Exploring the influence of UK Greenspaces on the COVID-19 pandemic

**Overall Summary:** We focused on urban green spaces as an ecosystem service, addressing the question of *whether urban green spaces can downregulate the impact of COVID19*. These spaces provide the main source of interaction with semi-natural ecosystems that most people will have had during the COVID19 lockdown, since travel has been restricted and the majority of people live in urban areas. We used google mobility data, OS greenspace data, COVID19 death rate data. We processed this data into consistent regional areas (so that, for example, greenspace information corresponded to the same areas for which mobility data was reported). We performed correlation and regression analyses to identify the drivers of health outcome, as represented by the COVID19 death rate. Our conclusion was that while the per capita death rate is higher in cities, this effect was somewhat reduced by the use of parks (according to mobility data). On average, *a 50% increase in use of urban greenspaces corresponded to a 15% reduction in the COVID19 death rate.*

**Data Collection (main.py, subroutines/prepdata.py)**

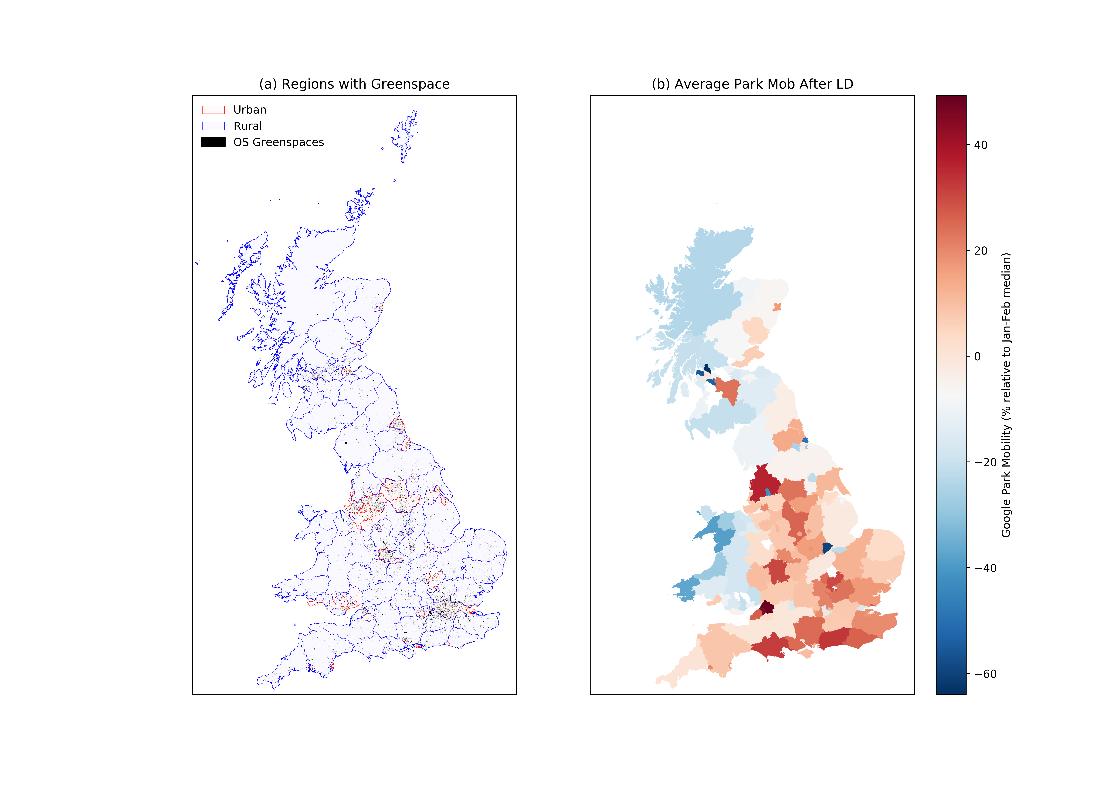
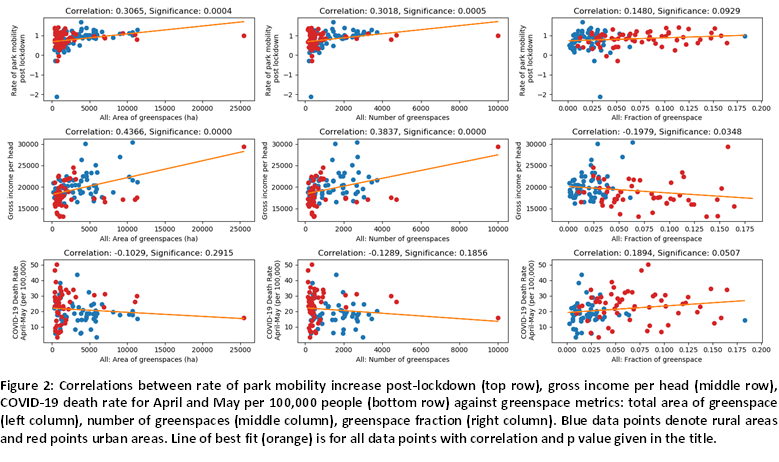
For this analysis we collected various datasets to encompass a range of influences that may affect the geographical variation of the susceptibility of COVID19. Firstly, we gathered mobility data from Google[[1]](#footnote-2). The mobility data is defined as percentage difference relative to the median of January-February. To get a single measure of mobility we used the average mobility and the gradient of the linear regression for a rate of change defined since the start of the UK lockdown (24/03). To define geographic shapes from the regions from the Google dataset we compared defined regions with shape files from the (International standardisation organisation) ISO and unitary authorities OS[[2]](#footnote-3). This allowed us to define mobility spatially (Figure 1). To define regions green space for this we used the OS greenspace data[[3]](#footnote-4). By comparing the intersecting regions, we can estimate the corresponding. In addition, we collected COVID19 death rates[[4]](#footnote-5) along with 2018 measurements of the gross income per capita for unitary authorities[[5]](#footnote-6).

