


Building Biodiversity Datasets for Invasive Species:  
A BLUE module for core data skills

Sara E. Hansen, Anna K. Monfils, Debra L. Linton  
Central Michigan University

CC BY 4.0



Hello everyone. My name is Sara Hansen and today I'll be sharing our education resource Building Biodiversity Datasets for Invasive Species. We initially developed this module with BLUE and then revised it as part of a 4DEE Faculty Mentoring Network.

# Biodiversity Literacy in Undergraduate Education

Diversity and Inclusion

Defining Core  
Biodiversity Data Skills



Integrating Data  
Science into Biology

Community  
Engagement

QUBES Group



[biodiversityliteracy.com](http://biodiversityliteracy.com)

CC BY 4.0

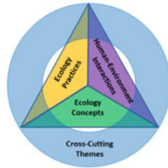


I want to provide a little bit of background on BLUE to give some context for our module and why we developed it. BLUE is Biodiversity Literacy in Undergraduate Education, and is a network that brings together experts in biodiversity data science and education. BLUE's main goals are centered about diversity and inclusion, integrating data science into biology and biology education, defining core biodiversity data skills for science professionals and upcoming scientists, and fostering community engagement and creating resources for the biodiversity and education community. I encourage you to visit our group on QUBES Hub to see our educational resources, or visit our website [biodiversity literacy dot com](http://biodiversityliteracy.com) to learn more about the BLUE network and our webinars and workshops.

# Acknowledgements

## Biodiversity Faculty Mentoring Network

Mentors Luanna Prevost, Teresa Mourad, Anna Monfils



This material is based upon work supported by the National Science Foundation under DBI-1730526 and DBI-2120678. Any opinions, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

CC BY 4.0



This module was made possible by the Biodiversity Faculty Mentoring Network we were able to participate in, so we want to acknowledge ESA and the 4DEE initiative as well as our NSF funding.

# Building Biodiversity Datasets for Invasive Species



<https://qubeshub.org/publications/2865/3>

Sara Hansen, Anna Monfils, and Debra Linton. August 2023. Building Biodiversity Datasets for Invasive Species. Teaching Issues and Experiments in Ecology, Vol. 19: Practice #3.  
[https://tiee.esa.org/vol/v19/issues/data\\_sets/hansen/abstract.html](https://tiee.esa.org/vol/v19/issues/data_sets/hansen/abstract.html)

## Supporting Docs

 [InstructorCopy\\_BuildingBiodiversityDatasets\\_20230711.docx](#) (*Instructors only*)

← Instructor Module Guide

 [StudentCopy\\_BuildingBiodiversityDatasets\\_20230711.docx](#)

← Student Module

 [ExampleDataTemplate.xlsx](#) (*Instructors only*)

← Dataset Examples (Excel key)

[Aggregated occurrence records of invasive European frog-bit \(Hydrocharis morsus](#)

← Online Dataset Link

CC BY 4.0



There are a couple different ways you can access our Building Biodiversity Datasets module. The QR code will take you to the QUBES Hub page, or you can also check out the TIEE article to get a little more background and supporting information. If you visit the QUBES page, these are the resources you'll see. You get an instructor copy that guides you through the module, including how you can split it up if you're short on time or want to just teach some of the concepts. You'll also see the student copy, an example data template for Excel, and a link to a publicly available dataset that's used in the module.

## The module is intended for a wide audience, including:



Undergraduates in life science programs



Graduate students in biology or biodiversity data science



Instructors (high school, undergraduate, and graduate)



Established researchers in biodiversity science



Conservation professionals generating and using biodiversity data

CC BY 4.0



BLUE is all about undergraduate education, so the primary intended audience for this module are undergraduates in the life sciences. We didn't assume any prior knowledge of data or of any particular computer program, because we wanted this to work at all levels of understanding. Even though this was created for the undergraduate level, this module could also be useful for graduate students who didn't have this kind of data-focused education in their undergraduate careers and instructors who want to integrate biodiversity data skills more into their curriculum and want to gain a deeper understanding for themselves first. This module is also for the established researchers and conservation professionals who are collecting and using biodiversity data all the time, but maybe never got formal training or want to learn more about best practices and conventions in this field. We really created this module with all of these people in mind, and we've actually used it in some of these ways ourselves.

