Medical Research Archives

European Society of Medicine

Published: April 30, 2024

Citation: Morales-Fajardo HM, Rodríguez-Arce J, et al., 2024. Simplifying COVID-19 Data Analytics for Efficient Pandemic Management: A Novel Approach, Medical Research Archives, [online] 12(4). https://doi.org/10.18103/mra.v12i4.5269

DOI
https://doi.org/10.18103/mra.v12i4.5269

ISSN: 2375-1924

RESEARCH ARTICLE

Simplifying COVID-19 Data Analytics for Efficient Pandemic Management: A Novel Approach

H. M. Morales-Fajardo, MSc[a], J. Rodríguez-Arce, PhD[a,b,c,*], B. E. Ruvalcaba-Ramos, PhD[d], S. Montes de Oca, PhD[c]

[a] School of Engineering, Universidad Autónoma del Estado de México, México
[b] School of Medicine, Universidad Autónoma del Estado de México, México
[c] Tecnologico de Monterrey, School of Engineering and Sciences, México
[d] Institute of Neurosciences, CUCBA, Universidad de Guadalajara, México

***Corresponding author:** jrodrigueza@uaemex.mx

ABSTRACT

This study presents a streamlined approach to pandemic management by simplifying COVID-19 data analytics. It focuses on the significant role of mobility patterns in forecasting case trajectories. Utilizing open mobility data from Google and Apple, a novel predictive model is proposed that aids health authorities in scenario projection and case monitoring. This model facilitates informed decision-making with minimal economic impact during future outbreaks.

Key findings highlight the profound link between mobility changes and COVID-19 case trends, emphasizing the necessity of integrating mobility data into predictive models. The model employing linear and polynomial regression analyses and incorporating the effective reproduction number, R_t, and the influence mobility changes have on population forecasts can be extended up to 90 days.

The study acknowledges limitations, particularly the reliance on mobility data that does not fully encompass all variables affecting virus transmission. Moreover, it explores the mental health implications of mobility restrictions, suggesting a broader impact of pandemic management strategies.

The proposed model is a practical tool for managing pandemics through mobility data analysis, underscoring the need for comprehensive studies on the broader effects of mobility changes to guide public health policies.

Keywords: COVID-19, Mobility Data, Predictive Modeling, Data Analytics, Public Health Policy

1. Introduction

The novel coronavirus, also known as COVID-19, first emerged in late 2019 and quickly garnered global attention. Initially detected in Wuhan, China, the virus rapidly evolved into a global pandemic characterized by symptoms ranging from mild to severe respiratory illness. Recognized symptoms included fever, chills, muscle pain, headache, sore throat, cough, and shortness of breath, with symptom onset occurring up to 14 days post-infection. The virus's rapid spread and diverse clinical presentations necessitated urgent international response and research efforts to understand and mitigate its impact [1, 2, 3, 4].

The World Health Organization (WHO), noting the escalating case numbers and the virus's transnational transmission, declared COVID-19 a Public Health Emergency of International Concern by early 2020. In response, nations worldwide implemented various public health measures, including social distancing, travel restrictions, and quarantine protocols, to curb the virus's spread and prevent healthcare system overloads [5, 6, 1, 7, 8, 9, 10]. Despite these efforts, the pandemic's dynamic nature and socioeconomic repercussions have highlighted the urgent need for adaptable and data-driven strategies to effectively manage such crises.

1.1. PREDICTIVE MODELS OF COVID-19 OUTBREAK

Predictive modeling has emerged as a pivotal tool in navigating the COVID-19 pandemic, offering insights into potential case trajectories and healthcare demands. Various modeling approaches, including compartmental models like SEIR (Susceptible, Exposed, Infectious, Recovered) and statistical time series analyses, have been employed to forecast infection dynamics and inform public health interventions. Notably, adaptations of traditional models to include quarantine effects and mobility data have shown promise in capturing the complex interplays of human behavior and disease spread [11, 12].

Despite advancements, challenges persist in predictive accuracy and applicability across diverse contexts. For instance, while models like the modified SEIR have provided valuable projections under varying containment levels, their dependence on extensive and accurate data limits their utility in real-time decision-making Furthermore, incorporating mobility data, as demonstrated by [15], by leveraging mobile phone datasets, presents an innovative avenue to enhance the relevance of models by directly linking human movement patterns with transmission dynamics.

Recent efforts have also explored machine learning (ML) algorithms for COVID-19 mortality prediction, utilizing patient data to identify high-risk individuals. Such approaches underscore the potential of ML in enhancing predictive precision, albeit contingent upon the availability of comprehensive datasets [16]. Further, spatiotemporal modeling techniques, like those based on the Hawkes process, offer nuanced insights into regional transmission risks, facilitating targeted intervention strategies [17].

This evolving landscape of COVID-19 modeling underscores a critical gap: the need for simplified yet robust analytical frameworks that can provide reliable insights with minimal data requirements. Addressing this gap, our study aims to harness publicly available mobility data to develop a predictive model that balances simplicity with analytical depth, offering a pragmatic tool for pandemic management. By integrating mobility trends with critical epidemiological parameters, the aim is to contribute a novel approach to understanding and mitigating the spread of COVID-19, thereby assisting decision-makers in navigating the challenges of pandemic response with greater agility and foresight.

1.2. CURRENT STUDY

This work proposes a novel predictive model that leverages reduced data complexity without compromising the reliability of its forecasts, a critical factor for effective pandemic management. The model aims to elucidate the relationship between mobility patterns and COVID-19 transmission rates, offering a streamlined tool for assessing the potential impacts of public health interventions. By focusing on mobility data from Google and Apple, a practical metric is provided for gauging the effectiveness of social distancing measures and their implications on disease spread. This approach offers health authorities a simplified yet powerful analytical tool to inform timely and proportionate responses to the evolving pandemic landscape.

Recognizing the limitations of existing models, this study introduces a novel approach to simplifying COVID-19 data analytics. By leveraging publicly available mobility data from platforms like Google and Apple [18, 19], this research aims to develop a more accessible and efficient predictive model. The proposed model focuses on understanding and forecasting COVID-19 case trends based on mobility patterns, examining the correlation between changes in mobility and case numbers in two distinct locations over 30 days. Extending predictions up to 90 days, the model employs linear and polynomial regression analyses. It incorporates the primary reproduction number R0 to enhance its predictive capability and a variable considering the mobility changes m_0.

1.3 FOUNDATIONAL CONCEPTS

In managing infectious diseases, various parameters are crucial for countries striving to control outbreaks. Among these, vaccination stands out as a primary strategy, possibly reducing the number of cases to manageable levels, ideally eliminating them. However, in situations involving novel pathogens, such as the SARS-CoV-2 virus responsible for COVID-19, the absence of an existing vaccination presents significant challenges for public health officials aiming to maintain control over case numbers. In these instances, the reliance shifts metric is a fundamental measure of an infectious disease's potential to spread, representing the average number of secondary infections generated by a single infected individual [20, 21, 22].

It is critical to recognize that the R number is not a static value; somewhat, it is influenced by a multitude of factors, including the inherent infectiousness of the disease, its progression over time, population mobility, containment measures, and levels of immunity, whether from previous infection or vaccination [23].

In the study of infectious diseases, a specific parameter known as R_0, or the basic reproduction number, is crucial during the initial stages of an outbreak. This number represents how many cases one infected person will likely cause in a completely susceptible population. As the disease progresses and interventions such as social distancing and vaccinations are implemented, another variant of R, known as the effective reproduction number or R_t becomes more relevant. This reflects the actual transmission rate after considering the effects of immunity and interventions within the population. Ideally, with effective control measures, R_t would be reduced to below 1.0, signifying a controlled spread of the disease [25,23,20].

