Article

Using an Eigenvector Spatial Filtering-Based Spatially Varying Coefficient Model to Analyze the Spatial Heterogeneity of COVID-19 and Its Influencing Factors in Mainland China

Meijie Chen ¹, Yumin Chen ^{1,*}, John P. Wilson ² , Huangyuan Tan ¹ and Tianyou Chu ¹

¹ School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China; chen_meijie@whu.edu.cn (M.C.); tanhuangyuan@whu.edu.cn (H.T.); chutianyou@whu.edu.cn (T.C.)

² Spatial Sciences Institute, University of Southern California, Los Angeles, CA 90089-0374, USA; jpwilson@usc.edu

* Correspondence: ymchen@whu.edu.cn; Tel.: +86-27-6877-8386

Abstract: The COVID-19 pandemic has led to many deaths and economic disruptions across the world. Several studies have examined the effect of corresponding health risk factors in different places, but the problem of spatial heterogeneity has not been adequately addressed. The purpose of this paper was to explore how selected health risk factors are related to the pandemic infection rate within different study extents and to reveal the spatial varying characteristics of certain health risk factors. An eigenvector spatial filtering-based spatially varying coefficient model (ESF-SVC) was developed to find out how the influence of selected health risk factors varies across space and time. The ESF-SVC was able to take good control of over-fitting problems compared with ordinary least square (OLS), eigenvector spatial filtering (ESF) and geographically weighted regression (GWR) models, with a higher adjusted R^2 and lower cross validation RMSE. The impact of health risk factors varied as the study extent changed: In Hubei province, only population density and wind speed showed significant spatially constant impact; while in mainland China, other factors including migration score, building density, temperature and altitude showed significant spatially varying impact. The influence of migration score was less contributive and less significant in cities around Wuhan than cities further away, while altitude showed a stronger contribution to the decrease of infection rates in high altitude cities. The temperature showed mixed correlation as time passed, with positive and negative coefficients at 2.42 °C and 8.17 °C, respectively. This study could provide a feasible path to improve the model fit by considering the problem of spatial autocorrelation and heterogeneity that exists in COVID-19 modeling. The yielding ESF-SVC coefficients could also provide an intuitive method for discovering the different impacts of influencing factors across space in large study areas. It is hoped that these findings improve public and governmental awareness of potential health risks and therefore influence epidemic control strategies.

Keywords: COVID-19; spatial heterogeneity; eigenvector spatial filtering; spatially varying coefficients

Citation: Chen, M.; Chen, Y.; Wilson, J.P.; Tan, H.; Chu, T. Using an Eigenvector Spatial Filtering-Based Spatially Varying Coefficient Model to Analyze the Spatial Heterogeneity of COVID-19 and Its Influencing Factors in Mainland China. *ISPRS Int. J. Geo-Inf.* **2022**, *11*, 67. <https://doi.org/10.3390/ijgi11010067>

Academic Editor: Wolfgang Kainz

Received: 30 November 2021

Accepted: 12 January 2022

Published: 15 January 2022

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

Copyright: © 2022 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 COVID-19 pandemic, initially reported in Wuhan, China in December 2019, is incredibly infectious and has had a large impact the world [1]. The World Health Organization (WHO) designated COVID-19 as a global pandemic on 11 March 2020 [2]. The virus has spread rapidly worldwide and confirmed cases have been found in almost all countries. By November 2021, global cumulative cases were above 256 million and deaths were above 5.1 million, and the number keeps rising [3].

Given this background, researchers have conducted a large number and variety of studies (e.g., clinical research, statistical modeling and behavior analysis). Many have focused on trend analysis and time-series prediction [4–8], which could effectively estimate both the outbreak point and turning point of the COVID-19 pandemic as well as help to

evaluate the effectiveness of measures and whether strategies should be strengthened [9–11]. However, these studies mostly have not considered the influence of external risk factors, which are also important in when it comes to epidemic analysis [12].

Researchers have explored the relationship between socioeconomic as well as infrastructural factors and the spread of the COVID-19 virus [13–17]. Some have found that COVID-19 has had a more significant impact on poor areas [18,19] while a few have found that the influence of income has not been significant in certain study areas [20,21]. Factors such as population density, population mobility and accessibility to hospitals have also been considered in relation to the spread of COVID-19 in many studies [22–26]. Additionally, a number of researchers have estimated the association between environmental factors and COVID-19 transmission, specifically the influence of air quality, wind, humidity as well as temperature on COVID-19 cases [27–33]. However, in these studies, the spatial relationships among cities or countries were underestimated. As the spread of infectious diseases is often spatially autocorrelated, traditional nonspatial statistical models are not so suitable because the data violates the independence criterion [12,34–36].

To solve the problem of spatial autocorrelation, many spatial models such as spatial error models (SEM) and spatial lag models (SLM) models have been utilized in the analysis of COVID-19 [20,37–39]. The studies above considered spatial autocorrelation from a global point of view. However, one problem is that spatial epidemic data often exhibit high spatial heterogeneity [34,40]. Global models are not so helpful for examining the effect of risk factors that may vary across space. Some researchers have used the GWR model and its variations [41,42] to consider spatial varying characteristics that lie in the health risks factors of COVID-19 [43–48]. However, there were often multi-collinearity bias in the coefficients, not to mention the problem of choosing the appropriate bandwidth in GWR modeling [49]. Griffith [50] developed an eigenvector spatial filtering-based spatially varying coefficients (ESF-SVC) method to take control of spatial heterogeneity. One benefit of this method is that it decomposes the spatial weights matrix into eigenvectors which contain information of different spatial patterns. Murakami et al. [51,52] developed an ESF-SVC modeling approach by considering random effects and found the ESF-SVC method was better than the GWR approach in terms of model accuracy. Another advantage of the ESF-SVC method is that it can detect whether the coefficients of independent variables need to be spatially varying or constant through the Bayesian information minimization criteria (BIC). The output of the ESF-SVC model includes the estimated coefficients (be it spatially varying or constant) as well as their statistical significance, which could help to detect spatial varying characteristics [53]. However, this method has not been used in the study of COVID-19.

The main research gap in the aforementioned COVID-19 studies is that many could not model spatial heterogeneity in an effective way. Therefore, an ESF-SVC model was constructed to reveal the spatially varying impact of certain socio and environmental factors on the spread of COVID-19. It deposited the spatial relationship into eigenvectors and combined them with selected health risk factors, and was therefore able detect whether the coefficients of health risk factors are spatially varying or constant. The main objectives of this paper were: (1) to explore how selected health risk factors are related the COVID-19 infection rate within different study extents; (2) to find out if the influence of selected health risk factors vary across space and time and how they vary. Considering data availability and rationality, 10 factors, including socio and environmental ones, were used as the initial health risk factors according to the literature reviews mentioned above. Socio factors included population density, human migration, hospital capacity, GDP and building density, while environmental factors included precipitation, wind speed, temperature, average altitude and air pressure. The ESF-SVC model results were compared with those of the OLS, ESF and GWR models, and the results showed that the proposed ESF-SVC was a promising method in the context of COVID-19 health risk modeling and the discovery of spatial varying characteristics. This study hopes to provide not only a feasible path to solve the problem of spatial autocorrelation and spatial heterogeneity in COVID-19

studies but also an intuitive way to discover spatial and temporal patterns that lie in the influencing factors.

2. Materials and Methods

2.1. Study Area

China, which is within the extension of longitude 73°33' E to 135°05' E, latitude 3°51' N to 53°33' N, is a large country with over 1.4 billion citizens. Hubei province, which is within the extension of longitude 108°21' E to 116°01' E, latitude 29°01' N to 33°16' N and has over 59.17 million residents, had a high rate of COVID-19 infections compared to other provinces in China during the early stages of the pandemic. To explore if the characteristic of influencing factors is different as study extent changes, Hubei province and mainland China were both taken as study areas. The experiments were conducted at the spatial resolution of city level (Figure 1). Considering data quality and availability, 17 and 362 cities were selected as study units in Hubei province and mainland China (excluding Hong Kong, Macao, Taiwan, Zhoushan and Sansha), respectively.

Figure 1. Two study areas (mainland China and Hubei province).

2.2. Data Resources and Pre-Processing

Daily confirmed cases of COVID-19 from January to April 2020 were collected. Natural influencing factors included average temperature (TEMP), precipitation (PRCP), wind speed (WDSP), air pressure (PRE) as well as altitude (DEM). Socioeconomic data included gross domestic product (GDP) and population density data (PDEN). Migration score (MS) was calculated using the daily outflow migration data. Infrastructure data, including the coordinates of hospitals and other buildings, was obtained with the Baidu map application programming interface (Baidu map API), a web map service that provides location data, route planning and other products, for the purpose of calculating hospital capacity (HOS) and building density (BD). The corresponding data sources are listed in Table 1.

Table 1. Data resources.

