B-Cubed: Leveraging Analysis-Ready Biodiversity Datasets and Cloud Computing for Timely and Actionable Biodiversity Monitoring

Quentin Groom[‡], Laura Abraham[‡], Tim Adriaens[§], Lissa Breugelmans[‡], David A Clarke[|,¶], Miguel Alejandro Fernández[#], Louise Hendrickx[‡], Cang Hui[¤], Sabrina Kumschick[¤], Matilde Martini[«], Melodie McGeoch[|,¶], Teodor Metodiev[»], Joe Miller[^], Damiano Oldoni[§], Henrique M. Pereira[#], Cristina Preda[˅], Tim Robertson[^], Duccio Rocchini[«], Hanno Seebens[¦], Heliana Teixeira[?], Maarten Trekels[‡], John Ross Wilson[ς], Nikol Yovcheva[»], Tsungai Zengeya[¢], Peter Desmet[§]

‡ Meise Botanic Garden, Meise, Belgium
§ Research Institute for Nature and Forest (INBO), Brussels, Belgium
| La Trobe University, Melbourne, Australia
¶ Securing Antarctica's Environmental Future, Melbourne, Australia
# German Centre for Integrative Biodiversity Research (iDiv), Leipzig, Germany
¤ Stellenbosch University, Stellenbosch, South Africa
« University of Bologna, Bologna, Italy
» Pensoft Publishers, Sofia, Bulgaria
^ Global Biodiversity Information Facility, Copenhagen, Denmark
˅ Ovidius University of Constanta, Constanta, Romania
¦ Justus-Liebig-Universität Gießen, Gießen, Germany
? CESAM & Department of Biology, University of Aveiro, Aveiro, Portugal
ς South African National Biodiversity Institute, Stellenbosch, South Africa
¢ South African National Biodiversity Institute, Cape Town, South Africa

Corresponding author: Quentin Groom (quentin.groom@plantentuinmeise.be)

Abstract

Effective biodiversity management and policy decisions require timely access to accurate and reliable information on biodiversity status, trends, and threats. However, the process of data cleaning, aggregation, and analysis is often time-consuming, convoluted, laborious, and irreproducible. Biodiversity monitoring across large areas faces challenges in evaluating data completeness and quantifying sampling effort. Despite these obstacles, unprecedented amounts of biodiversity data are being accumulated from diverse sources, aided by emerging technologies such as automatic sensors, eDNA, and satellite tracking.

To address these challenges, the development of tools and infrastructure is crucial for meaningful interpretations and deeper understanding of biodiversity data (Kissling et al. 2017). Furthermore, a significant delay exists in converting biodiversity data into actionable knowledge. Efforts have been made to reduce this lag through rapid mobilisation of biodiversity observations, digitization of collections (Nelson and Ellis 2018), and streamlined workflows for data publication (Reyserhove et al. 2020). However, delays still occur in the analysis, publication, and dissemination of data.

The B-Cubed project (Biodiversity Building Blocks for Policy)*[1] proposes solutions to overcome these challenges. It implements the concept of Occurrence Cubes (Oldoni et al. 2020), which aggregate occurrence data along spatial, temporal and taxonomic dimensions. Cube generation will be available as a new service provided by the Global Biodiversity Information Facility (GBIF). By leveraging aggregated occupancy cubes as analysis-ready biodiversity datasets, we aim to enhance comprehension and reduce barriers to accessing and interpreting biodiversity data. Automation of workflows will provide regular and reproducible indicators and models that are open and useful to users. Additionally, the use of cloud computing offers scalability, flexibility, and collaborative opportunities for applying advanced data science techniques anywhere. Finally, close collaboration with stakeholders will inform us of the requirements for tools, increase impact, and facilitate the flow of information from primary data to the decision-making processes.

Keywords

biodiversity indicator, data cubes, Global Biodiversity Information Facility, policy support

Presenting author

Peter Desmet

Presented at

TDWG 2023

Funding program

B3 (Biodiversity Building Blocks for policy) receives funding from the European Union's Horizon Europe Research and Innovation Programme (ID No 101059592). Views and opinions expressed are however those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the EU nor the EC can be held responsible for them.

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Kissling WD, Ahumada J, Bowser A, Fernandez M, Fernández N, García EA, Guralnick R, Isaac NB, Kelling S, Los W, McRae L, Mihoub J, Obst M, Santamaria M, Skidmore A,

Williams K, Agosti D, Amariles D, Arvanitidis C, Bastin L, De Leo F, Egloff W, Elith J, Hobern D, Martin D, Pereira H, Pesole G, Peterseil J, Saarenmaa H, Schigel D, Schmeller D, Segata N, Turak E, Uhlir P, Wee B, Hardisty A (2017) Building essential biodiversity variables (EBVs) of species distribution and abundance at a global scale. Biological Reviews 93 (1): 600-625. https://doi.org/10.1111/brv.12359

- Nelson G, Ellis S (2018) The history and impact of digitization and digital data mobilization on biodiversity research. Philosophical Transactions of the Royal Society B: Biological Sciences 374 (1763). https://doi.org/10.1098/rstb.2017.0391
- Oldoni D, Groom Q, Adriaens T, Davis AS, Reyserhove L, Strubbe D, Vanderhoeven S, Desmet P (2020) Occurrence cubes: a new paradigm for aggregating species occurrence data. bioRxiv https://doi.org/10.1101/2020.03.23.983601
- Reyserhove L, Desmet P, Oldoni D, Adriaens T, Strubbe D, Davis AJS, Vanderhoeven S, Verloove F, Groom Q (2020) A checklist recipe: making species data open and FAIR. Database 2020 https://doi.org/10.1093/database/baaa084

Endnotes

*1 https://b-cubed.eu/