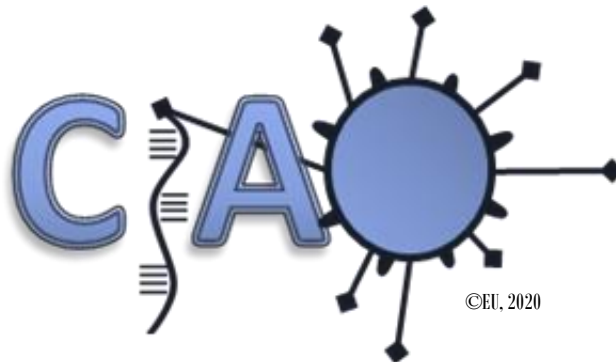




MODELING OF THE COVID-19 EPIDEMIC USING
GEOSTATISTICAL METHODS AND BASED ON THE
INVESTIGATION OF SPATIAL CORRELATIONS BETWEEN
ATMOSPHERIC POLLUTION, METEOROLOGICAL FACTORS,
AGE, SEX, HEALTH CONDITION, AND LOCKDOWN PHASES
WITH LEVELS OF VIRUS CONTAGION AND LETHALITY

Technical Report



Project Date:
01/06/2020 – 01/12/2020

MIRSINA MOUSAVI AGHDAM^{1, 2}, QUENTIN CROWLEY¹
1. Trinity College Dublin and 2. Crè-IQ

“Do what you can, where you are, with what you have.” – Theodore Roosevelt

“To all doctors, nurses, professionals and front line health workers of Ireland”

Abstract

The current COVID-19 pandemic is having a catastrophic effect on human health and global economies. In terms of groups most at risk, the elderly and those with underlying medical conditions are disproportionately affected by the virus, with a corresponding higher number of fatalities. So far, environmental factors affecting transmission and infection rates are under-investigated. Given that COVID-19 attacks the respiratory system, environmental factors that also impact respiratory health may leave an individual more prone to infection. To contain the spread of the COVID-19 virus, governments have enforced multi-phase restrictions on outdoor activities or even collective quarantine on the population. This proof of concept study aims to gain a better understanding of a) environmental exposures which may lead to increased Severe Acute Respiratory Syndrome infection rates, b) the impact of gender-specific lifestyle, health condition and vulnerability by age and sex, and also c) the Impact of setting lockdown and physical distancing. These environmental, statistical, and epidemiological parameters were considered as inputs of a model that aims to predict the contagion and lethality of Covid-19 in Ireland. A whole-system approach has been applied to millions of time series data based on the principal component regression (PCR) model. To develop the PCR model, the weekly contagion and lethality rates were set as the response variables. Forty-one explanatory parameters including, mean values of weekly atmospheric pollutants (i.e. PM10, PM2.5, O3, SO2 and NO2, and Indoor radon levels) metrological parameters (i.e. precipitation amount, air temperature, wet bulb air temperature, dew point air temperature, relative humidity, vapor pressure, mean sea level pressure, predominant hourly wind speed, and sunshine duration), Sociodemographic data (self-perceived health status, population, population density, distribution of sex and age profile) together with lockdown phases were introduced as predictors. The adjusted R-squared values which compare the explanatory power of the PCR regression models of the Covid-19 contagion and lethality were found to be 0.53 and 0.54, respectively. The rather good adjusted R-squared values express that through the prediction models presented in this research, one can explain more than 50 percent of the variability in contagion and lethality rates of the SARS-CoV-2 virus.

As a result of this study it was found that for the contagion model, the factors of PM10, PM2.5, O3, dew point air temperature, vapor pressure predominant hourly wind speed, sunshine

duration, health status, age profile and lockdown phases would bring significant information. The decrease of dew point air temperature and increase of vapor pressure has the highest standardized effect on the increase of covid-19 contagion rate.

On the other hand, for the lethality rate prediction model, the PM10, NO2, Indoor radon levels, air temperature, vapor pressure, mean sea level pressure, health status, age profile and lockdown phases) bring useful information that improves the model. Similar to the contagion model, a decrease of air temperature and increase of vapor pressure were found to have the highest standardized effect on the increase of covid-19 lethality rate.

A key deliverable of the project is an early warning system for the identification of future high-risk environmental conditions. This can be used as a tool-kit to support decision-making on local, regional, national, or international levels. Possible societal impacts are (1) improved protection of public health, (2) focused restrictions for most at-risk locations and durations, and (3) more informed decision making by governments and health agencies for planning and mitigation of future outbreaks of highly infectious and acute respiratory diseases.

Table of Contents

Abstract	2
List of Figures	6
List of Tables	8
1 Introduction.....	9
1.1 People most at risk of developing severe forms of the disease and contracting the infection.....	10
1.2 Effect of social distancing and population density	11
1.3 Effect of isolation and lockdown measures	11
1.4 Effect of atmospheric pollution	11
1.5 Effect of metrological parameters.....	12
2 Situation of Covid-19 in Ireland.....	13
2.1 Sociodemographic profile.....	16
2.2 Health profile	16
2.3 Weather information.....	17
2.4 Air pollution monitoring	20
2.5 Indoor Radon risk.....	20
2.6 Covid-19 Lockdown in Ireland	22
3 Methodology	23
3.1 Model Setting.....	24
3.2 Model validity.....	24
4 Results	26
4.1 Principal Component Analysis (PCA).....	26
4.2 Principle Components Regression (PCR)	28
4.3 Model validity.....	33

4.4 Comparison of observed and predicted weekly new cases and deaths of Covid-19 in Dublin County and the other Irish Counties.....	35
5 Discussion, Conclusions, and recommendations	38
6 Openness and Transparency	39
7 Acknowledgements.....	40
8 References.....	41
Appendix I: Supplementary data	44

List of Figures

Figure 1: Daily new confirmed COVID-19 cases (and 3-Day averages, bold line) in Ireland. Data source: www.data.gov.ie (01/02/2020 – 30/10/2020).....	13
Figure 2: Daily confirmed death (and 3-Day averages, bold line) due to COVID-19 in Ireland. Data source: www.data.gov.ie (01/02/2020 – 30/10/2020).....	14
Figure 3 Total number of COVID-19 infections per 100,000 population by county (03 April to 30 October 2020)	15
Figure 4 Total number of COVID-19 mortalities per 100,000 population by county (03 April to 30 October 2020)	15
Figure 5: Pie chart of the self-stated health condition of the Irish population	17
Figure 6: Weather and air pollution monitoring stations of Ireland	18
Figure 7: Weekly average of precipitation amount in mm, sunshine duration in hour (left y-axis), and Relative Humidity (rhum) in % (right y-axis) for County Dublin. Based on data driven from Met Éireann.	19
Figure 8: Weekly average of air temperature (temp), wet bulb air temperature(wetb), dew point air temperature(dewpt) in °C, wind speed (wdsp) in kt, Vapour Pressure in hpa (left y-axis), and mean Sea Level Pressure in hpa (right y-axis) for County Dublin. Based on data driven from Met Éireann.....	19
Figure 9: Average of atmospheric pollutants concentration ($\mu\text{g m}^{-3}$) for monitoring stations of County Dublin- the left vertical axis shows the values of PM10, PM2.5, NO3, and O3. The right axis expresses the SO2 and CO contents. Based on data driven from https://airquality.ie	20
Figure 10: Indoor radon map of Ireland- Data source EPA- www.epa.ie	21
Figure 11: Covid-19 Lockdown phases in Ireland. Figure courtesy Aine McMahon-PA Graphics.....	22
Figure 12: Scree plot showing the Eigenvalues and the Cumulative variability in PCA	26
Figure 13: Correlation circle of the factors in PCA	27
Figure 14: Correlation between predicted and observed values of Covid -19 weekly new cases.....	29

Figure 15 Correlation between predicted and observed values of Covid-19 weekly confirmed deaths ...	29
Figure 16: Chart of standardized coefficients of the model for prediction weekly new case of Covid-19 .	32
Figure 17: Chart of standardized coefficients of the model for prediction weekly deaths of Covid-19	33
Figure 18: Total number of predicted and observed cases of Covid-19. (3 April- 30 October).....	35
Figure 19: Total number of predicted and observed deaths of Covid-19. (3 April- 30 October).....	36
Figure 20: Predicted and Observed weekly new cases in County Dublin	37
Figure 21: Predicted and Observed weekly death in County Dublin	37

List of Tables

Table 1: Distribution of sex and the total population of Irish counties. Data Source: Central Statistics Office https://www.cso.ie/en/census/	16
Table 2: Population estimates (percent of the total population) for regions by age group- Data Source: Central Statistics Office https://www.cso.ie/en/census/	17
Table 3: Self-stated health status in percent of the total population- Data Source: Central Statistics Office https://www.cso.ie/en/census/	16
Table 4: Summary of the statistics and normality test of response and predictor variables	25
Table 5: Eigenvalue of the Principal Component Analysis (only 11 factors of 39 are shown here) - For more details, we refer to Appendix I	26
Table 6: Goodness of fit statistics	28
Table 7: Results of variance analysis.....	30
Table 8: Model parameters for prediction of Covid-19 weekly new cases and deaths.....	31
Table 9: Results of cross-validation of the observed and predicted values	34
Table 10: Mean square errors (MSE) and Root mean square errors (RMSE) of Log transformed weekly Covid-19 new cases (NC) and confirmed deaths (CD).....	34

1 Introduction

The severe acute respiratory syndrome Coronavirus-2 (SARS-CoV-2) is the name given to the new coronavirus of 2019 and COVID-19 is the name given to the disease associated with the virus. The coronaviruses (CoV) are a 'large family of respiratory viruses that can cause disease of mild to moderate level, from the common cold to respiratory syndromes such as MERS (Middle East Respiratory Syndrome) to SARS (severe acute respiratory syndrome). They are positive-stranded RNA viruses, with a corona-like appearance under the electron microscope.

Coronaviruses were identified in the mid-1960s and are known to infect humans and some animals (including birds and mammals). The primary target cells are epithelial, those of the respiratory tract and gastrointestinal.

To date, seven coronaviruses can infect humans. Some were identified several years ago (the first in the mid-1960s) and some in the new millennium. Common human coronaviruses are as follows (WHO, 2020):

- HCoV-229E (Alphacoronavirus)
- HCoV-NL63 (Alphacoronavirus)
- HCoV-OC43 (Betacoronavirus)
- HCoV-HKU1 (Betacoronavirus)
- MERS-CoV (Betacoronavirus that causes Middle East respiratory syndrome)
- SARS-CoV (Betacoronavirus that causes Severe acute respiratory syndrome)
- SARS-CoV-2 (Betacoronavirus that causes COVID- 19)

The virus responsible for the current epidemic is a new version of coronavirus never previously identified in humans. The appearance of new pathogenic viruses for humans, previously circulating only in the animal world, is a widely known phenomenon (called spillover or species jump) and it is thought that it may also be at the basis of the origin of the new coronavirus. The scientific community is currently trying to identify the source of the infection. Currently, available evidence suggests that SARS-CoV-2 has an animal origin and is not a constructed virus. The ecological reservoir of SARS-CoV-2 most likely resides in bats. (Decaro & Lorusso, 2020)

1.1 People most at risk of developing severe forms of the disease and contracting the infection

The disease is characterized by a lethality (number of deaths out of the total number of patients) lower than that observed for other coronaviruses responsible for epidemics in the past such as SARS and MERS, even if its contagiousness is greater than that observed in the two competing coronaviruses (Petrosillo et al., 2020).

Generally, older people (over 70 years old) and those with pre-existing diseases, such as arterial hypertension, heart problems, diabetes, chronic respiratory diseases, cancer, and immunosuppressed patients (due to congenital or acquired disease, transplanted or treated with immunosuppressive drugs) are more likely to develop severe forms of the disease (Flaherty et al., 2020).

Many coronaviruses can be transmitted from person to person, usually after close contact with an infected patient, for example among family members or in a healthcare setting. The novel coronavirus responsible for respiratory disease COVID-19 can also be passed from person to person via close contact with a probable or confirmed case (CDC, 2020).

Persons who live or who have traveled in areas at risk of infection by a new coronavirus or people who meet the criteria close contact with a confirmed or probable case of COVID-19 are at most risk of contracting the virus. Areas at risk for new coronavirus infection are those where local transmission of SARS-CoV-2 is present, as identified by the World Health Organization. These should be differentiated from areas where only imported cases are present.

Even health care workers may be at greater risk of contracting the disease because they come into contact with patients more often than the general population. The World Health Organization (WHO) recommends that health care providers implement adequate measures for prevention and control of infections in general and respiratory infections in particular.

The children thus far constitute a very small percentage of cases COVID-19 reported. Children appear to be more likely to be infected as adults but have a much lower risk than adults of developing symptoms or severe forms of the disease.

1.2 Effect of social distancing and population density

Since a vaccine does not currently exist, the response to the pandemic has focused on non-pharmaceutical interventions, involving the general reduction of social interactions, and in particular the isolation of persons with actual or suspected infection (te Vrugt et al., 2020). The recent studies explore that the spreading of COVID-19 is concerning population density and social distancing. Social distancing measures seem to have had the biggest impact in areas where people are living closest together (e.g. County Dublin). Therefore Adherence to the social measures might be most challenging in areas of high population density given the high levels of interpersonal contact (Tammes, 2020).

1.3 Effect of isolation and lockdown measures

Despite involving short-term economic costs, to prevent the spread of the Covid-19 pandemic, governments have started to apply bans under many social restrictions. Lockdown is at the forefront of these restrictions. The results of the statistical analysis show that lockdown plays an important role in preventing COVID-19 and the spread of the virus can be significantly reduced by preventive restrictions. The application of lockdown by governments is also believed to be effective in psychology, environment, and economy besides having an impact on Covid-19. (Atalan, 2020).

1.4 Effect of atmospheric pollution

Fine powders and aerosols can provide the possibility for viruses to attach themselves to fine dust present in the air and thus be transported by the wind for large distances or remain suspended in the air. A study concluded that the high air pollution loading could be a co-factor causing the high fatality rate due to the COVID-19 infection (Conticini et al., 2020). Prior exposure to air pollution may aggravate the health impacts of COVID-19 and increase the risk of death by suppressing immunity. Therefore, the effect of atmospheric pollution (i.e. PM10, PM2.5, O3, SO2, and NO2) should be considered as an additional co-factor for increasing the level of Covid-19 lethality. Among the studied atmospheric factors PM2.5, ozone (O3), sulfur dioxide (SO2), and to

a less extent also PM10 show a positive correlation with virus transmission (Conticini et al., 2020; Copat et al., 2020; Lolli et al., 2020).

Studies are also underway to verify the transmission of the virus by aerosol in the indoor environments: some medical procedures, such as those performed in a hospital setting in the care of COVID-19 patients, can produce very small droplet droplets (called aerosolized or aerosolized droplet nuclei) that can remain suspended in the air for a long time and can potentially be inhaled by other people if they are not wearing adequate personal protective equipment (Jayaweera et al., 2020). COVID-19 outbreaks have been reported in some closed environments, such as restaurants, nightclubs, places of worship, and work where people can scream, talk, or sing. In these outbreaks, aerosol transmission cannot be ruled out, particularly in closed, crowded, and poorly ventilated places where infected people spend long periods with others. Further studies are needed to investigate these episodes and evaluate their importance for the transmission of the virus (WHO, 2020b).

1.5 Effect of metrological parameters

A study on the spreading of COVID-19 in Saudi Arabia revealed that the number of COVID-19 positive cases increases due to the decrease of temperature or humidity, whereas an average decrease in the wind speed was also found to be associated with an elevation of the number of positive cases (Alkhowailed et al., 2020). Furthermore, a study on the effects of meteorological factors on COVID-19 cases in India showed a positive association between daily new cases of COVID-19 with temperature. Relative and absolute humidity was also found to have shown an association with daily new cases of COVID-19. According to these, it can be concluded that COVID-19 pandemic transmission prefers dry and cool environmental conditions, as well as polluted air. For those reasons, the virus might easier spread in unfiltered air-conditioned indoor environments (Lolli et al., 2020).

2 Situation of Covid-19 in Ireland

The COVID-19 pandemic reached the Republic of Ireland on 29 February 2020, and within three weeks, cases had been confirmed in all counties. The first wave of the virus was between February–May 2020 adding that the peak of daily new cases and confirmed death was reported in mid-April (Figure 1 and Figure 2). Later the National Public Health Emergency Team reported that lockdown and other measures had driven the growth rate of the pandemic. Due to a significant fall in the contagion and lethality rate, the Irish government began to easing COVID-19 restrictions in early May. The easing of restrictions continued until August 2020. By starting the second wave of virus spread and following significant increases of COVID-19 cases in the three counties (i.e. Kildare, Laois, and Offaly), a set of new restrictive measures were considered in early August. In October 2020, Ireland introduced its highest level of national restrictions as part of a lockdown that will last six weeks. To date, Ireland is passing the second week of the covid-19 pandemic. But as can be understood from the statistical data, it seems that the full lockdown has beaten the second Covid-19 wave in Ireland.