Name	Data Resources
COVID-19 infections	The National Health Commission of the People's Republic of China [54]
TEMP, PRCP, WDSP, PRE	National Centers for Environmental Information (NCEI) [55,56]
DEM	CGIAR Consortium for Spatial Information (CGIAR-CSI) [57,58]
Population, GDP	The statistical yearbook of 2018 [59]
Outflow migration data	The Baidu qianxi platform [60]

To explore the relationship between COVID-19 infection rates and health risk factors, the weekly average COVID-19 infection rate (denoted as IFR) was taken as the dependent variable. As the distribution of the COVID-19 infection rate data was skewed, a Box–Cox transformation method was used [61] to make the data nearly normally distributed and thus able to fit the basic data assumption for regression models [62]. The formula is shown in Equation (1):

$$IFRi = \begin{cases} \left(\frac{1}{7} \sum_{k=1}^7 \frac{confirmed_k}{pop} \right)^{pow} - 1, & (pow \neq 0) \\ \ln \left(\frac{1}{7} \sum_{k=1}^7 \frac{confirmed_k}{pop} \right), & (pow = 0) \end{cases} \quad (1)$$

where $IFRi$ is the transformed average COVID-19 infection rate at week i , $confirmed_k$ is the number of confirmed cases on the k th day of week i , pow is the corresponding Box–Cox transformation parameter estimated by the maximum likelihood estimation method, which not only considers the situation of power transformation but also squared transformation ($pow = 0.5$), log transformation ($pow = 0$) and reciprocal transformation ($pow = -1$). In this study, $pow = -0.3$ in Hubei province and $pow = 0.11$ in mainland China. pop was the number of residents within each city. Histograms before and after data transformation are included in the Supplementary Material (S1). Population density (PDEN) was calculated by population/area (km^2).

As the incubation period of COVID-19 cases generally ranges from 0 to 14 days, for the consideration of time lag effects, meteorological variables in the previous two weeks (week $i - 2$ and week $i - 1$) were averaged in the modeling of average COVID-19 infection rates in week i [63,64]. The meteorological data covered 376 stations with daily records. For cities that have stations within them, TEMP, PRCP and WDSP were averaged by every two weeks. For those cities that did not have stations within them, it was necessary to match the nearby stations within the 50 km buffer and use the inverse distance weighted method to yield average weather information.

The migration data obtained from Baidu qianxi map from 1 January 2020 to 1 February 2020 included two parts: the migration scale and migration proportion. The migration scale vector $M1$ ($1 \times m$) represents the total number of people that moved out of Wuhan to other places within m ($m = 32$) days, while the daily migration proportion ($M2$) is a $m \times n$ ($m = 32$, $n = 17$ in Hubei province; $m = 32$, $n = 362$ in mainland China) vector which shows the daily percentage of Wuhan migrants who moved to the other n cities within m days. The migration score vector MS ($1 \times n$) was calculated using $M1 \times M2$.

The infrastructure data, including location and type of hospitals as well as other buildings, was obtained with the Baidu map API. Other buildings, such as shopping malls, hotels, restaurants, and tourist attractions, also contributed to population gatherings and the spread of the epidemic, so the number of these types of buildings per km^2 in each city was calculated and selected as a risk factor, and was denoted as building density (BD). The hospital capacity (HOS) was calculated by a weighted sum of hospitals at different levels within a city, as is shown in Equation (2):

$$HOS = 0.5 * hos3A + 0.25 * hos3 + 0.15 * hos2 + 0.1 * hos1 \quad (2)$$

where $hos3A$ is the number of tertiary level-A hospitals, $hos3$ is the number of other tertiary hospitals. As tertiary hospitals, especially 3A hospitals, contain the most important medical resources in China, they have a relatively higher weight [65]. $hos2$ is the number of secondary hospitals and $hos1$ is the number of other hospitals.

All 10 risk factors mentioned above were normalized to fall between 0 to 1 for improve comparisons, and the normalization method is shown in Equation (3).

$$X_{normalized} = (X - X_{min}) / (X_{max} - X_{min}) \quad (3)$$

where $X_{normalized}$ is the normalized influencing factor, and X_{max} and X_{min} are the maximum and minimum values of factor X .

Then, Pearson correlation tests and multicollinearity diagnosis using the variance inflation factor (VIF) were applied before modeling. If the VIF value was smaller than 10, the multicollinearity within variables was not serious, otherwise variables were excluded using the stepwise regression method [66].

2.3. Methods

As the confirmed cases in most of cities within mainland China did not increase much from 10 March to 1 June (297 cities had zero new confirmed cases, 57 cities had new confirmed cases of less than 10 over 3 months), this research focused on a study period ranging from 22 January 2020 to 10 March 2020. The experiment was conducted over 7 weeks within two study extents (Hubei province and mainland China). Spatial varying coefficient maps were used for spatial pattern discovery and to determine how the effects of health risk factors vary across cities. The methodology included five main tasks: (1) spatial weights matrix construction; (2) eigenvector extraction; (3) variable selection and ESF-SVC model construction; (4) model assessment and comparison; and (5) spatial pattern discovery.

2.3.1. Spatial Weights Matrix Construction

The spatial weights matrix W used in the study is a K nearest neighbors contiguity matrix, which is a distance-based matrix and can deal with the problem of “isolation” units [67,68]. For each study unit i , first calculate the distance between the centroid of unit i and the other $n - 1$ units ($d_{i1}, d_{i2} \dots d_{i-1}, d_{i+1}, \dots, d_n$) and rank them by distance. Then find the k closest units to i and denote these as neighbors in contrast to the other units. When cities i and j are neighbors, W_{ij} takes on non-zero values (one, for a binary matrix), otherwise zero [69].

Then the spatial weights matrix W mentioned above is centered into a spatial weight matrix C as below in Equation (4):

$$C = \left(I - \frac{11^T}{n} \right) W \left(I - \frac{11^T}{n} \right) \quad (4)$$

where n represents the city polygon units in this study, $n = 17$ in the study of Hubei province, $n = 362$ in the study of mainland China, I is an n -dimension identity matrix and 1 is an n -by-1 vector of ones.

2.3.2. Eigenvector Extraction

The centered spatial weight matrix C was decomposed into eigenvectors and eigenvalues as in Equation (5):

$$C = E \Lambda E^T \quad (5)$$

where $E = (E_1, E_2, E_3, \dots, E_n)$ is the collection of eigenvectors, Λ is an $n \times n$ diagonal matrix of eigenvalues ($\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n$) which defines the corresponding eigenvectors' MC value, λ_1 is the largest eigenvalue and λ_n is the smallest eigenvalue.

To capture the positive spatial autocorrelation, the subset of eigenvectors C was first selected using the criterion $\frac{\lambda_i}{\lambda_1} > 0.25$ for the following model. A stepwise variable selection method was applied in further eigenvector extraction to avoid multicollinearity. The number of selected candidate eigenvectors was 3 in Hubei province and 85 in mainland China.

2.3.3. Variable Selection and ESF-SVC Model Construction

An extended ESF-based SVC model could be shown as:

$$\hat{\mathbf{Y}} \approx \left(\beta_0 \mathbf{1} + \sum_{k=1}^{K_0} E_k \beta_{0,k} \right) + \sum_{p=1}^P \left(\beta_p \mathbf{1} + \sum_{k=1}^{K_p} E_k \beta_{p,k} \right) \cdot \mathbf{X}_p + \varepsilon \quad (6)$$

where $\mathbf{X}_p$ is a $n \times 1$ vector of the p -th health risk factor, K_p is the number of eigenvectors within the selected eigenvector set that combined with health risk factor p , E_k is the k -th eigenvector within the eigenvector set that combined with health risk factor p , $\beta_0, \beta_p, \beta_{0,k}, \beta_{p,k}$ are the regression coefficients that were estimated using the restricted maximum likelihood method, ε represents random disturbance and “ $\cdot$ ” denotes the element-wise product operator. $E_k \beta_{p,k}$ yields the spatially varying coefficients of health risk factor p .

For modeling COVID-19 cases, Equation (6) can be expressed as:

$$\begin{aligned} IFR = & \beta_0^{R-ESF} + \beta_1^{R-ESF} PDEN + \beta_2^{R-ESF} GDP + \beta_3^{R-ESF} MS + \\ & \beta_4^{R-ESF} HOS + \beta_5^{R-ESF} BD + \beta_6^{R-ESF} TEMP + \beta_7^{R-ESF} PRCP \\ & + \beta_8^{R-ESF} PRE + \beta_9^{R-ESF} WDSP + \beta_{10}^{R-ESF} DEM + \varepsilon \end{aligned} \quad (7)$$

where

$$\beta_k^{R-ESF} = \beta_{k0} + \sum_{k=1}^k E \gamma_k$$

The coefficients β_k^{R-ESF} in Equation (7) include two parts: The first part, β_{k0} , represents the spatially constant effect of health risk factors on the COVID-19 infection rate and the second part, $E \gamma_k$, represents the spatial effect on health risk factors and their local influence on the COVID-19 infection rate.

2.3.4. Model Assessment and Comparison

Four models with Gaussian link functions—OLS, GWR, ESF and ESF-SVC—were built and compared with the corresponding formulas of the other three models (these are attached in the Supplementary Material (see S2–S4)). The comparison criteria were specified below.

Firstly, the adjusted R^2 , the AIC and the rooted mean square error (RMSE) was selected as performance criteria to assess model reliability. The adjusted R^2 and RMSE are two conventional criteria used to evaluate model fit in with regard to absolute value and proportion. The Akaike information criterion (AIC) is a criterion used to evaluate model complexity. The joint inclusion of these assessment criteria meant that the model's fitness could be assessed in a more comprehensive way. The corresponding calculation can be found in the Supplementary Material (S5).

Then cross validation method was utilized to further test model stability and fitting accuracy. Models with a better cross validation performance can be viewed as bearing less over-fitting problems [70,71]. Instead of 10-fold cross validation, a more time consuming but unbiased leave-one-out cross validation method (LOOCV) was applied to alleviate the problem of spatial autocorrelation that might exist within the test dataset [72,73]. That is, only one city with its corresponding variables acts as the validation sample and the other cities serve as the training set. The process was repeated until each city was taken as a validation sample and the average RMSE was taken as the assessment criteria.

