



HAL
open science

Development of a collaborative web platform documenting the diversity and extent of diatom deformities

Iane Vallanzasca, Sébastien Boutry, Martin Laviale, Eric Quinton, Soizic
Morin

► To cite this version:

Iane Vallanzasca, Sébastien Boutry, Martin Laviale, Eric Quinton, Soizic Morin. Development of a collaborative web platform documenting the diversity and extent of diatom deformities. *Botany Letters*, 2024, 171 (1), pp.50-53. 10.1080/23818107.2023.2183898. hal-04019816

HAL Id: hal-04019816

<https://hal.science/hal-04019816v1>

Submitted on 9 Mar 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Development of a collaborative web platform documenting the diversity and extent of diatom deformities

Vallanzasca, Iane^{1*}; Boutry, Sébastien¹; Laviale, Martin²; Quinton, Eric¹; Morin, Soizic¹

¹Inrae, UR EABX, 50 avenue de Verdun, 33612 Cestas cedex, France

²Université de Lorraine, CNRS, LIEC, F-57000 Metz, France

*corresponding author: iane.vallanzasca@inrae.fr

Development of a collaborative web platform documenting the diversity and extent of diatom deformities

Key words: open science; diatoms; teratologies; crowd sourcing

Introduction

In the implementation of the European Water Framework Directive (2000/60/EC), benthic diatoms are routinely used to assess the ecological status of streams and water bodies using morphological identification of taxa, based on microscopic observations (Morin et al. 2016). Diatoms are powerful ecological indicators, since diatom community structures reflect a gradient of general degradation in water quality (organic pollution, eutrophication, or pH) integrated over time (Lowe & Pan 1996; Stevenson & Pan 1999; Whitton & Rott 1996). Indices were designed to assess general pollution but fail to detect toxic pollution (e.g. Morin et al. 2016).

Deformities in diatoms are increasingly considered good indicators of water toxicity (Morin et al. 2012; Lavoie et al. 2017, Pandey et al. 2017). Indeed, recent works suggested that the types of teratologies, the species impacted by the deformities, and the severity of the deformity could be indicators of particular stress (Lavoie et al. 2017; Falasco et al. 2021) and/or of the intensity of toxic exposure (Olenici et al. 2017; Cerisier et al. 2019). The revised version of the Biological Diatom Index (BDI; Coste et al., 2009), which is the regulatory indicator used for the assessment of French rivers, includes the proportion of deformed diatoms, whose presence probability profile is an indicator of toxic stress. This information is rarely informed in diatom records, as there is no consensual objective approach for attributing a teratological character to a diatom when the deformation is subtle. Olenici et al. (2020) showed that the scores of the diatom Specific Polluosensitivity Index (SPI; Cemagref, 1982) were overestimated by 1-2 units when the percentage of teratologies was not taken into account, reaching up to

8 units in the case of high abundance of deformed individuals. They concluded that overlooking diatom teratologies in flora records may fail in detecting water impairment in these sites.

To ensure proper consideration of teratologies in the BDI, it is imperative to share knowledge on the types and severity of deformities observed in diatoms, associated with the environmental conditions at their sampling site. To this end, gathering an iconographic collection of deformed diatoms, with the help of all interested diatomists, will make it possible to document the most common teratologies.

In this context, we aim to develop a tool, designed and fed by diatomists, to assist in determining the teratological status of an individual diatom. In line with projects such as diatoms.org, designed for providing information on diatom identification, ecology, and distribution (Spaulding et al. 2021), the development of a collaborative web interface is currently being implemented to acquire a large database of teratology images, illustrating the specific diversity, or on the contrary the stability, of the types of deformations (Falasco et al. 2021) encountered for French mainland diatom species. In the long term, sharing knowledge on the deformations observed on the territory and its associated environmental conditions will be key to guaranteeing the quality of the existing BDI, by optimizing the use of diatom deformation for the assessment of toxic stress.

Hereafter, we present a prototype of the platform, which is currently under development. The platform has two main sections, one handling the collection of data and the other providing a reference guide documenting the teratologies already observed.

Collection of teratology data

Creating the database of deformed diatoms requires different types of data, which can be split into two categories: data specific to the diatom specimen, and sample-related environmental data. The user will be assisted at all steps. For any sample containing multiple teratology pictures, environmental data will be informed once. To simplify data entries and get the benefit of already existing information in national databases, the platform will be connected to the Naiades website gathering data on surface water quality (<https://naiades.eaufrance.fr/>) and the National Service for Water Data and Common Repositories Management: Sandre (<https://www.sandre.eaufrance.fr/>) through an Application Programming Interface (API). Our objective is to get the user to enter all the data we need without overwhelming them, or taking too much of their time.

For the diatom picture, any format and color characteristics of the picture will be accepted; afterwards, it will be automatically homogenized as black and white. It will be reframed by the user in such a way that the diatom of interest occupies most of the picture. The specimen identification will be retrieved either via OMNIDIA code or taxon name.

