
COVID-19 PANDEMIC R EFFECTIVE REPRODUCTION NUMBER BASED CORRELATION DYNAMICS

Nabil Arhab, Yazid Bounab, Nacer Boudjemia, Md Saroar Jahan

PhD student

University of Oulu

<nabil.arhab>, <Yazid.Bounab>, <nacer.boudjemia>@oulu.fi, mjahan18@edu.oulu.fi

ABSTRACT

A comprehensive investigation on the variation of the Covid-19 pandemic adequate reproduction number (R) in four critical European countries is provided in this study as a function of the variations related to the weather, mobility, government responses, and epidemiology. In this study, an open data set about the Covid-19 pandemic is used for the analysis. The data contains newly recorded components with different metrics since the appearance of the COVID19 pandemic. The aim is to investigate the impact of the weather, mobility, and government restriction on the R values and pandemic. The appropriate statistical analyses are used to reveal the association and the effect of the various attributes to the R . In addition, multiple predictive models were applied. The results show significant differences between studied countries and have different factors that influence the virus's spread.

Keywords Covid19 · R · weather · demography · mobility · governments restrictions.

1 Introduction

The COVID-19 widespread driven to a sensational misfortune of human life around the world and presents a bizarre challenge to open wellbeing, nourishment frameworks, and the world of work. The financial and social unsettling influence caused by the widespread is overpowering. As a result, numerous companies confront an existential danger to their jobs, reflecting on cutting a critical number of workers to outlive. In this manner, an overwhelming burden is put on the social security companies to supply a source of earnings for their clients who cannot bolster themselves and their families amid lockdowns.

The main challenges that all the governments around the world are facing nowadays are related to the quickness of the spread of the Covid19 virus in the population. This is due to its impact on the economy, society, and health. One of the measures that allow access to the evolution of this pandemic is the effective reproduction number know as R . It is a metric that gives the probability of an infected person passing the virus to another person. When the factor $R < 1$, one person is less likely to pass the virus to one other person; therefore, this metric is desperately essential to bring it down and control the infection rate.

The number R is not fixed, and it depends on different significant factors that contribute to its increase and the contamination speed of the virus, such as the biology of the virus, people's behavior, Social distancing, and population immunity, though, what about the weather, mobility and governments responses?

2 Related work

Several studies have been reported before about the evolution of the Covid-19 pandemic with different parameters in different countries. Therefore, Ramadhan Tosepu et al. [1], studied a correlation between the Covid-19 pandemic and the weather in Jakarta and Indonesia has been investigated. Their study reported a significant correlation between average temperature and the covid-19 pandemic, where the observation is used to reduce the incidence rate in Indonesia by strengthening or softening the restriction measures. In the work of Zohair Malki et al. [2], a machine learning

approach has been used for the estimation of the mortality rate as an association between the weather and the Covid-19 pandemic. The goal was to answer the question of whether high temperature during summer and low temperature during winter can slow down or brisk the spread of the Covid-19 pandemic as it is found in the past decades on seasonal flues. Therefore, a correlation was reported between the weather and the mortality rate. The same study has been carried out in Turkey by Mehmet Şahin [3], where he observed an important correlation between the spread of the virus and the weather (high wind speed and low temperatures). In addition, Other related studies have been carried out as well by Renato Pedrosa [4], where the dynamics of Covid-19: weather, demographics, and infection timeline have been investigated, in which the effects of temperature, absolute humidity, population density are taken into account for the United States when 100 cases were detected, and the study showed a correlation with the spread of the virus as those factors are high at lower temperatures. In the study of Ran Xu et al. [5], the study's focus was put on the impact of weather and air pollution on Covid-19 transmission in 1072 global cities. In this study, Ran Xu et al. observed that a warmer temperature and moderate outdoor ultraviolet may play an essential role in reducing the transmission of the Covid-19 pandemic.

In our case, the R metric was taken as the main study parameter in correlation with the weather (average temperature, relative humidity, and rainfall), mobility (see, Table 5), and government response (see, Table 7) for key European countries including Finland, France, Germany, and United Kingdom. The choices of those countries are made due to the ample cleavage that exists between the chosen attributes (weather, mobility, and government responses).

