Ultrasonics 98 (2019) 72–81

Contents lists available at ScienceDirect

Ultrasonics

journal homepage: www.elsevier.com/locate/ultras

Machine learning modelling for the ultrasonication-mediated disruption of recombinant *E. coli* for the efficient release of nitrilase

Kiran D. Bhilare[a], Mahesh D. Patil[a], Sujit Tangadpalliwar[b], Ashok Shinde[a], Prabha Garg[b], Uttam Chand Banerjee[a,*]

[a] Department of Pharmaceutical Technology (Biotechnology), National Institute of Pharmaceutical Education and Research, Sector-67, S.A.S. Nagar, 160062 Punjab, India
[b] Department of Pharmacoinformatics, National Institute of Pharmaceutical Education and Research, Sector-67, S.A.S. Nagar, 160062 Punjab, India

ARTICLE INFO

Keywords:
Machine learning model
Escherichia coli
Ultrasonication
Multi-layer perceptron
Nitrilase

ABSTRACT

The ultrasonication-mediated cell disruption of recombinant *E. coli* was modeled using three machine learning techniques namely Multiple linear regression (MLR), Multi-layer perceptron (MLP) and Sequential minimal optimization (SMO). The four attributes were cellmass concentration (g/L), acoustic power (A), duty cycle (%) and treatment time of sonication (min). For the three responses (nitrilase, total protein release and cell disruption) MLP model was found to be at par with RSM model in terms of generalization as well as prediction capability. Nitrilase release was significantly influenced by the cellmass concentration so was in case of total protein release. Fraction of cells disrupted was heavily influenced by acoustic power and sonication time. Almost 32 U/mL nitrilase could be released for 300 g/L cellmass concentration when sonicated at 225 W for 1 min with 20% duty cycle.

1. Introduction

Since many years, microbial cells have remained as good source of option for many proteins and other bioproducts. Extensive research carried out in the field of molecular biology has made large scale production of industrially relevant proteins very much possible [1]. Till date, most of the recombinant products reported in literature are produced intracellularly and are retained inside the cell itself [2]. In such cases, any unit operation that gives maximum disruption efficiency within minimal time and cost will always be desirable [3]. Very often, high power-low frequency ultrasound treatment is used to carry out cell disruption at laboratory scale [4–6]. It has low operating cost and does not require sophisticated equipment [7].

High energy sonic waves (frequency > 20 MHz) generated from the tip of sonication probe generate eddie currents that can intensely damage the cellular membrane of the microorganisms [8,9]. Nitrilases have wide applications such as chiral production of pharmaceutical intermediates, bioremediation etc. [10,11].

In the present report, the ultrasound-mediated *Escherichia coli* cell disruption was characterized for the following dependent variables (responses): (1) release of nitrilase (target product); (2) release of cytosolic proteins; and (3) extent of cell disruption. A full factorial central composite design (CCD) was designed with the help of Design Expert™. Software. The independent/operational factors were: cellmass concentration, sonication time, duty cycle and acoustic power. The resultant data obtained was used to compare the prediction efficiency of different machine learning models viz:- MLP, SMO and MLR. Also, an attempt was made to explain the effect of process variables of the sonication on the kinetics of cell-disruption process, nitrilase release and total protein release after cell lysis. Usage of machine learning-based prediction models for optimization studies of *E. coli* cell disruption in a laboratory scale ultrasonicator is the novelty of the present work.

2. Materials and methods

2.1. Microorganism and cultivation conditions

E. coli cells overexpressing nitrilase enzyme were cultivated in 14 L bioreactor (BioFlo 310) and harvested as per previously reported method [12].

2.2. Mechanical cell disruption by ultrasonication

E. coli cell suspension was subjected to ultrasonication through a horn-type sonicator (Vibra-cell 750; NJ, US). During sonication treatment, it was ensured that the titanium alloy tip was immersed at the

* Corresponding author.
E-mail addresses: prabhagarg@niper.ac.in (P. Garg), ucbanerjee@niper.ac.in (U.C. Banerjee).

https://doi.org/10.1016/j.ultras.2019.06.006
Received 20 August 2018; Received in revised form 15 May 2019; Accepted 9 June 2019
Available online 10 June 2019
0041-624X/

mid-point of the suspended cells inside the plastic tube. The parameters for cellmass concentration, total sonication time, acoustic power and duty cycle were as prescribed by the Design Expert™ 8.0 software (Stat-Ease Inc., Minneapolis, USA) for the thirty runs. The three desired responses were estimated as explained in the following sections.

2.3. Experimental designs

2.3.1. Response surface methodology

Based on the preliminary studies and data available in the literature, the maximum and minimum values of the factors were selected as mentioned in Table 1 [5,7,13–15]. A total of 30 independent runs were carried out. In the CCD design, a 2nd order polynomial equation was obtained from the Design Expert software thereby elucidating the relationship between the dependent variables (responses) and coded values of independent (affecting) variables.

2.3.2. Multi-layer perceptron (MLP) modeling

MLP classifier uses backpropagation for modelling nonlinear responses and to classify instances, whereas RSM approximates responses in terms of an explicit quadratic function of the independent experimental factors [16,17]. The MLP network can be built by hand, created by an algorithm or both. The network can also be monitored and modified during training time. The nodes in this network are the unthresholded linear units. Tool WEKA (Waikato Environment for Knowledge Analysis) version 3.6.13 was used for the development of the MLP (Multilayer Perceptron) model. WEKA is a state-of the-art data mining system which implements algorithms using JAVA. It can be used to implement algorithms for regression, classification, data preprocessing, etc.

In this research work, we used MLP which is feed-forward backpropagation artificial neural network (ANN) consisting of an input layer, a hidden layer and an output layer. A backpropagation algorithm was used for training the network, in which the weights of the connections between the nodes of the layers were repetitively adjusted so that the output value is as close as possible to the desired output. Four input variables viz. Cell-mass, Pressure and Number of passes were used to feed neural network. Each input value was normalized by normalize function present in weka [WEKA at http://www.cs.waikato.ac.nz/~ml/weka]. The output layer had three neurons corresponding to the Nitrilase, the total protein count and the cell disruption. The hidden layer was having two neurons. The nonlinear activation function sigmoid and learning rate 0.3 was used during training of network.

2.3.3. Multiple linear regression (MLR)

Multiple Linear regression works by estimating coefficients for a line or hyperplane that best fits the training data. It is a very simple regression algorithm, fast to train and can have great performance if the output variable for the given data is a linear combination of inputs [18]. Linear regression is an approach for modeling the relationship between a scalar dependent variable y and one or more explanatory variables (or independent variables) denoted X [19]. In this study, we used, Linear Regression method present in weka. In weka, class for using linear regression for prediction uses the Akaike criterion for model selection, and is able to deal with weighted instances. M5 method is used as attribute selection method during model development process. Algorithm involved behind these methods is: Removal of those attributes with the smallest standardized coefficient unless no improvement is seen in the estimate of the error given by the Akaike information criterion [20].

2.3.4. SMOreg (Support Vector Machine for regression)

Support vector machine (SVM) is globally acceptable supervised learning method for classification and regression problems in predictive modeling. Training of an SVM model requires solution of a large number of quadratic programming (QP) optimization problems. SMO algorithm breaks them into a series of smallest possible QP problems. These small QP problems are solved analytically rather than numerical QP optimization which is time consuming. SMOreg implements the support vector machine for regression [21]. The most popular "RegSMOImproved" algorithm was used as RegOptimizer for learning the models. In the present study, experimental data (Table 1) was used to train the SMOreg model using polykernel with cache size 250,007 and exponent 1.0 for the development of models.

2.4. Comparison between statistical and machine learning models

The models were compared for their generalization and predictive capabilities by assessing error parameters as reported in previous publications [5,11,22,23].

2.5. Analytical methods

2.5.1. Nitrilase activity

Nitrilase activity was determined as previously reported method [12]. Modified Berthelot method was used to assess the nitrilase activity. The mean values of the experiments in triplicate were used to train the models.

2.5.2. Total protein quantification

The cell lysate was cleared of the cell debris using high speed centrifugation (7000*g*, 5 min). Cell lysate protein concentration was estimated using Bradford assay [24] and the mean values are reported.

