Monoalgal and mixed algal cultures discrimination by using an artificial neural network

B.M. Franco[a],[*], L.M. Navas[a], C. Gómez[b], C. Sepúlveda[b], F.G. Acién[b]

[a] *Department of Agricultural and Forestry Engineering, University of Valladolid, 34005 Palencia, Spain*
[b] *Department of Engineering, University of Almeria, 04120 Almeria, Spain*

ABSTRACT

In this paper, a rapid methodology to elucidate microalgae species in suspensions has been developed and validated. To do this, microalgae spectral signatures from light absorption measurements of different algal species were analysed through an artificial neural network (ANN) in order to describe and classify them. Four important species were used: *Nostoc* sp., *Scenedesmus almeriensis*, *Spirulina platensis* and *Chorella vulgaris*. Absorbance from monoalgal and mixed algal cultures was the input data for training, testing and validating the ANN. The results show that the ANN was capable of distinguishing between monoalgal and mixed algal cultures, identifying the microalgae species in the monoalgal cultures and providing the approximate composition of mixed algal cultures. These results confirm that the application of spectral signatures with ANN is a suitable method for approximating the biological composition of microalgae cultures.

1. Introduction

Microalgae have been proposed for a wide range of applications, from the production of foods and animal feed, cosmetics, biofuels and wastewater treatment processes [1–4]. Despite the large variety of applications proposed, only a few are presently performed at the commercial scale using a limited number of algal strains. Examples of this are the production of carotenoids, beta-carotene and astaxanthin from *Dunaliella salina* and *Haematococcus pluvialis* [5], biomass for foods from *Chlorella vulgaris* and *Spirulina platensis* [6], and biomass for aquaculture from *Nannochloropsis gaditana*, *Tetraselmis suecica* and *Isochrysis galbana* T-ISO [7]. All of these applications require the maintenance of monoalgal cultures, whereas when focusing on biofuel production or wastewater treatment, the utilization of mixed cultures is usually acceptable. In the case of mixed algal cultures, the relative composition is gradually modified according to changes in environmental or operational conditions [8].

Monitoring the biological composition of microalgae cultures is a mandatory task, generally performed through routine microscopic examination. By means of light microscopy, an expert can distinguish the presence of a "contaminating" microorganism and whether the prevailing algal strain is close to the expected value. However, only highly skilled taxonomists are capable of correctly recognizing algal strains (their species and genera) using light microscopy observation solely based on morphology – this is because most of the strains are small round cells with similar features, only a few have easily recognizable morphology. Alternative methods, based on omics allow to accurately identify the prevailing microalgae strains in cultures [9] but these are expensive and require a lot of time (reducing time can be useful for making operation process decisions). As intermediate methods, biochemical analyses, such as the chlorophyll to carotenoid ratio and the fatty-acid profile have also been used as tools for verifying the biological composition of microalgae cultures; however, their precision is limited. These methods are unable to identify the presence of low-level contamination; likewise, they take time, although their cost is much lower than that for the omics methods [10,11].

Microalgae strain identification up to the phylum or class levels has been achieved based on the fluorescence properties of photosynthetic pigments using flow cytometry [12]. Employing light-emitting diode induced fluorescence analysis, it is possible to differentiate between *Anabaena* sp. and *Cylindrospermum* sp. cells by comparing the fluorescence spectra [13]. Microalgae have photosynthetic pigments providing different spectral signatures for different strains; thus, it is possible to build classes based on the presence of pigments. As a result, the relative content of chlorophyll, carotenoids and other pigments can be used to differentiate between diatoms, red/green/brown microalgae and cyanobacteria groups by means of their light absorption spectra [10].

Microalgae species can be distinguished by their spectral signature.

As an example, *Rivularia* M-216 exhibits a different absorbance signature to that of *Anabaena variabilis* - the heterocyst absorbance from *Rivularia* is more than double that from *A. variabilis* at wavelengths between 540 and 620 nm; this variation is a result of the different phycocyanin and chlorophyll contents [14]. The spectral signatures of *Botryococcus braunii*, *Chlorella* sp. and *Chlorococcum littorale* allow to clearly identify them by comparing their absorption indexes [15]. Moreover, the absorption spectrum in the 400 to 700 nm range is used to determine the extinction coefficient of the biomass, such as a specific microalgae strain characteristic or the culture conditions [16]. Based on this, absorption properties could be a possible approach to distinguish between microalgae species [17].

Artificial neural networks (ANNs) are a powerful tool for finding relationships between experimental data and the phenomena behind these data. ANNs have been used to predict harmful algal blooms in lakes [18,19], as well as microalgae growth and biomass concentration under laboratory conditions and in outdoor environments [20,21]. In microalgae identification using ANN, extracted features such as the perimeter, shape, area and Fourier transformations of microalgae micrographics were used to train a model capable of identifying the genera *Navicula*, *Scenedesmus*, *Microcystis*, *Oscillatoria* and *Chroococcus* [22]. Microalgae micrographic image processing and colour analysis were used with ANN to achieve taxonomic accuracy of up to 99% by first detecting a cell in the image and subsequently extracting the detected cell colour [17].

In this paper, we analysed the ability of ANN to differentiate between monoalgal and mixed algal cultures. To do this, we used only the light absorption spectra of the different cultures, without any additional equipment or image analysis algorithms. The objective of this work is to determine the feasibility of using such a methodology to elucidate the presence of contaminants in monoalgal microalgae cultures and to approximate the microalgae species contaminating the cultures - this would provide a fast and powerful tool for microalgae culture management at the commercial scale.