3 Objectives

This pandemic has other effects on social life in which the lockdown becomes too frustrating because of the imposed restrictions that lock people in their homes for a long time. Therefore, we wanted to identify the policies that do not impact the spread of the virus going back to everyday life, respecting the safety measures. In order to determine the significant factors that influence the rate of spread of COVID19; we introduced three fundamental research questions associated with three domains :

- **RQ1:** How the weather conditions have an impact on increasing/decreasing the metric R?
- **RQ2:** What kind of mobility that have an impact on increasing/decreasing the metric R?
- **RQ3:** What are the government restrictions that have an impact on decreasing the metric R?

The following three domains were selected to answer the research questions: Weather, Mobility, and governmental response. Each domain includes a set of specific factors used to examine the association and their effects on the metric R. After that, statistical and machine learning techniques are used to determine the factors that correlate with the metric R as features. Finally, the most significant features used to build a predictive model to forecast the R values. Four countries were selected to ensure the models' generalization (France, Finland, Germany, and United Kingdom).

- France: had a serious lockdown,
- Finland: have very harsh weather,
- Germany: is the most populated country in Europe, and
- United Kingdom: exposure to new Covid-19 variants and tough restrictions.

4 Data

4.1 COVID-19 Open-Data

The COVID-19 Open-Data [6] contains datasets of daily time-series data related to COVID-19 for over 20,000 distinct locations worldwide. The data is at the spatial resolution of states/provinces for most regions and county/municipality resolution for many countries such as Argentina, Brazil, Chile, Colombia, Czech Republic, Mexico, Netherlands, Peru, United Kingdom, and the USA. All regions are assigned a unique location key, which resolves discrepancies between ISO / NUTS / FIPS codes. The different aggregation levels are:

1. Country
2. Province, state, or local equivalent
3. Municipality, county, or local equivalent
4. Locality which may not follow strict hierarchical order, such as "city" or "nursing homes in X location".

Table 1: Datasets name and sources.

Datasets, Start date	Content	Source
Epidemiology, 2019.12.31	COVID-19 cases, deaths, recoveries and tests	Various sources
Mobility, 2020.02.15	Various metrics related to the movement of people	Google
Government Response, 2020.01.01	Government interventions and their relative stringency	University of Oxford
Weather, 2020.01.01	Dated meteorological information for each region	NOAA

Table 2: Datasets name and variables.

Datasets	Variables
Mobility	mobility of grocery pharmacy, parks, transit stations, retail recreation, residential and workplaces.
Government Response	school closing, workplace closing, cancel public events, restrictions on gatherings etc.
Weather	average, minimum and maximum temperature, rainfall, snowfall, dew point, relative humidity, etc.

The datasets carries mostly three types of information:

- (a) **The covid-19 outcome:** are related to tests, hospitalizations, deaths and recoveries, for a particular region and for a particular time.
- (b) **Static covariate:** such as population size, health statistics, economic indicators, geographic boundaries.
- (c) **Dynamic covariate:** such as mobility, search trends, weather, and government interventions.

In this study, the research questions are related to investigating the variation of the metric R parameter as a function of the weather, mobility, and government response. Therefore, four different types of datasets are used that is drawn from multiple sources, as listed Table 1, and stored in separate tables as CSV files grouped by context, which can be easily merged due to the use of consistent geographic keys which are available in every files.

Here, the epidemiology dataset is the outcome datasets of Covid-19, which has been used to calculate R. The rest of the datasets, mobility, government response, and weather, have been used for investigating how they have an impact on metric R. However, each dataset (mobility, government response, and weather) has multiple variables, as shown in Table 2; those variables have been used during the experiment to know which variables have the increasing/decreasing impact on R. Since datasets are collected from different sources, it has the different beginning of data collection date (Table 1) and updating time. Each data source has its own update schedule based on the latest data published by its corresponding source.

4.2 Data pre-processing

The focus was put on the pre-processing of the datasets that correspond to four key European countries, which include France, Finland, Germany, and England.

No missing values were found for datasets epidemiology, mobility, weather, and government response; therefore, no action was needed. The categorical variables were turned into numerical variables (e.g., governmental response has three types of categorical values that were replaced with numerical values). After the scattered plot of datasets, some possibility of outliers were noticed in mobility and weather data. However, After manual checking of those outliers, we have decided to keep those since those outliers could be considered actual data.

