Agriculture and Biodiversity: A Multi-scaled Approach for Conservation

By Emily Hastings
Spring 2022
Virginia Commonwealth University

Purpose

The purpose of this lesson is to understand the impact of agriculture on biodiversity and how a multi-scale approach for conservation is needed to minimize biodiversity loss. In order to understand the need for multi-scaled conservation approaches, we will conduct a meta-analysis using the R programming language.

Learning Objectives

- Understand meta-analysis and the variables involved
- Run linear models using R and understand why they're insufficient in this tutorial's context
- Run a linear mixed model using R to account for random effects
- Create a plot from our linear mixed model outputs

Prerequisites

Students should have an intermediate understanding of statistics which includes linear modeling and basic model outputs. We will be building on this knowledge by running progressively more complex statistical models.

Introduction

Biodiversity & Agriculture

Biodiversity is defined as the variety of life in a given habitat or ecosystem and it is incredibly important to agriculture. Every year thousands of different types of pollinators pollinate crops around the world. It is estimated that pollinators contribute more than $24 billion to the US economy every year by providing their ecological services (Office of the Press Secretary, 2014). Without pollinators, we would not have many of the foods that we consume today. Promoting biodiversity also bolsters other beneficial species that may consume agricultural pests and provide other ecological services to farmers.

Unfortunately, biodiversity is decreasing at an alarming rate largely due to agriculture. Currently, 38% of Earth's surface is in agricultural production (United Nations, 2020). This means that agriculture is a large contributor to biodiversity loss due to land clearing and loss of essential habitats. This is why it is important to understand the mechanisms of agricultural management and its impact on different species so that be can conserve these important resources to protect our food systems and our diverse and unique environments.

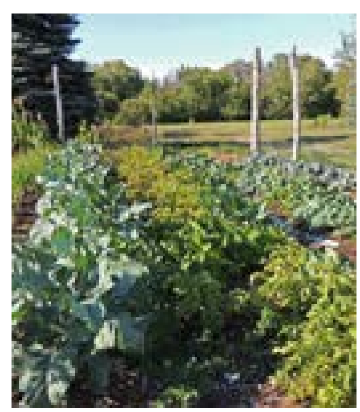
Local Intensity

The two main factors that determine the impact of agriculture on wildlife are local management intensity and landscape complexity. Local mangement intensity is a measure of how densely and intensively crops are being planted. Highly intensive management is typically seen in large-scale conventional agriculture that focuses on producing the most amount of one crop in a given field by planting as closely as possible. Because the crops are so closely planted, they must rely on pesticides and herbicides to stop agricultural pests and unwanted weeds from competing with the crops which would decrease their yield. This practice directly harms plant and invertebrate species with the frequent spraying. It also rids the area of different habitat types and favors generalist species (species that can survive in a wide variety of environmental conditions and make use of a variety of different resources) because they can adapt easily to these oversimplified environments with a lot of only one food source (Ramiadantsoa, 2018).

Local Management Intensity & Landscape Complexity

Intensive Agriculture with Less Landscape Complexity

Less Intensive Agriculture with More Landscape Complexity



Introduction

Landscape Complexity

Landscape complexity is a measure of how complex the landscape is, which can take into account the number of different vegetation types in the landscape, the arrangement of vegetation types in the landscape and the amount of potential habitats in an area for different species (Monmany, 2015). Many smaller-scale organic farms have more landscape complexity, as many different crops are planted in a given area and less or no herbicides are being utilize which allows many more plant species to be in the area to serve as habitat and food. This complexity allows for more habitats which can therefore support more species to exist in the area.

Implications for Conservation

Not all local and landscape factor combinations are created equal. Some species may be impacted more on the localized scale compared to the landscape scale. This is largely due to the species mobility. Mobility allows the species to move to a different area with the resources that they need to thrive. Plants are sessile so they can easily be impacted by intensive tillage and herbicides. Invertebrates are generally mobile but may be slow or not able to respond to pesticides fast enough to avoid their deadly effects. Vertebrates are the most mobile but require a large amount of habitat, food and resources to thrive.

Due to the potential for disproportionate impacts of local intensity and landscape complexity on different types of species, it is important to consider a multi-scale approach for species conservation. This would ensure that different types of species would have access to habitat, which could improve both species abundance (number of individual per species) and species richness (number of difference species) in a given area. Current policy strategies have mainly focused on changing local agricultural management practices but there is a clear need for broader, regional management to conserve habitats and improve biodiversity into the future.

