

Article

Yearly and Daily Relationship Assessment between Air Pollution and Early-Stage COVID-19 Incidence: Evidence from 231 Countries and Regions

Yuan Meng ¹, Man Sing Wong ^{1,2,*}, Hanfa Xing ^{3,4}, Mei-Po Kwan ^{5,6} and Rui Zhu ¹

¹ Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong, China; myuan.meng@connect.polyu.hk (Y.M.); felix.zhu@polyu.edu.hk (R.Z.)

² Research Institute for Sustainable Urban Development, The Hong Kong Polytechnic University, Hong Kong, China

³ School of Geography, South China Normal University, Guangzhou 510000, China; xinghanfa@scnu.edu.cn

⁴ College of Geography and Environment, Shandong Normal University, Jinan 250100, China

⁵ Department of Geography and Resource Management, and Institute of Space and Earth Information Science, The Chinese University of Hong Kong, Hong Kong, China; mpk654@gmail.com

⁶ Department of Human Geography and Spatial Planning, Utrecht University, 3584 CB Utrecht, The Netherlands

* Correspondence: ls.charles@polyu.edu.hk; Tel.: +852-3400-8959



Citation: Meng, Y.; Wong, M.S.; Xing, H.; Kwan, M.-P.; Zhu, R. Yearly and Daily Relationship Assessment between Air Pollution and Early-Stage COVID-19 Incidence: Evidence from 231 Countries and Regions. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 401. <https://doi.org/10.3390/ijgi10060401>

Academic Editors: Fazlay S. Faruque and Wolfgang Kainz

Received: 7 April 2021

Accepted: 4 June 2021

Published: 10 June 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract: The novel coronavirus disease 2019 (COVID-19) has caused significantly changes in worldwide environmental and socioeconomics, especially in the early stage. Previous research has found that air pollution is potentially affected by these unprecedented changes and it affects COVID-19 infections. This study aims to explore the non-linear association between yearly and daily global air pollution and the confirmed cases of COVID-19. The concentrations of tropospheric air pollution (CO , NO_2 , O_3 , and SO_2) and the daily confirmed cases between 23 January 2020 and 31 May 2020 were collected at the global scale. The yearly discrepancies of air pollutions and daily air pollution are associated with total and daily confirmed cases, respectively, based on the generalized additive model. We observed that there are significant spatially and temporally non-stationary variations between air pollution and confirmed cases of COVID-19. For the yearly assessment, the number of confirmed cases is associated with the positive fluctuation of CO , O_3 , and SO_2 discrepancies, while the increasing NO_2 discrepancies leads to the significant peak of confirmed cases variation. For the daily assessment, among the selected countries, positive linear or non-linear relationships are found between CO and SO_2 concentrations and the daily confirmed cases, whereas NO_2 concentrations are negatively correlated with the daily confirmed cases; variations in the ascending/declining associations are identified from the relationship of the O_3 -confirmed cases. The findings indicate that the non-linear relationships between global air pollution and the confirmed cases of COVID-19 are varied, which implicates the needs as well as the incorporation of our findings in the risk monitoring of public health on local, regional, and global scales.

Keywords: COVID-19; confirmed cases; air pollution; generalized additive model



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The novel coronavirus disease, known as COVID-19, is caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and the first confirmed case was found in Wuhan, China in December 2019 [1,2]. In the following months, COVID-19 spread across the country, and eventually the disease was transmitted rapidly on a global scale and evaluated as a global pandemic [3,4]. The World Health Organization (WHO) reported 1,610,909 global confirmed cases and 99,690 deaths, as of April 2020 [5]. To control the spread of COVID-19, measures such as government-enforced lockdowns and public gathering prohibitions have been implemented in many countries [6,7]. These strict measures

have brought about substantial environmental and socioeconomic changes, which in turn exert a great influence on human health [8,9].

These environmental and socioeconomic changes have led to alterations in energy uses in traffic, industries, and residences, and have further caused both short- and long-term changes in air pollutants such as Carbon Monoxide (CO), Nitrogen Dioxide (NO₂), Ozone (O₃), and Sulphur Dioxide (SO₂) [10,11]. Research has demonstrated that socioeconomic growth has been exacerbating anthropogenic emissions in the past decades, which suggests that there is an association between different uses of energy caused by economic activity and air pollution changes [12,13]. Previous studies have also found that air pollution is a risk factor for respiratory infection, which is associated with people's systematic immunity [14,15]. Therefore, it is important to evaluate the effect of COVID-19 related policies on air pollution changes and its relationship with human health.

Existing studies have investigated the impact of restriction policies for personal mobility and economic activity on air pollution trends, and the relationship between air pollutants and COVID-19 transmission [16,17] in the early stage. For instance, He et al. [18] evaluated the influence of the lockdown policy in China on Air Quality Index (AQI) and the concentrations of particulate matter with a diameter of less than 2.5 μm (PM_{2.5}), and the results reveal that the lockdown effect is greater in highly industrialized and developed cities. Venter et al. [19] further extended the research to lockdown policies at the global scale and found a relationship between the decline in global vehicle transportation and reduction in NO₂. On this basis, the research emphasis has focused on the effects of changes of air pollution trends on COVID-19 infection, and the results showed that there was a positive relationship between daily PM_{2.5}, PM_{2.5}, NO₂, and O₃ and daily confirmed cases and a negative between SO₂ and daily confirmed cases in China [20]. In addition, Wu et al. [21] found that an increase of 1 μg/m³ in PM_{2.5} is associated with an 8% increase in the death rate of COVID-19 in the U.S.

Although research has linked air pollution trends with COVID-19 transmission in the early stages, the estimations are mainly focused on linear relationships, either increasing or decreasing, between air pollution and the growing number of confirmed cases or death toll in individual countries. The non-linear association between air pollution and COVID-19 transmission and the comparison of non-linear associations in different countries at the global scale have not yet been discussed. Moreover, the emphasis was on the short-term impact of air pollution during a short period in previous studies, while the impact of long-term air pollution, i.e., yearly changes, on early-stage COVID-19 transmission was seldom discussed. In fact, the spread of COVID-19 in the early stage can not only reflect the process of the epidemic from the outbreak to containment but also be quantified without the control of vaccination.

To fill this gap, the association between yearly and daily air pollution, including CO, NO₂, O₃, and SO₂ and confirmed cases of COVID-19 at the global scale is investigated based on the generalized additive models (GAMs). For the yearly air pollution, discrepancies between different air pollution trends during four months in 2020 (during COVID-19) and the same period in 2019 (without COVID-19) are quantified through the Dynamic Time Warping (DTW) approach, and their association with the total confirmed cases of COVID-19 are analyzed. For daily air pollution, the non-linear relationships between daily air pollution concentrations and the daily confirmed cases of each country are analyzed.

The remainder of this paper is organized as follows. In Section 2, the materials and methods used in this research are introduced. In Section 3, the relationship between long-term and short-term air pollution and confirmed cases of COVID-19 is studied. In Section 4, the relationships between air pollution and the confirmed cases are analyzed and the limitations of this study are discussed. Finally, the conclusion is drawn in Section 5.

2. Materials and Methods

2.1. Study Area and Data

- Study area

231 countries and regions on a global scale have been selected for the yearly and daily relationship assessment of air pollution and confirmed cases of COVID-19. Due to the significant differences in environmental and socioeconomic conditions among countries, only the differences of air pollution between 2019 and 2020 were quantified in the yearly analysis, whereas the daily relationship analysis was proposed based on each country independently. Nine countries with the most confirmed cases of COVID-19 within this period were further selected as examples to illustrate the daily relationship in detail.

- Air pollution data

Four types of pollutants were obtained, including CO, NO₂, O₃, and SO₂ from TROPOspheric Monitoring Instrument (TROPOMI) in Sentinel-5 Precursor satellite [22]. Previous research has evaluated the quality of data on CO, NO₂, O₃, and SO₂ in different places, and confirmed their usability [23–25]. Air pollution data between 23 January 2019 and 31 May 2019 and between 23 January 2020 and 31 May 2020 were collected and processed on the Google Earth Engine platform, with global mean column concentration calculated each day at the country level. It is worth noting that missing values of data on air pollution time series were further processed.

- Meteorological data

Meteorological data including temperature, precipitation, and wind speed were collected from Global Surface Summary of the Day—GSOD in the Integrated Surface Hourly (ISH) dataset as the controlling variables. The collected data were aggregated into country level and further processed as daily data during 23 January 2020–31 May 2020. In addition, the missing values were processed using KalmanSmoother based on an autoregressive integrated moving average (ARIMA) model [26].