**Objective 1:**  
Explain the role of data in monitoring the spread and assessing impacts of non-native species



Figure 1. A thick mat of flowering European frog-bit in Michigan



Figure 2. Saginaw Bay, Michigan and surrounding counties

Module based on the invasive plant European frog-bit

Introduction to biodiversity data

How would a scientist know and be able to explain what is happening with EFB in Saginaw Bay? Would they have to guess based on other species they are familiar with? Would they have to sit in a wetland all year and watch changes happen? No and no. Scientists use **data** to answer their questions. Data are the units of information. Data provide context for information and records of knowledge. Much of what you know about science comes from **data** that were collected, put in the context of a particular question or topic (becoming **information**), and then disseminated to other members of the scientific community and the public (becoming **knowledge**; Brackett 2013). Data take many forms, including lists, tables, or even notes in a notebook. Data are fundamental in the study and preservation of biodiversity, as they allow us to inventory biological organisms or systems, track changes over time, and combine efforts from multiple researchers toward a common goal. By recording what we observe, we are making sure others can use the same information to verify or refute our conclusions.

**Data vs. Datum**

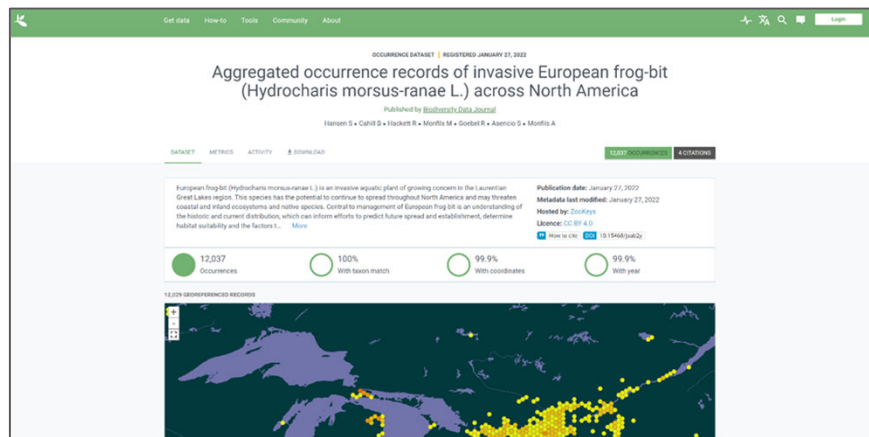
The word "**data**" is a plural noun used for multiple pieces of information. The word "**datum**" is the singular form of the noun used for a single piece of information. For example, one temperature measurement produces one datum, and temperature measurements at several points over time produce many data. You will almost always deal with many data, not just one individual datum, so "**data**" is the more commonly used form of the word.

CC BY 4.0



In these next several slides, I'm going to walk through the module's learning objectives, which we outline right at the beginning, and some of the activities that support each objective. Objective 1 is explaining the role of data in monitoring the spread and assessing impacts of non-native species. The module focuses on an aquatic plant, European frog-bit, which is invasive in the Great Lakes region of North America, and is actually a project we work on in our lab. We first explain the core concepts of biodiversity and invasive species, then explain the specific concerns associated with European frog-bit, and then explain how and why we need to use data in conservation related to this species. Some of our 4DEE themes here were biodiversity, biogeography, specifically species invasions, and resource management.

**Objective 2:**  
Access publicly  
available  
biodiversity data  
through a data  
aggregator



Students interact with a  
published dataset on GBIF

CC BY 4.0



One they get some context from the introduction, students jump right into exploring a dataset as part of the second objective. This a dataset we published in 2022 where we aggregated all publicly available records of European frog-bit in North America. So students can go onto the main GBIF page, get familiar with the interface, and learn a little more about the species ecology in the process. This is where we start to get into the more quantitative 4DEE concepts like data mining, as well as the human dimension of community or citizen science.

**Objective 3:**  
Visualize and  
interpret publicly  
available  
biodiversity data  
through mapping

Mapping by  
time period

Interpret maps relative to  
European frog-bit ecology

CC BY 4.0

As they're checking out this dataset online, we actually ask them to generate a few maps right on the GBIF interface to show how European frog-bit distributions are changing over time. They get a little more practice with the website and try out some data visualizations without needing to learn any software. They then get to think about what these maps are telling us about the invasive spread and how they would then make decisions about management, like which area they would prioritize based on what they're seeing. I like this activity because they're getting all four dimensions at once as they think about the ecology, practice some data skills, and imagine how this species could be most successfully managed.



