Article

Analysis of the Spread of COVID-19 in the USA with a Spatio-Temporal Multivariate Time Series Model

Rongxiang Rui [1] , **Maozai Tian** [2] , **Man-Lai Tang** [3,*] , **George To-Sum Ho** [4] and **Chun-Ho Wu** [4]

1 School of Statistics, Renmin University of China, Beijing 100872, China; raynerrui@ruc.edu.cn
2 College of Medical Engineering and Technology, Xinjiang Medical University, Ürümqi 830011, China; mztian@ruc.edu.cn
3 Department of Mathematics, Statistics and Insurance, Hang Seng University of Hong Kong, Hong Kong, China
4 Department of Supply Chain and Information Management, Hang Seng University of Hong Kong, Hong Kong, China; georgeho@hsu.edu.hk (G.T.-S.H.); jackwu@ieee.org (C.-H.W.)
* Correspondence: mltang@hsu.edu.hk

Abstract: With the rapid spread of the pandemic due to the coronavirus disease 2019 (COVID-19), the virus has already led to considerable mortality and morbidity worldwide, as well as having a severe impact on economic development. In this article, we analyze the state-level correlation between COVID-19 risk and weather/climate factors in the USA. For this purpose, we consider a spatio-temporal multivariate time series model under a hierarchical framework, which is especially suitable for envisioning the virus transmission tendency across a geographic area over time. Briefly, our model decomposes the COVID-19 risk into: (i) an autoregressive component that describes the within-state COVID-19 risk effect; (ii) a spatiotemporal component that describes the across-state COVID-19 risk effect; (iii) an exogenous component that includes other factors (e.g., weather/climate) that could envision future epidemic development risk; and (iv) an endemic component that captures the function of time and other predictors mainly for individual states. Our results indicate that maximum temperature, minimum temperature, humidity, the percentage of cloud coverage, and the columnar density of total atmospheric ozone have a strong association with the COVID-19 pandemic in many states. In particular, the maximum temperature, minimum temperature, and the columnar density of total atmospheric ozone demonstrate statistically significant associations with the tendency of COVID-19 spreading in almost all states. Furthermore, our results from transmission tendency analysis suggest that the community-level transmission has been relatively mitigated in the USA, and the daily confirmed cases within a state are predominated by the earlier daily confirmed cases within that state compared to other factors, which implies that states such as Texas, California, and Florida with a large number of confirmed cases still need strategies like stay-at-home orders to prevent another outbreak.

Keywords: columnar density of total atmospheric ozone; COVID-19; maximum temperature; minimum temperature; spatio-temporal multivariate time-series analysis; USA

1. Introduction

The pandemic due to the coronavirus disease 2019 (COVID-19), caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1], was the most disastrous incident in 2020, causing millions of deaths and resulting in economic activity worldwide falling sharply. According to the latest report of the World Health Organization (WHO), the cumulative cases around the world reached 28,637,952 and the cumulative deaths were 917,417 as of 13 September 2020 (https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200914-weekly-epi-update-5.pdf?sfvrsn=cf929d04_2). Furthermore, the World Bank suggested that most countries would be expected to suffer

from economic recession in 2020 (https://www.worldbank.org/en/news/feature/2020/06/08/the-global-economic-outlook-during-the-covid-19-pandemic-a-changed-world).

Although some countries and regions (e.g., North America, China, and Europe) are actively developing vaccines with some showing encouraging signs [2–4], it is almost impossible to provide sufficient effective vaccines to every person in the next few years. Hence, this pandemic will undeniably last several months or even a few years. As one of the most developed countries in the world, the United States declared a public health emergency on 31 January 2020, and preventive and proactive measures (e.g., suspending the entry and the quarantine of foreign nationals seeking entry) have been taken to control the spread of the virus and treat those affected. However, it has become one of the most severely affected nations as the respective numbers of confirmed cases and deaths approximately account for 1/5 (i.e., 6,386,832 and 191,809) of the whole global cases as of 13 September 2020. Even worse, as Chowell and Mizumoto [5] argued, states and territories with the largest proportions of older populations (such as Florida, Maine, and Puerto Rico) have become the places with the largest number of confirmed cases. The spread of this pandemic in the USA has become a global concern.

Typically, susceptible-infected-recovered (SIR) based models (e.g., SIRD [6]), first proposed by Kermack and McKendrick [7], are widely used models due to their simplicity and good performance. However, these models (such as susceptible-infected-recovered-susceptible (SIRS), susceptible-exposed-infected-recovered (SEIR), and susceptible-exposed-infected-recovered-susceptible (SEIRS)) only take into account the tendency of the related epidemic transmission corresponding to one single region, and other useful information can hardly be uncovered, which includes the impacts derived from the place itself, other areas, and other exogenous variables. Some other models that are utilized to characterize epidemic pervasion are based on time series models. For instance, seasonal autoregressive integrated moving average (SARIMA) models were employed for modeling infectious disease count data in Helfenstein [8] and Trottier et al. [9]. Recently, different time series models (e.g., auto-regressive integrated moving average (ARIMA), the Holt–Winters additive model, and HWAAS) and machine learning approaches (e.g., Prophet, DeepAR, and N-Beats) have been adopted to analyze and compare the prediction accuracy of the percentage of active cases per population based on the COVID-19 data from ten countries with the highest number of confirmed cases as of 4 May 2020 [10].

Held et al. [11] proposed a space-time multivariate time series model (denoted as the HHHmodel) that can be applied to model multiple-unit cases where the "unit" can be different geographical regions, different age groups, or different epidemics caused by different pathogens. Motivated by the HHH model, Paul et al. [12] and Paul and Held [13] developed a spatio-temporal framework to jointly model several epidemics by considering the spatial interaction effect, as well as the time autoregressive effect. Their models have been applied to analyze the transmission of dengue fever in Guandong Province in China in 2014 [14], malaria and cutaneous leishmaniasis analysis in Afghanistan [15], hemorrhagic fever with renal syndrome in Zhejiang Province of China [16], and the effect of containment measures for COVID-19 in Italy [17].

One drawback of the model proposed by Paul et al. [12] and Paul and Held [13] is that it mainly takes care of the connection between the current number of infected cases and the previous numbers of infected cases and the adjacent units/areas, which may ignore other exogenous predictors. As a result, it is limited in its interpretability and applicability for infectious diseases such as COVID-19. In fact, recent studies (e.g., [18,19]) have pointed out that some weather/climate related variables show statistically significant associations with the transmission of COVID-19.

To overcome this limitation for the analysis of state-level time series of the COVID-19 contagion effect, we consider a spatio-temporal framework based on the multivariate time series model proposed by Paul et al. [12] and Paul and Held [13], which however decomposes the COVID-19 risk additively into autoregressive, spatiotemporal, exogenous, and endemic components. Briefly, the autoregressive and spatiotemporal components

Int. J. Environ. Res. Public Health **2021**, 18, 774

3 of 18

respectively describe the within-state and across-state COVID-19 risk effects. The exogenous component includes other factors that could affect future epidemic development risk, while the endemic component captures the function of time and other predictors mainly about individual states.

Briefly, some weather/climate related variables are carefully selected as exogenous factors in our analyses. Indeed, some climate/weather related variables have been shown to be correlated with epidemic transmission in the relevant literature. For instance, in a study of the influence of weather on the foot-and-mouth disease epidemic spread from 1967 to 1968, Hugh-Jones and Wright [20] argued that wind and precipitation played a major role in the spread of the disease, especially wind. According to Tan et al. [21], the environmental temperature can influence the spread of SARS. Qi et al. [18] found that the daily average temperature and daily average relative humidity are significantly negatively associated with the daily confirmed cases of COVID-19 in Hubei, China. Similarly, Tosepu et al. [19] found that the average seasonal temperature was significantly correlated with COVID-19 in Jakarta, Indonesia.

The rest of this article is organized as follows. In Section 2, we elaborate our spatio-temporal multivariate time series model, including the sub-models of each coefficient in the model. In Section 3, we employ our model for analyzing the COVID-19 count data of the USA and show our main findings. In Sections 4 and 5, the discussion and conclusions are presented, respectively. We report the technical materials for parameter estimation in Appendix A.

2. Development of the Model

Models

Let $Y_{r,t}$ denote the number of infected cases in state r at time point t with $r = 1, \cdots, R, t = 1, \cdots, T$. Usually, $Y_{r,t}$ is assumed to follow a Poisson distribution [22–24]. Since the number of infected cases in each state is hardly totally observed (i.e., the existence of heterogeneity for different states), employing the Poisson assumption could underestimate the underlying dispersion. Here, we adopt the negative binomial distribution [11,12,14]. That is, suppose $Y_{r,t}$ follows a conditional negative binomial distribution, i.e., $Y_{r,t}|Y_{\cdot,t-l}, V \sim \text{NegBin}(\mu_{r,t}, \varepsilon_r)$ for $r = 1, \cdots, R, t = 1, \cdots, T$, with conditional mean $\mu_{r,t}$ and conditional variance:

$$\sigma^2_{r,t} = \mu_{r,t}(1 + \varepsilon_r \mu_{r,t}),$$

where $Y_{\cdot,t-l}$ indicates the vector consisting of the number of infected cases of all states at time point $t - l$, l is the time lag term satisfying $l \in \{1, \cdots, T - 1\}$, ε_r is the overdispersion parameter of state r, and V is a random effect vector with $V \sim N(0, \Sigma)$ with $\Sigma = \text{diag}\{\sigma^2_{(\lambda)}, \sigma^2_{(\psi)}, \sigma^2_{(\theta)}, \sigma^2_{(\zeta)}\} \otimes I_{R \times R}$, $\otimes$ being the Kronecker product and $I_{R \times R}$ an $R \times R$ identity matrix. It is easy to see that when ε_r equals zero, the distribution of $Y_{r,t}$ reduces to a Poisson distribution, whereas the larger the value of ε_r, the greater the overdispersion is. Thus, comparing with the Poisson assumption, the negative binomial assumption has wider applicability.

To embed other predictors in the distribution of $Y_{r,t}$, a hierarchical modeling procedure is employed here. In the first layer, the conditional mean $\mu_{r,t}$ is formulated as follows:

$$\mu_{r,t} = \lambda_{r,t} Y_{r,t-l} + \psi_{r,t} \Psi_{r,t-l} + \theta_{r,t} \Theta_{r,t-l} + \zeta_{r,t}, \tag{1}$$

where:

$$\Psi_{r,t-l} = \sum_{j \to r} \omega_{r,j} Y_{j,t-l}, \quad \Theta_{r,t-l} = \sum_j \eta_{r,j} x_{j,k,t-l}, \quad j = 1, \cdots, R, \ k = 1, \cdots, K.$$

Here, $x_{j,k,t}$ denotes the observation at time t of the k-th exogenous factor in state j, which could have an influence on state r. $\eta_{r,j}$ is an indicator with the value being $1/m_j$ if $x_{j,k}$ has influence on state r and zero otherwise, where m_j is the number of factors that have an influence on the number of cases of the j-th state. $j \to r$ indicates that states j and

r are neighbors that share the same border. $\omega_{r,j}$ is an indicator with the value being $1/n_j$ if state r is adjacent to state j and zero otherwise, where n_j is the number of states that have a common border with state j. Other choices of weights (i.e., $\omega_{r,j}$) are also available in [12–14] and the references therein.