2. Materials and methods

2.1. Microalgae species and growth conditions

The microalgae *Nostoc* sp., *Scenedesmus almeriensis*, *Spirulina platensis* and *Chorella vulgaris* were obtained from the culture collection of "Las Palmerillas" Research Center (Almería, Spain). These species were selected because they include two fast growing green microalgae - *Chlorella vulgaris* and *Scenedesmus almeriensis*, and two distinct cyanobacteria - *Nostoc* sp. and *Spirulina platensis*; therefore, they cover a wide range of possible microalgae species. Continuous cultures of the different microalgae were cultivated under laboratory conditions in 2 L bubble-column photobioreactors at an irradiance of $750\,\mu E \cdot m^{-2} \cdot s^{-1}$, adapting the described methodology [23]. Arnon medium was used for the four microalgae, although in the case of *S. platensis*, it was enriched with $16\,g \cdot L^{-1}$ of sodium bicarbonate. The temperature of the cultures was maintained at 25 °C while the pH was controlled at 8.0 by the on-demand injection of CO_2. To perform absorption measurements, samples from the reactors were collected and placed in batch mode inside 0.2 L flasks for ten days, with no CO_2 injection and at an irradiance of $250\,\mu E \cdot m^{-2} \cdot s^{-1}$ to simulate variations in pigments and composition that can be found in batch cultures of each of the selected microalgae.

2.2. Light absorption measurements

Samples of *Nostoc* sp., *S. almeriensis*, *S. platensis* and *C. vulgaris* were taken from the reactors over five consecutive days. Light absorption measurements were performed on the monoalgal cultures. Additionally, light absorption measurements were carried out on paired cultures in proportions of 50%/50% and 75%/25%. In total, 22 suspensions were prepared daily; these included both monoalgal and mixed algal cultures. Each suspension was measured 25 times. Over the five days, a total of 550 samples were measured to obtain sufficient data to train the ANN; the larger the amount of data, the lower the model error and the greater the accuracy of the developed ANN [24]. In order to validate the ANN, additional measurements were taken over two further days for monoalgal and mixed algal suspensions. Mixed algal suspensions were of the same and different proportions to those used in the training phase; in addition, mixtures of 3 algal species were used. Measurements were performed using a Minolta CM-3500d colorimeter - 31 light absorption data points were acquired per sample in the 400 nm to 700 nm range, in bandwidths of 10 nm. Each sample consisted of a 12 mL microalgae suspension, which was stirred and homogenized.

2.3. Artificial neural network and data analysis

The absorbance values were converted to their relative frequencies using Eq. (1), where ra_i is the relative absorbance for a specific bandwidth, a_i is the raw absorbance in a bandwidth, and Σa is the total light absorption of a microalgal suspension across the spectrum. Consequently, the relative absorbance represents the quantity of light absorption in each bandwidth without the cell concentration effect.

$$ra_i = \frac{a_i}{\sum a} \tag{1}$$

To develop the ANN, data were normalized in order to improve network performance, increase output accuracy and reduce computational cost [25,26]. The statistical, or Z-score, normalization technique (Eq. (2)) was applied to the relative absorbance data, where ra_i is the relative light absorption datum (for each absorbance band per sample), μ_i is the mean of the whole dataset, σ_i is the standard deviation and ra_i' is the normalized datum.

$$ra_i' = \left| \frac{(ra_i - \mu_i)}{\sigma_i} \right| \tag{2}$$

The ANN was a fully connected feed-forward neural network, programmed in Python 3.6.5 language using the TensorFlow 1.8.0 machine learning library; both are open source. After various configuration tests, the simplest structure with the least training and testing errors was selected. Hence, the neural network architecture was 31 neurons in the input layer, one neuron for each absorbance datum acquired per sample, 45 neurons in the hidden layer with the hyperbolic tangent activation function and 4 neurons in the output layer with the softmax activation function to weight each microalgae species in suspension; a schema is presented in Fig. 1.

The optimizer was the gradient descent algorithm with the cross-entropy cost function. In the training and testing phase, 10% of the input dataset was used for testing. The early stop, used to prevent overfitting, was applied before the testing error started to increase [27]. In order to validate the model, the ANN output was analysed by performing a regression using the expected microalgae concentration in suspension of the validation dataset.

