www.nature.com/scientificreports

scientific reports

Check for updates

OPEN

Impact of dual climatic and socioeconomic factors on global trends in infectious disease outbreaks

Zongxi Qu[1,2], Lingqing Zhang[1,2], Yongzhong Sha[1,2], Beidou Zhang[3] & Kequan Zhang[1,2]✉

A recurrent pattern has been observed in global infectious disease outbreaks in recent decades. Nevertheless, quantifying the impact of global change on infectious disease outbreaks remains a challenging endeavor. This study examines the spatiotemporal characteristics of global infectious disease outbreaks and the extent to which global climatic and socioeconomic factors influence them, as well as the relationships between these factors and outbreaks. The results illustrate that the global trend in infectious disease outbreaks is characterized by an upward trend followed by a levelling off, with spatial progression first from the mid-latitudes to the mid-latitudes and high latitudes of North America and South Asia and then clustering in the Central Asian and North American regions. Additionally, socioeconomic factors, such as GDP, population, and human development, contribute more to infectious disease outbreaks than do climatic factors, such as temperature and precipitation. Furthermore, the dual-factor influence of socioeconomic factors and climate change on the trend of infectious disease outbreaks has a complex nonlinear relationship, with an overall increasing trend in infectious disease outbreaks when the HDI increases from 0.19 to 0.55 and from 0.74 to 0.93, when the temperature increases from 1.9 °C to 23 °C and when precipitation increases from 6.75 mm to 800 mm and from 1,600 mm to 2,900 mm per year. The above findings can help countries formulate appropriate prevention and control policies that are adapted to their development contexts and optimize the allocation of public health resources.

Keywords Infectious disease outbreaks, Climatic factors, Socioeconomic factors, Global trends, Nonlinear relationship

The dawn of the 21st century has been marked by an alarming increase in the incidence of infectious disease outbreaks, which have had profound implications for both public health and economic stability worldwide[1,2]. The emergence of the SARS virus in 2003[3], the H1N1 influenza pandemic in 2009[4], the MERS-CoV outbreak in 2012[5], and the Ebola virus disease (EVD) outbreak in West Africa in 2013[6] have each served as a stark reminder of the potential for widespread devastation. The Zika virus disease in 2015[7] and the COVID-19 pandemic in 2020 further underscored the reality that these diseases can swiftly traverse geographical boundaries, impacting regions globally with varying intensities. This notable upswing in the number of disease outbreaks is not a random occurrence but rather a consequence of the transformative changes that have taken place over recent decades. We find ourselves at a unique crossroad in history, defined by the fourth phase of globalization, which is characterized by an unparalleled level of global interconnectivity[8]. The rise in the number of infectious diseases is shaped by a complex array of factors, including but not limited to urban expansion, escalating population numbers, poverty, rapid technological evolution, the ease of air travel, the intricacies of international commerce, social unrest, environmental deforestation, and the pressing reality of climate change[9–13]. The convergence of these elements signifies the onset of a novel epoch in the realm of infectious diseases. Pathogens, whether newly identified, re-emerging, or persistently present, are continually surfacing and propagating at alarming speeds. The forces of global integration, economic and societal shifts, and climatic alterations are all instrumental in the swift spread of these diseases. The collective impact of these phenomena poses a sophisticated and multifaceted challenge to the global public health infrastructure. Thus, the task at hand is to understand and mitigate the

[1]School of Management, Lanzhou University, Lanzhou 730000, China. [2]Research Center for Emergency Management, Lanzhou University, Lanzhou 730000, China. [3]College of Atmospheric Sciences, Lanzhou University, Lanzhou 730000, China. ✉email: zkq@lzu.edu.cn

impacts of these dual forces—climate and socioeconomic changes—on the dynamics of infectious disease outbreaks[8].

Our current understanding of how infectious disease outbreaks are affected by global changes remains limited[14–16]. Two decades ago, Harvell et al. reviewed the potential for the increased incidence of infectious diseases with climate warming[17]. Many prevalent human diseases are linked to climatic fluctuations, as climate change can lead to the emergence of pathogens whose vectors are sensitive to changes in environmental conditions such as temperature, rainfall, and severe weather events[18–21]. The replication, development, and transmission rates of these pathogens are strongly dependent on temperature[22]. Moreover, Mora noted that more than half of known human diseases are exacerbated by climate change. Furthermore, Rory Gibb et al.[23] noted that temperature is the primary factor influencing the distribution of diseases such as dengue fever.

However, some researchers argue that while the extent of the distribution of some infectious diseases changes with global warming, economic and social factors, especially poverty, can likewise affect the spread of infectious diseases[24]. Some studies have shown that outbreaks of emerging zoonotic infectious diseases may be triggered by the complex interplay of changing ecological, epidemiological, and socioeconomic factors. When predicting the risk of EVD outbreaks, multiple factors, including ecological, epidemiological, and socioeconomic variables, are considered. Thus, socioeconomic development plays a significant role in both the spread and control of infectious diseases[25]. Baker reported that in low- and lower-middle-income countries, the burden of infectious diseases remains at a high level, with high mortality and morbidity rates related to HIV infection, tropical diseases, tuberculosis, and malaria[8]. The global trend of emerging infectious diseases (EIDs) is more concentrated in low-latitude developing countries; it has also been reported that population density is highly correlated with the emergence of diseases but not with environmental or ecological variables. Bhutta et al.[26] described the disease burden, distribution, existing interventions, and coverage of poverty-related infectious diseases, including neglected tropical diseases (NTDs), malaria, tuberculosis, and HIV/AIDS, indicating that most infectious diseases in low- and lower-middle-income countries can be prevented or treated with existing drugs or interventions.