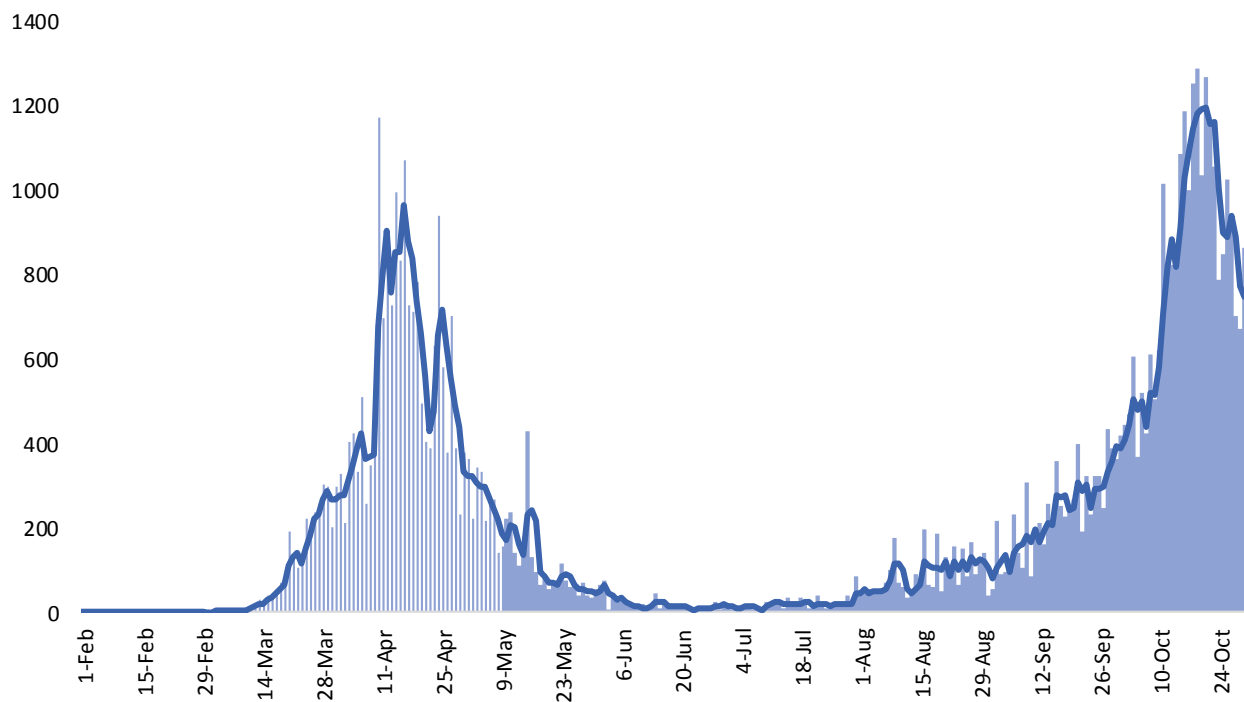


Figure 1: Daily new confirmed COVID-19 cases (and 3-Day averages, bold line) in Ireland. Data source: www.data.gov.ie (01/02/2020 – 30/10/2020)

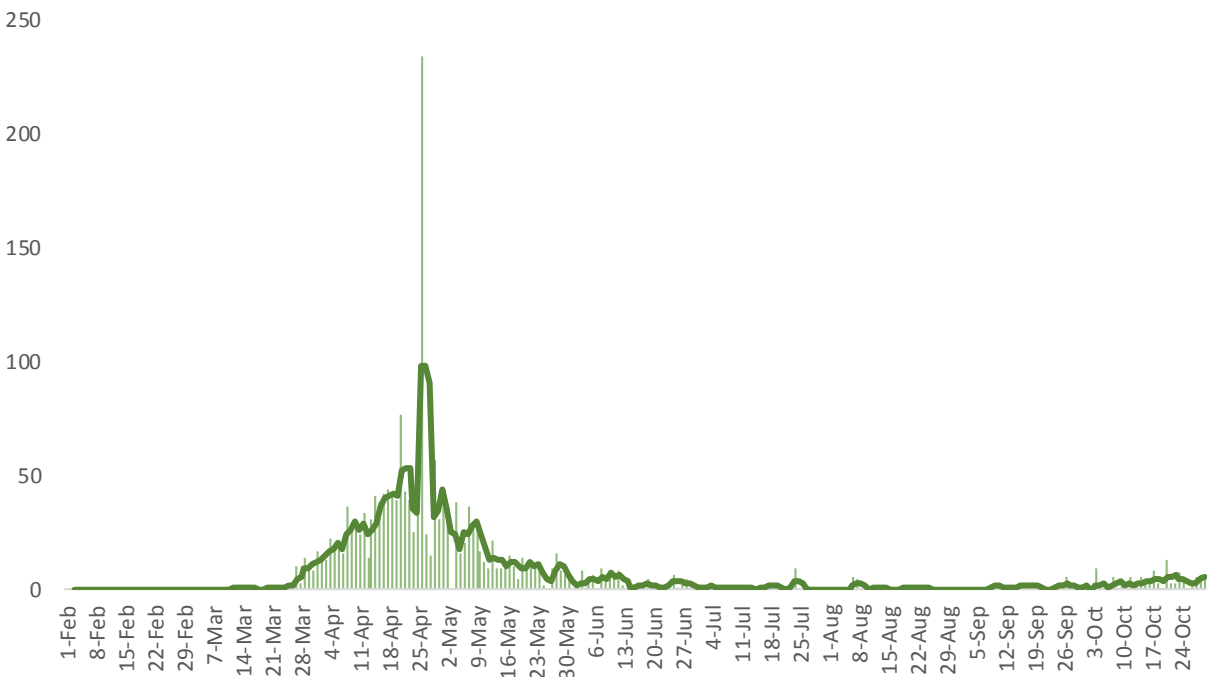


Figure 2: Daily confirmed death (and 3-Day averages, bold line) due to COVID-19 in Ireland. Data source: www.data.gov.ie (01/02/2020 – 30/10/2020)

Among all Irish counties, Meath, Cork, and Laois have the highest total weekly new cases per 100,000 population by county. Counties Laois, Limerick, and Wexford also have the highest total weekly death rates (For the period of 03 April to 30 October 2020). The number of cases in Ireland translates to a rate of around 1,430 cases per 100,000 population. In terms of deaths from the virus, Ireland has recorded 41.23 deaths per 100,000, which was the tenth highest in the EEA as of November 22, 2020. Like other western European countries, high COVID-19 morbidity and mortality (about 20 percent of total death) was observed among residents in long-term care facilities.

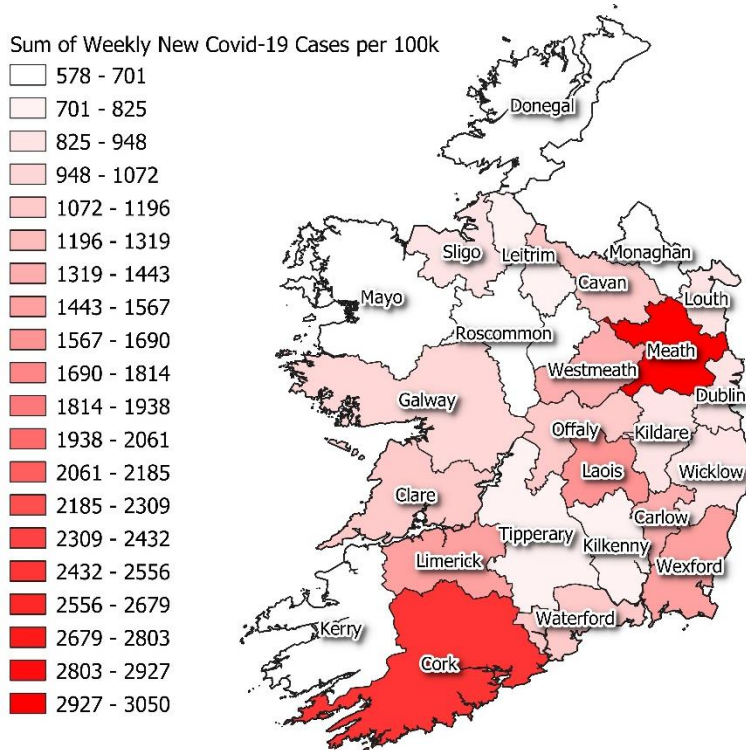


Figure 3 Total number of COVID-19 infections per 100,000 population by county (03 April to 30 October 2020)

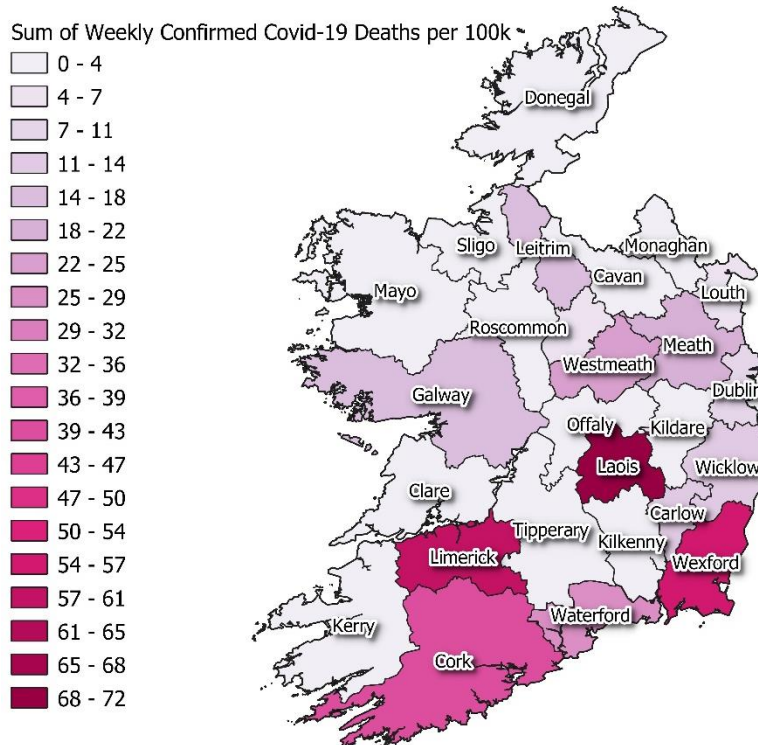


Figure 4 Total number of COVID-19 mortalities per 100,000 population by county (03 April to 30 October 2020)

2.1 Sociodemographic profile

According to the most recent census (2016), a total number of 4,689,921 people were resident in Ireland. The breakdown of the population estimates (percent of the total population) by region, sex, and age group are contained in Table 1 and Table 2. There is a heavy concentration of the population in Dublin and the Mid-East region that surrounds Dublin, with over 43% of the population living in that area. Overall, the population density is 72 people per Km². The proportion of the population aged 65 or older is a little over 14% while the proportion aged over 85 is just over 1.5% (Kennelly et al., 2020).

Table 1: Distribution of sex and the total population of Irish counties. Data Source: Central Statistics Office <https://www.cso.ie/en/census/>

County	Percent of Female	Percent of male	Population
Carlow	50.00	50.00	56932
Cavan	49.68	50.32	76176
Clare	50.52	49.48	118817
Cork	50.51	49.49	542868
Donegal	50.36	49.64	159192
Dublin	51.14	48.86	1347359
Galway	50.53	49.47	258058
Kerry	50.54	49.46	147707
Kildare	50.32	49.68	222504
Kilkenny	50.08	49.92	99232
Laois	49.45	50.55	84697
Leitrim	49.87	50.13	32044
Limerick	50.06	49.94	194899
Longford	49.63	50.37	40873
Louth	50.63	49.37	128884
Mayo	50.16	49.84	130507
Meath	50.38	49.62	195044
Monaghan	49.72	50.28	61386
Offaly	50.18	49.82	77961
Roscommon	49.84	50.16	64544
Sligo	50.61	49.39	65535
Tipperary	50.07	49.93	159553
Waterford	50.38	49.62	116176
Westmeath	50.34	49.66	88770
Wexford	50.74	49.26	149722
Wicklow	50.76	49.24	142425

Table 2: Population estimates (percent of the total population) for regions by age group- Data Source: Central Statistics Office <https://www.cso.ie/en/census/>

Age Group County	<20	20_24	25_29	30_34	35_39	40_44	45_49	50_54	55_59	60_64	65_69	70_74	75_79	80_84	85>
Carlow	28.63	5.84	5.95	7.26	7.90	7.55	6.87	6.34	5.70	5.03	4.33	3.32	2.38	1.59	1.30
Cavan	29.98	4.63	5.24	7.12	7.63	7.21	6.95	6.52	5.87	5.14	4.52	3.38	2.39	1.84	1.62
Clare	28.18	4.84	4.74	6.27	7.47	7.67	7.36	6.79	6.23	5.59	5.18	3.85	2.58	1.75	1.50
Cork	27.33	5.99	5.98	7.44	8.13	7.49	6.88	6.44	5.70	5.02	4.45	3.44	2.49	1.79	1.43
Donegal	28.83	4.74	5.07	6.17	7.30	7.17	7.11	6.47	5.89	5.53	5.14	4.08	2.87	1.92	1.69
Dublin	25.14	6.83	8.28	9.16	8.95	7.44	6.42	5.82	5.23	4.50	3.91	3.04	2.30	1.63	1.36
Galway	27.31	6.38	5.93	7.31	8.21	7.53	6.86	6.12	5.67	5.16	4.43	3.41	2.42	1.74	1.53
Kerry	25.66	4.55	4.82	6.58	7.46	7.24	7.05	6.87	6.63	6.20	5.81	4.43	3.00	2.05	1.66
Kildare	31.03	5.61	5.50	7.65	8.73	8.40	7.30	6.25	5.22	4.41	3.75	2.64	1.61	1.02	0.88
Kilkenny	28.53	4.70	5.33	6.75	7.86	7.73	6.92	6.57	6.07	5.37	4.67	3.65	2.49	1.77	1.57
Laois	30.86	4.85	5.78	7.89	8.51	7.52	7.02	6.10	5.46	4.68	3.84	2.98	1.97	1.41	1.14
Leitrim	27.47	4.04	4.47	6.19	7.28	7.28	6.91	6.64	6.58	6.25	5.45	4.14	3.01	2.13	2.14
Limerick	27.10	6.47	6.05	7.31	7.80	7.22	6.63	6.19	5.83	5.32	4.77	3.64	2.57	1.73	1.36
Longford	29.63	4.78	5.44	6.81	7.82	7.11	6.85	6.08	5.83	5.39	4.87	3.70	2.38	1.84	1.46
Louth	29.65	5.49	5.91	7.18	7.95	7.75	7.18	6.32	5.42	4.69	4.16	3.28	2.27	1.52	1.24
Mayo	26.74	4.44	4.72	6.21	6.79	6.91	6.84	6.82	6.72	6.23	5.85	4.36	3.10	2.26	1.98
Meath	31.75	4.81	5.01	7.19	8.63	8.53	7.46	6.34	5.19	4.45	3.79	2.84	1.81	1.19	1.02
Monaghan	29.28	4.85	5.65	7.11	7.67	7.24	6.71	6.37	5.91	5.25	4.68	3.43	2.42	1.88	1.57
Offaly	29.68	5.06	5.28	6.91	7.60	7.39	6.90	6.61	5.86	5.11	4.54	3.37	2.47	1.78	1.43
Roscommon	27.38	4.26	4.80	6.29	7.11	7.09	6.96	6.76	6.54	6.16	5.25	4.09	2.91	2.28	2.12
Sligo	27.06	5.49	4.89	6.51	7.27	6.89	6.63	6.71	6.51	5.84	5.43	4.03	2.86	2.05	1.85
Tipperary	27.95	4.86	5.08	6.50	7.40	7.34	7.06	6.76	6.17	5.59	4.96	3.87	2.75	2.05	1.67
Waterford	27.74	5.26	5.39	6.51	7.77	7.28	7.05	6.74	5.91	5.40	4.91	3.88	2.76	1.93	1.47
Westmeath	29.01	5.65	5.85	7.28	7.85	7.30	7.05	6.49	5.72	4.99	4.27	3.22	2.35	1.63	1.33
Wexford	29.01	5.02	5.09	6.78	8.07	7.99	7.25	6.88	5.81	5.06	4.50	3.51	2.32	1.49	1.22
Wicklow	28.54	4.79	5.20	6.53	7.57	7.44	7.17	6.71	6.01	5.36	4.85	3.87	2.77	1.80	1.39

2.2 Health profile

There were 44,531 people with at least one disability living in a communal establishment in 2016. Almost 70% of these were aged 65 or older. Table 3 and Figure 5 show the percentages of self-stated health conditions grouped by Irish counties. Just under 28% of the population report having a chronic illness or health problem. The percentage of the adult population that are smokers has declined steadily in recent years and now stands at 17%. Approximately 23% of the population in Ireland is obese (Kennelly et al., 2020). Life expectancy at birth was 82.2 years for the whole population in 2017, 84.0 years for females, and 80.4 years for males (OECD, 2017).

Table 3: Self-stated health status in percent of the total population- Data Source: Central Statistics Office <https://www.cso.ie/en/census/>

County	Very Good (VG)	Good (G)	Fair (F)	Bad (B)	Very Bad (VB)
Carlow	58.05	28.33	8.76	1.39	0.25
Cavan	60.60	26.84	8.51	1.25	0.28
Clare	58.39	28.95	8.25	1.31	0.32
Cork	60.97	27.15	7.62	1.20	0.28
Donegal	56.68	28.88	9.84	1.47	0.34
Dublin	59.58	26.56	7.41	1.33	0.30
Galway	58.67	28.60	8.22	1.21	0.26
Kerry	56.46	29.83	8.55	1.26	0.26
Kildare	63.10	26.40	6.77	1.14	0.25
Kilkenny	60.95	27.52	7.93	1.24	0.27
Laois	59.19	27.64	8.03	1.35	0.23
Leitrim	56.91	28.90	9.63	1.37	0.26
Limerick	56.92	29.18	8.72	1.54	0.33
Longford	55.91	29.36	9.75	1.70	0.35
Louth	58.21	28.15	8.71	1.51	0.34
Mayo	56.04	30.20	9.88	1.49	0.28
Meath	63.15	26.48	6.68	1.04	0.20
Monaghan	59.60	28.01	8.57	1.13	0.23
Offaly	57.26	29.17	8.95	1.50	0.34
Roscommon	57.36	29.25	9.49	1.53	0.32
Sligo	57.70	28.46	9.19	1.48	0.26
Tipperary	57.62	29.15	9.13	1.47	0.30
Waterford	58.86	28.15	8.28	1.30	0.31
Westmeath	58.57	28.70	8.23	1.35	0.28
Wexford	62.48	26.20	7.44	1.15	0.27
Wicklow	59.04	28.00	9.00	1.44	0.31

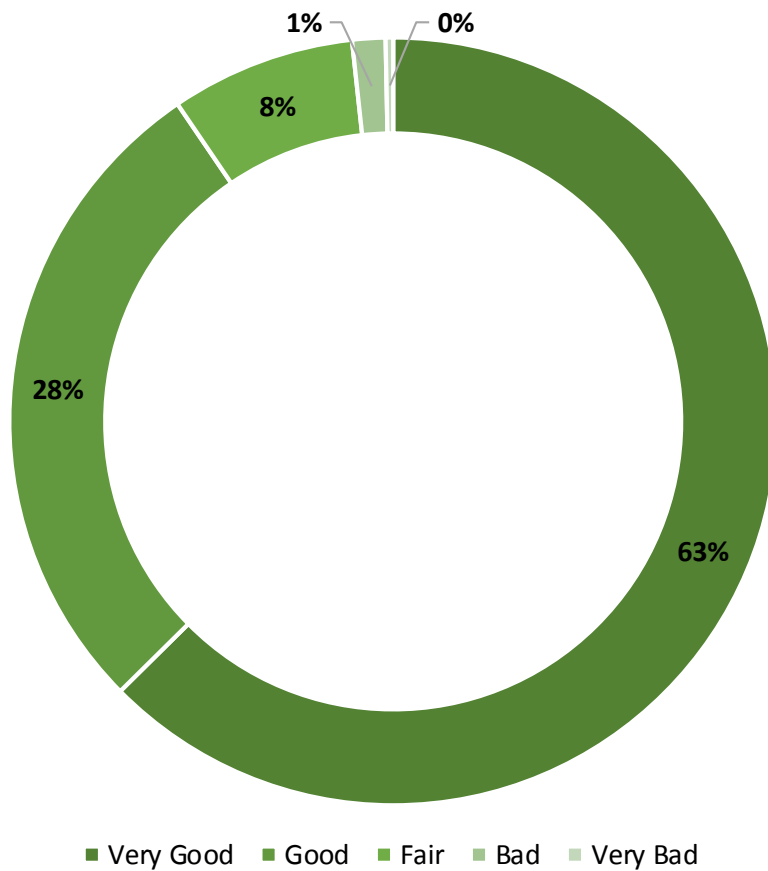


Figure 5: Pie chart of the self-stated health condition of the Irish population

2.3 Weather information

Met Éireann (www.met.ie), Ireland’s National Meteorological Service provides hourly data of 20 fully automatic weather stations across the country recording meteorological elements (i.e. Precipitation Amount, Air Temperature, Wet Bulb Air Temperature Dew Point Air Temperature, Vapor Pressure, Relative Humidity, Mean Sea Level Pressure, Mean Hourly Wind Speed, Predominant Hourly Wind Direction, Sunshine Duration, Visibility, Cloud Ceiling Height, Cloud Amount) on a minute-by-minute basis. There are also five manned weather stations in operation. Figure 6 shows the location of all 25 weather monitoring stations in Ireland. The averages of weekly monitored weather parameters for County Dublin are shown in Figure 7 and Figure 8.

