Algal Research 69 (2023) 102908

Contents lists available at ScienceDirect

Algal Research

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

Chemometric analysis unravelling the effect of key influencing factors on algal biochar yield

Aastha Kapoor, Nageshwari Krishnamoorthy, Abhijeet Pathy, Paramasivan Balasubramanian *

Department of Biotechnology & Medical Engineering, National Institute of Technology Rourkela, Odisha 769008, India

ARTICLE INFO

Keywords:
Chemometric analysis
Algal biochar yield
MDS
PCA
Pyrolysis
Correlations

ABSTRACT

This study attempted to comprehend the key influencing factors amongst 16 variables on algal biochar production by pyrolysis from 115 datasets using chemometric tools such as Principal Component Analysis (PCA) and Multidimensional scaling (MDS). PCA results revealed that the first 5 principal components (PCs) accounted for 81.75 % of available dataset variability. Results of both PCA and MDS marked out the pyrolysis parameters, C (%), H (%), O (%), N (%), H/C, N/C, O/C as the main determining factors of algal biochar yield. The results also tied the occurrence of high algal char yields to low temperature-less residence time-moderate heating rate regimen. The correlations validated with the supportive findings of MDS, further revealed that algal biochar production is a complex process, which is dependent on the interplay of multiple variables involved rather than any singular major factor. The results thus offered an insight into the condensed criteria that presents an important cue for research and industry to achieve high algal biochar yields.

1. Introduction

Algal biochar is a porous carbonaceous solid product of thermal treatment of algal biomass at 300–800 °C ranges in an oxygen- deficient or limited environment [1,2]. Conventional pyrolysis process has been the widely used production technique for algal biochar production. The advantages of pyrolysis include better control of heating rate, temperature and higher yields obtained as compared to newer techniques like gasification, hydrothermal carbonization. Though pyrolysis has been used to produce char but drawbacks like low yield, reduced physiochemical properties with elevated temperatures have hampered the efforts to project biochar as a fitting replacement for coal. This necessitates the need to either improve the pyrolysis production process efficiency by targeting the critically decisive factors that influence algal biochar yield and properties or developing newer, more feasible production technologies. Recently greener technologies like hydrothermal carbonization (HTC), torrefaction etc. and assisted-improved pyrolytic techniques like microwave assisted pyrolysis, catalytic pyrolysis have been shown to solve some of the trouble areas of biochar production like costs, calorific value, property enhancement etc. [2–4]. Bestowed with unique properties like high heating value (HHV), high surface area, high zeta potential, high O/C, H/C, N/C ratios, high pH and high Cation exchange capacity (CEC) that sets it apart from lignocellulosic biochar, algal biochar finds applications in areas of coal replacements, wastewater treatment, bioremediation, soil amelioration, capacitors etc. [3,5,6].

Broadly it is observed that the parent feedstock composition and pyrolysis process parameters dictate the structure, physiochemical properties and yield of algal biochar being produced [3,7]. High CEC, high HHV, alkaline pH and abundance of trace metals and nitrogen set algal biochar apart from the lignocellulosic biochar and makes the former a better candidate for applications such as bioremediation, soil ameliorant, supercapacitor, catalysts etc. The advantages of using algal biomass for biochar production by pyrolysis over the years have been marred with the low product yield of algal biochar that takes away from its potential as a green alternative to charcoal-based products. While manual tinkering of thermal parameters and genetic engineering efforts on algal biomass can help achieve high yield and enhanced properties of biochar the results are limited [8]. The results of the same efforts have not yet been adequate to provide information about the key factors influencing algal biochar yield as well as the effect of various thermochemical strategies on the properties and application potential of algal biochar [2,3]. While extensive research insights are present that mention type of feedstock and experimental parameters as the governing set of variables for algal biochar production [9,10], research that underlines and arranges these variables based on their importance

* Corresponding author.
E-mail address: biobala@nitrkl.ac.in (P. Balasubramanian).

https://doi.org/10.1016/j.algal.2022.102908
Received 15 February 2022; Received in revised form 1 November 2022; Accepted 16 November 2022
Available online 5 December 2022

seems to be scarce. This makes it important to devise a way to streamline the major variables affecting the biochar yield into significant determinants based on available experimental datasets. Statistical analysis of algal biochar production process which includes the feedstock elemental composition and thermal regimen variables can thus be analyzed for highlighting the key factors affecting yield. Not only such important factors, but a multivariate statistical analysis can also shed light on the intricate correlations and interdependencies that these variables share amongst themselves [11]. Chemometric techniques like Principal component analysis (PCA) and Multidimensional scaling (MDS) can therefore provide the primary insight into the prime determinants of algal biochar yield as well as visual representation of the variations and correlations within the datasets under study. PCA is used to reduce the dimensions while conserving the variability of dataset and deriving correlations between input variables and MDS is used to visualize and map the similarity between individual cases of dataset. Thus, these clarified correlations and prime influencing factors affecting the algal biochar yield can be focused on specifically to achieve higher yield and enhanced properties results by researchers and industry alike.

To deduce relations between variables and samples of any dataset multivariate or chemometric analysis tools are used. These not only help find a pattern in the dataset under study but also help construct new latent variables to better represent the dataset. As a descriptive and predictive analysis, chemometrics has already been used for biomass classification and characterization as potential feedstock for biodiesel production [12], quantifying changes in plant xylem after exposure to elevated CO_2 [13], differentiating between wood pellet residue based on near infrared spectroscopy [14], catalyst based total oxidation of propane (Pattiya et al., 2010), predicting biomass char yield from high heating rate devolatilization (Espensen et al., 2018) and also for researching the composition of biomass products of gasification [15,16]. To the best of author's knowledge, no previous studies has been reported so far to derive the correlations and criteria that can be applied to achieve high yields of algal biochar in laboratories and scaled up to meet industry demands. In this manuscript, principal component analysis (PCA) as a chemometric tool was used to determine the effect of various algal feedstock variables like elemental composition and ratios (C, H, N, O, O/C, H/C, N/C), proximate analysis criteria (volatile fraction, ash, fixed carbon content etc.) and pyrolysis process parameters (residence time, heating rate, temperature etc.) on the yield of algal char. As a chemometric analysis tool, PCA is used for feature extraction, thereby deducing a pattern from the available datasets [11,17]. The main objective of PCA is to conserve the variability of whole dataset while reducing its dimensionality and redundancy. In this study, PCA has been conducted using XL-STAT software to determine the most influential parameters of algal biochar yield and collaterally derive the correlations between these parameters so found. The results obtained from PCA have been supplemented with MDS as it has been used as a visualisation technique for representing the similarity between the individual algal biochar datasets. Besides considering the representation of composite algal biochar dataset, MDS has also been used to represent the correlations between variables based on proximity matrix and also distinguish the results for microalgae and macroalgae derived biochar. It is expected that these chemometrically derived correlations between feedstock composition variables and thermal regimen parameters can direct if not guarantee the achievement of biochar from algal biomass in high yield without the trade-off for its unique physiochemical properties.

2. Methodology

2.1. Data collection and pre-processing

To conduct chemometric analysis, 115 datasets for algal biochar production by pyrolysis method were curated from various research papers and experimental findings. The data was collected from diagrams, graphs, supplementary tables of available literature sources. While collecting datasets only those sources were considered which applied conventional pyrolysis as the method of production. The missing figures were imputed with standardized mode values of variables to achieve complete dataset for PCA [18,19]. A total of 16 quantitative and qualitative variables{C (%), H (%), O (%), N (%), ash (%), H/C, O/C, N/C, fixed carbon (%), volatile compound (%), carbohydrate (%), protein (%), lipid (%), residence time (min), heating rate (°C min^{-1}), temperature (°C)}and 2 supplementary variables (microalgae and macroalgae) were considered for this study (supplementary variables are used for illustrating and categorizing the results and deducing correlations). Resultant yield (%) of each of the data entry was itself categorized to low (0–17 %), medium (18–34 %), high (35–49 %) and very high (>49 %) categories respectively. The statistical analysis of the dataset included counts, mean, standard deviation, quartile values and maximum and minimum values of all parameters. For an outcome of n objects (categories of algal biochar yield) decided by p variables (like elemental composition, proximate analysis, heating regimen parameters etc.) firstly the initial raw dataset was standardized (to be able to compare all input variables having different units on a common scale) using mean and standard deviation (Eq. (1)) [20] to compare different variables on an identical scale,

$$z = \frac{value - mean}{standard\ deviation} \tag{1}$$

(where, z is the standardized value of data point under a given variable p)

2.2. PCA run on pre-processed datasets

From the standardized values of variables, a correlation matrix (Eq. (2)) was obtained, which was a $p * p$ symmetric matrix that had correlation values of all possible variable pairs.

$$R = \begin{bmatrix} 1 & r_{12} & r_{13} \cdots r_{1p} \\ r_{21} & 1 & r_{23} \ldots r_{2p} \\ r_{31} & r_{32} & 1 \cdots r_{3p} \\ \vdots & \vdots & \vdots \\ r_{p1} & r_{p2} & r_{p3} \ldots 1 \end{bmatrix} \tag{2}$$

where R the Pearson correlation coefficient between variable x_j and x_k as described in Eq. (3) [21].

$$r_{jk} = \frac{S_{jk}}{S_j S_k} = \frac{\Sigma_{i=1}^{n}\left(x_{ij} - \underline{x_j}\right)\left(x_{ik} - \underline{x_k}\right)}{\sqrt{\Sigma_{i=1}^{n}\left(x_{ij} - \underline{x_j}\right)}^2 \sqrt{\Sigma_{i=1}^{n}\left(x_{ik} - \underline{x_k}\right)^2}} \tag{3}$$

where S is the covariance between j^{th} and k^{th} variables out of total p variables.

