Supplemental material v1

August 10, 2020

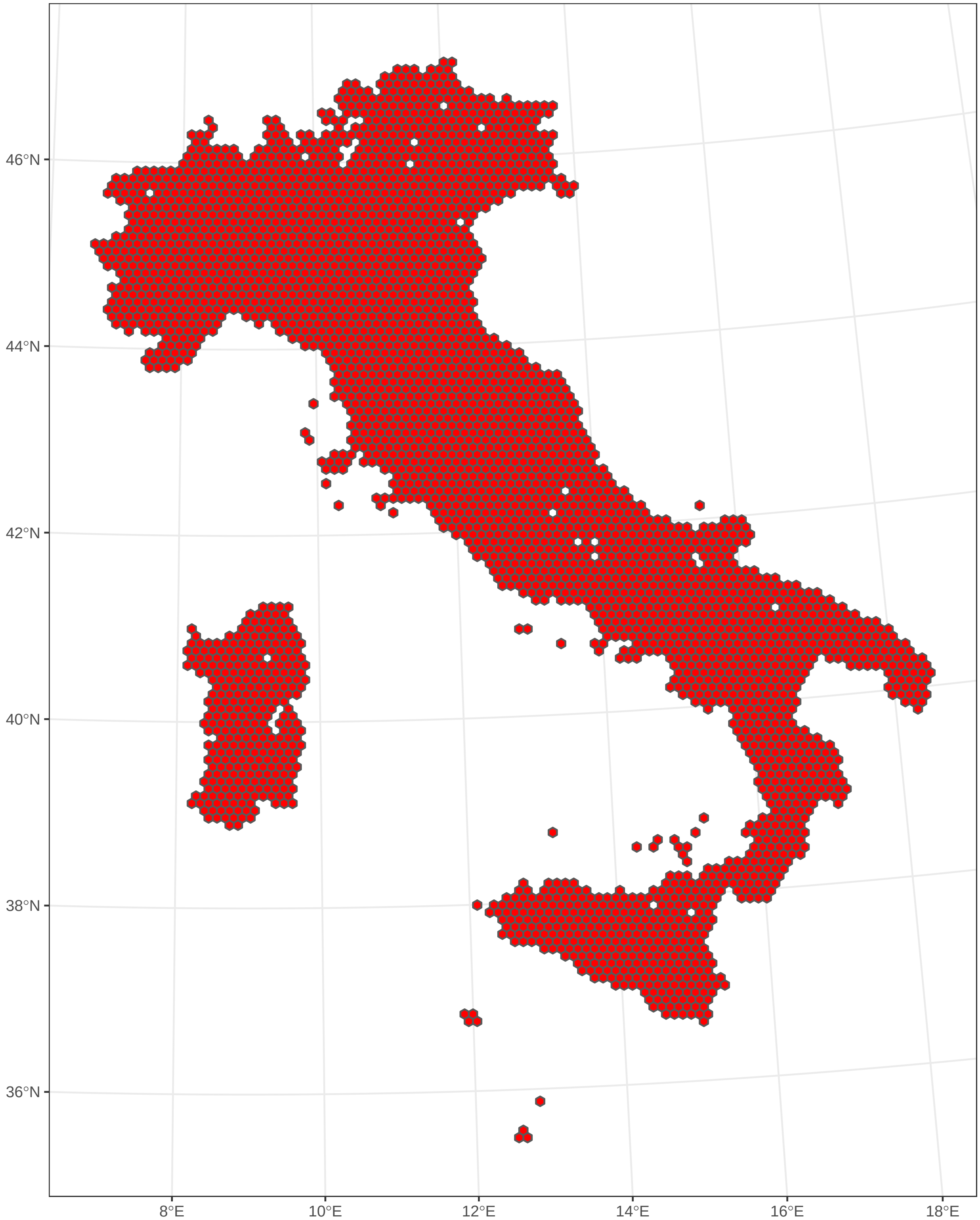
Contents

Supplemental material to the paper *The security implications of COVID-19 in times of recurrent respiratory viruses: environmental, social, and medical risk factors*

1 Hexagonal Grid

An hexagonal grid is randomly created to cover the entire Italian territory. Each grid is set to have a size (diameter) of 10-km. Variables associated to each cell are used as the unit of analysis for the regression models. The adoption of a grid of hexagonal cell is justified by the necessity of mapping statistics reported for geographies with significantly different shape and area.

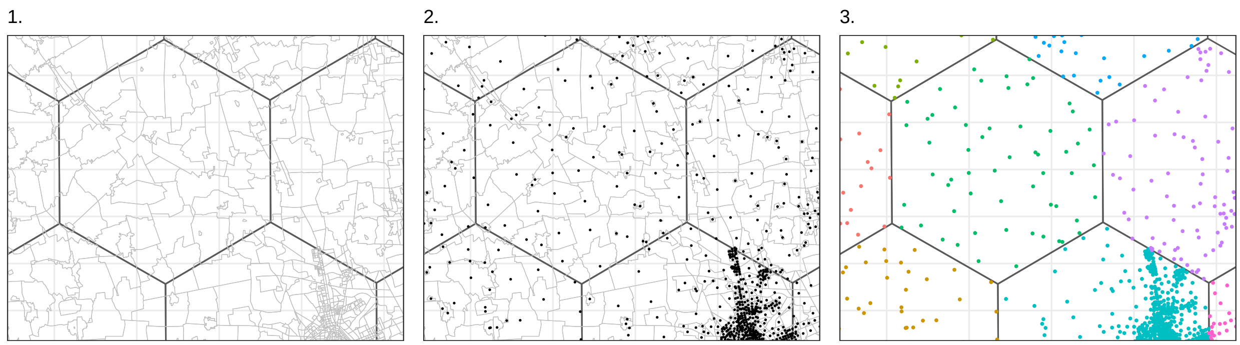
The following figure shows the grid of 3801 hexagonal cell used in the analysis. Notably, 271 cells have been dropped since based on the 2011 census are not inhabitated.


46°N
44°N
42°N
40°N
38°N
36°N
8°E
10°E
12°E
14°E
16°E
18°E

1.1 Interesection between census sections and hexagonal cells for count variables

To map census count variables measured at the level of census sections with hexagonal cells, we proceeded as following:

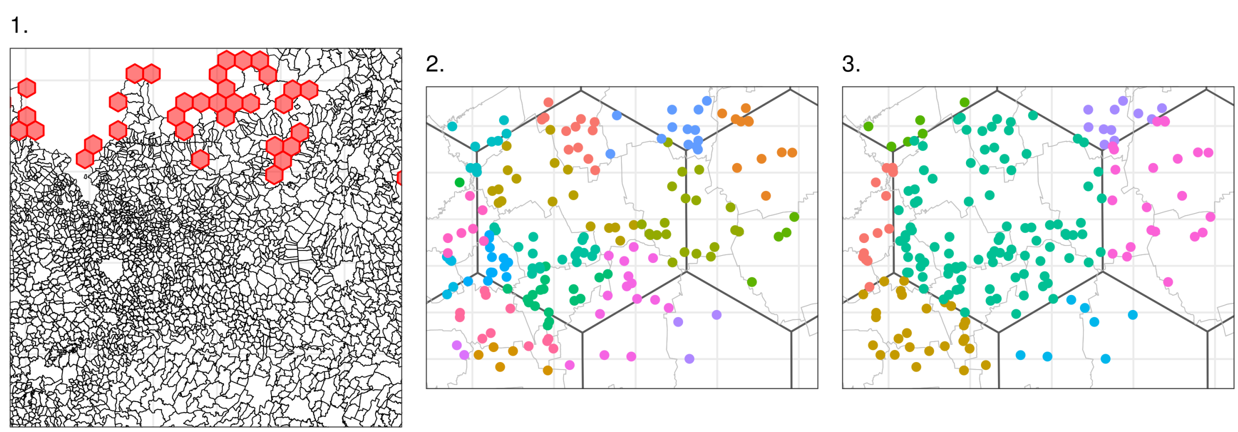
1. we computed the centroid, or its geographic centre, of the 402,678 census sections;
2. we intersected centroids and hexagonal cells;
3. for each haxagonal cell, we summed the count variables pertaining to the census sections with a centroid contained within the cell.



1.2 Interesection between comune and hexagonal cells for count variables

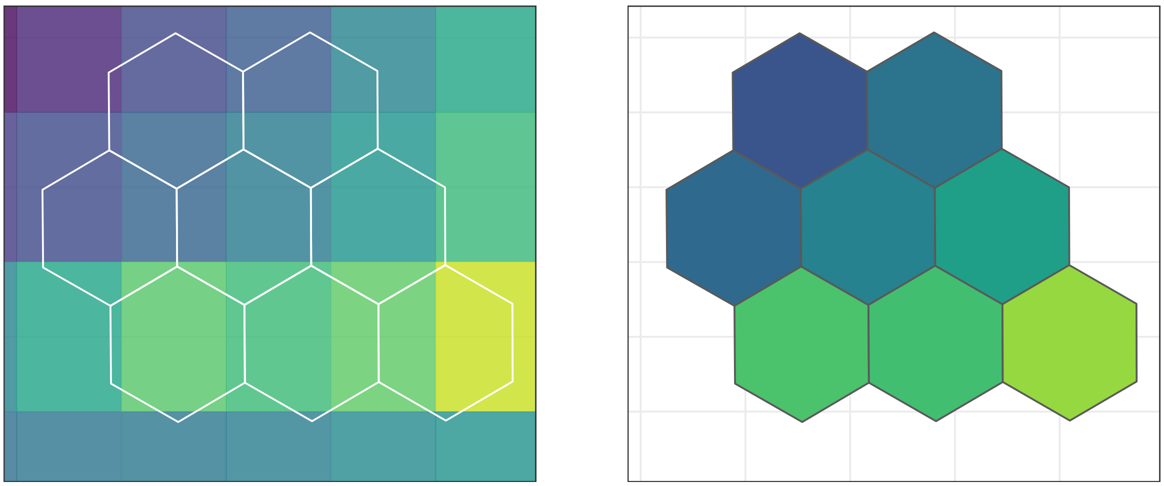
To map *comune* count variables onto hexagon cells, we proceeded as following:

1. we clipped unpopulated areas corresponding to the unpopulated hexagon cells identified before from the *comune*;
2. we randomly sampled a number of points within the remaining (populated) area of the *comune* corresponding to the reported count;
3. we counted how many points resulted within each hexagonal cell.



