Chapter 2

Literature Review

2. Literature Review

2.1. Mechanism of pollutant removal by microalgae

2.1.1. Carbon Assimilation

Carbon dioxide (CO_2) is one of the critical atmospheric pollutants that contribute to the level of greenhouse gas. The rapid development of industries and urban areas are considered to be significant sources of inorganic carbon [47]. Photosynthesis mediated by microalgae represents one of the practical approaches for CO_2 fixation [48], [49]. Microalgal biomass is nearly composed of 50% carbon of its total weight (in %) [50]. Microalgae utilize carbon either in autotrophic or heterotrophic mode. Figure 2.1 represents the flow diagram of carbon assimilation and fatty acid biosynthesis in both autotrophic and heterotrophic cultivation modes. In the autotrophic mode, CO_2 is fixed through a light and dark reactions (Calvin cycle) by using light energy and water molecules with the simultaneous release of oxygen [51], [52]. The schematics have been shown in equation (2.1) and (2.2) [53]:

$$2H_2O + 3ADP + 2NADP^+ + 3P \xrightarrow{light} 2H^+ + 3ATP + 2NADPH + O_2 \quad (2.1)$$

$$3CO_2 + 6NADPH + 9ATP + 6H^+ \xrightarrow{RuBisco} C_3H_6O_3 - phosphate + 6NADP^+ + 9ADP + 8P + 3H_2O \quad (2.2)$$

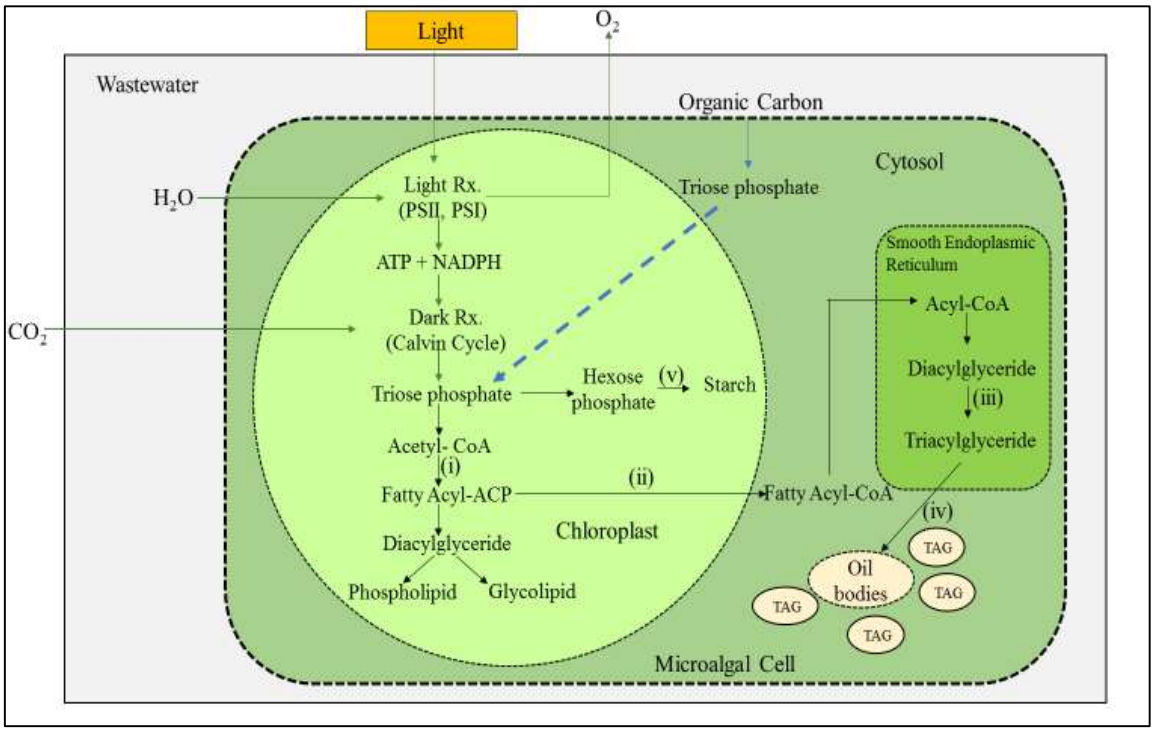

Figure 2.1. Flow diagram for carbon capture and lipid biosynthesis by microalgae. The green line represents the autotrophic mode, and the blue dotted line represents the heterotrophic mode. Number indicates: (i) FAS (fatty acid synthase) and ACCase (acetyl-CoA carboxylase); (ii) acyl-CoA synthetases and fatty acid thioesterases; (iii) Triacyl glyceride (TAG) biosynthesis enzymes, including acyl-CoA: DGAT (diacylglycerol acyltransferase); (iv) pathway of formation of oil bodies; and (v) Starch synthase and ADP-glucose pyrophosphorylase.

In the heterotrophic mode, microalgae utilize organic carbon sources. Figure 2.1 shows how the carbon fixed during photosynthesis is used for synthesis of polysaccharide, precursor fatty acid and other hydrocarbons. Synthesized fatty acids are then translocated to smooth endoplasmic reticulum and converted to triacylglyceride, and it buds off into oil bodies in the cytosol [54]. CO_2 is also required also for the maintenance of the pH in the medium [55]. Additionally, microalgae can also assimilate soluble carbonates for there carbon requirements. When the medium pH is low (5-7), microalgal cells uptake CO_2 through diffusion. At high pH (more than 7), bicarbonate (HCO_3^-) form of carbon is present in the solution. It is transported into the cells through active transport by the activity of external carbonic anhydrase [56]–[58].

2.1.2. Assimilation of nitrogen

Microalgae require nitrogen (N) as one of their essential elements for growth, which can be easily obtained from wastewater in the large amount [59]. It is present in various biological macromolecules including enzymes, proteins, genetic materials (DNA/RNA) and energy transfer units (ATP/ADP). Microalgae assimilate inorganic N (including NH_4^+, $NO_3^-$and NO_2^-) and convert them to organic N. Cyanobacteria also transform atmospheric nitrogen into ammonia through nitrogen fixation [60], [61]. Figure 2.2 shows the pathway for the assimilation of inorganic nitrogen by microalgae.

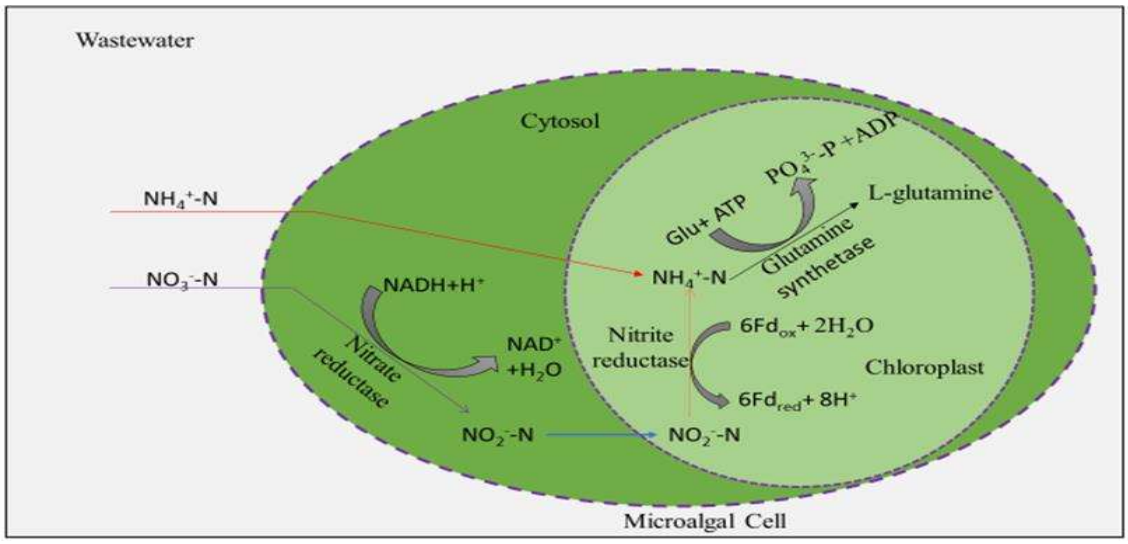

Figure 2.2. Diagrammatic representation of inorganic nitrogen assimilation by microalgae.