Finally, the Moran's I value, an index used to detect spatial autocorrelation [74], was also taken as a performance criterion to test if spatial autocorrelation in the residuals was successfully filtered out. The calculation of Moran's I value is shown in Equation (8).

$$\text{Moran's I} = \frac{n \sum_{i=1}^{i=n} \sum_{j=1}^{j=n} C_{ij} (Y_i - \bar{Y}) (Y_j - \bar{Y})}{\sum_{i=1}^{i=n} \sum_{j=1}^{j=n} C_{ij} \sum_{i=1}^{i=n} (Y_i - \bar{Y})} \quad (8)$$

where $n = 17$ in the extent of Hubei province and $n = 362$ in the extent of mainland China. C_{ij} is an entry in the centered spatial weight matrix C , which represents the weight of the spatial relationship between units i and j . Y is the response variable and in this study represents the model residual. $\bar{Y}$ is the mean value of Y . The Moran's I value ranges from -1 to 1 , and if the Moran's I value is high and significant, it suggests that the residuals are spatially autocorrelated, which violates the hypothesis of independence in linear regression.

Generally, the higher the adjusted R^2 , the lower the RMSE, AIC and the RMSE of LOOCV, the smaller and less significant the Moran's I for residuals and the better the model performance.

2.3.5. Spatial Pattern Discovery

The output of the ESF-SVC model contained the coefficients (be they spatially varying or constant) as well as their statistical significance, indicating how and to what extent risk factors affect the COVID-19 infection rate and how they vary across space. For better visualization, the coefficients of corresponding risk factors were mapped through 7 weeks. The quantile method was used to grade the legend, and each risk factor shared the same mapping dimension. Cities whose coefficient significance was less than 0.1 were labeled by scattered points. A total of 28 (four coefficients $\times$ 7 weeks) coefficient maps were displayed. By comparing the coefficients of different risk factors across space in 7 weeks, it was possible to explore how the effects on the COVID-19 infection rates vary in different cities over time. For better visualization, a flowchart of the research method and experimental procedures is shown in Figure 2.

Figure 2. Flowchart of research method and experimental procedures.

3. Results

3.1. Hubei Province

3.1.1. Correlation Analysis and Multicollinearity Diagnosis

Table 2 shows the Pearson correlation coefficients between health risk factors and COVID-19 infection rates in Hubei province. In the first 3 weeks, from 22 January to 11 February, all 10 risk factors did not show significant correlation with the COVID-19 infection rate, at 0.1. From week 4 to week 7, PDEN, MS, BD and WDSP began to show significant positive correlation, while PRE and DEM showed significant negative correlation at either 0.01 or 0.05. PRCP only passed the 0.05 significance test from weeks three to four. Other risk factors including HOS, GDP, PRCP and TEMP did not pass the significance test. The Pearson correlation coefficients that passed the significant test were mostly above 0.7 or below -0.55 . To test if there is multicollinearity that existed between significantly correlated variables, the VIF test was conducted. The VIF results are shown in Table 3. As some of the VIF values were larger than 10, with MS and BD being larger than 100, the multicollinearity problem was present. Therefore, a stepwise regression method was applied for variable selection to avoid multicollinearity. Finally, only PDEN and WDSP were maintained, and the other variables were screened out.

Table 2. Pearson correlation coefficients between IFR and risk factors in Hubei province.

Time Period	PDEN	HOS	MS	GDP	BD	PRCP	TEMP	WDSP	PRE	DEM
Week1	0.590	0.636	0.521	0.598	0.593	−0.075	0.513	0.356	0.453	−0.348
Week2	0.529	0.503	0.589	0.517	0.600	0.099	0.226	0.57	−0.383	−0.339
Week3	0.526	0.535	0.510	0.554	0.646	0.752	−0.058	0.475	−0.512	−0.516
Week4	0.847 ***	0.608	0.700 **	0.618	0.753 ***	0.754	0.02	0.587 *	−0.567 *	−0.607 *
Week5	0.865 ***	0.601	0.711 **	0.621	0.770 ***	0.462	−0.204	0.548 **	−0.594 **	−0.609 **
Week6	0.867 ***	0.614	0.715 **	0.625	0.774 ***	0.465	0.375	0.666 **	−0.604 **	−0.608 **
Week7	0.869 ***	0.618	0.719 **	0.629	0.778 ***	0.293	0.331	0.674 **	−0.607 **	−0.606 **

* Significant at 0.1 level; ** Significant at 0.05 level; *** Significant at 0.01 level.

Table 3. Variance inflation factor (VIF) of exploratory variables in Hubei province.

Time Period	PDEN	HOS	MS	GDP	BD	PRCP	TEMP	WDSP	PRE	DEM
Week1	55.270	29.969	101.467	30.543	185.873	1.876	14.756	11.737	18.651	7.879
Week2	39.863	32.478	147.708	21.589	172.429	1.864	4.246	3.970	11.681	20.239
Week3	43.940	28.406	152.646	13.499	198.129	6.236	1.743	3.125	9.395	10.919
Week4	45.979	31.843	155.576	16.063	213.393	6.254	1.919	2.632	10.383	12.529
Week5	44.517	25.721	121.521	17.097	190.554	6.275	1.876	3.350	9.033	10.554
Week6	56.205	25.420	138.165	15.531	223.637	6.681	4.008	3.735	11.868	9.383
Week7	40.733	26.986	134.071	21.618	197.476	3.823	5.437	2.491	15.059	11.191

3.1.2. Model Performance Comparison

The results of different models are shown in Table 4. In terms of model fitness, the ESF-SVC model performed the best, with an average adjusted R^2 of 0.770, 16.31%, 5.48% and 18.83%, which was higher than that of the GWR (0.662), ESF (0.730) and OLS (0.648) models. The ESF-SVC's weekly adjusted R^2 was also the highest from weeks one to five but was slightly smaller than that of the ESF model in weeks six and seven. In terms of model errors, the ESF-SVC model performed the best, with an average RMSE of 0.310 compared with the 0.327, 0.374 and 0.445 of the GWR, ESF and OLS models. But the AIC of the ESF-SVC model was larger than the other models, which was due to the spatially varying coefficients increasing the model's complexity.

Table 4. Model performance comparison in Hubei province.

	OLS			ESF			GWR			ESF-SVC		
	Adj. R ²	AIC	RMSE	Adj. R ²	AIC	RMSE	Adj. R ²	AIC	RMSE	Adj. R ²	AIC	RMSE
Week1	0.488	32.340	0.525	0.605	28.784	0.446	0.493	7.731	0.210	0.707	32.810	0.340
Week2	0.355	40.456	0.667	0.444	38.746	0.598	0.524	15.711	0.647	0.629	39.185	0.433
Week3	0.567	28.739	0.445	0.699	23.284	0.358	0.534	7.482	0.212	0.717	29.573	0.305
Week4	0.735	25.065	0.400	0.818	19.440	0.319	0.734	11.239	0.266	0.829	27.099	0.272
Week5	0.775	22.987	0.376	0.837	18.250	0.308	0.768	10.539	0.263	0.852	25.598	0.258
Week6	0.813	20.081	0.345	0.855	16.485	0.293	0.793	14.704	0.336	0.850	24.637	0.261
Week7	0.801	21.322	0.358	0.854	16.771	0.295	0.785	16.029	0.353	0.804	25.849	0.300
All	0.648	27.284	0.445	0.730	23.109	0.374	0.662	11.919	0.327	0.770	29.250	0.310

The LOOCV results are shown in Table 5. The average LOOCV RMSE for ESF-SVC, ESF, GWR and OLS models was 0.607, 0.675, 0.801 and 0.882, respectively. Therefore, in terms of the robustness of cross validation, the ESF-SVC model performed the best. The ESF model also performed better than GWR model.

Table 5. RMSE of cross validation result (LOOCV) in Hubei province.

	RMSE	OLS	ESF	GWR	ESF-SVC
Week1		2.739	1.648	2.313	1.398
Week2		0.830	0.695	0.793	0.615
Week3		0.534	0.464	0.479	0.405
Week4		0.469	0.405	0.415	0.378
Week5		0.452	0.406	0.430	0.426
Week6		0.425	0.479	0.453	0.446
Week7		0.723	0.646	0.723	0.586
Avg		0.882	0.675	0.801	0.607

The Moran's I values for model residuals are displayed in Table 6. The OLS residuals were significantly spatially autocorrelated during the first three weeks, but they did not show significant spatial autocorrelation in the final weeks. The GWR residuals only showed significant spatial autocorrelation in week two. None of the residuals for the ESF and ESF-SVC models were spatially autocorrelated, suggesting that they could better filter out the influence of spatial autocorrelation.

Table 6. Moran's I values for model residuals in Hubei province.

Moran's I	OLS	ESF	GWR	ESF-SVC
Week1	0.164 *	−0.006	−0.271	−0.264
Week2	0.145 *	−0.070	0.035 *	−0.263
Week3	0.194 **	0.011	−0.087	−0.113
Week4	0.054	−0.143	−0.034	−0.221
Week5	0.018	−0.146	−0.106	−0.266
Week6	−0.015	−0.193	−0.043	−0.245
Week7	−0.038	0.008	−0.048	−0.069

* Significant at 0.1 level; ** Significant at 0.05 level.

To intuitively analyze model fitness, the absolute of model average residuals are visualized in Figure 3. The map legend was graded by quantiles of all four models' residuals. The lighter the color, the smaller the average model error. The average residuals for the ESF-SVC and GWR models remained almost at the same level, with only four and five units in the deepest two colors respectively. Both the ESF-SVC model and GWR model performed better than those of the OLS and ESF models.

Figure 3. Residual maps for model fitted in Hubei province.

