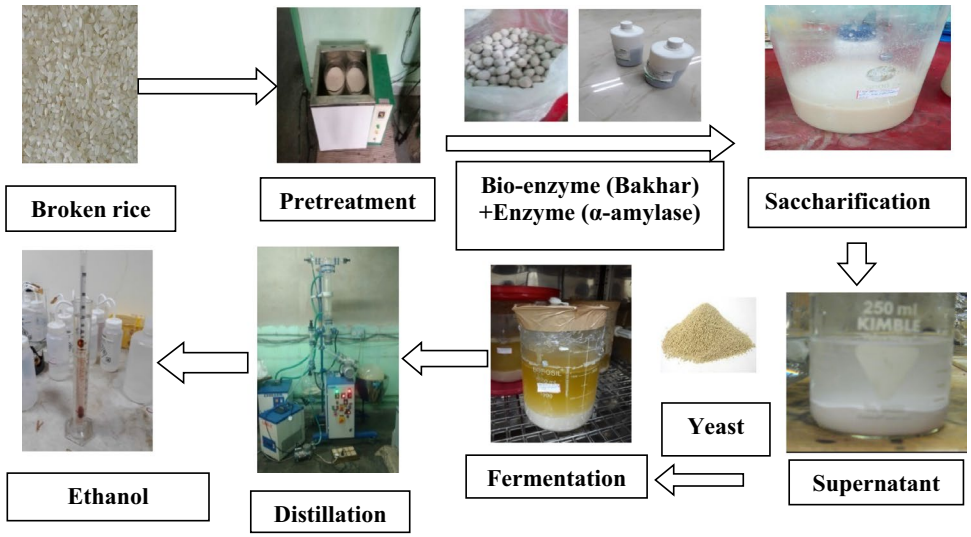


Fig. 1 Schematic diagram for the process of ethanol production from waste broken rice

where *'Y'* is the estimated yield, X_i be the input parameters (i is a counter for no of input) α_0 is a constant, α_i is the linear coefficient associated with X_i, α_{ii} expresses the coefficient of the squared terms, α_{ij} implies the cross coefficient of the cross-product of input terms for the development of generalized quadratic equation. The coefficients of the above second-order polynomial was estimated using the software Design Expert, Version 10. The corresponding experimental design setup is shown in Fig. 1.

Artificial neural network

An artificial neural network is a mathematical tool extensively used to predict the linear as well as the non-linear relationship between multiple inputs and output for a complicated process. ANN and RSM both may be used separately for optimizing the process parameters of the non-linear and highly interconnected input parameter-dependent bio-processes. ANN is considered to be more accurate compared to RSM due to the presence of its highly interconnected bundles of elements called neurons. The linkage between the neurons is represented by its unique parameter bias (b) and weight (w). Three different functions like tan-sig, log-sig and purelin (Hagan et al. 2014) are mostly used for a complex non-linear relationship. A multi-layer neural structure is comprised of the input layer, the hidden layers, and the output layer. In the present case, the predictive ANN model has been developed using temperature (°C), time (min), pH and enzyme dosages (mL/kg) as input keys and the total reducible sugar (g/g) yield as the output key. The signal generated from the hidden layer operation acts as the input to the output layer. The predicted response from the output layer and the actual yield obtained experimentally for the given set of the input data are employed to estimate mean-squared error considering all the experimental data sets. Back Propagation (BP) training algorithm is used to minimize the error function by adjusting the weights and biases (Nasab et al. 2019). A properly trained network can act like an expert (Ali et al. 2014) to predict the output for the new set of process inputs. The mean square error (MSE) is lowered to the extent of 0.00047 so that the average correlation coefficient (CC) is closer to 1. The MSE and coefficient of determination, R^2 values of the ANN model, are calculated as follows (Nasab et al. 2019):

$$MSE = \left(\frac{1}{n}\sum_{i=1}^{n}(y_{ie} - y_{ip})^2\right), \tag{3}$$

where n is the no. of runs and y_{ie} is the experimental data and y_{ip} is the predicted data getting from the model:

$$R^2 = 1 - \sum_{i=1}^{n}\left(\frac{(y_{ie} - y_{ip})^2}{(y_{ie} - y_m)^2}\right), \tag{4}$$

where y_m represents the average of the experimentally obtained reducing sugar yield. The value of R^2 close to 1 indicates that the degree of fitness of the model is good and the model prediction matches well with the experimental data.

Genetic algorithm

The genetic algorithm is used after constructing the ANN model to optimize all process parameters during saccharification. GA uses imaginary exploration to optimize the input parameters. It follows Darwin's theory of 'survival of the fittest' (Pasandideh et al. 2006). The dependent variable reducing sugar, was considered as a chromosome. Every chromosome consists of 4 independent input parameters genes, i.e., temperature, time, pH and enzyme dosages. In this GA tool

(MATLAB V 2014), the inputs are fed and the resulting reducing sugar is generated for different combinations. The chromosome population is derived, and the ANN model is used to evaluate its performance. The high performing chromosomes are replicated to segregate after the crossover of chromosomes and mutation of genes. The performance of each chromosome is evaluated using the ANN model. The process is continued for several generations to identify the optimum combination of all inputs for maximum yield of reducing sugar.

Result and discussion

Scanning electron microscopy analysis of untreated and treated broken rice

The SEM image (×250) of untreated broken rice shows (Fig. 2a) the presence of starch granules and fibers. The cellulose in the polymerized starch appears in a very systematic manner with a pyramidal structure with dimensions ranging from 22 ± 2 to 28 ± 2 µm for the lower and higher dimension, respectively. On further magnification (×2000) the individual granules and the surface layer become prominent (Fig. 2b) and it shows that the cellulose in starch is cross-linked with protein and lipid to form a continuous gel-like structure with a thick outer layer made up of proteins and lipids (Nawaz et al. 2016). These granular structures are disrupted after treating it with bio-enzyme and enzyme and also shown in the SEM images (Fig. 2c). This is due to the de-polymerization of cellulose by the enzyme treatment, where the surface layer of protein and lipids are removed by the action of the enzyme. Consequently, smaller granules with large surface areas are available which will be beneficial for the extraction of higher yield of reducible sugar during the saccharification process. The relatively more irregular shapes with higher porosity are evident from the SEM images on the higher magnification of the treated broken rice (Fig. 2d).

XRD analysis of untreated and treated broken rice

XRD analysis was carried out for both untreated and treated (bio-enzyme and enzyme) samples to study the crystallinity of the sample and the crystallinity index was determined. Figure 3 represents the XRD plots. Due to the presence of Van der Waals force and intermolecular hydrogen bonding between the cellulose molecules within the starch, the presence of a strong crystalline structure is detected for untreated samples (Chirayil et al. 2014). The presence of a prominent peak in the XRD diagram of untreated broken rice supports the crystalline structure of starch in the sample. However, when the sample is treated with bio-enzyme and enzyme, the crystalline structure is demolished, leading to the formation of the amorphous structure of cellulose, where more giant starch molecules disintegrated into smaller globules. This is evident from the XRD graph of the treated broken rice, where