This correlation matrix was used to assign magnitude (Eigenvalues) and direction (Eigenvectors) to each of the parent variable in dataset. Each variable, hence having its own value and effect on new dimensions being so created contributed to the output (algal biochar yield). Top k eigenvector values correspond to largest of k eigenvalues and these were arranged in descending order and chosen. These resultant new dimensions known as "Principal Components (PCs)" thus were a singular and linear or mixed combination of the p parent variables that consist all the compressed, compact information and variability of whole dataset within themselves. Upon application to our dataset besides presenting the PCs, PCA also highlighted the most decisive of correlations and interdependencies between the variables. A total of 16 quantitative and qualitative variables comprising of pyrolytic parameters and elemental composition (as mentioned in Section 2.1) were analyzed for yield (%). The algal biochar yield (%) for all the observations in dataset was further categorized to four ranges: low, medium, high and very high. Moreover, two supplementary variables of microalgae and macroalgae were used to better illustrate the dataset profiles based on the type of feedstock available. These supplementary variables were assigned the values of 0 and 1 so as to make them identifiable for data analysis. Correlation type PCA was conducted with standardized dataset using 5 filter factors on XLSTAT 2021.3.1.1149 software. The filter factors employed by the software focused on increasing the readability of observation charts by

using $\cos^2$ values of data entries as a screening criterion for dropping less useful observations. Correlation PCA was chosen over covariance as the article is aimed at studying the key factors influencing the algal biochar yield and how the 16 (quantitative and qualitative) variables are related amongst themselves rather than the difference between the variables. Since the variables under study had different dimensions and variant levels correlation type PCA was done to ascertain the extent to which the variables under study might be related to one another Results were further elucidated using OriginPro® software toolkit as it provided better representation of single variable datasets. MDS was carried out to further illustrate the results obtained by PCA. Based on distances and dissimilarities between different variables, a proximity matrix was obtained and thereby MDS was done using XLSTAT® 2021.3.1.1149 software. The degree of correspondence between the distances amongst points implied by MDS map and the matrix input has been measured (inversely) by a *stress* function (t) as mentioned in Eq. (4) [22,23].

$$t = \sqrt{\frac{\Sigma\Sigma\left(f\left(x_{ij}\right) - d_{ij}\right)^2}{scale}} \tag{4}$$

In the Eq. (4), d_{ij} refers to the Euclidean distance, across all dimensions, between points i and j on the map, $f(x_{ij})$ is some function of the input data, and *scale* refers to a constant scaling factor, used to keep stress values between 0 and 1. When the MDS map perfectly reproduces the input data, $f(x_{ij})$ - d_{ij} is for all i and j, so stress is zero. Thus, the smaller the stress, the better the representation, which has been described by Kruskal stress (Eq. (5)) [22].

$$w = \sqrt{\frac{\Sigma\Sigma\left(f\left(x_{ij}\right) - d_{ij}\right)^2}{\Sigma\Sigma d_{ij}^2}} \tag{5}$$

where w is the Kruskal stress for given dataset.

For better illustrating the correlations between 16 input variables and for comparing the dissimilarity between microalgae derived biochar and macroalgae derived biochar, MDS was run for composite (combined dataset that consists of both microalgae and macroalgae derived biochar) and segregated algal char datasets. A metric proximity matrix was obtained for datasets to clearly ascertain the distance and thereby dissimilarity in given observation points.

2.3. Data interpretation

Upon the pre-processing and standardization of data, correlation type PCA was conducted using 5 filter factors to edge out noisy data. PCA analysis results were evaluated and interpreted based on the standard statistical conventions.

2.3.1. Pearson correlation matrix

As a first step in PCA, a Pearson correlation matrix was devised that exhibits all the relations between all of the input variables in the ranges of −1 and +1. Based on these Pearson values (ranging from +1 to −1) it was ascertained whether the variables in question had a positive correlation or a negative correlation respectively [24]. The quantum of the coefficient decides how loosely or tightly, the two variables being studied are associated.

2.3.2. Selection of principal components

To begin with the interpretation of PCA results, firstly, the resultant PCs were evaluated according to Kaiser's rule [11] that states that only PCs with eigenvalues above 1 must be considered from PCA analysis. This gave significant insight upon the first 5 PCs that hold highest eigenvalues and consequently retain the most information from parent variables. In general, the PCs having variability score of ≥7 were taken into consideration as it is assumed that these PCs sufficiently reflect the majority of variance of the whole dataset [17].

2.3.3. Scree plot

The scree plot has been traditionally used to represent the eigenvalues and cumulative variability scores of the PCs graphically. By convention, in cases of good dataset fit for PCA, the scree plot tends to have a very distinct elbow -like turn from higher to lower PCs; this effect is known as the “Elbow effect” [25]. A more continuous scree plot points out to the dataset having many equally important PCs.

2.3.4. Eigenvectors, factor scores and contribution (%) of variables

The eigenvectors represent the extent to which a PC's direction is decided by the parent input variable [26]. Factor scores obtained in PCA results correspond to observed data points' coordinates on PCA dimensions under consideration. They are the underlying constructs that are latent to building up of PCs by variables [27]. The contribution percentage of variables to different PCs simply imply their role in determining a particular PC [26]. This has been graphically represented to portray the visually tangible contribution each variable makes to different PCs.

2.3.5. Squared cosines of variables

Squared cosine values showcase the representational quality of given variables on PCA axes which helps in avoiding interpretational errors that might creep in during result analysis. While dealing with multiple independent variables as in the PCA recorded, working with square of cosine and 1-cosine produces a result of related variance v/s unrelated variance which allows for better interpretation of variance of the factors involved. Commonly known as R^2 (correlation coefficient), the values of squared cosines between 0 and 1 are considered for they represent the best fit of data thereby giving us the proportion of variation in dependent variables caused by independent variables [26].

2.3.6. Correlations between variables and PCs, loading plot, Biplot

The correlation centroid has been used to represent the placement of various variables in accordance with their correlation to the corresponding PC axes. While horizontal axis describes the first PC and vertical axis describes the second PC, the length of line connecting the variable datapoint to origin exhibits how well the variable is represented on the factor map. Also, since the centroid has 4 different quadrants the placement of variable data points across these gives insights about how various variables correlate amongst themselves [26]. By convention positively correlated variables share the same quadrant while negatively correlated variables are housed in different quadrants. The loading plot has been a staple of PCA data representation. It simply represents the observations as plotted against the chosen PCs [28]. Generally occurring as distinct clusters of characteristic nearby values that can be grouped into categories, as an anomaly a homogenously mixed and scattered points on loading plot may point out to the inadequacy of PCA in reducing the dimensionality of the given dataset or the possibility of other PCs (than those in question) having a better share of information under them. A biplot combines the loading plot as well as the correlation centroid to give the user a correct representation of variables and observations and their correlations [29]. Amongst different types of biplots, distance biplot was used in this study to find the contribution of variable vectors from their length with respect to the PC axes.

2.3.7. Proximity measures

After the minimum number of dimensions have been chosen for representing the dataset proximity measures are done to quantify how “close” two objects are. *Dissimilarities* represent the distance between two objects measured as *Euclidean distance* and *Similarities* represent how close two objects are [30].

2.3.8. Kruskal's stress, MDS plot and Shepard diagrams

Kruskal stress value determines how a particular set of data has been represented by the model that the analysis imposes. Based on goodness of fit, Kruskal's stress value lower than 0.05 are considered as ideal with

Table 1
Statistical data of input variables as collected from 115 datasets on algal biochar production by pyrolysis.

	C (%)	H (%)	O (%)	N (%)	Ash (%)	Fixed carbon (%)	Volatile compound (%)	Carbohydrate (%)	Protein (%)	Lipid (%)	Residence time (min)	Heating rate (°C min^{-1})	Temperature (°C)	Algal biochar yield (%)
Count	107	104	101	107	107	89	89	63	62	63	85	79	114	97
Mean	43.58	6.29	39.58	6.60	13.28	14.16	67.59	26.79	47.73	10.34	60.12	20.96	457.68	34.79
Std. dev	7.06	0.98	10.87	2.99	9.00	6.09	9.24	13.74	14.55	13.35	34.27	22.06	102.00	10.41
Min	17.20	4.02	21.27	0.86	0.76	0.06	45.65	9.00	4.90	0.09	20.00	1.00	150.00	14.76
25 %	38.84	5.09	31.74	4.11	8.06	10.98	62.02	15.5	38.45	0.25	30.00	10.00	400.00	28.00
50 %	45.13	6.60	35.10	6.29	9.50	12.77	68.31	23.44	49.60	6.25	60.00	10.00	462.50	32.00
75 %	49.48	7.05	47.40	8.89	17.33	16.61	75.55	37.31	54.35	18.00	60.00	20.00	512.50	39.50
Max	55.79	7.99	62.44	11.19	43	31.13	82.03	65.82	72.00	42.90	180.00	100	650.00	66.07

0.025–0.000 ranged values being considered excellent fit [31]. The MDS plot have been obtained for two discrete dimensions only and represents the spread of 16 input variables along newly constructed 2 dimensions. Shepard diagrams are used to illustrate the dataset's goodness of fit into newer dimensions as formed upon multidimensional scaling. Used alongside PCA, Shepard diagram validated the PCA results by exhibiting how well the data points fit before and after dimensional reduction. A closely clustered Shepard plot indicate the low level of dissimilarity before and after data transformation and vice versa [32,33].