3.1.3. ESF-SVC Model Coefficients

The constructed ESF-SVC model coefficients are shown in Table 7. All 10 independent variables were normalized so that the coefficient parameter values reflected how risk factors affect COVID-19 infection rates in Hubei province.

Table 7. ESF-SVC model coefficients in Hubei province.

Time Period	Coefficients	PDEN	HOS	MS	GDP	BD	PRCP	TEMP	WDSP	PRE	DEM
Week1	coefficients Sig ($p < 0.1$)	0/17	0/17	0/17	0/17	0/17	0/17	0/17	0/17	0/17	0/17
Week2	coefficients Sig ($p < 0.1$)	0/17	0/17	0/17	0/17	0/17	0/17	0/17	0/17	0/17	0/17
Week3	coefficients Sig ($p < 0.1$)	0/17	0/17	0/17	0/17	0/17	0/17	0/17	0/17	0/17	0/17
Week4	coefficients Sig ($p < 0.1$)	0.966 17/17	0/17	0/17	0/17	0/17	0/17	0/17	0.857 17/17	0/17	0/17
Week5	coefficients Sig ($p < 0.1$)	0.857 17/17	0/17	0/17	0/17	0/17	0/17	0/17	0.892 17/17	0/17	0/17
Week6	coefficients Sig ($p < 0.1$)	1.077 17/17	0/17	0/17	0/17	0/17	0/17	0/17	1.356 17/17	0/17	0/17
Week7	coefficients Sig ($p < 0.1$)	0.875 17/17	0/17	0/17	0/17	0/17	0/17	0/17	1.582 17/17	0/17	0/17
Avg	coefficients Sig ($p < 0.1$)	0.944 10/17	0/17	0/17	0/17	0/17	0/17	0/17	1.172 10/17	0/17	0/17

The “\” represents factors that did not pass the significant test and were not added for modeling; the “svc/number” represents the average spatially varying coefficients; Sig ($p < 0.1$): number/17 represents the number of cities whose corresponding coefficient passed the significant test at 0.1 (p value < 0.1).

During the first three weeks, none of the 10 risk factors passed the Pearson correlation test and 0.1 significance test, but they did register on the spatial varying intercepts generated by eigenvectors, which represent spatial autocorrelation patterns. Therefore, no ESF-SVC coefficients for health risk factors were obtained. PDEN and WDSP passed the modeling significance test from weeks four to seven. WDSP, with an average coefficient of 1.172, played a more important role in the growth of the COVID-19 infection rate in Hubei province from week four onwards. No spatial varying coefficient maps were produced because the health risk factor coefficients were constant.

3.2. Mainland China

3.2.1. Correlation Analysis and Multicollinearity Diagnosis

Table 8 shows the covariates that were correlated with weekly average COVID-19 infection rates. In mainland China, all but two of the risk factors (PRCP and WDSP) passed the 0.05 significance test in 7 weeks. PRCP only passed the significant test from week three to week six and the coefficients were positive. WDSP only passed the significant test in week one. In general, there were significant positive correlations between PDEN, HOS, MS, GDP, BD, TEMP and IFR, whereas PRE and DEM were negatively correlated with IFR. Table 9 shows the VIF of risk factors. All the VIF values were less than 10, which means the multicollinearity among these factors was weak and could be added for modeling.

Table 8. Pearson correlation coefficients between IFR and risk factors in Mainland China.

Time Period	PDEN	HOS	MS	GDP	BD	PRCP	TEMP	WDSP	PRE	DEM
Week1	0.249 ***	0.216 ***	0.237 ***	0.278 ***	0.234 ***	−0.126 *	0.092 *	−0.130 *	0.045 *	−0.422 **
Week2	0.231 ***	0.161 ***	0.260 ***	0.223 ***	0.209 ***	−0.008	0.304 ***	−0.044	−0.145 **	−0.455 **
Week3	0.236 ***	0.182 ***	0.193 ***	0.232 ***	0.207 ***	0.130 **	0.325 ***	−0.035	−0.279 ***	−0.507 **
Week4	0.234 ***	0.178 ***	0.226 ***	0.229 ***	0.209 ***	0.148 **	0.302 ***	−0.082	−0.293 ***	−0.502 **
Week5	0.234 ***	0.178 ***	0.231 ***	0.229 ***	0.209 ***	0.122 **	0.273 ***	−0.039	−0.309 ***	−0.501 **
Week6	0.234 ***	0.178 ***	0.233 ***	0.229 ***	0.208 ***	0.116 *	0.299 ***	−0.042	−0.309 ***	−0.500 **
Week7	0.233 ***	0.178 ***	0.234 ***	0.229 ***	0.208 ***	0.078	0.261 ***	−0.096	−0.317 ***	−0.499 **

* Significant at 0.1 level; ** Significant at 0.05 level; *** Significant at 0.01 level.

Table 9. Variance Inflation Factor (VIF) of exploratory variables in mainland China.

Time Period	PDEN	HOS	MS	GDP	BD	PRCP	TEMP	WDSP	PRE	DEM
Week1	6.403	4.712	1.050	6.381	4.961	1.067	1.150	1.021	1.107	1.245
Week2	6.555	4.736	1.045	6.356	5.082	1.092	1.134	1.030	1.196	1.251
Week3	6.842	4.787	1.047	6.389	5.223	1.148	1.278	1.030	1.487	1.571
Week4	6.769	4.793	1.050	6.361	5.180	1.123	1.245	1.039	1.492	1.547
Week5	6.980	4.795	1.044	6.334	5.387	1.242	1.289	1.032	1.499	1.545
Week6	7.137	4.788	1.045	6.306	5.496	1.218	1.295	1.034	1.508	1.555
Week7	6.772	4.787	1.050	6.324	5.154	1.120	1.236	1.028	1.504	1.535

3.2.2. Model Performance Comparison

Table 10 shows the results for the four models in mainland China. In terms of model fitness, the ESF-SVC model performed the best, with an average adjusted R^2 of 0.624, which was 10.25%, 19.54% and 105.94% higher than that of GWR, ESF and OLS models, respectively. In terms of the AIC, the results produced with the ESF-SVC model were not always better than those produced with the other models, since the spatially varying coefficients increase the model's complexity. In weeks five and six, the ESF-SVC model's RMSE was smaller than that of the other three models. However, in other weeks, the RMSE of the GWR model was the lowest.

Table 10. Model performance comparison in mainland China.

	OLS			ESF			GWR			ESF-SVC		
	Adj. R^2	AIC	RMSE	Adj. R^2	AIC	RMSE	Adj. R^2	AIC	RMSE	Adj. R^2	AIC	RMSE
Week1	0.269	1208.950	1.261	0.380	1155.900	1.149	0.425	1020.251	0.871	0.480	1170.977	1.035
Week2	0.323	1266.708	1.365	0.582	1110.828	1.042	0.625	1117.044	0.816	0.667	1162.689	0.907
Week3	0.306	1486.751	1.850	0.522	1367.015	1.500	0.616	1220.195	1.162	0.622	1411.172	1.330
Week4	0.310	1423.892	1.696	0.547	1288.458	1.339	0.623	1150.409	1.050	0.653	1331.335	1.173
Week5	0.305	1418.737	1.684	0.538	1287.841	1.337	0.519	1254.774	1.280	0.651	1322.890	1.163
Week6	0.308	1414.939	1.675	0.541	1283.402	1.329	0.522	1250.099	1.272	0.647	1322.643	1.166
Week7	0.304	1414.346	1.674	0.543	1279.484	1.322	0.626	1123.939	0.993	0.646	1320.905	1.163
All	0.303	1376.332	1.601	0.522	1253.276	1.288	0.566	1162.387	1.063	0.624	1291.802	1.134

Although the average RMSEs of the ESF-SVC model were not always the smallest in mainland China, the ESF-SVC model's cross validation result outperformed the other three models, with average LOOCV RMSEs of 3.067, much lower than that of the GWR (5.732), ESF (3.422) and OLS (5.805) models (Table 11). The ESF-SVC's weekly LOOCV RMSEs were

also the smallest. Hence, the ESF-SVC model once again outperformed the other models in terms of LOOCV.

Table 11. RMSE of cross validation result (LOOCV) in mainland China.

RMSE	OLS	ESF	GWR	ESF-SVC
Week1	3.583	2.689	3.788	2.560
Week2	5.801	3.407	5.327	2.997
Week3	6.476	3.678	7.412	3.267
Week4	6.273	3.581	6.243	3.184
Week5	6.181	3.579	5.500	3.164
Week6	6.154	3.532	5.653	3.139
Week7	6.167	3.486	6.200	3.160
Avg	5.805	3.422	5.732	3.067

The Moran's I values for the model residuals are displayed in Table 12. All of the OLS model residuals were significantly spatially autocorrelated, and some of the ESF (week one) and GWR (weeks five and six) model residuals showed significant spatial autocorrelation as well. None of the Moran's I values for the ESF-SVC model residuals passed the significance test, suggesting that the ESF-SVC model could better filter out the influence of spatial autocorrelation.

Table 12. Moran's I values for model residuals in mainland China.

Moran's I	OLS	ESF	GWR	ESF-SVC
Week1	0.194 ***	0.045 *	−0.030	−0.027
Week2	0.290 ***	−0.003	−0.028	−0.070
Week3	0.246 ***	−0.001	−0.052	−0.077
Week4	0.279 ***	0.002	−0.051	−0.082
Week5	0.285 ***	0.012	0.064 **	−0.079
Week6	0.284 ***	0.018	0.062 **	−0.074
Week7	0.286 ***	0.012	−0.066	−0.070

* Significant at 0.1 level; ** Significant at 0.05 level; *** Significant at 0.01 level.

