Transportation Research Part C 165 (2024) 104730

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

How did international air transport networks influence the spread of COVID-19? A spatial and temporal modeling perspective

Chi Li [a,f], Linhao Yu [b], Jianfeng Mao [b,c,*], Wei Cong [d], Zibin Pan [a], Yuhao Du [e], Lianmin Zhang [f]

[a] School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, Guangdong, 518172, PR China
[b] School of Data Science, The Chinese University of Hong Kong, Shenzhen, Guangdong, 518172, PR China
[c] Guangdong Provincial Key Laboratory of Big Data Computing, School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, Guangdong, 518172, PR China
[d] Feeyo Technology Co., Ltd., No. 320 Qianshan Rd., Hefei, Anhui, PR China
[e] School of Management and Economics, The Chinese University of Hong Kong, Shenzhen, Guangdong, 518172, PR China
[f] Shenzhen Research Institute of Big Data, The Chinese University of Hong Kong, Shenzhen, Guangdong, 518172, PR China

ARTICLE INFO

Keywords:
COVID-19
Air transport
Spatial heterogeneity
Temporal dynamics
Omicron

ABSTRACT

The international air transport network is pivotal in the global propagation of emerging infectious diseases. Gaining insights into the nuances of this transmission mechanism can pave the way for more strategic and effective interventions. While previous studies have delved into the subject, an integrated spatial and temporal modeling framework, specifically tailored to distinct phases of COVID-19 and its variants, has yet to be fully explored. This research aims to address this gap by exploring the spatial and temporal impacts of the air transport network on the spread of COVID-19 and its Omicron variant. We introduce an improved effective distance metric to assess the spatial correlation between various distance metrics and the onset of infectious diseases in selected nations. Subsequently, we employ a network-based heterogeneous susceptible-unreported infectious-confirmed-recovered-death (SUCRD) mathematical model to delineate the temporal evolution of infections by country. Our findings underscore the air transport network's instrumental role in the pandemic's spatial dynamics. Moreover, our model has been validated, demonstrating robustness and reliability. Through rigorous validation and simulation experiments, we discern the significance of the timing and intensity of interventions in shaping the pandemic's trajectory. Notably, while the air transport network exerts a profound influence during the phases of both COVID-19 and Omicron, international travel restrictions exhibit diminishing returns once the disease achieves widespread prevalence. Through comparative analysis and discussion, we highlight the advantages of our experimental outcomes and methodological approach compared to previous studies. Based on our findings, we identify six key policy implications that offer critical perspectives for aviation stakeholders. This study illuminates the role of the air transport network in affecting the spatial accessibility and temporal dynamics of pandemic transmission, thereby providing valuable insights for informed policy-making in the aviation sector.

* Corresponding author at: School of Data Science, The Chinese University of Hong Kong, Shenzhen, Guangdong, 518172, PR China.
E-mail addresses: chili@link.cuhk.edu.cn (C. Li), linhaoyu@link.cuhk.edu.cn (L. Yu), jfmao@cuhk.edu.cn (J. Mao), congwei@variflight.com (W. Cong), zibinpan@link.cuhk.edu.cn (Z. Pan), yuhaodu1@link.cuhk.edu.cn (Y. Du), lmzhang@sribd.cn (L. Zhang).

https://doi.org/10.1016/j.trc.2024.104730
Received 17 October 2023; Received in revised form 27 March 2024; Accepted 17 June 2024
Available online 3 July 2024

1. Introduction

In December 2019, the emergence of a novel coronavirus, termed COVID-19 and caused by the SARS-CoV-2 coronavirus, first appeared in Wuhan, China. The disease expanded to 24 other countries worldwide within a few weeks (Wells et al., 2020). By March 11, 2020, the World Health Organization (WHO) declared COVID-19 a pandemic (Organization et al., 2020). By June 2023, there have been 690 million confirmed cases and over 6.88 million fatalities related to COVID-19. COVID-19 has developed into a prevalent pandemic in the past three years, posing a threat to global health and economies. What is worse, the new severe Omicron variant has spread to more than 100 countries and territories around the world since it was first detected in South Africa in November 2021 (Tong et al., 2022). As of now, the global aviation industry is in a recovery phase, with travel demand and various aviation markets steadily rebounding. Although there remains a gap compared to pre-pandemic levels, this presents a positive sign of progress in the aviation sector (Sun et al., 2023a). However, in the recovery process of the aviation industry, new variants of the virus continue to emerge. The aviation system should remain vigilant to the evolving characteristics of these new viruses (Sun et al., 2023b; Xu et al., 2023). By examining the transmission dynamics of COVID-19, especially the characteristics of the newly emerged variants, and understanding the influence of the aviation network in this spread, the aviation sector can be better equipped to mitigate substantial setbacks akin to those encountered during the pandemic's initial phase, even in the face of a continually evolving epidemic landscape.

Expanding air transport networks and fast-growing air traffic demand have intensively increased international trade and tourism across the globe. With the evolutionary development of COVID-19, the aviation industry has become one of the hardest-hit sectors due to the unprecedented reduction in air passenger demand caused by travel restrictions (Sun et al., 2020). Bao et al. (2021) examine these effects from the perspective of network structure and explain how the worldwide air transportation network yields an immense impact on the national and international economy and politics. From the statistics, Li (2021) reveals that the world scheduled capacity and global air international and domestic passengers declined 50% and 60% in 2020. Ivanov (2020) employs a simulation-based approach to comprehensively analyze and predict the impact of COVID-19 on the global aviation supply chain, offering managerial insights. Moreover, Sun et al. (2021d) conduct a systematic survey to analyze the impact of COVID-19 on the passenger-centric flight experience and its long-term influence on aviation.

Nevertheless, it is often ignored that the aviation sector is not only a victim of COVID-19 but also a crucial channel in the transmission of diseases, allowing a local pandemic to become a worldwide one (Sun et al., 2021d). Air travel is believed to play a pivotal role in hastening the spread of pandemics. The intrinsic connectivity of air transportation provides a conducive setting for minor epidemic outbreaks to rapidly escalate into widespread pandemics (Colizza et al., 2006; Sun et al., 2021a). In general, two analytical methods are employed to scrutinize the influence of air transport networks on pandemic dissemination, focusing on spatial accessibility and temporal evolution perspectives, respectively. On the one hand, to evaluate the role of transport network accessibility in the spread of COVID-19, Mahmud et al. (2021) verify a strong association between the presence of COVID-19 and the mechanism of transport networks with the application of Pearson correlation analysis. Besides, Brockmann and Helbing (2013) introduce a concept known as effective distance and they find the number of passengers is a reasonably accurate indicator of predicting the disease's arrival time. Adiga et al. (2020) and Kuo and Chiu (2021) apply this method to the case of COVID-19. However, they have not utilized the recent data containing flight suspensions or considered the impact of indirect flights. On the other hand, spatial mobility and control measures contribute substantially to the temporal dynamics of COVID-19 spreading (Kraemer et al., 2020; Lu et al., 2021). Chen et al. (2020) employ a network-based compartmental mathematical model to evaluate the temporal influence while Hu et al. (2021) investigate how human mobility varies in response to different policies. Meanwhile, many studies have also examined the impact of different policy interventions on the spread of COVID-19 (Li et al., 2021c; Wandelt et al., 2023; Meng et al., 2023). In addition to research on the early stages of the pandemic's spread, the impact during different epidemic phases (Yu and Chen, 2021) and the effects of different variants have also been taken into account (Sun et al., 2021b; Choi et al., 2022).

However, the majority of conclusions drawn from prior research tend to be narrowly focused and seldom integrate the two representative models to analyze both the spatial accessibility and temporal evolution effects. Little is also known about how the timing, intensity, and duration of travel restrictions and control measures affect the spread of the pandemic. Moreover, most of the research is independent of the specifics of COVID-19, which is emphasized in Sun et al. (2021d), and hardly involves the impact of the typical variant Omicron. Overall, by integrating the aforementioned two types of analytical models and addressing the stated limitations, we are dedicated to offering a spatial and temporal modeling framework to assess the role of aviation networks in the spread of the pandemic.

Moreover, it is pertinent to note that our chosen subject within the air transport network is constructed on a country-wise basis. Unlike a seminal work (Guimera et al., 2005) that employs network science techniques to analyze the structural characteristics and community features of global aviation networks and sub-networks, we aim to assess the role of air transport networks in international transmission from a macroscopic perspective. As stated in Brockmann and Helbing (2013) and Sun et al. (2021a), air transportation serves as the primary medium for virus diffusion between countries. Similarly, transportation-focused studies in Nikolaou and Dimitriou (2020), Sun et al. (2021a) and Sun et al. (2022a), which analyze structural characteristics and transmission roles from an air transport country network perspective, also underscore that such analytical approaches can mitigate the influence of domestic factors, such as other modes of transportation on transmission. Therefore, constructing networks from a country-wise angle is both rational and substantiated, whether viewed from an epidemiological perspective or that of transportation networks.

From a methodological standpoint, our study provides a combined spatio-temporal modeling framework. In assessing spatial implications, we compare the disease's initial arrival times with various effective distance metrics. By delving into the unique

characteristics of each metric, we elucidate the sources of discrepancies and evaluate the viability of using these metrics to forecast initial disease arrival times. For temporal evolution, our research expands the foundational network-based compartmental model by incorporating an often-overlooked segment: the unreported infected individuals. Furthermore, our model integrates a policy intervention function, accounting for the onset, intensity, and duration of policy measures. The model also captures the heterogeneity and temporal nuances of policy implementations across different countries/regions. Validation experiments and sensitivity analyses underscore the robustness and reliability of our model.

From the perspective of research questions, our analysis is multi-faceted, considering the distinct transmission characteristics of COVID-19 across various stages. Recognizing the unique propagation traits of the disease within air transport networks and data utilization, we employ stage-specific medical parameters reflecting the disease's transmission capability. Instead of the widely used yet often inaccurate traffic volume, we estimate passenger volume using load factor data, factoring in the potential influence of transfer flights on the transmission process. While we are currently in the recovery phase of COVID-19, the emergence of new variants remains a concern. Given our model's adaptability to different virus parameter characteristics, it can be extended to evaluate the impact of these new variants on the aviation industry, thereby aiding in the formulation of more informed policies.

To the best of our knowledge, this is among the first analytical frameworks that integrate both spatial and temporal models to evaluate the role of air traffic networks in the spread of COVID-19 and its Omicron variant. Our contribution can be illustrated as follows:

- We introduce an integrated spatio-temporal modeling framework, tailored to the unique characteristics of the pandemic, to evaluate the impact of air transport networks on COVID-19 transmission across multiple stages.
- We incorporate the influence of transfer flights on disease propagation, building upon traditional effective distance metrics. By comparing and meticulously analyzing various effective distance indicators, we discern their relationship with the initial arrival times of the disease, thereby shedding light on spatial implications.
- We enhance the conventional network-based compartmental model by integrating additional compartments. Synergistically combining a policy intervention function, heterogeneous disease transmission parameters across nations, and time-varying traffic flows, our model offers insights into temporal dynamics.
- We conduct an extensive array of validation and simulation experiments, underscoring the reliability and robustness of our model. These experiments also elucidate how the implementation of policies at varying intensities can modulate disease transmission.

The remainder of this study is organized as follows: Section 2 proposes a review of the extant literature for analyzing the role of air transport networks on the development of COVID-19 from two aspects: modeling methods and influence analysis. Section 3 describes the data and the related preprocessing. Section 4 provides the integrated models that are applied to our research. We then conduct the experiments of measuring the spatial and temporal influence in Section 5. In Section 6, we summarize the policy implications derived from our study from six aspects. Section 7 concludes our results, limitations, and provides some future research directions.

2. Literature review

We will summarize the literature relevant to our research from two perspectives. One is to provide an overview of the modeling methods used to depict the spread of COVID-19 through air traffic networks. The other is to outline the typical literature that examines the relationship and influence between COVID-19 and air transportation.

2.1. Modeling methods

2.1.1. Effective distance

The actual geographical distance commonly denotes the distance between two nodes in a specific network. However, employing the geographic distance to determine its correlation with the epidemic transmission via air transport networks is inconsistent with the natural evolution process of infectious diseases. For instance, an epidemic can be transmitted by busy flight schedules although two airports are far away from each other. As proposed by Brockmann and Helbing (2013), the innovative concept of effective distance matched a higher relationship with the arrival dates of various diseases by converting the air passenger data to a metric. Further, derived from the random walks theory, Iannelli et al. (2017) proposed general logarithmic metrics to predict the arrival time of disease on complex networks, which presented that diseases were more likely to spread via multiple paths.

In recent COVID-19 validation experiments, Adiga et al. (2020) apply the effective distance metric to measure the arrival time of COVID-19 in the early stages. Kuo and Chiu (2021) investigate and compare the relationship between three different distance metrics involving effective distance, shortest path distance, and random walk distance with the actual arrival time. Nevertheless, these results hardly consider the indirect links of air transport networks or involve current air passenger data during the pandemic. In a recent study, Choi et al. (2022) investigate several variants, measuring the relationship between their arrival times and the basic effective distance. An empirical model is also established to analyze the relationship between the disease's arrival time and policies.

2.1.2. Epidemiological models

The epidemiological models that predict the dynamic transmission of COVID-19 can be classified into four categories: statistical-based models, network-based models, agent-based models, and SI-based models.

Statistical-based models

Typically, a statistical model is stated as a mathematical representation that reflects a set of statistical assumptions regarding the relationship between one or more random variables and non-random variables. One of its broad applications is assessing the pandemic's importation risk. Wells et al. (2020) evaluated the country-level importation risk probability and investigated how border control measures affected the spread of the outbreak using Monte Carlo simulations. Some other researchers apply statistical models to estimate the role of air transport and high-speed rail in the propagation of COVID-19. Dai et al. (2021) utilized an Ordinary Least Squares (OLS) regression-based model to validate that air travel had a significant and favorable effect on the early stages of pandemic growth. Similarly, Mahmud et al. (2021) indicated a strong association between the structure of transport networks and the presence of COVID-19 with a high Pearson correlation coefficient. The random-effects panel data model (Su et al., 2022), economic model (Zhu and Guo, 2021), and regression model (Sokadjo and Atchadé, 2020) have also been proposed to measure the correlation of new daily confirmed cases and some related factors including flight suspension.

The statistical models are very adaptable in terms of the structures and characteristics of the input data, and they can facilitate the interpretation of the relationship between variables. However, they may not be suitable for analyzing the spatial and temporal effects of the pandemic based on the time-varying transport networks.

Network-based models

Metapopulation models are a concept that gives rise to network-based epidemiological models where linkages between regions indicate specific mobility flows (Ding et al., 2021). Network-based models are often utilized to generalize intricate interaction networks with complex network metrics so as to assess the probability and course of epidemic transmission. Ruan et al. (2021) employed Zipf's law, a specific discrete power-law distribution connected with social networks, to predict the transmission of the virus. In Coelho et al. (2020a), the authors applied a centrality metric named Eigenvector Centrality to determine the importance of a country in the connected network. Network-based models can also be utilized to evaluate the importation risk. Li et al. (2021d) established a high-speed rail (HSR) network and integrated the network metrics with probabilistic risk models to rank risks in different regions of China. Bao et al. (2021) uncovered that the impacts of pandemics on the worldwide air transportation network were aligned with global attacks via network robustness analysis.

Network-based models can more accurately replicate epidemic transmission and evaluate the risk if it is known how different populations/regions behave and interact. The analysis of infectious illnesses with individual detached structures is hampered by the fact that network-based models cannot describe micro-individuals (Li et al., 2021a).

Agent-based models

Agent-based models are based on individual modeling under the assumption that discrete individuals' contact activities and personal social networks contribute to disease spreading. We discover a small number of papers employing agent-based models to examine the diffusion of COVID-19. Chinazzi et al. (2020) proposed a Global Epidemic and Mobility Model (GLEAM) to analyze the effect of travel and quarantine on the dynamic evolution of the outbreak. Previous studies referred to Duan et al. (2013), Ciofi degli Atti et al. (2008) and Stefanoff et al. (2010).

Agent-based models have more adaptability and high predictability in simulation experiments. However, they are complicated to evaluate and predict due to the complex operating mechanism and individual heterogeneity.

SI-based models

The Susceptible-Infective (SI) based model, often referred to as compartmental models, presents an intuitive comprehension of how infectious diseases transmit among the crowd. The basic model contains two categories of people: those who are susceptible to infection (S) and those who are infected (I). There are numerous variations of SI-based models where the Susceptible-Infectious-Recovered (SIR) model and the Susceptible-Exposed-Infectious-Recovery (SEIR) model are the two most used approaches. Yang et al. (2020) developed an SEIR method to simulate the dynamic evolution of the disease. This framework can also incorporate human mobility and spatial factors, including traffic data and mobile phone signaling data. This is commonly referred to as the network-based compartmental model, which simulates and analyzes how populations in different states spread and transmit within a network or geographical space (Chen et al., 2020; Zhang et al., 2020; Sun et al., 2022d; Luo et al., 2022).

However, there is no unanimous agreement on the optimal approach to model the dynamic progression of COVID-19, since each methodology possesses its unique characteristics and application domains. In our research, we are committed to analyzing the influence of air traffic networks on the spread of COVID-19 from both spatial and temporal modeling perspectives. As such, the network-based compartmental model combined with effective distance emerges as a particularly apt modeling approach.

2.2. Relationship and influence analysis

There exists an extensive body of literature exploring the relationship and influence analysis between the COVID-19 pandemic and air transportation. Sun et al. (2021d) provide a comprehensive review, emphasizing the interplay between the COVID-19 disease and the aviation sector. They also highlight the role of aviation as a potential conduit for virus transmission. In their recent work, Sun et al. (2022b) delve into the direct and indirect interplay between COVID-19 and air transportation. Additionally, they distill key conclusions drawn from these studies and highlight policy implications geared toward building a pandemic-resilient aviation sector.

Drawing inspiration from Sun et al. (2021d) and Sun et al. (2022b), two primary themes emerge in the literature concerning the interrelation between the COVID-19 pandemic and air transportation. The first theme delves into the repercussions of the COVID-19 pandemic on the aviation industry. In contrast, the second theme centers on the ramifications of air transportation on the spread and dynamics of the COVID-19 pandemic.

2.2.1. Influence of the COVID-19 pandemic on air transportation

To date, there have been numerous articles analyzing the impact of COVID-19 on air transportation. Given that our paper focuses on the spatio-temporal effects of air traffic networks on the spread of COVID-19, we will list some representative literature that examines the various ways COVID-19 affects air transport at the operational level. For a more detailed discussion, readers can refer to these two reviews (Sun et al., 2021d, 2022b).

Table A.6 illustrates the impact of COVID-19 on air transportation at the operational level, primarily summarizing key findings from the literature in terms of network structural performance and the perspectives of traffic or passenger demand. Many articles employ time series analysis to observe the relationship between the growth sequence of COVID-19 cases and the reduction sequence of flights. They identify the synchronicity between the two sequences, further uncovering the underlying patterns of influence (Sun et al., 2021c; Wang et al., 2023; Sun et al., 2023b). Additionally, some studies have explored the relationship between the reduction in aviation demand and the number of confirmed COVID-19 cases using time series analysis, offering relevant policy guidance (Truong, 2021; Li et al., 2021b; Kim and Sohn, 2022). Moreover, the integration of complex network theory with time series analysis has emerged as a prevalent methodology. By examining the changes in the connectivity structure and network performance during the process of flight reductions, researchers analyze whether the aviation network possesses robustness (Sun et al., 2020; Zhou et al., 2021; Li et al., 2022). They also investigate if the centrality indicators of the aviation network changed before and after the outbreak of the pandemic (Sun et al., 2021a; Kuo et al., 2022).

2.2.2. Impact of air transportation on the COVID-19 pandemic

The escalating international connectivity and the momentum of globalization have accelerated the spread of infectious diseases far beyond what was previously anticipated. As highlighted in Sun et al. (2021d), while the enhancement of connectivity and the reduction in travel times are advantageous for passengers, they spell trouble for disease transmission. Historically, research has indicated that air travel acts as a significant catalyst for the global spread of infectious diseases (Colizza et al., 2006). However, when it comes to the unique characteristics of the COVID-19 virus, a thorough and scientific investigation is still required to gauge the spatio-temporal evolution of its spread influenced by air transportation. In the following sections, I will delve into how prior research has employed mathematical models to quantitatively measure this impact.

Li et al. (2020) employ multi-source data to analyze the efficacy of air transportation regulations in curbing the progression of the pandemic. The study indicates that restrictions on air traffic can help reduce the number of confirmed cases in major cities. However, as the intensity of these restrictions increases, their impact diminishes. Drawing from existing research and global spatial and cartographic information, Nakamura and Managi (2020) calculate the overall relative risk of COVID-19 import and export for each airport in municipalities worldwide. Sun et al. (2022a) utilize complex network and time series analysis to explain why African countries did not witness a significant surge in infection cases and delve into the role of air transportation in this transmission process. The study proposed by Mutascu and Sokic (2023), on the other hand, employs wavelet analysis to study the bilateral relationship between the pandemic situation in European countries and air traffic.

During different stages of the pandemic's development, air transportation tends to have varying impacts on COVID-19, and conversely, COVID-19 affects the aviation industry to different extents (Sun et al., 2023b). Yu and Chen (2021) employ a latent variable-based structural equation model (SEM) to analyze the influence of policies at different stages on the number of COVID-19 infections, revealing that implementing circuit breaker policies alone might not be very effective. On the other hand, the presence of different variants during various stages of the pandemic exhibits unique transmission characteristics, necessitating a meticulous scientific analysis of these effects. Sun et al. (2021b) use network-based compartmental models and time series analysis to study the relationship between the propagation time of COVID-19 variants and flight bans, concluding that national policies addressing variant spread were indeed delayed. Choi et al. (2022) specifically examine several particular variants, measuring their first arrival time in relation to the basic effective distance. Additionally, numerous articles approach the topic from a policy intervention perspective, researching how to formulate reasonable travel restriction measures to effectively curb the spread of COVID-19 (Truong, 2021; Zhang and Hayashi, 2022; Meng et al., 2023). It is essential to emphasize that many models' initial assumptions are based on passenger numbers. However, actual simulations and calculations often use flight frequencies or a uniform load factor as approximations, which can reduce the feasibility of the results. Sun et al. (2022c) and Choi et al. (2022) highlight potential sources of load factor data, such as Sabre Market Intelligence. They utilize the load factor data to identify abnormal flights executed by airlines during the pandemic and the number of passengers during this period.