1.3 Intersection between areas and hexagon cells for continuous variables

When continuous variables were reported for a grid of cells, as with air quality and atmospheric data, or when count variables were reported at the level of comune for multiple income brackets, as in the case of income data, the variable for each hexagonal cell was computed as the average of the values reported for the intersecting cells (or comune) weighted according the extension of the area of intersection.



1.4 Intersection between variables reported at the regional level and hexagon cells

For variables reported at the regional level, each hexagon cell assumes the value reported for the region where it is located. In the case of a hexagonal cell intersecting more than one region, the cell is assigned to the region with the largest intersection area.

2 Variables

2.1 Excess mortality

Excess mortality is based on Istat figures published in July 2020 (link).

Figures for deaths recorded in every calendar day for the period 2015-19 are reported for 7904 *comune*. Figures for deaths recorded for each calendar day for the first five months of 2020 are reported for 7357 *comune* (or 93.08% of the total). In 2019, these 7357 *comune* reported 94.89% of the number of deaths reported in Italy. The following table shows the proportion of the total number of deaths reported by the these 7357 *comune* by Region.

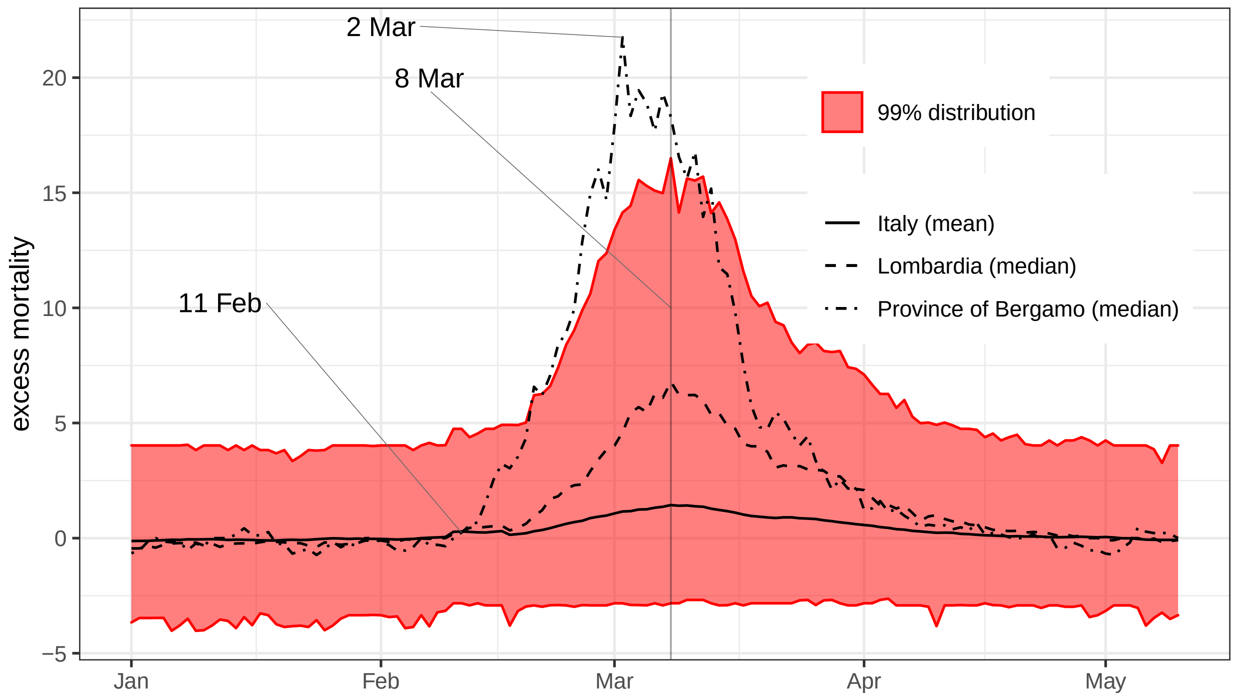
We compute excess mortality is computed through the following steps:

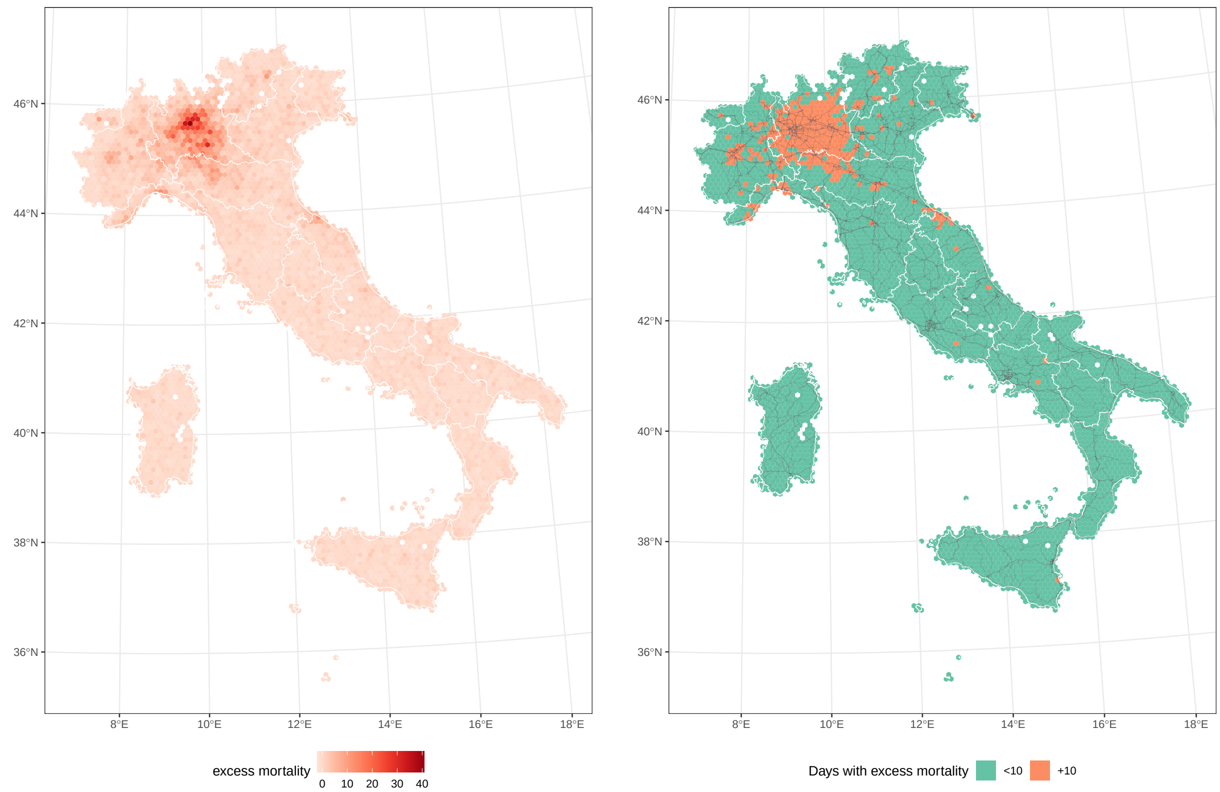
1. The daily number of deaths reported for each *comune* between 2015 and 2020 is mapped onto our hexagon cells with the methods detailed above for counted variables.
2. For each year between 2015 and 2020, we calculated for each hexagon cell a centered 7-day moving average of the number of reported deaths over 65 in each comune for each year from 2015 to 2020.
3. For each calendar day, the mean and the standard deviation of the moving averages from 2015 to 2019 are calculated.

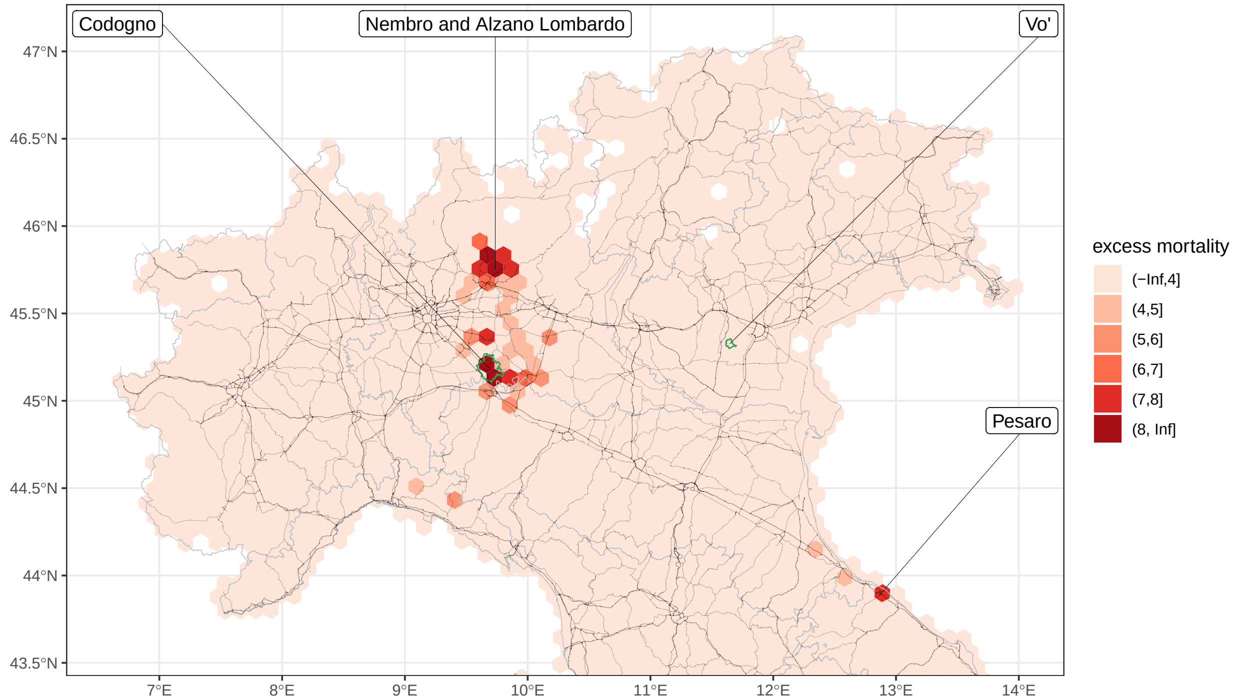
4. For each calendar day, the difference between the 2020 moving average and the 2015-2019 mean is divided by the 2015-2019 standard deviation. In doing so, we define the standard deviation as the unit of measurement and normalise differences across cells due to population density.
5. We shifted the resulting time series backward by 17 days to the estimated day of the infection resulting in death instead of the actual day of reported death.
6. Infinite and non-number values resulting from division by zero are assigned an excess mortality value of zero.