Legend

- + Weather Station
- ★ Air Monitoring Site
- City or Town

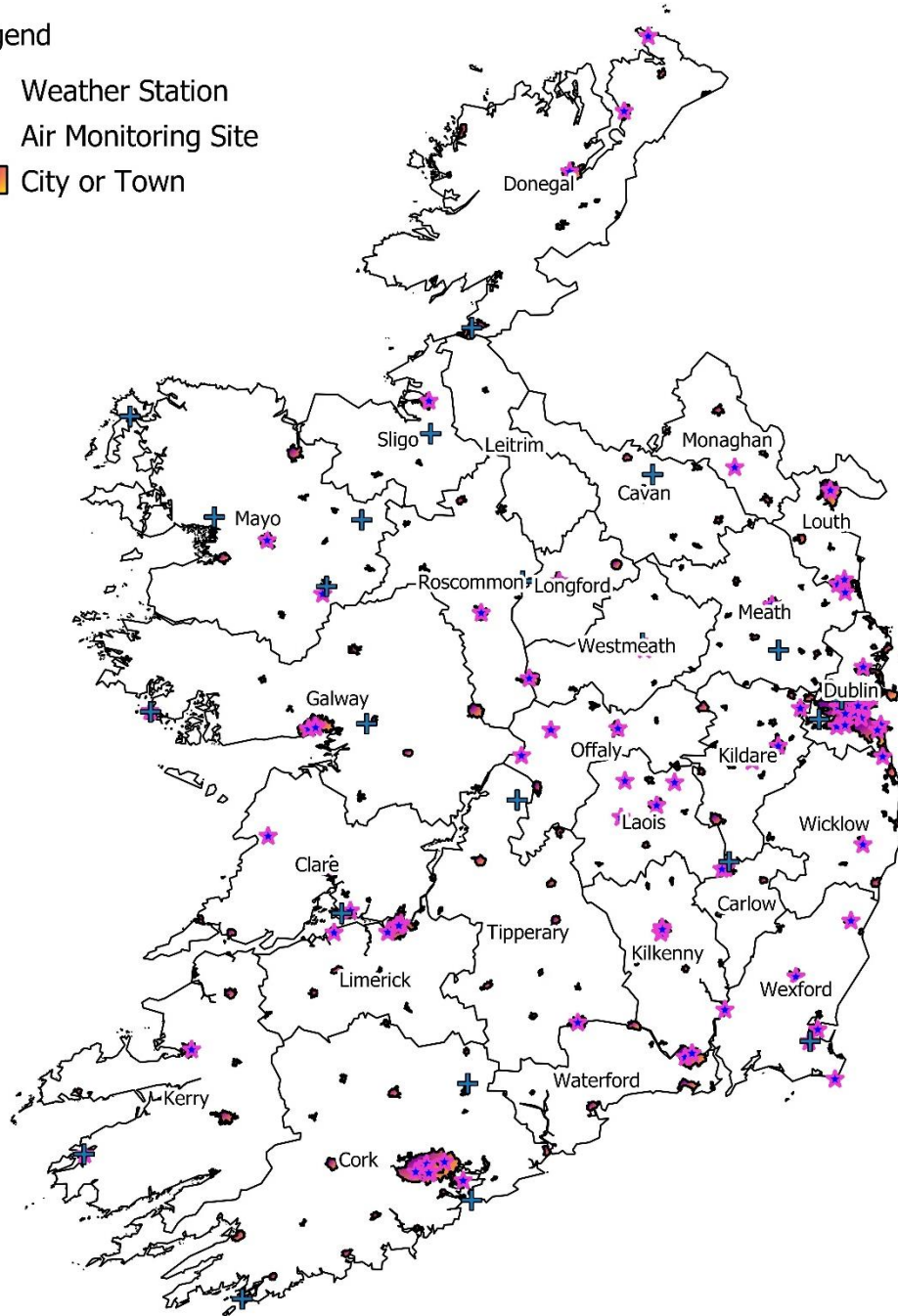


Figure 6: Weather and air pollution monitoring stations of Ireland

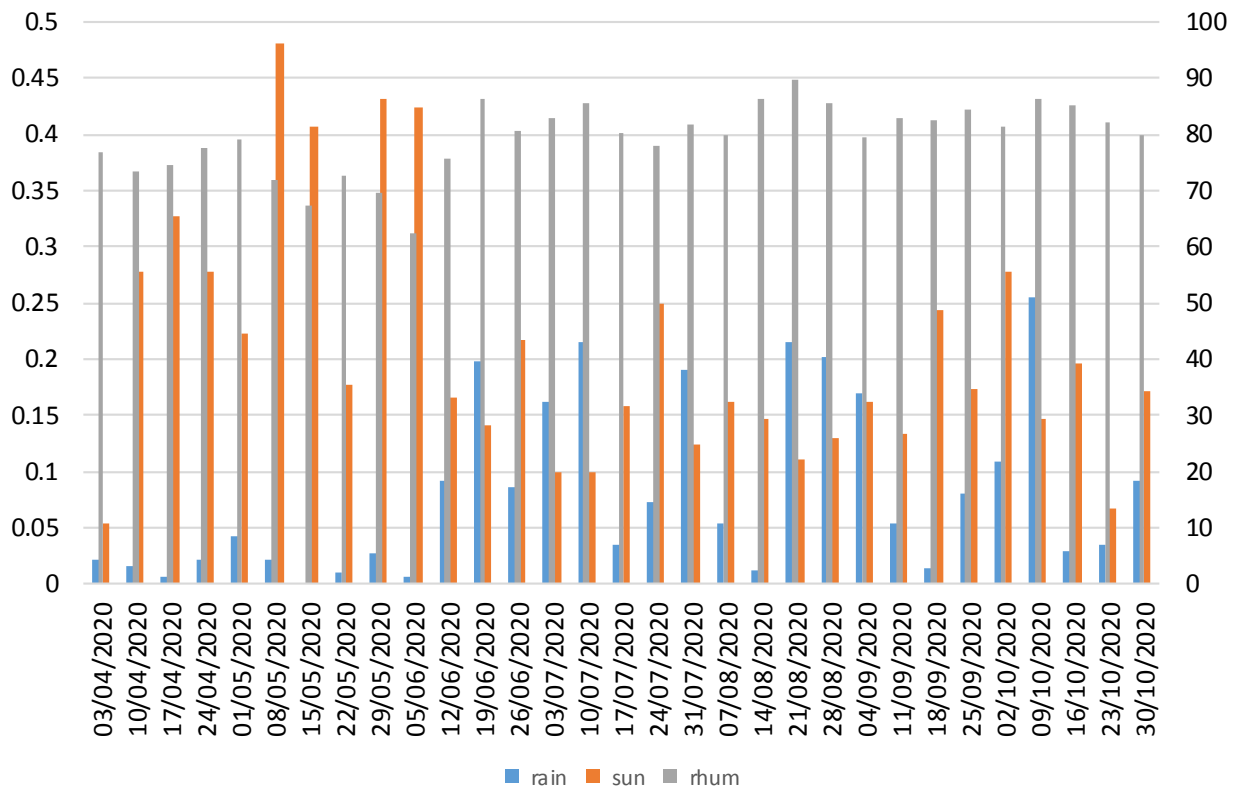


Figure 7: Weekly average of precipitation amount in mm, sunshine duration in hour (left y-axis), and Relative Humidity (rhum) in % (right y-axis) for County Dublin. Based on data driven from Met Éireann.

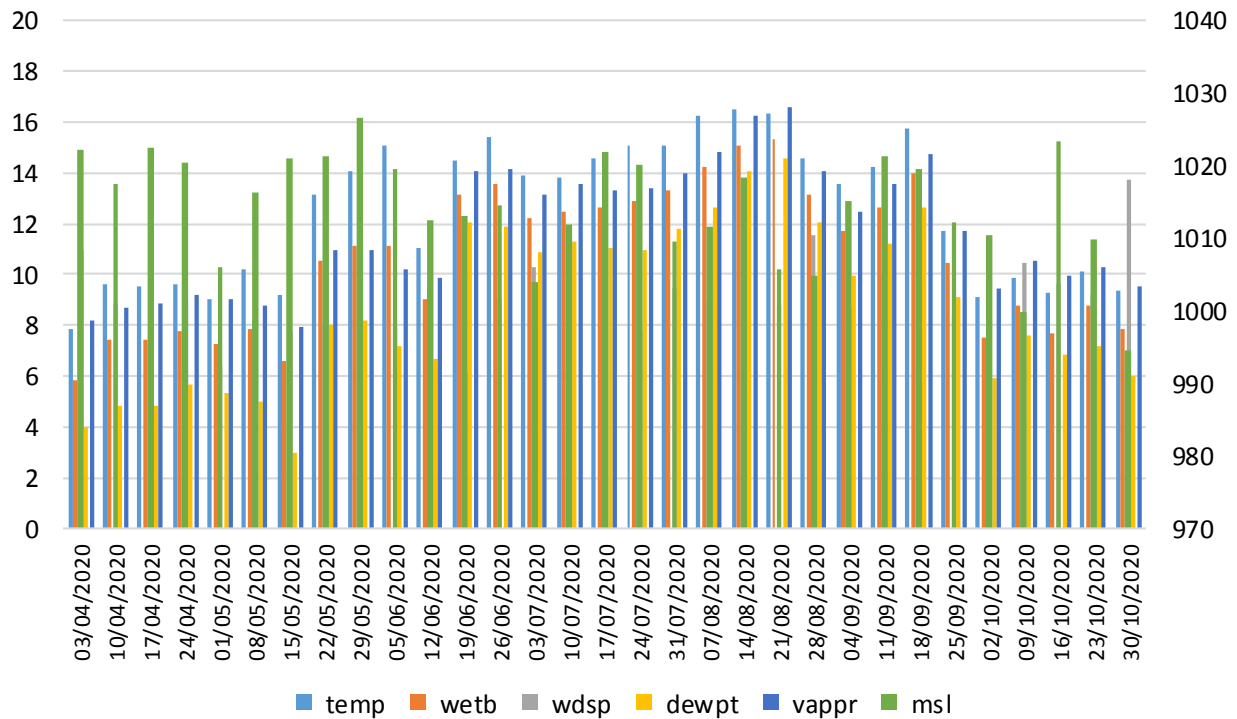


Figure 8: Weekly average of air temperature (temp), wet bulb air temperature(wetb), dew point air temperature(dewpt) in °C, wind speed (wdsp) in kt, Vapour Pressure in hpa (left y-axis), and mean Sea Level Pressure in hpa (right y-axis) for County Dublin. Based on data driven from Met Éireann.

2.4 Air pollution monitoring

The Environmental Protection Agency of Ireland (EPA) manages the national ambient air quality monitoring network (<https://airquality.ie/>). The hourly levels of several atmospheric pollutants in $\mu\text{g m}^{-3}$ are measured by 84 monitoring stations distributed over the state (Figure 6). The pollutants of most concern are Particulate matter (PM10 and PM2.5), Nitrogen dioxide (NO2), Ozone (O3), Sulphur dioxide (SO2), and Carbon monoxide (CO). Figure 6 shows the averages of the weekly concentration of atmospheric pollutants recorded by monitoring stations of County Dublin.

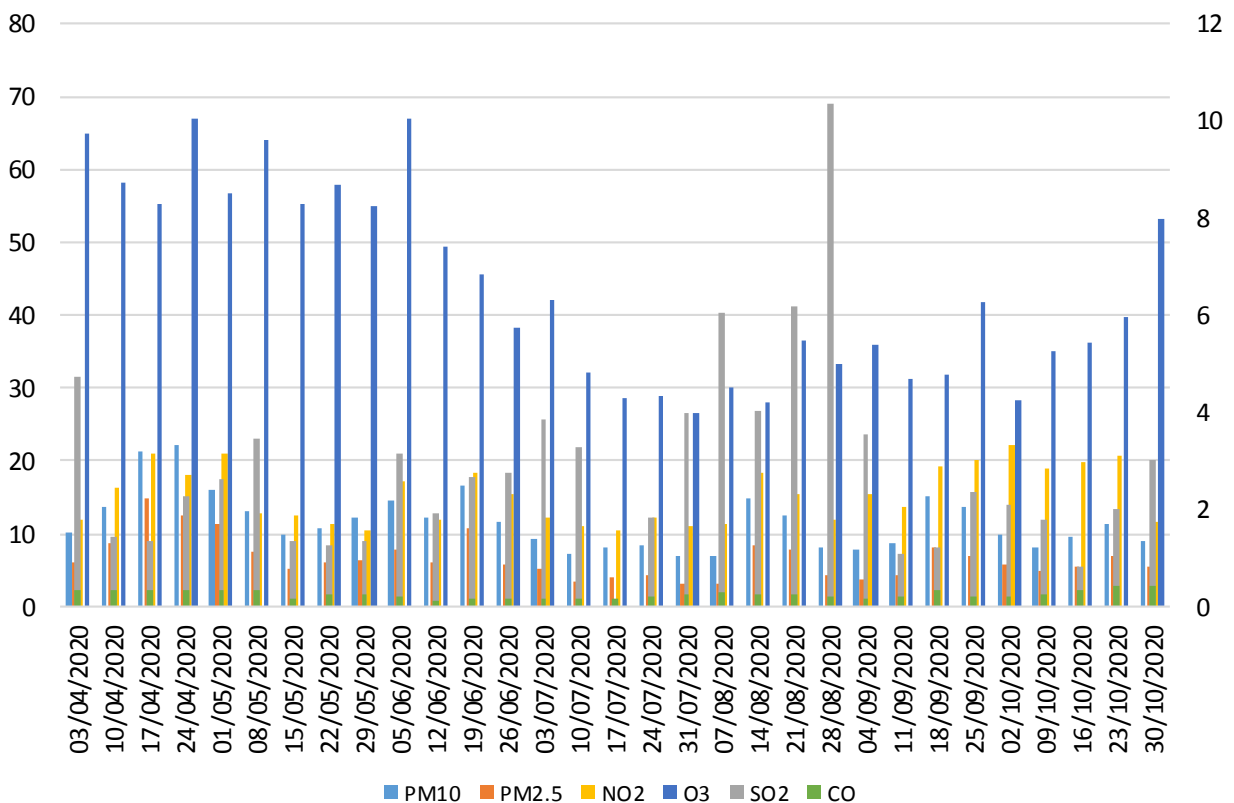


Figure 9: Average of atmospheric pollutants concentration ($\mu\text{g m}^{-3}$) for monitoring stations of County Dublin- the left vertical axis shows the values of PM10, PM2.5, NO3, and O3. The right axis expresses the SO2 and CO contents. Based on data driven from <https://airquality.ie>

2.5 Indoor Radon risk

Radon is a radioactive natural gas resulting from the decay of radium and enters our homes from foundation soil/rocks and building materials, all in a natural way. "Inhaled in excess quantities

and for prolonged periods, it can cause serious damage to the health system, in particular to the lungs. It is a threat to the lungs that existed long before the arrival of Covid-19. There are no scientific reports or technical evaluations between coronavirus and radon yet; however as both affect the respiratory system, the occurrence of enhanced levels of radon will expose extra pressure on the lungs. For example, in the Lombardy province of Italy, one of the regions most affected by Covid-19, there are approximately 200,000 homes at risk (4.1% of the total) (Zoran et al., 2020). Figure 10 shows the indoor radon map of Ireland. Results of previous studies indicate that those who lived in an area in which 10%-20% of households were above the national reference level (200 Bq/m³) were 2.9-3.1 times more likely to report a lung cancer diagnosis relative to those who lived in areas in which less than 1% of households were above the national reference level (Dempsey et al., 2018).

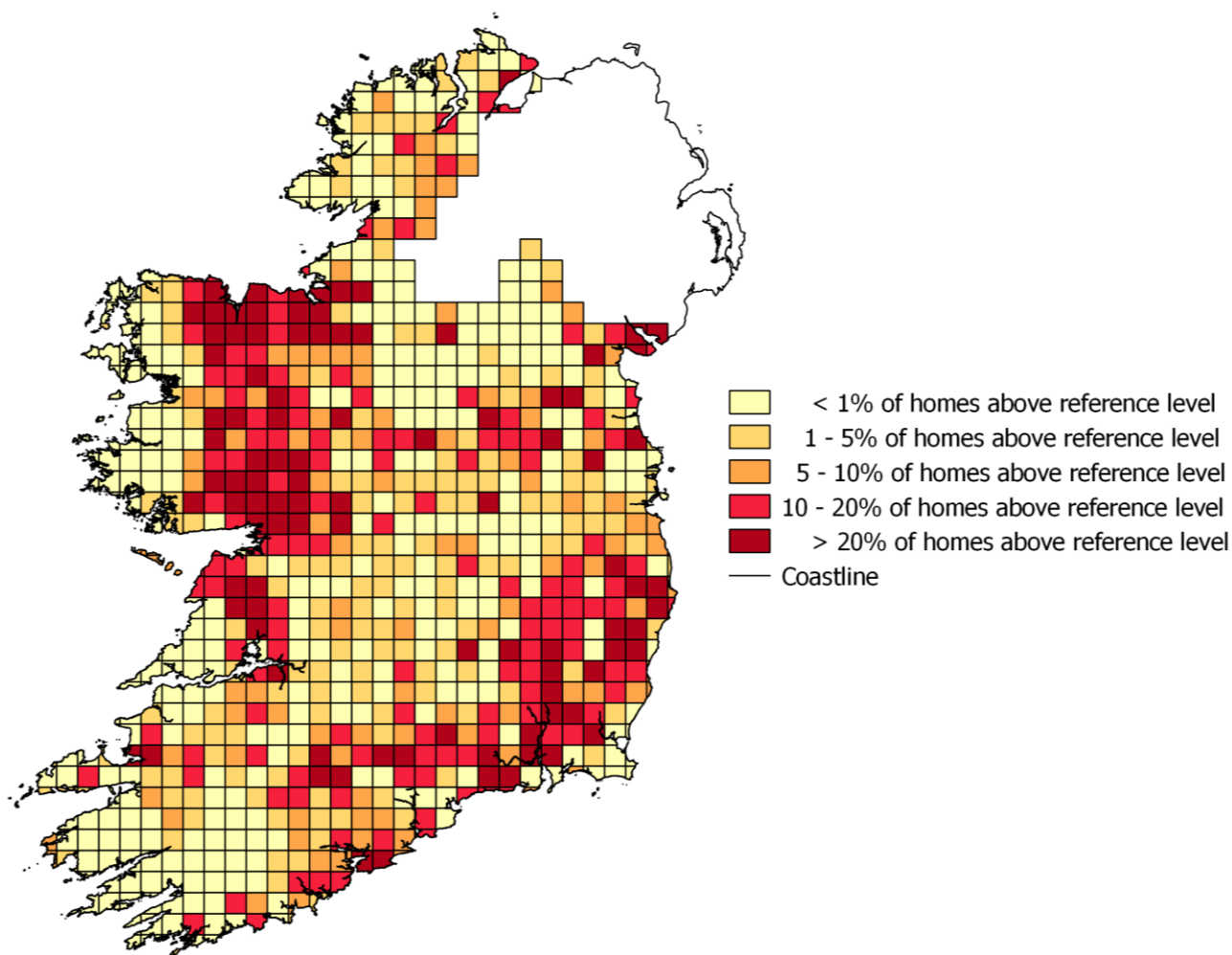


Figure 10: Indoor radon map of Ireland-Data source EPA- www.epa.ie

2.6 Covid-19 Lockdown in Ireland

On 12 March, the Irish government closed all schools, colleges, childcare facilities, and cultural institutions, and advised canceling large gatherings. On 24 March, almost all businesses, venues, facilities, and amenities were shut; but gatherings of up to four were allowed. Three days later on 27 March, the government imposed a stay-at-home order, banning all non-essential travel and contact with people outside one's home (including family and partners. The elderly and those with certain health conditions were told to cocoon. People were made to keep apart in public. A roadmap to easing restrictions in Ireland that includes five stages (Figure 11) was adopted by the government to be implemented by the beginning of 18th May 2020.