- COVID-19 confirmed cases

Confirmed cases of COVID-19 in each country between 23 January 2020 and 31 May 2020 were obtained from the dashboard established by the Johns Hopkins University Center for Systems Science and Engineering (JHU CCSE) (<https://github.com/CSSEGISandData/COVID-19>, accessed on: 30 June 2020) [27]. Specifically, data from the dashboard are mainly collected from the World Health Organization, regional health department, and online news services, and are updated every 15 min. From 22 January 2020, daily case reports, including the counts of confirmed, deaths, and recovered cases in each country/region are provided. To investigate the association between air pollution and the confirmed cases of COVID-19 at different temporal scales, both new daily confirmed cases and total confirmed cases as of May 2020 were collected at the country-level.

2.2. Measures of Variables

- Yearly changes of air pollution

To quantify yearly changes of CO, NO₂, O₃, and SO₂ between 2019 and 2020, we summarize daily CO, NO₂, O₃, and SO₂ among each country within 2 years (that is, between 23 January 2019 and 31 May 2019 and between 23 January 2020 and 31 May 2020, respectively) separately, and calculate the discrepancies of the time-series air pollution using DTW [28,29]. Specifically, the main purpose of identifying air pollution discrepancies using DTW is to find the optimal warping path, which refers to the series distances between the neighboring air pollution values and the minimum cumulative values. With the optimal path, the

warping cost between time series is minimized— $S_{2019} = s_{2019,1}, \dots, s_{2019,i}, \dots, s_{2019,n}$ and $S_{2020} = s_{2020,1}, \dots, s_{2020,i}, \dots, s_{2020,n}$:

$$\text{DTW}(S_{2019}, S_{2020}) = \underset{W=w_1 \dots w_k \dots w_K}{\operatorname{argmin}} \sqrt{\sum_{k=1, w_k=(i,j)}^K (s_{2019,i} - s_{2020,j})^2} \quad (1)$$

where $W = w_1 \dots w_k \dots w_K$ indicates the warping path aligned by time series S_{2019} and S_{2020} . The alignment of the minimum warping path is accomplished using a dynamic programming algorithm [30], and it is calculated as:

$$c(i,j) = d(s_{2019,i}, s_{2020,j}) + \min\{c(i-1, j-1), c(i-1, j), c(i, j-1)\} \quad (2)$$

where $c(i,j)$ indicate the cumulative distance between s_{2019} and s_{2020} . $d(s_{2019,i}, s_{2020,j})$ represents the distance between $s_{2019,i}$ and $s_{2020,j}$. In particular, higher values of cumulative distances indicate higher discrepancy between the two time-series.

Moreover, the average daily temperature, total precipitation, and wind speed during 23 January 2020–31 May 2020 were calculated as the controlling variables to reduce the biases in yearly air pollution COVID-19 confirmed cases estimation.

- Daily air pollution

Daily air pollution of CO, NO₂, O₃, and SO₂ are calculated based on the data on the TROPOMI platform within each country during 23 January 2020–31 May 2020, which are considered as independent variables in analyzing the relationship with confirmed cases of COVID-19. For the meteorological data as controlling variables, daily temperature, total precipitation, and wind speed in 2020 during 23 January 2020–31 May 2020 were adopted.

- Confirmed cases of COVID-19

To analyze the impact of yearly air pollution changes on the COVID-19 epidemic, the total confirmed cases during 23 January 2020–31 May 2020 are calculated within each country separately. Moreover, daily new confirmed cases at the country-level are also calculated to investigate how confirmed cases vary with air pollution changes.

2.3. Statistical Models

The non-linear relationship between both yearly and daily air pollution and confirmed cases of COVID-19 are analyzed. In particular, the GAMs are effective in quantifying the non-linear relationships between air pollution and the confirmed cases of COVID-19, and significant advantages of these models in association modelling compared with linear models are observed in past studies [31,32]. Thus, GAM with a Gaussian distribution is utilized to analyze the relationship between air pollution (including CO, NO₂, O₃ and SO₂) and confirmed cases of COVID-19 [33].

To investigate how yearly air pollution changes affect COVID-19 epidemic, the relationship between the yearly air pollution and the total confirmed cases of COVID-19 is analyzed with each country as samples. The discrepancies between air pollution on 23 January 2019 to 31 May 2019 and on 23 January 2020 to 31 May 2020 calculated in Section 2.2 are the independent variables, with average daily temperature, total precipitation, and wind speed as the controlling variables, and the total confirmed cases of COVID-19 in each country serve as the dependent variables. On this basis, the non-linear relationship analysis is conducted based on each independent variable separately using GAMs.

For daily analysis, the non-linear relationship between daily air pollution and daily confirmed cases of COVID-19 during 23 January 2020–31 May 2020 is estimated independently in each country. For the i th country, GAM is proposed as follows:

$$\log E(Y_i) = s(\text{CO}_i) + s(\text{NO}_2i) + s(\text{O}_3i) + s(\text{SO}_2i) + \text{PRCP}_i + \text{TEMP}_i + \text{WDSP}_i \quad (3)$$

where $\log E(Y_i)$ denotes the log-transformed expected values of the number of confirmed cases of COVID-19. $s(\cdot)$ denotes the penalized smoothing spline. CO_i , NO_{2i} , O_{3i} and SO_{2i} represent the summed column concentration of CO, NO₂, O₃ and SO₂ of the i th country. $PRCP_i$, $TEMP_i$ and $WDSP_i$ refer to the average daily precipitation, temperature, and wind speed in the i th country.

3. Results

3.1. Yearly Relationship Analysis

3.1.1. Yearly Air Pollution Discrepancies

Figure 1 displays the discrepancies between the air pollutions of CO, NO₂, O₃, and SO₂ during these two time-series between 23 January 2019 and 31 May 2019 and between 23 January 2020 and 31 May 2020. Higher values indicate larger distances and a higher degree of discrepancies. In specific, compared with CO column concentration in 2019, most significant changes in CO during the pandemic of COVID-19 are observed in Canada, Russia and China, followed by the U.S., India, Brazil, Australia, and Kazakhstan (Figure 1a). With regard to NO₂ column concentration, Figure 1b shows that China experienced the greatest changes during 2019–2020 followed by Russia, the U.S., and India, whereas most of the countries in Europe and Africa share lower discrepancies in terms of NO₂ trends. For the O₃ column concentration in Figure 1c, Russia presents the greatest difference between 2019–2020, followed by Canada, the U.S., and China. Figure 1c displays the discrepancy patterns of SO₂, in which the U.S., Russia, and China underwent the greatest changes during 2019–2020. It is worth noting that, as the air pollution trends are calculated based on the countries, the variety of country areas can affect the discrepancy patterns. However, because the confirmed cases of COVID-19 are analyzed based on the total number in each country, the biases caused in different country areas can be ignored.

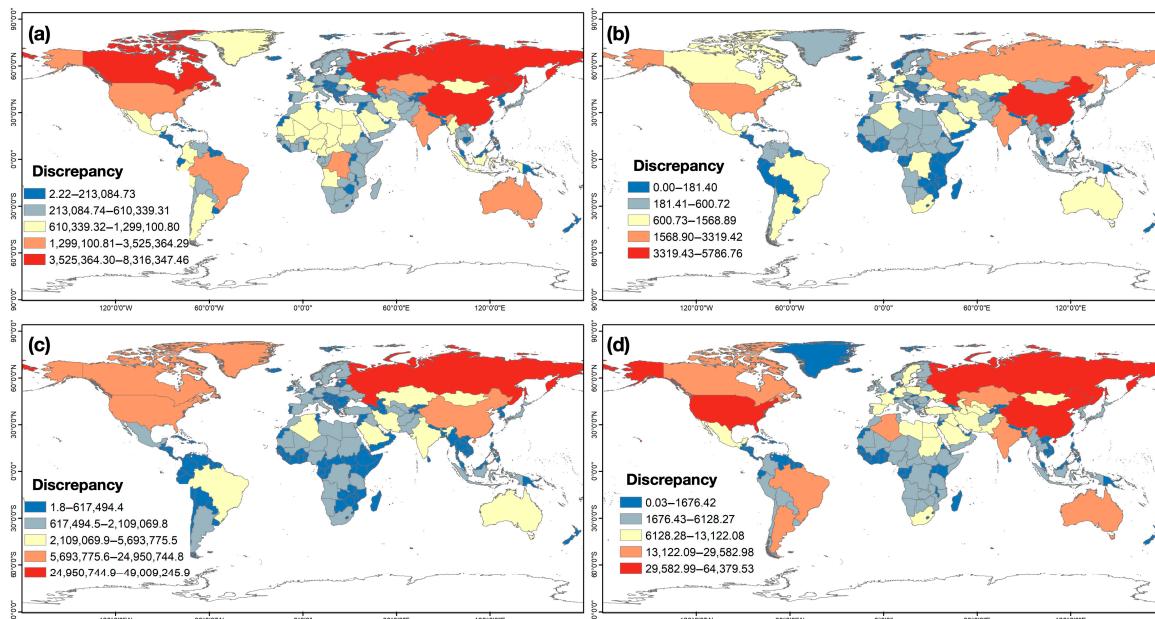


Figure 1. Air pollution discrepancies between two time periods, i.e., from 23 January 2019 to 31 May 2019 and from 23 January 2020 to 31 May 2020. (a) CO discrepancy. (b) NO₂ discrepancy. (c) O₃ discrepancy. (d) SO₂ discrepancy.

