



# Software sensors in the monitoring of microalgae cultivations

Luis Porras Reyes · Ivo Havlik · Sascha Beutel 

Received: 5 September 2023 / Accepted: 3 December 2023 / Published online: 10 January 2024  
© The Author(s) 2024

**Abstract** Microalgae are well-known photosynthetic microorganisms used as cell factories for the production of relevant biotechnological compounds. Despite the outstanding characteristics attributed to microalgae, their industrial-scale production still struggles with scale-up problems and economic feasibility. One important bottleneck is the lack of suitable online sensors for the reliable monitoring of biological parameters, mostly concentrations of intracellular components, in microalgae bioprocesses. Software sensors provide an approach to improving the monitoring of those process parameters that are difficult to quantify directly and are therefore only indirectly accessible. Their use aims to improve the productivity of microalgal bioprocesses through better monitoring, control and automation, according to the current demands of Industry 4.0. In this review, a description of the microalgae components of interest as candidates for monitoring in a cultivation, an overview of software sensors, some of the available approaches and tools, and the current state-of-the-art of the design and use of software sensors in microalgae

cultivation are presented. The latter is grouped on the basis of measurement methods used as software sensor inputs, employing either optical or non-optical techniques, or a combination of both. Some examples of software sensor design using simulated process data are also given, grouped according to their design, either as model-driven or data-driven estimators.

**Keywords** Microalgae cultivation monitoring · Software sensors · Data-driven estimators · Model-driven estimators · Optical methods · Non-optical methods

## 1 Introduction

Algae comprise a diverse group of photosynthetic organisms, unicellular (microalgae) and multicellular (macroalgae), commonly inhabiting terrestrial (e.g., deserts, stones, animals, plants) and aquatic environments (e.g., ponds, seawater), but also at sites where high concentrations of salts, long, persistent snow, and volcanic materials conform the environment (Bold and Wynne 1985; Avagyan 2018). Microalgae comprise an important group of microscopic photosynthetic eukaryotes and photosynthetic bacteria, e.g., cyanobacteria (formerly known as blue-green algae) (Andersen 2013; García et al. 2017). The capacity of microalgae to adapt and live under wide-ranging environmental conditions also offers interesting

---

L. Porras Reyes · I. Havlik · S. Beutel (✉)  
Institute of Technical Chemistry, Leibniz University  
of Hanover, 30167 Hanover, Germany  
e-mail: Beutel@iftc.uni-hannover.de

L. Porras Reyes  
e-mail: porras@iftc.uni-hannover.de

I. Havlik  
e-mail: havlik@iftc.uni-hannover.de

biotechnological alternatives for the energy, food, and pharmaceutical sectors.

Photosynthesis implies an attractive process for the development of biotechnological processes: there is an inexhaustible supply of energy (light) with which CO<sub>2</sub> capture, production of different biomolecules and biomass are performed (Masojídek et al. 2013). Unlike plants, microalgae have higher photosynthetic rates, do not require land used for agriculture, consume less water for growth, grow fast and are metabolically more flexible (Benedetti et al. 2018; Vecchi et al. 2020). Microalgae are well-known photosynthetic microbes used as cell factories for the production of relevant biotechnological compounds (e.g., pigments, lipids, carbohydrates, proteins, and vitamins) (Jacob-Lopes et al. 2020), as food sources (e.g., *Chlorella* sp. and *Arthrospira* sp.) or in environmental applications, e.g., wastewater treatment (Martínez-Roldán and Cañizares-Villanueva 2020; Tan et al. 2020). Traditionally, microalgae cultivation involves two major production systems: open (e.g., raceway ponds, circular ponds) and closed systems (e.g., tubular, flat panel, bubble column photobioreactors), however, recent research has also focused on biofilm-based cultivation (Fabris et al. 2020). The description of these systems is not the subject of this article; however, available scientific literature can be reviewed elsewhere (Borowitzka and Moheimani 2013; Zittelli et al. 2013; Tan et al. 2020; Chanquia et al. 2021).

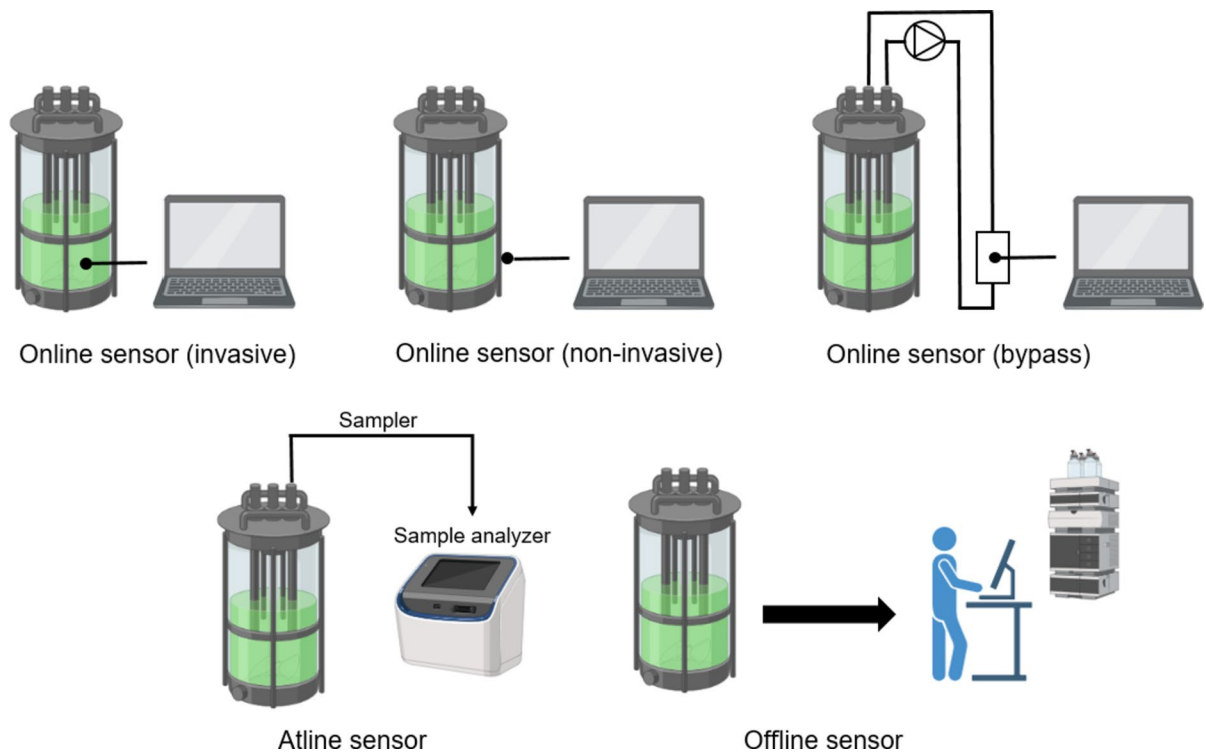
In a bioprocess, the molecules to be produced (e.g., biomass in the form of pure cultures or co-cultures, enzymes, secondary metabolites, recombinant products), the production strategy.

(e.g., batch, fed-batch, continuous system), as well as the type of organism employed (e.g., mammalian cells, yeast, bacteria, microalgae) define the parameters or variables to be monitored. These process parameters to monitor can be classified into three groups: physical, chemical and biological (Biechele et al. 2015).

In microalgae cultivations physical variables include temperature, light intensity, mixing intensity and light frequency within the culture. Chemical variables include pH, dissolved CO<sub>2</sub> and oxygen, extracellular products, intracellular products (e.g., lipids, carbohydrates, pigments, proteins, and vitamins), nitrogen, and phosphorus concentration (e.g., NO<sub>3</sub><sup>−</sup>, NH<sub>4</sub><sup>+</sup>, PO<sub>4</sub><sup>3−</sup>). Biological variables involve biomass concentration, cell morphology and viability,

photosynthetic efficiency or the presence of contaminants (Havlik et al. 2013a, 2022; Claßen et al. 2017). It is evident that, unlike other cell cultures, some of the parameters to be monitored in microalgae cultures are particularly distinctive (e.g., light monitoring given the autotrophic capacity of microalgae (Masojídek et al. 2009), the monitoring of intracellular products, or the presence of contaminants, e.g., rotifers or undesirable algae strains (Deore et al. 2020; Sauer et al. 2021)). Physical and chemical variables and light intensity can be easily measured using classical physicochemical sensors employed in chemical processes. On the other hand, biological variables can usually be measured only by applying sampling and tedious offline methods. Indirect online measuring techniques called software sensors, which are reviewed here, offer an alternative.

There are different terms related to process monitoring that are based on the way the sampling is done and the location of the sensor within the process (Havlik et al. 2022). It is common to find terms such as offline, online, inline, atline and in situ in the literature (Claßen et al. 2017; Reyes et al. 2022). It should be noted that there are slight differences between online and offline definitions found in the literature. Nevertheless, a broader classification is based on the terms offline, online and atline (see Fig. 1). When a sample is measured offline, it means that it is completely removed from the production system for further analysis in a separate laboratory. Offline measurements are laborious, not immediate, require qualified manpower, and are prone to an increased risk of contamination (Claßen et al. 2017; Noll and Henkel 2020; Rösner et al. 2022). On the other hand, online measurements refer to those measurements made by a sensor that are either in direct contact with the contents inside the reactor (invasive), the sensor is separated by the reactor walls (non-invasive), or in a bypass arrangement where the reactor contents are diverted, measured in a flow-through mode, and returned to the cultivation stream (Rösner et al. 2022). Online measurements are performed continuously (Reyes et al. 2022) and imply that the data is taken in real time or almost real time. In atline measurements, samples are continuously removed from the process through some device and analyzed in an analysis system (e.g., HPLC, GC, biochemical analyzer, MS) spatially close to the process site. Measurements are performed at defined time intervals, and a time



**Fig. 1** Types of sensor systems: Online, atline and offline. The online sensors can be arranged in different configurations: invasive (inside the vessel and in contact with the contents), non-invasive (outside the vessel but close to the vessel

surface), and bypass (the sample is routed, measured in a flow-through mode, and returned to the main stream). Partially created in Biorender.com

gap occurs between the time of sampling and sample analysis (Claßen et al. 2017).

Microalgae cultivation systems require advanced and efficient methods to assure improvements in productivity, control and automation processes. Industry 4.0, also known as the "Fourth Industrial Revolution" or the "Industry Internet of Things" (Turcu and Turcu 2018; Richter et al. 2023), i.e., an advanced manufacturing approach based on inter-machine communication technologies in which automation, sensors and machine learning converge to create self-adaptive manufacturing processes that adjust in real time to changes in the process itself, offers possibilities to pursue in microalgae cultures or biorefineries (Fabris et al. 2020). In this context, the application of software sensors, i.e., estimators based on mathematical and/or data-driven process models, using physical signal inputs from multiple sensors could complement the online sensors as it allows the estimation of the desired biological process variables in real or

almost real time, which are often hidden in the measured data and are only indirectly accessible.

In this review, an overview of those microalgae components derived from their cultivation that have a real or potential application is presented: biomass and its various biochemical components (e.g., lipids, pigments, vitamins, proteins, and carbohydrates). In addition, in some cases, their spectroscopic characteristics, i.e., maximum absorptions and fluorescence excitation/emission regions, used in conventional analytical methods or in the implementation of software sensors for their monitoring, are mentioned. Subsequently, the three different approaches used (model-driven, data-driven or hybrid model) in software modeling-assisted monitoring and control, i.e., software sensors, are described, as well as the main tools within each of these categories. The fourth section includes a list of the latest publications describing the design and use of different types of software sensors for the measurement of biological parameters

in microalgae cultures. These were grouped according to physical measurement methods employing optical techniques, non-optical techniques, or a combination of both. Another category comprises software sensors designed with the help of simulated input data, grouped on the basis of whether they are using model-driven or data-driven estimators.

The aim of this review is to present an overview of recent applications of the indirect measurement technique, usually called "software sensors", to microalgae cultivation. This technique, or tool, provides us with the possibility to estimate in almost real-time the values of process variables, as a rule biological ones, that are not accessible for sensors based on physicochemical principles. Application of software sensors to microalgal processes is now in the early stages and is being investigated under laboratory conditions, and applications are, up to this day, relatively few compared to the number of applications in other engineering branches (Perera et al. 2023). This review seeks to summarize the current state of affairs in the field of software sensor applications to microalgae cultivation, concentrating not on the development of general software sensor algorithms or their application in process control but on the application of these algorithms in the estimation of process parameters, thus supplementing the existing physicochemical sensors.

## 2 Biochemical process variables in microalgae cultivation

### 2.1 Biomass as a key parameter in microalgae cultivation

Biomass is the most relevant variable estimated in microalgae cultivation because almost all products obtained by microalgae cultivation are of intracellular nature and thus procured by biomass processing (Jacob-Lopes et al. 2020). Microalgal biomass is known to be a rich source of lipids, pigments, vitamins, amino acids, proteins, carbohydrates, e.g., polysaccharides and oligosaccharides, and essential minerals (Caporgno and Mathys 2018; Jacob-Lopes et al. 2020). Currently, the U.S. Food and Drug Administration (FDA) designates the microalgae *Arthrospira platensis*, *Auxenochlorella protothecoides*, *Chlorella vulgaris*, *Chlamydomonas reinhardtii*, *Dunaliella bardawil*, *Euglena gracilis*, *Haematococcus*

*pluvialis*, *Schizochytrium*, *Porphyridium cruentum*, and *Cryptocodinium cohnii* with a Generally Recognized as Safe (GRAS) status (García et al. 2017; Diaz et al. 2023). It is nowadays common to find commercial presentations of microalgal biomass marketed as powders, pills and capsules.

The growth of microalgae cultures can be determined directly by measuring the abundance of the cells using cell counting by various methods and/or the increase in biomass, either by dry weight (DW), ash-free dry weight or chlorophyll *a* concentration. Methods using optical properties (e.g., turbidity or absorbance) are also used as a measure of algae abundance, however, precise correlations with direct methods (e.g., cell count, DW) should be considered for proper interpretation (Borowitzka and Moheimani 2013). The possibility of measuring cell count online together with statistics on cell size has been investigated by employing an in-situ, flow-through microscope installed in a microalgae cultivation bypass (Havlik et al. 2013b). Biomass measurement is not only important for its association with valuable compounds but also for estimating culture metabolic state variables (e.g., specific growth rate) necessary in fed-batch strategies (Wechselberger et al. 2013). Moreover, industrial processes using microalgae require the biomass as a variable for the estimation of process productivity (e.g., g DW/L, ton DW/ha), yields, and economics, e.g., €/kg DW (Norsker et al. 2011). Therefore, the use of devices to monitor this variable has been of interest to researchers, given its relevance in commercial applications.

