A machine learning approach to analyzing spatiotemporal impacts of mobility restriction policies on infection rates

Annie Young Song[1], Seunghyeon Lee[2]*, S.C. Wong[3,4]

[1] *Department of International Relations, Yonsei University, Wonju, South Korea*
[2] *Department of Transportation Engineering, University of Seoul, Seoul, South Korea*
[3] *Department of Civil Engineering, the University of Hong Kong, Pokfulam Road, Hong Kong SAR, China*
[4] *Guangdong - Hong Kong - Macau Joint Laboratory for Smart Cities*
*Corresponding author: seunghyeon.lee@uos.ac.kr

ABSTRACT

This study analyzed the impact of a range of policies that restrict travel accessibility and mobility on infection rates for the original strain of the virus during the first year of the COVID-19 crisis. We constructed a multidimensional dataset and developed an effective data-driven predictive model to investigate causality between a policy, mobility, and an infection, drawing upon spatiotemporal perspectives. The multidimensional dataset included daily infections, daily restriction policies, and daily and hourly multimodal travel patterns. We quantified and normalized the dataset in relation to pre-COVID-19 policies and travel activities. A machine learning framework that integrated principal component analysis (PCA) and a Gaussian process regression (GPR) was formulated to evaluate the effectiveness of mobility restriction policies and their optimal implementation time during the infancy stage of the pandemic. In a case study, we selected Seoul in South Korea and Sydney in Australia for model calibrations and validations. Both countries deployed comprehensive urban restriction policies during the worldwide pandemic. The proposed model produced better performance than diverse non-parametric and parametric models to estimate the daily number of infections in the two areas. Furthermore, we discovered effective restriction policies and the best times to implement them to minimize the number of acquired COVID-19 cases by analyzing coefficients in PCA and GPR kernel functions. Our finding has far-reaching policy implications. First, the proposed methods can be used for formulating restriction policies for other regions with diverse population densities as the chosen cities in this case study. Second, our finding contributes to evidence-based policymaking.

Keywords: COVID-19, the stringency of restriction policy, multimodal travel patterns, machine learning approach, evidence-based policymaking.

1. Introduction

The advanced development of information and communication technologies (ICT) has allowed one to efficiently collect high-quality multivariate data from multiple sources in recent years. It has derived the advancement of a data-driven approach in transportation domains, which mainly depends on the quality and the amount of available data. Mitchell (1997) defined a machine learning (ML) approach as a technique of computer algorithms that allows computer programs to improve through experience automatically. The ML approach has played a significant role in forecasting and classifying a variety of quantified features due to its excellent performance in the tasks, which do not need to explain causality and sensitivity between variables and parameters. In particular, the ML approach has been adopted in the public policy area in recent years. Diverse ML techniques are used to assess the impacts

of policies on environmental management (Firebanks et al., 2022; Rana & Miller, 2019), to support healthcare with accurate information (Ashrafian & Darzi, 2018), to monitor financial fraud (Aziz & Dowling, 2019), to deliver public services and facilitate administrative decisions (Veale & Brass, 2019), to inform development aid decisions (Pincet et al., 2019), and to identify individual incentives for environment-friendly consumption decisions (de Rubens, 2019).

In the realm of macroscopic transportation domains, the ML has been used for modeling travel behaviors and demands (Koushik et al., 2020 and reference therein). For instance, they are increasingly used to examine activity-travel behavior to propose the optimal routes for the users using their previous choices (Cantarella & de Luca, 2005; Doherty & Mohammadian, 2003; Golshani et al., 2018; Hagenauer & Helbich, 2017). Furthermore, they can be employed to collect travel behavior data from various sources (Lin et al., 2009, Wang et al., 2018, and reference therein). In microscopic traffic modeling, Lee et al. (2019a, 2019b) proposed innovative methods to estimate queue lengths and lane changing maneuvers in an individual lane and reference therein.

While the ML techniques are frequently applied to conduct regression and classification problems in transportation studies, they are rarely used for transport policy-making processes. Particularly, this study focuses on the part of policy-making processes, and policy evaluation. To fill this research gap, we assess the relations between restriction policy and infection rates during the initial response to the COVID -19 crisis in Seoul and Greater Sydney. We aim to analyze the impact of a range of policies that restrict travel accessibility and mobility from spatiotemporal perspectives. In doing so, this study moves beyond the use of ML techniques as identifying patterns and predicting behaviors and extends its direct applicability to the evaluation of transportation policy.

Meanwhile, we note a caveat in its use in transportation policy analysis. Although ML can contribute to using an evidence-based policy, the policy decisions made with the ML techniques can reflect systematic biases in society as ML is based on the data gained from our experience (Coyle & Weller, 2020). This challenge relates to explainability. There is a growing demand to explain ML in policy analysis (Coyle & Weller, 2020). This means that policymakers need to be overt about policy goals that embed their political interests.

We subdivided multivariate datasets into a detailed level of travel patterns and policies to mitigate this challenge, including urban mobility and restriction policies under the COVID-19 crisis. We adopted the ML framework, including principal component analysis (PCA) and a Gaussian process regression (GPR), using a quantified and normalized data-set on pre-COVID-19 policies and travel activities. It enables one to maximize the explainability of the ML techniques in analyzing the impacts of instantaneously deployed policies from multiple perspectives. In a technical aspect, the PCA decreases a dimension of the time-series features to remain uncorrelated variables for modeling the GPR in the following stage. The GPR model was implemented to forecast the number of transmitted COVID-19 cases. Thus, we fused the PCA and GPR model of a pandemic causality of policy-mobility-infection to predict the number of infections. Moreover, the model identifies critical spatiotemporal features in restriction policies and multimodal mobility patterns to reduce locally confirmed cases of COVID-19 at the early stage of the pandemic before vaccinations. The study makes five primary contributions:

- We subdivided and normalized policies that restrict travel accessibility and mobility and multimodal travel patterns to investigate spatiotemporal aspects of each policy for the initial year of the COVID-19 crisis using a ML framework with maximized explainability.

- We calibrated and validated the ML approach using open data collected from two different metropolitan areas, Greater Sydney and Seoul, to cope with future pandemic scenarios in other metropolitan areas.

- Fusing PCA and GPR identifies which policies are effective and when policies should be imposed in response to the pandemic for the initial period to control instantaneously increasing rates of COVID-19 acquisitions.

- The framework identifies the specific policies and mobility patterns that significantly influence the number of locally transmitted COVID-19 cases.

- This study moves beyond the use of ML techniques as identifying patterns and predicting behaviors and extends its direct applicability to the evaluation of transportation policy.

The remainder of this paper is organized as follows. A literature review is provided in Section 2. We illustrate the policy-mobility-infection data profiles in Section 3. Section 4 describes the integrated method, including data normalization, the PCA, and Gaussian process regressions. Empirical results are presented and interpreted in Section 5. Finally, Section 6 presents the conclusions and suggests future research directions.

2. Related works

The World Health Organization (WHO) reported that there were 84.5 million confirmed cases and 1.9 million deaths in 2020 after the rapid spreading of the original strain of the virus, Coronavirus disease 2019 (COVID-19).[1] In response to the outbreak, several countries deployed a range of policies that restrict travel accessibility and mobility; these policies have transformed diverse aspects of society and daily life.

Policies in response to the COVID-19 crisis have included school and workplace closures, the cancelation of public events and gatherings, stay-at-home restrictions, and international and domestic travel restrictions (Han et al., 2020). While the policy responses share similarities in controlling the movement of individuals, countries have chosen to implement these policies with varying degrees of scope and stringency.[2] Some local governments have implemented stricter policies than those applied at the national level because the pandemic had especially severe impacts in their jurisdictions (Goolsbee and Syverson, 2021). Major cities such as New York, Paris, and London[3] introduced lockdown orders, which were the highest levels of travel restrictions over the last year in response to the COVID-19 outbreak.[4] The COVID-19 epidemic led to a reduction in the use of public transit. For instance, individuals changed their transport modes from public transit to others, including walking, cycling, or using private cars in Germany (Anke et al., 2021). Furthermore, there is a decrease in the percentages of individuals using public transit in China, whereas the percentages of those using cars and walking were increased (Jiang et al., 2020). However, the empirical evidence suggests that the demand for public transit has increased (Cho and Park, 2021), although the recovery may depend on psychological factors

[1] https://covid19.who.int/
[2] https://ourworldindata.org/policy-responses-covid#
[3] https://www.nytimes.com/2020/12/19/world/europe/coronavirus-uk-new-variant.html
[4] https://www.bbc.com/news/uk-56158405

(Zhao and Gao, 2022). Similarly, major South Korean cities such as Seoul and Daegu and major Australian cities such as Melbourne and Sydney designed and implemented restriction orders in response to the first confirmed cases of COVID-19 in January and February 2020, respectively.

We summarize the related works dealing with the causality between policies, mobility, and infections during the COVID-19 crisis. Then, we demonstrate the research gaps tackled by this study. Recent studies have focused on the economic effects of policies that restrict mobility. The stringency of such policies has been viewed as a trade-off between the economy and lives, as limiting individual mobility and access invariably affects economic activity. In this sense, lockdown policies appear to be the primary driver of economic recessions, although social distancing has also played a role in economic downturns (IMF, 2020). In addition to policy stringency, the timing of policy implementation has also affected economic costs. For example, a less stringent lockdown in the USA, allowing people to leave their homes for essential needs, lowered deaths "at a lower economic price than mandatory business shutdowns" during the early months of the pandemic (Arnon et al., 2020).[5]

2.1 Travel restriction policy during the COVID-19 crisis

Ku et al. (2021) presented that the several policies for controlling the spread of COVID-19, which are based on suppressing the movement of people and goods, have led to remarkable changes in mobility patterns in multimodal transport systems around the world. Gibbs et al. (2020) investigated the spatiotemporal changes in human mobility caused by travel restrictions in the early stages of the pandemic. Wei et al. (2021) stimulated the spatiotemporal characteristics of COVID-19 using a city-based epidemic and mobility model. Beck and Hensher (2020) examined infection data, policy data, and survey results to provide comprehensive insights into the impact of COVID-19 on household travel and activities in Australia. Hensher et al. (2021) emphasized that working from home (WFH) became the new normal during the COVID-19 crisis. They used a logit model linked to a zero-inflated Poisson regression model to identify the incidence of WFH and its effect on commuting trips. Hu et al. (2021) used mobile data of over 150 million trip records to assess mobility patterns during the COVID-19 crisis. Chan et al. (2021) modeled the relationship between mobility patterns, the number of locally acquired COVID-19 cases, and the stringency of government responses. They argued that Hong Kong communities reacted faster than implementing health interventions, although the government policies effectively decreased the number of local infections. Similarly, Liu and Yamamoto (2022) indicated that the government-declared state of emergency significantly reduced travel, despite the fact that mobility restriction policies (i.e., stay-at-home requests) were not mandatory in Japan. To further examine the effect of government policies on mobility behavior, Bian et al. (2021) quantified the time-lag effect reflected in transportation systems when the government implemented new restriction policies in response to the COVID-19 pandemic. They found that the national declaration of emergency had no time-lag effect reflected in transportation systems, although stay-at-home and reopening policies affected mobility. These studies indicate that stringent measures in response to the COVID-19 outbreak are closely related to fluctuations in travel behavior patterns in metropolitan areas.

Changing mobility patterns have fluctuated with the number of daily confirmed cases worldwide. Nouvellet et al. (2021) investigated the relationship between COVID-19 transmission and mobility in 52 countries. They found that transmission remarkably declined with the initial decrease in

[5] https://www.brookings.edu/bpea-articles/epidemiological-and-economic-effects-of-lockdown/

mobility in 73% of countries. In contrast, there was a decoupling of transmission and mobility following the relaxation of restriction measures for 80% of countries. Badr et al. (2021) used aggregated, and anonymized mobile phone data to analyze daily mobility patterns and found that mobility patterns were strongly correlated with decreased COVID-19 case growth rates for the most affected counties in the USA.

2.2 Machine learning approach

The Gaussian process has been used in non-parametric probabilistic modeling. Sun and Xu (2010) developed mixtures of Gaussian processes to reduce the deficiency of the computationally cubic complexity of a single Gaussian process regression (GPR) model for traffic flow prediction. Jabari and Liu (2013) developed a Gaussian approximation of a stochastic traffic flow model that includes covariance matrices of traffic variables computed using only a few parameters, regardless of system size. Zhang and Wang (2014) introduced a non-parametric distribution model with Gaussian kernel functions to model vehicle headways on urban multi-lane freeways. They extracted intrinsic patterns from observed headways using Gaussian kernel models, which performed better than parametric methods in diverse flow conditions in their experiments. Shin et al. (2017) used Gaussian processes to enhance segmentation accuracy in a real-time object segmentation algorithm for 3D point clouds collected from LiDAR sensors on a vehicle. The concept of multiplicative Gaussian white noise has been used to illustrate the acceleration profiles of the following vehicle toward the leading vehicles' speed profiles in the form of stochastic car-following models (Laval et al., 2014; Ngoduy et al., 2019; Lee et al., 2019; Lee et al., 2021).

In addition, a variety of studies have validated the excellent predictive performance of GPR compared to other models. Murça and Hansman (2018) used a GPR model to create an airport capacity prediction model capable of translating predictions of external features into probabilistic arrival capacity predictions for strategic time horizons. They compared the performance of the GPR model with that of Bayesian linear regression and random forest models. The GPR produced more accurate point predictions and uncertainty quantification. Rodrigues and Pereira (2018) proposed the use of heteroscedastic Gaussian processes to model non-constant observation noise in traffic speed modeling for imputing missing observations and real-time forecasting. Liu et al. (2019) used a GPR model to estimate spatially heterogeneous traffic volume and fleet compositions based on heterogeneous datasets, including taxi GPS data, license plate recognition data, and geographical features in a large-scale urban network. Gammelli et al. (2020) constructed a flexible non-parametric GPR model incorporating a censored likelihood function to handle the censoring problem inherent in travel demand forecasting models.

PCA is one of the most popular and oldest multivariate statistical techniques. The goals of PCA are (a) to extract the most critical information from a data table, (b) to compress the size of the dataset by keeping only this vital information, (c) to simplify the description of the dataset, and (d) to analyze the structure of the observations and the variables (Abdi and Williams, 2010).

Wang and Lien (2008) developed novel automatic vehicle detection methods based on local features located within three significant subregions of an image by using PCA to model the low-frequency components of the eigenspace. Qu et al. (2009) used probabilistic PCA to impute missing traffic flow data based on historical data mining. Their probabilistic PCA outperformed conventional methods in their experiments. García et al. (2010) used PCA to validate obstacle detection on high-speed railways and to improve the efficiency of the system. Guardiola et al. (2014) adopted functional

PCA to project the original dataset into a low-dimensional space. Jenelius and Koutsopoulos (2017) proposed a multivariate probabilistic PCA to develop a network travel time prediction methodology based on probe data. Lasisi and Attoh-Okine (2018) used PCA to find a low dimensional representation of a given dataset of track quality indices used to obtain an average-based assessment of track segments and schedule track maintenance. Wang et al. (2018) used PCA to reduce the dimensions of variables for the short-term trajectory prediction problem in a terminal maneuvering area. Saffari et al. (2020) presented a novel network-wide approach to identifying critical links and estimating average traffic flow and density. They reduced a dataset containing variables that were potentially interrelated to a smaller dataset of uncorrelated variables using PCA. Kumar and Khani (2020) evaluated special event transit demand using a robust PCA to reduce the dimension of automatic passenger count data.

2.3 Main findings from literature review

We derived the following two main findings from our review of the literature. First, recent studies, e.g., those of Guzman et al. (2021), Pawar et al. (2021), Chen et al. (2022), De-Toledo et al. (2022), and Zhao and Gao (2022), have used diverse methodologies based on both dynamic and static data sources to separately investigate the relationships during the early stage of the COVID-19 pandemic between restriction measures and mobility patterns, and between mobility patterns and infections. However, no recent studies have used ML algorithms and considered multiple time-lags to simultaneously investigate the relationships during this period between infections, restriction policies, and multimodal urban mobilities. Moreover, to the best of our knowledge, predictive models of locally transmitted COVID-19 cases based on the policy–mobility–infection cycle have not been captured from a spatiotemporal perspective. Furthermore, complex restriction policies have not been decomposed to identify the policies that are effective and when these policies should be imposed in response to the pandemic. In this study, we conducted real-world empirical modeling based on datasets collected from two regions that had implemented restriction policies to validate the effectiveness of our models in explaining policy–mobility–infection relationships under different circumstances.

Furthermore, we used GPR, a widely used ML algorithm, with PCA based on a spatiotemporal multivariate dataset. Rasmussen and Williams (2006) noted that "a Gaussian process is a generalization of the Gaussian probability distribution, which describes a finite-dimensional random variable, to functions" (pp. 13-16). Thus, a GPR framework enables the simple implementation of a flexible data-driven approach and stochastic forecasting with a confidence interval. As a result, GPR has exhibited excellent performance as a non-parametric kernel-based probability model in prediction procedures involving a small number of datasets. Because of its simplicity and reliability, PCA has been widely used in transport engineering to decrease the dimensions of original datasets and identify combinations of non-correlated variables.

