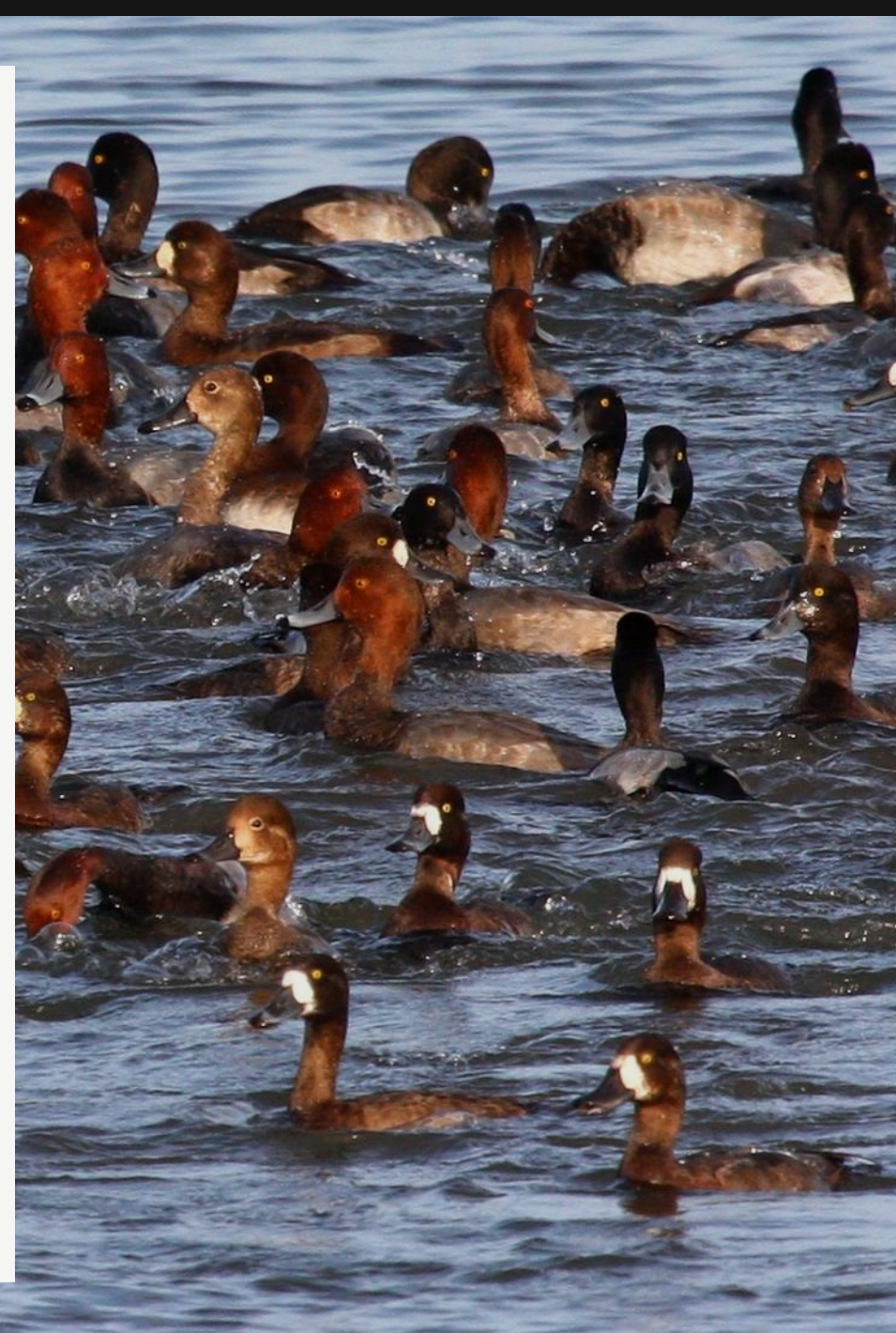


FW 599 Special Topics: Multivariate Analysis of Ecological Data in R

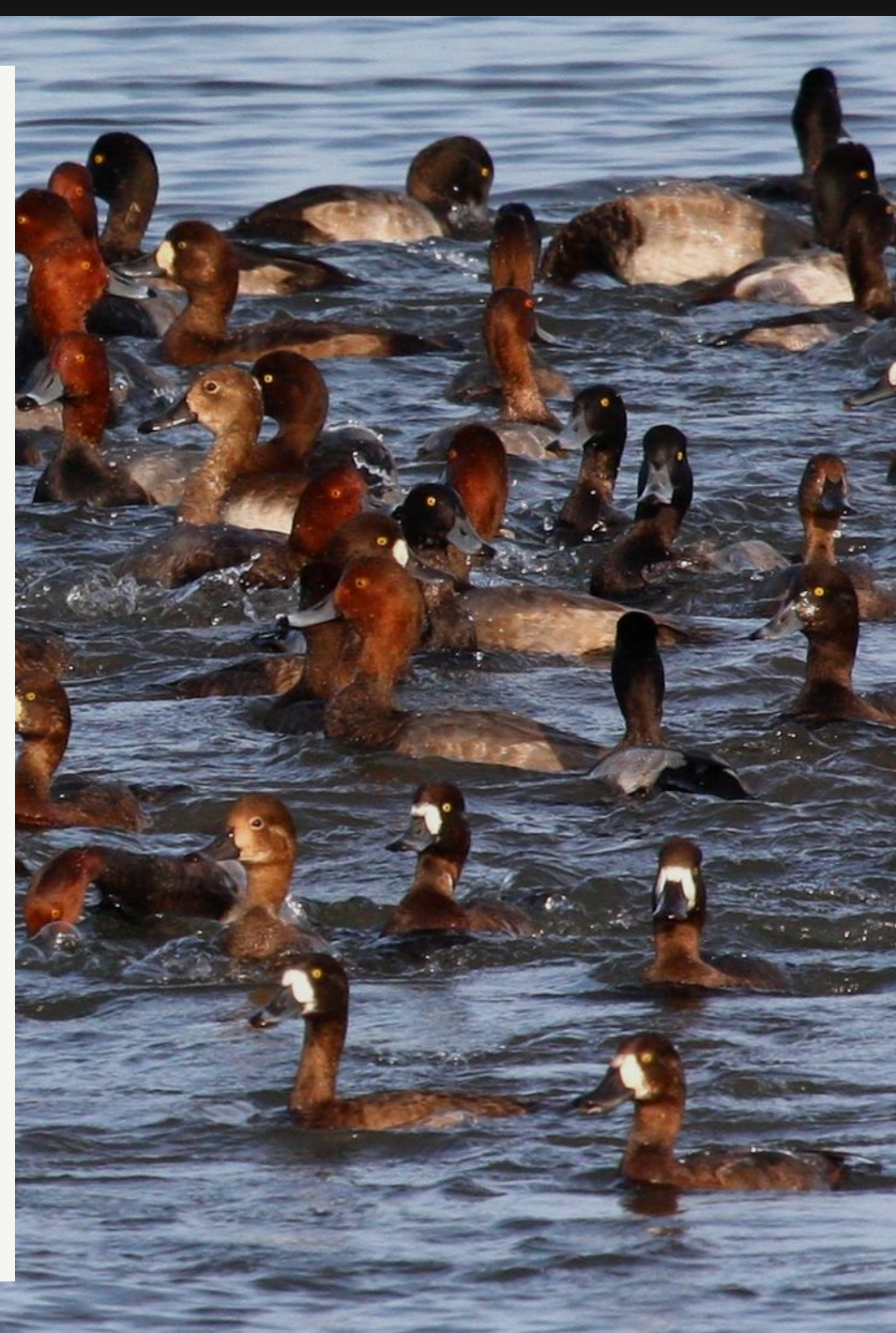
Lecture 7: Other Ordination Methods

Tuesday, October 22, 2024

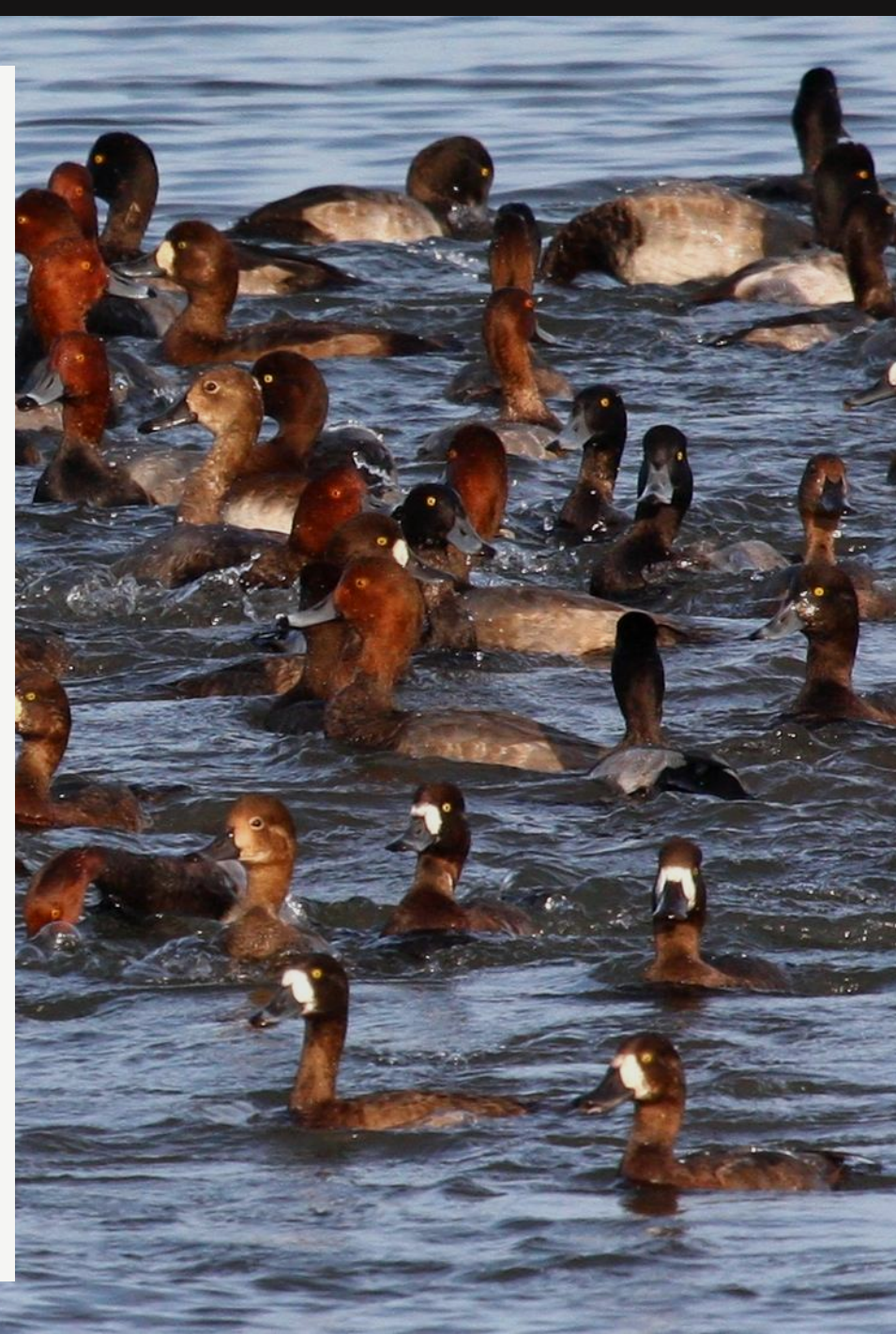


Lecture 7: Other Ordination Methods

- Correspondence Analysis (CA)
- Principal Coordinate Analysis (PCoA)
- Nonmetric Multidimensional Scaling (NMDS)



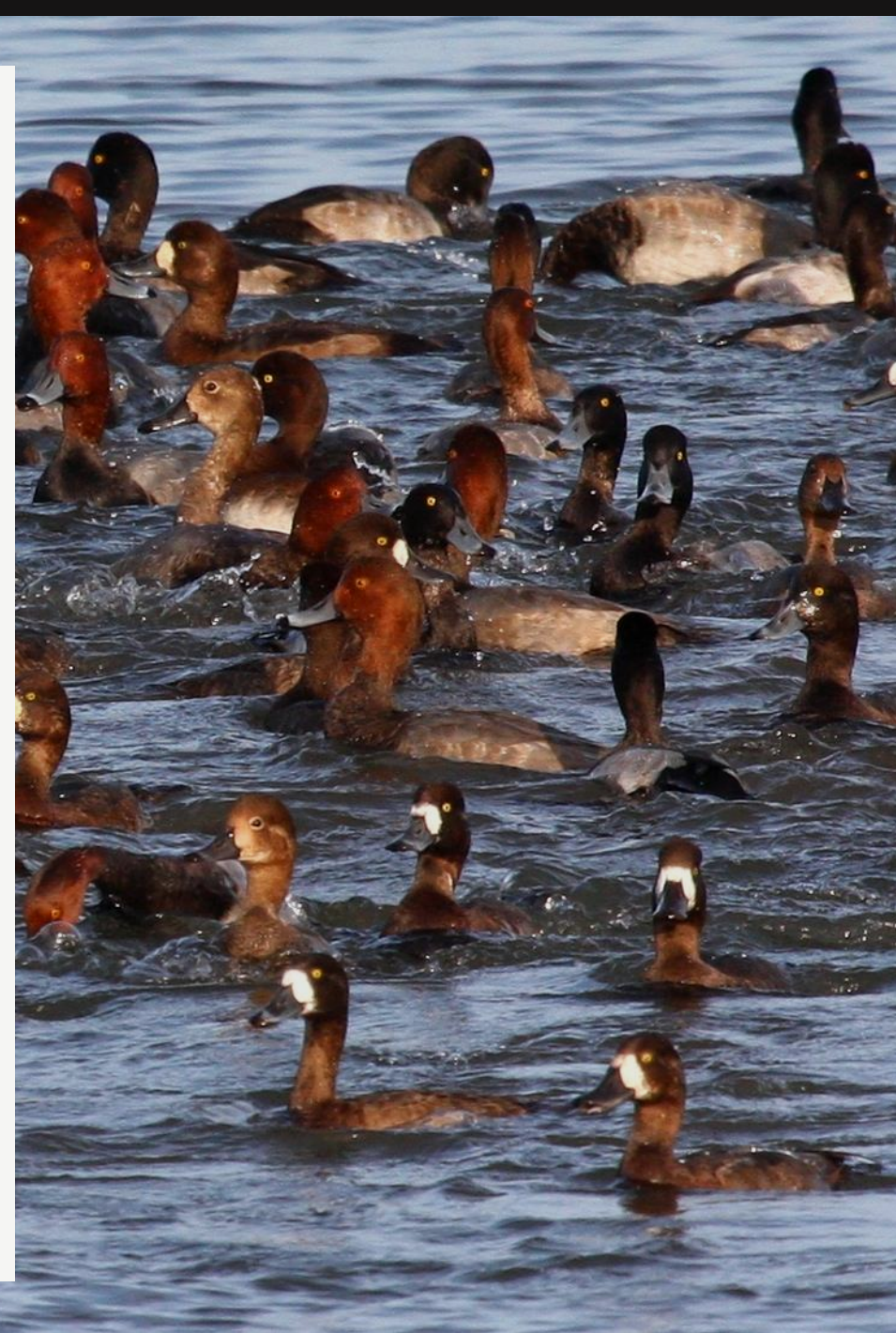
Recap: Principal Component Analysis



Recap: Principal Component Analysis

Principal Component Analysis uses eigenanalysis to reduce the dimensionality of large, ecological datasets while retaining as much information as possible.

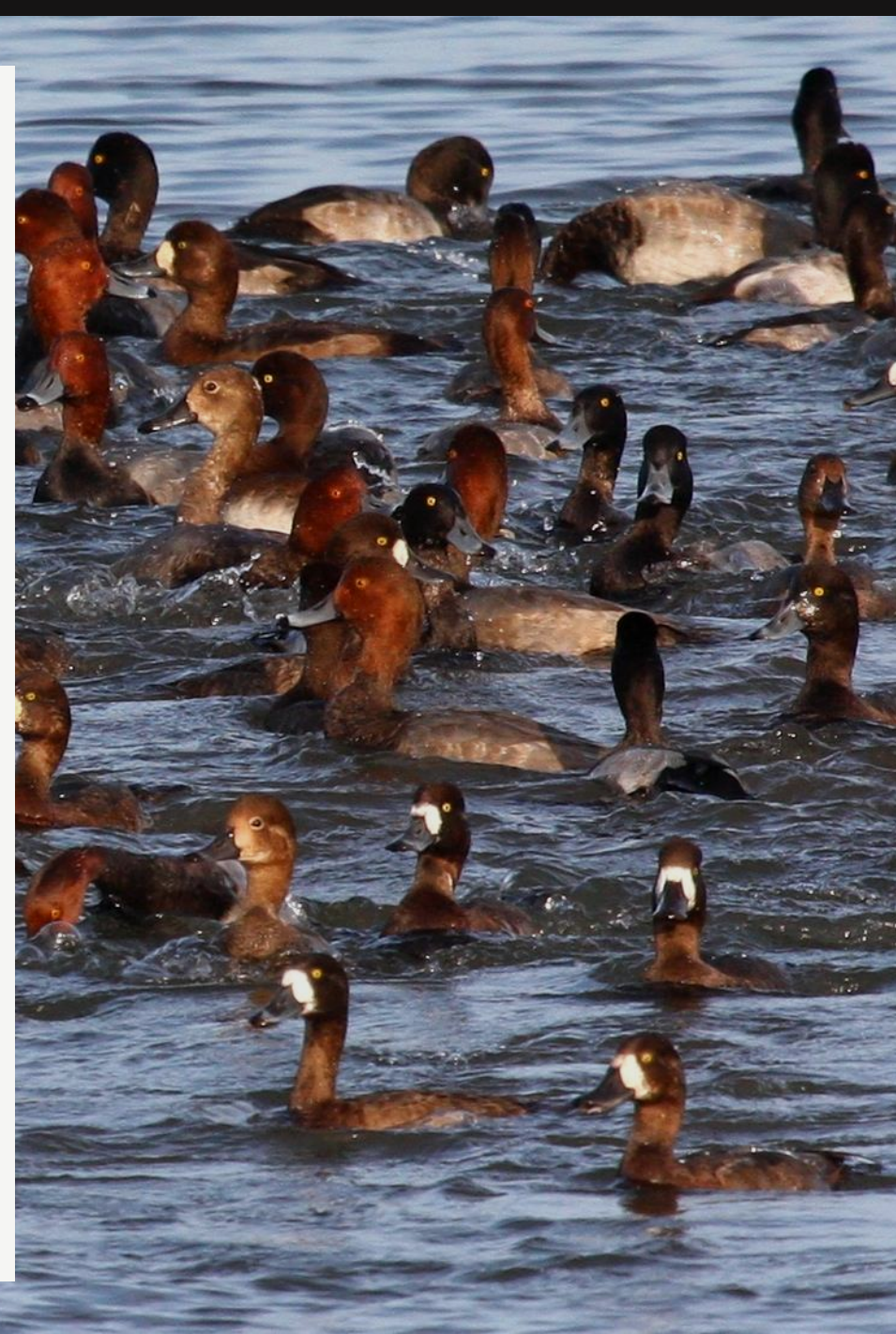
- Re-projects data in multidimensional space
- Maximizes the variance explained by the first principal axes (eigenvectors)



Recap: Principal Component Analysis

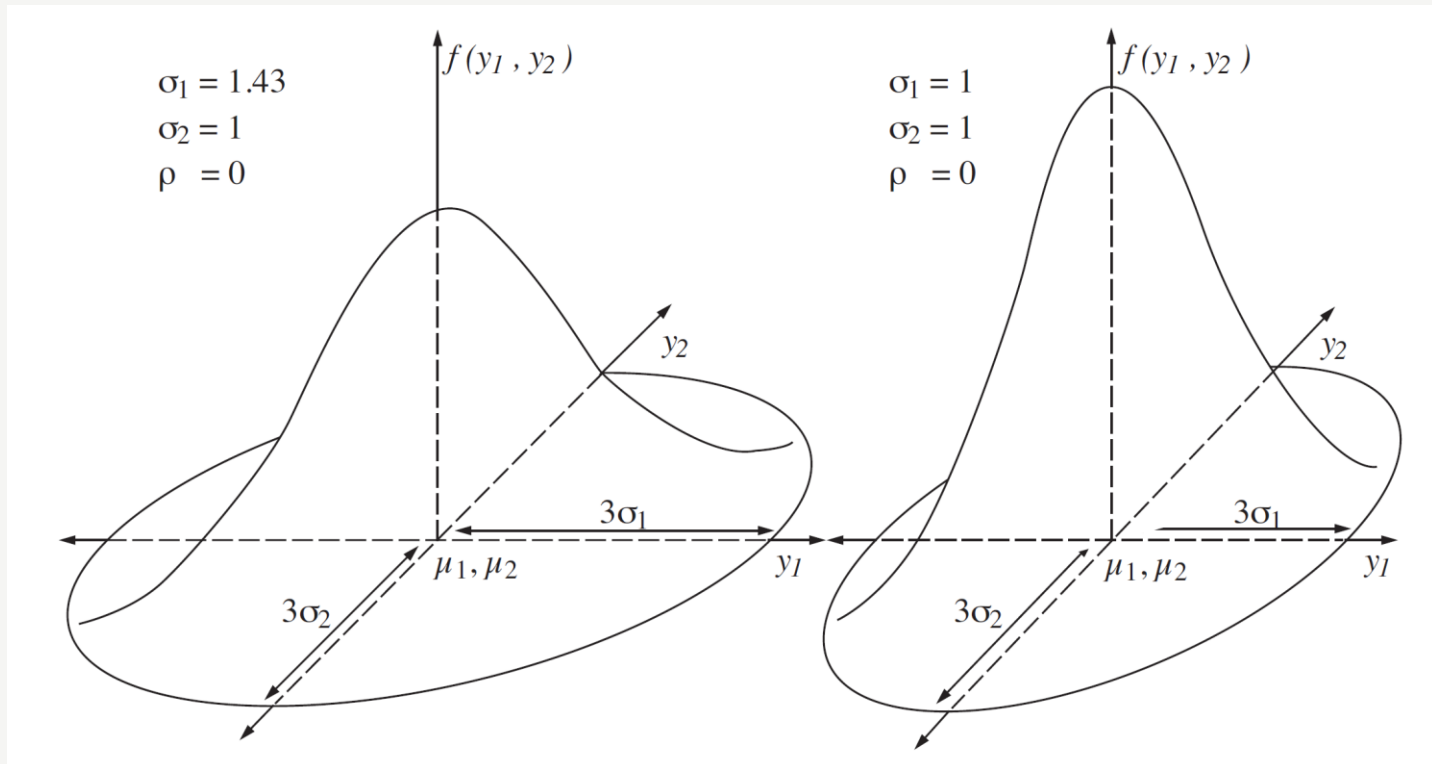
Many methods of multivariate analysis, including PCA, perform better when the response data distributions are **multivariate normal**. *Why?*

1. PCA Assumes the relationships between variables are linear
2. PCA depends on aligning the principal components with the directions of maximum variability
3. Interpretation is influenced by non-linear relationships, skewness, and outliers

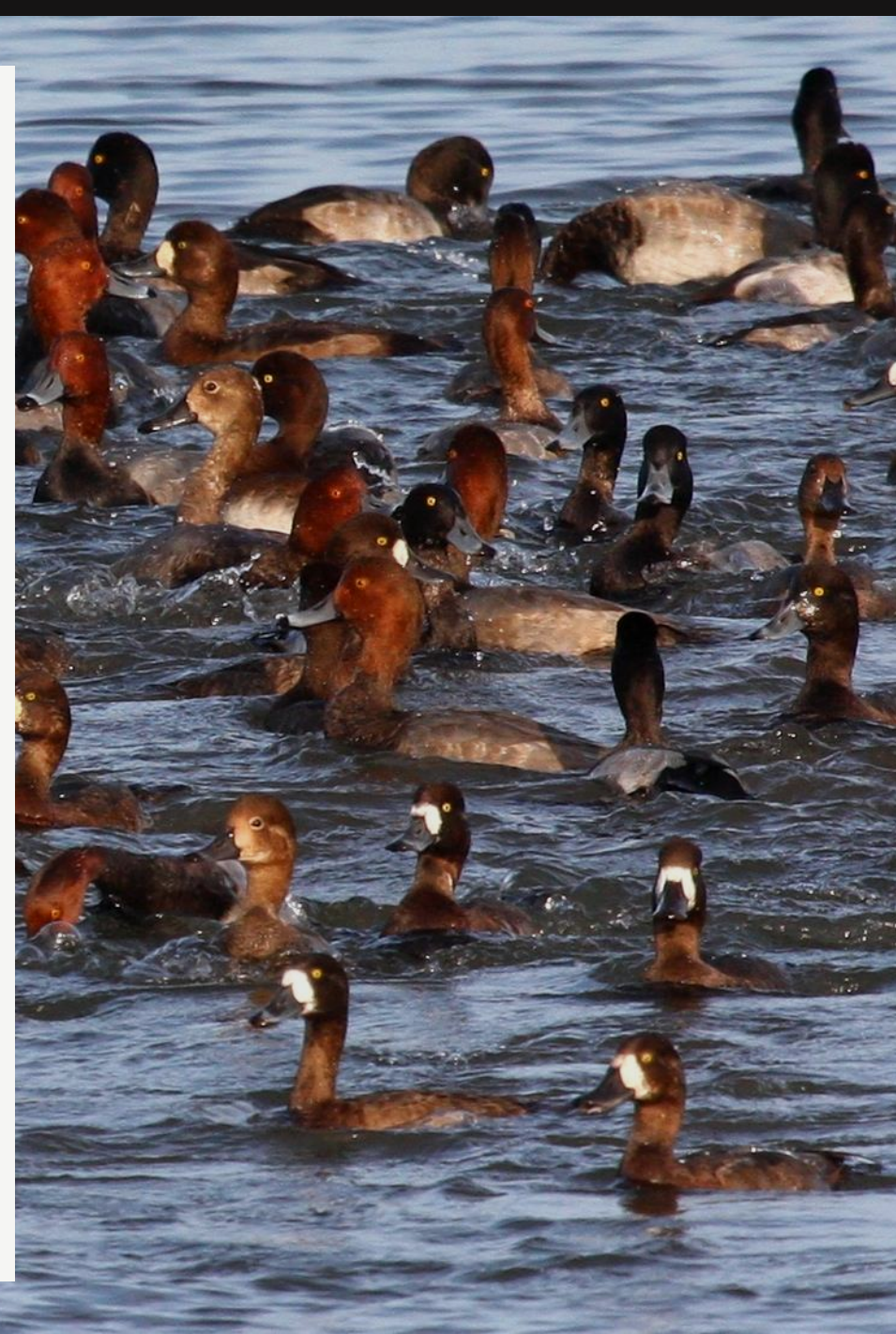


Recap: Principal Component Analysis

Multivariate normal distribution: Requires mean (μ), standard deviation (σ), and correlation (ρ).

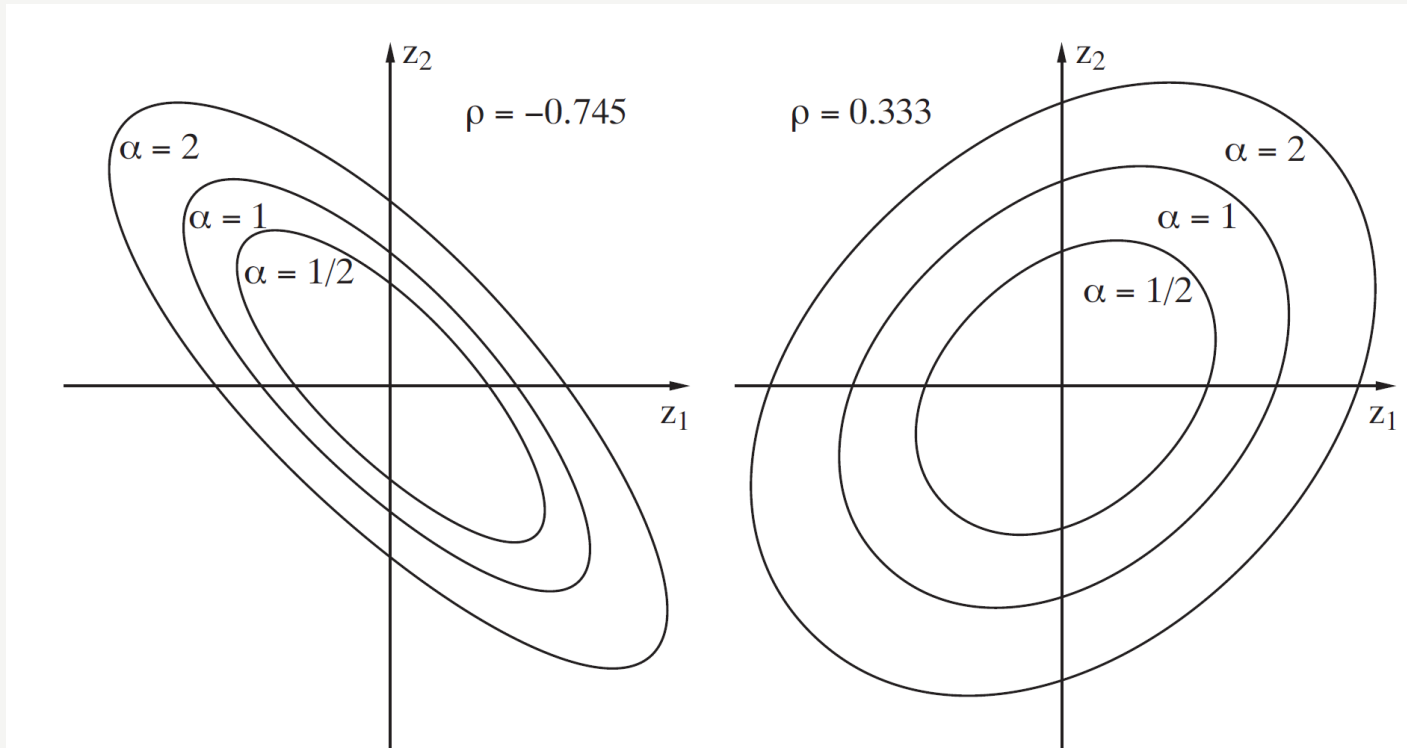


Legendre & Legendre Fig. 4.6

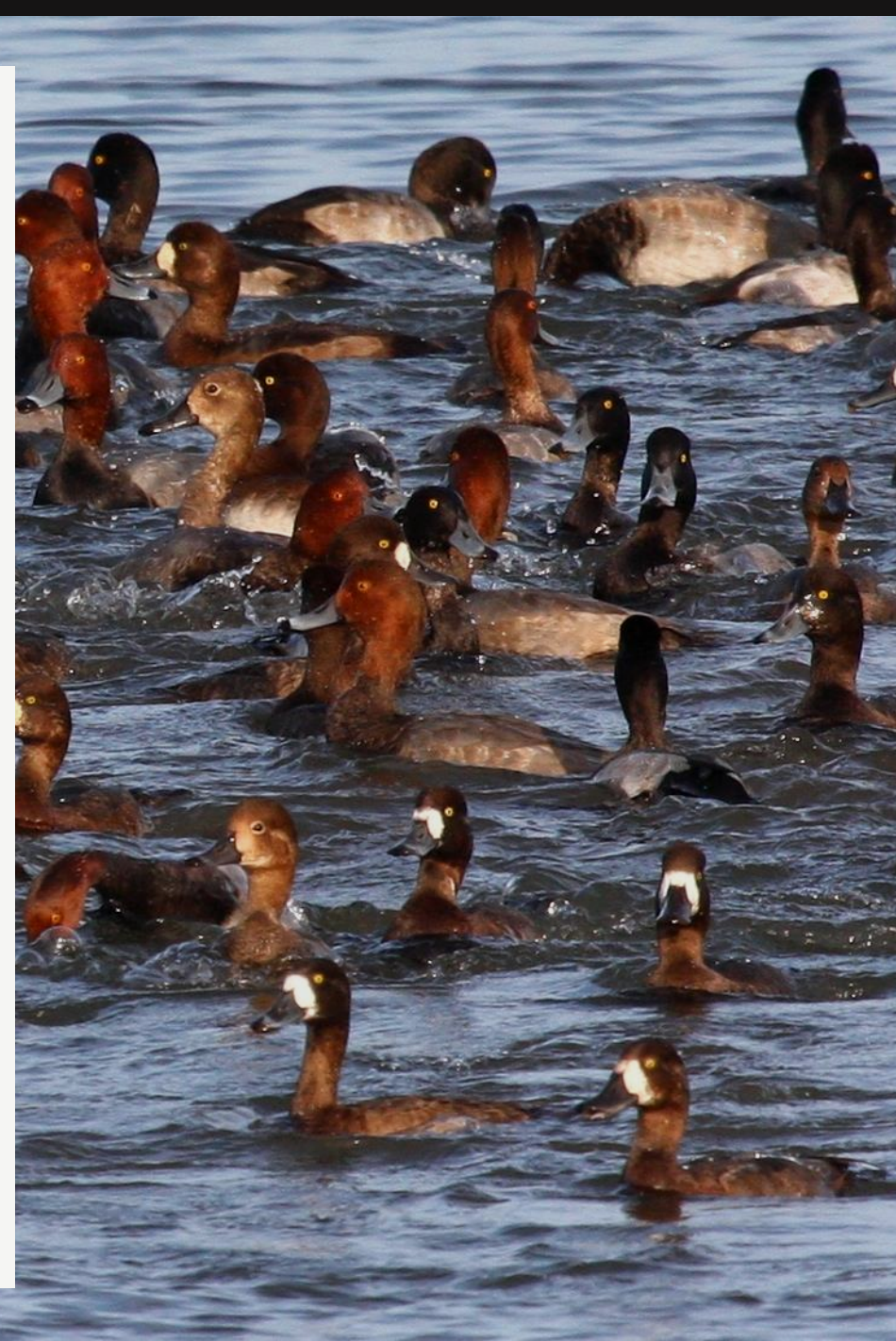


Recap: Principal Component Analysis

Multivariate normal distribution: Requires mean (μ), standard deviation (σ), and correlation (ρ).



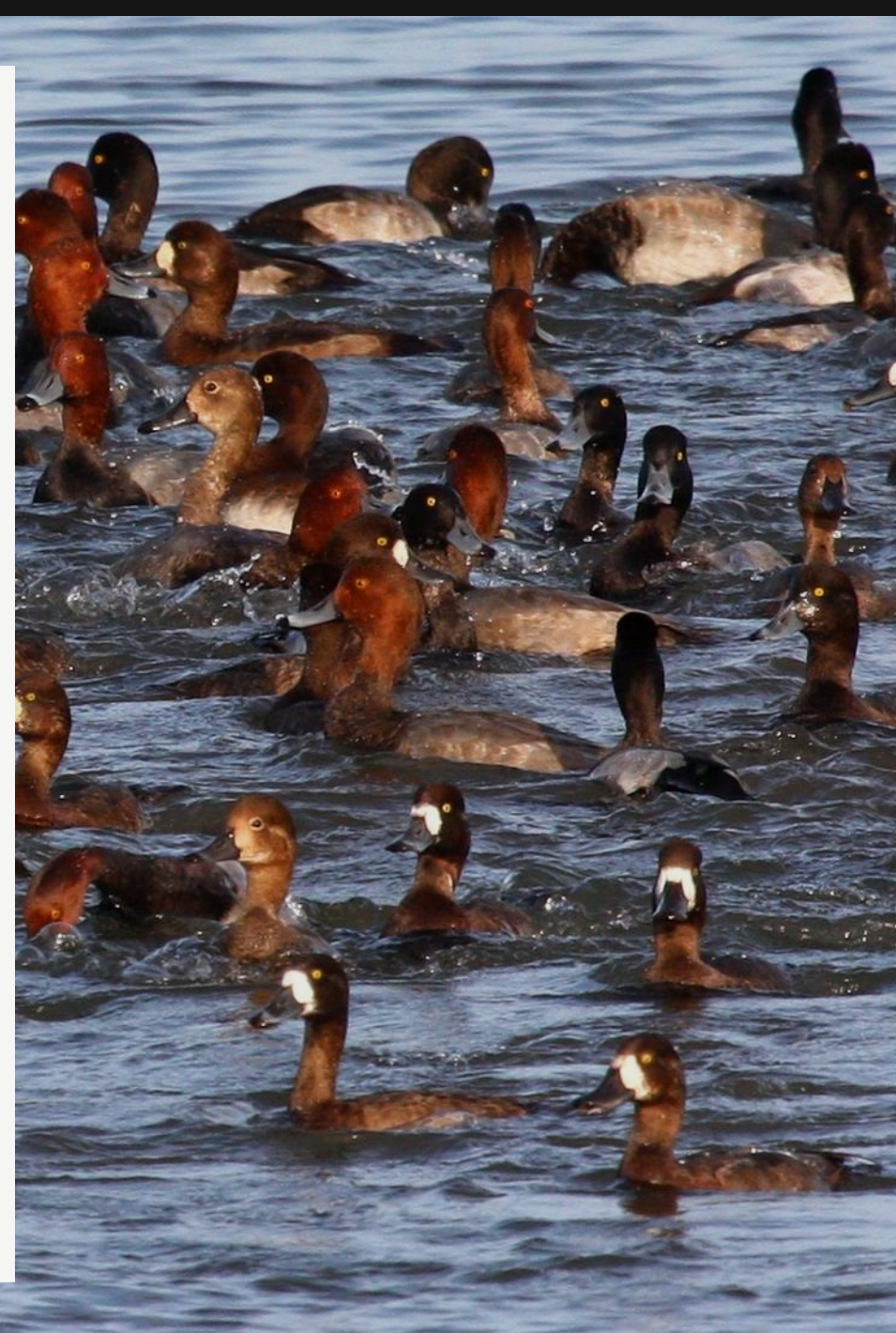
Legendre & Legendre Fig. 4.7



Recap: Principal Component Analysis

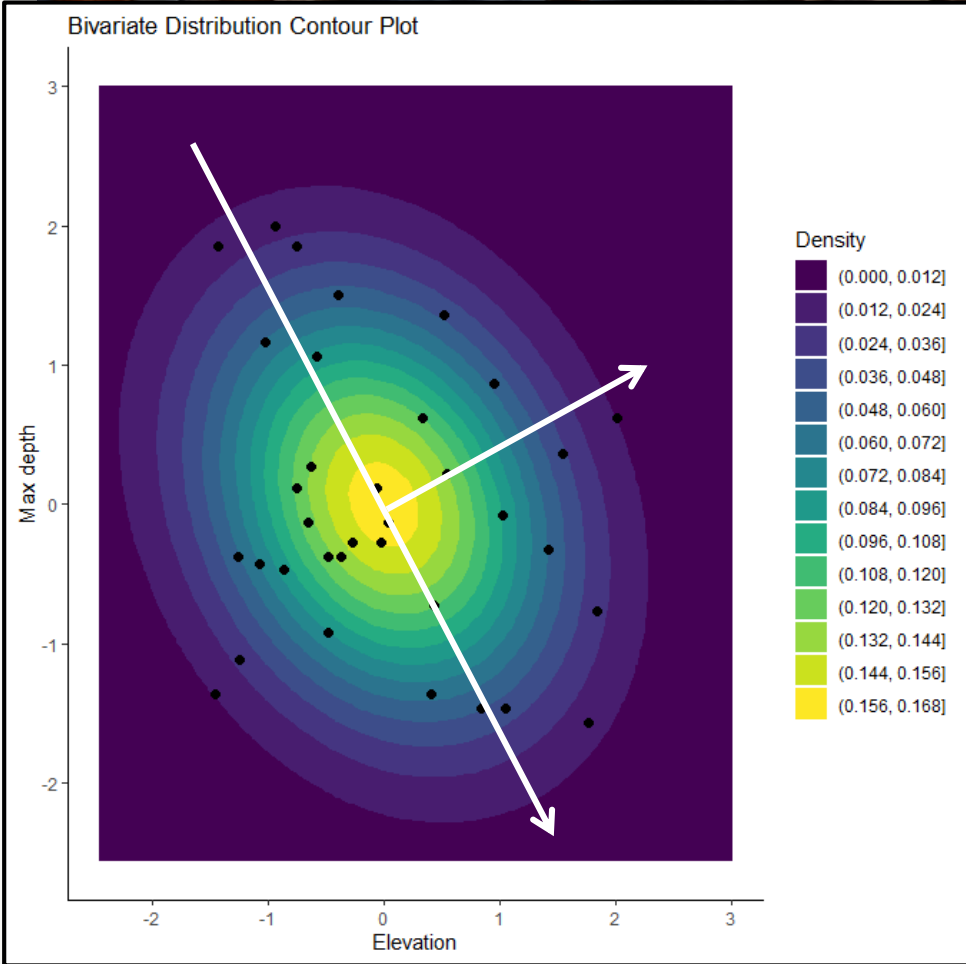
PCA depends on aligning the principal components (axes) with the **directions of maximum variability**.

- The first **principal axis** is the line that passes through the dimension of greatest variance of the ellipsoid.
- Each subsequent principal axis passes through dimensions of successively smaller variance.
- All axes are perpendicular to one another in hyperspace.