Figure 1Panel (a) Shows the Urban (red) and Rural (blue) google mobility regions in Great Britain along with OS Greenspaces (black). Panel (b) shows the average park mobility after lock down.

**Correlation analysis of park mobility, gross income per head, COVID19 death rate against various greenspace analysis (main.py, subroutines/plotting.py, subroutines/plotting\_secondary.py)**

We started by performing a linear regression analysis on the rate of increase of park mobility post-lockdown, gross income per head and the COVID19 death rate per 100,000 people for April and May against multiple greenspace metrics.

The greenspace metrics used were total area of greenspace for the region, the number of greenspaces in the region and the fractional area of the region that is assigned as a greenspace. These greenspaces can be further sub-divided into greenspaces for sport such as golf courses, tennis courts and greenspaces for non-sporting activities such as parks. The results of these analyses are summarised in Figure 2 below, and further plots such as sport-only greenspace metrics can be obtained using the Python script provided.

We find that the **rate of park mobility post lockdown** has a strongly significant (p values < 0.001) positive correlation to both the total area of greenspaces and number of greenspaces. Rural areas reflect these results even more strongly (p values < 0.00001). Urban areas on the other hand, have a significant (p value = 0.0504) positive correlation to greenspace fraction.

**Gross income per head** has a strongly significant (p values < 0.01) positive correlation to total area of greenspaces and number of greenspaces. However, a significant (p value = 0.0348) negative correlation to greenspace fraction. Both urban and rural areas individually show strongly significant (p values < 0.01) positive correlations to total area of greenspaces and number of greenspaces. However, rural areas have a strongly significant (p value = 0.0062) positive correlation to greenspace fraction, whereas, urban areas have no significant (p value = 0.5304) correlation.

The **COVID19 death rate** for April and May has a significant (p value = 0.0507) positive correlation to fraction of greenspace. (This is more strongly correlated for fraction of sport specific greenspaces). However, this seems to be an artefact of urban and rural areas counteracting each other as when separated there are no significant (p values > 0.25) correlations in the COVID-19 death rate to greenspace.