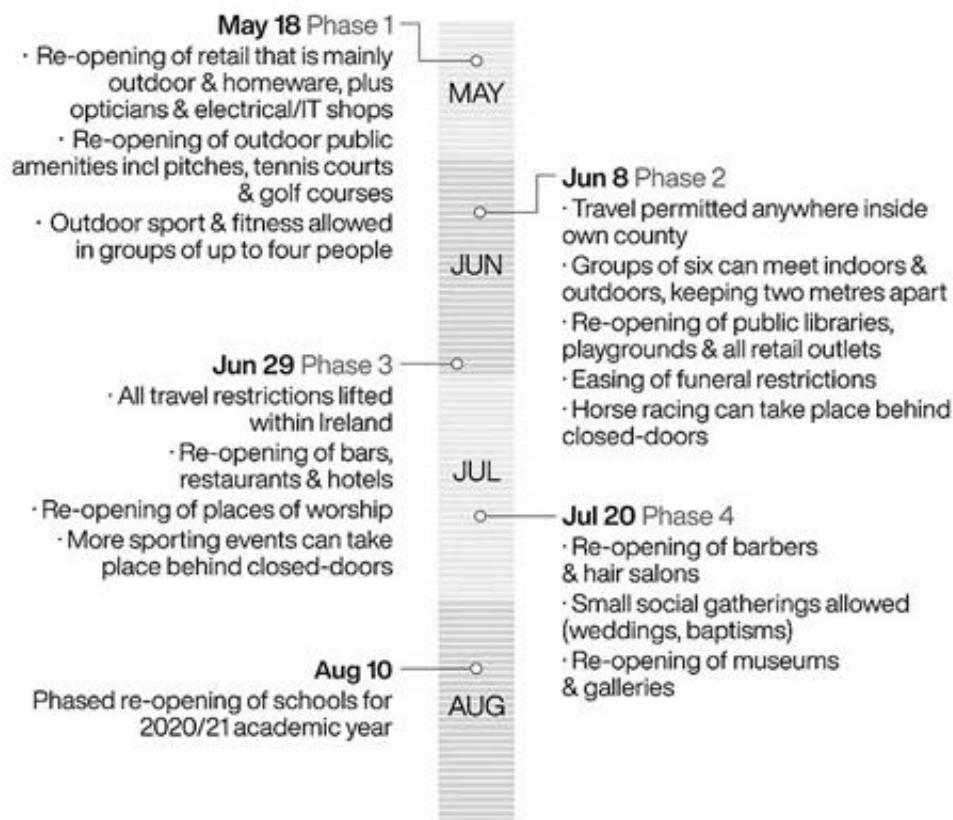


Figure 11: Covid-19 Lockdown phases in Ireland. Figure courtesy Aine McMahon-PA Graphics.

On 18 August, the Government of Ireland announced six new measures because of the growing number of confirmed cases, which remained in place until 15 September. On 15 September, the Government of Ireland announced a medium-term plan for living with COVID-19 that includes five levels of restrictions, with the entire country at Level 2 and specific restrictions in Dublin including the postponement of the reopening of pubs not serving food.

3 Methodology

Currently, disease forecasting technologies especially those related to COVID-19 are in high demand. Different types of regression have been developed and long used to evaluate the short-term associations of air pollution and weather with mortality or morbidity of non-infectious diseases (Imai et al., 2015). The application of regression approaches to infectious diseases like Covid-19 is somehow a new experiment and therefore is less well explored. In this study, we evaluate the spatial association between socio-demographic composition, atmospheric pollution, weather data, and COVID-19 deaths and cases by developing a comprehensive principal components regression (PCR) model.

PCR which was introduced and clearly explained by Jolliffe & Cadima, 2016 is a regression method that includes three main operations. The first is to run a PCA (Principal Components Analysis) on the table of the explanatory variables. Then running an Ordinary Least Squares (OLS) regression also called linear regression on the selected components and finally computing the parameters of the model that correspond to the input variables. PCA and PCR are among the highly recommended techniques for improving the regression models (Artigue & Smith, 2019). The regression analysis (PCR) of these PCs as independent variables will yield an appropriate estimation of the response parameter. As PCR is built on PCA, a great advantage of PCR regression over classical regression is the available charts that describe the data structure. PCR has many other advantages, it solves the dimensionality among the datasets and it is robust against collinearity between predictor variables (Dormann et al., 2013).

We used the PCR method using the first principal components that have a cumulative proportion of total variance of at least 95% to simulate daily new cases and confirmed deaths of Covid-19 as the dependent variables of the model. The equation used is as follows:

$$\text{Equation 1: } Y_{1,2} = a + b_1 Z_1 + b_2 Z_2 \dots + b_n Z_n$$

Y_1 = Covid-19 weekly new cases estimation and Y_2 = Covid-19 weekly deaths estimation

Z = Principal components

a = constant

b = regression coefficient Z with respect to Y

The main purpose of applying PCA in this study was to identify the socio-demographic composition, atmospheric pollution, and weather data (predictors) that are significantly correlated with the contagion and lethality rates of Covid-19. In particular, PCA accounts for interactions between all predictors and attributes these interactions to PCs. The PCs that explained most of the variance were included in the regression model (PCR).

3.1 Model Setting

To set the PCR model the first step is to define the response variables/s and specify the explanatory parameters. The weekly sum of COVID-19 cases and deaths (Response variables) were calculated based on the data retrieved from 03 April 2020 to 30 October 2020. 41 input parameters including socio-demographic composition, atmospheric pollution, and weather data were defined and weekly averages were calculated for the concerning time interval. To make the statistical analysis results from this data more valid, the weekly sums were log-normally transformed. Table 4 shows the input parameters and a summary of their statistics. The socio-demographic data for Ireland was extrapolated from the report of the last census (2016). Hourly data on atmospheric pollution as well as indoor radon activities were extrapolated from EPA. Met Éireann provided the hourly data of weather data. The prepared data was introduced to XLSTAT software (Addinsoft, 2019). The software used for statistical analysis, Principal Component Analysis, calculation of the correlation matrix, estimation of Cook's distance, determination of Eigenvalues, and interpretation of the results (See Appendix I).

3.2 Model validity

Standard deviation (σ) and standard error (SE), as well as the size of the root, mean square error (RMSE), mean square error (MSE), Mean absolute error (MAE), and the cross-validation between the observed and the predicted values of new cases and deaths of Covid-19 were used to evaluate the goodness of the PCR model. The SD, SE, RMSE, and MSE are defined as follows:

$$\text{Equation 2: } MAE = \frac{1}{n} \sum_{i=1}^n |Y_i - \hat{Y}_i|, \text{ Equation 3: } \sigma = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (Y_i - \bar{Y})^2} \text{ and Equation 4: } SE = \frac{\sigma}{\sqrt{n}}$$

$$\text{Equation 5: } MSE = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \text{ and Equation 6: } RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}$$

Where, n = number of data points, Y_i = observed values, \bar{Y} = mean and \hat{Y}_i = predicted values.

Table 4: Summary of the statistics and normality test of response and predictor variables

Variable	Minimum	Maximum	Mean	Std. deviation	Shapiro-Wilk p-value
Log Weekly New Cases	0.00	3.24	1.35	0.72	< 0.0001
Log Weekly Confirmed Deaths	0.00	2.11	0.12	0.34	< 0.0001
PM10 ($\mu\text{g m}^{-3}$)	2.62	41.88	11.41	4.83	0.00037
PM2.5 ($\mu\text{g m}^{-3}$)	1.27	23.88	6.59	3.75	0.00033
SO ₂ ($\mu\text{g m}^{-3}$)	0.00	17.38	4.67	4.08	< 0.0001
NO ₂ ($\mu\text{g m}^{-3}$)	0.73	79.66	33.01	20.58	0.00012
O ₃ ($\mu\text{g m}^{-3}$)	0.26	99.32	25.74	29.18	< 0.0001
CO ($\mu\text{g m}^{-3}$)	0.01	27.56	10.16	7.40	< 0.0001
Rain (mm)	0.00	0.59	0.13	0.10	< 0.0001
Temp (C)	7.25	17.65	13.12	2.79	0.00063
Sun (hour)	0.05	0.48	0.23	0.09	0.00029
wetb (C)	2.59	16.48	11.51	3.01	0.01581
Wdsp (Kt)	3.18	17.76	8.45	2.06	0.15573
dewpt (C)	2.99	15.76	9.98	3.48	0.15016
Vappr (hpa)	7.75	17.97	12.69	2.82	0.05541
rhum (%)	60.20	92.42	81.97	5.89	0.00267
msl (hpa)	989.67	1028.30	1014.75	7.08	0.12528
AGE<20T* (%)	25.14	31.75	28.44	1.55	< 0.0001
AGE20_24T (%)	4.04	6.83	5.16	0.69	< 0.0001
AGE25_29T (%)	4.47	8.28	5.44	0.72	< 0.0001
AGE30_34T (%)	6.17	9.16	6.96	0.64	< 0.0001
AGE35_39T (%)	6.79	8.95	7.80	0.50	< 0.0001
AGE40_44T (%)	6.89	8.53	7.45	0.38	< 0.0001
AGE45_49T (%)	6.42	7.46	6.98	0.23	< 0.0001
AGE50_54T (%)	5.82	6.88	6.49	0.28	< 0.0001
AGE55_59T (%)	5.19	6.72	5.91	0.43	< 0.0001
AGE60_64T (%)	4.41	6.25	5.30	0.53	< 0.0001
AGE65_69T (%)	3.75	5.85	4.70	0.57	< 0.0001
AGE70_74T (%)	2.64	4.43	3.59	0.45	< 0.0001
AGE75_79T (%)	1.61	3.10	2.51	0.36	< 0.0001
AGE80_84T (%)	1.02	2.28	1.77	0.29	< 0.0001
AGEGE_85T (%)	0.88	2.14	1.50	0.30	< 0.0001
HEALTH STATUS_VGT (%)	55.91	63.15	58.78	2.02	< 0.0001
HEALTH STATUS_GT (%)	26.20	30.20	28.23	1.09	< 0.0001
HEALTH STATUS_FT (%)	6.68	9.88	8.52	0.86	< 0.0001
HEALTH STATUS_BT (%)	1.04	1.70	1.35	0.16	< 0.0001
HEALTH STATUS_VBT (%)	0.20	0.35	0.29	0.04	< 0.0001
T Female (%)	49.45	51.14	50.25	0.39	< 0.0001
T Male (%)	48.86	50.55	49.75	0.39	< 0.0001
Population Density per m ²	0.0011	0.0071	0.0021	0.0012	< 0.0001
Rn_mean Bq m ⁻³	47.03	167.16	102.33	34.79	< 0.0001
Lock down phase	0**	5	-	-	< 0.0001
Population	32044	1347359	183149	252957	< 0.0001

*Total, **Shutdown level- Bold values express normal distribution.

4 Results

4.1 Principal Component Analysis (PCA)

Table 5 shows the eigenvalues calculated based on the PCA. In the PCA, the eigenvectors (principal components- we refer to Appendix I for more details) determine the directions of the new feature space, and the eigenvalues determine their magnitude which reflects the quality of the projection from the N-dimensional initial table (N=39 in this example) to a lower number of dimensions. According to this table, we can see that the first eigenvalue equals 12.66 and represents 30.87% of the total variability. This means that if we represent the data on only one axis, we will still be able to see the percent of the total variability of the data.

Table 5: Eigenvalue of the Principal Component Analysis (only 11 factors of 39 are shown here) - For more details, we refer to Appendix I

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11
Eigenvalue	12.66	6.70	4.51	3.27	2.17	2.03	1.92	1.27	1.11	0.87	0.73
Variability (%)	30.87	16.34	11.00	7.98	5.30	4.96	4.69	3.09	2.71	2.13	1.79
Cumulative %	30.87	47.21	58.21	66.19	71.49	76.45	81.13	84.23	86.94	89.07	90.86

Each eigenvalue corresponds to a factor and each factor to one dimension. A factor is a linear combination of the initial variables, and all the factors are uncorrelated ($R^2=0$). The eigenvalues and the corresponding factors are sorted by descending order of how much of the initial variability they represent (converted to %) (Figure 12)

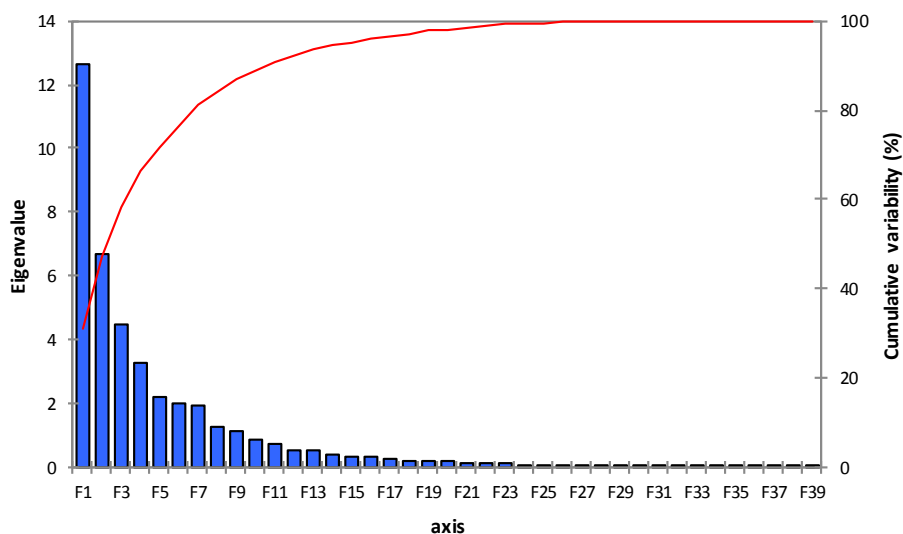


Figure 12: Scree plot showing the Eigenvalues and the Cumulative variability in PCA

Ideally, the first sets of eigenvalues will correspond to a high % of the variance, ensuring us that the results based on the first two or three factors are a good quality projection of the initial multi-dimensional table. In this example, the first eleven factors allow us to represent 90.86% of the initial variability of the data. This is a very good result, but we'll have to be careful when we interpret the factors as some information might be hidden in the next factors. We can see here that although we initially had 41 variables, the number of factors is 39. This is due to the elimination of two variables, which are negatively correlated ($R^2=-1$). The number of "useful" dimensions has been automatically detected by XLSTAT.

In Figure 13 the correlation circle (below on axes F1 and F2) shows a projection of the initial variables in the factors space. When two variables are far from the center, then, if they are: Close to each other, they are significantly positively correlated (R^2 close to 1); If they are orthogonal, they are not correlated (R^2 close to 0); If they are on the opposite side of the center, then they are significantly negatively correlated (R^2 close to -1).

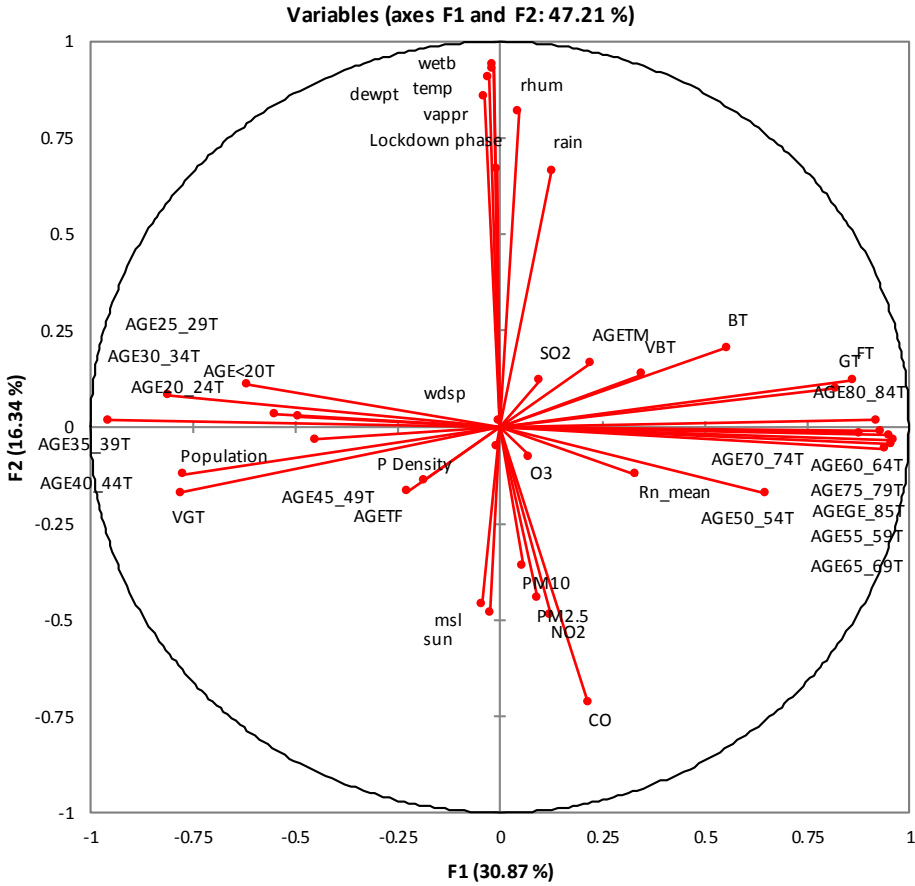


Figure 13: Correlation circle of the factors in PCA

When the variables are close to the center, some information is carried on other axes, and that any interpretation might be hazardous. For example, we might be tempted to interpret a correlation between the variables CO and Mean Radon although there is none. This can be confirmed either by looking at the correlation matrix or by looking at the correlation circle on axes F1 and F3. The correlation circle is useful in interpreting the meaning of the axes. In this example, the horizontal axis is linked with age categories and the vertical axis with weather data. These trends will help interpret the next parameter. To confirm that a variable is well-linked with an axis, take a look at the squared cosines table: the greater the squared cosine, the greater the link with the corresponding axis. The closer the squared cosine of a given variable is to zero, the more careful you have to be when interpreting the results in terms of trends on the corresponding axis. Looking at this table we can see that the trends for the health status would be best viewed on an F2/F3 combination.

