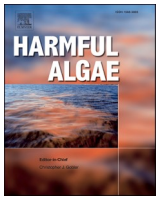
Harmful Algae 117 (2022) 102290

Contents lists available at ScienceDirect

Harmful Algae

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



Uncertainties of cell number estimation in cyanobacterial colonies and the potential use of sphere packing

Enikő T-Krasznai [a,*], Verona Lerf [a], István Tóth [a], Tibor Kisantal [a], Gábor Várbíró [a], Gábor Vasas [b], Viktória B-Béres [a], Judit Görgényi [a], Áron Lukács [a], Zsuzsanna Kókai [a], Gábor Borics [a,c]

[a] Centre for Ecological Research, Institute of Aquatic Ecology, Department of Tisza Research, Functional Algology Research Group, Bem square 18/c, H-4026 Debrecen, Hungary
[b] University of Debrecen, Faculty of Science and Technology, Department of Botany, Egyetem square 1, H-4032 Debrecen, Hungary
[c] University of Nyíregyháza, PO Box 166, H-4401 Nyíregyháza, Hungary

ARTICLE INFO

Editor: Linda Ann Lawton, Ph.D.

Keywords:
Biovolume
Inverted microscope
Water quality
Cell count

ABSTRACT

Cyanobacteria are notorious bloom formers causing various water quality concerns, such as toxin production, extreme diurnal variation of oxygen, pH, etc., therefore, their monitoring is essential to protect the ecological status of aquatic systems. Cyanobacterial cell counts and biovolumes are currently being used in water management and water quality alert systems. In this study, we investigated the accuracy of traditional colonial biovolume and cell count estimation approaches used in everyday practice. Using shape realistic 3D images of cyanobacterial colonies, we demonstrated that their shape cannot be approximated by ellipsoids. We also showed that despite the significant relationship between overall colony volume and cell biovolumes, because of the considerable scatter of cell count data the regressions give biased estimates for cyanobacterial cell counts. We proposed a novel approach to estimate cell counts in colonies that was based on the random close sphere packing method. This method provided good results only in those cases when overall colony volumes could be accurately measured. The visual investigation of colonies done by skilled experts has given precise but lower estimates for cell counts. The estimation results of several experts were surprisingly good, which suggests that this capability can be improved and estimation bias can be reduced to the level acceptable for water quality estimations.

1. Introduction

Despite the efforts that have been made to reduce the nutrient loads to surface waters, eutrophication still remains a major source of concerns for water management (Smith and Schindler, 2009). The process coincides with a drastic increase in phytoplankton biomass and undesirable compositional changes, reducing the stability of aquatic ecosystems and threatening the services they provide to society. Monitoring the quality of surface waters, therefore, is essential to avoid the negative consequences of eutrophication.

Water blooms caused by cyanobacteria and several eukaryotic algae (Reynolds and Walsby, 1975) are the most serious problems in eutrophic waters worldwide. While some of the eukaryotic algae are capable to form freshwater red tides (e.g. *Uroglenopsis americana* (G.N.Calkins) Lemmermann; Ochrophyta) (Ishikawa et al., 2005), others produce surface scums (e.g. *Botryococcus braunii* Kützing; Chlorophyta) (Paerl et al., 2001). The formed blooms can be skin irritating (*Gonyostomum semen* (Ehrenberg) Diesing; Ochrophyta) (Hongve et al., 1988), or toxic (e.g. *Prymnesium parvum* N.Carter; Haptophyta) (Vasas et al., 2012). Cyanobacteria are one of the most notorious groups of phytoplankton and are responsible for the vast majority of harmful algal blooms throughout the world (Pearl et al., 2001). Cyanobacterial blooms, apart from being aesthetically displeasing, can be toxic and can decrease water quality causing health problems both for animals and humans and thus reducing the usability of water (Carvalho et al., 2013a). Because of the globally increasing occurrence of cyanobacterial harmful algal blooms (CyanoHABs), several methods have been developed for estimating the biomass of cyanobacteria (e.g. qPCR, chlorophyll-a concentration, the concentration of microcystin, ESP; Alcántara et al., 2018; Seltenrich, 2014; Wang et al., 2015; in vivo phycocyanin fluorescence

* Corresponding author.
E-mail address: ekrasznai@gmail.com (E. T-Krasznai).

https://doi.org/10.1016/j.hal.2022.102290
Received 10 March 2022; Received in revised form 1 July 2022; Accepted 11 July 2022
Available online 20 July 2022
1568-9883/© 2022 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

(McQuaid et al., 2011). Because WHO (1999; 2003) uses cell counts for risk thresholds, analysis of samples requires traditional microscopy, identification of taxa, estimation of cell counts and taxon-specific biovolumes (Carvalho et al., 2013).

Biovolume estimation of phytoplankton is based on the microscopy measurements of linear dimensions of the units that are compared to geometric shapes (Hillebrand et al., 1999). Multiplying these values with volumetric cell counts (cells mL^{-1}) provides species-biovolume data, which serves as a basis for both scientific and management purposes (CEN 16695, 2015). This approach provides reasonable estimates for specific biovolumes in the case of simple-shaped unicellular taxa and multicellular filaments, but for complex shapes, it might give uncertain results (Borics et al., 2021).

Based on their morphological appearances, cyanobacteria can be distinguished into three major groups: i) unicellular, ii) filamentous, and iii) colony forming. While measuring the cell sizes and counting the numbers of unicellular or filamentous forms (considering them as cylinders) can be easily accomplished, estimation of cell numbers in colonial forms has an unknown uncertainty, especially in the case of coenobial (*Aphanocapsa, Coelomoron, Microcystis* or *Pannus*) and utricular (*Coelosphaerium, Snowella* or *Woronichinia)* forms. However, the problem is not new, and during the last decades, some alternative approximations and procedures were proposed to estimate the cell numbers in colonies. In the case of most methods, the disintegration of colonies to single cells is always the first step, which can be carried out by sonication, heating, boiling or using alkaline hydrolysis (Box, 1981). The second step is to count the cell numbers per sample, which can be carried out by hemocytometer (Joung et al., 2006), or FlowCAM (Wang et al., 2015). These methods, however, are proposed for blooms dominated by *Microcystis*. Reynolds and Jaworski (1978) estimated the concentration of cells per colony using a regression equation fitted to data derived from natural populations. Joung et al. (2006) applied a similar approach; after the disintegration of *Microcystis* colonies by boiling, they counted the cell numbers and calculated their concentrations in the colonies, which had been measured previously and were considered as spheres. The authors found a clear relationship between colony size and *Microcystis* cell numbers.

Uncertainty of biomass estimation of disintegrated colonies is low and is equal to that of the other unicellular microalgae, but these time-consuming approaches are not suitable for routine monitoring.