3.1.2. Relationship Analysis

Figure 2 demonstrates the non-linear relationship between the yearly air pollution discrepancies and the confirmed cases of COVID-19. The confirmed cases are quantified as the total number of cases on 31 May 2020 in each country and region. With each country as an independent sample, through the four GAM models, the relationship between CO, NO₂, O₃, SO₂, and total confirmed cases are revealed with the R square values of 0.541, 0.536, 0.367 and 0.82, respectively. The results indicate that the number of confirmed cases is rising with fluctuation with the increasing CO and O₃ discrepancies, while a significant peak is observed in the NO₂ discrepancies-confirmed cases relationship. A sharp increasing trend is shown in the SO₂ discrepancies-confirmed cases relationship.

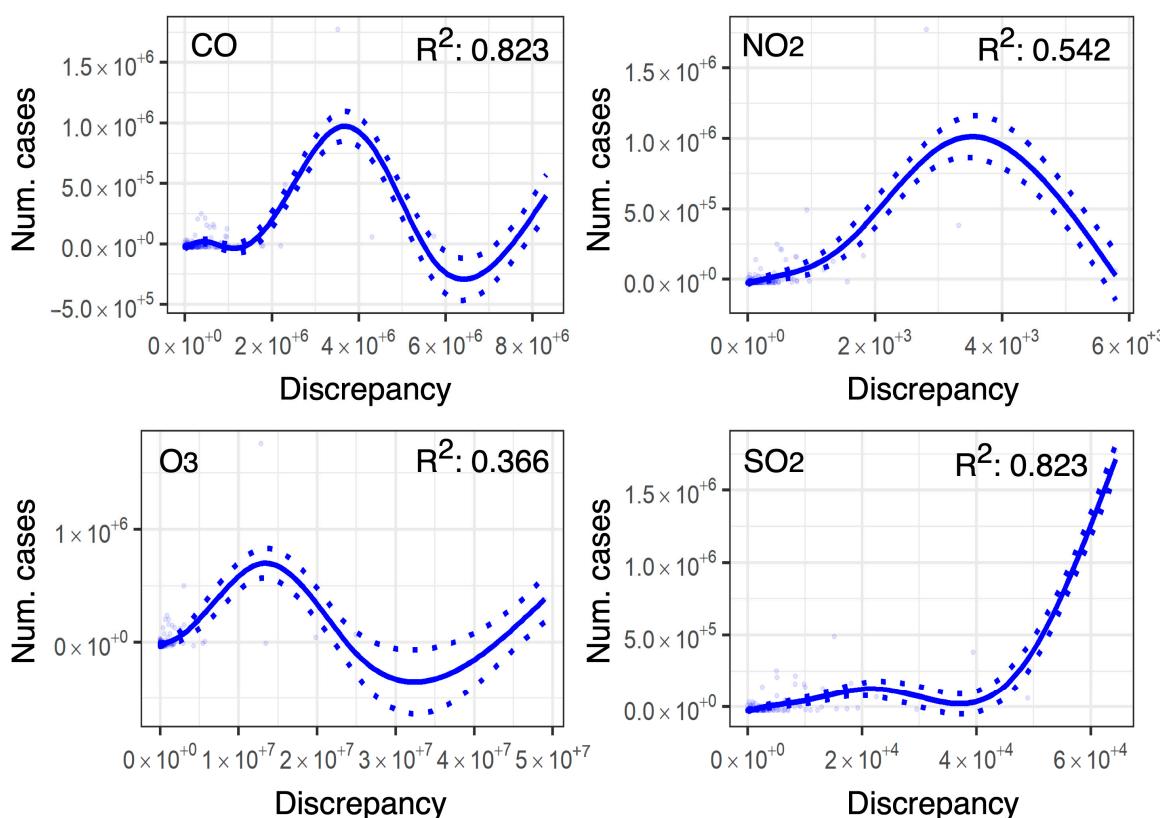


Figure 2. Non-linear relationship between air pollution discrepancies and confirmed cases of COVID-19. (a) CO discrepancy and confirmed cases. (b) NO₂ discrepancy and confirmed cases. (c) O₃ discrepancy and confirmed cases. (d) SO₂ discrepancy and confirmed cases.

3.2. Daily Relationship Analysis

3.2.1. Model Evaluation

Before analyzing the relationship between daily air pollution and the confirmed cases of COVID-19, the study performed a comparison between the linear regression model (LM) with the GAM to assess the modelling accuracies. Figure 3a,b show the R² values of LM and GAM, respectively. In particular, R² values of the LM range from 0–0.812, whereas R² values of GAM range from 0–0.948. Compared with the LM, through which only a few of countries, including the U.S., Canada, and Russia, show higher R² values greater than 0.5, more countries with higher R² values such as Germany, Brazil, India, and Mexico were obtained based on GAM and higher modelling accuracies are found.

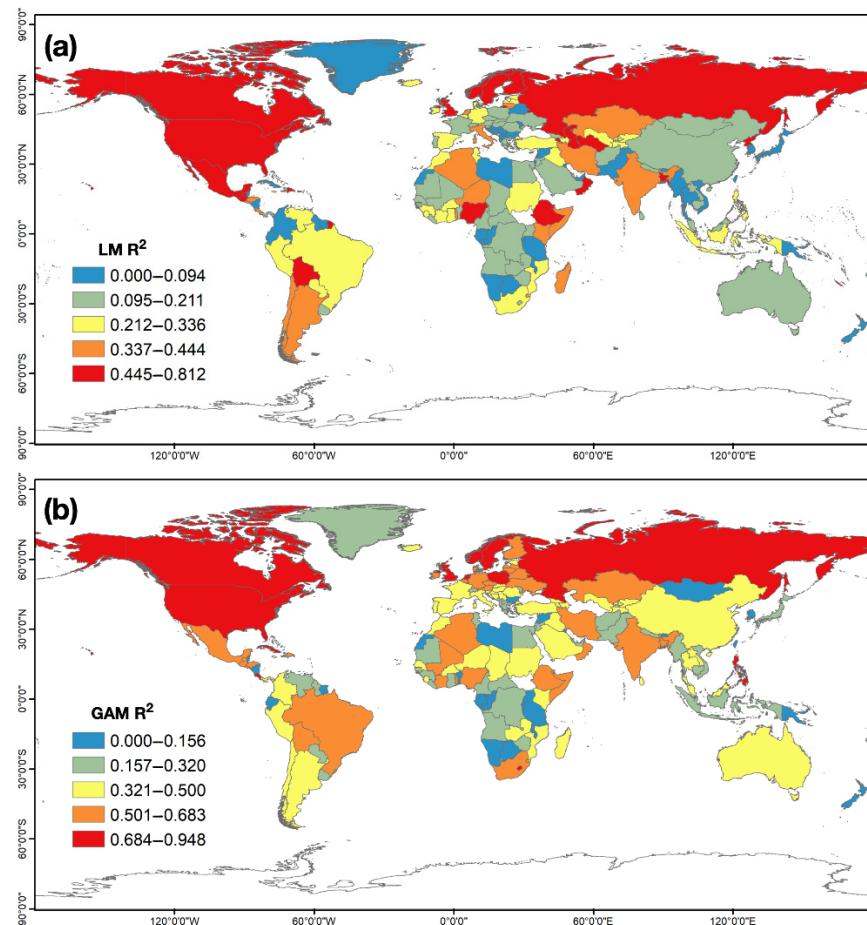


Figure 3. Distributions of R^2 values. (a) R^2 values of LM; (b) R^2 values of GAM.