4.2 Principle Components Regression (PCR)

Table 6 shows the goodness of fit coefficients of models for prediction of weekly new cases and confirmed deaths of Covid-19. The R^2 (coefficient of determination) indicates the percent of the variability of the dependent variable which is explained by the explanatory variables. R^2 values greater than 0.5 are good. The closer to 1 the R^2 is, the better the fit. In the case of this study, 53 and 54% of the variability of the contagion and lethality are explained by the predictors. The remainder of the variability is due to some effects (other explanatory variables like the degree of actual social distancing, blood groups, etc.) that have not been included in this analysis.

Table 6: Goodness of fit statistics

Variable	Log Weekly Confirmed Deaths	Log Weekly New Cases
<i>DF</i>	766	766
R^2	0.56	0.55
<i>Adjusted R²</i>	0.54	0.53
<i>MSE</i>	0.05	0.25
<i>RMSE</i>	0.23	0.50
<i>MAPE</i>	40.35	22.54
<i>DW</i>	1.17	1.41
<i>Cp</i>	40.00	40.00
<i>AIC</i>	-2307.49	-1089.73
<i>SBC</i>	-2119.81	-902.04
<i>PC</i>	0.49	0.49
<i>Press RMSE</i>	46.54	209.69

Figure 14 and Figure 15 show the correlations between predicted and observed values of Covid-19 weekly new cases and confirmed deaths. The 95 percent confidence level intervals are also shown in the charts of these figures. The values out of the space between confidence levels are potential outliers or might suggest that the normality assumption is wrong.

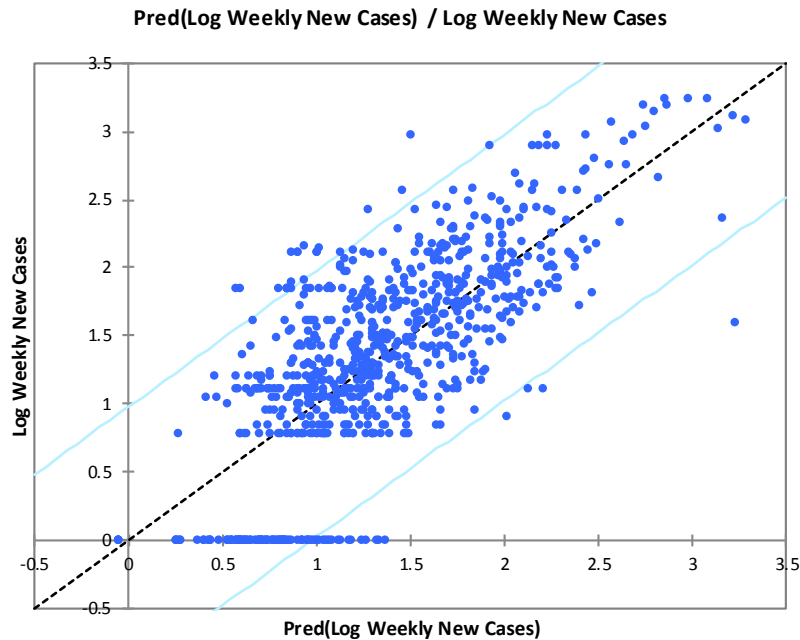


Figure 14: Correlation between predicted and observed values of Covid-19 weekly new cases

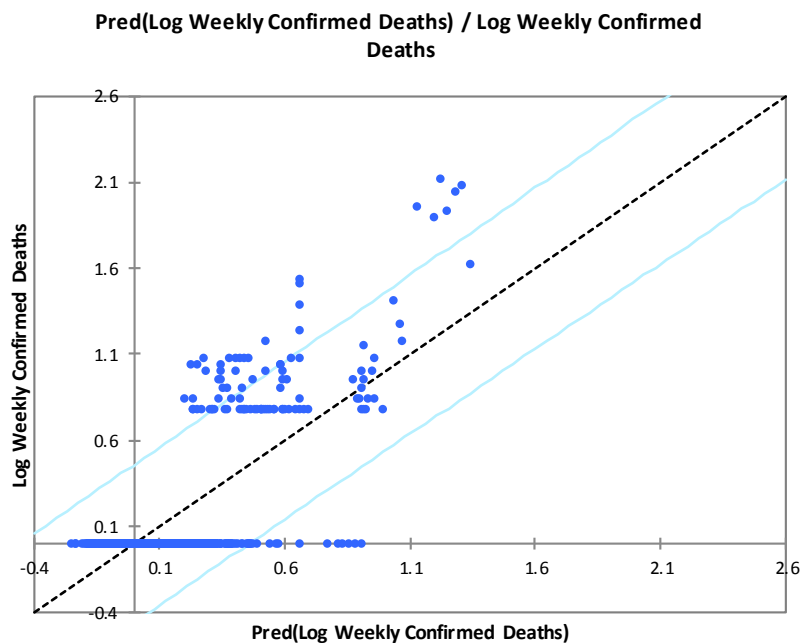


Figure 15 Correlation between predicted and observed values of Covid-19 weekly confirmed deaths

It is also important to examine the results of the analysis of variance (Table 7). The results enable us to determine whether or not the explanatory variables bring significant information (null hypothesis) to the model. In other words, it's a way of asking whether it is valid to use the mean to describe the whole population, or whether the information brought by the explanatory variables is of value or not. Fisher's F test is used. Given the fact that the probability corresponding to the F value is lower than 0.0001, it means that we would be taking a lower than 0.01% risk in assuming that the null hypothesis is wrong. Therefore, we can conclude with confidence that the forty-one variables do bring a significant amount of information to the models for the prediction of contagion and lethality rates.

Table 7: Results of variance analysis

Log Weekly Confirmed Deaths					
Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	39	53.103	1.362	25.026	< 0.0001
Error	766	41.676	0.054		
Corrected Total	805	94.780			
Log Weekly New Cases					
Model	39	232.975	5.974	24.234	< 0.0001
Error	766	188.823	0.247		
Corrected Total	805	421.798			

Table 8 gives vital details on the model parameters for the prediction of Covid-19 weekly new cases and deaths. This table is helpful for future predictions, or when it is needed to compare the coefficients of the model for given input data with the ones weekly observed (it could be used to compare the models for prediction of Covid-19 weekly new cases and deaths). As can be seen in this table, the variables with p-values lower than 0.05 bring significant information to the prediction model. According to this for the model for prediction of weekly new cases, PM10, PM2.5, O3, Sun duration, wind speed, dew point temperature, elder age category, percent of the population in fair and bad condition together with Lock downing are the most significant parameters. The p-value indicates also the degree of significance, the higher the p-value of a variable, the weaker effect in the model is anticipated. Based on what is mentioned above, we can say that the findings of our study confirm the positive association between virus transmissions with some of the atmospheric pollutants and metrological parameters. Also, the results are in agreement with the few studies that have been recently conducted (see section 1.4 Effect of atmospheric pollution and 1.5 Effect of metrological parameters

Table 8: Model parameters for prediction of Covid-19 weekly new cases and deaths

Source	Model parameters for the Log Weekly New Cases				Model parameters for the Log Weekly Confirmed Deaths			
	Value	St. error	t	Pr > t	Value	St. error	t	Pr > t
<i>Intercept</i>	-4.67E-01				3.81E+00			
<i>PM10</i>	2.23E-02	0.01	2.45	0.01*	1.69E-02	0.00	3.95	< 0.0001
<i>PM2.5</i>	3.11E-02	0.01	2.28	0.02	-9.76E-03	0.01	-1.52	0.13
<i>SO2</i>	1.30E-02	0.01	1.77	0.08	-4.34E-03	0.00	-1.26	0.21
<i>NO2</i>	1.76E-03	0.00	1.01	0.31	-3.06E-03	0.00	-3.72	0.00
<i>O3</i>	4.25E-03	0.00	2.30	0.02	-9.44E-04	0.00	-1.08	0.28
<i>CO</i>	-8.27E-03	0.01	-1.37	0.17	5.18E-03	0.00	1.83	0.07
<i>rain</i>	-8.23E-03	0.33	-0.03	0.98	2.83E-01	0.15	1.84	0.07
<i>temp</i>	5.31E-02	0.20	0.26	0.79	-2.14E-01	0.10	-2.25	0.02
<i>sun</i>	9.83E-01	0.33	3.01	0.00	9.10E-02	0.15	0.59	0.55
<i>wetb</i>	6.49E-03	0.07	0.09	0.93	4.34E-02	0.03	1.30	0.19
<i>wdsp</i>	5.62E-02	0.01	4.04	< 0.0001	-5.79E-04	0.01	-0.09	0.93
<i>dewpt</i>	-6.72E-01	0.23	-2.93	0.00	2.10E-02	0.11	0.19	0.85
<i>vappr</i>	6.13E-01	0.08	7.58	< 0.0001	1.62E-01	0.04	4.27	< 0.0001
<i>rhum</i>	3.22E-02	0.04	0.72	0.47	-3.73E-02	0.02	-1.78	0.08
<i>msl</i>	6.83E-03	0.00	1.74	0.08	4.54E-03	0.00	2.47	0.01
<i>AGE<20T</i>	-1.03E-01	0.09	-1.16	0.25	-7.90E-02	0.04	-1.90	0.06
<i>AGE20_24T</i>	1.57E-01	0.11	1.42	0.15	2.99E-02	0.05	0.58	0.56
<i>AGE25_29T</i>	-1.04E-01	0.19	-0.54	0.59	1.68E-01	0.09	1.87	0.06
<i>AGE30_34T</i>	5.40E-01	0.27	1.98	0.05	-1.79E-01	0.13	-1.40	0.16
<i>AGE35_39T</i>	-9.39E-01	0.66	-1.42	0.16	1.95E-01	0.31	0.63	0.53
<i>AGE40_44T</i>	7.07E-01	0.26	2.73	0.01	-8.32E-02	0.12	-0.68	0.49
<i>AGE45_49T</i>	1.13E+00	0.47	2.38	0.02	7.28E-01	0.22	3.28	0.00
<i>AGE50_54T</i>	-6.69E-01	0.47	-1.41	0.16	-2.22E-01	0.22	-1.00	0.32
<i>AGE55_59T</i>	-8.71E-01	0.63	-1.38	0.17	2.20E-01	0.30	0.74	0.46
<i>AGE60_64T</i>	-8.54E-01	0.84	-1.02	0.31	8.47E-01	0.39	2.16	0.03
<i>AGE65_69T</i>	1.46E+00	0.75	1.95	0.05	9.98E-02	0.35	0.28	0.78
<i>AGE70_74T</i>	-9.28E-01	0.89	-1.04	0.30	-9.23E-01	0.42	-2.20	0.03
<i>AGE75_79T</i>	6.93E-01	0.83	0.84	0.40	2.10E-01	0.39	0.54	0.59
<i>AGE80_84T</i>	1.34E+00	0.99	1.36	0.18	-9.73E-02	0.46	-0.21	0.83
<i>AGEGE_85T</i>	4.03E-01	0.72	0.56	0.58	-4.03E-01	0.34	-1.19	0.23
<i>VGT</i>	-4.01E-02	0.14	-0.28	0.78	-5.28E-02	0.07	-0.78	0.44
<i>GT</i>	-2.63E-01	0.20	-1.31	0.19	-2.65E-01	0.09	-2.81	0.01
<i>FT</i>	-5.24E-01	0.29	-1.78	0.07	1.93E-01	0.14	1.40	0.16
<i>BT</i>	2.48E+00	0.83	2.97	0.00	-8.23E-02	0.39	-0.21	0.83
<i>VBT</i>	-3.44E+00	2.04	-1.69	0.09	-1.65E+00	0.96	-1.72	0.09
<i>AGETF</i>	-7.66E-02	0.09	-0.82	0.41	3.65E-02	0.04	0.83	0.41
<i>AGETM</i>	7.66E-02	0.09	0.82	0.41	-3.65E-02	0.04	-0.83	0.41
<i>P Density</i>	-6.63E+01	52.62	-1.26	0.21	-1.08E+02	24.72	-4.36	< 0.0001
<i>Rn_mean</i>	-3.45E-03	0.00	-1.72	0.09	-3.47E-03	0.00	-3.68	0.00
<i>Lockdown phase</i>	1.69E-01	0.02	10.31	< 0.0001	-2.98E-02	0.01	-3.87	0.00

*Bold values express variables with statistical significance

For the model for prediction of weekly deaths, PM10, Temperature, vapor and mean sea level pressures, elder age category, percent of the population with a good health condition, background radon levels together with population density and Lock downing do bring useful information. We can see that the mean radon concentration has a positive effect on the increase of virus lethality. The equation of the model's parameter, which can be also used for future predictions (in the case of the inputs being provided), can be obtained using Equation 1 and the values mentioned in Table 8 In other words, the values of this table are the regression coefficient of Equation 1 (b1, b2... and b41) and the a coefficient is the Intercept value. Note that back-transformation of calculated logarithmic values to the original scale is also necessary.

To have a better understanding of the effect of explanatory parameters on Covid-19 contagion and Lethality rates, Figure 16 and Figure 17 show the standardized regression coefficients (sometimes referred to as beta coefficients). They allow to directly compare the relative influence of the explanatory variables on the dependent variable, and their significance. According to these figures, dew point temperature (negative correlation) and Vapor pressure (positive correlation) have the highest standardized significances in the model for the prediction of contagion. In the model for prediction of lethality rates, temperature (negative correlation), and Vapor pressure (positive correlation) showed the highest significance levels.

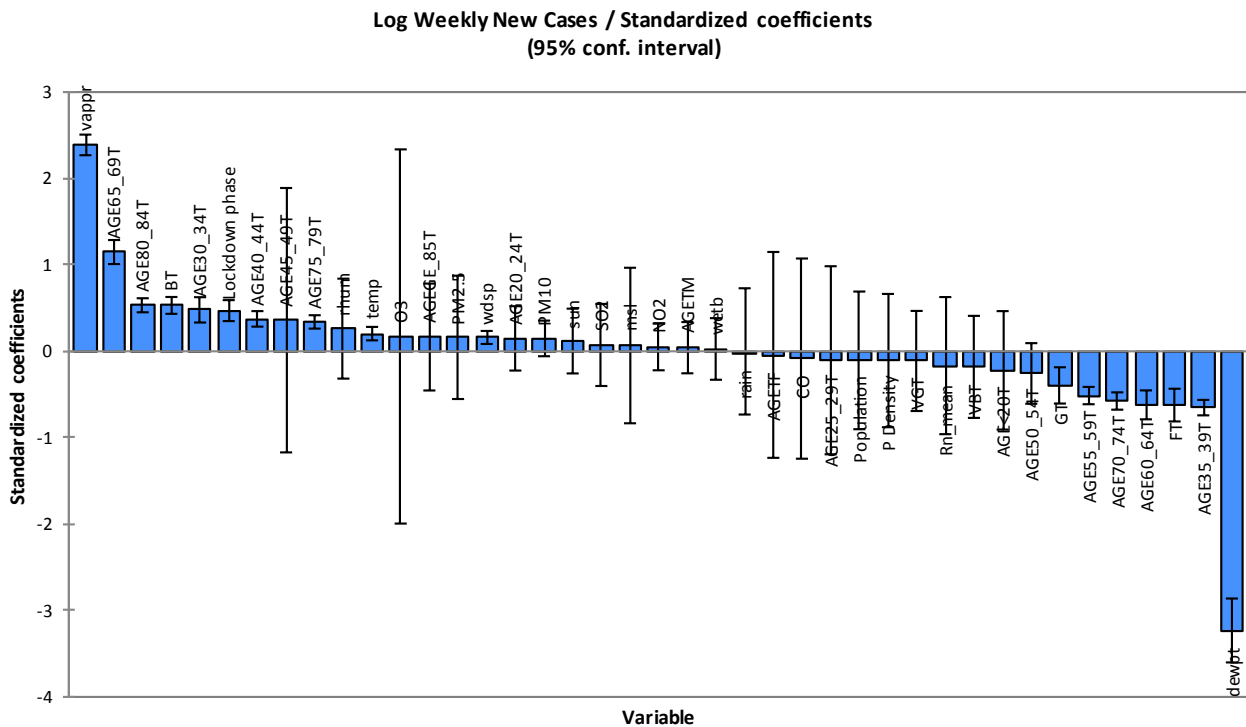


Figure 16: Chart of standardized coefficients of the model for prediction weekly new case of Covid-19

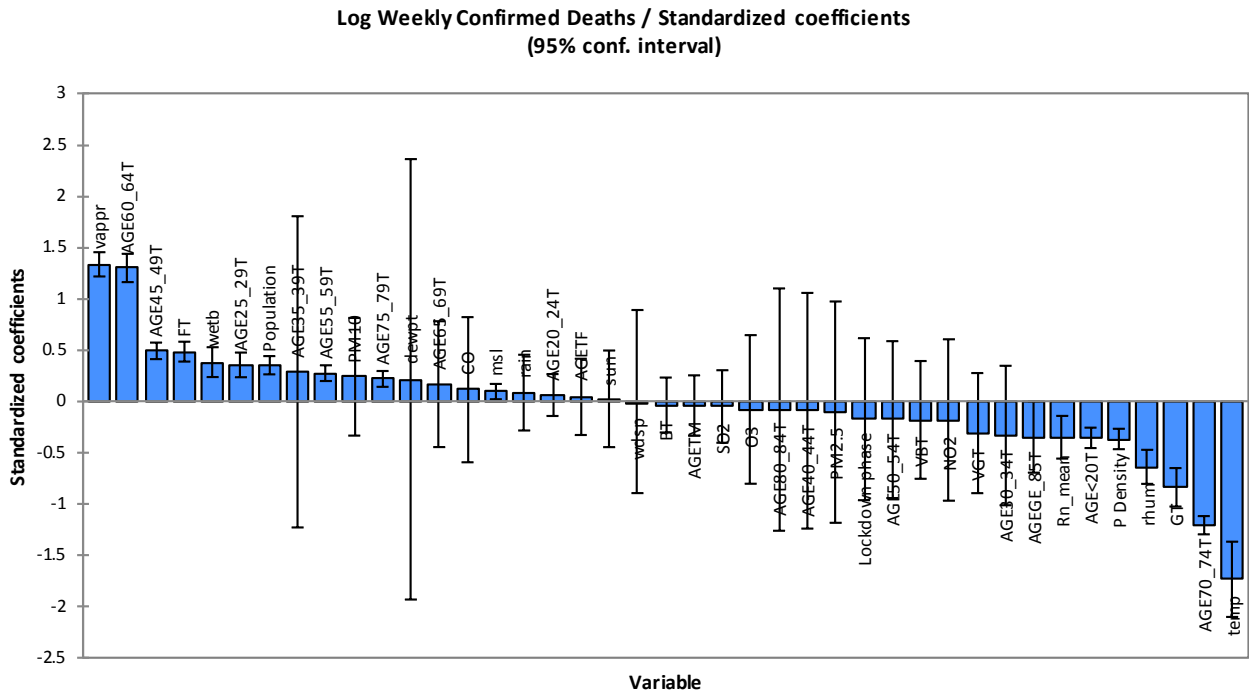


Figure 17: Chart of standardized coefficients of the model for prediction weekly deaths of Covid-19

4.3 Model validity

The results of cross-validation of observed and predicted models show that the predicted values match those measured (Table 9): the absolute mean prediction error weekly Covid-19 new cases (NC) and confirmed deaths (CD) ($MAE_{NC} = 0.38$ and $MAE_{CD} = 0.15$) and the mean standardized error ($MSE_{NC} = 0.0007$ and) are close to zero, indicating that the prediction method of both contagion and lethality rates are unbiased (centered on the true values) and the model is rather accurate; the average standard error ($SE_{CD} = 0.14$) is lesser than the root mean squared prediction error ($RMSE_{CD} = 0.23$), suggesting that the model slightly underestimates the lethality rates, however, $SE_{NC} (=0.77)$ is more than $RMSE_{NC} (=0.50)$ which indicates that the model overestimates the contagion rates.