As outlined previously, we will employ a method that combines effective distance metrics with the network-based compartmental model to measure the role of air transport networks on pandemic transmission. Table 1 offers a concise comparison of relevant literature, underscoring the various problems addressed by these studies.

Our research provides a detailed examination of this domain. Notably, no current studies simultaneously analyze the impact of air transportation networks on COVID-19 transmission from both spatial and temporal modeling perspectives. Among these models, very few pieces of literature incorporate load factor data, and there is a limited number of studies addressing the effects of different variants. In specific spatio-temporal models, our methodology is more reflective of practical modeling situations. For instance, in spatial models, we consider transfer flights and utilize multiple effective distance metrics. In our temporal models, we consider six different facets that are more attuned to real-world problem modeling, something other papers have not completely achieved, especially when considering the time-varying traffic flow and distinct medical transmission characteristics displayed by different variants during specific periods. It is important to acknowledge the challenges in accurately assessing the influence of air traffic networks on infection case counts in specific countries due to the complexity of the factors involved. As such, the methodology we present offers significant value, providing a reference point for the development of related policies and furnishing a solid framework for the modeling and analysis of new variant outbreaks.

Table 1
Summary of previous research on evaluation of the spatial and temporal effects of air transport networks on the spread of COVID-19.

Reference	Study area	COVID-19	Multiple	Data		Effective distance		Network-based compartmental models					
		variants	stages	Recent	Load factor	Transfer flights	Multiple metrics	Time varying	Heterogeneity	Policy intervention	Medical parameters	Evolutionary dynamics	Validation
Lau et al. (2020b)	4 Regions	–	–	–	–	✓	–	NA	NA	NA	NA	NA	NA
Daon et al. (2020)	6 Countries/Regions	–	–	✓	–	–	–	NA	NA	NA	NA	NA	NA
Coelho et al. (2020b)	5 Regions	–	–	✓	–	–	–	NA	NA	NA	NA	NA	NA
Adiga et al. (2020)	24 countries	–	–	–	–	✓	–	NA	NA	NA	NA	NA	NA
Chen et al. (2020)	The US	–	–	✓	–	NA	NA	–	✓	–	✓	✓	✓
Zhang et al. (2020)	China	–	–	✓	–	NA	NA	–	–	✓	✓	✓	✓
Nikolaou and Dimitriou (2020)	European countries	–	–	✓	–	NA	NA	–	–	✓	–	–	-
Sun et al. (2021b)	25 countries	✓	✓	✓	–	NA	NA	–	✓	✓	–	–	✓
Sheng et al. (2021)	10 countries	–	–	✓	–	NA	NA	–	✓	✓	–	–	-
Kuo and Chiu (2021)	29 countries	–	–	–	–	–	✓	NA	NA	NA	NA	NA	NA
Yu and Chen (2021)	6 Regions	–	✓	✓	–	NA	NA	–	–	✓	–	–	✓
Liu et al. (2022)	Singapore	–	✓	✓	–	NA	NA	✓	–	✓	–	✓	✓
Sun et al. (2022d)	9 Countries	✓	–	✓	–	NA	NA	–	–	✓	–	–	✓
Choi et al. (2022)	224 Countries	✓	✓	✓	✓	✓	–	NA	NA	NA	NA	NA	NA
Our work	**24 Countries/Regions**	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

[1] "–" indicates that the study did not consider this factor, "NA" signifies that the study did not take into account this type of model, while "✓" denotes that the study did consider this factor.
[2] "Study Area" refers to the geographical or thematic scope within the research's experimental validation. "COVID-19 Variants" indicates whether the research has considered the various mutations of the virus. "Multiple Stages" denotes whether the study considers the impacts beyond the initial phase of the COVID-19 outbreak. "Data" encompasses whether the research utilizes the most recent data available and if it incorporates load factor data or air passenger data for its analysis. "Effective Distance" entails whether the study takes into account the impact of transfer flights and if it considers multiple effective distance metrics in its experiments. "Network-based Compartmental models" encompasses whether the study considers time-varying air passenger data, heterogeneous parameters, medical parameters, and the impact of policy intervention. Additionally, it evaluates if the research compares simulation results with actual data and showcases the fit between the simulated curves and the actual infection count over time.

2.3. Research gap

Effective distance combined with network-based compartmental models offer useful insights into the pandemic's impact, considering various virus strains' transmission characteristics. However, even with the existing literature that aligns with our research question and methodology, there remain the following research gaps.

- To our knowledge, no existing studies present a modeling framework that concurrently addresses the spatial and temporal impacts of COVID-19. Spatially, the framework correlates the virus's initial arrival times with international passenger volumes. Temporally, it examines how alterations in the aviation network affect COVID-19 case fluctuations. This integrated approach offers a detailed perspective, providing valuable insights for analyzing the onset and progression of new virus outbreaks.
- The impact of air traffic networks on COVID-19 spread across its various phases has often overlooked the virus's evolving characteristics, such as changes in morbidity and infection rates (Arnaout and Arnaout, 2022). Notably, the transmissibility of variants like Omicron significantly diverges from the virus's initial strains. Although some studies consider different variants, they rarely integrate these variants' specific medical parameters into their models.
- Understanding the interaction strength between countries or regions is key to addressing networked transmission effects. Traditional methods, mainly gravity models or average air traffic data, often fail to account for daily passenger volume dynamics. This oversight can affect the accuracy of analyses, especially since COVID-19 case numbers are reported daily, underscoring the need for models that reflect real-time passenger flow to better assess the aviation network's role in COVID-19's spread.
- Utilizing network compartmental models to understand air transport networks' dynamic impact on pandemics primarily sees studies limited to basic simulations. While some research predicts discrepancies between model-projected disease arrival and actual onset, and others simulate COVID-19's trend within one country, extensive simulations across multiple nations are rare. Moreover, few studies simultaneously examine how the timing, intensity, and duration of travel restrictions affect disease transmission's progression.

3. Data

This section primarily outlines the data required for our experimental analysis. Section 3.1 introduces historical flight plan data to construct the air transport country network and analyze changes in the network's structure and performance. Section 3.2 presents the use of load factor data to estimate the flow of air passengers between countries. Finally, in Section 3.3, we describe the collection and processing of COVID-19 and its variants' data.

3.1. Historical flight plan data

As outlined in Section 1 regarding the motivation for establishing country-wise air transport networks, here we detail the process of utilizing historical flight plan data to construct the subject of our study. Understanding the network diffusion effects of disease transmission, we did not directly select representative countries such as China, the United States, and Europe-27 to construct the aviation network, as done in Sun et al. (2021a). Instead, we initially selected top-ranking airports based on the volume of arriving and departing flights. These airports are typically major aviation hubs and the primary transit points for international flights in each country, making them more likely to serve as the main conduits for disease spread within the international aviation network (Nikolaou and Dimitriou, 2020). Subsequently, we aggregated data from these selected airports to construct the corresponding national aviation networks.

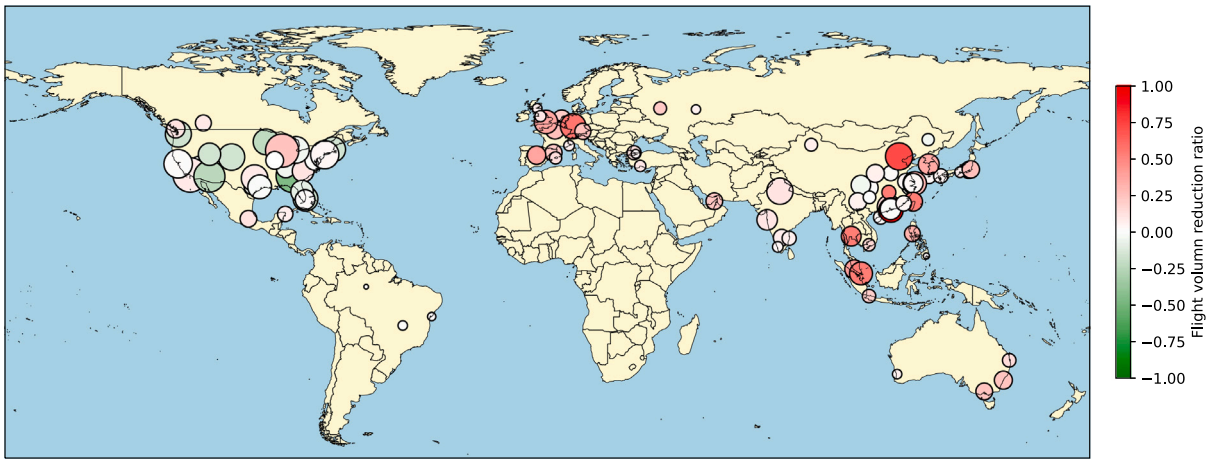


Fig. 1. The geographical distribution map of the selected 105 international airports. In this map, the circles' size represents each airport's takeoff and landing volume in 2019, while the color indicates the change in flight volume between 2019 and 2020. It is important to note that due to varying policies implemented by different countries, some airports actually experienced an increase in flight volume in 2020, resulting in a negative reduction percentage.

Specifically, using data provided by Airports Council International,[1] we selected the top 50 airports in terms of air passenger volume for the years 2019 and 2020, and merged them to form a consolidated database comprising 65 airports. Upon examining the geographical distribution of these 65 airports, it [revealed] that the distribution of airports in certain countries was relatively dense, potentially leading to an unrepresentative sample. Drawing inspiration from the selection criteria outlined in Sun et al. (2022a) and Nikolaou and Dimitriou (2020), we further considered the primary international flight hubs in these countries, taking into account the nuances of their population distribution. As a result, we expanded our initial list from 65 to a total of 105 international airports for our study. The geographical distribution of these airports is depicted in Fig. 1.

It is evident that China and the United States have the highest number of international airports, spanning a wide geographical range, symbolizing the two major aviation transportation subsystems (Sun et al., 2021a). Additionally, countries with significant aviation markets in Europe, such as the United Kingdom, France, Germany, and the Netherlands, also play pivotal roles in global air transportation. Moreover, the air transport networks of these 24 countries and regions represent the current mainstream level of the aviation transportation market. Based on GDP data provided by the World Bank[2] and the International Monetary Fund (IMF),[3] as well as annual flight take-off and landing data from the International Civil Aviation Organization (ICAO),[4] these countries collectively account for 72.3% and 71.3% of the global market,[5] respectively, holding a significant share. Therefore, these airports, spanning across 24 countries and regions, aptly reflect the global aviation network's operational prowess.

After constructing our country-level aviation network for the study, we then proceeded to process the flight plan data required for our experiments. The historical flight plan data comprises flight ID, callsign, aircraft type, airline, registration number, estimated and actual departure and arrival times, all of which are sourced from Feeyo Technology Company Limited.[6] We processed and categorized nearly 80 million flight records into historical flight flow data for the 105 airports under our study. Fig. 2 simultaneously displays the spatial distribution of these 105 international airports and the direction of flight flows between them. It can be observed that airports with a higher volume of takeoffs and landings tend to have closer connections with other airports. Moreover, some airports with a smaller volume of flights also have dense connections with hub airports. Given the intricate connectivity of the aviation network and the transit nature of flights, this suggests that their role in disease transmission is also significant and cannot be overlooked.

In Fig. 3(a), we provide a detailed representation of the monthly changes in the total flight volume of the selected 105 airports in 2019 and 2020. It shows that the flight volume in 2020 was generally lower than that in the corresponding months of 2019. The largest gap between the two years is observed in April and May, after which the difference gradually narrows. Additionally, we have specifically analyzed the changes in transfer rates.[7] The transfer rates in 2019 shows a relatively stable trend. However, there was a significant drop in April 2020, likely due to many countries implementing travel restrictions to curb the spread of COVID-19. In Fig. 3(b), we depict the volume of flights transferred between any two countries or regions. The United States has the highest volume of international flights, followed by the United Kingdom. Influenced by flight control policies, China's international flight volume decreased compared to the same period in previous years, but it maintains strong connections with the major aviation systems.

[1] www.aci.aero.
[2] https://www.piie.com.
[3] https://www.imf.org/.
[4] https://www.icao.int.
[5] GDP share (trillions of dollars): 63.482/87.751 = 72.3%; Flight volume share: 54630078 / 76598000 = 71.32%.
[6] www.variflight.com.
[7] Identification of transfer flights refer to Appendix B.

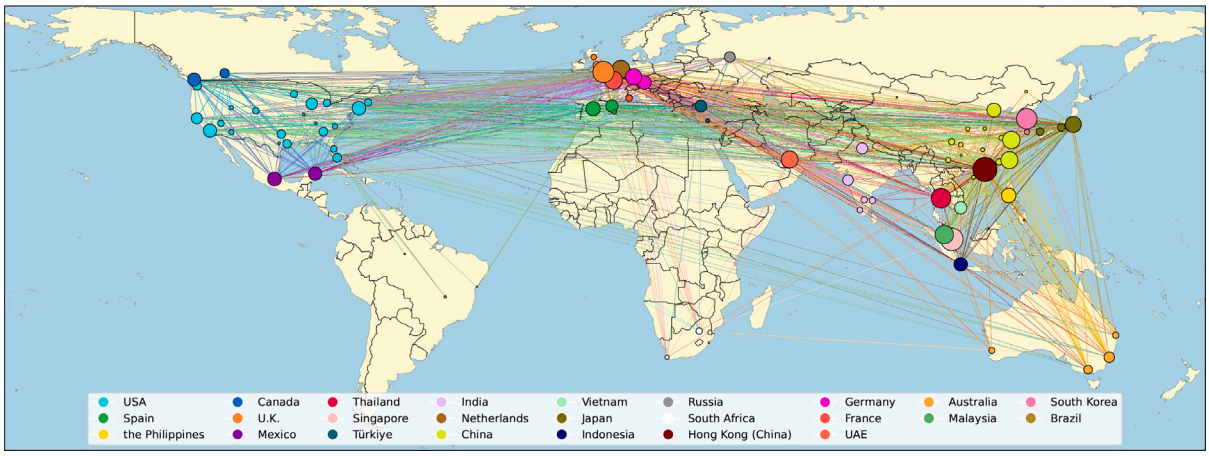


Fig. 2. The geographical distribution of the chosen 105 international airports and their flight volumes and interconnections as of January 2019. In this visualization, the size of the circles indicates the volume of takeoffs and landings at each airport. The color of the circles denotes which country or region the airport belongs to, with airports of the same color indicating they are from the same country or region. Additionally, the lines connecting the airports signify the existence of flights between them, with denser lines indicating a higher volume of flights. For clarity and to effectively illustrate the direction of flight flows, we use the color of the departing airport to represent the color of the connecting flight route.

3.2. Load factor data

In Section 2, it is mentioned that many previous studies directly used flight volume data for experiments. However, many modeling methods or evaluation indicators, such as effective distance and network compartmental models, rely on air passenger data. Directly using flight volume assumes that the load factor of each aircraft during the pandemic is the same. As described in Sun et al. (2022c) and Choi et al. (2022), the differences in the load factor data of aircraft are significant. Moreover, many airlines operate flights with minimal passengers to retain their time slots, so using a single load factor data can reduce the experiment's credibility. Therefore, based on the obtained flight plan data, we further estimated the air passenger data by combining the load factor data of airlines. We obtained the load factor data of the airlines for the countries under study from the CEIC database.[8] Fig. 4 shows the monthly trend of these load factor data from 2019 to 2021. It can be observed that before 2020, the load factor data of airlines was generally maintained at around 80%, with a few data showing significant deviations. After March 2020, there was a noticeable decline, followed by a slight increase, and finally, it stabilized at around 40%.

3.3. COVID-19 and variants data

The term "COVID-19" in the study can encompass all forms of the virus, including its variants (Wells et al., 2020; Sun et al., 2022b). However, it can also specifically refer to the virus during its initial phase of spread (Choi et al., 2022; Sun et al., 2021b). To avoid confusion and ensure clarity, in this Section and Section 5, we use the terms "COVID-19" and "Omicron" to distinctly represent the virus in its initial spread phase and the representative variant, respectively. When both terms are used side by side, they serve to differentiate between the initial virus and its subsequent variant. In other parts of the article, unless otherwise specified, "COVID-19" is used as a general term encompassing all types of viruses.

In the analysis of the first phase, we sourced country-level reported case data from the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University.[9] This dataset provides the cumulative daily numbers of confirmed, deceased and recovered individuals for each country. For the second phase, our focus shifts to the impact of air traffic networks on the spread of COVID-19 variants. Given that the World Health Organization (WHO) has designated the new SARS-CoV-2 Omicron variant as having a very high transmission risk[10] and it has had a significant impact on various industries over the past two years (Tong et al., 2022), we have chosen it as the subject of study for this phase.

Since its detection in South Africa in November 2021, the Omicron variant has spread to more than one hundred countries and territories worldwide and has recently become dominant in several countries. We obtained biweekly moving averages from Our World in Data[11] and procured weekly updated data on Omicron variant cases[12] from the GISAID Initiative.[13] Inspired by

[8] https://www.ceicdata.com.

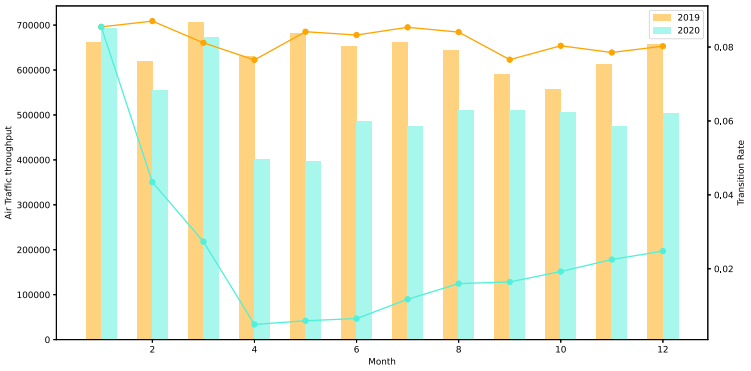
[9] https://github.com/CSSEGISandData/COVID-19.

[10] https://www.who.int/zh/news/item/28-11-2021-update-on-omicron.

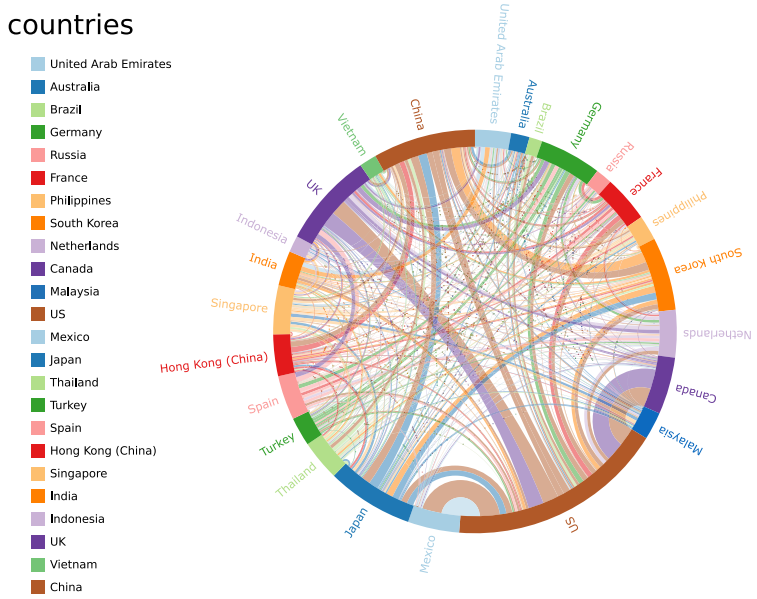
[11] https://ourworldindata.org/.

[12] We make the differentiation in our experiments that COVID-19 refers specifically to the early stages of the outbreak while Omicron is individually represented as its variant.

[13] https://www.gisaid.org/.



(a) Volatility in the targeted 105 airports' total throughput and transfer rates on a monthly basis.



(b) A chord diagram revealing air traffic flow transition between these 24 countries/regions in December 2020.

Fig. 3. Illustration of the monthly fluctuations in air traffic volume and the transfer of air traffic between countries/regions.

the visualization tools provided in Arnaout and Arnaout (2022), we selected countries that publicly [reported] variant infection case numbers, specifically the United States and South Africa, to showcase the evolution trend of their deaths per million versus cases per million in Fig. 5. Each colored counterclockwise loop represents different waves of variants, with dates labeled at the peak of each wave. The dynamics of the variant show significant variations between countries. Within the same country, different variants exhibit distinct characteristics. For instance, Omicron is characterized by a high infection rate but a low mortality rate. Thus, incorporating these medical characteristics is important when constructing mathematical models for analysis. The red curve indicates that the variant is gradually gaining traction, underscoring the need for continued vigilance against the potential impacts of these ever-evolving variants on our daily lives and productivity.