It became evident from Figure 2.2 that nitrate assimilation is carried out in two transport pathways mediated by two steps of reduction. Initially, nitrate is translocated in the cytosol through the plasma membrane. Thereafter, nitrate is reduced to nitrite in a reduction reaction catalyzed by Nitrate Reductase (NR) present in the cytosol. It takes two electrons from NADH (nicotinamide adenine dinucleotide- present in reduced form) and transfers it to nitrate. Nitrite is then translocated into the chloroplast, where its reduction to ammonium occurs in a reduction reaction catalyzed by Nitrite reductase (NiR) by the transfer of six electrons from a reduced form of ferredoxin (Fd). Finally, Glutamate synthetase catalyzes the merging of ammonium

into amino acid glutamine by using adenosine triphosphate and glutamate [62]. All types of inorganic N are first reduced to ammonium N before merging into amino acids within the cell [55]. It has been reported that when glutamate was added to the wastewater, it led to 70% further reduction of NH_4^+ by each cell of *Chlorella vulgaris* [63].

2.1.3. Assimilation of phosphorus

Microalgal cells also require phosphorus for their growth. It plays a vital role in controlling its biomass composition in freshwater. It is an integral part of the DNA, RNA, ATP, protein/amino acids and lipids/fatty acids present in the cell wall. It also occurs in intermediates of carbohydrates and fatty acid metabolism and cell membrane materials [64]. The absence or depletion of this nutrient can considerably affect the photosynthetic process [65]. Microalgae perform active transport at the plasma membrane for the uptake of ortho phosphorus in the forms of $H_2PO_4^-$ and HPO_4^{2-}. During algal metabolism, PO_4^{3-}-P is merged into the organic compounds by the following mechanism: (i) oxidative phosphorylation (ii) substrate-level phosphorylation; and (iii) photophosphorylation. These mechanisms include the production of ATP from ADP and energy input. In the first mechanism, energy is grabbed from the ETS (Electron Transport System) occurring in mitochondria. In the second mechanism, respiratory substrate is oxidized to provide energy input. In the third mechanism, ATP is obtained from the transformation of the light energy. The general reaction of phosphorylation has been represented in Eq. (2.3) [66]:

$$ADP + P_i \xrightarrow{Energy} ATP \quad (2.3)$$

When there is a shortage of inorganic phosphate, microalgal cells convert organic phosphate to orthophosphate by phosphatase present on the surface of the cell and utilize them. When there is an excess of phosphate, microalgal cells assimilate them and store them within the cells in the form of polyphosphate (volatin) granules. These granules are utilized for

continued growth during the shortage of phosphate in the growth medium/ environment [67]–[69].

2.2. Assimilation of heavy metals

Microalgae possess the ability to uptaking of heavy metals (HMs) from wastewater. Thus, the concentration of HMs in microalgal cells is higher in comparison to the surrounding medium [70], [71]. Microalgae perform uptake of HMs by different metabolic mechanism [72]. The uptake process generally includes two steps: (i) Initially, metals are rapidly sorbed at the cell surface and (ii) detoxification of HMs by slower metabolic process occurring within the cell. Advantages of using microalgae for the metal bioremediation process include: (i) metal uptake at faster rate in comparison to other adsorption techniques, (ii) time and energy saving, (iii) faster growth rates, (iv) can bind up to 10% of their biomass, (iv) application in both batch and continuous process (v) eco-friendly, recyclable/ reusable,, and (vi) highly applicable in wastewater treatment [73].

Dirbaz and Roosta (2018) examined four microalgae species that were: *Spirulina sp.*, *Parachlorella sp.*, *Nannochloropsisi sp.*, and *Scenedesmus sp.*, for the biosorption capability of Cd^{2+} ions from aqueous solution. *Parachloralla sp.* showed the highest biosorption capacity which was 90.72 mg/g (mass of sorbate/mass of sorbent) at 30°C and pH 7. It was 1.5-3 times greater than other biosorbents investigated. Biosorption by *Parachlorella* was further optimized and maximum uptake was 96.20 mg g^{-1} at 35°C, and pH 7 was reported. Effect of agitation rate was also studied. When the agitation rate was increased ≥ 250 rpm or higher, the uptake of heavy metals was reduced to less than half of the initial bioaccumulation rate [74]. Immobilization techniques are also used to enhance the uptake capacity of microalgal biomass. Ahmad, Bhat, and Buang (2018) investigated the use of both free and immobilized *C. vulgaris* biomass for the sorption of ferrous (Fe^{2+}), manganese (Mn^{2+}) and zinc (Zn^{2+}) ions. *C. vulgaris* biomass was trapped in calcium alginate beads. The authors also studied the effects of initial

metal ion concentration, contact time, biosorbent dosage, and pH. Immobilized biomass achieved maximum biosorption which was 129.83 mg/g for Fe^{2+}, 115.90 for Mn^{2+} and 105.29 for Zn^{2+} at optimal pH of 6.0, a biosorbent dosage of 0.4 g/L with a contact time of 5 h at 25°C [75]. Figure 2.3 represents the heavy metal bioremediation capability of various microalgal species.

Table 2.1. Metal uptake by various species of microalgae.

Metallic Species	**Microalgal Species**	**Initial Conc. of metal**	**Temp.(°C)**	**Medium pH**	**Amount of metal uptake (mg/g)**	**Reference**
Cd^{2+}	*Chaetoceroscalcitrans*	1 ppm	20-22	8	1055.27	[76]
	Desmodesmuspleimorphus (ACOI 561)	5.0 mg L^{-1}	25	4	85.3	[77]
	Desmodesmuspleimorphus(L)	5.0 mg L^{-1}	25	4	61.2	[77]
	Planothidiumlanceolatum	100 mg L^{-1}	25	7	275.51	[78]
	Tetraselmischuii	1 ppm	20-22	8	13.46	[76]
	Scenedesmus abundans	1 mg L^{-1}	25	7.8-8	11.5	[79]
	Ulva prolifera	100 µM	20	Na	100.633 ± 15.711 µg/g	[80]
Co	*Chlamydomonas reinhardtii*	15 µM L^{-1}	25	5.5	0.89	[81]
	Chlamydomonas reinhardtii (without cell wall)	15 µM L^{-1}	25	5.5	1.3	[81]
Cr^{3+}	*Spirulina sp.*	0.05-0.5 g/25ml	na	na	304	[82]

