QualityAirQualityCities Specs

Product Owner	Minneapolis Health Department (MHD)/ University of Minnesota	
Project Manager	Jake Ford	ford0161@umn.edu
ArcOnline Lead	Taylor Andersen-Beaver	and03449@umn.edu
PurpleAir & MHD Liaison	Rob Hendrickson	hendr862@umn.edu
Required Approvers	Jenni Lansing, MHD Senior Environmental Scientist Kelly Meulman, MHD Sustainability Program Coordinator Noah Bowlin, GreenCorps	Jenni.Lansing@minneapolismn.gov kelly.muellman@minneapolismn.gov noah.bowlin@minneapolismn.gov

Overview

Objectives
- Build a Full Stack Modeling System that:
  a. Streams near real-time air quality (AQ) data through the PurpleAir API for the 70 Community Air Monitoring Project sensors in Minneapolis to:
     - Open-source and ESRI data formats
     - A locally developed and maintained postgres/postGIS database of historical values
     - An automated quality assurance and quality control (QAQC) procedure to:
       - Generate Tier 0 (raw data), Tier 1 (calibrated data), and Tier 2 (QAQC data), Tier 3 (spatio-temporal interpolations and zonal statistics)
  b. Analyzes and models:
     - Historic trends in AQ by:
       - Multiple spatial Units
         - Individual Sensor
         - 50 meter resolution interpolated raster surface
       - Multiple Temporal Units
         - Times of day (rush hour, recess, workday)
         - Days of week (weekday vs. weekend, holidays)

- ○ Seasonally
- ● Spikes
  - ○ Time and duration
  - ○ Spatial extent
- ■ Concurrent weather data as a interactive layer
- ■ Potential Sources (industry/Traffic)
  - ● MPCA Data
- ■ Limitations of Data
- c. Automates and serves visualizations to ArcOnline of:
  - ■ Near real-time point, polygon, and raster representations of AQ
  - ■ Historic trends in AQ at different spatio-temporal units
  - ■ Uncertainty and limitations of interpolated results for comparative analysis
- d. Shares data through:
  - ■ A REST API
- e. Results in a presentation for the general public

Problems QAQC Will Address
- ● **Data Storage** -- Infrastructure requirements and maintenance costs can be large for long term real time data projects. Currently, principal investigators are coordinating a partnership between the GEMS Lab at the University of Minnesota and the City of Minneapolis.
- ● **Community Accessibility –** Currently PurpleAir Data is available through PurpleAir's API, requiring technical knowledge to access and interpret. This creates a barrier between communities hosting the sensors and the information they generate. No system yet exists to interpret and display this data.
- ● **Ease of use** -- PurpleAir Generates real time data in large quantities accessible through their proprietary API which can be difficult to navigate for users who want access to discrete blocks of specific time periods or aggregated data.
- ● **Data Integration** -- Currently there is no system to collect data from the Area of Interest (AOI) and display it with data from other related sources like MPCA pollution data, Traffic Volume and weather data.
- ● **Data Analysis** -- While the source data is available, no system is yet in place to conduct analysis on the data to create useful information for the communities impacted by the AOI's AQ and those involved in collection.
- ● **Data Visualization** – In depth analysis on near real time data requires centralized hub for visualization which does not yet exist for the AOI

Who are the stakeholders
- ● Community of Minneapolis and residents of Green Zones
- ● Citizen Scientists who have installed sensors
- ● Policy makers in City Government, MDH
- ● Data scientists, data reporters, professionals, and researchers

The raw data is out there and none of the products necessary for storage and analysis exist. Turning data into useful information requires robust and iterable programmatic analysis. In order for the PurpleAir data project to have the most impact these resources must be developed and must be publicly available, accessible and user friendly. Quality Air Quality Cities is passionate about geospatial data science and environmental justice, and is well suited to develop the required full stack web application.

Motivation

The unequal spatial distribution of poor air quality across the City of Minneapolis has come under increased scrutiny. Particulate Matter 2.5 (PM2.5) is one of the major pollutants of concern. The American Heart Association has established a causal link between PM2.5 particles and heart and lung disease. These impacts manifest themselves in missed days of school and work, hospitalizations, and premature mortality.

In 2021, the Minneapolis Health Department (MHD) launched a community air monitoring project to help better understand these disparities. Beginning in 2022, PurpleAir sensors were installed by volunteers throughout the city with a focus on the Green Zones. MHD will be our primary design partners.

