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

¹Enikő T-Krasznai, ¹Verona Lerf, ¹István Tóth, ¹Tibor Kisantal, ¹Gábor Várbíró, ²Gábor Vasas, ¹Viktória B-Béres, ¹Judit Görgényi, ¹Áron Lukács, ¹Zsuzsanna Kókai, ¹Gábor Borics

¹Centre for Ecological Research, Institute of Aquatic Ecology, Department of Tisza Research, Functional Algology Research Group, 18/c Bem square, H-4026 Debrecen, Hungary

²Department of Botany, Faculty of Science and Technology, University of Debrecen, Egyetem square 1, H-4032 Debrecen, Hungary

Corresponding Author: Enikő T-Krasznai, ekrasznai@gmail.com

Orcid:

Enikő T-Krasznai: 0000-0002-0012-1422

Gábor Várbíró: 0000-0001-5907-3472

Viktória B-Béres: 0000-0002-9632-2703

Áron Lukács: 0000-0003-3617-2843

Abstract

Cyanobacteria are notorious bloom formers causing various water quality concerns, such as toxin production, extreme diurnal variation of oxygen, or 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 colonial biovolume and cell count estimation approaches used in the everyday practice. Using real like 3-dimensional 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. Estimation results of several

experts were surprisingly good which suggest that this capability can be improved, and estimation bias can be reduced to the level acceptable for water quality estimations.

Key words: biovolume, inverted microscope, water quality

Highlights

Cyanobacterial colony – cell biovolume relationships provide biased estimates for cellbiovolumes.

Sphere packing approach provides good cellcount estimates if colony volumes are accurately measured.

Considering cyanobacterial colonies as ellipsoids gives inaccurate volume estimates.

Skilled experts slightly underestimate the cellcounts but dispersion of their estimates is low.

Introduction

Despite the efforts that have been made to reduce the nutrient loads to surface waters, eutrophication remains a major source of concerns for water management (Smith & Schindler, 2009). The process coincides with 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, which can be caused by cyanobacteria and several eukaryotic algae (Reynolds & 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*; Ochrophyta) (Ishikawa et al., 2005), others produce surface scums (e.g. *Botryococcus braunii*; Chlorophyta) (Paerl et al., 2001). The formed blooms can be skin irritating (*Gonyostomum semen*; Ochrophyta) (Hongve et al., 1988), or toxic (e.g. *Prymnesium parvum*; Haptophyta) (Vasas et al., 2012). Cyanobacteria as one of the most notorious groups of phytoplankton, 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, reduce water use (Carvalho et al., 2013). 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, concentration of microcystin, ESP; Alcántara et al., 2018; Seltenrich, 2014; Wang et al., 2015; *in vivo* phycocyanin fluorescence, McQuaid et al., 2011). However, giving the accepted thresholds for cell counts (WHO, 1999, 2003) requires the use of traditional microscopic analyses of samples with identification of taxa and estimation of cell counts and taxon-specific biovolumes (Carvallho 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 numbers of unicells and filaments (considering them as cylinders) can be easily accomplished, estimation of cell numbers in colonial forms has an unknown uncertainty of accuracy, especially in the case of coenobial (*Aphanocapsa*, *Coelomorion*, *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 number per colony. In the case of most methods, 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). 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 to use in the case of blooms with the dominance of *Microcystis* spp. 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 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.

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 propose 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 of cyanobacterial blooms is emerging, because the 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 3-D imagery of microalgae (Borics et al., 2021) and the practical application of the age-old sphere packing problem (Hifi & 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.

Material and methods

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 by Lugol's solution in the field. Samples were stored at 4 °C until the analyses.

Sample processing and giving the reference cellcounts

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 program.

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 cover glass. 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 among the cells a minimum of 20 measurements were taken. 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 was carefully desiccated from under the cover glass, resulting in flattening of the colonies until the cells lied in one layer (Fig. 1.). 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 comparison of the three cell count estimation approaches shown below. Based on the measurements on the cells, we counted cell volumes and overall cell biovolumes.