Figure 4 shows the absolute of average residuals in mainland China. Areas with zero infections were covered by slashes for the purpose of visualizing which model had a better fit on zero infected areas and which had a better fit on infected areas. The map legend was graded by quantiles. The OLS model performed the worst (Figure 4a) and showed large modeling errors in Hubei province and in the west of China. The ESF model performed better but also produced large modeling errors in the west of China (Figure 4b). The mean modeling errors for the GWR and ESF-SVC models were similar across mainland China as a whole. However, the ESF-SVC model (Figure 4d) produced relatively large errors in the northwest (especially cities with few reported infections) while the GWR model (Figure 4c) produced relatively large modeling errors in central China, where cities with relatively higher infection rates are located.

Figure 4. Residual maps for model fitted in Mainland China.

3.2.3. ESF-SVC Model Coefficients

The constructed ESF-SVC model coefficients are shown in Table 13. MS was positively correlated with the COVID-19 infection rate, with an average spatial varying coefficient of 171.81. The weekly average MS coefficient increased at first and reached its peak of 188.48 in week two, then it decreased to 170.41 and flattered around 171 after week 5. BD was also positively and significantly correlated with the COVID-19 infection rate, with all cities passing the significant test after week three with an average spatial varying coefficient of 4.69. DEM was negatively correlated with the COVID-19 infection rate, with an average spatially varying coefficient of -4.02 . Although the total average spatially varying coefficient of TEMP was -0.60 , its effect on the COVID-19 infection rate was more complicated because the coefficients included a mix of positive and negative values: The weekly average coefficient of TEMP was at first positive in weeks one and two but then became negative from week three onwards. PDEN and PRE only passed the model's significant test in the first week and both of their coefficients were constants.

Figure 5 shows the spatially varying coefficient map of four selected risk factors (MS, BD, TEMP and DEM) over the seven weeks. Figure 5 MS (a–g) shows that MS contributed more to COVID-19 in the central and west-north cities than other parts of mainland China. The average MS coefficients in provinces that contain or are near Wuhan ranged from 77.75 to 132.17, with a mean value of 106.16. While the average MS coefficients in provinces that are far away from Wuhan ranged from 204.11 to 227.34, with a mean value of 214.02, which was 101.06% larger than that was in provinces containing or near Wuhan. The cities with non-significant MS coefficients mostly occur near the Chinese border. BD was also positively correlated with the COVID-19 infection rate (Figure 5 BD (a–g)) and the coefficients were significant in the whole study area and relatively larger in cities in southern China than they were in other areas after week two. However, the variation of BD coefficients between cities was small, with the average value ranging from 4.41 to 4.91. The coefficients for temperature (Figure 5 TEMP (a–g)) were mostly positive during the first two weeks, with an average value of 0.555. However, the number of cities with negative coefficients increased, especially in southeast and northeast China as time passed. The DEM coefficients (Figure 5 DEM (a–g)) were mainly negative and cities that passed the significance test and had high model coefficients after week three mostly occurred in Yunan,

Guangxi, Guizhou, Sichuan, Neimeng, Xinjiang, Liaoning and Gansu provinces, with an average altitude of 1299.7 m.

Figure 5. ESF-SVC spatially varying coefficient maps of health risk factors in mainland China, each column represents a risk factor (MS, BD, TEMP, DEM) and each line (a–g) represents a week.

Table 13. ESF-SVC model coefficients Mainland China.

Time Period	Coefficients	PDEN	HOS	MS	GDP	BD	PRCP	TEMP	WDSP	PRE	DEM
Week1	coefficients Sig ($p < 0.1$)	3.860 14/362	0/362	SVC/148.39 211/362	0/362	SVC/5.76 265/362	0/362	SVC/0.58 112/362	0/362	0.070 11/362	−2.62 362/362
Week2	coefficients Sig ($p < 0.1$)	0/362	0/362	SVC/188.48 219/362	0/362	SVC/4.88 347/362	0/362	SVC/0.53 150/362	0/362	0/362	SVC/−3.89 241/362
Week3	coefficients Sig ($p < 0.1$)	0/362	0/362	SVC/177.35 217/362	0/362	SVC/4.76 362/362	0/362	SVC/−0.82 126/362	0/362	0/362	SVC/−5.04 107/362
Week4	coefficients Sig ($p < 0.1$)	0/362	0/362	SVC/176.16 208/362	0/362	SVC/5.06 362/362	0/362	SVC/−0.31 125/362	0/362	0/362	SVC/−4.32 94/362
Week5	coefficients Sig ($p < 0.1$)	0/362	0/362	SVC/170.42 207/362	0/362	SVC/4.83 362/362	0/362	SVC/−0.31 136/362	0/362	0/362	SVC/−4.11 89/362
Week6	coefficients Sig ($p < 0.1$)	0/362	0/362	SVC/171.13 201/362	0/362	SVC/4.79 362/362	0/362	SVC/−0.13 193/362	0/362	0/362	SVC/−3.96 83/362
Week7	coefficients Sig ($p < 0.1$)	0/362	0/362	SVC/170.72 201/362	0/362	SVC/4.59 362/362	0/362	SVC/−0.14 141/362	0/362	0/362	SVC/−4.22 97/362
Avg	coefficients Sig ($p < 0.1$)	3.860 2/362	0/362	SVC/171.81 208/362	0/362	SVC/4.96 346/362	0/362	SVC/−0.60 138/362	0/362	0.070 2/362	SVC/−4.02 153/362

The “\” represents factors that did not pass significant test and were not added for modeling; the “svc/number” represents the average spatially varying coefficients; Sig ($p < 0.1$): number/326 the represent the number of cities whose corresponding coefficient passed the significant test at 0.1 level (p value < 0.1)

4. Discussion

4.1. Improving Model Accuracy

According to the model assessment results, the ESF-SVC model performed better than the other three models in modeling COVID-19 infection rates.

In Hubei province, the average adjusted R^2 of the ESF-SVC model was 16.31%, 5.48% and 18.83% higher than that of GWR, ESF, and OLS models, respectively. When the study area expanded to mainland China, the average adjusted R^2 of the ESF-SVC model was 10.25%, 19.54% and 105.94% higher than that of GWR, ESF and OLS models, respectively.

The average RMSE value of the ESF-SVC model was much smaller than in the ESF and OLS models. Although its RMSE was slightly larger than that of the GWR in mainland China, the LOOCV results for the ESF-SVC model were the smallest. This suggests that the ESF-SVC model can better estimate the relationship between COVID-19 infection rates and health risk factors in different areas and took good control of the over-fitting problems that plagued the GWR model, thereby providing a more robust model [49,51,75].

The average residual maps showed that the ESF-SVC model generated relatively large modeling errors in northwest China, whereas the GWR model produced large errors in central China. In particular, the ESF-SVC model outperformed the other three models when only the infected areas were modeled. None of the MC generated with the ESF-SVC model residuals were significant, indicating that the ESF-SVC model could better filter out the influence of spatial autocorrelation across large areas.

4.2. Influence of Health Risk Factors

The spatial varying coefficient values of the ESF-SVC model reflect how corresponding risk factors possibly affected COVID-19 infection rates. If a risk factor passed the significance test during the modeling, the larger the absolute coefficient and the greater its contribution to the COVID-19 infection rate.

In Hubei province, all 10 risk factors did not show significant correlation with the COVID-19 infection rate in the first three weeks. This might be because the nucleic acid detection method was not mature and detection resources were not sufficient in high-risk areas at first, such that many patients were not detected as infected very quickly in some places (i.e., Hubei province). The Chinese Healthcare Commission announced an improved method for detecting confirmed cases on 12 February, which was when week 4 in this study began. In the following weeks, PDEN and WDSP passed the final significance test and were selected for modeling. WDSP, with a larger average coefficient than that of PDEN, contributed more to the increase in the infection rate after 12 February. Similar results were also found by Şahin [32]. This might be because the virus could spread more easily under

low or moderate wind speed situations [76,77]. However, as the study area expanded to mainland China, WDSP did not show a significant impact on the increase of COVID-19, indicating that the influence of wind speed is more likely to act on small but high-risk areas. Previous studies also observed no significant association between wind speed and COVID-19 infection rates when the study extent was at the country level [78,79].