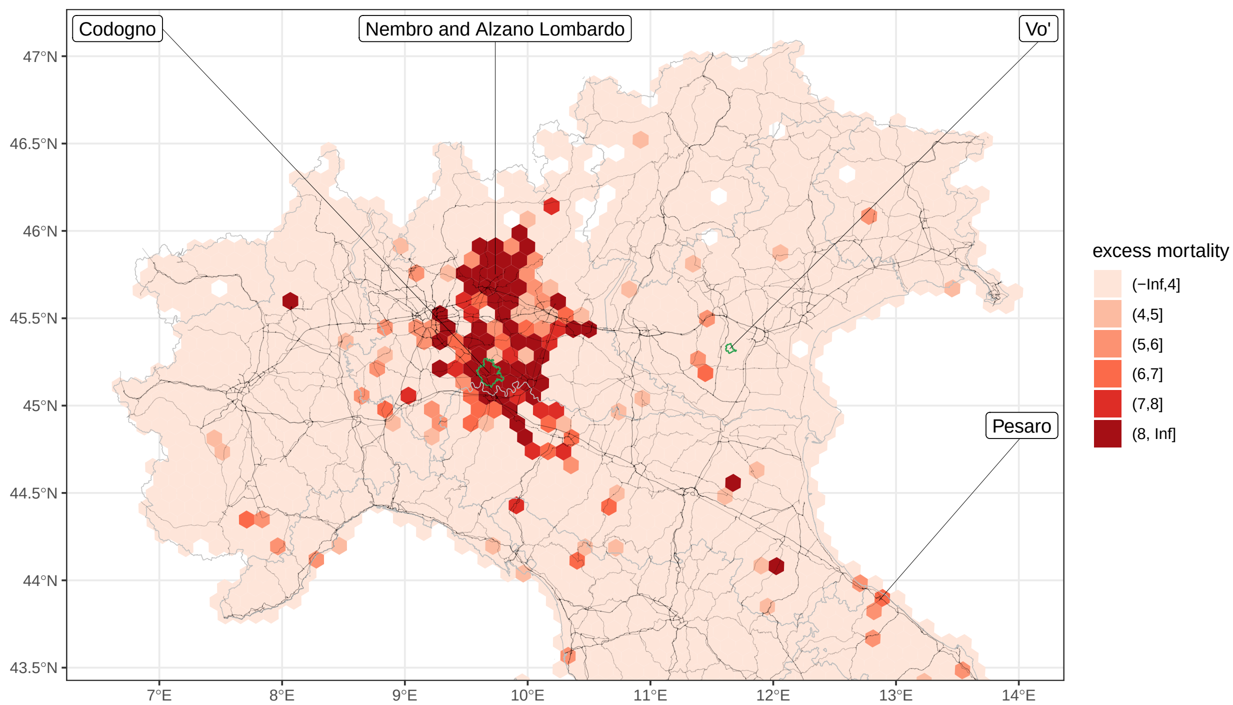
With this method we are able to estimate the excess mortality for 4072 hexagon cells and for each day between 01 January 2020 and 14 May 2020.

The distribution of the resulting excess mortality variable is detailed over time and over space in the following figures.



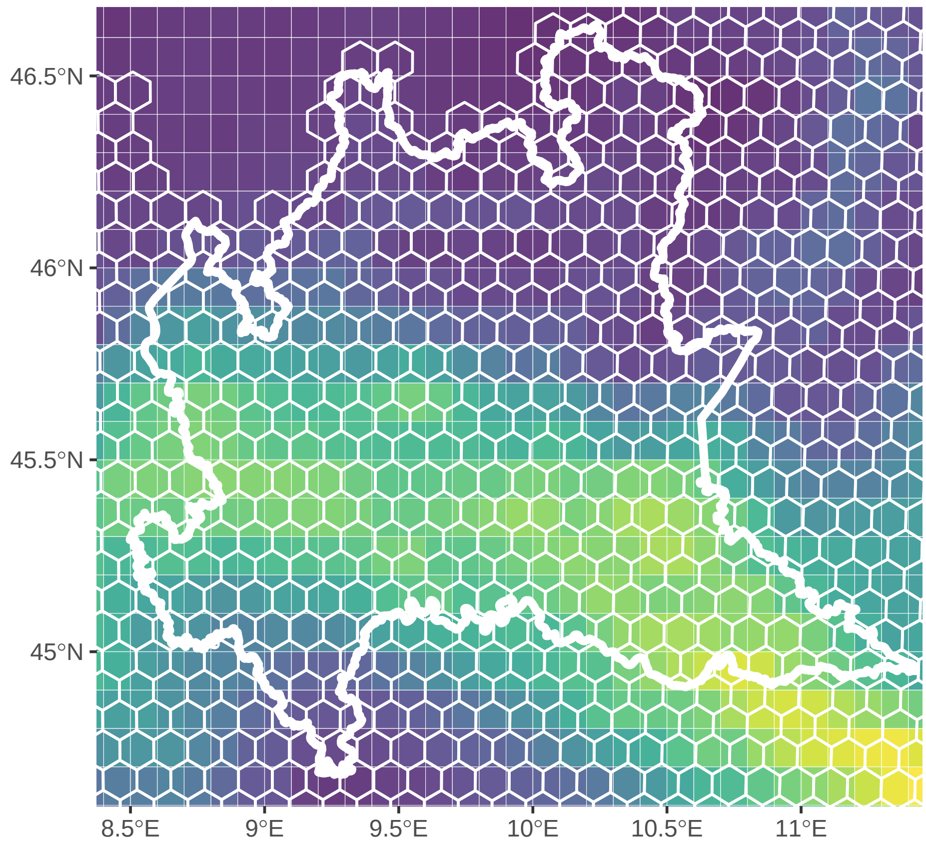

46°N
44°N
42°N
40°N
38°N
36°N
8°E
10°E
12°E
14°E
16°E
18°E
excess mortality
0
10
20
30
40
Days with excess mortality
<10
+10



