

Mobility and transmission of COVID-19 in Wuhan

This repository includes codes for reproducing the results in the manuscript:

High-resolution data on human behavior for effective COVID-19 policy-making

In the following sections, we will describe the usage of each script file for reproducing the results in the paper.

It should be kept in mind that the scripts and data in this repository are meant for reproduction of the results in the paper only.

Scripts were tested on **Python 3.6.0**.

Prerequisite Python packages

The following Python packages are needed for running scripts (tested versions of packages are listed in parenthesis):

- numpy (1.18.1)
- pandas (0.25.3)
- matplotlib (2.1.2)
- seaborn (0.8.1)
- pymc (2.3.7)

Descriptions for the scripts

All scripts are placed in the **Folder code**. The following guides will help you to reproduce the results in the paper.

1. Clone this project into your local directory or download the ZIP archive and unzip it.
2. Target the terminal into the **root directory** of this repository.
3. Estimate parameters through the MCMC methods (Optional). Run

```
python code/Params_sampler.py 0
python code/Params_sampler.py 1
python code/Params_sampler.py 2
python code/Params_sampler.py 3
python code/Params_sampler.py 4
```

in different terminals to estimate the parameters for different groups of subdistricts parallelly. Each command takes about **16-18 hours** to run. If everything goes well, the word **Finish!** will present on the terminal at the end. The estimates of parameter will be saved to the **Folder objs**.

Since it takes a long time to re-estimate the parameters, this repository contains the parameters corresponding to the results in the paper, and the step is optional.

4. run `python code/SEPIR_main_analysis.py` to reproduce the main analysis in **Fig.2**. The script will take about 30 minutes to run. If everything goes well, the word **Finish!** will present on the terminal at the end. Results will be generated in the **Folder result**, including:

- `figure_2a.pdf`, `figure_2b.pdf`, `extended_fig_5.pdf` are the same figures in the paper.
- `figure_2ab_data.xlsx` is the data plotted in Fig.2a and Fig.2b, which contains the number of cumulative exposed cases by different ways of transmission, including the 95% confidence interval.

5. run `python code/Scenarios_simulation_NPI.py` to reproduce the scenario simulation in **Fig.2c** (Non-pharmaceutical interventions). The script will take about 3 hours to run. If everything goes well, the word **Finish!** will present on the terminal at the end. Results will be generated in the **Folder result**, including:

- `figure_2c.pdf` is the same figure in the paper, which contains the contour plot between effective reproduce number and the two types of indicators for interventions implemented in Wuhan.
- `extended_fig_8.pdf` are the same figures in the paper, describing the simulation results for Scenario 1 and 2.
- `npi_simulation_sce1.xlsx`, `npi_simulation_sce2.xlsx`, `npi_simulation_reality.xlsx` contain the number of cumulative exposed cases by different ways of transmission under Scenario 1, Scenario 2 and reality of intervention, including the 95% confidence interval.

Note: Due to uncertainty of sampling, there may be slightly differences among the results of each run.

Descriptions for other folders

Folder data contains the main data used in this study.

Folder objs contains necessary parameters in this study.

Folder result contains outputs of the scripts.