3. Results and discussion

3.1. Statistical analysis of variables

From various sources of literature, a total of 115 datasets were curated for PCA analysis of variables affecting the algal biochar yield. All these datasets strictly accounted for only those observations that underwent pyrolysis as production technique and any of the missing values were imputed with mode as it is the most frequent value in any given dataset. The variables under study were classified as ultimate analysis data [C (%), H (%), N (%), O (%) and their ratios: H/C, O/C, N/C], proximate analysis data [ash content (%), fixed carbon (%), volatile compound (%)], biomolecular composition [carbohydrate (%), lipid (%), protein (%)] and pyrolysis parameters [temperature, heating rate, residence time]. These all were clubbed under quantitative input variables and algal biochar yield % was taken as qualitative output variable. It was categorized into very high (> 49 %), high (35–49 %), medium (18–34 %) and low (0–17 %) categories. Besides two supplementary variables of feedstock type; microalgae and macroalgae were also assigned binary values and considered for better representation of data. This proved to be helpful in representing the incidence of algal biochar yield with respect to type of parent algae used.

The basis of taking all these variables into account was their effect on algal biochar yield and properties. For instance, the proportion of carbohydrates, lipids and proteins decide the progression of pyrolysis (the rate of dehydration and decomposition into simpler compounds like anhydrosugars, amines and hydrocarbons) and the products formed (bio-oil, *syn*-gas and biochar) [34]. Demands of "designer char" having optimal H/C, O/C, N/C ratios that guarantees enhanced properties because the more the number of hydroxyl, amine and carboxyl groups the more freely catalysis or cation exchange take place on biochar surface highlights the importance of elemental ratios of feedstock in determining properties of algal biochar. [35–37]. Also, there is a notable decrease in yield and increase in biochar properties like graphitisation degree, surface area and electrical conductivity as the temperature increases from 350 °C to above 700 °C as most of carbon atoms become sp^2 hybridized, with strong π-π interactions, resulting in high electron transfer rates of biochar as reported in various literature [38,39]. This mandatorily places temperature and feedstock composition variables as main determinants of algal biochar production.

The statistical analysis of the collected data is given in Table 1, where the count shows the number of datasets under consideration for a particular variable. The distribution pattern of all the collected data is well defined by measures of mean and standard deviation. The maximum and minimum values of every given variable along with 4 quartile values define the whole range and data distribution of input and output parameters. The same have been illustrated with box plots that exhibit the spread of data in Fig. S1 (*supplementary*). The C and H content of algae is found to be in ranges of 17.2–55.79 % and 4.02–7.99 %, respectively, which is lesser than lignocellulosic biomass [39]. As seen in Fig. 1, the high N content of algal biomass, around 6.6 % can be attributed to presence of protein groups [35]. These protein groups also contribute significantly to the -NH_2 functional groups that confer the resultant char its basicity and catalytic properties [40]. Molar ratios of O/C and N/C are also quite high which indicates the abundance of proteins and carbohydrates in algal cells [35]. The mean ash content of

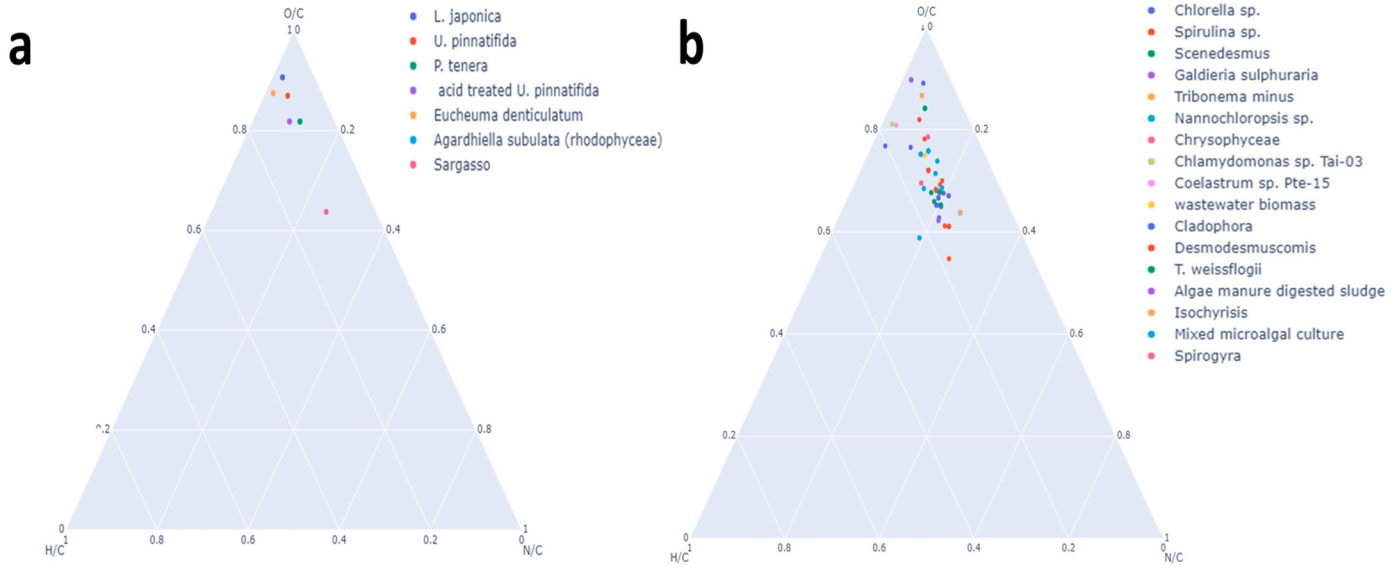

Fig. 1. Ternary phase diagrams for elemental ratio (H/C: O/C: N/C) of a) Macroalgal species b) Microalgal species.

Table 2
Pearson correlation matrix for 16 qualitative and quantitative and 2 supplementary variables collected from 116 algal biochar production by pyrolysis datasets.

Variables	C (%)	H (%)	O (%)	N (%)	Ash (%)	H/C	O/C	N/C	Fixed carbon (%)	Volatile compound (%)	Carbohydrate (%)	Protein (%)	Lipid (%)	Residence time (min)	Heating rate (°C/min)	Temperature (°C)	Microalgae	Macroalgae	Colour bar
C (%)	1	0.632	-0.755	0.632	-0.749	-0.523	-0.902	0.419	0.444	0.477	-0.096	-0.036	0.203	0.221	-0.140	0.063	0.460	-0.460	+1
H (%)	0.632	1	-0.766	1.000	-0.481	0.306	-0.748	0.959	0.295	0.418	-0.074	0.108	0.003	0.258	-0.156	0.071	0.481	-0.481	
O (%)	-0.755	-0.766	1	-0.767	0.627	0.143	0.944	-0.698	-0.381	-0.626	0.050	0.087	-0.212	-0.111	0.192	-0.126	-0.454	0.454	
N (%)	0.632	1.000	-0.767	1	-0.481	0.306	-0.748	0.959	0.296	0.418	-0.075	0.108	0.004	0.259	-0.156	0.072	0.481	-0.481	
Ash (%)	-0.749	-0.481	0.627	-0.481	1	0.386	0.741	-0.335	-0.298	-0.687	-0.058	-0.107	-0.108	0.093	0.454	-0.100	-0.205	0.205	
H/C	-0.523	0.306	0.143	0.306	0.386	1	0.350	0.138	-0.410	0.042	0.217	-0.213	0.068	-0.212	-0.118	-0.175	-0.091	0.091	
O/C	-0.902	-0.748	0.944	-0.748	0.741	0.350	1	-0.616	-0.417	-0.597	0.053	0.110	-0.229	-0.132	0.149	-0.103	-0.455	0.455	
N/C	0.419	0.959	-0.698	0.959	-0.335	0.138	-0.616	1	0.208	0.376	-0.076	0.134	-0.032	0.224	-0.128	0.089	0.435	-0.435	
Fixed carbon (%)	0.444	0.295	-0.381	0.296	-0.298	-0.410	-0.417	0.208	1	-0.112	-0.343	0.218	0.016	0.498	0.104	0.175	0.122	-0.122	
Volatile compound (%)	0.477	0.418	-0.626	0.418	-0.687	0.042	-0.597	0.376	-0.112	1	0.075	0.015	0.242	-0.300	-0.477	-0.009	0.211	-0.211	0
Carbohydrate (%)	-0.096	-0.074	0.050	-0.075	-0.058	0.217	0.053	-0.076	-0.343	0.075	1	-0.369	-0.412	-0.088	-0.299	-0.317	-0.213	0.213	
Protein (%)	-0.036	0.108	0.087	0.108	-0.107	-0.213	0.110	0.134	0.218	0.015	-0.369	1	-0.416	0.124	-0.119	0.374	-0.037	0.037	
Lipid (%)	0.203	0.003	-0.212	0.004	-0.108	0.068	-0.229	0.032	0.016	0.242	-0.412	-0.416	1	0.141	0.165	-0.050	0.197	-0.197	
Residence time (min)	0.221	0.258	-0.111	0.259	0.093	-0.212	-0.132	0.224	0.498	-0.300	-0.088	0.124	0.141	1	0.289	0.005	0.118	-0.118	
Heating rate (°C min^{-1})	-0.140	-0.156	0.192	-0.156	0.454	-0.118	0.149	-0.128	0.104	-0.477	-0.299	-0.119	0.165	0.289	1	0.036	0.038	-0.038	
Temperature (°C)	0.063	0.071	-0.126	0.072	-0.100	-0.175	-0.103	0.089	0.175	-0.009	-0.317	0.374	-0.050	0.005	0.036	1	0.102	-0.102	
Microalgae	0.460	0.481	-0.454	0.481	-0.205	-0.091	-0.455	0.435	0.122	0.211	-0.213	-0.037	0.197	0.118	0.038	0.102	1	-1.000	-1
Macroalgae	-0.460	-0.481	0.454	-0.481	0.205	0.091	0.455	-0.435	-0.122	-0.211	0.213	0.037	-0.197	-0.118	-0.038	-0.102	-1.000	1	

