



BIODIVERSITY
BUILDING
BLOCKS FOR
POLICY

M2 Training Materials

28/02/2025

Author(s): [Laura Abraham](#), [Sabrina Kumschick](#), [Maarten Trekels](#)



Funded by
the European Union

This project receives funding from the European Union's Horizon Europe Research and Innovation Programme (ID No 101059592). Views and opinions expressed are those of the author(s) 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.



Table of contents

Summary.....	3
List of abbreviations.....	3
1 Introduction.....	4
2. Guides, tutorials and training events.....	4
2.1. E-learning and online resources.....	4
2.1.1. Webinars.....	4
2.1.2. Tutorials.....	6
2.1.3. Coding club.....	6
2.2. Training courses.....	6
2.2.1. Courses.....	6
2.2.2. Hackathon.....	6
2.2.3. Live demonstration at COP16.....	7
2.2.4. Demonstrations at BioSpace25.....	7
3. Table of available guides and tutorials.....	8





Summary

The B3 project facilitates training and knowledge sharing on biodiversity indicators through webinars, tutorials and structured courses. Training materials, available under an open license, help researchers and developers create automated workflows on biodiversity indicators.

Key initiatives include a documentation website, GitHub-hosted tutorials, and an e-learning series covering tools like Docker, Google Earth Engine, and biodiversity data cubes. Hands-on hackathon and live demonstrations at global events like COP16 and BioSpace25 promote practical engagement.

Structured training courses, such as the South Africa mini-school, provide in-depth knowledge on biodiversity data cubes. The project also supports open science with interactive webinars and GitHub resources for data analysis and ecological modeling. These efforts collectively enhance accessibility, interoperability, and capacity building in biodiversity informatics.

List of abbreviations

EU European Union





1 Introduction

To ensure uptake of the products developed in B3, training events and tutorials are created to show trainees how to create and use data cubes for biodiversity data, as well as how to create automated workflows for biodiversity indicators. Guides and tutorials are available on the B3 documentation website under an open license. This document gives an overview of the currently available material. Many more training materials are in development and will be created once the products of B3 become available. These materials will be made available on the B3 website when ready.

2. Guides, tutorials and training events

The B3 project has a documentation website on GitHub where information can be found on how to put together tutorials (<https://docs.b-cubed.eu>). This should be followed by project partners when developing such materials.

2.1. E-learning and online resources

2.1.1. Webinars

A series of webinars are organised by and dedicated to B3 partners, aimed at conveying knowledge, exchanging ideas and gaining skills in specific areas related to the expertise of the consortium. The series is coordinated by Matilde Martini, University of Bologna, with a roughly monthly occurrence, throughout much of the duration of the project. Various subjects have been approached so far e.g. exploring Docker, simulation and visualisation of biodiversity data cubes, using Google Earth Engine.

Date: 19th of March from 10:30am to 12:00pm

Title: Go see a Docker! First help with containerizing.

Presenter: Christophe Van Neste (Meise Botanic Garden)

Abstract: "Dependency hell" is a common challenge faced by developers and end-users alike. You find a promising library or program, only to discover that navigating its prerequisites and installation process leads to frustration and wasted time. This is where Docker comes to the rescue. If you've never explored Docker before, or if you're struggling to deploy your tools across different systems, this workshop is designed for you. Discover how Docker containers can simplify software deployment and management, allowing you to focus on your work without getting bogged down in system configurations. Whether you're a developer seeking seamless collaboration or a proponent of reproducible science, this workshop offers practical insights and hands-on guidance. Join us on the Docker workshop and streamline your workflow.

Date: 15th of April from 10:30am to 12:00pm

Title: Unveiling Ecological Dynamics Through Simulation and Visualization of Biodiversity Data Cubes

Presenter: Ward Langerhaert (The Research Institute for Nature and Forest (INBO))

Abstract: Simulation studies offer numerous benefits due to their ability to mimic real-world scenarios in controlled and customisable environments. Ecosystems and biodiversity data are





very complex and involve a multitude of interacting factors. Simulations allow researchers to model and understand the complexity of ecological systems by varying parameters such as spatial and/or temporal clustering, species prevalence, etc. During the B-Cubed hackathon, we created a simulation framework for biodiversity data cubes using the R programming language. The framework starts by simulating a species distributed in a landscape over a temporal scope. In the second phase, the simulation of a variety of observation processes and efforts generates the actual occurrence datasets. Based on their (simulated) spatial uncertainty, occurrences can then be designated to a grid to form a data cube.