3. Results and discussion

3.1. Light absorption

The light absorption spectrum of a microalga is a function of the biomass concentration in the culture and the pigment content of the biomass. The light absorption spectrum measurement is presented in Fig. 2 (left). The results showed that no substantial variations in the absorption spectrum occurred for any of the microalgae over the time course measurements, the slight variations observed being due to differences in the biomass content - the higher the cell density, the greater the absorbance [28,29]. These variations were ruled out after the absorbance was converted to the relative absorbance for each measurement (Eq. (1)).
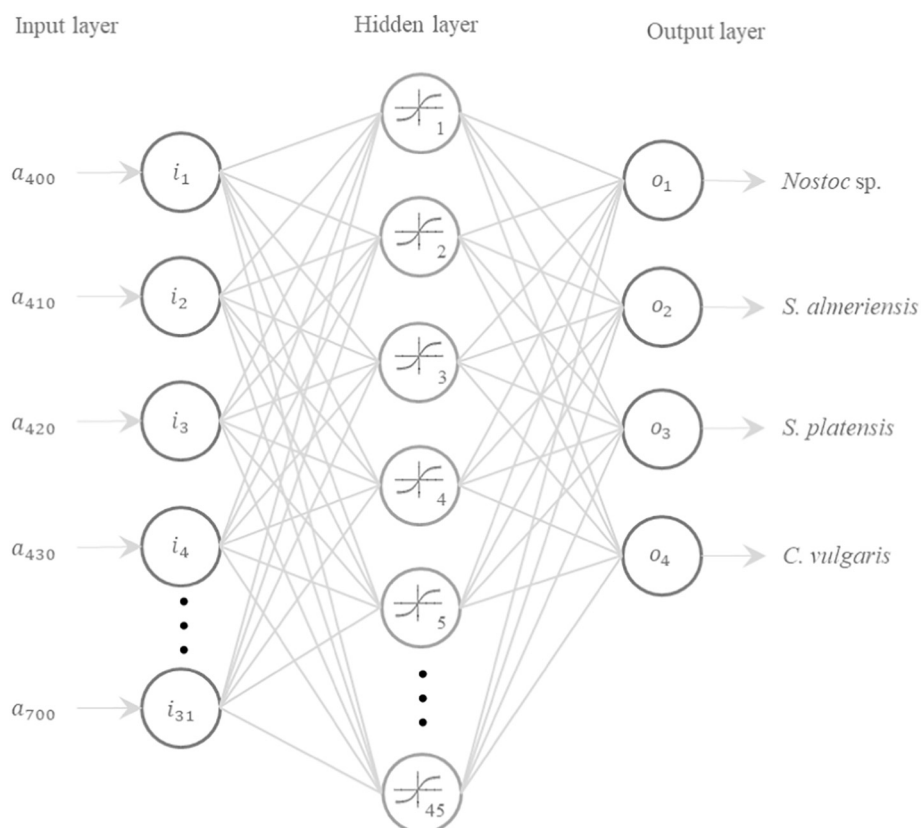
Fig. 1. ANN architecture schema, with 31 neurons in the input layer, 45 in the hidden layer and 4 in the output layer.

When normalizing to relative absorption, the spectral signature is a function of the pigment content; this varies depending on the algal species. One can observe how the lines overlap each other, thus making the spectral signature of the different microalgae used more comparable (Fig. 2 right). The results show that each microalga specie had a singular spectral signature: *S. almeriensis* and *C. vulgaris* had similar spectral signatures while *Nostoc* sp. and *S. platensis* had particular shapes that allow to distinguish one from the other. In the case of *Nostoc* sp., slight variations in spectral signature were observed over time, especially on days 4 and 5. This variation, especially in the chlorophyll absorbance region, is not due to changes in environmental conditions because these were kept constant under laboratory conditions, so it might be related to changes in the cells' biochemical composition.

Nutrient supply and light conditions are two major factors that modify the pigment content of a microalga strain. Thus, the pigment content is related to the physiological status of the cells [30]. Furthermore, the growth-cycle phase also modifies the pigment content [31]. Other pigment content variations are related to changes in environmental and operational conditions [32,33]. Despite the difference in days between sample measurements, other factors remained constant and the absorption spectrum remained uniform. Under different culture conditions, the spectral signature could present variations within the same microalgae species across the measurements.

The variation in a microalga species' spectral signature over time was minor compared to the inter-species signature variation. Similar light absorption spectra, with peaks at 450 nm and 680 nm, have been obtained for other microalgae, such as *Chlamydomonas reinhardtii* [34], *Thalassiosira pseudonana* [35] and *C. vulgaris* [36]. Absorbance peaks are caused by photosynthetic pigments. Thus, while chlorophylls present two distinct absorption maxima, one between 400 and 500 nm and the other between 600 and 700 nm, the maximum absorption for carotenoids can increase above 500 nm due to spectral shifts caused by different contributing pigments [37].

3.2. Artificial neural network and data analysis

Data from 550 samples were measured to train and test the ANN. Once the training and testing phases were concluded, the obtained model was used to determine whether or not the validation dataset samples were monoalgal or mixed algal cultures. The number of samples met the requirement of having more training samples than hidden neuron units [38], more than 10 times the present scenario. The final average cross-entropy (ACE) during the training and testing phases were 0.520 and 0.487, respectively. For the validation phase, the ACE was 0.527. The comparison between the experimental data and the ANN output is shown in Table 1.