The most common technique used in laboratory and industrial environments to determine microalgae growth involves the use of optical density (OD). Traditionally, the wavelengths at 750 nm (OD<sub>750</sub>) and 550 nm (OD<sub>550</sub>) have been employed for this purpose due to minimal interference with the cellular pigments that are present, e.g., chlorophyll and carotenoids (Borowitzka and Moheimani 2013; Wang et al. 2019).

On the other hand, authors have described the use of other wavelengths for cell growth estimation. In a short study, four different wavelengths comprised in a range between 677 and 688 nm (within the maximum chlorophyll absorption), using four different microalgae species, were employed for the purpose of correlating absorbance and cell concentration (Santos-Ballardo et al. 2015). In another study, the authors

employed several wavelengths, 480, 510, 630, 647, 647, 650, 664 and 750 nm, to correlate them with cell concentration, but also to compare OD measurements and wavelengths employed. The authors used seven microalgae strains in this task. Their results showed consistency between OD measurements regardless of the wavelength used (Pearson coefficient = 0.92–0.97) which suggests that the wavelength selected is not a determining factor (Chirivella-Martorell et al. 2018).

## 2.2 Lipids

Microalgae are an interesting source for different types of lipids. In general, the lipid profile of microalgae consists of neutral lipids (or non-polar lipids) such as triacylglycerols (TGA), free fatty acids (FFA) and carotenoids, i.e., lipidic pigments, and polar lipids mainly represented by glycolipids and phospholipids (Sarpal et al. 2015). The fatty acid profile in microalgae is characterized by a mixture of C16 and C18 saturated and unsaturated fatty acids, as well as long-chain (C20–C22) polyunsaturated fatty acids (PUFAS), e.g., eicosapentaenoic acid (EPA), docosahexaenoic acid (DHA), and docosapentaenoic acid (DPA) (Gouveia and Oliveira 2009; Barkia et al. 2019; Yang et al. 2020).

The type and amount of fatty acids present in microalgae can vary considerably among different phylogenetic groups, and even differ at the species level. Environmental conditions and growth conditions (e.g., pH, light intensity, nutrient limitation, and oxidative stress) are other factors that shape the lipid profile in microalgae (Andersen 2013; Borowitzka et al. 2016; Bi and He 2020; Morales et al. 2021). An extensive study of the long-chain fatty acid profile (C14–C24) in 2076 microalgae strains that investigate cyanobacteria and different phyla and classes of eukaryotic algae has been reported (Lang et al. 2011). Other studies on lipid characterization employing different microalgae can be found elsewhere (Yao et al. 2015; Shen et al. 2016).

Research with microalgae lipids started with the aim of biodiesel production (Barkia et al. 2019). However, this application has encountered several limitations in its step toward commercialization (Chisti 2013). The high costs of producing biofuels from microalgal biomass, both in open and closed systems, limit the economic success of the sector; however, scenarios that integrate biofuel production

processes with "next generation cultivation systems and processes" could favor the profitability of this activity (Richardson et al. 2014; Bi and He 2020). For this reason, biotechnological applications have currently been directed towards lipid products not for biofuels but products with higher added value within the pharmaceutical and food sectors due to their beneficial properties for human health, e.g., PUFAs, phytosterols, food supplements and infant formulas (Barkia et al. 2019; Fabris et al. 2020; Fernandes and Cordeiro 2021).

## 2.3 Pigments

As in plants, chlorophyll a (Chl a) constitutes the main pigment for photosynthesis in phytoplankton. However, these photosynthetic organisms also contain different accessory pigments (e.g., Chl b and c), carotenoids, and phycobiliproteins to carry out fundamental cellular processes, e.g., enhancement of light uptake and photoprotection (Andersen 2005). The pigments present in microalgae can vary between taxonomic groups and therefore be used to differentiate between them. For example, cyanobacteria, in general, produce Chl a, d and f, as well as the phycobiliproteins (e.g., phycocyanin, allophycocyanin, and phycoerythrin). An interesting exception within the cyanobacteria is the group of prochlorophytes, i.e., *Prochloron*, *Prochlorothrix*, and *Prochlorococcus*, which, in addition to Chl a, contains Chl b and lacks phycobiliproteins (Roche et al. 1996). Meanwhile, Chlorophyta, a group to which the genera *Chlorella*, *Dunaliella*, and *Haematococcus* belong, contains Chl a and b in addition to various types of carotenoids (e.g.,  $\beta$ -carotene) and various xanthophylls, e.g., astaxanthin, canthaxanthin, lutein, and zeaxanthin (Barkia et al. 2019).

There are three main classes of pigments present in microalgae and whose commercialization is of great interest: chlorophyll, carotenoids and phycobiliproteins (Silva et al. 2020).

### 2.3.1 Chlorophyll

Chlorophylls are responsible for the photosynthesis process as well as light energy harvesting in oxygenic photosynthetic organisms, i.e., plants, algae, and cyanobacteria. Currently, five types of chlorophyll have been described: Chl a, b, c, d, and f, obtained



from oxygenic photosynthetic organisms (Li and Chen 2015).

Chlorophylls exhibit variations in their side chains and/or reduced states. These structural variations in the rings and/or side chains give each chlorophyll distinctive characteristics in its absorption spectra. In general, chlorophylls present two main light absorption bands where their corresponding absorption maxima ( $\lambda$  max) are found; the short wavelength absorption band, i.e., the Soret band (in vitro and in vivo ~400–470 nm) and the long wavelength absorption band, i.e., the  $Q_y$  band (in vitro ~620–710 nm and in vivo ~640–710 nm) (Papageorgiou and Govindjee 2004; Chen 2014). Chl a in methanol exhibits a  $\lambda$  max of 436 and 665 nm. However, under in vivo conditions, Chl a, found in photosystem II, shows a  $\lambda$  max ~680 nm and  $\lambda$  max ~700 nm in photosystem I mainly attributed to the protein environment surrounding these molecules. On the other hand, Chl b under in vivo conditions presents values of  $\lambda$  max of ~650 nm (Chen 2014). It should be noted that different types of Chl c, present in golden-brown eukaryotic algae but absent in plants, present an additional  $\lambda$  max at ~580 nm under in vitro conditions (Zapata et al. 2006).

Interestingly, the Chl a is also almost the only chlorophyll performing fluorescence under in vivo conditions at ordinary temperatures. At room temperature, Chl a shows a heterogeneous behavior, however, a fluorescence band at 683–685 nm, which originates in PSII, and a small amount in the 710–760 nm region from the PSI antenna are observed (Papageorgiou and Govindjee 2004).

Despite being light- and oxygen-sensitive molecules, the biotechnological uses of chlorophylls are varied. For example, their use is common as dyes in the food, cosmetic and pharmaceutical sectors (Silva et al. 2020). In vivo evaluations of plant-derived chlorophyll extracts have demonstrated their protective capacity as antioxidants (Suparmi et al. 2016). A more recent work describes the process for stabilizing chlorophyll extracts with  $\text{Cu}(\text{NO}_3)_2$  in paints (Sulaiman et al. 2019).

### 2.3.2 Carotenoids

Carotenoids are lipid-soluble accessory pigments made up of isoprene units whose coloration mainly spans the visible light spectrum between yellow and

red, i.e., 400–600 nm (Langi et al. 2018; Silva et al. 2020). In addition to their role in light harvesting, carotenoids present a photoprotective function against oxidative stress and adverse environmental conditions (Barkia et al. 2019). According to their function, there are two major groups of carotenoids: carotenes (e.g.,  $\alpha$ -carotene,  $\beta$ -carotene, lycopene) and xanthophylls (e.g., astaxanthin, zeaxanthin, lutein, violaxanthin, canthaxanthin). Based on their chemical structure, carotenes are considered oxygen-deprived hydrocarbon compounds (C40 polyenes), whereas xanthophylls present oxygenated groups, i.e., hydroxyl and keto groups, toward the end rings. This feature confers a relative hydrophilic character to xanthophylls (Langi et al. 2018).

The properties and functions of carotenoids depend on their molecular structure. For example, carotenoids can present different isomeric configurations (trans and cis), resulting in variations in the melting point, solubility and stability of the molecule. In addition, the conjugated polyene chromophore, present in carotenoids, defines the properties of light absorption and light harvesting (Langi et al. 2018).

Among their biological activities, carotenoids have been associated with antioxidant, anti-inflammatory, and anticarcinogenic properties. The effects of different types of carotenoids on human health (e.g., cardiovascular protection, prevention of liver fibrosis, prevention against different types of cancer) can be widely consulted in the literature (Park et al. 2010; Yoshida et al. 2010; Milani et al. 2017; Langi et al. 2018; Barkia et al. 2019). In addition, the carotenoids are used in the food sector as a food coloring additive and nutraceutical, in the animal feed industry, and in cosmetology (Silva et al. 2020).

### 2.3.3 Phycobiliproteins

Phycobiliproteins are photosynthetic light-harvesting protein pigments naturally found in cyanobacteria, red algae, cryptomonads, and glaucophytes (Silva et al. 2020). These proteins are hydrophilic and are found in superstructures called phycobilisomes in the chloroplast stroma. A classification of the phycobiliproteins includes three main groups: allophycocyanin, phycocyanins, and phycoerythrins (Stadnichuk et al. 2015). Covalently attached to their polypeptide structure through cysteine residues, chromophore molecules called phycobilins are found (Kovaleski

et al. 2022). There are four types of phycobilins: phycocyanobilin (PCB, blue), phycoviolobilin (PVB, violet), phycoerythrobilin (PEB, red) and phycourobilin (PUB, yellow) (Dagnino-Leone et al. 2022). Other authors report only three types of phycobilins: PEB, PCB, and PVB (Kovaleski et al. 2022).

The amino acid sequence, the number of chromophores per monomer, and the type of chromophores present are criteria used to differentiate between phycobiliproteins. Based primarily on structural features and their absorption spectra, phycobiliproteins are divided into four groups: phycoerythrin ( $\lambda_{\max}=490\text{--}570\text{ nm}$ ), phycocyanin ( $\lambda_{\max}=610\text{--}620\text{ nm}$ ), phycoerythrocyanin ( $\lambda_{\max}=560\text{--}600\text{ nm}$ ), and allophycocyanin ( $\lambda_{\max}=650\text{--}655\text{ nm}$ ) (Kovaleski et al. 2022).

The online monitoring of phycocyanins using fluorescence spectrophotometry in the marine cyanobacterium *Synechococcus* sp. has demonstrated the feasibility of this approach to obtain direct information from cell cultures (Sode et al. 1991). Moreover, the offline methods of extraction and purification of phycobiliproteins are relatively simple and widely known, a fundamental step in the training and validation of a software sensor (Kovaleski, 2022). Therefore, all these elements raise the possibility of using methods based on fluorescence spectrophotometry for online monitoring and software sensor development for estimating phycobiliproteins. Considering that the range of fluorescence spectra of phycobiliproteins is between 585 and 665 nm (Stadnichuk and Tropin 2017), selective online monitoring of different types of phycobiliproteins, e.g., phycocyanin or phycoerythrin, in microalgae cultivation systems poses interesting challenges.

Cyanobacteria of the genus *Arthrospira* sp. and *Porphyridium* sp. are the most relevant microalgae in the industrial production of phycobiliproteins, specifically phycocyanin (blue) and phycoerythrin (red), respectively (Silva et al. 2020). Commercial interest in phycobiliproteins is associated with their bioactive properties, such as antioxidant, anti-inflammatory, anti-metabolic diseases, anti-cancer, anti-neurodegenerative, and antibiotic (Dagnino-Leone et al. 2022). Other applications related to the use of these molecules are related to protein markers, cell sorting, and phycobiliprotein-derived conjugates that take advantage of their properties as fluorescent probes (Tounsi et al. 2023). Recent applications employ the

fluorescent properties of phycoerythrin and phycocyanin to sense different analytes as part of nanoprobe or complexed with other molecules like DNA (You et al. 2020; Ghosh et al. 2020). Finally, the differential quenching effect of heavy metals such as silver and copper on phycocyanin fluorescence has demonstrated their potential use in selective monitoring for the presence of heavy metals (Bellamy-Carter et al. 2022).

Silva et al. (2020) determined, in a bibliometric study, that between 2009 and 2019, research was mainly focused on the study of the pigments phycocyanin, chlorophylls,  $\beta$ -carotene, and astaxanthin obtained from *A. platensis*, *C. vulgaris*, *D. salina*, and *H. pluvialis*, respectively. At present, there are reports where the use of software sensors has been employed for the estimation of chlorophyll and carotenoids (Sá et al. 2020b, a). However, the monitoring and estimation of phycobiliproteins in microalgae cultivation through the use of software sensors represents an interesting challenge to be explored.

## 2.4 Vitamins

Vitamins are organic compounds commonly required in low concentrations by organisms and indispensable in many vital cell processes. Vitamins are classified into two categories: fat-soluble molecules (e.g., vitamins A, D, E and K) and water-soluble molecules (e.g., B-complex vitamins and vitamin C) (Udayan et al. 2017).

Microalgae constitute a diverse source of different types of vitamins, e.g., vitamins A, D, E, K and several B vitamins, i.e., B<sub>1</sub> (thiamine), B<sub>2</sub> (riboflavin), B<sub>3</sub> (niacin), B<sub>5</sub> (pantothenic acid), B<sub>6</sub> (pyridoxine), B<sub>7</sub> (biotin), B<sub>9</sub> (folic acid), and B<sub>12</sub> (cyanocobalamin) (Del Mondo et al. 2020). For example, in *A. platensis* (formerly *Spirulina platensis*) (<http://www.algaebase.org/>) the presence of vitamins B<sub>1</sub>, B<sub>2</sub>, B<sub>3</sub>, B<sub>6</sub>, B<sub>9</sub>, B<sub>12</sub>, vitamin C, vitamin D and vitamin E has been reported (Jung et al. 2019). *C. vulgaris* constitutes another source of multiple vitamins, e.g., B<sub>2</sub>, B<sub>3</sub>, B<sub>9</sub>, and B<sub>12</sub> (Edelmann et al. 2019). *Dunaliella* sp. is rich in fat-soluble vitamins and moreover, interesting concentrations of vitamin B<sub>2</sub>, vitamin B<sub>12</sub>, folic acid, vitamin C, vitamin B<sub>3</sub> and vitamin E have been reported in *D. tertiolecta* (Udayan et al. 2017). More details on the production of different types of vitamins by

some microalgae genera can be found elsewhere (Del Mondo et al. 2020).