Directly on the picture, its scale ($\mu\text{m}/\text{pixel}$ ratio) will be recovered, and the degree of deformity will eventually be determined. For outline deformities, the quantification of the deformity will be based on geometric morphometry approaches (e.g. Olenici et al. 2017; Wishkerman & Hamilton 2017; Cerisier et al. 2019). The % surface of deformed ornamentations will be assessed by selecting the area where ornamentations are abnormal, compared to the total diatom area. This module could be used by users of any level of a priori knowledge in morphometrics.

Compliance with copyright international standards will be ensured by collecting the identity of the contributor.

Environmental data associated with the sample will be retrieved from other databases (if available), using sample identification (sampling location and sampling date). Physico-chemical information and diatom counts available in Naiades, the French database gathering water quality data, will be collected directly. In case the site is not referenced in Naiades, or the data is unavailable for any other reason, those pieces of information will be collected directly from the contributors. Moreover, the contributor can also mention any stressor suspected to occur at the sampling location and likely to explain the deformity. Among the stressors associated with the presence of diatom teratologies (reviewed in Lavoie et al., 2017, Millan et al. 2020 or Falasco et al. 2021), are chemical factors (extreme pH, lack of nutrients or silica, presence of organic or inorganic micro-pollutants), physical factors (UV radiations, temperature, desiccation, flow speed, specific environments such as springs, and radioactivity) or biological factors such as crowding.

Such information about potential external stressors will be crucial for further analysis of the relationships between teratology types and intensity with environmental data.

Participative assessment of diatom teratologies

The guide will allow browsing through all diatom taxa referenced on the Sandre website, a core part of the French water information system. Sandre guarantees interoperability by providing reference data sets that are made up of technical specifications and code lists. By using an API, data regarding diatom taxa are kept up to date and updated when needed.

For each taxon, information such as the author, contributor, taxonomic rank, reference, taxon type (teratological, reference, or synonym), taxon code... will be displayed, together with the abnormal forms related to that taxon added by diatomists

(following the steps described above). This iconographic resource of already uploaded abnormal specimens will make it possible for users doubting the teratological status of an individual (considered to be "slightly deformed") to confirm the deformation, if the morphology observed corresponds to a typical "attenuated" variant of the deformation(s) usually exhibited by this species, or invalidate it.

For each abnormal form that is uploaded, a comment section will allow users to discuss whether it is actually, or not, teratology. This discussion space is important to check the level of consensus on the teratology information displayed. This collaborative space will be open to contribution and discussion/validation by any registered operator, comparable to other collaborative tools using peer knowledge (e.g. Wikipedia).

Conclusions and perspectives

In its design, the French diatom index BDI takes into account the percentage of teratologies in diatom index calculations. Despite their relevance, abnormal forms are usually overlooked: it is difficult to assert whether or not a diatom is teratological objectively because of the lack of documentation about existing abnormal forms. This database of deformed diatom pictures, in the form of a web application, will serve as a reference guide for the identification and analysis of abnormal forms. This online database will be fed by all diatomists interested in contributing, leaning on a collaborative approach: although initially designed for the French practitioners, this tool intends to be useful for the overall diatomist community.

This collaborative web platform aims at gathering deformed diatom pictures together with their environmental characteristics, intending to draw relevant correlations between deformities and associated stressors in the long run. This aim will be achieved only if enough data are collected. To encourage participation and collect a great diversity of images, the web platform will be developed to be as user-friendly as

possible and some features may evolve according to users' feedback. A large set of existing literature dealing with teratologies with good iconographical material (and associated environmental or chemical data) could also be included in the database. Finally, artificial intelligence could be used to fully exploit image datasets. Indeed machine learning methods have recently proven their usefulness for the automatic analysis of diatom images and further morphometric measurements (Kloster et al. 2014, 2017, Venkataramanan et al. 2023). Algorithms could be trained to automate some steps that users should do manually, such as detecting the diatom on the image, segmenting an individual diatom, but also ultimately quantifying the deformities (Orenstein et al. 2022). The analysis of the intensity of the deformations could thus be systematized and compared to the gradient of deformations already available in the database of pictures. The acquisition of data on the intensity of the deformations and the species affected, in parallel with the environmental characteristics, will make it possible to analyze more finely the statistical relationships between teratologies and stress and to jointly define deformation limits (degree of divergence in the shape normal, minimum relative abundance) for the consideration or not of deformed individuals in the diatom records for the calculation of the BDI.

Acknowledgements

This project part of the AQUAREF program (French reference laboratory for monitoring aquatic environments), funded by the OFB (Office Français de la Biodiversité).

References

Cemagref, 1982. Etude des méthodes biologiques quantitative d'appréciation de la qualité des eaux, pp. 1-218.