The simulation framework can be used to assess multiple research questions under different parameter settings, such as the effect of clustering on occurrence-to-grid designation and the effect of different patterns of missingness on data quality and indicator robustness. Simulation studies can incorporate scenarios with missing data, allowing researchers to assess the impact of data gaps on analyses.

In this workshop, we will go over the results from the B-Cubed hackathon regarding this framework, show how users can utilize the R code in practice, and we can discuss how we can still improve the framework.

Date: 14th June 2024 from 10:30am to 12:00pm CEST

Title: Getting started with Google Earth Engine: Basic JavaScript rules, create, import, explore and visualise GEE datasets

Presenter: Sandra MacFadyen (Stellenbosch University)

Abstract: Sandra will present some Google Earth Engine basics with a practical course that should allow participants to:

1. Understand the basic layout of Google Earth Engine platform (incl. the Code Editor).
2. Understand basic JavaScript syntax rules.
3. Find and import datasets into the code editor.
4. Inspect a dataset in the console.
5. Visualise datasets in the interactive map explorer.
6. Use simple functions.
7. Know where to find help.

Date: 7th October 2024 from 11am. Practical session on 14th October 11am.

Title: EBVCube: Enhancing Biodiversity Data Sharing with Interoperable Geospatial Standards

Presenter: Lina Estupinan-Suarez, Henrique Pereira, Miguel Fernandez

Abstract: Gridded geospatial data on different dimensions of biodiversity, at different spatio-temporal scales are increasingly used to model and estimate biodiversity trends and their relationship with global change. There is an increasing need for an interoperable data standard that covers the complexity and scope of biodiversity observations in a flexible and accessible way for biodiversity researchers and practitioners. Building on efforts made by the climate community, we applied a hierarchical, multidimensional array structure based on the Network Common Data Form (netCDF) to the Essential Biodiversity Variables (EBV) framework promoted by GEO BON, which we call the EBVCube format. This data format allows for a more comprehensive data management and analysis and at the same time accommodates data diversity and scales efficiently. The EBVCube format is supported by the ebvcube R package which provides tools to create, access, and visualise these datasets, streamlining the process and reducing complexity. Together with the EBV Data Portal, this approach offers a complete solution to manage biodiversity data efficiently and follow best practices for data sharing.





2.1.2. Tutorials

Online tutorials and guides are available on our searchable and versioned documentation website <https://docs.b-cubed.eu/>. This website contains information on the cube structure and generation, but also a list of tutorials is available that people can use in a hands-on manner.

2.1.3. Coding club

The Research Institute for Nature and Forest (INBO) organises a coding club (<https://inbo.github.io/coding-club/>) every month. It is an open science initiative with a clear mission: it unites all ecologists, researchers, and coders who want to develop their R programming skills in a pleasant and supportive environment to replace the fear of coding with inspiration and motivation to learn. The INBO coding club is intended for INBO personnel, but external parties (including B3 partners) can ask for an invitation to join. This initiative was advertised during the B-Cubed General Assembly in May and other executive board meetings. People interested in participating can contact Damiano Oldoni by email or via Slack.

2.2. Training courses

2.2.1. Courses

A training course was held in South Africa on data cubes from 2nd to 23rd October 2024. The mini-school was aimed at academics, postgraduate students, managers of protected areas, environmental researchers, but is open to anyone interested in attending including to all partners of the B3 project. Training materials are available through the NITheCS YouTube channel (<https://www.youtube.com/@nithecs/videos>).

Title: Empowering Biodiversity Monitoring through Data Cubes: Techniques & Applications for Open Science

Presenters: Sandra MacFadyen, Maarten Trekels

Abstract: Presented by the NITheCS Research Programme Advancing Biodiversity Informatics & Ecological Modelling Research (<https://nithecs.ac.za/modelling-biodiversity>), this 4-lecture Mini-school aims to empower researchers, students, and practitioners with cutting-edge skills in biodiversity informatics and ecological modelling through the innovative application of data cubes. Participants will gain practical expertise in using data cubes for biodiversity monitoring and ecological analysis, fostering collaboration and openness within the scientific community. Each lecture is linked to one or more of the 17 Sustainable Development Goals (SDGs) (<https://sdgs.un.org/goals>).

