Kendall Greer

Water Chemistry Correlations in Santa Cruz Island, Galapagos: An R Studio Visual Analysis

Many physical, chemical, and biological factors and parameters interact to form diverse sets of complex aquatic ecosystems around the world (Liang, 2024). In smaller, more shallow aquatic systems, parameters such as dissolved oxygen, pH, ion concentrations, and temperature can fluctuate and interact in ways that create a diel pattern over time (Andersen, 2017). Specifically looking at these four parameters (dissolved oxygen, pH, specific conductivity, and temperature), it is expected that dissolved oxygen, pH, and temperature will all rise throughout the day, as sunlight fosters the release of oxygen through photosynthesis, and plummet after sundown with the release of carbon dioxide and acidic hydrogen ions overnight. Specific conductivity theoretically follows this pattern inversely, due to the release of high concentrations of hydrogen ions at peak respiration times; however, this pattern is not always observed (Reddy 1981). Certain questions regarding these patterns arise when looking at inland freshwater systems in the Galapagos, as they are largely unstudied due to their secluded nature. Looking at shallow ponds on Santa Cruz Island in the Galapagos begs the question: Does this system show the patterns of correlation and diel cycling seen in similar bodies of water?

To answer these questions, I used R Studio packages ggplot, patchwork, and plotly to create both 24-hour diel scatterplots and linear regressions analyzing parameter relations for three freshwater sites on Santa Cruz (Slater, 2019). I asked ChatGPT for a general outline of each of the graphs I wanted, then pasted the responses into R Studio and inserted my own data into the outlines. Generally, my first prompt was able to create a usable outline to create the graphs; however, I also used ChatGPT to troubleshoot by pasting my erroneous code in and asking ChatGPT to find any issues with it. Troubleshooting took anywhere from 1-5 attempts, depending on how many issues there were within my code.

My first figure was a 24-hour scatterplot of each individual parameter created using ggplot. These four plots were then combined into one, 4-panel figure using patchwork. My second figure mirrored the first in structure, but plotted two parameters against each other in a linear regression in order to assess correlation. I also fitted my linear regressions with a trendline equation and R-squared value for quantitative analysis of the parameter relationships. My final figure combined the ideas of the first two in the form of a 24-hour scatterplot created using plotly, with 2-parameter comparisons to visualize diel fluctuations. Plotly suited my data well because it is an interactive interface that allows figures to have a primary and secondary y-axis, which is useful for comparing two parameters with variable ranges (GeeksforGeeks, 2024) Eventually, my goal is to incorporate these plotly figures into an accessible database, where anyone could view and interact with the data visualizations.

Kendall Greer

Selected References

Andersen, M. R., Kragh, T., & Sand-Jensen, K. (2017-9-13). Extreme diel dissolved oxygen and carbon cycles in shallow vegetated lakes. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1862). https://doi.org/10.1098/rspb.2017.1427

GeeksforGeeks. 2024. Getting Started with Plotly in R. GeeksforGeeks. From www.geeksforgeeks.org/getting-started-with-plotly-in-r/

Liang, X., Jian, Z., Tan, Z., Dai, R., Wang, H., Wang, J.,…Li, T. (2024-10-29). Dissolved Oxygen Concentration Prediction in the Pearl River Estuary with Deep Learning for Driving Factors Identification: Temperature, pH, Conductivity, and Ammonia Nitrogen. *Water 2024, Vol. 16, Page 3090*, *16*(21). https://doi.org/10.3390/w16213090

Reddy, K. R., & Reddy, K. R. (1981-12-01). Diel variations of certain physico-chemical parameters of water in selected aquatic systems. *Hydrobiologia 1981 85:3*, *85*(3). https://doi.org/10.1007/BF00017610

R Core Team. 2024. R: A Language and Environment for Statistical Computing. from https://www.r-project.org/

Slater, L. J., Thirel, G., Harrigan, S., Delaigue, O., Hurley, A., Khouakhi, A.,…Smith, K. (2019/07/12). Using R in hydrology: a review of recent developments and future directions. *Hydrology and Earth System Sciences*, *23*(7). https://doi.org/10.5194/hess-23-2939-2019

Wong, M.-F., Guo, S., Hang, C.-N., Ho, S.-W., Tan, C.-W., Wong, M.-F.,…Tan, C.-W. (2023-06-01). Natural Language Generation and Understanding of Big Code for AI-Assisted Programming: A Review. *Entropy 2023, Vol. 25, Page 888*, *25*(6). https://doi.org/10.3390/e25060888