4. Model

In this section, we present how we integrate the method of effective distance with network-based compartmental models to construct a spatio-temporal analytical framework. In Section 4.1, we begin by refining the basic definition of effective distance to incorporate transfer flights, leading to the introduction of metrics such as shortest path distance and random walk distance. These

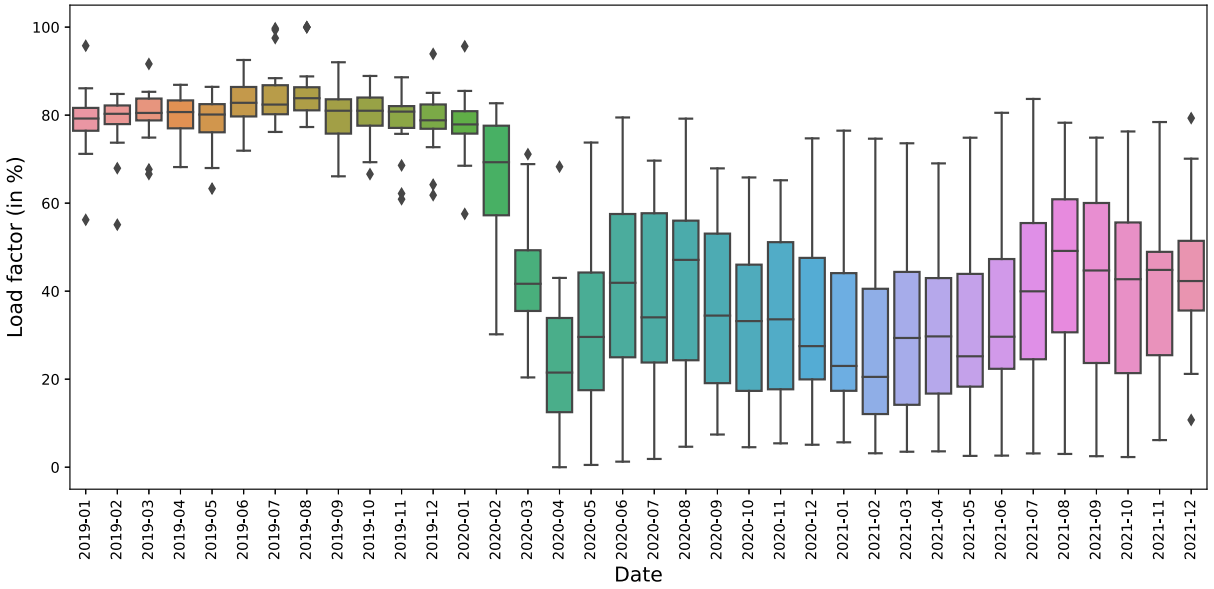
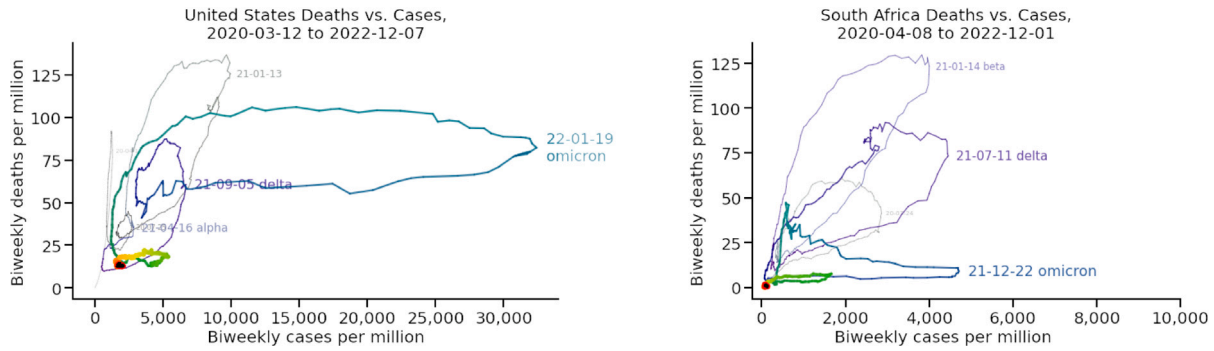


Fig. 4. Monthly variation of load factor data from 2019 to 2021.



(a) United States (b) South Africa

Fig. 5. The illustration of deaths per million versus cases per million for the United States and South Africa. Each colored wave defines the evolution of the particular variant from the rightmost end to the origin. In the temporal diagram, waves progress from gray to purple, followed by two successive blue waves, transitioning to green and culminating in two recent red waves. The intensity and label font increase as it nears the present. The flatter wave denotes the less deadly but more contagious virus, such as Omicron. Refer to https://github.com/rarnaout/Covidcycles.

metrics are employed to compare and analyze the spatial relationship between different indicators and the initial arrival of diseases. Subsequently, in Section 4.2, we introduce the IDVI metric, aimed at providing potential explanations for errors in predicting the initial arrival time of diseases. Finally, in Section 4.3, we elaborate on the modeling process of our SUCRD model, which is based on a system of ordinary differential equations (ODE) to simulate the progression of diseases. This includes detailing the model assumptions, model descriptions, and the method for fitting these parameters through experiments.

4.1. Distance metrics

4.1.1. Effective distance

We adapt the traditional effective distance metric developed by Brockmann and Helbing (2013) to account for the influence of indirect flights. As presented in Fig. 6, the direct air traffic passenger volume between departure airport M and arrival airport N is denoted by f_{mn}, and passengers are also capable of taking flights through transfer airports A and B. Thus, we define total passenger flow from airport M to airport N as

$$F_{mn} = f_{mn} + f_{man} + f_{mbn} \tag{1}$$

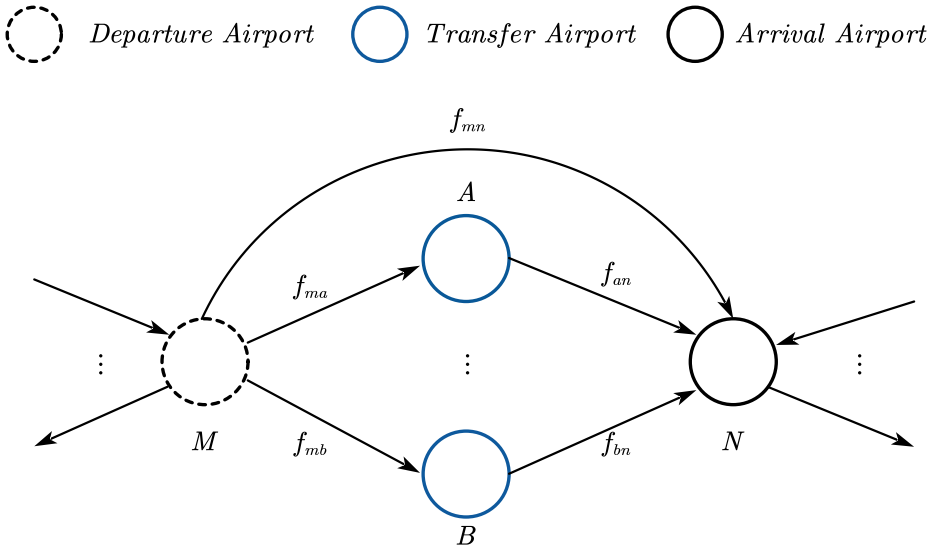


Fig. 6. An example of a simplified airport network. There are four airports in the network where airport M and airport N denoting the departure and arrival airport and the others are the transfer airports.

Then, the corresponding flux percentage P_{mn} can be formulated as

$$P_{mn} = F_{mn}/F_m, \quad 0 \le P_{mn} \le 1 \tag{2}$$

where $F_m = \sum_{m \neq n} F_{mn}$ is the sum of air traffic passenger volume from airport M to all other direct and indirect airports. Therefore, the effective distance D_{mn}^{EF} from airport M to airport N can be shown below:

$$D_{mn}^{EF} = 1 - \log\left(P_{mn}\right), \quad D_{mn}^{EF} \ge 1 \tag{3}$$

It can be seen that the larger the flow fraction P_{mn} is, the smaller the effective distance D_{mn}^{EF} will become, which implies a negative relationship.

It is important to clarify that the consideration of the impact of transfer flights on the initial arrival times of diseases is not an original concept. Choi et al. (2022), in its methodology for calculating effective distance, distinguish between countries connected to the origin country by direct flights and those without (limited to no more than two-stop connections), applying definitions from Brockmann and Helbing (2013) to compute effective distances and assess their correlation with the diseases' initial arrival times in these two groups of countries. Although they acknowledge the influence of transfer flights, our approach examines a more detailed scenario, not differentiating between the two categories of countries but assuming that the epidemic can propagate across any series of flights (Sun et al., 2021d). In our analysis of transfer flights (see Appendix B), we have identified up to seven-stop flights within a single day, indicating that countries directly connected to the origin country by flights are also subject to the importation risk of epidemics caused by transfer flights from other countries. This nuanced approach to calculating effective distance, inclusive of transfer flights, and its implications for accurately predicting initial arrival times are further explored in our experimental discussions. Moreover, our methodology encompasses more intricate effective distance metrics than those considered by Choi et al. (2022) (detailed in subsequent Sections 4.1.2 and 4.1.3). Our comparative analysis with existing literature, presented in Section 5.4, further elucidates the advantages of our fitting results.

4.1.2. Shortest path distance

The shortest path effective distance D_{mn}^{SP} is the smallest total length of the effective distance from airport M to indirectly connected airport N

$$D_{mn}^{SP} = \min_{\Gamma_{mn}} \sum_{(i,j)\in\Gamma_{mn}} D_{ij}^{EF} \tag{4}$$

where Γ_{mn} is the sum of all the possible paths from airport M and N and a path is a combination of consecutive link pairs (i, j).

4.1.3. Random walk distance

Using the shortest path distance, we tacitly consider that the disease is transmitted between countries via one route: the route that minimizes the topological distance between two countries. However, in reality, the pandemic can spread via multiple paths.

Thus, in Iannelli et al. (2017), the authors provide a more realistic approach that takes into account the effect of multiple paths. Starting from two paths, the authors find a generalized relation between multiple paths $\{\Gamma_{mn}\}$:

$$e^{-D_{mn}^{MP}} = \sum_{\Gamma_{mn}} e^{(-D^{\Gamma_{mn}})} \tag{5}$$

where $D^{\Gamma_{mn}}$ is the effective distance associated with path Γ_{mn}. It should be noted that the length of Γ_{mn} can be arbitrary. Following Eq. (5), we can determine the multiple path distance by

$$D_{mn}^{MP} = -\log\left(\sum_{t=1}^{t_{max}} e^{-t} F_{mn}(t)\right) \tag{6}$$

$$F_{mn}(t) = \sum_{|\Gamma_{mn}|=t} \prod_{(k,l)\in\Gamma_{mn}} P_{kl} \tag{7}$$

In Eq. (6), t refers to the path length of Γ_{mn} and $F_{mn}(t)$ is the total probability associated with the path Γ_{mn}. In Eq. (7), P_{kl} is the flux percentage representing the transmission probability from node k to node l.

While the multiple path distance D_{mn}^{MP} integrates information from various direct paths between nodes, it can become computationally cumbersome for large graphs due to the sheer number of potential paths between nodes. In an attempt to capture more nuanced transmission possibilities in complex networks, Iannelli et al. (2017) introduce the concept of random walks from the source to the target. A salient feature of random walks, in contrast to multiple paths, is that nodes can be revisited multiple times. This is accommodated by the function $H_{mn}(t)$, which is designed to permit such revisits. The resulting equation for the random walk distance is:

$$D_{mn}^{RW} = -\log\left(\sum_{t=1}^{\infty} e^{-t} H_{mn}(t)\right) \tag{8}$$

$$H_{mn}(t) = \sum_{k\neq n} P_{mk} H_{kn}(t-1) \tag{9}$$

Here, $H_{kn}(t)$ can be derived using matrix power operations on the flux percentage matrix after specific rows and columns are removed. Further details can be found in Appendix D.

However, it is crucial to note that with the introduction of random walks, the maximum path length becomes infinite. Our updated analyses have shown that the simpler shortest path effective distance D_{mn}^{SP} tends to yield more accurate predictions in our study's context, suggesting that while random walks offer a broader perspective, they might not always be the most appropriate measure for every scenario.

4.2. IDVI metrics

Developed by RAND corporation, the Infectious Disease Vulnerabilities Index (IDVI) is an indicator to identify different countries' vulnerabilities to infectious disease outbreaks (Moore et al., 2017). The metric is generated using factors from seven broad domains: demographic, health care, public health, disease dynamics, political-domestic, political-international, and economic. A higher IDVI score of a country indicates that it is more resilient to disease outbreaks. Additionally, for countries with a similar effective distance to the origin of the pandemic, IDVI provides another aspect of analyzing their difference in the delay of estimated arrival time and importation risks.

4.3. SUCRD models

Network-based compartmental models, which simulate population changes during disease transmission, incorporate several assumptions to simplify the complexities of individual behaviors and interactions (Kraemer et al., 2020; Liu et al., 2022). We will outline the model's assumptions, detailed in both the main text and the appendix. Subsequently, we will provide a detailed description of our modeling process and the simulation process for parameter estimation and optimization.

4.3.1. Model assumptions

We have listed the assumptions incorporated in our model in Table 2, summarizing the characteristics and functions of these assumptions from aspects such as their descriptions, aims, categories, and references.

4.3.2. Model descriptions

Our temporal analysis model draws inspiration from Zhang et al. (2020), which employ a network-based compartmental model to delineate five population states: susceptible, unconfirmed infectious, confirmed, recovered, and dead (SICRD). Utilizing a differentiable ODE solver by the frame of PyTorch, Zhang et al. (2020) solve for unknown parameters within the model, thereby quantifying the impact of controlling population mobility between cities on the progression of the epidemic across these urban areas.

While our model's core framework follows the conceptualization of Zhang et al. (2020) to a certain extent, there are several differences and improvements. First, we have generalized the population groups considered in the model, defining five compartmental states as susceptible, unreported infectious, confirmed, recovered, and death (SUCRD), replacing the unconfirmed infectious states

Table 2
Summary of the model assumptions.

Number	Assumption Description	Category	Aim	Reference
1	Unreported infectious individuals, with or without symptoms, are contagious and can transition to confirmed, recovered, or death states	Model definition	Assisting in defining and simplifying the model's state transition process to align with the disease propagation process	Zhang et al. (2020)
2	An individual in the confirmed state is isolated, with no chance of spreading infection but possible transitions to recovery or death	Model definition	Assisting in defining and simplifying the model's state transition process to align with the disease propagation process	Zhang et al. (2020) and Chen et al. (2020)
3	Individuals in the recovered state have been cured and are immune	Model definition	Assisting in defining and simplifying the model's state transition process to align with the disease propagation process	Zhang et al. (2020) and Chen et al. (2020)
4	There exists inflow and outflow in the susceptible, unreported infectious and recovered states	Model definition	Assisting in modeling the movement behavior of populations in different states to simulate the disease propagation process	Zhang et al. (2020) and Chen et al. (2020)
5	We calculate Omicron recoveries and deaths by multiplying confirmed cases with the corresponding COVID-19 recovery and death rates for that period	Data estimation	Assisting in estimating the recovery and death count for Omicron	/
6	We select cases that visited Wuhan before showing symptoms and assume they got infected on their first day there	Parameter estimation	Assisting in computing the values of parameters D_I and D_C	Zhang et al. (2020)
7	We assume that the time from infection to becoming infectious is negligible	Parameter estimation	Assisting in computing the values of parameters D_I and D_C	Zhang et al. (2020)

in SICRD with unreported infectious states. This adjustment accounts for the population that is unable to adequately detect during international flights (Lau et al., 2020a).[14] Second, we have further refined the model by incorporating the heterogeneity of disease transmission rates and control measures, as well as the time-varying number of international air travelers. These enhancements are both necessary and challenging, as Sun et al. (2021a,b), Yu and Chen (2021) and Lu et al. (2021) highlight the importance of reflecting the heterogeneity of national policies and infection levels, as well as the dynamic effects of actual travel demand. This also addresses the shortcomings of the SICRD model emphasized in Zhang et al. (2020) and identified as areas for improvement in future work. Third, due to the more detailed modeling of actual influencing factors in our SUCRD model, we find that the simulation framework proposed by Zhang et al. (2020) is not suitable for SUCRD. Therefore, we have established a set of parameter evaluation and simulation processes based on the Differential Evolution Algorithm, detailed in subsequent experiments and Appendix H. As stated in Guan et al. (2022) and Zhang et al. (2022), for disease transmission modeling, extending the model and designing an effective simulation algorithm are of equal importance. Therefore, under the premise of improving the model, being able to design a valid parameter evaluation method also constitutes a significant technical contribution. Finally, in Section 5.3, we demonstrate the effectiveness of our model, which accounts for time-varying passenger flows, in enhancing simulation results compared to the model that, akin to the SICRD model's setting, only considers average passenger flows. Fig. 7 depicts the SUCRD model's structure and transition between its five states.

We then describe the five states mentioned above in detail as follows:

- Susceptible state: This is a healthy state where people become infected if they contact an infectious individual.
- Unreported infectious state: People who are in this state may or may not have symptoms, yet they are contagious, and they can transfer to the confirmed state, recovered state, or death state.
- Confirmed state: We presume that a person in this state will be isolated, and there is no likelihood that he or she would spread the infection to others but could move to the recovered or death state.
- Recovered state: People who have been cured and are immune.
- Death state: The fatalities caused by COVID-19.

[14] Note that although our model's unreported state may encompass a broader group of individuals compared with the SICRD model, as Zhang et al. (2020) suggest that the "exposed" state in the SEIR model is included within the unconfirmed infectious state of the SICRD model, the principles of our model parameter estimation and training remain fundamentally unchanged.

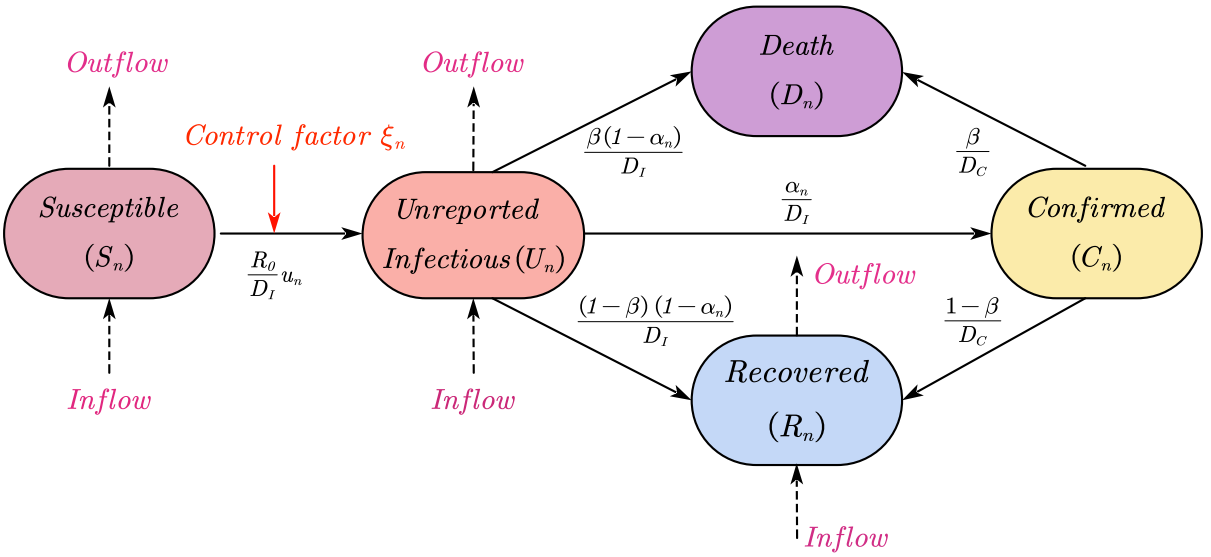


Fig. 7. The state transition diagram of the SUCRD model. S_n, U_n, C_n, R_n, and D_n are the state populations where numbers on the arrow denote the transition probability between states. We assume that there also exists inflow and outflow in the states S_n, U_n and R_n.

Our proposed SUCRD model can be expressed mathematically by the following ordinary differential equations.

$$\begin{aligned}
\frac{\partial s_n}{\partial t} &= -\frac{R_0}{D_I}\xi_n s_n u_n + \frac{\omega}{\gamma}\sum_{m\neq n} P^t_{mn}\left(s_m - s_n\right)\\
\frac{\partial u_n}{\partial t} &= \frac{R_0}{D_I}\xi_n s_n u_n - \frac{1}{D_I}u_n + \frac{\omega}{\gamma}\sum_{m\neq n} P^t_{mn}\left(u_m - u_n\right)\\
\frac{\partial c_n}{\partial t} &= \frac{\alpha_n}{D_I}u_n - \frac{1}{D_C}c_n\\
\frac{\partial r_n}{\partial t} &= \frac{(1-\beta)\left(1-\alpha_n\right)}{D_I}u_n + \frac{(1-\beta)}{D_C}c_n + \frac{\omega}{\gamma}\sum_{m\neq n} P^t_{mn}\left(r_m - r_n\right)\\
\frac{\partial d_n}{\partial t} &= \frac{\beta\left(1-\alpha_n\right)}{D_I}u_n + \frac{\beta}{D_C}c_n
\end{aligned} \tag{10}$$

where $s_n = S_n/N_n, i_n = I_n/N_n, c_n = C_n/N_n, r_n = R_n/N_n, d_n = D_n/N_n$ are the proportions of susceptible, unreported infectious, confirmed, recovered, and dead individuals in country n respectively while N_n represents the population of country n. P is the flux fraction matrix produced by the historical air passenger data (see Section 3), and especially P^t_{mn} denotes the corresponding proportion matrix from country m to country n at day t. All other parameters are detailed in Table 4. We will elaborate on the derivation and estimation of these parameters in the subsequent paragraphs and the appendixes. As Fig. 7 shows, people in states S_n, U_n and R_n could move to other countries and induce population mobility between countries.

In addition, inspired by Zhang et al. (2020), we propose an enhanced heterogeneous intervention term ξ_n in country n to replicate the intervention approach that the government employs to limit contact or interaction among people, including some social-distancing and isolation rules. The form of the intervention function is shown below:

$$\xi\left(t, t_0, t_m, \lambda\right) = \frac{1}{1+\exp\left(\lambda\left(t - t_0 - t_m/2\right)\right)} \tag{11}$$

$$\lambda = 2\frac{\log\left(\frac{1-\varepsilon}{\varepsilon}\right)}{t_m} \tag{12}$$

The intervention function represented as a Logistic model is affected by the following factors: the initiating time t_0 and the control strength λ determined by the ratio of an acceptable reproduction number ε and the duration time t_m. A larger value of λ (a relatively smaller t_m when fixing ε) means the disease can be controlled in a shorter period under a fixed value of ε. As described in Fig. 8, we set t_0 and t_m as different values to compare how these elements influence the function. Notably, since the intervention term ξ_n is multiplied by the infectious term, a smaller value of the intervention term will result in more rapid pandemic prevention. Moreover, in our simulation experiments, we design numerous integration of these parameters in Table 3 to perform different control policies ξ_n in real-world circumstances and evaluate how these various interventions influence the estimated confirmed cases.

Table 3
Parameters of the intervention control.

Parameter[a]	Description	Value[b]
t_0	The initial time of implementing intervention control	0–40
t_m	The periods needed for stabilizing the pandemic spreading	14–139
ε	The acceptable ratio of the reproduction number R_0	0.01–0.2

[a] These parameters can reflect when the intervention policy starts, how long it lasts, and what the control strength is.
[b] The units of t_0 and t_m are days. For example, China implemented a strong lockdown policy in Wuhan on January 23, 24 days after December 31, 2019. It is the initial date of our first validation experiment, then we have $t_0 = 24$. Moreover, if we restrict the reproduction number to the acceptable level εR_0, we could stop the policy implementation. All these parameters are derived from the actual policy taken in every individual country.

Table 4
Parameters of the SUCRD model.