According to Giuliani et al. [25], $\lambda_{r,t}Y_{r,t-l}$, $\psi_{r,t}\Psi_{r,t-l}$, and $\zeta_{r,t}$ are respectively called the epidemic-within component, epidemic-between component, and endemic component. In this paper, we adopt the terminology from Giuliani et al. [25], and we further call $\theta_{r,t}\Theta_{r,t-l}$ the epidemic-boosted component. It is noteworthy that $\Psi_{r,t-l}$, which is based on the space-time dimension, mainly includes the interaction information between one state and other states neighboring that state, while $\Theta_{r,t-l}$ contains the correlation information of other exogenous factors between one state and other states neighboring that state. Comparing with the model proposed in Paul et al. [12] and Paul and Held [13], our proposed model in (1) improves the interpretability, as well as the applicability.

In the second layer, for parameters $\lambda_{r,t}, \psi_{r,t}$, and $\zeta_{r,t}$, we adopt the same strategy as given in Paul et al. [12] and Paul and Held [13]; that is, each parameter assumes the following log-linear form:

$$\log(\cdot_{r,t}) = \alpha_r^{(\cdot)} + V_r^{(\cdot)} + \boldsymbol{\beta}^{(\cdot)\top} \boldsymbol{z}_{r,t}^{(\cdot)}, \tag{2}$$

where $\boldsymbol{V}^{(\cdot)}$ is assumed to have a multivariate normal distribution with zero mean and covariance matrix $\sigma_{(\cdot)}^2 \boldsymbol{I}_{R\times R}$, i.e., $\boldsymbol{V}^{(\cdot)} \sim N(0, \sigma_{(\cdot)}^2 \boldsymbol{I}_{R\times R})$. We further discuss the formulation of each parameter as follows.

We first consider autoregressive parameter $\lambda_{r,t}$. As suggested in Paul et al. [12], Cheng et al. [14], Adegboye et al. [15], and Wu et al. [16], $\lambda_{r,t}$ is formulated by the following log-linear form:

$$\log(\lambda_{r,t}) = \alpha^{(\lambda)} + V_r^{(\lambda)}, \tag{3}$$

where $\alpha^{(\lambda)}$ is related to the intercept term and $V_r^{(\lambda)} \sim N(0, \sigma_{(\lambda)}^2)$.

For $\psi_{r,t}$, it satisfies:

$$\log(\psi_{r,t}) = \alpha^{(\psi)} + V_r^{(\psi)} + \beta_1^{(\zeta)} \log(Pu_{r,t}), \tag{4}$$

where $\alpha^{(\lambda)}$ is related to the intercept term and $V_r^{(\lambda)} \sim N(0, \sigma_{(\lambda)}^2)$. For the choice of $Pu_{r,t}$, unlike Paul et al. [12] and Paul and Held [13], Giuliani et al. [25] argued that it should be a variable that reflects the possible heterogeneous influence for different regions, and their choice was the population of a state. However, we believe that people of different ages could be significantly divergent under the consideration of the infection effect from the population. In this regard, we define $Pu_{r,t}$ as the population size of people whose ages are under 65 in state r at time t based on the fact that this group of people is more likely to travel to other places.

For $\theta_{r,t}$, the log-linear formula may not be suitable as the influence of such exogenous variables could have a positive or negative effect on epidemic transmission, i.e., the sign of $\theta_{r,t}$ could be "+" or "−". Here, we suppose that $\theta_{r,t}$ follows a normal distribution with mean $\alpha^{(\theta)}$ and variance $\sigma_{(\theta)}^2$, i.e.,

$$\theta_{r,t} = \alpha^{(\theta)} + V_r^{(\theta)}, \tag{5}$$

where $V_r^{(\theta)} \sim N(0, \sigma_{(\theta)}^2)$.

Giuliani et al. [25] employed a second-order polynomial log-linear regression to evaluate the fluctuation of the number of confirmed cases from the perspective based on the time dimension. This is reasonable when the epidemic is in the early stage (i.e., t is relatively small). However, with the development of the pandemic, the reproduction number will inevitably tend to be small as the population for a specific state is limited.

Thus, we suggest to use the s-shaped growth curve—logistic growth model—which was also employed to study age-specific case-fatality rates of COVID-19 in China and Italy [26]. That is, we consider:

$$\log(\zeta_{r,t}) = \alpha^{(\zeta)} + V_r^{(\zeta)} + \beta_1^{(\zeta)} \log\left(\text{logit}(t)\right) + \beta_2^{(\zeta)} \log(Po_{r,t}), \tag{6}$$

where:

$$\text{logit}(t) = \left(1 + \exp\left\{-(\beta_3^{(\zeta)} + \beta_4^{(\zeta)}t)\right\}\right)^{-1}$$

is a logistic function and $Po_{r,t}$ is defined as the population size of people whose ages are over 65 years old in state r at time t.

3. Results

3.1. Data of Interest

Study area: Here, we consider the 50 states plus Washington, D.C. (DC), for our COVID-19 analyses. However, American Samoa, Guam, the Northern Mariana Islands, the Commonwealth of Puerto Rico, and the Virgin Islands are excluded from our study for simplicity.

COVID-19 data: We obtained the state-level confirmed cases data on COVID-19 in the USA from Kaggle, which are available from https://www.kaggle.com/sudalairajkumar/covid19-in-usa. Here, we are mainly interested in the cumulative positive cases, as this will be used for the calculation of the number of daily increased COVID-19 cases in the USA. On 14 March 2020, the U.S. President held a coronavirus conference, one day after he declared the pandemic a national emergency (https://www.rev.com/blog/transcripts/donald-trump-coronavirus-press-conference-transcript-march-14). For this reason, we consider those data starting from March 15th, 2020.

Weather/climate data: The state-level weather and climate data we use in this paper are openly available from Kaggle, which is fully powered by Dark Sky and can be downloaded from https://www.kaggle.com/eeemonts/weatherclimate-data-covid19?select=csv. The factors included in our analyses are the maximum temperature (MaT), minimum temperature (MiT), humidity (Hu), the probability of precipitation appearance (PA), the percentage of cloud coverage (CC), sea-level air pressure (AP), wind speed (WS), and the columnar density of total atmospheric ozone (CDTAO). Consistent with the COVID-19 data, we only consider weather/climate data starting from March 15th, 2020.
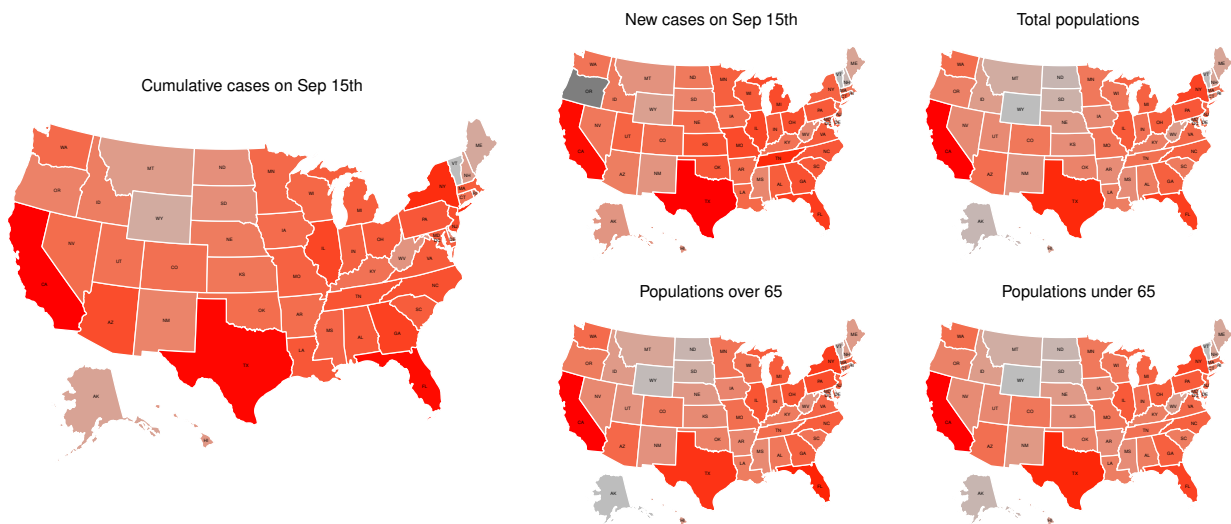
Population data: Both the state-level population data of 2019 and the state-level population percentage of people over 65 years old were collected from Population Reference Bureau (PRB), which is available from (https://www.prb.org/usdata/indicator/age65/snapshot). Since the direct accessibility of the population over 65 years old is denied, we simply used the state-level population of 2019 multiplying the related percentage over 65 years old to get the approximation of the state-level population over 65.

Figure 1 depicts the daily confirmed cases and the cumulative confirmed cases in the USA. One can see that the cumulative confirmed cases and daily confirmed cases in some states such as Connecticut (CT), New Jersey (NJ), and New York (NY) have eased up; some states like North Dakota (ND) became more and more severe; and others tended to recur.

Figure 2 shows state-level subplots of the cumulative confirmed cases on September 15th, the daily confirmed cases on September 15th, the total populations in 2019, the populations over 65, and the populations under 65. It appears that both the cumulative confirmed cases and the daily confirmed cases have a strong consistency with the populations under 65, as well as the populations over 65. The appearance of a high infection rate among individuals under 65 could mean that there is a trend of younger people having severe COVID-19 infections in the USA as warned by Kass et al. [27], which can cause much worse situations (e.g., infecting more older adults) [28].