Now that most of the 70 PurpleAir sensors are collecting data, there is a need to store this information locally in a database that is accessible to scientists, city workers, and developers. Upon creation of this data architecture, spatio-temporal analysis will be conducted to yield insights into the air quality in Minneapolis. Finally, approachable visualizations and summary stats will be served to a web map for the benefit and use of community members, city officials, health professionals, and policy makers.

Definitions

Particulate Matter 2.5 (PM2.5): generic particles of 2.5 microns (micrometers) in diameter created during the combustion processes of cars, energy production, manufacturing, and trash incineration ([source](#)).

Green Zones: Formally recognized by the city of Minneapolis, these are areas of interest in Minneapolis with high levels of environmental pollution and racial, political, and economic marginalization ([source](#)).

PurpleAir: A manufacturer of low-cost, real time PM2.5 sensors that also serves their collected data through an API. Founded on principles of openness, sharing, and community ([source1](#) - [source2](#)).

Application Programming Interface (API): The programming level protocol for interacting with remote servers. ([source](#))

Quality Assurance and Quality Control (QAQC): Best practices for monitoring data quality from collection to implementation.
([source](#))

Graphical User Interface (GUI): The primary user facing, user friendly interface for a piece of software or a web app.

Scope

Functional Requirements

Essential

- Store PurpleAir air quality data
- Interactive ArcOnline Webmap
- Report for MHD

Nice to have

- Layers of analysis related to air quality (weather, traffic, potential sources, etc)

Non-Functional Requirements

Essential

- Usability
- Access to PurpleAir Historical Data
- Accurately referencing PurpleAir products/functions
- ArcGIS Enterprise account access

Nice to have

- Different interfaces for the public vs researchers/scientists
- Accessibility testing

Optional Priorities

- Two separate interfaces for the public vs researches

Out of Scope Requirements

- Additional types of air quality sensors for more robust data capturing
  - AQ Mesh to capture VOCs
- GUI for Data Downloading
- Additional layers of analysis including but not limited to
  - Weather/climate impact
  - Traffic patterns
  - Industry markers
- Forecasting
- Automated reports for users in a dashboard format
  - categorical AQ risk levels per Air Quality Index standards
  - real time hotspots and problem areas
  - weekly/monthly/yearly top impacted areas

Persona Acceptance Criteria

As a developer I …
- Require clear communication so that I understand the purpose of the project
- Require continued adherence to the open source licensing agreement so that I can reproduce this product in the future

As an operator I…
- Require sustainable code so that I can manage the incoming bugs
- Require proper data architecting so that I can efficiently manage data

As a member of the city I…
- Require usability so that I can interpret the data that is being presented
- Require accessibility so that I can consult others about my concerns

As a researcher I…
- Require sufficient, reliable data so that I can study air quality cause and effect

Open Questions

What are assumptions we are making and known risks with respect to the feasibility of a project? This can be with respect to licensing, staffing, or how a particular requirement will be achieved. All open questions must be addressed by the design stage.
- Will the final database be hosted by GEMS at the University of Minnesota or on City of Minneapolis based servers?
- Will we develop the project on the ArcEnterprise License of the City or as a Partnership between the City and the U of M?

- Will we have guaranteed access to the data and the accompanying City of Minneapolis resources for the duration of the project build
- Should we include ALL of the active sensors currency in Minneapolis or only the 70 Community Air Monitoring Project sensors?
- Who will monitor and manage this project long-term?
- Is a single interface functional for all end users (public, researchers, policy makers)?
- How do we incorporate legitimate policy and health information to the results?
- Do we need to be concerned about the security of the PurpleAir data?

Dependencies

What dependencies does this specification have on other projects, components, or software? All dependencies must be highlighted here as every dependency is a potential risk.

- Access to the City of Minneapolis' ArcGIS Enterprise
- PurpleAir Sensor Hardware
- Access to the PurpleAir data storage
- Access to GEMS Laboratory resources and data storage infrastructure
- ESRI products:
  - ArcGIS Pro
  - ArcGIS Online
  - ArcPy Python Library
- Open Source Python Libraries

References

This subsection should provide a complete list of all documents referenced elsewhere in the SRS; This information may be provided by reference to an appendix or to another document.

[1] "API - PurpleAir." https://api.purpleair.com/ (accessed Feb. 08, 2023).

[2] "Esri QAQC Presentation Slides."

https://www.esri.com/content/dam/esrisites/en-us/about/events/media/UC-2019/technical-works

hops/tw-6391-1016.pdf (accessed Feb. 08, 2023).