In other words, R_0 measures the transmission potential of a disease in a completely susceptible population without any control measures. It estimates how quickly a disease could spread without interventions or pre-existing immunity. On the other hand, R_t represents the average number of new infections caused by an infected case at a specific time, considering current control measures and changes in population immunity. R_t is used to assess the effectiveness of public health interventions in real-time and to adjust policies as needed.

R_t number for SARS-CoV-2 exhibits significant variability across different regions, influenced by population behavior and density factors. In urban areas, where interactions and contact rates are higher, R_t values are expected to be greater than those in rural settings. Achieving control over infectious diseases like COVID-19 necessitates reducing the R_t below 1.0. An R_t-*value* of 1.0 suggests a stable state of the outbreak, where the number of new cases is constant over time. However, an R_t greater than 1.0 indicates a growing outbreak, potentially significantly burdening public health systems, especially if the pathogen is highly lethal. Figure 1 illustrates the dynamic behavior of an infectious agent, starting with only 10 cases at day 0 with different values of R_0 (without vaccination).

The World Health Organization (WHO) emphasizes social distancing as a primary strategy to mitigate the spread of novel viruses such as SARS-CoV-2. This involves maintaining a safe distance between individuals and avoiding crowded places, particularly public spaces [26]. Governments may implement various measures to enforce social distancing, from voluntary compliance encouraged by public campaigns to more stringent enforcement by authorities.

According to WHO, there are four phases in a pandemic: Interpandemic phase, Alert phase, Pandemic phase, and Transition phase, whereas the US Centers for Disease Control and Prevention (CDC) refers to 6 intervals: Investigation, Recognition, Initiation, Acceleration, Deceleration, and Preparation. Figure 2 describes the pandemic curve of any infectious agent and both WHO and CDC phases or intervals. In addition, the Community Mitigation Guidelines to Prevent Pandemic Influenza [27] recommends conducting a risk assessment to understand the deepness of measures such as the closure of schools or workplaces and the reduction of people's mobility in public spaces.

Table 1 shows the initial assessment regarding transmissibility and clinical severity. R_0, more significant than 1.8, means a moderate to high risk and a case fatality rate greater than 10%. Hence, predictive models are essential for health officials to make decisions to keep R_0 at low values. For a specific country or location, future case growth can be expressed as a function of its population mobility parameter called m_0 and its basic reproduction number R_0.

Table 2 shows different values of R_0 and their impact on future case growth, with only 10 cases on day 1, assuming R_0 is unaffected. According to pandemic mitigation, this action would be referred to as a *scenario with no intervention* [30].

Equation 1 shows the case growth based on R_0:

$$C_i = C_{i-1} * R_0 \qquad (1)$$

where:

C_i is the number of cases in interval i with $i > 0$.

R_0 is the basic reproduction number.

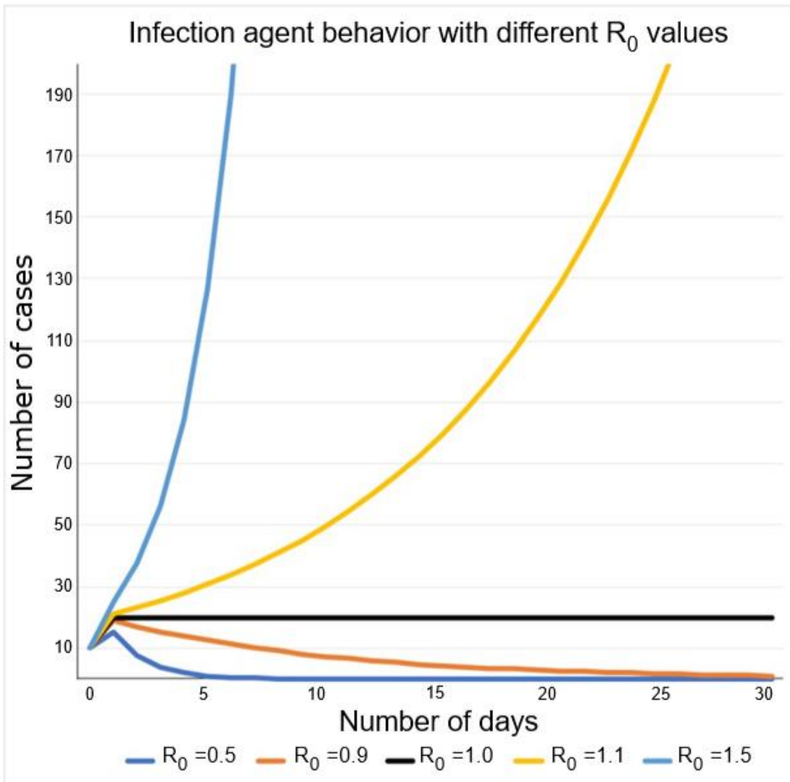

Figure 1: An infection agent such as SARS-COV-2 has a different behavior while reproduction number R_0 varies

Balancing public health concerns with a nation's socioeconomic well-being is crucial, especially during outbreak conditions where social distancing remains the primary intervention in the absence of treatment.

The pandemic is characterized by distinct phases outlined by WHO and the US Centers for Disease Control and Prevention (CDC), each requiring specific public health responses. Understanding the depth of interventions necessary, such as school or workplace closures and mobility restrictions, is essential for effective pandemic management [27].

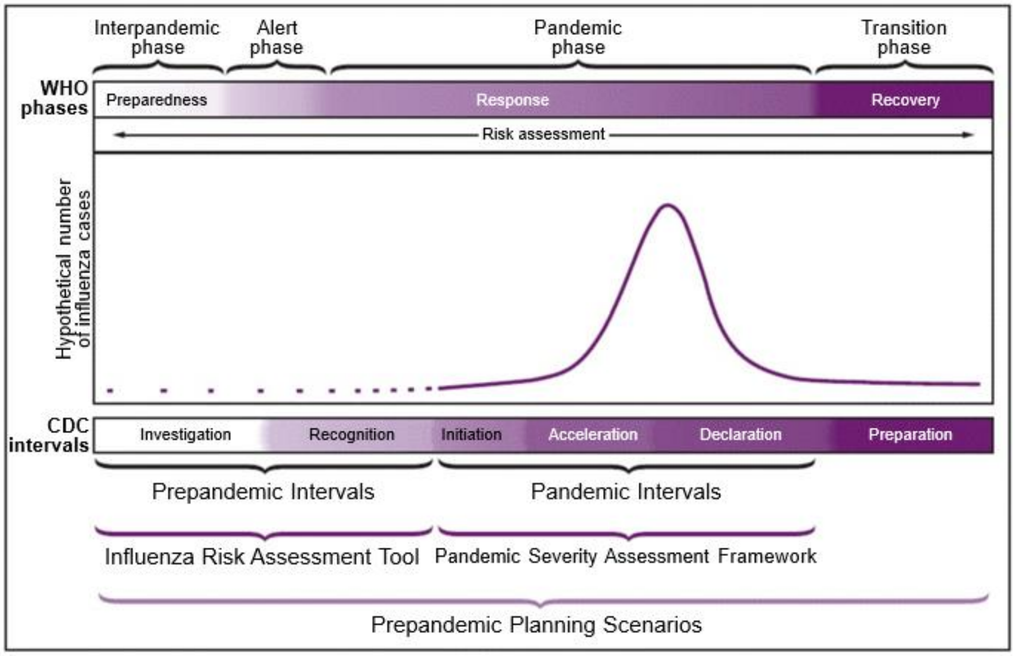

Figure 2: The pandemic curve of any infectious agent and both WHO and CDC phases or intervals, adapted from [28].

2. Methods

As explained in section 1.3, the main goal of this research is to find how mobility impacts future case growth and, eventually, be able to use mobility parameter m_0 as the controlling factor to reduce novel virus transmission, such as SARS-COV-2, in the form of case future growth reduction, given that there could not be any vaccination available.