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 and cell size and cell distance measurements (A2)
3. regression approach based on the colony volume and cell number relationships (A3)

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 experts to give an estimation for the cell numbers in the case of 100 colonies (Fig.2). We have shown at least three photographs of each colony taken in different layers to the experts who had 40 seconds for the visual investigations. Estimated cell counts were registered in paper, and these were tabulated electronically. The ratio between experts' estimations and reference cell numbers were used to illustrate estimation bias of each expert. Results of the experts are illustrated as boxplots.

Geometric approach (A2_{3D} and A2_{ell})

In our geometric approach, we considered colonial cyanobacteria as distantly packed spheres in a 3-dimensional space (Fig. 3). 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 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 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 packing density of loosely packed spheres is approximately 0.55 (Zamponi, 2008), overall cell volume of the enlarged cells will be the 55% of the colony volume. Volume of the enlarged cells can also be calculated, because their radius equals with the half of the distance between cyanobacterial cells (supposing that distances of cell centres have been measured; Fig. 3). 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.

$$V = N V_{cell}$$

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

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

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

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

171

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

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

172 V: overall volume of the cells in the colony

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

174 N: number of cells in the colony

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

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

177 V_{ec} : mean volume of a virtually enlarged cell

178 d_{cell} : mean diameter of the cyanobacterial cells in the colony

179 d_{mean} : mean diameter of the fattened cells, which is equal with the mean distances measured
180 between the cells' centers.

181 r: radius of the sphere (here the cells)

182

183 In this study, two overall colony volume (V_{colony}) values were used: A_{2D} measured on the 3D
184 models and A_{2ell} estimated by the experts using linear measurements on the colonies
185 (Alcántara et al., 2018) (Fig. 4). We made the geometric cell count estimations using the
186 results of both approaches separately and the estimation biases were expressed as
187 estimated/reference ratios.

188

189 Regression approach

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

193 To estimate the accuracy of colony volume estimations we created the real-like 3-D models of
194 each (100) studied colony and these served as references (Fig. 4). For the real-like 3-D
195 visualization, we used the free open source 3-D graphics application Blender 2.79 (Blender,
196 2020). The models were based on the photographs that were taken at different layers on the
197 colonies. Volume of the models were calculated using the NeuroMorf software toolset

(Jorstad et al., 2006). 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):

$$V = \frac{\pi}{6} a h b$$

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 experts (15) were asked to measure the linear dimensions of cells and colonies, moreover the mean distances of cells on 30 colonies (Fig. 4.). Cell diameter measurements were done on randomly selected five cells, while during cell distances 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 2-D CAD application) and all measurements were done using the ruler function of the program. Since cell diameters are 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 3-D approach. Estimation bias was expressed as the ratio between the estimated volumes (Alcántara et al., 2018) and volumes given by 3-D 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 were plotted against the reference colony volumes, and the relationship between variables was studied using OLS (Ordinary Least Square) analysis.

Results

We involved 100 cyanobacterium colonies into the analyses, which based on their morphology belonged to the *Microcystis* and *Chroococcus* genera. Real-like 3-D modelling, cell countings, 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 experts.

In the first step we compared the results of the traditional approach (A1) estimating the cell counts by experts; and the geometric (sphere packing) approaches (A2) estimating the cell numbers by using their average cell sizes and distances in colonies of which volumes were determined by 3-D modelling (A2_{3D}); and by treating them as ellipsoids (A2_{ell}) (Alcántara et al., 2018). We created ratios of these values and the reference ones, i.e. those given by direct counting after colony disintegrations. 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 basis for comparison (Fig.5 A-C).

The above mentioned approaches (A1, A2_{3D}, A2_{ell}) showed considerable differences both in terms of data dispersion and location of the mode. The first approach (Fig. 5A) was characterised by low data dispersion (-70 to 20%). Location of the mode indicated -30% estimation error. In the case of the A2_{3D} and A2_{ell} approaches, which are based on the overall colony volume estimation and cell distances, dispersion of data was larger. When colony volumes were estimated by using 3-D 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 the less accurate predictions (Fig. 5C). Estimated volumes dispersed from -80 to >120% percentage range and the mode fell outside the >110% value.

