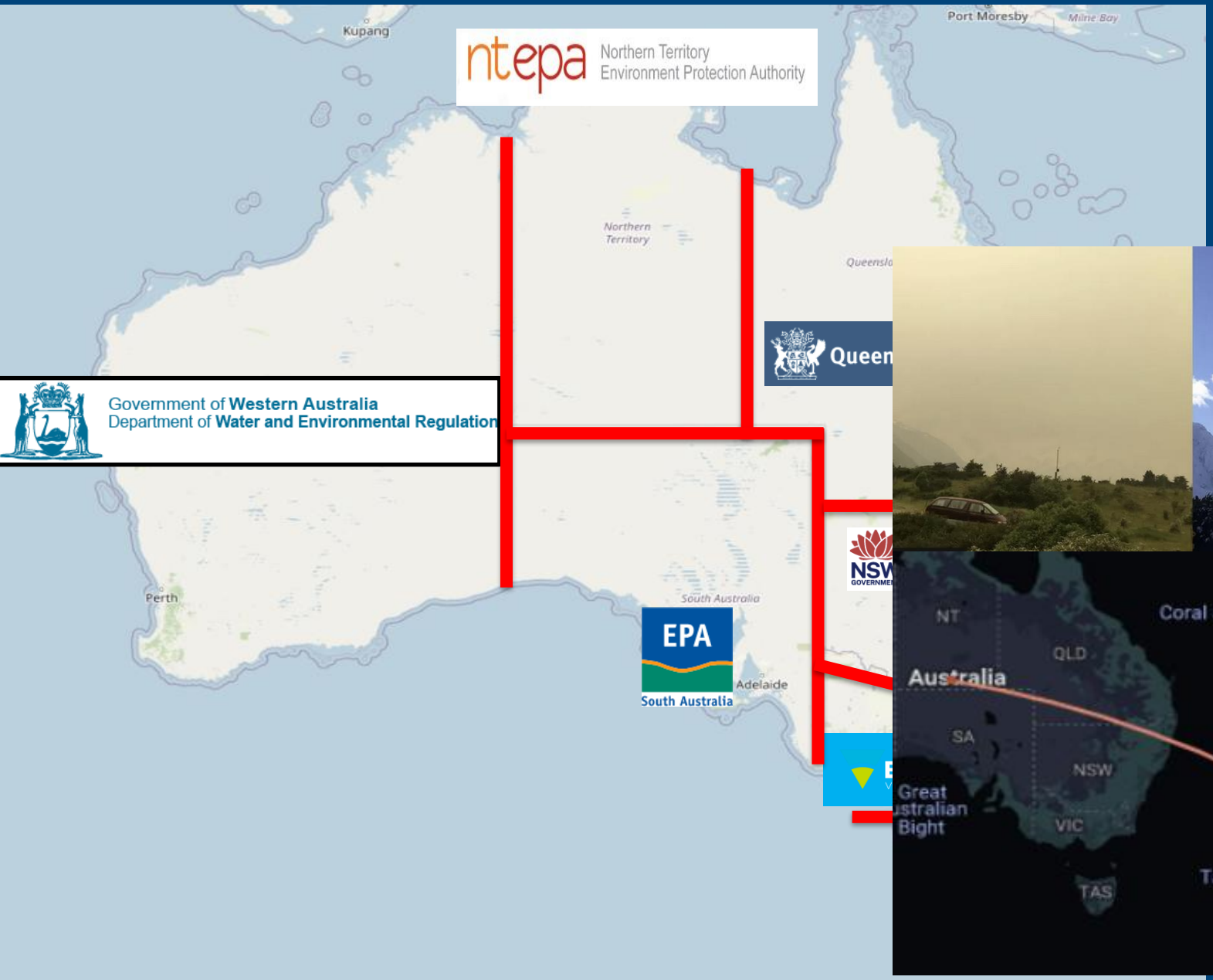
THE UNIVERSITY OF MELBOURNE

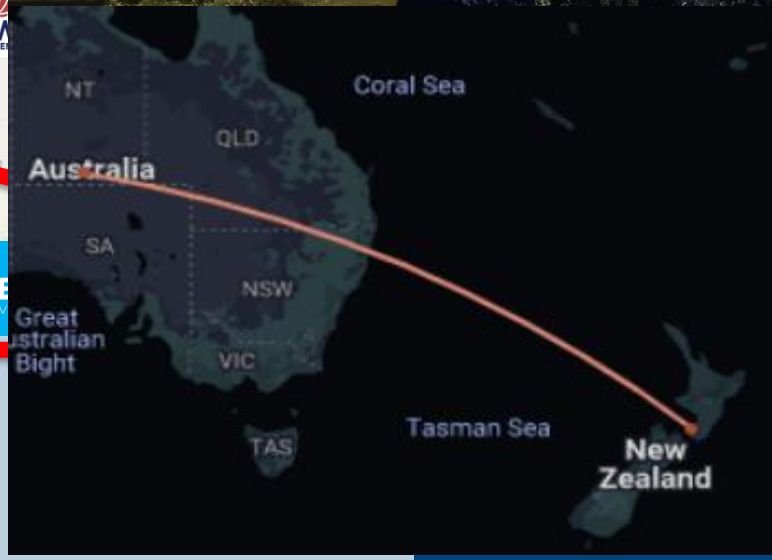
MELBOURNE RESEARCH

Establishment and Exploration of a National Air Quality Database for Australia

Prof Richard Sinnott, Philip Greenwood, Luca Morandini, Dr Hossein Pursultani, Assoc Prof Robyn Schofield, Robin Wan, Siqi Zhong
University of Melbourne
rsinnott@unimelb.edu.au