In general terms, case growth can be represented as a function of the number of cases in the period i (Ci) and its basic reproduction number (R_0), as described in Equation 1. If $i = 1$, then C_0 is the initial case of the pandemic. This study is the first documented case, indicating the first day of the epidemic curve.

When R_0 is more significant than 1.0, and there is no intervention, the number of cases tends to grow exponentially until the outbreak reaches its natural peak [30].

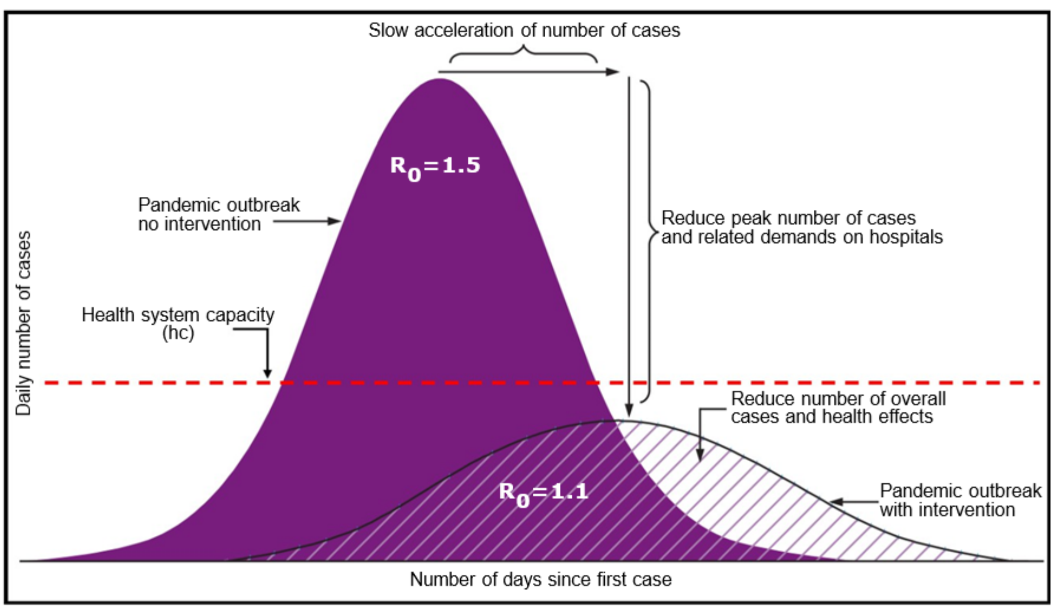

Figure 3: Comparison between pandemic outbreak with intervention and no intervention, adapted from [30].

Herein, the importance of R_0 is to be controlled with some intervention mechanism. In this case, the use of m_0 is proposed, and for that matter, the use of mobility data provided by mobile applications such as Google and Apple.

In this study, R_t is introduced as the temporal reproduction number, which is a rolling average value based on known R values. Consequently, R_0 can be represented as a function of m_0 to smooth the exponential growth of daily cases based on known R values. Equation 2 shows this expression:

$$R_0 = R_t + m_0 \qquad (2)$$

Table1: Initial assessment: scaled measures of influenza virus transmissibility and clinical severity (adapted from [29]).

Measures of transmissibility and clinical severity	Scale	
	Low to moderate	Moderate to high
Transmissibility		
Secondary attack rate, household	≤20%	>20%
Attack rate, school or university	≤30%	>30%
Attack rate, workplace or community	≤20%	>20%
R_0, basic reproductive number	1.0-1.7	≥1.8
Underlying population immunity	Some underlying population immunity	Little to no underlying population immunity
Emergency department or other outpatient visits for influenza-like illness	<10%	≥10%
Virologic characterization	Generic markers for transmissibility absent	Genetic markers for transmissibility present
Animal models, transmission	Less efficient or similar to seasonal influenza	More efficient than seasonal influenza
Clinical severity		
Upper bound of case-fatality ratio	<1%	≥1%
Upper bound of case-hospitalization ratio	<10%	≥10%
Deaths-hospitalizations ratio	<10%	≥10%
Virologic characterization	Genetic markers for virulence absent	Genetic markers for virulence present
Animal models, evaluation of morbidity and mortality	Less virulent or similar to seasonal influenza	More virulent than seasonal influenza

Where:

$$m_0 < 1$$

to make

$$R \leq 1$$

Consequently, equation 1 can be rewritten as:

$$C_i = C_i - 1 * (R_t + m_0)$$

R_0 and m_0 parameters are calculated in subsequent sections.

2.1. DATA SETS

The data sets for COVID-19 cases were sourced from the Journal "Our World in Data"[31]. Mobility data was obtained from public mobility data sets provided by Google and Apple [18, 19]. This analysis leverages daily COVID-19 cases and mobility changes across two distinct locations. The mobility data encapsulates variations in population movement, with positive values signalling enhanced mobility and negative values indicating reduced movement.

Acknowledging the potential biases inherent in utilizing mobility data derived exclusively from Google and Apple platforms [18, 19]. Such data may not represent all demographic segments uniformly, possibly skewing the analysis. This limitation merits consideration, and the study could benefit from discussing the implications of these biases. Further, exploring methodologies to mitigate or recognize these biases' impact would fortify the research's integrity and applicability.

Table 2: Case growth of any given infectious agent with A) R_0=1.1 B) R_0=1.5 and C) R_0=1.8.

Day	Daily cases	R_0	Day	Daily cases	R_0	Day	Daily cases	R_0
1	10	1.10	1	10	1.50	1	10	1.80
2	11	1.10	2	15	1.50	2	18	1.80
3	12	1.10	3	23	1.50	3	32	1.80
4	13	1.10	4	34	1.50	4	58	1.80
5	15	1.10	5	51	1.50	5	105	1.80
6	16	1.10	6	76	1.50	6	189	1.80
7	18	1.10	7	114	1.50	7	340	1.80
8	19	1.10	8	171	1.50	8	612	1.80
9	21	1.10	9	256	1.50	9	1102	1.80
10	24	1.10	10	384	1.50	10	1984	1.80
11	26	1.10	11	577	1.50	11	3570	1.80
12	29	1.10	12	865	1.50	12	6427	1.80
13	31	1.10	13	1297	1.50	13	11568	1.80
14	35	1.10	14	1946	1.50	14	20823	1.80
15	39		15	2919		15	37481	1.80

3. Results

Table 3, presents two locations, designated as Location A and Location B, showcasing their observed cases and mobility shifts over the initial 30-day span. Subsequent visual depictions, as illustrated in Figure 4, afford an analytical perspective on the COVID-19 case evolution and mobility fluctuations within these locales across the same timeframe:

COVID-19 Cases: The first plot (top left) illustrates the number of COVID-19 cases in both locations. Cases in Location A show a more pronounced increase compared to the relatively stable number of cases in Location B until a significant rise towards the end of the period.