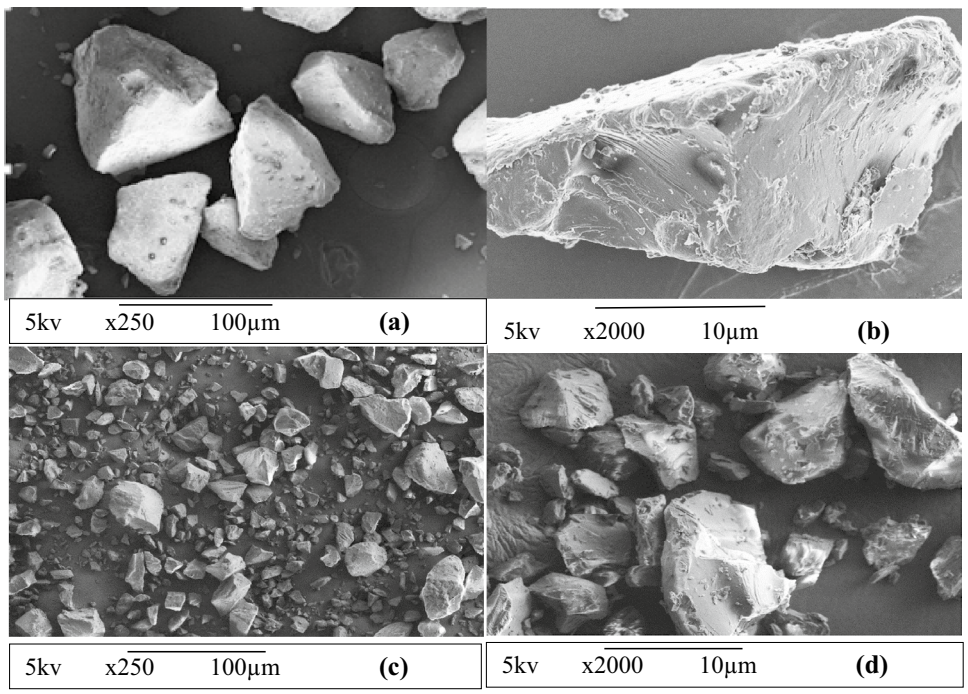


Fig. 2 Scanning electron microscopy image of **a** Untreated broken rice (UBR) ×250 image **b** UBR ×2000 image **c** Treated broken rice (TBR) ×250 image **d** TBR ×2000 image

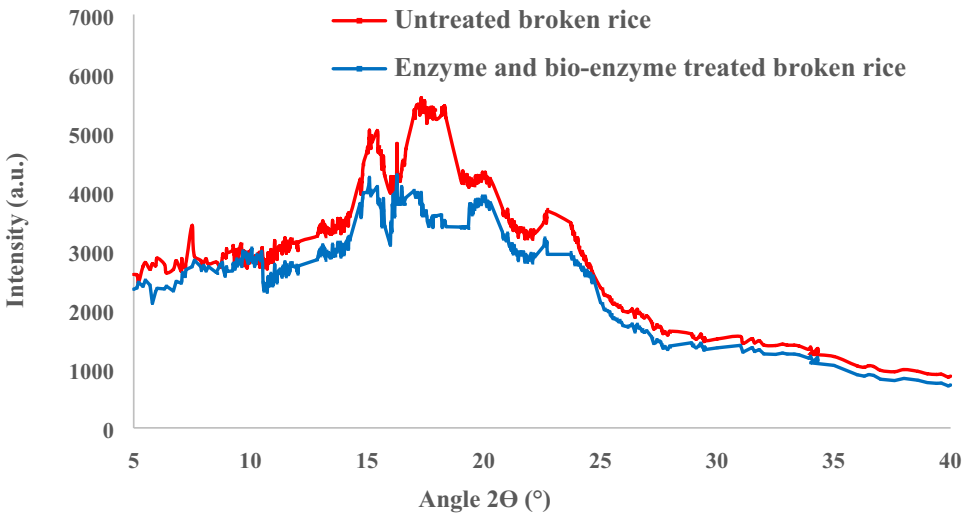


Fig. 3 XRD plots of untreated and treated broken rice

the intensity count hardly shows any peak; rather the entire profile becomes flat, indicating the wider amorphous region. The observation is in line with that observed by SEM images (Fig. 2c) of the treated sample, where small globules of starch were clearly visible. This transformation may also be attributed due to the partial solubilization and expansion of cellulose crystal during de-polymerization under the action of enzymes at modest temperatures (Zhang et al. 2014).

The CI is a quantitative parameter to evaluate the degree of the crystallinity and amorphousness of cellulose in starch. The CI value evaluated (Eq. 1) for untreated broken rice is found to be 33.62%, while that of treated rice is 20.07%. The crystallinity peaks are prominent at 13° and 18° in 2θ for the cellulose of untreated starch, but such prominent peaks are not pronounced for treated samples. Hence, it is clear that the degree of crystallinity has reduced by 13.55% on enzyme treatment due to the disruption of the regular shape of cellulose. A similar observation on the reduction of crystallinity was reported by Yoon et al. (2012) when the ionic liquid was used to treat the sugarcane bagasse.

FTIR analysis of untreated and treated broken rice

Untreated and treated broken rice were characterized by FTIR for observation of the presence of functional groups in Fig. 4. The intensity peak is detected at 3650-3200 cm^{-1} for untreated broken rice due to the presence of (-OH) group and the broader band is present in the untreated sample but it decreases when it is treated with bio-enzyme and enzyme. The absorption peak at 2885, 2930 to 2998 cm^{-1} corresponds to the presence of symmetric and asymmetric C–H bond of CH_3, CH_2, and CH, respectively. The absorption peak detected at 2934 indicates the presence of alkaen (C–H) groups in untreated broken rice (Das et al. 2015). The values are enhanced for treated broken rice due to the treatment of bio-enzyme and enzyme.

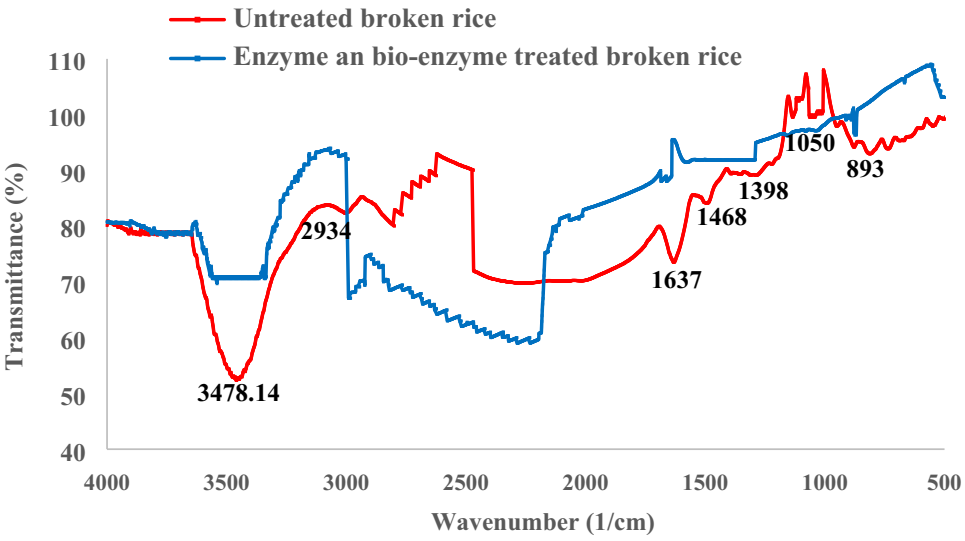


Fig. 4 FTIR plots of untreated and treated broken rice

In the sample untreated broken rice, an absorption peak at 1637 cm^{-1} is detected and it corresponds to the adsorbed water in the broken rice. The asymmetric bending of $–CH_3$ is detected at 1468 cm^{-1}. Another peak at 1398 cm^{-1} is the characteristic stretching of C=C linkages of starch due to the dehydration of cellulose. But the peak intensity almost disappears after treatment with bio-enzyme and enzyme, leading to the hydrolysis of the sample and formation of amorphous cellulose upon de-polymerization. The presence of C–O bonds in carbohydrate show the absorption peak at 1050 cm^{-1} the similar observation is also reported by Saha et al. (2018) for sugarcane bagasse (Saha et al. 2018). Another important information is also observed at 893-897 cm^{-1} for the enzyme-treated sample and indicates the presence of CO–C stretching of the β-(1, 4)-glycosidic linkage in cellulose; this is mainly the characteristic shown by glucose present in cellulose (Chirayil et al. 2014). It is significantly amorphous in nature and its presence increased after the treatment with bio-enzyme and enzyme.