Alcántara et al. (2018) estimated the colony volume of '*Microcystis aeruginosa* complex' using a geometrical approximation. Three theoretical morphologies of the colonies (sphere, prolate spheroid and ellipsoid) were compared, and the authors proposed to apply ellipsoid forms to estimate colonial biovolume of *Microcystis* spp. This approach can be useful to estimate colony size in routine monitoring but unfortunately does not provide any help in measuring cellular biovolumes.

Recently an increasing demand for monitoring cyanobacterial blooms is emerging, because climate change and the increasing anthropogenic pressures trigger the bloom development both in eutrophic (Huisman et al., 2018) and oligotrophic (Reinl et al., 2021) environments.

In this study, we aim to investigate the accuracy of the existing approaches and of a new one that merges the advantages of 3D imagery of microalgae (Borics et al., 2021) and the practical application of the age-old sphere packing problem (Hifi and M'hallah, 2009).

We hypothesised that colony volume and cell count estimation together give a better prediction for colonial cyanobacterial biovolume than the aforementioned regression approaches.

2. Material and methods

2.1. Sampling

To assess the accuracy of the above-mentioned approaches we collected phytoplankton samples from various Hungarian standing waters in 2020. Samples were collected from the photic zone of the waters and preserved with Lugol's solution in the field. Samples were stored at 4 °C until the analyses. Although Lugol's solution might cause shrinkage of cyanobacterial cells (Hawkins et al., 2005), we made every investigation on the same material, thus the shrinkage did not have any effect on the results of comparisons.

2.2. Sample processing and giving the reference cell counts

We investigated the colonies by Leica DMRB light microscope at 100× and 400× magnifications. Photos were taken by Canon 4000D camera and the colonies were measured by QuickPHOTO CAMERA 3.2 microscope software.

The studied colonies were randomly selected in the droplets. In order to know the real shape of the colonies we preferred those that could be rotated under the coverglass. We took several photos of these colonies in different layers, and these were used later to measure the colony size (length, width and depth), the cell size and then the distances of cells within the colonies. To estimate the mean distances amongst the cells we selected 4–5 cells randomly in the colonies and using the ruler tool of the QuickPHOTO CAMERA microscope software we measured the distances between the centres of the selected and the neighbouring cells. To get reliable estimates for the mean cell distances we performed a minimum of 20 measurements in each colony.

The mucilage of the colony was not measured, only the space filled by cells. In order to know the exact number of cells within the colony, the water between the slide and the coverglass was allowed to evaporate, resulting in the flattening of the colonies until the cells lay in one layer (Fig. 1). We checked the complete flattening of colonies under the microscope using the fine focus. Finally, we took a photo of the flattened colonies and counted the cells in the photographs (shown in electronic supplementary material, Fig. S1). Counted cells were marked by colour dots using Microsoft Paint. These cell numbers were considered as references during the comparison of the three cell count estimation approaches shown below. Based on the measurements of the cells, we counted cell volumes and finally, we gave overall colony biovolumes.

2.3. Cell count estimations

We compared the accuracy of three cell count estimations:

1 traditional cell count estimation carried out by experts (A1)
2 geometric approach, based on sphere packing, cell size and cell distance measurements (A2)
3 regression approach, based on the colony volume and cell number relationships (A3)

2.3.1. Traditional approach (A1)

During traditional microscopic analysis of phytoplankton samples, the analysers have to give an estimation on the number of cells in the colonies appearing in the field of the microscope. We asked 15 skilled experts (who are currently working for the Hungarian monitoring network as algologists) to give an estimation for the cell numbers in the case of 100 colonies (Fig. 2). We projected to the experts at least three high-resolution photographs of each colony taken in different layers. The experts had 40 sec for the visual investigations (based on a prior query, this is the time that the experts spent on cell counting during routine monitoring). Estimated cell counts were registered in a paper, and these were tabulated electronically. The ratio between experts' estimations and reference cell numbers was used to illustrate the estimation bias of each expert. The results of visual estimations by experts were illustrated as boxplots.

2.3.2. Geometric approach ($A2_{3D}$ and $A2_{ell}$)

In our geometric approach, we considered colonial cyanobacteria as

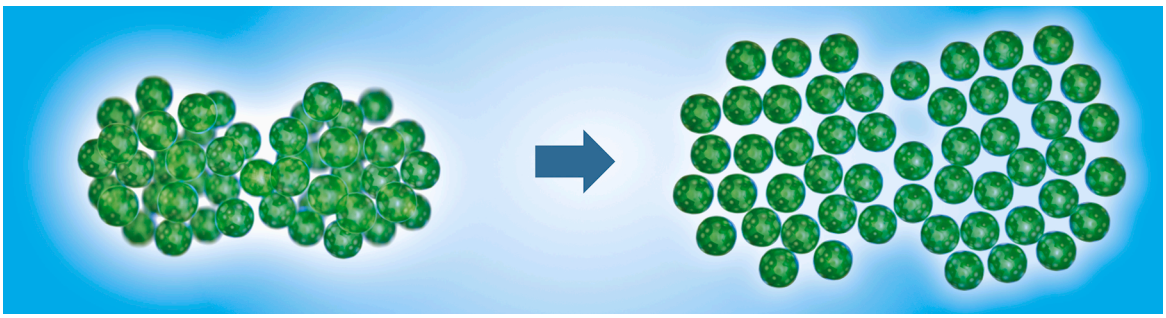
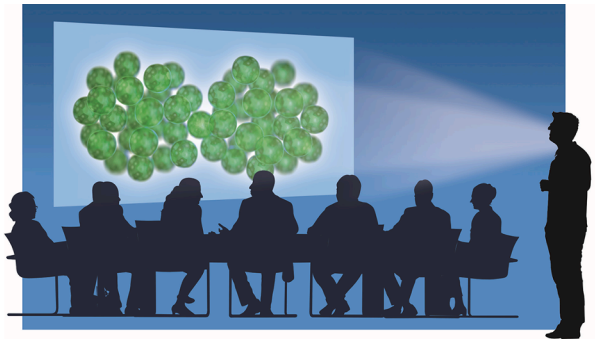


Fig. 1. Estimation of colonial cell numbers by flattening the colonies; (A) original colony, (B) flattened colony (after the evaporation of water from below the coverglass).