The computed RMSEs and MSEs of Log transformed weekly Covid-19 new cases (NC) and confirmed deaths (CD) predicted for Irish counties are shown in Table 10. The range of the errors for the PCR-NC model was 0.13 (for County Kerry) to 0.52 (for County Dublin), while the errors for PCR-CD ranged from 0.35 (for county Dublin) to 0.58 (for County Waterford). Based on the obtained values one can judge that the error of the model for prediction of contagion rates is slightly higher than the lethality prediction model.

Table 9: Results of cross-validation of the observed and predicted values

Variable	Log Weekly Confirmed Deaths	Log Weekly New Cases
Mean absolute error (MAE)	0.15	0.38
Standard Error of Mean (SEM)	0.0003	0.0007
Mean square error (MSE)	0.05	0.25
Root mean square error (RMSE)	0.23	0.50
Average standard error (SE)	0.14	0.77

Table 10: Mean square errors (MSE) and Root mean square errors (RMSE) of Log transformed weekly Covid-19 new cases (NC) and confirmed deaths (CD)

County	RMSE-CD	MSE-CD	RMSE-NC	MSE-NC
Carlow	0.20	0.04	0.49	0.24
Cavan	0.24	0.06	0.46	0.21
Clare	0.22	0.05	0.43	0.18
Cork	0.18	0.03	0.44	0.19
Donegal	0.24	0.06	0.41	0.17
Dublin	0.52	0.27	0.35	0.12
Galway	0.15	0.02	0.42	0.17
Kerry	0.13	0.02	0.58	0.34
Kildare	0.37	0.14	0.37	0.14
Kilkenny	0.17	0.03	0.41	0.17
Laois	0.19	0.03	0.50	0.25
Leitrim	0.17	0.03	0.37	0.14
Limerick	0.16	0.03	0.53	0.28
Longford	0.21	0.05	0.50	0.25
Louth	0.27	0.07	0.55	0.30
Mayo	0.21	0.04	0.48	0.23
Meath	0.23	0.05	0.48	0.23
Monaghan	0.24	0.06	0.53	0.28
Offaly	0.19	0.04	0.55	0.31
Roscommon	0.15	0.02	0.46	0.21
Sligo	0.14	0.02	0.57	0.32
Tipperary	0.19	0.03	0.52	0.28
Waterford	0.18	0.03	0.58	0.34
Westmeath	0.13	0.02	0.50	0.25
Wexford	0.22	0.05	0.54	0.29
Wicklow	0.23	0.05	0.43	0.18

4.4 Comparison of observed and predicted weekly new cases and deaths of Covid-19 in Dublin County and the other Irish Counties

Figure 18 shows the total number of cases observed and predicted for the study period (3 April to 30 October 2020), according to this figure, for most of the Irish counties the number of predicted values is less than the observed cases. This explains that there are hidden factors that affect the contagion rates, one of them could be the very higher contagiousness of the SARS-CoV-2 virus in hospitals where health care workers were infected. The average difference between the predicted and observed values is about 27 percent. Also, the cumulative percent of contagion rate of health workers for the period of study was about 18%, so it can be concluded that extreme contagiousness of the virus in sanitary environments would account for most of the difference between predicted and observed cases. Another parameter that would explain the remainder of the difference is the actual degree of social distancing considered by the majority of the population.

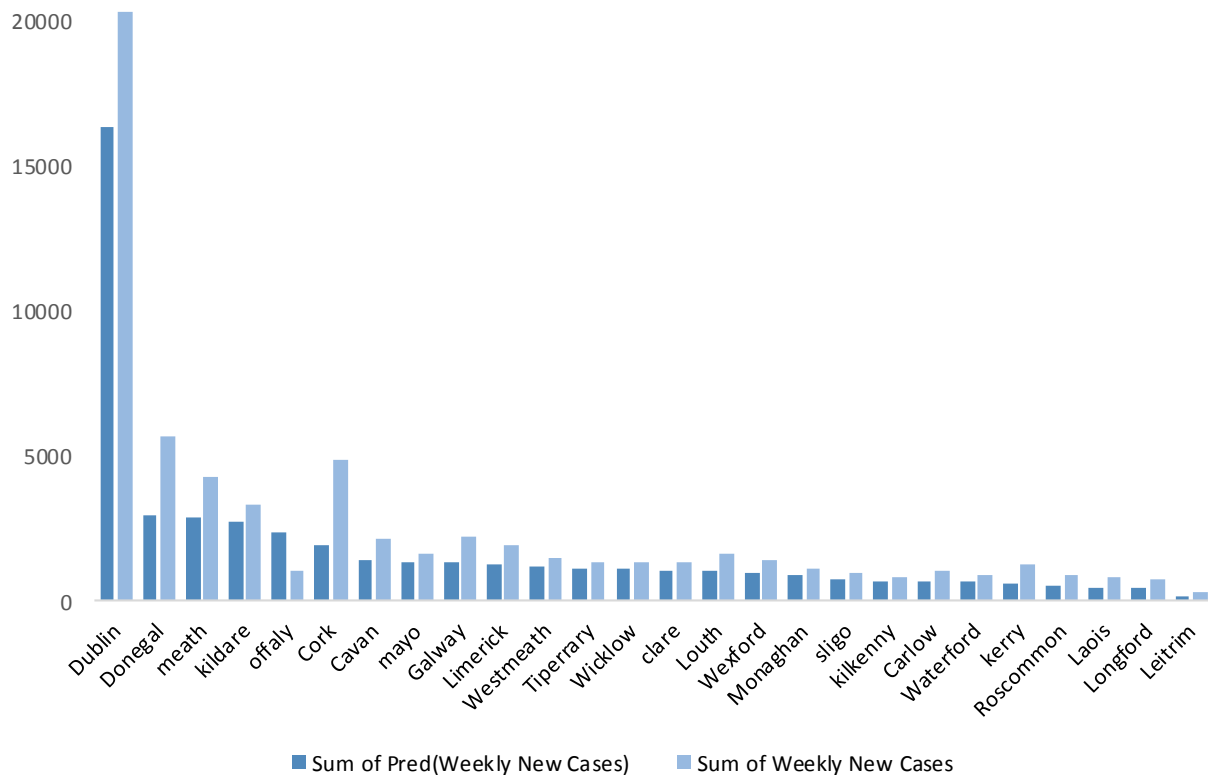


Figure 18: Total number of predicted and observed cases of Covid-19. (3 April- 30 October)

According to Figure 19, except for county Dublin, the total predicted and observed deaths due to Covid-19 from 03 April to 30 October 2020 are close to each other (predicted values are slightly lower). The extra death in County Dublin can be attributed to the residential and community care facilities including nursing homes for which significantly high death rates (62% total) were reported. Also, the cumulative death rate percent of health workers for the period of study was about 1%. As can be seen in Figure 19 the total predicted death for county Dublin is 37% of the observed ones. As most of the care facilities and hospitals are in County Dublin, it can be concluded that the additional deaths from health care facilities together with the death of health workers in hospitals can justify the high difference between the observed and predicted deaths (63% total).

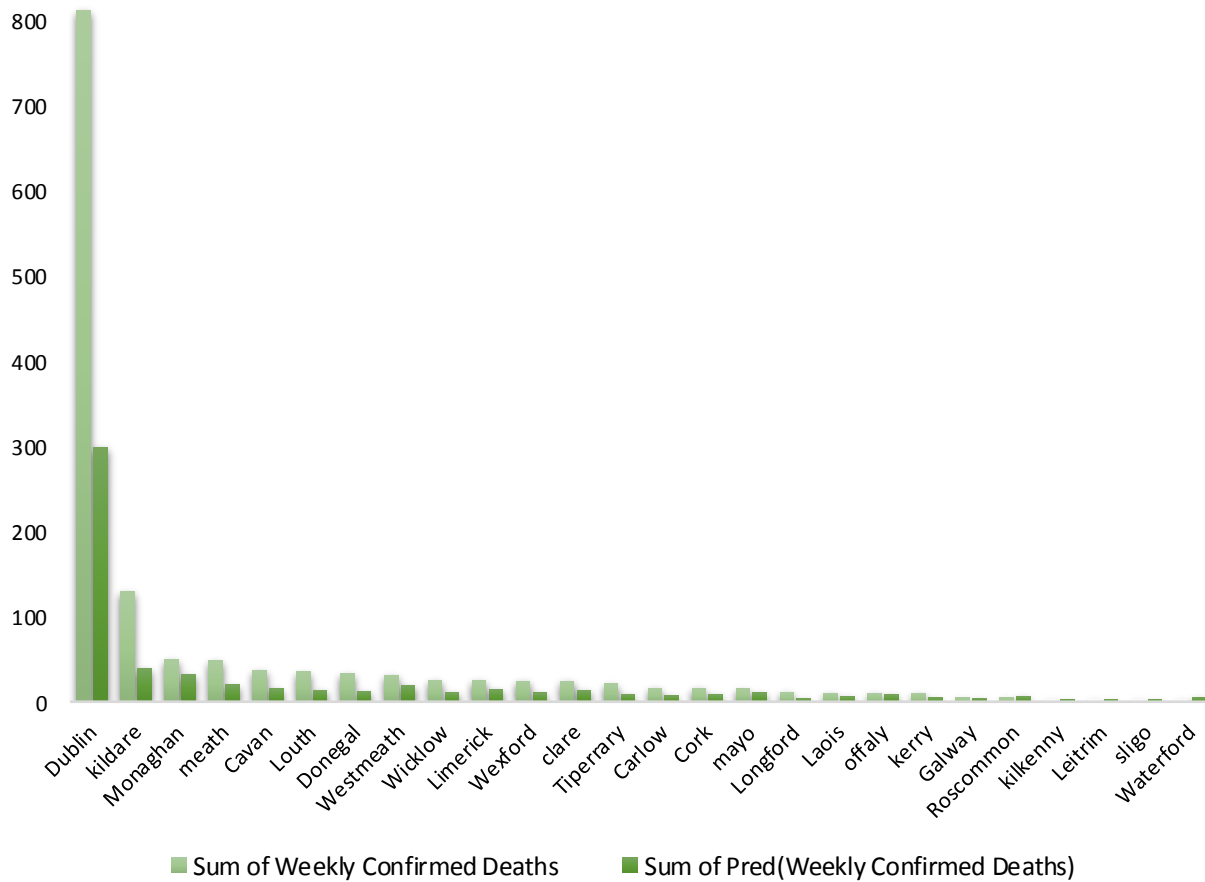


Figure 19: Total number of predicted and observed deaths of Covid-19. (3 April- 30 October)

Figure 20 shows the chart of predicted and observed Weekly cases of Covid-19 in County Dublin. There is a rather good agreement between predicted and observed cases.

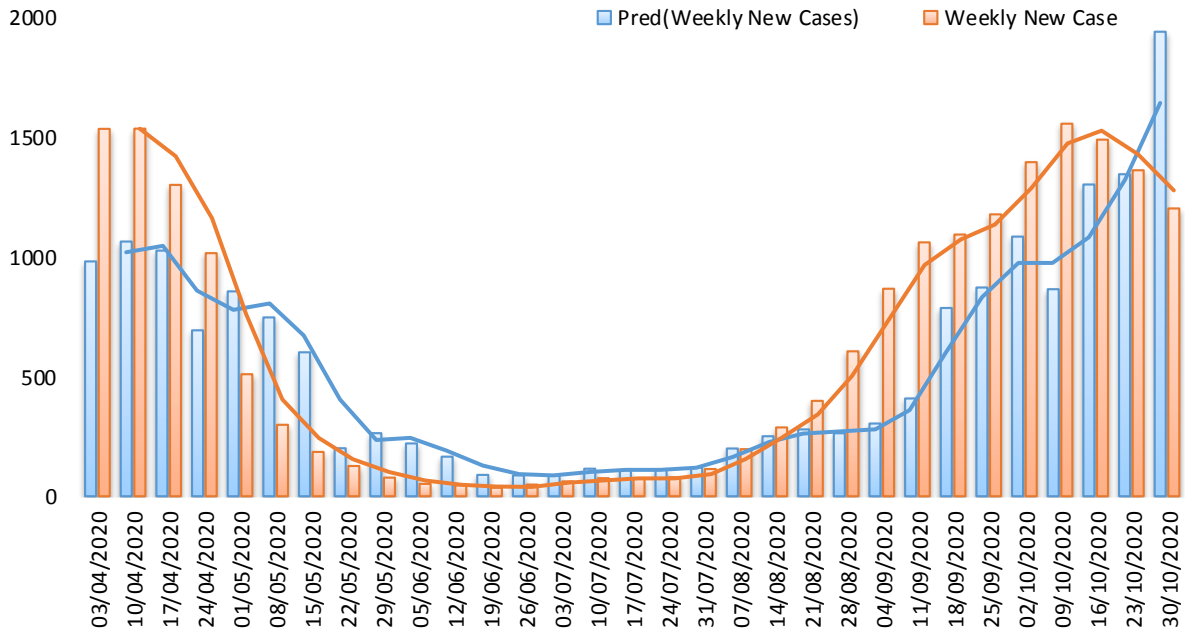


Figure 20: Predicted and Observed weekly new cases in County Dublin

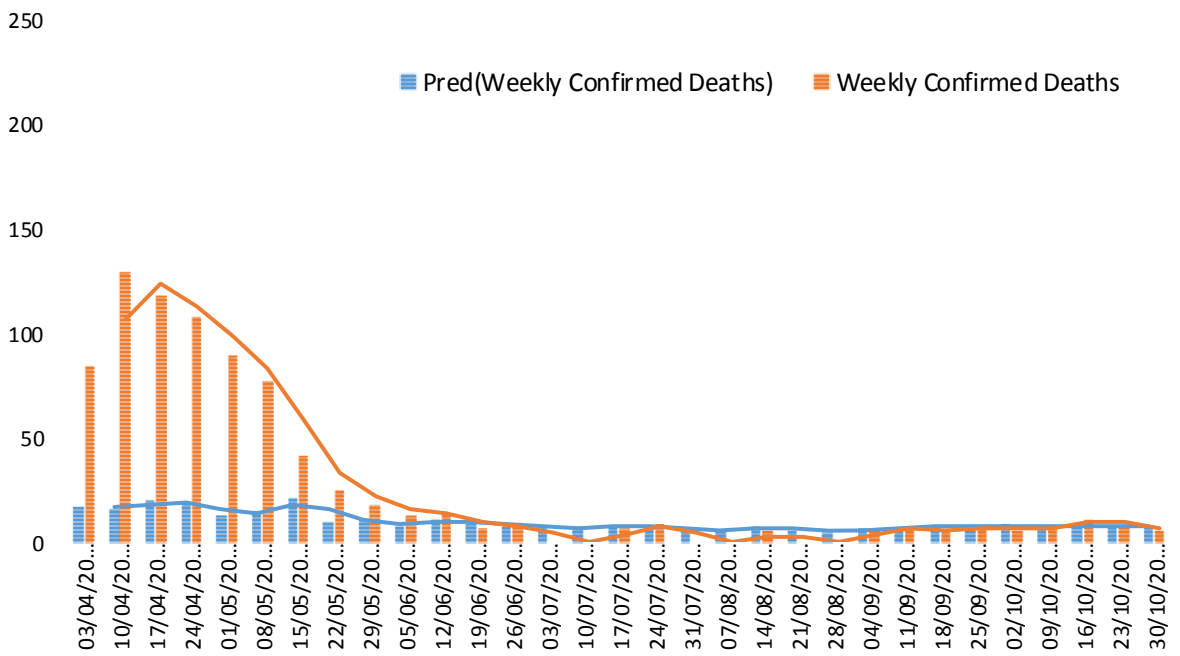


Figure 21: Predicted and Observed weekly death in County Dublin

According to Figure 21, except for the month of April and May 2020, the predicted values are close to the observed ones. High numbers of death (most of them attributed to health care facilities) in April and early May were suppressed as a result of activities of the Department of Health of Ireland.

5 Discussion, Conclusions, and recommendations

This research project studied the spatial correlations between atmospheric pollution, weather information, and social-demographic data with contagion and lethality rates of Covid-19 in Ireland. This country was selected as the case study because the extensive use of a large number of health and non-health technologies have been employed including diagnostic testing and the use of medical devices allowed us to have a better understanding of the parameter that might affect Covid-19 spread. According to the correlation circle of the factors in principal components analysis, two sets of parameters were identified; 1) those affect the health condition like aging and health status, and 2) those affect the spread of the virus such as temperature and humidity. A principal components regression analysis was carried out to a) develop a prediction model and b) to evaluate the contribution of input parameters on virus transmissions and its killing power. It was found out the atmospheric pollution contributes in two ways; a) by facilitating virus transmission and b) to less extent by forcing extra pressure on the respiratory system (e.g. background radon activity). The metrological parameters (especially air temperature and dew point temperature) were found to significantly affect both increases of contagiousness and killing power. Similar to what was anticipated, the elderly category was found to be the most vulnerable group to be hit by the virus. The health status as an indicator of the wellbeing of the immune system was found to have a key role in fighting against the virus. We found that the lockdown set by the Irish government significantly prevented the increase of covid-19 cases. Lockdown measures could have a secondary positive effect on the decrease of mortalities (i.e., by limiting the number of infected cases). However, the occupancy of hospital beds during lockdowns and also the unwillingness of people to not go to the hospital due to fear of contracting the virus may increase the mortality rate during the lockdowns.

The models we developed were successful to predict the portion of contagion and lethality rates that were related to environmental and sociodemographic data. It seems that there are additional effective parameters that are not included in this study since the input parameters introduced in the proposed model do not cover the whole variation in the predictions. Therefore we would suggest adding parameters like the number of death in care homes, percent of cumulative contagion and mortality among health workers, the degree of compatibility with social distancing, blood groups, and details about the number of tests per capita while developing

future models. The research methodology presented here can be effectively utilized for modeling Covid-19 contagion and lethality in other countries. The results of this project can be extended and used for modeling other viral infections like flu, Influenza, and Ebola.