Parameter[a]	Definition	Description	Value[Median][b]	
			COVID-19	Omicron
D_I	Time duration from infected to be confirmed	This is calculated based on patient medical records	8.4	4.0
D_C	Time duration from being confirmed to be recovered or dead	This is calculated based on patient medical records	9.2	5.8
β	The fatality proportion	The average ratio of deaths to the total number of infectious people	0.023	0.015
w	The average migration ratio	The total air traffic flux between countries divided by the total population	0.042%	0.042%
R_0	Basic reproductive number	The average number of cases directly infected by one infectious case	2.33	9.5
α_n	The rate of transferred confirmed cases among all cases in the country n	Estimated by fitting the data	0.93	0.88
γ	The ratio of air transportation in total travel modes	Estimated by fitting the data	90%	92%
$U_n(0)$	The initial number of unreported infectious cases in country n	Estimated by fitting the data	414	522

[a] Calculation of w, D_I and D_C refers to Appendices C and G.
[b] The estimated parameters are recorded under 50 parallel searches.

Detailed descriptions and estimation for parameters $D_I, D_C, \beta, w, R_0, \alpha, \gamma, U_n(0)$ are shown in Table 4, where D_I, D_C and β are retracted and calculated from patient clinical records while w is estimated in historical flight plan data (see Appendices G and C). It is evident that the medical characteristics of the COVID-19 virus in its early stages differ significantly from those of the Omicron variant. Specifically, Omicron has a shorter incubation period, and higher transmissibility, but a lower fatality rate compared to the initial strain of COVID-19. We have incorporated these factors into our model to better assess the impact of air transportation networks on the spread of different types of viruses. The remaining parameters R_0, α, γ, and $U_n(0)$ need to be estimated by fitting the number of confirmed cases for the corresponding countries in different scenarios. Thus, we define the loss function as

$$\min_{R_0, U_{n(0)}, \alpha_n, \gamma} L = \sum_n \sum_t \left[\left(c_n(t) N_n \right) - \left(C_n^*(t) \right) \right]^2 \tag{13}$$

where $c_n(t)$ is the predicted confirmed proportions in country n at time t and N_n denotes the population of country n, while $C_n^*(t)$ is the corresponding reported confirmed cases. In Appendix H, we provide a detailed derivation and estimation of parameters such as $U_n(0)$. We solve this optimization problem by applying the Differential Evolution Algorithm via the Python toolkit *Geat Py*.[15]

5. Experiment

In this section, we are dedicated to presenting our detailed experimental analysis. Section 5.1 introduces the software and libraries used in the experiments, as well as the execution process of the experiments. Section 5.2 displays the spatial effects experiment, including analysis of fitting effects and error analysis, while Section 5.3 presents the temporal effects experiment, encompassing detailed validation processes and simulation experiments. Section 5.4 summarizes our experimental results and provides a comparative analysis with previous literature findings.

[15] https://github.com/geatpy-dev/geatpy.

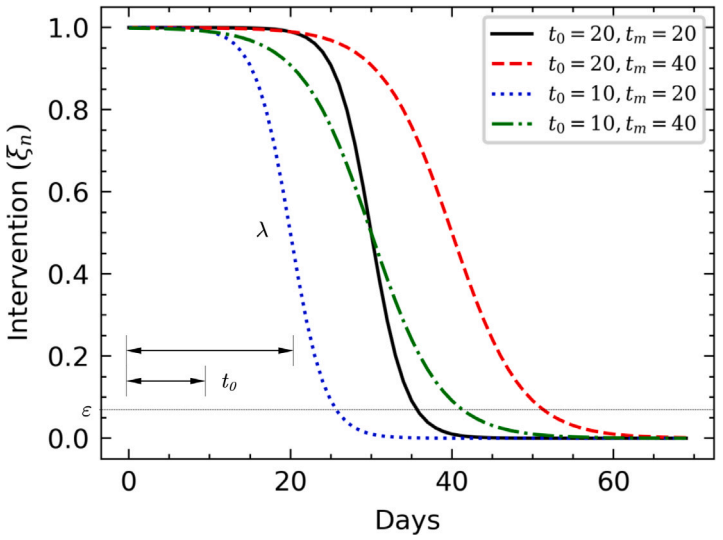


Fig. 8. The functional form of the intervention control. t_0 is the time that control interventions are to be implemented and t_m denotes the duration. We set different combinations of t_0 and t_m to examine how it affects the intervention function ξ.

5.1. Experimental setup

Our experiments were conducted on a local computer equipped with a 3.80 GHz Intel i7-10700K CPU, utilizing Python 3.8.8 as the primary coding framework. For the spatial effects analysis, we employed software packages such as *Numpy*, *Pandas*, and *Networkx* for analytical processing. In the temporal effects analysis, we further utilized the ODE-solving module from *Scipy* and the Differential Evolution Algorithm framework from *GeatPy* for parameter estimation and simulation implementation of the ODE system.

The implementation of our experiments is delineated as follows. Initially, we constructed various effective distance metrics by estimating air passenger numbers between different countries. These metrics were then correlated with the initial arrival times of diseases through linear regression, allowing us to analyze the outcomes and associated errors. Subsequently, we processed daily reported infection, recovery, and death figures from each country, amalgamating the daily transferring air passenger numbers between countries. This data was then inputted into the established ODE system for simulation experiments, during which parameters were optimized and evaluated (see Appendix H). Finally, the optimized parameters were inputted into the ODE-solving framework, facilitating a comparison between the simulated transmission curves and the actual changes in infection numbers.

5.2. The spatial effects analysis

The spatial effects of air transport networks on the spread of COVID-19 are analyzed by quantifying the first arrival time of the pandemic in the countries studied. Taking the initial phase of the pandemic as an example to illustrate our analysis process, we recognize China as the early epicenter of the COVID-19 outbreak. For any given country i, we define T_i^1 as the date of its first confirmed case. The interval $T_i = T_i^1 - T_{China}^1$ represents the time duration for COVID-19 to spread from China to the country i.[16]

The propagation time T_i could be influenced either by the geographic distance D^G or by the effective distances $\{D^{EF}, D^{SP}, D^{RW}\}$ from country i to China, as mentioned in Section 4.1. Utilizing the most recent air passenger data from January 2020 and taking into consideration the influence of transfer flights, we can calculate the effective distance metrics from China to other nations. Our utilization of linear regression to analyze the relationship between effective distance and the initial arrival times of diseases is informed by both foundational and recent research methodologies. Initially, the concept "effective distance" was explored in seminal works focusing on the spread of SARS and H1N1 (Brockmann and Helbing, 2013), where linear regression was utilized to simplify the complex spatio-temporal patterns of disease propagation into a more manageable wave-like model. It highlighted a stronger correlation between effective distance and disease arrival times than geographical distance, establishing a precedent for our analysis. In line with this, our research adopts China and South Africa as the initial outbreak points for COVID-19 and Omicron, respectively, mirroring the original article's approach of selecting specific locations. Furthermore, our construction of effective distance metrics, grounded in international air passenger numbers, aligns with the original methodology, aiming to compare different effective distances with the arrival times of the prevalent infectious diseases (Sun et al., 2021a). Recent literature (Adiga et al., 2020; Kuo and Chiu, 2021; Choi et al., 2022) also embraces linear regression to dissect the dynamics between COVID-19's arrival times and

[16] For the detailed data acquisition and correction process, see Appendix F.

Table 5
Considering vs. not considering transfer flights: effect on fitting performance for effective distances.

Distance metrics	COVID-19		Omicron	
	Without transfer	With transfer	Without transfer	With transfer
Effective distance	0.50	**0.67**	0.24	**0.27**
Random walk effective distance	0.63	**0.71**	0.32	**0.40**
Shortest path effective distance	0.75	**0.84**	0.36	**0.43**

effective distances, reinforcing the relevance and applicability of this analytical technique in current pandemic scenarios. Our study's alignment with these methodological precedents, coupled with our specific problem settings and metric constructions, justifies our use of linear regression. This method facilitates direct comparisons with prior findings and underscores the advantages of our refined metrics, providing a solid rationale for our analytical choices.

Fig. A.18 displays the correlation between the COVID-19 transmission time T_i and the four distance metrics. The effective distance is shown to be a more reliable predictor of an epidemic's initial arrival time compared to the geographic distance. Further analysis indicates that the shortest path effective distance, which is considered the most efficient path through potential indirect connections between airports, provides the most accurate predictions. This finding contrasts with previous research that posited diseases tend to spread via multiple routes (Iannelli et al., 2017). One possible explanation for this discrepancy is the rapid enactment of global travel restrictions following the initial COVID-19 case detection, which significantly altered the international air transport network.

Turning our attention to the Omicron variant, its transmission is traced back to South Africa.[17] Using a methodology analogous to our COVID-19 study, we evaluate the capacity of the geographic distance D^G, along with the effective distances D^{EF}, D^{SP}, and D^{RW}, to predict the first arrival time of Omicron. It is important to note that the countries selected for the COVID-19 and Omicron plots differ. This is primarily due to the limited availability of data for the Omicron variant in certain countries. Consequently, the sets of countries in the two analyses are not identical. As illustrated in Fig. A.19, the results indicate that the effective distance metrics, particularly the shortest path effective distance D^{SP} with an R^2 of 0.43, are more robust predictors of Omicron's first arrival time compared to the geographic distance, which has an R^2 of 0.22. This finding is consistent with our analysis of the original COVID-19, where the shortest path effective distance emerged as the most accurate predictor. The random walk effective distance D^{RW} also exhibits significant predictive power with an R^2 of 0.40.

However, the R^2 values for Omicron are notably lower than those observed for COVID-19. This reduction can be attributed to several factors, including the limited data availability and potential inaccuracies in reported initial arrival times for some countries. Furthermore, the rapid spread of the Omicron variant suggests a shortest-path propagation mechanism, much like the original COVID-19. Yet, given the challenges in data collection and the inherent uncertainties tied to the Omicron variant, these results should be interpreted with caution. For a detailed description of the data collection and processing specific to Omicron, readers can refer to Appendix E.

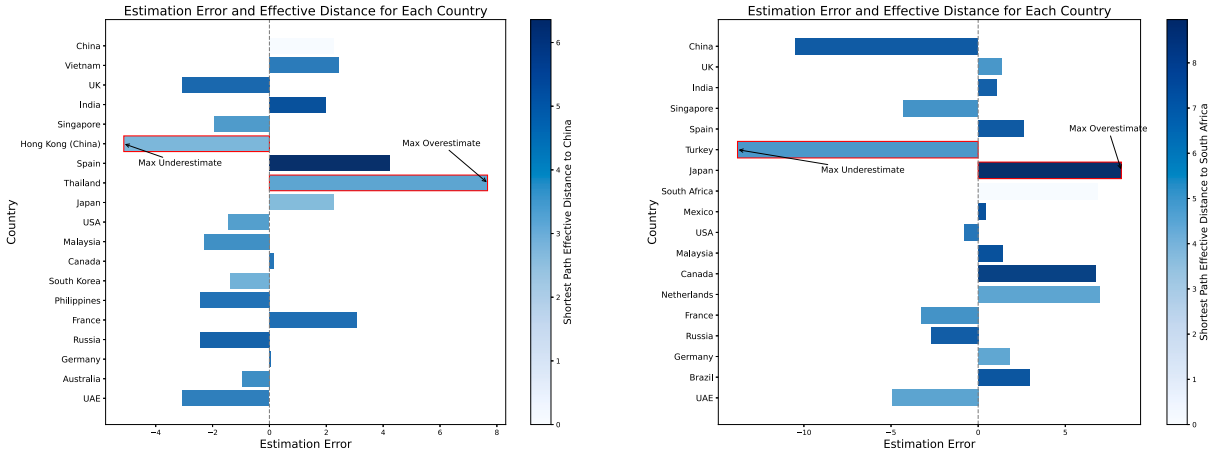
To validate the impact of incorporating transfer flights into our improved metrics on simulation outcomes, we have compared them against the original metrics that only consider direct flights. As shown in Table 5, we present the comparative fitting effects (R^2 values) of considering versus not considering transfer flights for the effective distance, random walk effective distance, and shortest path effective distance against the initial arrival time of diseases in both COVID-19 and Omicron scenarios. The fitting capabilities of the three metrics that account for transfer flights surpass those that only consider direct flights in both scenarios. Moreover, the shortest path effective distance continues to exhibit the best fitting effect even without considering transfer flights. Therefore, transfer flights play an indispensable role in the disease transmission process and can aid in accurately predicting the disease's initial arrival time.

In Fig. 9, we present the estimation error of the first arrival time for both the initial phase of COVID-19 and the Omicron variant periods. This error quantifies the disparity between the predicted arrival times, derived from the shortest path effective distance (SPED), and the actual reported times of arrival. Our choice to employ SPED for this analysis is grounded in its demonstrable accuracy for both the initial phase of COVID-19 and the Omicron variant in our tests.

For the initial wave of COVID-19, most countries demonstrate estimation errors that hover within a −3 to 3-day window. Yet, certain outliers, notably Hong Kong (China) and Thailand, deviate significantly from this norm. Transitioning our focus to the Omicron variant, the estimation landscape broadens. Countries, like Turkey and Japan, exhibit errors that surpass 6 days, underscoring a marked divergence between predictions and actual events.

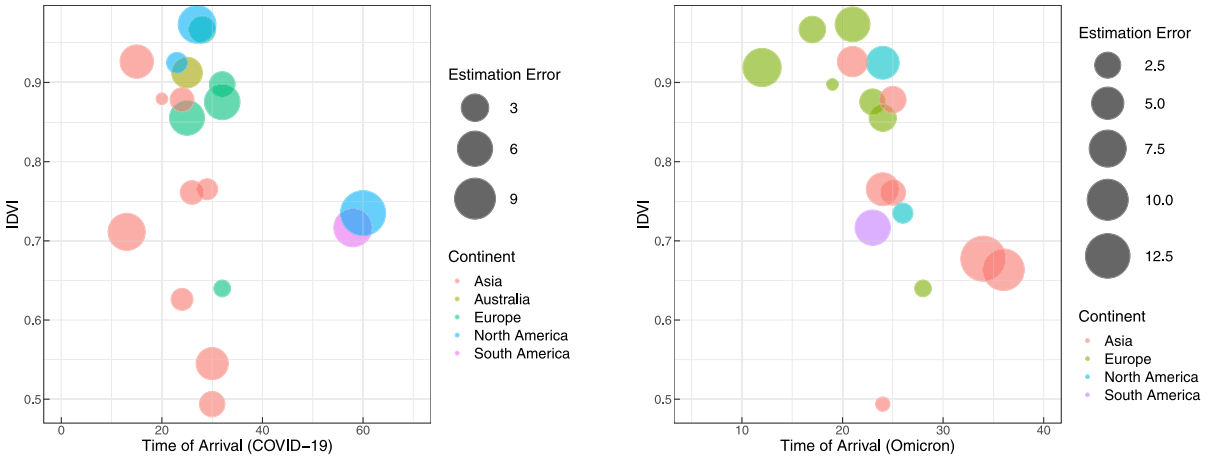
Expanding upon this analysis, and referencing the methodology from Adiga et al. (2020) using IDVI metrics to elucidate this manifested discrepancy, Fig. 10 highlights the intricate association between IDVI scores and the time of arrival. A discernible trend emerges, indicating a negative correlation between the two. This suggests that nations with robust health infrastructures, as reflected in higher IDVI values, tend to swiftly detect and report pandemics. In stark contrast, countries with diminished IDVI scores, such as Mexico and Brazil during the COVID-19 phase, grapple with reporting delays, leading to magnified estimation errors. As further detailed in Section 4.2, we highlight the synergy between the IDVI indicator and effective distance metric. Together, these metrics provide a valuable tool for authorities to assess disease importation risks and develop informed epidemic prevention strategies.

[17] For this analysis, we incorporated case and flight data from South Africa's top four international airports.



(a) COVID-19 (b) Omicron

Fig. 9. The plot presents the estimation errors in predicting the arrival time of both COVID-19 and the Omicron variant across various countries using the Shortest Path Effective Distance (SPED). For each variant, individual bars represent countries, with their lengths indicating the estimation error and their color gradient reflecting the SPED value. Countries with the most pronounced overestimation and underestimation errors are emphasized with red outlines.



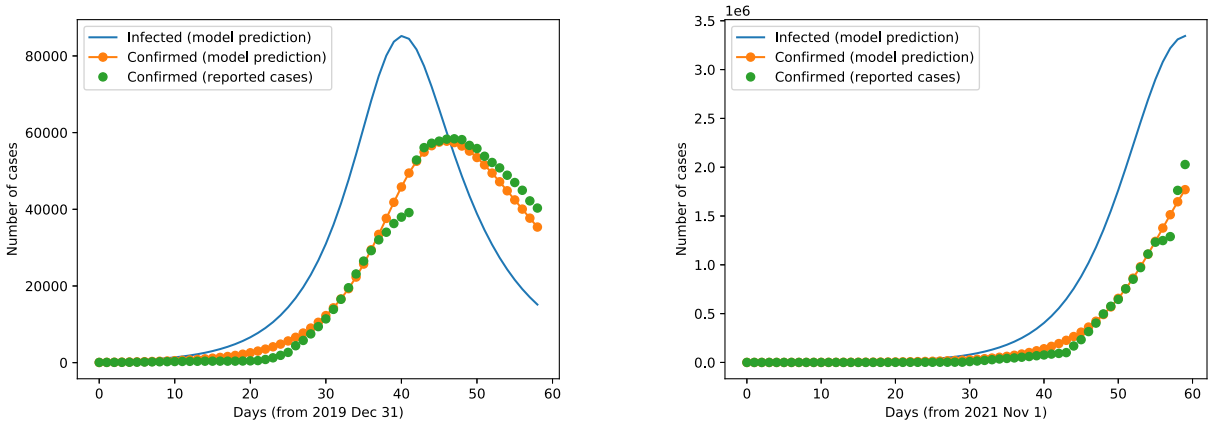
(a) COVID-19 (b) Omicron

Fig. 10. The relationship between IDVI metrics and reported arrival time for both COVID-19 and Omicron. Each circle represents a country or region, with the size of the circle determined by the absolute estimation error of the arrival time, and the color of the circle grouped by the continent to which the country belongs. This visualization facilitates a comparative analysis of the determined first arrival times, estimation errors, and their corresponding IDVI values for certain countries.

Concluding our spatial impact analysis, we discern that the initial arrival time of diseases correlates more significantly with the shortest path effective distance, likely due to the structural impact of travel restrictions on the air traffic network during the pandemic. This renders the random walk distance's concept of recurrent visits to specific nodes less applicable under current circumstances. Moreover, we validate that the fitting capabilities of our improved metrics, which consider transfer flights, surpass those of the original metrics that only account for direct flights. We employ the shortest path effective distance to predict the disease's initial arrival time and attempt to provide possible explanations using the IDVI metric. It is noteworthy that the fitting results in the Omicron scenario are not as satisfactory, largely due to the absence of precise Omicron data. Therefore, more detailed simulation experiments, including considering more complex transmission patterns during this scenario, require further exploration. We summarize related future directions in Section 7.2.

5.3. The temporal effects analysis

We assess the temporal influence of air transport networks on the spread of COVID-19, considering its progressive nature. To this end, we undertake a multifaceted analysis, encompassing both validation and simulation experiments, during the initial phase of



(a) COVID-19 (b) Omicron

Fig. 11. Validation of our model across all selected countries during both COVID-19 and Omicron stages. The x-axis represents the duration in days from the respective starting point. The predicted confirmed cases (orange dots), align closely with the actual reported cases (green dots). This alignment underscores the robustness and precision of our model.

COVID-19 and the Omicron period. As delineated in Section 3, for clarity, we use "COVID-19" and "Omicron" to denote these respective stages. Furthermore, we broaden our validation experiments to include the scenario with multiple outbreak origins, ensuring a further evaluation of our model's generalizability.

5.3.1. Validation experiments

We assess the accuracy of our model by comparing its estimated outcomes for the chosen countries/regions with the actual reported cases across three distinct stages.

Stage 1 (COVID-19): From 2019.12.31 to 2020.02.28

As depicted in Fig. 11(a), the predicted confirmed cases[18] from our proposed model are plotted against the actual reported cases for all the countries selected in our study, demonstrating a strong alignment. A notable inflection point in the green dotted curve may be attributed to the revised diagnostic criteria adopted in Wuhan to encompass all clinically diagnosed cases, while other cities continued to rely on the previous diagnostic tools. The infected cases, which include both confirmed and unreported infectious cases, are represented by the blue lines, indicating a significant discrepancy from the officially reported cases. This implies that during the early outbreak stages, many infectious individuals may have gone undiagnosed or were asymptomatic, largely due to the absence of effective diagnostic and treatment measures.

Additionally, it is important to recognize that the characteristics of cases and the policies implemented in response to the outbreak can differ across countries. To provide a more extensive validation of our assumptions, especially in the context of spatial mobility, we delve into detailed experimental results for each country, as depicted in Fig. 12. Given that our study encompasses 24 countries/regions, for clarity and representativeness in our validation, we have selected a subset of countries/regions for a detailed comparison of the evolution of case numbers. Our selection is primarily based on the following criteria: inclusion of the origin country of the outbreak, countries at different stages of infection severity, and typical aviation hubs. For instance, in Fig. 12, China is the epicenter of the outbreak, while the aviation systems of the United States and the United Kingdom serve as major global hubs. The other countries showcased represent varying levels of infection severity during the respective phase of the pandemic.

As depicted in Fig. 12, the SUCRD model incorporating P^t_{mn} generally delivers the most accurate results. However, in certain countries, its performance does not show a pronounced difference compared to the SUCRD model with P_{mn}.[19] This might be attributed to the fact that the temporal variation in air passenger demand in these countries is closely aligned with the average traffic demand. Furthermore, both models significantly outperform the SUCRD model that neglects spatial air traffic mobility, underscoring the pivotal role of the aviation network in the early stages of the COVID-19 outbreak. It is worth emphasizing that these findings are consistent across countries that exhibit a rising trend in confirmed cases, demonstrating the robustness of our model in fitting various epidemic trajectories.

However, in the case of China, we observe that the fitting curves of all three models are nearly identical, especially in the mid-to-late stages of the curve, closely matching the actual infection numbers. The distinguishing factor among these models is whether they consider (time-varying) air passenger data. This suggests that, during this period for China, the significance of considering spatial air mobility seems to diminish. Following Kraemer et al. (2020), this period of epidemic spread in China before February 10, 2020,

[18] The number of confirmed cases refers to the daily existing confirmed cases rather than the cumulative confirmed cases.

[19] The monthly average air traffic mobility matrix.