Although medical and technological advancements have promoted population health, disparities in medical development between countries and within countries remain prominent, and this imbalance can affect the distribution pattern of infectious diseases[27]. Mahon et al.[28] considered the impact of various global change drivers on the risk of infectious disease outbreaks and reported that among the drivers of global change, habitat loss/change leads to a significant reduction in the number of disease outbreaks, whereas chemical pollution, climate change, the introduction of species, and changes in biodiversity increase disease response or disease-related negative effects. Although it is relatively widely accepted that global changes affect the risk of infectious disease outbreaks, the full extent of this risk is still difficult to quantify because the drivers of global change are numerous and the pathways of influence are complex. There are also few studies that focus on incorporating both climate change and socioeconomic factors into global changes to analyze their nonlinear relationships and impacts on global infectious disease outbreaks. Because resources for infectious disease management are limited, if we do not understand which drivers of global change have the greatest impact on the risk of infectious diseases, then it may be difficult to set targets. Therefore, understanding the relative importance of climatic and socioeconomic factors can help governments and public health departments worldwide prepare for resource allocation and epidemic prevention and control, thereby improving the effective use of limited resources in disease control.

In this study, our goal is to assess the influence of two types of global change factors on the occurrence of infectious disease outbreaks and to determine their relative contributions. Climatic factors include temperature and precipitation, whereas socioeconomic factors include GDP, the Human Development Index (HDI), population, and the unemployment rate. We compile and organize data on infectious disease outbreaks and corresponding global change factors from 189 regions worldwide, spanning a period of nearly four decades. The extent to which these factors contribute to global infectious disease outbreaks is analyzed through a machine learning approach, and the nonlinear relationship between influencing factors and infectious disease outbreaks is investigated via a generalized additive model (GAM).

Results

(1) Temporal and spatial characteristics of global infectious disease outbreaks

We use the GIDEON Global Disease Database to analyze data on infectious disease outbreaks in 193 countries and regions worldwide from 1979 to 2021, focusing on the spatial and temporal changes in global infectious disease outbreaks over the past 40 years and mapping the global distribution of these infectious diseases (Fig. 1). Figure 1(a) illustrates the notable disparities in the incidence of infectious disease outbreaks across countries and regions worldwide. From 1979 to 2021, the majority of infectious disease outbreaks occurred in countries and regions situated in the mid-latitude region of the Northern Hemisphere, with the Eurasian and North American continents representing the most prominent geographic areas. The ten countries with the highest cumulative number of outbreaks during this period were the United States (4,378), the United Kingdom (1,196), India (994), Japan (931), Canada (855), China (834), Australia (787), Spain (716), France (616), and Brazil (588).

To further examine the influence of climate change, we map the spatial distribution of the number of global infectious disease outbreaks every five years between 1979 and 2021 (Fig. 1(b-i)). The data reveal an overall upward trend in the number of global infectious disease outbreaks over the past four decades, which can be broadly categorized into three phases. This shift may be attributed to a combination of factors, including climate change, the increased frequency of weather events in the mid-latitude region of the Northern Hemisphere, socioeconomic development, and global integration[12,29,30]. In the following, the spatiotemporal characterization of the impact of socioeconomic and climate change on infectious disease outbreaks over three phases is analyzed.

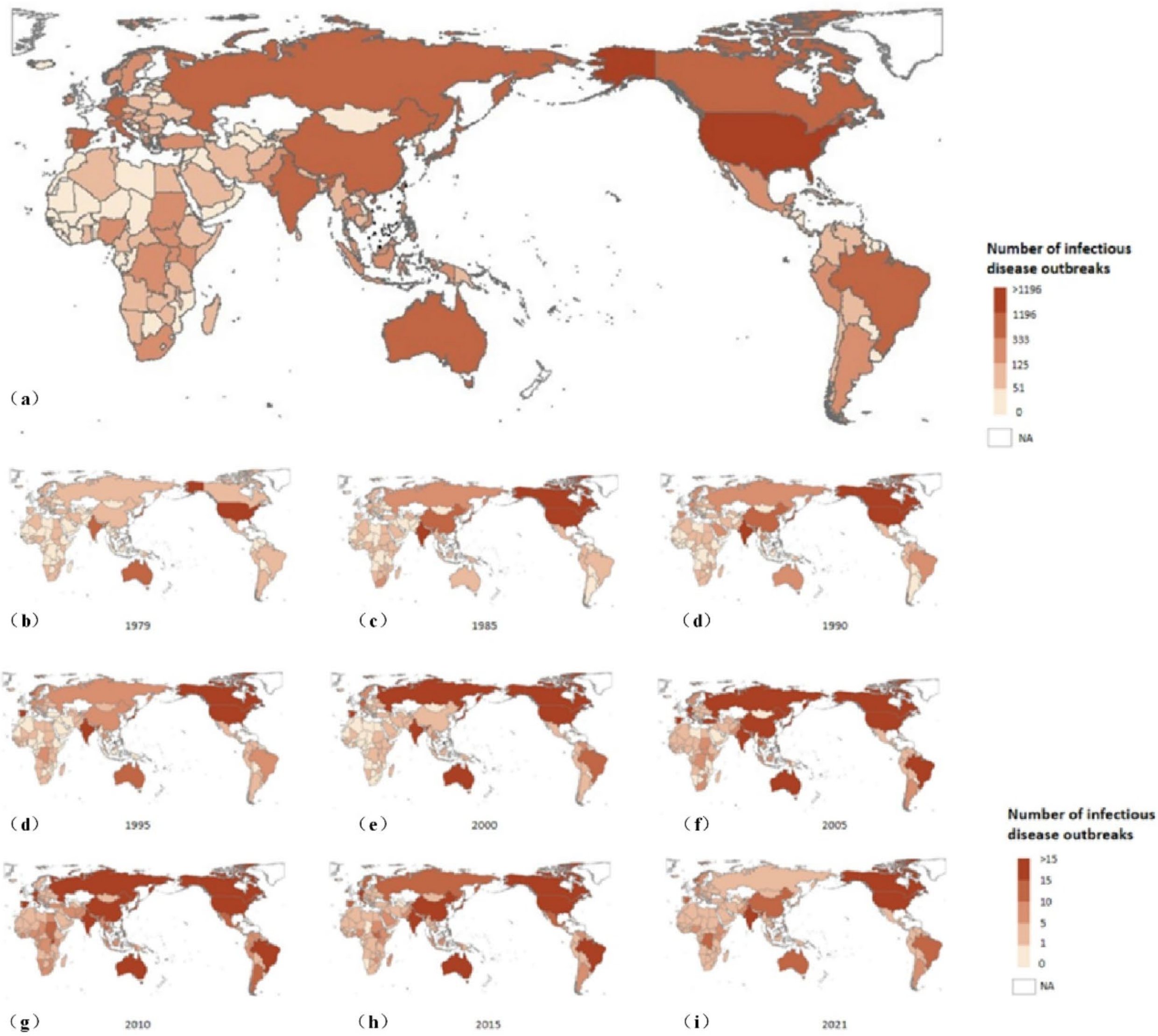

Fig. 1. Spatial distribution of global infectious disease outbreaks (**a**) Global spatial distribution of the cumulative number of infectious disease outbreaks since 1979–2021; (**b-i**) global spatial distribution of the cumulative number of infectious disease outbreaks per five-year period since 1979–2021.

