RESEARCH ENGINEER PLAN

April 10, 2019

1 Current System

1.1 Downloading Waze

For more details Sec. 3.

We have been downloading Waze data for over one year. We are using a Linux VM on azure which has a simple python script that calls the waze API every minute and stores the raw json file in file storage. We have defined a download region that covers both Cambridge and London.

1.2 Simple Model

We are developing multi-resolution Gaussian process models.

1.3 Technologies Used

1. Feature Extraction

    (a) PostgreSQL + PostGIS
    (b) Python
    (c) Bash

2. Model

    (a) Python + Tensorflow

2 Envisioned System

2.1 Pipeline

See Fig. 1.

2.2 Requirements

1. Data Download

    (a) For each data source be continually download and store data from the feeds
    (b) Store both historic, observed and predicted values from the feeds
    (c) Data should be accessible for other projects

2. Pipeline

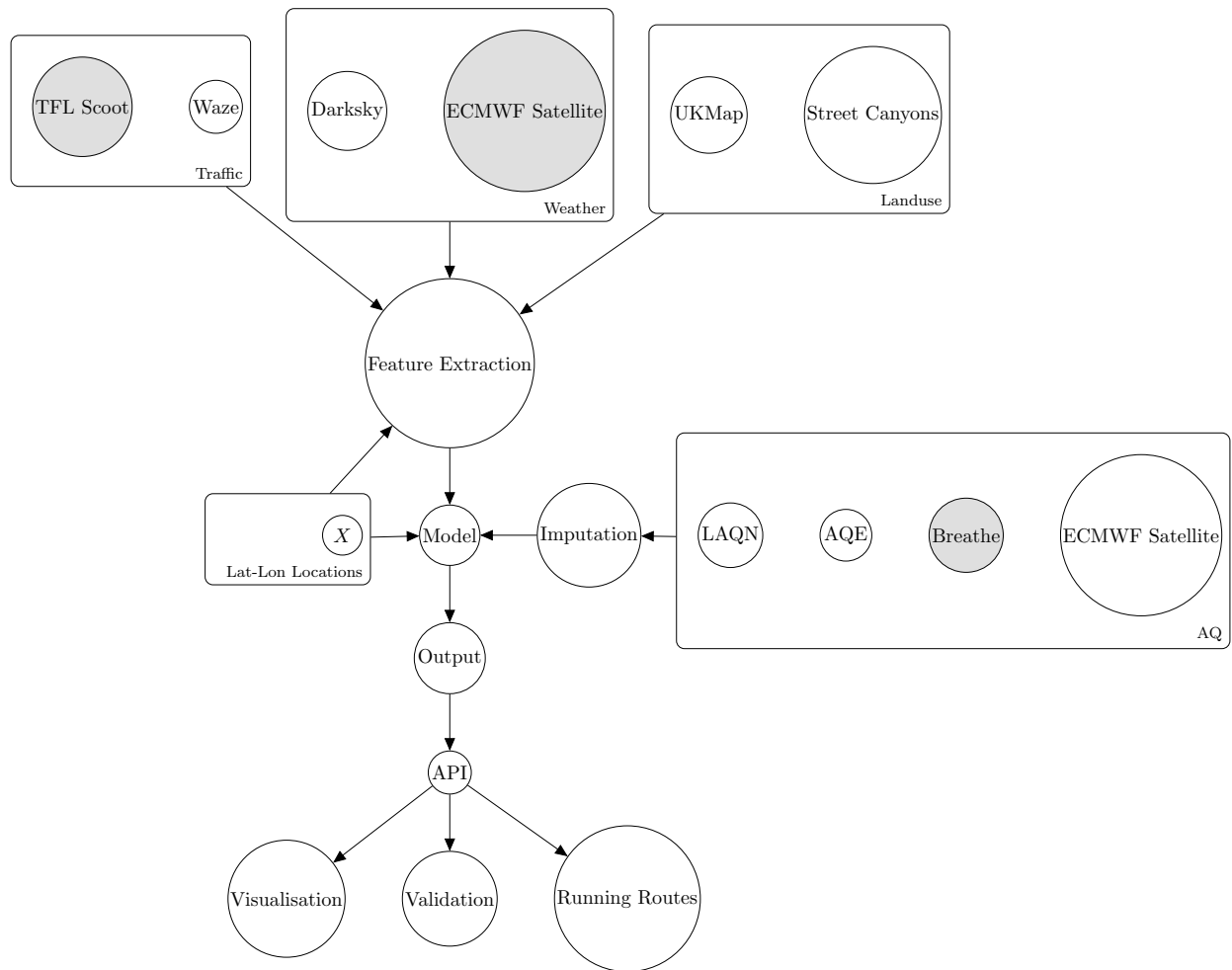   (a) Run the system everyday to provide 48 hour predictions across London