	Spirulina sp. (HD-104)	na	na	na	306	[83]
	Scenedesmus quaricauda	100mg L^{-1}	22.3	6	58.47	[84]
Cr^{+6}	*Chlamydomonas reinhardtii*	1000mg/L	22	2	25.6	[85]
	Spirulina sp.	0.05-0.5 g/25ml	na	na	333	[82]
$Cr_2O_7^{-2}$	*Spirulina sp.* (HD-104)	na	na	na	226	[83]
Cu^{2+}	*Anabenacylindrica*	450 µg L^{-1}	23	4.0-5.0	12.62	[86]
	Asterionella Formosa	450 µg/L	23	4.0-5.0	1.1	[86]
	Chlorella vulgaris	5 mg/L	24 ± 2	4.5	63.08	[87]
	Planothidiumlanceolatum	100 mg L^{-1}	25	7	134.42	[78]
	Spirulina platensis	2.55-3.8 mg L^{-1}	34	9	0.85	[88]
	Spirulina sp. (HD-104)	na	na	na	576	[83]
Fe^{3+}	*Chlorella vulgaris*	na	na	2	24.52	[89]
	Microcystis sp.	50 µg ml^{-1}	29 ± 2	9.2	0.03	[90]
Hg^{2+}	*Chlamydomonas reinhardtii*	25-500 mg/L	25	6	106.6	[91]
	Pseudochloro coccumtypicum	0 – 100 µg/L	20 ± 1	7	15.13	[92]
Ni^{2+}	*Chlorella miniate*	10 – 40 µg/ml	24 ± 1	7.4	1.37	[93]
	Chlorella vulgaris	1000 ppm	25	5	15.4	[94]
	Arthrospira platensis	0.5–3.0 mM	30 ± 1	5.0-5.5	20.78	[95]

	Spirulina	0.05-0.5 g/25ml	na	na	1378	[82]
Pb^{2+}	*Chlamydomonas reinhardtii*	25-500 mg/L	25	6	380.7	[91]
	Oscillatoria laete-virens	10 – 100 mg/L	25 ± 2	5	21.6	[96]
	Pseudochlorococcumtypicum	0 – 100 µg L^{-1}	20 ± 1	7	4.49	[92]
	Spirulina platensis	5 – 100 µg L^{-1}	25 ± 1	7	188	[97]
Zn^{2+}	*Cyclotella cryptica*	0.5 mg L^{-1}	25 ± 2	6	242.9	[78]
	Planothidiumlanceolatum	100 mg/L	25	7	118.66	[78]
	Scenedesmus subspicatus	0.5 mg/L	20 ± 2	6	72.06	[98]

Table 2.1 shows that *Chlorella*, *Scenedesmus* and *Spirulina* species are most widely applied for the uptake of metal ions from the liquid phase. Also, metal ion uptake capability is affected by pH of the medium.

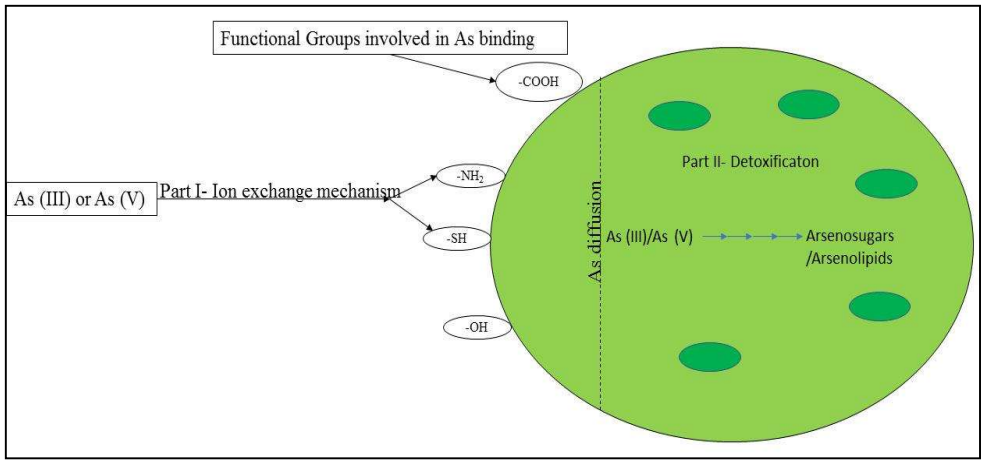

Figure 2.3. Flow diagram of As ions biosorption (Part I) and detoxification of As ions (Part II) by microalgae.

2.2.1. Interaction of HMs with the cell wall of microalgae

Interface of the HMs at the cell wall of microalgae have been proposed as the initial step of metal sorption as metal cation translocates through the microalgal cell wall. Metallic species react with the carbohydrates, lipid and protein present on the external surfaces of the cell wall [73], [99], [100]. During the interaction, HMs form a distinctive coordination complex with the functional groups such as sulphate, carboxyl and amino group of protein and polysaccharide, imidazole of histidine, nitrogen and oxygen of peptide bond. The pH of the medium is a crucial parameter in determining the extent of protonation. Some unprotonated carboxyl oxygen and sulfate also get electrostatically bonded to HMs [100]. Figure 2.3 (Part-I) represents the As uptake by microalgal cells through ion exchange mechanism [101].

2.2.2. Interaction with the plasma membrane

The transport and detoxification of the HMs (Figure 2.3; partII) through plasma membrane is a crucial process. Microalgal cells interact with the external environment through metal transporters. Generally, the transporters are placed into two groups: Group A and Group B [102]. Metal ions are translocated into the cytoplasm through the membrane by Group A transporters. Members of Cu-Transporter (CTR), Natural Resistance Associated Macrophages Proteins (NRAMP), Fe-transporter (FTR) and Zrt-, Irt- like proteins (ZIP), families constitute Group A transporters. This group also includes the assimilative transporters present in the membrane and they increase the concentration of HMs in the cytosol. The membrane of the vacuole also has Group A transporters and perform a similar function as assimilative transporters, but they uptake metal ions sent in within the cell. The function of Group B transporters is to decrease the concentration of metal ions present in the cytoplasm such as providing metal ions for binding to the metal-dependent proteins localized in the cell organelle. This group consists of members from the P1B-type ATPase, Ferro Porti N (FPN), Cation

Diffusion Facilitator (CDF) and Ca(II)-sensitive Cross-Complementer 1 (Ccc1)/ Vacuolar Iron Transporter 1 (VIT 1) families [102].

2.2.3. Physical Adsorption

The physical adsorption process is not dependent upon the metabolism of the microalgal cell, and also it is a reversible process offering several advantages. In this process, polyelectrolytes present on the cell wall bind to the metal ions through electrostatic interaction: Vander Waals forces, redox interactions, covalent bonding, and biomineralization and thus they achieve electroneutrality [103]. Ionic interactions occurring between the cell walls and the metal ions are responsible for the biosorption of cadmium, uranium and zinc [104]. In the same way, copper is physically adsorbed by alga *C. vulgaris* through Vander Waals forces [105].

2.2.4. Role of microalgal organic acids

The microalgal organic acids (e.g., lactic, citric, fumaric, oxalic, malic acids, and gluconic, and) perform two functions: (i) formation of metalloorganic molecules by chelating toxic metal ions and (ii) leaching and solubilization of metal components from the cell surfaces [103].

2.2.5. Precipitation

When the pH of the solution is low, functional sites of the cell wall are blocked by protons. Thus metal ions are restricted from binding due to repulsive forces. With the increase in pH of the medium, the protons are displaced by negative charges from functional sites. This results in an increase in the adsorption of HMs on the functional sites. The decrease in the solubility of metallic ions results in the reduction of their bioavailability. Thereafter, the precipitation of ions takes place subsequently [103]. Cellular metabolism plays a crucial role in precipitation, as it may or may not be dependent on the cellular metabolism: (i) When the precipitation depends upon the metabolism, microbes secrete specific compounds that cause precipitation when they encounter a metal ion and (ii) when the precipitation process is independent of

cellular metabolism, precipitation is due to interaction between the cell surface [106]. It was reported that precipitation of Cd^{2+} took place in the vacuole of *Tetraselmis suecica* [107].