5 Method of work

In this work, various statistical and machine learning-based regression models are applied to investigate the three research questions. Python programming language and appropriate data science libraries were used in the development

of the various techniques. The desired outcomes will bring value and reveal the association between the weather conditions, mobility attributes, government restriction, with the reproduction number R in four countries in Europe.

Different statistics techniques such as ANOVA tests and correlation matrices are used to determine each dataset's most correlated attributes with the R values. Then these attributes are used to examine their impact on R in different periods of the year. Furthermore, each dataset's best features or characteristics are applied in regression models to predict the R values. Moreover, different shifting periods were tested to provide better comparisons and expose the impact of these restrictions and factors on the pandemic reproduction number. Also, the period shifting will permit us to identify the appropriate incubation period of the Covid-19. Based on some related works[7][8][9], and health institutions[10][11], the incubation period was determined between 6 to 14 days. Consequently, based on these outcomes, we want to investigate multiple shifting periods to estimate or confirm the appropriate or most suitable incubation period. The selected shifting periods were: seven days lag, ten days lag, and 15 days lag. Figure 1 provides a high-level overview of the methodology.

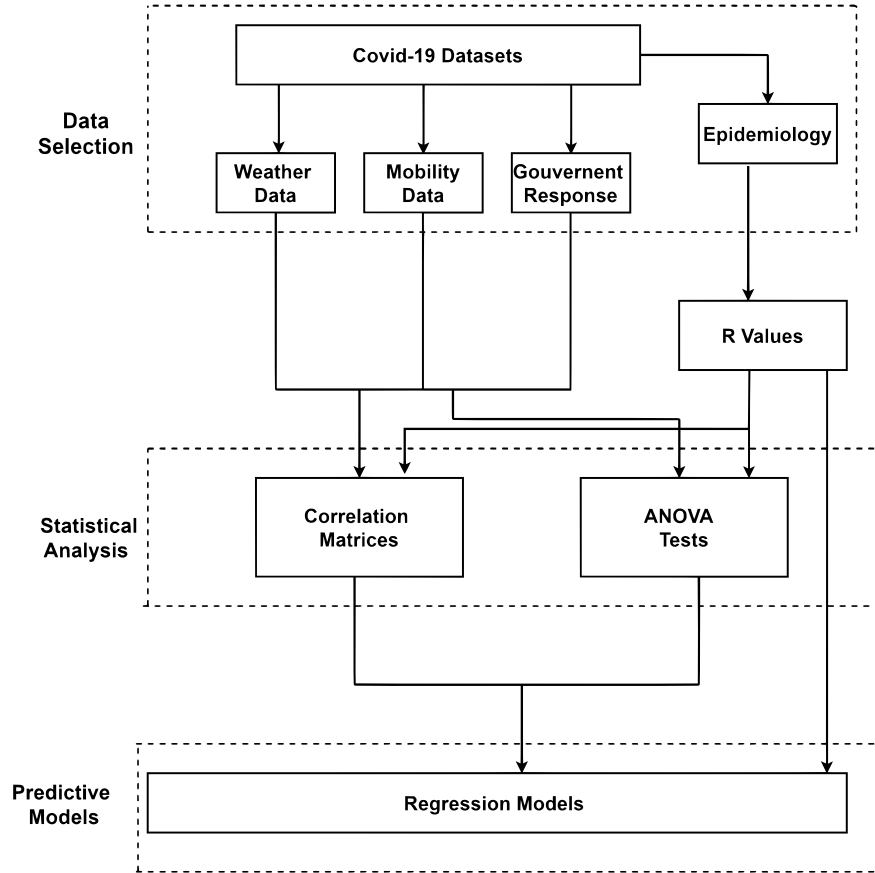


Figure 1: High level diagram of the methods

5.1 Reproduction number: R

The Reproduction (R) values are needed in order to perform the statistical and regression model. To estimates the effective reproduction number, we use the time series of reported case numbers of epidemics. To do this, we have followed the work done by *Huisman et al.*[12], *Cori et al.* [13]. And, the public python library Epyestim¹ is used for the calculations.

For demonstration purposes, Germany R values are visualized in Figure 2 and given as an example. The R values are presented with respect to the time index. The same outcomes are found for the remaining countries, France, Finland, and England.