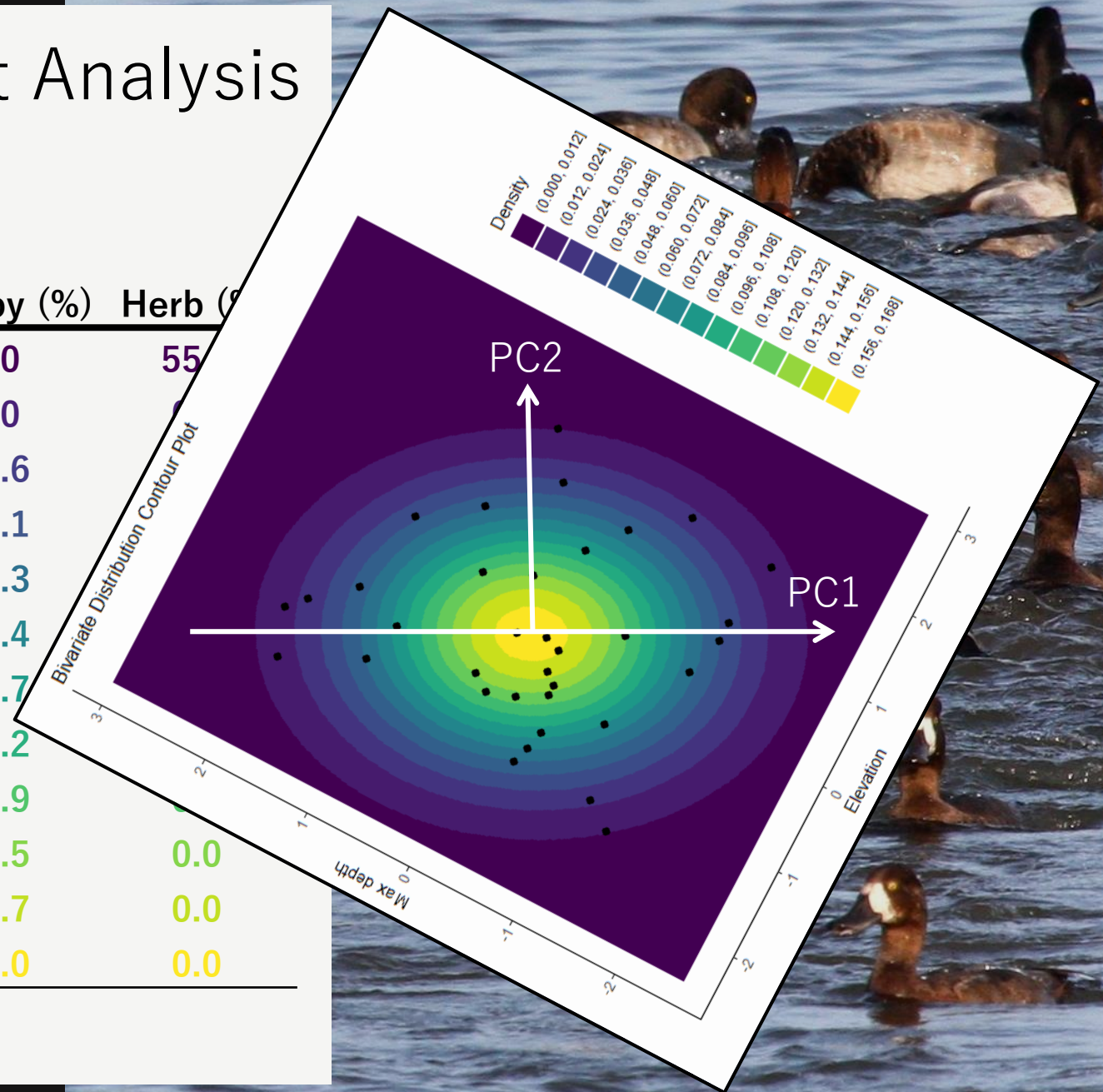
Recap: Principal Component Analysis

Site ID	Max Depth (m)	Gradient (%)	Elevation (m)	Canopy (%)	Herb (%)
Silvies-11	0.45	0.3	1439	0.0	55.1
Silvies-34	0.78	1.1	1487	0.0	0.0
Silvies-02	0.71	0.4	1372	29.6	0.0
Silvies-15	0.40	0.2	1471	41.1	0.0
Silvies-07	0.50	1.3	1547	52.3	0.0
Silvies-08	0.40	0.6	1492	51.4	0.0
Silvies-22	0.42	0.9	1555	54.7	0.0
Silvies-18	0.42	0.5	1510	46.2	0.0
Silvies-12	0.52	3.2	1658	51.9	0.0
Silvies-21	0.18	2.4	1713	37.5	0.0
Silvies-05	0.45	5.5	1565	46.7	0.0
Silvies-03	0.20	3.3	1634	59.0	0.0



Recap: Principal Component Analysis

Site ID	Max Depth (m)	Gradient (%)	Elevation (m)	Canopy (%)	Herb (%)
Silvies-11	0.45	0.3	1439	0.0	55.0
Silvies-34	0.78	1.1	1487	0.0	0.0
Silvies-02	0.71	0.4	1372	29.6	0.0
Silvies-15	0.40	0.2	1471	41.1	0.0
Silvies-07	0.50	1.3	1547	52.3	0.0
Silvies-08	0.40	0.6	1492	51.4	0.0
Silvies-22	0.42	0.9	1555	54.7	0.0
Silvies-18	0.42	0.5	1510	46.2	0.0
Silvies-12	0.52	3.2	1658	51.9	0.0
Silvies-21	0.18	2.4	1713	37.5	0.0
Silvies-05	0.45	5.5	1565	46.7	0.0
Silvies-03	0.20	3.3	1634	59.0	0.0



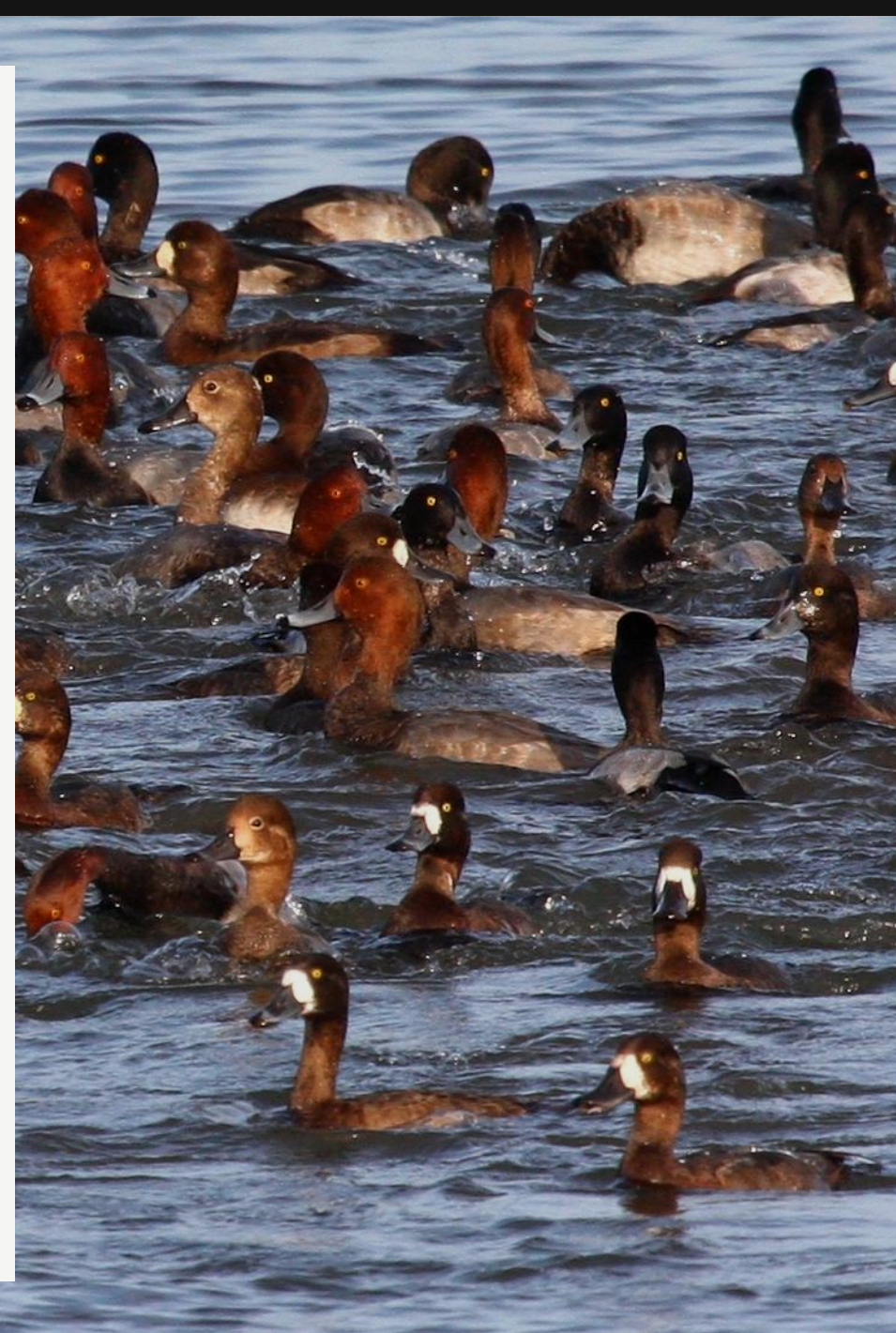
Recap: Limitations

PCA: To use or not to use?

- Optimal use calls for normalization of the data
- If the number of objects is smaller than the number of descriptors ($n < p$), negative eigenvalues will occur
- PCA is not useful for R-mode analysis
- PCA cannot incorporate multi-state descriptors
- Watch out for the double zero problem!



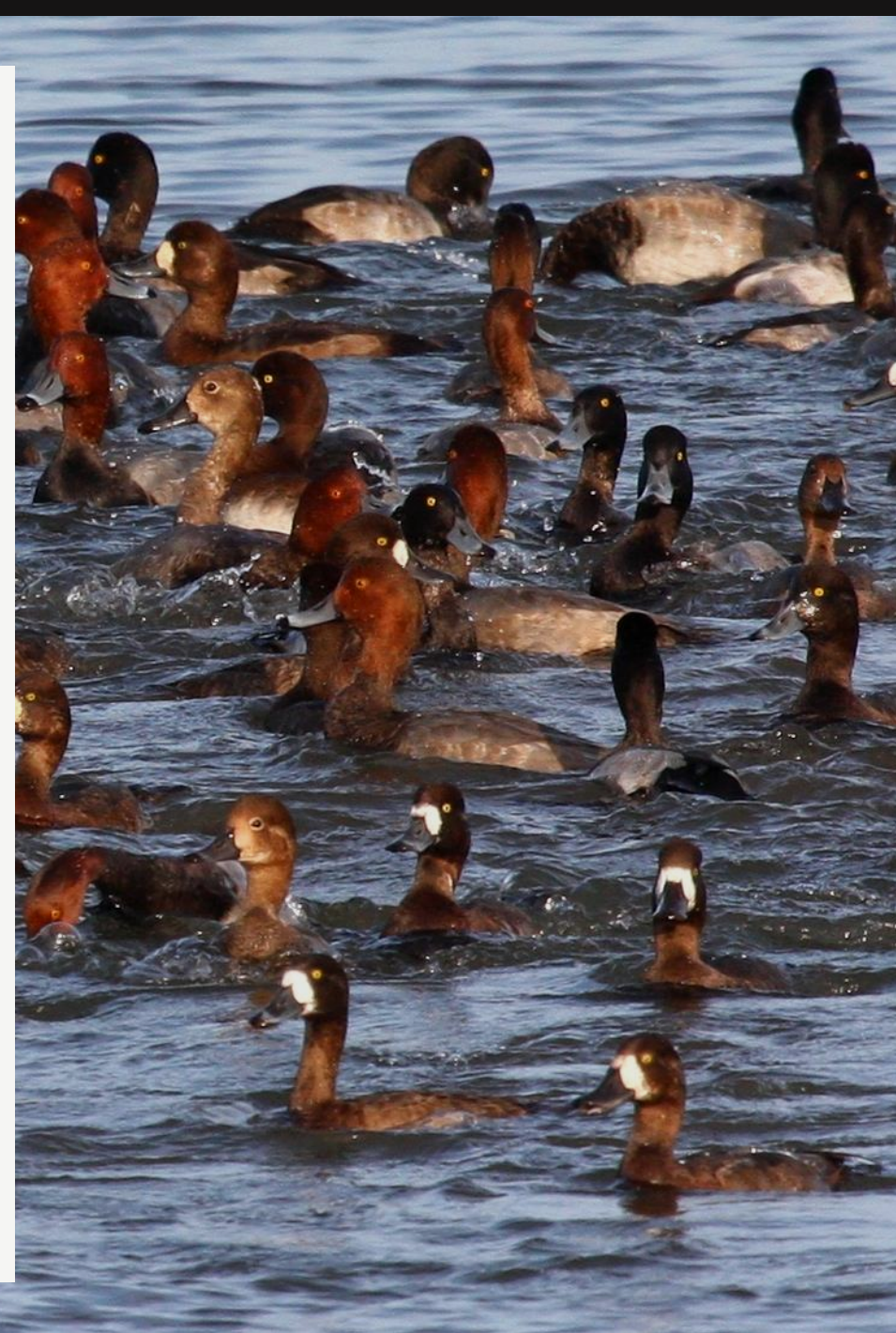
Correspondence Analysis (CA)



Correspondence Analysis (CA): Introduction

Species abundance and other kinds of count data are often:

- Non-linear
- Not multinormally distributed
- Highly zero-skewed

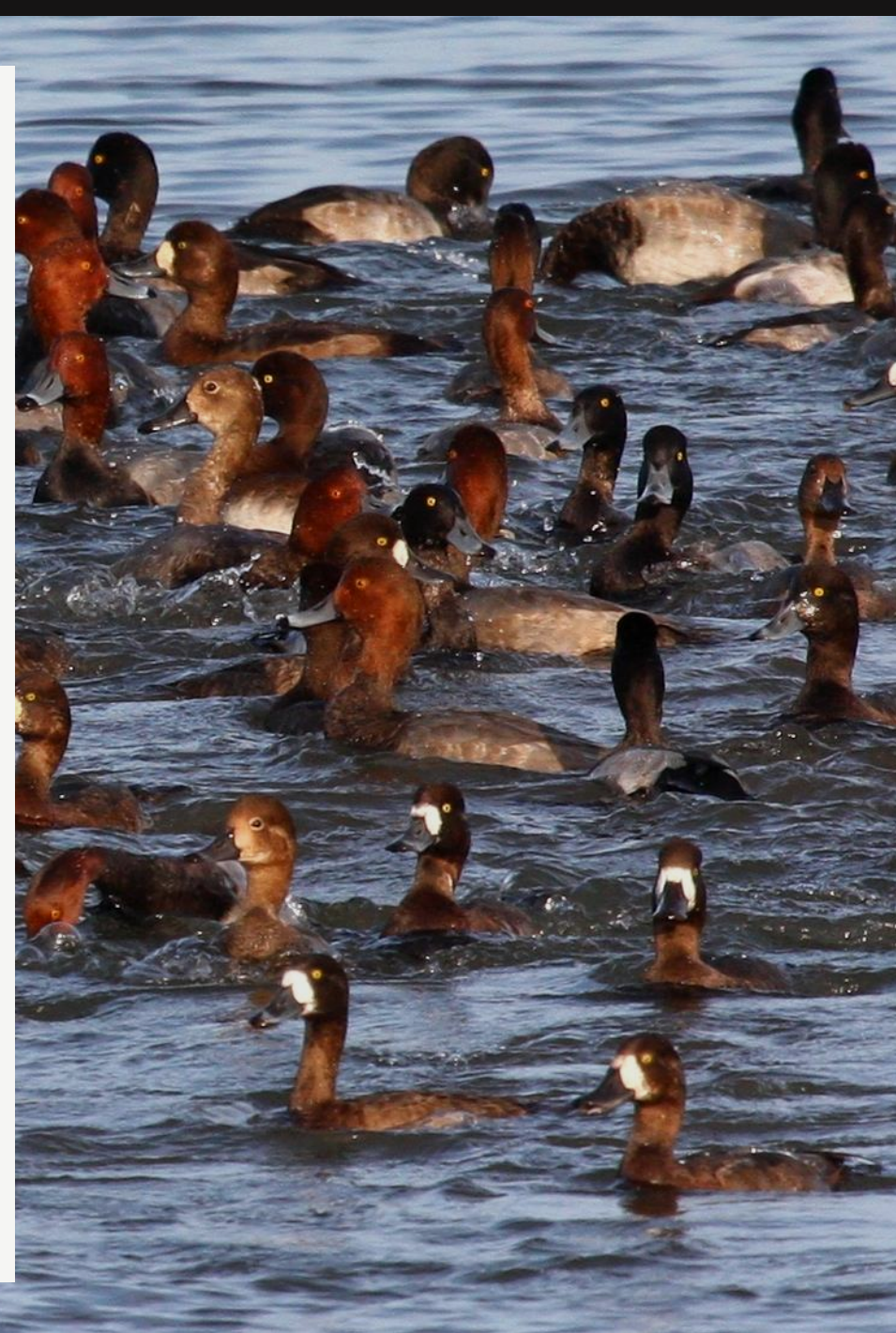


Correspondence Analysis (CA): Introduction

Species abundance and other kinds of count data are often:

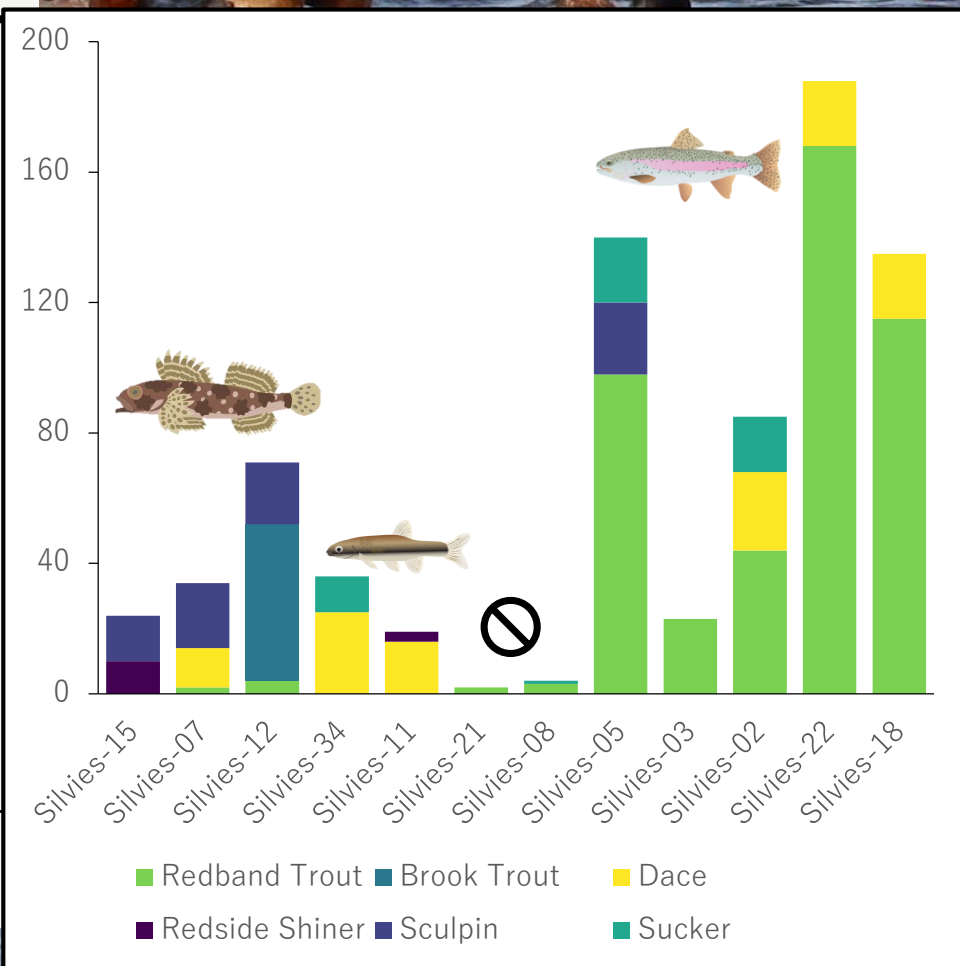
- Non-linear
- Not multinormally distributed
- Highly zero-skewed

Is there a better way?



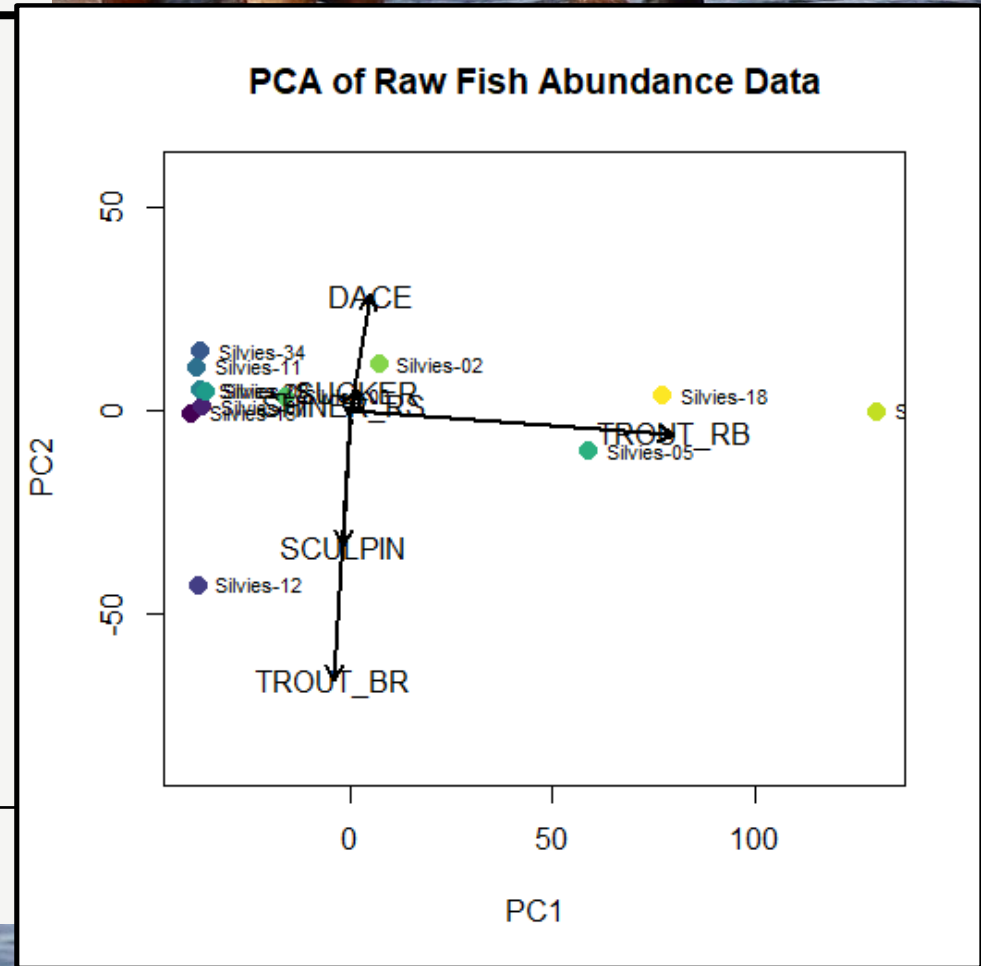
Correspondence Analysis (CA)

Site ID	Redband Trout	Brook Trout	Dace	Redside Shiner	Sculpin	Sucker
Silvies-15	0	0	0	10	14	0
Silvies-07	2	0	12	0	20	0
Silvies-12	4	48	0	0	19	0
Silvies-34	0	0	25	0	0	11
Silvies-11	0	0	16	3	0	0
Silvies-21	2	0	0	0	0	0
Silvies-08	3	0	0	0	0	1
Silvies-05	98	0	0	0	22	20
Silvies-03	23	0	0	0	0	0
Silvies-02	44	0	24	0	0	17
Silvies-22	168	0	20	0	0	0
Silvies-18	115	0	20	0	0	0



Correspondence Analysis (CA)

Site ID	Redband Trout	Brook Trout	Dace	Redside Shiner	Sculpin	Sucker
Silvies-15	0	0	0	10	14	0
Silvies-07	2	0	12	0	20	0
Silvies-12	4	48	0	0	19	0
Silvies-34	0	0	25	0	0	11
Silvies-11	0	0	16	3	0	0
Silvies-21	2	0	0	0	0	0
Silvies-08	3	0	0	0	0	1
Silvies-05	98	0	0	0	22	20
Silvies-03	23	0	0	0	0	0
Silvies-02	44	0	24	0	0	17
Silvies-22	168	0	20	0	0	0
Silvies-18	115	0	20	0	0	0

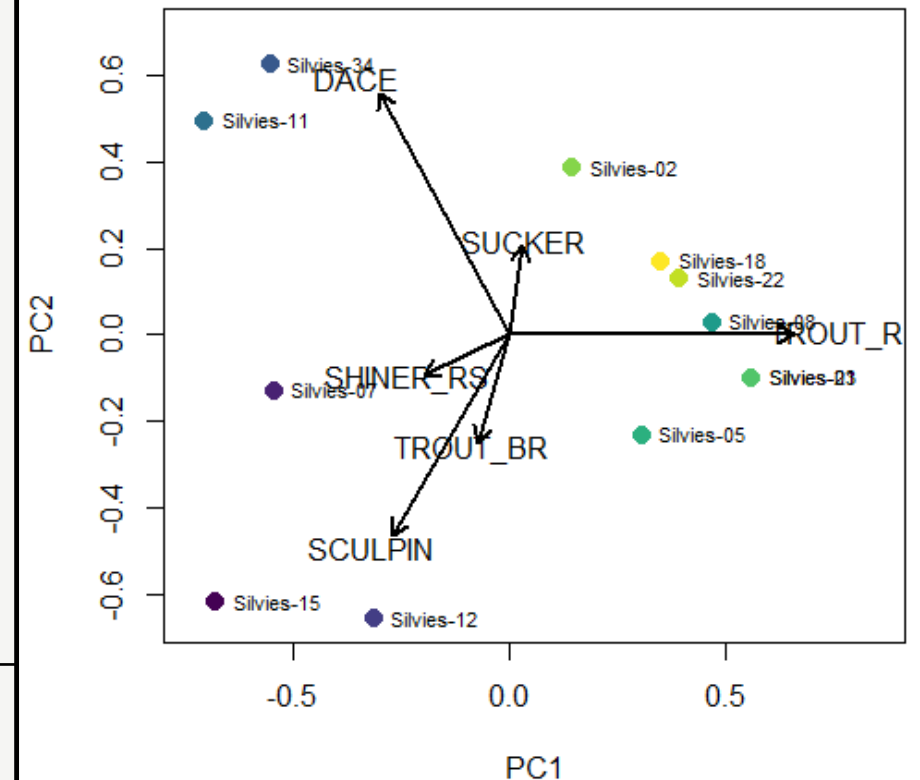


Correspondence Analysis (CA)

Site ID	Redband Trout	Brook Trout	Dace	Redside Shiner	Sculpin	Sucker
Silvies-15	0	0	0	10	14	0
Silvies-07	2	0	12	0	20	0
Silvies-12	4	48	0	0	19	0
Silvies-34	0	0	25	0	0	11
Silvies-11	0	0	16	3	0	0
Silvies-21	2	0	0	0	0	0
Silvies-08	3	0	0	0	0	1
Silvies-05	98	0	0	0	22	20
Silvies-03	23	0	0	0	0	0
Silvies-02	44	0	24	0	0	17
Silvies-22	168	0	20	0	0	0
Silvies-18	115	0	20	0	0	0

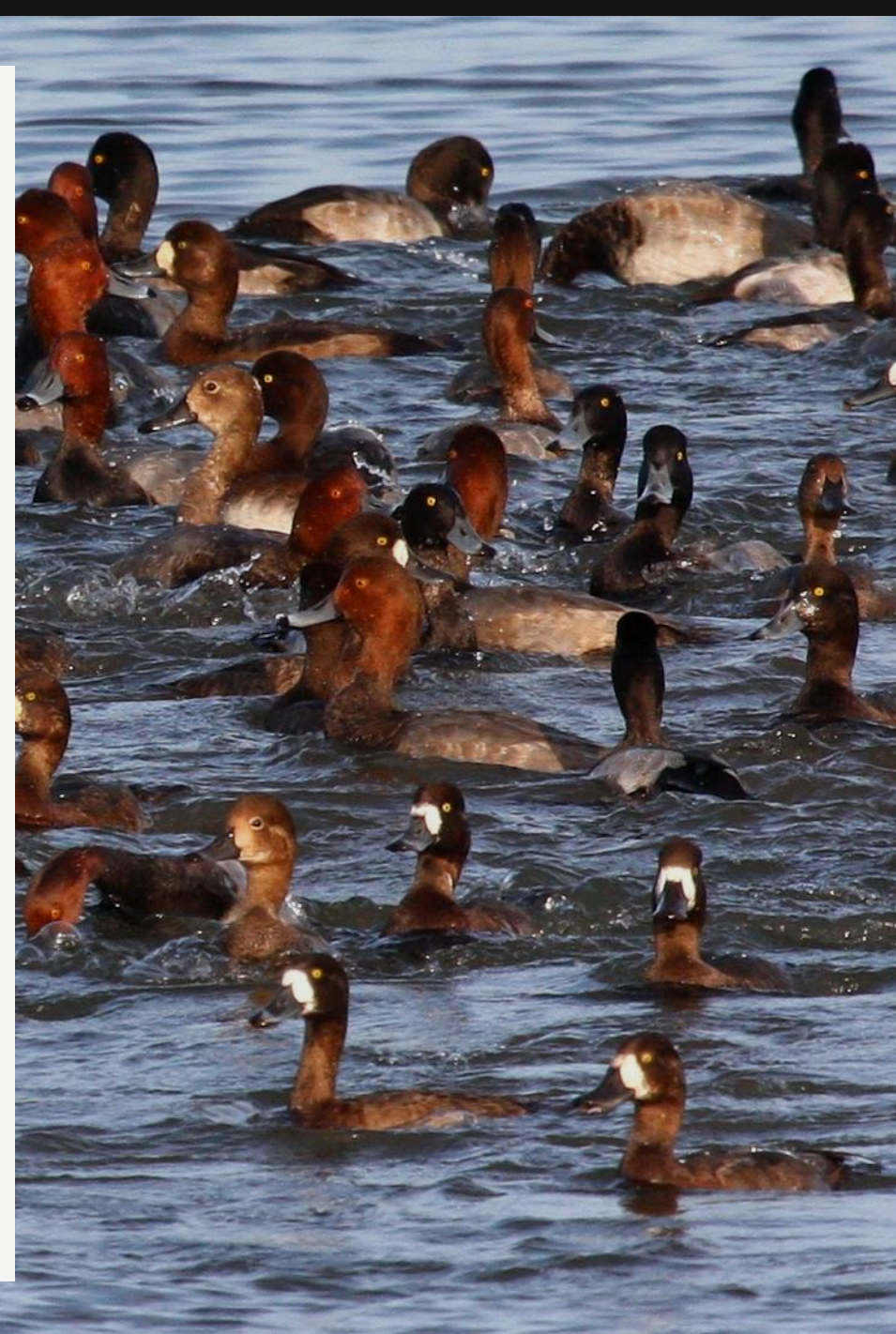


PCA of Hellinger Transformed Fish Abundance Data



Correspondence Analysis (CA): Introduction

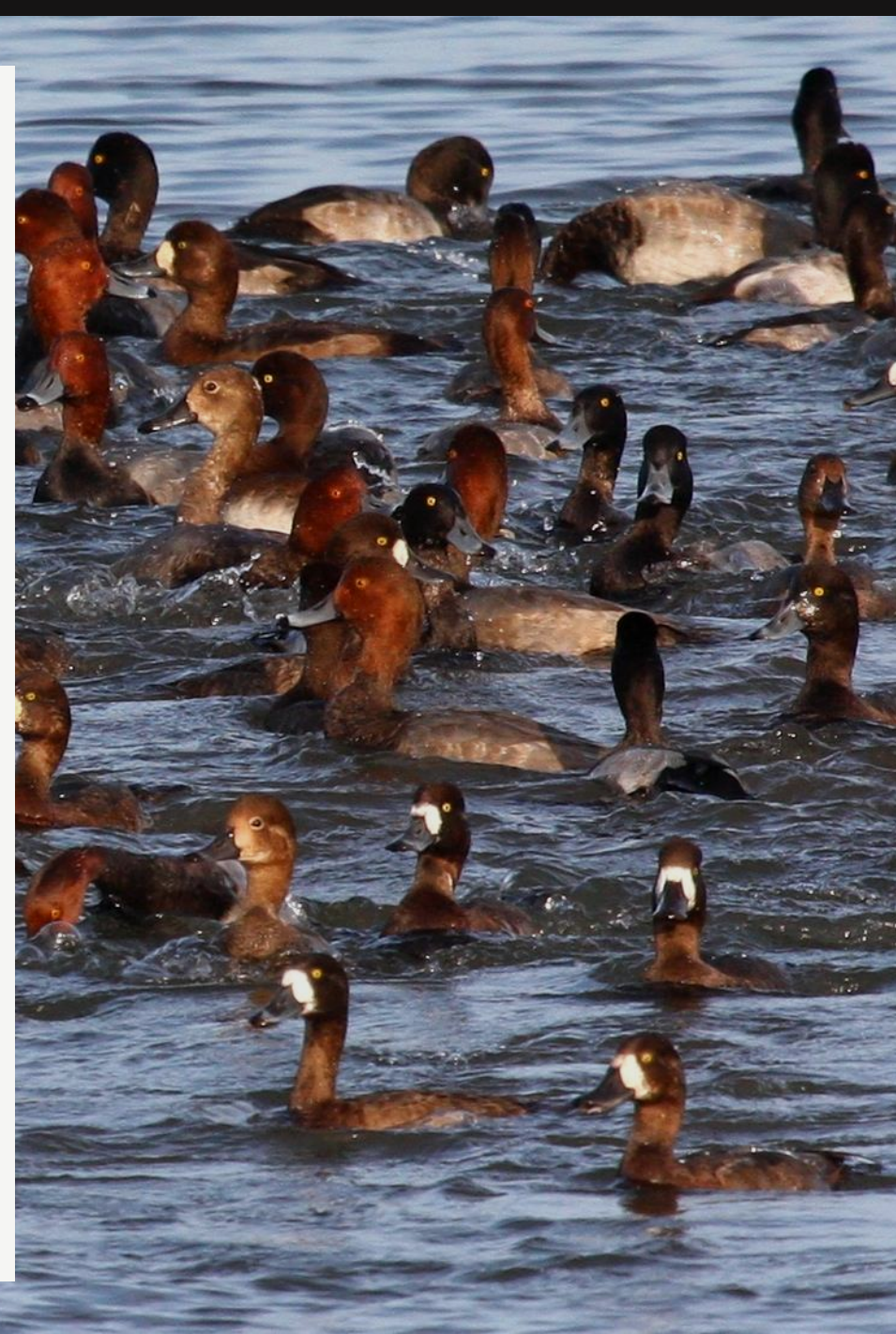
Correspondence Analysis is an alternative ordination method that preserves the chi-square distances (χ^2) among objects in the principal axes.



Correspondence Analysis (CA): Introduction

Correspondence Analysis is an alternative ordination method that preserves the chi-square distances (χ^2) among objects in the principal axes.

Designed for dimensionally homogeneous, non-negative data (such as counts).

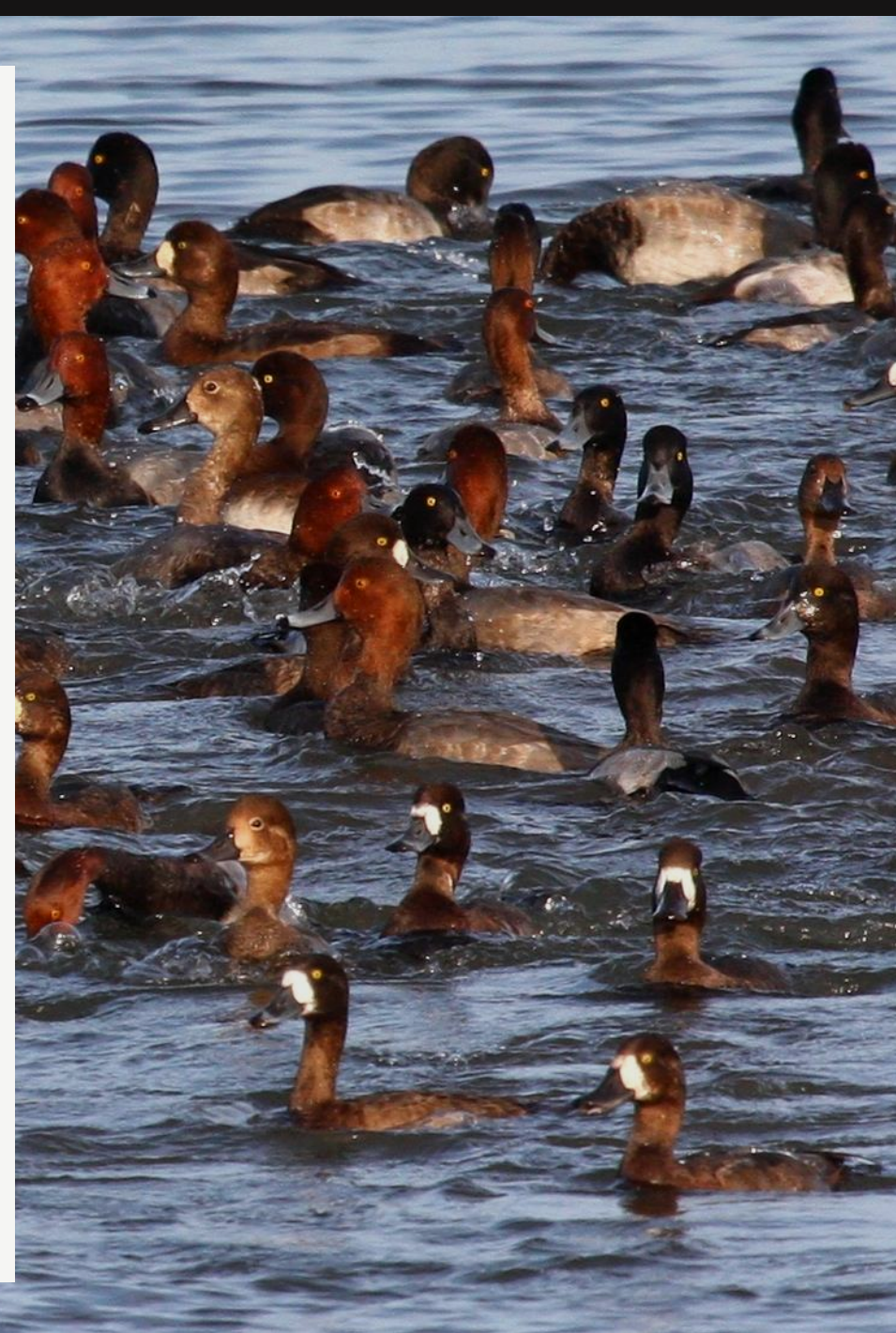


Correspondence Analysis (CA): Introduction

Correspondence Analysis is an alternative ordination method that preserves the chi-square distances (χ^2) among objects in the principal axes.

Designed for dimensionally homogeneous, non-negative data (such as counts).

Computed on a two-way contingency table instead of a dispersion matrix.



