



OPEN

# Geospatial data on the sediments of Lake Balaton

DATA DESCRIPTOR

Mihály Kocsis<sup>1,2</sup>, László Pásztor<sup>1,2</sup>, András Makó<sup>1,2</sup>, Piroska Kassai<sup>1,2</sup>✉, Kálmán Csermák<sup>3</sup>, Alice Csermák<sup>3</sup>, Erzsébet Aradvári-Tóth<sup>3</sup> & Gábor Szatmári<sup>1,2</sup>

Freshwater lakes in most inhabited areas of the world are threatened by water quality issues. Standard water conservation measures have shown efficiency in the past; however, polluted lakes have only partially recovered from eutrophication. Our knowledge is still incomplete about the sensitivity of these lakes to different anthropogenic sources and to the changes in their internal processes due to global warming. In this article, we present a database comprising sediment data from Lake Balaton (Hungary), which can facilitate further analysis helping to better understand the internal processes and changes occurring in the lake. The published dataset includes the following parameters measured in 4211 lake bed sediment samples: pH<sub>KCl</sub>, calcium carbonate (CaCO<sub>3</sub>), organic carbon (C<sub>org</sub>), total nitrogen (N<sub>total</sub>), soluble phosphorus (AL-P<sub>2</sub>O<sub>5</sub>) and soluble potassium (AL-K<sub>2</sub>O), magnesium (Mg<sup>2+</sup>), zinc (Zn<sup>2+</sup>), copper (Cu<sup>2+</sup>) and manganese (Mn<sup>2+</sup>). We are confident that our database serves as a strong basis for further research relating to freshwater lakes influenced by human activities.

