Chapter 1

Introduction

1. Introduction

Rapid industrialisation, urbanisation and globalisation in recent decades have led to an increase in the release of greenhouse gases and wastewater around the globe [1]. Many countries are in the race of strengthening their economy and overlooking these problems. Human activities globally generate 380 billion m^3 of wastewater each year. With the current rate of generation, wastewater production may increase by 24% in 2030 and 51% in 2050 [2]. Globally, 80% of generated wastewater is released into water bodies without proper treatment [3]. Low-income countries treat only 8% of the generated wastewater, lower-middle-income countries treat 28%, upper-middle-income countries treat 38%, and higher-income countries treat more than 70% of the generated wastewater [3]. Excessive discharge of wastewater in water bodies increases the nutrient level, causing eutrophication, a potential threat to the natural ecosystem [4]. Also, eutrophication possesses a threat to the economy by affecting fishing and real estate activities. Approximately, a loss of 2 billion dollars per year occurs due to eutrophication [5]. Currently, conventional activated sludge (CAS) treatment is used for wastewater treatment. CAS treatment can efficiently treat organic content from wastewater, but an additional treatment process is required for the remediation of inorganic nutrients. Such as, nitrogen is treated via an anammox process and phosphorus via chemical precipitation [6], [7]. But CAS with additional processes is expensive and require intensive energy, as it consumes 3% of the total electricity produced worldwide. The consumption of electricity increases with the increase in the volume of wastewater [8]. Also, wastewater treatment plants generate a large amount of greenhouse gases, that accounts for approximately 1.3% of total global greenhouse emission [9]. Report of IEA fuel combustion 2019 highlights that 2.2, 4.8 and 9.8 Metric gigatons of CO_2 was emitted by India, the United States and China, respectively [10]. According to the Intergovernmental Panel on Climate Change (IPCC, 2018), an increase in global temperature of approximately 0.8ºC-1.2ºC above the set level for industries has been noticed [11]. During the Paris Agreement on Climate Change, it was decided that countries should strengthen their efforts to reduce the emission of greenhouse gases and limit the temperature increase below 1.5°C [11]. However, according to the analysis of the United Nations Environment Programme (UNEP), current efforts will only lead to a decrease in global temperature by 0.01°C by 2050. So, accordingly, the temperature will rise by 3.2°C by the end of the century leading to global warming and a rise in the water level of oceans [12]. Other models also predicted that there is only a 50% chance that global temperature will only increase by 1.5°C [13], [14]. Therefore, it is necessary to develop sustainable technologies that can remediate wastewater with simultaneous sequestration of greenhouse gases and recovery of resources.

The research in the field of development of technologies that can minimise carbon emissions and recover resources from wastewater treatment plants has increased at a rapid rate. There has been progress in the development of various technologies such as biochar production, microbial electrolytic carbon capture, microbial electrosynthesis, constructed wetland and microalgae-based wastewater treatment process [15]. The biological wastewater treatment mediated by microalgae is a sustainable approach for treating wastewater with simultaneous recovery of nutrients in the form of microalgae biomass [1]. They can uptake nutrients at a higher rate from wastewater and with a high fixing potential of CO_2 through photosynthesis. Some of them can even grow heterotrophically in the absence of light, utilising organic carbon from effluents. Some of the essential nutrients including carbon, nitrogen and phosphorus are assimilated by microalgae from wastewater for the synthesis of valuable biomass [16], [17]. Advantages that microalgae offer when applied for wastewater treatment are: (i) the microbial biomass produced can be used for the production of protein supplements (single cell protein), animal feed and biofuel, (ii) the production of fertilisers from microbial biomass and (iii) evolution of oxygen during photosynthesis [18], [19].