Based on the GAMs, the effective degrees of freedom (EDF) in each country were calculated. Specifically, EDF is considered as a proxy to quantify the degree of non-linearity among the independent and dependent variables. $EDF = 1$ indicates a completely linear relationship, $EDF \geq 1$ and ≤ 2 indicate a weakly non-linear relationship, while $EDF \geq 2$ represents a highly non-linear relationship [34]. Figure 4a–d indicates the EDF of each country in the CO, NO₂, O₃, SO₂—confirmed cases of COVID-19 relationships. For the CO column concentration, countries such as Canada, Russia, and Australia show significantly non-linear patterns, while linear regressions are performed for countries such as Canada, China, Germany, and India in picturing the relationship of the CO-confirmed cases (Figure 4a). The relationship between NO₂ and confirmed cases reveals different patterns. A highly non-linear relationship is only found in a few countries and regions such as Russia, Canada, and Greenland, whereas a linear relationship is found in countries including Mexico, Australia, and several countries in South America, Europe, and Africa (Figure 4b). For the EDF of O₃, most of the countries represent highly or weakly O₃-confirmed cases relationship, while several countries and regions such as Greenland, U.S., Mexico, and Germany still show significant linear patterns (Figure 4c). The distribution of EDF of SO₂ indicates a highly non-linear relationship in North America, Europe, as well as several countries in Africa and Asia. In addition, weak non-linear relationships are found in Greenland, Argentina, and several countries in Africa (Figure 4d). The variations in EDF of CO, NO₂, O₃, and SO₂ reveal the non-stationary patterns of the air relationship between pollution and the confirmed cases of COVID-19.

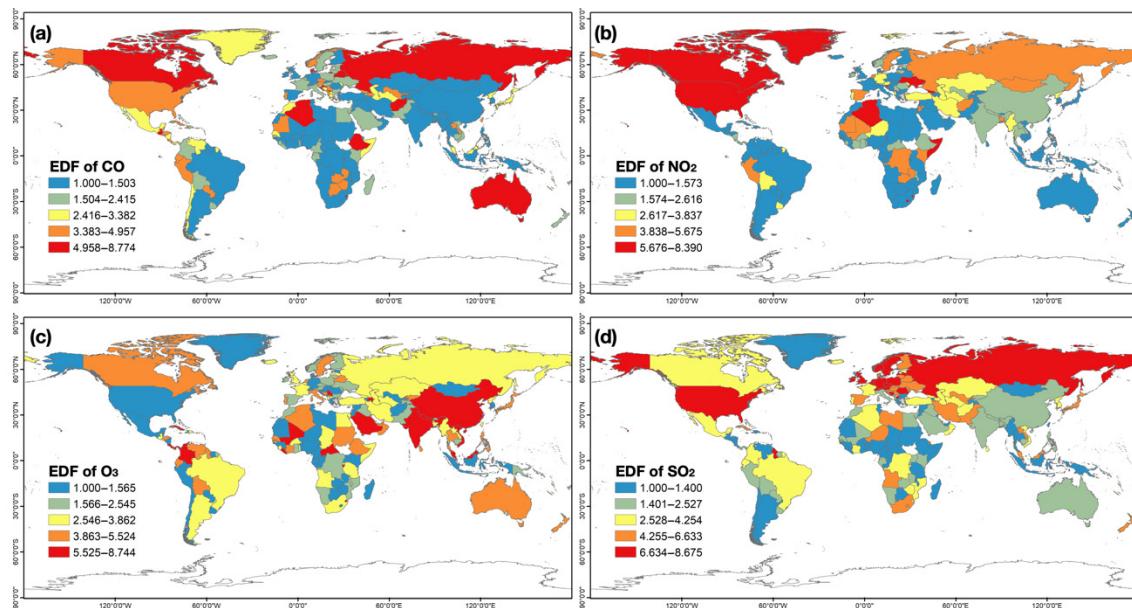


Figure 4. EDF of air pollutions estimated from GAM. (a) EDF of CO. (b) EDF of NO₂. (c) EDF of O₃. (d) EDF of SO₂.

3.2.2. Relationship Analysis

The variation in the confirmed cases of COVID-19 relative to the air pollution trends were further investigated based on the GAM. Specifically, countries with the most confirmed cases (calculated until May 2020) were selected for detailed analysis, including the U.S., Brazil, Russia, the U.K., Spain, Italy, India, France, and Germany.

Figure 5a shows the marginal effects of the CO column concentration on the daily confirmed cases of COVID-19 in nine countries. A significant trend of increasing daily confirmed cases is observed as the CO column concentration in the U.S. increases, whereas Brazil and India show trends of slightly declining daily confirmed cases. A positive CO-daily confirmed cases relationship is also found in Spain, Italy, and France, showing a stable tendency followed by a slightly growing trend of daily confirmed cases with the rising CO column concentration, respectively. Moreover, significantly non-linear variation is observed in Russia, indicating alternatively increasing and decreasing and finally ascending trends of daily confirmed cases under the ascending of CO column concentration. In addition, no significant changes in CO column concentration are observed in the U.K. and Germany.

The marginal effects of NO₂ column concentration on the confirmed cases of the selected countries are shown in Figure 5b. Only India represents a positive trend in NO₂-daily confirmed cases relationship. Six countries including the U.S., Russia, Spain, and Italy show significant trends of declining daily confirmed cases with the rising NO₂ column concentration. In particular, with the greater values of NO₂ column concentration, a slightly descending trend followed by a trend of dramatically decreasing daily confirmed cases are observed in the U.S., whereas a substantially decrease followed by a steady tendency are presented in Russia and Spain. In addition, a linear-decreasing relationship indicating lower NO₂ column concentration concerning increasing daily confirmed cases in Italy. The other four countries including Brazil, U.K., France, and Germany represent relatively unchanging trends with NO₂ column concentration changing.

Figure 5c displays the margin effects of the O₃ column concentration on the confirmed cases. Specifically, linear relationship is shown in the U.S., with stabilized trends of descending daily confirmed cases corresponding to rising O₃ column concentration. Six countries including Brazil, Russia, the U.K., Spain, Italy, India, and France represent non-linear variation patterns in terms of O₃-confirmed cases relationship. In particular, alternative increasing and decreasing variations combined with declining trends of daily confirmed cases with increasing O₃ column concentration are shown in Brazil and the U.K.,

whereas irregular variation followed by significant non-linear increasing patterns of daily confirmed cases are observed in Russia, Spain, Italy, India, and France. In addition, the non-significant variation of daily confirmed cases is shown in Germany.

