

'Mine the gaps' web application

A tool for dealing with geographically sparse datasets

Ann Gledson, Douglas Lowe, Manuele Reani, Caroline Jay, Dave Topping

The University of Manchester



#BritainBreathing

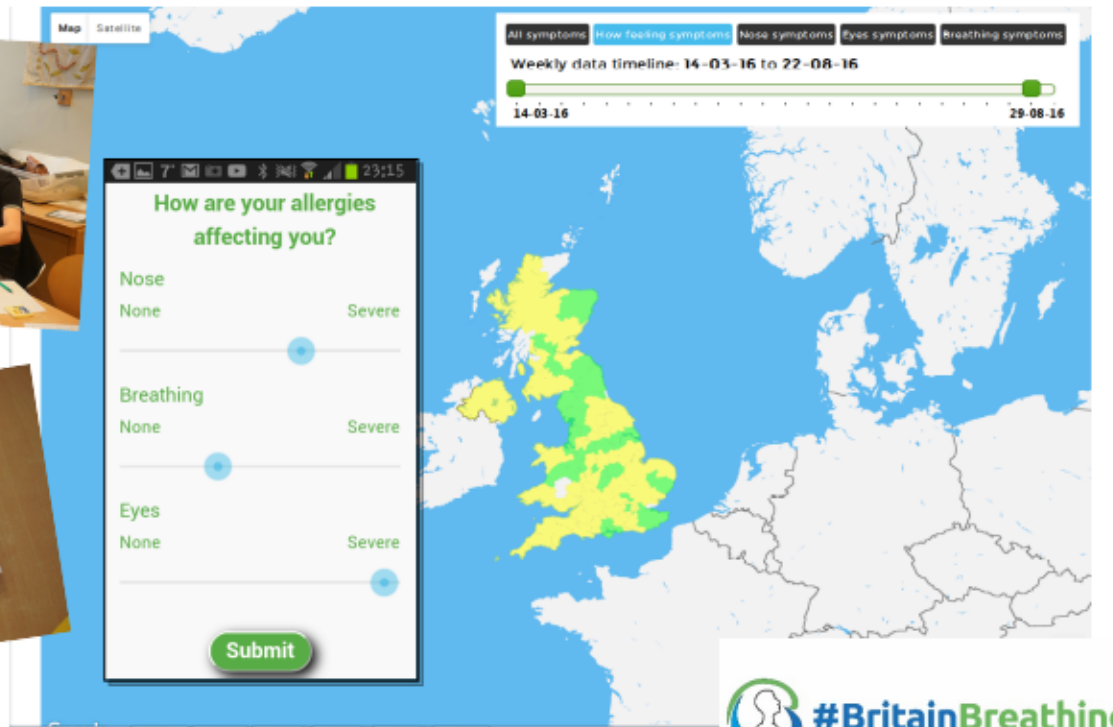
athing.org/index.php/data-visualisation/