6 Openness and Transparency

Openness and transparency are key ingredients to build accountability and trust. We see our work as a public good and are committed to making information open and accessible to academic bodies, health professionals as well as decision and policymakers at the administrative level.

We believe that a research organization like the European Open Science Cloud (EOSC) has a democratic duty to provide everyone with the information that allows them to better understand the COVID-19 spread, affecting factors, models, and preventive tools. According to the openness and open-source access policies of EOSC, we tried to at first maintain transparency in introducing the methodology, data, maps, and all related files. Any base data, map, or information used for modeling will be presented together with the technical report of the project. Anyone can reconstruct a model for the prediction of Covid-19 contagion and lethality rates by following the approach introduced here and providing the information on required parameters. Secondly, the technical report as well as all other related data can be shared on any bibliographic and scientific platform with the agreement and consent of EOSC. As an example, we are planning to submit an abstract as well as an oral presentation file for the EGU General Assembly 2021 (by 13th January 2021) (in the session: Covid-19 pandemic: health, urban systems, and geosciences). This meeting (this year will be virtual) has thousands of audience each year (please see <https://www.egu21.eu/> for details).

In general, as we strongly believe that any study that can help to fight against any disease that threatens human beings' health should have open access to the scientific society, our target will be free and public posting of our scientific findings because this research is not only useful for COVID-19 but also can be efficiently used to better understanding of the effect of environmental and sociodemographic parameters on other viral infections.

Recently scientists are sharing their Covid-19 experiences, results, and analyses in the form of published papers or even unpublished records. Because traditional academic publishing is too

slow – typically taking weeks, or even months from submission to publication, scientists are turning to a different form of open publishing – preprints.

Therefore, to increase the impact of the research carried out in this project, share the results rapidly and openly to inform the public health response and help save lives, we would offer to extract a draft of journal publication from the initial technical report in cooperation with the EOSC steering group and submit it to an open-access peer-review journal (e.g. the Open Infectious Diseases Journal, The MDPI international peer- journal of applied sciences, etc.). This would be subjected to a 25 - 30 % increase (1500 to 1800 euro) in the initial costs previously mentioned in the contract to cover the publication fees and Article Processing Charges.

7 Acknowledgements

The authors would like to thank the Irish Government, the Department of Health, Environmental Protection Agency, and Met Éireann as the sources of the raw data used as an input for the prediction model. We would like to also acknowledge Mr. Conor Rothwell for his contribution to curating the metrological dataset.

We thank Mr. Bernd Saurugger for his useful discussions and help. We also thank the referees of the European Open Science Cloud (EOSC) for the careful and insightful review of our proposal and the technical report manuscript.

This work was supported by European Open Science Cloud (EOSC). EOSCsecretariat.eu has received funding from the European Union's Horizon Programme call H2020-INFRAEOSC-05-2018-2019, grant Agreement number 831644.

8 References

- Addinsoft. (2019). *XLSTAT statistical and data analysis solutions*. <https://www.xlstat.com>.
- Alkhowailed, M., Shariq, A., Alqossayir, F., Alzahrani, O. A., Rasheed, Z., & Al Abdulmonem, W. (2020). Impact of meteorological parameters on COVID-19 pandemic: A comprehensive study from Saudi Arabia. *Informatics in Medicine Unlocked*, *20*, 100418. <https://doi.org/https://doi.org/10.1016/j.imu.2020.100418>
- Artigue, H., & Smith, G. (2019). The principal problem with principal components regression. *Cogent Mathematics & Statistics*, *6*(1), 1622190. <https://doi.org/10.1080/25742558.2019.1622190>
- Atalan, A. (2020). Is the lockdown important to prevent the COVID-9 pandemic? Effects on psychology, environment and economy-perspective. *Annals of Medicine and Surgery* (2012), *56*, 38–42. <https://doi.org/10.1016/j.amsu.2020.06.010>
- Conticini, E., Frediani, B., & Caro, D. (2020). Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy? *Environmental Pollution*, *261*, 114465. <https://doi.org/https://doi.org/10.1016/j.envpol.2020.114465>
- Copat, C., Cristaldi, A., Fiore, M., Grasso, A., Zuccarello, P., Signorelli, S. S., Conti, G. O., & Ferrante, M. (2020). The role of air pollution (PM and NO₂) in COVID-19 spread and lethality: A systematic review. *Environmental Research*, *191*, 110129. <https://doi.org/10.1016/j.envres.2020.110129>
- Decaro, N., & Lorusso, A. (2020). Novel human coronavirus (SARS-CoV-2): A lesson from animal coronaviruses. *Veterinary Microbiology*, *244*, 108693. <https://doi.org/https://doi.org/10.1016/j.vetmic.2020.108693>
- Dempsey, S., Lyons, S., & Nolan, A. (2018). High Radon Areas and lung cancer prevalence: Evidence from Ireland. *Journal of Environmental Radioactivity*, *182*, 12–19. <https://doi.org/https://doi.org/10.1016/j.jenvrad.2017.11.014>
- Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*,

- 36(1), 27–46. <https://doi.org/https://doi.org/10.1111/j.1600-0587.2012.07348.x>
- Flaherty, G. T., Hession, P., Liew, C. H., Lim, B. C. W., Leong, T. K., Lim, V., & Sulaiman, L. H. (2020). COVID-19 in adult patients with pre-existing chronic cardiac, respiratory and metabolic disease: a critical literature review with clinical recommendations. *Tropical Diseases, Travel Medicine and Vaccines*, 6(1), 16. <https://doi.org/10.1186/s40794-020-00118-y>
- for Disease Control, C., Prevention, & others. (n.d.). *Scientific brief: SARS-CoV-2 and potential airborne transmission*. 2020.
- Imai, C., Armstrong, B., Chalabi, Z., Mangtani, P., & Hashizume, M. (2015). Time series regression model for infectious disease and weather. *Environmental Research*, 142, 319–327. <https://doi.org/https://doi.org/10.1016/j.envres.2015.06.040>
- Jayaweera, M., Perera, H., Gunawardana, B., & Manatunge, J. (2020). Transmission of COVID-19 virus by droplets and aerosols: A critical review on the unresolved dichotomy. *Environmental Research*, 188, 109819. <https://doi.org/10.1016/j.envres.2020.109819>
- Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065), 20150202.
- Kennelly, B., O’Callaghan, M., Coughlan, D., Cullinan, J., Doherty, E., Glynn, L., Moloney, E., & Queally, M. (2020). The COVID-19 pandemic in Ireland: An overview of the health service and economic policy response. *Health Policy and Technology*, 9(4), 419–429. <https://doi.org/https://doi.org/10.1016/j.hlpt.2020.08.021>
- Lolli, S., Chen, Y.-C., Wang, S.-H., & Vivone, G. (2020). Impact of meteorological conditions and air pollution on COVID-19 pandemic transmission in Italy. *Scientific Reports*, 10(1), 16213. <https://doi.org/10.1038/s41598-020-73197-8>
- OECD. (2017). *Health at a Glance*.
- Petrosillo, N., Viceconte, G., Ergonul, O., Ippolito, G., & Petersen, E. (2020). COVID-19, SARS and MERS: are they closely related? *Clinical Microbiology and Infection : The Official Publication of the European Society of Clinical Microbiology and Infectious Diseases*, 26(6), 729–734. <https://doi.org/10.1016/j.cmi.2020.03.026>
- Tammes, P. (2020). Social distancing, population density, and spread of COVID-19 in England: a

longitudinal study. *BJGP Open*, 4(3), bjgpopen20X101116.

<https://doi.org/10.3399/bjgpopen20X101116>

te Vrugt, M., Bickmann, J., & Wittkowski, R. (2020). Effects of social distancing and isolation on epidemic spreading modeled via dynamical density functional theory. *Nature*

Communications, 11(1), 5576. <https://doi.org/10.1038/s41467-020-19024-0>

WHO (World Health Organization). (2020a). *Naming the coronavirus disease (COVID-19) and the virus that causes it.*

WHO (World Health Organization). (2020b). *Transmission of SARS-CoV-2: implications for infection prevention precautions: scientific brief, 09 July 2020.*

Zoran, M. A., Savastru, R. S., Savastru, D. M., & Tautan, M. N. (2020). Assessing the relationship between surface levels of PM_{2.5} and PM₁₀ particulate matter impact on COVID-19 in Milan, Italy. *Science of The Total Environment*, 738, 139825.

<https://doi.org/https://doi.org/10.1016/j.scitotenv.2020.139825>

Appendix I: Supplementary data

Chart I-1: Standardized Residuals

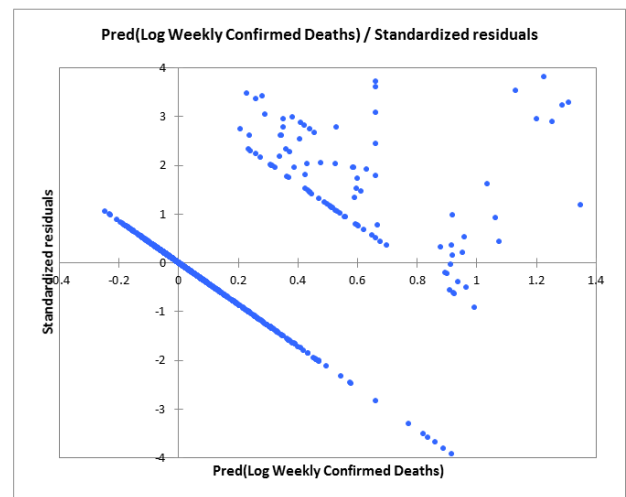
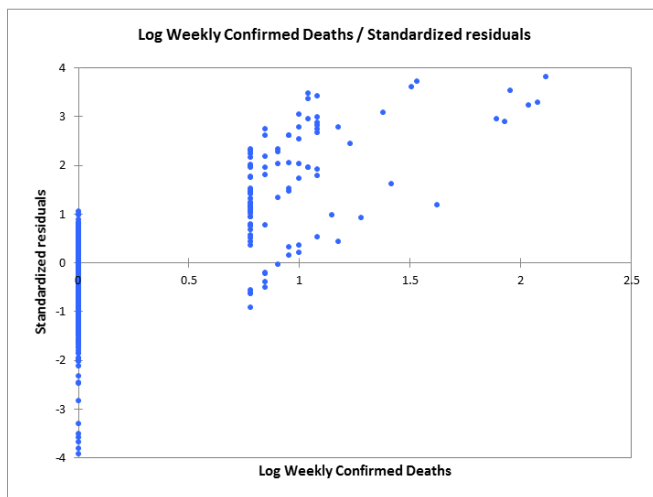
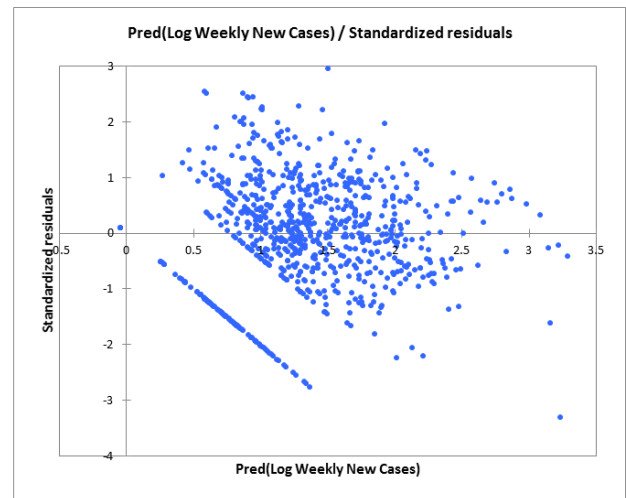
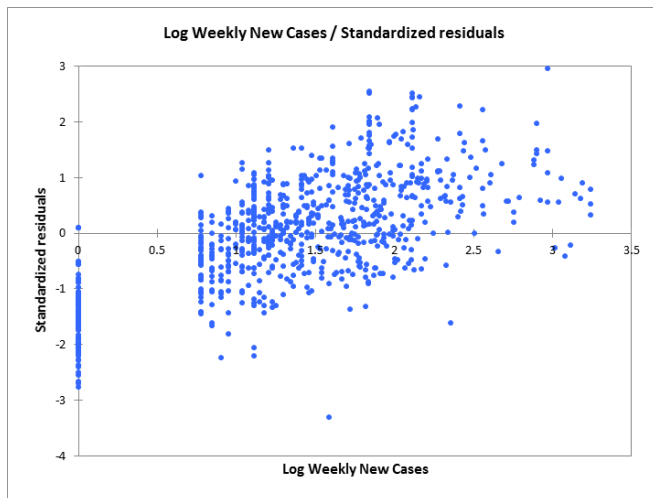



Table I-1: Correlation matrix

Variables	PM10	PM2.5	SO2	NO2	O3	CO	rain	temp	sun	wetb	wdsp	dewpt	vappr	rhum	msl	AGE<20T	AGE20_24T	AGE25_29T	AGE30_34T	AGE35_39T	AGE40_44T	AGE45_49T	AGE50_54T	AGE55_59T	AGE60_64T	AGE65_69T	AGE70_74T	AGE75_79T	AGE80_84T	AGE85T	VGT	GT	FT	BT	VBT	AGETF	AGETM	P Density	Rn_mean	Lockdown phas	Population	Log Weekly New	Log Weekly Cor
PM10	1.000	0.871	-0.130	0.003	0.155	0.582	-0.064	-0.224	0.139	-0.220	-0.154	-0.219	-0.187	-0.121	0.072	-0.010	-0.076	-0.022	-0.040	-0.007	-0.005	-0.005	0.036	0.042	0.046	0.051	0.031	0.036	0.030	-0.042	0.056	0.012	0.002	0.002	-0.031	0.031	P Density	0.005	-0.312	-0.009	0.250	0.306	
PM2.5	0.871	1.000	-0.102	0.048	0.218	0.683	-0.099	-0.339	0.139	-0.330	-0.185	-0.322	-0.291	-0.166	0.067	-0.036	-0.089	-0.051	-0.057	-0.046	-0.057	0.003	0.055	0.082	0.087	0.088	0.079	0.058	0.062	0.046	-0.049	0.084	0.010	-0.044	-0.040	-0.072	0.072	-0.011	0.017	-0.320	-0.023	0.277	0.307
SO2	-0.130	-0.102	1.000	0.024	-0.045	-0.094	0.153	-0.127	-0.315	-0.072	0.143	-0.015	-0.041	0.235	-0.294	-0.001	-0.112	-0.176	-0.177	-0.106	0.009	0.145	0.164	0.080	0.103	0.111	0.119	0.105	0.072	0.093	0.032	0.001	0.056	-0.055	-0.066	0.097	-0.097	0.034	0.155	0.528	-0.040	0.192	-0.241
NO2	0.003	0.048	0.024	1.000	-0.227	0.317	-0.232	-0.407	-0.057	-0.440	0.086	-0.456	-0.458	-0.426	0.134	-0.006	-0.143	-0.255	-0.238	-0.110	0.086	0.233	0.217	0.158	0.161	0.123	0.127	0.081	0.044	0.110	0.048	-0.006	0.047	-0.015	0.070	0.065	-0.065	0.206	0.138	-0.300	-0.120	-0.067	-0.033
O3	0.155	0.218	-0.045	-0.227	1.000	0.076	0.033	-0.012	0.108	-0.010	0.187	-0.016	-0.016	-0.031	0.051	-0.303	0.062	0.128	0.024	0.019	-0.127	-0.156	0.001	0.075	0.112	0.162	0.230	0.205	0.104	0.015	-0.045	0.034	-0.017	-0.178	-0.084	0.176	-0.176	-0.203	0.263	-0.155	0.134	0.089	0.112
CO	0.582	0.683	-0.094	0.317	0.076	1.000	-0.324	-0.606	0.195	-0.611	-0.085	-0.603	-0.580	-0.404	0.287	0.013	-0.174	-0.300	-0.305	-0.225	-0.026	0.086	0.253	0.223	0.218	0.223	0.179	0.136	0.135	0.134	-0.012	0.140	0.077	-0.076	-0.037	-0.090	0.090	0.033	0.109	-0.455	-0.243	0.146	0.220
rain	-0.064	-0.099	0.153	-0.232	0.033	-0.324	1.000	0.422	-0.488	0.502	0.212	0.557	0.559	0.677	-0.702	-0.038	-0.093	-0.071	-0.079	-0.093	-0.116	-0.036	0.020	0.105	0.117	0.119	0.127	0.101	0.096	0.081	-0.140	0.135	0.115	0.088	0.038	-0.092	0.092	0.004	0.015	0.455	-0.101	-0.093	-0.261
temp	-0.224	-0.339	-0.127	-0.407	-0.012	-0.606	0.422	1.000	-0.211	0.983	-0.164	0.956	0.952	0.545	-0.184	0.046	0.017	0.061	0.047	0.054	-0.024	-0.038	-0.100	-0.061	-0.047	-0.053	-0.027	-0.034	-0.035	-0.072	-0.085	0.049	0.024	0.092	0.078	-0.126	0.126	-0.016	-0.058	0.431	-0.069	-0.381	-0.357
sun	0.139	0.139	-0.315	-0.057	0.108	0.195	-0.488	-0.211	1.000	-0.289	-0.361	-0.354	-0.340	-0.538	0.430	0.030	0.048	0.021	0.027	0.008	-0.002	-0.016	0.020	-0.033	-0.045	-0.058	-0.054	-0.028	-0.011	-0.033	0.070	-0.041	-0.057	-0.069	-0.081	-0.048	0.048	-0.069	0.019	-0.416	-0.036	0.073	0.208
wetb	-0.220	-0.330	-0.072	-0.440	-0.010	-0.611	0.502	0.983	-0.289	1.000	-0.136	0.987	0.984	0.672	-0.257	0.053	0.004	0.050	0.039	0.045	-0.035	-0.046	-0.101	-0.055	-0.039	-0.048	-0.021	-0.030	-0.025	-0.059	-0.090	0.054	0.037	0.100	0.076	-0.145	0.145	-0.009	-0.068	0.489	-0.081	-0.359	-0.371
wdsp	-0.154	-0.185	0.143	0.086	0.187	-0.085	0.212	-0.164	-0.361	-0.136	1.000	-0.116	-0.138	0.006	-0.231	-0.098	0.111	0.046	-0.012	0.036	0.044	-0.026	-0.110	-0.040	-0.001	0.062	0.071	0.046	-0.029	-0.018	-0.051	0.032	0.041	0.001	0.067	0.232	-0.232	-0.026	0.071	0.083	0.146	0.151	-0.018
dewpt	-0.219	-0.322	-0.015	-0.456	-0.016	-0.603	0.557	0.956	-0.354	0.987	-0.116	1.000	0.996	0.767	-0.313	0.056	-0.009	0.038	0.026	0.033	-0.039	-0.040	-0.092	-0.046	-0.030	-0.039	-0.012	-0.023	-0.017	-0.051	-0.091	0.059	0.046	0.106	0.083	-0.147	0.147	-0.006	-0.073	0.547	-0.088	-0.334	-0.387
vappr	-0.187	-0.291	-0.041	-0.458	-0.016	-0.580	0.559	0.952	-0.340	0.984	-0.138	0.996	1.000	0.761	-0.309	0.065	-0.016	0.038	0.027	0.035	-0.042	-0.045	-0.100	-0.048	-0.031	-0.042	-0.015	-0.026	-0.017	-0.051	-0.097	0.062	0.052	0.115	0.087	-0.164	0.164	0.004	-0.086	0.530	-0.096	-0.318	-0.367
rhum	-0.121	-0.166	0.235	-0.426	-0.031	-0.404	0.677	0.545	-0.538	0.672	0.006	0.767	0.761	1.000	-0.503	0.051	-0.068	-0.021	-0.029	-0.034	-0.069	-0.038	-0.044	0.012	0.026	0.011	0.030	0.017	0.044	0.030	-0.081	0.071	0.086	0.112	0.072	-0.153	0.153	0.013	-0.085	0.619	-0.097	-0.117	-0.325
msl	0.072	0.067	-0.294	0.134	0.051	0.287	-0.702	-0.184	0.430	-0.257	-0.231	-0.313	-0.309	-0.503	1.000	0.008	0.004	-0.010	-0.004	0.018	0.037	0.007	0.031	-0.004	-0.009	-0.008	-0.010	-0.016	-0.020	-0.028	0.040	-0.018	-0.045	-0.046	-0.002	0.007	-0.007	0.007	-0.009	-0.421	-0.004	-0.039	0.205
AGE<20T	-0.010	-0.036	-0.001	-0.006	-0.303	0.013	-0.038	0.046	0.030	0.053	-0.098	0.056	0.065	0.051	0.008	1.000	-0.246	-0.201	0.054	0.333	0.572	0.581	-0.189	-0.534	-0.561	-0.583	-0.614	-0.707	-0.647	-0.576	0.485	-0.411	-0.303	-0.233	-0.233	-0.398	0.398	0.016	-0.483	0.000	-0.441	-0.031	-0.121
AGE20_24T	-0.076	-0.089	-0.112	-0.143	0.062	-0.174	-0.093	0.017	0.048	0.004	0.111	-0.009	-0.016	-0.068	0.004	-0.246	1.000	0.812	0.721	0.588	0.141	-0.417	-0.619	-0.576	-0.560	-0.510	-0.491	-0.364	-0.404	-0.474	0.172	-0.285	-0.419	-0.079	0.063	0.466	-0.466	-0.278	0.112	0.000	0.617	0.259	0.276
AGE25_29T	-0.022	-0.051	-0.176	-0.255	0.128	-0.300	-0.071	0.061	0.021	0.050	0.046	0.038	0.038	-0.021	-0.010	-0.201	0.812	1.000	0.904	0.683	0.091	-0.512	-0.775	-0.655	-0.632	-0.611	-0.561	-0.408	-0.371	-0.440	0.176	-0.387	-0.393	-0.054	0.080	0.313	-0.313	-0.197	-0.196	0.000	0.809	0.311	0.430
AGE30_34T	-0.040	-0.057	-0.177	-0.238	0.024	-0.305	-0.079	0.047	0.027	0.039	-0.012	0.026	0.027	-0.029	-0.004	0.054	0.721	0.904	1.000	0.855	0.326	-0.344	-0.808	-0.769	-0.778	-0.783	-0.779	-0.671	-0.606	-0.609	0.413	-0.573	-0.620	-0.275	-0.213	0.190	-0.190	-0.049	-0.316	0.000	0.725	0.302	0.409
AGE35_39T	-0.007	-0.046	-0.106	-0.110	0.019	-0.225	-0.093	0.054	0.008	0.045	0.036	0.033	0.035	-0.034	0.018	0.333	0.855	1.000	0.700	0.072	-0.724	-0.899	-0.892	-0.894	-0.875	-0.863	-0.853	-0.821	-0.680	-0.756	-0.840	-0.503	-0.333	0.240	-0.240	0.054	-0.240	0.000	0.550	0.266	0.301		
AGE40_44T	-0.005	-0.057	0.009	0.086	-0.127	-0.026	-0.116	-0.024	-0.002	-0.035	-0.044	-0.039	-0.042	-0.069	0.037	0.572	0.141	0.091	0.326	0.700	1.000	0.639	-0.205	-0.659	-0.685	-0.685	-0.689	-0.777	-0.856	-0.747	0.805	-0.688	-0.812	-0.641	-0.402	0.253	-0.253	0.168	-0.188	0.000	0.110	0.157	0.091
AGE45_49T	-0.005	0.003	0.145	0.233	-0.156	0.086	-0.036	-0.038	-0.016	-0.046	-0.026	-0.040	-0.045	-0.068	0.007	0.581	-0.417	-0.512	-0.344	0.072	0.639	1.000	0.366	-0.157	-0.170	-0.154	-0.137	-0.294	-0.412	-0.367	0.395	-0.202	-0.303	-0.327	-0.120	0.095	-0.095	0.035	0.004	0.000	-0.403	-0.049	-0.179
AGE50_54T	-0.005	0.055	0.164	0.217	0.001	0.253	0.020	-0.100	0.020	-0.101	-0.110	-0.092	-0.100	-0.044	0.031	-0.189	-0.619	-0.775	-0.808	-0.724	-0.205	0.366	1.000	0.746	0.669	0.686	0.675	0.615	0.519	0.492	-0.149	0.327	0.304	0.008	0.049	0.079	-0.079	0.055	0.377	0.000	-0.497	-0.227	-0.294
AGE55_59T	0.036	0.082	0.080	0.158	0.075	0.223	0.105	-0.061	-0.033	-0.055	-0.040	-0.046	-0.048	0.012	-0.004	-0.534	-0.576	-0.655	-0.769	-0.899	-0.659	-0.157	0.746	1.000	0.979	0.956	0.922	0.887	0.874	0.875	-0.606	0.696	0.688	0.347	0.136	-0.163	0.163	0.121	0.387	0.000	-0.407	-0.276	-0.262
AGE60_64T	0.042	0.087	0.103	0.161	0.112	0.218	0.117	-0.047	-0.043	-0.039	-0.001	-0.030	-0.031																														