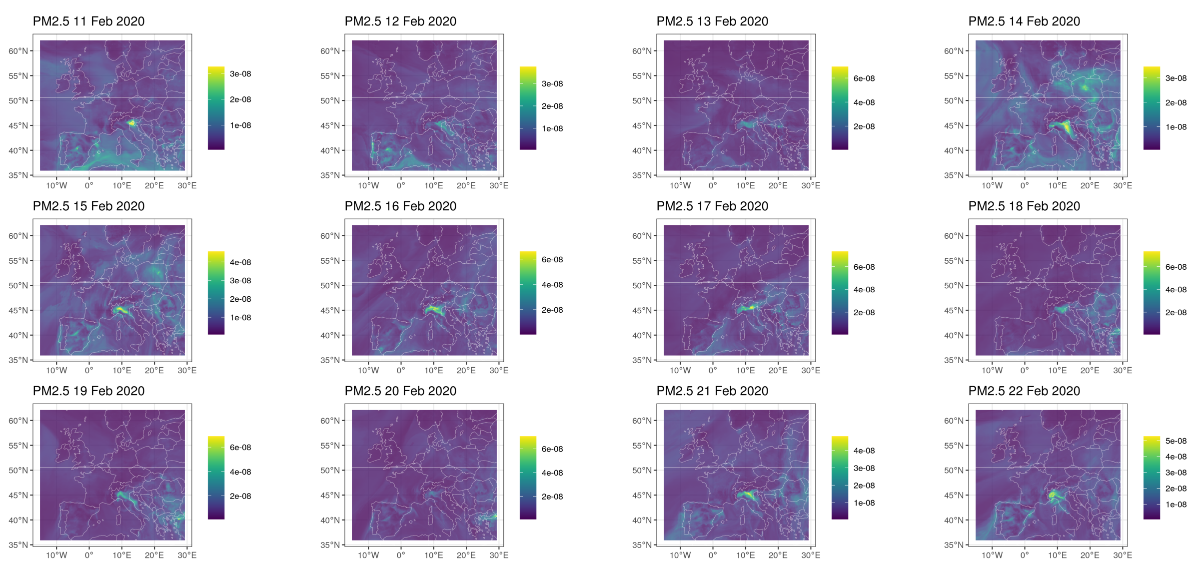


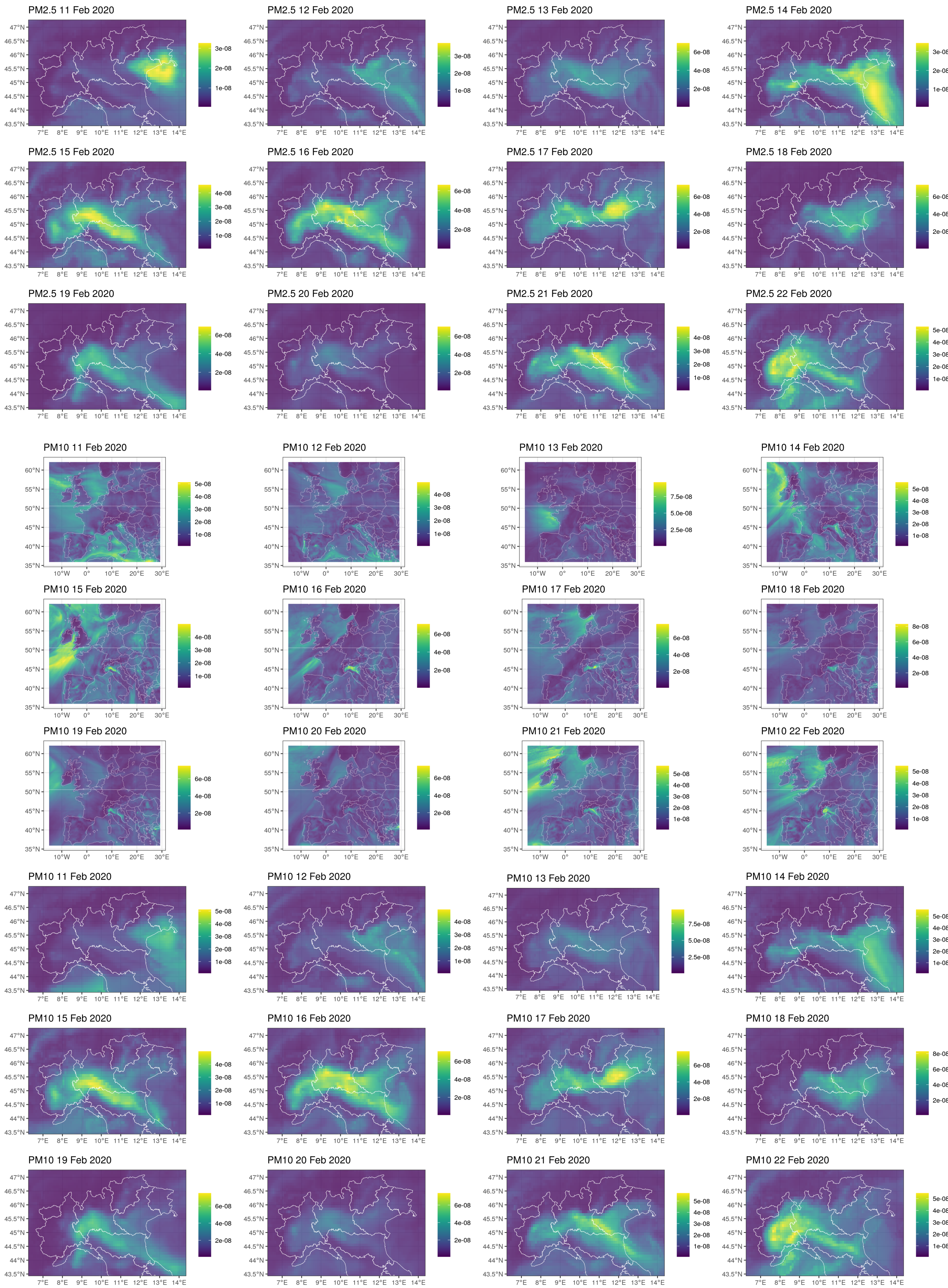
2.2 Air quality variables

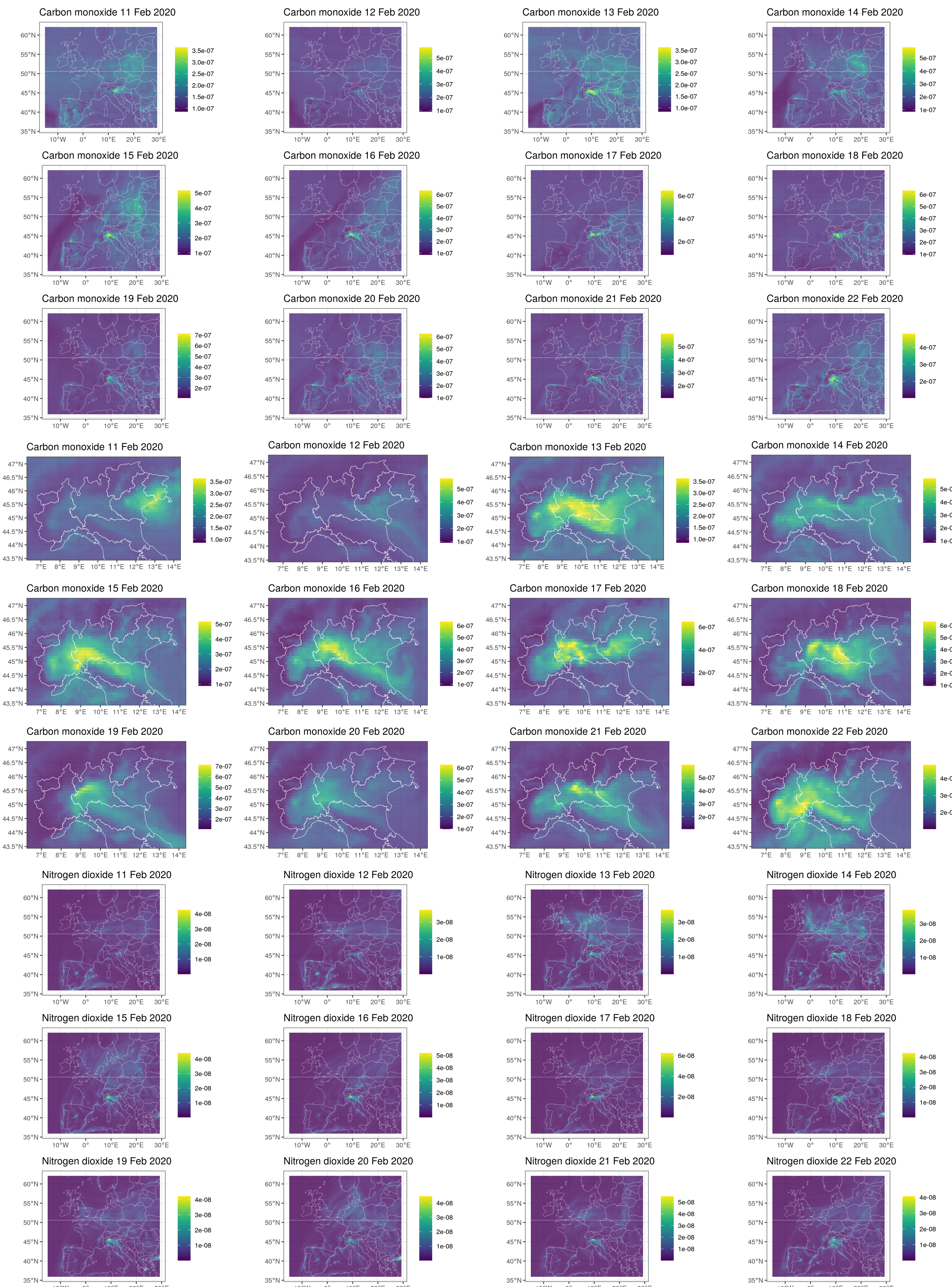
Air quality data are obtained from Copernicus Atmosphere Monitoring Service (CAMS, link) with a 10 kilometres' resolution. In the following figures, CAMS cells are overlayed to our hexagon cells.

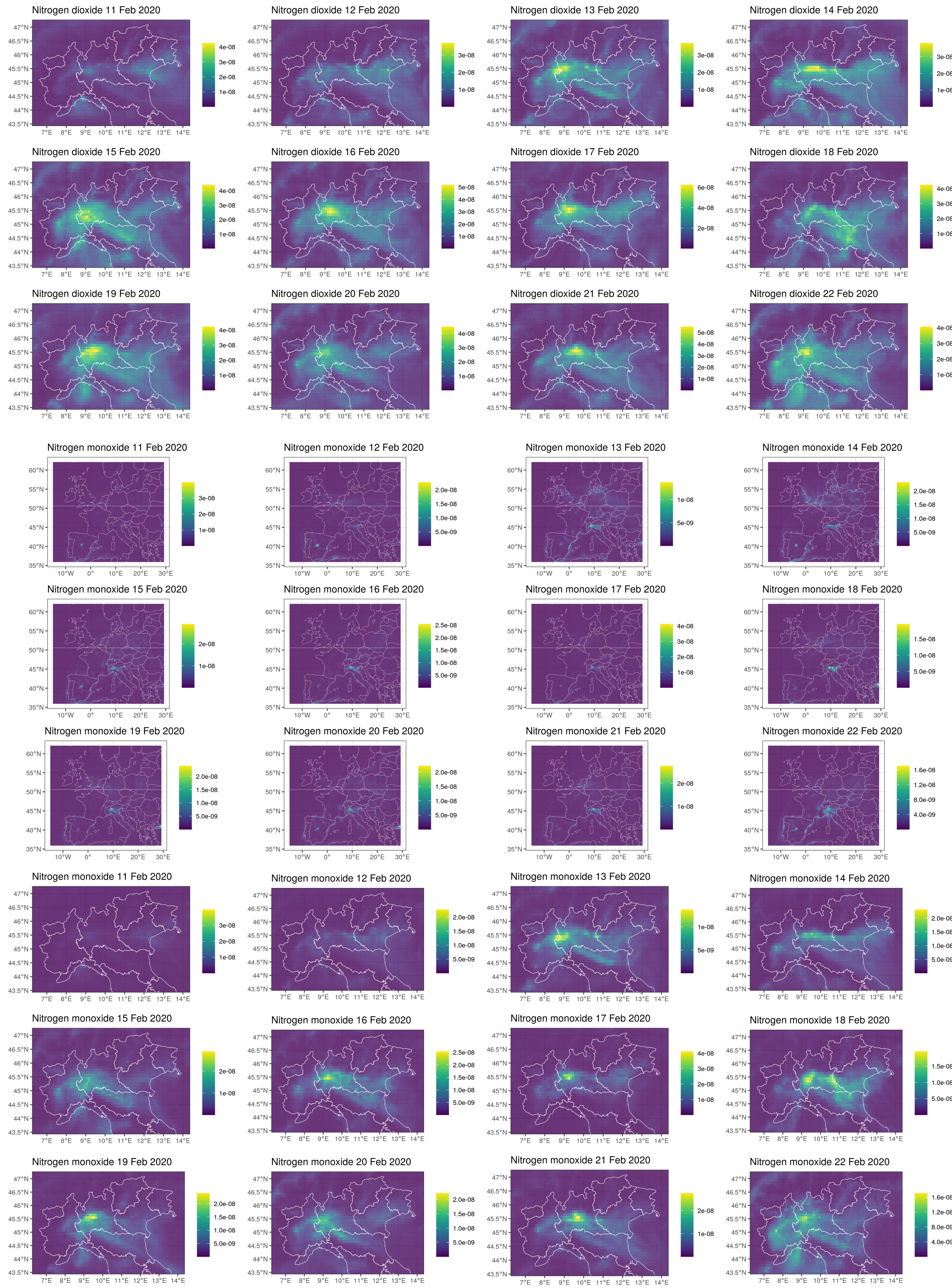


In the following figures we map instead the evolution of selected air quality measures over Europe and Northern Italy.