**Objective 4:**  
Identify ways humans impact the spread of non-native species and data about them

**Formative assessments provide context**

8. How would people in your county react to EFB invasion?

Specific questions to consider:


- Would people in your county care about EFB invasion?
- What would people in your county do to stop the spread of EFB or remove it from the water?
- Are there ways people in your county could accelerate the spread of EFB?
- Is there anything you could do to help prevent or stop the spread of EFB in your county?  
*(Remember, observations recorded by community scientists are significant sources of data for many conservation projects.)*

**Summative assessments check understanding**

a. **Non-native** species can be intentionally or unintentionally introduced by humans. They may be beneficial or become invasive and cause harm.

d. A lot of the publicly available data we have comes from **community scientists**, members of the public who contribute to scientific efforts.

CC BY 4.0



They get to do a little more open-ended practice on the GBIF website, and then we move more into the human dimension. We ask them to think about how humans and invasive species interact and affect one another, but in a very real way rather than just presenting the ideal. Realistically, not everyone in our local communities cares a lot about conservation, and management might not even start until an invasive species becomes a big problem. We want students to really reflect on how the people around them would interact with European frog-bit and what they might personally do to curb the spread in their local areas if it ever spread to them.

**Objective 5:**  
Describe the effects of invasive species on plant communities

**ACTIVITY III: INVESTIGATE LITERATURE ON THE IMPACTS OF INVASIVE AQUATIC PLANTS**

As you know, EFB is an aquatic plant with leaves that float on the surface of the water. Unlike many terrestrial (land) plants, its roots do not need to attach to soil and can hang down into the water. European frog-bit can reproduce asexually through **stolons**, or stems that grow horizontally out from a plant and form new plants (Figure 3). In dense populations, tangled roots and stolons can create thick mats of EFB plants, as you saw in Figure 1. Many of the negative effects of invasive aquatic plants are related to this habit.

In Michigan, EFB is considered invasive due to its potential to cause harm, even though its specific negative effects on native ecosystems are not fully understood. We can infer some potential effects based on what is known about other invasive aquatic plants.

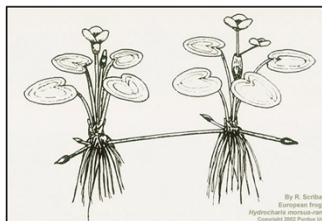


Figure 3. Two EFB plants connected by a stolon

Read about the effects of invasive species from the [National Wildlife Federation](#). Then, find at least one additional webpage or article that addresses the effects of aquatic invasive plants specifically. The webpage or article should come from a reputable source such as a governmental organization, university, or museum.

Online literature provides another way to engage

Finding reputable sources build skills

Objective 5 is describing the effects of invasive species on plant communities. There's a brief introduction to European frog-bit biology and reproduction, which is important for understanding its specific effects because one of the modes of asexual reproduction creates these dense mats that might reduce light and oxygen. They read an online resource about invasive species impacts and then find their own additional webpage or article about invasive plants specifically from a reputable source. Finding reliable information is obviously a very important skill for all scientists so we wanted to make sure we touched on that. Many of us are botanists so we're specific about having them learn more about plants rather than the "popular" invasive animals they might already know a lot about.

Basic terminology and examples are provided

**Objective 6:**  
Apply vocabulary and concepts related to biodiversity datasets

You used data to map the spread of EFB in Activity 1. The GBIF website provides a nice way of visualizing and interpreting data, but you didn't actually see any raw data. So what do data look like? One kind of data that you will encounter frequently are **tabular datasets**, which are groups of columns and rows with values representing each combination. In most cases, columns should contain the types of information collected (**fields**) and rows should contain each piece of information (**observations**).

*Columns store fields, or the types of information collected*

	Restaurant name	City	State	Year established
→	Cochon	New Orleans	Louisiana	2006
→	Waiahole Poi Factory	Kaneohe	Hawaii	2009
→	Indian Pueblo Kitchen	Albuquerque	New Mexico	1976
→	Turtle Tower	San Francisco	California	2000

*Rows store individual observations, or the pieces of information*

CC BY 4.0



The next objective is to apply vocabulary and concepts related to biodiversity datasets. Accessibility was a really important consideration when creating this module, so we didn't want to assume a level of understanding about the data that wouldn't be there for early undergraduates or even established professionals who never had formal training in biodiversity data or are looking to enter into this field for the first time. Throughout the module we explain the basic terms and core concepts that are needed to really engage with the rest of the data components in the module, and we also provide a glossary at the end. We really wanted this to be an introduction to biodiversity data, so even though European frog-bit was great for framing the ecology we wanted to expand our examples beyond it a little bit to increase students' exposure to different types of data.