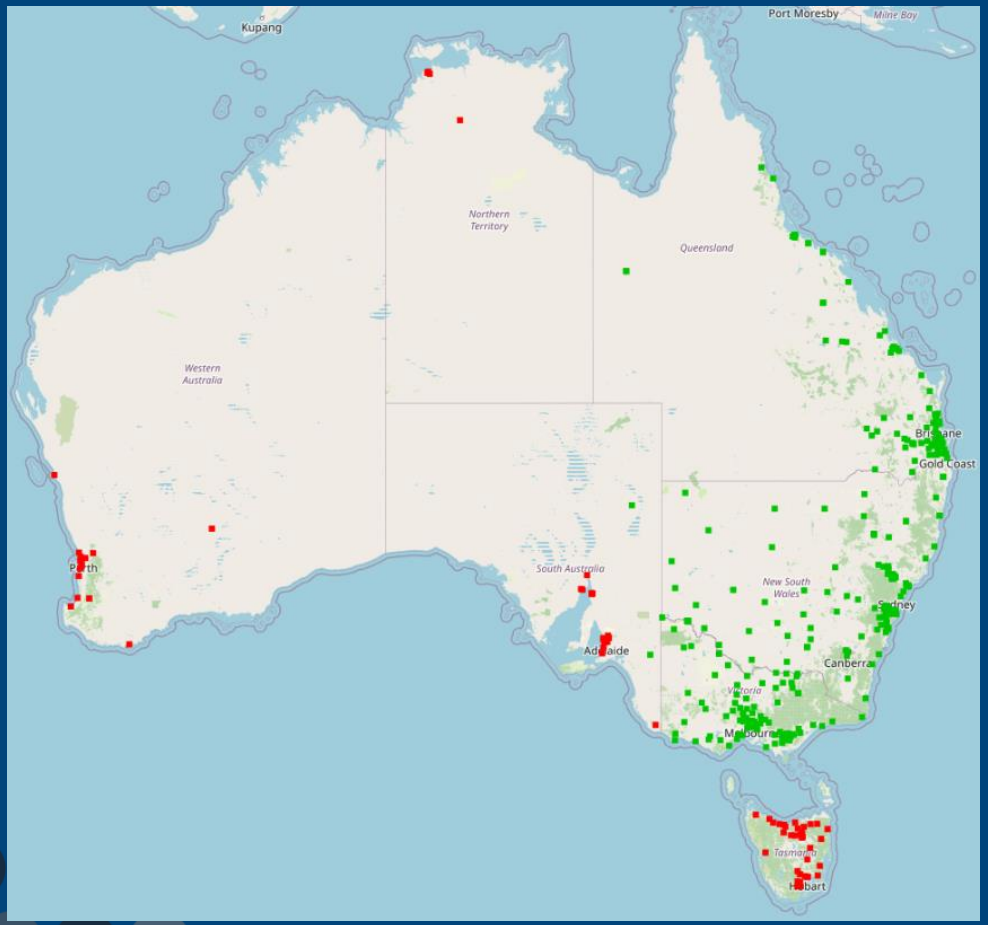
Air Quality in Australia





AURIN High Impact Project

- Establish National Air Quality Database (NAQD)
  - Programmatically access air quality data from official agencies (with API)
    - No web scraping