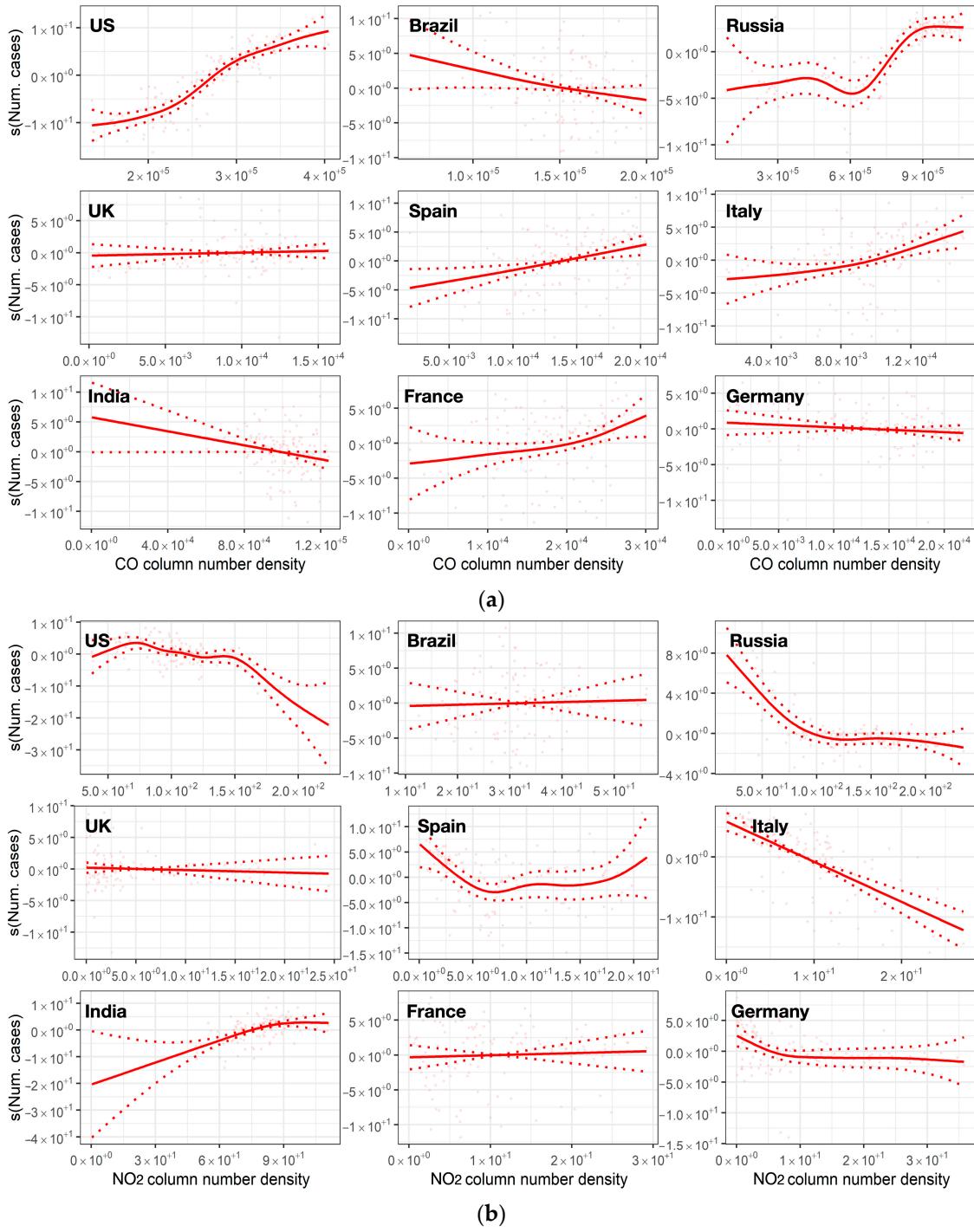


Figure 5. Cont.

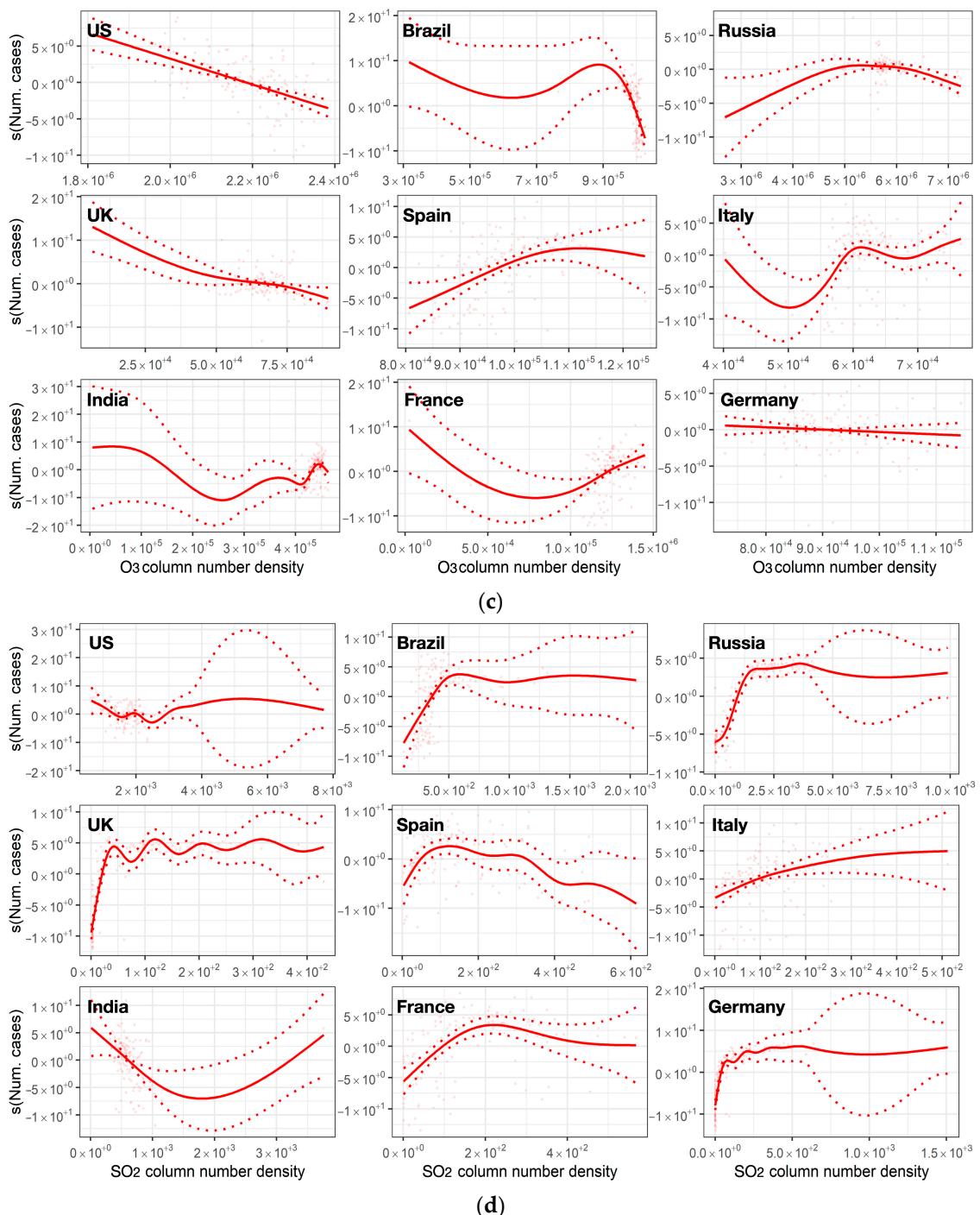


Figure 5. Margin effects (with confidence intervals) of air pollutions on confirmed cases in different countries. (a) CO column concentration; (b) NO₂ column concentration; (c) O₃ column concentration; (d) SO₂ column concentration.

Compared with CO, NO₂, and O₃, similar variation patterns among selected countries are shown in the margin effects of SO₂ column concentration on confirmed cases. As shown in Figure 5d, the U.K. presents the trend of dramatically increasing cases followed by relatively unvarying tendencies in SO₂-confirmed cases relationship. The margin effects are showing stable raising trends in Brazil, Russia, Italy, France, and Germany. In particular, only slight changes are observed in the daily confirmed cases with the increasing SO₂ column concentration in the U.S. The daily confirmed cases are rising and then declining at a relatively small scale with higher SO₂ column concentration in Spain, whereas opposite

patterns are revealed in India with significant declining variation, followed by a trend of growing daily confirmed cases.

Based on CO, NO₂, O₃, and SO₂ column concentrations, the daily confirmed cases in each selected country between 23 January 2020 and 31 May 2020 are predicted through the generalized additive model. Figure 6 displays the comparison between the observed and the predicted numbers of daily confirmed cases, and the period can be divided into two phrases, that is, non-case period and emerging-cases period, respectively. Among the selected countries, the observed and predicted number of daily confirmed cases are generally consistent between the non-case period and emerging-cases period in Russia and Germany. However, in other countries, the relatively satisfactory prediction is realized only during the non-case period, while abnormal cases exist in the predicted trends in the emerging-cases period. Considering the visualization of the observed-predicted cases plotting, several extreme abnormal cases in France, India, Italy, Spain, US, and Brazil are removed from Figure 6. It should be noted that the removal of the extreme abnormal cases does not change the overall comparison patterns of the observed and predicted cases.