algal feedstock is 13.28 % which is high enough to support the evidence of inorganic residues like K, Mg, Na, Si etc. in resultant biochar. These metal and non-metal traces not only enhance properties like graphitisation, adsorption and catalysis but also themselves help in pyrolysis process as auto-catalysts [41]. Fixed carbon content and volatile compound are pegged at mean 14.16 % and 67.59 %, respectively, which presumably give rise to a porous, high C content, algal biochar having high heating value (HHV) which is almost similar to bituminous coal. This can be attributed to the high rate of devolatilization and carbonization of biochar at higher pyrolysis temperatures. From the datasets under study high lipid content at an average of 10.34 % supports the fact that as temperature is risen from 400 to 600 °C high yield of bio-oils from algal biomass can be achieved successfully [42]. The pyrolysis temperature ranges from 150 to 600 °C, heating rate ranges from 8 to 100 °C min^{-1} and residence time ranges from 20 to 180 min. All these ranges are inclusive of most standard pyrolysis conditions so followed. Consequently, the algal biochar yield ranges from 14.76 % at its lowest to 66.07 % at the highest available value from dataset. With 64.34 % abundance in observed data points medium yield category of algal char yield has the highest number of entries amongst different yield categories as given in Table S1 (*supplementary*).

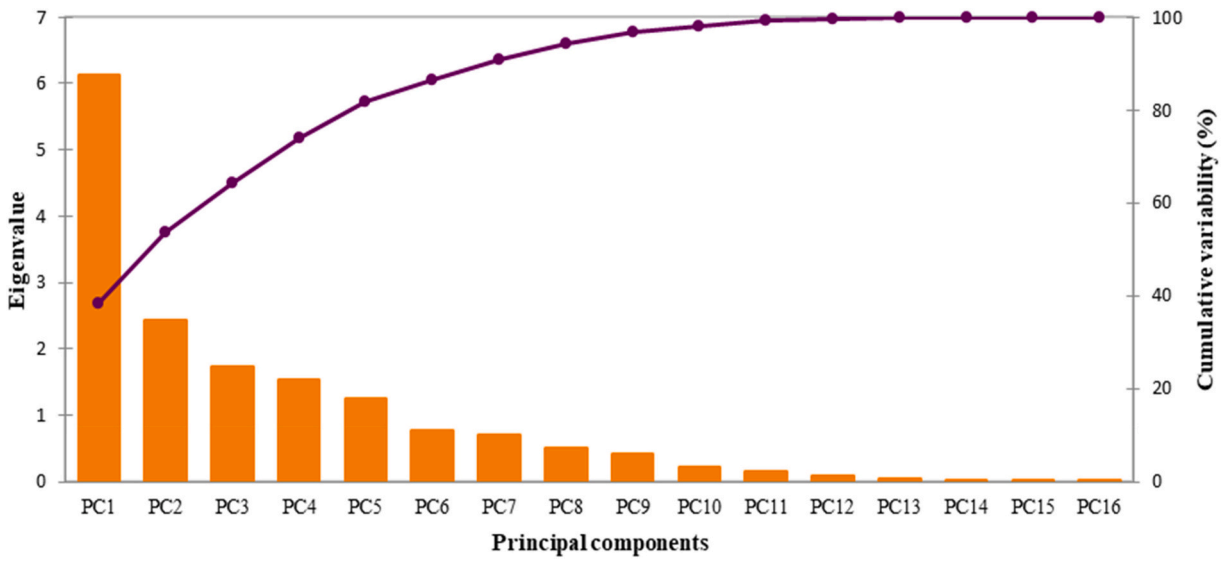

Fig. 2. Scree plot obtained from PCA of 16 variables influencing algal biochar yield showing 5 major principal components.

3.2. Principal components and correlations derived through PCA

3.2.1. Pearson correlation matrix

The factors influencing algal biochar yield tend to vary in synchrony with each other as efforts are made to optimise a single important factor for achieving high algal biochar yield. This necessitates the need to highlight the relationship various feedstock and pyrolysis variables share with one another and type of algal feedstock used. The Pearson correlation matrix given in Table 2 illustrates the relation input variables share amongst themselves. It can be observed in line with Yu et al. [2] that H/C and O/C ratios are positively correlated to macroalgae due to higher concentration of carbohydrates in macroalgal biomass but negatively correlated to microalgae. However, N/C ratio is positively correlated to microalgal biomass which is supported by the abundance of N containing functional groups in protein profile of microalgal feedstock. After pyrolysis, nitrogen content remains almost unchanged due to the presence of elevated ash concentrations. Further, alkali and alkaline earth metal traces result in higher pH of microalgae-derived algal biochar [43]. Ash content of microalgae is also exhibited to be higher than that of macroalgae and is also very positively related to N content [44]. Also, both fixed carbon content and volatile compound content are positively correlated to microalgae as well which explains the HHV and increasingly porous structure of algal char [45]. As mentioned in literature, C content increases with temperature, while H/C and O/C decrease thereby giving low yielding but high HHV algal biochar [46,47]. As seen from Table 2, residence time and heating rate also are closely related in positive aspect. Also fixed carbon content shares a positive relationship with temperature as well which substantiates the findings of Nizamuddin et al. [48].

3.2.2. Eigenvalues and scree plot

Upon PCA application the algal biochar yield dataset throws up first 5 PCs that have eigenvalues above 1 and also account for nearly 81.75 % of variance as given in Table S2 (*supplementary*). As an illustrative method of shortlisting PCs, a standard Scree plot is also used. Given in Fig. 2, the scree plot faintly earmarks first 5 PCs as having larger eigenvalues ($6.123 > 2.439 > 1.739 > 1.523 > 1.249$) than other PCs which are smaller and have comparable sizes (in the range of 0–0.7). Interestingly the eigenvalues, variability as well as the scree plot point towards an important cue from the dataset: that unlike the convention of first 2–3 PCs accounting for maximum variance of dataset, the results in this manuscript put forth just 53.51 % of cumulative variability from first 2 PCs which necessitates taking successive PCs under consideration. Due to the same the scree plot also tends to acquire an uncharacteristic curvature than the elbow like sharp turn from highest PC value. These observations illustrate that the given dataset has a lot of variance and dimensions that may not be covered adequately by PCA thereby laying potential for applying advanced chemometric analysis to our dataset besides reinforcing the literature findings that algal char yield is an intricate interplay of numerous variables [49,50].

3.2.3. Eigenvectors, factor loadings and contribution of variables

Eigenvectors in PCA define the directions of the feature space of PCs. These vector values provide the user with knowledge about the directions in which the information of the parent variables of dataset is compressed. The closer the vector value is to $+1$ the more that variable contributes to PC axis [26]. Values closer to -1 can be discarded as they don't contribute to PC axes. The more negative the eigenvector value of a variable in the PC score chart, the lesser is its contribution to that PC. The same has been corroborated by the correlation chart of PCs with variables wherein the negative values imply lower contribution, and the positive values represent higher contribution to PC as directed by their respective magnitude. Based on the results obtained it is seen that PC1's direction is determined by O (%), O/C, and ash (%). Similarly, the PC2 axis is defined by residence time, fixed carbon content (%), and heating rate (in decreasing order). Hence the difference in PC1 and PC2 axes is that the former depends largely on proximate and ultimate analysis variables whereas the latter depends on pyrolysis parameters. This emphasizes the influence that feedstock elemental composition and pyrolysis conditions wield over algal biochar yield. The highest recorded eigenvector for the complete dataset is for lipid (%) that contributes 0.59879 to PC3 axis followed by contribution of 0.446 by carbohydrate (%) to PC5 and contribution of 0.3987 by residence time to PC2 axis as mentioned in Table S3 (*supplementary*).