There are some intrinsic characteristics in the fluorescence of some vitamins of the B complex that could be useful for monitoring microalgae cultures with software sensors using optical measurement methods. To make use of these characteristics, the knowledge of the excitation and emission spectra of these molecules is essential. For example, vitamin B<sub>1</sub> in water presents a fluorescence region at  $\lambda_{ex}/\lambda_{em}=370/460$  nm (Yang et al. 2016). Vitamin B<sub>6</sub>, B<sub>2</sub> and B<sub>9</sub>, in aqueous solution, exhibit a fluorescence region at  $\lambda_{ex}/\lambda_{em}=330/380\text{--}390$  nm,  $\lambda_{ex}/\lambda_{em}=445/520$  nm, and  $\lambda_{ex}/\lambda_{em}=330/450$  nm, respectively (Parri et al. 2020). Vitamin B<sub>2</sub> also has a second fluorescence region that comprises  $\lambda_{ex}/\lambda_{em}=365/520$  nm (Faassen and Hitzmann 2015).

## 2.5 Other relevant microalgal compounds

Another component present in microalgal biomass is the protein fraction. Several microalgae species report high protein concentrations ranging from 42 to 70% in some cyanobacteria and up to 58% in *C. vulgaris* on a dry cell basis (Becker 2007; Barkia et al. 2019). Other species such as *D. salina*, *H. pluvialis*, *Nannochloropsis* sp., *A. platensis* and *A. maxima* show protein amounts of 49–57%, 29–45%, 50–55%, 46–63%, and 60–71% on a dry cell basis, respectively (Timira et al. 2022). A recent work shows a protein content of 51% and 64% (w/w) in two *Galdieria sulphuraria* extremophile strains (Canelli et al. 2023). The essential and non-essential amino acid profiles present are comparable with those of other protein sources, i.e., egg albumin, soybean and milk lactoglobulin (Williams and Laurens 2010).

Carbohydrates comprise between 12 and 64% of microalgal biomass (Becker 2007; Markou et al. 2012). These, present as mono-, oligo-, and polysaccharides, can be used directly in cellular metabolic activities, stored in internal structures called plastids, or be cellular structural components (e.g., the cell wall). Different types of polysaccharides produced may vary according to the microalgae group. For example, starch, floridean starch, and chrysolaminarin are usually found in green algae (Chlorophyta), red algae and diatoms, respectively (Barkia et al. 2019). Glycogen and paramylon constitute other

storage polysaccharides present in microalgae (Gouda et al. 2022).

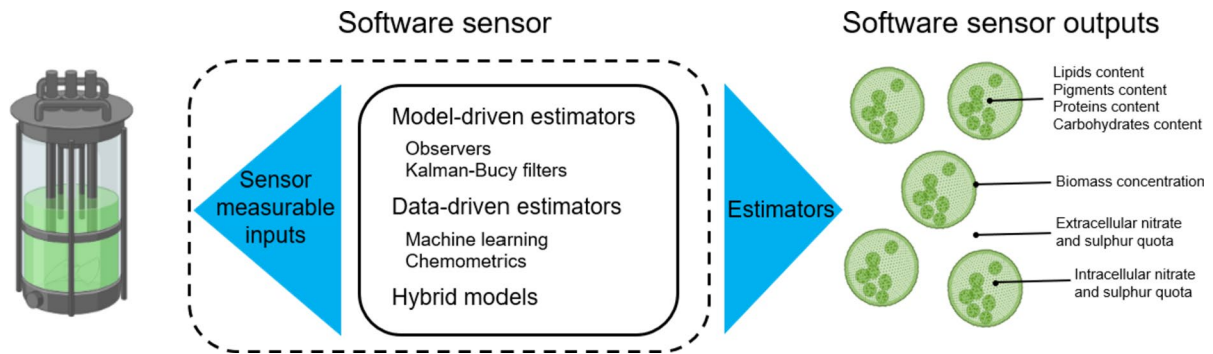
Applications of different microalgal polysaccharides are targeted towards the cosmetology industry either as hygroscopic agents or topical antioxidants (Barkia et al. 2019). Applications of exopolysaccharides in the food sector include their use as thickeners and gelling additives (García et al. 2017). Moreover, some microalgal polysaccharides and oligosaccharides reveal potential uses in the food sector as prebiotics. Xylooligosaccharides (XOS), galacto-oligosaccharides (GOS), alginate oligosaccharides (ALGOS), neoagaro-oligosaccharides (NAOS), galactans, arabinoxylans, and  $\beta$ -glucans are interesting molecules aimed at this type of application (Gouda et al. 2022).

## 3 Software sensors

A software sensor is a combination of process data (sensor input), measured by hardware sensors, and a model that analyzes these input data in order to obtain new process information (sensor output), in order to deliver information about the process similar to that delivered by hardware sensors (see Fig. 2) (Luttmann et al. 2012). Hence, a software sensor implies an indirect measurement of the desired variable (Brunner et al. 2021). Software sensors are commonly applied in all the engineering disciplines (automotive and aircraft control, wastewater treatment, industrial-scale chemical processes such as distillation), including bioprocess engineering (Luttmann et al. 2012). Generally, all software sensors are observers in the broad sense of the word in control theory (Luttmann et al. 2012; Aguilar-Garnica and García-Sandoval 2015).

The advantages of software sensors seem obvious: they enable the process operator to estimate values of process variables that either cannot be measured directly online in real time or can be measured only by sampling and an atline or offline analysis, which complicates things. There are, naturally, shortcomings as well: designers of software sensors used in bioprocesses have to cope with variable process lengths, multiple process phases and especially faults of physical sensors used as inputs. These and other challenges for software sensor developers are discussed in (Brunner et al. 2021) and (Luttmann et al. 2012).





**Fig. 2** General overview of a software sensor for microalgae cultivation. Measurable sensor inputs obtained from hardware components are processed by complex mathematical models

(software-based) to indirectly predict relevant microalgae cultivation variables. Partially created in Biorender.com

In general, the development of software sensors can be carried out using two different types of methodologies: model-driven estimators and data-driven estimators, which are terms generally accepted in the literature (Kadlec et al. 2009; Luttmann et al. 2012). In a model-driven estimator, also referred to as mechanistic, white-box, parametric or knowledge-based models, there is a defined structure originating from previous mathematical transformations (e.g., partial or ordinary differential equations) that describes a causal relationship in complex systems (e.g., biological systems) based on parameters that can be measured or inferred (Solle et al. 2017). These estimators use mass and energy balance models as well as kinetic models (Randek and Mandenius 2018). Under this approach, a prior knowledge of the system/process is essential to describe the variables of interest; nevertheless, detailed knowledge of the process mechanisms is not always available, or its use becomes challenging due to the intrinsic complexity of living cells (Wechselberger et al. 2013). Generally, these models have a wider range of application than data-driven estimators because of their greater extrapolation power (Solle et al. 2017).

Data-driven estimators, on the other hand, comprise the use of mathematical representations (e.g., artificial neural network(ANN)) and chemometric tools (e.g., principal component analysis (PCA), partial least squares (PLS), and principal component regression(PCR)) in order to obtain values of relevant process variables obtained from the data (Solle et al. 2017). Basically, these estimators are based on empirical observations of the process (Randek and

Mandenius 2018). The models in these estimators correspond to correlations between the measured variables and the unknown variables to be estimated (Havlik et al. 2022). This approach does not require prior knowledge of the process and is useful in contexts where process conditions do not vary considerably (Wechselberger et al. 2013). Unlike model-driven estimators, data-driven estimators do not require as much time for implementation as long as the data is available (Solle et al. 2017). The data-driven estimators could also be referred to as "black box" or empirical models. Generally, the ability of data-driven estimators to extrapolate outside the region of available data is poor.

A third methodology suggested in the literature for software sensor development concerns the use of hybrid models (Brunner et al. 2021). Essentially, this model is a combination of the model-driven and data-driven approaches (Simutis et al. 1992; Noll and Henkel 2020). A drawback of these models may be the difficulty in understanding and validating the parameters (Mandenius and Gustavsson 2015). The hybrid model approach could also be known as the "grey box" model.

When developing a software sensor, different techniques could be employed individually or in combination, in parallel or in sequence (Randek and Mandenius 2018; Rathore et al. 2021). For example, software sensors could employ individual approaches like ANN, deep neural network, adaptive interval observer, Luenberger observer, Kalman filter, and PLS regression model, but also multiple linear regression (data-driven model) in combination with a

mechanistic model, support vector regression (SVR) together with random forest regression (RFR), both data-driven approaches, among others (Rathore et al. 2021; Havlik et al. 2022). In general, the techniques described here are part of the two wider categories mentioned previously.

### 3.1 Model-driven estimators

#### 3.1.1 Observers

In a linear system, in order to estimate a state variable (e.g., biomass concentration, substrate concentration, and product concentration), a state vector (observer) is constructed based on the available outputs and inputs of the original system. The main task of an observer is the online state estimation of variables that cannot be measured directly (Flores et al. 2020). The simplest observer described is the Luenberger observer (Luenberger 1966). Several types of observers are described in the literature, including the integral proportional observer, adaptive observer, interval observer, high gain observer, and asymptotic observer (AO) (Gauthier et al. 1992; Ploix et al. 1997; Noll and Henkel 2020; Flores et al. 2020). High-gain observers, whose structure does not depend primarily on the input, have demonstrated satisfactory performance in nonlinear systems, especially in systems with biological and environmental processes, when estimating variables and disturbances with uncertainties. On the other hand, integral proportional observers have demonstrated in multiple applications a higher robustness to system input disturbances compared to high-gain and adaptive observers (Flores et al. 2020).

Examples of interval observers applied to microalgae cultivation can be found in literature (Goffaux et al. 2009; Mairet et al. 2014a). Adaptive interval observer have also been reported in microalgae cultivation (Mairet et al. 2014a, b). The Luenberger observer and high gain observers have also been applied to microalgae cultivation (Benavides et al. 2015; Flores et al. 2020).

#### 3.1.2 Kalman-Bucy filters (Kalman filters)

The Kalman filter, originally developed in the 1960s, is a model-driven estimator used to improve the reliability of estimated data and noise filtering (Kalman 1960; Kalman and Bucy 1961; Noll and Henkel

2020). The Kalman filters are useful and powerful estimators when a process is linear and a model is available (Rehm and Reed, 1991). Although this algorithm was originally used for the estimation of linear systems, subsequent improvements and extensions of the model facilitated the development of the so-called extended Kalman filter (EKF), which attempts to estimate system behavior based on nonlinear models and prior knowledge of the mean and covariance associated with the measurement error (de Assis and Filho 2000).

Other types of Kalman filters used for nonlinear systems are the unscented Kalman filter (UKF), the ensemble Kalman filter, and the cubature Kalman filter. However, the two most relevant algorithms for state estimation of nonlinear systems are the EKF and the UKF. It is worth mentioning that the different variants of Kalman filters differ in the way the error is calculated (Yousefi-Darani et al. 2021).

The Kalman filters have been used for the estimation of different biological parameters in microalgae cultures, e.g., biomass concentration, extracellular and intracellular sulfur concentration, and lipid concentration (Daboussy et al. 2014; Yoo et al. 2015; García-Mañas et al. 2019).

### 3.2 Data-driven estimators

#### 3.2.1 Artificial neural network

The ANNs were first introduced in 1943 by Warren McCulloch and Walter Pitts (Gerón 2019). Further theoretical principles behind ANN were described in the 1950s; however, its application in the context of bioprocesses became relevant in the early 1990s (Thibault et al. 1990). The ANN is a machine learning technique, i.e., an algorithm, inspired by the neural structure of the brain system in humans. A typical ANN is composed of three layers of interconnected nodes: an inner layer, a hidden layer, and an outer layer (Ning et al. 2022). Essentially, this technique allows the estimation of process variables without the requirement of a process model (Noll and Henkel 2020). ANNs are used to provide approximations to nonlinear systems (de Assis and Filho 2000). Briefly, the network must receive an input from a set of data (e.g., pH, nitrate concentration), which is then propagated by one or more hidden or intermediate layers to an outer layer, which finally predicts the variable of

interest, e.g., biomass concentration (Mowbray et al. 2021).

The ANNs are considered highly flexible, powerful, and scalable models since they provide the possibility to handle large amounts of parameters and demonstrably approximate any function (Cybenko 1989; Gerón 2019). Interestingly, the parameters of the ANN could be modified in an iterative manner using the model itself and a set of input and output data (commonly referred to as “training data”). In this way, the inputs correspond to the correct outputs, but this method also ensures that the estimates are approximations to the training data. Additionally, the network has the ability to “learn” in the presence of a newly supplied set of input data and predict an unobserved output. This is achieved based on the previous knowledge generated from the training data supplied in the network (Mowbray et al. 2021).

Several types and structures of ANN have been described in biological and engineering applications, e.g., feed-forward neural network, convolutional neural network (CNN), recurrent neural network (RNN), and deep neural network (DNN), among others (Beiroti et al. 2019; Mowbray et al. 2021; Ning et al. 2022; Xu et al. 2022).

The use of ANNs as software sensors in microalgae cultivation has been reported for the estimation of cell counts, contamination in monoalgal cultures, or biomass concentration using fluorescence spectra, multispectral absorption, and reflectance through offline monitoring (López Expósito et al. 2017; Franco et al. 2019; Liu et al. 2020a).

### 3.2.2 Chemometric models

Chemometrics, a concept initially introduced by Svante Wolf in 1971, is the method of extracting information from chemical systems by data-driven methods (Sarker and Nahar 2015; Biechele et al. 2015). The information from the data is extracted using multivariate statistics, applied mathematics, and computer science. When using chemometrics-based approaches, the data must be pretreated/transformed to ensure homogeneity and ensure that the model can improve its prediction capability (Biechele et al. 2015; Solle et al. 2017). For this purpose, several pre-processing methods are described, including filtration, centering, normalization, standardization, deviation, and weighing (Biechele et al. 2015).