Compositional study and yield of reducing sugar

In broken rice, cellulose content was estimated to be 79.89% (w/w). The total reducing sugar (fermentable sugar) theoretically obtainable from the cellulose is

$$\left(C_6H_{10}O_5\right)n + nH_2O \rightarrow nC_6H_{12}O_6; \left[79.89\% \times \frac{180}{162} = 88.77\%\right]$$

Hence, 0.89 g of reducing sugar may be obtained theoretically from 1 g of broken rice during the saccharification process

ANN-GA modeling and optimization

The performance of ANN solely depends on the experimental data for various process conditions of saccharification and corresponding yields of reducing sugar. The experimental data set presented in Table 1 has been utilized for training, while the newly performed experimental data set presented in Tables 2, 3 are used for testing and validation of the neural network. The neural network is constructed and trained to utilize 30 datasets (Table 1). MathWorksV14.0 is employed to develop the artificial neural network model.

The feed-forward neural network consists of three layers– the input layer, the hidden layers and the output layer. The weights and biases of the hidden layer are denoted by input weight matrix IW and input bias $b^{(1)}$. The weight matrix and bias of the output layer are denoted by OW and $b^{(2)}$, respectively. Levenberg–Marquardt backpropagation algorithm is used to train the neural network, and the resulting weights and biases have been reported. The stopping criteria for training are set to a very low MSE of 0.0005, and after 6^{th} iterations, the ANN model reached the error level and stopped. Figure 5 shows the variation of mean square error vs. iteration. It shows a very low MSE of 0.00047 after 6th iterations. The model predicted yield data for validation and test set are in good agreement with that from

Table 2 Experimental data of total reducing sugar for test data set of ANN

Run no.	Temperature (°C)	Time (min)	pH	Enzyme dosages (ml/kg)	Experimental yield (g/g)	Predicted yield (g/g)
1	85	100	6.00	1.00	0.548	0.568
2	80	150	6.50	1.75	0.395	0.374
3	90	200	6.75	1.50	0.456	0.475
4	95	125	7.00	1.00	0.498	0.481
5	80	200	7.25	1.5	0.422	0.403
6	95	175	6.25	1.75	0.379	0.351

Table 3 Experimental data of total reducing sugar for validation data set of ANN

Run no.	Temperature (°C)	Time (min)	pH	Enzyme dosages (ml/kg)	Experimental yield (g/g)	Predicted yield (g/g)
1	90	150	5.25	1.25	0.379	0.352
2	85	175	6.90	1.75	0.499	0.478
3	95	210	6.50	1.00	0.512	0.532
4	100	145	7.00	1.85	0.479	0.491
5	92	290	7.25	1.25	0.524	0.537
6	95	300	6.10	1.50	0.421	0.402

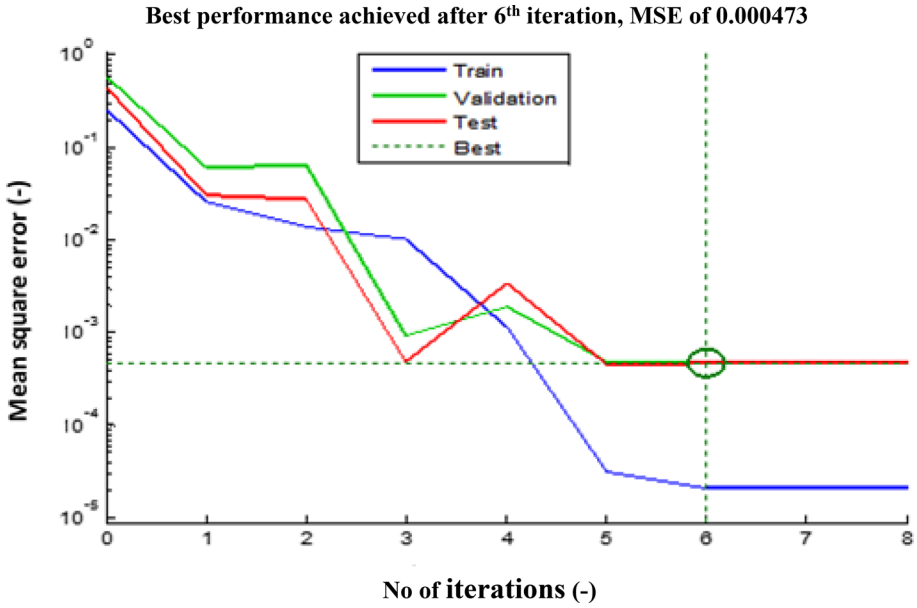


Fig. 5 Variation of mean squared error, MSE during the training phase of the ANN

Table 4 The ANOVA analysis for ANN model

Source	Sum of squares	Deg.of freedom	Mean square	*F* value	*p* value	R^2	Adj. R^2
Model	1.18	28	0.0035	271.14	<0.0001	0.999	0.995
Residual	0.00028	6	0.00012				
Total	1.18028	34					

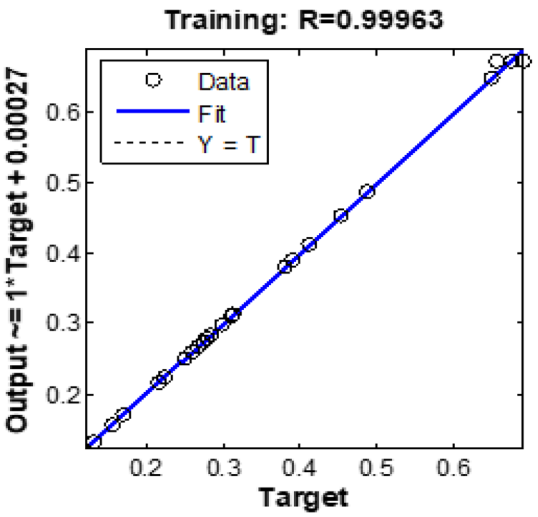


Fig. 6 Comparison of experimental and ANN predicted total reducible sugar yield for the training data set

Table 5 Weights and biases values for the trained ANN model

IW =	0.0878 −0.3592 −1.7345 0.5913 1.4517 1.2742 0.9847 −0.7007 0.3065 −1.9558 1.7608 0.9083 −1.0596 0.2301 0.8004 −1.4836 −0.4278 2.0697 −0.3536 1.0518 −1.2613 1.7161 1.2944 −0.4605 −0.4948 2.2196 2.0984 −0.6884 0.1095 1.7171 −0.9093 2.5512 0.4688 − 0.8484 2.1022 3.3096 0.9921 1.6734 2.0044 −0.3831	$b^{(1)}$ =	−3.5220 −2.7168 −1.1642 1.0978 −0.6914 −0.3112 1.5095 −1.5605 1.3733 2.0446
OW =	1.4988 −0.7060 −0.7995 0.2093 −0.2809 −0.3347 0.2356 −0.4894 0.6577 0.6475		
$b^{(2)}$ =	−1.2563		

experiments (Table 2, 3). The mean square error for validation and test set shows almost similar behavior with respect to an error, which indicates training is very successful and the final network can predict the experimental findings with a higher level of confidence.

The analysis of variance (ANOVA) is performed for the statistical analysis of the model developed and is presented in Table 4. The estimated *F* value of 271.14 and the *p* value of <0.0001, implying that the model is significantly acceptable for practical uses. Figure 6 compares the proximity between the ANN predicted data and that experimentally obtained reducing sugar yield for all the datasets used for training the ANN. The accuracy of the ANN model is further tested against the calculated other two statistical parameters, coefficient of determination ($R^2 = 0.999$)

and adjusted co-efficient (adj. $R^2 = 0.995$) which are also very near to unity, indicating a high confidence level of the ANN. The optimized weights and biases values are presented in Table 5. The root mean square error (RMSE) is also estimated by taking the square root of MSE. The same for training, validation, and test data set are 0.010, 0.013, 0.011, respectively, and the corresponding R^2-values are 0.997 for training, 0.993 for validation and 0.995 for the test data set.