	Expert 1	Expert 2	Expert...	Expert 15
Colony 1	500	600	750	450
Colony 2	200	250	180	350
Colony 3	70	80	100	60
Colony 4	350	300	270	280
Colony 5	1200	2000	1500	1000
Colony ...	600	800	1500	1200
Colony 100	80	100	120	150

Fig. 2. Visual estimation of colonial cell numbers. Data are tabulated and computerized.

distantly packed spheres in a 3-dimensional space (Fig. 3). The arrangement of non-overlapping spheres in a Euclidean space and the calculation of packing density is an age-old problem in geometry (Kepler, 1611). Several models have been developed to estimate the packing density of differently arranged spheres and results range between 0.5236 (cubic arrangement) to 0.7405 ($\pi/3 \times \sqrt{2}$, i.e. Kepler's conjecture). Although cyanobacterial cells in the colonies are small and do not or rarely touch each other, theoretically they can be enlarged to the size at which they densely fill the available space. Since the packing density of loosely packed spheres is approximately 0.55 (Zamponi, 2008), the overall cell volume of the enlarged cells will be 55% of the colony volume. The volume of the enlarged cells can also be calculated because their radius is equal to half of the distance between cyanobacterial cells in the colony (supposing that distances of cell centres have been measured; Fig. 3b). The colonial volume/enlarged cell volume ratio gives the cell numbers, and thus, using the volume of the original cyanobacterial cells their overall volume can be calculated.

$$V = N\ V_{cell}$$

$$V_{packed} = 0.55\ V_{colony}$$

$$V_{cell} = \frac{4\pi r^3}{3} = \frac{\pi d_{cell}^3}{6} = 0.523 d_{cell}^3$$

$$V_{ec} = 0.523 d_{mean}^3$$

$$N = \frac{V_{packed}}{V_{ec}} = \frac{0.55\ V_{colony}}{0.523 d_{mean}^3}$$

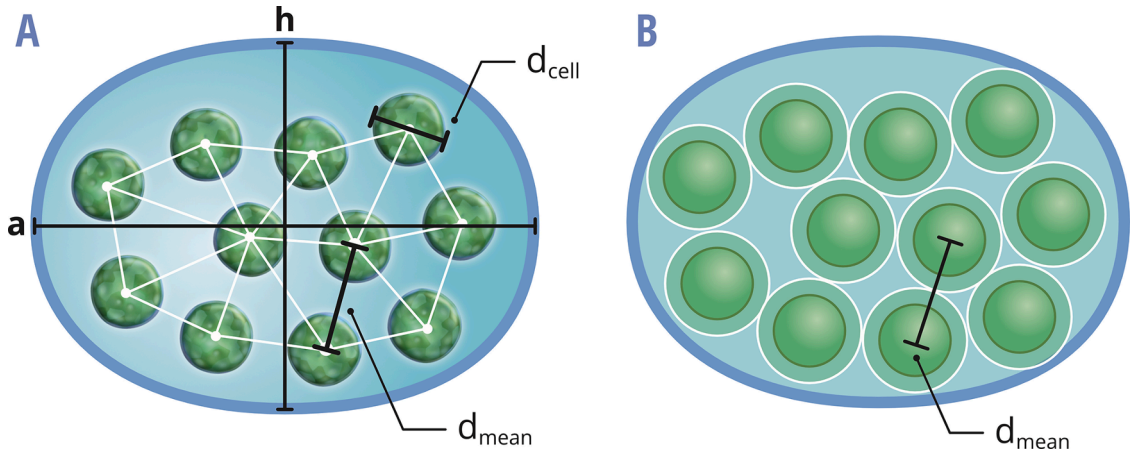


Fig. 3. Cell number estimation by close sphere packing. A: loosely arranged cells in the colony; black lines are cell distances; a: length, h: width of the colony; d_{cell}: average diameter of a cells, d_{mean}: average distance between the cell centres. B: colony with enlarged cells. Cell diameters are equal with average cell distances measured between the centre of cells.

$$V = N\ V_{cell} = \frac{0.55\ V_{colony}\ 0.523 d_{cell}^3}{0.523 d_{mean}^3} = \frac{0.55\ V_{colony}\ d_{cell}^3}{d_{mean}^3}$$

$$= 0.55\ V_{colony} \left(\frac{d_{cell}}{d_{mean}}\right)^3$$

$$V = 0.55\ V_{colony} \left(\frac{d_{cell}}{d_{mean}}\right)^3$$

V: overall volume of the cells in the colony

V_{cell}: median volume of the cells in the colony

N: number of cells in the colony

V_{packed}: volume of the colony filled with loosely packed enlarged spherical cells

V_{colony}: overall volume of the colony (cells and mucilage)

V_{ec}: mean volume of a virtually enlarged cell d_{cell}: mean diameter of the cyanobacterial cells in the colony d_{mean}: mean diameter of the fattened cells, which is equal to the mean distances measured between the cells' centres. r: radius of the sphere (here the cells)

In this study, two overall colony volume (V_{colony}) values were used: $A2_{3D}$ measured on the 3D models and $A2_{ell}$ estimated by the experts using linear measurements on the colonies (Alcántara et al., 2018) (Fig. 4). We made the geometric cell count estimations using the results of both approaches separately and the estimation biases were expressed as estimated/reference ratios.

2.3.3. Comparison of the results of A1, $A2_{3D}$ and $A2_{ell}$ approaches

The results of the estimations were compared to the reference cell numbers. We created ratios of these values and the reference ones, i.e. those given by direct counting after flattening of colonies. Multiplying these values by 100 we get percentages. Subtracting 100 from these values expresses the results as percentage differences. We used the number of estimations in the percentage difference categories as a basis for comparison.

2.3.4. Regression approach (A3)

To study the accuracy of the regression approach two questions have to be answered: what is the accuracy of colony volume estimations, and what is the accurate relationship between the colony volumes and overall cell biovolumes within the colonies.

To estimate the accuracy of colony volume estimations we created the shape realistic 3D models of each (100) studied colony and these served as references (Fig. 4B). For the realistic 3D visualization, we used the Blender 2.79 software (Blender, 2020), which is a free, open-source 3D graphics application that enables the creation of high-resolution mesh objects. The models were based on the photographs that were taken at different layers of the colonies. We calculated the models' volumes using the NeuroMorf software toolset (Jorstad et al., 2015). Since diameters of the colonies are known from microscopic measurements, colonial volumes provided by NeuroMorf were transformed to μm^3-scale.

We estimated the colony volumes using the formulae proposed by Alcántara et al. (2018):

$$V = \frac{\pi}{6} ahb$$

where, V: volume of an ellipsoid, a: length, h: width and b: depth of the colonies. We estimated the depth of colonies by the following formula:

$$\text{LogDepth} = 1.86 + 5 \times 10^{-4} \times \text{Length} + 1.54 \times 10^{-4} \times \text{Width}.$$

This estimation requires measuring the length and width of the colonies. Using the photographs that were previously used for cell count estimation of the colonies the fifteen experts were asked to measure the linear dimensions of cells and colonies, moreover the mean distances of cells in 30 colonies (Fig. 4C). Cell diameter measurements were done on randomly selected five cells, while during cell distance measurements, depending on the structure of the colonies, ten to twenty-five cell distances were measured.