From 1979 to 1990, the majority of infectious disease outbreaks occurred in the mid-latitudes of North America and South Asia, with an average of 30.06 outbreaks per year. In contrast, 94.8% of the other countries and regions (183) experienced fewer than 10 outbreaks, with an average of only 0.67 outbreaks. Notably, the United States, the United Kingdom, Japan, and India experienced a greater number of outbreaks. From an economic and social standpoint, the aforementioned regions are currently undergoing a period of significant restructuring within the global economic landscape. Human activities have intensified in all countries within the region. For example, the United States has become the world's largest economy; European economic integration has facilitated the accelerated growth of the United Kingdom, France, and other European countries; Japan is in the post-World War II period of economic takeoff; and India has achieved rapid development in manufacturing and agriculture[28]. From the perspective of climate change, these regions are situated in the mid-latitudes and are distinguished by complex climates with multiple regional climate zones and weather processes.

Since 2000, there has been a notable surge in the prevalence of global infectious disease outbreaks, with over 95.9% of the world's countries experiencing multiple outbreaks. First, infectious diseases spread to the high latitudes of the Northern Hemisphere and then further to the northern mid-latitudes until 2010, when they reached the Eurasian and American continents. From a socioeconomic perspective, this period coincides with a period of accelerated economic globalization and deepening regional cooperation. Following the conclusion of the Cold War, the acceleration of economic integration in Europe and the rapid growth of emerging market economies, most notably China, led to a significant expansion in global socioeconomic exchanges and cooperation[23]. This situation resulted in a global multipolar situation, which may have also contributed to the accelerated spread of infectious diseases worldwide[31]. From a climate change perspective, the global increase in temperature also exhibited an anomalous trend during this period, and the impact of climate extremes on human society intensified[32]. It can be surmised that the confluence of socioeconomic and climatic factors may have been a significant contributing factor to the global proliferation of infectious diseases during this period.

Between 2010 and 2021, infectious disease outbreaks were concentrated in Central Asia and North America, with a gradual decline at higher latitudes in the Northern Hemisphere. From a socioeconomic perspective, this period represents a phase of recovery and adjustment in the global postfinancial crisis era, with the world economy exhibiting bipolar characteristics. China continues to demonstrate rapid economic growth, functioning as a significant driver of global economic expansion[33]. The United States, conversely, maintains its status as the world's largest economy, exhibiting a high level of economic vitality. Consequently, outbreaks of infectious diseases are concentrated primarily in China, the United States, and neighboring countries[8].

It is evident that global infectious disease outbreaks are influenced by both climate change and socioeconomic factors. However, the precise nature of the relationship between these influences remains unclear. To address this, we employ a machine learning approach to analyze the degree of influence of these factors on global infectious disease outbreaks. Furthermore, we utilize a GAM to investigate the nonlinear relationship between the influencing factors and infectious disease outbreaks.

(2) Importance of socioeconomic and climatic factors for the impact of infectious diseases

Figure 2a shows the importance of socioeconomic and climatic factors for the impact of infectious disease outbreaks. Figure 2a shows that the gross domestic product (GDP), HDI and population density (POP) are the most important factors affecting infectious disease outbreaks in the importance ranking of features, indicating that socioeconomic factors have a greater impact on infectious disease outbreaks than do climatic factors. To determine how these characteristics affect the risk of infectious disease outbreaks, Fig. 2b shows a scatterplot of SHapley Additive exPlanations (SHAP) values for each characteristic for each sample. GDP, HDI, and POP are positively correlated with the number of infectious disease outbreaks, suggesting that the risk of infectious disease outbreaks is greater in countries or regions with a high level of socioeconomic development and high population density than in other countries or regions (Fig. 2). Although greater GDP may mean that more resources are available for public health and healthcare, if these resources are distributed unequally, then it may lead to an increased risk of disease in certain groups[34,35]. In addition, HDI and POP contribute more to infectious disease outbreaks than does GDP, implying that higher levels of urbanization and population density may increase the chances of infectious disease transmission[9,36]. In addition, temperature is negatively correlated with the number of infectious disease outbreaks, suggesting that countries or regions with lower temperatures have a greater risk of infectious disease outbreaks than do those with higher temperatures[20,37]. Climatic factors, such as temperature and rainfall, appear to contribute less to outbreaks of infectious diseases than do socioeconomic factors; however, this does not mean that these climatic factors are unimportant, and many studies have shown that climate change can influence outbreaks of infectious diseases by influencing changes in human activities[22,38].