¹<https://pypi.org/project/epystim/>

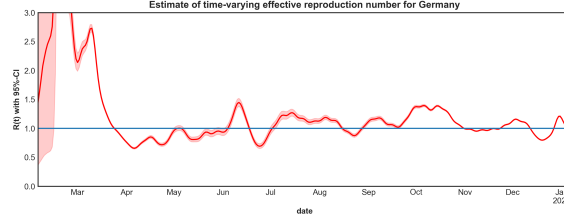


Figure 2: R values for Germany

5.2 Statistical techniques

After calculating the R values, we want to identify each domain's factors that impact the R values. For this reason, we used two statistical tests based on the data type of the factors in each domain. We applied the ANOVA test only on the governmental response since all the factors are categorical. However, we transformed the categorical data to numerical format using percentages to perform a correlation test. Let us have categorical data in the range of $[A, D]$. To compute each value's percentage, we divide the value index by the number of all categorical values. The percentages will be like this A:0.0, B:0.33, C:0.66, D:1.0.

If the p-values for a given factor are less than 0.05, this means that the factors have a statistically significant effect. The p-values in bold represent the factors that have a statistical significance with metric R.

The data type of the Weather and the Mobility domains is numerical. Therefore, the correlation test is a suitable method to identify the factors that correlate high with R values. Furthermore, the different shifting periods (0, 7, 10, 15) days lag are used.

5.3 Regressions models

Various predictive models are also used in order to predict the reproduction number R in the different countries using the four factors. The appropriate attributes are taken for each factor. The dataset is divided into a longer segment of 80% for the train, and the remaining segment of 20% is used in testing. Different regression models were applied with grid search. The models are as follows, Decision Tree(DT), Random Forest(FR), Support Vector Regressor (SVRF), XGBoost(XGB), GBoost(GB). Furthermore, period shifts were taken into account, where we have trained data with 0 days lag, seven days lag, ten days, and 15 days. Using different days lag aims to provide better comparisons with the data and reveal or identify the appropriate incubation period.

6 Results

In this section, the results, interpretations, and discussions about the research questions are considered. Three main subsections are included: a subsection for weather, the second one for mobility, and the last one about government restrictions.

6.1 Weather

The correlation matrices are used to identify the most important factors or the weather attributes that could impact the R values regarding the weather analysis.

A primary selection of the attributes was made for the weather by following the related studies' previous results[14]; we have selected the relative humidity, average temperature, and rainfall data as the more appropriate attributes to further investigations and interpretations. After Statistical analysis, regression models further deepen our analysis for R by developing predictive models to predict the R values. The weather data is investigated by following two methods. First, all the data is taken in bulk and analyzed with statistical analysis. In second, the weather seasons are taken into considerations by dividing the data into winter, spring, summer, and autumn.

6.1.1 First analysis: all-weather data

In this approach, we considered all the weather data and ignored the seasons' aspects of it. The goal is to comprehensively analyze the weather conditions and how they could broadly affect the R values by considering the weather as numbers only without further info about the characteristics and following the concept of 'let the data talks.'

Table 3: Weather data best correlation with R values.

Countries	Fr	Fi	DE	ENG
Days lag	15D	7D	7D	7D
Avg Temperature	6	6	14	5
Rainfall	6	4	18	16
Relative Humidity	4	22	7	6

Table 4: 7 days lag, weather seasons as a function of the R in %

Weather features	Fr	Fi	DE	ENG
Season	Spring	Spring	Spring	Spring
Average temp	38	37	47	57
Rain fall	32	8	21	21
relative humid	48	41	47	51
Regression model	SVR	DT	RF	RF
Model performance	70.85	82.63	60	69.05

Table 3 summarises the best correlation results obtained from the the analysis. All the values are presented in percentages in order to ease the interpretations, the more the percentage of correlation is high, the more the association or impact of the attribute to the R value. Based on the findings in the table, a shallow correlation or very low impact of the three weather attributes to the R value and this in all the four countries. The results have slightly higher values when the data is lagged 07 days, but the patterns are quite similar, and low in general. Another small reflection could be given to the relative humidity values in Finland where it has registered the highest correlation of 22%, but still low and does not reflect a real impact.

These results demonstrates a non-effectiveness of the weather condition towards the R values. But these results can not provide a good reflections of the effectiveness of the weather to the spread the Covid-19 virus and this due the non similarity between the weather data periods. For instance, the winter is characterise more with high rainfall with low temperatures, and summer is characterized with very high temperatures. So, analysing the weather seasons is very important and compulsory in order to come up with better conclusions and more understanding of the weather impact. Thus, in the following part, the season are explored and predictive models are also applied.