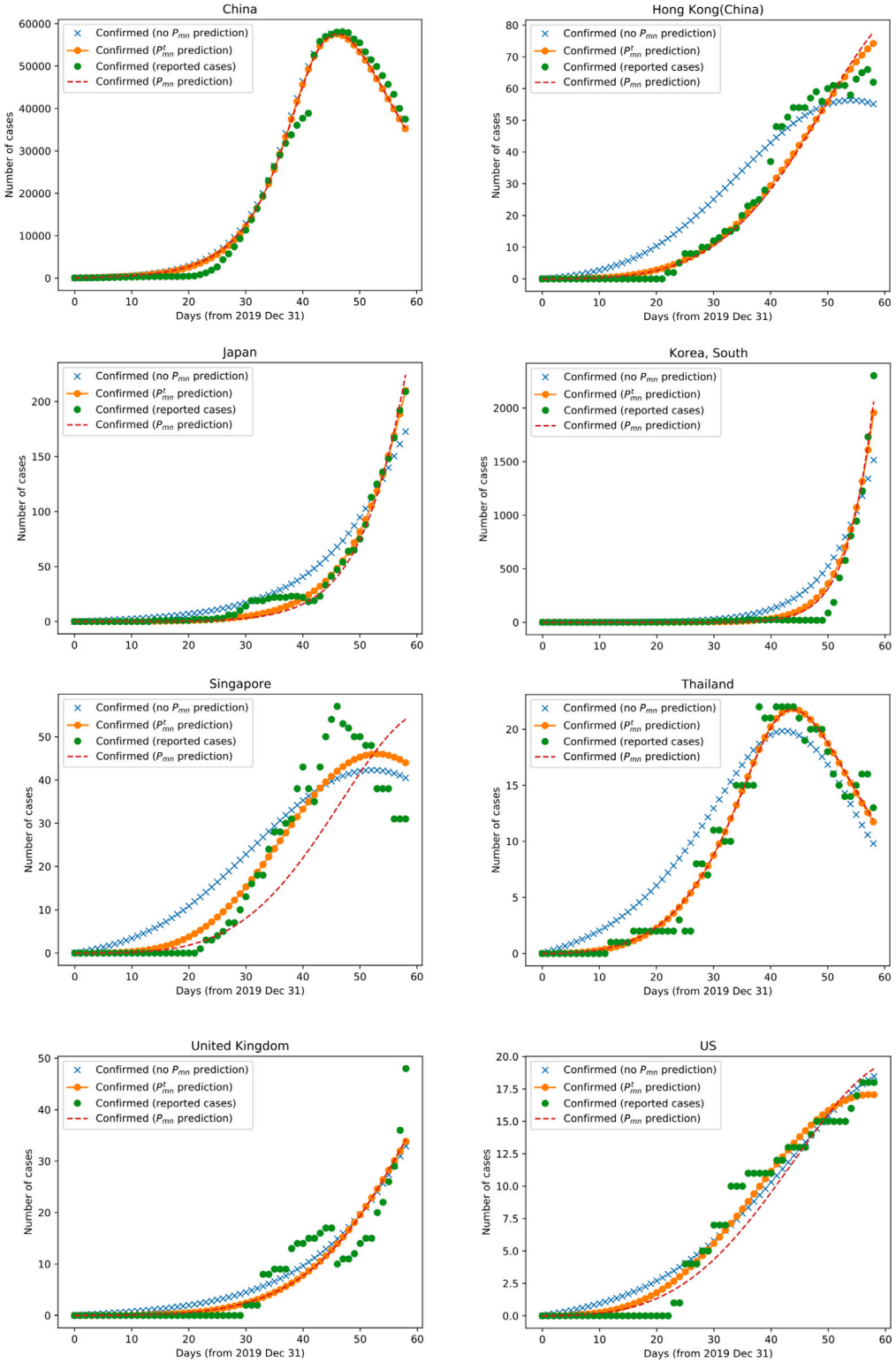


Fig. 12. Comparison between the simulation outcomes using three distinct models and the actual reported case numbers in China and other representative countries, as per the selection criteria mentioned earlier, during stage 1. The eight sub-figures presented correspond to the experimental results for the following selected countries: China, Hong Kong (China), Japan, South Korea, Singapore, Thailand, United Kingdom, and the United States.

is defined as the early epidemic phase, while the subsequent period is termed the "widely spread" phase. Further referencing (Chen et al., 2020; Murano et al., 2021), although they do not explicitly define the "widely spread" phase for domestic epidemics, they provide a relative definition—namely when the disease is distributed across all states within a country. Considering the distribution characteristics of the epidemic in China during this period and the segmentation method of Kraemer et al. (2020), we also define the simulation period after February 10 as China's "widely spread" phase. Therefore, the analysis suggests that during this phase, the impact of travel restrictions may become less significant and does not greatly influence the overall epidemic trend. Our findings are consistent with the conclusions drawn in previous studies proposed by Kraemer et al. (2020), Chen et al. (2020) and Murano et al. (2021). However, while these studies are based on domestic scenarios, our analysis focuses on the international aviation network, examining the impact of travel restrictions between countries on the spread of COVID-19. This perspective offers a broader understanding of the pandemic's dynamics in the context of global air travel and its implications for international policy-making.

Stage 2 (Omicron): From 2021.11.01 to 2021.12.30

We further validate our assumptions by comparing the estimated results with the reported cases using Omicron data. As depicted in Fig. 11(b), our predicted confirmed cases align closely with the reported confirmed cases, though the inferred number of infected individuals is higher than the previous estimates. To provide a detailed demonstration of our model's capability to fit the infection curve for individual countries, we continue to select representative countries for simulation experiments based on the criteria mentioned earlier: the origin of the outbreak, major aviation hubs, and countries representing different stages of the epidemic's progression.

Similar to the COVID-19 stage, simulation outcomes from the SUCRD model with P^t_{mn} generally align with the trend of the reported confirmed cases in these countries, as illustrated in Fig. 13. In some countries, there are subtle differences compared to the SUCRD model with P_{mn}. In most countries, models that do not consider air traffic mobility typically exhibit the poorest fit. However, in countries like the United States and the United Kingdom, the fit between the three different simulation models and the actual infection curve is nearly indistinguishable. Upon further investigation, it was confirmed that during this period, the Omicron variant had already spread widely across all regions of the aforementioned countries. Thus, we can conclude similar to that of Stage 1: During the epidemic's widespread phase, the role of air transport networks diminishes, and implementing travel restrictions may have limited impact in altering the course of the outbreak.

Stage 3 (Additional validation): From 2020.09.01 to 2020.10.30

Furthermore, we extend our model to address a complex scenario where a global outbreak originates from multiple sources and spreads extensively to gauge its accuracy and adaptability. We select a mid-phase period during the initial outbreaks of COVID-19 and Omicron (from 2020.09.01 to 2020.10.30), as mentioned in Sun et al. (2023b), which is influenced by multiple-origin outbreaks. It is important to note that our simulation process during this phase, including data processing, parameter estimation, and simulation validation, remains consistent with the previously described procedures. The challenge lies in the selection of initial parameters for the simulation phase. Given the complex and variable environment of this phase, determining the initial number of unreported infectious individuals and accurate medical parameters becomes difficult. We treat the unreported infectious numbers of each country as optimization variables and use the early phase model parameters as a preliminary attempt. We discuss this experiment in our future work in Section 7.2.

Given the scarcity and difficulty in obtaining and estimating parameters for this situation, we select four countries that show relatively good simulation results for demonstration. As depicted in Fig. 14, even in situations where data quality deteriorates and is susceptible to fluctuations, our model aligns reasonably well with the reported cases. This exploratory endeavor potentially demonstrates the model's capability to generalize across outbreak scenarios with multiple origins. In essence, our findings suggest that when a disease has become widespread, the influence of spatial air traffic mobility, representing the air transport network, becomes less consequential to the pandemic's spread. This is evident as we observed minimal discrepancies between estimated and actual reported cases across the three scenarios, regardless of whether we accounted for (time-varying) air traffic volumes.

5.3.2. Parameter sensitivity analysis

Furthermore, it is important to recognize that while certain parameters in our model, namely D_I, D_C, β, and w, are directly sourced from official statistics, others such as R_0, $U_n(0)$, α_n, and γ are derived through an optimization process tailored for data fitting. As a result, any adjustments to the directly sourced parameters can influence the derived parameters and, consequently, the outcomes of our simulations. To gauge the robustness and reliability of our proposed model, we delve into an analysis that examines how variations in the provided parameters impact the optimized parameters. Notably, since the parameter β solely influences the fatality count and does not impact the derived parameters, its variations do not need to be factored into the model's optimization process (Zhang et al., 2020). Our findings from the parameter sensitivity analysis, particularly concerning D_I, D_C, and w, are thoroughly presented in Tables A.7, A.8, and A.9.

The sensitivity analysis experiments reveal that R_0 is the most stable parameter, showing minimal fluctuations relative to other derived parameters. As depicted in Tables A.7 and A.8, while the parameter $U_{China}(0)$ does exhibit relative sensitivity to changes in D_I and D_C, referring to the simulated results of initial unreported cases in Fig. H.21, the overall fluctuations remain within an acceptable and reasonable range. Specifically, all parameters exhibit an upward trend with increasing D_C, with the magnitude of this change being more pronounced in Table A.8. This suggests that our derived parameters are particularly sensitive to variations in D_C. As presented in Table A.9, the optimized variables, on the whole, remain relatively consistent in the face of changes in w. Additionally, the estimated values for the parameter α_{China} consistently exceed 0.8, indicating a high proportion of individuals transitioning from the unreported infectious state to the confirmed state during the early stages of the outbreak. Another intriguing observation from our analysis is that the optimized value for γ consistently remains above 0.85. This reaffirms the notion that, even amidst the pandemic, air transportation remains the predominant mode of inter-country travel. In summation, our model and its estimated parameters exhibit strong resilience and robustness against fluctuations in other sourced parameters.

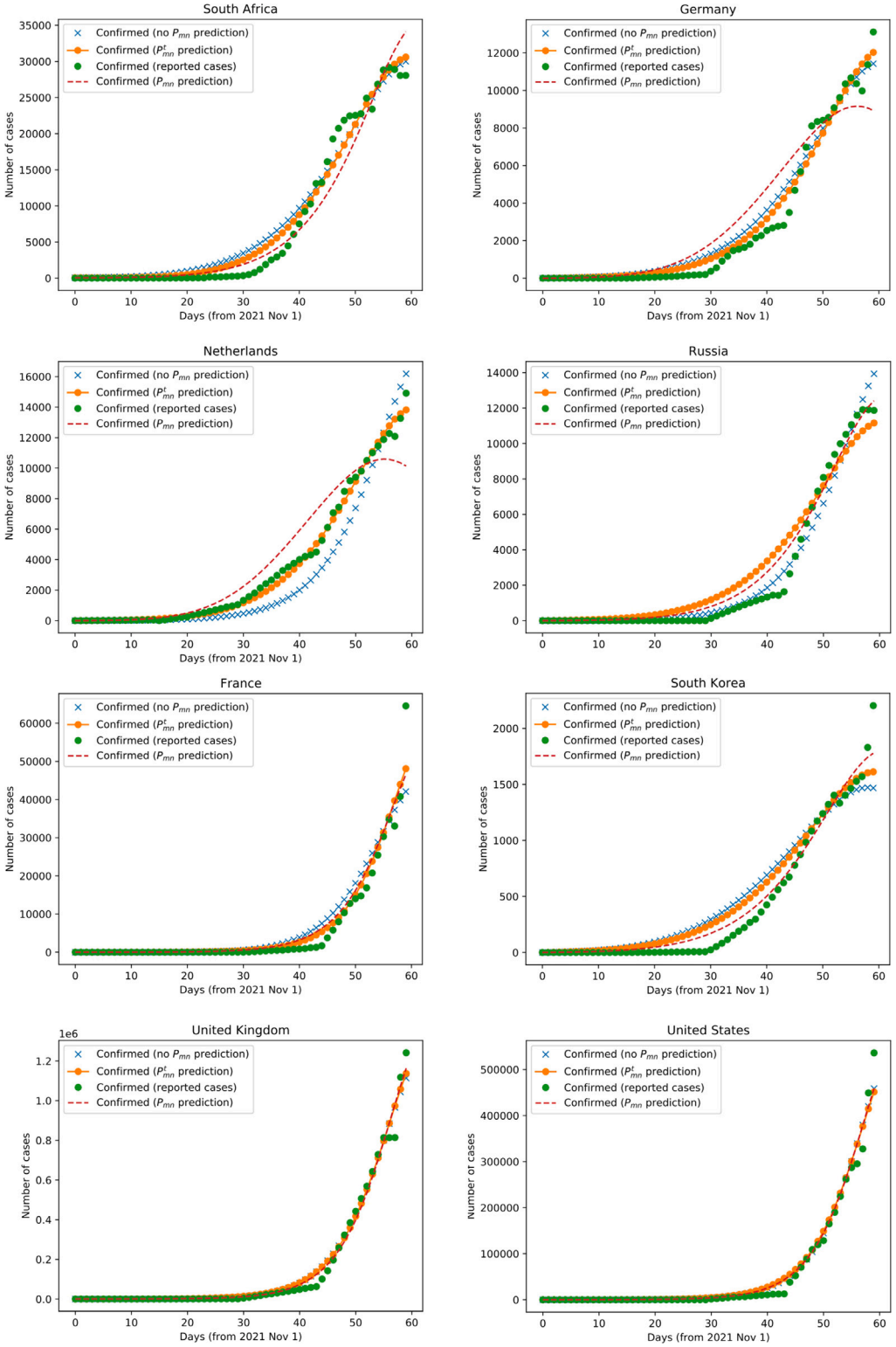


Fig. 13. Comparison between the simulation outcomes using three distinct models and the actual reported case numbers in South Africa and other representative countries, as per the selection criteria mentioned earlier, during stage 2. The eight sub-figures presented correspond to the experimental results for the following selected countries: South Africa, Germany, Netherlands, Russia, France, South Korea, United Kingdom, and the United States.

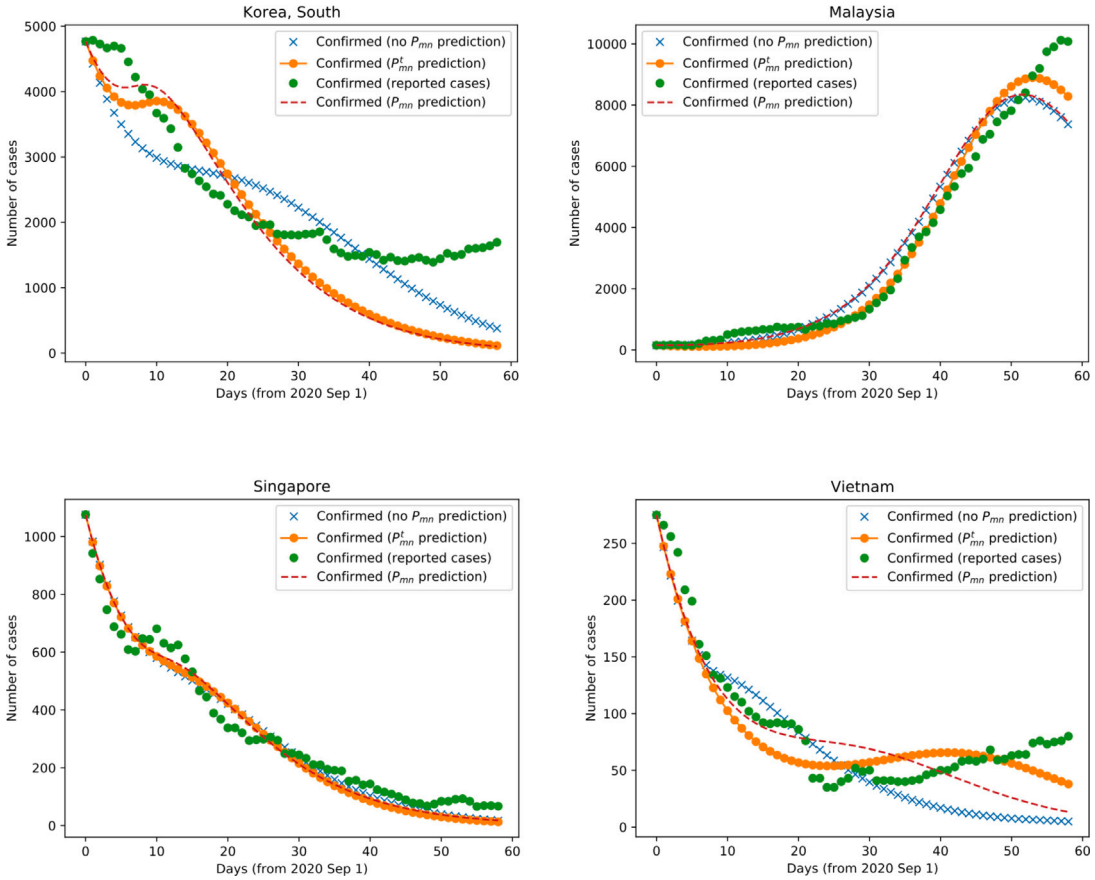


Fig. 14. Further comparison between the simulation outcomes using three distinct models and the actual reported case numbers in four selected representative countries, as per the selection criteria mentioned earlier, during stage 3. The four sub-figures presented correspond to the experimental results for the following selected countries: South Korea, Malaysia, Singapore, and Vietnam.

5.3.3. Scenario analysis

In this section, we delve into two scenarios encompassing both local policy interventions and the implementation of international travel restrictions. We aim to provide a detailed analysis of the efficacy of these containment strategies in curbing the progression of the epidemic.

As outlined in Section 4, we incorporate a control factor, denoted as ξ, into our model. Here, t_0 signifies the moment when an intervention is initiated, while t_m indicates its duration. It is crucial to note that the control strength, λ, is determined by both t_m and ε. With a fixed value of ε, a larger t_m suggests a more lenient intervention. In Fig. 15, we depict the trajectories of simulated confirmed cases[20] in the initial countries of both stages, namely China and South Africa, across various intervention start time t_0 and duration t_m. As illustrated by these diagrams, the timing and intensity of interventions play pivotal roles in mitigating the pandemic's spread.

Furthermore, drawing inspiration from Zhang et al. (2020), we have delineated four scenarios, each representing distinct combinations of t_0 and t_m. Among these, the "actual scenario" mirrors the parameter combination referenced in Section 4, used during our validation experiments. The earlier an intervention is implemented, the greater the potential to suppress the virus's spread. Conversely, even a stringent intervention, if imposed later, tends to be less influential. This is evident from Fig. 15, where late interventions, regardless of their intensity, result in a higher tally of confirmed cases. Additionally, we discern that the incremental benefit of reducing infections diminishes with prolonged measures. This underscores the need for a judicious balance between the intensity of controls and their associated costs.

Reducing air passenger volume on specific origin–destination (O-D) routes critically influences the dynamics of disease spread. In our study, we focus on two primary routes: China-South Korea and South Africa-United Kingdom. These routes were chosen based on the origin of the epidemic and the country with the highest number of infections, excluding the origin country, during the respective periods. We aim to delve into the ripple effects of reducing air traffic on these specific routes and how it impacts the

[20] The simulation experiments in this section are designed to run for 60 days, and Figs. 15 and 17 illustrate the simulation findings for the $60th$ day.

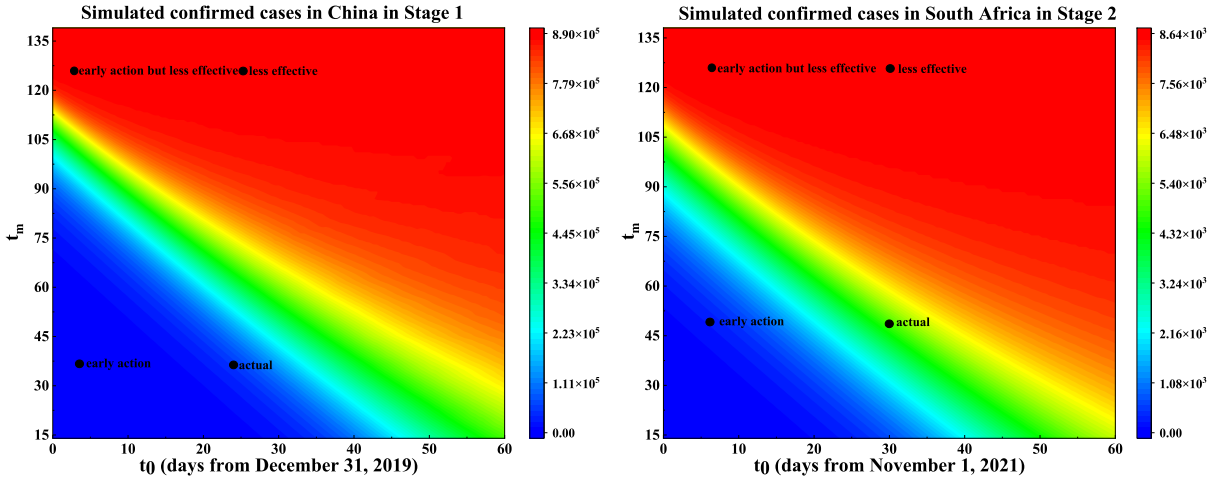


Fig. 15. The phase diagrams of the simulated number of confirmed cases under different interventions with starting dates t_0 and duration t_m, annotated to highlight four distinct intervention scenarios. The first sub-figure represents the simulated results of China during Stage 1 while the second delves into the outcomes for South Africa during Stage 2.

spread of the disease in other countries, particularly major aviation hubs like the United States and the United Kingdom. To provide a broader view, we also randomly selected two other countries for analysis.

Fig. 16 illustrates the disease spread trajectory in selected countries following a 30% reduction in air passenger volume on the mentioned O-D routes from the 24th day.[21] A key observation from the data is the latency effect of travel restrictions. While the immediate aftermath of the restrictions does not show a significant change in infection numbers, a noticeable reduction in subsequent cases is evident after a delay of 3–5 days. This suggests that while travel restrictions can be effective, their impact is not instantaneous. Furthermore, the reduction in air passenger volume on a specific O-D route does have repercussions on the number of confirmed cases in other countries. However, this impact is indirect and not as pronounced. The reason for this muted effect can be attributed to the complex and interwoven nature of the global air transport network, where multiple routes and connections can compensate for reduced traffic on a specific route. This conclusion aligns with previous research findings (see Sun et al. (2021b), Yu and Chen (2021) and Liu et al. (2022)), which have emphasized the limitations of implementing flight reductions on specific O-D routes. Therefore, while targeted flight reductions can be a part of the solution, a more holistic approach that considers the broader dynamics of global travel and connectivity is essential for effective pandemic control.