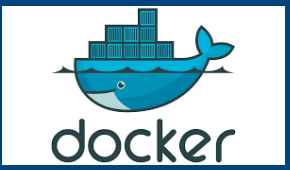


NAQD Data Diversity

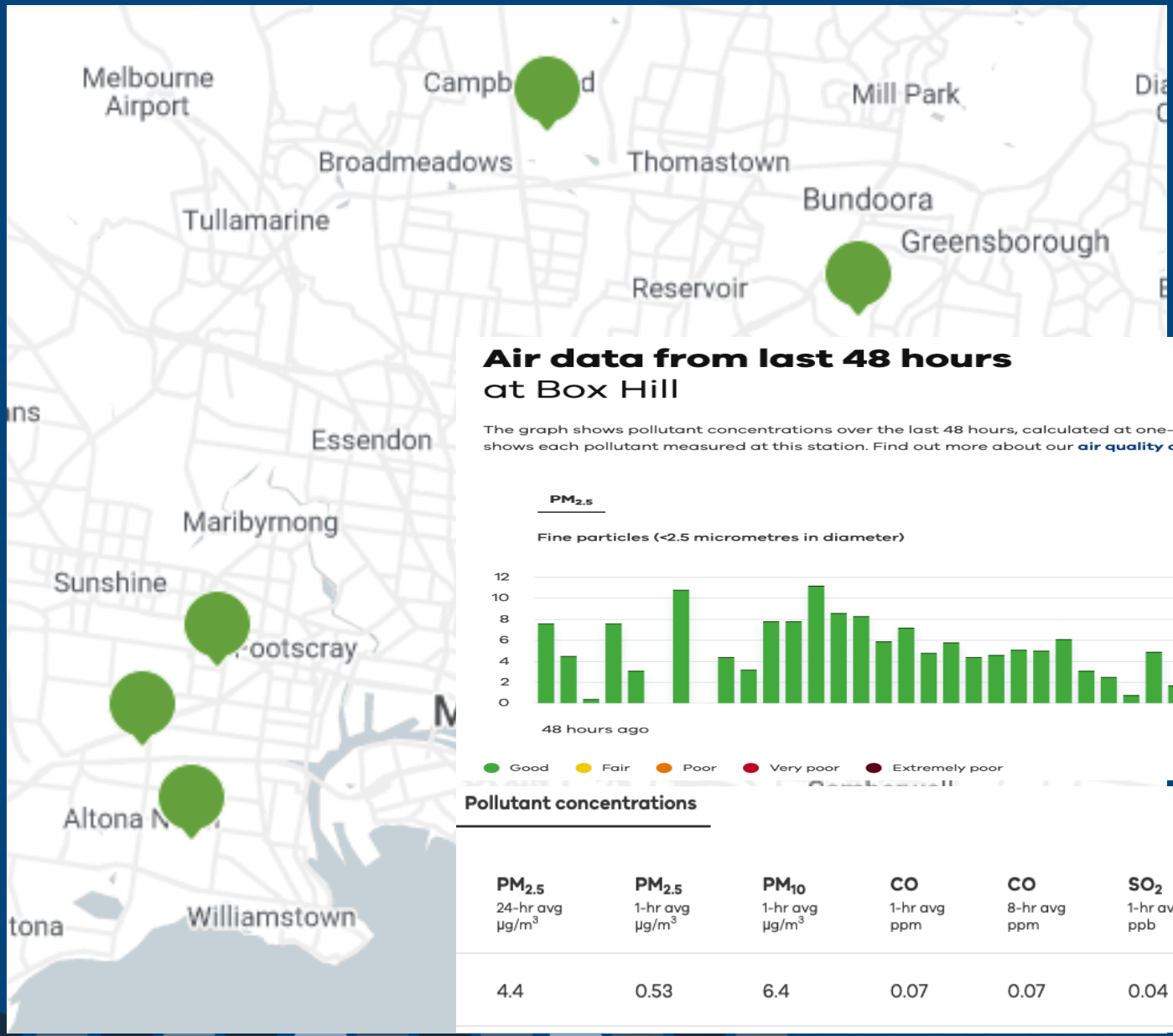
naqd_observation	observation_description	units	units_description	nsw	vic	qld	act
arsenic	Arsenic	µg/m³	micrograms per cubic metre			Y	
benzene	Benzene	ppb	parts per billion			Y	
cadmium	Cadmium	µg/m³	micrograms per cubic metre			Y	
carbon_m	Carbon Monoxide	ppm	parts per million	Y	Y	Y	Y
copper	Copper	µg/m³	micrograms per cubic metre			Y	
formaldehy	Formaldehyde	ppb	parts per billion			Y	
humidity	Humidity	%	percent	Y		Y	
hydrogen_s	Hydrogen Sulphide (30min)	ppb	parts per billion			Y	
lead	Lead	µg/m³	micrograms per cubic metre			Y	
neph	Nephelometer	$10^{-4}m^{-1}$	10 to the power of -4 per metre	Y			
ammonia	Ammonia	ppb	parts per billion	Y			
nitric_o	Nitric Oxide	ppb	parts per billion	Y			
nitrogen_d	Nitrogen Dioxide	ppb	parts per billion	Y	Y	Y	Y
nickel	Nickel	µg/m³	micrograms per cubic metre			Y	
ozone	Ozone	ppb	parts per billion	Y	Y	Y	Y
pm10	Particle PM10	µg/m³	micrograms per cubic metre	Y	Y	Y	Y
pm2p5	Particle PM2.5	µg/m³	micrograms per cubic metre	Y	Y	Y	Y
rainfall	Rainfall	mm/m²	millimetres per square metre	Y		Y	
sd1	Standard Deviation of Wind Direction	°	degrees (sigma theta)	Y			
sulphur_d	Sulphur Dioxide	ppb	parts per billion	Y	Y	Y	
solar	Solar Radiation	W/m²	watts per square metre	Y		Y	
temp	Temperature	°C	degrees Celsius	Y		Y	
toluene	Toluene	ppb	parts per billion			Y	
tspd	Particles TSP	µg/m³	micrograms per cubic metre	Y		Y	
wind_dir	Wind Direction	°	degrees	Y		Y	
wind_sp	Wind Speed	m/s	meters per second	Y		Y	
xylene	Xylene	ppb	parts per billion			Y	
zinc	Zinc	µg/m³	micrograms per cubic metre			Y	

NAQD Technology Stack

- Technology Stack
  - GeoNode
    - GeoServer
    - PostGIS
  - NeCTAR
    - Continuous data collection/processing/storage
  - Docker
    - Clients to access/clean data from agency APIs
  - Jupyter Notebooks
    - Deeper analytics via WFS endpoint