6.1.2 Second analysis: weather seasons

In this specific study, a correlation between the R number and the weather has been carried out. Correlation has been carried out between R and the weather as a function of the seasons for 2020 and 2021, including winter, spring, summer, and autumn. Therefore, a significant correlation has been observed only for the spring season for all four countries, as reported in Table 4. This reflects that the number of contamination increases in this season even though the weather features taken for the study, such as average temperature, rainfall, and relative humidity, are much more critical in the winter season than summer, autumn, or spring. One may wonder what the reason behind this strange observation is. One reason could be due to the fact that extreme government restrictions were taken into account during winter, and those measures were softened during the spring season. As for remark, to analyze the seasonality properly, one needs to have data from several years.

6.2 Mobility

In this section, a correlation between the R and mobility has been carried out. Correlation has been carried out between R and mobility as a function of the Retail recreation, Grocery pharmacy, Parks, Transit stations, and Residential. We have performed a correlation between 0, 7, 10, and 15 days of lag. However, Table 5, presented only the best results. Among all these four countries, R has shown the most negligible correlation with the mobility of Park than other mobility. We can observe R has shown a medium correlation with the grocery pharmacy and residential mobility. However, R has shown a high correlation with the mobility of transit stations and mobility retail recreation.

Five regression models (RF, GB, XGB, DT, and SVR) have been used with different days of lag (0, 7, 10, and 15) to predict R. Fifteen days lag performed best for France with the GB and Germany with the SVR model. However, seven days lag showed better performance for Finland and England with the RF model.

Table 5: correlation results of France, Finland, Germany and England mobility data with R.

Countries	Fr	Fi	DE	ENG
Days lag	15D	7D	15D	7D
Retail recreation	40	45	43	52
Grocery pharmacy	33	35	35	43
Parks	6	11	11	6
Transit stations	55	51	55	61
Workplaces	41	40	43	45
Residential	37	42	39	43
Regression Model	DT	RF	SVR	RF
Performance	84.74	90.1	87.2	89.7

Table 6: 10 days forecast with 7 days lag data of the mobility data for the months of September and December.

Country	Fr	Fi	DE	ENG
Month	September			
Model	SVR	SVR	SVR	DT
Performance	85.54	99.96	74.79	89
Month	December			
Model	DT	SVR	RF	XGB
Performance	97.3	78.98	84.23	72.95

Since for mobility, 7 days lag showed better results; therefore, we have rerun all the models with 7 days lag and predict the 10 days R values in which we selected two random months for September and December. Table 6 highlights the best results for the selected countries. The prediction of both cases showed decent accuracy (number of correct predictions total number of predictions). In September, the prediction accuracy was 99% for Finland, France 85.54%, England 89%, and Germany 74.7%. However, The predictions in December, France performed 97%, which is 11% better compared to September. We see the apparent difference between these two-month R prediction performances. For example, in September, Finland's prediction of R was 99.6%, but in December, it dropped to 78.9%. There could be many possible factors that may have impacted R. For example, Finland has the most challenging winter in December than September, which may have influenced people to move less in December.

6.3 Government restrictions

We employed two statistical analysis techniques for the government response: ANOVA and correlation tests to identify the most critical features used with various predictive models to predict the R values. Further, we adopt 10 days forecast to predict the R values in two different months.

6.3.1 ANOVA test vs Correlation

We performed an ANOVA test with the four countries in order to recognize the impacting factors used as a safeguard by governments. Table 7 summarises the outcomes of the ANOVA tests for the four countries.

Table 7: ANOVA Test outcomes.

Government Factors	Fr	Fi	DE	ENG
school_closing	4.58e-66	8.35e-53	0.017	1.32e-07
workplace_closing	3.52e-04	1.07e-21	0.863	1.17e-05
cancel_public_events	1.87e-05	1.98e-16	0.565	2.21e-03
restrictions_on_gatherings	3.13e-09	1.83e-04	0.036	4.30e-06
public_transport_closing	8.46e-01	0	0.600	5.92e-01
stay_at_home_requirements	5.12e-01	6.00e-02	0.558	4.68e-03
restrictions_on_internal_movement	9.71e-01	1.62e-04	0.967	1.41e-03
international_travel_controls	9.01e-03	7.65e-02	0.875	2.89e-02