3. Data

To examine the applicability of ML techniques to various policies, we chose Greater Sydney in Australia and Seoul in South Korea as cases to calibrate and validate the proposed ML framework for demonstrating causality between pandemic policy, mobility, and infection. Sydney and Seoul adapted their COVID-19 related policies to instantaneously restrict the transmission of the COVID-19 in the community under the varied circumstance. The geographic and demographic specifications of each city have shaped how they have deployed restriction policies in response to COVID-19. By comparing their policies, we identified the best (*i.e.*, most generalizable, flexible, and applicable) method for modeling the causality between policies, mobility, and infection rates.

3.1 Study area

In designing effective response policies, the two governments varied in their timings for policy implementation and the stringency of restrictions within the constraints of their unique conditions. The restriction policies on mobility were deployed by New South Wales (NSW) government based on the framework of the COVID-19 guidelines offered by the Australian government for the first year of the COVID-19 crisis. In contrast, the Seoul government mainly followed Korea's national COVID-19 guidelines, which sought to end transmission cycles in the middle via "Test-Trace-Treatment" (3T) strategies rather than placing stringent restrictions on mobility in the first year of the pandemic. Australia thoroughly and quickly banned international arrivals from China: the Australian government confirmed the first coronavirus case in Victoria on January 25, 2020, and announced a ban on air flight arrivals from all areas of China six days later. South Korea reported its first coronavirus case on January 20, 2020, and 15 days later, it imposed a travel ban on foreign nationals who had stayed in Hubei province in the previous 14 days. While international arrivals were still allowed to enter, Seoul utilized testing and tracing to minimize the transmission risks posed by these arrivals. Furthermore, domestic mobility restrictions were stricter for Sydney than for Seoul. After the first surge, Sydney announced stay-at-home orders and lockdown when the second wave of COVID-19 infections arrived. In contrast, Seoul announced more strict social distancing and mask-wearing rules rather than imposing stringent measures on mobility.

These differences in response policies are related to differences in the demographic and geographic conditions of the two regions. Greater Sydney and Seoul are profiled in Table 1. Greater Sydney covers an area 20 times larger than Seoul, although Seoul has approximately twice the population and a population density 37 times greater. Therefore, limited human resources, connectivity, and accessibility led Sydney to seek to inhibit the development of COVID-19 transmission cycles by deploying high levels of restrictions on mobility at an early stage, whereas Seoul's highly dense population and daily mobility—approximately ten times larger than Sydney's—(see Table 1) led to the imposition of 3T strategies to disrupt transmission cycles.

With some similarities in their restrictions, Australia and South Korea are listed among the top 10 best countries in terms of response to the COVID-19 crisis before vaccination, according to Bloomberg's Covid Resilience Ranking [6] based on comprehensive review processes [7]. Australia's ranking is in part because it closed its international borders to contain COVID-19. In contrast, South Korea did not implement a stringent mobility policy, such as a lockdown. Instead, it focused on contact tracing to prevent further community transmission. The information gained from contact tracing was rapidly disseminated to the public and encouraged voluntary compliance with the government's guidelines (You, 2020).

The availability of policy information and multimodal travel patterns for Sydney and Seoul made this study feasible. These extensive metropolitan areas play significant roles in politics, the economy, education, and transportation in their respective countries. We were easily able to mine the

[6] https://www.bloomberg.com/graphics/covid-resilience-ranking/
[7] https://www.bloomberg.com/news/articles/2020-11-24/inside-bloomberg-s-covid-resilience-ranking

required dataset from 33 and 25 local government areas in Sydney and Seoul, respectively, as their governments provide time-series infection, policy, and mobility data to the public on their websites. The NSW government also established archives of news and media releases related to COVID-19 crisis[8]. Transport for NSW (TfNSW) operates an open data website[9] that provides daily and hourly Opal card usage for all public transport modes and road traffic counts at designated on-road sensors. South Korea's Ministry of Health and Welfare provides webpages[10] detailing infections, restrictions, news and media releases, and government briefings. In Seoul, the Transport Operation and Information Service (TOPIS) records and provides high-resolution traffic volume data and public transport usage[11]. These open data sources through which government agencies document and deliver information to the public are highly reliable and valuable resources for developing and analyzing macroscopic estimation models based on ML algorithms.

Table 1. Profiles of the Greater Sydney and Seoul study areas

	Greater Sydney	Seoul
Map		
Country	Australia	South Korea (Republic of Korea)
Population	5,312,163 (2019)	9,575,355 (2021)
Area	12,367.7 km^2	605.2 km^2
Population density	423/km^2	15,821.8/km^2
Local government areas	33	25
Number of suburbs	658	424
CBD	The City of Sydney	Jongno-gu, Jung-gu, Yongsan-gu
First COVID-19 case	March 1, 2020	January 30, 2020
Analysis period	March 1, 2020 – February 28, 2021	February 1, 2020 – January 31, 2021
Pre-COVID-19 period for normalizing features	February 1, 2020 – February 29, 2020	January 1, 2020 – January 31, 2020

The policies in Sydney and Seoul differed in their timings for implementing the restrictions on international arrivals and the stringency of such approaches, but they were similar in that they introduced restrictions before vaccinations. The policies successfully evolved over the study period to control community transmission of COVID-19 under different constraints. Comparing the two cities can provide insights into how demographic and geographical conditions may influence the effectiveness of mobility policy decisions.

[8] https://www.nsw.gov.au/covid-19/news-and-media
[9] https://opendata.transport.nsw.gov.au/
[10] http://ncov.mohw.go.kr/en/
[11] https://topis.seoul.go.kr/

3.2 Multiple data profiles

We illustrate multivariate data profiles and structures used as input data to a ML framework. The data processing model is described in Figure 1. The multivariate time-series dataset involves data profiles on road traffic counts and transit passenger counts in multimodal transport systems, restriction policy profiles, and COVID-19 related profiles for Seoul, South Korea and Greater Sydney, Australia. The study period is an initial year from when the first locally acquired COVID-19 case was observed in each city. We set the dependent variable as the number of daily COVID-19 cases in the designated area. The independent variables contain the mobility and the policy-related variables defined by divisions, features, and dates. The normalization process required defining the independent features to numbers from 0.0 to 1.0 based on their ratio to pre-COVID-19 mobility behavior and policy stringency, respectively. We transform the normalized input dataset to the principal component set in a ML framework to reduce structural dimensions using PCA. It helps to improve the efficiency of the GPR model. In the data-driven model, we apply the integrated framework of the GPR and PCA methods, as devised by Lee and Chen (2022). The principal components are defined as the independent variables to estimate the number of daily acquired COVID-19 cases in Seoul and Sydney. We present the specific data pre-processing and mathematical frameworks in the Method section. Analysis periods for Greater Sydney and Seoul cases are presented in Table 2.

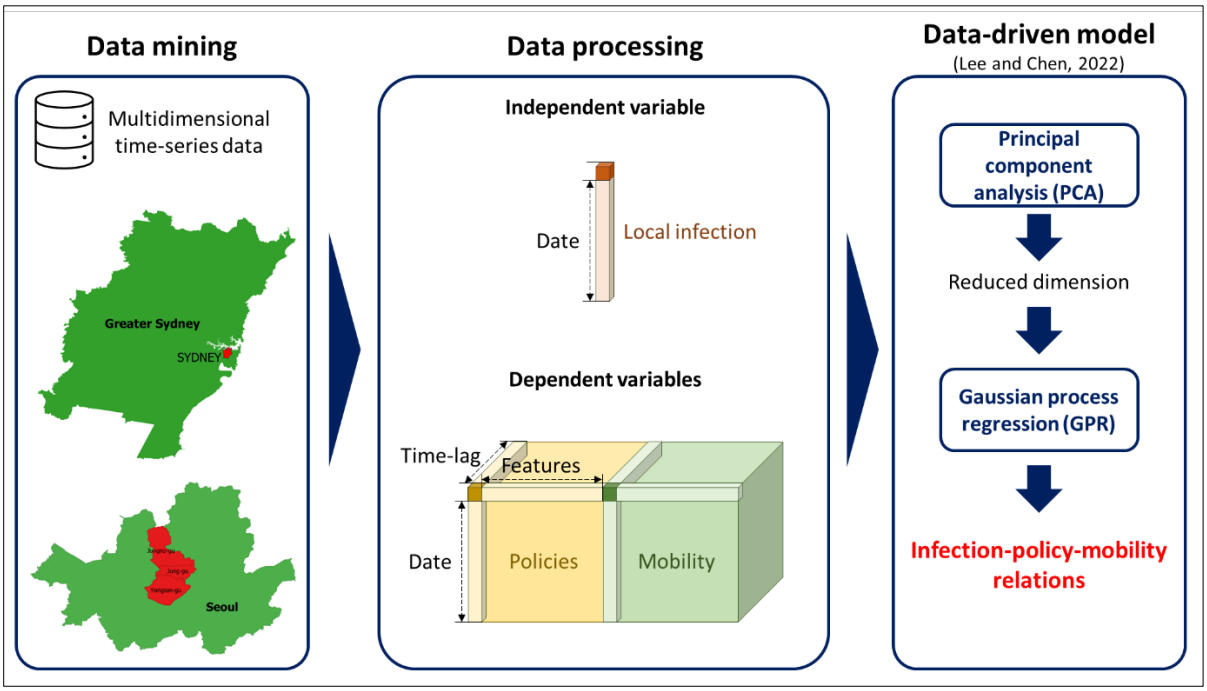


Figure 1. The overall framework for model construction

Table 2. Analysis periods for Greater Sydney and Seoul

	Greater Sydney	Seoul
First COVID-19 case	Mar 1, 2020	Jan 30, 2020
Analysis period	Mar 1, 2020 – Feb 28, 2021	Feb 1, 2020 – Jan 31, 2021
Pre-COVID-19 period for normalizing features	Feb 1, 2020 – Feb 29, 2020	Jan 1, 2020 – Jan 31, 2020

3.2.1 *Policies*

This section provides an overview of the travel restriction policies of Seoul and Greater Sydney and describes our procedure for coding their policies into ordinal indicators. We distinguished between policies for international and domestic travel restrictions to understand the similarities and differences between Australian and Korean policies.

The international travel restrictions of Seoul and Sydney varied in terms of timing and target areas. Australia implemented an international travel ban earlier than Korea, designating mainland China as a travel restriction area. The Australian government confirmed the first cases of coronavirus on January 25, 2020, in Victoria. Responding to the spread of the virus, Australia announced a ban on air flight arrivals from China on February 1, 2020. While it took six days from Australia's first case to its international travel ban from China, Korea announced a ban on arrivals from Hubei province 15 days after identifying its first case. Korea reported the first case of coronavirus in Daegu on January 20, 2020, and imposed a travel ban on foreign nationals who had stayed in Hubei province in the previous 14 days on February 4, 2020, later expanding this ban to the rest of China and other countries.

To suppress and control the local transmission of the COVID-19, Seoul and Sydney instantaneously organized and subdivided policies limiting mobility and accessibility. The unprecedented pandemic and uncertainty about the nature of the virus derived that the governments reactively modified and adjusted response strategies. While some of the orders and measures were specific and devised on an ad hoc basis after the transmission of infections (e.g., restrictions on corporate gatherings in Sydney and coin karaoke in Seoul), the Australian and Korean governments provided comprehensive guides to their responses to the outbreak.

On March 2, 2020, the Korean government launched a "Temporary Stop" campaign to encourage the public to engage in voluntary social distancing by avoiding face-to-face meetings, communicating online, and adhering to sanitation rules. On May 3, 2020, the government relaxed the level of social distancing to "social distancing in daily life" (i.e., this level prompts individual adherence to COVID Safe steps while maintaining social and economic activities and daily routines)[12] effective from May 6, 2020, because the number of infections had decreased. A social distancing framework was introduced on June 28, 2020, to differentiate three levels of social distancing—(1) daily life, (2) intermediate, and (3) strict—depending on the severity of the outbreak. In response to a surge in infections in August 2020, the government added more robust measures for intermediate social distancing into the existing framework. On November 7, 2020, the social distancing regulations were revised to five levels (1, 1.5, 2, 2.5, and 3; see Table 3). The regulations defined distancing requirements for each level, such as the capacity of indoor and outdoor activities and the closure or reopening of cultural and arts facilities and privately owned education facilities and gyms. We used publicly available data from the Seoul Metropolitan Government website to find indicators such as regulations on compulsory mask-wearing, social distancing rules, and capacity restrictions for indoor and outdoor household activities, restaurants, schools, and public gatherings.

Table 3. Criteria for social distancing levels in Seoul[1)]

[12] Central Disaster and Safety Countermeasures Headquarters (2020, May 3). Government announces the change to the social distancing level to 'social distancing in daily life'. *Republic of Korea Policy Briefing* https://www.korea.kr/news/policyNewsView.do?newsId=148872066

Social distancing level	1	1.5	2	2.5	3
	Low	Community-wide		Countrywide	
Severity of outbreak	Social distancing in daily life	Initial phase of community-wide outbreak	Transmission phase of the outbreak and initial phase of countrywide outbreak	Large-scale countrywide outbreak	Countrywide pandemic
Indicator	Fewer than 100 cases in the metropolitan area, fewer than 30 cases in other areas (fewer than 10 cases in Gangwon and Jeju)	More than 100 cases in the metropolitan area, more than 30 cases in other areas (more than 10 cases Gangwon and Jeju)	One of the following conditions to be met - a twofold increase of cases indicated in the 1.5 level - more than two areas persisting at the 1.5 level - more than 300 cases in total	More than 400-500 cases, or a surge of cases (e.g., a doubled rate)	More than 800-1,000 cases or a surge of cases (e.g., a doubled rate)

[1)]Source: Seoul Metropolitan Government (https://news.seoul.go.kr/welfare/archives/524124)

The Australian government nationally demonstrated a guideline detailing safety measures, including limited operation hours and capacity of public gatherings (see Table 4). In May 2020, it laid out a three-step guide for responding to the outbreak in daily life settings. Step 1 was critical first small steps, allowing friends and family to connect and groups of people to be together in homes and the community. Step 2 contains slightly larger gatherings and more businesses reopening, while higher-risk activities have tighter restrictions. In Step 3, it is suggested to reopen business and the community with minimal restrictions but underpinned by COVID-safe ways of living.

Table 4. Australian national guidelines in response to the COVID-19 outbreak[1)]

Step 1	Critical first small steps: allowing friends and family to connect and groups of people to be together in homes and the community.
Step 2	Slightly larger gatherings and more businesses reopening. Higher-risk activities have tighter restrictions.
Step 3	A commitment to reopening business and the community with minimal restrictions but underpinned by COVID-safe ways of living.

[1)]Source: Department of Health, the Australian government (https://www.health.gov.au/resources/publications/3-step-framework-for-a-covidsafe-australia)

While both governments provided step-wise guidance to respond to the outbreak, their urban policies differed. In the case of Australia, states and territories can implement changes based on their COVID-19 conditions, so state governments have the flexibility to adapt their policies. Rather than pinpointing the number of infections for each level of the outbreak as in Korea's policy framework, the Australian government offered more general recommendations to state governments. For this reason, the NSW government policy appears to be more detailed than the national policy in terms of how the restrictions were imposed. For instance, Sydney banned dance floors at weddings on December 20, 2020, whereas it had allowed dance floors with up to 20 dancers on September 24, 2020. We captured indicators for compulsory mask-wearing, social distancing, and capacity restrictions for indoor, outdoor, and household activities, restaurants, pubs, schools, and public gatherings; capacity restrictions for public transport, weddings, hospitality, funerals, and worship; and permissions for regional and interstate travels. We used binary coding for these indicators and coded full restrictions as 0 and 1 otherwise. For instance, the closure and reopening of facilities were coded as 0 and 1, respectively.

3.2.2 *Mobility*

TOPIS and TfNSW, which were open data web services in Seoul and Sydney, respectively, store and provide high-dimensional time-series mobility data via in multimodal urban transport systems. Opal patronage data is available in TfNSW for train, bus, ferry, and light rail transit (LRT). The data structure involves the number of Opal hourly taps on and taps off, the date of the trip, and modes. In addition, TfNSW categorizes travel insights regionally into the whole of NSW, Sydney CBD, Parramatta, Chatswood, Macquarie Park, North Sydney, Strathfield, Newcastle, and surrounds, Wollongong and surrounds, and Other. We describe the overall aggregated trips of each mode in Greater Sydney in Appendix I(a). We integrated 455 files of Opal patronages, organized by 685 rows and six columns, into one spreadsheet to establish the data structure of multimodal mobility in Sydney. The usage patterns in the transit systems have not been quickly recovered until the end of our study period due to the remaining 1.5 m social distancing rules and recommended WFH policies.