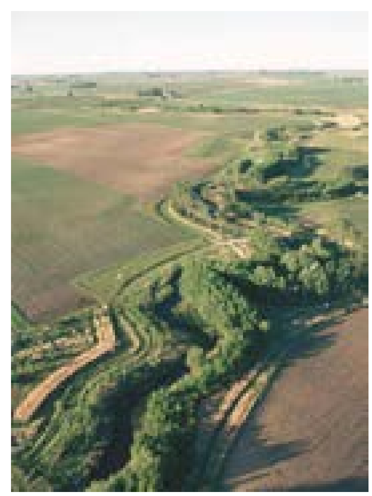
For more information on local intensity & landscape complexity, follow these links:

Local Intensity

Landscape & Ecosystem Complexity

For more information on agricultural conservation measures, follow this link:

Conservation Measures



An example of a riparian buffer to protect water resources from agricultural runoff and provide a wildlife corridor for species to live and move through the landscape.

Introduction

Overview of Meta-analysis

In this tutorial, we will be conducting a meta-analysis. A meta-analysis is a statistical analysis that combines data from multiple studies that address the same scientific question (Mikolajewicz, 2019). It can be used to perform a more complex, multi-scaled analysis that hasn't been conducted before. This is the case for this tutorial. We will be looking at the effect of local intensity, landscape complexity and crop type on species richness for 3 taxonomic groups: vertebrates, invertebrates and plants.

Meta-analyses are conducted differently than a conventional scientific study. Instead of utilizing raw data, the scientist must calculate effect sizes for key variables in order to be able to compare data from separately derived studies. The effect size is the magnitude of experimental effect. A larger effect size indicates a stronger relationship between two variables while a smaller effect size indicates a weaker relationship. In this tutorial, we have two effect size variables and they are measuring the magnitude of the relationship between local intensity and species richness and landscape complexity and species richness for the 3 taxonomic groups. Both of our effect size variables are calculated to be normally distributed.

There are different statistical models that can be applied to meta-analyses. In our case, we will be utilizing a linear mixed model with random effects. We are utilizing this model because we would be producing pseudoreplication by using a simple linear model since there are multiple data points coming from the same study. Pseudoreplication happens when data points are not independent from one another and could create statistically significant relationships where there shouldn't be. In our case, data points from the same study will be heavily related to one another (non-independent) relative to the other data points. We must account for this to ensure that we do not skew our results and create relationships where there shouldn't be.

Steps for a Meta-analysis

Formulate Research Question

↓

Identify Relevant Literature

↓

Extract and Consolidate Data

↓

Calculate Effect Sizes for Approprtiate Variables for Data Comparison Across Studies

↓

Run Models Based on the Nature of Your Data

↓

Synthesize Knowledge and Draw Conclusions

Introduction

Study Overview

The data for this analysis is derived from a meta-analysis conducted by Dr. David J. Gonthier et al. in 2014. The researchers sought to study the effect of local management intensity and landscape complexity on species richness and abundance for 3 taxonomic groups. We will be focusing our analysis on only those studies that address species richness.

Data Overview

The study contained observations of different species from 31 agricultural studies conducted in the USA and Europe. Below are some of the variables used in the study with information on how they were derived. For more information, see the data overview document.

Study

This variable is the number assigned to each of the agricultural studies. Some studies have multiple data points.

Taxonomic Group

This variable is the taxonomic group that was studied in each of the agricultural studies. The groups are vertebrates, invertebrates and plants.

Crop Type

This variable is the type of crops being studied by each of the agricultural studies. These included mixed, cereal, pasture/meadow, and fruit/vegetable crops.

Local Intensity

This variable is the log response ratio which is an effects size measure to quantify functional relationships in terms of proportionate change. This was derived using local intensity and species richness values.

Landscape Complexity

This variable is the Fisher's Z value which is the distribution of Pearson's correlation coefficient transformed to become normally distributed. This was derived using landscape complexity and species richness values.

Lead Researcher
Dr. David J. Gonthier

Dr. Gonthier is an Entomologist at the University of Kentucky who specializes in Agroecological Biodiversity, Ecosystem Services and Landscape Ecology.

Read the reference paper here:

Biodiversity conservation in agriculture requires a multi-scale approach.

Find the original data here:

The Data from Gonthier et al.

Tutorial

Research Question

Research Question:
Is there a difference in effect of local management intensity and landscape complexity combined with crop type on species richness by taxonomic group?

Students will use data derived from the Gonthier et al. study to explore the relationship between local intensity and landscape complexity (the predictor variables) and the 3 taxonomic groups (the response variable). We will be conducting a meta-analysis. This means that our predictor variables are effect sizes based on local and landscape factors with species richness, instead of typical raw data. We must keep this study type in mind, as it will inform what statistical choices we make to create the most robust model to fit our data.