Indeed, the impact of flight reductions on specific O-D routes varies across different countries, reflecting the heterogeneity in each country's epidemic development, control capabilities, and policy measures. For instance, as depicted in Fig. 16 for Stage 2, countries like France and the United Kingdom seem to be relatively unaffected by flight reductions, with their epidemic trajectories remaining largely consistent. In contrast, countries like the United States and the Netherlands exhibit subtle fluctuations in their subsequent curves, suggesting that they might be more sensitive to changes in international air traffic. This disparity underscores the importance of understanding the unique context and dynamics of each country when evaluating the effectiveness of international travel restrictions.

We delve deeper to discern how varying travel restriction scenarios might influence our simulation results. This involves adjusting the duration and the O-D flow reduction ratio to represent different control strengths. Importantly, we maintain the same O-D route from our prior experiment for this analysis. Subsequently, we scrutinize the simulated confirmed cases for another designated country during both the COVID-19 and Omicron phases. Therefore, for these two stages, we focus on South Korea and the United Kingdom, respectively.

As illustrated in Fig. 17, a clear trend emerges: the more significant the flight reduction and the longer its duration, the fewer the infections. Moreover, Fig. 17 underscores that imposing international travel restrictions becomes less influential when the disease has already achieved widespread transmission. This is evident in the case of the United Kingdom during Stage 2, where the disease had already permeated extensively. As a result, the simulated number of infections remains largely consistent, irrespective of the varying degrees of O-D flow reduction. Furthermore, our model's capability to estimate the variation in infection numbers under different flight reduction policies provides a foundation to analyze the associated costs. This can pave the way for a cost-effective analysis, enabling policymakers to devise optimal flight management strategies.

[21] Refer to the actual date when travel restrictions started to be implemented.

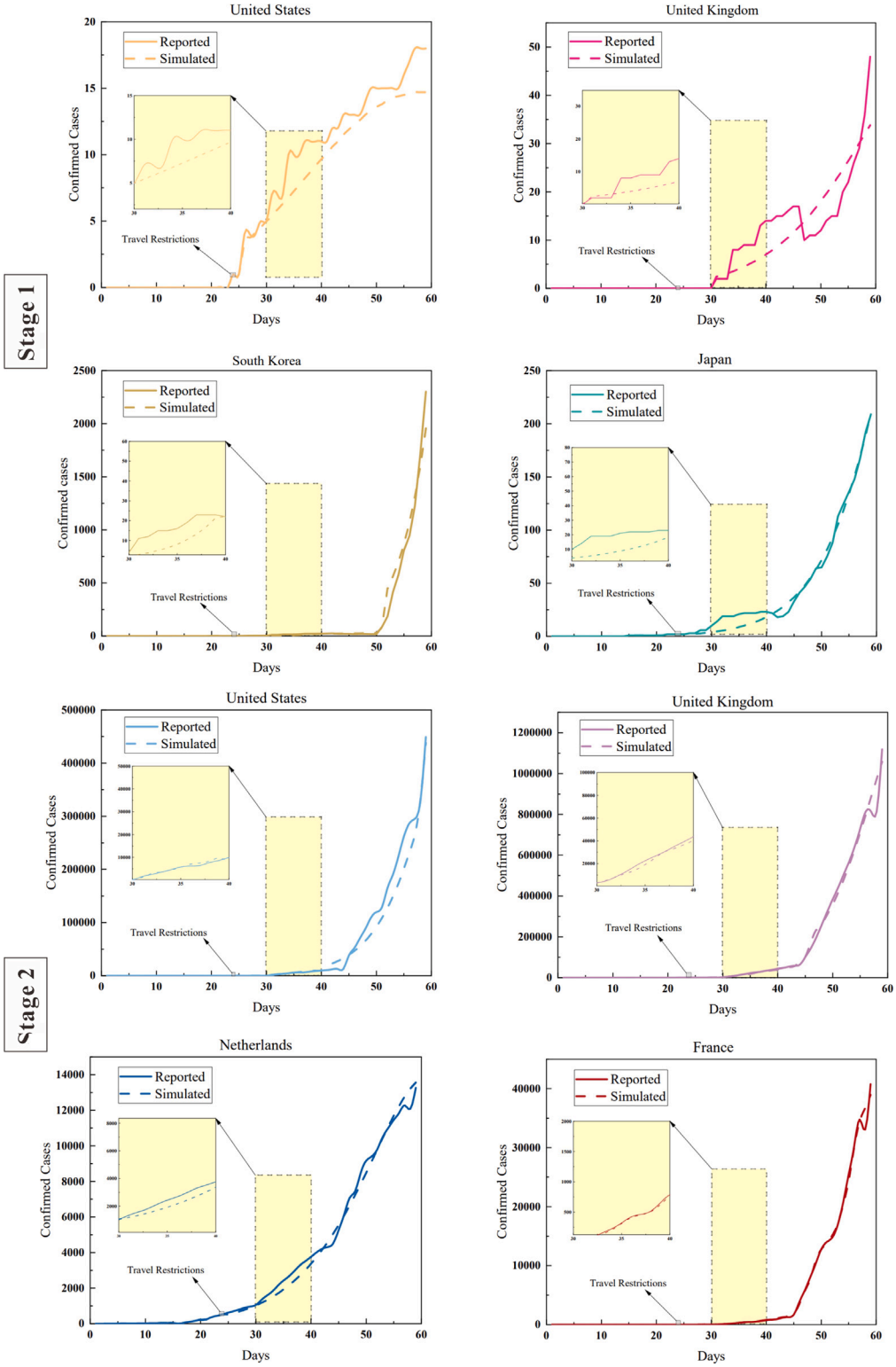


Fig. 16. Epidemic predictions for selected countries during the COVID-19 and Omicron phases, considering travel restrictions on a specific O-D route. The first four sub-figures depict the simulated outcomes for representative countries during Stage 1. In contrast, the subsequent four sub-figures illustrate the simulated outcomes for representative countries during Stage 2.

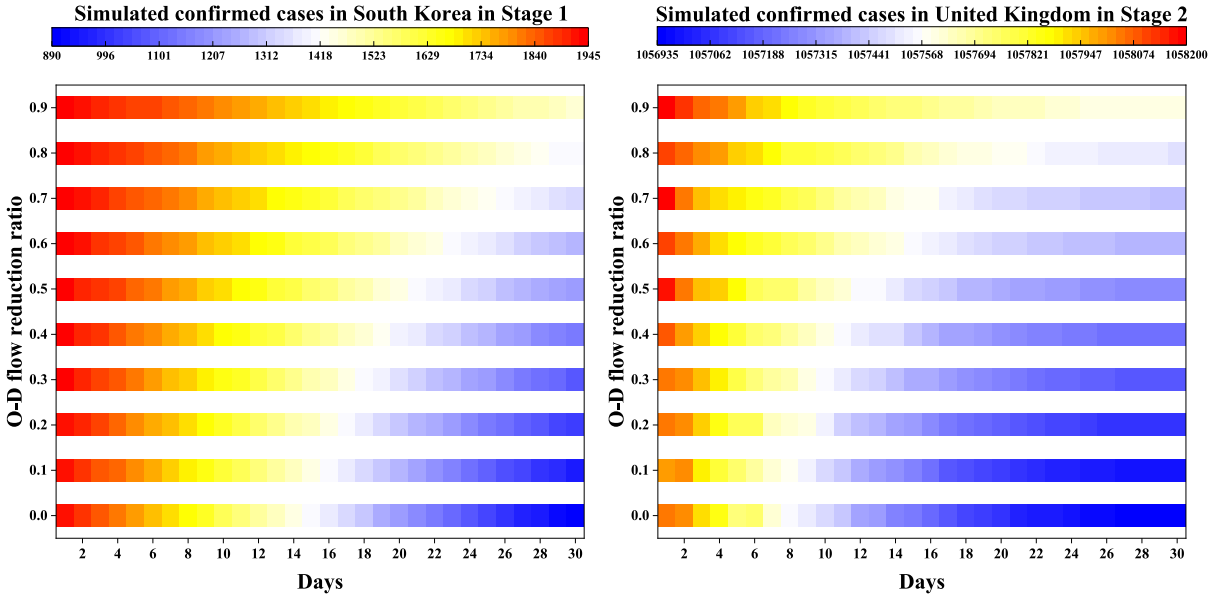


Fig. 17. The illustration of the simulated number of confirmed cases under a spectrum of air travel restriction scenarios, which are characterized by different combinations of reduction ratios and implementation duration. The first sub-figure represents the results for South Korea during Stage 1, while the second delves into the outcomes for the United Kingdom during Stage 2.

5.4. Experimental summary and discussion

In our spatial impact analysis, we have compared the relationship between four effective distance metrics and the first arrival times of COVID-19 and Omicron in different countries/regions. The results indicate that the shortest path effective distance metric performed optimally in both scenarios. This conclusion contradicts the findings of Iannelli et al. (2017), which suggested that diseases might spread more likely in a random walk manner, meaning they would probabilistically choose any possible path for transmission. In their study, they highlighted that the distinction between the random walk effective distance and the multiple path effective distance lay in whether nodes were revisited. Although both metrics consider all possible transmission paths, the former allows for nodes to be revisited, while the latter does not. The dissemination of COVID-19 and the subsequent implementation of travel bans and restrictions have significantly reshaped the aviation network's configuration (Sun et al., 2020). Furthermore, several nations have enforced stringent quarantine and isolation measures. As a result, the concept of recurrent visits to specific nodes may not accurately depict the prevailing circumstances. This observation implies that the random walk effective distance might not wholly capture the intricate transmission patterns of COVID-19.

Furthermore, our use of the shortest path effective distance for the initial arrival time of COVID-19, with an R^2 value of 0.84, yields results that surpass those reported by Adiga et al. (2020) (R^2: 0.79) and Kuo and Chiu (2021) (R^2: 0.39). This superiority might be attributed to the fact that, in contrast to Kuo and Chiu (2021), we incorporated the most up-to-date air passenger data and considered the impact of transfer flights. In comparison to Adiga et al. (2020), we employed the shortest path effective distance, a metric that more closely mirrors the transmission pathways of COVID-19. Choi et al. (2022) also examined the relationship between the basic effective distance and the first arrival time of different variants. While the paper did not provide direct R^2 results, our study distinctively contrasts the fit of multiple effective distance metrics.

In our temporal impact analysis, we have unveiled the pivotal role of air transport networks in the initial stages of the epidemic's spread. However, once the epidemic becomes widespread within a country, subsequent travel restrictions appear to have a diminished impact on the trajectory of the outbreak. Implementing travel restrictions on specific O-D pairs can produce effects, albeit limited. Furthermore, the restrictive effects often manifest with a delay, not producing immediate results. To achieve a more substantial containment of the epidemic's spread, travel restrictions need to be applied across the entire network. The impact of flight reduction policies varies across different stages of the epidemic. This variation is attributed to the evolving nature of the disease over time. Moreover, the influence of such policies is heterogeneous across countries. This disparity arises primarily due to the diverse stages of epidemic development and the differences in policy implementation across nations.

Moreover, compared to previous studies such as Zhang et al. (2020), Chen et al. (2020) and Liu et al. (2022), our model offers a broader scope. We have explored the impacts of the air transport network, constructed from 24 countries or regions, on the spread of COVID-19. Additionally, we delved into the effects of international flight reductions on the progression of the epidemic. In contrast, Zhang et al. (2020), Chen et al. (2020) and Liu et al. (2022) respectively focused solely on the impacts of domestic travel restrictions on the spread of the epidemic within China, the United States, and Singapore. When juxtaposed with Sheng et al. (2021), although the study selected 10 countries for research, it neither validated the accuracy of the model nor conducted

detailed sensitivity and scenario analysis experiments. Furthermore, in comparison with Sun et al. (2021b), that study did not employ time-varying air passenger data. The model in Sun et al. (2021b) mainly examined the accuracy of the disease's propagation time without practical exploration of the evolutionary dynamics. In summary, our model's validation results are robust. As evidenced by the sensitivity analysis experiments and multiple simulation outcomes, our estimates for relevant parameters, such as $U_n(0)$, are reliable.

6. Policy implications

We will summarize the policy implications reflected by our model and experimental part from the following six aspects.

Navigating Effective Distances Amidst Disease Dynamics: The limitations of using geographical distances are evident, as they do not accurately capture the complexities introduced by modern transportation dynamics. With the advent of COVID-19 and associated travel restrictions, the shortest path effective distance has emerged as a more precise predictor of disease spread. For aviation policymakers, this underscores the importance of closely monitoring key air routes, implementing thorough health screenings, enforcing quarantine protocols, and sharing real-time data to mitigate the risk of transmission.

Strategic Planning for Aviation Networks in Pandemic Response: The early spread of COVID-19 highlights the critical influence of aviation network structures. Our findings reveal that this influence wanes during later stages of widespread transmission, particularly under travel restrictions. These observations suggest that early in pandemics, strategic aviation network planning is essential for identifying and managing potential hotspots. Such preemptive measures protect public health while aiming to reduce the impact on global travel and trade, balancing health priorities with economic considerations.

Comprehensive Flight Restriction Strategies with Anticipatory Planning: Travel restrictions often have delayed effects, underscoring the importance of proactive policy planning. While targeting specific O-D pairs may provide short-term relief, the overall effectiveness in slowing the pandemic's spread may be modest. The complex interconnectivity of the aviation network means the expected benefits of such restrictions could fall short. Policymakers need to carefully evaluate the timing, intensity, and duration of these strategies, aiming to preempt potential outbreaks rather than in reaction to them.

Tailored Regional Strategies with Cost-Benefit Considerations: The impact of COVID-19 across aviation markets differs, and our model reflects these variances, accurately capturing developmental trends. It is crucial for policymakers to tailor interventions to their region's unique development stage and traits. Additionally, by simulating the impact of different policies on infection rates and comparing these to the costs of implementation, our model aids in finding an optimal balance between economic and health priorities.

Adaptive Policies for Evolving and Emerging Threats: Different pandemic phases necessitate distinct policies. Our adaptable model, informed by the changing nature of COVID-19, supports the development of both reactive and forward-looking measures. This flexibility arises from incorporating specific viral characteristics, enabling the simulation of spread patterns for new strains, given their medical parameters are known. Such capabilities afford critical insights for early interventions. Consequently, the effectiveness of similar travel restrictions may differ across stages, underscored by the evolving dynamics of the virus.

7. Conclusion

In this section, we first introduce the improvements made to our model, along with the corresponding results and analyses. Based on the findings from our experiments, we summarize the relevant policy implications. Subsequently, we discuss the limitations of our study and outline directions for future work.

7.1. Results and implication summary

In this paper, we develop a spatial and temporal modeling framework that aims to provide insights into the impact of the international air transport network on the spread of the COVID-19 pandemic. Our focus predominantly lies on two pivotal phases of COVID-19: its nascent stage and the subsequent Omicron era.

Adapting the conventional effective distance metric to include transfer flights and more recent air passenger data has enabled more accurate predictions of pandemics' initial onset. Our findings underscore that the effective distance offers a more precise measure than the geographical distance in predicting the spread of the disease. Notably, the shortest path effective distance emerged as the most optimal metric, consistently outperforming others. Moreover, the non-conformity to the random walk effective distance can be traced back to the structural shifts in the network, a direct consequence of the travel restrictions imposed during the early stages of the pandemic.

Transitioning to our network-centric heterogeneous SUCRD mathematical model, we have endeavored to encapsulate the temporal progression of infectious cases by individual countries. This model serves as an effective tool to evaluate the ramifications of supplemental intervention controls or travel curtailments on the trajectory of the disease. Our simulation results from this temporal model are not only robust but also demonstrate that our parameter estimations, such as $U_n(0)$, are reliable. The model's accuracy in capturing the disease's temporal dynamics, coupled with its robustness against parameter variations, underscores its potential as a valuable tool for policymakers.

The rigorous spatio-temporal analysis yields several critical implications for researchers and stakeholders:

- The international air transport network stands as a linchpin in the early dissemination of pandemics, underscoring the need for vigilant oversight of major air hubs.

- The shortest path effective distance emerges superior, advocating for a focus on travel patterns over mere geographical proximity in risk assessment.
- Transfer flights are not mere adjuncts; they are pivotal transmission vectors, necessitating a broader surveillance scope beyond direct flights.
- Early interventions wield paramount influence, emphasizing the essence of timely, decisive actions.
- The diminishing returns of travel restrictions over an epidemic's lifespan call for a judicious evaluation of their prolonged imposition, especially when a disease achieves widespread status.
- The mutable nature of diseases, as exemplified by the contrasting trajectories of COVID-19 and Omicron, necessitates adaptive, strain-specific interventions.
- Flight reduction policies cast a heterogeneous impact, influenced by a country's epidemic stage and its intrinsic control measures, advocating for a tailored rather than templated approach.
- Achieving a balance between health and economic outcomes presents a delicate challenge. As interventions intensify, the diminishing health benefits must be judiciously weighed against escalating economic costs.

7.2. Limitations and future works

Several limitations of this study are structured as follows: (1) There is a need for refinement in data collection to obtain more accurate data. While we utilized load factor data to estimate the number of passengers per flight, this may not fully reflect the actual operational level of flights. Access to real-time passenger data would provide more accurate experimental results for our research. Additionally, considerable effort was expended in acquiring Omicron data, estimating the daily increase in cases across countries. However, there is still a lack of publicly available, granular daily Omicron case data. (2) The simulation efficiency of the model needs enhancement. By incorporating heterogeneous model parameters and time-varying air passenger flow into the previous network-based compartmental model, we significantly increased the model's complexity and simulation difficulty. Although we adopted a new parameter evaluation and simulation framework for our model, its operational efficiency has decreased compared to the previous simulation process within the original model and requires further improvement. (3) The simulation experiments could be expanded to encompass a broader array of datasets and scenarios. While our current investigation assessed the influence of the air transport network comprising 24 countries on pandemic propagation, acquiring additional data would enable the model's application and testing across a wider selection of countries. Moreover, conducting validations for various COVID-19 variants would offer more compelling evidence for the model's applicability and effectiveness in understanding and managing pandemic dynamics. (4) While our study highlights the international air transport network's impact on pandemic transmission, including detailed domestic air travel studies and finer-grained network models could deepen our understanding of intra-country spread. This broader analysis would improve our knowledge of transmission dynamics.

For future research directions, our study suggests several promising avenues. (1) Developing simulation frameworks for multiple-origin outbreaks or more general scenarios: While our model has shown promising results in simulating the initial development curves of COVID-19 and Omicron, attempts at simulating multiple-origin outbreak scenarios necessitate further refinement due to the lack of accurate assessment of model parameters and initial conditions for such stages. Employing machine learning techniques to estimate initial parameters for simulation runs presents a forward-looking approach for more generalized scenarios. (2) Investigating the impact of various modes of transportation on pandemic spread: Although air travel dominates the international spread of pandemics (Sun et al., 2021a), other modes of transportation, such as maritime, also play a significant role that cannot be overlooked. Integrating data from these modes and adjusting the model structure to assess the roles of multi-modal transportation in the pandemic spread is a worthwhile endeavor. (3) Delving deeper into the role of air transport networks: Previous research and our work validate that effective distance metrics related to air passenger data correlate more closely with the initial arrival times of diseases than geographical distance. It is mentioned in Brockmann and Helbing (2013) that under the effective distance calculation framework, complex disease transmission behaviors can be simplified into wave propagation patterns. Exploring whether there exists a more complex, possibly nonlinear relationship between different effective distance metrics and initial arrival times presents an intriguing topic. (4) Studying the interplay between air transport networks and disease spread: Our model offers a feasible approach to quantitatively measure one aspect of this relationship. However, as suggested in Sun et al. (2021d) and Mutascu and Sokic (2023), the interaction between air transport networks and disease spread is mutual. Developing a quantitative model that can measure this bidirectional influence would provide a more comprehensive perspective for policy formulation, highly appreciated by researchers and stakeholders.

CRediT authorship contribution statement

Chi Li: Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Linhao Yu:** Data curation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Jianfeng Mao:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Supervision, Writing – review & editing. **Wei Cong:** Formal analysis, Investigation, Resources, Writing – review & editing. **Zibin Pan:** Data curation, Methodology, Software, Visualization. **Yuhao Du:** Data curation, Visualization, Writing – review & editing. **Lianmin Zhang:** Resources, Supervision.

Data availability

Data will be made available on request.

Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (NSFC) under grants U1733102 and 72394362; in part by the Guangdong Provincial Key Laboratory of Big Data Computing, PR China, The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), PR China under grant 2024SC0003; in part by the Shenzhen Science and Technology Innovation Committee, PR China under the Shenzhen Stability Science Program grant 2024SC0010; and in part by CUHK-Shenzhen, PR China under grant PF.01.000404.

Appendix A. Extended figures and tables

A.1. Illustration of the literature summary about the influence of the COVID-19 pandemic on air transportation

See Table A.6.

A.2. Plots of the correlation between first arrival time for COVID-19 and four distance metrics

See Fig. A.18.

A.3. Plots of the correlation between first arrival time for Omicron and four distance metrics

See Fig. A.19.

Table A.6
Summary of the literature about the influence of the COVID-19 pandemic on air transportation.

Reference	Influence area	Major findings
Sun et al. (2020)	Network structural performance	1. Flight restrictions predominantly targeted long-haul international flights, resulting in more pronounced effects on international routes than domestic; 2. Connectivity patterns varied heterogeneously per country and fluctuated according to the situation of COVID-19.
Sun et al. (2021a)	Network structural performance	1. The impact of COVID-19 was homogeneous across the United States, strongly heterogeneous across Europe, and rather short-lived in China; 2. Domestic air connections recovered much faster than international air connections, especially for the Chinese air transport system.
Sun et al. (2021c)	Network structural performance	1. The growth of COVID-19 infection cases in all countries exhibited a two-phase pattern; 2. Nearly all countries probably reacted too slowly in their decision to reduce flights; 3. Airlines responded differently to international versus domestic flight operations.
Li et al. (2021b)	Network structural performance	1. The rate of change in passenger throughput is closely aligned with the growth rate of confirmed cases; 2. The air transport reaction to the pandemic varied significantly across countries.
Zhou et al. (2021)	Network structural performance	1. Some connections primarily cater to local travel demand and remain robust against such disruptions; 2. The world airport networks are resilient to the disconnection of individual countries but are vulnerable when countries essential for international transfers are simultaneously disconnected.
Truong (2021)	Passenger/flow demand	Merely improving the pandemic situation will not guarantee an increase in air travel; economic factors must also be considered.
Kim and Sohn (2022)	Passenger/flow demand	Air passenger behavior is influenced by both internal and external factors, with travel demand increasing as global COVID cases decline and vaccine distribution progresses.
Kuo et al. (2022)	Network structural performance	1. The popularity rankings of most airports remained relatively stable during the outbreak; 2. Centrality measures for most hub airports saw significant declines due to government-imposed travel restrictions.
Li et al. (2022)	Network structural performance	The early control measures implemented had a positive impact on the subsequent recovery of the aviation industry and other sectors.
Wang et al. (2023)	Passenger/flow demand	Given the outbreak's unpredictability, future provincial outbreaks are probable, posing additional challenges to the air transportation system.