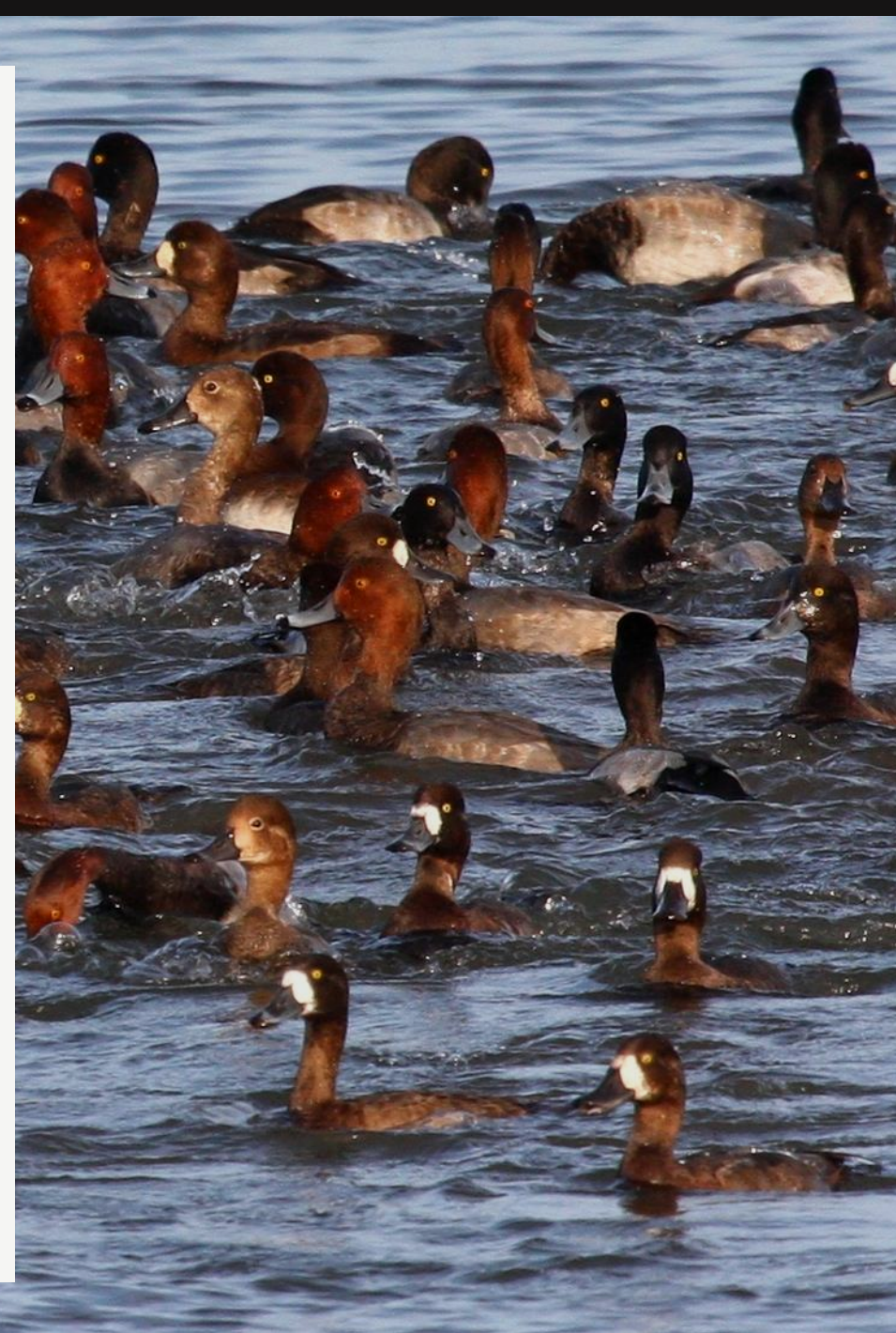
Correspondence Analysis (CA): Introduction

Correspondence Analysis is an alternative ordination method that preserves the chi-square distances (χ^2) among objects in the principal axes.

Designed for dimensionally homogeneous, non-negative data (such as counts).

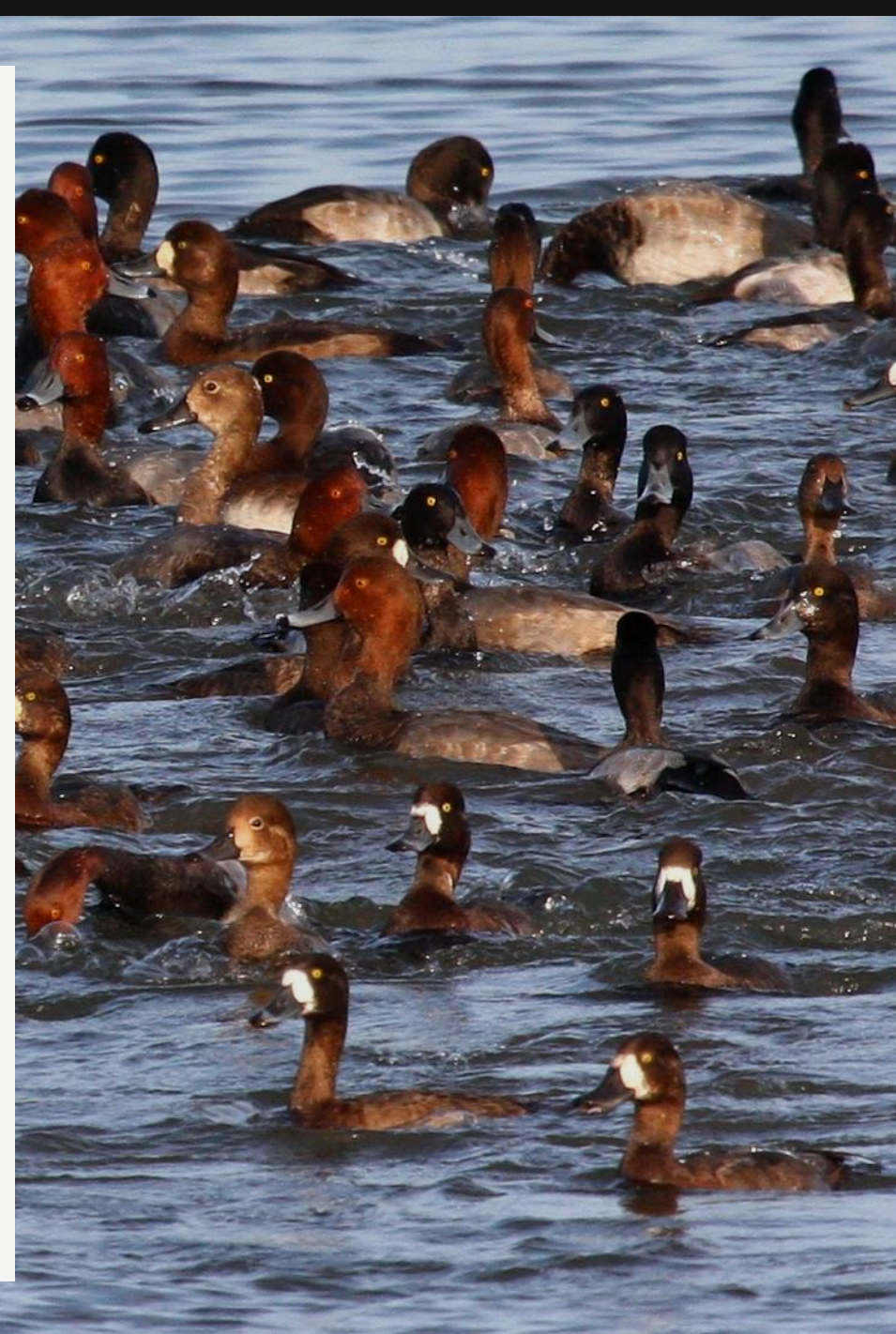
Computed on a two-way contingency table instead of a dispersion matrix.

Excludes double zeros!



Correspondence Analysis (CA): Steps

Step 1: Transform contingency table into a table of contributions (\bar{Q}) to the Pearson chi-square statistic (χ^2_P).

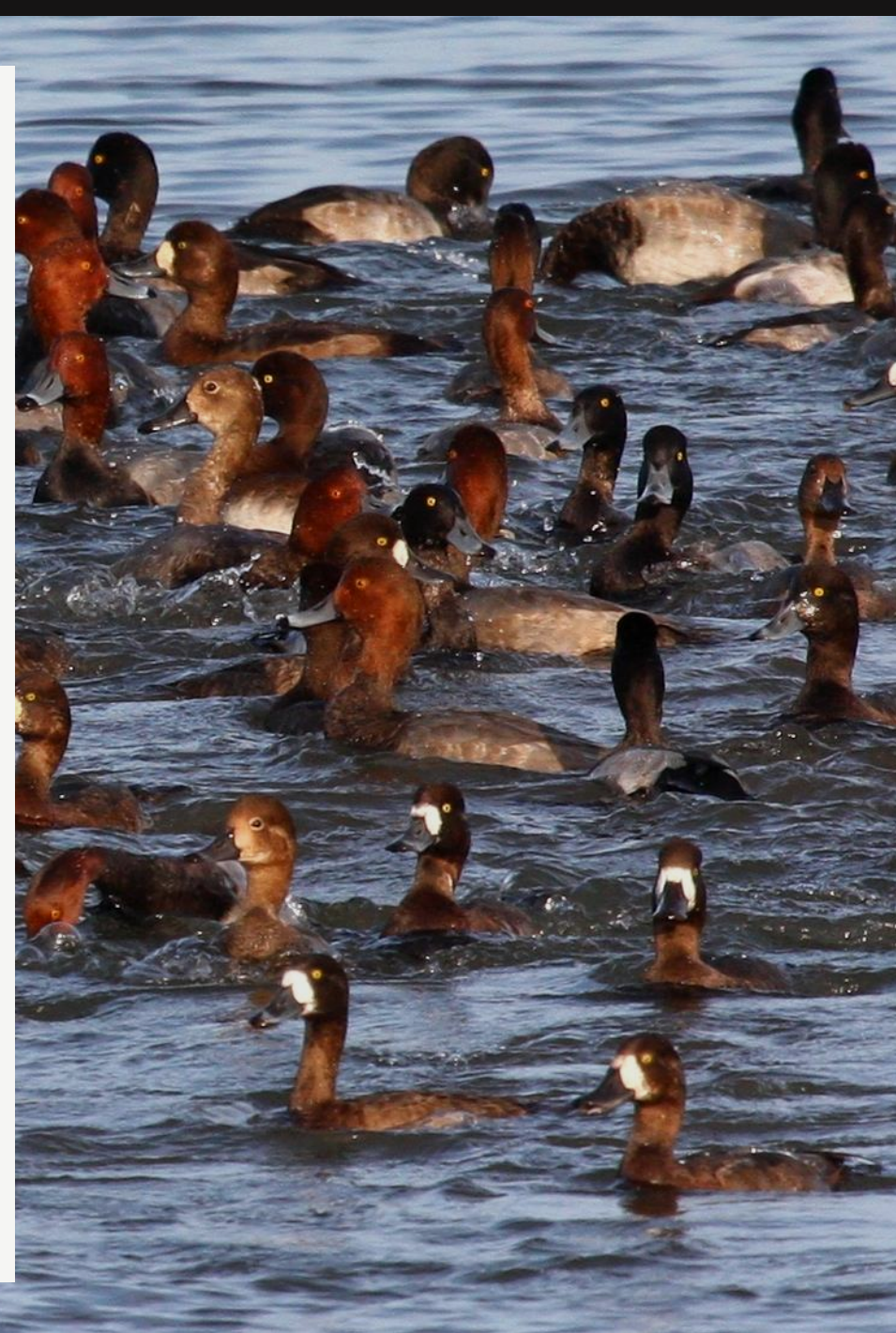


Correspondence Analysis (CA): Steps

Step 1: Transform contingency table into a table of contributions ($\bar{\mathbf{Q}}$) to the Pearson chi-square statistic (χ^2_P).

Table of frequencies, \mathbf{Q} :

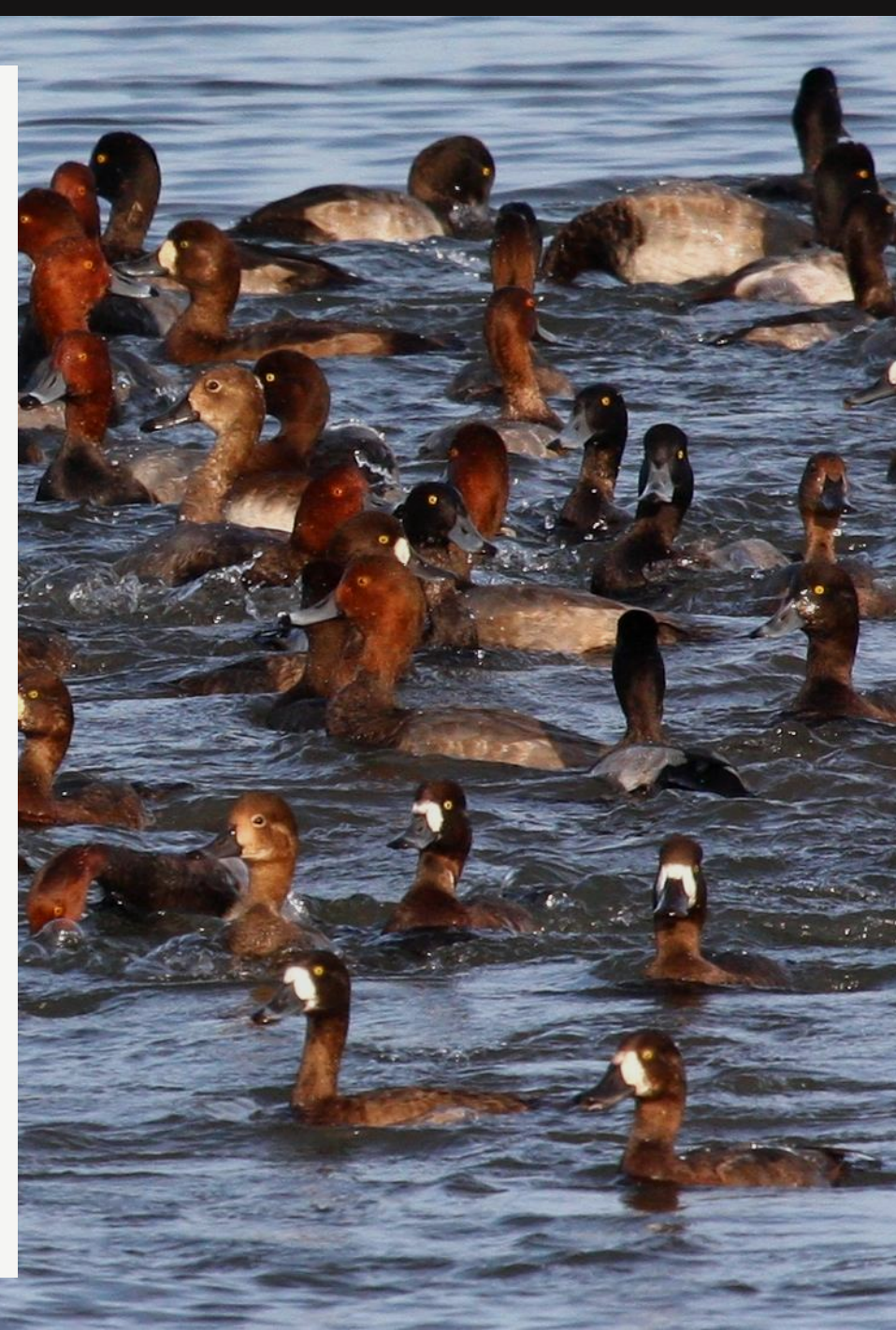
$$p_{ij} = f_{ij}/f_{++}$$



Correspondence Analysis (CA): Steps

Table Q

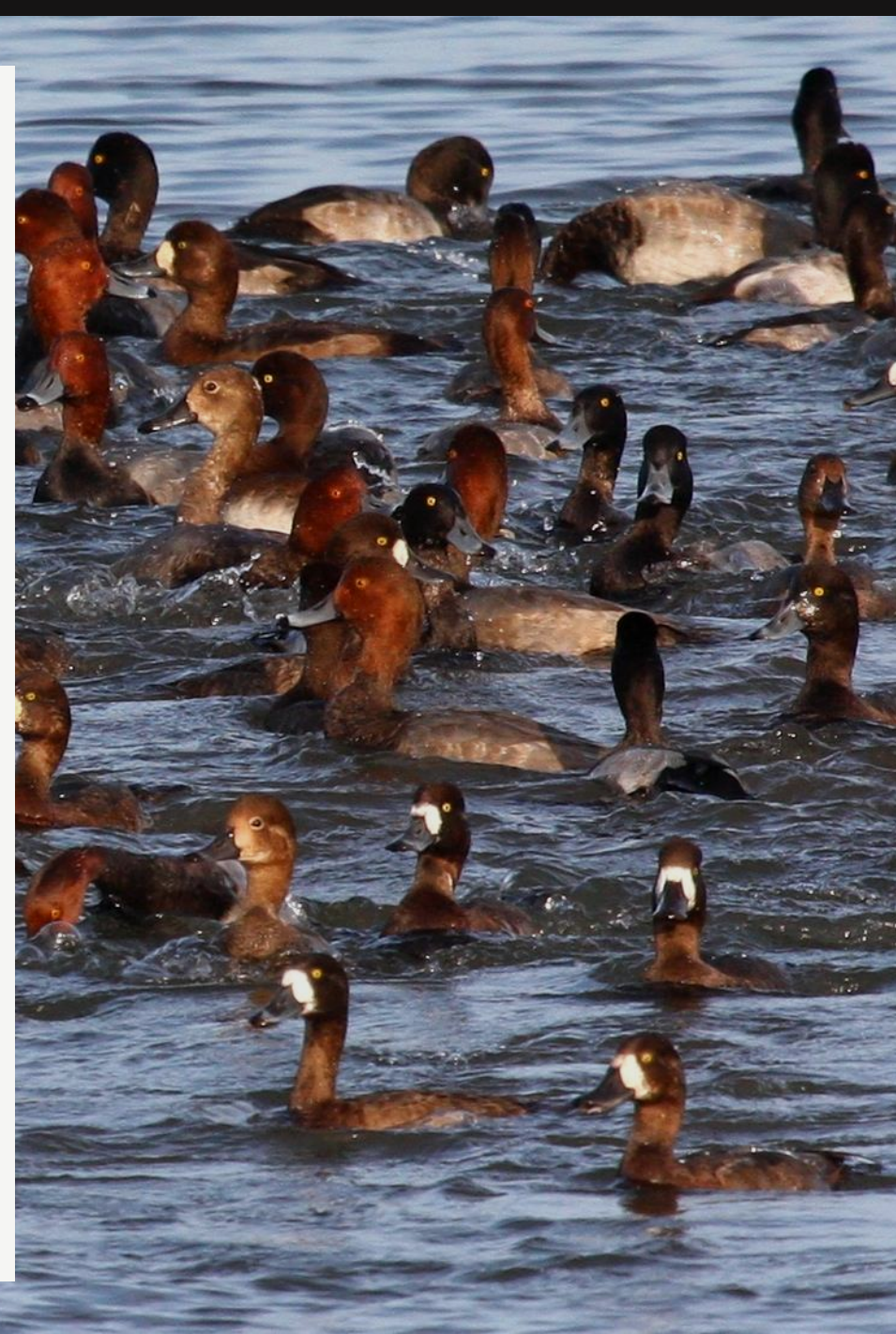
Site ID	Redband Trout	Brook Trout	Dace	Redside Shiner	Sculpin	Sucker
Silvies-15	0	0	0	0.013	0.018	0
Silvies-07	0.003	0	0.016	0	0.026	0
Silvies-12	0.005	0.063	0	0	0.025	0
Silvies-34	0	0	0.033	0	0	0.014
Silvies-11	0	0	0.021	0.004	0	0
Silvies-21	0.003	0	0	0	0	0
Silvies-08	0.004	0	0	0	0	0.001
Silvies-05	0.129	0	0	0	0.029	0.026
Silvies-03	0.030	0	0	0	0	0
Silvies-02	0.058	0	0.032	0	0	0.022
Silvies-22	0.221	0	0.026	0	0	0
Silvies-18	0.151	0	0.026	0	0	0



Correspondence Analysis (CA): Steps

Row weights: p_{i+}

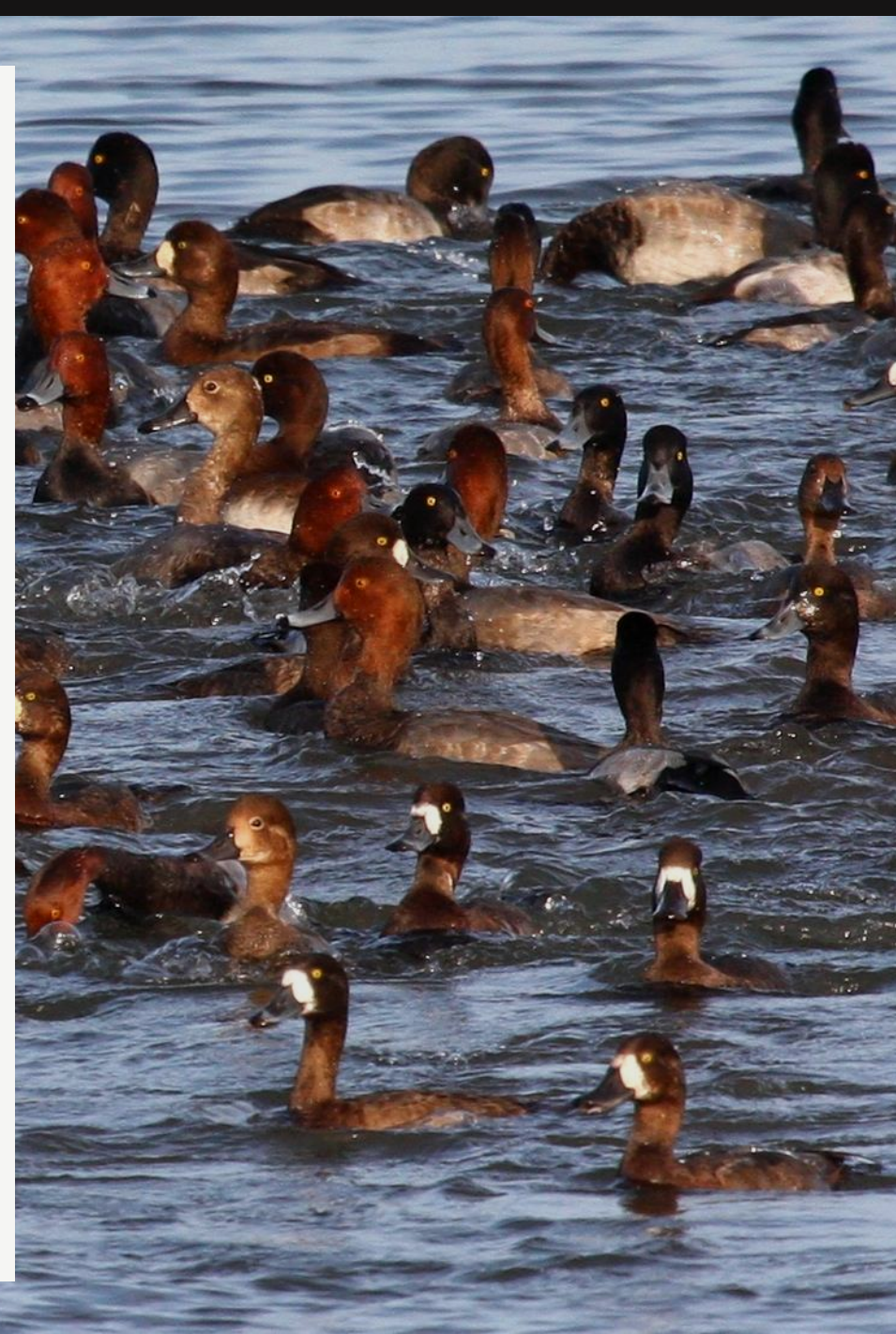
<u>Site ID (Row)</u>	<u>p_{i+}</u>
Silvies-15	0.032
Silvies-07	0.045
Silvies-12	0.093
Silvies-34	0.047
Silvies-11	0.025
Silvies-21	0.003
Silvies-08	0.005
Silvies-05	0.184
Silvies-03	0.030
Silvies-02	0.112
Silvies-22	0.247
<u>Silvies-18</u>	<u>0.177</u>



Correspondence Analysis (CA): Steps

Column weights: p_{+j}

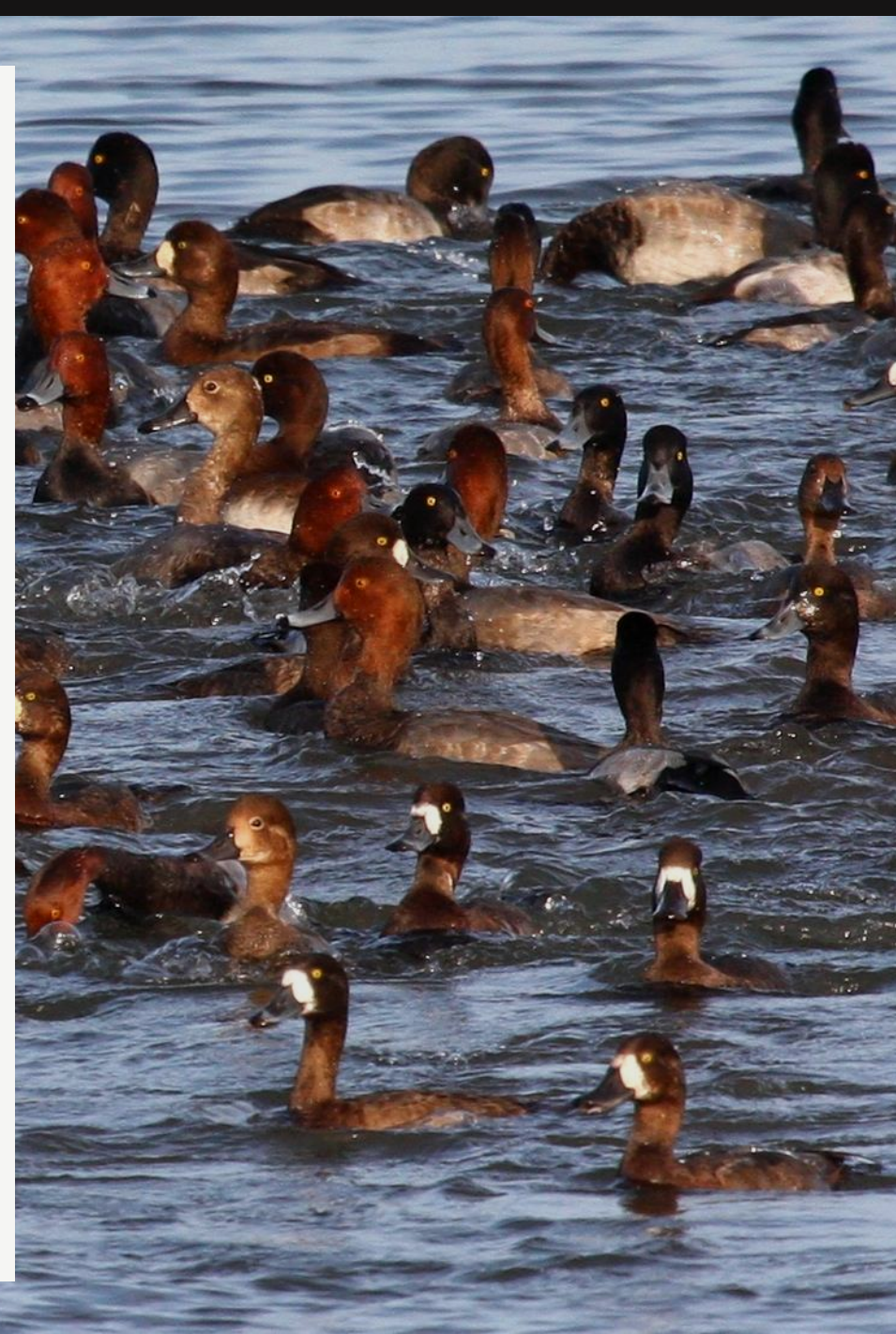
<u>Species (Column)</u>	<u>p_{+j}</u>
Redband Trout	0.603
Brook Trout	0.063
Dace	0.154
Redside Shiner	0.017
Sculpin	0.099
Sucker	0.064



Correspondence Analysis (CA): Steps

Step 1: Transform contingency table into a table of contributions (\bar{Q}) to the Pearson chi-square statistic (χ^2_p).

$$\chi_{ij} = \sqrt{f_{++}} \left[\frac{p_{ij} - p_{i+}p_{+j}}{\sqrt{p_{i+}p_{+j}}} \right]$$

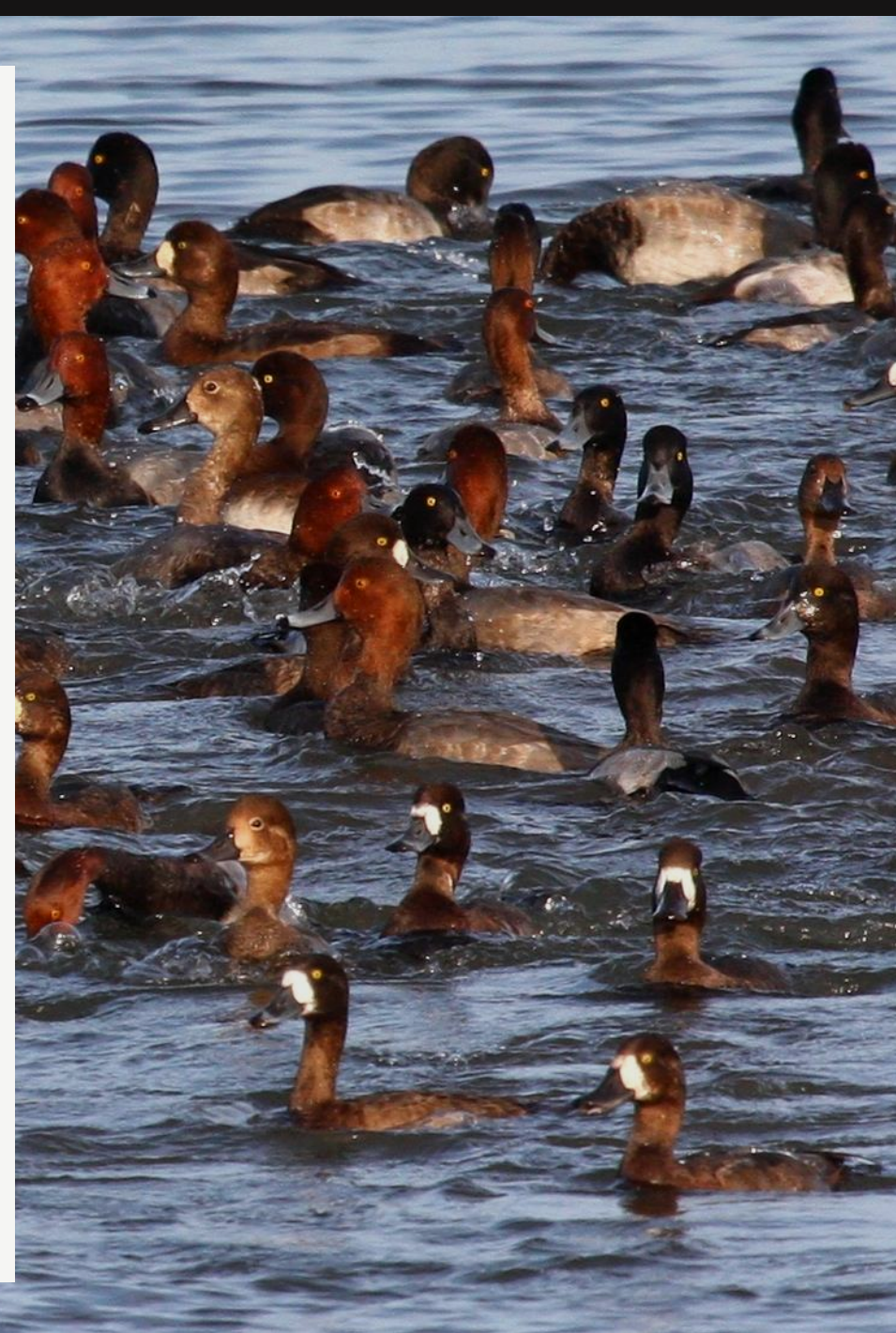


Correspondence Analysis (CA): Steps

Step 1: Transform contingency table into a table of contributions (\bar{Q}) to the Pearson chi-square statistic (χ^2_P).

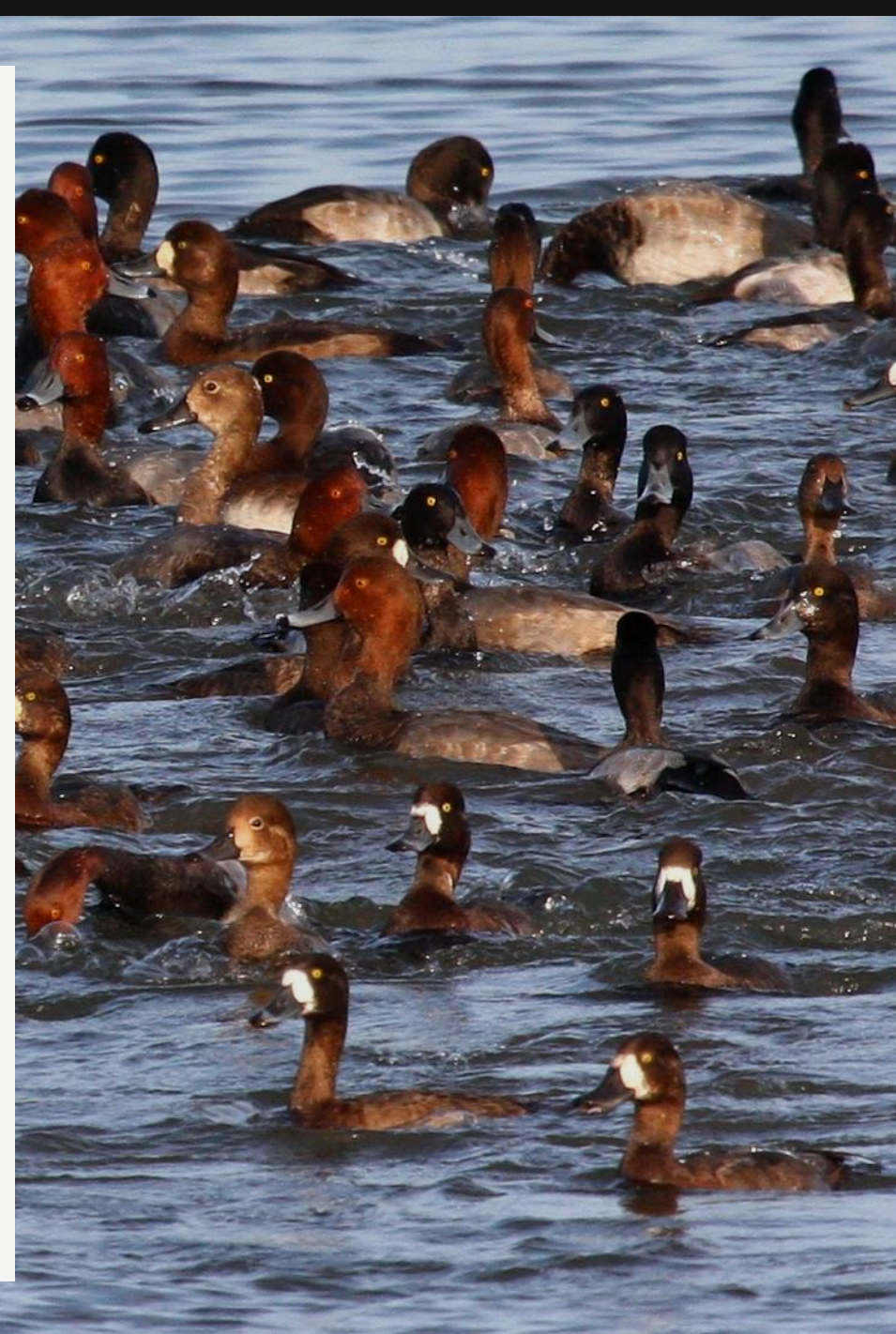
$$\chi_{ij} = \sqrt{f_{++}} \left[\frac{p_{ij} - p_{i+}p_{+j}}{\sqrt{p_{i+}p_{+j}}} \right]$$

$$\bar{Q} = \left[\frac{p_{ij} - p_{i+}p_{+j}}{\sqrt{p_{i+}p_{+j}}} \right]$$



Correspondence Analysis (CA): Steps

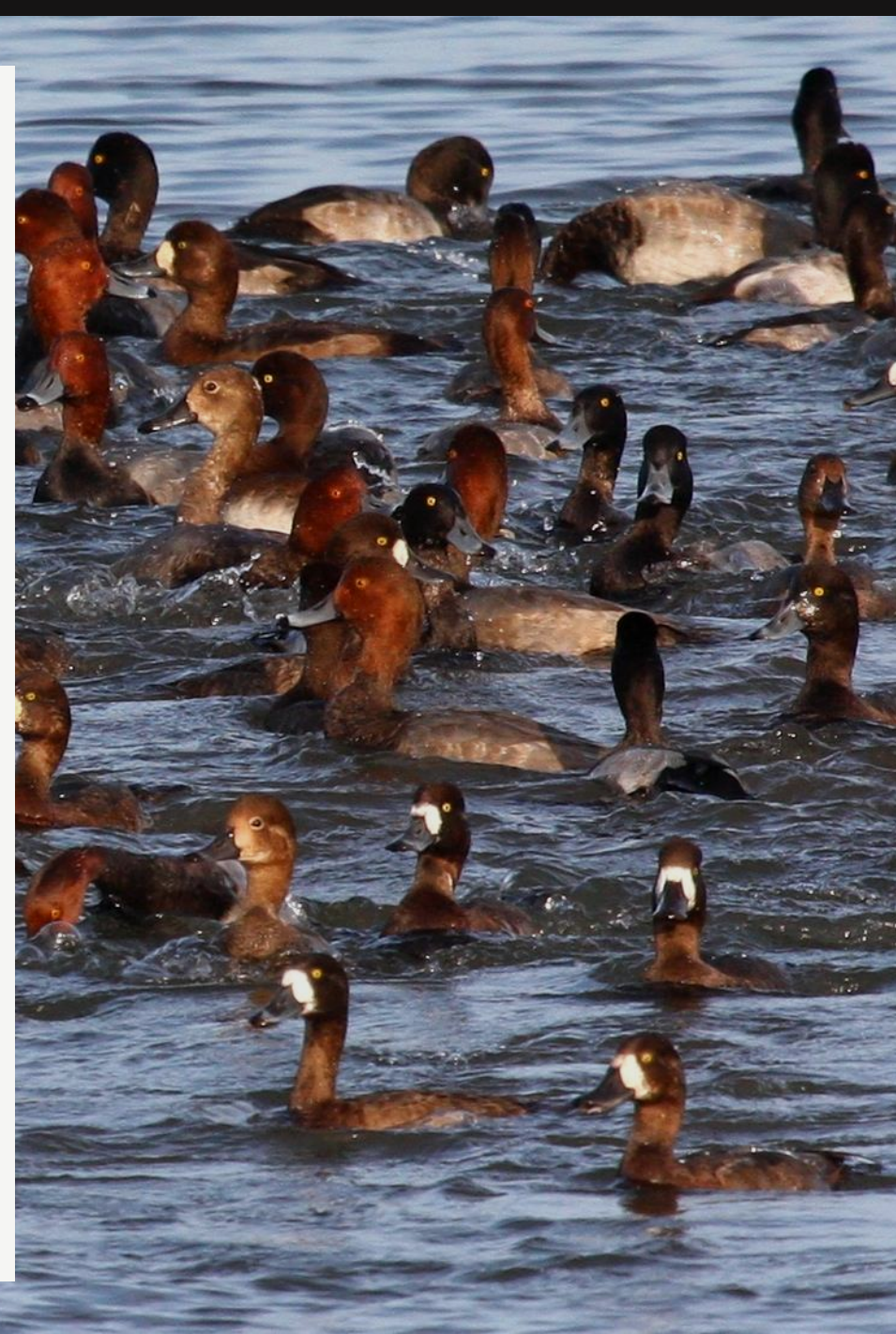
Step 2: Decompose transformed data table to obtain eigenvalues and eigenvectors.



Correspondence Analysis (CA): Steps

Step 2: Decompose transformed data table to obtain eigenvalues and eigenvectors.

Conduct eigenanalysis on the covariance matrix:
 $\bar{Q}'\bar{Q}$



Correspondence Analysis (CA): Steps

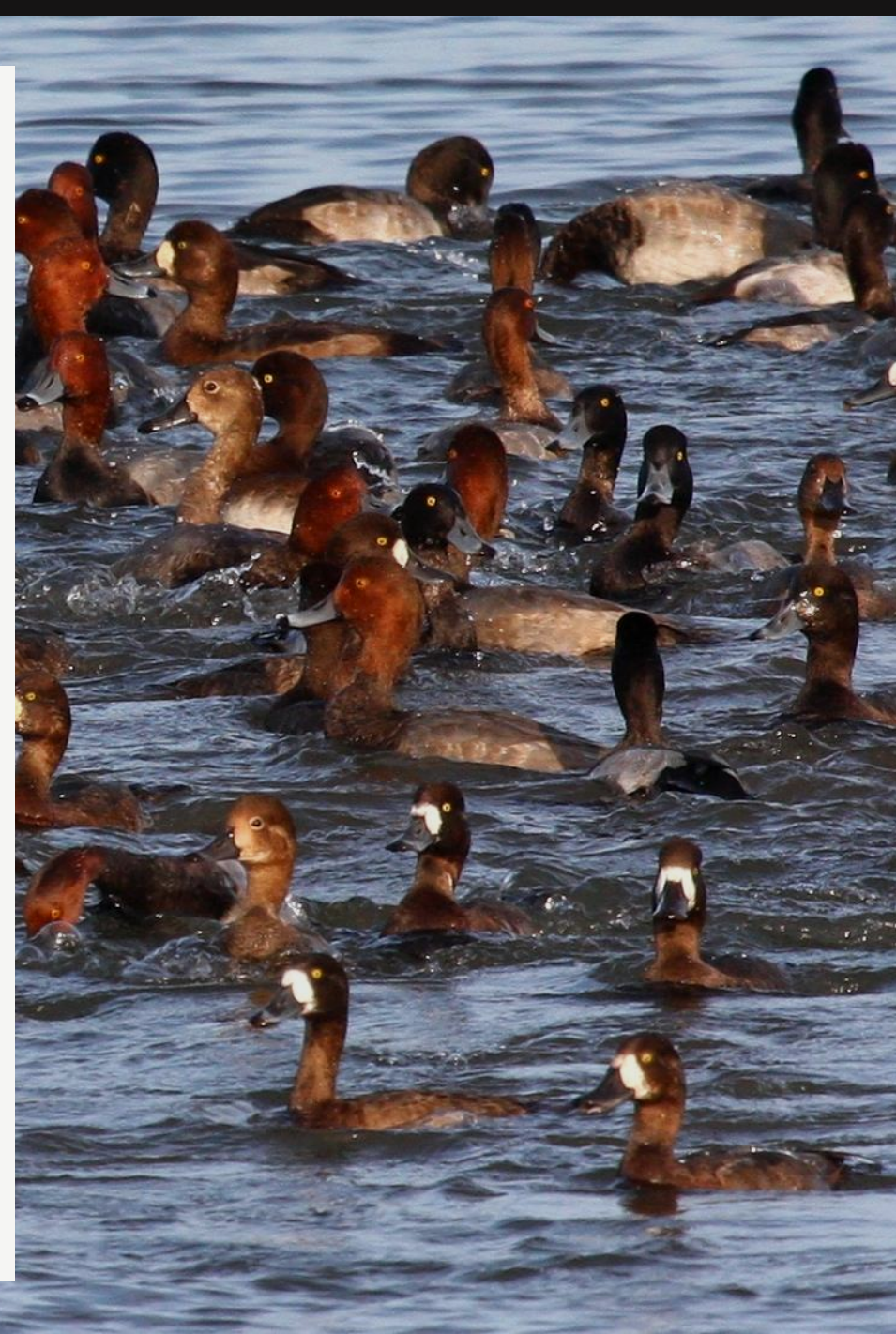
Step 2: Decompose transformed data table to obtain eigenvalues and eigenvectors.