The Error histogram plot is prepared for the estimation of the accuracy level of the ANN model. It is found that the errors for the training set lie in the range of −0.002 to 0.002. The maximum error for the validation set is found to be 0.03. It may be noted that the validation set is totally different than that of the training set. To eliminate the effect of the possible experimental error on neural network quality, repeated experiments were carried out for same process inputs.

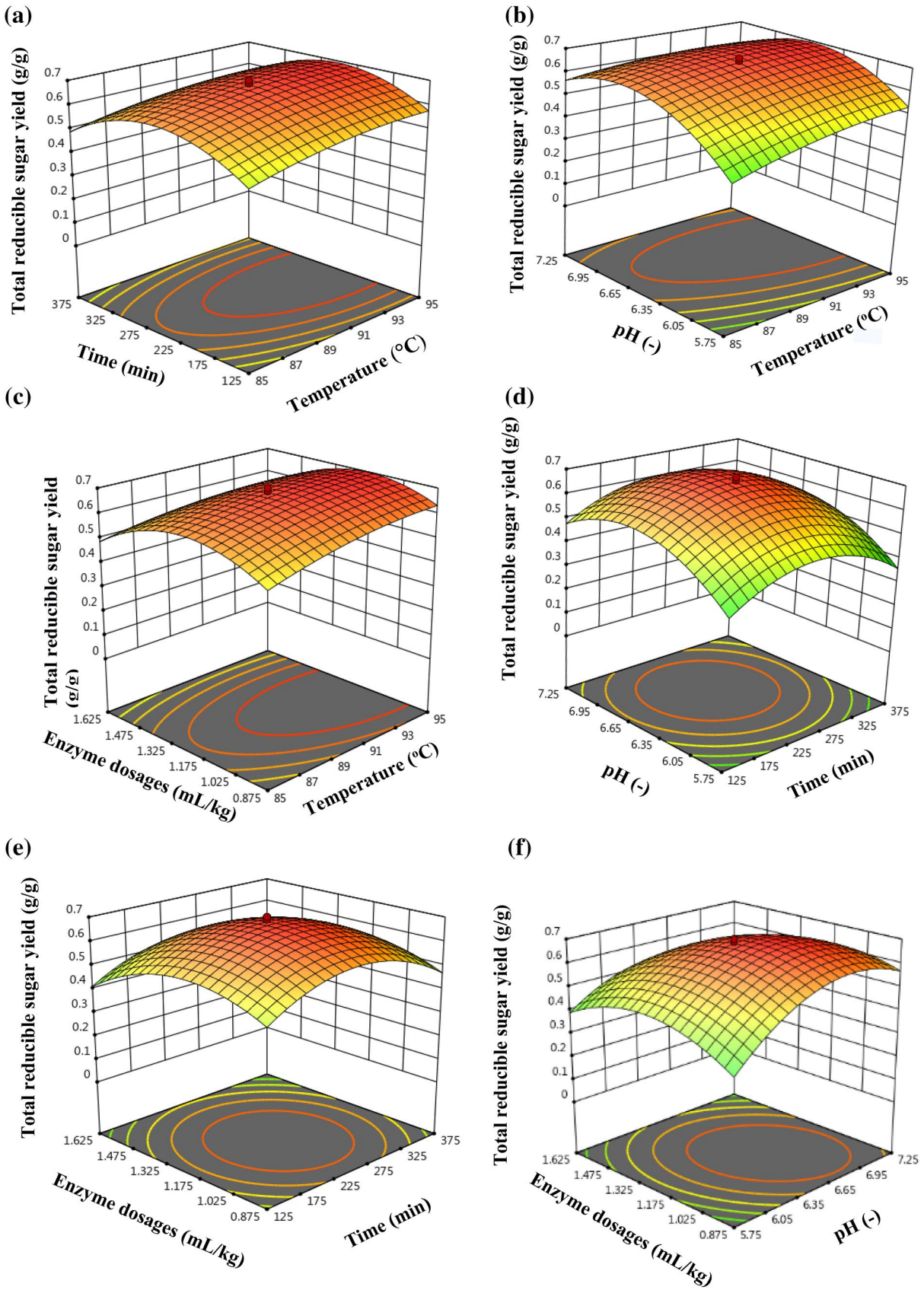
Each experiments has been repeated three times and the average of the these has been reported both for training, test, validation and confirmation data set. The maximum value of standard deviation was found to be 0.0098 from the experimental runs.

The overall performance and potentiality of the neural network model have further been enhanced by incorporating the following techniques -(i) normalization of input and output data, (ii) an optimal number of neurons in the hidden layer (10 neurons) and (iii) dividing the experimental data sets into training, validation and test sets. In the present study, 30 training sets (Table 1), six validation sets (Table 2) and six test sets (Table 3) experimental data have been employed for the development of the ANN model.

To optimize the combination of process parameters (temperature, time, pH and enzyme dosages) which would maximize the production of reducing sugar from the saccharification process using the ANN model, the Genetic algorithm has been used. To search the optimum condition by GA, few generation searches have been carried out. It has been found that after 61 generations, the GA reached the optimal point, beyond which no further improvement of the fitness function was observed. The optimal condition of the saccharification of bio-enzyme and enzyme-treated waste broken rice were: 93 °C (temperature), 250 min (time), 6.5 pH and 1.25 (mL/kg) enzyme dosages. The corresponding maximum total reducing sugar predicted at the optimal condition was found to be 0.704 g/g. The prediction accuracy of the maximized condition was verified by carrying out fresh experiments and a average value of 0.71 g/g was obtained. Hence, the optimal condition predicted by the combined ANN-GA model was found to be fairly acceptable with a deviation of (0.84%) less than 1%.

Table 6 ANOVA analysis for RSM model

Source	Sum of squares	Deg.of freedom	Mean square	Coef. sensitivity	*F* value	*p* value
Model	1.21	14	0.086		285.13	<0.0001 significant
A-temp	0.034	1	0.034	0.020	112.37	<0.0001
B-time	0.0008	1	0.0008	0.034	52.62	<0.012
C-pH	0.069	1	0.069	0.054	229.20	<0.0001
D-Enzyme dosages	0.028	1	0.028	0.030	93.73	<0.0001
AB	0.0008	1	0.0008	0.0006	2.72	0.1198
AC	0.013	1	0.013	0.0011	44.49	<0.0001
AD	0.0052	1	0.0052	0.0007	18.07	0.0007
BC	0.00003	1	0.00003	0.0016	0.098	0.7583
BD	0.00004	1	0.00004	0.001	0.15	0.7021
CD	0.034	1	0.034	0.0018	118.08	<0.0001
A^2	0.020	1	0.020	0.0004	69.17	<0.0001
B^2	0.45	1	0.45	0.0009	1565.43	<0.0001
C^2	0.50	1	0.50	0.0029	1736.01	<0.0001
D^2	0.35	1	0.35	0.0008	1202.42	<0.0001
Lack of fit	0.0032	10	0.0003		1.5	0.356 non significant
Std. Dev.	0.017		R^2		0.9964	
Mean	0.36		Adjusted R^2		0.9931	
CV (%)	3.74		Predicted R^2		0.9833	
PRESS	0.021					