TOPIS provides the number of hourly and daily taps on and taps off at individual bus and metro stations per bus and train. The dataset includes station number, station name, route number (or bus number), and hourly taps on and taps off for 24 hours. Appendix I(b) describes the daily bus and metro trips for the year after the first confirmed COVID-19 case and the pre-COVID-19 period to normalize the features. We integrated three bus patronage files organized by 469,133 rows and 56 columns and 17 metro patronage files organized by 202,281 rows and 25 columns into one spreadsheet to establish the mobility data structure for the transit system in Seoul. Although road traffic counts did not significantly fluctuate throughout the study period, three valleys in road traffic count patterns in off-peak hours correspond to patterns in policies and infections in Seoul. The use of the transit system had significantly recovered at the end of the study period, although the level of infections was higher than that in Greater Sydney.

In addition to daily and hourly patterns of passengers in transit systems, we demonstrated road traffic counts from TfNSW and TOPIS. TfNSW provides road traffic volume count data, including traffic direction, vehicle classification, date, year, month, day of week, holiday index, hourly road traffic counts for 24 hours, and road traffic count station data, including station ID, road name, intersection, road number, road hierarchy, number of lanes, road classification, region, LGA, suburb, post code, latitude-longitude coordinates, and direction. We utilized an application programming interface (API) to combine these data into one spreadsheet organized in 381,820 rows and 16 columns. Appendix I(c) illustrates the aggregated daily road traffic counts in Sydney CBD and Greater Sydney for morning peak hours from 6:00 am to 9:00 am, afternoon peak hours from 4:00 pm to 7:00 pm, and off-peak hours. Road traffic counts significantly decreased after the first travel restrictions were imposed and then recovered, at times exceeding the counts for the pre-COVID-19 period. Due to social distancing rules, Sydney residents appear to have changed their commuting mode from public transit to private vehicles on weekdays.

TOPIS also provides hourly and daily road traffic counts collected from 135 inductive loop detectors installed across Seoul. We downloaded road traffic count files organized by 8,371 rows and 30 columns for 13 months and integrated them into one spreadsheet. The files contained the date, location of data collection points, location ID, directions, road name, and hourly road traffic counts for 24 hours. Appendix I(d) shows the aggregated daily road traffic counts in the CBD and Seoul for morning peak hours from 7:00 am to 10:00 am, afternoon peak hours from 4:00 pm to 7:00 pm, and off-peak hours. Although road traffic counts did not significantly fluctuate across the study period, three

valleys in road traffic count patterns for off-peak hours correspond to patterns for policies and infections in Seoul.

We divided daily multimodal patronage data in the CBD and metropolitan areas into morning peak hours, afternoon peak hours, and off-peak hours according to taps on and off and by mode. The CBD and metropolitan road traffic counts were categorized into morning peak hours, afternoon peak hours, and off-peak hours according to the road hierarchy in both regions. The designated data structure is described as a set of independent variables in the following section.

3.2.3 *Infections*

Greater Sydney and Seoul announced daily COVID-19 news, including overseas and locally acquired COVID-19 cases, after observing the first case of the COVID-19 infection. A trend in COVID-19 infections in Sydney and Seoul is illustrated in Appendix II. Appendix II(a) presents that overseas and locally acquired COVID-19 cases surged dramatically in Sydney in the first wave of the COVID-19 infection, March 2020. After the initial wave of infections, locally acquired cases did not increase above 30 cases per day in Sydney in the following waves because the federal government closed the Australian borders to non-residents on March 20, 2020.

Appendix II(b) shows that Seoul avoided high infection rates for several months after the first confirmed cases in January 2020. The peak of the first wave was below 50 locally confirmed cases per day. However, locally transmitted cases in Seoul increased rapidly from August 2020. Daily infections were continuously above 100 cases per day after the peak of the third wave and over 500 cases toward the end of the study period. Overseas cases were below 15 cases per day during the first wave and below 10 cases per day afterwards. Appendix II(c) compares the total accumulated cases in Greater Sydney and Seoul, including overseas and locally transmitted COVID-19 cases, to illustrate the differences in infection trends between the two regions. Although the total number of locally acquired COVID-19 cases in Sydney was around 5,000 after an initial jump, the number in Seoul was approximately 25,000 after two jumps toward the end of the study period.

The number of locally transmitted COVID-19 cases in Seoul and Sydney was used as dependent variables to demonstrate causality between mobility, policy, and infections. The policy-related variables and mobility-related variables were normalized to explain the number of local transmissions in Sydney and Sydney through the integrated PCA and GPR model without a modeling bias.

4. machine learning framework

We describe the data normalization and the ML algorithms in this section. We developed the integrated PCA and GPR to predict the number of COVID-19 cases in Lee and Chen (2022). The developed framework was modified to explain the causality between infections, policies, and mobility in this study. The PCA is used to reduce data dimensions for modeling the GPR because the GPR model is more effective for low-dimensional data in an estimation process. As described in the previous section, the number of acquired COVID-19 cases in a local, $\mathbf{Y} = \{y_i\}$, $\forall i \in I$, is an outcome variable given policy and mobility features, $\mathbf{X} = \{x_{i,j,k}\}$, $\forall i, j, k \in I, J, K$, where I, J, and K denote the number of days since the first confirmed case, policy and mobility indicators, and time lags, respectively, in Greater Sydney and Seoul.

4.1 Data pre-processing: normalization

This section illustrates the data process and structure for normalizing the dependent and independent variables in the PCA and the GPR models. We constructed two sets of dependent and independent variables for the Seoul and Greater Sydney areas. The dependent variable is the number of locally acquired cases shown in Figure 1. The independent features are mobility and policy-related indicators.

4.1.1 *The first set of independent variables: policy indicators*

The Oxford COVID-19 Government Response Tracker proposed a stringency index (SI). Chan et al. (2021) used the SI to quantify and normalize a qualitative form of variables, the stringency of government responses to COVID-19 in Hong Kong. Hale et al. (2020) provided detailed procedures and organizations. The basic concept of the SI was enhanced in this study to identify the relatively effective restriction policies and best timings for imposing the identified policies to suppress local transmission. The index table of the original SI was reversed from 0.0 (no restriction) to 1.0 (full restriction). The reverse SI was applied to each individual category of policies. This new SI was defined as individual inverse SI (IISI) in this study.

Restriction policies in response to the COVID-19 crisis in NSW were categorized into 17 variables in Table 5. The IISI from 0.0 (full restriction) to 1.0 (no restriction) was utilized to normalize the stringency of an individual categorized policy. The number of classes is varied across the number of stringency levels in each policy. For example, the mask policy has three stringency levels: not mandatory, partially mandatory, and mandatory. Therefore, there were three classes in the IISI for the mask policy, with 1.0 indicating not mandatory, 0.5 partially mandatory, and 0.0 mandatory. Table 5 describes all policies and their stringency levels, imposition and release dates, and IISI values. The household capacity restriction had 11 classes in the IISI, whereas the capacity policies for public gatherings and transit had two classes.

Table 5. Inverse individual stringency index (IISI) for Greater Sydney

Policy variable	Stringency level	From	To	IISI (0.0-1.0)
Mask	Not mandatory		1/3/2021	1.00
	Mandatory	1/4/2021	1/29/2021	0.00
	Partially mandatory	1/30/2021		0.50
Social distancing	No social distance		3/17/2020	1.00
	Social distancing measures of 1.5 m (2.25 m^2)	3/18/2020	5/31/2020	0.50
	One person per 4 m^2	6/1/2020	12/25/2020	0.75
	Residents do not leave this zone (“stay-at-home” restriction)	12/26/2020	1/9/2021	0.00
	2 m^2 except gym	2/10/2021		0.25
Indoor (capacity)	Unlimited		3/17/2020	1.00
	100	3/18/2020	11/23/2020	0.75
	50	11/24/2020	12/17/2020	0.50
	0	12/18/2020	12/19/2020	0.00
	No singing/chatting/dancefloors	12/20/2020	12/25/2020	0.00
	0	12/26/2020	1/1/2021	0.00
	30	1/2/2021	2/23/2021	0.25
	50	2/24/2021		0.50
Household (capacity)	Unlimited		5/14/2020	1.00
	5	5/15/2020	6/10/2020	0.20
	20	6/11/2020	12/17/2020	0.60
	0	12/18/2020	12/19/2020	0.00
	10	12/20/2020	12/23/2020	0.40
	5	12/24/2020	1/1/2020	0.20
	0	1/2/2021	1/9/2021	0.00
	10	1/10/2021	1/28/2021	0.40
	30	1/29/2021	2/23/2021	0.60

	50	2/24/2021	3/28/2021	0.80
	Unlimited	3/29/2021		1.00
Restaurant (group)	Unlimited		5/14/2020	1.00
	10	5/15/2020	5/30/2020	0.33
	50	6/1/2020	7/16/2020	0.66
	10	7/17/2020	12/17/2020	0.33
	0	12/18/2020	12/20/2020	0.00
	10	12/21/2020		0.33
Pub (group)	Unlimited		5/31/2020	1.00
	50	6/1/2020	7/13/2020	0.75
	10	7/14/2020	10/22/2020	0.25
	30	10/23/2020	12/17/2020	0.50
	0	12/18/2020	12/20/2020	0.00
	30	12/21/2020		1.00
School	Unlimited		5/10/2020	1.00
	1	5/11/2020	6/14/2020	0.33
	2	6/15/2020	3/4/2021	0.66
	Unlimited	3/5/2021		1.00
Outdoor (group)	Unlimited		3/17/2020	1.00
	499	3/18/2020	5/14/2020	0.83
	10	5/15/2020	6/10/2020	0.00
	20	6/11/2020	10/22/2020	0.17
	30	10/23/2020	11/24/2020	0.33
	50	11/25/2020	12/29/2020	0.50
	30	12/30/2020	1/28/2021	0.33
	50	1/29/2021	3/28/2021	0.50
	200	3/29/2021		0.67
Public gathering (capacity)	Unlimited		6/30/2020	1.00
	20	7/1/2020		0.00
Transit (capacity)	Unlimited		6/30/2020	1.00
	50%	7/1/2020		0.00
Regional travel	Unlimited		3/30/2020	1.00
	No essential trip	3/31/2020	5/31/2020	0.00
	Unlimited	6/1/2020		1.00
Interstate travel	Unlimited		6/30/2020	1.00
	14 days self-isolation	7/1/2020	8/4/2020	0.33
	Hotel quarantine	8/5/2020	9/3/2020	0.00
	Permit required	9/4/2020	11/22/2020	0.66
	Unlimited	11/23/2020		1.00
International travel	Unlimited		3/17/2020	1.00
	14 days self-isolation	3/18/2020	7/7/2020	0.66
	450	7/8/2020	7/19/2020	0.33
	350	7/20/2020		0.00
Weddings	Unlimited		5/14/2020	1.00
	10	5/15/2020	5/31/2020	0.14
	20	6/1/2020	7/16/2020	0.29
	150	7/17/2020	9/23/2020	0.71
	20	9/24/2020	12/19/2020	0.29
	0	12/20/2020	1/1/2021	0.00
	100	1/2/2021	1/28/2021	0.57
	300	1/29/2021	3/28/2021	0.86
	Unlimited	3/29/2021		1.00
Hospitality	Unlimited		10/15/2020	1.00
	2 m^2	10/16/2020	12/19/2020	0.66
	4 m^2	12/20/2020		0.33
Funerals	Unlimited		5/14/2020	1.00
	30	5/15/2020	5/31/2020	0.00
	50	6/1/2020	7/16/2020	0.25
	100	7/17/2020	1/28/2021	0.50
	300	1/29/2021	3/28/2021	0.75
	Unlimited	3/29/2021		1.00
Worship	Unlimited		5/14/2020	1.00
	10	5/15/2020	5/31/2020	0.00
	50	6/1/2020	7/16/2020	0.25
	100	7/17/2020	10/20/2020	0.50

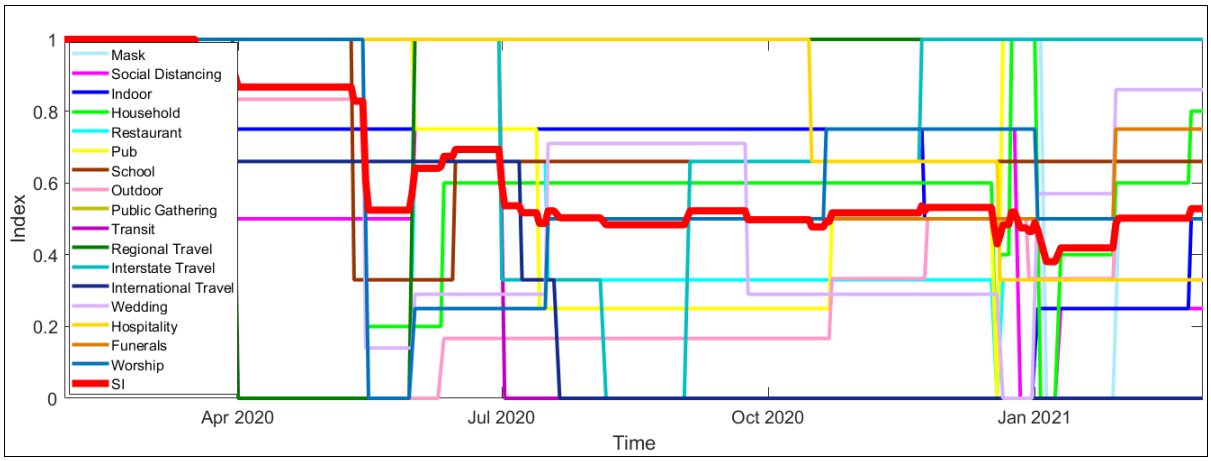
	300	10/21/2020	1/1/2021	0.75
	100	1/2/2021	3/28/2021	0.50
	Unlimited	3/29/2021		1.00

Our IISI table for restriction policies in Seoul is shown in Table 6. The Seoul Metropolitan Government followed the national social distancing regulations developed by the Korean government. Although the national social distancing regulations include a mask policy, the capacity of public gathering in complex facilities, the capacity of indoor facilities, private study and tutoring restrictions, childcare restrictions and restaurant restrictions, we separately developed the IISI for these six policies because the government imposed these policies individually when they observed COVID-19 clusters related to specific locations and facilities. Therefore, we introduced seven policy variables with corresponding IISI values to develop independent variables for the Seoul modeling. Table 6 contains the policy variables, stringency levels, corresponding imposition and release dates, and IISI value.

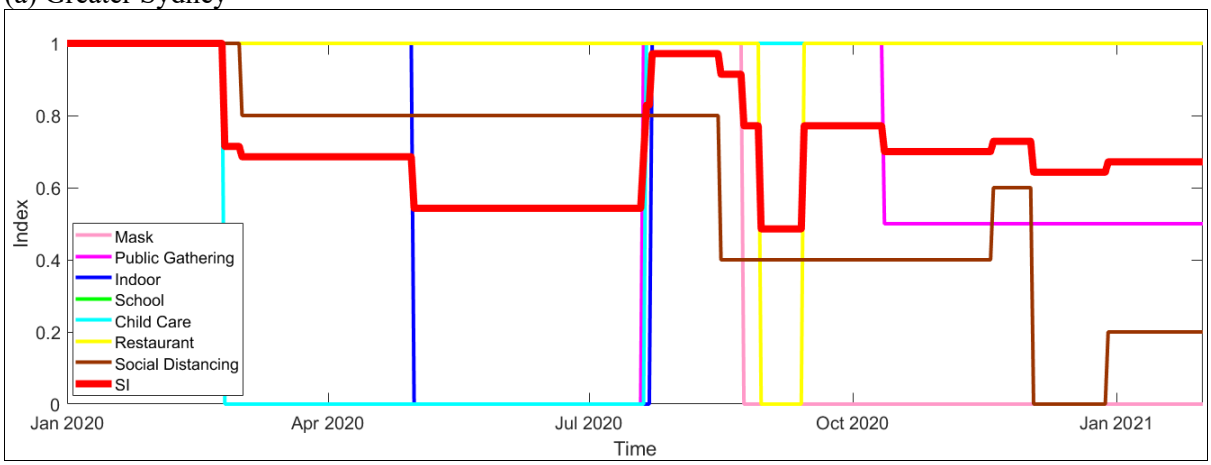
Table 6. Inverse individual stringency index (IISI) in Seoul

Policy variable	Stringency level	From	To	IISI (0.0-1.0)
Mask	0		8/23/2020	1.00
	1	8/24/2020		0.00
Public gathering in complex facilities	Unlimited		2/24/2020	1.00
	Closed	2/25/2020	7/19/2020	0.00
	Open	7/20/2020	10/11/2020	1.00
	Partially closed	10/12/2020		0.50
Seoul cultural center: indoor	Unlimited		4/30/2020	1.00
	Closed	5/1/2020	7/22/2020	0.00
	Open	7/23/2020		1.00
Private tutoring/study/sports/school	Unlimited		8/29/2020	1.00
	Closed	8/30/2020	9/13/2020	0.00
	Open	9/14/2020		1.00
Childcare	Unlimited		2/24/2020	1.00
	Closed	2/25/2020	7/20/2020	0.00
	Open	7/21/2020		1.00
Restaurant	Unlimited		8/29/2020	1.00
	Restricted	8/30/2020	9/13/2020	0.00
	Open	9/14/2020		1.00
Social distancing	Unlimited		3/1/2020	1.00
	Phase 1.0	3/2/2020	8/15/2020	0.80
	Phase 2.0	8/16/2020	11/18/2020	0.40
	Phase 1.5	11/19/2020	12/2/2020	0.60
	Phase 2.5	12/3/2020	12/28/2020	0.00
	Phase 2.0	12/29/2020		0.20

Figure 2 shows the IISI for an individual policy in Sydney and Seoul. The NSW and Seoul governments reactively updated the stringency levels of each policy to reduce the number of local acquisitions in Sydney and Seoul during the first year of the pandemic. The comprehensive inverse SIs for Sydney and Seoul, shown as thick red lines in Figure 2, have been maintained due to continuing a social distancing rule after the initial year of the COVID-19 crisis.