Conduct eigenanalysis on the covariance matrix:

$$\bar{\mathbf{Q}}'\bar{\mathbf{Q}}$$

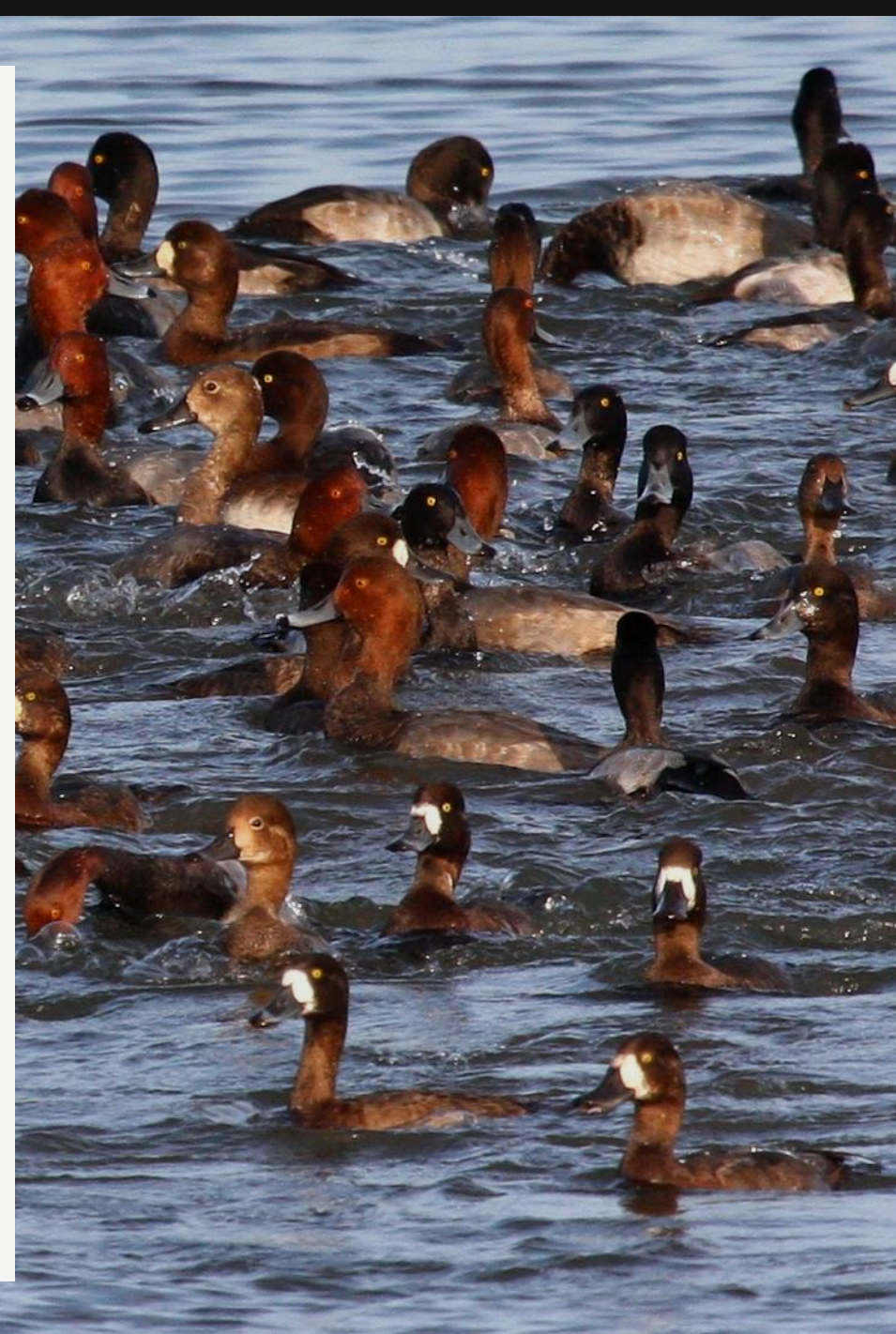
$$\lambda_1 = 0.77 \quad \lambda_2 = 0.48 \quad \lambda_3 = 0.32$$

$$\lambda_4 = 0.14 \quad \lambda_5 = 0.09 \quad \lambda_6 = 0.00$$



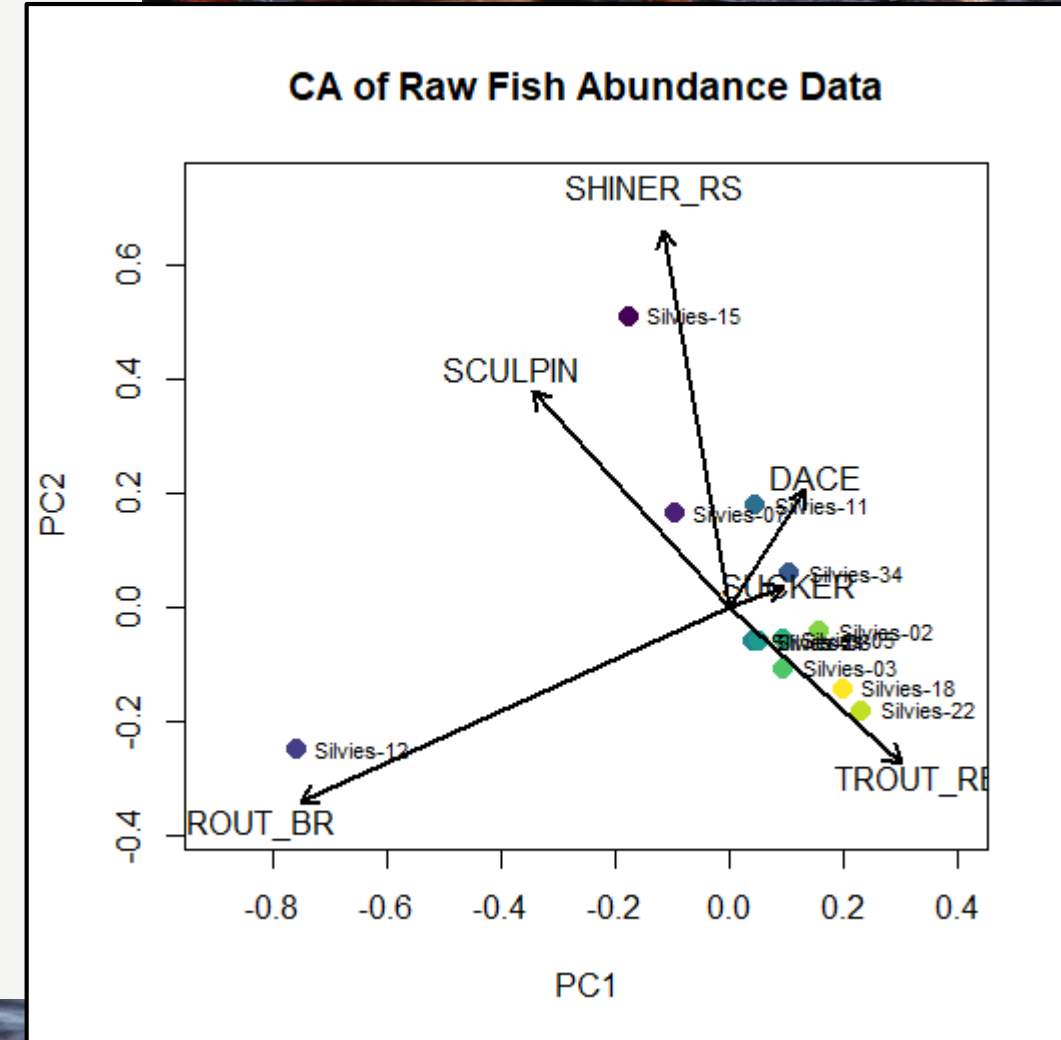
Correspondence Analysis (CA): Steps

Step 3: Scale as appropriate for plotting and data interpretation.



Correspondence Analysis (CA): Steps

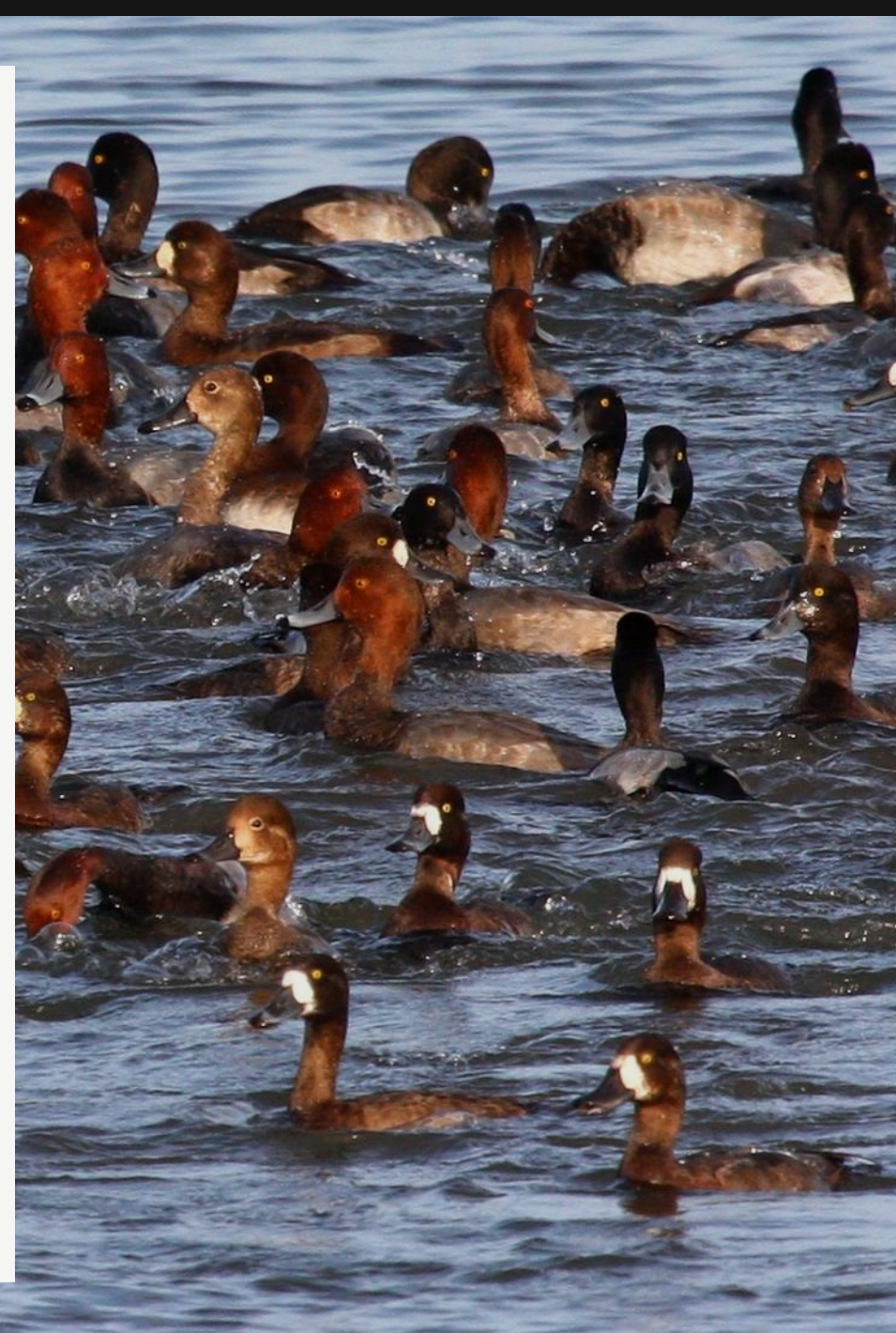
Step 3: Scale as appropriate for plotting and data interpretation.



Correspondence Analysis (CA): Steps

Step 3: Scale as appropriate for plotting and data interpretation.

Scaling 1: Draw sites at the centroids of the species – most appropriate for representing distances among sites

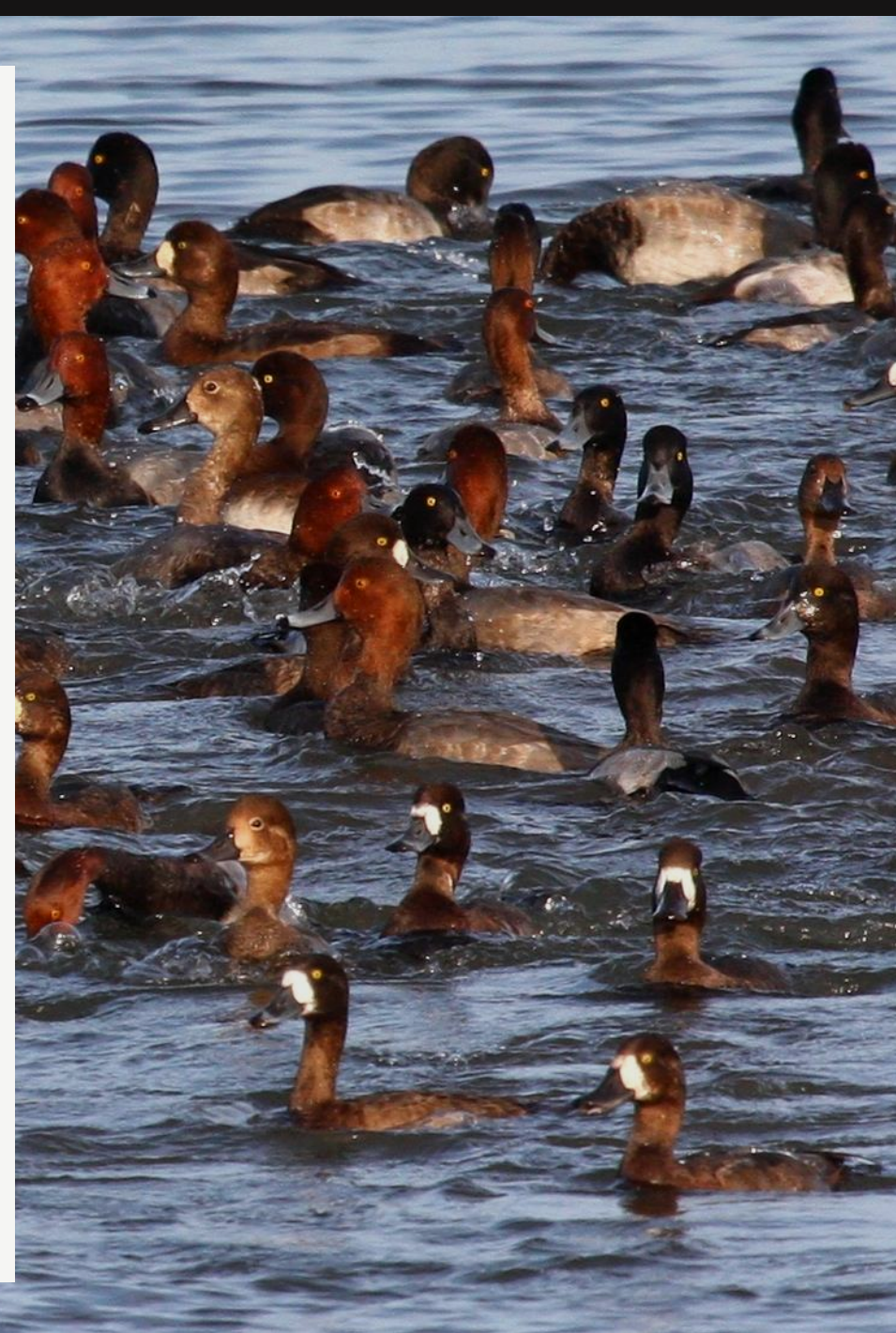


Correspondence Analysis (CA): Steps

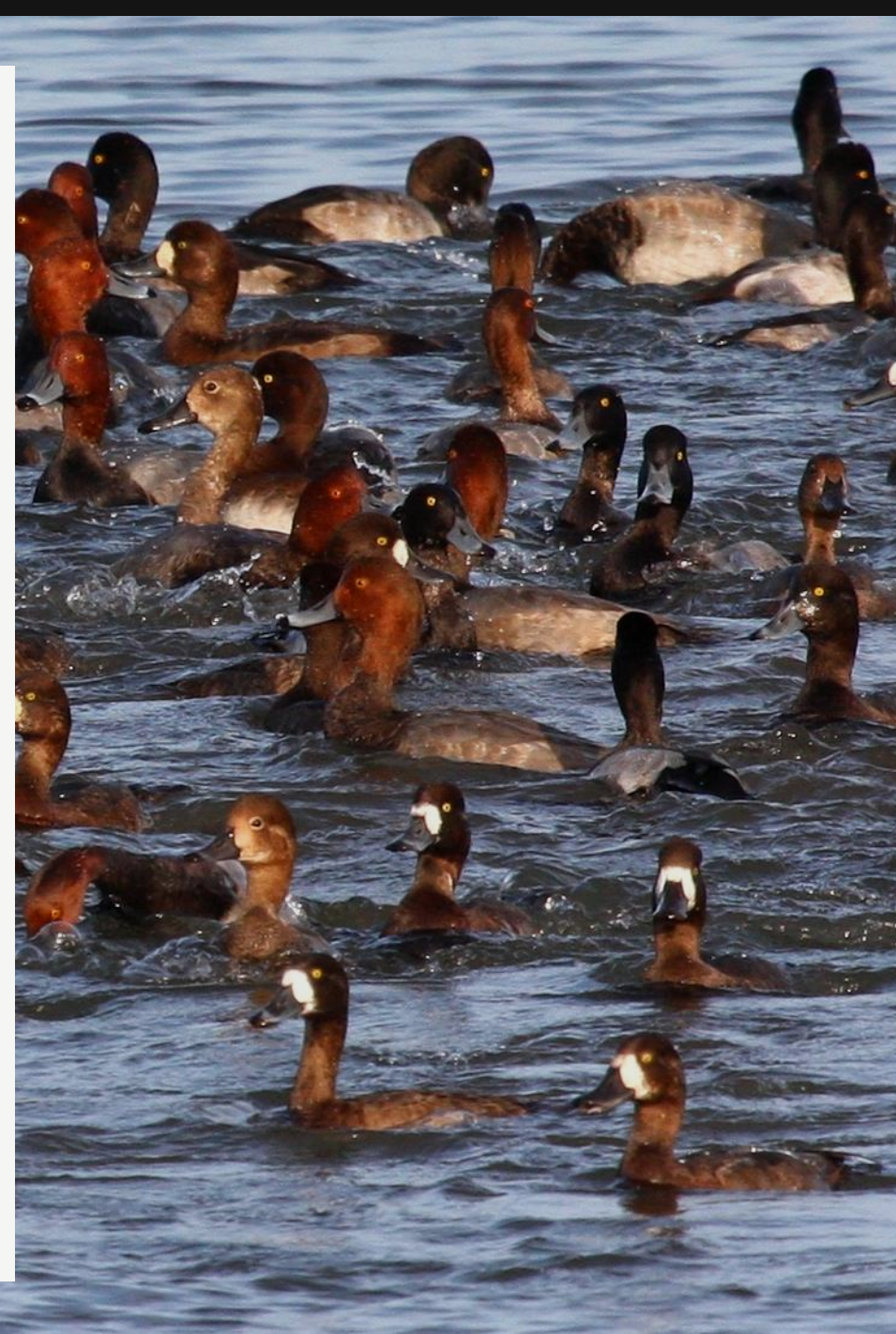
Step 3: Scale as appropriate for plotting and data interpretation.

Scaling 1: Draw sites at the centroids of the species – most appropriate for representing distances among sites

Scaling 2: Draw species at the centroids of the sites – most appropriate for representing distances among species

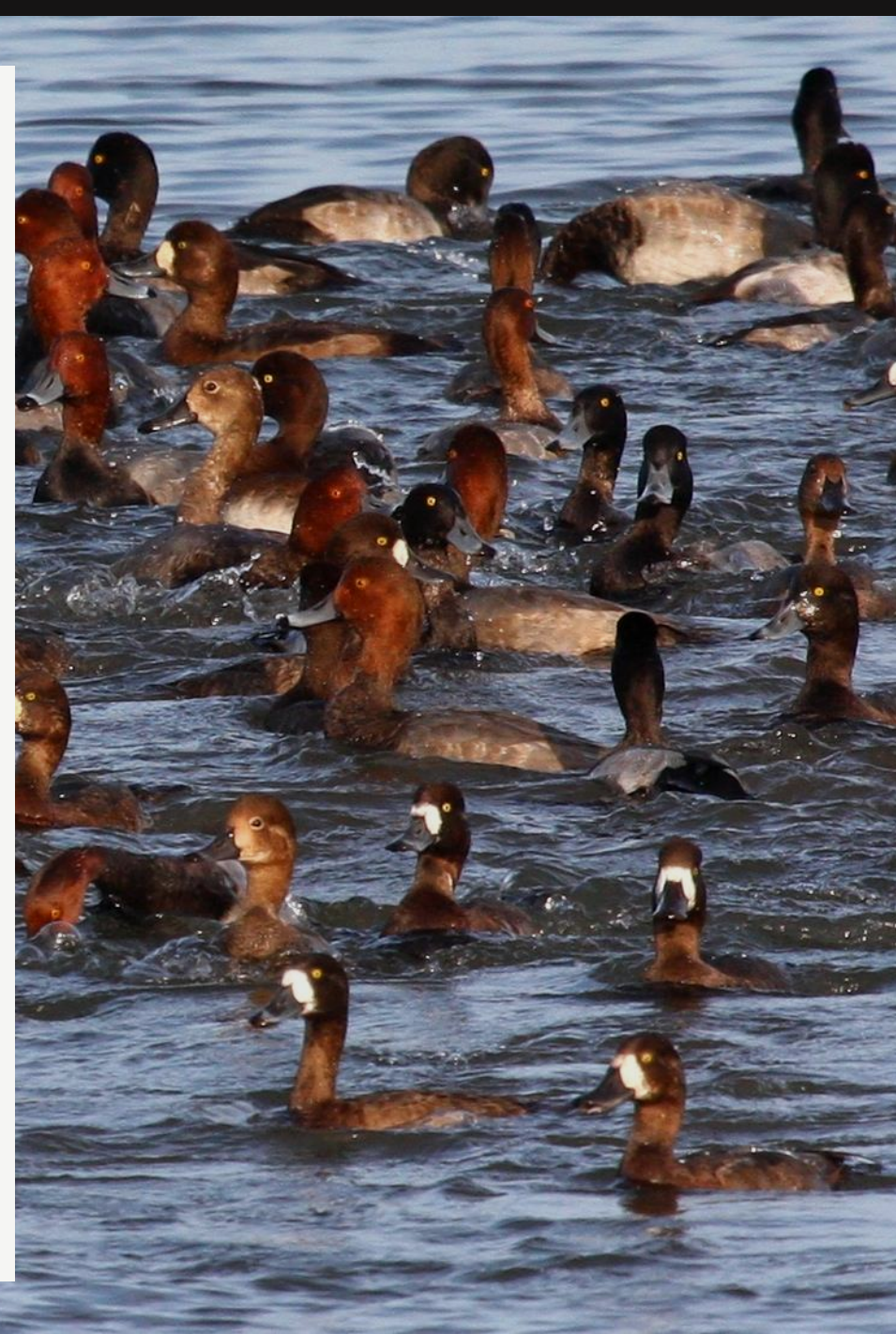


Correspondence Analysis (CA): Limitations



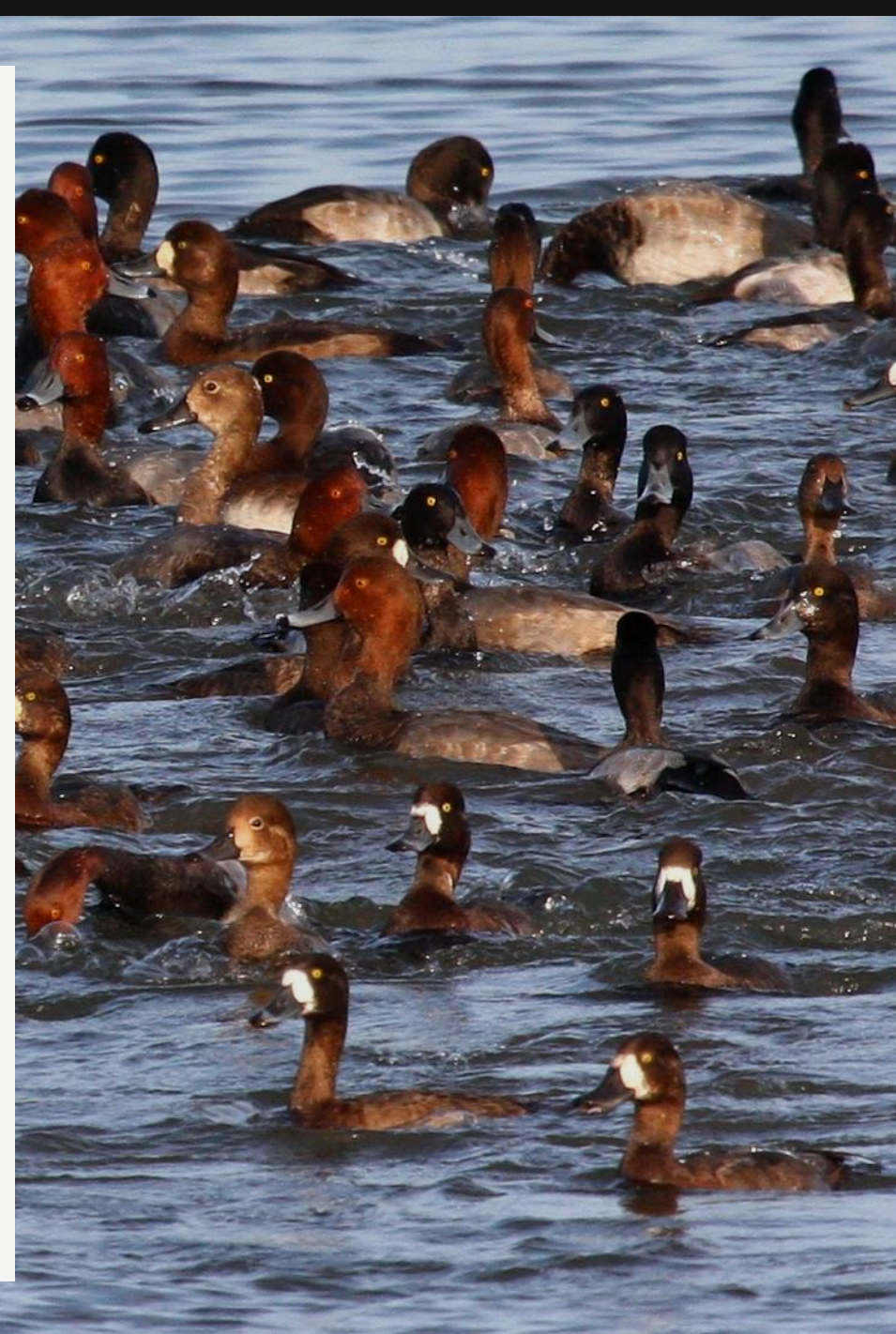
Correspondence Analysis (CA): Limitations

- Data must be dimensionally homogeneous and non-negative



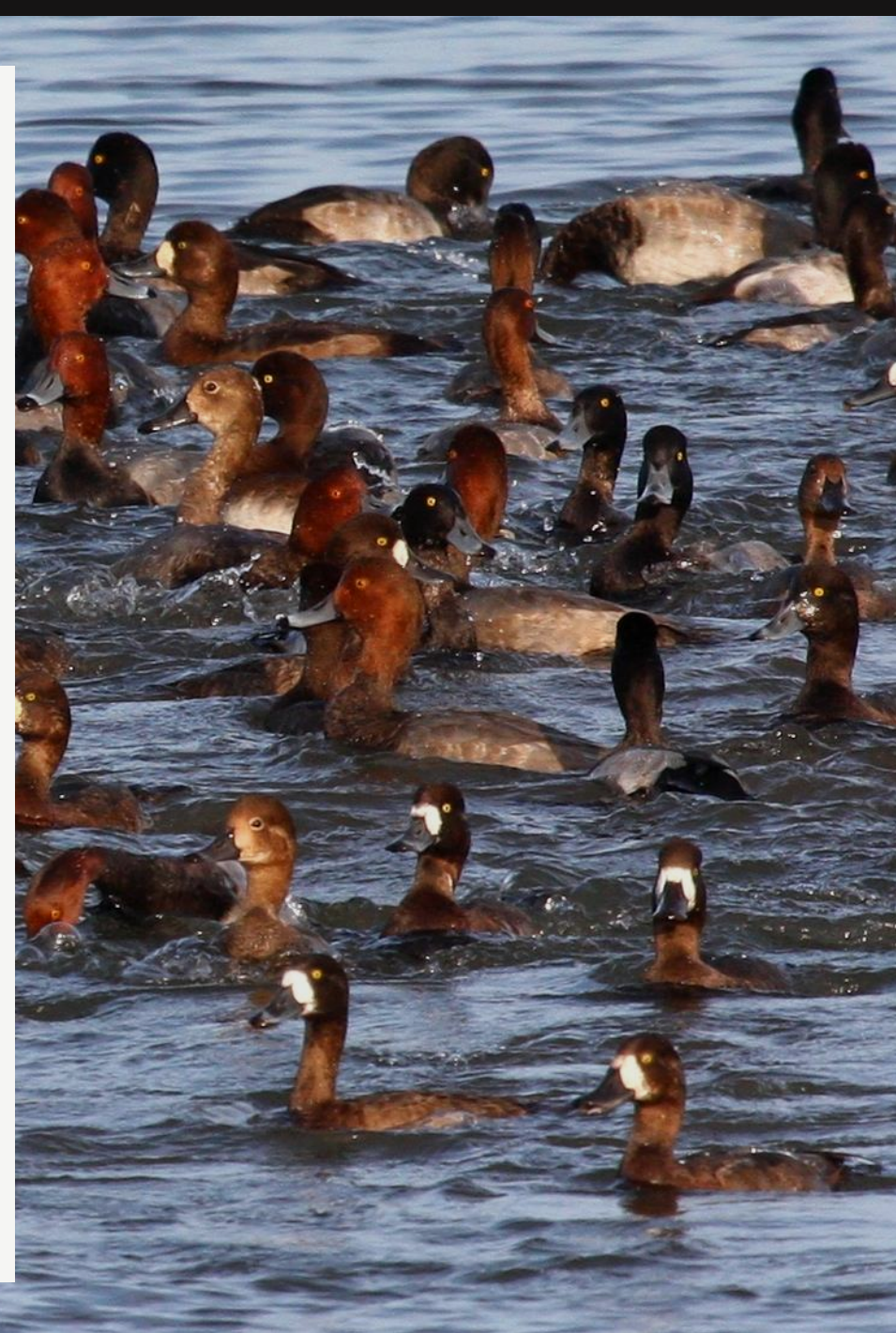
Correspondence Analysis (CA): Limitations

- Data must be dimensionally homogeneous and non-negative
- Assumes linear relationships with environmental gradients



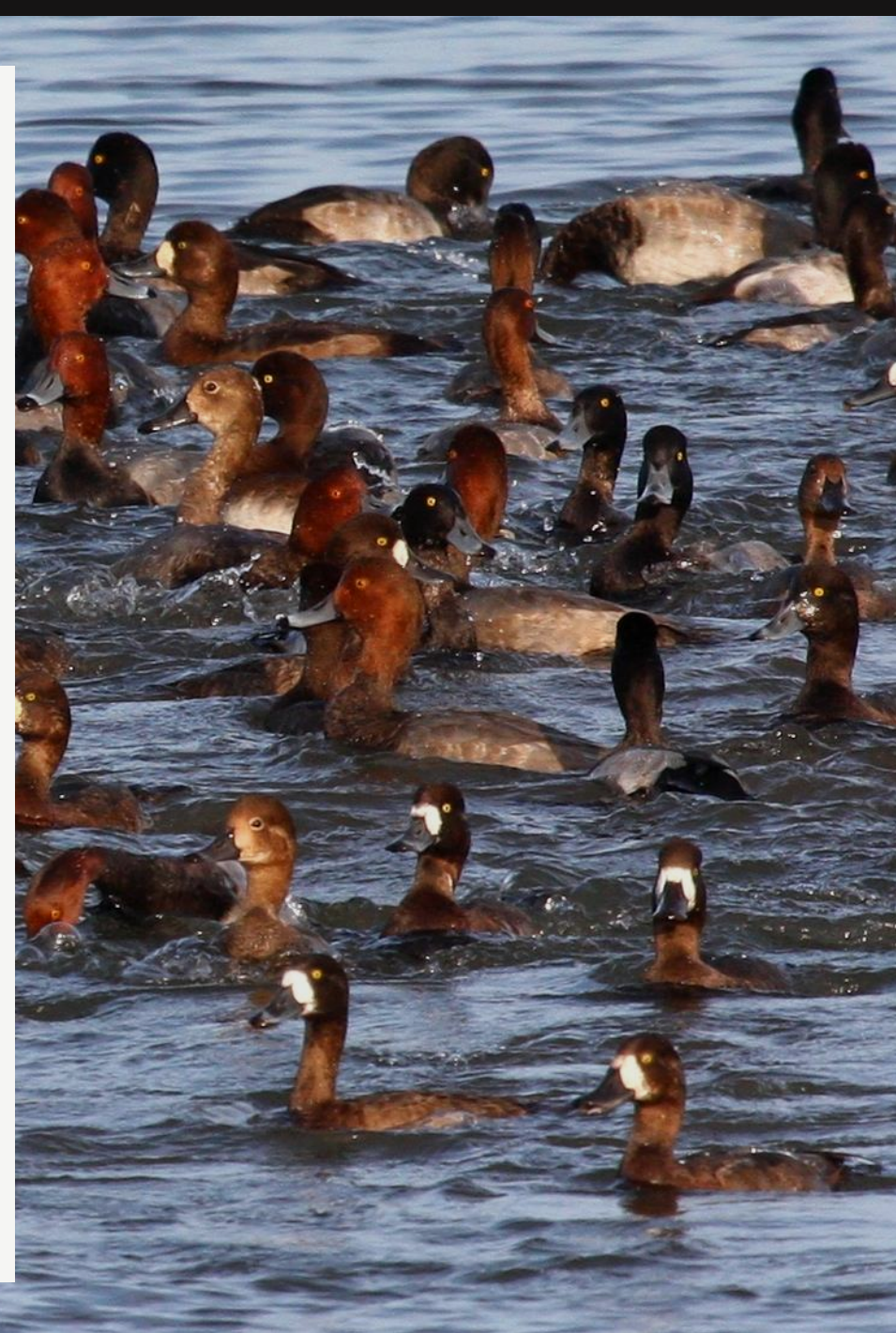
Correspondence Analysis (CA): Limitations

- Data must be dimensionally homogeneous and non-negative
- Assumes linear relationships with environmental gradients
- Assumes linear relationships between objects and descriptors



Correspondence Analysis (CA): Limitations

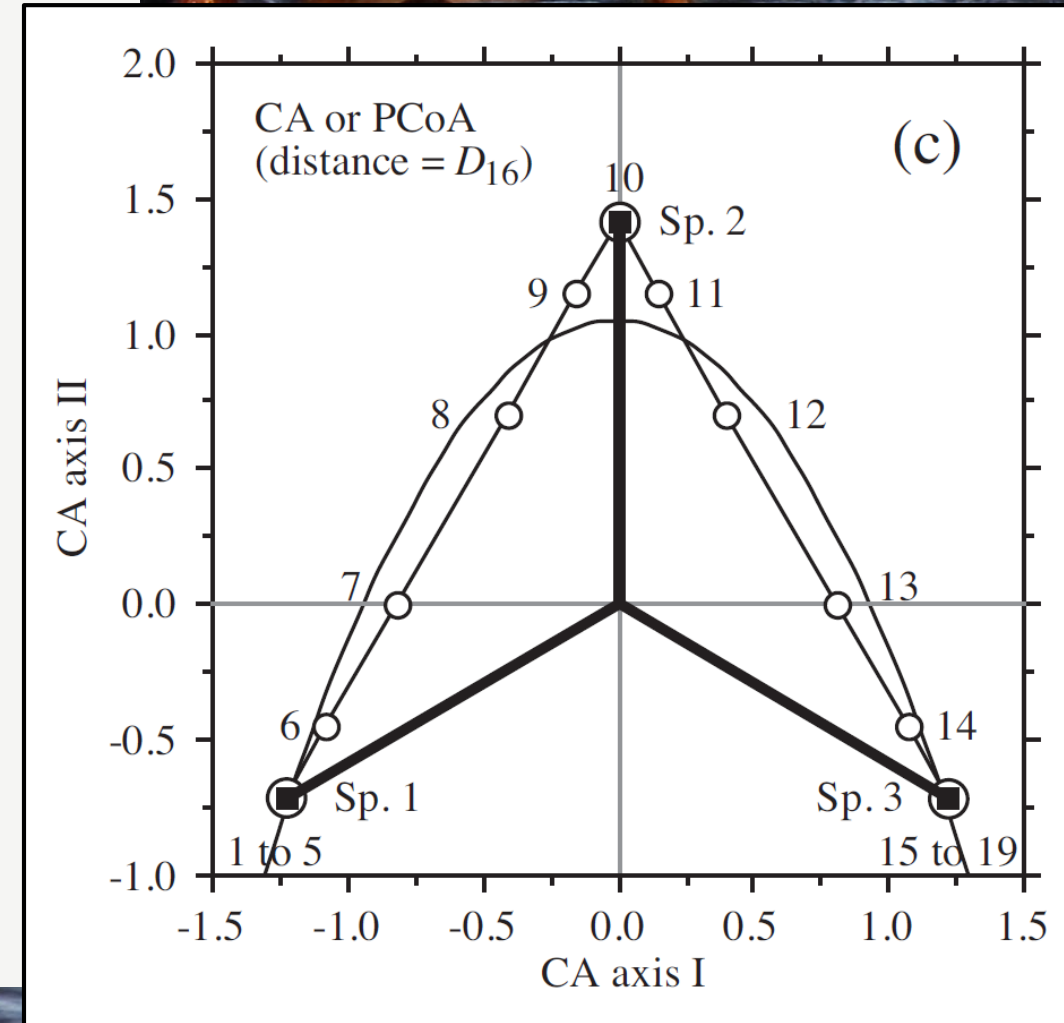
- Data must be dimensionally homogeneous and non-negative
- Assumes linear relationships with environmental gradients
- Assumes linear relationships between objects and descriptors
- Overly sensitive to rare species



Correspondence Analysis (CA): The Arch Effect

The **arch effect**, also known as the "horseshoe effect" or "Guttman effect," is a common artifact in ordination techniques.