(a)
Total reducible sugar yield (g/g)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
375
325
275
225
175
125
Time (min)
85
87
89
91
93
95
Temperature (°C)
(b)
Total reducible sugar yield (g/g)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
7.25
6.95
6.65
6.35
6.05
5.75
pH (-)
85
87
89
91
93
95
Temperature (°C)
(c)
Total reducible sugar yield (g/g)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1.625
1.475
1.325
1.175
1.025
0.875
Enzyme dosages (mL/kg)
85
87
89
91
93
95
Temperature (°C)
(d)
Total reducible sugar yield (g/g)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
7.25
6.95
6.65
6.35
6.05
5.75
pH (-)
125
175
225
275
325
375
Time (min)
(e)
Total reducible sugar yield (g/g)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1.625
1.475
1.325
1.175
1.025
0.875
Enzyme dosages (mL/kg)
125
175
225
275
325
375
Time (min)
(f)
Total reducible sugar yield (g/g)
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1.625
1.475
1.325
1.175
1.025
0.875
Enzyme dosages (mL/kg)
5.75
6.05
6.35
6.65
6.95
7.25
pH (-)

◂**Fig. 7** Effect of various process parameters on the yield total reducing sugar (TRS) predicted by RSM **a** Effect of time and temperature, **b** Effect of temperature and pH, **c** Effect of enzyme dosages and temperature, **d** Effect of pH and time, **e** Effect of enzyme dosages and time, **f** Effect of enzyme dosages and pH

Response surface methodology modeling and optimization

The central composite design has been utilized for the determination of factors that have an impact on the bio-enzymatic and enzymatic saccharification of broken rice. Table 6 represents the summary of the results obtained from the study of ANOVA. The acceptability of the model was determined by calculating the statistical parameters like *F* value and p value from ANOVA. The *p* value of less than 0.05 implies that the corresponding variable has a significant effect on model development (Ramaraj et al. 2019). In the present study, the model has a high *F* value (285.13) and very low *p* value (< 0.0001) which are in the acceptable range for the higher confidence level of correlation (Pasandideh et al. 2006). It is evident that the independent inputs like temperature, time, pH and enzyme dosages have a significant influence on the bio-enzymatic and enzymatic saccharification process for the production of total reducing sugar. Hence, the quadratic model developed using RSM is quite appropriate to predict the experimental data. The model has a coefficient of determination (R^2) of 0.996, indicating the accuracy of the model is quite higher. Both the values of the predicted R^2 (0.983) and the adjusted R^2 (0.993) which are in the acceptable range and *are* in reasonable agreement with that reported by others (Panichikkal et al. 2018). The lack of fit for the RSM model is found to be 0.356 which is non-significant; hence the quadratic model relating the inputs and output is reasonably acceptable with a higher degree of confidence level. The coefficient of variance (CV) is also used to identify the deviation of residual variation relative to the size of the mean. A lower value of CV indicates more precision and reliability on the experimental result. The standard deviation and mean values were estimated to be 0.017 and 0.36. Another statistical parameter, predicted residual sum of squares (PRESS), also describes the fitness and acceptability of the model. The lower the PRESS, the better is the fitness of data and higher acceptability of the model. For the second-order model developed in the present study, the value of the PRESS was estimated to be 0.021.

Final equation in terms of actual factors

$$\begin{aligned}\text{T RS yield} = &-28.202 + 0.267 * a + 5.213 \times 10^{-3} * b + 4.089 * c + 3.854 * d - \\ &1.121 \times 10^{-5} * a * b - 7.555 \times 10^{-3} * a * c - 9.630 \times 10^{-3} * a * d - \\ &1.420 \times 10^{-5} * b * c - 3.533 \times 10^{-5} * b * d - 0.164 * c * d - \\ &1.086 \times 10^{-3} * a^2 - 8.228 \times 10^{-6} * b^2 - 0.241 * c^2 - 0.801 * d^2\end{aligned}$$

a = temperature, b = time, c = pH, d = enzyme dosages.

The performance of the RSM model has been tested by comparing the RSM predicted reducing sugar yield and that obtained from the experiment. A considerable agreement with good fitness of the RSM model is observed. On plotting the predicted vs. the experimental sugar yield shows that the maximum number of the experimental data is located on or proximity of 45° diagonal (not shown here, but similar to Fig. 6) indicating that the predicted results obtained from RSM model are in excellent agreement with the experiment. Three-dimensional plots are better representation for the effect of various process parameters like temperature, time, pH and enzyme dosages on the total yield of reducing sugar.

The central composite design is an inbuilt module in RSM to provide the design of experiments depending on the no of input parameters. In the present study, the CCD suggested thirty numbers of experimental runs of different combinations of four input parameters, where six trials combinations are the same (Table 1). The non-linear model of RSM is mainly based on these experimental data sets. The six trials represent the replication of the central points and is the measure of the confidence level of the experimental data. The number thirty follows relationship: $2^n + 2n + p$, where n is the no of independent process parameters and p is the replication of the central points. For our present study n and p are 4 and 6, respectively. So the experimental data set consists of six numbers of replicated center point information regarding experiments and uses the same for evaluating the pure sum of the square error. The information is used during the development of a quadratic relationship by the software RSM (Ghelich et al. 2019).

The careful observation of experimental data relives that the maximum yield of reducible sugar is 0.689 g/g and the same is obtained at 90 °C temperature, 250 min time, 6.5 pH and 1.25 mL/kg enzyme dosages. The RSM predicted sugar yield at this input combination is 0.682 g/g. However, the RSM predicted the optimum process conditions for maximum sugar yield of 0.7025 at 93.5 °C temperature, 244.8 min time, 6.63 pH and 1.16 mL/kg enzyme dosages.

Figure 7 represents the variation of any two process parameters independently, keeping the other two at the optimum values as predicted by the RSM. The 3D surface plots (Fig. 7) provide better insight regarding the location of maximum sugar yield (0.7025) at the optimum process inputs. Figure 7a shows the variation of the total yield of reducing sugar as a function of temperature and time in the form of the 3D surface plot, while the pH and enzyme

dosages remain constant at 6.5 and 1.25 mL/kg, respectively. It is observed that the percentage yield is very low when the residence time and operating temperature of the saccharification process are maintained at very low values. The percentage yield slowly rises as both the temperature and time are increased and eventually, the yield shows maxima at 0.682 g/g for the combination of temperature and time to be 90 °C and 250 min, respectively. However, a further increase the factors like temperature and time, the yield starts falling due to the feedback inhibition of the system. The similar surface plots for variation of a pair of factors keeping the other two factors constant are presented in Fig. 7b–f.

Figure 7b represents the effect of the temperature and pH on the yield of reducing sugar for the constant time and enzyme dosages of 250 min and 1.25 mL/kg, respectively. At low temperature and low pH, the biochemical reaction rate becomes very slow due to the lower value of the rate constant at low temperature and negative catalytic effect at high acidic medium (pH 5-6.5). As the temperature and pH are increased, the biochemical reaction rate starts to rise, and the yield of reducing sugar gradually rises to show a maximum. A further rise in the factors, the recessive feedback effect may be predominant, and the yield starts to fall (Helle et al. 1993).

The effect of temperature and enzyme are assessed from the response surface equation and is depicted in Fig. 7c in the form of 3D surface plot. When the values of enzyme dosages and the temperature are very low, the yield is also meager due to slower reaction rate and a lower concentration of enzyme dosage, which cannot significantly modify the surface property of the substrate suitable for biochemical reaction. Gradually the yield increases with the increase of temperature and enzyme. The yield shows a maximum for enzyme dosages of 1.25 mL/kg; on the further rise of enzyme dosages, negative surface inhibition occurs and yield starts falling.

The action of pH and time are the other two important factors, and the effect is described in Fig. 7d. The other factors enzyme dosages and temperature (1.25 mL/kg and 90 °C) are kept constant. At lower pH and very low residence time, the yield is very low due to incomplete saccharification reaction. At higher residence time the progress of the biochemical reaction proceeds in the forward direction resulting in the higher yield; it is maximum at a residence time of 250 min. A further rise of residence time may lead to the simultaneous other parallel reactions using reducing sugar as a substrate and thereby, the yield eventually starts to fall.