(a) Greater Sydney



(b) Seoul

Figure 2. Inverse individual stringency index (SI) and comprehensive SI for the initial period of the pandemic

4.1.2 *The second set of independent variables: mobility indicators*

This section describes the mobility indicators for Sydney and Seoul. We categorized the overall study areas into metropolitan and CBD regions. Hourly multimodal patronage and road traffic counts were aggregated into four groups, including AM peak, PM peak, off-peak, and daily total. Moreover, the patronage data is grouped into boarding and alighting groups. The road traffic count data is categorized into five levels of road hierarchy: arterial, distributor road, local road, motorway, primary road, and aggregated counts from all roads. The categorization of mobility indicators is shown in Table 7.

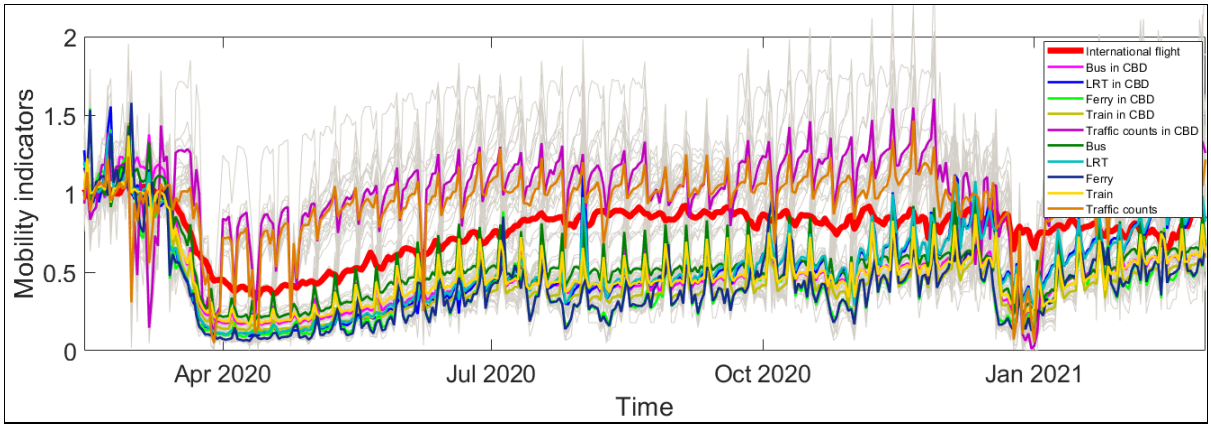
Table 7. Categorization of mobility indicators

Mobility type	Category 1	Category 2	Category 3	Category 4
Transit patronage	• Greater Sydney: train, bus, ferry, LRT • Seoul: metro, bus	Metropolitan CBD	AM peak PM peak Off-peak Total	Boarding Alighting
Road traffic count	All vehicles	Metropolitan CBD	AM peak PM peak Off-peak Total	Arterial Distributor Local road Motorway Primary road Aggregated

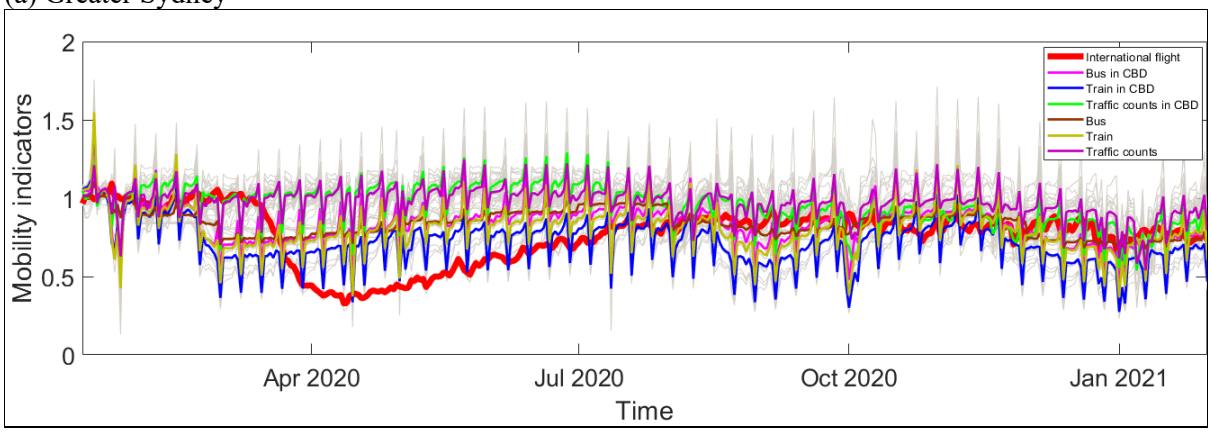
In addition to the local mobility indicators, we added the number of daily international flights worldwide to measure the significance of international transportation. Furthermore, we included 7-day moving averages in the number of international flights to minimize random variations in international travel behavior. The constructed variables are described in Appendix III. We used 93 and 58 variables to explain the causality between infection, mobility, and policy in Sydney and Seoul, respectively.

Mobility patterns before the era of the pandemic were used to normalize the established initial features. We defined the pre-COVID-19 periods as from February 1, 2020, to February 29, 2020, for Sydney modeling and from January 1, 2020, to January 31, 2020, for Seoul modeling. Mobility features were normalized by the average values for the pre-COVID-19 period corresponding features for weekdays and holidays. The normalized mobility features demonstrate how much each mobility feature varied during the pandemic compared to the pre-COVID-19 period. For example, 1.0 for a normalized mobility feature means that the level of the corresponding mobility was maintained even after the first confirmed case of the virus, whereas 0.5 means that the corresponding mobility halved during the COVID-19 crisis and 1.5 of the normalized mobility features means that the corresponding mobility increased by 50% compared to the pre-COVID-19 period.

The normalized mobility features for Greater Sydney and Seoul are given in Figure 3. The gray lines denote all normalized mobility features, and the colored lines represent the average values of normalized mobility features per mode, including international flights.



(a) Greater Sydney



(b) Seoul

Figure 3. Averages of normalized mobility features for each transport mode

Figure 3(a) demonstrates that road traffic counts in the CBD and metropolitan areas of Sydney dramatically decreased compared to the pre-COVID-19 period. They, however, increased to over 1.0, which means larger road traffic counts than during the pre-COVID-19 period before the arrival of the second wave of infections. Patterns of transit patronages constantly retained below 0.5 of mobility indicators due to the continuing social distancing rules in NSW. Figure 3(b) shows that road traffic counts in Seoul slightly increased to over 1.0 of mobility indicators and then decreased below 1.0 after October 2020 as the infection rates of COVID-19 continuously fell for the 2020-2021 winter season. Unlike in Greater Sydney, transport patronage in Seoul did not significantly decline because 1.5 m social distancing was not mandatory on public transport.

4.1.3 *Multivariate time-series features*

In this section, we construct multivariate time-series features. Our feature matrix, **X**, comprises I observations described by J variables for K time-lags. It is represented by the $I \times J \times K$ matrix **X**, whose generic element is $x_{i,j,k}$. In this study, I denotes 381 and 384 days for Greater Sydney and Seoul, respectively, which are the number of days since the first confirmed case. J indicates the 110 and 65 variables, including IISI for policy indicators and mobility indicators, for Greater Sydney and Seoul, respectively. The policy and mobility features dimension allow us to investigate which policies and mobility patterns significantly influence the number of local acquisitions. K is 14 in this study, which represents the time lag from days 1 to 14 before the instant at which the governments officially employ the number of locally transmitted COVID-19 cases as a dependent variable. The time-lag dimension enables us to identify the best moments for imposing various restriction policies to manage rapidly surging infections. The multivariate time-series features are described in Figure 4.

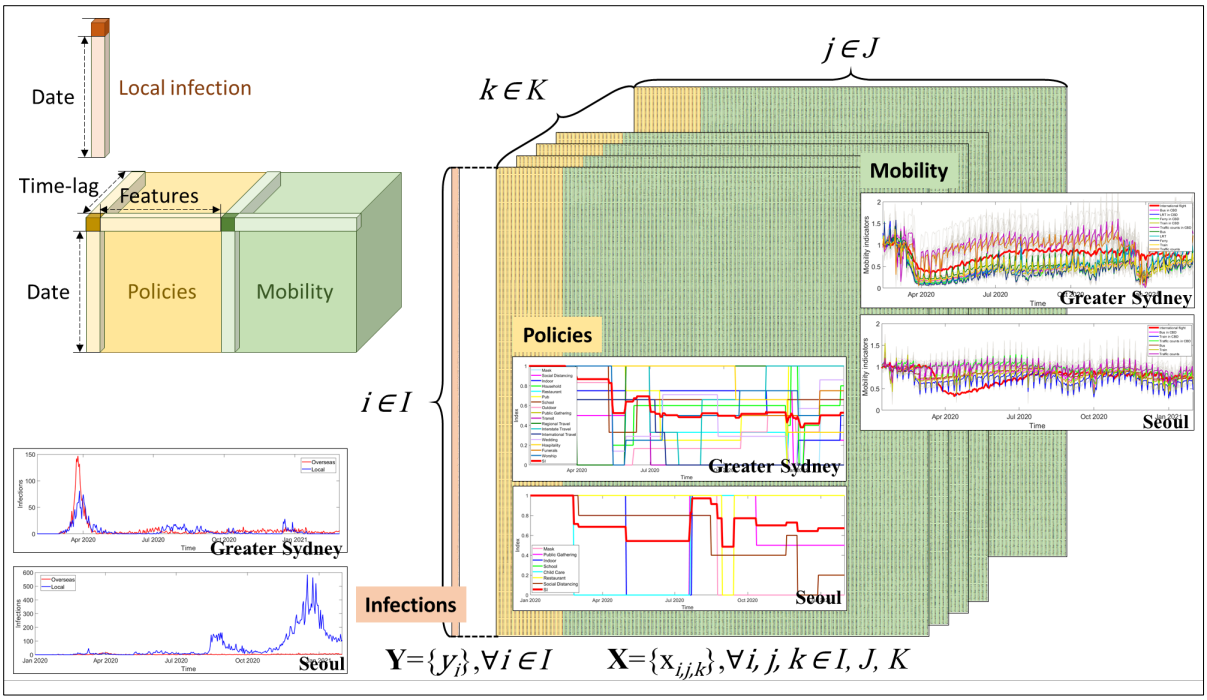


Figure 4. The framework of multidimensional time-series features (improved from Lee and Chen, 2022)

Figure 4 illustrates the framework of the multivariate time-series features linked to Figure 1. The red single column denotes daily infections, a set of dependent variables for model training and validating in the following sections. $\mathbf{Y} = \{y_i\}$, $\forall i \in I$ illustrates the set of dependent variables. The multivariate time-series features consist of IISI for all policies in the yellow-colored groups and

normalized indicators for all mobilities in the green sets. $\mathbf{X} = \{x_{i,j,k}\}$, $\forall i, j, k \in I, J, K$ denotes the set of independent features; 1,540 (110 × 14) and 910 (65 × 14) features are used to predict the daily number of local COVID-19 acquisitions in Sydney and Seoul, respectively.

4.2 Principal component analysis

We used PCA to reduce the dimensions of our constructed datasets, as PCA reconstructs the given data structure to maximize the variance over a set of linear combinations. We modified the traditional PCA mathematical framework described in Abdi and Williams (2010) to establish a set of indexes and parameters for the data structure proposed in this article. Rank L was defined for matrix $\mathbf{X} = \{x_{i,j,k}\}$, $\forall i, j, k \in I, J, K$, where $L \leq \min\{I, J, K\}$. The number of time-lags, K, is combined into the number of the independent variables, J, (i.e., $\mathbf{X} = \{x_{1,1,1}, \ldots, x_{1,1,K}, \ldots, x_{1,J,K}, \ldots, x_{I,J,K}\} \rightarrow X = \{x_{1,1\times1}, \ldots, x_{1,1\times K}, \ldots, x_{1,J\times K}, \ldots, x_{I,J\times K}\}$) to promote the mathematical processes in the PCA and GPR model. The mean of each column was located at zero to center the X columns. The matrix $\mathbf{X}$ has the following singular value decomposition:

$$\mathbf{X} = \mathbf{P}\boldsymbol{\Delta}\mathbf{Q}^{T} \quad (1)$$

where

$\mathbf{P}$ = the $I \times L$ matrix of left singular vectors,

$\mathbf{Q}$ = the $J \times K \times L$ matrix of right singular vectors, and

$\boldsymbol{\Delta}$ = the diagonal matrix of singular values.

The detailed mathematical methods are described in Lee and Chen (2022).

4.3 Gaussian process regressions

The GPR model was utilized to forecast the number of acquired cases given the matrix of features, $\mathbf{X}$, from the PCA. $\mathbf{Y} = \{y_i\}$, $\forall i \in I$ was defined as the set of dependent variables and $\mathbf{X} = \{x_{i,j,k}\}$, $\forall i, j, k \in I, J, K$ was defined as the set of independent variables in the PCA. We formulate a general form of a linear regression model as Equation (2) for the training set $\{(x_{i,j\times k}, y_i), \forall i \in I\}$, where $x_{i,j\times k} \in \mathbb{R}^d$ and $y_i \in \mathbb{R}$.

$$\mathbf{Y} = \mathbf{X}^{T}\boldsymbol{\beta} + \varepsilon, \quad (2)$$

where $\varepsilon \sim N(0, \sigma^2)$. Latent variables, $\boldsymbol{f}(x_{i,j\times k})$, $\forall i \in I$ were derived from a Gaussian process, and explicit basis functions, $\boldsymbol{h}$, to demonstrate the dependent variable in the GPR model. Rasmussen and Williams (2006) presented the overall mathematical equations for GPR, and Lee and Chen (2022) explained the calculations for the latent variables, $\boldsymbol{f}(x_{i,j\times k})$, on pp. 8–9.

5. Results

The results of the established ML approach are illustrated in this section. We used PCA to reduce the dimensions of the data structure and identify the most effective strategies to reduce local infections in Greater Sydney and Seoul. Furthermore, we verified the excellent performance of the proposed method in predicting locally transmitted COVID-19 cases in the two regions compared to other methods.

5.1 Cross-validation

Cross-validation methods (Stone, 1974) were used to assess the accuracy of the devised ML model over the given datasets to avoid overfitting and selection bias. The datasets from Sydney and Seoul were folded and shuffled five times (i.e., we used a five-fold cross-validation method) to reduce their variability and clarify the generalizability of the ML model over future unknown datasets. The root-mean-square-error (RMSE) was used to measure the accuracy of the models and was determined as follows:

$$\varepsilon^{RMSE} = \sqrt{\frac{\sum_{\forall i}(y_i - \hat{y}_i)_i^2}{I}} \tag{3}$$

where $i \in I$ and $\hat{y}_i$ are the predictions. We compared the RMSE of the ML model with those of other models (Table 8). Specifically, we used the regression learner app[13] in MATLAB R2020b to produce six groups of ML approaches: linear regression models, regression trees, support vector machines, Gaussian process regression models, ensembles of regression trees, and neural network regression models. Several configurations were designed, as indicated in the second column in Table 8. Moreover, we examined the accuracy of the predictions with and without PCA in the ML model and in other models for the Seoul and Greater Sydney areasj, respectively.