Search



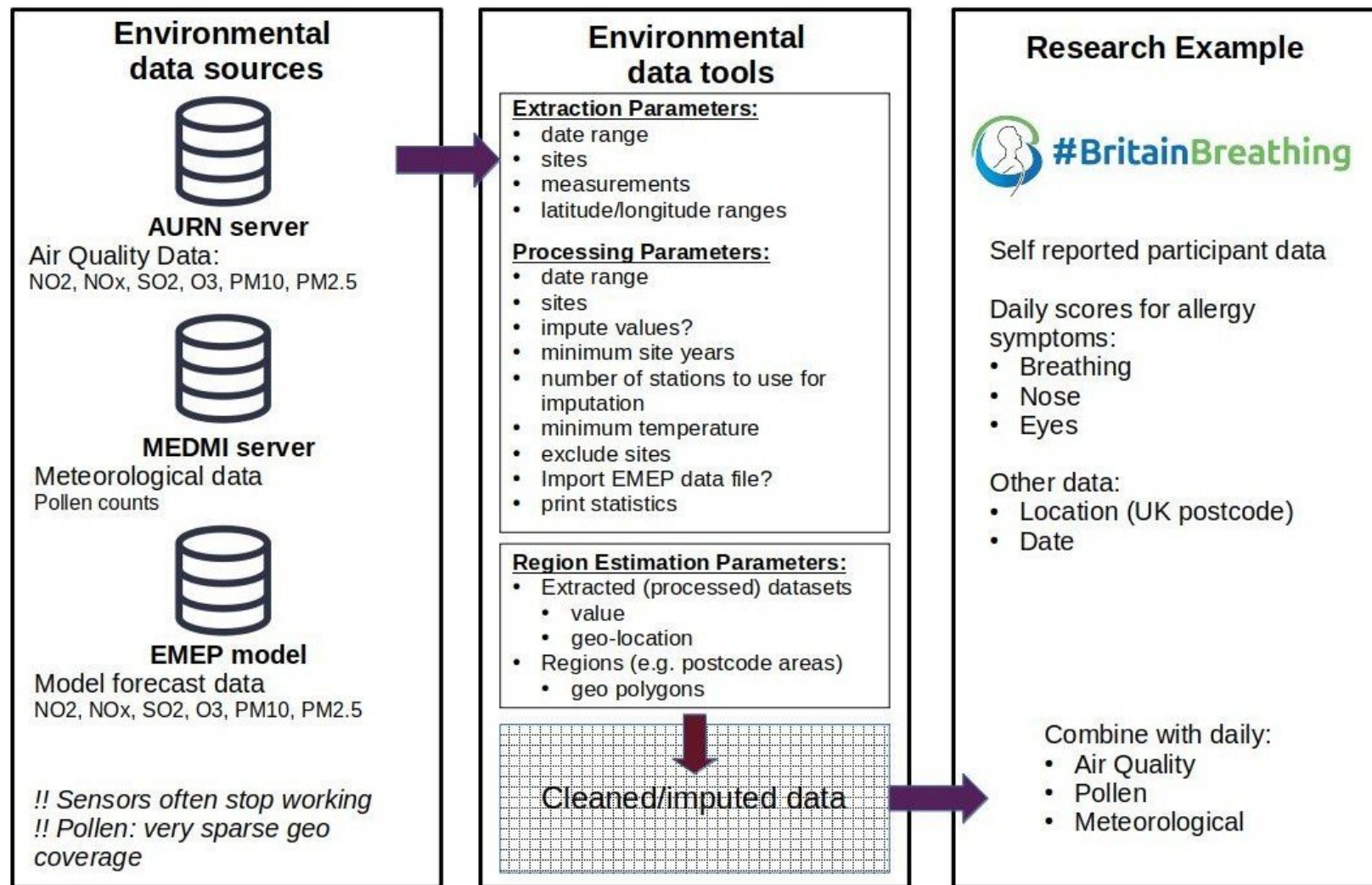
#BritainBreathing

HOME ABOUT US THE PROJECT NEWS IN ACTION FIND OUT MORE



#BritainBreathing

Filling the gaps



Visualisation tool

Timestamp: 2016-03-18

Select measurement *alnus*

Select estimation method *file*

Filter sites by...

Upload data and visualise

Upload geographical time-series data as a CSV file. The csv file must have the following columns:

Column 1: has header 'timestamp' and contains timestamps as strings, matching those available on the timestamp slider, Example: 2016-10-12

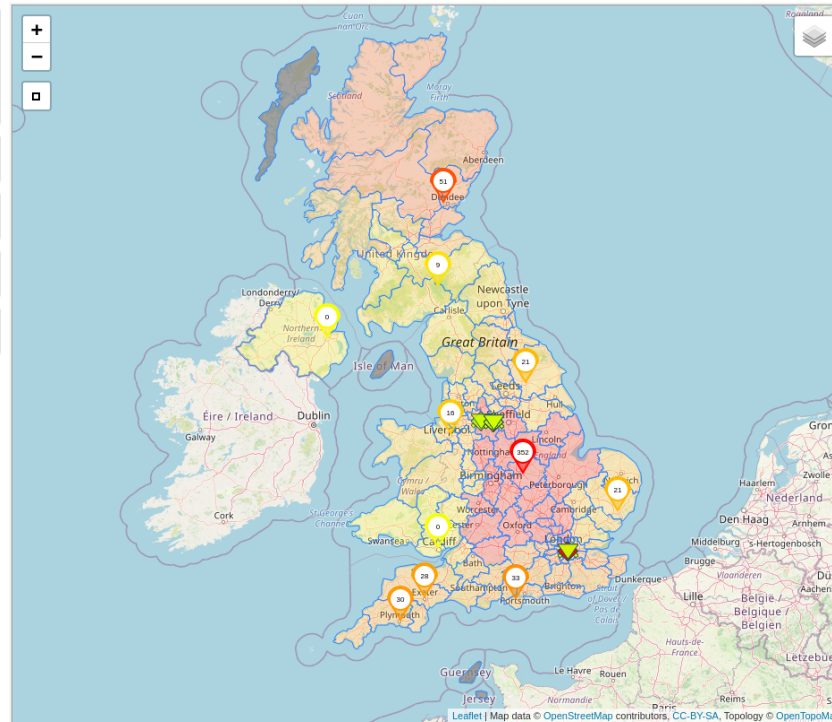
Column 2: has header 'geom' and contains geographical locations (currently points only), Example: POINT (-2.254 53.397)

Columns remaining: header(s) will be used as measurement name(s) and values should be numeric (non numeric will be converted to null)

Note that the uploaded file's data is stored locally on the browser, and only used for visualisation on the map, within this browser session.