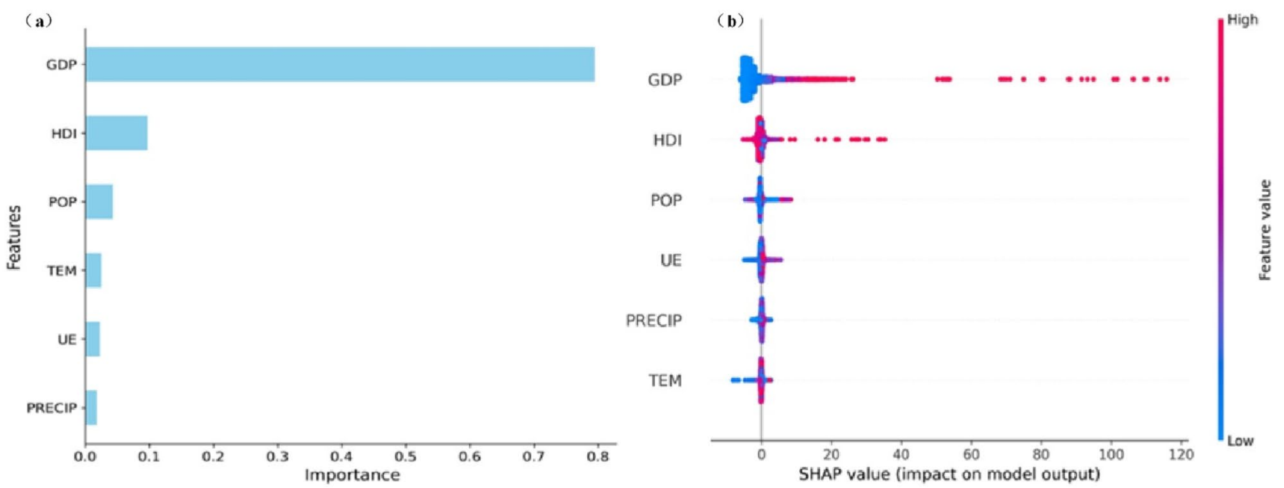

Fig. 2. Overview plot of feature impacts on the basis of SHAP values (**a**) Ranking of feature importance, a global ranking of feature importance obtained by averaging the absolute SHAP values for each feature. (**b**) A graph of the SHAP values for each feature for each sample. The vertical axis shows to which feature the point belongs; the color shows the value of the point under the feature, with red indicating a larger value and blue indicating a smaller value; and the horizontal axis shows whether a feature increases or decreases the predicted value.

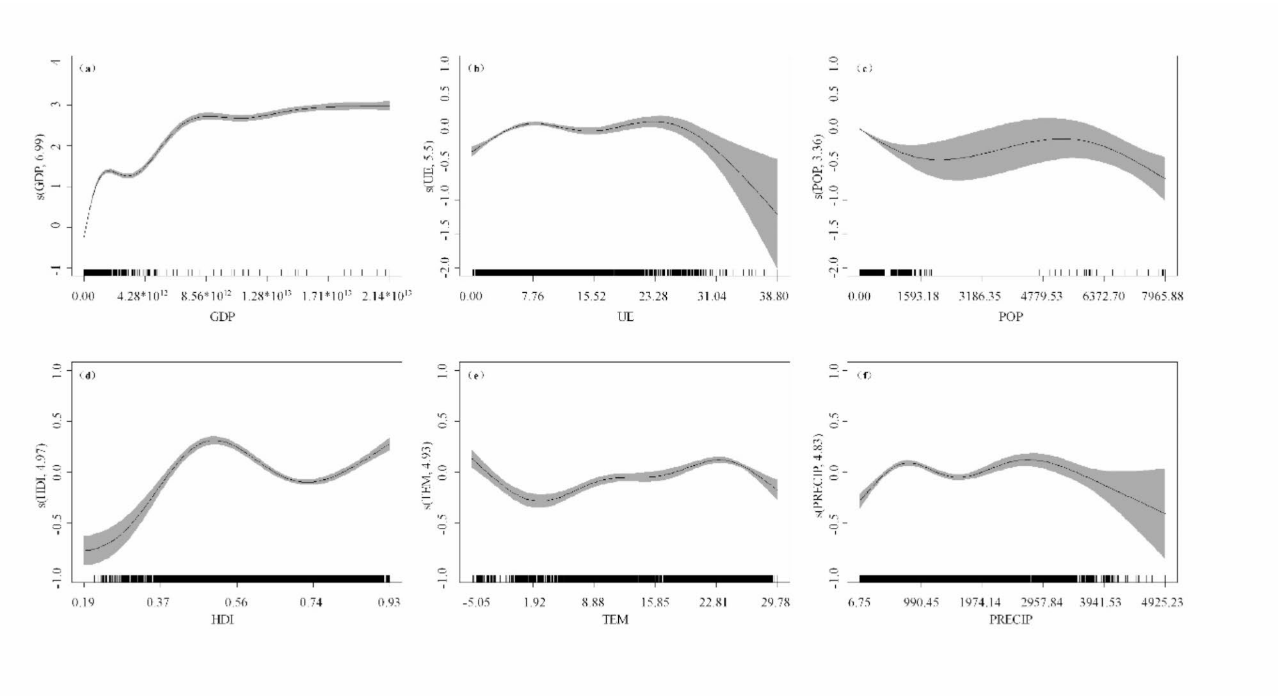

Fig. 3. Response curves of the number of infectious disease outbreaks and influencing factors in the GAM (the dashed interval in the graph represents the confidence interval for the fitted additive functions, whereas the solid line represents the smooth fit of the explanatory variable to the number of outbreaks. The horizontal axis represents the observed values of the dependent variable, the vertical axis represents the smooth fit values of the independent variable to the dependent variable, and the parentheses on the vertical axis denote the dependent variable and its corresponding estimated degrees of freedom).

(3) Relationships among climatic factors, socioeconomic factors and infectious disease outbreaks

We analyze the nonlinear relationships between socioeconomic and climatic factors affecting disease outbreaks via a GAM model, and the results are shown in Fig. 3. The statistics of the fitted parameters in the GAM model are shown in Supplementary Table S2, and all the edf values in the model are greater than 1 and pass the 99% significance test. In the following, we analyze the results in terms of both socioeconomic and climatic factors.

Figure 3(a-d) shows the relationships between infectious disease outbreaks and socioeconomic factors, including GDP, unemployment rate, population density, and the HDI. First, the number of infectious disease outbreaks shows an overall upward trend with increasing GDP; the number of infectious disease outbreaks increases rapidly with GDP when GDP < US\$8*10^{12}, and the number of infectious disease outbreaks is affected by GDP when GDP > US\$8*10^{12}. Peter's concept of "blue marble health", whereby most of the world's poverty-related neglected diseases, including NTDs and the "big three" (i.e., HIV/AIDS, tuberculosis and malaria), are actually most prevalent in G20 economies[27,34,35]. These diseases, specifically, occur primarily among poor people living in impoverished areas that are close to and amidst areas of wealth. These people are sometimes referred to as the "poor among the rich"[39]. These findings suggest that countries with higher GDPs may have a greater risk of contracting infectious diseases. Second, the number of infectious disease outbreaks increases with increasing unemployment when UE is less than 7.7% and increases but maintains a high level of infectious disease outbreaks when UE is 7.7% < UE < 25%. The number of infectious disease outbreaks increases with increasing population density, and the number of infectious disease outbreaks decreases with increasing population density when POP < 1600 people/km^2 or POP > 5100 people/km^2. The increase in population may lead to increased urbanization, which has economic, income, and scale effects; enriches healthcare resources; improves health needs; and creates health awareness. All these channels reduce the risk of infectious disease transmission[36].