The variation of sugar yield due to pH and enzyme dosages is depicted in Fig. 7e. The yield value is near about 0.022 (g/g) when the enzyme dosages 0.875 (mL/kg) and pH 5.75. As the pH and enzyme dosages values are increased the yield starts to rise, and at 6.5 pH and 1.25 mL/kg enzyme dosages, the yield becomes maximum. Beyond that both pH and enzyme dosing adversely affects the biochemical reaction leading to lower yield.

Figure 7f illustrates the effect of time and enzyme dosages on the yield of reducing sugar. The yield is low at the lower level of both the factors. The maximum yield is detected at the residence time of 250 min and 1.25 mL/kg enzyme dosages.The enzyme loading is one of the important factors for surface inhibition of the substrate, and maximum activity is found to be at 1.25 mL/kg. However, at higher levels of residence time, the yield starts to decrease due to the consumption of the reducing sugar by other side reactions (Helle et al. 1993).

Optimization using RSM technique

Using the RSM technique, all the individual system parameters are varied independently within the bound of each parameter. The optimum values are found to be- temperature 93.5 °C, time 244.8 min, pH 6.63 and enzyme dosages 1.16 (mL/kg), respectively. At these optimum inputs conditions, the maximum yield of total reducing sugar is found to be 0.7025 (g/g). To validate the RSM prediction, a fresh experiment is carried out with the above-mentioned process inputs. The same is repeated thrice, and the average of the three is found to be 0.715(g/g) for total sugar yield. The percentage deviation is calculated to be 1.8%. Hence, the above optimum condition predicted by RSM is reasonably acceptable.

Comparative performance of hybrid ANN-GA and RSM

Predictive performance

The predictive performance of the RSM and ANN model developed above is further tested against the fresh, experimental results. Five additional experimental runs (the confirmation dataset) were carried out which were the different combinations of inputs and had not been used in previous experimentations, as discussed in subsection "ANN-GA modeling and optimization" (Tables 1, 2, 3). The confirmation dataset and the model (RSM and ANN) predicted total reducing sugar yield is presented in Table 7. The coefficient of determination for this new data set is estimated to be 0.987 and 0.993 for RSM and ANN models, respectively. The RMSE is also calculated both for RSM and ANN models; the values are found to be 0.102 and 0.078. The corresponding RMSE and CC values are also presented in Table 8.

The relative mean error (RME) is also estimated for the design set separately for ANN and RSM and is observed to be 2.4% and 3.8%, respectively. It is evident from the

Table 7 Experimental data of total reducing sugar used as confirmation data set to test the ANN and RSM models

Run no	Temperature (°C)	Time (min)	pH	Enzyme dosage (ml/kg)	Total reducing sugar(g/g)		
					Experimental yield	RSM predicted yield	ANN predicted yield
1	70	15	6.0	1.0	0.092	0.090	0.094
2	70	90	6.0	1.0	0.112	0.119	0.116
3	85	200	7.5	1.5	0.279	0.268	0.291
4	85	300	7.5	2.0	0.189	0.182	0.187
5	75	300	7.0	1.0	0.449	0.459	0.442

Table 8 Comparison of root mean square error, relative mean error and correlation coefficient, R^2 for ANN and RSM

Parameters	Design data		Confirmation data	
	ANN	RSM	ANN	RSM
RMSE	0.010	0.012	0.078	0.102
RME (%)	2.4	3.8	2.6	3.7
Coefficient of determination, R^2	0.997	0.996	0.993	0.987
Correlation coefficient (CC)	0.999	0.998	0.982	0.975

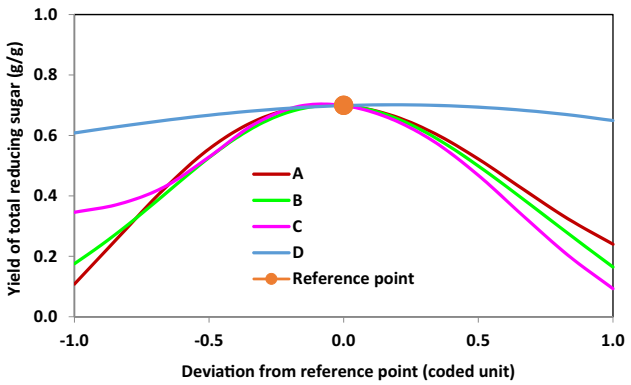


Fig. 8 Sensitivity analysis by perturbation method by ANN model. A-pH, B- time, C-enzyme dosages, D-temperature

analysis that ANN has better potential to replicate the saccharification process than that by RSM (Nasab et al. 2019). The better adaption by ANN is because it uses a non-linear Tansig function, which is sigmoidal and can approximate any non-linear relationship rather than the quadratic polynomial adapted by RSM.

Sensitivity analysis

The sensitivity analysis is one of the important techniques to detect the most sensitive input parameter for a given system. In the present saccharification study, four inputs have been identified as regulating parameters of the process. The ANOVA analysis provides the coefficient of sensitivity (coef. sensitivity column) for each parameter and is reported in Table 6. From the table, it can be observed that the pH (0.054) parameter has the highest impact on the present process. Then time (0.034), which is the second important parameter. After that, enzyme dosages are also important, which affects the process considering. The sensitivity analysis carried out from the ANN-GA hybrid model is presented in Fig. 8. The order of sensitivity is found to be pH > time > enzyme dosages > temperature.

Comparison of optimum process parameters

The optimum process conditions are examined by hybrid ANN-GA and RSM model separately. Table 9 compares the optimum process condition at which the maximum reducible sugar is obtained during the saccharification process treated with bio-enzyme and enzyme. The maximum yield of reducing sugar, as predicted by the ANN-GA hybrid model is 0.704 (g/g), while under similar process conditions, the experimentally obtained value is 0.71 (g/g). Similarly, the RSM predicted maximum value is found to be 0.7025 (g/g) against the experimental value of 0.715 (g/g). The percentage deviation was calculated to be 0.84% and 1.8% for ANN-GA hybrid and RSM, respectively. In both cases, the percentage error is very small and acceptable. However, the hybrid ANN-GA model adopts the experimental results better than the RSM model.

Table 9 Optimized process parameters for maximum TRS yield from ANN-GA and RSM models

Optimum by	Temperature (°C)	Time (min)	pH	Enzyme dosages (ml/kg)	Predicted yield (g/g)	Experimental yield (g/g)	Deviation (%)
ANN-GA hybrid	93	250	6.5	1.25	0.704	0.71	0.84
RSM	93.5	244.8	6.63	1.16	0.7025	0.715	1.8

Table 10 Economic and cost analysis

Basis: 10 kg of broken rice	Final product: 3500 mL of (70% (v/v) ethanol)
Input cost	Final product price
Reject broken rice Rs. 10 × 10 = 100.00/–	Rs.220/L (70% (v/v) ethanol)
Enzyme 1.25 mL Rs. 10 × 1.25 = 10.25/–	Selling price- Rs.220 × 3.5 = 770
Yeast & bakhar Rs. = 100.00/–	The selling price of fish feed = 20
Purification & other cost Rs. = 80.00/–	
Total cost Rs. = 290.25/–	
Production cost Rs./L 290.25/3.5 = 82.00/–	Product price = 790/–

Profit: (790–290.25)/10 = Rs.50.00/- per kg of broken rice processed

End-uses of total reducible sugar (TRS)

1 kg of broken rice contains about 79.89% of cellulose (experimentally determined) (Sluiter et al. 2012). After saccharification of broken rice with a bio-enzyme (bakhar) and enzyme (α-amylase), the maximum yield of TRS was observed to be 160 mg/mL (from spectroscopic analysis) for a total volume of 3500 mL saccharified solution. Hence, the solution contains (160 × 3500/1000 =) 560 g of TRS which corresponds to the conversion of TRS from starch is (560 × 100/798.9 =) 70.0% on fermentation of stained saccharified solution at 30 °C and pH of 6.5 using the commercial yeast (*Saccharomyces cerevisiae*) for 60 h the final solution found to contain 70 mg/mL of ethanol (3000 mL of solution). The two stages vacuum distillation process was employed at 10 millibar pressure to obtain 350 mL ethanol of concentration 558 mg/mL (detected in HPLC analysis). The yield of final ethanol is found to be (558 × 350/1000 =) 195 g. Hence the ethanol yield is found to be (195 × 100/560 =) 34.8%.