LECTURE 1 (2 Oct 2024) Introduction to Data Cubes: Fundamentals and Applications

LECTURE 2 (9 Oct 2024) Building Data Cubes with GBIF: A Hands-On Guide for Biodiversity Monitoring

LECTURE 3 (16 Oct 2024) Analysing Data Cubes: Techniques and Tools for Ecological Modelling

LECTURE 4 (23 Oct 2024) Open Science: Sharing and Disseminating Results from Data Cubes

2.2.2. Hackathon





The hackathon held in Brussels in April 2024 explored the potential applications of the cubes. This four-day event, organised at the Herman Teirlinck building as a hybrid meeting, attracted 86 participants from diverse fields, including biology, ecology, and data science, as well as stakeholders from sister projects such as AD4GD and FAIRiCUBE. The event featured a combination of lectures and practical sessions, where participants had the opportunity to experiment with the cubes, identify opportunities, and address challenges with guidance from five B3 technical coaches. Five keynote speakers from related fields and sister projects presented informative and inspiring sessions, setting the stage for participants to explore cube applications and address real-world challenges in data science and environmental research. Over the course of the event, participants worked in teams to develop and refine nine projects, all of which were shared on a dedicated GitHub repository titled “Hackathon-projects-2024.” Participants were encouraged to create separate repositories for their individual software solutions, fostering a collaborative and accessible environment. On the final day, each team presented its project to a jury panel of experts from universities and sister projects, allowing for constructive feedback and showcasing the diversity of innovative approaches explored. The PowerPoint presentations from the final session are available [here](#). The hackathon served as a training ground, promoting collaborative problem-solving and knowledge sharing. This immersive learning experience helped to build capacity among participants while expanding the project’s knowledge base. For further details, check out the [hackathon webpage](#).

2.2.3. Live demonstration at COP16

During the 16th meeting of the Conference of the Parties (COP16) to the Convention on Biological Diversity (CBD), MLU lead the session titled “Advancing NBSAP Reporting Using GEO BON's EBV Data Ecosystem: Live Demonstration and Training” at the GEO-GEO BON Pavilion on October 21, 2024 (Blue Zone). In this session, we showcased the mobilisation of data from GBIF to the EBV Data Portal using the species occurrence software developed for B3. Additionally, the session highlighted the importance of FAIR indicators, a key focus of the B3 Policy Brief led by Senckenberg.

The session aimed to strengthen capacity building and facilitate technology transfer in line with Target 20 of the Global Biodiversity Framework. By breaking down barriers to data access, it aimed to improve the sharing and availability of high-quality data and knowledge. In the process, we will support the broader objectives of Target 21 by ensuring better access to the best available information for biodiversity management and conservation.

2.2.4. Demonstrations at BioSpace25

At the BioSpace25 conference, two demonstration sessions were organized around GBIF data and biodiversity data cubes.

Title: GBIF: Effective use of GBIF-mediated data

Presenting: Andrew Rodrigues

Abstract: GBIF - the Global Biodiversity Information Facility - provides free and open access to over 3 billion species occurrence records to anyone with an account with the organisation. The use of this valuable data source is increasing year on year, with over 12000 peer-reviewed publications using GBIF-mediated data. It is a key data source for monitoring the state of





biodiversity. In this session, we aim to showcase the principles of good use of GBIF-mediated data and will address:

- Data Sources and Standards
- An introduction to GBIF data publication workflows and how the data is organised
- Data Quality - key data quality issues that users should be aware of, and how to deal with such issues in data use
- Data Access - the different mechanisms for accessing data including APIs, cloud computing and SQL downloads, and how to correctly cite data use

The session will be through a set of presentations, interspersed with guided navigation through GBIF resources on www.gbif.org, to support the effective use of GBIF-mediated data for all those who have used, or are planning to use, GBIF-mediated data.

Title: Biodiversity Data Cubes

Presenters: Maarten Trekels, Lissa Breugelmans, Rocio Beatriz Cortès Lobos

Abstract: This demo session introduces researchers and policymakers to the innovative use of data cubes in biodiversity informatics and ecological modeling, as they are developed in the B-Cubed project (<https://b-cubed.eu/>). Participants will gain insights into practical applications for biodiversity monitoring and the development of indicators, emphasizing collaboration and open science.

Session Outline:

1. Introduction to GBIF Data Cubes and the B-Cubed project (30 minutes). Participants will be introduced to biodiversity data cubes, their structure, and their applications in biodiversity monitoring. Key topics will include an overview of biodiversity data, common biases, and how the data cubes can be generated using the Global Biodiversity Information Facility (GBIF) APIs. The principles of open science will be highlighted through practical examples.
2. Using Biodiversity Data Cube for Indicators (25 minutes). This section will provide a live demonstration of building workflows to generate indicators out of the GBIF data cubes. During the demonstration, the Phylogenetic Diversity indicator will be explained and applied to a specific case.
3. Ecological Modeling and Simulated Data Cubes (25 minutes). Participants will explore advanced analytical techniques for ecological modeling, focusing on climate-related impacts and adaptation strategies. The session will demonstrate how Virtual Suitability Data cubes can be generated and used in modeling workflows.
4. Open discussion (10 minutes).