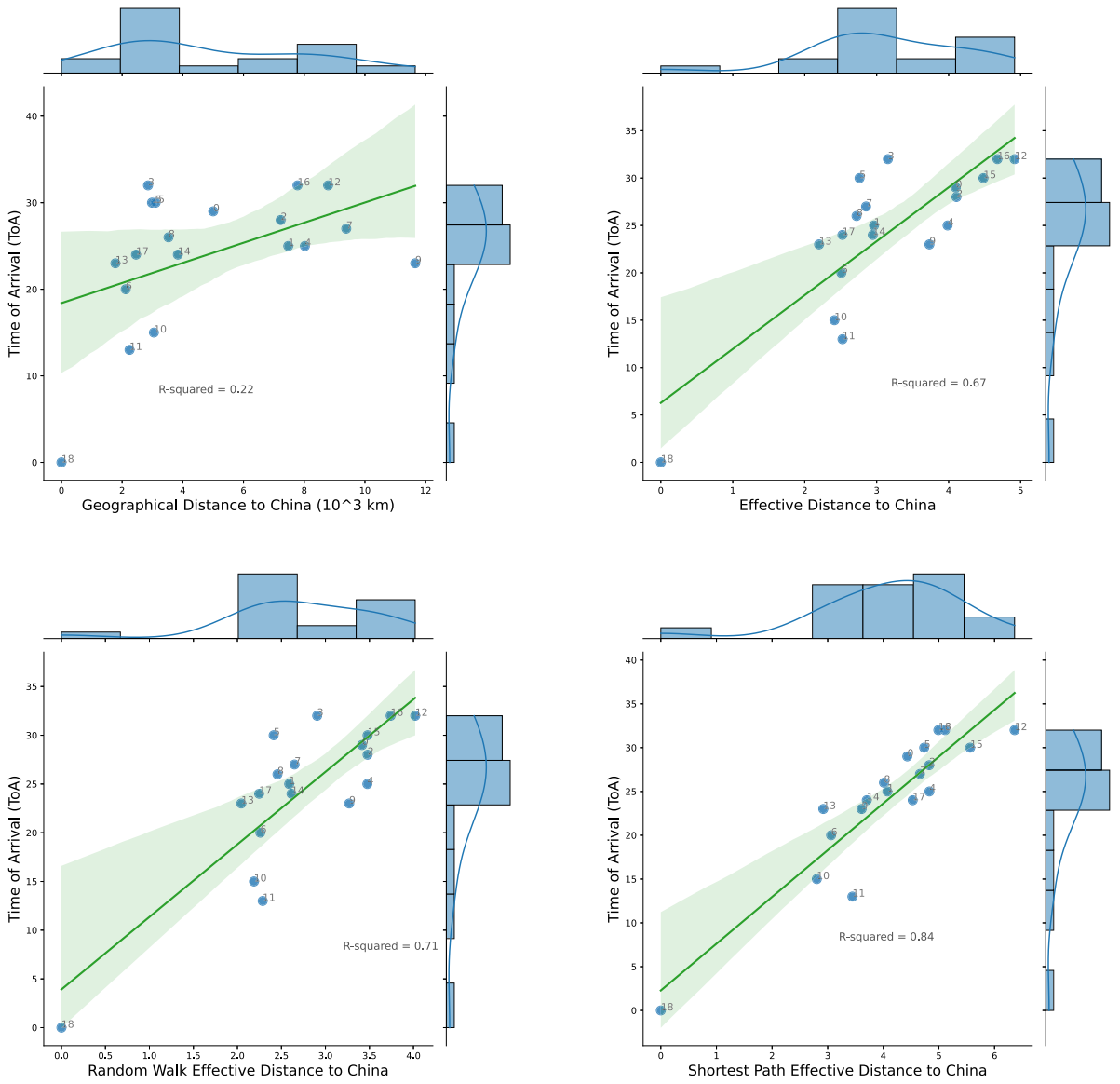


Fig. A.18. A combination of plots delineating the correlation between Time of Arrival (ToA) for COVID-19 and four distinct distance metrics: geographical, effective, random walk, and shortest path distance metrics. Each data point corresponds to a nation or region, identified by an index number. The side histograms depict the respective distributions of ToA and the distance metrics, illustrating the nuances of travel interconnections in the context of COVID-19's dissemination. Country index mapping: 0: UAE, 1: Australia, 2: Germany, 3: Russia, 4: France, 5: Philippines, 6: South Korea, 7: Canada, 8: Malaysia, 9: USA, 10: Japan, 11: Thailand, 12: Spain, 13: Hong Kong (China), 14: Singapore, 15: India, 16: UK, 17: Vietnam, 18: China.

A.4. Illustration of parameter sensitivity analysis

See Tables A.7–A.9.

Appendix B. Identification of transfer flights

In the context of air travel, a transfer flight refers to a journey where passengers change planes at a transit airport to reach their final destination (Chen and Li, 2019). The flight chains or flight itineraries describe the sequence of airports that an aircraft travels through Wu and Law (2019). Given the significant indirect impact of transfer flights on the spread of COVID-19, it is imperative to first identify the flight chains or flight itineraries to determine the transfer flights.

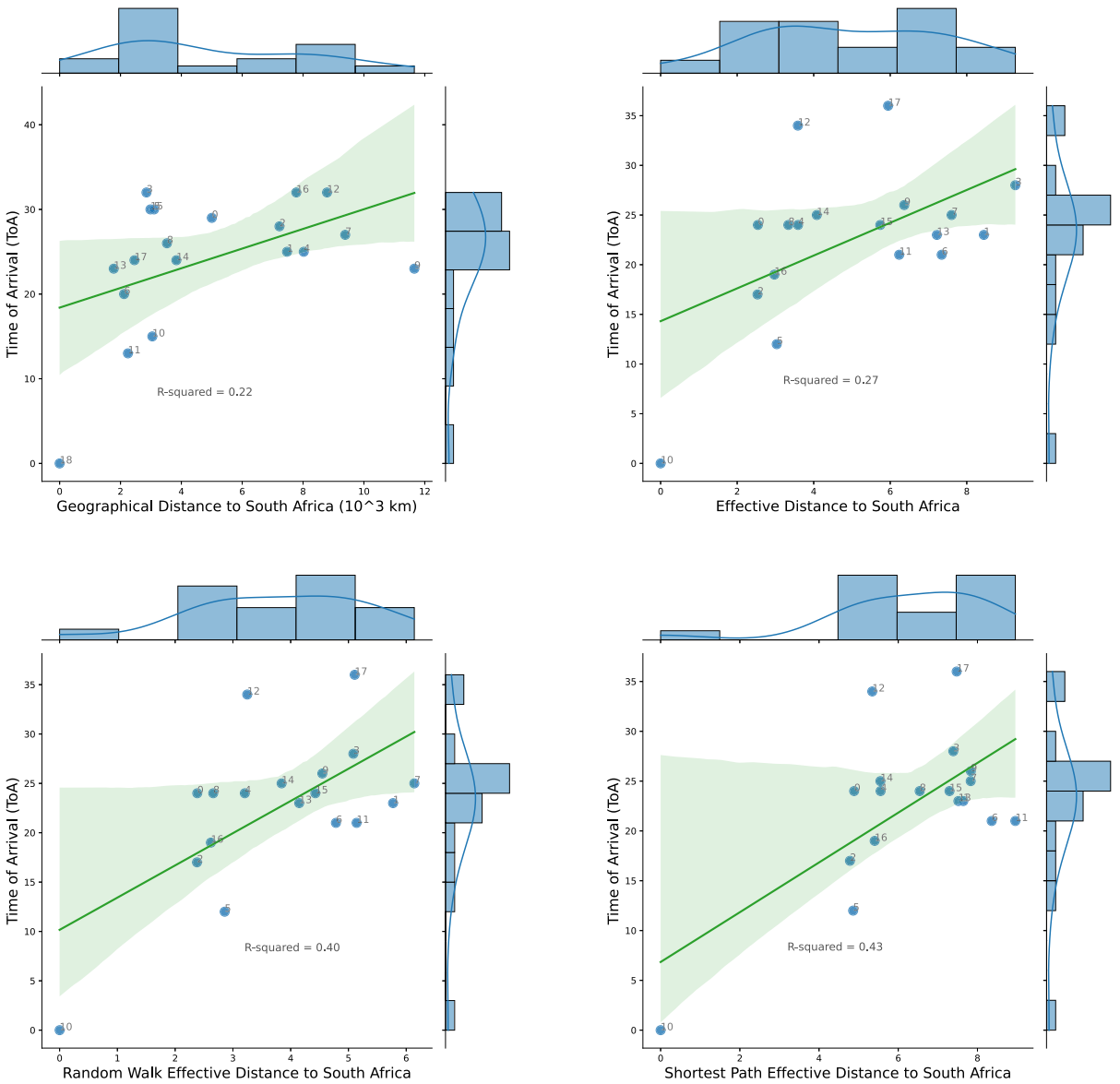


Fig. A.19. Comparative plots showcasing the relationship between the Time of Arrival (ToA) of the Omicron variant and four distance metrics: geographical, effective, random walk, and shortest path distance metrics. Each point represents a country, annotated with index numbers. Marginal histograms outline the distributions for each metric and ToA. Index-to-country mapping: 0: UAE, 1: Brazil, 2: Germany, 3: Russia, 4: France, 5: Netherlands, 6: Canada, 7: Malaysia, 8: USA, 9: Mexico, 10: South Africa, 11: Japan, 12: Turkey, 13: Spain, 14: Singapore, 15: India, 16: UK, 17: China.

Table A.7
Sensitivity analysis for D_I.

D_I	5	9.2 (In this paper)	15
R_0	2.35	2.33	2.32
α_{China}	0.85	0.93	0.89
$U_{China}(0)$	456	414	366
γ	0.95	0.90	0.92

Referring to Cai et al. (2021), a flight itinerary is defined as a sequence of flights that an aircraft undertakes from its origin to its final destination. The identification of a specific aircraft executing a particular flight itinerary can be discerned through its unique registration number. This method of determining aircraft routings via the same registration number is also emphasized in Kafle and Zou (2016) and Wu and Law (2019). Additionally, these sources highlight that such flight transfers occur within the same calendar

Table A.8
Sensitivity analysis for D_C.

D_I	4	8.4 (In this paper)	12
R_0	2.25	2.33	2.35
α_{China}	0.86	0.93	0.96
$U_{China}(0)$	329	414	488
γ	0.88	0.90	0.95

Table A.9
Sensitivity analysis for w.

w	0.01%	0.042% (In this paper)	0.12%
R_0	2.32	2.33	2.33
α_{China}	0.83	0.93	0.87
$U_{China}(0)$	353	414	389
γ	0.93	0.90	0.96

Table B.10
An example of identified transfer flight chains.

Reg number	Departure airport	Arrival airport	Departure time	Arrival time
9MAGH	WMKK	WBGR	2020-02-01 06:55	2020-02-01 09:20
9MAGH	WBGR	WMKK	2020-02-01 09:45	2020-02-01 12:00
9MAGH	WMKK	VTSP	2020-02-01 12:25	2020-02-01 14:00
9MAGH	VTSP	WMKK	2020-02-01 14:30	2020-02-01 16:00
9MAGH	WMKK	WBGS	2020-02-01 16:30	2020-02-01 18:35
9MAGH	WBGS	WMKK	2020-02-01 19:00	2020-02-01 20:55
9MAGH	WMKK	VOMM	2020-02-01 21:30	2020-02-01 01:25

day. Based on the methodologies proposed in the aforementioned literature, we have formulated the following rules for identifying flight chains in this study:

- All flights within the chain are scheduled to depart on the same calendar day.
- Each flight in the sequence must have an identical registration number.
- The airport at which one flight in the chain arrives should be the departure airport for the succeeding flight in the sequence. Furthermore, the departure time of the subsequent flight should be after the arrival time of the preceding flight.
- The arrival time of a flight in the sequence must precede the departure time of the next flight, ensuring a minimum time differential of 15 min. This means that the minimum turnaround time for flights at a specific airport is set at 15 min.

After identifying the flight chains, we further discern the transfer flights. For instance, consider a flight chain $A \rightarrow B \rightarrow C$. Here, the aircraft starts at airport A, stops at airport B, and finally reaches airport C. In this scenario, if a passenger boards the plane at airport A and disembarks at airport C, they have taken a transfer flight via airport B. However, if the flight chain is $A \rightarrow B \rightarrow A$, it means the aircraft starts at airport A, goes to airport B, and then returns to airport A. In this case, there are no transfer flights because the aircraft returns to its starting point. Table B.10 provides a clear example of how to identify transfer flights from a typical flight chain:

- The second and third records show a flight chain of $WBGR \rightarrow WMKK \rightarrow VTSP$. This indicates a transfer flight from $WBGR$ to $VTSP$ via $WMKK$.
- The first two records show a flight chain of $WMKK \rightarrow WBGR \rightarrow WMKK$. Since the aircraft returns to its starting point, there are no transfer flights in this chain.
- The example also includes transfer flights from $VTSP$ to $WBGS$ and from $WBGS$ to $VOMM$.

By associating each airport with its respective country, we can create a network that represents the transfer flights between different nations. This network can be invaluable in understanding the connectivity and flow of passengers between countries, especially in the context of studying the spread of diseases like COVID-19.

Appendix C. Estimation of w

w is defined as the average ratio between a country's daily total air passenger flux and its population. Thus, we can obtain $w = F/P$, where $F = \sum_{mn} f_{mn}$ denotes the total air passenger flux for all countries, and f_{mn} represents the absolute air passenger volume between country pairs m and n. Besides, $P = \sum_k p_k$ is the total population for all countries where p_k is the population of country k.

We calculate ω in the manner described above. Firstly, we estimate the number of passengers on each aircraft based on the acquired load factor data, as described in Section 3. Subsequently, we accumulate this data over time to calculate the number of air

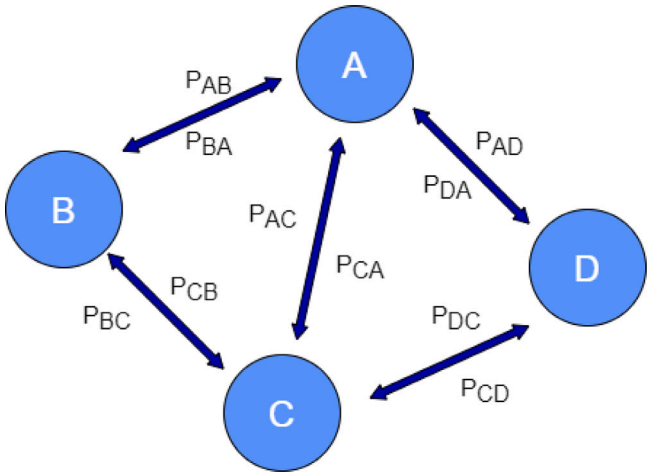


Fig. D.20. A toy example of an air traffic network.

passengers moving through these 105 international airports over a year. Ultimately, we estimate the total number of people moving through these airports in a year to be $F = 742319578$. The total population of these countries/regions is $P = 4842267302$. Thus, we ultimately have $\omega = F/(P * 365) = 742319578/(4842267302 * 365) = 0.042\%$.

Appendix D. Examples for calculating the effective distance

As illustrated in Section 4.1.1, the calculation of effective distance is simple and only requires the knowledge of the flux percentage of each node. In the example of Fig. D.20, with known flux percentages P_{AB}, P_{AC} and P_{AD}, one can easily calculate the effective distance between node A and B, A and C, A and D to be $1 - log(P_{AB})$, $1 - log(P_{AC})$ and $1 - log(P_{AD})$.

However, the computation of random walk effective distance is more challenging. Still using Fig. D.20 as an example, if we want to obtain the probability of A reaching D in a t-step random walk $H_{AD}(t)$, we first need to calculate the $(t-1)$th transition probability matrix. In the beginning, we already have the transition probability matrix:

$$\begin{bmatrix} 0 & P_{AB} & P_{AC} & P_{AD} \\ P_{BA} & 0 & P_{BC} & P_{BD} \\ P_{CA} & P_{CB} & 0 & P_{CD} \\ P_{DA} & P_{DB} & P_{DC} & 0 \end{bmatrix} \tag{D.1}$$

Since we want node D to be visited exactly at the t-th step, we need to remove the 4th column and row and get a reduced matrix H:

$$\begin{bmatrix} 0 & P_{AB} & P_{AC} \\ P_{BA} & 0 & P_{BC} \\ P_{CA} & P_{CB} & 0 \end{bmatrix} \tag{D.2}$$

Taking the $(t-1)$th power of matrix H, we will obtain the $(t-1)$th transition probability among A, B and C. To get the desired probability, we just need to time H^{t-1} with $\left[P_{AD}, P_{BD}, P_{CD}\right]^T$, and the result will be the probability of node A,B,C visiting node D in a t-step random walk. After that,

$$D_{mn}^{RW} = -\log\left(\sum_{t=1}^{\infty} e^{-t} H_{mn}(t)\right) \tag{D.3}$$

can be used to compute the random walk effective distance.

Appendix E. Data curation for Omicron data

Compared to COVID-19, the accurate data of daily confirmed cases for Omicron is limited. On the website of Our World In Data, the share of the Omicron variant in all analyzed sequences is presented every two weeks while we also have access to weekly updated Omicron confirmed cases in countries from the GISAID Initiative. Firstly, we use two weeks as a period and the Omicron data is recovered by the daily confirmed COVID-19 cases times the share of the Omicron variant in that period. For example, the share of Omicron in the United States from 2021.12.13 to 2021.12.27 is 59.17%. In that period (2021.12.13 ~ 2021.12.27), we find out the daily confirmed cases in the United States and time it with 59.17%. The resulting number will be the Omicron confirmed cases on that specific day. We also apply the GISAID data for double-checking to ensure that we obtain relatively accurate reported confirmed cases for Omicron. After obtaining the confirmed cases, we estimate the number of Omicron recovered and death cases by multiplying the Omicron confirmed cases by the COVID-19 recovery rate and death rate in that period.

Appendix F. Supplement and correction of the first case for COVID-19 and Omicron

The COVID-19 data is obtained from the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University. However, the datasets of Johns Hopkins University only contain data starting from 2020.01.22. For some countries close to China, the first case may arrive before 2020.1.22. Therefore, we correct the first case for COVID-19 by checking the news on the website of the health department in the corresponding country. Expressly, in China, the origin of the COVID-19 outbreak is set to be 2019.12.31. In Thailand, Japan, and South Korea, the first case of COVID-19 is 2020.01.13, 2020.01.15, and 2020.01.20, respectively. Since the Omicron data is obtained by the COVID-19 data timing a weekly share, the first case of Omicron in each country can have an error in 0∼ 7 days. Thus, the first case of Omicron is also revised by checking the news in each country.

Appendix G. Calculation of D_I and D_C

We apply the method utilized in Zhang et al. (2020) to calculate these two parameters. As mentioned in Section 4.3, D_I represents the number of days from the moment a case is infectious until it is confirmed, recovered, or dead. In order to estimate D_I, we choose cases that visited Wuhan prior to exhibiting symptoms and assume they were infected on their first day in Wuhan. We also assume that the time between infection and infectiousness can be disregarded. The case data shows that the average time from infection to the beginning of symptoms is $T_s = 4.0$. Also, $T_h = 5.9$ is the time from infection to hospital admission and the period between infection and confirmation is $T_L = 8.4$. In addition, we estimate that the interval between the onset of symptoms and recovery or death is $T_{s,\text{ end}} = 14$ days. Moreover, the time between hospital admission and death or recovery is $T_{h,\text{end}} = 10.2$ days. Besides, the time duration from being confirmed to be recovered or dead is $D_C = T_{\text{end}} - D_I$, where $T_{\text{end}} = \left[\left(T_s + T_{s,\text{ end}}\right) + \left(T_h + T_{h,\text{end}}\right)\right]/2 = 17.5$ days. Thus, we have $D_C = 9.2$ days.

We also employ the method for Omicron case data (Lewnard et al., 2022). On average, the duration from infected to onset symptoms $T_s = 3.0$ while the duration from infected to admission hospital is $T_h = 4.5$. Similarly, we have $D_I = 4.0$, $T_{s,\text{ end}} = 7.0$ and $T_{h,\text{end}} = 5.0$ days. The time duration from being confirmed to be recovered or dead is $D_C = T_{\text{end}} - D_I$, where $T_{\text{end}} = \left[\left(T_s + T_{s,\text{ end}}\right) + \left(T_h + T_{h,\text{end}}\right)\right]/2 = 9.8$ days. Thus, we get $D_C = 5.8$ days.

Appendix H. Derivation and estimation of $U_n(0)$

In our model, there are four free parameters that need to be estimated. These are:

- R_0 - the basic reproduction number.
- $U_n(0)$ - the initial unreported infectious cases in country n.
- α_n - the probability of confirmation for the infectious cases to become confirmed or recovered in country n.
- γ - the ratio of air transportation in total travel modes.