To study the differences among experts who performed the counting, estimated values were divided by the reference values 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 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 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 (Fig. 8). Volumes of the ellipsoids were calculated on the basis of measured length and width of the colonies and the calculated depth values. Means of the experts' volume calculations were divided with that of the reference (i.e. calculated by the Neuromorph on the 3D models). We plotted these values against the volume of the colonies (Fig. 8). The results showed large uncertainties in colony biovolume estimates. In the majority of cases considering the colonies as ellipsoids resulted in overestimation of the colony volumes. These uncertainties however did not depend on the volumes. Large overestimations occurred along the whole scale of colony volumes.

We found strong linear relationship between colony volumes (based on the 3D models) and cell volumes (using reference cell counts of disintegrated colonies) (Fig.9). The relationship was characterised by high R^2 value ($R^2=0.54$) and even scatter of residuals around the regression line (homoscedastic). However, the residuals cover a range of approx. one order of magnitude, which indicates low precision.

Discussion

Cyanobacteria are the most notorious bloom forming planktic photosynthesizers in freshwaters their biovolumes are frequently used bioindicators in lake quality assessment (Carvalho et al., 2013). While volume of single celled and filamentous forms can be measured accurately, in the case of colonial cyanobacteria the estimation of cell numbers has unknown bias, moreover it is also questionable whether ignoring the volume of mucilage is acceptable or not.

In the case of many cyanobacterial species, the gelatinous envelope covering the living cell, filament and colony has a prominent appearance in both microscopic and macroscopic examinations (De Philippis et al., 2001). Of the many names used, exopolysaccharide (EPS) is perhaps the most appropriate one for the gelatinous forms (Chen et al., 2016). In addition to being an important taxonomic marker for many species, it has many physiological, ecophysiological, and biogeochemical functions. Without wishing to be exhaustive, the protective function, metal binding capacity, its effect on the nutrient-cycles, and the tolerance of the organism to frost and temperature should be emphasized. As well as in many cases it serves as a substrate and habitat for microorganisms (heterotrophs and autotrophs too) and it has effects on the organization of the own cells of the cyanobacterial species (colony forming) (Liu et al., 2018; De Philippis & Vincenzini, 1998).

Considering the amount and significance of EPS, it seems evident to calculate with this product when calculating the biovolume of cyanobacterial species (Liu et al., 2018). However, the heterogeneity, visibility, unique species-, and environment-specific appearances of EPS complicate the issue. Overall, based on the EPS chemistry, it is a hygroscopic acidic heteropolysaccharide natural polymer, that also contains proteins (not to mention many other extracellular organic compounds produced by Cyanobacteria) (Chen et al., 2016). In addition, it can be classified into soluble and insoluble fractions, making it difficult to define exactly how to calculate with this material.

Overall, it is difficult to make a decision whether to consider the cyanobacterial EPS content in studies or not. Although in water quality monitoring the applied metrics calculate exclusively with the cell volumes, there can be several questions (concerning nutrient cycling, grazing, sinking properties) in the case of which investigation of EPS content cannot be set aside.

The study of colony volume / cell volume relationships gave us controversial results. Although the relationship was strong, the large amount of scatter in data implies that this approach is also prone to considerable error. The large scatter partly can be explained by that we involved different colonial species into the analyses. However, colony compactness (sum of cell volumes/overall colony volume ratio) may change even within a species, depending on the actual constraints of the environment (Reynolds & Jaworski, 1978; Wu et al., 2020). The experienced one order of magnitude differences in the residuals enables only rough estimation of cyanobacterial cell biovolumes.

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 approach according to which the colonies should be treated as ellipsoids (Alácantra et al., 2018) resulted in considerable bias in volume estimation. Colony shape is a species-specific property (Komárek & Komárková, 2002), but because of the high phenotypic plasticity of colonial cyanobacteria, size and morphology of colonies can change responding to various biotic and abiotic factors (Xiao et al., 2018). Infochemicals released from some grazing flagellates induce *Microcystis* production of mucilage and colony formation (Burkert et al., 2001; Yang & Kong, 2012). Among 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.