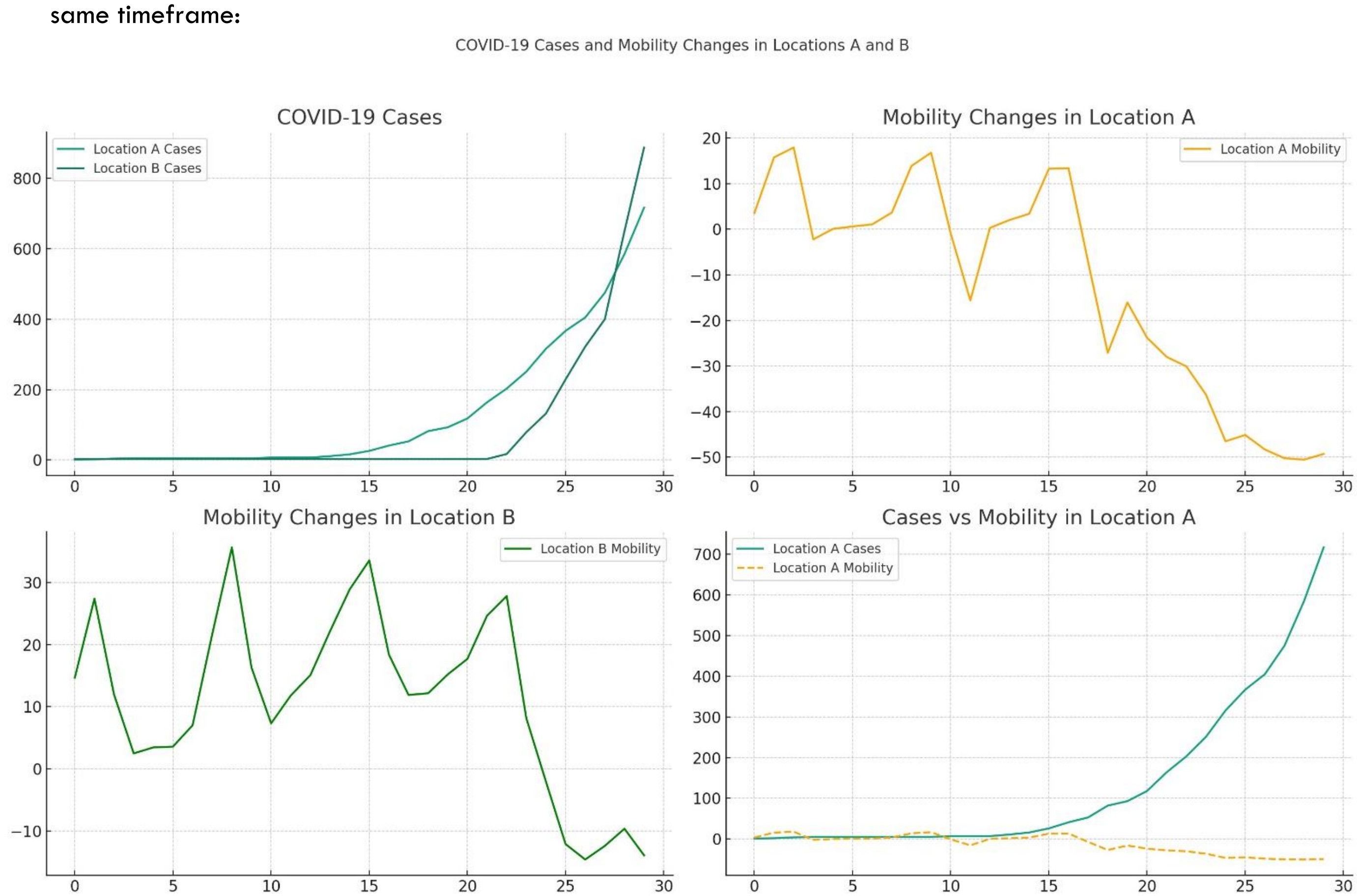

Figure 4: COVID-19 Cases and mobility changes for locations A and B.

Mobility changes in Location A: The second plot (top right) indicates the mobility changes in Location A, showing fluctuations with both positive and negative values. This suggests variations in how much the population moved around, with increased and decreased mobility periods.

Mobility changes in Location B: The third plot (bottom left) shows the mobility changes in Location B, which also fluctuates but remains mostly positive until a decline towards the end of the period. This indicates a reduction in mobility, possibly in response to increasing cases or other factors.

• **Cases vs mobility in Location A:** The final plot (bottom right) compares the number of cases and mobility changes in Location A. It illustrates how mobility trends do not directly correlate with the immediate rise in cases, suggesting that the impact of mobility on cases might be delayed or influenced by other factors.

Overall, these visualizations underscore the complexity of the relationship between mobility and COVID-19 case numbers, with Location A experiencing a more dramatic rise in cases and both locations showing varied patterns of mobility change.

3.1. PROCEDURE

In the proposed methodology, linear and polynomial regression analyses are employed to scrutinize n-day trends and compare predictions yielded by these models, which align to the theoretical expression described in equation 3. The core objective of this research is to ascertain the minimal duration required to accurately forecast case trajectories and mobility alterations, thereby furnishing actionable insights for public health policy formulation aimed at either pre-emptive preparation or mitigation of spread. A deliberate emphasis on simplicity, implementation feasibility, and a track record of reliability underpinned the selection of linear and polynomial regression models. These models are distinguished by their direct applicability and are pivotal in rapidly unfolding pandemic scenarios where expeditious decision-making is paramount. Furthermore, existing epidemiological modeling literature substantiates these regression techniques 'efficacy in generating dependable forecasts, even when constrained by limited data sets and computational resources, resonating with the current study's specific limitations and objectives.

Additionally, R_0 values were calculated for both locations, using it to infer the potential spread of the virus as a function of mobility changes and to provide a set of equations that can help to calculate the mobility change rates and predict cases effectively with early data.

For the 15-day use case, the regression equations describing the behavior of COVID-19 cases as a function of mobility changes for each location are as follows:

Location A: Linear Regression Equation:

$$Y = -0.1122x + 6.4504 \quad (4)$$

Location A: Polynomial Regression Equation:

$$Y = -0.0056x^2 - 0.0713x + 6.7845 \quad (5)$$

Location B: Linear Regression Equation:

$$Y = 3 \quad (6)$$

Location B: Polynomial Regression Equation:

$$Y = 3 \quad (7)$$

Table 3: Initial dataset, location A and B, cases and mobility changes

Location A Cases	Location B Cases	Location A Mobility Changes	Location A Mobility Changes
1	3	3.58	14.66
2	3	15.75	27.38
4	3	17.92	11.94
5	3	-2.22	2.50
5	3	0.06	3.46
5	3	0.61	3.59
5	3	1.05	7.02
5	3	3.64	21.65
5	3	13.88	35.62
5	3	16.77	16.31
7	3	-0.94	7.32
7	3	-15.62	11.80
7	3	0.29	15.07
11	3	2.02	22.13
16	3	3.38	28.85
26	3	13.24	33.51
14	3	13.39	18.39

Location A Cases	Location B Cases	Location A Mobility Changes	Location A Mobility Changes
53	3	-7.20	11.89
82	3	-27.12	12.16
93	3	-16.05	15.20
118	3	-23.78	17.69
164	3	-28.00	24.63
203	17	-30.06	27.79
251	79	-36.25	8.16
316	132	-46.50	-2.02
367	229	-45.12	-12.07
405	322	-48.29	-14.56
475	400	-50.22	-12.36
585	650	-50.53	-9.60
717	888	-49.27	-13.88

Figure 5 illustrates the actual and predicted cases of COVID-19 using linear and polynomial regression models for the first 15 days. For Location A, both models attempt to capture the trend of cases with slight differences in fit, indicated by the curvature in the polynomial regression. For Location B, due to the stable number of cases during the first 15 days, both the linear and polynomial models essentially predict a constant value, reflecting the limited variation in case numbers during this period.

