Open-Sourcing and Automating Air Quality Modelling

Anthony Walker

University of Glasgow

Colin Torney

University of Glasgow

February 2024**maps**

# Abstract

# Table of Contents

[**Abstract 1**](#_heading=h.avz9qzs55eu)

[**Table of Contents 1**](#_heading=h.vzjninsdind)

[**Introduction 2**](#_heading=h.ltm3a4tzczdc)

[Research Question 2](#_heading=h.unr6trgpbqev)

[Previous Research 2](#_heading=h.cwe5mp5nw95b)

[**Exploratory Analysis of Data Sources 2**](#_heading=h.kqvfl9mjsb41)

[OS Data Hub 3](#_heading=h.lvsnyg4zd48v)

[DEFRA air quality data - ukair/openair 3](#_heading=h.4cpt95z0tkfn)

[Meteorology - NOAA - worldmet. 3](#_heading=h.os02ax52lmvt)

[**Web Framework Development 3**](#_heading=h.zh5syg1n339y)

[**Methods and Functions Development 3**](#_heading=h.uq9jd7hmbff6)

[**Model Development and Validation Testing 3**](#_heading=h.w16rzazfp45)

[**User Testing 3**](#_heading=h.70w3apbs8vfh)

[**Deployment 3**](#_heading=h.yozabi7yfelf)

[**Conclusion 4**](#_heading=h.sf6jkd61udgi)

[**Bibliography 4**](#_heading=h.ftbifcccos5a)

# Introduction

## Research Question

**The project will involve the development of a Python based WebApp, that will be open-sourced, to allow users to conduct air quality concentration modelling without having to format and develop their own air quality modelling inputs.**

**Most current open-source air quality models rely on the users being able to understand and manually format a number of different geospatial, meteorological, and emission based data after retrieving them from a number of different sources.**

**The aim with this app is to automate the retrieval, formatting of a number of different air quality modelling inputs, as well as bespoke calculation of some specific geospatial characteristics such as a point specific canyon factor based on surrounding building and road characteristics.**

**These will then be used in the development of a number of different statistical modelling approaches to provide an air quality concentration output – using existing monitoring data available through APIs. This will allow a secondary exercise of exploring the best modelling approach given the input data with model performance analysis, and also ensure there is a useable output for the end user in the form of tabulated and visualised air quality data.**

Air quality modelling tools are often a myriad of input data files in varying format types and require an in-depth knowledge of the different parameters to be able to conduct a simple estimate of concentrations at a given point. Topography, meteorology, traffic, road layout, building layout, dispersion, all have an impact on air quality concentrations.[..ref..]

## Previous Research

# Exploratory Analysis of Data Sources

* + *Exploration of available open APIs/data sources*
    - *Python friendly HTTPS based APIs or other database integrations required?*
  + *Geospatial analysis of data sources to ensure sufficient coverage*
  + *Transformation of data required?*
  + *Summary statistics of underlying datasets – building heights/topography/road lengths/air quality data.*

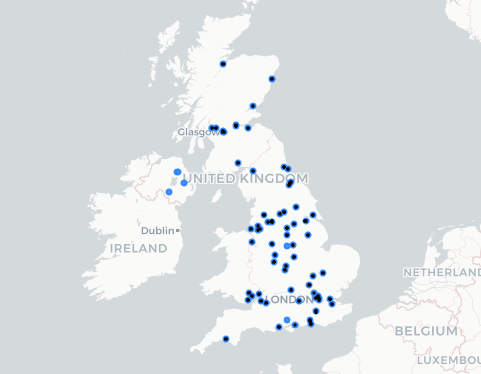
As stated, the aim of this project is to leverage available APIs and data sources to enable automated localised data gathering to enable formatted input data for air quality modelling. The following section will discuss the availability of API and data sources relevant to the project, their usage, an exploration of the data itself, and any transformations required to make it usable for the purposes of air quality modelling.

## DEFRA air quality data - ukair/openair

## OS Data Hub

### Building Data

One







## 

## Meteorology - NOAA - worldmet.

# Web Framework Development

* + Exploration of Webapp framework suitability
  + Develop initial webapp prototype.
    - Design/Sketch expected user interface.
    - Code and implement webapp

# Methods and Functions Development

* + Outline expected webapp functionality
    - Data integration methods
    - Data transformation methods
    - Data modelling methods

# Model Development and Validation Testing

* + Explore and implement statistical/machine learning models based on webapp transformed output data.
  + Compare model outputs to real world air quality monitoring data.

# User Testing

* + Implement user testing phase

# Deployment

* + Describe and manage app deployment and Continuous Integration plan.

# Conclusion

# Bibliography