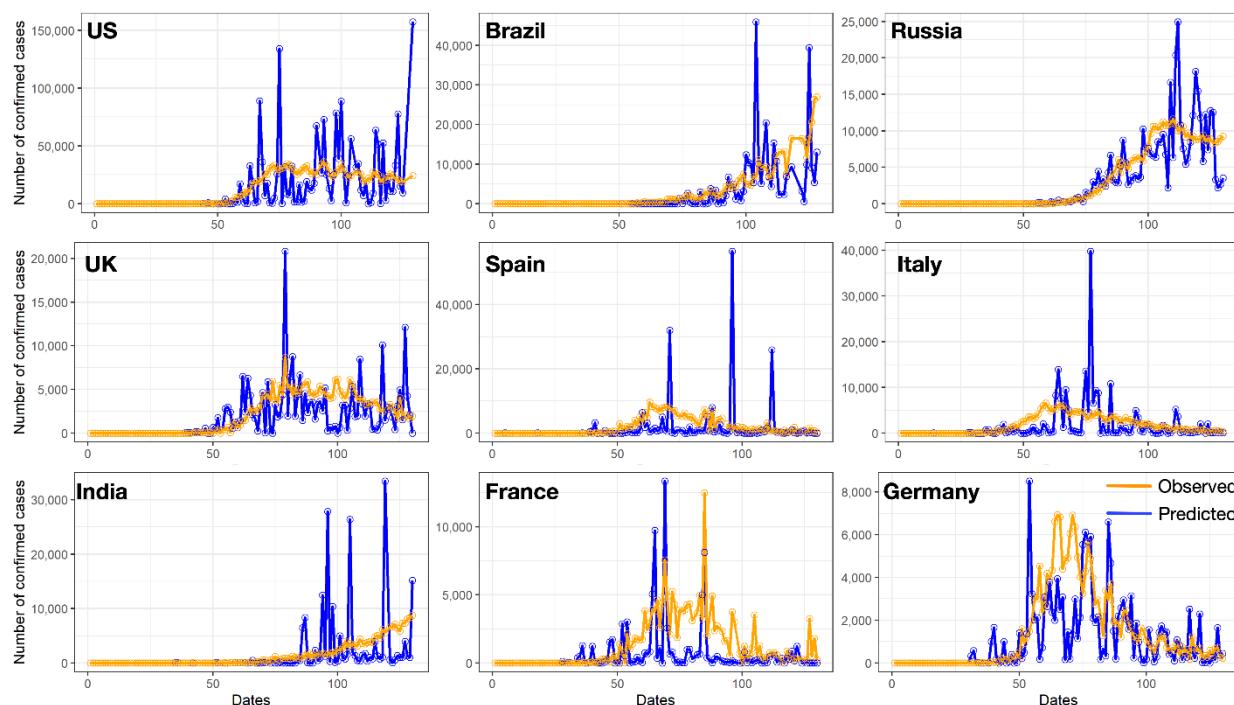


Figure 6. Observed and predicted number of confirmed cases in different countries.

4. Discussion

4.1. Linking Air Pollution Trends and New Confirmed Cases

Despite economic factors among countries not being involved in this study, our findings still provide facts that the variations of yearly and daily CO, NO₂, O₃, and SO₂ are closely associated with confirmed cases of COVID-19 at a global scale. In yearly relationship, a higher degree of CO, NO₂, O₃, and SO₂ changes between 23 January 2019 and 31 May 2019 and between 23 January 2020 and 31 May 2020 are associated with an increased number of total confirmed cases of COVID-19 at a global scale. The correlation between the changes in air pollution of SO₂ and the total confirmed cases is the most prominent, followed by the changes of CO, NO₂, and O₃. The yearly changes in air pollutions linked to COVID-19 can be explained by the lockdown policies which led to the changes in anthropogenic emissions, including traffic, industry, and residential energy uses to some extent [18,19].

For daily variation, we observed a non-linear relationship between the air pollution of CO, NO₂, O₃, and SO₂ and the confirmed cases in most of the countries, which is consistent

with the findings of previous research [35]. A comparison of two regression models (i.e., the LM in Figure 3a and the GAM model in Figure 3b) indicates that the potential of the non-linear model is found in fitting air pollution and daily confirmed cases with higher accuracy. The degree of non-linearity presented in Figure 4 further reveals that countries such as the U.S., Canada, China, and Brazil show much greater non-linearity among one or more types of air pollutions in CO, NO₂, O₃, and SO₂ and daily confirmed cases. The findings indicate that daily air pollution and the number of confirmed cases display a linear relationship and a non-linear relationship in different countries.

The emphasis is also put on the relationship between daily air pollution and the confirmed cases of COVID-19 in nine countries, including the U.S., Brazil, Russia, the U.K., Spain, Italy, India, France, and Germany. Margin effects proposed by GAMs in Figure 5a–d demonstrate that the growing number of daily confirmed cases are linearly or non-linearly correlated with the increasing concentration of CO and SO₂ in most of the selected countries, whereas a trend of non-linear decreasing NO₂ was observed. This may be related to the rapid lockdown as a response to the spread of COVID-19, which leads to the uncertain variation of energy uses of traffic and industry. On the other hand, alternative growing and declining trends of daily confirmed cases are associated with the increasing concentration of O₃, which could be explained by the dilution of economic activities during COVID-19.

The various non-linear patterns between air pollution and confirmed cases of COVID-19 might be explained by the environmental-socioeconomic changes caused by COVID-19 lockdown policies. For instance, the lockdown policies have not only reduced emissions of NO₂ generated by traffic flows, industries, and human mobilities, but also led to the declines in the number of daily confirmed cases due to self-quarantine. Thus, non-linear negative trends between daily NO₂ concentrations and confirmed cases have been observed in most of the countries.

4.2. Limitations and Future Work

Despite the above discussion, limitations still exist in the current research. First, CO, NO₂, O₃, and SO₂ from the satellite data, i.e., TROPOMI were utilized to explore the relationship between air pollution and the confirmed cases of COVID-19. Although TROPOMI shows the advantages in processing air pollution changes at a global scale, the pollutant level cannot be completely reflected compared with the ground-level air pollution [36]. As ground-level fine-scale concentrations cannot be sufficiently obtained at a global scale, the focus can be put on the individual countries in future research by integrating the satellite data from TROPOMI with the ground-level air pollution concentrations. Considering that the satellite data cannot precisely reflect the individual exposure to air pollution, this study mainly focuses on the association patterns between different air pollutants and COVID-19 confirmed cases, instead of assessing the causal relationship of how individual exposures respond to the COVID-19 incidence. Other shortages lies in the lack of particulate matter, which has proved to be a dominant role in affecting respiratory health [37]. Future research will consider the impact of particulate matter on air pollution-COVID-19 incidence responses. In addition, daily relationship analysis requires accurate reported data on COVID-19 cases. Although data from JHU CCSE can provide daily confirmed cases in different countries, the number of reported cases may be underreported in many countries due to insufficient case collection.

Local economic conditions play a crucial role in coronavirus spreading and leading COVID-19 infections. Jahangiri et al. [38] indicated that larger population size can enhance the transmission rate of COVID-19 in Iran. Although a more synthetic relationship could be gained by considering economic factors, findings proposed in this study can still provide insights in delineating the difference of air pollution-confirmed cases relationship among countries at the global scale.

Moreover, limitations are also shown on the spatial and temporal scales of the relationship analysis. As the country-level analysis cannot reflect the heterogeneity of air

pollutants and COVID-19 incidence within countries, it could lead to biases when investigating the association in specific countries. To reduce the biases, this study mainly focuses on the association variation among countries to provide an insight into the air pollution patterns on a global scale. Future research will investigate the air pollutants-COVID-19 cases responses in individual cities to reduce the biases caused by spatial heterogeneity. Regarding to the temporal scale, the selected time period—23 January 2020 to 31 May 2020—in this study is short, and a longer time period covering the year of 2020 is required to investigate the impact of seasonal air pollution changes on COVID-19 incidence. For instance, as the information threshold of O_3 is $180 \mu\text{g}/\text{m}^3$ (hourly average), concentrations below the information threshold, which is usually the case in spring and winter, may be not effectively linked to the individual health responses. However, the selected short time period could reduce the biases caused by the ongoing waves of COVID-19 and the seasonal patterns of air pollution. Thus, this study only focuses on the initial stage of the COVID-19 outbreak to investigate the changes of air pollution driven by the different meteorological and socioeconomic conditions among countries.

In addition, the proposed relationship model allows for the daily air pollution and confirmed cases as variables, while it has not considered the lag effect that confirmed case changes influenced by air pollution might have in the future [39]. Although the impact of historical air pollution trends on the COVID-19 pandemic progress in independent countries is investigated in some research [20,40], the lag effect at the global scale has not been fully discussed. The potentially lagged effects including single-day lag and multiple-day average lag [41] on the relationship between air pollution trends and the confirmed cases of COVID-19 should be studied.