[3] C. of Minneapolis, "Green zones."

https://www2.minneapolismn.gov/government/departments/health/sustainability-homes-environ

ment/sustainability/green-zones/ (accessed Feb. 08, 2023).

[4] "HPHC_factsheet_AirPollution.pdf."

https://hpforhc.org/wp-content/uploads/2021/11/HPHC_factsheet_AirPollution.pdf (accessed

Feb. 08, 2023).

[5] "PurpleAir | Real-time Air Quality Monitoring," PurpleAir, Inc. https://www2.purpleair.com/

(accessed Feb. 08, 2023).

[6] "What Is an API Endpoint? (And Why Are They So Important?)."

https://blog.hubspot.com/website/api-endpoint (accessed Feb. 08, 2023).

Appendix
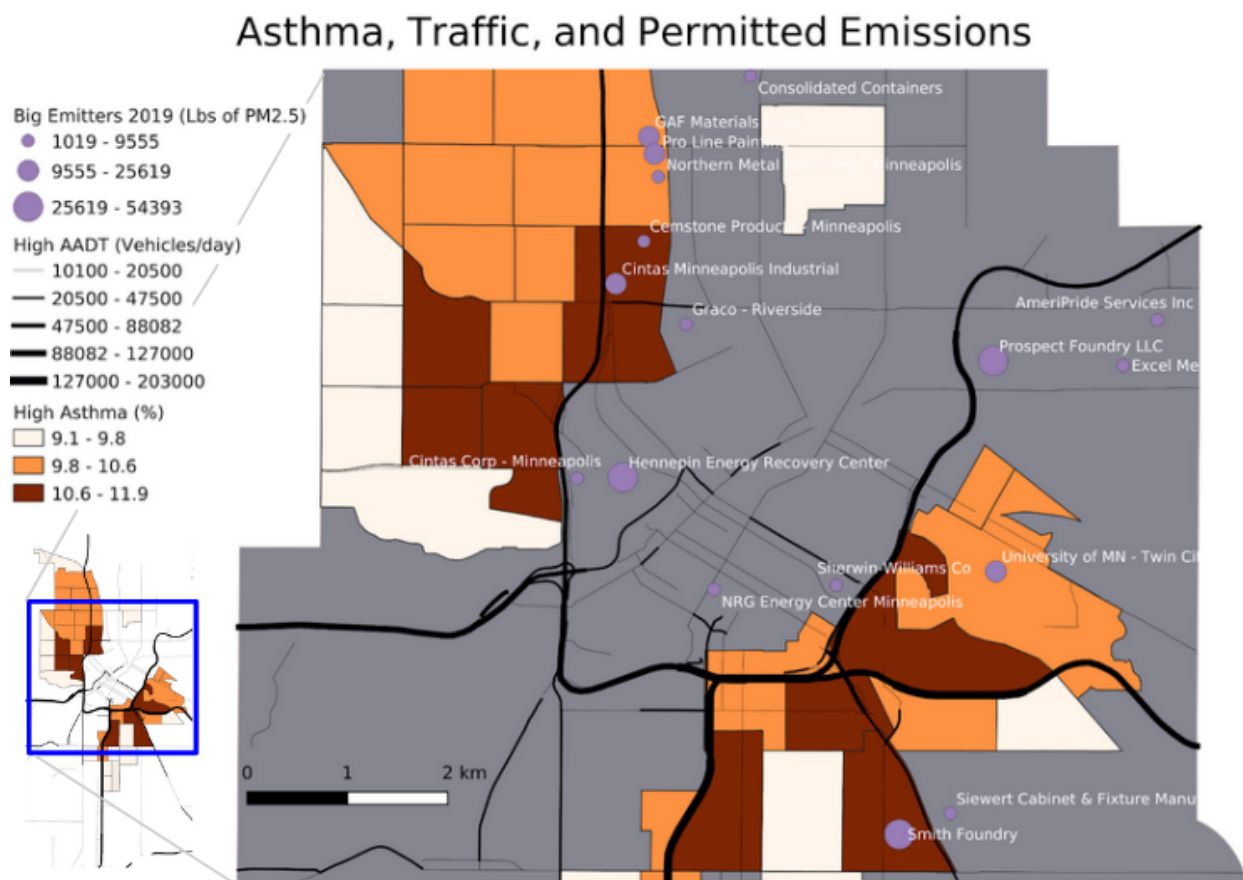
PurpleAir API Information: https://api.purpleair.com/

Example of a 6 Month Average PurpleAir Interpolation

Purple Air Interpolation