The photographs were imported to LibreCad software (an open-source 2D CAD application) and all measurements were done using the ruler function of the program. Since cell diameters were known from microscopic measurements, cell distance values provided by LibreCad could be transformed to µm-scale.

These estimated colony volumes were compared to that of the reference given by the 3D approach. Estimation bias was expressed as the ratio between the estimated volumes (Alcántara et al., 2018) and volumes given by 3D modelling of the colonies.

To study the second question i.e. "what is the relationship between the colony volumes and overall cell biovolumes within the colonies?" the sum of the reference cell biovolumes was plotted against the reference colony volumes and the relationship between variables was studied using OLS (Ordinary Least Square) analysis.

3. Results

We involved 100 cyanobacterial colonies in the analyses. The colonies based on their morphology belonged to the *Microcystis* and

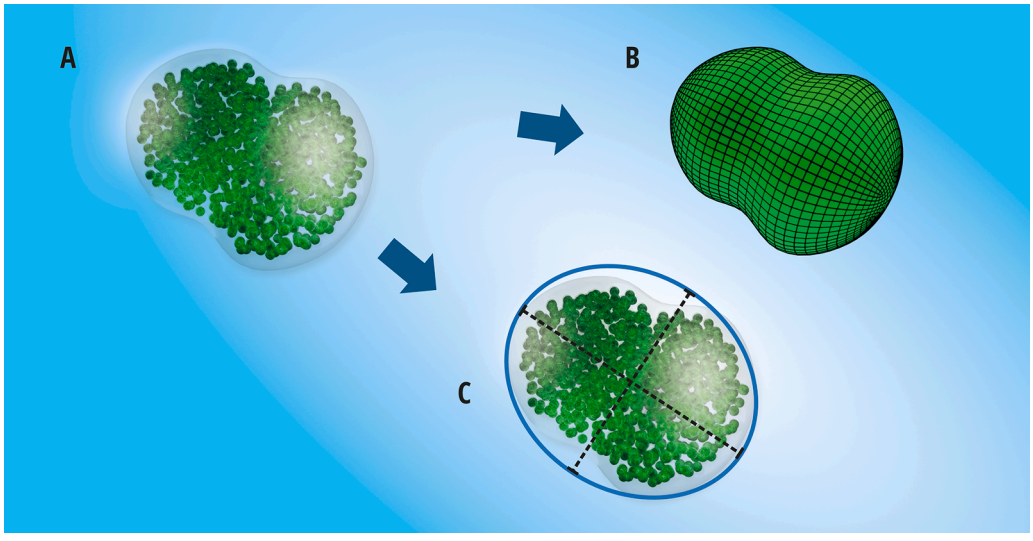


Fig. 4. Estimation of the volume of cyanobacterial colony (A), using shape realistic 3D modelling (B), and approximated by ellipsoid (C).

Chroococcus genera. Real-like 3D modelling, cell counting, cell size and distance measurements were carried out in the case of each colony as described in the Methods section. Cell distance, cell and colony size measurements were performed on a smaller (30) set of colonies by the fifteen experts.

The (A1, $A2_{3D}$, $A2_{ell}$) approaches showed considerable differences both in terms of data dispersion and location of the mode. The traditional approach (A1; estimating the cell counts by experts) (Fig. 5A) was characterised by low data dispersion (−70 to 20%). The location of the mode indicated −30% estimation error. Mode and median values were almost the same. In the case of the two geometric sphere packing approaches ($A2_{3D}$ and $A2_{ell}$), which were based on the overall colony volume estimation and cell distances, the dispersion of data was larger. The mode and median values were almost identical in the case of $A2_{3D}$, but considerably differed in the case of $A2_{ell}$ approach. When colony volumes were estimated by using 3D models ($A2_{3D}$) (Fig. 5B) data dispersed from −70 to 100%. The mode however located at −10%. The $A2_{ell}$ approach (when colonies were considered as ellipsoids) provided less accurate predictions (Fig. 5C). Estimated volumes dispersed from −80 to >120% percentage range and the mode fell outside the >110% value.

Then, to study the differences amongst experts who performed the counting, estimated values were divided by the reference ones and these were presented as boxplots (Fig. 6). Results revealed large individual differences in cell number estimations. Median values of estimated cell counts fell in the range of 0.5–1.5 in the case of 75% of experts. However, most experts underestimated the real cell numbers, and only three experts gave overestimations. Colonial cell numbers had a pronounced impact on the estimation results. In the case of colonies in which the number of cells was higher than approximately 300, overestimations only occasionally occurred (Fig. 7).

As it was shown in the methods, the regression approach (A3) is based on the overall colony volume/cell number relationships. This approach requires accurate colony volume estimation and a strong relationship between colony size and cell numbers. Our results revealed that considering cyanobacterial colonies as ellipsoids leads to severely biased estimates for their volumes. Volumes of the ellipsoids were calculated based on the measured length and width of the colonies and the calculated depth values. Means of the experts' volume calculations were divided by that of the reference (i.e. calculated by the Neuromorph on the 3D models). The standard errors (SE) calculated for these estimation biases showed great variability (0.01–1.79) and indicated large uncertainties in colony biovolume estimates. In the majority of cases, the experts (using the ellipsoid approximation method) overestimated the colony volumes. However, the uncertainties did not depend on the volume of the colonies. There was no correlation between colony size and SE of the estimation bias ($r = 0.081$, $p = 0.68$). Large overestimations were found along the whole scale of colony volumes.

We found a strong linear relationship between colony volumes (based on the 3D models) and cell volumes (using reference cell counts of the flattened colonies) (Fig. 8). The relationship was characterised by a high R-value ($R = 0.74$) and even scatter of residuals around the regression line was observed (homoscedasticity). However, the residuals covered a range of approximately one order of magnitude indicating low precision.