The proposed geometric approach, which is based on close sphere packing (A2), requires exact colony volume estimates. However, this can be done only by using the tedious 3-D 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 are performed accurately.

Comparing to the geometric (sphere packing) approach ($A2_{3D}$, $A2_{cell}$), mean values of experts' cell number estimations (A1) showed slightly lower accuracy but higher precision. These surprisingly good results can be accounted for by an innate ability of human mind and learned skills. We humans and some other species are capable of subitizing (Kaufman & 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, i.e. it cannot perform well on multiple sets, or on 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). 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 and 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 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 it and extrapolated to other sections. This approach was especially successful in the case of colonies with uniform density.

Although in terms of mean values the traditional cell count estimations of the experts provided good results, considerable differences were observed at 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 estimation for the spatial extension of a three dimensional object from two- dimensional images provided by the microscope. Some of the experts provided good results both in terms of accuracy (position of the median) and precision (data dispersion) and this indicates that cell count estimation is a learnable skill.

We investigated the accuracy of various cyanobacterial biovolume and cell count estimation approaches. We demonstrated that overall colony volume estimation from linear dimensions have considerable bias. Although the colony volume - cell number relationship was significant because of the large residual variability for any given value on the x axis the regression model provides only a rough estimate for cell counts. 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}). Cell counts were underestimated by the involved experts (A1), but precision (variance) of these estimations was low, and individual differences indicates that estimation competence can be developed.

Although all cyanobacterium biovolume estimation approaches aim to estimate total cell biovolumes, the volume of mucilage in which cells are embedded can also be important.

Acknowledgement

We would like to thank you the contributing experts.

Disclosure statement: No potential conflict of interest was reported by the authors.

Funding Authors

This work was supported by Hungarian Scientific Research Fund (NKFIH OTKA) project no.: K-132150.

Supplementary information

Supplementary Fig. S1. Photographs of original colonies, flattened colonies, the real-like 3-dimensional models and cellnumbers of each measured colony.

Author contributions

GB conceived the original idea, ETK and GB wrote the manuscript with input from, GV, VBB, JG, ÁL, ZsK. ETK carried out the experiments, VL performed the calculations, IT, TK carried out visualisation, VG performed the computations. All authors discussed the results and contributed to the final manuscript.

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. *Journal of Microbiological Methods*, **151**: 20-27.

387 Annadi, R., Chen, Y., Ranjan, V., Samaras, D., Zelinsky, G. and Hoai, M. (2020). *A Study of*
388 *Human Gaze Behavior During Visual Crowd Counting*. arXiv preprint
389 arXiv:2009.06502.

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

393 Blender, C.O. (2020). *Blender - a 3D modelling and rendering package*. Stichting
394 BlenderFoundation, Amsterdam Retrieved from <http://www.blender.org>.

395 Box, J. D. (1981). Enumeration of cell concentrations in suspensions of colonial freshwater
396 microalgae, with particular reference to *Microcystis aeruginosa*. *British Phycological*
397 *Journal*, **16**(2): 153-164.

398 Burkert, U., Hyenstrand, P., Drakare, S. & Blomqvist, P. (2001). Effects of the mixotrophic
399 flagellate *Ochromonas* sp. on colony formation in *Microcystis aeruginosa*. *Aquatic*
400 *Ecology*, **35**: 11–17.

401 Carvalho, L., Poikane, S., Solheim, A.L., Phillips, G., Borics, G., Catalan, J., De Hoyos, C.,
402 Drakare, S., Dudley, B.J., Järvinen, M. & Laplace-Tretyure, C. (2013). Strength and
403 uncertainty of phytoplankton metrics for assessing eutrophication impacts in lakes.
404 *Hydrobiologia*, **704**(1): 127-140.

405 Carvalho, L., McDonald, C., de Hoyos, C., Mischke, U., Phillips, G., Borics, G., Poikane, S.,
406 Skjelbred, B., Solheim, A.L., Van Wichelen, J. and Cardoso, A.C. (2013). Sustaining
407 recreational quality of European lakes: minimizing the health risks from algal blooms
408 through phosphorus control. *Journal of Applied Ecology*, **50**(2): 315-323.