For the 20-day use case, the regression equations describing the behavior of COVID-19 cases as a function of mobility changes for each location are as follows:

Location A: Linear Regression Equation:

$$Y = -1.3739x + 21.7562 \quad (8)$$

Location A: Polynomial Regression Equation:

$$Y = 0.0623x^2 - 1.1092x + 12.7433 \quad (9)$$

Location B: Linear Regression Equation:

$$Y = 3 \quad (10)$$

Location B: Polynomial Regression Equation:

$$Y = 3 \quad 11)$$

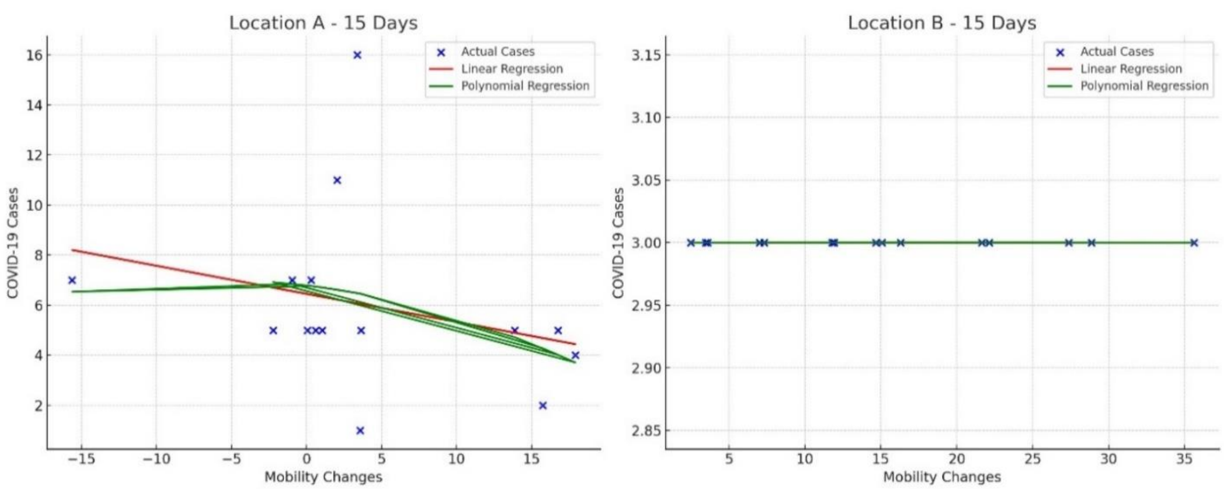

Figure 5: Prediction of cases using 15 days.

Figure 6 for the 20-day use case shows a distinct pattern, especially for Location A, where both linear and polynomial regression models indicate a more complex relationship between mobility changes and COVID-19 cases. The polynomial model captures a nonlinear trend that suggests a varied impact of mobility on case numbers. For Location B, the models remain unchanged, predicting a constant value, reflecting the stable number of cases in the first 20 days without much variation.

For the 30-day use case, the regression equations for both locations are as follows, illustrating the behavior of COVID-19 cases as a function of mobility changes with more comprehensive data:

Location A: Linear Regression Equation:

$$Y = -7.3555x + 41.7586 \quad (12)$$

Location A: Polynomial Regression Equation:

$$Y = 0.1871x^2 - 0.8755x - 4.8204 \quad (13)$$

Location B: Linear Regression Equation:

$$Y = -10.4420x + 209.1123 \quad (14)$$

Location B: Polynomial Regression Equation:

$$Y = 0.4665x^2 - 18.8234x + 155.1477 \quad (15)$$

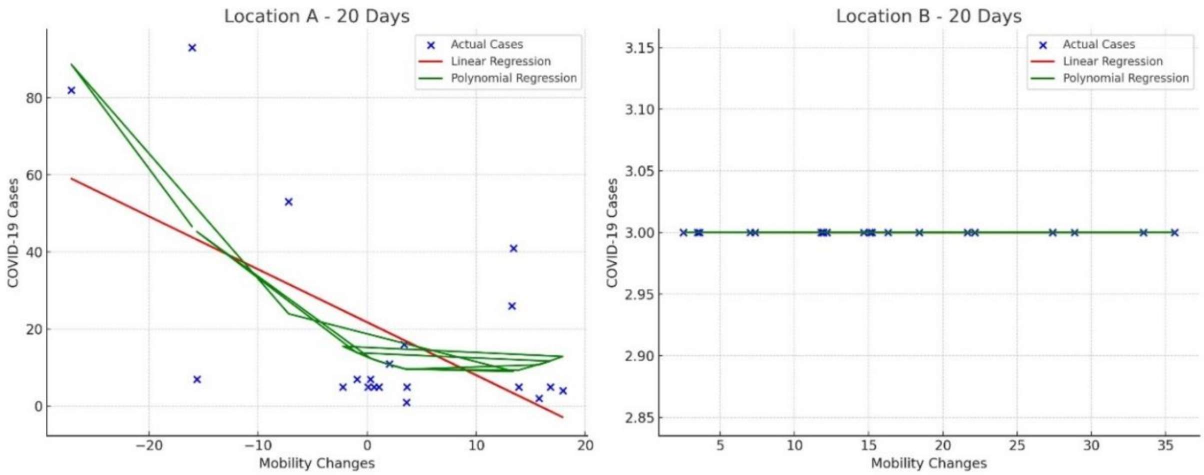

Figure 6: Prediction of cases using 20 days.

Figure 7 for the 30-day use case shows a pronounced differentiation between the linear and polynomial models, especially for Location B, which now incorporates the late surge in cases into its prediction model. This reveals critical insight into how mobility changes have significantly predicted COVID-19 cases as the data set increases.

For Location A, the polynomial model suggests a nonlinear relationship with a more significant variance in cases as a function of mobility changes, capturing the complex dynamics of case spread concerning mobility.

The linear and polynomial equations for Location B indicate a relationship between mobility changes and COVID-19 cases, moving away from the constant prediction seen in earlier use cases. This change reflects the increase in cases and suggests that mobility changes have become a more critical factor in predicting COVID-19 cases as the situation evolved.

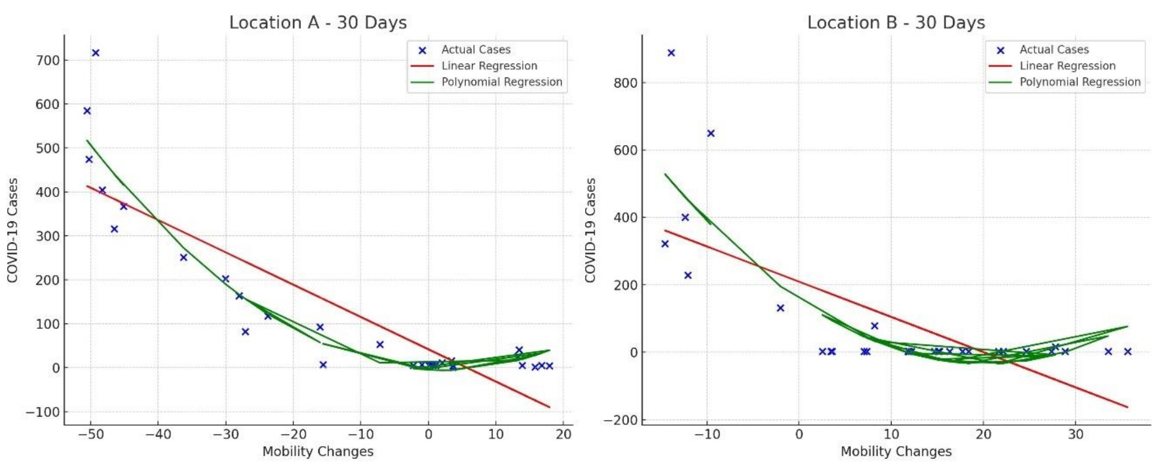

Figure 7: Prediction of cases using 30 days.

4. Discussion

4.1. REGRESSION ANALYSIS

The polynomial regression model offered detailed insight into the dynamics between mobility changes and COVID-19 cases. In Location A, it revealed a broader range of case fluctuations with mobility, highlighting the intricate spread patterns. Meanwhile, Location B accurately reflected the late surge in cases, indicating a heightened responsiveness to mobility changes.

Models were analyzed over 15, 20, and 30-day periods to identify the most effective one for depicting a decline in COVID-19 cases due to mobility changes.

Each model's fit to the observed data and capacity to represent complex, nonlinear interactions indicative of real-world phenomena were considered. The analysis underscored the significance of nonlinear trends, as demonstrated by polynomial regressions, in understanding the COVID-19 outbreak.