We Notice that there is high variability in terms of the reported p-values. In France, the p-values of public transport closing, stay-at-home requirements, and restrictions on internal movement are more significant than 0.05. Therefore, there is no statistical significance that affects the R by those three factors. In Finland, the p-values of public transport closing, stay-at-home requirements, and international travel controls are also greater than 0.05; this means that the three factors do not impact R. However, the p-values of the ANOVA test in Germany show no significance between the factors and R. lastly, in England, all the elements are statistically significant with R except the factor of public transport closing. Overall, we can say that school closing, workplace closing, cancel public events, and gatherings restrictions have the highest correlation with the metric R. Hence, the governments use these restrictions when the R is too high.

6.3.2 Regression models

We used different machine learning algorithms to predict the entire governmental response dataset's R values for the four countries. Also, we extracted the best day lag with the best features using correlation analysis. Table 8 shows the best performance with the associated day lags.

Table 8: Best correlation values and predictive models for the four countries

Countries	FR	FI	DE	ENG
Day lags	7D	10D	7D	7D
school_closing	25	25	25	51
workplace_closing	54	44	65	62
cancel_public_events	42	17	74	81
restrictions_on_gatherings	84	02	59	80
public_transport_closing	08	16	16	79
stay_at_home_requirements	30	12	22	41
restrictions_on_internal_movement	29	04	46	24
international_travel_controls	76	09	27	31
Regression Model	SVR	GB	SVR	DT
Performance (1 - MAE)	88,01	90,69	82,47	89,47

The results show that seven days lag the most significant for France, Germany, and England. On the other hand, ten days lag In Finland is the most significant. It is also noticeable that the different models' performances were close for France, Finland, and Germany, whereas Germany's best-reported performance was far from the other countries. Overall, the seven-day lag is the most optimal for predicting the R values regarding all four governments' adopted factors.

Moreover, after we adopted the seven-day lag, we run again all the models to predict the R values in which we selected two random months for demonstration purposes. Table 9 highlights the best results for the specified countries.

Table 9: 10 days forecast with 7 days lag data of the government response for the months of September and December.

Country	Fr	Fi	DE	ENG
Month	September			
Model	SVR	DT	XGB	RF
Performance	98.13	96.58	98.48	86.3
Month	December			
Model	SVR	XGB	XGB	XGB
Performance	92.77	96.64	83.13	95.41

The predictions of the first month demonstrate the feasibility of the models in which the performances were the highest for France, Finland and Germany except for England. However, The predictions for the second month were relatively low compared with the first month. Therefore, the government response had a tremendous significance and satisfactory results towards the prediction of R. Hence. It maintained some control over the pandemic.

7 Conclusion & Discussions

In this project, the Covid-19 pandemic was investigated by analyzing the impact of the weather, mobility, and government restrictions on the reproduction number R of related to the Covid-19 pandemic. A dataset related to the pandemic was used, and appropriate statistical and regression models were used in order to reveal any association.

Regarding the weather research question, a significant correlation has been observed only for the spring season for all four countries, with shallow to shallow correlation values in the remaining seasons. These results are surprising, as one would expect to see more significance and correlation in the winter seasons. The reasons may be due to the government restrictions that have been observed to affect the R number and confuse the observation on the actual correlation between other factors, for instance, the weather seasons and the number of contamination per a specific period. However, with these outcomes, we observe that the weather attributes in some environments, especially relative humidity, impact the spread of the COVID-19 pandemic when the government restrictions are softened.

Regarding mobility research Q2, we wanted to know what kind of mobility has an impact on R. Our findings have shown that different types of mobility have a distinct correlation towards R. For example, the mobility of transit stations and the mobility of retail recreation has shown a high correlation with R. One possible explanation for this could be that people may have moved from one place to another during pandemic times, spreading more viruses. Therefore, R has shown a high correlation with the mobility of transit stations. Other-hand the mobility of parks had the lowest correlation with R. People may not have used parks much during the pandemic time, spreading lesser viruses through park mobility.

Our study has found a meaningful relation between the different attributes of mobility and R. Furthermore, our predicted model has successfully predicted 10 days of future R with 74-99% accuracy, promotes the feasibility of using mobility data for forecasting the Covid-19 situation. This study may also suggest taking more precautions while moving places like transit stations and visiting recreation places since those kinds of mobility have shown the highest correlation with R.