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

410 Chen M, Xie M, Wu W, Yu F. & Li P. (2016). Chemical characteristics of capsular
411 polysaccharide and water-soluble released exopolysaccharide from *Microcystis*. *Journal*
412 *of Lake Sciences*, **28**: 609–615. Chinese.

413 De Philippis R. & Vincenzini, M. (1998). Exocellular polysaccharides from cyanobacteria
414 and their possible applications. *FEMS Microbiology Reviews*, **22**: 151–175.

415 De Philippis R., Sili C., Paperi R. & Vincenzini, M. (2001). Exopolysaccharide-producing
416 cyanobacteria and their possible exploitation: a review. *Journal of Applied Phycology*,
417 **13**: 293–299.

418 Hifi, M. & M'hallah, R. (2009). A literature review on circle and sphere packing problems:
419 models and methodologies. *Advances in Operations Research*, 2009: 150624,
420 <https://doi.org/10.1155/2009/150624>

421 Hillebrand, H., Dürselen, C. D., Kirschtel, D., Pollinger, U. & Zohary, T. (1999). Biovolume
422 calculation for pelagic and benthic microalgae. *Journal of Phycology*, **35**(2): 403-424.

423 Hongve, D., Løvstad, Ø. & Bjørndalen, K. (1988). *Gonyostomum semen*—a new nuisance to
424 bathers in Norwegian lakes: With 4 figures in the text. *Internationale Vereinigung für*
425 *theoretische und angewandte Limnologie: Verhandlungen*, **23**(1): 430-434.

426 Huisman, J., Codd, G.A., Paerl, H.W., Ibelings, B.W., Verspagen, J.M. & Visser, P.M.
427 (2018). Cyanobacterial blooms. *Nature Reviews Microbiology*, **16**(8): 471-483.

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

431 Joung, S. H., Kim, C. J., Ahn, C. Y., Jang, K. Y., Boo, S. M., & Oh, H. M. (2006). Simple
432 method for a cell count of the colonial cyanobacterium, *Microcystis* sp. *Journal of*
433 *Microbiology*, **44**(5): 562-565.

434 Kaufman, E. L. & Lord, M. W. (1949). The discrimination of visual number. *American*
435 *Journal of Psychology*, **62**(4): 498–525. <https://doi.org/10.2307/1418556>

436 Kepler, J. (1611). *Strena seu de nive sexangula*. Godefrid Tampach, Frankfurt.

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

439 Paerl, H. W., Fulton, R. S., Moisander, P. H. & Dyble, J. (2001). Harmful freshwater algal
440 blooms, with an emphasis on cyanobacteria. *The Scientific World Journal*, **1**: 76-113.

441 Li, M., Zhu, W., Gao, L. & Lu, L. (2013). Changes in extracellular polysaccharide content
442 and morphology of *Microcystis aeruginosa* at different specific growth rates. *Journal of*
443 *Applied Phycology*, **25**: 1023–1030.