Table 8. Comparing RMSE values of the constructed model and alternative models

		RMSE			
		Without PCA		With PCA	
Group	Model option	Seoul	Sydney	Seoul	Sydney
Linear regression models	Basic linear	208.340	32.516	41.327[1)]	6.277
	Interactions linear	n/a	n/a	261.280	32.111
	Robust linear	n/a	n/a	42.825	6.806
Regression trees	Fine	40.277	12.068	55.377	9.025
	Medium	39.210	11.480	50.074	8.319
	Coarse	57.510	10.202	65.829	9.885
Support vector machines	Linear	57.985	8.108	62.898	9.834
	Quadratic	37.931	6.081	44.123	8.064
	Cubic	43.847	6.146	39.534	7.169
	Fine Gaussian	101.870	11.818	74.626	11.103
	Medium Gaussian	50.633	8.656	93.439	11.702
	Coarse Gaussian	80.097	10.826	109.100	12.514
Gaussian process regression models	Rational quadratic	37.107	5.500	35.550	6.165
	Squared exponential	37.205	11.673	36.588	6.165
	Matern 5/2	40.514	5.460	35.624	6.096
	Exponential[2)]	32.477	5.448	30.059[3)]	4.508
Ensembles of regression trees	Boosted	29.270	7.662	38.387	6.791
	Bagged	31.371	8.012	50.794	7.756
Neural network regression models	Narrow	34.546	6.247	44.895	6.049
	Medium	37.282	6.007	41.855	5.915
	Wide	36.128	5.882	37.331	5.859

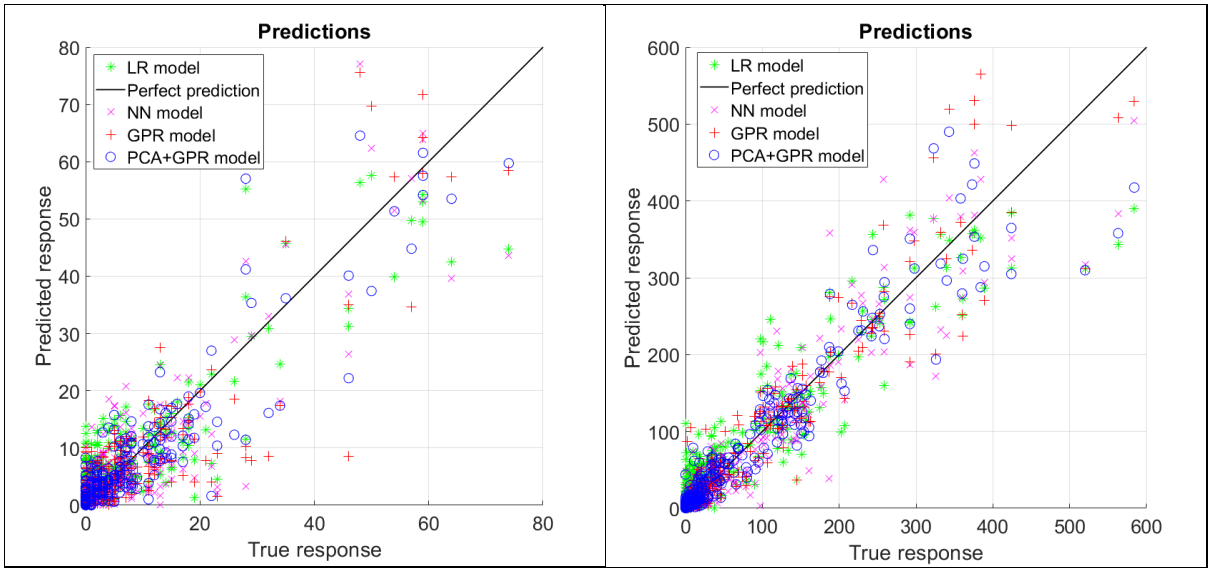
[13] https://au.mathworks.com/help/stats/choose-regression-model-options.html

Bilayered	35.970	5.813	45.151	6.137
Trilayered	<u>34.197</u>	<u>5.814</u>	43.769	8.758

[1] The bold notations present the proposed GPR model with exponential kernel functions in both areas.
[2] The underlined notations present the RMSE values of three models illustrated in Figures 5-6 and Table 9.
[3] The red notations present the RMSE values of the integrated PCA and GPR models in both areas.

As shown in Table 8, the integrated PCA and GPR models outperformed the other models and achieved the lowest RMSE (30.059 and 4.508 for Seoul and Sydney, respectively). The calculation times for the PCA+GPR model were 411.65 s and 204.58 s for Greater Sydney and Seoul, respectively. These results indicate that the dimensions of the original data structure were decreased, which improved the predication accuracy, and the integrated model was superior to the other ML algorithms in forecasting the number of local acquisitions of COVID-19.

The GPR model with exponential kernel functions also performed well, as indicated by the low RMSEs (32.477 and 5.448 for Seoul and Greater Sydney, respectively). The RMSE of the tri-layered neural network regression model without PCA (34.197 and 5.814 for Seoul and Greater Sydney, respectively) were lower than those of the linear regression models, regression trees, support vector machines, and ensembles of regression trees, which did not accurately predict the number of daily COVID-19 infections from the datasets. Figure 5 presents a comparison of the performances of the four best models: the basic linear regression model (the reference model), the tri-layered neural network regression model, the Gaussian process regression model, and out developed PCA+GPR model.



(a) Greater Sydney (b) Seoul

Figure 5. Scatter plots of predictions by the developed models. LR = basic linear regression, NN = trilayered neural network regression, GPR = Gaussian process regression, PCA+GPR = integrated PCA and GPR

In Figures 5(a) and (b), estimates generated by our framework are densely distributed on the periphery of the ideal prediction lines. In contrast, the other models produce estimates that are widely scattered around the graphs as the size of the predictions increases. Our methods produced R-squared values of 0.848 and 0.911 for Sydney and Seoul, respectively. The statistical results of the selected four models are presented in Table 9.

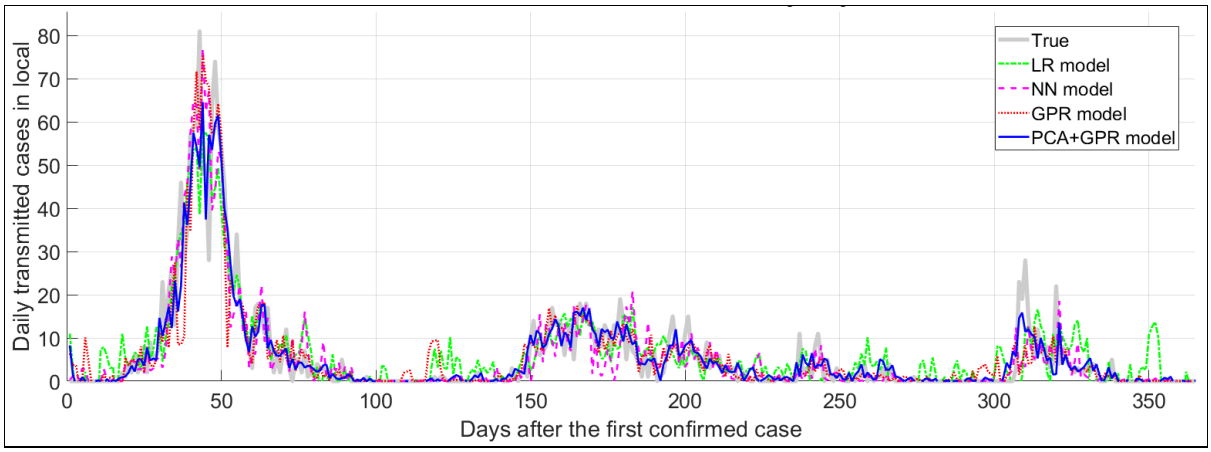
Table 9. Statistical results of the four best-performing models

	Greater Sydney	Seoul

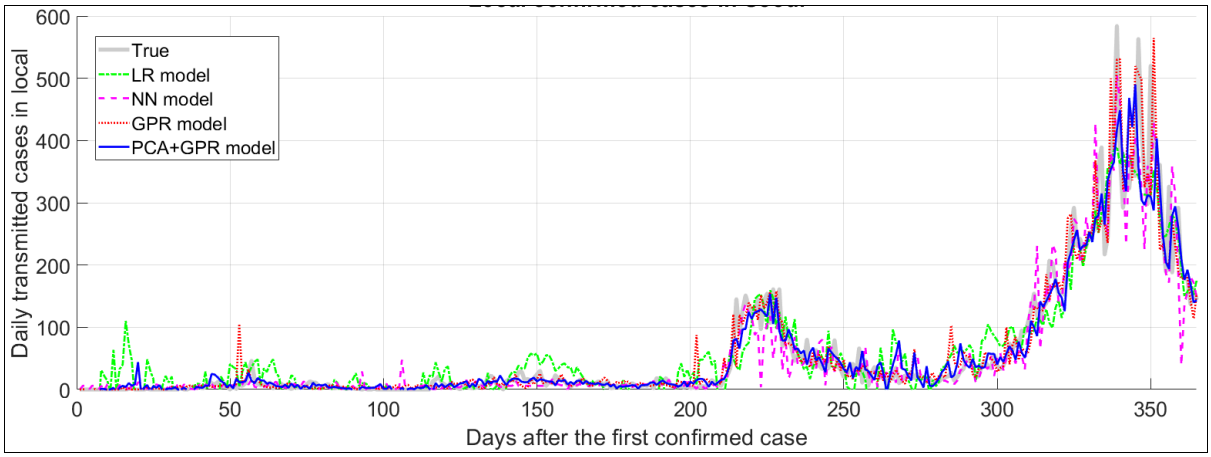
		Observation	LR	NN	GPR	PCA+GPR	Observation	LR	NN	GPR	PCA+GPR
MOE	RMSE	n/a	6.28	5.81	5.76	**4.51**[1)]	n/a	41.33	34.20	32.48	**30.06**
	R-squared	n/a	0.708	0.755	0.760	**0.848**	n/a	0.831	0.885	0.901	**0.911**
	Pearson Correlation	n/a	0.84	0.87	0.87	**0.92**	n/a	0.91	0.94	0.95	**0.95**
Stats	Observation			381					384		
	Variables	n/a	45	1,540	1,540	45	n/a	32	910	910	32
	Mean	5.71	5.57	5.37	5.64	5.55	60.68	59.04	57.35	62.03	59.86
	Standard error	0.59	0.54	0.57	0.57	0.53	5.16	4.83	4.92	5.27	4.81
	Median	1.00	3.27	1.05	1.84	1.72	14.00	31.08	9.38	13.25	14.37
	Standard deviation	11.54	10.49	11.12	11.18	10.36	100.69	94.29	96.07	102.82	93.93
	Sample variance	133.19	110.01	123.65	124.98	107.39	10139.2	8891.13	9229.42	10571.9	8823.15
	Kurtosis	14.80	8.51	15.76	19.04	13.88	6.15	2.61	4.90	7.24	4.81
	Skewness	3.57	2.53	3.74	4.13	3.52	2.41	1.70	2.29	2.58	2.24
	Range	81.00	68.60	78.02	77.66	65.98	584.00	469.85	506.83	566.03	497.61
	Minimum	0.00	-11.06	-0.97	-2.10	-1.43	0.00	-79.50	-2.34	-1.35	-7.59
	Maximum	81.00	57.54	77.06	75.56	64.55	584.00	390.36	504.48	564.68	490.02
	Sum	2177.0	2123.4	2045.9	2149.3	2115.7	23120.0	22494.9	21851.6	23634.7	22806.2
	Largest	81.00	57.54	77.06	75.56	64.55	584.00	390.36	504.48	564.68	490.02
	Smallest	0.00	-11.06	-0.97	-2.10	-1.43	0.00	-79.50	-2.34	-1.35	-7.59

[1)]The bold and red notations indicate the measure of effectiveness (MOE) values of the proposed model.

The RMSEs of the established model (4.51 for Greater Sydney and 30.06 for Seoul) are smaller than those of the other models. Moreover, compared with the R-squared values of the other models, those of the PCA+GPR model (0.848 in Greater Sydney and 0.911 in Seoul) are higher, which demonstrates its superiority. The high Pearson correlations of the PCA+GPR model (0.92 for Greater Sydney and 0.95 for Seoul) indicate that the values predicted by the PCA+GPR model are strongly correlated with the observed number of acquisition cases in Sydney and Seoul. Figure 6 compares the observations and estimates of daily COVID-19 cases with the model validations.



(a) Greater Sydney



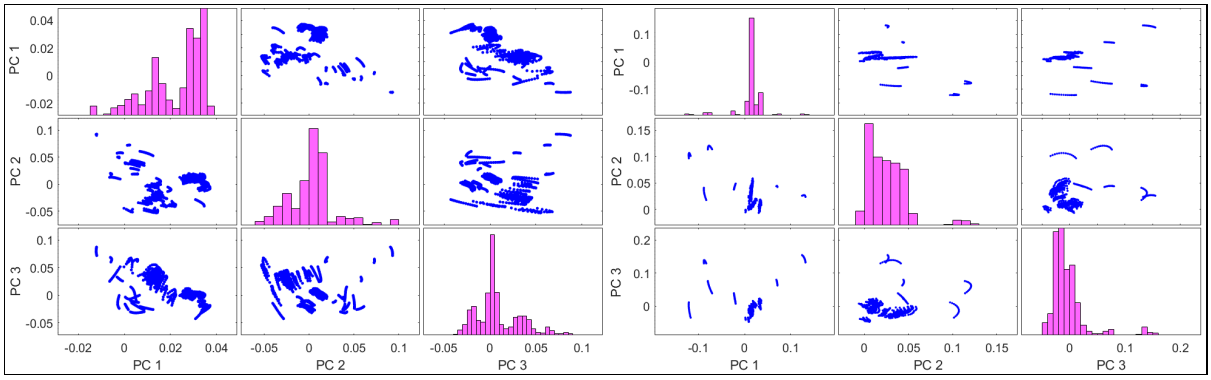
(b) Seoul

Figure 6. Daily locally transmitted COVID-19 cases, LR = basic linear regression, NN = tri-layered neural network regression, GPR = Gaussian process regression, PCA+GPR = integrated PCA and GPR

Our framework describes the global waves of COVID-19 case acquisitions and locally fluctuating waves of acquisitions in both regions during the first year of the COVID-19 pandemic with minor errors. The estimates yielded by the PCA+GPR model are identical to the observations in all periods. The linear regression model tends to overestimate the infections when the number of infections is small in both regions, and the neural network model underestimates the number of infections in the large waves in both areas. The comparison results demonstrate that our data processing, normalization, and estimation framework adequately estimates the number of acquired COVID-19 cases in Seoul and Sydney.

5.2 Principal component analysis

We used MATLAB R2020b packages to find the matrix **Q** based on the proposed data structure. The PCA computes 127 and 97 components for Greater Sydney and Seoul, respectively, to explain 99% of the inertia of the data structure. The top 10 principal components explain 81.8% of the variance (42.1%, 17.1%, 6.3%, 4.4%, 2.8%, 2.7%, 1.9%, 1.7%, 1.5%, and 1.3%, respectively) for Greater Sydney. For Seoul, the top 10 principal components explain 81.9% of the inertia (34.5%, 17.9%, 6.2%, 4.1%, 4.0%, 3.5%, 3.5%, 3.4%, 2.5%, and 2.3%, respectively). The scatter plots in Figure 7 illustrate how the top three principal components, explaining 65.5% and 58.6% of the variance in Greater Sydney and Seoul, respectively, capture the inertia of the given data structure.



(a) Greater Sydney (1,540 features) (b) Seoul (910 features)

Figure 7. The inertia of the data structure explained by the top three principal components

The 1,540 features for Greater Sydney and 910 features for Seoul are represented in a series of biplots by a vector in Figure 7. The first principal component explains the largest variability. We conducted a covariance analysis of the transformed variables, and the covariance matrixes are presented in Table 10. None of the principal components for Greater Sydney and Seoul were larger than zero except for those presented as diagonal values.

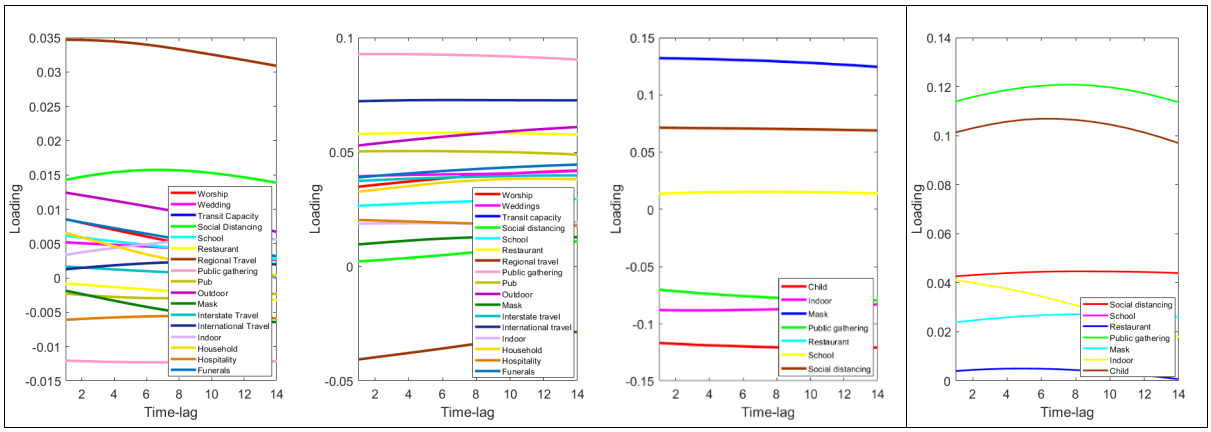
Table 10 (a). Covariance table of the top-10 principal components for Greater Sydney

	PCA01	PCA02	PCA03	PCA04	PCA05	PCA06	PCA07	PCA08	PCA09	PCA10
PCA01	48.21									
PCA02		19.62								
PCA03			7.19							
PCA04				5.06						
PCA05					3.21					
PCA06						3.10				
PCA07							2.16			
PCA08								1.90		
PCA09									1.69	
PCA10										1.54

Table 10 (b). Covariance table of the top-10 principal components for Seoul

	PCA01	PCA02	PCA03	PCA04	PCA05	PCA06	PCA07	PCA08	PCA09	PCA10
PCA01	11.01									
PCA02		5.69								
PCA03			1.99							
PCA04				1.30						
PCA05					1.27					
PCA06						1.11				
PCA07							1.10			
PCA08								1.07		
PCA09									0.80	
PCA10										0.73

We decomposed the first principal component to identify which and when mobility and policy indicators effectively reduced the number of locally transmitted cases. We compared the loading values in matrix **Q** to comprehensively understand the distributed loadings of each feature, although it is impossible to define causal relationships between features through the PCA. Figure 8 presents loading fluctuations in policy indicators with regard to a time-lag in the first and second principal components.