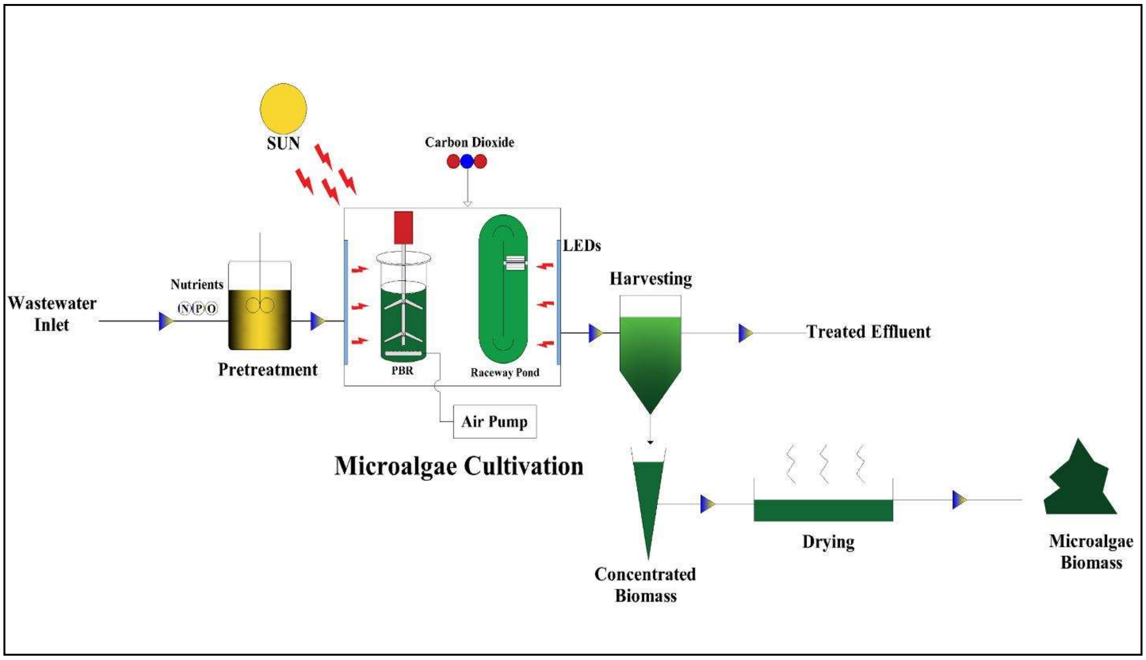

Figure 1.1. Process flow diagram depicting wastewater treatment by microalgae and simultaneous biomass production.

Figure 1.1 depicts an illustration of an integrated microalgae-based wastewater treatment process. Figure 1.1 demonstrates that microalgae are grown by utilising the nutrients from wastewater and CO_2 emerging out from industrial combustion or during respiration. A wide variety of microalgae species have been used to treat different kinds of wastewater (WW) emerging from domestic, municipal, industrial, agricultural and livestock sources, through biosorption, biodegradation and bioaccumulation processes [20]. Microalgae growth in wastewater is highly influenced by various parameters and environmental conditions such as CO_2 content in the inlet air, temperature, initial inoculum level, pH, light intensity, photoperiod, nutrient concentration in wastewater, etc. (Figure 1.1). These parameters are together termed predictor variables [21]. These variables show mutual interaction with each other and cooperatively affect the treatment process. In order to enhance the treatment capability and biomass productivity of microalgae, it becomes essential to optimize these variables and provide the right combination [22]. However, it is practically impossible to test all combinations of variables in a single study. Numerous studies have been published in the literature that attempted to optimise predictor variables for maximizing biomass productivity and wastewater treatment capability using traditional statistical techniques such as response surface methodology (RSM) [23]–[31]. However, model development from RSM requires the several experimental trials and the models only work for a narrow range of input parameters [32].

Instead of an experimental approach, computational techniques such as Machine Learning (ML) presents an inexpensive method of optimisation. ML algorithms can quickly analyse an extensive dataset published in the literature and determine the best combinations of predictor/predictor variables for increasing biomass productivity and wastewater treatment capability of microalgae. ML algorithms play a vital role in analysing large datasets by detecting significant patterns in the data in a minimal time [33]. Machine learning techniques are becoming an internal part of biotechnology research fields, including biochemical engineering [34], healthcare [35], agriculture [36], genomics [37] and environment [38]. In biotechnology, a massive amount of data is present, which cannot be analysed merely through physical modelling. There has been a gradual shift towards application of machine learning (ML) algorithms and adaptive data-driven modelling. ML algorithms can easily detect complex relationships in a large dataset [39]–[41]. This attractive application of ML has also fascinated researchers in the field of microalgae. In the Web of Science core collection group, 2705 publications have reported microalgae biomass production in wastewater, generating a large dataset (Figure 1.2). This dataset can be easily analysed using ML algorithms such as decision trees for developing different optimised models. Generated models provide multiple combinations of input parameters, leading to increased microalgae biomass productivity in wastewater. This will assist in the large-scale production of microalgae biomass, as wastewater offers a cost-efficient medium for the growth of microalgae.