**Regression analysis of factors controlling COVID death rate (multilinear\_analysis\_code.R)**

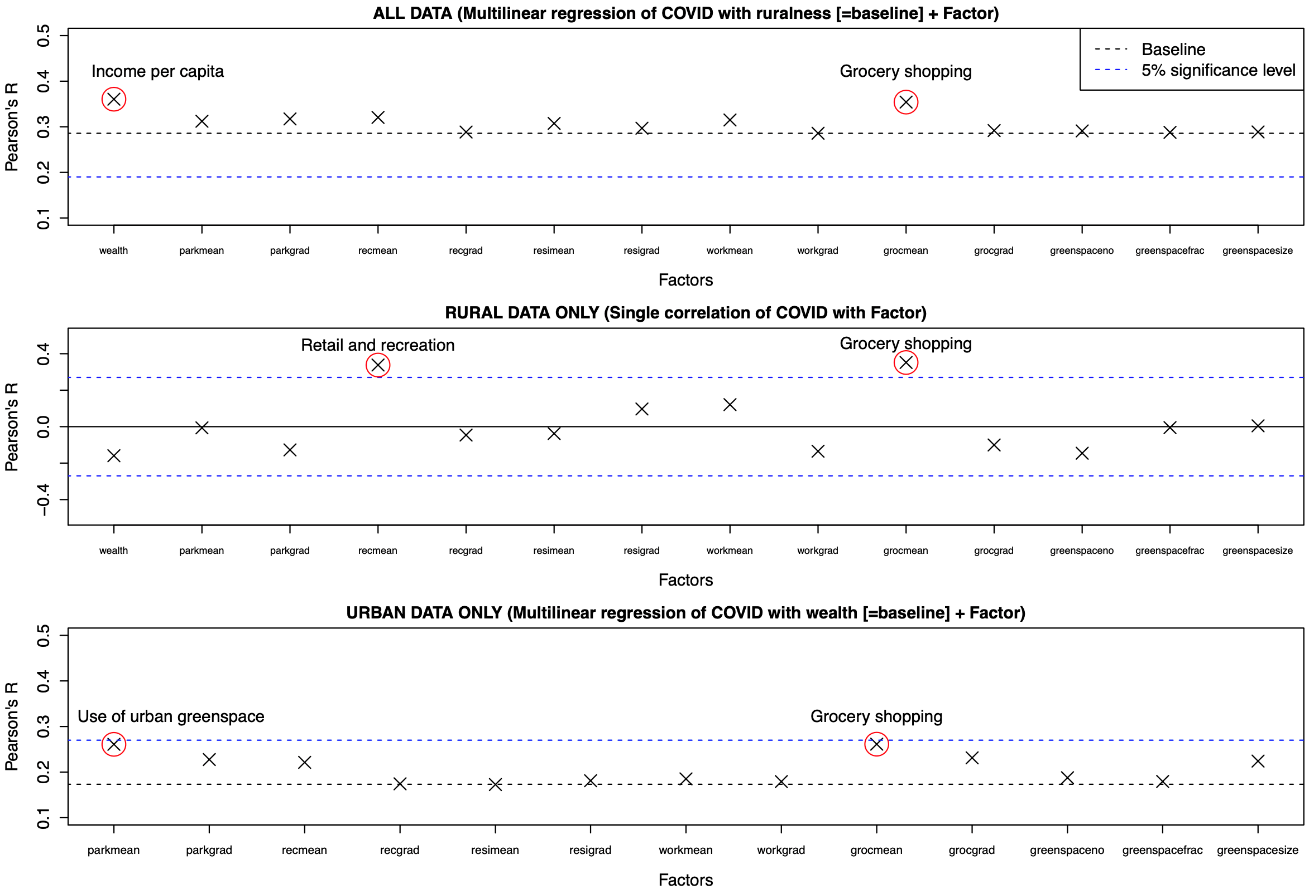
We performed multi-linear regression analysis of the COVID death rate data with the processed datasets described in the previous section. This allowed us to identify the relative importance of the different factors affecting COVID death rate. The most significant variable controlling COVID death rate was rural vs urban status. We therefore:

1) Performed a multilinear regression using “ruralness” (+1 for rural and 0 for urban) and the other variables. Using these regressions we calculated the Pearson correlation against COVID death rate to determine whether the additional variables added predictive power in addition to ‘ruralness’. The correlation coefficients are shown in Figure 3, upper plot. The two variables that added the most predictive power were wealth and grocery shopping. Regions with higher income showed lower death rates, and regions with higher rates of grocery shopping showed higher death rates.

2) We treated rural and urban data separately, shown in Figure 3 central and lower plots.

In rural areas, the amount that people went grocery shopping or travelled for “retail and recreation” were the largest factors, potentially because shops are the main locations in which large numbers of people come into contact in rural areas.

In urban areas, wealth was the most significant controlling factor, so we performed a multilinear regression using “income per capita” and the other variables. After wealth, Figure 3 (lower panel) shows that the two greatest factors controlling COVID death rate were the average “park mobility” during lockdown, and the amount of grocery shopping. Grocery shopping was an important controlling factor across the dataset, suggesting that supermarket use may be leading to COVID transmission. However, the park mobility data had an equally large – and opposite – impact on death rate in cities. This suggests that the more people are using parks/green spaces in cities, the lower their death rate. Note that we factored out the effect of wealth and considered all of the other mobility variables, none of which explain the COVID death rate to the same extent.



**Figure 3.** Results of multilinear regression analysis of variables against COVID19 death rate

From our regression analyses we conclude that rural vs urban status is the biggest factor in COVID death rate (death rate being higher in urban areas). Wealthand the amount of grocery shoppingpeople did during lockdown are the next two biggest factors (higher income corresponds to lower death rate; more grocery shopping corresponds to higher death rate). However, when we consider the urban data only, time spent in urban green spaces becomes the next most significant factor after those outlined above. According to our regression model,***a 50% increase in use of urban green spaces leads to a 15% reduction in COVID death rate***,compared to the average COVID death rate in urban areas. The other aspects of green spaces such as their total area, number, or average size, did not appear to be important.

**Conclusions:** We considered the urban green spaces as these provide the only (or main) interaction with semi-natural ecosystems that most people will have had during the COVID19 lockdown. Having factored out the other main controlling variables we found that use of these green spaces corresponded to a reduced death rate. This points to the health benefits of these green spaces as an ecosystem service, although further analysis is required to identify cause and effect. For example, weather and climate data should be used in further analysis. Rural vs urban environments have a strong impact on the COVID19 death rate. Likely this is not directly due to the ecosystem services that rural areas provide, but more due to the reduced contact between people in rural areas, which is supported by the fact that grocery shopping leads to a strongly increased death rate in rural areas.

1. Google LLC *"Google COVID-19 Community Mobility Reports"*.  
   https://www.google.com/covid19/mobility/ Accessed: 03/07/2020 [↑](#footnote-ref-2)
2. <https://gadm.org/data.html> Accessed: 03/07/2020, <https://www.ordnancesurvey.co.uk/business-government/products/boundaryline> Accessed: 03/07/2020 [↑](#footnote-ref-3)
3. <https://www.ordnancesurvey.co.uk/business-government/products/open-map-greenspace> Accessed: 03/07/2020 [↑](#footnote-ref-4)
4. <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/deathsinvolvingcovid19bylocalareasanddeprivation/deathsoccurringbetween1marchand31may2020> Accessed: 03/07/2020 [↑](#footnote-ref-5)
5. <https://www.ons.gov.uk/economy/regionalaccounts/grossdisposablehouseholdincome/bulletins/regionalgrossdisposablehouseholdincomegdhi/1997to2018> [↑](#footnote-ref-6)