Although the prediction of daily confirmed cases has been applied between 23 January 2020 and 31 May 2020, it is proposed to evaluate the fitness the GAM driven by the air pollution trends. Accurate estimation and forecasting of the daily confirmed cases of COVID-19 by integrating air pollution-driven factors are important. The advantages of mathematical modeling include probing the complexity of infectious diseases such as the Hawkes model and Susceptible-Exposed-Infectious-Removed (SEIR) model [42–44]. Thus, it is possible to understand the epidemiological patterns of COVID-19 by integrating air pollution-driven factors with mathematical models.

The effectiveness of the COVID-19 vaccination and the threats of mutation should also be considered. Vaccination can reduce susceptible social contacts and lower the risk of virus transmission [45]. It is necessary to evaluate the impact of air pollution on the changes of confirmed cases concerning the varied virus transmission rate. Moreover, the virus transmission may also be influenced by the COVID-19 mutation [46,47]. As the gene mutation is associated with the human immune system, biases can be caused in air pollution impact evaluation, regardless of these internal factors.

Despite the discussed uncertainties, the proposed results provide effective insights into the relationship between CO , NO_2 , O_3 , and SO_2 and the daily confirmed cases of COVID-19. Since air pollutants such as NO_2 are significantly influenced by human activities, based on the relationship analysis, some suggestions can be put forth for governments in different countries to monitor air pollution trends and propose effective air pollution control policies.

5. Conclusions

This study investigated the relationship between yearly and daily air pollution and confirmed cases of COVID-19 in 254 countries and regions at the global scale. Results show non-stationary variations between air pollution and confirmed cases of COVID-19. For the yearly relationship, the number of confirmed cases shows positive fluctuation with the increasing CO , O_3 , and SO_2 discrepancies, while a significant peak is observed in NO_2 discrepancies-confirmed cases relationship. For the daily variation among the selected countries, air pollution variations are non-linearly associated with daily conformed cases of COVID-19 in the early stage. Among CO , NO_2 , O_3 , and SO_2 , comparatively growing trends of daily confirmed cases are revealed related to the rising CO column concentration;

a non-linear decreasing pattern of daily confirmed cases is revealed with the increase of NO₂ column concentration in most of the countries; alternative increasing and decreasing changing trends of daily confirmed cases are observed in different countries with regard to the variation of O₃ column concentration; non-linear ascending numbers of confirmed cases are associated with increasing SO₂ column concentration in most of the countries. The findings suggest that the variation of CO, NO₂, O₃, and SO₂ may be significant factors to monitor confirmed cases of COVID-19 at the global scale.

Author Contributions: Conceptualization, Yuan Meng and Man Sing Wong; methodology, Yuan Meng; software, Yuan Meng; validation, Yuan Meng; formal analysis, Yuan Meng and R.Z.; investigation, Yuan Meng; resources, Yuan Meng; data curation, Yuan Meng; writing—original draft preparation, Yuan Meng; writing—review and editing, Yuan Meng, Man Sing Wong, Hanfa Xing, Mei-Po Kwan and Rui Zhu; visualization, Yuan Meng and Rui Zhu; supervision, Man Sing Wong; project administration, Man Sing Wong; funding acquisition, Man Sing Wong. All authors have read and agreed to the published version of the manuscript.

Funding: Man Sing Wong thanks the funding support from a grant by the General Research Fund (Grant no. 15602619), the Collaborative Research Fund (Grant no. C7064-18GF), and the Research Institute for Sustainable Urban Development (Grant no. 1-BBWD), the Hong Kong Polytechnic University. Mei-Po Kwan was supported by grants from the Hong Kong Research Grants Council (General Research Fund Grant no. 14605920; Collaborative Research Fund Grant no. C4023-20GF) and a grant from the Research Committee on Research Sustainability of Major Research Grants Council Funding Schemes of the Chinese University of Hong Kong. Hanfa Xing thanks the funding support from a grant by the National Natural Science Foundation of China (Grant no. 41971406).

Acknowledgments: The authors would like to thank the anonymous reviewers for their constructive comments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gorbatenya, A.E.; Baker, S.C.; Baric, R.S.; Groot, R.J.D.; Ziebuhr, J. The species Severe acute respiratory syndrome-related coronavirus: Classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **2020**, *5*, 536.
2. Li, R.; Pei, S.; Chen, B.; Song, Y.; Zhang, T.; Yang, W.; Shaman, J. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **2020**, *368*, 489–493. [[CrossRef](#)]
3. Gilbert, M.; Pullano, G.; Pinotti, F.; Valdano, E.; Poletto, C.; Boëlle, P.-Y.; D'Ortenzio, E.; Yazdanpanah, Y.; Eholie, S.P.; Altmann, M. Preparedness and vulnerability of African countries against importations of COVID-19: A modelling study. *Lancet* **2020**, *395*, 871–877. [[CrossRef](#)]
4. Wang, C.; Horby, P.W.; Hayden, F.G.; Gao, G.F. A novel coronavirus outbreak of global health concern. *Lancet* **2020**, *395*, 470–473. [[CrossRef](#)]
5. World Health Organization. *Coronavirus Disease 2019 (COVID-19): Situation Report*, 82; World Health Organization: Geneva, Switzerland, 2020.
6. Cohen, J.; Kupferschmidt, K. Strategies shift as coronavirus pandemic looms. *Science* **2020**, *367*, 962–963. [[CrossRef](#)] [[PubMed](#)]
7. Pepe, E.; Bajardi, P.; Gauvin, L.; Privitera, F.; Lake, B.; Cattuto, C.; Tizzoni, M. COVID-19 outbreak response: A first assessment of mobility changes in Italy following national lockdown. *MedRxiv* **2020**.
8. Peto, J.; Alwan, N.A.; Godfrey, K.M.; Burgess, R.A.; Hunter, D.J.; Riboli, E.; Romer, P.; Buchan, I.; Colbourn, T.; Costelloe, C. Universal weekly testing as the UK COVID-19 lockdown exit strategy. *Lancet* **2020**, *395*, 1420–1421. [[CrossRef](#)]
9. Lau, H.; Khosrawipour, V.; Kocabach, P.; Mikolajczyk, A.; Schubert, J.; Bania, J.; Khosrawipour, T. The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. *J. Travel Med.* **2020**, *27*, taaa037. [[CrossRef](#)]
10. Diffenbaugh, N.S.; Field, C.B.; Appel, E.A.; Azevedo, I.L.; Baldocchi, D.D.; Burke, M.; Burney, J.A.; Ciais, P.; Davis, S.J.; Fiore, A.M. The COVID-19 lockdowns: A window into the Earth System. *Nat. Rev. Earth Environ.* **2020**, *1*, 470–481. [[CrossRef](#)]
11. Sharma, S.; Zhang, M.; Gao, J.; Zhang, H.; Kota, S.H. Effect of restricted emissions during COVID-19 on air quality in India. *Sci. Total Environ.* **2020**, *728*, 138878. [[CrossRef](#)]
12. Rao, S.; Klimont, Z.; Smith, S.J.; Van Dingenen, R.; Dentener, F.; Bouwman, L.; Riahi, K.; Amann, M.; Bodirsky, B.L.; Van Vuuren, D.P. Future air pollution in the Shared Socio-economic Pathways. *Glob. Environ. Chang.* **2017**, *42*, 346–358. [[CrossRef](#)]
13. Zhang, W.; Wang, F.; Hubacek, K.; Liu, Y.; Wang, J.; Feng, K.; Jiang, L.; Jiang, H.; Zhang, B.; Bi, J. Unequal exchange of air pollution and economic benefits embodied in China's exports. *Environ. Sci. Technol.* **2018**, *52*, 3888–3898. [[CrossRef](#)] [[PubMed](#)]
14. Brunekreef, B.; Holgate, S.T. Air pollution and health. *Lancet* **2002**, *360*, 1233–1242. [[CrossRef](#)]
15. Kampa, M.; Castanas, E. Human health effects of air pollution. *Environ. Pollut.* **2008**, *151*, 362–367. [[CrossRef](#)]