2.5.3. Measurement of the extent of cell disruption

Percent cell disruption (F_d) was calculated as per previously reported method [25].

3. Results

3.1. RSM based optimization of ultrasonication-mediated disruption of E. coli

The individual responses measured were then correlated with the independent factors (coded values) using following equations:

$$Y_1 = 19.23 + 7.37A - 0.46B - 3.05C - 1.31D - 0.76AB - 0.27AC + 0.88AD - 1.67BC - 1.91BD - 0.63CD \quad (1)$$

$$Y_2 = 3.02 + 0.97A - 0.13B - 0.18C - 0.11D + 0.071AB - 0.018AC - 0.037AD - 0.15BC - 0.26BD - 0.23CD - 0.086A^2 + 0.16B^2 - 0.51C^2 + 0.15D^2 \quad (2)$$

$$Y_3 = 51.09 - 1.74A + 11.97B + 17.25C - 2.32D + 1.91AB - 4.72AC + 1.09AD + 0.01BC - 1.31BD - 1.81CD + 11.25A^2 - 2.58B^2 - 8.73C^2 + 0.83D^2 \quad (3)$$

In the above equations, Y_1 denotes nitrilase activity (U/mL); Y_2 denotes released intracellular protein (g/L); Y_3 denotes cells disruption (%); A is the sonication time (min); B is the cellmass concentration (g/L); and C is the acoustic power and D is the duty cycle of the sonicator. Analysis of variance (ANOVA) was used to assess the correctness of the models. [Supplementary Tables 1–3].

The ANOVA for nitrilase response is given in Supplementary Table 1. The value of correlation coefficient (R) was 0.988, indicating a good agreement between experimental and model-predicted values. The corresponding value of determination coefficient (R^2) was 0.9763 thereby implying approx. 98% of total variation in nitrilase response by the selected variables under investigation. With higher number of model terms and the relatively smaller sample size, the adjusted determination coefficient sometimes tends to achieve a value smaller than

Table 1
Factor values and responses in the various runs.

Run	*A* (min)	*B* (g/L)	*C* (W)	*D* (%)	Nit activity (U/mL)					Total protein release (g/L)					Cell disruption (%)				
					Actual	RSM predicted	MLP predicted	SMO predicted	MLR predicted	Actual	RSM predicted	MLP predicted	SMO predicted	MLR predicted	Actual	RSM predicted	MLP predicted	SMO predicted	MLR predicted
1	5	300	225	40	17.36	17.41	18.09	20.30	22.23	2.29	2.66	2.35	3.79	3.64	75.64	73.38	77.15	79.01	80.78
2	3	200	187.5	30	19.02	19.23	16.99	19.16	19.23	3.18	3.02	3.05	2.98	2.86	47.53	51.09	49.14	50.55	51.55
3	1	100	150	20	11.17	12.33	13.42	18.02	16.22	1.72	1.57	1.79	2.16	2.07	20.46	23.05	27.57	22.10	22.32
4	1	100	225	40	12.79	13.16	11.62	12.79	13.61	1.56	1.79	1.71	1.79	2.07	34.15	37.95	43.82	45.95	46.27
5	1	100	225	20	20.35	20.08	16.08	16.42	16.22	1.89	1.99	1.97	2.02	2.07	44.89	43.41	48.34	44.92	46.27
6	3	200	187.5	30	18.10	19.23	16.99	19.16	19.23	3.08	3.02	3.05	2.98	2.86	48.59	51.09	49.14	50.55	51.55
7	5	100	225	20	11.58	12.43	10.56	9.98	10.11	1.83	1.82	1.56	1.81	1.71	92.80	93.00	92.53	75.85	80.78
8	3	200	187.5	30	19.20	19.23	16.99	19.16	19.23	2.97	3.02	3.05	2.98	2.86	47.59	51.09	49.14	50.55	51.55
9	5	100	150	40	8.45	9.52	6.23	7.96	7.50	1.65	1.92	1.66	1.73	1.71	55.56	60.41	61.74	54.06	56.83
10	3	200	150	30	21.73	19.69	19.96	19.96	19.23	3.45	3.32	3.26	3.05	2.86	36.31	36.54	36.17	39.15	39.58
11	1	300	150	20	27.60	27.36	28.95	31.96	30.96	3.23	3.48	2.71	4.36	4.00	25.70	23.01	29.28	24.23	22.32
12	5	300	150	20	24.47	25.32	24.79	25.52	24.85	3.84	3.85	3.89	4.16	3.64	53.14	49.67	54.77	55.16	56.83
13	1	300	225	20	31.92	32.06	28.31	30.36	30.96	4.22	4.18	4.00	4.22	4.00	55.53	51.01	55.80	47.05	46.27
14	5	300	150	40	27.93	27.02	26.13	21.90	22.23	3.94	3.60	3.77	3.94	3.64	43.59	45.84	52.19	56.19	56.83
15	3	300	187.5	30	26.13	26.59	27.95	26.13	26.60	4.10	3.91	4.04	4.08	3.82	55.43	60.60	47.66	51.62	51.55
16	1	200	187.5	30	22.86	22.28	25.31	22.38	22.28	2.93	2.70	3.33	3.08	3.04	21.69	25.11	31.45	35.09	34.29
17	1	300	150	40	31.23	31.59	28.80	28.34	28.34	3.91	4.15	3.98	4.14	4.00	26.28	26.41	24.63	25.26	22.32
18	5	300	225	20	23.87	23.34	17.88	23.92	24.85	4.01	3.94	3.48	4.02	3.64	78.01	81.72	71.62	77.98	80.78
19	3	200	225	30	18.17	18.77	17.06	18.36	19.23	2.91	3.05	2.47	2.90	2.86	65.09	60.49	64.28	61.96	63.52
20	5	100	150	20	11.60	11.36	7.31	11.58	10.11	2.01	2.01	1.81	1.95	1.71	69.52	68.59	65.48	53.03	56.83
21	5	100	225	40	3.91	2.98	7.03	6.36	7.50	1.18	0.69	1.45	1.59	1.71	76.85	80.31	86.31	76.88	80.78
22	3	200	187.5	30	20.66	19.23	16.99	19.16	19.23	2.83	3.02	3.05	2.98	2.86	50.67	51.09	49.14	50.55	51.55
23	1	100	150	40	13.67	13.02	11.21	14.40	13.61	2.56	2.4	2.19	1.94	2.07	25.04	22.10	30.94	23.13	22.32
24	3	100	187.5	30	13.96	11.86	8.96	12.19	11.86	1.76	1.97	1.83	1.87	1.89	73.62	64.08	65.73	49.49	51.55
25	1	300	225	40	29.59	28.66	28.78	26.73	28.34	4.06	3.83	3.81	4.00	4.00	48.19	49.89	50.56	48.07	46.27
26	3	200	187.5	30	17.90	19.23	16.99	19.16	19.23	3.12	3.02	3.05	2.98	2.86	50.22	51.09	49.14	50.55	51.55
27	3	200	187.5	20	20.92	20.53	17.51	20.97	20.53	3.38	3.29	3.20	3.09	2.86	47.65	54.25	51.21	50.04	51.55
28	3	200	187.5	40	15.03	17.92	15.94	17.35	17.92	2.96	3.06	2.62	2.87	2.86	60.57	49.60	47.79	51.07	51.55
29	3	200	187.5	30	18.62	19.23	16.99	19.16	19.23	2.99	3.02	3.05	2.98	2.86	48.81	51.09	49.14	50.55	51.55
30	5	200	187.5	30	17.02	16.17	13.91	15.94	16.17	2.09	2.34	2.44	2.87	2.68	67.42	59.62	82.49	66.02	68.8

A, sonication time (min); *B*, biomass concentration (g/L); *C*, acoustic power (W); *D*, duty cycle (%).

R^2 [26–28]. The closer to unity value of adjusted determination coefficient (0.9817) and F-value of 78.27 vindicate high significance of the model. To put in other words, the probability of experimental noise (error) becomes close to 0.01%. The Fisher's *F*-test value is the marker for the data deviation around its mean value [29].