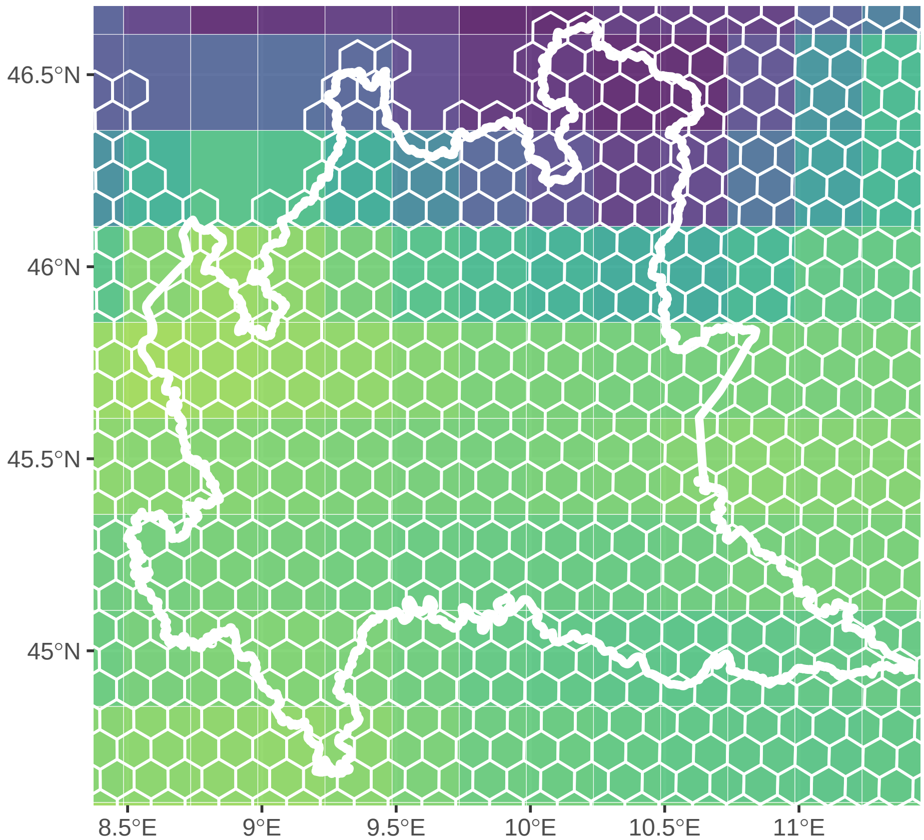

PM2.5 11 Feb 2020
PM2.5 12 Feb 2020
PM2.5 13 Feb 2020
PM2.5 14 Feb 2020
PM2.5 15 Feb 2020
PM2.5 16 Feb 2020
PM2.5 17 Feb 2020
PM2.5 18 Feb 2020
PM2.5 19 Feb 2020
PM2.5 20 Feb 2020
PM2.5 21 Feb 2020
PM2.5 22 Feb 2020
PM10 11 Feb 2020
PM10 12 Feb 2020
PM10 13 Feb 2020
PM10 14 Feb 2020
PM10 15 Feb 2020
PM10 16 Feb 2020
PM10 17 Feb 2020
PM10 18 Feb 2020
PM10 19 Feb 2020
PM10 20 Feb 2020
PM10 21 Feb 2020
PM10 22 Feb 2020
PM10 11 Feb 2020
PM10 12 Feb 2020
PM10 13 Feb 2020
PM10 14 Feb 2020
PM10 15 Feb 2020
PM10 16 Feb 2020
PM10 17 Feb 2020
PM10 18 Feb 2020
PM10 19 Feb 2020
PM10 20 Feb 2020
PM10 21 Feb 2020
PM10 22 Feb 2020


Carbon monoxide 11 Feb 2020
Carbon monoxide 12 Feb 2020
Carbon monoxide 13 Feb 2020
Carbon monoxide 14 Feb 2020
Carbon monoxide 15 Feb 2020
Carbon monoxide 16 Feb 2020
Carbon monoxide 17 Feb 2020
Carbon monoxide 18 Feb 2020
Carbon monoxide 19 Feb 2020
Carbon monoxide 20 Feb 2020
Carbon monoxide 21 Feb 2020
Carbon monoxide 22 Feb 2020
Carbon monoxide 11 Feb 2020
Carbon monoxide 12 Feb 2020
Carbon monoxide 13 Feb 2020
Carbon monoxide 14 Feb 2020
Carbon monoxide 15 Feb 2020
Carbon monoxide 16 Feb 2020
Carbon monoxide 17 Feb 2020
Carbon monoxide 18 Feb 2020
Carbon monoxide 19 Feb 2020
Carbon monoxide 20 Feb 2020
Carbon monoxide 21 Feb 2020
Carbon monoxide 22 Feb 2020
Nitrogen dioxide 11 Feb 2020
Nitrogen dioxide 12 Feb 2020
Nitrogen dioxide 13 Feb 2020
Nitrogen dioxide 14 Feb 2020
Nitrogen dioxide 15 Feb 2020
Nitrogen dioxide 16 Feb 2020
Nitrogen dioxide 17 Feb 2020
Nitrogen dioxide 18 Feb 2020
Nitrogen dioxide 19 Feb 2020
Nitrogen dioxide 20 Feb 2020
Nitrogen dioxide 21 Feb 2020
Nitrogen dioxide 22 Feb 2020


Nitrogen dioxide 11 Feb 2020
Nitrogen dioxide 12 Feb 2020
Nitrogen dioxide 13 Feb 2020
Nitrogen dioxide 14 Feb 2020
Nitrogen dioxide 15 Feb 2020
Nitrogen dioxide 16 Feb 2020
Nitrogen dioxide 17 Feb 2020
Nitrogen dioxide 18 Feb 2020
Nitrogen dioxide 19 Feb 2020
Nitrogen dioxide 20 Feb 2020
Nitrogen dioxide 21 Feb 2020
Nitrogen dioxide 22 Feb 2020
Nitrogen monoxide 11 Feb 2020
Nitrogen monoxide 12 Feb 2020
Nitrogen monoxide 13 Feb 2020
Nitrogen monoxide 14 Feb 2020
Nitrogen monoxide 15 Feb 2020
Nitrogen monoxide 16 Feb 2020
Nitrogen monoxide 17 Feb 2020
Nitrogen monoxide 18 Feb 2020
Nitrogen monoxide 19 Feb 2020
Nitrogen monoxide 20 Feb 2020
Nitrogen monoxide 21 Feb 2020
Nitrogen monoxide 22 Feb 2020
Nitrogen monoxide 11 Feb 2020
Nitrogen monoxide 12 Feb 2020
Nitrogen monoxide 13 Feb 2020
Nitrogen monoxide 14 Feb 2020
Nitrogen monoxide 15 Feb 2020
Nitrogen monoxide 16 Feb 2020
Nitrogen monoxide 17 Feb 2020
Nitrogen monoxide 18 Feb 2020
Nitrogen monoxide 19 Feb 2020
Nitrogen monoxide 20 Feb 2020
Nitrogen monoxide 21 Feb 2020
Nitrogen monoxide 22 Feb 2020

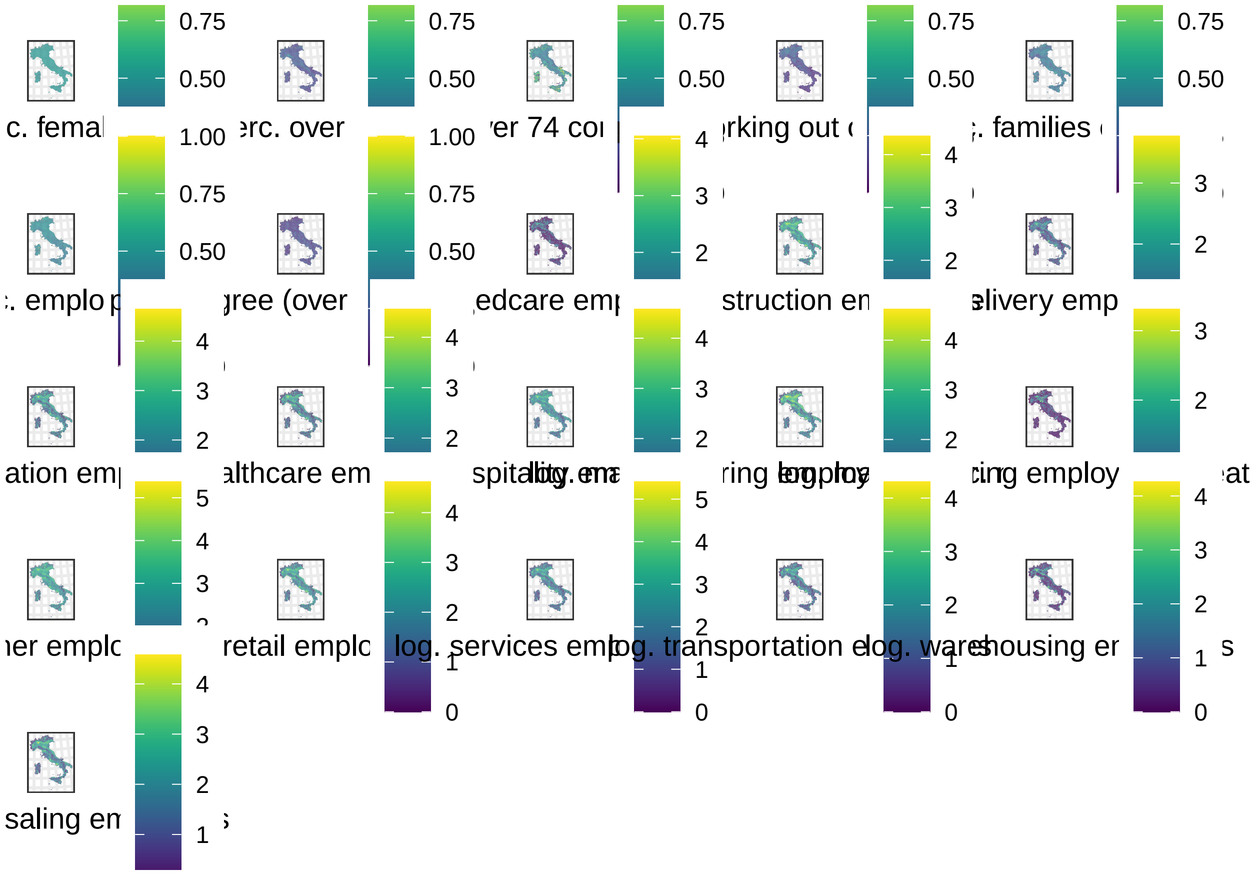
2.3 Atmosphere variables

Atmosphere data are derived from ERA5 (link) variables with a 30 kilometres' resolution. In the following figures, ERA5 cells are overlayed to our hexagon cells.