Legendre & Legendre Fig. 9.11

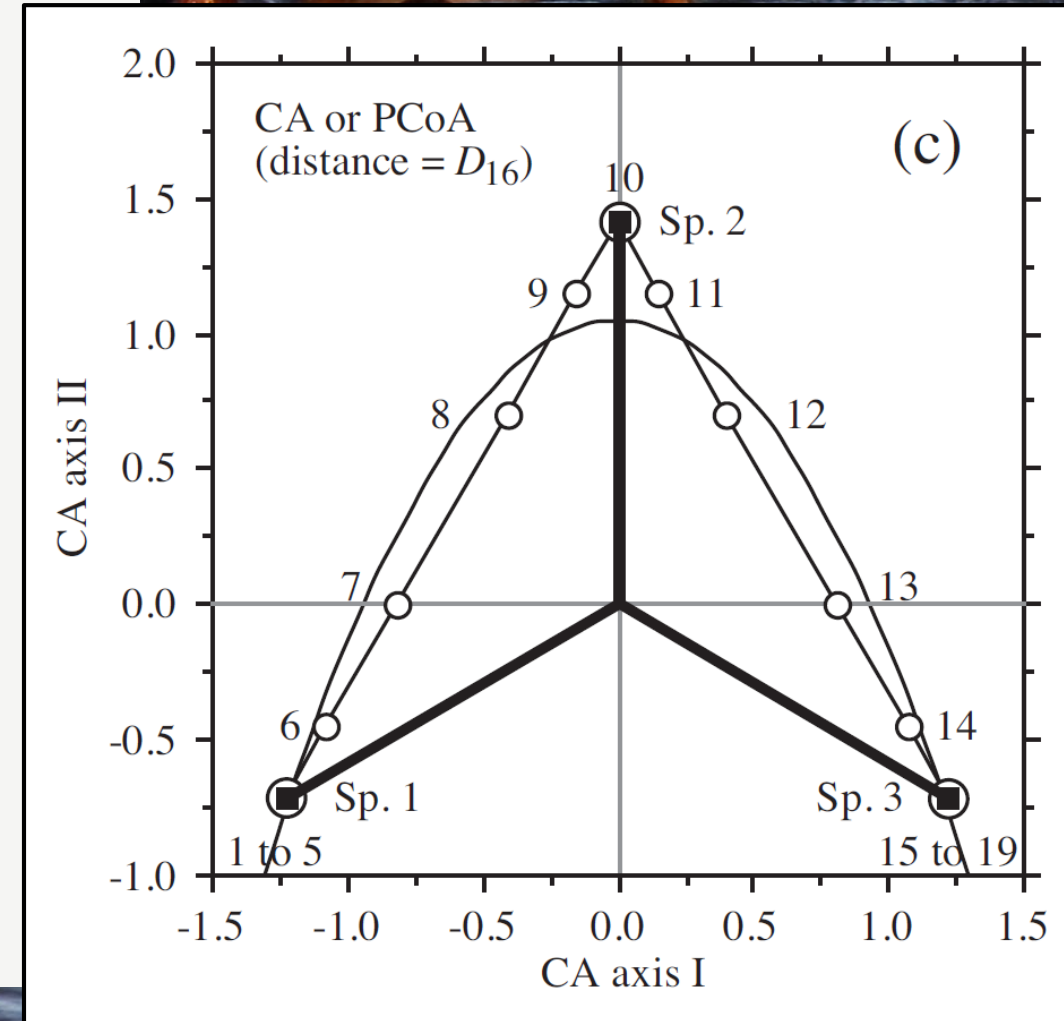


Correspondence Analysis (CA): The Arch Effect

The **arch effect**, also known as the "horseshoe effect" or "Guttman effect," is a common artifact in ordination techniques.

It occurs when the data points form a curved pattern or an arch-shaped (horseshoe) configuration in the ordination space.

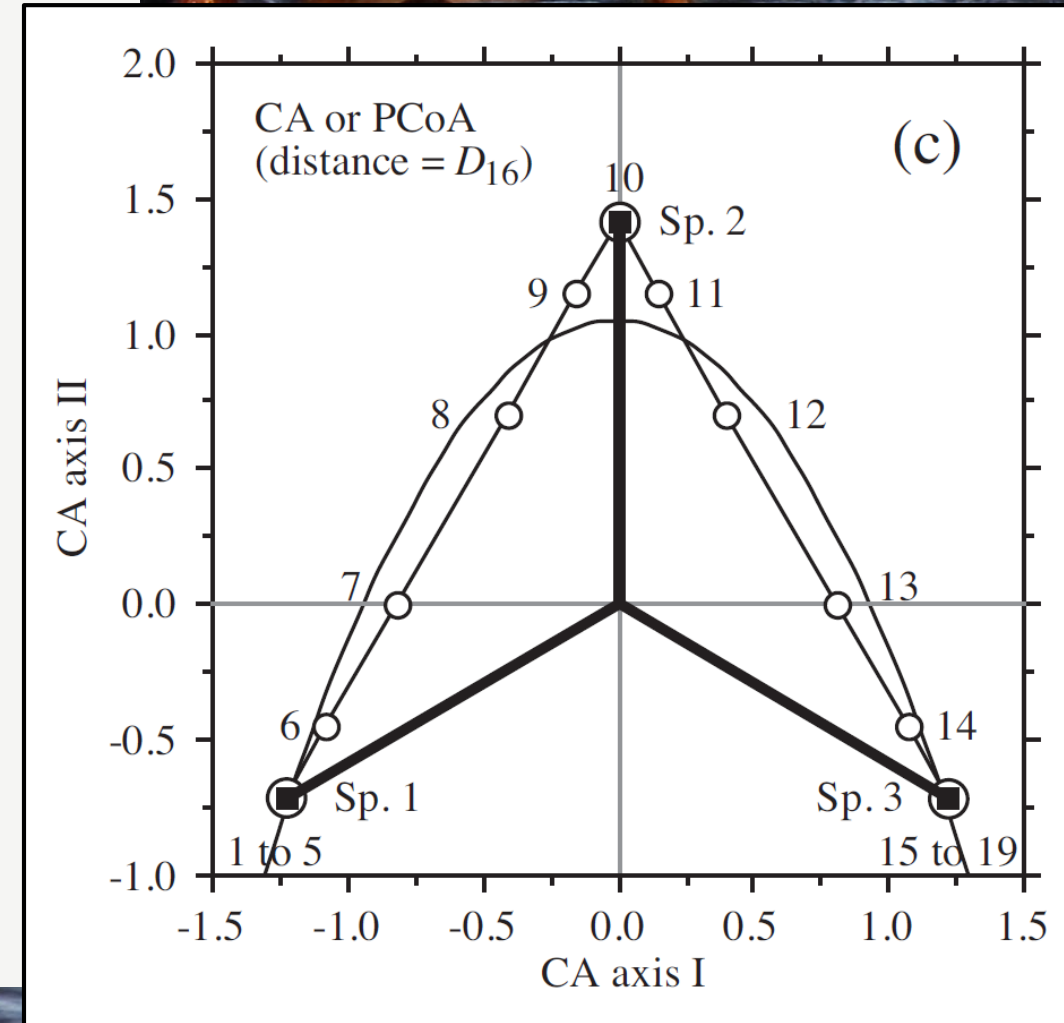
Legendre & Legendre Fig. 9.11



Correspondence Analysis (CA): The Arch Effect

The **arch effect**, also known as the "horseshoe effect" or "Guttman effect," is a common artifact in ordination techniques.
But why?

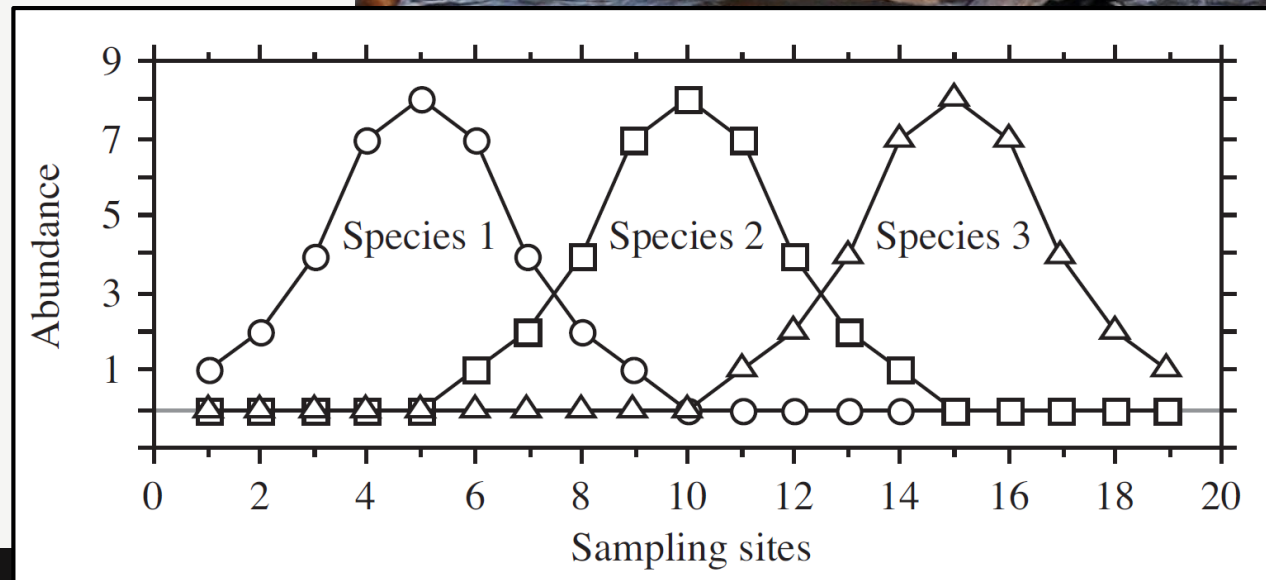
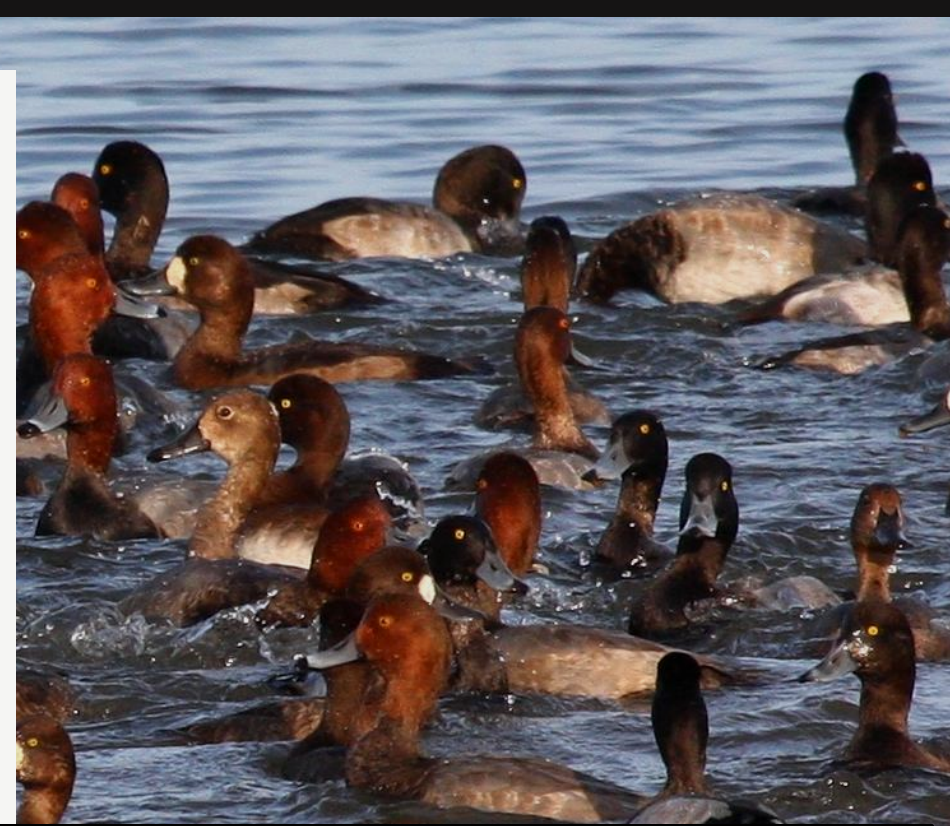
Legendre & Legendre Fig. 9.11



Correspondence Analysis (CA): The Arch Effect

The **arch effect**, also known as the "horseshoe effect" or "Guttman effect," is a common artifact in ordination techniques. **But why?**

- Non-linear (especially unimodal) relationships in the underlying data

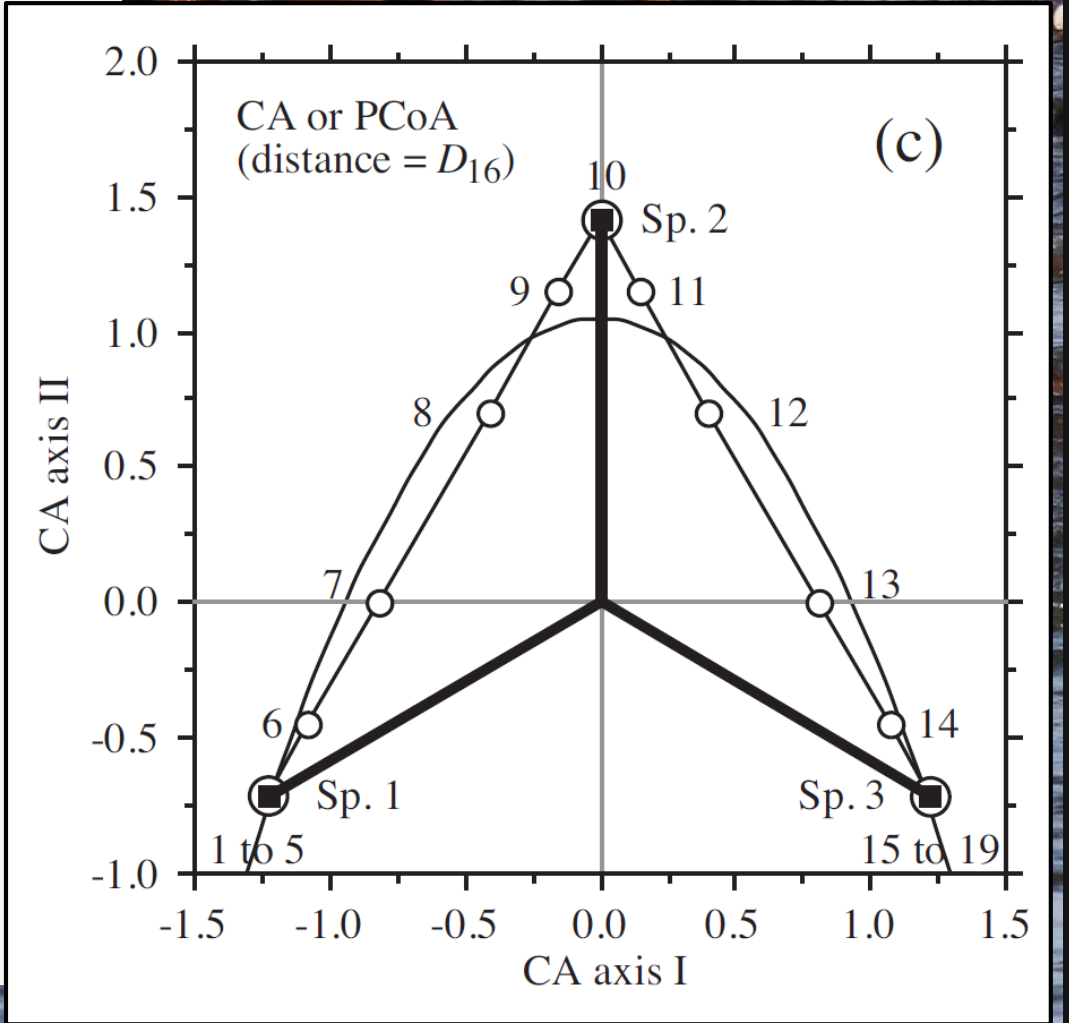


Legendre & Legendre Fig. 9.10

A group of Red-throated Loons swimming in water. The birds have dark brown heads and necks with a characteristic yellow patch around the eye. Their bodies are mostly submerged, with light-colored underparts visible. The water is a deep blue-grey color.

- **Misinterpretation:** Data points at the ends of the arch may appear closer together in the reduced-dimensionality space, suggesting a stronger similarity than actually exists.
- **Loss of Information:** Important environmental gradients or other underlying factors are not adequately represented in the first few ordination axes

Figure 1 is a CA or PCoA plot (distance = D_{16}) showing the relationship between CA axis I and CA axis II. The plot displays 19 numbered points (1 to 19) and three species (Sp. 1, Sp. 2, Sp. 3) marked with squares. The axes range from -1.5 to 1.5 on the x-axis and -1.0 to 2.0 on the y-axis. A thick line connects the three species, forming a triangle. A thin line connects points 1 to 5, 6 to 10, 11 to 14, and 15 to 19. A vertical line is drawn at CA axis I = 0.0, and a horizontal line is drawn at CA axis II = 0.0.

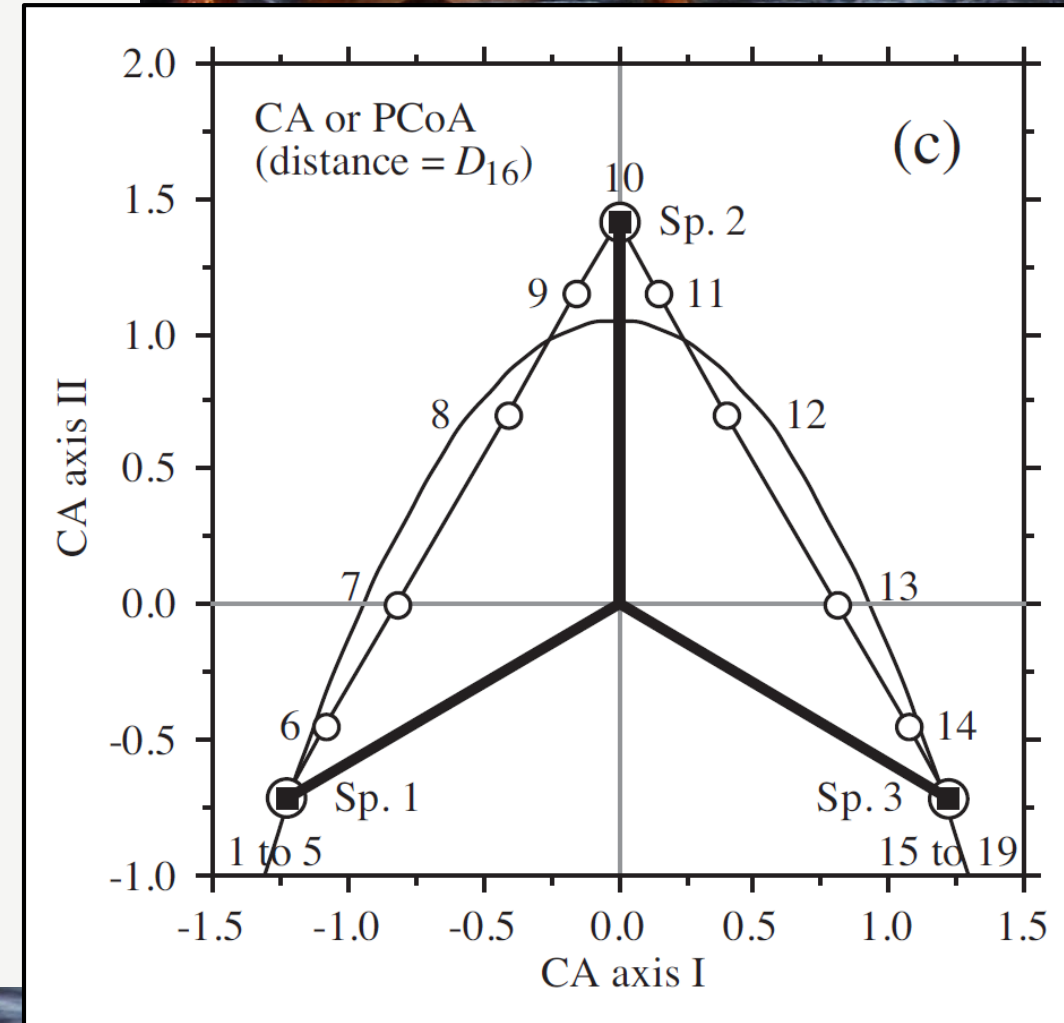


Correspondence Analysis (CA): The Arch Effect

Solutions:

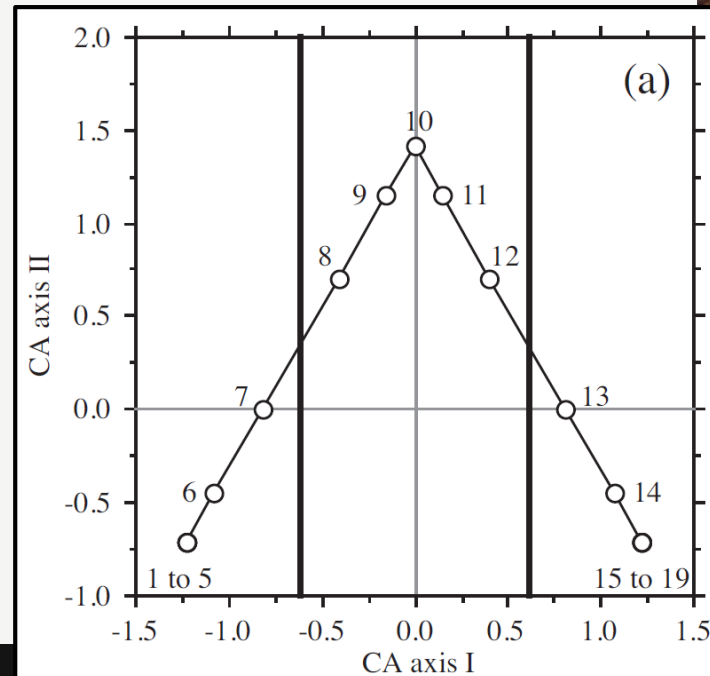
- **Detrended Correspondence Analysis (DCA):** Explicitly corrects for the curvature by removing the arch, providing a “more accurate” representation of the underlying gradients.
- **Non-metric Multidimensional Scaling (NMDS):** Is less prone to the arch effect because it focuses on preserving the rank order of distances rather than the exact distances.

Legendre & Legendre Fig. 9.11

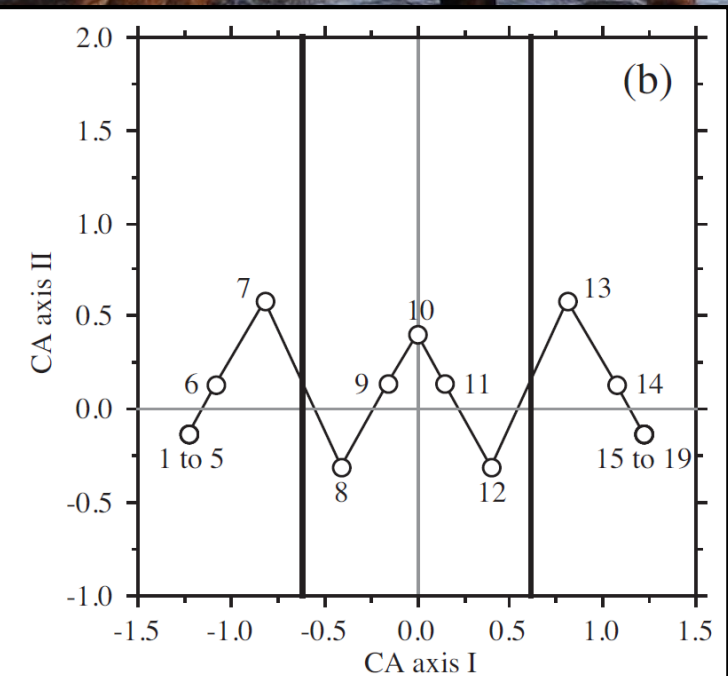


Correspondence Analysis (CA): Detrended Correspondence Analysis (DCA)

Detrending by segments: Axis 1 is divided into segments and the mean scores along axis 2 are set to zero.



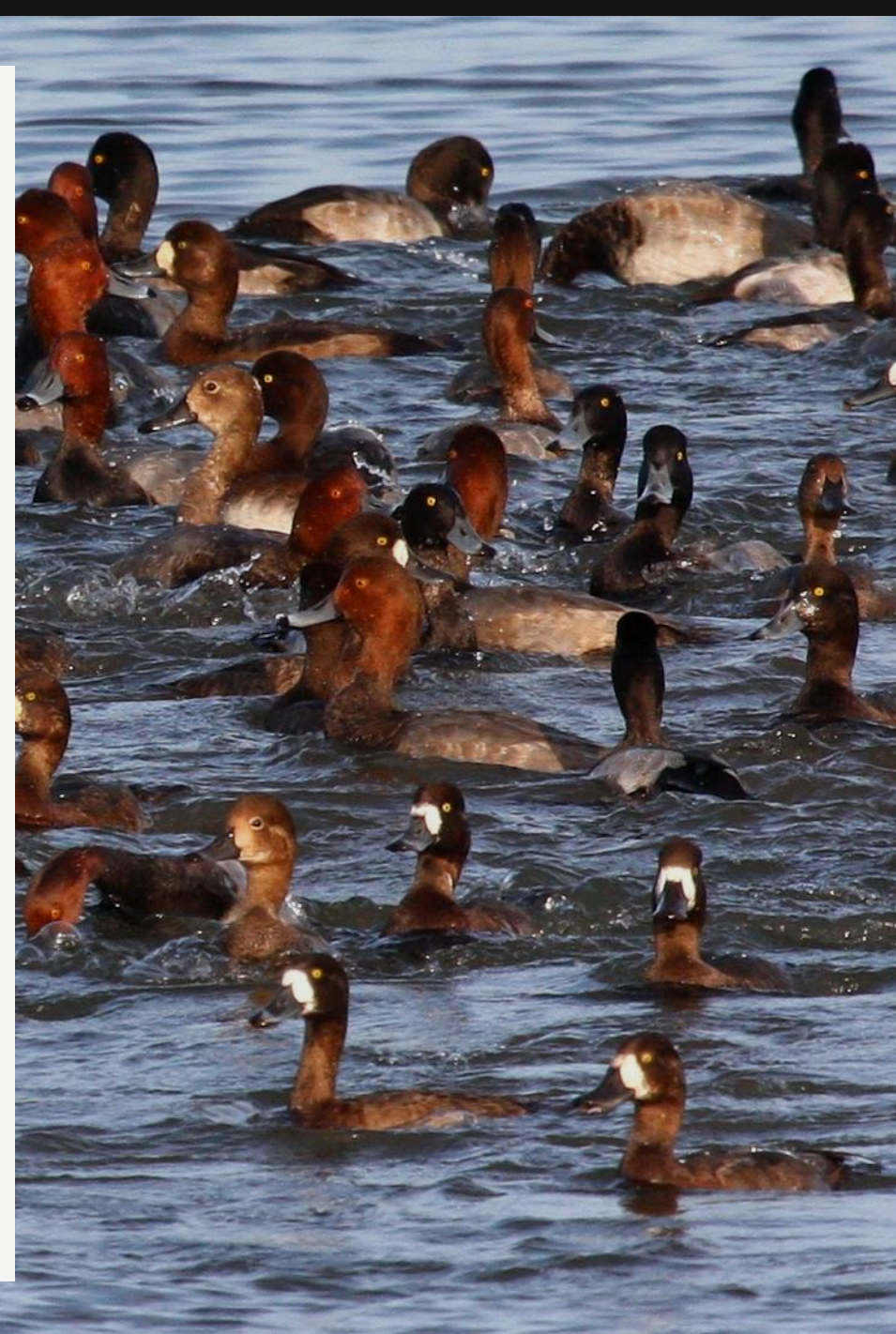
⇌



Legendre & Legendre Fig. 9.12

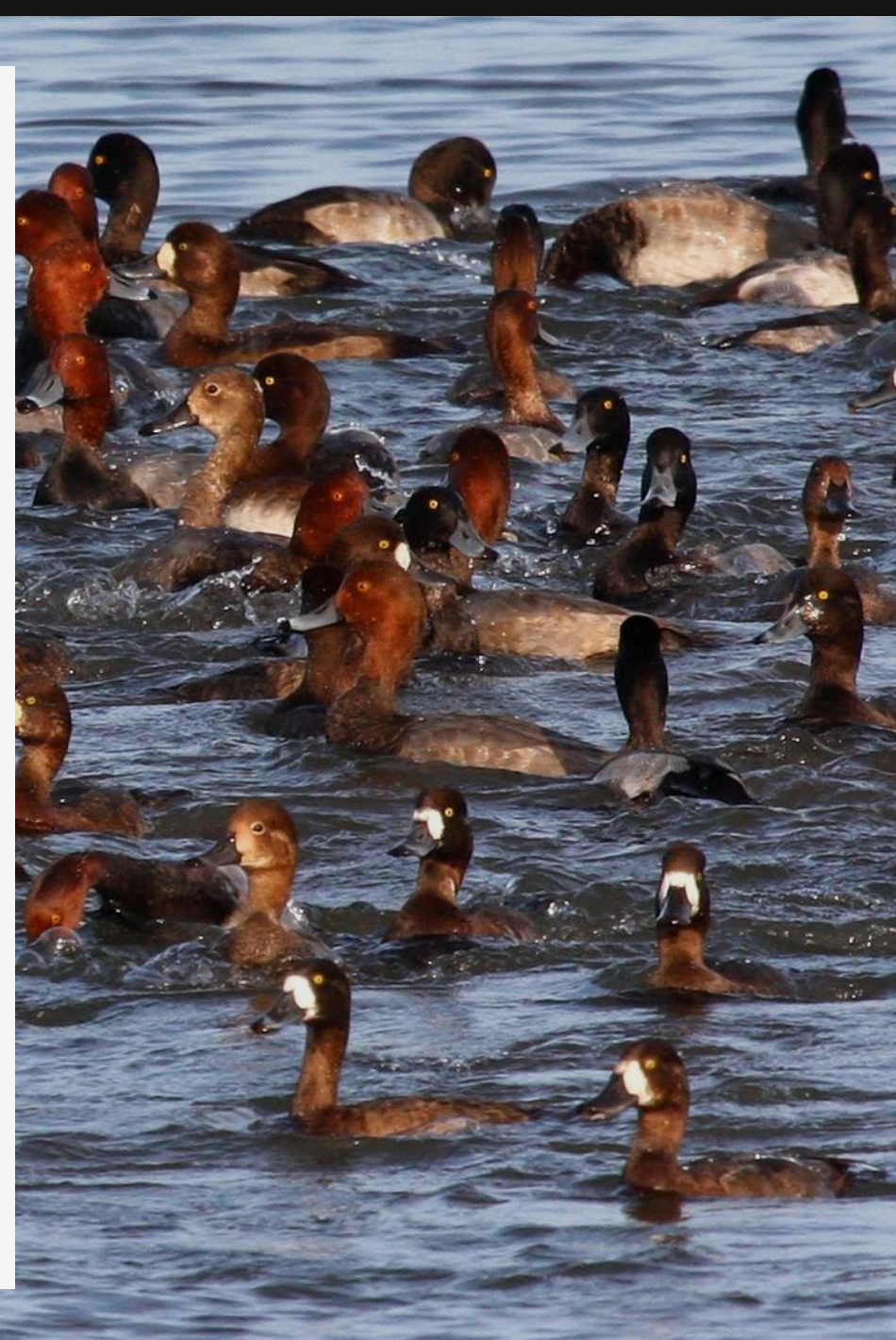
Correspondence Analysis (CA): Detrended Correspondence Analysis (DCA)

Detrending by polynomials: When species are distributed along a “long” environmental gradient, axis 2 is often a quadratic function of axis 1. Thus, a constant is imposed such that axis 2 is uncorrelated with axis 1, etc.

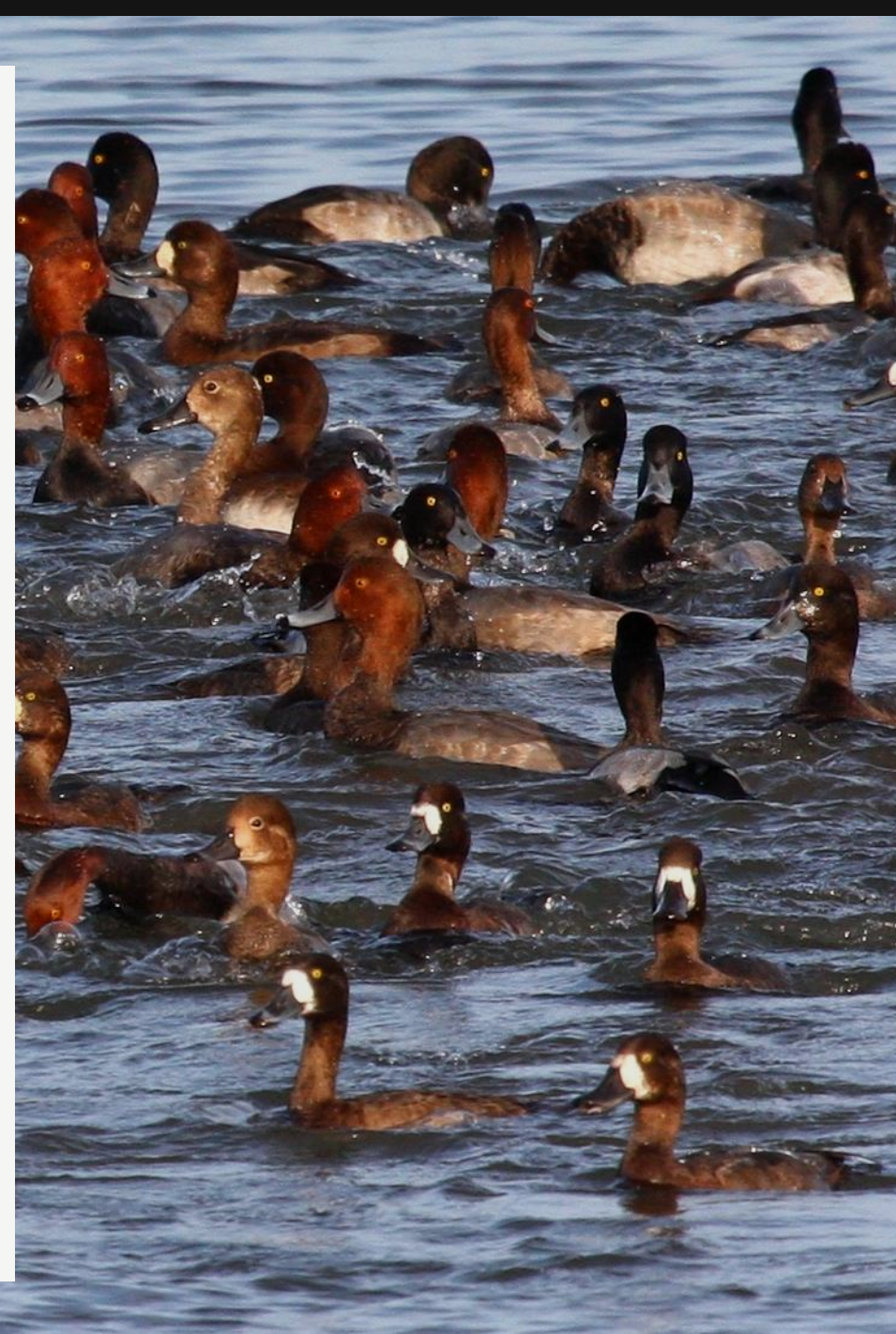


Correspondence Analysis (CA): Detrended Correspondence Analysis (DCA)

Just don't do it. (Use NMDS instead)

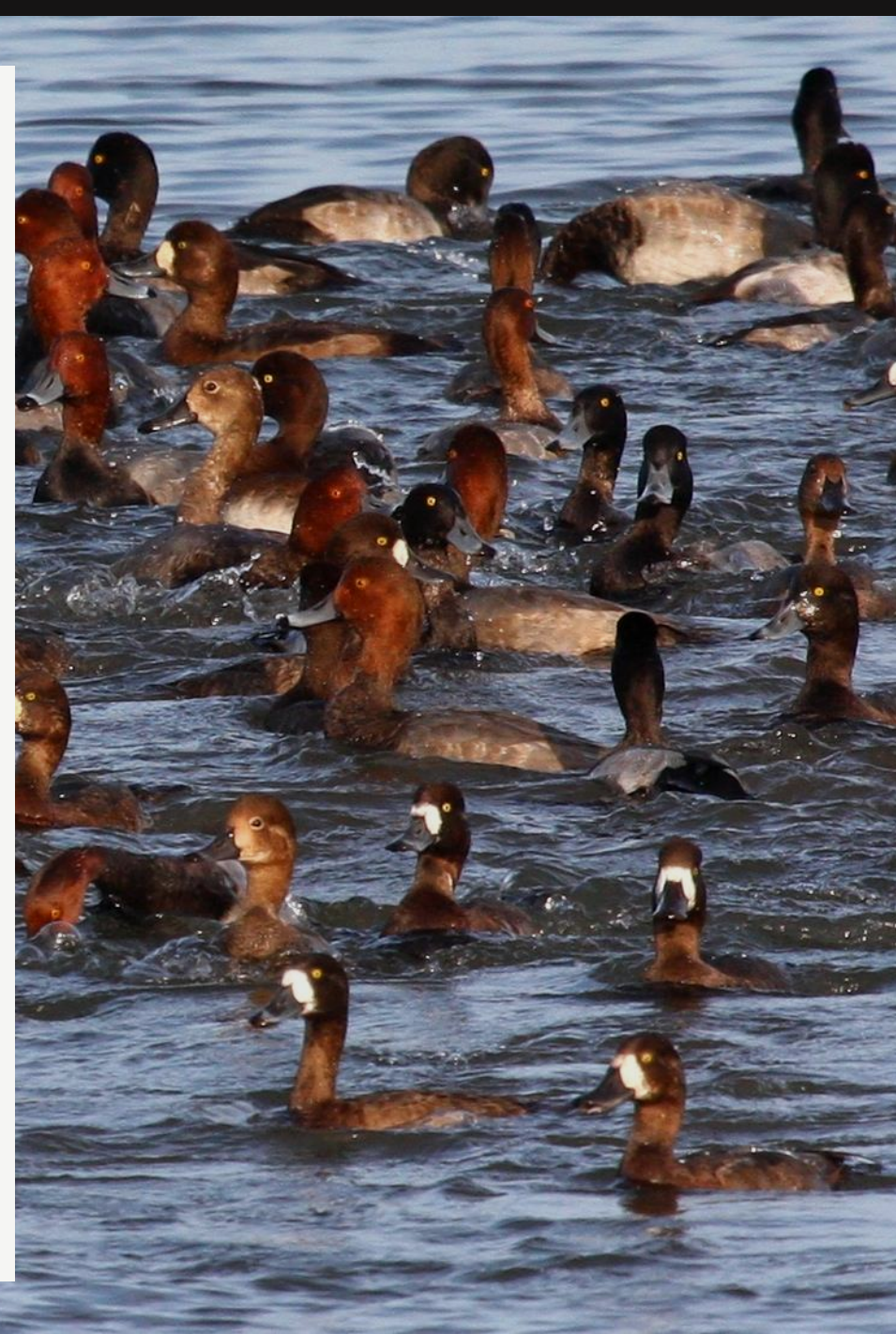


Principal Coordinate Analysis (PCoA)



Principal Coordinate Analysis (PCoA): Introduction

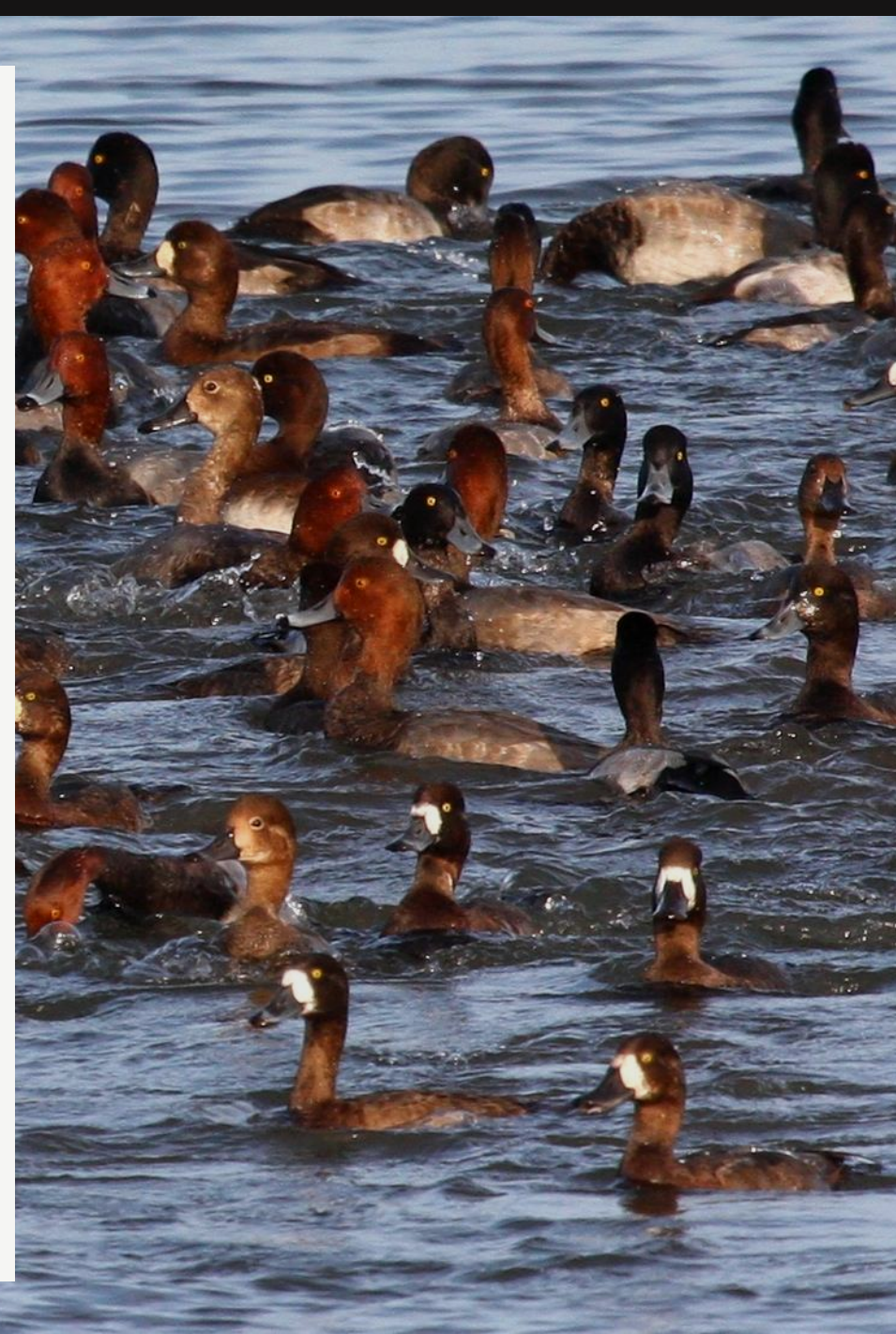
PCA preserves Euclidean distances and CA preserves χ^2 distances. What about other distance relationships?