Table I-3: Factor Loadings

	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15	F16	F17	F18	F19	F20	F21	F22	F23	F24	F25	F26	F27	F28	F29	F30	F31	F32	F33	F34	F35	F36	F37	F38	F39	
PM10	0.058	-0.359	0.016	-0.230	0.584	-0.391	-0.356	-0.271	-0.062	-0.029	0.071	-0.151	0.104	0.034	-0.011	-0.157	0.092	0.026	-0.013	0.037	0.048	-0.059	0.152	-0.079	-0.064	-0.022	0.032	0.001	-0.003	-0.007	-0.002	0.003	0.002	0.000	0.001	-0.001	0.000	0.000	0.000	0.000
PM2.5	0.093	-0.446	0.027	-0.218	0.578	-0.436	-0.333	-0.164	-0.104	-0.005	0.050	-0.004	0.042	-0.016	0.088	-0.057	-0.005	-0.003	0.049	0.028	0.057	0.097	-0.123	0.134	0.074	0.011	-0.041	-0.007	0.003	0.011	0.006	-0.003	-0.002	0.000	-0.001	0.000	0.000	0.000	0.000	0.000
SO2	0.098	0.120	0.112	0.459	-0.246	-0.384	0.057	0.206	-0.529	-0.142	0.264	0.022	0.254	0.009	-0.076	-0.039	-0.096	0.089	-0.184	-0.030	0.047	-0.037	-0.034	-0.009	0.004	0.007	-0.010	0.013	0.003	0.002	0.000	-0.002	0.001	0.000	0.001	0.000	0.000	0.000	0.000	
NO2	0.125	-0.487	0.160	0.174	-0.354	-0.043	0.077	-0.361	0.202	0.384	0.145	0.301	0.274	0.103	0.108	0.017	0.030	0.063	0.051	0.135	0.018	-0.045	-0.036	-0.023	-0.004	-0.007	0.004	-0.009	0.003	-0.001	-0.002	-0.001	0.000	-0.002	0.000	-0.001	0.000	0.000	0.000	
O3	0.074	-0.075	-0.298	0.161	0.493	0.011	-0.150	0.484	0.323	-0.303	0.128	0.273	0.218	-0.009	-0.049	0.132	-0.053	-0.006	0.108	-0.003	0.017	-0.018	0.045	-0.009	0.024	0.024	0.028	-0.010	-0.014	0.000	-0.002	0.001	-0.001	0.000	-0.003	0.001	-0.002	0.000	0.000	
CO	0.216	-0.715	0.170	-0.122	0.254	-0.268	-0.212	-0.069	-0.030	0.099	0.130	0.035	-0.227	0.183	-0.005	0.153	-0.141	-0.030	-0.033	-0.084	-0.157	-0.053	-0.091	-0.094	0.005	0.021	0.021	0.023	0.004	-0.010	-0.003	0.000	-0.002	-0.001	0.000	0.001	0.000	0.000	0.000	
rain	0.127	0.667	0.067	0.198	0.103	-0.429	-0.097	0.018	0.193	0.129	-0.312	0.020	0.021	0.030	0.028	0.120	0.107	0.300	-0.070	-0.103	-0.061	0.022	0.024	0.012	0.030	-0.005	0.002	-0.002	-0.003	0.000	0.000	0.001	0.000	-0.001	0.000	0.000	0.000	0.000	0.000	
temp	-0.037	0.856	0.062	-0.064	0.222	0.346	-0.064	-0.126	0.086	0.045	0.096	-0.026	0.093	0.059	0.077	-0.116	-0.103	0.017	-0.026	-0.053	-0.022	-0.007	-0.015	-0.002	0.016	0.006	0.008	0.008	0.001	0.008	-0.004	-0.008	0.001	-0.034	-0.004	-0.003	-0.001	0.001	-0.009	
sun	-0.043	-0.462	-0.025	-0.315	0.192	0.495	0.025	0.197	-0.179	-0.100	-0.310	-0.219	0.267	0.251	0.065	0.116	0.086	0.022	-0.050	0.081	0.009	-0.041	-0.080	-0.019	-0.004	0.005	-0.007	0.008	0.000	-0.002	-0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
wetb	-0.026	0.910	0.077	-0.042	0.224	0.250	-0.068	-0.112	0.069	0.035	0.090	-0.008	0.050	0.049	0.024	-0.067	-0.060	0.003	-0.034	-0.003	-0.005	-0.013	-0.027	-0.005	0.011	0.004	0.002	0.004	0.004	-0.005	0.006	0.034	-0.043	0.028	0.007	0.006	-0.001	0.000	0.000	
wdsp	-0.004	0.017	-0.204	0.344	-0.338	-0.364	-0.027	0.212	0.581	-0.163	0.166	-0.259	-0.078	0.203	0.074	-0.127	0.063	-0.030	-0.059	0.039	0.062	0.018	-0.062	-0.022	0.011	0.013	-0.002	0.013	0.004	-0.001	0.000	-0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	
dewpt	-0.015	0.940	0.089	-0.011	0.212	0.169	-0.077	-0.106	0.039	0.022	0.088	0.010	0.010	0.041	-0.009	-0.015	-0.024	-0.013	-0.028	0.030	0.000	-0.016	-0.036	-0.011	-0.001	0.002	0.003	0.006	0.000	0.001	-0.003	-0.006	0.000	-0.023	-0.003	-0.003	0.001	-0.003	0.014	
vappr	-0.014	0.932	0.096	-0.041	0.235	0.160	-0.085	-0.123	0.042	0.025	0.079	0.007	0.016	0.045	-0.013	-0.004	-0.023	-0.010	-0.016	0.018	-0.008	-0.015	-0.035	-0.013	-0.002	0.006	-0.005	-0.002	-0.002	-0.007	0.002	-0.024	0.047	0.035	-0.001	0.001	0.000	0.001	-0.001	
rhum	0.047	0.816	0.115	0.098	0.130	-0.281	-0.078	-0.034	-0.098	-0.038	0.016	0.081	-0.168	-0.019	-0.188	0.202	0.164	-0.076	-0.036	0.202	0.051	-0.030	-0.060	-0.015	-0.021	-0.010	-0.004	0.004	0.000	-0.001	0.000	-0.003	0.003	-0.008	-0.010	0.000	-0.001	0.001	-0.004	
msl	-0.024	-0.481	-0.020	-0.259	0.059	0.563	0.053	-0.053	-0.030	-0.147	0.484	0.004	-0.181	-0.003	-0.010	0.052	0.178	0.217	-0.045	-0.004	0.002	0.033	0.008	0.018	0.023	-0.006	0.001	0.001	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	
AGE<20T	-0.492	0.027	0.738	-0.265	-0.209	-0.045	-0.194	0.093	0.018	-0.088	-0.008	-0.012	0.009	0.040	-0.111	-0.098	-0.009	0.054	0.062	0.030	-0.071	-0.010	-0.024	0.018	-0.029	0.028	0.001	-0.045	-0.022	0.016	-0.014	-0.006	-0.007	0.002	-0.001	-0.002	0.001	0.004	0.000	
AGE20_24T	-0.550	0.034	-0.694	-0.005	-0.034	0.026	-0.088	0.043	-0.078	0.281	0.047	-0.079	-0.071	0.177	-0.054	0.118	-0.144	0.011	-0.045	0.042	0.099	0.066	0.066	0.079	0.016	-0.042	0.067	-0.015	-0.003	-0.020	0.002	0.004	0.005	-0.002	0.001	0.006	0.000	0.001	0.000	
AGE25_29T	-0.614	0.110	-0.721	-0.204	0.027	-0.073	0.077	-0.022	-0.044	-0.006	-0.009	0.084	0.022	-0.010	0.010	-0.005	0.006	0.010	-0.019	0.068	0.011	-0.092	0.035	-0.040	-0.059	0.060	0.017	0.043	-0.023	0.020	0.010	0.002	-0.009	0.002	-0.003	0.000	0.000	0.000		
AGE30_34T	-0.806	0.082	-0.456	-0.204	0.063	-0.093	0.189	0.017	-0.035	0.051	-0.020	0.035	-0.003	-0.116	0.068	-0.015	0.029	0.039	0.027	-0.014	0.058	-0.009	-0.051	-0.058	-0.015	0.039	0.023	0.056	-0.015	-0.017	0.042	-0.027	-0.017	0.003	-0.003	-0.009	0.003	0.000	0.000	
AGE35_39T	-0.955	0.019	-0.129	0.001	0.077	-0.048	0.040	-0.017	0.053	0.056	0.086	-0.010	0.096	-0.065	0.029	0.059	0.086	-0.052	-0.027	-0.069	-0.014	-0.013	-0.017	-0.004	-0.030	0.079	-0.032	-0.037	0.030	-0.040	0.000	0.021	0.016	-0.008	0.015	0.005	0.004	-0.009	-0.001	
AGE40_44T	-0.772	-0.123	0.400	0.300	0.000	0.060	-0.123	-0.119	0.080	-0.038	0.045	-0.115	0.021	-0.058	0.072	0.124	0.001	-0.084	-0.056	-0.014	-0.001	-0.116	0.082	-0.017	0.136	-0.067	-0.057	-0.018	-0.009	-0.010	0.014	-0.013	-0.007	0.000	-0.002	0.002	-0.003	0.001	0.000	
AGE45_49T	-0.183	-0.137	0.701	0.362	-0.137	0.146	-0.394	-0.111	0.039	-0.124	-0.026	-0.043	0.124	-0.162	0.080	0.069	0.037	-0.060	-0.057	0.031	-0.057	0.150	-0.010	-0.006	0.000	-0.008	0.081	0.059	0.012	-0.002	0.003	0.003	0.005	0.000	0.001	0.006	-0.003	-0.002	0.000	
AGE50_54T	0.649	-0.172	0.371	0.450	0.079	0.219	-0.094	-0.080	-0.102	-0.065	-0.162	0.106	-0.144	0.031	0.024	-0.059	-0.029	0.039	0.023	-0.036	0.186	0.065	-0.037	-0.108	0.044	0.011	-0.009	-0.022	-0.029	-0.036	-0.006	0.017	0.007	-0.001	-0.008	-0.005	0.003	-0.001	0.000	
AGE55_59T	0.943	-0.057	0.062	0.150	0.118	0.049	0.149	-0.014	-0.001	0.057	-0.037	0.004	-0.076	-0.037	0.065	-0.005	0.022	0.005	0.033	-0.028	0.096	-0.093	-0.006	0.007	0.012	0.031	0.012	0.041	-0.016	0.030	-0.013	-0.014	0.000	-0.004	0.018	0.028	0.001	0.000	0.000	
AGE60_64T	0.966	-0.034	0.023	0.114	0.099	0.015	0.139	-0.014	0.033	0.033	0.028	-0.016	0.008	-0.037	0.064	0.059	0.017	-0.033	-0.023	-0.024	0.021	-0.057	0.009	0.043	0.004	0.016	0.004	0.032	0.009	0.000	0.006	0.025	0.014	-0.003	0.016	-0.023	-0.003	0.011	0.001	
AGE65_69T	0.956	-0.044	-0.047	0.185	0.085	0.037	0.020	0.011	0.043	-0.011	0.021	-0.021	-0.037	-0.023	0.061	0.015	0.004	-0.005	-0.011	-0.051	-0.004	-0.083	-0.038	0.076	-0.094	-0.060	-0.008	0.011	-0.021	0.000	0.006	0.007	0.002	0.004	-0.015	-0.004	-0.012	-0.011	-0.001	
AGE70_74T	0.952	-0.020	-0.097	0.205	0.086	0.055	-0.020	-0.012	0.061	-0.060	0.028	0.022	0.057	-0.005	0.006	0.029	0.045	-0.042	-0.007	-0.057	-0.021	-0.005	-0.017	0.019	-0.047	-0.067	-0.008	-0.008	0.051	-0.008	0.007	-0.003	-0.005	0.001	-0.014	0.007	0.020	0.004	0.000	
AGE75_79T	0.931	-0.010	-0.245	0.157	0.081	0.050	0.075	-0.055	-0.002	-0.081	-0.018	0.014	0.030	0.032	-0.039	0.008	-0.034	0.009	0.009	0.013	-0.013	0.095	0.001	-0.055	0.008	-0.051	-0.009	-0.031	0.041	-0.002	-0.012	-0.032	-0.016	0.003	0.023	-0.012	-0.004	-0.006	-0.001	
AGE80_84T	0.924	0.019	-0.213	-0.029	0.072	-0.011	0.195	-0.022	-0.061	-0.050	-0.053	0.083	-0.030	0.047	0.009	-0.078	0.038	-0.033	-0.084	0.043	-0.069	0.060	0.023	-0.046	0.017	0.039	0.003	-0.059	0.015	0.012	0.043	0.004	0.006	-0.007	-0.006	0.014	-0.011	0.004	0.001	
AGEGE_85T	0.881	-0.014	-0.093	-0.003	0.067	-0.062	0.362	-0.097	-0.029	-0.056	-0.039	-0.003	-0.009	0.025	-0.017	-0.064	0.013	-0.027	-0.058	0.106	-0.097	-0.020	0.088	0.058																



Modeling of the Covid-19 epidemic using geostatistical methods and based on the investigation of spatial correlations between atmospheric pollution, meteorological factors, age, sex, health condition, and lockdown phases with levels of virus contagion and lethality

Technical Report