4. Discussion

As Cyanobacteria are the most notorious bloom-forming planktic photosynthesizers in freshwaters, their biovolumes are frequently used bioindicators in lake quality assessment (Carvalho et al., 2013a,b). While the volume of single-celled and filamentous forms can be measured accurately, in the case of colonial cyanobacteria the estimation of cell numbers has an unknown bias. In this study we assessed the

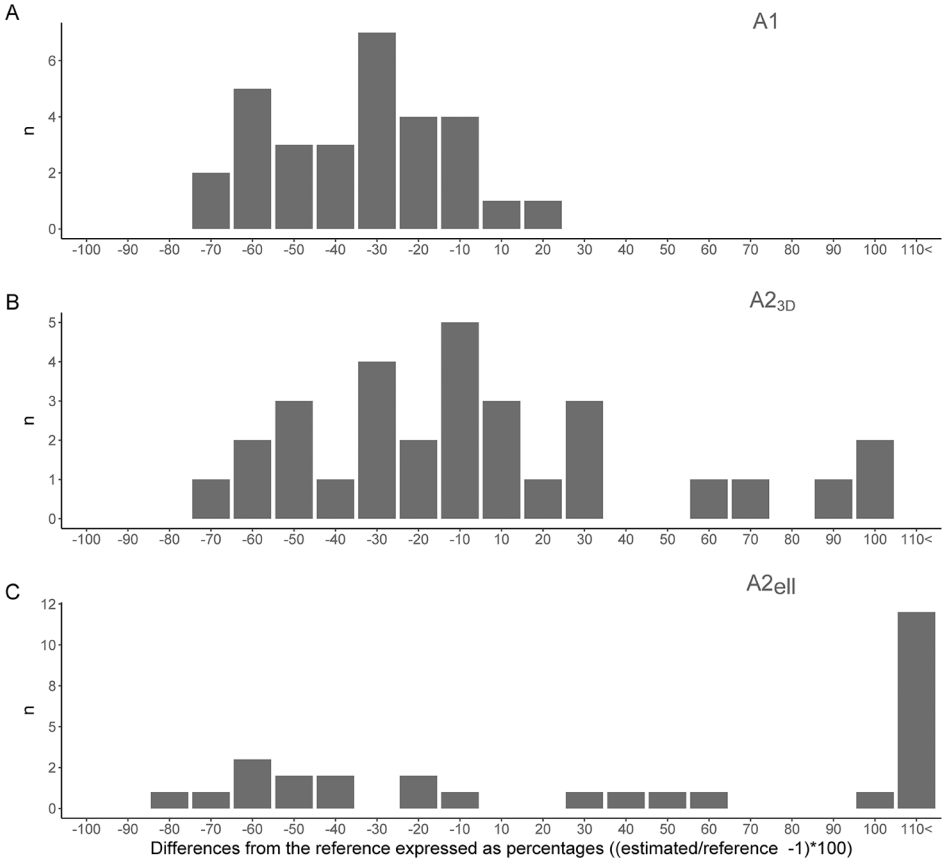


Fig. 5. Distribution of the average estimation bias for the A1, $A2_{3D}$, $A2_{ell}$ cell count estimation approaches; A: visual estimation of cell count by experts (A1); B: close sphere packing using colony volumes derived from 3D models of colonies ($A2_{3D}$); C: close sphere packing using colony volumes derived from the ellipsoid approximation approach ($A2_{ell}$). Values on the x axis are differences from the reference expressed as percentages ((estimated/reference −1) × 100). The columns indicate the number of observations in the given range.

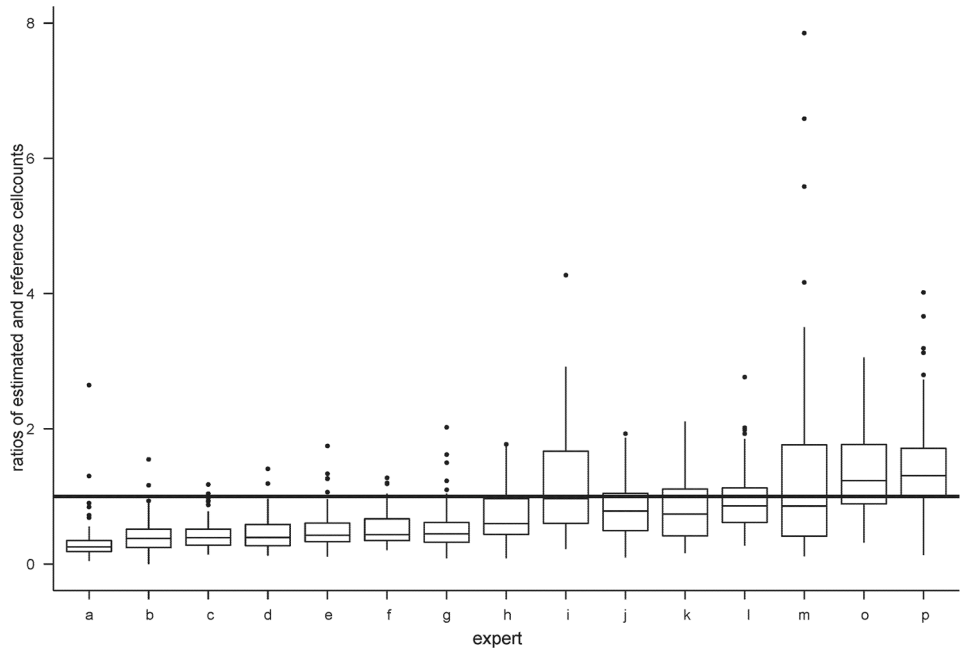


Fig. 6. Distribution of estimated cell counts provided by the 15 experts. Values on the y axis are the ratios of estimated and reference cell counts. A horizontal line at a ratio of 1 indicates the reference. The box plots represent the interquartile range containing the middle 50% of values. Whiskers extend to the highest and lowest values while dots show the outliers.

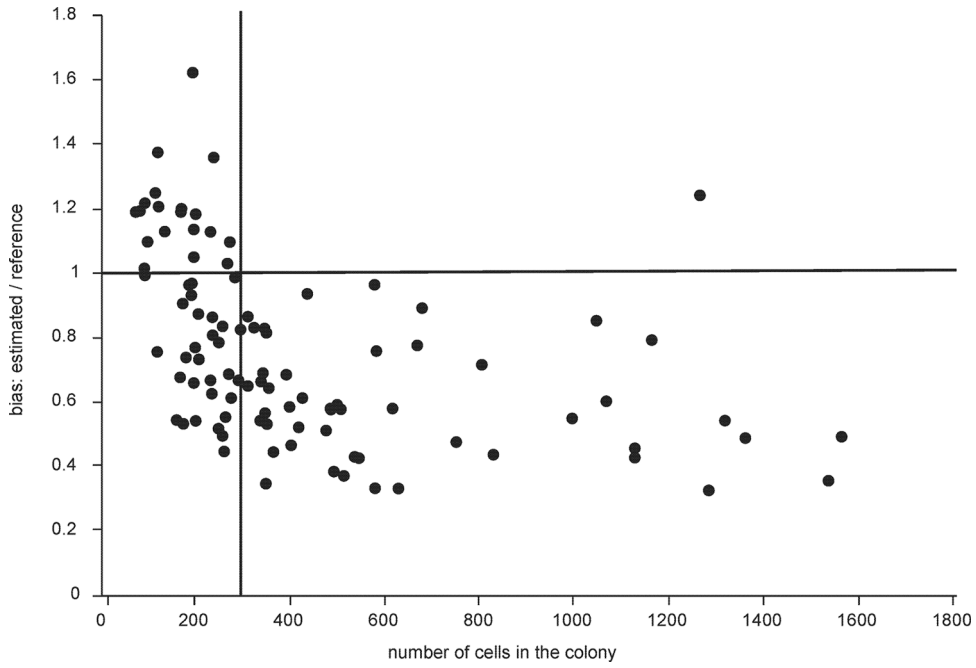


Fig. 7. Relationship between the number of cells in the colony and the estimation bias expressed as: estimated /reference (reference at the horizontal line). Cell numbers have been notoriously underestimated when colonial cell numbers exceeded 300 (vertical line).

accuracy of traditional cell counting (A1), geometric (A2) and regression (A3) approaches to help practitioners to decide which approach fits their needs better.