After the application of multivariate data analysis, qualitative and quantitative information can be obtained from the data provided. Some of the most commonly employed linear algorithms in a multivariate data analysis include PCA, PLS, PCR, and multiple linear regression (MLR) (Faassen and Hitzmann 2015; Mowbray et al. 2021). A brief explanation of PCA and PLS, two commonly used chemometric tools, is given below:

The PCA is a mathematical algorithm that reduces the dimensionality of the data while retaining most of the variation in the data set (Ringnér 2008). The model achieves this reduction by identifying directions or uncorrelated variables, i.e., latent variables, called principal components, in which the variation of the data is maximal (Ringnér 2008; Mowbray et al. 2021). Therefore, it is assumed that the variance of latent variables can represent the ‘hidden’ pattern within the original datasets and hence is the best way to distinguish different datasets (Mowbray et al. 2021). Briefly, a PCA allows plotting the data, visually assessing similarities and differences, and determining whether the data can be grouped (Ringnér 2008). PCA is a qualitative method commonly used to analyze the structure, variance, or distribution of data sets and to identify outliers (Biechele et al. 2015).

PLS regression is a quantitative method using a linear multivariate algorithm, that relates and models two data matrices, X and Y (Wold et al. 2001). Therefore, it is used when the goal of the analysis is to predict a set of variables (denoted Y) from a set of predictors (called X) (Kumar 2021). The PLS model not only captures the maximum variances linked to the predictor (e.g., spectroscopic data) and prediction (e.g., concentration data), but also ensures the maximum correlation between them (Abdi and Williams 2013). A favorable aspect of PLS is the ability to analyze data with many noisy, collinear (correlated), and incomplete variables in both matrices (Wold et al. 2001).

## 4 Measurement methods used as software sensor inputs in microalgae cultivation.

In the following section, a detailed set of recent literature comprising measurement methods used in microalgae software sensors is described. Measurement

methods are divided into optical and non-optical techniques, a combination of both, and those based on simulated data (see Fig. 3). Model-driven or data-driven estimators are commented on regardless of the previous classification according to the measurement method. References described in this section are summarized in Table 1, characterized by software sensor input, type of software sensor, and process variables estimated by the sensor (i.e., software sensor output).

#### 4.1 Optical methods

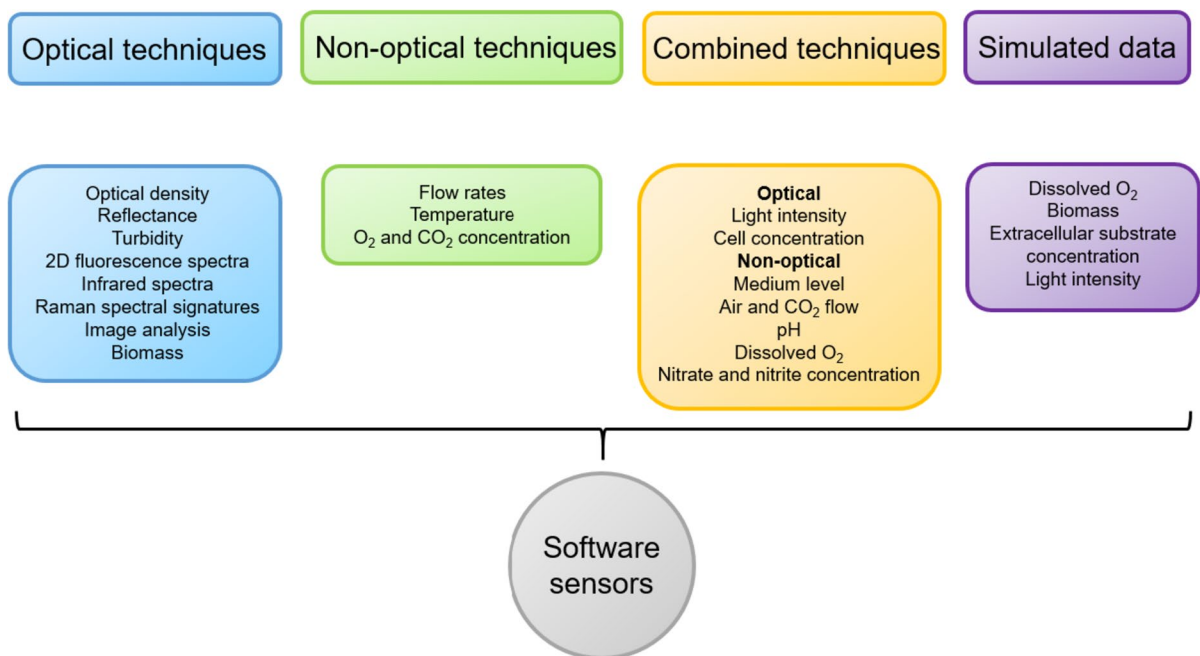
##### 4.1.1 UV–VIS spectroscopy

In biotechnology, UV–VIS spectroscopy uses the interaction of ultraviolet and visible light, i.e., wavelengths from 200 to 780 nm, with molecules in a biological sample to gather information from them (Roberts et al. 2018). In general, the interaction of electromagnetic radiation with the molecule could result in two different phenomena called absorbance and scattering. Fluorescence, in particular, is an emission phenomenon different from any of the two previously mentioned. Other physical interactions of light with matter include absorption, reflectance,

refraction, dispersion, and diffraction (Barsanti, 2014). These different processes are used as the measuring principle when referring to optical sensors.

Optical density, reflectance, and turbidity (a type of scattering) have been used in the development of software sensors using microalgae cultivation. In a work using flat-panel photobioreactors with *D. tertiolecta*, Benavides et al. 2015 reported the development of two software sensors using extended Luenberger observers. The first observer allowed the estimation of extracellular nitrate and intracellular nitrate content considering only biomass online monitoring (input), while the second observer benefited from biomass and extracellular nitrate (inputs) to estimate intracellular nitrate concentration. Biomass online monitoring was tracked with optical density. Nitrate online measurements were obtained through an immersed probe and a spectrometer.

Flores et al. (2020) used a turbidity sensor that allowed them to estimate biomass and substrate (glucose) after using a high-gain, robust nonlinear observer in mixotrophic cultures with the cyanobacterium *A. platensis*. The turbidity sensor was designed to be non-invasive and capable of monitoring microalgae flowing inside a borosilicate tube.



**Fig. 3** Measurement methods and measured variables for the development of software sensors in microalgae cultivation

**Table 1** Software sensors used for the estimation of relevant process parameters in microalgae cultivations

Software sensor inputs	Type of software sensor	Estimated variables	Reference
OD (biomass)	Extended Luenberger observer	Extracellular and intracellular $\text{NO}_3^-$ quota	Benavides et al. (2015)
OD (biomass), extracellular $\text{NO}_3^-$		Intracellular $\text{NO}_3^-$ quota	
Turbidity	Robust nonlinear observer	Biomass (g/L) Glucose	Flores et al. (2020)
Turbidity (biomass) Glucose	EKF UKF PF	Total lipids	Yoo et al. (2015)
Turbidity (biomass) Glucose	UKF	Total lipids	Yoo et al. (2016)
Multispectral absorption	ANN	Contamination	Franco et al. (2019)
2D fluorescence spectra	Chemometrics	Biomass (cell/mL), pigment content (chlorophyll), lipids content (saturated, unsaturated, and total fatty acids)	Sá et al. (2020a)
2D fluorescence spectra	Chemometrics	Biomass (cell/mL) and viability	Sá et al. (2017)
2D fluorescence spectra Climatic metadata	Chemometrics	Pigment content (chlorophyll and carotenoids)	Sá et al. (2020b)
2D fluorescence spectra	Chemometrics	Biomass (g/L) Fucoxanthin content	Gao et al. (2021)
Fluorescence spectra	ANN	Biomass (cell/mL)	Liu et al. (2020a, b)
Laser reflectance	SVR RFR	Biomass (g/L)	López Expósito et al. (2017)
Laser reflectance	ANN	Biomass (g/L)	López Expósito et al. (2016)
IR spectra (ATR-FTIR)	Chemometrics	Proteins, lipids, carbohydrates	Ferro et al. (2019)
Raman spectral signatures	Chemometrics	Physiological cell state	Lieutaud et al. (2019)
Image-based	Linear regression, RFR, XGBoost MobileNetV3Small (Deep Learning) MobileNetV3Large (Deep Learning)	Microalgae identification	Kaya et al. (2023)
Particle counter	ANN	Microalgae identification	Otálora et al. (2021)
Image-based	Chemometrics	Biomass (cell/mL)	Nguyen et al. (2023)
Wastewater flow rate	Linear polynomial, nonlinear polynomial model, Hammerstein-Wiener model (data-driven models) Hybrid model (data-driven + kinetic model)	pH	Paladino et al. (2022)
Dilution flow rate		Biomass (g/L) using pH computed by 1	
Temperature			
$\text{CO}_2$ and $\text{O}_2$ concentrations	EKF	Biomass (g/L) Extracellular and intracellular sulphur quota	Daboussy et al. (2014)
Solar radiation Medium level $\text{CO}_2$ supply	ANN	pH	Otálora et al. (2023)
Dissolved $\text{O}_2$ pH Gas injection (air flow and $\text{CO}_2$ flow) Solar radiation	EKF	Biomass (g/L)	García-Mañas et al. (2019)



**Table 1** (continued)

Software sensor inputs	Type of software sensor	Estimated variables	Reference
NO <sub>3</sub> <sup>-</sup> and NO <sub>2</sub> <sup>-</sup> concentrations	Adaptive interval observer	Lipids	Mairet et al. (2014a)
Particle counter		Carbohydrate quota	Mairet et al. (2014b)
Simulated dissolved O <sub>2</sub>	EKF, AO, EKF+AO	Biomass (kg/m <sup>3</sup> )	Arango Restrepo et al. (2022)
Simulated biomass and extracellular substrate concentrations	STO and AO	Substrate uptake rate, biomass growth rate, and internal substrate quota	Coutinho et al. (2019)
Simulated light intensity	High gain observer	Biomass (mg/L) Dissolved CO <sub>2</sub>	Farza et al. (2019)

*EKF* extended Kalman filter, *UKF* unscented Kalman filter, *PF* particle filter, *ANN* artificial neural network, *ATR-FTIR* attenuated total reflection–Fourier transform infrared spectroscopy, *SVR* support vector regression, *RFR* random forest regression, *STO* super twisting observers, *AO* asymptotic observer

Cultures were performed with a flat-panel PHB. Measurements of the device, which operated at OD<sub>560</sub>, were calibrated by an external determination of dry biomass weight. External glucose measurements were carried out with the DNS assay, a standard method for the determination of reducing sugars.

The evaluation of various estimators employing different observers, i.e., EKF, UKF, and particle filter (PF), was carried out for the determination of total lipids in *C. protothecoides* photobioreactor cultures. The authors used a turbidity sensor that allowed online estimation of biomass concentration (externally calibrated with DW) as well as offline HPLC analysis to measure glucose concentration. These two variables were the basis for the development of the lipid estimators in the cultures. The external determination of total lipids was carried out using the Red Nile stain test in order to compare the total lipid predictions generated by the models and the microalgae cultures. The comparison of the applicability of these three algorithms yielded favorable results for the use of UKF and PF compared to EKF (Yoo et al. 2015).

In a continuation of the previous research, the same authors used only UKF to estimate total lipids through turbidity and glucose measurements and used this data to perform model predictive control (MPC). MPC, based on successive linearization, was applied both in simulations and experimental validations for the purpose of optimizing biomass and lipid productivity in microalgal systems with photobioreactors. It was pointed out that significant improvements in biomass and lipid productivity were achieved when MPC calculation was applied, however, inaccuracies

as “lag phase” were observed due to the introduction of large amounts of inputs (Yoo et al. 2016).

Franco et al. (2019) reported the use of color spectrophotometry measurements (range of 400–700 nm), i.e., light absorption spectra, in combination with ANN to identify and distinguish among different microalgae species in monoalgal and mixed cultures. The research was carried out under laboratory conditions with the microalgae *Nostoc* sp., *Scenedesmus almeriensis*, *A. platensis* and *C. vulgaris* using continuous cultures in bubble-column photobioreactors. The research takes advantage of the personalized spectral fingerprint of each microalgae species to identify them in a monoculture and differentiate among them in a mixed culture in order to detect microalgae contaminants in commercial-scale production.

#### 4.1.2 Fluorescence spectroscopy

The phenomenon of fluorescence in a molecule occurs when the absorption of light energy promotes electrons into a temporary excited state (fluorescence lifetime), followed by a return to a basal state in which the excess energy is dissipated through the emission of photons, i.e., energy whose wavelength is larger than that which generated the excited state, due to energy loss by, e.g., vibration (Ishikawa-Ankerhold et al. 2012). Fluorescence offers a number of advantages in the monitoring of biological variables, such as its high sensitivity to monitor sub-nanomolar concentrations or even at the level of a single molecule, simplicity and speed, low background due to the differences between excitation and emission wavelengths, selectivity due to the ability to recognize

specific fluorophores within a mixture with non-fluorescent molecules, and its applicability to living organisms in a non-invasive and non-destructive way (Elson 2011; Sá et al. 2020a).

In biotechnology, fluorescence measurements are commonly used for the detection of biological molecules such as proteins, amino acids (e.g., tryptophan, tyrosine, and phenylalanine), vitamins, and cofactors, e.g., FAD and NADH (Faassen and Hitzmann 2015). In the case of microalgae cultivation, chlorophyll should also be considered a fluorophore used in the monitoring and control of these microorganisms. Their application, together with software sensor approaches using microalgae cultures, is described in further detail in the following paragraphs.

Offline fluorescent spectra based on Excitation-Emission Matrix (EEMs) together with chemometric models (PLS regression and N-PLS) was used to determine five biological parameters, i.e., the biomass as cell concentration, pigment content (chlorophyll), and lipid content (saturated, unsaturated, and total fatty acids) in *Nannochloropsis oceanica* cultivation. The determination coefficient ( $R^2$ ) of the predicted models developed was reported for cell concentration ( $R^2=0.66$ ), chlorophyll content ( $R^2=0.78$ ), total fatty acids ( $R^2=0.78$ ), saturated ( $R^2=0.81$ ), unsaturated ( $R^2=0.74$ ) (Sá et al. 2020a).

Cultivation using *D. salina*, 2D fluorescence spectroscopy, and chemometric models was employed for the monitoring of cell viability and cell number during the “green” growing phase and cell damage/disruption in the “orange” harvesting phase. The EEMs and mathematical modeling, i.e., PCA and PLS, allowed the development of model predictions for the desired process parameters. Models that predicted cell number and viability with green *D. salina* showed values of  $R^2=0.8$  and  $R^2=0.9$ , respectively, for training and validation. Also, the model prediction of cell damage for orange *D. salina* was achieved with an  $R^2=0.95$  for training and  $R^2=0.87$  for validation (Sá et al. 2017).