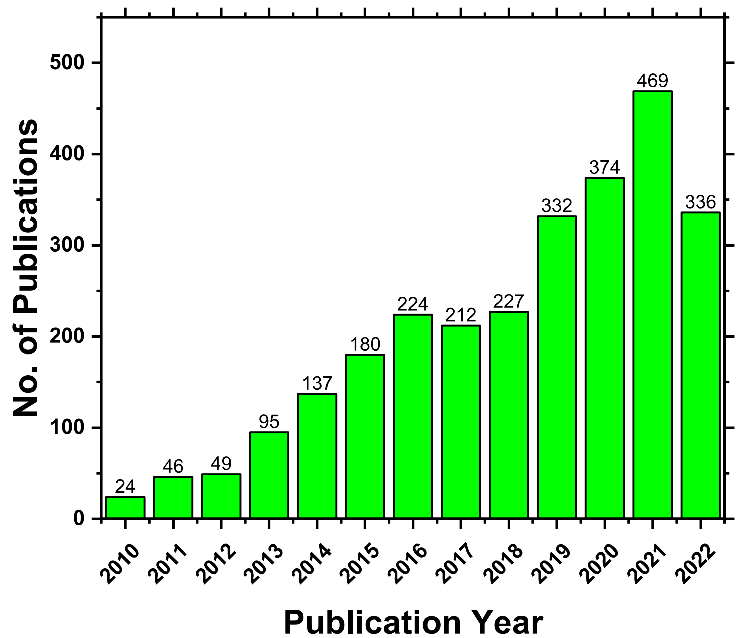

Figure 1.2. Chart representing the number of publications and years that appeared in search of "Web of Science Core Collection" for the keyword "microalgae", "wastewater" and "biomass".

Various authors have reported successful application of ML algorithms in the field of microalgae cultivation, such as predicting critical factors for microalgae biomass and lipid production [42], modelling of *Chlorella* sp. harvesting through flocculation [43], determining the ratio of different microalgae species in microalgae consortia [44], optimising the growth of microalgae in raceway pond [45]. Cosgun et al. (2021) optimised microalgae growth and predicted different combinations of variables leading to high biomass and lipid productivity using a Decision Tree [42]. However, a study that deals with the application of ML algorithms for analysing large datasets in the field of wastewater treatment by microalgae with simultaneous biomass production are not available. ML algorithms can quickly process large datasets and detect different patterns of predictor variables for increasing the nutrient removal capability of microalgae, including nitrogen and phosphorus, simultaneously increasing biomass productivity. Moreover, these algorithms can even generate new information from already reported data which is expected from new experiments. Patterns detected from these algorithms can assist in constructing high throughput experimental designs and assist in carrying out efficient wastewater treatment at a larger scale. With the prediction of different combinations of input variables, ML also provides various other functions that can be used to analyse the effects of predictor variables on microalgae more efficiently, such as the "predictorImportance" function and partial dependence plot (PDP). The predictor importance function predicts each input parameter's relative importance or significance in influencing microalgae biomass production. In contrast, the Partial Dependence function determines the functional relationship between the input parameters and biomass production [46].

Based on the discussed hypothesis, the current research was started first with data collection from a recently published research article (2010-2023). Thereafter, one of the ML algorithms, a decision tree was used for predicting the best combination of predictor variables leading to increase wastewater treatment capability and biomass production, which was applicable to all microalgae species. But different microalgae have different tolerance to the loading of predictor variables. Therefore, in the next objective, different datasets were prepared for two microalgae classes, namely, *Trebouxiophyceae* and *Chlorophyceae* were prepared. Individual datasets were again analysed using a Decision Tree and the best combination of predictor variables leading to increase wastewater treatment capability and biomass production for both classes were determined. In the third objective, another machine learning algorithm, association rule mining (ARM) was used for data analysis. ARM algorithm determined specific condition or a range of individual predictor variables that increases the growth of microalgae in wastewater. Rules or results derived from the Decision tree and ARM were then verified both computationally and experimentally in the fourth objective. At last, in the fifth objective, cost-benefit analysis (CBA) was performed during the application of artificial lights for microalgae cultivation.