In mainland China, eight of the ten risk factors showed significant correlation with the COVID-19 infection rate, and six passed the model significance test (PDEN, MS, BD, TEMP, PRES and DEM) and were selected to develop the final models. MS had the largest influence on the increase of COVID-19 infection rates. It was found that the migration out of Wuhan was less contributive and less significant in cities near Wuhan than cities further away. This might be because there was intense interaction between Wuhan and nearby cities, be it by private car or other travel methods that were not recorded in the Baidu qianxi platform. When traveling to cities that are further away, people are more likely to choose public transportation, so the corresponding migration score is more associated with infection rates. The influence of MS on COVID-19 infection rates reached a peak in week two (around 29 January 2020) and then decreased. This may be explained by the city lockdown policy announced on 23 January 2020, around three weeks after the first reported case of COVID-19 in Wuhan. This policy prevented people from leaving or entering Wuhan by any means of transportation until 8 April 2020, before which there are no increased confirmed cases for a period of 21 days. Public transportation, businesses and entertainment venues within Wuhan were closed to ensure rigorous home quarantine. After the announcement of the city lockdown policy, other cities, especially cities around Wuhan, adjusted their emergency response levels and suggested that people stay home as much as possible and cancel gatherings and events, which largely weakened the virus spread [28,33,80–82]. Building density (BD) had a relatively greater impact on cities in southern China after week three (around 5 February 2020), but the variation of coefficients between cities was small. This indicates that the clustering of entertainment venues accelerated the spread of COVID-19, but the size of its effect and the differences among cities shrank if social distancing was strong. In terms of temperature, the coefficients were mostly positive in the first two weeks, with an average temperature of 2.42 °C. As time passed, more cities, especially those in northeast and southwest China (e.g., cities within the Guangdong, Guangxi, Yunnan and Fujian provinces), had negative coefficients that passed the significant test, with an average temperature of 8.17 °C. The findings suggest that when temperature reached a certain point, the increase of temperature might have resulted in a decrease in the COVID-19 infection rate. Similar results about the influence of temperature were also found in some country-specific and worldwide studies [26,83,84]. As for altitude (DEM), cities that passed the significance test and had high model coefficients were mainly found in plateau regions, suggesting that high altitude areas may have less capacity for the virus to survive. Another study also indicated that people living in high altitude areas might have a better tolerance to hypoxia and might be more resistant to the COVID-19 virus [85]. Unlike in the extent of Hubei province, population density only showed a weak influence on COVID-19 infection rates in the first week and its coefficients were constants, indicating that in large study extents, population density could not explain the increase of COVID-19 infection rate very well when compared with migration outflow and the clustering of buildings [21]. Also, social distancing and traveling restrictions helped to shrink the influence of population density on the spread of COVID-19 [86].

4.3. Limitations

Although the constructed ESF-SVC model performed well in modeling spatial heterogeneity with improved model fitness and robustness in the context of the study of COVID-19, this study still has several limitations. First, some risk factors such as age structure, education level, government policy and community response, were not used. Second, although the time lag effect was considered by using the average of variables in the previous weeks, other temporal characteristics that may influence the COVID-19 infection

rate, such as temporal lag effects at the level of individual days should be considered as well [63,64]. Third, the impact of health risk factors may vary within different regions, different study scales and different periods of the COVID-19 wave. The migration score in this study only included the outflow from Wuhan until 23 January 2020 (the day after lockdown) but did not take population outflows from secondary infection sources into account [87]. In addition, as the migration connectivity was one of the key factors in the spread of COVID-19, conventional topology or a distance-based spatial weight matrix in spatial modeling might be insufficient. Therefore, a migration connectivity network matrix between study units should be taken into account.

5. Conclusions

In this paper, an ESF-SVC method was developed to explore how health risk factors influence the COVID-19 infection rate differently across space and time. It could simultaneously consider the influence of spatial autocorrelation and heterogeneity and could better control multicollinearity and over-fitting problems that plague the GWR model, with a higher average adjusted R^2 . Also, the ESF-SVC model's cross validation RMSEs were also largely lower than in the other three models, indicating that it can better estimate the relationship between COVID-19 infection rate and health risk factors within different areas, thereby providing a more robust model.

The ESF-SVC's spatial varying coefficients at different periods could discover spatial and temporal patterns of influencing factors. It was found that the effect of health risk factors was different as the study extents and study period changed. In Hubei province, WDSP contributed more to the increase in the infection rate than the other health risk factors after 12 February. When the study was extended to mainland China, migration score contributed the most to the COVID-19 infection rate, followed by building density, altitude and temperature, and all of them showed significant spatial varying characteristics. Migration score was less contributive and significant in cities near Wuhan than cities further away, while cities with larger building density coefficients were mainly found in southern China. DEM contributed to the decline of the COVID-19 infection rate and its influence became more significant in high altitude cities as time passed. The influence of temperature was at first positively correlated with COVID-19 infection rate, but after 11 February, the increase of TEMP showed was shown to have a weak but significant impact on the decrease in the COVID-19 infection rate in southeast and northwest cities.

These findings about the impact of health risk factors were also partially consistent with previous studies on COVID-19 and other respiratory infectious diseases, which could help increase public and governmental awareness of the potential health risks and therefore influence COVID-19 control strategies. For example, WDSP showed a significant impact in Hubei province; therefore, wearing a mask in this high-risk area while going out would be the preferred response, which was also recommended by other related studies within different study regions. Meanwhile, after around 29 January, the influence of MS and BD in mainland China decreased, suggesting that the lockdown and social distancing policies worked and could be referred to other areas.

As the proposed method was not limited by datasets, it could form a reference for other spatial epidemiology studies. In the future, we plan to add temporal variables and expand the ESF-SVC model into an ESF based spatiotemporal varying coefficient model, thereby exploring the spatiotemporal relationships between COVID-19 and health risk factors (e.g., the time lag effect of factors and the influence of secondary sources of infection). A dynamic migration connectivity weight matrix will be added to enhance the model. We will also consider how to expand the study area to the entire world and explore the spatiotemporal patterns of COVID-19 spread within different countries and continents.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijgi11010067/s1>, S1: descriptive statistics for power transformation; S2: ordinary least squares (OLS) regression model; S3: eigenvector spatial filtering (ESF) regression model; S4: geographical weighted regression (GWR) model; S5: model assessment criteria.

Author Contributions: Conceptualization, Meijie Chen, Yumin Chen and John P. Wilson; methodology, Meijie Chen and Yumin Chen; software, Meijie Chen and Huangyuan Tan; validation, Meijie Chen, Tianyou Chu and Huangyuan Tan; data curation, Meijie Chen, Tianyou Chu and Huangyuan Tan; writing—original draft preparation, Meijie Chen; writing—review and editing, Yumin Chen and John P. Wilson. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Key S&T Special Projects of China [Grant No. 2018YFB0505302]. The authors sincerely acknowledge their financial support for this research.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data and materials used are public accessed as illustrated in the article.

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

Abbreviations

AIC: Akaike information criterion; BD: building density; DEM: altitude; ESF: eigenvector spatial filtering regression; ESF-SVC: eigenvector spatial filtering-based spatially varying coefficient model; GDP: gross domestic product; GWR: geographically weighted regression; HOS: hospital capacity; IFR: infection rate; LOOCV: leave one out cross validation; MS: migration score; OLS: ordinary least square linear regression; PDEN: population density; PRCP: precipitation; PRE: air pressure; RMSE: root mean squared error; SEM: spatial error model; SLM: spatial lag model; TEMP: average temperature; VIF: variance of inflation; WDSP: wind speed.

References

1. He, J.; Chen, G.; Jiang, Y.; Jin, R.; Shortridge, A.; Agusti, S.; He, M.; Wu, J.; Duarte, C.M.; Christakos, G. Comparative infection modeling and control of COVID-19 transmission patterns in China, South Korea, Italy and Iran. *Sci. Total Environ.* **2020**, *747*, 141447. [\[CrossRef\]](#) [\[PubMed\]](#)
2. Cucinotta, D.; Vanelli, M. WHO Declares COVID-19 a Pandemic. *Acta Biomed.* **2020**, *91*, 157–160. [\[CrossRef\]](#)
3. World Health Organization (WHO). Weekly Operational Update on COVID-19. *Emergency Situational Updates*, 23 November 2021; 1–10.
4. Chadsuthi, S.; Modchang, C. Modelling the effectiveness of intervention strategies to control COVID-19 outbreaks and estimating healthcare demand in Germany. *Public Health Pract.* **2021**, *2*, 100121. [\[CrossRef\]](#)
5. Fanelli, D.; Piazza, F. Analysis and forecast of COVID-19 spreading in China, Italy and France. *Chaos Solitons Fractals* **2020**, *134*, 109761. [\[CrossRef\]](#)
6. Reiner, R.; Barber, R.; Collins, J.; Zheng, P.; Adolph, C.; Albright, J.; Antony, C.; Aravkin, A.; Bachmeier, S.; Bang-Jensen, B.; et al. Modeling COVID-19 scenarios for the United States. *Nat. Med.* **2021**, *27*, 94–105. [\[CrossRef\]](#)
7. ArunKumar, K.E.; Kalaga, D.V.; Sai Kumar, C.M.; Chilkoor, G.; Kawaji, M.; Brenza, T.M. Forecasting the dynamics of cumulative COVID-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-Regressive Integrated Moving Average (ARIMA) and Seasonal Auto-Regressive Integrated Moving Averag. *Appl. Soft Comput.* **2021**, *103*, 107161. [\[CrossRef\]](#)
8. Hu, B.; Ning, P.; Qiu, J.; Tao, V.; Devlin, A.T.; Chen, H.; Wang, J.; Lin, H. Modeling the complete spatiotemporal spread of the COVID-19 epidemic in mainland China. *Int. J. Infect. Dis.* **2021**, *110*, 247–257. [\[CrossRef\]](#)
9. ben Khedher, N.; Kolsi, L.; Alsaif, H. A multi-stage SEIR model to predict the potential of a new COVID-19 wave in KSA after lifting all travel restrictions. *Alex. Eng. J.* **2021**, *60*, 3965–3974. [\[CrossRef\]](#)
10. Guleryuz, D. Forecasting outbreak of COVID-19 in Turkey; Comparison of Box–Jenkins, Brown’s exponential smoothing and long short-term memory models. *Process Saf. Environ. Prot.* **2021**, *149*, 927–935. [\[CrossRef\]](#) [\[PubMed\]](#)
11. Desai, P.S. News Sentiment Informed Time-series Analyzing AI (SITALA) to curb the spread of COVID-19 in Houston. *Expert Syst. Appl.* **2021**, *180*, 115104. [\[CrossRef\]](#)
12. Jia, P.; Lakerveld, J.; Wu, J.; Stein, A.; Root, E.; Sabel, C.; Vermeulen, R.; Remais, J.; Chen, X.; Brownson, R.; et al. Top 10 Research Priorities in Spatial Lifecourse Epidemiology. *Environ. Health Perspect.* **2019**, *127*, 74501. [\[CrossRef\]](#)
13. Hawkins, R.B.; Charles, E.J.; Mehaffey, J.H. Socio-economic status and COVID-19-related cases and fatalities. *Public Health* **2020**, *189*, 129–134. [\[CrossRef\]](#)
14. Martins, L.D.; da Silva, I.; Batista, W.V.; de Fátima Andrade, M.; de Freitas, E.D.; Martins, J.A. How socio-economic and atmospheric variables impact COVID-19 and influenza outbreaks in tropical and subtropical regions of Brazil. *Environ. Res.* **2020**, *191*, 110184. [\[CrossRef\]](#) [\[PubMed\]](#)