Three linear (*A, C* and *D*) and four interactive term (*AB, AD, BC, BD and CD*) were found to be significant for nitrilase release response [Supplementary Table 1]. Only one single term (*B*) and one interactive term (*AC*) were not significant. For the total protein release response (Eq. (2)), linear terms (*A, B* and *C*); interactive terms (*BC, BD* and *CD*) and one quadratic term (C^2) were significant [Supplementary Table 2]. For cell disruption response (Eq. (3)), two linear (*B* and *C*), one interactive (*AC*) and two quadratic (A^2, C^2) terms were significant [Supplementary Table 3]. Most of the selected operational variables were significantly affecting nitrilase and total protein response as observed from the *P*-values of the quadratic model terms. Moreover, it can be concluded that sonication time and acoustic power were largely affecting cell disruption [Supplementary Table 3]. Three-dimensional response surface graphs were used to study the interactive effects between independent variables on any given responses. Response surface plots were created by changing the response at two independent variables while keeping other two factors at their middle level [Fig. 1 & Supplementary Figs. 1 and 2].

3.2. Multilayer perceptron (MLP) modeling

MLP is popular machine learning technique owing to the availability of large number of training algorithms [30]. MLP is a modified version of the standard linear perceptron which utilizes a supervised learning technique of back-propagation for the training purpose and can distinguish data that are not linearly separable [31]. It is a feed-forward network consisting of a general framework of input layer, output layer and one or more hidden layers for non-linear activation function [32,33]. The number of neurons in each layer differ which has different mathematical functions. The neurons in each layer are interconnected through weights and these neurons accept input from preceding layer. Thus, it can be stated that the predictions performed by MLPs are defined by the activation functions of neurons [34]. Owing to the adaptive nature of the system, the weights between the neurons continuously adjust the overall structure of the network based on the information flowing through the network during the training phase [35].

In the present study, experimental data obtained by the RSM model was adequate to generate effective MLP model. As shown in Fig. 2, the training dataset of the MLP model exhibited very good *R*-values of 0.944, 0.963 0.944, and 0.986 for nitrilase activity, total protein release, fraction of cells disrupted and combined dataset, respectively. Thus, it was concluded that trained MLP models could significantly predict all the three responses.

3.3. Sequential minimal optimization (SMO) model

SMO, originally developed by Platt et al., is a potential class of decomposition technique in the machine learning domain [36,37]. It is advantageous in a way that the large quadratic programming problem generated in the SVMs can be divided into a series of smaller set of

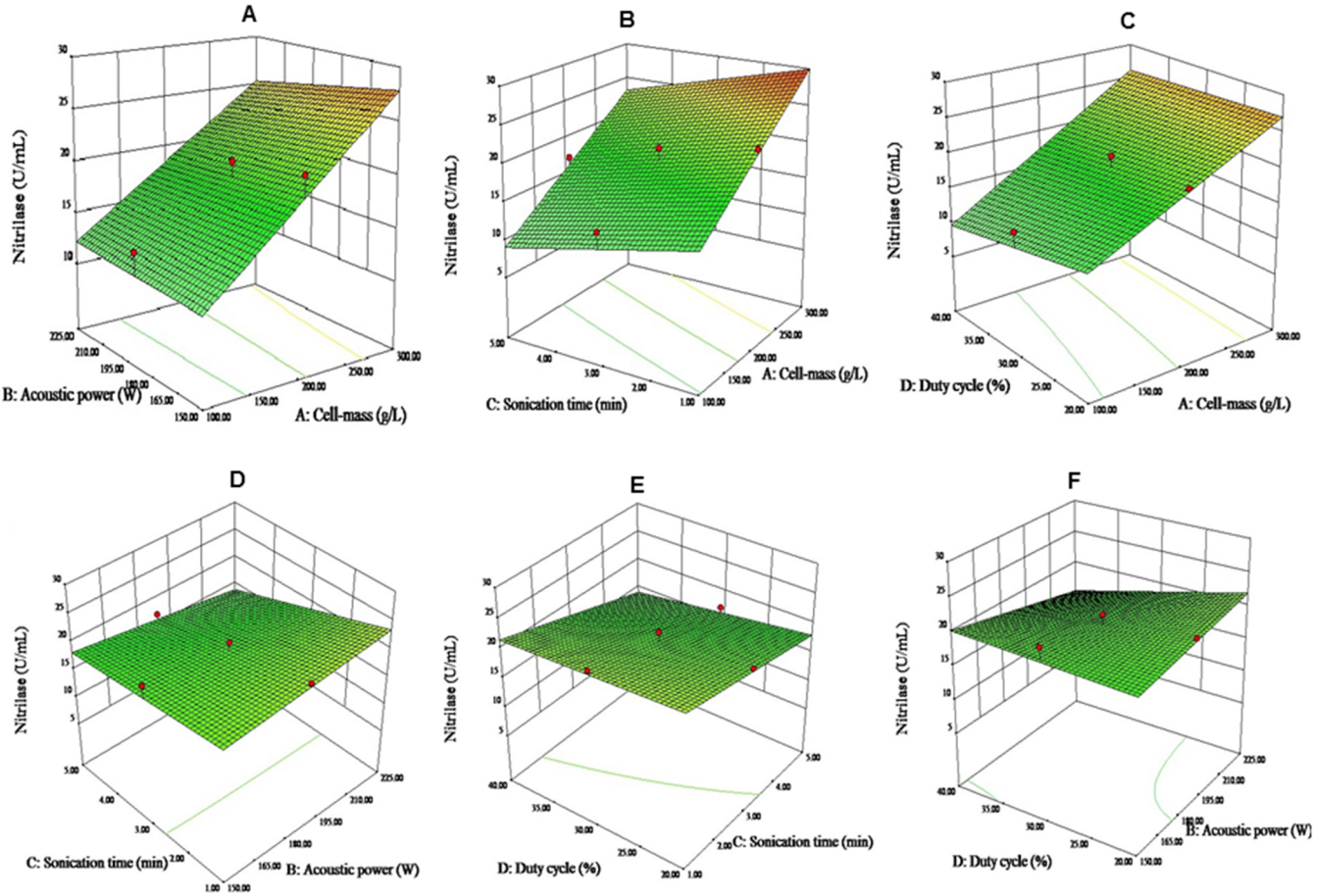

Fig. 1. Response surface plots showing the interactive effect of (A) acoustic power and cellmass concentration; (B) cellmass concentration and sonication time; (C) duty cycle and cellmass concentration; (D) acoustic power and sonication time; (E) sonication time and duty cycle; and (F) acoustic power and duty cycle on nitrilase release by ultrasonication-mediated cell disruption of *E. coli*.

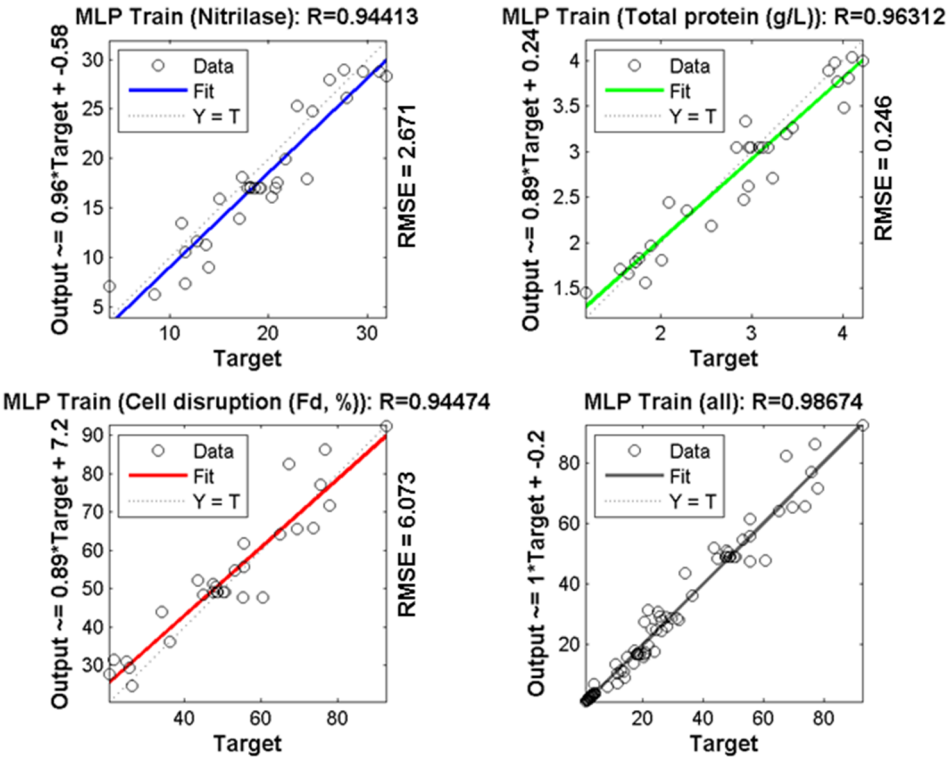

Fig. 2. Comparison between MLP-predicted and the measured responses of ultrasound-mediated cell disruption of *E. coli*.