Linear Modeling

We will start the tutorial by running linear models. Linear models are used to determine the relationship between two or more variables. In other words, they determine how well the predictor variable "predicts" the outcome of the response variable. Because our predictor variables are effect sizes calculated to be normally distributed, we do not need to first check for normality. We will be running these models separately for local versus landscape factors.

Local Management Intensity Linear Model

First, let's set up the code for our local factor model.

```
#Local intensity linear model
locallm <- lm(LocalIntense ~ TaxGroup + CropType,  data = conserve)
summary(locallm)
```

Notice that crop type is being added with our model, as different studies had different types of crops that we need to account for. Crop type can affect the types of species present in an agricultural area.

Tutorial

Now let's look at the local model outputs.

```
lm(formula = LocalIntense ~ TaxGroup + CropType, data = conserve)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.79533 -0.23423 -0.06077  0.19416  2.59234 

Coefficients:
                       Estimate Std. Error t value Pr(>|t|)    
(Intercept)             0.30249    0.10750   2.814 0.006468 ** 
TaxGroupplant           0.65134    0.16166   4.029 0.000149 ***
TaxGroupvertebrate     -0.03031    0.17924  -0.169 0.866237    
CropTypefruit/vegtable  0.14109    0.21949   0.643 0.522600    
CropTypemixed          -0.36140    0.18452  -1.959 0.054458 .  
CropTypepasture/meadow -0.22597    0.14454  -1.563 0.122818    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5063 on 65 degrees of freedom
Multiple R-squared:  0.2978,	Adjusted R-squared:  0.2437 
F-statistic: 5.512 on 5 and 65 DF,  p-value: 0.0002725
```

§ Pause § Look at the model outputs and what they say about the relationship between variables. This significance has important implications.

Landscape Complexity Linear Model

Now let's set up the code for our landscape factor model.

```
#Landscape complexity linear model
landscapelm <- lm(LandscapeComp ~ TaxGroup + CropType,  data = conserve)
summary(landscapelm)
```

Tutorial

Let's look at the landscape model outputs.

```
lm(formula = LandscapeComp ~ TaxGroup + CropType, data = conserve)

Residuals:
     Min       1Q   Median       3Q      Max
-0.77815 -0.18336 -0.04813  0.22513  0.57763

Coefficients:
                        Estimate Std. Error t value Pr(>|t|)
(Intercept)             0.149435   0.069853   2.139   0.0362 *
TaxGroupplant          -0.051162   0.105047  -0.487   0.6279
TaxGroupvertebrate      0.009254   0.116473   0.079   0.9369
CropTypefruit/vegtable -0.012492   0.142626  -0.088   0.9305
CropTypemixed           0.271077   0.119907   2.261   0.0271 *
CropTypepasture/meadow -0.005684   0.093922  -0.061   0.9519
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.329 on 65 degrees of freedom
Multiple R-squared:  0.08813,	Adjusted R-squared:  0.01798
F-statistic: 1.256 on 5 and 65 DF,  p-value: 0.2935
```

§ Pause § Look at the model outputs again and what they say about the relationship between variables.

We know that by not accounting for the non-independence of data points from our Study variable we are pseudoreplicating. Therefore, the significance or lack of significance of these relationships cannot be trusted. This goes to show that you must be careful to make sure all of your data is independent or corrected for non-independence accordingly. To account for non-independent data points, we must now run a linear mixed model with random effects. Again, we will be running models separately for local versus landscape factors.

Tutorial

Linear Mixed Modeling

Next, we will begin to set up our linear mixed models with random effects. Linear mixed models are an extension of basic linear models to allow for both fixed and random effects. In our case, we will be using fixed effects (taxonomic group and crop type) like our basic linear model with a random effect (study) because we must account for the non-independence in the data.

Local Management Intensity Linear Mixed Model with Random Effects

Let's set up the code for our local factor mixed model. Our random effect is the Study variable.

```
#Linear Mixed Model with Random Effect for Local Intensity
LocalLMM <- lmer(LocalIntense ~ TaxGroup + CropType + (1|Study), data =
conserve)
```

Now let's look at the model outputs.

```
Linear mixed model fit by REML ['lmerMod']
Formula: LocalIntense ~ TaxGroup + CropType + (1 | Study)
   Data: conserve
REML criterion at convergence: 111.1932
Random effects:
 Groups   Name        Std.Dev.
 Study    (Intercept) 0.0000
 Residual             0.5063
Number of obs: 71, groups:  Study, 31
Fixed Effects:
           (Intercept)          TaxGroupplant      TaxGroupvertebrate
               0.30249                0.65134                -0.03031
CropTypefruit/vegtable           CropTypemixed  CropTypepasture/meadow
               0.14109               -0.36140                -0.22597
optimizer (nloptwrap) convergence code: 0 (OK) ; 0 optimizer warnings; 1
```

§ Pause § Notice that this model understands that there are 31 studies out of 71 total observations. This means we are accounting for non-independence.