The experimental findings by Kumar et al. (2017) reveal that the maximum conversion of TRS (glucose) to ethanol via the fermentation route is about 47%. The recovery of reducible sugar from starchy biomass like rice and its derivatives solely depends on the nature of starch and the use of different enzymes/hydrolyzing agents like alpha-amylase, gluco-amylase, sulphuric acid etc. (Suryawanshi et al. 2018). Schneider et al. (2018) achieved 70% of reducible sugar from broken rice as raw substrate (80.1%) when enzymes alpha-amylase and gluco-amylase both were used. While the researchers Omar et al. (2016) reported the release of 23.34% of total reducing sugar from rice (91.43%) after using H_2SO_4 as a hydrolyzing agent.

The details on the fermentation process using the TRS from saccharification as discussed above and the economic analysis of the combined process of saccharification followed by fermentation, is aimed in our future communication.

Economical and cost analysis

For 1 kg of raw, broken rice used 70 mg/mL of ethanol was obtained immediately after fermentation. On two-stage vacuum distillation the finally 350 mL of ethanol 557 mg/mL, which is equivalent to 70% (v/v) alcohol solution (assuming density of ethanol 0.789 mg/mL at 30 °C). The production cost is found to be Rs.82/L, the details on cost data are presented in Table 10. Whereas, according to Kang et al. 2019, production cost was estimated to be US$1.31/L, which is equivalent to Rs. 93.5/L (1 US$ = 71.4 INR). Hence, the process proposed in the present article is economical.

Conclusion

The results for the production of bioethanol using waste broken rice in a laboratory-scale set up is proved to be one of the promising process routes due to the abundant availability of low-cost feedstock. The study is carried out using another low-cost bio-enzyme bakhar as a hydrolyzing agent for the hydrolysis of starch in broken rice. The use of bio-enzyme bakhar considerably reduces the use of alpha-amylase using the starch hydrolysis process. The four process parameters for the saccharification process are optimized using the RSM and ANN-GA hybrid model separately. The statistical parameters like the coefficient of determination, relative mean error, the *p* value for ANOVA analysis and RMSE are estimated for both the RSM and ANN-GA hybrid model. Both the models are found to be effective as an optimizing tool; however, the ANN-GA hybrid model better replicates the experimental findings of the bio-enzymatic and enzymatic saccharification process. The TRS so obtained subsequently will be utilized in the fermentation process using *Saccharomyces cerevisiae* to produce bio-ethanol. The economic analysis shows the method is cost-effective with a yield of reducible sugar 70.4% and the final ethanol yield of 195 g (195 × 100/560 =) 34.8%. Even though

the concentration of ethanol immediately after fermentation is lower (70 mg/mL), the concentration can further be enhanced by upgrading the industrial-grade *Saccharomyces cerevisiae* and optimizing the other process parameters during the fermentation process, and the same is aimed to elaborate in our future communication.

Acknowledgments The authors acknowledge the experimental infrastructure support received from the North-East Technical Development Group of CSIR-CMERI, Durgapur, India. The computational and analytical instrumentation support received from DST-FIST, GOI by Department of Chemical Engineering, National Institute of Technology, Durgapur, India.

Compliance with ethical standards

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

Research involving human participants and/or animals Not applicable.

Informed consent Not applicable.

References

Achinas S, Euverink GJW (2016) Consolidated briefing of biochemical ethanol production from lignocellulosic biomass. Electron J Biotechnol 23:44–53. https://doi.org/10.1016/j.ejbt.2016.07.006

Aditiya HB, Mahlia TMI, Chong WT, Nur H, Sebayang AH (2016) Second generation bioethanol production: a critical review. Renew Sustain Energy Rev 66:631–653. https://doi.org/10.1016/j.rser.2016.07.015

Ahmad F, Jameel AT, Kamarudin MH, Mel M (2011) Study of growth kinetic and modeling of ethanol production by *Saccharomyces cerevisae*. Afr J Biotechnol 16(81):18842–18846. https://doi.org/10.5897/ajb11.2763

Alam MA, Yuan T, Xiong W, Zhang B, Lv Y, Xu J (2019) Process optimization for the production of high-concentration ethanol with *Scenedesmus raciborskii* biomass. Biores Technol 294:122219. https://doi.org/10.1016/j.biortech.2019.122219

Ali KF, Sulaiman R, Elamir AM (2014) Implementations of Back Propagation Algorithm in Ecosystems Applications. Proceedings of the International conference on condensed matter physics 2014 (ICCMP 2014). https://doi.org/10.1063/1.4915861

Azhar SHM, Abdulla R, Jambo SA, Marbawi H, Gansau JA, Faik AAM, Rodrigues KF (2017) Yeasts in sustainable bioethanol production: a review. Biochem Biophys Rep 10:52–61. https://doi.org/10.1016/j.bbrep.2017.03.003

Balat M, Balat H (2009) Recent trends in global production and utilization of bio-ethanol fuel. Appl Energy 86:2273–2282. https://doi.org/10.1016/j.apenergy.2009.03.015

Bankole SA, Osho A, Joda AO, Enikuomehin OA (2005) Effect of drying method on the quality and storability of 'egusi' melon seeds (*Colocynthis Citrullus* L.). Afr J Biotechnol 4(8):799–803

Betiku E, Taiwo AE (2015) Modeling and optimization of bioethanol production from breadfruit starch hydrolyzate vis-à-vis response surface methodology and artificial neural network. Renew Energy 74:87–94. https://doi.org/10.1016/j.renene.2014.07.054

Chirayil CJ, Joy J, Mathew L, Mozetic M, Koetz J, Thomas S (2014) Isolation and characterization of cellulose nanofibrils from *Helicteres isora* plant. Ind Crops Prod 59:27–34. https://doi.org/10.1016/j.indcrop.2014.04.020

Chu-Ky S, Pham TH, Bui KLT, Nguyen TT, Pham KD, Nguyen HDT, Luong HN, Tu VP, Nguyen TH, Ho PH, Le TM (2016) Simultaneous liquefaction saccharification and fermentation at very high gravity of rice at pilot scale for potable ethanol production and distillers dried grains composition. Food Bioprod Process 98:79–85. https://doi.org/10.1016/j.fbp.2015.10.003

Claudia Conesa, Ibáñez Civera Javier, Lucía Seguí, Pedro Fito, Nicolás Laguarda-Miró (2016) An electrochemical impedance spectroscopy system for monitoring pineapple waste saccharification. Sensors 16(2):188. https://doi.org/10.3390/s16020188

Das S, Ganguly A, Dey A, Ting YP, Chatterjee PK (2014) Characterization of water hyacinth biomass and microbial degradation of the biomass under solid state fermentation using a lignocellulolytic fungus (*Alterneria Spp* NITDS1). J Chem Biol Phys Sci 4:2279–2293

Das S, Bhattacharya A, Haldar S, Ganguly A, Sai Gu, Ting YP, Chatterjee PK (2015) Optimization of enzymatic saccharification of water hyacinth biomass for bio-ethanol: comparison between artificial neural network and response surface methodology. Sustain Mater Technol 3:17–28. https://doi.org/10.1016/j.susmat.2015.01.001

de Schneider Rosana de Cassia S, Junior Célio S, Fornasier F, Diego de S, Corbellini Valeriano A (2018) Bioethanol production from broken rice grains. Interciencia 43(12):846–851

Devarapalli M, Atiyeh Hasan K (2015) A review of conversion processes for bioethanol production with a focus on syngas fermentation. Biofuel Res J 2(3):268–280. https://doi.org/10.18331/brj2015.2.3.5

