Link Prediction in Ecological Networks using WLNM with Directed Graph Support and Ecological Sampling

***Student:*** *Jorge Eduardo Castro Cruces*

***Supervisor:*** *Athen Ho Ma*

**Tables**

TBD

**Figures**

A graph with red and blue lines

AI-generated content may be incorrect. A graph of a stream

AI-generated content may be incorrect. A graph of a train

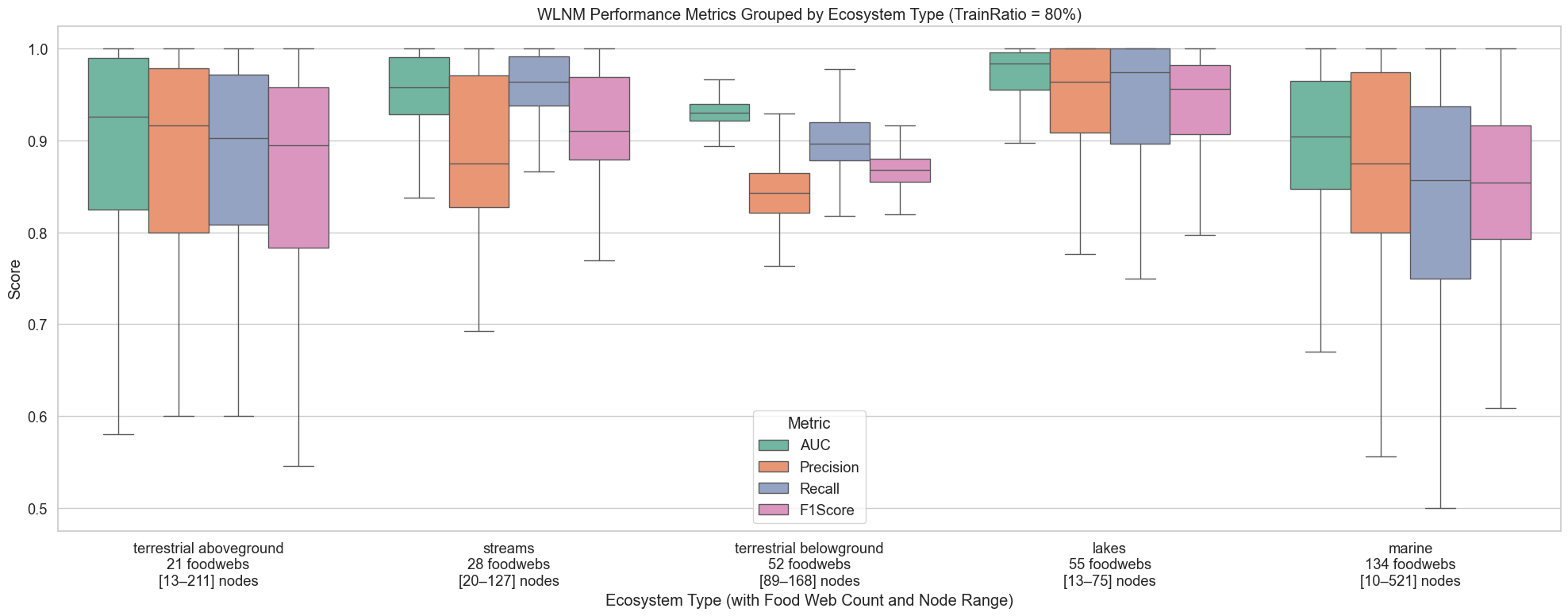
AI-generated content may be incorrect.

A graph with lines and numbers

AI-generated content may be incorrect.A graph with lines and numbers

AI-generated content may be incorrect.

**Figure 1**. Performance of WLNM across train–test splits. For each ecosystem (five panels), coloured lines show the mean AUC, Precision, Recall and F1 as the training fraction increases from 60% to 80% (markers at 60, 65, 70, 75, 80). Scores remain high and largely stable across this range, with only marginal gains towards 80%.



**Figure 2.** WLNM performance metrics grouped by ecosystem type at 80% train ratio. Boxplots summarize AUC, Precision, Recall, and F1-Score across terrestrial aboveground, streams, terrestrial belowground, lakes, and marine food webs. This figure highlights ecosystem-level variability in model performance.

A graph with different colored squares

AI-generated content may be incorrect.