2.4 Census variables

Demographic variables are computed based on census counts at the level of 2011 census sections. The geographic distribution of the variables used in the model is showed in the following figures.



2.5 Testing variables

Testing data are based on official daily figures published by the Italian government (link) and reported at the level of region. The first day for which figures are available is 24 February 2020. The case-to-test ratio is computed dividing the 7-day center-aligned moving average of number of positive tests divided by the 7-day center-aligned moving average of the number of tests.

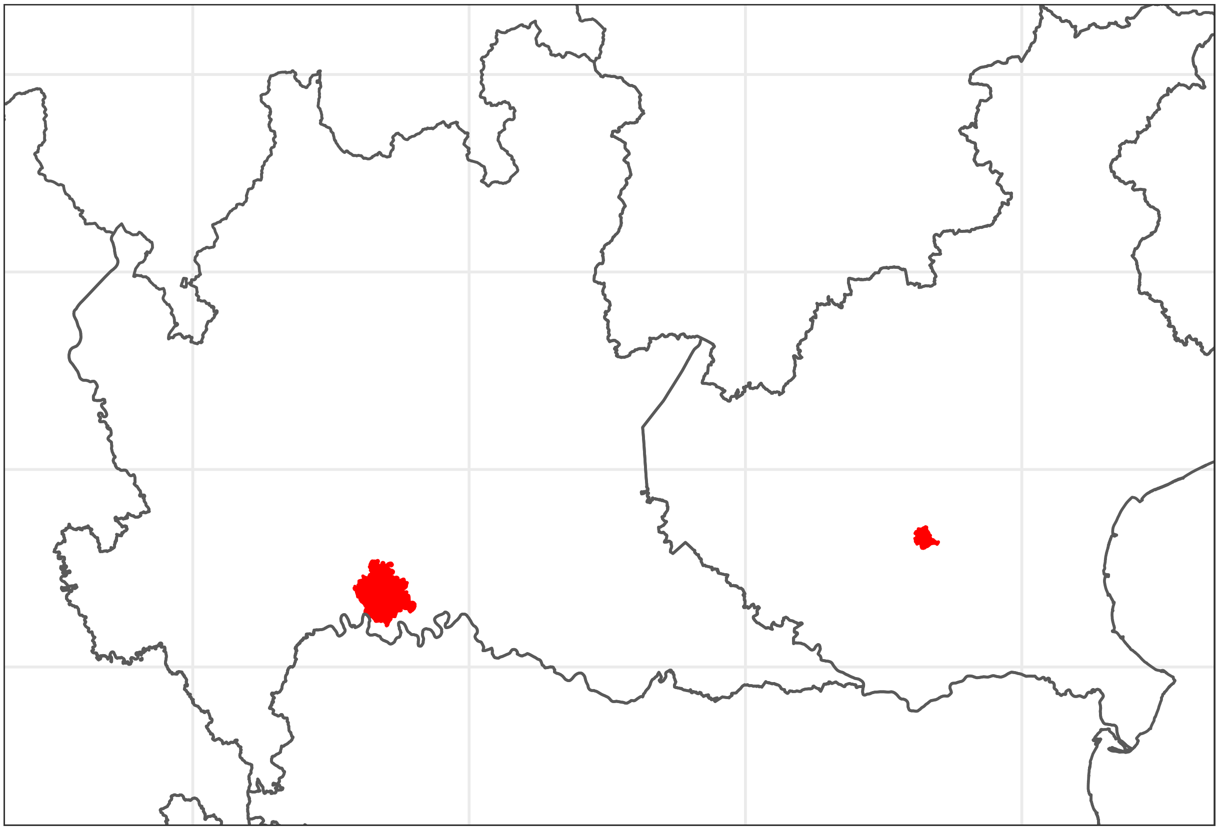
The following figure shows the evolution of the case-test ratio in the 20 Italian regions.

2.6 Lockdown

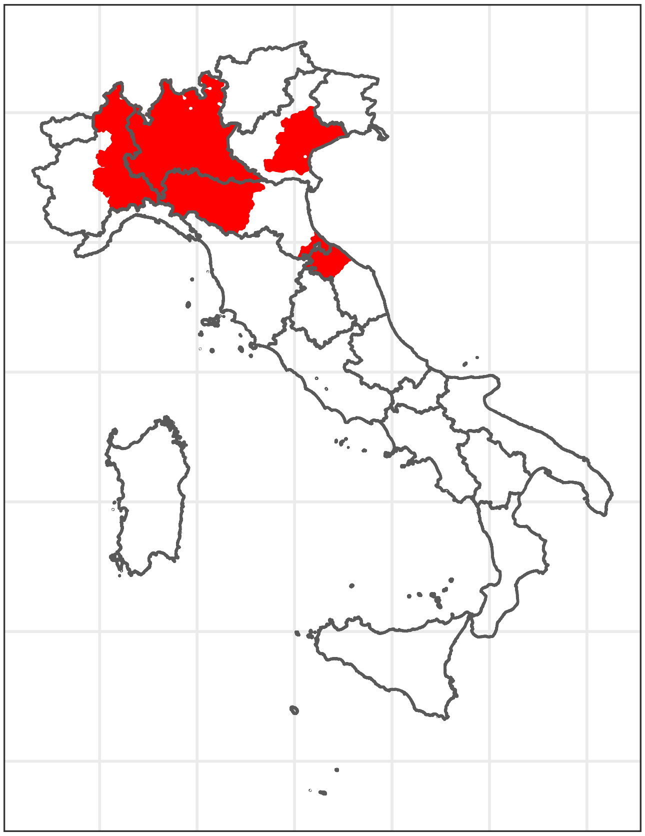
For each hexagon cell, the percentage of the population estimated to be living under lockdown measures in computed through the following steps:

1. The two areas resulting from the union of the area of the *comune* targetted by the lockdown measures established on 23 February 2020 and 8 March 2020 respectively (see below) are computed.
2. For each hexagon cell the percentage of the population under lockdown measures was estimated to be 0% (100%) if the area of the cell was included (excluded) entirely by the two lockdown areas.
3. When an hexagon cell was only partially covered by one of the two lockdown areas, the percentage of the population under lockdown measures was estimated by
4. summing the count of the population of all the census sections pertaining to a *comune* under lockdown measures and, based on the census section's centroid, intersecting with the hexagon cell.
5. dividing the resulting sum for the entire population estimated to live in the area of the hexagon cell (see above).

The following *comune* were subjected to lockdown measures on 23 February 2020: .



All the *comune* in Lombardy and in the the following *provincia*: Alessandria, Asti, Modena, Novara, Padua, Parma, Piacenza, Pesaro and Urbino, Reggio Emilia, Rimini, Treviso, Venice, Verbano-Cusio-Ossola, were subjected to lockdown measures on 08 March 2020.



2.7 Commuter-weighted network

A network connecting the hexagon cells and weighted based on the number of people communiting between pairs of cells is estimated based on a *comune*-to-*comune* asymmetric matrix populated with the count of commuters on census day 2011.

Consistently with the approach followed in computing other variables, we proceeded as following:

1. For each *comune*, we computed how its 2011 population distributes across different hexagon cells so to have the following table

Comune ID	Hexagonal cell ID	Population prop.
1	1	50%
1	2	50%
2	1	70%
2	2	30%
3	3	10%
3	4	10%
3	5	80%

2. We joined the resulting table with the long version version of the asymmetric matrix, so to have the following table

From comune ID	From hexagonal cell ID	From population prop.	To comune ID	N. commuters
1	1	50%	2	10

From comune ID	From hexagonal cell ID	From population prop.	To comune ID	N. commuters
1	1	50%	2	10
1	1	50%	2	10
1	1	50%	2	10
2	1	70%	3	15

3. We combined this last table with the table generated in the first step joining on `to Comune ID = Comune ID` resulting in