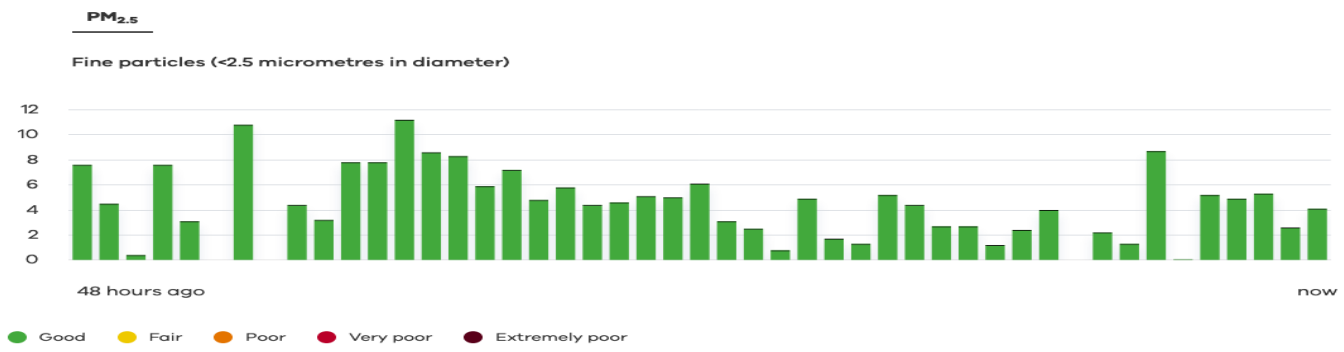


Air Quality around Melbourne



Air data from last 48 hours at Box Hill

The graph shows pollutant concentrations over the last 48 hours, calculated at one-hour averages. Using the labels above the graph, it shows each pollutant measured at this station. Find out more about our **air quality categories**.

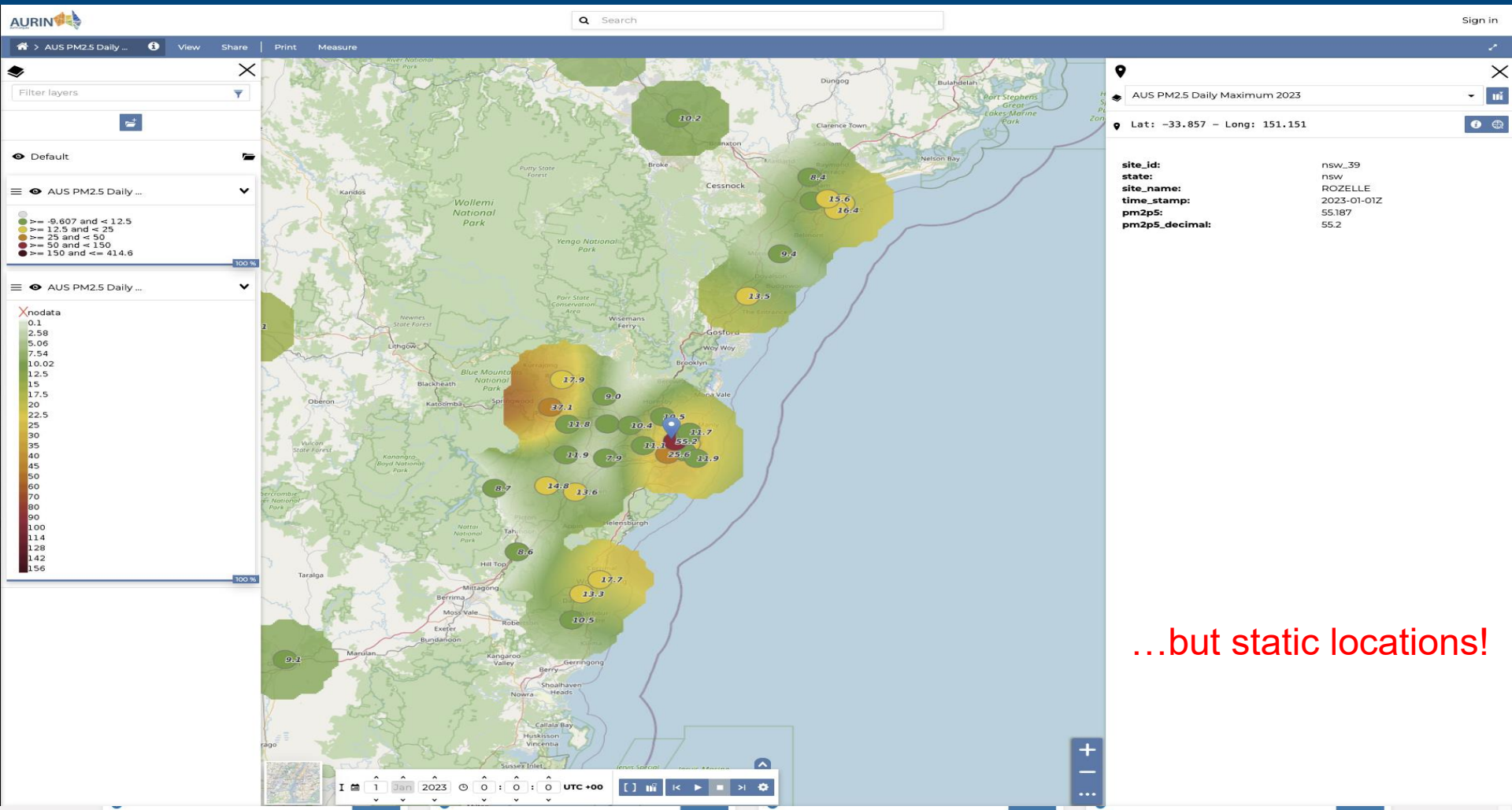


Pollutant concentrations

$PM_{2.5}$ 24-hr avg µg/m³	$PM_{2.5}$ 1-hr avg µg/m³	PM_{10} 1-hr avg µg/m³	CO 1-hr avg ppm	CO 8-hr avg ppm	SO_2 1-hr avg ppb	NO_2 1-hr avg ppb	O_3 1-hr avg ppb	Visibility 1-hr avg km
4.4	0.53	6.4	0.07	0.07	0.04	5	27	70.0