Int. J. Environ. Res. Public Health **2021**, *18*, 774

6 of 18
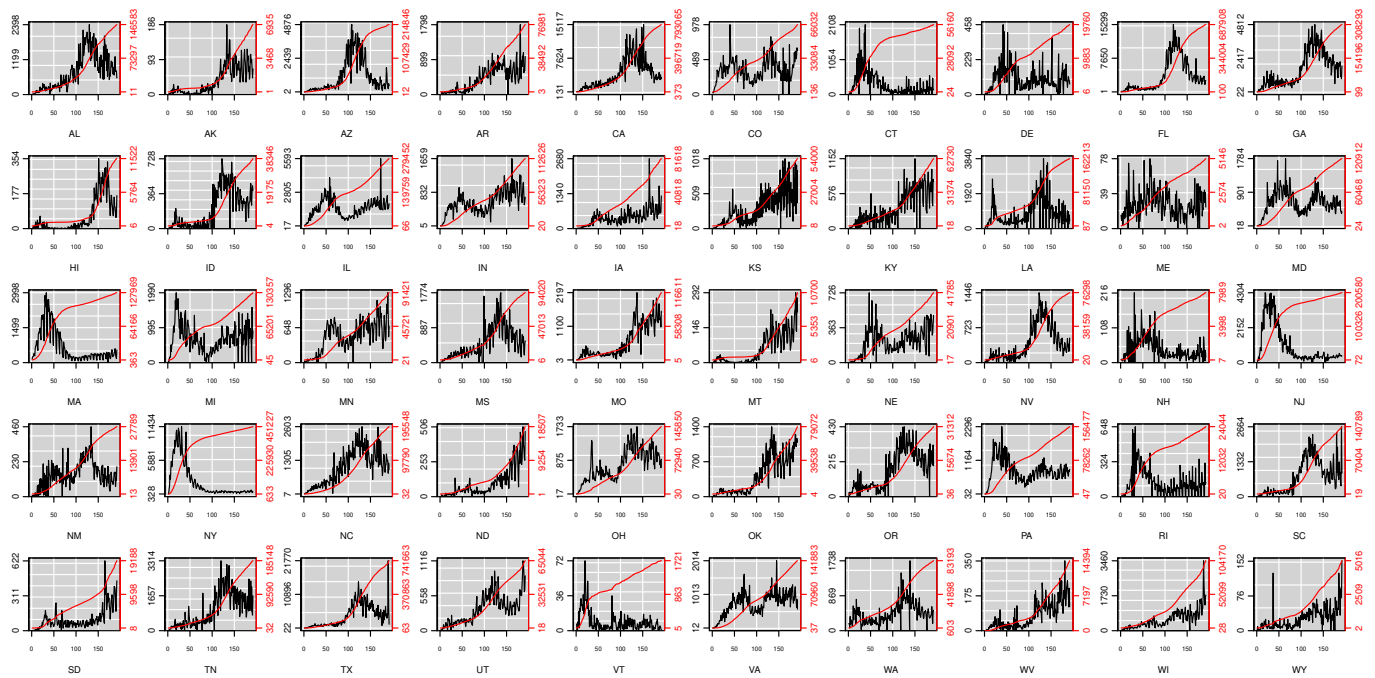
Figure 1. Original data of state-level daily positive test cases in the USA from https://www.kaggle.com/sudalairajkumar/covid19-in-usa (excluding Washington, D.C.). The black line indicates the related daily confirmed cases and the red line the cumulative confirmed cases. The names of all states are denoted by the related postal codes.


Figure 2. State-level population related information. We took the logarithm of the original data for better visualization.

### 3.2. Weather and Environmental Factors' Selection

To explore the association between weather/climate based factors and COVID-19 transmission, Tosepu et al. [19] applied the Spearman-rank correlation test, while Bashir et al. [29] utilized the Kendall and Spearman-rank correlation tests. Here, we first use the Kendall and Spearman-rank correlation tests to identify factors that have a significant correlation with daily confirmed cases.

Figure 3 shows the results from the Spearman and Kendall tests. According to the scatter plots, the *p*-values associated with MaT, MiT, Hu, CC, and CDTAO with respect to all states are mostly below 0.10, whereas those associated with AP and PA are mostly over 0.10. On the other hand, results from bar plots suggest that no particular factor has a remarkable association with the majority of the states. Briefly, AP and PA only have an influence on less than 25 states, while MaT, MiT, and CDTAO have a strong association

Int. J. Environ. Res. Public Health **2021**, *18*, 774

7 of 18

with more than 40 states. According to Figure 4, it can be safely concluded that MaT, MiT, and CDTAO are the major factors that contribute to the strong associations in most of the states and are taken into account for further modeling analyses.
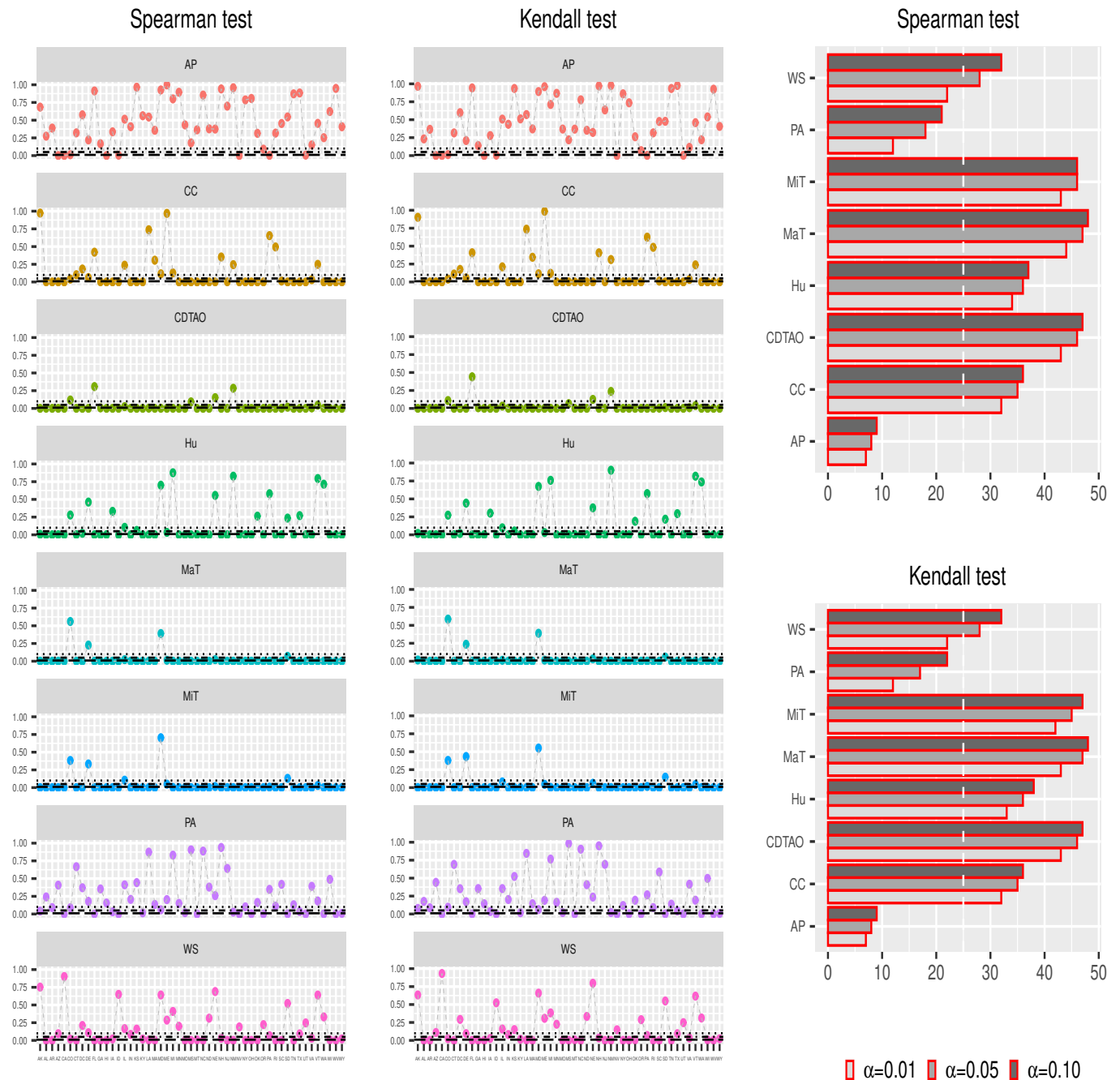
Figure 3. State-level Spearman and Kendall test results. Scatter plots show the related *p*-value, and bar plots show the related number of states that are significantly influenced by related factors with $\alpha = 0.01, 0.05$, and 0.10, respectively, which are drawn with dotted black lines. The white dotted lines in the bar plots are equal to 25, which is approximately half of the total number of analyzed states. Maximum temperature (MaT), minimum temperature (MiT), humidity (Hu), the probability of precipitation appearance (PA), the percentage of cloud coverage (CC), sea-level air pressure (AP), wind speed (WS), and the columnar density of total atmospheric ozone (CDTAO)
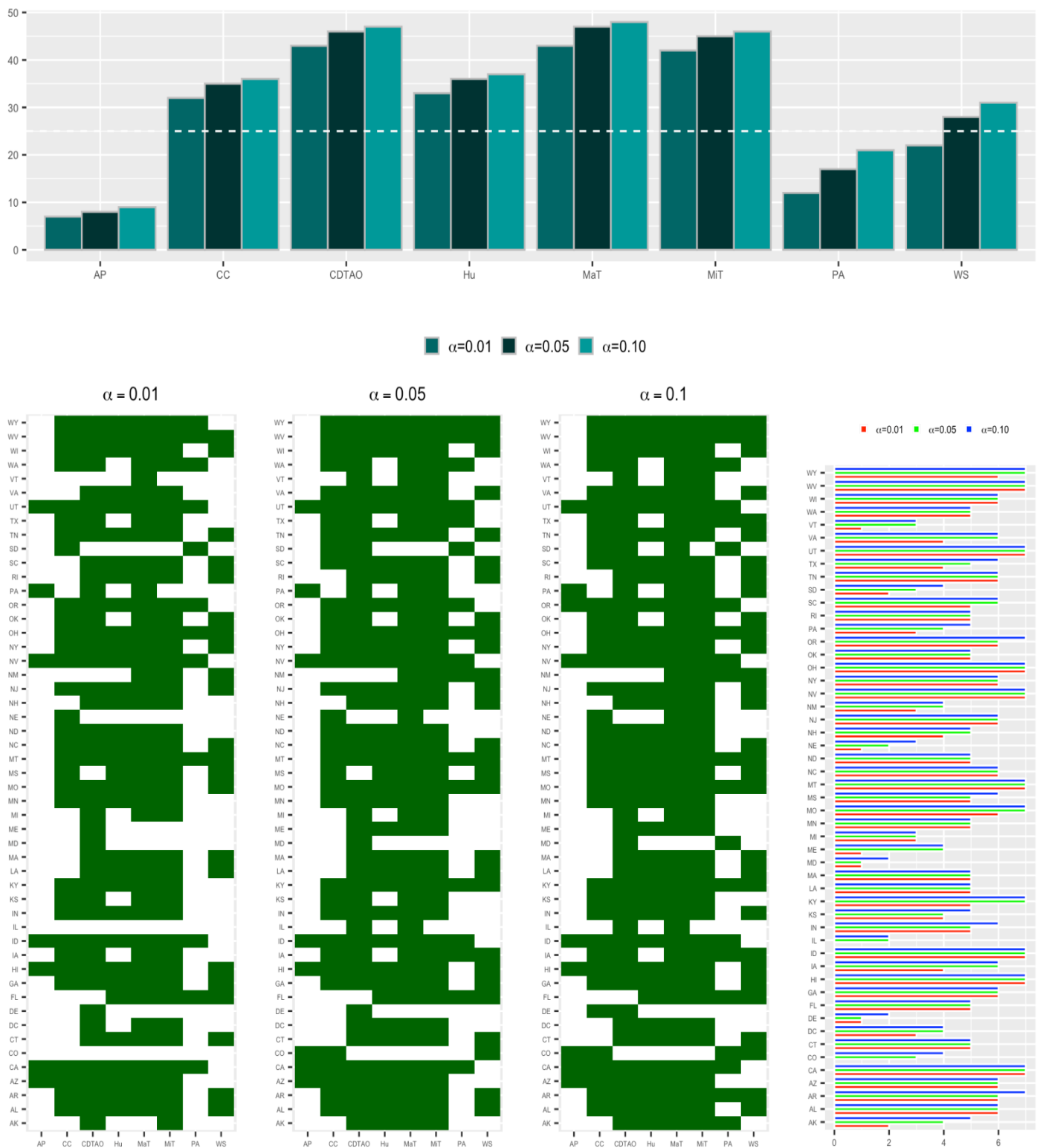
Figure 4. State-level combination of Spearman and Kendall tests. Two-dimensional barcode plots show whether a factor is significantly associated with a state by both the Spearman and Kendall tests (with squares in dark green color). The bar plot represents the number of states that are significantly associated with each factor with $\alpha = 0.01, 0.05$, and 0.10, respectively, by both the Spearman and Kendall tests. The horizontal line in the top subplot is equal to 25, which is approximately half of the total number of analyzed states.