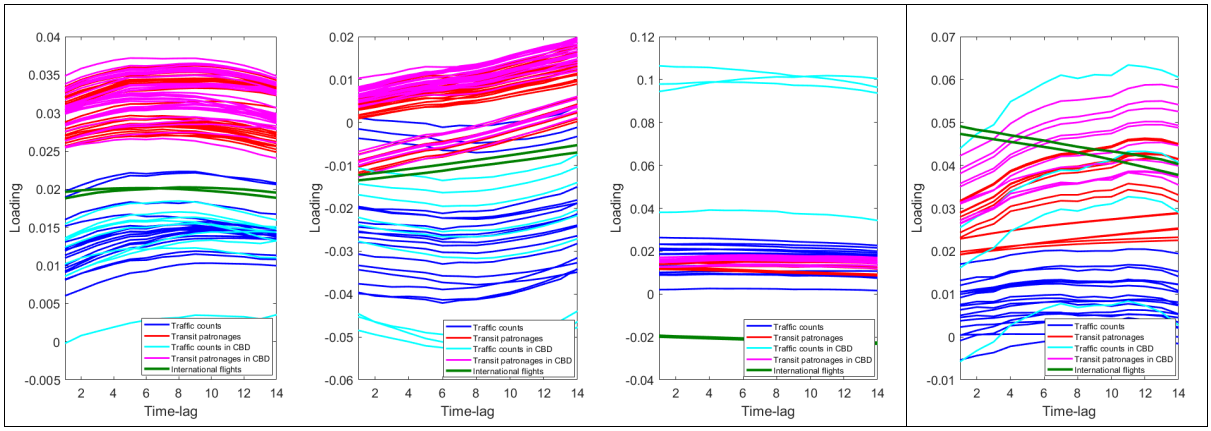


(a) PC 1 for Sydney (b) PC 2 for Sydney (c) PC 1 for Seoul (d) PC 2 for Seoul

Figure 8. Loading fluctuations in the policy indicators

Figure 8(a) shows that the regional travel ban, the social distancing rule, and the capacity limitation of outdoor activities significantly contributed to building the first principal component for Greater Sydney. These three policies were relatively more effective at reducing the number of local transmissions in Sydney than other policies. According to the loading tendency, the best timings for the regional travel ban, the social distancing rule, and the limitation of outdoor activities were one day, one day, and seven days prior to the base date, respectively, even if the loadings in these three policies did not change significantly as the time-lag increased. Figure 8(b) shows that capacity limitations on public transport and public gatherings, and the international travel ban contributed to the second principal component in Greater Sydney under the constraint of being orthogonal to the first principal component. The best timings for these three policies were two days, two days, and six days before the base date.

Figure 8(c) shows that the mask policy and social distancing rules effectively reduce the number of locally transmitted cases in Seoul, as indicated by the first principal component. The best timing for both policies is one day before the base date. Figure 8(d) shows that restrictions on public gatherings and child care significantly contributed to building the second principal component. Their best timings were eight and six days prior to the base date, respectively. Figure 9 illustrates the loading fluctuations for the mobility indicators.



(a) PC 1 for Sydney (b) PC 2 for Sydney (c) PC 1 for Seoul (d) PC 2 for Seoul

Figure 9. Loading fluctuations in the mobility indicators

Figure 9(a) shows that a decline in transit patronage in the CBD for five to ten days before the base date significantly decreased the number of locally acquired COVID-19 cases in Sydney.

Furthermore, decreases in transit patronage eight to nine days prior to the base date in the metropolitan area reduced the locally confirmed cases. Decreased road traffic counts had an insignificant impact on the number of infections in Greater Sydney. In Figure 12(b), the loading distribution in the second principal component is similar to that in the first principal component. Transit patronage patterns had a more significant influence on decreasing the number of infections than road traffic counts in Greater Sydney. In contrast, road traffic counts had a negative impact on the number of infections. We can assume that vehicular traffic increased because of the capacity limitations on public transport and personal preferences during the COVID-19 crisis in Greater Sydney.

As shown in Figure 9(c), reduced road traffic counts in the CBD for one to four days prior to the base date significantly decreased the number of local acquisitions of the COVID-19 in Seoul. In addition, declining transit patronage did not significantly reduce infections in the first principal component. Figure 9(d) shows that the second principal component mainly explains the influences of reduced transit patronages on infections. The reduced transit patronage in the CBD for 11–14 days before the base date significantly contributed to building the second principal component.

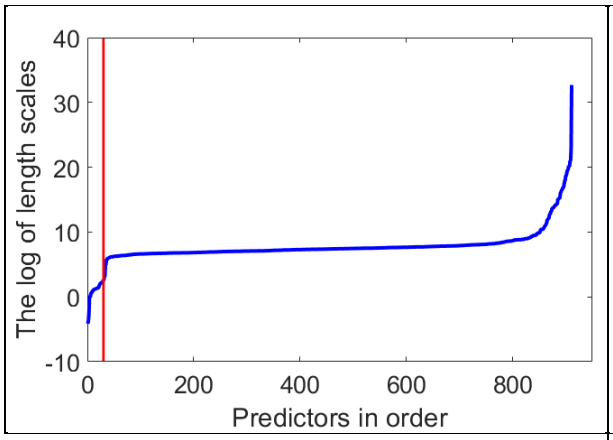
In summary, we used PCA to reduce the dimensions of the data structure. Moreover, we examined the loading values in the matrix **Q** to understand which and when policy and mobility indicators effectively reduced the number of COVID-19 transmissions in Sydney and Seoul. Our policy analysis suggests that the regional travel ban, the social distancing rules, and the limitation of outdoor activities imposed respectively one day, one day, and seven days before the base date were effective in Greater Sydney. In Seoul, the mask policy and the social distancing rules imposed one day before the base date were the most effective for reducing the number of locally transmitted cases. Our mobility analysis suggests that reduced transit patronage was the most critical mobility factor in reducing COVID-19 infections in Greater Sydney and Seoul.

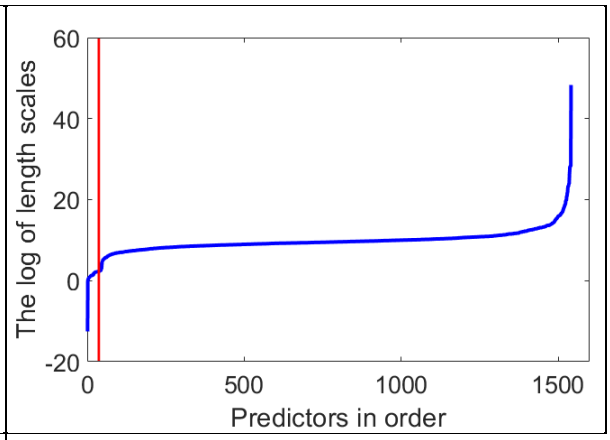
5.3 Gaussian process regressions

Using the Gaussian process regression model with the exponential option and without PCA, as illustrated in Table 8, we obtain the length scales σ_m and signal standard deviation σ_f. The set of optimized length scales based on the log of length scales enable determination of the features that decreased the number of locally transmitted COVID-19 cases in our study areas. The number of predictors below the inflection point in the graph of the log of length scales was 45 for Greater Sydney and 32 for Seoul. Complete lists of predictors that are adequate to explain infections according to the log of length scales are presented in Appendixes IV and V. Table 11 presents the log of length scales with the ten features most effective at reducing the number of infections in each region.

Table 11. A graph of the log of length scales of the top ten features most effective at reducing infection

Greater Sydney	Seoul





P/M	Feature	Time-lag	Log of length scale	P/M	Feature	Time-lag	Log of length scale
Policy	School closure	3	-12.54	Policy	Public gathering	14	-4.14
Policy	International travel	7	0.26	Policy	Public gathering	13	-3.61
Policy	International travel	3	0.26	Policy	Public gathering	12	-2.60
Policy	International travel	11	0.51	Policy	Childcare	12	0.04
Policy	International travel	13	0.68	Policy	Childcare	13	0.06
Policy	Regional travel	9	0.73	Policy	Social distancing	4	0.22
Policy	Regional travel	5	0.75	Policy	Mask	7	0.57
Policy	Social distancing	14	0.96	Policy	Social distancing	13	0.65
Policy	Social distancing	4	0.97	Policy	Social distancing	9	0.75
Policy	Social distancing	5	1.15	Policy	Social distancing	6	0.93

Table 11 shows that the school closure restriction, the international travel ban, the regional travel ban, and the social distancing rule effectively reduced the number of locally transmitted COVID-19 cases in Greater Sydney. The most effective timings for all policies are scattered between 3 and 14 days before the base date. In Seoul, the public gathering restriction, the childcare closure restriction, the social distancing rule, and the mask rule effectively reduced the number of infections. Their most effective dates are distributed between 4 and 14 days before the base date. Although mobility indicators are included in the complete lists of features provided in Appendixes I and II, their impacts were not significant relative to the influences of policy indicators.

6. Conclusions and discussion

In this study, we investigate the spatiotemporal relationships between COVID-19 infections, urban travel restriction policies, and multimodal travel behavior patterns for the original strain of the virus during the first year of the COVID-19 crisis before vaccinations. We established time-series multidimensional features involving daily local, quarantined, and international infections and death rates from the acquisition of COVID-19, several restriction levels of policies related to COVID-19, and daily and hourly travel patterns in multiple transport modes in selected areas, Greater Sydney and Seoul. Specifically, the ML approach describes spatiotemporal causality between restriction policies, mobility, and the acquisition of COVID-19. The PCA was employed to discover effective COVID-19 related policies and travel patterns at the first level. The PCA was also used to demonstrate the most effective timings for implementing such policies to reduce the number of transmitted COVID-19 cases in Sydney and Seoul.

The contributions of this study can be summarized as follows. 1) We subdivided and normalized policies that restrict travel accessibility and mobility and multimodal travel patterns to investigate the spatiotemporal aspects of each policy for the first year of the COVID-19 pandemics. 2) We calibrated and validated the ML approach using open data collected from two metropolitan areas, Greater Sydney and Seoul, to address future pandemic scenarios in other metropolitan areas. 3) PCA and GPR were integrated to identify the effective policies and the timings at which they should be imposed in response to the pandemic in the initial period to control instantaneously increasing rates of COVID-19 acquisitions. 4) Our framework could identify the specific policies and mobility patterns that considerably affect the number of locally transmitted COVID-19 cases.

The technical contributions of our ML approach can be summarized as follows. Open-source datasets were used to construct multi-dimensional time-series datasets incorporating mobility and restriction policy-related features. Because of the immense size of the data features, we applied PCA to decrease the dimensions of the data structure. PCA also allowed us to analyzing loadings in the first and second principal components, thereby enabling us to determine the policy and mobility indicators and the corresponding implementation timings that helped to decrease the number of local COVID-19 transmissions in Sydney and Seoul, respectively. Our policy analysis indicated that the regional travel ban, physical distancing rule, and limitations on outdoor activities imposed 1 day, 1 day, and 7 days before the base date, respectively, were effective in Greater Sydney. The mask policy and physical distancing rules imposed 1 day before the base date were the most effective policies for reducing the number of locally transmitted cases in Seoul. The mobility analysis indicated that reduced transit patronage was effective at reducing the spread of COVID-19 infections in Greater Sydney and Seoul.

A GPR model was used to define the effective restriction policy and mobility indicators and their most effective imposition dates in Greater Sydney and Seoul. The school closure, international travel ban, regional travel ban, and physical distancing rule effectively reduced the number of local acquisitions in Sydney. The most effective timings for all of the policies were ranged from 3 to 14 days before the base date. In Seoul, the public gathering restriction, childcare closure, physical distancing rule, and mask rule effectively reduced the number of infections. The effective dates were ranged from 4 and 14 days prior to the base date. Although mobility indicators are included in the lists of features provided in Appendixes I and II, their effects were not significant relative to the influence of the policy indicators.

The proposed method paves the way for the application of integrated ML algorithms to effectively impose restriction policies in response to the spread of COVID-19 in temporal and spatial dimensions. Moreover, we demonstrated the time-lag findings to decide when and which specific restriction policies should be imposed to respond to upcoming pandemic shocks depending on the demographic and geographic conditions of the target areas before introducing vaccinations. For example, sparsely populated countries with a vast geographic area, such as Australia, could impose an international travel ban, regional travel ban, and social distancing rules, in that order. In contrast, densely populated countries with highly developed technology infrastructure, such as Korea, should impose public gathering restrictions, childcare restrictions, mask rules, and social distancing rules, in that order.

Our findings, based on PCA and GPR modeling, have far-reaching policy implications. We identified the largest proportion of variance explained in each PC in the PCA by interpreting PCA coefficients. Moreover, based on the GPR framework, we demonstrated the set of optimized length scales using the log of length scales, which helped to identify the features that contributed to decreasing

the number of locally transmitted COVID-19 cases in our study areas. First, our methods can be used for formulating restriction policies for other regions with population densities different to those of the cities selected in this case study. This generalizability can help policymakers to better implement and enforce effective restriction policies. Second, our finding contributes to evidence-based policymaking. Restriction policies inherently create socio-economic challenges. By increasing the objectivity and effectiveness of restriction policies, our framework can help to alleviate any socio-economic challenges and increase the probability of achieving policy goals. Future research can be focused on the addition of vaccination indicators into the model to construct long-term policies and mobility responses to COVID-19 variants. Moreover, data pertaining to vaccines, daily percentages of citizens vaccinated, and daily infections of diverse variants can be added to the established multivariate time-series datasets.

Acknowledgement

This work was supported by the 2023 Research Fund of the University of Seoul for Seunghyeon Lee. This work was supported by the Guangdong - Hong Kong - Macau Joint Laboratory Program of the 2020 Guangdong New Innovative Strategic Research Fund, Guangdong Science and Technology Department (Project 2020B1212030009). The third author was also supported by the Francis S Y Bong Professorship in Engineering.

References

Abdi, H. and Williams, L.J., 2010. Principal component analysis. *Wiley Interdisciplinary Reviews: Computational Statistics*, *2*(4), pp.433-459.

Anke, J., Francke, A., Schaefer, L.M. and Petzoldt, T., 2021. Impact of SARS-CoV-2 on the mobility behaviour in Germany. *European Transport Research Review*, *13*(1), pp.1-13.

Ashrafian, H. and Darzi, A., 2018. Transforming health policy through machine learning. *PLoS Medicine, 15*(11), p.e1002692.

Aziz, S. and Dowling, M., 2019. Machine learning and AI for risk management. In T. Lynn, J.G. Mooney, P. Rasati, M. Cummins, M. (eds), *Disrupting finance,* Palgrave Pivot, Cham, pp. 33-50

Badr, H.S., Du, H., Marshall, M., Dong, E., Squire, M.M. and Gardner, L.M., 2020. Association between mobility patterns and COVID-19 transmission in the USA: A mathematical modelling study. *The Lancet Infectious Diseases*, *20*(11), pp.1247-1254.

Beck, M.J. and Hensher, D.A., 2020. Insights into the impact of COVID-19 on household travel and activities in Australia: The early days of easing restrictions. *Transport Policy*, *99*, pp.95-119.

Bian, Z., Zuo, F., Gao, J., Chen, Y., Venkata, S.S.C.P., Bernardes, S.D., Ozbay, K., Ban, X.J. and Wang, J., 2021. Time lag effects of COVID-19 policies on transportation systems: A comparative study of New York City and Seattle. *Transportation Research Part A: Policy and Practice*, *145*, pp.269-283.

Cantarella, G.E. and de Luca, S., 2005. Multilayer feedforward networks for transportation mode choice analysis: An analysis and a comparison with random utility models. *Transportation Research Part C: Emerging Technologies*, *13*(2), pp.121-155.

Chan, H.Y., Chen, A., Ma, W., Sze, N.N. and Liu, X., 2021. COVID-19, community response, public policy, and travel patterns: A tale of Hong Kong. *Transport Policy*, 106, pp.173-184.

Chen, W., Liu, X., Chen., X., Cheng, L., Wang, K. and Chen, J., 2022. Exploring year-to-year changes in station-based bike sharing commuter behaviors with smart card data. *Travel Behaviour and Society*, 28, pp. 75-89.