For the research question about the effect of governmental response, the outcomes vary from one country to another because of other hidden factors that were not considered. Overall, the study of governmental response's impact shows that there are common vital factors crucial for lowering the R-value, such as school closing, workplace closing, and cancel public events. On the other hand, other factors do not affect the R. The predictions estimated suggest that there is a space to ease restrictions, provide more social interactions and maintain control of the pandemic.

The study outcomes bring values and exciting findings that could help understand the pandemic and what factors could improve the situation and responses. For instance, the ten days forecast models are helpful to monitor the Covid-19 situations monthly by showing how efficient the restrictions made in the first 20 days and their impact in the following ten days. In comparison, the other regression models were more global and were trained in longer segments. Still, the findings provide some differentiation of the effectiveness of each condition or measure towards the pandemic, and analyzing them from the proper perspective will help suggest new measures or maintain the already existing ones.

Overall, the results from the three questions were overlapping with other works found in the literature in terms of the statistical answers, model performance, and best incubation periods of the Covid-19.

Acknowledgments

We would like to thank Eija Haapalainen for her kind help and support.

References

- [1] Ramadhan Tosepu, Joko Gunawan, Devi Savitri Effendy, Hariati Lestari, Hartati Bahar, Pitrah Asfian, et al. Correlation between weather and covid-19 pandemic in jakarta, indonesia. *Science of The Total Environment*, 725:138436, 2020.
- [2] Zohair Malki, El-Sayed Atlam, Aboul Ella Hassanien, Guesh Dagnew, Mostafa A Elhosseini, and Ibrahim Gad. Association between weather data and covid-19 pandemic predicting mortality rate: Machine learning approaches. *Chaos, Solitons & Fractals*, 138:110137, 2020.
- [3] Mehmet Şahin. Impact of weather on covid-19 pandemic in turkey. *Science of The Total Environment*, 728:138810, 2020.
- [4] Renato HL Pedrosa. The dynamics of covid-19: weather, demographics and infection timeline. *MedRxiv*, 2020.
- [5] Ran Xu, Hazhir Rahmandad, Marichi Gupta, Catherine DiGennaro, Navid Ghaffarzadegan, Heresh Amini, and Mohammad S Jalali. The modest impact of weather and air pollution on covid-19 transmission. *MedRxiv*, 2020.
- [6] O. Wahltinez et al. Covid-19 open-data: curating a fine-grained, global-scale data repository for sars-cov-2. 2020. Work in progress.
- [7] Robert B Noland. Mobility and the effective reproduction rate of covid-19. *Journal of Transport & Health*, 20:101016, 2021.

- [8] Stephen A Lauer, Kyra H Grantz, Qifang Bi, Forrest K Jones, Qulu Zheng, Hannah R Meredith, Andrew S Azman, Nicholas G Reich, and Justin Lessler. The incubation period of coronavirus disease 2019 (covid-19) from publicly reported confirmed cases: estimation and application. *Annals of internal medicine*, 172(9):577–582, 2020.
- [9] Nazar Zaki and Elfadil Abdalla Mohamed. The estimations of the covid-19 incubation period: a systematic review of the literature. *medRxiv*, 2020.
- [10] Finnish institute for health and welfare. transmission and incubation period of coronavirus. <https://thl.fi/en/web/infectious-diseases-and-vaccinations/what-s-new/coronavirus-covid-19-latest-updates/>.
- [11] World health organization. coronavirus disease 2019 (covid-19) situation report – 73. <https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200402-sitrep-73-covid-19.pdf>.
- [12] Jana S Huisman, Jeremie Scire, Daniel C Angst, Richard A Neher, Sebastian Bonhoeffer, and Tanja Stadler. Estimation and worldwide monitoring of the effective reproductive number of sars-cov-2. *medrxiv*, 2020.
- [13] Anne Cori, Neil M Ferguson, Christophe Fraser, and Simon Cauchemez. A new framework and software to estimate time-varying reproduction numbers during epidemics. *American journal of epidemiology*, 178(9):1505–1512, 2013.
- [14] The centre for evidence-based medicine, university of oxford. environmental weather conditions and influence on transmission of sars-cov-2. <https://www.cebm.net/covid-19/weather-conditions-sars-cov-2/>.