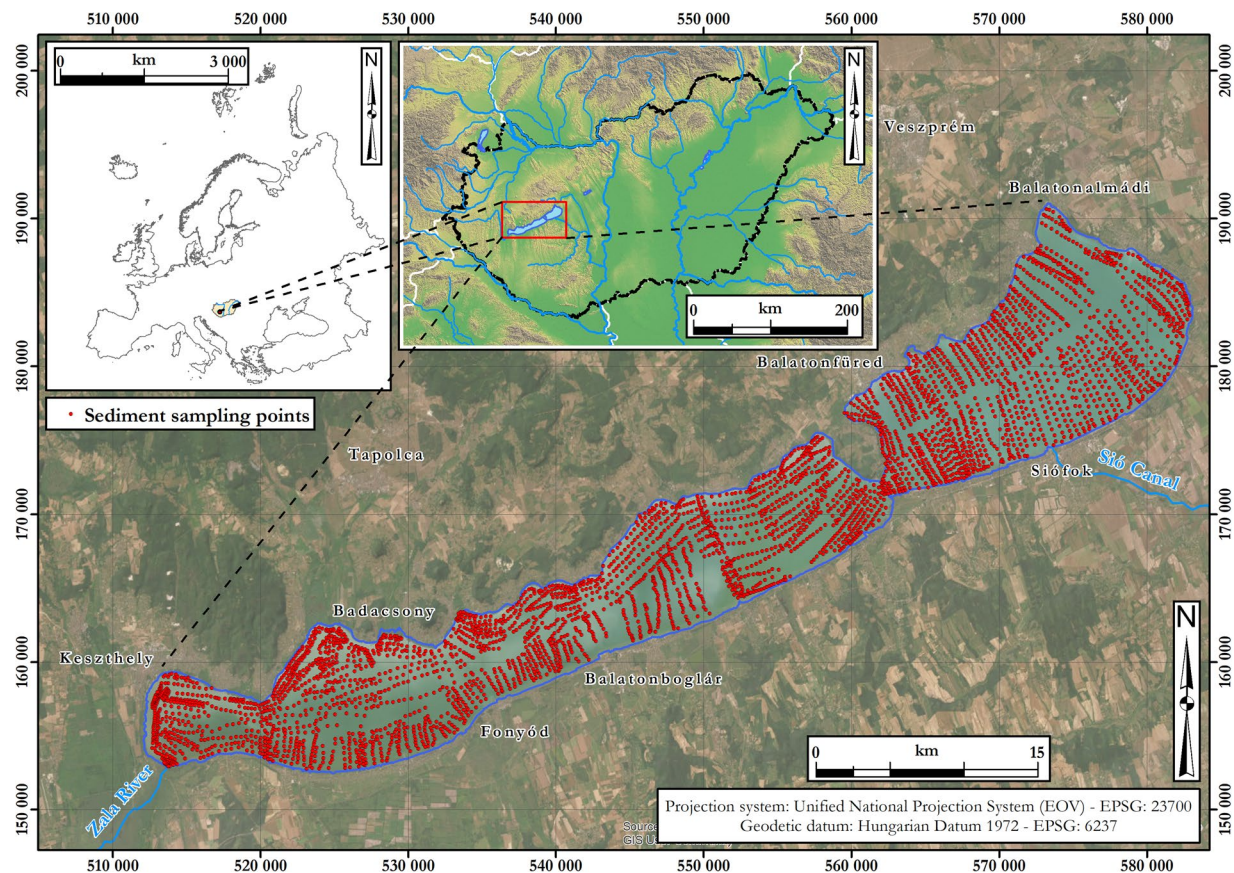
## Background & Summary

Lake Balaton is the largest freshwater lake in Central Europe. It is considered not only one of Hungary's most important natural treasures, but also a highly popular tourist destination with significant economic value. Throughout recent history, there have been major changes in the land use around the lake (urbanisation, railway construction, over-use of surrounding land for agriculture), which have had a significant impact on the extent of the water surface and the position of the lake shore. Unfortunately, these changes have led to the ecological degradation of the lake since the mid-20th century. In the 1970s and 1980s, the lake underwent a rapid eutrophication episode, causing a dramatic decline in water quality<sup>1</sup>. The main cause of eutrophication was the loss of function of the natural filtration/buffer area (now called Kis-Balaton) at the estuary of the lake's largest inflowing river (the Zala River) due to the drop in water levels. Unfortunately, the intensive, yield-oriented agricultural production during the socialist era further exacerbated nutrient overload and by the 1980s it had reached critical levels.

In the following decades, targeted environmental measures (reduction of external phosphorus load and restoration of Kis-Balaton) have halted eutrophication, leading to a substantial improvement in the ecological status of the lake<sup>2-4</sup>. However, despite all the previous efforts to protect water quality, significant algal blooms have been observed again in recent years. Istvánovics *et al.*<sup>5</sup> found that the recurrence of algal bloom events is most likely related to global warming, leading to significant changes in the internal lake processes. Numerous scientists believe that a warming climate affects and modifies the thermal structure of shallow lakes and reduces mixing by increasing the thermal stability of the water column<sup>6,7</sup>. This can lead to a rapid hypolimnetic near-bottom oxygen deficiency, which increases the release of phosphorus (P) from the sediments<sup>8,9</sup>. In the case of Lake Balaton, external load regulation is no longer sufficient to control eutrophication<sup>5</sup>. Since the internal P load is derived from the lake bed, it is very important to obtain and evaluate all available information on the physical and chemical properties of the lake bed sediments to better understand this process.

As sampling and assessment of lake sediments is usually very expensive and labour intensive, detailed surveys are rarely carried out. As climate change intensifies, legacy data can be extremely valuable in comprehending the underlying processes in our changing environment. A positive trend to collect data from previous surveys into open databases and making it accessible for further research, is a response to this demand. There are already numerous open databases created for this purpose. For example, the PLUTO geochemical database contains more than 120,000 records of sediments (lakes, streams and other sediments)

<sup>1</sup>Institute for Soil Sciences, HUN-REN Centre for Agricultural and Environmental Research, Budapest, Hungary. <sup>2</sup>National Laboratory for Water Science and Water Security, Budapest, Hungary. <sup>3</sup>Retired researcher: Kálmán Csermák, Alice Csermák, Erzsébet Aradvári-Tóth. ✉e-mail: [kassai.piroska@atk.hun-ren.hu](mailto:kassai.piroska@atk.hun-ren.hu)



**Fig. 1** Spatial location of Lake Balaton and the 4211 sample points of the sediment survey.

collected and analysed by the United States Geological Survey from across the United States<sup>10</sup>. Some states (Louisiana, Minnesota, or Wisconsin) have their own databases specifically for lake sediment data. Canada also has an open database<sup>11</sup> available for research that preserves historical and current sediment data, remediation, lake and shoreline management, habitat studies, and engineering projects. VARDA (Varved Sediments Database)<sup>12</sup> provides data specifically on annually laminated lake sediments from around the world. This database allows comparison of well-dated and annually laminated sediment records to reconstruct abrupt and regional temporal changes.