problems using heuristics [38–40]. These smaller problems are analytically solved using two Lagrangian multipliers instead of optimization by numerical quadratic programming which is quite time-consuming [41]. Owing to the iterative selection of subsets of only of 2, among which the equality constrain can be used to eliminate one of the two Lagrange multipliers, SMO functions as a faster method than traditional SVM [42]. Thus, the actual functioning of SMO can be divided into two parts (1) A set of heuristics for the efficient selection of Lagrange multipliers to split the training dataset into smaller manageable datasets and (2) Solving a quadratic problem using selection of a working set of two [42,43]. Also, one of the added advantages of SMO algorithm is its time saving, because it does not involve the accumulation of errors caused by iterative process. Furthermore, SMO algorithms do not store the nuclear matrix to its interior memory while other numerical solutions use nuclear matrices in each and every step. Thus, while the use of inner memory by other algorithms increases by square, its use by SMO only linearly increases with the increase in the size of sample dataset [44]. Thus, the problem of inner memory space is avoided with the use of SMO algorithms. Selection of suitable kernel, which define the feature space for the classification of the training datasets, is an important aspect, because the resultant outputs vary with the selection of differential kernel [45].

In the present case, a polykernel was used for model development with cache size 250,007 and 1.0 exponent value. WEKA "Normalize" function used to normalize data before model development. Three separate individual models were developed using input parameters for the outputs of nitrilase activity, total protein released, and the fraction cells disrupted. RegSMO improved method used for reg optimizer with tolerance and epsilon parameter 0.001 and epsilon 1.0E-12. As shown in Fig. 3, the *R*-values for the nitrilase activity, total protein, cell disruption and the combined training datasets were 0.933, 0.887, 0.901 and 0.979 respectively.

3.4. Multiple linear regression (MLR) model

MLR is a fundamental method of modeling used to derive the linear relationship between a dependent variable (also called as predict or response) and the independent variables (called as predictors) [46].

While simple linear regression is used to find the find the straight line that fits the data, MLR models are used to find the *plane* best fitting the data [47]. In the analysis of MLR models, the method of least squares is used for the estimation of the regression coefficients. The regression coefficients exemplify the contribution of predictors towards the prediction of the responses. MLR is based on the assumptions such as the normal distribution of the residues and equal variance across the data space of independent variables [48]. Thus, the flexible and insightful analysis by MLR allows the introduction of simple nonlinearities and interaction effects to extend the utility of this method. The most important advantage of MLR is its ability to assess the unique contribution of each individual independent variable within a set of variables [49].

In the present case, several models have been developed using M5 method and Greedy method as attribute selection method parameter in WEKA. Final model selected having 1.0E−8 ridge value and attribute selection method as M5 method. As shown in Fig. 4, the linear regression model with the training dataset had very good *R*-values of 0.933, 0.887 and 0.901 for nitrilase activity, total protein release and fraction of cells disrupted, respectively. The *R*-value for the entire dataset (i.e. the three datasets combined) was also good (=0.978).

3.5. Comparison of predictive and generalization capacities of the models

A set of six unidentical runs were carried out to evaluate the performance machine learning models. The unseen dataset and experimental values of the responses were compared with that of predicted values as shown in Supplementary Table 4. For statistical analysis of the unseen dataset error parameters as given in Table 2 were used. Upper panel indicates training dataset while the lower panel indicates test dataset.

Among the various empirical models built by machine learning tools, MLR performed marginally better than SMO, followed by MLP. Nonetheless, the R^2 values for all the models built by machine learning tools were comparable (Table 2, Upper panel). Error analyses also

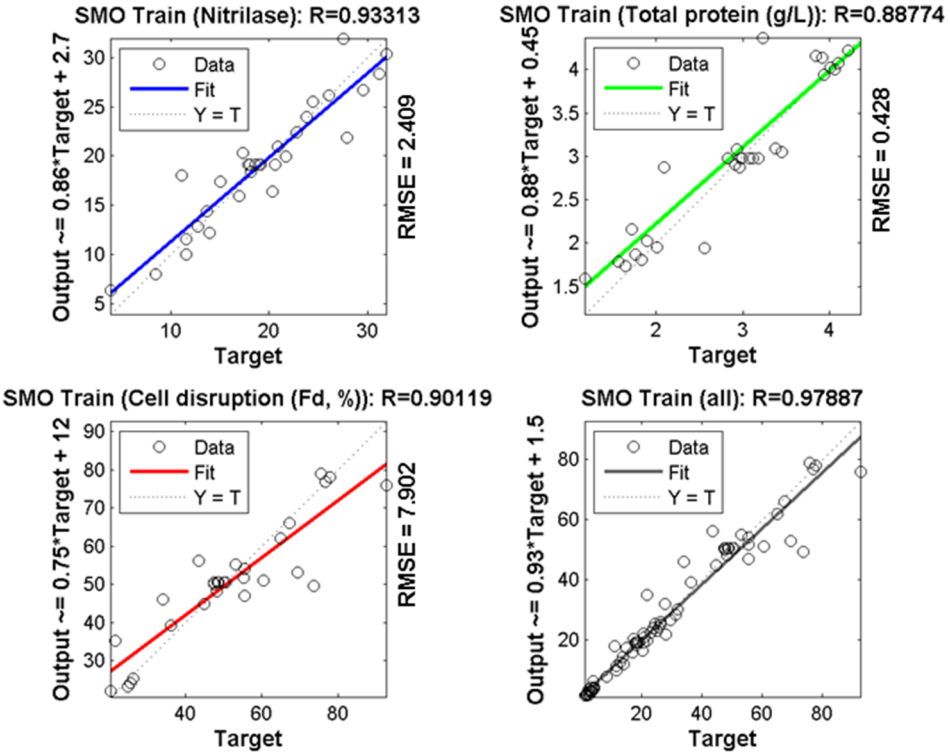

Fig. 3. Comparison between SMO-predicted and the measured responses of ultrasound-mediated cell disruption of *E. coli.*

revealed the comparable performances of the SMO and MLR. In case of cell disruption response, MLP was found to be most proficient due to its least error values amongst all. For the training dataset of total protein release as a response, values for R^2 and error parameter for machine learning model MLP was comparable to statistically designed RSM model.

Once the models are built through curve fit, their predictive capability is tested by comparing with additional "new" data sets that are not used for the training of the models. Because non-mechanistic, empirical models may be able to fit the same data used for their training, however, it is an important to show how these models would predict under 'new dataset'. Generally, new dataset having 15–20% of data-