In another interesting approach, the content of total pigments, total chlorophyll, total carotenoids and specific pigments (e.g., chlorophyll a-b, zeaxanthin,  $\alpha$ -carotene, all-trans- $\beta$ -carotene, and 9-cis- $\beta$ -carotene) was estimated using 2D fluorescence spectroscopy, climatic metadata (temperature, total precipitation, sunlight, clouds, fraction of photo-synthetically active radiation-FPAR and irradiance)

and chemometric tools with outdoors and indoors cultures of *D. salina*. To do this, predictive models were developed with PCA and PLS to correlate EEMs and climatological data with the pigments present using two offline calibration methodologies: spectrophotometry and HPLC analysis. Predictive models for chlorophyll content (total, a-b) showed  $R^2$  values between 0.6 and 0.9 for trained and validated data. Models for total carotenoids and specific carotenoids showed  $R^2$  values between 0.7 and 0.9 and 0.6 and 0.9 for trained and validated data, respectively (Sá et al. 2020b).

In an experimental set-up using pilot-scale outdoor flat-panel photobioreactors and a semicontinuous cultivation of the brown microalgae *Tisochrysis lutea* and *Phaeodactylum tricornutum*, the biomass concentration and the carotenoid fucoxanthin were predicted using 2D fluorescence spectroscopy and chemometric modeling. The preprocessed EEMs, done by removal of the Rayleigh scatter and inner filter effect, were correlated with cell concentration (DW) and fucoxanthin content using PLS and N-PLS. Interestingly, this study presents the estimation of both biological variables based on each microalgal species and on both combined. While the validation models for estimating the biomass concentration from each microalgae showed, in both cases, an  $R^2=0.96$ , those with *T. lutea* and *P. tricornutum* presented an  $R^2=0.93$ . On the other hand, the individual models in the validation of the estimation of the fucoxanthin content showed an  $R^2=0.63$  (*T. lutea*) and 0.64 (*P. tricornutum*), while the results of both showed an  $R^2=0.77$  (Gao et al. 2021).

Cell counts of *C. reinhardtii* were estimated in the range between  $2 \times 10^5$  and  $6.4 \times 10^6$  cells/mL using offline fluorescence measurements and ANN-based models. The wavelength used for excitation was 470 nm, while the defined emission spectrum was in the range between 660 and 760 nm. The fluorescence spectra were analyzed by ANN and optimized with Genetic algorithms (GA) given the non-linear relationship between the input (fluorescence spectra) and the output (cell count) variables. The application of GA produced a slight improvement in model prediction compared to the application of ANN alone, showing values of  $R^2=0.99$  and 0.98, respectively (Liu et al. 2020a).

#### 4.1.3 Reflectance measurements

Offline laser reflectance data, analyzed as cord length spectra (input), was used for biomass estimation of *C. sorokiniana* cultures using wastewater treatment in 5.5 L photobioreactors. In this case, the cord length spectra were analyzed together with two machine learning techniques: SVR and RFR. The models generated with SVR and RFR showed good accuracy in biomass estimation (dry biomass concentration—g/L) after presenting values of  $R^2=0.87$  and  $0.81$ , respectively (López Expósito et al. 2017).

In a similar approach, the biomass concentration (flocculated and raw cultivation) of *C. reinhardtii* was estimated by means of a focused beam reflectance probe (FBRM) using software sensors. The authors used the chord length distribution data and a feed-forward multilayer perceptron, i.e., ANN, to develop great biomass concentration estimators showing  $R^2=0.96$  and  $0.92$  for training and validation, respectively (López Expósito et al. 2016).

#### 4.1.4 Infrared spectrum

Infrared spectroscopic techniques use the electromagnetic radiation wavelengths from about 700 to 4000 nm to obtain relevant information from cultivated cells (Barsanti, 2014). Especially the near infrared (NIR) and the mid infrared (MIR) regions are used in the development of methods for bioprocessing cell monitoring, whereas the far infrared region (above 15  $\mu\text{m}$ ) is rarely used for these applications (Rösner et al. 2022).

Attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy was used for the quantification of species-specific banding patterns of proteins, lipids and carbohydrates in seven different strains (*C. vulgaris*, *Coelastrella* sp., *Coelastrum astroideum*, *Desmodesmus* sp., *Scenedesmus* sp., *Desmodesmus* sp. and *S. obliquus* UTEX 417). These strains were cultivated in flat-panel photobioreactors under nitrogen-limiting conditions. The spectral data were analyzed with three chemometric methods: univariate linear regression analysis (ULRA), orthogonal partial least squares (OPLS), and multivariate curve resolution—alternating least squares (MCR-ALS). After analysis with the OPLS method, the correlations between the experimental values and the model obtained for proteins and lipids were excellent ( $R^2$

$Y \geq 0.90$ ), while those for carbohydrates showed moderate correlations ( $R^2$   $Y=0.77$ ). The analysis with OPLS presented the best results compared to the other two methods for the prediction of proteins, lipids and carbohydrates (Ferro et al. 2019).

#### 4.1.5 Raman spectroscopy

Raman spectrometry is employed for the estimation of the physiological state of *C. reinhardtii* cultures grown in batch mode using an airlift photobioreactor (Lieutaud et al. 2019). The physiological stages of the culture were determined by offline measurements of cell concentration (cell/mL), optical density ( $\text{OD}_{750}$ ), ammonia and pigment (Chl a and  $\beta$ -carotene) concentrations, as well as Raman spectroscopic analysis. Raman spectral signatures were coupled to three physiological phases of the culture (i.e., exponential, deceleration and stationary phases) by means of a PCA and a factorial discriminant analysis. On average, physiological state predictions reached 81.4%, 84% and 95.3% for the exponential, deceleration and stationary stages, respectively. The 16-day prediction rate for the three physiological stages reached 89.2% on average. All Raman measurements are carried out in an offline mode by using cell deposits on gold surfaces, but a potential implementation of a Raman probe inside a PBR and continuous monitoring is suggested for future improvements.

#### 4.1.6 Image-based methods

Image-based analysis using machine learning approaches is useful for the monitoring of algae abundance, classification and identification in aquatic environments. There are several previous works where these tools have been used for this purpose. In (Li et al. 2017), morphologically similar algae were classified by several types of convolutional neural networks (CNN) trained on a dataset of Mueller matrices containing the shapes of eight species of microalgae and one species of cyanobacteria. A 97% classification accuracy was achieved. In comparison, the classification CNN outperforms the conventional PCA (principal component analysis) plus SVM (support vector machine) method by 19%. In (Mary et al. 2018), images of three genera of cyanobacteria and three genera of green microalgae were employed to extract shape features, reduce them by using PCA,

and use them for training an ANN. Microalgae of these genera could then be recognized with an accuracy of 92 to 100%. In (Salido et al. 2020), diatoms serving as a bioindicator of water quality were detected and classified by using an automated system consisting of a microcomputer-controlled optical microscope scanning the manually fed slides and a PC processing the acquired images. A database (now publicly available) of annotated images of 80 diatom species was built for training purposes for three types of deep learning networks of the fully convolutional network type. A maximum precision of 86% for diatom detection and 99.5% for diatom classification was achieved.

In another recent study (Kaya et al. 2023), a system for the classification of microalgae based on deep learning techniques was designed. Images of 12 microalgae species were classified using 8 different neural network models, comprising logistic regression, RFR, XGBoost, five different deep learning networks, and combinations of these methods. 6,300 images of the 12 species were preprocessed and converted into a training dataset. The authors highlighted the improvement obtained in their results after the hybrid application of the models over their individual use, however, the best accuracy results were obtained with the deep learning models that incorporated auxiliary layers. With the best model, the success rate of classifying algae cells was 92%.

Additionally, image analysis coupled with machine learning techniques could also be employed in applications involving microalgae cell monitoring during their cultivation and could potentially be applied to the detection of contaminants in microalgae bioprocesses. Some works in these directions are listed below:

In (Otálora et al. 2021), two neural network-based models were developed and validated capable of distinguishing two microalgae species, *C.vulgaris* and *S.almeriensis*, in a mixed sample and calculating the proportion of each species in the sample. As input data, images acquired with a flow cytometry FlowCAM device and the image features provided by FlowCAM for each of the detected particles (e.g., particle diameter, area, edge gradient, red, green, blue RGB intensity, shape parameters, etc.) were employed. The first ANN was trained using the numerical values of these image features, the second ANN was trained using the images themselves, after

appropriate image processing. After training, neural network models could determine the fraction of each species in the mixture with  $R^2=0.99$  and a maximum error of 6.5%.

By using the chemometrical approach and employing the least absolute shrinkage and selection operator (LASSO) regression method, (Nguyen et al. 2023) developed a monitoring system with a software-based human–machine interface to effectively monitor the microalgae density (cell/mL). Color images of microalgae suspensions of *C. vulgaris* in a transparent vessel are acquired by means of a camera, average pixel brightness values and their confidence intervals are obtained, and further power spectra and entropy as a measure of the image texture in all three color channels (RGB) are extracted. These data are then employed, together with the measured microalgae density, as training data for the LASSO regression. The resulting model is then used for microalgae density monitoring. The root mean square error (RMSE) results of this study were compared, in which the LASSO approach presented a RMSE value of 1.54 against 2.16 and 3.68 of the other reported methods. It is important to point out that the strategy adopted employed the texture features of the images which provide information on the spatial distribution of the pixels.

A simple and non-invasive image-based approach using RGB analysis was used to quantitatively estimate the biofilm growth thickness of *Ettlia* sp. and its 3D topographic features (Asgharnejad et al. 2021). The presented method showed a preliminary investigation of a possible approach that could be converted into a software sensor for automated online monitoring of biofilm growth.

Online monitoring of cell count and cell morphology that could be used for online estimation of cultivation productivity was carried out by using the in situ microscope (Havlik et al. 2013b; Marba-Ardebol et al. 2019).

#### 4.2 Non-optical methods

Other alternatives using software sensors for the determination of process parameters in microalgae cultivation include non-optical measurement methods as model inputs.

A recent report (Paladino et al. 2022) indicates the use of data-driven model approaches for pH

prediction employing linear polynomial, nonlinear polynomial, and Hammerstein-Wiener models in fed-batch cultures of *C. vulgaris* grown mixotrophically. For the model development, measured values of wastewater flow rate, dilution flow rate, and temperature were used as the model inputs. Measured values of pH, dissolved oxygen and electrical conductivity were used as the model outputs. The best model obtained for pH prediction, i.e. nonlinear polynomial model, presented a relative RMSE=1.29% and  $R^2=0.975$  for both training and test data. In a second objective, the concentration of microalgae was estimated with two-step "grey-box" models using the pH estimation produced by the data-driven model as an input to a kinetic model predicting the biomass concentration. Both models were applied to process control operations.

A software sensor using EKF was used for online estimations of biomass and sulfur content (intracellular quota and extracellular sulfur concentration) in experiments aimed at producing hydrogen in *C. reinhardtii* under light attenuation and limited sulfur conditions. The inputs feeding the observer, i.e.,  $\text{CO}_2$  and  $\text{O}_2$  concentrations, were quantified at the outlet of the system and in an online fashion using a mass spectrometer (Daboussy et al. 2014).

#### 4.3 Combined methods (non-optical and optical methods)

The pH dynamics over 12 h were estimated using a data-driven approach in two raceway ponds using freshwater and wastewater cultures (Otálora et al. 2023). The input variables incorporated into both ANN models (called nonlinear autoregressive with exogenous inputs, NARX) were based on measurements of solar radiation (optic method), medium level, and  $\text{CO}_2$  injection, all relevant for the pH response. The models using fresh water and wastewater cultures showed a fit of 71.3% and 73.75%, as well as an MSE of 0.02 and 0.01. The authors highlight the applicability of the study not only for production operations with microalgae but also for general use in the description of dynamic biological models in biotechnology.

Also using an EKF, the biomass of *S. almeriensis*, cultivated in an industrial raceway photobioreactor, was estimated for real-time measurements. In this research, the experimental measurements of dissolved

oxygen, pH, gas injection (air flow and  $\text{CO}_2$  flow), and solar radiation were used as state estimator inputs (García-Mañas et al. 2019).

The lipid content was estimated using an adaptive interval observer approach in cultures of *Isochrysis galbana* (Mairet et al. 2014b). Also using an adaptive interval observer, the same authors estimated neutral lipids and the carbohydrate quota in *I. galbana* cultivation (Mairet et al. 2014a). In both cases, the nitrate ( $\text{NO}_3^-$ ) and nitrite ( $\text{NO}_2^-$ ) concentrations were measured with a Technicon Auto-analyzer coupled to an automated data acquisition system and the biomass (cell/mL) with an optical particle counter using the principle of light blockage (Mairet et al. 2011). These two inputs were used by the observer to monitor the variables already mentioned.

#### 4.4 Simulated software sensors

In an approach with simulated data, the online measurements of dissolved oxygen (inputs) were used to predict the biomass concentration of *C. reinhardtii* in a closed photobioreactor under autotrophic and continuous cultivation (Arango Restrepo et al. 2022). An EKF, an AO, and a hybrid observer (a combination of both) were tested. The hybrid model showed the best performance and accuracy when referring to biomass estimation. Prospective work will be directed to the evaluation of the model under real experimental conditions in the laboratory, as well as its possible use in the estimation of the concentration of mineral nutrients for the monitoring and control of the bioprocess.

The internal substrate quota  $q$ , required in the Droop model, was predicted using super twisting observers (STO) and an AO in microalgae cultivation (Coutinho et al. 2019). The STOs were developed for the estimation of the substrate uptake rate and the biomass growth rate based on simulated biomass and extracellular substrate concentration measurements. The AO estimated  $q$  by fusing the results from STOs predictions. This approach shows the design of an observer before the use of real measurement data. The authors highlight the robustness of the developed software sensor and the advantage of not requiring a complete knowledge of the model parameters.

A study reported the use of a new high-gain observer to predict the microalgae biomass concentration and the dissolved carbon dioxide (Farza et al. 2019). The simulated input data consists of the



average light intensity irradiated over the reactor. A second observer, i.e., a standard high gain observer, was constructed and compared against the high-gain observer, in which the latter showed superior results.