3.3. Optimal Parameters' Determination

For better fitting the confirmed cases in different states, the optimal time lag l needs to be determined. Considering that the desirable range for l is dynamically increased, point-wise optimization seems too tedious and less efficient. Thus, we adopted the possible range from one to 14, which is the time interval that the CDC suggested to stay at home after one's last contact with a person who has COVID-19 (https://www.cdc.gov/coronavirus/2019-ncov/if-you-are-sick/quarantine.html). Besides, all weather/climate factors were respectively nondimensionalized by Studentization to mitigate the impact of the different units.

Table 1 summarizes the correlation between different time lags and the related penalized log-likelihood values. It is noticed that when $l = 2$, the average estimate of $l(\pi, v)$ has the largest value (i.e., 48,510,611.518) with the smallest sd (i.e., 301,043.787). Hence, $l = 2$ will be used in all subsequent analyses.

Table 1. Different time lags and the corresponding penalized log-likelihood values shown with the mean and standard deviation (sd). Thirty different initial processes for each lag are randomly implemented. The corresponding mean and sd are calculated based on 30 repetitions.

Lags	$l = 1$	$l = 2$	$l = 3$	$l = 4$	$l = 5$	$l = 6$	$l = 7$
mean$\left(l(\pi, v)\right)$	47512882.850	48510611.518	48188861.487	47895162.927	47783167.600	47998732.406	47976026.240
sd$\left(l(\pi, v)\right)$	1600987.450	301043.787	923652.391	849380.316	1297418.980	825285.208	1055666.220
lags	$l = 8$	$l = 9$	$l = 10$	$l = 11$	$l = 12$	$l = 13$	$l = 14$
mean$\left(l(\pi, v)\right)$	47747910.670	47946826.823	47493940.570	47630036.889	47306645.479	47309743.565	47115278.420
sd$\left(l(\pi, v)\right)$	1199672.290	643524.434	1288241.920	773951.669	920626.589	673304.655	1158747.060

Table 2 shows the estimates of π and σ based on the configuration for time lag $l = 2$. From the estimates of the dispersion parameters ε's, we observe that the daily confirmed cases of COVID-19 show obvious overdispersion in almost every state (especially in New Mexico (NM)), which confirms that using the negative binomial distribution is a sensible choice for analyzing the transmission of COVID-19 in the USA.

The estimates of σ characterize the heterogeneity of COVID-19 transmission across states. According to the results in Table 2, there is spatial variation concerning the epidemic-within component with $\hat{\sigma}_\lambda = 0.851$, the epidemic-between component with $\hat{\sigma}_{(\psi)} = 0.836$, the epidemic-boosted component with $\hat{\sigma}_{(\theta)} = 0.816$, and the endemic component with $\hat{\sigma}_\zeta = 0.853$. Therefore, we believe that there is significant spatial heterogeneity in the epidemic-within, endemic-between, epidemic-boosted, and endemic component.

Table 2. Optimal parameter estimates with time lag $l = 2$. One-hundred different initial processes are randomly implemented. The related mean and standard deviation (sd) are calculated based on $100\times$ outcomes.

Estimates	$\sigma_{(\lambda)}$	$\sigma_{(\psi)}$	$\sigma_{(\theta)}$	$\sigma_{(\zeta)}$	$\alpha^{(\lambda)}$	$\alpha^{(\psi)}$	$\alpha^{(\theta)}$	$\alpha^{(\zeta)}$
mean	0.851	0.836	0.816	0.853	0.220	−0.136	0.173	−0.233
sd	0.121	0.113	0.107	0.124	0.911	0.929	0.958	0.991
estimates	$\beta_1^{(\psi)}$	$\beta_1^{(\zeta)}$	$\beta_2^{(\zeta)}$	$\beta_3^{(\zeta)}$	$\beta_4^{(\zeta)}$	ε_{AL}	ε_{AK}	ε_{AZ}
mean	−0.300	0.232	−0.064	−0.119	−0.076	0.731	0.774	0.764
sd	0.873	1.068	1.004	1.180	1.085	0.285	0.286	0.269
estimates	ε_{AR}	ε_{CA}	ε_{CO}	ε_{CT}	ε_{DE}	ε_{FL}	ε_{GA}	ε_{HI}
mean	0.739	0.697	0.754	0.776	0.744	0.732	0.730	0.693
sd	0.286	0.296	0.284	0.270	0.285	0.301	0.275	0.297
estimates	ε_{ID}	ε_{IL}	ε_{IN}	ε_{IA}	ε_{KS}	ε_{KY}	ε_{LA}	ε_{ME}
mean	0.707	0.673	0.781	0.715	0.696	0.751	0.816	0.784
sd	0.253	0.278	0.284	0.285	0.292	0.281	0.296	0.290

Table 2. *Cont.*

Estimates	$\sigma_{(\lambda)}$	$\sigma_{(\psi)}$	$\sigma_{(\theta)}$	$\sigma_{(\zeta)}$	$\alpha^{(\lambda)}$	$\alpha^{(\psi)}$	$\alpha^{(\theta)}$	$\alpha^{(\zeta)}$
estimates	ε_{MD}	ε_{MA}	ε_{MI}	ε_{MN}	ε_{MS}	ε_{NJ}	ε_{NM}	ε_{NY}
mean	0.780	0.658	0.740	0.724	0.758	0.708	0.858	0.780
sd	0.301	0.291	0.296	0.289	0.308	0.291	0.273	0.306
estimates	ε_{MO}	ε_{MT}	ε_{NE}	ε_{NV}	ε_{NH}	ε_{NC}	ε_{ND}	ε_{OH}
mean	0.716	0.757	0.732	0.780	0.776	0.793	0.769	0.710
sd	0.299	0.281	0.298	0.303	0.284	0.273	0.284	0.301
estimates	ε_{OK}	ε_{OR}	ε_{PA}	ε_{RI}	ε_{TX}	ε_{SC}	ε_{SD}	ε_{TN}
mean	0.667	0.732	0.742	0.763	0.812	0.765	0.746	0.729
sd	0.294	0.277	0.306	0.283	0.296	0.279	0.273	0.290
estimates	ε_{UT}	ε_{VT}	ε_{VA}	ε_{WA}	ε_{WV}	ε_{WI}	ε_{WY}	ε_{DC}
mean	0.736	0.742	0.687	0.757	0.805	0.779	0.727	0.791
sd	0.291	0.287	0.327	0.283	0.287	0.300	0.289	0.296

3.4. Components Analysis

Figure 5 shows the state-level estimated random effects with respect to epidemic-within component, epidemic-between component, epidemic-boosted component, and endemic component. Clearly, heterogeneity appears in all components, and the random effects from the four components have a significant effect in most of the states. Here, we mainly focus on random effects from the epidemic-within component and epidemic-boosted component. From Subplot (a) in Figure 5, the estimates of $\alpha^{(\lambda)} + V^{(\lambda)}$ for most states are smaller than zero, which implies that community-level spread in most states has considerably alleviated from the perspective of epidemic-with component. Similarly, the estimates of $\alpha^{(\lambda)} + V^{(\lambda)}$ are smaller than the negative estimate of $\alpha^{(\theta)}$ (see Subplot (c) in Figure 5), which indicates higher values of weather/climate factors corresponding to less daily confirmed cases.
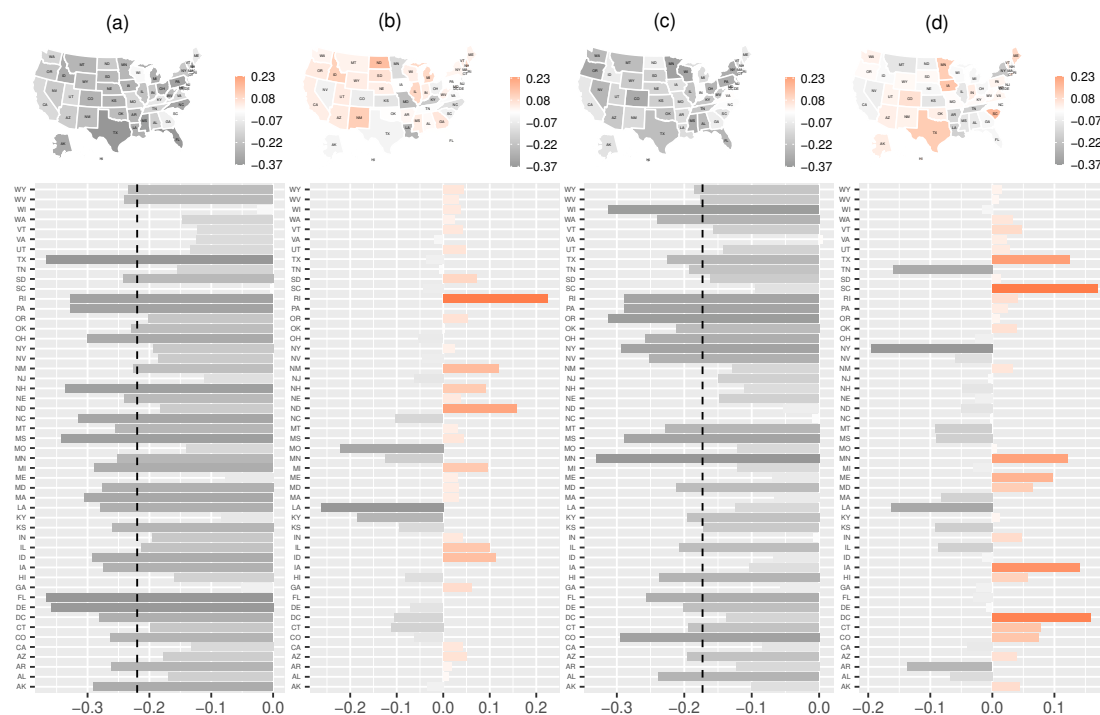

Figure 5. State-level estimated random effects in the multivariate time series model for: (**a**) epidemic-within component $V^{(\lambda)}$; (**b**) epidemic-between component $V^{(\psi)}$; (**c**) epidemic-boosted component $V^{(\theta)}$; and (**d**) endemic component $V^{(\zeta)}$. Vertical dotted lines in (**a**,**b**) are $-\alpha^{(\lambda)}$ and $-\alpha^{(\theta)}$. There is a strong variation in all four components, and different random effects for different states have various influences.