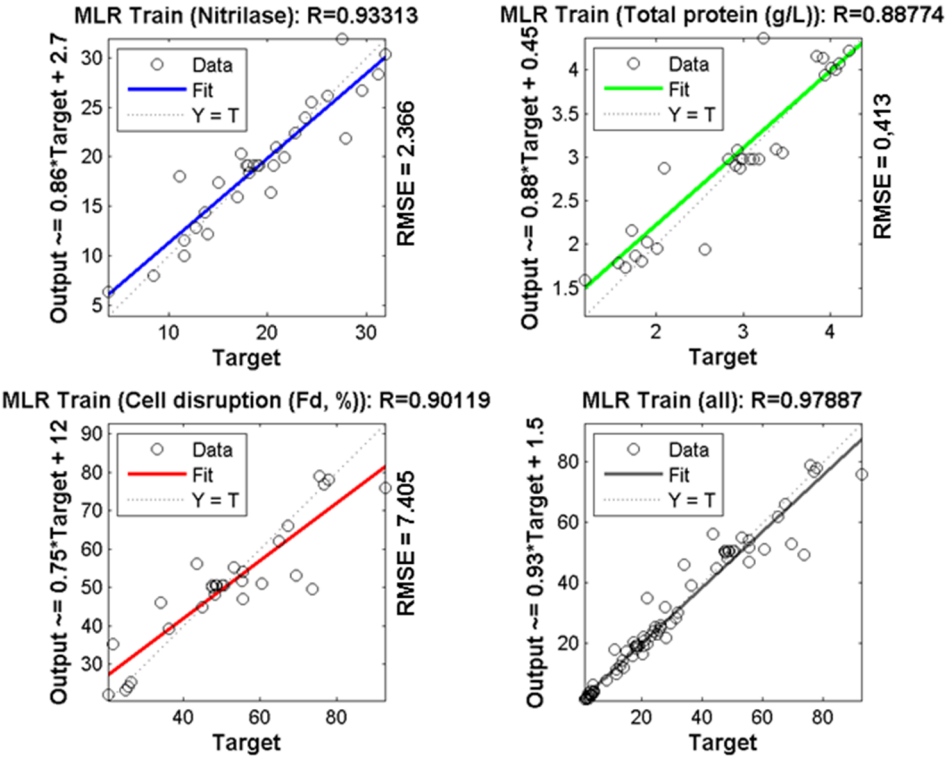

Fig. 4. Comparison between MLR-predicted and the measured responses of ultrasound-mediated cell disruption of *E. coli.*

Table 2
Validation and comparison with the unseen dataset.

Statistical parameter	Nitrilase activity (U/mL)				Total protein release (g/L)				Cell disruption (%)			
	RSM	MLP	SMO	MLR	RSM	MLP	SMO	MLR	RSM	MLP	SMO	MLR
Training dataset												
R^2	0.976	0.891	0.871	0.875	0.948	0.928	0.788	0.774	0.943	0.893	0.812	0.828
MSE	1.063	7.133	5.805	5.597	0.040	0.060	0.183	0.171	18.254	36.880	62.446	54.832
RMSE	1.031	2.671	2.409	2.366	0.199	0.246	0.428	0.413	4.272	6.073	7.902	7.405
SEP	5.362	13.891	12.531	12.305	6.962	8.600	14.998	14.472	8.288	11.780	15.329	14.364
RPD	5.444	15.073	10.702	12.389	7.044	7.478	10.808	12.800	7.073	10.612	10.507	11.938
Unseen Dataset												
R^2	0.880	0.802	0.005	0.749	0.734	0.840	0.077	0.749	0.846	0.796	0.650	0.605
MSE	0.892	1.546	91.996	4.878	0.056	0.030	1.046	0.076	20.554	37.226	84.582	122.753
RMSE	0.944	1.243	9.591	2.209	0.236	0.174	1.023	0.275	4.534	6.101	9.197	11.079
SEP	3.429	4.514	34.821	8.018	6.313	4.634	27.303	7.352	9.636	12.968	19.548	23.549
RPD	2.644	4.104	29.177	7.475	5.995	3.625	23.117	5.473	8.216	13.382	17.973	22.232

R^2, determination coefficient; MSE, mean square error; RMSE, root mean square error; SEP, standard error of prediction; RPD, relative percent deviation. *A*, Cell mass concentration; *B*, Operating pressure; *C*, number of passes.

points of the total dataset is used to validate the predictive capacities of the empirical models. In the present study, the generalization capabilities of the predictive models were evaluated based on the error analyses of the test dataset consisting of eight experiments which were not used in the training of the models (Table 2, Lower panel). Based on the values of determination coefficients and error parameters, the generalization capabilities of RSM and MLP models for the response of nitrilase activity were comparable; though RSM may be considered marginally better owing to its lower error values. For cell disruption response, the predictive capabilities of RSM for unseen dataset were superior to other models. Also, the generalization capability of MLP was superior among the other models developed by machine learning.

3.6. Kinetics of the three responses

In one of our previous cell disruption study with *Pseudomonas putida* microorganism cell disruption, total protein and enzyme release were found to follow first-order kinetics [5]. In the present study, similar profile was observed. The values of the rate constants are reported in Table 3.

4. Discussion

Ultrasonication has been reported as one of the most commonly employed method due to its low operation cost and simplicity of operation without much need of technical expertise [50–54]. In principle, ultrasonic devices can be scaled up and operated continuously. They are very much used in the chemical industry. Although many studies have successfully applied the ultrasonication-mediated cell disruption, its application is limited due to inadequate knowledge of the mechanism of cell disruption, product release mechanisms and unavailability of the predictive models [55].

The last few years have evidenced an emergence of several machine learning techniques like MLP, MLR and SMO as proficient alternatives to statistical models. Machine learning techniques have been applied widely for the modeling of complex relationships between input and output variables in non-linear responses [56]. Techniques, such as MLP offer an added advantage of prediction of future responses for any given set of the independent variables, notably without any explicit relationships between the inputs and the response [11,57]. Alternatively, supervised learning based-method such as SMO are widely used for the prediction of the dependent responses which involve small datasets [21,58]. The performance of machine learning techniques has oftenly been reported to outperform the statistically-designed models for the prediction of biological responses.

Till date, ultrasonication has been used for the recovery of intracellular enzymes from many microorganisms such as Gram-positive bacteria [59], Gram negative bacteria [5,52], yeast cell [60,61], microalgae [62,63]. Several theories have been proposed to explain the mechanism behind ultrasonication-mediated cell lysis. Cavitating conditions are considered essential for the cell wall disruption for the release of intracellular contents of the cell. During ultrasonication, acoustic cavitation lead to the rapid formation, growth and collapse of vapor filled microbubbles at the nucleation site to produce local shock waves in the vicinity (several magnitudes of atmospheric pressure) [64,65]. This leads to shear gradients imparting high kinetic energy to

Table 3
Rate constants for the nitrilase release, total protein release and cell disruption.

Run#	K_A	K_c	K_p	Run	K_A	K_c	K_p	Run	K_A	K_c	K_p
1	0.008	0.012	0.016	15	0.017	0.008	0.019	29	0.036	0.008	0.035
2	0.018	0.012	0.024	16	0.025	0.037	0.035	30	0.027	0.006	0.018
3	0.068	0.025	0.056	17	0.033	0.057	0.096				
4	0.076	0.066	0.043	18	0.011	0.010	0.022				
5	0.052	0.050	0.038	19	0.021	0.022	0.027				
6	0.019	0.022	0.028	20	0.026	0.004	0.021				
7	0.016	0.009	0.011	21	0.036	0.010	0.016				
8	0.021	0.016	0.028	22	0.017	0.015	0.031				
9	0.039	0.007	0.013	23	0.051	0.035	0.044				
10	0.023	0.010	0.028	24	0.036	0.017	0.035				
11	0.095	0.039	0.058	25	0.051	0.037	0.057				
12	0.016	0.006	0.014	26	0.039	0.011	0.018				
13	0.041	0.012	0.044	27	0.015	0.012	0.046				
14	0.017	0.007	0.013	28	0.022	0.010	0.032				

K_A, Rate constant for nitrilase release (s^{-1}); K_c, Rate constant for fraction of the cell disrupted (s^{-1}); K_p, Rate constant for total protein released (s^{-1}).

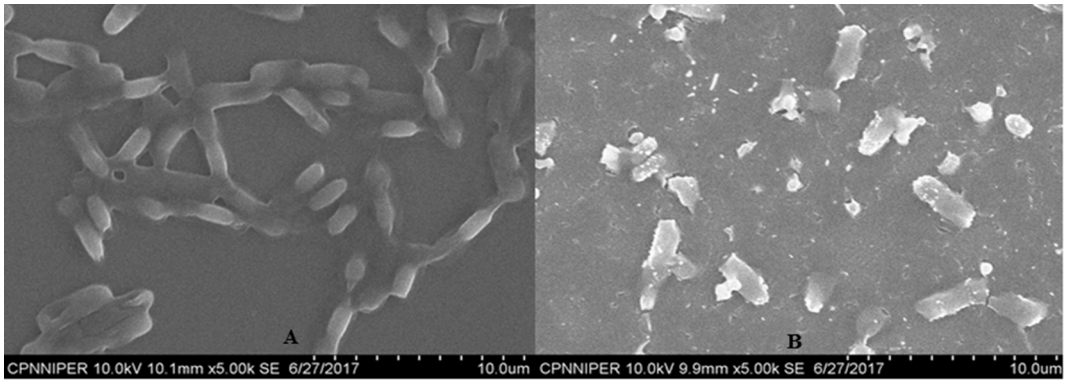

Fig. 5. Scanning electron microscopy images of *E. coli* cells A) before sonication and B) after sonication (5 min, 225 W, 100 g/L cellmass and 20% Duty cycle).

the cell surface rupturing the microbial cell wall [50].

