Group 6: Land-use Impacts on Biodiversity along FW-Coastal Gradients

Marc Los Huertos

2025-04-30

Introduction

The **Shannon-Weaver Index**, also known as the **Shannon-Wiener Index** or simply the **Shannon Diversity Index**, is a widely used metric in ecology for measuring biodiversity. It incorporates both species richness and evenness into a single value.

Historical Background

- **1948**: Claude Shannon introduced the concept of *entropy* in his paper *"A Mathematical Theory of Communication"*, where he quantified the uncertainty in a message source using:

$$H = -\sum p_i \log_2 p_i$$

- **Warren Weaver** later helped interpret and popularize this theory, leading to the combined attribution as the **Shannon-Weaver Index**.

- In the **1950s–60s**, ecologists adapted the entropy formula to describe biodiversity, treating species as symbols and their relative abundances as probabilities.

- Sometimes incorrectly referred to as the **Shannon-Wiener Index** due to confusion with Norbert Wiener.

Implemented in R

```
group6.csv = "/home/mwl04747/RTricks/00_Project_Group_Demos/Group6_FakeData.csv"

# Read in the data
group6 = read.csv(group6.csv, header = TRUE)

# check read data
head(group6)
```

```
##              X Athecate.dinoflagellates Thecate.dinoflagellates
## 1    San Onofre                      120                      90
## 2   Aliso Creek                       60                      45
## 3   Aliso Beach                      100                      80
## 4   Laguna Lake                       20                      15
## 5 Bonelli Park                       15                      10
```

```
##   Silicoflagellates Diatoms Coccolithophores Centrarchidae Salmonidae
## 1               35     200             140             5         30
## 2               10     100              25            12         18
## 3               30     180             110             7         22
## 4                5     140               8            90         10
## 5                2     130               5            85          8
##   Cyprinidae Serranidae Sciaenidae Ecosystem.Type  pH Temperature..C.
## 1          2         60         80      Saltwater 8.2            16.7
## 2         30          5         10     Freshwater 7.4            16.0
## 3          5         55         70      Saltwater 7.5            15.0
## 4         75          1          0     Freshwater 7.1            14.0
## 5         65          0          0     Freshwater 7.1            16.6
##   Salinity..ppt. Dissolved.Oxygen..mgL. Turbidity..NTU.
## 1             33                    4.8              40
## 2              0                    6.7              35
## 3             15                    3.9              48
## 4              0                    5.0              45
## 5              0                    5.0              45
```

```r
names(group6)
```

```
##  [1] "X"                     "Athecate.dinoflagellates"
##  [3] "Thecate.dinoflagellates" "Silicoflagellates"
##  [5] "Diatoms"               "Coccolithophores"
##  [7] "Centrarchidae"         "Salmonidae"
##  [9] "Cyprinidae"            "Serranidae"
## [11] "Sciaenidae"            "Ecosystem.Type"
## [13] "pH"                    "Temperature..C."
## [15] "Salinity..ppt."        "Dissolved.Oxygen..mgL."
## [17] "Turbidity..NTU."
```

Calculating Shannon-diversity index

Definition

The Shannon Diversity Index (also known as the Shannon-Wiener or Shannon-Weaver Index) is a measure of biodiversity that accounts for both species richness and species evenness. It is defined mathematically as:

$$H' = -\sum_{i=1}^{S} p_i \ln p_i$$

Where:

- H' is the Shannon diversity index.

- S is the total number of species.

- p_i is the proportion of individuals belonging to the i-th species.

- ln is the natural logarithm.

Calculation Steps

To calculate the index:

1. Count the total number of individuals across all species, denoted N.

2. For each species i, calculate the proportion $p_i = \frac{n_i}{N}$, where n_i is the number of individuals in species i.

3. Plug each p_i into the formula and compute the sum.

Alternate Formula Using Counts

Using species counts directly:

$$H' = -\sum_{i=1}^{S} \left(\frac{n_i}{N} \ln \frac{n_i}{N} \right)$$

Example Calculation in R

```
library(vegan)
```

```
## Loading required package: permute
```

```
## Loading required package: lattice
```

```
# Example data: rows = sites, columns = species counts
species_data <- matrix(
  c(5, 10, 15,
    8, 12, 10),
  nrow = 2,
  byrow = TRUE
)

# Calculate Shannon index
shannon_index <- diversity(species_data, index = "shannon")

# Print result
print(shannon_index)
```

```
## [1] 1.011404 1.085189
```

The `diversity()` function defaults to the Shannon index when `index = "shannon"`. It computes the index for each row (e.g., sample or site).

Example Calculation with Fake Data!

```
# Load the required library
library(vegan)
# Calculate the Shannon diversity index
shannon_diversity = function(data) {
  # Remove any rows with NA values
  data = na.omit(data)

  # Calculate the Shannon diversity index
  shannon_index = diversity(data, index = "shannon")
```

3

```
    return(shannon_index)

}
```

```
# Calculate the Shannon diversity index for phytoplankton and fish from group6 data
shannon_diversity_phytoplankton <- shannon_diversity(group6[, c("Athecate.dinoflagellates", "Thecate.di
```

```
##         1         2         3         4         5
## 1.2360911 1.1822503 1.2280767 0.7462206 0.6115810
```

```
shannon_diversity_fish <- shannon_diversity(group6[, c("Coccolithophores", "Centrarchidae", "Salmonidae"
```

```
str(shannon_diversity_fish)
```

```
##  Named num [1:5] 1.344 1.651 1.414 1.039 0.961
##  - attr(*, "names")= chr [1:5] "1" "2" "3" "4" ...
```

```
fish = as.data.frame(shannon_diversity_fish); fish
```

```
##   shannon_diversity_fish
## 1              1.3440054
## 2              1.6509059
## 3              1.4143310
## 4              1.0385452
## 5              0.9609695
```

```
fish$Location = group6$X; fish
```

```
##   shannon_diversity_fish      Location
## 1              1.3440054    San Onofre
## 2              1.6509059   Aliso Creek
## 3              1.4143310   Aliso Beach
## 4              1.0385452   Laguna Lake
## 5              0.9609695  Bonelli Park
```

```
## Let's add the location to dataframe
```

```
shannon_diversity_phytoplankton$Location = group6$X; shannon_diversity_phytoplankton
```

```
## Warning in shannon_diversity_phytoplankton$Location = group6$X: Coercing LHS to
## a list
```

```
## $`1`
## [1] 1.236091
##
## $`2`
## [1] 1.18225
##
## $`3`
## [1] 1.228077
##
## $`4`
## [1] 0.7462206
##
## $`5`
## [1] 0.611581
##
## $Location
```

```
## [1] "San Onofre"   "Aliso Creek"  "Aliso Beach"  "Laguna Lake"  "Bonelli Park"
```

Combine with WQ data! super neat dataset – great job!!

Hypotheses

How do phytoplankton and fish diversity vary with sampling location?

Plots