Factor loading of PCs gives the correlation coefficients between variables under study and PCs that they aid in constructing. They illustrate the combination from which a particular PC is constructed, the contribution being given by the magnitude of loading [26]. In the dataset used for this study, PC1 exhibited the highest loading values of 0.939, 0.908 and 0.753 from O/C, O (%) and ash (%), respectively, succeeded by PC3 which was constructed by lipid (%) (loading value of 0.790). PC2 and PC4 took into account 7 and 8 variables, respectively, for their construction from the initial set of 18 parent variables (16 quantitative-qualitative and 2 supplementary variables) as given in Table S4 (*supplementary*). The values of factor loadings tended to follow the same pattern as that of eigenvectors but differed in magnitude which

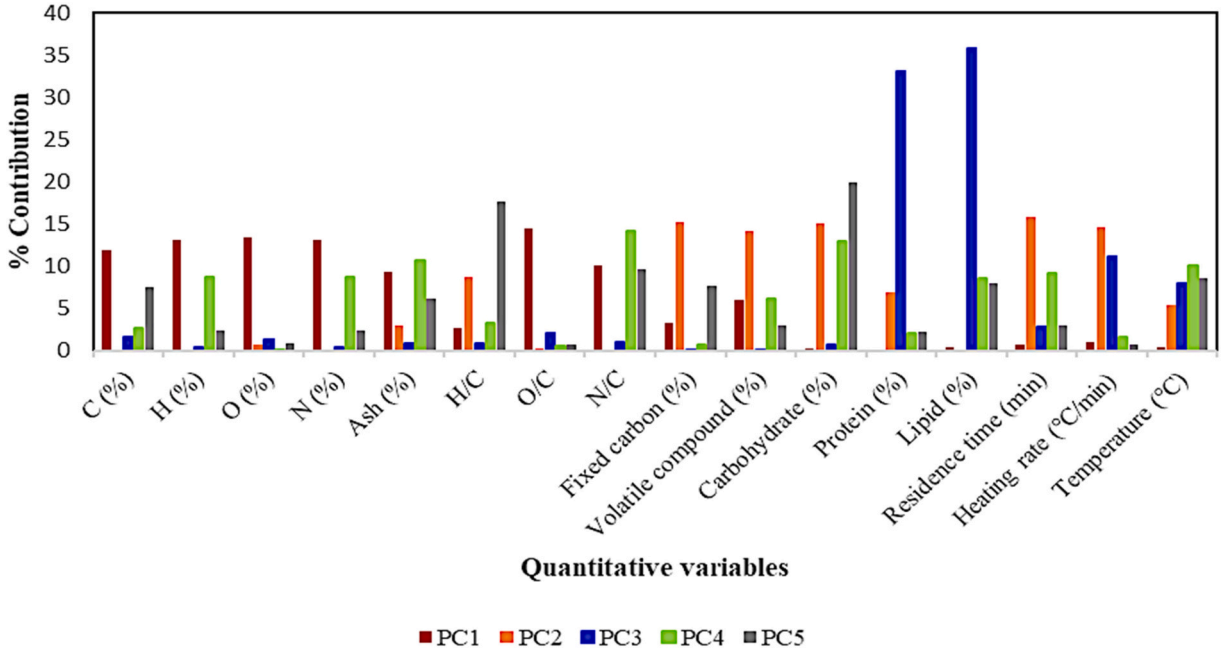

Fig. 3. Contribution of 16 variables to top 5 principal components influencing algal biochar yield.

Table 3
Squared cosines of variables showing their contribution to top 5 principal compnents for algal biochar yield.

Input variables	PC1	PC2	PC3	PC4	PC5	Colour bar
C (%)	0.7323	0.0014	0.0276	0.0412	0.0943	
H (%)	0.7990	0.0022	0.0075	0.1325	0.0295	
O (%)	0.8239	0.0176	0.0218	9.67E-05	0.0097	1
N (%)	0.7990	0.0023	0.0074	0.1322	0.0296	
Ash (%)	0.5676	0.0714	0.0143	0.1627	0.0764	
H/C	0.1660	0.2113	0.0147	0.0502	0.2204	
O/C	0.8811	0.0062	0.0347	0.0083	0.0087	
N/C	0.6165	0.0008	0.0178	0.2161	0.1202	
Fixed carbon (%)	0.2017	0.3698	0.0008	0.0111	0.0947	
Volatile compound (%)	0.3679	0.3437	0.0007	0.0931	0.0364	
Carbohydrate (%)	0.0141	0.3690	0.0119	0.1981	0.2483	
Protein (%)	0.0076	0.1666	0.5760	0.0312	0.0282	
Lipid (%)	0.0222	0.0001	0.6234	0.1312	0.0987	
Residence time (min)	0.0447	0.3877	0.0484	0.1385	0.0371	
Heating rate (°C min^{-1})	0.0580	0.3564	0.1926	0.0232	0.0094	
Temperature (°C)	0.0207	0.1315	0.1386	0.1526	0.1062	
Microalgae	0.2229	0.0046	0.0225	0.0036	0.0423	0
Macroalgae	0.2229	0.0046	0.0225	0.0036	0.0423	

concretizes the new feature space's dimensions. The contribution (%) of variables to different PCs is illustrated in Fig. 3 and Table S5 (*supplementary*).

3.2.4. Squared cosine of variables

Squared cosines determine the proportion of variation amongst dependent variables. In this study we chose R^2 values between 0 and 1 as they represent the best possible fit of data with respect to the chosen 5 PCs. Amongst the R^2 values obtained for each PC, PC1 has the highest values followed by PC3 and PC2 (Table 3). This clearly implies that PC1 has the vastest spread of data that fits optimally in the new feature space constructed by 5 PCs. Determined mostly by proximate and ultimate analysis variables PC1 exhibits the proportion of variation in algal biochar yield due to variables having values closer to 1. As discussed earlier the C (%), H (%), O (%), N (%), N/C, O/C, H/C, fixed carbon (%), ash (%), volatile compound (%), etc. decide the rate and progression of the pyrolysis process thereby dictating the yield of algal biochar. The thermal process parameters decide the rate of reactions like devolatilization, decomposition and carbonization, whereas the feedstock elemental composition decides the products of pyrolysis – bio-oil, syngas, and biochar, and their respective physiochemical properties.

3.2.5. Correlation between variables and principal components

The correlation chart gives insights about the correlation of different variables to PCs under consideration and the same is graphically represented by the correlation centroid. This tends to provide the relationship between older parent variables and newer dimensions that have been etched out of them. From Fig. 4a and Table 4, it is seen that the PC1 and PC2 contribute to just 53.51 % of dataset information which hints at the need for including more PCs as well as a potential to apply other multidimensional methods for obtaining significant correlations. While heating rate < ash (%) < O (%) < O/C ratio affect PC1 in increasing order, lipid (%) < temperature < residence time < fixed carbon (%) < H/C < N/C determine the PC2 in increasing fashion. Also seen very distinctly from the results is the higher tendency of microalgae and its char to be positively influenced by temperature and residence time and negatively influenced by heating rate (which is in agreement to the literature survey that exhibits a decrease in yield as temperature increases) [51,52]. Also, it is seen that heating rate is negatively correlated to temperature and residence time which supports the literature that states that biochar yields increase with slow heating-longer time period regimen [52–54]. Similarly, O (%), O/C and N/C, H/C are positively correlated within amongst themselves but negatively correlated and housed in opposite quadrants. N (%) is placed at an acute angle to protein (%) in quadrant 2 which finds coherence with the literature that says that majority of N comes from protein present in algal biomass [35]. Volatile compound (%) and H/C, residence time and ash (%) are at 90° to each other and hence are deduced to be unrelated. As indicated in literature [54] since residence time increment causes the vapors from char are to act as catalyst for better char yields, in centroid too the volatile compound (%) and residence time are inversely proportional or highly negatively correlated for they share an obtuse angle. Temperature also shares a close positive correlation with C (%) and fixed carbon content (%) which finds numerous proofs in literature that states that with increasing pyrolysis temperature carbon content of algal biochar increases [46,47]. The increase in carbon content is due to increase in

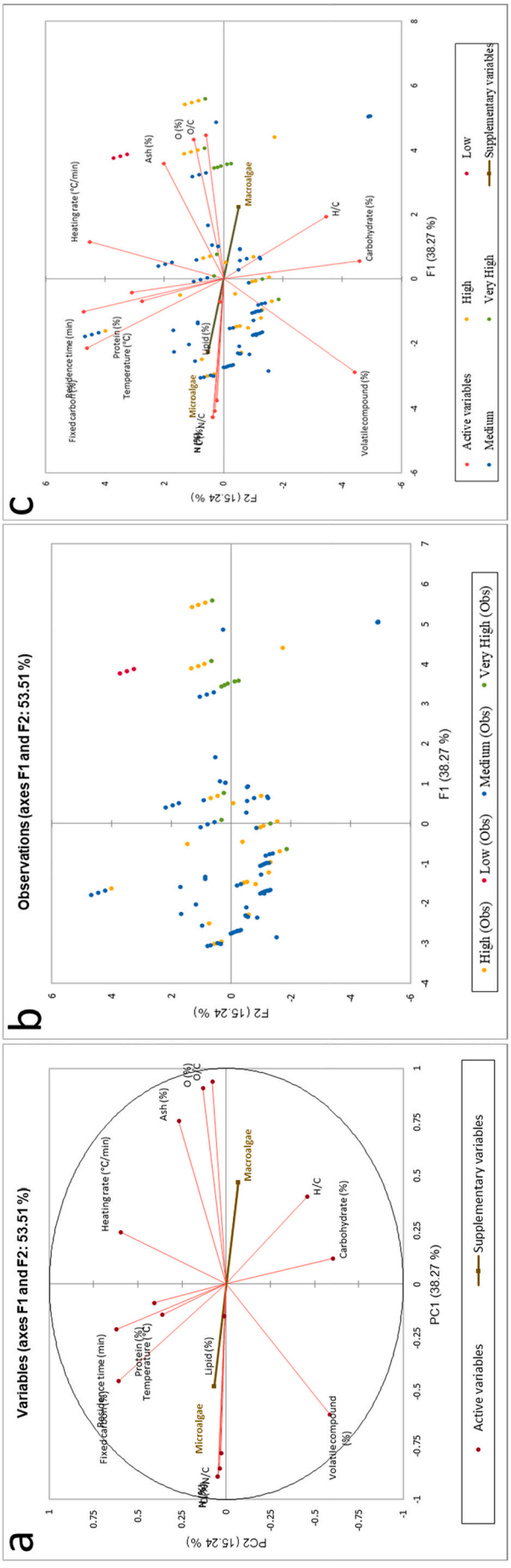

Fig. 4. PCA results a) Correlation centroid b) Loading plot c) Biplot showing the spread of algal biochar yield data points and correlations between 16 input variables spread across PC1 and PC2.