As evidenced from Supplementary Tables 1 and 2, the cellmass concentration had a significant contribution on the release of nitrilase as well as total protein. The more concentrated slurry released more protein (protein concentration = 1.89 mg/mL in Run 5 versus 4.22 mg/mL in Run 13) and hence more target product (nitrilase) as well.

Singh et al. demonstrated that during sonication of *E. coli* K-12 cells, there was a decline in the released aspartase activity upon exposure to increased sonication time interval (after 3 min) and acoustic power (beyond 150 W) [7]. However, the authors also did not found any noticeable effect of cellmass concentration on the enzyme release of *E. coli* (varied from 50 to 250 g/L) a phenomenon similarly reported by Feliu et al. [52]. In literature, many scientists have failed to find any significant effect of cellmass on target/total protein release. [52,59,66,67]. Our present results are in disagreement on this stance.

Optimum conditions for the maximal cell disruption vary from organism to organism. *E. coli* is gram negative cell and much smaller in size as compared to yeast cell that in turn has thicker cell wall. Despite using sonication time of 40–60 min, the maximum cell disruption obtained was 52–65% only in case of yeast cell. In the present conditions, maximum cell disruption of 92.8% was attained in Run 7 where 100 g/L cells were used with 20% duty cycle for total sonication time of 5 min with 225 W acoustic energy (Table 1). Also, maintaining cold conditions during ultrasonication is not only advantageous for maintaining the integrity of liberated cell contents but also leads to higher cell disruption due to intense shock created on account of violent collapse of vapor filled bubbles at lower temperature [64].

Comparing Runs 4 and 21 (Table 1), carried out with identical concentrations of cellmass (100 g/L), the shorter duration run (Run 4, 1 min duration) released more active nitrilase compared to the longer duration run (Run 21, 5 min duration). Excessive cell disruption using ultrasonic cavitation has been reported to produce micro-sized cell debris that can pose serious threat during separation in subsequent downstream processing [9]. Intense level of sonication has been reported to affect a few labile biomolecules that lie around the tip of the sonic horn [68]. A similar detrimental effect on nitrilase release was observed in the present case (Negative β-coefficient value of acoustic power in Eq. (1)).

From the P-values, it is intriguing to observe that sonication time remained as the single factor that could influence all the three responses. For longer sonication period, there was a clear evidence of ultrasound-mediated denaturation of nitrilase at higher acoustic power in Runs 6 and 19. These runs were identical except for the intensity of the sonicator horn probe.

The fraction of the cells disrupted increased from 43.6 to 75.6% in runs 14 and 1, respectively when the acoustic power had increased from 150 to 225 W. Our results are in accordance with the findings of Feliu et al. [52] and Ho et al. [14] who showed a linear relationship between intracellular product released from *E. coli* cells and acoustic power used for sonication. This is because, increased acoustic power leads to the generation of more number of smaller eddies finally resulting in the surge in disruptive eddie currents responsible for enhanced cell lysis [8]. However, it was found that, there was an inverse relationship between enzyme activity and acoustic power. With acoustic power set at 150 W (Run 10; F_d = 36%), the nitrilase activity was nearly 21.7 U/mL, while at 225 W, the same was reduced to 18.1 (Run 19; F_d = 65%).

Due to enhanced cell disruption, one would intuitively, expect for the increased total protein release and hence increased enzyme activity. The corresponding total protein concentration in Run 14 and 1 were 3.94 and 2.29 mg/mL respectively, while the activity was 27.93 and 17.36 U/mL respectively.

This anomaly can be attributed to the cellwall porosity created during sonication sufficient enough for expelling the intracellular components outside the cellwall. Nitrilase is present in the soluble fraction of the *E. coli* cell cytoplasm. Hence in runs where there is little/no cell disruption, nitrilase activity is still observed. This is because of the mild vibrations generated at lower acoustic power sufficient for the extraction of contents present inside the cell wall [63,64] as also reported by Lateef et al. [69] and Vargas et al. [54] in case of other microorganisms.

Scientists have used direct cell counting for the measurement of extent of cell disruption (intactness of cell wall). However, the process is tedious and time consuming if the samples involved are too many [14]. Acoustic power had a much profound effect on the fraction of the cells disrupted (Supplementary Fig. 2A, D and F) as indicated by its significantly low *p*-value (Supplementary Table 3). For example, an incremental increase in the acoustic power also increased the extent of cells disrupted (Runs 10, 8 and 19, Table 1). However higher acoustic power has been reported to create turbulent conditions that can potentially damage the components in the biological fluid [8,68]. Higher cellwall breakage capacity of acoustic power was very well evident from the scanning electron microscope (SEM) images as shown in Fig. 5.

In contrast with the literature [70], increase in the duty cycle (from 20 to 40%) did not necessarily resulted in the increased cell disruption, protein release and nitrilase release as observed in Table 1. With the present sonicator, the duty cycle could be adjusted from 0 to 99%. When the cycle stage is on, the probe generates disruptive ultrasonic waves that may also lead to thermal changes in the vicinity of sonic probe while at the off stage it remains shut and hence the heat gets dissipated in the surrounding. Hence, in the runs where higher duty cycle and sonication time is involved, the nitrilase activity of the cell lysate isn't proportionately increasing with the increased level of cell disruption. During the process of ultrasonication, free radicals are generated out of the intense vibrations from ultrasonic waves [71]. These can have detrimental effect of the liberated cellular contents (sonochemical changes). Hence, at lower duty cycle (less than 50%)

lesser amount of free radicals will be generated which in fact is an added economical advantage [61,62].

In our previous studies with high pressure homogenizer, 22.3 U/mL was the maximum amount of nitrilase activity that could be obtained. In the present case, 32 U/mL was the highest nitrilase activity of the cell lysate. Our results are in agreement with the findings of B. Balasaundaram et al. who also demonstrated the superior performance of ultrasonication (US) over high pressure homogenizer for the release of cytoplasmic proteins [72].

During sonication, there is a differential release of intracellular contents with time [73]. The release of total proteins from lysed cells tends to follow first-order kinetics. In the present scenario, the reaction rate constants for the three responses were averaged as follows: $K_A = 0.032 \pm 0.020\,s^{-1}$; $K_c = 0.020 \pm 0.016\,s^{-1}$; and $K_p = 0.032 \pm 0.018\,s^{-1}$.

5. Conclusion

Ultrasound-mediated cell disruption is commonly used for the homogenization of many organisms for the release of their intracellular contents. In terms of nitrilase and protein release from *E. coli*, ultrasonic treatment proved to be better compared to homogenization in a French press for the disruption of the same microorganism [11]. In the present study, maximum amount of disruption was found to be 92% for 100 g/L cellmass concentration with experimental conditions as follows: acoustic power of 225 W; sonication time 5 min; duty cycle 20%.

Error analysis of the training and test dataset revealed MLP model to be equally competent to RSM for all the three responses. The MLP model also had good generalization capability as compared to other models tested. 32 U/mL was the highest amount of nitrilase activity of the cell lysate at protein concentration of 4.2 g/L. The experimental conditions for the same were: cellmass concentration 300 g/L; acoustic power 225 W; sonication time 1 min and a duty cycle of 20%. Notable influence of experimental factors on the three individual responses were in the following order. For nitrilase release: Cellmass concentration = Sonication time > Duty cycle. For total protein release: Cellmass concentration > Sonication time > Acoustic power. For extent of cell disruption: Acoustic power > Sonication time. However, higher acoustic power had a damaging effect on nitrilase activity. Cell disruption obeyed first order kinetics.