## 5 Challenges and perspectives of software sensors in monitoring microalgae cultivation

In 2012, a report on soft sensors in bioprocessing was issued by the European Federation of Biotechnology, which identified critical needs and issues for the successful development of soft sensor methods in bioprocess research and industry and concluded with a set of eight recommendations highlighting areas requiring development (Luttmann et al. 2012). These recommendations are fairly comprehensive, and the last ten years could not add anything substantial to them.

The general challenges (or needs, as called in the above reference) in the construction and implementation of software sensors in bioprocessing that are summarized in the above reference do not essentially differ from what a process engineer expects of every hardware sensor. These challenges concern mainly the operational performance of the soft sensor (long-term stability, short-term response, easy recalibration, highly reliable operation, highly automated, multi-analyte capacity) and the process economics (productivity monitoring, low maintenance and operational cost, moderate capital investment). There are additional benefits of software sensors stemming from their computational nature (i.e., they are programs running on digital computers that can be adjusted and extended at will) that hardware sensors cannot easily fulfill: they could monitor seed and raw material quality, monitor the variability of products and other analytes, and monitor deviations from expected process behavior. They could also be used for process optimizations, either in process parts or overall.

The specific challenges and perspectives for soft sensors in microalgal bioprocesses include several groups of needs. The most important is the identification of process variables or characteristics whose measurement is currently carried out by sampling and manual processing, such as lipid and pigment concentration, identification of contamination, cell morphology, cell viability, and cell count of unusually shaped microalgae (e.g., *Spirulina*). Automated

online estimation of these process variables by a soft sensor would enable the assessing of process state (e.g., process phase, onset of product formation, onset of microorganism stress) or impending dangers (contamination, loss of viability, unusual morphology or cell clogging). For some of these process variables, e.g., viability, online spectroscopic methods using software sensors for signal processing are already being developed (Rösner et al. 2022). The second group of needs is the selection of hardware measurement methods whose signals can be processed by a software sensor to yield an estimation of the desired process variable. Here, preferred are optical methods that can be constructed as noninvasive, e.g., spectroscopy as Raman, terahertz, FTIR, fluorescence, further derivative spectroscopy, color measurement and hyperspectral imaging (Méléder et al. 2013; Podevin et al. 2018; Liu et al. 2020b; Havlik et al. 2022). The third group includes issues partially specific to microalgae cultivation due to the properties of microalgae cells (e.g., biofilm formation) as well as questions of sensor cleaning that occur in other bioprocesses (Vanrolleghem and Lee 2003). On the computing side, algorithms used as state observers, or in other words, software sensors, are already well-developed, be they model-driven algorithms such as Luenberger observer and Kalman filter in their many variants and refinements, or data-driven algorithms starting from simple correlations over chemometric models to artificial neural networks (ANN), again in many variants and refinements. This is due to the universal nature of these algorithms, which can be employed in different engineering fields. The main problem in the implementation of software sensors lies in the selection of process outputs, which have to deliver information about process variables of interest in the sense of observability, and the selection of appropriate physical sensors measuring these outputs. If there are no such sensors available, the challenge lies in the development of such sensors, preferably noninvasive and therefore measuring mostly optical or electrical properties.

An interesting approach is the method of Digital Twins (Richter et al. 2023). A Digital Twin is a virtual copy of the bioprocess consisting of digital models of its individual functional components as sensors, upstream processing units (media and culture preparation, bioreactor with growth kinetics and product generation, harvesting), downstream processing units

(cell disruption, product extraction, purification and concentration, etc.), and process control components as controllers and actuators. With the help of this virtual process copy, rapid and cost-effective development of various process strategies can be carried out, e.g., for sensor selection, process control optimization, automation, the implementation of conventional or advanced controllers, etc. The selected process setup developed in a Digital Twin can then be realized physically. Digital Twins can also be employed as an effective training and educational tool (Appl et al. 2021).

Software sensors that include data-driven process models are closely connected to machine learning. In this respect, the FAIR concept (Findable, Accessible, Interoperable, Reusable) for archiving suitably structured scholarly data on (bio)processes in order to facilitate data access by data mining and learning tools should be mentioned (Wilkinson et al. 2016). It is suggested that all process data be archived in FAIR compliant form so that data-driven software sensors can be more easily trained.

In summary, the main drivers of future development of microalgal process monitoring are optical measurement methods such as light absorption, reflection and scattering in connection with software sensors and machine learning. These methods can monitor biomass, microorganism stress and lipid and pigment accumulation (Podevin et al. 2018; Liu et al. 2020b; Solovchenko 2023; Thiviyanathan et al. 2024). It is to be hoped that in the perspective of a few years (or decades?) the reliability level of software sensors in bioprocessing including microalgae will reach the level now enjoyed in chemical industries, even when the challenges due to the multiphase conditions of bioprocesses are a lot higher. The basic uncertainty factor, as usual, is the amount of money and effort invested into the software sensor development.

## 6 Concluding remarks

Microalgae offer a set of characteristics in line with sustainable production systems. In addition, they offer a plethora of products in the energy, food, and pharmaceutical sectors. However, microalgae cultivation poses interesting challenges to improving the economically viable support of sustainable large-scale

production at the present state of technology. Currently, the monitoring of biological variables in microalgae cultivation employs labor-intensive, time-consuming, non-immediate offline analytical methods. In some cases, online spectroscopic methods are used, mainly for biomass estimation, however, their use is not necessarily widespread. At the present time, the use of software sensors, either model-driven or data-driven estimators, in microalgae cultivation has been developed mainly with spectroscopic measurement techniques as input data and only a small part with non-optical methods. There also lies the perspective for further development of software sensors to include in the array of possible hardware inputs other than spectroscopic or optical sensors. Generally, it is not too difficult to add a hardware sensor input to a software sensor. Rapidly progressing miniaturization of computers, together with their ever-increasing computing power, will increase the number of one-task software sensors, be it only for complex online signal processing with noise filtering. Software sensors possess, with suitable design, the property of providing not only estimates of the current state of the measured process variables but also a near-future prediction of their values. In microalgae cultivation, a suitable time point for harvesting or other action could thus be predicted. Therefore, online sensors together with software sensors represent a promising future alternative for the improvement of microalgae monitoring, especially in monitoring those intracellular products for which no hardware online sensors presently exist, e.g., lipids or proteins. Software sensor development and in situ implementation could lead to substantial improvements in the development of sustainable processes based on microalgae cultivation.

It is noteworthy that the use of hybrid models for the development of software sensors as well as the monitoring of microalgae biofilm growth coupled with soft sensing represent interesting areas to be explored in microalgae cultivation. Likewise, the spectroscopic qualities of other microalgal biomolecules, apart from those reported in this review, represent appealing potential targets for the future development of software sensors. In general, developing and using advanced monitoring tools, such as software sensors, gives the process controller a better overview of how the process is performing and helps him or her decide on suitable actions.

**Acknowledgements** This work was supported by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement N° 955520 (DigitAlgaesation)".

**Author contributions** LPR literature search, data analysis, structure and draft process of the paper. IH literature search and data analysis, critically revised the paper. SB idea of the review paper, critically revised the paper.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

## Declarations

**Conflict of interest** All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abdi H, Williams LJ (2013) Partial least squares methods: partial least squares correlation and partial least square regression. In: Reisfeld B, Mayeno AN (eds) Computational toxicology, vol II. Humana Press, Totowa, NJ, pp 549–579
- Aguilar-Garnica E, García-Sandoval JP (2015) Software sensors design and selection for the production of biodiesel from grease trap wastes. In: Gernaey KV, Huusom JK, Gani R (eds) Computer aided chemical engineering. Elsevier, Great Britain, pp 1589–1594
- Andersen R (2005) Algal culturing techniques. Elsevier Academic Press, New York, p 578
- Andersen RA (2013) The microalgal cell. In: Handbook of Microalgal Culture, Second edition. Wiley, pp 1–20
- Appl C, Moser A, Baganz F, Hass VC (2021) Digital Twins for bioprocess control strategy development and realisation. In: Herwig C, Pörtner R, Möller J (eds) Digital Twins: applications to the design and optimization of bioprocesses. Springer International Publishing, Cham, pp 63–94
- Arango Restrepo JP, Titica M, Tebbani S (2022) A hybrid asymptotic-Kalman observer for estimation of micro-algae growth in a closed photobioreactor. In: 2022 10th International Conference on Systems and Control (ICSC). pp 323–328
- Asgharnejad H, Sarrafzadeh M-H, Abhar-Shegoftteh O et al (2021) Biomass quantification and 3-D topography reconstruction of microalgal biofilms using digital image processing. Algal Res 55:102243. <https://doi.org/10.1016/j.algal.2021.102243>
- Avagyan A (2018) Algae to energy and sustainable development technologies, resources, economics and system analyses. New design of global environmental policy and live conserve industry. First edition. CreateSpace, p 196
- Barkia I, Saari N, Manning SR (2019) Microalgae for high-value products towards human health and nutrition. Mar Drugs 17:304. <https://doi.org/10.3390/md17050304>
- Barsanti L, Gualtieri P (2014) Algae anatomy, biochemistry, and biotechnology, 2nd edn. CRC Press, Boca Ratón, p 326
- Becker EW (2007) Micro-algae as a source of protein. Biotechnol Adv 25:207–210. <https://doi.org/10.1016/j.biotechadv.2006.11.002>
- Beiroti A, Aghasadeghi MR, Hosseini SN, Norouzian D (2019) Application of recurrent neural network for online prediction of cell density of recombinant *Pichia pastoris* producing HBsAg. Prep Biochem Biotechnol 49:352–359. <https://doi.org/10.1080/10826068.2019.1566153>
- Bellamy-Carter J, Sound JK, Leney AC (2022) Probing heavy metal binding to phycobiliproteins. FEBS J 289:4646–4656. <https://doi.org/10.1111/febs.16396>
- Benavides M, Coutinho D, Hantson AL et al (2015) Robust Luenberger observers for microalgal cultures. J Process Control 36:55–63. <https://doi.org/10.1016/j.jprocont.2015.09.005>
- Benedetti M, Vecchi V, Barera S, Dall'Osto L, (2018) Biomass from microalgae: the potential of domestication towards sustainable biofactories. Microb Cell Fact 17:173. <https://doi.org/10.1186/s12934-018-1019-3>
- Bi Z, He B (2020) Chapter 13 - Biodiesel from microalgae. In: Jacob-Lopes E, Maroneze MM, Queiroz MI, Zepka LQ (eds) Handbook of microalgae-based processes and products. Academic Press, First edition, pp 329–371
- Biechele P, Busse C, Solle D et al (2015) Sensor systems for bioprocess monitoring. Eng Life Sci 15:469–488. <https://doi.org/10.1002/elsc.201500014>
- Bold H, Wynne M (1985) Introduction to the algae: structure and reproduction, 2nd edn. Prentice Hall, Englewood Cliffs, p 720
- Borowitzka MA, Moheimani NR (eds) (2013) Algae for biofuels and energy. Springer, Netherlands, Dordrecht
- Borowitzka MA, Beardall J, Raven JA (eds) (2016) The physiology of microalgae. Springer International Publishing, Cham
- Brunner V, Siegl M, Geier D, Becker T (2021) Challenges in the development of soft sensors for bioprocesses: a critical review. Front Bioeng Biotechnol. <https://doi.org/10.3389/fbioe.2021.722202>
- Canelli G, Abiusi F, Vidal Garcia A et al (2023) Amino acid profile and protein bioaccessibility of two *Galdieria sulphuraria* strains cultivated autotrophically and