Figures 6 and 7 show the state-level percentage of daily confirmed cases of each component in every single day and the state-level means of fitted values based on (1). It is obvious that the estimated values of the epidemic-within component are the highest among all components in most of the states (e.g., CA) as time goes on. This phenomenon suggests that cross-state spread, weather/climate factors' influence, or other unobserved factors have little impact on COVID-19 transmission, whereas the previous state's related cases predominate the fluctuation of future infections in most states. For several states such as Connecticut (CT), New Hampshire (NH), and Rhode Island (RI), no particular component demonstrates a dominant effect on daily confirmed cases. According to Figure 1, we find that with relatively less confirmed cases, the endemic component seems to be more dominant in these states, which is also shown in CT, NH, and RI.
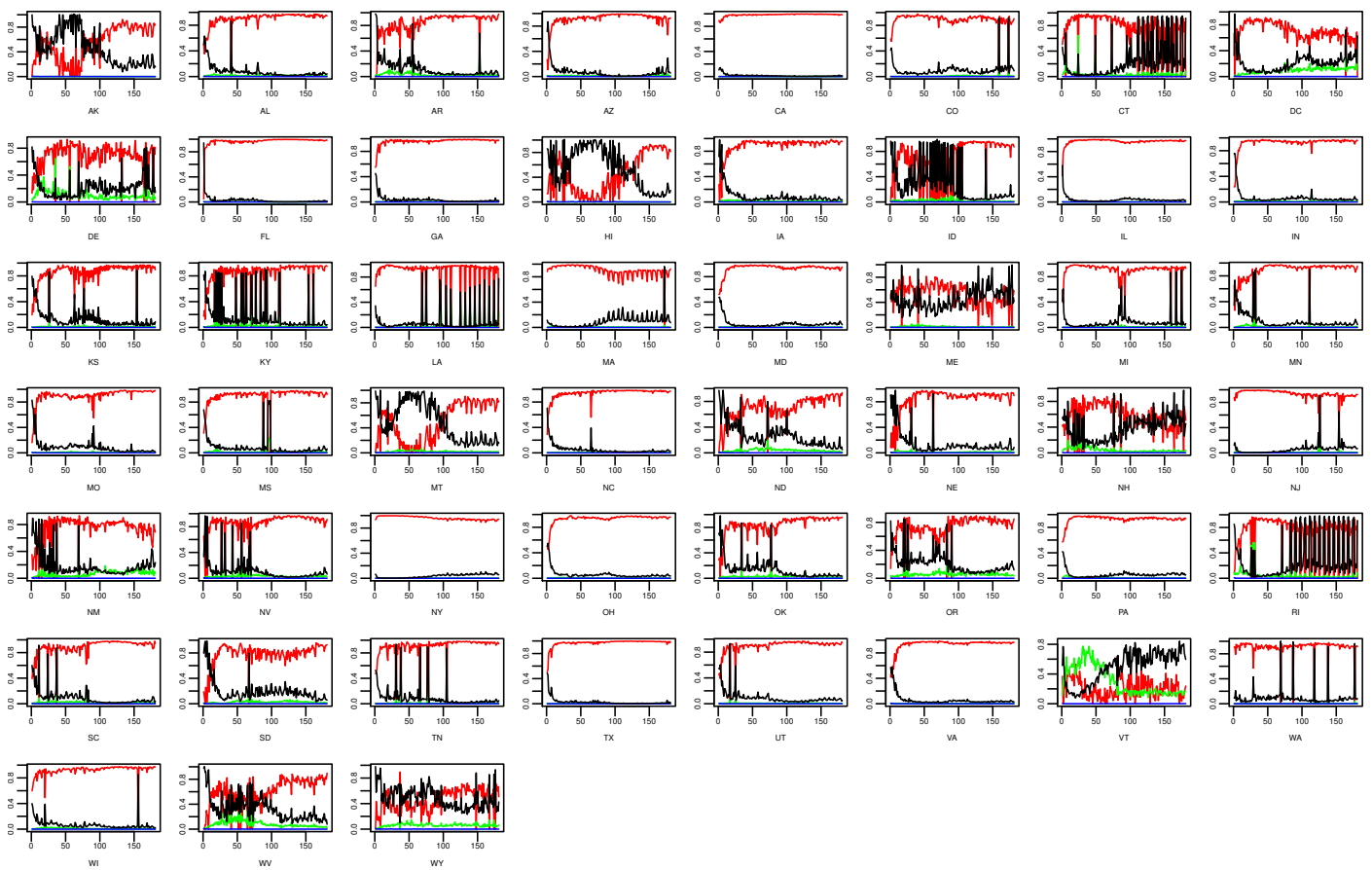
Figure 6. State-level daily percentage of daily confirmed cases with respect to the epidemic-within, endemic-between, epidemic-boosted, and endemic components, which are indicated by red, green, blue, and black lines, respectively.
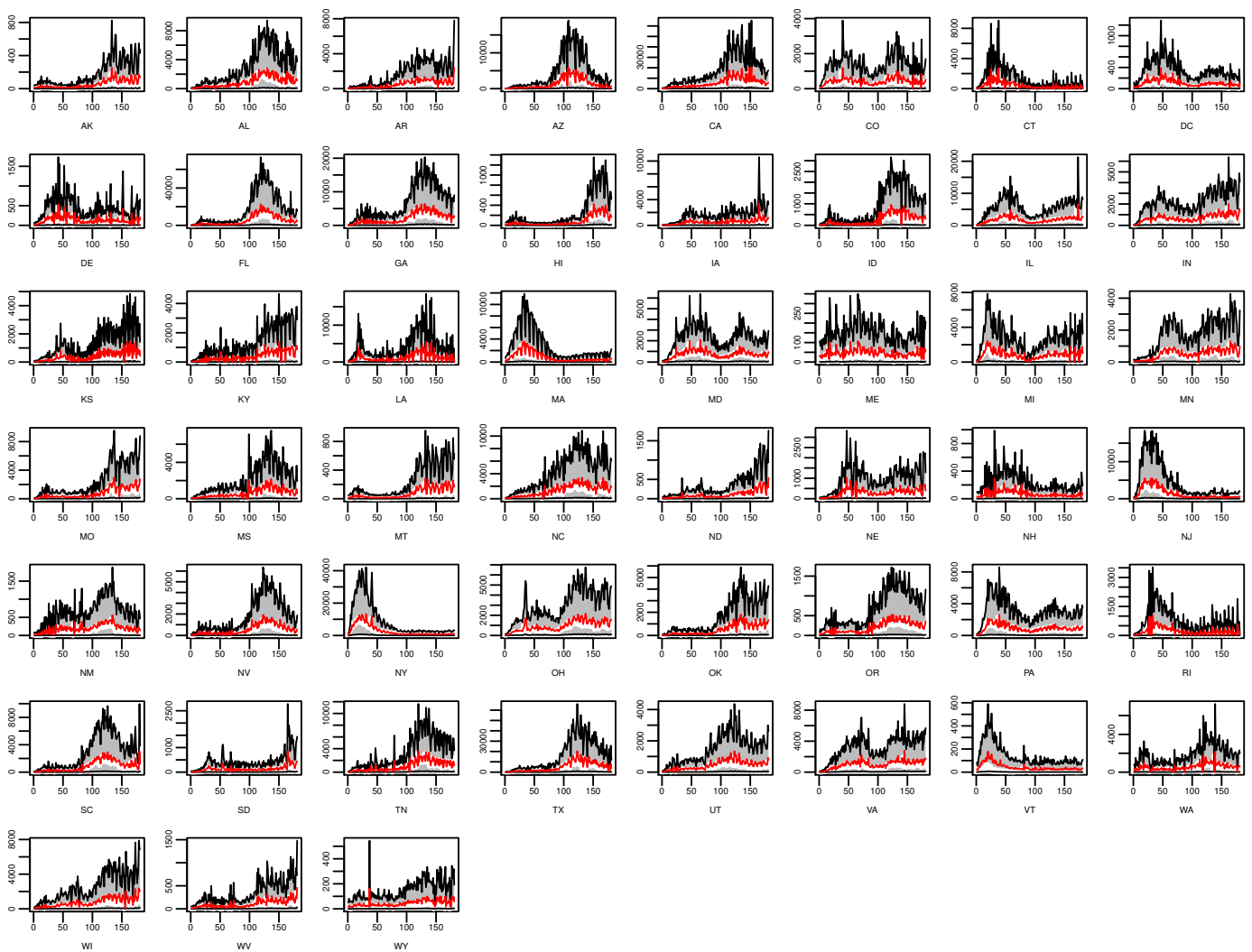
Figure 7. State-level estimated means of daily confirmed cases and related 90% confidence intervals. The gray band represents the 90% confidence interval, the white line the observations of the daily confirmed cases, and the red line the estimated mean values.

4. Discussion

In this article, we find that MaT, MiT, and CDTAO have statistically significant associations with daily confirmed cases in almost all the states in America, based on both the Kendall and Spearman-rank correlation tests. Furthermore, from the estimated coefficients of the epidemic-boosted component, we identify that this association is negative. That is, higher MaT, MiT, and CDTAO correspond to smaller daily confirmed cases. However, further analysis uncovers that the previous daily confirmed cases in one state itself are generally predominant for the next confirmed cases, which suggests that states with a large number of confirmed cases tend to cause more infections.