**Objective 7:**  
Translate plant diversity field sampling protocols to data collection template

Real sampling protocol for European frog-bit

Best practices: planning for data collection before going into the field

**Protocol for Sampling Plant Diversity in Saginaw Bay**  
Sampling will occur for a total of 24 days from June to August. There are 96 plots already set up for sampling, and you will visit each plot once.

1. When you arrive at the sampling plot, use a GPS device to record the latitude and longitude.
2. Record the temperature and amount of dissolved oxygen in the water.
3. Record the depth of the water at the center of the plot using a standard meter stick.
4. Note whether European frog-bit is present or absent in the plot. If present, record the estimated percentage of the plot that is covered, as aerial percent cover (<https://cnhp.colostate.edu/cnhpblog/2009/11/24/field-techniques-percent-cover-estimation/>).
5. Examine European frog-bit for flowers.
6. Record the estimated aerial percent cover of each plant species. Because aquatic plants often overlap, total percent cover may exceed 100%.

**List of fields (columns) to be collected**

<p>EFB flowering</p>	
----------------------	---

Field pictures

CC BY 4.0


In the last two activities, we’re bringing all of the ecology and the data together by asking students to translate our actual field sampling protocol for European frog-bit into a data collection template. So if you are someone that is really interested in the field work, you get to engage with the data in a way you might not have before, and if you’re someone really interested in the data, you get to see where those are coming from and how they connect to the actual biological system. They do a little work in Excel to introduce the program and they learn some best practices like how to name files, how to track versions, how to plan for your data before you go out and collect it. Many of us have our own ways of organizing files on our personal computers, which we like to call “highly personalized data management systems” and we know that sometimes that can cause problems downstream when we can’t find what we’re looking for. So by including the data management component in this module, we’re setting students up correctly from the beginning. Our main 4DEE concepts here are in the ecology practices, so more in the study design and computer skills areas.

**Objective 8:**  
Formulate  
research questions  
that could be  
addressed  
using data

3. You explored the spread and impacts of the non-native invasive species EFB. Now think about a different invasive species in your area (you can look online if needed). Imagine some ways this particular species would impact your area specifically. What is the species and what is one question you want to know that could be tested using data?

4. List the fields you would include in a data template to address your question.



**Broader applications: What questions do you have about other invasive species and how would you test them?**

CC BY 4.0



Our final objective is to formulate research questions that could be addressed using data. We touch on this a couple of times throughout the module, but in the final assessment students get to think about a new species of their choosing and formulate a new question, and then start the process of planning what data they would need. They're basically repeating some of the European frog-bit activities with a new species, which makes this final assessment pretty customizable because you can have them do the other parts like find data on GBIF and map the species if it makes sense.

# Building Biodiversity Datasets for Invasive Species

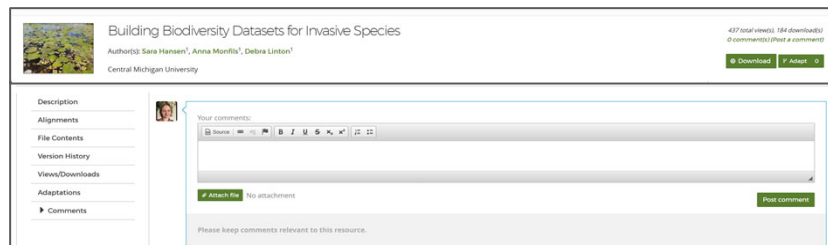


<https://qubeshub.org/publications/2865/3>

Sara Hansen, Anna Monfils, and Debra Linton. August 2023. Building Biodiversity Datasets for Invasive Species. Teaching Issues and Experiments in Ecology, Vol. 19: Practice #3.

[https://tiee.esa.org/vol/v19/issues/data\\_sets/hansen/abstract.html](https://tiee.esa.org/vol/v19/issues/data_sets/hansen/abstract.html)

Please comment using your QUBES account so we can continue to improve this resource!



CC BY 4.0



If you're interested in downloading this module, again here is a link and QR code, or you can search in the QUBES hub for Building Biodiversity Datasets or look for the TIEE publication. If you do use this module, please leave us a comment on QUBES so we can continue to make improvements to this resource. Thank you all very much for listening and I am happy to take your questions.