2.2.6. Metallothioneins

Valle and coworkers first characterized metallothioneins in the late 1950s [108]. Metallothioneins are proteinaceous, generally low molecular weight (approx 6-7 kDa), structurally diverse, and cysteine-rich compounds. They form complexes with HMs in the thiol cluster. These peptides are grouped into two categories: (i) Phytochelatins (class III metallothioneins or MtIII) are synthesized by enzymes in the form of short-chain polypeptides and found in certain fungi, algae and higher plants; (ii) MtII (Class II metallothioneins- reported in algae, cyanobacteria and higher plants), and MtI (class I metallothioneins) and observed in *Neurospora* and *Agaricus bisporus* (not identified in algae) found in most vertebrates - both are encoded by genes [109]. Several investigations have been performed by researchers confirmed that class III Mt are synthesized and are present in algae [110]–[112]. Also, *in vitro* studies have reported that a stable complex is formed when HMs binds to long chain MtIII [109], [113]. These molecules reduce the cytosolic free-metal ion concentration by chelating them such as Cd and other metallothioneins were believed to perform Zn and Cu homeostasis [112].

2.2.7. Role of the vacuole in metal sequestration

Chlorella salina was investigated to study the RE of three metal ion (Co, Zn and Mn) and it was detected that higher concentration of HMs was present in the vacuole instead of cytosol [114]. The mechanism of this phenomenon could be (i) HMs regulation within the cytoplasm or (ii) metal ions detoxification [73]. Microscopical and X-ray studies showed that metal ion was complexed with MtIII and then transported into the vacuoles of microalgal cells [109]. Few electron dense materials made up of cadmium and sulphur (in ratios between 2 and 2.4) were observed in the vacuoles of the microalga *Dunaliella bioculata* [115]. Other studies have

also reported the presence of Cd^{2+} in the vacuoles of green alga *Tetraselmis suecica* [107], and diatom *Skeletonema costatum* (with the simultaneous presence of Cu^{2+}) having sulfur to metal ratio of 1.5 [116]. Dark and spherical electron bodies were detected in the vacuoles of the three freshwater microalgal cells [*Scenedesmus quadricauda var quadrispina*, *Pseudochlorococcum typicum* (Chlorophyta) and *Phormidium ambiguum* (Cyanobacterium)] exposed to Pb^{2+} ions [92]. In this phenomenon, metal ions either bind or forms complexes with phytochelatin or forms metallo-iron, metallo-sulfur or metallo-phosphate complex and then transported from cytosol to vacuole. In the vacuole, a high concentration of organic acids is present which releases the metal from complex and peptide for returning back to the cytosol [92].

2.2.8. Role of the chloroplast and mitochondria

When microalgal species [*Scenedesmus quadricauda var quadrispina*, *Pseudochlorococcum typicum*, (Chlorophyta) *and Phormidium ambiguum* (Cyanobacterium),] isolated from freshwater were exposed to Hg^{2+}, Pb^{+2} and Cd^{2+}, it was observed that excessive starch was accumulated in the chloroplast (around the pyrenoids) in aqueous solution [92]. This study showed the possibility of accumulation of heavy metals in other organelles such as mitochondria. In another study, the accumulation of Cd^{2+} inside the chloroplast was also observed in the *Chlamydomonas reinhardtii* [117]. Sequestration of Cd^{2+} in mitochondria and chloroplast may occur due to any of the following processes: (i) Complex formation of MtIII with Cd^{2+} in the cytosol and then transfer of these complexes into the mitochondria and chloroplat; (ii) Cd^{2+} binds to the MtIII synthesized inside the organelle, which translocated as free ions and then forms HMW (High Molecular Weight) compounds; or (iii) Above two processes occurs at same time and MtIII are synthesized in all the three cellular compartments (chloroplast, mitochondria and cytosol) [109], [118]. When *Oocystis nephrocytioides* was grown in medium containing Cu^{2+}, a high concentration of Cu^{2+} was accumulated in the pyrenoids and thylakoids. Localization of Cu^{2+} suggests that its interaction with the ligands is

confined in the chloroplast [119]. In the other way, transportation of Cu^{2+} from the cytosol to the chloroplast can occur by the formation Cu^{2+}- ligand complex [109].

2.3. Predictor variables affecting microalgal growth in wastewater

A brief discussion about the effects of predictor variables on output variables during wastewater treatment by microalgae has been done in the upcoming sections. Such a brief discussion provides insight into the microalgae mediated wastewater treatment process and will assist in analyzing the machine learning results more efficiently.

2.3.1. Microalgae Class

Microalgae are characterised as microscopic photosynthetic organisms. They are placed in the category of thallophytes, i.e., lacking leaves, stem and roots. Their main photosynthetic pigment is chlorophyll a. They exist both in marine and freshwater habitats in the form of single cells, chains and flocs [120]. Their growth is operated by almost similar photosynthetic processes adapted by higher species of the plant kingdom, i.e., converting solar energy and CO_2 into microalgal biomass [121], [122]. They naturally grow as a suspension culture in water and have relatively faster growth rates. Moreover, they are not sensitive to seasonal variation and exhibit high photon-to-biomass conversion efficiency (Microalgae: 3%, Terrestrial Plants: <1%) [123], [124]. Due to high photosynthetic efficiency, microalgae are able to fix more CO_2. Microalgae can fix 1.83 g CO_2 on average, for the production of 1g of dry microalgae biomass [125]. In microalgal cells, there is no requirement of the vascular system for the transportation of nutrients, and each cell is capable of directly uptaking dissolved nutrients [122]. Microalgae can be easily genetically manipulated and exploited in mass culture for the production of biomass and carbon sequestration from the air [126]. They can survive for a prolonged period of time and can acclimatise to varying environmental conditions. This is due to the fact that microalgal cells can form a resting cyst, which remains in a dormant condition during an unfavourable environment [127]. Microalgae biomass is a rich source of carbohydrates, lipids

and proteins that serve as raw material for the production of third-generation biofuel and various high-value products such as vitamins and pigments [128]–[131]. Microalgae have the capacity to produce 200 times more oil/unit in comparison to conventional oil crops such as jatropha [132].

The biological and physiological characteristics of microalgae species determine their suitability for the application in wastewater treatment. Favorable characteristics include rapid adaptation to different types of wastewater and local environmental conditions, fast growth rate, high biomass productivity, and high nutrient removal efficiency. Various microalgae species from different classes, including *Chlorodendrophyceae* [133], *Chlorophyceae* [134], *Cyanophyceae* [135], *Eustigmatophyceae* [31], *Trebouxiophyceae* [136], and *Xanthophyceae* [137], have been used for the treatment of wastewater. *Chlorella vulgaris* has been widely used for wastewater treatment among all species due to its robustness, high growth rate and biomass productivity, and high removal efficiency (RE) of nutrients [138]. Other species reported for high nutrient RE were *Tetradesmus obliquus* [134], *Scenedesmus obliquus* [24], *Tetraselmis sp.* [28], *Tribonema sp.*, *Synechocystis* sp. [135], *Coelastrum* sp. [139], *Nannochloropsis oculate* [31], *Dunaliella* sp. [140] and *Botryococcus braunii* [141]. However, a single strain cannot meet all the mentioned criteria. In such circumstances, strain with a high growth rate and high nutrient RE should be selected [142]. Another approach to tackle this problem is using a consortium of microalgae mainly composed of well-adapted native strains [4].

