DATE: 20th March, 2024

TO: The Commissioner – Maine Department of Agriculture, Conservation and Forestry (DACF)

FROM: Parikshit Patil (ppatil25@uwo.ca)

Briefing Note: Analysis of Maine Lakes' Mercury Levels

Issue

There is concern about toxic chemicals entering the food chain. The government of Maine seeks to be able to predict the mercury levels in the lakes of Maine using the collectable lake characteristics related to mercury levels in fish. These predictable insights should aim to help guide sampling efforts and/or lake pollution control efforts.

Summary

Following a previous task, a predictive model is now implemented. Although the predictive accuracy is not very strong, key lake characteristics affecting mercury levels were identified. The mercury levels of the lake are estimated to increase with an increase in elevation and latitude, and a decrease in runoff factor. Some data preprocessing was done to improve the predictive model's performance. More data gathered in the right structure will help further improve performance. Lakes with high mercury levels, high elevations, high latitudes, and low runoff factors should be given priority to execute pollution control efforts.

Background

A previous task (Appendix A1) initiated an analysis of this issue and resulted in the following key findings:

- The mercury levels at Maine's lakes were said to not be high enough to raise concern at the moment
- The presence of human infrastructure for the lakes was said to not affect the mercury levels of the lakes
- The lake type was said to not affect the mercury levels of the lakes.

Following these results, the government of Maine has decided to further proceed with the mentioned issue by being able to predict the mercury levels.

Current Status

The following lake characteristics are most important in best predicting the mercury levels: Elevation, Runoff factor, and Latitude.

At present, the following relationship can be followed between the mercury levels and the above-mentioned lake characteristics:

- By moving higher up in elevation only, the mercury levels can be said to increase by approximately 0.008ppm (part per million) for every foot of increase in the elevation
- At a constant elevation and latitude, every unit decrease in the runoff factor can lead to an expected increase in mercury levels of 0.19ppm
- For every single degree moved North, the mercury levels are expected to increase by 0.1ppm while the elevation and runoff factor is kept unchanged

Key Considerations

The results are based on a modified version of the dataset by removing missing values and values too extreme to be plausible. The results are considered to be used only in relevance with lakes that possess a similar range of characteristics to those provided in the dataset. The model does not make highly accurate predictions due to the nature of the dataset. Identifying the most important lake characteristics affecting the mercury levels has been given more importance than more precisely predicting the mercury levels.

Options

Gathering data with fewer missing values and anomalies can potentially lead to an improved predictive model. Collecting data from randomly sampled Maine lakes will help make the model more relevant to all the lakes of Maine in general. Overall, collecting more data and ensuring that there are a similar number of records collected for every lake will further improve the predictive model's performance. All these options implemented have the potential to greatly improve the scope and performance of the model, especially in terms of mercury level prediction accuracy and lake characteristic analysis

Recommendations

First, all lakes with mercury levels greater than 0.43ppm should be the primary target for pollution control efforts. Next, lakes with mercury levels below but near 0.43ppm should be the target for pollution control efforts giving priority to lakes with the highest elevations and latitudes, and lowest runoff factors. Lastly, a more random data sample should be collected, especially in terms of elevation, latitude, and runoff factor. The number of data samples obtained should aim to be more or less equal for different levels of elevation, latitude, and runoff factor, ensuring an important balance in the dataset.