To estimate these parameters, we fit the time series data of reported confirmed cases for the selected countries/regions. Specifically, we estimate these combined parameters by solving the following optimization problem:

$$\min_{R_0, U_{n(0)}, \alpha_n, \gamma} L = \sum_n \sum_t \left[\left(c_n(t) N_n\right) - \left(C_n^*(t)\right)\right]^2 \tag{H.1}$$

s.t.

$$\begin{aligned}
\frac{\partial s_n}{\partial t} &= -\frac{R_0}{D_I}\xi_n s_n u_n + \frac{\omega}{\gamma}\sum_{m\neq n} P_{mn}^t \left(s_m - s_n\right) \\
\frac{\partial u_n}{\partial t} &= \frac{R_0}{D_I}\xi_n s_n u_n - \frac{1}{D_I}u_n + \frac{\omega}{\gamma}\sum_{m\neq n} P_{mn}^t \left(u_m - u_n\right) \\
\frac{\partial c_n}{\partial t} &= \frac{\alpha_n}{D_I}u_n - \frac{1}{D_C}c_n \\
\frac{\partial r_n}{\partial t} &= \frac{(1-\beta)\left(1-\alpha_n\right)}{D_I}u_n + \frac{(1-\beta)}{D_C}c_n + \frac{\omega}{\gamma}\sum_{m\neq n} P_{mn}^t \left(r_m - r_n\right) \\
\frac{\partial d_n}{\partial t} &= \frac{\beta\left(1-\alpha_n\right)}{D_I}u_n + \frac{\beta}{D_C}c_n
\end{aligned} \tag{H.2}$$

The objective is to minimize the sum of squared differences between the actual and estimated confirmed cases over all countries and days, subject to the constraints imposed by the aforementioned system of differential equations. We use the initial period of COVID-19 as an example to explain how these parameters were derived. In this scenario, the initial conditions are:

- $u_n(0) = U_{China}(0)/N_{China}$ for n = China, and 0 for other countries/regions. The initial unreported infectious cases need to be estimated.
- $c_n(0) = C_{China}(0)/N_{China}$ for n = China, and 0 for other countries/regions. The initial confirmed cases refer to Zhang et al. (2020).
- $s_n(0) = 1 - i_n(0) - c_n(0)$ for n = China, and 1 for other countries/regions.

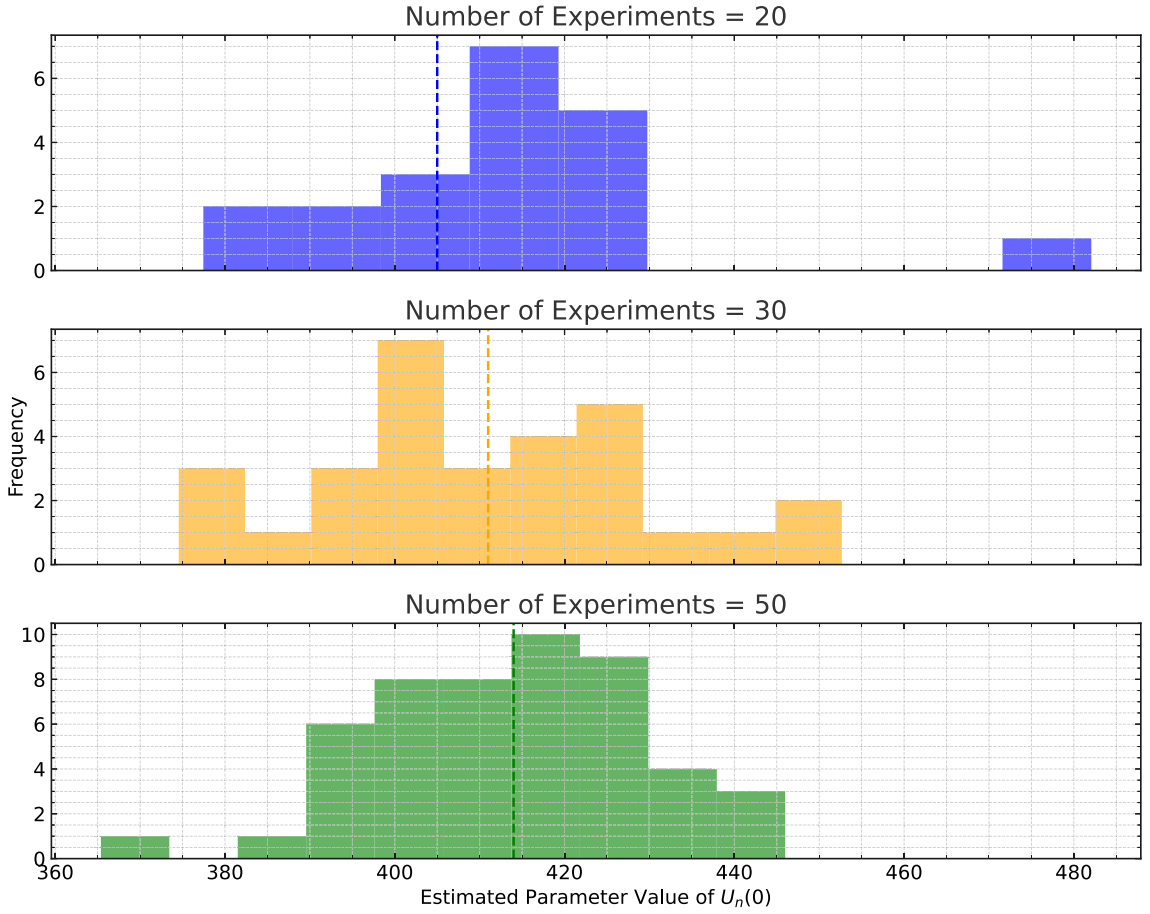


Fig. H.21. The distributions of estimated parameter $U_n(0)$ with 20, 30 and 50 experiments. The dashed lines show the medians of the distributions.

- $d_n(0) = r_n(0)$ for any countries/regions.

Solving this optimization problem is not straightforward as it deviates from traditional optimization paradigms. We employ the Differential Evolution Algorithm[22] to output the optimized parameter combinations, which are then input into the ODE solver.[23] This solver computes the estimated number of infections for different countries/regions at various time points. By comparing the cumulative error between these estimates and the actual infection numbers for these times and regions, we iteratively refine and optimize the parameters output by the Differential Evolution algorithm. Ultimately, the optimal parameter combination, corresponding to the minimal objective function value, is produced after a specified maximum number of iterations. To enhance search efficiency, we also incorporate parallel evaluation and multiple population evolution training techniques.

To validate the reliability of our parameter estimation method, we conducted 50 parallel searches and recorded the optimal parameter outputs from each simulation experiment, as well as the median values of the overall results in Table 4. Fig. H.21 presents histograms of estimated parameter values of $U_n(0)$ across three experiment counts: 20, 30, and 50. As the number of experiments increases, the distributions become more centered and narrower, indicating enhanced precision in parameter estimation. Notably, the overall distributions exhibit minimal shifts, underscoring the stability and consistency of our parameter estimates.

References

Adiga, A., Venkatramanan, S., Schlitt, J., Peddireddy, A., Dickerman, A., Bura, A., Warren, A., Klahn, B.D., Mao, C., Xie, D., et al., 2020. Evaluating the impact of international airline suspensions on the early global spread of COVID-19. MedRxiv.

Arnaout, R., Arnaout, R., 2022. Visualizing omicron: COVID-19 deaths vs. cases over time. PLoS One 17 (4), e0265233.

Ciofi degli Atti, M.L., Merler, S., Rizzo, C., Ajelli, M., Massari, M., Manfredi, P., Furlanello, C., Scalia Tomba, G., Iannelli, M., 2008. Mitigation measures for pandemic influenza in Italy: an individual based model considering different scenarios. PLoS One 3 (3), e1790.

[22] https://github.com/geatpy-dev/geatpy.

[23] https://github.com/scipy/scipy.

Bao, X., Ji, P., Lin, W., Perc, M., Kurths, J., 2021. The impact of COVID-19 on the worldwide air transportation network. R. Soc. Open Sci. 8 (11), 210682.
Brockmann, D., Helbing, D., 2013. The hidden geometry of complex, network-driven contagion phenomena. Science 342 (6164), 1337–1342.
Cai, Q., Alam, S., Duong, V.N., 2021. A spatial–temporal network perspective for the propagation dynamics of air traffic delays. Engineering 7 (4), 452–464.
Chen, J., Li, M., 2019. Chained predictions of flight delay using machine learning. In: AIAA Scitech 2019 Forum. p. 1661.
Chen, S., Li, Q., Gao, S., Kang, Y., Shi, X., 2020. State-specific projection of COVID-19 infection in the United States and evaluation of three major control measures. Sci. Rep. 10 (1), 1–9.
Chinazzi, M., Davis, J.T., Ajelli, M., Gioannini, C., Litvinova, M., Merler, S., Pastore y Piontti, A., Mu, K., Rossi, L., Sun, K., et al., 2020. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science 368 (6489), 395–400.
Choi, Y., Zou, L., Dresner, M., 2022. The effects of air transport mobility and global connectivity on viral transmission: Lessons learned from Covid-19 and its variants. Transp. Policy 127, 22–30.
Coelho, M.T.P., Rodrigues, J.F.M., Medina, A.M., Scalco, P., Terribile, L.C., Vilela, B., Diniz-Filho, J.A.F., Dobrovolski, R., 2020a. Exponential phase of COVID-19 expansion is driven by airport connections. MedRxiv.
Coelho, M.T.P., Rodrigues, J.F.M., Medina, A.M., Scalco, P., Terribile, L.C., Vilela, B., Diniz-Filho, J.A.F., Dobrovolski, R., 2020b. Global expansion of COVID-19 pandemic is driven by population size and airport connections. PeerJ 8, e9708.
Colizza, V., Barrat, A., Barthélemy, M., Vespignani, A., 2006. The role of the airline transportation network in the prediction and predictability of global epidemics. Proc. Natl. Acad. Sci. 103 (7), 2015–2020.
Dai, L., Tereshchenko, I., Hansen, M., 2021. Quantifying the impact of air travel on growth of COVID-19 pandemic in the United States. In: Fourteenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2021). p. 10.
Daon, Y., Thompson, R.N., Obolski, U., 2020. Estimating COVID-19 outbreak risk through air travel. J. Travel Med. 27 (5), taaa093.
Ding, Y., Wandelt, S., Sun, X., 2021. TLQP: Early-stage transportation lock-down and quarantine problem. Transp. Res. C 129, 103218.
Duan, W., Qiu, X., Cao, Z., Zheng, X., Cui, K., 2013. Heterogeneous and stochastic agent-based models for analyzing infectious diseases' super spreaders. IEEE Intell. Syst. 28 (4), 18–25.
Guan, J., Zhao, Y., Wei, Y., Shen, S., You, D., Zhang, R., Lange, T., Chen, F., 2022. Transmission dynamics model and the coronavirus disease 2019 epidemic: applications and challenges. Med. Rev. 2 (1), 89–109.
Guimera, R., Mossa, S., Turtschi, A., Amaral, L.N., 2005. The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. Proc. Natl. Acad. Sci. 102 (22), 7794–7799.
Hu, S., Xiong, C., Yang, M., Younes, H., Luo, W., Zhang, L., 2021. A big-data driven approach to analyzing and modeling human mobility trend under non-pharmaceutical interventions during COVID-19 pandemic. Transp. Res. C 124, 102955.
Iannelli, F., Koher, A., Brockmann, D., Hövel, P., Sokolov, I.M., 2017. Effective distances for epidemics spreading on complex networks. Phys. Rev. E 95 (1), 012313.
Ivanov, D., 2020. Predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case. Transp. Res. E 136, 101922.
Kafle, N., Zou, B., 2016. Modeling flight delay propagation: A new analytical-econometric approach. Transp. Res. B 93, 520–542.
Kim, M., Sohn, J., 2022. Passenger, airline, and policy responses to the COVID-19 crisis: The case of South Korea. J. Air Transp. Manage. 98, 102144.
Kraemer, M.U., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D.M., Open COVID-19 Data Working Group, Du Plessis, L., Faria, N.R., Li, R., et al., 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. Science 368 (6490), 493–497.
Kuo, P.-F., Chiu, C.-S., 2021. Airline transportation and arrival time of international disease spread: A case study of COVID-19. PLoS One 16 (8), e0256398.
Kuo, P.-F., Putra, I.G.B., Setiawan, F.A., Wen, T.-H., Chiu, C.-S., Sulistyah, U.D., 2022. The impact of the COVID-19 pandemic on OD flow and airport networks in the origin country and in Northeast Asia. J. Air Transp. Manage. 100, 102192.
Lau, H., Khosrawipour, V., Kocbach, P., Mikolajczyk, A., Ichii, H., Schubert, J., Bania, J., Khosrawipour, T., 2020a. Internationally lost COVID-19 cases. J. Microbiol. Immunol. Infect. 53 (3), 454–458.
Lau, H., Khosrawipour, V., Kocbach, P., Mikolajczyk, A., Ichii, H., Zacharski, M., Bania, J., Khosrawipour, T., 2020b. The association between international and domestic air traffic and the coronavirus (COVID-19) outbreak. J. Microbiol. Immunol. Infect. 53 (3), 467–472.
Lewnard, J.A., Hong, V.X., Patel, M.M., Kahn, R., Lipsitch, M., Tartof, S.Y., 2022. Clinical outcomes among patients infected with Omicron (B. 1.1. 529) SARS-CoV-2 variant in southern California. MedRxiv.
Li, Z., 2021. Air emergency transport under COVID-19: Impact, measures, and future. J. Adv. Transp. 2021, 5560994.
Li, J., Huang, C., Wang, Z., Yuan, B., Peng, F., 2020. The airline transport regulation and development of public health crisis in megacities of China. J. Transp. Health 19, 100959.
Li, T., Rong, L., Zhang, A., 2021d. Assessing regional risk of COVID-19 infection from Wuhan via high-speed rail. Transp. Policy 106, 226–238.
Li, Y., Wang, J., Huang, J., Chen, Z., 2022. Impact of COVID-19 on domestic air transportation in China. Transp. Policy 122, 95–103.
Li, J., Xiang, T., He, L., 2021a. Modeling epidemic spread in transportation networks: A review. J. Traff. Transp. Eng. (Engl. Ed.) 8 (2), 139–152.
Li, S., Zhou, Y., Kundu, T., Sheu, J.-B., 2021b. Spatiotemporal variation of the worldwide air transportation network induced by COVID-19 pandemic in 2020. Transp. Policy 111, 168–184.
Li, S., Zhou, Y., Kundu, T., Zhang, F., 2021c. Impact of entry restriction policies on international air transport connectivity during COVID-19 pandemic. Transp. Res. E 152, 102411.
Liu, J., Ong, G.P., Pang, V.J., 2022. Modelling effectiveness of COVID-19 pandemic control policies using an Area-based SEIR model with consideration of infection during interzonal travel. Transp. Res. A 161, 25–47.
Lu, J., Lin, A., Jiang, C., Zhang, A., Yang, Z., 2021. Influence of transportation network on transmission heterogeneity of COVID-19 in China. Transp. Res. C 129, 103231.
Luo, Q., Gee, M., Piccoli, B., Work, D., Samaranayake, S., 2022. Managing public transit during a pandemic: The trade-off between safety and mobility. Transp. Res. C 138, 103592.
Mahmud, K.H., Hafsa, B., Ahmed, R., 2021. Role of transport network accessibility in the spread of COVID-19-a case study in Savar Upazila, Bangladesh. Geospat. Health 16 (1), 954.
Meng, X., Guo, M., Gao, Z., Kang, L., 2023. Interaction between travel restriction policies and the spread of COVID-19. Transp. Policy 136, 209–227.
Moore, M., Gelfeld, B., Adeyemi Okunogbe, C.P., 2017. Identifying future disease hot spots: infectious disease vulnerability index. Rand Health Q. 6 (3).
Murano, Y., Ueno, R., Shi, S., Kawashima, T., Tanoue, Y., Tanaka, S., Nomura, S., Shoji, H., Shimizu, T., Nguyen, H., et al., 2021. Impact of domestic travel restrictions on transmission of COVID-19 infection using public transportation network approach. Sci. Rep. 11 (1), 3109.
Mutascu, M., Sokic, A., 2023. Air transportation under COVID-19 pandemic restrictions: A wavelet analysis. Transp. Policy.
Nakamura, H., Managi, S., 2020. Airport risk of importation and exportation of the COVID-19 pandemic. Transp. Policy 96, 40–47.
Nikolaou, P., Dimitriou, L., 2020. Identification of critical airports for controlling global infectious disease outbreaks: Stress-tests focusing in Europe. J. Air Transp. Manage. 85, 101819.
Organization, W.H., et al., 2020. WHO Director-General's opening remarks at the media briefing on COVID-19. 11 March 2020.
Ruan, Y., Luo, Z., Tang, X., Li, G., Wen, H., He, X., Lu, X., Lu, J., Wu, C.-I., 2021. On the founder effect in COVID-19 outbreaks: how many infected travelers may have started them all? Natl. Sci. Rev. 8 (1), nwaa246.

Sheng, H., Wu, L., Wu, T., Peng, B., 2021. Network dynamic model of epidemic transmission introducing a heterogeneous control factor. J. Med. Virol. 93 (12), 6496–6505.
Sokadjo, Y.M., Atchadé, M.N., 2020. The influence of passenger air traffic on the spread of COVID-19 in the world. Transp. Res. Interdiscip. Perspect. 8, 100213.
Stefanoff, P., Mamelund, S.-E., Robinson, M., Netterlid, E., Tuells, J., Bergsaker, M.A.R., Heijbel, H., Yarwood, J., on Standardization of Attitudinal Studies in Europe, V.W.G., et al., 2010. Tracking parental attitudes on vaccination across European countries: the Vaccine Safety, Attitudes, Training and Communication Project (VACSATC). Vaccine 28 (35), 5731–5737.
Su, M., Hu, B., Jiang, Y., Zhang, Z., Li, Z., 2022. Relationship between the Chinese main air transport network and COVID-19 pandemic transmission. Mathematics 10 (13), 2348.
Sun, X., Wandelt, S., Fricke, H., Rosenow, J., 2021a. The impact of COVID-19 on air transportation network in the United States, Europe, and China. Sustainability 13 (17), 9656.
Sun, X., Wandelt, S., Zhang, A., 2020. How did COVID-19 impact air transportation? A first peek through the lens of complex networks. J. Air Transp. Manage. 89, 101928.
Sun, X., Wandelt, S., Zhang, A., 2021b. Delayed reaction towards emerging COVID-19 variants of concern: Does history repeat itself? Transp. Res. A 152, 203–215.
Sun, X., Wandelt, S., Zhang, A., 2021c. On the degree of synchronization between air transport connectivity and COVID-19 cases at worldwide level. Transp. Policy 105, 115–123.
Sun, X., Wandelt, S., Zhang, A., 2022a. Air transportation as a puzzle piece of COVID-19 in Africa? Res. Transp. Bus. Manage. 43, 100780.
Sun, X., Wandelt, S., Zhang, A., 2022b. COVID-19 pandemic and air transportation: Summary of recent research, policy consideration and future research directions. Transp. Res. Interdiscip. Perspect. 16, 100718.
Sun, X., Wandelt, S., Zhang, A., 2022c. Ghostbusters: Hunting abnormal flights in Europe during COVID-19. Transp. Policy 127, 203–217.
Sun, X., Wandelt, S., Zhang, A., 2023a. A data-driven analysis of the aviation recovery from the COVID-19 pandemic. J. Air Transp. Manage. 109, 102401.
Sun, X., Wandelt, S., Zheng, C., Zhang, A., 2021d. COVID-19 pandemic and air transportation: Successfully navigating the paper hurricane. J. Air Transp. Manage. 94, 102062.
Sun, X., Wandelt, S., Zheng, C., Zhang, A., 2022d. On the effectiveness of flight ban strategies in the fight against COVID-19 variants of concern. Available at SSRN 4031490.
Sun, X., Zheng, C., Wandelt, S., Zhang, A., 2023b. Air transportation and COVID-19: A tale with three episodes. Transp. Econ. Manage. 1, 13–21.
Tong, C., Shi, W., Zhang, A., Shi, Z., 2022. Tracking and controlling the spatiotemporal spread of SARS-CoV-2 Omicron variant in South Africa. Travel Med. Infect. Dis. 46, 102252.
Truong, D., 2021. Estimating the impact of COVID-19 on air travel in the medium and long term using neural network and Monte Carlo simulation. J. Air Transp. Manage. 96, 102126.
Wandelt, S., Sun, X., Zhang, A., 2023. On the contagion leakage via incoming flights during China's aviation policies in the fight against COVID-19. J. Air Transp. Manage. 108, 102377.
Wang, J., Zhang, Z., Lu, G., Yu, B., Zhan, C., Cai, J., 2023. Analyzing multiple COVID-19 outbreak impacts: A case study based on Chinese national air passenger flow. Transp. Res. A 169, 103586.
Wells, C.R., Sah, P., Moghadas, S.M., Pandey, A., Shoukat, A., Wang, Y., Wang, Z., Meyers, L.A., Singer, B.H., Galvani, A.P., 2020. Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak. Proc. Natl. Acad. Sci. 117 (13), 7504–7509.
Wu, C.-L., Law, K., 2019. Modelling the delay propagation effects of multiple resource connections in an airline network using a Bayesian network model. Transp. Res. E 122, 62–77.
Xu, Y., Wandelt, S., Sun, X., 2023. A distributionally robust optimization approach for airline integrated recovery under in-flight pandemic transmission risks. Transp. Res. C 152, 104188.
Yang, J., Li, J., Lai, S., Ruktanonchai, C.W., Xing, W., Carioli, A., Wang, P., Ruktanonchai, N.W., Li, R., Floyd, J.R., et al., 2020. Uncovering two phases of early intercontinental COVID-19 transmission dynamics. J. Travel Med. 27 (8), taaa200.
Yu, M., Chen, Z., 2021. The effect of aviation responses to the control of imported COVID-19 cases. J. Air Transp. Manage. 97, 102140.
Zhang, J., Dong, L., Zhang, Y., Chen, X., Yao, G., Han, Z., 2020. Investigating time, strength, and duration of measures in controlling the spread of COVID-19 using a networked meta-population model. Nonlinear Dynam. 101 (3), 1789–1800.
Zhang, J., Hayashi, Y., 2022. Research frontier of COVID-19 and passenger transport: A focus on policymaking. Transp. Policy 119, 78–88.
Zhang, M., Wang, S., Hu, T., Fu, X., Wang, X., Hu, Y., Halloran, B., Li, Z., Cui, Y., Liu, H., et al., 2022. Human mobility and COVID-19 transmission: a systematic review and future directions. Ann. GIS 28 (4), 501–514.
Zhou, Y., Kundu, T., Qin, W., Goh, M., Sheu, J.-B., 2021. Vulnerability of the worldwide air transportation network to global catastrophes such as COVID-19. Transp. Res. E 154, 102469.
Zhu, P., Guo, Y., 2021. The role of high-speed rail and air travel in the spread of COVID-19 in China. Travel Med. Infect. Dis. 42, 102097.