Table 4
Correlations between 16 qualitative and 2 supplementary variables and top 5 PCs.

Input variables	PC1	PC2	PC3	PC4	PC5	Colour bar
C (%)	-0.856	0.038	0.166	-0.203	0.307	Positive correlation
H (%)	-0.894	0.048	-0.087	0.364	-0.172	
O (%)	0.908	0.133	-0.148	-0.010	0.099	
N (%)	-0.894	0.049	-0.086	0.364	-0.172	
Ash (%)	0.753	0.267	0.120	0.403	-0.277	
H/C	0.407	-0.460	0.121	0.224	-0.470	
O/C	0.939	0.079	-0.186	0.092	-0.094	
N/C	-0.785	0.029	-0.133	0.465	-0.347	
Fixed carbon (%)	-0.449	0.608	0.029	-0.106	0.308	
Volatile compound (%)	-0.607	-0.586	-0.027	-0.305	-0.191	
Carbohydrate (%)	0.119	-0.607	-0.109	0.445	0.498	
Protein (%)	-0.088	0.408	-0.759	-0.177	-0.168	
Lipid (%)	-0.149	0.013	0.790	-0.362	-0.314	
Residence time (min)	-0.211	0.623	0.220	0.372	0.193	
Heating rate (°C min^{-1})	0.241	0.597	0.439	0.153	-0.097	
Temperature (°C)	-0.144	0.363	-0.372	-0.391	-0.326	Negative correlation
Microalgae	-0.472	0.069	0.150	0.060	-0.206	
Macroalgae	0.472	-0.069	-0.150	-0.060	0.206	

dehydration and decomposition with an increase in temperature that also causes an increment in HHV of algal biochar.

3.2.6. PCA loading plot

A traditional PCA plot groups data points based on their shared similarity. Upon application of dataset to PCA, lack of well-defined, distinct clusters and limitations due to scarce and singular literature sources is seen to be very visible in the loading plot (Fig. 4b). In quadrant 1 high yield points are seen towards the higher end of PC1. The rest of quadrant 1 is pinned by medium to very high yield data points. Also, the lone 3 low yield data points of the whole dataset also occur in quadrant 1 alone. Quadrants 2 and 3 are found to be housing majority of medium to high yield data points in a slightly homogeneous manner. Quadrant 4 has the least number of data points from the whole plot that belong to medium (7), high (4) and very high (2) yield categories. Since PC1 is constructed majorly out of proximate and ultimate analysis variables and PC2 is constructed by pyrolysis conditions of temperature, heating rate and residence time the occurrence of very high char yield points can be attributed to high C (%), H (%), N (%), O (%) and their ratios and high to medium yield points can be linked to low to medial pyrolysis temperature, heating rate and residence time. The peculiar occurrence of data points in close groups of 2 and 3 can be traced to the similar experimental conditions used for achieving algal char in different research experiments. Also since only a fraction of data points are expressed on loading plot of PC1 and PC2 it can be assumed that other successive PCs (3,4,5) might hold the rest of the data points on their respective plots.

3.2.7. Biplot

A biplot combines the best loading PCA plot and correlation centroid to present insights on correlations of variables and observations. As seen in the biplot results obtained in this study (Fig. 4c), PC1 and PC2 axis houses those clusters of data points of low to very high char yield that are centered around points of O, N and their molar ratios. This is in agreement to the literature evidence that states that elemental proportion and molar ratio of O and N in particular does not affect the yield of algal biochar [10]. High ash content of feedstock also contributes to high yield of algal char as seen in quadrant1 and supported by literature that presence of inorganic elements tends to catalyse the reactions which convert carbohydrates in feedstock to char [55]. Fixed carbon content also was tightly positively correlated with algal char yield as stated by Maddi et al. [56] in works on algae from natural blooms whereby it has

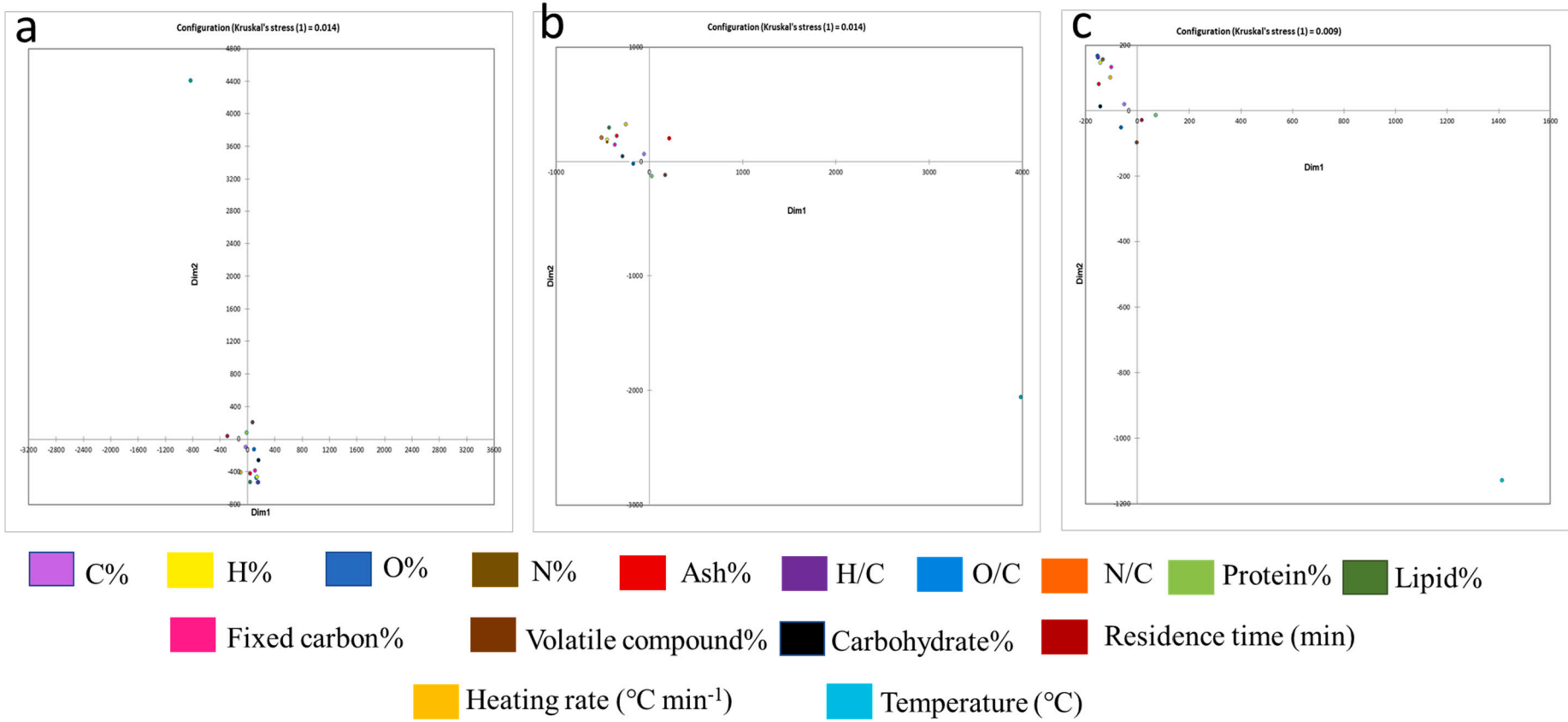

Fig. 5. Multidimensional scaling plots for 16 input variables across Dimension 1 and Dimension 2 for a) Composite b). Microalgal biochar dataset c) Macroalgal dataset.