Recent research [18,19] has shown some evidence of the correlation between weather/climate factors and COVID-19 transmission. Our work, based on an extended multivariate time series model, further confirms and quantifies the existence of a similar relationship. Unlike some existing models (e.g., [10]), our model successfully facilitates the interpretability and practicability by an additional term that characterizes the degree of the influence from some other external factors. However, one obvious drawback of our model is that a large number of unknown parameters need to be estimated, and the computational cost is therefore high. As a result, the effective sample size also needs to be sufficiently large. Recently, Kimball et al. [30] investigated a COVID-19 outbreak in a long-term care

skilled nursing facility (SNF) in King County, Washington, recognized on 28 February 2020, and discovered that screening the SNF residents based on the symptoms related to this epidemic could not to discover all SARS-CoV-2 infections since they found that 23 (30.3%) workers had SARS-CoV-2 positive tests even if they were asymptomatic or presymptomatic on the testing day. Such clustering based infections dramatically aggregate the transmission intensity, making the analysis based on not only our model, but others previously mentioned non-trivial challenges. More recently, Rader et al.'s analysis [31] unveiled that "epidemics in crowded cities are more spread over time, and crowded cities have larger total attack rates than less populated cities". Such phenomena further imply that an area with a higher population density could cause a much severer outbreak, which is also a situation that requires further investigation.

5. Conclusions

As COVID-19 has become the most disastrous health event in the world, especially in the USA by far, so understanding the transmission pattern of this pandemic has become more urgent. Our analyses of the COVID-19 surveillance data depict remarkably heterogeneous transmission across states during the COVID-19 outbreak in the USA from 15 March 2020 to 15 September 2020. The degree of heterogeneity is characterized by random effects parameter estimates. With the Kendall and Spearman-rank correlation tests, we explore the association between weather/climate factors and daily confirmed COVID-19 cases for each state, which is further used for the analysis of the spatial and temporal occurrence of COVID-19.

Some interesting findings are noteworthy. First, the heterogeneity of COVID-19 transmission across states is observed in all four components, which implies that there are different situations in different states and the same strategies may not work perfectly to contain this pandemic in all states.

Second, some weather/climate factors (i.e., CC, Hu, MaT, MiT, and CDTAO) demonstrate significant correlations with daily confirmed cases in many states. In particular, MaT, MiT, and CDTAO have a strong association with most states. Based on the estimated coefficients from the epidemic-boosted component, one can further find that these variables correspond to daily confirmed cases with a negative correlation in almost all states, i.e., higher MaT, MiT, and CDTAO correlate with less daily confirmed cases. This phenomenon suggests that climate change in the local and adjacent areas could affect the possibility of infection in this area.

Third, since the estimates of the epidemic-within component in most states are predominant, their corresponding values can represent the fluctuation of the daily confirmed cases. Since the estimated coefficients of the epidemic-within component in most states are smaller than one ($\exp\left(\alpha^{(\lambda)} + V_r^{(\lambda)}\right) < 1$), the community-level spread of COVID-19 in most states is remarkably mitigated and the transmission intensity is decreased. Furthermore, we believe that the relatively large number of daily confirmed cases in the current stage are mainly due to the previous large number of infected cases, and the number of new confirmed cases per day will gradually decrease as time goes on.

Fourth, for the future tendency of the daily confirmed cases, the influence of the previous confirmed cases is the most important among the four components. This means that for states like Texas (TX), which has a large number of confirmed cases, the risk of a sharp increase of the daily confirmed cases is still higher than other states with less confirmed cases. Therefore, there is no doubt that regulations like social distancing or wearing masks in public places are clearly necessary.

Final, since the endemic components for some states (e.g., Vermont (VT)) show obvious predominance, other possible variables that could influence COVID-19 transmission need to be further determined.

Appendix A. Materials for Parameter Estimation

Appendix A.1. Inference

We consider the process of model parameter estimation. Let $\boldsymbol{\sigma} = (\sigma_{(\lambda)}, \sigma_{(\psi)}, \sigma_{(\theta)}, \sigma_{(\zeta)})^\top$, $\boldsymbol{\varepsilon} = (\varepsilon_1, \cdots, \varepsilon_R)^\top$, $\boldsymbol{\alpha} = (\alpha^{(\lambda)}, \alpha^{(\psi)}, \alpha^{(\theta)}, \alpha^{(\zeta)})^\top$, and $\boldsymbol{\beta} = (\beta_1^{(\psi)}, \beta_1^{(\zeta)}, \beta_2^{(\zeta)}, \beta_3^{(\zeta)}, \beta_4^{(\zeta)})^\top$ represent all unknown parameters. Let $\boldsymbol{v} = (\boldsymbol{v}^{(\lambda)^\top}, \boldsymbol{v}^{(\psi)^\top}, \boldsymbol{v}^{(\theta)^\top}, \boldsymbol{v}^{(\zeta)^\top})^\top$ denote the observations of random effects $\boldsymbol{V}$.

Recall that $Y_{r,t}|\boldsymbol{Y}_{.,t-l}, \boldsymbol{V} \sim \text{NegBin}(\mu_{r,t}, \varepsilon_r)$, and $\boldsymbol{V} \sim N(0, \Sigma)$. The conditionally joint probability function is then given as:

$$
\begin{aligned}
f(\boldsymbol{y}_t&, \boldsymbol{v}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\varepsilon}, \boldsymbol{\sigma}|\boldsymbol{y}_{t-l}) \\
&= f(\boldsymbol{y}_{,t}; \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\varepsilon}, \boldsymbol{\sigma}|\boldsymbol{v}, \boldsymbol{y}_{t-l}) f(\boldsymbol{v}; \boldsymbol{\sigma}) \\
&= \left\{ \prod_r \frac{\Gamma(y_{r,t} + \frac{1}{\varepsilon_r})}{\Gamma(y_{r,t}+1)\Gamma(\frac{1}{\varepsilon_r})} \left(\frac{\mu_{r,t}\varepsilon_r}{1 + \mu_{r,t}\varepsilon_r} \right)^{y_{r,t}} \left(\frac{1}{1 + \mu_{r,t}\varepsilon_r} \right)^{\frac{1}{\varepsilon_r}} \right\} \\
&\quad \cdot (2\pi)^{-2R} |\Sigma|^{-1/2} \exp(-\frac{1}{2}\boldsymbol{v}^\top \Sigma^{-1} \boldsymbol{v}) \\
&\propto \left\{ \prod_r \frac{\Gamma(y_{r,t} + \frac{1}{\varepsilon_r})}{\Gamma(\frac{1}{\varepsilon_r})} \left(\frac{\mu_{r,t}\varepsilon_r}{1 + \mu_{r,t}\varepsilon_r} \right)^{y_{r,t}} \left(\frac{1}{1 + \mu_{r,t}\varepsilon_r} \right)^{\frac{1}{\varepsilon_r}} \right\} \\
&\quad \cdot |\Sigma|^{-1/2} \exp(-\frac{1}{2}\boldsymbol{v}^\top \Sigma^{-1} \boldsymbol{v}).
\end{aligned}
\tag{A1}
$$

or the logarithm of the conditional likelihood function is:

$$
\begin{aligned}
l(\boldsymbol{\pi}, \boldsymbol{v}) \quad \propto \sum_{r,t} &\left\{ \log \Gamma(y_{r,t} + \frac{1}{\varepsilon_r}) - \log \Gamma(\frac{1}{\varepsilon_r}) + y_{r,t} \log\left(\frac{\mu_{r,t}\varepsilon_r}{1 + \mu_{r,t}\varepsilon_r} \right) \right. \\
&\left. - \frac{1}{\varepsilon_r} \log(1 + \mu_{r,t}\varepsilon_r) \right\} - \frac{1}{2}\log|\Sigma| - \frac{1}{2}\boldsymbol{v}^\top \Sigma^{-1} \boldsymbol{v},
\end{aligned}
\tag{A2}
$$

where $\boldsymbol{\pi} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}^\top, \boldsymbol{\varepsilon}^\top)^\top$, $|\cdot|$ denotes the determinant of a matrix and $\Gamma(\cdot)$ represents the gamma function.

Since (A2) includes the random effects, maximizing the above equation with respect to $\boldsymbol{\pi}$ and $\boldsymbol{v}$ does not generate the log-likelihood, but the penalized log-likelihood or posterior mode estimates from a Bayesian viewpoint [32]. Furthermore, not only the sample observations $\boldsymbol{v}$ are unavailable, but also the values of parameters $\hat{\sigma}^2_{(\lambda)}, \hat{\sigma}^2_{(\psi)}, \hat{\sigma}^2_{(\theta)}$, and $\hat{\sigma}^2_{(\zeta)}$ are unknown. To address this issue, Paul and Held [13] proposed a Laplace

Int. J. Environ. Res. Public Health **2021**, 18, 774

15 of 18

approximation based method to estimate the covariance by maximizing the marginal likelihood with respect to π and ε. It is noteworthy that their approach can be applied to general cases. Note that the covariance matrix (i.e., Σ) is assumed to be a diagonal matrix and $V_r^{(\cdot)} \sim (0, \sigma_{(\cdot)}^2 I), r = 1, \cdots, R$. If R is sufficiently large, we can replace $\sigma_{(\cdot)}^2$ by its sample variance. In our case, R represents the number of states in the USA, which can be safely assumed to be sufficiently enough. Specifically, we have the sample variance as:

$$\hat{\sigma}_{(\cdot)}^2 = \frac{v^{(\cdot)\top} v^{(\cdot)}}{R-1}. \tag{A3}$$

The estimate of $v^\top \Sigma^{-1} v$ then satisfies:

$$\widehat{v^\top \Sigma^{-1} v} = v^\top \widehat{\Sigma}^{-1} v = 4(R-1)$$

and the estimate of $|\Sigma|$ is thus:

$$|\widehat{\Sigma}| = (\hat{\sigma}_{(\lambda)}^2 \hat{\sigma}_{(\psi)}^2 \hat{\sigma}_{(\theta)}^2 \hat{\sigma}_{(\zeta)}^2)^R$$
$$= \frac{[(v^{(\lambda)\top} v^{(\lambda)})(v^{(\psi)\top} v^{(\psi)})(v^{(\theta)\top} v^{(\theta)})(v^{(\zeta)\top} v^{(\zeta)})]^R}{(R-1)^{4R}}.$$

Therefore, (A2) can be simplified as:

$$\begin{aligned}
l(\pi, v) \propto \sum_{r,t} \Big\{ &\log \Gamma\Big(y_{r,t} + \frac{1}{\varepsilon_r}\Big) - \log \Gamma\Big(\frac{1}{\varepsilon_r}\Big) \\
&+ y_{r,t} \log\Big(\frac{\mu_{r,t}\varepsilon_r}{1 + \mu_{r,t}\varepsilon_r}\Big) - \frac{1}{\varepsilon_r} \log\big(1 + \mu_{r,t}\varepsilon_r\big) \Big\} \\
&- \frac{R}{2} \sum_{\lambda,\psi,\theta,\zeta} \log\Big(v^{(\cdot)\top} v^{(\cdot)}\Big).
\end{aligned} \tag{A4}$$