15-Day use case. The linear and polynomial models for Location A attempted to capture a downward trend with minimal data complexity. For Location B, the models predicted a constant value, reflecting stable case numbers and not capturing any decrease.

20-Day use case. The linear model for Location A showed a more pronounced negative slope, indicating a better capture of decreasing cases as mobility changes. The polynomial model indicated a nuanced understanding of the relationship, suggesting some non-linearity. Location B's models remained constant but did not effectively capture decreases.

30-Day use case. The linear regression models for both locations showed significant negative slopes, indicating a strong relationship between decreased mobility and decreased cases. The polynomial models for both locations captured more complex, non-linear trends, suggesting a detailed relationship between mobility changes and case numbers.

4.2. MOST ACCURATE USE CASE FOR DESCRIBING DECREASE IN CASES

The 30-day use case provides the most accurate and nuanced understanding of how mobility changes affect COVID-19 case numbers considering the following reasons:

- **Data completeness:** The 30-day period provides a fuller dataset, capturing more of the pandemic's dynamics, including rises and falls in case numbers.
- **Model complexity and fit:** The polynomial models, especially for the 30-day use case, indicate a better fit for the complex relationship between mobility and cases. This complexity is essential for accurately describing
- decreases in cases, as simple linear models may not capture the full scope of how behavioral changes impact pandemic trends.
- **Reflecting real-world trends:** The 30-day models, particularly with the polynomial regression, are better suited to reflect the non-linear dynamics observed in real-world data, where the relationship between mobility and case numbers can be influenced by various factors, including policy changes, population compliance with mobility restrictions, and the virus's natural spread.

Thus, the 30-day use case, with its polynomial regression model, offers the most accurate framework for predicting how changes in mobility could decrease COVID-19 cases. It captures both mobility's direct and indirect influences on pandemic trends.

Figure 8 shows the actual vs. predicted COVID-19 cases for Location A and Location B over an extended period of up to 90 days, based on the polynomial regression model developed from the initial 30-day data.

For Location A, the actual cases are plotted in blue, and the predicted cases based on mobility changes are plotted in red with a dashed line. The model predicts an increasing trend in cases following mobility changes. It's important to note that these predictions are extrapolations based on the model and the mean mobility change observed towards the end of the 30 days.

Similarly, the actual cases are shown in blue for Location B, and the predicted cases are in red with a dashed line. The model also predicts an increase in cases based on the trend in mobility changes.

These plots illustrate how changes in mobility might influence the predicted number of COVID-19 cases in both locations over time, according to the polynomial regression model. However, it's crucial to remember that these predictions are highly dependent on the accuracy and representativeness of the mobility data and on the assumption that past trends in mobility changes will continue without additional interventions or changes in population behavior.

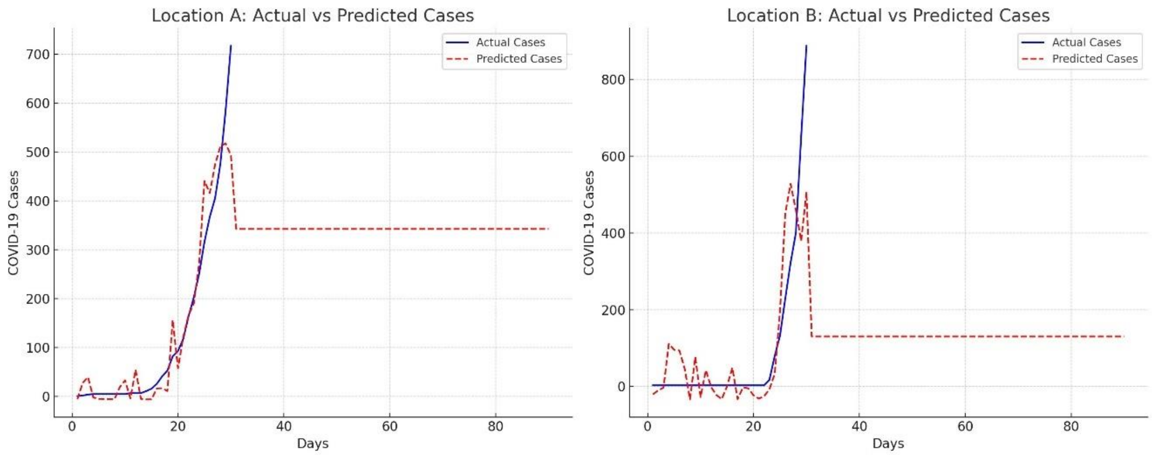

Figure 8: Prediction experiments for 45, 60, and 90 days.

4.3. BASIC REPRODUCTION NUMBER R_0

Utilizing the initial 30-day data, the basic reproduction number, R_0, for both locations can be approximated using a simplified method that relates R_0 to the epidemic's growth rate. This method assumes a homogeneously mixing population without any interventions.

The basic reproduction number, R_0, is a key epidemiological metric indicating the average number of secondary infections produced by one infected individual in a wholly susceptible population.
One common formula to approximate R_0 from observed case counts over time is given by:

$$R_0 = 1 + \frac{\ln\frac{I_t}{I_0}}{g} \qquad (16)$$

Where:
I_t is the number of cases at time t
I_0 is the initial number of cases,
g is the average generation time (in the same time units as t).

For simplicity, the average generation time (g) for COVID-19 was assumed to be about five days. This is a rough estimate, as the generation time can vary based on the population and the virus variant.

The R_0 value is calculated for both locations using the case counts from the first and 30th days, which are I0 and It, respectively.

R_0 was estimated at 2.31 for Location A, and for Location B, at 2.14. These values indicate the average number of secondary infections produced by one infected individual in a completely susceptible population, underlining the potential for spread within each location.

4.4. PREDICTIVE MODELS AND Rt

Using a conceptual model linking mobility changes to R_t and new case predictions, forecasts can be extended up to 90 days. The model predicts an upward trend in cases for both locations, emphasizing the importance of mobility in understanding COVID-19 spread dynamics, but can be extended to other novel or existing viruses.

To predict new COVID-19 cases as a function of mobility change and incorporate the basic reproduction number R_0, a conceptual model that links mobility changes to the effective reproduction number R_t (the average number of secondary cases per infection case at time t) and then to new case numbers can be used, as stated in equation 2. The effective reproduction number R_t adapts R_0 based on changes in behavior or interventions, like mobility changes.

4.5. CONCEPTUAL MODEL

The model could be conceptualized as follows:
Step 1: Relate mobility change to R_t. A simple approach assumes that R_t decreases linearly with decreased mobility, given that reduced mobility reflects reduced contact rates. However, the relationship might be more complex in reality.
Step 2: Relate R_t to new daily cases. Given R_t and the current number of infectious individuals, new cases can be estimated.

Mobility Change to R_t

$$R_t = R_0 \times (1 - kM) \qquad (17)$$

Where:
R_0 is the basic reproduction number,
M represents the mobility change (normalized between 0 and 1, where 0 is no change and 1 is complete lockdown),
k is a constant reflecting the sensitivity of R_t to mobility changes. R_t to New Daily Cases, given the

average generation time (g) and the current number of infectious individuals (I), the new daily cases ($Dnew$) can be approximated as:

$$Dnew = I \times \frac{R_t}{g} \qquad (18)$$

For each location, the R_0 values previously calculated can be used:
Location A: $R_0 = 2.31$
Location B: $R_0 = 2.14$

Though theoretical and simplified, the equations presented lay the groundwork for further refinement and empirical validation. The determination of the value of k necessitates analysis of historical data, specifically how mobility changes impact R_t and, in turn, case numbers.

A simple model is conceptualized with $k = 0.5$ for illustrative purposes. This setup allows exploration of the effects of a 50% reduction in mobility (hypothetically setting $M = 0.5$ on R_t described as m_0 in the introductory section) and the subsequent estimation of new cases based on the current count of infectious individuals. It's important to note that this example simplifies, and the dynamics in the real world are expected to be more intricate.