Currently, no common open database exists for lake sediments in the European Union, but some Member States have their own platform. No such database has been available in Hungary, despite the recurrent ecological problems associated with Lake Balaton. Although detailed research on the physicochemical properties of the sediments of Lake Balaton was carried out in the 1970s and 80s, at which time interval more than 5,000 samples were collected, it was only accessible to a narrow professional circle. Several results have already been published from the analysis of the database<sup>13,14</sup>, yet the full potential inherent in its use has not been fully realized. So, together with the authors of the survey, we decided to publish this dataset in the form of a digital, open access database. The aim of this study is to present this database and make it widely available for future use both nationally and internationally. On the Zenodo<sup>15</sup> platform, the following parameters are available from the original survey of Lake Balaton sediments:  $\text{pH}_{\text{KCl}}$ ,  $\text{CaCO}_3$ ,  $\text{C}_{\text{org}}$ ,  $\text{N}_{\text{total}}$ ,  $\text{AL-P}_2\text{O}_5$ ,  $\text{AL-K}_2\text{O}$ , Mg, Zn, Cu, Mn.

Potential users of this database are professionals, including hydrologists and limnologists, who are studying the effects of land use and climate change on freshwater lakes to prevent further ecological disasters.

## Methods

**Sampling strategy.** The sampling strategy and laboratory methods of the Lake Balaton sediment survey are summarized based on the work of Csermák & Máté<sup>16</sup>.

The sample density was adjusted to the intended cartographic scale of the planned survey maps, i.e., 1:10,000, which is the same scale as the topographic maps used to identify the lake shore. Based on their soil survey experience, Csermák & Máté<sup>16</sup> found, that mapping on this scale requires about 10 sediment samples per  $\text{km}^2$ . As the sediment was more homogeneous spatially in some areas than expected, the survey team worked with a density of 7–10 samples per  $\text{km}^2$  in the less disturbed inner area of the lake. Altogether, 4211 sediment samples were collected from a total of 592  $\text{km}^2$  of water-covered area (Fig. 1). The vertical sampling design took into account that the top 10 cm layer of sediment plays a more important role in water quality compared to the deeper layers. Since the vertical inhomogeneity of the upper sediment layers may not be significant due to regular mixing and bioturbation, mixed sediment samples were taken from the upper 10 cm layer (without more accurate depth information).

Parameter	Method	Unit
Organic carbon ( $C_{org}$ )	Tyurin method <sup>19</sup>	%
$pH_{KCl}$	Potentiometer	-
Calcium carbonate ( $CaCO_3$ )	Scheibler's calcimeter	%
Hydrolytic acidity ( $hy_1$ )	Kappen's method <sup>20</sup>	-
Total nitrogen ( $N_{total}$ )	Kjeldahl method <sup>21</sup>	%
Soluble phosphorus (AL- $P_2O_5$ )	AL extract, atomic adsorption spectrophotometry <sup>22</sup>	mg kg <sup>-1</sup>
Soluble potassium (AL- $K_2O$ )	AL extract, flame photometer <sup>22</sup>	mg kg <sup>-1</sup>
Magnesium ( $Mg^{2+}$ )	KCl extract, atomic adsorption spectrophotometry <sup>23</sup>	mg kg <sup>-1</sup>
Zinc ( $Zn^{2+}$ )	EDTA + KCl extract, atomic adsorption spectrophotometry <sup>24</sup>	mg kg <sup>-1</sup>
Copper ( $Cu^{2+}$ )	EDTA + KCl extract, atomic adsorption spectrophotometry <sup>24</sup>	mg kg <sup>-1</sup>
Manganese ( $Mn^{2+}$ )	EDTA + KCl extract, atomic adsorption spectrophotometry <sup>24</sup>	mg kg <sup>-1</sup>
Sodium ( $Na^+$ )	AL extract, flame photometer <sup>22</sup>	mg kg <sup>-1</sup>
Sulphate ( $SO_4^{2-}$ )	KCl extract, atomic adsorption spectrophotometry <sup>25</sup>	mg kg <sup>-1</sup>