2.3.2. Wastewater Source

Wastewater from various sources has been used to cultivate microalgae species, as shown in Table 2.2. Wastewater is one of the common wastes produced during the day-to-day activities of humans. The characteristics and composition of wastewater differ depending on the origin of discharge and generation process [4]. Wastewater contains ample organic and inorganic nutrients, including nitrogen, phosphorus, and other micronutrients, favouring microalgae

growth. Composition difference in wastewater influences microalgae's growth, pollutant removal rate and biomass composition of microalgae biomass (lipid, protein and carbohydrate) [143]. Wastewater sources commonly treated by microalgae can be broadly categorized into five categories based upon the source of discharge: (i) domestic, (ii) municipal, (iii) industrial, (iv) livestock, (v) agriculture [144].

Table 2.2. Wastewater sources used for microalgae cultivation with corresponding removal efficiency and biomass concentration.

Microalgae Species	**Wastewater Source**	**Experimental Conditions**	**Removal Efficiency (%)**		**Biomass Conc. (g/L)**	**Reference**
			NRE	**PRE**		
Chlorella vulgaris, Scenedesmus obliquus and Ourococcus multisporus	Municipal wastewater	500 mL serum bottles; IIL: 2%; 15% CO_2; 27 °C; 45-50 µmol photon m^{-2} s^{-1}; 16h/8h; 150 rpm;	>99	>99	0.29-0.31	[145]
Chlorella sorokiniana	Potato processing industry	24 ± 2.7 °C; 6000 lux; 12 h;	82.7	58.0	0.789	[146]
Chlorella sorokiniana	Pig manure	24 ± 2.7 °C; 6000 lux; 12 h;	>95	80.7	0.564	[146]
Microalgae Consortia	carpet industry wastewater + municipal wastewater	6% CO2; 15-25 °C; pH 7; 1 L Erlenmeyer Flasks; IIL: 0.1 g/L; 75-80 µmol photon m^{-2} s^{-1};	>96	>96	1.47	[147]

Chlorella sorokiniana	Municipal Wastewater	1 L Erlenmeyer Flasks; 120 rpm; 120 µmol photon m^{-2} s^{-1}; 16:8h; pH 7;	94.29	83.3	2.275	[148]
Chlorella pyrenoidosa	Starch wastewater + Alcohol wastewater	IIL: 0.5 g/L; pH 6-7; 2 L conical flasks; 127 µmol photon m^{-2} s^{-1}; 12:12 h; 25 °C;	91.64	90.74	3.01	[149]
Chlorella zofingiensis	Piggery wastewater	pH: 6.8; 1.37 L tubular bubble PBR; 5-6% CO_2; 25 ± 1 °C; 230 ± 20 µmol photon m^{-2} s^{-1};	82.70	98.17	2.96	[150]
Chlorella vulgaris UTEX-265	Brewery industry	500 mL Erlenmeyer flasks; 100; 12 h ;150 rpm; Air flow rate: 100 cc/min;	>90	>90	3.20	[151]
Neochloris oleoabundans	Dairy manure	50 L PBR; 200 µmol photon m^{-2} s^{-1}; IIL: 10%; 2-3% CO_2; 14:10 h: 23-25 °C;	90-95		0.704-0.768	[152]
Desmodesmus sp.	Facultative lagoon wastewater	20 L open batch reactor; 12 h;	80	38	0.58	[153]

	treatment plant					
Chlorella sorokiniana	Domestic wastewater	1 L Duran bottles; 0.2 vvm; 30 °C; 80 µmol photon m^{-2} s^{-1};	100		0.32	[154]
Chlamydomonas reinhardtii	Industrial centrate	15 L Biocoil; 25 °C; 220 µmol photon m^{-2} s^{-1}; Air flow rate: 1.8 L/min;	83	14.45	2	[155]
Chlorella pyrenoidosa	Riboflavin manufacturing unit effluent	83-278 µmol photon m^{-2} s^{-1}; 24 h; 28 ± 2 °C;	78.76	94.78	1.25	[156]
Chlamydomonas sp. TAI-2	Science industrial park	6 L glass columni-form flasks; 125 µmol photon m^{-2} s^{-1}; 120 rpm; 5% CO_2; Air flow rate: 165 mL/min;	100	33	1.5	[157]
Chlorella vulgaris	Municipal wastewater	2 L flasks; IIL: 0.1 g/L; 0.5 vvm; 2000-10000 lux;	100		0.832	[158]
Chlorella pyrenoidosa	Soybean processing wastewater	500 mL conical flasks; IIL: 0.3 g/L; 40.5 µmol photon m^{-2} s^{-1};	89.1	70.3	2-2.25	[159]

							14:10 h; 27 ± 1 °C;

Conc.: Concentration; NRE: Nitrogen Removal Efficiency; PRE: Phosphorus Removal Efficiency; IIL: Initial Inoculum Level;

2.3.2.1. Domestic Wastewater

Wastewater discharged from the residences and commercial locations constitute the domestic wastewater (DW). Other sources include public and private institutional facilities. Chemical constituents of DW include both organic (carbohydrates, fat, oil, protein, surfactants, etc.) and inorganic (nitrogen, phosphorus, sulphur, etc.) components [160]. As shown Table 2.2, various microalgal species has been successfully cultivated in domestic waster. Ren et al., (2019) have inspected the effect of ultrasonic treatment for RE of nutrients from non-sterile domestic effluent. Ultrasonic waves reduce the reaction time and improve mass transfer. Authors cultivated *Scenedesmus sp.* in bubble column reactor made up of polymethyl methacrylate (600 ml working volume) using DW as medium for 7 days and ultrasonic waves of different frequencies (ranging from 0 - 30 kHz), power (ranging from 0 - 50 W), time interval (ranging from 0 - 50 min) were exposed on the reactor. Maximum removal of TP and TN reached up to 97.7% and 96.8% respectively at the optimal ultrasonic treatment parameters that were: power - 20 W, frequency - 18 kHz, time - 10 min [161].

In another study, SBBR (Sequencing batch biofilm reactor) was modified by the incorporation of microalgae to construct a algal-bacterial symbiosis (ABS) system for enchancing the RE of nutrients from DW. SBBR was made up of glass (4 L working volume) with ceramic carriers fitted at the upper side of reactor that were favourable for both development of the ABS system and algae enhancement. This ABS system improved the removal efficiency of TN from 38.5% to 65.8% and of TP from 31.9% to 89.3% [162]. Enhancement of nutrient removal efficiencies is also done by using immobilization techniques.

Katam and Bhattacharyya (2019) compared the removal efficiency of two systems: (i) System A- suspended activated sludge and immobilized mixed microalgal culture system ; (ii) System B- suspended co-culture system. They immobilized the microalgal consortium in alginate polymer. System A showed higher nitrogen and phosphorus removal (91% and 93%) than system B (58% and 80%) [163].