3. Pipeline Update

Figure 1: Pipeline of the required system. Datasets that are still to be obtained are shaded.

   (a) To be able to add data sources

   (b) To be able to update model - perhaps via production github branch

4. API

   (a) To access historic and current predictions for a given region

   (b) To provide simple queries

      i. Which road is the most polluted for a given region and time period

   (c) Provide $A \rightarrow B$ Routes which minimise air pollution

5. Visualisation

   (a) Website which provides simple visualisation of the current 48 hour prediction - see Sec. 4

3 Dataset Descriptions

3.1 Air Quality Data

- LAQN

  - Managed by Kings College London
  - 70-100 sensors at hourly resolution
  - Publicly available via `https://www.londonair.org.uk/Londonair/API/`
  - variables: NOx, PM10, PM25, O3. SO2 (not all sensors have all variables)

- Diffusion Tubes

  - Managed by the London Boroughs
  - 700 sensors at a monthly resolution
  - Publicly available by scraping the air pollution reports each borough has to publish each year
  - variables: NO2

- AQE (similar quality to LAQN)

  - Publicly available here: `http://www.airqualityengland.co.uk/local-authority/?la_id=999`
  - variables: NOx, PM10, PM25, O3. SO2 (not all sensors have all variables)

- ECMWF

  - Aggregation of lots of EU weather/pollution models - mechanistic models
  - Free and publicly available at `https://atmosphere.copernicus.eu/catalogue#/`
  - Provides historic data and at least 48h forecasts on a grid size of 7-10km by 7-10km
  - variables: NO2, PM10, PM25 (more available in the catalogue)

3.2 Traffic Data

- Waze traffic jam

  - every minute
  - provides location of traffic jams
  - We have been storing on Azure for about a year (minus a few months due to technical issues)
  - We are part of the connected citizen program (https://www.waze.com/en-GB/ccp) which gives access to the API

- TFL Scoot Data

  - every minute
  - provides flows/occupancy/counts of cars travelling over the induction loops on the roads
  - Data from around July 2018 to November 2018
  - TFL uploads the data to their amazon buckets
  - Through GLA we have access to these buckets

3.3 Weather Data

- Darksky

  - Has a free API (`https://darksky.net/forecast/51.4786,-0.158/us12/en`)
  - Provides 48 hours forecasts and historic data
  - variables: wind speed, wind direction, temperature, pressure etc

3.4 Landuse Datasets

- from GLA - not public

- Ordnance Survey (https://www.ordnancesurvey.co.uk

  - Building - Provide geometries for buildings in London
  - Highways - Provide geometries and classification of Londons road network
  - Background - Provide landuse for London