Cho, S.H. and Park, H.C., 2021. Exploring the behaviour change of crowding impedance on public transit due to COVID-19 pandemic: before and after comparison. *Transportation Letters*, *13*(5-6), pp.367-374.

Coyle, D. and Weller, A., 2020. "Explaining" machine learning reveals policy challenges. *Science*, *368*(6498), pp.1433-1434.

de Rubens, G.Z., 2019. Who will buy electric vehicles after early adopters? Using machine learning to identify the electric vehicle mainstream market. *Energy, 172*, pp.243-254.

De-Toledo, K.P., O'Hern, S. and Koppel, S. 2022. Travel behaviour change research: A scientometric review and content analysis. *Travel Behaviour and Society*, 28, pp. 141-154.

Doherty, S.T. and Mohammadian, A., 2003. Application of artificial neural network models to activity scheduling time horizon. *Transportation research record*, *1854*(1), pp.43-49.

Firebanks-Quevedo, D., Planas, J., Buckingham, K., Taylor, C., Silva, D., Naydenova, G. and Zamora-Cristales, R., 2022. Using machine learning to identify incentives in forestry policy: Towards a new paradigm in policy analysis. *Forest Policy and Economics, 134,* p.102624.

Gammelli, D., Peled, I., Rodrigues, F., Pacino, D., Kurtaran, H.A. and Pereira, F.C., 2020. Estimating latent demand of shared mobility through censored Gaussian processes. *Transportation Research Part C: Emerging Technologies*, *120*, p.102775.

García, J.J., Urena, J., Mazo, M., Espinosa, F., Hernández, Á., Losada, C., Jiménez, A., De Marziani, C., Álvarez, F. and García, E., 2010. Sensory system for obstacle detection on high-speed lines. *Transportation Research Part C: Emerging Technologies*, *18*(4), pp.536-553.

Gibbs, H., Liu, Y., Pearson, C.A., Jarvis, C.I., Grundy, C., Quilty, B.J., Diamond, C. and Eggo, R.M., 2020. Changing travel patterns in China during the early stages of the COVID-19 pandemic. *Nature Communications*, *11*(1), pp.1-9.

Golshani, N., Shabanpour, R., Mahmoudifard, S.M., Derrible, S. and Mohammadian, A., 2018. Modeling travel mode and timing decisions: Comparison of artificial neural networks and copula-based joint model. *Travel Behaviour and Society*, *10*, pp.21-32.

Goolsbee, A. and Syverson, C., 2021. Fear, lockdown, and diversion: Comparing drivers of pandemic economic decline 2020. *Journal of Public Economics*, *193*, p.104311.

Guardiola, I.G., Leon, T. and Mallor, F., 2014. A functional approach to monitor and recognise patterns of daily traffic profiles. *Transportation Research Part B: Methodological*, *65*, pp.119-136.

Guzman, L.A., Arellana, J., Oviedo, D. and Aristiz´abal, C.A.M. 2021. COVID-19, activity and mobility patterns in Bogot´a. Are we ready for a '15-minute city'? *Travel Behaviour and Society*, 24, pp. 245-256.

Hagenauer, J. and Helbich, M., 2017. A comparative study of machine learning classifiers for modeling travel mode choice. *Expert Systems with Applications*, *78*, pp.273-282.

Hale, T., Petherick, A., Phillips, T., Webster, S., 2020. *Variation in Government Responses to COVID-19 (No. 31).* Blavatnik school of government working paper.

Han, E., Tan, M.M.J., Turk, E., Sridhar, D., Leung, G.M., Shibuya, K., Asgari, N., Oh, J., García-Basteiro, A.L., Hanefeld, J. and Cook, A.R., 2020. Lessons learnt from easing COVID-19 restrictions: An analysis of countries and regions in Asia Pacific and Europe. *The Lancet*, 369 (10261), pp.1525-1534.

Hensher, D.A., Beck, M.J. and Wei, E., 2021. Working from home and its implications for strategic transport modelling based on the early days of the COVID-19 pandemic. *Transportation Research Part A: Policy and Practice*, *148*, pp.64-78.

Hu, S., Xiong, C., Yang, M., Younes, H., Luo, W. and Zhang, L., 2021. A big-data driven approach to analysing and modeling human mobility trend under non-pharmaceutical interventions during COVID-19 pandemic. *Transportation Research Part C: Emerging Technologies*, *124*, p.102955.

International Monetary Fund. (2020). Chapter 2. The great lockdown: Dissecting the economic effects. pp. 65-84, In *World Economic Outlook: A Long and Difficult Ascent*. Washington, DC, October. https://www.imf.org/en/Publications/WEO/Issues/2020/09/30/world-economic-outlook-october-2020#Chapter%202:%20The%20Great%20Lockdown,%20Dissecting%20The%20Economic%20Effects

Jabari, S.E. and Liu, H.X., 2013. A stochastic model of traffic flow: Gaussian approximation and estimation. *Transportation Research Part B: Methodological*, *47*, pp.15-41.

Jenelius, E. and Koutsopoulos, H.N., 2017. Urban network travel time prediction based on a probabilistic principal component analysis model of probe data. *IEEE Transactions on Intelligent Transportation Systems*, *19*(2), pp.436-445.

Ku, D., Yeon, C., Lee, S., Lee, K., Hwang, K., Li, Y.C. and Wong, S.C., 2021. Safe traveling in public transport amid COVID-19. *Science advances*, *7*(43), p.eabg3691.

Koushik, A.N., Manoj, M. and Nezamuddin, N., 2020. Machine learning applications in activity-travel behaviour research: a review. *Transport Reviews, 40*(3), pp.288-311.

Kumar, P. and Khani, A., 2020. Evaluating special event transit demand: A robust principal component analysis approach. *IEEE Transactions on Intelligent Transportation Systems*, pp.1-13.

Lasisi, A. and Attoh-Okine, N., 2018. Principal components analysis and track quality index: A machine learning approach. *Transportation Research Part C: Emerging Technologies*, *91*, pp.230-248.

Laval, J.A., Toth, C.S. and Zhou, Y., 2014. A parsimonious model for the formation of oscillations in car-following models. *Transportation Research Part B: Methodological*, *70*, pp.228-238.

Lee, S. and Chen, F., 2022, February. Identifying the Effective Restriction and Vaccination Policies During the COVID-19 Crisis in Sydney: A Machine Learning Approach. In *Australasian Joint Conference on Artificial Intelligence* (pp. 356-367). Springer, Cham.

Lee, S., Ngoduy, D. and Keyvan-Ekbatani, M., 2019. Integrated deep learning and stochastic car-following model for traffic dynamics on multi-lane freeways. *Transportation Research Part C: Emerging Technologies*, *106*, pp.360-377.

Lee, S., Ryu, I., Ngoduy, D., Hoang, N.H. and Choi, K., 2021. A stochastic behaviour model of a personal mobility under heterogeneous low-carbon traffic flow. *Transportation Research Part C: Emerging Technologies*, *128*, p.103163.

Lin, H.Z., Lo, H.P. and Chen, X.J., 2009. Lifestyle classifications with and without activity-travel patterns. *Transportation Research Part A: Policy and Practice*, *43*(6), pp.626-638.

Liu, J., Han, K., Chen, X.M. and Ong, G.P., 2019. Spatial-temporal inference of urban traffic emissions based on taxi trajectories and multi-source urban data. *Transportation Research Part C: Emerging Technologies*, *106*, pp.145-165.

Liu, S. and Yamamoto, T., 2022. Role of stay-at-home requests and travel restrictions in preventing the spread of COVID-19 in Japan. *Transportation Research Part A: Policy and Practice*, *159*, pp.1-16.

Mao, Y. 2021. Political institutions, state capacity, and crisis management: A comparison of China and South Korea. *International Political Science Review*, 0192512121994026.

Mitchell, T.M. 1997. *Machine learning*. McGraw-hill, New York:

Murça, M.C.R. and Hansman, R.J., 2018. Predicting and planning airport acceptance rates in metroplex systems for improved traffic flow management decision support. *Transportation Research Part C: Emerging Technologies*, *97*, pp.301-323.

Ngoduy, D., Lee, S., Treiber, M., Keyvan-Ekbatani, M. and Vu, H.L., 2019. Langevin method for a continuous stochastic car-following model and its stability conditions. *Transportation Research Part C: Emerging Technologies*, *105*, pp.599-610.

Nouvellet, P., Bhatia, S., Cori, A., Ainslie, K.E., Baguelin, M., Bhatt, S., Boonyasiri, A., Brazeau, N.F., Cattarino, L., Cooper, L.V. and Coupland, H., 2021. Reduction in mobility and COVID-19 transmission. *Nature Communications*, *12*(1), pp.1-9.

Pawar, D.S., Yadav, A.K., Choudhary, P. and Velaga, N.R. 2021. Modelling work- and non-work-based trip patterns during transition to lockdown period of COVID-19 pandemic in India. *Travel Behaviour and Society*, 24, pp. 46-56.

Pincet, A., Okabe, S., Pawelczyk, M., 2019. Linking aid to the sustainable development goals – a machine learning approach. *OECD development cooperation working paper 52*.

Qu, L., Li, L., Zhang, Y. and Hu, J., 2009. PPCA-based missing data imputation for traffic flow volume: A systematical approach. *IEEE Transactions on Intelligent Transportation Systems*, *10*(3), pp.512-522.

Rana, P. and Miller, D.C., 2019. Machine learning to analyze the social-ecological impacts of natural resource policy: insights from community forest management in the Indian Himalaya. *Environmental Research Letters, 14*(2), p.024008.

Rasmussen, C.E. and Williams, C.K., 2006. *Gaussian Processes for Machine Learning* Cambridge, MA: the MIT Press.

Rodrigues, F. and Pereira, F.C., 2018. Heteroscedastic Gaussian processes for uncertainty modeling in large-scale crowdsourced traffic data. *Transportation Research Part C: Emerging technologies*, *95*, pp.636-651.

Saffari, E., Yildirimoglu, M. and Hickman, M., 2020. A methodology for identifying critical links and estimating macroscopic fundamental diagram in large-scale urban networks. *Transportation Research Part C: Emerging Technologies*, *119*, p.102743.

Shin, M.O., Oh, G.M., Kim, S.W. and Seo, S.W., 2017. Real-time and accurate segmentation of 3-D point clouds based on Gaussian process regression. *IEEE Transactions on Intelligent Transportation Systems*, *18*(12), pp.3363-3377.

Stone, M., 1974. Cross-validatory choice and assessment of statistical predictions. *Journal of The Royal Statistical Society: Series B (Methodological)*, *36*(2), pp.111-133.

Sun, S. and Xu, X., 2010. Variational inference for infinite mixtures of Gaussian processes with applications to traffic flow prediction. *IEEE Transactions on Intelligent Transportation Systems*, *12*(2), pp.466-475.

Veale, M. and Brass, I., 2019. Administration by algorithm? Public management meets public sector machine learning. In Yeung, K and Lodge M, (eds), *Public management meets public sector machine learning,* Oxford University Press, Oxford, UK, pp.121-149.

Wang, B., Gao, L. and Juan, Z., 2017. Travel mode detection using GPS data and socioeconomic attributes based on a random forest classifier. *IEEE Transactions on Intelligent Transportation Systems*, *19*(5), pp.1547-1558.

Wang, C.C.R. and Lien, J.J.J., 2008. Automatic vehicle detection using local features: A statistical approach. *IEEE Transactions on Intelligent Transportation Systems*, *9*(1), pp.83-96.

Wang, Z., Liang, M. and Delahaye, D., 2018. A hybrid machine learning model for short-term estimated time of arrival prediction in terminal manoeuvring area. *Transportation Research Part C: Emerging Technologies*, *95*, pp.280-294.

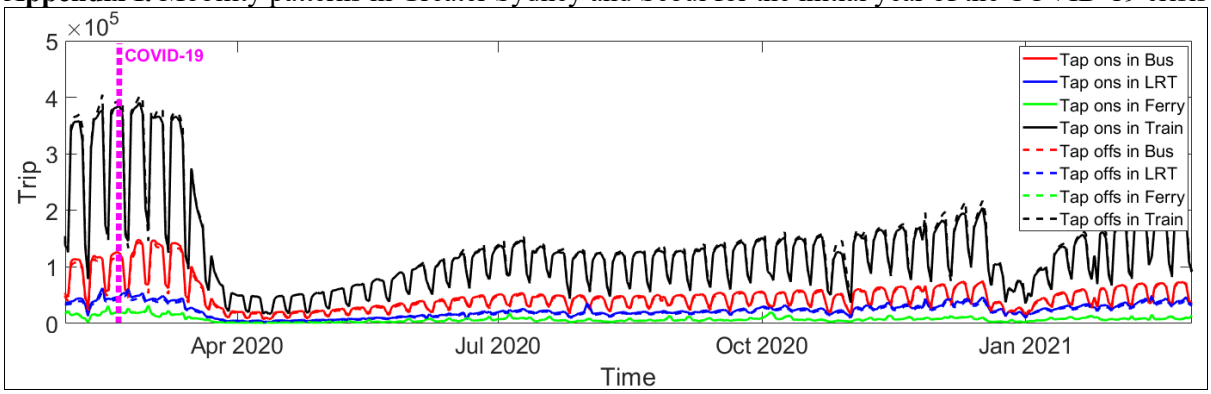
Wei, Y., Wang, J., Song, W., Xiu, C., Ma, L. and Pei, T., 2021. Spread of COVID-19 in China: Analysis from a city-based epidemic and mobility model. *Cities*, *110*, p.103010.

You, J. (2020). Lessons from South Korea's Covid-19 policy response. *The American Review of Public Administration*, 50(6-7), 801-808.

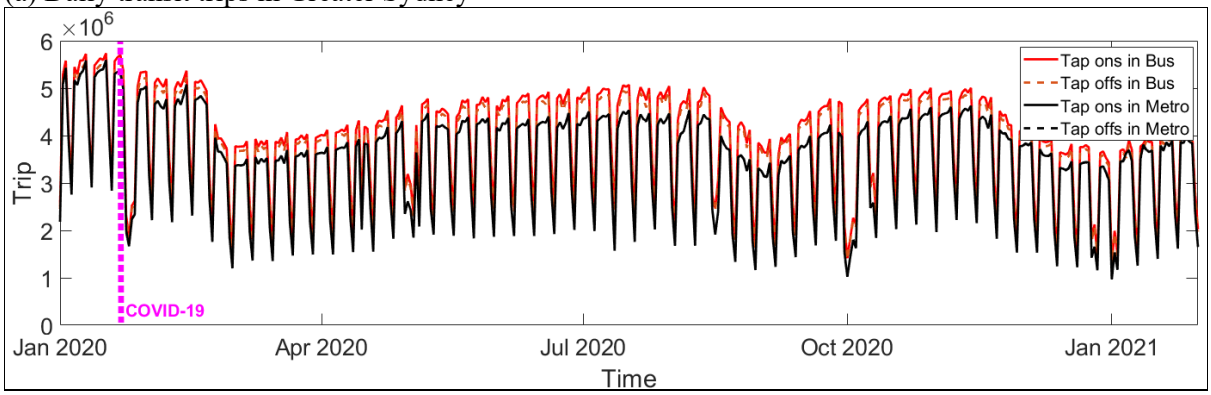
Zhang, G. and Wang, Y., 2014. A Gaussian kernel-based approach for modeling vehicle headway distributions. *Transportation Science*, *48*(2), pp.206-216.

Zhao, P. and Gao, Y., 2022. Public transit travel choice in the post COVID-19 pandemic era: An application of the extended Theory of Planned Behavior. *Travel Behaviour and Society*, 20, pp. 181-195.