Principal Coordinate Analysis (PCoA): Introduction

PCA preserves Euclidean distances and CA preserves χ^2 distances. What about other distance relationships?

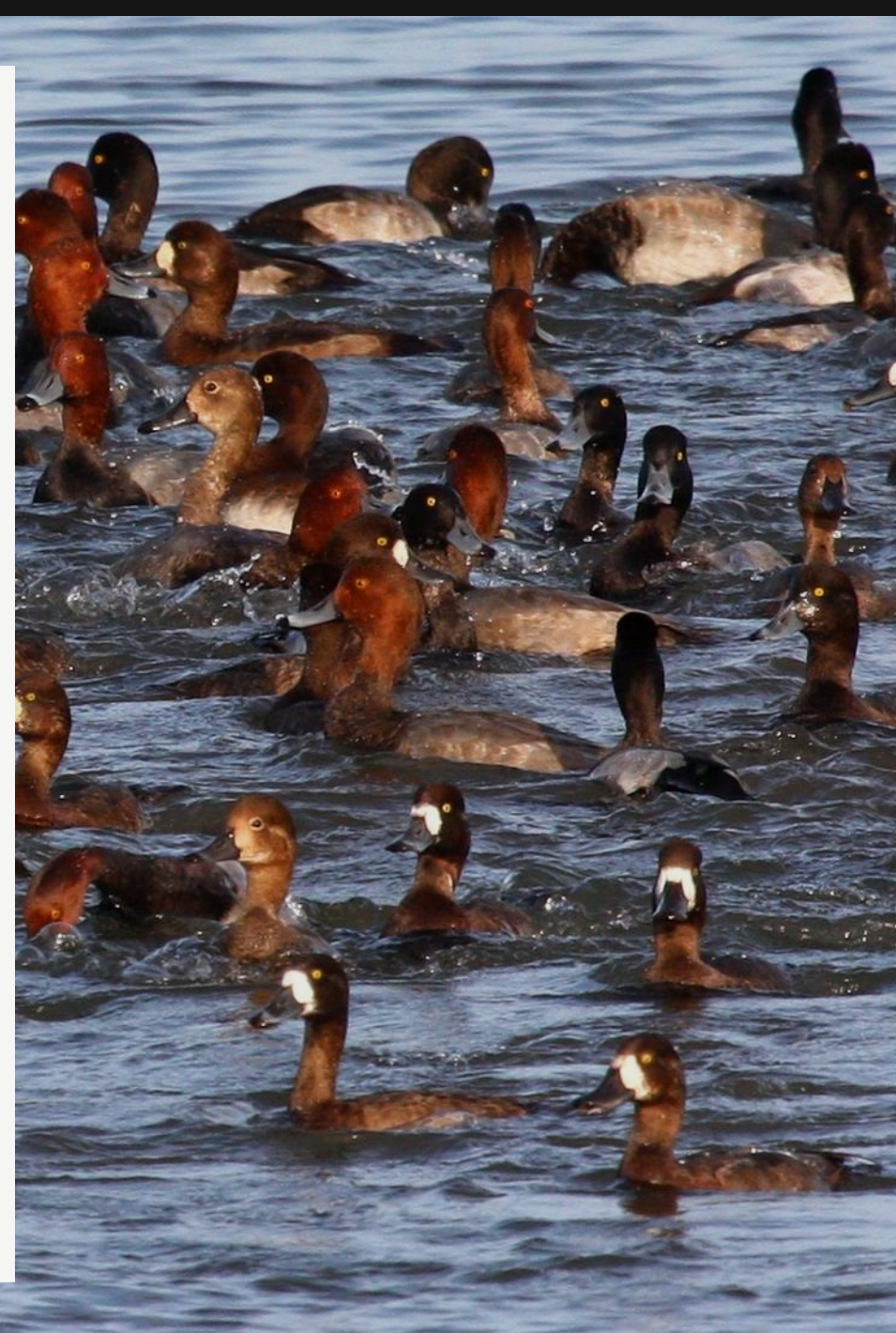
- Binary data
- Mixtures of quantitative and qualitative descriptors
- Pretty much any of the other dissimilarity metrics in Legendre & Legendre Chapter 7...



Principal Coordinate Analysis (PCoA): Introduction

Enter: **Principal Coordinate Analysis** or **Metric Multidimensional Scaling**.

Principal coordinates are mediated through a distance function that has been computed among objects.

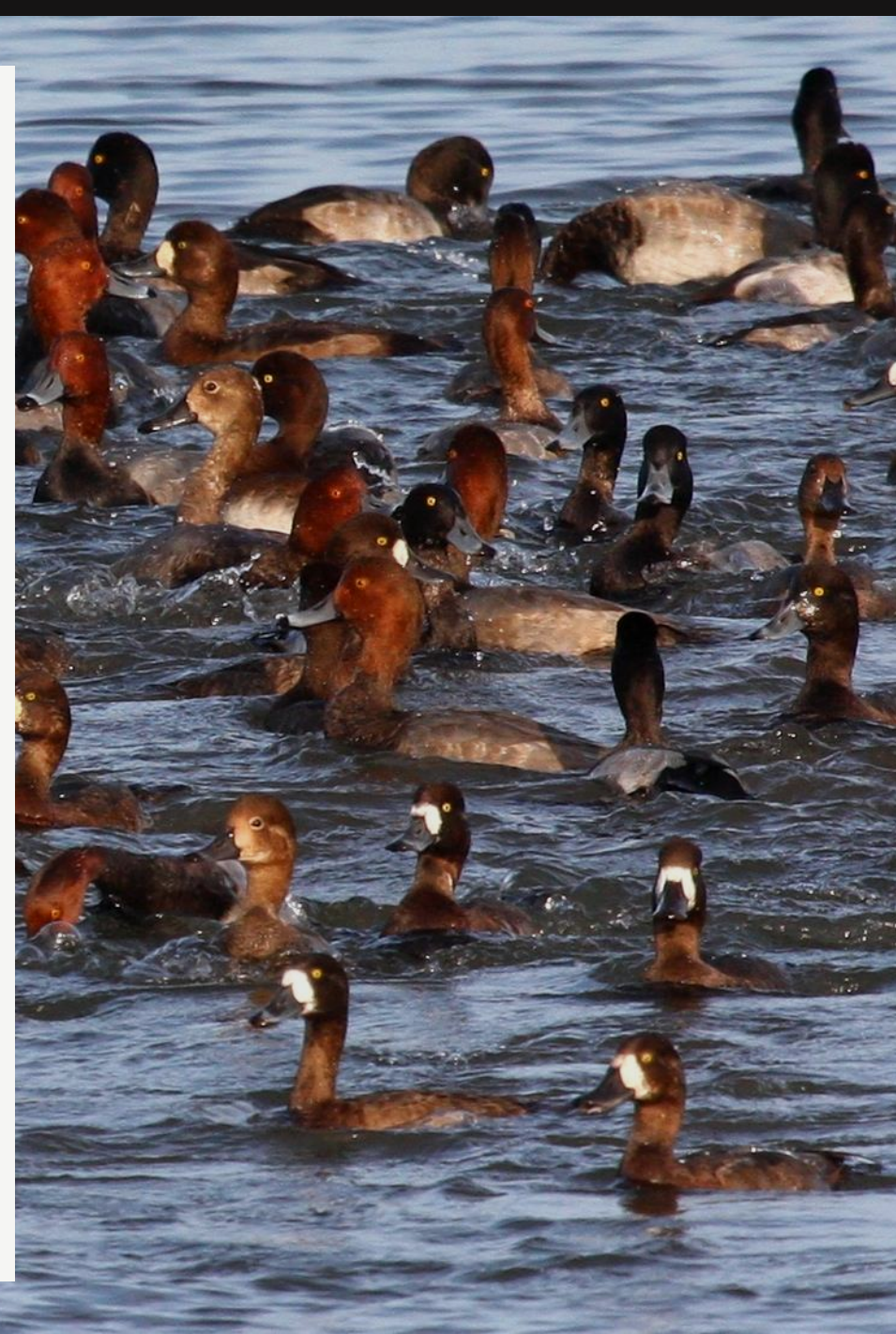


Principal Coordinate Analysis (PCoA): Introduction

Enter: **Principal Coordinate Analysis** or **Metric Multidimensional Scaling**.

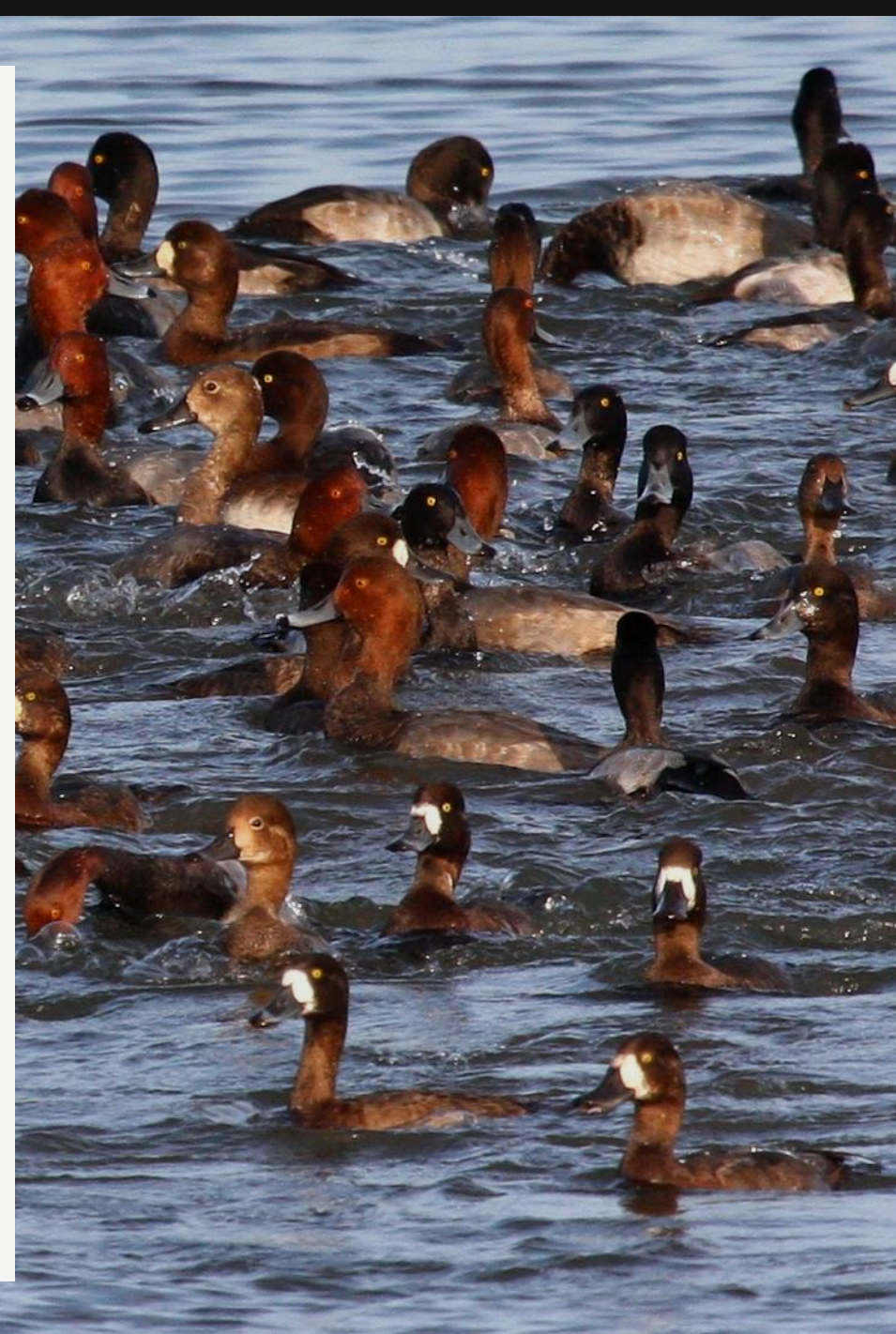
Principal coordinates are mediated through a distance function that has been computed among objects.

Improved performance if distances are **metric** (i.e., do not violate the “triangle inequality”)



Principal Coordinate Analysis (PCoA): Steps

Step 1: Calculate a dissimilarity/distance matrix



Principal Coordinate Analysis (PCoA): Steps

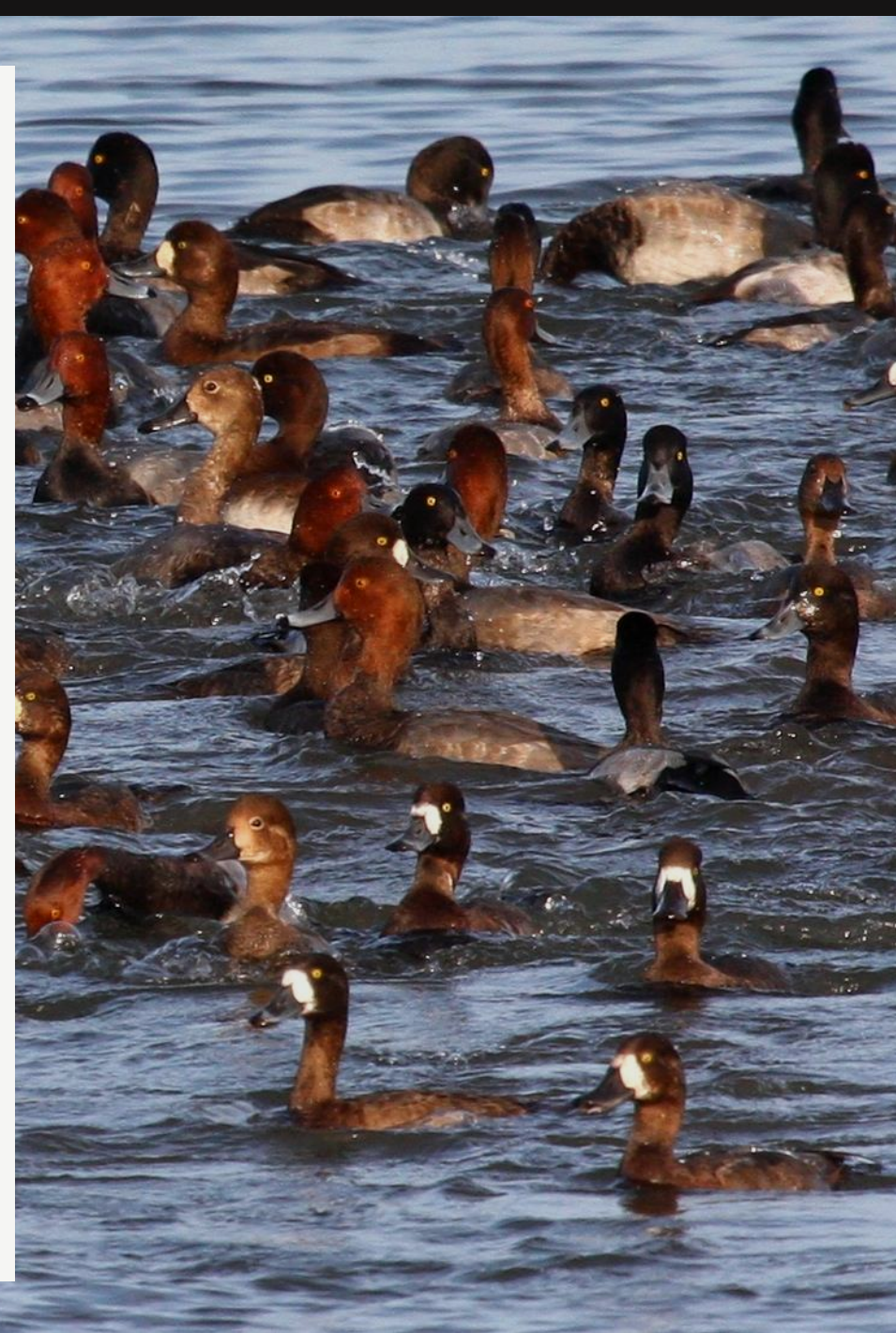
Step 1: Calculate a dissimilarity/distance matrix

Binary: Jaccard or Sørensen coefficients

Environmental: Euclidean or Manhattan distances (after standardization)

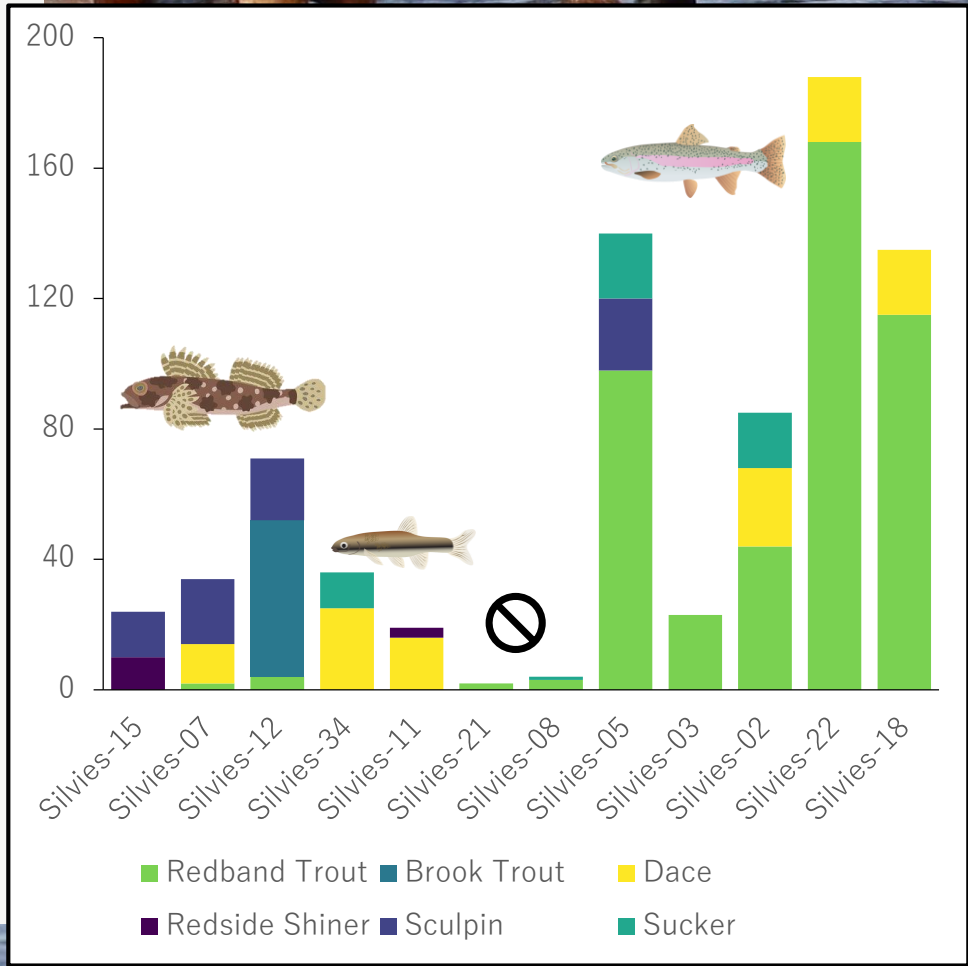
Species abundance: Chi-square distance or Percentage Difference/Bray-Curtis dissimilarity

Mixed data types: Gower's coefficient



Principal Coordinate Analysis (PCoA): Steps

	S-15	S-07	S-12	S-34	S-11	S-21	S-08	S-05	S-03	S-02	S-22
Silvies-07	0.52										
Silvies-12	0.73	0.72									
Silvies-34	1.00	0.62	1.00								
Silvies-11	0.88	0.55	1.00	0.54							
Silvies-21	1.00	0.87	0.88	1.00	1.00						
Silvies-08	1.00	0.88	0.89	0.93	1.00	0.33					
Silvies-05	0.75	0.72	0.76	0.80	1.00	0.90	0.81				
Silvies-03	1.00	0.91	0.93	1.00	1.00	0.80	0.73	0.36			
Silvies-02	1.00	0.64	0.95	0.50	0.58	0.86	0.74	0.42	0.41		
Silvies-22	1.00	0.76	0.96	0.81	0.75	0.90	0.86	0.38	0.38	0.41	
Silvies-18	1.00	0.73	0.95	0.79	0.70	0.88	0.83	0.31	0.29	0.34	0.10

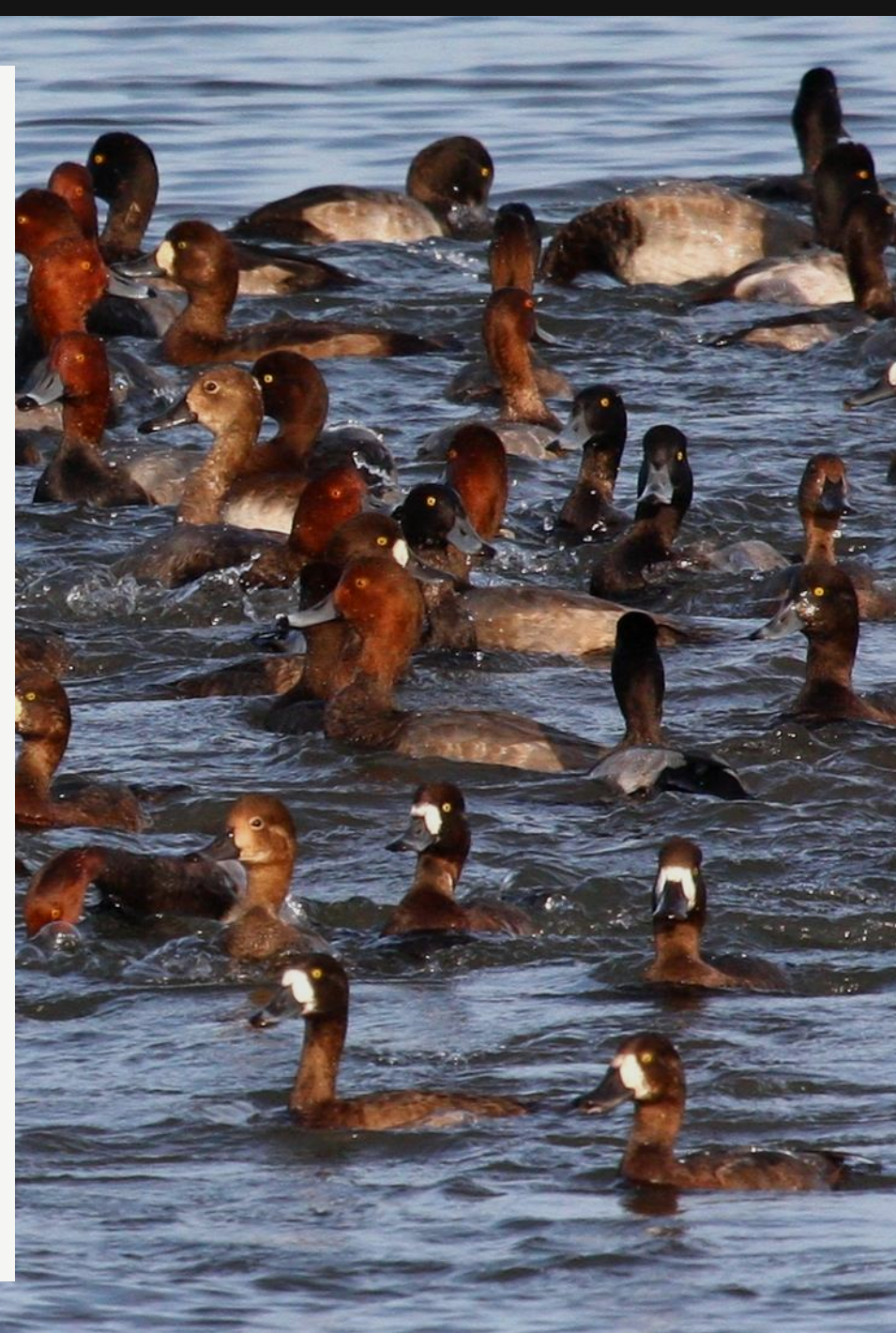


Principal Coordinate Analysis (PCoA): Steps

Step 2: Transform using equation:

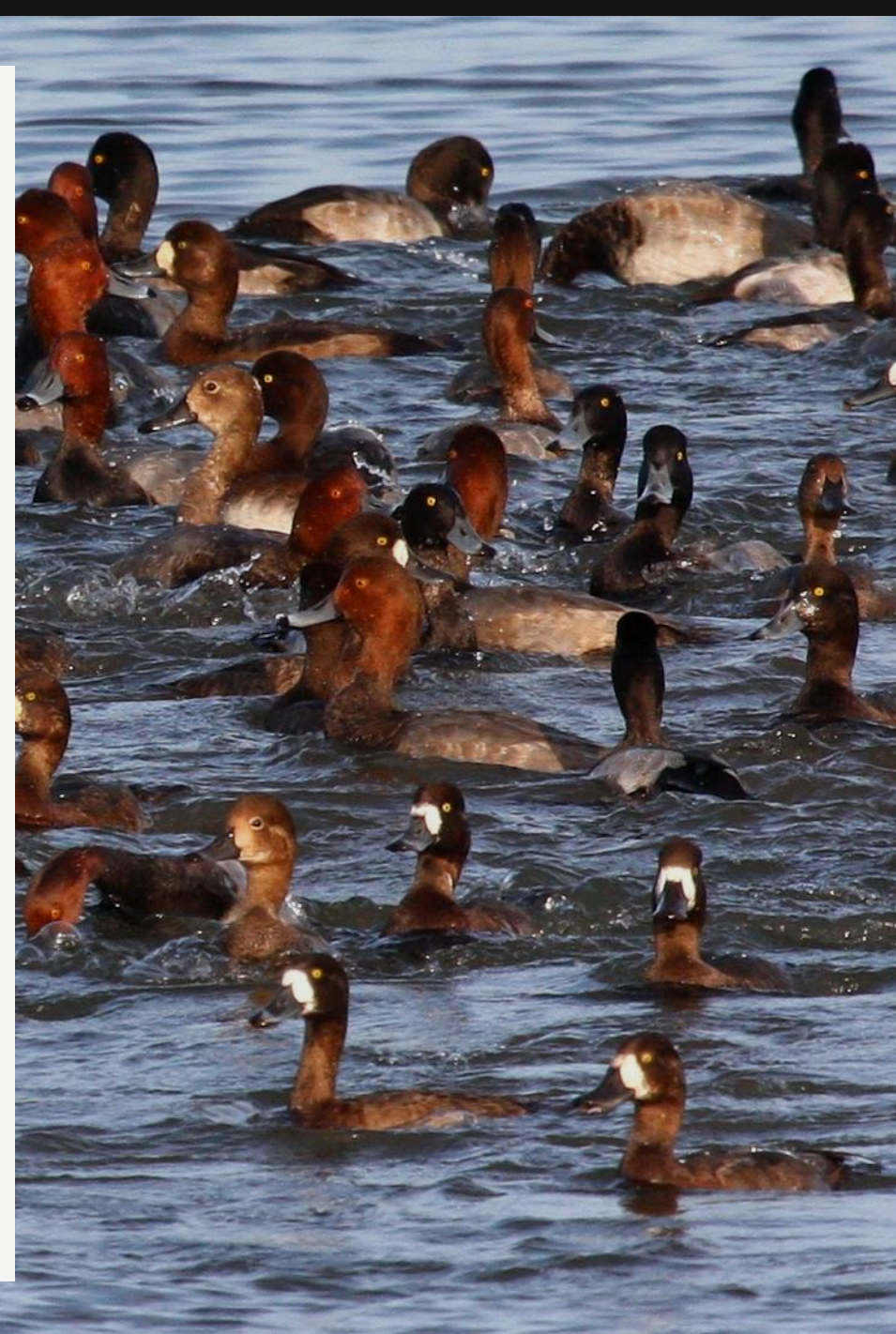
$$A = -1/2 * D^2$$

then double-center A by rows and columns



Principal Coordinate Analysis (PCoA): Steps

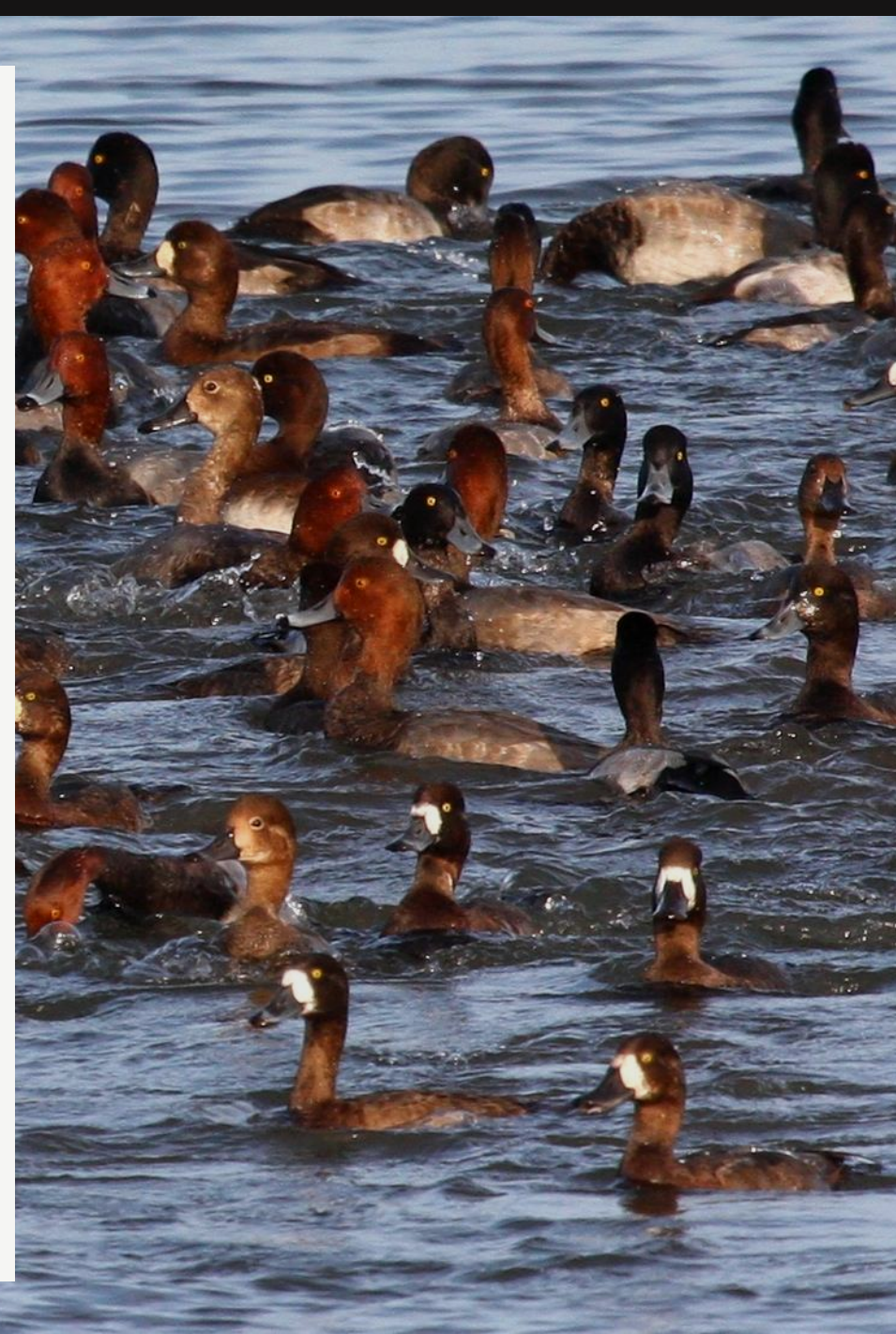
Step 3: Decompose centered dissimilarity matrix to obtain eigenvalues and eigenvectors.



Principal Coordinate Analysis (PCoA): Steps

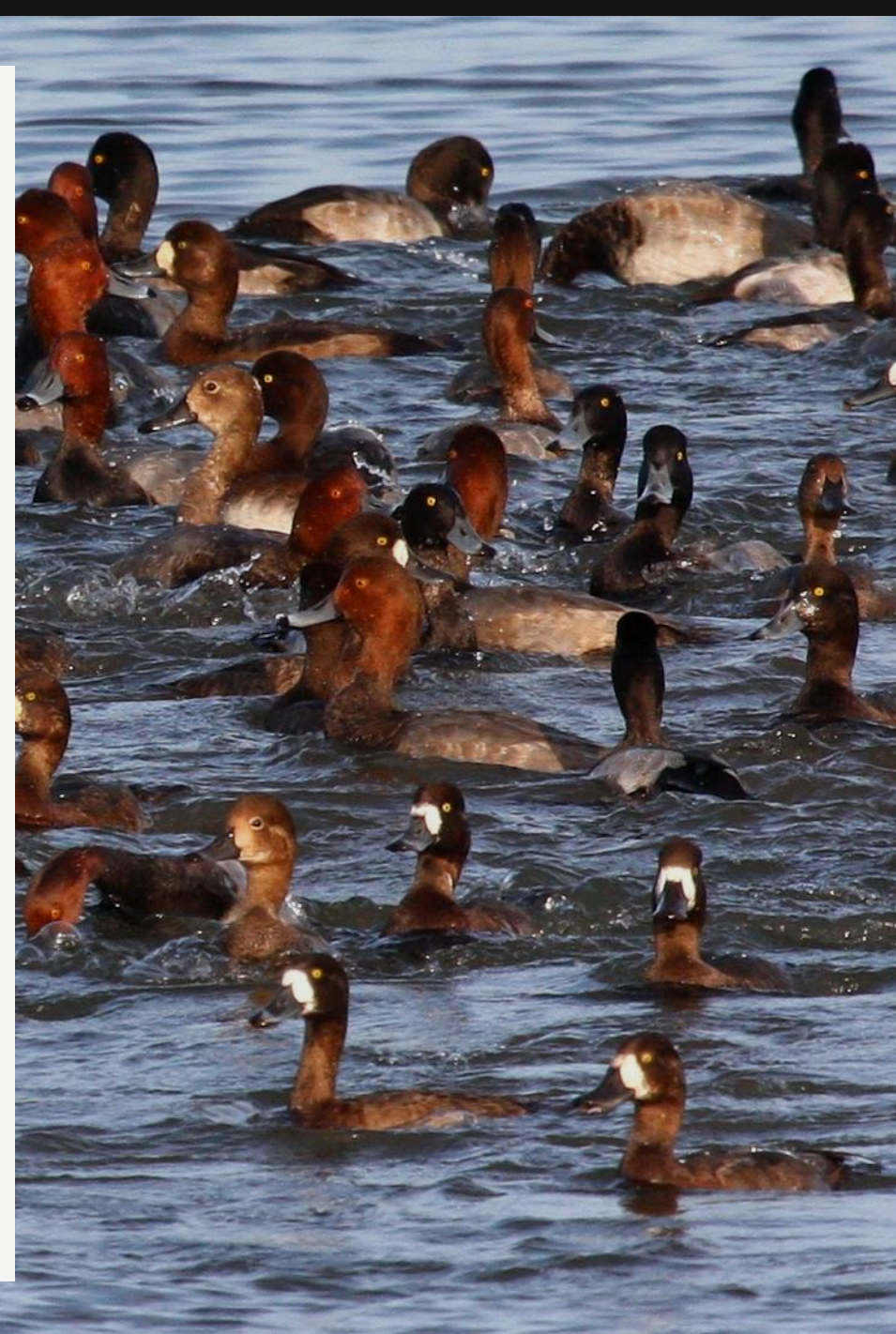
Step 3: Decompose centered dissimilarity matrix to obtain eigenvalues and eigenvectors.

Pretty much the same process as for PCA and CA. **Eigenvalues** correspond to explained variance and **eigenvectors** represent the coordinates of the objects in reduced-dimensional space.



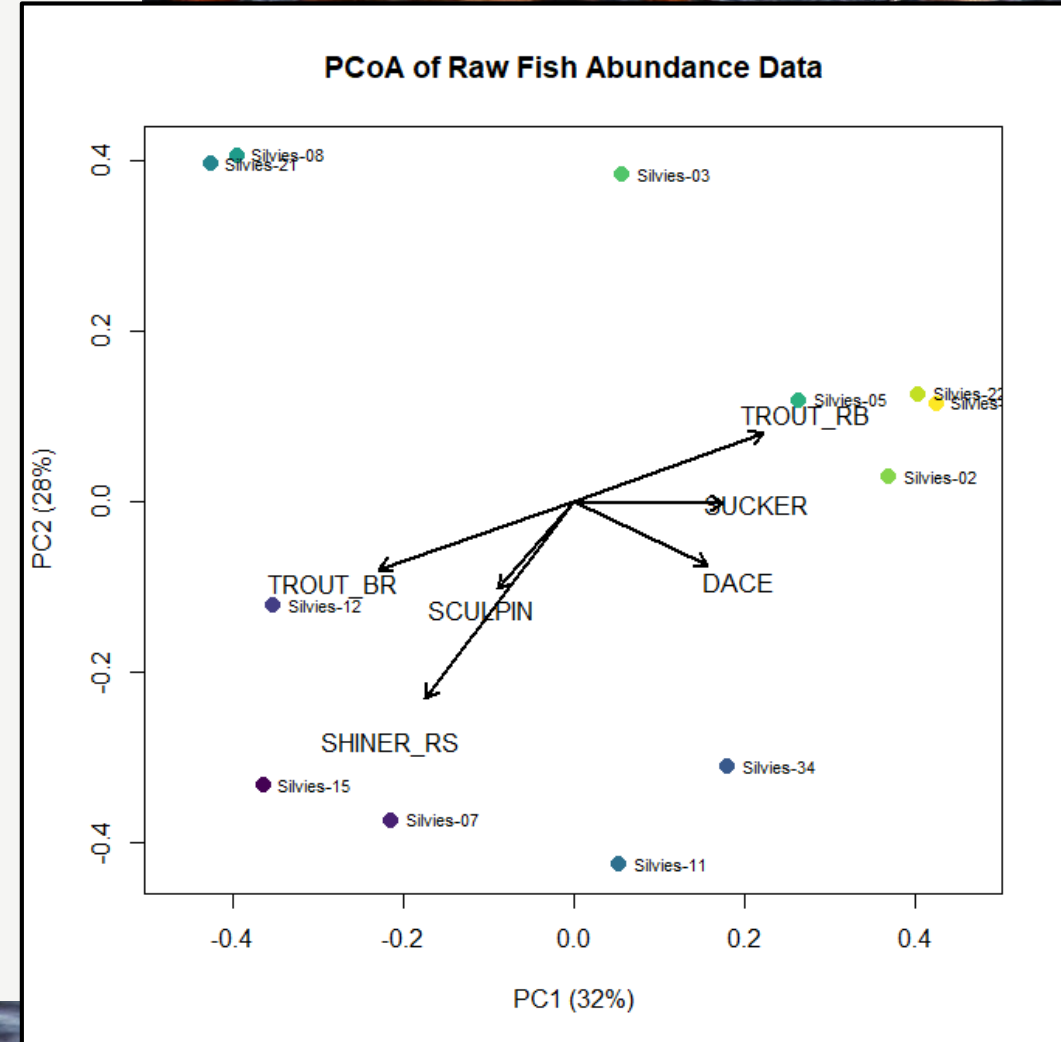
Principal Coordinate Analysis (PCoA): Steps

Step 4: Scale each eigenvector by multiplying by the square root of its eigenvalue.