Ghaffar SH, Fan M (2013) Structural analysis for lignin characteristics in biomass straw. Biomass Bioenerg 57:264–279. https://doi.org/10.1016/j.biombioe.2013.07.015

Ghelich R, Jahannama MR, Abdizadeh H, Torknik FS, Vaezi MR (2019) Central composite design (CCD)- Response surface methodology (RSM) of effective electrospinning parameters on PVP-B-Hf hybrid nanofibrous composites for synthesis of HfB_2-based composite nanofibers. Compos B 166:527–541. https://doi.org/10.1016/j.compositesb.2019.01.094

Gronchi N, Favaro L, Cagnin L, Brojanigo S, Pizzocchero V, Basaglia M, Casella S (2019) Novel yeast strains for the efficient saccharification and fermentation of starchy by-products to bioethanol. Energies 12(4):714. https://doi.org/10.3390/en12040714

Hagan MT, Demuth HB, Beale MH and Jesús OD (2014) Neural Network Design (2nd Edition)

Helle SS, Duff SJB, Coope DG (1993) Effect of surfactants on cellulose hydrolysis. Biotechnol Bioeng 42:611–617. https://doi.org/10.1002/bit.260420509

Hickert LR, Cunha-Pereira FD, Souza-Cruz PBD, Rosa CA, Ayub MA (2012) Ethanogenic fermentation of co-cultures of *Candida shehatae* HM 52.2 and *Saccharomyces cerevisiae* ICV D254 in synthetic medium and rice hull hydrolysate. Bioresource Technology 131:508–514. https://doi.org/10.1016/j.biortech.2012.12.135

Kang KE, Jeong JS, Kim Y, Min J, Moon SK (2019) Development and economic analysis of bioethanol production facilities using lignocellulosic biomass. J Biosci Bioeng 128(4):475–479. https://doi.org/10.1016/j.jbiosc.2019.04.004

Kumar R, Alak KG, Parimal P (2017) Fermentative energy conversion: renewable carbon source to biofuels (ethanol) using *Saccharomyces cerevisiae* and downstream purification through solar driven membrane distillation and nanofiltration. Energy Convers Manage 150:545–557. https://doi.org/10.1016/j.enconman.2017.08.054

Li H, Jiao A, Xu X, Wu C, Wei B, Hu X, Jin Z, Tian Y (2013) Simultaneous saccharification and fermentation of broken rice: an enzymatic extrusion liquefaction pretreatment for Chinese rice wine

production. Bioprocess Biosyst Eng 36(8):1141–1148. https://doi.org/10.1007/s00449-012-0868-0

Stöcker M (2008) Biofuels and Biomass-to-Liquid Fuels in the Biorefinery: catalytic Conversion of Lignocellulosic Biomass Using Porous Materials. Angew Chem Int Ed 47(48):9200–9211. https://doi.org/10.1002/anie.200801476

Myburgh MW, Cripwell RA, Favaro L, Zyl WHV (2019) Application of industrial amylolytic yeast strains for the production of bioethanol from broken rice. Biores Technol 294:122222. https://doi.org/10.1016/j.biortech.2019.122222

Nasab SG, Semnani A, Teimouri A, Yazd MJ, Isfahani TM, Habibollahi S (2019) Decolorization of crystal violet from aqueous solutions by a novel adsorbent chitosan/nanodiopside using response surface methodology and artificial neural network-genetic algorithm. Int J Biol Macromol 124:429–443. https://doi.org/10.1016/j.ijbiomac.2018.11.148

Nawaz MA, Gaiani C, Shu F, Bhandari B (2016) X-ray photoelectron spectroscopic analysis of rice kernels and flours: measurement of surface chemical composition. Food Chem 212:349–357. https://doi.org/10.1016/j.foodchem.2016.05.188

Omar KA, Salih BM, Abdulla NY, Hussin BH, Rassul SM (2016) Evaluation of Starch and Sugar Content of Different Rice Samples and Study their Physical Properties. Indian J Nat Sci 6(36):11084–11093

Pandey G, Zhang B, Chang AN, Myers CL, Zhu J, Kumar V, Schadt EE (2010) An integrative multi-network and multi-classifier approach to predict genetic interactions. PLoS Comput Biol 6(9):e1000928. https://doi.org/10.1371/journal.pcbi.1000928

Panichikkal AF, Prakasan P, Nair UK, Valappil MK (2018) Optimization of parameters for the production of biodiesel from rubber seed oil using onsite lipase by response surface methodology. 3 Biotech 8:459. https://doi.org/10.1007/s13205-018-1477-7

Pasandideh SHR, Niaki STA (2006) Multi-response simulation optimization using genetic algorithm within desirability function framework. Appl Math Comput 175(1):366–382. https://doi.org/10.1016/j.amc.2005.07.023

Ramaraj R, Unpaprom Y (2019) Optimization of pre-treatment condition for ethanol production from *Cyperus diformis* by response surface methodology. 3 Biotech 9:218. https://doi.org/10.1007/s13205-019-1754-0

Rao KJ, Kim CH, Rhee SK (2000) Statistical optimization of medium for the production of recombinant hirudin from *Saccharomyces cerevisiae* using response surface methodology. Process Biochem 35:639–647. https://doi.org/10.1016/S0032-9592(99)00129-6

Rosillo-Calle F, Hall DO (1987) Brazilian alcohol: food versus fuel? Biomass 12:97–128. https://doi.org/10.1016/0144-4565(87)90050-3

Saha K, Dwibedi P, Ghosh A, Sikder J, Chakraborty S, Curcio S (2018) Extraction of lignin, structural characterization and bioconversion of sugarcane bagasse after ionic liquid assisted pretreatment. 3 Biotech 8:374. https://doi.org/10.1007/s13205-018-1399-4

Sluiter A, Hames B, Ruiz R, Scarlata C, Sluiter J, Templeton D, and Crocker D(2012). Determination of structural carbohydrates and lignin. Laboratory Analytical Procedure (LAP) 2008

Suryawanshi OP, Khokhar D, Patel S (2018) Effect of different pretreatment methods on reducing sugar of rice substrate to enhance the ethanol yield. Int J Curr Microb Appl Sci 7(03):2715–2733. https://doi.org/10.20546/ijcmas.2018.703.314

Taheri S, Lakmehsari MS, Soltanabadi A (2017) Separation based adsorption of ethanol–water mixture in azeotropic solution by single-walled carbon, boron-nitride and silicon-carbide nanotubes. J Mol Graph Model 75:149–164. https://doi.org/10.1016/j.jmgm.2017.05.004

Teslić N, Bojanić N, Rakić D, Takači A, Zeković Z, Fišteš A, Bodroža-Solarov M, Pavlić B (2019) Defatted wheat germ as source of polyphenols—Optimization of microwave assisted extraction by RSM and ANN approach. Chem Eng Process 143:107634. https://doi.org/10.1016/j.cep.2019.107634

U.S. Energy Information Administration (2019). http://www.eia.gov/aeo

Wang MQ, Han J, Haq Z, Tyner WE, Wu M, Elgowainy A (2011) Energy and greenhouse gas emission effects of corn and cellulosic ethanol with technology improvements and land use changes. Biomass Bioenerg 35(5):1885–1896. https://doi.org/10.1016/j.biombioe.2011.01.028

Yoon LW, Ang TN, Ngoh GC, Chua ASM (2012) Regression analysis on ionic liquid pre-treatment of sugarcane bagasse and assessment of structural changes. Biomass Bioenerg 36:160–169. https://doi.org/10.1016/j.biombioe.2011.10.033

Zhang J, Wang Y, Zhang L, Zhang R, Liu G, Cheng G (2014) Understanding changes in cellulose crystalline structure of lignocellulosic biomass during ionic liquid pre-treatment by XRD. Biores Technol 151:402–405. https://doi.org/10.1016/j.biortech.2013.10.009