

Xie_et_all-full_script

Garland Xie

09/01/2022

Preamble

This document was last updated on January 09 2022.

Garland Xie created this annotated R markdown script. The goal of this script is to maximize transparency and reproducibility associated with the following PhD chapter of my dissertation: **Drivers of invasibility in urban meadow restoration (a field study)**.

This script is licensed under CC BY-NC-SA 4.0.

Session info and loading packages

It is important to note that certain functions are preceded by their package sources to help with understanding and reproducing the code (if necessary). This is denoted as “package::function()”.

Load packages

```
library(here)           # for creating relative file-paths
library(google sheets4) # for importing googlesheet files
library(dplyr)          # for manipulating data
library(knitr)           # for creating tables
library(kableExtra)     # for customizing tables
```

Seed bank

Methods

For the spring sampling season, we collected soil samples for 81 1-m² plots across nine sites in three different management regimes (i.e., seed drilling, undisturbed, and maintenance mow) in May 2021. Here, we located the centroid of each plot in the field using a handheld GPS. At each plot, we measured a distance of 30 cm away from the centroid. Before sampling, we removed the surface litter layer or gravel. I then took a soil sample using a 3.5-cm diameter soil core to a depth of 5 cm. This procedure was repeated five times to increase the precision and accuracy of the seed bank estimate (**citation**) and reduce the spatial autocorrelation of community abundance and species richness within plots (**citation**). I combined all five soil samples to one composite seed bank sample per plot into a plastic bag. It is possible to obtain a complete picture of the community dynamics of the seed bank by sampling twice within a growing season such as the spring and fall season (**citation**). I have also repeated the field measurement protocol for obtaining seed bank samples from October-November 2021. Because fall samples require cold temperatures to break winter dormancy,

all spring and fall samples were stored in a freezer at -20 °C for up to three weeks prior to processing at the greenhouse.

Prior to seed emergence methods, we processed soils to remove any litter, roots, rocks, or other debris using a 4 mm sieve. We then filled 81 plastic germination trays with 3 cm of sterilized potting soil. I then spread the field-collected soil samples (one composite sample per plot amounting to 240 cubic centimeter of soil) evenly over the potting soil to ensure complete germination. To serve as a control to track any contaminated seeds from either the greenhouse environment or the potting soil, we filled five additional trays with sterilized potting soil and four additional trays with unsterilized potting (n = 9 for control treatments). All of the trays were randomly dispersed throughout the greenhouse bench. To reduce any spatial effects of environmental heterogeneity (i.e., light intensity and temperature), we completely randomized the trays once a week. We also watered all the trays twice a week to ensure adequate moisture.

We monitored seed germination over a period of four months (July to October 2021). During this time, we scarified the soil if there was at least a week of no seed germination. Any species that were difficult to identify as either a seedling or a juvenile was also collected and placed into a separate tray and grown until flowering. After the fourth month, we removed all the remaining germinants. While previous seed bank study suggests a monitoring period of six months is required, we were primarily interested in species that germinated first and most abundantly as these have competitive advantage in establishment as propagule pressure for both exotic and native plant species. We plan to replicate the seedling emergence for the fall sampling season starting from November 2021 to February 2022.

Import data

Read in data from a Google spreadsheet API

```
# spring sampling season
sb_spr_link <- "https://docs.google.com/spreadsheets/d/101Ll_PsW3qKwdZ_xnTrDKT_kGnpLGvtUKBQ73zvvBBM/edit#gid=148551104"

seed_bank_spr <- googlesheets4::read_sheet(
  sb_spr_link,
  sheet = "raw_data"
)

# fall sampling season
sb_fall_link <- "https://docs.google.com/spreadsheets/d/1SWlk5eWdk3IOMFS9nv61p1HW-Tw4K3yR-tAIkAMlzs4/edit#gid=148551104"

seed_bank_fall <- googlesheets4::read_sheet(
  sb_fall_link,
  sheet = "raw_data"
)
```

Metadata

Descriptions of raw data files:

- **seed_bank_spring** - seedling emergent data from spring season and sampled in the greenhouse from July to October 2021
- **seed_bank_fall** - seedling emergent data from fall season and sampled in the greenhouse from July to October 2021

Descriptions of each variable in the data frames:

- **season**: a factor variable that indicates the sampling season (i.e., Spring or Fall)
- **section**: a factor variable that represents a specific section in the Meadoway (i.e., section 2 or section 4)
- **site**: a factor variable that represents a sampled site (40m X 40m area)
- **treatment**: a factor variable that represents a management regime (i.e., tilling, restored, or mowing) or control treatments (i.e., potting soil, sterilized potting soil)
- **plot**: a factor variable represent a sampled plot within a site
- **date_ymd**: adate of when seedling emergent was identified and recorded (using ISO 8601 format)
- **spp_code**: a character string that represents an abbreviation of the taxonomic species name (see Table X).
- **abund**: an integer variable that represents the seedling emergent abundance for a specific plot
- **sampled_by**: a character string that represents the initials of the surveyor
- **comments**: additional comments