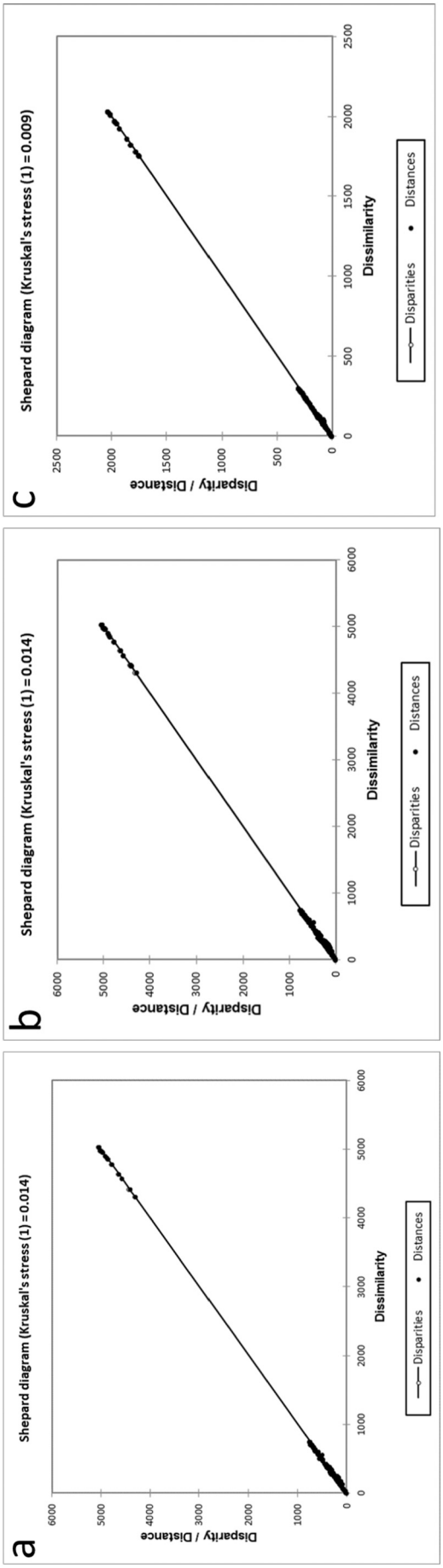

Fig. 6. Shepard plots obtained from Multidimensional scaling for a) Composite dataset b) Microalgal biochar dataset c) Macroalgal biochar dataset.

been stated that yield is high for species with high fixed carbon as upon increase in temperature as the volatiles vaporsise it's just the carbon residue that directly forms char. As provided for by numerous literature sources to this effect [52,57], increased temperature as well as reduced residence time as shown on PC2 tends to cause medium yield of algal biochar. As supported by earlier results, the correlation between pyrolysis temperature and residence time is only slightly positive which puts them in an almost inverse relationship for achieving high algal biochar yield. Overall, these observations place the pyrolysis temperature as the major determinant of char yield. Placement of heating rate on biplot displayed little insight on its effect on algal char yield though lower heating rates are known to favor char production and higher heating rates are known to aid in high bio-oil yield in literature [52,58]. Besides pointing to need for more PCs as discussed earlier the biplot observations also lay more strength to the premise of low temperature-low heating rate-median residence time of char production processes in giving high yield outputs [51]. This also corroborates with the literature review done that states that feedstock composition, more so of microalgae (as compared to macroalgae) is instrumental in achieving higher yields [2,3].

3.3. Support of PCA results by multidimensional scaling

3.3.1. Kruskal's stress and MDS plot

Based on the values of Kruskal stress [31], the best fit representation of data was found to be of macroalgae (0.009) > microalgae (0.014) ≥ composite (0.014) dataset. As documented by MDS plots (Fig. 5), the input variables were found to be clustered in 2nd quadrant for macroalgae as well as microalgae whereas the input variables for the composite dataset were clustered near origin and 4th quadrant. In support of PCA results the dimensions created in MDS for composite dataset also followed the same trend of N/C, O/C, H/C, carbohydrate being the chief contributors of dimension 1, and temperature, residence time, and volatile carbon compound were found to be the prime constructors of dimension 2. Upon comparison, it was found that mostly pyrolysis process parameters contributed to MDS dimensions in case of microalgae. For macroalgal biochar, 7 variables namely, temperature, residence time, N (%), C (%), H (%), N/C, H/C contributed to formation of newer dimensions. As confirmed by correlations deduced from PCA, ash (%) in microalgal biochar was again found to be one of the chief factors influencing dimension 2 of microalgal dataset. Similarly, volatile compound (%) was found to be chief contributor of dimension 1 for microalgal biochar dataset in accordance with PCA results.

3.3.2. Shepard plot

The Shepard diagram is a scatterplot of distances between points in the MDS plot against the observed dissimilarities or similarities. As seen in Fig. 6 the Shepard plot for composite, microalgal biochar and macroalgal biochar the plots showcased monotonic, linear scaling of pre-transformed and transformed data points of 16 variables. The distances between the scaling of data points are exhibited to be small in configuration and these points have a clustered scattering showing an increasing trend because dissimilarity was used. Since the Kruskal stress value was found to be least for macroalgae the Shepard plot for the same is most linear having best fit data points.

To summarise PCA and MDS results in a nutshell, pyrolysis process parameters (temperature, heating rate, residence time) and C (%), H (%), O (%), N (%) and their elemental ratios (H/C, O/C, N/C) are the major decisive factors influencing algal biochar yield from amongst all the 16 input variables so chosen. Additionally, microalgae are shown to have a better chance as feedstock for algal biochar when aiming for higher yield with better physicochemical properties. The knowledge extracted from the chemometric analysis can thus be used to address the research gap in achieving higher algal biochar yield in a targeted way whereby focused efforts on optimizing the most influential factors can be made. This can prove to be instrumental for the quest of introducing

algal biochar as a viable green alternative to charcoal in varied applications as the results can be used as important cues for designing char production process and feedstock.

4. Conclusion

Principal component analysis (PCA) was performed to find main determinants dictating the algal char yield as well as the correlations between these 16 parent variables. Out of the total 16 PCs produced by PCA, top 5 were chosen based on eigenvalues and these held 81.705 % of variability of the dataset under study. Both the correlation centroid and biplot highlighted the pattern followed by high to very high algal char yields to low temperature-less residence time-moderate heating rate regimen. Categorical classification of results also pointed out to the aptness of microalgae for biochar production over macroalgae as validated by MDS results. However, the results discussed above fall short of ranking-prioritizing the factors that influence algal biochar yield. Also, the inclusion of results from near similar experimental process parameters narrowed the range of dataset besides overruling the mean-based imputation. Hence, as a desired direction for future research, a more varied dataset for algal biochar production can be used. Further, experimental results of algal biochar produced by latest thermochemical techniques of hydrothermal carbonization, torrefaction, microwave assisted pyrolysis can also be included in the dataset to ascertain the linkage between process suitability and feedstock used with respect to algal biochar yield.

CRediT authorship contribution statement

Astha Kapoor: Data collection, Analysis, Writing- Original draft preparation

Nageshwari Krishnamoorthy: Analysis, Writing- Original draft preparation

Abhijeet Pathy: Writing- Original draft preparation

Balasubramanian Paramasivan: Conceptualization, Funding acquisition, Writing- Reviewing and Editing, Final approval

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgement

The authors thank the Department of Biotechnology and Medical Engineering of National Institute of Technology Rourkela for providing the necessary research facility. The authors greatly acknowledge the Ministry of Education of Government of India for sponsoring the Master of Technology (M.Tech.) programme of the first author.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.algal.2022.102908.

References

[1] C. Xia, A. Pathy, B. Paramasivan, P. Ganeshan, K. Dhamodharan, A. Juneja, K. Rajendran, Comparative study of pyrolysis and hydrothermal liquefaction of microalgal species: analysis of product yields with reaction temperature, Fuel 311 (2022), 121932.

[2] K.L. Yu, B.F. Lau, P.L. Show, H.C. Ong, T.C. Ling, W.H. Chen, J.S. Chang, Recent developments on algal biochar production and characterization, Bioresour. Technol. 246 (2017) 2–11.

[3] Y.D. Chen, F. Liu, N.Q. Ren, S.H. Ho, Revolutions in algal biochar for different applications: state-of-the-art techniques and future scenarios, Chin. Chem. Lett. 31 (10) (2020) 2591–2602.

[4] M. Sekar, T. Mathimani, A. Alagumalai, N.T.L. Chi, P.A. Duc, S.K. Bhatia, A. Pugazhendhi, A review on the pyrolysis of algal biomass for biochar and bio-oil–bottlenecks and scope, Fuel 283 (2021), 119190.

[5] M. Sevilla Solís, W. Gu, C. Falco, M.M. Titirici, A.B. Fuertes Arias, G. Yushin, Hydrothermal Synthesis of Microalgae-derived Microporous Carbons for Electrochemical Capacitors, 2014.

[6] C. Zhang, S.H. Ho, W.H. Chen, Y. Xie, Z. Liu, J.S. Chang, Torrefaction performance and energy usage of biomass wastes and their correlations with torrefaction severity index, Appl. Energy 220 (2018) 598–604.

[7] S. Aravind, P.S. Kumar, N.S. Kumar, N. Siddarth, Conversion of green algal biomass into bioenergy by pyrolysis. A review, Environmental Chemistry Letters 18 (3) (2020) 829–849.

[8] S.H. Ho, C. Zhang, F. Tao, C. Zhang, W.H. Chen, Microalgal torrefaction for solid biofuel production, Trends in Biotechnology 38 (9) (2020) 1023–1033.

[9] P. Das, V.P. Chandramohan, T. Mathimani, A. Pugazhendhi, A comprehensive review on the factors affecting thermochemical conversion efficiency of algal biomass to energy, Sci. Total Environ. 766 (2021), 144213.

[10] A. Pathy, S. Meher, P. Balasubramanian, Predicting algal biochar yield using eXtreme gradient boosting (XGB) algorithm of machine learning methods, Algal Res. 50 (2020), 102006.

[11] K. Kumar, Principal component analysis: most favourite tool in chemometrics, Resonance 22 (8) (2017) 747–759.

[12] B. Škrbić, J. Cvejanov, N. Đurišić-Mladenović, Chemometric characterization of vegetable oils based on the fatty acid profiles for selection of potential feedstocks for biodiesel production, J. Biobased Mater. Bioenergy 9 (3) (2015) 358–371.

[13] K. Kim, N. Labbé, J.M. Warren, T. Elder, T.G. Rials, Chemical and anatomical changes in Liquidambar styraciflua L. Xylem after long term exposure to elevated CO2, Environ. Pollut. 198 (2015) 179–185.