2.3.2.2. Municipal Wastewater

Extensive investigation has been done to estimate the potential and prospective application of microalgae for municipal wastewater treatment [164]. The growth of urbanization and the population increase have resulted in heightened municipal wastewater generation. Municipal wastewater contains a varying amount of domestic (80-95%) and industrial (5-20%) influents, which broadly depends upon the local activities. It includes inorganic elements such as ammonia and phosphates which supports microalgae growth, as well as they, contain micronutrients such as coppers, magnesium, required for there growth [148]. Successfully cultivation of *Chlorella sp.* was carried out in wastewater samples obtained from four different locations from the municipal wastewater treatment plant. Authors collected samples: (i) before primary treatment, (ii) after primary treatment, (iii) wastewater from activated sludge chamber, (iv) centrate (wastewater generated during centrifugation of sludge). Removal efficiency (RE) for PO_4^{3-}-P, COD, and NH_4^+-N was obtained in the range of 78.3-82.4%, 83.2-90.6% and 50.9-83.0% respectively [59]. In another study, *Chlorella sorokiniana* was grown in influent coming to the treatment plant and anaerobic tank centrate under both heterotrophic and mixotrophic cultivation mode. Effective NH_4^+-N (94.29%) and PO_4^{3-}-P (83.30%) removal were obtained under mixotrophic condition using anaerobic tank centrate [148]. Zhai et al. (2017) applied the response surface methodology (RSM) technique for the prediction of the optimal conditions to increase the RE of substrate by *Spirulina platensis* from synthetic simulated municipal wastewater and confirmed by conducting laboratory experiment. The optimal

parameters for growth were in the range of the light intensity of 3300-3400 lx , 8.8-8.9 for pH, when the temperature was set at 25 ± 1˚C with aeration rate at 0.5 vvm and the daily illumination time, was set to 12h. The removal efficiency obtained under the optimum conditions was 81.51 % for nitrogen and 80.52 % for phosphorus [165].

Photo-sequencing batch bioreactors were used for the treatment of municipal discharge with mixed co-culture of microalgae and bacteria. The RE of 98 ± 2% for TKN and 87 ± 5% for COD was obtained without forced aeration and studied the combined effects of microbial oxygen consumption and photosynthetic oxygenation [166]. For the reducing the cost of cultivation of microalgae at larger scale, *Chlorella zofingiensis* was cultivated in municipal wastewater by supplementing it with pig biogas slurry. Batch experiments were conducted with different proportions of municipal wastewater and pig biogas slurry. The results exhibited that, when 8% of pig biogas slurry was supplementated in municipal wastewater, it had remarkable effects on microalgae growth with 93% TN and 90% TP removal [167].

2.3.2.3. Agriculture Wastewater

Agricultural wastewater mainly constitutes an excess of nutrients transferred from agricultural lands to the water streams. Excess nutrients are accumulated due to the massive use of fertilizers, pesticides, and insecticides, leading to eutrophication and loss of biodiversity [168]. Khalid et al. (2019) cultivated a native strain *Chlorella sorokiniana* in agriculture effluent with a high concentration of nutrients (C: 2364 mg/L; N: 385 mg/L; P: 106 mg/L). The native strain was well adapted in the effluent with a growth rate and biomass productivity of 0.24/d and 100 mg/L/d and more than 80% of the nutrient was metabolized by strain [169].

2.3.2.4. Industrial Wastewater

Industrial wastewater characteristics and composition vary from one dumping site to other and includes a high concentration of decomposable and non-decomposable inorganic and organic

materials and growth inhibitory constituents depending upon the type of industry [47]. Various types of industrial wastewater, including effluent emanating from food processing unit [170], ethanol biorefinery [171], pyropia processing plant [172], starch processing unit and brewery industry [137], [173], tannery and meat processing unit [28], [174], soybean processing and vinegar production manufacturing division [175], [176], fertilizer production plant [30], palm oil mill effluent [177], textile industry effluent [178], petrochemical industry [179], gourmet powder factory [180] have been used successfully for the cultivation of microalgae coupled with high removal of nutrients and heavy metals. Textile effluent is composed of different organic materials, phosphorus and nitrogen material that can be utilized for microalgae cultivation. Influence of pH and various sources of phosphorus and nitrogen was investigated during the cultivation of *Chlorella sp.* G-23 in varying dilutions of textile wastewater. The highest NH_4^+-N RE wasm78 ± 3% at 0% dilution rate ad 84 ± 4% at 10% dilution rate with aeration at pH 9. For COD, peak RE (> 60%) was obtained at 0% dilution rate, without any effect of aeration. There were no notable effects of the type of nitrogen source on microalgae growth [181]. In another study, textile wastewater was treated by mixed microalgae consortia (*Chlorella* and *Scenedesmus sp.*) in a fed-batch reactor. Fed-batch reactor was operated for five cycles and the duration of the cycles were reduced (30 to 10 days) as the cycles were repeated. This led to the gradual adaption of microalgae in textile wastewater. RE of 70% of total nitrogen and 95% of total phosphorus was obtained throughout the operation [182].

Tannery wastewater is rich in nitrogenous compounds and are also composed of high carbon-based content which supports microalgae growth in both autotrophic and heterotrophic mode. Fontoura et al. (2017) used different concentrations (20-100%) of tannery wastewater for the cultivation of *Scenedesmus sp.* at different light intensities (20-200 μmolphotons $m^{-2}s$) at temperature 25°C and with constant aeration. Maximum RE of COD (80.33%), NH_4^+-N (85.63%), and PO_4^{3-}-P (96.78%) was obtained at a light intensity of 102.5 μmol photons $m^{-2}s$

and wastewater concentration of 88.4% [183]. During the cultivation of purple phototrophic bacteria (PPB) and microalgae in tannery wastewater, microalgae showed better removal efficiency. Both microalgae and PPB were cultivated on five different agro-industrial wastewater (poultry, pork, dairy, red meat and sugar processing industry). PPB showed moderate removal (up to 80% NH_4^+-N, 55% PO_4^{3-}-P, 74% COD), whereas microalgae showed higher RE (up to 91% NH_4^+-N, 73% PO_4^{3-}-P, 91% COD) [184].

2.3.2.5. Livestock Wastewater

Animal feeding operations generate a tremendous amount of dung and manure-contaminated wastewater. Recently from the past decades, livestock processes are carried out at larger scale, thus generating a large amount of effluent [185]. Livestock effluent is often rich in ammonium, organic phosphorus and nitrogen. Thus, providing vital nutrients for supporting microalgae growth [186]–[188]. In a study, *Botryococcus braunii* was cultivated in submerged membrane photobioreactor (SMPBR) and its ability was investigated to conduct tertiary treatment of livestock effluent. Semi-continuous photobioreactor was operated in three phases based on the hydraulic retention time (HRT) (3, 4 and 5 days) to evaluate nutrient removal efficiency. Results showed that shorter HRT (3 days) provided better removal efficiency of TP (85%) and TN (96%) [189]. Five microalgal species *C. vulgaris* (FACHB-1227), *Parachlorella Kesskeri* TY, *S. obliqus* (FACHB-417), *S. quadricauda* (FACHB-1468) and *Chlorococcum sp.* GD were cultivated in cattle far wastewater without dilution. *C. vulgaris* showed highest nutrient removal efficiency which was 98.69%, 81.16%, 83.59%, 85.29% and 62.30% for NO_3^--N, NH_4^+-N, NO_2^--N, TP and COD, respectively [190].

Piggery wastewater (PWW) is categorized as one of the most polluted wastewater due to its high values of BOD and COD because of the presence of high concentration of organic nitrogen and organic matter [191]. The applicability of microalgae for the treatment of PWW can prove to be an economical and effective way for the asssimilation of nutrients, organic

matter, emerging pollutants. A comparitive investigation was performed to evaluate the efficiency of algal-bacterial photobioreactors treating PWW under outdoor and indoor conditions. Four algal-bacterial photobioreactors (each of 3L and without closing lid) were operated under outdoor and indoor conditions for the treatment of diluted (10 and 20 times) PWW for four months and 26 days of HRT. The highest RE for TOC and TP (94 ± 1% and 100% respectively) was observed under indoor conditions for 10 times dilutions, while highest RE for TN (72 ± 1%) was obtained under the outdoor condition for 10 times dilutions. *Chlorella vulgaris* and *Proteobacteria* were the dominant species in the aforesaid removal [192]. For improving the nutrient RE, the blending of PWW is carried out with other wastewaters. *Chlorella vulgaris* was cultivated for 7 days in three different mixtures of PWW with brewery wastewater: (i) PWW and malt processing wastewater, (ii) PWW and brewing (saccharifying and fermenting) processing wastewater, (iii) PWW and packaging processing wastewater. The mixture of PWW and packaging processing wastewater showed the maximum RE for ammonia (100%), TN (96%), TP (90%), TOC (93%) at pH 7.0 and mixing ratio 1:5 [193]. In another study, PWW and winery wastewater were mixed in the ratios of 20:80, 50:50, 80:20, 100:0 and 0:100. Mixtures were then inoculated with the soil microalga, *Chlorella sp.* MM3 and were grown for 10 days. Mixture of 20:80 showed the maximum removal efficiency which was 100% for NH_3-N, 96% for TN, 90% for TP and 93% for COD, thus proving it to be an effective approach for phycoremediation of the mixture of piggery and winery wastewater [194].