HDI is a composite indicator for a comprehensive assessment of the overall level of development of a country, and when 0.55 < HDI < 0.74, which refers to moderate- and high-HDI countries and regions, the number of infectious disease outbreaks decreases with increasing HDI. This finding suggests that in countries and regions with medium and high HDI levels, countries and regions with higher HDI levels usually have stronger public health systems, healthcare resources, and health security, which effectively reduces the incidence of infectious diseases. The number of infectious disease outbreaks increases with increasing HDI when HDI < 0.55 or HDI > 0.74. The number of infectious disease outbreaks in low-HDI areas may be underestimated because of underreporting of infectious disease outbreaks in low-income countries due to the poor quality of their disease registration and reporting systems. Increases in HDI are usually accompanied by improvements in the healthcare system and enhanced reporting mechanisms, which means that outbreaks are being monitored and reported with improved capacity. As HDI increases, these diseases are more accurately recorded and reported, leading to an increase in the number of reports. However, in countries or regions with a relatively high HDI level, the risk of transmission of infectious diseases may instead increase due to factors such as high urbanization or high population mobility levels, despite a relatively high overall levels of development. This finding aligns with the research conducted by Zilidis et al.[40] and Shahbazi et al.[41], who report that mortality and morbidity from infectious diseases, including COVID-19, are more pronounced in countries with higher socioeconomic statuses. Their study also shows that both morbidity and mortality from these diseases increase with increasing HDI levels[42].

Figure 3(e-f) shows the relationships between infectious disease outbreaks and climactic factors, including TEM and PRECIP. When 1.9 °C < TEM < 23 °C, the number of infectious disease outbreaks increases with increasing temperature, providing a more suitable environment for the spread of those infectious diseases suitable for high-temperature environments, but temperatures above 23 °C are unsuitable for the spread of infectious diseases[43]. When the temperature is less than 1.9 °C, the number of infectious disease outbreaks decreases with increasing temperature, and an increase in temperature may have an inhibitory effect on the spread of those infectious diseases suitable for low-temperature environments. In terms of changes in precipitation, the number of infectious disease outbreaks shows a U-shaped trend of decreasing and then increasing when 800 mm < PRECIP < 2900 mm. When PRECIP was < 800 mm, the risk of infectious disease outbreaks increases with increasing precipitation. When PRECIP > 2900 mm, the risk of infectious disease outbreaks begins to decrease with increasing precipitation. Head et al.[44] suggest that morbidity in arid regions is more sensitive to fluctuations in precipitation and that increased precipitation in areas with low precipitation may increase the risk of infectious disease outbreaks by providing pathogens with more favorable conditions for reproduction.

Since HDI can reflect the overall socioeconomic development level of a country in a more holistic way, we further combine HDI, precipitation, and temperature to explore the trend of infectious disease outbreaks under the influence of both socioeconomic and climate change factors. According to the results in Fig. 3(d-f), the numbers of infectious disease outbreaks show an overall increasing trend when HDI increases from 0.19 to 0.55 and from 0.74 to 0.93, when the temperature increases from 1.9 °C to 23 °C, and when the precipitation increases from 6.75 mm to 800 mm and from 1,600 mm to 2,900 mm. The above results can help countries formulate reasonable prevention and control measures according to their own development situations and optimize the allocation of public health resources.

Discussion

The concentrated outbreaks of global infectious diseases are related to climate change

There is a relationship between global outbreaks of infectious diseases and climate change. Global outbreaks of infectious diseases peaked in 2010 and 2016. Studies have shown that in 2009 and 2010, various infectious diseases frequently broke out and spread globally; for example, the outbreak of H1N1 influenza posed a serious challenge to all continents. From 2015 to 2016, an abnormally strong El Niño event occurred, and studies have shown that El Niño events can indirectly and directly affect temperature and precipitation, affect the occurrence of extreme weather events[45], and increase the incidence of various infectious diseases, such as cholera and malaria. Therefore, the concentrated outbreaks of infectious diseases in different countries and regions in 2017 may be related to climate change.

Socioeconomic factors are more worthy of attention than climate factors

Since 2000, the spatial extent of global outbreaks of infectious diseases has rapidly expanded. In the 21st century, with the development of economic globalization, factors such as increased population mobility and accelerated urbanization have led to frequent outbreaks of infectious diseases internationally, with a larger scope and greater impact. Some viruses began to spread widely in multiple countries worldwide, resulting in a significant increase in the number of infectious diseases that broke out in 2000, such as H1N1 influenza and dengue fever[4]. From a socioeconomic perspective, countries with higher GDPs, higher population density, and higher development levels have a greater risk of infectious disease outbreaks (Fig. 3). In regions with higher GDPs, the risk of infectious disease outbreaks is even greater. On the one hand, this finding may be due to the fact that low-income countries often lack infectious disease monitoring systems that can report outbreaks of infectious diseases in a timely manner. Therefore, in regions and countries with lower GDPs, as GDPs increase, the number of infectious disease outbreaks also increases rapidly. On the other hand, economically developed regions attract many people from other regions, thus affecting the transmission and diffusion of infectious diseases[46].

The role of government epidemic prevention, control and intervention cannot be ignored

In early 2020, to combat the COVID-19 pandemic, the government and various organizations took several measures to counter the spread of the disease, including enacting travel restrictions and home quarantines[47–49]. Epidemic prevention policies related to COVID-19 may have affected mainly the transmission of other infectious diseases by reducing the degrees of interpersonal contact and environmental exposure. Studies have shown that compared with other years, in 2020, the incidence of infectious diseases other than COVID-19, especially respiratory and gastrointestinal infectious diseases, significantly decreased in countries and regions that implemented strict epidemic prevention measures[33]. These findings indicate that epidemic prevention policies and measures related to infectious diseases have had some impact. However, considering the number of infectious disease outbreaks in the past decade, the overall trend continues to increase.

