NM WRRI Student Water Research Grant Final Report

1. Student Researcher: Paige G. Tunby
   Faculty Advisor: Ricardo González-Pinzón

2. Project title:

   Informing River Corridor Transport Modeling by Harnessing Community Data and Machine Learning

3. Description of the research problem and research objectives.

   River corridors connect terrestrial and aquatic ecosystems and allow critical interactions between hydrological, biological, geological, and chemical processes. It is crucial to understand how water moves in river corridors. We use tracers, such as dyes or salts, to track changes in concentration over time at downstream locations. The combined concentration and time data are called breakthrough curves (BTCs) and are the most common methods for observing solute transport in rivers and understanding water flow and storage. Researchers collect these data in streams and rivers across the globe. However, it is difficult to contextualize their results with other studies due to the lack of standardization in how the data is collected and made publicly available. We are creating a database gathering a wealth of information from researchers around the globe to facilitate the recognition of scaling patterns through physics-informed machine learning and move us closer to having real-time estimates of solute transport processes in river corridors.

4. Description of the methodology employed.

   We have constructed a preliminary database in Microsoft Excel comprising approximately ~450 sheets, each with individual BTCs and all accompanying hydrogeomorphic measurements. We created an Entity Relationship Diagram (ERD) was created for the data. An ERD is a chart that illustrates how entities (e.g., people, concepts, objects) relate to each other through relationships. This schema was translated into a Logical Relational Schema, which defines the database's tables, how it is organized, and how the relations are associated. Using PostgreSQL, an open-source relational database system, we constructed and imported all the BTCs and associated data into the new database. Then, the undergraduate student I hired through this grant constructed a user interface (website) that allows users to filter the data for download and use. This allows for this research to incorporate ICON (Integrated, Coordinated, Open, and Networked) and FAIR (Findable, Accessible, Interoperable, and Reusable) principles through allowing for individuals to access the data for use in further research. The code for the UI is also publicly available on GitHub at https://github.com/aabusang/cwe.

5. Description of results; including findings, conclusions, and recommendations for further research.

We have completed the database and use HydroShare as the current public repository of database files and the location from which we will pull data for use with our User Interface (UI) (Figure 1). The UI provides information regarding the project and search and download features. The data can be filtered using dropdown menus of data types/features. The dropdown menus provide filtering by location, hydrologic, and geomorphic characteristics. Advanced filtering options will be by tracer type, geo-feature, flow rate, and other specific criteria (Figure 2). Filtering of data is enabled by classes and functions that enable class extraction based on specified filtering criteria. For example, the 'TimeSeries' class allows filtering and retrieving data from the database by parameters such as river name, tracer type, geo-feature, and others. Filtering occurs on the primary interface, allowing users to input search criteria and view results that match their criteria. Further work includes creating a feature allowing users to input their datasets, automatically provide a quality assessment of the data, and then standardize their data so that it can be integrated into the database. This database may further be used for models, such as machine learning and physics-informed neural networks, to predict the BTCs that occur in a river using available site characteristics. Due to the standardization and aggregation of data, this research will allow machine learning and neural networks to find unique relationships that humans cannot recognize or understand due to the overwhelming amount of data involved and make the data accessible to the wider public and research community.

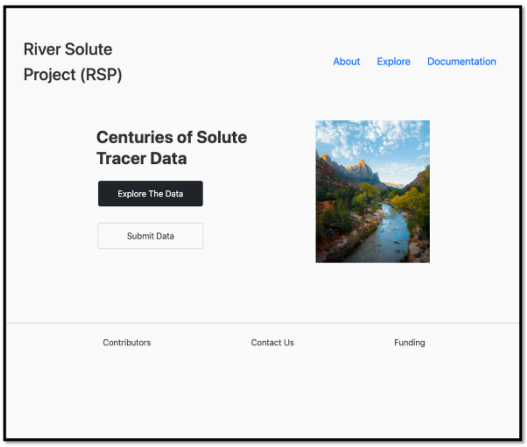

Figure 1. Front page for the website of the User Interface.