Figure 8 compares actual cases for the first 30 days and the predicted cases up to 90 days for both Location A and Location B, using a simplified model that links mobility changes directly to the effective reproduction number R_t and, consequently, to new case predictions.

- **Location A:** The blue line represents the actual cumulative cases for the first 30 days, and the red dashed line shows the predicted cumulative cases up to 90 days. The model predicts an upward trend in cases based on the mobility data and the calculated R_0.
- **Location B:** The blue line shows the actual cases, and the red dashed line depicts the predicted cases. The prediction also suggests an increase in cases, although the model's assumptions and simplifications limit the accuracy of these predictions.

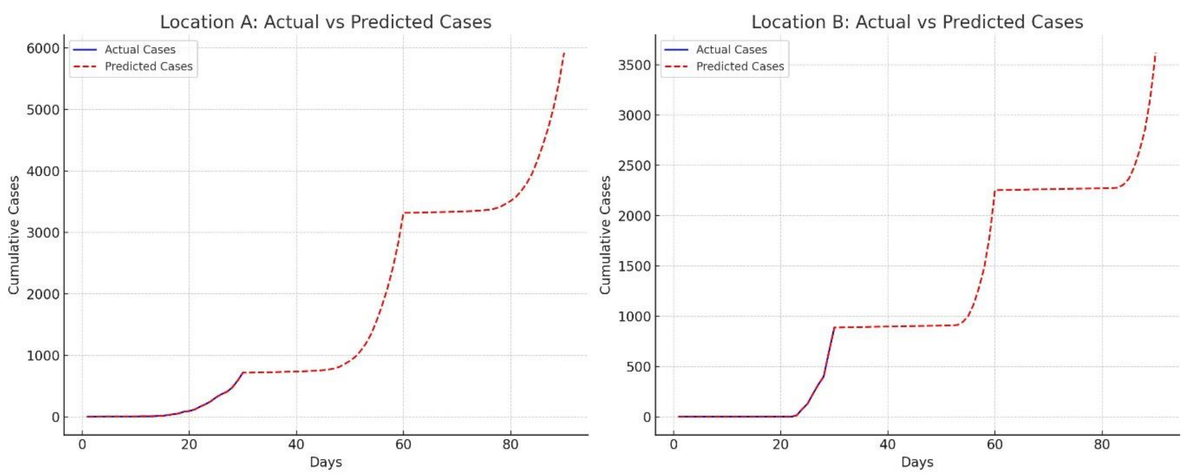

Figure 9: Prediction using the proposed model.

Figure 9 shows both locations' predictive model adjusted to R_t using mobility changes (with M=0.5 and k=0.5). As can be seen, these adjustments smoothen the predictive model's exponential case growth to help health authorities prevent or plan scenarios based on population behavior. These predictions highly depend on the model's assumptions, including the direct relationship between mobility changes and R_t and the constant k value representing R_t sensitivity to mobility changes. However, it provides a powerful predictive tool with early case data (as low as 30 days).

Numerous factors influence the dynamics of virus spread, such as COVID-19, including public health interventions, changes in population behavior over time, and the virus's biological characteristics. Therefore, while this model provides a conceptual framework for understanding potential trends, it should be refined with more detailed data for accurate forecasting.

4.6. MENTAL HEALTH CONSIDERATIONS

The pandemic has ushered in unprecedented changes to daily life, primarily due to social distancing, introduced mobility restrictions, and lockdowns implemented to curb the spread of COVID-19. These alterations have the potential to impact mental health significantly, exacerbated by the invisible yet pervasive threat of the virus [32].

A comprehensive survey by [32] involving 9,565 participants across 78 countries, including the USA, Spain, and Italy, delved into the interplay between various factors ranging from sociodemographic and lockdown variables to social and psychological influences and their impact on mental health. Key variables examined included the duration of quarantine, frequency of leaving home for work, and financial changes. Findings indicated that individuals who ventured out for work, even minimally, exhibited lower levels of stress and depression, coupled with higher well-being, compared to those confined to their homes. This suggests potential negative repercussions of remote work on mental health. However, the study did not establish a direct correlation between the length of quarantine, mobility, and mental health outcomes.

Conversely, a review by [33] on the psychological effects of quarantine underscored predominantly adverse outcomes, such as confusion, anger, and symptoms of post-traumatic stress. While both [32] and [33] hint at the detrimental impact of quarantine and lockdowns on mental health, it remains elusive a definitive link between reduced mobility and the adverse mental health consequences post-pandemic. Yet, there is evidence pointing to a significant global increase in the prevalence of major depression disorder by 27.6% [CI95 25.1, 30.3] and anxiety disorders by 25.6% [CI95 23.2, 28.0] during the COVID-19 outbreak [34]. Moreover, a systematic review by [35] highlighted a surge in anxiety, depression, and PTSD among the general populace, with specific demographics such as women, the youth, and those with existing chronic or psychiatric conditions being particularly susceptible.

Ongoing research is imperative to unravel the intricate connections between diminished mobility, social distancing protocols, and mental health repercussions amidst and after pandemics akin to COVID-19. It is vital to investigate potential causative links, like the effects of extended isolation on mental health, and to assess interventions designed to alleviate such impacts. Thus, a holistic model encapsulating the adverse implications of reduced mobility on mental health and economic downturns could pave the way for minimizing collateral damage. Future studies should aim to identify at-risk groups and formulate bespoke support mechanisms to bolster their mental health during crises. Grasping these dynamics is pivotal for devising public health strategies and interventions that fortify mental health resilience against forthcoming pandemics or analogous adversities.

Studies such as those conducted by [32] and [33] provide empirical evidence of the adverse effects of prolonged isolation and uncertainty, highlighting the urgent need for comprehensive mental health strategies within public health policies.

As the narrative progresses, it becomes increasingly evident that the repercussions of pandemic-induced mental health challenges extend far beyond the individual, exerting significant strain on the broader economic framework. The deterioration in mental health, marked by escalating instances of stress, anxiety, and depression, directly correlates with diminished productivity, heightened healthcare expenditure, and a distressing uptick in suicide rates. This confluence of factors underscores the imperative for a paradigm shift in policy formulation, advocating for a holistic approach that intricately weaves mental health considerations into the fabric of economic recovery and resilience-building measures.

To bridge the gap between the observed mental health impacts and actionable policy interventions, it is proposed that:

- Policies explicitly address the nexus between mental health and economic stability, crafting interventions cognizant of this interdependency.
- Future legislative frameworks incorporate mental health support mechanisms, ensuring economic recovery strategies include mental well-being initiatives.
- Research and policy development efforts are focused on thoroughly examining the comprehensive impacts of the pandemic and devising multifaceted strategies to mitigate adverse outcomes.

Moreover, the role of Infection Control Measures (ICMs) in mediating the relationship between mobility restrictions and mental health outcomes necessitates further exploration. Policymakers can tailor more nuanced and effective interventions by delineating how ICMs influence mental well-being.

In conclusion, integrating mental health considerations into the economic decision-making matrix is beneficial and essential for fostering a resilient society capable of withstanding the challenges posed by future pandemics or similar crises. By adopting a more integrated and empathetic approach to policy development, governments can ensure the enactment of comprehensive strategies that safeguard their populations' economic and mental health, thereby paving the way for a more robust and thriving global community in the post-pandemic era.

5. Conclusions

This study addressed the pressing need to simplify COVID-19 data analysis for effective pandemic management, underscoring the importance of incorporating mobility data into predictive models. The primary goal was to develop a predictive method using open mobility data to project potential scenarios and monitor case growth. This would aid health authorities in making timely decisions with minimal disruption to the country's economies during future outbreaks and reduce the mental health impact.