NAQD Deployment

- Web application demonstration
  - https://naqd.eresearch.unimelb.edu.au

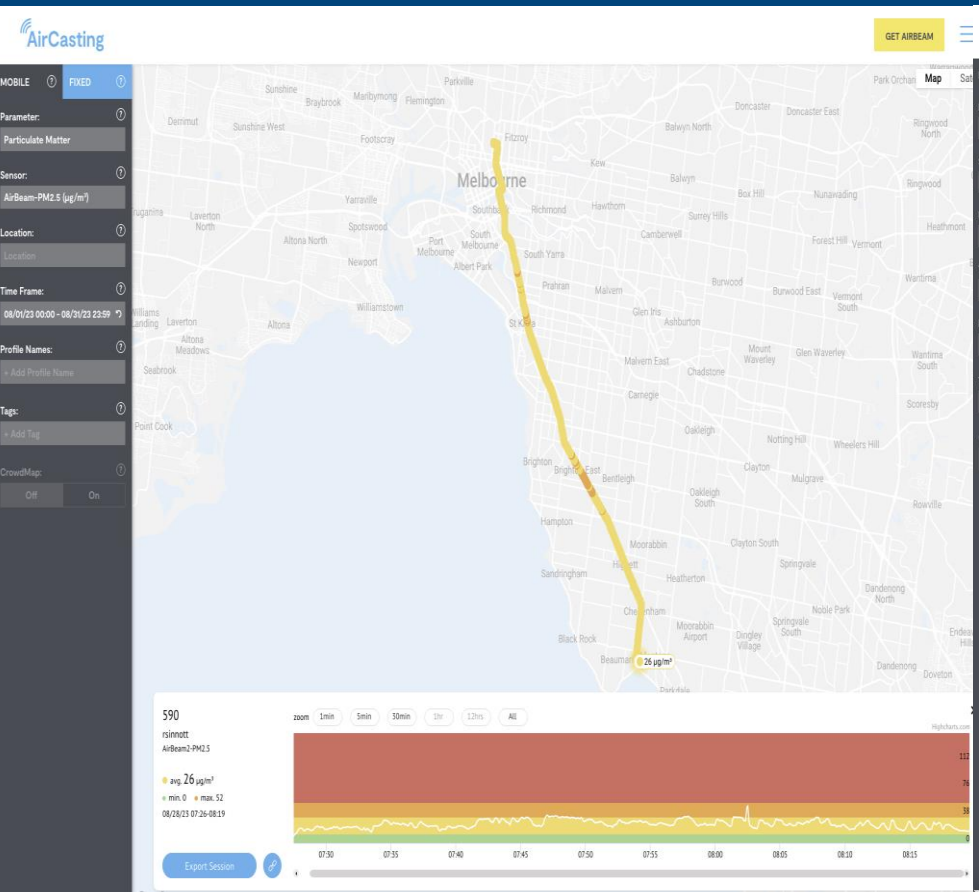


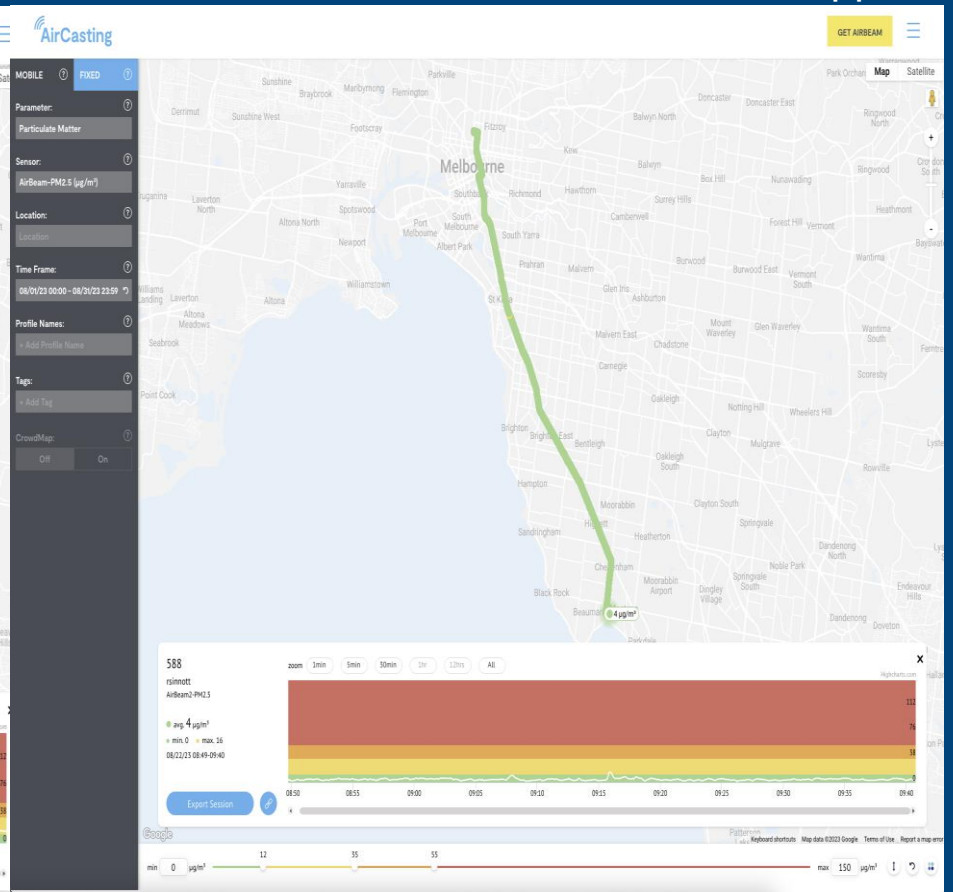
…but static locations!