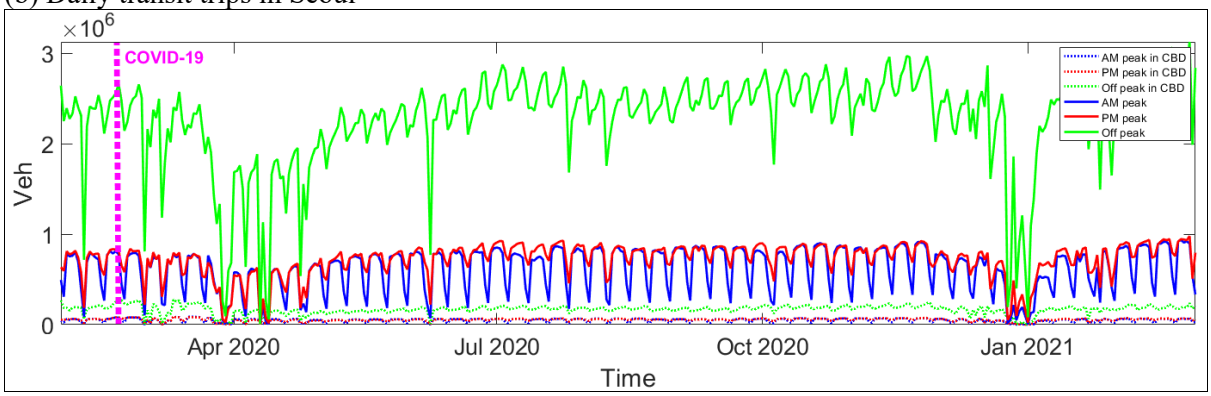
Appendix I. Mobility patterns in Greater Sydney and Seoul for the initial year of the COVID-19 crisis



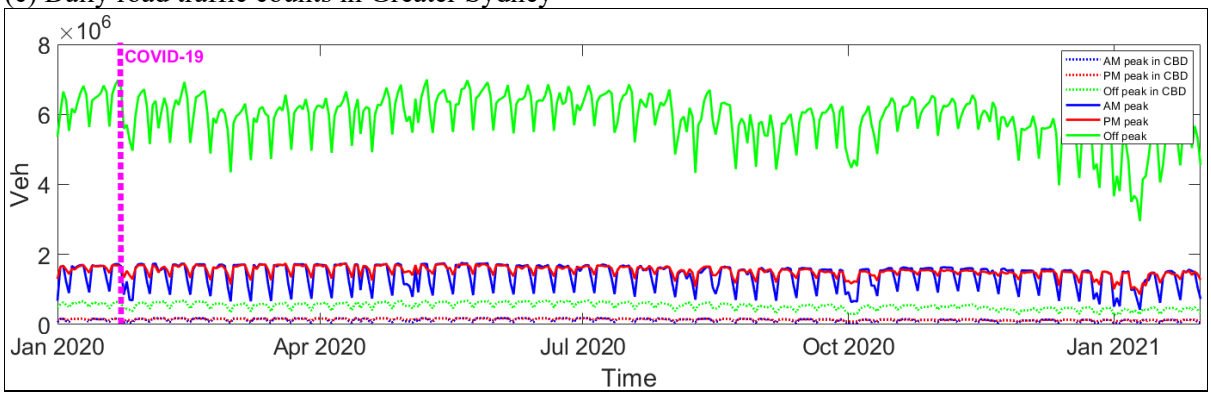
(a) Daily transit trips in Greater Sydney



(b) Daily transit trips in Seoul

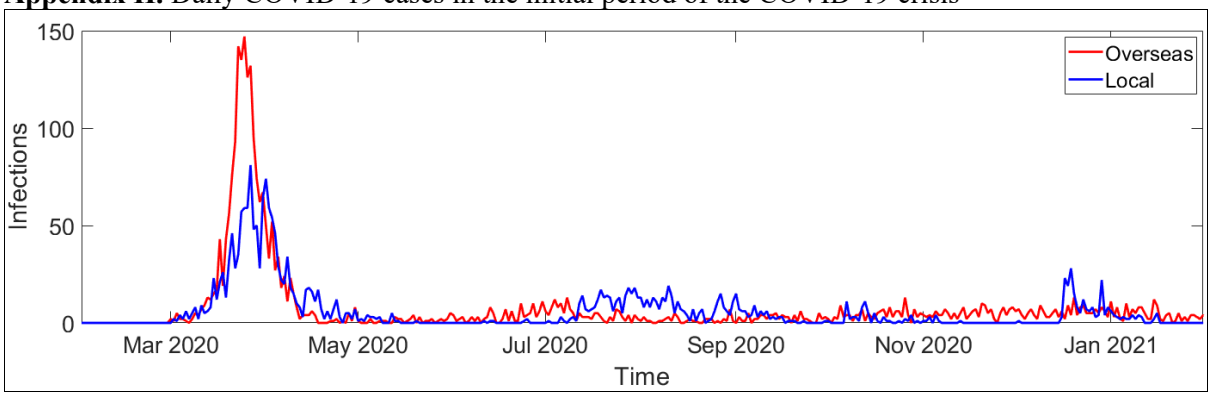


(c) Daily road traffic counts in Greater Sydney

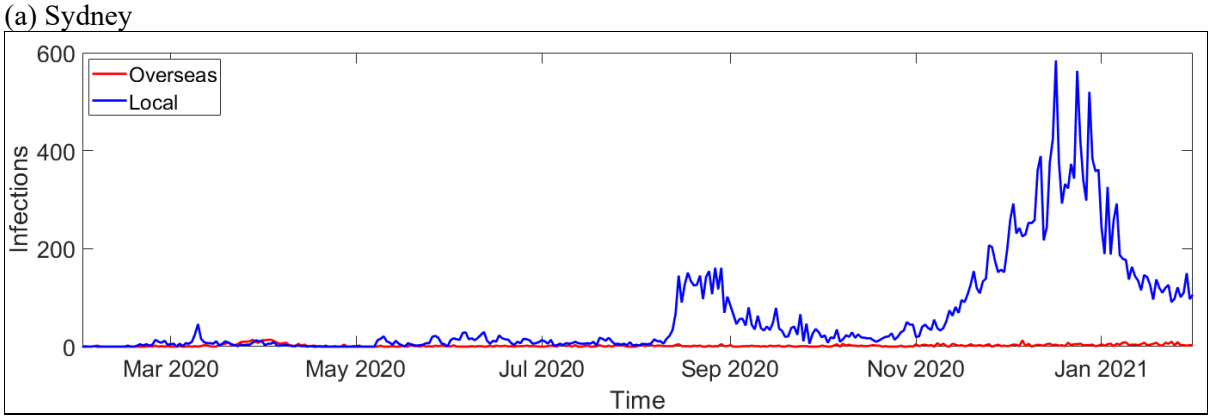


(d) Daily road traffic counts in Seoul

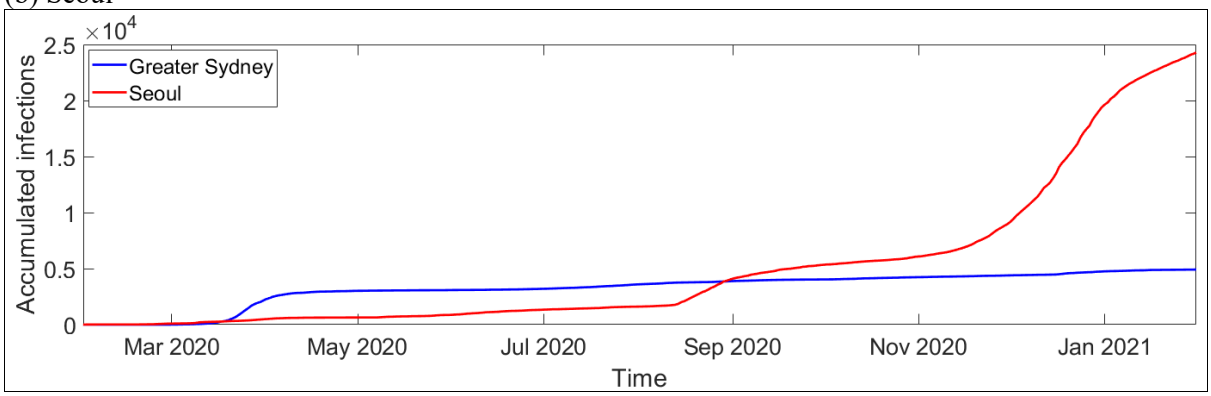
Appendix II. Daily COVID-19 cases in the initial period of the COVID-19 crisis



(a) Sydney



(b) Seoul



(c) Accumulated total infections

Appendix III. Constructed variables used for modeling the infection-mobility-policy cycle

Mode	Variable	Greater Sydney		Seoul	
		Metropolitan	CBD	Metropolitan	CBD
Number of daily international flights		O[1)]	X[2)]	O	X
7-day moving averages in international flights		O	X	O	X
Bus	Boarding AM peak hour	O	O	O	X
	Boarding PM peak hour	O	O	O	X
	Boarding off-peak hour	O	O	O	X
	Total boarding	O	O	O	O
	Alighting AM peak hour	O	O	O	X
	Alighting PM peak hour	O	O	O	X
	Alighting off-peak hour	O	O	O	X
	Total alighting	O	O	O	O
Train	Boarding AM peak hour	O	O	O	O
	Boarding PM peak hour	O	O	O	O
	Boarding off-peak hour	O	O	O	O
	Total boarding	O	O	O	O
	Alighting AM peak hour	O	O	O	O
	Alighting PM peak hour	O	O	O	O
	Alighting off-peak hour	O	O	O	O
	Total alighting	O	O	O	O
Ferry	Boarding AM peak hour	O	O	X	X
	Boarding PM peak hour	O	O	X	X
	Boarding off-peak hour	O	O	X	X
	Total boarding	O	O	X	X
	Alighting AM peak hour	O	O	X	X
	Alighting PM peak hour	O	O	X	X
	Alighting off-peak hour	O	O	X	X
	Total alighting	O	O	X	X
Light rail	Boarding AM peak hour	O	O	X	X
	Boarding PM peak hour	O	O	X	X
	Boarding off-peak hour	O	O	X	X
	Total boarding	O	O	X	X
	Alighting AM peak hour	O	O	X	X
	Alighting PM peak hour	O	O	X	X
	Alighting off-peak hour	O	O	X	X
	Total alighting	O	O	X	X
Vehicle counts for AM peak	Arterial	O	O	O	O
	Distributor	O	X	O	O
	Local	O	X	O	O
	Motorway	O	O	O	X
	Primary	O	X	O	X
	Aggregated	O	O	O	O
Vehicle counts for PM peak	Arterial	O	O	O	O
	Distributor	O	X	O	O
	Local	O	X	O	O
	Motorway	O	O	O	X
	Primary	O	X	O	X
	Aggregated	O	O	O	O
Vehicle counts for off-peak	Arterial	O	O	O	O
	Distributor	O	X	O	O
	Local	O	X	O	O
	Motorway	O	O	O	X
	Primary	O	X	O	X
	Aggregated	O	O	O	O

[1)]O indicates that the corresponding variable is included in the constructed mobility data structure.
[1)]X indicates that the corresponding variable is not included in the constructed mobility data structure.

Appendix IV. Predictors below the inflection point in a graph of the log of length scales in Greater Sydney

Log of length scale	Attributes of features						
	Time-lag	Policy/Mobility	Indicator	Tap on/off	Time	Region	Road hierarchy
-12.54	3	Policy	School	n/a	n/a	Metropolitan	n/a
0.26	7	Policy	International Travel	n/a	n/a	Metropolitan	n/a
0.26	3	Policy	International Travel	n/a	n/a	Metropolitan	n/a
0.51	11	Policy	International Travel	n/a	n/a	Metropolitan	n/a
0.68	13	Policy	International Travel	n/a	n/a	Metropolitan	n/a
0.73	9	Policy	Regional Travel	n/a	n/a	Metropolitan	n/a
0.75	5	Policy	Regional Travel	n/a	n/a	Metropolitan	n/a
0.96	14	Policy	Social Distancing	n/a	n/a	Metropolitan	n/a
0.97	4	Policy	Social Distancing	n/a	n/a	Metropolitan	n/a
1.15	5	Policy	Social Distancing	n/a	n/a	Metropolitan	n/a
1.16	1	Policy	Indoor	n/a	n/a	Metropolitan	n/a
1.21	1	Policy	Pub	n/a	n/a	Metropolitan	n/a
1.24	2	Policy	International Travel	n/a	n/a	Metropolitan	n/a
1.27	5	Policy	Interstate Travel	n/a	n/a	Metropolitan	n/a
1.27	10	Policy	Social Distancing	n/a	n/a	Metropolitan	n/a
1.33	4	Policy	Regional Travel	n/a	n/a	Metropolitan	n/a
1.33	12	Policy	Transit	n/a	n/a	Metropolitan	n/a
1.33	12	Policy	Public Gathering	n/a	n/a	Metropolitan	n/a
1.63	12	Policy	International Travel	n/a	n/a	Metropolitan	n/a
1.81	11	Policy	Regional Travel	n/a	n/a	Metropolitan	n/a
1.82	1	Policy	Interstate Travel	n/a	n/a	Metropolitan	n/a
1.94	12	Mobility	LRT	Tap on	PM peak	CBD	n/a
1.97	6	Policy	International Travel	n/a	n/a	Metropolitan	n/a
1.98	13	Mobility	Bus	Tap on	AM peak	CBD	n/a
2.09	3	Mobility	Ferry	Tap off	PM peak	Metropolitan	n/a
2.10	13	Mobility	Road traffic counts	n/a	PM peak	CBD	Total
2.14	5	Mobility	Road traffic counts	n/a	Off peak	Metropolitan	Primary
2.15	12	Mobility	Bus	Tap off	PM peak	CBD	n/a
2.18	2	Mobility	Ferry	Tap off	PM peak	Metropolitan	n/a
2.19	4	Mobility	Ferry	Tap on	PM peak	Metropolitan	n/a
2.20	11	Mobility	Ferry	Tap on	Off peak	CBD	n/a
2.20	6	Mobility	Road traffic counts	n/a	Off peak	Metropolitan	Primary
2.24	10	Mobility	LRT	Tap off	AM peak	CBD	n/a
2.29	3	Mobility	Ferry	Tap on	AM peak	Metropolitan	n/a
2.30	2	Policy	School	n/a	n/a	Metropolitan	n/a
2.31	14	Policy	Interstate Travel	n/a	n/a	Metropolitan	n/a
2.32	12	Mobility	Road traffic counts	n/a	Off peak	Metropolitan	Primary
2.34	7	Mobility	Road traffic counts	n/a	PM peak	Metropolitan	Primary
2.38	3	Mobility	Road traffic counts	n/a	Off peak	Metropolitan	Distributor
2.39	9	Mobility	Road traffic counts	n/a	Off peak	Metropolitan	Total
2.40	4	Mobility	Road traffic counts	n/a	PM peak	Metropolitan	Arterial
2.54	6	Mobility	Road traffic counts	n/a	AM peak	Metropolitan	Primary
2.79	11	Mobility	Road traffic counts	n/a	PM peak	Metropolitan	Total
2.83	2	Mobility	Road traffic counts	n/a	PM peak	Metropolitan	Primary
3.35	3	Mobility	Road traffic counts	n/a	PM peak	Metropolitan	Motorway

Appendix V. Predictors below the inflection point in a graph of the log of length scales in Seoul

Log of length scale	Attributes of features						
	Time lag	Policy/Mobility	Indicator	Tap on/off	Time	Region	Road hierarchy
-4.14	14	Policy	Public Gathering	n/a	n/a	Metropolitan	n/a
-3.61	13	Policy	Public Gathering	n/a	n/a	Metropolitan	n/a
-2.60	12	Policy	Public Gathering	n/a	n/a	Metropolitan	n/a
0.04	12	Policy	Childcare	n/a	n/a	Metropolitan	n/a
0.06	13	Policy	Childcare	n/a	n/a	Metropolitan	n/a
0.22	4	Policy	Social Distancing	n/a	n/a	Metropolitan	n/a
0.57	7	Policy	Mask	n/a	n/a	Metropolitan	n/a
0.65	13	Policy	Social Distancing	n/a	n/a	Metropolitan	n/a
0.75	9	Policy	Social Distancing	n/a	n/a	Metropolitan	n/a
0.93	6	Policy	Social Distancing	n/a	n/a	Metropolitan	n/a
1.00	11	Policy	Mask	n/a	n/a	Metropolitan	n/a
1.13	13	Mobility	Bus	Tap off	AM peak	Metropolitan	n/a
1.16	1	Policy	Social Distancing	n/a	n/a	Metropolitan	n/a
1.19	7	Mobility	Road traffic counts	n/a	Off peak	Metropolitan	Motorway
1.24	1	Mobility	Road traffic counts	n/a	AM peak	CBD	Local
1.25	11	Mobility	Road traffic counts	n/a	PM peak	Metropolitan	Distributor
1.31	10	Mobility	Road traffic counts	n/a	PM peak	Metropolitan	Arterial
1.36	11	Mobility	Road traffic counts	n/a	Off peak	Metropolitan	Arterial
1.37	14	Mobility	Bus	Tap off	AM peak	Metropolitan	n/a
1.37	2	Policy	Social Distancing	n/a	n/a	Metropolitan	n/a
1.38	6	Policy	Mask	n/a	n/a	Metropolitan	n/a
1.65	6	Mobility	Road traffic counts	n/a	AM peak	Metropolitan	Motorway
1.88	11	Mobility	Road traffic counts	n/a	AM peak	Metropolitan	Arterial
2.12	12	Policy	Mask	n/a	n/a	Metropolitan	n/a
2.14	5	Mobility	Bus	Tap off	AM peak	CBD	n/a
2.17	4	Policy	Restaurant	n/a	n/a	Metropolitan	n/a
2.17	4	Policy	School	n/a	n/a	Metropolitan	n/a
2.26	5	Mobility	Road traffic counts	n/a	AM peak	CBD	Local
2.60	12	Policy	School	n/a	n/a	Metropolitan	n/a
2.60	12	Policy	Restaurant	n/a	n/a	Metropolitan	n/a
3.12	11	Policy	Restaurant	n/a	n/a	Metropolitan	n/a
3.12	11	Policy	School	n/a	n/a	Metropolitan	n/a