2.3.3. Cultivation Type

Microalgae are commonly cultivated in three modes of cultivation based upon the type of energy and carbon source used: (i) autotrophic; (ii) heterotrophic; and (iii) mixotrophic. Among all three, autotrophic is the most common mode of cultivation. But heterotrophic mode favors high biomass productivity than autotrophic mode. There is no light dependency in

heterotrophic mode, and the growth rate remains constant as carbon and energy source remain available [195]. Biomass productivity of *Chlorella vulgaris*, *Chlorella protothecoides,* and *Chlorella sorokiniana* was 4.8 times, 3.4 times, and 3.3 times, respectively, higher when cultivated in heterotrophic mode than autotrophic mode [196]–[198]. Moreover, no significant difference was observed in nutrient RE when *Coelastrum* sp. was cultivated in both autotrophic and heterotrophic mode [139]. Thus, the heterotrophic mode of microalgae cultivation is indeed a better choice than the autotrophic mode due to its simplicity, cost-effectiveness, and less maintenance requirement [199].

Compared to autotrophic and heterotrophic modes, biomass productivity and nutrient RE are higher in the mixotrophic mode due to its light independency [200]–[202]. Mixotrophs utilize both inorganic and organic carbon sources and prevent the photo inhibitory effects of high light intensities [203], [204].

2.3.4. Culture Type

As mentioned earlier, microalgae are either cultivated in monoculture mode or in co-culture mode with other microalgae species, bacteria, and fungi [205]. Research on the application of microalgae-bacterial consortia for wastewater treatment is increasing rapidly [206]. Microalgae consume a large portion of inorganic nutrients, including nitrogen and phosphorus, while bacteria metabolize the organic nutrients present in the medium. Furthermore, microalgae and bacteria mutually support each other, as CO_2 generated by bacterial respiration is utilized by microalgae during the photosynthesis process while releasing O_2, which in turn used by bacteria [207]. Thus, microalgae can replace the mechanical aeration device utilized during activated sludge, decreasing energy requirement and operation cost significantly [208].

2.3.5. Pretreatment of wastewater

Wastewater contains various types of particulate matter, toxins, bacterial and fungal contamination and possesses high turbidity. In addition, foreign microbes compete with microalgae for nutrients in wastewater, while particulate matter and turbidity hinder the proper utilization of light. Therefore, pretreatment of wastewater is needed to reduce the population of undesirable microbes and particulate matter. Pre-treatment methods applied before microalgae cultivation are primary treatment methods, aerobic and anaerobic treatment, autoclaving, dilution, acidogenic fermentation, magnetic treatment, ozonation, and chlorination. Primary treatment method includes gravity settling, centrifugation and filtration [209].

In most cases, it has been reported that the primary pretreatment method is sufficient to make wastewater a suitable medium for microalgae cultivation. Anaerobic and aerobic pretreatment methods are applied when organic content in the wastewater is high. Anaerobic pretreatment generates biogas or other products having a high value, while aerobic pretreatment generates a high amount of carbon dioxide [170], [210]. Autoclaving has been suggested as the most effective method for decreasing microbial load in order to carry out the study on specific microalgae strains. However, autoclaving may decrease nutrient concentration and also requires a high amount of energy [148], [174]. Ozonation and chlorination have also been considered for the pretreatment of wastewater [171], [174]. Ozone pretreatment not only reduces microbial load but also increases the biodegradability of the wastewater by microalgae. Ozone pretreatment also decreases the turbidity of the wastewater, enhancing the penetration of light [171]. Thus, ozonation can be a suitable method for treating wastewater at large-scale microalgae cultivation [211].

2.3.6. Reactor Type

Microalgae cultivation is generally carried out either in an open system (raceway ponds) or a closed system (flasks, PBR) [212]. The main advantage of ponds is their simplicity in construction and operation, which leads to low operation and production cost [213]. However, they are highly vulnerable to bacterial and fungal contamination, thus not suitable for monoculture cultivation [213]. PBRs are illuminated culture vessels that offer better control over predictor parameters such as CO_2 content, temperature, pH, light intensity, etc., and permits high biomass concentration [212]. Various types of PBR have been used for microalgae cultivation that includes flasks (for low volume cultivation) [136], fluidized bed reactor [30], flat-panel PBR [214], tubular PBR [170], helical type PBR [212], stirred tank reactor [205]. Some of the disadvantages of PBRs are as follows: (a) it does not permit its application for large scale microalgae cultivation, especially for low-cost products; (b) it has a complex construction design with high operation-cum-maintenance cost [212].

2.3.7. CO_2 content

All microalgae and cyanobacterial species can assimilate inorganic carbon either in the form of CO_2 from the atmosphere or flue gas emission or soluble carbonates [205], [215]. There is an optimum level of CO_2 that should be used for the aeration of microalgae culture. Any deviation from the optimum level has adverse effects on microalgae growth [216]. When the concentration of CO_2 is increased above 5%(v/v), it results in the lower activity of carbonic anhydrase, decreasing its affinity towards CO_2 while increasing its sensitivity towards O_2 [217], [218]. Goncalves et al. (2016) cultivated *Chlorella vulgaris* and *Pseudokirchneriella subcapitata* at different concentrations of CO_2 (Air, 3, 5, 7, 9, 10 % (v/v)) using synthetic simulated wastewater. Both biomass concentration and nutrient RE increased up to 5% CO_2 (v/v), but after that, they start decreasing. Optimum CO_2 (v/v) concentration was found to be

5.35% for both species and almost 100% RE was obtained at that level (Gonçalves et al., 2016b).

2.3.8. Temperature

The optimum temperature level for most microalgae strains is around 25-30°C [136], but some strains can even survive up to 35°C [134] or below 15°C [220]. At high temperature, the algal photorespiration rate increases, while at low temperature, both photosynthesis and growth rate decrease [221], [222]. Temperature also affects intracellular enzyme activity and solubility of gases in the medium [223], [224]. Therefore, heating or cooling devices are needed for maintaining the optimum level of temperature. Chu et al. (2015) cultivated *Chlorella pyrenoidosa* in starch processing wastewater in outdoor conditions and predicted the influence of seasonal temperature variations. The temperature in summer (30.2 °C) provided the highest nutrient RE with a maximum biomass concentration of 2.26 g/l, while nutrient RE and biomass concentration (1.16 g/l) was lowest during the winter (9.3 °C) [220].

2.3.9. pH

Every microalga species has its optimum pH range for growth. However, most of them are active around a neutral pH range (6.5-7.5). pH affects the solubility of CO_2 and other nutrients as well as their form and transformation in the medium [225]. At low pH, microalgae directly uptake CO_2 through diffusion. At high pH main form of inorganic carbon in the medium is bicarbonate, and microalgae consume it through the activity of the carbonic anhydrase enzyme [57]. Various studies have been conducted that reported the effect of pH on biomass productivity and nutrient RE either by using chemical buffers or purging CO_2 at different concentrations [226]–[228].