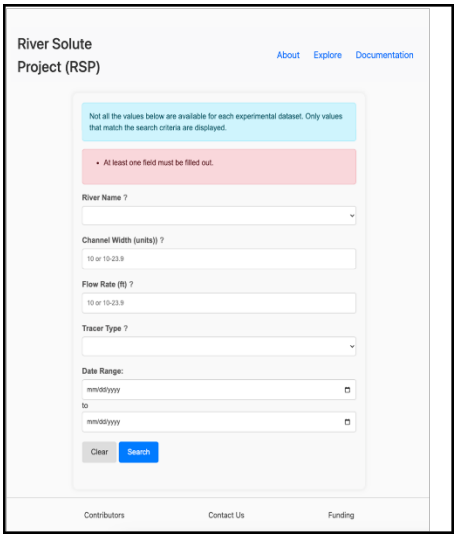

Figure 2. Filter and data download screen for the database.

6. Provide a paragraph on who will benefit from your research results. Include any water agency that could use your results.

Our work will benefit many stakeholder groups by creating a database that includes BTCs and associated hydrogeomorphic measurements. We will benefit the scientific community by providing a way for scientists to contextualize their experiments on a broader scale and

explore potential scaling relationships, similarities, and differences in river corridors across latitudes, ecological biomes, stream orders, and contrasting discharge. A public database will also improve the accessibility of scientific findings for the research community and the public, leading to improved resource management. Developing new solute transport models using this database will allow for better predictions of water flow using available site characteristics. It can improve forecasting of water flows and disturbance propagation that can be used by local water utilities that rely on surface water sources for their drinking water.

7. Describe how you have spent your grant funds. Also provide your budget balance and how you will use any remaining funds. If you anticipate any funds remaining after December 14, 2023, please contact Carolina Mijares immediately. (575-646-7991; mijares@nmsu.edu)

   The grant funds are being used to pay the salary of an undergraduate student from the Computer Science Department of the University of New Mexico, who was hired on May 15, 2023. He has provided immense support for the completion of the objectives of this grant, as he has experience in website development and working with databases. Additionally, funds were used for conference registration and poster printing to share this work at this year's NM Water Conference in Albuquerque, NM.

Costs and Purchases	Amount
Salary	
Undergraduate Student Salary to Date	$7,264.24
Other Staff Benefits Gen	$ 7.76
Conference Fees Gen	$150.00
Poster Printing Gen	$78.00
Total grant remaining	= $0.00

8. List presentations you have made related to the project.
   A poster was presented at the 68th Annual NM Water Conference.

9. List publications or reports, if any, that you are preparing. For all publications/reports and posters resulting from this award, please attribute the funding to NM WRRI and the New Mexico State Legislature by including the account number: NMWRRI-SG-FALL2022.

10. List any other students or faculty members who have assisted you with your project.

    Adam Abusang is an undergraduate student in the Computer Science Department at the University of New Mexico hired to work on this grant. Duncan Gardner is an undergraduate student in Civil Engineering who worked on populating the current database with BTCs and hydrogeomorphic measurements. He added ~200 additional sheets to the existing database.

11. Provide special recognition awards or notable achievements as a result of the research including any publicity such as newspaper articles, or similar.

12. Provide information on degree completion and future career plans. Funding for student grants comes from the New Mexico Legislature and legislators are interested in whether recipients of these grants go on to complete academic degrees and work in a water-related field in New Mexico or elsewhere.

    The student researcher is working on her Ph.D. with an intended graduation date of May 2025. After graduation, she plans to obtain a professional career with the US Geological Survey or at one of the National Laboratories researching the impacts of anthropogenic and natural disasters on water supplies. She would like to create tools that can forecast changes in water quality to ensure that communities relying on rivers for drinking water have safe water supplies.