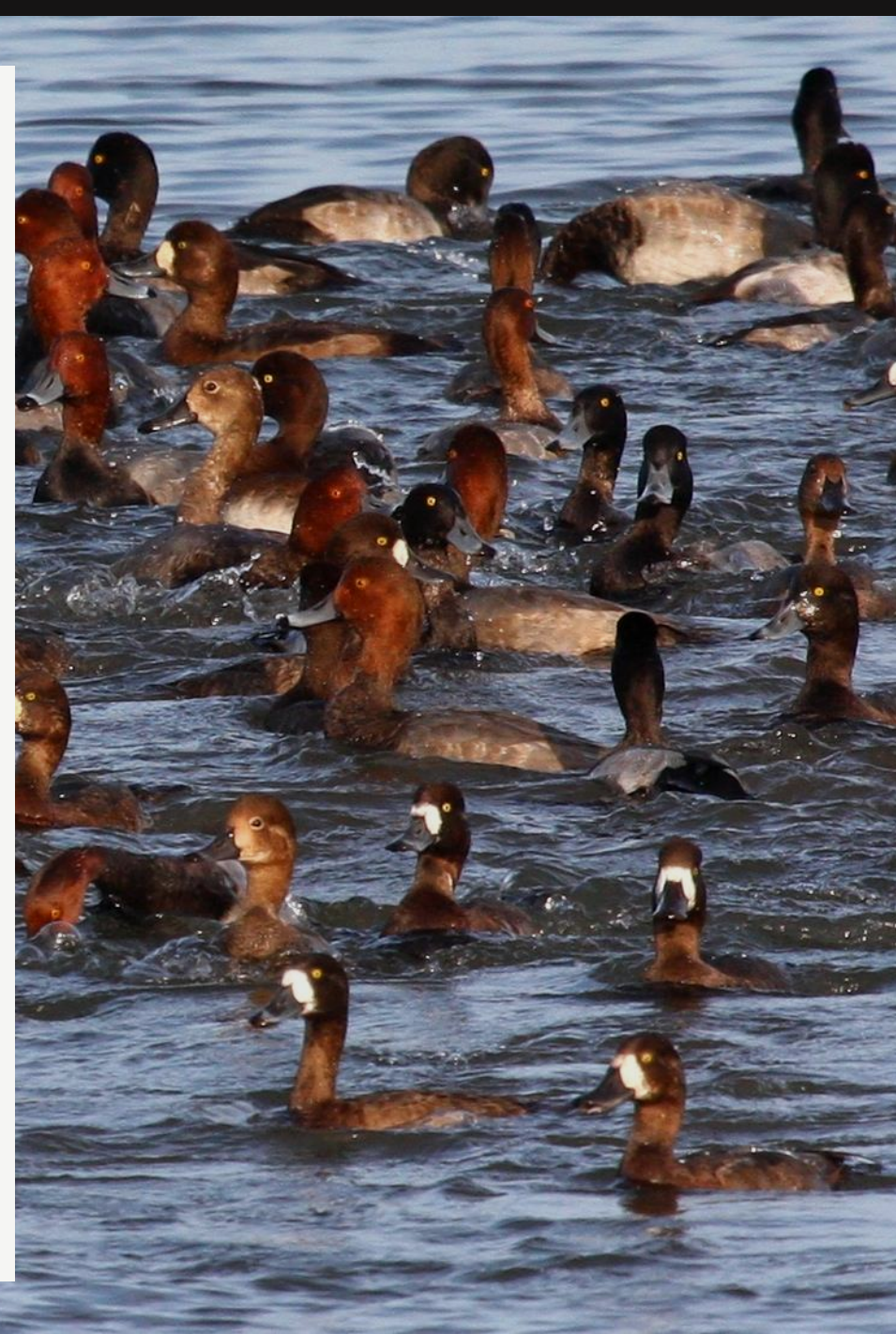
Principal Coordinate Analysis (PCoA): Steps

Step 4: Scale each eigenvector by multiplying by the square root of its eigenvalue.



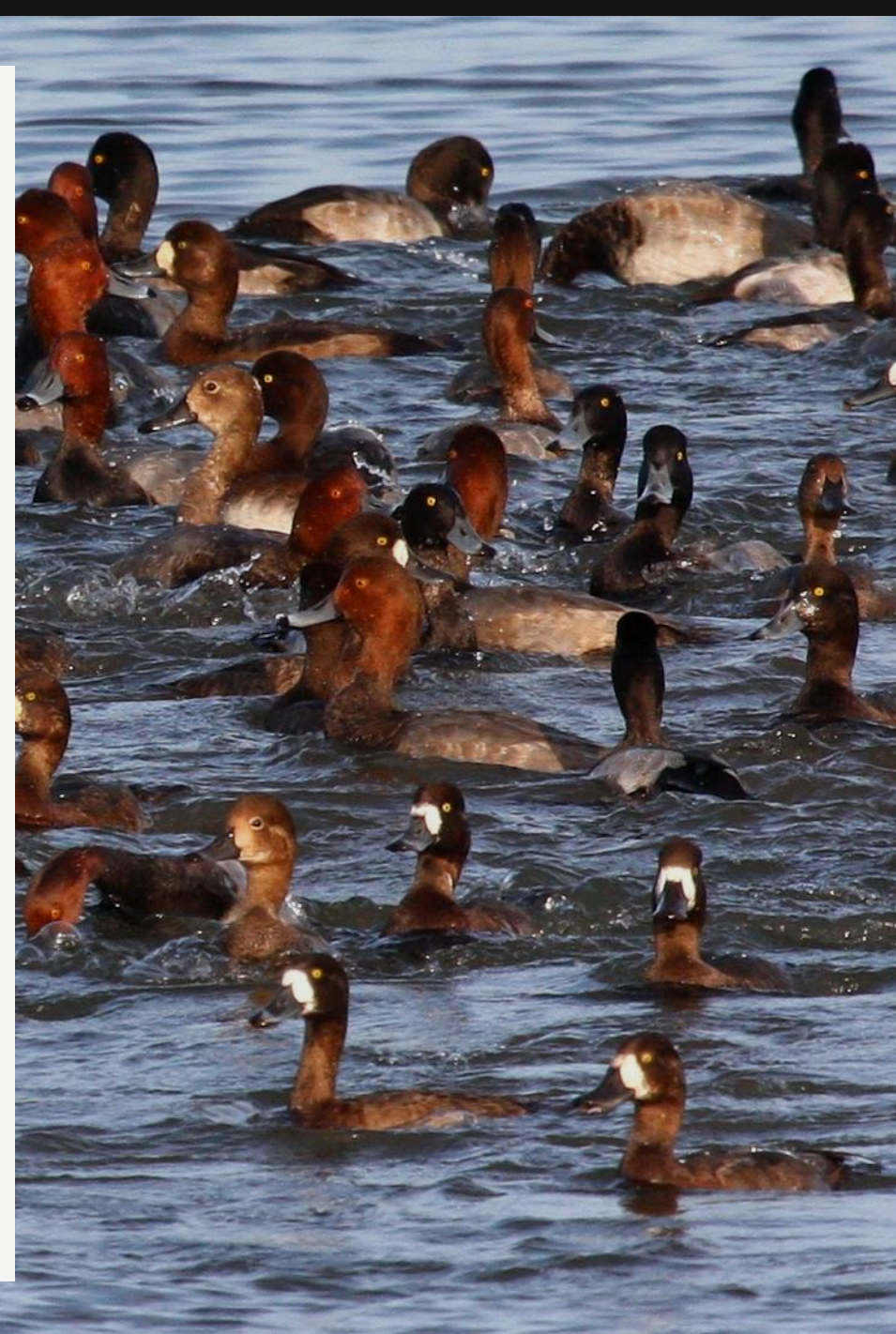
Principal Coordinate Analysis (PCoA): Limitations

- Outcome is highly dependent on the distance measure



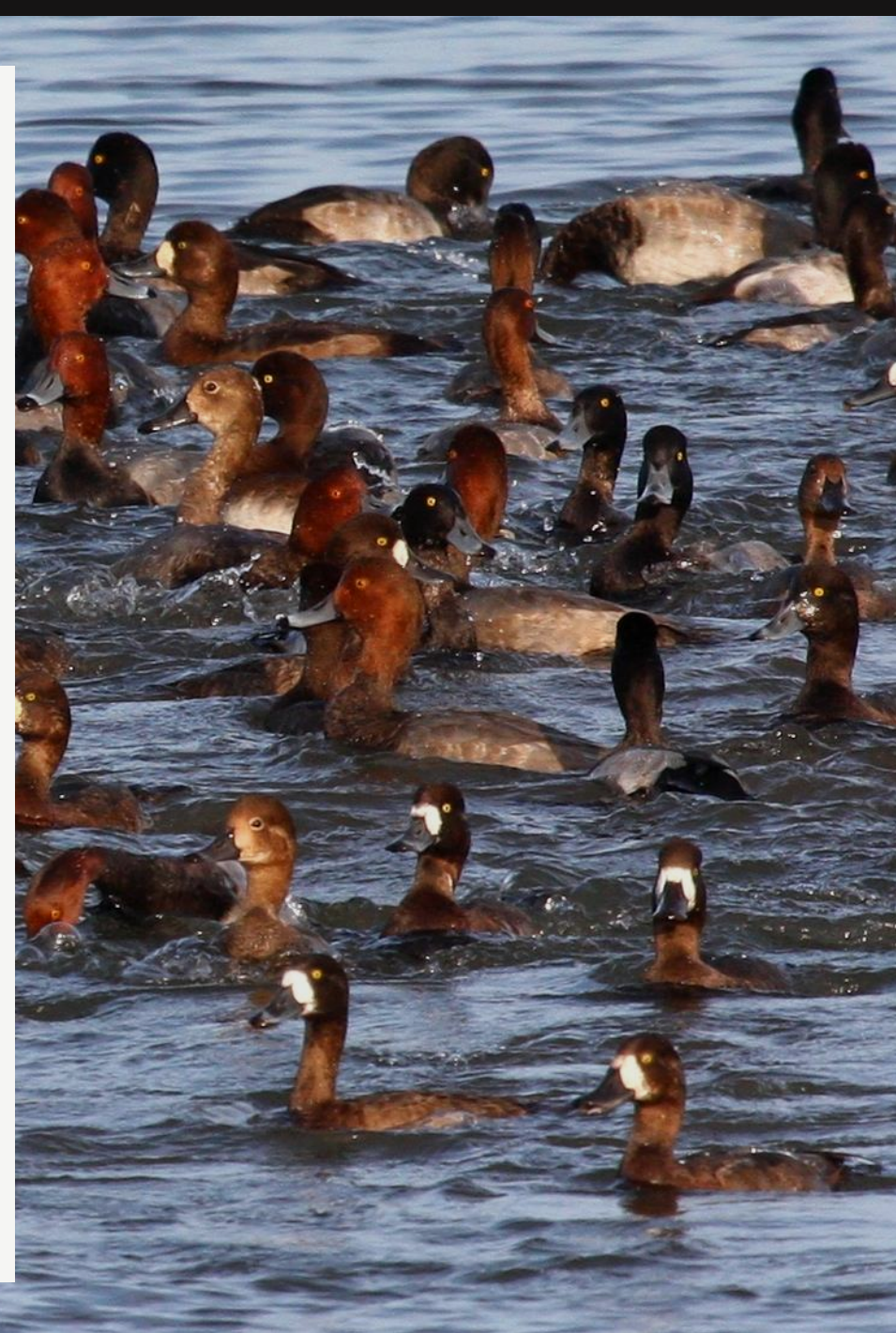
Principal Coordinate Analysis (PCoA): Limitations

- Outcome is highly dependent on the distance measure
- Produces of negative eigenvalues when the distance used is semi-metric



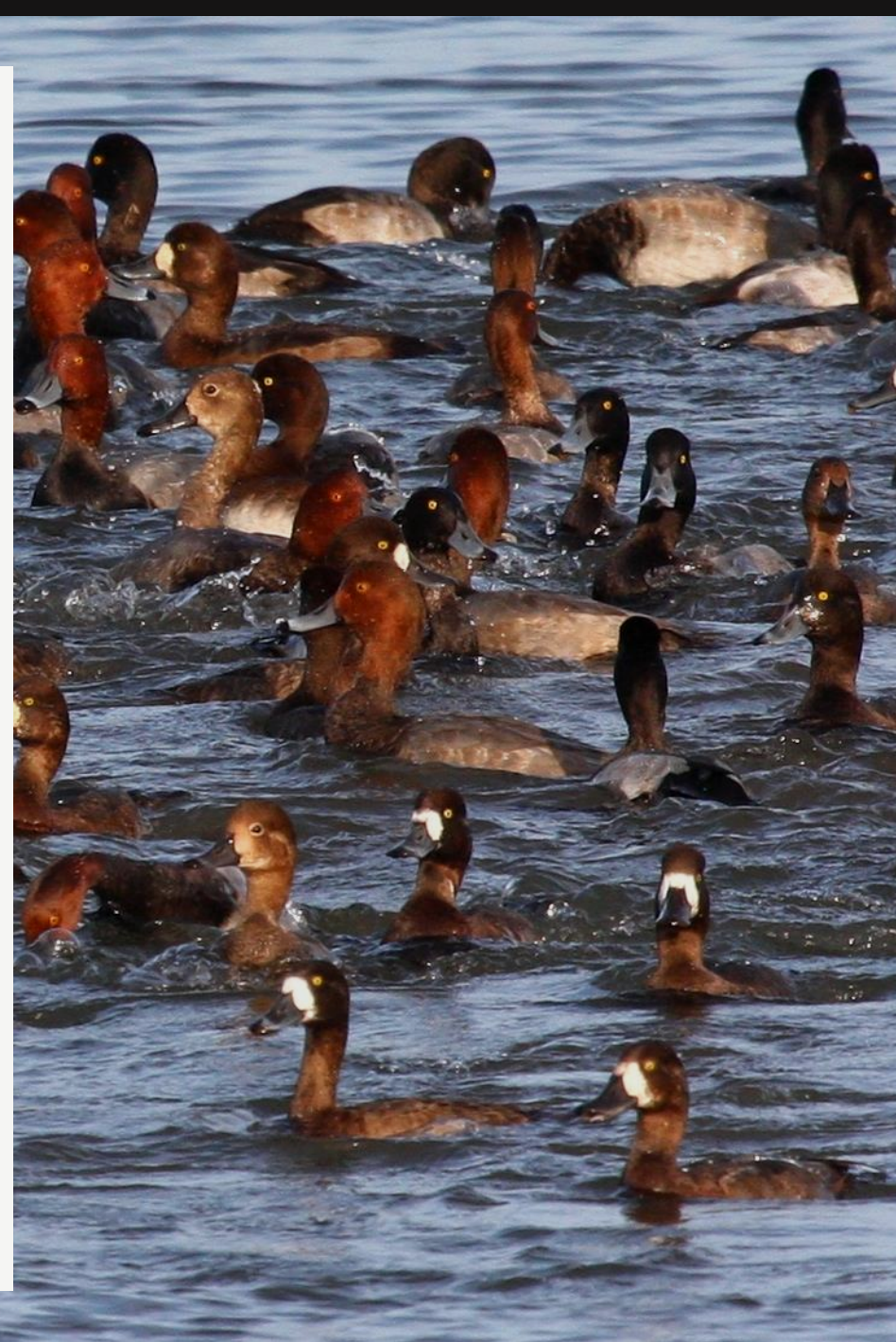
Principal Coordinate Analysis (PCoA): Limitations

- Outcome is highly dependent on the distance measure
- Produces of negative eigenvalues when the distance used is semi-metric
- Does not preserve patterns in the original data structure



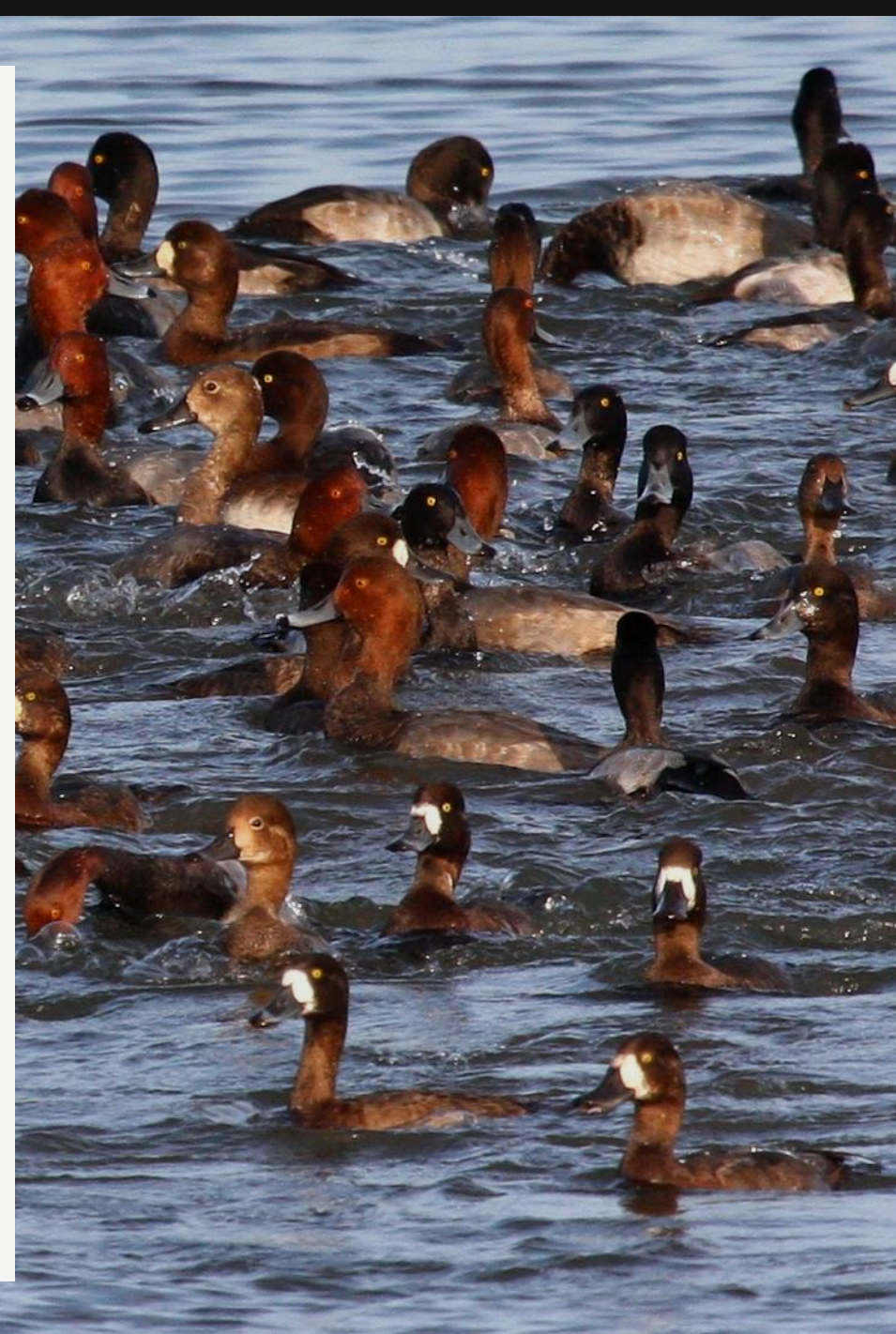
Principal Coordinate Analysis (PCoA): Limitations

- Outcome is highly dependent on the distance measure
- Produces of negative eigenvalues when the distance used is semi-metric
- Does not preserve patterns in the original data structure
- Sensitive to outliers

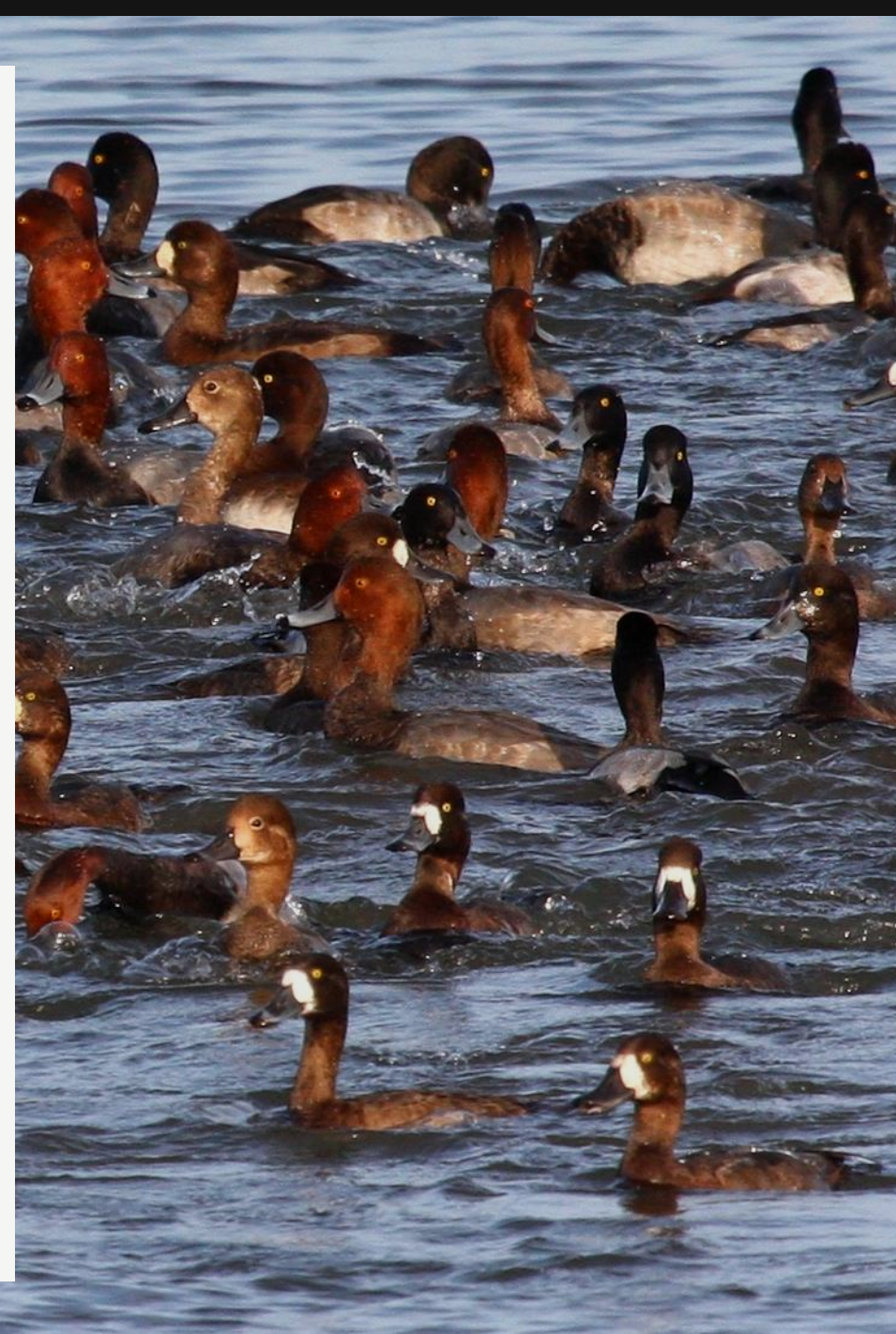


Principal Coordinate Analysis (PCoA): Limitations

- Outcome is highly dependent on the distance measure
- Produces of negative eigenvalues when the distance used is semi-metric
- Does not preserve patterns in the original data structure
- Sensitive to outliers
- Reduction in dimensionality may result in arch effect

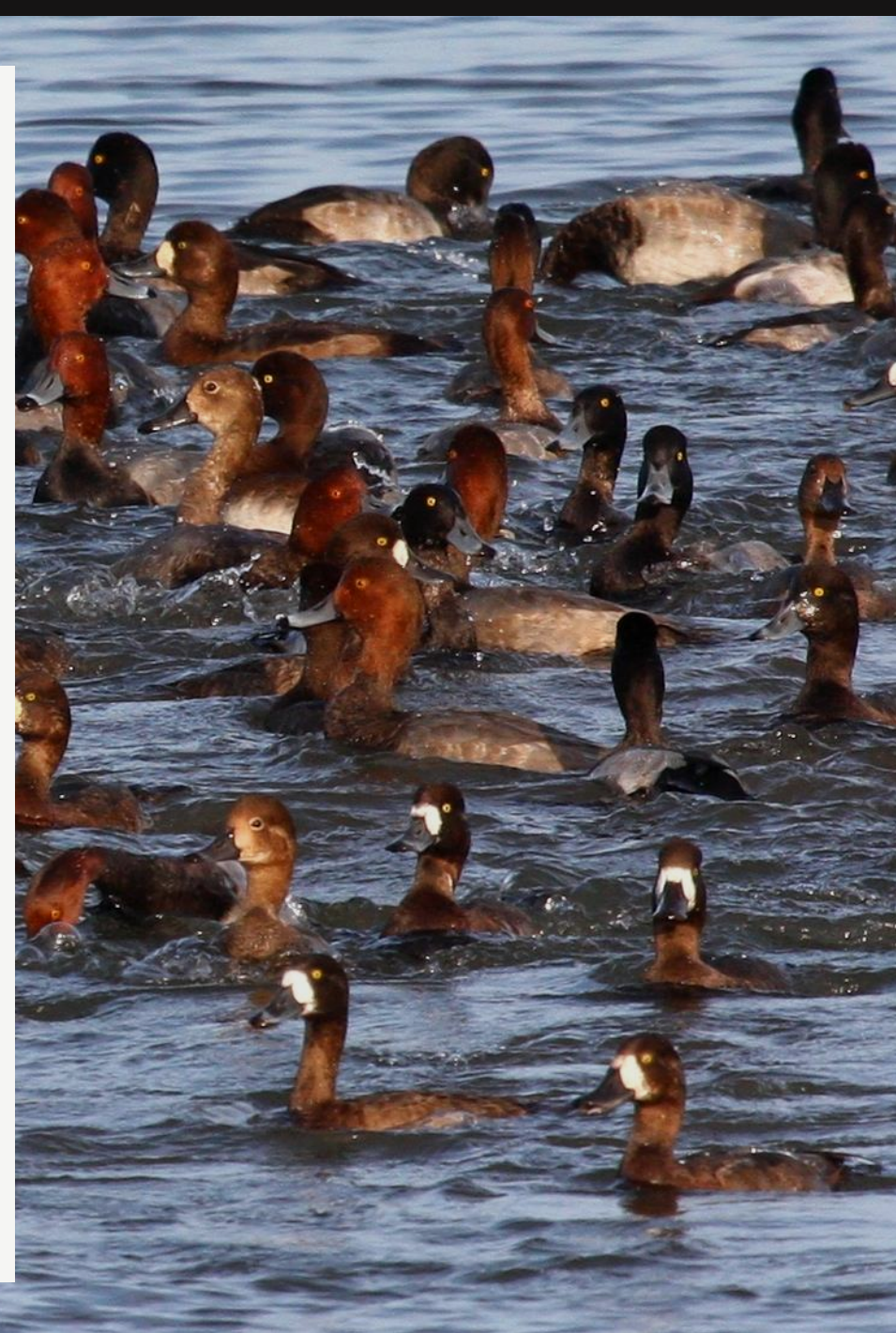


Nonmetric Multidimensional Scaling (NMDS)



Nonmetric Multidimensional Scaling (NMDS): Introduction

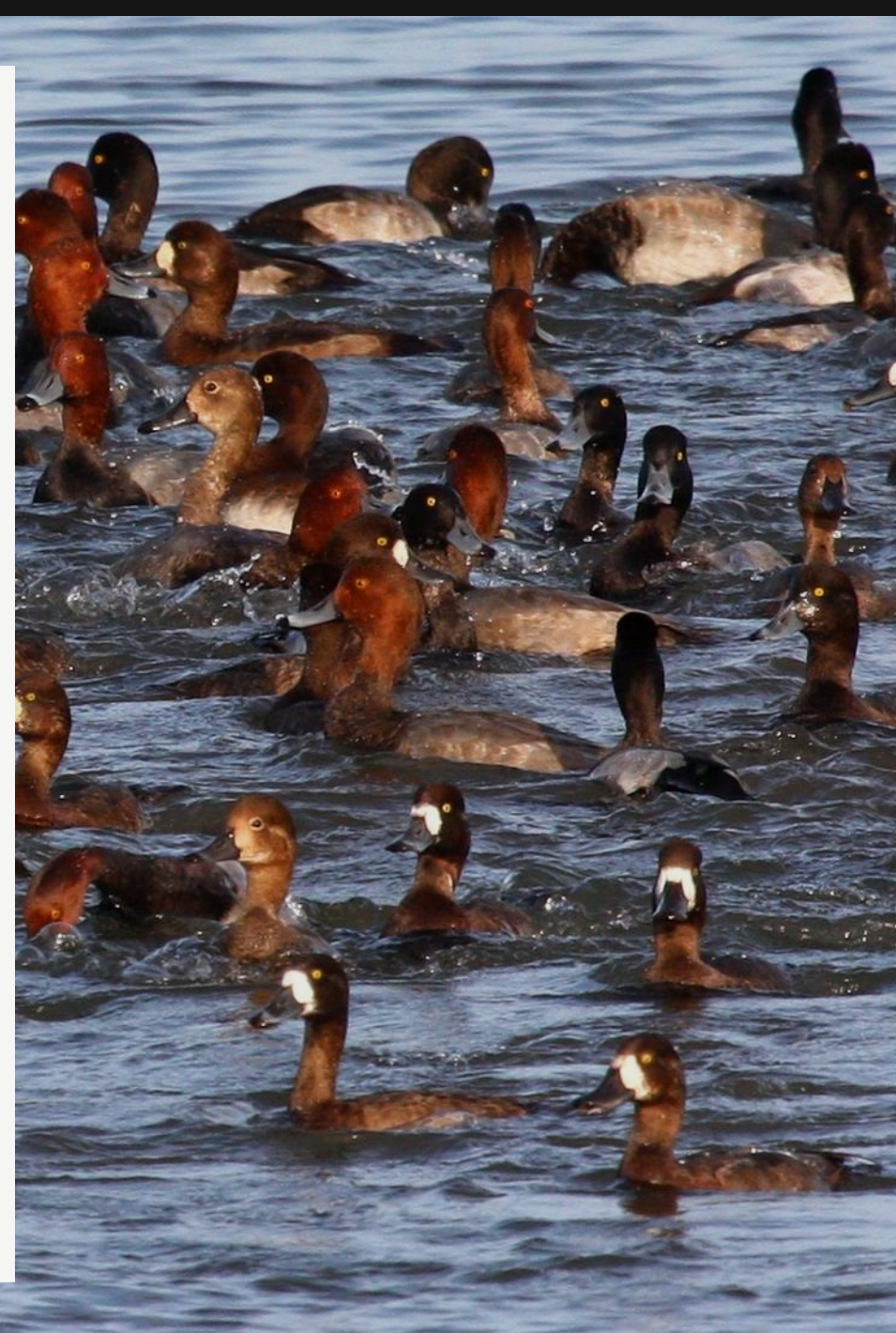
Nonmetric Multidimensional Scaling preserves rank order dissimilarities between objects in a low-dimensional space.



Nonmetric Multidimensional Scaling (NMDS): Introduction

Nonmetric Multidimensional Scaling preserves rank order dissimilarities between objects in a low-dimensional space.

Useful for examining *relative*, rather than *absolute*, distances among data points.

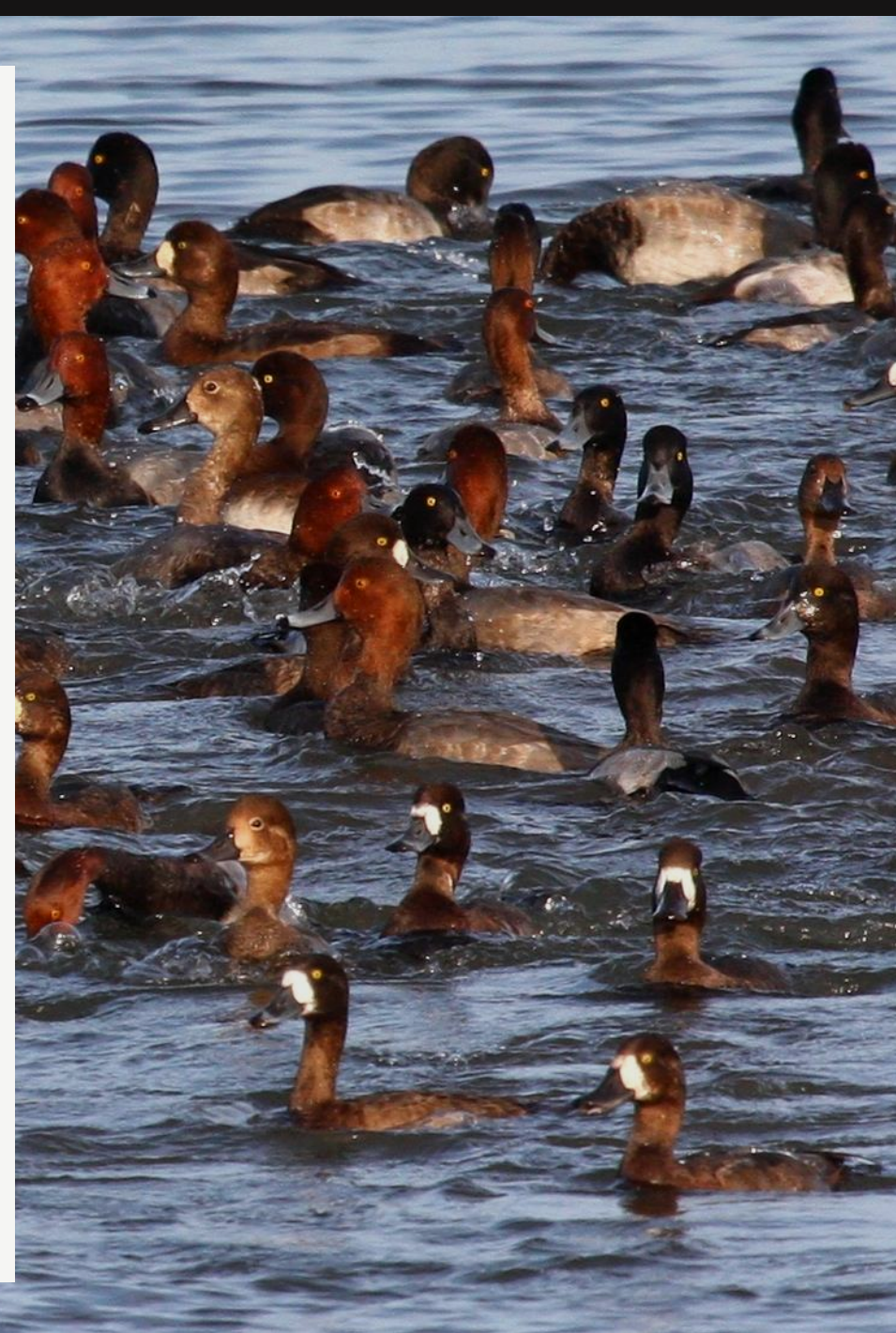


Nonmetric Multidimensional Scaling (NMDS): Introduction

Nonmetric Multidimensional Scaling preserves rank order dissimilarities between objects in a low-dimensional space.

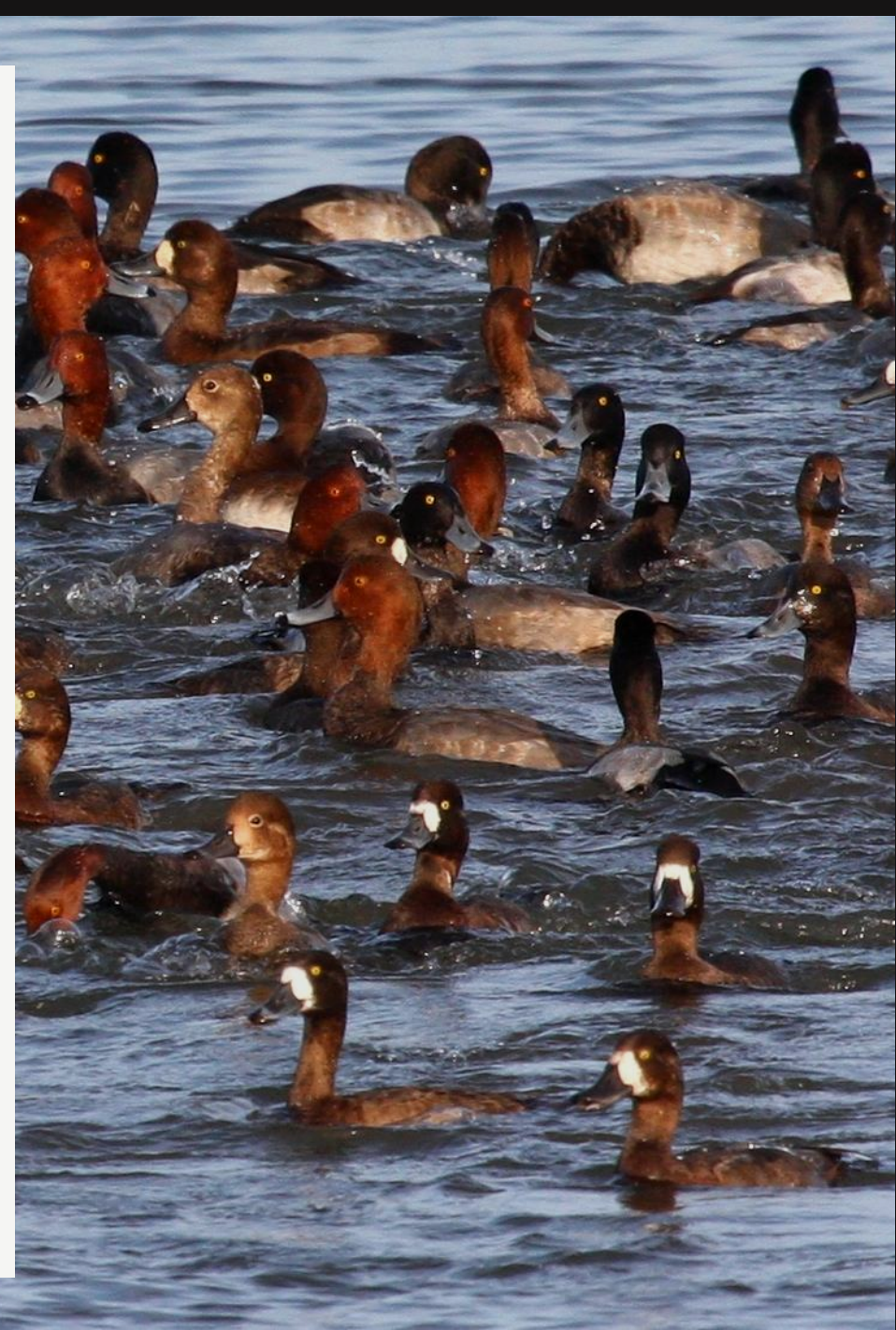
Useful for examining *relative*, rather than *absolute*, distances among data points.

NMDS axes *do not* maximize variability in space and are *arbitrary*!