Due to the difficulty of implementing the gamma function with a large input value (e.g., $\Gamma(200) = \infty$ in the R software), we replace the gamma function $\Gamma(\cdot)$ by Stirling's approximation [33], i.e., $\sqrt{2\pi} \exp(-u) u^{u-1/2}$. One noticeable advantage for this replacement is that one can directly calculate the logarithm value first rather than the value of the gamma function first (and the logarithm value next), which can significantly mitigate the data overflow phenomenon (e.g., $\log(\Gamma(u))|_{u=200} = \infty$, but $\log(\sqrt{2\pi} \exp(-u) u^{u-1/2})|_{u=200} = \frac{1}{2}\log(2\pi) - u + (u - \frac{1}{2})\log(u)|_{u=200} \approx 857.933$ in the R software). Accordingly, (A4) can be approximated as:

$$\begin{aligned}
l(\pi, v) \propto \sum_{r,t} \Big\{ &\Big(y_{r,t} + \frac{1}{\varepsilon_r}\Big) \log\Big(\frac{1 + y_{r,t}\varepsilon_r}{1 + \mu_{r,t}\varepsilon_r}\Big) \\
&+ y_{r,t} \log(\mu_{r,t}) - \frac{1}{2} \log(1 + \varepsilon_r y_{r,t}) \Big\} \\
&- \frac{R}{2} \sum_{\lambda,\psi,\theta,\zeta} \log\Big(v^{(\cdot)\top} v^{(\cdot)}\Big).
\end{aligned} \tag{A5}$$

To obtain estimates of π and v by maximizing (A5), we apply the Adam algorithm, which is sufficient and accessible for multi-parameter optimization and only requires first-order gradients with little memory requirement [34]. Its pseudo-code is shown in Algorithm A1 in Appendix A.3. Here, the initial values for v's, α's, and β's are generated from the standard normal distribution ($N(0, 1)$), while those for ε's are generated from the uniform distribution ($U(0.01, 1)$), which guarantees the non-negativity.

Appendix A.2. First-Order Partial Gradients of the Penalized Log-Likelihood Function

From (A5), we have:

$$\frac{\partial l(\boldsymbol{\pi}, \boldsymbol{v})}{\partial \mu_{r,t}} = \frac{y_{r,t}}{\mu_{r,t}} - \left(y_{r,t} + \frac{1}{\varepsilon_r}\right)\frac{\varepsilon_r}{1 + \mu_{r,t}\varepsilon_r},$$

$$\frac{\partial \mu_{r,t}}{\partial \lambda_{r,t}} = y_{r,t-l}, \ \frac{\partial \mu_{r,t}}{\partial \psi_{r,t}} = \Psi_{r,t-l},$$

$$\frac{\partial \mu_{r,t}}{\partial \theta_{r,t}} = \Theta_{r,t-l}, \ \frac{\partial \mu_{r,t}}{\partial \zeta_{r,t}} = 1, \ \frac{\partial \lambda_{r,t}}{\partial \alpha_r^{(\lambda)}} = \frac{\partial \lambda_{r,t}}{\partial v_r^{(\lambda)}} = \lambda_{r,t},$$

$$\frac{\partial \psi_{r,t}}{\partial \alpha^{(\psi)}} = \frac{\partial \psi_{r,t}}{\partial v_r^{(\psi)}} = \psi_{r,t} \ \frac{\partial \theta_{r,t}}{\partial \alpha^{(\theta)}} = \frac{\partial \theta_{r,t}}{\partial v_r^{(\theta)}} = 1,$$

$$\frac{\partial \zeta_{r,t}}{\partial \alpha^{(\zeta)}} = \frac{\partial \zeta_{r,t}}{\partial v_r^{(\zeta)}} = \zeta_{r,t}, \ \frac{\partial \psi_{r,t}}{\partial \beta_1^{(\psi)}} = \psi_{r,t}\log(Pu_{r,t}),$$

$$\frac{\partial \zeta_{r,t}}{\partial \beta_1^{(\zeta)}} = \zeta_{r,t}\log\left(\text{logit}(t)\right), \ \frac{\partial \zeta_{r,t}}{\partial \beta_2^{(\zeta)}} = \zeta_{r,t}\log(Po_{r,t}),$$

$$\frac{\partial \zeta_{r,t}}{\partial \beta_3^{(\zeta)}} = \beta_1^{(\zeta)}\left(1 - \text{logit}(t)\right), \ \frac{\partial \zeta_{r,t}}{\partial \beta_4^{(\zeta)}} = t\beta_1^{(\zeta)}\left(1 - \text{logit}(t)\right).$$

Thus,

$$\frac{\partial l(\boldsymbol{\pi}, \boldsymbol{v})}{\partial \varepsilon_r} = \sum_t -\frac{1}{\varepsilon_r^2}\log\left(\frac{1 + y_{r,t}\varepsilon_r}{1 + \mu_{r,t}\varepsilon_r}\right)$$

$$+ \left(y_{r,t} + \frac{1}{\varepsilon_r}\right)\left(\frac{y_{r,t}}{1 + y_{r,t}\varepsilon_r} - \frac{\mu_{r,t}}{1 + \mu_{r,t}\varepsilon_r}\right)$$

$$- \frac{y_{r,t}}{2(1 + \varepsilon_r y_{r,t})},$$

$$\frac{\partial l(\boldsymbol{\pi}, \boldsymbol{v})}{\partial v_r^{(\lambda)}} = \sum_t \frac{\partial l(\boldsymbol{\pi}, \boldsymbol{v})}{\partial \mu_{r,t}}\frac{\partial \mu_{r,t}}{\partial \lambda_{r,t}}\frac{\partial \lambda_{r,t}}{\partial v_r^{(\lambda)}} - \frac{v_r R}{|\boldsymbol{v}^{(\lambda)}|^2}$$

$$= \sum_t \lambda_{r,t}y_{r,t-l}\left(\frac{y_{r,t}}{\mu_{r,t}} - \frac{1 + y_{r,t}\varepsilon_r}{1 + \mu_{r,t}\varepsilon_r}\right) - \frac{v_r R}{|\boldsymbol{v}^{(\lambda)}|^2},$$

$$\frac{\partial l(\boldsymbol{\pi}, \boldsymbol{v})}{\partial v_r^{(\psi)}} = \sum_t \frac{\partial l(\boldsymbol{\pi}, \boldsymbol{v})}{\partial \mu_{r,t}}\frac{\partial \mu_{r,t}}{\partial \psi_{r,t}}\frac{\partial \psi_{r,t}}{\partial v_r^{(\psi)}} - \frac{v_r R}{|\boldsymbol{v}^{(\psi)}|^2}$$

$$= \sum_t \psi_{r,t}\Psi_{r,t-l}\left(\frac{y_{r,t}}{\mu_{r,t}} - \frac{1 + y_{r,t}\varepsilon_r}{1 + \mu_{r,t}\varepsilon_r}\right) - \frac{v_r R}{|\boldsymbol{v}^{(\psi)}|^2},$$

$$\frac{\partial l(\boldsymbol{\pi}, \boldsymbol{v})}{\partial v_r^{(\theta)}} = \sum_t \frac{\partial l(\boldsymbol{\pi}, \boldsymbol{v})}{\partial \mu_{r,t}}\frac{\partial \mu_{r,t}}{\partial \theta_{r,t}}\frac{\partial \theta_{r,t}}{\partial v_r^{(\theta)}} - \frac{v_r R}{|\boldsymbol{v}^{(\theta)}|^2}$$

$$= \sum_t \Theta_{r,t-l}\left(\frac{y_{r,t}}{\mu_{r,t}} - \frac{1 + y_{r,t}\varepsilon_r}{1 + \mu_{r,t}\varepsilon_r}\right) - \frac{v_r R}{|\boldsymbol{v}^{(\theta)}|^2},$$

$$\frac{\partial l(\boldsymbol{\pi}, \boldsymbol{v})}{\partial v_r^{(\zeta)}} = \sum_t \frac{\partial l(\boldsymbol{\pi}, \boldsymbol{v})}{\partial \mu_{r,t}}\frac{\partial \mu_{r,t}}{\partial \zeta_{r,t}}\frac{\partial \zeta_{r,t}}{\partial v_r^{(\zeta)}} - \frac{v_r R}{|\boldsymbol{v}^{(\zeta)}|^2}$$

$$= \sum_t \zeta_{r,t}\left(\frac{y_{r,t}}{\mu_{r,t}} - \frac{1 + y_{r,t}\varepsilon_r}{1 + \mu_{r,t}\varepsilon_r}\right) - \frac{v_r R}{|\boldsymbol{v}^{(\zeta)}|^2},$$

$$\frac{\partial l(\boldsymbol{\pi}, \boldsymbol{v})}{\partial \boldsymbol{\pi}} = \sum_{r,t} \frac{\partial l(\boldsymbol{\pi}, \boldsymbol{v})}{\partial \mu_{r,t}}\frac{\partial \mu_{r,t}}{\partial \boldsymbol{\pi}} = \sum_{r,t}\left(\frac{y_{r,t}}{\mu_{r,t}} - \frac{1 + y_{r,t}\varepsilon_r}{1 + \mu_{r,t}\varepsilon_r}\right)$$

$$\left(\lambda_{r,t}, \psi_{r,t}, 1, \zeta_{r,t}, \psi_{r,t}\log(Pu_{r,t}), \zeta_{r,t}\log\left(\text{logit}(t)\right),\right.$$

$$\left.\zeta_{r,t}\log(Po_{r,t}), \beta_1^{(\zeta)}\left(1 - \text{logit}(t)\right), t\beta_1^{(\zeta)}\left(1 - \text{logit}(t)\right)\right)^\top,$$

where:

$$\text{logit}(t) = \left(1 + \exp\left\{-(\beta_3^{(\zeta)} + \beta_4^{(\zeta)}t)\right\}\right)^{-1}.$$

Int. J. Environ. Res. Public Health **2021**, *18*, 774

17 of 18

Thus,

$$
\nabla_{(\boldsymbol{\pi},v)}l(\boldsymbol{\pi},\boldsymbol{v}) = \left(\left(\frac{\partial l(\boldsymbol{\pi},\boldsymbol{v})}{\partial \boldsymbol{\pi}} \right)^{\top}, \frac{\partial l(\boldsymbol{\pi},\boldsymbol{v})}{\partial v_r^{(\lambda)}}, \right.
$$
$$
\left. \frac{\partial l(\boldsymbol{\pi},\boldsymbol{v})}{\partial v_r^{(\psi)}}, \frac{\partial l(\boldsymbol{\pi},\boldsymbol{v})}{\partial v_r^{(\theta)}}, \frac{\partial l(\boldsymbol{\pi},\boldsymbol{v})}{\partial v_r^{(\zeta)}} \right)^{\top}.
\tag{A6}
$$

Appendix A.3. Pseudo Algorithm

Here is the algorithm for the parameters' optimization in (A5).