**Figure 3.** WLNM performance metrics at 80% train ratio across 290 food webs. Boxplots show the distribution of AUC, Recall, F1-Score, and Precision. This figure evaluates overall predictive performance when most of the observed links are used for training.

A screenshot of a graph paper

AI-generated content may be incorrect.

**Figure 4.** Original adjacency matrix (left) and predation matrix with training results (right) for the Baxver lake food web. Black dots represent observed links; in the predation matrix, green dots mark true positives, red squares false positives, black dots training links, and grey crosses false negatives. This figure evaluates how well the model reproduces observed predator–prey interactions when trained on this network.

A screenshot of a graph

AI-generated content may be incorrect.

**Figure 5.** Original adjacency matrix (left) and predation matrix with training results (right) for the TPP food web. Black dots indicate observed interactions; in the predation matrix, green dots are true positives, red squares false positives, black dots training links, and grey crosses false negatives. This figure illustrates the accuracy and errors of WLNM predictions on this network.

A screenshot of a graph

AI-generated content may be incorrect.

**Figure 6.** Original adjacency matrix (left) and predation matrix with training results (right) for the Isolated stream 877 August 2003 food web. Black dots mark observed links; in the predation matrix, green dots represent true positives, red squares false positives, black dots training links, and grey crosses false negatives. This figure visualizes the model’s predictive accuracy at the interaction level..

A screenshot of a graph

AI-generated content may be incorrect.

**Figure 7.** Original adjacency matrix (left) and predation matrix with training results (right) for the Big Hopu lake food web. Black dots correspond to observed predator–prey links; in the predation matrix, green dots mark true positives, red squares false positives, black dots training links, and grey crosses false negatives. This figure assesses WLNM’s predictive accuracy for this lake ecosystem.

A graph of different colored bars

AI-generated content may be incorrect.

**Figure 8.** AUC scores per food web, showing mean values with min–max ranges over 10 runs. Bars represent individual food webs, with error bars indicating variability across runs. This figure evaluates the stability and robustness of WLNM predictions at the food-web level.

A graph of different colored bars

AI-generated content may be incorrect.

**Figure 9.** Precision scores per food web, showing mean values with min–max ranges over 10 runs. Bars represent individual food webs, with error bars capturing variability across runs. This figure evaluates WLNM’s ability to avoid false positives at the food-web level.

A graph of a graph showing the difference between false and true

AI-generated content may be incorrect.

**Figure 10.** Precision–Recall curves for the AEW17 terrestrial belowground food web (N = 145). Solid line indicates predictions on seen links (training), while the dashed line shows unseen links (testing). This figure evaluates WLNM’s ability to generalize beyond the training set.

A graph with a line

AI-generated content may be incorrect.

**Figure 11.** Precision–Recall curves for the AEW03 terrestrial belowground food web (N = 122). The solid blue line corresponds to predictions on seen links (training), while the dashed orange line represents predictions on unseen links (testing). This figure evaluates WLNM’s predictive performance and generalization capacity.

A diagram of a positive sublocation

AI-generated content may be incorrect.

**Figure 12.** Top-2 positive enclosing subgraphs extracted from the WLNM pipeline. Red diamond nodes denote the target link, and orange circles represent surrounding nodes forming the enclosing subgraph. Counts indicate how many times each subgraph appeared. This figure illustrates typical local network patterns around true positive links.

A screenshot of a diagram

AI-generated content may be incorrect.

**Figure 13.** Two randomly sampled negative enclosing subgraphs. Red diamond nodes indicate the target non-link, while orange circles represent surrounding nodes forming the enclosing subgraph. These structures are used as negative samples for training WLNM.

A map of the world

AI-generated content may be incorrect.

**Figure 14.** Geographic distribution of the 290 food webs used in this study, grouped by ecosystem type. Symbols indicate food web locations and colors represent ecosystem categories (e.g., terrestrial aboveground, streams, lakes, marine, belowground). This figure provides a spatial overview of dataset coverage.

A black background with white squares and dots

AI-generated content may be incorrect.

**Figure 15.** Flowchart of the WLNM-based pipeline for link prediction in food webs. The diagram outlines the major steps: train/test split, negative sampling, enclosing subgraph extraction and encoding (WLNM), and machine learning classification. This figure provides a structured overview of the workflow.