2.3.10. Initial Inoculum Level

The initial inoculum level is considered as one of the significant factors during the initial stage of culture. The high inoculum level readily acclimatizes microalgae species in adverse conditions of wastewater, such as the presence of toxic chemicals. Thus, a high inoculum level supports high growth rate and high RE within less time as the lag phase duration is decreased [229]. However, it should be noted that at very high inoculum density, mutual shading of cells might reduce the penetration of light [230]. A high inoculum level of 30% (v/v) supported the acclimation of *Chlorella sorokiniana* during cultivation in tannery effluent containing toxic pollutants and relatively high biomass concentration (1.06 g/l), and increased nutrient RE was achieved in comparison to inoculum level above 30%(v/v) or below it [31].

2.3.11. Light Intensity & Photoperiod

The luminous intensity plays a vital role in biomass productivity and nutrient removal as it supplies the energy required for microalgae growth during autotrophic and mixotrophic cultivation. It promotes the conversion of inorganic nutrients present in the medium to organic microalgae biomass [231]. Low light intensity leads to low biomass productivity, whereas high light intensity causes photoinhibition or photo-oxidation due to hydrogen peroxide formation [232], [233]. Therefore, microalgae species require optimum light intensity to achieve maximum biomass productivity and high nutrient RE [234]. Chu et al. (2015) reported that high biomass productivity and nutrient RE was obtained during the cultivation of *Chlorella pyrenoidosa* in high intensity during the summer. In contrast, low biomass productivity and nutrient RE was obtained in low light intensity during winter [220].

Light intensity and photoperiod affect the microalgae growth interactively. Biomass productivity increases by increasing light intensity and photoperiod, but the high light intensity combined with longer photoperiod caused photoinhibition [235].

2.3.12. Initial Nutrient Concentration

The optimum level of N/P ratio is required for the high biomass productivity of microalgae. However, a high level of nutrients, especially ammonia, can be toxic to microalgae growth. Therefore, proper dilution is needed for optimizing the N/P ratio if the initial concentration of nutrients in the wastewater medium is high. Ge et al. (2016) cultivated *C. vulgaris* at different nutrient concentrations of centrate wastewater and observed that biomass productivity and nutrient RE increased with increasing the nutrient concentration. But after achieving maximum value, both started decreasing on further increase in nutrient concentration [236]. The choice of organic carbon source is also an essential factor during heterotrophic and mixotrophic cultivation mode. It has been reported that glucose and acetate facilitate high biomass productivity during mixotrophic cultivation [200].

2.4. Machine Learning

Machine learning is one of the subsets of Artificial Intelligence (AI), that was coined by Arthur Samuel in 1959 [237]. Machine learning is the combination of statistical tools and scientific knowledge that can analyse large dataset and generate new patterns/information in the dataset for the classification of new data [238], [239]. Machine learning algorithms construct mathematical models using training data (usually 80% of the constructed dataset), through which computers can predict new data without the need of comprehensive programming. There is no need to write code, generated mathematical models can predict new data itself [237]. They are widely used in various fields due to its high speed, high precision, automation, convenient extensibility and customization [240]. Machine learning algorithms can be broadly classified under three categories: (i) Supervised learning; (ii) Unsupervised learning; (iii) Reinforcement learning. In supervised learning, constructed labelled dataset is supplied to machine for training. Based upon the training, mathematical model is constructed which predicts new data.

In unsupervised learning, unlabelled and unsegregated dataset is supplied to the machine. Machine performs learning without any prior experience or supervision. In reinforcement learning, training of machine is performed in feedback-based mechanism, rewarding the desired move and penalizing the undesired move [237].

In the present research, decision tree (supervised learning algorithm) and association rule mining (unsupervised learning algorithm) was used for the optimization of microalgae-based wastewater treatment process.

2.4.1. Decision Tree

Decision tree, one of the popular supervised learning algorithms, can be used for both classification and regression problems. It has a tree like structure, where root node serve as the starting point and further growth of the tree is governed by if-then rule [241]. Decision tree generally has two nodes; (i) inner node; (ii) leaf node. Inner node, also known as decision node, used for making decision as per the analysis of dataset and has various branches. Second one is leaf node, also known as output node, provides the output but is has no further branches. In simple terms, decision tree splits the dataset on the basis of "yes or no" answer to the question [242]. Basically, there are two types of decision tree: (i) Categorical Decision tree: target variable is categorical; (ii) Continuous/Regression Decision Tree: target variable is continuous. In the present study, the classification Decision tree algorithm was used to analyse the dataset. The classification algorithm predicts or classifies the data into different classes of output variables. It cannot predict the numerical value of the output variable. The regression algorithms are used for the prediction of numerical values of the output variable. Common terminologies used in decision tree analysis are [237]:

(i) **Root Node:** Root Node serve as the starting point of the decision tree. From this point whole dataset splits into possible homogenous subsets.

(ii) **Leaf Node:** Leaf Node represents the final outcome of the tree and further no branching occurs from this point.

(iii) **Splitting:** It entails the process of further subdividing the main node into sub-nodes based on the supplied constraints.

(iv) **Sub Tree:** Splitting process results into further branching termed as sub tree.

(v) **Pruning:** In the pruning process, decision tree is optimized by eliminating unnecessary branching, without compromising the accuracy of the tree [243].

2.4.1.1. Attribute selection measures

ASM (Attribute selection measure) is the process of choosing the best spilling criterion that divides the dataset. The most two popular decision tree algorithms are Iterative Dichotomiser 3 (ID3) and Classification and Regression Tress (CART). ID3 uses Information gain spilt criterion while CART used Gini Index spilt criterion [243].

(i) **Information Gain:** This spilt criterion determines the information provided by the feature/predictor variable regarding the class. Selection of root node, spitting of the node and tree construction is done on the basis of information gain values. Information gain is calculated via Eq. (2.4);

$$\boldsymbol{Information\ Gain = Entropy(s) - [(Weighted\ Avg.) * Entropy(Every\ Feature)]} \quad (2.4)$$

Entropy: Entropy measures the amount of randomness presented in the dataset as follows:

$$\boldsymbol{Entropy(s) = -P(yes)log_2P(yes) - P(no)log_2P(no)} \quad (2.5)$$

where, S indicates the 'sample number', P(yes) indicates probability of getting yes and P(no) indicates probability of getting no.

(ii) Gini Index: In the present study, Machine learning used the CART (Classification and Regression Tree) algorithm for constructing Decision Tree. The CART algorithm uses the Gini index criterion to split the data and create different nodes. Gini index proceeds

by measuring the purity or impurity present in the dataset. Attributes/input variables with small Gini index are preferred the algorithm over attributes with large Gini index during decision making process [243].

Gini index is calculated by Eq. (2.6):

$$\boldsymbol{Gini\ index = 1 - \sum_{i=1}^{j} p_i^2} \tag{2.6}$$

where, j is the total number of classes present in the output variable, and p is the probability of selecting the data point of class i.

2.4.2. Association Rule Mining

Association rule mining is one of the famous data mining algorithms that detects association relationship between data points and valuable patterns in the dataset [244], [245]. The basket analysis problem, which helps to determine which things clients frequently purchase together and subsequently suggests a better layout for the business, was the initial driving force behind association rule mining [246]. Recently, association rule mining has been in the field of bioinformatics, healthcare, classification problems and many more. Patterns detected by association rule mining are concluded on the basis of support, confidence and lift values as explained in more detail in Chapter 5.