Acknowledgments

Kiran and Mahesh gratefully acknowledge the Department of Biotechnology (DBT), New Delhi, India for the award of Senior Research Fellowship. Authors are thankful to Mukesh Yadav, Shubham Chinchulkar and Shrikant Raskar for their technical support with the bioreactor runs.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of interest

None.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ultras.2019.06.006.

References

[1] N. Ferrer-Miralles, J. Domingo-Espin, J.L. Corchero, E. Vazquez, A. Villaverde, Microbial factories for recombinant pharmaceuticals, Microb. Cell Fact. 8 (2009) 1–8.
[2] T. Sauer, C.W. Robinson, B.R. Glick, Disruption of native and recombinant *Escherichia coli* in a high pressure homogenizer, Biotechnol. Bioeng. 33 (1989) 1330–1342.
[3] D. Foster, Optimizing recombinant product recovery through improvements in cell-disruption technologies, Curr. Opin. Biotechnol. 6 (1995) 523–526.
[4] C.R. Engler, C.W. Robinson, Effects of organism type and growth conditions on cell disruption by impingement, Biotechnol. Lett. 3 (1981) 83–88.
[5] M.D. Patil, M.J. Dev, S. Tangadpalliwar, G. Patel, P. Garg, Y. Chisti, U.C. Banerjee, Ultrasonic disruption of *Pseudomonas putida* for the release of arginine deiminase: kinetics and predictive models, Bioresour. Technol. 233 (2017) 74–83.
[6] M.D. Patil, A.S. Shinde, G. Patel, K.D. Bhilare, M.J. Dev, U.C. Banerjee, Combined effect of attrition and ultrasound on the disruption of *Pseudomonas putida* for the efficient release of arginine deiminase, Biotechnol. Prog. 34 (2018) 1185–1194.
[7] R.S. Singh, A comparative study on cell disruption methods for release of aspartase from *E. coli* K-12, Indian J. Exp. Biol. 51 (2013) 997–1003.
[8] M.S. Doulah, Mechanism of disintegration of biological cells in ultrasonic cavitation, Biotechnol. Bioeng. 19 (1977) 649–660.
[9] Y. Chisti, M. Moo-Young, Disruption of microbial cells for intracellular products, Enzyme Microb. Technol. 8 (1986) 194–204.
[10] S. Kaushik, U. Mohan, U.C. Banerjee, Exploring residues crucial for nitrilase function by site directed mutagenesis to gain better insight into sequence-function relationships, Int. J. Biochem. Mol. Biol. 3 (2012) 384.
[11] K.D. Bhilare, M.D. Patil, S. Tangadpalliwar, M.J. Dev, P. Garg, U.C. Banerjee, Machine learning modelling for the high-pressure homogenization-mediated disruption of recombinant *E. coli*, Process Biochem. 71 (2018) 182–190.
[12] K.D. Bhilare, M.D. Patil, M.J. Dev, G.A. Sharma, U.C. Banerjee, Production of thermostable nitrilase mutant from *Escherichia coli* and bioreactor studies, J. Adv. Biotechnol. Microbiol. 7 (2017), https://doi.org/10.19080/AIBM.2017.07.555713.
[13] L. Benov, J. Al-Ibraheem, Disrupting *Escherichia coli*: a comparison of methods, J. Biochem. Mol. Biol. 35 (2002) 428–431.
[14] C.W. Ho, T.K. Chew, T.C. Ling, S. Kamaruddin, W.S. Tan, B.T. Tey, Efficient mechanical cell disruption of *Escherichia coli* by an ultrasonicator and recovery of intracellular hepatitis B core antigen, Process Biochem. 41 (2006) 1829–1834.
[15] O. Ustun-Aytekin, S. Arisoy, A.O. Aytekin, E. Yildiz, Statistical optimization of cell disruption techniques for releasing intracellular X-prolyl dipeptidyl aminopeptidase from *Lactococcus lactis* spp. Lactis, Ultrason. Sonochem. 29 (2016) 163–171.
[16] N. Marchitan, C. Cojocaru, A. Mereuta, G. Duca, I. Cretescu, M. Gonta, Modeling and optimization of tartaric acid reactive extraction from aqueous solutions: a comparison between response surface methodology and artificial neural network, Sep. Purif. Technol. 75 (2010) 273–285.
[17] J.P. Maran, B. Priya, Comparison of response surface methodology and artificial neural network approach towards efficient ultrasound-assisted biodiesel production from muskmelon oil, Ultrason. Sonochem. 23 (2015) 192–200.
[18] L. Breiman, J.H. Friedman, Predicting multivariate responses in multiple linear regression, J. R. Stat. Soc. Series B 59 (1997) 3–54.
[19] S. Ye, Linear and Logistic Regression Analysis, https://rstudio-pubs-static.s3.amazonaws.com/116105_7c7d6099678e4419bf0d4209be33f3ad.html.
[20] P. Dutta, M.K. Naskar, O.P. Mishra, South Asia earthquake catalog magnitude data regression analysis, Int. J. Stat. Anal. 2 (2011) 161–170.
[21] U. Caydas, S. Ekici, Support vector machines models for surface roughness prediction in CNC turning of AISI 304 austenitic stainless steel, J. Intell. Manuf. 23 (2012) 639–650.
[22] A. Sarve, S.S. Sonawane, M.N. Varma, Ultrasound assisted biodiesel production from sesame (*Sesamum indicum* L.) oil using barium hydroxide as a heterogeneous catalyst: comparative assessment of prediction abilities between response surface methodology (RSM) and artificial neural network (ANN), Ultrason. Sonochem. 26 (2015) 218–228.
[23] M.D. Patil, G. Patel, B. Surywanshi, N. Shaikh, P. Garg, Y. Chisti, U.C. Banerjee, Disruption of *Pseudomonas putida* by high pressure homogenization: a comparison of the predictive capacity of three process models for the efficient release of arginine deiminase, AMB Express 6 (2016) 84.
[24] M.M. Bradford, A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding, Anal. Biochem. 72 (1976) 248–254.
[25] Y.J. Tam, Z.N. Allaudin, M.A.M. Lila, A.R. Bahaman, J.S. Tan, M.A. Rezaei, Enhanced cell disruption strategy in the release of recombinant hepatitis B surface antigen from *Pichia pastoris* using response surface methodology, BMC Biotechnol. 12 (2012) 1–15.
[26] M.D. Patil, M.J. Dev, A.S. Shinde, G. Patel, K.D. Bhilare, U.C. Banerjee, Surfactant-mediated permeabilization of *Pseudomonas putida* KT2440 and use of the immobilized permeabilized cells in biotransformation, Process Biochem. 63 (2017) 113–121.
[27] G. Patel, M.D. Patil, S. Soni, Y. Chisti, U.C. Banerjee, Production of mycophenolic acid by *Penicillium brevicompactum* using solid state fermentation, Appl. Biochem. Biotechnol. 182 (2017) 97–109.
[28] G. Patel, M.D. Patil, S. Soni, T.P. Khobragade, Y. Chisti, U.C. Banerjee, Production of mycophenolic acid by *Penicillium brevicompactum*—a comparison of two methods of optimization, Biotechnol. Rep. 11 (2016) 77–85.
[29] M.D. Patil, K.D. Shinde, G. Patel, Y. Chisti, U.C. Banerjee, Use of response surface method for maximizing the production of arginine deiminase by *Pseudomonas putida*, Biotechnol. Rep. 10 (2016) 29–37.
[30] T. Philippe, M.-C. Suhner, A new multilayer perceptron pruning algorithm for classification and regression applications, Neural Process Lett. 42 (2015) 437–458.
[31] K. Patel, J. Vala, J. Pandya, Comparison of various classification algorithms on iris