This study attempts to formulate policy recommendations for preventing and controlling infectious disease outbreaks from the perspective of both socioeconomic and climatic impacts. First, the complex interaction of climatic and socioeconomic factors may lead to an increased risk of infectious disease outbreaks, thus establishing the need for an integrated emergency response mechanism that includes health, the environment, and economic sectors to better coordinate resources, information, and actions. Second, for countries and regions with low levels of development, special attention should be paid to the early surveillance and early warning of infectious diseases. When primary healthcare institutions and community health organizations are supported, their disease surveillance and response capacity should be strengthened to ensure that potential outbreaks can be quickly identified and reported. In countries and regions with high levels of development, the focus should be on preventing widespread infectious disease outbreaks. Rapid preventive and control measures should be taken via a well-developed public health system, an advanced laboratory testing network and an efficient resource allocation system. Finally, as the trend of the globalization of infectious diseases intensifies, countries should pay more attention to regional cooperation and actively implement regional preventive and control measures to enhance their capacity to address infectious diseases.

This study seeks to fill knowledge gaps, unraveling the complex linkages between climatic and socioeconomic shifts and the infectious disease outbreak behaviors. These findings are essential for developing strategies for disease surveillance and intervention, ultimately assisting in a more effective global response to the ongoing challenge of infectious disease outbreaks. Although this study analyzes the long-term spatiotemporal distribution and factors influencing infectious disease outbreaks from multiple perspectives, several limitations remain. Infectious diseases are not classified according to type. Owing to the diversity of infectious diseases, the regional prevalence and trends of different types of infectious diseases differ. In addition to the climatic and socioeconomic factors considered in this study, numerous other factors that may influence the transmission and outbreak of infectious diseases, such as government policies and measures for epidemic prevention and national literacy in epidemic prevention, etc. Although these factors are not included in the model, they may still affect the results. In future research, we will incorporate additional factors influencing infectious disease outbreaks.

Methods

Analysis

We use SHAP values to calculate the importance of climatic and socioeconomic factors in predicting infectious disease outbreaks. The SHAP method is a method for interpreting the predictions of machine learning models, and its goal is to help us understand how models make their predictions[50]. In summary, the SHAP method acts as a translator, translating the "language" of the model into a human-understandable "language". In this way, we can not only know what the model predicts but also understand why the model makes such predictions[51]. The SHAP method works on the basis of the concept of game theory and evaluates the contribution of each feature to the prediction by calculating its SHAP value, which takes into account the different contributions of each feature in different combinations of features and distributes these contributions equally among the features. The SHAP value of each feature is then obtained by averaging all feature combinations to determine the importance ranking of each feature in a given set of features, which helps in interpreting the prediction results[52–55]. By plotting the importance of the features, important influences can be mined; by plotting the relationships and interactions of feature SHAP values with the values of all the sample features in the dataset, corresponding influences can be constructed on the basis of their trends.

In addition, we employ a GAM to explore the relationships between infectious disease outbreaks and climate and socioeconomic factors and identify the variations in these relationships. A GAM combines generalized linear models with additive models, allowing for the flexible handling of complex linear and nonlinear relationships between independent and dependent variables[56]. The basic form of a GAM is expressed as follows:

$$g\left(\mu_i\right) = \alpha_0 + f_1\left(x_{1i}\right) + f_2\left(x_{2i}\right) + \cdots + \varepsilon \tag{1}$$

where $\mathrm{f}_1\left(\mathrm{x}_{1\mathrm{i}}\right), \mathrm{f}_2\left(\mathrm{x}_{2\mathrm{i}}\right), \ldots$ represent various smoothing functions. GAMs do not impose requirements on the form of independent variables, allowing for the simultaneous fitting of both linear and nonlinear variables. For nonlinear variables, smoothing terms can be used for fitting to capture the relationship between nonlinear variables and the dependent variable as much as possible. In this study, the GAM is constructed via the mgcv package in R (Version 4.3.3). The model's quality is assessed through the check() function[57].

This work uses the Akaike information criterion (AIC) for model selection. The AIC is a statistical indicator used to balance the goodness of fit and complexity of a model. Through model comparison, the calculation formula is as follows:

$$AIC = -2L + 2k \tag{2}$$

where l represents the maximum likelihood estimation value and k represents the number of parameters in the model. A smaller AIC value indicates better model performance in terms of fitting data and explaining the research question.

In the process of model construction, we refer to the idea of full subset regression for optimal subset selection. By combining different variables to establish multiple GAMs, the best model combination is determined on the basis of the AIC minimum criterion, which is shown in Supplementary Table S1. The GAM established for the infectious disease outbreak data in this study is presented below. Here, s represents the smooth function for variables; $\mathrm{Y}_{\mathrm{i,\,t}}$ denotes the number of infectious disease outbreaks in country i in year t, where i represents different countries, t represents different years, ε represents the residual, and k represents the number of basic functions used in the smoothing term. β_{i} represents the overall average response. The variables include GDP, UE, POP, HDI, TEM, and PRECIP. Because $\mathrm{Y}_{\mathrm{i,\,t}}$ is a count variable, it is assumed to follow a Poisson distribution.

$$\begin{gathered} E\left(Y_{i,t}\right) = \beta_0 + s\left(GDP_{t,i}, k\right) + s\left(UE_{t,i}, k\right) + s\left(POP_{t,i}, k\right) + s\left(HDI_{t,i}, k\right) \\ + s\left(TEM_{t,i}, k\right) + s\left(PRECIP_{t,i}, k\right) + YEAR_t + Country_i + \varepsilon, family = poisson \end{gathered} \tag{3}$$

(2) Data

Infectious disease outbreak data

The infectious disease outbreak data for various countries and regions are sourced from the GIDEON database, which is a medical database that continuously updates data regarding the regional presence and prevalence of pathogens[58]. The records of epidemic outbreaks in the database are updated in real time through various sources, including the World Health Organization, ensuring the credibility and accuracy of the data. This study compiles and collects data on 244 infectious disease outbreaks from 193 countries for the period 1979–2021 as the research sample. The map in Fig. 1 was created using ArcGIS 10.7 (https://support.esri.com/en-us/products/arcmap).