15. Ma, Y.; Zhao, Y.; Liu, J.; He, X.; Wang, B.; Fu, S.; Yan, J.; Niu, J.; Zhou, J.; Luo, B. Effects of temperature variation and humidity on the death of COVID-19 in Wuhan, China. *Sci. Total Environ.* **2020**, *724*, 138226. [[CrossRef](#)] [[PubMed](#)]
16. Ravindra, K.; Goyal, A.; Mor, S. Does airborne pollen influence COVID-19 outbreak? *Sustain. Cities Soc.* **2021**, *70*, 102887. [[CrossRef](#)] [[PubMed](#)]
17. Damialis, A.; Gilles, S.; Sofiev, M.; Sofieva, V.; Kolek, F.; Bayr, D.; Plaza, M.P.; Leier-Wirtz, V.; Kaschuba, S.; Ziska, L.H.; et al. Higher airborne pollen concentrations correlated with increased SARS-CoV-2 infection rates, as evidenced from 31 countries across the globe. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2019034118. [[CrossRef](#)] [[PubMed](#)]
18. Messner, W. The Institutional and Cultural Context of Cross-National Variation in COVID-19 Outbreaks. *medRxiv* **2020**. [[CrossRef](#)]
19. Rahman, M.; Islam, M.; Shimanto, M.H.; Ferdous, J.; Rahman, A.A.-N.S.; Sagor, P.S.; Chowdhury, T. A global analysis on the effect of temperature, socio-economic and environmental factors on the spread and mortality rate of the COVID-19 pandemic. *Environ. Dev. Sustain.* **2021**, *23*, 9352–9366. [[CrossRef](#)]
20. Andersen, L.M.; Harden, S.R.; Sugg, M.M.; Runkle, J.D.; Lundquist, T.E. Analyzing the spatial determinants of local COVID-19 transmission in the United States. *Sci. Total Environ.* **2021**, *754*, 142396. [[CrossRef](#)] [[PubMed](#)]
21. Kwok, C.Y.T.; Wong, M.S.; Chan, K.L.; Kwan, M.-P.; Nichol, J.E.; Liu, C.H.; Wong, J.Y.H.; Wai, A.K.C.; Chan, L.W.C.; Xu, Y.; et al. Spatial analysis of the impact of urban geometry and socio-demographic characteristics on COVID-19, a study in Hong Kong. *Sci. Total Environ.* **2021**, *764*, 144455. [[CrossRef](#)]
22. Sirkeci, I.; Murat Yüceşahin, M. Coronavirus and migration: Analysis of human mobility and the spread of COVID-19. *Migr. Lett.* **2020**, *17*, 379–398. [[CrossRef](#)]
23. Li, J.; Yuan, P.; Heffernan, J.; Zheng, T.; Ogden, N.; Sander, B.; Li, J.; Li, Q.; Bélair, J.; Kong, J.D.; et al. Fangcang shelter hospitals during the COVID-19 epidemic, Wuhan, China. *Bull. World Health Organ.* **2020**, *98*, 830. [[CrossRef](#)]
24. Rocklöv, J.; Sjödin, H. High population densities catalyse the spread of COVID-19. *J. Travel Med.* **2020**, *27*, taaa038. [[CrossRef](#)]
25. Luo, M.; Qin, S.; Tan, B.; Cai, M.; Yue, Y.; Xiong, Q. Population Mobility and the Transmission Risk of the COVID-19 in Wuhan, China. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 395. [[CrossRef](#)]
26. Wu, X.; Yin, J.; Li, C.; Xiang, H.; Lv, M.; Guo, Z. Natural and human environment interactively drive spread pattern of COVID-19: A city-level modeling study in China. *Sci. Total Environ.* **2020**, *756*, 143343. [[CrossRef](#)]
27. Lim, Y.K.; Kweon, O.J.; Kim, H.R.; Kim, T.-H.; Lee, M.-K. The impact of environmental variables on the spread of COVID-19 in the Republic of Korea. *Sci. Rep.* **2021**, *11*, 5977. [[CrossRef](#)]
28. Bashir, M.F.; MA, B.; Shahzad, L. A brief review of socio-economic and environmental impact of COVID-19. *Air Qual. Atmos. Health* **2020**, *13*, 1403–1409. [[CrossRef](#)]
29. Qi, H.; Xiao, S.; Shi, R.; Ward, M.P.; Chen, Y.; Tu, W.; Su, Q.; Wang, W.; Wang, X.; Zhang, Z. COVID-19 transmission in Mainland China is associated with temperature and humidity: A time-series analysis. *Sci. Total Environ.* **2020**, *728*, 138778. [[CrossRef](#)]
30. 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](#)]
31. Gupta, S.; Raghuwanshi, G.S.; Chanda, A. Effect of weather on COVID-19 spread in the US: A prediction model for India in 2020. *Sci. Total Environ.* **2020**, *728*, 138860. [[CrossRef](#)]
32. Şahin, M. Impact of weather on COVID-19 pandemic in Turkey. *Sci. Total Environ.* **2020**, *728*, 138810. [[CrossRef](#)] [[PubMed](#)]
33. Wang, Q.; Dong, W.; Yang, K.; Ren, Z.; Huang, D.; Zhang, P.; Wang, J. Temporal and spatial analysis of COVID-19 transmission in China and its influencing factors. *Int. J. Infect. Dis.* **2021**, *105*, 675–685. [[CrossRef](#)] [[PubMed](#)]
34. Elliott, P.; Wartenberg, D. Spatial epidemiology: Current approaches and future challenges. *Environ. Health Perspect.* **2004**, *112*, 998–1006. [[CrossRef](#)]
35. Jia, P.; Dong, W.; Yang, S.; Zhan, Z.; Tu, L.; Lai, S. Spatial Lifecourse Epidemiology and Infectious Disease Research. *Trends Parasitol.* **2020**, *36*, 235–238. [[CrossRef](#)] [[PubMed](#)]
36. Tobler, W.R. A Computer Movie Simulating Urban Growth in the Detroit Region. *Econ. Geogr.* **1970**, *46*, 234. [[CrossRef](#)]
37. Huang, Z. Spatiotemporal Evolution Patterns of the COVID-19 Pandemic Using Space-Time Aggregation and Spatial Statistics: A Global Perspective. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 519. [[CrossRef](#)]
38. Sannigrahi, S.; Pilla, F.; Basu, B.; Basu, A.; Mölter, A. Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using spatial regression approach. *Sustain. Cities Soc.* **2020**, *62*, 102418. [[CrossRef](#)]
39. Yu, H.; Li, J.; Bardin, S.; Gu, H.; Fan, C. Spatiotemporal Dynamic of COVID-19 Diffusion in China: A Dynamic Spatial Autoregressive Model Analysis. *ISPRS Int. J. Geo-Inf.* **2021**, *10*, 510. [[CrossRef](#)]
40. Beale, L.; Abellan, J.; Hodgson, S.; Jarup, L. Methodologic Issues and Approaches to Spatial Epidemiology. *Environ. Health Perspect.* **2008**, *116*, 1105–1110. [[CrossRef](#)]
41. Fotheringham, A.S.; Yang, W.; Kang, W. Multiscale Geographically Weighted Regression (MGWR). *Ann. Am. Assoc. Geogr.* **2017**, *107*, 1247–1265. [[CrossRef](#)]
42. Brunson, C.; Fotheringham, A.S.; Charlton, M.E. Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Geogr. Anal.* **1996**, *28*, 281–298. [[CrossRef](#)]
43. Han, Y.; Yang, L.; Jia, K.; Li, J.; Feng, S.; Chen, W.; Zhao, W.; Pereira, P. Spatial distribution characteristics of the COVID-19 pandemic in Beijing and its relationship with environmental factors. *Sci. Total Environ.* **2021**, *761*, 144257. [[CrossRef](#)]
44. Mollalo, A.; Vahedi, B.; Rivera, K.M. GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. *Sci. Total Environ.* **2020**, *728*, 138884. [[CrossRef](#)]