Compared to the geometric (sphere packing) approach ($A2_{3D}$, $A2_{ell}$), the mean values of experts' cell number estimations (A1) showed slightly low accuracy but high precision. These surprisingly good results can be accounted for by an innate ability of the human mind and learned skills. We humans and some other species are capable of subitizing (Kaufman and Lord, 1949), i.e. to recognize a small group of objects without counting. This evolutionarily primary numerical ability is a fundamental skill in the development of number sense. Subitizing, however, has got its limitations, as it cannot perform well on multiple sets, or a high number of items (Liu et al., 2020). Cell number estimation however is analogous to the problem of crowd counting, which is an active research topic in computer vision (Ranjan et al., 2018; Annadi et al., 2020). The key idea behind crowd counting is simple: area times density. Enumeration of the items in regular crowds can be done by counting them in selected rows and columns, then multiple the values to obtain the final count. This approach is self-evident in the case of regular colonies characteristic for *Merismopedia* or *Eucapsis* genera. In the case of

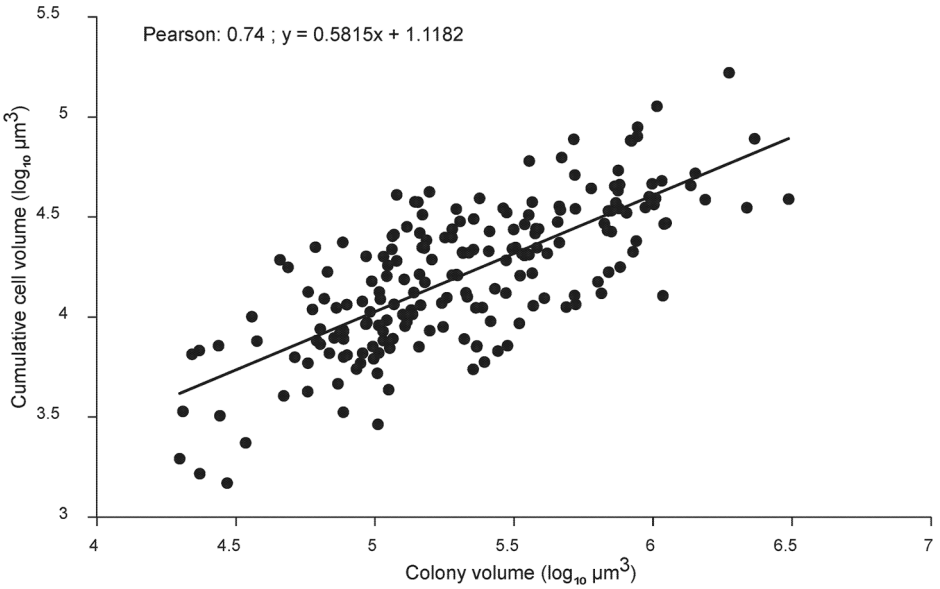


Fig. 8. Colony volume (log_{10} µm^3) and cumulative cell volume (log_{10} µm^3) relationship for cyanobacterial colonies. Colony volumes were determined using 3D images of colonies, while cell numbers were determined after disintegration of colonies. Cells were considered as spheres. Cell diameters were determined during microscopic measurements.

irregular forms, however, this approach cannot be applied. In this study, the experts used the same strategy: they focused on one section of the colony, tried to count the cells and extrapolated to other sections. This approach was especially successful in the case of colonies with uniform density.

Although in terms of mean values traditional cell count estimations of the experts provided good results, considerable differences were observed at an individual level. The fact that 80% of the observers notoriously underestimated the number of cells in the colonies can be accounted for by the difficulty to make an estimation for the spatial extension of a three-dimensional object from two-dimensional images provided by the microscope. Some of the experts have achieved good results both in terms of accuracy (position of the median) and precision (data dispersion), which indicates that cell count estimation is a learnable skill.

The results of the two geometric (sphere packing) approaches ($A2_{3D}$, $A2_{ell}$) showed considerable differences, which can be traced back to differences in the applied overall colony volume estimations.

The multi-layer photographs taken on the colonies and the subsequent 3D modelling of them revealed high diversity in colony shapes. Spherical, ellipsoid, tube-like or irregular reticulate forms also occurred. Therefore, not surprisingly the $A2_{ell}$ approach, i.e. the approach that treats cyanobacterial colonies as ellipsoids (Alácantra et al., 2018) resulted in considerable bias in volume estimation. Colony shape is a species-specific property (Komárek and Komárková, 2002), but because of the high phenotypic plasticity of colonial cyanobacteria, the size and morphology of colonies can change in response to various biotic and abiotic factors (Xiao et al., 2018). Infochemicals released from some grazing flagellates induce *Microcystis* to produce mucilage and form colonies (Burkert et al., 2001; Yang and Kong, 2012). amongst abiotic constraints low temperature and shortage of light (Li et al., 2013; Xu et al., 2016), high concentration of lead (Bi et al., 2013) or calcium (Wang et al., 2011) also induce the formation of mucilaginous colonies.

Summing up, these results demonstrated that the proposed geometric approach, which is based on close sphere packing (A2), requires exact colony volume measurements. This, however, can be done only by using the tedious 3D modelling of the colonies. Other colony volume estimations have considerable bias causing large mistakes in cell count estimations, despite cell size and cell distance measurements being performed accurately.