- mixotrophically in pilot-scale photobioreactors. *Innov Food Sci Emerg Technol* 84:103287. <https://doi.org/10.1016/j.ifset.2023.103287>
- Caporgno MP, Mathys A (2018) Trends in microalgae incorporation into innovative food products with potential health benefits. *Front Nutr*. <https://doi.org/10.3389/fnut.2018.00058>
- Chanquia SN, Vernet G, Kara S (2021) Photobioreactors for cultivation and synthesis: specifications, challenges, and perspectives. *Eng Life Sci* 22:712–724. <https://doi.org/10.1002/elsc.202100070>
- Chen M (2014) Chlorophyll modifications and their spectral extension in oxygenic photosynthesis. *Annu Rev Biochem* 83:317–340. <https://doi.org/10.1146/annurev-biochem-072711-162943>
- Chirivella-Martorell J, Briz-Redon A, Serrano-Aroca A (2018) Modelling of biomass concentration, multi-wavelength absorption and discrimination method for seven important marine microalgae species. *Energies*. <https://doi.org/10.3390/en11051089>
- Chisti Y (2013) Constraints to commercialization of algal fuels. *J Biotechnol* 167:201–214. <https://doi.org/10.1016/j.jbiotec.2013.07.020>
- Claßen J, Aupert F, Reardon KF et al (2017) Spectroscopic sensors for in-line bioprocess monitoring in research and pharmaceutical industrial application. *Anal Bioanal Chem* 409:651–666. <https://doi.org/10.1007/s00216-016-0068-x>
- Coutinho D, Vargas A, Feudjio C et al (2019) A robust approach to the design of super-twisting observers—application to monitoring microalgae cultures in photobioreactors. *Comput Chem Eng* 121:46–56. <https://doi.org/10.1016/j.compchemeng.2018.07.006>
- Cybenko G (1989) Approximation by superpositions of a sigmoidal function. *Math Control Signal Syst* 2:303–314. <https://doi.org/10.1007/BF02551274>
- Daboussy M, Titica M, Boillereaux L (2014) Parameter identification and state estimation of a microalgae dynamical model in sulphur deprived conditions: global sensitivity analysis, optimization criterion, extended Kalman filter. *Can J Chem Eng* 92:1378–1395. <https://doi.org/10.1002/cjce.22007>
- Dagnino-Leone J, Figueroa CP, Castañeda ML et al (2022) Phycobiliproteins: structural aspects, functional characteristics, and biotechnological perspectives. *Comput Struct Biotechnol J* 20:1506–1527. <https://doi.org/10.1016/j.csbj.2022.02.016>
- de Assis AJ, Filho RM (2000) Soft sensors development for on-line bioreactor state estimation. *Comput Chem Eng* 24:1099–1103. [https://doi.org/10.1016/S0098-1354\(00\)00489-0](https://doi.org/10.1016/S0098-1354(00)00489-0)
- de Martínez-Roldán A, J, Cañizares-Villanueva RO, (2020) Chapter 7 - Wastewater treatment based in microalgae. In: Jacob-Lopes E, Maroneze MM, Queiroz MI, Zepka LQ (eds) *Handbook of microalgae-based processes and products*. Academic Press, First edition, pp 165–184
- Del Mondo A, Smerilli A, Sané E et al (2020) Challenging microalgal vitamins for human health. *Microb Cell Fact* 19:201. <https://doi.org/10.1186/s12934-020-01459-1>
- Deore P, Beardall J, Noronha S (2020) A perspective on the current status of approaches for early detection of microalgal grazing. *J Appl Phycol* 32:3723–3733. <https://doi.org/10.1007/s10811-020-02241-x>
- Diaz CJ, Douglas KJ, Kang K et al (2023) Developing algae as a sustainable food source. *Front Nutr*. <https://doi.org/10.3389/fnut.2022.1029841>
- Edelmann M, Aalto S, Chamlagain B et al (2019) Riboflavin, niacin, folate and vitamin B12 in commercial microalgae powders. *J Food Compos Anal*. <https://doi.org/10.1016/j.jfca.2019.05.009>
- Elson EL (2011) Fluorescence correlation spectroscopy: past, present, future. *Biophys J* 101:2855–2870. <https://doi.org/10.1016/j.bpj.2011.11.012>
- Faassen SM, Hitzmann B (2015) Fluorescence spectroscopy and chemometric modeling for bioprocess monitoring. *Sensors* 15:10271–10291. <https://doi.org/10.3390/s150510271>
- Fabris M, Abbriano RM, Pernice M et al (2020) Emerging technologies in algal biotechnology: toward the establishment of a sustainable, algae-based bioeconomy. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2020.00279>
- Farza M, Rodriguez-Mata AE, Robles-Magdaleno J, M'Saad M (2019) A new filtered high gain observer design for the estimation of the components concentrations in a photobioreactor in microalgae culture. *IFAC-PapersOn-Line* 52:904–909. <https://doi.org/10.1016/j.ifacol.2019.06.177>
- Fernandes T, Cordeiro N (2021) Microalgae as sustainable biofactories to produce high-value lipids: biodiversity, exploitation, and biotechnological applications. *Mar Drugs* 19:573. <https://doi.org/10.3390/md19100573>
- Ferro L, Gojkovic Z, Gorzsa A, Funk C (2019) Statistical methods for rapid quantification of proteins, lipids, and carbohydrates in nordic microalgal species using ATR-FTIR spectroscopy. *Molecules*. <https://doi.org/10.3390/molecules24183237>
- Flores G, Rodriguez-Mata AE, Amabilis-Sosa LE et al (2020) A turbidity sensor development based on NL-PI observers: experimental application to the control of a Sinaloa's River *Spirulina maxima* cultivation. *Open Chem* 18:1349–1361. <https://doi.org/10.1515/chem-2020-0119>
- Franco BM, Navas LM, Gomez C et al (2019) Monoalgal and mixed algal cultures discrimination by using an artificial neural network. *Algal Res*. <https://doi.org/10.1016/j.algal.2019.101419>
- Gao FZ, Sa M, Teles I et al (2021) Production and monitoring of biomass and fucoxanthin with brown microalgae under outdoor conditions. *Biotechnol Bioeng* 118:1355–1365. <https://doi.org/10.1002/bit.27657>
- García JL, de Vicente M, Galán B (2017) Microalgae, old sustainable food and fashion nutraceuticals. *Microb Biotechnol* 10:1017–1024. <https://doi.org/10.1111/1751-7915.12800>
- García-Mañas F, Guzmán JL, Berenguel M, Acien FG (2019) Biomass estimation of an industrial raceway photobioreactor using an extended Kalman filter and a dynamic model for microalgae production. *Algal Res* 37:103–114. <https://doi.org/10.1016/j.algal.2018.11.009>
- Gauthier JP, Hammouri H, Othman S (1992) A simple observer for nonlinear systems applications to bioreactors. *IEEE Trans Autom Control* 37:875–880. <https://doi.org/10.1109/9.256352>



- Gerón A (2019) Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow. Second edition. O'Reilly Media, Sebastopol
- Ghosh T, Mondal A, Vyas A, Mishra S (2020) A 'one-tube' synthesis of a selective fluorescence 'turn off/on' DNA probe based on a C-phycocyanin-graphene oxide (CPC-GO) bio composite. *Int J Biol Macromol* 163:977–984. <https://doi.org/10.1016/j.ijbiomac.2020.06.286>
- Goffaux G, Wouwer AV, Bernard O (2009) Continuous - discrete interval observers for monitoring microalgae cultures. *Biotechnol Prog* 25:667–675. <https://doi.org/10.1002/btpr.167>
- Gouda M, Tadda MA, Zhao Y et al (2022) Microalgae bioactive carbohydrates as a novel sustainable and eco-friendly source of prebiotics: emerging health functionality and recent technologies for extraction and detection. *Front Nutr*. <https://doi.org/10.3389/fnut.2022.806692>
- Gouveia L, Oliveira AC (2009) Microalgae as a raw material for biofuels production. *J Ind Microbiol Biotechnol* 36:269–274. <https://doi.org/10.1007/s10295-008-0495-6>
- Havlik I, Lindner P, Scheper T, Reardon KF (2013a) On-line monitoring of large cultivations of microalgae and cyanobacteria. *Trends Biotechnol* 31:406–414. <https://doi.org/10.1016/j.tibtech.2013.04.005>
- Havlik I, Reardon KF, Unal M et al (2013b) Monitoring of microalgal cultivations with on-line, flow-through microscopy. *Algal Res* 2:253–257. <https://doi.org/10.1016/j.algal.2013.04.001>
- Havlik I, Beutel S, Scheper T, Reardon KF (2022) On-line monitoring of biological parameters in microalgal bioprocesses using optical methods. *Energies* 15:875. <https://doi.org/10.3390/en15030875>
- Hui Y, Xue X, Song ZX et al (2016) Study of fluorescence spectra of thiamine and riboflavin. *MATEC Web Conf* 63:03013. <https://doi.org/10.1051/mateconf/20166303013>
- Ishikawa-Ankerhold HC, Ankerhold R, Drummen GPC (2012) Advanced fluorescence microscopy techniques—FRAP, FLIP, FLAP, FRET and FLIM. *Molecules* 17:4047–4132. <https://doi.org/10.3390/molecules17044047>
- Jacob-Lopes E, Maroneze MM, Queiroz MI, Zepka LQ (eds) (2020) Handbook of microalgae-based processes and products, First edition. Elsevier Academic Press
- Jung F, Krüger-Genge A, Waldeck P, Küpper J-H (2019) *Spirulina platensis*, a super food? *J Cellular Biotechnol* 5:43–54. <https://doi.org/10.3233/JCB-189012>
- Kadlec P, Gabrys B, Strandt S (2009) Data-driven soft sensors in the process industry. *Comput Chem Eng* 33:795–814. <https://doi.org/10.1016/j.compchemeng.2008.12.012>
- Kalman RE (1960) A new approach to linear filtering and prediction problems. *J Basic Eng* 82:35–45. <https://doi.org/10.1115/1.3662552>
- Kalman RE, Bucy RS (1961) New results in linear filtering and prediction theory. *J Basic Eng* 83:95–108. <https://doi.org/10.1115/1.3658902>
- Kaya V, Akgül İ, Zencir Tanır Ö (2023) A novel hybrid model based on machine and deep learning techniques for the classification of microalgae. *Phyton* 0:1–16. <https://doi.org/10.32604/phyton.2023.029811>
- Kovaleski G, Kholany M, Dias LMS et al (2022) Extraction and purification of phycobiliproteins from algae and their applications. *Front Chem*. <https://doi.org/10.3389/fchem.2022.1065355>
- Kumar K (2021) Partial least square (PLS) analysis. *Reson* 26:429–442. <https://doi.org/10.1007/s12045-021-1140-1>
- Lang I, Hodac L, Friedl T, Feussner I (2011) Fatty acid profiles and their distribution patterns in microalgae: a comprehensive analysis of more than 2000 strains from the SAG culture collection. *BMC Plant Biol* 11:124. <https://doi.org/10.1186/1471-2229-11-124>
- Langi P, Kiokias S, Varzakas T, Proestos C (2018) Carotenoids: from plants to food and feed industries. *Methods Mol Biol* 1852:57–71. [https://doi.org/10.1007/978-1-4939-8742-9\\_3](https://doi.org/10.1007/978-1-4939-8742-9_3)
- Li Y, Chen M (2015) Novel chlorophylls and new directions in photosynthesis research. *Funct Plant Biol* 42:493. <https://doi.org/10.1071/FP14350>
- Li X, Liao R, Zhou J et al (2017) Classification of morphologically similar algae and cyanobacteria using Mueller matrix imaging and convolutional neural networks. *Appl Opt*, AO 56:6520–6530. <https://doi.org/10.1364/AO.56.006520>
- Lieutaud C, Assaf A, Goncalves O et al (2019) Fast non-invasive monitoring of microalgal physiological stage in photobioreactors through Raman spectroscopy. *Algal Res*. <https://doi.org/10.1016/j.algal.2019.101595>
- Liu JY, Zeng LH, Ren ZH et al (2020a) Rapid *in situ* measurements of algal cell concentrations using an artificial neural network and single-excitation fluorescence spectrometry. *Algal Res*. <https://doi.org/10.1016/j.algal.2019.101739>
- Liu JY, Zeng LH, Ren ZH (2020b) Recent application of spectroscopy for the detection of microalgae life information: a review. *Appl Spectrosc Rev* 55:26–59. <https://doi.org/10.1080/05704928.2018.1509345>
- López Expósito P, Blanco Suárez A, Negro C (2016) Estimation of *Chlamydomonas reinhardtii* biomass concentration from chord length distribution data. *J Appl Phycol* 28:2315–2322. <https://doi.org/10.1007/s10811-015-0749-4>
- López Expósito P, Blanco Suárez A, Negro Álvarez C (2017) Laser reflectance measurement for the online monitoring of *Chlorella sorokiniana* biomass concentration. *J Biotechnol* 243:10–15. <https://doi.org/10.1016/j.jbiotec.2016.12.020>
- Luenberger D (1966) Observers for multivariable systems. *IEEE Trans Autom Control* 11:190–197. <https://doi.org/10.1109/TAC.1966.1098323>
- Luttmann R, Bracewell DG, Cornelissen G et al (2012) Soft sensors in bioprocessing: a status report and recommendations. *Biotechnol J* 7:1040–1048. <https://doi.org/10.1002/biot.201100506>
- Mairet F, Bernard O, Masci P et al (2011) Modelling neutral lipid production by the microalga *Isochrysis* aff. *galbana* under nitrogen limitation. *Biores Technol* 102:142–149. <https://doi.org/10.1016/j.biortech.2010.06.138>
- Mairet F, Moisan M, Bernard O (2014a) Estimation of neutral lipid and carbohydrate quotas in microalgae using adaptive interval observers. *Bioprocess Biosyst Eng* 37:51–61. <https://doi.org/10.1007/s00449-013-0913-7>



- Mairet F, Moisan M, Bernard O (2014b) Interval observer with near optimal adaptation dynamics. Application to the estimation of lipid quota in microalgae. *Int J Robust Nonlinear Control* 24:1142–1157. <https://doi.org/10.1002/rnc.2934>
- Mandenius C-F, Gustavsson R (2015) Mini-review: soft sensors as means for PAT in the manufacture of bio-therapeutics. *J Chem Technol Biotechnol* 90:215–227. <https://doi.org/10.1002/jctb.4477>
- Marba-Ardebol AM, Emmerich J, Muthig M et al (2019) *In situ* microscopy for real-time determination of single-cell morphology in bioprocesses. *J vis Exp*. <https://doi.org/10.3791/57823>
- Markou G, Angelidaki I, Georgakakis D (2012) Microalgal carbohydrates: an overview of the factors influencing carbohydrates production, and of main bioconversion technologies for production of biofuels. *Appl Microbiol Biotechnol* 96:631–645. <https://doi.org/10.1007/s00253-012-4398-0>
- Mary AVA, Prabakaran G, Mohan S (2018) Freshwater microalga image identification and classification based on machine learning technique. *Asian J Comput Sci Technol* 7:63–67
- Masojídek J, Sergejevojá M, Rottnerová K et al (2009) A two-stage solar photobioreactor for cultivation of microalgae based on solar concentrators. *J Appl Phycol* 21:55–63. <https://doi.org/10.1007/s10811-008-9324-6>
- Masojídek J, Torzillo G, Koblížek M (2013) Photosynthesis in microalgae. In: *Handbook of Microalgal Culture*, Second edition. John Wiley & Sons, Ltd, pp 21–36
- Mélédér V, Laviale M, Jesus B et al (2013) *In vivo* estimation of pigment composition and optical absorption cross-section by spectroradiometry in four aquatic photosynthetic micro-organisms. *J Photochem Photobiol, B* 129:115–124. <https://doi.org/10.1016/j.jphotobiol.2013.10.005>
- Milani A, Basirnejad M, Shahbazi S, Bolhassani A (2017) Carotenoids: biochemistry, pharmacology and treatment. *Br J Pharmacol* 174:1290–1324. <https://doi.org/10.1111/bph.13625>
- Morales M, Aflalo C, Bernard O (2021) Microalgal lipids: a review of lipids potential and quantification for 95 phytoplankton species. *Biomass Bioenerg* 150:106108. <https://doi.org/10.1016/j.biombioe.2021.106108>
- Mowbray M, Savage T, Wu CF et al (2021) Machine learning for biochemical engineering: a review. *Biochem Eng J*. <https://doi.org/10.1016/j.bej.2021.108054>
- Nguyen L, Nguyen DK, Nguyen T et al (2023) Analysis of microalgal density estimation by using LASSO and image texture features. *Sensors* 23:2543. <https://doi.org/10.3390/s23052543>
- Ning H, Li R, Zhou T (2022) Machine learning for microalgae detection and utilization. *Front Marine Sci*. <https://doi.org/10.3389/fmars.2022.947394>
- Noll P, Henkel M (2020) History and evolution of modeling in biotechnology: modeling and simulation, application and hardware performance. *Comp Struct Biotechnol J* 18:3309–3323. <https://doi.org/10.1016/j.csbj.2020.10.018>
- Norsker N-H, Barbosa MJ, Vermuë MH, Wijffels RH (2011) Microalgal production—A close look at the economics. *Biotechnol Adv* 29:24–27. <https://doi.org/10.1016/j.biotechadv.2010.08.005>
- Otálora P, Guzmán JL, Acién FG et al (2021) Microalgae classification based on machine learning techniques. *Algal Res* 55:102256. <https://doi.org/10.1016/j.algal.2021.102256>
- Otálora P, Guzmán JL, Berenguel M, Acién FG (2023) Data-driven pH model in raceway reactors for freshwater and wastewater cultures. *Mathematics* 11:1614. <https://doi.org/10.3390/math11071614>
- Paladino O, Neviani M, Ciano D, De Francesco M (2022) Prediction of pH and microalgae growth in mixotrophic conditions by nonlinear black-box models for control purposes. *Biomass Conv Bioref*. <https://doi.org/10.1007/s13399-022-03634-3>
- Papageorgiou GC, Govindjee, (eds) (2004) *Chlorophyll a fluorescence: a signature of photosynthesis*, 1st edn. Springer, Netherlands, Dordrecht
- Park JS, Chyun JH, Kim YK et al (2010) Astaxanthin decreased oxidative stress and inflammation and enhanced immune response in humans. *Nutr Metab (lond)* 7:18. <https://doi.org/10.1186/1743-7075-7-18>
- Parri E, Santinami G, Domenici V (2020) Front-face fluorescence of honey of different botanic origin: a case study from Tuscany (Italy). *Appl Sci* 10:1776. <https://doi.org/10.3390/app10051776>
- Perera YS, Ratnaweera DAAC, Dasanayaka CH, Abeykoon C (2023) The role of artificial intelligence-driven soft sensors in advanced sustainable process industries: a critical review. *Eng Appl Artif Intell* 121:105988. <https://doi.org/10.1016/j.engappai.2023.105988>
- Ploix S, Maquin D, Ragot J (1997) Asymptotic observer for a non-linear descriptor system. *IFAC Proceedings Volumes* 30:639–643. [https://doi.org/10.1016/S1474-6670\(17\)42472-4](https://doi.org/10.1016/S1474-6670(17)42472-4)
- Podevin M, Fotidis IA, Angelidaki I (2018) Microalgal process-monitoring based on high-selectivity spectroscopy tools: status and future perspectives. *Crit Rev Biotechnol* 38:704–718. <https://doi.org/10.1080/07388551.2017.1398132>
- Randek J, Mandenius CF (2018) On-line soft sensing in upstream bioprocessing. *Crit Rev Biotechnol* 38:106–121. <https://doi.org/10.1080/07388551.2017.1312271>
- Rathore AS, Mishra S, Nikita S, Priyanka P (2021) Bioprocess control: current progress and future perspectives. *Life* 11:557. <https://doi.org/10.3390/life11060557>
- Reyes SJ, Durocher Y, Pham PL, Henry O (2022) Modern sensor tools and techniques for monitoring, controlling, and improving cell culture processes. *Processes* 10:189. <https://doi.org/10.3390/pr10020189>
- Richardson JW, Johnson MD, Zhang X et al (2014) A financial assessment of two alternative cultivation systems and their contributions to algae biofuel economic viability. *Algal Res* 4:96–104. <https://doi.org/10.1016/j.algal.2013.12.003>
- Richter J, Lange F, Scheper T et al (2023) Digitale Zwillinge in der Bioprozesstechnik—Chancen und Möglichkeiten. *Chem Ing Tec* 95:498–510. <https://doi.org/10.1002/cite.202200166>