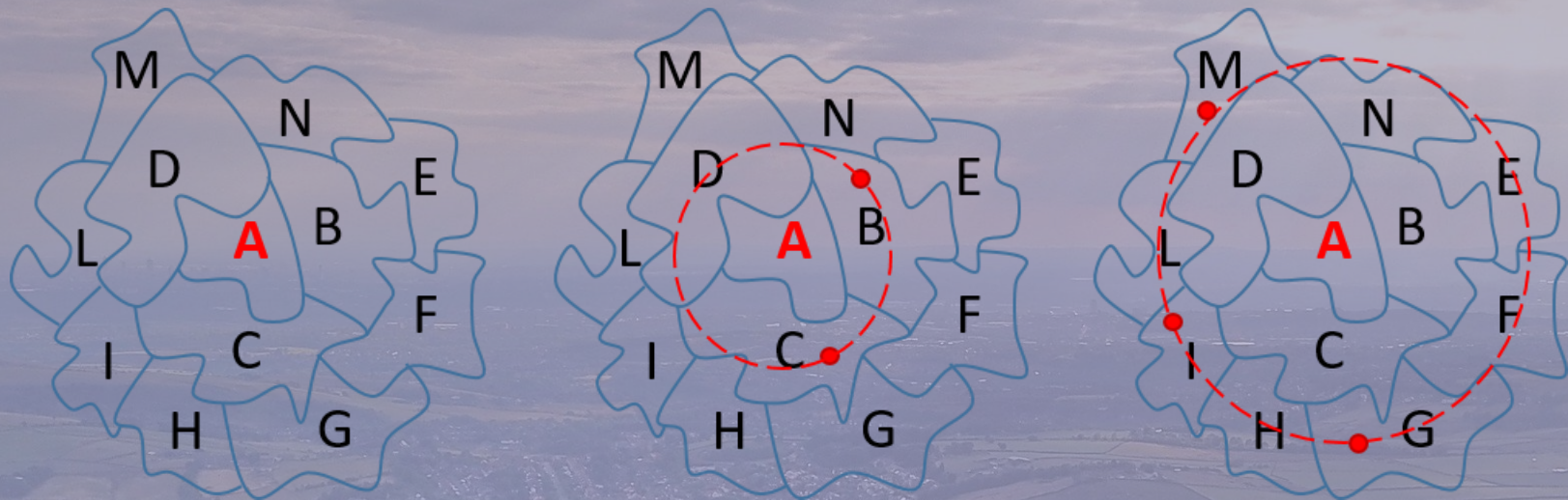
Choose file test_BB_use...ta_2016.csv

Download data



- use slider to find another timestamp
- use 'Select measurement' option to change measurements.

Regional estimations



Concentric Regions method illustrated on fictional postcode regions

- Regions where sensors exist: take mean
- Regions with no sensors: take mean of surrounding regions
 - Working outwards until sensors found

All code on Github (see links)

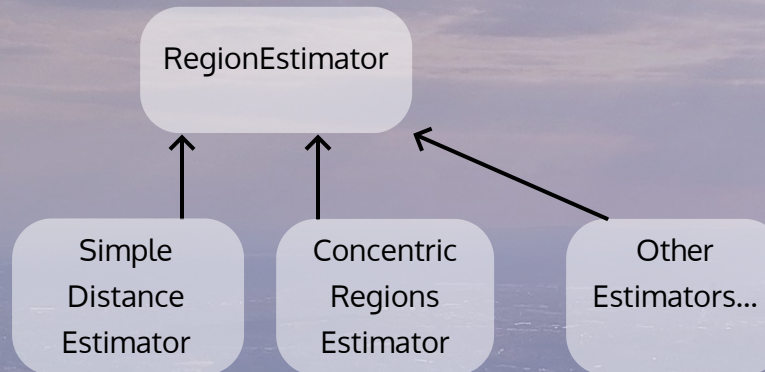
The screenshot displays the GitHub interface for the repository `UoMResearchIT / region_estimators`. The repository is currently on the `master` branch, which has 1 branch and 3 tags. The commit history shows a recent commit by `AnnAnnFryingPan` titled "Minor: remove comment and unused import" 4 days ago, with 132 total commits. The repository structure includes folders for `region_estimators`, `sample_input_files`, `scripts`, and `test`, along with files like `.gitignore`, `.travis.yml`, `LICENSE`, `README.md`, and `requirements.txt`.