**Table 1.** Summary of laboratory test methods applied to the sediment samples.

Sampling was carried out at higher spatial frequencies at natural and artificial water inlets and near wastewater treatment plants to detect local anomalies in sediment parameters due to possible contamination. Most samples were taken between 1978 and 1984 in May and August. The latest samples from the lake shore region were collected in three distinct areas and times: the sediments of the reedbeds in 1985–1986, the inflow estuaries in 1987–88, and the area around the sewage treatment plants in 1989. Sampling was carried out with an Ekman-Birge sediment sampler<sup>17</sup>. However, a simple spoon sampler was also used at some sampling points, as the Ekman-Birge grab was impractical for harder sediments.

**Laboratory analyses.** The following parameters were examined from the sediment samples in the laboratory:  $pH_{KCl}$ , hydrolytic acidity ( $hy_1$ ), calcium carbonate ( $CaCO_3$ ), organic carbon ( $C_{org}$ ), total nitrogen ( $N_{total}$ ), soluble phosphorus (AL- $P_2O_5$ ) and soluble potassium (AL- $K_2O$ ), magnesium ( $Mg^{2+}$ ), zinc ( $Zn^{2+}$ ), copper ( $Cu^{2+}$ ), manganese ( $Mn^{2+}$ ), sodium ( $Na^+$ ), sulphate ( $SO_4^{2-}$ ).

Due to the large number of samples, costly and labour-intensive procedures could not be applied. The analyses were carried out in the soil laboratories of the Plant Protection and Agrochemical Stations of Fejér and Zala County (Hungary) using well-established and partially automated methods of soil testing practices. Ammonium-lactate (AL) extractant was used for sample pre-treatment for nutrient content measurements (phosphorus, potassium) as it dissolves the plant-available forms. All test methods (Table 1.) are still commonly used in Hungary as these measures are included in the national standards.

**Geospatial (GIS) processing.** In the early 2000s, the GIS Laboratory of the Research Institute for Soil Science and Agricultural Chemistry, Hungarian Academy of Sciences (RISSAC HAS) digitally processed all sample sites and the related laboratory test results and organized them into a geospatial database. This information had previously been recorded on paper-based maps and protocols. Data from 4211 samples were included in a vector-based point database, from which preliminary digital sediment maps were generated using distance-based buffering (Thiessen polygon formation method) and ordinary kriging for the spatial extension of the measured chemical parameters<sup>16</sup>.

## Data Records

The dataset<sup>15</sup> is available on the Zenodo online repository at <https://doi.org/10.5281/zenodo.10076359>.

The full dataset is organized in a single CSV file that also contains the X Y coordinates defining the spatial location of the sample points. In the published database there are only 10 parameters out of the measured 13 due to significant data gaps. We also performed filtering on the data, which is described in detail in the Technical validation section. The final number of the samples of the published 10 parameters can be seen in Table 2. The coordinate system of the original map was converted from Hungarian Datum 1972 into WGS 1984 in the published database, but we also kept the original (HD 72) coordinates.

Missing values are replaced with ‘-2’ in the database for easier identification.

## Technical Validation

The laboratory tests were carried out in accredited soil laboratories (Plant Protection and Agrochemical Stations of Fejér and Zala County) following the current Hungarian standards. After the vectorization of the data points, we filtered the dataset in two steps: first, we identified data records resulting from digitization errors (e.g., mistyping) and removed them from the dataset. Secondly, we identified data records affected by measurement error using a spatial filtering approach, as follows: (i) we computed the averages of all 10 parameters for data points within 1 km of each other, (ii) we generated raster layers (of all parameters) with the calculated averages, (iii) we determined a threshold for each parameter from the difference between the measured parameter and the computed mean value based on descriptive statistics and distribution curves, and (iv) we identified and removed the data points where the deviation between the measured and the calculated parameter was larger than the threshold. The maximum allowable deviation (threshold value) was 0.25% for organic carbon content, 0.50%