Monoalgal cultures (S101 to S104) from the first validation experiment day were correctly identified by the artificial neural network, providing results above 98% purity. Regarding the mixed algal samples, S105 and S106 correspond to not-previously-performed combinations, thus, they are "unknown combinations" for the model. Despite this, the model was capable of approximating the composition of these samples. Therefore, these samples were close to the monoalgal cultures and the model was able to identify the most abundant strain and its relative proportion. The ANN considered that small amounts of other strains were also present in these samples; nevertheless, it was always capable of identifying the most prevalent strains and percentage composition. Regarding the samples with 50% of two different strains (S107 and S108), the artificial neural network was, likewise, capable of approximating the composition of these mixed algal samples, identifying the two prevailing strains and their approximate percentage. In no case were these samples identified as monoalgal cultures. The same validation protocol was repeated on the second day. In this case, when using monoalgal cultures (S201 to S204), the ANN provided results above 99% purity for these samples. For samples containing 90–95% of one strain and 10–5% of another (S205 to S209), these samples were "unknown combinations" for the model. Nonetheless, it was able to identify
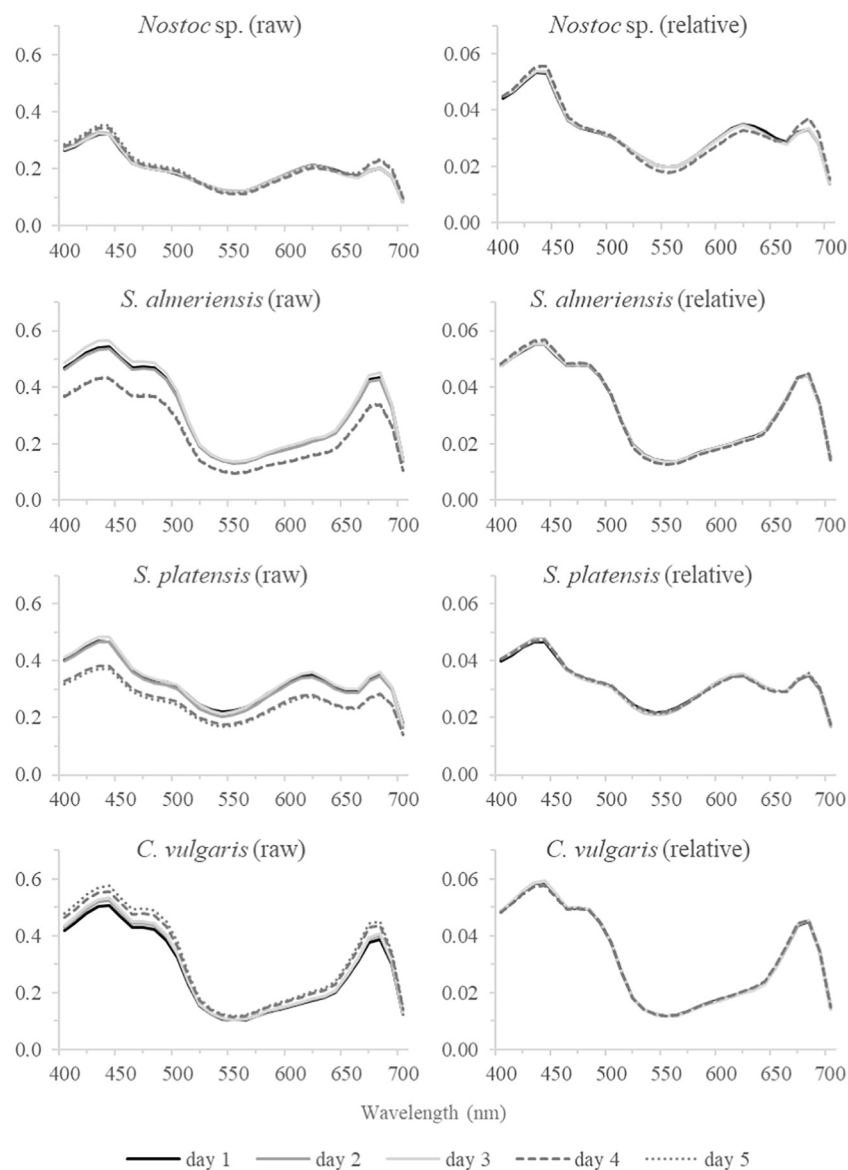
Fig. 2. Light absorption spectrum of monoalgal cultures as obtained from the colorimeter (left), and relative absorbance conversion (right).

the prevailing algal species present in each sample, with percentages similar to monoalgal cultures. The one exception was the 90% *S. platensis* and 10% *C. vulgaris* sample, for which an accurate estimate was given. Similar mixed algal samples to those used for the training process (S210 to S227) provided better results, with none of them being classified as monoalgal. The last sample (S228) also corresponded to an "unknown combination" for the model, and it was, likewise, not classified as monoalgal.

According to these results, mixed algal samples with 10% or less contamination are weighted as monoalgal suspensions. Mixed algal samples are less accurate when predicting their composition. In a study using flow cytometry with the SYTO9 stain, it was possible to classify *C. vulgaris*, *Scenedesmus obliquus*, *Chlamydomonas reinhardtii*, and *Navicula pelliculosa*, with errors ranging from 5 to 10%. However, the method misidentified microalgae cells in mixed algal samples [39]. Given that contamination by non-target microalgae is a serious problem for microalgae cultivation [40], our method could be a powerful alternative for supervising algal cultures. Additionally, the method can provide information about the relative composition of a sample, an advantage over traditional methods in which more steps and time are needed to reach similar conclusions.

The ANN gives the approximate composition of a sample based on data input processing; each absorbance bandwidth is weighted during the training phase. Confusion over the composition can be caused by microalgae species with similar spectral signatures and by the manner each absorbance bandwidth influences the ANN output. For instance, in S102, a *S. almeriensis* monoalgal sample, *C. vulgaris* obtained a 1.30% prediction compared to 98.70% for *S. almeriensis* - both microalgae species exhibited similar spectral signatures (Fig. 2 right). Furthermore, *C. vulgaris* was overestimated in mixed algal samples, whereas *S. platensis* was underestimated (S110, S216 and S221). This confusion may be due to the way the resulting spectral signature bandwidths of the mixed algal samples are weighted by the ANN.

To better show the accuracy of the developed ANN, a regression analysis of the experimental and modelled values was performed (Fig. 3). The results show that, regardless of the microalga species, the ANN fitted the experimental values, with the regression coefficient ranging from 0.951 to 0.970.