We found a linear relationship between a colony and cell volumes at the log-log scale (slope value = 0.58; A3 approach). The relationship was strong (Pearson correlation coefficient = 0.74, Fig. 8), indicating that larger overall colony volume coincides with large cumulative cell volumes. However, the amount of scatter in data implies that this approach is also prone to considerable error. The large scatter can be explained by two possible reasons. The first is that we involved different colonial species in the analyses and species-level differences in colony compactness (sum of cell volumes/overall colony volume ratio) do occur amongst the various cyanobacterium taxa (Komárek and, Anagnostidis, 2008). The other reason is that compactness may change even within a species, depending on the actual constraints of the environment (Reynolds and Jaworski, 1978; Wu et al., 2020). Therefore, even if the overall colony volume can be estimated relatively easily, the experienced one order of magnitude differences in the residuals enables only rough estimation of cyanobacterial cell biovolumes.

In the case of many cyanobacterial species, a gelatinous exopolysaccharide envelope (EPS) covering the living cells, filaments and colonies has a prominent appearance in both microscopic and macroscopic examinations (De Philippis et al., 2001; Chen et al., 2016). In addition to being an important taxonomic marker for many species, it has many physiological, ecophysiological and biogeochemical functions such as protective function, frost and temperature tolerance, metal-binding capacity, effect on the nutrient cycles, serving as a substrate and habitat for microorganisms, and a basic structure of colony-forming (Liu et al., 2018; De Philippis and Vincenzini, 1998). Although there can be several questions (concerning nutrient cycling, grazing and sinking properties) where investigation of EPS content cannot be set aside, in water quality monitoring, the applied metrics calculate exclusively with the cell volumes. Therefore, we think that ignoring or considering the volume of mucilage during investigations depends on the questions raised and the study's objectives.

5. Conclusions

The accuracy of various cyanobacterial biovolume and cell count estimation approaches was investigated. We demonstrated that using direct cell counting the involved experts underestimated the cell numbers (A1), but the precision of these estimations was high, and individual differences indicate that estimation competence can be

developed. The proposed geometric model (based on close sphere packing; A2) can be applied only in those cases when overall colony volume can be safely estimated ($A2_{3D}$). We highlighted that overall colony volume estimation from linear dimensions ($A2_{ell}$) has considerable bias. Although the colony volume/cell number relationship was significant (A3), because of the wide range of residuals the regression model provides only a rough estimate for cell counts.

Because of the high uncertainty of overall colony volume estimations and the large variability of colonial cell densities, those approaches can be proposed for monitoring purposes that focus directly on the estimations of cell counts within the colonies.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Hungarian Scientific Research Fund (NKFIH OTKA) project no.: K-132150 and by ÚNKP-21–4-I 232.

Author contributions

G.B. conceived the original idea, E.T.K. and G.B. wrote the manuscript with support from G.V., V.B.B., J.G., Á.L., ZsK., E.T.K. carried out the experiments, V.L. performed the calculations, I.T., T.K. carried out visualisation, V.G. performed the computations. All authors discussed the results and contributed to the final manuscript.

Supplementary information

Supplementary Fig. S1. Photographs of original colonies, flattened colonies, the shape realistic 3D models and counted cell numbers of each measured colony.

Declaration of Competing Interest

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

Data availability

The data is included on the supplematary document.

Acknowledgement

We would like to thank the contributing experts of the Hungarian water quality monitoring network.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.hal.2022.102290.

References

Alcántara, I., Piccini, C., Segura, A.M., Deus, S., González, C., de la Escalera, G.M., Kruk, C., 2018. Improved biovolume estimation of *Microcystis aeruginosa* colonies: a statistical approach. J. Microbiol. Methods 151, 20–27.

Annadi, R., Chen, Y., Ranjan, V., Samaras, D., Zelinsky, G., Hoai, M., 2020. A study of human gaze behavior during visual crowd counting. arXiv preprint arXiv: 2009.06502.

Bi, X.D., Zhang, S.L., Dai, W., Xing, K.Z., Yang, F., 2013. Effects of lead(II) on the extracellular polysaccharide (EPS) production and colony formation of cultured *Microcystis aeruginosa*. Water Sci. Technol. 67, 803–809.

Blender, C.O., 2020. Blender - a 3D modelling and rendering package. Stichting Blender Foundation, Amsterdam Retrieved from http://www.blender.org.

Borics, G., Lerf, V., T-Krasznai, E., Stanković, I., Pickó, L., Béres, V., Várbíró, G., 2021. Biovolume and surface area calculations for microalgae, using realistic 3D models. Sci. Total Environ. 773 https://doi.org/10.1016/j.scitotenv.2021.145538.

Box, J.D., 1981. Enumeration of cell concentrations in suspensions of colonial freshwater microalgae, with particular reference to *Microcystis aeruginosa*. Brit. Phycol. J. 16 (2), 153–164.

Burkert, U., Hyenstrand, P., Drakare, S., Blomqvist, P., 2001. Effects of the mixotrophic flagellate Ochromonas sp. on colony formation in *Microcystis aeruginosa*. Aquat. Ecol. 35, 11–17.

Carvalho, L., Poikane, S., Solheim, A.L., Phillips, G., Borics, G., Catalan, J., De Hoyos, C., Drakare, S., Dudley, B.J., Järvinen, M., Laplace-Treyture, C., 2013a. Strength and uncertainty of phytoplankton metrics for assessing eutrophication impacts in lakes. Hydrobiologia 704 (1), 127–140.

Carvalho, L., McDonald, C., de Hoyos, C., Mischke, U., Phillips, G., Borics, G., Poikane, S., Skjelbred, B., Solheim, A.L., Van Wichelen, J., Cardoso, A.C., 2013b. Sustaining recreational quality of European lakes: minimizing the health risks from algal blooms through phosphorus control. J. Appl. Ecol. 50 (2), 315–323.

CEN - EN 16695: 2015. Water quality – guidance on the estimation of microalgal biovolume.

Chen, M., Xie, M., Wu, W., Yu, F., Li, P., 2016. Chemical characteristics of capsular polysaccharide and water-soluble released exopolysaccharide from *Microcystis*. J. Lake Sci. 28, 609–615.

De Philippis, R., Vincenzini, M., 1998. Exocellular polysaccharides from cyanobacteria and their possible applications. FEMS Microbiol. Rev. 22, 151–175.

De Philippis, R., Sili, C., Paperi, R., Vincenzini, M., 2001. Exopolysaccharide-producing cyanobacteria and their possible exploitation: a review. J. Appl. Phycol. 13, 293–299.

Hawkins, P.R., Holliday, J., Kathuria, A., Bowling, L., 2005. Change in cyanobacterial biovolume due to preservation by Lugol's Iodine. Harmful Algae 4 (6), 1033–1043.

Hifi, M., M'hallah, R., 2009. A literature review on circle and sphere packing problems: models and methodologies. Adv. Oper. Res., 150624, 2009.

Hillebrand, H., Dürselen, C.D., Kirschtel, D., Pollingher, U., Zohary, T., 1999. Biovolume calculation for pelagic and benthic microalgae. J. Phycol. 35 (2), 403–424.