Parameters	Number of samples
Point identification	4 101
Sample identification	4 101
WGS 1984 X coordinate	4 101
WGS 1984 Y coordinate	4 101
HD 1972 X coordinate	4 101
HD 1972 Y coordinate	4 101
Organic carbon ( $C_{org}$ )	3 606
$pH_{KCl}$	3 776
Calcium carbonate ( $CaCO_3$ )	2 867
Total nitrogen ( $N_{total}$ )	4 030
Soluble phosphorus (AL- $P_2O_5$ )	3 340
Soluble potassium (AL- $K_2O$ )	3 665
Magnesium ( $Mg^{2+}$ )	3 963
Zinc ( $Zn^{2+}$ )	3 852
Copper ( $Cu^{2+}$ )	3 994
Manganese ( $Mn^{2+}$ )	3 973

**Table 2.** Number of samples by parameters in the Lake Balaton sediment database.

for  $pH_{KCl}$ , 5% for calcium carbonate content, 20 mg kg<sup>-1</sup> for soluble phosphorus (AL- $P_2O_5$ ) content, 0.50 mg kg<sup>-1</sup> for total nitrogen content, 100 mg kg<sup>-1</sup> for soluble potassium (AL- $K_2O$ ) content, 500 mg kg<sup>-1</sup> for magnesium content, 3 mg kg<sup>-1</sup> for zinc content, 5 mg kg<sup>-1</sup> for copper content, and 30 mg kg<sup>-1</sup> for manganese content. With this filtering procedure, the incorrect values resulting from sampling/measurement errors were removed.

An important question that needs to be addressed is to what extent the values of the measured parameters may have changed since the original survey. A few years after the original survey a resampling was carried out on some sample points and the results showed minimal percentage point deviations compared to the original measurements<sup>16</sup>. However, a number of water protection actions were implemented after the survey, which improved the quality of the settling sediment. As the sediment accumulation rate is 7–10 mm/year<sup>18</sup>, the new fresh sediment on the top layers most likely less polluted. Nevertheless, the internal physical factors (flow conditions, location of inlets) are unchanged, so it can be assumed that although the individual parameters can change over time (e.g., less phosphorus arrives), the spatial pattern remains the same. In addition, the database provides accurate information on the nutrient distribution of past sediment layers that currently positioned in deeper layers, but can be important, e.g., in predicting the environmental effects of planned dredging or other actions affecting the lake bed.

### Code availability

No custom code was used in creating the dataset.

Received: 5 September 2023; Accepted: 9 January 2024;

Published online: 18 January 2024

### References

- Herodek, S. in *Modelling and Managing Shallow Lake Eutrophication* (ed. Somlyódy, L. & van Straten, G.) Ch. 2 (Springer, 1986).
- Honti, M., Gao, C., Istvánovics, V. & Clement, A. Lessons learnt from the long-term management of a large (re)constructed wetland, the Kis-Balaton protection system (Hungary). *Water* **12**/3, 1–17 (2020).
- Hatvani, I. G. *et al.* Spatiotemporal changes and drivers of trophicstatus over three decades in the largest shallow lake in Central Europe, Lake Balaton. *Ecological Engineering* **151**, 105861 (2020).
- Kovács, J. *et al.* Delimiting sub-areas in water bodies using multivariate data analysis on the example of Lake Balaton (W Hungary). *Journal of Environmental Management* **110**, 151–158 (2012).
- Istvánovics, V., Honti, M., Torma, P. & Kousal, J. Record-setting algal bloom in polymictic Lake Balaton (Hungary): A synergistic impact of climate change and (mis)management. *Freshwater Biology* **67**, 1091–1106 (2022).
- Dokulil, M. T., *et al.* in *The impact of climate change on European Lakes. Aquatic ecology series 4* (ed. George, D. G., Springer, 2010).
- Livingstone, D. M. Impact of secular climate change on the thermal structure of a large temperate central European lake. *Climate Change* **57**, 205–225 (2003).
- Pettersson, K., Grust, K., Weyhenmeyer, G. & Blenckner, T. Seasonality of chlorophyll and nutrients in Lake Erken –effects of weather conditions. *Hydrobiologia* **506**, 75–81 (2003).
- Salmaso, N. Effects of climatic fluctuations and vertical mixing on the interannual trophic variability of Lake Garda, Italy. *Limnology and Oceanography* **50**(2), 553–565 (2005).
- Baedecker, P. A., Grossman, J. N. & Buttleman, K. P. *PLUTO geochemical database for the United States Data Series 47*, <https://doi.org/10.3133/ds47> (1998)
- Environment and Climate Change Canada. *Great Lakes Sediment Archive Database 1960–1975*, <https://catalogue.ec.gc.ca/geonetwork/srv/api/records/30fcc196-ea76-48c8-9f48-0ffed2aa3aba> (2021).
- Ramisch, A. *et al.* VARDA (VARved sediments DATabase) – providing and connecting proxy data from annually laminated lake sediments. *Earth System Science Data* **12**, 2311–2332 (2020).
- Kocsis, M. *et al.* Soluble phosphorus content of Lake Balaton sediments. *Journal of Maps* **18**/2, 142–150 (2022).
- Szattmári, G., Kocsis, M., Makó, A., Pásztor, L. & Bakacsi, Z. Joint spatial modeling of nutrients and their ratio in the sediments of Lake Balaton (Hungary): A multivariate geostatistical approach. *MDPI Water* **14**/3, 361 (2022).