Microalgae identification by means of ANN has been previously reported with taxonomic accuracy of up to 99%, using micrographic image analysis [17], similar accuracy to that achieved in the present study for monoalgal cultures. The method reported here demonstrated

Table 1
Comparison of the results from the experimental and the artificial neural network output of the validation experiment.

Day	Sample	Experimental				Model			
		Nostoc sp.	*S. almeriensis*	*S. platensis*	*C. vulgaris*	*Nostoc* sp.	*S. almeriensis*	*S. platensis*	*C. vulgaris*
1	S101	100.00%	0.00%	0.00%	0.00%	100.00%	0.00%	0.00%	0.00%
	S102	0.00%	100.00%	0.00%	0.00%	0.00%	98.70%	0.00%	1.30%
	S103	0.00%	0.00%	100.00%	0.00%	0.02%	0.00%	99.98%	0.00%
	S104	0.00%	0.00%	0.00%	100.00%	0.00%	0.14%	0.00%	99.86%
	S105	95.00%	5.00%	0.00%	0.00%	99.93%	0.05%	0.00%	0.02%
	S106	90.00%	10.00%	0.00%	0.00%	98.71%	1.19%	0.00%	0.09%
	S107	50.00%	0.00%	0.00%	50.00%	60.16%	0.91%	0.00%	38.93%
	S108	0.00%	50.00%	50.00%	0.00%	0.01%	30.01%	69.36%	0.62%
	S109	33.33%	33.33%	33.33%	0.00%	47.63%	36.42%	15.29%	0.66%
	S110	12.50%	50.00%	37.50%	0.00%	5.76%	42.75%	16.39%	35.10%
2	S201	100.00%	0.00%	0.00%	0.00%	99.83%	0.00%	0.17%	0.00%
	S202	0.00%	100.00%	0.00%	0.00%	0.00%	99.99%	0.00%	0.01%
	S203	0.00%	0.00%	100.00%	0.00%	0.10%	0.00%	99.89%	0.01%
	S204	0.00%	0.00%	0.00%	100.00%	0.00%	0.26%	0.00%	99.74%
	S205	90.00%	10.00%	0.00%	0.00%	99.68%	0.25%	0.01%	0.06%
	S206	0.00%	90.00%	10.00%	0.00%	0.02%	99.70%	0.21%	0.06%
	S207	0.00%	0.00%	90.00%	10.00%	0.01%	0.51%	89.19%	10.29%
	S208	0.00%	5.00%	0.00%	95.00%	0.00%	0.02%	0.00%	99.98%
	S209	10.00%	0.00%	0.00%	90.00%	0.01%	0.12%	0.00%	99.87%
	S210	75.00%	25.00%	0.00%	0.00%	87.42%	12.42%	0.00%	0.16%
	S211	75.00%	0.00%	25.00%	0.00%	73.57%	0.00%	26.43%	0.00%
	S212	75.00%	0.00%	0.00%	25.00%	68.98%	0.87%	0.00%	30.15%
	S213	0.00%	75.00%	25.00%	0.00%	0.09%	82.03%	17.85%	0.03%
	S214	0.00%	75.00%	0.00%	25.00%	0.00%	79.11%	0.00%	20.88%
	S215	25.00%	75.00%	0.00%	0.00%	32.55%	67.43%	0.02%	0.00%
	S216	0.00%	0.00%	75.00%	25.00%	0.00%	0.28%	60.20%	39.52%
	S217	25.00%	0.00%	75.00%	0.00%	19.39%	0.00%	80.59%	0.01%
	S218	0.00%	25.00%	75.00%	0.00%	0.01%	19.11%	80.44%	0.44%
	S219	25.00%	0.00%	0.00%	75.00%	7.92%	0.03%	0.01%	92.04%
	S220	0.00%	25.00%	0.00%	75.00%	0.00%	17.98%	0.00%	82.02%
	S221	0.00%	0.00%	25.00%	75.00%	0.06%	0.02%	9.61%	90.31%
	S222	50.00%	50.00%	0.00%	0.00%	62.62%	37.37%	0.01%	0.00%
	S223	50.00%	0.00%	50.00%	0.00%	49.55%	0.00%	50.45%	0.00%
	S224	50.00%	0.00%	0.00%	50.00%	48.69%	0.59%	0.02%	50.70%
	S225	0.00%	50.00%	50.00%	0.00%	0.04%	38.49%	61.39%	0.09%
	S226	0.00%	50.00%	0.00%	50.00%	0.01%	53.80%	0.00%	46.20%
	S227	0.00%	0.00%	50.00%	50.00%	0.00%	0.01%	56.01%	43.98%
	S228	20.00%	0.00%	40.00%	40.00%	0.79%	0.56%	36.22%	62.44%

its efficiency in discriminating mixed algal cultures whereas it was less efficient when there were smaller percentages of another strain in the samples. If the ANN output indicated 90% or less, there was a high probability that the examined sample was derived from a mixed algal culture; conversely, if it was more than 90%, the culture should be examined to confirm that it was monoalgal. Therefore, this decreases the number of chemical analyses required to monitor the biological composition of microalgae cultures.

Re-training the ANN with more samples of mixed algal cultures in a wider variety of relative composition or using a higher spectral resolution could improve the ANN precision. The re-training could also be applied to incorporate more microalgae species into the model and test the model capability to differentiate a greater number of microalgae. This is a relatively fast process and can take anywhere between 5 min to about 1 h depending on the computer hardware, training the ANN in a GPU is significantly faster than in a CPU only system [41].