[14] M. Mancini, Å. Rinnan, A. Pizzi, C. Mengarelli, G. Rossini, D. Duca, G. Toscano, Near infrared spectroscopy for the discrimination between different residues of the wood processing industry in the pellet sector, Fuel 217 (2018) 650–655.

[15] B.D. Škrbić, N. Đurišić-Mladenović, J. Cvejanov, Differentiation of syngases produced by steam gasification of mono-and mixed sources feedstock: a chemometric approach, Energy Convers. Manag. 171 (2018) 1193–1201.

[16] A. Smoliński, N. Howaniec, Analysis of porous structure parameters of biomass chars versus bituminous coal and lignite carbonized at high pressure and temperature—a chemometric study, Energies 10 (10) (2017) 1457.

[17] I.T. Jolliffe, J. Cadima, Principal component analysis: a review and recent developments, Philos. Trans. R. Soc. A Math. Phys. Eng. Sci. 374 (2065) (2016) 20150202.

[18] K. Nageshwari, V. Senthamizhan, P. Balasubramanian, Sustaining struvite production from wastewater through machine learning based modelling and process validation, Sustainable Energy Technol. Assess. 53 (2022), 102608.

[19] K.J. Nishanth, V. Ravi, Probabilistic neural network based categorical data imputation, Neurocomputing 218 (2022) 17–25.

[20] G.W. Milligan, M.C. Cooper, A study of standardization of variables in cluster analysis, J. Classif. 5 (2) (1988) 181–204.

[21] V.A. Profillidis, G.N. Botzoris, Chapter 5—statistical methods for transport demand modeling, in: V.A. Profillidis, G.N. Botzoris (Eds.), Modeling of Transport Demand, 2019.

[22] J.B. Kruskal, M. Wish, Multidimensional scaling (No. 11), Sage, 1978.

[23] R.N. Shepard, P. Arabie, Additive clustering: representation of similarities as combinations of discrete overlapping properties, Psychol. Rev. 86 (2) (1979) 87.

[24] P. Sedgwick, Pearson's correlation coefficient, BMJ 345 (2012).

[25] M. Zhu, A Simple Technique for Automatically Selecting the Number of Principal Components Via the Use of Profile Likelihood. Working Paper, 2004.

[26] H. Abdi, L.J. Williams, Principal component analysis, Wiley Interdiscip. Rev. Comput. Stat. 2 (4) (2010) 433–459.

[27] T. Kurita, Principal component analysis (PCA), in: Computer Vision: A Reference Guide, 2019, pp. 1–4.

[28] R. Bro, A.K. Smilde, Principal component analysis, Analytical methods 6 (9) (2014) 2812–2831.

[29] S.M. Holland, Principal Components Analysis (PCA), Department of Geology, University of Georgia, Athens, GA, 30602-2501, 2008.

[30] M.L. Davison, C.S. Ding, S.K. Kim, Multidimensional Scaling. New York, 1983.

[31] J.B. Kruskal, Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis, Psychometrika 29 (1) (1964) 1–27.

[32] P. Groenen, M. van de Velden, Multidimensional Scaling (No. EI 2004-15), 2004.

[33] L. Wilkinson, Multidimensional scaling, Systat 10 (2) (2002) 119–145.

[34] G. De Bhowmick, A.K. Sarmah, R. Sen, Zero-waste algal biorefinery for bioenergy and biochar: a green leap towards achieving energy and environmental sustainability, Sci. Total Environ. 650 (2019) 2467–2482.

[35] W. Chen, H. Yang, Y. Chen, M. Xia, X. Chen, H. Chen, Transformation of nitrogen and evolution of N-containing species during algae pyrolysis, Environ. Sci. Technol. 51 (11) (2017) 6570–6579.

[36] H.J. Cho, K. Baek, J.K. Jeon, S.H. Park, D.J. Suh, Y.K. Park, Removal characteristics of copper by marine macro-algae-derived chars, Chem. Eng. J. 217 (2013) 205–211.

[37] A.R. Zimmerman, Abiotic and microbial oxidation of laboratory-produced black carbon (biochar), Environ. Sci. Technol. 44 (4) (2010) 1295–1301, https://doi.org/10.1021/es90314.
[38] S.H. Ho, R. Li, C. Zhang, Y. Ge, G. Cao, M. Ma, N.Q. Ren, N-doped graphitic biochars from C-phycocyanin extracted spirulina residue for catalytic persulfate activation toward nonradical disinfection and organic oxidation, Water Res. 159 (2019) 77–86.
[39] C. Yang, R. Li, B. Zhang, Q. Qiu, B. Wang, H. Yang, C. Wang, Pyrolysis of microalgae: a critical review, Fuel Process. Technol. 186 (2019) 53–72.
[40] D. Saha, M.J. Kienbaum, Role of oxygen, nitrogen and sulfur functionalities on the surface of nanoporous carbons in CO2 adsorption: a critical review, Microporous Mesoporous Mater. 287 (2019) 29–55.
[41] S.H. Ho, Z.K. Yang, D. Nagarajan, J.S. Chang, N.Q. Ren, High-efficiency removal of lead from wastewater by biochar derived from anaerobic digestion sludge, Bioresour. Technol. 246 (2017) 142–149.
[42] D.C. Li, H. Jiang, The thermochemical conversion of non-lignocellulosic biomass to form biochar: a review on characterizations and mechanism elucidation, Bioresour. Technol. 246 (2017) 57–68.
[43] G. Binda, D. Spanu, R. Bettinetti, L. Magagnin, A. Pozzi, C. Dossi, Comprehensive comparison of microalgae-derived biochar from different feedstocks: a prospective study for future environmental applications, Algal Res. 52 (2020), 102103.
[44] X.J. Lee, H.C. Ong, Y.Y. Gan, W.H. Chen, T.M.I. Mahlia, State of art review on conventional and advanced pyrolysis of macroalgae and microalgae for biochar, bio-oil and bio-syngas production, Energy Convers. Manag. 210 (2020), 112707.
[45] H. Feng, Y. Jia, D. Shen, Y. Zhou, T. Chen, W. Chen, M. Wang, The effect of chemical vapor deposition temperature on the performance of binder-free sewage sludge-derived anodes in microbial fuel cells, Sci. Total Environ. 635 (2018) 45–52.
[46] Y.M. Chang, W.T. Tsai, M.H. Li, Chemical characterization of char derived from slow pyrolysis of microalgal residue, J. Anal. Appl. Pyrolysis 111 (2015) 88–93.
[47] K.W. Jung, T.U. Jeong, H.J. Kang, K.H. Ahn, Characteristics of biochar derived from marine macroalgae and fabrication of granular biochar by entrapment in calcium-alginate beads for phosphate removal from aqueous solution, Bioresour. Technol. 211 (2016) 108–116.
[48] S. Nizamuddin, H.A. Baloch, G.J. Griffin, N.M. Mubarak, A.W. Bhutto, R. Abro, B. S. Ali, An overview of effect of process parameters on hydrothermal carbonization of biomass, Renew. Sust. Energ. Rev. 73 (2017) 1289–1299.
[49] B.H. Cheng, R.J. Zeng, H. Jiang, Recent developments of post-modification of biochar for electrochemical energy storage, Bioresour. Technol. 246 (2017) 224–233.
[50] S.H. Ho, D. Wang, Z.S. Wei, J.S. Chang, N.Q. Ren, Lead removal by a magnetic biochar derived from persulfate-ZVI treated sludge together with one-pot pyrolysis, Bioresour. Technol. 247 (2018) 463–470.
[51] P. Basu, Pyrolysis and torrefaction, in: R. Zanol (Ed.), Biomass Gasification and Pyrolisis Practical Design and Theory, 2010, p. 1.
[52] M. Tripathi, J.N. Sahu, P. Ganesan, Effect of process parameters on production of biochar from biomass waste through pyrolysis: a review, Renew. Sust. Energ. Rev. 55 (2016) 467–481.
[53] K. Chaiwong, T. Kiatsiriroat, N. Vorayos, C. Thararax, Biochar production from freshwater algae by slow pyrolysis, Maejo Int. J. Sci. Technol. 6 (2) (2012) 186.
[54] D. Mohan, A. Sarswat, Y.S. Ok, C.U. Pittman Jr., Organic and inorganic contaminants removal from water with biochar, a renewable, low cost and sustainable adsorbent–a critical review, Bioresour. Technol. 160 (2014) 191–202.
[55] B.C. Huang, J. Jiang, G.X. Huang, H.Q. Yu, Sludge biochar-based catalysts for improved pollutant degradation by activating peroxymonosulfate, J. Mater. Chem. A 6 (19) (2018) 8978–8985.
[56] B. Maddi, S. Viamajala, S. Varanasi, Comparative study of pyrolysis of algal biomass from natural lake blooms with lignocellulosic biomass, Bioresour. Technol. 102 (23) (2011) 11018–11026.
[57] S. Li, S.H. Ho, C. Wang, Y.C. Lin, D. Nagarajan, J.S. Chang, N.Q. Ren, Integration of sludge digestion and microalgae cultivation for enhancing bioenergy and biorefinery, Renew. Sust. Energ. Rev. 96 (2018) 76–90.
[58] S. Pourkarimi, A. Hallajisani, A. Alizadehdakhel, A. Nouralishahi, Biofuel production through micro-and macroalgae pyrolysis–a review of pyrolysis methods and process parameters, J. Anal. Appl. Pyrolysis 142 (2019), 104599.