The `README.md` file is visible, titled "region_estimators package". It indicates that the build is passing and describes the package as a Python library for calculating regional estimations of scalar quantities. It provides an example of estimating the NO2 (pollution) level of a postcode/zip region, based on site data nearby. The first version of the package is initialised with 2 estimation methods:

1. ConcentricRegions: look for actual data points in gradually wider rings, starting with sites within the region, and then working in rings outwards, until sites are found. If more than one site is found at the final stage, it takes the mean.

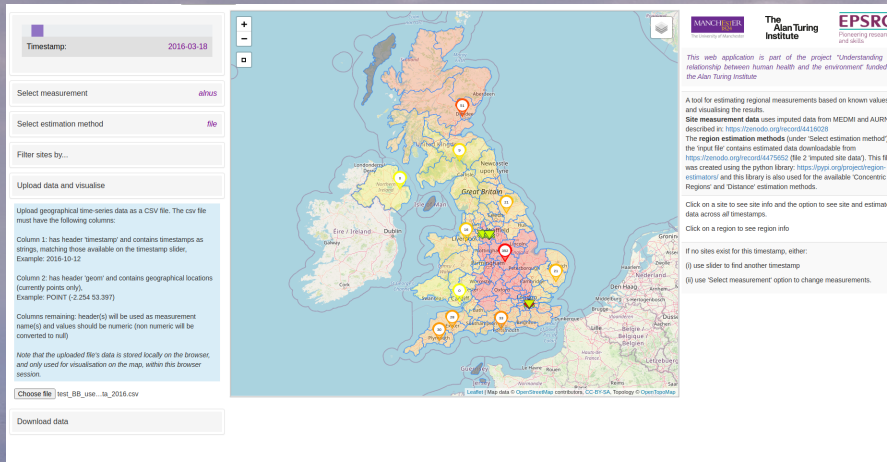
The right sidebar shows the repository's "About" section, describing it as a Python library for regional estimations, and lists the license as MIT. It also shows the "Releases" section with the latest release "Prevention abstract class ..." 4 days ago, and the "Packages" section, which currently has no published packages.

Regional estimations



- Current implementations are only baselines
- Open source MIT license
- Use as-is or extend as above
- Improve estimation techniques

Visualisation tool



- Quick overview of data
- visual patterns
 - estimation methods
 - filtering
- download data
- time-series
- Compare with own data

- Open source MIT license
- Usage:
 - Use web-app as-is (our data)
 - Contribute to and extend web-app code
 - Clone the repository and load in your data
 - Use Docker image to run/host on another machine

<http://minethegaps.manchester.ac.uk>

Timestamp: 2016-03-18

Select measurement

alrus

Select estimation method

file

Filter sites by...

Upload data and visualise

Upload geographical time-series data as a CSV file. The csv file must have the following columns:

Column 1: has header 'timestamp' and contains timestamps as strings, matching those available on the timestamp slider, Example: 2016-10-12

Column 2: has header 'geom' and contains geographical locations (currently points only), Example: POINT (-2.254 53.397)

Columns remaining: header(s) will be used as measurement name(s) and values should be numeric (non numeric will be converted to null)

Note that the uploaded file's data is stored locally on the browser, and only used for visualisation on the map, within this browser session.

Choose file

test_BB_use...ta_2016.csv

Download data

+

-

MANCHESTER

The University of Manchester

The Alan Turing Institute

EPSRC

Pioneering research and skills

This web application is part of the project "Understanding the relationship between human health and the environment" funded by the Alan Turing Institute

A tool for estimating regional measurements based on known values and visualising the results.

Site measurement data uses imputed data from MEDMI and AURN, described in: <https://zenodo.org/record/4416028>

The **region estimation methods** (under 'Select estimation method'): the 'input file' contains estimated data downloadable from <https://zenodo.org/record/4475652> (file 2 'imputed site data'). This file was created using the python library: <https://pypi.org/project/region-estimators/> and this library is also used for the available 'Concentric Regions' and 'Distance' estimation methods.

Click on a site to see site info and the option to see site and estimated data across *all* timestamps.

Click on a region to see region info

If no sites exist for this timestamp, either:

- use slider to find another timestamp
- use 'Select measurement' option to change measurements.

Links

- 2016-2019 environment datasets:
 - measurements (original and imputed)
 - <https://zenodo.org/record/4416028>
 - includes link to extraction and imputation tool set
 - regional estimations (from original and imputed)
 - <https://zenodo.org/record/4475652>
 - includes link to region_estimators tool
 - Scientific Data paper:
 - <https://www.nature.com/articles/s41597-022-01135-6>
- Visualisation Tool:
 - <http://minethegaps.manchester.ac.uk/>
 - <https://github.com/UoMResearchIT/mine-the-gaps>