4. Conclusions

It was demonstrated that microalgae light absorption spectra vary mainly as a function of the microalga strain, although variations due to environmental and operational conditions can also take place. When maintaining the cultures under similar conditions, the light absorption spectra can be used to develop an ANN that differentiates monoalgal from mixed algal cultures and identifies the prevailing strains. In addition, it is useful to be able to approximate the percentage of each strain in mixed cultures.

A major advantage of this method is that it does not require much time or chemical analysis, a single light absorption spectrum is enough to "quantify" the biological composition of the cultures. Re-training the ANN with more microalgae species, using monoalgal and mixed algal samples with more combinations and using more resolution in the spectral signature measurements are key to reinforcing the model and achieving more classification capability.

Declaration of contributions

The ideation of the work, the critical revision of the article and obtaining funding were performed by Acién, FG. The methodology for the microalgae light absorption measurement was conducted by Sepúlveda, C. The biological material was provided by Gómez, C. The administrative support and draft revision were carried out by Navas, LM. Franco, BM, performed the experiments, coded the neural network, analysed the data and wrote the manuscript.

Conflict of interest statement

The authors declare that there is no conflict of interest and give their informed consent. No financial or other interests influenced the outcome of the research. The authors recognise their authorship of the manuscript and agree to peer review.
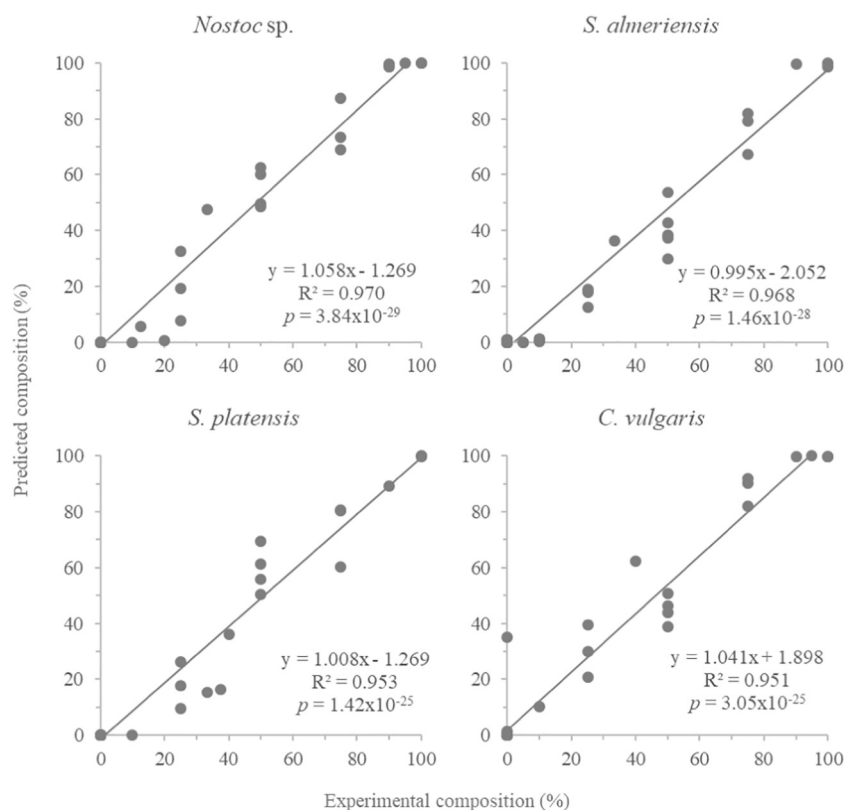
Fig. 3. Correlation between the experimental composition of the samples and that predicted by the artificial neural network developed for the samples contained in Table 1.

Statement of informed consent, human/animal rights

No conflicts, informed consent, human or animal rights applicable.

Acknowledgements

References

[1] M.A. Borowitzka, High-value products from microalgae-their development and commercialisation, J. Appl. Phycol. 25 (2013) 743–756, https://doi.org/10.1007/s10811-013-9983-9.

[2] Y. Chisti, Biodiesel from microalgae, Biotechnol. Adv. 25 (2007) 294–306, https://doi.org/10.1016/j.biotechadv.2007.02.001.

[3] C. Enzing, M. Ploeg, M. Barbosa, L. Sijtsma, Microalgae-based Products for the Food and Feed Sector: An Outlook for Europe, (2014), https://doi.org/10.2791/3339.

[4] E.J. Olguín, Dual purpose microalgae – bacteria-based systems that treat wastewater and produce biodiesel and chemical products within a biorefinery, Biotechnol. Adv. 30 (2012) 1031–1046, https://doi.org/10.1016/j.biotechadv.2012.05.001.

[5] E. Forján, F. Navarro, M. Cuaresma, I. Vaquero, M.C. Ruíz-Domínguez, Ž. Gojkovic, M. Vázquez, M. Márquez, B. Mogedas, E. Bermejo, S. Girlich, M.J. Domínguez, C. Vílchez, J.M. Vega, I. Garbayo, Microalgae: fast-growth sustainable green factories, Crit. Rev. Environ. Sci. Technol. 45 (2014) 1705–1755, https://doi.org/10.1080/10643389.2014.966426.

[6] M. Fradique, A.P. Batista, M.C. Nunes, L. Gouveia, N.M. Bandarra, A. Raymundo, Incorporation of *Chlorella vulgaris* and *Spirulina maxima* biomass in pasta products. Part 1: preparation and evaluation, J. Sci. Food Agric. 90 (2010) 1656–1664, https://doi.org/10.1002/jsfa.3999.