Tutorial

Landscape Complexity Linear Mixed Model with Random Effects

Now let's set up the code for our landscape factor mixed model.

```
#Linear Mixed Model with Random Effect for Landscape Complexity
LandscapeLMM <- lmer(LandscapeComp ~ TaxGroup + CropType + (1|Study), data =
conserve)
```

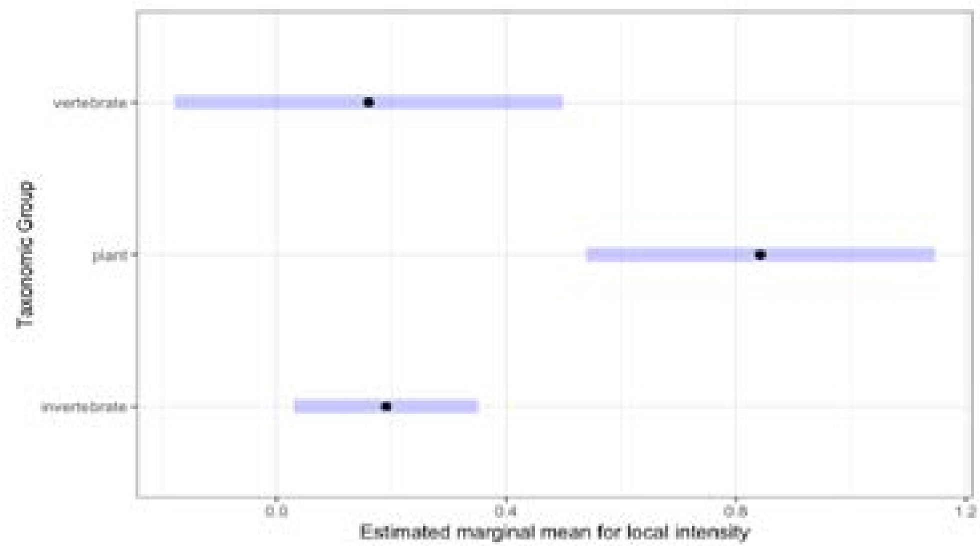
Let's look at the model outputs.

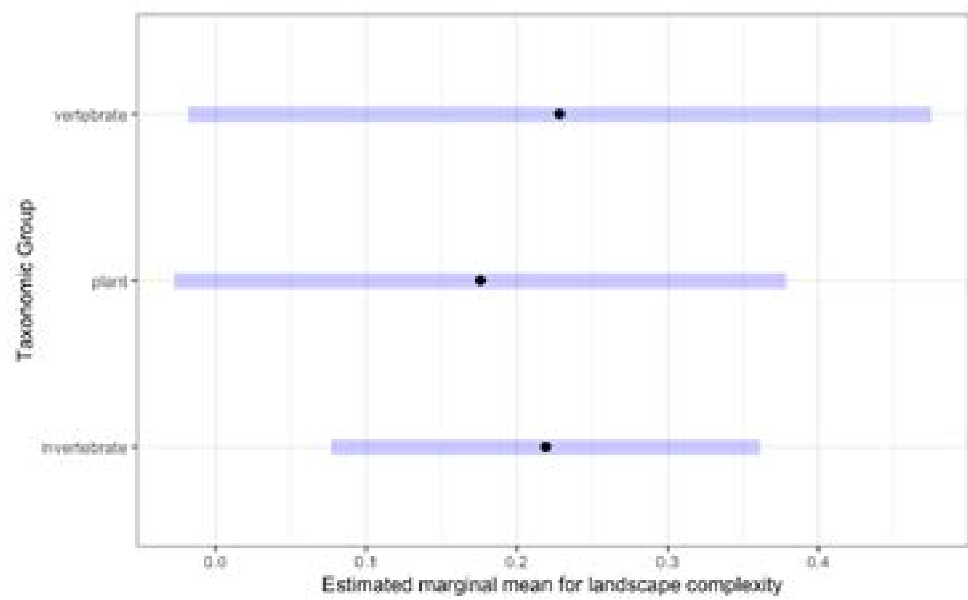
```
Linear mixed model fit by REML ['lmerMod']
Formula: LandscapeComp ~ TaxGroup + CropType + (1 | Study)
   Data: conserve
REML criterion at convergence: 47.8063
Random effects:
 Groups   Name        Std.Dev.
 Study    (Intercept) 0.2112
 Residual             0.2651
Number of obs: 71, groups:  Study, 31
Fixed Effects:
          (Intercept)           TaxGroupplant      TaxGroupvertebrate
             0.165434               -0.043307                0.009133
CropTypefruit/vegtable         CropTypemixed  CropTypepasture/meadow
            -0.053222                0.307305               -0.039297
```