Mobile Air Pollution Data

- Citizen Science
  - Airbeam Mobile Air Pollution Devices
    - https://www.habitatmap.org/airbeam

Uses Android app

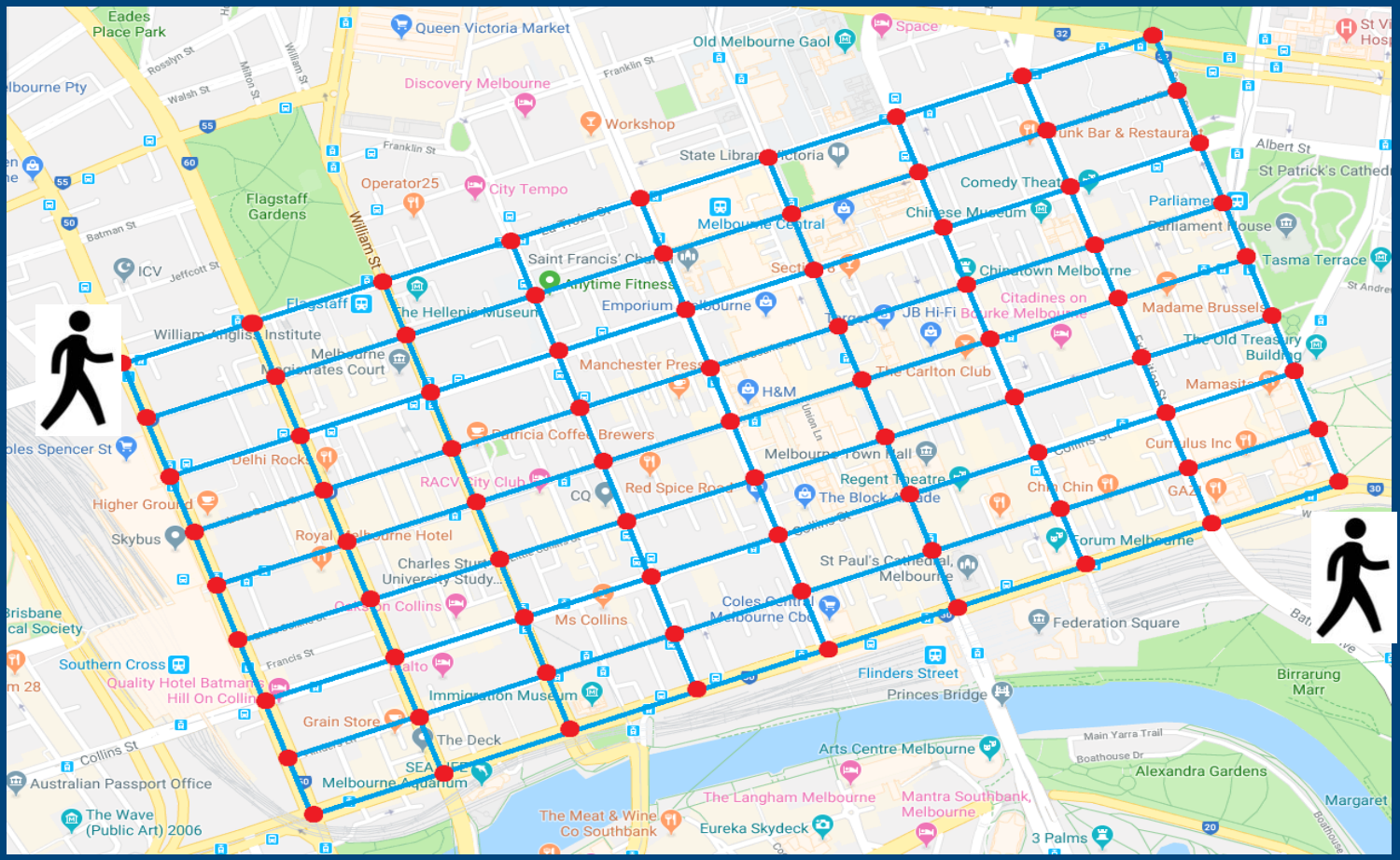




Mobile Air Pollution Data

	A	B	C	D	E	F
1	AirBeam2-PM2.5	AirBeam2:001896108040	Particulate Matter	micrograms per cubic meter	**Flinders St**	
2	Timestamp	geo:lat	geo:long	Value		
3	2019-01-20T16:49:21.217+1100	-37.81521493	144.9746181	2		-37.8152149321537,144.974618066912
4	2019-01-20T16:49:21.714+1100	-37.81521493	144.9746181	2		-37.8152149321537,144.974618066912
5	2019-01-20T16:49:22.712+1100	-37.81521457	144.9746185	2		-37.815214567,144.974618457
6	2019-01-20T16:49:23.709+1100	-37.81521593	144.9746188	2		-37.815215934,144.974618753
7	2019-01-20T16:49:24.707+1100	-37.81521713	144.9746168	2		-37.815217127,144.974616792
8	2019-01-20T16:49:25.704+1100	-37.81521974	144.9746169	2		-37.815219735,144.974616882
9	2019-01-20T16:49:26.702+1100	-37.81521974	144.9746169	2		-37.815219735,144.974616882
10	2019-01-20T16:49:27.699+1100	-37.81521974	144.9746169	1		-37.815219735,144.974616882
11	2019-01-20T16:49:28.697+1100	-37.81521974	144.9746169	1		-37.815219735,144.974616882
12	2019-01-20T16:49:29.694+1100	-37.81521974	144.9746169	1		-37.815219735,144.974616882
13	2019-01-20T16:49:30.692+1100	-37.81522603	144.9745866	1		-37.815226027,144.974586563
14	2019-01-20T16:49:31.690+1100	-37.81522791	144.9745764	1		-37.815227909,144.974576439
15	2019-01-20T16:49:32.687+1100	-37.81523108	144.9745658	1		-37.815231082,144.974565765
16	2019-01-20T16:49:33.686+1100	-37.81523443	144.974554	1		-37.815234431,144.974553993
17	2019-01-20T16:49:34.683+1100	-37.81523832	144.9745411	1		-37.815238323,144.974541087
18	2019-01-20T16:49:35.681+1100	-37.81524198	144.974527	2		-37.81524198,144.974526983
19	2019-01-20T16:49:36.678+1100	-37.81524587	144.9745111	2		-37.815245872,144.974511128