Climate data

The climate data used in this study are sourced from the Climate Change Knowledge Portal (CCKP) dataset published by the World Bank[59]. For this research, annual mean precipitation (mm) and temperature (°C) data for each country are compiled and obtained for the period 1991–2020 as the research sample.

Socioeconomic data

The socioeconomic data used in this study are derived from World Bank publications[60], encompassing GDP and population density data for various countries and regions for the period 1991–2020, along with unemployment rate data.

HDI

Additionally, this study adopts the HDI to measure the development level of different countries[61]. The specific HDI data and classification criteria used in this paper come from the Human Development Reports published by the United Nations Development Programme (UNDP) for the period 1991–2021, covering 189 countries or regions. The HDI, constructed by the UNDP, reflects the social and human development levels of various countries and regions worldwide. In 2021, among the 193 selected countries and regions, 65 had very high HDI values, 49 had high HDI values, 43 had medium HDI values, 32 had low HDI values, and 4 had missing HDI values.

Data availability

All data are available from open-access websites, and the datasets used and analyzed during the current study are available from the corresponding author upon reasonable request.

Received: 15 January 2024; Accepted: 16 December 2024
Published online: 08 May 2025

References

1. Jones, K. E. et al. Global trends in emerging infectious diseases. *Nature* **451**, 990–993 (2008).
2. Smith, K. F. et al. Global rise in human infectious disease outbreaks. *J. R Soc. Interface*. **11**, 20140950 (2014).
3. Heymann, D. L. The international response to the outbreak of SARS in 2003. *Philos. Trans. R Soc. Lond. B Biol. Sci.* **359**, 1127–1129 (2004).
4. Viboud, C. & Simonsen, L. Global mortality of 2009 pandemic influenza A H1N1. *Lancet Infect. Dis.* **12**, 651–653 (2012).
5. De Groot, R. J. et al. Commentary: Middle East Respiratory Syndrome Coronavirus (MERS-CoV): Announcement of the Coronavirus Study Group. *J. Virol.* **87**, 7790–7792 (2013).
6. Heymann, D. L. et al. Global health security: The wider lessons from the west African Ebola virus disease epidemic. *Lancet* **385**, 1884–1901 (2015).
7. Chang, C., Ortiz, K., Ansari, A. & Gershwin, M. E. The Zika outbreak of the 21st century. *J. Autoimmun.* **68**, 1–13 (2016).
8. Baker, R. E. et al. Infectious disease in an era of global change. *Nat. Rev. Microbiol.* **20**, 193–205 (2022).
9. Weiss, R. A. & McMichael, A. J. Social and environmental risk factors in the emergence of infectious diseases. *Nat. Med.* **10**, S70–S76 (2004).
10. Altizer, S., Ostfeld, R. S., Johnson, P. T. J., Kutz, S. & Harvell, C. D. Climate change and infectious diseases: From evidence to a predictive framework. *Science* **341**, 514–519 (2013).
11. Lafferty, K. D. The ecology of climate change and infectious diseases. *Ecology* **90**, 888–900 (2009).
12. Carlson, C. J. et al. Climate change increases cross-species viral transmission risk. *Nature* **607**, 555–562 (2022).
13. Ma, Y. et al. Linking climate and infectious disease trends in the Northern/Arctic Region. *Sci. Rep.* **11**, 20678 (2021).
14. Costello, A. et al. Managing the health effects of climate change. *Lancet* **373**, 1693–1733 (2009).
15. Whitmee, S. et al. Safeguarding human health in the Anthropocene epoch: Report of The Rockefeller Foundation–Lancet Commission on planetary health. *Lancet* **386**, 1973–2028 (2015).
16. Keesing, F. et al. Impacts of biodiversity on the emergence and transmission of infectious diseases. *Nature* **468**, 647–652 (2010).
17. Harvell, C. D. et al. Climate warming and disease risks for terrestrial and marine biota. *Science* **296**, 2158–2162 (2002).
18. Semenza, J. C. & Menne, B. Climate change and infectious diseases in Europe. *Lancet Infect. Dis.* **9**, 365–375 (2009).
19. Patz, J. A., Campbell-Lendrum, D., Holloway, T. & Foley, J. A. Impact of regional climate change on human health. *Nature* **438**, 310–317 (2005).
20. Morand, S., Owers, K. A., Waret-Szkuta, A., McIntyre, K. M. & Baylis, M. Climate variability and outbreaks of infectious diseases in Europe. *Sci. Rep.* **3**, 1774 (2013).
21. Wang, Q., Wang, J. & Gao, F. Who is more important, parents or children? Economic and environmental factors and health insurance purchase. *N. Am. J. Econ. Finance.* **58**, 101479 (2021).
22. Walther, G. R. et al. Ecological responses to recent climate change. *Nature* **416**, 389–395 (2002).
23. Gibb, R. et al. Interactions between climate change, urban infrastructure and mobility are driving dengue emergence in Vietnam. *Nat. Commun.* **14**, 8179 (2023).
24. Randolph, S. E. To what extent has climate change contributed to the recent epidemiology of tick-borne diseases? *Vet. Parasitol.* **167**, 92–94 (2010).
25. Redding, D. W. et al. Impacts of environmental and socio-economic factors on emergence and epidemic potential of Ebola in Africa. *Nat. Commun.* **10**, 4531 (2019).