[7] R.J. Shields, I. Lupatsch, Algae for aquaculture and animal feeds, Tech. – Theor. Und Prax. 21 (21) (2012) 23–37.

[8] I. de Godos, S. Blanco, P.A. García-Encina, E. Becares, R. Muñoz, Long-term operation of high rate algal ponds for the bioremediation of piggery wastewaters at high loading rates, Bioresour. Technol. 100 (2009) 4332–4339, https://doi.org/10.1016/j.biortech.2009.04.016.

[9] A. Godhe, D.M. Anderson, A.S. Rehnstam-Holm, PCR amplification of microalgal DNA for sequencing and species identification: studies on fixatives and algal growth stages, Harmful Algae 1 (2002) 375–382, https://doi.org/10.1016/S1568-9883(02)00049-5.

[10] B. Serive, E. Nicolau, J.B. Bérard, R. Kaas, V. Pasquet, L. Picot, J.P. Cadoret, Community analysis of pigment patterns from 37 microalgae strains reveals new carotenoids and porphyrins characteristic of distinct strains and taxonomic groups, PLoS One 12 (2017) 1–35, https://doi.org/10.1371/journal.pone.0171872.

[11] E.B. Sydney, T.E. da Silva, A. Tokarski, A.C. Novak, J.C. de Carvalho, A.L. Woiciecohwski, C. Larroche, C.R. Soccol, Screening of microalgae with potential for biodiesel production and nutrient removal from treated domestic sewage, Appl. Energy 88 (2011) 3291–3294, https://doi.org/10.1016/j.apenergy.2010.11.024.

[12] M. Cellamare, A. Rolland, S. Jacquet, Flow cytometry sorting of freshwater phytoplankton, J. Appl. Phycol. 22 (2010) 87–100, https://doi.org/10.1007/s10811-009-9439-4.

[13] C.L. Ng, Q.Q. Chen, J.J. Chua, H.F. Hemond, A multi-platform optical sensor for in vivo and in vitro algae classification, Sensors 17 (2017) 1–14, https://doi.org/10.3390/s17040912.

[14] S. Nozue, M. Katayama, M. Terazima, S. Kumazaki, Comparative study of thylakoid membranes in terminal heterocysts and vegetative cells from two cyanobacteria, *Rivularia* M-261 and *Anabaena variabilis*, by fluorescence and absorption spectral microscopy, Biochim. Biophys. Acta Bioenerg. 1858 (2017) 742–749, https://doi.org/10.1016/j.bbabio.2017.05.007.

[15] E. Lee, R.L. Heng, L. Pilon, Spectral optical properties of selected photosynthetic microalgae producing biofuels, J. Quant. Spectrosc. Radiat. Transf. 114 (2013) 122–135, https://doi.org/10.1016/j.jqsrt.2012.08.012.

[16] F.C. Rubio, F.G. Camacho, J.M.F. Sevilla, Y. Chisti, E.M. Grima, A mechanistic model of photosynthesis in microalgae, Biotechnol. Bioeng. 81 (2003) 459–473, https://doi.org/10.1002/bit.10492.

[17] P. Coltelli, L. Barsanti, V. Evangelista, P. Gualtieri, Algae through the looking glass, Microsc. Res. Tech. 80 (2017) 486–494, https://doi.org/10.1002/jemt.22820.

[18] F. Recknagel, A.N.N.A.–. Artificial Neural, Network model for predicting species abundance and succession of blue-green algae, Hydrobiologia 349 (1997) 47–57, https://doi.org/10.1023/a:1003041427672.

[19] W. Tian, Z. Liao, J. Zhang, An optimization of artificial neural network model for predicting chlorophyll dynamics, Ecol. Model. 364 (2017) 42–52, https://doi.org/10.1016/j.ecolmodel.2017.09.013.

[20] F. García-Camacho, L. López-Rosales, A. Sánchez-Mirón, E.H. Belarbi, Y. Chisti, E. Molina-Grima, Artificial neural network modeling for predicting the growth of the microalga *Karlodinium veneficum*, Algal Res. 14 (2016) 58–64, https://doi.org/10.1016/j.algal.2016.01.002.

[21] J. Sharon Mano Pappu, G.K. Vijayakumar, V. Ramamurthy, Artificial neural network model for predicting production of *Spirulina platensis* in outdoor culture, Bioresour. Technol. 130 (2013) 224–230, https://doi.org/10.1016/j.biortech.2012.12.082.

[22] M. AA Mosleh, H. Manssor, S. Malek, P. Milow, A. Salleh, A preliminary study on automated freshwater algae recognition and classification system, BMC Bioinf. 13 (Suppl. 1) (2012) S25, https://doi.org/10.1186/1471-2105-13-S17-S25.

[23] C. Gómez, R. Escudero, M.M. Morales, F.L. Figueroa, J.M. Fernández-Sevilla, F.G. Acién, Use of secondary-treated wastewater for the production of *Muriellopsis* sp. Appl. Microbiol. Biotechnol. 97 (2013) 2239–2249, https://doi.org/10.1007/s00253-012-4634-7.

[24] C. Ying-Jin, S. Davis, C. Chao-Kun, B. Xue, A study of sample size with neural network, Mach. Learn. Cybern. 2004. Proc. 2004 Int. Conf. 6, vol. 6, 2004, pp. 3444–3448, , https://doi.org/10.1109/ICMLC.2004.1380382.

