**Tarrik Quneibi** **(517) 206-8821** **tarrikq@umich.edu** **Portfolio**

## Education

**University of Michigan**
Bachelor of Science in Environmental Engineering
Ann Arbor, Michigan 48109
Graduation: December 2021

**University of Michigan**
Master of Science in Environmental Engineering
Ann Arbor, Michigan 48109
Graduation: December 2022

## Work Experience

**SustainaBase**
Data Engineer - Feb, 2023- Present

- Designed overall structure of storing and processing data from ingestion to visuals using various AWS tools such as S3, Lambda, Step functions, Glue, Athena, and Textract.
- Deployed pipelines for pre-processing (file specific), normalized transforms (schema specific), validation (checks and statistics), and error logging.
- Managed/optimized MySQL database operations including stored procedures, restructuring, and QA.
- Created various transformation and API connection modules to increase data ETL efficiency from weekly upload times to daily.

**University of Michigan**
Research Assistant/Project lead - Jan, 2022 - Feb, 2023

- Automated pipelines for real time updates to analysis report when entering new data during field work.
- Generated figures directly from data storage using API connections in R for reporting to stakeholders.
- Ran physicochemical and biological analysis on water samples throughout the distribution system.
- Analyzed DNA and water quality parameters before, and after UV disinfection.
- Managed project scheduling, task delegation, equipment implementation.

## Skills

| Technical: | Knowledge: | Modeling: |
|---|---|---|
| Amazon AWS (lambda, glue, S3, etc.) | Database Management/architecture | EPAnet |
| SQL | Water Resources | SUMO |
| Python | Carbon Accounting | Google Earth Engine |
| RStudio/RShiny | API connections | Modflow |
| Serverless | Machine Learning | ArcGIS |

## Projects

Prediction of harmful algal bloom concentration using deep learning neural network

- Scraped Lake Erie water quality data taken from buoys using Seagull.
- Performed multiple imputation via Amelia to fill in missing values for complete datasets.
- Ran Spearman correlation test to determine magnitude of correlations between individual water quality parameters and cyanotoxins concentrations.
- Determined feature importance using random forest classification to ensure model optimization.
- Built/trained deep learning neural network with portion of real data resulting in model predicting cyanotoxin concentration with around 89% accuracy.

California ground-level ozone analysis, modeling, and prediction

- Scraped ground level ozone concentrations from the California Air Resources Board.
- Converted tabular data to a spatial matrix allowing creation of interactive spatial maps, density plots, and 3-dimensional surface maps.
- Trained gradient boosted decision tree models to predict ground level ozone concentrations using various air quality parameters and solar irradiance resulting in performance of 90% accuracy.