Hongve, D., Løvstad, Ø., Bjørndalen, K., 1988. *Gonyostomum semen*—a new nuisance to bathers in Norwegian lakes: with 4 figures in the text. Verh. Int. Ver. Theor. Angew. Limnol. 23 (1), 430–434.

Huisman, J., Codd, G.A., Paerl, H.W., Ibelings, B.W., Verspagen, J.M., Visser, P.M., 2018. Cyanobacterial blooms. Nat. Rev. Microbiol. 16 (8), 471–483.

Ishikawa, K., Kumagai, M., Walker, R.F., 2005. Application of autonomous underwater vehicle and image analysis for detecting the three-dimensional distribution of freshwater red tide Uroglena americana (Chrysophyceae). J. Plankton Res. 27 (1), 129–134. https://doi.org/10.1093/plankt/fbh157.

Jorstad, A., Nigro, B., Cali, C., Wawrzyniak, M., Fua, P., Knott, G., 2015. NeuroMorph: a toolset for the morphometric analysis and visualization of 3D models derived from electron microscopy image stacks. Neuroinformatics 13 (1), 83–92.

Joung, S.H., Kim, C.J., Ahn, C.Y., Jang, K.Y., Boo, S.M., Oh, H.M., 2006. Simple method for a cell count of the colonial cyanobacterium, *Microcystis* sp. J Microbiol 44 (5), 562–565.

Kaufman, E.L., Lord, M.W., 1949. The discrimination of visual number. Am. J. Psychol. 62 (4), 498–525.

Kepler, J., 1611. Strena Seu De Nive Sexangula. Godefrid Tampach, Frankfurt.

Komárek, J., & Anagnostidis, K.C. 2008. Teil 1/Part 1: Chroococcales. Süßwasserflora von Mitteleuropa; Ettl, H., Gerloff, J., Heynig, H., Mollenhauer, D., Eds, 1–556.

Komárek, J., Komárková, J., 2002. Review of the European *Microcystis*-morphospecies (Cyanoprokaryotes) from nature. Fottea 2 (1), 1–24.

McQuaid, N., Zamyadi, A., Prévost, M., Bird, D.F., Dorner, S., 2011. Use of *in vivo* phycocyanin fluorescence to monitor potential microcystin-producing cyanobacterial biovolume in a drinking water source. J. Environ. Monit. 13 (2), 455–463. https://doi.org/10.1039/C0EM00163E.

Paerl, H.W., Fulton, R.S., Moisander, P.H., Dyble, J., 2001. Harmful freshwater algal blooms, with an emphasis on. Cyanobacteria. Sci. World J. 1, 76–113.

Li, M., Zhu, W., Gao, L., Lu, L., 2013. Changes in extracellular polysaccharide content and morphology of *Microcystis aeruginosa* at different specific growth rates. J. Appl. Phycol. 25, 1023–1030.

Liu, W., Zheng, P., Huang, S., Cicchini, G.M., 2020. Subitizing, unlike estimation, does not process sets in parallel. Sci. Rep. 10 (1), 1–8.

Liu, L., Huang, Q., Qin, B., 2018. Characteristics and roles of *Microcystis* extracellular polymeric substances (EPS) in cyanobacterial blooms: a short review. J. Freshw. Ecol. 33 (1), 183–193.

Ranjan, V., Le, H., Hoai, M., 2018. Iterative crowd counting. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 270–285.

Reinl, K.L., Brookes, J.D., Carey, C.C., Harris, T.D., Ibelings, B.W., Morales-Williams, A. M., De Senerpont Domis, L.N., Atkins, K.S., Isles, P.D., Mesman, J., North, R.L., 2021. Cyanobacterial blooms in oligotrophic lakes: shifting the high-nutrient paradigm. Freshw. Biol. http://archive-ouverte.unige.ch/unige:153778.

Reynolds, C.S., Jaworski, G.H.M., 1978. Enumeration of natural *Microcystis* populations. Brit. Phycol. J. 13 (3), 269–277.

Reynolds, C.S., Walsby, A.E., 1975. Water-blooms. Biol. Rev. 50 (4), 437–481. https://doi.org/10.1111/j.1469-185X.1975.tb01060.x.

Smith, V.H., Schindler, D.W., 2009. Eutrophication science: where do we go from here? Trends Ecol. Evol. 24, 201–207.

Seltenrich, N., 2014. Keeping tabs on HABs: new tools for detecting, monitoring, and preventing harmful algal blooms. https://doi.org/10.1289/ehp.122-A206.

Vasas, G., Márta, M., Borics, G., Gonda, S., Máthé, C., Jámbrik, K., Nagy, Z.L., 2012. Occurrence of toxic *Prymnesium parvum* blooms with high protease activity is related to fish mortality in Hungarian ponds. Harmful Algae 17, 102–110.

Wang, C., Wu, X., Tian, C., Li, Q., Tian, Y., Feng, B., Xiao, B., 2015. A quantitative protocol for rapid analysis of cell density and size distribution of pelagic and benthic *Microcystis* colonies by FlowCAM. J. Appl. Phycol. 27, 711–720.

Wang, Y.W., Zhao, J., Li, J.H., Li, S.S., Zhang, L.H., Wu, M., 2011. Effects of calcium levels on colonial aggregation and buoyancy of *Microcystis aeruginosa*. Curr. Microbiol. 62, 679–683.

WHO (World Health Organisation), 1999. In: Chorus, I., Bartram, J. (Eds.), Toxic Cyanobacteria in water: A guide to Their Public Health consequences, Monitoring and Management. E & F Spon, London.

WHO, 2003. Guidelines For Safe Recreational Water Environments. World Health Organisation, Geneva.

Wu, H., Yang, T., Wang, C., Tian, C., Donde, O.O., Xiao, B., Wu, X., 2020. Co-regulatory role of *Microcystis* colony cell volume and compactness in buoyancy during the growth stage. Environ. Sci. Pollut. Res. 27 (34), 42313–42323.

Xiao, M., Li, M., Reynolds, C.S., 2018. Colony formation in the cyanobacterium *Microcystis*. Biol. Rev. 93 (3), 1399–1420.

Xu, F., Zhu, W., Xiao, M., Li, M., 2016. Interspecific variation in extracellular polysaccharide content and colony formation of *Microcystis* spp. cultured under different light intensities and temperatures. J. Appl. Phycol. 28, 1533–1541.

Yang, Z., Kong, F., 2012. Formation of large colonies: a defense mechanism of *Microcystis aeruginosa* under continuous grazing pressure by flagellate *Ochromonas* sp. J. Limnol. 71, 61–66.

Zamponi, F., 2008. Packings close and loose. Nature 453 (7195), 606–607.