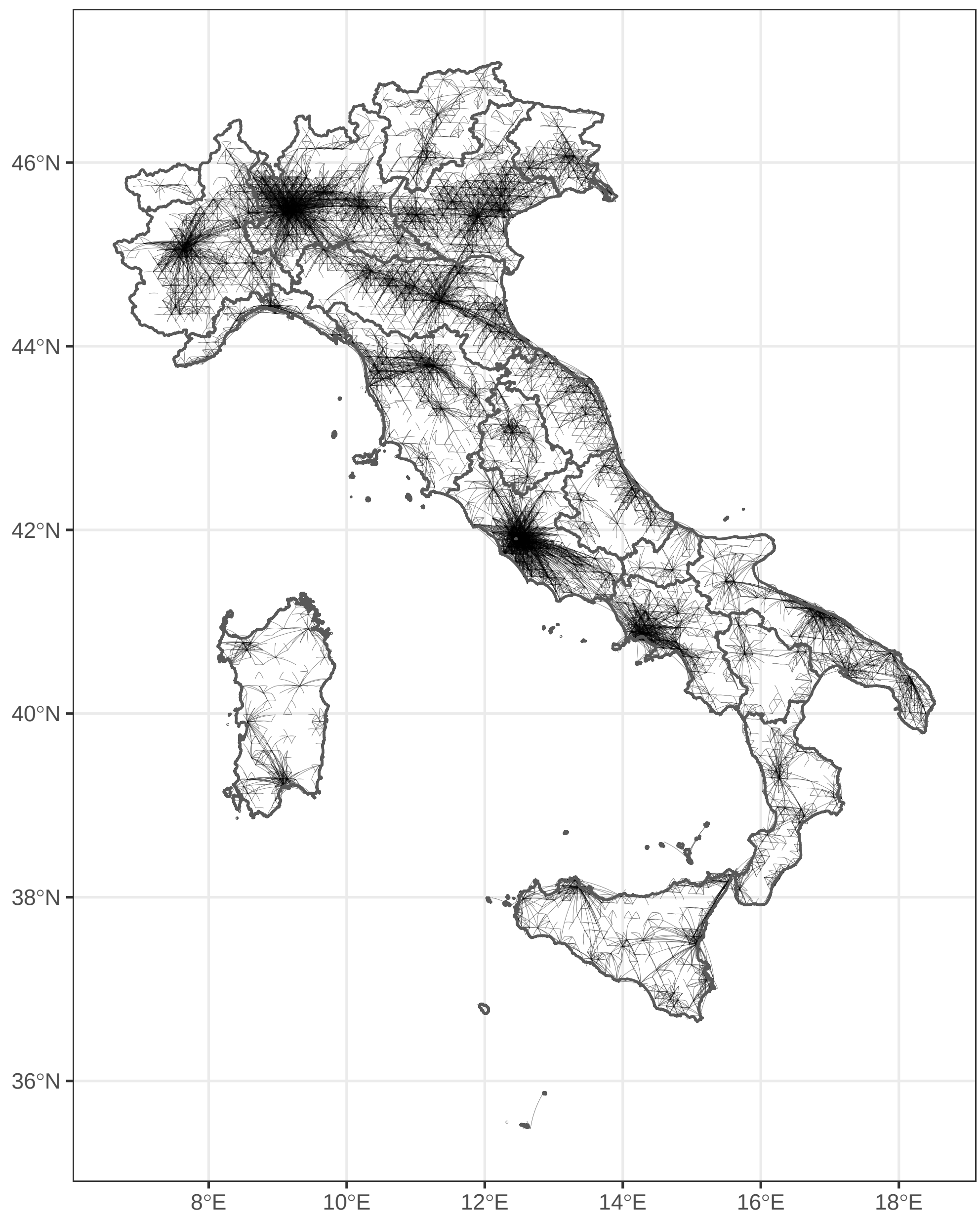
...	From hex. ID	From pop. %	...	N. commuters	To hex. ID	To pop. %
1	1	50%	2	10	2	30%
1	1	50%	2	10	3	10%
1	1	50%	2	10	4	20%
1	1	50%	2	10	5	10%
2	1	70%	3	15	4	100%

4. For each row of the resulting table, we then computed `weighted commuting` = `n. commuters` $\times$ `from pop. %` $\times$ `to pop. %` as the relationship between a pair of hexagonal cells definited by the communiting from $comune_A$ to $comune_B$.

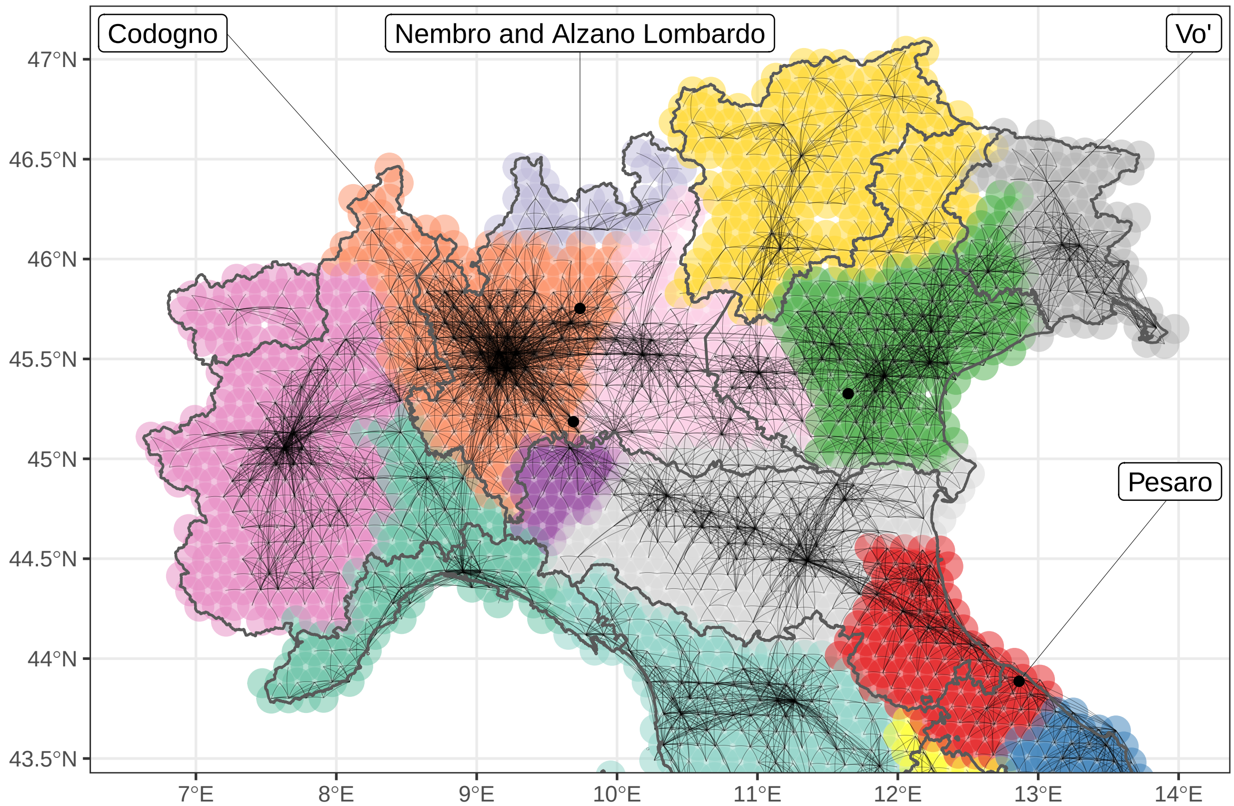
5. For each pair of hexagon cells A and B, we sum all the resulting `weighted commuting` to obtain $\texttt{weighted commuting}_{\vec{AB}}$ (which is different from $\texttt{weighted commuting}_{\vec{BA}}$).

6. We created direct network mapping the asymmetric relationships between pairs of hexagonal cells, where each relation was assigned weight $\texttt{weighted commuting}_{\vec{AB}}$

7. We simplify the directed network to obtain the undirected network with symmetrical relationship between pairs of hexagonal cell and weight $\texttt{weighted commuting}_{\overline{AB}}$ resulting from $\texttt{weighted commuting}_{\vec{AB}}$ + $\texttt{weighted commuting}_{\vec{BA}}$ illustrated by the following Figure.



The following Figure shows the same network and the communities identified by a community detection algorithm (A Clauset, MEJ Newman, C Moore: Finding community structure in very large networks, http://www.arxiv.org/abs/cond-mat/0408187).

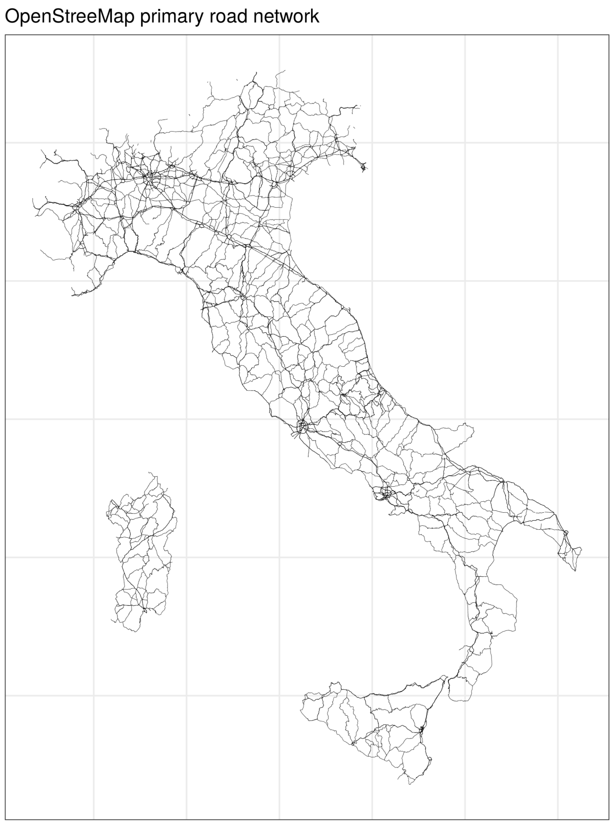


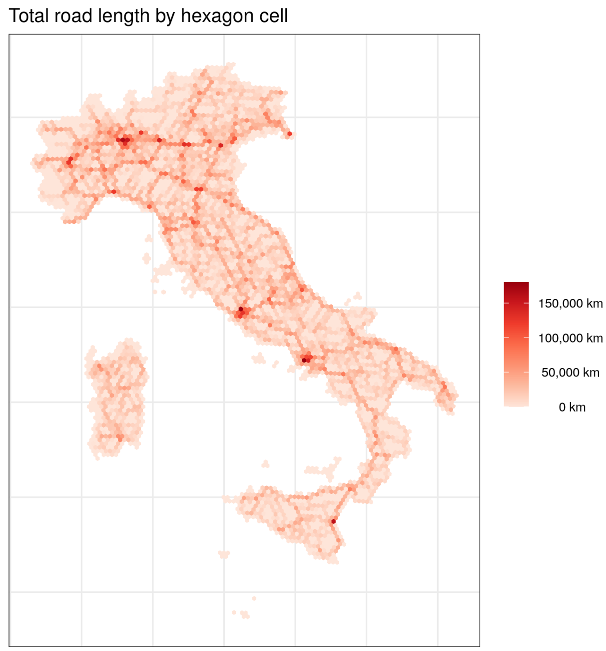
2.8 Neighboring excess mortality

We computed neighboring excess mortality based on the commuter-weighted network detailed above. For each day and for each cell, neighboring excess mortality was the average excess mortality in the other cells on that day wighted according to the corresponding `weighted commuting`$_{\overline{AB}}$.

2.9 Road density

The road density variable is computed based on the length of primary roads and based on OpenStreetMap data. For each hexagon cell, the total length of is computed by summing all the road segements contained in teh cell. Original data and length by cell is reported in the following figure.