Our findings indicated that mobility changes significantly impact COVID-19 case trends, highlighting the utility of mobility data in predicting future case trajectories. However, it is crucial to consider the limitations and assumptions inherent in these predictive models, suggesting a need for more comprehensive data and refined methodologies for accurate forecasting.

A notable finding from this study was the complex relationship between mobility changes and mental health, exacerbated by mobility restrictions and social isolation. Although a direct connection between reduced mobility and the adverse mental health effects observed post-pandemic was not established, evidence points to a significant global increase in the prevalence of Major Depression Disorder and Anxiety disorders during the COVID-19 outbreak.

This work underscores the necessity for future research to further investigate the intricate relationship between reduced mobility, social distancing measures, and mental health outcomes during and after pandemics like COVID-19.

Understanding these dynamics will be crucial for informing public health policies and interventions to promote mental health resilience in the face of future pandemics or similar challenges. Future work must focus on refining predictive models by incorporating a more comprehensive array of variables and data sources, which could enhance forecast accuracy and reliability in diverse demographic settings. Another promising direction is integrating advanced Artificial Intelligence techniques, such as machine learning and deep learning, to adapt the model based on real-time data dynamically. Furthermore, exploring mobility changes' psychological and social dimensions could yield insights into behavioral patterns that significantly influence public health outcomes.

Funding: No external funding.

Competing Interest: None

References

1. Mitchell Edith P. Corona Virus: Global Pandemic Causing World-Wide Shutdown *Journal of the National Medical Association*. 2020,112:113–114.
2. Li Qun, Guan Xuhua, Wu Peng, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus Infected Pneumonia *New England Journal of Medicine*. 2020,382:1199–1207.
3. Hu Zhiliang, Song Ci, Xu Chuanjun, et al. Clinical characteristics of 24 asymptomatic infections with COVID-19 screened among close contacts in Nanjing, China *Science China Life Sciences*. 2020,63:706–711.
4. Team Novel Coronavirus Pneumonia Emergency Response Epidemiology. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China *Zhonghua Liu Xing Bing Xue Za Zhi*.2020,41:145–151.
5. WHO coronavirus disease (COVID-19) dashboard Online 2020. Accessed 30 June 2020.
6. Lai Chih-Cheng, Shih Tzu-Ping, Ko Wen-Chien, Tang Hung-Jen, Hsueh Po-Ren. Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and coronavirus disease-2019 (COVID-19): The epidemic and the challenges *International Journal of Antimicrobial Agents*. 2020,55:105924.
7. Bedford Juliet, Enria Delia, Giesecke Johan, et al. COVID-19: towards controlling of a pandemic *The Lancet*. 2020,395:1015–1018.
8. Warren Michael S., Skillman Samuel W. Mobility Changes in Response to COVID-19 2020.
9. Kissler Stephen M., Tedijanto Christine, Goldstein Edward, Grad Yonatan H., Lipsitch Marc. Projecting the transmission dynamics of SARS- CoV-2 through the postpandemic period *Science*. 2020,368:860–868.
10. Kissler Stephen M, Tedijanto Christine, Lipsitch Marc, Grad Yonatan.Social distancing strategies for curbing the COVID-19 epidemic *medRxiv*.2020.
11. Lopez Leonardo, Rodo Xavier. The end of social confinement and COVID-19 re-emergence risk *Nature Human Behaviour*. 2020,4:746–755.
12. Lopez Leonardo R, Rodo Xavier. A modified SEIR model to predict the COVID-19 outbreak in Spain and Italy: simulating control scenarios and multi-scale epidemics *medRxiv*. 2020.
13. Li Lixiang, Yang Zihang, Dang Zhongkai, et al. Propagation analysis and prediction of the COVID-19 *Infectious Disease Modelling*. 2020,5:282–292.
14. Petropoulos Fotios, Makridakis Spyros. Forecasting the novel coronavirus COVID-19 *PLOS ONE*. 2020,15:e0231236.
15. Oliver Nuria, Lepri Bruno, Sterly Harald, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle *Science Advances*. 2020,6:eabc0764.
16. Moulaei Khadijeh, Shanbehzadeh Mostafa, Mohammadi-Taghiabad Zahra, Kazemi-Arpanahi Hadi. Comparing machine learning algorithms for predicting COVID-19 mortality *BMC Medical Informatics and Decision Making*. 2022,22:2.
17. Rambhatla Sirisha, Zeighami Sepanta, Shahabi Kameron, Shahabi Cyrus, Liu Yan. Toward Accurate Spatiotemporal COVID-19 Risk Scores Using High-Resolution Real-World Mobility Data *ACM Trans. Spatial Algorithms Syst*. 2022,8.
18. Google COVID-19 Community Mobility Reports Online 2020. Accessed 30 June 2020.
19. Apple's Mobility Trends Report Online 2020. Accessed 30 June 2020. Delamater Paul L., Street Erica J., Leslie Timothy F., Yang Y. Tony,
20. Jacobsen Kathryn H. Complexity of the Basic Reproduction Number (R0) *Emerging Infectious Diseases*. 2019,25:1–4.
21. Mahase Elisabeth. Covid-19: What is the R number? *The BMJ*. 2020.
22. Hartfield Matthew, Alizon Samuel. Introducing the Outbreak Threshold in Epidemiology *PLoS Pathogens*. 2013,9:e1003277.
23. Ridenhour Benjamin, Kowalik Jessica M., Shay David K. Unraveling R0: Considerations for Public Health Applications *American Journal of Public Health*. 2014,104:e32–e41.
24. Bar-On Yinon M, Flamholz Avi, Phillips Rob, Milo Ron. SARS-CoV-2 (COVID-19) by the numbers *eLife*. 2020,9.
25. Distante Cosimo, Pereira Igor Gadelha, Goncalves Luiz Marcos Garcia, Piscitelli Prisco, Miani Alessandro. Forecasting Covid-19 Outbreak Progression in Italian Regions: A model based on neural network training from Chinese data *medRxiv*. 2020.
26. Considerations for quarantine of individuals in the context of containment for coronavirus disease (COVID-19) Online 2020. Accessed 30 June 2020.
27. Qualls Noreen, Levitt Alexandra, Kanade Neha, et al. Community Mitigation Guidelines to Prevent Pandemic Influenza - United States,

2017 *MMWR Recommendations and Reports*. 2017,66:1–34.

28. Updated Preparedness and Response Framework for Influenza Pandemics Online 2014. Accessed 30 June 2020.
29. Reed Carrie, Biggerstaff Matthew, Finelli Lyn, et al. Novel Framework for Assessing Epidemiologic Effects of Influenza Epidemics and Pandemics *Emerging Infectious Diseases*. 2013,19:85–91.
30. Interim pre-pandemic planning guidance: community strategy for pandemic influenza mitigation in the United States-early, targeted, layered use of nonpharmaceutical interventions Online 2007. Accessed 30 June 2020.
31. Roser Max, Ritchie Hannah, Ortiz-Ospina Esteban, Hasell Joe. Coronavirus Pandemic (COVID-19) *Our World in Data*. 2020. https://ourworldindata.org/coronavirus.
32. Gloster A. T., Lamnisos D., Lubenko J., et al. Impact of COVID-19 pandemic on mental health: An international study *PLOS ONE*.2020,15:e0244809.
33. Brooks S. K., Webster R. K., Smith L. E., et al. The psychological impact of quarantine and how to reduce it: rapid review of the evidence *The Lancet*.2020,395:912-920.
34. Santomauro D. F., Mantilla-Herrera A. M., Shadid J., et al. Global prevalence and burden of depressive and anxiety disorders in 204 countries and territories in 2020 due to the COVID-19 pandemic *The Lancet*.2021,398:1700-1712.
35. Xiong J., Lipsitz O., Nasri F., et al. Impact of COVID-19 pandemic on mental health in the general population: A systematic review *Journal of Affective Disorders*. 2020,277:55-64.