45. Karaye, I.M.; Horney, J.A. The Impact of Social Vulnerability on COVID-19 in the U.S.: An Analysis of Spatially Varying Relationships. *Am. J. Prev. Med.* **2020**, *59*, 317–325. [\[CrossRef\]](#)
46. Snyder, B.F.; Parks, V. Spatial variation in socio-ecological vulnerability to COVID-19 in the contiguous United States. *Health Place* **2020**, *66*, 102471. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Mansour, S.; Al Kindi, A.; Al-Said, A.; Al-Said, A.; Atkinson, P. Sociodemographic determinants of COVID-19 incidence rates in Oman: Geospatial modelling using multiscale geographically weighted regression (MGWR). *Sustain. Cities Soc.* **2021**, *65*, 102627. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Maiti, A.; Zhang, Q.; Sannigrahi, S.; Pramanik, S.; Chakraborti, S.; Cerda, A.; Pilla, F. Exploring spatiotemporal effects of the driving factors on COVID-19 incidences in the contiguous United States. *Sustain. Cities Soc.* **2021**, *68*, 102784. [\[CrossRef\]](#)
49. Wheeler, D.; Tiefelsdorf, M. Multicollinearity and correlation among local regression coefficients in geographically weighted regression. *J. Geogr. Syst.* **2005**, *7*, 161–187. [\[CrossRef\]](#)
50. Griffith, D.A. Spatial-Filtering-Based Contributions to a Critique of Geographically Weighted Regression (GWR). *Environ. Plan. A* **2008**, *40*, 2751–2769. [\[CrossRef\]](#)
51. Murakami, D.; Yoshida, T.; Seya, H.; Griffith, D.A.; Yamagata, Y. A Moran coefficient-based mixed effects approach to investigate spatially varying relationships. *Spat. Stat.* **2017**, *19*, 68–89. [\[CrossRef\]](#)
52. Tan, H.; Chen, Y.; Wilson, J.P.; Zhang, J.; Cao, J.; Chu, T. An eigenvector spatial filtering based spatially varying coefficient model for PM2.5 concentration estimation: A case study in Yangtze River Delta region of China. *Atmos. Environ.* **2020**, *223*, 117205. [\[CrossRef\]](#)
53. Murakami, D. Spatial regression modeling using the spmoran package: Boston housing price data examples. *arXiv* **2017**, arXiv:1703.04467.
54. COVID-19 Epidemic Data in China. Available online: http://www.nhc.gov.cn/xcs/yqtb/list_gzbd.shtml (accessed on 21 June 2020).
55. Smith, A.; Lott, N.; Vose, R. The Integrated Surface Database: Recent Developments and Partnerships. *Bull. Am. Meteorol. Soc.* **2011**, *92*, 704–708. [\[CrossRef\]](#)
56. The Integrated Surface Database. Available online: <https://www.ncei.noaa.gov/products/land-based-station/integrated-surface-database> (accessed on 16 June 2020).
57. Reuter, H.I.; Nelson, A.; Jarvis, A. An evaluation of void-filling interpolation methods for SRTM data. *Int. J. Geogr. Inf. Sci.* **2007**, *21*, 983–1008. [\[CrossRef\]](#)
58. Jarvis, A.; Reuter, H.I.; Nelson, A.E.; Guevara, E. Hole-Filled Seamless SRTM Data V4, International Centre for Tropical Agriculture (CIAT). 2008. Available online: <https://srtm.csi.cgiar.org> (accessed on 20 June 2020).
59. Statical Year Book of 2018. Available online: <https://data.cnki.net/Yearbook> (accessed on 21 June 2020).
60. Baidu Qianxi Platform. Available online: <https://qianxi.baidu.com/> (accessed on 21 June 2020).
61. Box, G.E.P.; Cox, D.R. An Analysis of Transformations. *J. R. Stat. Soc. Ser. B* **1964**, *26*, 211–243. [\[CrossRef\]](#)
62. Griffith, D.A. Spatial Autocorrelation And Eigenfunctions Of The Geographic Weights Matrix Accompanying Geo-Referenced Data. *Can. Geogr./Le Géogr. Can.* **1996**, *40*, 351–367. [\[CrossRef\]](#)
63. Runkle, J.D.; Sugg, M.M.; Leeper, R.D.; Rao, Y.; Matthews, J.L.; Rennie, J.J. Short-term effects of specific humidity and temperature on COVID-19 morbidity in select US cities. *Sci. Total Environ.* **2020**, *740*, 140093. [\[CrossRef\]](#)
64. Ma, Y.; Cheng, B.; Shen, J.; Wang, H.; Feng, F.; Zhang, Y.; Jiao, H. Association between environmental factors and COVID-19 in Shanghai, China. *Environ. Sci. Pollut. Res.* **2021**, *28*, 45087–45095. [\[CrossRef\]](#)
65. Chen, Y.; Wang, B.; Liu, X.; Li, X. Mapping the spatial disparities in urban health care services using taxi trajectories data. *Trans. GIS* **2018**, *22*, 602–615. [\[CrossRef\]](#)
66. Casella, G.; Fienberg, S.; Olkin, I. *Springer Texts in Statistics*; Springer International Publishing: New Yorker, NY, USA, 2006; Volume 102, ISBN 9780387781884.
67. Gerkman, L.M.; Ahlgren, N. Practical Proposals for Specifying k-Nearest Neighbours Weights Matrices. *Spat. Econ. Anal.* **2014**, *9*, 260–283. [\[CrossRef\]](#)
68. Lesage, J.P.; Fischer, M.M. Spatial Growth Regressions: Model Specification, Estimation and Interpretation. *Spat. Econ. Anal.* **2008**, *3*, 275–304. [\[CrossRef\]](#)
69. Rogerson, P. *Statistical Methods for Geography*; SAGE Publications: Thousand Oaks, CA, USA, 2001.
70. Picard, R.R.; Cook, R.D. Cross-Validation of Regression Models. *J. Am. Stat. Assoc.* **1984**, *79*, 575–583. [\[CrossRef\]](#)
71. Ghojogh, B.; Crowley, M. The Theory Behind Overfitting, Cross Validation, Regularization, Bagging, and Boosting: Tutorial. *arXiv* **2019**, arXiv:1905.12787.
72. Refaeilzadeh, P.; Tang, L.; Liu, H. Cross-Validation BT. In *Encyclopedia of Database Systems*; Liu, L., Özsu, M.T., Eds.; Springer: Boston, MA, USA, 2009; pp. 532–538. ISBN 978-0-387-39940-9.
73. Brenning, A. Spatial cross-validation and bootstrap for the assessment of prediction rules in remote sensing: The R package sperrrest. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 5372–5375.
74. Griffith, D.A. The Moran coefficient for non-normal data. *J. Stat. Plan. Inference* **2010**, *140*, 2980–2990. [\[CrossRef\]](#)
75. Helbich, M.; Griffith, D.A. Spatially varying coefficient models in real estate: Eigenvector spatial filtering and alternative approaches. *Comput. Environ. Urban Syst.* **2016**, *57*, 1–11. [\[CrossRef\]](#)

76. Coşkun, H.; Yıldırım, N.; Gündüz, S. The spread of COVID-19 virus through population density and wind in Turkey cities. *Sci. Total Environ.* **2021**, *751*, 141663. [\[CrossRef\]](#)
77. Coccia, M. How do low wind speeds and high levels of air pollution support the spread of COVID-19? *Atmos. Pollut. Res.* **2021**, *12*, 437–445. [\[CrossRef\]](#)
78. Abdel-Aal, M.A.M.; Eltoukhy, A.E.E.; Nabhan, M.A.; AlDurgam, M.M. Impact of climate indicators on the COVID-19 pandemic in Saudi Arabia. *Environ. Sci. Pollut. Res.* **2021**. [\[CrossRef\]](#)
79. Saba, A.I.; Elsheikh, A.H. Forecasting the prevalence of COVID-19 outbreak in Egypt using nonlinear autoregressive artificial neural networks. *Process Saf. Environ. Prot.* **2020**, *141*, 1–8. [\[CrossRef\]](#) [\[PubMed\]](#)
80. Shi, W.; Tong, C.; Zhang, A.; Wang, B.; Shi, Z.; Yao, Y.; Jia, P. An extended Weight Kernel Density Estimation model forecasts COVID-19 onset risk and identifies spatiotemporal variations of lockdown effects in China. *Commun. Biol.* **2021**, *4*, 126. [\[CrossRef\]](#)
81. Lau, H.; Khosrawipour, V.; Kocbach, 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. [\[CrossRef\]](#) [\[PubMed\]](#)
82. Sun, Z.; Zhang, H.; Yang, Y.; Wan, H.; Wang, Y. Impacts of geographic factors and population density on the COVID-19 spreading under the lockdown policies of China. *Sci. Total Environ.* **2020**, *746*, 141347. [\[CrossRef\]](#) [\[PubMed\]](#)
83. Guo, C.; Bo, Y.; Lin, C.; Li, H.B.; Zeng, Y.; Zhang, Y.; Hossain, M.S.; Chan, J.W.M.; Yeung, D.W.; Kwok, K.; et al. Meteorological factors and COVID-19 incidence in 190 countries: An observational study. *Sci. Total Environ.* **2021**, *757*, 143783. [\[CrossRef\]](#)
84. Ujiie, M.; Tsuzuki, S.; Ohmagari, N. Effect of temperature on the infectivity of COVID-19. *Int. J. Infect. Dis.* **2020**, *95*, 301–303. [\[CrossRef\]](#) [\[PubMed\]](#)
85. Zhou, K.; Yang, S.; Jia, P. Towards precision management of cardiovascular patients with COVID-19 to reduce mortality. *Prog. Cardiovasc. Dis.* **2020**, *63*, 529–530. [\[CrossRef\]](#)
86. Yin, H.; Sun, T.; Yao, L.; Jiao, Y.; Ma, L.; Lin, L.; Graff, J.C.; Aleya, L.; Postlethwaite, A.; Gu, W.; et al. Association between population density and infection rate suggests the importance of social distancing and travel restriction in reducing the COVID-19 pandemic. *Environ. Sci. Pollut. Res.* **2021**, *28*, 40424–40430. [\[CrossRef\]](#)
87. Li, T.; Wang, J.; Huang, J.; Yang, W.; Chen, Z. Exploring the dynamic impacts of COVID-19 on intercity travel in China. *J. Transp. Geogr.* **2021**, *95*, 103153. [\[CrossRef\]](#)