[25] K. Gnana Sheela, S.N. Deepa, Neural network based hybrid computing model for wind speed prediction, Neurocomputing 122 (2013) 425–429, https://doi.org/10.1016/j.neucom.2013.06.008.

[26] J. Sola, J. Sevilla, Importance of input data normalization for the application of neural networks to complex industrial problems, IEEE Trans. Nucl. Sci. 44 (44(3)) (1997) 1464–1468, https://doi.org/10.1109/23.589532.

[27] A.P. Piotrowski, J.J. Napiorkowski, A comparison of methods to avoid overfitting in neural networks training in the case of catchment runoff modelling, J. Hydrol. 476 (2013) 97–111, https://doi.org/10.1016/j.jhydrol.2012.10.019.

[28] L. Helena, R. Rodrigues, A. Arenzon, M.T. Raya-rodriguez, N.F. Fontoura, Algal density assessed by spectrophotometry: a calibration curve for the unicellular algae *Pseudokirchneriella subcapitata*, J. Environ. Chem. Ecotoxicol. 3 (2011) 225–228.

[29] D.U. Santos-Ballardo, S. Rossi, V. Hernández, R.V. Gómez, R. del Carmen Rendón-Unceta, J. Caro-Corrales, A. Valdez-Ortiz, A simple spectrophotometric method for biomass measurement of important microalgae species in aquaculture, Aquaculture 448 (2015) 87–92, https://doi.org/10.1016/j.aquaculture.2015.05.044.

[30] I. Pancha, K. Chokshi, S. Mishra, Enhanced biofuel production potential with nutritional stress amelioration through optimization of carbon source and light intensity in *Scenedesmus* sp. CCNM 1077, Bioresour. Technol. 179 (2015) 565–572, https://doi.org/10.1016/j.biortech.2014.12.079.

[31] L. Lubián, O. Montero, I. Moreno-Garrido, I.E. Huertas, C. Sobrino, M. González-del Valle, G. Pares, Nannochloropsis (Eustigmatophyceae) as source of commercially valuable pigments, J. Appl. Phycol. 12 (2000) 249–255, https://doi.org/10.1023/A:1008170915932.

[32] I. Pancha, K. Chokshi, B. George, T. Ghosh, C. Paliwal, R. Maurya, S. Mishra, Nitrogen stress triggered biochemical and morphological changes in the microalgae *Scenedesmus* sp. CCNM 1077, Bioresour. Technol. 156 (2014) 146–154, https://doi.org/10.1016/j.biortech.2014.01.025.

[33] A. Solovchenko, O. Solovchenko, I. Khozin-Goldberg, S. Didi-Cohen, D. Pal, Z. Cohen, S. Boussiba, Probing the effects of high-light stress on pigment and lipid metabolism in nitrogen-starving microalgae by measuring chlorophyll fluorescence transients: studies with a δ5 desaturase mutant of *Parietochloris incisa* (Chlorophyta, Trebouxiophyceae), Algal Res. 2 (2013) 175–182, https://doi.org/10.1016/j.algal.2013.01.010.

[34] T. Isono, K. Yamashita, D. Momose, H. Kobayashi, M. Kitamura, Y. Nishiyama, T. Hosoya, H. Kanda, A. Kudo, N. Okada, T. Yagi, K. Nakata, S. Mineki, E. Tokunaga, Scan-free absorbance spectral imaging A(x, y, λ) of single live algal cells for quantifying absorbance of cell suspensions, PLoS One 10 (2015) 1–16, https://doi.org/10.1371/journal.pone.0128002.

[35] C.D. Hewes, The color of mass culture: spectral characteristics of a shallow water column through shade-limited algal growth dynamics1, J. Phycol. 52 (2016) 252–259, https://doi.org/10.1111/jpy.12393.

[36] J.A. Myers, B.S. Curtis, W.R. Curtis, Improving accuracy of cell and chromophore concentration measurements using optical density, BMC Biophys. 6 (2013) 4, https://doi.org/10.1186/2046-1682-6-4.

[37] A. Holzinger, M.C. Allen, D.D. Deheyn, Hyperspectral imaging of snow algae and green algae from aeroterrestrial habitats, J. Photochem. Photobiol. B Biol. 162 (2016) 412–420, https://doi.org/10.1016/j.jphotobiol.2016.07.001.

[38] G. Huang, Q. Zhu, C. Siew, Extreme learning machine: a new learning scheme of feedforward neural networks, IEEE Int. Jt. Conf. Neural Networks 2 (2004) 985–990, https://doi.org/10.1109/IJCNN.2004.1380068.

[39] G.T. Peniuk, P.J. Schnurr, D.G. Allen, Identification and quantification of suspended algae and bacteria populations using flow cytometry: applications for algae biofuel and biochemical growth systems, J. Appl. Phycol. 28 (2016) 95–104, https://doi.org/10.1007/s10811-015-0569-6.

[40] X. Wen, K. Du, Z. Wang, X. Peng, L. Luo, H. Tao, Y. Xu, D. Zhang, Y. Geng, Y. Li, Effective cultivation of microalgae for biofuel production: a pilot-scale evaluation of a novel oleaginous microalga *Graesiella* sp. WBG-1, Biotechnol. Biofuels 9 (2016) 1–12, https://doi.org/10.1186/s13068-016-0541-y.

[41] R. Ramachandran, D.C. Rajeev, S.G. Krishnan, P. Subathra, Deep learning in neural networks: an overview, Int. J. Appl. Eng. Res. 10 (2015) 25433–25448, https://doi.org/10.1016/j.neunet.2014.09.003.