datasets using WEKA, Int. J. Adv. Eng. Res. Dev. 1 (2014) 1–7.
[32] Y. Kumar, G. Sahoo, Analysis of Bayes, neural network and tree classifier of classification technique in data mining using WEKA, Comput. Sci. Inf. Technol. (2012) 359–369, https://doi.org/10.5121/csit.2012.2236.
[33] G. Sahoo, Y. Kumar, Analysis of parametric & non parametric classifiers for classification technique using WEKA, Int. J. Inf. Technol. Comp. Sci. 4 (2012) 43–49.
[34] Z.J. Cezary, Fuzzy decision trees: issues and methods, IEEE T. Syst. Man. Cyb. B 28 (1998) 1–14.
[35] N. Mallios, E. Papageorgiou, M. Samarinas, Comparison of machine learning techniques using the WEKA environment for prostate cancer therapy plan, IEEE 20th International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, IEEE, 2011, pp. 151–155.
[36] J. Platt, Sequential minimal optimization: A fast algorithm for training support vector machines, Technical Report MSR-TR-98-14; https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/tr-98-14.pdf.
[37] J.C. Platt, Using analytic QP and sparseness to speed training of support vector machines, Adv. Neural Inf. Process. Syst. (1999) 557–563.
[38] V. Vapnik, Methods of function estimation, The Nature of Statistical Learning Theory, Springer Science & Business Media, 2013, pp. 181–216.
[39] M.A. Heasrt, Trends and controversies-support vector machines, IEEE Intell. Sys. 13 (1998) 18–28.
[40] K.S. Sathiya, E.G. Gilbert, Convergence of a generalized SMO algorithm for SVM classifier design, Mach. Learn. 46 (2002) 351–360.
[41] G.W. Flake, S. Lawrence, Efficient SVM regression training with SMO, Mach. Learn. 46 (2002) 271–290.
[42] Y. Jing, X. Yang, J. Zhang, A parallel multi-class classification support vector machine based on sequential minimal optimization, in: First International Multi-Symposiums on Computer and Computational Sciences, vol. 1, 2006, pp. 443–446; http://doi.org/10.1109/IMSCCS.2006.20.
[43] C. Nello, S.-T. John, An Introduction to Support Vector Machines and Other Kernel-based Learning Methods, Cambridge University Press, Cambridge, UK, 2000.
[44] Q. Zhou, Z. Yong-Jie, Pu Han, Sequential minimal optimization algorithm applied in short-term load forecasting, in: International Conference on Machine Learning and Cybernetics, vol. 5, 2007, pp. 2479–2483. http://doi.org/10.1109/ICMLC.2007.4370563.
[45] B. Thai Pham, A comparative study of sequential minimal optimization-based support vector machines, vote feature intervals, and logistic regression in landslide susceptibility assessment using GIS, Environ. Earth sci. 76 (2017) 371.
[46] J. Mata, Interpretation of concrete dam behaviour with artificial neural network and multiple linear regression models, Eng. Str. 33 (2011) 903–910.
[47] S.H. Brown, Multiple linear regression analysis: a matrix approach with MATLAB, Ala. J. Math. 34 (2009) 1–3.
[48] B.K. Slinker, S.A. Glantz, Multiple linear regression: accounting for multiple simultaneous determinants of a continuous dependent variable, Circulation 117 (2008) 1732–1737.
[49] L.S. Aiken, S.G. West, S.C. Pitts, A.N. Baraldi, I.C. Wurpts, Multiple linear regression, Handbook of Psychology, second ed., John Wiley & Sons, New York, 2012, pp. 511–542.
[50] H. Kapucu, N. Gulsoy, U. Mehmetoglu, Disruption and protein release kinetics by ultrasonication of *Acetobacter peroxydans* cells, Biochem. Eng. J. 5 (2000) 57–62.
[51] M.A. Ganaie, U.S. Gupta, Recycling of cell culture and efficient release of intracellular fructosyltransferase by ultrasonication for the production of fructooligosaccharides, Carbohydr. Polym. 110 (2014) 253–258.
[52] J.X. Feliu, R. Cubarsi, A. Villaverde, Optimized release of recombinant proteins by ultrasonication of *E. coli* cells, Biotechnol. Bioeng. 58 (1998) 536–540.
[53] M. Klimek-Ochab, M. Brzezinska-Rodak, E. Zymanczyk-Duda, B. Lejczak, P. Kafarski, Comparative study of fungal cell disruption-scope and limitations of the methods, Folia Microbiol. 56 (2011) 469–475.
[54] L.H.M. Vargas, A.C.S. Piao, R.N. Domingos, E.C. Carmona, Ultrasonication effects on invertase from *A. niger*, World J. Microbiol. Biotechnol. 20 (2004) 137–142.
[55] S. Gao, G.D. Lewis, M. Ashokkumar, Y. Hemar, Inactivation of microorganisms by low-frequency high-power ultrasound: 1. Effect of growth phase and capsule properties of the bacteria, Ultrason. Sonochem. 21 (2014) 446–453.
[56] X. Li, J. Sha, Z. Wang, A comparative study of multiple linear regression, artificial neural network and support vector machine for the prediction of dissolved oxygen, J. Hydrol. Res. 48 (2017) 1214–1225.
[57] A. Witek-Krowiak, K. Chojnacka, D. Podstawczyk, A. Dawiec, K. Pokomeda, Application of response surface methodology and artificial neural network methods in modelling and optimization of biosorption process, Bioresour. Technol. 160 (2014) 150–160.
[58] A.J. Smola, B. Scholkopf, A tutorial on support vector regression, Stat. Comput. 14 (2004) 199–222.
[59] J.C. Caldeira, J.M. Cabral, Extraction of a steroid 1, 2 dehydrogenase by sonication of Arthrobacter simplex cells, Bioseparation 4 (1994) 271–278.
[60] P.B. Agrawal, A.B. Pandit, Isolation of α-glucosidase from *Saccharomyces cerevisiae*: cell disruption and adsorption, Biochem. Eng. J. 15 (2003) 37–45.
[61] D. Liu, X.-A. Zeng, D.-W. Sun, Z. Han, Disruption and protein release by ultrasonication of yeast cells, Innov. Food Sci. Emerg. Technol. 18 (2013) 132–137.
[62] J.A. Gerde, M. M-Lomboy, L. Yao, D. Grewell, T. Wang, Evaluation of microalgae cell disruption by ultrasonic treatment, Bioresour. Technol. 125 (2012) 175–181.
[63] R. Halim, R. Harun, M. Danquah, P. Webley, Microalgal cell disruption for biofuel development, Appl. Energy 91 (2012) 116–121.
[64] S.S. Save, A.B. Pandit, J.B. Joshi, Microbial cell disruption: role of cavitation, Chem. Eng. Biochem. Eng. J. 55 (1994) 67–72.
[65] I.Z. Shirgaonkar, R.R. Lothe, A.B. Pandit, Comments on the mechanism of microbial cell disruption in high-pressure and high-speed devices, Biotechnol. Prog. 14 (1998) 657–660.
[66] J.X. Feliu, A. Villaverde, An optimized ultrasonication protocol for bacterial cell distruption and recovery of b-galactosidase fusion proteins, Biotechnol. Techniques 8 (1994) 509–514.
[67] R. Kuboi, H. Umakoshi, N. Takagi, I. Komasawa, Optimal disruption methods for the selective recovery of b-galactosidase from *Escherichia coli*, J. Ferment. Bioeng. 79 (1995) 335–341.
[68] D. Liu, X.-A. Zeng, D.-W. Sun, Z. Han, Disruption and protein release by ultrasonication of yeast cells, Innov. Food Sci. Emerg. Technol. 18 (2013) 132–137.
[69] A. Lateef, J.K. Oloke, S.G. Prapulla, The effect of ultrasonication on the release of fructosyltransferase from *Aureobasidium pullulans* CFR 77, Enzyme Microb. Technol. 40 (2007) 1067–1070.
[70] X. Ren, D. Yu, L. Yu, G. Gao, S. Han, Y. Feng, A new study of cell disruption to release recombinant thermostable enzyme from *Escherichia coli* by thermolysis, J. Biotechnol. 129 (2007) 668–673.
[71] T.J. Mason, J.P. Lorimer, D.M. Bates, Y. Zhao, Dosimetry in sonochemistry: the use of aqueous terephthalate ion as a fluorescence monitor, Ultrason. Sonochem. 1 (1994) 91–95.
[72] B. Balasundaram, A.B. Pandit, Significance of location of enzymes on their release during microbial cell disruption, Biotechnol. Bioeng. 75 (2001) 607–614.
[73] A.J. Robertson, J.H. Gerlach, G.H. Rank, L.C. Fowke, Yeast cell wall, membrane, and soluble marker polypeptides identified by comparative two-dimensional electrophoresis, Can. J. Biochem. 58 (1980) 565–572.