

Data manipulation Lab

Trevor F, Monica K

Question 1: Define hypoxia. What is the dissolved oxygen concentration threshold that identifies hypoxia?

Hypoxia is an environmental phenomenon where the oxygen levels in water decreases to a point that cannot support the life of aquatic organisms. Hypoxia is detected if a dissolved oxygen concentration is less than or equal to 2 mg/L.

Question 2: Which species are affected? How are these species affected?

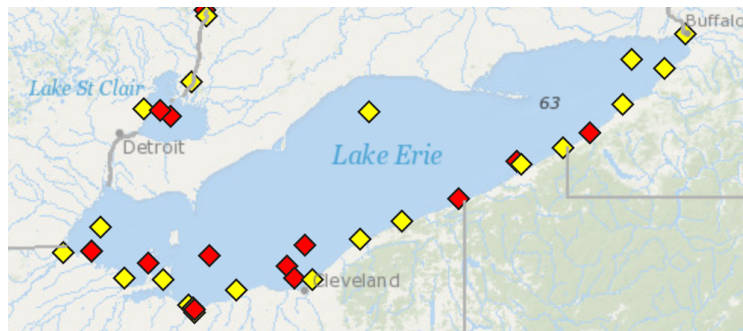
Fish, shrimp, crabs, and underwater fauna are affected. Fish, shrimp, and crabs typically migrate out of the location where hypoxia occurs to avoid harm, but the fauna are stuck in the area with hypoxia. As a result, the fauna experience stress or die from the lack of oxygen.

Question 3: Where and when is hypoxia likely to occur in Lake Erie? Either describe the general locations or provide an annotated map.

Hypoxia occurs in Lake Erie all the way from the southern coast from Port Clinton to Ashtabula, the northern coast from Arnherstburg to Port Stanley, and the area in between those two coast lines. Hypoxia will occur throughout late June to mid October. It typically occurs in the southwest area of the described region first, then fully encapsulates the described region, and finally ends with just the middle of the lake experiencing hypoxia.

Question 4: Are current NOAA buoys and sensors located in appropriate locations to monitor and surveillance of hypoxia? If so, which buoy/station IDs are in pertinent locations?

No, in lake Erie there are no NOAA-owned buoys or sensors. There are buoys not sponsored on Lake Erie, the buoys and sensors on Lake Erie are sponsored by NDBC, NOS, Limno Tech, GLERL, for example.



These buoys are in appropriate locations to monitor and survey hypoxia, as many of these stations (such as Station MRHO1 or Station GELO1) are located around the coast, and Station 45123 is located high center of Lake Erie; both of these locations are known to have low hypoxia levels throughout the year.

Question 5: Are there shipping lanes through the area that might provide sporadically-sampled monitoring data?

[Document says to skip this question.](#)

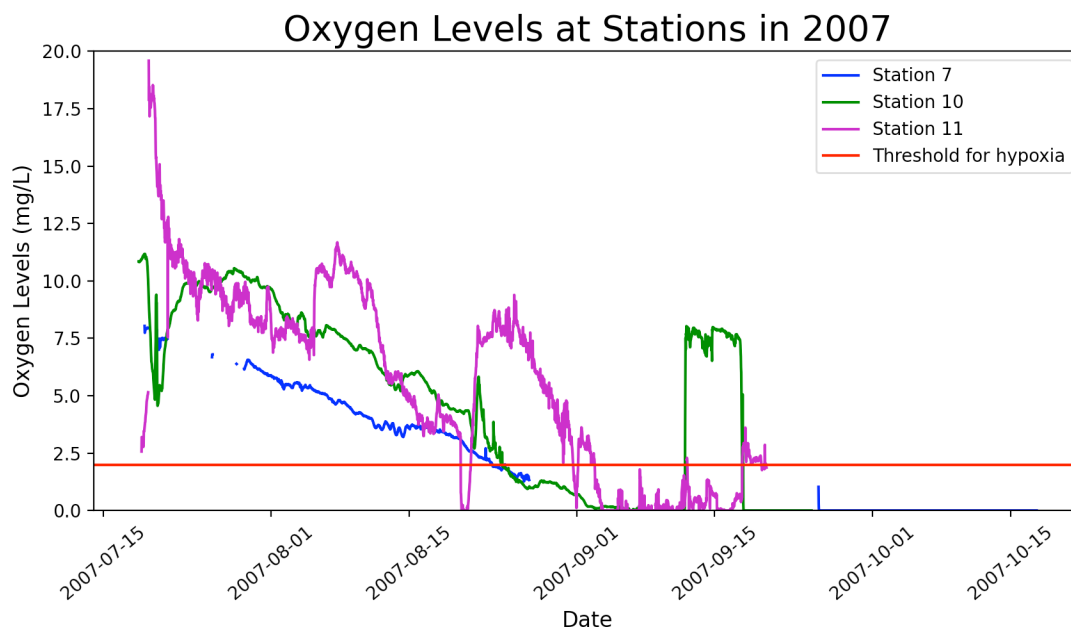
Question 6: These datasets contain very useful information (sensor depth, dissolved oxygen concentrations, temperature, turbidity, and chlorophyll) but were not always accurate (e.g., 2007 data). Describe the errors that are known to exist within the dataset.

[Across different datasets there seems to be missing data in multiple fields where the missing data is represented by a reading of -999.000. There also seems to be issues in certain data sets with negative readings for turbidity, chlorophyll, and dissolved oxygen levels which should not be possible. Also, some of the dissolved oxygen level readings seem to be too high to be valid readings, such as one dissolved oxygen level reading being 55.710 mg/L.](#)

Option A

Visualize hypoxia events in Lake Erie based on a high-dimension dataset. Create a histogram or similar visualization that details when the dissolved oxygen falls below 2 mg/L. Generate at least one chart for one location, covering data from the whole year at that location. Consider using matplotlib, ggplot2, or Tableau charts.

Example data: <https://www.glerl.noaa.gov/res/projects/ifyle/data/Mooring/ysi/2007/Y18.txt>
where day/time 206.0208 = 30 minutes, between day 205/365 to day 288/365



Deliverables:

Turn in a Word/text file containing your answers to Step 1 questions and Step 3 text + visualization. Also, include ANY computing code (e.g., R/Python scripts), Tableau workbooks (.twbx file), intermediate datasets (e.g., cleaned CSV files if small in size), etc. Upload a softcopy of these documents to Blackboard.