15. Kocsis, M. *et al.* Lake Balaton sediment database. *Zenodo*, <https://doi.org/10.5281/zenodo.10076359> (2003).
16. Csermák, K. & Máté, F. *A Balaton talaja* (Veszprémi Egyetem Georgikon Kar, Keszthely, ISBN 9639495441, 2004).
17. Schwoerbel, J. *Methoden der Hydrobiologie* (Kosmos, Stuttgart 95 pp, 1966).
18. Szesztay, K. *Investigation on the silting-up of Lake Balaton* (VITUKI, Budapest, 1963).
19. Tyurin, I. V. A new modification of the volumetric method of determining soil organic matter by means of chromic acid. *Pochvovedenie* **26**, 36–47 (1931).
20. Kappen, H. *Die Bodenazidität*. (Springer Verlag, 1929).
21. Kjeldahl, J. A new method for the determination of nitrogen in organic matter. *Zeitschrift für Analytische Chemie* **22**, 366–382 (1883).
22. Egnér, H., Riehm, H. & Domingo, W. R. Untersuchungen über die chemische Bodenanalyse als Grundlage für die Beurteilung des Nährstoffzustandes der Boden. II. Chemische Extraktionsmethoden zur Phosphor- und Kaliumbestimmung. *Kunglia Lantbrukshögskolans Annaler*, **26**, 199–215 (1960).
23. Mazaeva, M. M. The critical magnesium content of soils. *Agrohimija Moszkva* **10**, 93–105 (1967).
24. Viro, P. J. Use of ethylenediaminetetraacetic acid in soil analysis. *Soil Science* **80/1**, 69–74 (1955).
25. Baranyai, F., Fekete, A. & Kovács, I. *A magyarországi tápanyag-vizsgálatok eredményei* (Mezőgazdasági Kiadó, Budapest, 1987).

## Acknowledgements

The work presented in the article was carried out within the framework of the Széchenyi Plan Plus program with the support of the RRF 2.3.1 21 2022 00008 project. We would like to thank Ágota Horel for her help with English editing.

## Author contributions

Mihály Kocsis: methodology, visualization, writing-original draft preparation. László Pásztor: supervision, data curation, writing - review & editing, funding acquisition. András Makó: conceptualization, supervision, writing - review & editing. Piroska Kassai: conceptualization, writing - original draft preparation, writing - review & editing. Kálmán Csermák: conceptualization, investigation, data curation, Alice Csermák: investigation, data curation. Erzsébet Aradvári-Tóth: investigation, data curation. Gábor Szatmári: conceptualization, methodology, writing - original draft preparation

## Funding

Open access funding provided by HUN-REN Centre for Agricultural Research.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to P.K.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



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

© The Author(s) 2024