TSA FINAL REPORT:

Yuxiang_Ren

Spring 2023

Contents

Abstract

Introduction

Method (Data Processing)

Data

Three datasets used in the project were collected from EDI Data Portal, including:

1. North Temperate Lakes LTER: Phytoplankton - Madison Lakes Area 1995 – current;(Magnuson & H.Stanley, 2022)
2. North Temperate Lakes LTER: Physical Limnology of Primary Study Lakes 1981 – current;
3. North Temperate Lakes LTER: Chemical Limnology of Primary Study Lakes: Nutrients, pH and Carbon 1981 – current.

The three files respectively record the water body phytoplankton information, physical information and chemical information of multiple lakes in the Wisconsin range. We analyzed these data at the beginning stage to screen out suitable research subjects, including the target lake, and primary algae responsible for blooms. First, we chose Mendota Lake (ME) for this project, as it has more time measurement data compared to other lakes, which might be more conducive to time series analysis and obtaining more reliable results

(Table 1). Second, to obtain information on dominant species that may cause water blooms, we accumulated the biomass of algae from different divisions and considered the algae with the highest total biomass to be the main contributor to water bloom outbreaks. It is worth noting that the original data records the biomass of specific algal species on the observation day. Therefore, to obtain division-level data, we summed the biomass of all species within the same division on the same day to obtain the biomass information for the division. The result shows that the dominant division is Cyanophyta, which is also consistent with other studies (Table 2)[Brock (2012)](Beversdorf, 2015).

Table 1: Site Information

Site	Observation Date Count
ME	402
MO	355
WI	23
FI	1

Table 2: Division level Total Biomass (mg/L)

Division	Count	Total Biomass	Max Biomass	Min Biomass	Mean Biomass
Cyanophyta	8581	1824.25213	76.0000	0.00e+00	0.2125920
Bacillariophyta	1914	378.01888	13.4028	3.19e-05	0.1975020
Chlorophyta	4368	244.95205	84.6924	0.00e+00	0.0560788
Cryptophyta	1876	76.11009	2.9630	3.63e-05	0.0405704
Pyrrhophyta	415	29.12950	5.3194	0.00e+00	0.0701916

After identifying the target lake and algal division, we cleaned and combined the three data tables. The following are the data cleaning steps:

a. Integrate the phytoplankton data according to lakeid, sampledate, depth range, and division to obtain the biomass information of each division on the observation day. Then, filter out all data with a lake id of Mendota and a division of Cyanophyta.
b. Filter out the physical and chemical information of Lake Mendota. Considering that the original data records information at different depths on the same observation day, we calculated the average of all environmental data at depths of 0-8m, which correspond to the depths mentioned in the algae information. It is worth noting that on some dates, the depth of the algae information is 0-2m, and in these cases, we used the average environmental data for 0-2m.
c. Based on the sampling date and depth range, we combined these data together (Table 3).

Table 3: rawdata

lakeid	sampledate	total_biomass	Temperature	date_diff	TN	TP
ME	1995-01-24	0.0128965	NA	NA	NA	NA
ME	1995-03-28	0.0013183	NA	63	NA	NA
ME	1995-04-11	0.0017578	NA	14	NA	NA
ME	1995-04-24	0.0019157	NA	13	NA	NA
ME	1995-05-23	0.8052959	13.64444	29	0.7305	0.0895000
ME	1995-06-06	0.0738443	16.84444	14	0.7195	0.0756667

d. We averaged the data monthly and used the zoo function (na.approx, rule = 2) to fill in NA values. Due to this method is not suitable for filling in NA values at the beginning of data, data before May 1995 were removed.
e. The final dataset includes dates (from May 1995 to December 2020), temperature, total nitrogen, total phosphorus, and biomass (Table 4).

Table 4: Final Data

date	Temperature	TN	TP	Biomass
1995-05-01	13.64444	0.7305000	0.0895000	0.8052959
1995-06-01	18.45833	0.4156250	0.0627083	0.5145389
1995-07-01	23.27222	0.1007500	0.0359167	6.2401053
1995-08-01	25.31852	0.1928333	0.0252778	4.6301317
1995-09-01	20.22037	0.2849167	0.0681389	0.7740042

Result ()

ACF,PACF,+forecast

Discussion

Reference

```
#To add reference, you need first add reference in reference in reference.bib.
#Than use"[@uniqueID]" to site it.
```

Beversdorf, T. R. M., Lucas J; Miller. (2015). Long-term monitoring reveals carbon–nitrogen metabolism key to microcystin production in eutrophic lakes. *Frontiers in Microbiology*, *6*, 456.

Brock, T. D. (2012). *A eutrophic lake: Lake mendota, wisconsin*. Springer Science & Business Media.

Magnuson, S. R. C., J.J., & H.Stanley, E. (2022). North temperate lakes LTER: Phytoplankton - madison lakes area 1995 - current. Retrieved from https://portal.edirepository.org/nis/mapbrowse?scope=knb-lter-ntl&identifier=88&revision=31