20	2019-01-20T16:49:37.675+1100	-37.81524985	144.9744963	2		-37.815249848,144.974496301
21	2019-01-20T16:49:38.673+1100	-37.81525322	144.9744819	2		-37.815253224,144.974481893
22	2019-01-20T16:49:39.671+1100	-37.8152559	144.9744661	2		-37.815255901,144.974466111
23	2019-01-20T16:49:40.670+1100	-37.81525849	144.9744495	2		-37.815258494,144.974449521
24	2019-01-20T16:49:41.667+1100	-37.81526071	144.9744329	2		-37.815260706,144.974432933
25	2019-01-20T16:49:42.666+1100	-37.81526222	144.9744178	2		-37.815262219,144.974417788
26	2019-01-20T16:49:43.661+1100	-37.8152636	144.9744028	2		-37.815263602,144.974402765
27	2019-01-20T16:49:44.659+1100	-37.81526613	144.9743869	2		-37.815266134,144.974386919
28	2019-01-20T16:49:45.657+1100	-37.81527009	144.97437	2		-37.815270092,144.974370025
29	2019-01-20T16:49:46.654+1100	-37.81527415	144.9743531	2		-37.81527415,144.974353098
30	2019-01-20T16:49:47.651+1100	-37.81527693	144.9743353	2		-37.815276933,144.974335332

Melbourne CBD Adaptive Routing



- Data Collection – students!!!

Data Collection

Routing Analytics Workflow

Feature

- Wind Direction
- Wind Speed
- Temperture
- Relative Humidity
- Pressure
- Weather Condition
- Traffic flow
- Official PM2.5

Label

PM2.5 for each street measured by AirBeam

Gradient boosted trees algorithm - attempts to accurately predict a target variable by combining estimates of a set of simpler, weaker models.

Train → XGBoost Model → Predict → PM2.5 concentration data for each street (C)

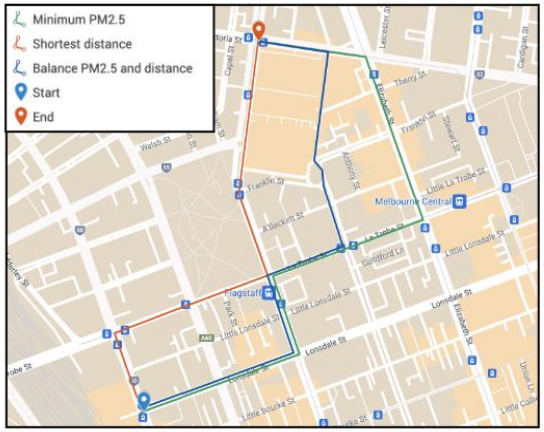
$$E_i = \frac{D_i}{V_j} \times R_j \times C_i$$

D = Road distance, V = Walking velocity,
R = Respiratory minute volume, C = PM2.5 Concentration

$$Weight_i = w \times \left(\frac{D_i - D_{\min}}{D_{\max} - D_{\min}}\right) + (1 - w) \times \left(\frac{E_i - E_{\min}}{E_{\max} - E_{\min}}\right)$$