- UKMap

  - Similar to OSHighways and OSBackground combined

- UKBuildings

  - Similar to OSBuildings

4 Visualisations

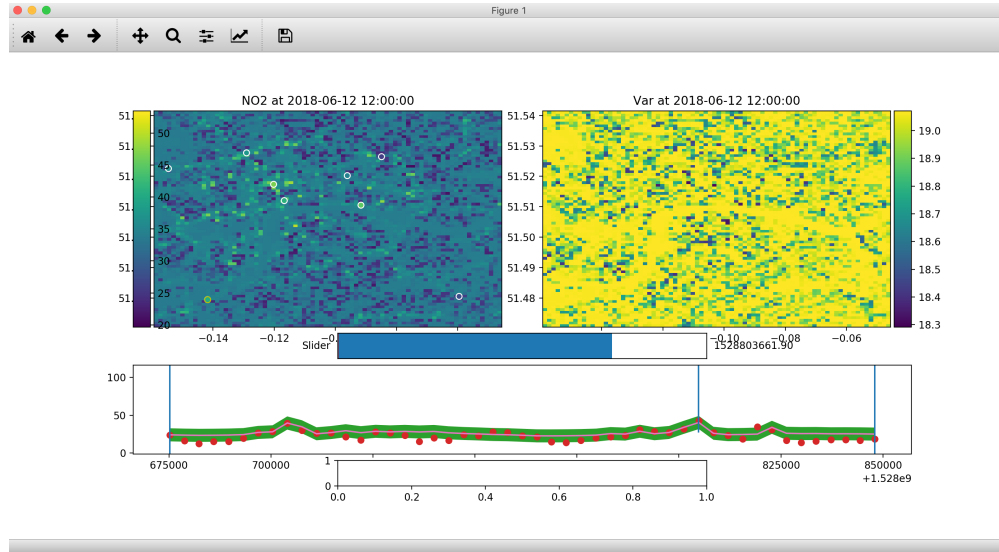### 4.0.1 Python Visualisation

4.0.2 Web Based Visualisation

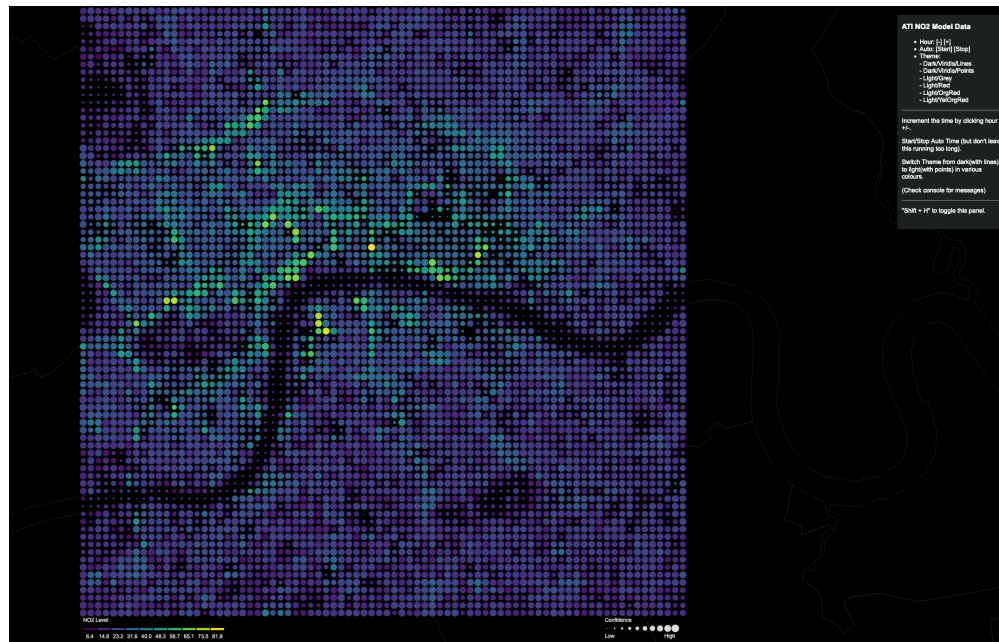Figure 2: Python based visualisation

Figure 3: Prototype web based visualisation developed by the GLA. The size of points shows the certainty of the predictions and the colors expresses level of NO2. See `https://gla-ati-no2-map.glitch.me/`.