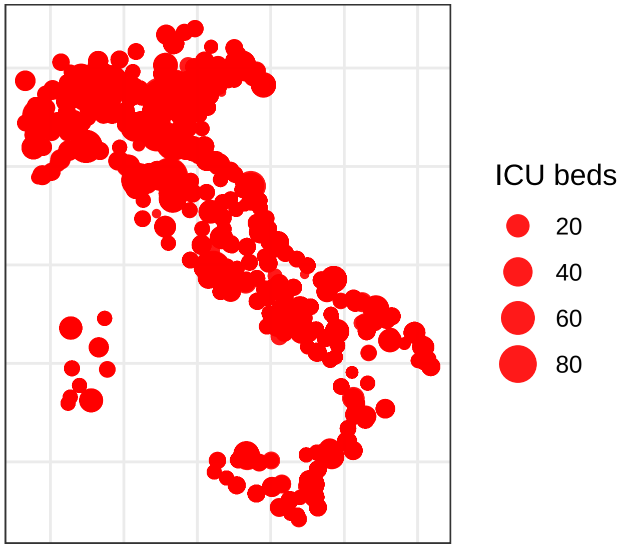


2.10 Hotel rooms

For each cell, we computed the number of hotel rooms based on data reported by Istat at the level of *comune* and according to the mapping method for count variables illustrated above.

2.11 Icu beds

The number of ICU beds ("terapia intensiva") was computed based on a list of health centres compiled by the Ministry of Health and containing the street address. From that list, we geocoded 2147 street address with at least one ICU beds using the Google Maps API. The distribution of ICU beds is mapped in the following Figure.


ICU beds
20
40
60
80

3 Modelling

3.1 Standardised coefficient table

Table 4:

	Dependent variable:			
	Excess mortality (backdated by 17 days)			
	OLS		*linear mixed-effects*	
	(1)	(2)	(3)	(4)
past excess mortality	0.202^{***} (0.005)	0.203^{***} (0.005)	0.199^{***} (0.005)	0.199^{***} (0.005)
perc. lockdown	0.025^{***} (0.006)	0.019^{***} (0.006)	0.026^{***} (0.006)	0.023^{***} (0.006)
neighbouring excess mortality	0.438^{***} (0.003)	0.439^{***} (0.003)	0.427^{***} (0.003)	0.428^{***} (0.003)
PM2.5 conc.	-0.026^{***} (0.003)		-0.016^{***} (0.004)	
PM10 conc.		-0.006^{**} (0.002)		-0.002 (0.002)
nitrogen monoxide conc.	0.213^{***} (0.015)	0.208^{***} (0.015)	0.229^{***} (0.015)	0.226^{***} (0.015)
ozone conc.	0.001 (0.004)	-0.005 (0.003)	-0.017^{***} (0.004)	-0.020^{***} (0.004)
carbon monoxide conc.	0.046^{***} (0.006)	0.031^{***} (0.005)	0.040^{***} (0.006)	0.030^{***} (0.005)
nitrogen dioxide conc.	0.042^{***} (0.010)	0.039^{***} (0.010)	0.025^{**} (0.010)	0.023^{**} (0.010)
sulphur dioxide	-0.0002 (0.002)	-0.002 (0.002)	-0.006^{**} (0.003)	-0.007^{***} (0.003)
temperature	0.00003 (0.003)	-0.004 (0.003)	-0.003 (0.003)	-0.006^{**} (0.003)
relative humidity	0.00004 (0.003)	-0.002 (0.003)	-0.009^{***} (0.003)	-0.011^{***} (0.003)
precipitation	-0.008^{***} (0.002)	-0.007^{***} (0.002)	-0.008^{***} (0.002)	-0.007^{***} (0.002)
perc. breast cancer incidence (reg.)	-0.003 (0.006)	-0.007 (0.006)	-0.040 (0.044)	-0.043 (0.045)
perc. colorectal cancer incidence (reg.)	-0.004 (0.003)	-0.003 (0.003)	0.011 (0.022)	0.012 (0.023)
perc. lung cancer incidence (reg.)	-0.001 (0.005)	-0.001 (0.005)	0.015 (0.030)	0.016 (0.031)
perc. diabetes (reg.)	-0.008^{*} (0.004)	-0.008^{*} (0.004)	-0.011 (0.036)	-0.012 (0.037)
perc. hypertension (reg.)	-0.023^{***} (0.004)	-0.023^{***} (0.004)	-0.028 (0.036)	-0.027 (0.036)
perc. bronchitis (reg.)	0.025^{***} (0.004)	0.024^{***} (0.004)	0.011 (0.030)	0.011 (0.030)
perc. allergies (reg.)	0.019^{***} (0.003)	0.019^{***} (0.003)	0.023 (0.023)	0.023 (0.023)
perc. heart cond. (reg.)	-0.008^{**} (0.004)	-0.007^{**} (0.004)	0.014 (0.029)	0.014 (0.030)
perc. smokers (reg.)	-0.020^{***} (0.004)	-0.020^{***} (0.004)	-0.028 (0.027)	-0.029 (0.027)
perc. ICU beds 50km radius	-0.018^{***} (0.002)	-0.018^{***} (0.002)	-0.011^{***} (0.002)	-0.011^{***} (0.002)
cases vs tests ratio (reg.)	0.040^{***} (0.003)	0.040^{***} (0.003)	0.037^{***} (0.003)	0.037^{***} (0.003)
median density	-0.008^{***} (0.003)	-0.008^{***} (0.003)	-0.013^{***} (0.003)	-0.014^{***} (0.003)
perc. females	0.0001 (0.003)	0.00001 (0.003)	0.002 (0.003)	0.002 (0.003)
perc. over 65	-0.002 (0.003)	-0.002 (0.003)	0.004 (0.003)	0.004 (0.003)
over 74 conc.	0.001 (0.002)	0.001 (0.002)	-0.004 (0.002)	-0.004^{*} (0.002)
perc. working out of comune	-0.018^{***} (0.003)	-0.018^{***} (0.003)	-0.020^{***} (0.003)	-0.020^{***} (0.003)
perc. families over 4	-0.001 (0.003)	-0.002 (0.003)	0.001 (0.003)	0.001 (0.003)
perc. employed	-0.001 (0.003)	-0.001 (0.003)	0.003 (0.003)	0.002 (0.003)
perc. degree (over workforce)	0.002 (0.003)	0.003 (0.003)	0.006^{**} (0.003)	0.006^{**} (0.003)
perc. low income	0.006 (0.004)	0.006 (0.004)	0.003 (0.005)	0.003 (0.005)
log. agedcare employees	0.030^{***} (0.003)	0.031^{***} (0.003)	0.026^{***} (0.003)	0.026^{***} (0.003)
log. construction employees	-0.026^{***} (0.007)	-0.026^{***} (0.007)	-0.021^{***} (0.007)	-0.021^{***} (0.007)
log. delivery employees	0.003 (0.005)	0.003 (0.005)	0.007 (0.005)	0.007 (0.005)
log. education employees	0.004 (0.006)	0.004 (0.006)	-0.0001 (0.006)	-0.001 (0.006)
log. healthcare employees	0.018^{***} (0.005)	0.018^{***} (0.005)	0.012^{**} (0.005)	0.012^{**} (0.005)
log. hospitality employees	-0.038^{***} (0.006)	-0.039^{***} (0.006)	-0.033^{***} (0.007)	-0.033^{***} (0.007)
log. manufacturing employees (exc. meat)	0.012^{**} (0.006)	0.012^{*} (0.006)	0.011^{*} (0.006)	0.010^{*} (0.006)
log. manufacturing employees (meat only)	0.021^{***} (0.002)	0.021^{***} (0.002)	0.020^{***} (0.002)	0.020^{***} (0.002)
log. other employees	-0.018^{***} (0.006)	-0.019^{***} (0.006)	-0.024^{***} (0.006)	-0.025^{***} (0.006)
log. retail employees	-0.0001 (0.010)	0.001 (0.010)	-0.0002 (0.010)	0.0003 (0.010)
log. services employees	0.019^{**} (0.010)	0.020^{**} (0.010)	0.024^{**} (0.010)	0.025^{**} (0.010)
log. transportation employees	0.006 (0.005)	0.005 (0.005)	0.013^{***} (0.005)	0.014^{***} (0.005)
log. warehousing employees	-0.008^{**} (0.004)	-0.007^{**} (0.004)	-0.006 (0.004)	-0.005 (0.004)
log. wholesaling employees	0.032^{***} (0.008)	0.033^{***} (0.008)	0.029^{***} (0.008)	0.030^{***} (0.008)
log. hotel beds	0.013^{***} (0.003)	0.013^{***} (0.003)	0.013^{***} (0.003)	0.014^{***} (0.003)
tot. length primary roads	0.044^{***} (0.003)	0.044^{***} (0.003)	0.045^{***} (0.003)	0.046^{***} (0.003)
past excess mort. X perc. lockdown	-0.071^{***} (0.006)	-0.072^{***} (0.006)	-0.074^{***} (0.006)	-0.075^{***} (0.006)
past excess mort. X neighb. excess mort.	0.009^{***} (0.0003)	0.009^{***} (0.0003)	0.010^{***} (0.0003)	0.010^{***} (0.0003)
neighb. excess mort. X perc. lockdown	-0.109^{***} (0.004)	-0.110^{***} (0.004)	-0.111^{***} (0.004)	-0.112^{***} (0.004)
past excess mort. X neighb. excess mort. X perc. lockdown	0.003^{***} (0.0005)	0.003^{***} (0.0005)	0.003^{***} (0.0005)	0.003^{***} (0.0005)
Constant	0.059^{***} (0.005)	0.062^{***} (0.005)	0.053^{***} (0.019)	0.055^{***} (0.019)
Observations	239,463	239,463	239,463	239,463
R^2	0.431	0.431		
Adjusted R^2	0.431	0.431		
Log Likelihood			$-330,453.100$	$-330,463.300$
Akaike Inf. Crit.			661,016.300	661,036.600
Bayesian Inf. Crit.			661,587.500	661,607.800
Residual Std. Error (df = 239410)	0.962	0.962		
F Statistic (df = 52; 239410)	$3,487.945^{***}$	$3,486.212^{***}$		
Note:				$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

3.2 Standardised coefficient plot