§ Pause § **Notice that this model too is accounting for the random effect of Study to statistically create independence.**

Tutorial

To get a better understanding of the relationships between local and landscape factors on the different taxonomic groups, it may be useful to plot marginal means with 95% confidence intervals. Marginal means are means extracted from a statistical model and represent the average of the response variable for each level of predictor variable. This will tell us if the relationships between our predictor and response variables in our models are significant. All confidence intervals that stay completely above or below zero are considered statistically significant.





§ Pause § **Notice which taxonomic groups have significant relationships between local and landscape factors. Consider why this calls for a multi-scale conservation approach in agriculture.**

Tutorial

Comprehension Questions

1) Why might high intensity farming have a greater affect on species richness compared to low intensity farming?

2) What implications does landscape complexity have for species richness? What factors in a complex landscape might contribute to higher species diversity?

3) How well do the local and landscape linear models explain the data? Is this the best method for assessing this data? Are we not accounting for something?

4) Why should we run a linear mixed model with random effects? Which random effect variable should we choose?

5) Which taxonomic groups have a significant relationship with local intensity?

6) Which taxonomic groups have a significant relationship with landscape complexity?

7) How might local intensity and landscape complexity affect the different taxonomic groups in different ways? Why is a multi-scale conservation approach essential?

8) What are some examples of conservation measures that may improve overall biodiversity and species richness in agricultural areas?

References

Chaplin-Kramer, R., O'Rourke, M. E., Blitzer, E. J., & Kremen, C. (2011). A meta-analysis of crop pest and natural enemy response to landscape complexity. Ecology letters, 14(9), 922-932.

Gonthier, D. J., Ennis, K. K., Farinas, S., Hsieh, H. Y., Iverson, A. L., Batáry, P., ... & Perfecto, I. (2014). Biodiversity conservation in agriculture requires a multi-scale approach. Proceedings of the Royal Society B: Biological Sciences, 281(1791), 20141358.

Gonthier, David J. et al. (2014), Data from: Biodiversity conservation in agriculture requires a multi-scale approach, Dryad, Dataset, https://doi.org/10.5061/dryad.mm6f4

Lael. (2014). Ecological complexity. The University of British Columbia. Retrieved from https://complexity.ok.ubc.ca/about/ecological-complexity/

Mikolajewicz, N., & Komarova, S. V. (2019). Meta-analytic methodology for basic research: a practical guide. Frontiers in physiology, 10, 203.

Monmany, A. C., Yu, M., Restrepo, C., & Zimmerman, J. K. (2015). How are landscape complexity and vegetation structure related across an agricultural frontier in the subtropical Chaco, NW Argentina?. Journal of Arid Environments, 123, 12-20.

Office of the Press Secretary (2014). Fact sheet: the economic challenge posed by declining pollinator populations. The Obama White House Briefings, Washington, DC, June, 20.

Pustejovsky, J. E. (2019). Log response ratio effect sizes. Lecture, Austin, Texas; University of Texas. https://www.jepusto.com/files/ABAI-2019-Log-response-ratios.pdf

Ramiadantsoa, T., Hanski, I., & Ovaskainen, O. (2018). Responses of generalist and specialist species to fragmented landscapes. Theoretical Population Biology, 124, 31-40.

Tschakert, P., Zimmerer, K., & King, B. (2016). Agricultural intensification. Penn State College of Earth and Mineral Sciences. Retrieved from https://www.e-education.psu.edu/geog30/node/139

United Nations. (n.d.). Conservation Agriculture. Food and Agriculture Organization . Retrieved from https://www.fao.org/conservation-agriculture/en/

United Nations (2020). Land Use in Agriculture by the Numbers. Retrieved from https://www.fao.org/sustainability/news/detail/en/c/1274219/

Zach. (2022). Fisher Z-transformation: Definition & Example. Statology. Retrieved from https://www.statology.org/fisher-z-transformation/