26. Bhutta, Z. A., Salam, R. A., Das, J. K. & Lassi, Z. S. Tackling the existing burden of infectious diseases in the developing world: Existing gaps and the way forward. *Infect. Dis. Poverty*. **3**, 28 (2014).
27. Hotez, P. J. Globalists versus nationalists: Bridging the divide through blue marble health. *PLoS Negl. Trop. Dis.* **13**, e0007156 (2019).
28. Mahon, M. B. et al. A meta-analysis on global change drivers and the risk of infectious disease. *Nature* **629**, 830–836 (2024).
29. Dimitrova, A. et al. Temperature-related neonatal deaths attributable to climate change in 29 low- and middle-income countries. *Nat. Commun.* **15**, 5504 (2024).
30. Ren, X., Xiao, Y., Xiao, S., Jin, Y. & Taghizadeh-Hesary, F. The effect of climate vulnerability on global carbon emissions: Evidence from a spatial convergence perspective. *Resour. Policy*. **90**, 104817 (2024).
31. Moore, F. C. et al. Determinants of emissions pathways in the coupled climate–social system. *Nature* **603**, 103–111 (2022).
32. Chitre, S. D. et al. The impact of anthropogenic climate change on pediatric viral diseases. *Pediatr. Res.* **95**, 496–507 (2024).
33. Geng, M. J. et al. Changes in notifiable infectious disease incidence in China during the COVID-19 pandemic. *Nat. Commun.* **12**, 6923 (2021).
34. Hotez, P. J. Blue marble health and the big three diseases: HIV/AIDS, tuberculosis, and malaria. *Microbes Infect.* **17**, 539–541 (2015).
35. Hotez, P. J., Damania, A. & Naghavi, M. Blue marble health and the Global Burden of Disease Study 2013. *PLoS Negl. Trop. Dis.* **10**, e0004744 (2016).
36. Yu, D. et al. Whether urbanization has intensified the spread of infectious diseases—Renewed question by the COVID-19 pandemic. *Front. Public. Health*. **9**, 699710 (2021).
37. Jaakkola, K. et al. Decline in temperature and humidity increases the occurrence of influenza in cold climate. *Environ. Health*. **13**, 22 (2014).
38. Mora, C. et al. Over half of known human pathogenic diseases can be aggravated by climate change. *Nat. Clim. Change*. **12**, 869–875 (2022).
39. Hotez, P. J. & Booker, C. S. T. O. P. Study, treat, observe, and prevent neglected diseases of poverty act. *PLoS Negl. Trop. Dis.* **14**, e0008064 (2020).
40. Zilidis, C., Papagiannis, D. & Kyriakopoulou, Z. Did Economic crisis affect mortality due to infectious diseases? Trends of infectious diseases mortality in Greece before and after economic crisis. *Cureus* https://doi.org/10.7759/cureus.13621 (2021).
41. Shahbazi, F. & Khazaei, S. Socio-economic inequality in global incidence and mortality rates from coronavirus disease 2019: An ecological study. *New. Microbes New. Infect.* **38**, 100762 (2020).
42. Hotez, P. J. Blue marble health redux: Neglected tropical diseases and human development in the Group of 20 (G20) nations and Nigeria. *PLoS Negl. Trop. Dis.* **9**, e0003672 (2015).
43. Zhang, X. et al. The effect of temperature on infectious diarrhea disease: A systematic review. *Heliyon* **10**, e31250 (2024).
44. Head, J. R. et al. Effects of precipitation, heat, and drought on incidence and expansion of coccidioidomycosis in western USA: A longitudinal surveillance study. *Lancet Planet. Health*. **6**, e793–e803 (2022).
45. Iyer, V. et al. Role of extreme weather events and El Niño southern oscillation on incidence of enteric fever in Ahmedabad and Surat, Gujarat, India. *Environ. Res.* **196**, 110417 (2021).
46. Zhang, X. et al. Spatiotemporal variations in the incidence of bacillary dysentery and long-term effects associated with meteorological and socioeconomic factors in China from 2013 to 2017. *Sci. Total Environ.* **755**, 142626 (2021).
47. Li, X., Hui, E. C. M. & Shen, J. Institutional development and the government response to COVID-19 in China. *Habitat Int.* **127**, 102629 (2022).
48. Qiao, S., He, M., Wang, J., Cai, J. & Zheng, J. Robust optimization for a dynamic emergency materials supply chain network under major infectious disease epidemics. *Int. J. Logist Res. Appl.* 1–36. https://doi.org/10.1080/13675567.2023.2269101 (2023).
49. Liu, Y. et al. Countermeasures against economic crisis from COVID-19 pandemic in China: An analysis of effectiveness and trade-offs. *Struct. Change Econ. Dyn.* **59**, 482–495 (2021).
50. Lundberg, S. M. & Lee, S. I. *A Unified Approach to Interpreting Model Predictions* (2017).
51. Su, Z., Cai, X. & Wu, Y. Exchange rates forecasting and trend analysis after the COVID-19 outbreak: new evidence from interpretable machine learning. *Appl. Econ. Lett.* **30**, 2052–2059 (2023).
52. Kavzoglu, T. & Teke, A. Predictive performances of ensemble machine learning algorithms in landslide susceptibility mapping using random forest, extreme gradient boosting (XGBoost) and natural gradient boosting (NGBoost). *Arab. J. Sci. Eng.* **47**, 7367–7385 (2022).
53. Kim, Y. & Kim, Y. Explainable heat-related mortality with random forest and SHapley Additive exPlanations (SHAP) models. *Sustain. Cities Soc.* **79**, 103677 (2022).
54. Morand, S. & Walther, B. A. Individualistic values are related to an increase in the outbreaks of infectious diseases and zoonotic diseases. *Sci. Rep.* **8**, 3866 (2018).
55. Sorci, G. et al. Ranking parameters driving siring success during sperm competition in the North African houbara bustard. *Commun. Biol.* **6**, 1–12 (2023).
56. Hastie, T. & Tibshirani, R. *Generalized Additive Models* (Chapman & Hall/CRC, 1999).
57. Morand, S. Emerging diseases, livestock expansion and biodiversity loss are positively related at global scale. *Biol. Conserv.* **248**, 108707 (2020).
58. Global Infectious Diseases and Epidemiology Network. *GIDEON*. https://www.gideononline.com/
59. Data Catalog |. Climate Change Knowledge Portal. https://climateknowledgeportal.worldbank.org/download-data
60. World Bank Open. Data | Data. https://data.worldbank.org.cn/
61. Home | Human Development Reports. https://hdr.undp.org/

Author contributions

Z.Q. provided the main ideas and designed the study. K.Z. provided technical and data support, and provided financial support. L.Z. designed relevant codes. Y.S. is responsible for proofreading and reviewing manuscripts. B.Z. is responsible for visualization and investigation. All authors reviewed the manuscript.

Funding

This work was supported by the National Natural Science Foundation project of China (72004086) and the Humanities and Social Science Project of the Ministry of Education (24YJC630170).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-024-83431-2.

Correspondence and requests for materials should be addressed to K.Z.