Algorithm A1 Adam based method for parameter optimization. Good default settings for the analyzed COVID-19 dataset are learning rate $\eta = 0.05$, exponential decay rates $b_1 = 0.1$ and $b_2 = 0.1$, and $\gamma = 1 \times 10^{-8}$. Algorithm tolerance $tol = 1 \times 10^{-4}$. All operations are element-wise.

Initialization: maxit = 200 (maximum iteration steps), flag = 0 (convergence indicator), $e_1 = 0$ (first moment vector), $e_2 = 0$ (second moment vector), $\epsilon = 0$ (iteration-step indicator), $(\boldsymbol{\pi}_0, \boldsymbol{v}_0)$, $a_0 = 0$, $\tau_0 = 0$.

Iteration process:

  while $\epsilon <=$ maxit **and** flag = 0 **do**

    $\epsilon = \epsilon + 1$

    $g_\epsilon = \nabla_{(\boldsymbol{\pi},v)}l(\boldsymbol{\pi}_{\epsilon-1}, \boldsymbol{v}_{\epsilon-1})$ (gradients of $l(\boldsymbol{\pi},\boldsymbol{v})$ shown in (A6) in Appendix A.2)

    $e_\epsilon^{(1)} = \frac{b_1 e_{\epsilon-1}^{(1)} - (1-b_1)g_\epsilon}{1 - b_1^\epsilon}$ (bias-corrected first moment estimate)

    $e_\epsilon^{(2)} = \frac{b_2 e_{\epsilon-1}^{(2)} + (1-b_2)g_\epsilon^2}{1 - b_2^\epsilon}$ (bias-corrected second raw moment estimate)

    $\tau_\epsilon = (\boldsymbol{\pi}_{\epsilon-1}, \boldsymbol{v}_{\epsilon-1}) - \eta \frac{e_\epsilon^{(1)}}{\sqrt{e_\epsilon^{(2)}} + \gamma}$ (temporarily updated parameters)

    $(\boldsymbol{\pi}_\epsilon, \boldsymbol{v}_\epsilon) = b_2 a_{\epsilon-1} + (1 - b_2)\tau_\epsilon$ (updated parameters)

    $a_\epsilon = (1 - \frac{1}{\epsilon})a_{\epsilon-1} + \frac{1}{\epsilon}\tau_\epsilon$ (averaged parameters for further iteration)

    if $\frac{|(\boldsymbol{\pi}_\epsilon, \boldsymbol{v}_\epsilon) - (\boldsymbol{\pi}_{\epsilon-1}, \boldsymbol{v}_{\epsilon-1})|}{\sqrt{5R+9}} < tol$ (convergence determination)

      flag = 1

  end while

  return $(\boldsymbol{\pi}_\epsilon, \boldsymbol{v}_\epsilon)$ (optimal estimates)

References

1. Gorbalenya, A.E.; Baker, S.C. The species Severe acute respiratory syndrome-related coronavirus: Classifying 2019-nCoV and naming it SARS-CoV-2. *Nat. Microbiol.* **2020**, *5*, 536.
2. Le, T.T.; Andreadakis, Z.; Kumar, A.; Roman, R.G.; Tollefsen, S.; Saville, M.; Mayhew, S. The COVID-19 vaccine development landscape. *Nat. Rev. Drug Discov.* **2020**, *19*, 305–306. [CrossRef] [PubMed]
3. Graham, B.S. Rapid COVID-19 vaccine development. *Science* **2020**, *368*, 945–946. [CrossRef] [PubMed]
4. Suganya, S.; Divya, S.; Parani, M. Severe acute respiratory syndrome-coronavirus-2: Current advances in therapeutic targets and drug development. *Rev. Med Virol.* **2020**. [CrossRef] [PubMed]
5. Chowell, G.; Mizumoto, K. The COVID-19 pandemic in the USA: What might we expect? *Lancet* **2020**, *395*, 1093–1094. [CrossRef]
6. Rui, R.; Tian, M. Joint Estimation of Case Fatality Rate of COVID-19 and Power of Quarantine Strategy Performed in Wuhan, China. *Biom. J.* **2020**, *63*, 46–58. [CrossRef] [PubMed]
7. Kermack, W.O.; McKendrick, A.G. A contribution to the mathematical theory of epidemics. *Proc. R. Soc. London. Ser. A Contain. Pap. Math. Phys. Character* **1927**, *115*, 700–721.
8. Helfenstein, U. Box-Jenkins modeling of some viral infectious diseases. *Stat. Med.* **1986**, *5*, 37–47. [CrossRef]
9. Trottier, H.; Philippe, P.; Roy, R. Stochastic modeling of empirical time series of childhood infectious diseases data before and after mass vaccination. *Emerg. Themes Epidemiol.* **2006**, *3*, 9. [CrossRef]
10. Papastefanopoulos, V.; Linardatos, P.; Kotsiantis, S. COVID-19: A Comparison of Time Series Methods to Forecast Percentage of Active Cases per Population. *Appl. Sci.* **2020**, *10*, 3880. [CrossRef]
11. Held, L.; Höhle, M.; Hofmann, M. A statistical framework for the analysis of multivariate infectious disease surveillance counts. *Stat. Model.* **2005**, *5*, 187–199. [CrossRef]

12. Paul, M.; Held, L.; Toschke, A.M. Multivariate modeling of infectious disease surveillance data. *Stat. Med.* **2008**, *27*, 6250–6267. [CrossRef] [PubMed]

13. Paul, M.; Held, L. Predictive assessment of a non-linear random effects model for multivariate time series of infectious disease counts. *Stat. Med.* **2011**, *30*, 1118–1136. [CrossRef] [PubMed]

14. Cheng, Q.; Lu, X.; Wu, J.T.; Liu, Z.; Huang, J. Analysis of heterogeneous dengue transmission in Guangdong in 2014 with multivariate time series model. *Sci. Rep.* **2016**, *6*, 33755. [CrossRef] [PubMed]

15. Adegboye, O.; Al-Saghir, M.; LEUNG, D.H. Joint spatial time-series epidemiological analysis of malaria and cutaneous leishmaniasis infection. *Epidemiol. Infect.* **2017**, *145*, 685–700. [CrossRef]

16. Wu, H.; Wang, X.; Xue, M.; Wu, C.; Lu, Q.; Ding, Z.; Zhai, Y.; Lin, J. Spatial-temporal characteristics and the epidemiology of hemorrhagic fever with renal syndrome from 2007 to 2016 in Zhejiang Province, China. *Sci. Rep.* **2018**, *8*, 1–14.

17. Dickson, M.M.; Espa, G.; Giuliani, D.; Santi, F.; Savadori, L. Assessing the effect of containment measures on the spatio-temporal dynamic of COVID-19 in Italy. *Nonlinear Dyn.* **2020**, *101*, 1833–1846. [CrossRef]

18. 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]

19. Tosepu, R.; Gunawan, J.; Effendy, D.S.; Lestari, H.; Bahar, H.; Asfian, P. Correlation between weather and Covid-19 pandemic in Jakarta, Indonesia. *Sci. Total Environ.* **2020**, *725*, 138436. [CrossRef]

20. Hugh-Jones, M.; Wright, P. Studies on the 1967–8 foot-and-mouth disease epidemic: The relation of weather to the spread of disease. *Epidemiol. Infect.* **1970**, *68*, 253–271. [CrossRef]

21. Tan, J.; Mu, L.; Huang, J.; Yu, S.; Chen, B.; Yin, J. An initial investigation of the association between the SARS outbreak and weather: With the view of the environmental temperature and its variation. *J. Epidemiol. Community Health* **2005**, *59*, 186–192. [CrossRef] [PubMed]

22. Verdasca, J.; Da Gama, M.T.; Nunes, A.; Bernardino, N.; Pacheco, J.; Gomes, M. Recurrent epidemics in small world networks. *J. Theor. Biol.* **2005**, *233*, 553–561. [CrossRef] [PubMed]

23. Hassen, H.B.; Elaoud, A.; Salah, N.B.; Masmoudi, A. A SIR-Poisson Model for COVID-19: Evolution and Transmission Inference in the Maghreb Central Regions. *Arab. J. Sci. Eng.* **2021**, *46*, 93–102. [CrossRef] [PubMed]

24. Read, J.M.; Bridgen, J.R.; Cummings, D.A.; Ho, A.; Jewell, C.P. Novel coronavirus 2019-nCoV: Early estimation of epidemiological parameters and epidemic predictions. *MedRxiv* **2020**. [CrossRef]

25. Giuliani, D.; Dickson, M.M.; Espa, G.; Santi, F. *Modelling and Predicting the Spatio-Temporal Spread of Coronavirus Disease 2019 (COVID-19) in Italy*; University of Trento: Trento, Italy, 2020.

26. Gao, X.; Dong, Q. A logistic model for age-specific COVID-19 case-fatality rates. *JAMIA Open* **2020**, *3*, 151–153. [CrossRef]

27. Kass, D.A.; Duggal, P.; Cingolani, O. Obesity could shift severe COVID-19 disease to younger ages. *Lancet* **2020**, *399*, 1544–1545. [CrossRef]

28. Harris, J.E. Data from the COVID-19 epidemic in Florida suggest that younger cohorts have been transmitting their infections to less socially mobile older adults. *Rev. Econ. Househ.* **2020**, *18*, 1019–1037.. [CrossRef]

29. Bashir, M.F.; Ma, B.; Komal, B.; Bashir, M.A.; Tan, D.; Bashir, M. Correlation between climate indicators and COVID-19 pandemic in New York, USA. *Sci. Total Environ.* **2020**, *728*, 138835. [CrossRef]

30. Kimball, A.; Hatfield, K.M.; Arons, M.; James, A.; Taylor, J.; Spicer, K.; Bardossy, A.C.; Oakley, L.P.; Tanwar, S.; Chisty, Z.; et al. Asymptomatic and presymptomatic SARS-CoV-2 infections in residents of a long-term care skilled nursing facility—King County, Washington, March 2020. *Morb. Mortal. Wkly. Rep.* **2020**, *69*, 377. [CrossRef]

31. Rader, B.; Scarpino, S.V.; Nande, A.; Hill, A.L.; Adlam, B.; Reiner, R.C.; Pigott, D.M.; Gutierrez, B.; Zarebski, A.E.; Shrestha, M.; et al. Crowding and the shape of COVID-19 epidemics. *Nat. Med.* **2020**, *26*, 1829–1834. [CrossRef]

32. Kneib, T.; Fahrmeir, L. A mixed model approach for geoadditive hazard regression. *Scand. J. Stat.* **2007**, *34*, 207–228. [CrossRef]

33. Pearson, K. Historical note on the origin of the normal curve of errors. *Biometrika* **1924**, *16*, 402–404. [CrossRef]

34. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.