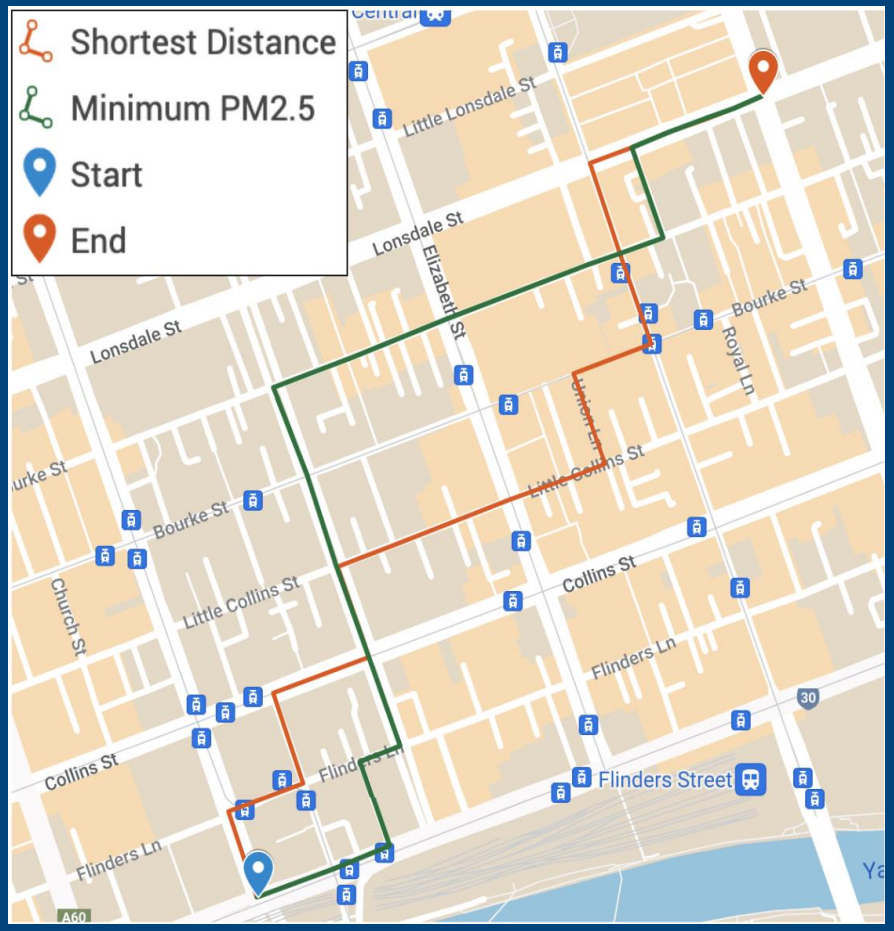
Dijkstra Algorithm

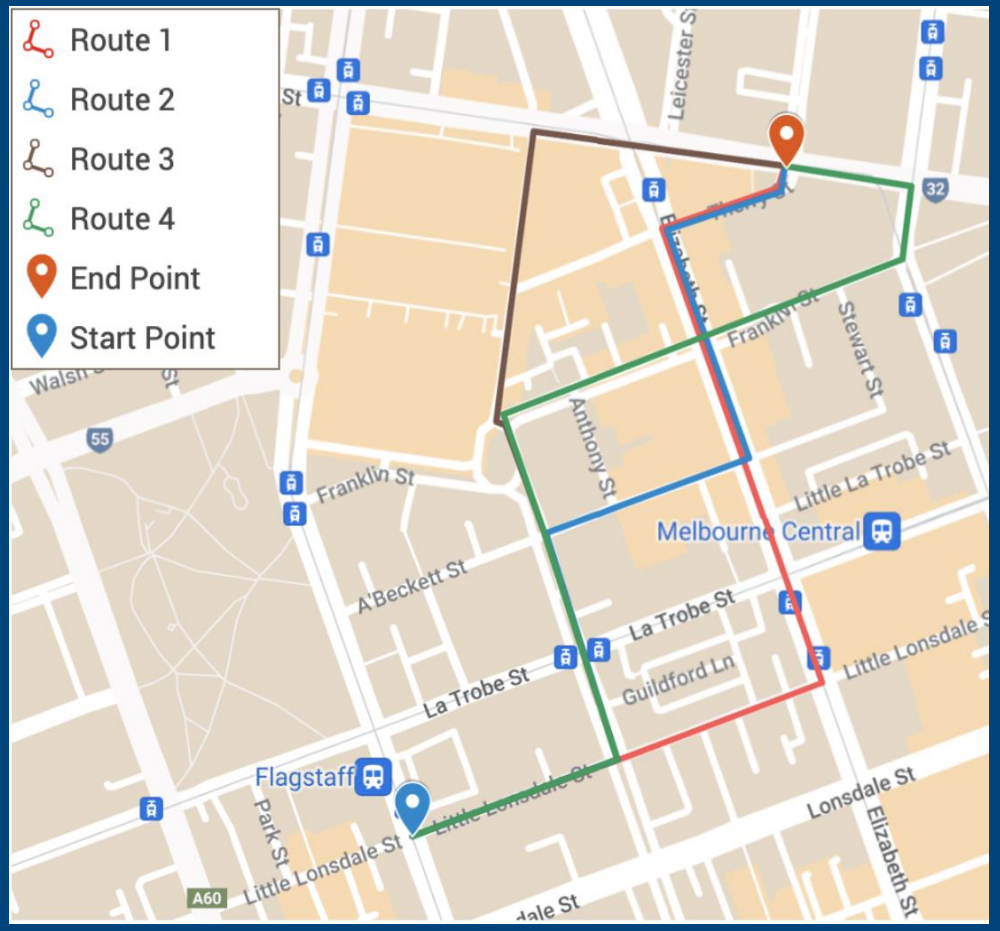
Trade-off PM2.5 exposure does and distance



Minimum PM2.5 exposure route and alternative option

Example Routing Scenarios





Route ID	Distance(m)	Distance Change	PM2.5(ug)	PM2.5 Change
Route 1	1109.84	-	0.83	-
Route 2	1110.79	0.09%	0.76	-8.43%
Route 3	1171.29	5.54%	0.73	-12.05%
Route 4	1219.83	9.91%	0.67	-19.28%

Conclusions and Future Work

- Scenario/data used for teaching
  - 400+ students of COMP90024
    - Will be supported indefinitely!
- Data feed to Bushfire Data Commons Dashboard
  - https://bdcpoc.eresearch.unimelb.edu.au
    - See other talk
  - Web service endpoint for data access (WFS)
- Evolution of air quality monitoring landscape