16. Li, L.; Li, Q.; Huang, L.; Wang, Q.; Zhu, A.; Xu, J.; Liu, Z.; Li, H.; Shi, L.; Li, R. Air quality changes during the COVID-19 lockdown over the Yangtze River Delta Region: An insight into the impact of human activity pattern changes on air pollution variation. *Sci. Total Environ.* **2020**, *732*, 139282. [[CrossRef](#)] [[PubMed](#)]
17. Chen, K.; Wang, M.; Huang, C.; Kinney, P.L.; Anastas, P.T. Air pollution reduction and mortality benefit during the COVID-19 outbreak in China. *Lancet Planet. Health* **2020**, *4*, e210–e212. [[CrossRef](#)]
18. He, G.; Pan, Y.; Tanaka, T. The short-term impacts of COVID-19 lockdown on urban air pollution in China. *Nat. Sustain.* **2020**, *3*, 1005–1011. [[CrossRef](#)]
19. Venter, Z.S.; Aunan, K.; Chowdhury, S.; Lelieveld, J. COVID-19 lockdowns cause global air pollution declines. *Proc. Natl. Acad. Sci. USA* **2020**, *117*, 18984–18990. [[CrossRef](#)]
20. Zhu, Y.; Xie, J.; Huang, F.; Cao, L. Association between short-term exposure to air pollution and COVID-19 infection: Evidence from China. *Sci. Total Environ.* **2020**, *727*, 138704. [[CrossRef](#)]
21. Wu, X.; Nethery, R.C.; Sabath, B.M.; Braun, D.; Dominici, F. Exposure to air pollution and COVID-19 mortality in the United States. *MedRxiv* **2020**. [[CrossRef](#)]
22. Veefkind, J.; Aben, I.; McMullan, K.; Förster, H.; De Vries, J.; Otter, G.; Claas, J.; Eskes, H.; De Haan, J.; Kleipool, Q. TROPOMI on the ESA Sentinel-5 Precursor: A GMES mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications. *Remote Sens. Environ.* **2012**, *120*, 70–83. [[CrossRef](#)]
23. Verhoelst, T.; Compernolle, S.; Pinardi, G.; Lambert, J.-C.; Eskes, H.J.; Eichmann, K.-U.; Fjæraa, A.M.; Granville, J.; Niemeijer, S.; Cede, A. Ground-based validation of the Copernicus Sentinel-5p TROPOMI NO₂ measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandoria global networks. *Atmos. Meas. Tech. Discuss.* **2021**, *14*, 481–510. [[CrossRef](#)]
24. Verhoelst, T.; Compernolle, S.; Granville, J.; Keppens, A.; Pinardi, G.; Lambert, J.-C.; Eichmann, K.-U.; Eskes, H.; Niemeijer, S.; Fjæraa, A.M. Quality assessment of two years of Sentinel-5p TROPOMI NO₂ data. In Proceedings of the EGU General Assembly Conference Abstracts, 4–8 May 2020; p. 15036.
25. Hubert, D.; Keppen, A.; Verhoelst, T.; Granville, J.; Lambert, J.-C.; Heue, K.-P.; Pedergnana, M.; Loyola, D.; Eichmann, K.-U.; Weber, M. Ground-based Assessment of the First Year of Sentinel-5p Tropospheric Ozone Data. In Proceedings of the Living Planet Symposium 2019, Milan, Italy, 13–17 May 2019.
26. Bishop, G.; Welch, G. An introduction to the kalman filter. *Proc. Siggraph Course* **2001**, *8*, 41.
27. Dong, E.; Du, H.; Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **2020**, *20*, 533–534. [[CrossRef](#)]
28. Berndt, D.J.; Clifford, J. Using dynamic time warping to find patterns in time series. In Proceedings of the KDD Workshop, Seattle, WA, USA, 31 July–1 August 1994; pp. 359–370.
29. Petitjean, F.; Ketterlin, A.; Gançarski, P. A global averaging method for dynamic time warping, with applications to clustering. *Pattern Recognit.* **2011**, *44*, 678–693. [[CrossRef](#)]
30. Ratanamahatana, C.A.; Keogh, E. Making time-series classification more accurate using learned constraints. In Proceedings of the 2004 SIAM International Conference on Data Mining, Lake Buena Vista, FL, USA, 22–24 April 2004; pp. 11–22.
31. Grüss, A.; Yemane, D.; Fairweather, T. Exploring the spatial distribution patterns of South African Cape hakes using generalised additive models. *Afr. J. Mar. Sci.* **2016**, *38*, 395–409. [[CrossRef](#)]
32. Sagarese, S.R.; Frisk, M.G.; Cerrato, R.M.; Sosebee, K.A.; Musick, J.A.; Rago, P.J. Application of generalized additive models to examine ontogenetic and seasonal distributions of spiny dogfish (*Squalus acanthias*) in the Northeast (US) shelf large marine ecosystem. *Can. J. Fish. Aquat. Sci.* **2014**, *71*, 847–877. [[CrossRef](#)]
33. Wood, S.N. *Generalized Additive Models: An Introduction with R*; CRC Press: Boca Raton, FL, USA, 2017.
34. Zuur, A.; Ieno, E.N.; Walker, N.; Saveliev, A.A.; Smith, G.M. *Mixed Effects Models and Extensions in Ecology with R*; Springer Science & Business Media: Berlin, Germany, 2009.
35. Sarfraz, M.; Shehzad, K.; Farid, A. Gauging the air quality of New York: A non-linear Nexus between COVID-19 and nitrogen dioxide emission. *Air Qual. Atmos. Health* **2020**, *13*, 1135–1145. [[CrossRef](#)] [[PubMed](#)]
36. Bucela, E.; Perring, A.; Cohen, R.; Boersma, K.; Celarier, E.; Gleason, J.; Wenig, M.; Bertram, T.; Wooldridge, P.; Dirksen, R. Comparison of tropospheric NO₂ from in situ aircraft measurements with near-real-time and standard product data from OMI. *J. Geophys. Res. Atmos.* **2008**, *113*, 1–14. [[CrossRef](#)]
37. Burnett, R.; Chen, H.; Szyszkowicz, M.; Fann, N.; Hubbell, B.; Pope, C.A.; Apte, J.S.; Brauer, M.; Cohen, A.; Weichenthal, S. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 9592–9597. [[CrossRef](#)]
38. Jahangiri, M.; Jahangiri, M.; Najafgholipour, M. The sensitivity and specificity analyses of ambient temperature and population size on the transmission rate of the novel coronavirus (COVID-19) in different provinces of Iran. *Sci. Total Environ.* **2020**, *728*, 138872. [[CrossRef](#)]
39. Braga, A.L.F.; Zanobetti, A.; Schwartz, J. The lag structure between particulate air pollution and respiratory and cardiovascular deaths in 10 US cities. *J. Occup. Environ. Med.* **2001**, *43*, 927–933. [[CrossRef](#)]
40. Berman, J.D.; Ebisu, K. Changes in US air pollution during the COVID-19 pandemic. *Sci. Total Environ.* **2020**, *739*, 139864. [[CrossRef](#)]
41. Kan, H.; London, S.J.; Chen, H.; Song, G.; Chen, G.; Jiang, L.; Zhao, N.; Zhang, Y.; Chen, B. Diurnal temperature range and daily mortality in Shanghai, China. *Environ. Res.* **2007**, *103*, 424–431. [[CrossRef](#)] [[PubMed](#)]

42. Heesterbeek, H.; Anderson, R.M.; Andreasen, V.; Bansal, S.; De Angelis, D.; Dye, C.; Eames, K.T.; Edmunds, W.J.; Frost, S.D.; Funk, S. Modeling infectious disease dynamics in the complex landscape of global health. *Science* **2015**, *347*, aaa4339. [[CrossRef](#)] [[PubMed](#)]
43. Daley, D.J.; Vere-Jones, D. *An Introduction to the Theory of Point Processes: Volume II: General Theory and Structure*; Springer: New York, NY, USA, 2008.
44. Yang, Z.; Zeng, Z.; Wang, K.; Wong, S.-S.; Liang, W.; Zanin, M.; Liu, P.; Cao, X.; Gao, Z.; Mai, Z. Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions. *J. Thorac. Dis.* **2020**, *12*, 165. [[CrossRef](#)] [[PubMed](#)]
45. Huang, B.; Wang, J.; Cai, J.; Yao, S.; Chan, P.K.S.; Tam, T.H.-W.; Hong, Y.-Y.; Ruktanonchai, C.W.; Carioli, A.; Floyd, J.R. Integrated vaccination and physical distancing interventions to prevent future COVID-19 waves in Chinese cities. *Nat. Hum. Behav.* **2021**. [[CrossRef](#)]
46. Grubaugh, N.D.; Hanage, W.P.; Rasmussen, A.L. Making sense of mutation: What D614G means for the COVID-19 pandemic remains unclear. *Cell* **2020**, *182*, 794–795. [[CrossRef](#)]
47. Law, S.; Leung, A.W.; Xu, C. COVID-19 mutation in the United Kingdom. *Microbes Infect. Dis.* **2021**, *2*, 187–188.