Nonmetric Multidimensional Scaling (NMDS): Steps

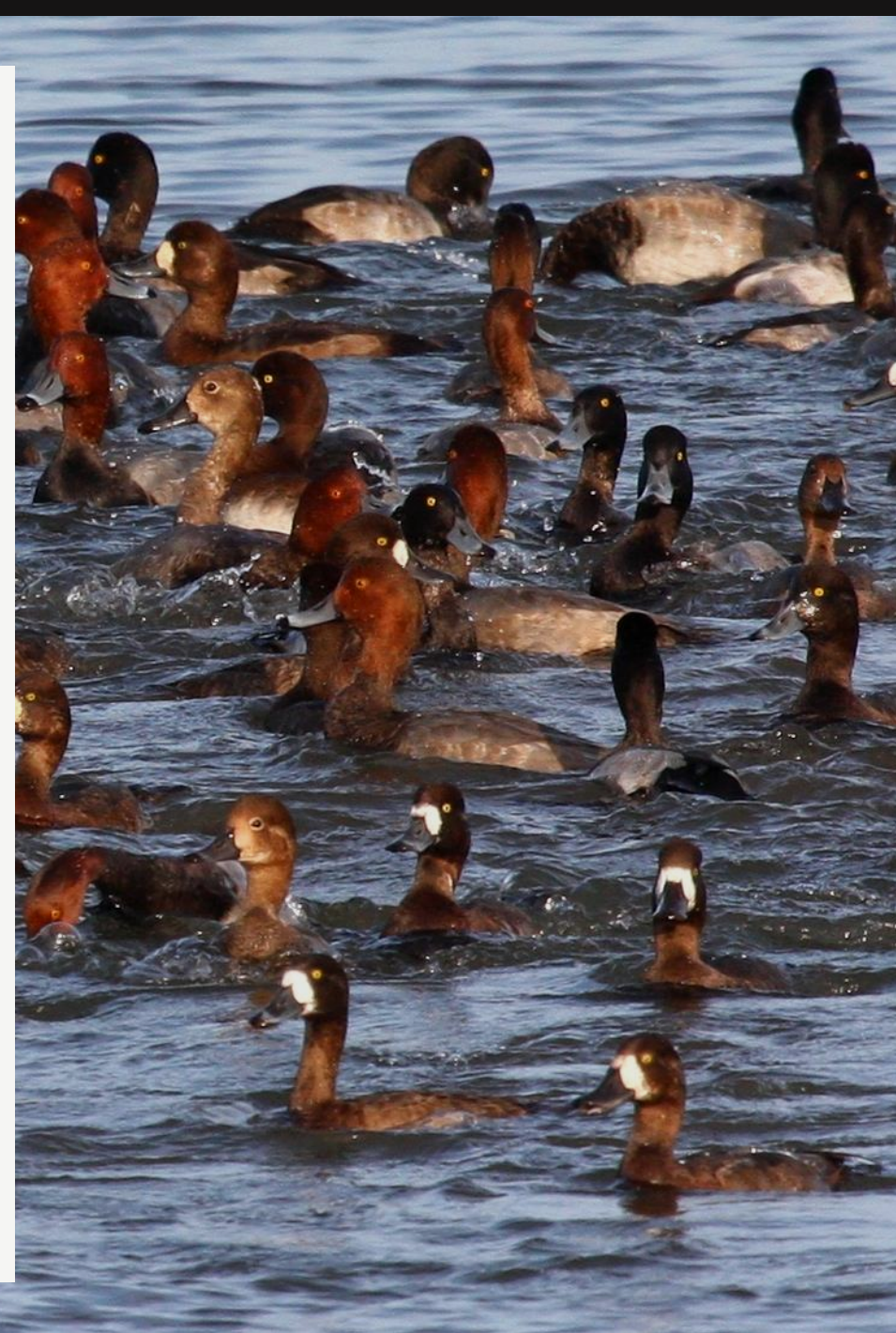
Step 1: Calculate distance/dissimilarity matrix



Nonmetric Multidimensional Scaling (NMDS): Steps

Step 1: Calculate distance/dissimilarity matrix

Step 2: Choose number of dimensions *a priori*

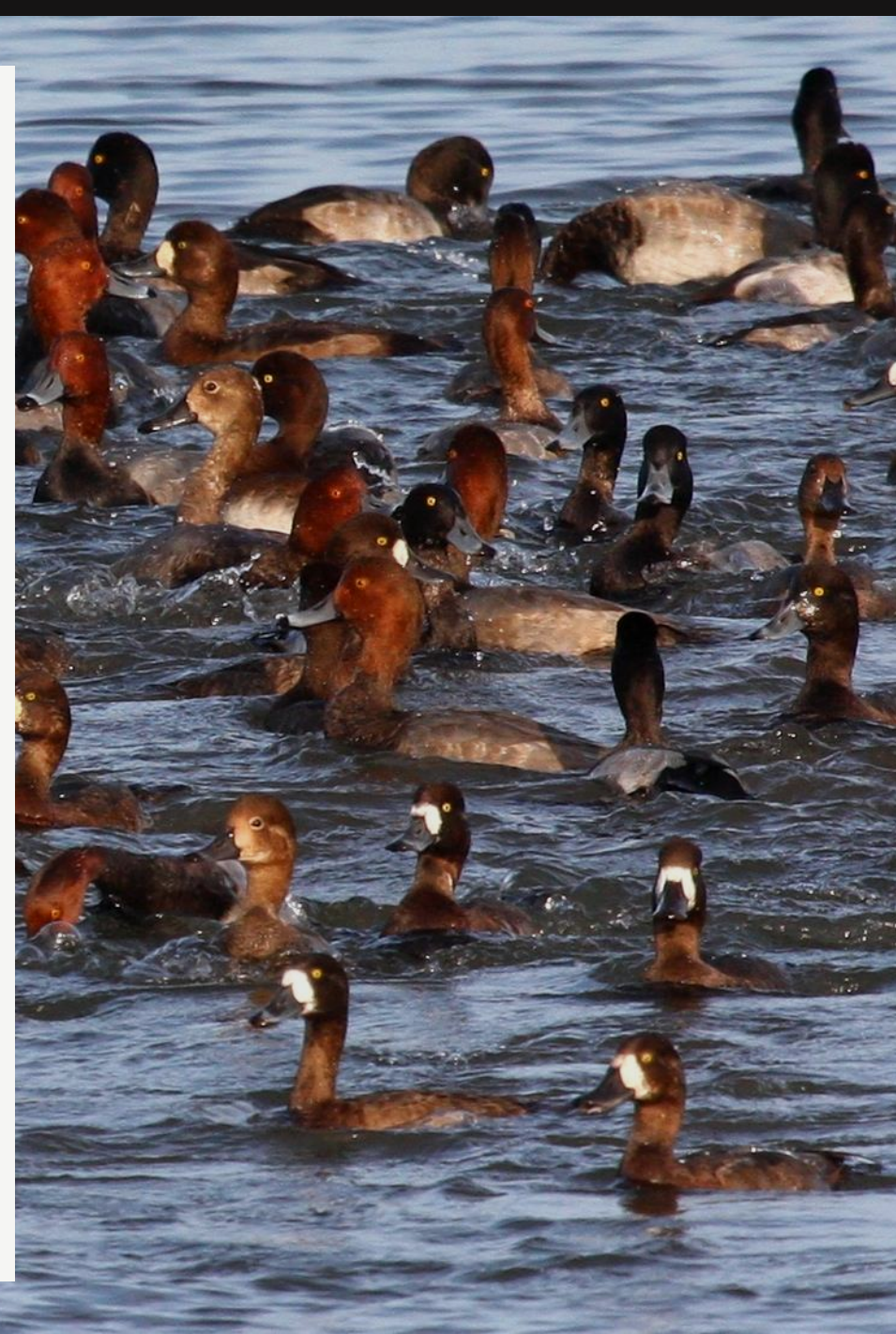


Nonmetric Multidimensional Scaling (NMDS): Steps

Step 1: Calculate distance/dissimilarity matrix

Step 2: Choose number of dimensions *a priori*

Step 3: Set initial configuration in m dimensions



Nonmetric Multidimensional Scaling (NMDS): Steps

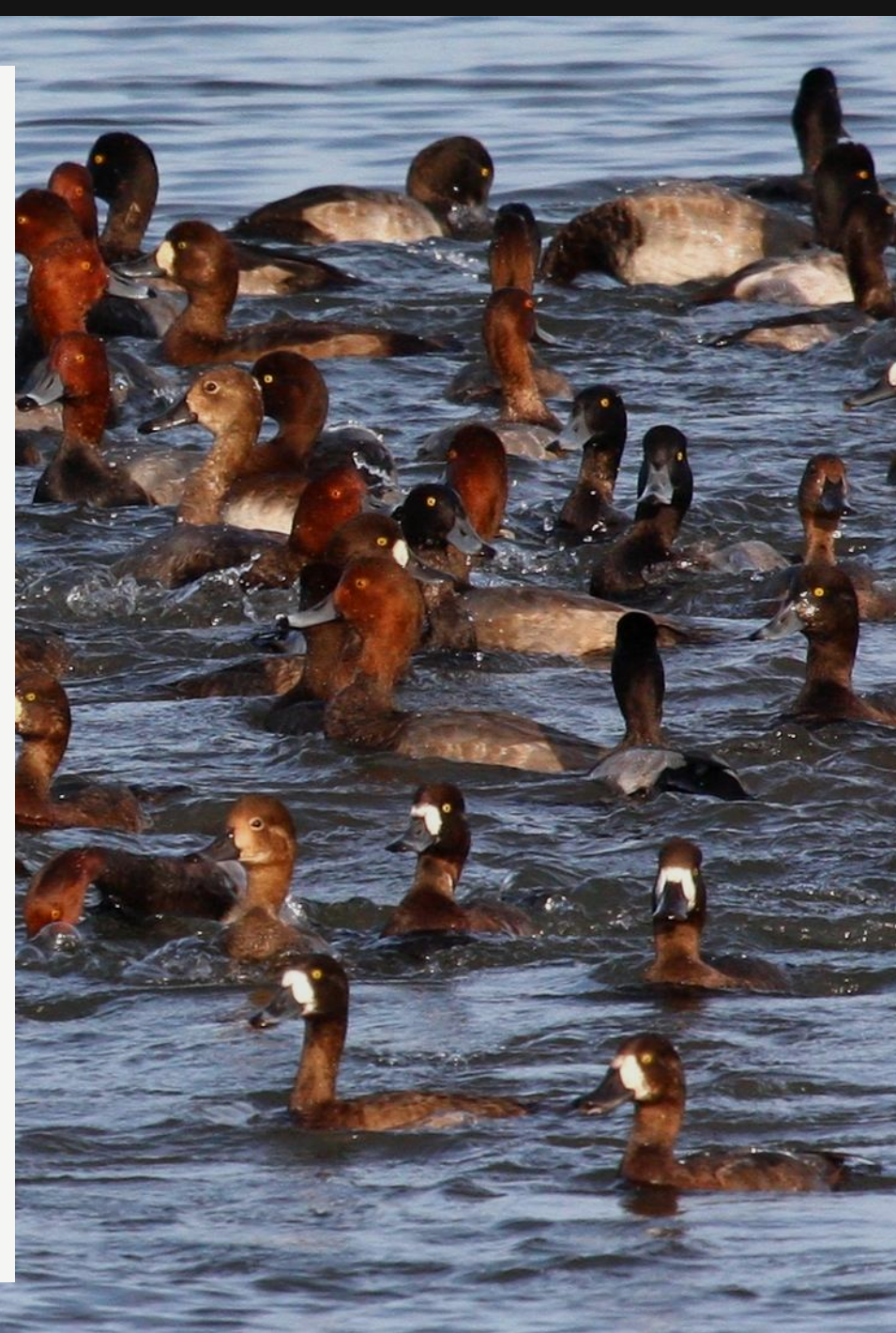
Step 1: Calculate distance/dissimilarity matrix

Step 2: Choose number of dimensions *a priori*

Step 3: Set initial configuration in m dimensions

Step 4: Iteratively optimize the configuration

Many iterations helps avoid the **local minima problem**.



Nonmetric Multidimensional Scaling (NMDS): Steps

Step 1: Calculate distance/dissimilarity matrix

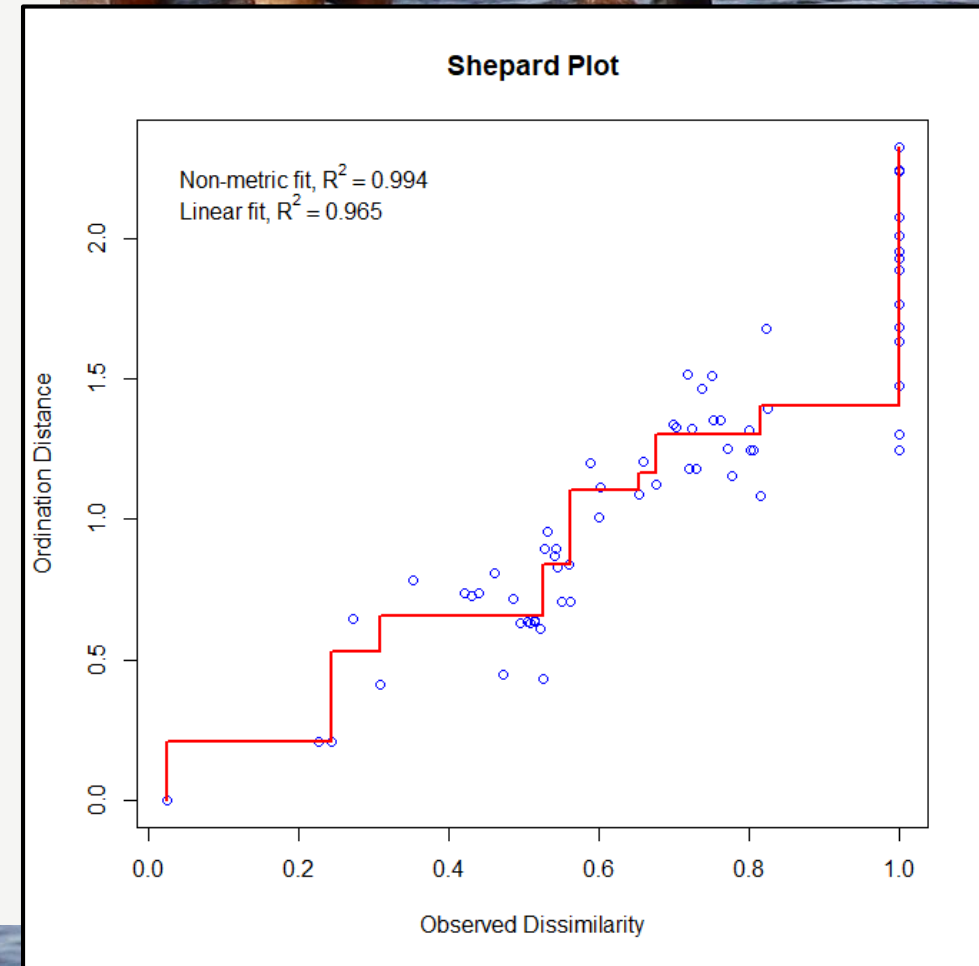
Step 2: Choose number of dimensions *a priori*

Step 3: Set initial configuration in m dimensions

Step 4: Iteratively optimize the configuration

Step 5: Evaluate **Shepard diagram** and **stress**

Stress values below 0.2 are borderline, below 0.1 are good, and 0.05 are excellent.



Nonmetric Multidimensional Scaling (NMDS): Steps

Step 1: Calculate distance/dissimilarity matrix

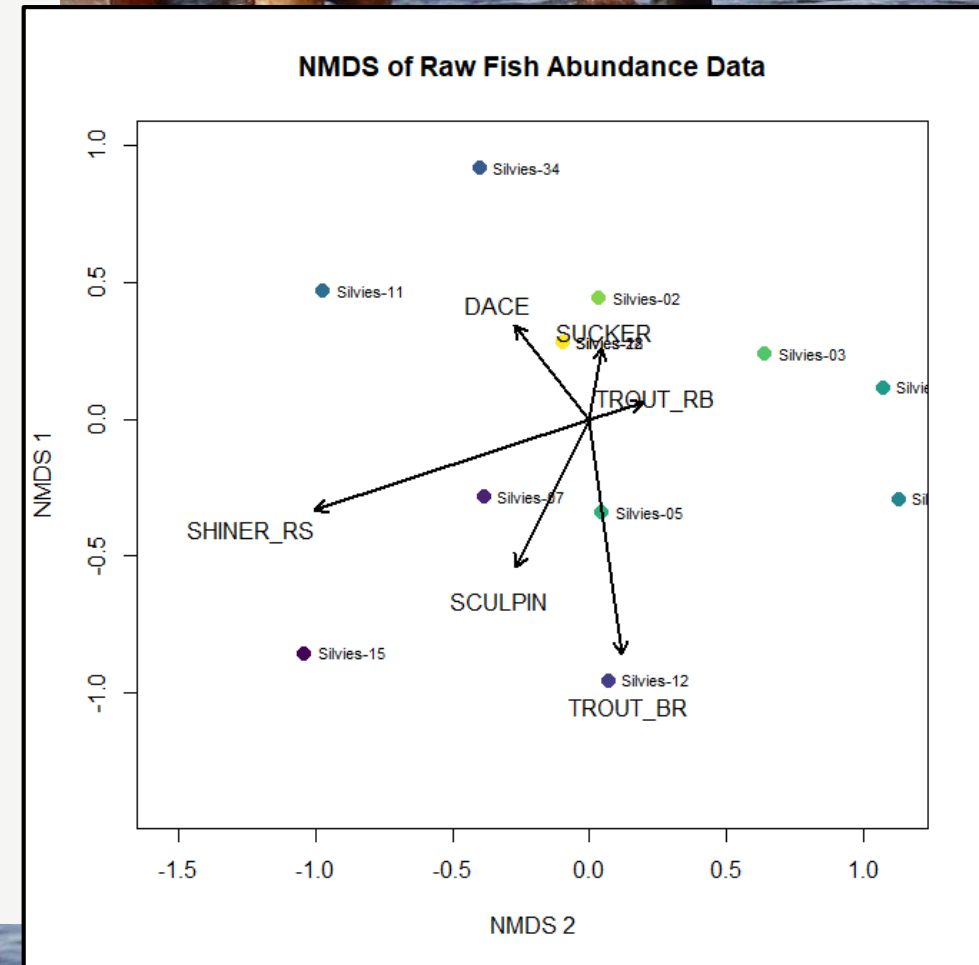
Step 2: Choose number of dimensions *a priori*

Step 3: Set initial configuration in m dimensions

Step 4: Iteratively optimize the configuration

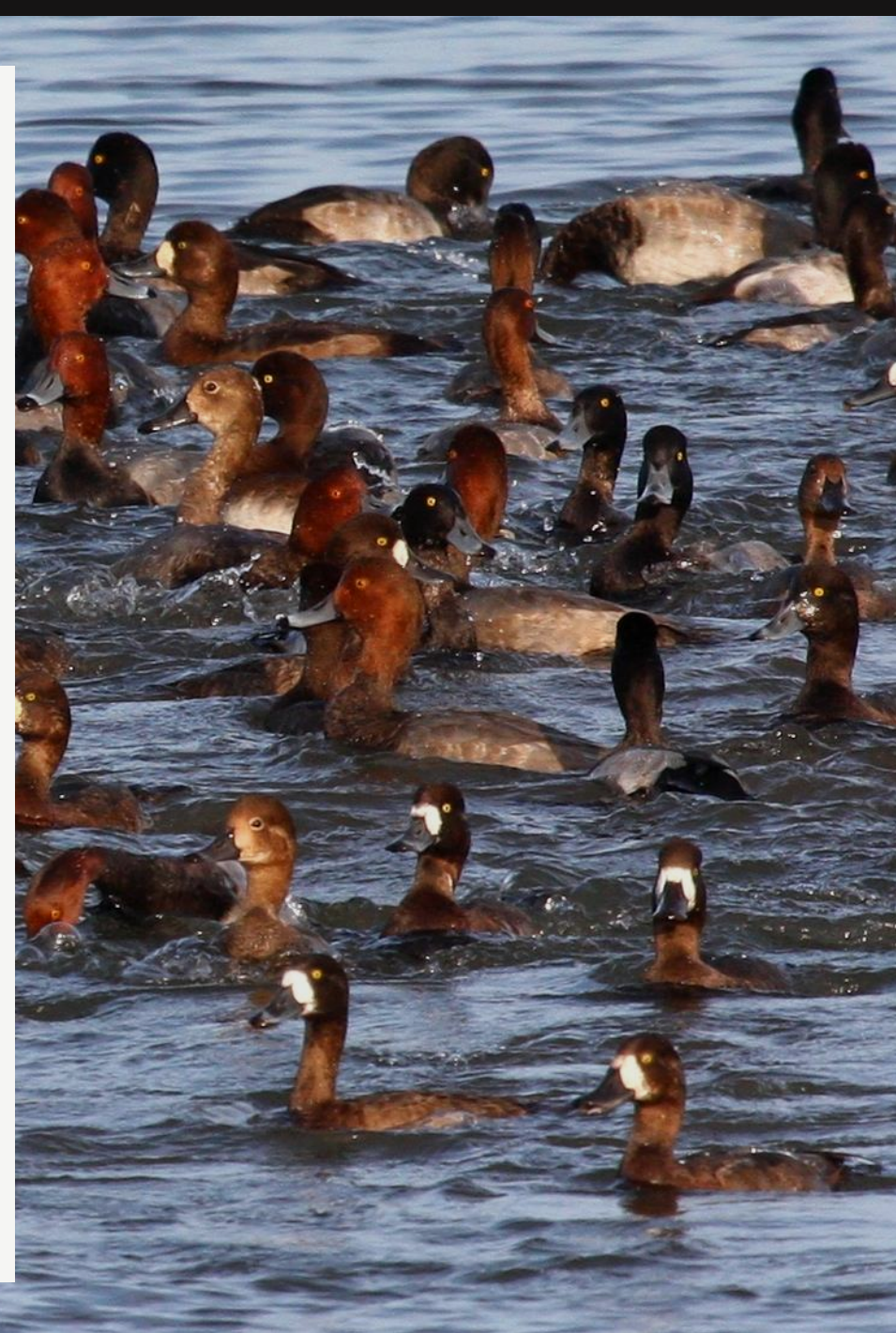
Step 5: Evaluate **Shepard diagram** and **stress**

Step 6: Plot NMDS biplot



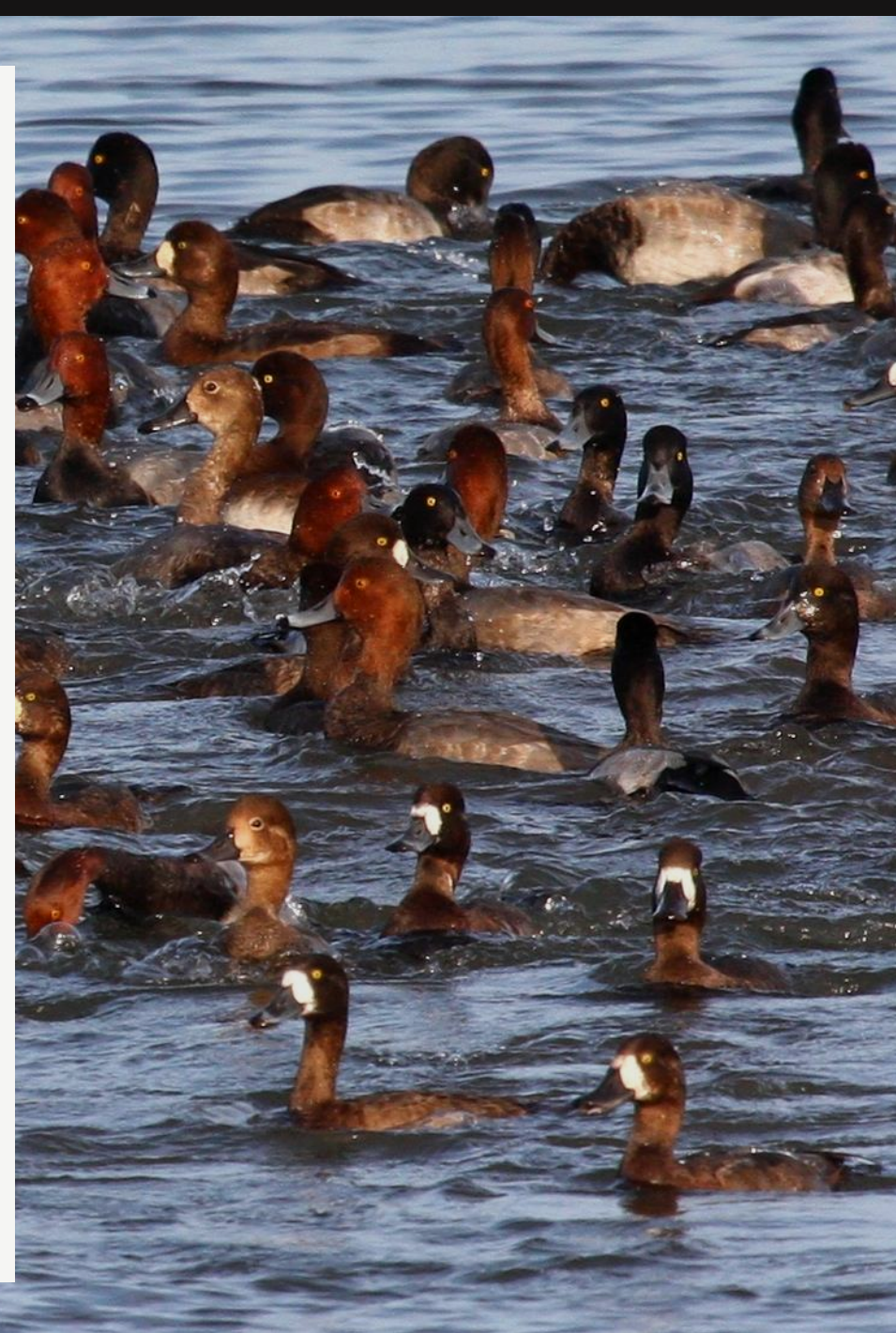
Nonmetric Multidimensional Scaling (NMDS): Limitations

- Ordination based on rank order rather than exact distances



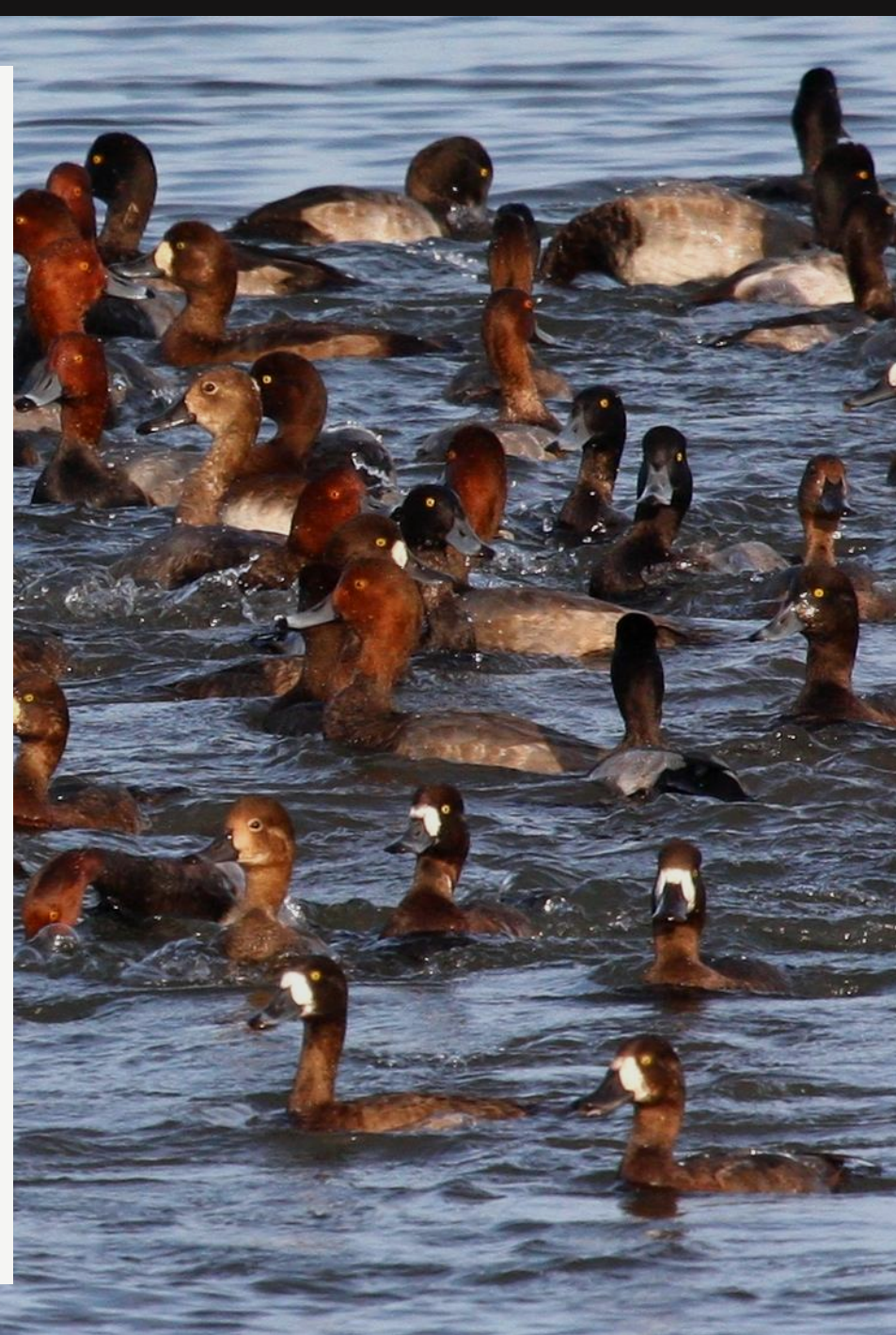
Nonmetric Multidimensional Scaling (NMDS): Limitations

- Ordination based on rank order rather than exact distances
- Computationally intensive for large datasets



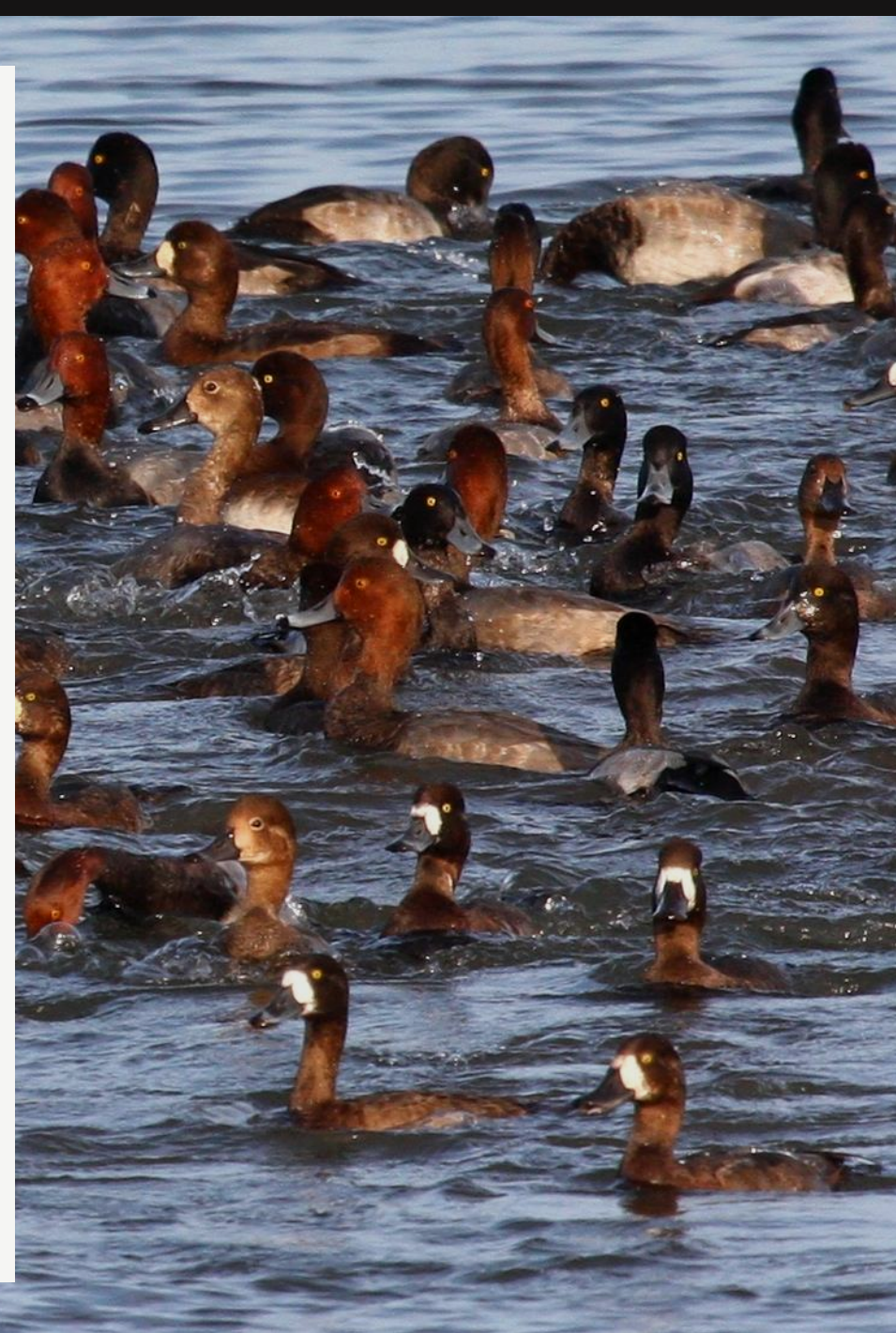
Nonmetric Multidimensional Scaling (NMDS): Limitations

- Ordination based on rank order rather than exact distances
- Computationally intensive for large datasets
- Can be sensitive to starting configuration



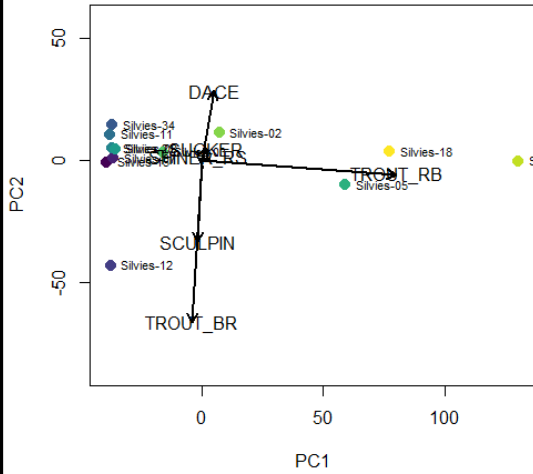
Nonmetric Multidimensional Scaling (NMDS): Limitations

- Ordination based on rank order rather than exact distances
- Computationally intensive for large datasets
- Can be sensitive to starting configuration
- Risk of overfitting

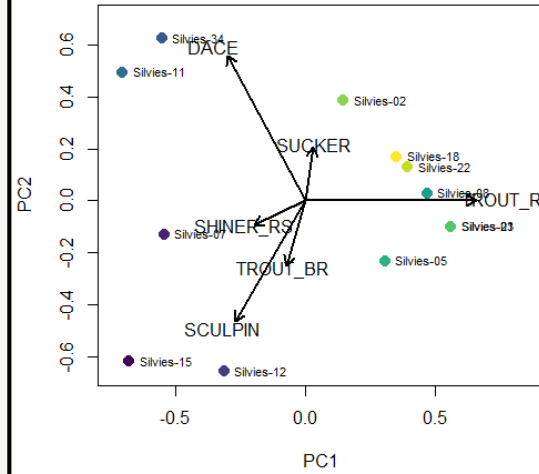


So Which One?

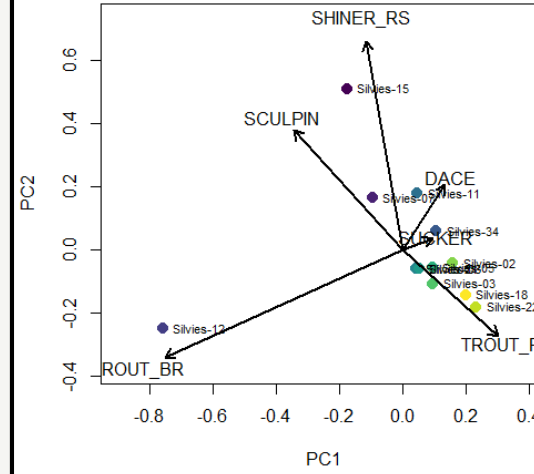
PCA of Raw Fish Abundance Data



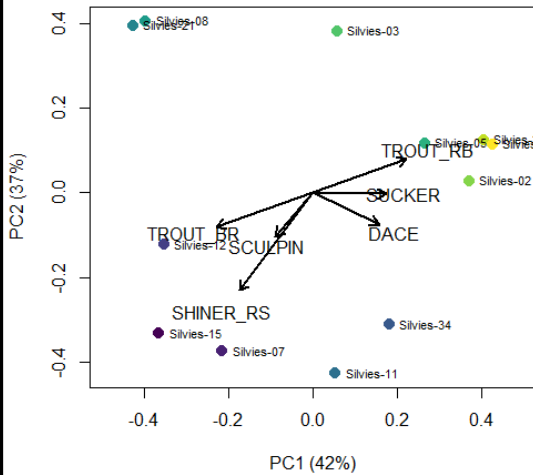
PCA of Hellinger Transformed Fish Abundance Data



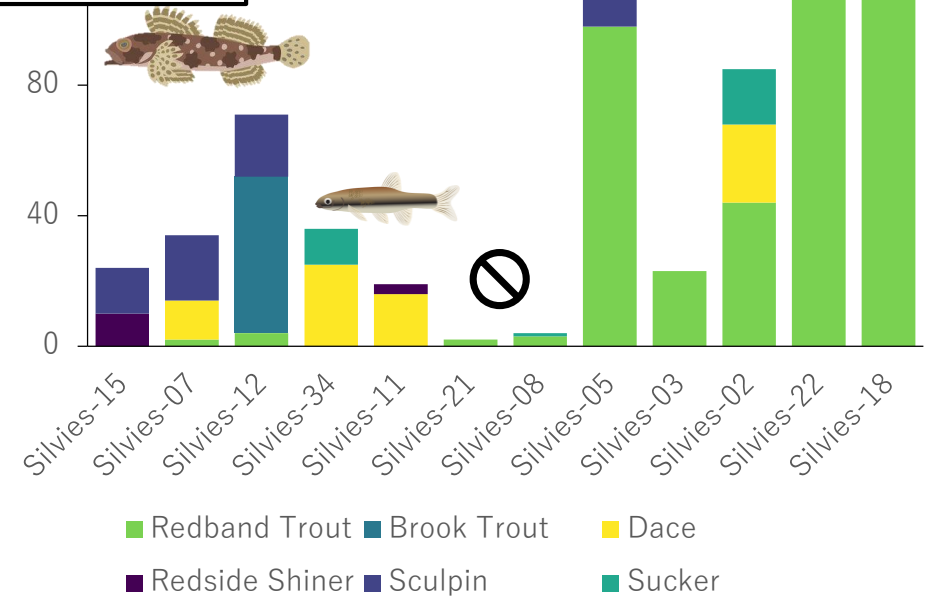
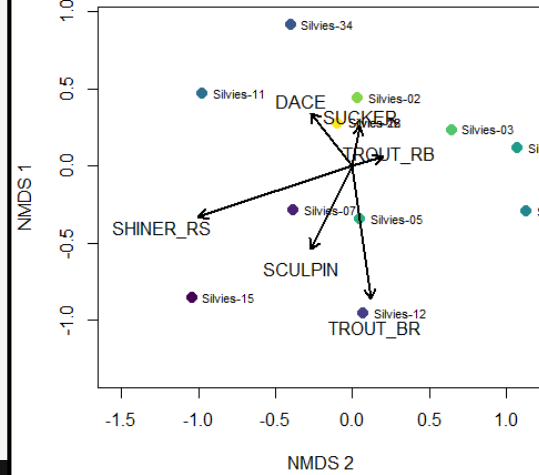
CA of Raw Fish Abundance Data



PCoA of Raw Fish Abundance Data

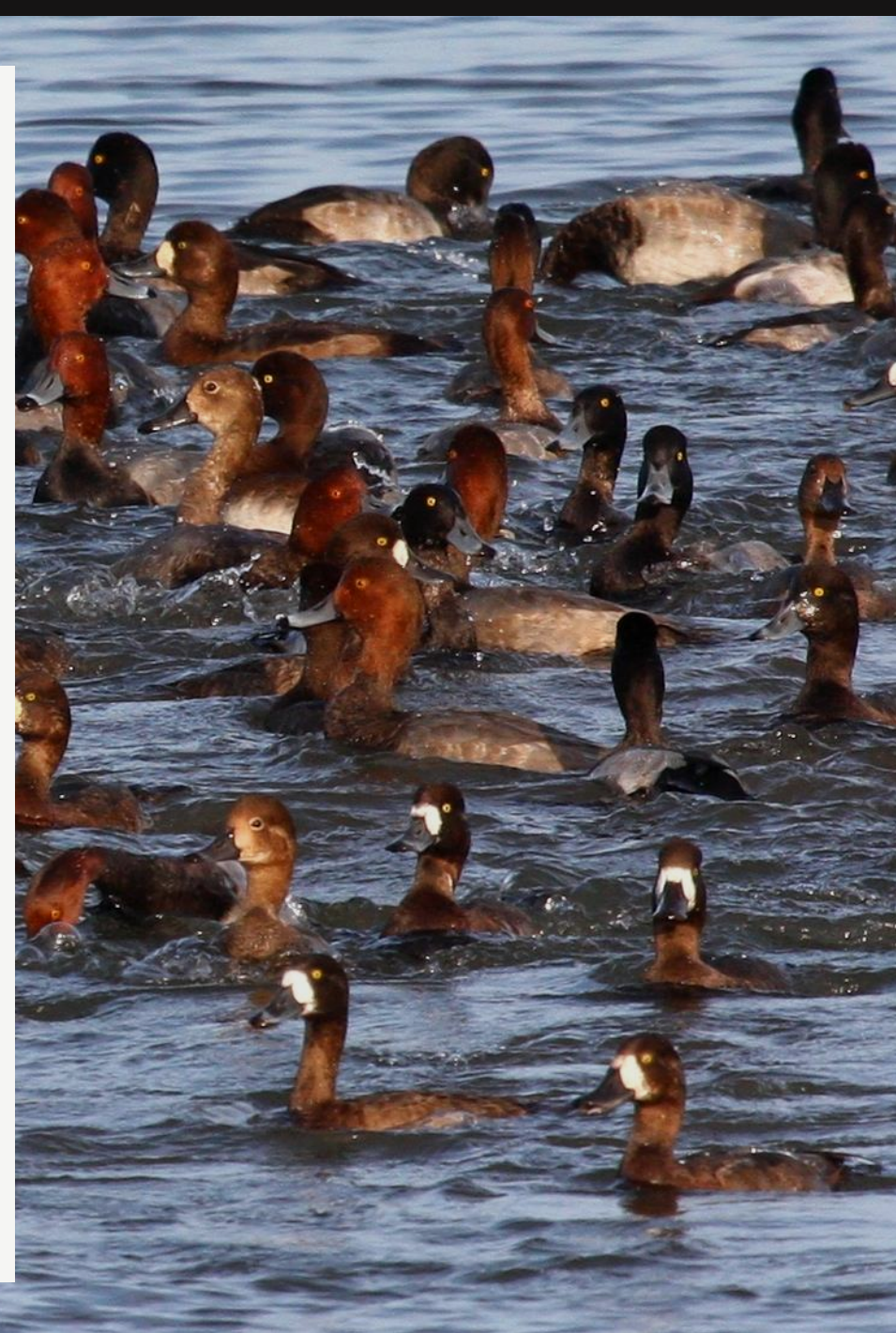


NMDS of Raw Fish Abundance Data



Conclusion: Summary of Key Points

- **Correspondence Analysis**
 - Relies on chi-square statistic
 - Designed for non-negative, frequency-based data like species abundances
 - Over-weighs importance of rare species
- **Principal Coordinate Analysis**
 - Can accommodate any distance matrix
 - May not capture all variance
 - More flexibility = more opportunities to make mistakes
- **Nonmetric Multidimensional Scaling**
 - Highly flexible and great for non-metric data
 - Axes are arbitrary



Questions?

