# Project Proposal

**Project Title:** Woodlands and Waterways Ecowatch

**Team Members:** Manthan Jani (0814650), Vishesh Patel (0792832)

**Date:** 2024-10-18

## Introduction

* 1. **Topic Statement**

This project focuses on exploring the functional diversity of benthic macroinvertebrate communities and assessing ecosystem health indices across various site characteristics in Haliburton County lakes over a 5-year period.

**1.2 Project Objective**  
Our objective is to develop predictive models for forecasting changes in benthic community composition using key environmental parameters such as elevation, drawdown, and headwater status, with an emphasis on biological responses to water chemistry.

**1.3** **Solution Overview**We propose to use statistical analyses and machine learning techniques to model the relationship between environmental factors and benthic community composition. By employing predictive modeling, we aim to uncover trends that can inform lake management decisions. This comprehensive analysis will include exploratory data analysis, functional diversity assessments, and the development of forecasting models for benthic community changes.

**1.4Motivation**  
The health of aquatic ecosystems is vital to biodiversity and water quality management. Understanding the factors that drive changes in benthic macroinvertebrate communities will provide valuable information for environmental management and conservation efforts. Our work will aid decision-makers in Haliburton County by delivering evidence-based insights into how lakes respond to natural and human-induced disturbances over time.

## 2. Literature Review

Ecosystem health and biodiversity assessment through benthic macroinvertebrate studies is well-documented in ecological research. Benthic communities serve as bioindicators of pollution and disturbance, making them essential for assessing water quality (Merritt et al., 2008). Functional diversity, which emphasizes the roles organisms play in an ecosystem, is increasingly recognized as a better indicator of ecosystem function than species richness alone (Petchey & Gaston, 2006).

Studies on predictive ecological modeling highlight the effectiveness of machine learning techniques such as random forest and generalized additive models in forecasting biological responses to environmental changes (Breiman, 2001; Hastie & Tibshirani, 1990). Research on water chemistry parameters like pH, dissolved oxygen, and nutrient load has also shown a significant influence on benthic communities and lake trophic status (Jones et al., 2003).

**References**

1. Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
2. Hastie, T., & Tibshirani, R. (1990). Generalized additive models. Chapman & Hall/CRC.
3. Jones, J. R., Obrecht, D. V., & Perkins, B. D. (2003). Impact of nutrients, temperature, and zooplankton on epilimnetic chlorophyll a: Regression analysis and modeling. Journal of Lake and Reservoir Management, 19(3), 261-273. <https://doi.org/10.1080/0743814030935394> 1
4. Merritt, R. W., Cummins, K. W., & Berg, M. B. (Eds.). (2008). An introduction to the aquatic insects of North America (4th ed.). Kendall Hunt Publishing Company.
5. Petchey, O. L., & Gaston, K. J. (2006). Functional diversity: Back to basics and looking forward. Ecology Letters, 9(6), 741-758. <https://doi.org/10.1111/j.1461-0248.2006.00924.x>

## 3. Dataset

**3.1 Data Overview**

The dataset spans five years of benthic macroinvertebrate surveys across various lakes in Haliburton County, provided by Woodlands and Waterways EcoWatch (WWEW). Key variables include:

1. Biological data: Benthic macroinvertebrate taxa and functional diversity indices.
2. Environmental data: Site characteristics such as elevation, water level drawdown, and headwater status.
3. Water chemistry data: pH, dissolved oxygen, conductivity, and nutrient concentrations.

**3.2 Source**

The data were collected by WWEW in partnership with U-Links Centre for Community-Based Research. The data are credible and well-suited for assessing ecological health trends over time.

**3.3 Assumptions**

We assume that environmental parameters like elevation, drawdown, and headwater status significantly influence benthic community composition. Additionally, water chemistry factors such as pH and dissolved oxygen are assumed to be primary drivers of functional diversity and ecosystem health.

## 4. Methodology

**4.1 Analytical Techniques**

We will begin with exploratory data analysis (EDA) to identify trends and outliers. For functional diversity assessments, indices like the Rao’s quadratic entropy and functional dispersion will be used. Statistical techniques such as Principal Component Analysis (PCA) will help identify relationships between environmental drivers and community composition.

For predictive modeling, we will explore machine learning algorithms such as:

* **Random Forest**: For its ability to handle complex, non-linear interactions between environmental factors and benthic community composition.
* **Generalized Additive Models (GAMs)**: To model the non-linear effects of variables like water chemistry on functional diversity indices.

**4.2 Justification**

Predictive models are ideal for this project because they allow us to forecast future changes in benthic communities based on environmental variables. These models will provide valuable insights into how different lakes may respond to ongoing environmental changes.

**4.3 Alternative Approaches**

Other approaches, such as linear regression or spatial analysis, were considered but were found less suited for capturing the complex interactions between environmental parameters and biological responses. Predictive models like random forest and GAMs are better equipped to handle the intricacies of ecological data.

The project will follow a structured timeline, as outlined below:

|  |  |  |
| --- | --- | --- |
| **Task** | **Oct '24** | **Nov '24** |
| Data Collection & Preprocessing | X |  |
| Exploratory Data Analysis | X |  |
| Functional Diversity Analysis | X |  |
| Predictive Model Development |  | X |
| Model Validation |  | X |
| Report Writing & Poster Prep |  | X |

The Gantt chart breaks down the major tasks into phases, from data preprocessing to final report writing and poster preparation for the March 2025 presentation.

## Conclusion

This project will provide valuable insights into how environmental factors influence the functional diversity and ecosystem health of benthic macroinvertebrate communities in Haliburton County lakes. The results will help inform lake management strategies and contribute to the broader understanding of freshwater ecosystem health in response to both natural and anthropogenic influences.