- Cerisier A., Vedrenne J., Lavoie I., Morin S. (2019). Assessing the severity of diatom deformities using geometric morphometry. *Botany Letters* 166:32-40. DOI: 10.1080/23818107.2018.1474800.
- Coste M., Boutry S., Tison-Rosebery J., Delmas F. (2009). Improvements of the Biological Diatom Index (BDI): Description and efficiency of the new version (BDI-2006). *Ecological Indicators* 9: 621–650.
- Falasco E., Ector L., Wetzel C.E., Badino G., Bona F. (2021). Looking back, looking forward: a review of the new literature on diatom teratological forms (2010–2020). *Hydrobiologia* 848: 1675-1753.
- Kloster M., Kauer G., Beszteri B. (2014). SHERPA: an image segmentation and outline feature extraction tool for diatoms and other objects. *BMC Bioinformatics* 15:218.
- Kloster M., Esper O., Kauer G., Beszteri B. (2017). Large-Scale Permanent Slide Imaging and Image Analysis for Diatom Morphometrics. *Applied Sciences* 7:330.
- Lavoie I., Hamilton P.B., Morin S., Kim Tiam S., Gonçalves S., Falasco E., Fortin C., Gontero B., Heudre D., Kojadinovic-Sirinelli M., Manoylov K., Pandey L.K., Taylor J.C. (2017). Diatom teratologies as biomarkers of contamination: are all deformities ecologically meaningful? *Ecological Indicators* 82:539-550.
- Lowe R.L., Pan Y. (1996). Benthic algal communities as biological monitors. In *Algal Ecology: Freshwater Benthic Ecosystems*. Aquatic Ecology Series, Academic Press Inc., San Diego, CA, USA, pp. 705–739.
- Millan F., Izere C., Breton V., Voldoire O., Biron D.G., Wetzel C.E., Miallier D., Allain E., Ector L., Beauger A. (2020). The effect of natural radioactivity on diatom communities in mineral springs. *Botany Letters* 167:95-113.
- Morin S., Cordonier A., Lavoie I., Arini A., Blanco S., Duong T.T., Tornés E., Bonet B., Corcoll N., Faggiano L., Laviale M., Pérès F., Becares E., Coste M., Feurtet-Mazel A., Fortin C., Guasch H., Sabater S. (2012). Consistency in diatom response to metal-contaminated environments. In: *Emerging and Priority Pollutants in Rivers*, Springer, Heidelberg. *Handbook of Environmental Chemistry* 19: 117–146.
- Morin S., Gómez N., Tornés E., Licursi M., Tison-Rosebery J. (2016). Benthic diatom monitoring and assessment of freshwater environments: Standard methods and future challenges. A.M. Romani; H. Guasch; M.D. Balaguer. *Aquatic Biofilms:*

- Ecology, Water Quality and Wastewater Treatment, Caister Academic Press, pp.111-124.
- Olenici A., Blanco S., Borrego-Ramos M., Momeu L., Baciú C. (2017). Exploring the effects of acid mine drainage on diatom teratology using geometric morphometry. *Ecotoxicology* 26:1018–1030.
- Olenici A., Baciú C., Blanco S., Morin S. (2020). Naturally and environmentally driven variations in diatom morphology: Implications for diatom-based assessment of water quality. In: *Modern Trends in Diatom Identification*, p39-50.
- Orenstein E.C., Ayata S.-D., Maps F., Becker É.C., Benedetti F., Biard T., de Garidel-Thoron T., Ellen J.S., Ferrario F., Giering S.L.C., Guy-Haim T., Hoebeke L., Iversen M.H., Kiørboe T., Lalonde J.-F., Lana A., Laviale M., Lombard F., Lorimer T., Martini S., Meyer A., Möller K.O., Niehoff B., Ohman M.D., Pradalier C., Romagnan J.-B., Schröder S.-M., Sonnet V., Sosik H.M., Stemmann L.S., Stock M., Terbiyik-Kurt T., Valcárcel-Pérez N., Vilgrain L., Wacquet G., Waite A.M., Irisson J.-O. (2022). Machine learning techniques to characterize functional traits of plankton from image data. *Limnology Oceanography* 67:1647-1669.
- Pandey L.K., Bergey E.A., Lyu J., Park J., Choi S., Lee H., Depuydt S., Oh Y.T., Lee S.M., Han T. (2017). The use of diatoms in ecotoxicology and bioassessment: Insights, advances and challenges. *Water Research* 118: 39-58.
- Spaulding S. A., Potapova M. G., Bishop I. W., Lee S. S., Gasperak T. S., Jovanoska E., Furey P. C., Edlund M. B. (2021). Diatoms.org: supporting taxonomists, connecting communities. *Diatom Research* 36:291-304.
- Stevenson R.J., Pan Y. (1999). Assessing environmental conditions in rivers and streams with diatoms. In *The Diatoms – Applications for the Environmental and Earth Sciences*. Cambridge University Press, Cambridge, UK, pp. 11–40.
- Venkataramanan A., Faure-Giovagnoli P., Regan C., Heudre D., Figus C., Usseglio-Polatera P., Pradalier C., Laviale M. (2023). Usefulness of synthetic datasets for diatom automatic detection using a deep-learning approach. *Engineering Applications of Artificial Intelligence* 117:105594.
- Whitton B.A., Rott E. (1996). *Use of Algae for monitoring rivers II*, Studia Student. G.m.b.H.
- Wishkerman A., Hamilton P.B. (2017). DiaCurv: a value-based curvature analysis application in diatom taxonomy, *Diatom Research* 32: 351-358.