- Ringnér M (2008) What is principal component analysis? *Nat Biotechnol* 26:303–304. <https://doi.org/10.1038/nbt0308-303>
- Roberts J, Power A, Chapman J et al (2018) The use of UV-Vis spectroscopy in bioprocess and fermentation monitoring. *Fermentation* 4:18. <https://doi.org/10.3390/fermentation4010018>
- Roche JL, van der Staay GWM, Partensky F et al (1996) Independent evolution of the prochlorophyte and green plant chlorophyll a/b light-harvesting proteins. *Proc Natl Acad Sci* 93:15244–15248. <https://doi.org/10.1073/pnas.93.26.15244>
- Rösner LS, Walter F, Ude C et al (2022) Sensors and techniques for on-line determination of cell viability in bioprocess monitoring. *Bioengineering* 9:762. <https://doi.org/10.3390/bioengineering9120762>
- Sá M, Monte J, Brazinha C et al (2017) 2D Fluorescence spectroscopy for monitoring *Dunaliella salina* concentration and integrity during membrane harvesting. *Algal Res* 24:325–332. <https://doi.org/10.1016/j.algal.2017.04.013>
- Sá M, Bertinetto CG, Ferrer-Ledo N et al (2020a) Fluorescence spectroscopy and chemometrics for simultaneous monitoring of cell concentration, chlorophyll and fatty acids in *Nannochloropsis oceanica*. *Sci Rep* 10:7688. <https://doi.org/10.1038/s41598-020-64628-7>
- Sá M, Ramos A, Monte J et al (2020b) Development of a monitoring tool based on fluorescence and climatic data for pigments profile estimation in *Dunaliella salina*. *J Appl Phycol* 32:363–373. <https://doi.org/10.1007/s10811-019-01999-z>
- Salido J, Sánchez C, Ruiz-Santaquiteria J et al (2020) A low-cost automated digital microscopy platform for automatic identification of diatoms. *Appl Sci* 10:6033. <https://doi.org/10.3390/app10176033>
- Santos-Ballardo DU, Rossi S, Hernandez V et al (2015) A simple spectrophotometric method for biomass measurement of important microalgae species in aquaculture. *Aquaculture* 448:87–92. <https://doi.org/10.1016/j.aquaculture.2015.05.044>
- Sarker SD, Nahar L (2015) Chapter 19 - Applications of high performance liquid chromatography in the analysis of herbal products. In: Mukherjee PK (ed) Evidence-based validation of herbal medicine. Elsevier, Boston, pp 405–425
- Sarpal AS, Teixeira C, Silva PRM et al (2015) Determination of lipid content of oleaginous microalgal biomass by NMR spectroscopic and GC-MS techniques. *Anal Bioanal Chem* 407:3799–3816. <https://doi.org/10.1007/s00216-015-8613-6>
- Sauer JS, Simkovsky R, Moore AN et al (2021) Continuous measurements of volatile gases as detection of algae crop health. *Proc Natl Acad Sci USA*. <https://doi.org/10.1073/pnas.210688211811of9>
- Shen P-L, Wang H-T, Pan Y-F et al (2016) Identification of characteristic fatty acids to quantify triacylglycerols in microalgae. *Front Plant Sci*. <https://doi.org/10.3389/fpls.2016.00162>
- Silva SC, Ferreira ICFR, Dias MM, Barreiro MF (2020) Microalgae-derived pigments: a 10-year bibliometric review and industry and market trend analysis. *Molecules* 25:E3406. <https://doi.org/10.3390/molecules25153406>
- Simutis R, Havlik I, Lubbert A (1992) A fuzzy-supported supported extended Kalman filter: a new approach to state estimation and prediction exemplified by alcohol formation in beer brewing. *J Biotechnol* 24:211–234. [https://doi.org/10.1016/0168-1656\(92\)90033-6](https://doi.org/10.1016/0168-1656(92)90033-6)
- Sode K, Horikoshi K, Takeyama H et al (1991) On-line monitoring of marine cyanobacterial cultivation based on phycocyanin fluorescence. *J Biotechnol* 21:209–217. [https://doi.org/10.1016/0168-1656\(91\)90042-T](https://doi.org/10.1016/0168-1656(91)90042-T)
- Solle D, Hitzmann B, Herwig C et al (2017) Between the poles of data-driven and mechanistic modeling for process operation. *Chem Ing Tec* 89:542–561. <https://doi.org/10.1002/cite.201600175>
- Solovchenko A (2023) Seeing good and bad: optical sensing of microalgal culture condition. *Algal Res* 71:103071. <https://doi.org/10.1016/j.algal.2023.103071>
- Stadnichuk IN, Tropin IV (2017) Phycobiliproteins: structure, functions and biotechnological applications. *Appl Biochem Microbiol* 53:1–10. <https://doi.org/10.1134/S0003683817010185>
- Stadnichuk IN, Krasilnikov PM, Zlenko DV (2015) Cyanobacterial phycobilisomes and phycobiliproteins. *Microbiol* 84:101–111. <https://doi.org/10.1134/S0026261715020150>
- Sulaiman L, Abidin ZHZ, Yunus SAM et al (2019) Dual functions of Cu(NO<sub>3</sub>)<sub>2</sub> as antifungal and colour stabilizer for coating paint film consisting of chlorophyll. *Pigm Resin Technol* 50:177–182. <https://doi.org/10.1108/PRT-11-2017-0086>
- Suparmi S, Fasitasari M, Martosupono M, Mangimbulude JC (2016) Comparisons of curative effects of chlorophyll from *Sauropus androgynus* (L) Merr leaf extract and Cu-chlorophyllin on sodium nitrate-induced oxidative stress in rats. *J Toxicol* 2016:e8515089. <https://doi.org/10.1155/2016/8515089>
- Tan JS, Lee SY, Chew KW et al (2020) A review on microalgae cultivation and harvesting, and their biomass extraction processing using ionic liquids. *Bioengineered* 11:116–129. <https://doi.org/10.1080/21655979.2020.1711626>
- Thibault J, Van Breusegem V, Chérut A (1990) On-line prediction of fermentation variables using neural networks. *Biotechnol Bioeng* 36:1041–1048. <https://doi.org/10.1002/bit.260361009>
- Thiviyanathan VA, Ker PJ, Hoon Tang SG et al (2024) Microalgae biomass and biomolecule quantification: optical techniques, challenges and prospects. *Renew Sustain Energy Rev* 189:113926. <https://doi.org/10.1016/j.rser.2023.113926>
- Timira V, Meki K, Li Z et al (2022) A comprehensive review on the application of novel disruption techniques for proteins release from microalgae. *Crit Rev Food Sci Nutr* 62:4309–4325. <https://doi.org/10.1080/10408398.2021.1873734>
- Tounsi L, Ben Hlima H, Hentati F et al (2023) Microalgae: a promising source of bioactive phycobiliproteins. *Mar Drugs* 21:440. <https://doi.org/10.3390/md21080440>
- Turcu CO, Turcu CE (2018) Industrial Internet of Things as a challenge for higher education. *Int J Adv Comput Sci Appl* 9(11):55–60

- Udayan A, Arumugam M, Pandey A (2017) Chapter 4 - Nutra-ceuticals from algae and cyanobacteria. In: Rastogi RP, Madamwar D, Pandey A (eds) *Algal green chemistry*. Elsevier, Amsterdam, pp 65–89
- Vanrolleghem PA, Lee DS (2003) On-line monitoring equipment for wastewater treatment processes: state of the art. *Water Sci Technol* 47:1–34. <https://doi.org/10.2166/wst.2003.0074>
- Vecchi V, Barera S, Bassi R, Dall'Osto L (2020) Potential and challenges of improving photosynthesis in algae. *Plants (basel)* 9:67. <https://doi.org/10.3390/plants9010067>
- Wang Q, Peng H, Higgins BT (2019) Cultivation of green microalgae in bubble column photobioreactors and an assay for neutral lipids. *J vis Exp*. <https://doi.org/10.3791/59106>
- Wechselberger P, Sagmeister P, Herwig C (2013) Real-time estimation of biomass and specific growth rate in physiologically variable recombinant fed-batch processes. *Bioprocess Biosyst Eng* 36:1205–1218. <https://doi.org/10.1007/s00449-012-0848-4>
- Wilkinson MD, Dumontier M, Ij A et al (2016) The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3:160018. <https://doi.org/10.1038/sdata.2016.18>
- Williams PJLB, Laurens LM (2010) Microalgae as biodiesel & biomass feedstocks: Review & analysis of the biochemistry, energetics & economics. *Energy Environ Sci* 3(5):554–590. <https://doi.org/10.1039/B924978H>
- Wold S, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. *Chemom Intell Lab Syst* 58:109–130. [https://doi.org/10.1016/S0169-7439\(01\)00155-1](https://doi.org/10.1016/S0169-7439(01)00155-1)
- Xu L, Xu L, Chen Y et al (2022) Accurate classification of algae using deep convolutional neural network with a small database. *ACS EST Water* 2:1921–1928. <https://doi.org/10.1021/acsestwater.1c00466>
- Yang Y, Du L, Hosokawa M, Miyashita K (2020) Total lipids content, lipid class and fatty acid composition of ten species of microalgae. *J Oleo Sci* 69:1181–1189. <https://doi.org/10.5650/jos.ess20140>
- Yao L, Gerde JA, Lee S-L et al (2015) Microalgae lipid characterization. *J Agric Food Chem* 63:1773–1787. <https://doi.org/10.1021/jf5050603>
- Yoo SJ, Jung DH, Kim JH, Lee JM (2015) A comparative study of soft sensor design for lipid estimation of microalgal photobioreactor system with experimental validation. *Bioresour Technol* 179:275–283. <https://doi.org/10.1016/j.biortech.2014.12.029>
- Yoo SJ, Jeong DH, Kim JH, Lee JM (2016) Optimization of microalgal photobioreactor system using model predictive control with experimental validation. *Bioprocess Biosyst Eng* 39:1235–1246. <https://doi.org/10.1007/s00449-016-1602-0>
- Yoshida H, Yanai H, Ito K et al (2010) Administration of natural astaxanthin increases serum HDL-cholesterol and adiponectin in subjects with mild hyperlipidemia. *Atherosclerosis* 209:520–523. <https://doi.org/10.1016/j.atherosclerosis.2009.10.012>
- You Y, Cheng S, Zhang L et al (2020) Rational modulation of the luminescence of upconversion nanomaterials with phycocyanin for the sensing and imaging of myeloperoxidase during an inflammatory process. *Anal Chem* 92:5091–5099. <https://doi.org/10.1021/acs.analchem.9b05468>
- Yousefi-Darani A, Paquet-Durand O, Hitzmann B (2021) The Kalman filter for the supervision of cultivation processes. *Adv Biochem Eng Biotechnol* 177:95–125. [https://doi.org/10.1007/10\\_2020\\_145](https://doi.org/10.1007/10_2020_145)
- Zapata M, Garrido JL, Jeffrey SW (2006) Chlorophyll c pigments: current status. In: Grimm B, Porra RJ, Rüdiger W, Scheer H (eds) *Chlorophylls and bacteriochlorophylls: biochemistry, biophysics, functions and applications*, 1st edn. Springer, Netherlands, Dordrecht, pp 39–53
- Zittelli GC, Biondi N, Rodolfi L, Tredici MR (2013) Photobioreactors for mass production of microalgae. In: *Handbook of Microalgal Culture*, Second edition. John Wiley and Sons, Ltd, pp 225–266

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.