Outcome:

Attendees will leave with an understanding of the potential of data cubes for biodiversity and ecological analysis, insights into open science practices, and inspiration to apply these techniques in their own work.

The Jupyter notebooks that were used during the demo session are available on GitHub under <https://github.com/b-cubed-eu/biospace25-demo>

3. Table of available guides and tutorials





Table 1: Overview of available training material

WP + task	Title	Responsible(s)	Type	Source URL	Documentation website
WP3 T3.1	Software development guide	Peter Desmet, Pieter Huybrechts, Maarten Trekels, Laura Abraham	Guides		https://docs.b-cubed.eu/dev-guide/
WP2 T2.1	Species occurrence cube specification	Peter Desmet, Damiano Oldoni, Matthew Blissett, Tim Robertson	Guides		https://docs.b-cubed.eu/occurrence-cube/specification/
WP2 T2.2	GBIF SQL download API	Matthew Blissett and John Waller	Guides		https://docs.b-cubed.eu/gbif-sql-download-api/
WP1 T1.4	FAQ helpdesk	Laura Abraham	Guides	https://docs.google.com/document/d/1eb89zuPnBUK6QxhygEBEiMvY6mIDcniCM4Z9DZ2ZMk/edit?usp=sharing	
WP4 T4.5	gcube R package	Ward Langeriaert	Software		https://docs.b-cubed.eu/gcube/
WP1 T1.4	Spatial data analysis in R	Sandra MacFadyen	Tutorials	https://docs.b-cubed.eu/tutorials/spatial-data-analysis-r/	https://docs.b-cubed.eu/spatial-r/
WP1 T1.4	Species distribution modelling in R	Sandra MacFadyen	Tutorials	https://docs.b-cubed.eu/tutorials/species-distribution-modelling-r/	https://docs.b-cubed.eu/sdm-r/
WP1 T1.4	Ecological modelling with Google Earth Engine	Sandra MacFadyen	Tutorials	https://docs.b-cubed.eu/tutorials/ecological-modelling-google-earth-engine/	https://docs.b-cubed.eu/google-earth-engine/
WP4 T4.1	Suitability cube GitHub	Rocio Beatriz Cortès Lobos	Tutorials	https://github.com/b-cubed-e	





				u/virtual-suitability-cube	
WP4 T4.1	Specification for suitability cubes and their production	Rocio Beatriz Cortès Lobos, Michele Di Musciano, Matilde Martini, Sandra MacFadyen, Duccio Rocchini	Guides	https://docs.b-cubed.eu/guides/suitability-cube/	
WP4 T4.1	Specification for dissimilarity cubes and their production	Sandra MacFadyen, Mukhtar Yahaya and Cang Hui	Guides	https://docs.b-cubed.eu/guides/dissimilarity-cube/	
WP4 T4.1	Specification for invisibility cubes and their production	Sandra MacFadyen, Maarten Trekels, Mukhtar Yahaya, Sabrina Kumschick, Cang Hui	Guides	https://docs.b-cubed.eu/guides/invisibility-cube/	
WP4 T4.2	dissMapR for the production and analysis of dissimilarity cubes	Sandra MacFadyen, Mukhtar Yahaya and Cang Hui	Software	https://docs.b-cubed.eu/software/dissmapr/	
WP4 T4.3	invasiMapR for the production and analysis of network invisibility cubes	Sandra MacFadyen, Maarten Trekels, Mukhtar Yahaya, Sabrina Kumschick, Cang Hui	Software	https://docs.b-cubed.eu/software/invasimapr/	
WP3 T3.3	Data mobilisation from GBIF to the EBV data portal	Lina Estupinan-Suarez	Tutorials		
WP4 T4.5	Scripts to explore the conditions that determine the reliability of models, trends and status comp-unstructured-data	Ward Langeriaert, Emma Cartuyvels, Toon Van Daele		https://github.com/b-cubed-eu/comp-unstructured-data	
WP5 T5.1	General indicators b3gbi	Shawn Dove	Guides	https://github.com/b-cubed-eu/b3gbi	
WP5 T5.2	Phylogenetic indicators GitHub pdindicatoR	Lissa Breugelmans	Software	https://github.com/b-cubed-eu/pdindicatoR	