- 444 Liu, W., Zheng, P., Huang, S. & Cicchini, G.M. (2020). Subitizing, unlike estimation, does
445 not process sets in parallel. *Scientific reports*, **10**(1): 1-8.
- 446 Liu, L., Huang, Q. & Qin, B. (2018). Characteristics and roles of *Microcystis* extracellular
447 polymeric substances (EPS) in cyanobacterial blooms: a short review. *Journal of*
448 *Freshwater Ecology*, **33**(1): 183-193.
- 449 Ranjan, V., Le, H. & Hoai, M. (2018). Iterative crowd counting. In Proceedings of the
450 European Conference on Computer Vision (ECCV): 270-285.
- 451 Reinl, K.L., Brookes, J.D., Carey, C.C., Harris, T.D., Ibelings, B.W., Morales-Williams,
452 A.M., De Senerpont Domis, L.N., Atkins, K.S., Isles, P.D., Mesman, J. & North, R.L.
453 (2021). Cyanobacterial blooms in oligotrophic lakes: Shifting the high-nutrient
454 paradigm. *Freshwater Biology*. <http://archive-ouverte.unige.ch/unige:153778>
- 455 Reynolds C.S. & Jaworski, G.H.M. (1978). Enumeration of natural *Microcystis* populations.
456 British Phycological Journal, **13**(3): 269-277.
457 <http://dx.doi.org/10.1080/00071617800650331>
- 458 Smith, V.H. & Schindler, D.W. (2009). Eutrophication science: where do we go from here?
459 *Trends in Ecology & Evolution*, **24**: 201-207.
- 460 Seltenrich, N. (2014). Keeping tabs on HABs: new tools for detecting, monitoring, and
461 preventing harmful algal blooms. <https://doi.org/10.1289/ehp.122-A206>
- 462 Vasas, G., Márta, M., Borics, G., Gonda, S., Máthé, C., Jámbrík, K. and Nagy, Z.L. (2012).
463 Occurrence of toxic *Prymnesium parvum* blooms with high protease activity is related
464 to fish mortality in Hungarian ponds. *Harmful Algae*, **17**: 102–110.
- 465 Wang, C., Wu, X., Tian, C., Li, Q., Tian, Y., Feng, B., & Xiao, B. (2015). A quantitative
466 protocol for rapid analysis of cell density and size distribution of pelagic and benthic
467 *Microcystis* colonies by FlowCAM. *Journal of Applied Phycology*, **27**: 711–720.
468 <https://doi.org/10.1007/s10811-014-0352-0>
- 469 Wang, Y. W., Zhao, J., Li, J. H., Li, S. S., Zhang, L. H. & Wu, M. (2011). Effects of calcium
470 levels on colonial aggregation and buoyancy of *Microcystis aeruginosa*. *Current*
471 *Microbiology*, **62**: 679–683.

- WHO (World Health Organisation) (1999). *Toxic Cyanobacteria in water: A guide to their public health consequences, monitoring and management*. (Chorus, I. & Bartram, J., editors), 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. *Environmental Science and Pollution Research*, **27**(34): 42313-42323.
- Xiao, M., Li, M. & Reynolds, C.S. (2018). Colony formation in the cyanobacterium *Microcystis*. *Biological Reviews*, **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. *Journal of Applied Phycology*, **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. *Journal of Limnology*, **71**: 61–66.
- Zamponi, F. (2008). Packings close and loose. *Nature*, **453**(7195): 606-607.

Legend for Figures

Fig. 1. Estimation of colonial cell numbers by flattening the colonies; (D) original colony before desiccation, (E) flattened colony.

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

Fig. 3. Cellnumber estimation by close sphere packing. A: loosely arranged sells 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.

Fig. 4. Estimation of the volume of cyanobacterial colony (A), using real like3-D modelling (B), and approximated by ellipsoid (C).

Fig. 5. Distribution of estimation bias for the three ($A1$, $A2_{3D}$, $A2_{\text{ell}}$) cell count estimation approaches; a: visual estimation of cell count by experts ($A1$); b: close sphere packing using colony volumes derived from 3-D models of colonies ($A2_{3D}$); c: close sphere packing using colony volumes derived from the ellipsoid approximation approach ($A2_{\text{ell}}$). Values on the x axis are differences from the reference expressed as percentages $((\text{estimated}/\text{reference} - 1) \times 100)$. 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.

Fig. 7. Relationship between the number of cells in the colony and the estimation bias expressed as: estimated /reference. Cell numbers have been notoriously underestimated when colonial cell numbers exceeded 500.

Fig. 8. Relationship between the overall colony volume estimation bias and colony volumes (μm^3). Estimation bias is given as the ratio between estimated and reference colony volumes. Estimated volumes were calculated using the ellipsoid approximation approach (Alcántara, et

al., 2018), while volumes based on computer calculations on real like 3-D images of colonies, were considered as reference.

Fig. 9. Colony volume (μm^3) and cumulative cell volume (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.

Figures

Fig.1

Fig.2

Fig.3

Fig.4

549

550

551

Fig.5

Fig.6

Fig.7

Fig.8

Fig.9